2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)

Development and evaluation of a test setup to investigate distance differences in immersive virtual environments

Stephan Fremerey, Muhammad Sami Suleman, Abdul Haq Azeem Paracha and Alexander Raake
Department of Audio Visual Technology, Technische Universitt Ilmenau, Germany
Email: [stephan.fremerey, muhammad-sami.suleman, abdul-haq-azeem.paracha, alexander.raake]@tu-ilmenau.de

***Abstract*—Nowadays, with recent advances in virtual reality technology, it is easily possible to integrate real objects into virtual environments by creating an exact virtual replication and enabling interaction with them by mapping the obtained tracking data of the real to the virtual objects. The primary goal of our study is to develop a system to investigate distance differences for near-field interaction in immersive virtual environments. In this context, the term distance difference refers to the shift between a real object and the respective replication of the real object in the virtual environment of the same size. This could occur for a number of reasons e.g. due to errors in motion tracking or mistakes in designing the virtual environment. Our virtual environment is developed using the Unity3D game engine, while the immersive contents were displayed on an HTC Vive Pro head-mounted display. The virtual room shown to the user includes a replication of the real testing lab environment, while one of the two real objects is tracked and mirrored to the virtual world using an HTC Vive Tracker. Both objects are present in the real as well as in the virtual world. To find perceivable distance differences in the near-field, the actual task in the subjective test was to pick up one object and place it into another object. The position of the static object in the virtual world is shifted by values between 0 and 4 cm, while the position of the real object is kept constant. The system is evaluated by conducting a subjective proof-of-concept test with 18 test subjects. The distance difference is evaluated by the subjects through estimating perceived confusion on a modified 5-point absolute category rating scale. The study provides quantitative insights into allowable real-world vs. virtual-world mismatch boundaries for near-field interactions, with a threshold value of around 1 cm.**

***Index Terms*—virtual reality, near field interaction, distance difference, offset, immersive virtual environments**

I. Introduction

Nowadays, both in the research community as well as in industry, a variety of immersive virtual environments (IVE) are developed. The level of immersion and presence when using Virtual Reality (VR) content could be increased by tracking objects present in the real world in all 6 Degrees of Freedom (6DoF) (rotation and translation), and presenting them in the virtual world, enabling real-life like interaction with these objects. Due to tracking errors, which could occur in Head-Mounted Display (HMD) systems, or mistakes in the design process of such environments, an offset between the real and the virtual environment could lead to problems when trying to interact with these objects. Hence it is important to identify the offset values between the real and the virtual objects that are still tolerable by subjects. Accordingly, in this study we want to find the offset threshold at which the subject's grade of confusion increases. Further, we want to investigate whether VR users are still able to fulfill simple tasks such as a pick-and-place actions at certain offset levels. Previously, most of the studies for distance perception in IVEs were conducted at distances of more than 5 m (cf. [1, 3]). Little research has been conducted to examine near-field ego-centric distance estimation (0 m to maximum arms reach) in IVEs.

Hence the following research questions are defined: From which offset between a tracked real and a virtual object does the perceived grade of confusion increase? To what extent does this offset level have an influence on the task completion time and performance? Could some kind of learning effect be observed? To answer these questions we developed a novel test framework and conducted a test with 18 subjects using an HTC Vive Pro HMD to display the IVE. Subjects are able to interact with one tracked object, while this object needs to be picked and placed into another, fixed object. For the fixed object, we introduced several offsets between the real and the virtual object. Further, we measured the grade of confusion and task completion time at several offset levels.

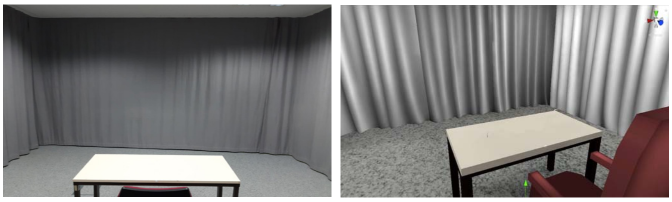
II. Related Work

A few studies on the quality of depth perception in IVEs and connected distance compression have been reported in the literature [10, 1]. Armbrüster et al. [1] explored the performance of distance perception in IVEs between 40 cm and 500 cm in three different IVEs of different depth information under two conditions, with either one or multiple target objects. For some parts of the subjective tests, a metric aid was used as orientation for the subjects while distance estimation was done verbally. Regardless of the condition, subjects underestimated distances in the virtual world.

