2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)

TEllipsoid: Ellipsoidal Display for Videoconference System Transmitting Accurate Gaze Direction

Taro Ichii [*]
Tokyo Institute of Technology

Hironori Mitake[†]
Tokyo Institute of Technology

Shoichi Hasegawa[‡]
Tokyo Institute of Technology

ABSTRACT

We propose ”TEllipsoid”, an ellipsoidal display for video conference systems that can provide not only accurate eye-gaze transmission but also practicality in conferences, namely the convenience to use and the preservation of the identity of the displayed face.

The display comprises an ellipsoidal screen, a small projector, and a convex mirror, where the bottom-installed projector projects the facial image of a remote participant onto the screen via the convex mirror. The facial image is made from photos shot from 360 degrees around the participant. Moreover, the image is modified to improve identity. The gaze representation is implemented by projecting the 3D model of eyeballs onto a virtual ellipsoidal screen.

We evaluated the gaze transmissibility of the display in conference situations. As a result of experiments, we concluded that accurate gaze transmission is available in conferences when the angular distance of the adjacent participants is more than 38.5 degrees.

Index Terms: Human-centered computing—Human computer interaction (HCI)—Interaction devices—Displays and imagers; Hardware—Communication hardware, interfaces and storage—Displays and imagers;

1 INTRODUCTION

In spite of the development of many kinds of information and communication technology (e.g., telephone, e-mail, video conference, and SNS), face-to-face conversation is the most frequent mode of communication. The technologies that make remote communication as realistic as face-to-face communication have huge advantages such as reducing labor/travel costs required for face-to-face decision-making.

However, the planar display, which is commonly used in current video conference systems, cannot accurately transmit the nonverbal information (e.g., eye gaze or head movement), which plays an important role in communication. Thus, smooth communication is not possible. It is also reported that turn-taking in multi-party conversations is adjusted by eye gaze or gesture [3,10]. The problem of accurate gaze transmission is because of the Mona Lisa Effect [15], which makes it difficult to understand who is spoken to by the remote participant.

Many studies have tried to eliminate this effect by using a volumetric display, such as a multiview 3D display [9] or face-shaped display [12]. However, the devices require actuated mechanisms that cause mechanical noises and maintenance issues. As a more reasonable method, it is reported that a curved display can transmit eye gaze more accurately than a planar display [14].

Furthermore, we believe that a practical videoconference system must not only transmit gaze direction accurately. It must be convenient to use and a displayed face of the system must be similar to the real face of the remote participant in the conference.

Figure 1: Intended usage.

For practical use, high versatility is needed because any number of people might participate in a conference as remote participants. The need for individual equipment, such as the tailor-made facial screen which is used on face-shaped display [12], makes it inconvenient for users. Moreover, lower manufacturing costs and higher reliability also are needed for practical application. It is obvious that the fewer the driving mechanisms in a system, the lower the risk of failure and manufacturing costs. However, for table meetings, such mechanisms are normally needed because the face will turn by 90 degrees left and right to look at the participants. The mechanisms are not necessary on the vertical-axisymmetric surface of revolution shape display, which can express horizontal head rotation only with changes of image.

The similarity of the displayed face to the real face also is important for practical video conference system because, in today’s society, communication is incomplete unless it is face to face. In other words, the interlocutor of the video conference talks face-to-face with the partner at least once. In that case, if the displayed facial image is not like the real face of the interlocutor, it can be difficult for the user to recognize the image as the interlocutor [8]. It has been reported in human cognition studies, that we identify interlocutors by the overall arrangement of the parts and contour of the face [1]. Hence, a display that can show the arrangement of the parts of the face naturally from any angle is necessary for practical video conference systems.

Therefore we propose ”TEllipsoid”, an ellipsoidal display for video conference systems. It has a curved, vertical-axisymmetric surface and the affinity with a human face, to provide not only accurate eye-gaze transmission but also practicality in conferences, namely the convenience to use and the preservation of the identity of displayed face. This display is designed for the small-scale video conference in which the participants are 3 to 15 people as shown in Fig. 1.

In this paper, we describe the design of TEllipsoid and the evaluation of eye-gaze transmissibility. As a result of our experiments, we concluded that accurate gaze transmission is available in conferences when the angular distance of the adjacent participants is more than

---

[*] e-mail: ichitaro@haselab.net
[†] e-mail: mitake@haselab.net
[‡] e-mail: hase@haselab.net

2642-5254/20/$31.00 ©2020 IEEE
DOI 10.1109/VR46266.2020.000-3

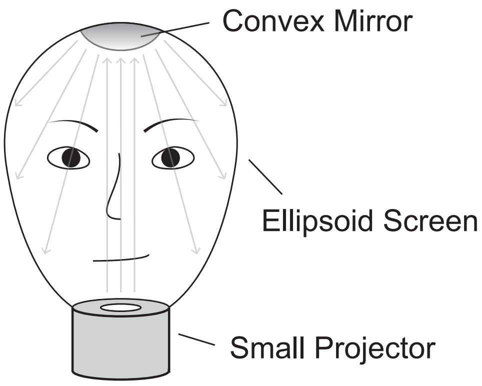


Figure 2: System overview of TEllipsoid.

