A Geometric Comparison of Algorithms for Fusion Control in Stereoscopic HTDs

Technical Report GIT-GVU-01-08a

7/5/01

Zachary Wartell, Larry F. Hodges, William Ribarsky

© 2001. IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

This work has been submitted to the IEEE for possible publication. Copyright may be transferred without notice, after which this version may no longer be accessible.

Author's Note (7/5/01): Following IEEE copyright, the original version of this technical report entitled "An Analytic Comparison of α-False Eye Separation, Image Scaling and Image Shifting in Stereoscopic Displays" (GVU Technical Report 00-09) is no longer available. The present document represents the revised version. The full citation to the official paper will be added to this document upon final publication.

A Geometric Comparison of Algorithms for Fusion Control in Stereoscopic HTDs

Zachary Wartell, Larry F. Hodges, William Ribarsky

Abstract

This paper concerns stereoscopic virtual reality displays in which the head is tracked and the display is stationary, attached to a desk, tabletop or wall. These are called stereoscopic HTDs (Head-Tracked Display). Stereoscopic displays render two perspective views of a scene, each of which is seen by one eye of the user. Ideally the user's natural visual system combines the stereo image pair into a single, 3D perceived image. Unfortunately users often have difficulty fusing the stereo image pair. Researchers use a number of software techniques to reduce fusion problems. This paper geometrically examines and compares a number of these techniques and reaches the following conclusions. In interactive stereoscopic applications, the combination of view placement, scale and either false eye separation or α-false eye separation can provide fusion control geometrically similar to image shifting and image scaling. However, in stereo HTDs image shifting and image scaling also generate additional geometric artifacts not generated by the other methods. We anecdotally link some of these artifacts to exceeding perceptual limitations of human vision. While formal perceptual studies are still needed, geometric analysis suggests that image shifting and image scaling may be less appropriate than the other methods for interactive, stereo HTDs.

Index Terms: virtual reality, stereoscopic display, head-tracking, distortion

1 Introduction

Virtual environments aim to perceptually place the user in a computer-generated world. A key component of creating this illusion is interactive 3D imagery. To generate this imagery, a typical VR system has a location and orientation tracking device, an image generator and one or more displays. The tracking device determines the positions of the user's head and/or eyes and of the displays. The image generator computes the image that each eye would see on a display surface if the eye and the display existed inside the virtual world at their tracked positions. This image is then fed to the display. A VR system is typically configured either as a head-mounted display (HMD) or as a head-tracked display (HTD). In a HMD, the display is attached to a helmet or headset worn by the user, so both the eye points and the display surfaces are in continuous motion. In a HTD, the display is stationary, attached to a desk, tabletop, or wall. Hence only the eye points move. HTD examples are the CAVE [1], desktop VR [4][29] and the responsive workbench [13].

Many VR systems generate a pair of images, one for each eye. This stereoscopic imagery provides a true 3D image so virtual objects appear to exist in front of and behind the physical display surface. Software methods for stereoscopic display are well known [8][18][24]. Stereoscopic display for virtual reality has been shown to improve user depth perception and task performance in a variety of tasks [3][20] [29]. This is not surprising since real world experience shows that stereopsis is an important depth cue, especially for objects within the user's personal space (1.5 meters) [2].

Stereoscopic HTDs add additional challenges to interface design. A key issue is maintaining good stereoscopic viewing conditions. This requires balancing several goals [35] including the fusibility of the stereoscopic images. Both experience [14] and experimental studies [25][42]

show that users with normal stereoscopic vision often have trouble fusing stereo image pairs into a single 3D image even when the viewing geometry is modeled exactly. Users may experience headaches, eyestrain and/or fatigue. At extremes they may be unable to fuse the images into a single 3D perception. A primary cause for these problems is the accommodation/vergence conflict. When the eyes fixate on a point in the physical world, two processes simultaneously occur. First, the eyes' lenses change shape or *accommodate* in order to create a clear image on the retina. Second, the eyes rotate or *verge* so that the image of the fixation point falls on the central region of the eye where visual acuity is greatest. When viewing a stereoscopic 3D object, however, the eyes verge to the synthetic object depth but accommodate elsewhere, nominally to the display screen depth. This violates the natural coupling of accommodation and vergence.

When a user cannot fuse a stereo image pair into single 3D image, she experiences diplopia (double vision). In a stereoscopic display the occurrence of diplopia is related to various physical attributes of the display system and the geometry of the display environments [8]. Some relevant geometric aspects are:

- the distance of the displayed virtual object relative to the display surface
- the eye separation value used in computing the viewing transformation
- the distance of the user's eyes to the display surface

Figure 1 illustrates two important measurements. The eyes are on the left and a point on a virtual object is on the right. This point is projected onto two points on the projection plane in the middle. The two projected points which correspond to a single 3D point are called homologous points. The screen parallax, *p*, associated with a virtual 3D point is the distance between these 2D image points. It's typically reported in meters or centimeters. Related to the

screen parallax is the horizontal visual angle or HVA. In Figure 1 the HVA is the visual angle, θ, subtended by the screen parallax, p. It's typically reported in degrees. For virtual points in front of the screen, screen parallax and HVA are defined to be negative while for virtual points behind the screen, screen parallax and HVA are defined to be positive. The negative values are called "crossed" values since eyes which are fixated on the screen must cross in order to fixate on a point in front of the screen. The positive values are called "uncrossed" values since eyes which are fixated on the screen must uncross to fixate on a point behind the screen.

Research has shown that if the HVA of homologous points is outside a certain range then diplopia occurs and the 3D depth illusion collapses [42]. Yeh and Silverstein found an average range of [-4.93,1.57] degrees for a stereo CRT. Even if diplopia does not occur, additional problems such as fatigue, eyestrain and headaches may occur [14]. Mon-Williams, Wann, and Rushton [16] show that using a stereoscopic display can temporarily alter the visual system's internal coupling of accommodation and vergence. Collectively, we refer to these problems as 'image fusion problems.'

To minimize image fusion problems, the software can dynamically adjust the user's view of the environment as she travels through and manipulates the virtual world. Numerous degrees of freedom control the view. We partition these into: view placement, view scale and view optics. View placement refers to the location and orientation of the projection window. The projection window is the virtual representation of the HTD's physical display surface in the virtual world. View placement does not refer to eye point locations because in a HTD the user's head position is a physical parameter controlled by the user and is not under software control. View scale is a single degree of freedom that represents the viewer's size in the world. View optics includes all other parameters modeled by the pin-hole camera model in interactive computer graphics. This

includes modeled eye separation, the position of the near and far clipping planes, field of view, and other distortions such as depth compression or expansion.

For controlling image fusion problems in non-head-tracked displays researchers use a variety of techniques [38]. This paper focuses on:

(1) 'false eye separation' – This method sets the modeled eye separation to an underestimated value either statically [8] or dynamically [30][31].

(2) 'α-false eye separation' – In early stereo cinema and photography, the lack of head-tracking causes the perceived 3D image to warp and shear with head motion [26][27]. Real-time image generation using head-tracking and proper eye separation can theoretically remove this effect. However, using false eye separation reintroduces the problem even with perfect head-tracking [34]. The α-false eye separation technique removes the shearing effects due to head motion parallel to the screen [36].

(3) 'image scaling' –This method scales down the projected 2D images about the center of the screen (called frame magnification in [23]).

(4) 'image shifting' – This method translates the left and right eye projected 2D images towards each other. Its use for altering the stereo image dates back to early stereo photography of the 1800's [27, p201-202]. Hodges and McAllister [10] discuss it with respect to computer graphics and Siegel, Tobinaga, and Akiya [21] use it to manage image fusion in non-head-tracked, stereoscopic video.

(5) 'view scaling' - This scales the view in order to *uniformly* scale the perceived world so that geometry is at a comfortably fusible distance. For non-head-tracked displays, Ware et al. [30][31][33] mix view scaling with underestimated eye separation. For certain applications, view scaling and view placement are sufficient [32][35].

Understanding the distortion effects of view scaling is straightforward because this algorithm explicitly specifies a 3D transformation. However, understanding the distortion effects of the other methods is non-trivial since they work indirectly on the 2D projected images or on the view optic parameters. Empirically, the indirect methods tend to have much more subtle and less noticeable effects on the monoscopic image properties. These properties can be seen when viewing only a single eye's image. False eye separation, for example, has negligible monoscopic effects since it involves only an inward movement of the modeled eye points. Image shifting basically has no effects on the monoscopic images.

While any other view optic technique for reducing fusion problems will distort the perceived image, the question is what are the artifacts, or the distracting and undesirable aspects, of a technique's distortion, and how do these artifacts compare with those of other techniques. Some artifacts maybe static and noticeable with no head motion while other artifacts maybe dynamic and only noticeable with head motion. Also note that in a stereoscopic HTD the user is free to view the display from an arbitrary angle. While often a user's eye axis remains roughly parallel to the screen in a single display system, this is not guaranteed. Moreover, in many multi-display systems the eye axis will always be non-parallel to some display surface. For these reasons, a general purpose method for fusion control should have minimal artifacts in the non-parallel case too.

This paper takes a theoretic, geometric approach to comparing image scaling and image shifting with false eye separation and α-false eye separation. The goal is to compute an equivalent 3D transformation that when applied to the virtual scene would yield the same displayed result as each of these four methods. These equivalent 3D transformations indicate what geometric distortions each of these four techniques induce on the virtual scene. While we

have not performed formal human factors experiments to examine the perceptual effects of these distortions, the theoretic analysis along with informal observation by the authors yield some interesting results. When the eye axis is parallel to the screen, image scaling and α-false eye separation yield distortions that change with head motion perpendicular to the scene. In contrast image shifting and false eye separation yield distortions that change with head motion both perpendicular and parallel to the screen. Image shifting has a translation component which can conflict with the standard view location parameter. Image scaling has a uniform 3D scale component which can conflict with the view scale parameter. When the eye axis is not parallel to the screen, the distortions of image shifting and image scaling exhibit further artifacts not present in either false eye separation method. Some of these additional artifacts are anecdotally linked to aggravating fusion problems. This paper's geometric analysis is an important step and guide toward future empirical studies and application development.

2 Background and Previous Work

False eye separation, image shifting and image scaling reduce image fusion problems by bringing corresponding projected 2D image points closer together. This reduces the screen parallaxes of the stereo image pair and reduces image fusion problems. These methods were undoubtedly experimented with by stereoscopic photographers and cinematographers as far back as the 1800's [15][27]. In the context of computer graphics, these screen parallax adjustments represent additional parameters in the viewing geometry. Software can also dynamically vary and optimize a given method's parameters based on the virtual scene content [30][31][33].

In [34] we observe and analyze the effects of false eye separation (either underestimated or overestimated) in a stereo HTD. We derive and discuss an analytic description of this distortion.

