2015 IEEE/SICE International Symposium on System Integration (SII)
December 11-13, 2015. Meijo University, Nagoya, Japan

Look at You, Look at Me: Detection and Analysis of Mutual Gaze Convergence in Face-to-Face Interaction

Chidchanok Thepsoonthorn, Takahiro Yokozuka, Jinhwan Kwon, Robin Miao Sin Yap, Shunsuke Miura, Ken-ichiro Ogawa, and Yoshihiro Miyake

Abstract— **In human communication, mutual attentiveness between individuals is essential for social interaction. The most significant and reliable indicator of mutual attentiveness is gaze. Despite the availability of diverse exorbitant eye-tracking devices, all claim highly accurate gaze positioning, which exceeds the necessity for average researches and applications. We propose an alternative model for mutual gaze convergence detection using economical video capturing devices and straightforward techniques that can investigate natural gaze behaviors in human-human interaction. We conducted a task supposing lecture in face-to-face interaction and applied our model to investigate (1) average percentage of total mutual gaze convergence, (2) duration of mutual gaze convergence, (3) the domination relationship between lectures and students during short lectures, including (4) the effects of attention span, which claimed to last for 10-20 minutes, toward short lecture task (5 minutes). Our results showed that the average total mutual gaze convergence between the lecturer and the student is 52.83%. It reveals the fact that in both long and short lecture, humans have momentary gaze fixation at a particular direction. We also found out that attention span has effect on different durations of lecture task, resulting in different dominance between lecturer and student, be it student for long lecture and lecturer for short lecture. We believe that our developed model implementation and findings exhibit practical applicability and worthy of contribution to the communities of related research areas.**

I. Introduction

Human communication is achieved by the interaction between speaker and listener. Richardsona et al.'s study [1] revealed that two individuals can efficiently interact when there exists mutual attentiveness and increased enthusiasm, resulting in greater success of the communication. Mutual attentiveness in speaker-listener interaction can be expressed in many behavioral means, both verbal and non-verbal, especially, gaze. Henderson [2] and Just [3] asserted that gaze is the most significant, reliable, and observable indicator of mutual attentiveness. Moreover, many previous studies found that gazing conveys socially relevant information and mutual gazes in face-to-face communication can furthermore be interpreted as a modality of hostility, anger [4], romantic attraction [5], interest and attention [6].

However, detection of mutual gazes is a challenging task, due to the need of bilateral eye-tracking devices capable of detecting gaze directions and their convergence. Despite diversity of available exorbitant eye-tracking devices, they all focus on achieving accurate gaze positioning [8], [9]. Employing these devices in a mutual gaze convergence detection task or only direction concerned studies becomes excessive, since the functionality is not necessary in said applications and therefore can be thought of as overcomplicating the solution. In addition, the learning curve for using these devices can be steep, especially in calibration. Calibrating the device requires knowledgeable person and complex techniques, which may resultantly cause frustration and time consumption when using the device.

In this paper, we propose an alternative model for gaze convergence detection, in which mutual gaze direction can be detected using a simple heuristics of eye detection techniques and economical video capturing devices. The detection results together with analysis will provide evidences on natural human gaze behaviors in communication, which will be beneficial for developing face-to-face interaction and gaze-friendly applications or human-robot interaction to approach natural human behaviors. As a proof-of-concept study, we conduct the experiment on short lecture task and investigate 4 following questions in our study. We believe that the answers for these 4 questions are steps closer to have insight understanding of natural human gaze behaviors in interaction.

- *Question 1: Average percentage of total mutual gaze convergence between lecturers and students in short lecture task*

The amount of total mutual gaze or mutual attention highly varies across different tasks. For this question, we expect to derive an insightful trend of total gaze convergence in human behaviors. This analysis will be useful for making human-robot interaction more adaptive and more natural.

- *Question 2: Duration of mutual gaze convergence during the short lecture task*

In this study, we asked our participants to interact with one another naturally. The gaze direction of each participant is based on his or her natural behavior. We hypothesize that the duration of mutual gaze convergence in short lecture task occurs for a short time more frequently during the lecture task as similar to what previously examined for longer task by Bunce et al. [10].

