SURE 静岡大学学術リポジトリ
Shizuoka University REpository

Gaze Analysis on the Effect of Intervention on Ruminative Web Browsing

メタデータ	言語: English 出版者: 公開日: 2022-02-28 キーワード (Ja): キーワード (En): Rumination, Advertisement, Cognitive modeling 作成者: Raj, Giri Basanta, Morita, Junya, Pitakchokchai, Thanakit メールアドレス: 所属:
URL	http://hdl.handle.net/10297/00028624

Gaze Analysis on the Effect of Intervention on Ruminative Web Browsing

Giri Basanta Raj[1], Junya Morita[2], and Thanakit Pitakchokchai

Shizuoka University, 3-5-1 Johoku, Naka-ku, Hamamatsu, JAPAN,
giri.basanta.raj.19@shizuoka.ac.jp[1]
j-morita@inf.shizuoka.ac.jp[2]

Abstract. In the current highly developed information society, having a habit of rumination (repetitive and negative thinking) can be dangerous to our mental health. To prevent rumination during web browsing, the authors' previous study built an advertisement system that is regulated by a computational cognitive model and users' heart rate variability (HRV). To validate and extend the system, this paper presents analyses of behavioral and physiological data obtained from two studies where participants engaged in two successive tasks: mood induction and main tasks. Study 1 aimed to develop a detector for rumination distraction utilizing behavioral and physiological data. Owing to the SVM classification, a large contribution of gaze extracted from the facial movie was verified. To validate these findings, Study 2, in which a small number of participants engaged in the same tasks as Study 1, was conducted with a well-established eye-tracking system. Analyses of the gaze data obtained with this device confirmed high consistency with the data obtained from the facial movies, and also confirmed the influence of advertisements on participants' attention during web browsing. Summarizing the results of these studies, the current paper indicates the validity of distracting rumination by presenting prompts regulated with a personalized cognitive model.

Keywords: rumination, advertisement, cognitive modeling

1 Introduction

Recently, we have been connected directly and indirectly with technology, which is now a part of our lives. Under such a situation, web browsing has become a daily life activity that positively and negatively influences our mood. In addition to the several advantages of this technology, we are prone to its negative effects as well. Among the several side effects of this technology, we focus on rumination, which is repetitive, negative thinking about one's unpleasant experiences caused by depressive moods. It has been revealed that high engagement with internet browsing is correlated with higher levels of anxiety [10]. Thus, with the advancement of this technology, we suffer from mood disorders, such as depression, bipolar disorders, and anxiety, which have direct effects on our mood and emotions [12].

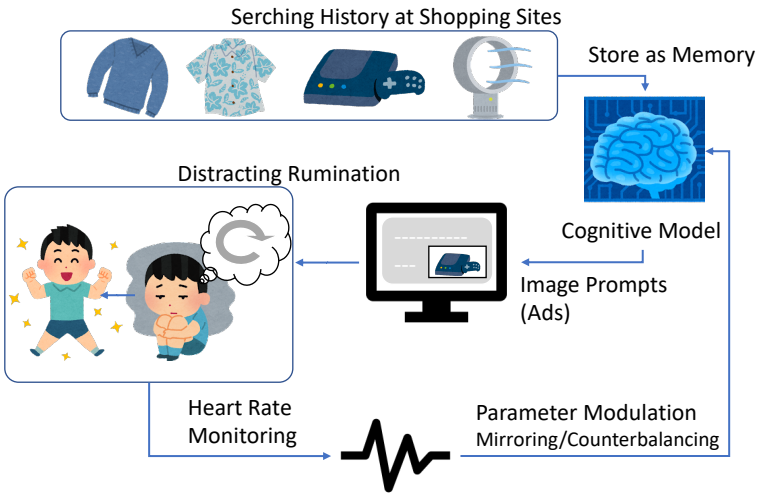

Fig. 1. System overview.

To address this negative side effects of technology, the discipline of affective computing, a multidisciplinary field spanning computer science, psychology, and cognitive science, is garnering attention to enable computations connected with human emotion [4]. Technologies in this field include recognizing the emotional state of a user with various types of sensing devices and response generation by performing specific, predefined product/service features [11]. Following such a method of affective computing, our previous study [8] developed a system to prevent rumination during web browsing. In this paper, we attempt to reveal the effect of the system on distracting rumination by analyzing gaze information.

