3D TASK PERFORMANCE USING
HEAD-COUPLED STEREO DISPLAYS

By

Kevin Wayne Arthur

B. Math., University of Waterloo, 1991

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

DEPARTMENT OF COMPUTER SCIENCE

We accept this thesis as conforming

to the required standard

....

....

THE UNIVERSITY OF BRITISH COLUMBIA

July 1993

Abstract

"Head-coupled stereo display" refers to the use of a standard graphics workstation to display stereo images of three-dimensional scenes using perspective projections defined dynamically by the positions of the observer's eyes. The user is presented with a virtual scene located within or in front of the workstation monitor and can move his or her head around to obtain different views.

We discuss the characteristics of head-coupled stereo display, the issues involved in implementing it correctly, and three experiments that were conducted to investigate the value of this type of display.

The first two experiments tested user performance under different viewing conditions. The two variables were (a) whether or not stereoscopic display was used and (b) whether or not head-coupled perspective was used. In the first experiment, subjects were asked to subjectively rank the quality of the viewing conditions through pairwise comparisons. The results showed a strong acceptance of head-coupled stereo and a preference for head-coupling alone over stereo alone. Subjects also showed a positive response to head-coupled stereo viewing in answers to questions administered after the experiment.

In the second experiment, subjects performed a task that required them to trace a path through a complex 3D tree structure. Error rates for this task showed an order of magnitude improvement with head-coupled stereo viewing compared to a static display, and the error rates achieved under head-coupling alone were significantly better than those obtained under stereo alone.

The final experiment examined the effects of temporal artifacts on 3D task performance under full head-coupled stereo viewing. In particular, the effects of reduced frame

rates and lag in receiving tracker data were investigated. The same path tracing task was performed under a set of simulated frame rate and lag conditions. The results show that response times of subjects increased dramatically with increasing total lag time, and suggest that frame rate likely has less impact on performance than does tracker lag.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

The goal of creating truly three-dimensional displays has long been pursued by scientists and engineers. With true 3D displays we could take advantage of our natural abilities to interact in three dimensions and avoid having to interpret 3D scenes using intermediate 2D displays.

In recent years, much progress has been achieved towards this goal through the use of computer graphics displays and real-time tracking technology. This thesis deals with one such type of 3D display technique, which we refer to as *head-coupled stereo display*. We define this technique as the use of a computer graphics workstation to display stereoscopic images of a 3D scene on a standard workstation monitor, with the images updated in real time according to an observer's eye positions. The scene appears stable and 3D in the sense that the observer can move around to obtain different views with binocular parallax and head-controlled motion parallax cues aiding depth perception.

Initial implementations of head-coupled stereo display systems using conventional workstation monitors were reported in the early 1980's, and several research implementations have been discussed since then. However, this type of display has yet to be accepted and put to practical use in general settings.

We describe previous work with this type of display and compare it with other 3D display techniques. The technical and human factors issues involved in implementing

and using head-coupling and stereo are outlined. In addition, we describe three experimental studies that were conducted to evaluate the effectiveness of head-coupled stereo, to compare the relative effectiveness of head-coupling and stereo as 3D depth cues, and to investigate the effects of temporal artifacts in the display.

The results from the first experiment show a high degree of user acceptance for the technique, as measured by subjective user preference tests comparing different viewing conditions. The second experiment shows, through objective measurements of subject performance on a 3D tree tracing task, that head-coupled stereo provides a significant improvement over a static workstation display, and that the depth cues from head-coupling are superior to those from stereo for tasks of this type. The third study provides an indication of how temporal artifacts in the tracking and display effect user performance. In particular the results show a serious degradation in response times as lag is increased, even at relatively low lags of approximately 200 milliseconds.

1.1   Historical Context

Various techniques have been developed to provide 3D display, ranging from those using optical and mechanical elements to those using computer graphics.

Optical techniques such as holography are best suited for creating static 3D images of objects. While some progress has been made recently in generating holograms using computers [14], the goal of updating high resolution holograms in real time is not expected to be attained for several years. The computational expense in doing so will most likely make the technique much less attractive, even though holography wouldn't require users to wear special glasses or tracking devices.

Techniques using complex electrical and mechanical components have also been developed for 3D display. The varifocal mirror technique displays volumetric images using

a vibrating mirror synchronized with the video output of a computer to display objects or pixels at varying depths [26][42]. The image is viewable from any angle in a reasonably large range. However the technique has some serious drawbacks: no occlusion effects are possible and all the objects appear semi-transparent. Rotating screen devices provide true 3D display in a cylindrical volume through the use of a rotating LED array or a rotating passive screen projected onto from below by a laser [9]. The technique is limited, however, with drawbacks such as the lack of occlusion effects.

One of the most popular methods for adding three-dimensionality to images is to display stereoscopic images, that is, to provide separate images to the left and right eyes [29][30]. Usually, special glasses are worn, and contain either coloured filters (anaglyphic) or polarized filters to direct the proper image to each eye. Stereo images can be presented without the use of glasses by employing lenticular arrays; the technique effectively displays different images depending on the angle at which the screen is viewed by using optical elements placed between the viewer and the screen. Stereoscopic imagery alone, however, suffers from some artifacts. In particular, when the image is viewed from the incorrect viewpoint (and especially if the observer is moving) the image appears to distort.

The advance of interactive 3D graphics has provided numerous techniques for simulating three-dimensionality and providing depth cues. It is possible to generate images of scenes using perspective, shading, shadows and motion, among other techniques, to indicate depth. Interactive 3D graphics is used widely today in various application domains, such as scientific visualization and computer aided design.

Traditional graphics displays employ a very simplified geometric model of the user. When displaying a 3D scene, the user's eyes are effectively modeled as a single point located at some arbitrary distance translated directly out from the center of the screen. Hence the display is really only correct if it is viewed from this one position. Of course,

people are accustomed to viewing 3D scenes from a physically incorrect angle through experience watching television or movies, but the effect is still one of viewing a 3D scene through a 2D medium. When the computer takes into account the positions of the observer's eyes it becomes possible to present a stable 3D scene which behaves correctly as the observer's eyes move around.

The idea of immersing the user in computer-generated 3D environments led to the concept of *virtual reality*, which was first introduced by Ivan Sutherland in the 1960's [55][56]. In a virtual reality system, the user typically wears a *head-mounted display* that contains two small screens and optics to stretch the images over a wide field of view. The user is separated visually from the real world and is immersed in a *virtual world*. The head-mounted display is connected to a head-tracker, and the host computer generates images for the two eyes depending on the position of the user's head and eyes.

Research in virtual reality progressed at various research labs [5][22][27], and by the mid 1980's the technology had advanced far enough that off-the-shelf systems started to appear on the market [3][32][46][57].

While most uses of head tracking technology have concentrated on the use of head tracking with head-mounted display systems, a few researchers have experimented with using head tracking with monitor-based graphics displays. Head tracking (and in effect eye position tracking) allows the computer to generate what we will call *head-coupled perspective*, meaning that the images displayed on the screen are computed with perspective projections defined by the positions of the observer's eyes.

We will use the term *head-coupled display* to refer to monitor-based systems using head-coupled perspective. Many such systems are also *stereoscopic*, or simply *stereo*, meaning that two images are presented to the observer, one for each eye. Combining the two techniques gives us what we will call *head-coupled stereo display*, meaning a display system employing both head-coupled perspective and stereoscopic images.

Another term sometimes used for head-coupled stereo display is *fish tank virtual reality* because the effect of the display is to present a small (fish tank-sized) virtual world to the user. For clarity, we will use the term *immersive virtual reality* when referring to systems that use head-mounted displays. Another, more descriptive, term for this would be *head-mounted stereo display.*

1.2   Overview

In the next chapter we outline previous work directly related to monitor-based head-coupled stereo display, studies of depth perception in computer graphics, and experiments to measure and evaluate temporal accuracy in virtual reality systems.

In Chapter 3, the requirements for implementing head-coupled perspective and for drawing stereoscopic images are discussed. A summary of the factors affecting performance under head-coupled stereo viewing, a list of issues related to accuracy and calibration, and the use of auxiliary depth cues are all discussed.

Chapter 4 describes the experimental procedures and results from the three experiments that were conducted. Chapter 5 discusses the relevance of the experiments to applications of interactive 3D graphics. In the final chapter we summarize the contributions of this work and discuss future extensions.

Chapter 2

Related Work

Monitor-based head-coupled stereo display is a technique that has been implemented in the past by various researchers. This chapter surveys their work as well as work related to studies of depth cues in head-coupled displays and issues of temporal accuracy and artifacts.

2.1 Head-Coupled Stereo Display

Head-coupled stereo display shares common elements with immersive virtual reality systems. In particular, both techniques employ head tracking and stereopsis with the goal of presenting a realistically stable computer-generated scene to the user.

Various terms have been used to describe monitor-based head-coupled stereo display, among them *fish tank virtual reality*, *viewpoint dependent imaging*, and *virtual integral holography*.

The earliest reported display of this type was the "Stereomatrix" system developed by Kubitz and Poppelbaum [36]. The user viewed a large (3 foot by 4 foot) projection screen illuminated from behind by lasers. Head tracking was performed by using photo detectors to track an infrared source worn on the user's head. A similar display, using a computer monitor, was reported by Diamond *et al.* [17]. Molecular data made up of line segments were viewed with head-coupled perspective (without stereo). A video camera was used to track a small light bulb worn on the user's forehead.

Similar early systems that employed head-coupling and stereo are described by various researchers [21][45][54][61]. These systems were typically limited to displaying wire-frame images because of the computational cost of displaying shaded objects. An alternative is to display from a set of precomputed perspective images, as suggested by Fisher [21]. Venolia and Williams proposed a system using precomputed perspective images accounting for only the side-to-side horizontal movements of the user, and not vertical movements or movements in depth, so as to minimize the number of images that need to be precomputed [61].

Codella *et al.* describe the multi-person "Rubber Rocks" simulator that uses a head-coupled stereo display interface [10]. Users wear a tracker attached to a baseball cap, as well as a glove, to interact with objects displayed stereoscopically on monitors or large projection screens. Multiple users, each viewing a different screen, can participate over a network to interact with the same scene.

Most early reports of work with head-coupled stereo displays focus entirely on implementation issues of performing the tracking and image generation correctly. Recent advances in tracking hardware have made it feasible and quite straightforward to create head-coupled stereo displays and hence more recent research has begun to deal with evaluating the effectiveness of the technique and the level of realism achieved by it.

McKenna reports on three experimental real-time graphics display systems using head-coupling [40]. His goal was to determine how effective head-coupling was with a monitor-based display, either fixed or movable. The first display system used a fixed high resolution monitor with the perspective projection coupled to the user's head position. The second display used the same monitor, but this time the monitor's position and orientation was also tracked so that it could be tilted or swiveled to obtain different views (head movements were tracked as well). The third display was a handheld LCD screen that could be freely moved. Both screen position and head position were tracked

and used in computing the images. Stereo was not used in any of the three displays. McKenna describes the results of an informal target selection experiment undertaken to evaluate the first of the three displays (the fixed-monitor, head-coupled display). Subjects controlled a 3D cursor using a handheld tracker. In each trial, a cube was displayed in the scene and the subject was asked to align the cursor with the cube. Three viewing conditions were employed in the experiment: fixed view, view controlled by mouse movements, and view coupled directly to head position. The results showed that under the head-tracked condition, subjects could match the target more rapidly, and that the mouse control was of virtually no benefit over the fixed view display for this task. No studies with the other two movable displays were reported.

Deering presents the most complete analysis to date of the issues that must be addressed to correctly implement head-coupled stereo display [15][16]. He discusses the importance of several factors, including fast and accurate head tracking, a correct model for the optics of the human eye, the use of physically correct stereo viewing matrices, and corrections for refraction and curvature distortions of CRT displays. Deering describes an implementation that achieves sub-centimeter registration between the virtual scene and the real environment surrounding it.

The refraction and curvature distortions inherent in most CRT displays make the image plane appear curved in 3-space, and the distortion changes depending on the location of the observer. Implementing the full correction for the distortion is not practical due to the high computational expense. The method suggested by Deering provides a first order approximation to the full correction by adjusting only the four points at the corners of the viewport (as opposed to adjusting the entire image). Another possibility is to correct for a particular point in the image where the observer is looking. This point might correspond to, for example, the location of a 3D mouse being used to interact with the scene.

Deering also describes the issues involved in implementing an immersive head-coupled display that doesn't use a head-mounted display [16]. The *virtual portal* system displays head-coupled stereo images on three large projection screens covering most of the observer's field of view. The level of realism achieved with the system exceeds current head-mounted displays in resolution and registration and is less physically intrusive.

2.2   Comparison of Depth Cues

Aside from evaluating the effectiveness of full head-coupled stereo viewing, our research was aimed at assessing the relative merits of stereopsis and head-coupled perspective as depth cues. Several studies comparing depth cues have been reported in the literature, although none directly comparing head-coupling and stereo are known at this time.

Sollenberger and Milgram compared the relative effectiveness of stereo and rotational depth cues [53]. They conducted two experiments, using a 3D tree tracing task (the same task that we employed for our experiments). The first experiment was a 2×2 study with the variables being the presence or absence of stereo and the presence or absence of rotational motion of the scene. The rotation, about a vertical axis in the center of the screen, was controlled by the user holding down a mouse button, with the direction of rotation defined by the mouse's position – forward rotation if on the right half of the screen, backwards if on the left side. The results showed a greater benefit from motion alone than from stereo alone and the best results were obtained with both depth cues combined. (Subject performances were: 56.9% correct for neither cue, 74.6% for motion, 66.6% for stereo, and 89.1% for both.) Their second experiment compared continuous motion (rotation) with viewing from multiple static views (rotations of the scene about a vertical axis), with the view changing automatically every 5 seconds. Display with stereo and multiple viewing angles was found to be no more effective than rotational motion

alone, but less effective than rotational stereo display.

