2011 IEEE/RSJ International Conference on
Intelligent Robots and Systems
September 25-30, 2011. San Francisco, CA, USA

Mutual Telexistence Surrogate System: TELESAR4
- telexistence in real environments using autostereoscopic immersive display -

Susumu Tachi, Kouichi Watanabe, Keisuke Takeshita, Kouta Minamizawa,
Takumi Yoshida, and Katsunari Sato, *Keio University*

Abstract— **Our aim is to afford a remote person the opportunity to participate virtually in some event by using a surrogate robot to communicate with local participants while moving about freely at the venue. Accordingly, we propose a mutual telexistence surrogate robot system, called TELESAR4, which was designed and constructed by development of the following: an immersive audiovisual system; an omnidirectionally mobile robot with a robot arm and hand; an omnidirectional stereo camera system (VORTEX); a head with a retroreflective screen for embodiment of the remote participant; and a retroreflective projection system for local participants. This paper describes the development of the TELESAR4 system, whose efficacy has been verified through demonstration experiments.**

I. Introduction

AT parties, meetings, or gatherings where each participant communicates with many other participants while walking freely about the venue, it would be convenient and convivial for a person not physically present at the venue still to have the ability to participate fully in the event.

In order to attain this goal through the use of a human-robot system, it is essential for the system to have the following capabilities in addition to the fundamental remote speaking and hearing capabilities:

(1) The system must provide the remote participant with the ability to see the venue and local participants as naturally as though actually present at the event; i.e., a wide-range, life size stereo view (ideally 360°) of the event venue in real time.
(2) Face-to-face communication must be provided. This means that the local participants in the event see the face and expressions of the remote participant in real time. A local participant in front of the robot must see the full face of the remote participant, while another local participant seeing the robot from its side must see the facial profile of the remote participant.
(3) The system must allow the remote participant not only to move freely about the venue by means of the surrogate robot but also to perform some manipulatory tasks such as handshakes or gestures.

The above conditions are not satisfied by the present commercially available teleconference systems, such as the Polycom Telepresence System, or telepresence surrogate robot systems, such as TiLR, Vgo, and QB. Nonetheless, there have been several attempts to satisfy the above conditions [1] - [10].

A method for mutual telexistence based on projection of real-time images of the operator onto a surrogate robot with retroreflective projection technology (RPT) was first proposed in 1999 [7]. The feasibility of this concept was demonstrated by the construction of experimental mutual telexistence systems using RPT in 2004 [8] and in 2011 [9]. However, the demonstrated systems had only a six degree of freedom (6-DOF) head mechanism and no manipulation mechanism.

In 2005, a mutual telexistence master-slave system called TELESAR2 was constructed for the Aichi World Exposition. In addition to conventional verbal communication, actions of nonverbal communication such as gestures and handshakes could be performed, because a master-slave manipulation robot was used as the surrogate for a person [10]. Moreover, a person who remotely visited a place where the surrogate robot was located could be seen naturally and simultaneously by several people standing around the surrogate robot. Although the system had an autostereoscopic 3D display, the field of view was less than 60° in the horizontal plane and no sense of immersion was realized.

In order to move freely about an open space with several other participants, an immersive autostereoscopic 3D display is necessary, one with a visual range of at least 200° in the horizontal plane and 360° when the head rotation is ideal.

In this study, we developed the mobile mutual telexistence system TELESAR4, which is equipped with a master-slave manipulation capability and an immersive omnidirectional autostereoscopic 3D display with a 360° field of view. Thus, TELESAR4 satisfies the abovementioned three preliminary conditions that previous TELESARs could not.

II. General Design of TELESAR4 System

The first condition can be satisfied by using a head-mounted display (HMD) [11], a cave automatic virtual environment (CAVE) [12], or the telexistence wide-angle immersive stereoscope (TWISTER) [13], [14], because each of these displays can provide an immersive environment with a 360° field of view. However, the second condition requires face-to-face communication and thus necessitates a real-time capture of the facial expression of the remote participant, especially around the eyes. Thus, the visual display must be

Manuscript received March 14, 2011.

