An Embodied Group Entrainment Characters System Based on the Model of Lecturer's Eyeball Movement in Voice Communication

Yoshihiro Sejima (Yamaguchi University)
Graduate School of Science and Engineering,
Computer Science and Design Engineering,
sejima@yamaguchi-u.ac.jp

Tomio Watanabe (Okayama Prefectural University)
Faculty of Computer Science and Systems Engineering,
Department of Systems Engineering,
watanabe@cse.oka-pu.ac.jp

Mitsuru Jindai (University of Toyama)
Graduate School of Science and Engineering,
Mechanical and Intellectual Systems Engineering,
jindai@eng.u-toyama.ac.jp

Atsushi Osa (Yamaguchi University)
Graduate School of Science and Engineering,
Computer Science and Design Engineering,
osaa@yamaguchi-u.ac.jp

Yukari Zushi (Yamaguchi University)
Graduate School of Science and Engineering,
Perceptual Sciences and Design Engineering,
s019ff@yamaguchi-u.ac.jp

***Abstract*—We have developed a speech-driven embodied group entrained communication system called "SAKURA" for enabling group interaction and communication. In this system, speech-driven computer-generated (CG) characters called InterActors with functions of both speakers and listeners are entrained as a teacher and some students in a virtual classroom by generating communicative actions and movements. In this study, for enhancing group interaction and communication, we analyze the eyeball movements of a lecturer communicating in a virtual group by using an embodied communication system with a line-of-sight measurement device. On the basis of the analysis results, we propose an eyeball movement model that consists of a saccade model and a model of the lecturer's gaze at the audience, called "group gaze model." The saccade model reveals eyeball movement with a delay of 0.20 s with respect to the lecturer's head movement. A group gaze model reveals the rate of the lecturer's gaze (Center: 60%, Left-side: 27%, Right-side: 13%). Then, we develop an advanced communication system in which the proposed model is used with SAKURA. Using this system, we perform experiments and carry out sensory evaluation for determining the effects of the proposed model. The results reveal that the proposed model is effective for group interaction and communication in the speech-driven embodied group entrainment characters system.**

Keywords-Human Interface; Human Interaction; Embodied Communication; Group Interaction; Eyeball Movement.

I. INTRODUCTION

With the advancements in the field of information technology, it is now becoming possible for humans to use CG characters called avatars to communicate in a 3D virtual space over a network [1]. Because the avatars express nonverbal behavior based on keyboard commands, current systems do not simulate embodied sharing using the synchrony of embodied rhythms such as nodding and body movements in human face-to-face communication. In such communications, not only verbal messages but also nonverbal behavior such as nodding, body movements, gaze, and facial expressions are rhythmically related and mutually synchronized between talkers [2]. This synchrony of embodied rhythms in communication is called entrainment, and it generates the sharing of embodiment in human interactions [3].

Focusing on the entrainment of embodied communication in our previous work, we analyzed the entrainment between a speaker's speech and a listener's nodding motion in a face-to-face communication and developed InterRobot Technology (iRT), which generates a variety of communicative actions and movements such as nodding and body movements using a speech input [4]. In addition, we developed an interactive CG character called "InterActor" and demonstrated that it can effectively support human interactions and communication. We also developed a speech-driven embodied group entrained communication system called "SAKURA" for enabling group interaction and communication in which InterActors are entrained as a teacher and some students in a virtual classroom. Furthermore, we demonstrated that the developed system could effectively support human interactions and communication [5].

In group communication, not only the lecturer's body movements but also the line-of-sight of the lecturer, such as gaze and eye contact, play an important role in enhancing the embodied interaction and communication [6]. Furthermore, it has been reported that the line-of-sight is important for enhancing the embodied interaction in an avatar-mediated

 ISBN: 978-1-61208-325-4

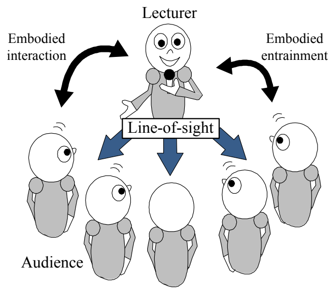


Figure 1. Research concept.

communication. For example, the teleconference system was developed by using some CG characters with the technology of mixed reality, and demonstrated that the line-of-sight of talkers was influenced by the position and direction of CG characters [7]. In addition, the interactive communication system which controls the line-of-sight among three avatars was developed based on the talker's voice information such as sound pressure and pitch, and demonstrated that the interaction of voice was enhanced by modulating the line-of-sight of the talker [8]. However, it is difficult to enhance such embodied interaction because the characteristics of a lecturer's line-of-sight in group communication have not been established thus far. Therefore, it is essential to develop a group embodied communication system that has the characteristics of a lecturer's line-of-sight in order to enable smooth communication during an embodied interaction [9] (Figure 1).

