Robot Mediated Round Table: Analysis of the Effect of Robot's Gaze

Michita Imai, Takayuki Kanda, Testuo Ono, Hiroshi Ishiguro, Kenji Mase
ATR Media Information Science Labs.
2-2 Seika, Soraku-gun, Kyoto
619-0288 Japan
E-mail michita@ics.keio.ac.jp

Abstract

This paper investigates the relationship between the direction of a robot's face and its gaze as noticed by a person. Non-verbal information such as gaze movement or gestures is crucial to conversation between people. In particular, the effect of a gaze is most important in non-verbal communication because gaze movements convey various information such as timing of taking turns or a target indicated by a person. This paper argues that the important role of a gaze also applies to interaction between a human and a robot. In spite of the importance of the gaze, there has been no study on the relationship between a robot's head orientation and its gaze. We have conducted an experiment to investigate this relationship in an experimental environment named RM-RT, where the robot's head orientation is fixed to a particular person during a conversation. The results indicate that people notice the robot's gaze according to the robot's head orientation.

I. Introduction

Following the development of Honda's P-3/ASIMO [4] or Sony's AIBO [3], the dream of a communication-robot appears to have come true. To improve the communication skills of robots, this paper deals with the eye-contact between a person and a robot. In other words, it concentrates on a robot's gaze toward a person.

There are several difficulties in getting a robot to achieve eye-contact with a person. That is, the robot must obtain a person's information from a vision sensor: where is the location of the person, where is the position of his/her face, and where does he/she gaze?

Several studies have investigated the difficulties in achieving eye-contact. For example, a social robot named Kismet [2] has the ability to find a person's face location to engage in social interaction with him/her. In a humanoid named SIG [8], audio and visual data streams are integrated in search of the exact location of a speaker person because the location only detected from one of the data streams is less accurate. In addition, a humanoid named Infanoid [6] attempts to recognize the exact face position and the direction in which the person pays attention.

However, even if the above technologies are fully developed, an important issue remains: the effective zone of the robot's gaze toward a person. In short, where should the robot's head turn to make the person notice its gaze? This issue is significant for success in achieving eye-contact. However, the effective gaze zone of the robot to the person has not yet been investigated.

This paper conducts a psychological experiment to investigate the effective gaze zone. Strictly speaking, it investigates the orientation of the robot's head that makes the person notice its gaze. The effective gaze zone is examined from the viewpoints of both the person to whom the robot's head turns and the person to whom it does not.

In addition, this paper develops an experimental environment named RM-RT (Robot-Mediated Round Table) for the psychological experiment. RM-RT is a communication environment where eight subjects sit down and surround a communication robot named Robovie. While they communicate with each other (except for Robovie), Robovie tries to communicate with one of the subjects by turning its head to face him/her depending on the experimental condition. Robovie gazes at the subjects during an actual conversation. Since they do not concentrate on the robot's gaze intentionally, RM-RT can examine the unconscious effective gaze zone.

The remainder of the paper is organized as follows. Section 2 explains the robot's gaze and proposes a hypothesis. Section 3 describes the RM-RT environment, and Section 4 explains a psychological experiment conducted under RM-RT to confirm the hypothesis. Section 5 discusses the robot's effective gaze zone in terms of the experimental results. Section 6 concludes the paper with a summary.

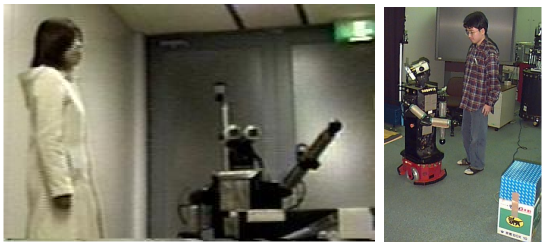


Fig. 1. Left: the robot refers to a poster on a wall with gaze movements and hand gestures. Right: the robot achieves eye-contact with the person.

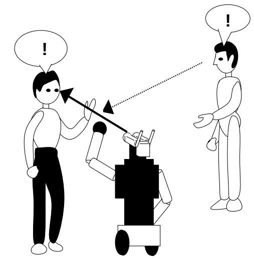


Fig. 2. The robot turns its head to face one person, and the other person observes the behavior.

II. Robot's Gaze

A. Gazing at real world objects

In a conversation between people, a real world object (other than the speaker and hearer) frequently becomes a main topic. If the object is in sight, they refer to it by pointing at it with gaze movements and hand gestures. These ordinary behaviors are also essential for natural interaction between a person and a robot.

