Special Feature

Mirror Interface—An Interactive Communication System Utilizing the Metaphor of a Mirror

Eiichi Hosoya[†], Ikuo Harada, Hidenori Sato, and Akira Onozawa

Abstract

The Mirror Interface is a novel interactive communication system that provides users with a virtual shared space using realtime video and ubiquitous information. The shared space, displayed using the metaphor of a mirror, provides users with an interactive environment where they can communicate with each other seamlessly across the border between real and virtual as well as the border between local and remote, enabling natural non-verbal communication. Icons of real objects and accessible information are also displayed and can be manipulated by "virtual touch". This article introduces the system and describes its implementation.

1. Introduction

Although the development of broadband networks has created various services, including videophones and videoconferencing, these have not spread as widely as expected. Meanwhile, a variety of services are being developed based on ubiquitous computing technologies. This article introduces a new human interface for communication that utilizes both broadband-based visual information and ubiquitous information.

In person-to-person communication, natural non-verbal expressions based on gestures, body language, and facial expressions are as important as direct linguistic expressions using voice and text information. To achieve such kind of communication between physically separated users, it is important to produce an environment in which the communicating users can feel as if they are sharing the same space at the same time. Conventional video conferencing systems have "seams" between the images of users' spaces and workspaces such as their document areas and whiteboards. These seams are clearly obstacles to smooth collaboration between separated users [1].

Many approaches for constructing a seamless shared space over separated users and workspaces have been proposed. ClearBoard [2] lets a user share a whiteboard with others while supporting gaze awareness. VIDEOPLACE [3], [4] and HyperMirror [5] extract the image of users at the local site using chroma-key so they can integrate the image at the remote site, thus enabling local users to share the real space at the remote site with remote users. CAVE [6] generates a shared space by extracting a human body image using 3D cameras and combining it with a workspace image synthesized by computer graphics. Agora [7] can share a real workspace on the desk while enabling users to give instructions naturally with natural body arrangements.

Those systems considered some important features for remote cooperative working such as workspace sharing and real-space sharing by having users "jump" to another site. In addition to those features, though, we think that it is natural and important to introduce the surroundings around the users equally into both spaces. This equality means that the shared space should reflect the ambience of both sides and that facilities at one site should be available to users at the other site. This would make the space sharing bidirectional.

Our system transfers various kinds of information including image and ubiquitous information such as

† NTT Microsystem Integration Laboratories
  Atsugi-shi, 243-0198 Japan
  E-mail: hosoya@aecl.ntt.co.jp

that obtained from sensors bidirectionally and constructs a shared space based on the metaphor of a mirror at each site. The shared space integrates both the local and remote site images and the workspaces such as documents or images of real desktops. Users can communicate bidirectionally and smoothly through this shared space, in which information from the local and remote locations and cyberspace information are integrated seamlessly. As a result, each user can provide instructions and perform various operations at the other site [8], [9] by using natural non-verbal expressions. We refer to the human interface forming the key component of this system as the "Mirror Interface".

In this article, we introduce the method of communication via the Mirror Interface in section 2 and describe its implementation in section 3. We explain some possible applications of the Mirror Interface in section 4.

2. Mirror Interface

2.1 Overview

The Mirror Interface is a real-world-oriented interface that enables users to operate objects existing in both local and remote real spaces as well as information in cyberspace without any distinction. It is constructed as a type of "desktop" by projecting real-world and cyber-world images onto a screen. By displaying the local and remote real-world images and cyber-world images together, the Mirror Interface facilitates smooth dialogue. In short, it acts as a communication tool that seamlessly integrates the real world and the cyber world, as well as remote and local locations.

The first idea for reducing the seams between the images of users' spaces and workspaces is to have real objects in two real spaces available to both sets of users, as illustrated in **Fig. 1**. A large display unit resembling a mirror is placed in front of the users. The camera placed above the center of this display faces the users and captures images of the users and their surroundings, which are converted to mirror images and shown on the display. There is an identical system at the remote side to capture similar images. These two images are transferred over the network, and the image of the remote location is superimposed as a translucent image on the image of the local location. As a result, objects and users at the two locations are integrated into one "room", giving both sets of users the impression of looking at the same mirror. This enables users to point out things with demonstrative words like this, that, here, and there.

The second point to consider is that users should be able to interact with shared objects in a natural way. The display not only shows mirror images of the real world, but also provides information about objects seen in the image. The users can access that information by using a hand as a pointer. Thus, the Mirror Interface enables users to interact with objects existing in real space as well as ones in cyberspace without any distinction.

