Combining Multiple Types of Eye-gaze Information to Predict User's Conversational Engagement

Ryo Ishii*, †, Yuta Shinohara††, Yukiko, I. Nakano††, Toyoaki Nishida*

* Dept. of Intelligence Science and Technology
Kyoto University
Kyoto-Shi 606-8501, Kyoto, Japan
{ishii, nishida}@i.kyoto-u.ac.jp

† NTT Cyber Space Laboratories
NTT Corporation
Yokosuka-Shi 239-0847, Kanagawa, Japan
ishii.ryo@lab.ntt.co.jp

†† Dept. of Computer and Information Science
Seikei University
Musashino-shi 180-8633, Tokyo, Japan
y.nakano@st.seikei.ac.jp

ABSTRACT
In face-to-face conversations, speakers are continuously checking whether the listener is engaged in the conversation by monitoring the partner's eye-gaze behaviors. The goal of this study is to build an intelligent conversational agent that can recognize user's engagement from multiple types of eye-gaze information. In our previous work, we developed a method of estimating the user's conversational engagement using user's gaze transition patterns. However, it was not accurate enough. In this study, we added new variables in our gaze analysis: occurrence of mutual gaze, gaze duration, and distance of eye movement. Then, based on the results of the analysis, we used the variables as estimation parameters, and set up prediction models which consist of different combinations of parameters. To test which parameters are effective, the performance of these models were compared. As the result, it was revealed that a model using the gaze transition patterns, occurrence of mutual gaze, gaze duration, distance of eye movement, and pupil size as prediction parameters performed the best and was able to predict the user's conversational engagement quite well.

Author Keywords
Empirical study, eye-gaze behavior, conversational engagement, Wizard-of-Oz experiment

ACM Classification Keywords
H5.2. Information System: User Interfaces

General Terms
Algorithms, Design, Human Factors

INTRODUCTION
In face-to-face conversations, not only the speaker presents communicative signals accompanying speech, but also the listener display nonverbal signals, such as eye-gaze and head nods as feedback to the speaker. Argyle & Cook [1] showed that a listener's eye-contact expresses his/her attention toward the conversation.

We expect that exploiting eye-gaze information in designing user interfaces improves the naturalness of human-computer interaction specifically in conversational agent systems. If the system can judge the user's engagement in the conversation, it becomes possible to change the agent's actions and responses according to the user's attitudes. However, there has not been enough research in mechanisms of judging the user's engagement and implementation of such mechanisms into fully automatic conversational agents.

In our previous work [2-4], we demonstrated the close connection between user's conversational engagement and gaze transition patterns: patterns of shifting her/his attention. Then, we developed a method of estimating user's conversational engagement using the gaze transition patterns. Then, we implemented the method in a dialogue system. As the results of evaluation experiment, we found that in interacting with the implemented system, users felt that the agent's conversational functions were improved. However, we still have not been satisfied with the estimation accuracy.

In this study, to improve the engagement estimation accuracy, we re-analyze the user's gaze behaviors by adding new variables: occurrence of mutual gaze, gaze duration, distance of eye movement, and pupil size. For this purpose, this paper addresses the following issues:

- Analyze human-agent conversation corpus to investigate whether user's engagement is correlated with gaze transition patterns, mutual gaze, gaze duration, distance of eye movement, and pupil size.
- Evaluate the estimation methods and compare which combination of the parameters performs the best.

In the following sections, first we discuss related work, and then explain our collected data in an experimental setting. Based on the analysis of the data, our new engagement estimation methods will be proposed. Finally, the methods will be evaluated and compared, and then we conclude our empirical study.

2nd Workshop on Eye Gaze in Intelligent Human Machine Interaction.
February 13, 2011, Palo Alto, California, USA.

RELATED WORK

In communication science and psychology, many studies investigated functions of eye-gaze in face-to-face communication. Kendon [5] very precisely observed eye-gaze behaviors by employing the ethnomethodological method and discussed various kinds of eye-gaze functions. Psychological studies reported that eye gazing, specifically accompanied by head nods, serves as positive feedback to the speaker, and demonstrates that the listener is paying attention to the conversation [1, 6]. This kind of mutual gaze also contributes to smooth turn-taking [7]. On the contrary, when conversational participants share the same physical environment and their task requires complex reference to, and joint manipulation of physical objects, joint attention between the participants is a positive signal of conversational engagement [8-10].

These findings were later used as the basis of conversational humanoids. Nakano et al [11] proposed a gaze model for nonverbal grounding in ECAs that judges whether the user understands what the agent said or not. They used this model to control human-agent conversations. In human-robot communication research, Sidner et al [12] proposed a gaze model for conversational engagement, and implemented a model using a head tracker to judge whether the user is engaged in the conversation or not from the user's head movements.

