Estimation of Mind State with Body Movements Observed using Video Cameras

with Frequency Domain Analysis

Kyota Aoki
Graduate School of Engineering,
Utsunomiya University,
Utsunomiya, JAPAN
kyota@is.utsunomiya-u.ac.jp

Souksakhone Bounyong*
Graduate School of Engineering,
Utsunomiya University,
Utsunomiya, JAPAN
souksakhone3@gmail.com

***Abstract*— This paper proposes a method for estimating the state of mind on observation of the body motion from the front and the side of a subject using 2 cameras. There are many researches that studied about the estimation of a psychological condition using facial features. For example, there are studies about facial expressions, eyes' motion and the facial thermal images using a specific camera. Observation of the movement of a body is easier and takes lower cost than one of the facial images. We especially focused on the situation when people feel bored and pleasant. We used the correlation analysis method to find which body motion correlates to the states of a mind. To estimate the state of a mind the authors have focused on the average amount of body motion, and the power (amplitude) of motion in each frequency range. Our experiment shows that when subject feels bored, stretching body up and down, and movements between left and right tend to increase that can be observed from both cameras. Especially, the result of both cameras has shown that the amplitude of stretching body up and down correlates to the state of mind in frequency range from 2.8Hz to 4.0Hz. The varieties of the subjects show many possible dependencies between body motions and the state of a subject's mind.**

***Index Terms*— body motion, video observation, mind state estimation**

I. INTRODUCTION

We are able to use cameras and image processing technology in many ways to make our life more comfortable, such as in education, medical treatment, helping disabled people, etc. We try to understand what people are feeling or what they want to do without speaking or writing with information and communication technologies.

There are many researches that studied about the estimation of the psychological condition using facial features [1-5]. For example, they are facial expressions, eyes' motions, and the facial thermal images using a specific camera. In some cases, it is difficult to observe facial expressions. However, we can observe the body's motion more easily. Other work uses the keystrokes [6]. The authors assume that the state of people's mind affects the one's body movement.

We proposed a method to estimate the state of mind with the observation of body's motions [7]. The proposed method is based on the observation of the body motion using 2 cameras. The experiment is designed to estimate subjects' state of mind, especially feeling about boring or pleasant. To distinguish those 2 psychological conditions, subjects are requested to read some short English novels and comics. After that, subjects are asked to fill the questionnaire about those English novels and comics to obtain the states of the mind.

During the experiment, 2 cameras are located at the front and at the side of a subject to capture the body's motion. The image measurement is non-contact measuring method. With the proposed method, we can observe subject's natural movement in natural psychological condition, because it is unnecessary to set any sensor on the subject's body and do not have to locate the cameras in front of the subject that can make the subject to feel uncomfortable.

In our experiment, we find the corner points and calculate optical flow at the corner points to measure the movement of the subjects' body. This paper focused on the average amount of total body motion and the average amount of power (amplitude) in each frequency range of the body's motion for estimating the mind states of a subject.

We analyze the correlation between the observed body movements and the results on the questionnaire. The authors expect as follows. When subject is reading an interesting story, he must keep his body still with the increasing of his concentration. This makes the amount of body movements to be small. By contrast, if the story is difficult and the subject does not understand it, the amount of body movements must be large, because subject cannot concentrate on reading it.

The experiment shows that the motions observed from 2 cameras describe the state of mind. For example, when subject feels bored, stretching body up and down, and moving to left to right tend to increase that can be observed from both cameras. Especially, the result of the both cameras shows that the amplitude of stretching body up and down correlates to the state of mind in frequency range from 2.8 Hz to 4.0 Hz of the body motion. There are apparent differences between Japanese subjects and other country's subjects. This shows that there may be the differences in the body motions among cultures.

First, this paper describes the mind state estimation method with video images. Then, this paper discusses the 2 types of

*This author is now in
Nishi-kadoma District, Corporate R & D Division, Panasonic Corporation, 1006 Kadoma, Osaka 571-8501, Japan.