The third consideration is to enable users to perform the above interactions in real time. Since the images are transferred in DV (digital video) format without compression, the latency is small. The pointing recognition can also be done in real time. These features are described in section 3.

Summarizing the above, the Mirror Interface can be regarded as a generalization of the concept of "desktop" to the real world. While the ordinary graphical user interface (GUI) using an ordinary computer desktop uses a mouse as a pointing device to perform by selecting and clicking an icon, the Mirror Interface
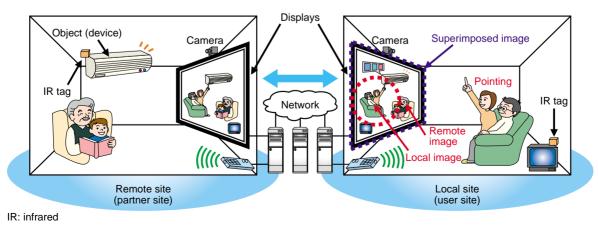

Fig. 1. Overview of the Mirror Interface.

uses the user's hand as a pointer to operate equipment in the real world. In this sense, the Mirror Interface can be regarded as a "real-world-oriented GUI". It consists of the real-world desktop and the pointing interface, as described below.

2.2 Real-world-oriented desktop

A real-world-oriented desktop provides the user not only with information in a computer but also with an interactive space consisting of real-world information. The real-world-oriented desktop displays images of the local site and those of the remote site in the form of the image seen in a mirror. An icon is displayed on each of the operable real objects (devices, etc.) on the desktop. Icons are also used to symbolize the usable cyber objects in computers. By pointing at these icons, the user can not only operate computer functions but also activate services in the real world. Since the user is also present in this image and can operate the objects by simply overlaying his or her hands in the image, he or she can gain a natural feeling of operating these devices by "touching them". An example of a desktop created using the Mirror Interface is shown in **Fig. 2**.

On a real-world-oriented desktop, operable objects and their icons must be arranged in corresponding positions on the screen. To enable the system to identify the positions of devices, they are given IR (infrared) tags [10], [11].

2.3 Pointing function for real-world-oriented GUI

On a conventional desktop, a mouse, touch panel, etc. are used for direct pointing. For things beyond immediate reach, devices like remote control units are used for indirect pointing. To enable natural pointing to real objects, several types of interface have been proposed; they all use a 3D camera to detect the direction in which the user is pointing [12]-[14]. These methods aim to create a function for pointing directly at objects within a room. Reference [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.

The Mirror Interface enables the user to indicate a point in the real world simply by indicating the corresponding point on the real-world-oriented desktop. The Mirror Interface detects the intended "pointing" when the user's hand in the image overlaps with the icons on the real-world-oriented desktop. The performance of this method was evaluated in an experiment [15]. Here, 28 subjects, who were novice IT users, were asked to touch a target (icon) on the screen displayed randomly at one of 18 locations after touching a fixed "Start" icon by using both our pointing method and an ordinary mouse. The time spent on this operation was measured in both cases for four different target sizes. The results show that our pointing method required less time than a mouse in general. Furthermore, the time taken did not depend on the target size for our pointing method while it did for the mouse. These results indicate that our method was more effective for novice IT users. This suggests that a system capable of robust recognition of the user's hand position can be used as a pointing interface.

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 action is performed. For more complex operations, the Mirror Interface can display a popup menu for the icon in a hierarchical manner. After the command has been executed, the menu disappears and the screen returns to normal, waiting for the next command. To improve the usability, we are also investigating a pointing method using a bare hand using computer vision technologies [16].
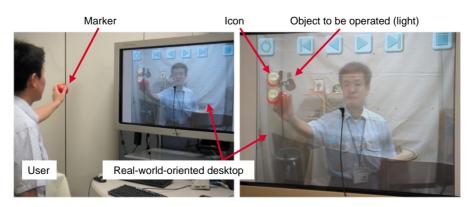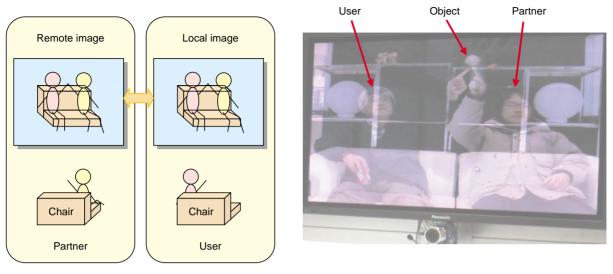

Fig. 2.   Real-world-oriented GUI.

(a) Placement of users and objects

(b) Superimposed translucent images

Fig. 3.   Superimposed translucent images (taken at NTT Group Communication EXPO 2005).

