Ghost-Hack AR: Human Augmentation Using Multiple Telepresence Systems for Network Communication

Hiroaki Tobita
Advanced Institute of Industrial Technology
Tokyo, Japan
tobby@aiit.ac.jp

ABSTRACT
We present Ghost-Hack AR, an approach that enhances network communication with a simple interaction technique by which remote users can change telepresence systems depending on the situation and purpose. Various telepresence systems, each with its own advantages and disadvantages, support communication between remote and local users. However, since remote users of conventional systems primarily use one telepresence system to communicate with local users, their activities and interactions depend on the system and are thus limited. By contrast, our system allows remote users to change telepresence systems via a simple interaction technique in order to reap the benefits of each system.

Author Keywords
Multiple telepresence, teleoperation, face-to-face, network communication, interactive system, augmented reality

ACM Classification Keywords
H5.2 [Information interfaces and presentation]: User Interfaces. - Graphical user interfaces

INTRODUCTION
Various telepresence systems support network communication between remote and local users. Some focus on face-to-face communication via video streaming, in which users share facial expressions and make eye contact, whereas others represent remote users' presence via robotics and display systems. When local users communicate via multiple systems, they can communicate more directly with remote users; however, in using only one system, remote users' activities remain limited.

Conventional systems provide several functions (e.g., display, gesture, mobility) and are divided roughly into two types: static and dynamic. Typical features of static systems are a large screen and display (e.g., wall-type screens, tabletop surfaces) used for formal meetings. Since the screen resolution is high, local users can clearly see remote users, share digital information, and even draw pictures on the screen. However, because the display is fixed, remote users can move only in the display space. Moreover, to initiate a meeting, local users have to move to where the system is. By contrast, dynamic systems are mobile and allow remote users to move around the local environment [1, 2, 3]. Since remote users can find and approach local users to initiate communication, the systems make them more active than static systems do. However, the display cannot show information comparable to what static systems can. Thus, remote users cannot reap the benefits of either type of system given the lack of interaction techniques to manipulate different telepresence systems.

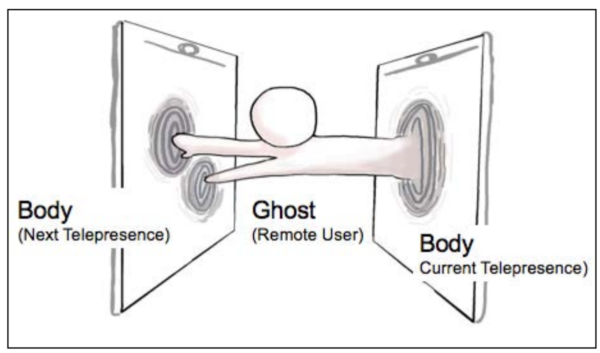

Figure 1: Concept of Ghost-Hack AR, in which a remote user (i.e., ghost) moves from one telepresence system (i.e., body) to another.

In response to those disadvantages, we developed Ghost-Hack AR, a set of unique interaction techniques for using multiple telepresence systems to enable remote users to enhance their presence and activities (Fig. 1). Our approach provides simple interactions for using multiple telepresence systems and their functions instead of a complex telepresence system with multiple functions. First, we prepared several telepresence systems (e.g., static, mobile, wearable), each connected to the network and involving a physical body with a camera, display, microphone, and speaker (Fig. 2 (left)).

PerDis '17, June 07-09, 2017, Lugano, Switzerland

ACM ISBN 978-1-4503-5045-7/17/06□$15.00
http://dx.doi.org/10.1145/3078810.3078827

