The Importance of Accurate Head Registration for Fine Motor Performance in VR

by

David William Sprague

B.Sc., Queen's University, 1998
B.Sc., Queen's University, 2001

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

Master of Science

in

The Faculty of Graduate Studies

(Computer Science)

The University Of British Columbia

February 2006

© David William Sprague 2006

Abstract

Many virtual reality researchers consider exact head registration and an exact multisensory alignment between real world and virtual objects to be a critical factor for effective motor performance in a virtual environment. Calibration procedures for head-mounted displays, however, can be error prone, time consuming and sometimes impractical to perform. To better understand the relationship between head registration and fine motor performance, we conducted a series of reciprocal tapping tasks under four conditions: real world tapping, virtual reality with correct head registration, virtual reality with mildly perturbed head registration, and virtual reality with highly perturbed head registration. As might be expected, virtual reality performance was worse than real world performance. There was no effect of head registration perturbation on motor performance in the tapping tasks. We believe that sensorimotor adaptation enabled subjects to perform equally well in the three virtual reality conditions despite the incorrect head registration in two of the conditions. This suggests that exact head registration may not be as critically important as previously thought, and that extensive per-user calibration procedures may not be necessary for some virtual reality tasks.

Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank Dr. Kellogg Booth for funding me and for being my supervisor and advisor over the past two and a half years. His suggestions improved my research either directly, or by forcing me to defend my plans and ideas. Our dissenting opinion about VR also led to this thesis topic, which it has been my privilege to examine.

I would like to thank Dr. Michiel van de Panne for being my second reader and for advice he provided. I would also like to thank my committee members Dr. Jim Enns and Dr. Brian Fisher for their advice. They saved me time and always brought a fresh perspective on my research.

There is no way that I could have finished this degree without the moral and financial support of my parents, Bill and Geraldine Sprague. They may not have always understood what I was talking about, but they always gave the most logical advice possible and acted as the ultimate sounding board for my ideas. They may be half a country away but they were always with me.....even if I am "out of the will."

I would like to thank all the people at UBC that made this degree manageable, either by giving me advice, helping with my experiments, forcing me to relax and socialize, or just listening to my complaints. I would particularly like to thank Dr. Barry Po (my proxy supervisor), Colin Swindells (my mentor for what graduate studies should be), and Tiberiu Popa (who made Linux tolerable). Their sage advice was invaluable and without their help I would have required another 2.5 years to graduate.

Thanks everyone. I finally did it, and I **_may_** have learned something in the process.

David Sprague

The University of British Columbia

February 27, 2006

Chapter 1

Introduction

Virtual reality is used in a variety of applications today including pilot training, interactive design, teleoperation, and medicine [22]. Virtual reality environments can permit users to familiarize themselves with a situation or environment and allow them to perform comparable physical tasks that may be difficult, dangerous, or costly to practice in the real world. Real world tasks such as airplane piloting and endoscopic surgery require extensive training. Virtual reality can supply a safe learning environment [8]. The motor skills acquired during training are affected by the perceived virtual environment. Inaccurate virtual environments can lead to errors when real world tasks are attempted. Knowing what factors affect fine motor tasks can be critical for a virtual reality system's success.

1.1 Motivation

Realism in an immersive virtual reality (VR) environment requires precise calculations and measurements to present a correct pair of virtual images. Deering's work on high resolution VR clearly demonstrates the benefits that precise measurements and attention to calibration detail have on motor ability in VR [12]. The impact that each VR factor has on the user, however, is not immediately apparent. Parameters such as field of view, image quality, system lag, and depth cues may have varying importance for a particular task.

Ensuring that each VR parameter is correct takes time, so cost/benefit trade-offs arise, especially when some calibrations are difficult or impractical to perform. Some parameters may even be mutually exclusive. For example, improving image quality (towards photo-realism) can increase system lag. Better image quality and low lag

have both been shown to improve VR usability [14, 54]. Identifying the importance of a VR parameter provides subsequent researchers with guidelines for improving VR system development times and costs, and a better understanding of design trade-offs. Correct mathematical models of the virtual environment are not the only aspects to consider; human perception factors invariably impact perceived VR quality.

1.2 Research Overview

This thesis focuses on the importance of head registration for motor performance in a virtual environment. Specifically, our research addresses how incorrect head registration affects movement speed and accuracy in a Fitts-like reciprocal tapping task performed in VR. We calculated the optimal head registration for a subject using a pinhole headband (discussed in Chapter 3). We then asked subjects to perform a series of virtual reality tapping tasks in three different registration quality conditions: optimal head registration, moderately perturbed head registration, and greatly perturbed head registration (see Figure 1.1). A real world tapping condition was also tested.
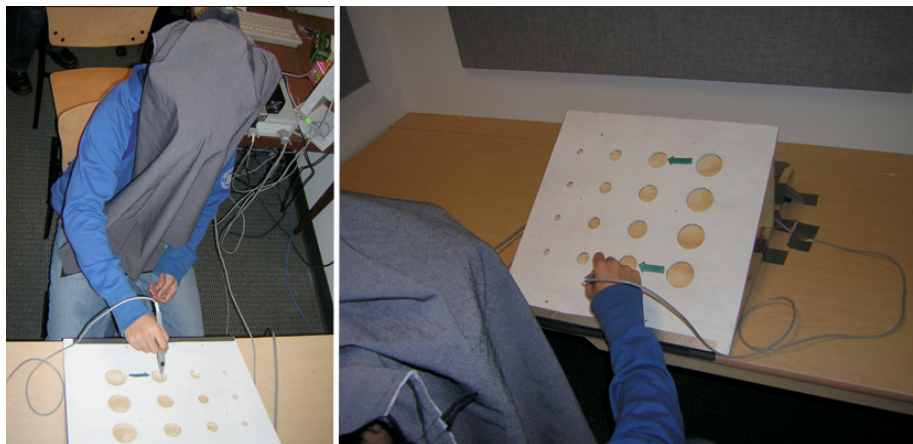
Figure 1.1: Two perspectives of the reciprocal tapping task being performed during an immersive VR condition.

A block of tapping trials required subjects to move a magnetically tracked stylus

back and forth between two target holes on a wooden board (tapping board) until told to stop. Target sizes and the distance between targets varied between trial blocks. Twenty-four subjects each performed a minimum of 1200 tapping trials (4 conditions $\times$ 12 blocks $\times$ 25 taps per block) and we analyzed tapping movement times and accuracy to examine how head registration affects motor skill.

1.3 Terminology

We will begin by defining a series of frequently used terms. Some terms, such as calibration, can have multiple definitions. The preferred definition will be identified in all such cases and used exclusively in the remainder of the thesis.

Virtual Reality (VR): Computer generated sensory information intended to be perceived as part or all of a user's environment. Virtual reality occludes some part of the real world environment from the user with computer generated elements. VR is unlike augmented reality which seeks to enhance or alter the real world. Virtual reality environments are primarily visual experiences frequently perceived using shutter glasses & desktop monitors (fish tank VR), VR caves, or head-mounted displays, but sound and touch can also be part of the virtual experience [49].

Virtual Environment (VE): The environment simulated by a virtual reality system. An immersive virtual environment is a sub-class of VEs where visual perception of the real world is obscured almost completely by the virtual world [39].

Stereopsis: The perception of depth based on the each eye receiving a slightly different view of the world (retinal disparity). Stereopsis is discussed in Chapter 2.

Head-Mounted Display (HMD): An HMD is any device worn on the head that has one or two monitors or displays positioned in front of the wearer's eyes. Subjects with binocular vision perceive stereopsis depth cues when each eye sees a different image of the VE.

Head-coupled Perspective (HCP): Tracking the position and orientation of a subject's head to present VR information appropriate to the subject's perspective and the

position of the view screen. Head movements result in the same perspective shift in VR as a person would experience in the real world. This is also called *head tracking (HT)* in this thesis.

First Nodal Point of the Eye: The location in the lens of the eye where light converges to a single point. This is discussed in Chapter 3.

View Screen: The real world monitor or display inside an HMD optical module (eye piece) on which the virtual environment is displayed. The view screens in this experiment's HMD are each 10.1 mm high and 14.1 mm wide.

View Frustum: The bounding area of the virtual environment that is visible to a virtual camera. A virtual camera has a 4 sided viewing pyramid expanding out (to infinity) from the center of projection in the camera's gaze direction. The view frustum is defined by two parallel planes intersecting the viewing pyramid. The six sides of the frustum each define a clipping plane determining what objects are displayed in the scene (see Figure 1.2). For more information on the synthetic-camera model, see any textbook on computer graphics, such as the standard textbooks by Angel[2] or Foley et al. [20].

Image Plane: The 2D representation of the 3D world. This region is defined by the intersection of a plane and the viewing pyramid. OpenGL always makes the image plane equal to the front clipping plane of the view frustum. Therefore, all light from a point in the frustum to the virtual camera must intersect the image plane. The image plane is a 2D projection of the objects in the frustum. The image plane should correspond in position and shape with a real world view screen if our head registration measurements are correct.

Registration: The procedure of aligning and synchronizing sensory elements, either within one sensory modality or across several sensory modalities [7]. In our experiment we register the position of a real world tapping board with its corresponding VR representation so that the VR image of the board corresponds with real world tactile, and auditory information perceived by the subject.

Head Registration (HR): The measurement of a subject's head and eye positions in order to present a realistic head-coupled perspective in VR. HR for this research requires knowing the vector between a head tracking sensor and each eye, and the vector from each eye to the corresponding view screen presenting the VR images.

Ideal head registration requires that the cameras in the virtual world exactly correspond with the subject's eyes in the real world, and the size, dimensions, and positions of the image planes in VR correspond with the HMD's view screens in the real world (see Figure 1.2).
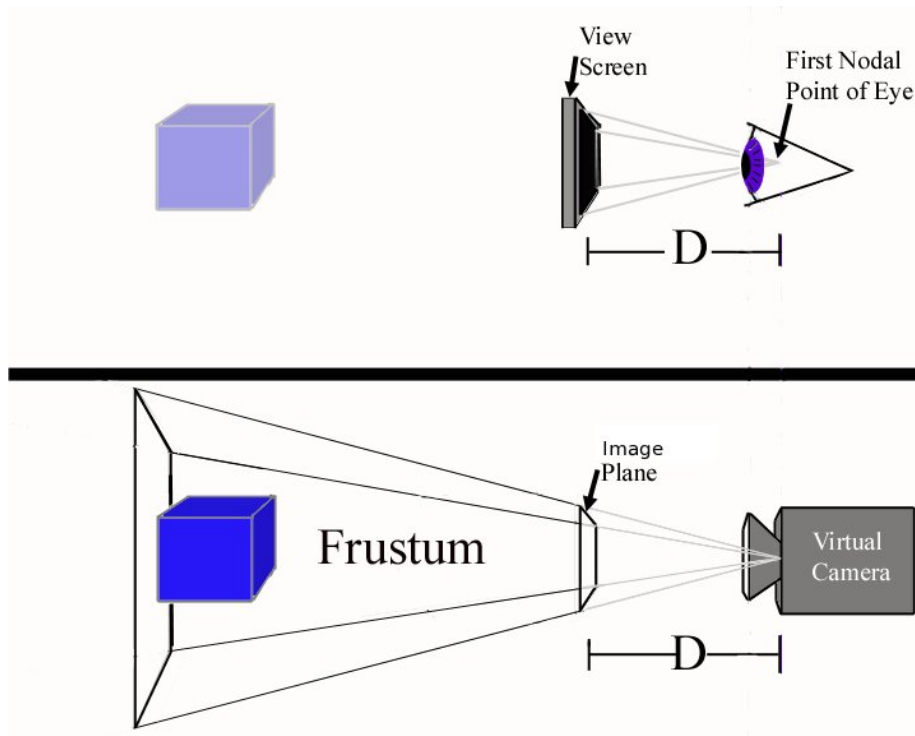
Figure 1.2: Head registration and the synthetic-camera model. This diagram shows the correspondence between real world eye and view screen positions and the camera and view frustum in the virtual world for ideal HR.

Optimal Head Registration (OR): The most accurate HR we can accomplish using the calibration methodologies we present in Chapter 3.

Calibration: To precisely adjust data or objects for a particular function. Head registration values are calibrated to be optimal using the method described in Chapter 3. Sensor calibration refers to the process of standardizing sensor information so that systematic errors or deviations in data can be determined and a proper correction factor applied.

Proprioception: Perception of a body part's position relative to other parts of the body without the use of external stimuli. Proprioception (often used interchangeably with the term *kinesthesia*) does not provide information about the outside world but provides feedback based on the position of the body. Proprioception is critical for speed and precision during motor tasks and is the fundamental sense required for body memory [32].

Passive Haptic Feedback (PHF): Passive haptic feedback in VR is proprioceptive and tactile information available to the user as a result of interactions with a real world object or input device. PHF means the virtual object is perceived to have tactile or physical characteristics such as texture, weight, temperature and hardness. The tapping board used in this experiment provides PHF as subjects both see a virtual object and feel when the stylus contacts the board (see Figure 1.1).

System Lag: The time between when a user's action occurs and the time a result is presented to the user.

Update Rate: The frequency that the virtual scene is changed or updated. We will also refer to update rate as the frame rate. We calculated the update rate by taking the average time interval between virtual scene drawings/renderings calls over a one second interval (equivalent to the number of render calls per second). The average update rate during the experiment was 40 frames per second.

Refresh Rate: The frequency that the view screen image is updated. The HMD screens have a refresh rate of 60Hz. For our system, refresh rate is independent of the update rate.

Display Field of View (DFOV): The subtended angle from the subject's eye to the edges of the display/view screen [11]. Vertical and horizontal DFOVs are independent. Field of view (FOV) will refer to DFOV.

Geometric Field of View (GFOV): The subtended angle from the virtual camera to the edges of the view frustum [11]. This relates to how much of a virtual scene can be observed from a given VR camera perspective. Vertical and horizontal GFOVs are independent. If head registration is ideal, horizontal and vertical GFOVs should be equivalent to the corresponding DFOVs. Perturbing HR during the experiment increased the GFOVs while DFOV values were constant.

1.4 Thesis Outline

It is important to investigate how incorrect registration affects immersive VR use because exact HR can be difficult to perform, error prone, and problematic to maintain. Head-mounted displays can shift during head movements and the focal center of the human eye can be difficult to precisely locate [12]. Despite this difficulty, many VR researchers claim that HR registration is critical for an effective VR system [7, 15, 60]. In order to examine the effects of HR quality on performance we need a registration methodology, an appropriate environment in which to test subjects, and a theory on which to base our research.

This thesis is separated into five sections. Chapter 2 outlines the related psychology, kinesiology, and computer science literature. Chapter 3 briefly describes how head-coupled immersive VR works and provides a novel methodology for fast, accurate and precise HR. Chapter 4 describes an experiment developed to examine the effects of HR quality on fine motor performance. Finally, Chapter 5 presents future work to be done based on our research and conclusions we made based on our results.

Chapter 2

Related Literature

Previous research seems to make contradictory claims about the importance of HR quality for VR use. Clarifying this ambiguity is the foundation of our research. A brief overview of human vision, depth perception, and techniques for simulating depth with an HMD is provided to clarify the role of head registration in VR systems. Previous research investigating VR registration and calibration will be discussed in the context of head-coupled HMD VR. Registration problems using an HMD will also be identified. Next, we will discuss known human perceptual factors that affect performance in VR (motor, perceptual, and presence) and how such factors can be independent of the quality of mathematical modeling underlying the VE. The passive haptic feedback literature is discussed in terms of registration and the need for exact HR. The sensori-motor adaptation literature contradicts the need for ideal HR, so this field of research is examined. Finally, the fine motor (reciprocal tapping) task used in this experiment is described. Our experimental hypotheses are based on previous findings in the literature.

2.1   Human Vision and Immersive Virtual Reality

Immersive head-coupled VEs rely heavily on precise mathematical models and a thorough understanding of human perception. If a person's eye positions and orientations can be tracked at all times, and if perspectively correct VR images can be presented to each eye independently (obstructing the real world) then a virtual environment can be presented, almost indistinguishable from the real world.

In our research, we are primarily concerned with motor performance in the VE. If motor performance metrics in two experimental conditions are the same, we claim that

the differences between conditions do not affect performance. The VE only needs to provide sufficient visual cues so motor performance in VR is equivalent to the same performance in the real world. Even with this reduced criteria, we found no studies where VR motor performance was equivalent to real world performance [8, 22, 35, 52, 54, 56]. This suggests that VE depth cues needed for visually guided 3D movements may be incorrect or insufficient compared to real world images.

2.1.1 Binocular and Monocular Depth Cues

Effective 3D virtual environment interactions require robust depth cues. Monocular depth cues such as linear perspective, relative size, familiar size, lighting and shadow, texture, interposition, and image clarity are available via 3D graphics on conventional computer monitors. Head tracked HMDs, however, provide two features not typically available: binocular depth cues (stereopsis and convergence), and head-coupled perspective (motion parallax depth cues).

Stereopsis is perception of depth based on retinal disparity. Each eye sees visual cues slightly differently (retinal disparity) and this disparate information is processed by the occipital lobe to provide depth information about the world. Stereoscopic displays provide slightly different images to each eye to simulate binocular depth cues and the brain interprets depth from these two 2D images.