In Fig. 2, an icon on the screen is being touched by the user holding his hand (with a marker) over it. This operation will cause the light at the remote side to be turned on.

2.4   Integration of local and remote images

The reflected image of the user is also used for communication with the partner at the remote site. The image of the partner and the room captured at the remote site is integrated with that of the local site by converting both images to translucent ones and superimposing them. The resulting image can be obtained by changing the pixel value p at each pixel to "$p = \alpha \cdot p_L + (1 - \alpha) p_R$" where p_L and p_R are the pixel values at the local and remote sites, respectively, and α is the degree of translucency. This superimposition process is performed independently at the two locations to generate independent images. An example of the real-world-oriented desktop created by this process is shown in **Fig. 3**. In this example, a user at the local site faces a person at the remote site and objects (such as shelves and a light) at both locations are displayed on a single screen. In other words, the two rooms are combined into one space. The superimposed translucent images show all objects at both locations without hiding overlapped objects; that is, users can recognize everything at both locations. During communication, the degree of translucency can be adjusted independently at each location.

2.5   Collaboration in a shared space

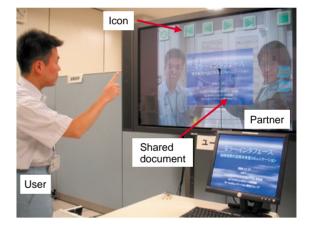To show how the shared space can be used, we built

Fig. 4.   Example of shared document in shared space.

a document sharing function in it. An example in which a translucent document is superimposed over the real-world-oriented desktop is shown in **Fig. 4**. By "touching" an icon, users at either location can turn the page forward or backward.

In Fig. 4, the user and partner at two separate locations are conversing while pointing with their fingers at the map shown in the slideshow. Each person can see the other pointing in the correct orientation on his own screen. Demonstratives such as "this" and "that" are understandable to both parties since both users see the identical mirror image.

Because the Mirror Interface allows two parties to share the same space, a 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. In our experiments, we even observed a case in which the local user accidentally "touched" the other person at the remote side and that person reacted by jerking away from the "hand".

3. System implementation

We have implemented the Mirror Interface to connect two locations. We set up two terminal environments, each with a personal computer (PC), plasma display panel (PDP), camera, microphone, speaker, and other devices (**Fig. 5**). The systems at both locations were constructed on local networks and were connected from separate ports via a network called the external network, which could be another LAN or the Internet. In addition, a PC for serving the slideshow images was connected to the external network. We set IPv4 local addresses for TCP/IP and UDP/IP communications between the PCs (TCP/IP: transmission control protocol, internet protocol, UDP: user datagram protocol).

The system captures images by cameras and voice by microphones, converts them into DV data format, and assembles the data into fixed-length UDP packets for transfer. If the object selected by the user is located at the remote site, the corresponding command must be performed at the remote site. Furthermore, coordinate data for registered objects must be exchanged when necessary to reflect the movement of objects. The data used for this purpose is very small in volume relative to the image data, so the transmission frequency can be about the same as the

frame rate required to ensure the proper display of object movement. To prevent packet loss and to ensure that all commands reach the appropriate destination, TCP/IP was used for this type of data transfer.

For the slideshow, data generated in Microsoft PowerPoint was multicast to two locations through the external network. The slideshow images were captured at preset intervals and converted to DV format for transmission. Therefore, even a paper document can be superimposed on the real-world-oriented desktop if a document camera is used.

We verified the behavior of this system. Although some communication delays ranging from approximately 100 to 150 milliseconds occurred due to the superimposition and codec processing, smooth dialogue was possible with no significant problems.

4. Applications

Here, we list some possible application areas for the Mirror Interface.
(1) Applications using superimposed document images
  • Video conferencing and interactive remote lectures: The ease with which users at remote locations can understand demonstrative words enables effective pointing to shared documents and/or real objects.
(2) Applications using superimposed images of a remote location
  • Remote instruction for exercising or operating a machine: Users can take advantage of a shared space, where one body can be superim-
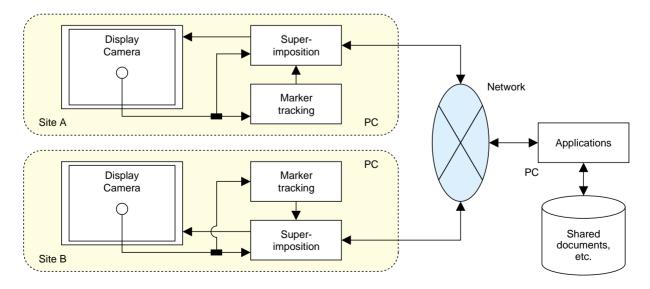

Fig. 5. System implementation with the Mirror Interface.

posed on another.