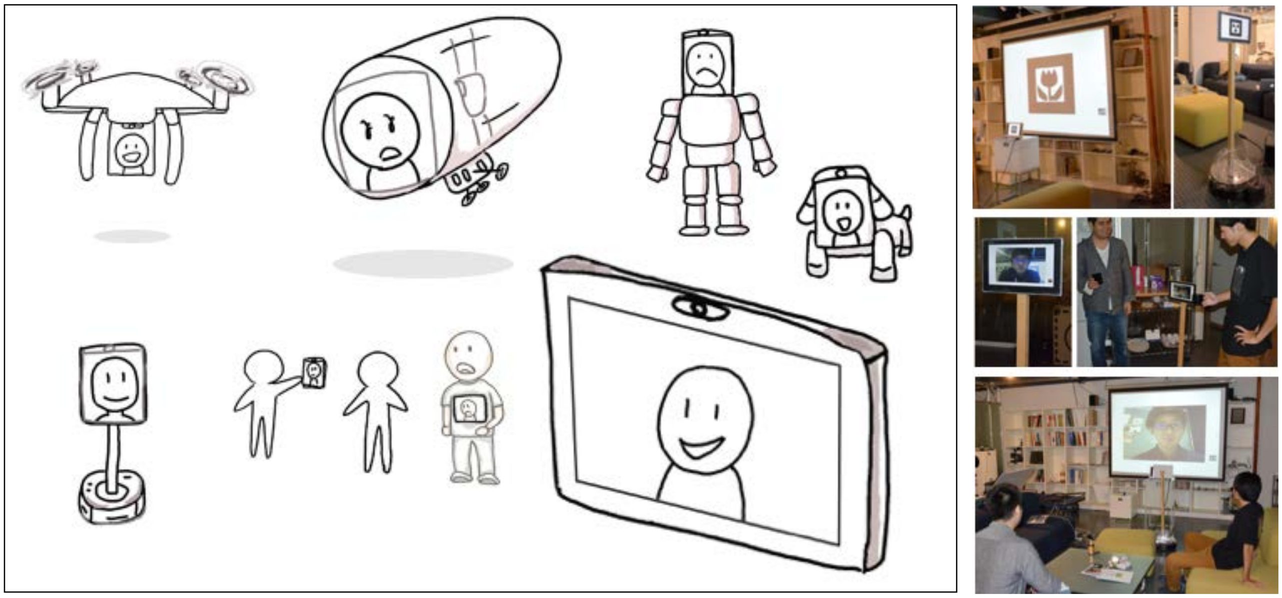

Figure 2: Ghost-Hack AR system overview, in which each telepresence system consists of a display and physical body with which remote and local users communicate face-to-face via the display (left)—for example, from a screen telepresence system to a wearable system and then to a mobile system (right).

Second, we developed simple interactions by using an augmented reality (AR) marker with a unique identity (ID) and position in the real world [4, 5]. Our system uses the AR markers to trigger movement from the current telepresence system to the next system so that a remote user can find a marker displayed on another telepresence system and initiate it in the graphic user interface (GUI). The interactions enable remote users to move around in a local environment via the mobile telepresence system. When the mobile telepresence system locates in front of a large-display system, the remote user can move from the mobility system to the large-display system fixed nearby in order to communicate in greater depth (Fig. 2, (right)).

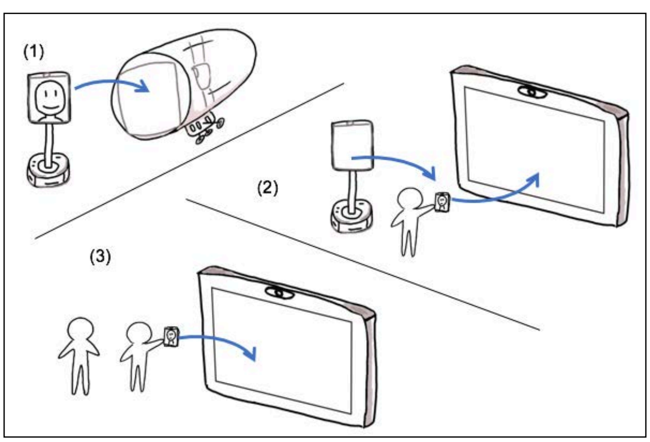

Figure 3: Communication with Ghost-Hack AR, in which a remote user moves from one telepresence system to another depending on the purpose.

Scenarios of Ghost-Hack AR Communication

We now introduce three situations in which a remote user can change from one telepresence system to another.

Mobility-Based Activity

As shown in Fig. 3(1), as the systems move around in the local environment, remote users can use different mobile systems to solve two problems. One, they can change their telepresence system if the battery dies, without which the system stops functioning. Two, they can avoid obstacles to moving in the local environment. Conventional mobile telepresence systems support both terrestrial movement and flying so that remote users can change the mobile system depending on the obstacle.

