XXI IMEKO World Congress “Measurement in Research and Industry”
August 30 – September 4, 2015, Prague, Czech Republic

EVALUATION OF USABILITY FOR CURSOR CONTROL FROM ELECTROENCEPHALOGRAPHY

Osamu Koga [1], *Natsue Yoshimura* [1 2], *Hiroyuki Kambara* [1], *Yasuharu Koike* [1 2]

[1] Precision and Intelligence Laboratory, Tokyo Institute of Technology, Yokohama, Japan,
yoshimura@, kambara@, koike@, pi.titech.ac.jp
[2] Integrative Brain Imaging Center, National Center of Neurology and Psychiatry, Kodaira, Japan

Abstract **–** Evaluating usability of user interfaces from biosignals has attracted attentions. This study aims to evaluate usability of non-user-friendly interfaces through electroencephalography (EEG) signals that may reflect users' frustration with the interfaces. Using a cursor control task with irregularly moving cursors to elicit frustration, our results showed accuracy of 86.3% in classifying brain activities during using irregularly and normally moving cursor.

Keywords: Brain-machine interface, human-computer interaction, usability, EEG, emotion

1. IINTRODUCTION

Emotion recognition is one of the sophisticated ways to evaluate human-computer interaction quality. Among surrounding numerous number of computers in daily life, sometimes we are forced to use complicated, unfriendly and irritating interfaces. Lack of methods for measuring feelings related to usability may have the shared responsibility of such interfaces. Therefore, there is a real need for establishing an objective and quantitative evaluation method for interfaces.

How can we measure feelings aroused from usability problem? Human activities can be modeled as a decision making system defined by many parameters. Humans decide their actions, e.g. press a button or pull a lever, to interact with an interface, which results in changes of the interface's state, e.g. a screen changes or a machine starts. This change will affect to humans' parameters via recognition and they will plan a next action to cope with the modified external state. Emotions can be considered as one of parameters humans have in this model [1]. For example, one would feel comfortable with an easy-to-use interface, and on the other hand, one would be irritated with an ill-behaving interface. If we could evaluate these emotional state appropriately it would be an effective measurement for usefulness.

There have been numerous studies concerned with emotion recognition from biological signals [2]. Among these studies, the use of electroencephalography (EEG) has recently attracted attentions [3]. EEG is considered to be capable to measure emotional states directly than other methods since the brain controls human activities. Li and Lu focued on high gamma band EEG and showed that “Happiness” and “Sadness” can be distinguished at over 90% accuracy [4]. In addition it is known that there are asymmetry of brain activity in frontal region when emotions are avoked [5], and Petrantonakis and Hadjileontiadis utilized this feature to distinguish 6 emotions [6, 7]. However, these studies did not cover emotions caused by non-user-friendly interfaces, and it seems to be necessary to establish an evaluation method of such emotions to improve usability of interfaces.

In this paper, we focused on emotions related to usability, such as “Unpleasant” or “Irritating”, which represents a state that interfaces do not work as intended and users are feeling frustration with them. Our goal is to distinguish such frustrating states from calm state, which enables us to detect unfriendly designed interfaces and provides a cue to improve them. Using a target-reaching task that was designed not to be user-friendly, we evaluated EEG data during the task using a machine learning classification method.

2. METHODS

2.1. Participants

Three healthy individuals (2 males and 1 female, 22-25 years old, right-handed) participated in this study. Written informed consent was obtained from all the participants prior to the experiment. The experimental protocol was approved by the ethics committee of Tokyo Institute of Technology.

2.2. Procedures

A target-reaching task on the computer was employed to evaluate cursor control usability. A red cursor was located on the screen and participants can control the cursor using a mouse. Participants were instructed to move the cursor onto a target denoted as a blue circle.

During the task, cursor was not exactly at the mouse pointer's location but transformation was applied depending on the task. There are to types of transformation: “Rotation” and “Acceleration.” “Rotation” applies rotation around the origin; when mouse pointer is at (X, Y) and rotation angle is θ then the cursor is shown at ($X\cos\theta + Y sin\theta$, $Y cos\theta - X sin\theta$) (Fig. 1(a)). In “Acceleration” task, moving speed of cursor become faster as mouse pointer is more distant from the origin. When the distance of mouse pointer is $d = \sqrt{X^2 + Y^2}$, the cursor is shown at (KdX, KdY) where K denotes acceleration coefficient (Fig. 1(b)).

