IEEE Trans on Systems Many and Cybernetics (in press)

DYNAMIC ADJUSTMENT OF STEREO DISPLAY PARAMETERS

Colin Ware, Cyril Gobrecht and Mark Paton
Faculty of Computer Science
University of New Brunswick
P.O. Box 4400 Fredericton, NB.
CANADA E3B 5A3 cware@unb.ca

Abstract- A review of the facts about human stereo vision leads to the conclusion that the human stereo processing mechanism is very flexible in the presence of other depth cues. Stereopsis seems to provide only local additional depth information, rather than defining the overall 3D geometry of a perceived scene. This paper reports on an experimental approach to adjusting stereo parameters automatically and thereby providing a low eye strain, easily accommodated stereo view for computer graphics applications. To this end the concept of virtual eye separation is defined. Experiment 1 shows that dynamic changes in virtual eye separation are not noticed if they occur over a period of a few seconds. Experiment 2 shows that when subjects are given control over their virtual eye separation, they change it depending on the amount of depth in the scene. Based partly on these results, an algorithm is presented for enhancing stereo depth cues for moving computer generated 3D images. It has the effect of doubling the stereo depth in flat scenes and limiting the stereo depth for deep scenes. It also reduces the occurrence of double images and the discrepancy between focus and vergence. The algorithm is applied dynamically in real time with an optional damping factor applied so the disparities never change too abruptly. Finally Experiment 3 provides a qualitative assessment of the algorithm with a dynamic “flight” over a digital elevation map.

I. INTRODUCTION

When observing a mountain range at a distance of 30 km stereo vision contributes almost nothing to our understanding of the spatial shape. However, if we create a stereo pair of images with the viewpoint separated by 5 km we will obtain a useful enhanced "hyper stereo" image. This technique is, used extensively in stereo photogrammetry. Conversely, if we wish to understand the spatial structure of an object only meters away, we will be better off with our normal eye separation.

Consider a display in which the stereo image is based on the projection of a 3D object onto a screen from a particular viewpoint. Given that the image is actually viewed from that viewpoint, the result is a geometrically correct perspective view of some object in front of or behind the screen (at least if we ignore screen curvature [1]). We can do this for each of the two eyes given a time multiplexing display such as CrystalEyes^{TM} glasses, coupled with an SGI workstation. The result is a correct stereo view of some virtual scene. This paper develops the hypothesis that dynamically adjusting certain stereo parameters may improve the stereo view, reduce the occurrence of double images and reduce eye strain due to the vergence-focus discrepancy. Our goal has been to achieve a system in which we can "fly through" an artificial 3D computer graphics environment while maintaining a good stereo image. Before discussing the problem in more detail we introduce some of the basic geometry and terminology of stereo display systems.

Figure 1 illustrates the simplest stereo display. Both eyes are fixated on the vertical line **a.** A second line **b** is closer to **a** in the left eye's image than in the right eye's image. The brain resolves this discrepancy by perceiving the lines as being at different depths as shown.

Retinal disparity is the difference between the angular separation of **a** and **b** at the two eyes (disparity $= \alpha - \beta$). *Screen disparity* is the distance between parts of an image on the screen. *Vergence* is the degree to which the two eyes converge to fixate a target (this is also called phoria).

If the disparity between the two images becomes too great then *diplopia* occurs. Diplopia is the appearance of the doubling of part of a stereo image when the visual system fails to fuse the images. The area near the convergence point is called *Panum's Fusional Area.* In the worst case Panum's fusional area has remarkably little depth. At the fovea the maximum disparity before fusion breaks down is only one tenth of a degree, whereas at 6 degrees eccentricity (of the retinal image from the fovea) the limit is one third of a degree [2]. However, the size of Panum's fusional area is highly dependent on a number of visual display parameters such as the exposure duration to the images and the size of the targets. Moving targets can be fused at greater disparities. Depth judgments can be made outside of the fusion area, although these are less accurate.

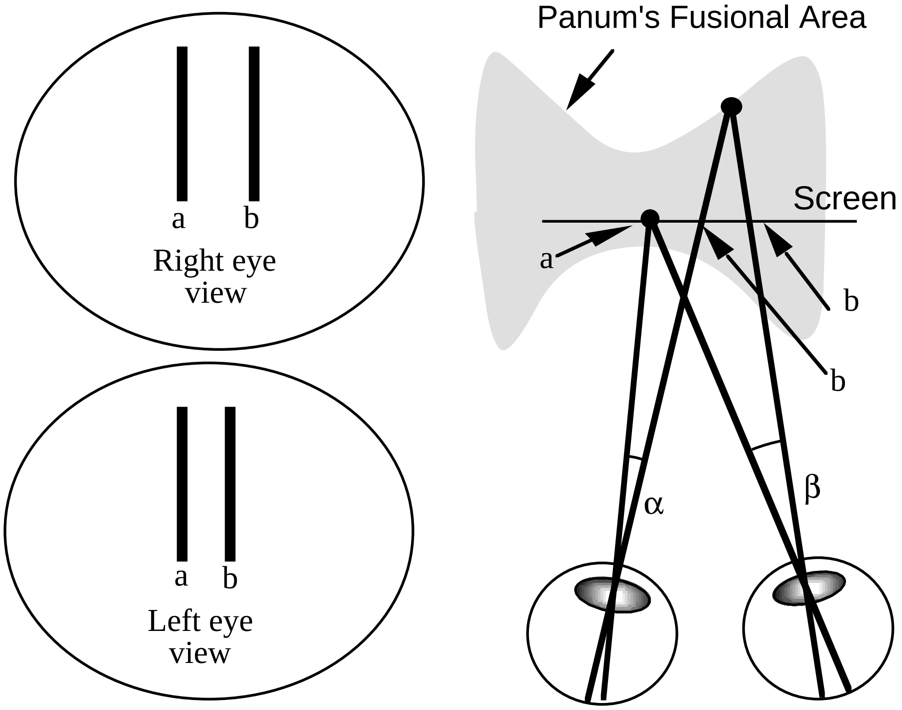

Figure 1. A simple stereo display. Different images for the two eyes are shown on the left. A top down view shows how the brain interprets this display. The vertical lines a and b are projected onto the screen as shown.

