Immersion Discriminated from Browsed Information in Writing Document Referring Web Pages

Takahisa Oe
Graduate School of Information Science and Engineering
Ritsumeikan University
Shiga, Japan
Email: oe2002 @de.is.ritsumei.ac.jp

Shinya Yonekura
Graduate School of Information Science and Engineering
Ritsumeikan University
Shiga, Japan
Email: shinya @de.is.ritsumei.ac.jp

Hiromitsu Shimakawa
College of Information Science and Engineering
Ritsumeikan University
Shiga, Japan
Email: simakawa @de.is.ritsumei.ac.jp

***Abstract*—This study proposes a method to distinguish student immersion in writing on computers without physical sensors. To improve the work efficiency of writing, students need support, such as warning when they have been distracted a long time from writing. A model using the Random Forest algorithm discriminates the immersion, examining windows of their operation target on the top of the display. In our experiment, the model discriminates the immersion of 5 subjects with the accuracy of 0.65 or higher in the F-measure, where placement of a specific window on the top of the screen turns out to be the most important feature. Various kinds of information is presented on the screen of the PCs of students. It includes not only information necessary for writing, but also entertainment information such as movies and games. The experiment result indicates that students tend to exclude entertainment information from their vision when they are under immersion in writing. It suggests that the student distraction from writing can be warned without extra effort from students, if we examine the top of the screen.**

Keywords–Sensorless detection; Immersion; Distraction; Machine learning.

I. INTRODUCTION

The popularization of the Internet has brought us the power of easy access to a wide variety of desired information. At the same time, we can also access information related to entertainment, which disturbs our concentration on works or activities. Students often write documents, such as technical reports and presentation materials using computers. They usually collect information necessary for the writing, using Web browsers. At the same time, their concentration is disturbed by Web-based information related to entertainment. A method to distinguish their immersion in the writing task from distraction is required to help students keep their concentration. For example, when a student has been distracted from writing for a long time, a self-management tool should send a warning to the student. Physical sensors are indispensable in existing methods to distinguish student immersion in their writing tasks [1]–[4]. However, it is not practical to use physical sensors for every writing task.

This study proposes a method to distinguish student immersion in writing, using only data which can be acquired without physical sensors. To achieve it, the method pays attention to the type of digital documents opened on the computers where the students engage in writing. Based on the bag-of-words algorithm, the proposed method figures out a document vector for each one of the documents used in the writing. They involve not only the target document, but also the ones accessed by Web browsers. We refer to the latter as browsed documents.

The association degree of browsed documents with the target one is calculated using the vectors. Browsed documents are categorized into 2 groups based on the degree of association. One group contains reference information, which is related to the target document. Students write a target document using information a from reference document. The other group contains supplemental information, which facilitates the writing of the target document, and entertainment information, which is unrelated to the writing. The contents of supplemental information are not directly related to the contents of the target document. However, they promote the writing, because they explain the knowledge to write comprehensive documents, such as how to organize technical documents in a general way, and how to write mathematical expressions in digital documents.

The proposed method distinguishes immersion of a student in writing by features of interactions the student is taking. The first feature is whether the student provides the computer with inputs, such as moving mouse cursor, clicking the mouse, or typing on the keyboard. In this feature, it is also important to know whether the inputs are used for the target document, the first group of browsed documents, or the second group. The second feature is related to windows handling documents. On the top of the display, the student places a window for either the target document, the first group of browsed documents, or the second group. The feature gives importance to what window is placed on the top of the display.

This study assumes students keep a specific behavior during their immersion. For example, they are typing characters in the target documents to write down their ideas. Others may stop the typing to consider their plans on target documents, or continue to search relevant information. This study examines if students engaging in writing keep one kind of behavior in terms of the above features of interaction during their

 ISBN: 978-1-61208-580-7

immersion in writing.

To distinguish between their immersion in the writing and their distraction from the writing with the method, we conducted an experiment where five students worked on a task to summarize a specific technology. The method discriminated student immersion with an accuracy exceeding 0.65. The method trained the discriminator taking student interaction with computers and the continuity of the specific behavior as their explanation variables. The method discriminates the immersion of students, using a machine learning method. The method trains a discriminator with data sampled from students. The data consists of student interaction with computers and the continuity of the specific behavior in the interaction. This paper discusses which explanation variables are effective for the discrimination from the viewpoints of machine learning.