We conducted an experiment to confirm the importance of gaze movements when a robot pointed at an object [5] (left picture of Figure 1). In the experiment, the robot generated an utterance "please look at this" while pointing at a poster on a wall with gaze movements and hand gestures.

The results indicated the effect of gaze movements on the pointing hand gesture. For instance, subjects realized that the robot referred to the poster by the utterance when the robot turned its head to them and carried out eye-contact with them. On the other hand, they did not become aware of the reference when it did not carry out eye-contact.

The experimental results thus suggest that eye-contact is also important in the interaction between a person and a robot. To achieve obvious eye-contact between them, this paper examines the condition where the person notices the robot's gaze.

B. Robot's gaze in eye-contact

The right picture of Figure 1 shows the eye-contact between a person and the robot. The following sequence of the robots behaviors achieves distinct eye-contact.

1. The robot finds the location of a person via a vision sensor and turns its head in that direction.
2. The robot recognizes the direction of the person's gaze and meets his/her gaze.

However, does the robot need to achieve such distinct eye-contact in usual interaction?

It is relatively easy for the robot to obtain the location of the person. For example, if it detects a skin color at the position where there is a difference between two frames by using an omni-directional camera, there is a good chance that the person will be there. The advantages of this method are low computational costs and simplicity of development. The main disadvantage is low precision of the angle at which the robot's head turns.

On the other hand, a technique for capturing a person's gaze direction already exists [7]. With this technique, it is possible for the robot to meet the person's gaze and to achieve distinct eye-contact. However, since the cameras on the robot move frequently, it is difficult to use this technique on an actual robot.

In comparison with the first method of conventional location detection, the second one is superior in terms of precision. However, the first method is appropriate for a robot in view of its vision system. To determine which one is appropriate, this paper investigates the required precision of a robot's gaze.

C. Power of gaze

This paper concentrates on the relationship between a robot's head orientation and its gaze to investigate the required precision of the robot's gaze. This paper uses the term "the effective gaze zone" for the range of a robot's head direction angle where a person can recognize the robot's gaze.

Figure 2 shows the relationship between the robot's head orientation and its gaze. The figure demonstrates two possibilities of a person noticing the robot's gaze: by the person to whom the robot's head turns (left person in Fig. 2) and by the person to whom the robot's head does not turn (right person in Fig. 2).

It is essential in human-robot interaction that the person toward whom the robot's head is facing notices the robot's gaze. To achieve the essential gaze, this paper investigates what angle of the robot's head can make the person notice its gaze.

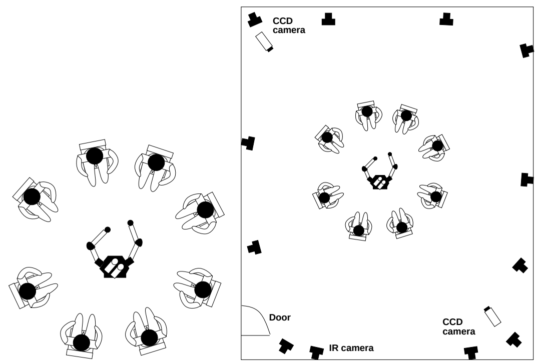


Fig. 3. Experimental environment RM-RT

In addition, it is vital that the person who is not gazed notices the robot's gaze toward the other person because information about who is the speaker and who is the hearer is crucial to the interpretation of a generated utterance [1]. Therefore, regardless of whether a person is gazed at, he/she must notice the robot's gaze to interact with the robot. This paper also investigates whether a person notices the robot's gaze when it turns its head to face another person.

This paper puts forward the following hypothesis in terms of the relationship between the robot's head orientation and its gaze.

Hypothesis 1: When two persons are near each other, the person to whom the robot turns its head notices the robot's gaze toward him/her and the other person to whom it does not turn its head does not notice the gaze toward him/her. Also, if the robot turns its head between them, they notice the gaze toward them simultaneously.
This paper confirms the hypothesis through interaction between persons and an actual robot.

III. Robot Mediated Round Table

A. RM-RT

This paper has developed an experimental environment named RM-RT to investigate the relationship between a robot's head orientation and its gaze (left figure in Fig. 3). In the figure, a communication robot named Robovie is in the center of RM-RT. The eight subjects sit down in a circle surrounding Robovie.