A. Virtual Eye Separation

Figure 2. illustrates the concept of virtual eye separation and demonstrates how the apparent depth of an object decreases if the virtual viewpoint uses a wider eye separation than the actual viewpoint. We consider only a single point in the virtual space. If E_V is the virtual eye separation and E_a is the actual eye separation of some observer, then the relationship between depth in the virtual image (z_V) and in the viewed stereo image (z_S) is a ratio.

$$\frac{E_v}{E_a} = \frac{z_s(z_v + S)}{z_v(z_s + S)} \qquad (1)$$

where S represents the distance to the screen. Rearranging terms we can get the stereo depth expressed as a function of the virtual depth and virtual eye separation.

$$z_s = \frac{SE_v z_v}{E_a z_v + E_a S - E_v z_v} \qquad (2)$$

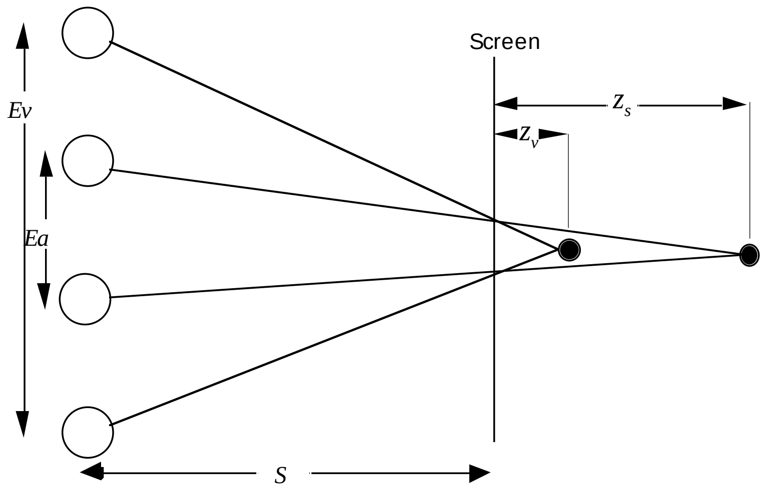

Figure 2. The geometry of virtual eye separation

If the virtual eye separation is larger than the actual eye separation, stereo depth is increased. If the virtual eye separation is smaller than the actual eye separation, stereo depth is decreased VR designers often try to achieve $E_v = E_a$ for "correct viewing." When $E_v = 0.0$, both eyes get the same image, as in single viewpoint graphics. Note that stereo depth and perceived depth are not always equal. The brain is an imperfect processor of stereo information and other depth cues may be much more important in determining the perceived depth.

B. Other Depth Cues and Conflicts

Occlusion is one of the major depth cues. In a perspective view, close objects occlude (i.e. cover up) more distant objects. This can cause a problem in stereo display. When disparity information causes an object to appear in front of a screen display, the edge of the screen may appear to occlude the object. Since occlusion is the stronger depth cue, the conflict is resolved perceptually in favor of occlusion, destroying the illusion of depth. Valyrus [3] called this the frame cancellation effect.

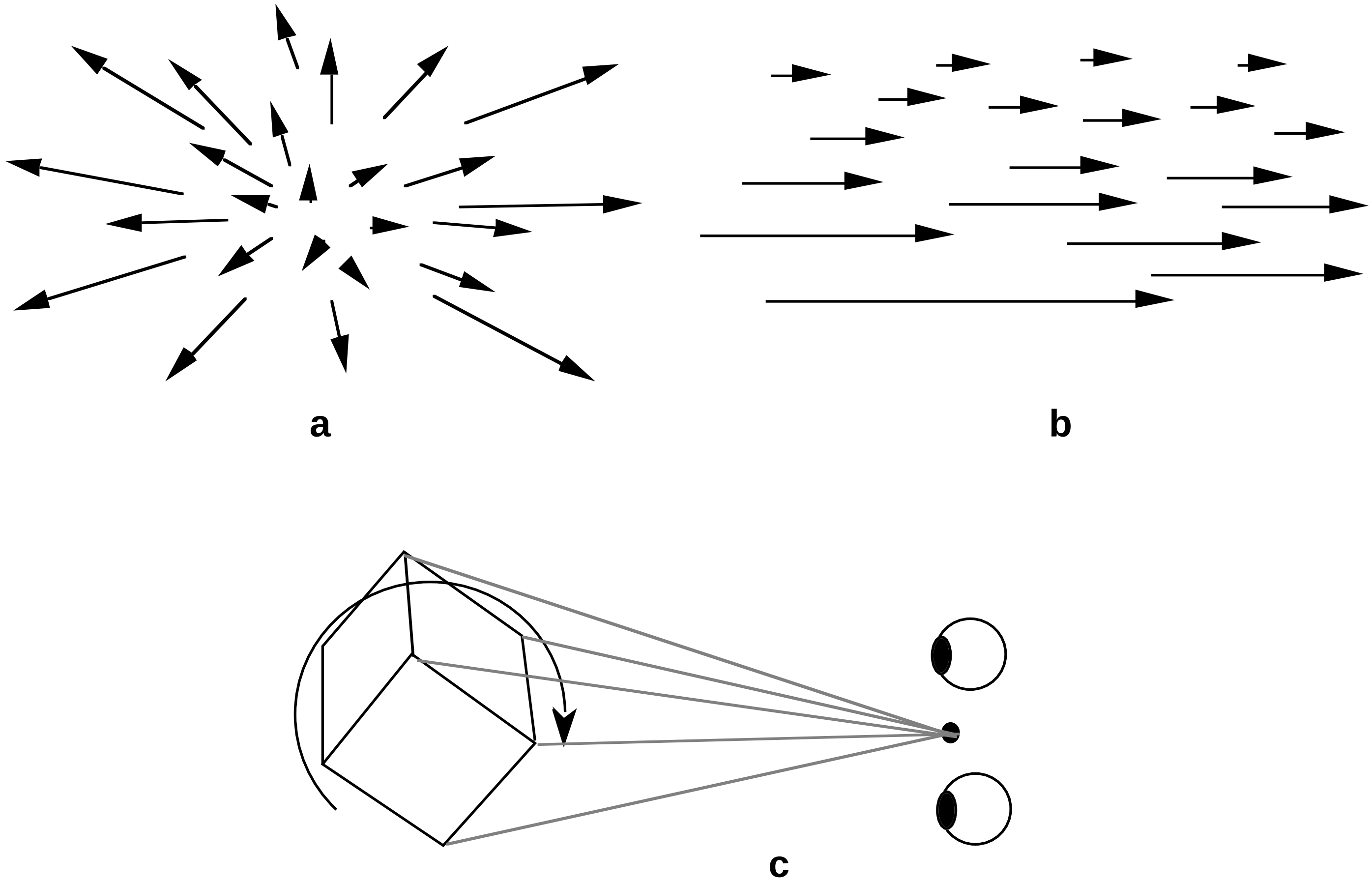

Figure 3 Different kinds of structure from motion information. (a) The vectors illustrate the kind of visual flow field that results from forward motion through a 3D environment. (b) The vectors illustrate the motion parallax that results from lateral translation with respect to a horizontal surface. (c) When the image of a rotating wire frame object is projected onto a flat screen the 3D shape is perceived. This is called the kinetic depth effect.

