

Multiparty Videoconferencing at Virtual Social Distance: MAJIC Design

Ken-ichi Okada, Fumihiko Maeda, Yusuke Ichikawaa, and Yutaka Matsushita

Department of Instrumentation Engineering
Keio University

3-14-1 Hiyoshi, Kohoku-ku, Yokohama, 223 Japan

Tel: +81-45-563-1141, Fax: +81-45-562-7625

E-mail: okada@inst.keio.ac.jp

ABSTRACT

This paper describes the design and implementation of MAJIC, a multi-party videoconferencing system that projects life-size video images of participants onto a large curved screen as if users in various locations are attending a meeting together and sitting around a table. MAJIC also supports multiple eye contact among the participants and awareness of the direction of the participants' gaze. Hence, users can carry on a discussion in a manner comparable to face-to-face meetings. We made video-tape recordings of about twenty visitors who used the prototype of MAJIC at the Nikkei Collaboration Fair in Tokyo. Our initial observations based on this experiment are also reported in this paper.

KEYWORDS: MAJIC, multi-party videoconferencing, multiple eye contact, gaze awareness, groupware, networked realities, tele-presence

INTRODUCTION

When we have discussions in face-to-face situations and people approve of a statement, we can tell by their attitude, tone, eye movements, gestures, and so forth, whether or not they approve wholeheartedly. It is difficult, on the other hand, to estimate how strongly they approve when we read only the minutes without attending a meeting. Hence, one of the purposes and/or advantages of face-to-face meetings is that all of the participant are aware of the speaker's intent and the other listeners' reactions based on both verbal and nonverbal communication. This is also very important in videoconferencing systems, but existing multi-party videoconferencing systems have the following shortcomings:

- (1) It is difficult to support multiple eye contact among participants.
- (2) Since the images of the participants are often small, it is difficult for users to read participants' facial

expressions and gestures, and it is impossible to achieve a sense of reality.

- (3) When multi-displays are used to support multi-party videoconferencing, there are boundaries between portraits, and consequently, participants can not achieve a sense of reality or the feeling of attending a meeting together.

In order to communicate with a sense of realism in videoconferencing, we have developed "Multi-Attendant Joint Interface for Collaboration" (MAJIC), which works in a multi-site and multi-user environment. MAJIC projects life-size video images of participants onto a large curved screen without boundaries between them, and consequently users feel as though they are attending a meeting together and sitting around a table with the other participants. It is estimated that 60 percent of conversation involves gaze, and 30 percent involves mutual gaze [16]. Therefore eye contact and gaze awareness should be supported in video-conferencing systems. With MAJIC, users can make eye contact with an individual participant and be aware of the direction of each participant's gaze; that is, they can recognize who is looking at whom. MAJIC achieves eye contact without the half transparent mirror which is used in most existing systems to support eye contact, and as a result, it can provide a work space at the center of participants.

Face-to-Face Meetings

A meeting is traditionally a process in which people gather at the same time and the same place and exchange views or information to solve problems. "Face-to-face," "distance" and "shared space" are very important factors for making a meeting creative. The best way may be for all of the participants to sit or stand around a round table and discuss topics in a meeting room in which they can see each other and in which the distance between them is not very great. The reasons for this are that we always communicate nonverbally as well as verbally. For example, when we are attempting to persuade someone, we look directly into their eyes and speak enthusiastically with expansive gestures. When we lack confidence, on the other hand, we cast our eyes down. Gaze is also used to control a meeting.

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association of Computing Machinery. To copy otherwise, or to republish, requires a fee and/or specific permission.

CSCW 94- 10/94 Chapel Hill, NC, USA

© 1994 ACM 0-89791-689-1/94/0010..\$3.50

A chairperson sometimes gazes at participants to urge them to speak when there is silence in a meeting.

Nonverbal communication is often unconscious. If we are interested in a certain person's opinion, we usually face him or her symmetrically. On the other hand, if we oppose or are unhappy with that person, we turn our face away or rest one elbow on the table; that is, we often posture asymmetrically [15]. In this way we sense the atmosphere in the meeting room and the aura of the participants, and, consequently, we can understand the opinions of the participants clearly and make the meeting productive [14].