The area of telerobotics and remote manipulation have benefitted from the techniques of stereo and head-coupling. Pepper *et al.* [47] describe studies of the use of stereo in telepresence applications, showing a clear advantage of stereo display over fixed viewpoint non-stereo display. In addition, they describe studies to compare the effectiveness of stereo under different conditions of motion: a study was conducted to investigate whether any depth perception is provided by the induced stereo movement seen when the viewer moves his or her head when viewing a stereo image. The results showed no change in the subject's perception of 3D and no change in the ability to perform 3D tasks. A preliminary report is given on true head-coupled display using a head-mounted display coupled isomorphically with the cameras recording the scene, and hence providing true motion parallax cues. The results under this viewing condition show a significant improvement over the stereo non-moving condition in measurements of stereoacuity.

Cole *et al.* conducted experiments to evaluate the benefits of motion parallax (through head-coupling) in a teleoperation task, performed with and without stereo [11]. Subjects viewed a real scene recorded through video cameras that moved according to their head movements. The video images were displayed on a monitor. The results showed a significant increase in performance when motion parallax was added to monocular views, but not when motion parallax was added to stereoscopic views. The reason cited for this lack of improvement with head-coupling and stereo is that for their experiments, the subjects' performance probably peaked with stereo alone and no further improvement was possible.

In addition to the two depth cues of binocular and motion parallax, there are several other well-known cues which aid in the perception of depth, such as occlusion, shading, and shadows [28]. Wanger *et al.* report on the relative effectiveness of several cues, including shading, textures, perspective projections, and different types of shadows on

3D perception and object placement tasks [63][64].

2.3   Temporal Accuracy

Of the various deficiencies present in current virtual reality interfaces, the issue that is perhaps most often raised is the problem of lag in acquiring tracking data. Lag is usually cited as being more severe than other problems such as low spatial resolution and low frame rates (temporal resolution). Although most related work on temporal accuracy of trackers has been done in the context of immersive virtual reality, the issue is the same for trackers used in monitor-based head-coupled displays, so results concerning lag probably apply to both immersive and non-immersive displays.

With respect to tracking, lag or latency is by definition the delay between movement of the tracker and the resulting change in the display. Lag can be classified as arising from three primary sources [25]. The first is the lag in receiving and processing tracker records, including performing smoothing algorithms on the data. The second component is the lag in the display loop (the time taken to compute and display a frame). The final source of lag arises from minor delays introduced by variations in the system load (caused, for example, by network or operating system activity).

The problem of lag has been studied by various researchers, primarily for the purpose of measuring and counteracting the lag in a system.

Liang, Shaw and Green measured the lag in the Polhemus IsoTrack magnetic tracker to be approximately 110 ms when receiving records at 20 Hz [37]. Their method used a video camera to record a swinging pendulum with a tracker attached, placed in front of a display screen showing a time stamp. The computer kept a log of tracker records with times so that the lag between these values and the time stamps seen in the video could be determined by comparing the delay for a specific reference position. They discuss the

use of Kalman filtering, an established technique in signal processing for smoothing and prediction [34]. Kalman filtering for virtual reality systems has also been discussed by other researchers [25].

Similar experiments to measure lag are reported by Bryson and Fisher [6]. They define two types of lag, *transmission lag* and *position lag*, and describe experiments to model the lag characteristics of several trackers. Transmission lag time is defined as the time between the first movement of a tracker at rest and the first movement of the cursor (or 3D object) being controlled by the tracker. The other type of lag, position lag, is the difference between the distance the tracker has moved and the distance the cursor controlled by the tracker has moved (measured in the same coordinate system). Hence position lag depends on the velocity at which the tracker is moved. Experiments were conducted using video recording to measure and establish a relationship between these two lags and the graphic update time. A model is proposed for the dependence of position lag on transmission lag time, velocity, and graphic update time.

A more precise general testbed for measuring lag in trackers is described by Adelstein *et al.* [1]. They used a large motorized rotary swing arm to move a tracker through a known motion at a controlled frequency, and measured lags of Ascension, Logitech, and Polhemus trackers.

Although careful measurements of lag have been reported and methods have been developed to help reduce the amount of lag in systems, little is known about the extent to which lag is a problem; few systematic studies have been undertaken to characterize the effects of lag on user perception and performance.

Recently, MacKenzie and Ware reported on an experiment to study the effect of lag on a 2D Fitts' law target selection task [39]. Subjects performed the usual task which involves moving a cursor as quickly as possible to a target area on the screen. To simulate lag coming from the mouse tracker, mouse records were buffered for different numbers of

frame times to generate different experimental conditions. The resulting response times were analyzed with respect to this lag. The results showed that a model in which lag has a multiplicative effect on Fitts' index of difficulty accounted for 94% of the variance in their data. This is better than alternative models that propose only an additive effect for lag.

Aside from lag, the other primary temporal artifact to consider is that of reduced frame rate. Various researchers have suggested that frame rate is relatively unimportant in comparison to lag [1][59], although no direct studies are known that confirm or refute this hypothesis.

Chapter 3

Head-Coupled Stereo Display

The technique of head-coupled stereo display shares common characteristics with traditional graphics workstation displays and with immersive virtual reality systems. Head-coupling and stereo provide extra depth cues not available with traditional displays, and at the same time do not suffer from the problems of current technology head-mounted displays. There are a number of reasons why, at least in the near term and for many applications, monitor-based head-coupled stereo display is more practical than immersive virtual reality using head-mounted displays.

One obvious reason for considering head-coupled display is that head-mounted display technology is currently very primitive and limiting compared with workstation monitors, as well as being expensive. Current head-mounted displays typically have low resolution, on the order of 400×600 pixels, and the optics that are used to stretch the image over a wide field of view create distortions in the image that are difficult to correct for efficiently [32][49]. While the technology has been improving rapidly, one can expect that it will be several years before the visual acuity afforded by monitor-based head-coupled display is matched by that of head-mounted displays.

A more fundamental argument, and one that will not change as technology improves, is that the property of immersion is probably not necessary for many applications. For instance, in the area of medical visualization, scientists viewing 3D scans of patients would probably not want to be immersed in the data. There are cases where immersion

is desirable, such as architectural walkthroughs for example, but it can be argued that these form a small subset of applications which use 3D computer graphics. Immersion has the disadvantage that it disconnects the user from the real world; he no longer has easy access to his keyboard or other standard input devices and cannot easily interact with his work environment and colleagues. See-through head-mounted displays and *computer augmented reality* interfaces alleviate this problem somewhat [2][20].

The following sections outline the components of head-coupled stereo display and the issues involved in implementing and using it correctly.

3.1 Stereo Display

The basic concept underlying stereo display is to provide the user with a sense of depth through *binocular parallax*. In the real world, the disparity between the two views seen by our two eyes allows us to judge distances to objects. In computer generated scenes, the same effect can be created by presenting different images to the two eyes, computed with an appropriate disparity. To correctly view a 3D scene stereoscopically, each image should be created using a perspective projection that corresponds exactly to the position of the user's eye when he is viewing the scene. Of course, without any head tracking capability, an assumption must be made about where the user is located. Hodges describes the software requirements for displaying computer generated stereo images using correct off-axis projections [30]. For simplicity, many stereo applications (that are not head-coupled) generate the left and right images by simply rotating or translating the scene (and displaying with two on-axis projections) [29][43][58]. While such approaches do provide a reasonable set of disparities for stereo images and can provide stereopsis effects, the images are not physically correct in general.

When viewed from a position other than the intended one, stereo images will appear
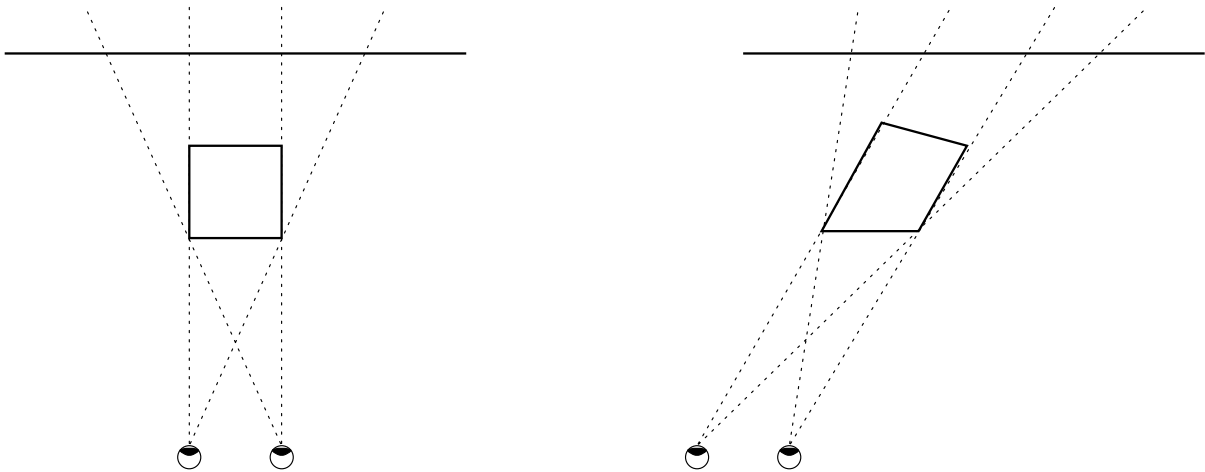
Figure 3.1: The effect of induced stereo movement on an image computed with fixed perspective projections. The solid line represents the image plane and the dotted lines represent the projection from the eyes to the image. The object being drawn, a box, appears to distort as the user moves to an off-axis viewpoint.

distorted. Stereo imagery suffers from an interesting artifact when the user's head moves, often referred to as *induced stereo movement* [60]. The scene will appear to bend about the image plane, following the user's eyes (see Figure 3.1). This effect can be distracting and degrades the illusion that the virtual scene is stable and three-dimensional. This artifact is also present when viewing non-stereo images but is not as distracting.

Induced stereo movement is only present when viewing a static (fixed viewpoint) stereo display and the effect disappears when head position is taken into account and the perspective projection is correctly coupled to head position, as will be discussed in the next section.

Another problem with stereo displays is the conflict between accommodation and convergence. When our eyes fixate on an object in the real world, they converge inward and the focal length of the lenses adjusts so that the fixated object is in focus, while objects at other depths are not in focus. This effect is known as *depth-of-field*. When viewing stereoscopic displays the eyes will converge according to a fixation point in the

scene, but they must always focus on the image plane, which is at a fixed depth. Hence all objects in the scene will appear in focus, regardless of depth. This effect can be distracting and degrades the level of realism exhibited by the virtual scene. It may also be physiologically harmful to the eyes, although little is known about this. This problem is less severe the farther away the display screen is situated from the user, and the closer the virtual objects are to the screen. In a head-mounted display, the optics make the image plane appear approximately 40 cm from the user's eyes [49]; when viewing a workstation monitor, a users's eyes are typically 80 cm from the screen surface.

3.2   Head-Coupled Perspective

A necessary part of the geometry pipeline for rendering 2D images of 3D scenes is the projection that maps the graphic primitives in 3-space onto the view plane in 2-space. Most commonly in computer graphics, orthographic or single-point perspective projections are employed. For a comprehensive discussion of the different types of projections the reader is referred to the survey paper by Carlbom and Paciorek [8] or to standard computer graphics textbooks [23][24][43][44][50].

A parallel projection (usually orthographic, or sometimes oblique) maps points in 3-space directly onto the view plane along a perpendicular direction. Alternatively, a perspective projection is often used to scale the scene's horizontal and vertical coordinates with depth, thus providing a sense of depth. The projection is called *on-axis* because the viewpoint is chosen to be a point along the z-axis (where the z-axis is perpendicular to the screen plane). The projection is given by a viewing pyramid defined by a viewpoint and the four corners of the screen (see Figure 4.2). The image would be physically correct if the image is viewed monocularly from the particular viewpoint used to create it. Viewing an image from the incorrect viewpoint causes the virtual scene to distort in various ways

[51]. This is the same problem encountered with induced stereo movement, although it is not as severe without stereo, and in fact the human visual system has evolved to compensate for incorrect viewpoints as demonstrated by our willingness to view cinema or television from incorrect viewpoints [13].

In a head-coupled display system, the perspective projection is dynamically coupled to the current positions of the user's eyes and the projection is necessarily an off-axis projection in the general case (see Figure 4.2).

The effect of head-coupled perspective is illustrated by the four screen photographs in Figure 3.2. In all cases the program is displaying the same 3D model of an automobile positioned at the center of the screen, level with respect to the monitor. Two different perspective projections and two corresponding camera angles are employed (resulting in the four photographs). Only in the two photographs where the camera position matches the perspective projection does the object appear three-dimensional and undistorted. In the other two photographs, where the camera position does not match the perspective projection, the object appears distorted.

Most graphics libraries contain functions for generating perspective projections, although not all support off-axis perspective. Our implementation uses the Silicon Graphics GL library (see Appendix B). Alternatively, one could directly compute the viewing matrix required to transform points according to the projection defined by an arbitrary viewpoint and viewing plane [15].

