Human Joystick: Wii-Leaning to Translate in Large Virtual Environments

Alyssa Harris, Rhodes College
Kevin Nguyen, Pomona College
Preston Tunnell Wilson, Rhodes College
Matthew Jackoski, Rhodes College
Betsy Williams*, Rhodes College

Abstract

In this work we present an inexpensive method of exploring a large immersive virtual environment (VE). More specifically, we describe a method of physically leaning to explore a VE on a Nintendo Wii Fit Balance Board, Wii–Leaning. In two experiments described in this paper, we directly compare our Wii-Leaning method to two other methods of exploring a large VE. In Experiment 1, the feasibility and accuracy of our Wii-Leaning is tested by directly comparing it to joystick navigation. In a spatial orientation task, we find that when participants explore the environment by physically leaning that their spatial representation of the VE is more accurate. In Experiment 2, we then directly compare our Wii-Leaning method to "walking in place" using a spatial orientation task. We find that spatial orientation is similar in the two conditions and that Wii-Leaning was more preferred 2:1 by users.

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

Keywords: locomotion, head-mounted display, spatial updating

1 Introduction

Research has shown that virtual environments (VEs) are best explored on foot [Ruddle and Lessels 2006]. However, when a person uses their own physical locomotion to explore a VE, they are limited by the range of the tracking system or the size of the physical room housing the tracking system. Thus, foot exploration is not suitable for exploring a large VE. Moreover tracking systems are expensive, and thus not well-suited for widespread use. This paper presents a new method of exploring a large VE using a head-mounted display (HMD). Specifically, we examine the idea of physically leaning or tilting as a means of translating the user forward in virtual space. This method does not suffer from space limitations. Moreover, our leaning method, Wii–Leaning, is implemented using two inexpensive Nintendo Wii Fit™ Balance Boards (100 USD). In Experiment 1 presented in this work, we test the effectiveness of this method by directly comparing the spatial awareness when users locomote with our lean method and a joystick. In both methods, physical rotations correspond to virtual rotations and users translate forward in the yaw angle they are facing. We show that spatial orientation is increased when users physically lean as compared to joystick exploration. Since we show that our Wii-Leaning method is better than a joystick, we directly compare Wii–Leaning to another inexpensive method of exploring a large VE which has been shown to out perform the joystick, "walking in place." Thus, in Experiment 2, we compare Wii–Leaning to a walking in place (WIP) algorithm implemented using two Kinect sensors[Wilson et al. 2014]. We show that WIP and leaning are similar in terms of spatial awareness. However, most participants prefer leaning. Moreover, leaning may have more desirable characteristics as discussed in Section 6. The present work adds to the body of research indicating that there is value added when the whole body moves to translate in a VE even if that movement does not match normal bipedal locomotion. Some examples of this type of work are WIP [Williams et al. 2011] and scaling the translational gain of walking [Williams et al. 2006].

*e-mail:sandersb@rhodes.edu

VRCAI 2014, November 30 – December 02, 2014, Shenzhen, China.

http://dx.doi.org/10.1145/2670473.2670512

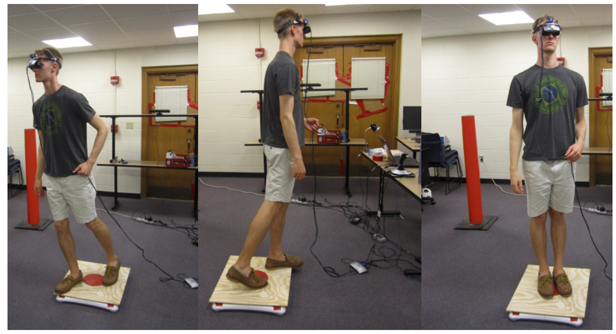

Figure 1: *Demonstration of our Wii–Lean method. During the experiment, the lights were off and there was no natural light.*

In this work we directly compare participants' spatial orientation in two experiments. Spatial orientation refers to the natural ability of humans to maintain their body orientation and position relative to the surrounding environment. An accurate sense of spatial orientation is necessary to successfully navigate through an environment. Spatial orientation is used to evaluate navigation because learning the layout and the information in the environment is often the goal of a virtual system. For example, a student walking home from school must identify his or her location and direction within the surrounding environment before determining in which direction to proceed. This sense of spatial orientation relies heavily on visual information and whole-body information while moving in an environment [Wartenberg et al. 1998]. In other words, both environmental cues and path integration (the process of integrating self-motion cues over time) inform spatial orientation. Thus, spatial orientation is generally tested by performing experiments that require participants to combine the use of environmental cues and path integration. We assess a user's spatial orientation in a VE by measuring turning errors and latencies in tasks where subjects are asked to turn–to–face previously learned target objects [May 1996].

