DESIGN AND EVALUATION OF METHODS FOR MOTOR EXPLORATION IN LARGE VIRTUAL ENVIRONMENTS WITH HEAD-MOUNTED DISPLAY TECHNOLOGY

By

Betsy Williams

Dissertation
Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements
for the degree of
DOCTOR OF PHILOSOPHY
in
Computer Science
August, 2008
Nashville, Tennessee

Approved:
Professor Bobby Bodenheimer
Professor Gautam Biswas
Professor Benoit Dawant
Professor John Rieser
Professor Bill Thompson

I would like to dedicate this thesis to my beloved parents and grandparents for their continual love and support. I also dedicate this thesis to my advisor, Bobby Bodenheimer, for his guidance, patience, and wisdom.

ACKNOWLEDGMENTS

My foremost thanks goes to my advisor Bobby Bodenheimer. Without him this dissertation would not be possible. I thank him for his patience and encouragement that have carried me through difficult times, and his thoughtful insights and suggestions that have helped shape my research skills. Through the years he has helped, guided, and influenced with amazing wisdom. I am forever grateful for everything that I learned from you, Bobby.

I would like to thank John Rieser for his lively psychology discussions. I can always count on him to discuss the smallest details of a problem. I thank him for being a great teacher, mentor and role model.

I would like to thank Benoit Dawant, Gautam Biswas and Bill Thompson, for their extremely valuable guidance and advice on this dissertation. I am grateful to Benoit and Gautam who helped me start my graduate student life at Vanderbilt. It has been a pleasure to have Bill Thompson on my committee. His visionary thoughts and energetic working style have been very valuable to this dissertation.

Last, I would like to thank the members of the Learning In Virtual Environments (LIVE) lab specifically, Gayathri Narasimham, Bjoern Rump, Tim McNamara, Bernhard Riecke, and Tom Carr. I have been lucky to work with such an extraordinary group. I specifically thank Gayathri for her help, encouragement, and collaboration.

Last of all, I thank my friends and family for supporting this journey.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

This thesis presents the development and evaluation of a system that allows people to explore a large virtual environment with a head-mounted display (HMD) when the size of the surrounding physical space is small. More specifically, this thesis focuses on exploring an HMD-based virtual environment by physically walking, i.e., bipedal locomotion. Bipedal locomotion is a highly effective method for learning the locations of things when exploring virtual environments, and seems to result in better spatial orientation than other locomotor interfaces such as joysticks [Williams et al. 2006a]. Bipedal locomotion within a virtual environment is easily accomplished as long as the physical space housing the tracking system and HMD are roughly the same size as the virtual environment. The issue becomes how to fit physical bipedal locomotion in a large virtual environment into a much smaller physical space while preserving a user's spatial orientation.

Virtual reality provides people with opportunities to experience places and situations remote from their actual physical surroundings. They potentially allow people to learn about an environment which, for reasons of time, distance, expense, and safety, would not otherwise be available. Virtual reality systems could have a huge impact in education, entertainment, medicine, architecture, and training, but they are not widely used because of their expense and delicacy. However, HMD technology may become readily available to the public within the next several years. Other immersive virtual technologies, such as virtual caves, are less likely to achieve commodity status since they often involve greater expense in the form of large screens, projectors, and a locomotion device such as a treadmill or bicycle that allows a user to move about the environment. Since HMD systems hold the promise of being readily available to the public, constraints of such a system need to be identified and addressed. In particular, to test the effectiveness of an HMD-based virtual environment, the perceptual constraints need to be understood.

People explore spaces in the real world by physically rotating and translating their perspectives as they examine the environment. Thus, this thesis seeks to leverage the natural ability of people to maintain spatial awareness of an HMD–based virtual environment when rotation and translation are provided by bipedal locomotion. Although there is no problem rotating by physically turning in a small physical space to explore a large virtual space, there is a problem translating. This problem occurs when trying to fit long distances walked through virtual space into a small physical space. Therefore, devising methods whereby large virtual environments can be explored while preserving a user's spatial representation of the environment is an important problem.

Thus, this work is important because it seeks to develop an effective system that would allow a person to seamlessly explore large virtual environments. The system envisioned here could be based in an office or small lab. In particular, if immersive virtual environments are to realize their potential as commodity-level components, a perceptually accurate interface that allows locomotion through them within the constraints of everyday space must be developed.

The ultimate goal of a virtual system is to immerse the user in a computer–generated environment. There are interesting issues associated with the user interacting or interfacing with that environment. Ideally, performing tasks or moving about in a virtual environment would be identical to the real physical environment. For example, when immersed in a virtual environment, a user might feel a virtual object, walk and feel the terrain, smell things, and hear accurate sounds. To train a person to throw a bowl on a pottery wheel using a virtual environment, the person would need to feel as if they were molding and shaping the clay with their hands. They would need to feel mass and act appropriately, and the virtual object would have to respond appropriately. Obviously, virtual environment technology is not yet at this stage.

This thesis limits the discussion to tracking systems. Such systems track the user's position in the physical environment and allow the computer to update the graphics of the corresponding position in the virtual environment. An HMD presents the user with a stereo view of this virtual environment.

Interestingly, human-computer interaction (HCI) issues associated with a virtual system are quite different than traditional 2-D or 3-D HCI [Lok and Hodges 2004]. First, the user experiences the virtual environment from a first person point of view. Moreover, the interface seeks to present a high level of fidelity between physical user action and the virtual interpretation of that action. Thus, the virtual system seeks to mimic how users interact with the real physical environment. This type of human-computer interaction is better examined using methods from cognitive science than most traditional user interface design problems. There are many variables that can be manipulated when interfacing into a virtual environment. Some are consistent with our experience in the real world while some are less so. For example, as Lok and Hodges point out, some virtual actions that are possible in the virtual environment have no correlation to a real action. We explore this idea later in this thesis by exploring different modalities of walking and changing the normal viewing eyeheight from which a user experiences the virtual environment.

Virtual systems that satisfy the high fidelity interactions can be important tools for learning and training in virtual environment. These computer generated environments allow a user to experience places that would be expensive, dangerous or infeasible in the real world. Virtual flight simulators are an example of a good interaction of real action and virtual experience and are commonly used to train pilots. This thesis presents work that would allow a user to experience a virtual environment by physically walking, thus looking specifically at the tracking interface. My work seeks to create virtual environments that are general learning environments while addressing a limitation of current virtual technology. The potential uses of such environments are limitless, and some include testing evacuation plans before a structure is built, experiencing historical sites such as the Pompeii, assessing search and rescue efforts of firefighters, and gaming environments.

I.1 Contributions

This thesis makes the following contributions in addressing the problem of exploring large virtual environments in small physical spaces:

1. **Determines the functional similarity in spatial reasoning between real and virtual environments.** This work analyzes the similarity in people's reasoning and learning about spatial representations of physical environments and of virtual environments viewed on an HMD, a phenomenon called functional similarity ([Williams et al. 2007b], Chapter III). The functional similarity work of Chapter III lays the groundwork for understanding if it is reasonable to build virtual environments and expect people to be able to navigate and orient themselves in the virtual world just as people do in the real world. Given that functional similarity exists between the two environments, this thesis then looks at how to exploit our own locomotion to explore virtual spaces larger than the tracking limits of the virtual system.

2. **Shows that scaling the translational gain of walking is a viable method of exploring a large virtual environment.** Chapter IV investigates increasing the translational gain of walking (where one step forward in physical space carries one several steps forward in virtual space) as a viable method to explore a large virtual environment [Williams et al. 2006a]. Two experiments presented in this Chapter show that the translational gain of bipedal walking can be scaled, and that this type of locomotion is a more efficient interface than using a joystick. However, these experiments limit the scale of translational gain to a factor of ten, since head movements and other small movements become distracting at higher gains.

3. **Examines the limits to which translational gain can be scaled while maintaining spatial awareness.** To scale gain higher than ten, it becomes necessary to investigate

ways of minimizing distracting motions. Chapter V finds such a method and shows that using a nonlinear scaling method is significantly superior to scaling gain linearly. Additionally, this work examines how high we can expect to scale gain in a virtual environment and still maintain a reasonable spatial awareness. More specifically, results of the experiments in Chapter V show that people can maintain good spatial awareness with translational gains up to 50 using a technique also presented in this chapter to minimize the distracting effect of head movements.

4. **Develops an effective method of navigating virtual space that accommodates the limits of physical space.** Inevitably, the physical limits of the tracking system will be reached no matter how high the gain is scaled. Therefore, Chapter VI investigates three different methods to "reset" users when they reach the end of their physical space by changing their location in physical space while their location in the virtual environment is ideally the same before and after resetting [Williams et al. 2006b; Williams et al. 2007a]. In other words, resetting involves manipulating optical flow in such a way that allows users to move away from a physical obstruction such as a wall while experiencing a continuous sense of their location in the virtual environment. This thesis develops a method that produces a minimal amount of disruption to a user's sense of presence in the virtual environment.

5. **Evaluates the advantages of manipulating user eyeheight at various rates of translational gain.** In addition to scaling gain, it may also be advantageous to scale eyeheight as well. More specifically, Chapter V investigates whether a person's spatial representation is improved when eyeheight is increased while locomoting through a virtual world at high rates of translational gain. If eyeheight is increased, more of a map-like overview of the terrain will be experienced. This thesis finds no significant advantage with respect to the user's spatial orientation when eyeheight is scaled.

I.2 Significance

The main contribution of this work to the field of computer science is to develop a tested interface using an HMD system that allows people to explore a large virtual environment in a small physical space. This thesis contributes novel engineering techniques to solving this human-computer interface problem. This system presents visual information based on user action by leveraging the natural ability of humans to spatially update using physical locomotion. This work is important since it seeks to make technology, that is, HMD technology, more useable and accessible.

The work presented in this thesis presents an interdisciplinary approach, using both computer science and psychology, to engineer a system that allows users to effectively explore a large environment with an HMD. Computer science is a broad discipline that systematically studies computing systems and computation involving the development of algorithms, tools, and methods for using a computer. Although the main contribution of this work uses computer science techniques to implement design prototypes, methods from psychology are also employed to evaluate and leverage the design of the system. Figure I.1 shows the relationship between computer science and perceptual psychology used in this research. First, computer graphics is used to investigate perception in the system. This involves developing virtual environments suitable for perceptual experiments, designing and implementing algorithms to explore virtual environments that leverage people's natural perceptual affordances, and systematically evaluating these solutions. Conversely, in this research, visual perception must also use computer graphics to identify important perceptual aspects of HMD technology. The presentation of computer graphics in the HMD system is validated using perceptual and learning methods borrowed from psychology. These psychology studies are leveraged to form conjectures about computer science solutions that could afford a more compelling virtual reality experience.

The high level interface created and tested in this work uses a combination of methods to achieve its goal. First, this work systematically scales the translational gain of walking

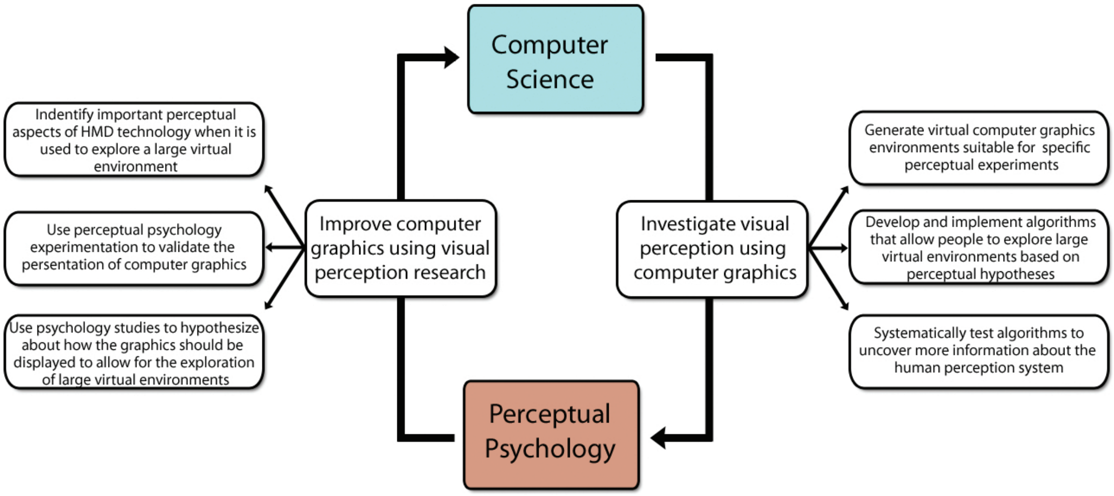

Figure I.1: This figure shows the cyclic process involved in virtual reality research.

(Chapters IV and V). When gain is scaled to larger amounts it becomes necessary to minimize these small head movements, and an appropriate method is developed and tested. This thesis then examines several parameters of the interface, seeking to better preserve a user's spatial orientation. We determine a threshold where people's spatial awareness begins to deteriorate as translational gain is scaled. And, we examine the modality of eyeheight and find what trade-off exists for spatial orientation while exploring a virtual environment at different eyeheights and different translational gains.

Scaling the translational gain of walking does not completely solve the problem, however, since the physical limits of the tracking system will inevitably be reached no matter how high gain is scaled. Therefore, we develop a method to remedy this that we called "resetting." We evaluate three plausible engineering methods to "reset" users when they reach the end of their physical space by changing their location in physical space while their location in the virtual environment is the same before and after resetting (Chapter VI). We categorize the methods that produce minimal disruption to a users' sense of locomotion in the virtual environment. We also show the number of resets that a person can undergo in a virtual environment and still maintain a reasonable spatial awareness. These findings allow others to create virtual reality systems based on parameters clearly defined in this thesis.

A contribution of this work is to extend virtual environment research beyond emphasizing low-latency photorealism. This thesis is aimed at designing an interface to a virtual world that allows users to interact with the graphics in a way that promotes user acceptance and usability of the system. Technology alone does not create user acceptance and usability, but it is the experience of the user which is key. Typically a virtual world is modeled after the real world, where experiencing the virtual world is similar to experiencing the real world. When users experience virtual environments, they are not bound to only common experiences of the real world. While the photorealism of the virtual environment may play an important role in the effectiveness of a virtual environment, it seems that the human-computer interaction model of the system is equally important. The software, the interactive devices, and the platform used are all part of the design of virtual environments. The most important thing about the virtual environment is the user experiencing the virtual environment. Since human locomotion can drive the interaction between human and computer, it becomes important to look at ways of engineering environments where technology is tailored to fit the physical and psychological needs of the participant. This thesis manipulates human translations and rotations so that the physical locomotion fits into the confines of a small physical space– allowing the user to get locomotive feedback while accomplishing the goal of seamlessly exploring a large virtual environment in a small physical space. It also examines the exploration of a virtual environment when a user's eyeheight is manipulated, so that the graphics in the system are viewed as if the users is taller than he or she actually is. That this manipulation could aid in human spatial awareness is an interesting way of incorporating human locomotion with computer graphics.

In the past, making a significant computer science contribution to virtual reality research involved developing algorithms that render computer graphics faster and more effectively so that they could be presented to a user in realtime. We are nearing the threshold of what hardware needs to do in realtime and should to look at other means of exploiting virtual reality technology besides improving the computational limitations. This work incorporates psychology with computer graphics so that psychology suggests and evaluates engineering improvements to computer graphics. In this manner, I focus my research more on what the user experiences and how this experience can be improved. Thus, the significant contributions of this work are closely linked to both computer science and psychology.

Thus, this work aims to contribute cognitive finding as well. Chapter III examines how people understand space in virtual reality and its relationship to how they understand space in the real world. The finding is significant because it shows that people maintain spatial awareness in both environments in a functionally similar manner. The idea that the translational gain of walking can be scaled was based on the psychological studies that show that people can recalibrate translation [Rieser et al. 1995] in the physical environment. This prior work used a much smaller scaling factor. This work does not look specifically at whether subjects can recalibrate to a new translational gains but investigates whether they can adapt and maintain spatial awareness at higher translational gains. Therefore, a significant contribution of this work is to show that translational gain can be scaled much higher, and that spatial orientation can be maintained with the aid of nonlinear scaling techniques. The resetting techniques discussed in Chapter VI manipulate rotation and translation, and test people's ability to maintain spatial orientation by relying on visual cues. When the gain of walking is manipulated and these resetting techniques are employed, locomotion is manipulated, and the human updating system is fooled. Therefore, an interesting psychological contribution is that this work shows that even if locomotion is manipulated it does help aid in spatial awareness. Finally, another significant psychological contribution of this work examines spatial awareness as a function of user eyeheight, or the height from which the virtual world is viewed.

I.3 Overview

The remainder of this thesis is organized as follows. The next chapter, Chapter II, provides background information and relates previous work to the present work. Chapter III examines the similarities and differences in spatial reasoning in real and virtual environments. Chapters IV, V, and VI then discuss how to develop a system to explore a virtual environment larger than the physical limits of the tracking system. Specifically, Chapters IV and V looks at scaling the translational gain of walking, while Chapter VI examines methods of resetting when people reach the limits of the explorable tracking space. Then, Chapter VII concludes this work and explores future directions.

CHAPTER II

BACKGROUND AND RELATED WORK

II.1 What is a virtual environment?

Virtual reality is a term used to describe the interaction between a human via sensory input with a computer generated environment. Virtual reality is the concept of perceiving a synthetic environment just like a real, physical environment. Equivalent terms such as "artificial reality" and "cyberspace" also refer this experience. Generally, a virtual reality environment is experienced visually– viewed on a computer screen or through a special stereoscopic display. However, some virtual reality systems also employ audio cues, [Lokki and Grohn 2005], and with an increase in technology, it may be possible to incorporate taste, touch, and smell [Sherman and Craig 2003].

In a virtual reality system, users interact with a virtual environment using some type of human computer interface (HCI). This human-computer interaction involves two steps. First, the computer presents information to the user via computer renderings of the virtual environment. Then, the human communicates to the computer by using some sort of device such as a keyboard, joystick, dataglove, treadmill, or tracking system. The range of communication devices from human to computer is more varied. Both the sophistication and diversity of these devices will continue to increase in the future [Sherman and Craig 2003]. Finding the best way to connect the senses to a virtual environment is a difficult problem and is commonly referred to as the "human factors" problem [Sherman and Craig 2003].

Virtual reality systems are different from other computer-based systems because they are thought to induce a certain feeling of immersion or presence. Heim [1998] summarizes virtual reality with "three I's" : immersion, interactivity, and information intensity. Immersion is a psychological effect that arises from "devices that isolate the senses sufficiently enough to make a person feel transported to another place" [Heim 1998]. Immersion is

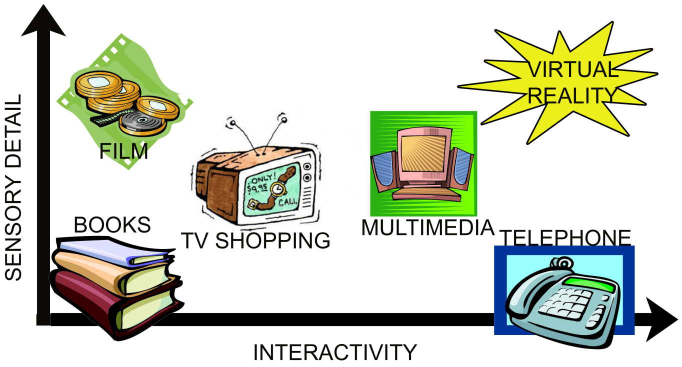

Figure II.1: This figure shows the two step process of the human computer interaction in a virtual reality system. The computer presents information to the user using computer generated graphics, and the human communicates to the computer by physical action such as walking or moving a joystick.

difficult to define because, it is a complicated and elusive phenomenon. Witmer and Singer [1998] define immersion as "a psychological state characterized by perceiving oneself to be enveloped by, included in, and interacting with an environment that provides a continuous stream of stimuli and experiences." They suggest that factors that affect immersion include "isolation from the physical environment, perception of self-inclusion in the virtual environment, natural modes of interaction and control, and the perception of self-movement." Heim explains that there are different degrees of immersion analogous to the different levels of involvement various media deliver in regards to "sensory detail" and "amount of interactivity" as shown in Figure II.1. Interactivity describes the computer's ability to deliver real-time images of a user's change in position and orientation in a 3–D environment based upon user input. Therefore, the computational speed of the computer must be fast enough to convert data from a person's sensory input to the virtual environment and vice-versa. Finally, "information intensity" is the idea that a virtual world can offer qualities such as "telepresence and artificial entities that show a certain degree of intelligent behavior" [Heim 1998]. Information intensity suggests that virtual reality may provide the user with different quantities and qualities of information. No virtual reality system is suitable for all applications, thus, virtual reality systems can be specialized in the information they present. Telepresence is an example of an information specific system where a participant is able to interact with a real physical environment remote from his or her current location, (for more information see [Sherman and Craig 2003]).

Virtual reality provides a new communication medium for human computer interaction [Ellis 1994]. Moreover, the synthetic nature of virtual reality allows for the incorporation of visual, auditory, and haptic interaction modes not possible in real physical environments. Virtual environments have the potential to make a broad impact on education, entertainment, medicine, architecture, training, etc. [Bricken 1991]. Virtual reality was developed to allow people to easily deal with information. An example of its near 40 year success in learning and task performance has been shown in the training Air Force personnel [Furness 1978].

II.2 Visual Experience of Virtual Reality

This work deals with the most popular means of experiencing a virtual environment–visually. Current technology does not allow us to develop a virtual reality system that creates photorealistic virtual images in real time from sensory input. Virtual reality emerged from research on interactive graphics and vehicle simulation in the late 1960s and early 1970s. Due to technology advances making real-time graphics feasible, in 1980s and 1990s "virtual reality" became a part of our everyday vocabulary.

To make the 2–D images displayed in virtual environments, 3–D object are rendered. These images are presented to the user in a continuous manner, and must be fast enough to provide the illusion of seamless motion. The rate at which these images are displayed is referred to as the temporal resolution or frame rate. There is also a graphics latency tolerance that a virtual system must maintain to preserve the illusion of real-time interaction. Image resolution must also be taken into consideration when developing a realistic virtual environment.

People naturally develop a sense of the 3–D world around them without consciously thinking about it. It is not well understood how the brain processes such information. In the real physical world, humans perceive depth using stereopsis or binocular disparity, monocular or pictorial cues, and motion depth cues. Monoscopic or pictorial depth cues are those obtained from a static image, such as that of a photograph and involve such things as interposition or occlusion, shading, linear perspective, relative size, height in the visual field, atmospheric effects, texture, etc. Stereoscopic depth cues depend on parallax, or the difference in position of objects viewed from the left and right eye. Motion depth cues are derived from parallax created from self motion, object motion, or a combination of the two. Depth cues will discussed under a few different modalities of visual input in virtual systems (such as desktop screens, head-mounted displays (HMDs), etc.). No one cue dominates depth perception in every scenario, and no one cue is necessary to correctly judge depth [Cutting and Vishton 1995]. For a detailed discussion on depth cues see [Goldstein 2006] and [Cutting and Vishton 1995].

Portraying depth in the virtual world is achieved using a number of real world cues. Virtual worlds usually mimic the real world, but because of computation costs, tradeoffs between realism and interactivity capacity must be considered by the designer. For example, the depth-accuracy of a virtual manufacturing or medical application where subjects grasp objects may be more important than a virtual architectural walk-through. Most of the current 3–D animation films use a computationally expensive rendering technique called ray tracing to obtain realistic images with complex shadows, reflections, refractions, and shadows giving the user many depth cues. To achieve a reasonable frame rate, such a technique is currently not feasible in complex virtual environments. However, the graphics of most virtual environments contain such depth cues as occlusion, size, linear perspective, motion parallax, and some sort of shading depending on the computing power. This shading usually includes flat shading, texture mapping, specular highlighting, and sometimes primitive shadowing. Texture mapping is a technique most commonly used in virtual environments as a computationally efficient way of adding realism. Texture mapping is the process of mapping an image onto a surface [Foley et al. 1990]. For example, to make a wooden desk look more realistic, a digital photograph can be mapped by the renderer onto the polygons of the desk. The texture can be repeated on the surface so that the scale and size of the wooden texture mimics that of a real desk. Texture maps are also used as backdrops for areas in virtual environments not intended for exploration such as a sky image or a group of buildings at a distance.

II.3 Types of Visual Displays

Current research on virtual reality is vast and rapidly changing. Visual perception of virtual environments is the most popular and perhaps the most important method of acquiring information about a virtual space. Therefore, this section focuses on presenting various types of visual display systems to get a sense of the range of technical capabilities.

II.3.1 Desktop

In desktop virtual reality, a user experiences a virtual environment by viewing high quality graphs on a standard computer monitor. These graphics are generally rendered by a higher-end workstation. The main limitation of such a system is the locomotion interface. Moving around the virtual environment is accomplished by using some sort of haptic device such as a keyboard, joystick or mouse. Using some sort of haptic device or passive interface to navigate the environment provides no vestibular feedback which has a direct affect on spatial orientation and navigation. Another limitation is that users may be viewing the real and virtual environment simultaneously. This dual representation may be a cognitive burden which could diminish the effectiveness of the system. Because of these limitations, the immersiveness of desktop virtual reality is generally thought of as considerably limited [Tarr and Warren 2002].

II.3.2 Caves

The Cave Automatic Virtual Environment or CAVE, which was introduced in 1992 by [Cruz-Neira et al. 1992], is an immersive virtual environment that involves using projectors to display images on three, four, five, or six walls of a room-sized cube. The locomotion interface in this environment is typically either a treadmill [Mohler et al. 2004], bicycle [Plumert et al. 2004], or some sort of haptic device. Using a haptic device lacks vestibular feedback which is generally accepted to enhance a user's spatial orientation [Chance et al. 1998; Witmer et al. 1996]. Most CAVE-based treadmill systems involve a single walking direction treadmill. Such a system lacks free exploration. Thus, some systems allow users to change their orientation in the virtual environment by rotating their torso while their feet remain in the same orientation on the treadmill [Vijayakar and Hollerbach 2002]. Thus, the action of torso rotation and walking on the treadmill allows navigation along a curved path. A few omnidirectional treadmill systems,[Darken et al. 1997; Iwata 1999; Schwaiger et al. 2007], have been developed that cancel the user's displacement and allow them to walk in any direction. The CyberWalk, [Schwaiger et al. 2007], platform is made of 25 conventional treadmills which are all chained together and can move in one direction while the individual treadmills move at right angles relative to the direction the chain is moving. This gives the user a 4.5m by 4.5m area to walk or jog on and is the first omnidirectional treadmill that allows for near natural walking. However, these systems are expensive to construct and are not robust. Another short coming of the CAVE system is the presence of an open area or place lacking visual feedback of the virtual environment. In other words, the real world can be seen while viewing the virtual world. However, a 5-sided CAVE is quite immersive. In general, CAVE systems suffer from being delicate and also are considerably more expensive than the HMD-based virtual systems discussed in the next section.

II.3.3 HMD

A head head-mounted display is a helmet mounted with two small screens that allow the user to view the virtual environment. The HMD is attached to a workstation that renders different graphics for both eyes simultaneously enabling stereo or motion depth cues. An HMD typically has a baffle that blocks out any view of the outside world. An HMD system can include position and orientation tracking to update the users' position and orientation in virtual space as they navigate in the real world.

One of the drawbacks of HMD technology is the limited field of view or viewing range of the virtual environment. The resolution of the two displays inside the HMD is also limited. Currently, a $1,500 Emagin Z800 HMD has a resolution of 800x600 per display and a field of view of 40° diagonally. The weight of the HMD may be distracting, as well as the tether which connects the HMD to the graphics-rendering workstation. Therefore, when a person is moving around while wearing the HMD, they must be careful not to trip over the tether. Despite the limitations, the advantage of these systems over CAVE systems is that they more robust, much less expensive, and most importantly the locomotion interface to this environment is your own. As Tarr et al. [2002] point out, with a 480x600 resolution HMD, users "rarely, if ever, have any sense of where they are in the physical room and respond appropriately when faced with 50-foot cliffs, spinning tunnels and carousels."

II.4 Spatial Orientation

This work relies on the investigation of humans' spatial orientation in virtual environments to build a virtual system. Spatial orientation refers to the natural ability of humans to maintain their body orientation and position relative to the surrounding environment. Spatial orientation relies heavily on visual information and whole-body information while moving in an environment [Wartenberg et al. 1998]. Spatial orientation refers to the natural ability of humans to maintain their body orientation and position relative to the surrounding environment.

After exploring and learning about physical environments, people can judge the perspectives available at various locations within the familiar place, even if the perspective is one that they have never directly experienced [Rieser et al. 1994]. Consider three conditions that influence the relative ease of such judgments. First, judging new perspectives is done more effectively when people physically locomote as if moving to the new perspective. This situation occurs when people close their eyes and locomote relative to their actual remembered surroundings [Loomis et al. 1999; Rieser et al. 1986]. In addition, when people stand in one place and pretend to be somewhere else, they are able to imagine new perspectives [Rieser et al. 1994; May 1996b; May 2004]. Second, people find it easier to judge changes in perspective when the geometry of the change is a translation rather than a rotation [Rieser 1989; Presson and Montello 1994; Philbeck et al. 2001]. Third, the difficulty of making a perspective change is a monotonic function of the degree of disparity between the original perspective and the to-be-judged perspective [Rieser 1989; May 2004]. By *disparity* we mean the difference in direction between targets from the new, to-be-judged point of observation and the one from which they were originally learned, a measure discussed further later in this section. In the related two experiments in Chapter III, subjects were asked to judge new perspectives across variations in locomotion mode (physical or imagined), geometry of perspective change (translation or rotation), and amount of disparity.

First, to investigate functional similarities in spatial orientation in virtual environments and physical environments, Experiment 1 of Chapter III examines imagined and physical rotation in both environments. Many studies [Easton and Sholl 1995; Farrell and Robertson 1998; May 1996b; Presson and Montello 1994; Rieser 1989; Rieser et al. 1986; Wraga 2003] have shown that updating spatial orientation is much harder with imagined movement than physical locomotion. In Rieser [1989], subjects were asked to point to a target after imagining facing an object and after physical locomotion to an object. The study showed that performance was slower for imagined rotations than physical rotations, and the response latency for the imagination condition increased as a function of the angle from the participant's facing direction to the imagined location. A possible explanation is that imagined rotations lack proprioceptive feedback. In Experiment 1 similar responses in the virtual environment are measured and compared to the physical environment.