The distortion is a non-affine collineation. While straight lines are mapped to straight lines, neither distances, angles nor parallelism are preserved. Moreover, as the user moves her head laterally (i.e. parallel to the projection plane) the perceived 3D image will shear laterally. As the user moves her head perpendicular to the screen (i.e. towards or away from the screen), the perceived 3D image will compress and expand. The latter two effects are most disappointing for a stereo HTD. One of the reasons for adding head-tracking to a stereoscopic display is to remove these types of distortions which had been observed in earlier non-head-tracked systems. A user viewing a static, non-tracked stereo image that is rendered using correct eye separation will perceive the 3D image to warp and shear as she moves her head [8][26]. Head-tracking with correct eye separation removes this warping and yields a rigid 3D image under user head movement. False eye separation on a stereo HTD induces stereo movement all over again!

In [36] we decompose the matrix representing this distortion and extract and invert the shear component. We then dynamically place the inverted shear on the matrix stack at run-time. This inverted shear predistortion cancels out the shear component of the basic distortion. The result is that the perceived 3D image no longer shifts side to side with lateral head movement. This augmented false separation method is called α-false eye separation.

Image scaling scales down the projected points about the center of the screen while image shifting translates the left and right images toward one and another. If done carefully this reduces the overall maximum screen parallax in the scene. Both methods were developed on non-head-tracked systems. While we are aware of published literature addressing stereoscopic display distortions [4][8][9][17][27][30][34][41], none address the 3D effects on the 3D displayed image due to image scaling or image shifting with respect to stereo HTDs. It has been qualitatively noted, however, that image shifting alters the perceived 3D depth of virtual objects

[5][11][27] and Hodges and McAllister [10] quantitatively describes how image shifting translates the stereoscopic window.

3 Methodology

To derive an analytic description of stereoscopic distortions, we must concisely describe the viewing model used in stereoscopic HTDs. A typical viewing model consists of a coordinate system hierarchy. The top coordinate system is the Platform Coordinate System. Manipulating this coordinate system moves the user through the virtual space. Directly attached to this coordinate system is the Projection Plane Coordinate System and the Tracker Coordinate System. The Projection Plane Coordinate System contains the projection plane in its XY plane with the projection window centered about the origin. The Tracker Coordinate System simply represents the tracker's emitter. Attached to the Tracker Coordinate System is the Head-Sensor Coordinate System and attached to that is the Eyes Coordinate System. The two eye points are on the x-axis of the Eyes Coordinate System and are symmetric about the origin.

The position and orientation of each child coordinate system relative to its parent is measured physically from the physical display setup as are the projection window dimensions. The Platform Coordinate System's mapping to virtual world coordinates defines the mapping of the physical space of the real world to the virtual space of the virtual world. In addition to specifying the position and orientation, the Platform Coordinate System can also be uniformly scaled. This causes the virtual world to appear to grow and shrink.

This scale factor requires that we distinguish between physical and virtual distance measurements. We assume that all measurements such as eye separation and window size are made in physical units. Each of these distances has a corresponding value in virtual space which

is computed by multiplying the physical value by the platform scale factor. For example if the eye separation is measured as 0.065 m and the platform scale is 1000, then the virtual eye separation is 65 m. In this paper, whenever, we discuss distance measurements such as eye separation, we are always referring to the physical values not the virtual ones. Importantly, the term false eye separation refers to a discrepancy between the true physical eye separation and the modeled physical eye separation and *not* the discrepancy between corresponding physical and virtual measurements.

When analyzing stereoscopic distortions we assume that all the important physical measurements are correct unless otherwise noted. This includes those made dynamically by the tracking system. We also assume any distortion due to curvature of the screen or any optics is negligible or accounted for by other means [4]. Additionally, we assume that the natural change of the eye separation as measured between the entrance pupils of the human eyes [19] is also relatively negligible or accounted for by proper eye tracking or eye gaze estimation [4].

These assumptions allow us to independently analyze the distortions induced by the fusion control technique under investigation. The methodology is to derive a 3D transformation, Δ. Transforming 3D "model space" by Δ and then projecting this new "displayed space" onto the image plane yields the same screen images as projecting the original modeled space onto the image plane and then manipulating the 2D images themselves. Δ provides a better understanding of how a fusion technique geometrically distortions model space to display space.

Being able to compare these induced geometric distortions is an important first step towards empirical investigation of their effects on the perceived space. While we do not investigate perceived space here, for completeness we briefly discuss the various operational definitions of perceived space found in the literature. Psychologists are interested in the discrepancies between

the geometry of physical environments and the geometry perceived by a human subject. Psychologists operationally define *perceived space* using several experimental formats [28]. In magnitude estimation subjects report the perceived measure such as size or distance of some target object. Subjects are given a reference object to define the reported units. In category estimation subjects place the target object's measure in one of a given set of categories. In perceptual matching the subject matches a reference stimuli to a target stimuli based on some measure. For example, they might view multiple target stimuli and have to pick the one of the same size as a reference stimulus. In mapping tasks, the subject constructs a small model, perhaps via hand drawing, of the environmental stimuli. Psychologist's often perform such studies and try to statistically fit the data to a variety of suggested theoretical models for perceived space. Researchers in virtual reality often operationally define perceived space quite differently [43]. They define perceived space through registration experiments. The subject views a virtual object at some depth and then positions a physical marker at the same perceived depth as the virtual object. The points of the perceived virtual scene are defined to be where the subject would position a physical marker to coincide with each virtual point. In such experiments the computational display model is set as accurately as possible and the virtual objects are kept to a range where they can be fused into single images. Fusion comfort is not as high priority as long as the images are fusible at all. Yoshida et. al. [43] perform such an experiment using a large projected stereo display. To avoid head-tracking errors the subject's head is held stationary during each trial. Eye-screen distance varied between 70 cm to 100 cm and eye-object distance varied between 30 and 50 cm. Positioning errors are generally under 1 cm and the authors present a numerical technique that theoretically can reduce the RMS (root mean square) errors to within 0.013 cm for the tested subjects.

Clearly, perceived space is an empirical, operationally defined space. This differs from the theoretically defined displayed space introduced earlier. However, the correlation between modeled and (registered) perceived space found by Yoshida et. al. [43] is fairly strong and shortly we'll illustrate that a fusion technique's effect on the 2D projected images is often geometrically equivalent to a 3D transformation of modeled space. To the degree that a fusion technique theoretically distorts a modeled point to a displayed point by a distance greater than the perceptual error range cited above (roughly 1 cm), then we'd expect the geometric distortion to be empirically measurable by formal registration experiments. Therefore, this paper's geometric analysis is an important step and guide toward further empirical studies.

4 Geometric Construction

In this section we develop specific geometric constructions describing the distortion due to image scaling and image shifting. These will lead to the analytic descriptions in Section 5.

4.1 Image scaling

The suggested implementation for image scaling (called frame magnification in [23]) is to compose the following matrices:

$$M = M_{scr}\, S_{mag}\, M_{proj}\, M_{view}\, M_{model} \tag{1}$$

In our notation, $\boldsymbol{M_{model}}$ maps model coordinates to world coordinates. $\boldsymbol{M_{view}}$ maps world coordinates to view coordinates. $\boldsymbol{M_{proj}}$ maps view coordinates to the canonical projection coordinates and $\boldsymbol{S_{mag}}$ is the image scaling. Note, $\boldsymbol{S_{mag}}$ only scales the x and y coordinates by a common scale factor. A scale factor less than 1 scales down the x and y coordinates which is equivalent to enlarging the modeled projection window size. Hence the name "frame magnification." $\boldsymbol{M_{scr}}$ maps canonical projection coordinates to the device dependent screen

coordinates. (Note, Southard's notation uses row vector notation so our presentation is the reverse order of his, and also he combines the image scaling scale, $\boldsymbol{S_{mag}}$, and $\boldsymbol{M_{proj}}$ into a single matrix which he labels $\boldsymbol{N_{proj}}$.)-

$\boldsymbol{M_{scr}}$ contains scales and translations [6, pg278] and is invertible. So:

$$M_{scr}\, S_{mag} = (M_{scr}\, S_{mag}\, M_{scr}^{-1})\, M_{scr} \qquad (2)$$

The latter equation is simply a scale about the center of the final window in screen coordinates. Assuming that all components of the viewing hierarchy are correctly measured, a scale about the screen window center is equivalent to a 2D scale of the projected image about the origin of Projection Plane Coordinate System (Figure 2).

We can describe the stereoscopic distortion induced by image scaling with the geometric construction illustrated in Figure 2. Figure 2 is an abstract diagram of a user viewing a stereoscopic display. The diagram is drawn from an overhead point of view looking down on the user. The user is represented by her eye points in blue and is looking at a display screen which is drawn as the vertical black line labeled 'Projection Plane.' The illustrated coordinate system is the Projection Plane Coordinate System. The projection plane is in the XY plane and the projection window is centered about the origin. (Note, Figure 2 only shows a portion of the projection window so the window does not actually appear centered in the diagram). A single modeled point on a virtual object is shown projected onto the projection plane once for each eye via its projectors in black. **I** is the eye axis center. **D** is the left eye displaced by vector **d**. **A** is the right eye displaced by vector **–d**. **E** is a modeled point on a virtual object. The modeled point is first projected on the projection window to points **H** and **G**. Image scaling then scales these points by factor *s* about the origin of the coordinate system. This yields points $\mathbf{H'}=s\cdot\mathbf{H}$ and $\mathbf{G'}=s\cdot\mathbf{G}$. These scaled points are the points actually displayed to the user. Ideally, the user's

natural visual system then reconstructs the perceived 3D point at location **F**. Unfortunately image scaling introduces a problem which is not evident in this 2D diagram: the red rays **AG′** and **DH′** do not generally intersect when the eye axis is not parallel to the projection plane. This is because the direct image manipulation violates the epipolar constraint of stereo imaging [22,pg 458]. To deal with this complication, we analyze the parallel case and the more general, non-parallel case separately.

First we verify that in the parallel case the rays **AG′** and **DH′** do intersect. Figure 3A illustrates the parallel case construction in 3D. **O** is the origin of the Projection Plane Coordinate System. **E** projects onto **H** and **G**. Since **AD** is parallel to the plane, **HG** is parallel to **AD**. (This occurs since for a line *l* (here **AD**) parallel to a plane *p* (here the projection plane), any plane *q* (here **ADGH**) containing *l* intersects plane *p* in another line parallel (here **HG**) to *l*.) Next, line **H′G′** is parallel to line **HG** because a uniform scale preserves angles. By transitivity **AD** is then parallel to **H′G′**. Hence there is a plane containing **AD** and **H′G′** and lines **AG′** and **DH′** are coplanar. Since coplanar lines intersect, **AG′** and **DH′** intersect. (Note, we are assuming a projective geometry where even parallel lines intersect at their ideal point). The construction defines a mapping on projective 3-space.

