Feature Analysis of PC Operation Logs and Biological Information for Estimating Users' Focus of Mind

Chikako Ishizawa ^{a,*}, Yuki Usami^a, Yoichi Kageyama^a, Motonari Shirasu^b

^aAkita University, 1-1 Tegata Gakuen-machi, Akita-shi, Akita 010-8502, Japan

^bJapan Business Systems, Inc., Toranomon Hills Mori Tower 16F, 1-23-1 Toranomon, Minato-ku, Tokyo 105-6316, Japan

*Corresponding Author: ishizawa@ie.akita-u.ac.jp

Abstract

There is a growing demand for better understanding operators' behavior while using personal computers (PCs). This study assumes that an operator's focus is on work while it progresses smoothly. To estimate whether an operator's mind remains focused on work, we obtained PC operation logs for keyboard activity during the work shift. In addition, we measured biological information.

Experimental results showed that it is possible to estimate whether the operator's mind is focused on work by investigating keyboard logs. It became clear that stress increased and body movement other than operating the PC was reduced when an operator's mind was focused on work.

Keywords: focus of mind, keyboard operation log, pulse, body movement

1. Introduction

With the increased use of personal computers (PCs) in schools and offices, there is a growing demand for understanding the behavior of operators. For example, when creating a report or document, a teacher or a boss wants to know whether the entry work of a student or a worker is smooth. If the entry work is not smooth, it is necessary to identify the problem and to provide guidance. Although there are various reasons why entry work might not be smooth, it is generally accepted that entry errors increase when the operator is impatient or anxious. Impatience may occur when there is insufficient time or if the workload is heavy. Anxiety may occur if a worker lacks confidence. Thus, understanding the behavior of PC operators contributes to early detection of problems.

There is little research on the behavior of PC operators. Although there is a study that determined whether interaction with a computer is by a machine or a human⁽¹⁾, there is no study to determine whether interaction with a computer changes in conjunction with change in the operator's mind.

It has been shown that there are strong correlations between (1) an increase in interruptions to the operator's work and the operator's stress level and (2) an increase in operator error with increasing stress⁽²⁾. Therefore, under conditions where there is almost no stress, the assumption is that work will progress smoothly. However, when work is progressing smoothly, or when the operator's mind is focused on work, it is not clear whether the operator feels stress.

It has been shown that when children are enthusiastic about listening to a story, their head movements decrease⁽³⁾. Similarly, if an operator's mind is focused on work, it is likely that movement other than that used to operate the PC is reduced.

This study assumes that the operator's mind is focused on work while the work is progressing smoothly. The goal of this study is to propose a method to estimate whether an operator's mind is focused on work by monitoring PC operation logs. To begin developing this method, we obtained PC operation logs and analyzed various features when an operator's mind was focused on entering sentences into documents. Additionally, we measured biological information concurrently with the logging of PC operations.

2. Experimental Method

2.1 Experimental conditions

The experiment was conducted in a quiet room. The room's layout is shown in Fig. 1. Subjects were asked to

performed the task of reading paper documents and entering the text into word processing software on the PC for 30 minutes. The collection of PC operation logs and the measuring subjects' biological information were performed together. There were four types paper documents, and subjects rested for one minute each time they finished inputting one type of document.

Although there are individual differences in the length of time that a subject can sustain focus on a task, it may be difficult for the subjects to maintain focus for a span of 30 minutes. In order to analyze the PC operation logs and biological information while the subject is focused on the task, it is necessary to detect what we refer to as the "time zone of focusing." To detect the time zone of focusing, we interviewed subjects about the text (i.e., the sentences or phrases of the document) they entered after they completed a document. Additionally, we recorded video with a camera placed at an angle where the PC monitor screen and the keyboard were both captured. The subjects were 23 students (college students and graduate students), designated by the letters A to W, who were in their 20s and used PCs on a daily basis.

2.2 Method to detect the time zone of focusing

First, the interview answers were confirmed with the electronic documents typed and the videos taken of the subjects. Next, the estimated time zone was determined based on the video time stamp of the subject entering the text corresponding to the interview answer. To determine the time zone of focusing we investigated the PC operation logs corresponding to the estimated time zone. When there was evidence of entering the interview answer text in the PC operation logs, the time zone of focusing was set to the estimated time zone.

2.3 Method to obtain and analyze PC operation logs

To investigate the sentence input of the operators, the keyboard logs, window logs, and mouse logs were obtained. All logs were obtained using the Microsoft® Windows® API⁽⁴⁾. Keyboard logs indicate when a key was pressed. Mouse logs indicate when a mouse click occurred. Information about the active window was captured by the window logs approximately every two seconds. Items recorded in each of the logs are listed in Table 1.

The time that the word processing software was running was determined using the window and mouse operation logs. Keyboard logs were selected for the time that the word processing software was running. Finally, the

selected keyboard logs were counted for each five second interval, and the variance value was calculated. Since keyboard logs indicate when a key is pressed, the number of logs counted equals the number of keystrokes.