Experiment 2 of Chapter III further investigates functional similarity by comparing two different imagined geometric movements, translation, i.e., changes in self location keeping the current facing direction, and rotation, i.e., changes in current facing direction while retaining the same location. Subjects are tested in both the physical and virtual worlds to test for functional similarity. A number of studies, [Easton and Sholl 1995; May 2004; Presson and Montello 1994; Rieser 1989], have shown imagined translations are faster and more accurate than imagined rotations. Furthermore, many have shown that response times and errors increase as a function of the angle of imagined rotation [Easton and Sholl 1995; Rieser 1989; Wraga 2003]. In these experiments, translations and rotations are decoupled, in contrast to other work, e.g., Klatzky et al. [1998].

The experiments on functional similarity in Chapter III use a spatial orientation task where subjects turn to face a direction, similar to the pointing task of Rieser [1989]. More specifically, the assessment of spatial knowledge is based upon turning errors and latencies in tasks where subjects are asked to turn and face an object the location of which they had already learned. Other comparisons of spatial updating and pointing in virtual environments have been done, although none have replicated the conditions of the present work. Chance et al. [1998] report that for subjects walking through a virtual maze, physical translation and rotation allowed subjects to update better than physical translation and joystick rotation. Klatzky et al. [1998] report that optic flow without locomotion in an HMD was not sufficient to induce spatial updating for turn responses, although Riecke et al. [2005] report the opposite result for large field-of-views. May and Klatzky [2000] find a functional similarity between real and virtual environments in that in both environments irrelevant movement had greater effects on a path integration task than verbal or cognitive distractions. Wraga et al. [2004] also studied spatial updating in virtual displays and report

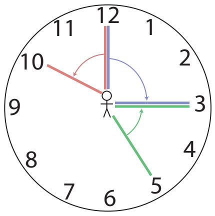

Figure II.2: An example of rotational disparity. The subject is facing the 12. The instruction "turn to face the 3 as if facing the 5" is given. The subject imagines facing the 5 while physically facing the 12. The subject must imagine the clock rotated such that the 5 is directly in front and then turn to face the 3 from this new perspective. Thus, to give the correct turn response to this example, the subject must turn to the left 60° as shown in red. The disparity is the difference between the real facing direction (12) and the target (3), 90° (blue), and the imagined facing direction (5) and target (3), $-60°$ (green). Thus, the disparity is $90 - (-60) = 150°$. The disparity is also equal to the difference between the final and facing position (10) and the target (3).

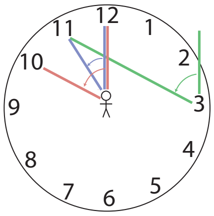

Figure II.3: An example of translational disparity. The subject is facing the 12 when the instruction "turn and face the 11 as if standing at the 3" is given. The correct motion is to turn to face the 10 (shown in red). The disparity is the difference between the real facing direction (12) and the target (11), -30° (shown in blue), and the imagined facing direction (3) and target (11), -60° (shown in green). Thus the disparity is $-30 - (-60) = 30°$. This quantity is also the same as the difference between the new facing direction (10) and the target (11).

that active rotation, in which the subjects rotate themselves, has advantages over passive rotation, in which the subjects were rotated by the experimenter. Note that they break down locomotion in two ways that this work does not. In Experiment 1 of Chapter III, subjects were passively rotated, while in the second, subjects actively rotated. However, the advantage found by Wraga et al. [2004] was small and we did not pursue that classification of rotation further. Waller et al. [2004] examined directional knowledge in the real world and virtual environments and also found a functional similarity between real and immersive HMD environments. Thus, the prevailing view in the literature seems to be that there are similarities between physical and virtual environments, and that locomotion in a virtual environment can help spatial updating, two views that we explore further in this thesis.

May [2004] argues that response latencies and errors due to perspective change are related to the conflict between the two perspectives. Similar research, [Brockmole and Wang 2003; May 1996b; Newcombe and Huttenlocher 2000; Wang 2005a] has supported this interference hypothesis. This effect is a function of a quantity called disparity, which is illustrated in Figures II.2 and II.3. For the purposes of the experiments presented in Chapter III, disparity is defined as the difference in angle when turning to face a given target relative to one's actual position in the physical or virtual room from the angle to-be-turned given the to-be-imagined facing direction (the location where the object is thought to be). In Figure II.2, the subject is facing the 12, and then imagines rotating to the 5. Now the subject is asked to point to the 3. When the subject is facing the 12, the 3 is on his right; when the subject is facing the 5, the 3 is on his left. These two relationships will interfere with each other as the subject makes a decision about the correct location of the 3 facing the 5. The effects of this interference increase as the magnitude of this difference, the disparity, increases. In this dissertation, we use the term "response disparity" or simply "disparity" to mean the "object direction disparity" as defined by May [2004]. For rotations, the amount of disparity is equal to the angle of imagined rotation as shown in Figure II.2. For translations, the amount of disparity is equal to the difference between the angle from the facing direction to target object, and the angle from the imagined source object location to target object as shown in Figure II.3.

Spatial orientation is also accessed in experiments presented in Chapters IV, V, and VI. However, the spatial orientation tasks are less complicated since we make the assumption from the results of Chapter III that people reason about space in the real and virtual worlds in a functionally similar way. Thus, in these three chapters, spatial knowledge is assessed by measuring errors and latencies in tasks where subjects turn to face a remembered object from their position in the virtual environments. Since there is no imagined facing direction in these type experiments, disparity is defined as the difference between the actual facing direction and the direction needed to face the target.

This thesis discusses aspects of the relation between perception, representation, and action. More specifically, we compare the spatial perception of rotations and translations in the virtual environment and the physical environment. We also closely examine the ability of people to perceive and represent space while actively locomoting through a large virtual environment using a limited amount of physical space. Some investigation of the relation between perception and action in virtual environments has been conducted, e.g., [Mohler et al. 2004; Mohler et al. 2007a; Kay and Warren, Jr. 2001]. In particular, much recent work [Loomis and Knapp 2003; Thompson et al. 2004; Willemsen and Gooch 2002; Witmer and Sadowski 1998], has studied the issue of the similarities and difference in distance estimation between real and virtual environments. This work has found that subjects underestimate distances in virtual environments. The precise reasons for this are not known, but several factors have been examined. There have been different empirical findings on how field of view in an HMD leads to an underestimation of distance. Wu et al. [2004] show that vertical field of view, FOV, of 21° or less leads to an underestimation of distance. Knapp and Loomis [2004] found that a reduced vertical FOV similar to that of an HMD has no influence in the real environment. The weight of the HMD itself may also cause problems with distance perception [Willemsen et al. 2004]. Thompson et al. [2004] show that distance perception in real and virtual environments is not due to the lack of realistic graphics. While these distance discrepancies exist in the virtual environment seen though the HMD, Plumert et al. [2005] found that time to walk estimates in real environment and virtual large-screen immersive display environment were highly similar. Oman et al. [2002] tested the ability of subjects to learn objects' spatial relationship and to predict their location as their bodies were specified in different 3D orientations. They found that body positions with respect to gravity had a minor but significant effect on locating objects, and that performances in the real world were functionally equivalent to those in the virtual.

The immersive qualities of virtual environments are getting better and ultimately we expect there to be no significant differences in operating on knowledge gained by exploring and learning about physical environments and their virtual renderings (*immersion* is a concept difficult to define, but see [Witmer and Sadowski 1998]). However, current HMD technology is limited in expressing rich visual representations. As mentioned above, several groups have reported inaccuracies of judging distances in immersive environments [Loomis and Knapp 2003; Thompson et al. 2004; Willemsen and Gooch 2002; Witmer and Sadowski 1998]. A number of factors may contribute to this discrepancy, such as the limited field of view, spatial resolution, subtle errors in rendering, and the weight of the HMD.

As discussed in Chapters IV and V, the gain of physical translation is scaled because Rieser et al. [1995] and Mohler et al. [2007b] have shown that people can quickly recalibrate to a new mapping between their own physical translation and visual input. However, the scaling factor of the translational gain in these recalibration studies was significantly smaller than that which is proposed in this thesis. Richardson and Waller [2005a; 2005b] showed that subjects adapted according to explicit feedback about the accuracy of their distance judgments in the virtual environment. They found that subjects accurately judged distances after receiving feedback (as opposed to pre-test). Kuhl [2004] and Pick et al. [1999] have shown that people can also recalibrate rotations. A compelling reason to manipulate translations instead of rotations is that research shows that physical changes in direction are more important than physical translation in the development of spatial knowledge [Presson and Montello 1994; Rieser 1989; Rieser et al. 1995]. By scaling translation and leaving rotation alone, we are decoupling rotation and translation, and no research has investigated what happens when people walk paths combining physical rotational locomotion with scaled translational locomotion. However, Riecke and Bülthoff [2004] found that people have some separation of visual and vestibular cues.

II.5 Navigating Large Environments

Previous research has explored various techniques of navigating a large scale virtual environment. Haptic devices, such as a joystick or keyboard, allow users to virtually explore large environments [Ruddle et al. 1999; Bowman et al. 1999; Waller et al. 1998; Darken and Sibert 1996; Pausch et al. 1997]. However, studies have shown that using physical bipedal locomotion rather than haptic devices produces significantly better spatial orientation [Chance et al. 1998; Ruddle and Lessels 2006; Lathrop and Kaiser 2002]. Specifically, Pausch et al. [1995] showed people were significantly more effective at search tasks involving a tracked HMD versus an untracked HMD. Ruddle et al. [1999] showed that exploring large virtual environments with a tracked HMD is significantly faster than exploring the same environment with a desktop display. Suma et al. [2007] show that using position and orientation tracking with an HMD is significantly better than using a system that combines the orientation tracking and a haptic device for translations.

Templeman et al. [1999] and Slater et al. [1995] have participants “walk in place” to move through large virtual environments, but this technique lacks the same proprioceptive cues of walking. Another method of navigating a large virtual environment is manipulating rotation such that the locomotion of the subject fits within the limits of the tracking system [Razzaque et al. 2001; Nitzsche et al. 2004]. Razzaque et al. [2001] examine subjects ability to locomote to a series of five targets they call “waypoints”. In this study, the virtual room is slightly rotated while the subject walks to the waypoint, and then to a greater degree as the subject searches for these waypoints. This method requires a large tracking area for the rotational manipulation to be imperceivable, and is not a complete solution because a situation could easily occur in which the physical limits of the tracking system are reached. Virtual flying [Usoh et al. 1999] and teleporting are other ways of exploring large virtual environments, yet they lack locomotive feedback.

Interrante et al. [2007] propose a method called “seven league boots” in which they scale gain based on wand control. In a pilot study, they compare scaling the translational

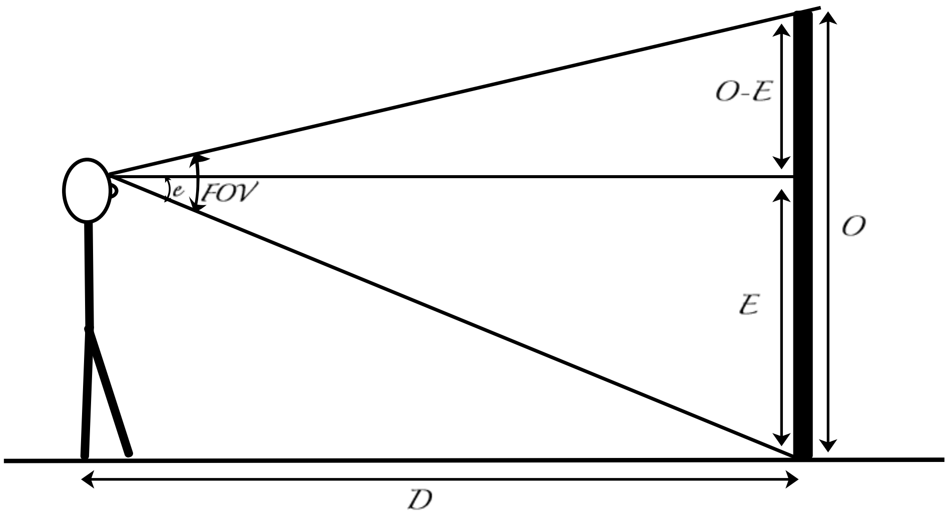

Figure II.4: This figure shows the side view of the geometry of a person's viewing angle and their associated eyeheight when viewing and object, *O*. The horizon line of gaze is represented by the line that connects the eye of the observer with *O* and is parallel to the ground. The eyeheight of the observer and the field of view are denoted *E* and *FOV*, respectively. The distance from the person to the observed object is distance *D*, and *e* represents the angle of declination from the horizon line to the ground.

gain of walking by ten, joystick locomotion, normal walking, and using "seven league boots" or a wand to activate translational gain scaled by ten. They then ask users to rate which method they prefer after walking down a 25 foot hallway in each of the conditions. Interrante et al. report that people seem to prefer "seven league boots" over normal walking, joystick, and scaling the translational gain by ten. This work suggests that it may be helpful to allow users some sort of control over optical flow rate. We explore this idea in Chapter V.

Other systems involve large screen caves with a locomotion input such as a bicycle or treadmill. Cave based system are expensive, and most only contain three virtual walls. Treadmill systems are difficult and expensive to construct with enough degrees of freedom to allow for free exploration.

II.6 Eyeheight

This work examines the role of eyeheight when experiencing a virtual environment. Mainly, it addresses the question of whether it is advantageous to scale eyeheight (or the height from which an environment is viewed). More specifically, eyeheight *E* refers to the distance from the viewer's visual horizon to the ground as shown in Figure II.4. For more information, see Sedgwick [1973; 1980] .

In our everyday lives, we humans constantly change our viewing perspective by sitting, standing, etc., yet the perceived relative size of objects remains the same. As Wraga [1999] points out, this may be because of familiar size or previous knowledge about size and shape [Gogel 1977; Rock 1975].

Additionally, the angle of declination *e* from the horizon line to the ground also provides another source of information as seen in Figure II.4. People use this information to recalibrate the relative sizes of objects at different eyeheights [Wraga 1999]. Warren et al. [1987] had subjects judge whether they could walk through doorways of varying widths. As a condition of the experiment the floor was raised relative to the floor that they were standing on, and subjects systematically overestimated the passibility of narrow doorways. It can be seen from Figure II.4 (after [Wraga 1999]), that if the floor is raised, then the perceived eyeheight *E* may become smaller. However, when the participants made the same judgments with the false floor and the real floor within the same block of experimental trials, no differences were observed. Moreover, Ooi et al. [2001] show that with a known eyeheight individuals can use the angular declination from the horizon to calculate absolute distances. Gardner et al. [2001] also perturb the vertical gaze distance by using prisms and they also provide evidence that the humans use vertical gaze angle as a distance cue. However, the Gardner et al. [2001] involved table top locations within reaching distance.

Wraga et al. [1999] compare seated, standing and ground–level prone observations and find that seated and standing observations are similar, but prone observations are significantly less accurate. Warren [1984] find that people judged whether they could sit on a surface according to whether the surface height exceeded 88% of their leg length. Moreover, people choose to climb or sit on a surface according to the relationship between the surface's height and their eyeheight [Mark 1987].

CHAPTER III

FUNCTIONAL SIMILARITY

III.1 Introduction

Important applications of virtual environments are based on the assumption that what people learn from exploring physical environments is functionally similar to what they learn from exploring virtual renderings of them. Functional similarity refers to the variables influencing responses based on representations learned through experiencing a physical environment having similar influence on responses based on representations learned through experiencing a virtual environment. Such functional similarity is important whenever people hope to explore and learn about virtual environments to apply their knowledge when planning or acting in the physical environments that they represent. Examples abound where evaluating scenarios in the physical world would be expensive and difficult so that virtual environments are used instead — for example — in simulations of behind-the-lines military operations, piloting a ship or plane, or learning to navigate a human body to prepare for surgery. In two experiments this chapter assesses the similarities and differences in how and how accurately people judge the directions toward objects after learning their locations by exploring physical environments versus virtual renderings of them displayed on an HMD [Williams et al. 2007b]. The results demonstrate functional similarities in access to spatial knowledge in real and virtual environments. Moreover, these functional similarities suggest that it is practical to explore a virtual environment using an HMD and maintain a spatial representation similar to the real world.

In these two experiments, subjects were asked to judge new perspectives across variations in their locomotion mode (whether their locomotion was physical or imagined), geometry of perspective change (whether their motion consisted of a translation or a rotation), and amount of disparity. In both experiments, subjects learned a novel environment by freely exploring it in each of two visual input conditions: in one people viewed their actual physical surroundings and in the other they viewed a graphically rendered simulation of the same surroundings presented over a head-mounted display (HMD). The purpose of the experiments was to investigate the functional similarity of the representations derived from viewing the physical environment and the virtual environment.

Thus, by functional similarity, we mean similarities or differences in performance when judging new perspectives based on the locomotion mode, the geometry of the perspective change, and the amount of disparity. These judgments were assessed by measuring the errors and latencies (time to complete the task) of the perspective changes. Both measures are free to vary in the experiments and both provide a measure of subjects' access to their spatial knowledge gained through one of the visual input modes.

For reasons discussed previously, we expect performance judging perspectives after exploring the virtual environment over the HMD to be less accurate or slower to process than after viewing the physical environment. Hence, a significant or not significant effect of the visual input type on these experiments is interesting, and potentially indicates the quality of immersion, but is not the primary goal here. Instead, an analysis of functional similarities involves determining whether the conditions that influence responses from representations learned while viewing the physical environment have similar influence on responses from representations learned while viewing virtual environments.

III.2 Experimental Design

Two experiments evaluated the functional similarities in using knowledge learned from exploring a large room-sized physical environment and a virtual construction of the same environment presented over an HMD and explored freely on foot. In each experiment subjects were asked to explore and learn the spatial layout of eight target objects. After learning the target locations from the perspective at one point of observation, they were asked to close their eyes and make knowledge-based judgments of the target directions from a new perspective. An important note about these experiments is that the FOV of a subject in the physical environment was not restricted to be the same as in the virtual environment, as some have done, e.g., [Thompson et al. 2004]. The reason for this decision is that we wanted to test the subjects when they learned in settings that were as unencumbered as possible. For a virtual environment presented through an HMD, this setting involves limitations on FOV, but does not in the physical environment. Thus, the emphasis in these studies is on the functional similarities in spatial reasoning despite the many differences in input for learning.

Experiment 1 was designed to determine whether physical locomotion facilitates the speed and/or accuracy of judging new perspectives after simple rotations in facing direction. In the *locomotion* condition the subjects were blindfolded (or the HMD was darkened), they were physically guided to face the new direction, and they were asked to judge the target directions. In the *imagination* condition the subjects were blindfolded and asked to imagine turning to face the new direction. Thus, the three factors of the study were *visual input* (visually exploring the physical environment versus its virtual environment rendition) × locomotion method (physical locomotion versus imagination) × disparity (that is, there were eight different amounts of rotation in perspective to be judged, ranging from 0° to 180°). The type of visual input was varied across subjects (one-half studied the physical environment and one-half studied the virtual environment), the locomotion method was blocked and varied within repeated trials of each subject, and the disparity was randomly ordered within each block for each subject. The blocking adds an additional condition that allows us to assess whether learning occurs between blocks.

Experiment 2 was designed to determine whether the geometry of the imagined changes in perspective (there was no physical locomotion) influence the difficulty. Again subjects visually explored either a physical room or a virtual rendering of the same room and then were asked judge target directions after imagining changes in perspective that consisted of simple rotations (that is, subjects were asked to imagine they had turned to change their facing direction while standing in the same spot in the room) or simple translations (that is, subjects were asked to imagine keeping their facing direction constant while moving straight to a new spot in the room). The amount of to-be-imagined change in facing direction varied across both the rotation trials and the translation trials. The type of visual input was varied across the subjects, the geometry of the perspective change was blocked within the repeated trials of each subjects, and the disparity was randomly varied within these blocks. The blocking was again a condition within the experiments that was considered.

People typically move through their environment using a combination of both rotation and translation, and Experiment 2 decouples them. In comparing translations and rotations, the underlying difficulty of the a perspective change depends on the disparity, as discussed previously. In the type of experiment presented here (modeled after the studies of Rieser [1989]), disparities for both translation and rotations fall into natural and principled values. However, these values are not identical, and, thus for the design, the range of disparities for translations will be less than the range of disparities for rotations, and the natural values of disparities for translations will be different than the natural values of disparities for rotations. When comparing rotational and translational disparities prior work (e.g., May [2004]) has typically clustered or "binned" values. Statistical power can be lost through this binning, and modern statistical methods [Harrell, Jr. 2001; Myers et al. 2002] do not require binning. Thus, the analyses later in this chapter will use the true, unbinned values of disparity. However, for didactic purposes, values of disparity are displayed in the figures, as it is easier to visually interpret this style of presentation than a scatterplot, and, as we will see, doing so does not misrepresent the data.

Thus, both experiments were designed to find out the relative difficulty of knowledge-based judgments of perspective changes as a function of whether the knowledge was learned from visually exploring a virtual environment versus a physical one. In addition, they allow us to investigate the degree to which the relative difficulty of the judgments was influenced by whether the perspective change was accompanied by locomotor movements, by the disparity in perspective that was to be judged, and by whether the geometry of the change consisted of a simple rotation or simple translation in perspective. To assess relative difficulty, subjects were asked to turn and face the target objects from the new perspectives quickly and accurately. Both the turning errors and the latencies can be expected to vary as a function of difficulty, except in situations where subjects trade speed for accuracy by, for example, responding rapidly and inaccurately some of the time and slowly and accurately some of the time. As long as latencies and accuracies are positively correlated (that is, people are not trading speed for accuracy), then both are reasonable measures of relative difficulty. However, in conditions where latencies and accuracies are negatively correlated, then the meaning of both measures is ambiguous and cannot be used to sort the relative difficulty across conditions.

In this experiment, the goal was to assess similarities of spatial knowledge between physical and virtual environments. More specifically, for both environments, we investigate the relative difficulty of imagining rotations and whether physical rotation facilitates access to spatial structure. This experiment mimics Experiment 2 of Rieser et al. [1989] with the additional condition of being conducted in a virtual or physical environment.

III.3 Experiment 1: Imagined Rotations and Physical Rotations

III.3.1 Method

Participants

Sixteen Vanderbilt University students participated in the experiment. Subjects were unfamiliar with the experimentation room and the virtual reality equipment.

Materials

In both the real and virtual conditions, participants viewed eight targets (ball, rubber duck, bottle, telephone, vase, videotape, scissors, and clock) that were arranged in an evenly spaced circle. Each of the targets varied in shape and color. Identical objects in the same location were used for the virtual and real world condition, and were of similar size. A

Figure III.1: Laboratory or "physical" environment for conducting experiments. Shown is the circular array of eight objects used in the experiments. For both Experiment 1 and Experiment 2, subjects stand in the center of the array. This particular environment was used in Experiment 2, and Experiment 1's virtual environment was similar.

Figure III.2: This figure shows the virtual environment and the circular array of eight objects used in the experiments.

virtual room, as seen in Figure III.2, of the same scale and layout was designed to mimic the real environment, shown in Figure III.1 (the views are not exactly similar because we could not replicate the exact camera parameters between the physical and virtual cameras).

The virtual world was viewed through a full color stereo Virtual Research Systems V8 HMD with 640 × 480 resolution per eye, a field of view of 60° diagonally, and full binocular overlap. The HMD also weighed approximately 1kg. An InterSense IS-900 tracker was used to update the participant's rotational movements. The rendered field of view was matched to the nominal displayed field of view (otherwise, performance errors can result [Psotka et al. 1998]).

Procedure

One-half of the subjects performed the experiment in the virtual world, and the other half performed the experiment in the real world. In both conditions, subjects stood in the center of the array of objects, which was roughly the center of a 5.8m by 6.4m room. Each target was approximately 2m from the participant. After approximately three minutes of study, the experimenter tested the subjects by asking them to close their eyes and point to randomly selected target objects. This testing and learning procedure was repeated until both the experimenter and subject were confident that the configuration had been learned.

Next participants were asked to imagine themselves at a novel direction of observation, and point to a target from that point. Specifically participants were instructed "face the ⟨ *target name* ⟩ as if you are facing the ⟨ *source object name* ⟩." There were two conditions: a locomotion condition and an imagination condition. In the locomotion condition, subjects were physically turned by the experimenter to face the *source object* during the verbal instruction. To move the subject, the experimenter grasped the subject by the shoulders and rotated them. In the imagination condition, subjects followed the instruction by imagining turning to face the object that identified the new facing direction. Note that in some of the trials, the subject was actually facing the source object, so that no imagined or physical locomotion occurred. Subjects were asked to not simply turn their heads, but to locomote their entire body, although the response was measured from the HMD based on the head-facing direction. Note also that as response angles were measured through the HMD, in all conditions subjects wore a darkened HMD when turning, which served as a blindfold. The physical room was darkened as well. The instructions for the imagination and locomotion conditions were explained before participants saw the experimental layout. After every trial, subjects rotated themselves back to a starting position facing the front of the room. If there was any error after this rotation between their facing direction and the correct starting position the experimenter rotated them to the correct starting position. The subjects completed 56 test trials for each condition in two blocks of 28. The order of the instructions for each block was randomized. Half of the subjects were tested first for imagination, and the other half were tested first for the locomotion condition. To compare the angles of the correct responses across different conditions, the same trials were used for the locomotion and imagination modes, and for the virtual and real world environments.

The subjects indicated to the experimenter that they were facing the target by verbal instruction, and the experimenter recorded their time and rotational position. The time was recorded using a stopwatch, and the rotational position was recorded using the InterSense tracker. Subjects were encouraged to respond as rapidly as possible, while maintaining accuracy. If subjects were confused, they were allowed to look again at the array during the experiment, between trials.

III.3.2 Results

The dependent variables in this experiment were the errors and latencies in turning to face the targets. There were four independent variables: visual input type, locomotion method, amount of disparity, and the learning block (either first or second) in which the particular trial was conducted. The type of input was either virtual or physical, depending on whether the subject learned the target locations in the virtual space or physical space. The locomotion method was either imaginary (imagined rotation) or physical (physically turned). The amount of rotation varied as repeated within-subject trials. The results were analyzed by learning block to determine if order effects (i.e., learning) are an important factor in the experiment.

Figure III.3 shows the mean turning error, and Figure III.4 shows the mean latencies as functions of visual input mode and locomotion method. Turning error was defined as the unsigned error, i.e., the unsigned difference between the actual target distances and the observed distances. The average turning errors as functions of disparity and locomotion method are presented separately for the physical input mode, Figure III.5, and the virtual input mode, Figure III.6. Likewise, the average latencies as a function of disparity and locomotion method are shown in Figure III.7 (physical input mode), and Figure III.8 (virtual input mode). Note that disparity is a continuous variable and has values between 0 and 180 determined by the geometry of the experimental setup. Following the practice of May [2004], for ease of presentation in these figures the disparities have been clustered to

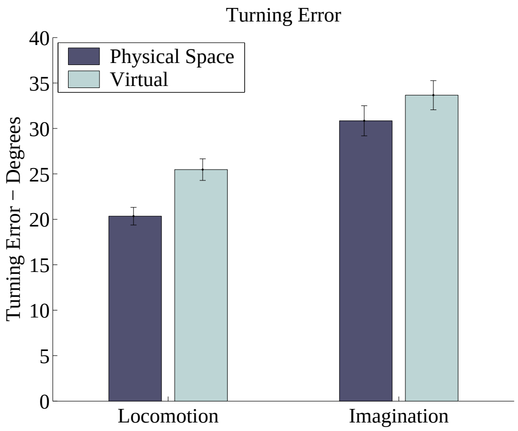

Figure III.3: Mean turning error for both the locomotion and imagination method in the virtual and physical conditions in Experiment 1. Error bars indicate the standard error of the mean.

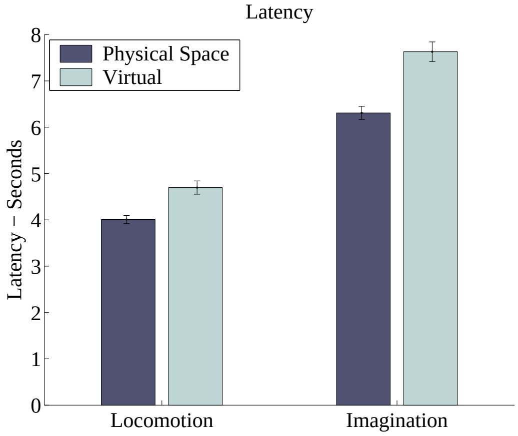

Figure III.4: Mean latency for both the locomotion and imagination task in the virtual and physical conditions in Experiment 1. Error bars indicate the standard error of the mean.

their closest 45° amount. In the statistical analysis described below, however, unclustered "true" disparity values were used. The mean turning error and latency by learning block are shown in Figures III.9 and III.10, respectively.

Since the amount of disparity is a quantitative independent variable, a least squares regression was used to fit a generalized linear (mixed) model to the data, and conducted an analysis of variance to test the significance of the independent variables and associated interactions using techniques described in Harrell [2001] and Myers et al. [2002]. Four independent variables were used to predict the turning error and latency, respectively, and modeled all two-way and three-way interactions of the independent variables. The resulting model had 14 degrees of freedom. This type of model allows us to test if disparity was better modeled as a quadratic, but the quadratic effect was not significant. Tables III.1 and III.2 show the analysis of variance results for this experiment. Results were considered significant if $p < .05$. The "effective" degrees of freedom are reported [Satterthwaite 1941; Satterthwaite 1946]. Since correlation may exist over repeated values of disparity, the residual errors in the regression may not be independent, and thus the variance estimates may be wrong. To correct this problem, the degrees of freedom are adjusted.

