



# Proposal of Remote Face-to-Face Communication System with Line of Sight Matching Based on Pupil Detection

Kiyotaka Fukumoto<sup>(✉)</sup>, Yoshiyuki Yamamoto,  
and Yoshinobu Ebisawa

Graduate School of Integrated Science and Technology, Shizuoka University,  
Hamamatsu 432-8561, Japan  
{fukumoto.kiyotaka, ebisawa.yoshinobu}@shizuoka.ac.jp

**Abstract.** In general, remote face-to-face communication systems via the Internet, it is impossible to match the eyes of a user and a communication partner with each other because their face positions displayed in PC screens, at which they look, are different from the positions of the cameras for shooting their faces. To solve this problem, we proposed a remote head-free system which was based on the combination of a wide-angle color camera located behind a 45-deg inclined magic mirror, our pupil detection technique, image deformation techniques, and so on. The same two systems were assigned to the user and the partner. The narrow region image including a whole of the user's face image was cut out from the image obtained from the wide-angle color camera in the user's system, was displayed in the partner's system so that the midpoint of the detected right and left pupils of the user's face image virtually located at the position of the color camera in the partner's system, and vice versa. In addition, to deal with large back and forth or lateral head movements, the displayed face image of the user was adjusted in its size and distorted by using the projection transformation so that it always looked like the same size as the user's face and looked like the front face as seen from the partner, and vice versa. The questionnaire results obtained from experiment showed the tendency that the subjects' lines of sight matched each other, and the other questionnaire items about the system showed high evaluation.

**Keywords:** Remote communication · Pupil detection  
Projection transformation

## 1 Introduction

Recently, remote face-to-face communication systems via the Internet (e.g., Skype) has spread. In such a system, it is impossible to match users' lines of sight with each other because a position on a PC screen, at which the user looking, is different from the position of a camera for capturing the face image of the user. To solve this problem, a communication system with lines of sight matching by using a magic mirror was proposed [1]. By installing the camera behind the 45-deg inclined mirror, the system matched the position of the camera and the gaze position on the screen. However, the system did not allow large head movements of the users.

The present paper proposes a novel head-free remote communication system based on the combination of our pupil detection technique [2], a color wide-angle camera behind an 45-deg inclined magic mirror, which captures the user’s face image, the image deformation technique, and so on. In the system, the face image is cut out from the wide-angle color image, resized, and deformed according to the user’s pupil positions. Hence, even when a large back and forth or lateral head movement occurs, the face image of each other’s looks like the same size as the actual face and looks like the front face image seen from the other user.

## 2 Method

### 2.1 Overview of the Proposed System and Pupil Detection Method

Figure 1 shows the overview of the proposed system. The display was installed horizontally and its screen image was reflected by the 45-inclined magic mirror (transmission; approximately 10%, reflection; approximately 50%). A user looked the virtual screen of the display screen by the mirror. A wide-angle color camera for capturing the image including the user’s face was installed behind the magic mirror. Here, the optical axis of the camera was perpendicular to the virtual screen and passed through the center of the screen. By displaying the partner’s face image so that the midpoint of both pupils in the partner’s face image coincides with the camera position, it was possible to match each user’s line of sight.

In order to detect the pupils, three optical systems, each of which was composed of a black-and-white video camera, near-infrared LED light sources, near-infrared pass optical filter, and so on, were installed at the gap between the display and the magic mirror. The pupils were detected from the difference image created by subtracting a dark pupil image from a successive bright pupil image [2]. 3D pupil coordinates were obtained by stereo matching using detected pupil coordinates.

When the user wore eyeglasses, frames and lens of the glasses produced reflection images in the camera image (glass reflections). If the glass reflections overlapped the pupil image, the pupil detection and stereo matching were impossible. By using the