2.4 Method to measure and analyze biological information

Pulse and body movements were measured. Pulse was measured by having the operator wear the pulse meter⁽⁵⁾ shown in Fig. 2 on his or her arm. Body movements were measured by an accelerometer⁽⁶⁾, shown in Fig. 3, attached to the chair. It is assumed that while the subject is focused on the input task, the arm moves but the posture barely changes. The accelerometer attached to the chair captured changes in posture. The sampling interval of the accelerometer was 0.5 seconds.

(a) Analysis of pulse

In a general analysis of heart beats, R-R intervals (RRI) are used^(7,8). The R wave is the wave with the largest amplitude, and the time interval between adjacent R waves

Fig. 1 Experimental room

Table 1 Items recorded in PC operation logs

Type of log	Item
Keyboard	Type of key pressed
Window	Position in the monitor, Size of window
Mouse	State of mouse button, Cursor position
All types	Date and time, Window title, Window handle, Window class

Fig. 2 Pulse meter⁽⁵⁾

Fig. 3 Accelerometer⁽⁶⁾

is calculated. In evaluating stress by heart beats, fluctuations of RRI are used. This study used two methods to evaluate stress by RRI measurements.

The first method uses the Lorentz plot analysis⁽⁹⁾. In the Lorentz plot analysis, the increase or decrease of RRI is plotted in twodimensions. Specifically, RRI is calculated from waveform data, and the k-th RRI is drawn on the x-axis, and the k+1-th RRI is drawn on the y-axis. An example of the Lorentz plot shape is shown in Fig. 4. Usually, the shape is like an ellipse. The ratio of the major axis to the minor axis of the ellipse is used as an indicator of the sympathetic nerve activity. It is generally accepted that the ratio is smaller if there is more stress.

The second method uses the ratio of the standard deviation of RRI and the mean of RRI. This ratio is called the coefficient of variation of R-R interval (CV-RR)⁽⁷⁾, and is used as an indicator of parasympathetic activity. CV-RR is calculated by the following formula:

$$CV-RR = \text{Standard deviation of RRI} / \text{Mean of RRI} \quad (1)$$

It is generally accepted that the value of CV-RR increases as relaxation increases.

In this study, the Lorentz plot analysis was applied to the measured pulse data, and the ratio of the major axis to the minor axis of the ellipse was calculated. In addition, the value of CV-RR was also calculated.

(b) Analysis of body movement

The feature amount of body movement was calculated in the following three steps. First, the values of the x-axis, y-axis, and z-axis of the measured data were synthesized using formula (2), to calculate the amount of chair movement.

$$\text{Movement amount} = (x^2 + y^2 + z^2)^{1/2} \quad (2)$$

x, y, z : Acceleration data

Next, the difference in the movement amount was calculated every 0.5 seconds. Specifically, the difference between the n-th movement amount and the n+1-th movement amount was calculated, and if the difference was not zero it was

counted as a chair movement. Finally, the number of chair movements and the amount of chair movement were combined by formula (3). In this study, the result of calculation by formula (3) was defined as the feature amount of body movement.

$$F = \text{Movement number} \times \text{Movement amount} \quad (3)$$

F : Feature amount

3. Experimental Results and Discussion

3.1 Time zone of focusing detected

An example of detecting the time zone of focusing is shown by the red-colored areas in Fig. 5. There is a time zone of focusing for the work of entering each document.

Most subjects finished the experiment while entering the third document. However, there were cases where there was no time zone of focusing while entering the first and third documents. The reason may be that the subjects are not accustomed to the keyboard used in the experiment and are tired near the end of the experiment. Therefore, it is considered that many subjects were not able to focus on the task when the near the start and the end of the experiment.

The time zone of focusing is increased while entering the second document. Fifteen of the 23 subjects focused on the task while entering the second document (the subjects designated A to O). Therefore, the following discussion uses the PC operation logs and biological information while subjects were entering the second document.

3.2 Analysis result of the keyboard operation logs

The variance value of the number of keystrokes is shown in Table 2. The green cells show that the variance value at the time zone of focusing is higher than the whole input time for the second document. The subjects displaying this result are 11 among 15.

In the experiment, the subjects repeat two tasks of reading the document and entering the text. It is thought that the number of keystrokes decreases while the subjects read the document. Thus, it is considered that the number of keystrokes every 5 seconds repeatedly increases and decreases. When the subject's mind is focused on the task, the amount of reading increases. Along with this, the amount of text input also increases. Thus, it is considered that as the range of increases and decreases expands, and the variance value is higher.

These results show that it is possible to estimate whether the operator's mind is focused on work by

Fig. 4 Example of a Lorentz plot

Fig. 5 Example of the time zone of focusing (Subject A)

Table 2 Variance value of keystrokes

Subject	Variance value of keystrokes	
	Time zone of focusing	Whole input time
A	35.540	33.403
B	80.716	48.496
C	24.818	18.014
D	21.299	19.943
E	29.135	28.794
F	140.140	127.732
G	41.149	32.236
H	23.923	22.364
I	338.010	116.238
J	12.527	11.980
K	16.132	14.611
L	15.842	24.290
M	21.758	37.235
N	44.153	47.819
O	9.270	23.341

More than whole input time

monitoring keyboard operation logs.