2 Background

Other methods of exploring large VEs exist, such as walking in place (WIP) [Williams et al. 2011], redirected walking [Hodgson et al. 2008], and scaling the translational gain of walking[Williams et al. 2006]. Like redirected walking, scaling the translational gain of walking (where one step forward carries one several steps forward in virtual space) requires an expensive tracker. WIP and redirected walking both require a walking movement from the user. Thus, walking in a large environment might take considerable time and physical energy. WIP can be implemented inexpensively and has been shown to be equivalent in terms of spatial orientation to walking [Williams et al. 2011]. However, the user can only

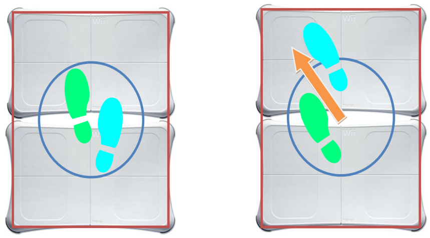

Figure 2: *When users' center of mass is more than 7in from the center of the platform, they translate in the yaw gaze direction.*

move in whole step increments. Joysticks lack the proprioceptive cues of the other methods mentioned above. Valkov et al. [2010] present a Wii Fit Balance Board VE exploration method where forward/backward leaning results in forward/backward translation, while left/right leaning rotates the VE in that direction. A user study testing feasibility of this algorithm has not been reported.

In the current work we directly compare spatial learning while WIP to our Wii–Leaning technique. This WIP algorithm was implemented from skeletal tracking data obtained from two Microsoft Kinect sensors (100 USD), a method we refer to as (WIP–K) [Wilson et al. 2014]. This WIP–K algorithm works by translating the user forward when the angle of their knee reaches a certain threshold as indicated from the Kinect skeletal tracking data. Other systems to explore a large VE involve large screen caves with a locomotion input such as a bicycle or treadmill. Treadmill systems are difficult and expensive to construct with enough degrees of freedom to allow for free exploration. Other expensive hardware solutions to explore a large VE include the omni directional treadmill Cyber Walk [Schwaiger et al. 2007].

3 Wii-Leaning Algorithm

Our algorithm works by tracking the user's center of mass across a plywood platform. If the user's center of mass moves from the center of the board, then our algorithm detects a lean and translates the user 3m/s in the direction the user is gazing until their center of mass is back in the center of the board. In our system, physical rotations represent virtual rotations. Thus, we had to be careful not to detect a lean if the user was simply moving to turning around and unintentionally shifting their center of mass from the center of the board. Our 24in by 24in lean platform as seen in Figure 1 was made from two Nintendo Wii Fit Balance Boards. The Wii Balance Board (20.1in by 12.4in) contains four pressure sensors at each of the four corners. Thus, our Wii–Leaning method uses the weight measurements obtained from each of the four pressure sensors. We felt that one balance board was too small to implement a reasonable leaning algorithm. Thus, to give the participant more space to lean, we placed two Wii Balance Boards placed side by side. We used industrial Velcro to fixate a single piece of plywood (24in by 24in) to the top of the both boards. Centered on top of this platform was a round piece of cardboard with a thickness of 1/8" covered with red duct tape. This raised piece of cardboard was positioned in the center so that user could feel the center of the board while wearing the HMD. The cardboard center serves as a "home base" and helps users to stop translating. To translate forward, users gaze in the desired direction, take a step, and lean in that direction.

We calculated the user's center of mass, or centroid, using the data obtained from all 4 corners of each of the Wii Boards. A translation required that the user's center of mass be outside of a radius of 7in from the center of the board (as shown in Figure 2). Additionally, to detect a lean, the lean direction (displacement of weight from the center of the board) must be within 30° of the viewing yaw angle. The user had to look and lean within 30° for .25sec before translating. Additionally, to prevent unintentional translation while turning, if the yaw viewing angle and leaning direction are changing at rates faster than 20°/sec and 120°/sec, respectively, then our algorithm does not translate the user. These measures were evaluated in pilot studies as successfully preventing unintentional movement while maximizing intentional lean detection. Once a lean is detected, users move forward at a speed of 3 m/s until their mass centriod is within 7in of the center of the platform. We used a large testing environment and in pilot studies users found slower speeds boring and inconsistent with the method of leaning. Thus, we chose a jogging pace of 3m/s. We also investigated the idea of varying the speed based on the extent to which the user was leaning, but this proved difficult to control due to the relatively small size of the platform. While translating, the user can adjust their direction by moving their head slightly to the left or right. Slight movements of less than 30° over .2sec results in a change of translation direction in the users' yaw gaze direction. A change in yaw angle faster than 30° over .2sec would not change the user's translation direction, but instead allows them to look around as they translate.