The reminder of this paper comprises of a review of the system and descriptions of the two studies. The first study is a reanalysis of the data obtained in our previous study to construct a state detector related to the distraction of rumination. The second study is conducted to strengthen the findings obtained in the first study. In this study, we collect fine-grained gaze information in a small experiment. Summarizing these two studies, we discuss the effectiveness of our approach in the final section.

2 System of Model-based Advertisement

Our previous study [8] developed a system of model-based advertisements that distracted ruminative web browsing. The basic concept is illustrated in Figure 1. When the user ruminates, the system presents advertisements that divert attention from ruminating information on the web. To realize this concept, the

system combines a human memory model to retrieve product images and a heart rate sensor to monitor the state of the user that relates to rumination.

Specifically, we utilize ACT-R (Adaptive Control of Thought-Rational) cognitive architecture [2] as the human memory model. The product images observed in the user's daily lives are stored in the model's declarative memory. The retrieval of memory is governed by the ACT-R's memory theory. The basic principle of this is recency and frequency effects, claiming that humans usually remember recently obtained and frequently retrieved images [1]; however, such effects are changed (strengthened/weakened) by the participants' mood. In a stressful situation, people usually retrieve a small set of memories, such as rumination [13] while in a relaxed state, memory search is expanded [5, 7].

Our system connects the mood effects of memory with real-time monitoring of heart rate. There are two variant models of the connection: normal (NOR) and inverted (INV) models. The normal model exhibits ruminative behavior while the user is in a stressful state, in which the "memory noise parameter" (a random factor against the frequency and recency effects) is calculated corresponding to HRV. Thus, in the normal model, memory noise is assumed to be high when HRV is high, and vice versa. Inversely, the inverted model is designed to act as a counterbalance to the normal model that corrects rumination behavior. Therefore, the memory noise parameter of the inverted model is high when the HRV is low, and vice versa.

While users utilize the system, the ACT-R model retrieves the product images collected in their daily lives.[1] When the user enters a stressful state and ruminates (low HRV), the normal model tends to retrieve images that the user can recognize well (frequent recent images), while the inverted model tends to retrieve images that the user can barely remember. In contrast, when the user relaxes, the models behave inversely, the normal model tends to retrieve images that the user cannot recognize, and the inverted model retrieves images that the user knows very well. We assumed that the user ignores the images that he/she knows well and pays attention to/gets distracted by those that he/she has rarely or not seen before; we hypothesized that the inverted model will carry the relaxed state on as the user ignores the illustrated images and is distracted from rumination after observing the images.

The previous study [8] conducted an experiment in which two groups of participants utilized each system implementing the NOR/INV model and obtained supportive results indicating the effectiveness of the INV model in the subjective evaluation of distracting rumination; however, behavioral data clearly indicating this effect was not obtained. In this paper, we indicate further evidence of the effectiveness of the system by analyzing behavioral data, especially gaze information.

[1] We developed a browser extension to collect images on the browser when the user visits specific shopping sites (Amazon and Rakuten).

3 Study 1: Detector of Rumination Distraction

This study aims to construct a detector for the effect of the system introduced by the INV model. We consider that constructing such a detector is useful not solely to validate the effect of the system but also to develop a function that monitors users' rumination during web browsing. In the rest of this section, we will relay the results of this analysis after briefly describing the experiment conducted in the previous study.

3.1 Dataset

In the previous study [8], a behavioral experiment was conducted with 12 Japanese undergraduate students. The experiment comprised the following two sessions.

- Mood-induction task (MI): the participants recalled their recent bad memories and wrote the stories on a word processor software.
- Main task (MT): the participants searched for goods that can be useful for their future on the internet, while the system presented advertisements when the participants ruminated.