- *Question 3: Who dominate the amount of mutual gaze convergence in short lecture task*

Broz et al. [8] discovered that there are two kinds of

All authors are with the Tokyo Institute of Technology, Yokohama, JAPAN (corresponding author to provide phone: +81(45)924-5646; E-mail: chidchanok@myk.dis.titech.ac.jp)

978-1-4673-7241-1/15/$31.00 ©2015 IEEE

participants: high gazer and low gazer. The *high gazers* look at their partner more frequently, while the *low gazers* control the amount of mutual gazes, henceforth called *the dominant*. We, therefore, aim to investigate the dominants that control the amount of mutual gazes in short lecture tasks.

- *Question 4: Effects of attention span toward short lecture task*

To address this question, we investigate natural human-human interaction, in which attention span may affect human behaviors differently in short and long lecture task. Previous studies [11]-[13] showed that the attention span of ordinary humans lasts for 10 to 20 minutes. In our experiments, we eliminate the variable of attention span by limiting the duration of lecture tasks to approximately five minutes. This is considered a short duration task and it is still covered by the attention span aforementioned. Our discussion lays a possible roadmap to a further analysis on optimal course designs and time allocation.

The rest of the paper is organized as follows. Section II describes our mutual gaze convergence detection model and the evaluation method. Section III explains experimental design to address these questions. Experimental results will be presented in section IV and discussed in section V. We finally will conclude this paper in section VI.

II. Detection Method and Evaluation

A. *Types of gaze behavior*

There are two types of gaze behaviors in our study: mutual gaze convergence and non-mutual gaze convergence. Fig. 1 shows all possible gaze scenarios in our study. We used two web cameras and two glasses cameras as a means to capture gaze directions of both lecturers and students.

The behaviors that we take into account as mutual gaze convergence can be divided into two categories: full and partial. Full mutual gaze convergence occurs when the web cameras detect that both individuals are looking straight at each other while the glasses cameras detect the Partner Face, which is the partner's face on the opposite side, at the middle of the frame. In the case of partial gaze convergence, the web cameras see that both individuals' gaze direction is toward their partner. However, they might tilt their head either aside, up, or down causing the glasses camera to fail to capture the Partner Face on the opposite side at the middle of the frame. In other words, mutual gaze convergence occurs when two individuals' gaze directions point towards each other.

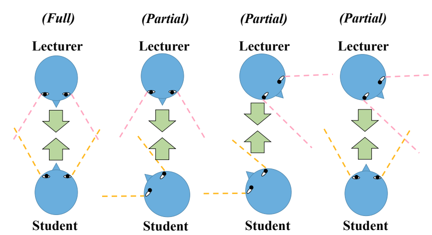
We consider both-side averted and one-side averted, including whether the head is tilting too high or too low vertically as non-mutual gaze convergence. In these cases, the web camera cannot capture the looking straight behavior regardless to the Partner Face detection from their glasses cameras.

[1] eyeLike, a webcam based gaze tracking framework can be found at https://github.com/trishume/eyeLike.git

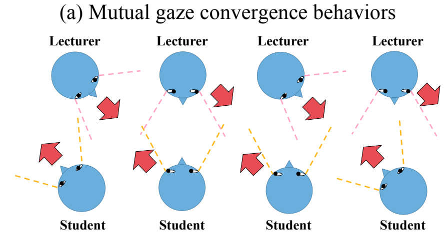
B. *Mutual gaze convergence detection model*

Our model for gaze convergence detection consists of five steps: eye detection, looking straight identification, Partner Face detection, gaze direction detection and mutual gaze convergence inference. The method for eye detection is as follows. First, we analyze with the web camera video files using eyeLike[1] framework (OpenCV C++) to perform eye detection. The eye detection process starts with face detection. *Face_cascade* function is employed for identifying facial features such as eyes and mouth, as shown in Fig. 2(a). Once the face is detected, the eye regions of both left and right eyes are extracted from the face-bounding box. Finally, eye center localization is performed using image gradients to locate the location in each region where most gradient vectors intersect (Fig. 2(b)) [14].

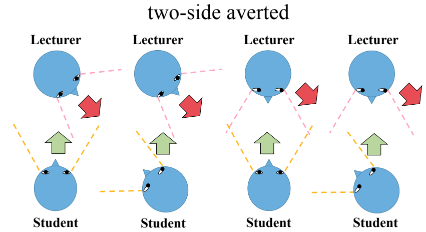
During the eye detection process, we also detect straight gaze detection for each eye in addition to eye center localization. To indicate straight gazes, we use adaptive threshold according to the detected eye region size to marginalize the effects of size difference and head movement.