	$d.f.$	$PartialSS$	MS	F	p
Disparity*	7	26732	3819	4.5	$< .01$
Visual Input	7	11319	1617	1.9	.06
Locomotion*	7	53921	7703	9.1	$< .01$
Block*	7	12789	1827	2.2	.03
Disparity × Visual Input	3	2965	988	1.2	.32
Disparity × Locomotion	3	5780	1927	2.3	.08
Visual Input × Locomotion	3	764	255	0.3	.82
Disparity × Block	3	2196	732	0.9	.46
Visual Input × Block	3	2710	903	1.1	.36
Locomotion × Block*	3	9414	3138	3.7	.01
Disparity × Visual Input × Locomotion	1	2	2	0.0	.97
Disparity × Visual Input × Block	1	1928	1928	2.3	.13
Disparity × Locomotion × Block	1	11	11	0.0	.91
Visual Input × Locomotion × Block	1	289	289	0.3	.56

Table III.1: Analysis of Variance for Turning Error in Experiment 1. The $d.f.$ column gives the number of coefficients of terms including that term in the regression model. Other columns are standard statistical measures. Significant main effects and interactions are indicated with an asterisk (*).

	$d.f.$	$PartialSS$	MS	F	p
Disparity*	7	329	47	4.6	$< .01$
Visual Input*	7	557	80	7.9	$< .01$
Locomotion*	7	3243	463	45.8	$< .01$
Block*	7	488	70	6.9	$< .01$
Disparity × Visual Input	3	8	3	0.3	.85
Disparity × Locomotion*	3	103	34	3.4	.02
Visual Input × Locomotion	3	56	19	1.9	.13
Disparity × Block	3	10	3	0.3	.81
Visual Input × Block	3	57	19	1.9	.13
Locomotion × Block	3	64	21	2.1	.10
Disparity × Visual Input × Locomotion	1	5	5	0.5	.49
Disparity × Visual Input × Block	1	3	3	0.3	.57
Disparity × Locomotion × Block	1	0	0	0.0	.87
Visual Input × Locomotion × Block	1	9	9	0.9	.35

Table III.2: Analysis of Variance for Latency in Experiment 1. The $d.f.$ column gives the number of coefficients of terms including that term in the regression model. Other columns are standard statistical measures. Significant main effects and interactions are indicated with an asterisk (*).

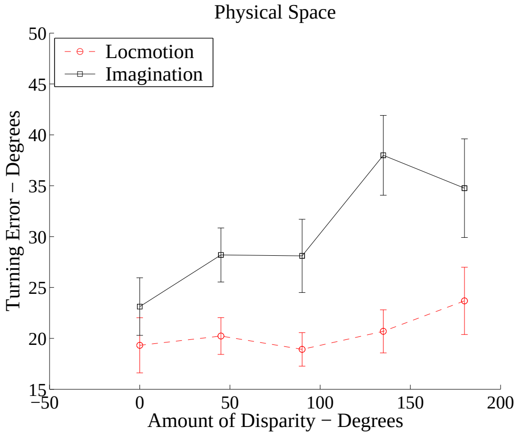

Figure III.5: Mean turning error as a function of disparity in the real world condition for Experiment 1. The red line represents the turning error in the locomotion repositioning task; the black line shows the turning error in the imagination repositioning task. Error bars indicate the standard error of the mean.

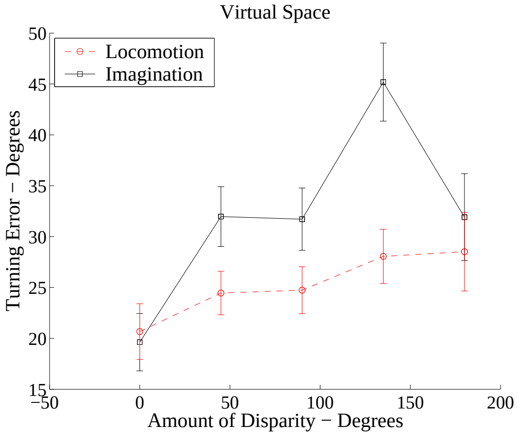

Figure III.6: Mean turning error as a function of disparity in the virtual condition for Experiment 1. The red line represents the turning error in the locomotion repositioning task; the black line shows the turning error in the imagination repositioning task. Error bars indicate the standard error of the mean.

From these tables, we see that there are a significant main effects of disparity, locomotion mode (physical or imagined), and learning block on the turning errors and latencies. The main effect on disparity indicates that lower disparities lead to more accurate, faster performance. Subjects were faster and more accurate in their judgments when turned (the locomotion mode) than when they have to imagine turning. Also, the performance of subjects in block 2 was faster and more accurate than performance in the block 1.

For turning error, there is a significant interaction of locomotion on learning block; turning errors decreased in the physical locomotion condition from block 1 to block 2, while they increased slightly in the imagination condition from block 1 to block 2. Subjects seemed to benefit from the chance to learn in the locomotion condition (performance improved), but in the more difficult imagination condition, subjects seemed to become fatigued and their errors increased. This interaction therefore does not seem to add much to the interpretation of the main effect.

For the latencies there is a significant main effect of visual input (real world or virtual).

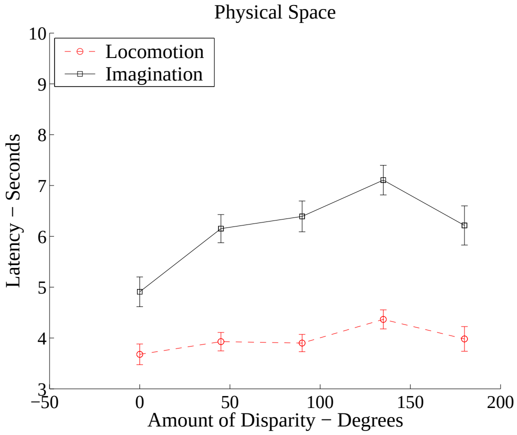

Figure III.7: Latency as a function of disparity in the real world condition for Experiment 1. The red line represents the latency in the locomotion repositioning task; the black line shows the latency in the imagination repositioning task. Error bars indicate the standard error of the mean.

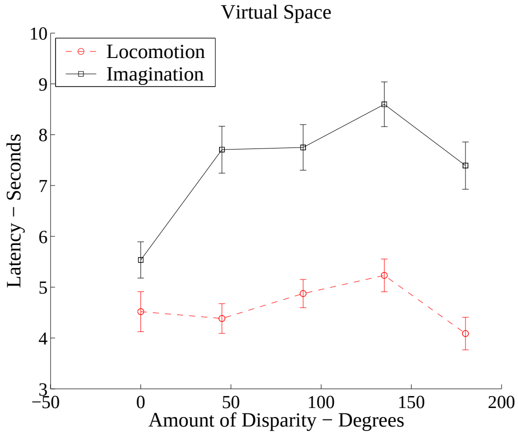

Figure III.8: Latency as a function of disparity in the virtual condition for Experiment 1. The red line represents the latency in the locomotion repositioning task; the black line shows the latency in the imagination repositioning task. Error bars indicate the standard error of the mean.

This effect is only marginal for the turning errors. Subjects were slower in completing the task if their visual input was through the HMD rather than through the physical environment. This is discussed further in the discussion, Section III.5, at the end of this Chapter, but this effect may be due to various factors involving the immersion that the subjects felt in the virtual environment. There is a significant interaction between disparity and locomotion, meaning that disparity has a larger impact in the imagination condition than in the physical locomotion condition. This interaction is one that is expected and consistent with the literature, e.g., [May 2004]. In particular, subjects had more difficulty turning accurately and speedily in the imagination condition than the locomotion condition for both environments, and the amount of difficulty they had depended on the amount of disparity in the task. The main conclusion drawn from this experiment is that there exists a functional similarity in subject's access to spatial knowledge in both real and virtual environments.

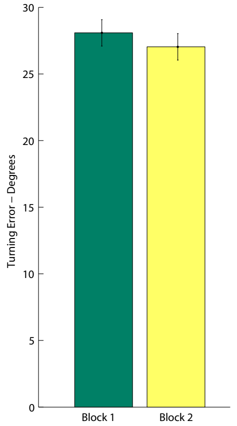

Figure III.9: This figure shows the mean turning error in the first block and second block of Experiment 1. The error bars represent standard errors of the mean.

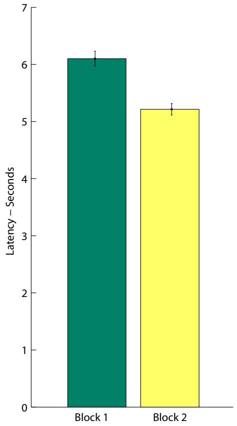

Figure III.10: This figure shows the mean latency in the first block and second block of Experiment 1. The error bars represent standard errors of the mean.

III.4 Experiment 2: Imagined Rotations and Imagined Translations

In this experiment, we replicated and elaborated the design of Experiment 3 from Rieser [1989]. Subjects were asked to learn the spatial layout of targets either by exploring a physical environment or a virtual rendering of it via a tethered HMD. Like the earlier experiment, subjects were asked to judge self-to-object directions from novel points of observation that were either simple rotations of their original point of simple translations of it.

III.4.1 Method

Participants

Fourteen Vanderbilt University students and two non-student adults participated in the experiment. All were unfamiliar with the experimentation room and the virtual reality equipment.

Materials

The HMD and tracker used in Experiment 1 were used again in Experiment 2. The array of objects used in this experiment were a sneaker, a videotape, a book, a hairbrush, a vase, a telephone, a soda can, and a coffee mug. These objects existed in the same locations in both the real and virtual environments as seen in Figures III.1 and III.2, and were spaced in 45 degree increments in a circle about 2.25 m from the center. Again, the room was approximately 5 m by 5 m.

Procedure

One-half of the subjects performed the task in the virtual world, the other half in the physical world. In both the virtual and physical environments, there were two conditions: rotation and translation. Subjects imagine translating or rotating to a specific spatial point, then they are asked to face an object from that spatial location. The rotation condition was the same as Experiment 1, and participants were asked to "face the ⟨ *target name* ⟩ as if you are facing the ⟨ *source object name* ⟩." In the translation condition, subjects were asked to "face the ⟨ *target name* ⟩ as if standing at the ⟨ *source object name* ⟩." The participants were tested using four alternating blocks of 28 trials. One-half of the subjects did a rotation block first, the other half did a translation block first. The procedures of the study were carefully explained to the subjects before they entered into the test room. Once the experimenter thought that the subject demonstrated understanding of the tasks, the subjects were led into the center of the array with their eyes closed. Next, the subjects were instructed to open their eyes and learn the locations of the objects in the array. The study phase was similar to that of the previous experiment and lasted about 2-5 minutes until the experimenter felt that the subject was familiar with the locations of the objects.

III.4.2 Results

Again in this experiment, the dependent variables were the errors and latencies in turning to face the targets, defined as in Experiment 1. There were four independent variables: visual input type, geometry of perspective change, amount of disparity, and the block (either first or second) in which the particular trial was conducted. Visual input, disparity, and block were as in Experiment 1. The geometry of perspective change indicates whether

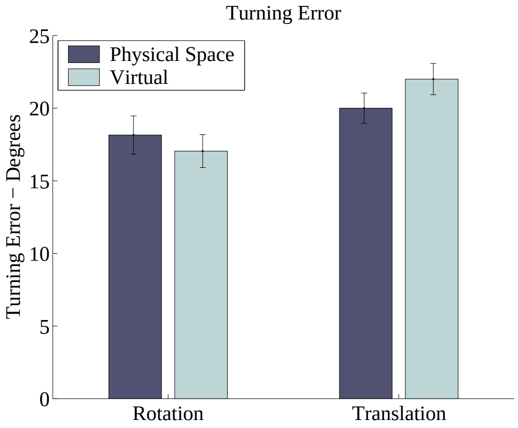

Figure III.11: Mean turning error for both the translation and rotation repositioning task in the virtual and real conditions for Experiment 2. Error bars indicate the standard error of the mean.

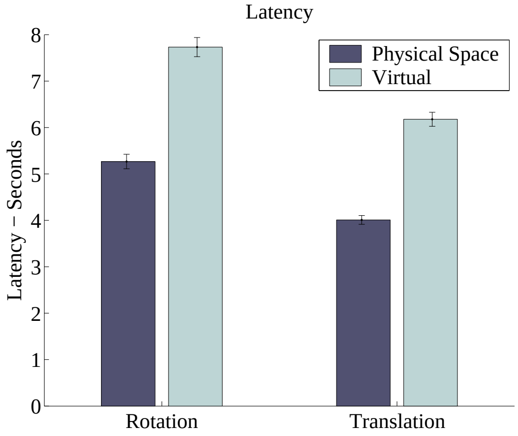

Figure III.12: Mean latency for both the translation and rotation repositioning task in the virtual and real conditions for Experiment 2. Error bars indicate the standard error of the mean.

the subject was asked to conduct a perspective change by turning (rotation) or by linear movement while maintaining heading direction (translation). Both the amount of rotation and the direction of translation varied as repeated within-subject trials. Results were again analyzed by learning block to determine if order effects (i.e., learning) are an important factor in the experiment. Data from two of the subjects were not calculated in the results as they failed to understand the task after repeated explanation. This failure was evident in comments they made after the experiment, and was evident in that they both performed worse than chance in the rotation and translation conditions.

Figures III.11 and III.12 show the mean turning errors and latencies, respectively. Mean turning error as a function of disparity is shown in Figures III.13 and III.14. Mean response time as a function of disparity is shown in Figure III.15 and III.16. In the figures, the disparities are clustered into groups for purpose of presentation, but analysis results are done on the measured values. Note that the circular arrangement of objects constrains the maximum translational disparity that can occur, in contrast to arrangements in May [2004]. Finally, mean turning error and latency as a function of learning block are shown in Figures III.17 and III.18.

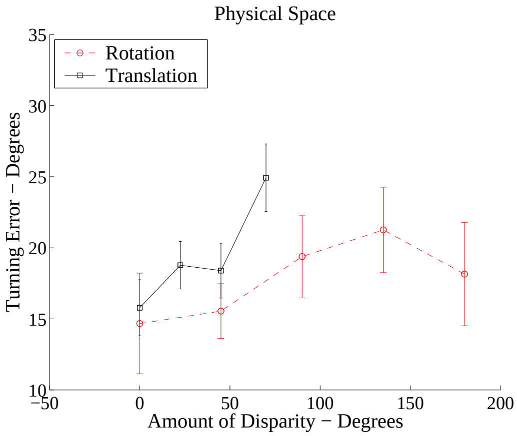

Figure III.13: Mean turning error as a function of disparity in the real world condition for Experiment 2. The red line represents the turning error in the rotational repositioning task; the black line shows the turning error in the translational repositioning task. Error bars indicate the standard error of the mean.

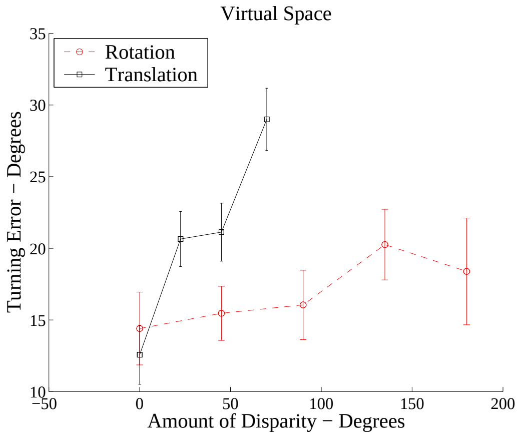

Figure III.14: Mean turning error as a function of disparity in the virtual condition for Experiment 2. The red line represents the turning error in the rotational repositioning task; the black line shows the turning error in the translational repositioning task. Error bars indicate the standard error of the mean.

To analyze these results, a model similar to that used in Experiment 1 was employed, with the substitution of a predictor of the geometry of perspective change rather than the locomotion method, and having the same number of degrees of freedom. Results were again considered significant if $p < .05$. Tables III.3 and III.4 show the analysis of variance results for this experiment.

Analogous to Experiment 1, there are significant main effects of disparity and geometry on both the turning errors and latencies. Like Experiment 1, lower disparities lead to faster and more accurate performance. The main effect of geometry is, however, an ambiguous finding, since turning errors increase but latencies decrease, indicating that speed-accuracy trade-offs are being made, a point discussed earlier. For latencies, there are significant main effects of the visual input mode and of learning block. As measured by latency, performance in the real world is significantly faster than performance in the virtual environment. For turning errors, there is a significant interaction between disparity and the geometry of the perspective change. This interaction indicates that disparity has a larger effect on the

	$d.f.$	*PartialSS*	*MS*	*F*	*p*
Disparity*	7	18463	2638	4.6	$< .01$
Visual Input	7	3610	516	0.9	.51
Geometry*	7	22802	3257	5.6	$< .01$
Block	7	6022	860	1.5	.17
Disparity × Visual Input	3	1096	365	0.6	.60
Disparity × Geometry*	3	8654	2885	5.0	$< .01$
Visual Input × Geometry	3	2410	803	1.4	.25
Disparity × Block	3	599	200	0.3	.79
Visual Input × Block	3	1576	525	0.9	.44
Geometry × Block	3	798	266	0.5	.71
Disparity × Visual Input × Geometry	1	729	729	1.3	.26
Disparity × Visual Input × Block	1	272	272	0.5	.49
Disparity × Geometry × Block	1	221	221	0.4	.54
Visual Input × Geometry × Block	1	557	557	1.0	.33

Table III.3: Analysis of Variance for turning error in Experiment 2. The $d.f.$ column gives the number of coefficients of terms including that term in the regression model. Other columns are standard statistical measures. Significant main effects and interactions are indicated with an asterisk (*).

	$d.f.$	*PartialSS*	*MS*	*F*	*p*
Disparity*	7	287	41	3.8	$< .01$
Visual Input*	7	2482	355	32.8	$< .01$
Geometry*	7	382	55	5.1	$< .01$
Block*	7	430	61	5.7	$< .01$
Disparity × Visual Input	3	41	14	1.3	.28
Disparity × Geometry	3	14	5	0.4	.74
Visual Input × Geometry	3	25	8	0.8	.51
Disparity × Block	3	44	15	1.4	.25
Visual Input × Block	3	44	15	1.4	.25
Geometry × Block	3	54	18	1.7	.17
Disparity × Visual Input × Geometry	1	1	1	0.1	.71
Disparity × Visual Input × Block	1	17	17	1.6	.21
Disparity × Geometry × Block	1	5	5	0.5	.49
Visual Input × Geometry × Block	1	24	24	2.2	.14

Table III.4: Analysis of Variance for latency in Experiment 2. The $d.f.$ column gives the number of coefficients of terms including that term in the regression model. Other columns are standard statistical measures. Significant main effects and interactions are indicated with an asterisk (*).

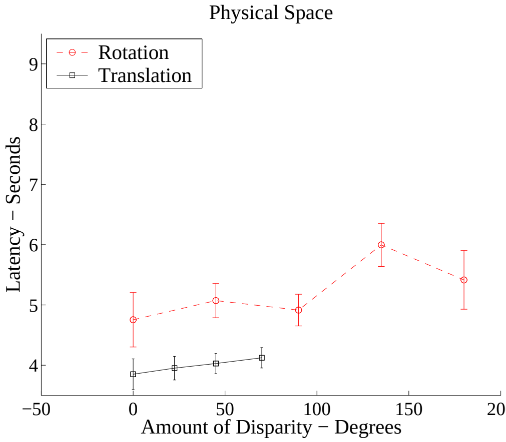

Figure III.15: Latency as a function of disparity in the real world condition for Experiment 2. The red line represents the latency in the rotational repositioning task; the black line shows the latency in the translational repositioning task. Error bars indicate the standard error of the mean.

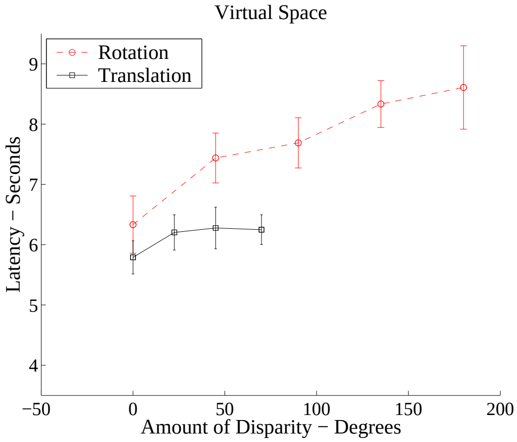

Figure III.16: Latency as a function of disparity in the virtual condition for Experiment 2. The red line represents the latency in the rotational repositioning task; the black line shows the latency in the translational repositioning task. Error bars indicate the standard error of the mean.

translation condition than the rotation condition.

III.5 Discussion

This chapter presents two experiments that were conducted to assess the degree to which representation-based judgments of perspective were functionally similar across conditions where the input to the representation resulted from freely viewing the physical environment versus a virtual rendering of the same environment viewed over an HMD. Judgments were assessed in terms of errors and latencies, both of which were free to vary. These measures are reasonable indicators of performance provided they are not traded off against one another.

By functional similarity, we mean that variables influencing responses based on representations learned through experiencing a physical environment have similar influence on responses based on representations learned through experiencing a virtual environment. Functional similarity does not refer to the presence or lack of a main effect of the visual

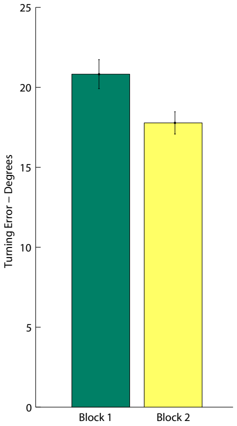

Figure III.17: This figure shows the average turning error in the first block and second block for Experiment 2. The error bars represent standard errors of the mean.

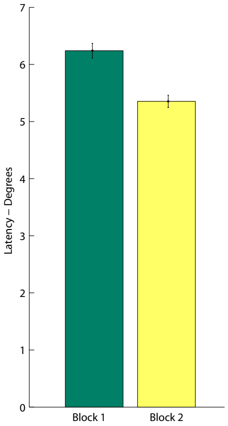

Figure III.18: This figure shows the average latency in the first block and second block for Experiment 2. The error bars represent standard errors of the mean.

input type on the errors or latencies. An HMD has technical limitations that make it reasonable to expect performance in judging perspectives to be inferior when learned through an HMD than when experienced in a physical environment. Indeed, errors were marginally higher after learning a representation through a virtual environment in Experiment 1, and latencies were reliably longer in both Experiments 1 and 2. However, the visual input type had no interactions with any of the other independent variables in either experiment. Thus, the functional similarities or differences between physical and virtual environments are the similarities or differences in responses to the two input types based on measures of mode of locomotion, geometry of perspective change, and disparity.

Thus, this work demonstrates that spatial learning is functionally similar in both real and virtual environments. Specifically, whether the spatial representation was gained through a virtual display or by sight in the physical world, physical locomotion was a strong aid in computing perspective changes. Moreover, perspective changes are harder with increasing disparity in both the physical and virtual environments. This result is well-known for physical environments, but had not been demonstrated before in virtual environments. Also, proprioceptive feedback from locomotion facilitates understanding of the spatial representation in virtual environments as well as physical ones. If this finding can be successfully exploited, it may prove useful in the design of navigation and way-finding interfaces for virtual simulations.

Some differences between our results and the results of Rieser [1989] exist. The present study shows differences in the locomotion and imagination conditions for the no-disparity conditions, whereas Rieser [1989] found identical performance. We believe that this difference occurs because in the experiments presented here subjects were primed to expect a change in the to-be-imagined facing direction during the block of trials, so that it took them longer even when the instruction was to imagine facing their actual facing direction. Also, response times and errors in Rieser [1989] were generally lower than my response times and errors. Note that in the experiments shown here, subjects were slower in both the physical environment and the virtual environment, and this can be attributed to differences in experimental design: Rieser [1989] used a swivel-mounted pointer that was manipulated by hand, whereas in my experiments subjects turned their bodies to a facing direction, and subjects in this experiment were always wearing an HMD in the testing condition. We have tested this experimental setup ourselves and find that it is difficult to move significantly more quickly than our results indicate.

Earlier studies show that when operating in the physical environment, adults imagine simple translations in perspective more accurately and/or rapidly than simple rotations [Presson and Montello 1994; Rieser 1989]. However, in the present experiment the results are ambiguous. On the one hand, people judged the to-be-imagined translations more rapidly than the rotations, but their errors were larger. We hypothesize that the different pattern of results for errors and latencies reflect the strategies that some subjects reported. For the translation conditions, most subjects reported they were able to imagine themselves physically standing at the new observation point; we assume their judgments were rapid because they based them directly on their representation, and we assume they were inaccurate because they misjudged the distance of the needed translations. For the rotation condition, on the other hand, subjects reported they were not able to imagine actually facing in that direction. And so instead, they computed their answers by figuring the difference between their actual facing angle and the to-be-judged angle. The fact that they needed to figure is consistent with their slower latencies.

A final note about the main effects in performance between virtual environment and physical environment should be made, however. While there were no statistically significant differences in the accuracy with which people turned in the physical or virtual environments, people took significantly longer to accomplish a perspective change in the virtual environment than in the physical environment. This poorer performances in the virtual environment may be a quality of "immersion" related to the more limited field of view in the HMD, the poorer quality of rendering in the HMD (as opposed to vision), or the limited resolution of the HMD. Although a sense of immersion is difficult to define, gaining strong spatial representations in virtual environments is likely to be a critical component of it, and this thesis progresses in this direction.

CHAPTER IV

SCALING TRANSLATIONAL GAIN

IV.1 Introduction

Navigating through large virtual environments using a head-mounted display (HMD) is difficult due to the spatial limitations of the tracking system. This chapter addresses the issue of exploring a virtual environment that is larger than the physical boundaries of the tracking system by manipulating the translational gain of walking. Specifically, this chapter examines whether scaling the translation gain of walking is plausible idea, and limits the investigation of spatial orientation on scaling gain to a factor of ten.

This thesis seeks to leverage the natural ability of people to maintain spatial awareness of an environment when translation through the environment is provided by bipedal locomotion. This modality is natural for the HMD since HMD technology often uses a head tracker that measures changes in orientation and position of the user's head within the physical environment. The display of the HMD is updated using the user's 3D location in the physical space so that movement in the virtual world is equal to movement in the physical world. Unfortunately, the finite range of the HMD tracking system, or, more importantly, the limited amount of space a commodity level user may have to devote to an HMD, limits the size of space that can be freely explored using bipedal locomotion. Of course, using a joystick to translate might be a solution, as some have proposed (e.g., [Bowman et al. 1999]), and we first address the issue of how well that works.

IV.2 Experimental Design

Increasing the translational gain of bipedal walking is useful if it is a superior method of exploring large virtual environments with an HMD. The logical choice of comparison to the scaled gain locomotion is joystick locomotion. Therefore, Experiment 3 aims at comparing joystick navigation with normal bipedal locomotion (the 1:1 condition) and with bipedal locomotion scaled by ten (the 10:1 condition). Virtual rotations are equal to the physical rotation in both joystick and physical locomotion conditions. Thus, in the 1:1 locomotion condition, subjects' position in the virtual world is limited to the physical limits of the HMD tracking system, whereas, in the 10:1 condition, subjects were able to explore well outside that limit. Subjects' spatial orientation under physical locomotion and joystick locomotion was compared by having them locate target objects in the room with eyes closed, recording their error and response latency. Our hypothesis that participants would orient themselves well using bipedal locomotion in a virtual environment is based on work published on the advantages of locomotion on spatial orientation under normal translation conditions [Easton and Sholl 1995; Farrell and Robertson 1998; May 1996b; Presson and Montello 1994; Rieser 1989; Wraga et al. 2000; Wraga 2003; Williams et al. 2007b]. Moreover, bipedal locomotion gives the subject proprioceptive cues allowing for more accurate distance and direction estimation as shown by Loomis et al. [1992].

Experiment 4 further examines spatial learning and updating orientation when the translational gain of bipedal locomotion is scaled. More specifically, the following three translation conditions are compared within subjects: translational gain scaled by one (1:1), translational gain scaled by two (2:1), and translational gain scaled by ten (10:1). In the 2:1 and 10:1 conditions, users are allowed to virtually walk beyond the physical boundaries of the tracking system. In all three of the conditions, the subjects' spatial orientation is tested by having them turn to face targets in the room with eyes closed. Their response latencies and turning error were recorded. In all three conditions, the subjects' physical rotation corresponded to the same rotation in virtual space. An issue to note with our framework is that people must be able to adapt to increases in these gains within the 10-15 minutes it takes to perform that portion of the experiment. However, it might take longer, so as a sub-study we examined whether people who play video games, where high rates of optic flow are the norm, performed better than people who did not play video games.

IV.3 Experiment 3 : Joystick versus walking translations

The first experiment compares locomotion interfaces that depend on two different motor actions to translate the subject's perspective in virtual space, contrasting bipedal locomotion in one condition with joystick manipulation in the other. The results of the study compare learning and orientation under physical rotation combined with joystick translation versus physical rotation combined with walking in the 1:1 and 10:1 gain conditions. To test subject orientation, the subjects were asked to remember the location of seven objects in the room, then were asked to move themselves (using joystick or walking) to a new point of observation and instructed to turn to face the targets from memory without vision.

IV.3.1 Method

Participants

Sixteen subjects, twelve Vanderbilt University students and four non-student adults, participated in the experiment. Subjects were unfamiliar with the experiment and the virtual environments. Subjects were given compensation for their participation.

Materials

The virtual world was viewed through a full color stereo Virtual Research Systems V8 Head Mounted Display with 640 x 480 resolution per eye, a field of view of 60° diagonally, and a frame rate of 60 Hz. The HMD weighs approximately 1 kg. An InterSense IS-900 tracker was used to update the participant's rotational movements around all three axes. Position was updated using two optical tracking cameras with an accuracy of < 0.5 cm over 3 x 3 x 3 m volume and an update rate of 60 Hz. The type of joystick used in this experiment was the Logitech Attack 3.

The size of the physical room in which the experiments were performed was approximately 5m by 6m, and within the room the limits of the video position tracking system was approximately 5m by 5m. The size of the 1:1 room corresponded to the physical limits of the tracking system. The size of the 10:1 room was ten times the size of the 1:1 room, such

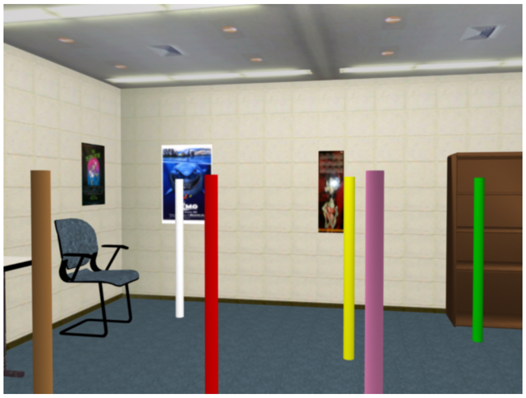

Figure IV.1: This figure shows the virtual environment of the 1:1 condition used in both Experiment 3 and 4. The target objects in this experiment were the different color cylinders, some of which can be seen in this image.

