Expressing Observation Direction through Face and Body Rotation in a Multi-user Conversation Setting

Satoshi Mieda, Shiro Ozawa, Munekazu Date, Hideaki Takada,
Yoshiaki Kurokawa, and Akira Kojima

NTT Media Intelligence Laboratories, NTT Corporation
{mieda.satoshi,ozawa.shiro,date.munekazu,takada.hideaki,
kurokawa.yoshiaki,kojima.akira}@lab.ntt.co.jp

Abstract. In this paper we clarified the range of observing direction by rotating the 2D human image and it is possible to express the observing direction by face direction. We conducted two subjective experiments about direction expression of the person on an image. In the first experiment, we compared two types of human image expression, rotated 2D human image of rotated 2D and direction correct. In the second experiment, we evaluated the effect of human image rotation and the criterion for judging the direction. We showed that the direction of the user's face is the main factor in expressing the observation direction. Results clearly showed that it is possible to express the observation direction, which is required for effective communication, by using only the rotation of human facial image.

Keywords: communication, remote, human expression.

1 Introduction

We are interested in enhancing communication between persons in locations remote from each other through the transmission of nonverbal information, such as the direction in which users are looking and the hand gestures they make. In order to show users and their objects of interest, it is very important to convey the direction in which they are looking when they are observing other users. With current video conference systems, it is difficult to accurately transfer the observation directions when two or more users are involved, so remote site users may feel that other users are looking at themselves.

2 Related Work

One approach to solving this problem is the use of multiple cameras to acquire and display multiple gaze directions [1]. Another is the use of a 3D display to show the directions [2] [3] [4]. However, these approaches require the use of many cameras or a special display device.

M. Kurosu (Ed.): Human-Computer Interaction, Part II, HCII 2014, LNCS 8511, pp. 273–280, 2014.

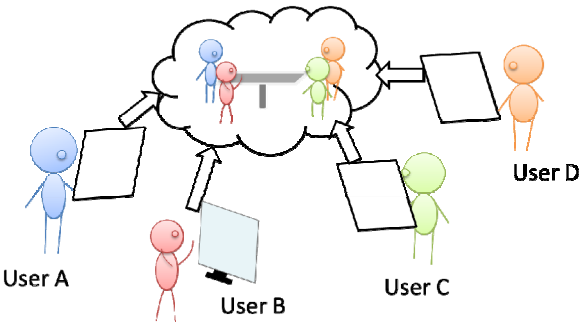

Fig. 1. Proposed communication system. Multiple users are arranged in the virtual space and can communicate with each other as if they were actually meeting together.

In another approach that has been proposed, remote participants are arranged in an actual space and the gaze direction is presented by giving a physical direction [5]. However, it requires special displays with a pan and tilt mechanism for remote participants. It is said that the normally 2D human image is perceived as always facing to the observer [6]. This is called the Mona Lisa effect.

The approach we propose involves a multiuser communication environment we are researching, which is arranged in a virtual space (Fig. 1). It comprises a desktop conference system in which simple equipment (a web camera and display) is used and only images showing a user's front view can be taken. In this paper, we report an experiment we conducted to validate our approach, in which a remote user's observation direction is expressed by using his or her front side facial image in a simple setting and rotating the image in the virtual space.

3 Evaluation of User's Observing Direction

Our goal is to build the environment where multiple people join and talk smoothly. So, we conduct a subjective experiment about the person's direction expression. We evaluated the effect the direction in which the remote user is thought to be observing and the effect of human image rotation, the criterion for judging the remote user's observing direction. We conduct two experiments. As first experiment, we confirm the effect of rotating human image. Secondly, we evaluate the effect of the difference the face and the body image.

