Supporting the Initiation of Remote Conversation by Presenting Gaze-based Awareness Information

Aoto Tanokashira[1], Ikkaku Kawaguchi[2], Buntaro Shizuki[2], and Shin Takahashi[2]

[1] University of Tsukuba, Tsukuba, Japan tanokashira@iplab.cs.tsukuba.ac.jp
[2] University of Tsukuba, Tsukuba, Japan
{kawaguchi, shizuki, shin}@cs.tsukuba.ac.jp

Abstract. Conversations among remote people cannot be easily initiated. A lack of non-verbal information in video calls contributes to the difficulty in initiating conversations. Conversely in face-to-face situations, conversations are initiated by exchanging gaze information, particularly mutual gaze. However, the mutual gaze cannot be easily established between remote locations. In this study, we proposed a voice call system with robots to initiate conversations between remote people by exchanging gaze information. The evaluation results showed that the proposed system was effective in reducing the psychological burden of initiating conversations between remote people.

Keywords: remote communication · awareness · gaze interaction · informal communication · communication robot.

1 Introduction

Initiating conversations, particularly informal communication, between remote people is difficult. Informal communication occurs incidentally, without a fixed schedule or agenda [2]. To initiate informal communication between people in remote locations, it is necessary to reduce the burden of initiating conversations and make it easier to talk to remote people. Initiating conversations will become easy by understanding whether the remote person is available for conversations. Previous research [4,7] has realized this assumption by providing awareness information about the remote person's availability for conversations. By contrast, in a face-to-face situation, people use non-verbal information to present the detailed state (e.g., to what extent is the other person focused on a task) of each other, so that they can initiate conversations without disturbing the other person's task. In particular, the exchange of gaze information plays an important role in initiating conversations. When people initiate conversations, they exchange gaze. Mutual gaze is a trigger for initiating conversations. In this study, based on the sociological knowledge about initiating conversations, we reproduced the exchange of gaze that leads to the initiation of conversations between remote

Fig. 1. Actual system setup. Users exchange gaze through the proposed system. When a mutual gaze is established, a voice call starts.

people (Fig. 1). We designed a voice call system for remote workers to initiate one-on-one conversations. This system is composed of robots that can exchange gaze between remote locations. When the mutual gaze is established, a voice call starts. We conducted an experiment to clarify the effectiveness of the proposed system.

2 Related Work and Our Approach

To initiate conversations between remote people, whether the remote person is available for conversations need to be determined. Previous research has focused on solving this problem by providing awareness information about the remote person's availability for conversations [4, 7]. However, only a few works have presented the detailed state of availability that typically precedes conversations in face-to-face situations.

When people initiate conversations in face-to-face situations, they use non-verbal information to present their detailed state to each other. In particular, gaze information plays an important role in initiating a conversation. According to Salvadori [11], before a person initiates a conversation, the initiator assesses the colleague's progress of work and availability. The initiator may use proximity, movement, or gaze to get the colleague's attention. The use of non-verbal information by the initiator allows the colleague to finish his/her task before acknowledging the initiator's request for attention. The colleague displays the termination of his/her tasks by orienting toward the initiator or looking toward the initiator. Thereafter, they start a conversation. Kendon [5] claimed that gaze is a fundamental feature of face-to-face interactions. The shifts of gaze coordinate the timing of speech production. Kendon [6] also showed that mutual gaze can trigger a conversation. Some previous systems [3, 8] for presenting awareness information can present not just the availability of a remote user but also non-verbal information. For example, Roussel et al. [10] proposed an always-connected video media space that continuously presents the remote user with non-verbal information. Although these systems are capable of providing non-verbal information, privacy issues still arise because of the use of a video [9].

The exchange of gaze plays an important role in initiating conversations. However, exchanging gaze between remote people without using a video is dif-

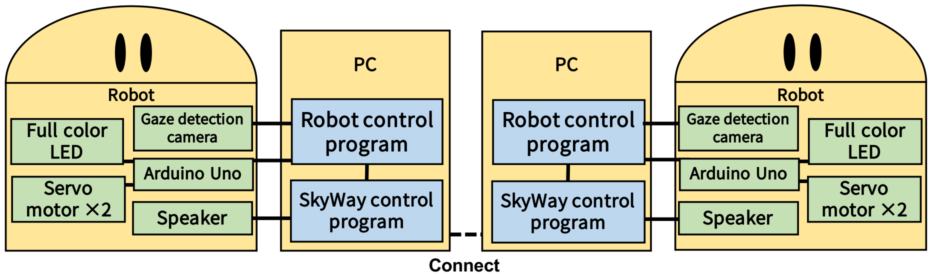

Fig. 2. System configuration.

