2024 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)

Eye Tracking Performance in Mobile Mixed Reality

Satyam Awasthi*[†] Vivian Ross*[‡] Sydney Lim*[§] Michael Beyeler[¶] Tobias Höllerer[‖]
University of California, Santa Barbara
Santa Barbara, CA, USA

Head-constrained Body-constrained Screen-Stabilized Walking World-Stabilized Walking Hallway

Figure 1: Tasks from our user study performed on augmented reality (AR) headsets to test the in-built eye-tracker' accuracy

ABSTRACT

Implementing and evaluating eye tracking across multiple platforms and use cases can be challenging due to the lack of standardized metrics and measurements. Additionally, existing calibration methods and accuracy measurements often do not account for the common scenarios of walking and scanning in mobile AR settings. We conducted user studies evaluating eye tracking on the Magic Leap One, the Meta Quest Pro, and the HoloLens 2. Our results reveal that the degree to which locomotion influenced eye tracking performance depended on the headset, with the HoloLens 2, which features a retractable visor, displaying the greatest decrease in accuracy during locomotion.

Index Terms: Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Mixed / augmented reality; Human-centered computing—Human computer interaction (HCI)—HCI design and evaluation methods—User studies; Human-centered computing—Ubiquitous and mobile computing—Ubiquitous and mobile devices—Mobile devices;

1 INTRODUCTION

Recent advancements in eye-tracking technology have made it more accessible and affordable for a wider range of applications [3]. As such, it is important that developers and users understand the capabilities and limitations of their eye tracking devices during use. With the rise of head-worn augmented reality (AR) displays such as the HoloLens 2, Magic Leap, Meta Quest Pro, and the Apple Vision Pro, eye trackers have been integrated into various platforms and use cases. However, optimally utilizing eye-tracking technology on multiple platforms and standardizing their performance is a difficult and complex problem.

Specific metrics, such as accuracy and precision, are often used to evaluate the performance of eye tracking systems [1, 3, 4]. Although existing work has used these metrics to examine the accuracy of eye trackers while stationary, there has been little research into how locomotion affects eye tracking, even though there is considerable interest in mobile eye tracking for specific applications, such as assessing behaviors for shopping, navigation, or wayfinding [2, 5]. Current calibration methods for eye tracking devices rely on users keeping their heads stationary, without incorporating any head or body movement. It is unclear whether these methods are effective when a user intends to use eye tracking while walking.

Our objective is to establish improved testing practices for mobile eye tracking development in AR headsets, particularly during locomotion. To achieve this goal, we conducted three user studies to assess the spatial accuracy and precision of the gaze signal captured by the integrated eye tracker of the Magic Leap, Meta Quest Pro, and HoloLens 2. We present the results of our study and provide an analysis of the eye tracking signal quality, focusing on accuracy across various AR scenarios.

The insights gained from this study contribute to our understanding of the variations in eye tracking accuracy observed across different tasks in our test suite, which can be attributed to the hardware design of the respective headsets. Both datasets will be made publicly available for further research and analysis.

2 USER STUDIES

Three separate user studies were carried out to evaluate the performance of eye tracking on different devices: the Magic Leap 1 ($n = 36$), the Meta Quest Pro ($n = 29$), and the HoloLens 2 ($n = 54$), with slight modifications to adapt the test suite to differences in the three headsets. We outline the tasks presented in the three studies below.

Head-Constrained Static (HCS) Users rest their head on a chin rest and watch a tracking stimulus on the screen, which shifts to random positions after a brief delay.

Head-constrained Moving (HCM) Users rest their head on a chin rest and watch a tracking stimulus, which moves along a path on the screen inside their field of view (FOV).

Body-constrained (BC) Users sit and turn their head to watch a tracking stimulus, which moves along a path in the world frame of reference (FOR) and wider than their FOV.

Screen-Stabilized Walking (SSW) Users walk in circles around a table while watching a tracking stimulus, which moves along a path on the screen inside their FOV.