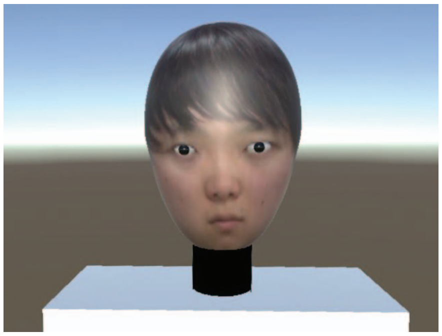


Figure 3: The prototype made in the VR environment.

38.5 degrees.

2 PROPOSED METHOD

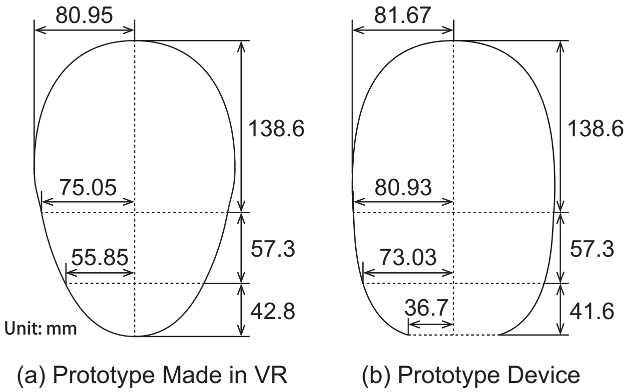
We proposed the ellipsoidal display for video conferences, which we named "TEllipsoid" (Tele + Ellipsoid). TEllipsoid comprises an ellipsoidal screen, a small projector, and a convex mirror (Fig. 2). The bottom-installed small projector projects the facial image onto the screen via the convex mirror. The displayed facial image is not a real-time but a pre-generated static facial image, with parametrizable rendered eyes for gaze representation. Moreover, the shape of the screen is designed based on the average face dimensions of young Japanese males [11].

2.1 Prototype made in VR

First, we made a prototype in the Virtual Reality environment (Fig. 3) to consider the screen shape, the projection image, and the method of gaze representation before making the real device. The VR environment was made with Unity (2019.1.2f1) and built on Oculus Quest, the head-mounted display made by Oculus.

The screen shape was designed from the average dimensions of the faces of young Japanese males [11]. We generated a 3D model of the ellipsoid of revolution by rotating a simplified face shape (Fig. 4(a)) along the vertical axis.

The facial image projected onto the screen (Fig. 6(a)) was generated as the panorama composite photo from the 425 frames of the movie shot 360 degrees around the person, by concatenating cropped facial images horizontally (Fig. 5). The eyes in the image also were cropped for the gaze representation, which is described later. Next, the facial image was inverse cylindrical-projected on the virtual screen as a texture, which preserves vertical arrangement.



(a) Prototype Made in VR (b) Prototype Device

Figure 4: Side view of the ellipsoidal screen.

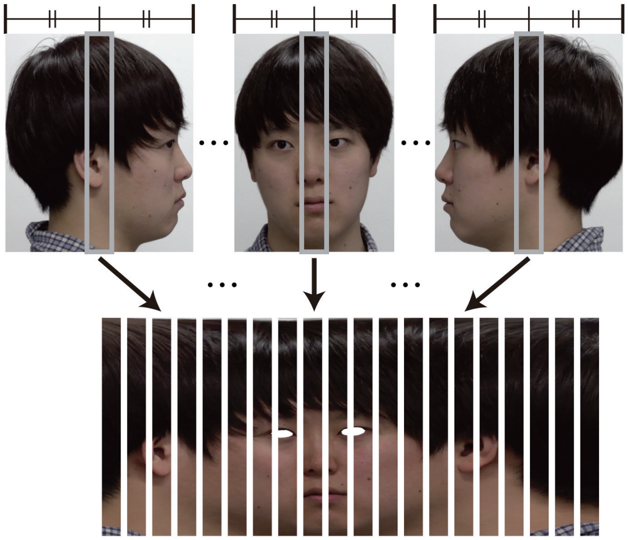


Figure 5: Method to generate facial image.

2.2 Gaze representation

We implemented the gaze representation because a participant gazes someone not only by turning the head but also by rotating the eyes in an actual conference. The method of gaze representation is shown in Fig. 7.