3.1 Experimental Settings

Figure 2 shows the experimental setting. In this experiment we assume the environment where four users talk surrounding the round table. In order to display a remote person at actual size, we use 40 inch LCD (1,920 x 1,080 pixels, 880 x 500 mm). We set the distance between the display to the user 600mm, and the distance to the facing

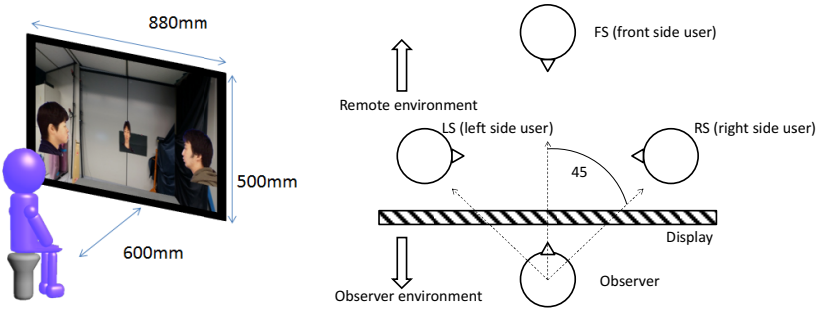

Fig. 2. Arrangement of users in experimental setting. In the left side, arrangement of observer and a display is shown. In the right side, The FS (front side user) is located in front of the observer, while the LS (left side user) and RS (right side user) are located at a 45 degree angle from the observer.

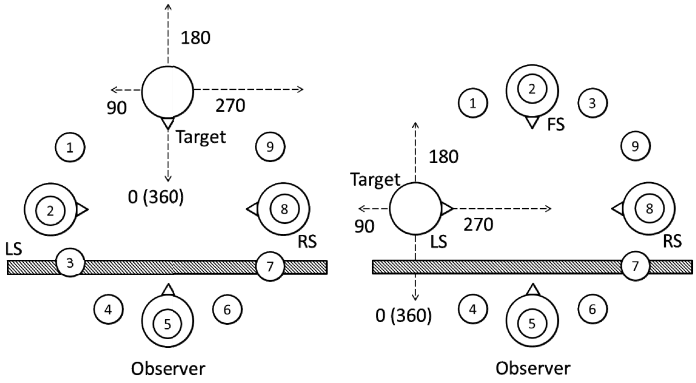

Fig. 3. The left side figure shows the answer direction when the target user is located in front of the observer. The right side figure shows the answer direction when the target user is located at left side of the observer.

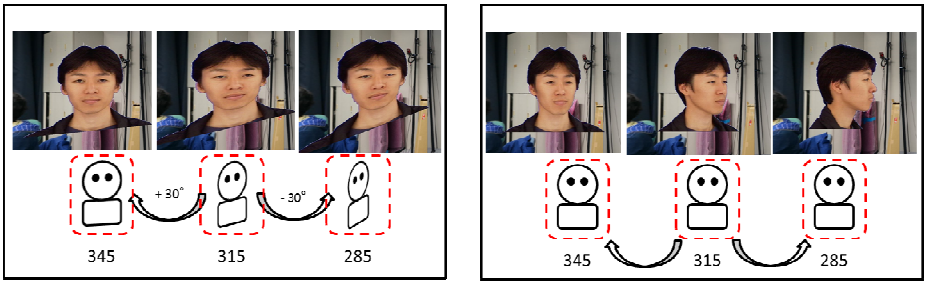

Fig. 4. Images of remote user located at a 45 degree angle to observer in a 3D virtual environment. Images in the left box are obtained by rotating human images; those in the right box are obtained with human image of directional correct.

partner who is displayed in the monitor is virtually 1,200 mm, which is possible distance to talk natural. The subject is seated in front of the display which is placed on 820mm high table. We construct 3D virtual space using OpenGL, and human images are arranged at the 3D position in the 3D virtual space. In the experiment we don't consider about displaying 3D such as binocular parallax or motion parallax because we are interested in plane and three-dimensional rotations of 2D images which can show high resolution and use easily.

