Comparison of Locomotion Interfaces for Immersive Training Systems

Linda E. Sibert[1] James N. Templeman[1] Roy M. Stripling[1] Joseph T. Coyne[1]
Robert C. Page[1] Zina La Budde[1,3] Daniel Afergan[2]
[1]Naval Research Laboratory, Washington, DC
[2]Strategic Analysis, Inc., Arlington, VA
[3]Currently at Alion Science and Technology, McLean, VA

The ability to stay oriented in an environment is an important skill for urban combat. Warriors must systematically clear areas of responsibility while moving tactically, scanning for potential targets, and engaging threats. A locomotion interface for an immersive virtual environment urban combat training system should enable the development of navigational skills. One aspect of navigation is path integration, in which a person estimates current position and orientation along a pathway from velocity and acceleration. Path integration allows people to stay oriented in low visibility or the dark, which is important in many tactical situations. This study compared the performance of three locomotion interfaces on three path integration tasks. Analysis revealed that type of locomotion interface has a significant impact on performance.

INTRODUCTION

An immersive virtual environment (VE) to train urban combat requires a locomotion interface that allows the user to move through the virtual world. Here, an immersive VE consists of a head-mounted display (HMD), a real-time 6-degree-of-freedom motion capture system, and multiple computers. Urban combat demands a high level of navigational skill. Trainees must learn to maintain their orientation while moving tactically, scanning for potential targets, and engaging threats. *Path integration* is one aspect of navigation and allows people to estimate their current position and orientation relative to an origin (Loomis & Knapp, 1999). It is not based on fixing position through landmarks or maps, but uses velocity and acceleration signals from kinesthetic and vestibular input to update an estimate of self-position along a pathway. Path integration is what allows people to stay oriented in low visibility or the dark, which is important in many tactical situations. A locomotion interface for a VE to train urban combat skills should enable path integration. The present study evaluated three locomotion interfaces used to direct the movement of the user's avatar, the representation of the user's body in the virtual world. The three path integration tasks, point-to-origin (with visual cues), blind walking, and blind rotation, are commonly used to assess spatial knowledge.

BACKGROUND

The three locomotion interfaces were programmed in-house; one is a research prototype and two represent commercially available approaches. They provide different vestibular and kinesthetic feedback, which might affect path integration performance.

Gaiter (Templeman, Denbrook & Sibert, 1999; Templeman & Sibert, 2006) is a high-end system in which the user's body motion directly drives the user's avatar. The user walks or runs in place to translate through the virtual world and physically turns to rotate. A harness is used to center the user because people tend to drift forward while walking in place. The tracked motion of the user's legs determines the direction and timing of the virtual steps, and raising and lowering one leg equals one virtual step. Because Gaiter is a leg-based interface, it provides kinesthetic feedback from muscles, tendons, and joints similar to real walking, with some vestibular sensing from the minor cyclical motion of the head when stepping in-place. It does not include any translational actions or forces that occur with walking. Since turning is achieved by turning the body, users receive full vestibular and kinesthetic feedback.

The *Rifle-Mounted Joystick (R-M joystick)* is similar to Quantum3D's ExpeditionDI and Atlantis Cyberspace, Inc.'s Immersive Group Simulator (IGS). The user controls the direction and rate of translation by pushing the thumb-joystick mounted on the handguard of the rifle prop in the desired direction, relative to where the rifle is pointed. The thumb-joystick is pushed forward to go forward, up to go right, back to go backwards, and down to go left. The user physically turns in the real world to rotate in VE and is centered by a circular railing. The R-M joystick combines an indirect control for translation with real turning. The user receives limited kinesthetic feedback from the thumb on the joystick and full vestibular sensing and kinesthetic feedback from turning.

10.1518/107118108X348882

The *desktop joystick* employs the type of joystick used in flight simulator games. The user controls all locomotion (rotation, translation, and direction of movement) with the joystick, which sits on a desk and is operated from a seated position. Direction of movement is in the direction of view and the velocity of movement (rate) is proportional to the amount the joystick is deflected. A hat control tilts the view and a reset button re-centers it. The desktop joystick provides only limited kinesthetic feedback from the hand operating the joystick and no vestibular feedback.

The path integration tasks required the subjects to use movement-based information to update their estimates of location. Two tasks, *point-to-origin* and *blind walking,* have been shown to assess path integration in the real world (Loomis, Klatzky, Golledge, and Philbeck, 1999). The *blind rotation task* isolated an important component of navigation, turning without an external frame of reference.