It was assumed that the first task induced a negative mood state (rumination), and that the second task caused the participants to recover from the negative mood state and induced a positive mood state (i.e., distracting rumination). The participants were seated in front of the computer during the experiment, and the face movement was recorded using a web camera. During the experiment, physiological arousal was also measured by a program that collected heart rate data from a heart rate monitor (Uniontool's myBeat). After the ACT-R model was activated in the MT session, the system displayed the image as an advertisement based on the user's current emotional condition.

3.2 Method of State Detection

Measurement of Heart Rate and Eye Gaze In our study, the ACT-R models repetitively output an image every several seconds on the computer screen during the main task. We assumed that such a presentation of images implicitly intervenes in rumination caused by the mood-induction task. This study attempted to detect the effects of the intervention.

To detect this effect, the previous study [8] analyzed physiological arousal using heart rate, facial expression, and self-report questionnaires. Among them, eyes and their movements are assumed to be important in expressing a person's desires, needs, and emotional states [9]. The state and movement of eyes can also express people's cognitive and emotional states based on their content of interest [6]. Thus, this study focused on eye-gaze movement. Simultaneously, HRV was used as a signal to demonstrate the current emotional state.

Figure 2 illustrates an overview of the proposed method. The method was used to classify the normal and inverted models in both tasks using heart rate

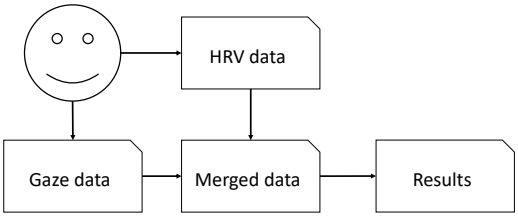

Fig. 2. Overview diagram of the state detection.

and eye gaze data. If the accuracy of classification between the two groups was higher in the main task than in the mood induction, we can conclude that the rumination of the participants was distracted by the advertisement on a computer screen in the main task.

Data Preparation Among the two different data in Figure 2, the HRV was measured utilizing Uniontool's MyBeat device and eye gaze data were extracted from the video with OpenFace software [3].

For each participant, 714 features were extracted from the recorded video using OpenFace. Out of them, we selected eight eye-gaze features. Generally, the participant's recorded video was 30 min long, of which approximately 15 minutes for the mood-induction task and approximately 15 min for the main task.

Thus, for each task and each participant, the 15-min long video contained 900 samples; each sample was 30 frames on average. The intervals between each heartbeat, that is, R-R interval data measured in the experiment, were also averaged for each second and concatenated with the gaze data. Each signal contained three beats. To resize the RRI data as the gaze data, we calculated the standard deviation of R-R intervals (SDNN) for each row in the data and merged this with the eye gaze data to make the final dataset and labeled the NOR and INV models.

Classification Method We adopted support vector machines (SVMs) as a supervised machine learning algorithm. For both MI and MT, SVM classified the two groups where the participants were exposed to advertisements using the different models (the normal and inverted models). A leave-one-out five-fold cross-validation method was adopted with a linear kernel and 10 regularization parameters C to calculate the accuracy of the classification between the results of the NOR and INV models in both tasks.

3.3 Results and Discussion

Data Visualization Figure 3 illustrates examples of fluctuations of the HRV in the MI and MT. The red line in each graph indicates the HRV, while the