Event-Oriented Activity

Remote users can attend all sorts of social events, including meetings, while changing from the static telepresence system to the mobile one, as shown in Fig. 3(2). For example, although a video conference is useful for in-depth discussions, during coffee breaks, for example, remote users cannot participate. Since many important ideas are developed in casual chats, with our system remote users can change their mobile telepresence system after the formal discussion and join casual conversations or spontaneous meetings during coffee breaks.

User-Mediated Activity

Remote users can change telepresence systems not only by themselves, but also with the local user's support, as Fig. 3(3) shows. If the local user has a smartphone, then the remote user can first move to the smartphone and then to another telepresence system. In such a scenario, the local user changes the angle and position of the smartphone so that the remote user can locate and easily relocate to the next system. In particular, by using a T-shirt telepresence, a remote user can be part of the local user's body so that both users can move together and share information with the same view and activities (Fig. 4).

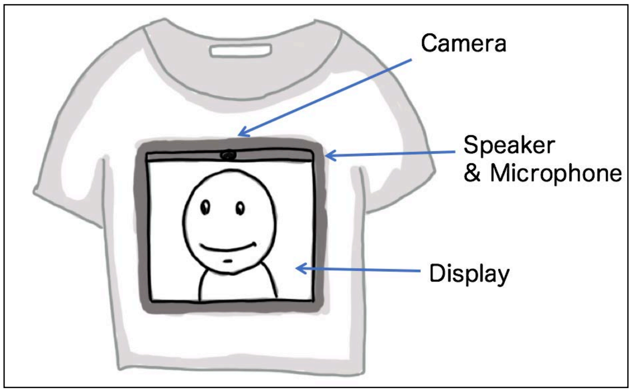

Figure 4: Example of a telepresence system integrating a tablet computer with a T-shirt.

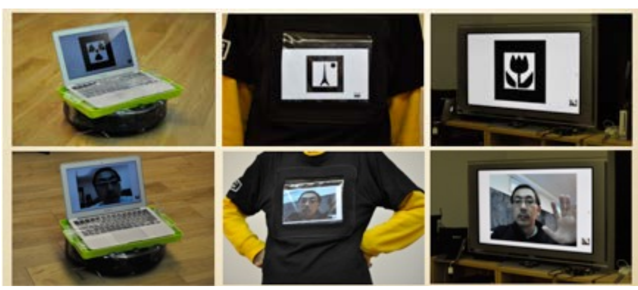

Figure 5: The system displays an augmented reality marker during wait mode (top panel) and streams video during active mode (bottom panel).

Interaction

To establish connections among different telepresence systems, we use an AR marker common in marker-based AR applications. Since each telepresence system has a display, it is easy to show both the AR marker and video stream. To switch from the marker to the stream, each display has two modes: wait and active (Fig. 5). In wait mode (Fig. 5, top panel), a device shows the AR marker in the display area, meaning that the device is ready for a remote user to connect. In active mode (Fig. 5, bottom panel), the display shows the face of the remote user, who can control the device if it allows interactions. By using both modes, local and remote users can recognize the wait or active mode of the system.

Interaction with Telepresence Systems

Figure 6 demonstrates the remote user's movement from the current to another telepresence system. In Figure 6.1 and 2, a remote user moves from a T-shirt to a TV screen-type system. First, the local user wears the remote user through a T-shirt system, and the remote user locates another system that shows the AR marker in the display area from the video camera on the T-shirt (Fig. 6.3). A virtual tower is superimposed on the AR code if the system recognizes it (Fig. 6.4), after which the remote user pushes the GUI button to begin moving to the TV screen from the T-shirt. Support information from the human model appears on the screen and changes size depending on the distance to the next telepresence system (Fig. 6.5 and 6). Lastly, the remote user completes the change (Fig. 6.7) so that his or her face is displayed on the TV screen and retrieves local information from the video camera. The T-shirt display changes the content from the video stream to a 2D marker, which prepares the T-shirt for telepresence.