To provide correct head-coupled perspective, the system must know where the user's eyes are located. The eye positions are typically found by tracking head position and orientation and estimating the positions of the eyes with respect to a reference point on the tracker. We assume that eye position is sufficient and that rotation of the eye has no effect. This assumption is incorrect, however. The effective viewpoint of the eye, the first nodal point in the eye's optical system, is located approximately 0.6 cm in
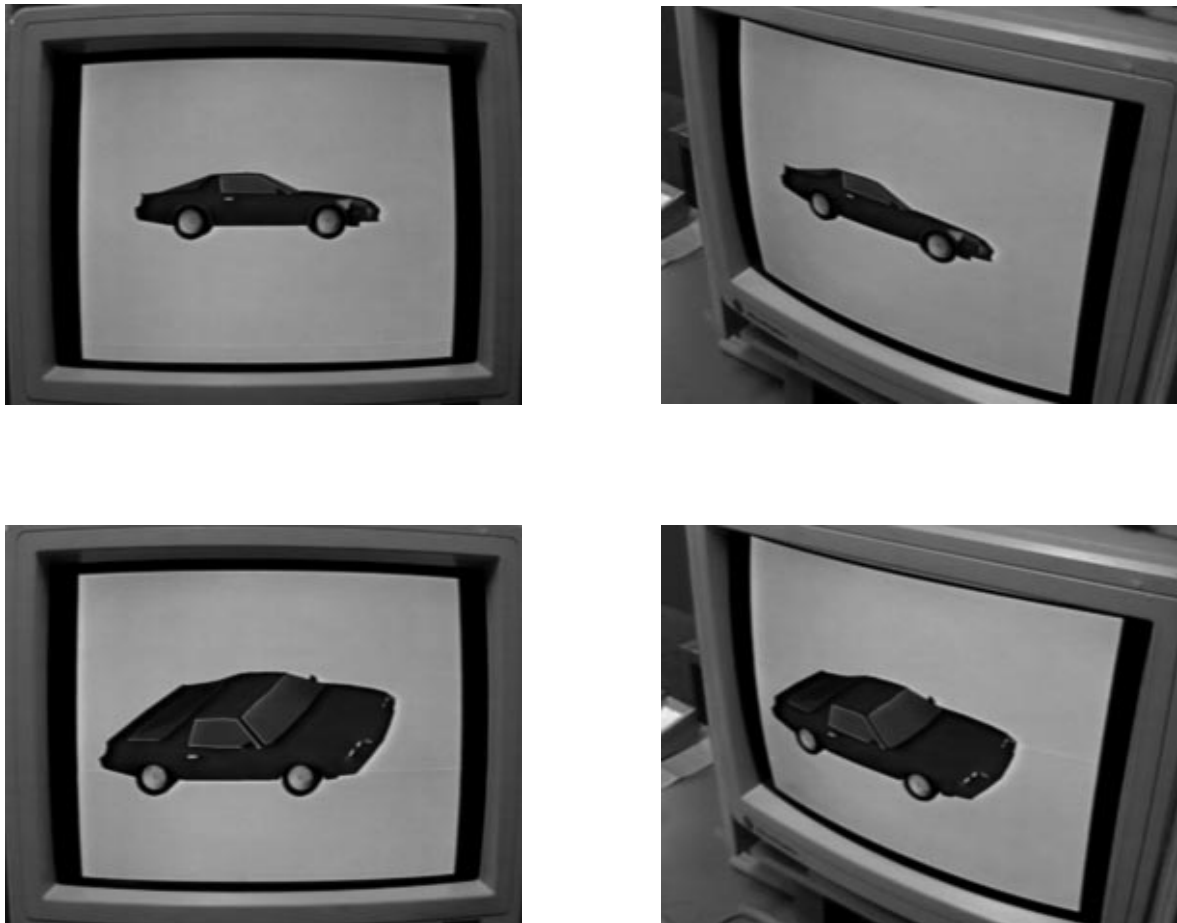
Figure 3.2: An illustration of the effects of head-coupled perspective. The program is displaying images of a car positioned in 3-space at the center of the screen. In the top row the image is computed with an on-axis perspective projection and in the bottom row with an off-axis projection. The left column shows the screen when viewed from a position on-axis and the right column shows the screen when viewed from a position off-axis. Only in the top-left and bottom-right photographs does the perspective projection match the viewing position, resulting in a realistic image that does not appear distorted.

front of the eye's center of rotation, for an adult with normal vision, and so it moves as the eye rotates [15]. Thus to provide the correct perspective, one would have to track the user's eye movements and adjust the computed viewpoint accordingly. While eye-tracking equipment is available and has been used in human-computer interaction [33], it is generally too costly and awkward to build into a head-tracking system. In actuality, the inaccuracy caused by having a slightly incorrect eyepoint for the perspective computations is unlikely to be larger than the errors arising from inaccuracies in the head-tracking and from distortions produced by the CRT screen. The effect may be worse for head-mounted displays since the screens are much closer to the eyes than a conventional monitor would be.

Note that head-mounted displays (as well as head-coupled displays) require off-axis perspective projections even though the relationship between the eyes and the screens is fixed. The screens in head-mounted displays are usually not perpendicular to the line of sight, but are usually angled away from the face [49]. There are also other distortion corrections that must be implemented in a head-mounted display and these can be quite expensive and difficult to perform in real time.

3.3   Six Degree-of-freedom Tracking

In order to employ head-coupled perspective, the system must track the user's eye positions, or at least the user's head position. Various technologies have been developed to perform 6 degree-of-freedom (position and orientation in 3D) tracking [41]. There are four main categories: magnetic, acoustic, optical, and mechanical.

Magnetic trackers, such as those sold by Polhemus and Ascension, are typically small and lightweight. However, they suffer from interference problems when operated in the vicinity of electrical objects.

Acoustic trackers transmit ultrasonic frequencies between one or more transmitters, usually worn on the user, and a common receiver. An example of an acoustic tracker is the Logitech 3D mouse. One drawback of acoustic trackers is that a clear line of sight must be maintained between the transmitter and receiver.

A variety of optical tracking methods have been experimented with, although optical trackers are not yet commercially available. Optical tracking holds much promise due to its unintrusive nature, although it does have the same line-of-sight restriction as acoustic tracking. Optical tracking is common for applications requiring less than 6 degrees of freedom, such as Krueger's *Videoplace*, which uses simple bitmap operations to trace the silhouette of a participant's head [35].

Mechanical trackers, such as the ADL-1 from Shooting Star Technology and the tracker used in the BOOM (Binocular Omni-Oriented Monitor) display from Fake Space Labs [38], use mechanical linkages and electronics for measuring angles (goniometers) to obtain very fast and accurate measurements. While mechanical trackers typically suffer negligible lag compared to other types of trackers, a disadvantage is that they can be heavier and less comfortable to use.

Implementing head-coupled display is independent of the tracking technology employed. With our system we have been using the ADL-1 mechanical tracker. This tracker uses potentiometers in each of 6 joints to measure angles, and returns the position and orientation of the end joint. The rated absolute positional accuracy of the ADL-1 is 0.51 cm and its repeatability is better than 0.25 cm. This device has a lag of less than 3 ms, which is shorter than the lag introduced by other factors such as the time taken for reading the input buffer through the RS-232 port. We have made no precise measurements of the tracker accuracy ourselves; the cited values are from the manufacturer. During our studies we used the raw data provided by the tracker and performed no prediction or smoothing. A simple backwards averaging method for smoothing was implemented for

test purposes, although it was not enabled during the experiments so as not to introduce additional lag into the system.

3.4 Factors Affecting Performance

In addition to the primary functional components necessary for implementing head-coupled stereo display, there are a number of issues that have effects on the quality of the system and on user performance. For general references on the topic of human factors for virtual reality systems, the reader is referred to the survey article by Ellis [18] and a collection of papers he edited [19].

3.4.1 Calibration

To maintain a high degree of realism with the display, the system should be calibrated carefully to take into account the physical locations and characteristics of the monitor, the head tracker and the user. With regard to the monitor, the primary parameters are the positions (in the real world) of the corners of the viewport. The glass in most monitors is not flat, but is shaped like a section of a sphere or cylinder. The image plane is distorted by this curvature, and also by refraction effects from the glass. Deering derives equations describing the distortions for any point in the image [15]. Given this function, he adjusts the four corners of the viewport to appear physically correct for the current eye position. This produces in effect a linear approximation to the distortion correction over the image. Deering's approximation has been implemented in our system, although it was not employed during the experiments to reduce computational costs.

Another important factor to correct for is any distortion in the tracker's measurements. Most trackers exhibit some distortion and this can be corrected for using a number of methods, such as approximating the inverse distortion with polynomials or

using lookup tables [7]. Although we have not implemented any correction schemes for our ADL-1 tracker, some distortion has been observed as the tracker nears the outer range of its operating volume, and in particular as the user moves very close to the screen.

There are also user-dependent parameters to adjust for, such as the spacing between the eyes and the location of the eyes with respect to the tracker (when it is worn on the user's head). In our current system, no provision for interactive calibration for new users is made; "average" values of eye spacing and position have been chosen and hence there will be some inaccuracies for different users.

3.4.2   Temporal Accuracy

Inaccuracies resulting from timing delays due to communication and processing time either in the graphics pipeline or prior to that in the computation of the viewing parameters for the virtual world must be dealt with in any system. The two primary temporal variables are *lag* and *frame rate*.

Lag is inherent in all instances of human-machine interaction, including virtual reality and telerobotics. Lag in head-coupled stereo display systems arises from delays in the tracker transmitting data records, hardware or software delays while processing the tracker data to perform smoothing or prediction, and additional processing time spent in the main display loop prior to displaying the scene. Although lag is recognized as an important factor in all VR interfaces and work has been done on techniques to compensate for it, there has been little experimental study of the perceptual and performance effects of lag in virtual reality systems reported in the literature.

The *frame rate* in an interactive graphics system is defined as the number of frames displayed per second. A low frame rate, and the resulting delay between frames, will not only contribute to the total lag in the system, but very low frame rates of around 10 frames per second or less will make the scene appear very jittery and the effect of

head-coupling will be less natural. Low frame rates are a standard problem in computer graphics when displaying complex scenes. For virtual reality systems, the problem may be more significant as the image changes in response to the user's natural movements and thus artifacts in the continuity may be more disturbing than temporal artifacts in traditional interactive graphics applications.

3.4.3   Auxiliary Cues

As with conventional graphics display, there are a number of techniques available for providing convincing depth cues. Particular techniques include shading, both Lambertian and specular, and the use of shadows to suggest the shape and relative positions of objects. A perspective projection provides depth information by scaling the extent of objects in the x and y directions according to z (depth). These and other well known cues are described in most standard graphics texts [23][24][43][44].

Chapter 4

Experiments

Three experiments were carried out to investigate different aspects of head-coupled stereo display. The primary purposes of these experiments were:

1. evaluate the effectiveness of head-coupled stereo display in general;

2. compare the relative performance of head-coupling and stereopsis as depth cues;

3. investigate the effects of temporal artifacts when using head-coupled stereo display.

This chapter describes the experiment system and general experimental procedure, followed by the details of the three studies. The next chapter will discuss the results and implications of them.

4.1 Experiment System Configuration

Figure 4.1 shows a photograph of the system used to conduct the experiments. The program is running on a Silicon Graphics Iris 4D 240/VGX workstation. The subject is wearing StereoGraphics Crystal Eyes glasses and a Shooting Star Technology ADL-1 head tracker. The glasses are synchronized with an interlaced video monitor which is refreshing at 120 Hz, and the LCD shutters in the glasses alternate to provide an effective 60 Hz update to each eye. The head tracker is mounted above the screen by a wooden frame attached to the sides of the monitor. We would have preferred to attach the tracker directly to the top of the monitor but this was not possible due to the range of operation

Figure 4.1: The head-coupled stereo display system. The subject's head position is measured by the ADL-1 mechanical tracker. StereoGraphics glasses are worn to provide different images to the left and right eyes, and the display monitor is synchronized with the glasses to provide an effective 60 Hz to each eye.

of our tracker; it was necessary to mount the tracker approximately 40 centimeters above the monitor. The monitor was raised so that the center of the screen was level with the subjects' eye positions and the mouse and pad were positioned comfortably for the subject. The distance from the screen to the subjects' eyes was approximately 50 cm.

4.2 General Experimental Procedure

Five basic viewing conditions were employed in the experiments. These are listed in Table 4.1 with the labels that are used to refer to them subsequently. They are shown schematically in Figure 4.2. For "non head-coupled" conditions, the perspective image was computed once according to the subject's initial head position, whereas for head-coupled conditions the image changed dynamically as the user moved his or her head. In stereo conditions (conditions STE and HCS), different images were displayed according to the estimated left and right eye positions of the viewer. In the "non stereo" conditions

(conditions PIC, HCM, and HCB), the same image was presented to both eyes. In conditions PIC and HCB the image was computed for the "cyclopean" eye position, that being the position midway between the two eyes. For the other non stereo condition, the "head-coupled monocular" condition (HCM), the image was computed correctly for the right eye, and the subjects were asked to close or cover the left eye (with their hand or a piece of paper over the stereo glasses).

In Experiments 1 and 2, the viewing condition varied randomly among the five conditions. In Experiment 3, the full head-coupled stereo condition was always employed and temporal artifacts were introduced by simulating tracker lag and reduced frame rates.

PIC	Picture
STE	Stereo only
HCM	Head coupled monocular
HCB	Head coupled binocular
HCS	Head coupled with stereo

Table 4.1: The five possible viewing conditions used in the experiments.

