Exploring the use of Mobile Devices as a Bridge for Cross-Reality Collaboration

Rishi Vanukuru* Ellen Yi-Luen Do†

ATLAS Institute
University of Colorado Boulder, USA

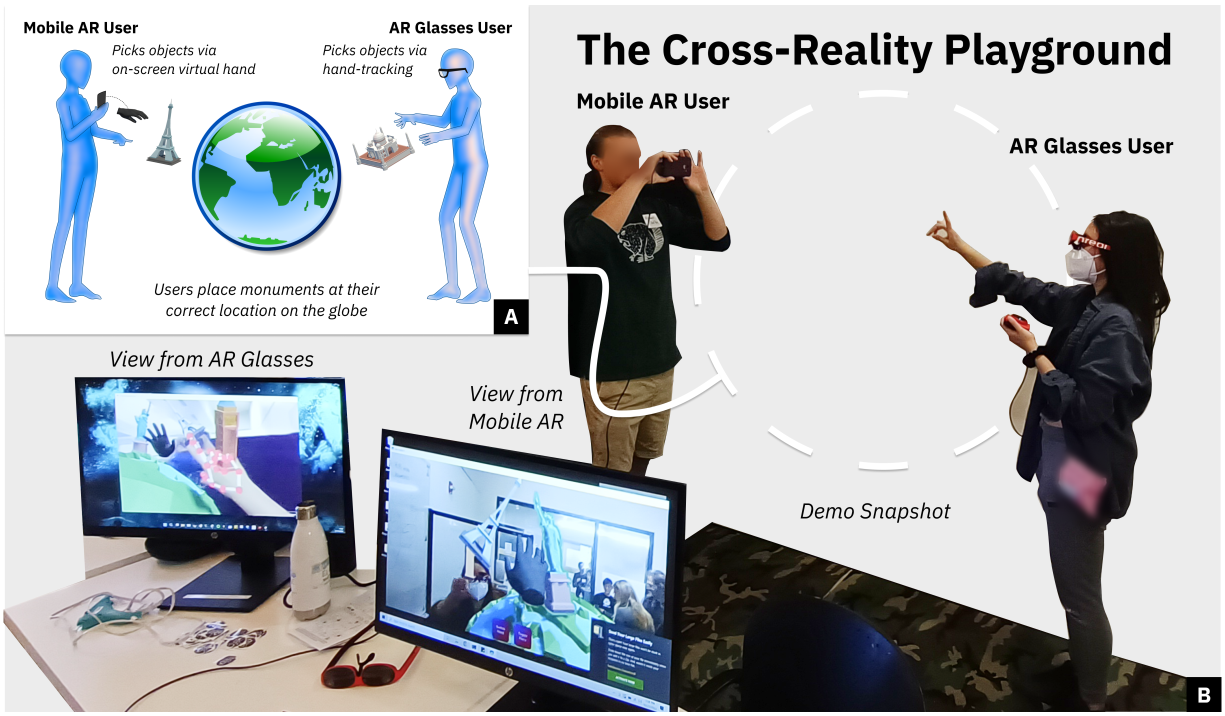

Figure 1: (A) An illustration of the key interactions in the Cross-Reality Playground. Users can pick up 3D models of monuments and place them on the virtual globe, either using a virtual hand (Mobile AR) or their real hands (AR Glasses) (B) A photograph from the demonstration of the prototype, showcasing how two users work together across different devices. The view seen by each user is displayed on screens to the bottom-left.

Abstract

Augmented and Virtual Reality technologies enable powerful forms of spatial interaction with a wide range of digital information. While AR and VR headsets are more affordable today than they have ever been, their interfaces are relatively unfamiliar, and a large majority of people around the world do not yet have access to such devices. Inspired by contemporary research towards cross-reality systems that support interactions between mobile and head-mounted devices, we have been exploring the potential of mobile devices to bridge the gap between spatial collaboration and wider availability. In this paper, we outline the development of a cross-reality collaborative experience centered around mobile phones. Nearly fifty users interacted with the experience over a series of research demo days in our lab. We use the initial insights gained from these demonstrations to discuss potential research directions for bringing spatial computing and cross-reality collaboration to wider audiences in the near future.

