www.s2is.org
International Journal on Smart Sensing and Intelligent Systems
ISSN 1178-5608

RANGE IMAGE SENSOR BASED EYE GAZE ESTIMATION BY USING THE RELATIONSHIP BETWEEN THE FACE AND EYE DIRECTIONS

Haruya Tamaki[1], Tsugunosuke Sakai[1], Yosuke Ota[1], Fusako Kusunoki[2],
Shigenori Inagaki[3], Ryohei Egusa[3, 4], Masanori Sugimoto[5], Hiroshi Mizoguchi[1]
[1]Department of Science and Technology, Mechanical Engineering,
Tokyo University of Science, 2641 Yamazaki, Noda-shi, Chiba, Japan
[2]Department of Information Design,
Tama Art University, 2-1723, Yarimizu, Hachioji, Tokyo, Japan
[3]Graduate School of Human Development and Environment,
Kobe University, Yayoigaoka 6, Sanda-shi, Hyogo, Japan
[4]JSPS research fellow, 5-3-1, Kojimachi,
Chiyoda-ku, Tokyo, 102-0083, Japan
[5]Graduate School of Information Science and Technology,
Hokkaido University, Kita 15, Nishi 8, Kita-ku, Sapporo, Hokkaido, Japan
Emails: 7515636@ed.tus.ac.jp, 7515624@ed.tus.ac.jp, 7516609@ed.tus.ac.jp,
kusunoki@tamabi.ac.jp, inagakis@kobe-u.ac.jp, 126d103d@stu.kobe-u.ac.jp, sugi@ist.hokudai.ac.jp,
hm@rs.noda.tus.ac.jp

Submitted: Oct. 15, 2016 ***Accepted: Nov. 2, 2016*** ***Published: Dec. 1, 2016***

Abstract- In this paper, we estimate the eye gaze by using the face direction to estimate the object of interest in an extensive area such as an educational display or museum exhibit. However, the direction of the gaze may differ from that of the face. Therefore, our approach was to use a method that utilizes both the face and eye directions to improve the accuracy of our estimation compared to only using the face direction. The first part of our study involved apprehending the relationship between the direction of the face and the eyes. The second part entailed estimating the eye gaze using this relationship. We then evaluated the effectiveness of this method in comparison to using the relationship with the face direction only.

Index terms*:* Range image sensor, Eye gaze direction, Face direction, Relationship between face and eyes, Active Appearance Models.

I. INTRODUCTION

The study we present in this paper involves determining the object of a person's interest at a particular time and location, for example in situations in which a person is viewing an educational interface or appreciating an exhibit in a museum. Knowledge of the object of interest would be useful to improve teaching materials and to set up additional objects of interest. Determination of these objects depends on the perceptiveness and experience of the teacher. However, the sheer number of educational interfaces and museum exhibits require quantification with a machine. Eye information is one of the indications that provide a measure of the objects in which people are interested [1]. Eye information accounts for 80[%] of the information humans receive through a sensory organ. Measuring the object of interest is known as eye gaze estimation. An estimation of the eye gaze of a learner and an estimation of the point of gaze would enable us to determine the objects at which they are gazing at a particular moment in time. Figure1 shows a diagram illustrating this system. As the first step toward designing this system we considered a method that would allow eye gaze estimation in an extensive area such as an educational interface or museum. One of the methods to calculate gaze direction involves fitting a device such as an eye camera and measuring the position of the eyeball [2], [3]. Most of eye camera estimate point of gaze by mapping the pupil center from image plane of camera and defined point of screen[4], [5]. However, this method has two limitations. The first is that the human would need to wear a device such as an eye camera. The second is that this method can only estimate the gaze direction at a distance not exceeding 500 [mm] from the camera. Owing to these limitations, researchers have proposed a method to estimate the gaze direction from the face direction, thereby making it possible to calculate the gaze direction without fitting a device [6]. However, the direction of the face may differ from that of the eye. Therefore, a method utilizing both the face and eye directions would provide a clearer estimation than when only using the face direction. The work included in this paper investigates the relationship between the face and eye directions and estimates the gaze direction by using this relationship.