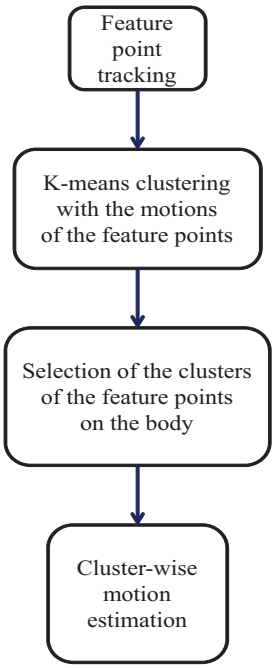

Fig. 1. Motion acquisition from videos.

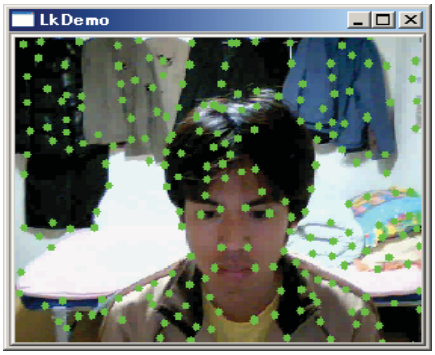

Fig. 2. Feature points in a front camera image.

Fig. 3. Feature points in a side camera image.

experiments. They are text reading experiments and video watching experiments. And last, we conclude this work.

II. Mind State Estimation Method

A. Measurement of Body Movements

In our experiments, subjects sit on a chair. There may be motions on legs. However, it is difficult to observe the legs with videos. The observation system contains 2 cameras. Using 2 cameras, we can easily measure the 3-dimensional body motions. One camera was located in front of the subject and the other was located at the side of the subject in order to capture the movement of upper half of the body.

Our body motion measurement system is a construction of 4 major steps as shows in Fig. 1. There is a work that uses Lucas & Kaneda method for fast tracking of a human head [8]. For tracking the movement of the body, we use the Lucas & Kaneda method implementation based on the LKdemo that is a sample program distributed by OpenCV [9]. Lucas & Kaneda technique uses the optical flow equation for a group of adjacent pixels and assuming that all of them have the same velocity. The optical flow computation task is reduced to solving a linear system. LKdemo recognizes and tracks every corner point that included in an image. Fig. 2 shows the tracked feature points on the original frame in the front camera video image. Fig. 3 does the feature points on the original frame in the side camera video image. In the images, there are feature points on both of the subject's body and background. However, we need to select only the points on the subject's body. To solve this problem, we used the K-means clustering technique on the motions of the feature points. It shows large motion the feature points that exist on the subject's body. Using K-means clustering on motions of the feature points, we select the clusters that show the large motions as parts of the body. Figure 4 shows the selected cluster points in the front camera image. Figure 5 does the selected cluster points in the side camera image. In the figures, we confirm that the selected clusters show the part of the subject's body. The amount of a body movement at frame number i is defined as (1).

$$\mathrm{ABM} = \sum_{k \in AFP} |Pfp_i(k) - Pfp_{i-1}(k)| \,/ |AFP| \qquad (1)$$

In the equation, ABM is the amount of a body movement. AFP is the set of all feature points in the area. An area may show the head. $Pfp_i(k)$ is the position of the feature point k in the i-th frame. ABM is the average amount of feature points' movements on a body between the frame i and the frame $i+1$.

B. Estimation of Mind State

We use the correlation analysis method to analyze the relation between the answer to the questionnaire and the observed body motions. With this analysis method, we have the

correlation coefficients. We use these correlation coefficients as the measure of reliability of the estimation.

The questionnaire used in the experiments is shown as follows (Rated 1 = low to 5 = high):

Q1. Interest in the story
Q2. Difficulty of story
Q3. Concentration while reading
Q4. Understanding of story
Q5. Willing to read the rest of the story
Q6. Relaxing while reading
Q7. Length of session
Q8. Satisfaction of the overall

If the Q1 (interest in the story) and Q5 (willing to read the rest of the story) were high rated, there is a possibility that degree of concentration (Q3) is high too. If Q6 (relaxing while reading) was high rated, subject is in the state of relaxed, we predict that subject would rarely change his pose. On the other hand, if Q2 (difficulty of story) and Q4 (understanding of story) were high rated, which means subject could not understand the story, because it is difficult. In this case, subject would impair an ability to concentrate and body motion tends to increase. If Q5 (willing to read the rest of the story) and Q7 (length of session) were high rated, that means subject does not want to read the story anymore, and felt that experiment time (10 minutes) is too long. These states of mind could reflect to the body motion.