S. Tachi, K. Watanabe, K. Takeshita, K. Minamizawa, T. Yoshida, and K. Sato are with the Keio University, 4-1-1 Hiyoshi, Kohoku-ku, Yokohama, JAPAN (telephone: +81-45-564-2499; e-mail: {tachi, kouichi, takeshita, kouta, takumi, sato}@tachilab.org).

autostereoscopic, otherwise the remote participant appears with an HMD or special eye glasses, like shutter glasses, which are unsuitable for face-to-face communication.

Hence, an HMD or a CAVE is not a suitable display to satisfy both of the first two conditions. TWISTER is suitable, because this is an immersive omnidirectional full-color autostereoscopic 3D booth that was designed according to a face-to-face telecommunication concept, a so-called mutual telexistence, whereby people at distant locations can communicate as though present in the same three-dimensional space. Each user occupies a cylindrical booth, which displays live full-color 360° autostereoscopic images without the need for any eyewear. At the same time, the booth captures images of the user from every angle. With multiple booths, people at remote locations can interact with each other as though face-to-face. Moreover, the images of participants can be captured at eye level using the rail camera system of TWISTER without interrupting their fields of view.

The third condition can be satisfied by using a mobile surrogate robot with manipulation capability in addition to audiovisual capabilities. The surrogate robot must move about freely and capture 360° full-color stereoscopic scenery, which makes it necessary to design a 360° stereo camera system that can be mounted on an omnidirectional vehicle.

Figure 1 shows a system that satisfies all three conditions simultaneously by means of telexistence [15]. Remote participant A in cockpit [A] can telexist through his surrogate robot A′, while local participants B and C occupy the real environment [B]; e.g., an event venue. Cockpit [A] has the capability to display 3D images of environment [B] in 360° around participant A, who can enjoy the scenery of the real environment [B] without using any eyewear; i.e., autostereoscopically by using TWISTER. This 360° stereo image is captured by a 360° stereo camera system (VORTEX) mounted on surrogate robot A′. Using measurement and control devices installed in the cockpit, remote participant A can control his surrogate robot while having a sense of presence in the real environment [B]. The control system consists of a motion capture system, data glove, and joystick. Local participants B and C can project onto robot A′ the images taken respectively by cameras B′ and C′, thereby enabling them to see remote participant A seemingly exist inside the robot through RPT (retroreflective projection technology). The movable cameras on a rail installed outside TWISTER are controlled so that local participants B and C see remote participant A with respect to their directions from surrogate robot A′. The positions of participants B and C are measured by a motion capture system on the ceiling of environment [B], and the appropriate directional images of participant A are transmitted to participants B and C. Thus, mutual telexistence is realized.

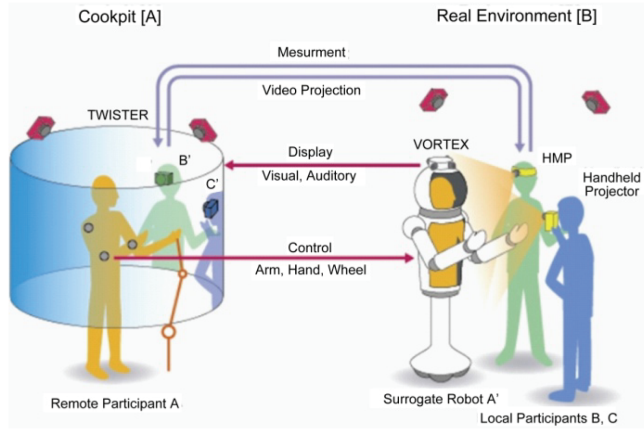


Fig. 1. Concept of robotic mutual telexistence in a real environment.

III. TELESAR4 System Configuration

Figure 2 shows the overall configuration of the TELESAR4 system: a telexistence remote cockpit system, telexistence surrogate robot system, and RPT viewer system.

The telexistence remote cockpit system consists of the immersive 360° full-color autostereoscopic display TWISTER, a rail camera system installed outside TWISTER, an omnidirectional speaker system, an OptiTrack motion capture system, a 5DT Data Glove 5 Ultra, a joystick, and a microphone.