Stereopsis depth cues are ineffective at close distances (less than 3 cm) or at great distances [44]. Virtual objects too close to the virtual cameras can induce physiological diplopia (double vision) and eye strain in subjects [55]. Patterson & Martin [38] state that qualitative stereopsis needs a minimum image disparity of 2 degrees, meaning a subject with a 6 cm interocular distance can reliably establish depth information from stereopsis as far as a depth of 172 cm. All visible objects within the virtual world were within a 100 centimeter radius of the subject.

Ocular convergence cues derive from the level of muscular tension in each eye as they point at a single object in focus. Ocular convergence has an effective distance of less than 1 meter so we expect stereopsis to provide the predominate binocular cues during our study[31]. For a more detailed description of monocular and binocular depth

cues, see Schwartz [44], Patterson & Martin [38], or Ware & Balakrishnan [54].

Not all visual depth cues are available in virtual reality systems. Accommodation is a monocular cue that derives depth information from the focal distance of each eye to a target object [44]. Accommodation cues rarely provide the same depth information as other visual cues in HMD VR, because the planar surface of a view screen displaying the VE is a constant distance from the user. Many graphical elements provided in VR, such as image resolution and DFOV, are inferior to the same visual cues experienced in the real world [31]. Thus, an ideal image in VR with truly realistic depth cues is not possible with current hardware.

Ware and Balakrishnan [54] found that precision tapping movements using fish tank VR were significantly slower toward/away from the subject rather than side to side. This suggests that depth cues are important to motor performance in VR and incorrect/insufficient depth information may hinder a subject's performance. Eggleston, Janson, and Aldrich [14] found that perceptual abilities in VR only equaled real world perception when both viewing conditions were impoverished. Thus we expect fine motor task scores in the real world to be significantly better than scores doing the same task in VR. Disparity between visual cues can lead to eyestrain, headaches, dizziness and cybersickness [18]. For this reason, all VR sessions throughout the experiment lasted less than 20 minutes each.

2.1.2   Head Registration and Simulating Depth in VR

Three dimensional computer graphics software systems such as OpenGL use a synthetic-camera model for scene rendering. A virtual camera has a location in VE coordinates while a view frustum corresponding to this camera defines what VE objects are seen by the camera. Any light ray from a visible object to the center of projection (camera) must intersect the image plane. Taking all light information at the image plane provides a 2D image of the virtual scene that can be displayed on a monitor via a view window. If the view window is the same size and location as a monitor screen (presuming the monitor is perfectly flat) then the monitor displays the VE from the perspective of the camera's theoretical position in the real world. Matching the camera's position with a

subject's tracked eye position (or more specifically the first nodal point or focal point of the eye) results in a non-distorted portal from the real world to the virtual environment [12]. Subjects can look into the virtual world the same way as they would look at the real world and linear perspective, relative size, familiar size, lighting and shadow, texture, interposition, and image clarity cues will all be correct (provided these depth cues are provided). A VR system with a second camera, view frustum, and display screen can present a portal to a person's other eye if it too is being tracked. Having two portals or windows into the VE automatically provides stereopsis and convergence information to the user. This second camera/view frustum/monitor set can be provided using a single display by rapidly covering/uncovering a person's eyes (so only one eye sees the display at a time) while alternating the images displayed. This technique is used in many HMD designs and for fish tank VR [53]. The HMD used during our study provides independent displays for each eye, and thus we did not need to be concerned about alternating the VR image. A far more in-depth analysis of 3D graphics and the synthetic-camera model can be read in books by Angel[2] and Foley et al. [20]. Chapter 3 discusses implementing head-coupled VR in more detail.

It is crucial to note that if a display's dimensions, position, or orientation are incorrect, or if a person's eyes are not correctly tracked the resultant image seen by that person will not be as realistic as possible. Deering [12] confirms this claim and clearly demonstrates VR motor benefits of exact calibration and registration. If we consider the human observer as the final step in a series of mathematical transformations to generate a virtual environment, then any inaccuracy in display or eye tracking will lead to a less realistic VE, and a presumed decrease in motor performance due to degraded visual conditions. Some authors such as Yokokohji et al.[60], advocate the need for extremely accurate head registration for this reason.

2.2 Head and Eye Measurement Techniques

Optimal HR using an HMD requires tracking the precise location of the first nodal point for each of the subject's eyes and the position of the HMD's view screens. Surprisingly, there are few published methodologies for HR in immersive virtual reality

using an HMD. Stevenson's [47] fish tank VR registration technique uses the screw holes in two magnetic tracking sensors and a third magnetic tracker attached to the head to extrapolate line of sight information. Subjects aligned the screw holes of two sensors along their eye's gaze direction so a far away object could be observed. Taking sensor position information for several sensor alignments allowed the software to determine the first nodal point of the subject's eye relative to the head sensor. Tracker signal noise requires best fit approximations to be used, and this technique was not sufficiently precise for the registration we required. There are numerous HR techniques for augmented reality (AR) displays and VR tables [6, 7, 48], but they rely on subjects aligning real world and virtual objects. The optical modules of the HMD occlude most of the real world so these automated AR registration techniques cannot be used. We, instead, made precise and accurate head measurements using a methodology discussed in Chapter 3.

2.2.1 Data Cleaning

Head tracking using a magnetic tracking device can often be error prone and can require data cleaning. Kalman filters are often used with magnetic tracker systems to reduce sensor jitter and remove white sensor noise by providing adaptive position and orientation data predictions. A Kalman filter is a set of mathematical equations used to estimate the state of a process based on the observation history [57]. The filter works in two stages: a prediction stage and a correction stage. Predictions are made based on the weighted fitness of previous predictions compared against the data observed. The predicted value is on the best polynomial fit for the data observed. The a priori estimate model is then modified based on the difference between the current data and the estimate made. This data filtering can add to system lag (especially since Kalman filters are recursively calculated) and filtering may not substantially improve VR system quality if sensor measurement errors are minimal or systematic in nature.

Systematic errors in sensor data have been calibrated in other VR systems using either piecewise linear interpolation or higher polynomial interpolation of sensor data [21, 27]. Magnetic trackers frequently require interpolation to correct their data if a

large metal object is nearby or the field of movement for the sensors is large.

No data cleaning or calibration was performed for the experiment described in Chapter 4. Instead, the subject's head position from the transmitter was typically under 1 meter, most/all metal was removed from around the tapping board, and registration testing indicated minimal random or systematic signal errors. Piecewise linear interpolation was used at an earlier stage in our research (when a large metal object was present) but the adjusted data was not sufficiently precise for our study. Extensions and continuations to this research will probably require data cleaning to remove our constraints to the experimental design.

2.3 Human Perception Factors & VR Effectiveness

The Psychology literature suggests that the human observer is not simply the end point of a VR system. Many neurological and psychological factors affect the sense of presence (the feeling of "being there" in a virtual environment), perception, and motor abilities in a virtual environment. Boritz and Booth [9] found no relationship between head-coupled stereo and performance, but stereoscopic pointing performance was significantly better than monoscopic pointing performance in fish tank VR.

Arthur, Booth and Ware [5] found that head-coupled VR and stereo depth cues greatly improved performance during a graph tracing task using a fish tank VR system. Head tracking was shown to significantly improve performance when stereo vision cues where available but head coupling alone did not show a significant improvement in task performance compared to the baseline (no head tracking, no stereo vision). Stereo vision alone significantly reduced performance error rates. For Arthur et al.'s [5] motor task, head-coupled stereo aided a subject's perception of 3D objects and the combination of the two factors was observed to be more effective than either factor in isolation.

Research by Arsenault and Ware [4] further supports the importance of eye-coupled perspective and stereo in a fish tank VE with passive haptic feedback. Stereopsis was again determined to be more important than eye-coupled perspective for their Fitts-like tapping task. Binocular vision cues and head coupled perspectives of the VE are

used in all VR conditions during our experiment to make the virtual environment as immersive as possible and hopefully more susceptible to registration errors.

2.3.1 Field of View

A VR system's FOV has been shown to affect motor abilities in VR. Wells & Venurino [58] demonstrated that larger DFOV values for helmet mounted displays allowed subjects experiencing a military simulation to hit targets faster and reduce the time they were threatened. Larger DFOVs required less head movement and subjects moved their heads faster when they did move. Czerwinski et al.[11] reported that sex differences in 3D path navigation abilities disappeared when the GFOV and DFOV were increased. The impact GFOV has on our study will be discussed in Chapter 4. Distance perception was also shown to degrade when an HMD DFOV was less than 60 degrees [14]. This is important to note since the current HMD's DFOV is $40°$ horizontal ($60°$ binocular) and $30°$ vertical compared to the average human non-foveal FOV of approximately $170°$ horizontal and $126°$ vertical [13].

2.3.2 System Lag and Frame Rate

System lag has been linked to reduced perception, motor skill, and immersion scores when using a VR system. Ware and Balakrishnan [54] found an exponential relationship between system lag and graph tracing speed. The r^2 fit between lag and tracing speed was also better than the correlation between frame rate and tracing speed, suggesting the overall lag between a user's action and the result is more problematic to motor functioning than the update rate of the image seen. Update rates of 10 Hz or less were highly noticeable by subjects and a 70ms (14 Hz) lag was used as the minimum lag they could achieve for their study. Mania et al. [36] found that subjects had a 15ms just noticeable difference in perceived system lag, which was independent of the number of objects in the VE scene. Hence, perception of system lag does not seem dependent on what is being displayed. Mania et al. also reported that subjects only reliably noticed lags greater 43ms. Our experiment's VR system is intentionally designed to keep system lag almost constant and to a minimum, averaging approximately

60ms.

2.3.3 Other Factors

Other factors shown to affect VR performances (perception, motor skills, or immersion) include luminance and colour [14], cursor to object size correspondence [50], the use of a background grid [51], and the occurrence of longitudinal parallel lines [37]. All of this evidence confirms the need to consider human perception when designing VR systems.

2.4 Passive Haptic Feedback in VR

To measure the effects of errors in HR on VR motor performance, we used passive haptic feedback (PHF), where correct registration is deemed crucial in the previous literature [3, 32, 60]. Furthermore, PHF has been shown to enhance a user's sense of presence and improve motor task abilities, which should make the VR and real world conditions in our experiment more comparable [8, 33, 51]. A study by Arsenault and Ware [3] examining the effects of head tracking and touch feedback on a tapping task found a 9% improvement in tapping time when head tracking was present, and a 14% time improvement when passive haptic feedback (using a Phantom force-feedback device) was used. This suggests that having PHF may even be more important than experiencing a head-coupled perspective.

Lindman et al. [33] discuss mobile and stationary object registration and tracking for real world PHF devices, including how they registered flat planar surfaces using a magnetic tracking device. We used similar registration techniques for our experiment.

Multi-sensory information such as PHF requires object registration to ensure sensory correspondence (when you see a tap, you feel a tap). Some researchers suggest a need for a *"What You See Is What You Feel"* or *WYSIWYF* approach to VR haptic design. Yokokohji et al. [60] suggest that an exact correspondence between the various sensory modalities is required. If an object being manipulated is seen directly in front of the subject, proprioception should indicate that the object is in exactly that

location. WYSIWYF requires HR to be as precise as possible to ensure that the same haptic/visual correspondence achieved in the real world is provided in VR. Rolland [42] also supports the need for exact sensory correspondence. Wang and MacKenzie's [50] work on haptic/visual size correspondence suggest that real (tactile) and virtual (visual) object correspondence affects object manipulation speed. Subjects were more efficient at an object alignment task when the real world controller matched the size and shape of the virtual object (cursor) being manipulated. In a VR object alignment task, virtual objects co-located with a matching real world object performed significantly better than an alignment task without real and virtual objects being co-located [56]. Finally, a subject's egocentric (head) reference frame also affects object alignment times in VR [52]. Ware et al. found that the degree of axis misalignment between the subject's frame of reference and a prop object's frame of reference was associated with a quadratic increase in movement time. These findings suggest that the WYSIWYF approach to VR design may be justified.

2.5 Sensorimotor Plasticity/Adaptation

Experiments using prism glasses have shown that the correspondence between motor functions and the senses is dynamic [24]. Subjects can adapt to a variety of alterations to their vision so that after repeated self-produced movements, a subject can function normally in the altered state [25, 41]. This adjustment to systematic sensory alterations is known as *sensorimotor adaptation* or *sensorimotor plasticity*. Multi-modal sensory information is dynamically adjusted so the perturbed sense is eventually recalibrated to the other senses and motor functioning. When the sensory perturbation is removed, temporary movement errors in the opposite direction of the alteration are found. This is known as a *negative aftereffect* [41]. Negative aftereffects are traditionally used to indicate that sensorimotor adaptation has occurred.

Sensorimotor adaptation provides clear benefits to animals, allowing them adapt to growth, injuries, and systematic alterations to perception (such as wearing glasses for humans). Rossetti et al. [43] suggest that sensorimotor adaptation occurs as two systems: a fast adapting short-term adaptation and a long-term multiple state system

for repeated regular perceptual changes. For VR systems, only short-term adaptations are relevant.

Research suggests that sensorimotor plasticity occurs in virtual and augmented reality environments, not just in the real world. Ellis et al.'s[17] immersive VR visual tracking experiment with control display misalignment clearly demonstrated a training or sensorimotor adaptation effect as relative error rates in misaligned conditions approached the perfect alignment scores over time. Adaptation has been seen in video-based augmented reality (AR). Video-based AR systems use video cameras to display real world information to the user, and hence human perspective of the real world is based on the camera positions. Subjects showed performance improvements in their pointing accuracy with prolonged exposure to the system, and a negative aftereffect was demonstrated when subjects removed the AR display suggesting that subjects adapted to the altered vantage point of the real world[42]. Groen and Werkhoven [22] demonstrated that sensorimotor adaptation occurs in virtual reality as well, and they found real world negative aftereffects when they displaced a subject's hand representation by a small amount in VR. They found no significant performance differences between the misaligned and the correctly aligned VR hand positions.

Our research investigates how readily sensorimotor adaptation occurs with passive haptic feedback, following the approach of Groen and Werkhoven. We examined a range of systematic head registration errors to study how perturbation magnitude affects motor performance.

2.6 Fitts's Law and Fine Motor Tasks

Many previous experiments examining motor abilities in virtual reality have used a Fitts-like tapping task. Such task behaviors are well defined, well documented, and well understood in many research communities [16, 23, 34, 35, 61]. These precision motor ability tests come in two main formats: a discrete tapping task, and a reciprocal tapping task [19, 30]. Both tests require subjects to point to or tap target objects as quickly and as accurately as possible. The discrete tapping task requires subjects to tap a target at an unknown position as soon as the target becomes visible, while reciprocal

tapping requires subjects to tap back and forth between two targets until indicated to stop (see Figure 2.1 for the original experiment diagram by Fitts). If the subject taps the target, the trial is successful. Tapping anything other than the target is considered an error.
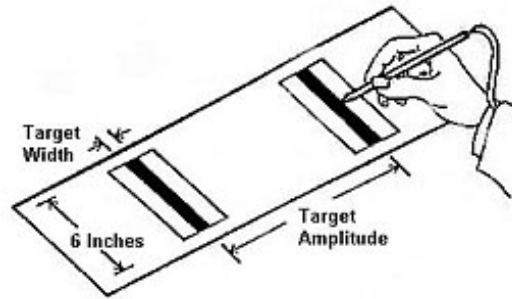
Figure 2.1: The original reciprocal tapping task design by Fitts [19]. Subjects tap back and forth between the two known targets until told to stop. When the stylus contacts a target plate, an electric circuit is completed indicating that a tap occurred.

By constraining the permitted error rate (to approximately 4%), a single time-based measure of task difficulty, or index of difficulty, can be determined. Fitts's experimental findings further identified that the logarithm of the quotient of distance traveled (A) divided by the width of the target (W) is linearly related to measured movement time (MT) [19, 30]. Variables C_1 and C_2 are empirically determined constants and the logarithmic term is referred to as the index of difficulty (ID). Thus there is a linear relationship between the ID and the tapping time. For this experiment we will be using the formulation proposed by S. MacKenzie (Equation 2.1) and a reciprocal tapping task described in Chapter 4. Other formulations such as the Welford formulation and Fitts's formulation are reviewed by Soukoreff & S. MacKenzie [46] and Mandryk [35].

$$MT = C_1 + C_2 * log_2(\frac{A}{W} + 1.0) \qquad (2.1)$$

Fitts-like tapping tasks are frequently used to compare between different experimental conditions [34]. Correlation coefficients, y-intercepts and linear regression

slopes, and information processing throughput values (discussed in Chapter 4) are used as cross-condition metrics. Conditions with faster movement times and greater information processing rates are considered better for target acquisition and motor performance in general. The well defined nature of the task enables detailed experimental hypotheses, verification of experimental results (target acquisition values should conform to Fitts's law), and detection of potential flaws in an experiment's design. For this reason, Fitts's law and Fitts-like tapping tasks are a standard method used in Human-Computer Interaction research to compare and evaluate different input devices, different viewing conditions, and different interaction styles [34].

2.7 Summary

Based on previous research, it is not obvious how important head registration accuracy is in an immersive head-coupled VE with passive haptic feedback. Some researchers claim that sensory information must correspond precisely in order for a VR system to be effective. Sensorimotor adaptation research, however, casts doubts on this assertion. Previous research also demonstrates that human perception is an important factor in VR use, and modelling the virtual world more accurately does not necessarily translate into improved performance. An experiment testing the effects of HR accuracy on fine motor performance is a first step in understanding how HR affects a virtual experience.