Videoconferencing

Face-to-face meetings are the best way to come to decisions, but it is sometimes difficult to assemble participants at the same time and same place. There have been many studies on tele-communication support systems using video images [3, 4, 7], and we think videoconferencing systems could be developed in two categories. One attaches importance to portability, in order to be able to communicate with anyone, anytime and anyplace. An example of this is desktop conferencing systems [1, 13, 19]. We will be able to have meetings using multi-media notebook computers equipped with radio network facilities in the near future.

The other category attaches importance to reality, and, as a result, usually requires a large space. Some topics or keywords in studies of these systems are

- (1) informal communication support [2, 5, 11],
- (2) simulation of a multi-way round-table meeting [16, 20],
- (3) eye contact support [8, 12],
- (4) integration of communication space and work space [9, 18], and
- (5) virtual reality [17].

In order to communicate with each other by videoconferencing with a sense of reality, one of the most important problems is how to support eye contact among participants and provide life-size portraits of them. Many systems use a half transparent mirror to support eye contact, while other systems use a large screen to provide life-size portraits, but it is difficult to provide both. Moreover, in the case of multi-party videoconferencing, multiple eye contact should be supported, and to support a sense of reality or feeling of togetherness, there should be no boundaries between the pictures.

DESIGN OF MAJIC

We identified the following four design requirements to implement "Multi-Attendant Joint Interface for Collaboration" (MAJIC):

- (1) multi-way (more than two-way) round-table meetings and multiple eye contact between participants must be supported,

- (2) users can become aware of the gaze of one participant toward another participant; in other words, users maintain awareness of who is visually attending to whom,
- (3) life-size video pictures of the participants are shown without boundaries to achieve a sense of reality, and
- (4) a shared work space must be provided at the center of the participants.

Figure 1 illustrates the first draft of MAJIC. It shows four persons working in Tokyo, New York, London and Paris offices respectively, and having a meeting one day. Hence, they are going to use MAJIC at each of their offices. There is a shared work space and a large curved screen in front of each user, and the other three participants are shown on the screen as though attending the meeting together and sitting around the work space. Each user can select a background of portraits in order to relax them or inspire them.

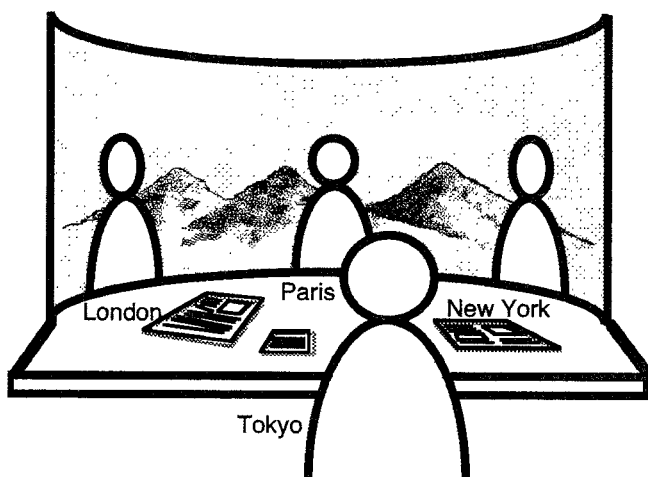


Figure 1: The first draft of MAJIC.

Life-size Portraits

There are two important factors in achieving a sense of reality during videoconferencing. One is the size of participants' images on the screen. For example, we may feel participants are far away if their portraits are smaller than life-size. Moreover, it would be difficult to read facial expressions or gestures if participants are shown on a small display or displays. Another factor is the background behind the portraits. Although there are some systems that use multi-displays to provide for multi-party conferencing, to achieve a feeling of togetherness, there should be no boundaries between the portraits. If users are surrounded by other participants with a seamless background, they can feel as though they are together.

Multiple Eye Contact

Multiple eye contact among participants should be supported to make multi-party conferencing effective, since without eye contact or calling out the person's name,

it is difficult to speak to a specific participant in a multi-party environment. We usually become aware of one participant gazing toward another through eye movements and head turns in face-to-face meetings. With the Hydra system, it has been reported that movements in the periphery on the screen lose the ability to attract recipients' attention because of the problem of small screens [16]. Life-size portraits arranged around a table produce a situation in which the eye movements and head turns of the users attract the attention of the recipients.

