Understanding the Effects of Physical Interaction Types and Distance on the Performance of Target Selection Task

Category: Research

Abstract

One of the advantages of mixed reality is that users can walk around and interact with virtual elements in their surrounding physical environment, allowing seamless interactions between virtual and physical worlds. However, most studies in mixed reality focused on indirect object manipulation with hand-based interactions, which could be unfamiliar for most of the users, whereas body-based interaction such as walking towards a target for direct manipulation is what users have been trained for a lifetime on a daily basis. In this study, we conducted a user study with 20 participants to investigate the effect of distance between a target and a user on the performance of the target selection task, and if the impact varies depending on the interaction type (hand vs. body). As a result, we found that while one-meter difference in distances does not always result in a longer task completion time in hand-based interaction, the time increased linearly in every meter for body-based interaction. Moreover, while hand-based interaction was preferred over body-based interaction in general due to physical demands, some preferred body-based interaction when the target is a meter away since manipulating a target that is too close with hand-based interaction was difficult.

Index Terms: Human-centered computing—Interaction design—Empirical studies in interaction design—; Human-centered computing—Human computer interaction (HCI)—Interaction paradigm—Mixed / augmented reality

1 Introduction

Mixed Reality (MR) is a combination of the physical and digital worlds, which can unlock the limit of interaction between human, computer, and environment [14]. It enables users to seamlessly interact in both worlds with a high level of immersion. However, most prior studies mainly focus on indirect object manipulation using hand-based interaction [6, 7, 10], which requires additional training, while body-based interaction with direct manipulation such as walking towards an object to have a close look could be more familiar for users since users have been trained for this for a lifetime on a daily basis. Since one of the most common disadvantages of head-mounted displays (HMD) is its lack of supporting an accurate sense of depth, which leads to perceptual bias and impacting the task performance [3,22], body-based interaction could be more ideal than hand-based interaction, especially for distant targets.

To provide users with the precise sensation of depth, multiple depth cues such as *pictorial depth cues*, *kinetic depth cues*, *physiological depth cues*, and *binocular disparity* have been proposed [8]. However, little has been studied how target depth impacts the users' performance. Thus, to identify the type of interaction that is more applicable to the target selection task, which is the most common task performed before every object manipulation tasks, varying in distance, we compared two types of interaction: hand-based interaction, the most widely used interaction in an MR environment, and body-based interaction as it is relatively easy to perform than hand-based interaction.

To answer this question, we conducted a user study with 20 participants, where we requested the participants to perform a target selection task on a virtual object with hand-based interaction, by using a laser pointer emitted from the tip of the participants' index finger, and with body-based interaction, by walking straight to the object until they hit the object which varied in distance. As a result, we found that the distance of an object affected the participants' task completion time in both hand-, and body-based interactions. While one-meter difference in distance does not always result in a longer task completion time in hand-based interaction, we found that the task completion time with body-based interaction was significantly longer for every meter increased in distance. Moreover, subjective feedback revealed that more participants preferred body-based interaction than hand-based interaction when the target is only a meter away since manipulating a target that is too close is difficult with hand-based interaction.

The contributions of our work include (1) the empirical analysis on the performance differences between hand-, and body-based interaction for target selection task, (2) the confirmation of the effects of distance on the performance of target selection task, and (3) design implications for hand-, and body-based interaction with a 3D virtual object in mixed-reality environments.

2 Related Work

Various studies have focused on understanding the hand-based interactions with extended reality (XR) devices [1, 10, 11, 15, 17, 20]. Shim *et al.* [20], for instance, proposed a system using AR HMD that provides users with multiple gestural and tangible interactions to interact with the AR objects. Similarly, Ababsa *et al.* [1] introduced a new device called Leap Motion Controller (LMC) to understand enriched gestural interactions that are used in Hololens, such as AirTap and bloom gestures.

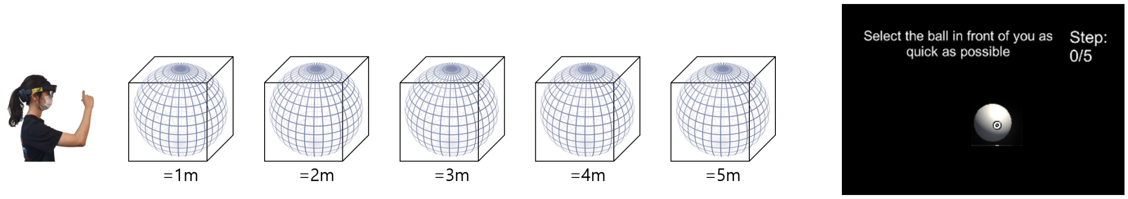
Other studies also have been conducted to understand interactions with free hands, which do not require additional handheld controllers or hand-worn devices for users to interact with virtual objects in a virtual environment [9,16]. For example, Kyto *et al.* [9] implemented a multimodal input system, where hand gesture input is combined with other types of input such as head movement and eye gaze, to enable users to make a precise selection in an AR environment.