In Section 2, we describe the necessity to determine the concentration in writing and existing methods to determine concentration. In Section 3, we describe how to determine concentration in writing without a physical sensor. In Section 4, we describe the procedure of the experiment and the experimental result. In Section 5, we discuss experimental results. In Section 5, we describe the conclusion and future research tasks.

II. The Judgement of Concentration of Students Attention on Writing

Here, we present the necessity to determine the concentration in writing and existing methods to evaluate concentration.

A. Problem of writing on PC

Nowadays, we can easily obtain the information we want using the Internet. At the same time, we can also access information related to entertainment, which disturbs our concentration on the work or activity. The survey of the Ministry of Internal Affairs and Communications in Japan says about 40% of college students are netaholic [5]. They are caught by Web-based videos, social media, or net games for an excessively long time. Coker's study proved that extreme Web-surfing during work time degrades the progress of the work [6]. The use of the Internet prevents them from concentrating on their tasks. Focusing their attention on specific target tasks, people would enter into a state where they can show the best performance on their tasks. In this study, we refer to this state as immersion [7][8]. On the contrary, we refer to states where people cannot concentrate on their target tasks as distraction.

Using computers, students make documents, such as technical reports and presentation materials. During the writing, their concentration is disturbed by Web-based information related to entertainment. Focus of student attention on tasks other than writing degrades the efficiency to finalize documents, because the students run into distraction from the writing. It is not desirable for students to run into the distraction from writing. Students struggle with entertainment information on the Internet, trying to concentrate their attention on the writing. If we can detect the distraction from the writing, we can avoid the inefficiency problem students suffer from. The distraction from writing corresponds to the immersion in information related to entertainment. It is difficult to distinguish which information the students focus their attention on. This paper addresses a method to distinguish student immersion in writing from the distraction.

B. Necessity to evaluate concentration

Students engage in writing when they make technical reports, presentation materials, and so on. This study assumes the following tasks as writing. Students edit documents on a computer, using text editors, authoring tools for presentation materials, and so on. Students collect information necessary for the writing, using Web browsers on a computer. In the following, we refer to students engaging in the writing as users. The users gather necessary information for the writing from the Web. The Internet is full of Web pages related to entertainment. Since they are attractive to the users, the users are likely to be prevented from concentrating on the writing. When the users are prevented from concentrating, the work efficiency decreases. They might use up the time for other tasks, such as extra-curricular activities.

Users should be supported so that they can immerse into writing. For example, when a user has been distracted from writing for a long time, a self-management tool should send a warning to the user. It is necessary to determine whether the users concentrate their attention on the writing or they are distracted from it. This study aims to increase the learning efficiency of students, having them concentrate on writing. Writing is a really common task for students. To get high practicability, it is preferable to minimize efforts and costs in the method to evaluate concentration. We should avoid any method which imposes extra effort and costs on students, such as wearing physical sensors, to determine the concentration of their attention on writing.

C. Existing methods

Sarrafzadeh et al. measured emotional states from bio-signals like facial expressions [9]. Jang et al. identified human intention by eyeball movement patterns and pupil size variation [10]. Kapoor et al. predicted students' quit puzzle by facial expressions, dermal activity, posture, and mouse pressure [11]. Jraidi et al. proposed a method to presume human immersion in a task using the skin conductance, the heart rate and the electroencephalography [1]. Nacke and Lindley also used physical sensors measuring from orbicularis oculi and zygomaticus major to know human immersion [2]. Leelasawassuk et al. tried to identify the target of attention using a Google Glass and an Eye Tracke [3]. Lee et al. measured emotional state from bio-signals like EGG signals [12]. They also measured the degree of human immersion from pupil movement and eye blinks using a webcam and an Eye Tracker. All of these methods need to use physical sensors [4]. Mello detected students who were bored, disengaged or zoning out by an eye tracker [13]. However, the necessity of physical sensors for every writing imposes extra efforts and costs on users. These methods should be avoided for the evaluation of immersion in writing. A method free from physical sensors is required to easily evaluate immersion into writing.

III. Concentration Judgment Using Behavior Logs

Here, we present how to evaluate concentration on writing without physical sensors.

