# TeleGate: Immersive Multi-User Collaboration for Mixed Reality 360° Video

Hyejin Kim* Jacob Young* Daniel Medeiros† Stephen Thompson* Taehyun Rhee*

Figure 1: Users collaborating using the TeleGate system. A shared immersive display is used to prevent crucial facial and bodily cues from being obscured by VR headsets. Virtual objects are realistically blended into live or pre-recorded 360° video using real-time lighting estimation, allowing this display to act as a "window" into remote spaces that can be manipulated using handheld controllers.


## Abstract

When collaborating on virtual content within 360° mixed reality environments it is often desirable for collaborators to fully immerse themselves within the task space, usually by means of a head-mounted display. However, these socially isolate any co-located collaborators, removing the ability to communicate through important gestural, facial, and body language cues. We present TeleGate, a system that instead utilises a shared immersive display to allow collaboration within remote environments between an arbitrary number of users, keeping collaborators visible while allowing immersive and interactive collaboration within remote environments.


**Index Terms:** Human-centered computing—Mixed / augmented reality; Human-centered computing—Collaborative interaction;

## 1 Introduction

In an increasingly connected world, people often find it desirable to collaborate together within remote mixed reality spaces to accomplish some shared goal. Engineers may wish to survey disaster areas without endangering themselves, furniture stores may wish to show customers what their products would look like in their homes, or builders may wish to overlay blueprints on their construction sites.

Many systems [8, 13] propose virtual reality as the solution, allowing collaborating users to fully immerse themselves within a shared space. This can often induce a significant sense of spatial presence within the remote environment [10] due to the ability to obtain novel viewpoints [5, 13], but does so at the expense of isolating users from the real world and any co-located collaborators. This can obscure important gestures and body language that are crucial in understanding a statement's intent [1, 11], making coordination between collaborators more difficult [2] and subsequently task completion time [3, 4, 6]. The wisdom of sharing HMDs in a post-COVID-19 world must also be considered in public spaces such as schools and department stores where users are unlikely to supply their own headsets.

## 2 System Overview

We present TeleGate, an immersive collaborative system that places co-located users in front of a large semi-encompassing immersive display, allowing for more intuitive and efficient communication through gestures, gaze, and other body language [3, 11]. Multiple users may come together and interact within a remote environment captured live by a 360° camera, placing and interacting with virtual objects which are dynamically lit based on the conditions detected within the 360° video using MR360 [7]. As users' view of the real world is no longer occluded by an HMD they are now free to communicate as they would for real-world tasks, though with the benefit of introducing realistic virtual content and real-life remote spaces. The immersive display thus acts as a kind of "window" into the mixed reality content through which an arbitrary number of viewers can peer, allowing scalability without increasing equipment costs and avoiding hygiene concerns in a post-COVID-19 world.

### 2.1 Camera Control

To enhance the illusion of the immersive display being a window into the virtual world, the properties of the virtual camera are dynamically calculated and adjusted based on the display's real-world dimensions. An HTC Vive tracker is attached to the corners of the display, as shown in Fig. 2, and the horizontal field of view of the virtual camera is calculated based on the display's aspect ratio and its distance from the world origin. This distance is also used to set the position of the near clipping plane so that virtual objects are removed from view at a distance that users expect.

The direction of this virtual camera can be controlled by users to increase their spatial presence within [5, 13] and spatial understanding of [9] the presented 360° environment. Reorientation is performed through a "scene in hand" technique [12], where any rotations of an HTC Vive controller while its grip buttons are depressed

*Hyejin Kim, Jacob Young, Stephen Thompson, and Taehyun Rhee are with the Computational Media Innovation Centre at Victoria University of Wellington. e-mail: {betty.kim, jacob.young, stephen.thompson, taehyun.rhee}@vuw.ac.nz

†Daniel Medeiros is with the University of Glasgow. e-mail: Daniel.PiresdeSaMedeiros@glasgow.ac.uk

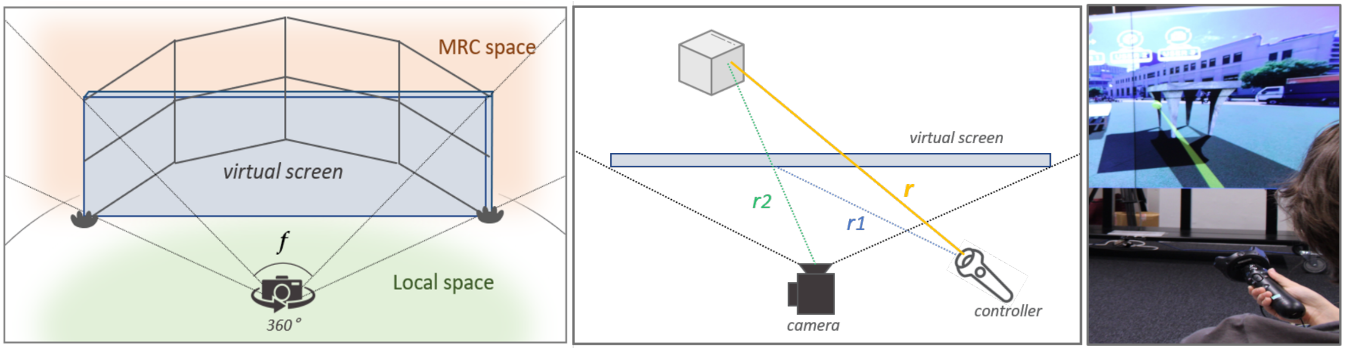


Figure 2: (Left): The parameters of the render camera are calculated from the dimensions of the real-world display. (Centre and Right): The user's pointer is redirected to match their perspective and the intended pointing direction by raycasting from their real-world position.

are reflected in the rotation of the virtual camera.