One of the most important depth cues comes from the dynamic flow of information across the retina [4]. When we are driving along a highway, we have a very strong sense of space although most objects are beyond the range at which stereo disparity is an effective depth cue. Figures 3a and 3b illustrate motion parallax. Figure 3a shows the visual flow field that results from forward motion. Figure 3b shows a motion gradient like that which occurs when looking sideways out of a moving vehicle. The top parts of the scene move slower as indicated by the shorter arrows . The whole scene is perceived as a horizontal carpet moving laterally with respect to the observer. Figure 3c illustrates a related dynamic depth cue called the Kinetic depth effect [5,6]. When a rotating wire frame object is projected onto a screen an observer perceives a three-dimensional cube rather than a set of connected lines expanding and contracting. For this information to be interpreted spatially, the perceptual system must assume that the lines correspond to a rigid 3D structure. Motion parallax and kinetic depth cues both come under the general heading of *structure from motion* (SFM) cues.

We are interested in automatically adjusting parameters such as virtual eye separation. Therefore the following question arises: what happens in moving scenes with constantly changing stereo depth parameters? Users should perceive a rubbery distortion of the scene if the brain is a perfect stereo geometry processor. Alternatively, other depth cues may dominate so that changes in stereo depth are not noticeable.

Stereo disparity provides relative depth information that must be calibrated relative to a perceived fixation distance and this calibration is a dynamic process [6]. One possibility is that vergence information provides an approximate calibration for stereo depth [9,10]. Structure from motion provides accurate shape information, assuming object rigidity, but the near/far relations between the object and the observer are not specified [5]. Hence depth in a kinetic depth display can spontaneously reverse. An extreme hypothesis presented by Ogle [7] proposes that stereopsis provides non-metric, ordinal depth information, while SFM and other cues provide spatial layout. This hypothesis is countered by evidence that stereopsis affects the degree of apparent depth in a SFM display [8]. Overall, it seems clear that depth derived solely from disparity information is inaccurate [9], and can be recalibrated dynamically [10,11,12], although ordinal information provided from stereo can be extremely precise. This suggests that dynamically changing stereo depth parameters is probably viable and certainly worth studying.

C. The Vergence Focus Problem

When we fixate objects at different depths, two things happen: the convergence of the eyes change (called vergence) and the focal length of the lenses in the eyes accommodate to bring the objects into focus. The vergence and the focus mechanism are coupled in the human visual system. If one eye is covered the vergence and the focus of the *covered* eye changes as the uncovered eye focuses on objects at different distances. This illustrates vergence being driven by focus. The converse also occurs; a change in vergence can drive a change in focus.

In a stereoscopic display all objects lie in the same focal plane regardless of their apparent depth. However, accurate disparity and vergence information may fool the brain into perceiving them at different depths. Screen based stereo displays provide vergence and disparity information but no focus information. There is some evidence that the failure to correctly present focus information may cause a form of eye strain presumably because of the coupling described above [12,13]. This problem is present in both stereoscopic head-mounted systems and monitor-based stereo displays.

In view of the above observations, how may we reduce the problems associated with the decoupling of focus and vergence in stereo displays? One solution is to reduce screen disparities. Valyrus [3] found experimentally that the focus and vergence discrepancy should not be more than 1.6 degrees. He proposed that the screen disparity should be less than 0.03 times the distance to the screen. Veron et al. [14] used this formula to derive the guideline that screen based stereo displays should be placed 2.3 meters from the viewer to give fusable images assuming a maximum eye separation of 6.9 cm and that virtual objects are always behind the screen.

Based on a different analysis of the problem Williams and Parrish [15] concluded that a practical viewing volume falls between -25% and + 60% of the viewer to screen distance. They proposed a method whereby objects at different depths can optimally use the available disparity range. Specifically they showed how objects at two or more different distances can be brought into the useful viewing volume. Their scheme parcels out the available disparity so that depth ranges containing objects are enhanced stereoscopically, while others are reduced. For example, in a scene with two objects, the distance between the front and back of each object is allocated a large disparity range, while the empty space between them is made devoid of disparity. The algorithm presented in the latter part of the present paper represents another solution to the same problem for dynamically changing computer graphic displays.

D. The Flexibility Of The Stereo Processing System

The evidence presented in the preceding sections suggests that the stereo processing system in human vision is extremely flexible in the presence of other depth cues. Perception of large scale 3D space is derived from depth cues such as occlusion, shape from motion and linear perspective. Stereo disparity primarily provides localized information about relative depths. If this theory is correct, it may be reasonable to devise algorithms that dynamically adjust disparity information to optimal values for a particular situation, even when this produces conflicts with other depth cues. We reason that the other cues will dominate depth perception on the large scale and that changes in stereo parameters will hardly be noticed, or at least will not be objectionable.

Such adjustments provide several advantages. We may be able to set optimal stereo viewing conditions regardless of the size of the virtual scene we can reduce vergence focus conflicts by reducing the screen disparities. In addition, we can simultaneously reduce the incidence of double images.

In this paper we first present two experiments designed to explore the usefulness of dynamic stereo adjustments. Following this we present our algorithm to dynamically set the effective eye separation during continuous motion of the viewpoints through a computer generated scene. This is done in a manner that both optimizes disparities and reduces vergence focus conflicts (and hence eyestrain). Finally we present an experimental evaluation of the algorithm.

II. EXPERIMENT 1: RATE OF CHANGE OF VIRTUAL EYE SEPARATION

The perceptual motor system is capable of re-calibrating the disparity and vergence depth cue mechanism in the presence of other depth cues, such as motion parallax. Consequently, the disparity mechanism is insensitive to low frequency change. The following study determines how rapidly virtual eye separation can be changed before it is noticed in a display with SFM and stereo depth. Changes in virtual eye separation should result in a sensation of changing depth if the brain were to rely primarily on disparity information but this would be in conflict with the rigidity assumption given the linear perspective and motion flow information. The experiment is exploratory in nature - we had no a priori hypothesis about the results.

A. Equipment Used

Experiments 1 and 2 used an Indigo^2 Extreme graphics workstation with the Cyberscope™. A Cyberscope consists of a hood that can be placed over a small monitor allowing for the stereo viewing of properly constructed images. The Cyberscope uses front surface mirrors to displace and rotate the images presented to the two eyes as shown in Figure 4. The image obtained using the Cyberscope was 25 cm wide and 15 cm high and the viewing distance was 33 cm.