A. Method overview

This study proposes a method to distinguish user immersion in writing without physical sensors. Figure 1 illustrates a use case of the method. The system gets a behavior log, which consists of applications and Web pages displayed on the screen, as well as inputs from the mouse and the keyboard along with their targets.

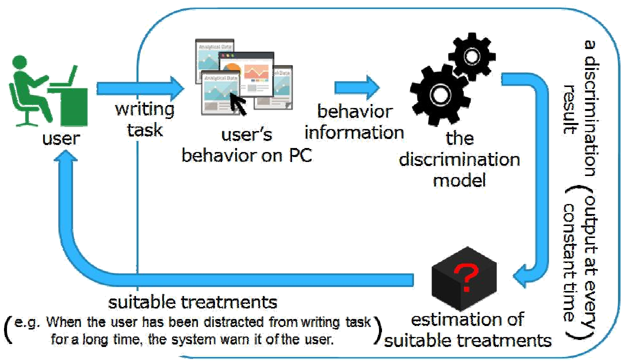

Figure 1. A use case of the method.

The discrimination model takes behavior logs of users to examine whether they get immersed at every constant period. Users may take a long time for writing. However, it is useless to provide discrimination results for them in a long period. It is desirable for the users to get immersed during the writing task as much as possible. To achieve it, the study proposes a method to provide a discrimination result at every short constant time. It enables provision for users while writing task to get discrimination results at every constant time. For example, when users have been distracted from writing tasks for the time being, the method warns it of them.

The data flow in the discrimination model is shown in Figure 2.

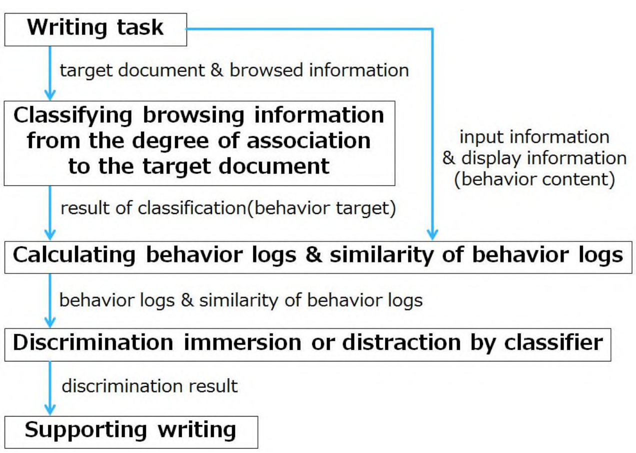

Figure 2. Discrimination Model.

A user usually makes a target document, referring to various kinds of information, such as Web pages. The discrimination model first clusters browsed information with its association degree with the target document. A behavior log is composed of results of the clustering, inputs with the mouse and the keyboard. It also contains what applications and Web pages are placed on the top of the screen, which indicates information the user is referring at the time point.

This study assumes users keep the same behavior during their immersion. For example, some users may go on typing characters in target documents to embody specific ideas in their minds. Others may stop the typing to consider their plans on target documents, or continue to search relevant information. Even if they are immersed into tasks other than writing, such as enjoying movies, they keep the same behavior. Based on this assumption, the discrimination model evaluates if users get immersed when they keep the same behavior. To distinguish immersion in the writing from that in other tasks, the discrimination model examines the information they refer to. Through the process above, it can detect distraction from writing without physical sensors.

B. Classification of browsed information

The study regards a target document as an electronic document being edited. It assumes only one document is addressed in a writing task. Browsed information means the information browsed by the Web browser. In this study, it is categorized into 3 types: reference information, supplemental information, and entertainment information, Reference information means any information, which contributes to the contents of a target document. In the study, it is assumed to be written in the same language as the target document. Taking a report assignment to explain the multiple regression analysis as an example, Web pages explaining the multiple regression in the same language as the target document, are examples of reference information. Supplemental information corresponds to information which is helpful to form the target writing, but not related to its contents. Back to the example of the report assignment on the regression analysis, supplemental information are Web pages explaining how to write a reference list, how to denote mathematical expression with TEX, and so on. Entertainment information involves unnecessary information for the writing. It is usually browsed by users' preference. In the report assignment example, Web pages on news article unrelated to the multiple regression, videos for users' entertainment and so on are categorized into entertainment information. For the writing task, reference information and supplemental information are necessary information, while entertainment information is unnecessary information. If we examine words, the association of reference information with the target document seems to be high, because reference information seems to have words similar to the target document. The association of supplemental information or entertainment information with the target document would be low, because supplemental information and entertainment information seem to have few words related to the target document. Measuring the degree of the association based on words, browsed information can be classified into the two groups: the group of reference information and the group of supplement information and entertainment information. This study classifies supplemental information and entertainment information into one group, because it is difficult to distinguish supplemental information from entertainment information by the word-based degree of association.