In this study, we analyze a lecturer's behavior in a virtual group communication. In particular, by focusing on the lecturer's line-of-sight, the eyeball movements are measured by using a line-of-sight measurement device, and the characteristics of the lecturer's line-of-sight such as group gaze are analyzed. On the basis of the analysis results, we propose an eyeball movement model that consists of a saccade model and a model of the lecturer's gaze at an audience, called "group gaze model," for enhancing group interaction and communication. In order to evaluate the effects of the proposed model on group interaction and communication, we develop an advanced communication system in which the model is used with SAKURA. The effectiveness of the proposed eyeball movement model is demonstrated for performing the communication experiments with a sensory evaluation using the developed system.

II. ANALYSIS OF LINE-OF-SIGHT IN GROUP COMMUNICATION

A. *InterActor*

In order to support human interaction and communication, we developed a speech-driven embodied entrainment character called InterActor, which performs the functions of both speaker and listener (Figure 2). The listener's interaction model includes a nodding reaction model that estimates the nodding timing from a speech ON-OFF pattern and a body reaction model linked to the nodding reaction model [4]. The timing of nodding is predicted using a hierarchy model consisting of two stages: macro and micro (Figure 3). The macro stage estimates whether a nodding response exists or not in a duration unit, which consists of a talkspurt episode $T(i)$ and the following silence episode $S(i)$ with a hangover value of 4/30 s. The estimator $M_u(i)$ is a moving-average (MA) model, expressed as the weighted sum of unit speech activity $R(i)$ in (1) and (2). When $M_u(i)$ exceeds a threshold value, nodding $M(i)$ also becomes an MA model, estimated as the weighted sum of the binary speech signal $V(i)$ in (3).

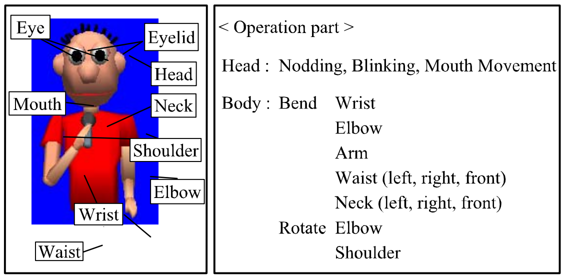


Figure 2. InterActor.

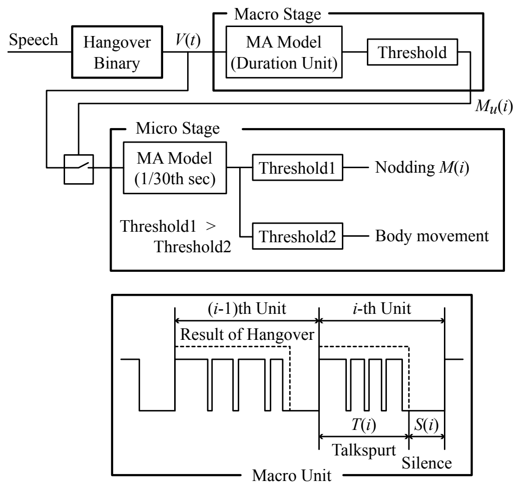


Figure 3. Listener's interaction model.

$$M_u(i) = \sum_{j=1}^{J} a(j)R(i-j) + u(i) \tag{1}$$

$$R(i) = \frac{T(i)}{T(i)+S(i)} \tag{2}$$

$a(j)$: linear prediction coefficient

$T(i)$: talksqurt duration in the i - th duration unit

$S(i)$: silence duration in the i - th duration unit

$u(i)$: noise

$$M(i) = \sum_{j=1}^{K} b(j)V(i-j) + w(i) \tag{3}$$

$b(j)$: linear prediction coefficient

$V(i)$: voice

$w(i)$: noise

The body movements are related to the speech input because the neck and one of the wrists, elbows, or arms, or the waist are operated when the body threshold is exceeded. The threshold is set lower than that of the nodding prediction of the MA model, which is expressed as the weighted sum of the binary speech signal to nodding. In other words, when InterActor functions as a listener for generating body movements, the relationship between nodding and other movements is dependent on the threshold values of the nodding estimation.

The body movements in the case of a speaker are also related to the speech input by operating both the neck and one of the other body actions at the timing over the threshold, which is estimated by the speaker's interaction model as its own MA model of the burst-pause of speech to the entire body motion [4]. Because speech and arm movements are related at a relatively high threshold value, one of the arm actions in the preset multiple patterns is selected for operation when the power of speech is over the threshold.