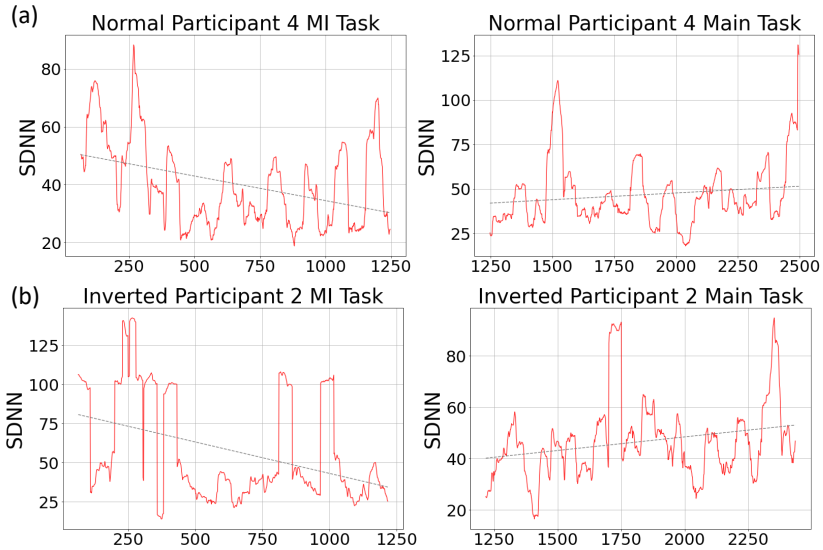

Fig. 3. Example of HRV data pattern in each group (a: Normal model b: Inverted model).

dotted gray line represents the trend of HRV. From the figures, we can observe the decreasing trends of HRV in the left panels (the mood-induction task) of Figures 3 (a) and (b), indicating that the participants went into the negative thought (ruminated). However, in the MT, the participants' rumination gets distracted by advertisements more in the INV group than in the NOR group, as illustrated in Figure 3 (b), although it highly depends upon the individuals.

Figure 4 illustrates the visualization of the data pattern of the MI and MT. Among the ten features of the dataset, SDNN and "gaze_0_x" were arbitrarily chosen to indicate the data distribution in each task. In Figure 4 (a), while the NOR INV models have similar data patterns in the MI, they are distributed differently in the MT. Similarly, in Figure 4 (b), the gaze movements for the NOR and INV models largely overlap in the MI, but we can observe the difference between the two peaks corresponding to the two models, suggesting the distraction of the rumination caused by the presentation of advertisements by the INV models.

Classification Result Table 1 shows the accuracy of SVM classification of the two groups for each task. In this analysis, we applied two different feature sets, indicating the impact of eye movement and heart rate on the classification. Although all the accuracies in the table exceeded the chance level (.50) reflecting individual differences between participants in the NOR and INV groups, we can observe differences in accuracies between the two tasks for each feature dataset.

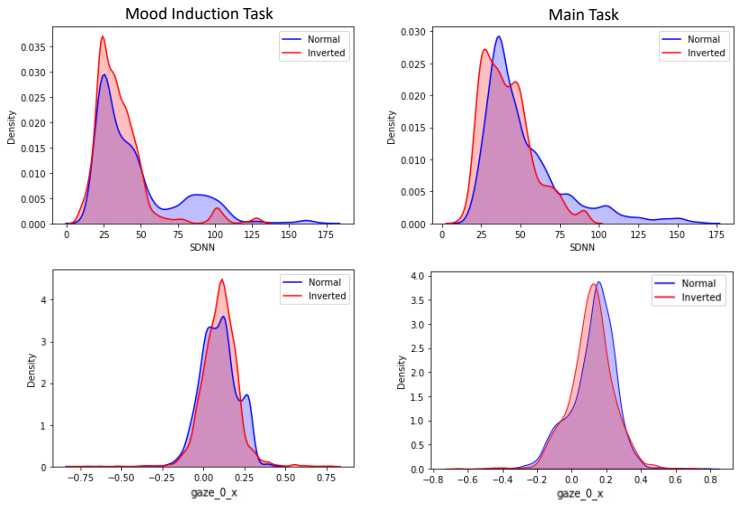

Fig. 4. Example density plots of SDNN (Top) and "gaze_0_x" (Bottom).

Table 1. Accuracy of Mood-induction and Main Task

Tasks	Data	Accuracy	Number of features
Mood-Induction task	Gaze data only	0.66	9
	Merged data	0.72	10
Main task	Gaze data only	0.74	9
	Merged data	0.78	10

This result indicates the possibility of detecting the effects of mood induction using our method.

We can also observe a larger difference between tasks in the gaze data only condition than the merged data, suggesting a large contribution of gaze pattern discriminating the two tasks. This result can be interpreted in an experimental setting. Because the mood-induction task made the participants write a text while recalling their recent unpleasant experiences, there seemed to be relatively fewer eye movements. However, in the main task, the participants utilized a web browser displaying advertisements on a computer screen. Thus, the eye gaze movements were assumed to be different, as they were prone to seeing the image displayed. Therefore, the classification accuracy of the gaze-data only condition is lower in the MI than in the MT.