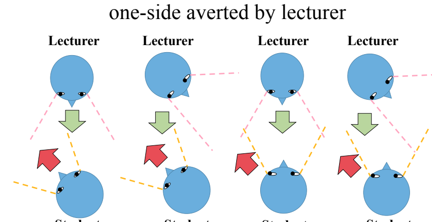
(a) Mutual gaze convergence behaviors



(b.1) Non-mutual gaze convergence behaviors: two-side averted

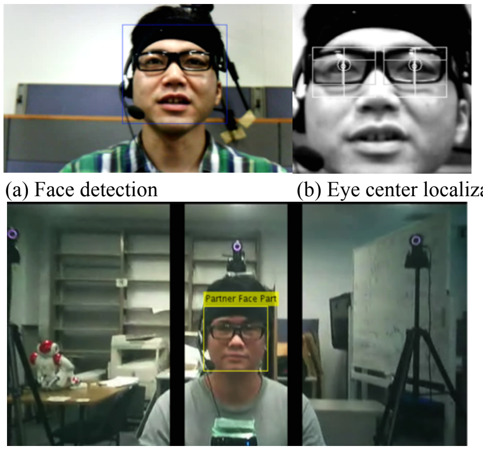


(b.2) Non-mutual gaze convergence behaviors: one-side averted by lecturer



(b.3) Non-mutual gaze convergence behaviors: one-side averted by student

Fig. 1 Types of gaze direction behavior



(a) Face detection (b) Eye center localization

(c) Partner Face detection

Fig. 2 Processes of mutual gaze convergence detection model

The adaptive threshold of looking straight was set by observing the average of straight gaze behavior from 4 sample participants from the video files compared with the detected eye region and eye center. We observed that there are 3 main ranges. If the detected eye region size is less than 80 pixels, the threshold of looking straight is set to be ± 8 pixels from the middle point of the detected eye region. For detected eye region size that larger than 80 but less than or equal to 90 pixels, and more than 90 pixels, the threshold is set to be ±12, and ±15 pixels, respectively. If the eye center localization is detected within the threshold, we assume that the participant is looking straight in that particular frame. The eye center localization is performed for both eye regions. The result of left and right straight gaze detection is the final product of eye detection process: looking straight or look elsewhere.

Consequently, we infer the looking-straight behavior of each individual from the straight gaze of his left and right eyes. Though ordinary human eyes have a line of sight symmetry, we need two separate eyes comparison to compensate the accuracy of the eye detection result. This process compares left and right eyes' straight gaze detection results. If one side of the eyes is detected as looking straight, we consider that the participant is looking straight in that particular frame; otherwise, he is looking elsewhere. Finally, we converted looking straight identification result from 30 frames to one-second unit by assigning the most occurrence looking straight identification result from each 30 frames as the looking straight identification result for each second.

For glasses camera files, we execute the Partner Face detection process using MATLAB to detect if the Partner Face is located within the threshold of the middle of the video frame. *Vision.CascadeObjectDetector* function in MATLAB is applied as a tool to detect the Partner Face. The ±100 pixels threshold from the middle of the frame is set as we observed that, in average when the participant is looking straight to the partner, the Partner Face is detected not exceeding ±100 pixels from the middle of the frame (Fig. 2(c)). We can affirm the threshold of every frame midpoint by having the same frame size and control seating position of both participants. If the detected Partner Face is within the threshold of that particular frame, we assume that the participant is turning his head directly toward his partner. Finally, 30 frames to one-second conversion was done similarly to the process used for web camera video files.

We process video files for both lecturers and students. After we have both results from looking straight identification and Partner Face detection of each individual, we continue to the process of gaze direction detection. This process compares the looking straight identification result and the Partner Face detection result of each individual. If both data are positive, we consider that the participant is now having the straight gaze direction at his partner in that particular second.

Once we have the gaze direction detection result of both lecturer and student, mutual gaze convergence inference process is executed by comparing the gaze direction results of the lecturer and the student. If both gaze directions indicate that they are both having the straight gaze direction at each other in a particular second, we assume this behavior as mutual gaze convergence. Otherwise, we consider it as non-mutual gaze convergence.