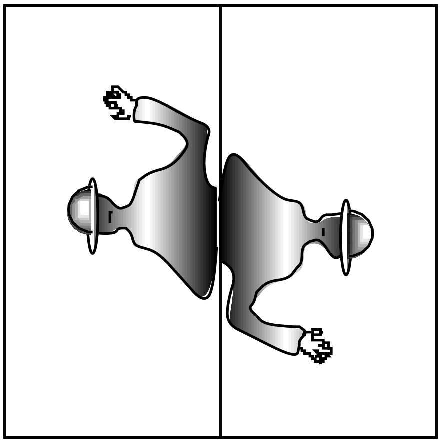

Figure 4. The Cyberscope optically rotates the images from the two halves of the screen, 90 deg clockwise and counter clockwise respectively, and superimposes them. This is done using front surface mirrors to provide optical clarity.

Figure 5. A stereo Cyberscope image of the moving carpet display used in Experiments 1 and 2.

B. Method

A scene was constructed in which a moving carpet dotted with truncated pyramids moved perpetually towards the observer. An image configured for the Cyberscope is illustrated in Figure 5. The pyramids were 5 cm wide at the base and 2.5 cm high. The depth of the plane was 4 meters and the angle to the vertical plane was 70 deg.

On each trial the frequency at which the virtual eye separation oscillated was gradually increased until it became noticeable. At this point the subject pressed a mouse button and the next trial was started after a two second interval. This method was chosen because it most closely mimicked the practical application environment where normally the virtual eye separation would not change rapidly. The acceleration was 0.04π radians per sec^2. This means that at 50 seconds the virtual eye separation was being changed at 1 Hz; at 100 seconds the frequency was 2 Hz, and so on. There was also a random temporal offset to the start of oscillation so that subjects could not anticipate this in their responses. There were nine viewing conditions, three values for virtual eye separation: 6.3 cm, 4.2 cm and 2.1 cm. with three different amplitudes of oscillation applied to each: 10%, 20% and 30%. For example, a 10% amplitude applied to a 6.3 cm eye separation meant that the eye separation varied sinusoidally between 6.3 cm and 90% of 6.3 cm (=5.67cm). There were also three randomly chosen practice trials given at the start of the session following which the entire set was given in a random. The trial block was repeated twice. The design was within subjects, fully randomized repeated measures.

Subjects: Nine subjects who were all undergraduate or graduate students were used as observers. All could see depth in Julesz stereograms. Five of the subjects were paid to participate, the rest were unpaid volunteers.

C. Results

The mean frequency at which subjects detected the oscillation is plotted in Figure 6. An analysis of variance revealed a significant effect of amplitude $F(2,72) = 7.97$, $p < 0.01$, and virtual eye separation $F(2,72) = 13.06$, $p < 0.01$. The frequency that is detectable varies inversely with the amplitude of oscillation, and virtual eye separation. There is no significant interaction between amplitude and virtual eye separation. The worst case is for the maximum amplitude and the maximum eye separation, in which case the average frequency at which the oscillation is detected is 0.3 Hz. This corresponds to a period of 3.3 seconds.

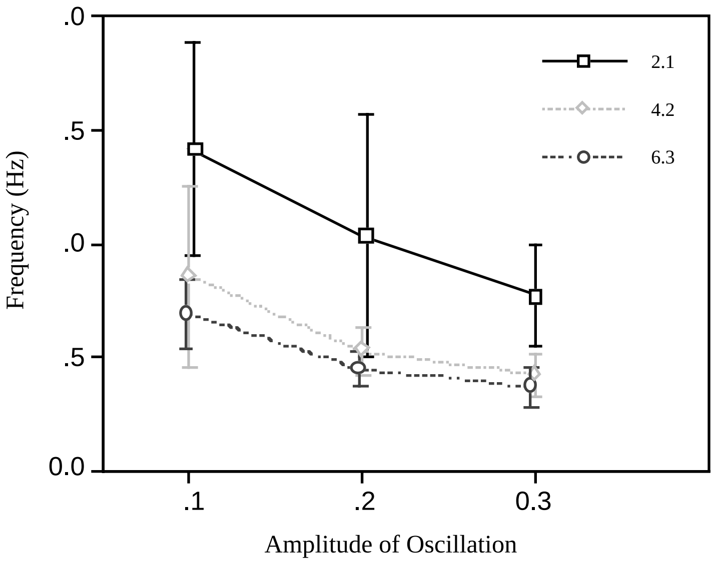

Figure 6. Results from Experiment 1 show how the frequency at which oscillating eye separations are detected varies with different amplitudes and virtual eye separations. The virtical bars represent 95% confidence intervals.

One possible hypothesis is that threshold is determined by the maximum rate of change of virtual eye separation. If this were the case there should be no difference between the conditions if the response thresholds were transformed into velocities. To test this hypothesis we transformed the data from frequencies to peak velocities using the formula.

$$V_p = 2\pi f a E_v \qquad (3)$$

Where V_p is the peak velocity, f is frequency, a is the amplitude and E_v is the virtual eye separation. The data from Experiment 1, transformed into velocities is plotted in Figure 7. An analysis of variance revealed that even after this transformation there are significant effects of amplitude F(2,72) =6.69, $p < 0.01$, and virtual eye separation F(2,72) = 4.15, $p < 0.05$. Thus peak velocity is not the primary determinant of response thresholds. However, this analysis does suggest that in practical situations the rate of change of virtual eye separation should be kept below 0.2 cm/sec.

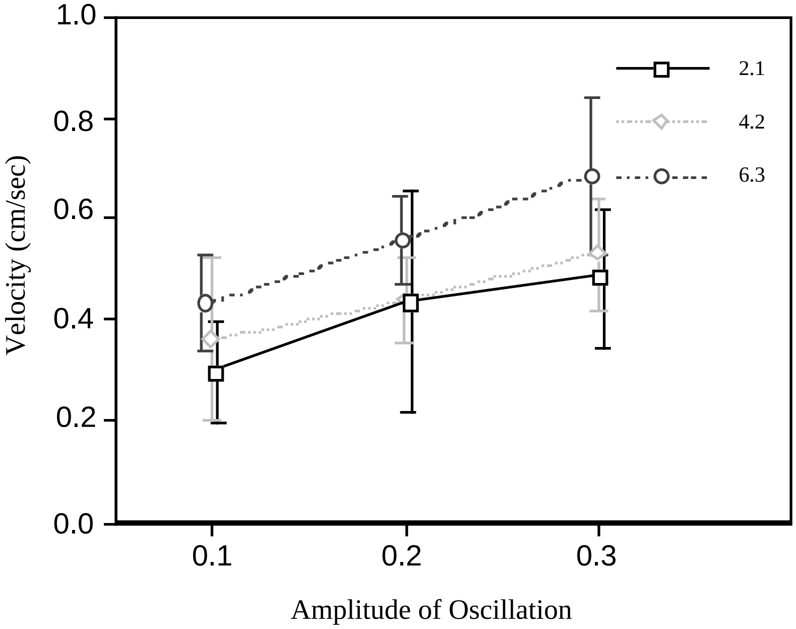

Figure 7. Results from Experiment 1 showing the data transformed to threshold peak velocity. The virtical bars represent 95% confidence intervals.