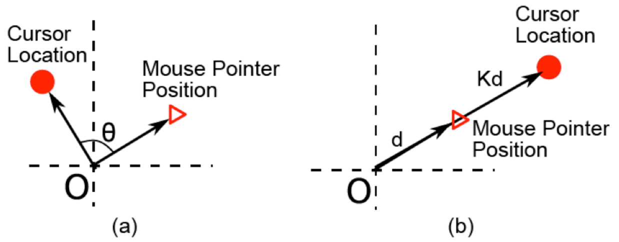

Fig. 1. Axis transformation between mouse pointer and cursor. (a) "Rotation" transformation, (b) "Acceleration" transformation.

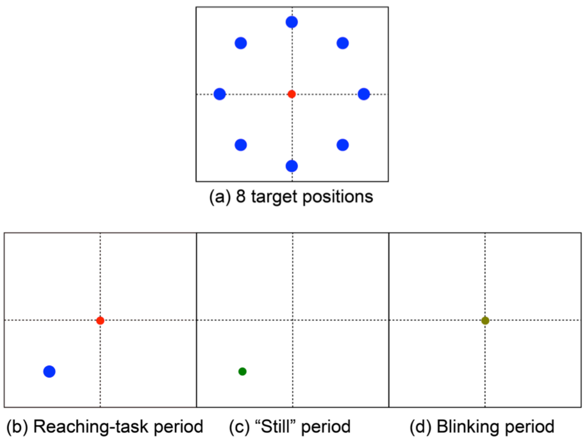

Fig. 2. Positions of the 8 reaching targets (a) and appearances of the three experimental periods. (b) Reaching-task period, (c) "Still" period, and (d) Blinking period.

We used 4 types of transformation for experiment: "Rotation ($\theta = \pi/3$)", "Rotation ($\theta = -\pi/2$)", "Acceleration (K = 0.02)" and "Normal (no transformation)." The main reason to use these transformations was to elicit differently arousing feeling caused by non-user-friendly interfaces. Since "Rotation" is easy to learn the rotation pattern and may not work effectively after a few sessions, we employed 2 "Rotation" tasks of different angle to avoid learning effect.

We had 5 sessions for each transformation, resulting in 20 sessions per participant. One session consists of 8 trials, and 1 reaching task was performed on a trial. In each trial, target position was randomly chosen from 8 points located every 45 degrees apart in a 300 pixel radius circle centered at origin (Fig. 2(a)), and all of the 8 points were chosen once in 1 session. The cursor was initially placed at the origin (Fig. 2(b)). After each reaching was completed, a 2-second "Still" period immediately followed (Fig. 2(c)). During the "Still" period, participants were instructed to gaze at the same position until the reached target was disappeared. Then a 3-second "Blink" period followed (Fig. 2(d)), which participants were allowed to blink freely. We used EEG signals during the "Still" period for analysis to avoid using signals containing motion intention. We expect feelings caused by the tasks last during the period.

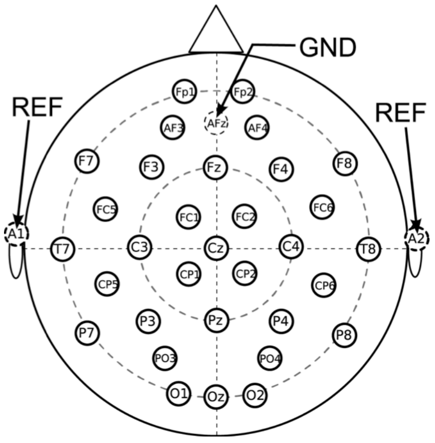

Fig. 3. EEG electrode positions

2.3. Experimental setup

A track-ball type mouse was provided to control the cursor. Participants were instructed to manipulate mouse by right hand and not to change the orientation of the mouse or their hands during the task.

EEG was simultaneously recorded during the task. We used g.tec USBamp for EEG recording with 32 electrodes placed according to the extended international 10-20 system as shown in Fig. 3. Sampling frequency was 256 Hz and 50 Hz notch filter was applied during recording. The head position was restrained using a chin rest.

2.4. Signal processing

EEG signals during the 2-second "Still" period were used for further analysis. We used a method based on [8] for feature extraction. A 256-point FFT was applied to the EEG signals for each 250 ms windows, which were overlapped by 125 ms windows, then the first 40 bins (ranged 1-40 Hz) of mean spectrum was used as features for classification. This feature extraction was performed on each signal from single electrode as well as differential signals between neighbouring electrode pairs.

