Prototyping Mixed Reality Large Screen Mobile Telepresence Robots

Ian Gonsher
Brown University
ian_gonsher@brown.edu

Yuxin Han
Rhode Island School of Design
yhan02@alumni.risd.edu

Karthik Desingh*
University of Washington
kdesingh@cs.washington.edu

Aaron Gokaslan*
Cornell University
aarongokaslan@gmail.com

***Abstract*—In this paper, we discuss the development of prototypes for remote communication and collaboration that apply large, mobile screens to immersive and embodied telepresence experiences. We discuss several prototypes being developed by our design research team. The first, the Large Screen Mobile Telepresence Robot gives users the ability to embody the robot at nearly full human scale, and to move about in a remote environment. The second, the Mixed Reality Window, uses both forward facing and rear facing cameras and screens, at human scale, to create a passthrough effect, which can be augmented and overlaid with additional visual content, while also offering both forward facing and rear facing perspectives to remote users in non-contiguous spaces. We also discuss prior work in the development of Situated Robots, which employ projection onto the build environment, as well as speculative designs for future work that integrate mobile large screens into the built environment.**

***Index Terms*—mixed reality, telepresence robots, telecommunication**

I. INTRODUCTION

As more and more of our face to face interactions with each other are mediated by technology, the development of new design modalities for videoconferencing, and for collaboration at a distance more generally, is becoming increasingly urgent. The COVID19 pandemic has accelerated these trends, making remote collaboration and communication even more important. Over the past few years, Zoom, Skype, and other videoconferencing apps have become fully integrated into our personal and professional habits. And yet, not much attention has been given to the design of these experiences beyond the conventional formula of the laptop or desktop screen and camera. The constraints these conventions place on the telepresence experience limits the quality of the interpersonal interactions users have with each other, and is ripe for critique and redesign.

We have seen many social telepresence robots over the years [1]–[4]. Often referred to as Mobile telepresence robots (MTR) that create a connection between two different environments to facilitate social interactions. MTRs are developed for environments such as office, schools, research and healthcare [5], [6]. Additionally, their potential use case is to assist people with physical or cognitive disabilities, i.e., they can be co-located with the MTR which will be accessed by their family member or healthcare professional. In addition to providing the basic videoconference mode of communication, these robots come with services for older adults such as reminding them of appointments, exercise programs, alert mechanisms for safety, and to communicate with healthcare professionals. Recently extensive research studies [7] have been performed to understand the impact of MTRs involving older adults before and during the COVID. The results showed that new tools such as MTRs are accepted for their usability. However, to implement them widely there are barriers to be overcome such as complexity in the interface, steep learning curve to use these tools, and the potential effects on the mental health. Our design research has explored new paradigms for expanding beyond these conventions. In prior work, we have explored interfaces that replace the screen with projection. This provides a user experience that has no frame, such as one finds on the edges of a computer screen, and is therefore more integrated into the built environment. With projection, the image can also be made much larger and more immersive when projected onto the nearby walls, especially when applied in low light environments [8], [9]. In this paper we discuss several prototypes of Large screen telepresence devices being developed by our design research team. Specifically, we discuss about 1) a Large Screen Mobile Telepresence Robot (LSMTR) and 2) a Mixed Reality Window (MRW).

Our Large Screen Mobile Telepresence Robot (LSMTR) (see Fig: 1) builds upon the above listed innovations. The videoconferencing experience in particular, and other collaborative remote experiences more generally, are completely reimagined. Users are given the ability to move within a remote environment. The large screen, which is roughly human scale, also helps to establish a more embodied experience than most telepresence robots currently on the market [1]–[4]. This affords an experience somewhere between projected video integrated into the built environment explored in prior work, and the conventional laptop or desktop screen. When contrasted against a fixed laptop, this addition of high mobility

*Work done during their time at Brown University

(a) Remote user and local user walking side by side

(b) Robot being piloted remotely

Fig. 1: Large Screen Mobile Telepresence Robot (LSMTR)

(a) Tbo - tablebot

(b) Tbo being a furniture

(c) Tbo moves to start a call

(d) Tbo projects the remote user

Fig. 2: Robotic furniture: Tbo is a "Situated Robot." It allows users to beam into furniture for videoconferencing, and then returns to being simply furniture once the call is over. [project webpage]

and a large screen gives users a more embodied experience when communicating or collaborating at a distance.