Ruddle, Payne, and Jones (1997) used a point-to-origin technique to assess subjects' spatial knowledge of the layout of a building in a VE. Our version had subjects walk through a maze consisting of three rooms connected by two straight hallways before pointing back to the origin in the final room. Each maze required an intervening turn in the middle room comprising different angles. The layout was adapted from a study by Klatzky, Loomis, Beall, Chance, and Golledge (1998). The blind walking task, in which a person looks at an object and walks to that location without vision, taking either a direct or indirect path (taking a path obliquely oriented to the target location before turning and walking to that location), is a popular approach for studying perceptually directed action in both the real world (Rieser, Ashmead, Talor, & Youngquist, 1990; Philbeck, Loomis, & Beall, 1997) and VE (Loomis & Knapp, 2003; Thompson et al., 2004; Interrante, Anderson, & Ries, 2006; Richardson & Waller, 2007). While the results in the real world have produced good correspondence, the results in VE have shown a consistent underestimation of distance. There is recent evidence that the underestimation can be practically eliminated (to approximately 95%) if the VE is either a high fidelity, low latency representation of the real environment (Interrante, Anderson, & Ries, 2006) or when feedback is provided between two blocks of blind walking, by having the subject walk to the object with vision (Richardson & Waller, 2007). Performance in the post-feedback block showed significantly less underestimation. The two exceptions suggest the importance of calibration. In the present study, subjects were given time to familiarize themselves with the experimental space. Because this study compares locomotion interfaces, subjects also had to calibrate to the interface itself (most blind walking studies in VE use real walking for locomotion). Using a blind walking task to compare locomotion interfaces is novel because of the need to calibrate to both the environment (to lessen underestimation) and the interface. In general, we expected that interfaces providing more vestibular and kinesthetic feedback would lead to better performance on the path integration tasks.

METHOD

Participants

Thirty U.S. Naval Research Laboratory (NRL) employees and contractors volunteered to participate without compensation. Four withdrew because of problems with the equipment and could not reschedule. Two withdrew due to simulator sickness but both indicated minor discomfort before beginning. Consequently, 24 subjects (21 male and 3 female) completed the experiment. All subjects had normal or corrected to normal vision.

Apparatus

All subjects wore an NVIS NVISOR SX HMD and Sennheiser HD 280 Precision headphones to limit sound localization. For Gaiter, the subjects' legs, torso, and HMD were tracked using a Vicon 612 passive optical tracking system, providing position and orientation information for each tracked body segment. A centering harness, built by Dr. Roger Kaufmann from the George Washington University, kept the subject in the tracked area. Subjects held a conventional gamepad to register trigger presses (trigger presses were used to mark events). For the R-M joystick, the subject's torso, HMD, and rifle prop were tracked using a PhaseSpace active optical tracking system that provided the position and orientation of the tracked segments. The AirSoft rifle prop was modified to register trigger presses and a thumb-joystick was attached to the left side of the handguard. Subjects stood within a circular railing 89 cm in diameter at waist level. With both Gaiter and the R-M joystick, all cables were routed overhead to allow the subject to turn freely. The desktop joystick used a Thrustmaster Top Gun Fox 2 Pro that was fixed to the table to maintain forward alignment and correspondence to the VE. The subject could move the chair to achieve a comfortable position. The HMD was not tracked and eye height in the VE was set to 166 cm.

Tasks

The *point-to-origin task* required subjects to walk through a maze consisting of three rooms (two 3 m square rooms with a 3 m diameter circular room in the middle) connected by two straight hallways (the first 9 m and the second 5 m). Each maze required a 45º, 90º, or 135º intervening turn in the circular room before pointing back to the origin in the final room. Each angle was repeated twice. The VE included full access to visual cues. The walls were simple texture maps designed to minimize landmarks but provide optic flow and connecting doors were invisible when closed. The *blind walking task* took place in a 108 m by 60 m virtual warehouse that had three 1.25 m wide paths marked on the floor with colored stripes. A diagonal path in the middle of the warehouse and one at a 60º angle to it were used in the task. During the task, the stripes forming the relevant path were visible to help the subjects maintain lateral stability. The subject faced down the diagonal path to view a 180 cm steel plate of a combatant similar to those used in real world training. Targets were placed on the path at 3.33 m, 6.66 m, or 10 m from the start location (repeated three times each for direct and indirect). The target was shown resting on the floor because it has been shown that angular elevation provides a strong distance cue (Philbeck & Loomis, 1997). The *blind rotation task* took place in the same virtual warehouse. Rotations were 45º, 90º, or 180º left and right (each angle/direction repeated two times).