In addition to the gaze direction, some other information obtained from "eye(s)" is also useful in HCI. Qvarfordt et al [13] developed an interactive system, iTourist, for city trip planning, which encapsulated knowledge of eye-gaze patterns and duration gained from studies of human-human collaboration systems. User study results showed that it was possible to sense users' interest based on the combination of eye-gaze patterns and gaze duration. Iqbal et al. [14, 15] investigated the use of task-evoked pupillary response to provide a measure of mental workload for interactive tasks. Their results showed that a more difficult task demands longer processing time, induces higher subjective ratings of mental workload, and reliably evokes greater pupillary response at salient subtasks. Eichner et al [16] used eye-trackers in detecting an object which the user was interested in, and integrated an eye-tracker as a component of interactive interfaces.

Based on the discussion above, we believe that exploiting multiple types of information in eyes improves human-agent conversational engagement.

CORPUS DATA

Wizard-of-Oz experiment in human-agent conversation

We conducted a Wizard-of-Oz experiment where the agents' speech and behavior were controlled by an operator [3]. In the experiment, two subjects participated in each session. One of the subjects participated as a user (called the "user"), and the other subject as an observer (called the "observer"). The user listened to the agent's explanation lasted for about 3 to 5 minutes for each 6 cell phone (see Figure 1). The user can communicate with the agent using speech. A push-button device was given to both the user and the observer. The user was instructed to press the button if the agent's explanation was boring and the user would like to change the topic. The observer was instructed to press the button when the user looked bored and distracted from the conversation. In this study, these button pressing behaviors were used as the human judgment of user's conversational engagement.

Collected corpus

We collected 10 conversations whose average length was 16 minutes, and built a multimodal corpus consisting of verbal and nonverbal behaviors shown below;

- Verbal data: The user's speech was transcribed from the recorded speech audio, and the agent's utterances were extracted from the log of the Wizard-of-Oz system. The total number of the agent's utterances was 951 and that of the user's was 61.
- Nonverbal data: Since the agent's behaviors were pre-set in the Wizard-of-Oz system, the agent's gestures and gaze behaviors were able to be extracted from the system log. As for the user's nonverbal behaviors, we collected user's gaze data using Tobii X-50 eye-tracker in 50Hz.
- Human judgment of engagement: When the user and/or the observer pressed her/his button, lights went on in another room, and these lights were recorded as video data.

All these data were integrated in xml format, and can be visually displayed using the Anvil annotation tool [17].

ANALYSIS

As a preliminary analysis, we examine four factors in eye gaze: occurrence of mutual gaze, gaze duration, distance of eye movement, pupil size. Then, we will investigate the correlation between these factors and human judgments of engagement.

We regarded the user as disengaged from the conversation when either the user or the observer pressed the button. We

Figure 1. Conversational agent projected on a screen

found some cases that the observer pressed the button slightly after the user displayed their disengagement gaze behaviors. So, we annotated the user's disengagement time period from 10 second before the observer's pressing the button to the time that s/he released it.

In following sections, we analyze a correlation between gaze transition patterns and human judgment of user's engagement. Then, we investigate other types of eye-gaze information (i.e. mutual gaze, gaze duration, distance of eye movement, and pupil size).

Analysis of Gaze 3-gram

Defining 3-gram patterns

By following our previous study [2-4], gaze transition patterns were defined as gaze 3-gram: three consecutive gaze behaviors. The constituents of gaze 3-gram are the following three types of eye gaze.

- *T*: look at the target object of the agent's explanation, which is the discourse focus.
- *A*: look at the agent's head
- *F*: look at non-target objects, such as other cell phones and an advertisement poster ($F1 \neq F2 \neq F3$).

Since the agent was designed to look at the current target object most of the time, it is presumed that joint attention is established between the user and the agent when the user's gaze label is *T*.

Since the eye-tracker cannot measure the pupils movement during blinks, small blanks often occurred in gaze data. So, we counted two consecutive gaze data as one continuous eye-gaze if the same object was continuously looked at in both data and the data blank was less than 200 msec. On the contrary, if the focus of the attention changed after the data blank or the blank was longer than 200msec, these two gaze data were not merged. Suppose that the user's gaze direction shifts in the following order: *T*, (100msec blank), *A*, (50msec blank), *A*, (150msec blank), and *F1*. There is a 50msec data blank between two consecutive *A*s. Thus, these two data are merged into one block. As the result, a 3-gram constructed from this sequence is *T-A-F1*. If the data loss is longer than 1sec, such data was discarded as an incomplete 3-gram.