4 Experimental 1: Wii–Leaning vs Joystick

This experiment compares two different methods to translate the subjects in virtual space: Wii–Leaning and Joystick. In each condition, we asked subjects to remember the locations of six different objects in a large virtual outdoors environment. Then, to test spatial orientation, we asked subjects to move (by Leaning or Joystick) to a new point of observation and instructed them to turn so that they would face in the direction of the target from memory without vision. The speed and accuracy of turing to face these remembered target objects without vision tells us how spatial oriented the subjects are under a locomotion condition.

Ten subjects participated in this experiment. Subjects were unfamiliar with the experiment and virtual environments. The VE was viewed through an eMagin Z800 head-mounted display which had a resolution of 800x600, a field-of-view of 40° diagonally, and an optical frame rate of 60Hz. A black cover was placed around the HMD to block the participants' view of the real world. An InterSense IS-900 tracker was used to update the participant's rotational movements around all three axes. The WiiYourself! C++ library was used to acquire data about the user's weight distribution from the balance boards. The joystick used in this experiment was a Logitech Extreme 3D Pro. The VE used in the experiment was a 50m by 50m grassy plain with a generic backdrop depicting the sky and surrounding countryside. For each condition, six different objects were placed at specific positions in the plain. These six target objects were arranged in a particular configuration, such that the configuration in both conditions varied only by a rotation about the center axis. However, the particular target objects changed across conditions. In this manner, the angles of correct response were preserved across both conditions. The random order of trials and the different objects concealed the fact that the arrangement was the same throughout the experiment. Examples of some of the target objects were a dog, a bookshelf, a piggy bank, and a harp. Objects were similar in size and height. Additionally, we used one object in the field to familiarize the subject with the set up of the experiment before the target memorization phase. We used a cylinder and an arrow to indicate the position and orientation, respectively, for a particular trial.

Each of the ten participants explored each of the environments under the two different translational locomotion conditions, Wii–Leaning and joystick locomotion. In both conditions, physical rotation matched rotation in the VE. In the joystick condition, participants rotated physically and moved in the direction of their gaze by joystick translation. In both conditions, the user translated forward at an optical flow of 3 m/s. We chose a jogging speed because our testing environment was considerably large. Since there were two orders of the two locomotion conditions, half subjects of the subjects completed the Wii–Leaning condition first, while the other half completed the joystick condition first. The experimental procedure was fully explained to the subjects prior to seeing the VE. Before the learning phase of each condition, the participant was shown the environment with a red cylinder and a single object that did not appear in our test set as an example. Participants performed several practice trials in this environment so that they were familiar with the setup and the experimental design. After the subject understood the task and the condition, the practice target object disappeared. During the learning phase, subjects were asked to learn the positions of the six target objects while freely moving around the VE. The participant practiced the locomotion condition by moving to various target objects. After a period of study, the experimenters tested the subject by having them move towards certain objects and turn to face other objects. This testing and learning procedure was repeated until the subject felt confident that the configuration had been learned and the experimenters agreed. The experimenters were the same across all the subjects so this process was consistent.

Participants' spatial knowledge was tested from six different locations. A given testing position and orientation was indicated to the subject by the appearance of a red cylinder and a red arrow in the environment. Participants were instructed to locomote to the cylinder until it turned green (indicating that they were in the correct testing location) and then to turn to face the arrow until it turned green (indicating that they were in the correct testing orientation). When the subject reached this position, the room and the objects were hidden so that the participant only saw the cylinder and the arrow in a black environment. From this testing location, the subject was told by verbal instruction to "turn to face the <target name>". In the VE, the cylinder (testing location) and arrow (orientation) disappeared. Once the participant indicated that they had turned to face the object, the angle turned, the angle of correct response, and the latency of turning to face the object were recorded. The cylinder and arrow reappeared. The participant then turned back to face the arrow until it turned green again. The subject was instructed to face another target object. At each location, the subject completed three trials by turning to face three different target objects in the room, totaling in 18 trials per condition. After completing the three trials at a particular testing location, the participant was asked to look at the arrow until it turned green so that they would not receive any feedback on their performance. Then the room and objects were displayed again. Subjects were encouraged to re–orient themselves after completing a testing location. After the room and objects were shown again and the participants had re–oriented themselves, the cylinder and arrow were moved to the next target location. The order of testing locations was randomized per condition per subject so that the subject could not tell that the conditions were similar.