In RM-RT, the subjects must discuss two topics. Within the discussion, Robovie tries to communicate with one of the subjects by turning its head to face him/her. The head orientation is determined according to the experimental condition. In addition, Robovie's body does not turn during the discussion. Also, the subjects must not swap their chairs. Therefore, the combination of each head orientation and each position of the subjects is fixed in RM-RT.

RM-RT is also designed to examine the effect of Robovie's gaze when the subjects do not focus on its gaze intentionally. Strictly speaking, the interaction with Robovie itself is not included in the discussion topics.

B. Robovie

The Robot in the right picture of Fig. 1 is Robovie by itself. Robovie has two 4-DOF hands and a 3-DOF head to generate hand gestures or gaze with head motions. It also has a movable base and several types of sensors: touch sensors to notice being touched by persons, ultrasonic distance sensors to detect obstacles, and two types of vision sensors. One of the vision sensors is an omni-directional vision sensor able to capture panoramic scenes around Robovie. Robovie obtains the direction of a person's location by detecting skin color movements in an image captured by the omni-directional vision sensor. The other vision sensor includes two CCD cameras on Robovie's head for stereo vision. Robovie is also equipped with speech recognition and generation software.

In RM-RT, Robovie turns its head to only a few subjects to communicate with them. The duration of the interaction with one subject is from 5 to 10 minutes. The interaction consists of vocal and gesture communication.

IV. Experiment

[Subjects]

There were 36 subjects (men and women), and their average age was 19.8 years old. They were divided into five groups. Three groups consisted of eight subjects and the remaining two groups consisted of six subjects. In the latter two groups, the two seats behind Robovie were vacant.

[Experimental environment]

The right figure in Fig. 3 shows the experimental environment of RM-RT. This environment is an experimental room in our labs. RM-RT was set in the center of the room. Also, there were twelve IR cameras around RM-RT to capture the head orientations of the subjects and Robovie. The subjects wore caps on which markers were for tracking by the IR cameras. In addition, there were two CCD cameras to record the experimental scenes.

[Experimental conditions]

Figure 4 shows the position of each subject and Robovie as well as the direction in which it turns its

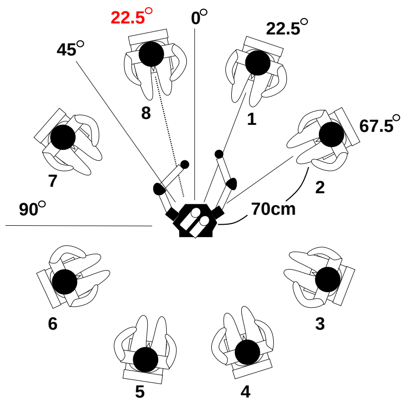


Fig. 4. Position of each subject and Robovie

Fig. 5. Experimental scene in RM-RT

head. In Fig. 4, each seat of the subjects was numbered from 1 to 8 to analyze the experiment. The subjects sat 70 cm from Robovie, and this distance from Robovie made it possible for them to touch Robovie.

Robovie turns its head in the several directions: 45 degrees to the left side, 90 degrees to the left side, 22.5 degrees to the right side, and 67.5 degrees to the right side. In short, Robovie turned its head to face two subjects sitting at number 1 and 2 seats. Also, it turned its head between numbers 1 and 8, between numbers 7 and 8, and between numbers 6 and 7. However, there was one experimental session where Robovie turned its head to the subject sitting at number 8. As a result, there were four experimental conditions. The subjects were also divided into four types according to the conditions: the subjects to whom Robovie turned its head (11 persons), the subjects near a position to which it turned its head (15 persons), the subjects next to a person to whom it turned its head (5 persons), and the subjects to whom it had never turned its head (5 persons).

[**Experimental sequence**]

In the experiment, all subjects played with Robovie for 10 minutes to become familiar with Robovie before beginning RM-RT. After the play session, the subjects sat down in the RM-RT seats at random and wore the caps with IR markers. After preparing RM-RT, the subjects discussed two topics: current problems of Japan and something like a game. After each topic was finished, the subjects answered questionnaires about the topics. Also, at the end of the two topics, the subjects were asked about the robot's gaze.

[**Evaluation method**]

The relationship between the orientations of Robovie's head and Robovie's gaze was investigated through the questionnaires. The questions were the following items.

- Did you notice the gaze from Robovie?
- Did you notice Robovie's gaze toward the other subjects? If any, who were they?

