Latency Detection and Illusion in a Head-Worn Virtual Environment

Collin Roth*
High Fidelity Virtual Environments Lab
Computer & Information Science
University of Mississippi

Ethan Luckett†
High Fidelity Virtual Environments Lab
Computer & Information Science
University of Mississippi

J. Adam Jones‡
High Fidelity Virtual Environments Lab
Computer & Information Science
University of Mississippi

Abstract

Through the history of virtual environments research there has been significant interest in understanding how latency in a system affects a user's experience. Though latency cannot be avoided, previous work has observed that there may be ranges within which small latencies are not discernible. However, the majority of the work examining latency detection thresholds was conducted using hardware and software that are no longer commonly used in contemporary research. In the current study, we examine whether similar latency tolerances exist for modern, off-the-shelf systems. We also look at the effect of increasing and decreasing latency on such tolerances. This revealed evidence of a "latency illusion" that presents in cases of decreasing latency resulting in subjects perceiving less latency than is actually present in the environment.

Index Terms: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual Reality; I.4.8 [Scene Analysis]: Depth Cues; H.5.1 [Information Systems]: Multimedia Information Systems—Artificial, Augmented, and Virtual Realities; H.1.2 [Information Systems]: User/Machine Systems—Human Factors

1 Introduction

As virtual environments (VEs) continue to become more accessible to the general public, the need to understand the interplay between users' perception and the technology's inherent limitations grows more important. One particular limitation of these technologies is system latency, sometimes referred to as the motion-to-photons delay. This delay is ever-present across essentially all platforms with varying degrees of severity. This is not a new problem and has been examined in the past [1, 3, 7]. In the present work, we replicate the procedures used by Adelstein et al. [1] but using a modern, commodity system.

Earlier works found that users were quite sensitive to variable latency and were able to detect differences as low as 16ms [1,3]. Our primary goal is to determine if similar sensitivities will be observed when using a modern, commodity system, specifically the HTC Vive. The Vive is reported to have a base latency of 22ms with the screen refreshing every 11ms [6, 9, 10]. As such, the granularity of our testing will be limited by these parameters. In the study that follows, we examine whether or not users of this system can detect latency presented at five levels ranging from 0 to 4 screen refreshes.

2 Related Work

Adelstein et al. [1] performed the study from which the procedures for the current work were adapted. The minimum perceptible level of latency was measured while an observer turned from side to side. They hypothesized that people use "image slip" to tell if an environment is delayed. Image slip is defined as "motion of the VE caused by system time lags" [1]. They also hypothesized that going from constant movement to randomized incremental movements would disrupt a subject's ability to discriminate latency in the VE using image slip. It was found that there was a minor reduction in the ability to detect latency in the VE when head movements were presented in randomized intervals, meaning that image slip was likely used to discriminate delay. They also found subjects were able to detect latency starting around 17ms.

In Ellis et al. [3], they studied subjects' abilities to discriminate differences in latency using lateral hand movements. To accomplish this, subjects were placed into a VE where they were asked to move a real-world ball. There was a virtual representation of the ball present within the VE as well. Subjects moved the ball at arms-length, in a 48 degree arc, in beat with a metronome set to 72 beats per minute (BPM). There were three base levels of latency: 27ms (system time latency), 93ms and 197ms. The latency would vary in increments of 16.7ms, up to 100.2ms higher than the base level. They found that subjects were able to consistently differentiate levels of latency as small as 16.7ms.

Meehan et al. [7] examined the effects of latency in a VE, specifically in regard to simulator sickness. Subjects were placed into two groups, one for the low latency environment (50ms), then another for the high latency environment (90ms). The environment used in this study was comprised of a "non-threatening" training room to help get the subjects acclimated to the environment, and a "threatening" room that contained a two-story-deep pit in which subjects were standing on a slim board peeking over the edge. Participants were instructed to drop bean bags onto targets in the pit room to elicit physiological responses such as an increased heart rate. This study found that the low latency environment resulted in subjects having a greater sense of presence and higher changes in physiological responses, and that there was no difference in the levels of simulator sickness caused by the environments.

Newman and Greeley [8] examined helmet-mounted displays in order to create a design guide for engineers of future displays. In their study, they looked at issues such as image stabilization, field of view (FOV) vs. resolution tradeoffs, design considerations of the helmets themselves, and display latency. They state that latency negatively affects image definition and recommended that latency be no more than 20ms.

In Frank et al. [5], a driving simulator was used to examine visual and physical latency in connection to simulator sickness. The simulator had no history of inducing motion sickness. It was modified to be able to introduce delay into the visual presentation and physical movement of the environment independently. Three levels of visual delay and three levels of movement delay were present in the environment. They found that visual delay was more damaging to a control performance than motion delay.