Other studies such as [4, 7] also show that people underestimate distances in the near field. Additionally, Interrante et al. [3] suggest that distance perception appears to not be significantly compressed in the virtual environment if it represents an exact virtual replica of the subject's actual real environment.

Niehorster et al. [5] investigated the tracking performance of Valve's Lighthouse 1.0 tracking system and an HTC Vive

978-1-7281-5965-2/20/$31.00 ©2020 IEEE



(a) Real test room (b) Virtual test room

Fig. 1: Real and virtual test room

HMD. They found that tracking dropouts can lead to incorrect pitch and roll measurements, changed height measurements and a tilt of the reference plane.

In addition, Suznjevic et al. [9] evaluated the pick-and-place task performance for two different HMDs: HTC Vive and Oculus Rift. A subjective test was conducted with 13 subjects, in which subjects had to pick and place several objects from one location to another. These objects were only present in the virtual and not in the real environment, while interaction with these objects was done using hand-held controllers. Due to the more robust tracking system, HTC Vive was found to outperform Oculus Rift on both subjective and objective measures.

Furthermore, Ogawa et al. [6] conducted a study in which the offset between a virtually generated hand and the person's hand was altered. They concluded that if the offset between the real and the virtual hand is too high, users recognize this and could become aware of e.g. a remapping of the hand. However, a maximum offset between the real and the virtual hand, e.g. in cm, is not provided.

To summarize, it becomes clear that there is a lack of research investigating the impact of offsets between real and virtual objects on human perception and action. This shows that there is a need for a detailed investigation of near-field egocentric distance differences experienced by the user.

III. Experimental Setup

A. IVE

To minimize the influence of the IVE on the distance perception of the user and to increase the level of immersion, we created an exact virtual representation of our test room with the same size as suggested in the study by Interrante et al. [3] using Unity3D 2019.2.0f1. In Figures 1a and 1b, the real and the virtual test room is shown, respectively. We used Unity3D's built-in shaders and lighting feature to make the curtains and the floor look more realistic. To increase the reproducibility of the paper, the developed Unity application will be published along with the paper.[1]

B. Technical Setup

To display the IVE, an HTC Vive Pro VR system using two Lighthouse 1.0 base stations for tracking was connected to a high-end VR-PC equipped with an Intel Core i7-6700 processor, a fast Samsung M.2 SSD and a NVIDIA GeForce RTX 2080 Ti graphics card. To ensure a highly reliable tracking, both base stations were put at the end of each table placed 4 m apart. The VR system and the IVE were calibrated before every test run as proposed in [5]. In the test scenario shown in Figure 2, two objects were both physically and virtually located on the table, with one of the objects (open cylinder, 8.5 cm diameter, 3 cm height) being fixed to the table and the other one (cylinder, 8 cm diameter, 12 cm height) being fully interactable. To track the 6DoF movements of the interactable cylinder in real time, an HTC Vive Tracker[2] was mounted onto it. During the tests, subjects were seated on a fixed chair.

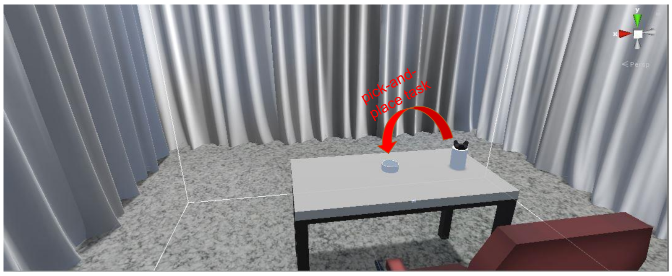


Fig. 2: Test setup for pick-and-place task

C. Test Design and Test Procedure

To evaluate at which offset between the physical and virtual objects subjects start getting confused, we conducted a test with 18 subjects. Each test run took about 20 minutes per person. Based on the successful design of this proof-of-concept test we plan for a higher number of subjects in future, more refined versions of the paradigm. At first, subjects were screened for accurate vision using Snellen charts (20/20). To ensure that there is no general problem with the visual comfort of the IVE, following a modified approach of Singla et al. [8], before and after every test subjects had to complete three short questions on their level of simulator sickness. A 5-point Absolute Category Rating (ACR) scale [2] was used, while subjects were asked to give ratings on their level of dizziness, headache and eyestrain. To familiarize the subjects with the IVE, subjects were told to watch a two minute long video on a virtual screen shown in the IVE before the test.

