VR-Enabled Telepresence as a Bridge for People, Environments, and Experiences

Brennan Jones
Department of Computer Science
University of Calgary
bdgjones@ucalgary.ca

Yaying Zhang
Microsoft Vancouver
yayzhang@microsoft.com

Priscilla N. Y. Wong
University College London & Microsoft Research Cambridge
ngoi.wong.13@ucl.ac.uk

Sean Rintel
Microsoft Research Cambridge
serintel@microsoft.com

Yasamin Heshmat
School of Interactive Arts and Technology
Simon Fraser University
yheshmat@sfu.ca

Social VR: A New Medium for Remote Communication and Collaboration Workshop at ACM CHI 2020, Honolulu, HI, USA
https://www.socialvr-ws.com/

Abstract

Rather than thinking of Virtual Reality (VR) as enabling geo-distant people to go to a common place, we have a more expansive view of VR as a bridge between people, environments, and experiences. We are interested in using VR and the related technology of Augmented Reality (AR), as tools to (1) bring a distant part of the physical world to someone, and (2) bring someone to a distant part of the physical world. This workshop paper discusses two design explorations we conducted to explore these subtle but important distinctions. We first discuss a unidirectional VR robotic telepresence interface that immerses a remote user in a distant environment, which was investigated for its value in outdoor exploration. We then discuss a bi-directional asymmetric VR-AR system that adds a photorealistic avatar for the remote user onto VR robotic telepresence. The avatar is viewed in third person by the local user as superimposed over the robot, and viewed in first person by the remote user. While the underlying technologies still need much development, we see promise in the 'VR as bridge' concept as a way to open up the design space to a critical need for flexible, diverse, and inclusive user needs.

Figure 1: Beam telepresence robot with 360° camera (circled) attached. This camera streams to a remote user who can view the robot's surroundings in VR.

Figure 2: The remote user operating the telepresence robot using a VR interface (top), immersing themself in the faraway location (bottom).

Introduction

Virtual Reality (VR) research over the years has focused on many different avenues, including heightening immersion (e.g., [3,17]), creating compelling virtual worlds (e.g., [23]), and enabling rich social experiences through such virtual worlds (e.g., [18,19]). VR technologies provide connection opportunities to many people who would otherwise not have them. For example, virtual worlds such as those provided by commercial apps like Rec Room [24], VRChat [8], and AltspaceVR [1] bring people from around the world, from many different backgrounds, together in one shared virtual space, where they are able to portray themselves however they like, interact with whomever they want, and explore wherever they want. This can be empowering, especially for people who are motor-impaired or face other health challenges that make them unable to move around or get as much benefit from the real world, or for people who struggle with their own identities. We are interested in using VR and the related technology of Augmented Reality (AR) to push this empowerment even further, to enable flexible *bridges* for connecting people, environments, and experiences in both the virtual and physical world.

VR-Enabled Telepresence

Telepresence is the feeling of being in a physical space other than one's own [15]. It differs from *virtual presence*, which is the feeling of being in a virtual space [21]. Presence (both tele- and virtual) can take on different forms, including *spatial presence* (feeling like you are *in* the space [11,15]), *social presence* (feeling like you are *with* someone [2,14]), *co-presence* (feeling like you are *in* a space *with* someone [7]), and *self-presence* (feeling like your 'self' or embodiment in the space is *indeed* yourself [4,5,11]).

Telepresence can be enabled by many different technologies, including traditional video communication (e.g., Skype), immersive video (e.g., from 360° cameras [6,12,13,22]), drones as embodiments [20] and camera sources [9,20], and telepresence robots (e.g., [25,26]). Telepresence robots are mobile robots that have a screen at the top for a talking-heads video-chat interface and wheels at the bottom. They provide a 'video-chat-on-wheels' experience. While they have not yet been widely adopted for use in the home, they are increasingly visible in workplaces (e.g., [27]).