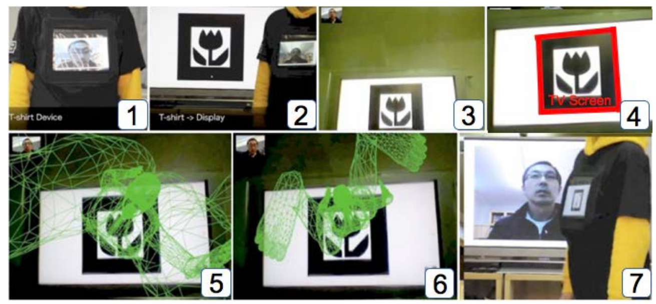

Figure 6: Interaction with Ghost-Hack AR, in which the remote user locates the augmented reality code displayed on another telepresence screen (1–4), after which the user moves from the current telepresence to another (5–7).

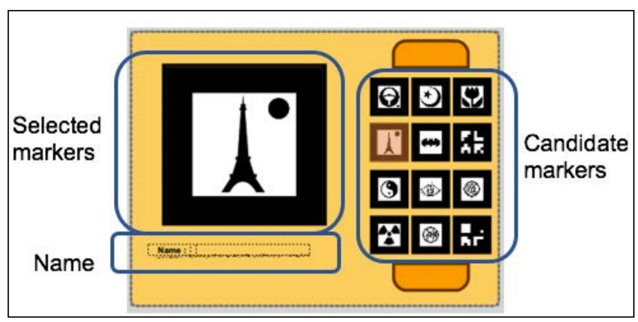

Figure 7. Graphic user interface to register a new telepresence system, which contains candidate markers.

Registering the Telepresence System

The telepresence system allows local users to easily add new telepresence systems to the server computer. The system provides GUIs (Fig. 7) so that local users can select a marker from candidate markers on the right side of the window and set the name of the device. Next, the marker ID, name, and media access control address of the device are sent to the server. This GUI is also used to change the 2D marker of current telepresence systems. The server stores that information on a database and refers to it in order to establish a connection between the remote user's PC and the next telepresence system.

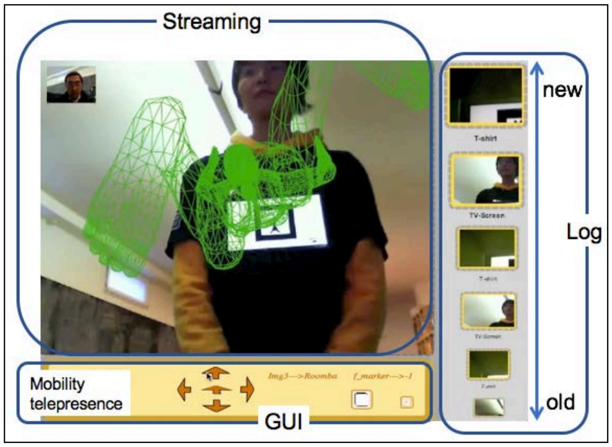

Figure 8: Remote user's interface, divided into three parts: the streaming window, flexible graphic user interfaces, and the transfer log.

Interface of Remote Users

Remote users manipulate GUIs to control telepresence systems in the local environment with an interface divided into three areas: the streaming window, flexible GUIs, and the transfer log (Fig. 8). Remote users can use GUIs with a PC, tablet, or smartphone.

Streaming Window

Displaying the video stream, the streaming window is used to communicate with local users, browse the local environment, and locate AR markers. The streaming window displays video captured with a camera mounted on a telepresence system in the local environment. When a remote user locates an AR marker displayed on a system in that environment, visual information related to the system such as ID and name appears on the marker. Such information is superimposed on the streaming window, and if the user starts to move from one device to another, then the virtual 3D computer-generated character—that is, the ghost of the remote user—appears. The 3D character is also superimposed on the streaming window. When the ghost of the remote user appears in the new telepresence system, the video stream changes from the previous device to the new one, after which the remote user retrieves the stream from the new telepresence system.

Flexible GUI

Our approach supports different types of telepresence systems with different functions, meaning that GUI elements change depending on the functions of the system. For example, if the system supports drawing, then pen and palette GUIs appear. If the system is mobile, then arrow buttons to control movement appear. Therefore, a remote user can move into the local environment with the mobile system by manipulating the control buttons. If the 2D marker is recognized, then the telepresence button on the left side of the GUI area can be used to move from one telepresence system to another.