For each of the experiments we ensured that each subject could perceive depth using stereopsis, and that each subject moved his or her head around throughout the experiment so that the effect of head-coupled perspective could be experienced. It is estimated that a small proportion of people cannot achieve the benefits of stereopsis, due to irregularities in the eyes. To confirm stereopsis, each subject was shown a stereo test scene prior to performing the experiment. Figure 4.3 shows the test scene. The background of the scene was blue with two black squares, one above the other. Inside of the black squares were red squares, offset from the center slightly, either to the left or to the right depending on the eye. When the images are viewed stereoscopically, the red square on top appears to be be located in front of the screen and the red square on the bottom appears to be

Non-Stereo   **Stereo**

PIC   STE

Non-Head-Coupled

HCM

HCS

Head-Coupled

HCB

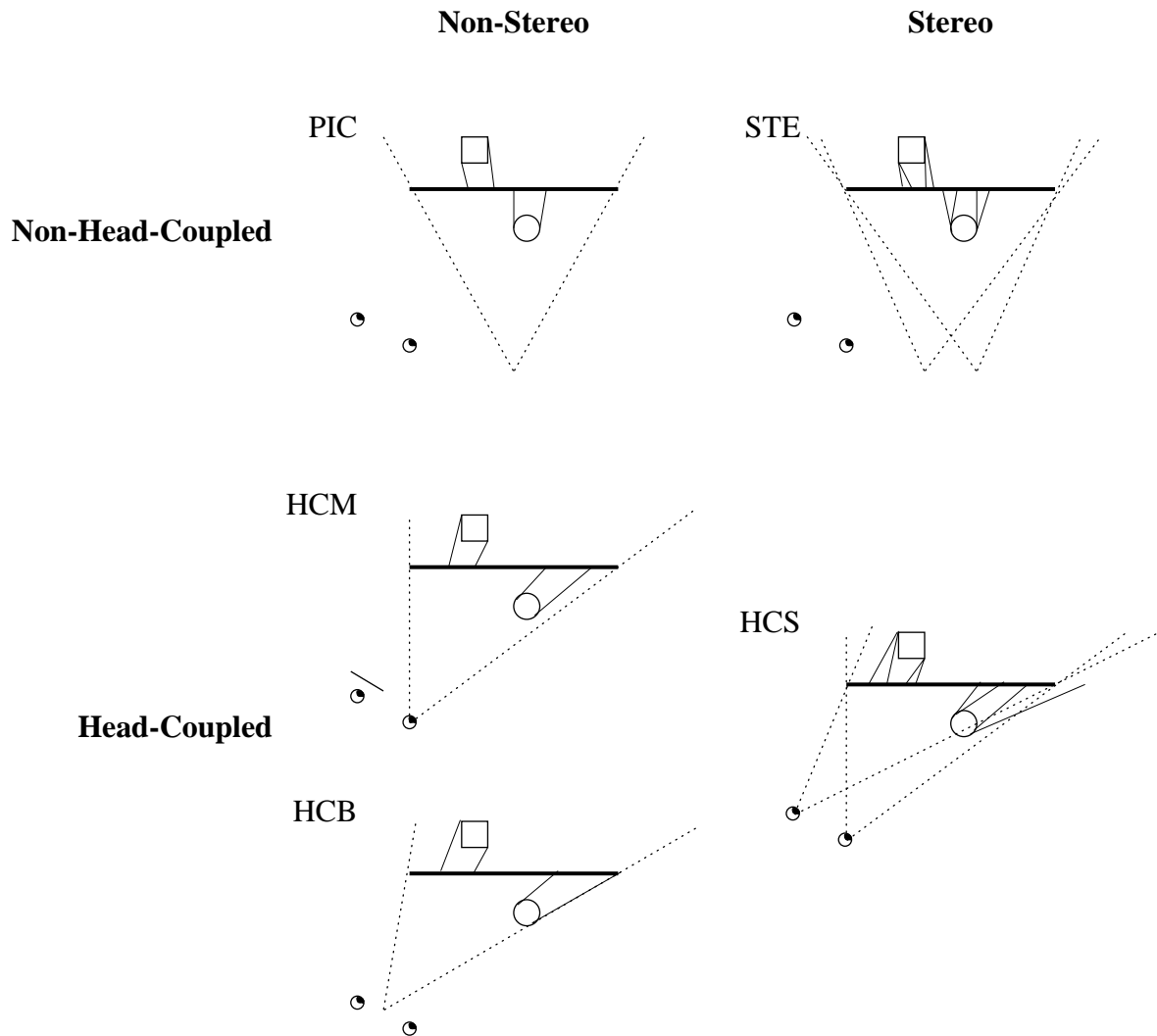Figure 4.2: The five viewing conditions used in Experiments 1 and 2 (see Table 4.1). In each of the diagrams the image plane is represented by a bold horizontal line, and virtual objects are shown in front of and behind the screen with the projection onto the image plane indicated by solid lines. The dotted lines indicate the perspective projections employed, each defined by an eyepoint and the corners of the screen.

Figure 4.3: The left eye and right eye images for the stereo test scene.

located behind the screen. The only difference between the left and right eye images was this horizontal displacement, and so stereopsis was the only cue that could lead to depth perception in the image (other cues such as perspective were not used). The static image was shown with no head-coupling. Each subject was asked to describe what he or she saw, and in all cases the subjects responded correctly that the top rectangle appeared to be coming out of the screen and the bottom one appeared to be going into the screen.

4.2.1 Experiment Scenes

In Experiment 1, two different scenes were shown to subjects to obtain their subjective evaluations of the value of stereo and head-coupling (see Figure 4.4). In both scenes we wanted to provide as much depth cueing information as possible and yet still maintain a 60 Hz update rate. The first scene contained an approximated sphere casting a pre-computed fuzzy shadow drawn on a striped ground plane. The scene was smooth shaded with specular highlights. The second scene consisted of a bent tube object, similar in shape to the Shepard-Metzler mental rotation objects [4][52]. Again, the scene was rendered with smooth shading and specular highlights, however a shadow and ground plane were not included for the tube scene as it was not possible to render it reliably at 60 Hz. Colours were chosen to minimize ghosting effects due to slow phosphor decay times of
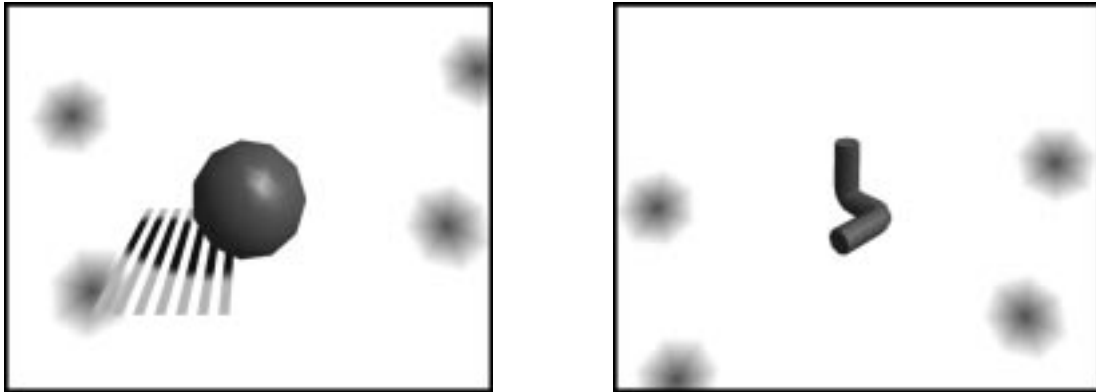
Figure 4.4: The sphere and bent tube displays used in Experiment 1. Hardware lighting was used to achieve the specular reflection. The blurry cast shadow was pre-computed and subsequently texture-mapped onto the floor. The colours and background have been modified for black and white reproduction.

the monitor. In particular, we chose colours with a relatively small green component.

The background is a *vection background* and was composed of a random field of fuzzy discs drawn on a blue background. The term "vection" is usually used to refer to the feeling of self movement when a large field display is moved with respect to an observer. Recent evidence indicates that the effect can be achieved with even a small field of view [31]. Howard and Heckman suggest that one of the important factors in eliciting vection is the perceived distance of a moving visual image, with images that are perceived as furthest away contributing the most. In the experiments, we desire the observer to perceive the monitor as a window into an extensive space. We created the background out of discs displayed as though they were an infinite distance from the user (with respect to their position, not their size). The edges of the discs were blurred to give the illusion of depth-of-field. The discs are not intended to be focussed on; they are intended to give a feeling of spaciousness when objects in the foreground are fixated.

For Experiments 2 and 3, a 3D tree tracing task was employed. The scene contained the same vection background as in Experiment 1, and two purple trees consisting of straight line segments in the foreground. The construction of the trees will be discussed

in the section describing the Experiment 2 procedure. A sample pair of trees is shown in Figure 4.5.

4.3 Experiment 1: Subjective Impression of Three-dimensionality

The goal of the first two experiments was to obtain subjective user preferences and performance measurements under various viewing conditions (shown in Figure 4.2), in an effort to evaluate head-coupled stereo display and compare the relative merits of stereo and head-coupling as depth cues. Experiment 1 obtained subjective rankings of the different viewing conditions using two arbitrary scenes. The two scenes were the sphere and bent tube displays discussed earlier (see Figure 4.4). An experimental protocol that involves comparison of randomly selected pairs of conditions was implemented to obtain the rankings.

4.3.1 Procedure

In a given trial, a subject compared the impression of three-dimensionality given by two viewing conditions randomly selected from the five conditions shown in Figure 4.2 and Table 4.1. Two icons, a triangle and a square, were shown in the top left corner of the screen, representing two conditions, with a circle around the icon representing the current condition. The triangle and square icons were used to make it easier to keep track of which condition was active. By pressing the space bar, the subject could change the viewing condition (and the highlighted icon). The subjects were asked to continue toggling between the two conditions until they made a decision as to which condition gave them a better sense of three-dimensionality. At this point they would click on either the left or the right mouse button (marked with a triangle and a square respectively) to indicate the preferred condition. During the trials, the conditions were not identified by

name to the subjects (they were identified only by icon). However, the conditions were described to subjects prior to the experiment.

To further judge the subjects' feelings about head-coupling and stereo, each was asked a set of questions after completion of the experiment.

4.3.2 Design

The construction of the experiment blocks was as follows. Each of the 5 viewing conditions was compared with all the others, making a total of 10 different pairs. The assignment of the two conditions to either the triangle or square icon was random. The 10 pairs were shown once for the sphere scene and once for the bent tube scene. A trial block consisted of these 20 trials in random order. The experiment consisted of two blocks of 20 trials for each subject (a different ordering was used for each block).

Following the comparison trials, each subject was presented with the following set of questions.

> **All of the following questions relate to the quality of the 3D spatial impression.**
>
> *Is head-coupling as important, more important or less important than stereo?*
>
> *Is the combination of head-coupling and stereo better than either alone?*
>
> *Is head-coupling alone worthwhile? (If you had the option would you use it?)*
>
> *Is stereo alone worthwhile? (If you had the option would you use it?)*
>
> *Is head-coupling with stereo worthwhile? (If you had the option would you use it?)*
>
> *Do you have other comments on these methods of displaying 3D data?*

Seven subjects performed the experiment. The subjects were graduate or undergraduate students at the University of British Columbia. All of the subjects were male, and

four of the subjects were familiar with high performance graphics systems.

4.3.3 Results

There were no systematic differences between the data from the sphere scene and the data from the tube scene and so these two sets of data were merged.

Tables 4.2 and 4.3 summarize the combined results from all subjects. Each entry in Table 4.2 corresponds to a pair of viewing conditions. The value is the percentage of the trials in which the row condition was preferred over the column condition. Hence corresponding percentages across the diagonal sum to 100%. For example, the value 89% in row 4 and column 2 means that condition HCB was preferred to condition STE in 25 out of all 28 comparisons (4 responses from each of the 7 subjects). The value of 11% in row 2, column 4 accounts for the other 3 responses in which condition STE was preferred over condition HCB.

Viewing Condition		PIC	STE	HCM	HCB	HCS
PIC	Picture		43%	4%	0%	7%
STE	Stereo only	57%		7%	11%	0%
HCM	HC monocular	96%	93%		29%	61%
HCB	HC binocular	100%	89%	71%		68%
HCS	HC & stereo	93%	100%	39%	32%	

Table 4.2: Pairwise comparison results from Experiment 1. The values in each row correspond to the frequency a particular condition was preferred over each of the other conditions.

The most interesting result apparent from the data is that head-coupling without stereo was preferred over stereo alone by a wide margin of 91% to 9% (averaging the monocular and binocular results).

Table 4.3 shows for each viewing condition the percentage of times it was preferred

Viewing Condition		Frequency
PIC	Picture	13%
STE	Stereo only	19%
HCM	HC monocular	70%
HCB	HC binocular	82%
HCS	HC & stereo	66%

Table 4.3: Summary by viewing condition of the results from Experiment 1. The value in each row corresponds to the frequency a condition was preferred in all of the trials in which it was present.

over all the trials in which that condition was present. The values in the second column sum to $n/2 \times 100\% = 250\%$, where the number of viewing conditions n is 5 in our experiment. Head-coupled display without stereo (both monocular and binocular) was preferred somewhat more than head-coupled display with stereo (although this preference is likely not statistically significant).

The responses to the questions also showed a strong preference for head-coupling. All users said they would use it if it were available. In response to the first question (*"Is head-coupling as important, more important or less important than stereo ?"*), two of the seven subjects stated that they thought stereo was more important than head-coupling. However, these same subjects preferred head-coupling over stereo in the direct comparison task. One subject complained about the awkwardness of the apparatus and pointed out that this would be a factor in how often it would be used. The complete set of responses is included in Appendix A as Tables A.1 through A.6.