Chapter 3

Head Registration Methodology

To test the effects of accurate head registration on motor performance, we first need a methodology to quickly find the optimal HR for any given subject. To determine where a subject's eyes and the view screens are located in the virtual world, we need to know how the eyes and screen are related to a tracked object position (a head tracking sensor). Immersive HMDs prevent most of the real world from being observed so we are unable to align the virtual world with real world visual cues to infer head tracker to eye vectors (augmented reality and non-immersive VR can use this technique). Instead we designed a unique method to measure tracker to eye and tracker to view screen vectors.

3.1 Head-Coupled Perspective

Determining the appropriate head-coupled perspective requires tracking the position and orientation of a real world object (head tracker) and inferring VR perspective based on the rigid relationship between objects.

3.1.1 Polhemus Fastrak

Optimal HR in a VE requires an accurate and precise way of tracking a subject's head position and exact measurements to infer eye position and orientation from tracker data. For our experiment, we used a six degree of freedom magnetic tracking device called a Polhemus Fastrak [40] including a receiver (or sensor), a stylus receiver, and a transmitter. The tracked points of the receivers were clearly marked (using masking tape) according to the Polhemus's specifications. Fastrak position and orientation records

are collected at a maximum of 120Hz divided by the number of sensors used. Thus, if two sensors are being tracked, data is collected at 60Hz (with a 4ms sensing lag). Each packet of sensor data provides 3D position information in a right-handed coordinate frame and yaw, pitch, roll Euler angles relative to the transmitter. The transmitter was consistently aligned so that the y-axis pointed up, the positive x-axis pointed to the subject's right, and the positive z-axis pointed from the tapping board towards the subject(see Figure 3.1). The three transmitter axes are orthogonal to each other. The head tracking sensor was mounted on the back of the HMD to minimize tracking signal noise due to metal and electronic interference produced by the optical modules.
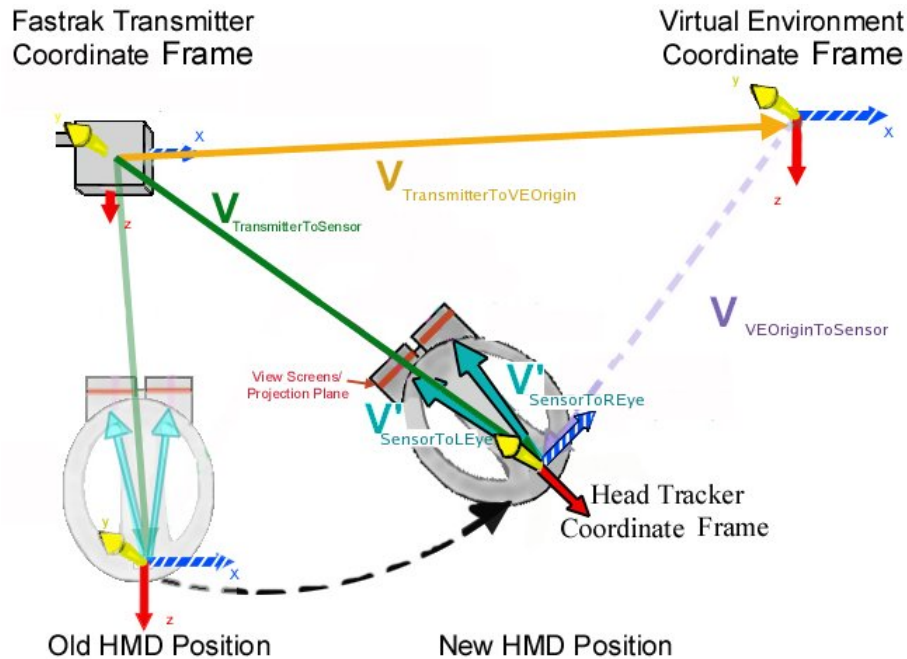

Figure 3.1: The affine transformations required to calculate the eye and view screen positions in the VR coordinate system.

3.2 Calculating Head-Coupled Perspectives

Ideal immersive VR using an HMD requires tracking a subject's eyes and the HMD view screens to present images of the VE from the correct perspective. Tracking the individual eyes of a subject when they are wearing an HMD, however, is difficult and prohibitively expensive with current technologies. Head tracked virtual reality using an HMD is traditionally accomplished by tracking the position and orientation of the helmet and inferring a subject's eye position from these data.

For this experiment we attached a Fastrak receiver to the HMD and inferred a subject's vantage point via a series of affine transformations. Affine transformations also enable a matching between the position and orientation of the VR camera and image plane with the real world eye and view screen data (Figure 3.1). First, we transform between the Fastrak's coordinate system and the VE's reference frame. For simplicity the orientation of the VE coordinate frame was aligned with the transmitter coordinate frame. Subtracting the tracker to VE vector ($\mathbf{V_{TransmitterToVEOrigin}}$) from the head sensor's position ($\mathbf{V_{TransmitterToSensor}}$) provides the head sensor position in the VE coordinate space ($\mathbf{V_{VEOriginToSensor}}$). If the VR coordinate frame and the transmitter coordinate frame are not aligned, then a 4×4 matrix can convert a vector in the tracker coordinate frame to a vector in the VE frame of reference.

Our head registration method provides four main vectors in the head sensor's coordinate system: $\mathbf{V_{SensorToLEye}}$, $\mathbf{V_{SensorToREye}}$, $\mathbf{V_{SensorToLScreen}}$, and $\mathbf{V_{SensorToRScreen}}$. The subject's eye positions in the VE coordinate frame ($\mathbf{V'_{SensorToLEye}}$ and $\mathbf{V'_{SensorToREye}}$) are calculated by rotating the sensor to eye vectors using the tracker's current orientation information. Adding vector $\mathbf{V_{VEOriginToSensor}}$ to these vectors give us the eye positions in the virtual world ($\mathbf{V_{VEOriginToLEye}}$ and $\mathbf{V_{VEOriginToREye}}$). This is equivalent to rotating the eye positions around the head tracker. The rectangular view screens are also rotated around the head sensor by the tracker's Euler angles. These transformations can also be accomplished by multiplying $\mathbf{V_{SensorToLEye}}$, $\mathbf{V_{SensorToREye}}$, $\mathbf{V_{SensorToLScreen}}$, and $\mathbf{V_{SensorToRScreen}}$ by a single 4×4 matrix that translates and rotates the vectors. Mathematical formulations for our head tracking method are reported in Appendix A.1 and program code can be supplied by the author upon request.

3.3 HMD Measurements and Markings

A series of precise 3D component measurements were taken prior to the experiment to determine the vectors from the head tracker to key positions on the HMD. All HMD positions are assumed to be rigidly connected to each other. All HMD component measurements/vectors are aligned with the same right hand coordinate frame. The x-axis of the HMD is perpendicular to the plane separating the HMD optical modules, the y-axis is perpendicular to the base of the head tracking sensor and the bottom of the HMD headband, and the z-axis is orthogonal to the first two orthogonal axes (see Figure 3.2). The optic modules were opened to measure the size and position of each view screen (see Figure 3.4). The screen locations were then marked on the outside of each eyepiece using masking tape to simplify adjustments and measurements (see Figure 3.3).
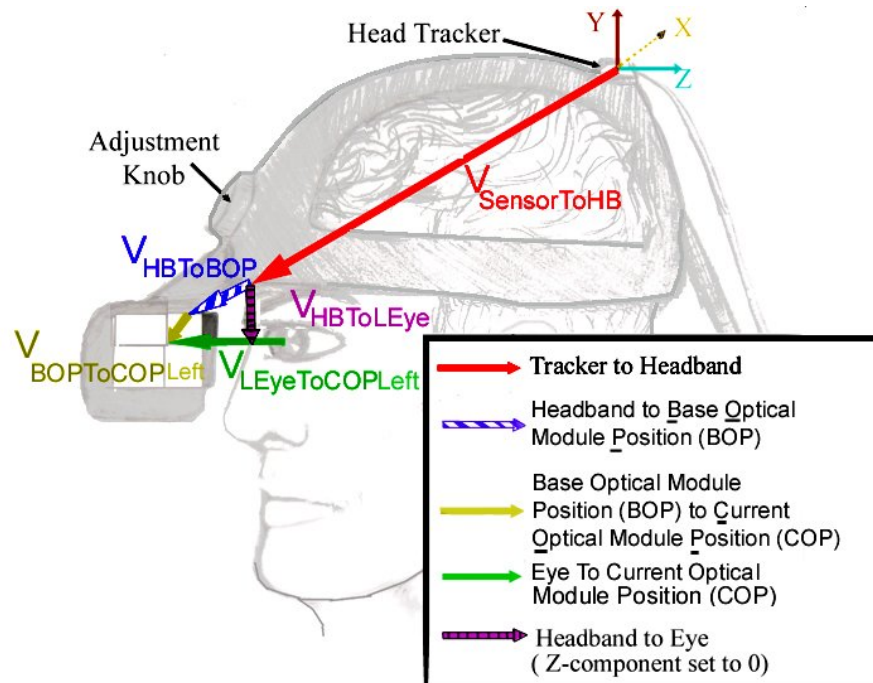
Figure 3.2: Component head measurements required to calculate the vector from the head tracking sensor to each eye and from the sensor to the view screens.

Figure 3.3: The Kaiser ProView XL50 head mounted display. The vertical line on the tape attached to the left optical module indicates the position of the z-position of view screen used.

The component distance from the attached head tracker to the HMD's headband location ($\mathbf{V_{SensorToHB}}$) was measured first. The point position of the HMD headband was defined to be at the bottom of the headband (y-position) where the subject's forehead rests (z-position) and located at the center of the two view screens (x-position) (Figure 3.2). The x-distance between the headband point position and the sensor was asserted to be zero. Second, we measured the component distance from the head tracking sensor to each "base" optical module position (*BOP*) for the HMD ($\mathbf{V_{sensorToLBOP}}$ and $\mathbf{V_{sensorToRBOP}}$). A BOP was designated as the center of a view screen when the HMD interocular distance is minimal and the optical modules are at their highest position. A
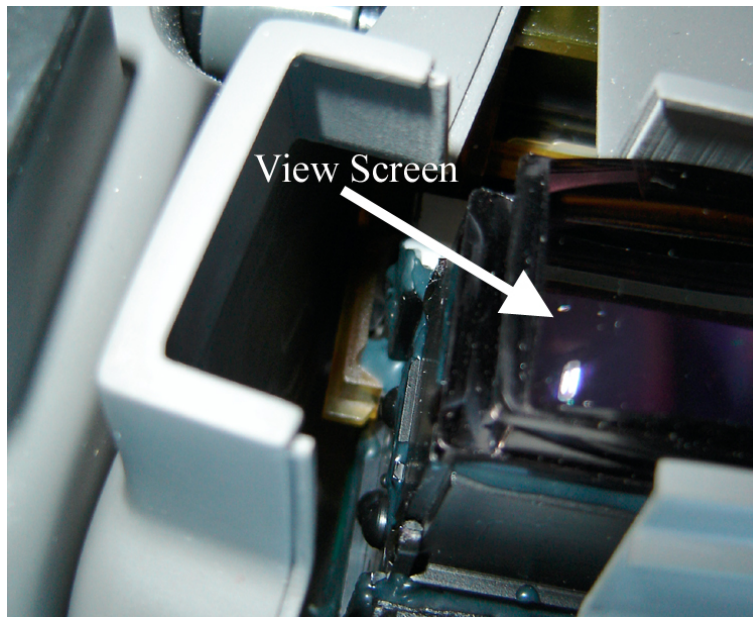
Figure 3.4: An open head-mounted display optical module displaying the LCD view screen. The view screen position was marked on the outside of the module with masking tape and the screen's width and height were recorded. Measurements were performed using calipers and a ruler and have a total precision of approximately 2mm.

discretized knob adjusts the y-axis and z-axis location of the HMD's optical modules. The y and z position of the optical modules at each wheel click were measured relative to the base optical module positions and stored in a look-up table (see Appendix B). All measurements were made using a ruler and calipers and were estimated to have a total precision of 3mm.

3.4 Pinhole Headband and Subject Measurements

The methodology presented above enables us to measure many of the constant component vectors required for head tracking, but individual subject measurements are required to make the link between eye position and tracker data. Determining these

vectors associated with individual subject differences requires an efficient and accurate measurement technique, and we decided to use a ruler and a pinhole headband (described next) to expedite this process (see Figure 3.5). First, we found the x and y components of the vector from the HMD headband point to each of the subject's eyes. We then adjusted the HMD appropriately for the subject and measured the z-distance from the view screen to the first nodal position of the subject's eyes (discussed later).
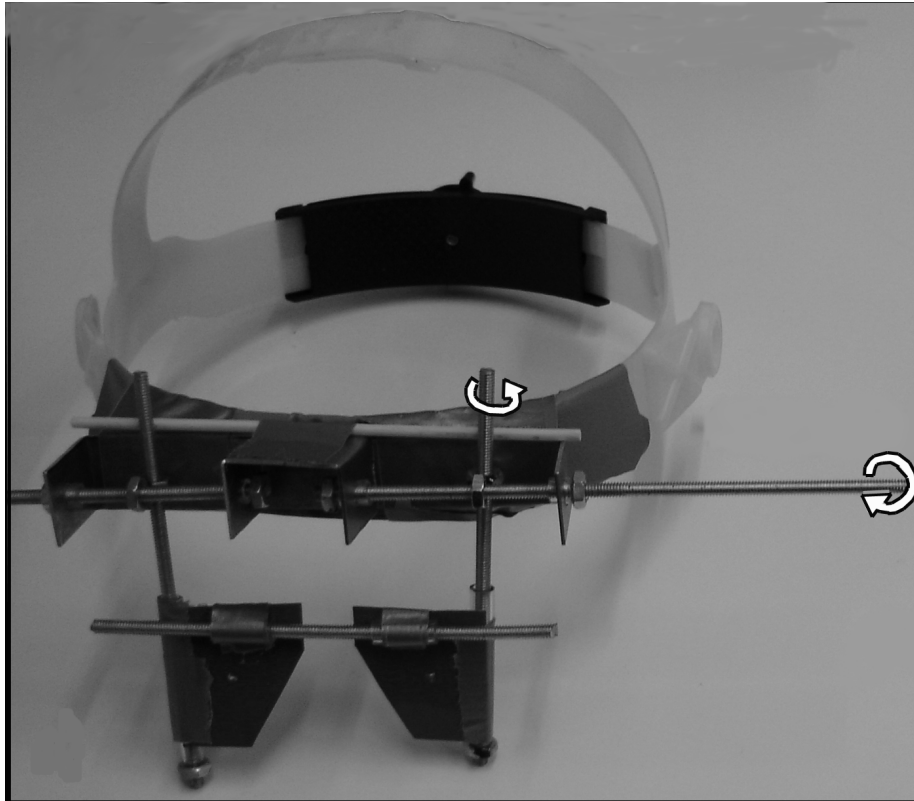
Figure 3.5: The pinhole headband (PHH) used to measure eye position and interocular distance. Arrows indicate how the adjustment bolts operate.

3.4.1 The Pinhole Headband

The pinhole headband (PHH) is used to measure the x and y components of vectors $\mathbf{V_{HBToLEye}}$ and $\mathbf{V_{HBToREye}}$. The PHH consists of pairs of orthogonal bolts connected to each other via two bonded nuts (see Figure 3.5). The horizontal bolt is loosely attached to a metal frame so that it can rotate without translating. This means that the nut threaded on this bolt moves left or right as the bolt turns. The vertical bolt indirectly connected to this nut is free to translate up and down when rotated. The metal frame is attached to a simple adjustable headband worn by the subject. An eye cover with a pinhole is attached to each vertical bolt. In this way, the covers blocking a subject's view can be precisely moved with two degrees of freedom. The PHH was adjusted until subjects were able to see a far away object through the pinholes. Measuring the distance from the headband to each pinhole and the distance between the pinholes allows us to calculate the location and separation of a subject's eyes. The estimated measurement resolution using the PHH is approximately 3mm.

3.4.2 Measuring Eye Positions

The first nodal point of a subject's eye relative to a worn HMD is extremely difficult to measure directly, so its position was calculated as a series of composite measures (see Figure 3.6). The x and y component measures for vectors $\mathbf{V_{HBToLEye}}$ and $\mathbf{V_{HBToREye}}$ can be quickly and accurately measured if the subject observes a far away object through small holes in an otherwise non-transparent surface (pinholes). Binocular vision of far away objects require a parallel eye gaze. If we assert that the pinholes are positioned directly in front of a subject's eyes and the far off object is seen with binocular depth, we claim that each pinhole shares the same x and y coordinate position with its corresponding eye. Knowing the x and y position of the pinholes relative to the HMD headband, infers the x and y component distances between the head tracker and each of the subject's eyes.

Vectors $\mathbf{V_{HBToLEye}}$ and $\mathbf{V_{HBToREye}}$ (with z=0) provide adjustment information for the HMD. Interocular distance or the distance between HMD view screens should equal the distance between pinholes. Using our look-up table, the y distance between the
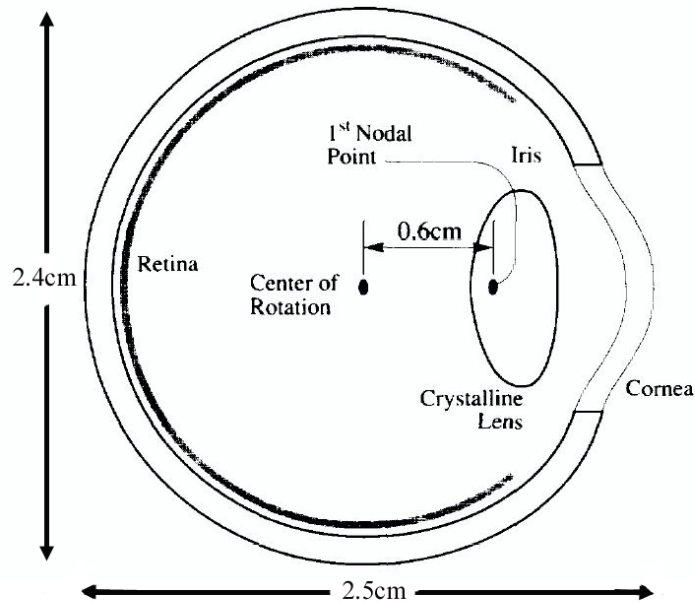
Figure 3.6: An illustrated cross section of the human eye and its first nodal point [12].