Correlation between 3-gram patterns and engagement

The results of the 3-gram analysis in 8 users are shown in Figure 2. In this graph, the probability of 60% means that this pattern co-occurs with the human disengagement judgment (i.e. overlap with the time period of pressing the buttons plus ten seconds before this) 60% of the time. For each 3-gram pattern, we calculated the probability of co-occurrence with the disengaging judgment.

As shown in Figure 2, the probability of disengagement judgment is different depending on the types of 3-gram. For example, *A-F1-F2* has the highest probability over 60%. This means that the user was judged as disengaged over 60% of the time when this pattern occurred. On the other hand, the probability for the *A-F1-T* 3-gram is only 13.3 %. These results suggest that the 3-grams with higher probability violate proper engagement gaze rules, and those with lower probability contribute to conversational engagement.

Analysis of mutual gaze

In face-to-face conversations, it is obvious that mutual gaze serves as feedback to the speaker. Mutual gaze gives a good opportunity for the listener to give feedback to the speaker, and for the speaker to receive the feedback and determine or perform conversational actions. Therefore, focusing on mutual gaze, we analyze correlation between mutual gaze and user's engagement.

To distinguish mutual gaze from use's gaze at the agent, the gaze category *A* (looking at the agent) was subcategorized into the following two labels: *M* (mutual gaze; the user establishes eye contact with the agent) and *A* (non-mutual gaze). Then, the probability of co-occurrence with disengagement judgment was calculated.

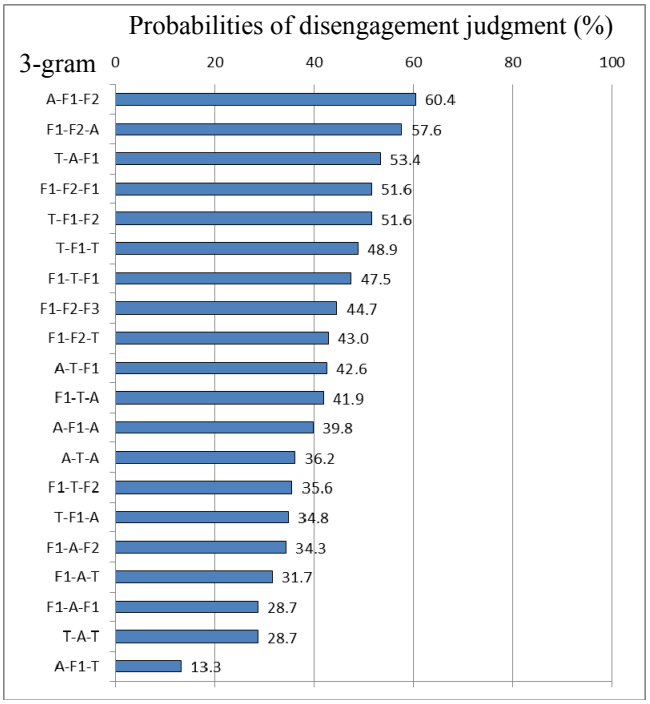

Figure 2. Eye-gaze 3-grams and probabilities of disengagement judgment

As shown in Figure 3, the new 3-grams with *M* label are indicated with red diagonal line's box. Some of these 3-grams have very high or low probability of disengagement judgment such as 100% or 0%. For example, *A-F1-A* 3-gram has the probability of 39.8% in Figure 2. By distinguishing mutual gaze from gazing at the agent, *A-F1-A* 3-gram is divided into four types of 3-grams, such as *A-F1-A*, *M-F1-M*, *M-F1-A*, and *A-F1-M* whose probabilities of disengagement judgment are 37.1%, 59.7%, 9.8%, and 100% respectively in Figure 3. Thus, by considering mutual gaze, the correlation between 3-grams and dis/engagement judgment becomes clearer. This result suggests that considering mutual gaze in categorizing gaze behaviors may improve the accuracy of estimating the conversational engagement.

Analysis of gaze duration

Our previous work [3] analyzed correlation between gaze duration and engagement, and suggested that looking at the target object and looking at the agent would be the positive signs of user's engagement. On the other hand, looking at non-target objects for a long time signals the user's disengagement from the conversation.