Although we have not been able to provide a theory based explanation for these data, we believe that their practical significance is considerable. This remarkable tolerance to changing disparities means that eye separation can be changed dynamically with a moving scene as long as it is done gradually over the course of several seconds.

III. EXPERIMENT 2: ADJUSTING VIRTUAL EYE SEPARATION

The second experiment was designed to determine the potential for changing virtual eye separation in terms of the depth present in the scene. Subjects were given control over the virtual eye separation and instructed to increase the eye separation until diplopia occurred and then move it back to a comfortable value. They were asked to use this procedure to set "the maximum comfortable setting" . The idea was to provide a set of dynamically changing scenes containing different amounts of depth and see if subjects behaved in some consistent manner. Our hypothesis was that subjects would vary their eye separation depending on the depth in the scene. We hoped to use this data to generate a method for automatically changing the virtual eye separations to suit different scenes.

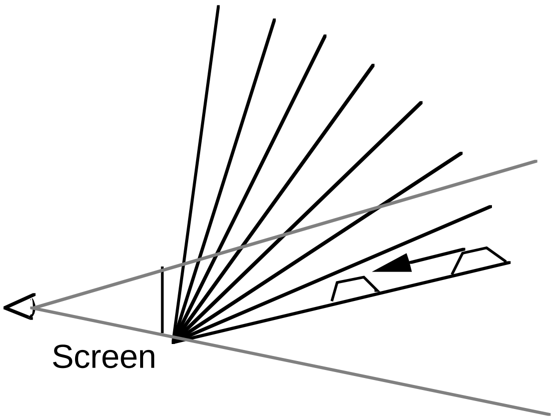

Figure 8. Schematic diagram illustrating the moving carpet of truncated pyramids as it was presented in stereo and at different angles to the vertical plane of the monitor screen (not to scale).

A. Method

The moving carpet display was used again for this study. To vary the scene depth the computer graphics model of the moving carpet was set at 8 different angles ranging between 10 deg. and 80 deg with respect to the monitor (as shown in Figure 8). Two buttons, one of which increased eye separation and the other that decreased it were provided as controls. Subjects made two settings at each angle and in addition there were two practice trials at the start of the session. The design was within subjects, fully randomized repeated measures.

Subjects: Twelve subjects who were all undergraduate or graduate students were used as observers. All could see depth in Julesz stereograms. Four of the subjects were paid to participate, the rest were unpaid volunteers.

B. Results

The results are plotted in Figure 9. The confidence intervals show that there was considerable variation between subjects. Nevertheless, all subjects increased their eye separation as the depth in the scene decreased. There is a very high negative correlation between the angle of the moving carpet and the average virtual eye separation setting ($r^2 = 0.99$). This relationship is described by the regression line

$$E_v = 18.5 - 0.149\theta$$

where θ is the angle given in degrees. The nearly linear relationship between the angle of the plane and the mean virtual eye separation setting was entirely unexpected and is very intriguing. To the best of our knowledge, this relationship has not been described before. However, in the present paper we are concerned with the practical application of these results.

The empirical relationship between subjects settings of virtual eye separation and the depth in the scene can be straightforwardly converted into a algorithm designed to change eye separations depending on the depth structure of a virtual scene. We have applied the result directly in stage 2 of the algorithm presented in the following section. It is also clear that there are large individual differences with respect to the amount of disparity that can be tolerated suggesting that users of stereo displays should be able to customize a disparity parameter for their own comfort.

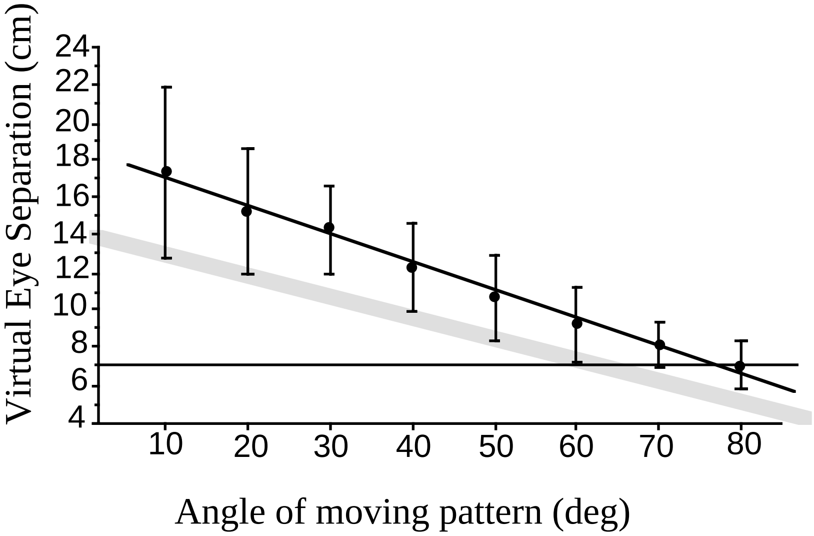

Figure 9. Mean "maximum comfortable" settings for subjects changing their eye separation as a function of scene angle. Zero degrees mean that the scene is vertical in the plane of the screen while 90 degrees denotes a horizontal plane. The vertical bars denote 95% confidence intervals using the t distribution. The gray line is the function used in the stereo adjustment algorithm. The horizontal line shows eye separation for the average human observer.

IV ALGORITHM FOR DYNAMIC DISPARITY ADJUSTMENT

In this section we present an algorithm that automatically changes the stereo viewing parameters depending on the depth structure of the 3D scene to be displayed. The purpose of this two stage algorithm is to reduce conflicts in vergence and focus while maintaining optimal disparities for stereo viewing. The algorithm as a whole relies on the flexibility of the disparity processing mechanism to ensure that shape distortions are not obtrusive when virtual eye separation is changed while viewing continuously moving scenes. The second stage of the algorithm is built directly on the results of Experiment 2.

A. Stage 1: Cyclopean Scale

The first stage algorithm is a scale about a point midway between the observers eyes. We call this a cyclopean scale after the Cyclops of mythology who had a single, central eye.

Stage 1 consist of two steps.

Step 1: determine the nearest point in the scene in viewing coordinates (with the origin at the viewpoint). In order to do this we use the depth buffer which is built in to many advanced computer graphics systems. Depth buffers are normally provided to solve the problem of hidden surface elimination but we use it in a new way. Since the depth buffer contains a depth image of the entire screen (albeit stored in a non-linear transformation) by samping the depth buffer and inverting the transform we can obtain an approximation of the nearest point in the scene currently being displayed.