Figure IV.2: This figure shows the virtual environment of the 10:1 condition used in both Experiment 3 and 4. The target objects in this experiment were the different color tables, some of which can be seen in this image.

that scaling gain by ten enabled exploration of the entire 10:1 room. Thus, the 1:1 room was 5m by 5m, and the 10:1 room was 50m by 50m. The two environments are shown in Figures IV.1 and IV.2. In each environment, subjects were asked to memorize the location of seven objects differing by a randomly selected color (red, yellow, orange, green, blue, purple, pink, brown, white, gray, black). The targets in the 1:1 and 10:1 environments were cylinders (.1 x .1 x 1.7 m) and tables (1.1 x .7 x 1.2 m), respectively. These seven target objects were arranged in a particular configuration, such that the configuration in the 1:1 and 10:1 conditions varied only in scale (1 and 10, respectively), and by a rotation about the center axis. In this manner, the seven objects were arranged similarly in the two environments so that the angles between the target objects were preserved. Other objects were also included in the rooms in different orientations to give the subject a sense of the size and scale of the environment. The 1:1 room contained six posters, two bookshelves, two tables, two chairs, doors, and a computer. In the 10:1 room, there were 14 posters, a refrigerator, a fish tank, three sofa areas, two bookshelves, a group of six chairs, a computer desk, a computer, doors, a group of slot machines, and a pool table.

Procedure

One-half of the subjects performed the experiment in the 1:1 environment, the other half in the 10:1 environment. In both environments, there were two locomotion conditions, physical bipedal walking and joystick translation. In both environments, subjects rotated their position by turning on foot. Translation was accomplished by walking or by using the joystick. The physical walking condition of the 1:1 environment involved regular walking, while walking in the 10:1 environment involved a scaled translational gain of ten. Translational gain was defined as the rate of translational flow in the virtual environment that mapped onto a given amount of motor activity. The motor actions of walking have a natural metric, and in the 1:1 walking condition, the geometry of the system was arranged so that each meter of distance walked mapped onto a meter's worth of translation in the virtual environment. In the 10:1 condition, the translation in the virtual environment was increased by a factor of ten, so that one step in physical space corresponded to a distance of ten steps in virtual space.

In the joystick condition, participants used physical rotation and moved in the direction of gaze by joystick translations. Using a joystick does not have a natural metric; that is, a given angle of the joystick does not map onto any corresponding amount of translation. To create a reasonably natural-seeming locomotion mode, we reasoned that pushing the joystick to its furthest extent should map onto a rapid, but relatively comfortable, walking speed. In the 1:1 environment, the maximum joystick translation rate was that of normal walking, 1 m/s, while the translation rate of the 10:1 environment was 10 m/s. Subjects could go slower with the joystick just as subjects could walk more slowly than normal in the locomotion condition. One-half of the subjects did the physical walking task first, the other half did the joystick task first. The procedures were carefully explained to the subjects before they saw the virtual environments. Once the experimenter and the subject agreed that the subject understood the task, the subjects saw the layout of the virtual environment from the center of the virtual room. The subjects were instructed to learn the locations of the seven target objects without moving from this center location.

Participants' spatial knowledge was tested from six different locations. A given testing position and orientation was indicated to the subject by the appearance of red and yellow spheres in the environment. Subjects were instructed to locomote to the red sphere, position themselves underneath it and face the yellow sphere, which also occluded their view of the room. At each location, the subject completed four trials by turning to face four different target objects in the room, making 24 trials per condition. Specifically, subjects were instructed "close your eyes and turn to face the ⟨*target name*⟩." After each trial, subjects were instructed to rotate back to their starting position facing the yellow sphere. To compare the angles of correct responses across conditions, the same trials were used for each condition. The testing location and target locations were analogous in both conditions, and target colors varied randomly across the environments. The trials were designed so that the disparity was evenly distributed in the range of $20-180^{\circ}$. Once the subject reached a testing location (the red sphere), they were not allowed to look at the target objects since the objects were made invisible. They were, however, encouraged to re-orient themselves after finishing each testing position.

To assess the degree of difficulty of updating orientation relative to objects in the virtual environment, latencies, and errors were recorded. Latencies were measured from the time when the target was identified until subjects said they had completed their turning movement and were facing the target. Unsigned errors were measured as the absolute value of the difference in initial facing direction (toward the yellow sphere) minus the correct facing direction. The subjects indicated to the experimenter that they were facing the target by verbal instruction, and the experimenter recorded their time and rotational position. The time was recorded using a stopwatch, and the rotational position was recorded using the InterSense tracker. Subjects were encouraged to respond as rapidly as possible, while maintaining accuracy.

	Means (Std. Error)			
	1:1		10:1	
	Walking	Joystick	Walking	Joystick
Turning Error (°)	19.47(1.52)	25.27(2.26)	26.36(2.53)	39.50(3.06)
Latency (s)	3.93(0.18)	3.72(0.19)	4.31(0.14)	4.78(0.25)

Table IV.1: Means and standard errors of the mean for turning error and latency in the joystick and walking conditions of Experiment 3.

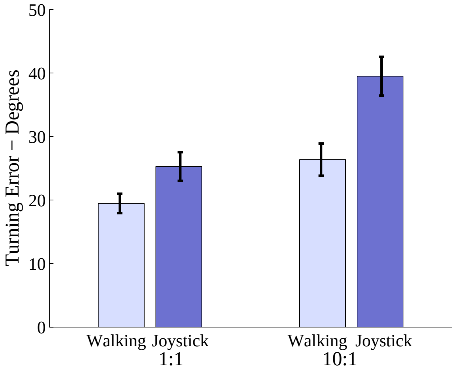

Figure IV.3: Mean turning error for the 1:1 Walking, 1:1 Joystick, 10:1 Walking, and 10:1 Joystick conditions of Experiment 3 . Error bars indicate the standard error of the mean. As discussed in Section IV.3.2, gain has a significant effect on turning error.

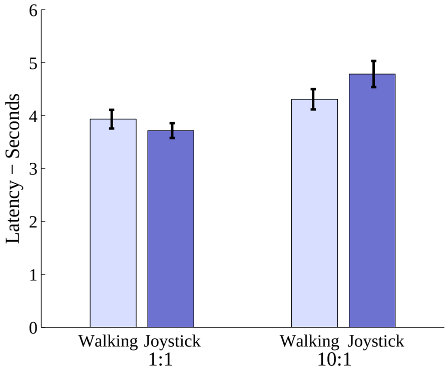

Figure IV.4: Latency for 1:1 Walking, 1:1 Joystick, 10:1 Walking, and 10:1 Joystick conditions of Experiment 3 . Error bars indicate the standard error of the mean.

IV.3.2 Results

Table V.3 shows subjects' mean turning errors and latencies by locomotion condition in the two virtual environments. A visual representation of the turning errors and latencies are shown in Figure IV.3 and IV.4, respectively. Error and latency were significantly correlated, $r = .18$, $p < .001$. Therefore errors increased as response time increased. Mean turning error as a function of disparity is shown in Figure IV.5. Mean response time as a function of disparity is shown in Figure IV.6. Note that disparity is a continuous variable and has values between 20 and 180 determined by the geometry of the experimental setup, but following the practice of May [2004], the disparities have been clustered to their closest 36° amount.

The independent variables included locomotion (walking versus manipulating a joy-

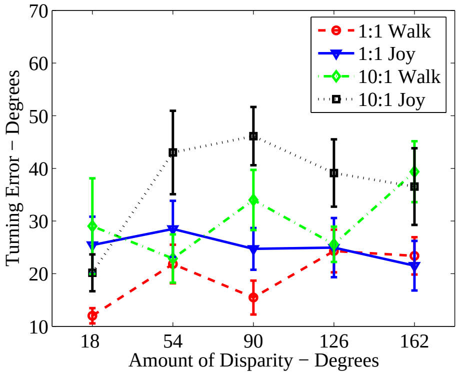

Figure IV.5: Mean turning error as a function of disparity in the 1:1 Walking, 1:1 Joystick, 10:1 Walking, and 10:1 Joystick conditions of Experiment 3 . Error bars indicate the standard error of the mean.

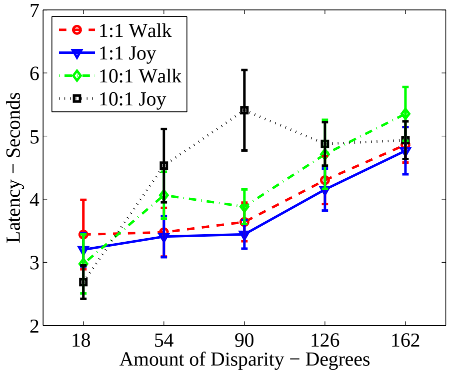

Figure IV.6: Latency as a function of disparity in the 1:1 Walking, 1:1 Joystick, 10:1 Walking, and 10:1 Joystick conditions of Experiment 3 . Error bars indicate the standard error of the mean.

stick), translational gain (1:1 versus 10:1; between-subjects), and the disparity (five categories, each spanning 36 degrees and centered, respectively, around 18°, 54°, 90°, 126°, and 162°). An analysis of variance on error, looking at effects of locomotion and translational gain revealed significant main effects and interactions. Locomotion was significant, $F(1,13) = 5.8$, $p < .05$. People made fewer errors if they explored the virtual environment physically than with the joystick. The translational gain condition was a significant factor, $F(1,13) = 9.8$, $p < .01$. Participants were more accurate in the 1:1 gain than the 10:1 gain. Finally, disparity was not significant, $F = 2.0, p = .1$—participants were not affected by the angular disparity as shown in Figure IV.5. The two-way interactions were non-significant, but the three-way interaction (of locomotion x translational gain x disparity) approached significance, $F(4,52) = 2.2$, $p = .08$.

The analysis of variance on latency for effects of locomotion, translational gain, and disparity also revealed several interesting effects. Locomotion was not significant—participants were not faster in any mode of locomotion, walking or using the joystick. The translational gain had no main effect either, $F(1,13) = .1$, $p = .9$. Participants were not faster in any gain condition—the 1:1 or 10:1 environment. There was a main effect of disparity,

$F(4,52) = 3.2$, $p < .05$. Participants' response times were affected by the angular disparity. No two-way interactions were significant, but the three-way interaction (disparity x locomotion x translational gain) was highly significant, $F(4,52) = 6.5$, $p < .001$.

To check for speed-accuracy tradeoffs, a separate analyses of covariance was done, looking at the effect of one variable on the other. The analysis of covariance on error with latency as the covariate and the same independent variables as the ANOVA revealed a main effect of the covariate, $F(1,759) = 18.6$, $p < .001$. There were significant main effects of locomotion and translational gain, $F(1,759) > 8.0$, $p < .01$. No interactions were significant.

The ANCOVA on latency with turning error as the covariate and the same independent variables as the ANOVA revealed a main effect of the covariate, $F(1,759) = 18.6$, $p < .001$. Surprisingly, there was a significant effect of translational gain, $F(1,759) = 8.8$, $p < .01$, and a significant three-way interaction, $F(1,759) = 26.8$, $p < .001$.

IV.4 Experiment 4: Effects of translational gain and subject expertise

In the second experiment, the goal was to assess how well subjects could maintain spatial awareness when the gain of translation in the virtual environment was varied relative to translation in the physical environment. More specifically, a subjects' spatial knowledge was tested in each of the three translational gain conditions: 1:1, 2:1, and 10:1. To see if experience with fast visual flow mattered, the results of six people who regularly play video games were compared to six people who do not in a sub-study. Similar to the first experiment, user orientation was tested by having subjects memorize the location of seven target objects in the room and identifying them from different positions in the room with eyes closed.

IV.4.1 Method

Participants

Eighteen subjects, thirteen Vanderbilt University graduate and undergraduate students and five non-student adults, participated in the experiment. Subjects were unfamiliar with the experiment and the virtual environments. Subjects were given compensation for their participation.

Materials

The HMD and tracker used in Experiment 3 were used again in Experiment 4. The 1:1 and the 10:1 virtual environments of Experiment 3 as shown in Figures IV.1 and IV.2 were also used in this experiment. Additionally, a 2:1 virtual environment was created specifically for the 2:1 gain condition. The 2:1 room was 10m by 10m, twice the size of the 1:1 room, and is shown in Figure IV.7. In all three environments, subjects were asked to memorize the location of seven objects differing by a randomly selected color (red, yellow, orange, green, blue, purple, pink, brown, white, gray, black). Similar to the 1:1 and 10:1 environments, the targets objects in the 2:1 environment were chairs (.8 x .6 x 1.2 m). The positions of these seven target objects were similar in the 1:1, 2:1, and 10:1 conditions, varying only in scale (1, 2, 10, respectively), and by a rotation about the center axis. Therefore, the angles between the target objects for each of three conditions were equivalent. The 2:1 room contained items to give the user a sense of scale: 12 posters, two bookshelves, doors, and a computer.

Procedure

Each of the 18 participants explored each of the environments under three different gain conditions, 1:1, 2:1, and 10:1. In all three conditions, rotation in the virtual environment matched rotation in the physical environment. In the 1:1, 2:1, and 10:1 conditions, the translational gain of the tracker was scaled by one, scaled by two, and scaled by ten, respectively. Since there were six orders of the three gain conditions, three subjects were

Figure IV.7: This figure shows the virtual environment of the 2:1 condition. The target objects in this experiment were the different color chairs, some of which can be seen in this image. The 1:1 and 10:1 environments of Experiment 4 are shown in Figures IV.1 and IV.2, respectively.

tested in each order in a counter-balanced fashion. The experimental procedure was fully explained to the subjects prior to seeing the virtual environments. During the learning phase, subjects were asked to learn the positions of the seven colored target objects while freely walking around the virtual environment. After about three minutes of study, the experimenter tested the subject by having them walk to various targets, close their eyes, and point to randomly selected targets. This testing and learning procedure was repeated until the subject felt confident that the configuration had been learned and the experimenter agreed.

The experimental design was similar to Experiment 3, yet only five testing positions were used, and the participant located three targets from each test position for a total of 15 trials. The location and orientation of the subject for a given testing position was controlled by the yellow and red spheres similar to Experiment 3 . The subject was not allowed to look at the target objects in the room when he or she was located underneath the red sphere. The testing location and target locations for each condition were analogous, and target colors varied across environments. Like Experiment 3 , these trials were also designed so that the disparity was evenly distributed in the range of 20 – 180°.

	Means (Std. Error)		
	1:1	2:1	10:1
Turning Error (°)	24.22(1.95)	28.27(2.07)	28.29(2.20)
Latency (s)	3.64(0.16)	3.95(0.15)	4.18(0.17)

Table IV.2: Means and standard errors of the mean for turning error and latency in the 1:1, 2:1, and 10:1 conditions of Experiment 2.

As a sub-study, the results of six gamers were compared to the results of six non-gamers. A gamer was defined as people who self-report that they play five or more hours of video games per week. Non-gamers were people who report that they do not currently play video games, and have never played first-person video games. For this experiment, there were six possible orders. When the original subject pool was divided to balance these orders, for a six gamer versus six non-gamer comparison, three gamer cells were empty. Three additional gamers were recruited to fill them. If while selecting six gamers from each cell, and six non-gamers from each cell, more than one subject in a given order met the gamer or non-gamer requirement, the subject from that cell was randomly selected.

IV.4.2 Results

The independent variables were translational gain (three levels, namely 1:1, 2:1, and 10:1), disparity (five levels, clustered as in Experiment 3), and subject expertise (gamer versus non-gamer). Dependent variables in this experiment were the errors and latencies in turning to face the targets. The amount of rotation varied as repeated within-subject trials and varied up to 180°.

Table V.2 shows the mean turning error and latencies as functions of translational gain. Figures IV.8 and IV.9 show a visual representation of the turning error and latency, respectively. Turning error was defined as the unsigned error, i.e., the unsigned difference between the actual target distances and the observed distances. Error and latency were significantly correlated, $r = .18$, $p < .001$. Thus errors increased as the response time on the trials increased.

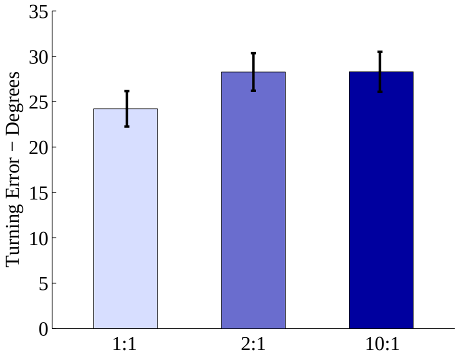

Figure IV.8: Mean turning error for the 1:1, 2:1, and 10:1 scaled translational gain conditions in Experiment 4. Error bars indicate the standard error of the mean. As discussed in the Section IV.4.2, turning errors show no significant main effect.

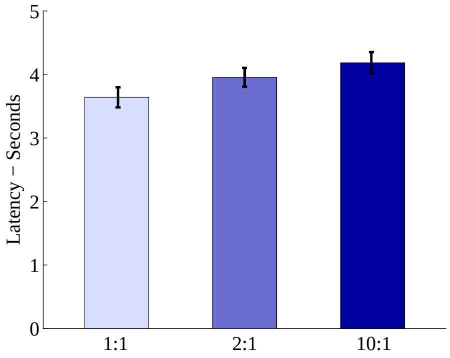

Figure IV.9: Latency for 1:1, 2:1, and 10:1 scaled translational gain conditions in Experiment 4. Error bars indicate the standard error of the mean.

Figure IV.10 shows the average turning errors as a function of disparity and translational gain. Likewise, the average latencies as a function of disparity and translational gain are shown in Figure IV.11. The analysis of variance on error with repeated measures on translational gain and disparity revealed no main effect of gain $F < 1$, $p = .63$, but a significant effect of disparity, $F(4,68) = 9.2$, $p < .001$. Participants performed equally well on the three translational gain conditions, but the angular disparity affected their accuracy. The analysis of variance on latency with repeated measures on gain and disparity showed no main effect of gain as before with error, $F = 2.1, p = .14$; disparity was significant, $F(4,68) = 8.7$, $p < .001$. Changes in the translational gain did not affect response times, but angular disparity affected participants' latencies.

With gamers and non-gamers, an analysis of variance for turning error with repeated measures on gain and disparity revealed no main effect of gain, but a significant effect of disparity, $F(4,40) = 5.1$, $p < .01$. The interaction between gain and disparity approached significance, $F(8,80) = 1.8$, $p = .08$. Participants' accuracy was affected by the angular disparity, but these effects were different in different gain conditions. There was no effect

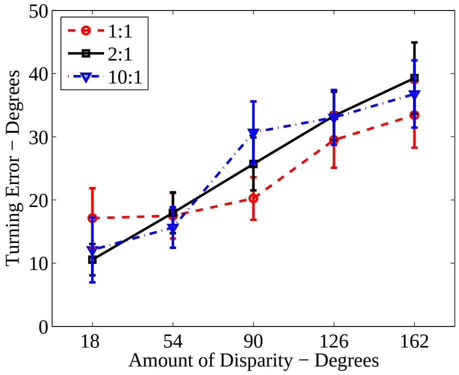

Figure IV.10: Mean turning error as a function of disparity in the 1:1, 2:1, and 10:1 conditions of Experiment 4. Error bars indicate the standard error of the mean.

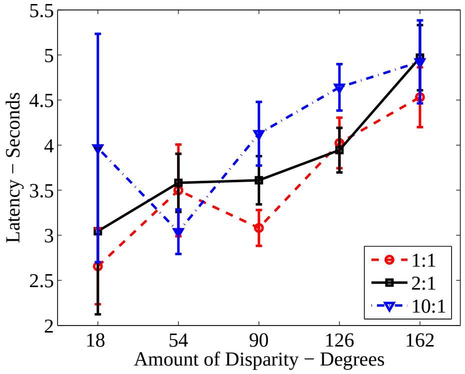

Figure IV.11: Latency as a function of disparity in the 1:1, 2:1, and 10:1 conditions of Experiment 4. Error bars indicate the standard error of the mean.

of experience (i.e., gamer or non-gamer). Thus experience with computer games did not enable participants to be more accurate on the different gain environments, nor at different angles of disparity. The analysis of variance for latency with repeated measures on gain and disparity showed no effects of gain, a main effect of disparity, $F(4,40) = 10.4$, $p < .01$, but no effect of experience. There were no significant interactions of either independent variable with experience. Thus while angular disparity affected participants' response times, being a gamer did not help subjects respond faster.

IV.5 Discussion

This chapter addresses the topic of how to to engineer systems that optimize the precision and ease with which people can fit motor exploration of large virtual environments within the confines of smaller physical rooms housing an HMD. If HMD technology is to become widespread, such an issue is important, because many potential users will not have large areas to devote to using an HMD. Research shows that spatial learning and orientation are good when people explore a virtual environment by physically turning to rotate and walking to translate their perspectives [Williams et al. 2007b].

The two experiments reported here investigated two different solutions to the problem of exploring a large virtual space in the confines of a small physical environment. Given the evidence that walking facilitates updating spatial orientation relative to physical environments [Presson and Montello 1994; Rieser 1989; Wraga et al. 2000] and virtual environments [Williams et al. 2007b], two alternative methods that varied in the amount of physical walking were examined. Both methods included physical turning to look around from a single location and rotate one's perspective. In the joystick translation condition, subjects used a joystick to translate their position through the virtual environment and physically turned to control their rotation. Thus, to explore a virtual environment they would turn and face their destination and then use the joystick to translate to it. The advantages of the joystick condition are that it includes physical movement for the rotations since these easily fit into the smallest rooms, and includes joystick translations, which also fit into the smallest rooms. In the walking translation condition people physically turned to rotate and walked to translate their perspective.

The results of Experiment 3 show two things about the locomotion interface. First, it shows that the joystick translation condition is viable, and resulted in reasonably accurate and rapid judgments. The errors in the joystick condition averaged 25.27° in the 1:1 condition and 39.50° in the 10:1 condition, which is much better than the 90° errors expected by chance. And the latencies did not reliably differ from the latencies in the walking translation condition. Second, it showed that there is value added in the walking translation condition compared to the joystick condition.

Gain, the rate at which a given action with the joystick or walking would translate the subject's perspective through the virtual environment, was also manipulated. In Experiment 3 , 1:1 and 10:1 gains were manipulated across subjects in two different conditions, and in Experiment 4 the gains (1:1, 2:1, 10:1) were manipulated across the repeated trials experienced by each subject. The effects of gain varied somewhat across the two experiments. In Experiment 3, gain exerted a significant effect on the errors, but not on the latencies. In Experiment 4, gain did not exert a significant main effect on errors or latencies. We conclude from this that varying gain is a feasible technique to use to fit the walking exploration of large virtual environments within the confines of small physical environments. We conjecture two reasons for the result that gain exerted a statistically significant effect on the errors in Experiment 3 but not in Experiment 4. One hypothesis is simply that it reflects error variance. Another hypothesis is that the difference reflects differences across the design of the two experiments. Experiment 1 tested gains between groups and Experiment 4 tested them within subjects. The higher errors in the 1:1 gain condition of Experiment 4 compared to Experiment 3 may be due to the subjects' additional experience of the higher gain conditions in Experiment 4. However this discrepancy is resolved by future work, both together demonstrate that manipulating gain is a useful way to fit large virtual environments into smaller physical environments.

For experiments 3 and 4, room sizes were scaled identically to translational gain. However, the targets across the environment remained relatively constant. There were also objects scattered throughout the 1:1, 2:1, and 10:1 environments giving the user a sense of size and scale. Thus, the optical flow was different across each environment.

Disparity, that is, the difference in the subject's facing direction at the start of a trial and the correct facing direction, exerted highly lawful, linear effects on the latencies in Experiment 3 and Experiment 4 (it did not significantly affect errors). This result makes sense from multiple perspectives; for example, it is consistent with it taking longer times to figure larger changes in angular direction and it is consistent with it taking longer times to turn through larger degrees of angle to face the target.

The gamer versus non-gamer subjects in Experiment 4 varied in their general experience with some of the features of the learning and test situation, though they did not vary in experience with the specific features of this system. Consider two ways that practice with video games could have mattered. First, the gamers' experience with first-person games could have facilitated their sensitivity to the rotations and translations in perspective that were rendered in virtual environments. Second, gamers practice controlling the locomotion interfaces in the context of their games, interfaces that typically involve manipulating a joystick or console. However, these empirical results show that the gamers and non-gamers did not significantly differ on either the latencies or errors. It is known that practice with first person games hugely facilitates the speed and accuracy of performance. Unlike Lathrop and Kaiser [2005], the results in this chapter indicate that the skills underlying these benefits do not transfer from the joystick/console interfaces and small-screen virtual environments typically provided by the games to the immersive HMD/walking system assessed in Experiment 4.

The implications for these findings in the development of artificial learning systems consisting of HMD displays, tracking systems, and walking interfaces are as follows. The results of these studies show that there is value added by using bipedal walking as the locomotion interface, compared with using a joystick. And in addition, we show that manipulating gain is a viable method to assist people in fitting exploration of large virtual environments within the confines of small physical spaces. However, we assume that there are limits to how far one can scale gain— generally comparable results for gains varying from 1:1 to 10:1 were found, but to explore a battlefield or city, one would need to use much larger gains or one would need to use an additional strategy. In this chapter, gain was limited to a factor of 10 since small head movements become distracting at gains higher than ten. Thus, the next chapter, Chapter V investigates how high gain can be scaled when a method of filtering is employed to control small local movements. However, we hypothesize that there is some limit to how high gain can be scaled. Thus, Chapter VI addresses one such additional strategy, to “reset” subjects when they walk and reach the end of their physical space. The technique presented in this thesis and a technique to viably “reset” a person’s position would present a compelling interface for the use of virtual environment technology in small physical spaces.

CHAPTER V

THE LIMITS OF SCALING TRANSLATIONAL GAIN

V.1 Introduction

This chapter further examines the issue of exploring a virtual environment that is larger than the physical boundaries of the tracking system using scaled translational gain. It looks at finding the limit to which translation gain can be scaled, and it investigates whether scaling eyeheight proportionally to gain increases spatial awareness.

When the translation gain of walking is scaled higher than 10, small body movements become more noticeable and distracting. Thus, this chapter expands the findings of the previous chapter, Chapter IV, and examines how far translational gain can be increased with the aid of engineering solutions to alleviate problems of small head movements. In a typical HMD system, the device that is used to update the position is mounted on top of the HMD. The HMD system discussed in this thesis uses a four camera tracking system that tracks an LED light that is mounted on top of the HMD. For example, at high translational gains small locomotive movements become disorienting, making it difficult to position near stationary objects in the virtual environment. For example, at a translational gain scaled by 100, one inch of movement results in approximately eight feet of movement in the virtual environment. Therefore, when people locomote at high rates of gain, a strategy needs to be developed that allows people to move locally in a natural way. Additionally, small head movements when examining the virtual environment from single location also become distracting. It is difficult and unnatural to maintain a fixed head position and rotate about that axis with the HMD. Consider the head movement of a user examining the contents of a virtual environment from a center location as in Figure V.1 where locomotion in the physical space matches locomotion in the virtual space. In Figure V.2, this same physical movement is replicated, yet the translational gain is scaled by a factor of 20. In this example, simply

Figure V.1: This figure shows a top-down view of virtual environment that is approximately 5m x 5m. A user is instructed to examine the environment and moves his head around slightly to view the world. The head motion of the user in the virtual environment is shown in yellow (emphasized by the orange arrow). Movement in this example is 1:1, which means that movement in physical space corresponds to that same movement in virtual space.

Figure V.2: This example shows the same virtual environment and user physical movement as Figure V.1. However, movement in physical space is scaled by a factor of twenty so that the small movement in physical space reflected by the yellow line in Figure V.1 becomes the yellow line in the current figure.

turning to view the contents of the room amounts to considerable locomotion in the virtual environment. Therefore, small head movements when the user is not locomoting to a new position also need to be filtered or somehow minimized.

When users experience virtual environments, they are not bound to only common experiences of the real world. For example, humans could navigate a virtual environment by flying and learning about that environment using a map-like overview. Does changing the eyeheight while locomoting with scaled translational gain aid a person in learning an environment? Does changing the eyeheight of a person change the limits to which we can expect to scale gain? Very short creatures, like ants, ordinarily perceive about 0.5 cm of translational optic flow for each step, and very tall creatures, like the mythical giant in the story "Jack and the Beanstalk" may perceive about 50m of translational optic flow for each step. Increasing the eyeheight to explore a large virtual environment could be useful when exploring an outdoors environment such as a large city. This strategy allows users to develop spatial orientation based on a map-like overview, and still gives users the proprioceptive feedback of walking unlike that virtual flying.

Several studies, [D'Zmura et al. 2000; Lambrey and Berthoz 2003], have examined human cognition and perception in non-real world situations. Foo et al. [2005] report that configurational or topological information (often referred to as survey knowledge) is inaccurate and non-euclidean. Therefore people reason about topological information or map-like information incorrectly. However, their work suggests that although people are able to construct a global sense of space using survey knowledge, they prefer to update their knowledge constantly using landmarks or route information. Thus, this method of scaling eyeheight would allow users to have a topological overview and allow updating. Thus, this chapter investigates eyeheight scaled proportionally to gain. That eyeheight is potentially an important factor is motivated by the work of Warren [Warren 1984], who studied the relation of eyeheight, the perception of the environment, and a subject's action system.

V.2 Method

To investigate how high gain can be scaled, a method of scaling gain while minimizing these disorienting movements was devised. Informal user studies of participants at unfiltered high gain (100:1 and 50:1) revealed that small head movements were disorienting. More specifically, disorientation seemed to occur when the user's locomotion was minimal and they were simply trying to either perform a local task such as move a few feet, or observe the environment. Participants also reported that large gain factors seem more natural and much less disorienting if their own physical locomotion was above a certain rate. Thus, we sought a method to minimize this effect by targeting the problem of disorientation when gain is scaled by large factors at slow speeds.