To compare the angles of correct response, we used the same trials for each condition. The testing location and target locations were analogous in all conditions, and target order varied randomly. The trials were designed so that the disparity was evenly distributed in the range of 20–180°. The testing locations were positioned in such a way that they would never turn to face a target object closer than 0.8m. To assess the degree of difficulty of updating orientation relative to objects in the VE, we recorded the latencies and turning errors. 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 of the final facing direction minus the correct facing direction. The subjects indicated to the experimenter that they were facing the target by verbal instruction. The latency was recorded by computer, and the rotational position was recorded using the InterSense tracker. Subjects were encouraged to respond as rapidly as possible while maintaining accuracy.

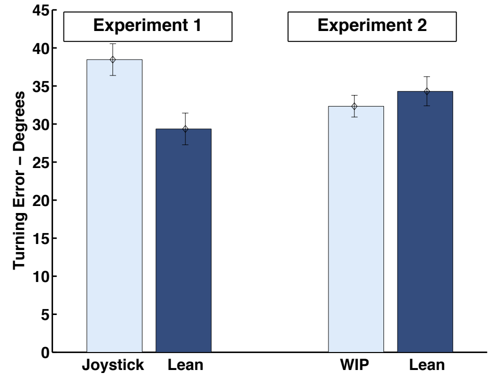

Figure 3: *Mean turning error of Experiments 1 and 2. Error bars show standard errors of the mean.*

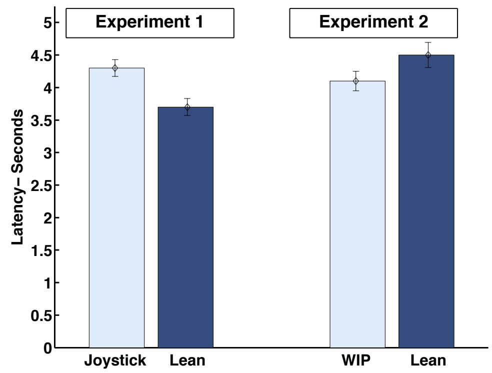

Figure 4: *Mean latency of Experiments 1 and 2. Error bars show standard errors of the mean.*

Figures 3 and 4 show the subjects' mean turning errors and latencies by locomotion condition in the VE. We compared mean turning error on the two conditions using a repeated measures ANOVA with order between subjects. The analysis revealed a significant effect of condition, $F(1, 11) = 5.8, p = .03$. Thus, participants in the Lean condition responded more accurately than those in the joystick condition. We did not find an effect of latency across the conditions. We also analyzed turning errors and latency and did not find an effect of order. Subjects' performance did not improve based on which condition they did first.

5 Experiment 2: Wii–Leaning vs WIP

This experiment compares Wii–Leaning and WIP with the Kinect (WIP–K) in terms of spatial orientation and user preference. Twelve subjects participated in this within–subject experiment. The details of the WIP–K algorithm can be found in [Wilson et al. 2014]. Each of the 12 participants explored the environments under the two different counterbalanced translational locomotion conditions: WIP and Wii–Leaning. In both conditions, physical rotation matched rotation in the VE, and participants moved in the direction of their gaze. In the WIP–K condition subjects were told not to physically translate and to stay on top of a .75m by .75m cardboard board that was positioned on the floor. We used this to make sure participants did not drift out of the range of the Kinect trackers. Participants were asked to take a break between each condition to better facilitate the transition from the cardboard mat (used to keep participants in range of the Kinect sensors) to the Wii Board platform. WIP–K steps result in 3 m/s optical flow.

Figures 3 and 4 show the subjects' mean turning errors and latencies by locomotion condition in the VE. We compared mean turning error on the two conditions using a repeated measures ANOVA with order between subjects. We did not find an effect of condition on errors. However, we compared latency on the two conditions using a repeated measures ANOVA with order between subjects. The analysis revealed a significant effect of condition, $F(1,13) = 8.1, p = .02$. Thus, participants in the WIP–K condition responded faster than those in the Lean condition. Additionally, users were asked which method they preferred and users preferred the Lean-Wii condition 2:1. That is, eight of the twelve participants preferred the Wii–Leaning method.