On the other hand in the *non*-parallel case the rays **AG′** and **DH′** typically do *not* intersect. This is best explained in terms of the epipolar geometry of stereoscopic imaging [22,pg 458]. Figure 3B illustrates the construction when the projection plane and eye axis are not parallel. Assume **G** and **H** are not collinear with **O**. Scaling **G** and **H** to **G'** and **H'** yields a line **G'H'** which is parallel to **GH** since uniform scales preserve parallelism. Since lines **G'H'** and **GH** are parallel, they have no points in common; in particular **G'H'** cannot intersect the epipole **J**.

So **G'H'** is not an epipolar line. Rather **G'** and **H'** lie on separate epipolar lines **JG'** and **JH'**. Hence **AG'** and **DH'** are in separate epipolar planes and are skew.

This lack of an intersection complicates analysis of the complete 3D distortion. However, there is an atypical subcase where **AG′** and **DH′** will intersect even when the eye axis is not parallel to the projection plane. The subcase occurs whenever an epipolar line intersects the origin, **O**. When scaled, points on such an epipolar line move along the same epipolar line preserving the epipolar constraint. So 3D points on an epipolar plane that contains the origin will not yield intersection problems. Unless the epipole coincides with the origin, only one such epipolar line and plane will exist. Section 5.2 will analytically examine a case where one epipolar line intersects the origin and will show that the planar geometric distortion does not preserve lines. Section 5.2 will also examine the vertical visual angles (VVA) between the epipolar planes of **G'** and **H'** and compare these values against human VVA fusion limits.

4.2 Image shifting

Image shifting translates the two projected stereo images toward each other. Image shifting is particularly effective when the scene only contains distant geometry behind the screen. The images can be translated so that the minimum positive parallax becomes zero. All other screen parallaxes in the scene will be reduced. If geometry exists in front of and behind the screen, image shifting may both reduce screen parallax for some modeled points and increase screen parallax for other modeled points. If done carefully, however, this technique can be used to reduce overall *maximum absolute* screen parallax. Figure 4 illustrates this possibility. In Figure 4A two eye points view three virtual points **A**, **B**, and **C**. The projectors for these points are drawn as black lines distinguished by different line styles. In Figure 4B, image shifting is applied to the projected images of these points. The translation vector is **V** for the left eye

image and –**V** for the right eye image. In this example we choose **V** to equal half the screen parallax of point **C**. As a result the new point, **C′**, now has zero screen parallax. Figure 4B shows the effect of this image shift on all three points. Note that while point **B**'s screen parallax increases, from zero to 2**V**, the overall maximum absolute screen parallax goes down. In Figure 4B absolute screen parallax of **A'** is smaller than that of **A** in Figure 4A. Hence image shifting can potentially be used to reduce fusion problems in scenes with geometry on both sides of the projection plane.

To understand how image shifting distorts displayed 3D space for a stereo HTD we must be careful about the direction of the translation. An intuitive choice is to translate parallel to the perpendicular projection of the eye axis onto the projection plane. Figure 6A shows this projected axis as a dashed-dotted line. Let **T** be unit vector on this axis pointing in the direction of the left eye and let τ be the magnitude of the desired translation. The left eye image is translated by vector $-\tau\mathbf{T}$ and the right eye image by $\tau\mathbf{T}$.

We can describe the stereoscopic distortion induced by image shifting with the geometric construction illustrated in Figure 5. Figure 5 is a highly abstract diagram of a user viewing a stereoscopic display. The color and labeling conventions follow Figure 2. Modeled point **E** is projected onto the projection plane to points **H** and **G**. Image shifting translates these to points $\mathbf{H'}=\mathbf{H}-\tau\mathbf{T}$ and $\mathbf{G'}=\mathbf{G}+\tau\mathbf{T}$. These are the points displayed to the user. Ideally **H′** and **G′** could also be generated by mapping modeled point **E** to displayed point **F** and then projecting**.** **F** would be at the intersection of the lines **AH′** and **DG′**. Unfortunately, the lines **AG′** and **DH′** only generally intersect if the eye axis is parallel to the projection plane. Due to this complication we analytically examine the parallel case and the non-parallel case separately in Section 5.

First, we verify that in the parallel case lines **AG′** and **DH′** generally intersect (Figure 6A). Since **AD** is parallel to the projection plane, all epipolar lines are parallel. So point **E**'s epipolar line **HG** is parallel to the projected eye axis (dashed-dot line). Because **HG** is parallel to the projected eye axis when **H′** and **G′** are shifted along **T,** they remain on the same epipolar line and plane. This guarantees the lines **AG′** and **DH′** intersect or are parallel. By treating 3-space as a projective space, this construction defines a mapping on 3-space.

In the non-parallel case (Figure 6B), however, epipolar line **HG** is not generally parallel to the projected eye axis along which the 2D images are translated. Consequently, applying the translations $-\tau\mathbf{T}$ and $\tau\mathbf{T}$ to the left and right eye images yields points not on the line **HG**. Typically, **G′** and **H′** will be on separate epipolar lines and planes which violates the epipolar constraint and leaves lines **AG′** and **DH′** skew. A rare exception occurs for the one epipolar line that happens to be parallel to the projected axis. Section 5.2 will show that in this case where **AG′** and **DH′** do intersect, the resulting distortion does not preserve lines. Section 5.2 also examines the vertical visual angles (VVA) between the epipolar planes of **G'** and **H'** and compare these values against VVA fusion limits from prior studies.

5 Analytic Descriptions

This section presents analytic descriptions of image scaling and image shifting based on the geometric constructions of Section 4. The detailed derivations are found in [37]. All the results were derived and/or verified using Mathematica, a commercial mathematics tool [40]. As mentioned, for both techniques it is necessary to distinguish between the parallel eye axis case and the more general, non-parallel eye axis case.

5.1 Parallel Case

Section A1 of [37] uses Mathematica [40] to obtain the analytic distortion of image scaling for an eye axis which is parallel to the projection plane. The resulting equation is parameterized on the central eye position, **I**, the vector to the left eye, **d**, and the scale factor *s*, as shown in Figure 2. Using column vector notation the matrix in Projection Plane Coordinates is:

$$\Delta_{sc}^{p} = \begin{bmatrix} s & 0 & 0 & 0 \\ 0 & s & 0 & 0 \\ 0 & 0 & s & 0 \\ 0 & 0 & \frac{s-1}{I_z} & 1 \end{bmatrix} \tag{3}$$

As a non-affine homology Δ_{sc}^{p} will preserve straight lines but not parallelism. Like α-false eye separation, Δ_{sc}^{p} contains no translation nor dynamic shearing components. Δ_{sc}^{p} also varies with head to screen distance, I_z, as does α-false eye separation. Δ_{sc}^{p} contains a static uniform scale component indicated by the repeating *s* along the diagonal. The authors observed that objects will appear not only to be compressed in depth, but also to change in overall size. For example a virtual box anchored to the display plane will shrink in all dimensions and appear as a tiny box. More generally, a modeled scene shown at true life-size may be shrunk and then appear to be a small toy model. In a sense, this also alters the effective field of view. For instance, imagine that in the original scene, a portion of the façade of a building fills up the screen. Image scaling may shrink the scene so that the user sees the entire building at a smaller scale and she may see around the sides of the building as well. Neither α-false eye separation, standard false eye separation nor image shifting have these effects. This empirical observation correlates with the fact that these other technique's distortion matrices either do not have a uniform scale component or the uniform scale component is not dependent on the technique's

fusion parameter (such as modeled eye separation or translation factor). Therefore while the other three techniques allow manipulation of the regular view scale factor independent of the fusion technique, image scaling does not. Note, that the presence or absence of a uniform scale is an affine transformation component. This is separate from the fact that when used for homogenous coordinate transformation, a 4 by 4 matrix is only unique up to a scale factor. The latter is true because when a 4 by 1 homogenous coordinate vector is multiplied by a matrix and then reduced to 3 space, the division by W cancels out any common multiplicative factor.

Section B1 of [37] uses Mathematica [40] to find an expression for the distortion induced by image shifting when the eye axis and projection plane are parallel. The distortion is parameterized on the eye axis center, **I**, the vector to the left eye, **d**, and the translation distance τ (Figure 5). Using column vector notation the matrix in Projection Plane Coordinates is:

$$\Delta_{sh}^{p} = \begin{bmatrix} Q & 0 & -\frac{I_x}{I_z}\tau & I_x\tau \\ 0 & Q & -\frac{I_y}{I_z}\tau & I_y\tau \\ 0 & 0 & Q-\tau & I_z\tau \\ 0 & 0 & -\tau/I_z & \tau+Q \end{bmatrix} \tag{4}$$

where

$$Q = \sqrt{d_x^{\,2} + d_y^{\,2}}$$

Δ_{sh}^{p} is a non-affine collineation and hence does not preserve parallelism. Δ_{sh}^{p} contains a translation component in the fourth column. This is to be expected since Figure 4 showed that points in the projection plane are moved out of the plane. This can be detrimental for applications which utilize the physical plane of the screen as a work surface for two-dimensional interactions such as precise curve drawing or laying route points on a map. Finally there are dynamic, head position dependent shearing components $-I_x/I_z\ \tau$ and $-I_y/I_z\ \tau$. Neither of these

artifacts occur in α-false eye separation nor image scaling. Also, note that while Δ_{sh}^{p} contains x,y and z scale factors only the z scale factor varies with the fusion parameter, τ, and the x,y factors only depend on head orientation. The authors observed that image shifting does not lead to the same overall uniform scaling effect as does image scaling although the depth dimension is noticeably altered by both techniques.

A pictorial comparison of the distortion artifacts of image scaling, image shifting, false eye separation and α-false eye separation is illustrated in Figure 7 and Figure 8. All these diagrams share the same color coding and format. Figure 7 illustrates how perpendicular (i.e. perpendicular to the projection plane) head motion effects the displayed 3D image for each fusion control method, while Figure 8 illustrates the effects of lateral head motion. All the diagrams are an overhead view looking down on the user. The projection plane is the middle black line. The eyes are at the top in blue. Instead of illustrating how a single modeled point is mapped to a single displayed point, the diagrams show how an entire modeled grid is mapped to a displayed grid. The modeled grid is square in black and the displayed grid is trapezoidal in red. While the system only presents the displayed grid (red), we overlay it with the modeled grid (black) for comparison. All images assume a true eye separation of 6.5 cm, a typical average value [14,p44]. In A and B, image scaling is applied with scale of 0.5. In C and D, α-false eye separation is used. The eye separation is underestimated by one-half its true value. The underestimating of the eye separation is illustrated by coloring the true eyes dark blue and modeled eyes light blue. In E and F, standard false eye separation is used. In G and H, image shifting is used with τ=1.625 cm. All fusion method parameters are set to reduce the nominal, maximum positive parallax of 6.5 cm to 3.25 cm, one half of its value. (Recall that in a properly calibrated system the maximum possible positive parallax equals the modeled eye separation.)

