

# Psychophysical Analysis of Delay Detection in a VR Avatar's Standing-up Motion

M. Olimov<sup>1</sup> and Y. Goto<sup>1</sup> and S. Okamoto<sup>1</sup>

<sup>1</sup> Department of Computer Science, Tokyo Metropolitan University, Japan

---

## Abstract

*Discrepancies between an avatar's movements in virtual space and participants' movements in the real world can degrade the quality of the virtual reality (VR) experience. One prominent form of such a discrepancy is delay. Many previous studies have investigated the acceptable delay between head-tracking and landscape rendering, or the delay of the seen user's hand movements. However, the minimum detectable delay during full-body movements, particularly those involving significant changes in viewpoint, has not yet been fully investigated. In this study, the detection threshold for delays between participants' real-world movements, where the head and viewpoint positions move substantially, and corresponding avatar movements in virtual space was investigated. In the experiment, participants wearing VR goggles performed stand-up motions. Corresponding stand-up motion of the avatar in the VR space involved the delay of up to 300 ms. Participants looked at the avatar's movements through the mirror placed in front of him/her in the VR space. The detection thresholds of five individuals were investigated using psychophysical method of constant stimuli. In the experiment, the participants answered whether the avatar's movements delayed or did not delay comparing with their own movements. The mean detection threshold, at which the participant reports the presence of delay for 50% of all the time, was found to be 129.70 ms, with a 95% confidence interval of 31.59 ms. These findings provide some insights for designers of VR applications.*

## CCS Concepts

• **Human-centered computing** → **Virtual reality**;

---

## 1. Introduction

In immersive virtual reality (VR) environments, an avatar, which is a computer graphics representation of the operator, acts in synchronization with the operator. VR goggles experience a delay of at least a few tens of milliseconds before updating displayed images after detecting the operator's head motion [TM21]. This delay depends on factors such as communication bandwidth, computer performance, and the complexity of VR environments. Substantial delays deteriorate bodily awareness toward the avatars and the computer graphics representation of the operator's body parts [IA15, SQH09, WSH\*16], whereas small delays are not recognizable. Hence, earlier researchers investigated the impact of delay on the perceptual quality of VR spaces and human behaviors within these spaces [WSH\*16, MAEH04, SPQ15, KZD\*23, CMG19]. For example, Waltemate et al. [WSH\*16] reported that operators noticed the delayed avatar motion nearly 50% of the time when the delay was 210 ms. Furthermore, the sense of agency and body ownership toward the avatar substantially declined with delays over 300 ms. In terms of task performance, Caserman et al. [CMG19] reported that reaching tasks were interfered with by delays in an immersive virtual reality setting. Outside of immersive VR settings,

sensory feedback delays have been also extensively studied by researchers (e.g. [OKST09, HK17]).

Earlier researchers pursued the detection of delay between head motion or eye motion and displayed scenery images [MAEH04, APLK17], as well as the motion delay of hand images [SHO05, SQH09, FFG\*01, Leu03]; however, the avatar's whole body motion has been rarely targeted. In contrast, Waltemate et al. [WSH\*16] performed a study similar to ours. They investigated the absolute threshold of avatar's motion delay when assessors mainly exercised their upper bodies through experiments using a cave automatic virtual environment. However, the results of the study by Waltemate et al. [WSH\*16] could have been biased by their specialized protocols.

In their experiment, the operators watched two superimposed avatars. One was used for the instruction of exercise, and the other was the operator's own avatar. These two avatars were superimposed, and their motions could be misaligned. Hence, the avatar motions could appear delayed relative to the instructive avatar, which might have influenced the detection of the avatar's motion delay. They reported that the delay was not detected 12.5% of the time even with a delay of 350 ms. As described in Section 3, in



**Figure 1:** Settings in the actual space. Participant wearing VR goggles sat on a chair.



**Figure 2:** VR images seen by participant. Participant looked at its avatar through a mirror positioned in front of him/her. Height of the avatar was matched that of the participant.

our setting, the delay of 300 ms was nearly perfectly detected by participants.

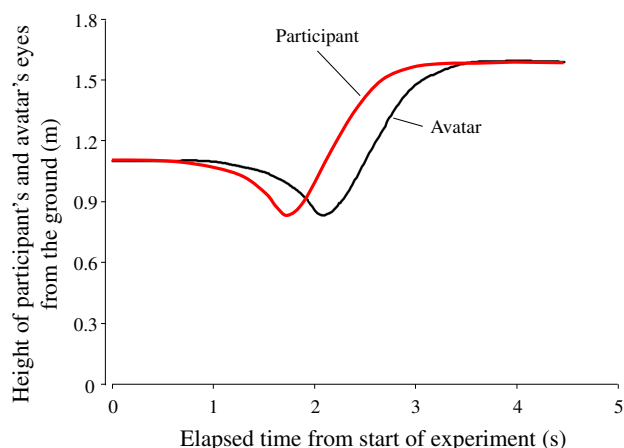
We target the whole-body motion of an avatar in an immersive VR environment using VR goggles. For this purpose, standing-up motions are employed in this study. During these motions, the operator's viewpoint, specifically the head position, changes drastically. This setting differs from that of Waltemate et al. [WSH\*16], where primarily the hands or arms were exercised.