In addition, we can observe that in the merged data (the eye gaze and the RRI data), the classification accuracy is higher than the eye gaze only data condition, indicating that the RRI also plays a role in the classification between the normal and inverted models. This result suggests some difference in recovering the mood

induced in the mood-induction task; the inverted participants gradually became relaxed, while the normal groups tended to continue the mood induced in the previous task.

4 Study 2: Detailed Gaze Analysis

This study was conducted to support the results of the first study, following the experimental setup and experimental tasks with a few changes in the tasks. We utilized a remote eye-tracking device (Tobii pro nano) to monitor eye behavior in addition to a web camera. Using data obtained with this device, we conducted a detailed analysis of eye movements.

4.1 Data collection

Participants Four participants who are international master's degree students at Shizuoka University were asked to participate in the experiment. All of them reported that they usually visit shopping websites (Amazon) a few times a week.

Design In the previous study, for the main task, the participants were divided into two groups that experienced different cognitive models: NOR and INV models. As indicated in (3.3), these two-model behaviors are distinguishable. Therefore, in this study, we conducted an experiment using solely the counterbalance model (INV model) to examine the actual effects of advertisements on the participants and gaze behavior during system use.

Procedure As aforemantioned, the procedure of this study is approximately the same as the previous study, with a few changes in tasks and experimental devices. In this study, the mood induction task (MI) asked participants to recall unpleasant experiences and utilize the internet to search related thoughts instead of writing a report. This modification was made to mirror the previous mood induction task (MI), where the participants also engaged in an internet search for future life. We set up all data collection devices including Tobii's eye tracker. After confirming that all the devices worked well, the procedure was explained to the participants. The entire experiment took approximately 55 min.

4.2 Results

In this study, we present an analysis of gaze information to support the results obtained in the first study. Particularly, we indicate the difference in gaze patterns between the two tasks and the consistency of gaze information obtained with the Tobii eye tracker and OpenFace.

Gaze fixation on the advertisement area We obtained gaze point variables from the Tobii eye tracker. It extracted the gaze point variables as the raw gaze coordinates for each eye individually which are pixels display area co-ordinates system (DACS), which is a coordinate system with its origin in the top left corner of the screen (stimuli area), with Y pointing downwards. Accordingly, the gaze point variables extracted by the tracker are "Gaze point X [DACS px]," and "Gaze point Y [DACS px]" corresponding to the x and y-coordinates, respectively. The screen recording of the Tobii tracker presentation resolution was 1920 × 1080 (Pro Spectrum, Pro TX300), and the display resolution utilized for the experiment was 1920 × 1080. We calculated the total gaze point on the advertisement area for each task by calculating the x-coordinate between 1500 and 1800-pixels, and y-coordinate between 200 and 400 pixels which resembles the top right corner of the screen, that is, the advertisement area on a computer screen.

Table 2 presents the fixation on the screen and total fixation on the advertisement area for each participant. From this table, we can observe that the total fixation on the advertisement area in the main task is higher than that in the mood induction task.

Table 2. Total fixations and fixations on the advertisement area during the two tasks.

	Mood induction task		Main Task	
	Total fixation	Ads area	Total Fixation	Ads area
P1	53840	96	53692	1782
P2	54036	0	54032	674
P3	54034	31	54028	82
P4	53839	0	53942	1377

We also calculated the correlation between the total image displayed and the total fixation on the advertisement area in the main task. As explained in Section 2, the model-based advertisement system alters product images on the display based on the monitoring of users' HRV. The numbers of total and unique displayed images are indicated in Table 3, while the correlation between these numbers and the fixation on the advertisement area is indicated in Table 4. From the figure, we found a significant correlation between the total image displayed and total fixation in the main task. This result indicates that the participants' gaze movements were influenced by the advertisement.

Correspondence between Tobii and OpenFace As in the previous study, we utilized a web camera to record the video of the participants' face during the experiment and the OpenFace to extract the gaze data. For these video-based data, we examined the consistency with the data obtained using the eye tracker. The Tobii eye tracker extracted approximately 1000 samples per second,

Table 3. Number of total and unique displayed images.