HMD headband and each pinhole informs us how much to turn the HMD's adjustment knob (Appendix B). A measurement along the z-axis from an adjusted optic module position to the subject's eye is the final measurement required (Figure 3.7). The vectors from the head tracking sensor to the subject's eyes are the differences between the head sensor to current view screen vectors (known) and the vectors from the current view screens to the subject's eyes (measured) (Figure 3.2).

3.5 Head Registration Procedure

Subjects were asked to put on the HMD and adjust it so that it felt comfortable and secure on their head. Subjects then placed a finger on their forehead just below the HMD's headband, the HMD was removed and a small piece of electrical tape was placed where their finger was located. This provided a point of reference for aligning
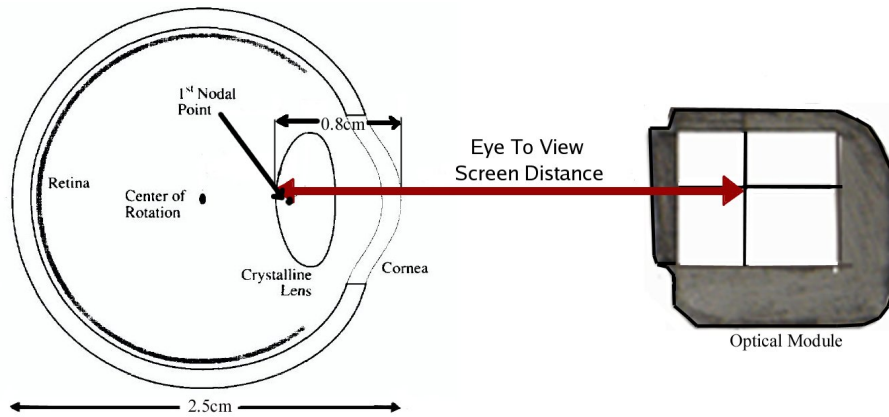
Figure 3.7: Measured distance from an HMD view screen to the estimated first nodal point of the subject's eye.

the PHH and the HMD. HMD and PHH positions were estimated to be within 2mm of each other along the z and y axes. The tape reference needs to remain on a subject's head until they were finished using the HMD.

The pinhole headband was placed on the subject's head just above the tape reference. The experimenter, with the subject's verbal feedback, adjusted each eye cover until the participant could view the light switch on the far wall of the room (approximately 3 meters away) through the two pinholes. The distance between the pinholes (interocular distance) and the distance from the pinhole to the headband (y axis distance) was taken for each eye. The HMD was then adjusted based on the average PHH y-axis distance and placed back on the subject's head. The distance from the view screen (marked on the HMD) to the first nodal point of the subject's eye was then measured using a ruler [12, 47]. The average human eye is approximately 25mm in diameter and the first nodal point is 6mm in front of the rotational center of the eye [12]. The focal point of a subject's eye was approximated to be 8mm in from the cornea and we estimate this measure to be accurate to within 3mm (see Figure 3.7).

Chapter 4

Head Registration & Motor Performance Experiment

A laboratory experiment was conducted to examine the relationship between head registration accuracy and fine motor performance in a virtual reality environment. Based on previous experimental work, we believe that HR may not be as critically important as previous VR literature suggests, and that sensorimotor adaptation enables people to compensate for incorrect HR. We devised an experiment with four different viewing conditions: real world (RW), VR with optimal HR (OR), VR with moderately perturbed HR (MR) and VR with highly perturbed (worst) HR (WR)(see Figure 4.1). We tested the following experimental hypotheses:

H1 There will be significantly slower movement times and a greater error rate in the three VR conditions compared to the RW condition because VR elements such as FOV, lighting, depth cues, and lag result in fine motor performance in virtual reality being slower and more error prone than the same motor task performed in the real world.

H2 Movement times and tapping accuracy in the WR condition will be slower and more error prone than performance in the other two VR conditions. There will be no movement time or accuracy differences between OR and MR conditions because small perturbations can be adapted to quickly. Extremely incorrect HR will take longer to adapt to, so a motor performance decrease will be observed in the WR condition.

H3 Movement time will decrease and accuracy scores increase with time (trial num-

ber) during each experimental viewing condition because of sensorimotor adaptation. Performance will not improve with condition number because we do not expect cross condition training effects.

To test the importance of accurate HR on VR motor performance, we calculate a subject's optimal head registration. Subjects then performed a series of Fitts-like motor tasks under four HR perturbation levels. The experiment consisted of four main phases: screening, head registration, practice trials, and experimental trials. A 4 (viewing condition) $\times$ 3 (target distance) $\times$ 4 (target width) within subjects design was used. The experiment required subjects to move a magnetically tracked stylus back and forth between two target holes (a target pair), attempting to successfully place the stylus in each hole as quickly and accurately as possible. Each placement attempt was an experimental trial. Each of the 12 trials (one block for each target pair) involved a minimum of 25 tapping trials. Each subject had completed 1200 trials at the end of the experiment.

4.1 Method

4.1.1 Participants

Twenty-four subjects (13 female, 11 male) ages 18-30 were recruited from a university undergraduate and graduate population. Viewing condition order was fully counterbalanced across subjects. All subjects were right handed, had normal or corrected to normal vision, and tested positive for binocular vision (in particular, for stereopsis) using the Titmus Stereo Fly test [38, 45]. One subject failed the colour vision test but was able to participate without any difficulties. No subjects were dismissed for not matching the criteria. Interocular distances ranged from 4.2cm to 7.0cm with a mean of 5.74cm. The distance from the view screen to the subject's first nodal point ranged from 4.3cm to 6.3cm and the mean distance was 5.61cm. No subject had more than two previous immersive VR experiences and all subjects used a computer daily. Each experimental session took approximately one hour. Subjects were financially compensated for their time.

A computer crash at the end of one experimental session resulted in no tapping data

Real World
Condition (RW)

Optimal Head
Registration
Condition (OR

Moderately
Perturbed
Registration
Condition (MR)

Highly
Perturbed
Registration
Condition (WR)
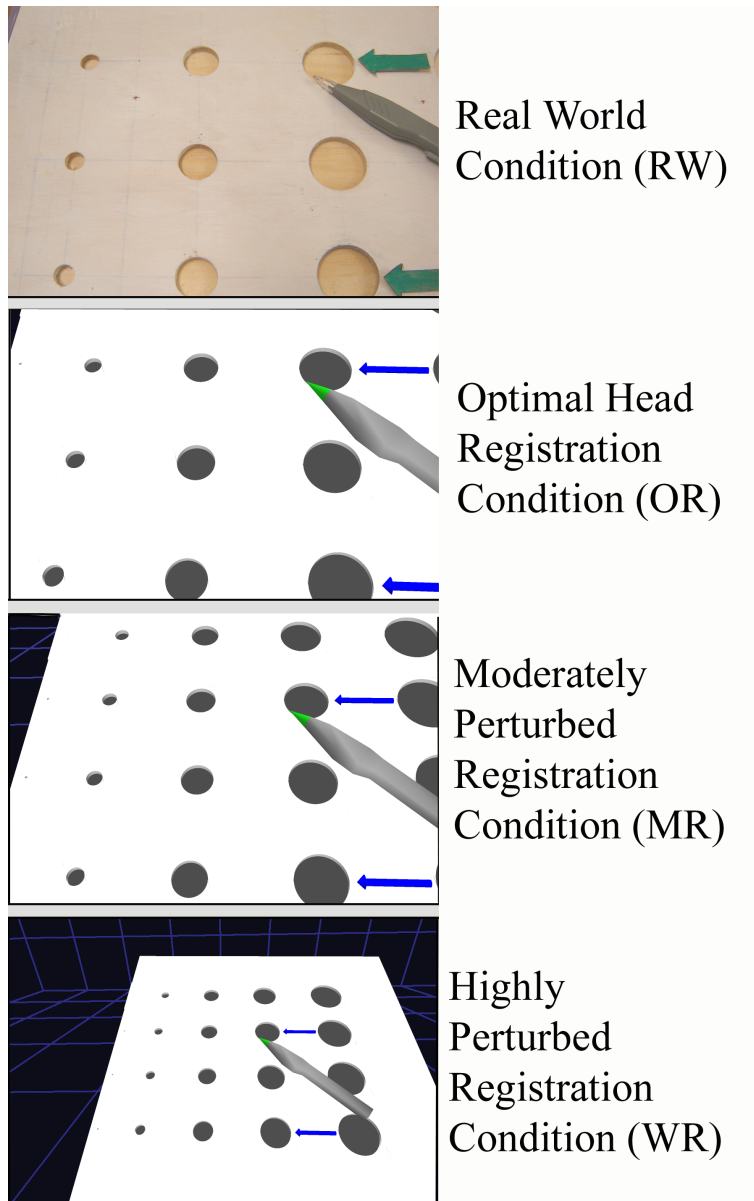
Figure 4.1: The effects of head registration accuracy on the virtual reality image. All images were taken from the same perspective.

being recorded for that participant. The subject was paid but no data from this session was used in our analyses. A new subject was recruited to be the same subject number, ensuring that condition order remained fully counterbalanced.

4.1.2 Apparatus

Subjects were tested for colour-blindness and stereoscopic vision at the beginning of the experimental session using Ishihara's Design Charts for Colour-Blindness of Un-lettered Persons and the Titmus Stereo Fly test, respectively [29, 45]. Subject eye measurements and head registration was performed using a pinhole headband (PHH) and a straight-edge ruler as described in Chapter 3. This experiment used a head-mounted display, a Polhemus Fastrak, a tapping board, and a PC to run the VR system.

Tapping Board

A 457mm×457mm wooden board was used as the tapping surface (see Figure 4.2). The board's surface was inclined at a 30-degree angle from horizontal to reduce subject neck strain. The structure was securely attached to a wooden experiment table and there was no metal within a 0.5 meter radius of the board's surface. During the experiment the only source of metal or electric distortion within this radius was from the HMD and potentially the subject. Holes in the board's surface were 12.4mm, 25.4mm, 38.1mm, and 50.8mm in diameter. Each hole was 0.3mm deep.

The location and orientation of the tapping board in the VE coordinate frame was calculated in advance by sampling the Fastrak stylus as it touched 6 pre-defined locations on the board. The surface normal of the board was then confirmed by taking a large (100+) sample of surface points (see Appendix A.3). In VR, the experimenter moved the stylus around the inside circumference of each VR tapping hole, noting any times the stylus intersected the tapping board. This approach confirmed tapping board and VR image registration error to be less than 2mm on the left side of the board and less than 3mm on the right side. Precision is limited to the 1mm resolution of the Fastrak sensors. Registration was confirmed before each experimental session to ensure the transmitter or tapping board did not shift.
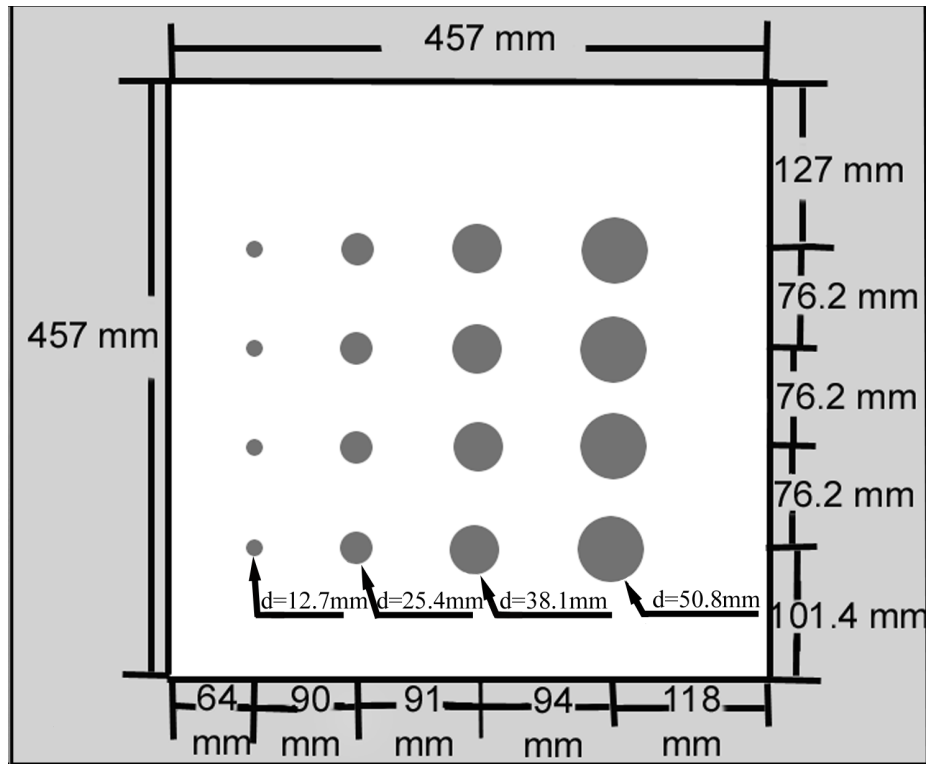
Figure 4.2: The experimental tapping board.

Magnetic Tracker

A Polhemus Fastrak monitored subject movements during the experiment using an HMD-mounted sensor for head tracking (discussed in Chapter 3), and a stylus pen for tapping [40]. A VR representation of the stylus was seen in the VR conditions, which matched the real object's position and orientation relative to the Fastrak transmitter (example views are shown in Figure 4.1. See Appendix A.2 for registration details). The tip of the stylus is the tracked point and the point used to detect tapping events. Therefore, minor real world/VR registration errors should not have affected performance. Signal accuracy of magnetic sensors degrade with distance from the transmitter. The optimal sensor range is less than 1 meter. The Fastrak transmitter was attached below the tapping board on the far left side providing maximum precision and accuracy

for the smallest target holes. Each Fastrak sensor was sampled at 60Hz by a separate processing thread in the VR program. This ensured that the sampling rate was not dependent on the update rate.

Head Mounted Display

A Kaiser ProView XL50 head mounted display with a 1024×768 resolution, 60Hz vertical refresh rate, $40°$ (horizontal) $\times 30°$ (vertical) DFOV, and dual input video channels (one for each view screen) were used in the experiment [28]. HMD dimensions and measurements were collected as discussed in Chapter 3. The HMD's optical modules allowed subjects to wear glasses if required. A large piece of grey cloth was sewn onto the HMD and placed in front of the subject's face to remove real world peripheral information and to reduce light glare on the HMD optics. The HMD weighs approximately 1 kilogram.

The Virtual Reality System

A 3 GHz dual CPU Linux machine with 1GB of memory, and an nVIDIA Quadro4 980 video card (128MB RAM with dual video out) was used to create the virtual environment using OpenGL.

Fastrak data collected by the Fastrak polling thread was stored with a corresponding time values in an array. All sensor data was then processed by the system before the VR image was updated. Multi-threading the application may have decreased the update rate, but it ensuring that the most accurate movement timing was used for collision detection and timing, while images seen by subjects were always based on the current system state.

Before the experiment was run, update rate and system lag measurements were taken using on-screen debug information. The average update rate was 40 frames per second and was always above 35 frames per second throughout the experiment. Update rates were calculated by taking the system time at the start of each render call and comparing that value to the previous system time recorded. This is the time t between render calls. The update rate is $\frac{1}{t}$. Update rates were averaged over one second intervals

to reduce variability.

System lag was estimated using a formula discussed by Ware and Balakrishnan [54]:

$$SystemLag = DeviceLag + (1.5 \times MachineLag) \tag{4.1}$$

$$SystemLag = (FastrakUpdateRate) + (1.5 \times UpdateRate) \tag{4.2}$$

Assuming our Fastrak driver's lag time is negligible (under 5ms), a Fastrak sensor's update rate is 60Hz or 16.7ms between samples. Machine lag can be roughly approximated using the update rate values collected. Using a worst case update rate, system lag was approximately 60ms, 10ms less than Ware and Balakrishnan's [54] best system lag, and 17ms longer than the minimal perceivable lag suggested by Mania et al.[36]. Keller and Colucci[31] claim that head tracked HMD systems have an average lag of 60-90ms. System lag during a rapid tapping task means that the position of the VR object seen by the subject may not match the real world object being moved. Tapping accuracy may suffer or subjects may compensate for the lag by slowing their movements ensuring that real and virtual objects are synchronized. Our system lag was noticeable, but subjects only occasionally commented on a perceived system lag during the smallest ID tapping condition. Movement times during this condition were 100-200ms and a 60ms lag may have affected subject movements. Some system lag will always be present in a VR system and previous literature suggests that our lag was reasonable for our task.

The image refresh rate and the update rate were not synchronized during this experiment. OpenGL does not provide vertical sync functionality for dual monitor systems using back buffering such as the current VR system. If two large image buffers are used for a 2 monitor display (instead of 2 front and 2 back image buffers), vertical sync signals no longer indicate that the buffers are safe for swapping; the other monitor could be drawing the scene when the sync signal arrives. Frame tearing occurred using our system but system lag was not affected by the monitor refresh rates.