ficult. In this study, we propose a voice call system with robots to convey gaze information to each other between remote people. The proposed system reproduces the exchange of gaze information that occurs when initiating conversations in face-to-face situations without using a video. The system design is presented as follows: (1) The initiator sends gaze information to the remote person to get the remote person's attention, as in a face-to-face situation. (2) The remote person's response to the initiator's gaze (return gaze or not) is conveyed to the initiator. (3) When a mutual gaze is established, a voice call starts. Based on this system design, the proposed system conveys gaze information to each other through robots placed on each location, which can detect and represent gaze. The initiator can then present the degree to which he/she wants a response based on the length of time he/she sends gaze to the robot.

3 Proposed System

We developed the voice call system with robots that can detect and present gaze and start a voice call. In this section, we describe the hardware, software, and interaction designs of the proposed system. An overview of the proposed system configuration is shown in Fig. 2.

3.1 Hardware Implementation

OMRON's HVC-P2 [3] was used for the gaze detection. HVC-P2 can detect the direction of the face and gaze. In the proposed system, we used the face direction detection to detect the user's gaze direction, because the accuracy of the gaze direction sometimes becomes unstable and the reliability is insufficient. HVC-P2 was controlled by a control PC. The detected face direction values were sent to the robot. Two servomotors and full-color LEDs were used to present the gaze. These devices were controlled by Arduino UNO. The robot was equipped with a small speaker used for voice calls. The robot body was fabricated using a three-dimensional printer, and its total height was approximately 180 mm. We created two units of robots to place one robot on each desk with two remote workers.

[3] https://plus-sensing.omron.co.jp/product/hvc-p2.html

3.2 Software Implementation

We implemented two functions for the proposed system: the SkyWay[4] control program and the robot control program. The SkyWay control program is used for sending/receiving gaze information and voice calling between remote locations. The SkyWay control program sends and receives gaze information and also makes voice calls according to the gaze information.

The robot control program obtains the face direction of a local user from HVC-P2. The robot control program sends the obtained gaze information to the SkyWay control program to convey the gaze information to the remote location. WebSocket is used to communicate with the SkyWay control program. The remote user's gaze information received from the SkyWay control program is sent to the Arduino Uno using serial communication to present the gaze.

3.3 Interaction Design

At the beginning of the interaction, the initiator looks toward the robot. The length of time the initiator sends gaze to the robot conveys the degree to which the initiator wants the remote person's response (Fig. 3-1). For example, if the initiator wants to talk to the remote person who is busy, then the initiator will look toward the robot for a long time. By contrast, if the initiator wants to speak to the remote person when he/she is not busy, then the initiator will look toward the robot for a short time. The remote person returns the gaze if he/she can respond and does not return if he/she cannot. He/she can recognize the degree to which the initiator wants a response from the gaze information presented by the robot (Fig. 3-2). He/she can decide whether to respond or not depending on that degree. If he/she returns the gaze, then a mutual gaze is established through the robot and a voice call starts (Fig. 3-3).

The proposed system changes the color of the eyes according to the gaze information of the users. Basically, the eye color is set to red. When only the initiator is looking toward the remote person, the eye color is yellow. When a mutual gaze is established, the eye color changes to green.

4 Evaluation

4.1 Experimental Design

This experiment aims to evaluate whether the proposed system has more advantages than previous systems in terms of presenting awareness information. Accordingly, we set three hypotheses on the effects of the proposed system.

- **H1.** It reduces the psychological burden of initiating conversations.
- **H2.** It reduces the burden of operating the system to initiate conversations.
- **H3.** It increases the number of attempts to initiate conversations.

[4] https://webrtc.ecl.ntt.com/

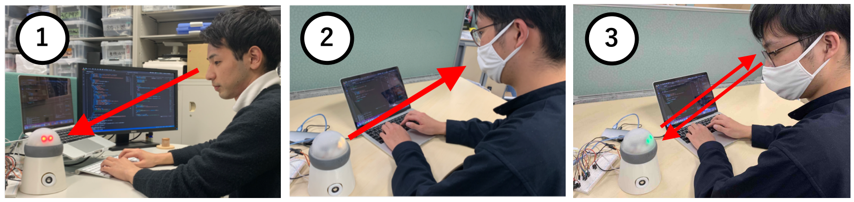

Fig. 3. interaction steps: (1) The initiator sends gaze to the robot. (2) The remote person receives the gaze. (3) The remote person returns the gaze, and a mutual gaze is established.

