# 5-6 An Interactive Communication System Implementation on JGNⅡ

**HOSOYA Eiichi, HARADA Ikuo, SATO Hidenori, OKUNAKA Junzo, TANAKA Takahiko, ONOZAWA Akira, and KOGA Tatsuzo**

A novel interactive communication system over JGNⅡ is proposed. Its most important component, named as the Mirror Interface, provides users a virtual shared space using real-time video and ubiquitous information such as the one from sensor tags. The shared space provides users an interactive environment where they can communicate each other seamlessly beyond the border of real and virtual as well as that of local and remote, enabling a natural non-verbal communication using gestures and positioning regarding meeting partners in both local and remote locations. Such information is implicitly utilized and is crucial in human communications. The shared space is shown as a metaphor of a mirror. The images from both locations are integrated into one in which users and objects seem to be virtually located in the same space. Icons of operable objects and accessible information are also displayed. Those icons are to be pointed at for operations by so-called “virtual touch”. In this report, the overview of this system, its implementation on JGNⅡ and related applications are described.

***Keywords***
Human-Computer Interaction (HCI), Video conference, Mirror interface, Shared space, Pointing

## 1 Introduction

Advances in research and development of broadband networks have led to the proposal of various services using broadband content, such as those dealing with video data. Examples of such services include videotelephony and videoconferencing, but these applications are not catching on as widely as many people had expected. Meanwhile, in line with the u-Japan concept, we have seen great progress in the area of ubiquitous computing technologies, which gather small fragments of information of various types distributed throughout real space, as in the case of sensor and tag information. A variety of services are in the process of development based on such technologies.

This paper proposes a new human interface that applies both broadband-based visual information and ubiquitous tag and sensor information for use in communications. In person-to-person communications, it is commonly held that the feeling of sharing a single space, in which natural non-verbal communications can take place based on gestures, body language, and facial expressions, is as essential as direct linguistic expressions using voice and text information. To accomplish this sort of communication, it is important to produce an environment in which the communicating parties feel as if they were present in the same space at the same time.

There have been a number of conventional approaches to the construction of such a shared space, including a system[1] enabling operation of a shared application through display screens (showing the users as if they were facing each other through a window), systems[2]-[4] that extract an image from one location (using silhouette or chroma-key techniques) for integration into a two-dimensional space viewed by the other party, a system[5] that generates a shared space by extracting the human body using 3D cameras and HMD, and a system[6] enabling natural instruction and body arrangement.

The system we propose in this paper, on the other hand, transfers various events and phenomena in the real world bidirectionally using both broadband-based visual information and ubiquitous information (such as that contained in tags) and constructing a virtual shared space based on the metaphor of a mirror. Users separated by distance can communicate smoothly using this shared space, in which information from the real world and information from the virtual world is integrated and expressed seamlessly. As a result, each user can provide instructions and conduct various operations in the other party's space[7]-[9]. We refer to the human interface forming the key component of this system as the "mirror interface". Here we will introduce a method of communication via the mirror interface and describe the construction of a system using this interface over the JGNII network.

We provide an overview of the mirror interface in section **2** below and describe its method of use over JGNII in section **3**. We explain some of the presently assumed applications of the mirror interface in section **4**, and offer a summary and prospects for future development of this technology in section **5**.

## 2 Overview of the mirror interface and related works

### 2.1 Overview of the mirror interface

The mirror interface is a real-world-oriented interface that enables people (users) to operate items (equipment, devices, items, etc.) existing in real space as well as information located in computers and on a network and databases. We are proposing the use of the mirror interface to construct a type of "desktop" by projecting real-world images onto the screen. By displaying the user and his or her counterpart in a distant location together as part of the real world, the mirror interface can serve as a communication tool for smooth dialogue, with none of the inconvenience of separation over distance. In short, the mirror interface offers a communication tool providing seamless integration between the real world and the virtual world, as well as between remote and local locations.