C. *Evaluation*

In this study, the measurement is based on the items from the paper-based test. The 10-question test evinces the success of the lecture task in each trail by considering it as a success if the student can correctly answer all of the questions, or failure, otherwise.

III. EXPERIMENTAL DESIGN

A. *Participants*

There were 12 participants in this experiments. All participants are native Japanese male students from Tokyo Institute of Technology with ages ranging from 21 to 47 years old. We grouped the participants into 6 pairs, each consisting of one lecturer, and one student. The protocols and procedures used in this experiment were approved by the Ethical Review Board for Epidemiological Studies of the Tokyo Institute of Technology.

B. *Task*

Our study focuses on one-to-one interaction between a human lecturer and a human student in a 5-minute lecture task. In the experiment, the lecturer teaches the student using a prepared article. We provided the article to the lecturer participants 2-3 days prior to the experiment date and asked them to study the article and practice teaching since we will evaluate their understanding on the article and affirm that they will be able to give a thorough lecture in lecturer training session before the actual lecture session.

C. *Content*

The lecture content is an article entitled "Naps Clear Brain's Inbox, Improve Learning" in Japanese version taken from the National Geographic's site. The article describes how sleeping and napping improve learning, especially in memorization and information recall.

D. Setup

The experimental setup is located in an environment similar to an office room in order to resemble real-world scenarios. We suggested the participant to focus on the partner during the interaction. Lecturers and students were seated in front of each other with a table between them, 1.5 meters apart from each other. Two web cameras was set to be 50 cm apart from the participants in order to capture the whole face of the participants regardless to their head tilting. Both participants had to wear a glasses camera to capture the Partner Face during the lecture task, as illustrated in Fig. 3. The specification of the cameras used in the experiment are listed in TABLE I.

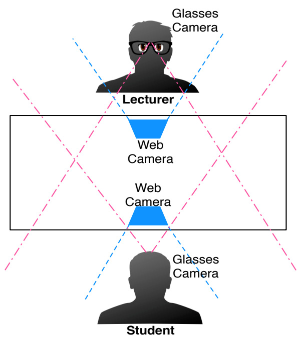


Fig. 3 Experimental Setup

TABLE I
EXPERIMENTAL TOOLS

Specification	Web Camera	Wearable Glasses Camera
Model	SANWA Supply CMS-V35BK, Japan	SPYDER E231, Japan
Resolution	2048x1536 pixels	1280x720 pixels
Frame Rate	30 frames per second	30 frames per second

E. Procedure

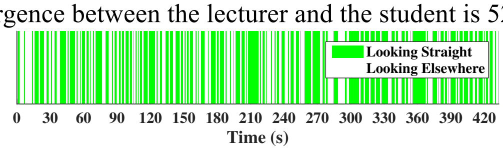
The experiment consists of three main sessions: lecturer training, lecture session, and evaluation. Prior to the lecture session, we assessed the lecturer by having the lecturer to practice his teaching with the experimenter and answer a 10-question test to ensure that the lecturer understood the lecture article and was capable of giving a smooth and thorough lecture. At the end of the lecture session, the student was evaluated with a different 10-question test for his understanding and success of learning.

IV. RESULTS

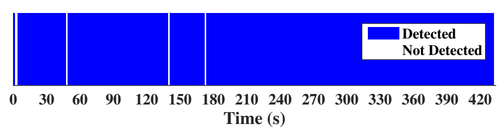
Data from 12 participants, 6 pairs in total was gathered. The gaze direction behaviors were inferred from the video recordings of the participant's interaction during the experiment session.

Fig. 4 shows an example of data retrieved from video processing. In Fig. 4(a), the example result of looking straight identification process (web camera video file analysis) is shown. The green bars indicate detected looking-straight behaviors of the participant in a particular second, whereas the blank spaces indicate looking elsewhere behaviors. Fig. 4(b) illustrates an example result from his Partner Face detection process (glasses camera video file analysis). The blue bars indicate the particular second that the Partner Face is detected in the range of the frame midpoint threshold, implying that the participant head is facing toward the partner on the opposite side. Fig. 4(c) shows detected mutual gaze convergence inferred from the comparison of the lecturer and the student's gaze direction results. The green bars display the particular second that both participants have straight gaze direction. In other words, they are looking at each other, performing mutual gaze convergence. The possible scenarios for mutual gaze convergence are represented in Fig.1.a. The blank spaces, on the other hand, show the non-mutual gaze convergence. This scenario can occur in three possible cases: both-side averted (Fig.1.b.1), one-side averted by the lecturer (Fig.1.b.2), and one-side averted by the student (Fig. 1.b.3).

