See Through With The Hands: Seamless Occluded Interactions in Augmented Reality

Wei-Tang Hsu
National Taiwan University
Taipei, Taiwan
b11902033@csie.ntu.edu.tw

Shan-Yuan Teng
National Taiwan University
Taipei, Taiwan
tengshanyuan@csie.ntu.edu.tw

Figure 1: (a) Our view can be occluded in many daily life contexts (e.g., maintaining a motor bike by unscrewing the drain plug, making certain manipulations more challenging and prone to fatigue (b) We propose a wearable device with a wrist-mounted depth camera that captures the hand view which is overlaid in Augmented Reality (AR), enabling the user to see-through the obstruction and manipulate the occluded objects with vision.

Abstract

Augmented Reality (AR) offers the ability to overlay images onto real-world views. See-through views have been explored for spatial understanding in larger scenes, such as through drone-mounted cameras. In this paper, we propose a wrist-mounted device that constructs real-time see-through views for occluded manual interaction without requiring external tracking. The depth image from the camera is transposed to align with the headset view, enabling intuitive body positioning during object manipulation. We combine this with a custom actuated mount that adapts the depth camera's angle to different hand poses. We show that our prototype, despite its simplicity, enables fine manipulation tasks such as aiming at and screwing an occluded nut.

ACM Reference Format:

Wei-Tang Hsu and Shan-Yuan Teng. 2026. See Through With The Hands: Seamless Occluded Interactions in Augmented Reality. In *The 39th Annual ACM Symposium on User Interface Software and Technology (UIST Adjunct '26), November 02–05, 2026, Detroit, MI, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3830397.3842461

UIST Adjunct '26, Detroit, MI, USA

ACM ISBN 979-8-4007-2855-6/26/11
https://doi.org/10.1145/3830397.3842461

1 Introduction

Human vision can perceive object characteristics (e.g., distance, shape) that support effective hand manipulation, including aiming at objects and adjusting grasp to fit the object, i.e., pre-shaping. However, in many daily tasks, the view of interest can be occluded by tight spaces or other constraints, making manipulation more difficult. In these cases, one might need to turn their head or twist their body, or even manipulate objects entirely eyes-free.

Prior work has leveraged Augmented Reality (AR) to render see-through views, or "X-ray vision," that reveal occluded objects. Examples include seeing behind walls, synthesized using cameras mounted on drones to examine a space [2]. Various strategies have also been explored for visualizing see-through views: Sandor et al. displayed see-through vision based on visual saliency to maximize legibility [4], while Avery et al. explored visualizations at different depths to support spatial understanding [1].

For hand manipulation specifically, strategies have been compared in a simulated AR environment [3]. A user study found that a view aligned with the line of sight provides the most intuitive perspective for manipulating objects behind a wall, such as plugging in a cable. However, this work relies on external body and object tracking with predefined 3D models. Inspired by this study, we ask: could a real-time wearable system support hand manipulation without such infrastructure, even outdoors?

We propose a wearable system combining a hand-mounted depth camera and a see-through headset. The depth camera captures an image of the hand as well as the view in front of it. Tracked in 3D, the camera's reconstructed image is transposed in AR to align with the user's eye perspective, creating the illusion that the user can see through blocking objects. Because the hand-mounted camera's limited field of view can be blocked by different hand poses, we deployed a servo-actuated camera mount that automatically adjusts the camera's angle to keep the user's hand tracked. We demonstrate that this wearable depth camera, though simple, provides enough spatial information to perform tasks in a novel way, including aiming at small screws behind a machine.

2 Wearable Prototype

Our prototype consists of an AR headset and a wrist-mounted depth camera with a custom servo-actuated mount (Figure 2).

Augmented Reality headset. We use a Meta Quest 3 and its video passthrough AR function with stereoscopic images.

Depth camera. We use a RealSense D405, which captures depth images from 7–50 cm, coupled with RGB images, with 87° × 58° field of view. In our current prototype, the camera is tracked by a Quest controller; other tracking mechanisms, such as integrating multiple IMUs, could be used to minimize the device's footprint.

Actuated mount. We mount the camera on the palm side of the hand, since this allows it to capture the hand and any held object across most grasping directions. However, this placement can limit the field of view. We therefore developed an actuated mount with a servo motor (MG90S) and 3D-printed parts, controlled via an Arduino microcontroller (Serial), enabling a camera pitch range of 0–90°. The mount straps onto the user's wrist and can be folded when not in use to minimize interference with daily tasks.

Occluded vision reconstruction. When the headset determined that the hand is not in view (built-in hand tracking), our prototype then initiates the see-through view by first unfolding the mount and starting the USB-connected depth camera. The RGB-D captured is then smoothed, filtered, holes-filled, and clamped to a distance of 7–35 cm to reduce clutter; this processing is done in Python. The resulting point cloud is then transmitted over the local network to a Unity program that renders the AR view on Quest. The point cloud is aligned with the AR camera and rendered in half visual transparency, allowing the foreground to remain visible and also preventing the reconstructed occluded scene from self-occluding. The end-to-end frame rate is 23 Hz.