The telexistence surrogate robot system consists of an omnidirectional mobile system, the 360° stereo camera system VORTEX, a retroreflective screen, an omnidirectional microphone system, a robot arm with hand, and a speaker.

The RPT viewer system consists of a RPT handheld projector, RPT head-mounted projector, and OptiTrack position and posture measurement system.

The three fundamental functions described in section 1 are installed in the TELESAR 4 system: i.e., (i) presentation of the real environment—e.g., party venue—including local participants’ figures and faces in the venue to a remote participant in real time to satisfy condition (1); (ii) RPT display of the remote participant’s face to the local participants from any direction at the venue in real time to satisfy condition (2); and (iii) mobility and manipulation capability of the surrogate robot by the remote participant to satisfy condition (3).

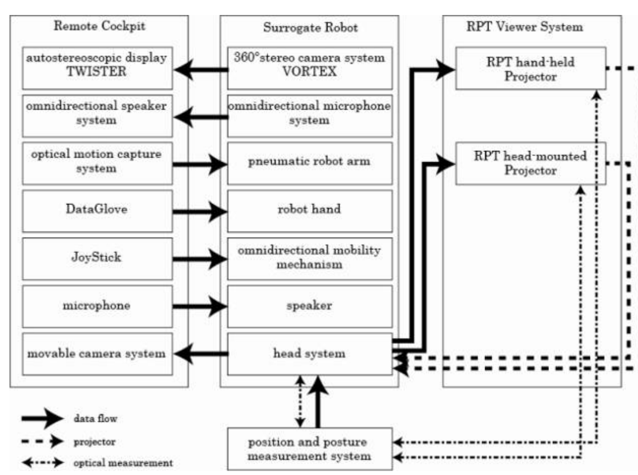


Fig. 2. System configuration of TELESAR4.

Stereo images captured by the 360° stereo camera system VORTEX are fed to a computer (PC1) mounted on the mobile robot. Real-time compensation and combination of the images are performed to transform these into 360° images for TWISTER. The processed images are sent to TWISTER and unprocessed sound from the omnidirectional microphone system is sent directly to the corresponding speaker system in TWISTER.

Cameras are controlled to move on the circular rail and to imitate the relative positions of the surrogate robot and the local participants at the venue. The images taken of the remote participant by these cameras outside TWISTER are fed to a computer (PC3) and sent to another computer (PC2) mounted on the mobile surrogate robot. These images are sent to the handheld and head-mounted projectors after being adjusted using the posture information acquired by the position and posture measurement system (OptiTrack), which consists of seven infrared cameras installed on the ceiling of the venue. These processed images are in turn projected onto the retroreflective screen atop the mobile robot. The voice of the remote participant is also sent directly to the speaker of the mobile surrogate robot.

The joystick is located inside TWISTER and is controlled by the left hand of the remote participant. The data of the joystick are fed to PC3 and sent to PC1 via the network. Based on the information received, PC1 generates and executes a motion instruction to have the mobile surrogate robot move in any direction or turn freely on the spot; i.e., omnidirectional locomotion is possible. Using the motion capture system (OptiTrack) installed on the ceiling of TWISTER, the position and posture of the right arm of the remote participant are obtained. The position and posture data are also fed to PC3 and sent to PC2 via the network. Based on the received data, PC2 controls the robot arm to imitate the arm motion of the remote participant. In addition, hand motion data acquired by the 5DT Data Glove 5 Ultra are sent to the hand of the surrogate robot to imitate the hand motion of the remote participant. Figure 3 shows a general view of the constructed TELESAR4 system.

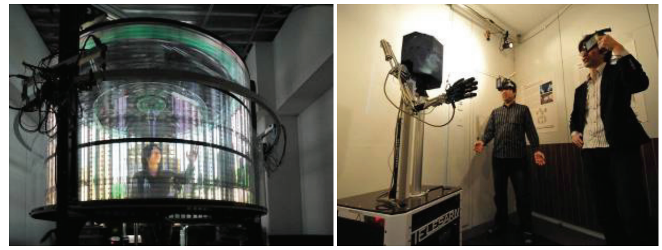


Fig. 3. General view of TELESAR4 system.

