

Conveying Gaze-Awareness by Using a Faint Light with a Video-Conferencing System

Mikifumi Shikida

Abstract—A method for conveying gaze-awareness by using a faint light is proposed. It is useful for conference-room meetings in which only a few participants are using a commercial video-conferencing system at a remote site. The proposed method involves the following steps. First, the gaze direction of the remote participant is detected; second, and a faint light is shone on the corresponding participant at the local site. The proposed method was experimentally evaluated in the case of two types of meetings under the three conditions: face-to-face, video conference, and video conference with gaze-awareness. When the video-conferencing system supports gaze-awareness by utilizing a faint light, the local participants gaze at the remote participant in the same way as they gaze at other local participants. The method supports natural communication via a commercial video-conferencing system at low implementation cost without the need for redesigning the shared display.

Index Terms—Eye contact, gaze-awareness, shared display, video conference.

I. INTRODUCTION

Commercial video-conferencing systems have become popular due to the spread of high-speed networks and high-quality video-transmission technology. Although these systems can transmit pictures and voices clearly, they cannot sufficiently transmit non-verbal information (cues). Non-verbal information is important to support communication [1]-[3].

Eye information (e.g., information about eye-contact frequency, eye-contact duration, or gaze direction) is one of the most important pieces of non-verbal information. Vertegaal *et al.* suggest, “when users cannot observe eye contact, they cannot accurately estimate whether they are being addressed or expected to speak, posing problems in floor management” [4]. In the case of current video-conferencing systems, cameras are mounted on top of a display or set beside it. Users look at the interlocutor in their display, not their camera. Therefore, direct eye-contact (including mutual gaze) is impossible through a traditional video-conferencing system. If users cannot make eye-contact, they do not know exactly whom they are talking to. Mutual gaze affects communication processes such as turn-taking [5]. And users obtain clues from eye-contact, and those clues make them aware of the intentions of other people or their interest in topics. The current speaker uses eye-contact to decide whether the current topic is suitable for the interlocutors. Thus, when people discuss something without

using eye-contact, communication does not flow smoothly.

From the “shared-display” viewpoint, video-conferencing systems can be classified into two types: “own-display” type or “shared-display” type. As for the former, the display is not shared by users. Each user has their own desktop PC or small display, and the users participate as a “multiparty.” As for the latter, in almost all cases, users participate as two or a few parties from separate locations such as a main office or a branch office. And the users communicate by sharing one big display for each party. This type has become popular for business communication. Several shared-display-type products have been developed and commercialized [6]. However, facilitating eye-contact is an important issue in regard to both types.

Some own-display-type systems convey gaze direction or support gaze-awareness (e.g., [7], [8], etc.). These systems can detect gaze direction and represent each user’s gaze direction because each user has their own exclusive display with a camera and can participate independently.

On the other hand, in regard to the shared-display-type systems, some issues about eye-contact or gaze-awareness remain to be addressed. The shared-display-type system is often used for meetings between a few parties (such as workers in a main office and a branch office). The number of participants in one party tends to differ from that in the other party; that is, the participants at the main office usually outnumber those at the branch office. When almost all users participate at the main office, the number of remote users at the branch office tends to be much lower. This situation affects the feeling of presence.

Eye-contact is important to enhance the feeling of communicative presence [9], which Böcker *et al.* define as “the capacity of a system to transfer mutual communicative signals of interlocutors” [10]. Shared-display-type systems cannot yet sufficiently facilitate eye-contact. Consequently, communicative presence between users at the main office and a small number of remote users is insufficient. Therefore, the majority of users at the main office tend to only talk actively among themselves in their “local world.” They obtain non-verbal information and presence information naturally through face-to-face communication in the real world because they are together in one room. Thus, it is hard for the few remote users to get a chance to take a turn at talking.

This study focuses on gaze-awareness in the case of a commercial videoconferencing system using two shared displays (each shared by one of two parties) when the number of remote users is one. A method for representing the gaze direction of the remote user is proposed. In detail, it uses a faint light to make the users at the other party aware of the gaze direction of the remote user. The aim of this proposal is

to provide gaze-awareness naturally (i.e., not intrusively) and avoid disturbing the talking and communication flow of the current speaker.

II. RELATED WORKS

Much research has studied conveying eye information captured by re-designed devices, such as displays/screens and cameras (e.g., [11]-[16], etc.). Major techniques of re-designing include changing the direction and angle of each user's display according to their head movement or eyes direction, using one-sided or two-sided transparent displays, using special screens, using life-sized or wall-sized displays, increasing the number of cameras, or a combination of these techniques. These techniques make it possible to convey eye information. Bondareva *et al.* implemented a videoconferencing system for supporting direct eye-contact and evoking a high level of social presence by combining of a large back-projection screen and half-silvered mirroring techniques. Their system provides life-sized images of remote users on the screen. However, the number of users is limited [17].