Meanwhile, multiple research has been conducted to understand interaction techniques using full-body with XR devices [2,12,13,18]. For instance, Crowell *et al.* [2] conducted a study to investigate if physically collaborating in a virtual environment can enhance the socialization ability of children with autism and discovered that such interaction provides them the opportunity to initiate social actions. Also, Roupe *et al.* [18] conducted a study comparing computer mouse movement and VR body-based interaction, and discovered that participants perceived body-based interaction as less demanding than computer mouse interaction. Similarly, Lindgren *et al.* [13] also identified that full-body interaction in an MR environment enhances the users' understanding and satisfaction.

Inspired by these prior works that examined the benefits of body-based interaction, we compared body-based interaction with hand-based interaction for target selection task in terms of task performance and participants' subjective responses.

3 Method

To understand if the impact of the target distance on the performance of target selection task differs depending on the interaction type (hand vs. body), we conducted a user study with 20 participants.



(a) Target distance from 1m to 5m.

(b) Presented target.

Figure 1: Experimental settings: (a) the distance parameters for target selection task, and (b) a screenshot example of a target selection task for the first target within a round at 2 m distance. All targets were identical, but distant targets seemed smaller than close targets.

3.1 Conditions

In this study, our main focus is to assess the performance difference between hand-, and body-based interaction techniques, and the impact of the target's distance if any.

> **Interaction Type (Hand vs. Body)**: The target can be selected with a virtual pointer under hand-based interaction condition at a static position whereas a participant needs to move towards the target and hit it with their physical body under body-based interaction condition.
>
> **Distance (1m vs. 2m vs. 3m vs. 4m vs. 5m)**: The distance between the target and a participant.

3.2 Participants

We recruited 20 participants (9 male, 11 female) by distributing an online Google Form to a university and by word of mouth. The participants' average age was 23.0, ranged between 19 and 27 and no one had difficulties in dexterity or mobility. While none of the participants had experience with mixed reality devices including Microsofts' Hololens, 11 participants had prior experience with VR and AR (two of them had extensive experience). All but two participants were right-handed.

3.3 Apparatus

As for the HMD device, we used Microsoft Hololens 2. It was connected with a computer through a wireless network, and the computer had an Intel i7-9750H (2.60GHz) CPU with 16GB of RAM and a Geforce RTX 2060 graphics card. As for the software, we used Unity 2017.4.20f1 with MRTK (Mixed Reality Toolkit). The software program displayed a switch button, which is a "start" button and is selected with the *AirTap* gesture[1], for each target, whose radius was set to 1m, in order to avoid the problem of selecting the target by mistake. The virtual targets appeared one at a time in various distances but identical in color, size, and height. The selection of the target differed depending on the interaction type condition. While the same *AirTap* gesture was used for selecting each target under the hand-based interaction condition, we monitored the collision event between the participant and the target and considered it as a selection interaction under the body-based interaction condition as the participant walked through the target with the HMD on. To do so, we calibrated the height of the target for each participant to make sure the HMD directly collided with the target. During the task, the software measured the task completion time from the press of the "start" button to the target selection.

3.4 Data and Analysis

Through the user study, we measured the task completion time for each target (a total of 1000 data; 20 participants $\times$ 2 interaction types $\times$ 5 distance type $\times$ 5 repetitions). However, for the analysis, we excluded a total of 7 data points (6 from hand-based interaction) that are beyond 3SD away from the mean. Perceived workload was also collected using the unweighted NASA TLX questionnaire in a 7-point Likert scale [5]. Additionally, we collected the interview responses after each task and after all the tasks were finished.

[1]It is a mid-air pinch gesture, which is a default selection gesture of Hololens 2.

3.5 Procedure

After signing the consent form, we explained the overall process of the user study. Next, we asked each participant to perform the target selection task with both hand-, and body-based interactions where the presentation order of the interaction type was balanced. For each interaction type, a practice session was provided to the participants before the actual task to get familiar with the task. Then the participants were asked to perform five rounds of tasks where each round had five targets with different distances between 1m to 5m with the interval of a meter, presented in random order. Participants were allowed to take a break between tasks if needed. After completing each condition, we had a semi-structured interview with the participants to receive subjective feedback and collected the responses based on the unweighted NASA TLX questionnaire. We also conducted a short interview at the end of the user study about their preference and their overall perception of the target distances.