3.2 Task and Experimental Design

We conducted a judging the person's observing direction task in a virtual environment. The experimental factors were displayed human image (rotating image, direction correct image), position of target user to answer (left side, front) and user directions. We simply show user's front image which is rotated depending on observing direction. We also show the direction correct user's image as observing direction to compare. The number of target user's positions was two because we considered about symmetry. The user directions were when the user was the front, two targets, and, when the user was the left side, three targets. We showed the human image which angle of images rotated at from -45 deg. to 45 deg. every 15 deg. to each target (7 patterns to each targets). Figure 4 shows two kinds of human image and target user's positions. Using all combinations, subjects did 70 trials three times for a total of 210 trials. We presented the each image randomly. Before starting experiment, the seat height was adjusted to the eye height of human image. In the task, a subject answered which direction a specified human image is observing by a number. Figure 3 shows the relation of the number and direction. Subjects included 4 men ranging from 25 to 30 years old, and the total time was about 20 minutes. Figure 5 shows the example of actual experiment.

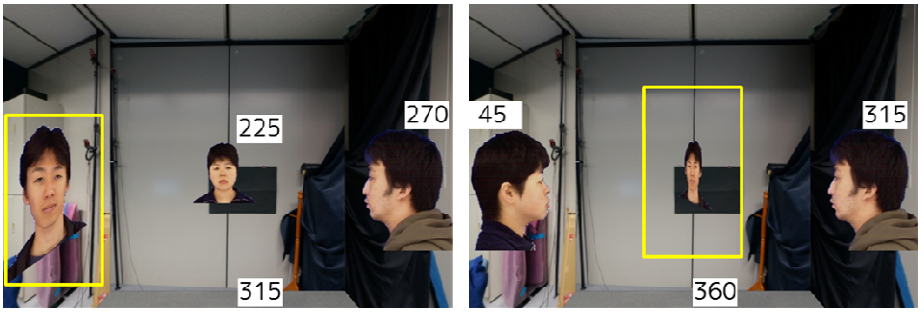

Fig. 5. Images used in actual experiment. The subject should answer the user observation direction surrounded yellow rectangle.

3.3 Result of Sotating Human Image

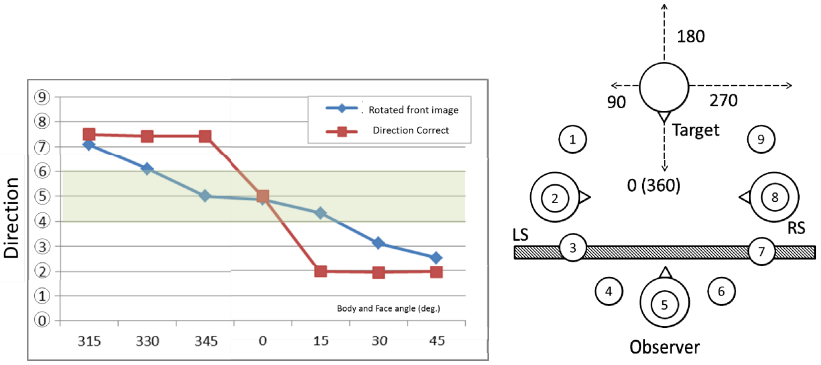

Fig. 6. The graph at left shows the perceived user direction when the target user is located in front of the observer and looking around the observer. The figure at right side shows the relation between answer number and observing direction.

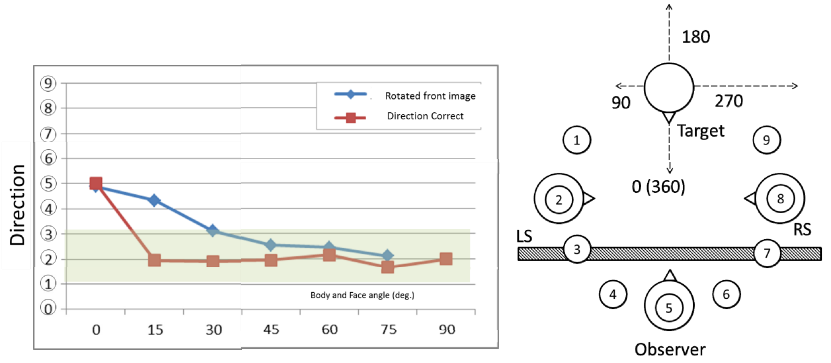

Fig. 7. The graph at left shows the perceived user direction when the target user is located in front of the observer and looking around the observer. The figure at right side shows the relation between answer number and observing direction.

