Spring 2018

A Comparison of Head-Mounted Displays vs. Large-Screen Displays for an Interactive Pedestrian Simulator

Sophia Mallaro
University of Iowa

A COMPARISON OF HEAD-MOUNTED DISPLAYS VS. LARGE-SCREEN DISPLAYS FOR AN INTERACTIVE PEDESTRIAN SIMULATOR

by

Sophia Mallaro

A thesis submitted in partial fulfillment of the requirements
for graduation with Honors in the Electrical Engineering

Joe Kearney
Thesis Mentor

Spring 2018

All requirements for graduation with Honors in the
Electrical Engineering have been completed.

Xiodong Wu
Electrical Engineering Honors Advisor

A Comparison of Head-Mounted Displays vs. Large-Screen Displays for an Interactive Pedestrian Simulator

Sophia Mallaro

Abstract

This investigation compared how people performed a complex perception-action task – crossing traffic-filled roadways – in a CAVE vs. an HMD virtual environment. Participants physically crossed a virtual roadway with continuous cross traffic in either a CAVE-like or an HTC Vive pedestrian simulator. The 3D model and traffic scenario were identical in both simulators, allowing for a direct comparison between the two display systems. We found that participants in the Vive group accepted smaller gaps for crossing than participants in the CAVE group. They also timed their entry into the gap more precisely and tended to cross somewhat more quickly. As a result, participants in the Vive group had a somewhat larger margin of safety when they exited the roadway than those in the CAVE group. Participants in the CAVE group focused their gaze further down the road and had more variability in their gaze distances. The results provide a foundation for future studies of pedestrian behavior and other tasks involving full-body motion using HMD-based VR.

Introduction

Pedestrian injuries and fatalities in the U.S. have increased at an alarming rate over the last several years, with states reporting 2,660 pedestrian fatalities in the first six months of 2016 [Macek 2017]. This represents the steepest increase in pedestrian fatalities since record keeping began. Virtual environment technology offers tremendous potential for safely and systematically investigating risk factors associated with the interactions of pedestrians and vehicles, and to study interventions to mitigate these risks. Recent technological advances in displays, tracking systems, and simulation software platforms have dramatically reduced the cost and difficulty of creating highly immersive, interactive pedestrian simulators.

In addition to the intrinsic interest in understanding risks associated with road crossing and the influence of strategies to mitigate risk, road crossing is a good model system for testing how different virtual environment technologies influence performance in complex perception-action tasks. This paper presents an experiment that leverages a pedestrian road-crossing task to compare the two primary VR display technologies, large-screen and head-mounted display (HMD) systems.

Related Work

A number of research labs have developed interactive pedestrian simulators built on virtual reality technology. One of the challenges in designing a pedestrian simulator is to create a realistic experience of road crossing in a confined space. Researchers have devised a variety of ways to enable participants to virtually cross a road in such systems.

One way to expand the virtual travel distance in a large screen display system is to use a treadmill as an interface for locomotion. Banducci et al. [Banducci et al. 2016] placed a self-powered treadmill in a CAVE virtual reality system, which they used to study the influence of cell phone conversations and texting on pedestrian road crossing [Neider et al. 2011, 2010]. While this approach is effective in expanding travel distance, initiating and maintaining walking motion on a self-powered treadmill is somewhat effortful and unnatural.

*Significant portions of this report appear in: Sophia Mallaro, Pooya Rahimian, Elizabeth E. O'Neal, Jodie M. Plumert, and Joseph K. Kearney. 2017. A Comparison of Head-Mounted Displays vs. Large-Screen Displays for an Interactive Pedestrian Simulator. In Proceedings of ACM Symposium on Virtual Reality Software and Technology, Gothenburg, Sweden, November 8-10 2017 (VRST'17), 4 pages.