Step 2: scale the scene about a point corresponding to the midpoint between the observer's two eyes. This scaling factor is calculated so that the nearest part of the scene comes to be located just behind the monitor screen as shown in Figure 10.

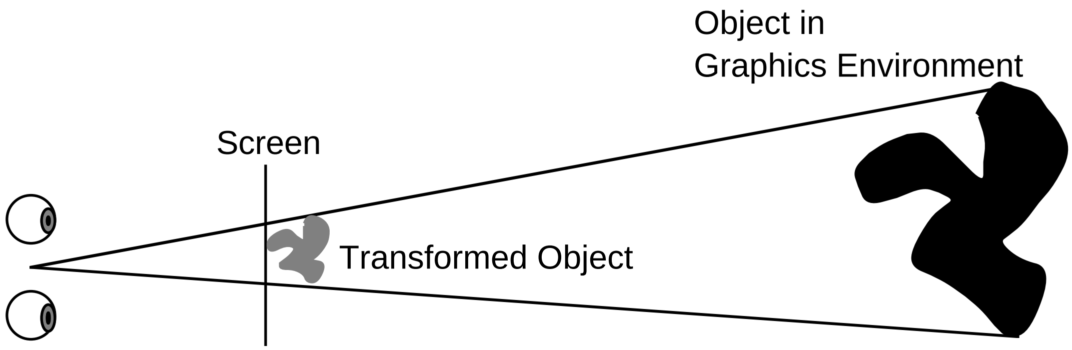

Figure 10. The cyclopean scale

This algorithm has two beneficial effects. The first is that it increases the eye separation *relative* to distant images. A distant large object, such as a mountain will have no useful disparity under normal viewing conditions. However, after a cyclopean scale the eye separation relative to object will be increased. The second benefit is that vergence-focus conflicts are reduced because objects are brought closer to the screen and hence into the area where the eye can be expected to focus.

An interesting point from a perceptual point of view is that if the user is flying forwards into a scene, the stereo disparity information is (paradoxically) always consistent with an object located just behind the monitor screen (since the scaling is done dynamically and continuously). Conversely, the motion parallax information is consistent with the actual movement of the observer. Nevertheless, despite the conflicting cues the observer always experiences forward movement with respect to the computer image. Thus the motion parallax dominates the perception of self motion, although there is no doubt that the stereo information adds to the perception of three dimensional structure.

B. Dynamic implementation

If we were to apply the transformation described above in a straightforward way we would have to do it in two steps. In the first step we would draw the scene solely to obtain the near point from the depth buffer. In the second step we would re-draw it with the cyclopean scale transformation applied. This is not suitable for real-time fly through animation and so we implement the following variation in order to create a dynamic implementation of the above algorithm. It has the following steps.

1 Draw image for the right eye without transformation
2 Sample the depth buffer (at 100 locations) to estimate the closest part of the scene.
3 For the left eye view
    3.1 scale the scene about the right eye viewpoint (not about the center point between the eyes).
    3.2 move the left eye viewpoint sideways by an amount corresponding to the eye separation. We use a value of 6.4 cm to correspond to the average observer.

Note that the Stage 1 algorithm is self contained, in many cases it may be all that is necessary or desirable. The Stage 2 algorithm which follows can be regarded as an enhancement.

C. Stage 2: Adjustment Of Eye Separation

Experiment 2 showed that users adjust their virtual eye separation depending on the depth (or depth gradient) present in the scene. Stage 2 of the algorithm applies this result to increase the disparities for scenes containing little depth and decrease them where there is considerable depth. In what follows, we always assume that Stage 1 has already been applied to bring the nearest point of the scene to the screen.

Algorithm 2 is a transformation based on the nearest and farthest points in the scene as sampled from the depth buffer. As in Algorithm 1, we first draw the view for one eye. Next we sample the depth buffer to estimate the nearest and farthest points. We move the near point of the scene to the screen by a scale as in Algorithm 1. All that remains is to change the virtual eye separation parameter as a function of the maximum depth in the scene. What follows is a description of how we obtained this function from the experimental results.

D. Eye Separation as a Function of Depth

We obtained our eye separation function by approximating the lower 95th percentile from Experiment 2 with a straight line. This is the wide gray line shown in Figure 8. We reasoned that this would provide a comfortable virtual eye separation function for most observers. This creates a virtual eye separation of 14 cm for a very flat scenes and one that is only 4 cm when the depth is extreme (i.e. a horizontal plane). The equation describing this line is :

$$E_v = -0.114\theta + 14.0 \qquad (4)$$

where E_v is the eye separation in cm and θ the angle in degrees. The relation the gradient of the moving carpet and the angle of the plane is given by

$$\theta = \arctan\left(\frac{dz}{dh}\right) \qquad (5)$$

dz/dh is the inverse of the change in height (*dh*) with change in depth (*dz*). Now by substituting for θ we obtain the function relating virtual eye separation to the depth gradient.

$$E_v = -.114\arctan\left(\frac{dz}{dh}\right) + 14.0 \qquad (6)$$

However, most virtual scenes do not consist of simple planes such as that used in the experiment. Our requirement was to create a general purpose algorithm that would adjust the virtual eye separation based on the near and far points of the scene. The cyclopean scale ensured that the near point was always at the screen. Thus we required an algorithm based on the far point, after cyclopean scale had been applied. Accordingly we substitute the height of the screen (S_h) for *dh* in equation 6, and the maximum sampled depth (after scaling) $Z_{\max}$ for *dz* to obtain our stage 2 algorithm.

$$E_v = -.114\arctan\left(\frac{Z_{\max}}{S_h}\right) + 14.0 \qquad (7)$$

E. Properties of the Algorithm

If we substitute the values 0 and ∞ for $Z_{\max}$ in equation 5 we find that the virtual eye separation can vary between 4.0 and 14.0 cm.

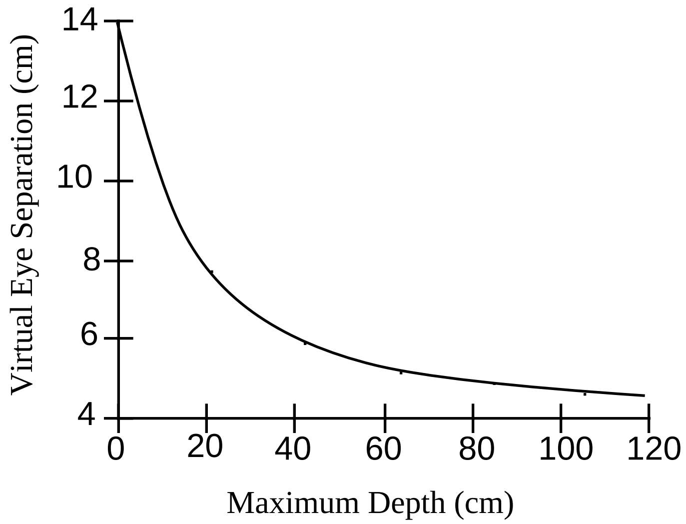

Figure 11. The function relating virtual eye separation to scene depth.