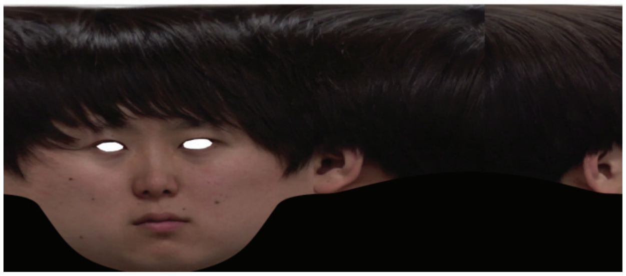
The eyeball models are placed on the eye section of the virtual screen. The vertical/depth position and size are adjusted to be consistent with the real face. The lateral position and the pivot are aligned to the center of the orbit.

The eyes gaze at the target position by turning the pupillary axis toward the target. Subsequently, the eyeballs are projected onto the screen surface using parallel projection along the pupillary axis. This projection method is adopted because the geometry looks consistent with the eyeball models from the target point, which theoretically enables the perception of eye contact from the target position. This projection method is also similar to previous studies on spherical display [2, 7].

2.3 Prototype device

We made a prototype of the system as shown in Fig. 8. The screen with a thickness of 0.4 mm was fabricated with a 3D printer using gray PETG filament. We chose a portable laser projector, Laser Beam Pro C200 made by Cremotech, because projection distance

(a) Prototype Made in VR



(b) Prototype Device

Figure 6: Facial images projected on the screen.

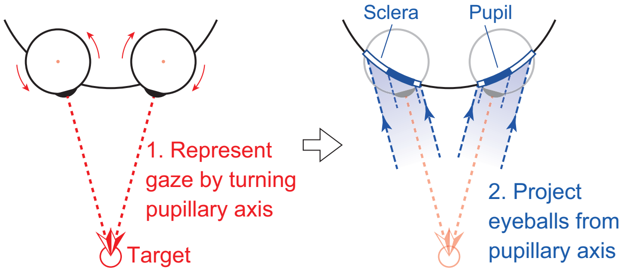


Figure 7: The method of gaze representation.

changes depending on position. The convex mirror is a commercial convex spherical mirror made of plated plastic with a diameter of 110 mm and a curvature radius of 102.5 mm.

We modified the shape of the screen based on the feedback from the evaluation of the VR prototype. We enlarged the lower part to improve the geometries of the nose and mouth and the naturalness of the profile(Fig. 4(b)).

The facial image also was modified. Part of the neck was filled with black, which is equivalent to no projection, to express the contour of the jaw(Fig. 6(b)). This improvement makes the head contour more natural and has the potential to evoke the Wollaston Illusion [18], in which the perceived direction of the gaze is influenced by the orientation of the face. The shape of the black fill was carefully adjusted to make the frontal face appear slender and the contour of the jaw appear natural. Such adjustments can be made to other facial images. The image also was trimmed to adjust the geometry of the parts of the face.

The transformed facial image for the projection was generated by the optics setup virtually assembled in Unity. The projector in the device was replaced by a virtual camera in which the field of view is consistent with the throw ratio of the projector. As a result, we could acquire a consistent stereoscopic image by projecting the image captured by the virtual camera on the setup.

Figure 8: Overall view of the prototype.

3 EXPERIMENT

To evaluate the performance of gaze transmission, we conducted four experiments.

- Experiment 1 was performed to investigate whether we can perceive eye contact only from a specific direction and acquire the relationship between gaze angle and observation angle that gives the perception of eye contact.
- Experiment 2.1 was performed to acquire the precise gaze angle where participants strongly feel eye contact for each observation angle.
- Experiment 2.2 was performed to investigate the range of observation angle where participants still feel eye contact, for each gaze angle acquired in Experiment 2.1.
- Experiment 3 was performed to investigate whether participants can recognize the displayed image as the remote person and how they feel about the image.

3.1 Experiment 1

This experiment was performed in the VR environment before making the hardware device to develop the system more efficiently.

The definitions of gaze angle θ and observation angle ϕ are shown in Fig. 9. The gaze angle is defined as the angle from the front to the look-at point around the display. The look-at point was placed 1 m from the VR prototype. The observation angle is defined as the angle from the front to the observation point around the display.

The participants sat and observed the display from 1 m away in a VR environment using Oculus Quest. The height of the participant's viewpoint was adjusted to match the level of the eye on the display. The participants were 11 people aged between 22 and 68.

The gaze angle was allocated in increments of 8 degrees from -40 to +40 degrees, where the pupil can be displayed naturally. The observation angle was allocated from -90 to +90 degrees in increments of 10 degrees because the display can turn its face only to 90 degrees left or right in the case of table meetings. By combining gaze angle $\theta \in \{-40^\circ, -32^\circ, -24^\circ, -16^\circ, -8^\circ, 0^\circ, 8^\circ, 16^\circ, 24^\circ, 32^\circ, 40^\circ\}$ and observation angle $\phi \in \{-90^\circ, -60^\circ, -30^\circ, 0^\circ, 30^\circ, 60^\circ, 90^\circ\}$, we generated 77 stimuli for the experiment.