Fig. 1 shows the system for implementation of the mirror interface. A large display unit resembling a mirror is placed in front of the user. The compact camera installed at the center of this display faces the user and cap-

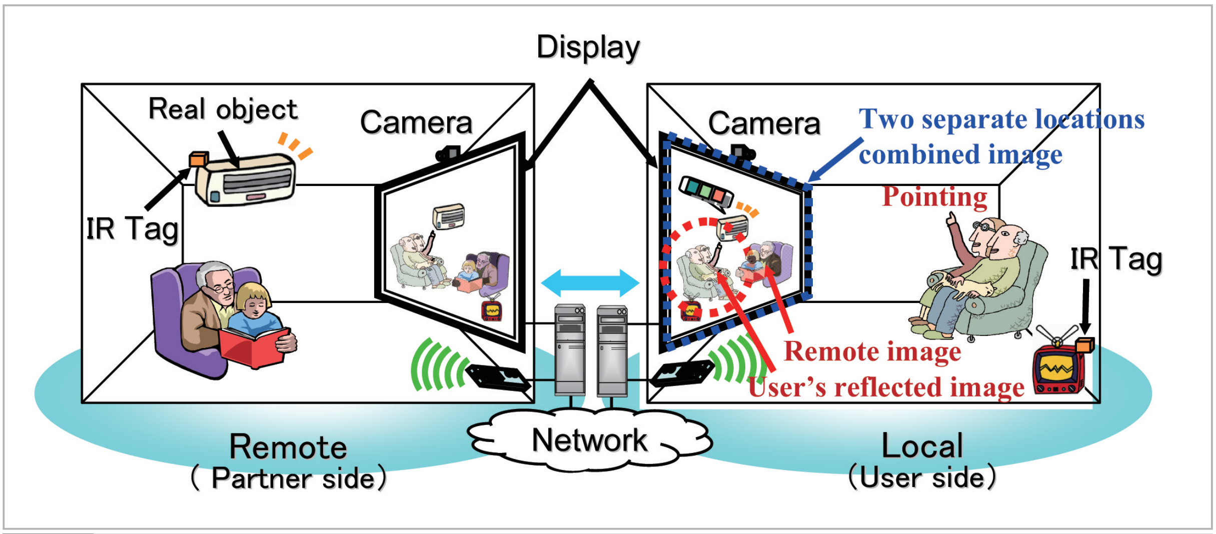



**Fig.1** Overview of the mirror interface

tures images of the user and his or her immediate surroundings. This image is converted to a mirror image and shown on the display. An identical system is set up on the remote side to capture an image of the corresponding user and his or her immediate surroundings, converting this image for mirror display on the remote side. These two images are then transmitted over the network, with the image of the remote location superimposed as a translucent image on the local display. As a result, items and persons at the two locations are integrated into one “room”, giving both users the impression of being in the same space, looking in a mirror.

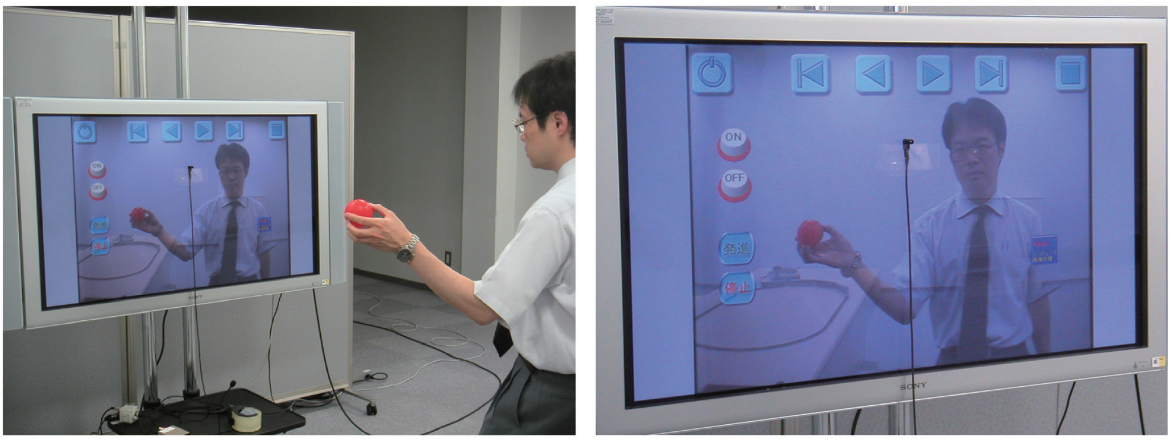
The display not only shows a mirror image of the user, but also provides information relating to items contained in the image; icons to operate these items are superimposed on the corresponding devices. The user can interact with the information and icons using his or her hand as a pointer.