4.4   Experiment 2: Performance on a 3D Tree Tracing Task

The second experiment compared the same viewing conditions used in Experiment 1 as measured by performance on a 3D task. This task is based on one used by Sollenberger

and Milgram [53] to study the ability of subjects to trace arterial branching in brain scan data under different viewing conditions. Subjects were asked to answer questions that required tracing leaf-to-root paths in ternary trees in 3-space. The stimulus trees were generated randomly by computer.

4.4.1 Procedure

The experiment stimulus consisted of a scene constructed as follows. Two ternary trees consisting of straight line segments were constructed in 3-space and placed side-by-side so that a large number of the branches overlapped (see Figure 4.5). One leaf of one of the trees was highlighted and the subject was asked to respond as to whether the leaf was part of the left tree or part of the right tree. For each trial, we chose as the highlighted leaf the one whose x coordinate was nearest the center of the screen. The reason for this was to ensure that the task would be reasonably difficult under all viewing conditions.

In each experimental trial, the subject was presented with a scene and asked to click on the left or right mouse button depending on whether the distinguished leaf appeared to belong to the left or the right tree. The bases of the two trees were labeled on the screen with a triangle (the left tree) and a square (the right tree). The corresponding left and right mouse buttons were similarly labeled with a triangle and a square as an additional aid to help subjects remember the labeling.

The trees were recursively defined ternary trees. A trunk of 8.0 cm was drawn at the base of the tree, connected to the root node. Nodes above the root were defined recursively, with the horizontal and vertical positions of the children placed randomly relative to the parent. There were three levels of branches above the root, resulting in 27 leaves for each tree. The following recurrence relation gives a precise specification for one tree. This assumes a right-handed coordinate system with y pointing upwards and x pointing right.

$$X_{base} = Y_{base} = Z_{base} = 0.0$$

$$VerticalSpacing_{root} = 8.0 \text{ cm}$$

$$HorizontalSpacing_{root} = 8.0 \text{ cm}$$

$$X_{root} = X_{base}$$

$$Y_{root} = Y_{base} + VerticalSpacing_{root}$$

$$Z_{root} = Z_{base}$$

$$VerticalSpacing_{child} = 0.7 \times VerticalSpacing_{parent}$$

$$HorizontalSpacing_{child} = 0.7 \times HorizontalSpacing_{parent}$$

$$X_{child} = X_{parent} + HorizontalSpacing_{child} \times Rand()$$

$$Y_{child} = Y_{parent} + VerticalSpacing_{child} \times (1.0 + 0.25 \times Rand())$$

$$Z_{child} = Z_{parent} + HorizontalSpacing_{child} \times Rand()$$

The function $Rand()$ returns a uniform random number in the range $[-1, +1]$. The two trees constructed for each trial were displayed side-by-side separated by a distance of 1.0 cm.

The visual complexity of the trees was tested beforehand, with the goal of making the task difficult enough that depth perception was a factor, but not so difficult that an extreme number of errors would be made by a typical subject. This resulted in the specific parameters that were selected. Figure 4.5 shows an example of the experiment stimuli for one trial.

The experiment tested the same five viewing conditions as in Experiment 1 (see Table 4.1) and subjects wore the stereo glasses and head tracking equipment throughout the experiment. Ten undergraduate and graduate students (nine male and one female), most of whom had experience with computer graphics workstations, served as subjects for the experiment. They were instructed that their error rates and response times were
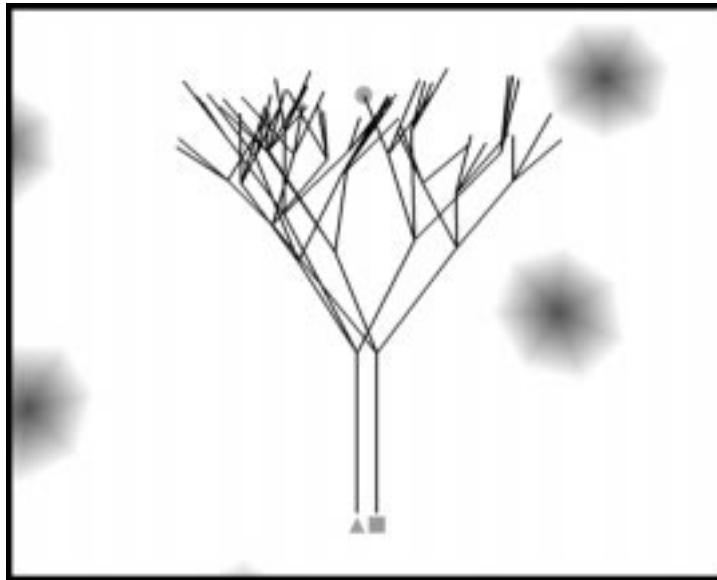
Figure 4.5: An example of the tree display used in Experiments 2 and 3. The colours and background have been modified for black and white reproduction.

being recorded and that they should be most concerned with making as few errors as possible.

4.4.2 Design

A new pair of trees was randomly generated for each trial. The viewing condition was held constant for each group of 22 random trials. The first two trials of each group were designated as practice trials to familiarize the subject with the condition. A trial block consisted of all 5 groups given in a random order, and the entire experiment consisted of 3 such blocks, resulting in a total of 60 trials in each of the 5 experimental conditions. A practice group of 10 trials (two in each condition) was given at the start of the experiment. The stereo test scene (see Figure 4.3) was presented to each subject prior to the experiment to verify the subject's ability to use stereopsis to perceive depth.

4.4.3   Results

The results from Experiment 2 are summarized in Table 4.4. The timing data show that the head-coupled stereo condition was the fastest, but that head-coupling without stereo was slow. There are significant differences at the 0.05 level between condition HCM and condition HCS and between condition HCB and condition HCS, by the Wilcoxon Matched Pairs Signed Ranks Test. The only other difference that is significant is between conditions HCB and PIC.

Viewing Condition		Time (sec)	% Errors
PIC	Picture	7.50	21.8
STE	Stereo only	8.09	14.7
HCM	HC monocular	8.66	3.7
HCB	HC binocular	9.12	2.7
HCS	HC & stereo	6.83	1.3

Table 4.4: Experiment 2 timing and error results

The error data in Table 4.4 provide more significant results, with errors ranging from 21.8% in the static non-stereo condition without head-coupling to 1.3% for the head-coupled stereo condition. All of the differences are significant in pairwise comparisons except for the difference between conditions HCM and HCB, the two head-coupled conditions without stereo.

4.5   Experiment 3: Effects of Lag and Frame Rate

In a head-coupled display system the delay in the display update arises from two primary sources. The first is the delay in receiving and processing physical measurements from the tracker to produce eye position and orientation data. The processing delay is typically due to communication delay and smoothing algorithms, implemented either within the

tracker hardware itself, or on the host computer. The second lag is the delay between receiving the eye positions and updating the display, that is, the time required to compute and render the scene using a perspective projection that takes into account the latest tracker measurements (or two perspective projections when displaying in stereo). This second lag is hence directly related to the frame rate. There is usually a third lag component present due to variations in system load. This component is more difficult to predict and measure and for our purposes we effectively eliminated it as a factor by restricting network access to the workstation during the experiment.

Experiment 3 was designed to investigate the effects of lag and reduced frame rates on performance of a 3D task under head-coupled stereo viewing. In particular we wanted to determine how response times were affected by increasing lag and to compare the relative importance of lag and frame rate.

4.5.1 Procedure

The 3D tree tracing task was used again for this experiment. All of the experimental trials were conducted under the full head-coupled stereo viewing condition (condition HCS). Subjects were informed that the accuracy of their responses and their response times would be recorded. They were instructed to perform the task as quickly as they could without seriously sacrificing the accuracy of their responses. Note that this is different from the instructions given to subjects in Experiment 2, where error rate was considered most important. The reason for this change of focus is due primarily to the low level of difficulty of our task and the fact that the trials were always performed under head-coupled stereo viewing. We reasoned that measuring response times would be most relevant when dealing with the addition of temporal artifacts and that error rates would not vary significantly as presumably a large degree of depth perception can still be obtained through stereopsis and motion (even in high lag conditions where the motion is

not coupled accurately with head movements).

The subjects were ten male graduate students, all of whom had some prior experience using graphics workstations.

4.5.2 Design

The two variables in the experiment were frame rate and simulated tracker lag. Frame rates of 30 Hz, 15 Hz, and 10 Hz were used, and tracker lags of 0, 1, 2, 3, or 4 frame times were simulated. Hence there were $3 \times 5 = 15$ conditions in total. Table 4.5 shows the total lag times resulting from these values. Total lag is defined as

$$TotalLag = TrackerLag + 1.5 \times FrameInterval.$$

The program was synchronized with the internal system clock to run at a maximum frame rate of 30 Hz. The frame rates of 15 Hz and 10 Hz were generated by redrawing frames once or twice, respectively. Tracker lags were simulated by buffering tracker records for a number of frame times. The actual frame rates and lags that were achieved were measured during the experiment to verify the accuracy of the software. The measured times were found to be within 3 milliseconds of the predicted values in all cases.

Subjects were presented with 15 blocks of 22 trials, with lag and frame rate kept constant within blocks. The first two trials in each block were designated as practice trials to enable the user to become familiar with the block's lag and frame rate. The blocks were presented in random order, and an additional block of 22 practice trials with moderate lag and frame rate (15 Hz and 233.3 msec lag) was given at the start of the experiment. The stereo test scene (shown in Figure 4.3) was presented to each subject prior to the experiment to verify the subject's ability to use stereopsis to perceive depth.

Frame Rate			Tracker Lag		
FR (Hz)	Frame time	Base Lag	# frames	Lag	Total Lag
30	33.3	50.0	0	0.0	50.0
30	33.3	50.0	1	33.3	83.3
30	33.3	50.0	2	66.6	116.6
30	33.3	50.0	3	100.0	150.0
30	33.3	50.0	4	133.3	183.3
15	66.6	100.0	0	0.0	100.0
15	66.6	100.0	1	66.6	166.6
15	66.6	100.0	2	133.3	233.3
15	66.6	100.0	3	200.0	300.0
15	66.6	100.0	4	233.3	333.3
10	100.0	150.0	0	0.0	150.0
10	100.0	150.0	1	100.0	250.0
10	100.0	150.0	2	200.0	350.0
10	100.0	150.0	3	300.0	450.0
10	100.0	150.0	4	400.0	550.0

Table 4.5: Experiment 3 conditions (all times are in msec)

4.5.3   Results

Figure 4.6 shows a plot of average response times over all trials and subjects for each of the 15 experimental conditions. The horizontal axis measures the total lag time, and the points are marked according to the different frame rates. Response times ranged from 3.14 to 4.16 seconds. On average, subjects responded incorrectly in 3.4% of the trials. The distribution of errors across conditions showed no distinguishable pattern; there was no significant correlation between errors and total lag ($F(1, 13) = 2.91$, hypothesis is not rejected at $p = 0.10$). A plot of error rates is shown in Figure 4.7.

Figure 4.6: Plot of response time versus total lag for Experiment 3. Each point corresponds to an experimental condition with a particular lag and frame rate (see Table 4.5). The line is the best fit to the linear regression model involving total lag only.

A regression analysis was performed to compare the effect of total lag and frame rate, and in particular to determine whether total lag, frame rate, or a combination of both would best account for the data. Three models were tested using linear regression on the
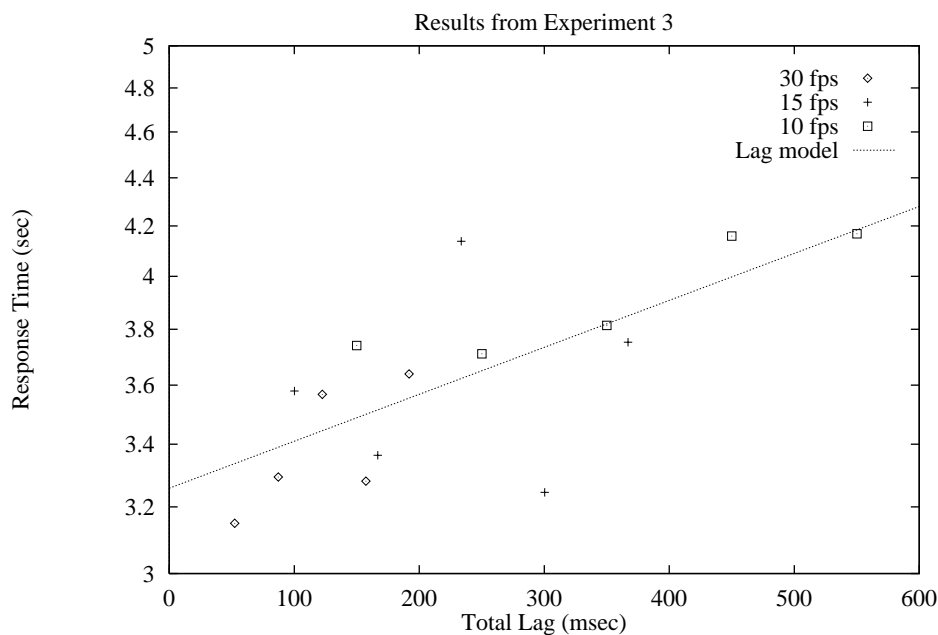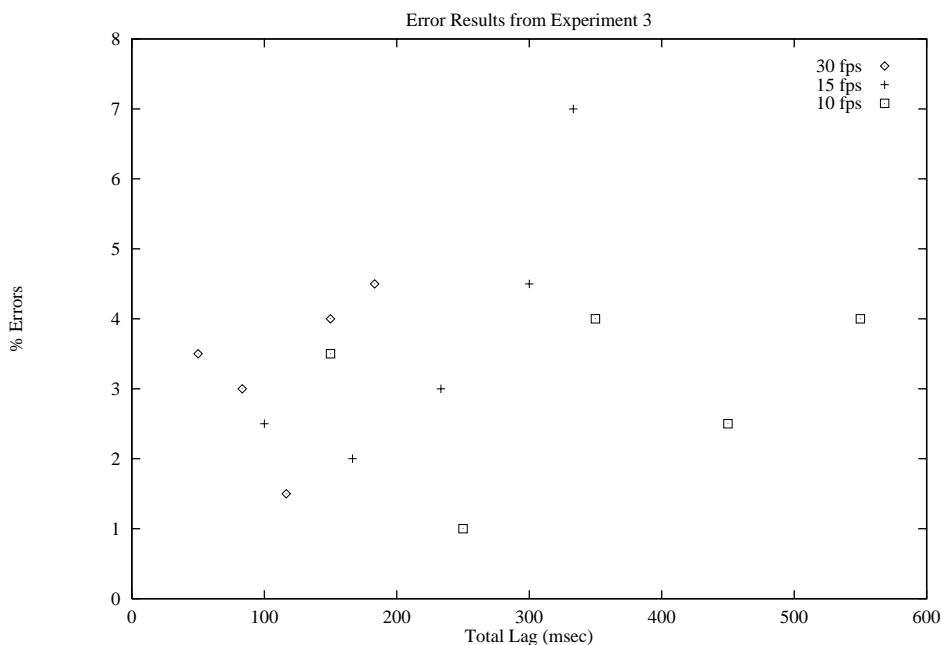
Figure 4.7: Plot of error rates versus total lag for Experiment 3. Each point corresponds to an experimental condition with a particular lag and frame rate (see Table 4.5).