III. EXPERIMENT

A. Text Reading Experiment

There were 6 participants carried out the 32 sets of the experiments. Each experiment set is reading a comic or an English textbook. These works change the mind state on the subject. Participants are Chinese, Gabonese, Japanese, Lao, and Malaysian. They are not the native English speakers. 20 sets of the experiments are with a comic. The others were with English textbook. Each set takes 10 minutes and contains 10000 to 17000 frames.

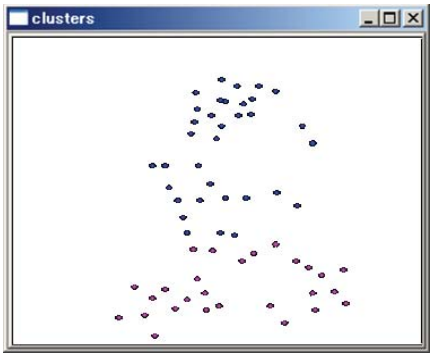

Fig. 4. Selected feature points in a front camera image.

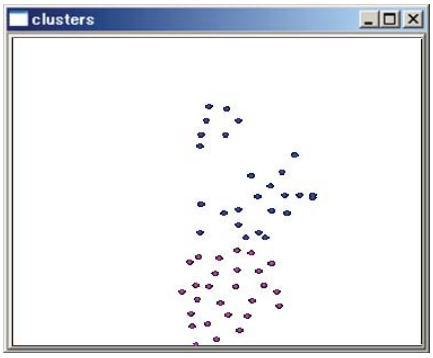

Fig. 5. Selected feature points in a side camera image.

B. Mind state estimation on text reading experiments

1) Questionnaires

There is a dependency among questionnaires. Table I shows the correlation matrix among questionnaires.

The Q1 has large correlations with Q2, Q3, Q4, Q5, Q7 and Q8. The Q6 has no large correlations with others. This shows that the Q1, Q2, Q3, Q4, Q5, Q7 and Q8 depend on enjoyment

TABLE I. Correlation Matrix among Questionnaires

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
Q1	1.00	-0.88	0.75	0.89	0.93	0.48	-0.87	0.91
Q2	-0.88	1.00	-0.80	-0.91	-0.85	-0.47	0.78	-0.85
Q3	0.75	-0.80	1.00	0.74	0.72	0.44	-0.71	0.81
Q4	0.89	-0.91	0.74	1.00	0.80	0.39	-0.73	0.79
Q5	0.93	-0.85	0.72	0.80	1.00	0.56	-0.91	0.88
Q6	0.48	-0.47	0.44	0.39	0.56	1.00	-0.52	0.54
Q7	-0.87	0.78	-0.71	-0.73	-0.91	-0.52	1.00	-0.90
Q8	0.91	-0.85	0.81	0.79	0.88	0.54	-0.90	1.00

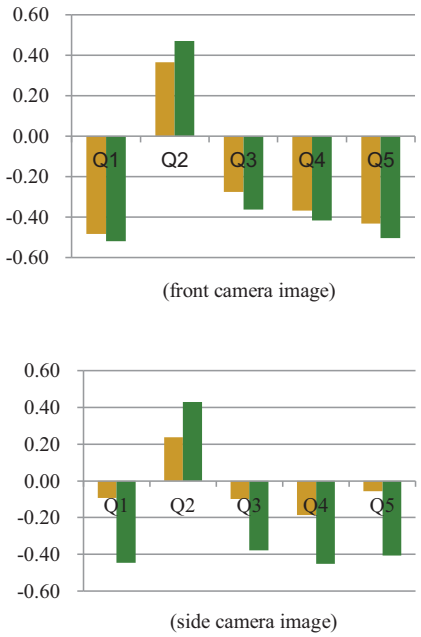

Fig.6. Correlations between amount of body motions and questionnaires

of a subject. We expected that all the questionnaires have large correlations. However, the Q6 differs from other questionnaires. Subjects' relaxing may be other factors.