Workspaces

In face-to-face meetings, participants usually sit around a table on which material is piled up in front of them; some material is for individual participants and other material is shared. Of course eye contact is very important in communicating with one another, as mentioned above, but especially in Japan it is impolite to look into someone's eyes for a long time. What should we look at? Sometimes we look at other participants one after another, or we look at the material on the table. A space in front of participants provides each person with a personal workspace and a shared workspace on which to place material; at the same time, it gives participants a reason for averting their eyes.

Distance

Distances between people vary according to their relationships and are classified as follows [6]:

- (1) Intimate distance: 0 - 1.5 feet
This is the distance for lovers, couples and close friends. Coming within intimate distance of others makes them feel uncomfortable or that the person is aggressive.
- (2) Personal distance: 1.5 - 4 feet
We spontaneously assume at least a personal distance from a stranger in our daily lives.
- (3) Social distance: 4 - 12 feet
Business activities such as face-to-face meetings are usually carried on at social distances.
- (4) Public distance: more than 12 feet
When we assume a public distance from an object, we do not have personal involvement with it. This is the distance for speeches at halls or plays in theaters.

Virtual Distance

How do participants feel when they are looking at video images of others on a display(s) during video teleconferencing? Perhaps there is a sensed distance among participants, which is not equal to the spatial distance. The sensed distance is influenced by spatial distance from the display, the size and quality of the video images, backdrops, voice fidelity and so on. We call the sensed distance between participants in networked environments "virtual distance." The best distance among each participant in business meetings using the video teleconferencing system is "virtual social distance," which allows participants to feel as though they are talking at social distances.

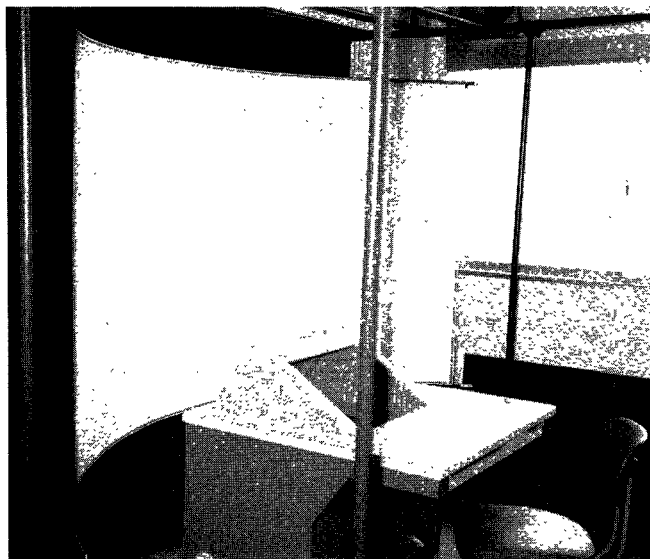


Figure 2: A prototype of MAJIC (front view).

IMPLEMENTATION OF MAJIC

Figure 2 shows a prototype of MAJIC; there is a desk equipped with a workstation and a curved screen, 4 feet long by 8 feet wide, forming an arc with a radius of 4 feet. Video images of other participants are projected on the screen by video projectors. When users sit at the center of the arc, the distance between them and other participants is around 4 feet. We have concluded that around 4 feet may be the best distance for face-to-face meetings with 3 or 4 colleagues, since the distance from people who work together tends to be shorter at a social distance, or it is sometimes at a personal distance. Figure 3 is a picture of the back of the screen in which you can see the desk through the screen.

How to Support Eye Contact

The MAJIC screen is a thin transparent film with a large number of very small hexagons printed on both faces. As shown in Figure 4, white hexagons are printed on the front side (Face A) and black hexagons are printed on the back side (Face B). Since the white hexagons reflect light, and it is dark behind the screen, Face A can be used as a normal screen (see Figure 2). The black hexagons, on the other hand, do not reflect light at all, and it is relatively light in front of the screen; consequently, one can see the other side through the screen from Face B (see Figure 3). The transmissibility of the screen depends on the size and the number of the hexagons. We have used a 40% transmissibility screen for the MAJIC prototype.