	Total images	Unique images
P1	158	70
P2	127	91
P3	85	59
P4	140	88

Table 4. Spearman's correlation between images displayed and total fixation on advertisement area. Numbers in the parenthesis indicate p-value.

	Total images	Unique images
MI	.32 (.68)	-.74 (.26)
MT	1.00 (.00)	.20 (.80)

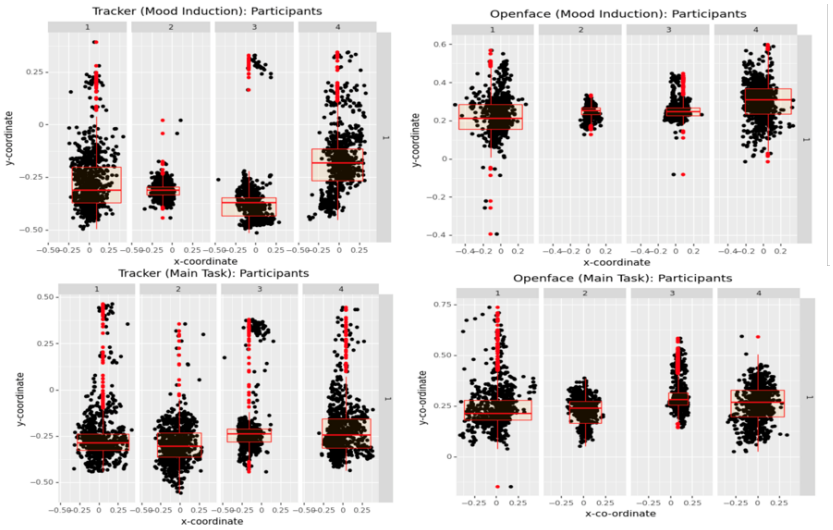

Fig. 5. Gaze direction of x-y coordinates in the two tasks for each participant (Top: mood induction task, bottom: main task; left: tracker, right: OpenFace).

whereas OpenFace extracted approximately 30 samples per second. Therefore, we calculated the average of each second sample for each variable, outputting a total of 900 samples for a 15 min video.

Among the several variables obtained by both methods, we focused on the gaze direction. Tobbi eye tracker extracted the gaze direction variables as the unit vector for the direction of the gaze such as ["Gaze direction left X," " Gaze direction left Y," "Gaze direction left Z," "Gaze direction right X," "Gaze direction right Y," "Gaze direction right Z"]. From this, for each participant in each task, we aggregated "Gaze direction left X" and "Gaze direction right X" to make x-coordinate and "Gaze direction left Y" and "Gaze direction right Y" to make the y-coordinate. Similarly, we aggregated the gaze direction of the OpenFace gaze direction to make the x and y-coordinates of the gaze direction.

We plotted the scatter and box for each task data of gaze direction in x-y coordinates, as illustrated in Figure 5. From the figures, we observed several differences between the two gaze extraction methods. That is, the median of the

y-coordinates of the tracker is negative, and that of OpenFace is positive. This is because of the different locations of each device (i.e., the web camera is at the top of the desktop, and the eye tracker is placed parallel to the bottom of the desktop). In addition, the tracker gaze direction moves toward the left, and the OpenFace gaze direction moves toward the right, reflecting different orientations of the OpenFace and eye tracker. Regarding the eye tracker, the orientation is the observer (i.e., participants who are sitting in front of a desktop), but in the case of OpenFace, the orientation is the web camera.

Despite these differences, the two extraction indicated similar individual differences. For example, both devices indicated that participant-2 (p2) and participant-3 (p3) have a negligible variance. To confirm the consistency between the two methods, we calculated the Pearson's correlation coefficient between the Tobii eye tracker and OpenFace gaze data. Table 5 presents the significant correlations between both x and y coordinates in all the tasks, although the correlation for the x-coordinate is negative because of the difference in the origin of the gaze direction, as described above. These significant correlations suggest that there is consistency between OpenFace and the eye tracker.

Table 5. Pearman's correlation coefficient between Openface and Eye Tracker data. Numbers in the parenthesis indicate p-value.