*e-mail: rishi.vanukuru@colorado.edu
†e-mail: ellen.do@colorado.edu

Index Terms: Human-centered computing—Mixed/augmented reality; Human-centered computing—Collaborative Interaction; Human-centered computing—Mobile computing

1 Introduction

Collaboration has long been the focus of spatial computing tools based on Augmented and Virtual Reality [6]. Some of the earliest research projects in the field concerned the use of handheld AR displays for 3D object inspection in teams [4], and collaborative scientific visualization using head-mounted displays [7]. Today, AR and VR headsets are growing more available in certain parts of the world, enabling a rich range of collaborative experiences both in-person (such as games on the Tilt Five[1]) and remotely (through social VR applications like Rec Room[2] and VRChat[3]). However, access to these devices is not widespread, and the interfaces are unfamiliar to many. In contrast, AR experiences on mobile devices have introduced millions of people worldwide to the power of spatial computing through tools, games, and social experiences. Researchers have studied how groups of users can work together with mobile AR [3, 8], but the possibilities for interaction become much more in-

[1]Tilt Five: tiltfive.com
[2]Rec Room: recroom.com
[3]VRChat: vrchat.com

teresting when asymmetric collaboration is considered—where users can leverage the advantages of diverse spatial computing devices, and share interfaces to work together. Addressing this need, the field of Cross-Reality interaction research has explored ways in which collaboration between users wearing head-mounted displays, and using handheld devices, can be better supported. The ShareVR project demonstrated how users in VR could collaborate and play games with external users via tracked tablets [2]. Researchers have studied how such asymmetric device configurations affect pair-learning [1], and enable transitional interactions for collaboration [5]. Inspired by these approaches, we have been developing a cross-reality collaborative experience called the *Cross-Reality Playground*, for demonstrations within our research community. While there were pragmatic reasons for this choice—our lab had a limited number of AR glasses, and single-person experiences are insufficient when engaging larger audiences during demo days—we were also interested in observing how people used the asymmetric interaction possibilities to inform future research. In the next section, we describe the design of the *Cross-Reality Playground*, followed by a brief discussion of the findings from a series of demonstrations of the prototype.

2 Developing the Cross-Reality Playground

We developed the *Cross-Reality Playground* to help broaden participation in demonstrations that showcased hand-tracking interactions on AR Glasses (in this case, the Xreal Light[4]). We created an interactive puzzle experience where users were presented with a 3D globe, and a number of world monuments. The goal was to place the monuments at their correct location on the globe. The virtual globe was 0.8 meters wide, and positioned at a height of 1 meter from the ground. The monuments were between 0.2 meters and 0.4 meters across their largest dimension. This was an ideal task for hand-tracking, involving grasping, positioning, and placement, and users could freely walk around the globe. We then designed an equivalent AR experience for mobile devices, but replaced hand-tracking with on-screen touch interactions. Mobile users would see a virtual hand at a fixed distance in front of the screen. They could move their phone and virtual hand close to the monuments, touch the screen to pick up the object, and let go when they wished to place it. To address the issue of view-locking in screen-based AR interfaces (where the orientation of picked objects is often fixed relative to the phone), we incorporated an interaction technique where once an object has been picked by touching the screen, the virtual hand can be twisted by moving the user's finger around the screen (illustrated in Fig. 2). This was included to help mobile users place monuments correctly without having to maneuver the phone in odd angles.

Both the AR Glass and Mobile Phone experiences were created as a single application using Unity 2022.2[5] and the respective device AR SDKs (AR Core[6] and NRSDK[7]). The application adapts to the device it runs on. We then incorporated the ability for multiple users to operate on the same set of monuments, by networking the movement of players and objects using the Photon Fusion SDK[8] for Unity. The position and orientation of the globe is synchronized using a visual marker, after which the position and orientation of each object is transmitted across all devices, based on the user who is moving it. Users can also hand-off objects between each other across AR glasses and mobile devices, and the latency of the system is low enough (between 50 ms and 200 ms) for this to be a viable interaction. AR Glass users can see the virtual hand attached to mobile devices, while mobile users only sees the monuments moving as if picked up directly by the AR glass users. We also used a PC client to reset the position of the monuments when a new

[4]Xreal Light: xreal.com
[5]Unity: unity.com
[6]Google AR Core developers.google.com/ar
[7]XReal SDK developer.xreal.com
[8]Photon Fusion: photonengine.com