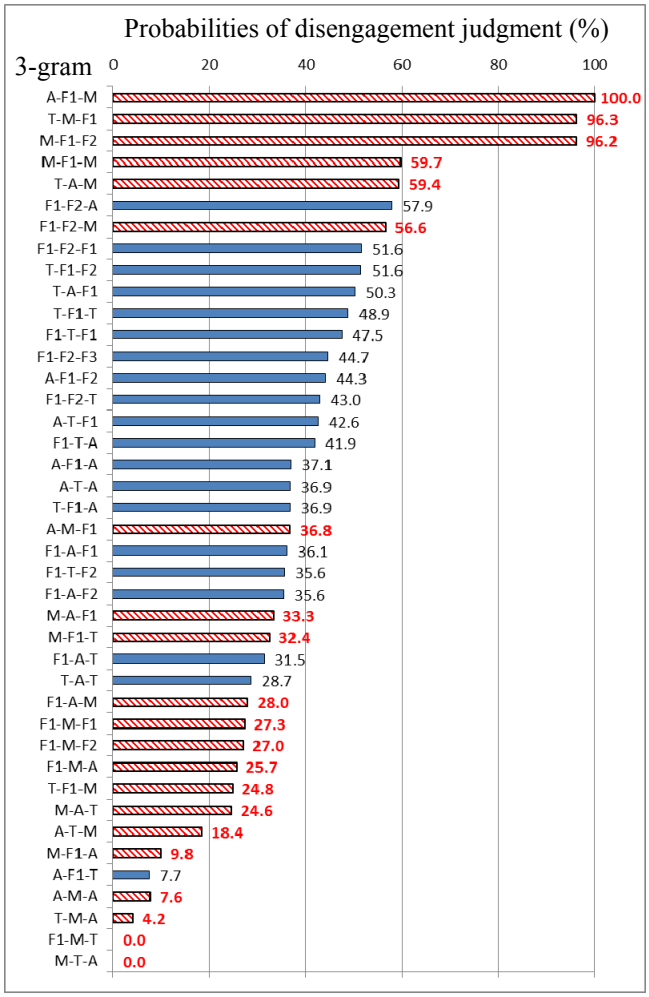

Figure 3. Probabilities of user's button pressed in mutual gaze or not situation

Based on these results, in this study, we considered the gaze duration in categorizing 3-grams. In the current analysis, we only considered the duration of the third constituent of a given 3-gram. We will extend this analysis by considering the duration for all the constituents in a 3-gram. According to the length of the third constituent, we classified 3-grams into long 3-grams and short ones. When the duration of the third constituent of a given 3-gram is shorter than the threshold, we labeled it with an extension "*_S*" at the end of the label. When the duration of the third constituent is longer than the threshold, we labeled it with an extension "*_L*". The thresholds were set at the values that clearly distinguish engagement from disengagement. For example, in *T-A-F1*, the probability of disengagement judgment was 0% when the duration of *F1* was 0 to 34 msec. The probability was 100% when the duration *F1* was 34 to 78 mesc. In this case, if the threshold is set at 34 msec, in long 3-grams (longer than 34 msec), the button pressing probability is 100%. By applying this threshold, we calculated the average probability of disengagement judgment for *T-F1-F_S* and *T-F1-F1_L*. Then, we found that the button pressing probability for *T-F1-F_S* (18.18%) was much lower than that for *T-F1-F_L* (100%).

Figure 4 shows the probability of co-occurrence between 3-grams with duration information and disengagement judgment. For example, *A-F1-F2* has the probability of 44.3% in Figure 3. By considering the duration information, a 3-gram of *A-F1-F2* is divided into *A-F1-F2_S* and *A-F1-F2_L* whose probabilities are 18.2% and 100% respectively. The new 3-grams containing *_S* or *_L* label at the end have very high or low probabilities of disengagement judgment such as 100% and 0% compared to the probabilities shown in Figure 3. Therefore, considering the duration in categorizing 3-grams is useful in distinguishing engagement patterns from disengagement ones. In particular, 3-grams whose probabilities are close to 100% or 0% almost always have "*_L*" label (indicated with box with orange diagonal lines). In other words, when the user looks at an object for a long time in the third constituent of a 3-gram, the user's engagement can be judged by looking at the preceding gaze behaviors (the first two gaze behaviors in a given 3-gram). Thus, these results suggest that considering gaze duration in categorizing gaze 3-grams may improve the accuracy of estimating conversational engagement.