Another way to expand the virtual travel distance and allow natural walking motion in a CAVE is to extend length of the side displays. Taking advantage of wide-screen display technology, our lab has constructed two CAVE-like display systems with side screens that are 4.33m long and 2.44m tall. This makes it possible to render a one-lane road that crosses through the display volume that meets American Association of State Highway and Transportation Officials (AASHTO) design standards (2.74m - 3.66m for urban, local roads). In a series of experiments, we have examined how a variety of factors influence road crossing including age, texting, vehicle-to-pedestrian (V2P) communication systems, and crossing with a partner [Jiang et al. 2016; O'Neal et al. 2017; Rahimian et al. 2016].

Figure 1: The large-screen (CAVE) pedestrian simulator.

Researchers at the French Institute of Science and Technology for Transport (IFSTTAR) have also built a full-scale pedestrian simulator that allows participants to physically cross a virtual roadway. The display system consists of 10 rear-projection screens forming an elongated corridor [Dommés et al. 2015]. Participants can walk up to 7 m (23 ft), enough to cross a two-lane road. They have used this system to investigate how older adults cross roads with traffic.

Recent advances in head-mounted display (HMD) technology with integrated tracking systems, such as the Oculus Rift and the HTC Vive, have led to dramatic improvements in both the quality and cost of virtual reality systems. These offer the potential for low-cost, portable pedestrian simulators [Deb et al. 2017; Feldstein et al. 2016; Morrongiello and Corbett 2015; Morrongiello et al. 2015a,b]. The chief advantage of head-mounted systems as compared to large-screen systems is that they are highly immersive and much less costly to build. Potential disadvantages are that they typically have a restricted field of view and viewers cannot see their own body or objects in their hands (e.g., a cell phone). This, along with the encumbrance of a head-mounted display and the necessity of managing a cable that connects the display to a computer, may lead to unnatural movement. For example, Morrongiello et al. [Morrongiello et al. 2015b] found that children crossed (virtual) roads at a slower rate when wearing a head-mounted display system than they did when walking without wearing the display.

A number of studies have examined the differences in perception between HMD vs. CAVE virtual environment display systems. This includes investigations of distance perception and user interaction [Bowman et al. 2002; Steed and Parker 2005]. One recent study examined how people perform a complex perception-action task that involved a virtual train boarding task [Grechkin et al. 2014]. Participants attempted to step from a train platform onto a moving flatcar as it passed by the platform. A gate at the front of the platform periodically opened and closed. Participants were required to synchronize their boarding to both the opening of the gate and the approach of the flatcar. Participants performed the task in either an nVIS nVisor ST HMD or a CAVE. The main influence of display type was on the timing of motions - those in the CAVE condition walked more slowly but synchronized their movement with the arrival of the flatcar more tightly than those in the HMD condition.

Here, we present the results of an experiment comparing performance in the HTC Vive to a large-screen display system using the everyday task of crossing a traffic-filled roadway. The study differs from the train boarding study in both the task and the technology. Our CAVE improves on the previous study by using stereo imaging and head tracking to set the viewpoint for rendering. Importantly, the longer side screens of our CAVE allow participants to physically cross a one-lane road. The Vive represents the latest generation of consumer-level HMDs. It is much lighter than the nVIS system and packaged with a lowlatency tracking system that has a larger workspace than the electro-magnetic system used in the train boarding study. The low cost and portability of the Vive make it attractive for studies of pedestrian behavior and for training applications.

Methods

Virtual Environments

CAVE. The CAVE-like virtual environment consisted of three screens at right angles to one another, forming a three-walled room that was 3.05m wide x 4.33m long x 2.44m tall (see Fig. 1). Three DPI MVision 400 Cine 3D high-resolution projectors were used to back-project stereo images on the three walls. A fourth projector front-projected stereo images on the floor. A stereo sound system was used to generate spatialized traffic sounds.