Jones et al. [6] measured the native latency of the HTC Vive, the display used in the following experiment. They used a method involving a pendulum-like apparatus to detect differences between real and virtual movement. They found the latency of the Vive to be 21.7±3ms: In Niehorster et al. [9], they looked at the latency of the Vive as well. This was done by recording changes in the image displayed in the head-mounted display using a high-speed camera as it was moved in the physical world. The recording was analyzed by counting the total number of screen refreshes that elapsed before

*e-mail: croth@go.olemiss.edu
†e-mail: eelucket@go.olemiss.edu
‡e-mail: jadamj@acm.org

978-1-7281-6532-5/20/$31.00 ©2020 IEEE
DOI 10.1109/VRW50115.2020.00046

Figure 1: A screenshot of the VE used for this experiment.

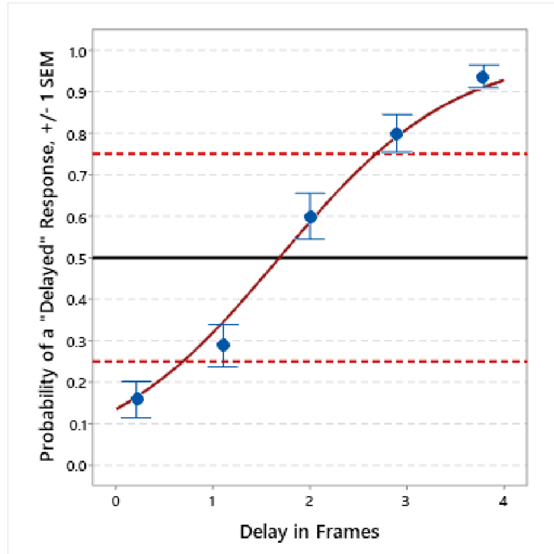

Figure 2: Latency detection responses fit with a psychometric curve. Error bars represent ± 1 standard error of the mean.

physical movements were reflected in the display's graphics. The latency of the Vive was found to be 22ms on average.

3 Experiment

3.1 Participants

Ten volunteers participated in this study, all of whom were experienced VE users (1 female, 9 males). The procedures used in this experiment were reviewed and approved by the University of Mississippi Institutional Review Board. All participants were made aware of the purpose of the experiment prior to participating. We focused primarily on experienced VE users as Adelstein et al. [1] found that users with more experience were most sensitive to the introduction of latency to a VE.

3.2 Apparatus

3.2.1 Hardware & Software

The study was developed on a desktop PC running Windows 10, with an NVIDIA GeForce GTX 1080 GPU, and an Intel i7-6800k processor. Unity3D was used to create and run the environment. The HTC Vive was used in conjunction with the SteamVR plug-in, allowing for easy integration with Unity3D. All scripts were written in C#.

3.2.2 Artificially Induced Latency

Generally, latency is seen as something to be avoided. As such, the SteamVR plug-in does not natively support features that can be used to deliberately increase latency. We worked around this limitation by utilizing both the SteamVR camera rig and a custom camera rig, referred to as the *delayed cameras*. The 3D transformations reported by the SteamVR camera were queued on a frame-per-frame basis. The environment was culled in such a way as to only allow the scene's geometry to be visible to the delayed cameras. The delayed cameras' view was rendered to a texture and then mapped to a plane positionally locked and only visible to the SteamVR cameras. The oldest transformation in the queue was applied to the delayed cameras thereby allowing the SteamVR cameras to see a delayed view of the scene. This procedure added a single frame of delay to the rendering process in addition to size of the delay queue. As such, when displaying a scene with no latency, it was rendered directly to the SteamVR cameras. The delay method was only applied to delays of one frame or greater. Since the display refresh rate of the HTC Vive is 90Hz, we were limited to testing latencies in increments of 11ms.

This method was appealing as it allowed delays to be introduced by queuing at most four relatively small 4x4 transformation matrices while only utilizing two rendered textures, one for each eye. This resulted in a low over-head method of artificially introducing latency, but delays were only applied to the user's view of the environment and not to the environment itself. However, this was not an issue for the current experiment as participants viewed an unchanging environment. A perhaps more intuitive method would be to directly queue rendered frames and render only the oldest. We did test this approach, but it resulted in unpredictable delays as a result of significantly reduced system performance. This was due to rendering, queuing, and retrieving high resolution renderings of the scene on a frame-per-frame basis. It is possible, however, that such an approach would be more feasible on higher-end hardware than that which was currently available at the time of the experiment.