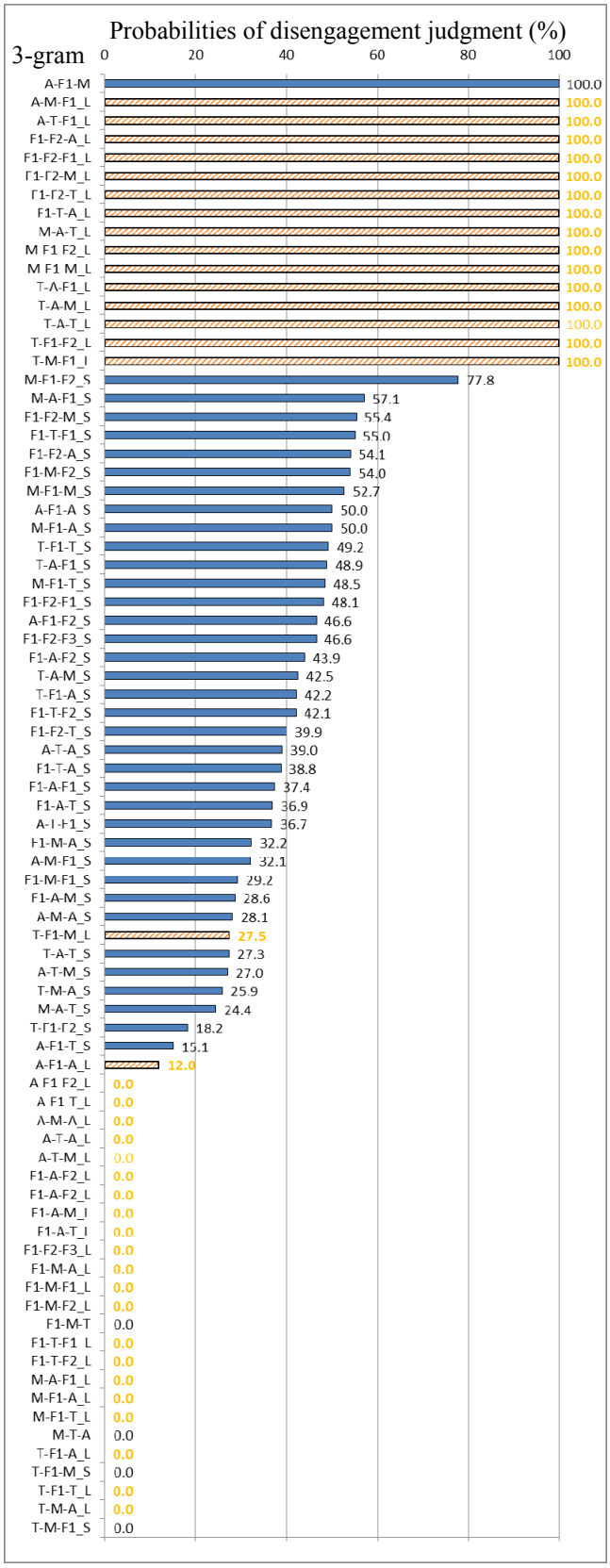

Figure 4. Eye-gaze 3-grams with duration and probabilities of button pressing

Analysis of distance of eye movement

We assume that user's conversational engagement is somehow related to the distance of eye movement during the conversation. For example, if the user is positively engaged in the conversation, distance of eye movement may be short because the user carefully looks at the object explained by the agent. On the other hand, if the user is not engaged, the distance of eye movement may become longer because the user looks at the objects which are not focused in the discourse, such as advertisement and cell phones currently not-explained. Therefore, we analyzed correlation between distance of eye movement and user's conversational engagement.

Figure 5 shows the moving average of 400 msec window for the distance of eye movement. The graph shows the proportion of occurrence of each distance value (this actually indicates the frequency of each distance value), the difference between engagement (dashed blue line) and disengagement (solid red line) is not very big, and average for each graph is 12.15 and 12.61 pixels, respectively. Although the peak of these graphs are not very different (Engagement: 5.25~5.50 pixels, Disengagement: 5.00~5.25 pixels), the distribution in disengagement situation is skewed to the right (the right tail is longer) compared to the distribution for engagement situation (the blue graph has larger values when the distance is 0 to 8 pixels). In other words, when the user is actively engaged in the conversation, the distance of eye movement is shorter than that in disengagement situation.

Thus, the results suggest that the distance of eye movement might be a weak predictor in estimating the conversational engagement.

Analysis of pupil size

It has already been known that pupil size becomes larger when people look at something interesting and exiting [18]. Therefore, it may be a reasonable assumption that the user's conversational engagement is somehow related to pupil size. For example, if the user is engaged in the conversation, the pupil size may become larger because the user carefully looks at the object explained by the agent. On the contrary, if the user is not engaged, the pupil size becomes smaller because the user doesn't seriously look at object. To examine this hypothesis, we analyzed the correlation between pupil size and user's conversational engagement.