In contrast to a completely virtual experience, where users meet each other in a metaverse or some other completely virtual space, our prototype explores the viability of using these kinds of telepresence robots for Mixed/Augmented Reality applications. Virtual/Augmented/Mixed (VAM) Reality can be described as existing on a continuum between the purely virtual and the purely physical [10]. On the one end of the spectrum you might find devices such as VR headsets, which give the user a completely immersive field of view within a completely virtual environment. On the other end of the spectrum you might find something more quotidian; a physical monitor sitting on a physical table with the physical body of the user positioned somewhere between. Both modalities embody certain possibilities and limitations. Our prototypes fit somewhere in the middle of this spectrum. The Large Screen Mobile Telepresence Robot (LSMTR) has at least two important innovations: It is mobile and can therefore navigate physical spaces at a distance, and the screen is human scale, thereby offering a more embodied experience. The Mixed Reality Window (MRW) (see Fig: 3) builds upon this integration of large screens into mobile platforms, by integrating bidirectional cameras and bidirectional screens into a Mixed Reality experience between non-contiguous spaces.

We have also explored the integration of these features into furniture and other elements of the built environment. Users can beam into these objects, and then beam out, allowing the robot to return to its prosaic role as furniture when it is no longer called upon for videoconferencing or other roles. These kinds of robotic furniture, or Situated Robots (see Fig: 2) as we call them, give users more mobility within remote environments, while also creating a virtual portal to distant spaces.

II. Prototyping

We have developed a family of prototypes that investigate the viability of integrating mobile large screens into the built environment, as well as their potential for Mixed Reality applications.

The Large Screen Mobile Telepresence Robot (LSMTR) was the first in this series. It explores novel design paradigms for the implementation of telepresence between local and remote users. In essence, the mobility of the large screen allows the LSMTR to become the physical body of the virtual user. The form factor of the LSMTR evokes an oversized smartphone, and like most smart phones, has the potential to be equipped with both forward facing and rear facing cameras to

Fig. 3: Prototype of the Mixed Reality Window.

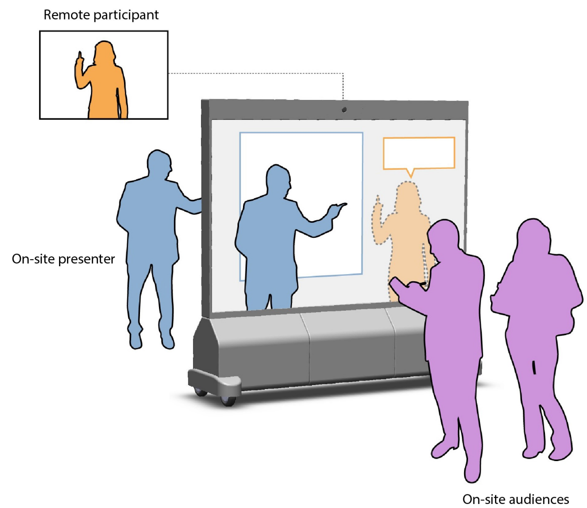


Fig. 4: Concept diagram for the Mixed Reality Window at full scale. Remote participants interact with local users on both sides of the window. Notes and presentation content are virtually overlaid onto both sides of the Mix Reality Window (MRW)

create affordances for Mixed Reality applications. These features were further developed in subsequent iterations. Unlike smartphones, which are usually tethered to the hands of users, and at a far smaller scale, LSMTR's size and mobility offer opportunities for interactions with remote users and in remote environments that feel more immersive and embodied. While LSMTR can accommodate conventional videoconferencing, where the unadulterated image from the camera of the remote user, at full scale, is directly translated onto the screen of the local user, we have also expanded upon these features towards the development of prototypes that can overlay the image of the remote user onto the background captured by the rear facing camera, thereby creating a passthrough effect ideal for AR/MR applications.