(3) Applications using the pointing interface
- Futuristic presentations in exhibitions and showrooms.

In some situations, these three areas can be combined for even more effective communications. We plan to evaluate the effectiveness of the Mirror Interface in some of those cases. We also plan to improve the current system considering practical NTT services while developing fundamental technologies such as those needed for tracking a bare hand.

References

[1] H. Ishi, "CSCW and Groupware," Ohmsha, 1994 (in Japanese).

[2] H. Ishi and M. Kobayashi, "ClearBoard: A Seamless Media for Shared Drawing and Conversation with Eye-Contact," Proc. CHI'92, ACM SIGCHI, Monterey, 3-7 May 1992.

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

[4] M. W. Krueger, "Artificial Reality II," Addison-Wesley Publishing Co. Inc., pp. 34-45, 1990.

[5] O. Morikawa, J. Yamashita, Y. Fukui, and S. Sato, "Soft Initiation in HyperMirror-III," INTERACT'01, pp. 415-422, 2001.

[6] 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.

[7] H. Kuzuoka, J. Yamashita, K. Yamazaki, and A. Yamazaki, "Agora: A Remote Collaboration System that Enables Mutual Monitoring," Proc. CHI'99, pp. 190-191, ACM SIGCHI, 15-20 May 1999.

[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 2005, A-16-20, pp. 296, 2005 (in Japanese).

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

[11] S. Mutoh, F. Morisawa, J. Terada, Y. Satao, and Y. Kado, "Stick-on Communicator and Its Broadband Applications—A Compact Wireless Device with a Built-in Sensor and Actuator Connecting the Real World and the Network," NTT Technical Review, Vol. 2, No. 2, pp. 53-58, 2004.

[12] 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.

[13] 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.

[14] Y. Kanatsugu, S. Nagashima, M. Yamada, and T. Shimizu, "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).

[15] M. Kitabata, H. Nojima, E. Hosoya, and A. Onozawa, "Evaluation of object selection methods on a mirror interface system," FIT2003, pp. 589-590, 2003 (in Japanese).

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

Eiichi Hosoya
Research Engineer, Home Communication Research Group, Ubiquitous Interface Laboratory, NTT Microsystem Integration Laboratories.
He received the B.E. and M.E. degrees in electrical and electronics engineering from Chiba University, Chiba, in 1991 and 1993, respectively. He joined NTT LSI Laboratories, Kanagawa, in 1993. He has been involved in R&D of LSI design technologies, computer vision and human computer interaction. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) of Japan and the Information Processing Society of Japan (IPSJ).

Ikuo Harada
Senior Research Engineer, Supervisor, Home Communication Research Group, Ubiquitous Interface Laboratory, NTT Microsystem Integration Laboratories.
He received the B.S., M.S., and Ph.D. degrees in electronic engineering from Osaka University, Osaka, in 1981, 1983, and 1998, respectively. In 1983, he joined the Atsugi Electrical Communication Laboratories, Nippon Telegraph and Telephone Public Corporation (now NTT), Kanagawa. There he engaged in the research of VLSI layout algorithms. In 2000, he moved to ATR Media Integration & Communications Research Laboratories, Kyoto, and researched Kansei information processing. He returned to NTT in 2002. His current research interests include human computer interactions and CG content creation. He is a member of IEEE, IEICE, and IPSJ.

Hidenori Sato
Senior Research Engineer, Service Deployment Strategy Project for Business Promotion, NTT Service Integration Laboratories.
He received the B.S. and M.S. degrees in physics from Tohoku University, Miyagi, in 1987 and 1989, respectively. In 1989, he joined NTT LSI Laboratories, Kanagawa, where he worked on LSI CAD. His current research interests are human computer interaction systems using computer vision. He is a member of IEEE, IEICE, IPSJ, and the Institute of Image Information and Television Engineers.

Akira Onozawa
Senior Research Engineer, Supervisor, Home Communication Research Group, Ubiquitous Interface Laboratory, NTT Microsystem Integration Laboratories.
He received the B.E. and M.E. degrees in electronic communication engineering and the Ph.D. degree in information and computer science all from Waseda University, Tokyo, in 1983, 1985, and 2002, respectively. He joined NTT Atsugi Electrical Communication Laboratories in 1985 and since then he has been working for R&D of LSI CAD algorithms and systems. His current interests include human computer interaction and computer vision. He is a member of the Association for Computing Machinery, IEEE, IEICE, and IPSJ.