4 Findings

4.1 Task Completion Time

We conducted a two-way ANOVA with repeated measures and found a significant interaction effect between *Interaction Type* (2-level: hand vs. body) and *Distance* (5-level: 1, 2, 3, 4 and 5m); $F_{(1,4)}$ = 37.57, $p < .001$. As shown in Fig. 2, the task completion time was almost the same across two interaction types when the target was a meter away but it was significantly longer with body-based interaction than with hand-based interaction beyond a meter ($p <$ .001 for all). Moreover, we found that the performance difference between targets with 1m difference is smaller with hand-based interaction than body-based interaction. To be specific, with hand-based interaction, we found that the average task completion time was significantly different across different distances with one-way ANOVA ($F_{(4)}$ = 9.62, $p < 0.01$). The post hoc pairwise comparisons showed that the difference was statistically significant between all pairs except for 1m & 2m, and 4m & 5m; see Table 1. This suggests that a longer time is needed for selecting a target that is farther away in general, but a meter difference for relatively close-ranged targets ($<=$ 2m) and far-ranged targets (4m $<=$ 5m) has no impact on the task performance.

Similarly, one-way ANOVA revealed that there is a significant difference across all the different distances for body-based interaction ($F_{(4)}$ = 233.99, $p < .001$). However, unlike the hand-based interaction, the performance difference between every pair was found to be significant with post hoc analysis; $p < .001$ for all pairs. This indicates that the participants took a significantly longer time to select a target as the distance becomes longer for every meter.

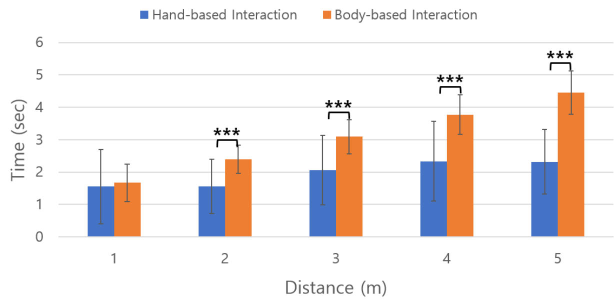


Figure 2: The average task completion time (*sec.*) per distance for both types of interactions. The error bars indicate standard deviation. The performance of hand-based interaction was significantly better than body-based interaction for all distances except for 1*m*. The '***' mark indicates p value under .001.

4.2 Regression Analysis on Task Completion Time

As a secondary analysis, we performed a linear regression on task completion time. The R^2 coefficient value was 0.73 for hand-based interaction with the slope of 0.576, and 0.95 for body-based interaction with the slope of 0.965; note that y-intercept was set to 0 for each case. We also ran a logarithmic regression analysis. However, the coefficient values were much lower: 0.09 for hand-based interaction and 0.72 for body-based interaction.

4.3 Perceived Workload

In addition to the task performance, we also collected the participants' perceived workload with an unweighted NASA TLX questionnaire on a 7-point Likert scale. As shown in Fig. 3, the average mental demand of body-based interaction was lower than hand-based interaction, while the rest was the opposite. However, Wilcoxon signed rank tests showed that the differences of individual perceived workloads between the two interaction types were not significant except for physical demand ($Z = 2.94$, $p = .003$).

4.4 Subjective Feedback Regarding Target Distance

Reflecting the perceived physical demand, 15 participants answered that farther targets are more difficult than closer targets because it is physically tiring to walk long distances for the body-based interaction. Yet, the trend was similar for hand-based interaction as well. More than half of the participants (N = 13 out of 20) commented that it was more difficult as the position of the target becomes farther with hand-based interaction. To be specific, one participant mentioned,

> *"It took a lot of time to adjust my pointer [the laser pointer] to the ball [the target] with my finger when it was far from me."*

Interestingly, three other participants felt that it was most difficult when the target was the closest at a 1-meter distance, which was not an issue raised with the body-based interaction. One of them mentioned that it was difficult to point the target because it was too close.