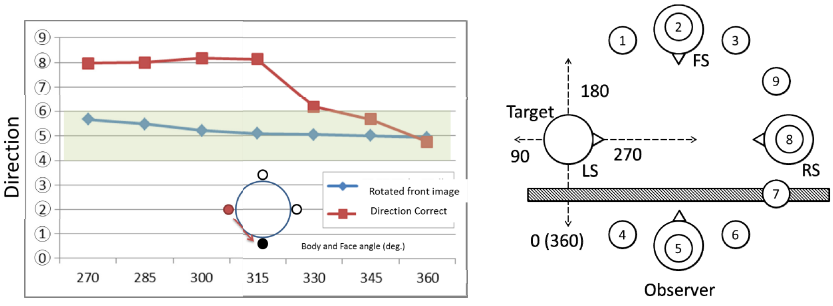

Fig. 8. The graph at left shows the perceived user direction when the target user is located at left side of the observer and looking around the observer. The figure at right side shows the relation between answer number and observing direction.

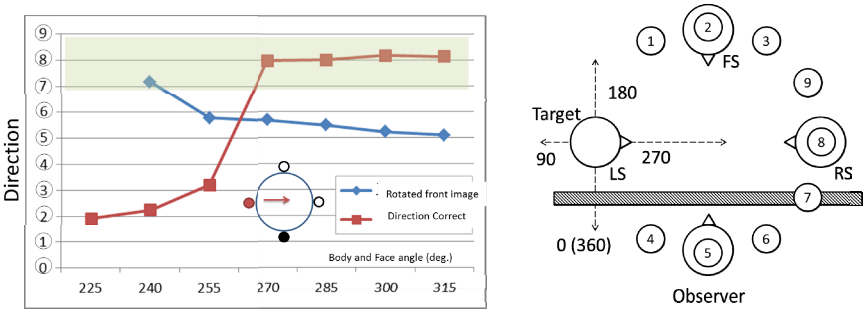

Fig. 9. The graph at right shows the perceived user direction when the target user is located at left side of the observer and looking around the observer. The figure at right side shows the relation between answer number and observing direction.

Figure 6 and 7 show the result of the answered direction when the target user is in front of subjects and figure 8 and 9 show the result of the answered direction when the target user is left side of subjects. Experiment 1 clarified that when a remote user is positioned such that he or she appears to be directly in front of the observer, the user's observation direction can be expressed only by rotating a 2D human image about z axis so that it is arranged to a 3D virtual space. Since it is difficult to acquire and carry out segmentation of the user's front view, we verify which element is most effective with respect to the observation direction of a person.

4 Evaluation of the Different Direction Face and Body

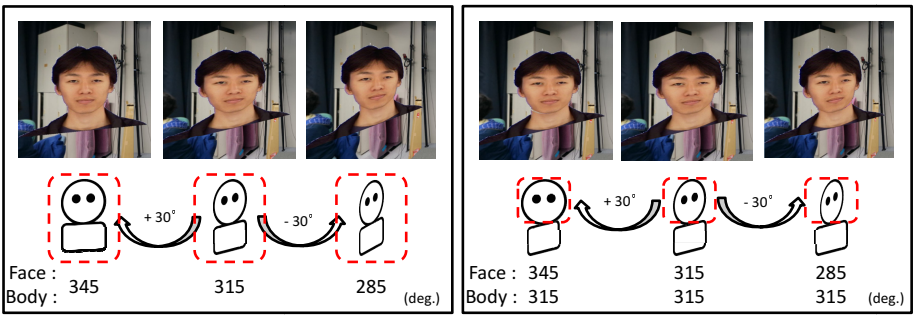

Fig. 10. Images of a remote user located at a 45 degree angle to the observer in a 3D virtual environment. Images in the left box were obtained by rotating the face and body together; those in the right box were obtained with our technique that rotates the face and body separately.