IV. Functional Experiments on TELESAR4

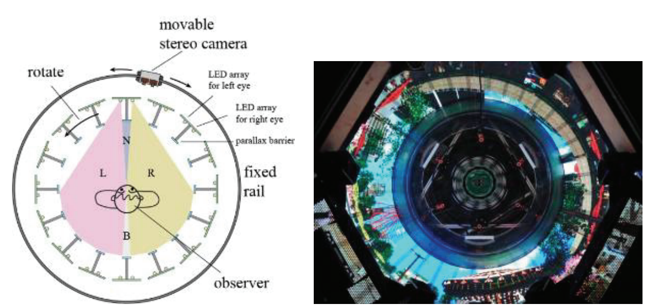
A. Capture and display of 360 degree stereo images

TWISTER (telexistence wide-angle immersive stereoscope) is an immersive omnidirectional full-color autostereoscopic 3D booth that was designed for the face-to-face telecommunication concept of mutual telexistence, whereby people in distant locations can communicate as though present in the same three-dimensional virtual space. Each user occupies a cylindrical booth, which displays live full-color 360° autostereoscopic images without the need for any eyewear. At the same time, the booth captures images of the user from every angle. With multiple booths, people at remote locations can interact with each other as though face-to-face.

This concept was proposed in 1996 [13], and five prototype models, TWISTERs I through V, were constructed on that basis. TWISTER V displays binocular stereo images using 36 display units, each consisting of two LED arrays (one per eye) and a parallax barrier that rotates at high speed (1.7 rps). This system captures 3D images of the observer by simultaneously employing stereo cameras installed outside the rotating cylinder. Thus, the observers inside two TWISTERs can view 3D images of each other in real time [14]. The horizontal resolution is 3168 pixels for a 360° horizontal field of view, and the vertical resolution is 600 pixels for a 61.9° vertical field of view. Figure 4(a) shows the principle of the rotating parallax barrier method and rail camera attachment, while Fig. 4(b) shows a general view of TWISTER.

Since TWISTER is the only immersive full-color autostereoscopic display with a 360° omnidirectional view, its use as the visual display is quite reasonable. In order to use TWISTER as the control cockpit of a mobile surrogate robot, it was necessary to develop a compatible 360° stereo camera system.

As the capture device of the omnidirectional image for TWISTER, we developed the 360° stereo camera system VORTEX. The principle of the 360° stereo camera system is shown in Fig. 5(a). As shown in the left half of Fig. 5(a), the eyes of a participant move around the rotational axis as the head is turned. The eyes lie on the circumference of a circle, maintaining the same distance from each other. If we arrange an infinite number of stereo camera pairs along the circumference of the circle, as implied by the right half of Fig. 5(a), then we attain the ideal omnidirectional stereo camera system. The VORTEX system arranges eight stereo camera pairs (16 individual cameras) on the circumference of a circle with a 120 mm radius to match the head and eye arrangement of the average person. The distance between



a b

Fig. 4. (a) Principle of rotating parallax barrier method and movable stereo camera attachment. (b) General view of TWISTER.

each two cameras was set to 65 mm, which is the average distance between two human eyes. Figure 5(b) shows the design concept of VORTEX, Fig. 5(c) shows the dimensions of the VORTEX design, and Fig. 5(d) shows the constructed VORTEX.

For each camera, the Firefly MV (Point Grey Research) with a lens of focal length 2.9 mm was used. Its horizontal and vertical fields of view are 53.2° and 67.8°, respectively. The resolution is 480 × 640 pixels (horizontal × vertical), which fits the resolution of TWISTER, i.e., 3168 × 600 pixels, through the use of eight camera images combined by blending.

Each camera is taken offline to conduct the calibration of the intrinsic parameters; i.e., the lens distortion and the relative orientation between the lens and CCD. The transformations from camera coordinates to cylindrical TWISTER coordinates were executed in real time as were the realizations of high dynamic range images that considered the lighting environment. Once processed, the images from the cameras were displayed on TWISTER.

