Accuracy of Commodity Finger Tracking Systems for Virtual Reality Head-Mounted Displays

Daniel Schneider* Coburg University
Alexander Otte† Coburg University
Axel Simon Kublin‡ Coburg University
Alexander Martschenko§ Coburg University
Per Ola Kristensson¶ University of Cambridge
Eyal Ofek‖ Microsoft Research
Michel Pahud** Microsoft Research
Jens Grubert†† Coburg University

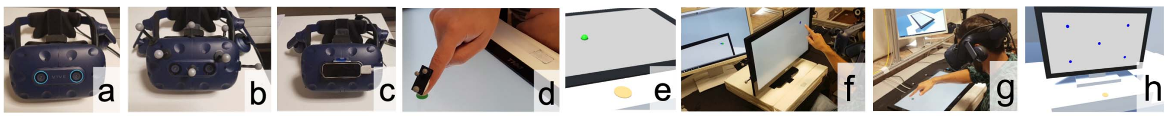

Figure 1: The HTC Vive Pro used for condition a: VIVE. b: OPTITRACK. c: LEAP MOTION. d: User touching the physical screen at the center target position. e: virtual finger tip (green sphere) relative to the target position (dark green) Participant conducting the task in f: VERTICAL g: HORIZONTAL orientation. h: view on all target positions in VR (only a single one at a time is highlighted during the experiment). The yellow disc in e and h indicates the starting position.

ABSTRACT

Representing users' hands and fingers in virtual reality is crucial for many tasks. Recently, virtual reality head-mounted displays, capable of camera-based inside-out tracking and finger and hand tracking, are becoming popular and complement add-on solutions, such as Leap Motion.

However, interacting with physical objects requires an accurate grounded positioning of the virtual reality coordinate system relative to relevant objects, and a good spatial positioning of the user's fingers and hands.

In order to get a better understanding of the capabilities of Virtual Reality headset finger tracking solutions for interacting with physical objects, we ran a controlled experiment ($n = 24$) comparing two commodity hand and finger tracking systems (HTC Vive and Leap Motion) and report on the accuracy of commodity hand tracking systems.

1 INTRODUCTION

The ability to see one's hands and finger movement inside Virtual Reality (VR) space opens opportunities to use tools in VR in a natural fashion and is crucial for many tasks [1].

Lately, commercial VR head-mounted displays (HMDs) are progressing to 'inside-out' tracking using multiple built-in cameras. Inside-out tracking allows a simple setup of the VR system, and the ability to work in un-instrumented environments. Tracking the user hands in real time is a potentially useful capability, that is already being proposed by two of-the-shelve platforms: Vive Pro from HTC, and the Leap Motion sensor, while Oculus Quest[1] and the upcoming HoloLens 2 expect to support it in the near future.

*e-mail: Daniel.Schneider@hs-coburg.de
†e-mail: Alexander.Otte@hs-coburg.de
‡e-mail: Axel-Simon.Kublin@stud.hs-coburg.de
§e-mail: Alexander.Martschenko@stud.hs-coburg.de
¶e-mail: pok21@cam.ac.uk
‖e-mail: eyalofek@microsoft.com
**e-mail: mpahud@microsoft.com
††e-mail: jens.grubert@hs-coburg.de

[1] https://www.oculus.com/blog/introducing-hand-tracking-on-oculus-quest-bringing-your-real-hands-into-vr/ Last access November 21st, 2019.

The applicability of hand and finger tracking has a crucial dependency on the accuracy of tracking the user's fingers: When mixing the interaction between physical input devices and 3D hand tracking, there is a special importance to the relations between those two very different techniques of finger sensing, starting from their comparative accuracy, the ability to select an object and the minimal distance between such an object and nearby ones to prevent selection errors.

The performance of the Leap Motion controller was already studied in various context (e.g., accuracy measurements for pen-based interaction [6], for desktop-based setups, i.e. mounted in front of a monitor [5], or analysis of gesture recognition, e.g., [3]). Recently, it was also evaluated in a VR setup (i.e. mounted in front of a VR HMD) [2], but no metric accuracy results were reported. For recent headset-integrated tracking solutions (such as the HTC Vive) we are not aware of accuracy evaluations.