The participants were asked to indicate if the display gazes at them by pressing the button of the controller (Oculus Touch) for the 77 stimuli presented randomly. To acquire a more intuitive response, each stimulus was switched to the next one every 5 seconds.

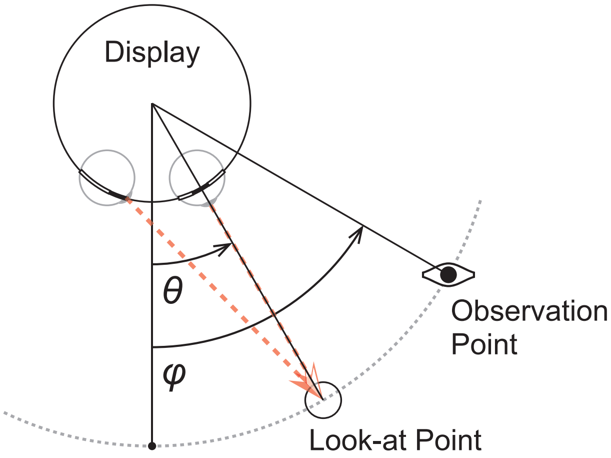
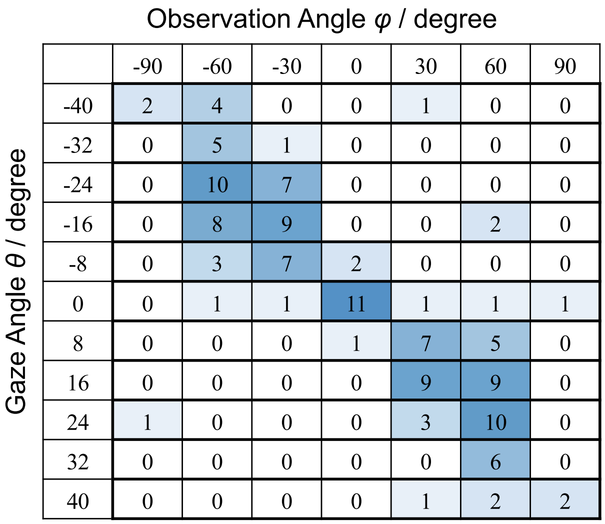


Figure 9: The gaze angle and observation angle.



	-90	-60	-30	0	30	60	90
-40	2	4	0	0	1	0	0
-32	0	5	1	0	0	0	0
-24	0	10	7	0	0	0	0
-16	0	8	9	0	0	2	0
-8	0	3	7	2	0	0	0
0	0	1	1	11	1	1	1
8	0	0	0	1	7	5	0
16	0	0	0	0	9	9	0
24	1	0	0	0	3	10	0
32	0	0	0	0	0	6	0
40	0	0	0	0	1	2	2

Figure 10: Number of participants who perceived eye contact for each stimulus.

3.2 Experiment 2

3.2.1 Experiment 2.1

This experiment was performed in a dark room to maximize the visibility of the prototype device. Because of the light emission from TEllipsoid and incident light from other rooms, objects in the environment, such as a table, chair, and the wall were visible. The definitions of the gaze angle and the observation angle were the same as in Experiment 1. The participants sat on height-adjustable chair and observed the device, where the height of the participant's viewpoint was adjusted to match the level of the eye on the device. The distance between the participant and the device was fixed at 1.5 m, which is in the range of social distance for the small-scale conference reported in the proxemics study [6]. The participants were 11 people aged between 22 and 38.

The observation angle was allocated from -90 to +90 degrees in increments of 30 degrees. This increment corresponds to 79 cm at a distance of 1.5 m, which is close to the smallest distance of adjacent participants. The change of the observation angle was substituted with equivalent head rotation. The participants were instructed to adjust the gaze angle to where they felt eye contact the strongest for each observation angle, which were presented randomly. The adjustment was performed by pressing buttons to change the gaze angle by fixed angles of 2 degrees and 0.5 degrees. Each time the participants pressed the button, the projection blinked for 0.5 seconds to reduce the effects of adaptation.

3.2.2 Experiment 2.2

This experiment was conducted just after Experiment 2.1 on the same participant using the result of the experiment. In this experiment, the participants were instructed to adjust the observation angle to the point where they no longer felt eye contact by shifting the observation angle from a given angle.