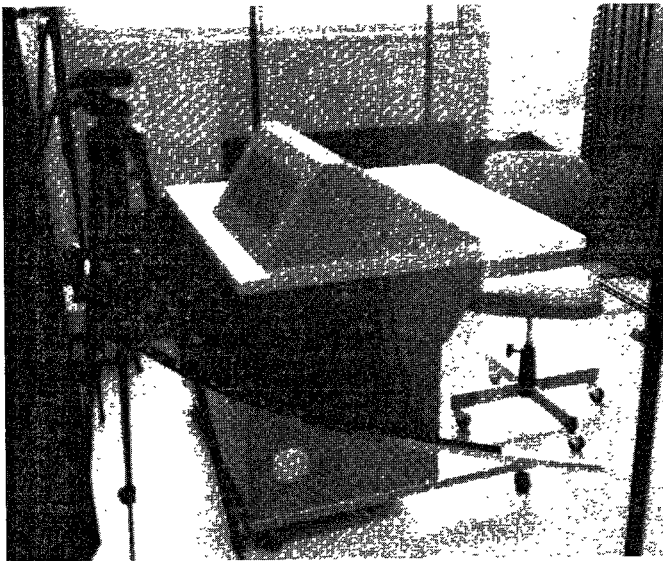


Figure 3: A prototype of MAJIC (back view).

Figure 5 illustrates how MAJIC supports eye contact. A video projector located above User A projects a life-sized video image of User B on the screen, and a video camera located behind the screen takes a picture of User A, and so on.

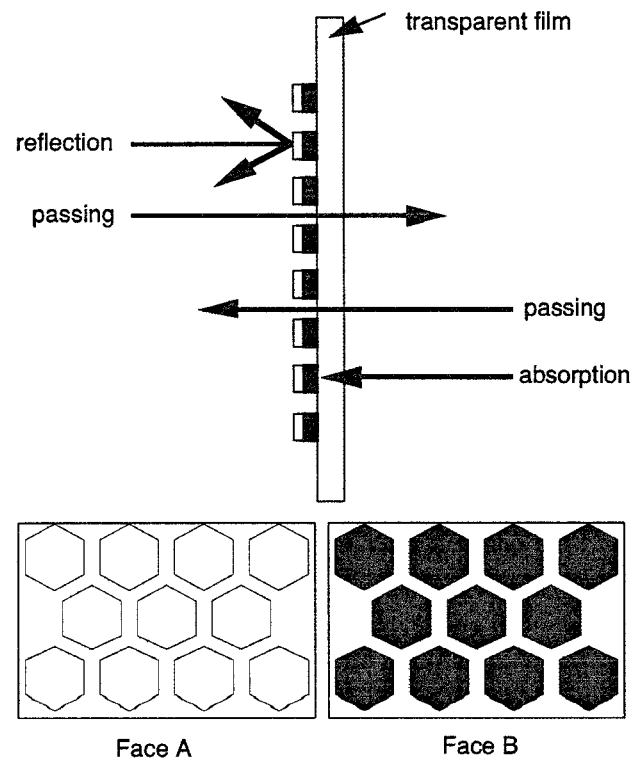


Figure 4: MAJIC screen.

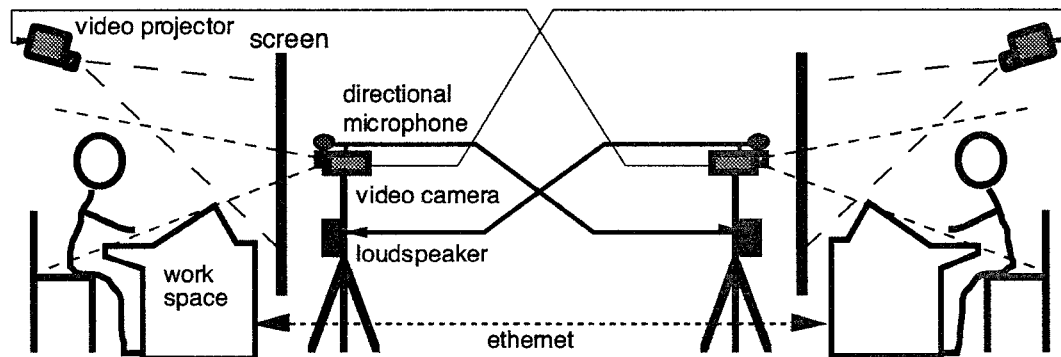


Figure 5: System architecture of MAJIC.

In order to take clear pictures of users, the video projectors are set up so that they do not enter the field of vision of the video cameras. Since each video camera is set up at the center of a partner's face on the video image, Users A and B can make eye contact. Moreover, since video images of participants are life-size, users can read their partner's facial expressions and observe the direction of their partner's gaze.