B. Experimental System

The experimental setup is shown in Figure 4. In this experiment, for using InterActor as a virtual listener, three isomorphic displays (I･O DATA LCD-AD203G) were used. InterActor was represented with each display having a resolution of 1600 x 1200 pixels; it was generated using Microsoft DirectX 9.0 SDK and a Windows XP workstation (CPU: Corei7 2.93 GHz, Memory: 3 GB, Graphics: NVIDIA Geforce GTS250). The frame rate at which InterActor was represented was 30 fps. The three displays were synchronized using the image distributor (ELECOM VSP-A2). The lecturer and the left and right displays made up one corner each of an equilateral triangle having a side length of 2 m. The positions and angles of the lecturer's body movements were measured using four magnetic sensors (Polhemus FASTRAK) placed on the top of the lecturer's head, both wrists, and the back of the lecturer's body. The image of the lecturer's eyeball was measured using a line-of-sight measurement device [10] (Figure 5) and was input to

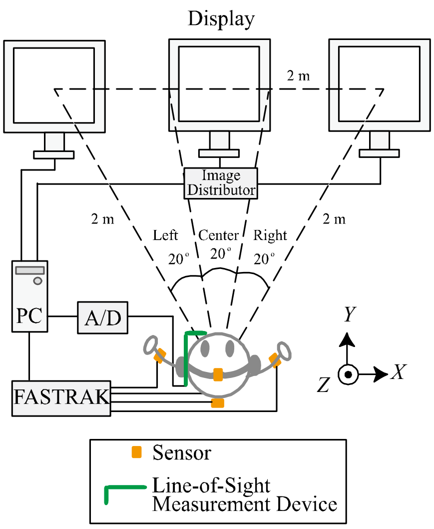


Figure 4. Experimental setup.

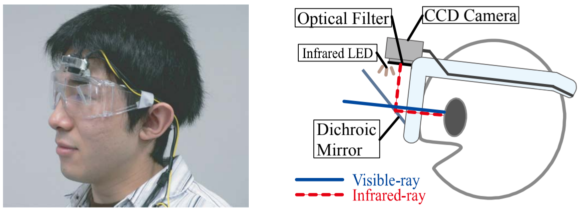


Figure 5. Line-of-sight measurement device.

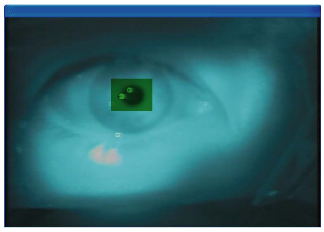


Figure 6. Eyeball image.

Figure 7. Experimental setup.

the PC through an A/D converter (CANOPUS ADVC110). The angle of the lecturer's eyeball movement was calculated by the template matching of the cornea (Figure 6). The voice was sampled using 16 bits at 11 kHz with a headset (SONY DR-260DP). The measured data were recorded on an HDD in real time.

The experimental process was as follows: First, the lecturer used the system for 1 min for the calibration of his eyeball movement. Next, the lecturer was told to talk on general conversational topics to the three InterActors for 5 min (Figure 7). The conversational topics were not specified, and the instructor was told that the rate of his gaze was to be equally divided between the three InterActors. The three InterActors behaved as virtual listeners by nodding and making body movements in real time. Ten male students played the role of the lecturer.

C. Analysis of Line-of-Sight in Group Communication

First, the rate of the lecturer's gaze was analyzed for the three InterActors. In this analysis, using the data on head movement, we defined the lecturer's gaze as Right-side, Center, or Left-side, as shown in Figure 4. An example of the time change of the head movement is shown in Figure 8. The figure shows that the lecturer mainly gazed at the center and the duration for which the lecturer subconsciously gazed at the Left-side was longer than that for the Right-side. The average duration of the lecturer's gaze is given in Table 1. The average duration of the Center gaze accounted for approximately 60% of the time, and the duration of Left-side (27%) was twice that of the Right-side gaze (13%). This

 ISBN: 978-1-61208-325-4

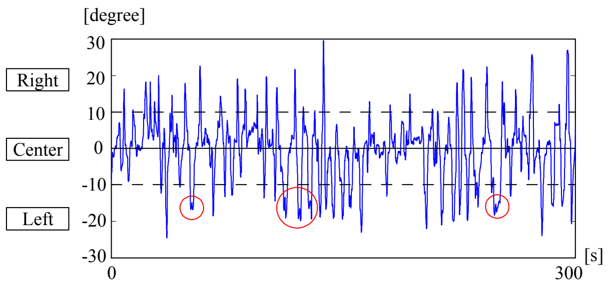


Figure 8. Example of time change of head movement.

result is consistent with the fact that the Left-side is

Table 1. Result of gaze duration.

	Left	Center	Right (%)
Average	27.6	59.5	12.9
Standard deviation	20.0	24.9	8.7

predominant over the Right-side in the spatial cognitive functions of humans [11] [12].