World-Stabilized Walking (WSW) Users walk in circles around a table while watching a tracking stimulus, which moves along a path in the world FOR above the table.

In addition to these five common tasks, only the user study performed on the HoloLens 2 incorporated one additional unique task, the hallway task. This task was added due to participant feedback that the existing tasks, which incorporated walking in circles around a table, felt unnatural and unintuitive. Our goal was to create a walking task that might more closely resemble real-life use-cases of augmented reality with locomotion.

---

*Awasthi, Ross, and Lim contributed equally to this work

[†] e-mail:satyam@ucsb.edu
[‡] e-mail:vivianross@ucsb.edu
[§] e-mail:sydneylim@ucsb.edu
[¶] e-mail:mbeyeler@ucsb.edu
[‖] e-mail:holl@cs.ucsb.edu

979-8-3503-7449-0/24/$31.00 ©2024 IEEE
DOI 10.1109/VRW62533.2024.00321

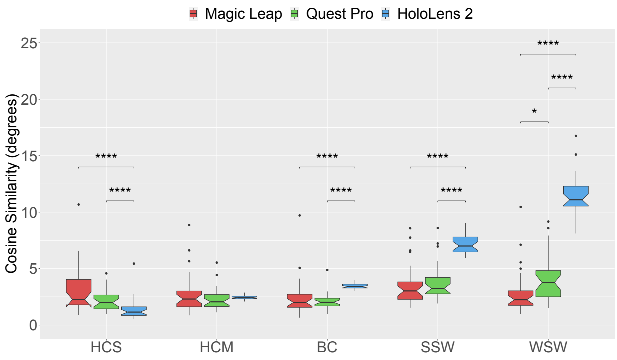

Figure 2: Cosine similarity per task for the Magic Leap, Meta Quest Pro, and HoloLens 2. Notches show 95% confidence interval, * denotes $p < 0.05$, and ** denote $p < 0.01$.

Hallway (H) Users walk down a hallway while watching a tracking stimulus, which moves down the hallway in front of them.

3 RESULTS

Results for all three studies are summarized here. Exactly like [1], we employed the inverse cosine of the dot product between the gaze and stimulus vectors to establish the cosine error.

Fig. 2 shows a graph of cosine error split by task for all three devices. Due to the different numbers of participants in each user study, the data from 29 random participants was selected from the Magic Leap One and HoloLens 2 studies in order to match the quantity of data available from the Meta Quest Pro and allow for a better comparison of the three studies. An ART ANOVA with independent variables “task” and “device” revealed a significant difference in cosine error between tasks ($F = 341, p < 2.22 * 10^{-16}$), a significant difference in cosine error between devices ($F = 49, p < 8.77 * 10^{-15}$), and an interaction between task and device ($F = 147, p < 2.22 * 10^{-16}$). For the Magic Leap One and Meta Quest Pro, a Bonferroni-corrected Mann-Whitney U posthoc test reveals a difference in cosine error in the world-stabilized walking task ($p = 0.0024$). For the Magic Leap One and HoloLens 2, the same posthoc test reveals a difference in cosine error in the head-constrained static task ($p = 1.94 * 10^{-6}$), the body-constrained task ($p = 1.01 * 10^{-6}$), the screen-stabilized walking task ($p = 1.32 * 10^{-7}$), and the world-stabilized walking task ($p = 4.02 * 10^{-10}$). For the Meta Quest Pro and HoloLens 2, the posthoc test reveals a difference in cosine error in the same tasks: the head-constrained static task ($p = 7.5 * 10^{-5}$), the body-constrained task ($p = 3.6 * `0^{-9}$), the screen-stabilized walking task ($p = 6.48 * 10^{-8}$), and the world-stabilized walking task ($p = 2.4 * 10^{-10}$).

Our overall results for the HoloLens 2 indicate that spatial accuracy for the hallway task is lower than the error on either walking task, but higher than the error on non-walking tasks.