Procedure

Subjects were randomly assigned to one of three experimental groups by interface. After the subjects were introduced to the experiment team, they were asked to sign an informed consent form, approved by NRL's Institutional Review Board, and complete a background questionnaire. Their inter-pupillary distance was measured to provide a customized stereo image in the HMD. The experimenters reviewed the operation of the locomotion interface. Subjects using the desktop joystick were reminded that the HMD was not tracked and the image would not change with head movement. R-M joystick subjects were told that an effective strategy for maintaining a stable frame of reference was to keep the rifle locked under the arm and pointed forward. Subjects were first immersed in a virtual combat town to practice using their interface until they felt comfortable, which took at least 5 min. A 5 min rest period was provided after each immersion. During the rest period following a task, the subjects were asked to explain their strategies. The instructions during the tasks were prerecorded and automated. The task order was point-to-origin, blind walking, and blind rotation.

For *point-to-origin*, practice consisted of subjects being shown pictures of interior views of the maze (not a top down map) before being immersed. In the last room, crosshairs appeared in the image of the right eye (to eliminate the need to fuse a stereo image) when the subjects reached the center. Subjects rotated to point the crosshairs back at the origin and pressed the trigger. They were then virtually transported back without vision to start the next trial. Subjects completed six experimental trails.

The next two tasks, path integration and blind rotation, took place in a virtual warehouse without a break so subjects could maintain their calibration to the environment. To encourage calibration, subjects practiced by moving through the warehouse filled with clutter. The objects included a manikin of the same height as the steel plate combatant target used in the blind walking task and the paths used in both.

For *blind walking*, direct and indirect trials were interspersed so subjects could not pre-program their responses. The subjects were instructed to acquire a good sense of the location of the target and keep track of its location as they moved quickly and decisively to it. After viewing the target, the image faded and only the parallel stripes delineating the initial path were visible. The stripes went out to infinity to not provide a distance cue. Subjects were instructed to either go forward (direct) or turn left and go (indirect). For direct, the stripes remained visible throughout the task. For indirect, subjects turned to the left and moved down the visible path until a tone sounded 5 m from the start location. Subjects then turned toward the location of the target object as the stripes disappeared, moved to the object in darkness, and pressed the trigger. Subjects did not receive feedback about their accuracy, and after each trial, they were transported without vision to begin the next trial. Before beginning the eighteen experimental trials, subjects practiced direct and indirect blind walking three times each.

For *blind rotation*, after each trial, subjects were virtually transported back to the start position in the virtual warehouse, but their orientation in the physical room did not change. The subjects practiced two times before beginning the twelve experimental trials.

At the end of the experiment, subjects were asked to respond to common symptoms of simulator sickness adapted from Ruddle, Payne, and Jones (1999).

Analysis

For the point-to-origin task, the accuracy of the subjects' performance was calculated by taking the absolute value of the difference between the subjects' direction judgment and the actual angle to the origin. For the blind walking task, only the direct walking data were analyzed due to technical difficulties with data collection during indirect trials. The distance traveled was calculated by projecting the subjects' location at trigger press onto the centerline of the path, similar to Loomis, Da Silva, Fujita, & Fukusima (1992). For the blind rotation task, the accuracy of the subjects' performance was calculated by taking the absolute value of the difference between the subjects' rotation angle (calculated by taking the difference between the subjects' initial and final orientations) and the requested angle. Data from all three tasks were skewed, so they were normalized using an ln(x+1) transformation, following Howell (1992).

RESULTS

Point-To-Origin Task

A mixed factorial ANOVA was performed on the data with locomotion interface as the between subjects variable and turn angle as the within subjects variable. There was a significant main effect for locomotion interface ($F(2, 21) = 4.25$, $p = .028$). The data for locomotion interface are in Figure 1. Bonferroni post hoc comparisons indicated that the desktop joystick subjects had significantly less angular error than those using the R-M joystick.