3.3 Procedures

Participants were instructed to view a relatively simplistic VE that consisted of a sparsely furnished model of the lab space where the experiment took place. A screenshot of the environment can be seen in Fig. 1. Participants sat in a swiveling desk chair and were asked to rotate themselves side-to-side in pace with the tone of a digital metronome set to 50 BPM.

Five levels of latency were present in the environment ranging from zero to four frames delayed, or 0-44ms of delay. Each level of latency was repeated eight times resulting in a total of 40 trials per participant. Using a computer mouse, a participant would indicate if they believed the environment was either artificially delayed or not using the left and right mouse buttons respectively. They were not given a time limit to make a choice but were instructed to respond as quickly and accurately as possible.

Once a button on the mouse was pressed, the view in the HMD would briefly black out for 500ms before beginning the next trial. This was done to prevent participants from noticing any abrupt changes in the latency of the environment. A restricted random shuffle, meaning the same latency could not occur in back-to-back trials, was used to determine the order for the 40 trials. This was done to ensure that the same level of latency couldn't occur in succession.

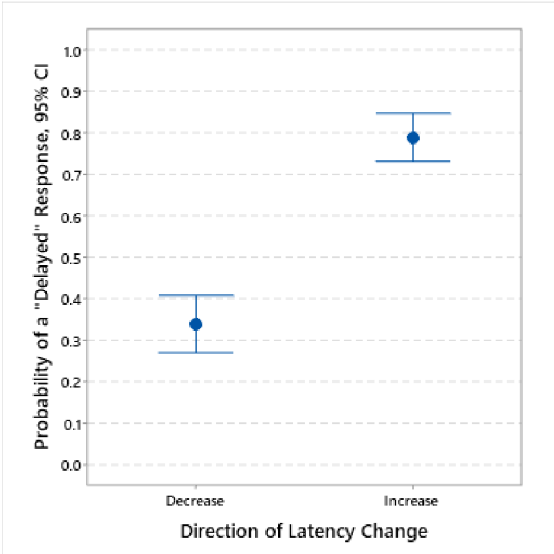

Figure 3: Latency detection responses for trials where the delay of the previous trial was either increasing or decreasing. Error bars represent 95% Confidence Intervals.

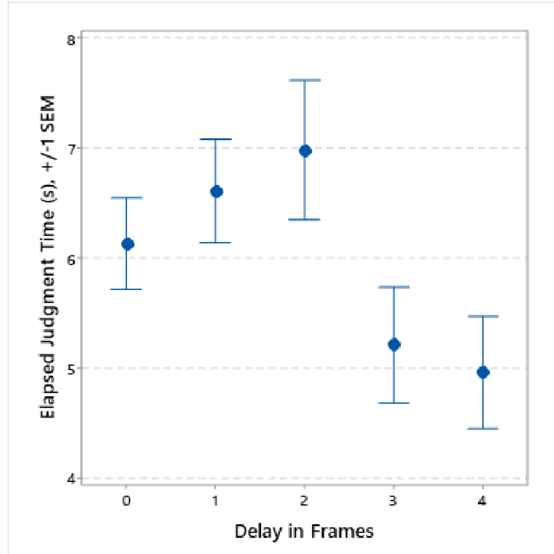

Figure 4: Mean response times for each level of delay. Error bars represent ± 1 Standard error of the mean.

4 Results

Fig. 2 shows the probability of a response of "delayed" for each tested latency. These data have been fit with a psychometric curve. We found that the point of subjective equality (PSE) where subjects could not reliably detect a delay between their physical and visual movements was 18.81ms or 1.71 frames. This is consistent with the observations of Adelstein et al. [1] who found that subjects were able to detect latencies in the range of 15 to 20ms. The just noticeable difference (JND) is defined by the amount of latency required to elicit a $\pm 25\%$ change in probability about the PSE. The $+25\%$ JND threshold was 7.51ms or 0.68 frames while the -25% threshold was 29.5ms or 2.68 frames. These thresholds are very close to ± 1 frame about the PSE. Probability of a "delayed" response significantly differed across levels of delay (ANOVA: $F_{4,36} = 29.711, p < 0.001$). Post-hoc analysis using Tukey pairwise comparisons indicated that delays of 0 and 1 frame as well as 3 and 4 frames did not significantly differ from each other while all other comparisons did differ ($p < 0.05$).

Additionally, we found that there was a relationship between the participants' ability to detect the presence of latency in a given trial and the direction of the change in latency seen relative to the immediately previous trial. As can be seen in Fig. 3, participants were found to be less sensitive to decreasing latency and more sensitive to increasing latency (ANOVA: $F_{1,9} = 84.474, p < 0.001$). It is important to keep in mind that the task subjects were performing required them to report if the visual scene was delayed relative to their physical movement and not whether latencies were increasing or decreasing.