In the experiments presented in this chapter, users "ramp-up" to high gain based on the magnitude of their velocity, or speed. When users are not moving, but simply observing an environment, then their speed is low and the translational gain is also low. As they begin to locomote, their speed is increasingly scaled up to the desired gain. We refer to this method as nonlinear translational gain. In this nonlinear condition, once users reach a critical speed threshold all movements are scaled linearly by a scaling factor (or simple linearly scaled translational gain). Speeds below the critical threshold are scaled nonlinearly according to a pre-specified function. Thus, for physical speeds between zero and the critical threshold speed, virtual speed is obtained by scaling physical speed according to this function. Suitable functions should be strictly monotonically increasing with an initial value equal to zero (for zero speeds) and value at the threshold equal to the threshold multiplied by the high gain scaling factor. An example of such a function is seen in Figure V.3. In this figure, speeds above the critical threshold of 0.5 m/s are scaled by a factor of 100. Speeds below 0.5 m/s are scaled according to a cubic function. User speed is calculated every time the graphics is updated, which was 60 Hz. Speed is defined as the distance between the user's position at the time of the graphics refresh (p_x, p_z) and the position of the preceding graphics refresh (p'_x, p'_z) divided by the refresh rate, *refreshRate* . To calculate the distance traveled we simply use the user's position in the *x* and *z* directions and do not take into account the *y* direction, which represents the user's viewing height. Thus, speed is calculated as follows:

$$speed = \frac{\sqrt{(p_x - p'_x)^2 + (p_z - p'_z)^2}}{refreshRate}. \tag{V.1}$$

It is important to note in "high gain mode" when gain is linearly scaled, calculating the new virtual position involves scaling the speed by the high gain amount. Thus, in high gain mode the virtual position in the new *x* and *z* positions in virtual space, v_x and v_z, can be obtained from positions of the user at the previous and current frames:

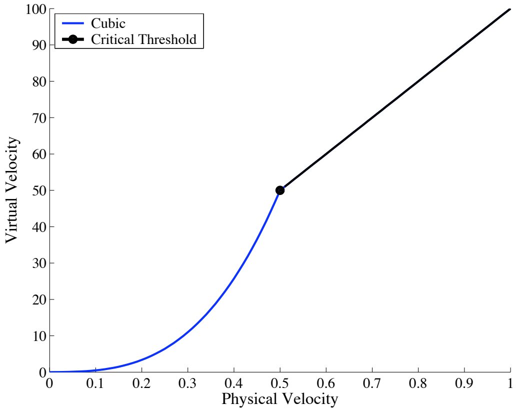

Figure V.3: This is figure shows a ramping cubic function. The critical threshold of 0.5 m/s is shown in black. After users reach above this threshold speed, gain is linearly scaled by 100.

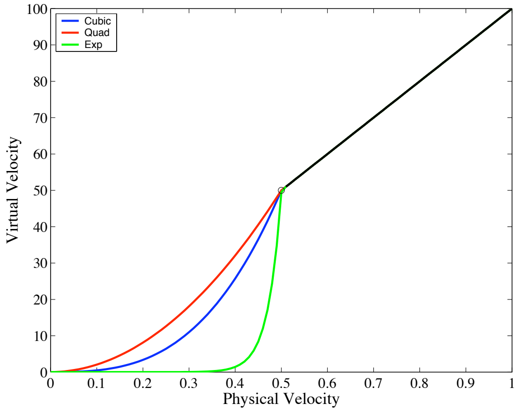

Figure V.4: This shows all three ramping functions that were evaluated in Experiment 6. In this figure, high gain was equal to 100, so at the threshold physical speed value of 0.5 m/s, virtual speed was 50 m/s. For speeds above 0.5m/s gain was scaled by 100.

$$v_x = v'_x + (p_x - p'_x) * scale, \tag{V.2}$$

$$v_z = v'_z + (p_z - p'_z) * scale, \tag{V.3}$$

where v'_x and v'_z represent virtual position from the previous frame.

There are many functions that meet the requirements for a ramping function, and beyond these requirements our goal was to select one which was most pleasing from a user's perspective. Additionally, the value of the critical threshold itself needs to be determined. Two experiments were designed to validate engineering choices for both the threshold and ramping functions. First, Experiment 5 examines the critical speed threshold at which a user should enter into linearly scaled high gain or linear gain. Experiment 6 evaluates three plausible functions used to scale speeds smaller than the critical threshold: an exponential, a cubic polynomial, and a quadratic polynomial.

Before discussing the experiments, formulas for the ramping functions are derived.

However, in the experiments there is a chicken-and-egg problem in that a ramping function cannot be derived without knowing the critical threshold, and determining a critical threshold assumes the use of some form of ramping function. In this work we do not examine this question exhaustively. Rather, we assume a cubic ramping function to determine the critical threshold, then assume this threshold is the best value for testing different ramping functions.

For both Experiment 5 and Experiment 6, gain was set at 100:1 because informal pilot studies indicated that this gain was a reasonable guess of the largest scale factor that subjects could maintain reasonable spatial awareness. If this method worked well for 100:1, then we assumed that it would work for smaller gains easily.

V.2.1 Mathematical Derivation of Ramping Formulae

The details are now discussed. The mathematical details below describe the simple cubic function that is pictured in Figure V.3. Below the critical threshold, the virtual speed, s_v, is described in terms of physical speed, s_p as follows:

$$s_v = s_p + c_1 (s_p)^3, \tag{V.4}$$

where c_1 is a constant whose value depends on the gain level. Thus, the value of c_1 changes with each gain level. Above the critical threshold gain is scaled is directly by the high gain amount. We use this form of the cubic since it has a desirable slope and it passes through (0,0) as shown in Figure V.3. In other words, at physical speed of 0, virtual speed is also 0. As an example we solve for c_1 at 100:1 gain and a critical threshold value of 0.5m/s. The refresh rate of the graphics and tracking system has a direct impact on the values of the constants found in the above equation. For purposes of this example, let us assume that tracking updates every 1s. At 0.5m/s speed should be scaled by 100, and values under 0.5m/s should be scaled according to the cubic function. We know that at a physical speed of 0.5m/s the virtual speed should be 50m/s ($0.5m/s * 100$). Thus, plugging in two known values, $s_p = 0.5$ and $s_v = 50$ gives us

$$50 = 0.5 + c_1(0.5)^3, c_1 = 396. \tag{V.5}$$

Thus, we scale gains lower than 0.5m/s according to the following function:

$$s_v = s_p + 396(s_p)^3, \tag{V.6}$$

which, again, is plotted in Figure V.3. In our system, the graphics are refreshed every 60 Hz. Therefore the constants change. Let us look again at the cubic function at 100:1 gain. Since we are updating every 1/60 of a second the graphics we would like a speed of $\frac{1}{60} * 0.5$ (or 0.0083) to map to $\frac{1}{60} * 50$ (or 0.8333) since each frame is $\frac{1}{60}$ of a second. Thus we solve for c_1 with these values $s_p = 0.0083$ and $s_v = .8333$ and find that the value of c_1 at 100:1 gain, a critical threshold of 0.5m/s and a refresh rate of 60Hz is $1.4256e+06$.

The constants for the quadratic and exponential ramping functions at each of the gain levels are found in a similar manner. The quadratic function we evaluated was:

$$s_v = s_p + c_1(s_p)^2, \tag{V.7}$$

and the exponential had the form

$$s_v = s_p + c_1 e^{c_2 s_p} - c_1. \tag{V.8}$$

We chose an exponential function simply by finding a function that scaled speed very little at low speeds and then drastically increased. We wanted the exponential function to be flat or have a small slope at small speeds so that gain would be scaled by a minimal amount.

The values of all the constants for a 1/60 refresh rate are shown in Table V.1. The three functions are plotted in Figure V.4 .

Table V.1: Values of the Constants

Gain	Cubic	Quad	Exp
10	$c_1 = 129600$	$c_1 = 1080$	$c_1 = 1/433.794, c_2 = 433.794$
25	$c_1 = 345600$	$c_1 = 2880$	$c_1 = 1/575.341, c_2 = 575.341$
50	$c_1 = 705600$	$c_1 = 5880$	$c_1 = 1/677.594, c_2 = 677.594$
100	$c_1 = 1.4256e+06$	$c_1 = 1.1880e+04$	$c_1 = 1/776.954, c_2 = 776.954$

V.3 Experiment 5: Finding the Critical Threshold

The purpose of this experiment was two-fold. First, this experiment investigates the speed at which users should switch from speed scaled by a function (resulting in a scaling less than the high-gain scaling value) to the linearly scaled high-gain speed (or linearly scaled translational gain). This experiment examines two critical speed threshold values: 0.5 m/s and 1 m/s and compares these results to linearly scaled translation gain where there are no critical values and gain is simply scaled by the high-gain amount. Thus, the second objective of this experiment is to formally evaluate the use of this "ramp-up" function and investigate whether users feel that problems with disorienting small head movements have become negligible with the proposed method.

In this experiment the high gain value or the highest scaled value of translational gain was fixed at 100:1 (where one step in physical space corresponded to 100 steps in virtual space). The scaling function used to scale speeds lower than the critical threshold speed value was a cubic polynomial as shown in Figure V.3.

Subjects were asked find and read three different Snellen eye charts as shown in Figure V.5, which were arranged on the sides of buildings in a large outdoors environment. An example of the Snellen eye chart on the side of a building in the environment is shown in Figure V.6. The ease of reading these charts allowed subjects to report a subjective measurement of the ease of localized movements or local locomotion in each condition. They were also asked to find and walk to a series of seven objects in the virtual environment that were a considerable distance apart. This task allowed subjects to report the ease of large-scale locomotion through the entire environment, which is referred to as global locomotion.

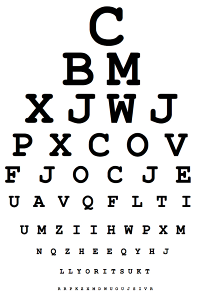

Figure V.5: This is an example of a randomly generated Snellen eye chart used in both Experiment 5, Section V.5, and Experiment 6, Section . Font size decreases with each row. Thus, the chart becomes harder to read after each row. This type of chart is commonly used for human eye exams.

Figure V.6: This shows an example of a Snellen chart in the virtual environment.

They were also asked to report if they felt nauseous or sick and if they felt any sense of unbalance.

Upon completing the experiment, they were asked to indicate which condition they liked best. They were also asked a specific question about the linear/nonlinear gain conditions. Informal pilot studies indicated that scaling gain at high speeds was not disorienting, or speeds above a certain threshold. To solidify this finding, subjects were specifically asked if side-to-side movement while walking at high gain was disorienting. For example, during normal walking, people tend to shift their body from side to side. At high translational gains this side-to-side movement could cause some disorientation and make it difficult to walk a straight path in the virtual environment.

V.3.1 Participants

Six subjects participated in the experiment. Subjects were unfamiliar with the experiment and the virtual environment. Subjects were given compensation for their participation.

Figure V.7: This figure shows the virtual environment used in Experiments 5, 6 and 7. The (x,z) position at which this image was obtained is the same in Figures V.8, V.9, V.10 and V.11. This figure represents a y position of normal eyeheight (1:1) which we approximate at 1.67m.

V.3.2 Materials

The virtual world was viewed through a full color stereo NVIS nVisor Head Mounted Display with 1280 x 1024 resolution per eye, a field of view of 60° diagonally, and a frame rate of 60 Hz. The HMD weighs approximately 1 kg. An InterSense IS-900 tracker was used to update the participant's rotational movements around all three axes. Position was updated using two optical tracking cameras with an accuracy of < 0.5 cm over 3 x 3 x 3 m volume and an update rate of 60 Hz.

The size of the physical room in which the experiments were performed was approximately 5m by 6m, and within the room the limits of the video position tracking system was approximately 5m by 5m. The same 650m x 650m large outdoors environment was used in each of the conditions. The size of the Snellen eye charts that participants were instructed to read were approximately 0.7m by 0.7m and they were randomly located on the sides of buildings that appeared in the environment. The Snellen eye chart was randomly generated for each trial. The environment is pictured in Figure V.7. Buildings and other objects were scattered throughout the environment. These objects were of natural shape and size and were items that you would expect to see outdoors. These objects gave the user a sense of size and scale. Larger objects are positioned further away from the center of the environment and smaller objects were closer to the center enabling the viewing of all objects from the center of the room. The seven target objects that the subjects had to walk to varied by trial but were such things as the front door of the cathedral, the water tower, the swing set, the entrance to the Panera, the front of the hotel, the parking meter, the police car, the entrance to the cathedral, etc.

V.3.3 Procedure

There were three conditions in this experiment: 0.5 m/s critical threshold speed, 1 m/s critical threshold speed, and linear gain scaled by 100 (or, a critical threshold speed of 0 m/s). Two conditions use a cubic polynomial to scale gain until a critical threshold speed is reached, then gain is simply scaled by 100. If speed drops below the critical value, then gain is again scaled according to the cubic function. Again, speed was calculated every 60Hz, which was the refresh rate of the tracking system. Each of the six participant explored each the environment under the three different critical thresholds (0.5 m/s, 1 m/s, and 0 m/s). In all three conditions, rotation in the physical environment matched that of the physical environment. Since there were six orders of three different critical threshold speeds, one subject was tested in each order in a counter-balanced fashion. The experimental procedure was explained to the participant prior to viewing the virtual environment. Subjects were told what condition they were experiencing and were instructed to walk freely around the environment to familiarize themselves with the gain and the critical threshold of that condition. When the subject indicated to the experimenter that they felt comfortable with the environment, they were instructed to find the first Snellen eye chart and read as many lines down the Snellen chart as they felt comfortable. The subjects were allowed to position themselves as close to the Snellen chart as possible, and reading the smallest rows generally required subjects to be about two feet away from the Snellen chart. After they had read as many rows as possible, they were instructed to find the second Snellen chart and read that set of letters, and so on for the third Snellen chart.

After they had read as much of that chart as possible, they were then asked to find and locomote to seven different objects in the environment. The objects were far enough apart so that subjects were required to reach above the critical threshold speed and locomote at high gain to reach the objects. If subjects walked too slowly in the environment to reach an object, a situation could occur where they could not reach that object because they reached the limits of the tracking system first (or reached a physical wall). We refer to this error as an out-of-range target error. When this error occurred, the experimenter would slowly lead the subject backward in the physical environment so that they were moving at low gains backward in the virtual environment. This was done until the experimenter felt that the subject had enough tracking space to reach the target object. This issue only had the potential to occur in the nonlinear conditions (or when there was a critical value equal to 0.5 m/s or 1 m/s). The frequency of this occurrence was recorded.

The speed and accuracy of reading the Snellen chart was also recorded. The subject indicated to the experimenter that they were ready to read the chart. The experimenter then began timing the subject reading the Snellen chart and stopped the timer when the subject was finished reading the chart or when they indicated that they could no longer read the rest of the chart. Time was recorded using a stopwatch and the positional accuracy was obtained using the tracking system.

After completing each condition, subjects were asked to rate the following on a scale from 1 to 10 local control, global control, sense of sickness, and sense of balance. Upon completing all three trials and the post-trial questions, subjects were asked to indicate what condition they preferred. They were also asked specifically if they found the scaling of side-to-side movement at high gain disorienting.

Critical Threshold	Mean User Ratings			
	Local Control	Global Control	Sickness	Unbalanced
0	1.5 (0.5)	7.2 (1.5)	5.8 (2.4)	4.1 (1.8)
0.5	7.8 (1.3)	8.2 (1.3)	1.3 (0.5)	1.8 (0.7)
1	8.1 (1.1)	6.1 (2.3)	2.1 (0.7)	2.4 (0.9)

Table V.2: This table shows the mean ratings of the post-condition test. More specifically, after experiencing each condition varying by critical threshold value, subjects were asked to rate the local control of their movement, the global control of their movement, their feeling of sickness, and their feeling of unbalancedness on a scale from 1 to 10. One represents a feeling of "No" local control, global control, sickness, or unbalancedness. Ten represents a strong feeling of local control, global control, sickness, or unbalancedness. Standard errors are indicated by parentheses.

V.3.4 Results

The results of the post-condition tests are shown in Table V.2. In the 0.5 m/s critical threshold condition, subjects felt the highest global control or sense of being able to control traveling around the environments for greater distances. They also felt control over local movements or locomotion needed to travel short distances. Participants felt the highest control over local movements with a 1 m/s critical threshold speed, yet their sense of global control was considerably less using the 0.5 m/s critical threshold. The linearly scaled gain (or 0 m/s critical threshold speed) provided very little local control and reasonable global control. The linearly scaled gain condition made people feel nauseated and altered their sense of balance. People rarely felt these effects in the other two nonlinear gain conditions.

When asked to rate which method they prefer best, four of the six participants preferred a critical threshold of 0.5 m/s, while the other two preferred the 1 m/s critical threshold. One of the subjects that preferred the 1 m/s over the 0.5 m/s condition found reading the Snellen charts easier in the 1 m/s condition yet preferred 0.5 m/s for walking long distances. Overall, subjects found the 0.5 m/s felt "most natural" for doing both local and global locomotion.

Interestingly, four of the six subjects in the 1 m/s condition had problems reaching their target objects in a few of their trials because they did not travel fast enough and ran out of tracking space. This out-of-range target error only occurred once in the 0.5 m/s critical threshold condition across all of the subjects.

As for reading the Snellen charts, in the 0.5 m/s condition, it took participants on average 105 seconds to read the chart with an average of 0.3 mistakes. Which meant on average, subjects did not make a mistake reading the chart. However, after reading approximately three charts, they would be likely to make a mistake. Similarly, for the 1 m/s critical threshold value, Snellen charts were read at an average of 111 seconds and were done so with an average of 0.28 mistakes per chart. In the linearly scaled gain condition, no subject was able to read the last three lines of the Snellen chart. On average, they could complete a few letters on the fourth to last line, but usually stopped because they felt uncomfortable.

At the end of the experiment subjects indicated whether they felt side-to-side movements while walking at high gain was disorienting. None of the subjects found this disorienting or thought any method of filtering needed to be employed.

We find that a critical value of 0.5 m/s is best since it provides a nice compromise between global and local control. Users can travel longer distances with little physical space, yet small head movements are not as distracting and disorienting as the linearly scaled gain. We also found that the 0.5 m/s threshold resulted in little or no sickness. Users also had the best sense of balance as compared to the 1 m/s and no threshold value. Thus, we use a critical value of 0.5 m/s for the remainder of this Chapter. Future work involves using a more exhaustive experiment to find a more precise value of the critical threshold. However, given the good user evaluations of this method, we feel that 0.5 m/s represents a reasonable critical threshold.

V.4 Experiment 6: Finding the "Ramping" Function

V.4.1 Participants

Six subjects participated in this experiment and were given compensation for their participation. The subjects were unfamiliar with the experiment and the virtual environment.

Ramping Function	Mean User Ratings			
	Local Control	Global Control	Sickness	Unbalanced
Quadratic	6.9 (1.9)	8.3 (1.1)	3.4 (1.8)	1.4 (0.5)
Cubic	7.9 (1.5)	8.1 (1.2)	1.4 (0.4)	1.8 (0.5)
Exponential	8.3 (1.3)	8.5 (0.9)	1.3 (0.4)	1.7 (0.5)

Table V.3: This table shows the mean ratings of the post-condition test of Experiment 6. More specifically, after experiencing each condition using one of the ramping functions, subjects were asked to rate the local control of their movement, the global control of their movement, their feeling of sickness, and their feeling of unbalancedness on a scale from 1 to 10. Standard deviations are noted in parentheses. One represents a feeling of “No” local control, global control, sickness, or unbalancedness. Ten represents a strong feeling of local control, global control, sickness, or unbalancedness.

V.4.2 Materials

The materials used in this condition were exactly the same as Experiment 5.

V.4.3 Procedure

The procedure for this experiment was almost the same as Experiment 5. However, the difference was that participants experienced different ramping functions in each of the three conditions. The critical threshold speed was fixed at 0.5 m/s. Additionally, in this experiment they were not told which condition they were experiencing. They were again asked to read three Snellen charts and locomote to seven target objects. After each condition, subjects rated their experiences. After completion of all three conditions, subjects indicated which condition they preferred best.

V.4.4 Results

The results of the post-condition questionnaire are presented in Table V.3. In all of the conditions, subjects felt a high amount of global control and local control. The quadratic function had the lowest local control. From observing the three functions in Figure V.4, we can see that gain is scaled higher at smaller speeds for the quadratic function than the other two functions. People felt a slight sense of sickness in the quadratic condition as well, an effect that was not observed with the cubic and exponential functions.

Since subjects where not told what condition they were experiencing, they were asked which condition they like best in the order that they experienced them. More specifically, did they prefer the "first," "second," or "third" condition? Four of the Six participants preferred the exponential function, while the other two preferred the cubic function.

The average time to read the Snellen chart in the exponential condition was 112 seconds and the average time to read the cubic was 109 seconds, which were almost the same. On average, in the quadratic condition, participants were unable to complete the reading of the last line of the chart.

Again, subjects were asked about the side-to-side movement when speed is linearly scaled in high gain and it was also not an issue in the experiment.

Overall, the exponential function performs best: as compared to the other two methods, it seems to give the user the highest amount of global and local control. If you look at the graph of the exponential curve in Figure V.4 as compared to the other two functions, the exponential has smaller slope at small speeds which gives it an increased local control. Thus, our nonlinear scaling method involves an exponential "ramping" function with a 0.5 m/s critical threshold.

V.5 Experiment 7

Having selected the ramping function and threshold, we are now in a position to examine the limits of scaling translational gain. Thus, in this experiment, the goal was to assess how well subjects could maintain spatial awareness when the gain of translation in the virtual environment was varied relative to translation in the physical environment. More specifically, we wanted to find the limit to which gain can be scaled under three different conditions: linearly scaled gain, nonlinearly scaled gain, and linearly scaled gain with eyeheight scaled. The subjects' spatial knowledge was tested in each of the five translational gain conditions: 1:1, 10:1, 25:1, 50:1, and 100:1. To test subject orientation, the subjects were asked to remember the location of five objects in the environment, then were asked to move themselves to a new point of observation and instructed to turn to face the targets from memory without vision. Each subject performed the task in each of the five gain scales under one of three conditions: linearly scaled gain, nonlinearly scaled gain, and linear gain scaled proportionally to eyeheight.

V.5.1 Participants

Forty-five subjects participated in the experiment. Subjects were unfamiliar with the experiment and the virtual environment. Subjects were given compensation for their participation.

V.5.2 Materials

The same HMD system that was used in Experiments 5 and 6 was used in this experiment. Also the same 650m x 650m large outdoors environment was in this experiment as well for all of the gain conditions. Figures V.7, V.8, V.9, V.10 and V.11 show the virtual environment used in this experiment. These figures give a glimpse of the virtual environment at each of the different scaled eyeheights. The explorable region of the virtual environment changed according to the size of the gain in each of the different conditions. The size of the explorable region in the 10:1 condition was 50m by 50 m or 10 times the size of the explorable region in the 1:1 condition. Similarly, the virtually explorable region for the 25:1, 50:1, and 100:1 conditions was 125m x 125m, 250m x 250m, and 500m x 500m, respectively. In each environment, subjects were asked to memorize the location of five objects differing in shape and size. An example of one of the five objects in the 1:1 environment was a fire hydrant. Example objects in the 10:1, 25:1, 50:1, and 100:1 environments include a picnic table, an 18-wheel truck, a church, and a tall hotel, respectively. These five target objects were arranged in a particular configuration, such that the configuration in the 1:1, 10:1, 25:1, 50:1, and 100:1 conditions varied only in scale (1,10, 25, 50, and 100, respectively), and by a rotation about the center axis. In this manner, the five objects were arranged similarly in the two environments so that the angles between the target objects were preserved.

Figure V.8: This figure shows the virtual environment used in Experiments 5, 6 and 7. The (x,z) position at which this image was obtained is the same in Figures V.7, V.9, V.10 and V.11. This figure represents a y position of 10 times normal eyeheight (10:1) which is approximate at 16.7m. Gaze is directed downward by about 20°.

Figure V.9: This figure shows the virtual environment used in Experiments 5, 6 and 7. The (x,z) position at which this image was obtained is the same in Figures V.7, V.8, V.10 and V.11. This figure represents a y position of twenty-five times normal eyeheight (25:1) which is approximate at 41.7m. Gaze is directed downward by about 30°.

V.5.3 Procedure

One-third of the subjects performed the experiment in the linearly scaled gain condition, one-third performed the experiment in the nonlinearly scaled gain condition, and the last third performed the experiment with linear gain and eyeheight scaled proportionally. Translational gain was defined as the rate of translational flow in the virtual environment that mapped onto a given amount of motor activity. In all three conditions, rotation in the virtual environment matched rotation in the physical environment. In the 1:1, 10:1, 25:1, 50:1, and 100:1 conditions, the translational gain of the tracker was scaled by one, scaled by 10, scaled by 25, scaled by 50 and scaled by a 100, respectively. Since there were 120 orders of the five gain conditions, subjects were tested in a pseudo-balanced fashion. More specifically, we counterbalanced the orders using a Latin square design. A Latin square is an $n \times n$ array, where each cell is in the array contains one of the n conditions such that each symbol

Figure V.10: This figure shows the virtual environment used in Experiments 5, 6 and 7. The (x,z) position at which this image was obtained is the same in Figures V.7, V.8, V.9 and V.11. This figure represents a y position of fifty times normal eyeheight (50:1) which is approximate at 83.5m. Gaze is directed downward by about 35°.

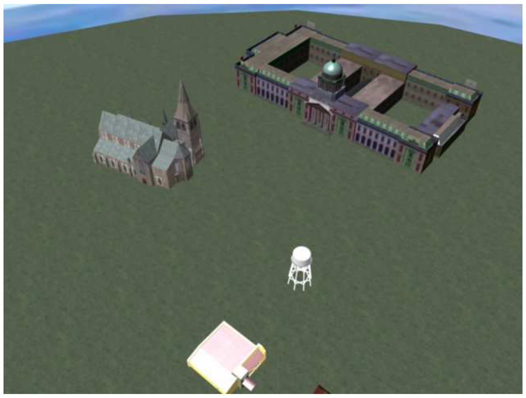

Figure V.11: This figure shows the virtual environment used in Experiments 5, 6 and 7. The (x,z) position at which this image was obtained is the same in Figures V.7, V.8, V.9 and V.10. This figure represents a y position of fifty times normal eyeheight (100:1) which is approximate at 167m. Gaze is directed downward by about 40°.

occurs only once in each column and only once in each row. We had five gain conditions and had 15 subjects, thus, we used three Latin squares to counterbalance our testing.

The experimental procedure was fully explained to the subjects prior to seeing the virtual environments. After about three minutes of study, the experimenter tested the subject by having them walk to various targets, close their eyes, and point to randomly selected targets. This testing and learning procedure was repeated until the subject felt confident that the configuration had been learned and the experimenter agreed.

Participants' spatial knowledge was tested from five different locations. A given testing position and orientation was indicated to the subject by the appearance of tall red rod and an avatar in the environment. Subjects were instructed to locomote to the red rod, position themselves near it and face the avatar. At each testing location, the subject completed three trials by turning to face three different target objects in the environment, making 15 trials per condition. Specifically, subjects were instructed "close your eyes and turn to face the ⟨*target name*⟩." After each trial, subjects were instructed to rotate back to their starting

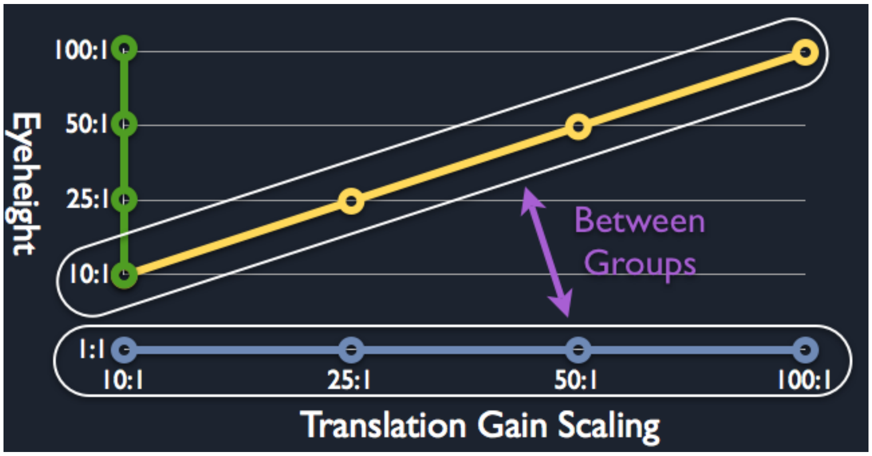

Figure V.12: This figure shows the five eyeheight levels versus the four different scaled translational gains. We choose to do a between groups analysis between scaling gain at natural eyeheight as indicated by the blue line and scaling gain proportional to eyeheight as shown in yellow. The green line represents a possible experimental design.

position facing the avatar. To compare the angles of correct responses across conditions, the same trials were used for each condition. The testing location and target locations were analogous in all conditions. The trials were designed so that the disparity was evenly distributed in the range of $20 - 180°$. Once the subject reached a testing location (the red rod), they were not allowed to look at the target objects as the objects were made invisible. They were, however, encouraged to re-orient themselves after finishing each testing position and locomoting to the next test position.

In the eyeheight condition, gain was scaled proportionally to eyeheight. In the 10:1, 25:1, 50:1, and 100:1 conditions users experienced the environment from a new viewing height. The target objects appeared smaller to the user since their eyeheight was elevated. Moreover, targets were observed by looking down. In this experiment eyeheight and gain were coupled. Figure V.12 represents eyeheight versus gain at the various scaling levels and shows different potential experimental designs that we considered. We chose to run the experiment plotted in yellow; that is, scaling gain proportionally to eyeheight. There are advantages and disadvantages to running this design. We could have fixed gain, running an experiment such as the one plotted in green, but findings would have been specific to a particular gain. Thus, an advantage of investigating eyeheight scaled proportionally to gain is that we are not limiting ourselves to findings relative to a particular gain. Another possible experimental design was to fix eyeheight and vary the gains, but Experiment 7 already gives us results for eyeheight fixed at one eyeheight, natural eye level (blue line in Figure V.12). Thus, we felt that we could gain the most knowledge in a practical experiment by scaling gain proportional to eyeheight since the line plotted in yellow adds the most amount of information to the diagram. However, the disadvantage of running this experiment is that eyeheight and gain are confounded.