Using an avatar or a tele-presence robot is one method for conveying eye information. Inoue developed a "mixed-reality meeting system" that displays avatars of users wearing a head-mounted display [18]. As for this system, the physical position and body direction of remote participants can be reflected by the avatars. The participant's eyes directions are analyzed. However, this system can be only applied as an own-display-type because each user needs to use their display to participate in virtual space. In the case of the shared-display-type, each party is in the real world. Each user not only talks with remote users but also talks with their neighbors in face-to-face situations in each party. If users use head-mounted displays, they cannot see their neighbor's eyes and facial expressions. Consequently, non-verbal information is insufficient despite users participating in the same room.

Adalgeirsson *et al.* developed a tele-presence robot, called MeBot [19], with a 3-DOF (degrees of freedom) neck and two 3-DOF arms. MeBot allows a remote operator to express some of their non-verbal behaviors, such as hand gestures, head gestures, and eye-contact. Eye-contact can be established by a camera-embedded display. In the case of this kind of tele-presence robot system, local users can easily pay attention to the tele-presence robot while it is roaming. However, the remote operator has to make full use of the robot, and they might be not able to concentrate on only communicating. Moreover, the roaming action of the robot leads to tracking of attention unnaturally.

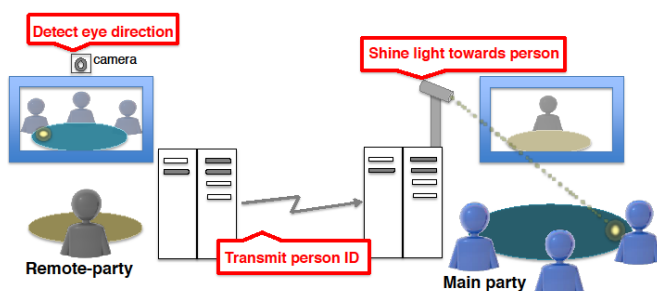


Fig. 1. Overview of proposed method.

Novel 3D displays for conveying tele-presence, instead of a flat image of a remote user, have been developed. In the case of a shared-display-type, a flat shared display is only visible from the front [20]. Developed by Misawa *et al.*, LiveMask is a tele-presence surrogate system with a human-face-shaped 3D screen. It accurately conveys the user's gaze direction in 3D space on the basis of the results of tracking a remote user's faces and extracting their head motion [2], [21]. Pan and Steed preserve accurate cues of gaze direction by using a cylindrical screen with an array of cameras and an array of projectors [20]. Their system provides the same angle of view for all users sitting at a round table in the same local room. It was developed at low cost and is used for asymmetric conversations like the ones targeted in the present study. Local users are able to easily recognize eye information and tele-presence information of remote users via a 3D display. However, the amount of co-presence information conveyed in a local room is much more than amount of tele-presence information conveyed through the current videoconferencing system. Accordingly, in the authors' opinion, if local users don't pay attention to the 3D display, it will be hard for them to maintain the gaze direction of the remote user.

The approach taken in the present study is described as follows. Re-designing devices incurs high costs, and complicated systems or wearable devices, e.g., head-mounted displays (HMDs), are not suitable for casual or common types of business meetings. The aim of this study is to support natural communication while relieving much of the cognitive load or manipulation of a surrogate of the user (e.g., a tele-presence robot). Moreover, a system must be easy for an inexperienced user to operate. Based on a commercial videoconferencing system, the proposed system incurs low implementation cost without the need for re-designing. In short, it can be implemented by adding some functions to the existing system to make local users in the main party aware of the gaze direction of the remote-party user.

III. CONVEYING GAZE-AWARENESS WITH FAINT LIGHT

In this section, we propose a method for conveying gaze-awareness based on faint light. It is useful for conference-room meetings in which only a few participants are using a video-conferencing system at a remote site.

It is assumed that a creative meeting with several participants, of which one or two participate in the meeting from a remote site via a video-conferencing system, takes place. All participants at the local site can discuss things in the same room naturally. They are not aware of the gaze of the participants from the remote site. In particular, they are not made aware of the remote-site-participants' gaze directly; instead, they can become aware of a cue from the remote site given by a faint light.

