Updating orientation in large virtual environments using scaled translational gain

Betsy Williams,* Gayathri Narasimham, Tim P. McNamara,* Thomas H. Carr,* John J. Rieser,* and Bobby Bodenheimer*
Vanderbilt University

Abstract

Navigating through large virtual environments using a head-mounted display (HMD) is difficult due to the spatial limitations of the tracking system. We conducted two experiments to examine methods of exploring large virtual spaces with an HMD under translation conditions different than normal walking. Experiment 1 compares locomotion in the virtual environment using two different motor actions to translate the subject. The study contrasts user learning and orientation of two different translational gains of bipedal locomotion (not scaled and scaled by ten) with joystick locomotion, where rotation in both locomotion interfaces is accomplished by physically turning. Experiment 2 looks further at the effects of increasing the translational gain of bipedal locomotion in a virtual environment. A subject's spatial learning and orientation were evaluated in three gain conditions where each physical step was: not scaled, scaled by two, or scaled by ten (1:1, 2:1, 10:1, respectively). A sub-study of this experiment compared the performance of people who played video games against people who did not.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual Reality;

Keywords: Virtual environments, head-mounted display, locomotion, spatial updating

1 Introduction

Immersive virtual environments provide people with opportunities to experience environments remote from their actual physical surroundings. Currently, the technology for this experience is not at the commodity level; however, it seems that head-mounted display (HMD) technology may be at this level in the next several years. Other immersive virtual technologies, such as virtual caves, are less likely to achieve commodity status since they involve greater expense in the form of large screens, projectors, and often a locomotion device such as a treadmill or bicycle that allows a user to move about the environment. Moreover, integrating a commodity level virtual reality system with a commodity level motion capture system such as that described by Chai and Hodgins [2005] could have a huge impact on the use of virtual environments in learning, training, and entertainment. Since the HMD holds the promise of being readily available to the public, our focus is on understanding and dealing with the underlying constraints of this technology. We

---
*{betsy.williams, gayathri.narasimham, t.mcnamara, tom.carr, j.rieser, bobby.bodenheimer}@vanderbilt.edu

believe that the most pressing of these, when HMDs are to be used at commodity level, are the spatial constraints of the HMD tracking system. This paper addresses the issue of exploring a virtual environment that is larger than the physical boundaries of a tracking system by manipulating the translational gain of walking.

More generally, devising methods whereby large virtual environments can be explored while preserving a user's spatial representation of the environment is an important problem. In this paper, we seek 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.

One way to increase the amount of space explored is to change the translational gain of the system, so that one step forward in the physical environment carries one several steps forward in the virtual. Rieser et al. [1995] showed subjects can recalibrate, although meaningful translational gain to explore large virtual environments is beyond the limits of what they found. If translational gain is scaled, two issues arise: (1) what to do about rotations, and (2) is increasing translational gain a complete solution? Pick et al. [1999] and Kuhl [2004] have shown that rotations can be scaled and recalibrated as well, but there are reasons to think such scaling is not desirable. First, rotations cost nothing in terms of space, as one can rotate in place. From a cognitive standpoint, rotations are more difficult to compute and can lead more easily to disorientation than can translations [Presson and Montello 1994; Rieser 1989; Rieser et al. 1995]. Finally, in most HMD setups both translation and rotation are measured relative to the helmet. We want people to be able to look around and examine their environments naturally, which could prove difficult or disorienting if rotations are scaled.

The second question is more interesting. How high can we increase the translational gain? This study limits it to a factor of ten, as head movement and other small motion becomes distracting at higher gains, although filtering or other processing techniques might alleviate those problems. However, inevitably one will run out of physical space or tracking area no matter how high the gain is scaled. Obviously, then, some method of resetting the environment in a cognitively friendly way is needed. That question is not addressed in this paper, although we return to it in the discussion of Section 5.

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 1 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 1996; Presson and Montello 1994; Rieser 1989; Wraga et al. 2000; Wraga 2003; Williams et al. 2006]. Moreover, bipedal locomotion gives the subject prioperceptive cues allowing for more accurate distance and direction estimation as shown by Loomis et al. [1992].

Experiment 2 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.