To answer our research questions, we developed a special test scenario already described in Section III-B. Test subjects were always given the following task: "Please place the cylinder inside the other open cylinder as fast as possible." This task is also illustrated in Figure 2. We measured the time needed to perform the given task. Between two successive tasks, we changed the position of the fixed virtual cylinder while keeping the position of the other, interactable cylinder constant. We used the following offset levels, which were increasing over time: 0 cm (ideal condition), 1 cm, 2 cm, 3 cm and 4 cm. At this point it has to be noted that between trials the translational directions of the offset levels were randomized, whether the order of offsets presented to subjects was not.

[1] https://github.com/Telecommunication-Telemedia-Assessment/distance_differences_ives

[2] https://www.vive.com/de/vive-tracker

After the subjects fulfilled the task, they were asked about their perceived grade of confusion while performing the task using a modified version of the 5-point ACR scale [2] (no confusion - slightly confused - confused - highly confused - extremely confused).

IV. Results and Discussion

A total of 18 subjects, 14 male and 4 female, all recruited from the university (mean age = 26.31, median age = 26), participated in the test. Mean dizziness, headache and eyestrain values before and after the test were very low (1.1 before and 1.2 after the test). Also taking into account the very low amount of motion presented in the IVE, we conclude that there was no general problem with the perceived visual comfort of the IVE.

In Figure 3, the mean confusion and task completion time are shown for the specific offset values between the real world and the IVE, together with the 95 % confidence intervals (CI).

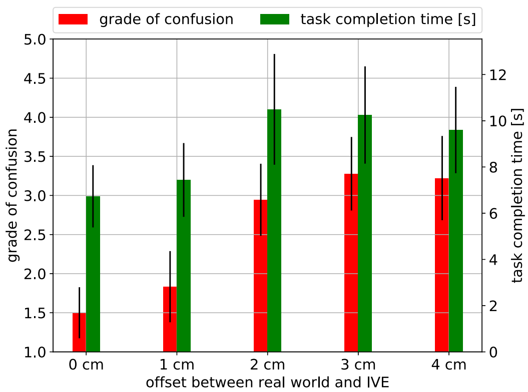


Fig. 3: Offset value vs. mean grade of confusion and task completion time

In general, every subject was able to fulfill the task at every offset level between the real and the virtual open cylinder. The condition without an offset between the physical and the virtual world created the lowest amount of confusion of 1.5, while subjects needed, on average, 6.7 s time to put the one cylinder into the other one. With an offset of 1 cm, subjects still were not confused (mean 1.8) and the mean time needed to fulfill the task is still acceptable (7.4 s).

Further, it can be concluded that starting from an offset of 2 cm, subjects experienced a higher degree of confusion (mean 2.9). This is supported by the fact that starting from a shift of 2 cm, the average time needed by the subjects to accomplish the task is increasing (7.4 s for 1 cm vs. 10.5 s for 2 cm). At higher offsets (3 cm and 4 cm), the grade of confusion (mean 3.3 and 3.2) and the time needed to fulfill the task (mean 10.3 and 9.6) is not considerably increasing any further, even though for an offset of 3 cm and 4 cm the grade of confusion was rated slightly higher. At an offset of 4 cm the confusion level and time needed to fulfill the task was even found to be slightly lower compared to an offset of 3 cm. This could be related to the learning effect of the subjects, enabling them to solve the task easier after several repetitions of the task. In general, it is interesting to note that there is an increase of confusion even in spite of the not randomized offset levels per trial and the connected learning effect. This could be related due to the fact that the translational direction of the offset between the real and the virtual open cylinder was randomized. Hence, another hypothesis is that between two conditions, subjects expected the object to shift into the same translational direction than before, but as the translational directions were randomized, with larger offset values subjects overall apply a different strategy to solve the task. Further, this could mean that subjects need a while to learn that offsets are integrated in the IVE, while this learning effect evolves over time.

