Spatial Consistency Perception in Optical and Video See-Through Head-Mounted Augmentations

Alexander Plopski, Kenneth R. Moser, *Student Member, IEEE*, Kiyoshi Kiyokawa, *Member, IEEE*, J. Edward Swan II, *Member, IEEE* and Haruo Takemura, *Member, IEEE*

ABSTRACT

Correct spatial alignment is an essential requirement for convincing augmented reality experiences. Registration error, caused by a variety of systematic, environmental, and user influences decreases the realism and utility of head mounted display AR applications. Focus is often given to rigorous calibration and prediction methods seeking to entirely remove misalignment error between virtual and real content. Unfortunately, producing perfect registration is often simply not possible. Our goal is to quantify the sensitivity of users to registration error in these systems, and identify acceptability thresholds at which users can no longer distinguish between the spatial positioning of virtual and real objects. We simulate both video see-through and optical see-through environments using a projector system and experimentally measure user perception of virtual content misalignment. Our results indicate that users are less perceptive to rotational errors over all and that translational accuracy is less important in optical see-through systems than in video see-through.

Index Terms: H.5.1 [[Information Interfaces and Presentation]: Multimedia Information Systems]: Artificial, augmented, and virtual realities—; H.5.2 [[Information Interfaces and Presentation]: User Interfaces]: Ergonomics, Evaluation/methodology, Screen design—

1 INTRODUCTION

At a fundamental level, the requirement of most Augmented Reality (AR) applications is to display virtual information, text, or geometry so that it appears statically aligned with, or registered to, existing physical objects in the environment. Improper spatial alignment due to modeling errors, calibration errors, or tracking problems produce application failures, at high levels, and low perceptual quality and utility, in general, due to confusion and misunderstandings [2, 6]. Mechanisms to maintain or improve registration between virtual and real content vary between systems and display types, with some techniques being applicable to only certain device technologies.

Hand-held and head-mounted video see-through (VST) devices are able to precisely control every portion of the scene visible to the user, including the appearance of physical objects as they are captured by the camera integrated into the device. As such, misalignment errors caused by improper modeling between the physical and virtual camera in the rendering engine can be minimized to negligible levels due to highly robust, accurate, and widely available camera calibration and vision based tracking methods. Unfortunately, these same corrective techniques are not viable for optical see-through (OST) AR systems since content displayed on the HMD screen must be aligned, not from the perspective of a camera, but with that of the user's eye. Current solutions to estimate the parameters of the user's perspective, involve a calibration step that aligns pixels on the HMD with known points in the environment [1, 5, 7]. These spatial calibration methods, though, are not able to achieve consistently accurate results due to user induced and modeling errors. As a result, registration quality in OST AR applications is often noticeably low and may even further degrade over time. Provided that a level of registration error will always persist in an OST HMD system, it is necessary to understand user perception of this error and the just noticeable levels at which discrepancies between virtual and real content are realized.