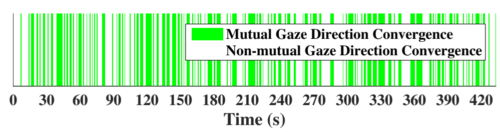
From the video analysis of all 6 pairs, we calculate the percentage of total mutual gaze convergence between the lecturers and the students. The percentage represents a relative value that compares the analysis results across all pairs to normalize the difference in time duration of the lecture task in each trial. In this short lecture task, we discovered that the average percentage of total mutual gaze convergence between the lecturer and the student is 52.83.



(a) Looking Straight Identification Result (Web Camera)



(b) Partner Face Detection Result (Glasses Camera)



(c) Mutual Gaze Convergence Detection Result
(Lecturer-Student Gaze Direction Comparison)

Fig. 4 An Example of Analysis Results

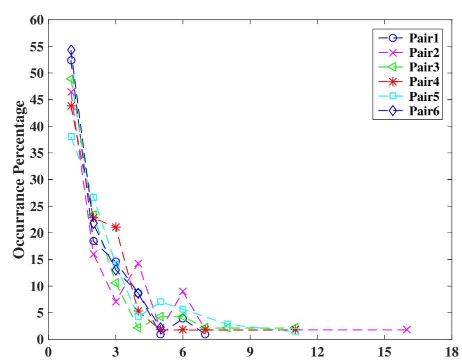


Fig. 5 The Cumulative Distribution of Mutual Gaze Convergence Duration in Percentage

The cumulative distribution of mutual gaze convergence during the short lecture task is shown in Fig. 5. It is conspicuous that for every pair, the mutual gaze convergence is characterized as momentary and recurrent, as supported by the duration-occurrence percentage relationship.

As regards to the success of the task, 2 of the 6 pairs failed in the lecture task since both pairs did not achieve full score on the 10-question test. Furthermore, we also conducted a T-test analysis on the straight gaze direction percentage of all 6 pairs. According to the T-test results, the students' straight gaze direction percentage is significantly higher than the lecturers' ($t(6) = 3.186$, $p < 0.05$). The conclusion can be drawn from the T-test result that the students have more straight gaze direction percentage than the lecturers. In other words, the students were looking at the lecturers more frequently than the other way around during the lecture task.

V. Discussion and Future Works

A. Mutual Gaze Convergence Detection Model Advantages and Limitations

As we conducted the experiment and obtained the detection results, we learned some advantages and limitations of our proposed model. The good points of this model are that the implementation cost is low and easy to use. In addition, it is efficient enough to detect the gaze direction and does not overcomplicate the solution. However, the limitations are that the lighting needs to be controlled and data processing cannot be done in real-time. TABLE II summarizes the advantages and limitations of our model comparing with other existing eye-tracking devices in general[2].

TABLE II
Performance Comparison

Performance	Our Model	Other Eye-tracking Devices
Implementation cost	Low	High to very high
Focus	Gaze direction	Gaze point and intensity
Time unit	Second	Millisecond
Processing	Post processing	Real-time processing
Lighting	Affected by different lighting	Robust with different lighting
Ease of use	Easy to apply and user-friendly tools	Require complex technical skills and tools

B. *Average percentage of* total *mutual gaze convergence between lecturers and students in short lecture task*

Previous studies raised the question of how long lecturers and students pay attention to each other during a lecture. Our study tries to answer this question by measuring the average percentage of total mutual gaze convergence between lecturers and students during the lecture tasks. The analysis result shows the average of 52.83% chance where mutual gaze convergence occurs between both participants during the lecture task. We further compared the percentage of total mutual gaze convergence with the evaluation on the task success. The results show that two pairs who failed the evaluation are also the pairs who have total mutual gaze convergence lower than 50%. According to this comparison, it is within reason to conclude that the more percentage of total mutual gaze convergence, the more mutual attentiveness between the lecturers and the students. It implies that increased mutual attentiveness and interaction to each other can lead to a higher possibility of achieving success in learning.