6 Discussion

Prior work has shown that users can have difficulty maintaining spatial orientation in a VE [Péruch et al. 2000]. Moreover, under certain circumstances users take significantly longer to learn VEs than comparable real environments [Richardson et al. 1999], and often produce large random and systematic errors in VE navigation [Riecke 2008]. The reasons for this are not well understood, but many cues are known to affect how well people can maintain their spatial orientation. If the potential for immersive VEs is to be realized, a perceptually accurate interface that allows virtual within the constraints of everyday space must be developed. Finding ways to spatially navigate in VEs that perform comparably to the way we navigate in the real world is a challenging and important problem. We show that physically leaning to explore a VE does aid in spatial awareness over joystick exploration. We also show that spatial awareness is similar to WIP–K. This is an important finding because it could inform the design of future VE locomotive devices.

As mentioned previously, we had promising prior results with WIP and have begun exploring its potential as a feasible solution for exploring large VEs. However, there are problems with WIP. The WIP algorithm used in this experiment limits the user to explore a space in whole step increments. Second, WIP may not be a suitable solution to exploring a significantly large VE because it would take considerable time and physical energy. Thus, WIP to explore a town or city might be exhausting. Leaning does not seem to suffer from these constraints. However, we would like to improve our leaning method by allowing users to change their speed based on how much they lean. We think we will have to look at other hardware solutions for inexpensively implementing a leaning algorithm. In the experiment that directly compared the Wii–Leaning and WIP–K algorithms, we found that subjects responded faster when they were in the WIP–K condition. We hypothesize that the reason for this is that people were worried about slipping off of the Wii Balance Board platform.

Users preferred to navigate the VE by leaning 2:1 over WIP–K, and in many cases they self reported that they *felt* more oriented when they leaned to explore the environment. While one participant expressed that the faster speed was more congruent with the Leaning method than with WIP, we hypothesize that the majority of the participants who cited Leaning as their preferred method felt that way because our Leaning algorithm enabled more localized movement. Additionally, leaning to explore a VE, while possibly similar to a Segue, is not a method that is employed by most people on a daily basis. Walking, even when staying in the same place, however, is something with which people have much more experience. Thus, in order to employ a WIP algorithm for exploring a VE that will win users' preference, that algorithm must meet as much as possible users' expectation of the real life equivalent of that method. In the future, we plan to improve our WIP algorithm by allowing steps of various sizes and allowing users to lean to look closer at their virtual surroundings. We think these realistic features will combine the usefulness of WIP and the user friendliness of Lean-Wii into a robust new method of exploring VEs.

References

HODGSON, E., BACHMANN, E., AND WALLER, D. 2008. Redirected walking to explore virtual environments: Assessing the potential for spatial interference. *ACM Trans. Appl. Percept. 8*, 4 (Dec.), 22:1–22:22.

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

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.

RICHARDSON, A. E., MONTELLO, D. R., AND HEGARTY, M. 1999. Spatial knowledge acquisition from maps and from navigation in real and virtual environments. *Memory & cognition 27*, 4, 741–750.

RIECKE, B. E. 2008. Consistent left-right reversals for visual path integration in virtual reality: More than a failure to update one's heading? *Presence: Teleoperators and Virtual Environments 17*, 2, 143–175.

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.

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.

VALKOV, D., STEINICKE, F., BRUDER, G., AND HINRICHS, K. H. 2010. Traveling in 3d virtual environments with foot gestures and a multi-touch enabled wim. In *Proceedings of Virtual Reality International Conference (VRIC 2010)*, 171–180.

WARTENBERG, F., MAY, M., AND PÉRUCH, P. 1998. Spatial orientation in ves: Background considerations and experiments. In *Spatial Cognition, An Interdisc. Appr. to Rep. and Proc. Spat. Knowledge*, Springer, London, 469–489.

WILLIAMS, B., NARASIMHAM, G., MCNAMARA, T. P., CARR, T. H., RIESER, J. J., AND BODENHEIMER, B. 2006. Updating orientation in large ves using scaled translational gain. In *3rd Symp. on Appl. Percept. in Graph. and Vis.*, ACM, 21–28.

WILLIAMS, B., BAILEY, S., NARASIMHAM, G., LI, M., AND BODENHEIMER, B. 2011. Evaluation of walking in place on a wii balance board to explore a ve. *ACM Trans. Appl. Percept. 8*, 3 (Aug.), 19:1–19:14.

WILSON, P. T., NGUYEN, K., HARRIS, A., AND WILLIAMS, B. 2014. Walking in place using the microsoft kinect to explore a large ve. In *VRCAI '14: Proceedings of the ACM SIGGRAPH International Conference on Virtual–Reality Continuum and its Applications in Industry*, ACM, New York, NY, USA, to appear.