From the size of the CIs shown in Figure 3, it is visible that the time needed to fulfill the task and the grade of confusion is quite different per subject, especially for confusion levels higher than 1 cm. This could be related to the different individual experiences of subjects with VR technologies and the relatively low number of subjects.

V. Conclusions and Future Work

This paper presents a novel test method in the area of IVE research and respective publicly available IVE to measure the effect of distance offsets between physical and virtual objects in VR. From the proof-of-concept test conducted with 18 subjects, it can be concluded that the IVE can in principle be utilized to measure the effect of these differences. It was observed that starting from an offset of 2 cm between the virtual and the real object, the confusion grade of the subjects was increasing. Also, people needed more time to fulfill the task at a 2 cm offset compared to a 1 cm offset. However, for larger offsets than 2 cm, both the grade of confusion and the time needed to fulfill the task did not increase. We observe that for offset values larger than 2 cm, the time needed to accomplish the task is even slightly decreasing, which could be related to the learning effect of the subjects resulting from the used increasing offset levels. However, due to the randomization of the translational direction of the offset, this learning effect seems to take some time to have a larger positive effect on the task completion time.

The results of this study are especially important for researchers and designers creating IVEs. If a real object should be tracked and displayed in VR with the same size, attention should be paid to make sure that the shift between the virtual and the real object is equal to or below 1 cm.

In the future, we will conduct a test with more subjects to investigate offset values between 0 and 2 cm with more gradations, while also introducing larger offsets than 4 cm. Here, the order of offsets presented to subjects will be randomized between trials, repeating the same conditions twice in two different test phases to explicitly study the learning effect. This ideally will result in a Just Noticeable Difference (JND) for offsets between the motion-tracked real and the corresponding virtual objects of the same size. Further, we plan to extend the developed IVE by hand and arm tracking, enabling a higher grade of self-embodiment.

REFERENCES

[1] C. Armbrüster, M. Wolter, T. Kuhlen, W. Spijkers, and B. Fimm. “Depth perception in virtual reality: distance estimations in peri-and extrapersonal space”. In: *Cyberpsychology & Behavior* 11.1 (2008), pp. 9–15.

[2] International Telecommunication Union. “ITU-T P. 910”. In: *Subjective video quality assessment methods for multimedia applications* (2008).

[3] V. Interrante, B. Ries, and L. Anderson. “Distance perception in immersive virtual environments, revisited”. In: *IEEE virtual reality conference (VR 2006)*. IEEE. 2006, pp. 3–10.

[4] P. E. Napieralski et al. “Near-field distance perception in real and virtual environments using both verbal and action responses”. In: *ACM Transactions on Applied Perception (TAP)* 8.3 (2011), pp. 1–19.

[5] 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”. In: *i-Perception* 8.3 (2017), p. 2041669517708205.

[6] N. Ogawa, T. Narumi, and M. Hirose. “Effect of Avatar Appearance on Detection Thresholds for Remapped Hand Movements”. In: *IEEE Transactions on Visualization and Computer Graphics* (2020).

[7] A. Peer and K. Ponto. “Evaluating perceived distance measures in room-scale spaces using consumer-grade head mounted displays”. In: *2017 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE. 2017, pp. 83–86.

[8] A. Singla, R. R. R. Rao, S. Göring, and A. Raake. “Assessing Media QoE, Simulator Sickness and Presence for Omnidirectional Videos with Different Test Protocols”. In: *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE. 2019, pp. 1163–1164.

[9] M. Suznjevic, M. Mandurov, and M. Matijasevic. “Performance and QoE assessment of HTC Vive and Oculus Rift for pick-and-place tasks in VR”. In: *2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE. 2017, pp. 1–3.

[10] P. Willemsen and A. A. Gooch. “Perceived egocentric distances in real, image-based, and traditional virtual environments”. In: *Proceedings IEEE Virtual Reality 2002*. IEEE. 2002, pp. 275–276.