The sparse virtual environment displayed a background grid and representations of the tapping board and the tapping stylus. The VE background colour was dark blue

(almost black) and all objects were viewed inside a virtual 10 meter $\times$ 10 meter dark blue gridded wireframe cube that provided orientation cues to subjects. During pilot studies several subjects believed the VR system had crashed when the stylus and tapping board left their field of view. The gridded cube alleviates these concerns. Figure 4.1 shows the physical tapping board as a subject would see it in the RW condition, and the virtual images that would be seen in each of the three VR conditions for the same head location and orientation. The VR representation of the tapping stylus was grey, like the real world object. The tip was coloured bright green to improve visibility (following suggestions by pilot subjects). When a tap occurred, or when a trial set was completed, the entire stylus turned dark grey and remained that colour until the next tapping trial was initiated (Figure 4.3). The tapping board did not change colour during the experiment.

The experimenter sat behind the subject throughout the experiment and was able to see the subject's perspective in VR using 2 LCD monitors. These monitors received their video signal through the HMD's video out connection. Command line output from the program was written to a nearby computer via a secure shell connection so the experimenter could monitor system events during the experiment without informing the subject. This output was also stored in a subject-specific text file in case it was needed for our experimental analyses.

4.1.3 Procedure

Subjects signed a consent form (Appendix C) at the start of the experiment, provided they believed they matched our screening criteria. Tests for stereopsis and colour vision were then administered and subject history including previous VR experience and frequency of computer use data were collected. Subjects then took 20 minutes to complete the subject-specific registration procedure outlined in Chapter 3. Data from the head registration was recorded and used as command line input for the VR tapping application.

The Tapping Task

For this experiment we used a Fitts-like reciprocal tapping task. Subjects were asked to repeatedly tap back and forth between two target holes of the same diameter as quickly and accurately as possible, contacting the board with the provided stylus for each trial. A tapping event occurred when either the stylus entered a hole, or came in contact with the top of the tapping board. A polygon collision algorithm determined when the stylus contacted the board. Successful taps were detected when the stylus entering the bounding box of the target hole without a collision being detected (see Appendix A.4). This meant that the stylus intersected the top planar surface of the hole but did not intersect any part of the tapping board (see Figure 4.3). If the stylus entered the bounding box of the target hole and the VR representation of the stylus intersected the tapping board, an unsuccessful trial was detected. The first point of intersection between the stylus and the target hole's bounding box is referred to as the tapping point. This point is, by definition, on the top planar surface of the tapping board.

All tapping events were calculated using the tapping board's frame of reference with the board's x-axis aligned with the hole rows, the y-axis aligned with the columns of holes and the z-axis perpendicular to the board. Aligning tapping points with the plane of the board meant that vectors from the intersection point to the target center could be represented as a 2D (x & y) vector, simplifying our analyses. Aligning the stylus with the tapping board'frame of reference also simplified the bounding box intersection calculations. The next trial began when the tapping event ceased. For each tapping event time, position (in the board's frame of reference), and trial success were recorded.

Subjects were given tapping task instructions and were shown how to perform the reciprocal tapping task. Subjects were told that for each trial they should contact the tapping board with the stylus and they should always aim to hit the center of a target hole, although anywhere in the hole was considered a successful trial. Participants were asked to tap as quickly and accurately as possible and were told to have approximately a 4% error rate (one error in 25 trials) [46]. Subjects had no trouble understanding the task. The experimenter demonstrated the real world reciprocal tapping task and
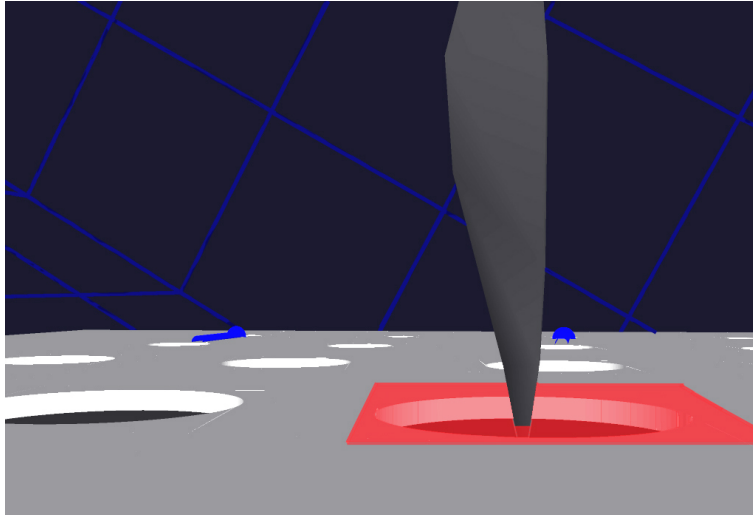
Figure 4.3: The stylus intersecting the plane of the tapping board indicating the end of a trial. The red rectangle in the figure indicates the planar surface of the board and did not appear in the experiment.

asked each subject to do the same on a different pair of targets to make sure they understood the instructions. Subjects put on the HMD (seeing the VE with optimal HR) and quickly checked for gross HR errors. By adjusting their head so that part of the tapping board was seen in VR and part of the tapping board was seen in the periphery, subjects indicated if the real and virtual boards appeared approximately in the same position. A grey cloth sewn onto the HMD was then lowered in front of the subject's eyes before tapping trials began. Subjects then performed a series of approximately 30 VR practice (or dynamic registration) trials in the OR condition using only targets in the left-most column of the tapping board (the smallest targets).

VR Practice Trials / Dynamic Registration

Practice trials served two purposes: they provided practice for participants, acclimatizing them to virtual reality, and they tested subjects to see if there was a systematic discrepancy between the HR measurements and what a subject perceived to be correct

via proprioception. After 30 tapping trials, if the distance between the mean tapping position and the target center is greater than both the standard deviation of the tapping trials and the target radius, we considered that the registration did not match the subject's perception and the HR needed adjustment (see Equation 4.3). We adjusted the VR camera or view frustum accordingly and asked the subject to perform 30 more trials. If the distance between the mean tapping position and the target center was less than the radius, we considered the HR to be correct and no more practice trials were required. If the distance between the mean tapping position and the target center was greater than the radius but less than the standard deviation of the tapping distribution, we asked subjects to do more trials until the ambiguity was resolved.

Head registration could be dynamically adjusted depending on the systematic nature of the tapping position errors. Tapping position shifts to the left or right of the target would result in the x-components of the sensor to eye vectors ($\mathbf{V}'_{\mathbf{SensorToLEye}}$ and $\mathbf{V}'_{\mathbf{SensorToREye}}$) and sensor to view screen vectors ($\mathbf{V}'_{\mathbf{SensorToLScreen}}$ and $\mathbf{V}'_{\mathbf{SensorToRScreen}}$) being altered. Tapping position shifts closer or further than the target hole meant we would adjust the eye to image plane distances (to make the board appear closer or further away). Subjects were asked to perform tapping trials until no image changes were required. This dynamic registration method assumes that sensorimotor abilities are not highly adaptable and that proprioception has a major influence on visually guided motor tasks. We hypothesized that sensorimotor adaptation readily occurs in VR and subjects would not require dynamic head registration. No systematic errors in tapping accuracy were expected.

$$Max(SD_x, Width_{target} < |X_{tappingposition} - X_{targetCenter}| \tag{4.3}$$

$$Max(SD_y, Width_{target} < |Y_{tappingposition} - Y_{targetCenter}| \tag{4.4}$$

Trial Blocks

A series of reciprocal tapping trials for a pair of targets constituted a block of trials. For each block, one hole (the bottom target) was always in the row closest to the subject

(see Figure 4.4). The top hole was one of the three remaining targets in the same column. This provided 12 possible target pair combinations and 9 distinct index of difficulty (ID) values ranging from 1.32 to 4.24 (see Figure 4.5). For each block of trials performed, blue arrows (real or virtual) on the board identified the two target holes (4.1). Subjects were asked to tap back and forth until they were instructed to stop by the experimenter. Participants indicated when they were ready to begin a new set and the experimenter pressed the enter key initiating the set. The end of each set of trials during the VR conditions was indicated by the virtual stylus remaining dark grey in colour. The first four tapping trials were not used as they were deemed additional practice trials required to get the subject up to speed. The trial data was recorded to a text file for each of the next 21 reciprocal taps. Subjects were required to always see the target hole they were attempting to tap (preventing subjects from tapping based exclusively on body memory) and this required subjects to repeatedly move their head position during the experimental trials.

Trial blocks were grouped so that all trials with the same target width were grouped together. The block ordering within this group was randomized. The order of the four groups (columns) was also randomized. Block order was determined prior to the experiment and stored in text files, enabling block order to be analyzed after the experiment.

When all twelve trial blocks were completed, the end of the condition was indicated by a black screen and subjects were given up to a five minute break to take off the HMD (when applicable), relax, stretch and ask questions. This break helped reduced the chance of VR induced nausea (cybersickness) [18].

Experimental Conditions

The distance between the VR camera and the frustum was adjusted to match the distance from the subject's eyes to the view screen in the real world. Reducing the distance between the eyes and the view frustum thus perturbed head registration, altering the correspondence between the DFOV and the GFOV. The three VR conditions differed only in how a subject's head registration data was perturbed. The image plane/view
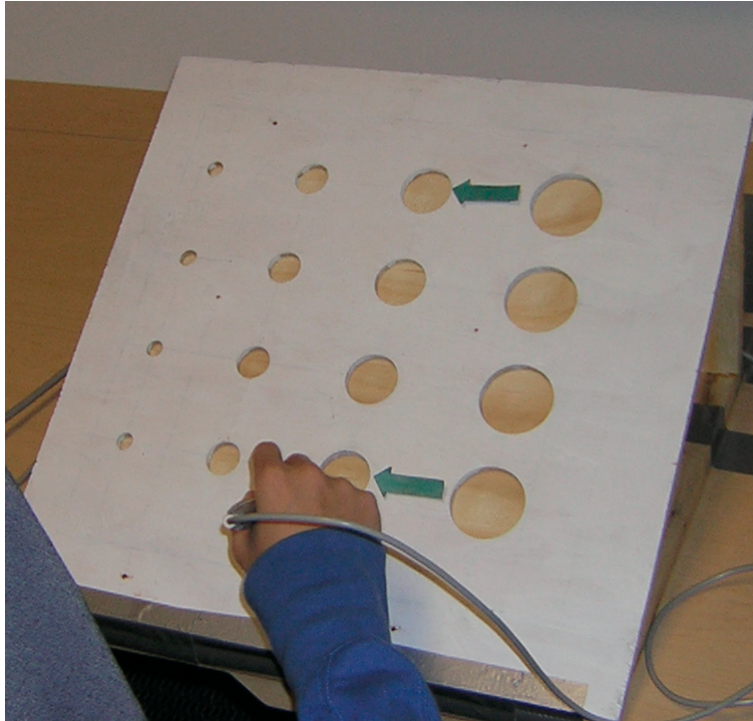
Figure 4.4: The tapping board with a subject performing a reciprocal tapping task. In the real world condition, target pairs are identified by blue cardboard arrows.

frustum in the MR condition was 0.5cm closer to the camera than in the optimal condition, and 2.0cm closer to the camera in the WR condition (see Figure 4.6). Camera-to-the-frustum perturbation was systematically altered because the view screen-to-eye distance was deemed the most difficult and most error-prone measurement to collect, making it the most likely source of registration error.

The size of an object on the HMD view screen for a viewing condition is:

$$Object_{XRScreenSize} = \frac{Object_{ORScreenSize} \times Distance_{XRCameraToImagePlane}}{Distance_{ORCameraToImagePlane}} \quad (4.5)$$

Therefore, adjustments to the GFOV result in a linear size scaling of the objects seen proportional to the change in the screen position.
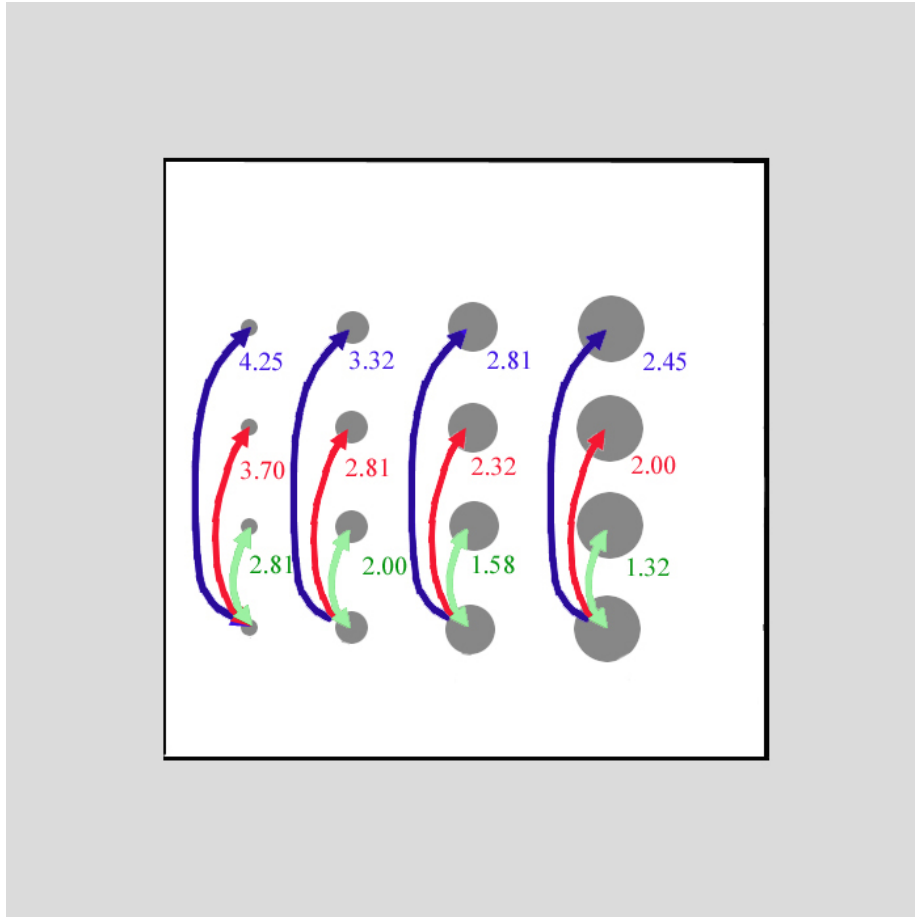
Figure 4.5: Index of difficulty values for pairs of tapping targets. A series of reciprocal taps between a target pair constitutes a block of trials.

The real world condition did not require any headgear to be worn. Coloured cardboard arrows were positioned on the tapping board by the experimenter before each group of pair trials began. The number of trials remaining, and the target holes to tap were displayed to the experimenter via private monitors.
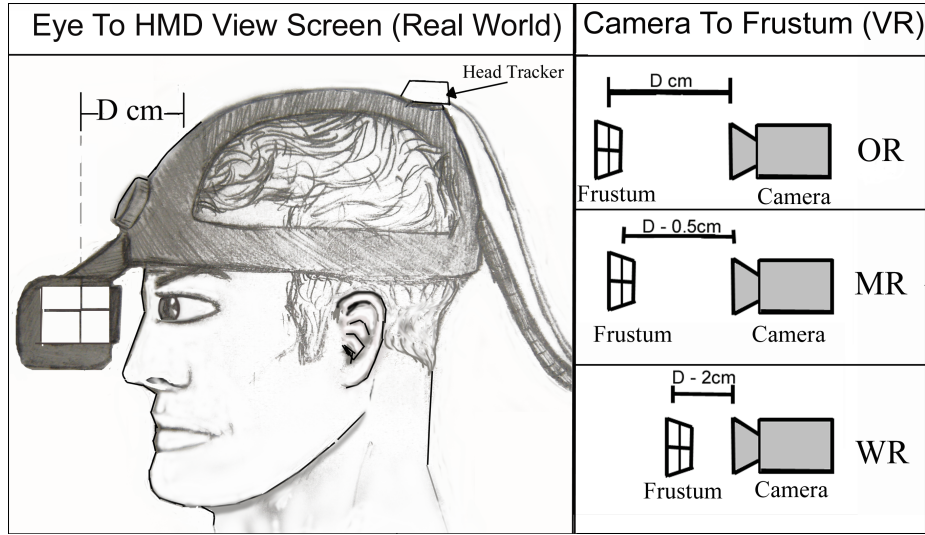
Figure 4.6: The real world/virtual reality correspondence between the eye to view screen distance (D) and the camera to image plane distance in each VR condition.

4.2 Results

To determine what impact head registration accuracy has on fine motor performance, we looked at primary measures of movement time, task correctness and information processing throughput. Movement time is the interval from the start of a tapping trial to the end of the trial (when a tap occurs). Task correctness was measured in two ways: root-mean-square (RMS) distance from the target center, and percent of trials tapped successfully (percent correct). The RMS distance from the target center is the distance between the stylus's intersection point and the target hole's center. Percentage correct refers to the percentage of included trials that were successful in a given trial pair.

We calculated information processing throughput values according to S. MacKenzie's formula [46]:

$$TP = \frac{1}{(m \times n)} \sum_{i=1}^{m} \left(\sum_{j=1}^{n} \frac{ID_{ij}}{MT_{ij}} \right) \tag{4.6}$$

where m is the number of subjects, n is the number of movement conditions, and MT_{ij}

and ID_{ij} are the movement time and index of difficulty (respectively) for subject i tapping in trial block j. This metric provides a general measure of information processing ability (or ability to do the task) for each of our experimental conditions and is measured in *bits per second (bps)*.