4.5 Preference

When asked the overall preference of interaction type, 80% of the participants (N = 16 out of 20) answered that they preferred hand-based interaction because it is more convenient to select the target without having to walk to the target. The rest who preferred body-based interaction over hand-based interaction (N = 4) commented that having to precisely point the target was too difficult. Meanwhile, we also asked the participants if their preference differed depending on the target distance and the answers were as in Table 2. Half of the participants preferred either hand-, or body-based interaction regardless of the target distance (N = 8 and N = 2, respectively). However, the rest ten participants' preferences differed depending on the distance. For instance, for targets that are far away, nine out of ten participants preferred hand-based interaction while only one participant preferred body-based interaction. For targets that are near, six of them preferred body-based interaction, three had no preference between the two interaction types, and only one participant preferred hand-based interaction.

m	1	2	3	4	5
1	-	.946	<.001	<.001	<.001
2		-	<.001	<.001	<.001
3			-	.002	.005
4				-	.442
5					-

Table 1: Pairwise post hoc analysis results with p values of hand-based interaction. Note that p values for every pair in body-based interaction were all under .001 and thus not presented in this table.

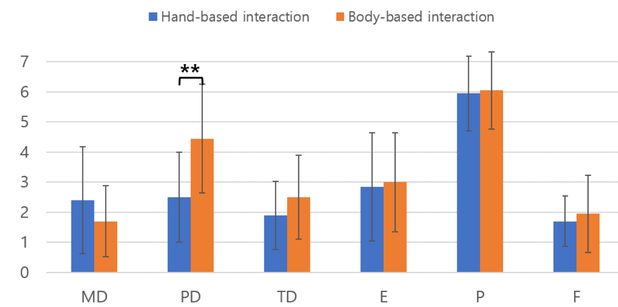


Figure 3: Unweighted NASA TLX questionnaire responses on average: Mental Demand (MD), Physical Demand (PD), Temporal Demand (TD), Effort (E), Performance (P), Frustration (F). The '**' mark indicates p value under .01.

5 Discussion

5.1 The Effect of Distance on Target Selection Performance

We identified how target distance affects participants' performance in the selection task for both hand-, and body-based interactions. While the task performance of body-based interaction significantly increased by every meter, the increment was weak for hand-based interaction. Also, while there was no performance difference between the two interaction types when the target is a meter-away from the participant, hand-based interaction significantly outperformed body-based interaction for all other cases. This suggests that having to physically walk to the target is more time-consuming than having to point a small-sized target that is farther than a meter. It would be interesting to study if the trend still exists for targets that are closer

	Head	Body	None
Near Distance	1	6	3
Far Distance	9	1	0
Any Distance	8	2	0

Table 2: The number of responses for interaction type preference depending on the target's distance.

than a meter or father than $5m$. Moreover, to understand if or when body-based performance outperforms hand-based interaction, further investigation is needed for distances with smaller intervals than a meter. We can also confirm the prior study [19] which presents that the distance of the target affected the participants' task completion time.

5.2 Linear Relation Between Distance and Task Performance

While we expected to see a logarithmic relationship between the distance and the hand-based interaction considering Fitts's Law [4], our secondary analysis revealed a higher R^2 coefficient value for the linear than the logarithmic regression, confirming a prior finding, which presents that changes in depth have a linear impact on the task completion time for target pointing task [21]. The most plausible explanation would be the difference in how we set the target distance. To be specific, since Fitts's Law is modeled based on the distance to the target center (d) and the width of the target (W) on a 2D display, the depth of the target should be converted to fit into a 2D space so that as the target distance becomes twice longer in z-axis (the depth), the target width should decrease by half ($0.5W$) instead of increasing the target distance to $2d$. If we assume that participants always start at the same location, then the distance to the target center remains the same. In addition, since Fitts's Law is a base-2 logarithmic function, our resulting data would show a linear relationship. It suggests that pixel-level distance is more important for estimating the task completion time for the target selection task than the actual distance in depth.

5.3 Perceived Workload and Preference

Although hand-based interaction outperformed body-based interaction for most cases, we found that the distance also affected the participants' preference in interaction type, especially for the target at a meter distance. As some of the participants commented, pointing a close-up target that is relatively large can be as difficult as pointing a smaller target that is far away. Moreover, sudden changes in depth may have forced participants to adjust their visual focus. Thus, we plan to study the impact of the target size in terms of the area of the pixels on a display along with the depth of the target for hand-based interaction.

6 Conclusion

In this paper, we investigated the impact of target distance on users' performance and found that the distance had a different impact on hand-, and body-based interaction although the task completion time tends to increase in both interaction types. Also, our findings show that hand-based interaction outperformed body-based interaction for targets that are farther than a meter even if the size of the displayed target becomes smaller to be selected with hand-based interaction. Yet, some participants preferred body-based interaction for a one-meter target since it was difficult to select a target with hand-based interaction when the target is too close. While we focused on target selection task when comparing hand-, and body-based interaction, we plan to investigate the performance difference between the two interaction types for various object manipulation tasks such as translation and rotation of objects.