Figure 6 shows the distribution of left eye pupil size data in engagement and disengagement situations. The x-axis indicates the pupil size, and the y-axis indicates the proportion of occurrence for each pupil size value. The average of pupil size for engagement is 5.20 cm and that is 5.16 cm for disengagement. Similar to the result of eye movement, although the averages of pupil size are not very different between engaged and disengaged situations, the distribution for engagement slightly shifts to the right compared to that for disengagement. As reported in previous studies, this result suggests that pupil size

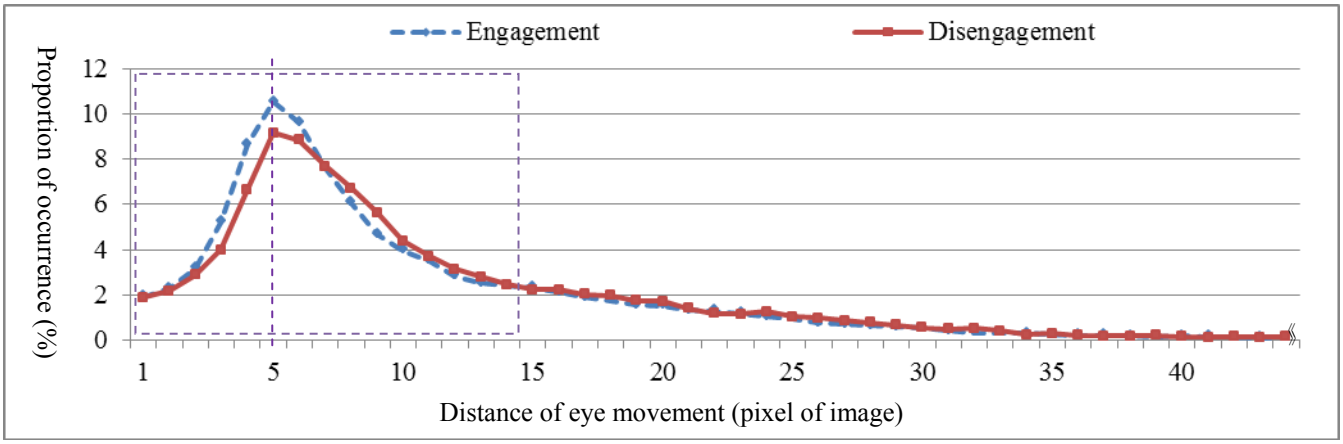

Figure 5. Distribution of distance of eye movement

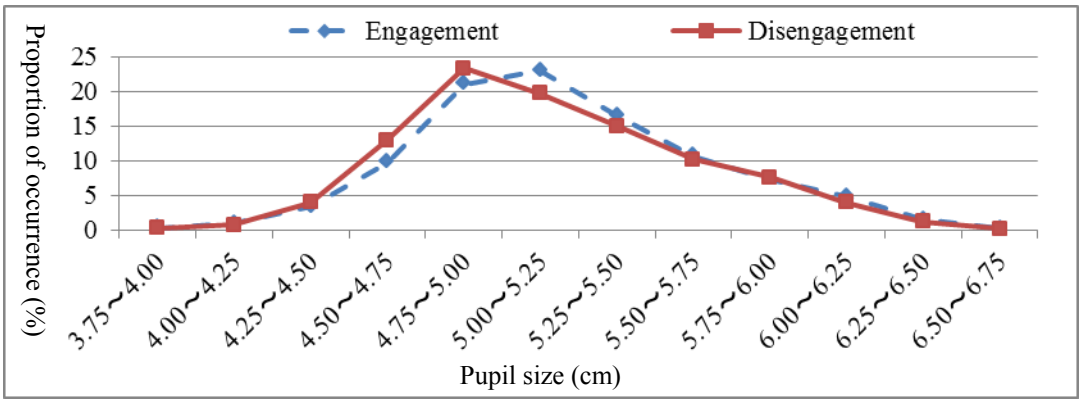

Figure 6. Distribution of pupil size

becomes larger when the user is engaged in the conversation and becomes smaller when the user is disengaged. Thus, pupil size may be a useful predictor of conversational engagement.