Figure 7 illustrates that all four techniques exhibit dynamic artifacts under perpendicular head motion. In contrast, in Figure 8 image scaling (A,B) and α-false eye separation (C,D) do not exhibit dynamic artifacts under lateral head motion but false eye separation (E,F) and image shifting (G,H) and do exhibit dynamic artifacts. The exact same parameters are reused as in the previous paragraph. Note that since these are 2D diagrams only a side to side head movement is shown, but similar results occur for head movement in any direction parallel to the screen.

In the parallel case, α-false eye separation and image scaling both have fewer dynamic distortion artifacts than image shifting and standard false eye separation. Image scaling yields dynamic distortion artifacts under the same head motions as α-false eye separation. However, image scaling induces a uniform scale component that can change the apparent size of the world and hence the amount of the scene seen by the user. In contrast, α-false eye separation has a much smaller effect on the relative field of view and it does not change noticeably. If an application designer desires a fusion control method to alter the displayed scene as subtly as possible then either false eye separation technique is a better choice than image scaling due to the lack of an overall scale. Additionally, if it is important to control the uniform scaling component independently of the view optic adjustment such as in [33] or [35], then false eye separation techniques are also a better choice since they allow more independent control of the uniform scale factor through the Platform Coordinate System scale.

5.2 Non-parallel Case

This section considers the analytic distortions in the general case when the eye axis is not parallel to the screen. First we discuss image scaling. As discussed in 4.1 in the non-parallel case, image scaling in general violates the epipolar constraint. An exception occurs when an epipolar plane contains the origin of the scale. This section first examines such a special case

when the modeled point and the eyes are in the XZ plane. This restriction is somewhat arbitrary but it allows simpler computation and serves to illustrate the induced curvature.

For the restricted XZ plane case, using Mathematica [40] section A2 of [37] finds the analytic distortion. The coordinate equations of the displayed point **F** are 2nd degree rational polynomials in terms of the modeled point **E** so image scaling maps lines in modeled space to curves in displayed space. Space limitations prevent the inclusion of the coordinate expressions in this text, but Figure 9 illustrates the distortion. Figure 9 is an abstract overhead view looking down on the user. The eyes are at the top in blue. The projection plane is the black horizontal line. Figure 9 shows the modeled black linear grid (black) is mapped to a curved displayed grid (red).

While the curvature is quite noticeable in Figures C and D is it less noticeable in Figures A and B. The authors viewed a similar grid on a desktop-VR display using a similar scale factor. We did not observe curvature. In part, this perhaps is because in situations like Figures C and D, one is too close to the screen to see anything anyway. When farther from the screen, as in Figures A and B, the curvature is more subtle. Our inability to perceive the curvature from these larger distances is also less surprising if we consider the following. While image scaling distorts 3D space curvilinearly, image scaling only distorts the individual 2D images by a scale which of course preserves lines. So the 2D projected image of a straight line must remain a straight line. Any curvature imparted to the displayed 3D curve exists solely in the depth, or the Z component. For example in Figure 10, a line *l* in modeled space is distorted into a displayed curve *l'* by such a distortion; however, the curvature of l'_{xy} is always 0. Humans can perceive a curved surface whose curvature is indicated purely by stereopsis. Julesz [11] illustrates a number of hyperbolic paraboloid and cosine surfaces whose curvature is indicated only by stereopsis cues. However perceiving this curvature probably involves a different mechanism than that used to distinguish

curves from straight lines drawn on a piece of paper. Additionally variations in shading and texture provide strong cues to surface shape and in VR environments these cues are based on the modeled geometry not the stereoscopically distorted geometry. Since we could not observe the induced curvature in a simple wireframe scene when we were explicitly looking for such curvature, we suspect that in a more complex, shaded scene a typical user would probably not notice these curvatures either. Clearly, however, only human factors studies over a wider variety of display configurations can more fully explore this issue.

We found similar results for image shifting. As discussed in 4.2 in the non-parallel case, image shifting in general violates the epipolar constraint. An exception occurs for the epipolar line parallel to the direction of the image shift. One such case occurs when the eyes and modeled points are restricted to the XZ plane. This restriction is somewhat arbitrary but it allows simpler computation. The coordinate equations of displayed point **F** are rational quadric ones in terms of modeled point **E** (see section B2 of [37]). However, the authors did not observe the curvature when viewing grids similar Figure 9 on a desktop-VR setup. We suspect this occurred for the same reasons cited in the previous paragraph concerning image scaling.

Potentially more troublesome is the fact that image scaling and image shifting violate the epipolar constraint by displacing homologous points vertically with respect to the eyes. This situation, called vertical parallax, aggravates image fusion problems because humans can fuse only a very small range of vertical parallax. Experiments with random dot stereograms show that if vertical parallax is slowly increased until fusion breaks down, the average limit for vertical visual angle (VVA) is only 20 minutes of arc. Additionally, once breakdown does occur VVA must be reduced back to 6 minutes of arc for fusion to reoccur [5][11].

First we determine the VVA for image scaling as shown in Figure 3B. **H** and **G** are the projections of a 3D modeled point onto the projection plane. **H'** and **G'** are **H** and **G** scaled about the window center, **O**. **J** is epipole. These points form epipolar planes **DJH**, **DJH'**, and **DJG'**. While **H** and **G** form one common plane, **H'** and **G'** lie in separate epipolar planes. The VVA separating **H'** and **G'** is the angle between epipolar planes **DJH'** and **DJG'** as measured about the line **AD**.

We will graph the VVA for image scaling using the following general approach which is illustrated in Figure 11. We pick some specific viewing configuration that includes eye positions, screen position and screen size. Next we pick some target plane in space with a fixed z coordinate and we compute the x and y extents of the region in this plane that is viewable from both the left and right eye. In Figure 11, this area is delimited by horizontal marks on the target plane. The purpose of using a fixed z coordinate is to help show how VVA varies across the screen which is parallel to the target plane. We project points on the target plane onto the projection plane once for each eye (**H** and **G** in Figure 3B) and apply the fusion control technique to these projected points (**H'** and **G'** in Figure 3B). We then compute the VVA for corresponding pairs of altered projected points. This yields a VVA value for each point on the target plane. Rather than plotting the VVA against the X and Y coordinates of the points on the target plane, we plot VVA against the X, Y coordinates of the left eye's altered projected 2D image point (**H'** in Figure 3B). This makes it easy to correlate VVA values with screen positions. The procedure for image shifting is quite similar.

The following examples assume an upright 2 x 2 meter screen. The eye separation is 6.5 cm. The eye axis is twisted 30 degrees looking towards the left of the screen while remaining in the XZ plane. The axis center is located at (0,0.524,1) in projection plane coordinates. This

accounts for a user standing 1 meter from the screen with eyes 1.524 meters (5 feet) off the ground. The target plane is at z = -10 (i.e. 10 meters behind the screen). Figure 11 shows an overhead view of the eye screen configuration. Note that distances are not to scale in order to make all relevant aspects legible. In order to approximate how much screen parallax may be generated beyond the recommended HVA limit, the traditional method assumes the eye axis is parallel to the screen and uses the perpendicular distance to the screen to map HVA to desirable screen parallax. In the above configuration, this yields a maximum possible positive screen parallax of 6.5 cm. An HVA limit of 1.57° then yields a desirable parallax of 2.74 cm. This implies an image scale factor of 0.4216 (desired/maximum) and image shift of 1.88 cm ((maximum-desired)/2). For arguments sake, we'll actually account for the rotated eye axis. Since the eye axis is not actually parallel to the screen it is no longer true that the maximum possible parallax equals 6.5 cm. By plotting parallax for very distant target planes ($z = -10^{13}$), we find the maximum possible parallax along the line of sight approaches 7.4 cm. As shown in Figure 11, at the given line of sight distance of 1.154 m (1/cos30°) to the screen, the recommended 1.57 degree HVA limit yields a desirable screen parallax of 3.65 cm. (HVA is measured within a given epipolar plane). This yields an image scale factor of 0.4938 and image shift of 1.87 cm. These values are more liberal than the values calculated with the traditional assumption since they are closer to the scale factor of 1.0 and a translation of 0. If these more liberal values yield VVA beyond the desirable VVA limits then so will the more stringent traditional values. We use the more liberal calculation.

Shortly we'll show that reducing screen parallax in the above viewing configuration using the calculated scale and shift values will create VVA values greater or equal to the VVA 20 arc minute limit. This is the acceptable threshold for vertical disparities found in random dot

stereograms [5][11]. An important question is whether such VVA values are displayable in pixels. Assume the screen is 1000 by 1000 pixels so that pixels are 2 mm tall. 20 arc minutes will subtend slightly different screen sizes depending what part of the screen is viewed because the eye to screen distance varies slightly. However, even at the closest distance of 1 meter, 20 arc minutes subtends 5.8 mm which is roughly 3 pixels. Hence, a VVA greater or equal to 20 arc minutes is significant pixel-wise. Figure 12A shows the VVA graph for image scaling along with a horizontal plane at the 20 arc minute threshold. The other axes, X and Y, are the coordinates on the 2 x 2 meter screen. A scale factor value of 0.4938 achieves the desired reduction of screen parallax. The shape of the plot is typical. The surface touches the VVA=0 plane along a valley that runs diagonally across the screen. This corresponds to the epipolar line which intersects the origin of the scale. As discussed earlier this is a special case where the epipolar constraint is preserved and hence VVA equals zero. Moving away from the valley along Y the VVA grows larger in either direction, but reaches a higher peak in the positive Y direction. VVA shrinks as the x-coordinate moves in the direction of head twist. In Figure 12A, the user's head is oriented to the left and the VVA shrinks as the x-coordinate moves to the left. This plot shows VVA meeting or exceeding the 20 arc minute limit for some y-position for every x-position. For the central view direction where x=0.5774, the VVA reaches 33 arc minutes at the top of the screen. The 6 arc minute re-fusion limit is far exceeded for an even larger portion of the space. Larger target plane distances such as z = -10000 km yield a very similar graph but with slightly steeper slopes and slightly higher VVA values; however, 10 m is sufficient to exceed the VVA threshold in the example configuration. Figure 12B graphs VVA for the same eye-screen configuration while using image shifting. The horizontal gray plane shows the 20 arc minute threshold. An image shift value of 1.87 cm is used. The surface touches

the VVA=0 plane along a valley that runs horizontally across the screen. This corresponds to the epipolar line which is parallel to the image translation direction. Moving away from the valley along Y the VVA grows larger rising at similar rates in either the positive or negative Y direction. VVA shrinks as the x-coordinate moves in the direction of head twist. For the central view direction where x=0.5774 (tan 30°), the VVA peaks at around 28 arc minutes.