To assess the degree of difficulty of updating orientation relative to objects in the virtual environment, latencies and errors were recorded. Latencies were measured from the time when the target was identified until subjects said they had completed their turning movement and were facing the target. Unsigned errors were measured as the absolute value of the difference in initial facing direction (toward the avatar) minus the correct facing direction. The subjects indicated to the experimenter that they were facing the target by verbal instruction, and the experimenter recorded their time and rotational position. The time was recorded using a stopwatch, and the rotational position was recorded using the InterSense tracker. Subjects were encouraged to respond as rapidly as possible, while maintaining accuracy.

V.5.4 Results

Figures V.13 and V.14 show the mean errors and latency collapsed across gain in the linearly scaled gain, nonlinearly scaled gain, and eyeheight condition. Figures V.17, V.18, V.15, V.16, V.19, and V.20 show the mean turning error and latency across different subjects, in the different experiment conditions (linear and nonlinear), and with different levels of translational gain (1:1, 10:1, 25:1, 50:1, and 100:1).

The linear and nonlinear gain data of this experiment were analyzed with five gain con-

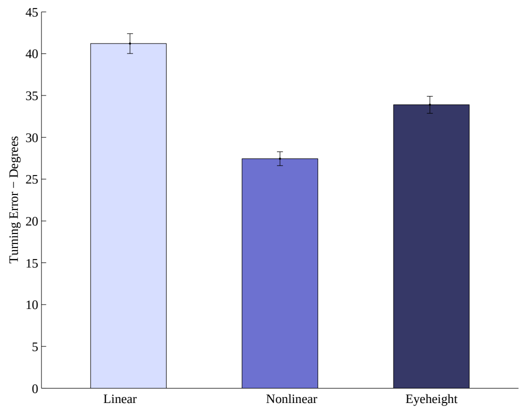

Figure V.13: This figure represents the mean turning error of each of the three different experimental conditions: Linear, Nonlinear, and scaled Eyeheight. Error bars show standard errors of the mean.

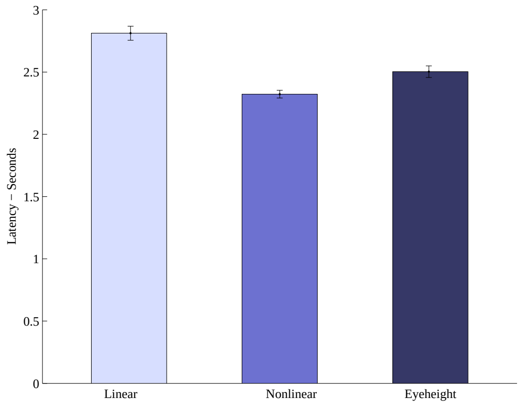

Figure V.14: This figure represents the mean latency of each of the three different experimental conditions: Linear, Nonlinear, and scaled Eyeheight. Error bars show standard errors of the mean.

ditions. We first examine the effects of the levels of translational gain in the two different experimental conditions of linear and nonlinear gain. All subjects were tested on different levels of translational gain, hence gain was a within-subjects factor; subjects were tested in one of the two experimental conditions, hence experiment condition was between-subjects. Separate analyses were done for each of the two dependent variables, turning error and latency. A multivariate repeated measures analysis on mean turning error showed main effects of gain, $F(4,112) = 10.6$, $p < .001$, experiment condition, $F(1,28) = 13.3$, $p = .001$, and a significant interaction of the two, $F(4,112) = 2.6, p = .05$. Participants errors were greater in the 1:1 and 100:1 gain levels, as well as in the linear gain experiment condition, than in other gain levels or in the nonlinear gain condition. Planned comparisons revealed that in the nonlinear gain condition, turning errors in the 1:1 gain level were significantly different from errors in the 10:1, 25:1, and 50:1 levels, but not from the 100:1 level. Interestingly, in the linear gain condition, errors on the 1:1, 10:1, 25:1, and 50:1 were all significantly different from errors on the 100:1 gain level. A similar within subjects analyses on mean latency showed a main effect of gain, $F(4,112) = 3.7$, $p < .05$, a marginal

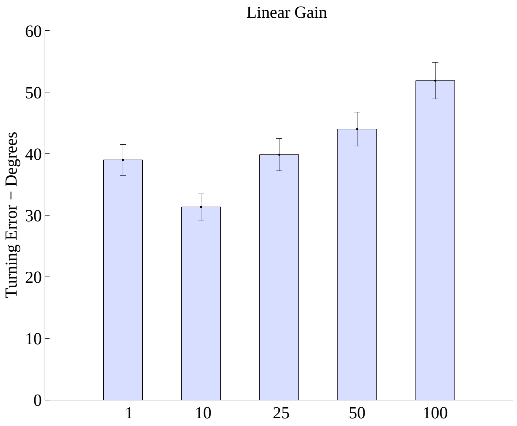

Figure V.15: This figure shows the mean turning errors in the Linear Gain condition for each of the translational gains (1, 10, 25, 50, 100). Error bars represent standard errors of the mean.

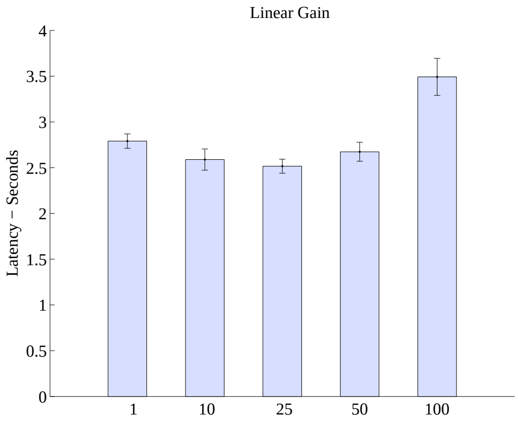

Figure V.16: This figure shows the mean latencies in the Linear Gain condition for each of the translational gains (1, 10, 25, 50, 100). Error bars represent standard errors of the mean.

effect of the experiment condition, $F(1,28) = 3.9$, $p = .06$, and no significant interaction. In both the linear and nonlinear gain, participants were faster on the 10:1, 25:1, and 50:1 gain levels, and slower on the 1:1 and 100:1 levels. These differences were significant in the nonlinear gain condition but not in the linear gain condition.

Analyses with order, experiment condition, and gain levels follow. We used three Latin squares to complete a counterbalanced array for 15 subjects at 5 different conditions. Thus, three subjects from each group had performed the experiment first in a given condition. A mixed model analysis on the dependent variable turning error, with translational gain levels (1:1, 10:1, 25:1, 50:1, and 100:1) and order (1:1 first, 10:1 first, 25:1 first, 50:1 first, 100:1 first) within group, and experiment condition (eyeheight, linear, nonlinear) between groups, showed a main effect of gain $F(4,120) = 9.7, p < .001$; a main effect of order $F(4,30) = 2.6$, $p = .05$, and a main effect of condition $F(2,30) = 7.4$, $p < .005$. Only the gain by condition interaction was significant, $F(8,120) = 2.9$, $p < .05$. Participants were liable to make more errors on the 1:1 and 100:1 gain levels, more errors when they had the 10:1 gain level first in the eye-height condition, (one-way $F(4,10) = 4.1$, $p < .05$); and the 50:1 gain level first in the linear gain condition, (one-way $F(4,10) = 5.5$, $p < .05$). Overall

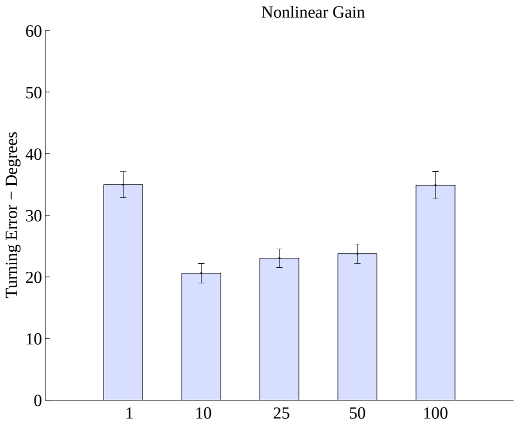

Figure V.17: This figure shows the mean turning errors in the Nonlinear Gain condition for each of the translational gains (1, 10, 25, 50, 100). Error bars represent standard errors of the mean.

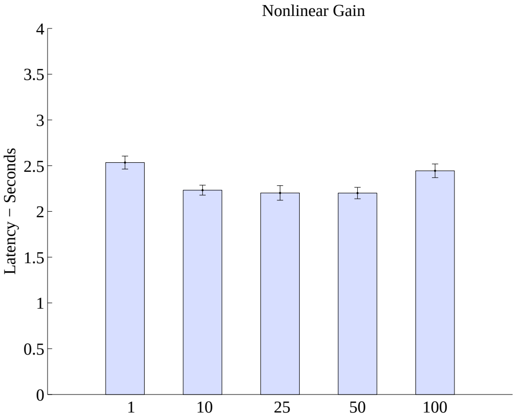

Figure V.18: This figure shows the mean latencies in the Nonlinear Gain condition for each of the translational gains (1, 10, 25, 50, 100). Error bars represent standard errors of the mean.

participants made the fewest errors in the nonlinear gain condition. When we repeated the analyses without the 1:1 gain level (i.e., with only 4 gain levels), we obtained similar main effects of gain, order, and condition but no interactions were significant. A similar analysis on latency as the dependent variable showed a main effect of gain, $F(4,120) = 4.1$, $p = .02$, but no effect of order or condition. The gain by order interaction was significant, $F(16,120) = 3.6$, $p = .001$. There were no other significant interactions. In general participants were slower in responding on the gain levels that they first performed, however overall most participants took longer to respond when they started with the 100:1 and 10:1 gain levels. These results did not change when we removed the 1:1 gain level from the analyses.

We report the effects of three experimental conditions (linear, nonlinear, and eyeheight) analyzed without the 1:1 data in all of the conditions. We started by testing for effects of the levels of translational gain (4), in the three different experimental conditions. All subjects were tested on different levels of translational gain, hence gain was a within-subjects factor; subjects were tested in one of the three experimental conditions, hence experiment condition was between-subjects. Separate analyses were done for each of the two depen-

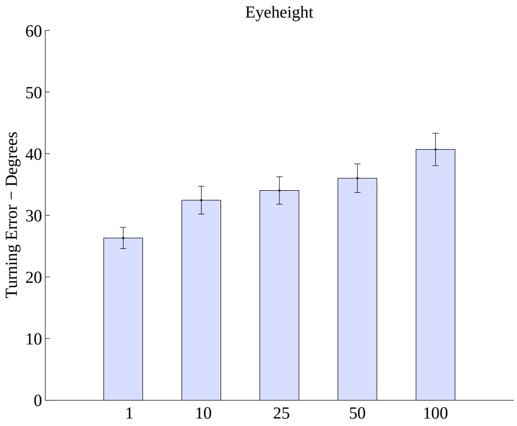

Figure V.19: This figure shows the mean turning errors in the scaled Eyeheight condition for each of the translational gains (1, 10, 25, 50, 100). Error bars represent standard errors of the mean.

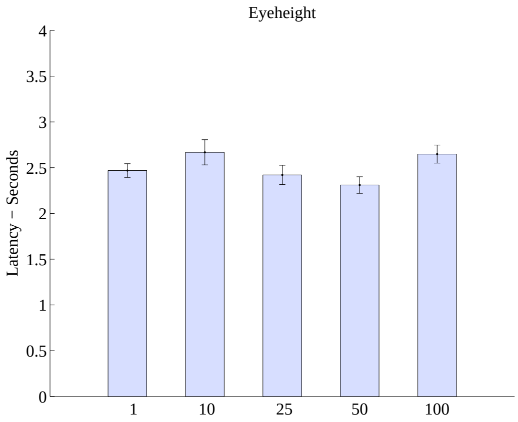

Figure V.20: This figure shows the mean latencies in the scaled Eyeheight condition for each of the translational gains (1, 10, 25, 50, 100). Error bars represent standard errors of the mean.

dent variables, turning error and latency. A multivariate repeated measures analysis on mean turning error showed main effects of gain, $F(3,126) = 11.4$, $p < .001$, experiment condition, $F(2,42) = 7.6$, $p = .002$, but no significant interaction. Participants errors were less in the 10:1 gain level, and increased as gain increased; participants errors were also less in the nonlinear gain condition than in the other two experimental groups. Planned comparisons revealed that errors in the 10:1 gain level were significantly lower than errors in the 50:1 ($t(44) = -2.4$, $p < .05$), and errors in the 10:1, 25:1 and 50:1 gain levels were all lower than errors in the 100:1 gain level (all $t > 3$, $p < .001$). A similar within subjects analyses on mean latency showed a main effect of gain, $F(3,126) = 3.9$, $p < .05$, no significant effect of the experimental condition, and no significant interaction. Similar to error, planned comparisons revealed that participants were faster to respond on the 10:1, 25:1, and 50:1 gain levels, than on the 100:1 gain level, all $t > 2$, $p < .05$.

Below we list all the possible 2 way ANOVAs (i.e., comparing 4 gain levels in any 2 of the experiment conditions). Results for turning error are presented first, followed by results of latency.

1. **Turning error for linear versus nonlinear gain experiment conditions:** A multi-

variate repeated measures analysis on mean turning error in the linear and nonlinear conditions showed main effects of gain, $F(3,84) = 13.3$, $p < .001$, experiment condition, $F(1,28) = 16.9$, $p < .001$, but no significant interaction. Participants errors were less in the lower gain levels, and increased as gain increased; participants errors were also less in the nonlinear condition than in the linearly scaled gain condition.

2. **Turning error for nonlinear gain versus eyeheight experiment conditions:** A multivariate repeated measures analysis on mean turning error in the nonlinear and eyeheight conditions showed main effects of gain, $F(3,84) = 5.8$, $p < .005$, experiment condition, $F(1,28) = 7.8$, $p < .005$, but no significant interaction. Participants errors were less in the lower gain levels, and increased as gain increased; participants errors were also less in the nonlinear gain condition than in the eyeheight condition.

3. **Turning error for eyeheight versus linear gain experiment conditions:** A multivariate repeated measures analysis on mean turning error in the eyeheight and linearly scaled gain conditions showed main effects of gain, $F(3,84) = 5.8$, $p < .005$, experiment condition was not significant and there was no significant interaction ($Fs < 1.5$, ns). Participants errors were less in the lower gain levels, but there were no significant differences in errors in the eyeheight versus linearly scaled gain experimental groups.

4. **Latency for linear versus nonlinear scaled gain experiment conditions:** A within subjects analyses on mean latency showed a main effect of gain, $F(3,84) = 4.2$, $p < .05$, a marginal effect of the experimental condition, $F(1,28) = 3.96$, $p = .056$, and no significant interaction. Just as with error, participants took longer to respond as gain increased; however there were no reliable differences between experimental conditions (i.e., participants were not faster with linear or nonlinear), although the trend was lower latencies in the nonlinear group.

5. **Latency for nonlinear gain versus eyeheight experiment conditions:** A within

subjects analyses on mean latency showed no main effects of gain levels or experiment groups, and no significant interaction. Participants were likely to be equally fast or slow on the different gain conditions, and with either experiment group.

6. **Latency for eyeheight versus linearly scaled gain experiment conditions:** A within subjects analyses on mean latency showed a marginal effect of gain, $F(3,84) = 3.0$, $p = .062$, and no effect of the experimental condition, or interaction. Participants were likely to be equally fast or slow on the different gain conditions, and with either experiment group (although the trend was faster responses in the 10:1 than other levels, and eyeheight than the linear gain group).

Results from t-tests comparing turning error and latency within each gain level against any two experiment conditions is show in Table V.4. Note that a significant t means that the mean for the first group was higher in value than the second; a minus sign for the t value indicates that the second value is larger.

V.6 Discussion

In Chapter IV, gain was limited to a translational scaling of 10 since small head movements become distracting for gains much higher. This chapter expands the findings of the previous chapter and looks at how high gain can be scaled. Increasing the user's eyeheight proportional to gain was added as an extra factor in the experimental design. We felt that scaled eyeheight could potentially aiding in spatial awareness and felt that it warranted further investigation. The results of this chapter suggest further techniques on how best to build a virtual HMD system when the size of the tracking space is small.

Three experiments were presented in this chapter. The first two experiments investigate the method of minimizing small head movements when gain is scaled higher than 10. A user study indicated tow movements that were particularly distracting in high gain, simply looking around the environment and localized movements. Thus the method of ramping up to high gain as discussed in this chapter minimizes these effects.

	$Significant?$	$t(28)$	p
TE for eyeheight vs. nonlinear gain:			
1:1	Yes	-2.3	$< .05$
10:1	Yes	2.3	$< .05$
25:1	Yes	2.6	$< .05$
50:1	No		
100:1	No		
TE for eyeheight vs. linear gain:			
1:1	Yes	-3.1	$< .005$
10:1	No		
25:1	No		
50:1	No		
100:1	Marginally	-1.997	$= .056$
TE for nonlinear vs. linear gain:			
1:1	No		
10:1	Yes	-2.9	$< .01$
25:1	Yes	-2.8	$< .01$
50:1	Yes	-3.4	$< .005$
100:1	Yes	-3.3	$< .005$
LT for eyeheight vs. nonlinear gain:			
1:1	No		
10:1	No		
25:1	No		
50:1	No		
100:1	No		
LT for eyeheight vs. linear gain:			
1:1	No		
10:1	No		
25:1	No		
50:1	No		
100:1	No		
LT for nonlinear vs. linear gain:			
1:1	No		
10:1	No		
25:1	No		
50:1	No		
100:1	Marginally	-1.8	$= .075$

Table V.4: Results from t-tests comparing turning error (TE) and latency (LT) within each gain level against any two experiment conditions. A significant t means that the mean for the first group was higher in value than the second. A minus sign for the t value means that the second condition value was more than the first.

Experiment 5 reported that subjects preferred the 0.5 m/s critical threshold because they were able to control local and global movements. This critical speed threshold was found using a cubic function to move into a linearly scaled translational gain. In Experiment 6, the critical threshold value was fixed at 0.5 m/s, and we found that subjects preferred the exponential ramping function. Although the critical value was not found using an exponential function, we feel, given the experimental results, that 0.5 m/s represents a reasonable value.

The results of Experiment 5 suggest that using this ramping function was an effective method of minimizing small head movements. We test this more closely in Experiment 3 using four different gain values (10:1, 25:1, 50:1, 100:1). Experiment 7 further revealed that using the ramping function results in better spatial orientation than simply scaling gain linearly. Turning errors in this condition were significantly better than the linearly scaled gain. There was also a marginal effect of nonlinearly scaling gain on latency. This marginal effect of faster responses in the nonlinear gain condition could suggest that people were more spatially oriented, but definitely shows that people were not making speed accuracy trade-offs. Experiment 7 also shows that scaling eyeheight proportionally to gain did not aid in spatial awareness as compared to linearly scaling gain.

We report strange effects in the 1:1 gain across each of the conditions. Errors are quite large. By design the 1:1 condition was identical in all three conditions. However, the responses were different. We believe that the reason for this phenomenon is a as follows. Objects in the 1:1 condition were generally found below eye level, and required users to look down about 35° to view them in the virtual environment, in contrast to objects in nonlinear and linear conditions. Thus, this experiment unwittingly confounded viewing angle in an HMD with translational gain. The mechanics of the HMD may make it more difficult to view objects low to the ground, a limitation of HMDs not heretofore reported. More investigation of this phenomenon is needed. More generally, does viewing angle have an effect on spatial orientation. In the nonlinear and linear conditions, 1:1 was not significantly different. That is, 1:1 turning errors and latencies are effectively the same in these two conditions. Eyeheight at 1:1 is different. We suggest that people were better at it because in terms of viewing angle because they were looking down at all times. It was not “different” among the gains. Moreover, subjects in the eyeheight condition seemed to prefer the 1:1 condition most. More work is needed to resolve this issue.

Interrante et al. [2007] propose a method called “seven league boots” in which they scale gain based on wand control. This work suggests that it may be helpful to allow users some sort of control over optical flow rate. We allow user more control over scaling using this nonlinear scaling technique.

This chapter shows that scaling gain nonlinearly is an effective method of explore a large virtual environments for gains up to 50. According to results of Experiment 7, turning errors and latencies get significantly worse at 100:1, making 100:1 an unreasonable chose for allowing users to explore a virtual environment and expecting them to maintain spatial orientation. At 50:1, turning errors and latencies are statistically the same as the 10:1, 25:1. Interestingly, we have better results at the 50:1 gain than at 1:1 gain.

CHAPTER VI

RESETTING

VI.1 Introduction

This chapter looks at “resetting” subjects when they reach the limits of the tracking system. Resetting involves manipulating optical flow in a way that allows users to move away from a physical obstruction such as a wall while experiencing a continuous sense of their location in virtual space.

As previously discussed, HMD-based virtual environments are often explored on foot. This type of exploration is useful since the inertial cues of physical locomotion aid in spatial awareness. The size of the virtual environment that can be explored is limited to the dimensions of the tracking space of the HMD unless some other method of exploration is used. Chapters IV and V present work that manipulates the translational gain of walking, so that one step forward in the physical environment corresponds to several steps forward in the virtual environment. Two experiments show that increasing the translational gain of walking is a useful method of navigating large virtual spaces, and that it is superior to joystick exploration. However, the physical limits of the tracking system may be reached no matter how high gain is scaled. Thus, this chapter presents work on an additional strategy that resets subjects when they walk and reach the end of the physical space. This strategy assesses the ability of people to rely on visual information for spatial updating during these resets. This chapter develops and evaluates three methods of resetting position while subjects walk in small physical tracking spaces in order to explore large virtual spaces.

After completing a reset, users travel along the same virtual path they had been traveling. In the Freeze-Backup method, the user obtains more space for virtual exploration by taking steps backwards while frozen in a fixed position in the virtual environment. In the other two methods, Freeze-Turn and 2:1-Turn, users overcome physical obstruction by physically turning 180° and maintaining their same position in virtual space before and after the turn. During a Freeze-Turn reset, the orientation of the user is frozen while the subject turns 180°. In the 2:1-Turn condition, the gain of the turn is doubled so that 180° turn in physical space corresponds to 360° turn in virtual space.

This work is important because one future goal discussed in Chapter VII is to extend and integrate the results of this chapter with the work on scaling the translational gain of walking (Chapters IV and V). The resulting system should allow a person to seamlessly explore large virtual environments. The system envisioned here could be based in an office or small lab. In particular, if immersive virtual environments are to realize their potential as commodity-level components, a perceptually accurate interface that allows locomotion through them within the constraints of everyday space must be developed.

VI.2 Method

Three resetting methods are evaluated. Resetting involves physical locomotion with optical flow manipulated in such a way that the user's sense of where they are relative to objects in their virtual environment is not changed. The three resetting methods are called Freeze-Backup, Freeze-Turn, and 2:1- Turn are explained as follows.

1. **Freeze - Backup.** In this method, the computer indicates to the user that they have reached the boundaries of the tracking system and needs to reset. The tracking system is no longer used to update the position of the subject in the virtual environment , so that the user's position in virtual space is no longer updated with movement in physical space. The user is then instructed to take steps backwards in physical space while user's position in virtual space remains fixed or frozen. When enough steps are taken, the computer indicates for the user to stop, the displays are unfrozen, and the user is allowed to continue along the same path that they were walking before the reset. During the backward walking, orientation tracking is active so that the user can look around.

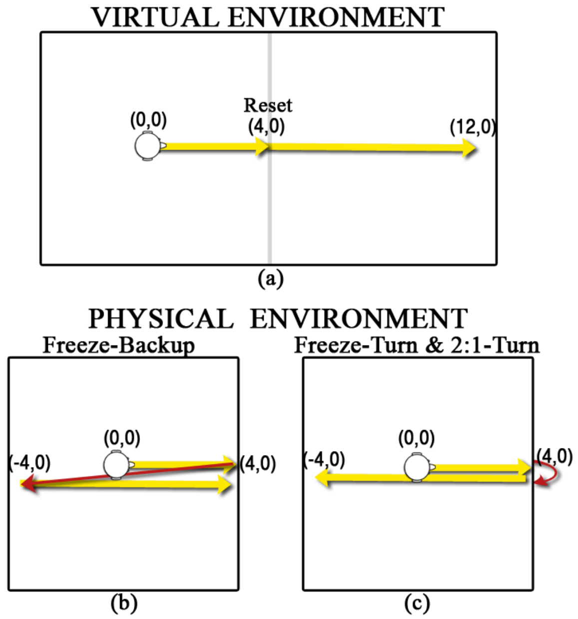

Figure VI.1: In (a), the path a subject perceives they have taken in the virtual environment is shown, while the path in physical space that the subject takes under the different resetting methods is shown in (b) and (c). In this example, a person position (0,0) in physical space views the virtual environment at position (0,0). In (a), the person walks forward in the virtual environment where they are alerted by a signal at (4,0) indicating they are near the tracking limits and need to reset their position in physical space. The person then continues walking to (12,0) in the virtual environment. The corresponding paths in the physical environment for the three resetting methods are shown in (a) and (b). Red arrows indicate physical movement during a reset.

The algorithm first initializes the reset offset *resetOffset* so that

$$resetOffset_x = 0 \tag{VI.1}$$

$$resetOffset_z = 0 \tag{VI.2}$$

After a reset, the position of the user in the virtual space must be calculated by offsetting the physical position of the user by some amount. Therefore, the virtual position (*vePos*) and orientation (*veOri*) at any point in time while the user is not undergoing a reset can be calculated as

$$vePos_x = currPos_x + resetOffset_x \tag{VI.3}$$

$$vePos_y = currPos_y \tag{VI.4}$$

$$vePos_z = currPos_z + resetOffset_z \tag{VI.5}$$

$$veOri(x,y,z) = currOri \tag{VI.6}$$

where vectors *currPos* and *currOri* indicate the current position and orientation of the user. Before the first reset, the user's position in x, y, z space and orientation (pitch, yaw, and roll) is equal to that of the physical space. Once the user has reached a boundary, a message automatically appears requesting that the user stop walking. Once the user stops moving, their position in the virtual position is fixed and their current location in the virtual environment must be recorded and the reset offset updated accordingly.

$$resetOffset_x = currPos_x + resetOffset_x \tag{VI.7}$$

$$resetOffset_z = currPos_z + resetOffset_z \tag{VI.8}$$

During the reset, the user takes steps backwards, yet the virtual position is not updated. However, virtual orientation is updated, enabling the user to look around from a fixed position while backing up. Therefore while the user is undergoing a reset, the user's position in the virtual world must be calculated:

$$vePos_x = resetOffset_x \tag{VI.9}$$

$$vePos_y = currPos_y \tag{VI.10}$$

$$vePos_z = resetOffset_z \tag{VI.11}$$

$$veOri(x,y,z) = currOri \tag{VI.12}$$

The user stops backing up when the system indicates that he or she has backed up enough. Then to complete the reset and enable to continue along the path he or she was traveling prior to the reset the following calculation is made:

$$resetOffset_x = resetOffset_x - currPos_x \tag{VI.13}$$

$$resetOffset_z = resetOffset_z - currPos_z \tag{VI.14}$$

Figures VI.1(a) and VI.1(b) show an example of the process. The rectangle shown in Figure VI.1b represents the physical limits of the tracking system while the larger rectangle shown in Figure VI.1a represents the virtual environment. In this example, the user starts at position (0,0) in both the real and virtual environments. The user then proceeds to (4,0) but cannot explore further because the limits of the physical space have been met. Therefore, the user undergoes a reset, and their position is frozen at (4,0) as they follow the red path and back up to (-4,0). During the backup phase, the user is instructed to simply walk backward and told when to stop, and is not guided backward. Thus, the user does not typically walk a straight path directly behind them as in this example. Once the user reaches (-4,0), the system instructs the user to stop, and the user continues along the the yellow path until they reach (4,0). The corresponding path in virtual environment from (0,0) to (12,0) is show on the right.

The physical position of a user in x, y, z space using a right-handed coordinate system is obtained from the tracking system. The position in the center of the room on the floor is (0, 0, 0). The x, y, and z directions while standing in the center of the room facing to the front correspond to front–to–back movement, user height, and right–to–left movement, respectively. Movement is limited in the x and z directions due to the finite range of the tracking system. Since the y-direction indicates movement perpendicular to the ground pane, this value typically represents the user's eyeheight, and does not limit the exploration of the virtual environment. Orientation is obtained from the rotational sensor located on HMD which updates rotation about the *x*-axis (pitch), *y*-axis (yaw), and *z*-axis (roll).

2. **Freeze - Turn.** In this method, when the tracking device finds that the subject has reached a boundary, the computer indicates to the participant that they need to reset by turning around. The display of the HMD is frozen, freezing the participant's position and yaw angle in virtual space, and the participant turns 180 degrees. The display is unfrozen, tracking is updated, and the subject is able to continue traveling along his route.