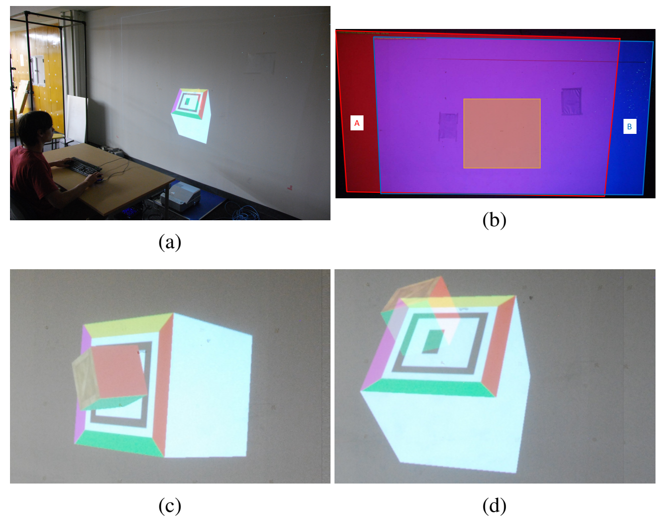


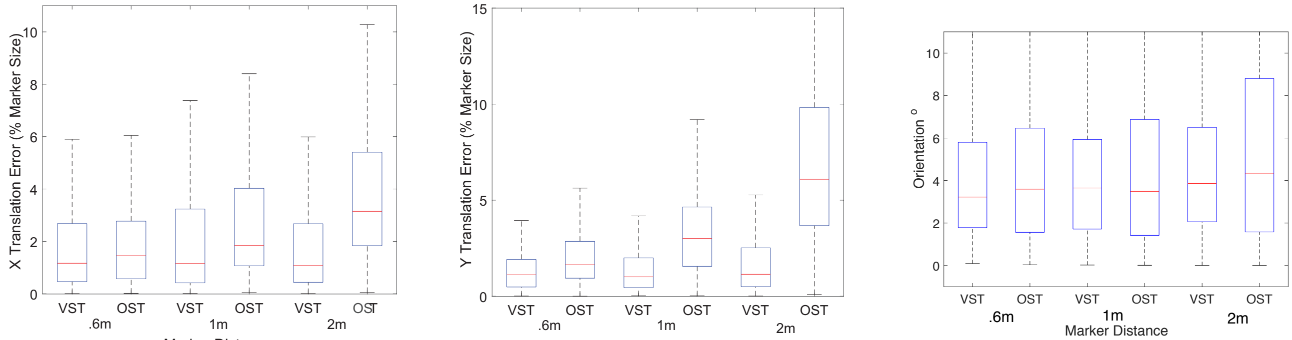
Figure 1: (a) View of our experimental environment, showing the location of subjects relative to the projected images. (b) The overlap produced by our dual projector system. Images in both conditions are only rendered within the overlapping region. Sample images, as seen by subjects, for the (c) VST and (d) OST condition.

The objectives of our study are to obtain quantifiable user tolerance levels to position and rotational error and, additionally, identify if these just noticeable thresholds differ between HMD presentation methods, VST and OST. Achieving these goals will greatly benefit AR system designers by revealing acceptable base-line tolerances which will help to avoid unnecessary accuracy and calibration over requirements. Furthermore, these levels will allow researchers to more easily identify when an OST system calibration must be repeated or once sufficient accuracy is achieved.

1.1 Display Method

Simulating OST AR in a VR environment is common practice for studies investigating latency effects on user performance [3, 4]. Within the VR environment, systematic noise and influencing factors can be controlled, or at the least, equalized across conditions. Using this same reasoning, We provide simulated VST and OST view to the users using overlapping projector images. We use two SANYO PDG-DWL2500J, 1280×800 native resolution set to display at 1920×1080, with a maximum contrast ratio of 2000:1 and brightness of 2500 lumens. The two projectors are positioned side-by-side to illuminate a single wall. Figure 1 (a) and (b) shows the position, relative to the subject, and overlap of both projectors.

VST Display Mode We simulate the VST AR condition using a straightforward implementation. Normal VST content simply consists of a combined camera image and computer generated augmentations. Our VST mode mimics this by using one projector to display a single image containing both the tracking marker and virtual object. Figure 1 (c) demonstrates how the projected image appears to our subjects.



(a) (b) (c)

Figure 2: (a) X translation error and (b) Y translation error by marker distance. Values along the y axis represent error magnitudes in terms of visual angle. (c) Angular error by marker distance. Values along the y axis represent the rotational difference between quaternion orientations.

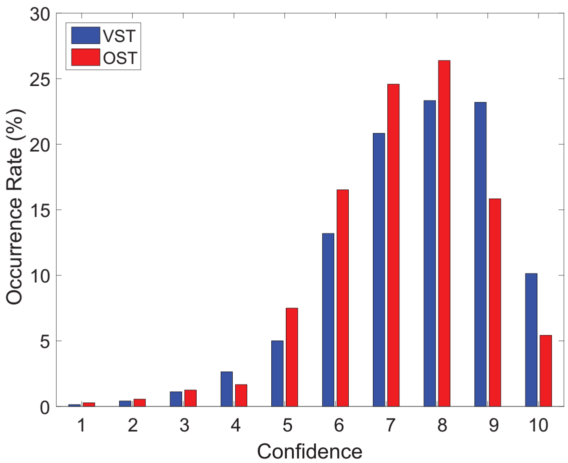


Figure 3: Confidence distributions for user responses in the VST and OST display sets.