An OptiTrack motion capture system tracked the movements of participants. The tracking system consisted of 17 Flex-13 infrared cameras mounted on the top and back of the CAVE. Participants wore a helmet with reflective markers to track their head position and orientation. The participant’s eye point was estimated from the head data and used to render the scene for the participant’s viewpoint. The OptiTrack system was also used to approximate gaze angle.

Participants viewed stereo images rendered for the estimated position and orientation of their eyes. A fixed inter-pupillary distance of 6.5 cm was used to render images of the left and right eyes. Participants wore Volfoni ActiveEyes stereo shutter glasses that were synchronized with the displays.

VIVE. The Vive virtual environment was based on the HTC Vive head-mounted display system using SteamVR. Participants wore stereo headphones to hear spatialized traffic sounds. The walking area was 7m wide x 3.92m long. The Vive comes with two LED emitting lighthouse boxes. These boxes flash and allow receivers on the headset to calculate its position and orientation. We used this orientation data to approximate gaze angle.

The virtual environment software for both systems was based on the Unity3D gaming platform. In-house code generated traffic and recorded the positions and orientations of vehicles and the participant during the experiment for later analysis.

Design and Procedure

We used a between-subjects design to compare road crossing behavior in the two types of virtual environments. Half of the participants were randomly assigned to the CAVE group and half to the Vive group. In both groups, participants physically crossed a one-lane virtual road with continuous traffic traveling from left to right. Cars travelled at the local, residential speed limit of 42.23 km/h (25 mph), with randomly ordered temporal gaps between cars ranging from 2.5-5 seconds with half-second intervals.

After a brief introduction to the virtual neighborhood, participants performed a single practice road-crossing trial with the experimenter. The experimenter instructed participants to watch the traffic and cross when they thought the gap between cars was large enough for them to reach the other side of the road without getting hit by a car. The traffic ceased to be generated after participants reached the other side of the road. Participants then walked back to the starting place and a new trial commenced. Participants then completed 20 road-crossing test trials on their own. The duration of the experiment was approximately 30 min.

Measures

We focused on three main aspects of road crossing: gap selection, movement timing, and gaze. Scores for each measure represented the average across the 20 test trials.

Gap Selection

Number of gaps seen represented how many gaps passed from the start of a trial before crossing, including the gap crossed.

Gap size was the temporal size (in seconds) of the gap selected for crossing.

Movement Timing

Standing position was the distance (in meters) from the standing position of the participant to the center of the roadway.

Timing of entry was the time (in seconds) between the pedestrian and the rear of the lead car in the gap at the moment the pedestrian entered the path of the traffic.

Crossing time was the amount of time that the participant took to cross the road.

Time to spare was the time (in seconds) between the participant and the front of tail car at the time the participant cleared the path of the traffic.

Collisions occurred when time to spare was less than 0. These were extremely rare and were not analyzed further.

Gaze

Roadway focus distance was the distance down the roadway that the participant focused their gaze. Negative values represent the left side of the road and positive values represent the right side of the road.

Distance variability was the variability in distance down the roadway that the participant gazed.

Because participants in the Vive and Cave stood different distances from the roadway, gaze was measured in distance down the road instead of angle. The distance was calculated by multiplying the angle of the head by the standing distance. Figure 2 shows this calculation.

Figure 2: Road distance calculation. Road distance, r, was calculated by multiplying tangent of angle x by distance to the center of the road, d.

Participants

The participants were 32 undergraduate students. There were 16 participants (6 male, 10 female) in the CAVE group and 16 participants (10 male, 6 female) in the Vive group. Participants received course credit for their participation.

Results

Group differences were analyzed in one-way Analyses of Variance (ANOVAs) with condition (CAVE, Vive) as a between-subjects factor. As a complement to the ANOVAs, mixed-effects logistic regression analyses were conducted to evaluate gap size and group as predictors of gap choices. Table 1 shows means and standard deviations for all performance measures.