Next, the relationship of the timing between the lecturer's head movement and the eyeball movement during the communication experiment was analyzed. The angle of Z axis $p(i)$ was measured at a sampling rate of 30 Hz using the magnetic sensor placed on the lecturer's head; this angle was then used for calculating the head movement $x(i)$ by measuring the difference between the following and the previous data $[p(i+1)-p(i-1)]$. The coordinate value $c(i)$ was measured at the sampling rate 30 Hz using the line-of-sight device on the lecturer's head; $c(i)$ was then used for calculating the eyeball movement $y(i)$ by taking the difference between the following and the previous data $[c(i+1)-c(i-1)]$. The relationship of time changes was analyzed using the following cross-correlation function $C(\tau)$.

$$C(\tau) = \frac{\sum_{i=1}^{n-\tau}\{x(i)-\mu_x\}\{y(i+\tau)-\mu_y\}}{\sqrt{\sum_{i=1}^{n}\{x(i)-\mu_x\}^2}\sqrt{\sum_{i=1}^{n}\{y(i)-\mu_y\}^2}} \quad (4)$$

μ_x, μ_y : average of x, y

n : number of data

τ : time delay

Figure 9 shows an example of time change of the cross-correlation function $C(\tau)$ in an analysis period of 30 s over 5 min. A strong positive correlation between the lecturer's head movement and eyeball movement was confirmed for $\tau = 0.20$ s. The average delay time of the eyeball movement was 0.196 ± 0.06 s over 5 min. This result showed that the eyeball movement had a delay time with respect to the head movement of 0.20 s during group communication; this delay time was similar to the latent time of a saccade in the body functions of a human (0.20 s) [13] [14].

The obtained results can be summarized as follows: The lecturer mainly gazed at the center, and he gazed at the Left-side twice as much as the Right-side. Further, the eyeball movement had a delay time with respect to the head movement of 0.20 s.

III. EYEBALL MOVEMENT MODEL

In this research, we have proposed an eyeball movement model that generates a lecturer's eyeball movement for enhancing group communication on the basis of the characteristics revealed by the abovementioned analysis. This model consists of a saccade model and a model of the lecturer's gaze at an audience called the "group gaze model." An outline of the proposed model is as follows:

A. Saccade model

The main characteristic of the saccade model is an eyeball movement with a delay of 0.20 s with respect to the lecturer's head movement. First, the angle of the lecturer's head movement was calculated in a virtual space. If the lecturer's head moved, the eyeball moved with a delay of 0.20 s with respect to the head movement in the same direction (Figure 10).

B. Group gaze model

The characteristic of the group gaze model is the duration of the lecturer's gaze (Center: 60%, Left-side: 27%, Right-side: 13%) [9].The lecturer's gaze is generated stochastically with an exponential distribution based on the abovementioned analysis. An example of the lecturer's gaze

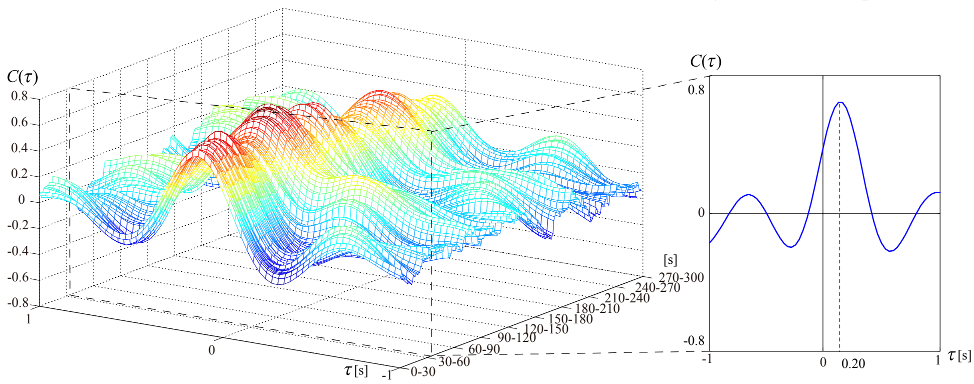


Figure 9. Example of time change of cross-correlation $C(\tau)$ between a lecturer's head movement and his own eyeball movement.

 ISBN: 978-1-61208-325-4

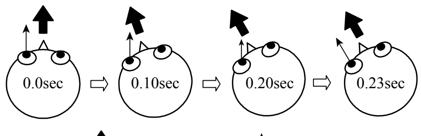


Figure 10. Example of the saccade model.

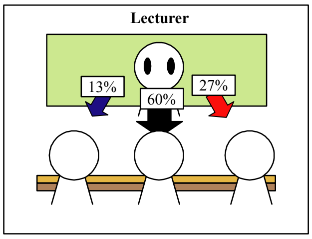


Figure 11. Group gaze model.

is shown in Figure 11. It is expected that the lecturer's gaze will be effective for group communication when the gaze duration is varied.