2 USER STUDY

In order to get a better understanding of the capabilities of hand tracking solutions for VR headsets, we ran a controlled experiment, comparing two commodity hand tracking systems. As a high-accuracy baseline, often used for prototyping interactions in VR and AR, we also included a stationary outside-in tracking system (OptiTrack Prime 13). We highlight, that it is expected that his baseline would outperform commodity tracking systems.

The experiment was a 3x2 within subjects design. The *independent variables* (all counterbalanced) were:

INTERFACE with three levels: VIVE which was using the HTC Vive built-in hand tracking system and LEAP, which was using a front mounted Leap Motion controller (add-on sensing system) on an HTC Vive Pro and OPTITRACK, which was using an OptiTrack Prime 13 tracking system. OPTITRACK was chosen to get baseline data of a high-accuracy tracking system, even though it is not intended for mobile use.

TARGET ORIENTATION with two levels: VERTICAL (see Figure 1, f) and HORIZONTAL (see Figure 1, g).

The investigated *dependent variable* is the accuracy (Euclidean distance between the centers of the touch point and the target).

2.1 Task, Procedure and Apparatus

The participants (10 female, 14 male, mean age 30.8 years, sd = 7.9, mean height 174.1 cm, sd = 7.9) were told to conduct a target acquisition task. The participants had to place the index finger of their preferred hand on a starting position (yellow spot in Figure 1, e and h). They were asked to hit a target with their virtual finger tip as accurately and fast as possible as soon as it appeared on the touchscreen monitor and to hold this position two seconds (to collect

978-1-7281-6532-5/20/$31.00 ©2020 IEEE
DOI 10.1109/VRW50115.2020.00253

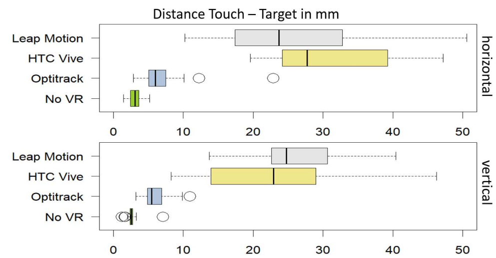

Figure 2: Average distances (over all target positions) between the centroid of the physical touch point and the target center in mm.

multiple tracking samples). After this, participants had to return their index finger to the starting position. One trial consisted of hitting five circular targets (diameter: 16 mm), which appeared in a fixed order (see Figure 1, h): top left, top right, center, bottom left, bottom right. The distance between the center target and the starting position is 25.9 cm for the VERTICAL orientation and 20.1cm for HORIZONTAL. The other four targets have an offset from the center target of 15.2 cm to the left (or right) and 8.6 cm to the top or bottom. Participants were asked to conduct five trials.

After welcoming, the volunteers were asked to fill out a demographic questionnaire. At the beginning of each block (being either HORIZONTAL or VERTICAL ORIENTATION), the participants were asked to conduct the task in a baseline NOVR to collect ground truth data, in which they did not wear an HMD. In this condition, they pursued the same procedure as in the later conditions. After each block they were asked to fill out a preference questionnaire. Finally a semi structured interview and debriefing was conducted.

As HMD the the HTC VIVE Pro (one modified with retro-reflective markers for the OptiTrack system and another one was equipped with the Leap Motion controller) was used. In VR, users' index finger tip was visualized as a sphere (following a recommendation of Grubert et al. [1]) with a diameter of 16 mm. For tracking the finger tip in OptiTrack, we created a rigid body (set of retro-reflective markers) and attached it to the user's index finger, see Figure 1, d. This rigid body was not present in the other conditions. The system was implemented in Unity 2019.2 and deployed on a PC (Intel Xeon E5-1650 processor, 64 GB RAM, Nvidia GTX 1070 graphics card) running Windows 10. The Monitor (Dell s2340t, screen: 56.2cm x 34cm) was spatially tracked. The OptiTrack outside-in tracking system used 13 cameras (positional accuracy of 0.24mm).

2.2 Results

Data from one participant had to be excluded due to an operator error.

Results for the euclidean distance (averaged over all targets) between the centers of the touch point and the target are depicted in Figure 2. An omnibus test on log-transformed data revealed that there were significant main effects of INTERFACE ($F_{2,44} = 306.52$, $p < 0.001$, $\eta_p^2 = 0.93$, $1-\beta = 1.0$).Holm-Bonferroni adjusted post-hoc testing level $\alpha = 0.05$ revealed that for INTERFACE there were significant differences between OPTITRACK and VIVE ($p < 0.001$), OPTITRACK and LEAP ($p < 0.001$) as well as between VIVE and LEAP ($p = 0.04$). No significant interactions have been indicated between between INTERFACE and ORIENTATION ($F_{2,44} = 0.34$, $p = 0.71$, $\eta_p^2 = 0.14$, $1-\beta = 0.10$). We did not observe any asymmetrical effects [4].