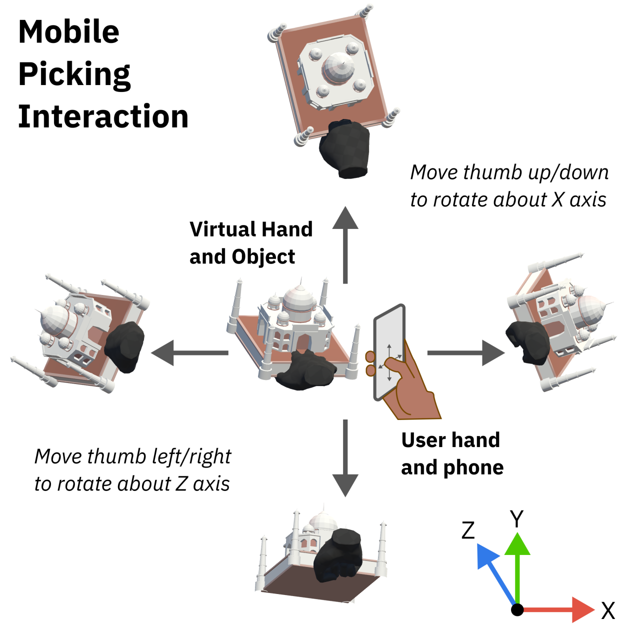

Figure 2: An illustration of the mobile interaction technique to avoid view-locking. Once an object is picked, users can slide their finger on the screen to rotate their virtual hand. This helps them to inspect and place the objects with a greater degree of control, similar to wrist-based manipulation in the real-world.

set of users began the experience. For the purposes of the demo, we ensured that two users, each with either a mobile device or AR Glasses, would interact with each other at the same time. However, during internal testing we verified that up to six users can work on the same set of objects, with any combination of mobile phone or AR glass. Fig. 1A illustrates the key functions of the prototype.

3 Discussion

We demonstrated the *Cross-Reality Playground* across two major demo days, and numerous in-lab sessions over the course of four months. We estimate about 50 users actively tried the demo (Fig. 1 B), and many more observed their friends and family engage with it from outside AR. At any point, two people were using the demo, one via AR glasses, and one via Mobile AR. Thus, the main interaction configurations were two users together, or one user working with one researcher. The purpose of this demonstration was to provide a glimpse into AR technology and collaboration possibilities, and it was not designed to be a formal evaluation of the prototype. These informal engagements helped us include a much wider range of participants, from young children to their grandparents, with different levels of prior experience with AR. We were able to observe users' reactions in a more natural setting, and this encouraged more open discussions about their thoughts and feedback.

For many, the *Cross-Reality Playground* was their first experience with head-worn AR. In past demonstrations of AR glasses, we noticed that users were excited by the novel interface, but continued to think of it as something from the distant future. In contrast, being able to perform the same functional task—picking and placing monuments on the globe—using both the AR glasses and mobile devices helped ground their understanding of head-worn AR. While direct interactions using hand-tracking felt more natural than operating the virtual hand via mobile screens, users indicated that they could envision how the capabilities of phones today might translate to headsets in the future. This experience also seemed to elevate their opinion of mobile AR as being useful and comparable to head-worn alternatives. For those who had prior experience with AR/VR, the prototype demonstrated a way in which friends without AR glasses could interact more actively, in contrast to existing methods such as watching the first-person view of AR games cast onto screen, or participating in a non-spatial manner using laptops. These users also drew upon their experience with AR to compare interactions across modalities. One user mentioned how they felt the mobile device was at a disadvantage because there was direct control of only one hand, while users with AR Glasses were able to use two hands at the same time, to perform gestures such as pointing while also moving objects. The small size of the phone screen, and the fatigue induced when holding the phone for longer than five minutes, were common issues that users wanted future versions of the experience to improve upon. That being said, most users were able to appreciate how the option of using mobile devices could help scale such experiences to larger groups of people.

The familiarity of mobile devices presented new configurations for social interaction, and helped include users who might otherwise not be able to work with AR glasses. The ways in which families interacted with the *Cross-Reality Playground* provided some of the strongest examples for this. Parents who were initially wary about the isolating effect of AR/VR headsets remarked that they would be more comfortable with their children using headsets if they could also participate using mobile devices. Grandparents who could not place the AR glasses over prescription lenses, or who were unable to make the hand gestures required due to limited range of motion, could participate with their grandchildren using mobile devices instead. Younger children, whose hands were too small to be accurately tracked by the AR glasses, could view the globe and monuments via phones, and direct their taller siblings to locations at the bottom of the globe which were easier for them to see. Taken together, the simple inclusion of a mobile device, and the opportunity to experience interactions in a cross-reality setting, can potentially help a wider range of users imagine a future where spatial computing is more integrated into everyday life.

4 Future Work

Encouraged by the response to the *Cross-Reality Playground*, we plan to more formally investigate cross-reality interfaces centered around the mobile phone. The demo days have shown how mobile devices can act as a bridge interface, not just between different modalities of AR, but also between users' notions of screen-based computing today, and immersive computing in the near-future. While the *Cross-Reality Playground* supports co-located interactions where digital objects are the focus, mobile devices are likely to help in a broad range of scenarios, such as remote communication in social VR, or collaborative 3D content creation. We envision mobile devices as being the connecting link between different device ecosystems (desktop and spatial), group configurations (co-located and remote), and group purposes (human interaction, 3D content manipulation) (Fig. 3 top). Building upon this, we have started expanding the prototype to include interactions on laptops and in VR, both in-person and remotely (Fig. 3 bottom). By maximizing the potential for spatial interactions using mobile devices, and building cross-reality ecosystems with stronger connections between mobile devices, PCs, and immersive setups, we can better support meaningful spatial collaboration between people, no matter where they are located, or what devices they have access to.