References

[1] F. Ababsa, J. He, and J.-R. Chardonnet. Combining hololens and leap-motion for free hand-based 3d interaction in mr environments. In *International Conference on Augmented Reality, Virtual Reality and Computer Graphics*, pp. 315–327. Springer, 2020.

[2] C. Crowell, B. Sayis, J. P. Benitez, and N. Pares. Mixed reality, full-body interactive experience to encourage social initiation for autism: Comparison with a control nondigital intervention. *Cyberpsychology, Behavior, and Social Networking*, 23(1):5–9, 2020.

[3] D. Drascic and P. Milgram. Perceptual issues in augmented reality. In *Stereoscopic displays and virtual reality systems III*, vol. 2653, pp. 123–134. International Society for Optics and Photonics, 1996.

[4] P. M. Fitts. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology*, 47(6):381, 1954.

[5] S. G. Hart and L. E. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, vol. 52, pp. 139–183. Elsevier, 1988.

[6] H. J. Kang, J.-h. Shin, and K. Ponto. A comparative analysis of 3d user interaction: how to move virtual objects in mixed reality. In *2020 IEEE conference on virtual reality and 3D user interfaces (VR)*, pp. 275–284. IEEE, 2020.

[7] S. Kim, A. Jing, H. Park, G. A. Lee, W. Huang, and M. Billinghurst. Hand-in-air (hia) and hand-on-target (hot) style gesture cues for mixed reality collaboration. *IEEE Access*, 8:224145–224161, 2020.

[8] E. Kruijff, J. E. Swan, and S. Feiner. Perceptual issues in augmented reality revisited. In *2010 IEEE International Symposium on Mixed and Augmented Reality*, pp. 3–12. IEEE, 2010.

[9] M. Kytö, B. Ens, T. Piumsomboon, G. A. Lee, and M. Billinghurst. Pinpointing: Precise head-and eye-based target selection for augmented reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–14, 2018.

[10] J. Y. Lee, G. W. Rhee, and D. W. Seo. Hand gesture-based tangible interactions for manipulating virtual objects in a mixed reality environment. *The International Journal of Advanced Manufacturing Technology*, 51(9-12):1069–1082, 2010.

[11] H. Liang, J. Yuan, D. Thalmann, and N. M. Thalmann. Ar in hand: Egocentric palm pose tracking and gesture recognition for augmented reality applications. In *Proceedings of the 23rd ACM international conference on Multimedia*, pp. 743–744, 2015.

[12] R. Lindgren, M. Tscholl, and J. M. Moshell. Meteor: Developing physics concepts through body-based interaction with a mixed reality simulation. In *Physics Education Research Conference-PERC*, vol. 13, pp. 217–220, 2013.

[13] R. Lindgren, M. Tscholl, S. Wang, and E. Johnson. Enhancing learning and engagement through embodied interaction within a mixed reality simulation. *Computers & Education*, 95:174–187, 2016.

[14] Microsoft. What is mixed reality?

[15] S. Mohatta, R. Perla, G. Gupta, E. Hassan, and R. Hebbalaguppe. Robust hand gestural interaction for smartphone based ar/vr applications. In *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 330–335. IEEE, 2017.

[16] T. Piumsomboon, A. Dey, B. Ens, G. Lee, and M. Billinghurst. The effects of sharing awareness cues in collaborative mixed reality. *Frontiers in Robotics and AI*, 6:5, 2019.

[17] I. U. Rehman, S. Ullah, and M. Raees. Two hand gesture based 3d navigation in virtual environments. *IJIMAI*, 5(4):128–140, 2019.

[18] M. Roupé, P. Bosch-Sijtsema, and M. Johansson. Interactive navigation interface for virtual reality using the human body. *Computers, Environment and Urban Systems*, 43:42–50, 2014.

[19] M. Serrano, D. Hildebrandt, S. Subramanian, and P. Irani. Identifying suitable projection parameters and display configurations for mobile true-3d displays. In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services*, pp. 135–143, 2014.

[20] J. Shim, Y. Yang, N. Kang, J. Seo, and T.-D. Han. Gesture-based interactive augmented reality content authoring system using hmd. *Virtual Reality*, 20(1):57–69, 2016.

[21] E. Triantafyllidis and Z. Li. The challenges in modeling human performance in 3d space with fitts' law. *arXiv preprint arXiv:2101.00260*, 2021.

[22] D. Van Krevelen and R. Poelman. A survey of augmented reality technologies, applications and limitations. *International journal of virtual reality*, 9(2):1–20, 2010.