   In this resetting condition, the user turns around with the virtual screen frozen until he or she feels that they have turn approximately 180°, and then the screen is unfrozen and the user continues along their path. Thus, head movement about the y-axis must be manipulated. This manipulation is controlled by θ_y, and is initialized with the reset offset in the x ($resetOffset_x$) and z ($resetOffset_z$) directions. The two variables $rotAxis_x$ and $rotAxis_z$ specify the origin of the transformation. Thus, the variables in this resetting conditions are initialized by equations VI.1, VI.2, and the following:

$$\theta_y = 0 \tag{VI.15}$$

$$rotAxis_x = 0 \tag{VI.16}$$

$$rotAxis_z = 0 \tag{VI.17}$$

   To calculate the users' position in the virtual environment while they are not resetting, the current physical location of the user in the x and z directions must be translated by $resetOffset_x$ and $resetOffset_z$ and rotated about the y-axis. Thus, the rotation matrix is defined as

$$R = \begin{bmatrix} cos(\theta_y) & sin(\theta_y) \\ -sin(\theta_y) & cos(\theta_y) \end{bmatrix} \tag{VI.18}$$

Current virtual position and orientation is calculated as follows:

$$\begin{bmatrix} vePos_x \\ vePos_z \end{bmatrix} = \begin{bmatrix} currPos_x - rotAxis_x \\ currPos_z - rotAxis_z \end{bmatrix} R + \begin{bmatrix} resetOffset_x \\ resetOffset_z \end{bmatrix} \tag{VI.19}$$

$$vePos_y = currPos_y \tag{VI.20}$$

$$veOri_x = currOri_x \tag{VI.21}$$

$$veOri_y = currOri_y + \theta_y \tag{VI.22}$$

$$veOri_z = currOri_z \tag{VI.23}$$

When the tracker senses the user out of bounds, the computer alerts the user by message on the HMD display instructing them to stop locomoting. To reset, the user turns around while frozen in their current position. Therefore, to start the reset the following calculations are made:

$$startAngle_y = currOri_y \tag{VI.24}$$

$$\begin{bmatrix} resetOffset_x \\ resetOffset_z \end{bmatrix} = \begin{bmatrix} currPos_x - rotAxis_x \\ currPos_z - rotAxis_z \end{bmatrix} R + \begin{bmatrix} resetOffset_x \\ resetOffset_z \end{bmatrix} \tag{VI.25}$$

$$rotAxis_x = currPos_x \tag{VI.26}$$

$$rotAxis_z = currPos_z \tag{VI.27}$$

The variable $startAngle_y$ stores the y-direction orientation of the user at reset. Thus, during a reset, virtual position is calculated using equations VI.9, VI.10, and VI.11, and orientation is calculated as:

$$veOri_x = currOri_x \tag{VI.28}$$

$$veOri_y = startAngle_y + \theta_y \tag{VI.29}$$

$$veOri_z = currOri_z \tag{VI.30}$$

Position is frozen in the x and z directions and orientation is frozen about the y axis. To end the reset θ_y must be updated:

$$\theta_y = \theta_y - (currOri_y - startAngle_y) \tag{VI.31}$$

Figures VI.1(a) and VI.1(c) show an example of the area walked in the physical space and the corresponding area walked in the virtual space. In this example, the user starts in position (0,0) in both the physical and virtual environments viewing the virtual world through the HMD. The user walks to position (4,0) where there is no more physical space and desires to continue along this same path. Thus, the user turns around with a display frozen in that y-direction at position (4,0) to reset. When the screen is unfrozen, the user has turned 180° and continues walking to (-4,0) in the physical space. The yellow path in Figure VI.1a shows the locomotion that the user perceived in the virtual environment.

3. **2:1 - Turn.** In this method, when the subject reaches the boundaries of the tracker, the computer indicates that they should turn and keep turning until completing a visually full turn in the virtual environment. The rotation gain of the yaw angle during this turn is scaled by two, such that the user rotates 180° in the physical environment, but rotates 360° in the virtual environment.

   The algorithm for this resetting condition is exactly the same as the Freeze-Turn condition, with the exception of equation VI.29 which calculates virtual orientation around the *y*-axis during the resetting phase:

$$veOri_y = (currOri_y - startAngle_y) * 2 + \theta_y \tag{VI.32}$$

Figure VI.1(a) shows an example of a path taken in a virtual environment where one reset is undergone. The corresponding physical path is seen in Figure VI.1(b). Note that the path taken by the Freeze-Turn and 2:1 conditions are similar since they both involve the user turning around to reset.

VI.3 Experimental Evaluation

Since all three of these methods are a priori reasonable forms of resetting, an experiment was conducted to evaluate which one worked best. Additionally, we wanted to determine what the cognitive cost of a reset was in each method and assess if users become increasingly disoriented during longer uses of the system. A priori, we can make several observations about the performance of the various methods. First, the backup method requires walking backward in an HMD, an action that is less stable than walking forwards. The 2:1-Turn condition switches users between a “normal” (1:1) rotational gain and a 2:1 rotational gain, which may prove disorienting. The Freeze-Turn system disassociates proprioceptive cues from optical flow, which may also be disorienting.

VI.3.1 Materials

Twelve naive subjects participated in this study. The virtual world was viewed through a full color stereo NVIS nVisor SX Head Mounted Display with 1280 x 1024 resolution per eye, and a field of view of 60° diagonally. The size of the physical room in which the experiments were performed was approximately 5m by 6m, and within the room the limits of the four camera video position tracking system was approximately 5m by 5m. The virtual room was 50m by 50m shown in Figure VI.2, ten times the size of the physical limits of the tracking system. Objects were placed in the room in different orientations

Figure VI.2: This figure shows the virtual environment used in the experiment.

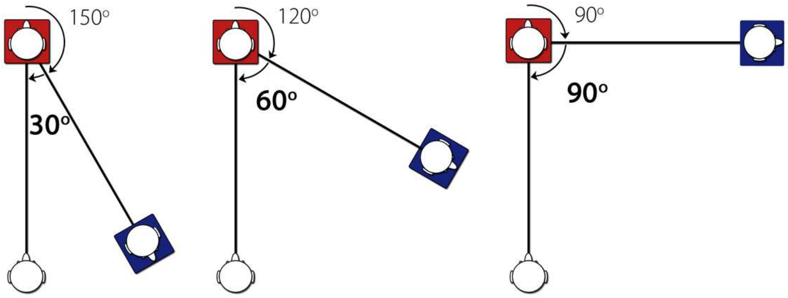

Figure VI.3: This figure shows a top down view of the three different angles (30°, 60°, and 90°) of the two segment path followed by each participant. The subject walks to a red chair then turns 150°, 120°, or 90°, respectively, and walks to the blue chair.

to give the subject a sense of the size and scale of the environment. The environment contained 7 different color tables scattered throughout the environment, 14 posters on the wall, a refrigerator, a fish tank, three sofa areas, two bookshelves, a group of six chairs, a computer desk, a computer, doors, a group of slot machines, and a pool table.

VI.3.2 Procedure

The goal of this experiment was to assess how well subjects maintain spatial awareness of an environment after undergoing resets. We tested subject's spatial knowledge in each of the three resetting conditions. Since there were six orders of three reset conditions, two

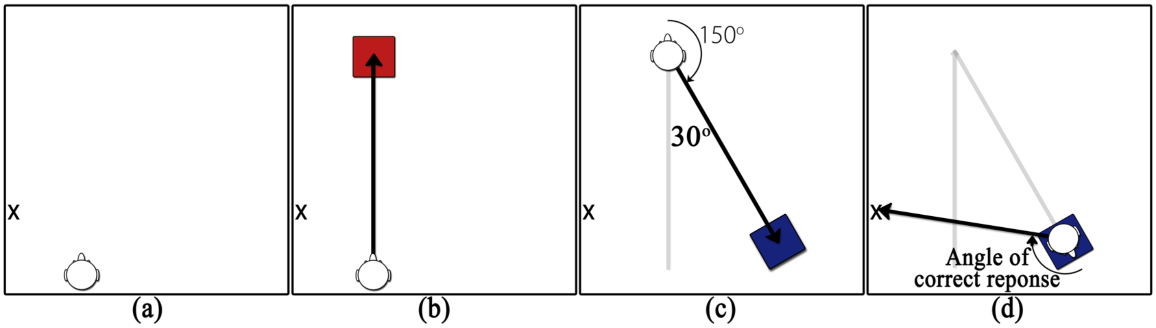

Figure VI.4: A trial consists of walking a two segment path and turning to face a remembered target object. In this example, the subject is asked to remember the location of an object denoted ‘x’ as shown in (a). Once the subject indicates to the experimenter that they have memorized the position, the red chair appears and the subject is instructed to walk to it as indicated by (b). Once the subject has reached the red chair, it disappears, and the blue chair appears (c). In this example the experimenter instructs the participant to find the blue chair on their right, requiring them to turn 150°. Once the subject reaches the blue chair (d), they are asked to turn to face object ‘x’ with eyes closed. The correct angle of response is shown in (d).

subjects completed each of the six different orders in a counter-balanced fashion. During each testing condition, the participant completed a total of eighteen trials. A trial consisted of walking a path and then turning to face a remembered target object. Before each trial participants were placed in a starting position, and then asked to remember the location of one object or a set of three objects. Trials involving three objects were included so that subjects needed to keep in mind all three objects during the walk to the test position. In this condition, three objects were named at the start of the trial and subjects were told that they would be asked to turn and face any one of the three after they walked to the test position. Participants were given about sixty seconds to remember the objects locations and freely rotate around from the starting position to view them before traveling the path. Objects were selected so that they did not appear along the participants’ path. The correct angle of response from the facing position at the end of the path to the object that the subject was asked to face varied from 30° to 180°.

The travel path consisted of a two segment route, where subjects walked to a red chair

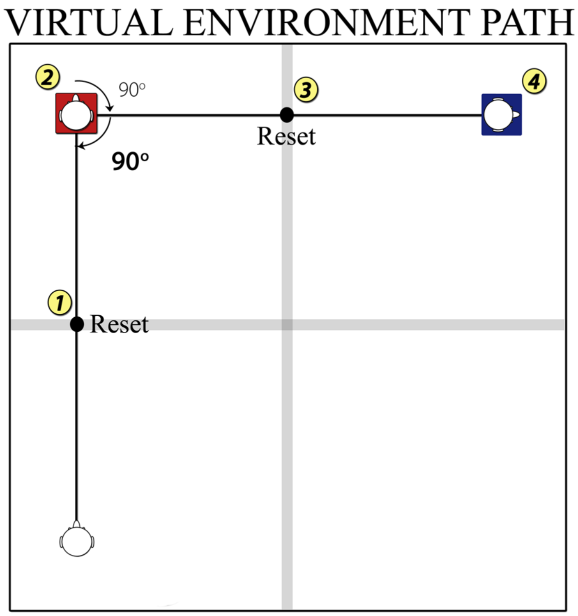

Figure VI.5: This figure shows a two segment path traveled by a participant in the virtual environment during a two reset trial. In this particular example, the angle of the path is 90°. The subject starts the trial positioned in the left corner. The subject walks to the red chair, and is reset once at position 1 along their path to the red chair. Once the subject reaches the red chair at position 2, the subject turns 90° to the right to find the blue chair and walks towards it. Along the way to the blue chair the subject is reset at position 3 and then continues to reach the blue chair at position 4. The resets do not change the position and orientation of the user in the virtual environment.

and then to a blue chair. The red and blue chairs were meant to serve only as signs, showing the way they should walk to reach the test position. The angle between the starting point, red chair, and blue chair was either 30°, 60°, or 90° as shown in Figure VI.3. Figure VI.4 shows an example of a trial. After the subject memorized the location of the object or objects (Figure VI.4(a)), a red chair appears and the participant is instructed to walk to the red chair (VI.4(b)). Once the subject had arrived at the red chair, the red chair disappeared and a blue chair appeared(VI.4(c)). The experimenter instructed the subject on which direction to turn (right or left) to find the blue chair. The subject was instructed not to look around at the target object or objects while walking the two segment path. At the end of the path, the experimenter instructed the subject to close his eyes and turn to face a remembered target object (VI.4(d)). Time was recorded using a stopwatch and

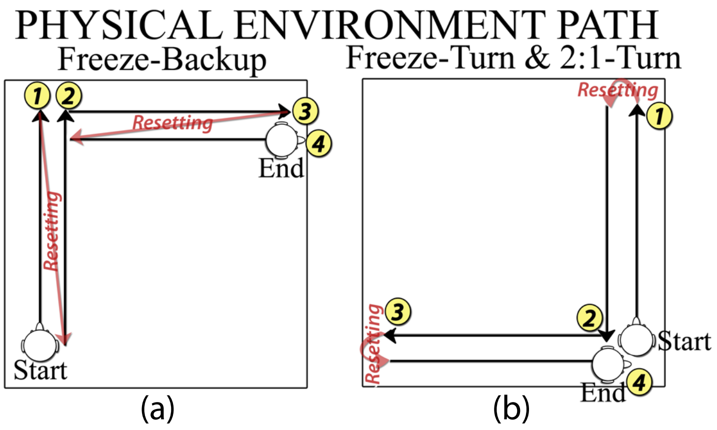

Figure VI.6: This figure shows the physical path taken by the subjects when traveling the virtual path of Figure VI.5. In the Freeze-Backup shown in (a) , participants walk toward the red chair and are reset at position 1. To complete a reset, they take steps backwards as shown by the red arrow. During the reset, their position is frozen in virtual space. After they have taken enough steps backwards, the screen unfreezes and they continue along their path to the red chair at position 2. Next the participant walks to the blue chair, is reset at 3 then continues along until reaching position 4. The physical path followed during this trial in 2:1-Turn condition and the Freeze-Turn condition is the same as shown in (b). The subject is reset at position 1 and turns 180° to continue to the red chair positioned at 2. The subject turns 90° to the right and continues toward the blue chair and is reset at position 3.

the rotational position was recorded by the computer. The average distance from the final location subject at the blue chair to the target object was approximately 20m and ranged from 3m to 40m. The starting position was varied randomly within 10m of the center of the room and the orientation varied randomly by 90°.

While locomoting along the path, the subject was reset zero times, two times, or four times depending upon the length of the path. In the zero reset condition, the subject completed two 4m segments. Note that the zero reset condition is the same under all three resetting methods; it was included in the experimental design to provide a baseline across trials. In the two and four reset conditions the subject traveled two 8m paths and two 12m paths, respectively. The position of the reset on the path was engineered so that they were spaced an equal distance apart. For example resets in the 8m path of the two reset condition occurred at 4m. Likewise, in the 12m segment length of the four reset condition, resets occurred at the 4m and 8m. Figures VI.5 and VI.6 show an example of a trial where the subject is reset twice. Figure VI.5 represents the path traveled in the virtual environment and Figure VI.6 shows the paths traveled under the three different resetting condition. Since there were two different numbers of objects to remember, three path angles, and three different numbers of resets, there were eighteen trials per condition representing each possible combination. Each condition took approximately 45 minutes to complete and thus were completed on consecutive days.

Participants completed zero reset trials during each of the resetting conditions indicating ideal behavior or baseline performance for the subject. Zero resets for each condition are equivalent. Since the subject did not undergo resets during the the zero reset condition, we did not use this condition to find an effect of method as discussed in the next section.

The path angle was varied by 30°, 60°, and 90° so that subjects did not use strategies based upon the the path angle. The participant had to memorize the location of one object or three objects. We used three objects as well as one object to see how difficult the task becomes when the cognitive burden is higher. In the case of memorizing one object participants could use strategies, but in the case of three objects subjects generally are forced to spatially update along their path.

VI.4 Results

The results of the experiment were analyzed in terms of the errors and latencies in turning to face the targets. Turning error is defined as the difference in angle when turning to face a given target relative to one's actual position in the virtual room. The angle of correct response to the target object from the end of the path varied as repeated within-subject trials from 30° to 180°. Latency was measured from the time the subject was given the object to face until the subject came to rest at a final position. The independent variables in this experiment were reset condition (Freeze-Backup, Freeze-Turn, and 2:1-Turn), number of resets (2 or 4), number of target objects (1 or 3) and angle of turn (30°, 60° or 90°). All independent variables were within-subjects. As noted above, the zero reset condition was identical across all reset conditions, to provide a baseline under an ideal condition. Thus it was not included in the statistical analysis conducted below.

Graphs of the mean turning errors and mean latencies collapsed across various factors are shown in Figure VI.7 through VI.12. Figures VI.7 and VI.8 show the mean turning error and latency as a function of reset condition with the zero reset condition included as a baseline, respectively. Figures VI.9 and VI.10 break this information out by number of resets, and Figures VI.11 and VI.12 break this information out further by number of objects.

A multivariate repeated measures analysis of variance on the turning error found a main effect of reset condition $F(2,22) = 5.4$, $p < .05$. Participants made fewer errors with Freeze-Backup than with other reset conditions. There were no other main effects or interactions. A repeated measures analysis on latency show a main effect of number of objects to remember, $F(1,11) = 29.9$, $p < .01$. Participants were faster when they had to remember fewer objects. There was a significant interaction of reset condition × number of objects, $F(2,22) = 9.8$, $p < .05$. Subjects were fastest when they had to remember fewer objects but were generally faster in the 2:1-Turn condition (see Figure VI.12).

VI.5 Discussion

This chapter examines methods for exploring large HMD-based virtual environments when the physical space housing the HMD is limited. It studies three methods for resetting a user's location in physical space while hoping they could maintain their spatial orientation in the virtual space. In the Freeze-Backup method, the user obtains more space for virtual exploration by taking steps backwards while frozen in a fixed position in the virtual environment. In the other two methods, Freeze-Turn and 2:1-Turn, users overcome physical obstruction by physically turning 180° and maintaining their same position in virtual space before and after the turn. During a Freeze-Turn reset, the orientation of the user is frozen

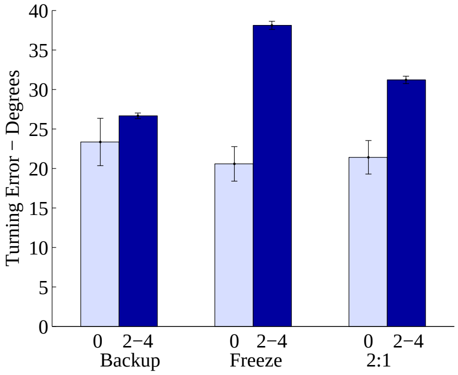

Figure VI.7: Mean turning error as a function of condition and number of resets. Under each resetting condition the mean of the zero resets is compared to the mean of the two and four resets combined.

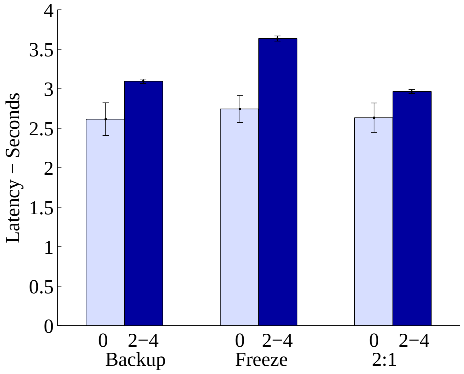

Figure VI.8: Latency as a function of condition and number of resets. Under each resetting condition the mean of the zero resets is compared to the mean of the two and four resets combined.

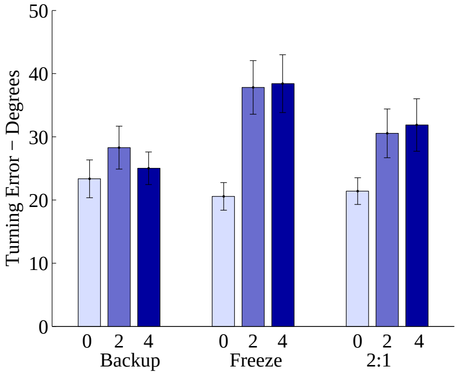

Figure VI.9: Mean turning error as a function of condition and number of resets. Under each resetting condition the mean is categorized by number of resets: zero, two, or four.

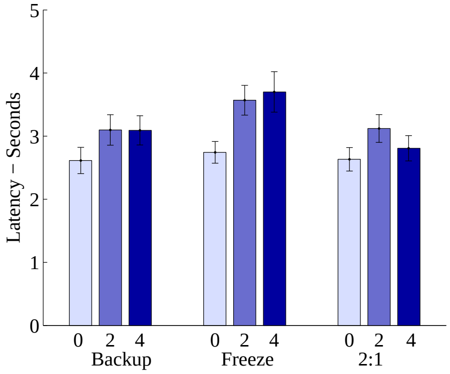

Figure VI.10: Latency as a function of condition and number of resets. Under each resetting condition the mean is categorized by number of resets: zero, two, or four.

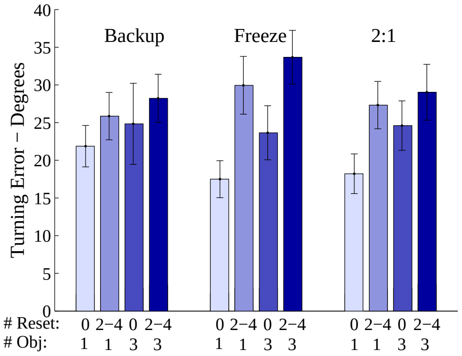

Figure VI.11: Mean turning error as a function of condition, number of resets, and number of objects memorized. Under each resetting condition the mean is grouped into six different categories representing each of the possible combinations of number of resets and number of objects.

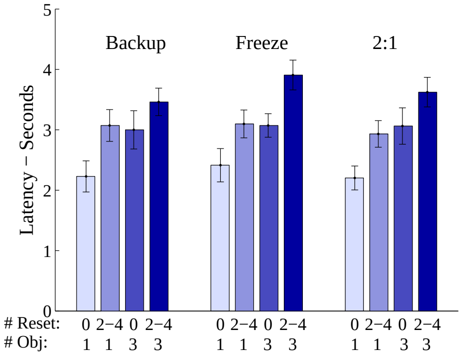

Figure VI.12: Latency as a function of condition, number of resets, and number of objects memorized. Under each resetting condition the mean is grouped into six different categories representing each of the possible combinations of number of resets and number of objects.

while the subject turns 180°. In the 2:1-Turn condition, the gain of the turn is doubled so that 180° turn in physical space corresponds to 360° turn in virtual space.

These results indicate that the lowest errors occur in the Freeze-backup condition, while latencies were lowest for the 2:1 condition. There are several interesting observations about these results. First, updating one's position in the Freeze-Backup condition involves ignoring proprioceptive cues that would result in the change of perspective being a geometric translation. According to the interference literature ([Rieser 1989; May 1996a; Waller et al. 2002]) and the functional similarity work discussed in Chapter III, there is a conflict between physical locomotion and imagined locomotion. In particular, imagined perspective changes are more difficult physical perspective changes. That is, spatial orientation is updated much faster and more accurately when physical locomotion is involved. This literature also shows that imagined perspective switches are much easier that imagined rotational perspective changes. We find that errors and the latencies were the worst for the Freeze-Turn condition. The other two conditions were both better than the Freeze-Turn condition. In the Freeze-Turn condition, users physically turn and are required not to update their orientation in virtual space. In other words, they make a physical perspective change and not update. The interference literature suggests that this change causes conflict, especially since it is a change of rotational perspective. In contrast, Freeze-Backup is a translational perspective change, which could be why people are better at the Freeze-Backup. Moreover, if the total disparity (as discussed in Chapter II) is calculated during a reset, as compared to Freeze-Turn and Freeze-Backup, we see that there is less disparity or potentially less conflict when users undergo a Freeze-Backup reset. The fact that 2:1-Turn was almost as good as the Freeze-Backup suggests that we update based on physical locomotion.

Kuhl [2004] shows that people can recalibrate rotations in a virtual environments. This finding may provide insight as to why people prefer the 2:1-Turn over the Freeze-Turn. However, we do not want people to use the 2:1-Turn resetting to recalibrate all subsequent turns in the virtual environment. Whether there is some recalibration when subjects turn-to-face target objects in the 2:1-Turn condition is not known. To examine this issue, we would need to record the direction of the reset turn and the direction the subject turned to face the object. These measures were not recorded in the present experiment. The fact that subjects may over-turn in the direction of the resetting in the 2:1-Turn condition is an interesting hypothesis that we plan to investigate.

We generally feel that the physical room was not important and people did not keep a dual representation of the real and virtual environment. Informal observations reveal that users do not know where they are in the physical room after completing the experiment. However, we did not do any tests involving spatial perception in the physical environment. Moreover, if people were maintaining a dual representation of both environments and were localizing objects in the virtual environment using the real physical environment, we would expect to see turning errors in excess of the errors we did observe. Specifically, in the 2:1-Turn and Freeze-Turn conditions, if the subject was using the physical environment to reference their spatial knowledge of virtual objects, then we would expect either large turning errors (on the order of 180°, rather than the 15° from the baseline of zero resets) or large latencies as they try to compute the locations of objects. Moreover, representing two environments would put a huge burden on working memory. Yet, in all three conditions when working memory was loaded with three objects instead of one, relatively low turning errors and latencies are still reported. Since there was not a huge impact on turning errors or latencies for the number of objects, we believe that people are not representing both environments. There has been some research on dual representations, [Brockmole and Wang 2003; Waller et al. 2002; Wang 2005b]. These studies investigate whether people update nested environments. An example of a nested environment would involve learning the locations of objects from a laboratory and bringing into memory the locations of objects from a familiar place such as the kitchen of a home. The authors conclude that when performing a physical rotation, people updated both the immediate and imagined environment when rotating to face an object in imagined environment. However, people did not update their position in the imagined environment when turning to face an object in the immediate environment. The fact that people did not update the non-immediate environment when immersed in the immediate environment could suggest that in our experiment people did not update their position in the non-immediate environment, the physical environment.

Even though the Freeze-Backup condition has the lowest turning errors, a satisfying design for a commodity-level HMD system would likely consist of either the Freeze-Turn or 2:1-Turn methods. The disadvantages of the Freeze-Backup method are the potential danger of backing into a wall or tripping over the tether, and the longer length of time and walking involved in resetting. However, since the Freeze-Backup condition is the best resetting condition, we envision its use in training applications where spatial orientation is important, where time is not an issue, and trainees are able to have a guide to make sure that they do not back into a wall or get stuck in a corner. For example, this method could be used to test emergency exits for proposed architectural designs in case of a fire, and other emergency type applications. We generally prefer the 2:1-Turn method since resets are relatively fast, and gain could be further exploited so that the amount a user turns during a reset could easily be manipulated to provide maximum physical space for forward exploration.

The results of these experiments also show that there is a cost to resetting in terms of a user's orientation to remembered objects. To mollify this out, a complete system would likely involve a method or provision for the users to reorient themselves periodically in the virtual environment. Given the current state of virtual environments, the need for reorientation is not a severe drawback, although it is one we would like to eliminate. There is ample evidence that people have difficulty maintaining orientation in virtual environments [Ruddle 2001; Allen and Singer 1997; Péruch et al. 2000]. Typically, these difficulties are attributed to poor idiothetic cues, such as the absence of proprioception and other sources of information provided by self locomotion (in the case of desktop virtual environments) and the limited field of view of HMDs.

CHAPTER VII

CONCLUSION AND FUTURE DIRECTIONS

VII.1 Conclusion

This thesis develops novel methods for humans to experience and learn in virtual environments. The potential uses of such environments are limitless, and some include testing evacuation plans before a structure is built, training construction workers to operate expensive and dangerous machinery, assessing search and rescue efforts of firefighters, and allowing medical professionals to practice complicated procedures.

Creating virtual environments has become much easier over the last few years. 3D models of smaller objects can be quickly acquired using laser scanners. Crude methods, such as Frueh and Zakhor [2003], exist to quickly generate 3D models from large real world scenes. With hi-resolution satellite images being more rapidly available, these type of virtual models will be built faster and less expensively. Snavely et al. [2006] generate a 3D model from a large collection of photographs that are either from a personal collection or from internet photo sharing sites. Given the potential ease with which building virtual environments, it seems that there will be a demand for virtual systems to explore these environments.

Chapter II discussed a number of different ways to explore virtual environments. As these virtual models become easier to develop and more accessible as a general learning tool, systems to explore these environments need to be readily available. We focused our research on head-mounded display (HMD) since it seems to be the most promising technology that can be used to explore these types of environments in the near future. HMD technology allows for the exploration of a virtual environment using your own locomotion or physically walking. Physical locomotion is a highly effective method of learning the locations of things when exploring virtual environments, and seems to result in better spatial orientation than other locomotor interfaces such as joysticks. Exploring a virtual environment using physical locomotion can be easily accomplished with this system as long as the size of the virtual environment is no larger than the HMD tracking system. However, exploring virtual environments larger than the physical constraints of the tracking system becomes an issue.

First, in Chapter III, we examined whether or not humans can explore virtual environments and develop a sense of space similar to the real world. That is, do people reason about space and spatial relationships between themselves and other objects in a virtual space similar to a physical space? This study mimicked Rieser [1989], which looks specifically at how people reason about space when undergoing perspective changes, i.e., rotational perspective changes and translational perspective changes. Thus, we repeated Rieser's [1989] experiments with an additional conditions in the virtual environment. We find that users perform the spatial orientation tasks or the experiments in a functionally similar manner in the real world and in the virtual world. That is, errors and latencies in Experiment 1 and 2 of Chapter III tell us that users reason about space when undergoing a perspective change in both the real world and the virtual world in a similar way.

Since Chapter III finds that we could expect users to explore a virtual environment and develop a useful sense of space and their orientation within that space, the rest of this thesis dealt with finding ways to solve the problem of exploring a large virtual environment in a small physical space. We proposed several techniques and evaluated the users spatial orientation.

In Chapter IV, we manipulated the translational gain of walking so that one step in physical space carries one several steps forward in virtual space. This chapter limited the scaling of translational gain to a factor of 10 since small head movements becoming distracting and disorienting for gains much larger than ten. In Experiment 3 of this chapter, we directly compared scaling translational gain of walking to joystick locomotion and found that scaling gain results in significantly better spatial orientation ([Williams et al. 2006a]).

This result was interesting because even though the proprioceptive cues of walking did not match the visual experience in the 2:1 and 10:1 condition, just having these cues had a positive impact on spatial orientation. In Experiment 4 we compared spatial orientation when translation gain was not scaled, scaled by 10, and scaled by 25. We found that the accuracy and speed in locating remembered target objects was essentially the same across the three different gain conditions. That is, people were just as spatially aware of their environment no matter what gain they experienced. The results of this chapter tell us that we can scale gain up to ten and expect people to be just as spatially oriented as they would in a 1:1 environment. Thus, we have increased the amount of space that the user can explore using their own locomotion to a size that is ten times the size of the tracking space of the HMD.

Since 10:1 translational gain leads to reasonable spatial orientation, we wanted to investigate this issue further and find how high we can expect to scale gain without greatly degradation spatial awareness. To scale gain higher than ten, an additional method must be employed. The results of Experiments 5 and 6 of Chapter V identified a useful method of minimizing distracting head movements. This method "ramps" users up to high-gain so that localized movements are not scaled as much as global movements. Experiment 7 of Chapter V evaluated this method. We found that scaling gain nonlinearly allowed people to explore a virtual environment with more spatial awareness than simply scaling translational gain without a method of minimizing these small head movements (simply scaling gain linearly). We show that using this nonlinear scaling method, people can maintain a reasonable spatial orientation for gains up to 50. Thus, with a tracked HMD system, one can expect to explore a virtual space 50 times the size of the tracked space. For example, a 5m by 5m tracked HMD space allows users to explore a virtual space that is 250m by 250 m. This is a huge space gain.