We used a factorial analysis of variance with repeated measures (ANOVA-RM) over three independent factors: experimental condition, trial pair order, and condition order. A one-way between subjects ANOVA was used to investigate how subject-specific factors affected the dependent measures. Subject data was coalesced for each experimental condition, and a test of correlation and linear regression between index of difficulty (ID) and movement time was performed. ID was calculated using the Shannon formulation [46]. From these linear regressions slopes, y-intercepts, and R^2 adjusted correlation coefficients were examined (Figure 4.7).

4.2.1 Data Cleaning

When subject data was initially examined, extreme outlying data points were quickly apparent. For a given target pair, any trials more than three standard deviations from the mean movement time were removed. Because we wanted to examine stable performance times, we tested for a performance plateau during a block of trials. A plateau was identified when three consecutive trial times did not improve by more than 600 milliseconds. All trials before a plateau were removed. If performance did not plateau, the final trial of the group was used. Data cleaning was intentionally conservative because we only wanted to remove atypical trials. During the real world condition approximately 0.1% of the trials were removed, while 3.5% of the OR condition trials, 1.7% of the MR condition trials, and 2.9% of the WR condition trials were removed.

4.2.2 Linear Regression and Fitts's Law

Tapping trials in all four conditions followed a Fitts's law linear regression pattern as expected. All linear regressions had R^2 adjusted coefficients greater than 0.97 and all y-intercept values were greater than 200 milliseconds (Figure 4.7) [46]. Data was also tested using Welford's model and Fitts's original formulation, but the Shannon

$$MT_{RW} = 117.25*ID - 56.96, \quad R^2 = 0.986$$
$$MT_{OR} = 238.65*ID - 148.59, \quad R^2 = 0.994$$
$$MT_{MR} = 229.95*ID - 133.78, \quad R^2 = 0.975$$
$$MT_{WR} = 249.65*ID - 181.71, \quad R^2 = 0.991$$
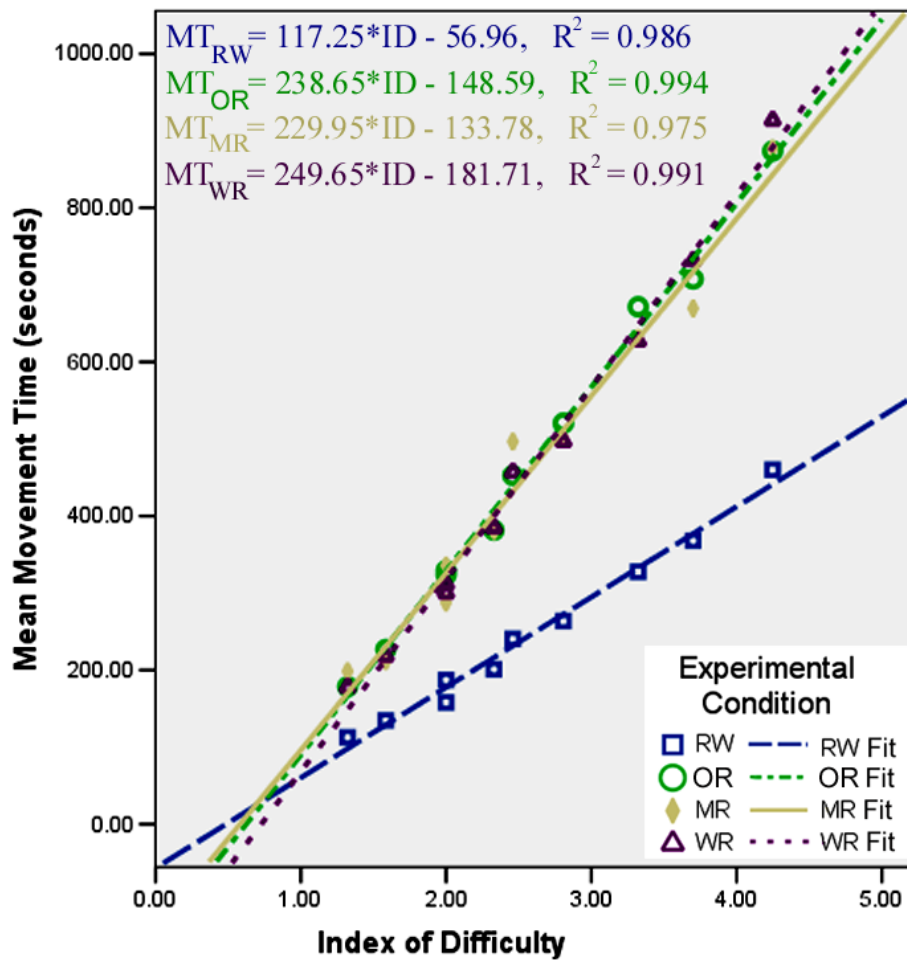
Figure 4.7: The correlation between index of difficulty and trial movement times for real world (RW), optimal VR registration (OR), moderate VR registration (MR), and worst VR registration (WR).

formulation showed better R^2 correlation coefficients than these other mathematical models. Error rates for subjects were all less than 4% after data cleaning which is consistent with how subjects were asked to perform.

4.2.3 Virtual Reality vs. The Real World

Our primary repeated measures ANOVA found a statistically significant difference between viewing conditions in terms of movement times $[F(3,21) = 43.5, p < .001]$, percent correct $[F(3,21) = 8.4, p < .001]$, and the RMS distance to the target center for top targets $[F(3,21) = 10.4, p < .001]$ and bottom targets $[F(3,21) = 12.6, p < .001]$. Mean trial times across all tapping targets were: $TrialTime_{RW} = 248$ms, $TrialTime_{OR} = 473$ms, $TrialTime_{MR} = 466$ms, and $TriallTime_{WR} = 466$ms. Percent correct rates were 99%, 97%, 98%, and 97% for the real world, OR, MR, and WR conditions, respectively. Information processing throughput values were 0.0160 bps for the RW condition, 0.00866 bps for the OR condition, 0.00880 bps for the MR condition, and 0.00888 bps for the WR condition.

Condition	Near Target RMS Variance(cm)	Far Target RMS Variance (cm)	Error Rate (%)	Sample Drop Rate (%)
RW	0.28	0.63	0.12	0.1
OR	0.319	0.79	1.5	3.5
MR	0.304	0.74	0.62	1.7
WR	0.349	0.72	1.5	2.9

Table 4.1: The average variance, error rates (% of kept samples), and % samples dropped based on experimental conditions.

A repeated measures ANOVA found no significant differences between the VR viewing conditions in terms of movement time or percent correct. A partial Eta-squared test for effect size was 0.046 for the effect of VR condition on movement time. To put this into context, Cohen suggests a partial Eta-square value of .01 indicates a small effect, .06 indicates a medium effect size, and a partial Eta-squared value of .14 is a large effect size [10].

4.2.4 Practice Effects

Subjects showed a significant tapping time improvement based on trial pair order within an experimental condition with the mean first pair's time being 0.511ms and the twelfth pair's mean time equaling 0.349ms $[F(11,13) = 4.8, p < .05]$ (see Tables 4.2 and 4.3). There was no significant effect of trial order on task correctness, nor were there any significant effects of condition order on movement time or task correctness. These results suggest that subject performance improves with exposure to an experimental condition but practice effects are not seen across conditions.

Condition Order Number	Movement Time (milliseconds)	Target RMS Variance (cm)	Number Successful (out of 21)
Condition 1	422	0.720	20.46
Condition 2	425	0.720	20.22
Condition 3	424	0.688	20.62
Condition 4	396	0.714	20.60

Table 4.2: Experiment performance metrics based on the number of experimental conditions experienced.

4.2.5 Between Subject Factors

Finally, subject-specific characteristics showed no significant effect on tapping performance. A between subjects one-way ANOVA revealed no significant relationship between sex, stereopsis scores, colour vision, measured eye position, interocular distance, or VR experience with any of the performance metrics. Dynamic adaptation was not required for any of the subjects and only two subjects required 40 practice trials rather than the minimum 30 trials.

Trial Set Number	Movement Time (milliseconds)	Target RMS Variance(cm)	Number Successful (out of 21)
Set 1	511	0.663	19.99
Set 2	463	0.672	20.67
Set 3	445	0.689	20.53
Set 4	448	0.706	20.14
Set 5	440	0.691	20.27
Set 6	371	0.704	20.65
Set 7	398	0.686	20.51
Set 8	416	0.724	20.46
Set 9	422	0.716	20.44
Set 10	371	0.742	20.71
Set 11	363	0.753	20.85
Set 12	349	0.778	20.53

Table 4.3: Experimental performance metrics based on the trial set order within an experimental condition.

4.3   Discussion

Two main results can be observed from this experiment's data. First, Fitts-like reciprocal tapping in a VE is significantly slower and more error prone than tapping in the real world. This result confirms **H1** and could be due to a combination of factors including: system lag, fewer depth cues in VR, the VR update rate, and VR image quality. Subjects may have also felt freer to move their heads without the constraints of the tethered head mounted display. Willemsen et al.[59] suggest that mechanical aspects of the HMD (such as mass) may have affected performance . The virtual environment also did not produce a shadow for the VR stylus making the stylus to tapping board distance more difficult to estimate [26]. Our results, however, suggest that the differences between VR and the real world are not due to head registration difficulties as previous research suggested [14]. The difference between the VR and real world condition

scores also demonstrate that the similar scores of VR viewing conditions are not due to subjects relying exclusively on body memory, proprioception, or tactile feedback.

The second result from our experiment is the lack of difference between VR conditions. A small effect size is defined to mean that 1% of the data's variability can be attributed to the condition grouping, while condition based variability is 6% for moderate effect sizes and 14% for large effect sizes. Thus, VR conditions have at most a small to moderate effect on trial times and less than 6% of the total movement time variability can be attributed to the VR condition. Cohen argues that a lack of significance and a small partial Eta-squared effect size indicate either little or no effect is present or the data collected is noisy. High R^2 correlation values suggest that our data is not noisy and HR perturbation has no effect on subjects performing a Fitts-like reciprocal tapping task.

This result does not agree with **H2**, but it is consistent with sensorimotor plasticity theories. If sensorimotor adaptation is occurring, it happens more rapidly than we expected. This rapid adaptation also meant that there was no detectable difference in adaptation time between viewing conditions. The trial performance improvements over time within a condition, but not between conditions, are also consistent with a sensorimotor adaptation explanation of our results and confirms **H3**.

There is a second possible explanation for our failure to observe an effect of VR condition. By moving the view frustum closer to the camera, we increase the GFOV and reduce the size of the objects on the view screens. Eggleston et al. [14] found that movement time for a Fitts-like tapping task decreased when the GFOV was increased and targets had a moderate difficulty level. In our experiment, target distances were 7.62cm, 15.24cm, and 22.86cm, and subjects viewing the board from 30cm away could see approximately 17.7cm of the board in the OR condition, 20.4cm in MR, and 37.2cm in WR. Thus, fewer head movements were required as the GFOV increased and performance may have improved because of this. The zoom-out effect of our HR perturbation, however, reduces image details and the DFOV no longer matches the GFOV, possibly reducing tapping performance. Hence, GFOV and VR image quality may act as opposite factors, producing the lack of effect we observed. Nevertheless, either explanation of our results suggests that other virtual reality factors are more important

than HR for motor performance in VR.

4.3.1 Potential Sources of Experimental Error

This experiment has several potential sources of error that could have affected results. HR measurement errors were minimized but may have arisen. If all three head registration qualities had a substantial measurement error, relatively small condition-induced perturbations wouldn't make much of a difference to the accuracy of the HR. Subjects were asked to informally confirm the real world/VR correspondence at the start of the experiment, making this possibility unlikely. We believe the eye to view screen distance is the most probable error source using our HR methodology. If the view screen to eye measurement was taken incorrectly, the actual optimal registration may have been bracketed by the OR and MR conditions and no difference would be found between these conditions. Subject performance was similar in all three VR conditions, however, suggesting that bracketing is not the issue.

System lag may have affected tapping times requiring subjects to slow down movements in all VR conditions. Subjects could then compensate for HR perturbations at these reduced speeds. Visual lag may also prompt participants to rely exclusively on body memory to do the tapping task in the VR conditions. System lag during this experiment was 60ms and not considered abnormally slow, so these explanations are improbable.

The significant within condition practice effect observed may be due to subject motivation, or subjects adapting to system lag rather than sensorimotor adaptation. Subjects may need to warm up to the tapping task after each 5 minute break or they may be hesitant to tap quickly in a new environment. These alternative explanations for the trial order performance effects observed can not be accounted for by our study.

Effective Target Width

Fitts-like tapping tasks rely on subjects constraining their error rates to approximately 4% so that movement time is the only factor influenced by the task difficulty. If a subject's error rate exceeds 4%, the target width can be enlarged so that 96% of the

tapping trials are within this new target. This adjusted target size is the effective target width. Replacing target widths with effective target widths (W_e) for index of difficulty calculations as suggested by Soukoreff and S. MacKenzie [46] provided a worse linear regression data fit than using the actual target diameter. This was unexpected since W_e is adjusted to fit the optimum error rate. The naturally low error rates in this experiment meant that W_e was smaller than the original target size. Subjects were also instructed to aim for the center of a target during each trial, potentially reducing the standard deviation of the intersection point distribution. If tapping points were tightly clustered around the center of a target, the effective target width would greatly over-estimate the difficulty of the tapping task and W_e values would be inappropriate for index of difficulty calculations. We used the raw target diameter for our calculations instead.

Tapping Detection Method

Virtual object based collision detection is not the ideal way to determine the end of a trial and this may have resulted in inaccurate data. Real world objects are the only concrete constraints to subject movements in VR environments, therefore tapping events should be based on real world object interactions. If the real world and VE are not perfectly aligned, subjects may not be able to put the stylus in a (visual) target hole, if the stylus contacts the top of the real board instead. Virtual objects were designed to precisely emulate the real world objects used. System bugs and measurement errors can result in the VR objects not interacting as expected. The chances of incorrectly identifying a trial result increases as the target size decreases, because we assume a normal distribution around the target hole's center. Tapping detection using virtual objects is an error source with no predictable effect on our results.

This experiment's initial tapping detection method used an electrical signal sent via the computer's parallel port to indicate a trial's completion. A metal stylus completed a (9V) circuit by touching the top or bottom aluminum plate of the tapping board. This tapping detection based on real world objects guaranteed that trial success/failure were correctly identified and based exclusively on physical constraints. Unfortunately, sufficient quantities of aluminum warp electro-magnetic fields so a VR/real world object

correspondence could not be established (the Polhemus Fastrak is well known to work poorly in the presence of electro-magnetic fields). We attempted to compensate by using piecewise linear interpolation to adjust sensor data. This approach was partially successful but sensor data remained distorted by up to 5mm, insufficient precision for our experiment. If different motor tasks such as a steering task or more elaborate interfaces are used in future experiments, real world collision detection may be necessary to ensure valid data [1].

4.3.2 Experimental Conclusions and VR Strategies

This experiment's results suggest that precise head registration is not a critical factor for VR motor performance and other elements of a VR system should perhaps take a higher priority. If time and money are not constraints, optimal registration and calibration should always be performed to guarantee realistic VR. Sensorimotor adaptation between the real world and virtual world should be avoided to reduce negative aftereffects in real world tasks.

This research is an initial examination into the effects of HR on VR usability. Exact HR is not always possible or practical and our results suggest that accurate HR is not necessary for quick or informal VR use. Approximate measurements may suffice. People may not be able to adapt to specific or multi-axial head registration errors and future experiments are needed before more general conclusions can be made. We believe that interocular distance should continue to be precisely measured because it is relatively easy to find (a ruler in front of the subject's eyes is sufficient), is constant for individual subjects, and correct measurements may reduce eye strain and reduce the possibility of cybersickness [18].

Approximating head registration in future immersive VR systems should be done with sensorimotor adaptation issues in mind. Users should be made aware of real world negative aftereffect issues after using an immersive VR system (like they are currently warned about cybersickness). If HR values are approximated, practice time in the VE may help subjects quickly adapt to the environment. Ideally this practice would require participants to move and manipulate objects in the virtual world with

visual feedback about their body's position. Self-generated movements with sensory feedback are required for sensorimotor adaptation to occur [24]. We also suggest that system lag should be minimized and exact object registration should be attempted if PHF is being provided. This should ensure that the sensorimotor feedback required for adaptation is as efficient as possible.

Chapter 5

Conclusions and Future Work

5.1 Future Work

The experiment described in Chapter 4 is an initial step to understanding the effects of registration and calibration accuracy on virtual reality usability. Our experiment did not provide us with a clear explanation for HR's lack of effect so other registration errors should be investigated to ensure our results generalize. Lateral shifts in vision are a logical next condition to test since they are the most common disturbance in prism adaptation experiments [14]. Repeating this experiment using a lateral shift instead of a camera/view frustum distance change would also mean that the GFOV is not altered. If no significant differences between perturbation amounts were found, the opposite factors explanation would not be supported.