Figure 6 shows a MAJIC eye contact experiment in which the video projector projects a video image taken by the video camera behind the screen.



Figure 6: A verification of eye contact.

How to Support Multiple Eye Contact

Figure 7 illustrates three-way videoconferencing using MAJIC. Each MAJIC system is equipped with a curved screen, a workstation, two video projectors, two video

cameras, two directional microphones, and two loudspeakers. Sitting at the center of MAJIC, the participants form an isosceles triangle with each other in a virtual meeting space. When User A turns his head to the right to look straight at User B, User B sees User A full-face and User C sees the left profile of User A. In other words, User C becomes aware of User A gazing toward User B, and User B becomes aware of the head-turning of User A toward her and can make eye contact with User A if she likes. When User A and User C both look at User B at the same time, User B makes multiple eye contact among them, as shown in Figure 8. When User A and User C are looking at each other and talking about something, User B becomes aware of this situation, as shown in Figure 9. These two photos were taken from the face position of User B. Thus the MAJIC system supports multiple eye contact and gaze awareness.



Figure 8: Multiple eye contact.

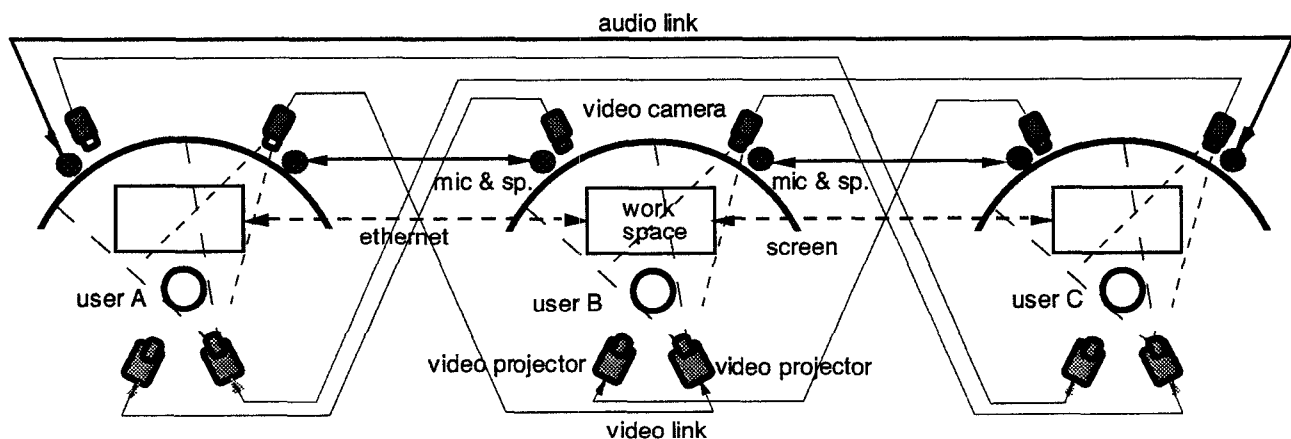


Figure 7: Three-way videoconferencing using MAJIC.

Voice also plays an important role in MAJIC, both in making communication seem real and in making the meeting proceed smoothly. We set a directional microphone and a loudspeaker beside the video camera, as shown in Figure 5. Since each set is positioned immediately behind the video images of the participants, it becomes easy for users to know who is speaking. For example, in Figure 7, User B listens to the voice of User A from the left side and the voice of User C from the right side. When User A sits upright and speaks to both Users B and C, User B hears User A's voice as loudly and clearly as User C does. On the other hand, when User A turns his head toward User B to speak to her, User B hears User A's voice slightly more loudly and clearly.