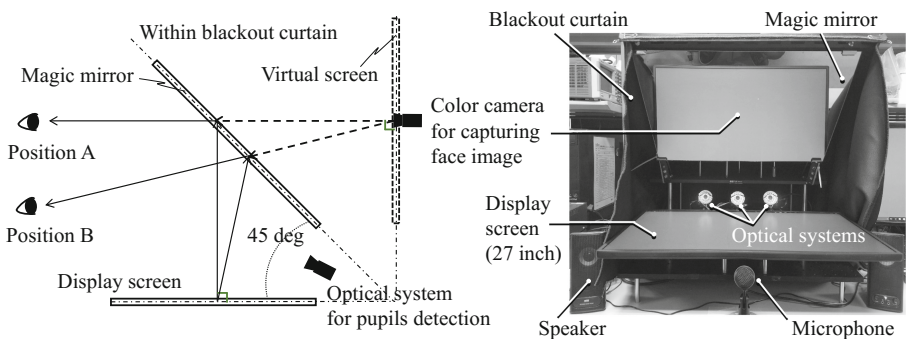
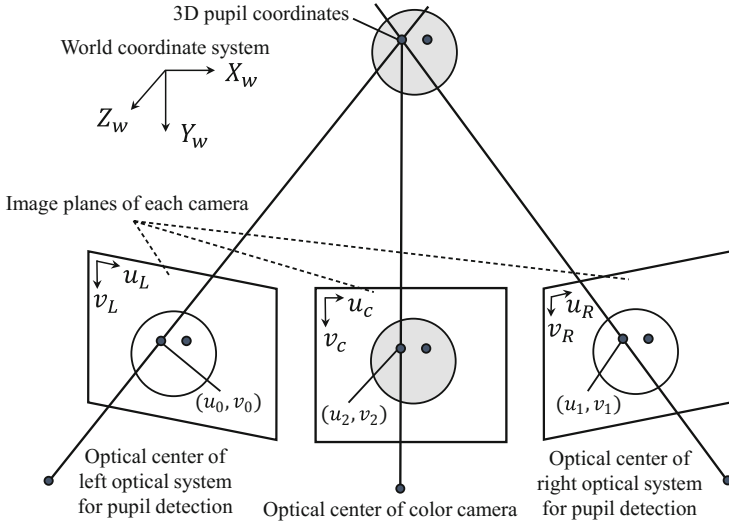


Fig. 1. Overview of the proposed head-free remote communication system.



**Fig. 2.** Estimation method of pupil coordinates in color camera image using our pupil detection technique.

three (more than two) optical systems, stereo matching was performed using the detected pupils from the two camera images even if the pupils were not able to be detected from other one camera image.

The three optical systems and the color camera were calibrated simultaneously by Zhang’s camera calibration method [3]. Therefore, pupil coordinates in the color camera image were able to estimate by projecting the 3D pupil coordinates to the color camera image (Fig. 2).

**2.2 Cutting Out Face Image from Color Camera Image**

As shown in Fig. 3, the color camera images, 3D pupil coordinates, pupil coordinates in color images, and voices obtained from microphones, were mutually transmitted and received between the two systems via TCP communication. A partner’s face region within the received color image was cut out into a rectangle of a certain range using the distance between both pupils based on the pupil coordinates in the color image. When the user moved the head forward and backward, the distance between both pupils also changes at the same ratio along with the change of the face size in the color camera image. Therefore, the width and height of the face image were determined based on the distance between both pupils.

**2.3 Deformation of Face Image Using Projection Transformation**

To deal with the user’s head movements, the partner’s face image was distorted by projection transformation so that the face image was observed as the front face image as seen from the user. In Fig. 4, the quadrangle  $A'B'C'D'$  on the display screen plane

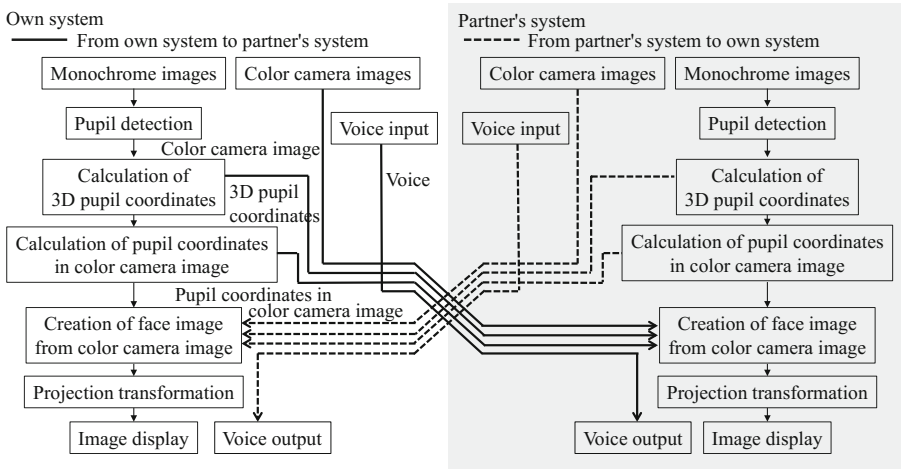


Fig. 3. Data transmission/reception between systems via TCP communication.