Fine motor performance ability does not necessarily correspond with other usability factors. Future experiments should also test HR effects on perception, subject preference, and a sense of presence in VR. Participants in our experiment only noticed the change in registration in the worst condition where they stated that the tapping board looked further away. This suggests that either insufficient linear perspective cues were displayed or subjects were not consciously aware of the distortion. Perception and motor task studies investigating HR quality and VE complexity would help clarify if the present study's minimalist VE affected our results. Finally, other motor tasks should be investigated to ensure that reciprocal tapping was not a special case where sensorimotor adaptation occurs. Fine motor tasks requiring less movement, such as a steering task, may be affected by HR quality since adaptation will not occur as readily [1].

Other studies involving collaboration in VR are planned, to investigate how shared

VR perspectives are perceived. In collaborative VEs, such as a VR cave, multiple people can simultaneously view the same virtual object but only one viewpoint can have the correct perspective. This research provides evidence that people not seeing a correct perspective image will not be highly disadvantaged provided the perturbations to their ideal view are constant and they have time to adapt.

5.2 General Conclusions

Our experiment examined the effect of head registration perturbations on motor performance in virtual reality environments. We utilized a low cost, efficient and precise methodology for immersive VR head registration using a pinhole headband to calculate the optimal HR for subjects. We then asked subjects to perform a series of Fitts-like reciprocal tapping tasks under four viewing conditions: real world, optimal registration, moderate perturbed registration, and highly perturbed registration. We found that motor performance in an immersive VR system is significantly worse than performance in the real world. We have provided evidence that perturbing the distance between the camera and the view frustum in VR does not affect subject motor performance, even when the perturbation is extreme. We believe that this lack of an effect is due to sensorimotor adaptation, and that subjects are able to adjust to most reasonable head registration errors with enough practice. Exact head registration may not be a crucial factor in virtual reality use and other VR factors may be more critical for fine motor performance.

This experiment provides further evidence that the human element should not be taken for granted in VR systems. Yokokohji et al.'s [60] assumption that improved HR results in VR quality improvements is not supported by our results so the benefits of optimal HR may not always be worth the costs. If time and money are not constraints, optimal head registration and calibration should always be performed to guarantee realistic VR. However, optimal head registration may constrain participant movements and may require costly, time consuming, or obtrusive head registration techniques. If optimal HR is not possible, then we believe participants should be provided with VR practice tasks. Five minutes of VR practice may be far more beneficial than spending

twenty minutes getting the head registration exact.

Bibliography

[1] J. Accot and S. Zhai. Beyond Fitts' law: models for trajectory-based HCI tasks. In *CHI '97: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 295–302, New York, NY, USA, 1997. ACM Press.

[2] E. Angel. *Interactive Computer Graphics with OpenGL, 3rd Edition*. Addison-Wesley, 2003.

[3] R. Arsenault and C. Ware. Eye-hand co-ordination with force feedback. In *CHI '00: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 408–414, New York, NY, USA, 2000. ACM Press.

[4] R. Arsenault and C. Ware. The importance of stereo, eye coupled perspective and touch for eye-hand coordination. *Presence: Teleoperators and Virtual Environments*, 13(5):549–559, 2004.

[5] K. Arthur, K. Booth, and C. Ware. Evaluating 3D task performance for fish tank virtual worlds. *ACM Trans. Inf. Syst.*, 11(3):239–265, 1993.

[6] R. Azuma and G. Bishop. Improving static and dynamic registration in an optical see-through HMD. In *SIGGRAPH '94: Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, pages 197–204, New York, NY, USA, 1994. ACM Press.

[7] R. T. Azuma. Course notes on "registration" from course notes #30: Making direct manipulation work in virtual reality. SIGGRAPH '97: Proceedings of the 24th annual conference on computer graphics and interactive techniques, 1997. http://www.cs.unc.edu/~azuma/sig97reg.pdf.

[8] W. Barfield and T. A. Furness III, editors. *Virtual environments and advanced interface design*, chapter 11. Oxford University Press, Inc., New York, NY, USA, 1995.

[9] J. Boritz and K. S. Booth. A study of interactive 3D point location in a computer simulated virtual environment. In *VRST '97: Proceedings of the ACM symposium on Virtual reality software and technology*, pages 181–187, New York, NY, USA, 1997. ACM Press.

[10] J. Cohen. *Statisical Power Analysis for the Behavioral Sciences, 2nd Edition*, pages 273–288. Academic Press, New York, NY, USA, 1977.

[11] M. Czerwinski, D. S. Tan, and G. G. Robertson. Women take a wider view. In *CHI '02: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 195–202, New York, NY, USA, 2002. ACM Press.

[12] M. Deering. High resolution virtual reality. In *SIGGRAPH '92: Proceedings of the 19th annual conference on Computer graphics and interactive techniques*, pages 195–202, New York, NY, USA, 1992. ACM Press.

[13] M. Deering. The limits of human vision, 1998. http://www.swift.ac.uk/vision.pdf.

[14] R. Eggleston, W. Janson, and K. Aldrich. Virtual reality system effects on size distance judgments in a virtual environment. In *VRAIS '96: Proceedings of the 1996 Virtual Reality Annual International Symposium (VRAIS 96)*, page 139, Washington, DC, USA, 1996. IEEE Computer Society.

[15] R. G. Eggleston, W. P. Janson, and K. A. Aldrich. Field of view effects on a direct manipulation task in a virtual environment. In *Proceedings of the Human Factors and Ergonomics Society, 41st Annual Meeting*, pages 1244–1248, 1997.

[16] D. Elliott. The influence of visual target and limb information on manual aiming. *Canadian Journal of Psychology*, 42(1):57–68, 1988.

[17] S. R. Ellis, M. Tyler, W. S. Kim, and L. Stark. Three-dimensional tracking with misalignment between display and control axes. *SAE Transactions: Journal of Aerospace*, 100(1):985–989, 1991.

[18] M. Finch and P. A. Howarth. Virtual reality, nausea, virtual simulation sickness, vision, adverse symptoms, head movements and VR sickness; a comparison between two methods of controlling movement within a virtual environment. Technical report, Department of Human Sciences, Loughborough University, September 1996. http://www.lboro.ac.uk/departments/hu/groups/viserg/9606.htm.

[19] P. M. Fitts. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47:381–391, 1954.

[20] J. Foley, A. Van Dam, S. Feiner, and J. Hughes. *Computer Graphics: Principles and Practice in C, 2nd Edition*. Addison-Wesley systems programming series, 1990.

[21] M. Ghazisaedy, D. Adamczyk, D. J. Sandin, R. V. Kenyon, and T. A. DeFanti. Ultrasonic calibration of a magnetic tracker in a virtual reality space. In *VRAIS '95: Proceedings of the Virtual Reality Annual International Symposium (VRAIS'95)*, pages 179–188, Washington, DC, USA, 1995. IEEE Computer Society.

[22] J. Groen and P. J. Werkhoven. Visuomotor adaptation to virtual hand position in interactive virtual environments. *Presence: Teleoperators and virtual environments*, 7(5):429–446, 1998.

[23] T. Grossman and R. Balakrishnan. Pointing at trivariate targets in 3D environments. In *CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 447–454, New York, NY, USA, 2004. ACM Press.

[24] R. Held. Plasticity in sensory-motor systems. *Scientific American*, 213(5):84–94, 1965.

[25] R. Held and S. Freedman. Plasticity in human sensorimotor control. *Science*, 142:455–462, 1963.

[26] H. Hu, Amy Gooch, S. Creem-Regehr, and W. Thompson. Visual cues for perceiving distances from objects to surfaces. *Presence: Teleoperators and Virtual Environments*, 11(6):652–664, 2002.

[27] M. Ikits, J. D. Brederson, C. D. Hansen, and J. M. Hollerbach. An improved calibration framework for electromagnetic tracking devices. In *VR '01: Proceedings of the Virtual Reality 2001 Conference (VR'01)*, page 63, Washington, DC, USA, 2001. IEEE Computer Society.

[28] Kaiser Electro-Optics Inc. ProView XL35 & XL50, 2006. `http://www.rockwellcollins.com/optronics/ProductsAndServices` `/DisplayProducts/HeadMountedDisplays/ ProViewXL50/pages738.html`.

[29] S. Ishihara. *Ishihara's design charts for colour-blindness of unlettered persons*. Kanehara Shuppan Co, Tokyo, Japan, 1990.

[30] P. M. Keele and J. R. Peterson. Information capacity of discrete motor responses. *Journal of Experimental Psychology*, 67:103–112, 1964.

[31] K. Keller and D. Colucci. Perception in HMDs: What is it in head mounted displays (HMDs) that really make them all so terrible? In R. Lewandowski, L. Haworth, and H. Girolamo, editors, *Proc. SPIE Vol. 3362, Helmet and Head Mounted Displays III '98*, volume 3362, pages 46–53, 1998.

[32] L. Kohli and M. Whitton. The haptic hand: providing user interface feedback with the non-dominant hand in virtual environments. In *GI '05: Proceedings of the 2005 conference on Graphics Interface*, pages 1–8. Canadian Human-Computer Communications Society, 2005.

[33] R. W. Lindeman, J. L. Sibert, and J. K. Hahn. Towards usable VR: an empirical study of user interfaces for immersive virtual environments. In *CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 64–71, New York, NY, USA, 1999. ACM Press.

[34] I. S. MacKenzie. Fitts' law as a research and design tool in human-computer interaction. *Human-Computer Interaction*, 7:91–139, 1992.

[35] R. L. Mandryk. Using the finger for interaction in virtual environments. Master's thesis, School of Kinesiology, Simon Fraser, 2000.

[36] K. Mania, B. D. Adelstein, S. R. Ellis, and M. I. Hill. Perceptual sensitivity to head tracking latency in virtual environments with varying degrees of scene complexity. In *APGV '04: Proceedings of the 1st Symposium on Applied perception in graphics and visualization*, pages 39–47, New York, NY, USA, 2004. ACM Press.

[37] K. Nemire, R. Jacoby, and S. Ellis. Simulation fidelity of a virtual environment display. *Human Factors*, 36(1):79–93, 1994.

[38] R. Patterson and W. L. Martin. Human stereopsis. *Human Factors*, 34(6):669–692, 1992.

[39] R. Pausch, D. Proffitt, and G. Williams. Quantifying immersion in virtual reality. In *SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 13–18, New York, NY, USA, 1997. ACM Press/Addison-Wesley Publishing Co.

[40] Polhemus. Polhemus Fastrak (magnetic motion capture equipment), 2005. http://www.polhemus.com.

[41] G. M. Redding and B. Wallace. Effects of movement duration and visual feedback on visual and proprioceptive components of prism adaptation. *Journal of Motor Behavior*, 26(3):257–266, 1994.

[42] J. P. Rolland, F. A. Biocca, T. Barlow, and A. Kancherla. Quantification of adaptation to virtual-eye location in see-thru head-mounted displays. In *VRAIS '95: Proceedings of the Virtual Reality Annual International Symposium (VRAIS'95)*, page 56, Washington, DC, USA, 1995. IEEE Computer Society.

[43] Y. Rossetti, K. Koga, and T. Mano. Prismatic displacement of vision induces transient changes in the timing of eye-hand coordination. *Perception & Psychophysics*, 54(3):355–364, 1993.

[44] S. Schwartz. *Visual Perception: A Clinical Orientation, 3rd Edition*, chapter 11. McGraw-Hill, Medical Publishing Division, New York, NY, USA, 2004.

[45] W. W. Somers and M. J. Hamilton. Estimation of the stereoscopic threshold utilizing perceived depth. *Ophthalmic and Physiological Optics*, 4(3):245–250, 1984.

[46] R. W. Soukoreff and I. S. MacKenzie. Towards a standard for pointing device evaluation: perspectives on 27 years of Fitts' law research in HCI. *Int. J. Hum.-Comput. Stud.*, 61(6):751–789, 2004.

[47] A. Stevenson. Calibrating head-coupled virtual reality systems. Master's thesis, Department of Computer Science, The University of British Columbia, 2002.

[48] V. Summers, K. Booth, T. Calvert, E. Graham, and C. MacKenzie. Calibration for augmented reality experimental testbeds. In *SI3D '99: Proceedings of the 1999 symposium on Interactive 3D graphics*, pages 155–162, New York, NY, USA, 1999. ACM Press.

[49] J. Vince. *Introduction to Virtual Reality*. Springer, 2004.

[50] Y. Wang and C. L. MacKenzie. Object manipulation in virtual environments: relative size matters. In *CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 48–55, New York, NY, USA, 1999. ACM Press.

[51] Y. Wang and C. L. MacKenzie. The role of contextual haptic and visual constraints on object manipulation in virtual environments. In *CHI '00: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 532–539, New York, NY, USA, 2000. ACM Press.

[52] C. Ware and R. Arsenault. Frames of reference in virtual object rotation. In *APGV '04: Proceedings of the 1st Symposium on applied perception in graphics and visualization*, pages 135–141, New York, NY, USA, 2004. ACM Press.

[53] C. Ware, K. Arthur, and K. Booth. Fish tank virtual reality. In *CHI '93: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 37–42, New York, NY, USA, 1993. ACM Press.

[54] C. Ware and R. Balakrishnan. Reaching for objects in VR displays: lag and frame rate. *ACM Trans. Comput.-Hum. Interact.*, 1(4):331–356, 1994.

[55] C. Ware and D. Fleet. Integrating flying and fish tank metaphors with cyclopean scale. In *CGI '97: Proceedings of the 1997 Conference on Computer Graphics International*, page 39, Washington, DC, USA, 1997. IEEE Computer Society.

[56] C. Ware and J. Rose. Rotating virtual objects with real handles. *ACM Trans. Comput.-Hum. Interact.*, 6(2):162–180, 1999.

[57] G. Welch and G. Howarth. An introduction to the Kalman filter. Technical report, Department of Computer Science, University of North Carolina at Chapel Hill, April 2004. http://www.cs.unc.edu/∼welch/media/pdf/kalman_intro.pdf.

[58] M. J. Wells and M. Venturino. Performance and head movements using a helmet-mounted display with different sized fields-of-view. *Optical Engineering*, 29:870–877, 1990.

[59] P. Willemsen, M. Colton, S. Creem-Regehr, and W. Thompson. The effects of head-mounted display mechanics on distance judgments in virtual environments. In *APGV '04: Proceedings of the 1st symposium on Applied perception in graphics and visualization*, pages 35–38, New York, NY, USA, 2004. ACM Press.

[60] Y. Yokokohji, R. L. Hollis, and T. Kanade. WYSIWYF display: a visual/haptic interface to virtual environment. *Presence: Teleoperators and Virtual Environments*, 8(4):412–434, 1999.

[61] S. Zhai, W. Buxton, and P. Milgram. The "Silk Cursor": investigating transparency for 3D target acquisition. In *CHI '94: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 459–464, New York, NY, USA, 1994. ACM Press.

Appendix A

Object Tracking and Affine Transformations

Symbols used:

$\mathbf{M}_m$ = a matrix m,

$\mathbf{V}_n$ = a vector n,

$\mathbf{P}_o$ = a plane o, and

$\mathbf{Pt}_p$ = a point (or vertex) p.

Some useful formulas commonly used in this appendix are:

$Mag(\mathbf{V}_a) = \sqrt{\mathbf{V}_a \cdot \mathbf{V}_a}$ How a vector's magnitude is calculated.

$Norm(V_c) = V_c/Mag(\mathbf{V}_c)$ Normalize a vector so the magnitude is 1.0

A.1   Calculating Head-Coupled Perspectives

The series of affine transformations used to generate a head-coupled perspective is described in the following series of equations. A Fastrak sample consisted of a position relative to the transmitter's origin, and Euler angles representing the orientation of the sensor relative to the transmitter.

The rotation matrix $\mathbf{M}_{rotate}$ is calculated using Euler angles ϕ (z-axis rotation), θ (x-axis rotation), and σ (y-axis rotation).

$$\mathbf{M}_{rotateZ}(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi & 0 & 0 \\ \sin\phi & \cos\phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{M}_{rotateX}(\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{M}_{rotateY}(\sigma) = \begin{pmatrix} \cos\sigma & 0 & \sin\sigma & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\sigma & 0 & \cos\sigma & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{M}_{rotate} = \mathbf{M}_{rotateZ}(\phi)\mathbf{M}_{rotateX}(\theta)\mathbf{M}_{rotateY}(\sigma) \tag{A.1}$$

and to reverse this rotation we do the following:

$$\mathbf{M}_{rev\_rotate} = \mathbf{M}_{rotateY}(-\sigma)\mathbf{M}_{rotateX}(-\theta)\mathbf{M}_{rotateZ}(-\phi) = \mathbf{M}_{rotate}^{-1} \tag{A.2}$$

The head sensor's position in the virtual environment is achieved by adding the transmitter's position in the virtual environment to the sensor's location relative to the transmitter:

$$\mathbf{V}_{VEOriginToSensor} = \mathbf{V}_{TransmitterToSensor} - \mathbf{V}_{TransmitterToVEOrigin} \tag{A.3}$$

$\mathbf{V}_{SensorToLEye}$ is the constant vector from the head sensor to the left eye position relative to the HMD sensor in the 0,0,0 Euler angle position. To get this vector in the virtual environment coordinate frame we simply rotate it by the current orientation of the head sensor:

$$\mathbf{V}'_{SensorToLEye} = \mathbf{M}_{rotate} \times \mathbf{V}_{SensorToLEye} \tag{A.4}$$

$$\mathbf{V}_{VEOriginToLEye} = \mathbf{V}'_{SensorToLEye} + \mathbf{V}_{VEOriginToSensor} \tag{A.5}$$

The same calculations are also performed on vector $\mathbf{V}_{SensorToREye}$ (from the sensor to the right eye).