As described above, the mirror interface recreates the real world on a desktop. While the ordinary GUI uses a mouse (cursor) as a pointing device, enabling execution of an action by selecting an icon with the pointer, the mirror interface relies on the user’s hand as a pointer, allowing for operation of equipment in the real world through manual indication of an item on the desktop. In this sense, the mirror interface can be considered a real-world-oriented GUI.

### 2.2 Real-world-oriented desktop

A real-world-oriented desktop constituting the similarly described GUI provides the user not only with information stored in a computer (offered by conventional GUIs) but also provides an interactive space incorporating information from the real world. The real-world-oriented desktop displays the user and the surrounding area and a real-time image (in moving pictures) of the remote side within a mirror image. An icon is placed on each of the operational devices and objects in the computer displayed on the desktop and these icons are used to control the operations of these devices and objects. By pointing at these icons, the user can operate not only computer functions but can also activate services in the real world. Since the user is also present in this image (i.e., in the interactive space), he or she gains a natural sense of operating these devices, which appear to be arranged in his or her immediate space. Figure 2 shows an example of a desktop created by the mirror interface.

On a real-world-oriented desktop, operational devices and their icons must be arranged in corresponding positions on the screen (i.e., in a two-dimensional space). If a device is stationary in real space, its position can be registered in the system in advance for




(a) User and the real-world-oriented desktop. (b) Outlook of the real-world-oriented desktop.

**Fig.2** *Real-world-oriented desktop*

positional correlation. If a device may be moved, its position on the screen must be detected. Accordingly, a system[10] that detects the position of tags and IDs within the screen using an infrared camera is used for positional tracking of operational devices. The authors plan to introduce this system at the Tsukuba JGNII Research Center in fiscal 2006 and to incorporate it into the development of motion-based applications.

### 2.3 Pointing function using user's reflected image

On a conventional desktop, the mouse and keyboard serve as user interfaces to initiate operations. In addition to these tools, button switches, touch panels, and other devices are commonly used for devices within arm's reach. For devices beyond immediate reach, remote controllers are used.

To enable natural pointing to items by the user, several types of interface are proposed, each of which uses a 3D camera to detect the direction in which the user is pointing[11][12]. These methods are investigated to establish functions involving direct pointing toward items within a room. The paper[13] shows a method based on a configuration in which the user stands in front of the screen and points at a location within the area shown.

Unlike methods proposed by others, the mirror interface enables the user to operate equipment in the real world simply by pointing at the equipment shown on the real-world-oriented desktop. The real-world-oriented desktop also shows a mirror image of the user. The mirror interface detects and determines the intended "pointing" when the hand of the user's mirror image overlaps with the icon of the target equipment on the real-world-oriented desktop.

The superiority of the method of "touching" an item within the image using the reflected hand was verified in an experiment using human subjects, as described in paper[14]. In this experiment, the time spent in pointing at a specified position on the screen using the superimposed self-image was compared with the time spent in pointing by a mouse. The results indicated that the reflected image was an effective pointing tool even for general users lacking particular IT awareness. This suggests that a system capable of robust recognition of the user's hand position can in fact be used as a practical interface.

As stated above, the pointing mechanism used in a real-world-oriented interface must be able to detect the overlapping of the user's reflected image and a target item within the image. The current mirror interface detects a marker of a specific color held in the user's hand instead of the hand itself. When the user holds the marker over an icon, the corresponding equipment begins operation. For a more complex command, the mirror interface can display the icon in the form of a hierarchical menu. After an operational command is executed, the menu disappears and the screen returns to the normal dialogue mode to enable input of the next command. In terms of usability, some argue that pointing with the bare hand would be better than requiring the use of a colored marker. Accordingly, we are investigating a method using CV (computer vision) technology[15].

In Fig. 3, an icon on the screen is touched. The user is holding his hand (with a marker) over the button-type icon near the right edge. In our experimental system, this operation activated the light (in the lower-right corner of the screen) on the remote side.