Building upon the insights we gained in developing the large, mobile screen of the LSMTR, we explored multidirectional screens and cameras. The Mixed Reality Window (MRW) features both a forward facing camera, which displays on a backward facing screen, and a backward facing camera which displays on a frontward facing screen, producing a passthrough effect, as if looking through a window from both sides. Like the LSMTR, this provides users with affordances for overlaying visual content onto the screen (e.g. slide presentations, annotations, etc.) while also giving users on both sides the ability to look through the screen at each other and the surrounding environment as if through a window. The whole structure is on wheels, which allows it to be moved easily around the space, manually, if not mechanically. Casters offer a simple and inexpensive solution to mobility in situations where autonomy or remote control are not necessary (see Fig. 3, 4).

These prototypes inspired us to consider the ways the scale of these screens could be greatly increased. By working at this architectural scale, and drawing from the formal languages of furniture, we can imagine new strategies for Mixed Reality and ubiquitous computing that are always already situated in the built environment. This prototype allowed us to take AR/MR off the body (e.g. Google Glass, Hololens, Oculus, etc.) and place it in the environment, incorporating it directly into the objects and spaces that are always already around us.

III. Scenarios

Over the course of development of the LSMTR and the MRW, we have explored user scenarios that help us better understand how larger, mobile screens, and their application to Mixed Reality, might be further developed.

Of particular interest were scenarios that explored the future of work; in particular collaborative interactions at a distance. We began with the question, "How might we design an embodied telepresence experience that allows users to feel as if they were standing in the same space as a remote colleague?" How might this improve collaboration? How might the integration of a full range of motions and gestures provide affordances for better communication and empathy with collaborators? Our demo video for the LSMTR prototype presents a scenario where a virtual CAD part can be sent to a collaborator, and physically printed with the assistance of the originator of the part[1],[2]. In addition to instructing the remote user how to use the 3D printer, the pilot can give instructions for how the part might fit spatially, using gesture and motion.

Another scenario that is of particular interest is the application of these emerging design conventions to remote learning. Both the MRW or LSMTR can be applied to remote and hybrid learning, for example. The COVID19 pandemic has challenged designers and teachers to rethink the development of hybrid courses, which can equally serve the needs of both local and remote students. We asked how these kinds of robots

[1] Large Screen Mobile Telepresence Robot [video]
[2] Large Screen Mobile Telepresence Robot Demo [video]

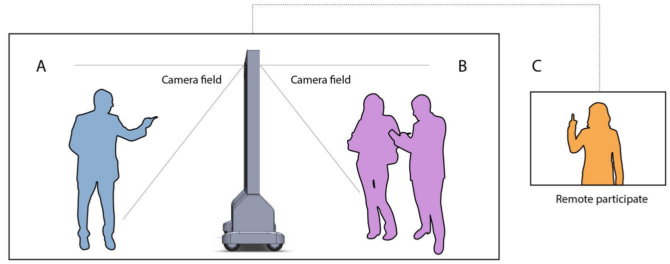
Fig. 5: An example of remote instruction with the Large Screen Mobile Telepresence Robot.

Fig. 6: Example of gestural communication between local and remote users.

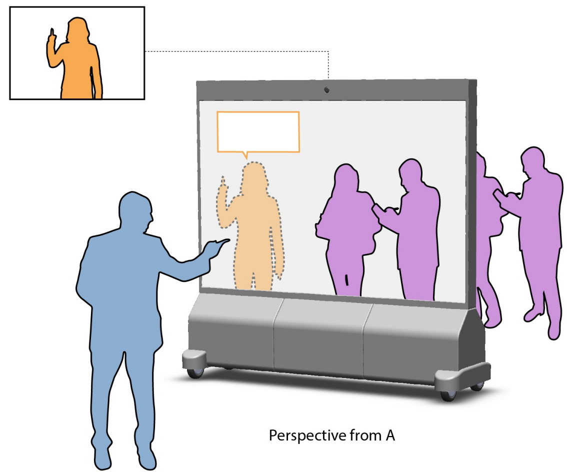
might offer new paradigms for sharing content and facilitating discussion.