2) Using Average Amount of Body Motion

We show the correlations between the answers to the questionnaires and the body movements observed in Fig. 6. The left bar represents the horizontal motion. And, the right bar does the vertical motions. In Fig. 6, there shows only Q1 to Q5. In Q6 to Q8, there are little correlations. The vertical motion correlates with every question item more than horizontal motion does. The vertical motion has a strong positive correlation coefficient with Q2 (difficulty of story), Q7 (length of session) and has a strong negative correlation coefficient with Q1 (interest in the story), Q5 (willing to read the rest of the story) and with Q8 (satisfaction with the overall). These absolute values of correlation coefficients are greater than 0.5. We also have correlation coefficients with Q3 (concentration while reading) and Q4 (understanding of story). These absolute values are greater than 0.4. In contrast, all most correlation coefficients with horizontal motion are smaller than 0.4 except with Q2 and with Q7.

The lower graph in Fig. 6 shows the result that observed from the side camera. The vertical motion correlates with every question item more than back and forth motion does. The observed vertical motion must be same in the front camera and the side camera. The result in the side camera is similar to the result obtained from the front camera. The difference is the correlation between back and forth motion. The correlation between questionnaire and the body motion between back and forth is weak.

The vertical motion of a body must be similar with any camera position. However, the observable feature points differ

TABLE II. FREQUENCY ZONES.

Zone	Lower limit(Hz)	Upper limit(Hz)
A	0	1
B	1	1.4
C	1.4	2
D	2	2.8
E	2.8	4
F	4	5

TABLE III. CORRELATIONS IN FREQUENCY DOMAINS.

	Hz	*Q1*	*Q2*	*Q3*	*Q4*	*Q5*	*Q6*	*Q7*	*Q8*
Front Camera x-axis	0~1.0								
	1.0~1.4								
	1.4~2.0						-0.449		
	2.0~2.8	-0.452			-0.462	-0.452			-0.449
	2.8~4.0	-0.454	**0.550**	-0.453	-0.451	-0.449		0.443	-0.443
	4.0~5.0	-0.444		-0.449	-0.439			0.434	
Side Camera y-axis	0~1.0				-0.442	-0.444			
	1.0~1.4								
	1.4~2.0								
	2.0~2.8								-0.447
	2.8~4.0		**0.539**	-0.443	-0.452				
	4.0~5.0			-0.455	-0.437				

with the camera positions.

From these results, the average amounts of stretching body up and down obtained from the front camera and the side camera correlates with the state of mind, especially the front camera. The state of mind may be 'enjoyment'. If a large amount of stretching up and down is observed with the front camera, the subject feels bored with what he does. The results were close to our expectation.

The body movements between back and forth carries a little information about the subjects' mind state. For the purpose to estimate the mind state, this experiments shows that the single front camera is enough.

3) Using Amplitudes of each Frequency Component of Body Motion

We discuss the correlation coefficients between questionnaire results and the amplitudes of each frequency range of body motion. Some body movements may represent the state of mind well. There are many types of body movements. The direction is one category. The speed of a body's motion is one of the other categories. In our experiments, a chair restricts the range of body motions. As a result, a fast body motion leads the fast change of body motion. The speed of a body's motion is described in the frequency domain.

We divide the frequency into 6 zones. Table II shows the zones. Table III shows the correlations between frequency zone and questionnaires. In table III, there are frequency zones where the correlation is larger than 0.4. The '-' shows that the correlation is reverse correlations. The block character shows that the correlation is larger than 0.5. Both of horizontal and vertical motions from the front camera, all question items have correlation with the amplitude of body motion in the frequency range from 2.8Hz to 4.0Hz. The amplitude of horizontal motion also correlates to all question items in frequency range from 2.0Hz to 2.8Hz while the vertical motion does not. The amplitudes of body motion correlate with the state of mind in the frequency range from 2.8Hz to 4.0Hz. This enables to estimate the state of mind, such as "Interest", "Difficulty", "Concentration" and "Understanding" with the observation of the amplitudes of a body motion from the front camera. For instance, estimate for the difficulty of what subject is reading can be done with the observation of the level of amplitude of moving body between left and right, and stretching body up and down in the frequency range from 2.8Hz to 4.0Hz.

