

# A Mixed Reality Telepresence System for Dissimilar Spaces Using Full-Body Avatar

Leonard Yoon  
Dongseok Yang  
lyoon@kaist.ac.kr  
dsyang@kaist.ac.kr

Korea Advanced Institute of Science and Technology

Choongho Chung  
Sung-Hee Lee  
thegenuine@kaist.ac.kr  
sunghee.lee@kaist.ac.kr

Korea Advanced Institute of Science and Technology



**Figure 1:** Our system allows for different room environments by retargeting the user's placement and deictic gesture to the remote avatar so as to preserve the user's context of placement and interaction. Small figures are user's egocentric views.

## ABSTRACT

We present a novel mixed reality (MR) telepresence system, which enables a local user to interact with a remote user through full-body avatars in their own rooms. If the remote rooms have different size and furniture arrangement, directly applying the user's motion to the avatar leads to mismatch of placement and deictic gesture. To overcome this problem, we retarget the placement, arm gesture and head movement of the local user to the avatar in remote room to preserve the environment and interaction context of the local user. This allows avatars to utilize real furniture and interact with the local user and shared objects as if they were in the same room.

## CCS CONCEPTS

• **Human-centered computing** → **User interface toolkits.**

## KEYWORDS

telepresence, avatar, mixed reality, social interaction

## ACM Reference Format:

Leonard Yoon, Dongseok Yang, Choongho Chung, and Sung-Hee Lee. 2020. A Mixed Reality Telepresence System for Dissimilar Spaces Using Full-Body

Avatar. In *SIGGRAPH Asia 2020 XR (SA '20 XR)*, December 01–09, 2020. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3415256.3421487>

## 1 INTRODUCTION

Recent advances in technology facilitated room-scale telepresence research in a MR setting, which allows a user to communicate with a remote partner augmented as an avatar. If the two rooms have identical furniture arrangements, augmenting a virtual copy of a remote user at the same location in the local space can clearly deliver the context of user's motion [Orts-Escolano et al. 2016].

However, in the case of using a whole space of the two rooms with different shapes and object arrangements, directly placing an avatar with respect to the spatial relation between a user, an avatar, and shared objects of interest may cause a discrepancy, for instance, an avatar sitting in the air or penetrating some objects. To avoid these artifacts, the avatar may be only allowed to be placed on limited areas such as a free space [Lehment et al. 2014] or a sofa [Pejsa et al. 2016]. The drawback is that an available space of communication is restricted to only a sub-space of the whole space. Alternatively, using a partial representation (e.g., upper body) of avatar can reduce the mismatch of environmental context between different spaces [Spatial 2020]. This is effective for the typical tasks in office meeting, but the absence of full-body motion limits the scope of interaction and reduces the presence of partner.

The placement of avatar from the different environment of two spaces have been studied by [Jo et al. 2015; Yoon et al. 2020]. However, to our knowledge, a telepresence system that places an avatar to an optimal location and animate it to keep the interaction context has not been developed yet. This paper presents our work to

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

SA '20 XR, December 01–09, 2020, Virtual Event, Republic of Korea

© 2020 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-8111-6/20/11.

<https://doi.org/10.1145/3415256.3421487>

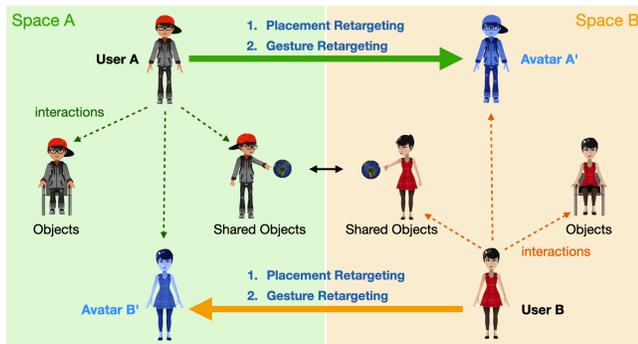


Figure 2: Telepresence system for two dissimilar spaces.

contribute to this goal. Specifically, we exploit our prior work [Yoon et al. 2020] for the avatar placement. In addition, this paper newly introduces our avatar gesture retargeting method, which detects a deictic pointing gesture of user and retargets it to avatar. Combining the two enables users to move between different locations of a small room to sit and stand while interacting with a partner using shared objects. Figure 1 describes a use case scenario.