The adjustments were performed by the participants in the same manner as in Experiment 2.1. The participants were instructed to change the head rotation, which is equivalent to changing the observation angle, in one direction from the strongest-gazed representation they had in the earlier experiment to the direction in which they no longer felt eye contact. The adjustments were made to the left and to the right for every observation angle, which were presented randomly.

3.3 Experiment 3

In this experiment, we asked the participants to describe how they felt about the displayed facial image and whether they could recognize the image as the model person. The participants were instructed to rate the displayed image with the Visual Analogue Scale (VAS), in which the participants indicate the position corresponding to their impression along a line between a pair of words. All the participants had met the model person at least once.

We asked the participants, "How do you feel about the face?" using these word pairs: Beautiful–Ugly, Normal–Creepy, Friendly–Unfriendly, and Humanlike–Robot-like. We also asked them, "Can you recognize that person?" using the scale, Easy to recognize–Hard to recognize.

The observation conditions were similar to those in Experiment 2. The participants were allowed to change the gaze angle and the head rotation by pressing the buttons. Their changes were reflected in the image immediately. The participants were 11 people aged between 22 and 28.

4 RESULTS

4.1 Experiment 1

Of the 11 participants, the number who perceived eye contact for each stimulus is shown in Fig. 10. For both left and right, there was a trend: the bigger the gaze angle, the bigger the observation angle in which eye contact was likely to occur. However, the gaze angle and the observation angle were not always consistent. Especially when the gaze angle was ±8, ±16, ±24 degrees, there were multiple observation angles in which the number of participants who perceived eye contact was relatively high. Participants also commented that the displayed face looked shy or suspicious.

4.2 Experiment 2

4.2.1 Experiment 2.1

The gaze angle for which the strongest eye gaze was perceived for each observation angle is shown in Fig. 11. There was a trend like that in the VR prototype: the bigger the gaze angle, the bigger the observation angle in which eye contact was likely to occur. Also, as in Experiment 1, the gaze angle and the observation angle were not always consistent.

The standard deviation was relatively large, up to 7.25 degrees when the observation angle was +60 degrees. When the observation angle was 0 degree, the standard deviation reached a minimum value

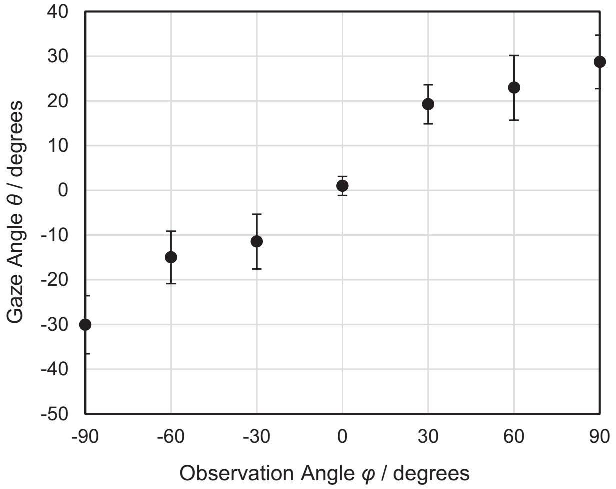


Figure 11: The gaze angle in which the strongest eye gaze was perceived for each observation angle.

of 2.12 degrees, which is still much larger than that for recognizing real faces [4].

4.2.2 Experiment 2.2

The range of observation angles at which the participants perceived eye contact with the strongest-gazed representation adjusted in Experiment 2.1 is shown in Fig. 12. The result of the experiment is also shown in Table 1.

The highest angular range width from the given observation angle was 38.5 degrees, which corresponds to 99 cm at a distance of 1.5 m. Note that the left (right) edges of -90(+90) degrees were ignored because there is no participant behind the device in the conference. Some angular widths were less than 30 degrees, which corresponds to the smallest distance between adjacent participants.

The angular range width was larger to the left at ±90 and ±60 degrees, and larger to the right at ±30 degrees. The bias of the width is relatively small at 0 degrees.

4.3 Experiment 3

Each mark on the Visual Analogue Scale (VAS) was quantified as a distance from the left end and converted into a ratio of the total length of the line. The average, standard deviation, and median of the ratio for each question are shown on Table 2.

The averages of the ratio on the question, “How do you feel about the face?” were almost neutral for all word pairs. Meanwhile, the average of the ratio for “Can you recognize this person? was 0.357, which was closer to “Easy to recognize.” Furthermore, the median for the second question was 0.271, which is smaller than average.

5 DISCUSSION

5.1 Eliminating the Mona Lisa Effect

The results of Experiment 1 show that a facial image gazing front evokes the perception of eye contact only from near the front. Similarly, the facial image not looking forward evokes the perception of eye contact from a specific range of observation angles.