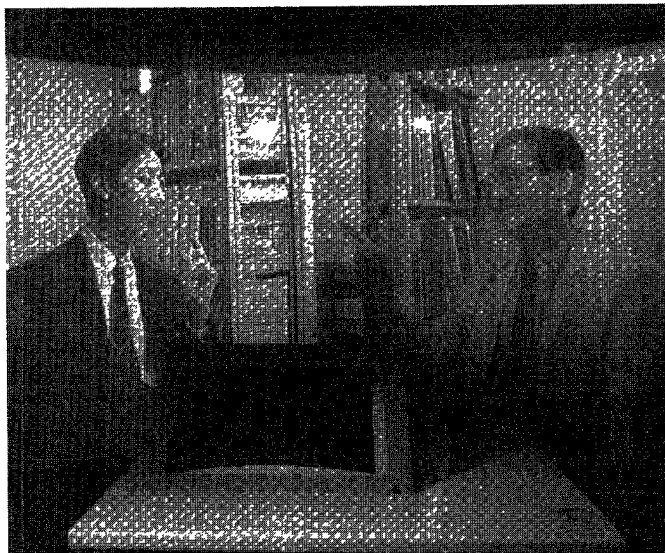


Figure 9: Gaze Awareness.

How to Support the Feeling of Togetherness

What should the system support to achieve a feeling of togetherness during videoconferencing? We have attached great importance to the size of video images and the continuity of pictures. Life-size video images of each participant projected on a curved screen are face-to-face with the user, and edges of pictures are overlapped.

Figure 10 shows multi-party videoconferencing using MAJIC. Since the workstation in the desk is tilted, the users can see both their own material in the display and their participants on the screen at the same time [10]. A few systems have achieved integration of the communication space and workspace, but we separate them in MAJIC. We observed face-to-face meetings in the real world in designing MAJIC and carefully noted participants' behavior in terms of when and how they made eye contact and averted their eyes. We were particularly interested in how they averted their eyes in round-table

meetings. The behavior observed most frequently was turning their eyes downward to the material on the table. Based on these initial observations of eye movements, we positioned a workstation in front of the users to give them a chance to avert their eyes by turning their eyes down to it.

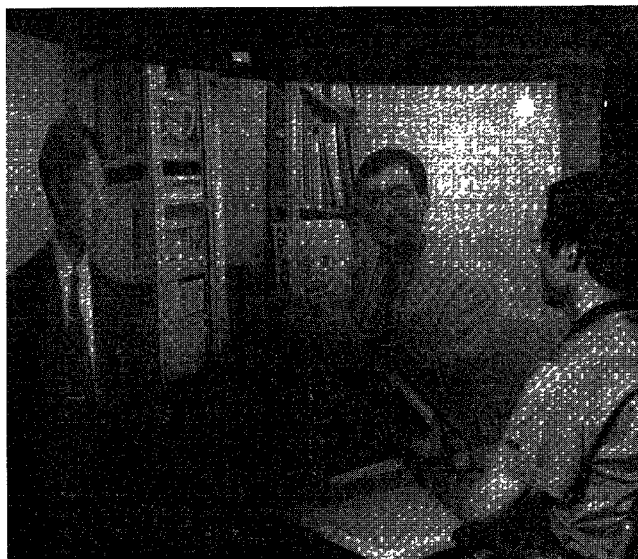


Figure 10: MAJIC in use.

EXPERIMENT AND USER FEEDBACK

We demonstrated the prototype of MAJIC at a collaboration fair held on December 14 and 15, 1993, at the Nikkei Hall in Tokyo. The fair was sponsored by Nihon Keizai Shinbun (Japan Economy Newspapers Inc.). More than 100 visitors came to our booth to see the demonstration of MAJIC, and about 20 visitors participated in three-way videoconferencing experiments with our demonstrators using MAJIC. They played a multi-user poker game developed for the collaboration fair. We were interested in how people play poker in a distributed environment, to see, for example, whether they can see through a pokerface or not.

Since we did not have enough space to set up more than two MAJICs at the collaboration fair hall (the prototype of MAJIC is eight feet by eight feet), we prepared a MAJIC for visitors as well as two simple systems constructed from a half transparent mirror, displays and video cameras for the demonstrators. These simple systems supported eye contact. The demonstrators were students from our laboratory; one was Japanese and another Korean. We recorded visitors on video tapes using the MAJIC video cameras and the microphones placed behind the screen; that is, the video and audio signals of the MAJIC network were recorded. We are analyzing the video tapes now.

Eye Contact

As soon as visitors sat in the MAJIC chair, they usually turned their heads to look both ways two or three times and made eye contact with the demonstrators. When one of the demonstrators explained MAJIC, they looked straight at him and sometimes nodded their heads. After they asked both of the demonstrators to look at each other or to look at them to confirm multiple eye contact, they turned their eyes downward at the workstation display. They occasionally looked at the demonstrators, but they made eye contact for less than one second at a time. In particular, while playing the poker game, they seldom made eye contact with the demonstrators. We asked them if they could make multiple eye contact and become aware of the gazes of each of the demonstrators, and all of their answers were "yes."