Gaze is one of the most important pieces of non-verbal information. It is important for smooth communication, but it is not necessary for participants to know the precise point of gaze to understand others. The proposed method provides only an approximate point of gaze by using a faint light. This weak light does not bother the person talking, but the person being gazed at is made aware of the attention of the speaker.

The proposed method is overviewed in Fig. 1 and involves

the following steps. First, the gaze direction of the remote participant is detected; second, and a light is shone on the corresponding participant at the local site. These steps are summarized as follows.

- 1) As for determining the gaze direction of the remote participant, a full-faced eye camera is not used because the expressions of the participants are important non-verbal information. Some products [22], [23] can detect gaze direction by camera.
- 2) The person at the local site whom the remote participant gazed at is detected.
- 3) A faint light is shone on the corresponding person. The light is focused on the table in front of the person. Fig. 2 shows an example of the faint light. The light does not interrupt talking because it is not dazzling.



Fig. 2. Appearance of faint light.

IV. EXPERIMENT

We describe an experiment to evaluate our method.

A. Design of Experiment

The proposed method was experimentally evaluated in the case of two types of meetings. The first type is a meeting with a scenario. The scenario for each participant describes turn-taking opportunities and gaze-direction changes. Whether participants in the meeting room can be made aware of the gazes of remote participants with a faint light was evaluated. The other type of meeting is a negotiation during consensus games such as “Desert Survival” and “NASA”.

B. Experiment Conditions

The experiments were performed under the following three meeting conditions.

- 1) **Face-to-face meeting (“Face”)** All participants are seated at a four-way round table in a room (see Fig. 3).
- 2) **Video conference (“Video”)** One of the participants is seated in another room. The other participants are seated at the table. They use the video-conferencing system to communicate with the remote participant (see Fig. 4).
- 3) **Video conference with gaze-awareness (“Light”)** The proposed method is used for conveying gaze-awareness in a video conference. Fig. 2 shows the faint light used in

this experiment.

C. Participants and Apparatus

Four subjects participated in the experiment. During the video conference, one of them (the remote participant) had a seat in another room. The other participants had seats at a four-way round table in the meeting room. For the video conference, the video-conferencing-system display was placed at the vacant seat of the remote participant. All of the participants had the chance to be at the remote site because we had four meetings for each meeting condition.



Fig. 3. Face-to-face meeting.



Fig. 4. Video-conferencing meeting.

The video-conferencing system used was a Polycom HDX-8000[6] with a 50-inch display. The expressions of the remote participant can be recognized because the quality of the video is high definition (1080p). The remote participant talks naturally in the same way as those talking in the meeting room because the system controls their voice.

D. Results for Scenario Meeting

The results of a questionnaire in the case of meetings with a scenario are listed in Table I. The questions in the questionnaire are based on the five-point Likert scale. According to these results, the participants at the local site cannot be made aware of the gaze directions of the remote participant by the video conference only (see “Video” column in Table I). However, they can be made aware of the gaze directions of the remote participant by video conference using the proposed method (see “Light” column in Table I). The results in the case the proposed method was applied are close to that obtained under the face-to-face meeting condition.

E. Results for Consensus Meeting

The conditions of the consensus meetings in the experiment are the same as those for the “Video” and “Light” scenario

meetings, because the same participants could not discuss the same game, so two groups of participants were formed. The results of the questionnaire in the case of consensus-game meetings are listed in Table II. The questions in the questionnaire are based on the five-point Likert scale. The values in the table are averages, and the standard deviations are shown in parentheses.

TABLE I: RESULTS FOR MEETING WITH SCENARIO

Answerer	Question	Face	Video	Light
Local	Was aware of gazes	4.50	2.25	4.25
Remote	Felt they were aware	-	2.00	4.00

TABLE II: RESULTS FOR CONSENSUS MEETING

Answered by local participants		
Question	Video	Light
(Gaze of remote participant)		
L1. Was aware when listening	2.2 (1.17)	3.7 (0.82)
L2. Was aware when talking him/herself	1.7 (0.52)	3.7 (1.03)
L3. Payed attention because of gaze	2.5 (1.22)	3.7 (0.82)
(Talk by remote participants)		
L4. Was aware when neighbor is talking	3.7 (1.03)	4.4 (1.21)
L5. Was aware when talking him/herself	3.2 (0.98)	3.9 (0.98)
L6. Did Not talk only by local participants	3.7 (1.03)	4.5 (0.55)
Answered by remote participant		
Question	Video	Light
(Awareness by local participants)		
R1. Felt attention when local person talked	2.0 (1.41)	3.5 (0.71)
R2. Felt attention when I talked	1.5 (0.71)	4.5 (0.71)
(Talk by remote participants)		
R3. Joined in talk easily	3.5 (0.71)	4.5 (0.71)
R4. Felt they were aware of me talking	2.5 (0.71)	4.0 (0.00)
R5. Did not miss a chance to talk	2.5 (0.71)	4.0 (0.00)
R6. Did not feel it was hard to talk	3.0 (0.00)	5.0 (0.00)
R7. Did Not talk only by local participants	4.5 (0.71)	5.0 (0.00)