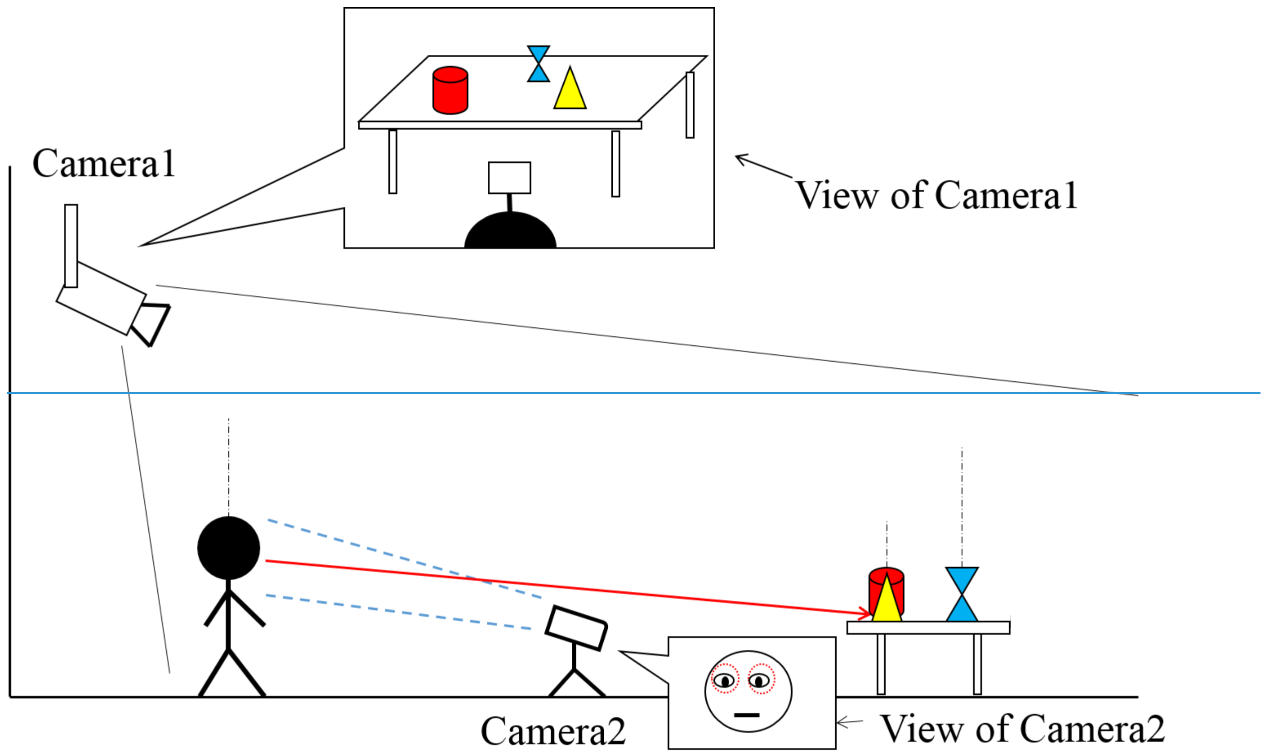

Figure 1. Illustration of the system

II. EYE GAZE ESTIMATION USING THE RELATIONSHIP BETWEEN THE FACE AND EYE DIRECTIONS

a. Conventional Eye Gaze Estimation

Usually, methods for eye gaze estimation mainly relied on using a device such as an eye camera [7] or using a contactless device to detect eyeball movement [8]. Although these methods can provide highly accurate eye gaze estimations, they have some demerits. First, the use of an eye camera causes mental stress. Second, sensors capable of detecting an eyeball are very expensive and can only be used with short-distance settings. These disadvantages prompted us to investigate whether the eye gaze could be estimated from the face direction as this approach would allow an extensive area to be included in the estimation range and it would be contactless.

b. Relationship between Face and Eye Directions

Estimation of the eye gaze relating to an educational interface or museum exhibit to determine the object of interest has two requirements. First, the sensor needs to be contactless. Second, the distance between the sensor and the person viewing the object needs to be greater than is presently the case. A method that uses the direction of the face to estimate the eye gaze would need to satisfy both of these two factors. If the face and eye direction were the same, we would be able to estimate the eye direction from the face direction. However, there are differences between the eye and face direction because the eyeball is able to move independently from the face [9]. Our approach avoids this difference by using the face direction to estimate the eye gaze by utilizing the relationship between the eye and face directions. Figure 2 illustrates this difference by way of an example.