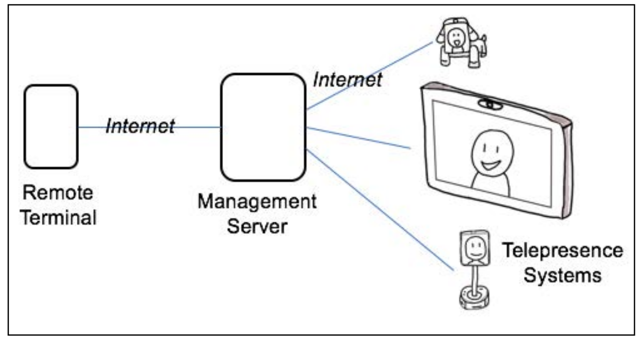

Figure 9: Implementation, with all systems connected to the network.

Transfer Log

The transfer log refers to the list of telepresence systems to which users have moved and allows users to move back to them. As mentioned, each telepresence system is connected to a Wi-Fi network so that the system can change its streaming channel without an AR marker. Since the log appears along a timeline, the top of the log lists the most recent telepresence system and the bottom lists the oldest. When a user moves to another system, a new icon denoting the previous system appears at the top of the list. Users can manipulate the list by clicking each icon to move from the current to another system. At the same time, the GUI changes depending on the features. In our prototype, each icon disappears at a regular interval or depending on the distance between the previous and current telepresence systems.

IMPLEMENTATION

All systems are connected to the network and work on different computers (Fig. 9), while the management server handles streaming, telepresence information, and AR markers. We use the Open Source software Red5 to implement the video stream server and database server to treat AR marker information. In the server, information about the AR marker (e.g., ID, name, media access control address) is treated as XML files.

Each client and telepresence system accesses the server via the Internet. Each telepresence system uses the same software that enables the recognition of the 2D marker and shows the video stream. We use Java to implement client applications of remote users, which enables them to work as desktop and mobile applications.

To support face-to-face communication, each telepresence system needs several elements (e.g., display, camera, microphone, speaker). We simply use mobile devices (e.g., mobile phone, tablet, PC) that contain the elements (Fig. 10). Those mobile devices support wireless communications such as Wi-Fi and Bluetooth low energy (BLE). Wi-Fi is used to communicate with the management server, whereas BLE is used to control mobile telepresence systems via Arduino.