3.3 Analysis result of pulse

The elliptical ratios of the Lorenz plot analysis result are shown Fig.6. The results of the CV-RR analysis are shown Fig. 7. In both Figs. 6 and 7, the red bars indicate the time zone of focusing and the blue bars indicate the whole input time for the second document.

In Fig. 6, there are many cases where the elliptical ratio at the time zone of focusing is small. The number of subjects displaying this result is 11 out of 15. This result shows that the sympathetic nerve is active. In Fig. 7, there are many cases where the CV-RR at the time zone of focusing is small. The number of subjects displaying this result is 10 out of 15. This result shows that the parasympathetic activity is small.

By analyzing the pulse data, it became clear that stress

Fig. 6 Ratio of Lorenz plot ellipse (L/T)

Fig. 7 Coefficient of variation of R-R interval (CV-RR)

increases at the time zone of focusing. It is considered that the stress at the time zone of focusing is not large enough to cause many input errors. However, sustained stress can cause operators to make mistakes and is unhealthy for them. Therefore, monitoring the length of the time zone of focusing is expected to predict a smooth progress of the work and the health of the PC operator.

3.4 Analysis result of the body movement

The feature amount of body movement is shown Fig. 8. The red bars indicate the time zone of focusing and the blue bars indicate the whole input time for the second document.

Fig. 8 Body movement rate

To investigate the difference of body movement between the whole input time and the time zone of focusing, the feature amount of body movement is normalized so that the amount of the whole input time become 100% in Fig. 8.

There are many cases where the feature amount of body movement at the time zone of focusing is small. The number subjects displaying this result is 10 out of 15. On the other hand, subjects B and F have zero acceleration data. That is, these two subjects do not always change their posture to a point that the chair moves.

By analyzing body movement, it became clear that movement other than that of operating the PC is reduced when the operator's mind focuses on the work.

4. Conclusions

This study obtained the keyboard logs, window logs, and mouse logs and analyzed them to identify when the operator was entering sentences into documents. Additionally, the pulse and body movements of the operator were measured. The experimental results were as follows:

- (1) The number of keystrokes per unit time repeatedly increased and decreased, and the variance value during the time zone of focusing was higher than that during the whole input time.
- (2) The ratio of Lorenz plot ellipse and value of CV-RR were small at the time zone of focusing and showed that there was stress.
- (3) The change in the operator's posture as chair movements decreased were indicative of when the

operator's mind was focused on the work.

Future studies will propose a method to automatically detect time zones of focusing during task performance.

Acknowledgment

This research was supported by JSPS KAKENHI, Grant Number JP18K02852.

References

- (1) Nikhil Kumar Singh, Deepak Singh Tomar, Bhola Nath Roy: "An Approach to Understand the End User Behavior through Log Analysis," *International Journal of Computer Applications*, Vol. 5, No. 11, pp. 27-34, 2010
- (2) Satoshi Hashimoto, Takahiro Tanaka, Kazuaki Aoki, Kinuya Fujita: "Improvement of Interruptibility Estimation during PC Work by Reflecting Conversation Status," *IEICE Transactions on Information and Systems*, Vol. E97-D, No. 12, pp. 3171-3180, 2014
- (3) Shigeo Kaneda, Kazuya Kosaka, Kazuhiro Imaki, Takahiro Mitsumoto, Kimio Shintani, Aki Kono: "A Concentration Analysis Approach for Storytelling using Stereo Cameras and Accelerometers," *The Special Interest Group Technical Reports of IPSJ*, Vol. 2010-CE-104, No. 2, 2010
- (4) MSDN, <http://msdn.microsoft.com> (Accessed 2019/06/04)
- (5) EPSON, <https://www.epson.jp/products/pulsense/ps600b/spec.htm> (Accessed 2019/06/04)
- (6) GRAPHTEC, http://www.graphtec.co.jp/site_instrument/instrument/gl100/index.html (Accessed 2019/06/04)
- (7) Shingi Mitake: "Physiological measurement and data analysis know-how for product development and evaluation: Characteristics of physiological index, measurement method, experimental plan, interpretation and evaluation method of data," NTS Co., Ltd.
- (8) Yoshiaki Matsumoto, Nobuaki Mori, Ryo Mitajiri, Zhongwei Jiang: "Study of Mental Stress Evaluation based on analysis of Heart Rate Variability : Considerations of Evaluation by Pattern Recognition," *Journal of Life Support Engineering*, Vol. 24, No. 2, pp. 62-69, 2012
- (9) Naomi Dodo: "Effects of Sound Tempos on Autonomic Nervous System Functions," *Journal of psychological science Health Sciences University of Hokkaido*, No. 8, pp. 7-13, 2012