This paper is organized as follows. Section 2 discusses related work and places the present work in context. The research in this paper consisted of two main experiments. The first experiment is described and its results reported in Section 3; the second is described and results reported in Section 4. We assess the impact of these experiments in Section 5.

2 Background

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], but 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]. 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]. 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 feed-

back. 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.

In our study, we chose to scale the translational gain of physical translation because Rieser et al. [1995] and Mohler et al. [2006] have shown that people can quickly recalibrate to a new mapping between their own physical translation and visual input, and that this recalibration had no effect on the perception of rotational locomotion. However, the scaling factor of the translational gain in these recalibration studies was significantly smaller than that which is proposed in this paper. 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.

Our method of assessing the accuracy of a person's spatial knowledge is to measure their errors and latencies in tasks where they are asked to turn and face an object the location of which they had previously learned. In these types of experiments, response latencies and errors are related to the difference between actual facing direction and direction needed to face the target, a variable called disparity [May 2004]. For our purposes, 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. Therefore, the angle correct response when asked to face a target object from a given location is the disparity. We will assess the impact of disparity in the experiments that follow.

3 Experiment 1: 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.

3.1 Method

3.1.1 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.

3.1.2 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 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 1 and 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.
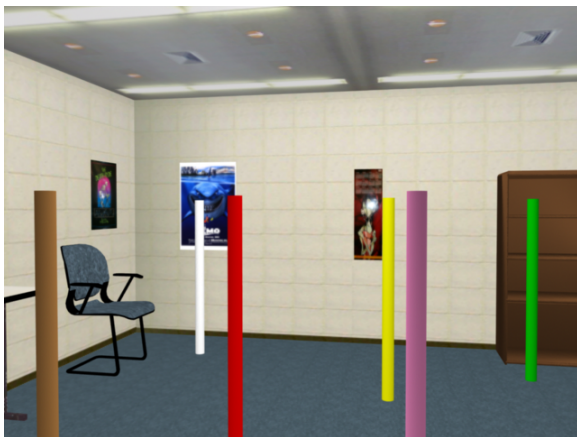

Figure 1: This figure shows the virtual environment of the 1:1 condition used in both Experiment 1 and 2. The target objects in this experiment were the different color cylinders, some of which can be seen in this image.

3.1.3 Procedure

One-half of the subjects performed the experiment in the 1:1 environment, the other half in the 10:1 environment. In both envi-

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

ronments, 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.

Partipants' 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

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

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 $20 - 180°$. 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.

3.2 Results

Table 1 shows subjects' mean turning errors and latencies by locomotion condition in the two virtual environments. A visual representation of the locomotion errors are shown in Figure 3. 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 4. Mean response time as a function of disparity is shown in Figure 5. 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], we have clustered the disparities to their closest $36°$ amount.