Ideally geometric analysis would show either that image scaling and shifting yield acceptable VVA for all display configurations or that image scaling and shifting yield unacceptable VVA for all configurations. Unfortunately, this is not the case because VVA depends on many factors. There are a large number of independent variables including screen size, eye axis orientation, eye axis position, and target plane depth. As with any screen based visual angle, VVA shrinks with farther viewing distances. VVA grows with larger eye axis angles and with greater target plane distances. Algorithmic variations further complicate geometric analysis. An image scaling algorithm can use either a fixed scale factor or a variable scale factor. A variable scale factor could be chosen to either avoid crossed diplopia or uncrossed diplopia or both. Similarly an image shifting algorithm can use either a fixed translation factor or a variable translation factor. All these factors have to be considered in order to determine whether the image manipulation will geometrically exceed VVA limits for a particular application with a particular display configuration. These geometric results indicate that image shifting and image scaling could have detrimental effects in stereoscopic HTDs. Further empirical studies are needed in order to examine the subjective effects of such conditions on user viewing comfort.

In contrast to image shifting and image scaling, neither false eye separation nor α-false eye separation induce VVA at all. False eye separation moves the centers of projection along the baseline (eye axis). This has the effect of moving all the projected image points along their

original epipolar lines. Therefore, the epipolar constraint is always maintained. α-false eye separation simply transforms geometry by an additional 3D transformation, the α predistortion matrix [36]. This transform is applied to virtual space as a whole and not to the individual left and right eye images so it cannot create epipolar constraint violations.

6 Conclusions and Future Work

A VR system that interactively generates the presented imagery and displays it on a stereoscopic HTD has great control of all viewing parameters at the time of viewing. In contrast, in stereoscopic video and photography most of the viewing parameters cannot be altered once the photo is taken. Image shifting and scaling are primary mechanisms to alter the stereo image at view time in these non-interactive displays. Given the greater flexibility of VR, it makes sense to geometrically analyze image shifting and image scaling to see how the implied 3D distortion compares to and interacts with simpler view parameter adjustments such as view position and view scale as well as false eye separation methods. Again, in VR these simpler parameters can be manipulated at view time due to the interactive image generation.

This paper presented geometric and analytic descriptions of the theoretic distortion induced by image shifting and image scaling with respect to stereoscopic HTDs. These geometric descriptions were compared to those of false eye separation and α-false eye separation and we examined how the image methods interact with the simpler view placement and scale parameters. This paper's geometric analysis is an important step and guide to future human factor studies and application development.

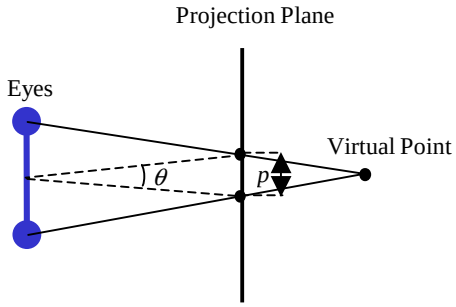


Figure 1: Illustration of the projection of a virtual point onto the projection plane for the two eyes of a user. The screen parallax is *p*, the distance on the screen between the stereo images of the virtual point. The horizontal visual angle is *θ*, the visual angle subtended by *p*.

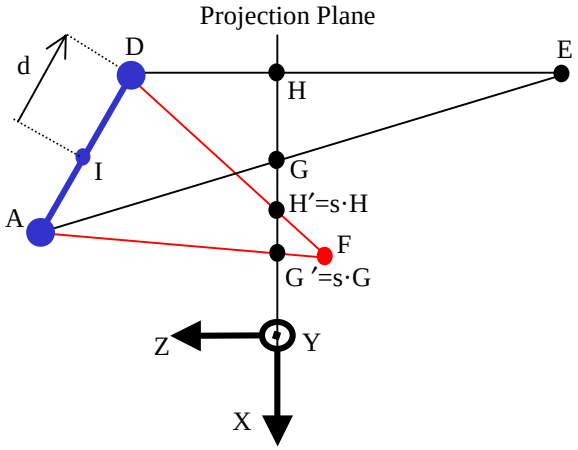


Figure 2: This is an abstract diagram of a user viewing a stereoscopic HTD. **I** is the center of the eye axis. **D** is the left eye, displaced by **d** from **I**. **A** is the right eye displaced by **-d** from **I**. The projection plane is embedded in the X-Y plane of the illustrated Projection Plane Coordinate System. **E** is a point of modeled geometry. It is projected onto points **H** and **G**. Image scaling scales **H** and **G** to points *s*•**H** and *s*•**G**. These altered image points could also be generated by projecting the equivalent displayed point **F.**

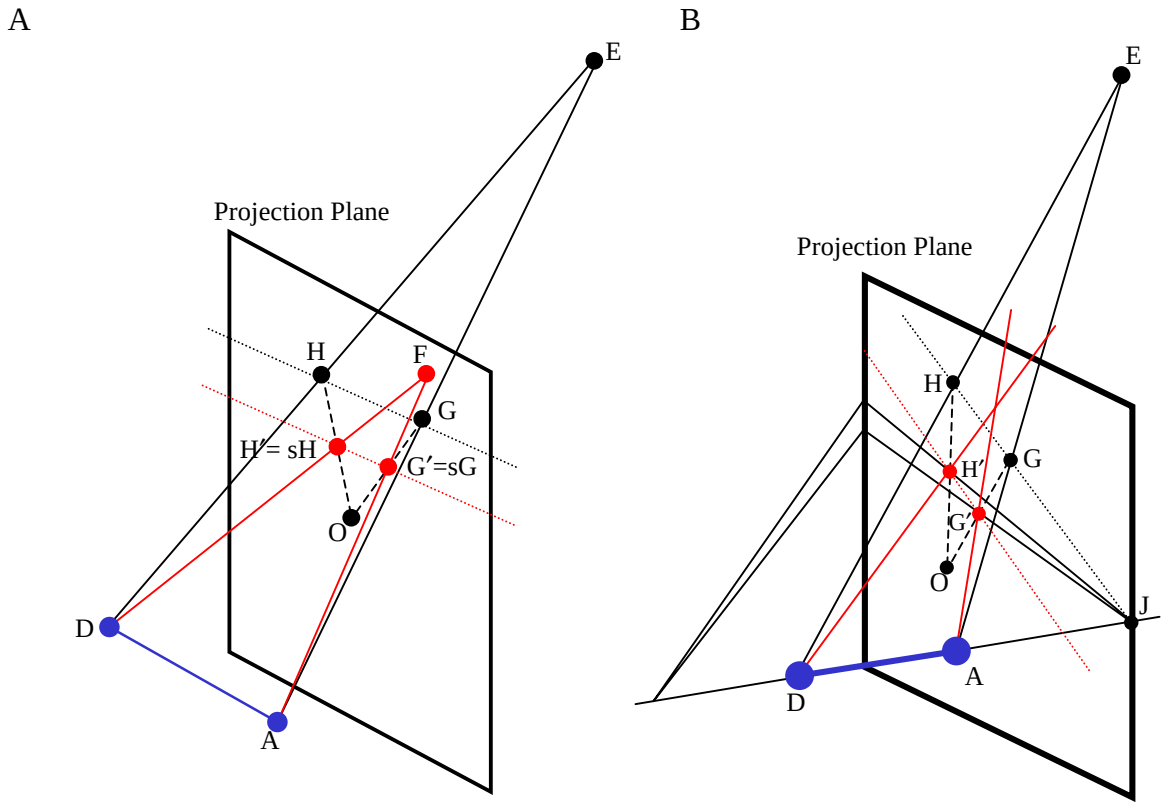


Figure 3: These figures illustrate the effects of image scaling for a modeled point **E**. The eyes, **A** and **D**, are in blue and the projected points are **H** and **G**. **H** and **G** are scaled about the center of the projection plane, **O**. The new points *s***H** and *s***G** are the points that the user sees. If the eye axis is parallel to the projection plane as in (A) the reconstructed point **F** is well-defined. If the eye axis is not parallel to the projection plane as in (B), the reconstructing rays (red) typically do not intersect. **H'** and **G'** are displaced by the vertical visual angle (VVA) measured between epipolar planes **DJG'** and **DJH'**. This leaves **F** undefined by this construction.

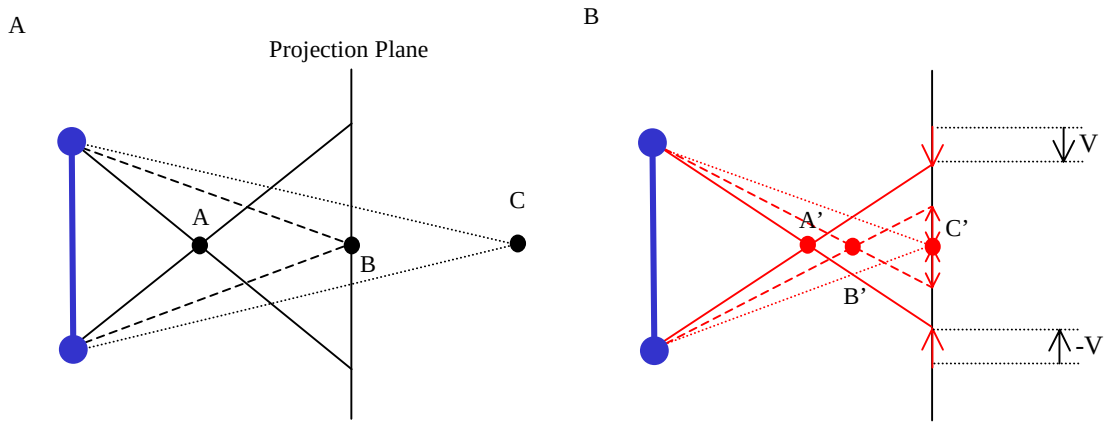


Figure 4: These figures illustrate how image shifting can be used to reduce the maximum absolute screen parallax of a scene. Figure A contains 3 points, **A**,**B** and **C**, at their modeled position with projectors indicating their left and right eye images. Figure B translates the points' left and right eye images by **V** and **–V** to yield different perceived points, **A′**, **B′** and **C′**. The maximum absolute screen parallax given by **A** is reduced.