In Experiment 7 of Chapter V, we also looked at spatial orientation when eyeheight was scaled proportionally to gain. Our motivation for doing this was that virtual reality allows user to experience environments in ways that they couldn't normally do in the real world and that manipulating eyeheight could give the user an advantage when exploring a large city where the user would have a map-like overview of the environment. However, we found that scaling the eyeheight proportionally to gain does not result in better spatial orientation than scaling gain using the user's normal eyeheight. Raising the eyeheight did raise an interesting issue about viewing angle with HMDs and its role on our ability to be spatial oriented in an environment.

Now that we have shown that we can scale gain up to 50, is there a way to explore a virtual space larger than 50 times the size of the tracking system? Inevitably, the physical limits of the tracking system will be reached. In Chapter VI presented methods that we developed to reset users when they reach the end of their physical space by changing their location in physical space while their location in the virtual environment was the same before and after resetting [Williams et al. 2006b; Williams et al. 2007a]. Specifically, we evaluated three resetting conditions Freeze-Turn, Freeze-Backup, and 2:1-Turn. The evaluation of these three different methods in this chapter found the smallest errors in locating remembered targets in the Freeze-Backup condition, while latencies were lowest for the 2:1-Turn condition. Updating one's position after a reset in the Freeze-Backup condition involves ignoring proprioceptive cues resulting while taking steps backward while position are frozen in virtual space. The lower turning errors is generally consistent with prior literature indicating that it is easier to judge changes in perspective when the geometry of the change is a translation rather than a rotation [Rieser 1989; Presson and Montello 1994; Philbeck et al. 2001]. Given this fact, we were surprised to find the 2:1-Turn condition is almost as good as Freeze-Backup condition. The Freeze-Turn condition reported the highest errors and latencies. Although the Freeze-Backup condition has the lowest turning errors, a commodity-level HMD system would likely consist of either the Freeze-Turn or 2:1-Turn methods since the Freeze-Backup method has particular disadvantages such as the longer length of time involved in the resetting, and the potential to back into a wall or trip over the tether. We generally prefer the 2:1-Turn method since resets are relatively fast, and gain resetting turn (which was 2:1) could be manipulated so that the amount a user turns during a reset would result in them facing the direction that allows them the most forward walking space.

This thesis thus develops engineering solutions that allows people to effectively explore virtual environments that are much larger than the physical tracked space while maximizing their spatial orientation. We explore several engineering solutions to designing this human-computer interface while using psychological experimentation to evaluate these techniques. The techniques presented in this thesis can be implemented and scale easily to any tracked HMD system.

Space will always be a limiting factor when using bipedal locomotion to explore virtual environments, especially if they are to be used in homes, schools, and offices where there is not a huge amounts of space readily available. Therefore, overcoming this limitation is important. Physical locomotion will always result in the best spatial orientation so leveraging these proprioceptive cues to explore virtual environments is desirable. Treadmill gain could also be scaled which would reduce the amount of time and energy required to explore a large virtual environment.

We would like to deploy immersive virtual environments widely, for learning and training, and it seems likely that physical space is a constraint that must be overcome for their widespread adoption. The end goal of this work is to create a system that allows people to explore a virtual environment of any size. Given any environment, the system would find the best parameters that would allow the person to explore the virtual environment. Thus, we intend to combine techniques discussed in this dissertation to develop an integrated, effective virtual system. The system could be based in an office or small lab. In particular, if immersive virtual environments are to realize their potential as commodity-level components, a perceptually accurate interface that allows locomotion through them within the constraints of everyday space must be developed.

VII.2 Future Directions

We would like to revisit many of the topics discussed in this thesis in future work. Our research will continue to look at how environments can be used to aid in education and in learning. More specifically, we intend to build a program to develop a better understanding of the cognitive capabilities of humans in virtual environments, and do so in a way that informs the design process about virtual environments and our understanding of how humans reason about space. Even state-of-the-art virtual environments are usually unconvincing, and people have difficulty organizing their spatial knowledge of them and moving around in them. Our research will improve our understanding of how people reason about space in a virtual environment and how that understanding can be technically leveraged into an improved interface.

One obvious extension of this work that we plan to look at is to combine scaled translational gain and resetting into a fully integrated system. We have shown in Chapters V and VI that it is possible to maintain spatial orientation with scaled translational gain scaled up to 50 and when resets occur. Thus, an issue of future work is to address how to trade gain against resets. We believe that a system that combines them both in a reasonable manner can be found. This will either involve a system with a high gain and few resets, or a low gain with many resets. This tradeoff may depend upon the size of the physically tracked space. Thus, we plan to also investigate what role the size of the tracked space plays on developing such a system.

Experiment 7 of Chapter V raised some interesting questions regarding the role of eyeheight on spatial orientation in a virtual environment. We would like to revisit this topic in future work. Specifically, we would like to fix eyeheight relative to different gains. We feel that increasing eyeheight proportionally to gain in our experiments resulted in participants being too high in the virtual environment. Additionally, as a follow up of Experiment 7 of Chapter V, we would like to look specifically at people's ability to learn the spatial layout of objects at different viewing angles by having subjects memorize objects of different heights across conditions. We believe that the 1:1 condition in Experiment 7 was confounded because the objects appeared on the ground and users had to look downward to view and memorize the locations of the objects in this condition.

With regard to the resetting results of Chapter VI, we would like to test and develop a method of resetting that results in the largest explorable area. In other words, the 2:1-Turn method could be manipulated so that the user turns more or less than 180° depending on what turn would give the most amount of forward-space to explore. Also, as an extension of this resetting work, we plan to test participants in the 2:1-Turn condition to see how well they perform the resetting with no experimenter intervention.

Chapter V shows that we can explore a virtual environment with a gain of 50:1 using the nonlinear scaling method discussed in the chapter. We would like to compare this 50:1 gain with other methods of exploring a large virtual environment such as a joystick or the walking-in-place method of Slater et al. [1995]. An advantage that this method has over the walking-in-place method is that it allows users to explore virtual environments with no additional interventions.

In Chapter IV, we examine results of gamers and non-gamers as a sub-study of Experiment 3. The results were inconclusive, but perhaps with more data one might be able to draw a more interesting conclusion. We intend to specifically look at how well gamers do at 50:1 gain versus non-gamers. In the study reported in this thesis, the gain was 10:1. This may not be high enough for experience to matter.

Long term, we plan to develop interfaces that leverage cues that people use to navigate in virtual environments. This research would quantify the types of disorientation that people experience and determine what specific cues are lacking or diminished in traditional virtual representations that allow this disorientation to occur. We intend to build scaffolding to overcome the constraints of virtual environment technology. An example of this scaffolding would be a method of rendering the virtual environment to make it easier for people to find their way in a virtual environment. However, a virtual compass would be an example of an interface that we would not want to provide since it gives information that a person might need to rely upon when performing a task in the physical world. Thus, it would not be an example of the scaffolding we are seeking. Additionally, we want to examine whether self-representation improves the virtual experience. Most virtual environment simulations do not support any sense of self-representation, although some research has shown it to be helpful. For rendering, we intend to investigate whether techniques such as stroke-based and painterly rendering can enhance people's spatial awareness. This research will determine whether non-photorealistic rendering (NPR) effects can be fruitfully used as scaffolding that may compensate or alleviate problems that people have in learning and orienting themselves in virtual environments.

In conclusion, we plan to look ahead to issues that will necessarily arise if virtual environment technology is to be used in large-scale, commodity-level applications. The results of this research could be employed to make virtual environments better tools for learning and training. Our particular focus is on issues of learning in large spaces, where wayfinding and orientation are critical components. The spatial reasoning abilities of people are not completely understood for physical environments, and even less is known when simulations of space are presented through virtual environments technology. We plan to significantly advance our understanding of these issues. The impact of this research will be to broaden the application of virtual environments by making the technology better and more accessible.

REFERENCES

[1997]ALLEN, R. C., AND SINGER, M. J. 1997. Landmark direction and distance estimation in large scale. In *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting*, 1213–1217.

[1999]BOWMAN, D. A., DAVIS, E. T., HODGES, L. F., AND BADRE, A. N. 1999. Maintaining spatial orientation during travel in an immersive virtual environment. *Presence 8*, 6, 618–631.

[1991]BRICKEN, M. 1991. Virtual reality learning environments: potentials and challenges. *SIGGRAPH Comput. Graph. 25*, 3, 178–184.

[2003]BROCKMOLE, J. R., AND WANG, R. F. 2003. Changing perspectives within and across environments. *Cognition 87*, B59–B67.

[1998]CHANCE, S. S., GAUNET, F., BEALL, A. C., AND LOOMIS, J. M. 1998. Locomotion mode affects updating of objects encountered during travel: The contribution of vestibular and proprioceptive inputs to path integration. *Presence 7*, 2, 168–178.

[1992]CRUZ-NEIRA, C., SANDIN, D. J., DEFANTI, T. A., KENYON, R. V., AND HART, J. C. 1992. The cave: audio visual experience automatic virtual environment. *Commun. ACM 35*, 6, 64–72.

[1995]CUTTING, J. E., AND VISHTON, P. M. 1995. *Perception of Space and Motion*, vol. 5 of *Handbook of perception and cognition*. Academic Press, San Diego, CA, ch. Percieving layout and knowing distances: The integration, relativie potency, and contexual use of different information and depth, 69–117.

[1996]DARKEN, R., AND SIBERT, J. 1996. Navigating in large virtual worlds. *The Int. J. of Human-Computer Interaction 8*, 1, 49–72.

[1997]DARKEN, R. P., COCKAYNE, W. R., AND CARMEIN, D. 1997. The omni-directional treadmill: a locomotion device for virtual worlds. In *UIST '97: Proceedings of the 10th annual ACM symposium on User interface software and technology*, ACM Press, New York, NY, USA, 213–221.

[2000]D'ZMURA, M., COLANTONI, P., AND SEYRANIAN, G. D. 2000. Visualization of events from arbitrary spacetime perspectives. In *Proc. SPIE Vol. 3960, p. 35-40, Visual Data Exploration and Analysis VII, Robert F. Erbacher; Philip C. Chen; Jonathan C. Roberts; Craig M. Wittenbrink; Eds.*, R. F. Erbacher, P. C. Chen, J. C. Roberts, and C. M. Wittenbrink, Eds., vol. 3960 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, 35–40.

[1995]EASTON, R. D., AND SHOLL, M. J. 1995. Object-array structure, frames of reference, and retrieval of spatial knowledge. *J. Exp. Psych. 21*, 483–500.

[1994]ELLIS, S. R. 1994. What are virtual environments? *IEEE Computer Graphics and Applications 14*, 1, 17–22.

[1998]FARRELL, M. J., AND ROBERTSON, I. H. 1998. Mental rotation and the automatic updating of body-center spatial relationships. *J. Exp. Psych: Learn., Mem., Cog. 24*, 993–1005.

[1990]FOLEY, J., VAN DAM, A., FEINER, S., AND HUGHES, J. 1990. *Computer Graphics: Principles and Practice*, 2nd ed. Addison-Wesley.

[2005]FOO, P., WARREN, W. H., DUCHON, A., AND TARR, M. 2005. Do humans integrate routes into a cognitive map? map versus landmark based navigation of novel shortcuts. *J. Exp. Psych: Learn., Mem., Cog. 31*, 195–215.

[2003]FRUEH, C., AND ZAKHOR, A. 2003. Constructing 3D city models by merging ground-based and airborne views. *IEEE Computer Graphics and Applications,édition spéciale* (Dec).

[1978]FURNESS, T. 1978. Visually coupled information systems. In *ARPA Conference on Biocybernetic Applications for Military Systems*.

[2001]GARDNER, P. L., AND MON-WILLIAMS, M. 2001. Vertical gaze angle: absolute height-in-scene information for programming of prehension. *Exp Brain Res 136*, 379–385.

[1977]GOGEL, W. 1977. *Stability and constancy in visual perception: Mechanisms and processes*. Wiley, New York, ch. The metric of visual space, 129–182.

[2006]GOLDSTEIN, E. B. 2006. *Sensation and Perception*, 7 ed. Wadsworth, Pacific Grove, CA.

[2001]HARRELL, JR., F. E. 2001. *Regression Modeling Strategies*. Springer Series in Statistics. Springer-Verlag, New York.

[1998]HEIM, M. 1998. *Virtual Realism*. Oxford UP.

[2007]INTERRANTE, V., RIES, B., AND ANDERSON, L. 2007. Seven league boots: A new metaphor for augmented locomotion through moderately large scale immersive virtual environments. In *IEEE Symposium on 3D User Interfaces 2007*, 167–170.

[1999]IWATA, H. 1999. The torus treadmill: Realizing locomotion in ves. *IEEE Comput. Graph. Appl. 19*, 6, 30–35.

[2001]KAY, B. A., AND WARREN, JR., W. H. 2001. Coupling of posture and gait: mode locking and parametric excitation. *Biological Cybernetics 85*, 2 (aug), 89–106.

[1998]KLATZKY, R. L., LOOMIS, J. M., BEALL, A. C., CHANCE, S. S., AND GOLLEDGE, R. G. 1998. Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. *Psychological Science 9*, 4 (July), 293–298.

[2004]Knapp, J. M., and Loomis, J. M. 2004. Limited field of view of head-mounted displays is not the cause of distance underestimation in virtual environments. *Presence 13*, 572–577.

[2004]Kuhl, S. A. 2004. Recalibration of rotational locomotion in immersive virtual environments. In *APGV '04: Proceedings of the 1st symposium on Applied perception in graphics and visualization*, 23–26.

[2003]Lambrey, S., and Berthoz, A. 2003. Combination of conflicting visual and non-visual information for estimating actively performed body turns in virtual reality. *Int. J. of Psychophysiology 50*, 1–2, 101–115.

[2002]Lathrop, W. B., and Kaiser, M. K. 2002. Perceived orientation in physical and virtual environments: Changes in perceived orientation as a function of idiothetic information available. *Presence 11*, 1, 19–32.

[2005]Lathrop, W. B., and Kaiser, M. K. 2005. Acquiring spatial knowledge while traveling simple and complex paths with immersive and nonimmersive interfaces. *Presence 17*, 3, 249–263.

[2004]Lok, B., and Hodges, L. 2004. *Encyclopedia of Human Computer Interaction*. Berkshire, Barrington, MA, ch. Human Computer Interaction in Virtual Reality.

[2005]Lokki, T., and Grohn, M. 2005. Navigation with auditory cues in a virtual environment. *IEEE MultiMedia 12*, 2, 80–86.

[2003]Loomis, J. M., and Knapp, J. M. 2003. *Virtual and Adaptive Environments*. Erl-Baum, Mahwah, NJ, ch. Visual perception of egocentric distance in real and virtual environments, 21–46.

[1992]Loomis, J. M., DaSilva, J. A., Fujita, N., and Fukusima, S. S. 1992. Visual space perception and visually directed action. *J. Exp. Psych: Hum. Perc. Perf. 18*, 4, 906–921.

[1999]Loomis, J. M., Klatzky, R. L., Golledge, R. G., and Philbeck, J. W. 1999. *Wayfinding: Cognitive mapping and other spatial processes*. Johns Hopkins University Press, Baltimore, MD, ch. Human navigation by path integration, 125–151.

[1987]Mark, L. 1987. Eyeheight-scaled information about affordances. a study of sitting and stair climbing. *jphh 13*, 361–370.

[2000]May, M., and Klatzky, R. L. 2000. Path integration while ignoring irrelevant movement. *J. Exp. Psych: Learn., Mem., Cog. 26*, 1 (Jan.), 169–186.

[1996a]May, M. 1996. Cognitive and embodied modes of spatial imagery. *Psychologische Beitr age 38*, 418–434.

[1996b]May, M. 1996. Thinking outside the body: An advantage for spatial updating during imagined versus physical self-rotation. *Psychologische Beitrge 38*, 418–434.

[2004]May, M. 2004. Imaginal perspective switches in remembered environments: Transformation versus interference accounts. *Cognitive Psychology 48*, 163–206.

[2004]Mohler, B. J., Thompson, W. B., Creem-Regehr, S., Pick, H. L., Warren, W. H., Rieser, J. J., and Willemsen, P. 2004. Visual motion influences locomotion in a treadmilll virtual environment. In *Symposium on Appliced Perception in Graphics and Visualization*, 19–22.

[2007a]Mohler, B. J., Thompson, W. B., Creem-Regehr, S. H., Willemsen, P., Herbert L. Pick, J., and Rieser, J. J. 2007. Calibration of locomotion resulting from visual motion in a treadmill-based virtual environment. *ACM Trans. Appl. Percept. 4*, 1, 4.

[2007b]Mohler, B. J., Thompson, W. B., Creem-Regehr, S. H., Willemsen, P., Herbert L. Pick, J., and Rieser, J. J. 2007. Calibration of locomotion resulting from visual motion in a treadmill-based virtual environment. *ACM Trans. Appl. Percept. 4*, 1, 4.

[2002]Myers, R. H., Montgomery, D. C., and Vining, G. G. 2002. *Generalized Linear Models*. Wiley Interscience, New York.

[2000]Newcombe, N. S., and Huttenlocher, J. 2000. *Making Space: The development of Spatial representation and reasoning*. MIT Press.

[2004]Nitzsche, N., Hanebeck, U., and Schmidt, G. 2004. Motion compression for telepresent walking in large target environments. *Presence 13*, 1, 44–60.

[2002]Oman, C. M., Shebilske, W. L., Richards, J. T., Tubr, T. C., Beall, A. C., and Natapoff, A. 2002. Three dimensional spatial memory and learning in real and virtual environments. *Spatial Cog. and Comp. 2*, 355–372.

[2001]Ooi, T. L., Wu, B., and He, Z. J. 2001. Distance determination by the angular declination below the horizon. *Nature 414*, 197–200.

[1995]Pausch, R., Proffitt, D., and Williams, B. 1995. Quantifying immersion in virtual reality. In *SIGGRAPH 95*, 399–400.

[1997]Pausch, R., Burnette, T., Brockway, D., and Weiblen, M. 1997. Quantifying immersion in virtual reality. In *Computer graphics and interactive techniques 1995*, 399–400.

[2000]Péruch, P., Belingard, L., and Thinus-Blanc, C. 2000. Transfer of spatial knowledge from virtual to real environments. In *Spatial Cognition II*, C. Freska, Ed. Springer, Berlin.

[2001]Philbeck, J. W., Klatzky, R. K., Behrmann, M., Loomis, J. M., and Goodridge, J. 2001. Active control of locomotion facilitates nonvisual navigation. *J. Exp. Psych: Hum. Perc. Perf. 27*, 141–153.

[1999]PICK, H. L., RIESER, J. J., WAGNER, D., AND GARING, A. E. 1999. The recalibration of rotational locomotion. *J. Exp. Psych: Hum. Perc. Perf. 25*, 5, 1179–1188.

[2004]PLUMERT, J. M., KEARNEY, J. K., AND CREMER, J. F. 2004. Child's perception of gap affordances: Bicycling across traffic-filled intersections in an immersive virtual environment. *Behavior Reseach Methods, Instruments, and Computers 75*, 1243–1253.

[2005]PLUMERT, J. M., KEARNEY, J. K., CREMER, J. F., AND RECKER, K. 2005. Distance perception in real and virtual environments. *ACM Transactions on Applied Perception 2*, 216–233.

[1994]PRESSON, C. C., AND MONTELLO, D. R. 1994. Updating after rotational and translational body movements: coordinate structure of perspective space. *Perception 23*, 1447–1455.

[1998]PSOTKA, J., LEWIS, S. A., AND KING, D. 1998. Effects of field of view on judgments of self-location. *Presence 7*, 4 (Aug.), 352–369.

[2001]RAZZAQUE, S., KOHN, Z., AND WHITTON, M. C. 2001. Redirected walking. *Eurographics Short Presentation.*

[2005a]RICHARDSON, A., AND WALLER, D. 2005. The effect of feedback training on distance estimation in virtual environments. *Applied Cog. Psych. 19*, 1089–1108.

[2005b]RICHARDSON, A., AND WALLER, D. 2005. Interaction with an immersive virtual environment corrects users' distance estimates. *J. of the Hum. Fact. and Ergonomics Soc. 49*, 3, 507–517.

[2004]RIECKE, B. E., AND BÜLTHOFF, H. H. 2004. Spatial updating in real and virtual environments: contribution and interaction of visual and vestibular cues. In *APGV '04: Proceedings of the 1st symposium on Applied perception in graphics and visualization*, 9–17.

[2005]RIECKE, B. E., HEYDE, M. V. D., AND BÜLTHOFF, H. H. 2005. Visual cues can be sufficient for triggering automatic, reflexlike spatial updating. *ACM Trans. Appl. Percept. 2*, 3, 183–215.

[1986]RIESER, J. J., GUTH, D. A., AND HILL, E. W. 1986. Sensitivity to perspective structure while walking with vision. *Perception 15*, 173–188.

[1994]RIESER, J. J., GARING, A. E., AND YOUNG, M. F. 1994. Imagery, action, and young children's spatial orientation: It's not being there that counts, it's what one has in mind. *Child Development 65*, 1262–1278.

[1995]RIESER, J. J., PICK, H. L., ASHMEAD, D. A., AND GARING, A. E. 1995. The calibration of human locomotion and models of perceptual-motor organization. *J. Exp. Psych: Hum. Perc. Perf. 21*, 480–497.

[1989]RIESER, J. J. 1989. Access to knowledge of spatial structure at novel points of observation. *J. Exp. Psych. 15*, 6, 1157–1165.

[1975]ROCK, I. 1975. *An introduction to perception.* MacMillan.

[2006]RUDDLE, R. A., AND LESSELS, S. 2006. For efficient navigational search, humans require full physical movement, but not a rich visual scene. *Psychological Science 17*, 6 (June), 460–465.

[1999]RUDDLE, R. A., PAYNE, S. J., AND JONES, D. M. 1999. Navigating large-scale virtual environments: What differences occur between helmet-mounted and desk-top displays? *Presence 8*, 2, 157–168.

[2001]RUDDLE, R. A. 2001. Navigation: Am i really lost or virtually there? *Engineering Psychology and Cognitive Ergonomics 6*, 135–142.

[1941]SATTERTHWAITE, F. E. 1941. Synthesis of variance. *Psychometrika 6*, 5, 319–316.

[1946]SATTERTHWAITE, F. E. 1946. An approximate distribution of estimates of variance components. *Biometrics Bulletin 2*, 6, 110–114.

[2007]SCHWAIGER, M., THUMMEL, T., AND ULBRICH, H. 2007. Cyberwalk: An advanced prototype of a belt array platform. *Haptic, Audio and Visual Environments and Games. HAVE 2007. IEEE International Workshop, 12-14 Oct.*, 50–55.

[1973]SEDGWICK, H. A. 1973. *The visible horizon: A potential source of visual information for the perception of size and distance.* PhD thesis.

[1980]SEDGWICK, H. A. 1980. *The Perception of Pictures.* Academic P, New York, ch. The geometry of spatial layout in pictorial representation, 33–90.

[2003]SHERMAN, W. R., AND CRAIG, A. B. 2003. *Understanding Virtual Reality: Interface, Application, and Design.* Morgan Kaufmann.

[1995]SLATER, M., USOH, M., AND STEED, A. 1995. Taking steps: The influence of a walking technique on presence in virtual reality. *ACM Trans. on Human Interaction 2*, 3, 201–219.

[2006]SNAVELY, N., SEITZ, S. M., AND SZELISKI, R. 2006. Photo tourism: Exploring photo collections in 3d. In *SIGGRAPH Conference Proceedings*, ACM Press, New York, NY, USA, 835–846.

[2007]SUMA, E., BABU, S., AND HODGES, L. 2007. Comparison of travel techniques in complex, mulit-level 3d environment. In *IEEE Symposium on 3D User Interfaces*, 149–155.

[2002]TARR, M. J., AND WARREN, W. H. 2002. The cave: audio visual experience automatic virtual environment. *Nature Neuroscience 5*, 1089–1092.

[1999]TEMPLEMAN, J. N., DENBROOK, P. S., AND SIBERT, L. E. 1999. Virtual locomotion: Walking in place through virtual environments. *Presence 8*, 6, 598–617.

[2004]THOMPSON, W. B., WILLEMSEN, P., GOOCH, A. A., CREEM-REGEHR, S. H., LOOMIS, J. M., AND BEALL, A. C. 2004. Does the quality of the computer graphics matter when judging distances in visually immersive environments. *Presence 13*, 560–571.

[1999]USOH, M., ARTHUR, K., WHITTON, M. C., BASTOS, R., STEED, A., SLATER, M., AND BROOKS, F. P. 1999. Walking > walking-in-place > flying, in virtual environments. In *SIGGRAPH 99*, 359–364.

[2002]VIJAYAKAR, A., AND HOLLERBACH, J. M. 2002. A proportional control strategy for realistic turning on linear treadmills. *H aptics 00*, 231.

[1998]WALLER, D., HUNT, E., AND KNAPP, D. 1998. The transfer of spatial knowledge in virtual environment training. *Presence 7*, 2, 129–143.

[2002]WALLER, D., MONTELLO, D. R., RICHARDSON, A. E., AND HEGARTY, M. 2002. Orientation specificity and spatial updating of memories for layouts. *J. Exp. Psych: Learn., Mem., Cog. 28*, 6, 1051–1063.

[2004]WALLER, D., BEALL, A. C., AND LOOMIS, J. M. 2004. Using virtual environments to assess directional knowledge. *J. Exp. Psych. 24*, 105–116.

[2005a]WANG, R. F. 2005. Beyond imagination: Perspective change problems revisted. *Psicolgica 26*, 25–38.

[2005b]WANG, R. F. 2005. Beyond imagination: Perspective change problems revisted. *Psicolgica 26*, 25–38.

[1987]WARREN, W. H., AND WHANG, S. 1987. Visual guidance of walking through apetures: Body-scaled information for affordances. *J. Exp. Psych: Hum. Perc. Perf. 13*, 371–383.

[1984]WARREN, W. H. 1984. Perceiving affordances: Visual guidance of stair climbing. *J. Exp. Psych: Hum. Perc. Perf. 10*, 683–703.

[1998]WARTENBERG, F., MAY, M., AND PÉRUCH, P. 1998. Spatial orientation in virtual environments: Background considerations and experiments. In *Spatial Cognition, An Interdisciplinary Approach to Representing and Processing Spatial Knowledge*, Springer-Verlag, London, UK, 469–489.

[2002]WILLEMSEN, P., AND GOOCH, A. A. 2002. Perceived egocentric distances in real, image-based and traditional virtual environments. In *IEEE Virtual Reality Conference*, 79–86.

[2004]WILLEMSEN, P., COLTON, M. B., CREEM-REGEHR, S. H., AND THOMPSON, W. B. 2004. The effects of head-mounted display mechanics on distance judgements in virtual environments. In *APGV '04: Proceedings of the 1st Symposium on Applied perception in graphics and visualization*, 36–48.

[2006a]WILLIAMS, B., NARASIMHAM, G., MCNAMARA, T. P., CARR, T. H., RIESER, J. J., AND BODENHEIMER, B. 2006. Updating orientation in large virtual environments using scaled translational gain. In *APGV '06: Proceedings of the 3rd symposium on Applied perception in graphics and visualization*, ACM Press, New York, NY, USA, 21–28.

[2006b]WILLIAMS, B., NARASIMHAM, G., RUMP, B., MCNAMARA, T. P., CARR, T. H., RIESER, J., AND BODENHEIMER, B. 2006. Exploring large virtual environments with an hmd on foot. In *APGV '06: Proceedings of the 3rd symposium on Applied perception in graphics and visualization*, ACM Press, New York, NY, USA, 148–148.

[2007a]WILLIAMS, B., NARASIMHAM, G., RUMP, B., MCNAMARA, T. P., CARR, T. H., RIESER, J. J., AND BODENHEIMER, B. 2007. Exploring large virtual environments with an hmd when physical space is limited. In *APGV '07: Proceedings of the 4th symposium on Applied perception in graphics and visualization*, ACM Press, New York, NY, USA.

[2007b]WILLIAMS, B., NARASIMHAM, G., WESTERMAN, C., RIESER, J., AND BODENHEIMER, B. 2007. Functional similarities in spatial representations between real and virtual environments. *ACM Trans. Appl. Percept. 4*, 2, 12.

[1998]WITMER, B. G., AND SADOWSKI, W. J. J. 1998. Nonvisually guided locomotion to a previously viewed target in real and virutal environments. *Human Factors 40*, 478–488.

[1998]WITMER, B. G., AND SINGER, M. J. 1998. Measuring presence in virtual environments: A presence questionnaire. *Presence 7*, 3 (jun), 225–240.

[1996]WITMER, B. G., BAILEY, J. H., KNERR, B. W., AND PARSONS, K. C. 1996. Virtual spaces and real world places: transfer of route knowledge. *Int. J. Hum.-Comput. Stud. 45*, 4, 413–428.

[2000]WRAGA, M., CREEM, S. H., AND PROFFITT, D. R. 2000. Updating displays after imagined object and viewer rotations. *J. Exp. Psych: Learn., Mem., Cog. 26*, 151–168.

[2004]WRAGA, M., CREEM-REGEHR, S. H., AND PROFFITT, D. R. 2004. Spatial updating of virtual displays during self- and display rotation. *Memory and Cognition 32*, 3 (Apr.), 399–415.

[1999]WRAGA, M. 1999. Using eye height in different postures to scale the heights of objects. *J. Exp. Psych: Hum. Perc. Perf. 25*, 518–530.

[2003]WRAGA, M. 2003. Thinking outside the body: An advantage for spatial updating during imagined versus physical self-rotation. *J. Exp. Psych: Learn., Mem., Cog. 29*, 993–1005.

[2004]WU, B., OOI, T. L., AND HE, Z. J. 2004. Perceiving distance accurately by a directional process of integrating ground information. *Nature 428*, 73–77.