At a regular interval, the proposed method analyzes morphologically the target document at a specific time point and all documents browsed in the interval. It categorizes the browsed documents in the following way.

1) With the bag-of-words method, it generates the document vectors of the target document and each of the browsed ones using nouns other than pronouns, non-autonomous words, suffixes, and numerals.
2) It calculates the degree of the association with the cosine similarity between the document vector of the target document and that of the browsed ones.

3) It clusters browsed documents into 2 clusters based on the degree of the association with the k-means method [14]. It regards the cluster of high degree of the association as the group of reference information, while the one of low degree of the association as the group of supplemental information and entertainment information.

Pronouns, non-autonomous words, suffixes, and numerals are removed from elements of document vectors, because these nouns do not seem to represent the peculiarity of the target document and browsed ones. The bag-of-words considers the peculiarity of sentences from only the occurrence of specific words in the sentence [15][16]. In this study, each element of a document vector corresponds to a specific noun. It takes 1 if the noun appears in the document, otherwise 0 [See Figure 3].

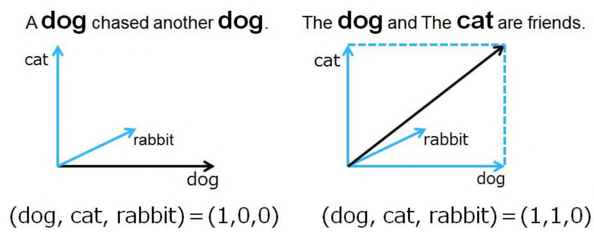

Figure 3. The bag-of-words.

The cosine similarity is calculated with (1), where the document vector of the target document is $\vec{m}$, while that of the browsed one is $\vec{r}$.

$$\cos(\vec{r}, \vec{m}) = \frac{\vec{r}\cdot\vec{m}}{|\vec{r}||\vec{m}|} \quad (1)$$

The proposed method classifies browsed documents into 2 clusters based on the degree of the association in the following k-means method.

a). It assigns each browsed document into 2 clusters randomly.
b). It calculates the centroid in each cluster.
c). It reclassifies each browsed document into the new cluster whose centroid is nearest to the browsed document.
d). It repeats b) and c) until the result of reclassification are the same.

C. Behavior log

The proposed method detects users' immersion in writing using the behavior log. At a regular interval, it examines the target document, documents in the group of reference information, and the ones in the group of supplemental information and entertainment information. It tests whether the user provides something for what is examined as its input, as well as whether the user places what is examined on the top of the screen. A behavior log contains binary values, which indicate the test results, as depicted in Figure 4.

To discriminate the user immersion in a short period, the duration to figure out the behavior log is far shorter than that to classify browsed information. The method regards the target document or the browsed information takes something as its

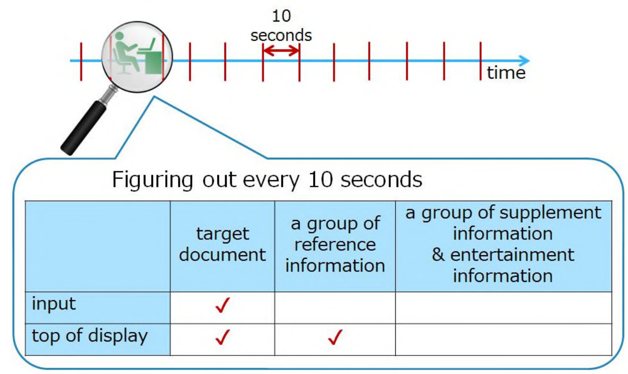

Figure 4. Behavior log.

input, if the user gives it typing with the keyboard, moving the mouse cursor or clicking the mouse. The method regards the target document or the browsed information is placed on the top of the screen, if any part of the window tool presenting it appears on the top of the screen. The target document or the browsed information is regarded not to be placed on the top, when the whole part of the tool has been hidden with other windows, or when the tool has been closed throughout the duration.