indicates the deformed face image-displayed area observed by the user, which was created based on the face image obtained from the color camera image of the partner. This image was distorted by projection transformation so that it can be seen as the original rectangular image without distortion (front face) as seen from the user position where was expressed by the midpoint  $M$  between both pupils. Specifically, the rectangle  $ABCD$  was defined as the front face image plane, which was perpendicular to the line  $O'M$  connecting the reference point  $O'$  (corresponding to the camera position) on the quadrangle  $A'B'C'D'$  and the point  $M$ . It was assumed that the pixels of the partner's face image were aligned in this image plane. In the front face image plane, the point  $O$  was the position of the midpoint between both pupils of the partner. The distance  $OM$  was a constant value. Here,  $xyz$  and  $XYZ$  coordinate systems with the origin  $O$  and  $O'$  were defined, respectively. In the  $XYZ$  coordinate system, the intersection  $P(X_i, Y_i, Z_i)$  between the image plane and the line connecting arbitrary pixel  $P'(X_i, Y_i, 0)$  and the point  $M$  was obtained while scanning within the display screen plane. The point  $P$  was transformed into a point  $p(x_i, y_i, 0)$  in the  $xyz$  coordinate system. When the point  $p$  existed within the  $ABCD$ , the pixel value of  $P'(X_i, Y_i, 0)$  in the display screen plane was calculated from the values of the pixels near  $p(x_i, y_i, 0)$  using the bilinear interpolation method. Thus, even when the user's head moved laterally from the front of the display screen or when the distance between the head and display screen plane changed, the user was able to look the partner's front face image with a certain size. Furthermore, the  $xyz$  coordinate system was rotated around the  $z$ -axis so that the intersections where the perpendicular lines were drawn from the left and right pupils to the image plane existed on the  $x$ -axis of the  $xyz$  coordinate system. Even if the user rolled the head, the partner's face image on the display screen rotates along with the user's head rotation, the retinal image hardly rotated.

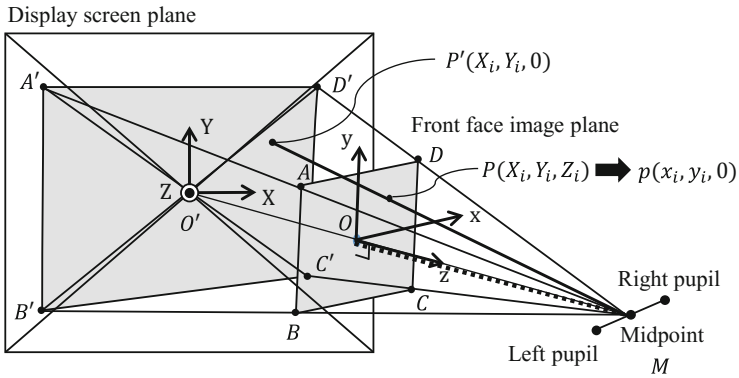


Fig. 4. Deformation of display image based on pupil position using projective transformation.



(a) Subject moved his head to left

(b) Subject moved his head upward

(c) Subject moved his head forward

(d) Subject rolled his head to left

Fig. 5. Each camera image and displayed face images without proposed methods.

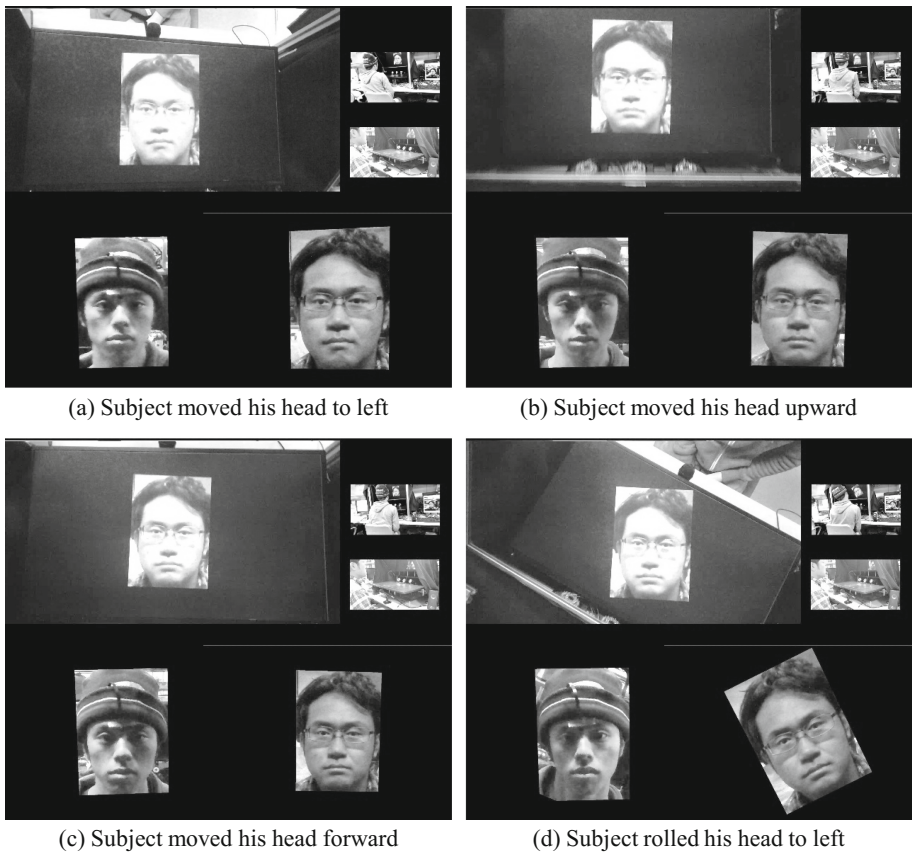
### 3 Experiments

#### 3.1 System Evaluation Using Camera Images

Subjects A and B were asked to be seated approximately 70 cm from the display screen and to move the head to left, right, upward, downward, forward, and backward and to roll to left and right. An eye camera for capturing the PC screen was attached to the middle of the eyebrows of the subject A.