It is instructive to look at the transformed stereoscopic depth of points compared to the original distance when the algorithm is applied.

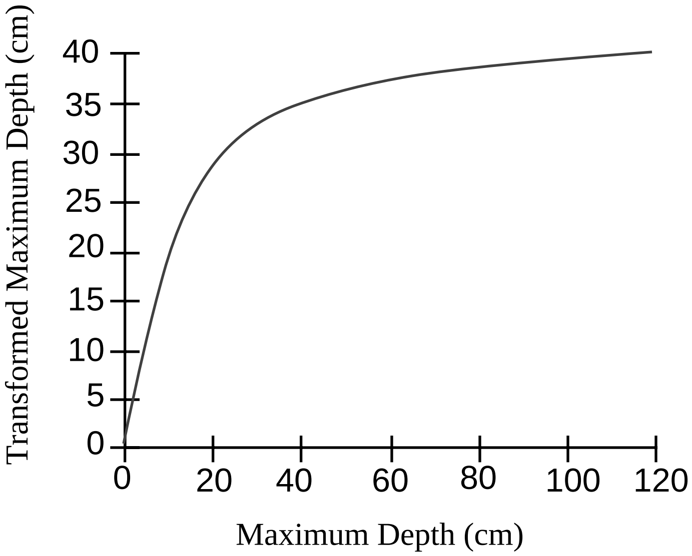

Figure 12. The function showing how the depth range is transformed by the algorithm.

If we substitute equation 7 into equation 2 we obtain a function relating the maximum depth after stage 2 transformation to the maximum depth before transformation (but after cyclopean scale). This function is plotted in Figure 12 This graph show two distinct trends. For flat scenes, where the depth is less than 30 centimeters, the transformed stereoscopic depth is about twice as great as the original depth, so the object will be stretched along the z axis. For scenes having a large depth range, the opposite occurs and stereo depth is decreased. If we consider the case where depth is infinite equation 2 simplifies to

$$Z_{\max} = \frac{SE_v}{E_a - E_v} \qquad (8)$$

For the parameters we use this means that in the most extreme case the stereo depth is equal to twice the distance of the observer to the screen.

E. Customizing for individual differences

In order to allow for individual differences we added the following parameterization to equation 6.

$$E_v = \left(\frac{4.0 - MaxSep}{90.0}\right)\arctan\left(\frac{Z_{\max}}{S_h}\right) + MaxSep \qquad (9)$$

Now by changing MaxSep we can change the virtual eye separation function in such a way that the minimum virtual eye separation is always 4.0cm and the maximum is whatever value is specified. Intervening values scale proportionately. As a default we use 14.0 cm for *MaxSep.*

F. Reducing the rate of change of eye separation.

Experiment 1 suggested that the rate of change of virtual eye separation should be kept low. We have investigated two methods. The first is to maintain a running average of the last five values for E_v and use this to specify the eye separation. This tends to smooth out rapid changes. The second method is to put a threshold on 0.2 cm/sec on the rated of change of virtual eye separation. Both of these techniques suffer from the same problem. If there is some abrupt change in the scene, for example an object pops up in the foreground, then a double image may occur because the algorithm only allows for a gradual change to the new eye separation values. The alternative is that eye separation values are not smoothed in which case a sudden change in eye separation will occur. However, this may be less noticeable than the double image effect. At present we have returned to the unfiltered version because despite the occasional jump in disparity we feel that this is the lesser of the two types of problem. Fortunately, most of the data we deal with is smooth and continuous and rapid jumps are rare.

V. EXPERIMENT 3: DISTORTION WITH CHANGES IN EYE SEPARATION

The purpose of this experiment was to assess the algorithm in terms of the amount of distortion that was perceived when the algorithm was applied; also to assess the extent to which double images are seen.

To recap, if the brain were a perfect geometry processor then dynamically changing virtual eye separations should cause the virtual environment to distort in a rubbery fashion. However, evidence from previous research suggests that the brain may be able to recalibrate the vergence and/or disparity processing mechanisms with the result that observers may not be aware of changing stereo parameters where the changes are gradual. This experiment was designed to find out the extent to which distortion and diplopia are perceived when eye separations are dynamically changed using our algorithm with a more natural scene.

A. Method

A scene was constructed that consisted of a small section of digital elevation map representing a small portion of the floor of the ocean. This section was an 80 by 80 grid of height values and these were color coded by height and shaded, with cast shadows.

A flyby was constructed so that the observer appeared to fly into the scene while looking directly down on it, then move to a horizontal viewing position and then fly back out again. This flight path was designed to result in large changes in viewing distance and large changes in the relative depths in the scene.

The six conditions were

1) No algorithm. In this case observers were presented with a stereo view of the scene with the eye separation set at 6.4 cm. This approximates the norm for humans.

2) Cyclopean scale only. This is the stage 1 algorithm described above.

In the remaining four conditions the two stage algorithm was applied using different values for the Maximum eye separation applied using equation 9.

3) 6.4 cm
4) 12.8 cm
5) 19.2 cm
6) 25.6 cm

Procedure:

Subjects were first trained by being shown a subset of three of the six conditions and the task was explained to them. They were told that they would be expected to report whenever double images (diplopia) occurred and to note if the overall shape of the surface appeared to change as the animation progressed.

Following the training, observers were shown the same animation under each of the six conditions (in a different random order for each subject) and asked to provide three pieces of information.
1) Do you see double images (yes/no)?
2) Rate the amount of distortion giving it a value between 0 and 4
3) Rate the amount of eye strain giving it a value between 0 and 4

Subjects: Seven volunteers who were either graduate students, undergrads or paid assistants .

B. Results

All subjects reported double images in condition 1 with no algorithm applied. This occurred when the scene was distant and the entire image became doubled. Of the other conditions only in condition 6 did two observers report any double images. This provides strong support for the practical utility of the algorithm in reducing double image formation.