The independent variables included locomotion (walking versus manipulating a joystick), 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 4. The two-way interactions were non-significant, but the three-way interaction (of locomotion x translational gain x disparity) approached
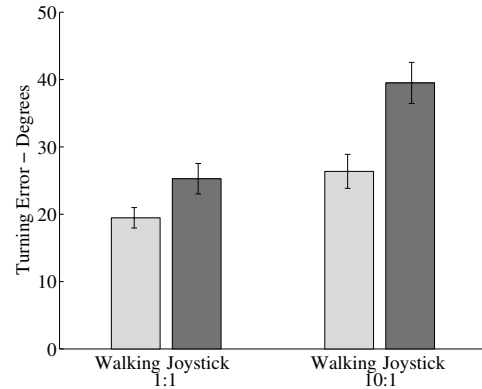
Figure 3: Mean turning error for the 1:1 Walking, 1:1 Joystick, 10:1 Walking and 10:1 Joystick conditions of Experiment 1. Error bars indicate the standard error of the mean. As discussed in Section 3.2, gain has a significant effect on turning error.
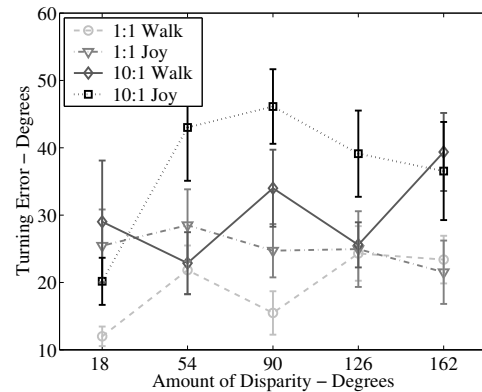
Figure 4: Mean turning error as a function of disparity in the 1:1 Walking, 1:1 Joystick, 10:1 Walking, and 10:1 Joystick conditions of Experiment 1. Error bars indicate the standard error of the mean.
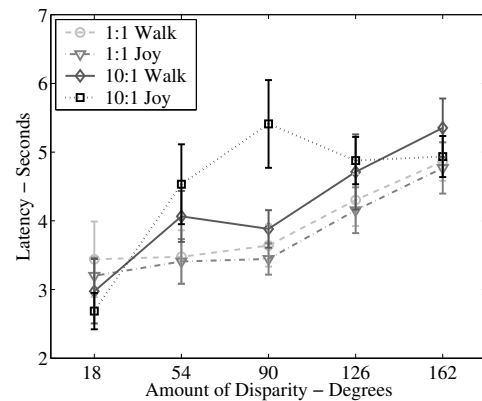
Figure 5: Latency as a function of disparity in the 1:1 Walking, 1:1 Joystick, 10:1 Walking, and 10:1 Joystick conditions of Experiment 1. Error bars indicate the standard error of the mean.

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, we did separate analyses of covariance 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, $F(1,759) = 26.8$, $p < .001$.

4 Experiment 2: Effects of translational gain and subject expertise

In the second experiment, our 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 tested a subjects' spatial knowledge 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.

4.1 Method

4.1.1 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.

4.1.2 Materials

The HMD and tracker used in Experiment 1 were used again in Experiment 2. The 1:1 and the 10:1 virtual environments of Experiment 1 as shown in Figures 1 and 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 6. In all three environments, subjects were asked to memorize the location

Figure 6: 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 2 are shown in Figures 1 and 2, respectively.

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.

4.1.3 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 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 1, 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 1. 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 1, these trials were also designed so that the disparity was evenly distributed in

	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 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.

the range of $20 - 180°$.

As a sub-study, we compared the results six gamers and six non-gamers. We define gamers 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 we divide our original subject pool to balance these orders, for a six gamer versus six non-gamer comparison, three gamer cells were empty. We recruited three additional gamers 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.

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 1), 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 2 shows the mean turning error and latencies as functions of translational gain. Figure 7 shows a visual representation of the turning error. 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 8 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 9. 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 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, $F(4, 40) =$
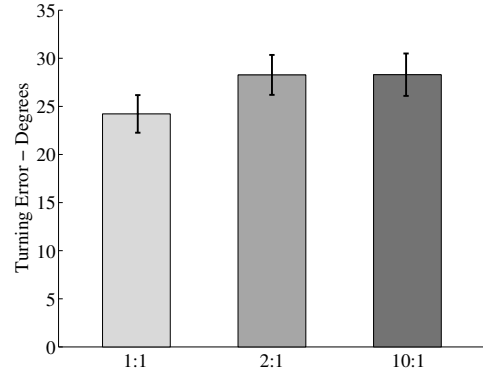

Figure 7: Mean turning error for the 1:1, 2:1, and 10:1 scaled translational gain conditions in Experiment 2. Error bars indicate the standard error of the mean. As discussed in the Section 4.2, turning errors show no significant main effect.
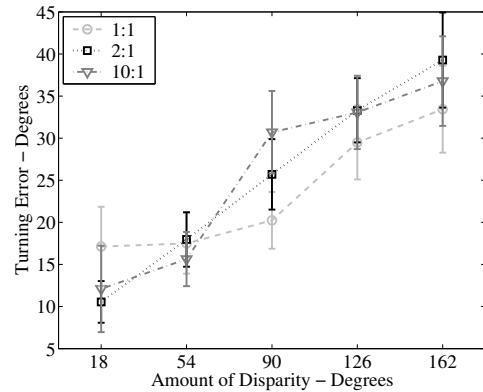