From the side camera, we can see, Q2 (difficulty of story) correlates to amplitudes of body motion in all range, that is different from the result of the front camera. Moreover, in the

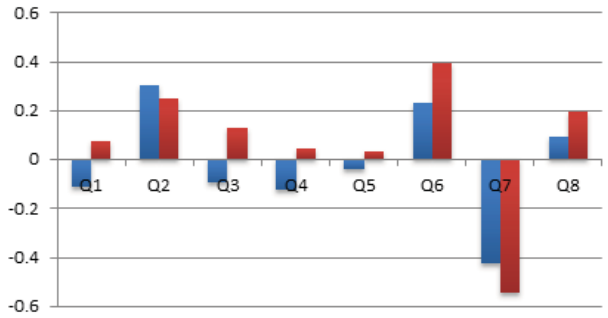

Fig. 7. The correlations between the amount of body motions and the answers of the questionnaires in a front camera.

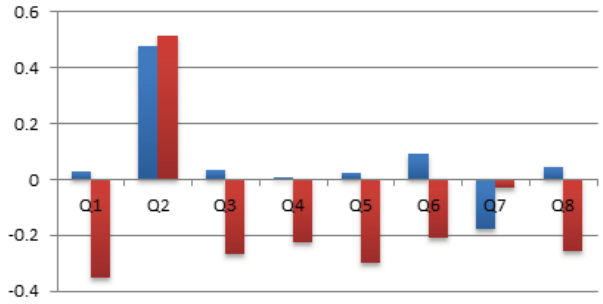

Fig. 8. The correlations between the amount of body motions and the answers of the questionnaires in a side camera.

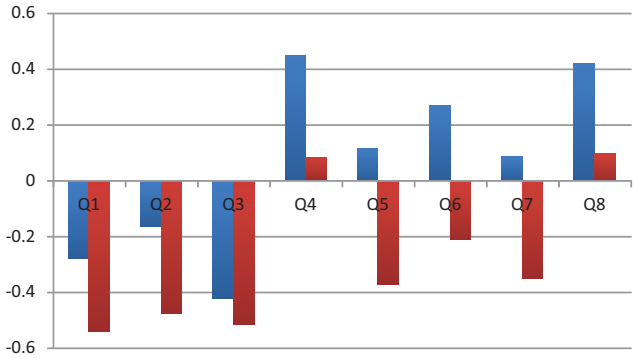

Fig. 9. Side camera results of Japanese students (2.8~4Hz).

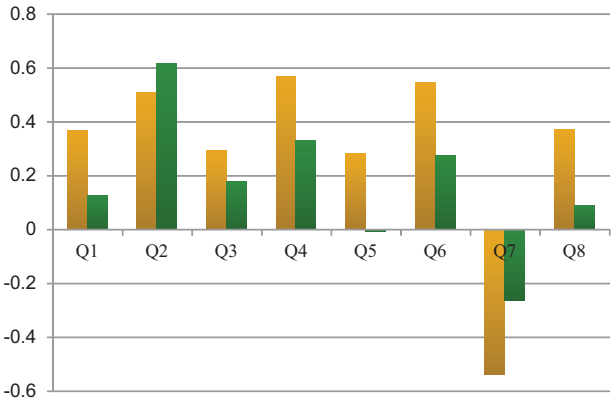

Fig. 10. Side camera results of non-Japanese students (2.8~4Hz).

frequency range from 1.0Hz to 1.4Hz, the amplitude correlates with the question items from Q2 to Q5. That means, if we focus on the low-frequency motion (less than 1.4Hz), it is possible to estimate of "Interest", "Difficulty", "Concentration" and "Understanding" with the observations of the level of stretching body up and down from the side camera.

C. Video watching experiments

In this experiment, each subject watches the English lecture video and the animation in 10 minutes. 2 video cameras observe the subject from the front and the side of the subject. After watching the videos, the subject answers the same questionnaires in the text reading experiment.

There are 11 subjects. Participants are Chinese, Gabonese, Japanese, Lao, and Malaysian. 6 subjects are Japanese.

D. Experimental results on video watching experiments

In the experiments, there are the high correlations between the answers to the questionnaires and the measured body movements comparing with the former experiments. There is over 0.7 correlations at the group of Japanese subjects. There are the differences between Japanese subjects and other subjects. This may represent the cultural differences.