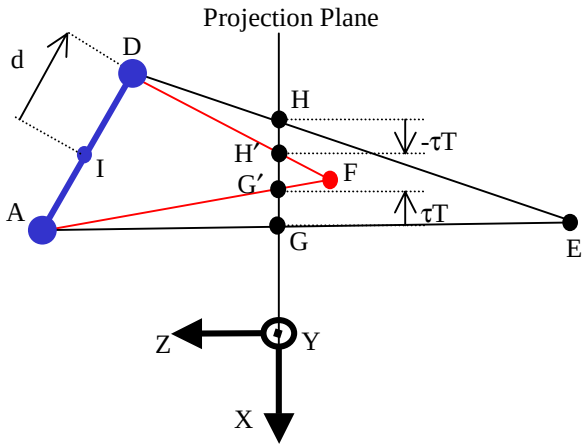


Figure 5: This figure illustrates the construction describing how image shifting is ideally equivalent to mapping a modeled point **E** to a displayed point **F**.

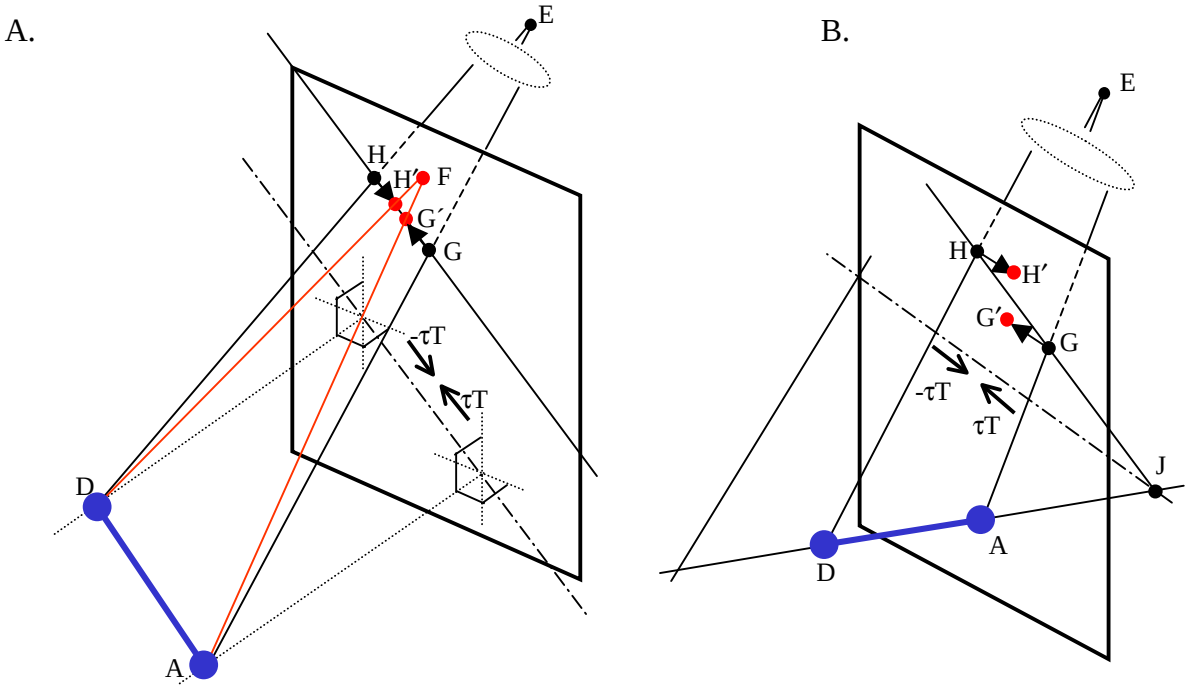


Figure 6: In image shifting, the images are shifted along -***τ*T** and ***τ*T** which are parallel to the perpendicular projection of the eye axis **AD**. This projected line is the dashed-dotted line. (A) shows that if the eye axis is parallel to the screen, then the shifted points **H′** and **G′** remain on the same epipolar line as the original points **H** and **G** and preserve the epipolar constaint (B) shows that if the eye axis is not parallel to the screen, the shifted points **H′** and **G′** end up on separate epipolar lines. This creates a vertical visual angle displacement which is the angle between epipolar planes **AJH′** and **AJG′**.

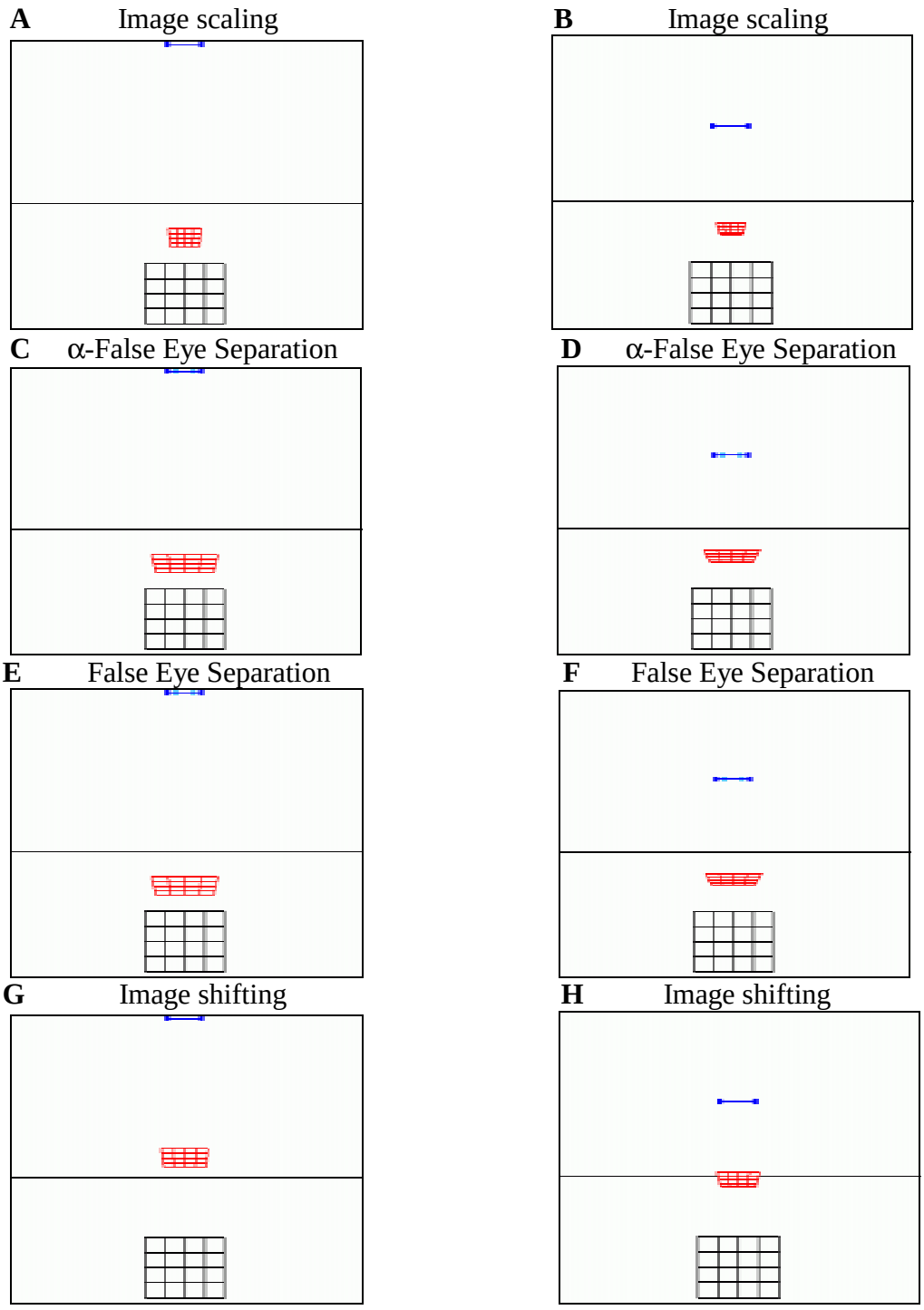


Figure 7: This figure compares the distortion due to forward/backward head movement for several image fusion control methods. Eye separation is 6.5 cm. (A) and (B) use image scaling with scale factor 0.5. (C) and (D) use α-false eye separation with eye separation ratio 0.5. (E) and (F) use false eye separation with eye separation ratio 0.5. (G) and (H) use image shifting (see Section 5) with translation magnitude 1.625 cm.