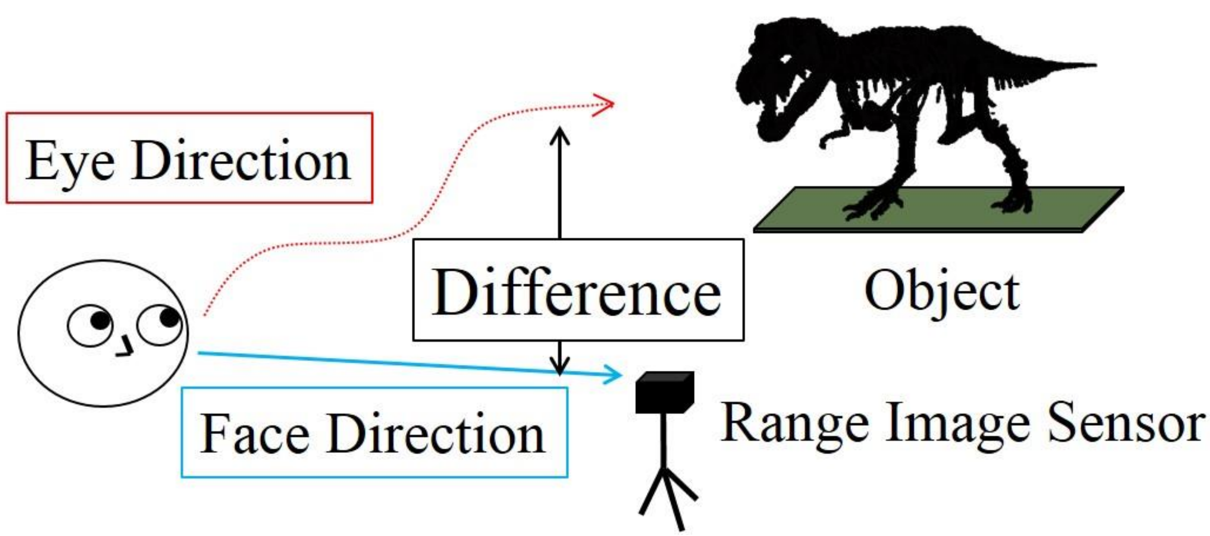

Figure 2. Example of difference between face and eye gaze directions.

c. Method for Measuring Face Direction

In this study, we use a Kinect sensor, which is a range image sensor, to measure the face direction. The Kinect sensor has previously been used to obtain head coordinates [10]. It measures the minutiae relating to the contours of the face, eye, eyebrows, nose, and mouth, a total of 100 three-dimensional coordinates. Figure 3 shows an example of extracted minutiae. We use Active Appearance Models (AAM) [11] [12] to measure 100 three-dimensional coordinates of minutiae. This method can be used to express a face model by determining the shape and brightness values of minutiae, which are learned by using Principal Component Analysis. Based on the results measured for the 100 three-dimensional coordinates we obtain the angle around the coordinate system, and the origin is located at the head coordinate. Figure 4 shows this coordinate system. The Kinect sensor has a different coordinate system. Figure 5 shows the relationship between these two coordinate systems. The face direction is expressed around the yaw, roll, and pitch axes.