D. Similarity of behavior logs

The method addresses the 3 kinds of documents: the target document, a group of reference information, and a group of supplemental information and entertainment information. To make behavior logs, it examines 2 characteristics in each of them: their input and their appearance on the top of the screen. The behavior logs consist of 6 kinds of binary data items.

The method distinguishes users' immersion in writing using each of the similarity of behavior logs under the assumption that users keep the same behavior during their immersion. To figure out the similarity of behavior logs, the 2 characteristics of the 3 groups of the information is examined, which means the proposed method consider 6 kinds of binary data items for the similarity. The similarity of each binary data item is the number of matching behavior logs within a time window of a fixed length. The similarity for the k-th binary data item is figured out with (2).

$$similarity_k = \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} (a_{i;k}, a_{j;k}) \quad (2)$$

where a_i denotes the i-th behavior log, and $a_{i;k}$ denotes the k-th binary data item inside a_i. Function is defined as (3).

$$(a_{i;k}, a_{j;k}) = \begin{cases} 1 & (a_{i;k} = a_{j;k}) \\ 0 & (otherwise) \end{cases} \quad (3)$$

Apparently, this value gets large when the user keeps the same behavior.

IV. Experiment to Evaluate the Proposed Method

Here, we present the procedure of the experiment and the experimental results.

A. Purpose and procedure of experiment

The experiment was conducted to verify whether it is possible to distinguish users' concentration for the writing from the similarity of behavior logs. Five students in a university worked on the task where they summarized investigation results on the will power [17] in a report for 150 minutes. They worked under a dual display environment in order to bring it closer to their usual working environment. Every 10 seconds, their computers recorded both of the windows, which become the input target and the windows, which are placed on the top of the screen even a moment during the period. The intermediate state of the target document is also recorded every 1 hour. The working of the subjects was recorded with a video camera during the experiment. After the experiment, the subjects answered the questionnaire, which asked the period during which they were immersed in the writing, watching the video.

B. Results of the experiment

The length of time window to calculate the similarity of behavior logs was set to 120 seconds. The discrimination models are constructed by the random forest, support vector machine and gradient boosting to distinguish whether the subjects get immersed in the writing. Figure 5 illustrates the explanatory variables in these models.

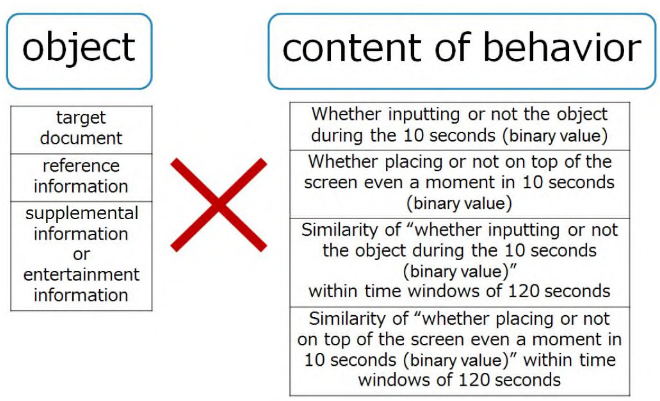

Figure 5. The explanatory variables.

5-fold cross validation was used for the evaluation. The precision, the recall, and the F-measure of the evaluation result for the 5 subjects are shown in Tables I, II, and III.

TABLE I. THE DISCRIMINATION RESULT BY THE RANDOM FOREST

	distraction from writing task	immersion in writing task
precision	0.63	0.71
recall	0.80	0.51
F-measure	0.70	0.59

Note that the F-measure value the distraction from the writing is 0.70 in all of these discrimination. The study to presume human immersion from the skin conductance, the heart rate and the electroencephalography using physical sensors [1] reported the accuracy of the discrimination of the distraction was 0.76, though the experimental environments and the situations are different. The result suggests that it is possible to distinguish the distraction from the writing using the above explanatory variables instead of physical sensors, although the accuracy decreases slightly. We can detect users who do not concentrate on the writing for a long time.