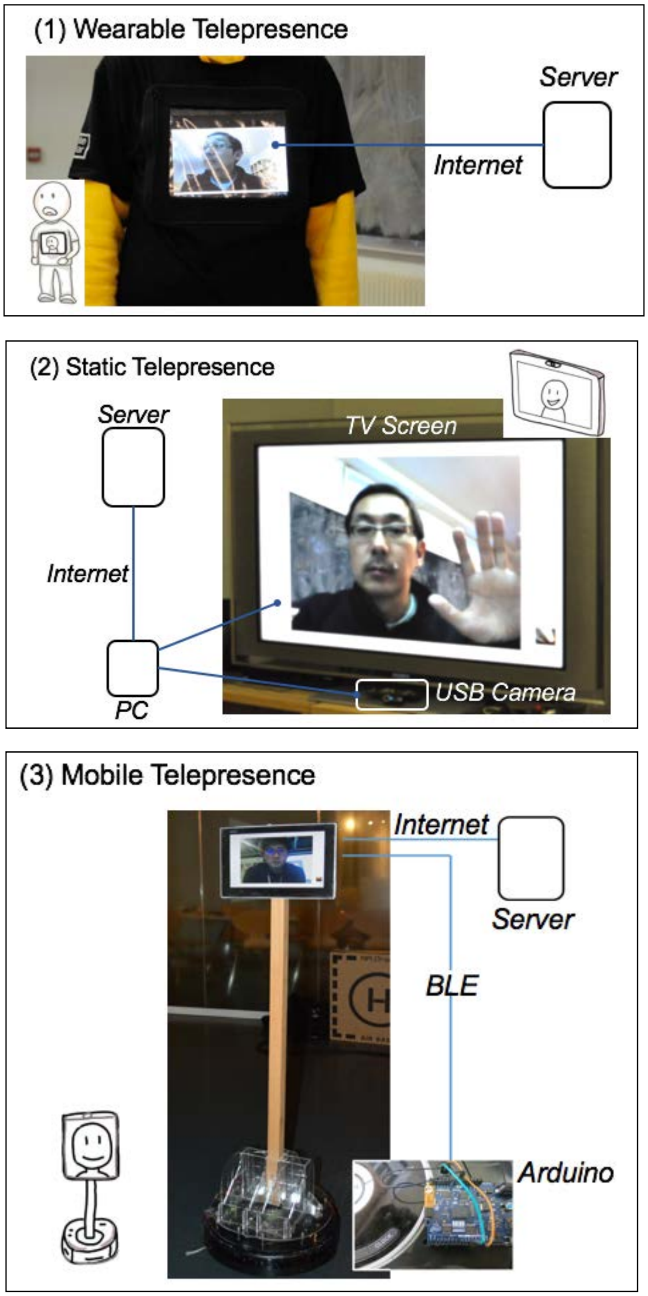

Figure 10: Example of telepresence implementation with (1) wearable, (2) static, and (3) mobile telepresence systems.

DISCUSSION

We have had several opportunities to demonstrate our system. In this section, we discuss our prototype of Ghost-Hack-AR based on demonstrations and comments from trial users.

Interaction of GH-AR

In our demonstrations, we prepared three interactions: TV screen to talking head, talking head to T-shirt, and TV screen to talking head with a local user's support. Most trial users understood our concept of allowing remote users to change systems, though some commented that it was difficult to locate the next system in the local environment from the remote view. It was also difficult for remote users to identify detailed local information, even though local user support facilitated that process. Some trial users who manipulated telepresence systems remotely requested help with locating and changing to the next system, and we observed that some local users controlled the position and angle of the current or next system. With such support, it became easier to locate and move to the next system, because local users knew the local environment better than remote users and could suggest the next system and navigate the current one. Thus, if local users physically supported teleoperation, then remote users could move from one system to another and had the feeling of being close to the local users.

Future Directions

We plan to further develop Ghost-Hack AR for education in which telepresence systems are used. In e-learning classes, communication and collaboration between remote and local users has become increasingly important. However, interaction via e-learning is limited by the use of fixed display with video streams and chat platforms. Our approach is useful for both teachers and students, since remote teachers can visit each classroom via the telepresence system and change telepresent bodies depending on the purpose. Teachers can also change their telepresence bodies via simple interaction. At the same time, remote students can attend real classes by using telepresence bodies, which by changing they can engage a wide variety of activities and easily select different viewpoints.

Although our system allows a single user to move from one telepresence device to another, it is possible to handle multiple users on one device. We can further develop our system to support several remote users on one telepresence device at the same time. In such an environment, remote users can share the same device, workspace, and view. It will be interesting to see them sharing one mobile telepresence device. Each user can share the same camera view mounted onto the device. Thus, remote users can share experiences through the network. Since our focus is face-to-face communication, it is important for local users to have a layout of multiple remote users' faces. Since our system is for video streaming, we can simply use the layout supported in conventional video chat systems.

Related Work

In most telepresence systems, remote users work as ghosts; they control local systems remotely and use telepresent bodies instead of their actual bodies [6, 7, 8]. Similarly, in Porta-Person [9], a telepresence system is placed on or near a table in a conference room, whereas Remote Handshaking [10] is a system for enhancing communication between local and remote users by using a real hand model. By contrast, Physical Telepresence [11] integrates face-to-face communication with a 3D display that can flexibly change an image's shape. Several other systems involve robotics for network communications [12, 13]. In Texai [1], for example, the systems move on the ground, whereas BitDrones [14], manipulated by users, float in the air.

AR research has generated several systems that integrate real and virtual worlds [15], recognize real-world information, and display virtual information related to the results of recognition. Some of those systems focus on interaction techniques between real and virtual worlds—for example, by projecting virtual information onto a real-world surface (e.g., tabletop, wall, floor) to allow users to interact with the information. Those systems also recognize a user's behavior via sensors such as cameras so that they can interact with the systems by using natural gestures.

Using AR markers is a highly popular method of connecting real and virtual worlds, since each marker has a unique real-world ID and position. Some systems have introduced the design of an AR marker, whereas others have provided applications by using those markers. Markers are designed with simple dot patterns, and though their design is simple, their advantages are many. Markers are easy to recognize via simple image processing [4], so that users can print them out and use them. Several methods that make 2D markers invisible have also been developed; one is implemented by using invisible LEDs, whereas another is based on transparent polarization filters [16, 17]. Since the markers have different features, users can choose from among them depending on their needs.