IV. A SPEECH-DRIVEN EMBODIED GROUP ENTRAINMENT CHARACTERS SYSTEM

A. *SAKURA*

A speech-driven embodied group entrained communication system called "SAKURA" has been developed for enabling group interaction and communication [5]. Five InterActors play the role of students, and one InterActor plays the role of a teacher; they are arranged in a virtual classroom where they are entrained on the basis of only a speech input. By using SAKURA, talkers can communicate with a sense of unity through the entrained InterActors by using only a speech input.

B. *Speech-driven embodied group entrainment system*

An advanced communication system was developed in which the proposed model was used with SAKURA (Figure 12). The virtual space was generated by Microsoft DirectX 9.0 SDK and a Windows XP workstation (HP workstation xw4200: CPU: Pentium4 2.8 GHz, Memory: 1 GB, Graphics: NVIDIA Quadro FX3400). The voice was sampled using 16 bits at 11 kHz. The frame rate at which the CG characters were represented was 30 fps.

When a lecturer's speech is fed into the system as an input, the lecturer's InterActor generates communicative body movements and actions, and an eyeball movement based on the proposed model. An example of the communicative actions and movements of the lecturer is

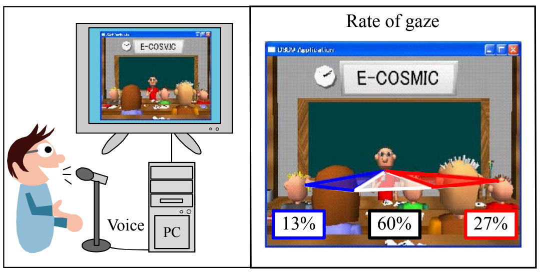


Figure 12. Configuration of the developed system.

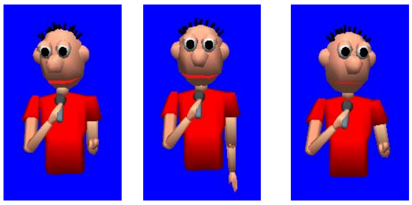


Figure 13. Example of InterActor's body motion and gaze as lecturer.

shown in Figure 13. The mouth operation of the InterActor was realized by the switching operation synchronized with the burst-pause of voice. The audience characters respond to the utterances with appropriate timings by means of their entire body movements, including nodding, blinking, and communicative actions in the manner of listeners [4]. As a result, the lecturer's InterActor gives the audience a natural line-of-sight and generates a communication environment in which the sense of unity is shared by the embodied entrainment.

V. EVALUATION OF EYEBALL MOVEMENT MODEL

A. *Evaluation experiment of saccade model*

A preliminary experiment was performed to evaluate the saccade model. The main characteristic of a saccade model is an eyeball movement having a delay of 0.20 s with respect to the lecturer's head movement. We performed an experiment in which the delay time of the eyeball movement had three patterns: 0.00 s, 0.20 s, and 0.40 s. An experimental communication scene using the system is shown in Figure 14. Ten male subjects were used for evaluating the three patterns. The result showed that the most common delay was 0.20 s; it was observed in eight out of 10 subjects. Therefore, in this research, the delay time for the eyeball movement was set to be 0.20 s.

B. *Evaluation experiment of group gaze model*

The effectiveness of the saccade model was demonstrated in the foregoing section. In this section, the group gaze model is evaluated.

1) Experimental Method: In this experiment, the subjects were asked to watch a video. The video was made by recording the display using a system input to the recorded speech data in 2 min. The speech content was an opinion on consumption tax. In this experiment, three modes were compared: in the first mode, a lecturer gazed at the center of the audience (mode (A)); in the second, the lecturer gazed at the entire audience equally (mode (B))

Figure 14. Subject watching video to evaluate modes.

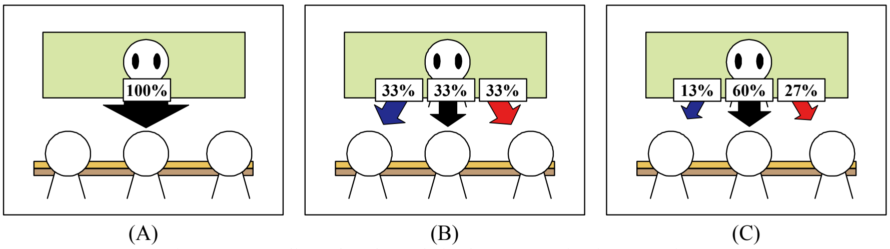


Figure 15. Outline of each mode in the communication experiment.