A preliminary directional discrimination experiment using cardioid unidirectional microphones (Radius 45 HD, Monitor Audio Ltd.) and speakers revealed a significant difference between systems composed of four and six of the microphone and speaker combinations, while there was no difference between systems composed of six and eight.

Thus, six microphones for the omnidirectional sound system were sufficient for our purposes. The six-microphone system was designed and placed on top of VORTEX, as shown in Fig. 5(e). The six speakers were placed on the stationary part of TWISTER at positions corresponding to those of the microphones.

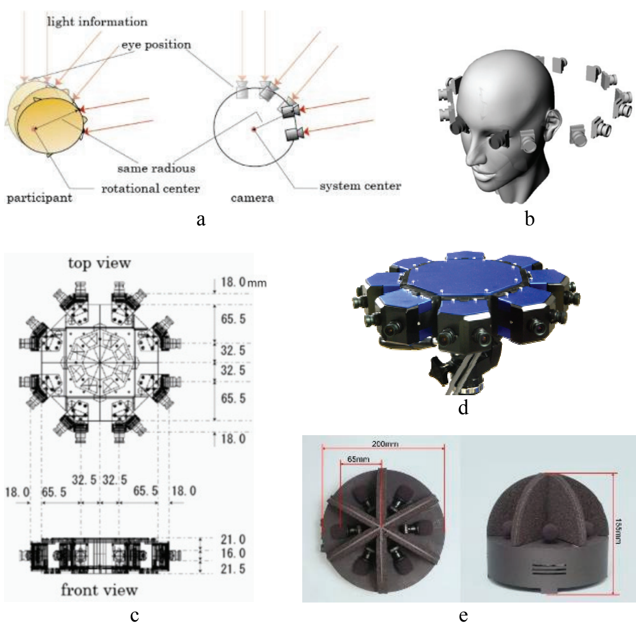


Fig. 5. 360° stereo camera system VORTEX: (a) human eye trajectory as head rotates; (b) design concept of VORTEX; (c) dimensions of VORTEX design; (d) constructed VORTEX; and (e) microphone arrangement.

As a demonstration experiment, we selected a room (90 m^2) as the remote environment and introduced the mobile surrogate robot. The robot moved about the room to capture 360° images and sound. More than ten men and women, aged between 23 and 64 years, interacted with this system. All reported an ability to experience the realistic 360° stereo color scenery and to feel a strong presence as though the person and scenery in the remote place were in the immediate vicinity. Figure 6 shows the scenery taken by VORTEX.

B. Projection of remote participant on surrogate robot

In order to realize mutual telexistence, a technology utilizing a retroreflective material as the projection surface has been pursued. The original concept was proposed in 1999 [7]. In order to display the face of the remote participant with the TELESAR4 system, this method for projecting the image of the human on the robot is used. Local participants at the venue use the RPT handheld projector [9] or the RPT head-mounted projector with a half-mirror to view a real-time image of the remote participant. The principle of RPT is shown in Fig. 7(a), while the structures of the RPT handheld and head-mounted projectors are shown in Figs. 7(b) and 7(c), respectively.

The head of the mobile surrogate robot is covered with retroreflective material, which is used as a screen. To capture the image of the remote participant, stereo cameras are installed on a rail outside TWISTER. These stereo cameras are controlled to imitate the relative positions of the mobile surrogate robot and the local participants. Each stereo camera corresponds to one local participant at the venue. The camera can look through to the remote participant as TWISTER rotates, as shown in Fig. 3. Thus, the camera serves to capture the image of the remote participant from any perspective outside the display. An active projection corresponding to the movement of the local participant is enabled by controlling the camera unit according to position data on the participant and the mobile surrogate robot.

The robot captures omnidirectional images and sound while the rail camera unit outside TWISTER captures images

Fig. 6. Scenery taken by VORTEX.