ESTIMATING USER'S ENGAGEMENT

Based on the analyses in previous sections, we found that 3-grams, mutual gaze, duration of eye-gaze, distance of eye movement, and pupil size may be useful as predictors of user's engagement in conversations. In this section, we estimate the user's engagement by employing SVM (Support Vector Machine) [19], and evaluate the accuracy and the effectiveness of each parameter. We used e1071 package[1] with R for SVM. The settings of SVM are RBF (radial basis function) kernel and default parameter of C (C = 1.0). The data used in SVM contains the user's conversational engagement state as a class, and the five eye-gaze parameters as features. Since we measured the user's behaviors 50 times per second (50Hz), the time resolution of the training data is also 50 Hz. The detailed explanation about each parameter is shown below;

- User's conversational engagement state: We have two classes in this model; engagement and disengagement. A disengagement state was a time period from ten seconds before the observer pressed button to when the button is released.
- 3-gram: The original types of gaze transition 3-gram
- 3-gram with mutual gaze: Distinguishing mutual gaze (*M* label) and gaze at the agent (*A* label) in constructing 3-grams.
- Duration of eye-gaze: The time duration from the start to the end of the third constituent of a given 3-gram.
- Distance of eye movement: The sum of the distance of eye movement for the last 400 msec.
- Pupil size: The average of both-eyes' pupil size data measured by an eye-tracker.

We tested the following estimation models that use these parameters above:

- 3-gram-based model (3-gram): Using original types of 3-gram without considering mutual gaze.

[1] http://cran.r-project.org/web/packages/e1071/

Model	State	Precision	Recall	F-mesure
3-gram	Engagement	**0.634**	**0.955**	**0.762**
	Disengagement	**0.532**	**0.085**	**0.147**
3-gram + M	Engagement	**0.645**	**0.950**	**0.768**
	Disengagement	**0.579**	**0.115**	**0.192**
3-gram + M + Dr	Engagement	**0.721**	**0.873**	**0.790**
	Disengagement	**0.658**	**0.419**	**0.512**
3-gram + M + Ds	Engagement	**0.674**	**0.877**	**0.762**
	Disengagement	**0.560**	**0.269**	**0.363**
3-gram + M + PS	Engagement	**0.664**	**0.937**	**0.777**
	Disengagement	**0.627**	**0.183**	**0.283**
ALL	Engagement	**0.769**	**0.918**	**0.837**
	Disengagement	**0.793**	**0.534**	**0.638**

Table 1. Results of evaluation

- 3-gram+M: 3-gram with considering mutual gaze.
- 3-gram+M+Dr: 3-gram with mutual gaze and the duration of eye-gaze.
- 3-gram+M+Ds: 3-gram with mutual gaze and the distance of eye movement.
- 3-gram+M+PS: 3-gram with mutual gaze and pupil size
- All parameters model (ALL): 3-gram with mutual gaze, the duration of eye-gaze, the distance of eye movement, and the pupil size.

The results of SVM are shown in Table 1. As the overall evaluation, F-measure of All parameters model (ALL) is 0.837, which is the best score among the other models. This result suggests that all the parameters contribute to estimate the user's conversational engagement. In comparing 3-gram and 3-gram+M, we did not find a big difference between them. In disengagement judgment, F-measure for 3-gram model is 0.147, and that for 3-gram+M is 0.192. This is because the agent rarely looks at the user in this corpus, and we did not collect enough number of 3-gram data accompanied by mutual gaze (indicated by *M* label). The performance of 3-gram+M+Dr is much better than 3-gram+M, specifically the performance of estimating the disengagement is improved (F-measure goes up from 0.192 to 0.512). This result suggests that duration of gaze is a strong predictor of user's engagement. In comparing a 3-gram+M and a 3-gram+M+Ds, the F-measure was 0.192 in 3-gram+M, and that was 0.363 in 3-gram+M+Ds. Moreover, in comparing 3-gram+M and 3-gram+M+PS, F-measure of 3-gram+M+PS (0.283) are better than that of 3-gram+M (0.192). This suggests that distance of eye movement and pupil size are useful in estimating the conversational engagement.

DISUCUSSION

The evaluation results showed that all the parameters proposed in this study are useful in estimating user's attitude towards the conversation. However, the parameters may need to be adjusted according to the user, situation, and conversational content or theme. For example, pupil size may change depending on the user's emotion and the brightness of visual stimuli. Individual difference is also not very small. Moreover, the distance of eye movement may different depending on the place and layout of visual stimuli. Therefore, to estimate the user's attitude more precisely, it is necessary to establish a situation adaptive model. On the contrary, 3-gram itself is not seriously affected by these environmental factors. It is suggested that because of such robustness, the performance of 3-gram+Dr model is better than 3-gram+Ds and 3-gram+PS model.

CONCLUSION

Aiming at estimating user's conversational engagement from eye-gaze information, this paper analyzed different kinds of gaze information: mutual gaze, gaze duration, distance of eye movement, and pupil size.