Portrait Size

When we played back the video tapes, we were surprised by scenes in which most visitors greeted the demonstrators on the screen, since this is unusual at computer fairs in Japan, at least in videoconferencing demonstrations. Visitors usually said "Excuse me?" or "Can you hear me?" to the demonstrators, but they seldom said "Good afternoon!" or bowed their heads in greeting to the demonstrators. This may be an effect of the life-size video images and the sensed distances among them. Some visitors felt pressure when we zoomed in on a demonstrator's face. On the other hand, the students in our laboratory, who are naturally friends of the demonstrators, preferred video images larger than life-size. Perhaps they like to talk at personal distances.

Image Quality

The quality of the video images was not very good because of the MAJIC screen at the visitors' site and the half transparent mirrors of the simplified system at the demonstrators' site. The transmissibility of a half mirror is less than 50% and the reflectivity of the MAJIC screen is less than 60%. Consequently, the video image is not very bright. Moreover, since the MAJIC screen is constructed of a large number of small hexagons, the resolution of the video images is not very high. No visitors complained about the quality of image, however, and the reasons they did not complain are as follows:

- (1) there are only life-size video images on the screen,
- (2) they may have supplemented the images on the screen with their imagination, and
- (3) other information is displayed clearly on the workstation in front of them.

A few visitors pointed out that the reflections on the screen were slightly glaring. The cause of the reflection is that MAJIC uses transparent film as the screen.

Impressions

A large number of visitors were interested in MAJIC, and they enjoyed using it. Most of them told us that they were satisfied with virtual meetings and that MAJIC was more effective for meetings than desktop conferencing systems.

A couple of them, however, pointed out that the prototype of MAJIC was too big to install in a Japanese office. We asked the visitors if they had a feeling of togetherness and why they felt so. Many of them gave positive answers, and the chief replies were that they were impressed by the life-size portraits and the capability for multiple eye contact.

DISCUSSION

We will take the opinions of the visitors into consideration, and we are planning to refine the prototype of MAJIC for practical use.

The prototype of MAJIC has the following shortcomings:

- (1) Only one user at each site can make eye contact among participants.
- (2) A user should not move around. Limitations (1) and (2) are caused by the video cameras set behind the screen. They are mounted on a tripod and fixed to sight on their target, that is, a user sitting at the center of the arc of the screen. In order to overcome these limitations, we are planning to introduce a video camera that pursues a person.
- (3) The number of sites is limited, four-way video conferencing being the maximum possible with the prototype. The number of the sites depends on the size of the screen and video images. When the number of participants increases, the distance between them becomes greater at real meetings. MAJIC could overcome this problem by projecting smaller video images, but it would become difficult to support gaze awareness.
- (4) The quality of images is not very good.
- (5) The size of the prototype is very large. It is difficult to solve problems (4) and (5). We have changed a print pattern of the screen and are conducting projecting experiments. Although the resolution is limited since MAJIC uses this kind of screen, the impressions of visitors at the collaboration fair were affirmative. Projecting life-size video images is the cause of the large size, but since MAJIC uses a thin film as its screen, it will be possible to make a folding MAJIC.
- (6) There is no seam between pictures, but there is a seam between backgrounds. It is very easy to support seamless backgrounds by chromakey (although we have not introduced it yet because of a problem of cost). Figure 11 shows an example of seamless backgrounds. We took this photo by adjusting the angle of the video cameras. It will be interesting to observe virtual meetings in various backgrounds. For example, what is the effect of the background, or what kind of background relaxes users or inspires them?



Figure 11: Seamless background.

CONCLUSION

We have described the design and implementation of MAJIC, a multi-party videoconferencing system which supports multiple eye contact, life-size portraits, and awareness of the direction of participants' gaze. It can provide an atmosphere in which users at different locations seem to be attending a meeting together and sitting around a table at social distances. We expect our approach will provide the next generation of teleconferencing systems.