2.5. Classification

To evaluate the possibility for extracting feelings caused by non-user-friendly interface, we conducted classification analysis. Each trial data was labelled according to the axis transformations and the labels were predicted using a machine learning technique. Note that because there were two types of "Rotation" trials, we mixed those trials and randomly eliminated half of them. Using support vector machine (SVM) [9], 10-fold cross validation was performed to calculate classification accuracy for the prediction. Since there were 40 trials for each class, 36 trials × 2 classes (Normal vs. Rotation, or Normal vs. Acceleration) were adopted for training and 4 trials × 2 classes were for verification, respectively. We focused on classifications using only frontal and parietal electrodes according to previous researches regarding negative emotions [5, 10, 11, 12]. Net accuracy was calculated by averaging classification

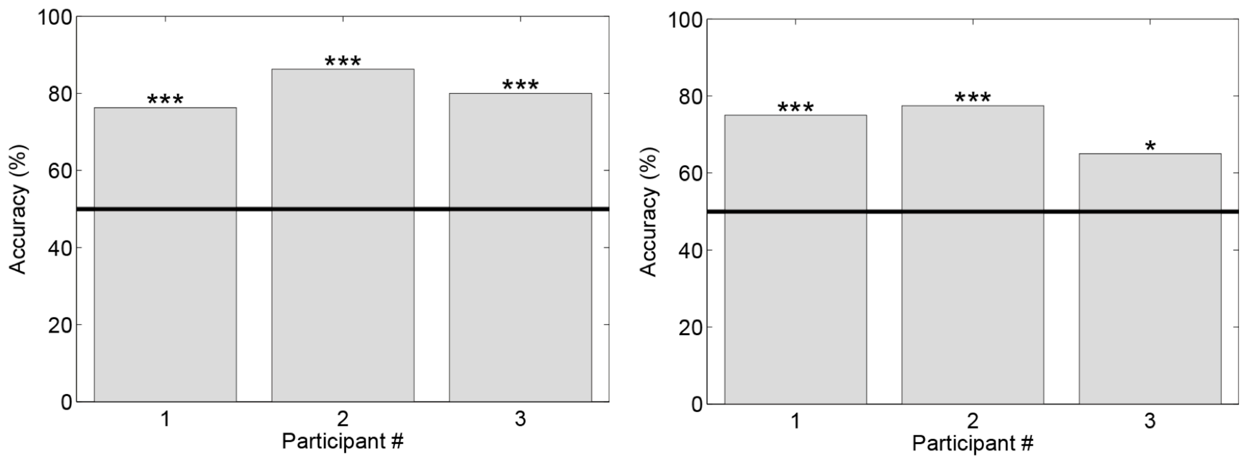

Fig. 4. "Normal" vs. "Acceleration" classification accuracies. (left) Using C3 – Cz. (right) Using AF3 – AF4. Black horizontal lines indicate chance level of binary classification. **p* < 0.05, ***p* < 0.01, ****p*<0.001.

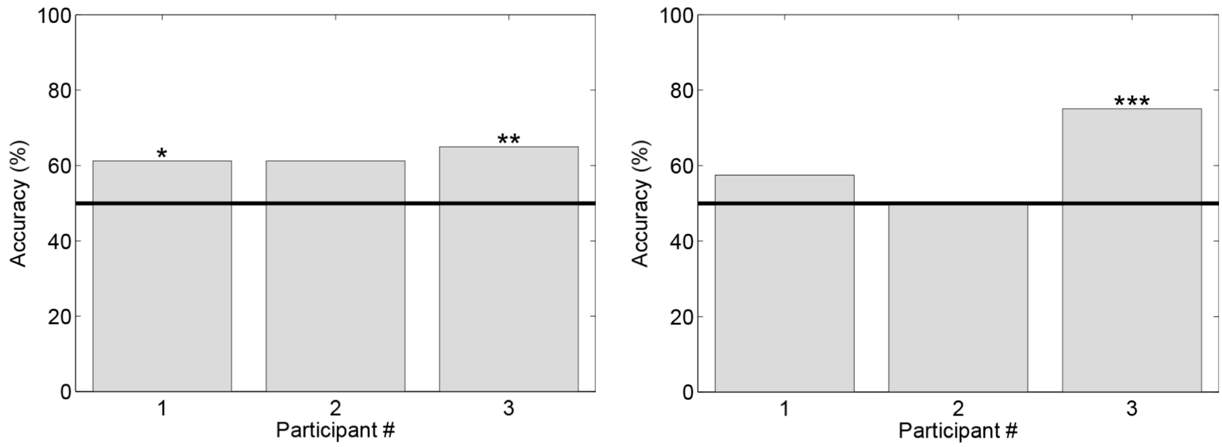

Fig. 5. "Normal" vs. "Rotation" classification accuracies. (left) Using C3 – Cz. (right) Using AF3 – AF4. Black horizontal lines indicate chance level of binary classification. **p* < 0.05, ***p* < 0.01, ****p*<0.001.

accuracy of respective iterations. Welch's t-test was conducted on the results to test significance.