Based on the results of the analysis, we integrated all kinds of eye-gaze information investigated in our empirical study, and applied them to SVM to test whether each parameter contribute to the model or not. As the results of testing various combinations of these parameters, it was revealed that a model with all the parameters performs the best, and can predict the user's conversational engagement quite well.

We have already implemented a fully autonomous conversational agent that incorporates the user adaptive engagement estimation method. As future work, we will improve the estimation method by adding more information by other nonverbal behaviors, such as head movement, facial expression, and posture.

ACKNOWLEDGMENTS

We would like to express great thanks to Yoshihiko Suhara in NTT Cyber Solution Laboratories for this professional advice about SVM technique.

This work is partially funded by the Japan Society for the Promotion of Science (JSPS) under a Grant-in-Aid for Scientific Research in Priority Areas "i-explosion" (21013042), and a Grant-in-Aid for Scientific Research (S) (19100001).

REFERENCES

1. Argyle, M. and Cook, M., *Gaze and Mutual Gaze*. (1976), Cambridge: Cambridge University Press.
2. Ishii, R and Nakano, Y.I., *Estimating User's Conversational Engagement Based on Gaze Behaviors.* in *the 8th international conference on Intelligent Virtual Agents (IVA'08).* (2008): Springer. pp. 200-207.
3. Ishii, R. and Nakano, Y.I., *An Empirical Study of Eye-gaze Behaviors: Towards the Estimation of Conversational Engagement in Human-Agent Communication*. in 2010 International Conference on Intelligent User Interfaces (IUI2010), Workshop on Eye Gaze in Intelligent Human Machine Interaction. (2010).
4. Nakano, Y.I. and Ishii, R., *Estimating User's Engagement from Eye-gaze Behaviors in Human-Agent Conversations*. in 2010 International Conference on Intelligent User Interfaces (IUI2010). (2010).
5. Kendon, A., *Some Functions of Gaze Direction in Social Interaction.* Acta Psychologica, (1967). 26: pp. 22-63.
6. Clark, H.H., *Using Language*. (1996), Cambridge: Cambridge University Press.
7. Duncan, S., *Some signals and rules for taking speaking turns in conversations.* Journal of Personality and Social Psychology, (1972). 23(2): pp. 283-292.
8. Argyle, M. and Graham, J., *The Central Europe Experiment - looking at persons and looking at things.* Journal of Environmental Psychology and Nonverbal Behaviour, (1977). 1: pp. 6-16.
9. Anderson, A.H., et al., *The effects of face-to-face communication on the intelligibility of speech.* Perception and Psychophysics, (1997). 59: pp. 580-592.
10. Whittaker, S., *Theories and Methods in Mediated Communication*, in *The Handbook of Discourse Processes*, Graesser, A., Gernsbacher, M., and Goldman, S., Editors. (2003), Erlbaum, NJ. pp. 243-286.
11. Nakano, Y.I., et al. *Towards a Model of Face-to-Face Grounding*. in *the 41st Annual Meeting of the Association for Computational Linguistics (ACL03)*. (2003). Sapporo, Japan. pp. 553-561.
12. Sidner, C.L., et al., *Explorations in engagement for humans and robots.* Artificial Intelligence, (2005). 166(1-2): pp. 140-164.
13. Qvarfordt, P. and Zhai, S., *Conversing with the user based on eye-gaze patterns. In Proceedings of the SIGCHI Conference on Human Factors in Computing System.* CHI '05. ACM Press, New York, NY, 221-230, (2005).
14. Iqbal, S.T. Adamczyk , P.D. Zheng, X.S. and Bailey B.P., *Towards an Index of Opportunity: Understanding Changes in Mental Workload during Task Execution*. *Proc. CHI'05.* Portland, OR: ACM, 311-320, (2005).
15. Iqbal, S.T. Zheng, X.S. and Bailey, B.P., *Task-Evoked Pupillary Response to Mental Workload in Human-Computer Interaction*. *Proc. CHI'04.* Vienna: ACM, 1477-1480, (2004).
16. Eichner, T., et al. *Attentive Presentation Agents*. in *The 7th International Conference on Intelligent Virtual Agents (IVA)*. (2007). pp. 283-295.
17. Kipp, M. *Anvil - A Generic Annotation Tool for Multimodal Dialogue*. in *the 7th European Conference on Speech Communication and Technology*. (2001). pp.1367-1370.
18. Hess, E.H., *Attitude and Pupil Size.* Scientific American, (1965). 212: pp. 46-54.
19. Vapnik, V. *The Nature of Statistical Learning Theory*. Springer-Verlag, (1995).