Acknowledgments

We would like to thank the members of the ACME lab for their help in testing, documenting, and demonstrating the prototypes, and the staff at ATLAS for organizing the demo days. This project was supported in part by a grant from Ericsson Research, and by the National Science Foundation under Grant No. IIS-2040489.

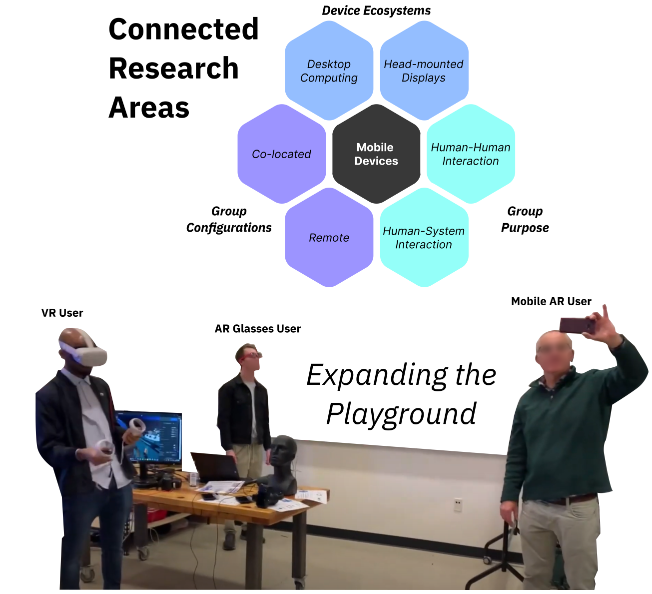

Figure 3: (Top) A representation of the broader research domain for future work, with mobile devices connecting different device ecosystems, group configurations, and group purposes. (Bottom) A photograph of an updated version of the prototype, with the addition of VR interactions.

References

[1] T. Drey, P. Albus, S. der Kinderen, M. Milo, T. Segschneider, L. Chanzab, M. Rietzler, T. Seufert, and E. Rukzio. Towards collaborative learning in virtual reality: A comparison of co-located symmetric and asymmetric pair-learning. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22. Association for Computing Machinery, New York, NY, USA, 2022. doi: 10.1145/3491102.3517641

[2] J. Gugenheimer, E. Stemasov, J. Frommel, and E. Rukzio. ShareVR: Enabling Co-Located Experiences for Virtual Reality between HMD and Non-HMD Users. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pp. 4021–4033. Association for Computing Machinery, New York, NY, USA, May 2017. doi: 10.1145/3025453.3025683

[3] L. P. Poretski, J. Lanir, R. Margalit, and O. Arazy. Physicality As an Anchor for Coordination: Examining Collocated Collaboration in Physical and Mobile Augmented Reality Settings. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2):470:1–470:29, Oct. 2021. doi: 10.1145/3479857

[4] J. Rekimoto. Transvision: A hand-held augmented reality system for collaborative design. In *Proceeding of Virtual Systems and Multimedia*, vol. 96, pp. 18–20, 1996.

[5] J.-H. Schröder, D. Schacht, N. Peper, A. M. Hamurculu, and H.-C. Jetter. Collaborating across realities: Analytical lenses for understanding dyadic collaboration in transitional interfaces. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23. Association for Computing Machinery, New York, NY, USA, 2023. doi: 10.1145/3544548.3580879

[6] M. Sereno, X. Wang, L. Besançon, M. J. McGuffin, and T. Isenberg. Collaborative Work in Augmented Reality: A Survey. *IEEE Transactions on Visualization and Computer Graphics*, 28(6):2530–2549, June 2022. Conference Name: IEEE Transactions on Visualization and Computer Graphics. doi: 10.1109/TVCG.2020.3032761

[7] Z. Szalavári, D. Schmalstieg, A. Fuhrmann, and M. Gervautz. "Studierstube": An environment for collaboration in augmented reality. *Virtual Reality*, 3:37–48, 1998.

[8] T. Wells and S. Houben. CollabAR - Investigating the Mediating Role of Mobile AR Interfaces on Co-Located Group Collaboration. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, pp. 1–13. Association for Computing Machinery, New York, NY, USA, Apr. 2020. doi: 10.1145/3313831.3376541