The results from the rating of distortion are summarized in Figure 13 for each of the six conditions. As can be seen the maximum distortion was reported in conditions 5 and 6 which is not surprising given the extreme virtual eye separations that could occur. Condition 1 where there should have been no distortion was reported to have high distortion by some of the subjects. However, this may have been related to the difficulty of fusing the images. The lowest level of distortion was reported in condition 3 where the virtual eye separations did change dynamically but by relatively small amounts. We interpret these results as adding support to the theory that stereo cues do not provide absolute depth information, at least not in the presence of shape from motion information. These results also add support for the utility of the algorithm since with more moderate settings no more distortion is reported than with correctly applied stereo perspective.

Eye strain did not appear to be a problem since only one subject reported any eye strain in any of the conditions. However, the entire experiment was completed in less that 20 minutes and eye strain problems are more likely to occur with prolonged observation.

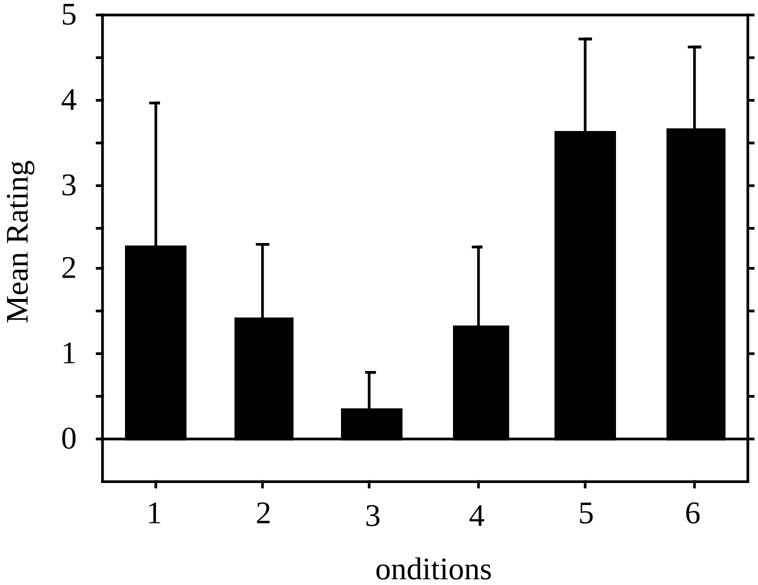

Figure 13. The mean distortion ratings obtained in the different conditions of Experiment 3. The vertical bars represent two standard errors of the mean.

VI. CONCLUSION

The algorithm described in this paper has been implemented in a stereo data visualization package called Fledermaus. This system uses a flying interface to allow users to rapidly gain different viewpoints within the scene [16]. Extensive use has proved the algorithm to be successful in providing a vivid sense of depth while minimizing the occurrence of diplopia under a wide range of conditions.

Our algorithm builds on the prior work of Williams and Parrish [15]. We believe that the most significant advantage of our algorithm lies with the treatment of scenes with small depth ranges. Their algorithm rescales the scene depth range to a fixed depth range which causes depth in shallow scenes to be greatly exaggerated. Our algorithm has the advantage that it exaggerates depths by no more than a factor of two for shallow scene and compresses it for deep scenes.

That these manipulations do not cause large perceptual distortions in the scene supports the hypothesis that people’s understanding of the global layout of objects in space does not come primarily from stereoscopic depth cues. Instead, kinetic depth and linear perspective cues are more important. Thus, we can undertake radical transformations (e.g. changing the effective eye separation by large amounts) without causing disortion. We speculate that there may be a penalty to such depth manipulations when depth judgments are critical but this remains to be tested. An interesting possibility that has yet to be explored is that virtual reality systems may benefit from the application of non-veridical eye separations. This runs counter to the usual approach which is to simulate the correct stereo pair as closely as possible.

In the future we would like to reexamine the way we sample the depth buffer to obtain the near and far point in the scene. The present algorithm assumes that the environments consist mainly of large extended surfaces. Small and narrow features such as lines may be occasionally missed by the sampling which can cause a transitory change in the near or far estimate. Another method to find the near and far points of the scene relying on bounding boxes might be advantageous. However, an algorithm that used bounding boxes would probably be considerably more complex than the simple z buffer method that we have implemented.

REFERENCES

[1] Deering, M. (1992) High Resolution Virtual Reality. Proceedings of SIGGRAPH '92. Computer Graphics, 26, 2 (July 1992), pp. 195-202.

[2] Patterson, R., and Martin, W.L. Human Stereopsis, Human Factors, 1992, 34(6) 669-692.

[3] Valyrus, N.A. Stereoscopy Focal Press, London, 1966.

[4] Cutting, J.E. (1986) Perception with an eye for motion. MIT Press, Cambridge, Mass.

[5] Wallach, H., & O'Connell, D.N. (1953) The kinetic depth effect. Journal of Experimental Psychology, 45, 205-217.

[6] Richards W, 1985 Structure from stereo and motion. Journal of the Optical Society of America A 2 343-349

[7] Ogle, K.N., 1952 "The visual space sense" Science **135** 763-771.

[8] Tittle, J.S. and Braunstein, M. L. 1993 Recovery of 3-D shape from binocular disparity and structure from motion, Perception and Psychophysics, 54 (2) 157-169.

[9] Todd J.T, Norman J F, Perotti V J, Tittle J S 1993 The discrimination of 3-D length from motion and stereopsis, Investigative Opthalmology and Visual Science, Suppliment 34 1131.

[10] Fisher, S.K, Ebenholtz M S, 1986 Does perceptual adaptation to telestereoscopically enhanced depth depend on the recalibration of binocular disparity? Perception and Psychophysics 40 101-109.

[11] Judge, S.J. and Miles, F.A. (1985) Changes in the coupling between accomodation and vergence eye movments induce in human subjects by altering the effective interocular separation. Perception 14, 617-629.

[12] Wallack, H, and Karsh, E. (1963) The modification of stereoscopic depth perception based on oculomotor cues. Perception and psychophysics. 11, 110-116.

[13] Noro, Kageyu, (1993) Industrial Aplication of Virtual Reality and Possible Health Problems. Japanese Journal of Ergonomica, v. 29, 126-129.

[14] Veron, H. , Southard, D.A. Leger, J.R. and Conway, J.L. 1990 Stereoscopic Displays of Terrain Database Visualization, Proceedings of SPIE Stereoscopic Displays and Applications, Santa Clara, 124-135.

[15] Williams, S.P, and Parrish, R.V. 1990 New computational control techniques and increased understanding for stereo 3D displays. Proceedings of SPIE Stereoscopic Displays and Applications, Santa Clara, 73-82.

[16] Ware, C. and Slipp, L. (1991) Using Velocity Control to navigate 3D graphical environments: a comparison of three interfaces. Proceedings of Human Factors Society Meeting, San Francisco, Sept. 35, 300-304.

[17] Ware, C. (1995) Dynamic Stereo Displays. ACM SIGCHI. Proceedings May, Denver. 310-316.