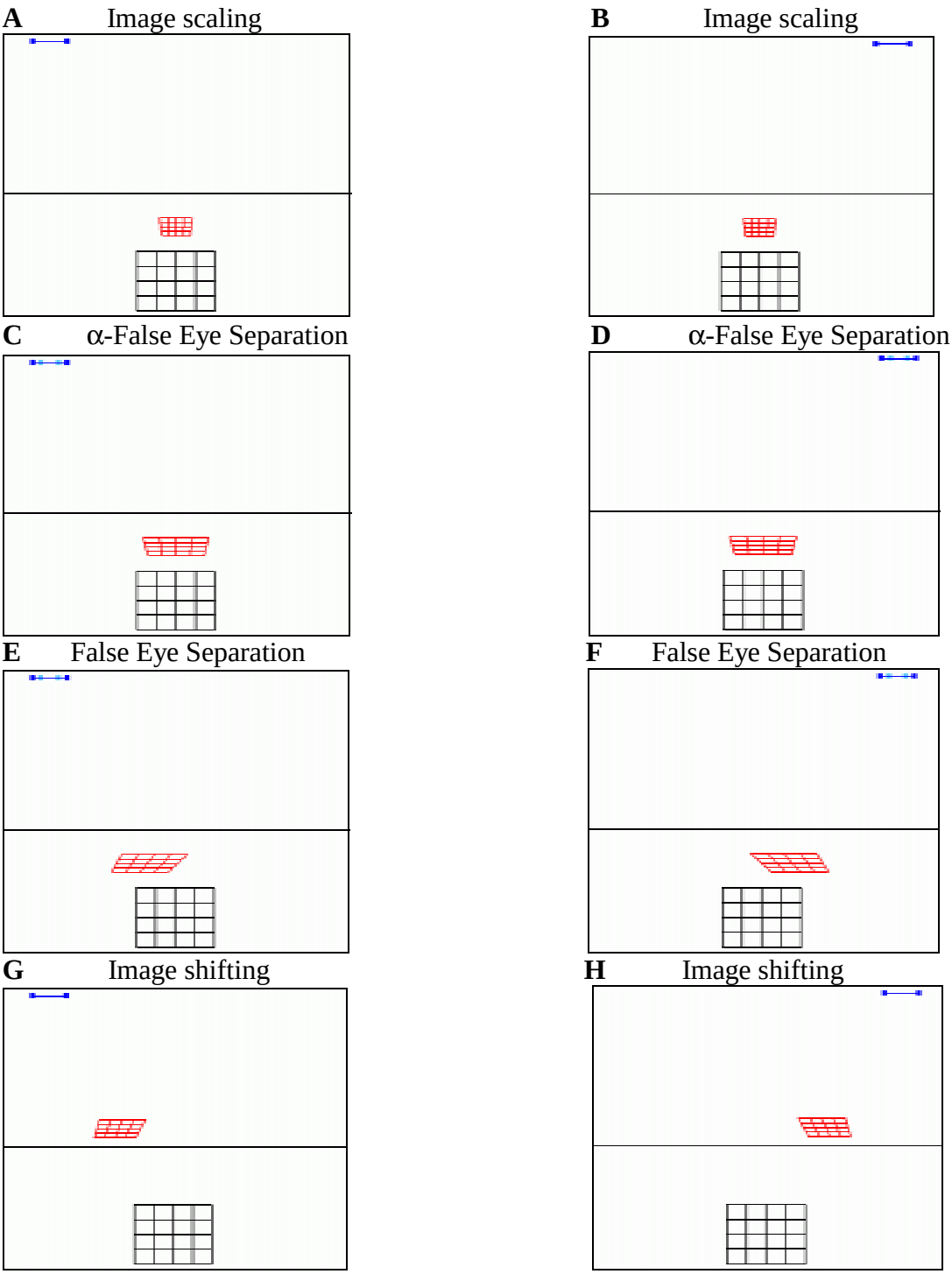


Figure 8: This figure compares the distortion due to left/right head movement for several image fusion control methods. (A) and (B) use image scaling with scale factor 0.5. (C) and (D) use α-false eye separation with eye separation ratio 0.5. (E) and (F) use false eye separation with eye separation ratio 0.5. (G) and (H) use image shifting with translation magnitude 1.625 cm.

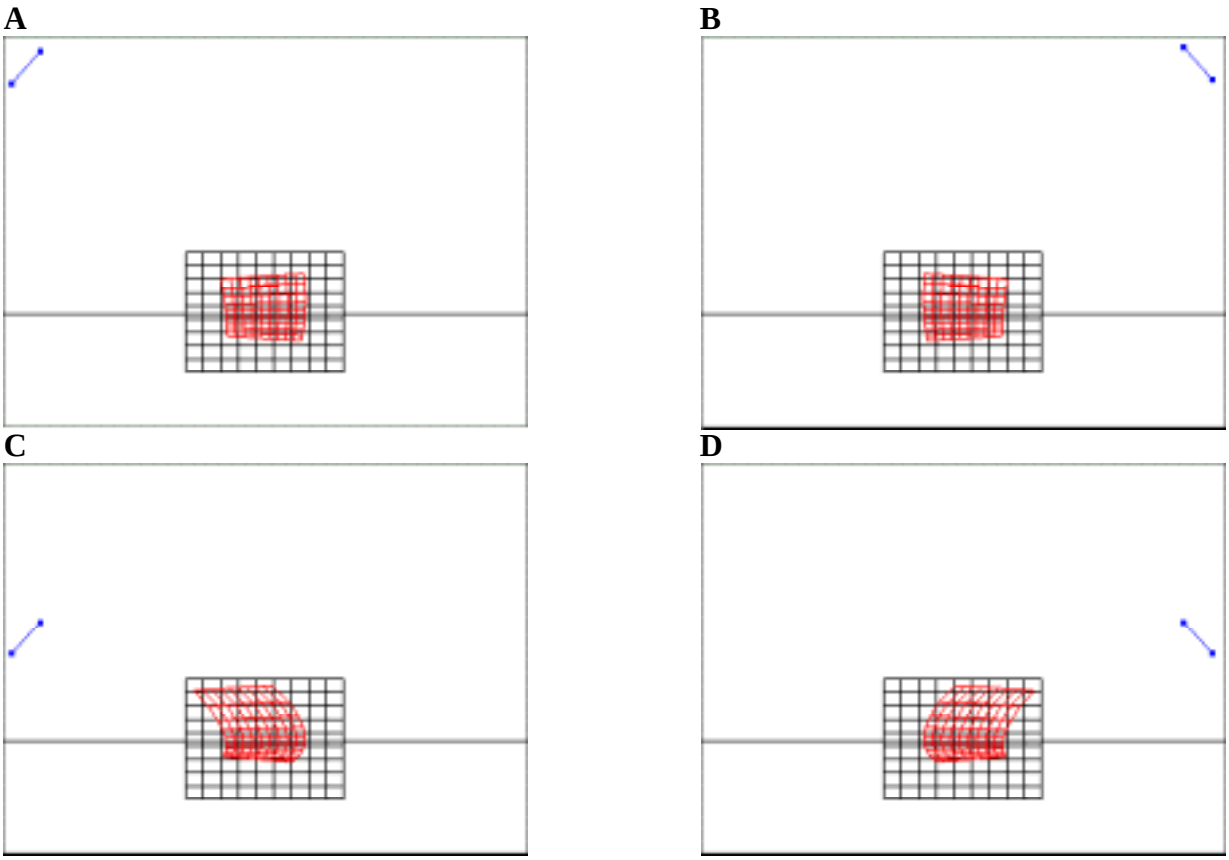


Figure 9: Illustration of the effect on the equivalent geometric displayed 3D image when image scaling is applied to an eye axis at an arbitrary orientation in the XZ plane. These diagrams are an overhead view looking down on the user. The projection plane is the black middle line. The eyes are in blue near to the top. The rectilinear black grid is the modeled grid while the curved red grid is the displayed geometry.

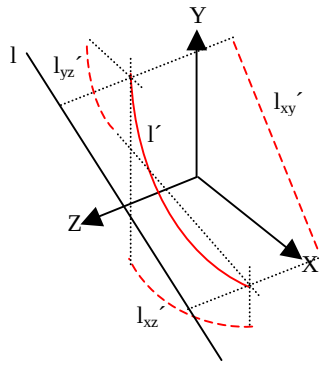


Figure 10: Some of fusion techniques map a modeled line *l* to the equivalent displayed curve *l´*. However, the straightness of l_{XY}´, the XY projection of the displayed line *l´*, is preserved. It is the straightness of the XZ and YZ projection (i.e., l_{XZ}´and l_{YZ}´) that is not preserved. The curvature exists only in depth.

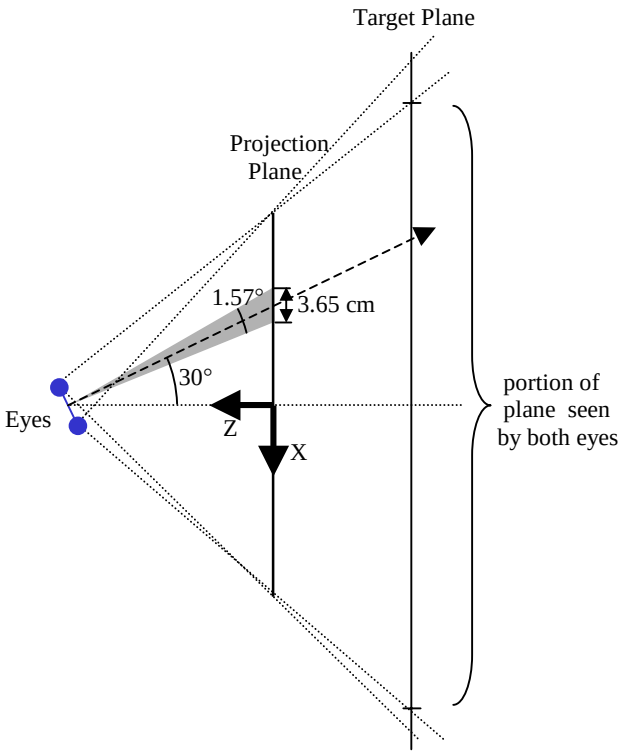
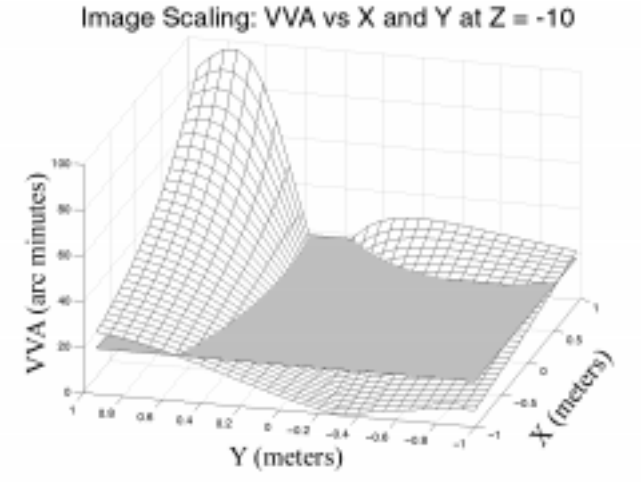


Figure 11**:** Illustration of the view configuration used to create a 3D plot of VVA for a given plane in model space. The plane of the diagram is assumed to be parallel to the XZ plane but at the height of the eyes (see text). This diagram is not drawn to scale.

A



B

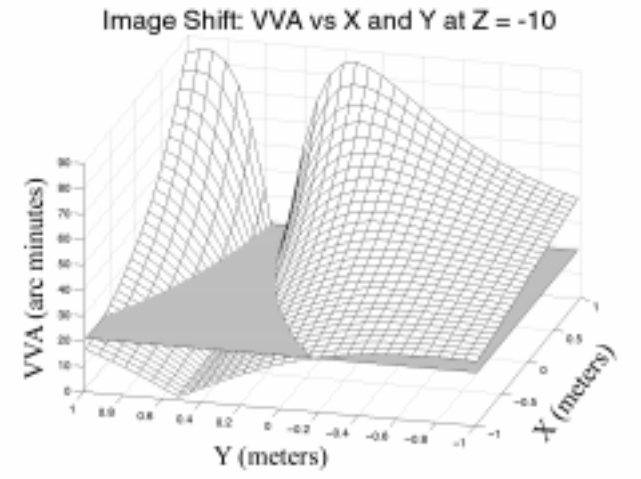


Figure 12: Plot of VVA for image scaling (A) and image shifting (B) for a plane of points in space at Z=-10.

References

[1] C. Cruz-Neira, D.J. Sandin, T.A. DeFanti, "Surround-screen projection-based virtual reality: the design and implementation of the CAVE," *SIGGRAPH 93 Conference Proceedings, Annual Conference Series*, ACM SIGGRAPH: Addison Wesley, Anaheim, CA, August 1993, pp. 135-42.

[2] J.E. Cutting, "How the eye measures reality and virtual reality," *Behavioral Research Methods, Instruments & Computers*, vol. 29, no. 1, Feb. 1997, pp. 27-36.