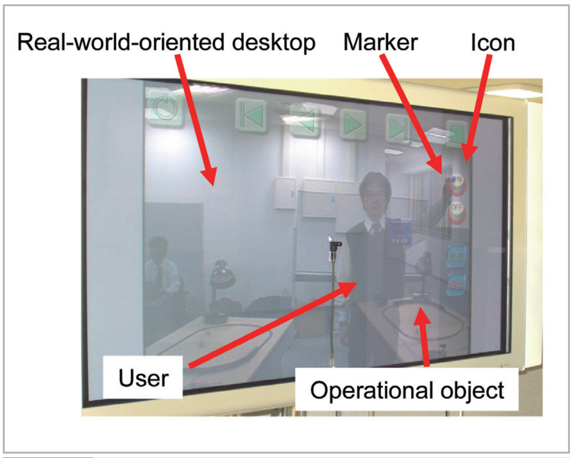



**Fig.3** *Example of equipment operation using the mirror interface*

## 2.4 Composition of the user image with the image on the remote side

When the mirror interface is used, the reflected image of the user becomes a pointing device and is also used for communication with the partner on the remote side. While visually noting the gestures of the other displayed on-screen via the mirror interface, the user and the partner can communicate through gestures while also operating the equipment or services shown on the real-world-oriented desktops.

The image of the partner and the room captured on the remote side is integrated with the image of the user and his or her room by converting both images to translucent images and superimposing them. The combined image can be obtained by changing the pixel value $p$ at each pixel to "$p = \alpha \bullet p_A + (1 - \alpha) p_B$", where $p_A$ and $p_B$ are the pixel values on the local side and remote side and $\alpha$ is the degree of translucency. This superimposition process is performed independently at the two locations to generate separate images. Figure 4 shows an example of the real-world-oriented desktop created by the superimposition process. In this example, two individuals on the local side face a person on the remote side, and items (such as shelves and other features) at both locations are displayed on a single screen. In other words, the two rooms are combined into one space.

As shown in the example in Fig. 4, the superimposed translucent images show all items in both locations without hiding overlapped items, thus allowing visual recognition of all equipment and devices present at both locations. The degree of translucency can be varied during communication, enabling the adjustment of appearance in accordance with the shift of focus in the conversation. Since images are superimposed independently at the two locations, the image can be adjusted separately at each location to suit the specific configuration of each site.

## 2.5 Communication in a shared space

As an example of the use of the shared space created by the mirror interface for communication, we assumed a meeting in which a visual presentation such as a slideshow would be given. Figure 5 shows an example in which a document is overlapped on the top of the screen displayed using the mirror interface system. The display serves as a slideshow screen. By pointing at an icon, users at either location can turn the page forward or backward.

In Fig. 5, the user and partner at two separate locations are conversing while pointing with their fingers at the map shown in the slideshow. Since the superimposed image is a