Furthermore, the results of Experiment 2.1 show that there is a range of gaze angles that evoke the perception of eye contact when observed not only from the front but also from oblique directions. Therefore, we can conclude that the Mona Lisa effect was eliminated by our method.

5.2 Relationship between gaze and observation angles

In Experiments 1 and 2.1 and contrary to reality, the observation angle when eye contact is perceived was not same as the gaze angle. The observation angle was always larger than the gaze angle.

The depth position and curvature of eyeballs, which were adjusted arbitrarily, may be the reason for this inconsistency. However, it is expected that the inconsistency will be reduced by further optimization of the position and curvature. In addition, the difference of facial flatness between a real face and the display, which influences the projection on gaze representation, also is considered to be a reason for the inconsistency.

More detailed evaluation and modeling of this relationship are necessary for further development so that we have effective rendering method of eyes for any direction.

5.3 Cone of gaze as the observation angle

The result of Experiment 1 shows that multiple people are likely to perceive eye contact from a single gaze representation when the gaze angle is ±8, ±16, ±24 degrees. Thus, a more precise angular range of gaze perception was investigated in Experiment 2. That investigation implies that one person is presumed to perceive eye contact when the angular distance between participants is more than 38.5 degrees. It is difficult in conferences with more than 6 participants to satisfy this condition. The range can be narrowed by optimizing gaze representation methods such as changing the eyeball position.

Furthermore, the participants in the experiment were instructed to identify the direction from which they no longer feel eye contact. It has been reported in studies of perception that the cone of gaze, defined as the range of gaze angle where mutual gaze occurs, has no clear border but Gaussian probability distribution of the mutual gaze perception as the function of gaze angle [5]. Therefore, it is presumed that similar distributions can be obtained on our method, and a narrower range of gaze perception, which is psychologically significant, can be derived because the border from the results of this study corresponds to the edge of the distribution.

5.4 The identity of the displayed facial image

The results of Experiment 3 show that the impressions of the displayed face are almost neutral against the four word pairs. This suggests that at least the displayed image does not prevent further interaction with negative impressions. However, the impression on Normal–Creepy has to be reconsidered and improved because there is a concern that the display might fall into the uncanny valley [13], which makes the interlocutor uncomfortable.

The results also show that more than half of the participants can easily identify the facial image as that of the remote participant. Therefore, we can conclude that our method provides a high level of identity between the displayed facial image and the real image.

However, the individual variance of recognition is relatively large. The difference in the faces the participants have seen in their lives is presumed to be a reason for the variance. It has been proposed as the prototype effect [16, 17] that we recognize who a person is by seeing differences between this face and the average faces which is generated from the people we have met and easy for us to recognize. The familiar face differs for each person, which leads to the variance on the impression of identity.

6 CONCLUSION

We have described TEllipsoid, an ellipsoidal display for video conference systems, which can provide both accurate eye-gaze transmission and practicality for conferences, namely the convenience to use and the preservation of the identity of displayed face.

We claimed three requirements for a practical videoconference system,

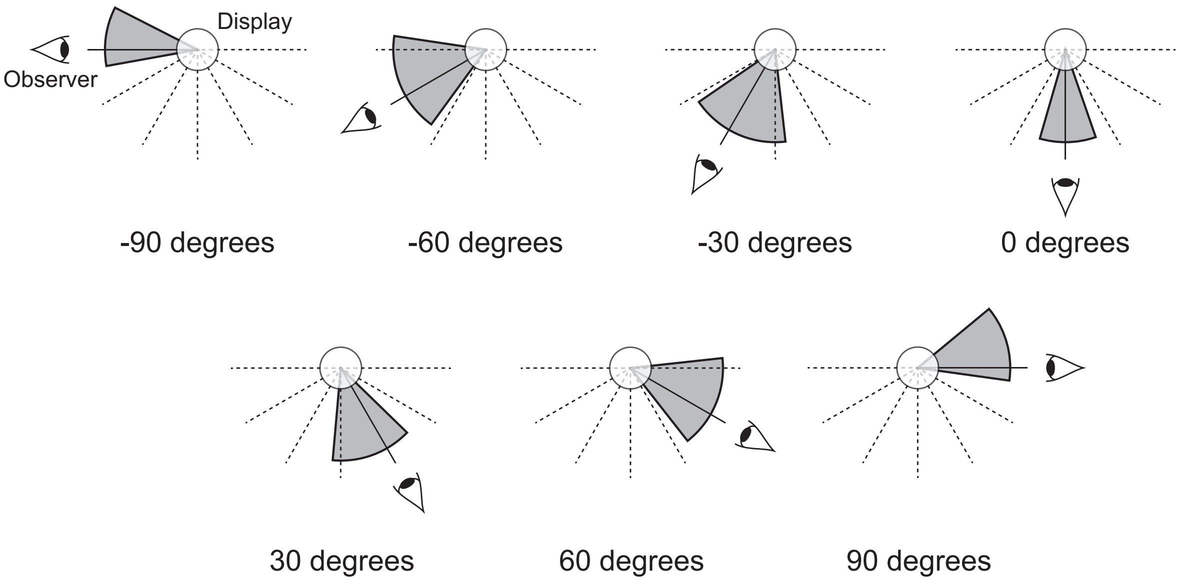


Figure 12: The ranges of observation angles at which the participants perceived eye contact with the strongest-gazed representation.