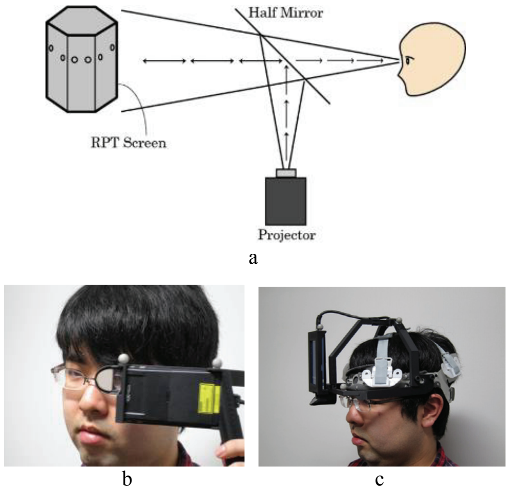


Fig. 7. (a) Principle of RPT. (b) RPT handheld projector. (c) RPT head-mounted projector.

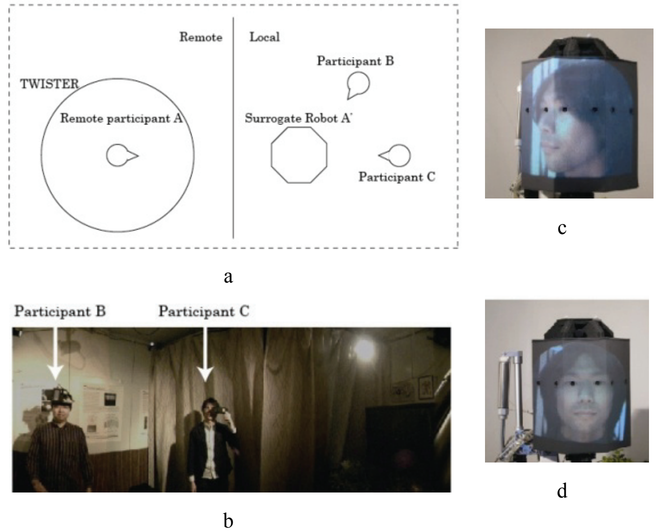


Fig. 8. Experimental results for projection of remote participant on RPT head: (a) remote and local arrangements; (b) local participants at venue; (c) remote participant image from head-mounted projector; and (d) remote participant image from handheld projector.

of the remote participant to project on the head of the surrogate robot. At the same time, the voice of the remote participant from the microphone is sent to the speaker on the surrogate robot. Figure 8 shows the experimental results for RPT viewing. In particular, Fig. 8(c) shows a front view and Fig. 8(d) shows a side view. The black dots in Figs. 8(c) and 8(d) correspond to holes in the retroreflective screen head at the lens positions of the 360° stereo camera system. Figures 8(c) and 8(d) show that two of these 16 holes coincide with the eye positions of the remote participant on the screen; i.e., the local participants are seen from the eye positions of the surrogate robot. Thus, face-to-face communication was simulated. When each of the local participants looked into the eyes of the remote participant on the robot head, virtual eye contact was thereby established.

C. Mobility of surrogate robot with an arm

Omnidirectional mobility is necessary for a surrogate robot to move about freely on a floor crowded with people, such as an event venue. Therefore, a truck-type mobile robot that can move omnidirectionally was designed and developed. This robot has two drive wheels and one passive wheel. The drive wheels are controlled by motors to adjust the speed and steering angle. Figure 9 shows the arrangement of these wheels. We constructed a locomotion model of the robot and controlled the drive wheels through software that applies this model to achieve the omnidirectional motion.

A joystick was selected as the input device because its operation is intuitive. The control of the drive wheels for rigid movements of translation and rotation is applied using the locomotion model through rotations of the instantaneous center. In addition, the model considers the range of steering motion and applies an inverse wheel rotation when the desired motion exceeds the steering range.

In order to achieve the embodiment of the remote person in the surrogate robot, an arm is necessary to get active motion in addition to mobility. Since contact between a person and the surrogate robot occurs frequently, the robot arm must be safe. A pneumatic robot arm was used for the TELESAR4 system. This robot arm works by using the compressibility of air and has mechanical compliance [16], which is effective when the robot interacts with a person through handshaking at an event. Even if the control program malfunctions, there is little possibility of the robot harming someone. The arm has 7 DOF and a mass of 2.6 kg. Its upper-arm and forearm lengths are 300 mm and 220 mm, respectively.