TABLE II. THE DISCRIMINATION RESULT BY THE SUPPORT VECTOR MACHINE

	distraction from writing task	immersion in writing task
precision	0.65	0.74
recall	0.82	0.53
F-measure	0.72	0.62

TABLE III. THE DISCRIMINATION RESULT BY THE GRADIENT BOOSTING

	distraction from writing task	immersion in writing task
precision	0.70	0.66
recall	0.73	0.60
F-measure	0.72	0.63

C. Important factors in discrimination

It is necessary to investigate whether the discrimination is caused by the similarity of behavior logs. The variable importance of the Random Forest algorithm indicates which explanatory variable is important in the discrimination. It says that the most important explanatory variable is not the similarity of behavior logs, but the binary variable representing whether supplemental information or entertainment information is placed on the top of the screen during the period of 10 second. To confirm the importance of the variable, let us regard the user runs into distraction from the writing when supplemental information or entertainment information is placed on the top of the screen, while he stays in the immersion in the writing when the information is not placed on the top of the screen. The data set collected in the experiment was discriminated again based on the simple criterion. The precision, the recall, and the F-measure of the discrimination are shown in Table IV.

TABLE IV. THE DISCRIMINATION RESULT BY SUPPLEMENTAL INFORMATION AND ENTERTAINMENT INFORMATION

	distraction from writing task	immersion in writing task
precision	0.67	0.72
recall	0.76	0.63
F-measure	0.71	0.67

Both of the F-measure values of the immersion and the distraction are higher than those of Section 4.*B*. Whether supplemental information or entertainment information is placed on the top of the screen is quite important among the explanatory variables in the experiment. On the contrary, the similarity of behavior logs is not as important as the placement of supplemental information or entertainment information.

V. DISCUSSION

The experimental result suggests the placement of supplemental information or entertainment information on the top of

 ISBN: 978-1-61208-580-7

the screen is important to distinguish whether concentration of users leads to their immersion in writing. Users who want to concentrate on the writing seem to exclude entertainment information from their vision under their immersion. On the contrary, they place entertainment information on the top of the screen when they are distracted. The following tendency was manually confirmed from the videos during the experiment. The subjects displayed entertainment information when they got distracted when they cannot concentrate on the writing any more. Meanwhile, they hid the entertainment information, switching the tabs on their browsers, or placing other windows covering the entertainment information when their mode changes from distraction to immersion.

Since supplementary information is not related to the distraction, the simple criterion assumed in Section 4.*C* may lead to erroneous discrimination. After the experience, the subjects manually classified browsed information into 3 categories: reference information, supplemental information, and entertainment information. The total number of entertainment information browsed by each subject in the experiment is 299, while that of supplemental information is 99. Accordingly, the entertainment information is approximately 3 times larger than the supplemental information. Since the supplemental information is far less than the entertainment information on this experiment, the accuracy of the discrimination is high, although the supplemental information is treated in the same group of the entertainment information. It is expected the accuracy is improved, distinguishing supplemental information from entertainment information.

The significance of the experiment results should be proved by increasing the number of the subjects. Also, the accuracy of the discrimination should be improved more for the practical usage. It would be improved, adding explanatory variables, such as the frequency of typing and clicking, the contents of the typed information, and so on. Additionally, it is necessary to confirm the validity of the questionnaire to form training data. They were determined through the answers of the subjects for questions asking the period where they were immersed into the writing. It is necessary to improve the validity of the questionnaire and the method of evaluation by using existing research to measure human immersion.

VI. CONCLUSION

In this paper, we have proposed a method to distinguish user immersion in writing without physical sensors. The method has been suggested to distinguish users' immersion in writing using each of the similarity of behavior logs under the assumption that users keep the same behavior during their immersion. It has turned out the most important explanatory variable is the binary variable representing whether supplemental information or entertainment information is placed on the top of the screen during the period of 10 seconds. From the videos during the experiment, it has been manually confirmed the subjects would display entertainment information when they have run into a distraction from writing. It implies they are likely to place entertainment information on the top of the display, when they cannot concentrate on the writing any more. Meanwhile, they would hide the entertainment information, switching the tabs on their browsers, or placing other windows covering the entertainment information, when their mode changes from distraction to immersion.

In the future, the significance of the experimental results should be proved by increasing the number of the subjects. Additionally, in actual environments, it may be possible for users to engage in multiple tasks at the same time. For example, while they are engaging in writing, they also have to handle interrupts of their work, such as responding to incoming calls. The range of adaptation of the method should be expanded to deal with such multitasking.