[15]; and in the third mode, the lecturer gazed at the audience by using the group gaze model (mode(C)) (Figure 15).

The experimental procedure was as follows: first, the subjects watched the video in each mode. Next, they were instructed to perform a paired comparison of the modes. In the paired comparison experiment, based on their preferences, they selected the better mode. Finally, they watched the video in each mode and evaluated each mode on a seven-point bipolar rating scale ranging from -3 (not at all) to 3 (extremely); 0 denoted moderation. The subjects were 30 Japanese students (15 females and 15 males) presented with the abovementioned three modes in a random order.

2) Results of Sensory Evaluation: The result of the paired comparison is shown in Table 2. Further, Figure 16 shows the calculated results of the evaluation provided in Table 2 and based on the Bradley-Terry model given in (5) and (6) [16].

$$p_{ij} = \frac{\pi_i}{\pi_i + \pi_j} \qquad (5)$$

$$\sum_i \pi_i = const.(=100) \qquad (6)$$

π_i : Intensity of i

p_{ij} : probability of judgment that i is better than j

The consistency of the matching of the modes was confirmed by performing a test of goodness of fit ($x^2(1,0.05) = 3.84 > x_0^2 = 0.02$) and the likelihood ratio test ($x^2(1,0.05) = 3.84 > x_0^2 = 0.02$). The proposed mode (C) was evaluated to be the most affirmative, followed by the (B) equal-gaze and (A) center-gaze modes.

Table 2. Result of paired comparison.

	(A)	(B)	(C)	Total
(A)		5	2	7
(B)	25		7	32
(C)	28	23		51

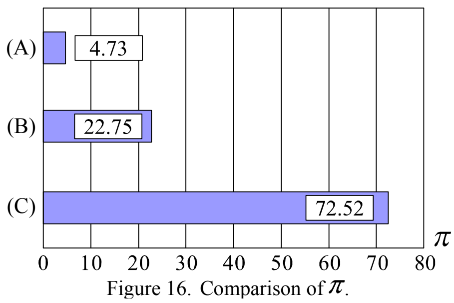


Figure 16. Comparison of π.

The questionnaire result is shown in Figure 17. From the results of the Friedman signed-rank test, we found that "Interaction," "Natural line-of-sight," "Unification," "Realistic sensation," "Vividness," and "Preference," had a significance level of 1% between mode (C) and mode (A). Further, "Lecturer's gaze," had a significance level of 1%, and "Interaction," and "Preference," were at 5% between mode (C) and mode (B). In both the experiments, mode (C)

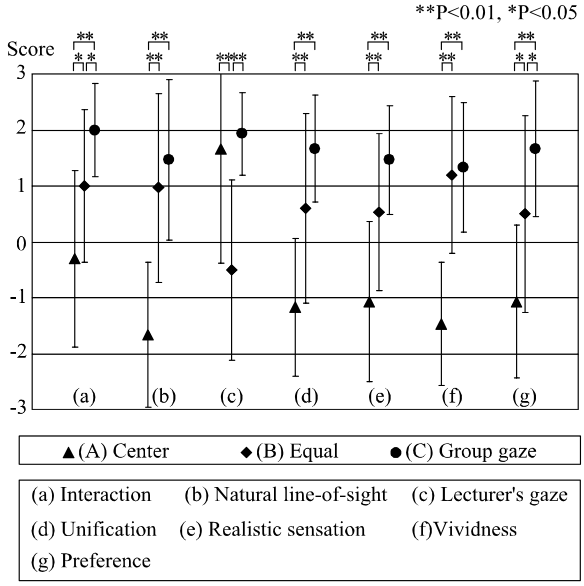
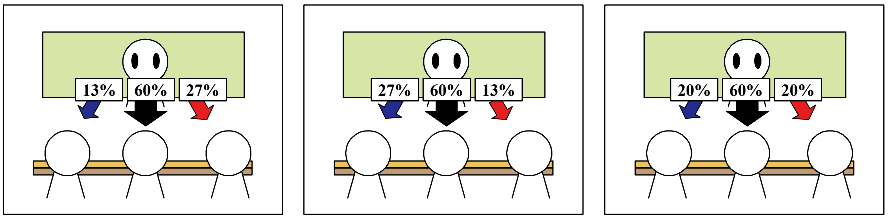


Figure 17. Seven-point bipolar rating.

of the proposed eyeball movement model was most often evaluated to be the best with respect to group interaction and communication. These results indicated that the proposed model in which a lecturer gazed at the center of the audience moderately (60%) was the best of the models considered in this study.

C. Evaluation of group gaze rate

The effectiveness of the group gaze model in which a lecturer gazed at the center of the audience 60% of the time was demonstrated in the foregoing section. In this section, the gaze rate of the group gaze model is evaluated by



(C) (D) (E)

Figure 18. Outline of each mode in the communication experiment.

changing the gaze rate of the remaining 40% of the gaze duration.