A robot hand with five fingers, electrically driven by DC motors, was attached to the end of the pneumatic arm. Its thumb has 3 DOF, while the four remaining fingers have 1 DOF each, and the hand itself has 1 DOF for abduction, yielding a total of 8 DOF. The hand is similar in size to that of a typical human, with length 185 mm, width 100 mm, and thickness 35 mm. The mass of the hand is 0.5 kg. All mechanisms such as motors, gears, and encoders are packed inside the hand, which is connected by a single cable to its control system; i.e., a servo amplifier and computers.

To achieve locomotion of the surrogate robot and operation of the robot arm simultaneously, the left hand of the remote participant controls the movement of the surrogate robot through the joystick, while the right hand controls the robot arm and hand through the OptiTrack system and the data glove. Specifically, the pneumatic robot arm is controlled by

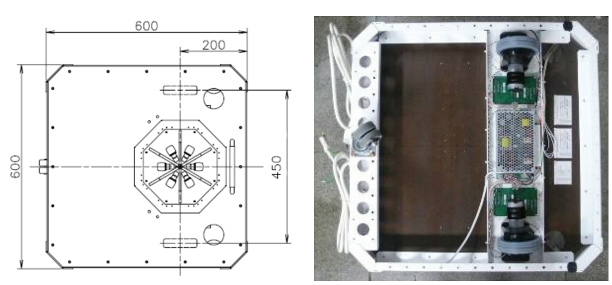


Fig. 9. Omnidirectional mobility mechanism.

the motion capture system (OptiTrack) installed within TWISTER, while the hand is controlled by the data glove. Figure 10 shows examples of handshaking and gesturing with the pneumatic robot arm and electrically driven robot hand.

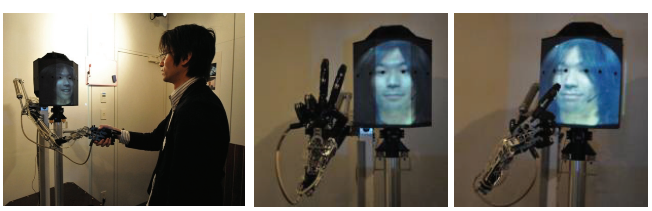
Figure 11 shows a demonstration experiment. A remote participant inside TWISTER joined a gathering by means of his surrogate robot and moved freely in a room with two local participants, who were able not only to see his face in real time through RPT handheld and head-mounted projectors but also to communicate with him naturally. The local participant in front of the robot saw the full face of the remote participant, while the local participant on the right side of the robot saw his right profile at the same time. The remote participant was able not only to see and hear the environment as though he were there but also to communicate face-to-face with each local participant. Moreover, he was able to shake hands and express his feelings with gestures.

V. Conclusions

We propose the mutual telexistence surrogate mobile robot system TELESAR4 to afford a remote person the opportunity to virtually participate in some event, such as a party, a meeting or a gathering, by using a surrogate robot to communicate with local participants of the event while moving about freely in the venue. TELESAR4 is constructed from an immersive audiovisual system, an omnidirectional mobile robot with a robot arm and hand, and RPT projection of the remote participant's image to the robot.

The TELESAR4 system was implemented as a prototype, and feasibility experiments confirmed that the system provided the remote participant with the ability to see the event venue and local participants as naturally as though at the event; i.e., a life size 360° wide-range stereo view of the event venue was observed in real time.

It was also confirmed that face-to-face communication was provided, as the local participants at the event were able to see the face and expressions of the remote participant in real time.



a b

Fig. 10. (a) Handshake. (b) Gestures.

a b

Fig. 11. (a) Remote participant virtually attending an event. (b) Local participants and surrogate robot in the venue.

A local participant in front of the robot saw the full face of the remote participant, while another local participant seeing the robot from its side saw the facial profile of the remote participant.

It was further confirmed that the system allowed the remote participant not only to move freely about the venue by means of the surrogate robot but also to perform some manipulatory tasks such as a handshake and several gestures.

Acknowledgment

This work was partly supported by the Strategic Information and Communications R&D Promotion Program (SCOPE).