Besides calculating the euclidean distance to the target we were also interested in the depth offset between the target plane and the virtual finger tip after calibration. This measure can serve as basis for determining minimal Z distances for differentiation between 3D objects given the evaluated hand tracking systems. To this end, we determined the 95% confidence intervals and average distances as depicted in Table 1.

In other words, as expected, OPTITRACK resulted in significant

Table 1: Average Z-distances and 95% Confidence intervals (in brackets) between the virtual finger tip and the target plane in mm before and after calibration.

	OptiTrack	Vive	Leap Motion
Horizontal	3.0 [2.7, 3.3]	24.5 [20.8, 28.2]	13.3 [10.0, 16.5]
Vertical	3.4 [3.1, 3.7]	53.9 [49.1, 58.7]	12.6 [9.3, 15.9]

higher accuracy results compared to VIVE and LEAP, but also LEAP resulted in significantly higher accuracy than VIVE.

From the 24 participants 19 indicated that OPTITRACK was overall their favorite interface, one participant favored it in HORIZONTAL orientation (LEAP in VERTICAL), one participant favored OPTITRACK and LEAP equally. Three participants favored LEAP overall. Two participants who declared LEAP as their favorite HMD mentioned the lack of a fiducial on their finger, which was described as "irritating" in both occasions.

3 Discussion

The purpose of the experiment was to quantify the performance of mobile finger tracking solutions available today. It was expected, that those systems would result in lower accuracy compared to a stationary outside-in tracking system. Our results indicate that the VIVE system results in an average error of over 30 mm (sd = 8.98) relative to physical target location on horizontal surfaces such as keyboards or tablet screens, with the average Z-distance being 24.5 mm (sd = 9.1) and 53.9 mm (sd = 11.7) for vertical surfaces. The LEAP system still results in an average error of over 26 mm (sd = 11.20) relative to the horizontal physical target location, with the average Z-distance being 13.3 mm (sd = 7.8).

A limitation of our work is the partially reflective touch surface and therefore it should be investigated how accurate the systems work on other surfaces.

4 Conclusions

To gain a better understanding of the capabilities of virtual reality headset finger tracking solutions for interacting with physical objects, we ran a controlled experiment ($n = 24$) comparing two commodity hand and finger tracking systems (HTC Vive and Leap Motion). Our results indicate that when interacting with horizontally aligned surfaces in walk-up-and-use scenarios, the Vive hand tracking system exhibits Z-errors over 30 mm and the Leap Motion system over 26 mm. For vertically aligned surfaces, this Z-error is over 50 mm for Vive and over 10 mm with Leap motion. These errors have to be taken into account when designing VR experiences for interacting with physical objects with commodity hand tracking systems. In future work, we plan to investigate the performance of upcoming finger tracking systems and work on approaches to mitigate the accuracy issues of today's commodity hand tracking systems for VR headsets.

References

[1] J. Grubert, L. Witzani, E. Ofek, M. Pahud, M. Kranz, and P. O. Kristensson. Effects of hand representations for typing in virtual reality. In *2018 IEEE VR*, pp. 151–158. IEEE, 2018.

[2] S. Lindsey. *Evaluation of Low Cost Controllers for Mobile Based Virtual Reality Headsets*. PhD thesis, 2017.

[3] G. Marin, F. Dominio, and P. Zanuttigh. Hand gesture recognition with jointly calibrated leap motion and depth sensor. *Multimedia Tools and Applications*, 75(22):14991–15015, 2016.

[4] E. Poulton and P. Freeman. Unwanted asymmetrical transfer effects with balanced experimental designs. *Psychological Bulletin*, 66(1):1, 1966.

[5] P. P. Valentini and E. Pezzuti. Accuracy in fingertip tracking using leap motion controller for interactive virtual applications. *IJIDeM*, 11(3):641–650, 2017.

[6] F. Weichert, D. Bachmann, B. Rudak, and D. Fisseler. Analysis of the accuracy and robustness of the leap motion controller. *Sensors*, 13(5):6380–6393, 2013.