Figure 8: Mean turning error as a function of disparity in the 1:1, 2:1, and 10:1 conditions of Experiment 2. Error bars indicate the standard error of the mean.
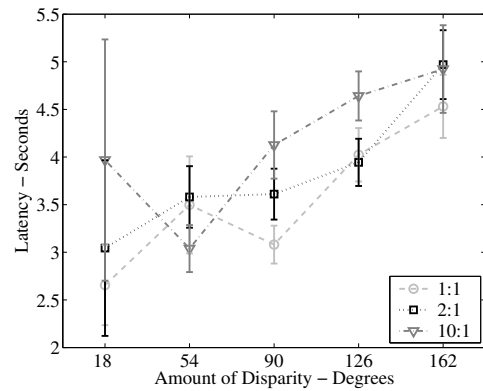

Figure 9: Latency as a function of disparity in the 1:1, 2:1, and 10:1 conditions of Experiment 2. Error bars indicate the standard error of the mean.

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.

5   Discussion

This paper 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.

People explore spaces, both virtual and physical, by rotating and translating their perspectives as they examine the environment. 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. 2006]. Although there is no problem in rotating by physically turning in a small physical room to explore a large virtual space, there is a problem in translating. This problem occurs when trying to fit long distances walked through virtual space into a small physical space.

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. 2006], we studied two alternative methods that varied in the amount of physical walking. 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 1 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 1, 1:1 and 10:1 gains were manipulated across subjects in two different conditions, and in Experiment 2 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 1, gain exerted a significant effect on the errors, but not on the latencies. In Experiment 2, 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 1 but not in Experiment 2. 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 2 tested them within subjects. The higher errors in the 1:1 gain condition of Experiment 2 compared to Experiment 1 may be due to the subjects' additional experience of the higher gain conditions in Experiment 2. 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.

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 1 and Experiment 2 (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 2 varied in their general experience with some of the features of our learning and test situation, though they did not vary in experience with the specific features of our 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, our 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], our results 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 we assessed in Experiment 2.

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, they 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—we found generally comparable results for gains varying from 1:1 to 10:1, but to explore a battlefield or city, one would need to use much larger gains or one would need to use an additional strategy. Our current research is focused on one such additional strategy, to "reset" subjects when they walk and reach the end of their physical space. We are working to engineer different methods so that when the person comes close to the physical walls, they are alerted by a signal, and then change their position in the physical room, while maintaining their perspective relative to the virtual environment. We believe that the technique presented in this paper 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. An advantage of the walking-in-place method of Slater et al. [1995] is that it allows users to explore large environments with no additional interventions. Therefore, future work should compare their method with our scaled translational gain technique.

Acknowledgments

References

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.

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

CHAI, J., AND HODGINS, J. K. 2005. Performance animation from low-dimensional control signals. In *SIGGRAPH 05*, 15–22.

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.

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

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.

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.

KUHL, S. A. 2004. Recalibration of rotational locomotion in immersive virtual environments. In *APGV 2004*, 23–26.

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.

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.

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.

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

MAY, M. 2004. Imaginal perspective switches in remembered environments: Transformation versus interference accounts. *Cog. Psych. 48*, 163–206.

MOHLER, B., THOMPSON, W. B., CREEM-REGEHR, S. H., WILLEMSEN, P., PICK, H., AND RIESER, J. J. 2006. Calibration of locomotion due to visual motion in a treadmill-based virtual environment. *ACM Trans. on Applied Perception*. To appear.

NITZSCHE, N., HANEBECK, U., AND SCHMIDT, G. 2004. Motion compression for telepresent walking in large target environments. *Presence 13*, 1, 44–60.

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.

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

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

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 2004*, 9–17.

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.

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

RUDDLE, R. A., AND LESSELS, S. 2006. For efficient navigation search, humans require full physical movement but not a rich visual scene. *Psych. Science*. To appear.

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.

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.

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

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.

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

WILLIAMS, B., NARASIMHAM, G., WESTERMAN, C., RIESER, J., AND BODENHEIMER, B. 2006. Functional similarities in spatial representations between real and virtual environments. *Trans. on Applied Perception*. To appear.

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.

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.