	Full task	MI	MT
X-coordinate	-.72 (< .001)	-.69 (< .001)	-.77 (< .001)
Y-coordinate	.60 (< .001)	.60 (< .001)	.62 (< .001)

4.3 Discussion

This study presents a detailed analysis of gaze behavior that is consistent with the two gaze extraction methods, indicating the validity of detecting mood changes with the OpenFace tracking system. We consider the advantage of video-based gaze extraction (OpenFace) to include a system of model-based advertisements. The cost of introducing OpenFace is very low because users can utilize a face camera in most web browsing situations. The study presented here indicates the possibility of monitoring users' attention to ruminating/distracting information using this method.

5 Conclusion

The presented studies indicate the successful application of the proposed method of detecting rumination and distraction of rumination based on a computational

cognitive model. The normal model was hypothesized to synchronize participants' rumination, while the inverted model was hypothesized to act as a counterpart to distracting participants' rumination by advertisements. The obtained results confirm these hypotheses. However, the data were limited to this study. In addition, further investigation of the application of this study and the limitations of the proposed method must be determined. In the future work, we will also develop a system integrating the advertisement system and the detector of rumination distraction developed in the first study.

References

1. Anderson, J.R., Schooler, L.J.: Reflections of the environment in memory. Psychological Science 2(6), 396–408 (1991)
2. Anderson, J.R.: How can the human mind occur in the physical universe? Oxford University Press (2007)
3. Baltrusaitis, T., Zadeh, A., Lim, Y.C., Morency, L.P.: Openface 2.0: Facial behavior analysis toolkit. In: 2018 13th IEEE international conference on automatic face & gesture recognition (FG 2018). pp. 59–66. IEEE (2018)
4. Calabrese, B., Cannataro, M.: Sentiment analysis and affective computing: Methods and applications. In: International Workshop on Brain-Inspired Computing. pp. 169–178. Springer (2015)
5. Dancy, C.L., Ritter, F.E., Berry, K.A., Klein, L.C.: Using a cognitive architecture with a physiological substrate to represent effects of a psychological stressor on cognition. Computational and Mathematical Organization Theory 21(1), 90–114 (2015)
6. He, H., She, Y., Xiahou, J., Yao, J., Li, J., Hong, Q., Ji, Y.: Real-time eye-gaze based interaction for human intention prediction and emotion analysis. In: Proceedings of Computer Graphics International 2018, pp. 185–194 (2018)
7. Itabashi, K., Morita, J., Hirayama, T., Mase, K., K., Y.: Evaluation of photo slideshow implementing a cognitive model and physiological indices. In: Proceedings of the 19th International Conference on Cognitive Modeling. p. 93–99 (2019)
8. Morita, J., Pitakchokchai, T., Raj, G.B., Yamamoto, Y., Yuhashi, H., Koguchi, T.: Regulating ruminative web browsing based on the counterbalance modeling approach. Frontiers in Artificial Intelligence 5 (2022), https://www.frontiersin.org/article/10.3389/frai.2022.741610
9. Ohno, T.: One-point calibration gaze tracking method. In: Proceedings of the 2006 symposium on Eye tracking research & applications. pp. 34–34 (2006)
10. Panova, T., Lleras, A.: Avoidance or boredom: Negative mental health outcomes associated with use of information and communication technologies depend on users' motivations. Computers in Human Behavior 58, 249–258 (2016)
11. Picard, R.W.: Affective computing: challenges. International Journal of Human-Computer Studies 59(1-2), 55–64 (2003)
12. Valstar, M.: Automatic behaviour understanding in medicine. In: Proceedings of the 2014 Workshop on Roadmapping the Future of Multimodal Interaction Research including Business Opportunities and Challenges. pp. 57–60 (2014)
13. Van Vugt, M.K., Taatgen, N.A., Sackur, J., Bastian, M., Borst, J., Mehlhorn, K.: Modeling mind-wandering: a tool to better understand distraction. In: Proceedings of the 13th International Conference on Cognitive Modeling. pp. 252–257. University of Groningen Groningen, Netherlands (2015)