VR-enabled telepresence is a tool to bring the real world to people and people to the real world, while preserving as many of the benefits of VR in virtual worlds. Our goals are to enable (1) the ability to freely roam the environment, (2) the feeling that the environment belongs to telepresent users as much it belongs to those who are physically present, and (3) the ability to freely express oneself in the environment. This paper provides a brief overview of two design explorations which pursue these goals.

Concept 1: VR Telepresence Robot Interface

We first explored VR-Enabled telepresence [6] by attaching a 360° camera to the top of a Beam telepresence robot (Figure 1) and live-streaming the 360° video to a mobile VR headset, allowing a remote user to drive the robot while immersed in a 360° view of the robot's surroundings (Figure 2).

We took this robot outside into a park and had remote participants accompany a local participant in a search for geocaches. Participants enjoyed the immersive view of the environment and the ease of looking around naturally, as opposed to having to reorient the robot to

Figure 3: VROOM: the remote user's avatar superimposed over the telepresence robot. Viewable through the local user's HoloLens.

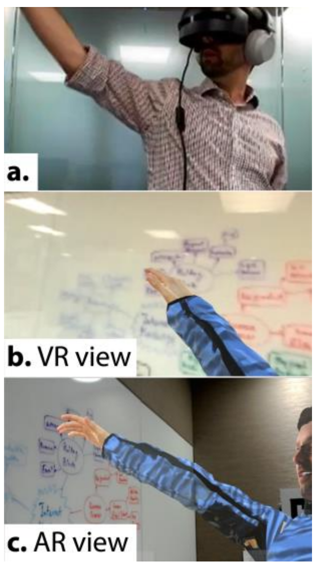

Figure 4: VROOM: the remote user gesturing at a white board. **a.** Remote (VR) user's action. **b.** Remote (VR) user's view. **c.** Local (AR) user's view.

look around. Some participants were so invested in the immersive view of the outdoor environment that they temporarily forgot that they were in an office using a VR system.

That said, participants found it more difficult to control the robot in the immersive view. This could have been due to a variety of factors, including the position of the 360° camera, the lack of driving-guidance markers (which are available in standard telepresence-robot interfaces) in our interface, or the fact that it was easy for the remote user to lose their orientation in the VR view, and thus their orientation with respect to the robot. Furthermore, it was not always easy for remote participants to refer to or gesture toward things in the environment, even though they had a full immersive view into it. While the local participant could see their remote partners on the screen of the robot, they were not always able to understand the remote participant's pointing gestures in their own space or match them to the outdoor space.

This VR setup brought the distant location closer to the remote user, making the space feel a little bit more like 'theirs' to explore freely (within the physical limits of the telepresence robot). It thus addressed goals of free movement and ownership outlined in the introduction. However, we wanted to improve this experience and additionally address the goal of free expression.

Concept 2: VROOM

In addition to an immersive 360° view of the environment from a telepresence robot's viewpoint, we designed *VROOM* (*Virtual Robot Overlay for Online Meetings*) [10] to provide the remote user with an avatar in the remote space (Figure 3). This avatar can be seen in third-person by the local user when they put on a HoloLens AR headset, superimposed on the robot wherever it goes. A first-person view of the avatar body can be seen by the remote user in the VR view (Figure 4b), creating identification between the user's physical body and the avatar body, and thus reinforcing the perception of *self-presence* in the distant space.

Through VROOM, the remote user controls not only the robot, but also their own avatar. The local user sees the remote's avatar's head orientation match that of the remote user, providing an indication of what the remote user is looking at in the activity space. The VR system's motion controllers control the avatar's arms, so the remote user is able to make basic arm gestures (like waving and pointing) which can be seen by the local user (Figure 4). The avatar's mouth moves when the remote user speaks, providing an attention cue for the local user. Finally, the avatar goes into a walking animation when the remote user drives the robot. This provides a social cue for the local user to recognize that the robot is an embodiment of a remote person and not simply a machine. It also provides an embodiment cue for the remote user, who sees the walk animation in first person, and thus creates an association of robot movement with their movement.