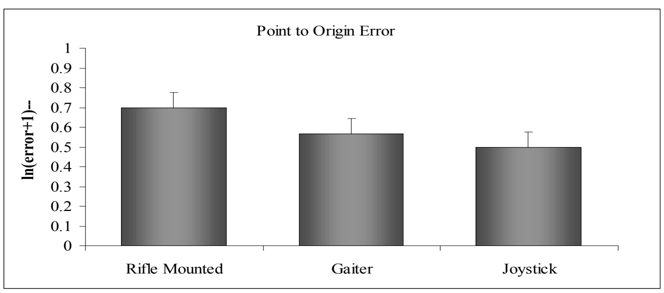

Figure 1. Effect of locomotion interface on the point-to-origin task. Error bars represent one standard error of the mean.

There was a significant effect of turn angle ($F(2, 42) = 3.374$, $p = .044$) on the point-to-origin data. However, Bonferroni post hoc comparisons did not reveal any significant differences between the different angles. The interaction of locomotion interface and turn angle was not significant.

Blind Walking Task

A mixed factorial ANOVA was performed on the data with locomotion interface as the between subjects variable and distance as the within subjects variable. Analysis revealed a significant main effect for locomotion interface ($F(2, 21) = 3.71$, $p = .042$), see Figure 2. Bonferroni post hoc comparisons indicated that the R-M joystick participants were more accurate than Gaiter.

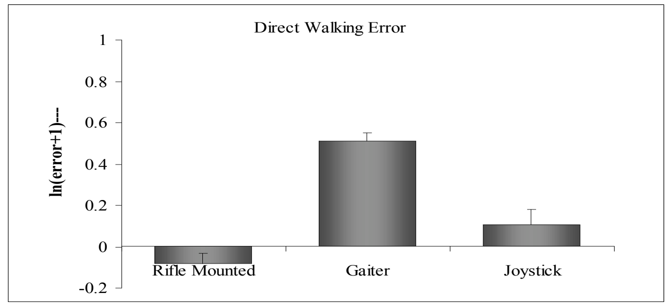

Figure 2. Effect of locomotion interface on the direct walking task. Error bars represent one standard error of the mean.

There was no significant main effect of distance or locomotion by distance interaction in the direct walking data.

Blind Rotation Task

A mixed factorial ANOVA was performed on the data with locomotion interface as the between subjects variable and angle as the within subjects variable. There was a main significant effect of locomotion interface ($F(2, 21) = 5.68$, $p = .011$), see Figure 3. Bonferroni post hoc comparisons indicated that the desktop joystick subjects had significantly more angular error than subjects using either Gaiter or the R-M joystick.

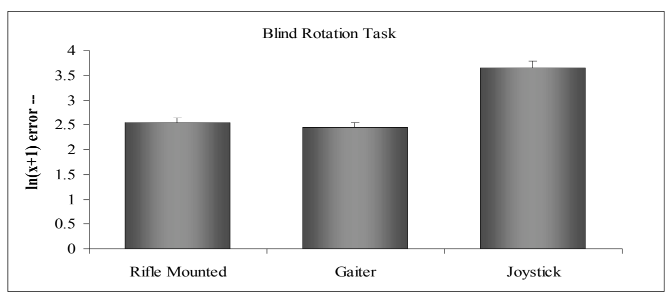

Figure 3. Effect of locomotion interface on the blind rotation task. Error bars represent one standard error of the mean.

There was a significant main effect of angle on the rotation task error ($F(2, 42) = 4.132$, $p = .023$). The error for blind rotation at 180 degrees (M =3.14 SD = 1.18) was significantly higher then the error at 45 degrees (M =2.74 SD = 1.12). Rotation at 90 degrees (M = 2.78 SD = 1.25) was not significantly different from the other two rotations. However, there was no significant interaction of interface and angle for the blind rotation task.

DISCUSSION

The overall results do not suggest one best locomotion interface for enabling path integration. All subjects reported only minor simulator sickness symptoms (less than .5 on a 0 to 3 scale), making all interfaces viable solutions. It was expected that more vestibular and kinesthetic feedback would lead to better performance. However, when visual cues were present during the maze task, subjects performed well with the desktop joystick, the purely indirect interface. The final response occurred in a square room that provided some cues for judging angular offsets so subjects might not have relied on path integration. When visual cues were not available for the rotation task, Gaiter and the R-M joystick outperformed the desktop joystick as expected, consistent with previous research (Chance, Gaunet, Beall, & Loomis, 1998). Gaiter and the R-M joystick provided vestibular sensing and kinesthetic feedback from real turning. For blind walking (direct), subjects using the R-M joystick accurately estimated distances. Further research is needed to examine the individual contributions of the subjects' ability to calibrate to the experimental space versus the interface to determine whether one compensated for the other. Gaiter subjects significantly overshot the distance. It was expected that subjects using the desktop joystick would overshoot the distance because it is a rate control device that works best with visual feedback. Gaiter subjects, however, were expected to do well because it is a leg-based control similar to real walking. Comments from the subjects suggested they misinterpreted an in-place step (one leg up and down) as a half step rather than a full step and traveled twice as far as necessary. After rescaling the Gaiter data, the results were comparable with the R-M joystick's results, suggesting that subjects might have made a mental calibration error.