Gap Selection

Preformance Measures	CAVE(N=16)	Vive(N=16)
Standing Position (m)	1.85 (0.04)	1.94 (0.09)
Number of Gaps Seen	3.30 (1.22)	2.46 (0.80)
Mean Gap Size Taken (s)	4.63 (0.19)	4.44 (0.29)
Timing of Entry (s)	0.97 (0.14)	0.65 (0.14)
Road Crossing Time (s)	2.13 (0.21)	2.03 (0.19)
Time to Spare (s)	1.54 (0.29)	1.75 (0.37)
Roadway Focus Distance (m)	-12.19 (9.58)	-5.86 (4.33)
Distance Variability	12.53 (10.21)	6.73 (5.29)

Table 1: Means (and standard deviations) for all performance measures for participants in the CAVE and

Number of Gaps Seen Before Crossing: There was a significant effect of condition for number of gaps seen, F (1, 30) = 5.31, p = 0.03, η^2 = 0.15. Those in the CAVE group saw more gaps before crossing than did those in the Vive group.

Gap Size There was a significant effect of condition for gap size, F (1, 30) = 4.49, p = 0.04, η^2 = 0.13. Those in the CAVE condition took larger gaps on average than those in the Vive condition.

Mixed-effects logistic regression analyses indicated that participants in both conditions were more likely to cross through larger than smaller gaps, z = 10.16, p < 0.001, with 121.56 increased odds of accepting a gap with each .5 second increase in gap size (Fig. 2). In addition, participants in CAVE condition had higher gap acceptance thresholds than participants in the Vive condition, z = 2.36, p < 0.001. The Vive group had 4.78 increased odds of accepting a gap of any size compared to those in the CAVE condition, indicating that those in the CAVE condition were more conservative in their gap acceptance.

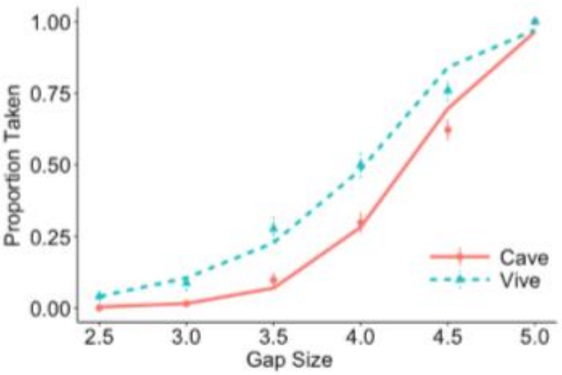

Figure 3: Logistic regression curves depicting the likelihood of accepting gaps of different sizes in the CAVE and Vive conditions.

In addition to threshold differences, condition marginally moderated gap size selection, z = -1.82, p = 0.07. Those in the Vive condition had 42.89 increased odds of gap acceptance with each .5 second increase in gap size, z = 7.93, p < 0.001, compared to those in the CAVE condition, who had 98.69 increased odds of accepting a gap with each .5 second increase in gap size, z = 11.29, p < 0.001. This indicates that participants in the CAVE condition were more discriminating in their gap choices, avoiding more of the smaller gaps and taking more of the larger gaps.

Movement Timing

Average standing position: There was a main effect of group for standing position, F (1, 30) = 8.36, p = 0.01, η^2 = 0.22, with those in the CAVE group standing closer to the virtual roadway before crossing than those in the Vive group.

Timing of entry: There was also a significant effect of condition for timing of entry, F (1, 30) = 38.28, p < 0.001, η^2 = 0.56. Those in the Vive group timed their entry into the gap more tightly than those in the CAVE group.

Road crossing time: There was no effect of condition for crossing time, F (1, 30) = 1.97, ns.

Time to spare: A marginal effect of condition emerged for time to spare, F (1, 30) = 3.24, p = 0.08, η^2 = 0.10. Participants in the CAVE condition had less time to spare than participants in the Vive condition.