Each view screen can be represented as three points (defining a rectangle) with points 1 and 3 representing opposite corners of the screen. Performing a rotation and translation (Equations A.4 and A.5) on points representing planes $\mathbf{P}_{SensorToLScreen}$ and $\mathbf{P}_{SensorToRScreen}$ gives us $\mathbf{P}_{VEOriginToLScreen}$ and $\mathbf{P}_{VEOriginToRScreen}$. Therefore, we have all the data required for a head-coupled perspective of the VE.

A.1.1 OpenGL Implementation Issues

The equations above should theoretically enable head-coupled perspectives to work, however, OpenGL only positions view frustums relative to the virtual camera's co-ordinate frame, not relative to the VE origin. Thus, we did not need to calculate $\mathbf{P}_{VEOriginToLScreen}$ and $\mathbf{P}_{VEOriginToRScreen}$ so instead of rotating the camera and view frustums, we rotated the world around the camera instead. We positioned our virtual camera at the origin of the VE, and set the projection plane of the view frustums to match $\mathbf{P}_{LEyeToLScreen}$ and $\mathbf{P}_{REyeToRScreen}$ (which are constant for a subject). The frustum depth was set to 10 meters. Finally, to generate the correct perspective we translated the world by $-\mathbf{V}_{VEOriginToLEye}$ (for the left eye) and rotated the virtual world around the camera by $\mathbf{M}_{rev\_rotate}$.

Rotating the world about the camera and view frustum is more conceptually complicated. In terms of the mathematics, rotating the world around the camera and view frustum is the equivalent to rotating the camera and view frustum in the world. Hence, in the main text our head-coupled perspective algorithm is presented in terms of mobile cameras and frustums to keep the concept simple.

A.2 Stylus Registration

The VR representation of the Fastrak stylus was intentionally modeled to eliminate real world/VR registration requirements (unlike an earlier experimental design, where the metal stylus required an extensive registration methodology). The VR stylus with zero

rotation is modeled to match the real world stylus with all Euler angles equaling zero. Therefore, the tip of the VR stylus is where the real stylus tip is located (in VE coordinate) and the current Euler angles of the real object is what the virtual object needs to be rotated by. Before rendering the VR stylus, the current matrix on OpenGL's model matrix stack ($\mathbf{M}_{model}$) simply needs to be multiplied by a translation matrix and a rotation matrix and pushed on the stack. After the stylus is rendered, the stack is popped.

$$\mathbf{M}_{rotate\_Stylus} = \mathbf{M}_{rotateZ\_Stylus}(\phi)\mathbf{M}_{rotateX\_Stylus}(\theta)\mathbf{M}_{rotateY\_Stylus}(\sigma)$$

$$\mathbf{M}_{model} = (\mathbf{M}_{rotate\_Stylus}\mathbf{M}_{translate\_VEOriginToStylus})\mathbf{M}_{model} \tag{A.6}$$

Thus for all vertices (points) $\mathbf{Pt}_i$ (in the stylus reference frame) that compose the stylus image:

$$\mathbf{P}_{VRStylus\_i} = (\mathbf{M}_{rotate\_Stylus}\mathbf{M}_{translate\_VEOriginToStylus})\mathbf{P}_i \tag{A.7}$$

A.3 Tapping Board Registration

Registering the tapping board requires us to calculate the tapping board's position and orientation in the VE.

A.3.1 Calculating the Planar Surface

The surface normal of the tapping board was calculated by sampling a large number of points on the board's surface. Three points were randomly selected from this sample array to determine a planar surface $\mathbf{P}$ (with points $\mathbf{Pt}_1$, $\mathbf{Pt}_2$, and $\mathbf{Pt}_3$):

$$\mathbf{V}_{planeNormal} = (\mathbf{Pt}_2 - \mathbf{Pt}_1) \times (\mathbf{Pt}_3 - \mathbf{Pt}_1) \tag{A.8}$$

A point's distance from the plane $\mathbf{P}$ is calculated by getting the distance from the point $\mathbf{Pt}_i$ to a point on the plane $\mathbf{Pt}_1$ multiplied by the cosine of the angle between a co-planar vector and the vector from $\mathbf{Pt}_1$ to $\mathbf{Pt}_i$:

$\mathbf{V}_{1,3} = \mathbf{Pt}_3 - \mathbf{Pt}_1$ A vector on the plane.

$\mathbf{V}_{1,i} = \mathbf{Pt}_i - \mathbf{Pt}_1$ A vector from a point on the plane to point i.

The angle between our two vectors is:

$$\cos\theta = (\mathbf{V}_{1,3} \cdot \mathbf{V}_{1,i})/(Mag(\mathbf{V}_{1,3}) \times Mag(\mathbf{V}_{1,i})) \tag{A.9}$$

$$D_{PtToPlane}(\mathbf{Pt}_i, \mathbf{P}_j) = Mag(\mathbf{Pt}_i - \mathbf{Pt}_1) \times \cos\theta \tag{A.10}$$

The squared (perpendicular) distance of each point from plane $\mathbf{P}$ was summed to give a planar fitness value for the current plane $\mathbf{P}_k$.

$$PlanarFitness = \sum_{j=1}^{n}(D_{PtToPlane}(\mathbf{Pt}_j, \mathbf{P}_k)^2) \tag{A.11}$$

This procedure was repeated 30 times and the plane with the best fitness was chosen ($\mathbf{P}_{VEPlaneTB}$). The axis-angle difference between the normals of $\mathbf{P}_{VEPlaneTB}$ and the plane z = 0 is calculated and stored as a rotation matrix $\mathbf{M}_{TBToVEPlane}$

Calculating the angular difference between two vectors is done by getting 1) the axis we are rotating around, and 2) the amount of rotation (using the right hand rule).

$$\mathbf{V}_{rotation\_axis}(\mathbf{V}_1, \mathbf{V}_2) = Norm(\mathbf{V}_1\mathbf{V}_2), \tag{A.12}$$

$$Angle\theta = \arccos \mathbf{V}_1 \cdot \mathbf{V}_2 \tag{A.13}$$

A.3.2 Planar Registration Points

Tapping board registration also required the Fastrak stylus to locate six pre-defined points on the board (*RWPoints*). Six points in the tapping board's frame of reference were stored in the text file with the board's layout information (*TBPoints*). A vector between any two *TBPoints* matches a corresponding vector between *RWPoints* and these vectors differ by two rotations: $\mathbf{M}_{TBToVEPlane}$ and an axis-angle rotation of σ (a rotation about the plane's ($\mathbf{P}_{VEPlaneTB}$) normal). These two rotations can be combined to a single rotation matrix $\mathbf{M}_{VEBoardOrient}$. For each rotation matrix calculated, we tested how well it fit our observed data points (we rotated all TBPoints and summed the absolute differences between the RWPoints and *TBPoints'*). The rotation with the

best fit was chosen as the board's orientation ($\mathbf{M}_{VEBoardOrient}$). The position of the tapping board in the VE could then be calculated:

$$\mathbf{V}_{TBPos} = \mathbf{V}_{RWPoint\_i} - (\mathbf{M}_{VEBoardOrient} * \mathbf{V}_{TBPoint\_i}) \tag{A.14}$$

This means the tapping board now has an orientation and position in the virtual environment and the virtual object is registered to the real world object.

A.4   Tapping Events Detection

To calculate tapping events, the tapping board's frame of reference was used. The board's origin was defined to be the top lower left corner of the object. All vertices in the VR stylus were translated by $\mathbf{V}_{VEOriginToBoardOrigin}$ and then rotated around the board's origin by the reverse of the tapping board's orientation in the VE (the matrix $\mathbf{M}_{Reverse\_VEBoardOrient}$). Thus for a stylus vertex $\mathbf{V}_i$:

$$\mathbf{V}_{i\_tap} = \mathbf{M}_{Reverse\_VEBoardOrient}(\mathbf{V}_i - \mathbf{V}_{VEOriginToBoardOrigin}) \tag{A.15}$$

A series of line segment/plane intersection calculations quickly determined if edges of the VR stylus intersected the bounding box of the tapping board. Since the tapping board's reference frame is being used, a line segment intersected the bounding box if at least one of the end points was inside the range of the bounding box (from the point at xMin, yMin, zMin up to the point at xMax, yMax, zMax). Each hole in the bounding box was a smaller object (a C++ object of type SMHoleBoard) defined by a smaller bounding box used for detecting intersections:

$PointInBoundingBox(\mathbf{Pt}_i) =$

true if $(xMin < \mathbf{Pt}_i\_x < xMax) \wedge (yMin < \mathbf{Pt}_i\_y < yMax)$

$\wedge (zMin < \mathbf{Pt}_i\_z < zMax)$

This bounding box intersection method permits line segments that skewer the bounding box to go unnoticed. However, the VR stylus' line segments were smaller than any bounding box dimensions, there were a large number of line segments to detect any collision, and the end of most tapping motions were expected to be predominately

translations along the virtual world's y-axis. Thus, we do not believe any intersections were missed using this method, and intersection calculation speed was greatly facilitated.

If any line segment/bounding box intersections were found for a hole, line segment/triangle intersection calculations tested the intersection of stylus edges with triangles composing the image of the particular SMHoleBoard. This enables us to know the exact location of a tap while minimizing system lag (since line segment/triangle intersections were only calculated when absolutely necessary).

First, a test to see if the line segment intersects the plane defined by the triangle being tested. Calculating if a line segment intersects a plane uses the "point distance from plane" method described earlier. We find the distance from each end of the line segment to a plane. If one distance is positive and the other is negative, we know the segment (with end points $\mathbf{Pt}_1$ and $\mathbf{Pt}_2$) intersects the plane:

IntersectPlane =

> false if $(D(\mathbf{Pt}_1 > 0 \wedge D(\mathbf{Pt}_2) > 0) \vee (D(\mathbf{Pt}_1 < 0 \wedge D(\mathbf{Pt}_2) < 0)$

> true otherwise

Next, to find out if the line segment intersects a triangle, we find the point of intersection with the plane $\mathbf{Pt}_{intersect}$ and get vectors from each point in the triangle ($T1$, $T2$, & $T3$) to $\mathbf{Pt}_{intersect}$. The sum of the angles from a triangle points to $\mathbf{Pt}_{intersect}$ and back to different triangle point should sum up to 2π.

> $\mathbf{Pt}_{intersect} = ((D_{Pt1ToPlane}/D_{Pt1ToPt2}) \times \mathbf{V}_{Pt1ToPt2}) + \mathbf{Pt}_1$

> $SumAngles = \arccos(\mathbf{Pt}_{intersect}, \mathbf{Pt}_{T1})$

>> $+ \arccos(\mathbf{Pt}_{intersect}, \mathbf{Pt}_{T2})$

>> $+ \arccos(\mathbf{Pt}_{intersect}, \mathbf{Pt}_{T3})$

IntersectTriangle =

> false, if IntersectPlane = false,

> false, if SumAngles $\neq 2\pi$,

> true otherwise.

Appendix B

Head-Mounted Display Adjustment Look-Up Table

Headband to Eye Y-Position (cm)	# Clicks	Y Offset From BOP	Z Offset From BOP
-2.20	0	0.0	0.0
-2.40	1	-0.2	-0.1
-2.60	2	-0.4	-0.2
-2.80	3	-0.6	-0.3
-3.00	4	-0.8	-0.3
-3.20	5	-1.0	-0.35
-3.40	6	-1.2	-0.35
-3.60	7	-1.4	-0.35
-3.80	8	-1.6	-0.35
-4.00	9	-1.8	-0.35
-4.20	10	-2.0	-0.35
-4.50	11	-2.3	-0.35
-4.80	12	-2.6	-0.35

Table B.1: The HMD look-up table.

The above table determines how much the HMD needs to be adjusted (number of clicks) by using the eye position of the subject relative to the headband position $((Y_{HBToLEye} + Y_{HBToREye})/2)$. This value is collected directly using the Pinhole headband and rounded to the nearest entry in the table. The HMD set to zero clicks means

that the optical modules have the least amount of extension possible. This position is known as the Base Optical module Position (BOP). A ceiling/floor value is used if a *Headband to Eye Y-Position* value exceeds the upper or lower limit of the HMD's range. No subjects reached this limit during our experiment.

Y From BOP and *Z From BOP* columns were not used directly by the experimenter (they were used by the system) but were included in case that information was useful. These offsets are current optical module positions relative to the BOP.

Appendix C

Experiment Consent Form

THE UNIVERSITY OF BRITISH COLUMBIA

October 4, 2005

Department of Computer Science
201-2366 Main Mall
Vancouver, B.C., Canada V6T 1Z4

Tel: (604) 822-9289 Fax: (604) 822-5485

www.cs.ubc.ca

Consent Form:
Collaborative Visualization and Interaction in Ubiquitous Computing Environments

Principal Investigator:

Dr. Kellogg S. Booth, Professor, Department of Computer Science
Email: ksbooth@cs.ubc.ca Tel: (604) 822-8193

Co-Investigators:

Dr. John Dill, Professor, School of Engineering, Simon Fraser University
Email: dill@cs.sfu.ca   Tel: (604) 291-3574
Dr. Sidney Fels, Associate Professor, Department of Electrical and Computer Engineering
Email: ssfels@ece.ubc.ca   Tel: (604) 822-5338
Dr. Sheryl Staub-French, Assistant Professor, Department of Civil Engineering
Email: ssf@civil.ubc.ca   Tel: (604) 827-5118
Dr. Tamara Munzner, Assistant Professor, Department of Computer Science
Email: ssfels@ece.ubc.ca   Tel: (604) 827-5200
Dr. Brian Fisher, Adjunct Professor, Department of Computer Science
Email: fisher@cs.ubc.ca   Tel: (604) 822-8158
Dr. Lyn Bartram, Research Associate, Department of Computer Science
Email: lyn@cs.ubc.ca   Tel: (604) 879-4678
Dr. Barry Po, Postdoctoral Fellow, Department of Computer Science
Email: po@cs.ubc.ca   Tel: (604) 827-3993
Dr. Melanie Tory, Postdoctoral Fellow, Department of Computer Science
Email: melanie@cs.ubc.ca   Tel: (604) 822-2218
Anthony Tang, Ph.D. Student, Department of Electrical and Computer Engineering
Email: tonyt@ece.ubc.ca   Tel: (604) 822-5338
Colin Swindells, Ph.D. Student, Department of Computer Science
Email: swindell@cs.ubc.ca   Tel: (604) 822-2218
J. Karen Parker, Ph.D. Student, Department of Computer Science
Email: parker@cs.ubc.ca   Tel: (604) 822-2218
Regan Mandryk, Ph.D. Student, School of Computing Science, Simon Fraser University
Email: rlmandry@cs.sfu.ca   Tel: (604) 822-8851
Mani Golparvar Fard, Ph.D. Student, Department of Civil Engineering
Email: mgolparvar@civil.ubc.ca   Tel: (604) 827-5118
David Sprague, M.Sc. Student, Department of Computer Science
Email: dsprague@cs.ubc.ca   Tel: (604) 822-2218
Lior Berry, M.Sc. Student, Department of Computer Science
Email: berry@cs.ubc.ca   Tel: (604) 822-2218
Sherman Lai, M.Sc. Student, Department of Computer Science
Email: s.lai@fisheries.ubc.ca   Tel: (604) 822-2218
Ima Hajshirmohammadi, M.A.Sc. Student, School of Engineering, Simon Fraser University
Email: ihajshir@sfu.ca   Tel: (604) 291-4371
Alex Merritt, Undergraduate Research Assistant, Department of Computer Science
Email: lecaran@cs.ubc.ca   Tel: (604) 822-2218

THE UNIVERSITY OF BRITISH COLUMBIA

This study is intended to show how collaborative computing applications are used in practice. You will be asked to use a prototype computer application to accomplish a computing task. Your participation will help us assess the usability of this prototype application. Computer images such as schematic diagrams or rendered blueprints will be presented on a computer display and you will be asked to manipulate these images through the use of an input device. We will record your performance and analyze how the application is used.

Audio and video recordings may be made of your performance. You may also be asked to complete a questionnaire that will help us assess your experience with computer technology and your impressions of the prototype application. By consenting to participate in this study, you also consent to participation in the audio and video recordings, and the completion of the questionnaire form. Your total participation time will be no longer than about two hours. In exchange for your participation, you will be provided with access to research computing applications and tools to assist you while you are being observed.

We will ensure that all recorded data are accessible only by project investigators and are kept secure in a locked faculty office. All data from individual participants will be coded so that your anonymity will be protected in any publicly available reports, papers, and presentations that result from this work.

We intend for your experience in this study to be pleasant and stress-free. If you are uncomfortable, or are unhappy participating in this study, you are free to withdraw at any time, without any repercussions to you whatsoever. We would be pleased to explain to you the purpose and methods used in this study in academic detail after your participation has concluded, and to furnish you with our results when they become available.

This research study is funded by the operating and strategic grant programs of the Natural Sciences and Engineering Research Council of Canada (NSERC). Portions of this research will be used for graduate theses. If you have any questions or desire further information with respect to this study, you may contact Dr. Kellogg S. Booth or one of his associates at (604) 822-8193. If you have any concerns about your rights or treatment in this or any other UBC experiment, you may contact the Research Subject Information Line in the UBC Office of Research Services at (604) 822-8598. You have been given a copy of this consent form for your records. You do not waive any legal rights by signing this consent form.

I consent to participate in this study under the above conditions:

Name (Please Print): _____

Signature: _____

Witness: _____

Date: _____

Reference Number: UbiqRef002; Version 1.1 / October 4, 2005
Page 2 of 2