OST Display Mode In order to create the simulated OST view for our experiment, we utilize both projectors to simultaneously produce one image for world content and the second for virtual content. The result, as provided in Figure 1 (d), creates an image that possesses the same transparency, varying brightness, and color consistency effects one would see from an OST HMD. We perform a brief calibration of the experimental environment to ensure that content, from both conditions, is properly displayed from the user's perspective with accommodation for distortion and skew produced by the projectors' orientations and positions.

2 Experiment Task and Analysis of Results

16 subjects (11 male, 5 female), between the ages of 21 and 33 (mean age of 24.8 years, stddev 3.6 years), participated in our experiment. 7 subjects claimed to have little to no prior experience with AR. Each subject was confirmed to have normal, or corrected to normal, vision and were monetarily compensated for their time. Subjects were instructed to align a virtual 6 sided cube with a $10cm \times 10cm$ target marker shown at 8 rotation orientations over 3 discrete distances. Through a simple docking procedure, subjects were able to transform the location of the cube in 5DOF, X and Y translation with 3 axis rotation, until they were satisfied that the cube's orientation best fit that of the marker. We recorded the translation and rotation offset between the subject responses and marker ground-truth positions, as well as a metric denoting subject confidence in each.

Figures 3 and 2 provide user confidence and response errors for both OST and VST conditions. The distribution of confidence values shows that subjects were able to more easily resolve position mismatch between the cube and marker for VST images compared to the OST display mode. The highest three confidence levels, 8-10, combined, yield over half, 56.67%, of the total responses for the VST mode compared to only 47.6% of the total responses for the OST trial sets. Likewise, the X and Y translation error show an over all trend for higher error in OST matches, especially when the marker was rendered at greater distances. Interestingly, orientation error did not show as strong a deviation between display modes.

As part of future studies, we want to include visualizations with stereo imagery to more properly model depth error. We also believe that an additional investigation of how transparency level impact the perceived alignment of the virtual and real objects is also warranted.

Acknowledgements

This research was funded in part by Grant-in-Aid for Scientific Research (B), #15H02738 from Japan Society for the Promotion of Science (JSPS), Japan.

References

[1] Y. Itoh and G. Klinker. Interaction-free calibration for optical see-through head-mounted displays based on 3d eye localization. In *Proc. IEEE Symposium on 3D User Interfaces (3DUI)*, pages 75–82, 2014.

[2] B. M. Khuong, K. Kiyokawa, A. Miller, J. La Viola, T. Mashita, and H. Takemura. The effectiveness of an ar-based context-aware assembly support system in object assembly. In *Virtual Reality (VR), 2014 iEEE*, pages 57–62, March 2014.

[3] C. Lee, S. Bonebrake, T. Hollerer, and D. A. Bowman. A replication study testing the viability of ar simulation for vr in controlled experiments. In *Proc. IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 203–204, 2009.

[4] C. Lee, S. Bonebrake, T. Hollerer, and D. A. Bowman. The role of latency in the validity of ar simulation. In *Proc. IEEE Virtual Reality Conference (VR)*, pages 11–18, 2010.

[5] A. Plopski, Y. Itoh, C. Nitschke, K. Kiyokawa, G. Klinker, and H. Takemura. Corneal-imaging calibration for optical see-through head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics (Proceedings Virtual Reality 2015)*, 21(4):481–490, April 2015.

[6] C. M. Robertson, B. MacIntyre, and B. N. Walker. An evaluation of graphical context as a means for ameliorating the effects of registration error. *IEEE Transactions on Visualization and Computer Graphics*, 15(2):1–10, March/April 2009.

[7] M. Tuceryan, Y. Genc, and N. Navab. Single-point active alignment method (spaam) for optical see-through hmd calibration for augmented reality. *Presence*, 11(3):259–276, 2002.