Gaze

Roadway focus distance. Results revealed a main effect of condition on where participants focused their attention on the roadway, t (32) = -2.49, p = .02, with those in the CAVE condition (M = -12.19, SD = 9.58) looking further down the roadway on average compared to those in the Vive condition (M = -5.86, SD = 4.33).

Distance variability. A significant main effect of condition emerged, t (32) = 2.02, p = .05, with those in CAVE condition (M = 12.35, SD = 10.21) being more variable in where they looked on the roadway than those in the Vive condition (M = 6.73, SD = 5.29).

Discussion

This study compared how people cross traffic-filled roadways in virtual environments using an HTC Vive HMD vs. a large-screen CAVE to display images. The results show small, but statistically significant differences in performance between the Vive and CAVE conditions. Participants in the Vive condition were less conservative and less discriminating in their gap choices and timed their entry into the roadway more tightly than those in the CAVE condition. Participants in the Vive condition had somewhat more time to spare when exiting the roadway, despite choosing smaller gaps and standing further from the roadway. This is likely due to the fact that they timed their entry into the gap more tightly and crossed the road slightly faster. Participants in the CAVE condition focused their gaze further down the roadway than participants in the Vive condition. The CAVE has a wider field of view than the Vive, leading us to hypothesize that those in the Vive would look further down the roadway and have more variability. Further experiments will help determine why this hypothesis was false.

While the results do not reveal the underlying causes for these differences in performance, there are several factors that may have contributed to both more risky and more skilled behavior in the HTC Vive. One possibility is that the lack of a body representation may make the environment less threatening and led to a willingness to take greater risks, in this case a lower gap acceptance threshold. Another contributing factor may be that the HMD provided superior cues for time to arrival judgments, possibly related to greater pixel density per visual angle, leading to tighter timing of motion initiation. Additional research is needed to better understand how display modality influences performance in complex, whole-body perception and action tasks.

The results contrast with those of Grechkin et al. [Grechkin et al. 2014] who found better timing in a train boarding task for participants in a CAVE as compared to an NVIS nVisor ST HMD. Differences in the task and the technology make it difficult to determine what may have caused this difference in results between their study and ours. However, the NVIS HMD is considerably more unwieldy than the HTC Vive which may have contributed to the timing differences.

Overall, both environments were effective in presenting a virtual road crossing task. We had no reports of simulator sickness in either the Vive or CAVE and no drop outs. This contrasts with an experiment conducted by Deb et al. [Deb et al. 2017] that also used an HTC Vive to study road crossing. They report an 11.5% rate of withdrawal due to simulator sickness.

Overall, the results provide promising support for using consumer level hardware to conduct research on pedestrian road-crossing behavior. These HMDs also offer promise for developing widely deliverable interventions related to pedestrian road crossing behavior, as well as an inexpensive platform for research on pedestrian behavior that is portable, easy to use, and simple to maintain.

Acknowledgements

This research was supported by National Science Foundation awards BCS-1251694 and CNS-1305131, and by the US Department of Transportation, Research and Innovative Technology Administration, Prime DFDA No. 20.701, Award No. DTRT13-G-UTC53.

Thank you to Pooya Rahimian, Elizabeth E. O'Neal, Jodie M. Plumert and Joseph K. Kearney

References

Sarah E Banducci, Nathan Ward, John G Gaspar, Kurt R Schab, James A Crowell, Henry Kaczmarski, and Arthur F Kramer. 2016. The effects of cell phone and text message conversations on simulated street crossing. *Human Factors* 58, 2 (2016), 150–162.

Doug A Bowman, Ameya Datey, Young Sam Ryu, Umer Farooq, and Omar Vasnaik. 2002. Empirical comparison of human behavior and performance with different display devices for virtual environments. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 46. SAGE Publications Sage CA: Los Angeles, CA, 2134–2138.