Figure 5 shows the examples of the images when the original color camera images were displayed (without the proposed methods). The subject moved the head to left, upward, and forward and rotated to left, respectively. In each of (a)–(d), the upper left shows the image of the eye camera, the upper right is the image captured from behind each subject, and the lower two images show the face images presented to each subject. The eye camera images showed the partner’s face images which had the various size and slope corresponding to the head position and rotation of the subject.

Figure 6 shows the examples of the images with the proposed methods when the subject moved the head to left, upward, and forward and rotated to left, respectively. In



**Fig. 6.** Each camera image and displayed face images with proposed methods.

**Table 1.** Questionnaire concerning the proposed system.

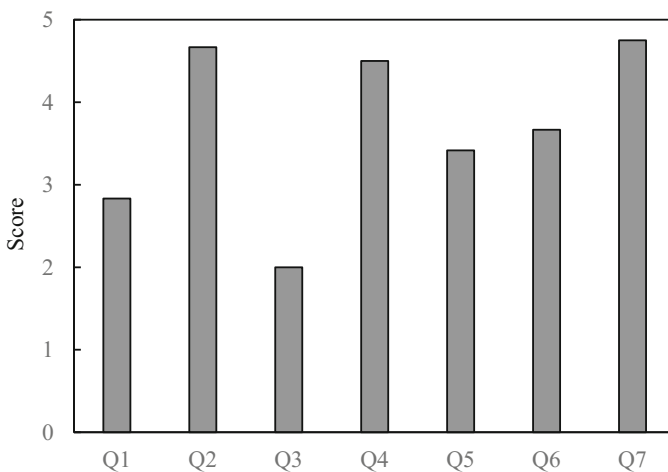
No.	Questions
1	Did you match the line of sight with partner? (without deformation by projection transformation)
2	Did you match the line of sight with partner? (with deformation by projective transformation)
3	Did you observe the front face image? (without deformation)
4	Did you observe the front face image? (with deformation)
5	Did you observe the bigger or smaller face compared to the actual face?
6	Did you feel the appropriate face size?
7	Did you feel that partner's face did not rotate when you rotated your head?

these conditions, the similar face images were captured from the eye camera. When the head was rotated, the presented face image also rotated, therefore, the subject A was able to observe the standing face image of the subject B.

### 3.2 Questionnaire About Proposed System

Twelve subjects were asked to use the system and to answer the seven questions about the system shown in Table 1. The subjects sat approximately 70 cm from the display and moved the head several times to left, right, upward, and downward within the range where both pupils were able to be detected (about  $\pm 15$  cm). All questions were evaluated in five grades. For questions except questions 5 and 6, "Yes" is set to 5, "No" to 1, and "Neither" to 3. In question 5, "Bigger" is set to 5, "Smaller" to 1, and "Neither" to 3. In question 6, "Big" is set to 5, "Small" to 1, and "Just right" to 3.

Figure 7 shows the results of the questionnaire. When the deformation of the face image by the projection transformation was performed, the high scores were obtained

**Fig. 7.** Results of questionnaire.

in the questions 2 and 4 about matching the line of sight and looking like the front face image compared to the condition without the deformation. Therefore, these results suggest that the proposed system improved realistic feeling compared to the conventional communication system.

## 4 Conclusions

In the present paper, we proposed the head-free face-to-face remote communication system with line-of-sight matching based on our pupil detection technique [2], a wide-angle color camera located behind a 45-deg inclined magic mirror, the image deformation techniques, and so on. In the experiment, we confirmed the realistic displayed partner's face images obtained from the eye camera images attached to the middle of the eyebrows. The results of the questionnaire suggested the improvements of usability and reality compared to the conventional system.

## References

1. Yaskawa information systems: NetSHAKER TalkEye. <https://www.ysknet.co.jp/>
2. Ebisawa, Y., Fukumoto, K.: Head-free remote eye-gaze detection system based on pupil-corneal reflection method with easy calibration using two stereo-calibrated video cameras. *IEEE Trans. Biomed. Eng.* **60**(10), 2952–2960 (2013)
3. Zhang, Z.: A flexible new technique for camera calibration. *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(11), 1330–1334 (2000)