We ran a mixed-methods study to understand how this bi-directional asymmetric VR-AR system affects people's collaborative and social interactions, how remote users make use of the VR view and express themselves through the avatar, and how local users understand and perceive the remote's avatar. Pairs of participants completed a game in which they looked around for clues and played a word-guessing game with each other. For each pair, one participant was local in

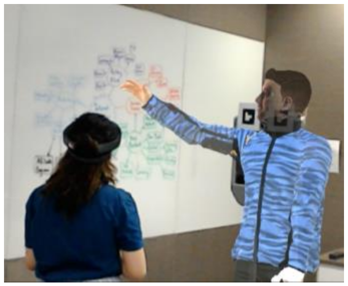

Figure 5: A local (left) and remote (right) user using VROOM to collaborate on a whiteboard.

ACKNOWLEDGEMENTS

We thank our colleagues who have helped us with the system design and study of the initial VR telepresence-robot interface: Xiong Xiaoxuan (Anty), Carman Neustaedter, Anthony Tang, Bernhard E. Riecke, and Lillian Yang. Additionally, we thank our Microsoft colleagues who have helped us tremendously in developing the VROOM prototype: James Scott, Xu Cao, He Huang, Minnie Liu, Zhao Jun, Matthew Gan, and Leon Lu.

the activity space, while the other was remote, in a separate room, and drove a telepresence robot in the activity space. Participants completed two instances of the activity: one using VROOM, and the other using standard robotic telepresence.

Our preliminary findings suggest that local users were able to understand their remote partner's gestures and gaze attention well through the remote's avatar. Some remote participants also mentioned that seeing an avatar body that they had control of in first-person through the VR view in relation to the local space, and knowing that their local partners could see the same avatar in the same spot, made them feel more like the avatar was an embodiment of *themself* in the local space. Many local participants identified the remote's avatar as being *them*.

However, after seeing what their avatar looked like to the local participant, some remote participants reported that they did not identify with their avatar. One reason for this could be that, due to technical reasons, the research team had to make the remote user's avatar rather than allowing the participant to make and customize their avatar before participation. Furthermore, we attempted to make each avatar a photorealistic representation of the person using a photo that the participant sent us. This could have triggered an 'uncanny valley' [16] reaction. Allowing users to make their own avatars, or even using cartoon-style avatars (e.g., Xbox avatars), may enable users to identify with their avatars more closely.

Conclusion

Our design explorations attempt to bring VR experiences into the physical world, and to bring one part of the physical world to another through VR. They attempt to deliver many of the same benefits of virtual worlds by blending VR experiences with the real world.

We believe that the 'VR as bridge' concept opens the Social VR design space up to address critical user needs for flexibility, diversity, and inclusion that go beyond mere visual representation. People should not be restricted from participation in work or social events because travel is difficult for reasons such as disability, travel/visa restrictions, or lack of transportation infrastructure. While pure VR systems enable a single place for geo-distant people to connect, which is clearly very valuable, a lack of access to the real world may have the unintended consequences of impoverishing people's breadth of experiences, excluding and disconnecting people from 'really' being part of remote teams, social groups, and events. If a colleague or friend is never really 'here' in local physical space, but only 'there' in virtual space, we may still be privileging the local. There will always be people who feel uncomfortable in purely VR spaces, and it is vital that we connect all people, especially if, as many in this field hope, we are to increase the sustainability of global work and personal life.

References

[1] AltspaceVR. Be there, together. *AltspaceVR Inc*. Retrieved February 18, 2020 from https://altvr.com/.

[2] Frank Biocca. 1997. The Cyborg's Dilemma: Progressive Embodiment in Virtual Environments. *Journal of Computer-Mediated Communication* 3, 2.

[3] Doug A. Bowman and Ryan P. McMahan. 2007. Virtual Reality: How Much Immersion Is Enough? *Computer* 40, 7: 36–43.

[4] Paul W. Eastwick and Wendi L. Gardner. 2009. Is it a game? Evidence for social influence in the virtual world. *Social Influence* 4, 1: 18–32.

[5] James Paul Gee. 2008. Video Games and Embodiment. *Games and Culture* 3, 3–4: 253–263.