1) Using Average Amount of Body Motion

Figure 7 and 8 show the correlations between the total amount of body motions and the answers to the questionnaires. In the figures, the left bar is the horizontal motion, and the right bar is the vertical motion. In this experiment, there is an apparent correlation between the Q7 and the body motion in the front camera. The correlation represents that the subject with more body motion does not feel the session length. This is different from the previous text reading experiments. The difference of the lengths of the sessions may affect the results.

2) Using Amplitudes of each Frequency Component of Body Motion

In the frequency domain, there is an apparent difference between Japanese subjects and the other subjects. Figure 9 and Fig. 10 show the correlation between questionnaires and the power of the body motions in the frequency zone from 2.8 Hz to 4 Hz of the side camera. This frequency zone shows strong correlations in the previous experiments. In Q1, Q2 and Q3, there is a reverse features in Japanese subjects and non-Japanese subjects. Japanese subjects move less in interesting, difficult and concentrating. However, the other subjects move more in interesting, difficult and concentrating. On the ability of English, a non-Japanese subject is better than a Japanese subject. This may affect the result of this experiment.

IV. CONCLUSION

This paper proposes the method to estimate the state of mind based on the motions of a body with video observations. The experimental results show that some body movements correlate some mind states. The video observation doesn't affect the state of the mind of a subject. There is no need of attached sensors or markers. There is no calibration beforehand. The proposed method may estimate the mind states of many persons simultaneously with a small number of video cameras. There will be the vast range of applications of the proposed method.

The resulting correlations between body movements and mind states in a frequency domain differ from the ones in the average amount of body motions. This shows that the average amount of body movements is not enough for representing the mind states.

The experiment shows the differences between Japanese students and non-Japanese students. This result is interesting. This may be the difference based on the English ability. This may be the difference based on the culture. We will make much more experiments about the differences between Japanese subjects and others.

ACKNOWLEDGMENT

This work was partially supported by Japan Society for Promotion of Science (JSPS) under Grants-in Aid for Scientific Research (B) (No. 22330260).

REFERENCES

[1] Hiroyasu Sakamoto, Sayuri Matsuoka, Shigehiro Fukuchima, Hiroshi Kadoka, Toshio Sakata, "Recognizing Emotions by Analyzing Facial Thermal Images Using Subspace Methods", IWAIT2010, ID 185, 2010.

[2] Tomoya Kurokawa, Kiyoshi Nosu, Kiyoyuki Yamazaki, "A Research on the Time Sequential Estimation of Emotions of e-Learning Users by the Face and Biometric Information", The Institute of Image Information and Television Engineers, 61, 12, pp.1779- 1784, 2007.

[3] M. Yeasin, B..Bullot, R. Sharma, "From facial expression to level of interest: a spatio-temporal approach", 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2, pp.922 -927, 2004.

[4] Masood Mehmood Khan, Michael Ingleby, Robert D. Ward, "Automated Facial Expression Classification and affect interpretation using infrared measurement of facial skin temperature variations", ACM Transactions on Autonomous and Adaptive Systems, 1, 1, pp.91-113, September, 2006.

[5] Timo Partalaa, Veikko Surakkaa, Toni Vanhalaa, "Real-time estimation of emotional experiences from facial expressions", Interacting with Computers , 18, 2, pp. 208–226, , March 2006.

[6] Clayton Epp, Michael Lippold, Regan L. Mandryk, "Identifying emotional states using keystroke dynamics", CHI '11 Proceedings of the 2011 annual conference on Human factors in computing systems, pp. 715-724, 2011.

[7] Souksakhone Bounyong, Kyota Aoki, "Estimation of State of Mind Based on Body Motion Using Cameras", Proceedings The 2012 International Workshop on Advanced Image Technology, pp.420-425, Ho Chi Minh City, Vietnam, 2012.

[8] Jianzheng LIU, Chang LIU, Zheng ZHAO, "Head Gesture Recognition Based on LK Algorithm and GentleBoost", Advances in information Sciences and Service Science, 4, 5, pp. 158-167, 2012.

[9] "OpenCVWiki", http://opencv.willowgarage.com/wiki/, August 2011.