3. RESULTS

3.1. Normal vs. Acceleration

Classification accuracies for "Normal" vs. "Acceleration" transformation are shown in Fig. 4. For all participants, accuracies using differential signal between parietal electrodes (C3 – (minus) Cz) were significantly higher than the chance level. Feature vector of (C3 – Cz) worked well on every participants (Participant #1: 76.3% ($p = 2.97\times10^{-4}$), Participant #2: 86.3% ($p = 3.50\times10^{-6}$), Participant #3: 80.0% ($p = 1.01\times10^{-4}$). Differential signal between frontal electrodes (AF3 – AF4) also showed significantly higher accuracy than chance level for all participants even though the accuracies were lower than that of C3 – Cz. (Participant #1: 75.0% ($p = 4.39\times10^{-5}$), Participant #2: 77.5% ($p = 8.66\times10^{-5}$), Participant #3: 65.0% ($p = 0.0330$)).

3.2. Normal vs. Rotation

Classification accuracies for "Normal" vs. "Rotation" transformation are shown in Fig. 5. Participant 2 did not show significance for both differential signals (C3 – Cz: 61.3% ($p = 0.0967$), AF3– AF4: 50.0% ($p = 0.50$)). Other participants also showed low classification accuracy except AF3 – AF4 for participant #3 (75.0%, $p = 1.96\times10^{-4}$).

4. DISCUSSION

4.1. Is it certainly frustration?

Since our results showed significantly high classification accuracy in Normal vs. Acceleration classification, the results suggest that brain activity varied across the axis transformations of the cursor movement. In addition, we used EEG signals that were not contaminated by artifacts such as hand movements or electrooculograms, therefore our results would reflect sole activity elicited by the tasks.

4.2. How appropriate electrode positions are?

Under the Pleaeure-Arousal-Dominance (PAD) model [13] and its two-dimensional variation [14, 15], which employes "Arousal" (Intense – Calm) and "Valence" (Positive – Negative) as its orthogonal axes, emotions related to unusefulness can be considered as high arousal and low valence. In this paper, C3 – Cz showed high accuracies for all participants. This result seems to be comparable to several researches that show the relation between "Unpleasant" image stimulation and parietal regions [11, 12].

Electrodes in frontal regions, AF3 – AF4, were also found to be effective for classification in this research, even its degree of contribution varied across participants. Regarding this perspective, Kostyunina and Kulikov [10] reported that power of α wave significantly increase in F3, T4 and O1 when feeling "Anger", which is said to be stronger feelings than "Annoyance" or "Irritation" [16]. "Disgust" feeling, which also has high arousal and low valence, also can be distinguished from calm state by right frontal electrodes [5]. That may be the reason why AF3 – AF4 signals showed relatively high accuracy in our results because the differential signals enhanced activity difference between the right and left hemisphere.

These facts support our hypothesis that the target-reaching tasks used in this study were appropriately non-user-friendly and elicited negative emotions.

4.3. Difference between "Rotation" and "Acceleration"

Despite of the participants' report that "Rotation" was more frustrating, classification performance of "Rotation" was apparently worse than "Acceleration." This result might be because our experimental design. Note that "Rotation" is difficult at the initial movement by uncertainty of controlling, while "Acceleration" is becoming difficult towards the end of the task period because it is hard to stop the cursor exactly on the target due to its high speed. In our experiment, EEG data used for evaluation was collected immediately after reaching tasks were completed, thus it would include frustration elicited by "Acceleration" but not by "Rotation."

5. CONCLUSIONS

In this paper, we aimed to evaluate the usability from EEG. Using a target-reaching task based on cursor control, we applied axis transformations onto cursor movement to introduce non-user-friendliness. Analysis using an FFT-based feature extraction and SVM classification revealed that electrodes in parietal and frontal regions, which were supposed to related with negative emotions, were capable to distinguish the feelings during non-user-friendly tasks from normal tasks.