1) Experimental Method: In this experiment, three modes were compared: in the first, a lecturer gazed at the audience with the group gaze model (mode (C)); in the second, a lecturer gazed at the audience for reversing the gaze rate between the right-side and the left-side (mode (D)); and in the third mode, a lecturer gazed at the audience for equalizing the gaze rate between the right-side and the left-side (mode (E)) (Figure 18). The experimental procedure was the same as that detailed in Section V.B.1. The subjects were 30 Japanese students (15 females and 15 males) other than the ones mentioned in the foregoing section.

2) Results of Sensory Evaluation: The result of the paired comparison is shown in Table 3. Further, Figure 19 shows the calculated results of the evaluation provided in Table 3.

The consistency of the matching of the modes was confirmed by performing a test of goodness of fit ($x^2(1, 0.05) = 3.84 > x_0^2 = 1.40$) and the likelihood ratio test (x^2 $(1, 0.05) = 3.84 > x_0^2 = 1.41$). The proposed mode (C) was evaluated to be the most affirmative, followed by the (E) equal-gaze and (D) reverse modes.

The questionnaire result is shown in Figure 20. From the results of the Friedman signed-rank test, we found that "Unification," had significance level of 5% between mode (C) and mode (D). Further, "Unification," had a significance level of 5% between mode (D) and mode (E). In both the experiments, mode (D) of reversing the gaze rate between

Table 3. Result of paired comparison.

	(C)	(D)	(E)	Total
(C)		24	14	38
(D)	6		10	16
(E)	16	20		36

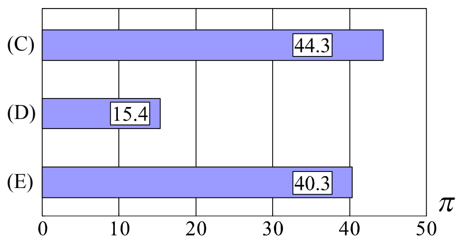


Figure 19. Comparison of π.

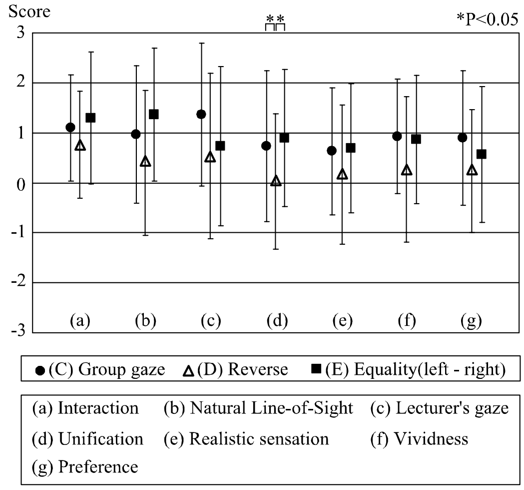


Figure 20. Seven-point bipolar rating.

the right-side and the left-side was not evaluated for group interaction and communication. This result is consistent with the theory that the left-side is predominant over the right-side in the case of the spatial cognitive functions of humans [11] [12].

D. *Effectiveness of group gaze model*

In the foregoing section, mode (D) in which the gaze rate was reversed between the left-side and the right-side was not evaluated with respect to group communication. In this section, the effect of group gaze is evaluated by comparing mode (C) to mode (E).

1) Experimental Method: In this experiment, two modes were compared: in the first, a lecturer gazed at the audience using the group gaze model (mode (C)); in the second mode, a lecturer gazed at the audience for equalizing the gaze duration between the right-side and the left-side (mode (E)). The experimental procedure was the same as that detailed in Section V.B.1. The subjects were another 20 Japanese students (10 females and 10 males).

2) Results of Sensory Evaluation: The result of the paired comparison is shown in Figure 21. The questionnaire

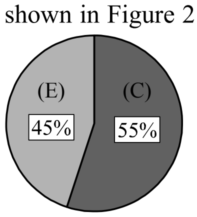


Figure 21. Result of paired comparison.

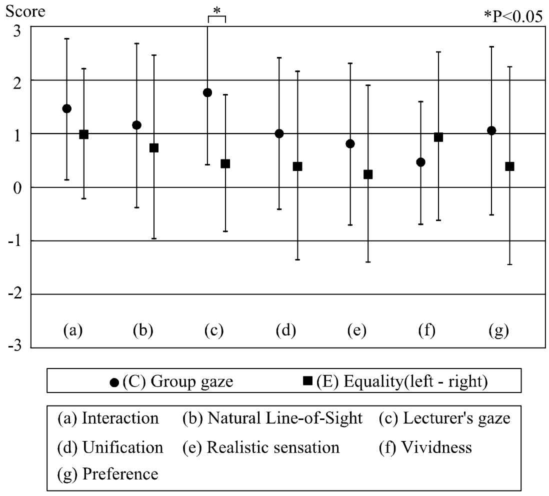


Figure 22. Seven-point bipolar rating.

result is shown in Figure 22. In accordance with the results of the Wilcoxon signed-rank test, the parameter "Lecturer's gaze," had a significance level of 5% between mode (C) and mode (E). This result indicates that the system with the proposed model is effective in group interaction and communication.