The objective of this study is to investigate the detectability of an avatar's delay when head motion is involved, using a standard protocol that minimizes any biases on the detection threshold of delay. The findings will contribute to the body of knowledge in the literature of perceptual sciences in VR.

## 2. Method

### 2.1. Experimental Settings

The experimental setup is shown in Fig. 1. The participant repeatedly stood up and sat down on an office chair with a height of 50 cm while wearing virtual reality goggles (Meta Quest 2, Meta



**Figure 3:** Eyes position of participant and avatar at the delay of 0.30 s. Participant stood up from the chair.

Platforms, Inc., USA). Throughout the experiment, the participant maintained a feet width equivalent to their shoulder width.

The virtual environment was developed using Unity 2022.3.6f1 (Unity Software Inc., USA). An avatar matching the participant's height was placed within the virtual space. The avatar's posture was synchronized with the participant's movements, such that when the participant stood up from a chair, the avatar mirrored this action. The transition from a seated to a fully standing position was animated [mix24], with the animation speed adjusted based on the participant's head position. The avatar's movements were visible to the participant through a mirror positioned 1 meter in front, a technique commonly used to enhance bodily awareness in virtual environments [IK21].

### 2.2. Delayed Avatar's Motion

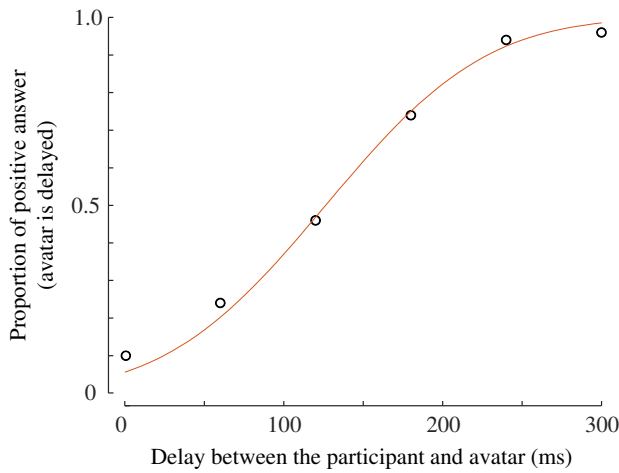
The avatar's standing-up animation was synchronized with the participant's head position. When no delay was introduced, the head positions of the avatar in the mirror and the participant matched. It should be noted that even under this minimal delay condition, a delay of a few tens of milliseconds can still occur [TM21].

In the delayed condition, the participant observed the delayed motion of the avatar through the VR goggles. This delay was constant throughout each trial. To achieve the delayed motion, the participant's head positions were buffered for a specified period, and the avatar's animation was based on these past head positions.

Figure 3 shows an example of the two head positions with a 300-ms delay. Six levels of nominal delays were tested: 0, 60, 120, 180, 240, and 300 ms. These delay values were determined through a preliminary study involving the authors and their colleagues. A delay of 300 ms was almost perfectly perceptible.

### 2.3. Experimental Procedures

We employed the psychophysical method of constant stimuli to determine the detection thresholds. The participant wore virtual real-



**Figure 4:** Normal cumulative distribution function fitted to the mean detection proportions across all the participants.

ity goggles and sat on a chair with their feet positioned at shoulder width. Upon receiving a cue from the experimenter, the participant stood up. After standing, they judged whether the avatar in the virtual space moved with or without delay compared to their own movement in reality. The participant then sat back down.

The delay magnitude was one of the six levels described in Section 2.2. The six delay levels were presented in a randomized block design, with the six levels appearing in a randomized order within each block. Each block consisted of six trials, and a total of 10 blocks were conducted. Thus, the procedure was repeated 60 times in total, with 10 repetitions for each delay level.

#### 2.4. Participants

Five university students in their 20s participated in the study after providing written informed consent. They were not informed of the study's objective beforehand.

#### 2.5. Ethical Statement

The protocol of this study was approved by Institutional Review Board, Tokyo Metropolitan University (R6-009).

#### 2.6. Data Analysis

For each participant, the proportions of detections at different delay values were approximated by a normal cumulative distribution function:

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) dt \quad (1)$$

where  $x$  denotes the delay in milliseconds. The parameters  $\mu$  and  $\sigma$  were estimated by the likelihood estimation method [WH01].