[Predictions]

This subsection predicts the experimental outcomes from the hypothesis and the experimental conditions. The following three items are predictions.

1. The subjects (sitting at seat numbers 1 ,2, and 8 when Robovie turns to 8) toward whom Robovie turns its head will notice Robovie's gaze.
2. The subject (sitting at seat number 3) next to the person to whom Robovie turns its head will not notice Robovie's gaze.
3. The subjects (sitting at seats numbers 6, 7, and 8) on either side when Robovie turns its head to the space between them will notice Robovie's gaze.

[Outcome]

The two upper photos of Figure 5 show experimental scenes in RM-RT. The lower photo shows the captured image from twelve IR cameras. Figures 6 and 7 show the outcome, where each graph represents the percentage of the persons who noticed the gaze (left bar) and who did not (right bar).

Figure 6 shows the answers to the question "did you notice the gaze from Robovie?" In other words, the number of subjects who noticed Robovie gazing at him/her. Each graph (a, b, and c) represents a comparison between the experimental conditions. The outcome in graph b shows there is the significance between the subjects to whom Robovie turned its head and the subjects next to a person to whom it turned its head.

Figure 7 shows the answers to the question “did you notice the Robovie’s gaze toward the other subjects?” In other words, the number of subjects who noticed Robovie gazing at the other subjects. Each graph (a, b, and c) also represents a comparison between the experimental conditions. The outcome in the graph c shows the significance between the subjects to whom Robovie turned its head and the subjects to whom it never turned its head.

[Verification of prediction 1.]

Figure 6 a. indicates that almost all subjects noticed Robovie’s gaze when Robovie turned its gaze to them, and Fig. 7 a. indicates the same outcome. This result suggests that prediction 1 is true.

[Verification of prediction 2.]

Figure 6 b. indicates that the subjects to whom Robovie turned its head noticed Robovie’s gaze but the subjects next to a person to whom Robovie turned its head did not notice Robovie’s gaze. This result suggests that prediction 2 is true. However, even though Figure 7 b. shows the same tendency as Figure 6 b., a few subjects noticed Robovie’s gaze at the subjects next to a person to whom Robovie turned its head.

[Verification of prediction 3.]

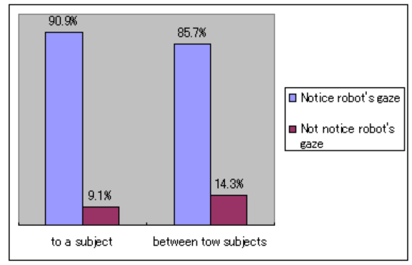
Figures 6 a. and 7a. do not show significance. This fact indicates that the two subjects between whom Robovie turned its head both noticed Robovie’s gaze. This result suggests that prediction 3 is true.

From prediction 1 to 3, the proposed hypothesis is verified.

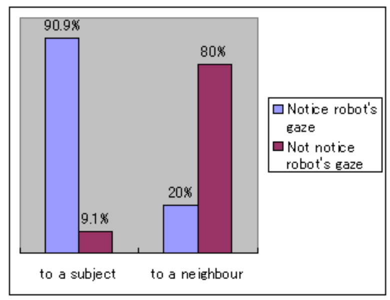
V. Discussion

This section discusses the robot’s head motion during a conversation from the viewpoint of the experimental outcome. Figures 6 a. and b. suggest the robot’s head motion when the robot does not determine who communicates with it. According to the outcome, the robot should turn its head between persons in such a situation. Also, if the robot intends to communicate with a specific person, it turns its heads to him/her exactly. With such an exact gaze, the people near the specific person do not notice its gaze.

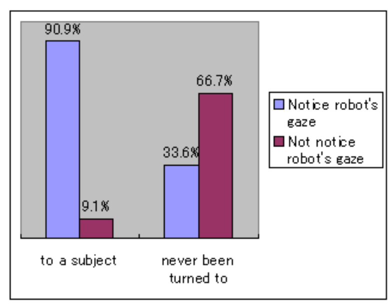
Moreover, the experimental outcome suggests that the robot can make the person notice its gaze even



a) comparison between subjects to whom Robovie turned its head and subjects near a position to which it turned its head. ($P > .1, U = 73$)