15 averaged points. The models were

Model 1: $\log time = c_1 + c_2 \times TotalLag$

Model 2: $\log time = c_1 + c_2 \times FrameInterval$

Model 3: $\log time = c_1 + c_2 \times TotalLag + c_3 \times FrameInterval$

c_1, c_2, and c_3 are constants. In models 2 and 3, *Frame interval* was used instead of frame rate (frame interval = 1/frame rate) since both lag and frame interval measure time (whereas frame rate has dimensions of 1/time). Model 3 is an additive model which takes both total lag and frame interval into account. The regression line for Model 1 is plotted along with the timing data in Figure 4.6.

Linear regression was performed and the regression constants were found to be the following.

Model 1: $\log time = 0.51 + 0.20 \times TotalLag$

Model 2: $\log time = 0.49 + 0.98 \times FrameInterval$

Model 3: $\log time = 0.49 + 0.13 \times TotalLag + 0.52 \times FrameInterval$

The effectiveness of the regression fit to the data can be measured by the coefficient of determination r^2, which measures the fraction of the variance which is accounted for by the regression model. The r^2 values for each of the three models are listed below. The F-test statistics that are given are for the test of significance of the regression (specifically, the test that the correlation coefficient differs significantly from zero). All three regression tests showed significant correlation.

Model 1: $r^2 = 0.50, F(1, 13) = 13.0, p < 0.005$

Model 2: $r^2 = 0.45, F(1, 13) = 10.6, p < 0.01$

Model 3: $r^2 = 0.57, F(2, 12) = 7.95, p < 0.01$

Model 3 involves multiple linear regression with two variables. A test of significance was performed to determine the strength of this model over Models 1 and 2, which each involve only one variable. Model 3 shows no significant improvement over Model 1 ($F(1, 12) = 1.95$, hypothesis is not rejected at $p = 0.10$), whereas Model 3 does show a moderately significant improvement over Model 2 ($F(1, 12) = 3.35, p < 0.10$). Thus the model which incorporates both total lag and frame interval does not perform significantly better than the model with total lag alone, although it is probably better than the model with frame interval alone.

The three models can be rewritten in terms of tracker lag and frame interval instead of total lag and frame interval, using the relation that

$$TotalLag = TrackerLag + 1.5 \times FrameInterval.$$

The result is the following.

Model 1: $\log time = 0.51 + 0.20 \times Tracker Lag + 0.30 \times Frame Interval$

Model 2: $\log time = 0.49 + 0.98 \times Frame Interval$

Model 3: $\log time = 0.49 + 0.13 \times Tracker Lag + 0.72 \times Frame Interval$

Chapter 5

Discussion

The results from the three experiments suggest a number of interesting conclusions with respect to the relative merits of head-coupling and stereopsis, and the effects of temporal artifacts on 3D task performance.

5.1 Subjective Evaluation of Head-Coupled Stereo Display

Both the comparison results and the positive response of the subjects to the concept of head-coupled display provide evidence for the value of the technique and suggest that applications that use computer graphics to display 3D data could benefit from its use.

An unexpected result from the comparison trials is that on average subjects preferred head-coupled non-stereo display over head-coupled stereo display. This is likely due to the ghosting present when displaying stereo images. The monitor we used, which is typical of common monitors used for computer graphics, is not optimized for stereo display: the phosphor decay times are longer than is desirable and hence there is some cross-talk between the left and right eye images. When objects are displayed without stereo, the image tends to appear sharper than stereo images because there is no ghosting. Subjects might have preferred the sharpness of the non-stereo images to the ghosted stereo images, despite the added advantage of stereopsis.

Subjects also mentioned the discomfort of the head tracker and this may have affected subjects responses to the questions following the comparison trials. Two of the

46

subjects said that they thought stereo was more important than head-coupling yet the same subjects preferred head-coupling in the comparison task. Aside from the tracker discomfort, another possible reason for this apparent bias towards stereo is the fact that stereoscopic 3D is a technique which is already well known to most people, either from similar use with graphics workstations or through 3D movies. In comparison, the head-coupling technique is much less well known. The awkwardness of head tracking is likely to become less of a problem with advances in tracking, and in fact many other currently available trackers are less intrusive than the mechanical tracker we used [41].

5.2 3D Task Performance

Experiment 2 provides objective evidence of the value of head-coupled stereo display. Error rates in the tree tracing task using head-coupled stereo were significantly lower than the error rates obtained under any of the other viewing conditions. The results also show that head-coupling alone is significantly better for this type of task than stereo viewing alone. Hence the results suggest that, if possible, head-coupling and stereo should both be implemented, but if only one of the two techniques must be chosen, then head-coupling should be given preference for this type of task.

Another factor to consider when choosing between head-coupling and stereo is the relative computational expense of the two techniques. To implement head-coupling in an interactive graphics application, all that is required is that the program change the perspective projection with each frame update, and hence the frame rate of the program is not reduced appreciably. Stereo requires that two images be generated and drawn for each frame and thus halves the frame rate of the program. Many techniques for creating stereo images, including using shutter glasses and field sequential display, also have the effect of reducing the vertical resolution of the frame buffer by half. As display hardware

becomes faster, this same factor of two will remain, but the time required to adjust the perspective projection will almost vanish for a given investment in hardware cost.

The fact that motion parallax (through head-coupling) outperforms binocular parallax (stereo) is not surprising and is supported by theories of visual perception promoted by Gibson and others [28]. The results are also similar to those obtained by Sollenberger and Milgram in their comparison of stereo and rotational depth cues [53]. Overall, the error rates obtained in the tree tracing task are lower than those obtained by Sollenberger and Milgram, but the pattern is very similar despite the differences in the stimulus trees, the viewing conditions and the experimental protocols. Both studies found motion to be more important than stereo, even though our motion was due to head-coupling rather than simple rotations of the object, as was the case in the study by Sollenberger and Milgram. Both studies found combined motion and stereo to be more effective than either in isolation. However, our data does not provide very much information about the extent of the benefit of combined stereo and head-coupling. Because the error rates from Experiment 2 were too close to zero in the head-coupled and head-coupled stereo conditions, subjects' performance could not increase much further.

It can be argued that the improvements seen with head-coupling in the tree tracing task are not due to the head-coupling as such, but rather to motion-induced depth perception [62]. Our current evidence does not counter this objection. However, it is likely that the image motion produced by dynamic head-coupled perspective is less distracting than techniques such as rocking the scene back and forth about a vertical axis, which is commonly done in commercial molecular modelling and volume visualization packages.

A more complete study would compare the benefits of dynamic head-coupled perspective with the benefits of motion, rotational or otherwise, under stereo and non-stereo conditions. One would expect the performance of stereo viewing to improve with the introduction of motion, and that the difference between head-coupled with motion and

stereo with motion would not be as pronounced as it was in our study.

Our timing data from Experiments 2 and 3 show an apparent inconsistency. Because the head-coupled stereo condition of Experiment 2 is comparable to the best lag and frame rate condition of Experiment 3, one would expect the response times to be comparable. However, the best response times from Experiment 3 are approximately half of the best response times from Experiment 2. This discrepancy is likely due to the fact that subjects were given slightly different instructions for Experiment 3 because response time was more important than error rate. In Experiment 2, subjects tended to be more careful to minimize errors at the expense of response time. This is supported by the fact that in Experiment 2 with head-coupled stereo the average error rate was 1.3%; in Experiment 3 this grew to 3.4%.

Another problem with Experiment 2 concerns the selection of stimulus trees. For each trial, the program generated a new random tree and thus the set of trees used under one condition would be different from the set used for another, although on average they should be roughly comparable in difficulty. There are two primary reasons why a more careful selection procedure was not used. When conducting the first two experiments we were having difficulties with communications with the tracker: in a small percentage of the trials, the records would become corrupted and the display would be unpredictable. When this occurred, the tracker was reset and the trial was restarted, with new randomly generated trees. Hence some of the trees would be thrown out and we could not rely on a precomputed set of trees. The problem with the tracker was solved before Experiment 3 was conducted and so the difficulties did not occur there. The second reason for using randomly selected trees was due to the large amount of time that would have to be spent to select "good" trees, and also the difficulty of deciding what in fact are good trees, without introducing any bias towards trees that were "better" in some conditions compared to the others. In our randomly selected trees, occasionally the solution would

be almost immediately recognizable, with very few overlapping branches, but on the other hand, some scenes would be very difficult, with a very dense overlapping of branches. This is not to say that a sophisticated method for selecting trees would be impossible, but rather that we decided that the random selection method was most appropriate for the scope of this study.

5.3 Effects of Lag and Frame Rate

Experiment 3 provides information about the importance of lag and frame rate on 3D task performance. Of the three regression models, the model that accounts for the most of the variance is Model 3, which takes both lag and frame rate into account, as expected. The fact that this only accounts for 57% of the variance is likely due to the random nature of the stimulus trees leading to a wide variance in response times within conditions.

The comparison of the relative importance of the two variables of lag and frame rate showed no significant difference between the strength of the model which took only lag into account (Model 1) and the model which took both lag and frame rate into account (Model 3), whereas there was a moderately significant difference in the relative strength of Model 2, which involved frame rate only, and Model 3. This suggests that lag itself accounts reasonably well for the performance degradations observed, and that lag is probably a more important temporal artifact than frame rate in its effect on the performance of similar tasks.

Given the data describing performance fall-off as lag increases, it is useful to obtain some measure of what level of lag becomes prohibitive. Specifically we would like to know the lag value that makes performance under head-coupled stereo viewing worse than the performance would be without head-coupling or stereo. We can compare the results from Experiments 2 and 3 to obtain an approximate cut-off value by finding a point on the

regression line in Figure 4.6 where response time is the same as for the static viewing condition.

The analysis becomes complicated because our range of response times for Experiment 3 was lower than that for Experiment 2 due to the differing instructions given to subjects, as was discussed in the previous section. In Experiment 2 we found a best case response time of 6.83 seconds, whereas in Experiment 3 under the same conditions the response time was 3.25 seconds, which is a factor of 2.10 less. The Experiment 2 average response time for static viewing was 7.50 seconds. If we scale this by the same factor of 2.10, we find that it corresponds to a response time of 3.58 seconds under the conditions of Experiment 3. From the plot of the first regression model (Figure 4.6), this corresponds to a total lag of 210 milliseconds. This suggests that for tasks similar to the tree tracing task, lag above 210 milliseconds will result in worse performance, in terms of response time, than static viewing. Note that due to the large variance in our data it is difficult to say how accurate or significant this lag cutoff value is, and how relevant it is for other tasks.

The error rates for Experiment 3 remained low under all conditions, averaging to 3.4%, in contrast to Experiment 2 where the number of errors rose significantly in the non-head-coupled conditions. This suggests that even in the presence of large lags and low frame rates, head-coupling provides some performance improvement. This is not surprising, however, because the effects are likely due to motion-induced depth; while we do not have the data to verify this, we suspect that performance is similar to what it would be if the scene were moving independent of the user, even without head-coupled viewing.

Systems that use predictive methods such as Kalman filtering must make a compromise between the size of the prediction interval and noise artifacts that become worse as this interval increases. Introducing prediction into a system will effectively flatten out

the low-lag portion of the curve in Figure 4.6 and hence there will be a cut-off point beyond which the lag caused by filtering artifacts becomes unacceptable.

The level of lag in commercial trackers can be expected to improve in the future through hardware improvements and improved prediction techniques. However, our results and estimates given by other researchers suggest that lags as low as 100 milliseconds can be disruptive. While the quality of commercial trackers can be expected to improve in the future with respect to lag, the current technology is such that many trackers introduce on the order of 75 ms or more. Even a very fast system by today's standards with a 20 Hz frame rate and a tracker with 25 ms lag will generate a total lag of at least 100 ms. Studies of lag are even more relevant for applications such as telerobotics where long communication lines can create delays of several seconds.