The results of L1 to L3 show that the participants at the local site could not be made aware of the gaze of the remote participants by the video conference only. The awareness while they talked themselves is the minimum because they focused on their own talks. The results of R2 show that remote participants felt a large difference between the “Video” and “Light” experiments. However, the results of L3 show that the local participants did not feel such a large difference. The results of L4 to L6 show that the local participants did not feel a large difference between the meetings using the “Video” and “Light” systems when they were talking. However, the results of R4 to R6 show that the remote participants felt it was hard to start talking.

V. DISCUSSION

In the experiments, a popular video-conferencing system was used. It can provide high-quality video stream on a 50-inch display. Such a system is usually used as a shared-display type in a conference room. Moreover, there were only three participants at the local site. When they look at the display intently, they can determine whether the gaze direction of the remote participant is to the left or to the right. However, the results presented in Section 4 show that they were not aware of the gaze. This means that they did not look at the display. It is easy to be aware of real motions in front of oneself, but it is not easy to be aware of virtual motions to the side of a display. In addition, the local participants did not consider this fact a big problem, although the remote

participant considers it a big problem. The difference between the local and the remote situations makes uncomfortable for all participants.

In another experiment, the faces of all the participants were recorded on video, and the gaze directions of all the local participants were analyzed. The results of this analysis are listed in Table III, which indicates the number of times the local participants gaze at the person talking (while they do not talk). The numbers in the table shows gaze times per person on average. In the meetings using “Video” system only, the number of times a local participant looked at another local participant seated at the same table is three times the number of times a local participant looked at the remote participant (35 and 12, respectively). In contrast, in the meetings using the “Light” system, the numbers of gaze to the local and remote participants are almost the same (27 and 24, respectively).

TABLE III: NUMBER OF TIMES LOCAL PARTICIPANTS GAZE WHILE NOT TALKING

Whom	Video	Light
Another person in local	35/person	27/person
Remote	12	24
Total	82	78

Some researches [20], [21] provide extraordinary display mechanisms that convey gaze directions on an advanced display easier than on an ordinary display. However, gaze directions of participants in the real world are much easier to understand. Moreover, it is not certain whether participants often look at the advanced display. On the other hand, the proposed method does not convey exact gaze directions. Instead, it provides a kind of awareness to pay attention to remote participants via an ordinary shared display.

Although video-conferencing systems provide high-quality video streams, they cannot help local participants pay attention to remote participants. It is not fair discussing place because local participants do not realize the problem. The proposed method can provide an environment for fair discussions via a typical commercial video-conferencing system with a shared display.

VI. CONCLUDING REMARKS

A method for conveying gaze awareness (by using a faint light with a typical video-conferencing system) is proposed. Many meetings are usually held via video-conferencing systems. Most participants are seated at a table in a conference room, but a few participants are in a remote location and use the shared display of the video-conferencing system. In this situation, the participants in the room mostly talk among themselves, because of a lack of non-verbal information from the remote site.

In the experiments, under normal conditions, the local participants gaze at another local participant seated at same table three times more than they look at the remote participant. When the video-conferencing system supports gaze-awareness by utilizing a faint light, the local participants gaze at the remote participant in the same way as they gaze at other local participants. The proposed method supports natural communication via a commercial video-conferencing

system at low implementation cost without having to redesign the shared display.

Currently, the method does not improve the appearance of the remote participants on the shared display. In future work, the proposed method will be evaluated in combination with an advanced display.

ACKNOWLEDGMENT

We would like to thank all our participants in the experiments for their time and effort. We would like to thank Chie Kadowaki for her academic advices.