4 DISCUSSION

Overall, the HoloLens 2 results differed the most from those of the other two headsets, with significantly higher errors, especially in walking tasks. This may be due to the difference in display quality which sometimes made the tracking stimulus difficult to see during the task. Although the hallway task was more accurate than either of the other two walking tasks, it was still less accurate than any of the tasks without locomotion.

The only task for which the Meta Quest Pro and Magic Leap have significantly differing eye tracking performance is the world-stabilized walking task, in which the Meta Quest Pro has a lower accuracy than the Magic Leap One. This may be due to the Meta Quest Pro using a video pass-through display instead of an optical-see-through display. Seeing the world through an imperfect video feed while navigating around a center table that is to be kept in focus is more taxing to the human visual system than using an optical-see-through display.

The differences between the accuracy of the HoloLens 2 compared to the Magic Leap One and Meta Quest Pro may largely be explained by the fit of these headsets. The Magic Leap One rests on the nose, forming a secure fit around a user’s eyes, and is the only headset having its compute unit separated from the headset, making it relatively stable while moving. The Meta Quest Pro uses a band that wraps around the forehead, with the display fixed in place. In contrast, the HoloLens 2 has a visor housing the display that can be flipped up, which can result in a lack of stability between the display and the eyes. This leads to more possibility of headset movement and thus, lower accuracy for tasks when head and body motion increase. The lack of a stable fit of the HoloLens 2 may contribute to its significantly lower accuracy of the eye tracker unless the head is kept still, whereas the Magic Leap One and Meta Quest Pro headsets show more consistent eye tracking accuracy throughout.

In fact, for the head-constrained static task task, the HoloLens 2 exhibited better eye tracking performance than the Magic Leap and Meta Quest Pro. This can be due to the greater distance between the display and the user’s eyes on the HoloLens 2, which may have allowed its eye tracking cameras to better capture eye gazes at the periphery of the display. Differences in the built-in calibration procedure may also have an effect here.

5 CONCLUSION AND FUTURE WORK

In this work, we compare the eye tracking capabilities of the Magic Leap One, Meta Quest Pro, and HoloLens 2 during locomotion. The included tasks account for when a user is sitting with their head held still, sitting with their head freely moving, or walking, as well as whether a user is viewing screen-stabilized, body-stabilized, or world-stabilized content.

Overall, our work provides future researchers with a means of investigating the accuracy of the eye tracker of their AR device. This may allow researchers to gain valuable insights, improving the performance of their eye tracker in mobile mixed reality applications.

ACKNOWLEDGMENTS

This work was supported in part by NSF grant IIS-2211784 as well as ONR grants N00014-19-1-2553 and N00014-23-1-2118. Thanks also go to Dr. Tom Bullock for helpful discussions and loaning of a chin rest.

REFERENCES

[1] S. Aziz and O. Komogortsev. An assessment of the eye tracking signal quality captured in the HoloLens 2. In *2022 Symposium on Eye Tracking Research and Applications*. ACM, June 2022. doi: 10.1145/3517031.3529626

[2] J. de Winter, P. Bazilinskyy, D. Wesdorp, V. de Vlam, B. Hopmans, J. Visscher, and D. Dodou. How do pedestrians distribute their visual attention when walking through a parking garage? An eye-tracking study. *Ergonomics*, 64(6):793–805, June 2021.

[3] A. T. Duchowski. *Eye tracking methodology: Theory and practice*. Springer, 2017.

[4] S. Kapp, M. Barz, S. Mukhametov, D. Sonntag, and J. Kuhn. ARETT: Augmented reality eye tracking toolkit for head mounted displays. *Sensors*, 21(6), 2021. doi: 10.3390/s21062234

[5] P. Kiefer, I. Giannopoulos, M. Raubal, and A. Duchowski. Eye tracking for spatial research: Cognition, computation, challenges. *Spat. Cogn. Comput.*, 17(1-2):1–19, Jan. 2017.