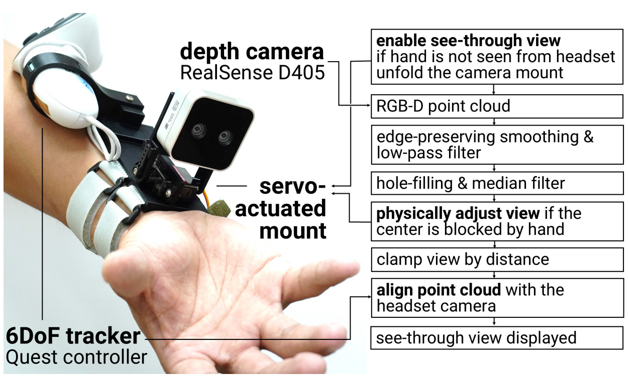

Figure 2: Our hardware prototype and the system diagram.

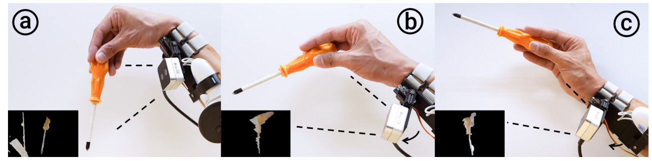

Figure 3: The wrist mount adjusts its view based on the pose of the hand, from (a) almost flat, (b) tilted, to (c) upright.

3 Application Scenarios

We envisioned the device for use in everyday scenarios, even outdoors. The wearable form factor is well suited for more cluttered and tight spaces. We showcase applications depicted in Figure 4.

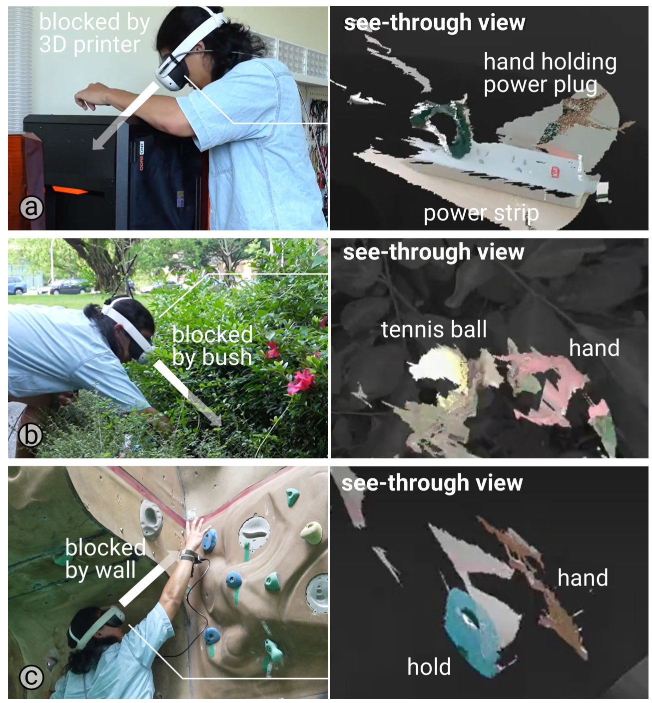

Figure 4: See-through views allow (a) plugging a power plug that is blocked by a 3D printer, (b) picking up a tennis ball in a bush, and (c) finding the hold when bouldering.

References

[1] Benjamin Avery, Christian Sandor, and Bruce H. Thomas. 2009. Improving Spatial Perception for Augmented Reality X-Ray Vision. In *2009 IEEE Virtual Reality Conference*. 79–82. doi:10.1109/VR.2009.4811002 ISSN: 2375-5334.

[2] Okan Erat, Werner Alexander Isop, Denis Kalkofen, and Dieter Schmalstieg. 2018. Drone-Augmented Human Vision: Exocentric Control for Drones Exploring Hidden Areas. *IEEE Transactions on Visualization and Computer Graphics* 24, 4 (April 2018), 1437–1446. doi:10.1109/TVCG.2018.2794058

[3] Klemen Lilija, Henning Pohl, Sebastian Boring, and Kasper Hornbæk. 2019. Augmented Reality Views for Occluded Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. doi:10.1145/3290605.3300676

[4] Christian Sandor, Andrew Cunningham, Arindam Dey, and Ville-Veikko Mattila. 2010. An Augmented Reality X-Ray system based on visual saliency. In *2010 IEEE International Symposium on Mixed and Augmented Reality*. 27–36. doi:10.1109/ISMAR.2010.5643547