REFERENCES

- [1] D. T. Nguyen and J. Canny, "More than face-to-face: Empathy effects of video framing," in *Proc. the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 2009, pp. 423, 432.
- [2] K. Misawa, Y. Ishiguro, and J. Rekimoto, "LiveMask: A telepresence surrogate system with a face-shaped screen for supporting nonverbal communication," *Journal of Information Processing Society of Japan*, vol. 21, no. 2, pp. 295, 303, 2013.
- [3] M. Shikida and Y. Masuda, "Awareness of turn-taking for remote group communications," *Journal of Information Processing Society of Japan*, vol. 56, no. 1, pp. 126, 136, 2015.
- [4] R. Vertegaal, I. Weevers, C. Sohn, and C. Cheung, "GAZE-2: Conveying eye contact in group video conferencing using eye-controlled camera direction," in *Proc. the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 2003, pp. 521, 528.
- [5] A. Kendon, "Some functions of gaze-direction in social interaction," *Acta Psychologica*, vol. 26, pp. 22, 63, 1967.
- [6] Polycom. Hdx series. [Online]. Available: <http://www.polycom.com/products/services/hd-telepresence-video-conferencing/realpresenceroom/realpresence-room-hdx-series.html>
- [7] R. Vertegaal, "The GAZE groupware system: Mediating joint attention in multiparty communication and collaboration," in *Proc. the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1999, pp. 294, 301.
- [8] T. Ohno, "Weak gaze awareness in video-mediated communication," *CHI'05 Extended Abstracts on Human Factors in Computing Systems*, ACM, 2005, pp. 1709, 1712.
- [9] J. Sun and H. Regenbrecht, "Implementing three-party desktop videoconferencing," in *Proc. the OzCHI'07*, ACM, 2007, pp. 95, 102.
- [10] M. Böcker and L. Mühlbach, "Communicative presence in videoconferences," in *Proc. the Human Factors and Ergonomics Society 37th Annual Meeting*, 1993, pp. 249, 253.
- [11] K. Okada, F. Maeda, Y. Ichikawaa, and Y. Matsushita, "Multiparty videoconferencing at virtual social distance: MAJIC design," in *Proc. the 1994 ACM Conference on Computer Supported Cooperative Work*, ACM, 1994, pp. 385, 393.
- [12] D. Nguyen and J. Canny, "MultiView: Spatially faithful group video conferencing," in *Proc. the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 2005, pp. 799, 808.
- [13] Y. Bondareva, L. Meesters, and D. Bouwhuis, "Eye contact as a determinant of social presence in video communication," in *Proc. the International Symposia on Human Factors in Telecommunication*, 2006.
- [14] D. Nguyen and J. Canny, "MultiView: Improving trust in group video conferencing through spatial faithfulness," in *Proc. the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 2007, pp. 1465, 1474.
- [15] A. Jones, M. Lang, G. Fyffe, X. Yu, J. Busch, I. McDowall, M. Bolas, and P. Debevec, "Achieving eye contact in a one-to-many 3D video teleconferencing system," *ACM Trans. Graph.*, vol. 28, no. 3, pp. 64:1-64:8, 2009.
- [16] I. Kawaguchi, H. Kuzuoka, and Y. Suzuki, "Study on gaze direction perception of face image displayed on rotatable flat display," in *Proc. the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 2015, pp. 1729, 1737.
- [17] Y. Bondareva and D. Bouwhuis, "Determinants of social presence in videoconferencing," in *Proc. the AVI'04 Workshop on Environments for Personalized Information Access*, 2004, pp. 1, 9.
- [18] T. Inoue, "Mixed reality meeting system enabling user to keep and share interpersonal distance in the real world," *Journal of Information Processing Society of Japan*, vol. 50, no. 1, pp. 246.253, 2009.
- [19] S. O. Adalgeirsson and C. Breazeal, "MeBot a robotic platform for socially embodied telepresence," in *Proc. the 5th ACM/IEEE International Conference on Human-Robot Interaction*, IEEE Press, 2010, pp. 15, 22.
- [20] Y. Pan and A. Steed, "A gaze-preserving situated multiview telepresence system," in *Proc. the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 2014, pp. 2173, 2176.
- [21] K. Misawa, Y. Ishiguro, and J. Rekimoto, "LiveMask: A telepresence surrogate system with a face-shaped screen for supporting nonverbal communication," in *Proc. the International Working Conference on Advanced Visual Interfaces*, ACM, 2012, pp. 394, 397.
- [22] SmartEye. Eye tracking systems. [Online]. Available: <http://smarteye.se/products/category/eye-tracking-systems/>
- [23] The EyeTribe. The eye tribe tracker. [Online]. Available: <http://theeyetribe.com/products/>



Mikifumi Shikida received a Ph.D. in engineering from Tokyo Institute of Technology in 1995.

He is a professor at Research Center for Advanced Computing Infrastructure, Japan Advanced Institute of Science and Technology, Japan.

His research interests are computer supported cooperative work, office information systems, and large-scale network services.