Fig. 4 shows the mean response time (RT) for each level of delay. RT is frequently used as a means of gauging the relative difficulty of tasks. Typically, increases in RT are associated with increases in cognitive load or task difficulty [2, 4]. When examining RT, we found that it significantly differed as a function of latency (ANOVA: $F_{4,36} = 5.701, p = 0.001$). Subjects took longer, on average, to detect latency for delays of 2 frames or less as compared to delays of 3 and 4 frames. Post-hoc analysis using Tukey pairwise comparisons indicated that subjects were significantly faster in response to delays of 4 frames compared to 2 frames ($p < 0.05$). As compared to the zero frames of delay condition, subjects took approximately 10% longer to respond to the 1 and 2 frames of delay while they were 13% and 21% faster to respond in the 3 and 4 frames of delay conditions, respectively. As one would expect, the larger the latency, the easier it was for subjects to detect.

5 Discussion

We found that subjects were able to detect changes in latency within a range similar to that reported by Adelstein et al. [1] and Ellis et al. [3]. We found that as delays approached the PSE, the relative difficulty of detecting latency increased and then decreased beyond the PSE, implying that delays greater than 22ms were easier to detect. Though this is not terribly surprising, it does indicate that similar latency related work conducted on legacy hardware may still be applicable to modern hardware. However, we found that a user's detection threshold may be influenced by whether or not latency increased or decreased relative to the previous trial. Interestingly, participants were significantly less likely to detect delays between physical and visual movement if the change in latency between trials was decreasing. Effectively, this appears to result in an illusory condition where a subject perceives less latency than is actually present in the system. This latency illusion could potentially be an advantage for systems with high latency as it implies that users may become less sensitive to the presence of delays if first presented with artificially large delays before being stepped downward to the base, system latency. It is important to note that further research is required to better quantify these effects.

There are several limitations to this study, however. First, the study was not designed specifically to examine whether or not the direction of the change in latency affected a subject's ability to detect delays between physical and visual motion. Since conditions were presented based on a random shuffle, we were not guaranteed the same number of increasing (194) and decreasing (186) trials. However, the size of the difference between these two conditions was quite large despite this asymmetry. None-the-less, a study designed to look more closely at this particular factor is necessary. Second, the number of subjects in this study was somewhat low, and all subjects had prior experience with VEs. It could be argued that these results may not be generalizable given the subjects' prior experiences. However, Adelstein et al. [1] found that experienced users were more capable of detecting delays. As such, it may be the case that utilizing experienced subjects provides a more conservative lower bound for system latency.

REFERENCES

[1] B. D. Adelstein, T. G. Lee, and S. R. Ellis. Head tracking latency in virtual environments: psychophysics and a model. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 47, pp. 2083–2087. SAGE Publications Sage CA: Los Angeles, CA, 2003.

[2] E. Cooper-Martin. Measures of cognitive effort. *Marketing Letters*, 5(1):43–56, 1994.

[3] S. R. Ellis, M. J. Young, B. D. Adelstein, and S. M. Ehrlich. Discrimination of changes of latency during voluntary hand movement of virtual objects. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 43, pp. 1182–1186. SAGE Publications Sage CA: Los Angeles, CA, 1999.

[4] H. Finney and J. A. Jones. Asymmetric effects of the ebbinghaus illusion in a virtual environment. In *IEEE Virtual Reality, 2020. Proceedings.*, p. in press. IEEE, 2020.

[5] L. H. Frank, J. G. Casali, and W. W. Wierwille. Effects of visual display and motion system delays on operator performance and uneasiness in a driving simulator. *Human factors*, 30(2):201–217, 1988.

[6] J. A. Jones, E. Luckett, T. Key, and N. Newsome. Latency measurement in head-mounted virtual environments. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 1000–1001. IEEE, 2019.

[7] M. Meehan, S. Razzaque, M. C. Whitton, and F. P. Brooks. Effect of latency on presence in stressful virtual environments. In *IEEE Virtual Reality, 2003. Proceedings.*, pp. 141–148. IEEE, 2003.

[8] R. L. Newman and K. W. Greeley. Helmet-mounted display design guide. Technical report, CREW SYSTEMS CONSULTANTS SAN MARCOS TX, 1997.

[9] D. C. Niehorster, L. Li, and M. Lappe. The Accuracy and Precision of Position and Orientation Tracking in the HTC Vive Virtual Reality System for Scientific Research. *Iperception*, 8(3):2041669517708205, 2017.

[10] A. Vlachos. Advanced VR Rendering, Mar. 2015. https://www.gdcvault.com/play/1021 771/Advanced-VR, Accessed: Sep 2018.