## 2 SYSTEM OVERVIEW

Figure 2 illustrates the concept of user actions and avatar animation in our telepresence system. The goal of our system is to enable a user to make use of his or her own space including existing furniture while interacting with a remote person represented as a full-body avatar, and sharing a screen or augmented virtual objects. To deliver the user’s environmental and interaction context in the local space to the remote space, the user’s motion needs to be retargeted to his or her avatar. With this goal in mind, we designed our system as following:

- The avatar is placed to match the environment context of the remote user and the spatial relation with the partner.
- The avatar’s upper body gesture is modified to preserve the interaction context of the remote user if the user is paying attention to interaction targets.
- Candidate interaction targets include the partner’s head, a shared screen, and augmented virtual objects for pointing and gazing activities.

We use an avatar placement algorithm from our prior work [Yoon et al. 2020], which finds the avatar’s placement in a dissimilar space as *similar* as possible to the user’s placement. If the user engages in interaction with the partner or a shared virtual object, the avatar’s movement is decoupled from the user’s and made to realize the interaction context in the remote space by controlling the avatar’s head and arm to point the same targets.

By integrating placement and deictic gesture retargeting for avatars in dissimilar spaces, we can reduce confusion and ambiguity in interpreting both spatial and deictic context of the other user who is communicating from a different physical space through a virtual avatar.

Our system consists of two identical hardware setups for each distant space. We integrated a RGB-D camera (ZED mini) on HTC Vive Pro headset for MR rendering and real time occlusion between

real and virtual objects. Additional Vive trackers are worn by a user to track joints and reconstruct the whole body pose with an IK solver (Final IK). For accurate finger action of a remote avatar, a user wear Noitom Hi5 VR Gloves. The system is implemented with Unity3D engine (2019 2.6f1) and SteamVR framework.

## 3 EXPERIENCE

Please refer to the supplementary video for demonstration.

### 3.1 Interactive Learning

The participants can converse about the Sun, Moon, and Earth displayed on the screen while sitting in the chairs, and then they can stand around a table for a deeper discussion with 3D dynamic objects [Dong 2017] augmented above a table for both rooms.

### 3.2 Collaborative Painting

The participants can perform digital painting while sitting in a chair or standing in front of a screen using various colors and shapes. During each task, the participants will be encouraged to talk about the contents of their drawings.

## 4 LIMITATION AND FUTURE WORK

The proposed system work well with the common room settings with chairs and tables. However, the difference of seating arrangement or shape of a room between two sites may place an avatar in front of the different interaction target compared with the user in remote space. In extreme cases, an avatar can be placed in front of a TV while a user is looking at a partner. Also, the distance between two interaction targets (e.g., a partner and a TV) may vary depending on the position of a user. In such cases, the resulting motion of gesture retargeting algorithm might not look natural or feasible. For future work, we plan to identify those undesirable situations and guide the users to avoid them.

## ACKNOWLEDGMENTS

This work was partly supported by Samsung Science and Technology Foundation (SRFC-IT1701-14).

## REFERENCES

- Dee Dong. 2017. Solar-System-Unity3D. <https://github.com/PatwinchIR/Solar-System-Unity3D>.
- Dongsik Jo, Ki-Hong Kim, and Gerard Jounghyun Kim. 2015. SpaceTime: Adaptive Control of the Teleported Avatar for Improved AR Tele-conference Experience. *Comput. Animat. Virtual Worlds* 26, 3-4, 259–269. <https://doi.org/10.1002/cav.1645>
- Nicolas H. Lehment, Daniel Merget, and Gerhard Rigoll. 2014. Creating automatically aligned consensus realities for AR videoconferencing. In *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 201–206. <https://doi.org/10.1109/ISMAR.2014.6948428>
- Sergio Orts-Escolano, Christoph Rhemann, Sean Fanello, Wayne Chang, Adarsh Kowdle, Yury Degtyarev, David Kim, Philip L Davidson, Sameh Khamis, Mingsong Dou, et al. 2016. Holoportation: Virtual 3d teleportation in real-time. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. 741–754.
- Tomislav Pejša, Julian Kantor, Hrvoje Benko, Eyal Ofek, and Andrew Wilson. 2016. Room2Room: Enabling Life-Size Telepresence in a Projected Augmented Reality Environment. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (San Francisco, California, USA) (CSCW ’16)*. ACM, New York, NY, USA, 1716–1725. <https://doi.org/10.1145/2818048.2819965>
- Spatial. 2020. Retrieved May 06, 2020 from <http://spatial.io/>
- Leonard Yoon, Dongseok Yang, Jaehyun Kim, Choongho Chung, and Sung-Hee Lee. 2020. Placement Retargeting of Virtual Avatars to Dissimilar Indoor Environments. *IEEE Transactions on Visualization and Computer Graphics* (2020). <https://doi.org/10.1109/TVCG.2020.3018458>