Comparative Analysis of Augmented Reality Devices for Surgical Applications

Alexis Bader[1,2,3]✉, **Yao Li**[2,3], **Sergey Drobinsky**[4], **Matías de la Fuente**[3], **Jan Egger**[5], **Behrus Puladi**[2,3]

[1]Department of Systems Engineering, [2]University of Waterloo, Canada Department of Oral and Maxillofacial Surgery, University Hospital RWTH Aachen, [3]Institute of Medical Informatics, University Hospital RWTH Aachen, [4]Institute of Medical Engineering, RWTH Aachen University, [5]Institute for Artificial Intelligence in Medicine, University Hospital Essen; ✉a3bader@uwaterloo.ca

Introduction

At this time, a definitive standard for an optimal display for surgical applications in augmented reality (AR) head-mounted displays (HMDs) has yet to be established. Consequently, the suitability of devices for specific applications must be determined through evaluation. In the surgical context, it is imperative that devices do not impede the surgeon's field of view, accurately represent depth for virtual projections, and have sufficient resolution to render the intricate details of images. Therefore, the objective of this study is to evaluate four AR-HMDs to determine if there are differences in user depth perception, see-through contrast, and resolution.

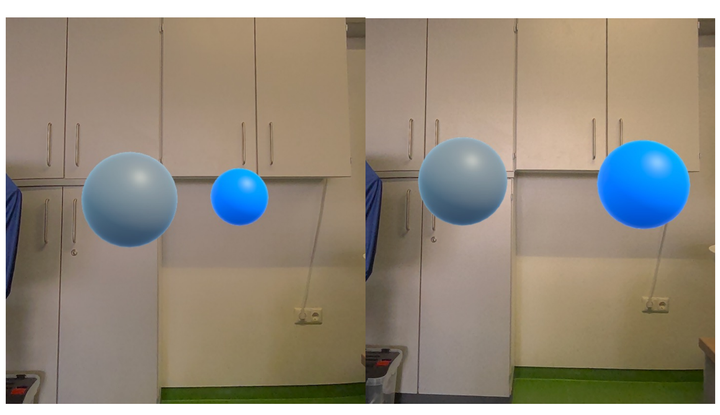
Methods

A Unity project was developed and deployed to four HMD devices, namely the Apple Vision Pro (AVP) ($3499), Meta Quest 3 ($499), HoloLens 2 ($3500), and Varjo XR3 ($5995). The experiment was divided into three sections and performed for each of the four devices for two subjects.

The setup of a participant for the test with Meta Quest 3 and keyboard

I. Depth Perception

In the study, two virtual cubes, spheres, and cylinders 10 cm wide were rendered between 30 cm and 100 cm away from the participant: one as the target object and the other movable by the participant. The task was to align the depth of the movable object to the target object.



Comparison of the initial position (left) and the aligned position (right) of the two virtual 3D objects

II. Display Resolution

In the next section, three horizontal, vertical, and diagonal lines were rendered in front of the participant. The subject was required to reduce the thickness of the lines until they were no longer visually discernible on the display.

Comparison of the initial thickness of the vertical line resolution test (left), and the minimal thickness needed to maintain visibility (right).

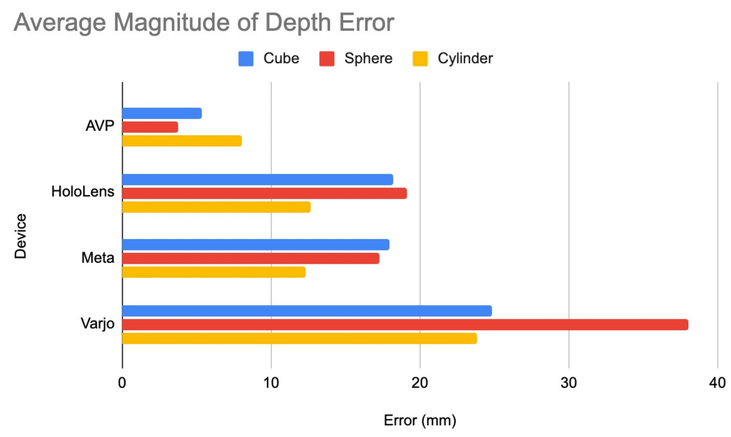
III. Contrast Detection

Images of head and neck surgeries were analyzed to extract a palette of colors typically seen in such procedures. These colors were used to generate images with random combinations of letters. Participants were then asked to read the letters from five of these different printed images for each device.