Lag tends not only to have an effect on task performance but can also contribute to motion sickness. This is a problem that is very important for the practical use of virtual reality, in particular for immersive systems where the conflict between the senses is more apparent.

5.4 Applications

The tree tracing task used in our experiments maps directly to several current applications of interactive 3D graphics. The first, and in fact the application out of which the task arose, is medical visualization of brain scan data where doctors may wish to trace paths through complex networks of blood vessels [53]. The tree structure is also similar to software visualization techniques that display modules with nodes connecting them to represent object dependencies [48]. Our results are also likely to be applicable to many other 3D tasks that can benefit from an extra sense of depth.

The head-coupling technique is not limited exclusively to systems that use computer

graphics. As reported by various researchers [11][47], there is potential for teleoperation using cameras that are isomorphically coupled to the user's head. In fact, teleoperation is one application where techniques such as rotating the scene may be difficult or even impossible to implement if dealing with a large working scene (in the real world), whereas head-coupling is relatively easy since it only involves moving cameras.

Chapter 6

Conclusions

This thesis presents a discussion of the technique of head-coupled stereo display, an examination of the issues involved in implementing it correctly, and experimental studies to investigate its effectiveness.

Head-coupled stereo display, or "fish tank virtual reality", is similar to immersive virtual reality systems in that head tracking and stereoscopy are employed to provide an added sense of three-dimensionality. A number of reasons, including the state of current head-mounted display technology and the impracticality of immersion in some situations, make monitor-based head-coupled display a more practical choice for many applications.

While the technique can be implemented with commercially available hardware, there are a number of important issues, including accurate calibration of the system and the minimization of temporal inaccuracies, that are important for properly implementing the technique. The effect of lag is particularly relevant to both immersive and non-immersive virtual reality systems.

Three experiments were conducted to evaluate the effectiveness of head-coupled stereo display and to investigate the effects of temporal artifacts. The results of the first two experiments showed strong evidence for the value of the technique, both through sub-jective user preference tests and objective measures of 3D task performance. For the 3D tree tracing task we tested, the results suggest that the head-coupling technique alone is more beneficial than is stereo alone. Combined head-coupling and stereo provided the

most improvement.

The third experiment provides an indication of the seriousness of the effect of lag on user performance. Subjects' response times increased dramatically as lag increased and compared with the effect of low frame rates, it appears that lag alone accounts reasonably well for the degradation. This would suggest that designers of virtual reality systems should make it a priority to employ very low-lag tracking devices and effective prediction methods. The advantage of faster tracking is likely more significant than the advantage of having a very fast graphics display.

6.1   Future Work

The studies presented here represent initial attempts at characterizing user performance using head-coupled stereo displays and hence are necessarily limited in scope. There are a number of issues regarding 3D performance using the technique that require further study.

6.1.1   Experimental Studies

The first two experiments compared the relative benefits of head-coupling and stereo. In most cases (in particular with the stimuli for Experiment 1) we endeavoured to provide as many depth cues from techniques *other than* head-coupling and stereo as possible, including specular highlights, shadows and a vection background. We neglected the depth cues possible from motion-induced depth, however, and the scenes we displayed were all static in space. This is somewhat unrealistic: typically an application that can display scenes at a high frame rate (high enough that head-coupling can be employed) will take advantage of motion and allow the scene to be moved by the user. A more complete analysis of the benefits of head-coupling would compare performance with motion and

without. It is likely that with a moving scene, the performance difference between head-coupling alone and stereo alone would not be as large as the difference seen in our study. The relative effects of motion will of course depend on how the motion is controlled, whether through automatic or manual control by the user, and also on what type of device is used to control the motion.

Another important area for future studies is to investigate the performance of other tasks under head-coupled stereo viewing and how the effects vary with the size of the display (and whether the display is immersive).

Our third experiment gives some initial indication of the effects of lag and frame rate on user performance. However, the design of the experiment, and the selection of the experimental conditions, did not permit us to obtain a clear comparison between the effects of frame rate and lag. There is a need for closer investigation of the cut-off values where lag becomes prohibitive; this could be accomplished better by choosing a finer set of lag conditions to focus in on particular regions.

6.1.2 Extensions to Head-Coupled Stereo Display

An important problem that has not been fully addressed in our implementation is proper calibration of the system. Our system has been calibrated approximately and the display parameters are not adjusted individually for different users. There is a need for techniques to calibrate efficiently and accurately for distortions both from the tracker and from the screen, and to interactively calibrate for different users, adjusting for different eye spacings and eye positions relative to the head tracker. There would also be value in studies that evaluate just how accurate the display needs to be, so that an appropriate balance could be met between the accuracy and the expense of different calibration methods.

There are many interesting possibilities for extending head-coupled stereo displays beyond the implementation described here. With larger screens it is possible to obtain

the effects of head-coupling and stereo and also provide a sense of immersion [12][16]. Large or multiple screen displays have the advantages of immersion yet do not suffer from the problems of head-mounted displays, such as wide angle distortions and physical discomfort.

One drawback to head-coupled display in comparison to traditional display is that the display can only be viewed effectively by one user at a time. The stereo images and perspective distortions become distracting to other people looking on from the side. A possible solution to this is to use the same technology as field sequential stereo to effectively multiplex the display between different users (instead of different eyes), and track multiple head positions. While the cost may be too expensive to make this technique practical in all but very high-end applications, it may be feasible in limited situations where only two or three users are presented with non-stereo head-coupled display.

Another interesting technique, suggested by McKenna's work [40], is to use small movable displays that are tracked. A small LCD display in an augmented reality-type application might be preferable to see-through head-mounted displays, as the display could be shared between users and wouldn't have to be worn.

By definition, head-coupled display implies using head position to adjust the perspective projection used in displaying 3D scenes. Given that the application knows where the user's head is, it can make further use of this information. An interesting interaction technique which we have experimented with is head-coupled scene rotation. Objects are made to rotate in a matter analogous to 2D virtual trackball techniques, but with the rotation defined by head position rather than mouse position. As the user moves his head, the scene can be made to rotate in an opposite direction to give him a more complete view of it.

Appendix A

Experiment Results

The following 6 tables list the answers given by the seven subjects in Experiment 1 in response to questions administered after the experiment.

	Is head-coupling as important, more important or less important than stereo?
1	head coupling is more important
2	head coupling is more important
3	head coupling is less important than stereo
4	head coupling is more important
5	head coupling is more important
6	head coupling is more important
7	Less important

Table A.1: Experiment 1 subject comments from Question 1.

	Is the combination of head-coupling and stereo better than either alone?
1	Yes
2	Head coupling and head coupling with stereo seem roughly the same
3	Yes
4	No, I prefer head coupling alone
5	Yes
6	It is a close call between head coupling alone and head coupling with stereo.
7	Yes definitely

Table A.2: Experiment 1 subject comments from Question 2.

	Is head-coupling alone worthwhile? (If you had the option would you use it?)
1	Yes
2	Yes
3	Yes
4	Yes
5	Yes
6	Only problem is discomfort. For some visualization tasks head coupling would be worthwhile.
7	Yes

Table A.3: Experiment 1 subject comments from Question 3.

	Is stereo alone worthwhile? (If you had the option would you use it?)
1	Yes
2	No
3	Yes
4	No
5	Yes
6	Yes. But only sparingly. The glasses are less of a problem than the head mount.
7	Yes

Table A.4: Experiment 1 subject comments from Question 4.

	Is head-coupling with stereo worthwhile? (If you had the option would you use it?)
1	Yes
2	Yes
3	Yes
4	No – stereo is too much of a hassle, it dims the view and does not add much
5	Yes
6	Given head linking I would not bother with stereo. This is mainly because of the problem wearing both the glasses and the head mount.
7	Yes

Table A.5: Experiment 1 subject comments from Question 5.

Do you have other comments on these methods of displaying 3D data?
1	
2	Motion is important for 3D
3	Hard to tell difference between stereo and head coupling
4	In general stereo made the images less crisp. When choosing between a crisp non-moving image and a fuzzy stereo image which was moving, the fuzzy stereo images was chosen.
5	Background gives a good feeling of space as did the shading
6	Found head coupling very effective. Very positive first impression. One eye was sometimes better than both. Ghosting is worse on the sphere scene.
7	did not notice a difference between one eye and two eye conditions

Table A.6: Experiment 1 subject comments from Question 6.

	Subjects									
Condition	1	2	3	4	5	6	7	8	9	10
Picture	20	14	15	12	12	10	15	9	7	17
Stereo	7	9	5	11	6	13	11	13	3	10
HC Monocular	8	0	2	4	1	1	0	1	3	2
HC Binocular	6	3	1	2	0	3	0	0	2	2
HC Stereo	3	0	0	1	0	0	1	0	1	2

Table A.7: Experiment 2 errors by subject.

	Subjects									
Condition	1	2	3	4	5	6	7	8	9	10
Picture	8.02	6.02	6.02	4.88	7.28	5.70	8.77	7.39	10.68	8.23
Stereo	12.32	5.56	5.22	7.04	6.49	4.83	7.07	8.87	10.47	7.03
HC Monocular	11.74	8.63	7.11	6.22	6.68	5.81	9.75	9.21	9.12	11.18
HC Binocular	14.52	10.21	7.51	6.00	8.62	5.95	10.19	7.20	8.08	11.82
HC Stereo	10.96	6.97	4.48	5.51	5.60	4.34	6.15	8.55	7.70	7.92

Table A.8: Experiment 2 response times by subject (for correct responses only).

		Subjects									
FR	Lag	1	2	3	4	5	6	7	8	9	10
30	0	3.30	2.57	3.15	3.22	2.42	2.68	2.55	4.17	4.33	4.58
30	1	3.12	2.27	2.59	3.44	2.71	2.64	2.99	4.28	4.20	6.06
30	2	3.99	2.65	3.51	2.90	2.56	2.75	3.26	5.10	5.20	5.13
30	3	3.52	2.19	2.64	2.88	2.59	2.61	3.11	4.80	4.50	5.17
30	4	2.82	2.51	4.15	4.51	2.72	2.46	3.11	5.21	5.67	4.76
15	0	3.93	3.09	3.37	3.43	2.68	2.99	2.83	5.13	3.71	5.65
15	1	2.48	2.45	3.05	3.29	2.62	2.71	3.11	5.07	4.53	5.71
15	2	4.81	2.41	3.55	8.64	2.24	2.88	3.81	6.40	4.57	5.15
15	3	2.37	2.35	3.38	3.64	2.20	2.72	3.87	3.08	5.86	4.23
15	4	4.81	2.42	3.58	4.19	2.55	2.71	3.34	4.59	5.54	5.39
10	0	4.19	2.62	3.25	4.04	3.47	2.63	3.64	5.06	4.26	4.98
10	1	3.36	3.22	3.00	3.36	2.00	3.53	3.96	4.62	5.03	6.89
10	2	4.13	2.41	3.23	5.09	2.72	2.94	3.31	4.58	5.60	5.68
10	3	7.60	3.39	3.12	4.21	2.50	3.30	3.77	6.76	4.23	5.29
10	4	5.10	3.44	4.44	3.53	2.30	3.06	3.47	5.51	5.39	7.28

Table A.9: Experiment 3 response times by subject (for correct responses only).

Appendix B

Head-Coupled Stereo Display Software

The head-coupled stereo display experiment and demonstration software was implemented using the Silicon Graphics GL library. The software has been run on SGI workstations as well as IBM RS/6000 workstations. The C function `draw_hc_stereo_scene()` displays a scene in head-coupled stereo. It assumes that there is a function `get_tracker_eyepos()` that returns the positions of the user's eyes obtained through a head tracker, and that there is a function `draw_scene()` that draws the scene centered at the origin. The function `draw_view()` is called to draw a single head-coupled view for each eye, using the GL `window()` function.

```
#include <gl.h>
#include <stdio.h>

#define Lx  0
#define Ly  0
#define Hx  1280
#define Hx  1024

#define YMAXSTEREO 491
#define YOFFSET 532


/* Draw a scene using head-coupled stereo display.
This assumes the monitor is already in stereo mode */

void draw_hc_stereo_scene()
{
    float L_eye[3], R_eye[3];

    get_tracker_eyepos(L_eye, R_eye);
```

```
        /* draw right eye view */

    viewport(0, XMAXSCREEN, 0, YMAXSTEREO);
    draw_view(R_eye);


        /* draw left eye view */

    viewport(0, XMAXSCREEN, 0, YMAXSTEREO);
    draw_view(L_eye);
}



/* Draw a view for a single eye position */

void draw_view(float eye[3])
{
    Coord left, right, bottom, top, near, far;
    static Matrix Identity =
        {1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1};


        /* set up the off-axis perspective projection, with the eye at the origin looking
           down the positive z-axis */


    left   = Lx - eye[0];
    right  = Hx - eye[0];
    bottom = Ly - eye[1];
    top    = Hy - eye[1];

    near = eye[2];

      /* far clipping plane – 10000 is arbitrary */

    far  = 10000.0 + eye[2];

    loadmatrix(Identity);

    window(left, right, bottom, top, near, far);


        /* draw the background */


    cpack(0x00000000);
    clear(); zclear();
```

/* move the clipping plane out of the screen */

```
scale(4.0, 4.0, 4.0);
```

/* move view frustum according to eye position by doing the opposite translation; this moves the center of the viewport to the world origin */

```
translate(-eye[0], -eye[1], -eye[2]);

draw_scene();
}
```

Bibliography

[1] Adelstein, Bernard D., Eric R. Johnston, and Stephen R. Ellis. "A testbed for characterizing dynamic response of virtual environment spatial sensors". *Proceedings of the ACM Symposium on User Interface Software and Technology*, pp. 15–22, 1992.