Figure 3. Extracted minutiae

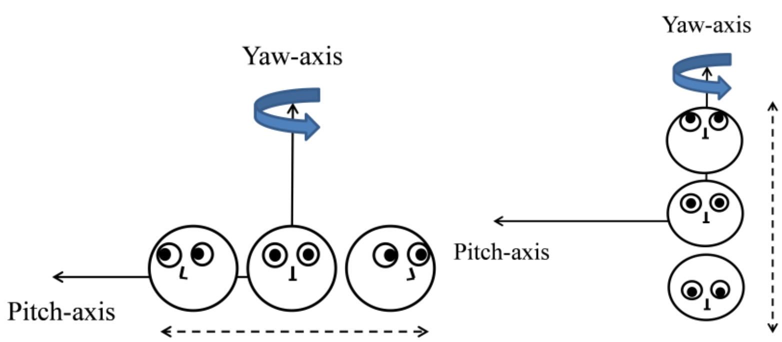

Figure 4. Head coordinates

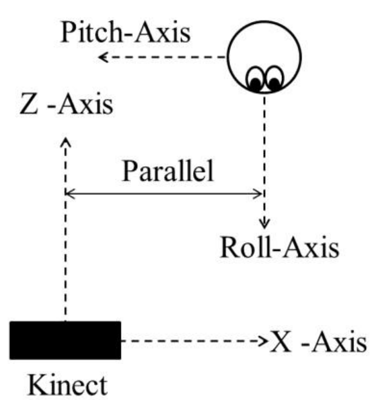

Figure 5. Relationship between the two coordinate systems

d. Gaze Direction Estimation using Relationship between Face and Eye Directions

In this study, the face and eye directions are divided laterally and lengthwise. These directions are expressed in terms of angles around the yaw (θ[deg]) and pitch (φ[deg]) axes. It is necessary to measure the angle of the face and the eye direction when the person gazes at the same objects to understand the relationship between the face and eye directions. We set up an object in both the lateral and lengthwise directions. In the case of the lateral direction, the object is set up at -30[deg] $\leqq \theta \leqq$ 30[deg] every 10[deg], with a total of seven points. In the case of the lengthwise direction, the object is set up at -20[deg] $\leqq \phi \leqq$ 20[deg] every 10[deg], with a total of five points. We measured the face angle using a Kinect sensor for the object that was viewed, and it was expressed by θ^M[deg], φ^M[deg]. The eye angle was fixed at the angle of the object when the human gazed at the object, and it was expressed by θ^T[deg], φ^T[deg]. Ten subjects were selected for this study. We calculated the average value of the face angle for each of the subjects. Figure 6 and 7 show the experimental environments for the lateral and lengthwise directions, respectively. We estimated the relational expressions using the calculated average value. Figure 8 and 9 show the relation between the face and eye angle of each direction. These relations are expressed by the following two numerical expressions ((1) and (2)). Based on these expressions we calculate the gaze direction from the face direction.