C. Question 2: Duration of mutual gaze convergence during the short lecture task

From the results, we also discovered that the participants have momentary mutual gaze convergence toward each other. We also found that the participants often lost their fixation at a particular direction, not only in long but also in short lectures. As we further investigated the video files, we found that the participants mostly lost their fixation at a particular direction when they had thinking process: they were either nodding or trying to recall or understand the lecture content. These led to the loss of straight gaze toward their partner, causing either one-side averted or both-side averted scenarios.

D. Question 3: Who dominate the amount of mutual gaze convergence in short lecture task

In our study, we observed that in 5 out of 6 pairs, the lecturers were the ones who have less straight gaze towards the students. According to the T-test result, there is significant difference between the percentage of straight gaze direction of lecturers and students. We further investigated into the possible reasons why the lecturers have less gaze towards the students during the lecture. We learned that during the lecture task, the lecturers were the ones who spent time trying to recall the information of the lecture to teach the students. Therefore, they varied their gaze more frequently than the students did. The analysis results also show that the students were the ones who fixed their gaze toward the lecturer most of the time during the lecture, and the mutual gaze convergence is formed once the lecturers had their gaze toward the students. According to Broz et al. [8] study, they revealed that low gazers, who look at their partner less, is the dominant and control the amount of mutual gaze. We can conclude from our study that the lecturer is the dominant of mutual gaze convergence in short lecture task.

E. Question 4: Effects of attention span toward short lecture task

To address this question, we compared our results with the previous studies. Many studies revealed that the attention span duration of ordinary human is 10-20 minutes. Middendorf [11] and Souna [12] state that the lecturer might see the beginning effects of attention decline after 10 minutes of lecturing. According to these studies, it can be implied that student is the dominant influence of the interaction in long lecture task. They control the amount of mutual gaze convergence duration during the interaction by returning their gaze to the lecturer and avert it when they lose attention or feel uncomfortable.

Nevertheless, in our study on short lecture tasks, our results contradict the claims of the previous studies conducted in

[2] Tobii: http://www.tobii.com and RightEye: http://www.righteye.com

longer lectures. We discovered that in short lecture task, the lecturers dominate the mutual gaze convergence instead of students. Because we conducted 5-minute lecture task, the duration of our lecture task is less than the attention span. Based on these findings, we can infer that the students still concentrated and focused within the attention span, while the lecturers instead shifted their focus to recalling the content of the lecture. As for long lectures, the duration exceeds the attention span, causing the students to easily lose their attention. The percentage of losing attention by the students can overcome the percentage of losing attention by the lecturers. Therefore, it shows that the time duration definitely affects the mutual gaze convergence and mutual attentiveness, especially, when it exceeds the attention span.

The analysis results from this study can be further applied to human-robot interaction. The average behavior of human-human interaction in our study indicates some natural behaviors in the lecture tasks. Instead of having robots always fix their gaze toward human it is interacting with, possibly creating a sense of artificialness and unease, the robot can learn from human and adapt its gaze fixation to achieve more natural behavior. Teaching assistant robot, tutor robot [15], and interaction robot [16], [17] researches, for instance, can benefit from this study.

At the current state, our study focuses on one-to-one interaction. Subsequent investigation on one-to-many is necessary to approach the actual lecture scenario where there are many students vs. one lecturer. Additionally, only Japanese male participants were observed. Gender and cultural differences can therefore be further investigated to increase diversity and obtain an insightful understanding of possible effects that is applicable to the robot behaviors, making it be able to adapt to any individual it interacts with.

VI. Conclusion

We have demonstrated the use of straightforward techniques and economical devices (i.e. web camera and glasses cameras) for detecting the mutual gaze convergence between lecturers and students. We implemented an alternative model using the eyeLike framework and image processing functions in MATLAB. We also discussed the natural behaviors of human-human interaction for lecturer-student interaction in short lecture tasks.