When  $x = \mu$ ,  $p(x) = 0.5$ , indicating that the delay is recognized by chance. This value of  $x$  is defined as the 50% detection threshold. When  $x = \mu + \sigma$ ,  $p(x) = 0.84$ , and this  $x$  value is considered as the 84% threshold in this study.

**Table 1:** Detection (50%) thresholds and 84% thresholds of participants in milliseconds.

Participant	50% threshold $p = 0.50$	84% threshold $p = 0.84$
A	161.43	259.28
B	102.17	205.56
C	172.31	222.19
D	90.17	174.52
E	122.42	159.96
Mean	129.70	204.30
95% Conf. int.	31.59	34.50

### 3. Results

Table 1 lists the 50% and 84% detection thresholds for each of the five participants (Participants A–E). The average value of the 50% thresholds was 129.70 ms, with a 95% confidence interval of  $\pm 31.59$  ms. For the 84% thresholds, the average value was 204.30 ms, with a 95% confidence interval of  $\pm 34.50$  ms.

Figure 4 shows the mean positive detection proportions at different delays, along with the approximated probability distribution function. It should be noted that these proportion values represent the means across participants, while the thresholds were determined individually for each participant.

### 4. Discussion

Here, we compare our results with those reported by Waltemate et al. [WSH\*16], where the detection thresholds of avatar's delay were investigated when whole-body motions were involved. They determined the 50% and 84% detection thresholds to be approximately 210 ms and 310 ms, respectively. These values are substantially greater than those identified in this study: 129.7 and 204.3 ms. It should be noted that the actual delay values are greater than these by a few tens of milliseconds due to the intrinsic delay of the system [TM21]. Nonetheless, there exist substantial distinctions between these two studies, and we speculate that they are attributed to the type of whole-body motions and VR systems used in the studies.

In the study by Waltemate et al. [WSH\*16], the motions tested in the experiment were largely limited to the upper body and arms. In contrast, the stand-up motion used in this study was associated with the large vertical head's position change. The perspectives also change with such head's movement. According to Mania et al. [MAEH04], the scenery images' delay toward the head's smooth yaw motion could be detected even when the delay was as small as approximately 30 ms. It should be noted that in their study, assessors could refer to a reference stimulus of minimum delay: 12.5 ms, and under such conditions, the assessors would have been more sensitive. Adelstein et al. [ALE03] reported the thresholds of 40–60 ms by using a similar protocol. Without a reference stimulus, the detection threshold could range 180–320 ms [AHJ\*01]. These values are comparable to or even greater than those in this study. When the head position or perspective is stationary, delayed hand motion is detected at approximately 200 ms (50% threshold) [SHO05] or

160–200 ms (86% threshold) [Leu03]. Hence, the difference of detection thresholds could originate from the difference in head motion during the tasks between our study and that of Waltemate et al.

In discussing the detection threshold of delay, we may need to concern judgment bias. In our experiment, even with the minimum delay, the delay was reported for 10% of all the time. Similarly, in the experiment by Waltemate et al. [WSH\*16], this value was 15% for the minimum delay of 45 ms. For such cases, the detection threshold cannot be discussed merely on the basis of the proportions of operators' reports [YIS07]; however, indices explicitly considering response biases are necessary. In future, we introduce such indices to analyze the detection thresholds.

## 5. Conclusion

We investigated the minimally noticeable delay between the motions of a VR goggle wearer and their VR avatar. Previous studies have rarely focused on motions involving large head position or perspective changes; however, this study employed standing-up motions, which are accompanied by substantial head movements. The analysis of data acquired using the psychophysical method of constant stimuli identified the 50% detection threshold as 123.70 ms. It should be noted that this value does not account for system latency, meaning the actual threshold is likely higher. This threshold is smaller than those reported in previous studies, which primarily tested upper-body motions with minimal head movement. These findings provide important insights for designers of VR applications.

For more reliable conclusions, a larger sample size is needed. Additionally, the minimum system latency in our experimental setup, which is estimated to be at least 18 ms [TM21], remains unknown. Determining this latency is crucial for accurately discussing detection thresholds.