$$\theta^M = 0.52\theta+0.29 \qquad (1)$$

$$\varphi^M = 0.41\varphi+0.18 \qquad (2)$$

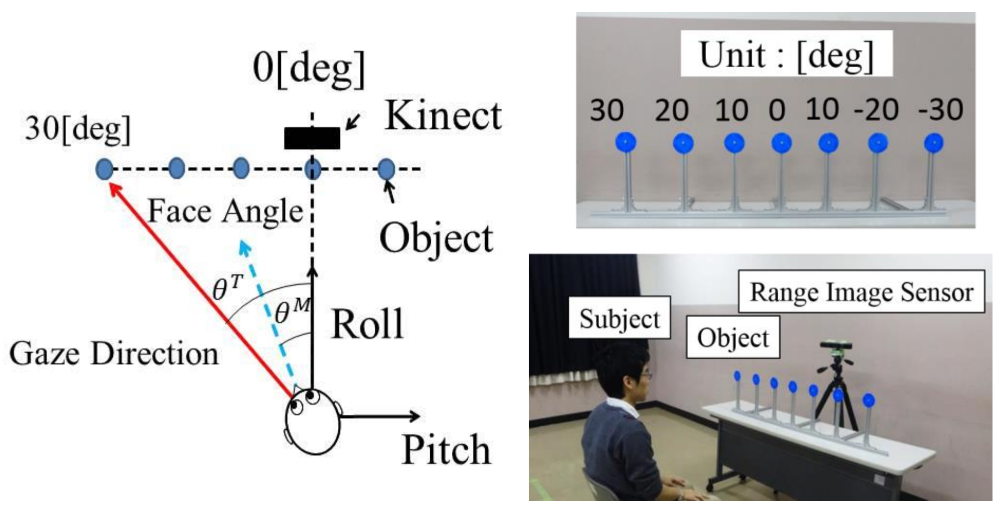

Figure 6. Lateral direction

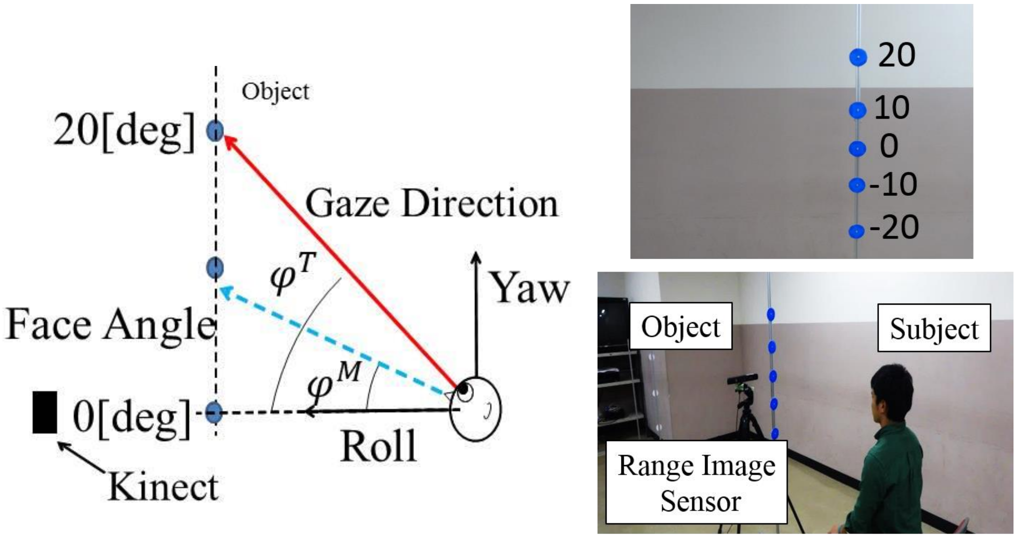

Figure 7. Lengthwise direction

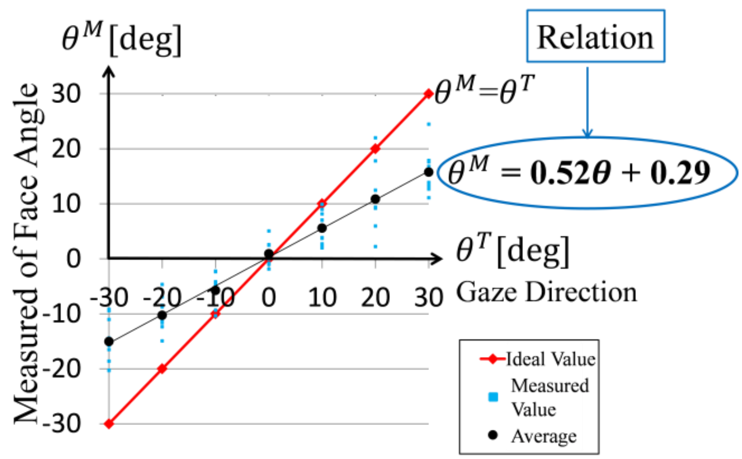

Figure 8. Lateral relation

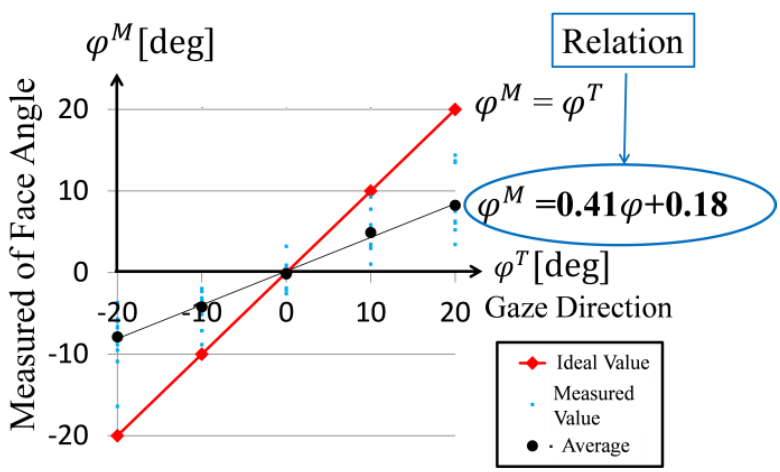

Figure 9. Lengthwise relation

III. CALCULATION OF EYE GAZE DIRECTION

a. Experimental Method

We calculate the eye gaze direction by utilizing the relational expressions between the face and eye directions for antecedent apprehension. Ten new subjects were used in this experiment. We set up a total of 35 objects in front of the subjects. Figure 10 shows the experimental environment. Subjects were asked to gaze at all the objects. We obtained two kinds of values, the first of which was the measurement of the face angle (θ^M[deg], φ^M[deg]) when gazing at the objects. The second value was obtained by calculating the eye gaze angle (θ^C[deg], φ^C[deg]) by utilizing the relational expressions between the face and eye directions. Figure 11 shows an example to illustrate the values that were used to calculate the value of the eye gaze angle. We evaluated the relational expression by comparing the face angle (θ^M[deg], φ^M[deg]) with the value calculated for the eye gaze angle (θ^C[deg], φ^C[deg]). For evaluation purposes, we specified a quota around the object to define the ideal value of the eye gaze angle. The range of the quota is within ±5[deg] from the object in the lateral and lengthwise directions. We specified this quota for all 35-unit objects. We then compared the number of values within the quota. Figure 12 shows the basis for this evaluation method. If the number of values calculated for the eye gaze angle exceeds those measured for the face angle, the relational expression is effective.