A comparison of contrast perception between a physical image (left) and its view through a virtual see-through device (right) from the surgical colour palette (top).

Results and Discussion

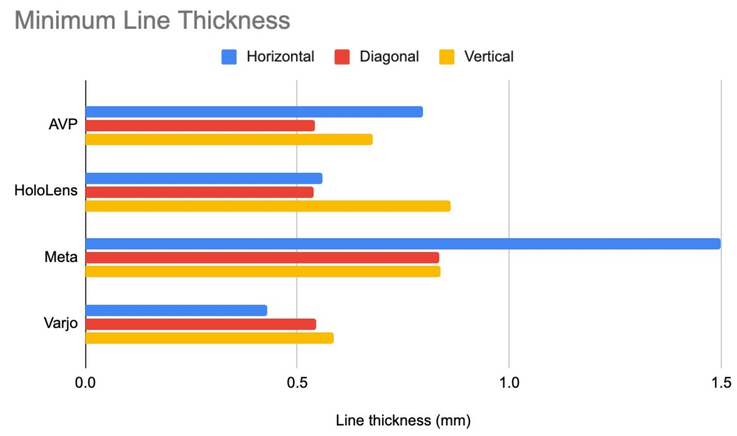
I. Depth Perception



Average magnitude of the relative error of each moveable and target object depth for each device.

An ANOVA analysis revealed significant differences in the depth capabilities the devices. A Fisher LSD test revealed that there is a significant difference between of all devices tested, with the exception of the HoloLens and Meta devices, which exhibited similar performance. The AVP exhibited the lowest depth error, demonstrating superior performance compared to the other devices. Conversely, the Varjo exhibited the highest magnitude of depth error. No dominant trend was identified in the error associated with each specific shape.

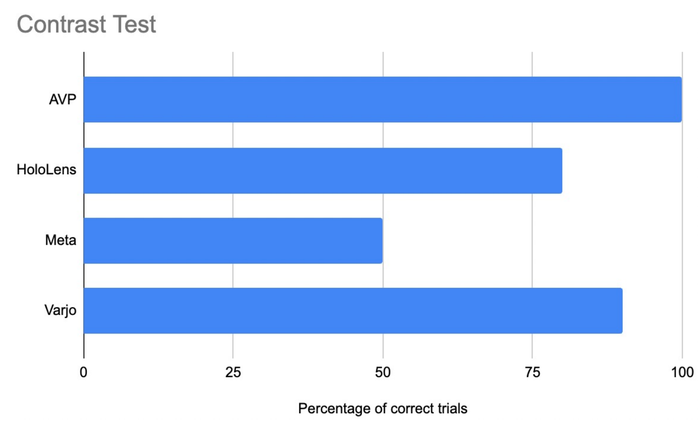
II. Display Resolution



Minimum visible line thickness for each line orientation for each device.

An ANOVA analysis revealed no significant differences in the resolution capabilities among the devices. While there are slight differences in the minimum line thickness between the devices, one device does not offer an advantage in resolution capabilities. Additionally, there is no dominant trend in the relationship between the direction of lines and the minimum display thickness across the various devices.

III. Contrast Detection



Percent correctness of letter reporting for each device.

The results of the contrast test indicate that the AVP is the best display for visualizing low-contrast environments with an HMD AR device. Following the AVP, the Varjo and HoloLens devices also demonstrated comparable performance, whereas the Meta device exhibited a markedly inferior outcome in this assessment.

Conclusion

Among the four AR-HMD devices evaluated in this study, the Apple Vision Pro is most promising for use as a surgical tool. It significantly outperforms the other devices in simulating accurate depth cues, allowing users to better perceive the depth of objects in the virtual space. Although the AVP does not exhibit a notable distinction in resolution when compared to the other devices, it attains an optimal rating on the contrast detection test. This provides surgeons with near-perfect visibility of their surgical field while also benefiting from high-resolution virtual projections.