References

[1] S. Tachi, K. Tanie, K. Komoriya, and M. Kaneko: Tele-Existence (I): Design and Evaluation of a Visual Display with Sensation of Presence, In *Proc. of the 5th Symposium on Theory and Practice of Robots and Manipulators (RoManSy '84)*, pp. 245-254, Udine, Italy, (Published by Kogan Page, London), June 1984.

[2] S. Tachi: Real-Time Remote Robotics — Toward Networked Telexistence, *IEEE Computer Graphics and Applications*, vol. 18, pp. 6-9, 1998.

[3] R. Tadakuma, Y. Asahara, H. Kajimoto, N. Kawakami, and S. Tachi: Development of Anthropomorphic Multi-D.O.F. Master-Slave Arm for Mutual Telexistence, *IEEE Transactions on Visualization and Computer Graphics*, vol. 11, no. 6, pp. 626-636, 2005.

[4] M. Shoji, K. Miura, and A. Konno: U-Tsu-Shi-O-Mi: The Virtual Humanoid You Can Reach, *Siggraph2006 Emerging Technologies*, 2006.

[5] K. Watanabe, I. Kawabuchi, N. Kawakami, T. Maeda, and S. Tachi: TORSO: Development of a Telexistence Visual System Using a 6-d.o.f. Robot Head, *Advanced Robotics*, vol. 22, pp. 1053-1073, 2008.

[6] P. Lincoln, G. Welch, A. Nashel, A. Ilie, A. State, and H. Fuchs: Animatronic Shader Lamps Avatars, *8th IEEE International Symposium on Mixed and Augmented Reality*, pp. 27-33, 2009

[7] S. Tachi: Augmented Telexistence, *Mixed Reality*, pp. 251-260, Published by Springer-Verlag, 1999.

[8] S. Tachi, N. Kawakami, M. Inami, and Y. Zaitsu: Mutual Telexistence System Using Retro-Reflective Projection Technology, *International Journal of Humanoid Robotics*, vol. 1, no. 1, pp. 45-64, 2004.

[9] K. Takeshita, K. Watanabe, T. Yoshida, K. Minamizawa, and S. Tachi: Transmission of Existence by Retro Reflective Projection Technology Using Handheld Projector, In *Proc. of IEEE International Symposium on Virtual Reality Innovations (ISVRI 2011)*, pp. 71-74, 2011.

[10] S. Tachi, N. Kawakami, H. Nii, K. Watanabe, and K. Minamizawa: TELEsarPHONE: Mutual Telexistence Master-Slave Communication System Based on Retroreflective Projection Technology, *SICE Journal of Control, Measurement, and System Integration*, vol. 1, no. 5, pp. 1-10, 2008.

[11] I. Sutherland: A Head-mounted Three Dimensional Display. In *Proc. of Fall Joint Computer Conference, AFIPS Conf. Proc.*, vol. 33, pp. 757-764, 1968.

[12] C. Cruz-Neira, D. Dandin, and T. Defanty: Surround-screen Projection-based Virtual Reality: The Design and Implementation of the CAVE, In *Proceedings of SIGGRAPH '93*, pp. 135-142, 1993.

[13] S. Tachi, T. Maeda, Y. Yanagida, M. Koyanagi, and H. Yokoyama: A Method of Mutual Tele-Existence in a Virtual Environment, In *Proc. of the 6th International Conference on Artificial Reality and Tele-Existence (ICAT'96)*, pp. 9-18, Japan, 1996.

[14] S. Tachi: TWISTER: Immersive Omnidirectional Autostereoscopic 3D Booth for Mutual Telexistence, In *Proc. of ASIAGRAPH 2007*, vol. 1, pp. 1-6, 2007.

[15] S. Tachi: Telexistence, *World Scientific*, ISBN-13 978-981-283-633-5, 2010

[16] K. Watanabe, H. Nagayasu, N. Kawakami, and S. Tachi: Mechanical Compliance Control System for a Pneumatic Robot Arm. In *SICE Annual Conference 2008*, pp. 2789-2794, 2008.