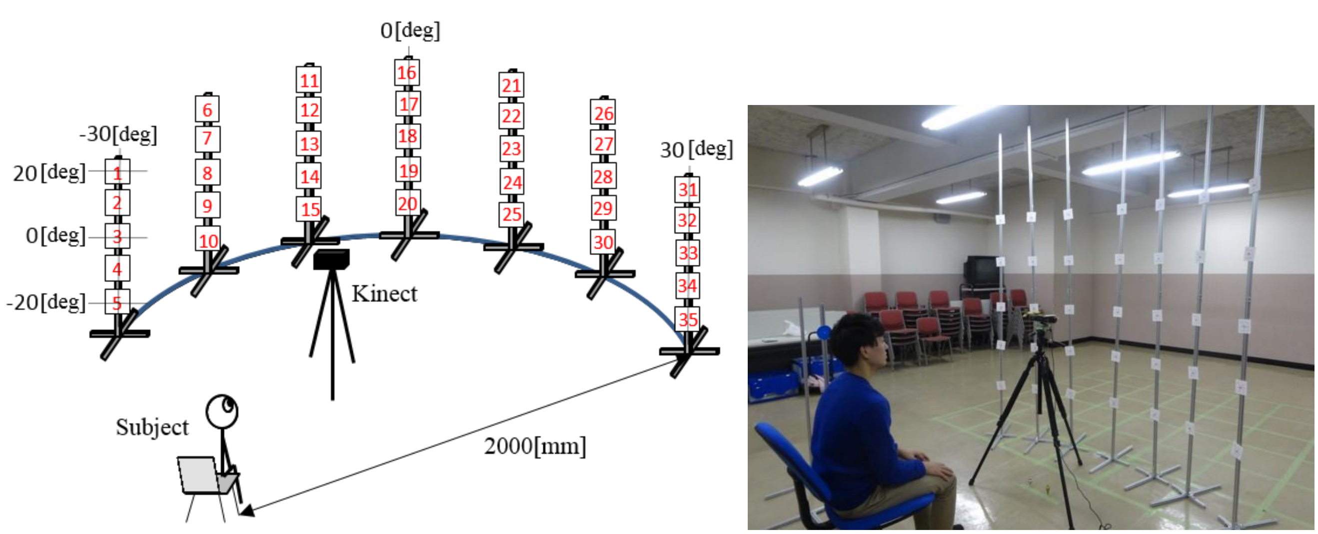

Figure 10. Experimental environment

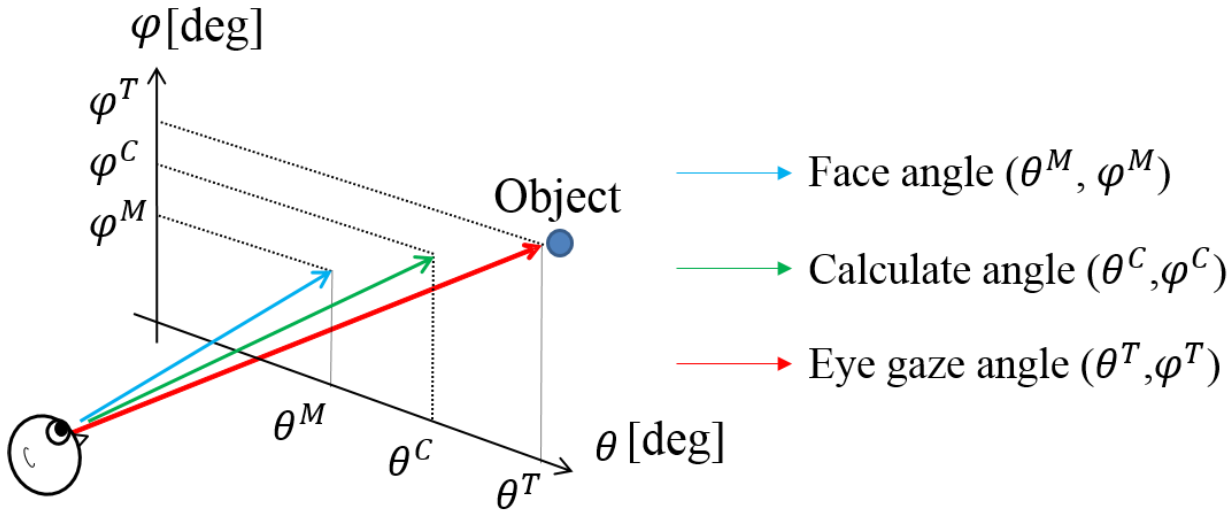

Figure 11. Example showing calculation of the value of the eye gaze angle

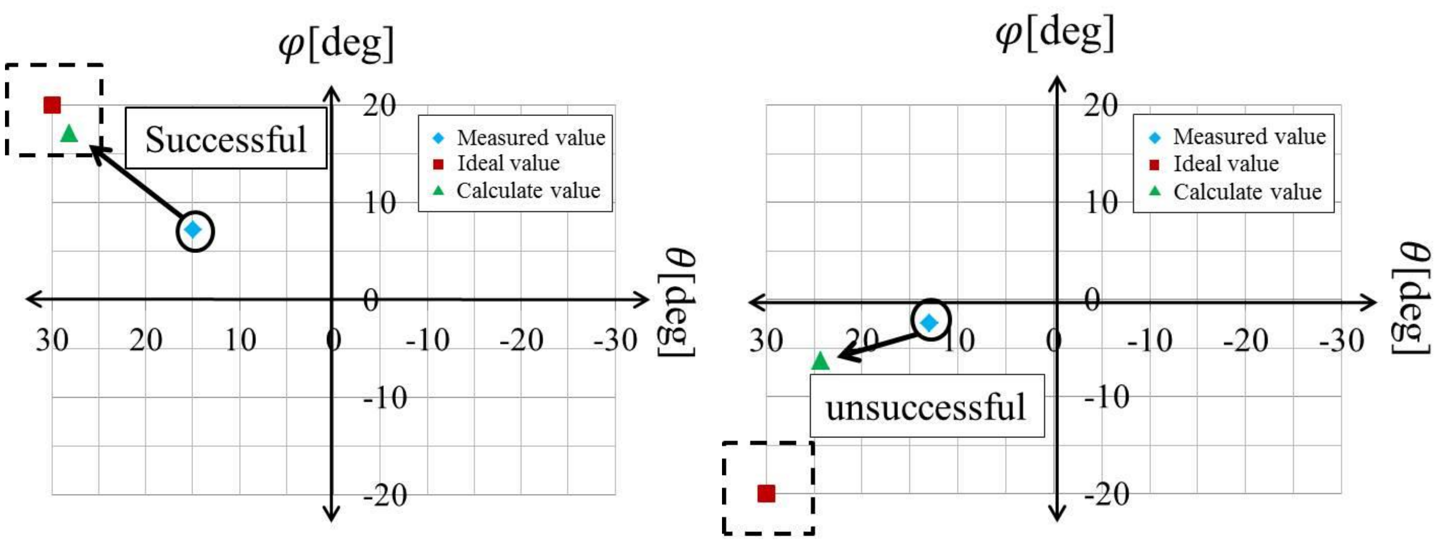

Figure 12. Basis for evaluation

b. Experimental Results

Figure 13 shows the result of the face angle measurements and the calculated value of the eye gaze angle of all the objects. We compare the number of values within the specified quota. Figure 14 shows the experimental results. When only the face angle is used, the number of successful units is 2. However, when utilizing the relationship, the number of successful units is 16. Thus, the experimental results indicate that eye gaze estimation utilizing the relationship between the face and eye direction is effective.