CONCLUSION

We have described Ghost-Hack AR by focusing on its design concept, system implementation, and interaction. Our system has two functions—face-to-face communication and multiple telepresence-based interactions—so that users can reap the benefits of each function. In demonstrations, both local and remote users felt that each other were closer than with conventional video streaming systems. Moreover, remote users could choose a telepresence system depending on the purpose. Thus, we conclude that our system is effective for network communication.

REFERENCES

1. Texai.

   http://www.willowgarage.com/pages/texai/ (2017)

2. Cisco TelePresence.

   http://www.cisco.com/c/en/us/products/collaboration-endpoints/index.html (2017)

3. Lee, M. K., and Takayama, L. "Now, I Have a Body": Uses and Social Norms for Mobile Remote Presence in the Workplace. In *ACM CHI '11* (2011*)*, 33–42.
4. Rekimoto, J. Matrix: A Realtime Object Identification and Registration Method for Augmented Reality. In *APCHI '98* (1998), 63–69.
5. Kato, H., and Billinghurst, M. Marker Tracking and HMD Calibration for a Video-Based Augmented Reality Conferencing System. In *IEEE IWAR '99* (1999), 85–94.
6. Buxton, W. Telepresence: Integrating Shared Task and Person Spaces. In *Graphics Interface '92* (1992), 123–129.
7. Venolia, G., Tang, J., Cervantes, R., Bly, S., Robertson, G., Lee, B., and Inkpen, K. Embodied Social Proxy: Mediating Interpersonal Connection in Hub-and-Satellite Teams. In *ACM CHI '10* (2010), 1049–1058.
8. Sellen, A., Buxton, B., and Arnott, J. Using Spatial Cues to Improve Videoconferencing. In *ACM CHI '92* (1992), 651–652.
9. Yankelovich, N., Simpson, N., Kaplan, J., and Provino, J. Porta-Person: Telepresence for the Connected Conference Room. In *ACM CHI '07 Extended Abstracts* (2007), 2789–2794.
10. Nakanishi, H., Tanaka, K., and Wada, Y. Remote Handshaking: Touch Enhances Video-Mediated Social Telepresence. In *ACM CHI '14* (2014), 2143–2152.
11. Leithinger, D., Follmer, S., Olwal, A., and Ishii, H. Physical Telepresence: Shape Capture and Display for Embodied: Computer-Mediated Remote Collaboration. In *ACM UIST '14* (2014), 461–470.
12. Asada, M., Hosoda, K., Kuniyoshi, Y., Ishiguro, H., Inui, T., Yoshikawa, Y., Ogino, M., and Yoshida, C. Cognitive Developmental Robotics Toward Understanding and Designing Body, Brain, and Mind. *IEEE Transaction on Autonomous Mental Development*, vol. 1, no. 1 (2009), 1–24.
13. Fernando C. L., Furukawa, M., Kurogi, T., Hirota, K., Kamuro, S., Sato, K., Minamizawa, K., and Tachi, S. TELESAR V: TELExistence Surrogate Anthropomorphic Robot. In *ACM Siggraph '12 Emerging Technology* (2012).
14. Gomes, A., Rubens, C., Braley, S., and Vertegaal, R. BitDrones: Towards Using 3D Nanocopter Displays as Interactive Self-Levitating Programmable Matter. In *ACM CHI'16* (2016), 770–780.
15. Milgram, R., and Kishino, F. A Taxonomy of Mixed Reality Visual Display. *IEICE Trans. Information and Systems,* vol. E77-D, no. 12 (1994), 1321– 1329.
16. Matsushita, N., Hihara, D., Ushiro, T., Yoshimura, S., Rekimoto, J., and Yamamoto, Y. ID CAM: A Smart Camera for Scene Capturing and ID Recognition. In *IEEE ISMAR '03* (2003), 227–236.
17. Koike, H., Nishikawa, W., and Fukuchi, K. Transparent 2-D Markers on an LCD Tabletop System. In *ACM CHI '09* (2009), 163–172.