### 2.2 Object Control

Each user's handheld controller can also be used to manipulate and interact with virtual objects augmented onto the 360° video. A ray is cast from their controller's tip (r1 in Fig. 2) and the point at which this ray intersects with the display is calculated based on its orientation and position relative to the display's corners, then offset by a ray from the camera to this point on the display (r2). The resulting ray is rendered, allowing users to point to areas of interest.

When a virtual object is pointed to, it can be picked up by pulling the trigger and moved by pointing to the desired destination. Pressing the touchpad also allows the carried object to be scaled or rotated by moving the controller vertically or horizontally, respectively.

## 3 Evaluation

A preliminary user study was conducted using a within-subjects design with 22 participants to compare the spatial and social presence induced by TeleGate compared to a more conventional HMD-based system. Wilcoxon signed-rank tests ($\alpha = 0.05$) revealed an expected significant decrease in spatial presence ($p < 0.0001$, $Z = -4.076$) and no significant difference in social presence ($p = 0.071$, $Z = -1.802$), though an increase emerges when the question "(I—My partner) was easily distracted from (my partner—me) when other things were going on)" was excluded ($p = 0.038$, $Z = -2.09$) which suggests HMDs as useful for focusing on conversation, even if it potentially becomes constrained in the process.

## 4 Conclusion

We presented TeleGate, an immersive collaborative system that avoids the pitfalls of VR teleconferencing by leaving collaborators visible to each other through use of a large immersive display. This allows intuitive conversational cues such as gestures, facial expressions, and body language to be used, and also avoids hygiene concerns of public headset use in a post-COVID-19 world. The display acts as a window to a remote 360° space with virtual objects seamlessly blended in using real-time lighting estimation which can be intuitively manipulated using handheld controllers.

## Acknowledgments

The authors wish to thank the other members of the Computational Media Innovation Centre for their support throughout this project.

## References

[1] N. Flor. Side-by-side collaboration: a case study. *Internation Journal of Human-Computer Studies*, 49(3):201–222, 1998.

[2] S. R. Fussell, R. E. Kraut, and J. Siegel. Coordination of Communication: Effects of Shared Visual Context on Collaborative Work. In *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work*, pp. 21–30, 2000. doi: 10.1145/358916.358947

[3] S. R. Fussell, L. D. Setlock, J. Yang, J. Ou, E. Mauer, and A. D. I. Kramer. Gestures over Video Streams to Support Remote Collaboration on Physical Tasks. *Human-Computer Interaction*, 19(3):273–309, Sept. 2004. doi: 10.1207/s15327051hci1903

[4] S. Gauglitz, B. Nuernberger, M. Turk, and T. Höllerer. World-stabilized Annotations and Virtual Scene Navigation for Remote Collaboration. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, UIST '14, pp. 449–459. ACM, New York, NY, USA, 2014. doi: 10.1145/2642918.2647372

[5] H. Jo and S. Hwang. Chili: Viewpoint Control and On-Video Drawing for Mobile Video Calls. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, pp. 1425–1430. 2013.

[6] S. Kim, G. Lee, N. Sakata, and M. Billinghurst. Improving Co-Presence with Augmented Visual Communication Cues for Sharing Experience through Video Conference. In *ISMAR 2014 - IEEE International Symposium on Mixed and Augmented Reality - Science and Technology 2014, Proceedings*, pp. 83–92, 2014. doi: 10.1109/ISMAR.2014.6948412

[7] T. Rhee, L. Petikam, B. Allen, and A. Chalmers. MR360: Mixed Reality Rendering for 360° Panoramic Videos. *IEEE Transactions on Visualization and Computer Graphics*, 23(4):1379–1388, 4 2017. doi: 10.1109/TVCG.2017.2657178

[8] T. Rhee, S. Thompson, D. Medeiros, R. dos Anjos, and A. Chalmers. Augmented Virtual Teleportation for High-Fidelity Telecollaboration. *IEEE Transactions on Visualization and Computer Graphics*, 26(5):1923–1933, 2020. doi: 10.1109/TVCG.2020.2973065

[9] M. A. Schnabel and T. Kvan. Spatial Understanding in Immersive Virtual Environments. *International Journal of Architectural Computing*, 1(4):435–448, 2003.

[10] Y. Shu, Y.-Z. Huang, S.-H. Chang, and M.-Y. Chen. Do Virtual Reality Head-Mounted Displays Make a Difference? A Comparison of Presence and Self-Efficacy between Head-Mounted Displays and Desktop Computer-Facilitated Virtual Environments. *Virtual Reality*, 23(4):437–446, 2019.

[11] P. Taylor, R. E. Kraut, S. R. Fussell, and J. Siegel. Visual Information as a Conversational Resource in Collaborative Physical Tasks. *Human-Computer Interaction*, 0024(917808986):13–49, 2009. doi: 10.1207/S15327051HCI1812

[12] C. Ware and S. Osborne. Exploration and virtual camera control in virtual three dimensional environments. *SIGGRAPH Comput. Graph.*, 24(2):175–183, Feb. 1990. doi: 10.1145/91394.91442

[13] J. Young, T. Langlotz, M. Cook, S. Mills, and H. Regenbrecht. Immersive Telepresence and Remote Collaboration using Mobile and Wearable Devices. *IEEE Transactions on Visualization and Computer Graphics*, 25(5):1908–1918, 2019.