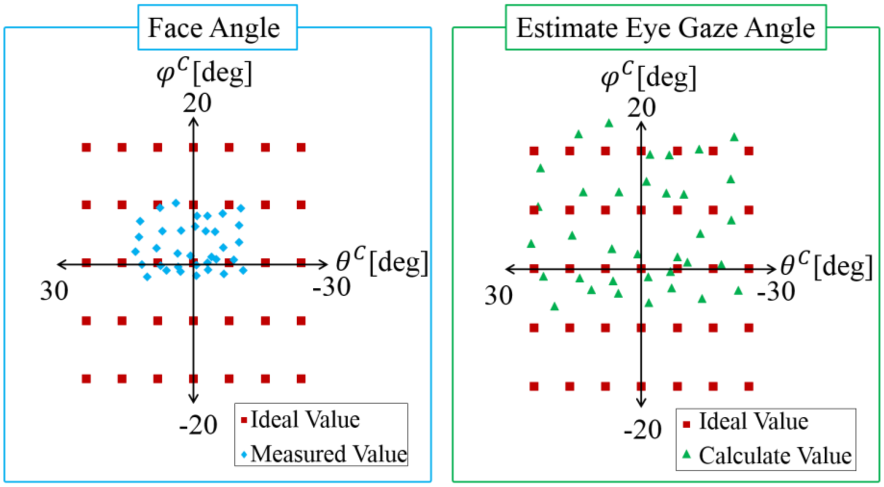

Figure 13. Results for all the objects

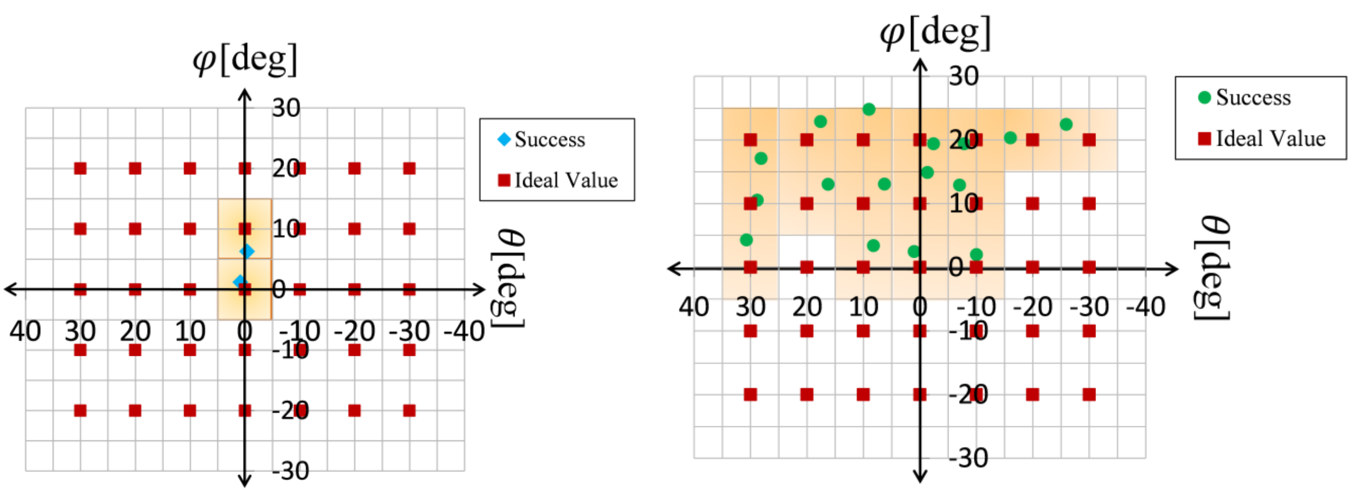

Figure 14. Experimental results

IV. CONCLUSION

In this paper, we described the relationship between the face and eye directions and the estimation of the gaze direction using this relationship. Our experimental results indicated that the use of this relationship produced a greater number of successful results compared to using only the face direction. We therefore confirmed that using the relationship between the face and eye direction to determine the gaze direction is effective.

ACKNOWLEDGEMENT

This work was supported by JSPS KAKENHI Grant Nos. 15H02936 and 26560129.

REFERENCES

[1] Brandherm Boris, Helmut Prendinger, and Mitsuru Ishizuka, "Interest estimation based on dynamic Bayesian networks for visual attentive presentation agents." Proceedings of the 9th International Conference on Multimodal Interfaces, ACM, pp. 346-349, 2007.

[2] Newman R, Matsumoto Y, Rougeaux S, and Zelinsky A, "Real-time stereo tracking for head pose and gaze estimation." Automatic Face and Gesture Recognition, Proceedings Fourth IEEE International Conference on, IEEE, pp. 122-128, 2000.

[3] Huang Chi-Wu, and Chun-Wei Hu, "An wearable eye tracker for coordinates of POG estimation using 3D model." Consumer Electronics-Taiwan (ICCE-TW), IEEE International Conference on, pp. 1-2, 2016.

[4] Shih-Chen Tseng, Zong-Sian Jiang, Chun-Wei Hu, "Projective Mapping Compensation for the Head Movement during Eye Tracking", Consumer Electronics — Taiwan (ICCE-TW), IEEE International Conference on, pp. 131-132, 2014.

[5] D. W. Hansen, Q. Ji, "In the Eye of the Beholder: A Survey of Models for Eyes and Gaze", IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 32, no. 3, pp. 478-500, 2010.

[6] Doshi Anup and Mohan M. Trivedi, "Attention estimation by simultaneous observation of viewer and view." Computer Vision and Pattern Recognition Workshops (CVPRW), IEEE Computer Society Conference on, pp. 21-27, 2010.

[7] Tursky, B. "Recording of human eye movements." Bioelectric Recording Techniques Part C, Academic Press, 1974.

[8] Hess, B.J.M. "Dual-search coil for measuring 3-dimensional eye movements in experimental animals." Vision Reseach 30.4 , pp. 597-602, 1990.

[9] Poole, Alex, and Linden J. Ball. "Eye tracking in HCI and usability research." Encyclopedia of human computer interaction 1, pp. 211-219, 2006.

[10] Zhang Zhengyou, "Microsoft Kinect Sensor and its Effect," MultiMedia, IEEE, Vol.19, No.2, pp.4-10, 2012.

[11] T.F.Cootes, G.J Edwards, and C.J Taylor "View-based Active Appearance Model" Image and Vision Computing 20, pp.657-664 , 2002.

[12] Stegmann, Mikkel Bille. "Active appearance models." Diss. PhD thesis, Technical University of Denmark, 125, 129, 2007.