The main contributions of this study are as follows:

- We presented an alternative model of detecting mutual gaze convergence that can be executed using the simple devices and techniques instead of exorbitant eye-tracking devices.
- We revealed the fact that humans have momentary gaze fixation at a particular direction both in short and long lecture tasks. We also discovered that there is a time duration effect on who will dominate the amount of mutual gaze convergence in a lecture task. In short lecture task (5 minutes), which is less than the attention span (10-20 minutes), the lecturers dominate the amount of mutual gaze convergence. In contrast, the students are dominant in the tasks that are longer than the attention span.

Ultimately, our developed model implementation is based on readily available products that are easily accessible and applicable. The work is shown to be practical and worthy of contributing to the related research area communities.

Acknowledgment

This study was supported by Center of Innovation Science and Technology based Radical Innovation and Entrepreneurship Program Grant Number JY250158, and Japan Society for the Promotion of Science (KAKENHI) on Scientific Research on Innovative Areas Program Grant Number 26560114 and Scientific Research (A) Program Grant Number 15H01771.

References

[1] D. C. Richardsona, and R. Daleb (2005). Looking To Understand: The Coupling Between Speakers' and Listeners' Eye Movements and its Relationship to Discourse Comprehension. *Cognitive Science*, *29*, 39-54.

[2] J. Henderson, and F. Ferreira (2013). The interface of language, vision, and action: Eye movements and the visual world. Psychology Press.

[3] M. A. Just, and P. A. Carpenter (1976). Eye fixations and cognitive processes.*Cognitive psychology*, *8*(4), 441-480.

[4] M. Argyle, and M. Cook (1976). Gaze and mutual gaze. *Cambridge University Press*.

[5] J. Kellerman, J. Lewis, and J.D. Laird (1989). Looking and loving: The effects of mutual gaze on feelings of romantic love. *Journal of Research in Personality*, *23*, 145–161.

[6] C.L. Kleinke (1986). Gaze and eye contact: A research review. *Psychological Review*, *100*, 78–100.

[7] C. N. Macrae, B. M. Hood, A. B. Milne, A. C. Rowe, and M. F. Mason (2002). Are you looking at me? Eye gaze and person perception. *Psychological Science*, *13*(5), 460-464.

[8] F. Broz, H. Lehmann, C. L. Nehaniv, and K. Dautenhahn (2013). Automated analysis of mutual gaze in human conversational pairs. In *Eye Gaze in Intelligent User Interfaces* (pp. 41-60). Springer London.

[9] F. Broz, H. Lehmann, C. L. Nehaniv, and K. Dautenhahn (2012, September). Mutual gaze, personality, and familiarity: Dual eye-tracking during conversation. In *RO-MAN, 2012 IEEE* (pp. 858-864). IEEE.

[10] D. M. Bunce, E. A. Flens, and K. Y. Neiles, (2010). How long can students pay attention in class? A study of student attention decline using clickers. *Journal of Chemical Education*, *87*(12), 1438-1443.

[11] J. Middendorf, and A. Kalish, (1996, January). The "change-up" in lectures. In*Natl. Teach. Learn. Forum* (Vol. 5, No. 2, pp. 1-5).

[12] D. A. Sousa (2011). How the brain learns. Corwin Press.

[13] C. C. Bonwell, and J. A. Eison (1991). *Active Learning: Creating Excitement in the Classroom. 1991 ASHE-ERIC Higher Education Reports*. ERIC Clearinghouse on Higher Education, The George Washington University, One Dupont Circle, Suite 630, Washington, DC 20036-1183.

[14] F. Timm, and E. Barth (2011). Accurate Eye Centre Localisation by Means of Gradients. In *VISAPP* (pp. 125-130).

[15] A. L. Vollmer, K. S. Lohan, K. Fischer, Y. Nagai, K. Pitsch, J. Fritsch, and B. Wredek (2009, June). People modify their tutoring behavior in robot-directed interaction for action learning. In *Development and Learning, 2009. ICDL 2009. IEEE 8th International Conference on* (pp. 1-6). IEEE.

[16] N. Wang, and J. Gratch (2010, April). Don't just stare at me!. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1241-1250). ACM.

[17] L. Cavedon, C. Kroos, D. Herath, D. Burnham, L. Bishop, Y. Leung, and C. J. Stevens (2015). "C′ Mon dude!": Users adapt their behaviour to a robotic agent with an attention model. *International Journal of Human-Computer Studies*, *80*, 14-23.