4.1 Experimental Setting

The arrangement of the users in the experimental setting is shown in Fig. 2. Four users join the environment, and one of them (the observer) observes the other three via a monitor. The observer is shown images of the three persons and reports the

observation direction he or she perceives for a specific person. The face and body are given different rotation angles in the image. The facial angle, body angle, and the position of the target user were given as experimental factors. Figure 10 shows an example of images used in the experiment.

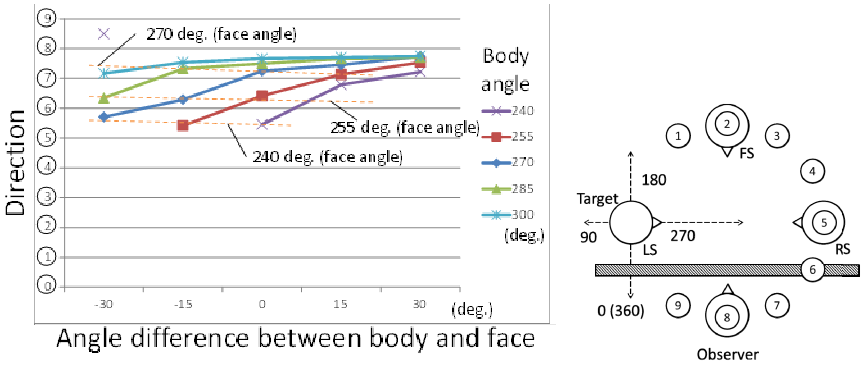

Fig. 11. The graph at left shows the perceived gaze direction relative to body angle with rotation angle difference given to the face and body. The dotted lines indicate the angle of the same face. The image at right shows the relations among the observation direction, the answer number, and the LS rotation angle.

4.2 Result of Express the Different Direction Face and Body

Figure 11 shows the experiment results obtained when the target person was the LS. It shows that the gaze direction of the LS is expressed to the observer, if the direction of the image is rotated to the observer side rather than to the direction of the RS, who is located in front of the LS, and if the direction of the front of the LS is widened, the direction other than that the observer will be shown. The face direction is dominant in expression of the observation direction; the body direction is largely irrelevant to it.

5 Discussion and Impact

We aim at achieving an environment in which four users communicate in virtual space using one camera and one display for each user (Fig. 1). We showed that the direction of the user's face is the main factor in expressing the observation direction, and that the observation direction can be shown without acquiring the image of a user's whole body.

Experiment results clearly showed that it is possible to express the observation direction, which is required for effective communication, by using only the rotation of human facial image, and to achieve virtual space communication among users coming together from remote locations in a communication system arranged in a virtual space where images of the users' front view can be taken.

References

1. Nguyen, D., Canny, J.: Multiview: Spatially faithful group video conferencing. In: Proc. CHI 2005 (2006)
2. Feldmann, I., Waizenegger, W., Atzpadin, N., Schreer, O.: Real-time depth estimation for Immersive 3D video vonferencing. In: Proc. 3DTV-CON 2010, pp. 1–4 (2010)
3. Jones, A., Lang, M., Fyfe, G., Yu, X., Busch, J., Mcdowall, I., Bolas, M., Debevec, P.: Achieving eye contact in a one-to-many 3D video teleconferencing system. ACM Transactions on Graphics 28(3), Article 64 (2009)
4. Iso, K., Date, M., Takada, H., Andoh, Y., Ozawa, S., Matsuura, N.: Video conference 3D display that fuses images to replicate gaze direction. In: Proc. IAS 2011, pp. 1–6 (2011)
5. Otsuka, K., Mucha, K.S., Kumano, S., Mikami, D., Matsuda, M., Yamato, J.: A system for reconstructing multiparty conversation field based on augmented head motion by dynamic projection. In: Proc. MM 2011, pp. 763–764 (2011)
6. Kendon, A.: Some functions of gaze direction in social interaction. Acta Psycologica 26, 22–63 (1967)