[3] E.T. Davis and L.F. Hodges, "Human Stereopsis, Fusion, and Stereoscopic Virtual Environments," *Virtual Environments and Advanced Interface Design*, Woodrow Barfield and Thomas A. Furness III, ed., New York/Oxford: Oxford University Press, 1995, pp. 145-175.

[4] M. Deering, "High Resolution Virtual Reality," *Computer Graphics (SIGGRAPH 92 Conference Proceedings)*, vol. 26, July 1992, pp. 195-202.

[5] D. Fender, B. Julesz, "Extension of Panum's Fusional Area in Binocularly Stabilized Vision," *Journal of the Optical Society of America*, vol. 57, no. 6, June 1967, pp. 819-831.

[6] J. D. Foley, A. Van Dam, S. K. Feiner, J. F. Huges. *Computer Graphics: Principles and Practice*. Addison-Wesley Publishing Company. 1992.

[7] Larry F. Hodges, David F. McAllister. Rotation algorithm artifacts in stereoscopic images. *Optical Engineering*. 29(8): 973-976, August 1990.

[8] L.F. Hodges, "Tutorial: Time-Multiplexed Stereoscopic Computer Graphics," *IEEE Computer Graphics and Applications*, vol.12, no.2, March 1992, pp.20-30.

[9] L.F. Hodges, E.T. Davis, "Geometric Considerations for Stereoscopic Virtual Environments," *Presence*, vol. 2, No 1, Winter 1993, pp. 34-43.

[10] L.F. Hodges, D. McAllister. Computing Stereoscopic Views. In David McAllister, editor, *Stereo Computer Graphics and Other True 3D Technologies*. Princeton University Press-Princeton, New Jersey. 1993.

[11] B. Julesz, *Foundations of Cyclopean Perception*, The University of Chicago Press. Chicago and London, 1971.

[12] J. Konrad, "Enhancement of viewer comfort in stereoscopic viewing: parallax adjustment," *IS&T/SPIE Symposium on Electronic Imaging, Stereoscopic Displays and Virtual Reality Systems*, San Jose, CA, 1999, pp. 179-190.

[13] W. Krüger. C.-A. Bohn, B. Fröhlich, H. Schüth, W. Strauss, G. Wesche, "The Responsive Workbench: A Virtual Work Environment," *IEEE Computer*, vol. 28. no. 7. July 1995. pp 42-48.

[14] L. Lipton, *Foundations of the Stereoscopic Cinema: A Study in Depth*, Van Nostrand Reinhold, 1982.

[15] D. McAllister, "Introduction," *Stereo Computer Graphics and Other True 3D Technologies*, David McAllister, ed., Princeton University Press-Princeton, New Jersey. 1993.

[16] M. Mon-Williams, J. P.Wann, and S. Rushton, “Design Factors in stereoscopic virtual-reality displays,” *Journal of the SID,* vol. 3, no. 4, Dec. 1995, pp. 207-210.

[17] W. Robinett, J.P. Rolland, “A Computational Model for the Stereoscopic Optics of a Head-Mounted Display,” *Presence*, vol. 1, no. 1, Winter 1992. pp 45-62.

[18] W. Robinett, R. Holloway, “The Visual Display Transformation for Virtual Reality,” *Presence,* vol. 4, no. 1, Winter 1995, pp. 1-23.

[19] J.P. Rolland and W. Gibson, “Toward Quantifying Depth and Size Perception in Virtual Environments,” *Presence*, Vol 4., No 1., Winter 1995, pp. 24-49.

[20] L.B. Rosenberg, “The Effect of Interocular Distance upon Operator Performance using Stereoscopic Displays to Perform Virtual Depth Tasks,” *IEEE Virtual Reality Annual International Symposium 93*, Seattle, WA, 1993, pp. 27-32.

[21] M. Siegel, Y. Tobinaga, and T. Akiya, “Kinder Gentler Stereo,” *IS&T/SPIE Symposium on Electronic Imaging, Stereoscopic Displays and Virtual Reality Systems*, San Jose, CA, 1999, pp. 18-27.

[22] Milan Sonka, Vaclav Hlavac, Roger Boyle, *Image Processing Analysis and Machine Vision*, Brooks/Cole Publishing Company, 1999,

[23] D.A. Southard, “Transformations for Stereoscopic Visual Simulation,” *Computer & Graphics*, vol. 16, no. 4, Winter 1992, pp.401-410.

[24] D.A. Southard, “Viewing Model for Virtual Environment Displays,” *Journal of Electronic Imaging*, vol. 4, no. 4, Oct. 1995, pp 413–420.

[25] R.T. Surdick, E.T. Davis, R.A. King, L.F. Hodges, "The Perception of Distance in Simulated Displays," *Presence*, vol. 6, no. 5, Oct. 1997, pp. 513-531.

[26] C. W. Tyler, "Induced Stereomovement," *Vision Research*, vol. 14, August 1974, pp. 609-613.

[27] N. A. Valyus, *Stereoscopy*, The Focal Press, London and New York, 1966.

[28] M. Wagner, "The metric of visual space," Perception & Pscyhophysics, 985, vol. 38, no. 6, pp. 483-495.

[29] C. Ware, K. Arthur and K.S. Booth, "Fish Tank Virtual Reality," In proceedings of *InterChi '93*, Apr. 1993, pp. 37-41.

[30] C. Ware, C. Gobrecht and M. Paton, "Algorithm for dynamic disparity Adjustment," *Proceedings of the SPIE - The International Society for Optical Engineering. Stereoscopic Displays and Virtual Reality Systems II*, vol. 2409, Feb. 1995, pp. 150-156.

[31] C. Ware, "Dynamic Stereo Displays," *CHI'95 Mosaic of Creativity*, ACM Press: Addison Wesley, Denver, CO, 1995, pp. 310-316.

[32] C. Ware, and D. Fleet, "Integrating Flying and Fish Tank Metaphors with Cyclopean Scale," Proceedings of Computer Graphics International, Hasselt and Diepenbeek, Belgium, 23-27 June 1997, pp.39-46.

[33] C. Ware, C. Gobrecht, and M.A. Paton, "Dynamic Adjustment of Stereo Display Parameters," *IEEE Transactions on Systems, Man and Cybernetics—Part A: Systems and Humans*, vol. 28, no. 1, Jan. 1998, pp. 56-65.

[34] Z. Wartell, L. Hodges, and W. Ribarsky. "The Analytic Distortion Induced by False-Eye Separation in Head-Tracked Stereoscopic Displays," Graphics, Visualization and Usability Center Tech Report 99-01, Computer Science Dept., Georgia Institute of Technology, Atlanta, GA, 1999. (see http://www.gvu.gatech.edu/gvu/reports/index.html)

[35] Z. Wartell, W. Ribarsky, and L. Hodges, "Third-Person Navigation of Whole-Planet Terrain in a Head-Tracked Stereoscopic Environment," *Proceedings of IEEE Virtual Reality '99 Conference*, IEEE Computer Society Press, Houston, Texas, 1999, pp. 141-148.

[36] Z. Wartell, L.F. Hodges, and W. Ribarsky, "Balancing Fusion, Image Depth, and Distortion in Stereoscopic Head-Tracked Displays," *SIGGRAPH 99 Conference Proceedings, Annual Conference Series*, ACM SIGGRAPH, Addison Wesley, Los Angeles, CA, 1999, pp. 351-357.

[37] Z. Wartell, L.F. Hodges, and W. Ribarsky. "A Geometric Comparison of Algorithms for Fusion Control in Stereoscopic HTDs: Supplementary Derivations," GVU Tech Report,GIT-GVU-01-08b, Computer Science Dept., Georgia Institute of Technology, Atlanta, GA, 2001. (see http://www.gvu.gatech.edu/gvu/reports/index.html)

[38] Z. Wartell, L.F. Hodges, W. Ribarsky, "Characterizing Image Fusion Techniques in Stereoscopic HTDs," to appear in proceedings of *Graphics Interface 2001*, Ottawa, Ontario, Canada 7-9 June 2001.

[39] S.P. Williams and R. V. Parrish, "New computational control techniques and increased understanding for stereo 3-D displays," *Stereoscopic Displays and Applications,* Proceedings

of the SPIE - The International Society for Optical Engineering, Santa Clara, CA, 1990, pp. 73-82.

[40] Mathematica, Wolfram Research Inc., 1988-1996.

[41] A. Woods, T. Docherty, R. Koch, "Image Distortion in Stereoscopic Video Systems," *Proceedings of the SPIE – The International Society for Optical Engineering, Stereoscopic Displays and Applications IV*, vol. 1915, 1993, pp. 36 – 48.

[42] Y.-Y. Yeh and L.D. Silverstein, "Limits of Fusion and Depth Judgements in Stereoscopic Color Displays," *Human Factors*, vol. 32, no. 1, Feb. 1990, pp. 45-60.

[43] S.Yoshida, S. Miyazaki, T. Hoshino, J. Hasegawa, T. Ozeki, T.Yasuda, S.Yokoi, "A technique for precise depth representation in stereoscopic display," Computer Graphics International, 1999. Proceedings , 1999.

Affiliation of Authors

Zachary Wartell: Zachary Wartell is a Ph.D. student in Computer Science at Georgia Institute of Technology within the Graphics, Visualization and Usability Center. He received his B.S. in Computer Science from the Georgia Institute of Technology in 1995. For several years he worked for the virtual environments group in the Human Interface Technology Center at NCR Corporation. His dissertation research analyzes and develops techniques for managing human-computer interface issues in stereoscopic HTDs.

Larry F. Hodges: Larry F. Hodges is Associate Professor in the College of Computing and Head of the Virtual Environments Group in the Graphics, Visualization and Usability Center at Georgia Tech. He holds a Ph.D. from North Carolina State University in Computer Engineering (1988), a M.S. in Computer Science from NCSU (1984), a M.A. in Religion from Lancaster Theological Seminary (1978), and a B.A. with a double major in Mathematics and Physics from Elon College (1974). Dr. Hodges research interests are in software and algorithm development, experimental quantification, and application development for virtual reality systems.

William Ribarsky: William Ribarsky is a Senior Research Scientist in the College of Computing and Associate Director for Service and Computing in the GVU Center. He received his Ph.D. in Physics from the University of Cincinnati. Dr. Ribarsky leads the Data Visualization Group, which has the charge of exploring general data and information visualization methods for a broad range of applications and in environments including workstation-based, collaborative, and virtual reality. He is a member of the IEEE Computer Society, the American

Physical Society, and on the executive committee of the IEEE Technical Committee for Computer Graphics.

Acknowledgements: Support for this work was provided by the National Science Foundation award number IIS-9977325 and by an Office of Naval Research AASERT contract N00014-97-1-0082. We thank Frank Jiang, Jeff Wilson and David Krum for useful discussions on various aspects of this work and we thank the anonymous reviewers for their useful comments and suggestions for improving this paper.