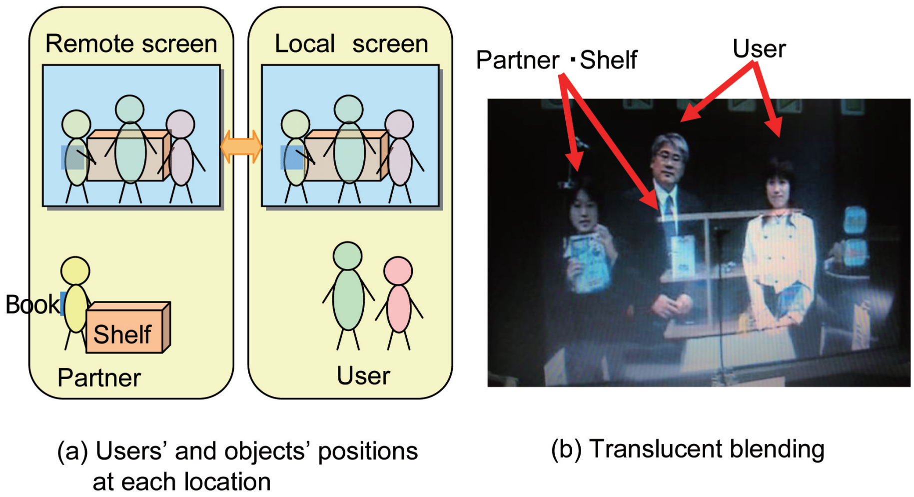



**Fig.4** *Process for creating superimposed translucent image*

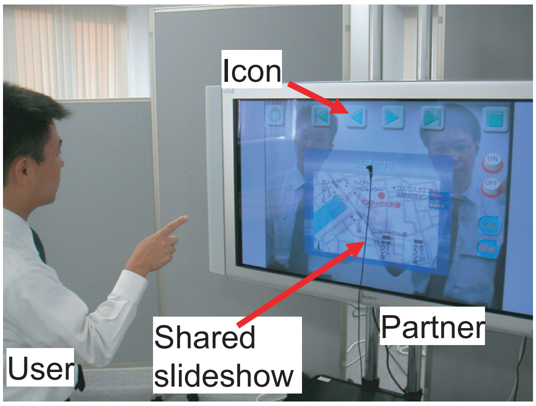


**Fig.5** Example of shared space: slideshow imposed on a real-world-oriented desktop

mirror image of both the local side and remote side, each party can see the other pointing in the correct orientation on his or her own screen. Demonstrative pronouns such as “this” and “that” are understandable to both parties, and phrases that express relative positions such as “upper right” and “upper left” indicate the same positional orientation. When images are shown in normal orientation, as in videoconferencing systems, the person on the remote side pointing at the “upper-right” location appears to be pointing at the “upper-left” location on the local side, causing confusion and inconvenience in many instances. This problem is prevented by converting the image on the remote side to a mirror image.

Moreover, since the images of both locations overlap, the users can communicate using both hands—pointing, for example, or via other gestures and body language. In addition, by setting the first-come-first-served rule for the operation of equipment and devices on the local and remote sides, it is possible to obtain or transfer operating privileges smoothly based on dialogue between the user and partner (i.e., through oral and behavioral indications). This can eliminate the cumbersome procedure of transferring the mouse operation privilege, which is normally required when an application is shared through a conventional videoconferencing system.

Because the mirror interface allows two parties to share the same “space”, including their reflected images, the user can see how his/her own gestures, body language, and facial expressions are seen by the partner; therefore, the user can make appropriate adjustments during the conversation. The user can also maintain the same positional relation and distance from the partner as he would in a real-world situation. In our experiments, we even observed a case in which the hand of the user on the local side moved accidentally and “contacted” the body of the person on the remote side and the person on the remote side reacted by jerking his body away from the “hand”. We believe that the space recreated by the mirror interface thus provides the user with a significant sense of realism.

## 3 JGNII network

### 3.1 Network configuration

Using JGNII, we constructed a system that used the mirror interface to connect two locations. We set up two terminal environments, each with a PC, a PDP, and other devices, at the Tsukuba JGNII Research Center, and these were connected via a circuit routed through Kitakyushu or Otemachi. We then conducted an operation verification experiment (Fig. 6). The systems at both locations were constructed on local networks, and were connected from separate ports via JGNII connection hubs (compatible with Gigabit Ethernet). In addition, the PC set up on the local side to distribute slideshow images was connected to the JGNII connection hub.

This JGNII network appeared as an Ethernet network when viewed from the hub, and we set IPv4 local addresses for TCP, UDP/IP communications between PCs.

### 3.2 Transfer of images, item coordinates and commands

The system we developed and used in our experiments integrated data from images captured by terminal cameras and voice data picked up by microphones, converted this information into DV data format, and assem-

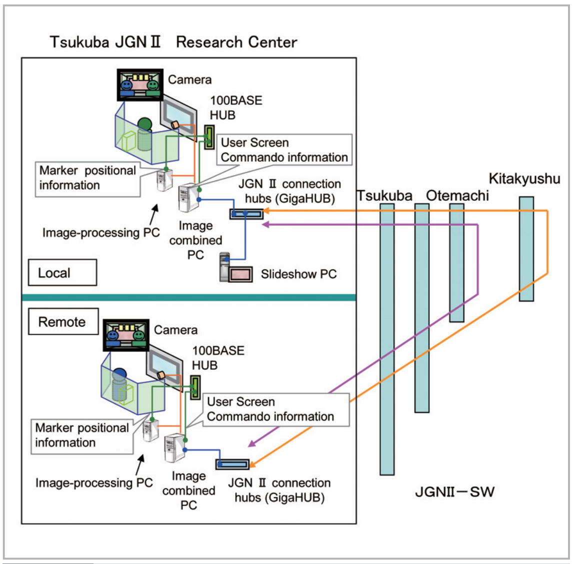



**Fig.6** *JGNⅡ network connection for mirror interface experiments*

bled the data block into fixed-length UDP packets for transfer.

When the item selected by the user is located on the remote side, the mirror interface must perform the necessary operations on the remote side. Furthermore, coordinate data of registered items must be exchanged when necessary in order to reflect the relocation of items. The data used for this purpose is very small in volume relative to image data, and the transmission frequency can be about the same as the frame rate required to ensure the proper display of item movement. As a result this data represents a very small proportion of total data transferred. To prevent packets from being lost and to ensure that all commands reach the appropriate destination, TCP/IP was used for data transfer and reception.