Table 1: Results of Experiment 2.2.

	Observation Angle [degrees]	-90	-60	-30	0	30	60	90
Left Border	Average [degrees]	-117.0	-98.5	-55.6	-16.1	-5.0	38.1	82.1
	Standard Deviation [degrees]	15.5	14.5	13.5	5.5	12.7	10.0	3.5
	Angular Range Width [degrees]	27.0	38.5	25.6	16.1	35.0	21.9	7.9
Right Border	Average [degrees]	-79.8	-36.1	6.4	19.1	45.8	96.2	129.6
	Standard Deviation [degrees]	3.9	8.3	17.3	8.9	6.9	14.1	21.0
	Angular Range Width [degrees]	10.2	23.9	36.4	19.1	15.8	36.2	39.6

Table 2: Results of Experiment 3.

	Average	SD	Median
Questionnaire: How do you feel about the face?			
Beautiful (0)–Ugly (1)	0.499	0.123	0.504
Normal (0)–Creepy (1)	0.453	0.175	0.421
Friendly (0)–Unfriendly (1)	0.481	0.219	0.504
Humanlike (0)–Robot-like (1)	0.405	0.201	0.329
Questionnaire: Can you recognize that person?			
Easy to recognize (0)–Hard to recognize (1)	0.357	0.262	0.271

- Elimination of the Mona Lisa effect.
- Convenience of the system such as high versatility, high reliability, and low cost.
- Identity between the displayed face and the real face.

Therefore, we propose TEllipsoid, which satisfies all of them. TEllipsoid projects a facial image from a small projector to the ellipsoidal screen via a convex mirror. The prototypes were implemented in both the Virtual Reality environment and the real environment, in which the facial image was modified and gaze representation was implemented.

We evaluated the system on the elimination of the Mona Lisa effect and the gaze transmissibility in a conference situation. As a result of our experiments, we concluded that accurate gaze transmission is available in conferences when the angular distance of the adjacent participants is more than 38.5 degrees.

Our plan for future work is to implement communication functions which are necessary to establish the system as a usable videoconference system, such as:

- Representing facial expressions, blinking, and lip-sync,
- Capturing head and facial movements to represent expression
- Developing an immersive video transmission system from the local conference room to remote participants.

Further evaluation of the system also is necessary as future work, such as evaluating with another facial image, evaluating whether participants other than the intended recipient can comprehend who is gazed from the device, and evaluating the real face as the reference for comparison. The vertical dependence of gaze transmissibility also must be evaluated because the participants can be taller or shorter than TEllipsoid, and eye height can be different. Furthermore, it is necessary to evaluate the effect of the eye convergence parameter on gaze transmissibility because the geometry of TEllipsoid is different from that of the real face. This may make gaze transmissibility optimal when the parameter is different from an exact distance.

The device is still not bright enough for use in a normal conference room because of the low power of the current laser projector. We expect that the power will be improved in a few years.

7 RELATED WORKS

One-to-many 3D video teleconferencing system using an autostereoscopic horizontal-parallax 3D display. [9] is a solution for eliminating the Mona Lisa effect. The face of the remote participant is projected using a high-speed projector and rotating tent-shaped aluminum sheet. The authors evaluated eye contact accuracy and reported that the error of eye gaze ranged between 3 and 5 degrees.

A face-shaped display for the telepresence system. [12] projects the facial image onto a face-shaped screen made from the mold of an actual human face. This is another solution for eliminating the Mona Lisa effect. However, the system is not very versatile since an individual screen is needed for each user. Furthermore, a driving mechanism is needed to move or rotate the face.

A cylindrical multiview videoconference system. [14] was proposed as a system that preserves gaze direction by displaying correct perspective images for multiple viewpoints. It is reported that the gaze transmission accuracy of that method was better than the planar display and also that a single facial image projected to the cylinder surface had more accuracy than the planar display. This method can provide both the elimination of the Mona Lisa effect and convenience. However, the consistent facial parts arrangements, which guarantees the identity, can be observed only from limited viewing angle.

A head robot with a curved surface display. [7] was proposed to guide a human gaze to the target, with control not only of the gaze but also of the facial expression. The gaze representation method is similar to our method, except that the depth positions of eyeball models are different. In the other method, the eyeballs are placed inside the screen, in contrast to protruded eyeballs oi our method. In our method, the projected image is not the line-drawn face but a real face so the eyeballs must protrude according to the real geometry of eyes which protrude from the orbit.