## References

- [AHJ\*01] ALLISON R., HARRIS L., JENKIN M., JASIOBEDZKA U., ZACHER J.: Tolerance of temporal delay in virtual environments. In *Proceedings IEEE Virtual Reality 2001* (2001), pp. 247–254. doi:10.1109/VR.2001.913793. 3
- [ALE03] ADELSTEIN B. D., LEE T. G., ELLIS S. R.: Head tracking latency in virtual environments: Psychophysics and a model. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 47, 20 (2003), 2083–2087. doi:10.1177/154193120304702001. 3
- [APLK17] ALBERT R., PATNEY A., LUEBKE D., KIM J.: Latency requirements for foveated rendering in virtual reality. *ACM Transactions on Applied Perception* 14, 4 (2017), 1–13. doi:10.1145/3127589. 1
- [CMG19] CASERMAN P., MARTINUSSEN M., GÖBEL S.: Effects of end-to-end latency on user experience and performance in immersive virtual reality applications. In *Entertainment Computing and Serious Games*. Springer International Publishing, 2019, pp. 57–69. doi:10.1007/978-3-030-34644-7\_5. 1
- [FFG\*01] FRANCK N., FARRER C., GEORGIEFF N., MARIE-CARDINE M., DALÉRY J., D'AMATO T., JEANNEROD M.: Defective recognition of one's own actions in patients with schizophrenia. *American Journal of Psychiatry* 158, 3 (2001), 454–459. doi:10.1176/appi.ajp.158.3.454. 1
- [HK17] HACHISU T., KAJIMOTO H.: Vibration feedback latency affects material perception during rod tapping interactions. *IEEE Transactions on Haptics* 10, 2 (2017), 288–295. doi:10.1109/TOH.2016.2628900. 1
- [IA15] IMAIZUMI S., ASAI T.: Dissociation of agency and body ownership following visuomotor temporal recalibration. *Frontiers in Integrative Neuroscience* 9 (2015). doi:10.3389/fnint.2015.00035. 1
- [IK21] INOUE Y., KITAZAKI M.: Virtual mirror and beyond: The psychological basis for avatar embodiment via a mirror. *Journal of Robotics and Mechatronics* 33, 5 (2021), 1004–1012. doi:10.20965/jrm.2021.p1004. 2
- [KZD\*23] KHAN T., ZHU T. S., DOWNES T., CHENG L., KASS N. M., ANDREWS E. G., BIEHL J. T.: Understanding effects of visual feedback delay in ar on fine motor surgical tasks. *IEEE Transactions on Visualization and Computer Graphics* 29, 11 (2023), 4697–4707. doi:10.1109/TVCG.2023.3320214. 1
- [Leu03] LEUBE D.: The neural correlates of perceiving one's own movements. *NeuroImage* 20, 4 (2003), 2084–2090. doi:10.1016/j.neuroimage.2003.07.033. 1, 4
- [MAEH04] MANIA K., ADELSTEIN B. D., ELLIS S. R., HILL M. I.: Perceptual sensitivity to head tracking latency in virtual environments with varying degrees of scene complexity. In *Proceedings of the 1st Symposium on Applied perception in graphics and visualization* (2004), APGV04, ACM. doi:10.1145/1012551.1012559. 1, 3
- [mix24] Mixamo, 2024. URL: <https://www.mixamo.com/>. 2
- [OKST09] OKAMOTO S., KONYO M., SAGA S., TADOKORO S.: Detectability and perceptual consequences of delayed feedback in a vibrotactile texture display. *IEEE Transactions on Haptics* 2, 2 (2009), 73–84. doi:10.1109/TOH.2009.17. 1
- [SHO05] SHIMADA S., HIRAKI K., ODA I.: The parietal role in the sense of self-ownership with temporal discrepancy between visual and proprioceptive feedbacks. *NeuroImage* 24, 4 (2005), 1225–1232. doi:10.1016/j.neuroimage.2004.10.039. 1, 3
- [SPQ15] SAMARAWEEERA G., PERDOMO A., QUARLES J.: Applying latency to half of a self-avatar's body to change real walking patterns. In *IEEE Virtual Reality* (2015), pp. 89–96. doi:10.1109/VR.2015.7223329. 1
- [SQH09] SHIMADA S., QI Y., HIRAKI K.: Detection of visual feedback delay in active and passive self-body movements. *Experimental Brain Research* 201, 2 (2009), 359–364. doi:10.1007/s00221-009-2028-6. 1
- [TM21] TACHIBANA R., MATSUMIYA K.: Accuracy and precision of visual and auditory stimulus presentation in virtual reality in python 2 and 3 environments for human behavior research. *Behavior Research Methods* 54, 2 (2021), 729–751. doi:10.3758/s13428-021-01663-w. 1, 2, 3, 4
- [WH01] WICHMANN F. A., HILL N. J.: The psychometric function: I. fitting, sampling, and goodness of fit. *Perception amp; Psychophysics* 63, 8 (2001), 1293–1313. doi:10.3758/BF03194544. 3
- [WSH\*16] WALTEIMATE T., SENNA I., HÜLSMANN F., ROHDE M., KOPP S., ERNST M., BOTSCH M.: The impact of latency on perceptual judgments and motor performance in closed-loop interaction in virtual reality. In *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology* (2016), VRST'16, ACM. doi:10.1145/2993369.2993381. 1, 2, 3, 4
- [YIS07] YAIRI S., IWAYA Y., SUZUKI Y.: Estimation of detection threshold of system latency of virtual auditory display. *Applied Acoustics* 68 (2007), 851–863. doi:10.1016/j.apacoust.2006.12.005. 4