In the case of the slideshow described earlier, data generated in Microsoft PowerPoint was multicast to the terminals at the two locations by means of an image-distribution PC connected to the JGNⅡ network. The slideshow images were captured in preset intervals and converted to DV format for transmission. Owing to this specification, a paper document can be superimposed on top of a real-world-oriented desktop using a document camera for on-screen display.

We constructed the system as explained above and verified operation. Although some communication delays took place due to image-superimposition processing and CODEC processing, smooth dialogue was enabled with no significant problems. The delay time in the JGNⅡ circuit was approximately 10 ms each way, through Kitakyushu according to actual measurements of the video distribution time conducted in a separate experiment. In the experiment on mirror interface operation, we observed no delays that would result in adverse effects preventing normal conversation.

## 4 Applications

The concept of the mirror interface and its implementation were described in the above sections. This paper also discussed the advantages of communication using the mirror interface in the context of a slideshow-based example exchange.

Below we list areas of application in which the advantages of the mirror interface can be maximized. Based on the potential uses of the display of superimposed images and the equipment-operating interface, we anticipate the following three main areas of application.

(1) Use of display with superimposed document images
- Video conferencing, interactive remote lectures, remote counseling: in these applications presentations are conducted from a remote location. These uses take advantage of the easy understanding of demonstrative pronouns such as “this” and “that” (described earlier in this paper).

(2) Use of display with superimposed image of remote location
- Exercise guidance, remote guidance for physical therapy: instruction related to postures and motions
- Workshops: The instructor can “participate” in events (shared spaces) with multiple groups held at different remote locations. These applications take advantage of the superimposition of translucent images.

(3) Use of equipment-operating interface

- Remote equipment operation: monitoring and remote maintenance of equipment and devices at remote locations
- Walkthroughs: Walkthrough in the real world by enabling the local side to control the selection and operation of cameras installed on the remote side
- Presentations: Enhanced presentations in showrooms and other sites

In some situations, the three areas of use described above can be combined for even more effective communications. We plan to construct some of these situations and to conduct verification experiments, with the participation of local communities, in order to examine the full effectiveness of communication using the mirror interface.

## 5 Conclusions

This paper described research on an interactive communication system constructed on the JGNII network currently underway at the Tsukuba JGNII Research Center. We explained the functions achieved to date with the current system and also presented our view of possible applications.

We plan to expand the current two-location communication system to a multi-location system connecting three or more sites, and to conduct further research on fundamental technologies required for a bare hand pointing interface. We also intend to assess system characteristics through various experiments testing usability. Furthermore, we intend to promote the development of useful applications and to examine ways to apply the results of this R&D through specific experiments conducted in cooperation with local communities.

## Acknowledgments

We would like to thank Ms. Miki Kitabata at NTT Microsystem Integration Laboratories (presently Plala Networks Inc.) for her valuable cooperation in constructing the system, Dr. Hisao Nojima at NTT Microsystem Integration Laboratories (presently a Professor in the Faculty of Social Innovation at Seijo University) and Ms. Noriko Shingaki (Assistant Professor in the Faculty of Social Innovation at Seijo University) for their very informative discussions on user interfaces, and Dr. Haruhiko Ichino at NTT Microsystem Integration Laboratories for his kind guidance on research procedures. We would also like to express our appreciation to Professor Koji Munakata and Assistant Professor Sayuri Hashimoto at the University of Tsukuba for their generous guidance and support.

## *References*


1 H. Ishii and M. Kobayashi, "ClearBoard : A Seamless Media for Shared Drawing and Conversation with Eye-Contact", Proc. CHI'92, ACM SIGCHI, pp. 525-532, 1992.

2 M. W. Krueger, T. Gionfriddo, and K. Hinrichsen, "VIDEOPLACE -- An Artificial Reality", Proc. CHI-85 Human Factors in Computing Systems, pp. 35-40, 1985.

3 M. W. Krueger, "Artificial Reality II", Addision-Wesley Publishing Company. Inc., pp. 33-64, 1990.

4 O. Morikawa and T. Maesako, "HyperMirror:a Video-Mediated communication style that includes reflected images of users", IPSJ Tech. Report HI, No. 72, pp. 25-30, 1997. (in Japanese)

5 C. Cruz-Neira, D. J. Sandin, and T. A. DeFanti, "Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE", COMPUTER GRAPHICS Proceedings, Annual Conference Series, pp.135-142, 1993.