Shuchisnigdha Deb, Daniel W Carruth, Richard Sween, Lesley Strawderman, and Teena M Garrison. 2017. Efficacy of virtual reality in pedestrian safety research. *Applied Ergonomics* 13 (2017).

Aurelie Dommés, Tristan Le Lay, Fabrice Vienne, Nguyen-Thong Dang, Alexandra Perrot Beaudoin, and Manh Cuong Do. 2015. Towards an explanation of age-related difficulties in crossing a two-way street. *Accident Analysis & Prevention* 85, 8 (2015), 229–238.

Ilja Feldstein, André Dietrich, Sasha Milinkovic, and Klaus Bengler. 2016. A Pedestrian Simulator for Urban Crossing Scenarios. *IFAC-PapersOnLine* 49, 12 (2016), 239–244.

Timofey Y Grechkin, Jodie M Plumert, and Joseph K Kearney. 2014. Dynamic affordances in embodied interactive systems: The role of display and mode of locomotion. *IEEE Transactions on Visualization and Computer Graphics* 20, 16 (2014), 596–605.

Yuanyuan Jiang, Elizabeth O'Neal, Pooya Rahimian, Junghum Paul Yon, Jodie M Plumert, and Joseph K Kearney. 2016. Action coordination with agents: crossing roads with a computer-generated character in a virtual environment. In *Proceedings of the ACM Symposium on Applied Perception*. ACM, 57–64.

Kara Macek. 2017. Pedestrian Fatalities Projected to Surge 11% in 2016. *Governors Highway Safety Association* Washington, DC, 1 (2017).

Barbara A Morrongiello and Michael Corbett. 2015. Using a virtual environment to study child pedestrian behaviours: a comparison of parents' expectations and children's street crossing behaviour. *Injury Prevention* 21, 10 (2015), 291–295.

Barbara A Morrongiello, Michael Corbett, Melissa Milanovic, and Jonathan Beer. 2015a. Using a virtual environment to examine how children cross streets: Advancing our understanding of how injury risk arises. *Journal of Pediatric Psychology* 41, 11 (2015), 265–275.

Barbara A Morrongiello, Michael Corbett, Melissa Milanovic, Sarah Pyne, and Robin Vierich. 2015b. Innovations in using virtual reality to study how children cross streets in traffic: Evidence for evasive action skills. *Injury Prevention* 21, 9 (2015), 266–270.

Mark B Neider, John G Gaspar, Jason S McCarley, James A Crowell, Henry Kaczmarski, and Arthur F Kramer. 2011. Walking and talking: dual-task effects on street crossing behavior in older adults. *Psychology and Aging* 26, 4 (2011), 260.

Mark B Neider, Jason S McCarley, James A Crowell, Henry Kaczmarski, and Arthur F Kramer. 2010. Pedestrians, vehicles, and cell phones. *Accident Analysis & Prevention* 42, 3 (2010), 589–594.

Elizabeth E O'Neal, Yuanyuan Jiang, Lucas J Franzen, Pooya Rahimian, Junghum Paul Yon, Joseph K Kearney, and Jodie M Plumert. 2017. Changes in Perception-Action Tuning Over Long Time Scales: How Children and Adults Perceive and Act on Dynamic Affordances When Crossing Roads. *Journal of Experimental Psychology: Human Perception and Performance* (2017). https://doi.org/10.1037/xhp0000378

Pooya Rahimian, Elizabeth E O'Neal, Junghum Paul Yon, Luke Franzen, Yuanyuan Jiang, Jodie M Plumert, and Joseph K Kearney. 2016. Using a virtual environment to study the impact of sending traffic alerts to texting pedestrians. In *Virtual Reality (VR), 2016 IEEE*. IEEE, 141–149.

Anthony Steed and Chris Parker. 2005. Evaluating effectiveness of interaction techniques across immersive virtual environmental systems. *Presence: Teleoperators and Virtual Environments* 14, 15 (2005), 511–527.