VI. CONCLUSION

In this study, we analyzed the characteristics of a lecturer's eyeball movement in the case of group communication by using an embodied communication system with a line-of-sight measurement device. On the basis of the analysis results, we proposed an eyeball movement model that consists of a saccade model and a model of the lecturer's gaze at an audience, called "group gaze model," for enhancing group interaction and communication. The proposed model could be summarized as follows: the main characteristic of a saccade model is an eyeball movement with a delay of 0.20 s with respect to the lecturer's head movement, and a group gaze model has the following durations of the lecturer's gaze: Center 60%, Left-side 27%, and Right-side 13%. Further, we developed an advanced communication system in which the proposed model was applied to a speech-driven embodied group entrained communication system. By using this system, we performed communication experiments and carried out a sensory evaluation. The effectiveness of the proposed model was demonstrated during group interaction and communication in a speech-driven embodied group entrainment characters system.

ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Number 24700536, 22300045, 24118707, 25330239.

REFERENCES

[1] Dilley, C., "Virtual World Summit 2007 Report," ASCII, 2008, pp. 94-97.

[2] Condon, W.S. and Sander, L.W., "Neonate movement is synchronized with adult speech,"Science, no. 183, 1974, pp. 99-101.

[3] Watanabe, T., "Human-Entrained Embodied Interaction and Communication Technology for Advanced Media Society,", Proceedings of 16th IEEE International Symposium on Robot and Human Interactive Communication(RO-MAN2007), 2007, pp. 31-36.

[4] Watanabe, T., Okubo, M., Nakashige, M., and Danbara, R., "Interactor: Speech-driven embodied interactive actor,"International Journal of Human-Computer Interaction, vol. 17, 2004, pp. 43-60.

[5] Watanabe, T. and Okubo, M., "SUKURA: Voice-Driven Embodied Group-Entrained Communication System," Proceedings of the 10th International Conference on Human-Computer (HCI International 2003), vol. 2, 2003, pp. 558-562.

[6] Kendon, A., "Some Functions of Gaze-direction in Social Interaction," Aca Psychologica, vol.26, 1967, pp. 22-63.

[7] Inoue, T., "Mixed Reality Meeting System Enabling User to Keep and Share Interpersonal Distance in the Real World," Transactions of Information Processing Society of Japan, vol.50, no.1, 2009, pp. 246-253.

[8] Ishii, R., Miyajima, T. and Fujita, K., "Avatar's Gaze Control to Facilitate Conversation in Virtual-Space Multi-User Voice Chat System, " Transactions of Human Interface Society, vol.10, no. 1, 2008, pp. 87-94.

[9] Sejima, Y., Watanabe, T., Jindai, M., Osa, A. and Zushi, Y., "A Speech-driven Embodied Group Entrainment System with the Model of Lecturer's Eyeball Movement," Proceedings of the 21st IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN2012), 2012, pp. 1086-1091.

[10] Sejima, Y., Watanabe, T. and Jindai, M., "An Embodied Communication System Using Speech-Driven Embodied Entrainment Characters with an Eyeball Movement Model," Transactions of the Japan Society of Mechanical Engineers, Series C, vol.76, no.762, 2010, pp.340-350,.

[11] Kato, T.,"Psychology of a Sense of Space (in Japanese),"Shin-Yo-Sha Ltd., 2001, pp.37-47.

[12] Nehira, K., "The Left and Right of Animate Nature (in Japanese)," KYORITSU SHUPPAN CO. LTD., 1998, pp.103-106.

[13] Fukushima, J., "Disturbance of Eye Movements in Schizophrenic Patients," Bulletin of Hokkaido University Collection of Scholarly and Academic Papers, vol.2, 1989, pp.1-13.

[14] Ohno, T., "What Can Be Learned From Eye Movement?: Understanding Higher Cognitive Processes From Eye Movement Analysis," Transactions of the Japanese Cognitive Science Society, vol.9, no.4, 2002, pp.565-579.

[15] Aramaki, M., "Skill of Presentation (in Japanese)," SANNO Institute of Management Publication dept., 2010, pp.39-42.

[16] Luce, R. D., "Individual Choice Behaviour: A Theoretical Analysis," New York: J. Wiley, 1959.