Imagine an instructor at the front of a lecture hall with a MRW in front of her. The instructor can see her students and the students can see the instructor through the frame of the screen, which is on wheels. The instructor can layer her presentation slides, technical details, annotations, and other relevant content - text, image, and video - onto the screen facing the students. She can also access content on her side, such as student names or lecture notes. This approach, employing bidirectional cameras and screens, provides affordances for a Mixed Reality experience that can be overlaid with relevant content. A third user group – remote users – can see both what the instructor is seeing and what the class is seeing, providing a complete sense of the interactions taking place in the remote environment(see Fig: 7).

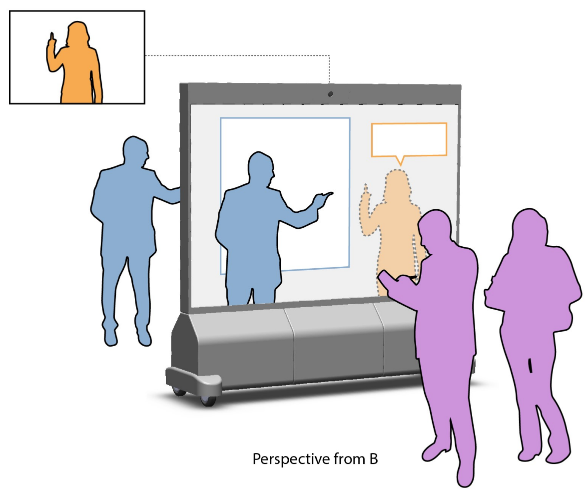
The MRW offers the potential to radically change classrooms and work spaces. This approach to Mixed Reality experiences that are integrated into the built environment fluently connects remote users with local users, as well as offering those learning and working together in the same physical space better tools for collaboration. Deriving from insights gained in the development of Tbo, and Situated Robots generally, the integration of these Mixed Reality Windows, using screens



(a) The on-site instructor and students interact via MRW's past-through screens. The remote student's live feed is overlaid onto both sides of the screen.



(b) The instructor can see the on-site students, the virtually present remote student, and augmented notes on MRW's display.



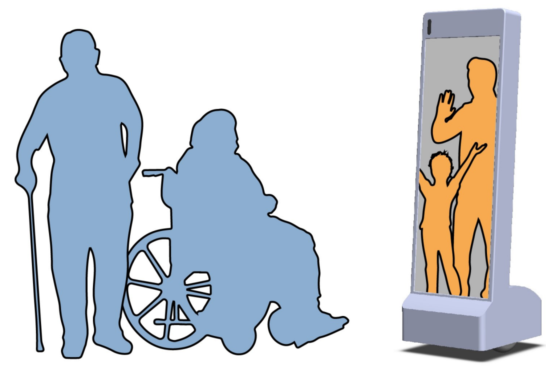
(c) The on-site students can see the instructor, the virtually present remote student, and augmented notes on MRW's display.

Fig. 7: Concept diagram of MRW in a lecture scenario

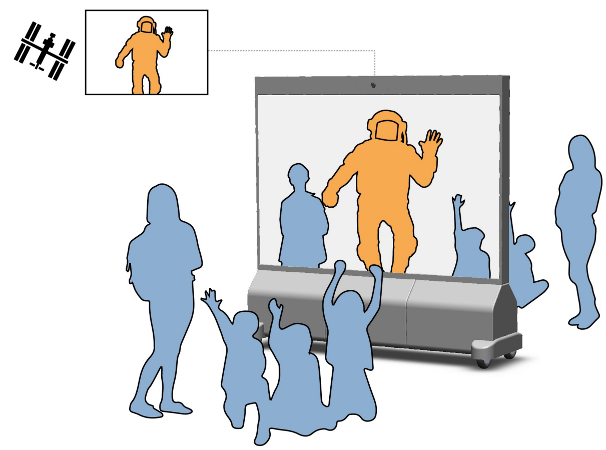
instead of projection, can be directly built into the furniture and architecture of a space.