6 H. Kuzuoka, J. Yamashita, K. Yamazaki, and A. Yamazaki, "Agora: A Remote Collaboration System that Enables Mutual Monitoring", Proc. CHI'99, ACM SIGCHI, pp.190-191, 1999.

7 E. Hosoya, M. Kitabata, H. Sato, I. Harada, H. Nojima, F. Morisawa, and S. Mutoh, "A Mirror Interface for Real World Interaction", Interaction 2003, pp. 95-96, 2003. (in Japanese)

8 E. Hosoya, M. Kitabata, H. Sato, I. Harada, H. Nojima, F. Morisawa, S. Mutoh, and A. Onozawa, "A Mirror Metaphor Interaction System: Touching Remote Real Objects in an Augmented Reality Environment", ISMAR2003, pp. 350-351, 2003.

9 E. Hosoya, M. Kitabata, H. Sato, I. Harada, and A. Onozawa, "Interactive Communication using Mirror Metaphor Interface", IEICE General Conference 05, A-16-20, p. 296, 2005. (in Japanese)

10 F. Morisawa, S. Mutoh, J. Terada, Y. Sato, and Y. Kado, "Interaction with remote objects using Sticl-on Communicator", Interaction 2003, pp.223-224, 2003. (in Japanese)

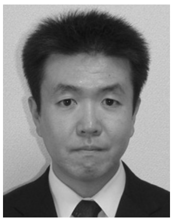
11 Y. Yamamoto, I. Yoda, and K. Sakaue, "Arm-Pointing Gesture Interface Using Surrounded Stereo Cameras System", Proc. of 17th International Conference on ICPR2004, Vol. 4, pp. 965-970, 2004.

12 E. Hosoya, H. Sato, M. Kitabata, I. Harada, H. Nojima, and A. Onozawa, "Arm-Pointer: 3D Pointing Interface for Real-World Interaction", in Proc. of ECCV 2004 Workshop on HCI, LNCS 3058, pp.72-82, 2004.

13 Y. Kanatsugu, S. Nagashima, M. Yamada, and T. Shimuzu, "A Method for specifying cursor position by finger pointing", IEICE Tech. Report ITS, ITS-2001, Vol. 101, No. 625, pp. 55-60,2002. (in Japanese)

14 M. Kitabata, T. Ikenaga, K. Uchimura, and K. Yamashita, "A Consideration for Interactions with Character Agents---An Evaluation of Interface using a Self-Image---", .The 5th SIGNOI, pp. 11-16, 2002. (in Japanese)

15 M. Kitabata, E. Hosoya, H. Sato, I. Harada, and A. Onozawa, "A Bare Hand Tracking Algorithm for Mirror Metaphor Interface", IEICE General Conference 05, D-12-35, p. 185, 2005. (in Japanese)



***Hosoya Eiichi***

*Research Engineer, Microsystem Integration Laboratories, Nippon Telegraph and Telephone Corp.*

*Human Computer Interaction, Image Processing*

***Harada Ikuo**, Dr. Eng.*

*Supervisor, Senior Research Engineer, Microsystem Integration Laboratories, Nippon Telegraph and Telephone Corp.*

*Human Computer Interaction, Computer Graphics, Kansei Information Processing, LSI CAD*

***SATO Hidenori***

*Senior Research Engineer, Microsystem Integration Laboratories, Nippon Telegraph and Telephone Corp.*

*Human Computer Interaction, Image Processing, Computer Vision*

***OKUNAKA Junzo***

*Expert Researcher, Tsukuba JGNII Research Center, Collaborative Research Management Department*

*Information Terminal*

***TANAKA Takahiko***

*NTT Communications Corporation*

*Human Computer Interaction*

***ONOZAWA Akira**, Dr. of Information and Computer Science*

*Supervisor, Senior Research Engineer, Microsystem Integration Laboratories, Nippon Telegraph and Telephone Corp.*

*Human Computer Interaction, Computer Vision, LSI CAD*

***KOGA Tatsuzo**, Ph.D.*

*Expert Researcher, Tsukuba JGNII Research Center, Collaborative Research Management Department*

*Application of HCI Technology and GMPLS Network Operation and Management*