We would like to point out the role of eye movements, eye contact, and gaze at meetings, especially at the meetings in which a few members or colleagues are having heated discussions about the creation of something or the solution to a difficult problem. Eyes can convey the emotions of the speakers, listeners' interest, or the relationships among the members and can also regulate the flow of conversation. We hardly ever say "no" directly in official meetings in Japan, since we usually perform "nemawashi," or have negotiated before the meeting. However, even when no one says "no" clearly, sometimes decisions are not reached, and we must be able to perceive this. This is Japanese culture. As the proverb says "The eyes are more eloquent than the lips."

ACKNOWLEDGMENTS

We thank Giseok Jeong, Masato Kudoh, and Noriyuki Abe in our laboratory for their help in implementing MAJIC and demonstrating it. We also thank Toppan Printing Co., Ltd. and NTT software laboratories for their support.

REFERENCES

1. Ahuja, S. R. The Rapport multimedia conferencing system. In *Proceedings of Office Information Systems*, ACM, N.Y., 1988, pp. 1-8.
2. Bly, S. A., Harrison, S. R. and Irwin, S. MediaSpaces: Bringing people together in a video, audio, and computing environment. *Communications of the ACM*. 36, 1 (Jan. 1993), 29-47.
3. Brittan, D. Being There: The promise of multimedia communication. *Technology Review* (May/June 1992), 42-50.
4. Buxton, W. Telepresence: Integrating shared task and person spaces. In *Proceedings of Graphic Interface '92*, Morgan Kaufman Publishing, Los Altos, 1992, pp. 123-129.
5. Fish, R. S., Kraut, R. E. and Chalfonte B. L. The VideoWindow system in informal communications. In *Proceedings of CSCW'90*, ACM, N.Y., 1990, pp. 1-11.
6. Hall, E. T. *The Hidden Dimension*. Doubleday & Company Inc., N.Y., 1966.
7. Hollan, J. and Stornetta, S. Beyond being there. In *Proceedings of CHI'92*, ACM, N.Y., 1992, pp. 119-125.
8. Ishii, H., Kobayashi, M. ClearBoard: A seamless medium for shared drawing and conversation with eye contact. In *Proceedings of CHI'92*, ACM, N.Y., 1992, pp. 525-532.
9. Ishii, H., Kobayashi, M. and Grudin, J. Integration of inter-personal space and shared workspace: Clearboard design and experiments. In *Proceedings of CSCW'92*, ACM, N.Y., 1992, pp. 33-42.
10. Kuwana, E. and Sakamoto, Y. Designing computer supported meeting environments. *IPSJ SIG Notes*, 93-GW-2, Tokyo, 1993, 33-40.
11. Matsuura, N., Fujino, G., Okada, K. and Matsushita Y. An approach to encounters and interaction in a virtual environment. In *Proceedings of ACM Computer Science Conference*, ACM, N.Y., 1993, pp. 298-303.
12. Nakazawa, K. Proposal of a new eye contact method for teleconferences. *IEICE Transaction on Communication*, J76, 1, (January 1993), 618-625.
13. Nishimura, T., Arikawa, T., Masaki, S. and Yamaguchi, H. Communication environment over conference using a personal multimedia-multipoint

- teleconference system and its human machine interface. *IPSJ SIG Notes*, 92-GW-1, Tokyo, 1992, 75-82.
14. Okada, K. and Matsushita, Y. Hierarchy collaboration model for groupware design. *IPSJ SIG Notes*, 93-GW-4, Tokyo, 1993, 87-94.
 15. Sakamoto, N. *Stage Craft of Meetings*. Diamond Executive, Tokyo, 1992, 76-81.
 16. Sellen, A. J. Speech patterns in video mediated conversations. In *Proceedings of CHI'92*, ACM, N.Y., 1992, pp. 49-59.
 17. Takemura, H. and Kishino, F. Cooperative work environment using virtual workspace. In *Proceedings of CSCW'92*, ACM, N.Y., 1992, pp. 226-232.
 18. Tang, J. C. and Minneman, S. L. Video Whiteboard: Video shadows to support remote collaboration. In *Proceedings of CHI'91*, ACM, N.Y., 1991, pp. 315-322.
 19. Watabe, K., Sakata, S., Maeno, K., Fukuoka H. and Ohmori, T. Distributed multiparty desktop conferencing system: MERMAID. In *Proceedings of CSCW'90*, ACM, New York, 1990, pp. 27-38.
 20. Wellner, P. Interacting with paper on the z, *Communications of the ACM* 36, 7 (July 1993), 86-97.