An animated face that is retro-projected on a mask-shaped screen was compared with that projected on a hemisphere screen, the real face, and the human face displayed on a flat screen [2]. The authors concluded that gaze transmission accuracy of the hemisphere is not better than other methods.

ACKNOWLEDGMENTS

This work was supported in part by a grant from JSPS KAKENHI 19K12284.

REFERENCES

[1] R. Cabeza, V. Bruce, T. Kato, and M. Oda. The prototype effect in face recognition: Extension and limits. *Memory & Cognition*, 27(1):139–151, Jan 1999. doi: 10.3758/BF03201220

[2] F. Delaunay, J. de Greeff, and T. Belpaeme. A study of a retro-projected robotic face and its effectiveness for gaze reading by humans. In *Proc. HRI*, pp. 39–44. IEEE, Osaka, Japan, March 2010. doi: 10.1109/HRI.2010.5453271

[3] S. Duncan. Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology*, 23(2):283–292, Aug. 1972. doi: 10.1037/h0033031

[4] M. Gamer and H. Hecht. Are you looking at me? measuring the cone of gaze. *Journal of experimental psychology. Human perception and performance*, 33(3):705–715, Jun. 2007. doi: 10.1037/0096-1523.33.3.705

[5] J. Hakala, J. Kätsyri, T. Takala, and J. Häkkinen. Perception of stereoscopic direct gaze: The effects of interaxial distance and emotional facial expressions. *Journal of vision*, 16(9):5:1–15, Jul. 2016. doi: 10.1167/16.9.5

[6] E. T. Hall. *The hidden dimension*. Anchor Books, Garden City, N.Y., United States, 1969.

[7] M. Hashimoto, H. Kondo, and Y. Tamatsu. Effect of emotional expression to gaze guidance using a face robot. In *Proc. RO-MAN*, pp. 95–100. IEEE, Munich, Germany, Aug 2008. doi: 10.1109/ROMAN.2008.4600649

[8] M. Heike, K. Fujita, and T. Tanaka. Personalized and deformed avatars for realizing avatar-mediated real person-to-person communication. In *Proc. OCSC*, pp. 207–215. Springer-Verlag, Berlin, Heidelberg, 2009. doi: 10.1007/978-3-642-02774-1_23

[9] A. Jones, M. Lang, G. Fyffe, X. Yu, J. Busch, I. Mcdowall, M. Bolas, and P. Debevec. Achieving eye contact in a one-to-many 3d video teleconferencing system. *ACM Transactions on Graphics*, 28(3):64:1–8, Aug. 2009. doi: 10.1145/1531326.1531370

[10] A. Kendon. Some functions of gaze-direction in social interaction. *Acta Psychologica*, 26(1):22–63, Jan. 1967. doi: 10.1016/0001-6918(67)90005-4

[11] M. Kouchi and M. Mochimaru. AIST anthropometric database. National Institute of Advanced Industrial Science and Technology, H16PRO 287.

[12] K. Misawa, Y. Ishiguro, and J. Rekimoto. Livemask: A telepresence surrogate system with a face-shaped screen for supporting nonverbal communication. In *Proc. AVI*, pp. 394–397. ACM, New York, NY, USA, 2012. doi: 10.1145/2254556.2254632

[13] M. Mori, K. F. MacDorman, and N. Kageki. The uncanny valley [from the field]. *IEEE Robotics Automation Magazine*, 19(2):98–100, Jun. 2012. doi: 10.1109/MRA.2012.2192811

[14] Y. Pan and A. Steed. A gaze-preserving situated multiview telepresence system. In *Proc. CHI*, pp. 2173–2176. ACM, New York, NY, USA, 2014. doi: 10.1145/2556288.2557320

[15] S. Rogers, M. Lunsford, L. Strother, and M. Kubovy. *The Mona Lisa effect: Perception of gaze direction in real and pictured faces*, pp. 19–24. Lawrence Erlbaum, Jan. 2003.

[16] T. Valentine and V. Bruce. The effects of distinctiveness in recognising and classifying faces. *Perception*, 15(5):525–535, Oct. 1986. PMID: 3588212. doi: 10.1068/p150525

[17] T. M. Valentine and V. Bruce. Recognizing familiar faces: the role of distinctiveness and familiarity. *Canadian journal of psychology*, 40(3):300–305, Sep. 1986.

[18] W. H. Wollaston. Xiii. on the apparent direction of eyes in a portrait. *Philosophical Transactions of the Royal Society of London*, 114(1):247–256, Jan. 1824. doi: 10.1098/rstl.1824.0016