While the present study does not conclusively recommend one best interface for enabling path integration, it suggests that an interface with more vestibular sensing and kinesthetic feedback provides a better overall solution. As immersive VE systems are adopted for training, it becomes important to develop new approaches for studying the interface's effect on performance, realizing the answer might not be simple.

ACKNOWLEDGEMENTS

The work was supported by the Office of Naval Research under its Virtual Technologies and Environments (VIRTE) program and the Naval Research Laboratory.

REFERENCES

Chance, S .S., Gaunet, F., Beall, A. C., & Loomis, J.M. (1998). Locomotion mode affects the updating of objects encountered during travel: The contribution of vestibular and proprioceptive inputs to path integration. *Presence*, *7*, 168-178.

Howell, D. C. (1992). *Statistical methods for psychology* (Third ed.). Belmont, CA: Duxbury Press.

Interrante, V., Anderson, L., & Ries, B. (2006). Distance perception in immersive virtual environments, revisited. *Proceedings of IEEE Virtual Reality 2006*, 3-10.

Klatzky, R. L., Loomis, J. M., Beall, A .C., Chance, S. S., & Golledge, R. G. (1998). Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. *Psychological Science*, *9*, 293-298.

Loomis, J. M., Da Silva, J. A., Fujita, N., & Fukusima, S. S. (1992). Visual space perception and visually directed action. *Journal of Experimental Psychology: Human perception and Performance*, *18*, 906-921.

Loomis, J. M., Klatzky, R. L., Golledge, R. G., & Philbeck, J. W. (1999). Human navigation by path integration. In R. G. Golledge (Ed.), *Wayfinding Behavior: Cognitive Mapping and Other Spatial Processes* (pp. 125-151). Baltimore, MD: Johns Hopkins University Press.

Loomis, J. M. & Knapp, J. M. (2003). Visual perception of egocentric distance in real and virtual environments. In L. J. Hettinger & M. W. Haas (Eds.), *Virtual and Adaptive Environments* (pp. 21-46). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Philbeck, J. W. & Loomis, J. M. (1997). Comparison of two indicators of perceived egocentric distance under full-cue and reduced-cue conditions. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 72-85.

Philbeck, J. W., Loomis, J. M., & Beall, A. C. (1997). Visually perceived location is an invariant in the control of action. *Perception & Psychophysics*, 59, 601-612.

Richardson, A. R. & Waller, D. (2007). Interaction with an immersive virtual environment corrects users' distance estimates. *Human Factors, 49*, 507-517.

Rieser, J. J., Ashmead, D. H., Taylor, C. R., & Youngquist, G. A. (1990). Visual perception and the guidance of locomotion without vision to previously seen targets. *Perception*, *19*, 675-689.

Ruddle, R. A., Payne, S. J., & Jones, D. M. (1997). Navigating buildings in "desk-top" virtual environments: Experimental investigations using extended navigational experience. *Journal of Experimental Psychology: Applied*, *3*, 143-159.

Ruddle, R. A., Payne, S. J., & Jones, D. M. (1999). Navigating large-scale virtual environments: What differences occur between helmet-mounted and desk-top displays? *Presence*, *8*, 157-168.

Templeman, J. N., Denbrook, P. S., & Sibert, L. E. (1999). Virtual locomotion: Walking in place through virtual environments. *Presence, 8,* 598–617.

Templeman, J. N. & Sibert, L .E. (2006). Immersive simulation of coordinated motion in virtual environments. In G. Allen (Ed.), *Applied Spatial Cognition: From Research to Cognitive Technology* (pp. 339-372). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Thompson, W. B., Willemsen, P., Gooch, A. A., Creem-Rehehr, S. H., Loomis, J. M., & Beall, A. C. (2004). Does the quality of the computer graphics matter when judging distances in visually immersive environments? *Presence*, *13*, 560-571.