To initiate conversations between remote locations, the burden of initiating conversations should to be reduced. Specifically, we focused on the psychological burden and the burden of operating the system. Reproducing the non-verbal interaction in a face-to-face situation before initiating a conversation, it may become easier to talk to the remote person. Thus, the psychological burden of initiating conversations is considered to be reduced (H1). Moreover, by starting voice calls with only gazing, the usability of the system is improved, and the burden of operating the system to initiate conversations is reduced (H2). As a result, the number of attempts to initiate conversations will increase (H3). To test the hypothesis, the following two experimental conditions were set.

– **C1. Gaze presentation condition** This condition was the proposed system in this study. We presented the gaze as the user's awareness information through the robot. When the user wanted to initiate conversations, he/she looked at the robot, and if mutual gazing was established, then a voice call starts.
– **C2. Light presentation condition** In this condition, we used the same system with C1, but did not use gaze detection and presentation function. The users operated the system through the GUI on the PC and made/received a call. The user's availability was presented through light emission. To make it look like a simple LED indicator rather than the actual eyes of the robot, only one of the LEDs in the robot's eye was turned on. Basically, the light was set to red. The light changed to yellow when a voice call was being requested, and if a voice call started, then the light changed to green. This specification was based on the previous system [4, 7] providing awareness information about the remote person ' s availability for conversations.

In this experiment, the participants were asked to do tasks by using a PC according to the a task list(e.g. make a table from multiple data). The participants could interact with the experimenter in another room at any time via the system. The reason for employing an experimenter was to control the contents of the conversation and reduce the effect except of conditions. To evaluate the impression of initiating conversations, the participants needed to talk to the experimenter. Thus, we set several tasks that triggered questions in the task list.

To evaluate the impressions of being talked to, the experimenter talked to the participants four times in each trial. The participants sat in front of the PC for the task, and the robot was placed on the left side of the PC (45° from the front of the user). The experiment was conducted in a between-subject design. A total of 20 participants (14 males and 6 females, mean age 22.1 = years, and college students) participated in the experiment, 10 for each condition.

In this study, three assessment items were set to test the hypotheses. To test H1, we conducted the evaluation of impressions related to the psychological burden. We used a questionnaire for evaluating the emotional benefits and costs of communication systems proposed by Yarosh et al [12]. We used one benefit scale (Presence-In-Absence) and two cost scales (Feeling Obligated and Unmet Expectations) from the questionnaire. We set these scales because we thought that improving the benefit scale(Presence-In-Absence) and decreasing the cost scales(Feeling Obligated and Unmet Expectations) would lead to a reduction in the psychological burden. To test H2, we conducted the evaluation of system usability. We used the System Usability Scale (SUS) [1], which evaluates the usability of a system. To test H3, we analysed the number of attempts to initiate a conversation. We used video recordings of the work for analysis.

4.2 Results and Discussion

In this section, we present and discuss the evaluation results. The first two participants in this experiment were not used in the analysis because of the instability of the system.

Impressions related to the psychological burden Based on the questionnaire results, Welch's t-test was conducted on the participants' mean scores on each scale. Fig. 4 shows a graph of the mean scores for each experimental condition for each scale.

For the Presence-In-Absence scale, a significant difference was found between the conditions ($t(15.63) = 2.44$, $p = .027$). The experiment results showed that presenting gaze significantly improved the effectiveness of conveying the presence of the remote user to the local user. For the Feeling Obligated scale, a significant difference was found between the conditions ($t(11.83) = 2.50$, $p = .028$). The results showed that presenting gaze significantly increased the sense of obligation to communicate. By contrast, the scores were in the positive range (< 4), implying that a negative impression was not given on the sense of obligation in both conditions. For the Unmet Expectations scale, no significant difference was found between the conditions ($t(14.08) = 1.19$, $p = .25$).

System usability Based on the questionnaire results, Welch's t-test was conducted on the participants' scores. The mean SUS scores for each condition are shown in Fig. 5. The test results showed no significant difference between the conditions ($t(14.71) = .70$, $p = .49$). Conversely, the mean SUS score (gaze: 81.67, light: 78.06) was high in both conditions [1].

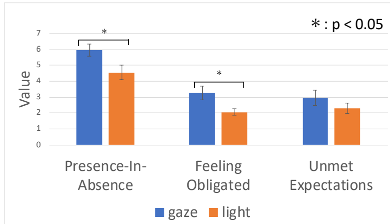

Fig. 4. Results of thr questionnaire related to psychological burden. Error bars indicate the standard error.