[6] Yasamin Heshmat, Brennan Jones, Xiaoxuan Xiong, et al. 2018. Geocaching with a Beam: Shared Outdoor Activities Through a Telepresence Robot with 360 Degree Viewing. *Proceedings of the 2018 ACM Conference on Human Factors in Computing Systems*, ACM, 359:1–359:13.

[7] Wijnand A. IJsselsteijn, Jonathan Freeman, and Huib De Ridder. 2001. *Presence: Where are we?* Mary Ann Liebert, Inc.

[8] VRChat Inc. VRChat. Retrieved February 18, 2020 from https://www.vrchat.com/.

[9] Brennan Jones, Kody Dillman, Richard Tang, et al. 2016. Elevating Communication, Collaboration, and Shared Experiences in Mobile Video Through Drones. *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*, ACM, 1123–1135.

[10] Brennan Jones, Yaying Zhang, Priscilla N. Y. Wong, and Sean Rintel. 2020. VROOM: Virtual Robot Overlay for Online Meetings. *Extended Abstracts of the 2020 ACM Conference on Human Factors in Computing Systems*, ACM.

[11] Kwan Min Lee. 2004. Presence, Explicated. *Communication Theory* 14, 1: 27–50.

[12] Zhengqing Li, Shio Miyafuji, Toshiki Sato, Hideki Koike, Naomi Yamashita, and Hideaki Kuzuoka. 2018. How Display Shapes Affect 360-Degree Panoramic Video Communication. *Proceedings of the 2018 Designing Interactive Systems Conference*, ACM, 845–856.

[13] Zhengqing Li, Shio Miyafuji, Erwin Wu, Hideaki Kuzuoka, Naomi Yamashita, and Hideki Koike. 2019. OmniGlobe: An Interactive I/O System For Symmetric 360-Degree Video Communication. *Proceedings of the 2019 on Designing Interactive Systems Conference*, ACM, 1427–1438.

[14] Matthew Lombard and Theresa Ditton. 1997. At the Heart of It All: The Concept of Presence. *Journal of Computer-Mediated Communication* 3, 2.

[15] Marvin Minsky. 1980. Telepresence.

[16] Masahiro Mori. 1970. The uncanny valley. *Energy* 7, 4: 33–35.

[17] Randy Pausch, Dennis Proffitt, and George Williams. 1997. Quantifying immersion in virtual reality. *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, ACM Press/Addison-Wesley Publishing Co., 13–18.

[18] Ralph Schroeder. 1996. Possible worlds: the social dynamic of virtual reality technology. Westview Press Boulder, CO.

[19] Ralph Schroeder. 2012. The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments. Springer Science & Business Media.

[20] Hanieh Shakeri and Carman Neustaedter. 2019. Teledrone: Shared Outdoor Exploration Using Telepresence Drones. *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing*, Association for Computing Machinery, 367–371.

[21] Thomas B. Sheridan. 1992. Musings on telepresence and virtual presence. *Presence: Teleoperators & Virtual Environments* 1, 1: 120–126.

[22] Anthony Tang, Omid Fakourfar, Carman Neustaedter, and Scott Bateman. 2017. Collaboration with 360° Videochat: Challenges and Opportunities. *Proceedings of the 2017*

Conference on Designing Interactive Systems, ACM, 1327–1339.

[23] Mary C. Whitton. 2003. Making virtual environments compelling. *Communications of the ACM* 46, 7: 40–47.

[24] Rec Room. *Rec Room*. Retrieved February 18, 2020 from https://recroom.com.

[25] BEAM - From Here to Anywhere. *Beam*. Retrieved December 17, 2019 from https://suitabletech.com/.

[26] Double Robotics - Telepresence Robot for Telecommuters. Retrieved December 17, 2019 from https://www.doublerobotics.com/.

[27] R&D Lab - Microsoft Research. Retrieved December 21, 2019 from https://suitabletech.wistia.com/medias/eg82wq4x5b?embedType=iframe&seo=false&videoWidth=560?wtime=0.