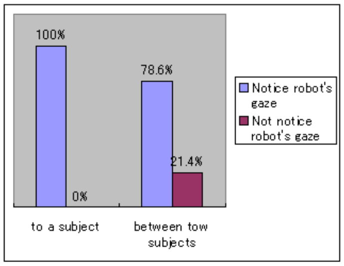


b) comparison between subjects to whom Robovie turned its head and subjects next to a person to whom it turned its head. ($*P < .05, U = 8$)

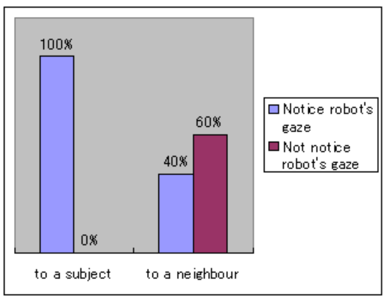


c) comparison between subjects to whom Robovie turned its head and subjects to whom it never turned its head . ($P < .1, U = 14$)

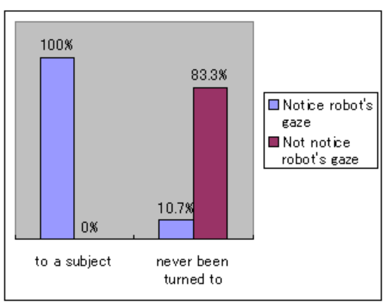
Fig. 6. Answers to the question “did you notice the gaze from Robovie?” Each graph represents the percentage of persons who noticed the gaze (left bar) and who did not (right bar).



a) comparison between subjects to whom Robovie turned its head and subjects near a position to which it turned its head. ($P > .1, U = 60.5$)



b) comparison between subjects to whom Robovie turned its head and subjects to whose next to a person to whom it turned its head. ($P < .1, U = 11$)



c) comparison between subjects to whom Robovie turned its head and subjects to whom it never turned its head. ($*P < .01, U = 5.5$)

Fig. 7. Answers to the question "did you noticed Robovie's gaze toward the other subjects?" Each graph represents the percentage of persons who noticed the gaze (left bar) and who did not (right bar).

thought the robot cannot obtain an exact location of the person or their gaze direction. In particular, the robot can select a person to communicate with even by using an easy gaze control. In short, the communication robot does not always recognize a person's gaze direction, but location information of the person is sufficient when interacting with people.

Also, RM-RT employs IR cameras to capture head orientations of the subjects and Robovie. In future work, we will try to analyze the data and investigate the relationship between head motions and Robovie's gaze much more precisely.

VI. Conclusions

This paper investigated the relationship between a robot's head orientation and its gaze. For the investigation, this paper developed an experimental environment named RM-RT, where subjects discussed topics with each other, and conducted a psychological experiment within RM-RT.

The outcomes indicated that the subjects to whom the robot turned noticed the robot's gaze. On the other hand, the subjects next to him/her did not notice the robot's gaze. If the robot turned its head between two subjects, they both noticed the robot's gaze.

Acknowledgement

This research was supported in part by the Telecommunications Advancement Organization of Japan.

References

[1] J. Barwise and J. Perry. *Situations and Attitudes.* Mit Press, 1983.
[2] C. Breazeal and B. Scasselati. A context-dependent attention system for a social robot. In *IJCAI-99*, pages 1146–1151, 1999.
[3] M. Fujita and K. Kageyama. An open architecture for robot entertainment. In *Proc. of Autonomous Agent'97*, pages 435–442, 1997.
[4] K. Hirai. Research of Honda humanoid robot. In *Proc. of JSAI '98*, pages 1–4, 1998.
[5] M. Imai, T. Ono, and H. Ishiguro. Physical relation and expression: Joint attention for human-robot interaction. In *In Proc. of 10th IEEE International Workshop on Robot and Human Communication (RO-MAN2001)*, pages 512–517, 2001.
[6] H. Kozima and A. Ito. Towards mindreading by an attention-sharing robot. In *Proc. of third int. Symp. on Artificial Life, and Robotics (AROBO III' 98)*, pages 478–481, 1998.
[7] Y. Matsumoto and A. Zelinsky. An algorithm for real-time stereo vision implementation of head pose and gaze direction measurement. In *Proc. of IEEE Int. Conf. on Face and Gesture Recognition (FG2000)*, pages 499–505, 2000.
[8] K. Nakadai, K. Hidai, H. Mizoguchi, H. G. Okuno, and H. Kitano. Real-time auditory and visual multiple-object tracking for humanoids. In *IJCAI2001*, pages 1425–1432, 2001.