While consciously referencing windows, the form factor of the MRW also intentionally draws on smart boards, and their antecedents such as chalk boards. The ability to augment content by making thinking visible, through both prepared and extemporaneous content, offers opportunities for sharing and archiving lectures and discussions in ways not yet possible. Furthermore, by allowing remote users to participate, and overlaying their images and content directly onto the bidirectional window, it greatly expands accessibility to these spaces in ways that are not strictly bound by physical proximity.

We imagine these types of large screen, mobile Mixed Reality platforms might be applied to a range of user scenarios beyond our initial investigation outlined in the scope of this paper. For example, particularly promising are applications for seniors who wish to age in place, connecting with family and caregivers without the need to move to an assisted living facility. The mobility of these robots acts as a surrogate when physical mobility is limited. The need for these kinds of applications has become even more urgent since the beginning of the pandemic, where isolation of seniors has had a significant psychological and physical impact(see Fig: 8-a). One can imagine even more speculative scenarios, beyond these: an astronaut in low earth orbit teaching a lesson to a class of children on earth, or workers collaborating in hazardous environments where physical presence is not practical or safe(see Fig: 8-b).



(a) LSMTR allows for remote interactions between older family members and their relatives.



(b) A speculative scenario where the MRW could be used to create an immersive classroom experience between students on earth and astronauts in low earth orbit.

Fig. 8: Large screen telepresence scenarios and applications

IV. Conclusion

With larger, bidirectional screens and cameras, and the further development of both autonomous and manual movement, we envision the possibility of an even richer integration of these kinds of technologies into the built environment. Not only does this offer an alternative paradigm for ubiquitous computing by integrating Mixed Reality experiences directly into the built environment, it allows designers, architects, and engineers to fundamentally rethink the ways interactions between non-contiguous spaces and collaboration at a distance might occur. These prototypes validate these design principles, and provide models upon which further design research can be developed.

References

[1] "Robot Beam. AWABOT beyond Robotic." https://awabot.com, [accessed on 19 September 2020].

[2] "Robot Double. Doublerobotics." https://www.doublerobotics.com, [accessed on 19 September 2020].

[3] "Robot Cutii. CareClever." https://www.cutii.io, [accessed on 21 September 2020].

[4] "Robot Kompai. Kompairobotics," http://kompairobotics.com/fr, [accessed on 25 September 2020].

[5] J. M. Beer and L. Takayama, "Mobile remote presence systems for older adults: acceptance, benefits, and concerns," in *Proceedings of the 6th international conference on Human-robot interaction*, 2011, pp. 19–26.

[6] A. Kristoffersson, S. Coradeschi, and A. Loutfi, "A review of mobile robotic telepresence," *Advances in Human-Computer Interaction*, vol. 2013, 2013.

[7] B. Isabet, M. Pino, M. Lewis, S. Benveniste, and A.-S. Rigaud, "Social telepresence robots: a narrative review of experiments involving older adults before and during the covid-19 pandemic," *International Journal of Environmental Research and Public Health*, vol. 18, no. 7, p. 3597, 2021.

[8] I. Gonsher and J. Y. Kim, "Robots as furniture, integrating human-computer interfaces into the built environment," in *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, 2020, pp. 215–217.

[9] I. Gonsher, "Beyond the keyboard, mouse, and screen: New paradigms in interface design," in *Proceedings of the Future Technologies Conference*. Springer, 2021, pp. 107–113.

[10] P. Milgram, H. Takemura, A. Utsumi, and F. Kishiro, "A class of displays on the reality-virtuality continuum," *Telemanipulator and Telepresence Technologies, Boston, MA, USA: SPIE*, p. 282, 1992.