ACKNOWLEDGMENT

We gratefully appreciate the help of Prof. Yusuke Kajiwara who gave us a lot of advice and suggestions for conducting this research and writing this paper. We also express our gratitude to all members and colleagues of Data Engineering Laboratory from Ritsumeikan University who supported us and cooperated with the experiment.

REFERENCES

[1] I. Jraidi, M. Chaouachi, and C. Frasson, "A Dynamic Multimodal Approach for Assessing Learners' Interaction Experience," ICMI '13 Proceedings of the 15th ACM on International conference on multi-modal interaction, 2013, pp. 271–278.

[2] L. Nacke and C. A. Lindley, "Flow and Immersion in First-Person Shooters: Measuring the player's gameplay experience," Future Play '08 Proceedings of the 2008 Conference on Future Play: Research, Play, Share, 2008, pp. 81–88.

[3] T. Leelasawassuk, D. Damen, and W. Mayol-Cuevas, "Estimating Visual Attention from a Head Mounted IMU," ISWC '15 Proceedings of the 2015 ACM International Symposium on Wearable Computers, 2015, pp. 147–150.

[4] G. Lee, A. Ojha, and M. Lee, "Concentration Monitoring for Intelligent Tutoring System Based on Pupil and Eye-blink," HAI '15 Proceedings of the 3rd International Conference on Human-Agent Interaction, 2015, pp. 291–294.

[5] Ministry of Internal Affairs and Communications Institute for Information and Communications Policy, "Survey findings report on Internet use and dependency trend among young people," 2013, URL: http://www.soumu.go.jp/iicp/chousakenkyu/data/research/survey/telecom/2013/internet-addiction.pdf [accessed: 2016-12-21].

[6] B. L. S. Coker, "Workplace Internet Leisure Browsing," Human Performance, vol. 26, 2013, pp. 114–125.

[7] M. Csikszentmihalyi, Finding Flow: The Psychology of Engagement with Everyday Life. Basic Books, 2008.

[8] J. Saiki and E. Inoue, "Relationship between Concentration and Tempo-ral Duration Estimation: Implications for Flow Experience," Psycholo-gia, vol. 54, 2011, pp. 206–221.

[9] A. Sarrafzadeh, S. Alexander, F. Dadgostar, C. Fan, and A. Bigdeli, ""How do you know that I don't understand?" A look at the future of intelligent tutoring systems," Computers in Human Behavior, vol. 24, no. 4, 2008, pp. 1342–1363.

[10] Y.-M. Jang, R. Mallipeddi, S. Lee, H.-W. Kwak, and M. Lee, "Human intention recognition based on eyeball movement pattern and pupil size variation," Neurocomputing, vol. 128, 2013, pp. 421–432.

[11] A. Kapoor, W. Burleson, and P. Rosalind W, "Automatic prediction of frustration," International Journal of Human-Computer Studies, vol. 65, 2007, pp. 724–736.

[12] G. Lee, M. Kwon, S. K. Sri, and M. Lee, "Emotion recognition based on 3D fuzzy visual and EEG features in movie clips," Neurocomputing, vol. 144, 2014, pp. 560–568.

[13] S. D 'Mello, A. Olney, C. Williams, and P. Hays, "Gaze tutor: A gaze-reactive intelligent tutoring system," International Journal of Human-Computer Studies, vol. 70, 2012, pp. 377–398.

[14] J. MacQueen, "Some Methods For Classification and Analysis of Multivariate Observations," Proc. Fifth Berkeley Symp. on Math. Statist. and Prob, vol. 1, 1967, pp. 281–297.

[15] S. Dumais, J. Platt, D. Heckerman, and M. Sahami, “Inductive Learning Algorithms and Representations for Text Categorization,” CIKM ’98 Proceedings of the seventh international conference on Information and knowledge management, 1998, pp. 148–155.

[16] S. M. Weiss, C. A. andFred J. Damerau, D. E. Johnson, F. J. Oles, T. Goetz, and T. Hampp, “Maximizing Text-Mining Performance,” IEEE Intelligent Systems archive, vol. 14, 1999, pp. 63–69.

[17] K. McGonigal, The Willpower Instinct: How Self-Control Works, Why It Matters, and What You Can Do To Get More of IT. Avery, 2011.