[2] Bajura, Michael, Henry Fuchs, and Ryutarou Ohbuchi. "Merging virtual objects with the real world: Seeing ultrasound imagery within the patient". *Computer Graphics (SIGGRAPH '92 Proceedings)*, Vol. 26, No. 2, pp. 203–210, July 1992.

[3] Blanchard, Chuck, Scott Burgess, Young Harvill, Jaron Lanier, Ann Lasko, Mark Oberman, and Michael Teitel. "Reality Built For Two: A Virtual Reality Tool". *Computer Graphics (1990 Symposium on Interactive 3D Graphics)*, Vol. 24, No. 2, pp. 35–36, March 1990.

[4] Booth, Kellogg S., M. Phillip Bryden, William B. Cowan, Michael F. Morgan, and Brian L. Plante. "On the parameters of human visual performance: An investigation of the benefits of antialiasing". *IEEE Computer Graphics and Applications*, Vol. 7, No. 9, pp. 34–41, September 1987.

[5] Brooks, Frederick P., Jr. "Walkthrough — A dynamic graphics system for simulating virtual buildings". *Proceedings of 1986 Workshop on Interactive 3D Graphics*, pp. 9–21, October 1986.

[6] Bryson, S. and S. Fisher. "Defining, Modeling and Measuring System Lag in Virtual Environments". *Proceedings of the 1990 SPIE Conference on Stereoscopic Displays and Applications*, Vol. 1256, pp. 98–109, 1990.

[7] Bryson, S. "Measurement and Calibration of Static Distortion in Three-Dimensional Magnetic Trackers". *Proceedings of the 1992 SPIE Conference on Stereoscopic Displays and Applications III*, Vol. 1669, 1992.

[8] Carlbom, I. and J. Paciorek. "Planar geometric projections and viewing transformations". *ACM Computing Surveys*, Vol. 10, pp. 465–502, December 1978.

[9] Clifton, T.E. III and Fred L. Wefer. "Direct volume display devices". *IEEE Computer Graphics and Applications*, Vol. 13, No. 4, pp. 57–65, July 1993.

[10] Codella, Christopher, Reza Jalili, Lawrence Koved, J. Bryan Lewis, Daniel T. Ling, James S. Lipscomb, David A. Rabenhorst, Chu P. Wang, Alan Norton, Paula Sweeney, and Greg Turk. "Interactive simulation in a multi-person virtual world". *Proceedings of CHI '92 Conference on Human Factors in Computing Systems*, pp. 329–334, April 1992.

[11] Cole, Robert E., John O. Merritt, Richard Coleman, and Curtis Ikehara. "Teleoperator performance with virtual window display". *Proceedings of the 1991 SPIE Conference on Stereoscopic Displays and Applications II*, Vol. 1457, pp. 111–119, 1991.

[12] Cruz-Neira, Carolina, Daniel J. Sandin, Thomas A. DeFanti, Robert V. Kenyon, and John C. Hart. "The CAVE Audio Visual Experience Automatic Virtual Environment". *Communications of the ACM*, pp. 64–72, June 1992.

[13] Cutting, J.E. "On the efficacy of cinema, or what the visual system did not evolve to do". *Pictorial communication in virtual and real environments*. pp. 486–495. Taylor and Francis, 1991.

[14] Dallas, W.J. "Computer-generated holograms". *Topics in Applied Physics Volume 41, The Computer in Optical Research: Methods and Applications*. pp. 291–366. Springer-Verlag, 1980.

[15] Deering, Michael. "High resolution virtual reality". *Computer Graphics (SIGGRAPH '92 Proceedings)*, Vol. 26, No. 2, pp. 195–202, July 1992.

[16] Deering, Michael F. "Making virtual reality more real: Experience with the Virtual portal". *Proceedings of Graphics Interface '93*, pp. 219–226, May 1993.

[17] Diamond, R., A. Wynn, K. Thomsen, and J. Turner. "Three-dimensional perception for one-eyed guys, or the use of dynamic parallax". *Computational Crystallography*, pp. 286–293, 1982.

[18] Ellis, Stephen R. "Nature and origins of virtual environments: A bibliographic essay". *Computing Systems in Engineering*, Vol. 2, No. 4, pp. 321–347, 1991.

[19] Ellis, Stephen R., M.K. Kaiser, and A.J. Grunwald, editors. *Pictorial communication in virtual and real environments*. Taylor and Francis, 1991.

[20] Feiner, Steven, Blair MacIntyre, and Doree Seligmann. "Annotating the real world with knowledge-based graphics on a see-through head-mounted display". *Proceedings of Graphics Interface '92*, pp. 78–85, May 1992.

[21] Fisher, Scott S. "Viewpoint dependent imaging: An interactive stereoscopic display". *Processing and Display of Three-Dimensional Data, Proc. SPIE Int. Soc. Opt. Eng.*, Vol. 367, pp. 41–45, 1982.

[22] Fisher, S.S., M. McGreevy, J. Humphries, and W. Robinett. "Virtual environment display system". *Proceedings of 1986 Workshop on Interactive 3D Graphics*, pp. 77–87, October 1986.

[23] Foley, J.D. and A. van Dam. *Fundamentals of Interactive Computer Graphics*. Addison-Wesley Publishing Company, 1982.

[24] Foley, J.D., A. van Dam, Steven K. Feiner, and John F. Hughes. *Computer Graphics: Principles and Practice*. Addison-Wesley Publishing Company, second edition, 1990.

[25] Friedmann, Martin, Thad Starner, and Alex Pentland. "Device synchronization using an optimal linear filter". *Computer Graphics Special Issue (1992 Symposium on Interactive 3D Graphics)*, Vol. 26, pp. 57–62, March 1992.

[26] Fuchs, H., S.M. Pizer, L.C. Tsai, S.H. Bloomberg, and E.R. Heinz. "Adding a true 3-D display to a raster graphics system". *IEEE Computer Graphics and Applications*, Vol. 2, pp. 73–78, September 1982.

[27] Furness, T.A. "Harnessing virtual space". *Proceedings of the 1988 SID International Symposium*, pp. 4–7, 1988.

[28] Gibson, J.J. *The ecological approach to visual perception*. Houghton Mifflin, Boston, 1979.

[29] Grotch, S.L. "Three-dimensional and stereoscopic graphics for scientific data display and analysis". *IEEE Computer Graphics and Applications*, Vol. 3, No. 8, pp. 31–43, November 1983.

[30] Hodges, Larry F. "Tutorial: Time-multiplexed stereoscopic computer graphics". *IEEE Computer Graphics and Applications*, Vol. 12, No. 2, pp. 20–30, March 1992.

[31] Howard, I. P. and T. Heckman. "Circular vection as a function of the relative sizes, distances and positions of two competing visual displays". *Perception*, Vol. 18, No. 5, pp. 657–665, 1989.

[32] Howlett, Eric M. "Wide angle orthostereo". *Proceedings of the 1990 SPIE Conference on Stereoscopic Displays and Applications*, Vol. 1256, pp. 210–223, 1990.

[33] Jacob, Robert J.K. "What you look at is what you get: Eye movement-based interaction techniques". *Proceedings of CHI '90 Conference on Human Factors in Computing Systems*, pp. 11–18, April 1990.

[34] Kalman, R.E. and R.S. Bucy. "New results in linear filtering and prediction theory". *Transactions of ASME (Journal of basic engineering)*, Vol. 83d, pp. 95–108, 1961.

[35] Krueger, M.W. *Artificial Reality II*. Addison-Wesley Publishing Company, 1991.

[36] Kubitz, W.J. and W.J. Poppelbaum. "Stereomatrix, an interactive three dimensional computer display". *Proceedings of the Society for Information Display*, Vol. 14, No. 3, pp. 94–98, 1973.

[37] Liang, Jiandong, Chris Shaw, and Mark Green. "On temporal-spatial realism in the virtual reality environment". *Proceedings of the ACM Symposium on User Interface Software and Technology*, pp. 19–25, 1991.

[38] MacDowall, I., M. Bolas, S. Pieper, S. Fisher, and J. Humphries. "Implementation and integration of a counterbalanced CRT-based stereoscopic display for interactive viewpoint control in virtual environment applications". *Proceedings of the 1990 SPIE Conference on Stereoscopic Displays and Applications*, Vol. 1256, 1990.

[39] MacKenzie, I. Scott and Colin Ware. "Lag as a determinant of human performance in interactive systems". *Proceedings of INTERCHI '93 Conference on Human Factors in Computing Systems*, pp. 488–493, April 1993.

[40] McKenna, Michael. "Interactive viewpoint control and three-dimensional operations". *Computer Graphics Special Issue (1992 Symposium on Interactive 3D Graphics)*, Vol. 26, pp. 53–56, March 1992.

[41] Meyer, Kenneth and Hugh L. Applewhite. "A survey of position trackers". *Presence*, Vol. 1, No. 2, pp. 173–200, 1992.

[42] Muirhead, J.C. "Variable focal length mirrors". *Review of Scientific Instruments*, Vol. 32, pp. 210–211, 1961.

[43] Newman, William M. and Robert F. Sproull. *Principles of Interactive Computer Graphics*. McGraw-Hill, 1973.

[44] Newman, William M. and Robert F. Sproull. *Principles of Interactive Computer Graphics*. McGraw-Hill, second edition, 1979.

[45] Paley, W.B. "Head-tracking stereo display: Experiments and applications". *Proceedings of the 1992 SPIE Conference on Stereoscopic Displays and Applications III*, pp. 84–89, 1992.

[46] Pausch, Randy. "Virtual reality on five dollars a day". *Proceedings of CHI '91 Conference on Human Factors in Computing Systems*, pp. 265–270, April 1991.

[47] Pepper, R.L., R.E. Cole, and E.H. Spain. "Influence of camera separation and head movement on perceptual performance under direct and TV-displayed conditions". *Proceedings of the Society for Information Display*, Vol. 24, No. 1, pp. 73–80, 1983.

[48] Robertson, G. G., J. D. Mackinlay, and S. K. Card. "Cone trees: animated 3D visualizations of hierarchical information". *Proceedings of CHI '91 Conference on Human Factors in Computing Systems*, pp. 189–194, April 1991.

[49] Robinett, Warren and Jannick P. Rolland. "A computational model for the stereoscopic optics of a head-mounted display". *Presence*, Vol. 1, No. 1, pp. 45–62, 1992.

[50] Rogers, D.F. and J.A. Adams. *Mathematical Elements for Computer Graphics*. McGraw-Hill, 1976.

[51] Sedgwick, H.A. "The effects of viewpoint on the virtual space of pictures". *Pictorial communication in virtual and real environments*. pp. 460–479. Taylor and Francis, 1991.

[52] Shepard, R.N. and J. Metzler. "Mental rotation of three-dimensional objects". *Science*, Vol. 171, pp. 701–703, 1971.

[53] Sollenberger, Randy L. and Paul Milgram. "A comparative study of rotational and stereoscopic computer graphic depth cues". *Proceedings of the Human Factors Society 35th Annual Meeting*, pp. 1452–1456, 1991.

[54] Suetens, P., D. Vandermeulen, A. Oosterlinck, J. Gybels, and G. Marchal. "A 3-D Display System with Stereoscopic, Movement Parallax and Real-time Rotation Capabilities". *Proceedings of the SPIE, Medical Imaging II: Image Data Management and Display (Part B)*, Vol. 914, pp. 855–861, 1988.

[55] Sutherland, Ivan. "The ultimate display". *Proceedings of IFIP Congress*, pp. 506–508, 1965.

[56] Sutherland, Ivan. "A head-mounted three dimensional display". *Fall Joint Computer Conference, AFIPS Conference Proceedings*, Vol. 33, pp. 757–764, 1968.

[57] Teitel, Michael A. "The Eyephone, a head mounted stereo display". *Proceedings of the 1990 SPIE Conference on Stereoscopic Displays and Applications*, Vol. 1256, pp. 168–171, 1990.

[58] Tessman, Thant. "Perspectives on stereo". *Proceedings of the 1990 SPIE Conference on Stereoscopic Displays and Applications*, Vol. 1256, pp. 22–27, 1990.

[59] Tharp, G., A. Liu, and L.W. Stark. "Timing considerations in helmet mounted display performance". *Proceedings of the SPIE Conference on Human Vision, Visual Processing and Digital Display III*, Vol. 1666, 1992.

[60] Tyler, William. "Induced stereo movement". *Vision Research*, Vol. 14, pp. 609–613, 1974.

[61] Venolia, D. and L. Williams. "Virtual integral holography". *Proceedings of the SPIE, Extracting Meaning from Complex Data: Processing, Display, Interaction*, Vol. 1259, pp. 99–105, 1990.

[62] Wallach, H. and D. H. O'Connell. "The kinetic depth effect". *Journal of Experimental Psychology*, Vol. 45, pp. 205–217, 1953.

[63] Wanger, Leonard. "The effect of shadow quality on the perception of spatial relationships in computer generated imagery". *Computer Graphics Special Issue (1992 Symposium on Interactive 3D Graphics)*, Vol. 26, pp. 39–42, March 1992.

[64] Wanger, Leonard R., James A. Ferwerda, and Donald P. Greenberg. "Perceiving spatial relationships in computer-generatedimages". *IEEE Computer Graphics and Applications*, Vol. 12, No. 3, pp. 44–58, May 1992.