On the other hand, one of the axis transformation conditions could not be sufficiently distinguished by those electrode positions. It would be because of our experimental design and its improvement is required considering the needs for building a general model, which is useful in cross-participant evaluation or building easy-to-use application.

ACKNOWLEDGMENTS

This research was partially supported by JSPS KAKENHI, Grant-in-Aid for Scientific Research on Innovative Areas from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (26112004 and 15H01659), and the Center of Innovation Program from Japan Science and Technology Agency, JST.

REFERENCES

[1] P. Rani, N. Sarkar, C. A. Smith, and J. A. Adams, "Affective communication for implicit human-machine interaction", *IEEE International Conference on Systems, Man and Cybernetics 2003*, IEEE, vol. 5, pp. 4896-4903, Oct. 2003.

[2] S. Jerritta, M. Murugappan, R. Nagarajan, and K. Wan, "Physiological signals based human emotion recognition: a review", *2011 IEEE 7th International colloquium on Signal Processing and its Applications*, IEEE, pp. 410-415, March 2011.

[3] S. G. Mason, A. Bashashati, M. Fatourechi, K. F. Navarro, and G. E. Birch, "A comprehensive survey of brain interface technology designs", *Annals of Biomedical Engineering*, vol. 35, no. 2, pp. 137-169, Springer, Feb. 2007.

[4] M. Li, and B. L. Lu, "Emotion classification based on gamma-band EEG", *Annual international Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1223-1226, Sept. 2009.

[5] R. J. Davidson, P. Ekman, C. D. Saron, J. A. Senulis, and W. V. Friesen, "Approach-withdrawal and cerebral asymmetry: emotional expression and brain physiology", *Journal of personality and social psychology*, vol. 58, no. 2, pp. 330-341, Feb. 1990.

[6] P. C. Petrantonakis and L. J. Hadjileontiadis, "A novel emotion elicitation index using frontal brain asymmetry for enhanced EEG-based emotion recognition", *IEEE transactions on information technology in biomedicine*, vol. 15, no. 5, pp. 737-746, Sept. 2011.

[7] P. C. Petrantonakis and L. J. Hadjileontiadis, "Adaptive emotional information retrieval from EEG signals in the time-frequency domain", *IEEE transactions on signal processing*, vol. 60, no. 5, pp. 2604-2616, May 2012.

[8] O. Alzoubi, R. A. Calvo, and R. H. Stevens, "Classificatoin of EEG for Affect Recognition: An Adaptive Approach", *AI 2009: Advances in Artificial Intelligence, Lecture Notes in Computer Science, vol. 5866*, pp. 52-61, Springer-Verlag Berlin Heidelberg, Dec. 2009.

[9] C. C. Chang and C. J. Lin, "LIBSVM: A library for support vector machines", *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, pp. 1-27, Apr. 2011, Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

[10] M. B. Kostyunina, and M. A. Kulikov, "Frequency characteristics of EEG spectra in the emotions", *Neuroscience and Behavioral Physiology*, vol. 26, no. 4, pp. 340-343, Jul. 1996.

[11] B. Güntekin and E. Baar, "Event-related beta oscillations are affected by emotional eliciting stimuli", *Neuroscience letters*, vol. 483, no. 3, pp. 173-178, Oct. 2010.

[12] N. Martini, D. Menicucci, L. Sebastiani, R. Bedini, A. Pingitore, N. Vanello, M. Milanesi, L. Landini, A. Gemignani, "The dynamics of EEG gamma responses to unpleasant visual stimuli: from local activity to functional connectivity", *Neuroimage*, vol. 60, no. 2, pp. 922-932, Apr. 2012.

[13] A. Mehrabian, "Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in Temperament", *Current Physiology*, vol. 14, no. 4, pp. 261-292, Dec. 1996.

[14] J. A. Russell, "Affective space is bipolar", *Journal of Personality and Social Psychology*, vol. 37, no. 3, pp. 345-356, Mar. 1979.

[15] T. Eerola and J. K. Vuoskoski, "A comparison of the discrete and dimensional models of emotion in music", *Psychology of Music*, vol. 39, no. 1, pp. 18-49, Aug. 2010.

[16] J. R. Averill, "Studies on anger and aggression: Implications for theories of emotion", *American Psychologist*, vol. 38, no. 11, pp. 1145-1160, Nov. 1983.