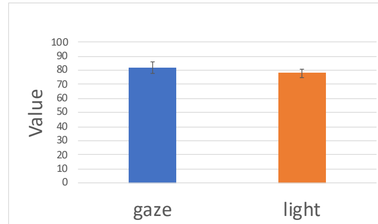

Fig. 5. Result of SUS. Error bars indicate the standard error.

Number of attempts to initiate a conversation A Welch's t-test was conducted using the data from the video recording analysis for H3. In the light presentation condition, the mean number of attempts was 2.67 (SE = 0.44). In the gaze presentation condition, the mean number of attempts was 2.78 (SE = 0.55). The test results showed no significant difference between the conditions ($t(15.31) = .16$, $p = .88$).

Discussion In this study, we set three hypotheses (H1 - H3). For H1, the presence of the remote person could be felt, and it became easy to talk to them. Moreover, the sense of obligation to communicate increased, but not to the extent that a negative impression was produced. These results suggest that the system was partially effective in reducing the psychological burden of initiating conversations, so H1 was partially supported. For H2, although it was not supported, the burden of operating the system to initiate conversations was considered to be small because the system usability was high. For H3, the number of attempts to initiate a conversation did not improve, so H3 was not supported.

The evaluation results suggest that the proposed system was partially effective in reducing the psychological burden of initiating conversations. However, it was not effective in facilitating the initiation of conversations, probably because of the experimental environment. In this experiment, all participants performed a preset task list. Conversations were limited to only the minimum necessary questions in the task, and the number of conversations might depend on the task in the experiment. In addition, an experimenter controlled the contents of the conversation. In a real environment, contents and the number of conversations are not controlled. Hence, in the future, we will conduct experiments in a real environment. Moreover, in this study, whether the results are due to the gaze detection or the gaze presentation is unclear. Thus, we will evaluate the effects of gaze detection and gaze expression separately. The proposed system may result in wrong interpretations among users because the system can only provide limited awareness information through the gaze channel. We will incorporate other non-verbal information into our system.

5 Conclusion

In this paper, we proposed a voice call system with robots that supports initiating conversations between people in remote locations by presenting gaze information. The proposed system can exchange gaze between people in remote locations. When a mutual gaze is established, a voice call starts. We conducted an experiment to clarify the effectiveness of the proposed system. The evaluation results suggest that the proposed system was partially effective in reducing the psychological burden of initiating conversations. Based on the results, it can be said that the burden of operating the system to initiate a conversation is small. However, no significant difference was found between the conditions about the number of attempts to initiate a conversation, which can be attributed to the experimental environment. Thus, int the future, we will conduct experiments in a real environment.

References

1. Brooke, J.: Sus: a "quick and dirty'usability. Usability evaluation in industry p. 189 (1996)
2. Daft, R., Lengel, R.: Information richness. a new approach to managerial behavior and organization design. Research in Organizational Behavior (1983)
3. Dou, M., Shi, Y., Frahm, J.M., Fuchs, H., Mauchly, B., Marathe, M.: Room-sized informal telepresence system. In: 2012 IEEE Virtual Reality Workshops (VRW). pp. 15–18 (2012). https://doi.org/10.1109/VR.2012.6180869
4. Greenberg, S.: Peepholes: Low cost awareness of one's community. In: Conference Companion on Human Factors in Computing Systems. pp. 206–207 (1996)
5. Kendon, A.: Some functions of gaze-direction in social interaction. Acta psychologica **26**, 22–63 (1967)
6. Kendon, A.: Conducting interaction: Patterns of behavior in focused encounters, vol. 7. CUP Archive (1990)
7. Kuzuoka, H., Greenberg, S.: Mediating awareness and communication through digital but physical surrogates (01 1999). https://doi.org/10.1145/632716.632725
8. Kuzuoka, H., Kodama, Y., Xu, J., Myodo, E., Harada, E., Osawa, H.: Telepresence robot's salutations to trigger informal conversation with teleworkers. In: Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing. pp. 233–236 (2018)
9. Neustaedter, C., Greenberg, S., Boyle, M.: Balancing privacy and awareness for telecommuters using blur filtration (2003)
10. Roussel, N., Evans, H., Hansen, H.: Mirrorspace: using proximity as an interface to video-mediated communication. In: International Conference on Pervasive Computing. pp. 345–350. Springer (2004)
11. Salvadori, F.A.: Open office interaction : initiating talk at work(doctoral dissertation). In: King's College London (2016)
12. Yarosh, S., Markopoulos, P., Abowd, G.: Towards a questionnaire for measuring affective benefits and costs of communication technologies. pp. 84–96 (02 2014). https://doi.org/10.1145/2531602.2531634