International Journal of Medical, Medicine and Health Sciences
ISSN: 2517-9969
Vol:5, No:11, 2011

Identification of Arousal and Relaxation by using SVM-Based Fusion of PPG Features

Chi Jung Kim, Mincheol Whang and Eui Chul Lee

***Abstract*—**In this paper, we propose a new method to distinguish between arousal and relaxation states by using multiple features acquired from a photoplethysmogram (PPG) and support vector machine (SVM). To induce arousal and relaxation states in subjects, 2 kinds of sound stimuli are used, and their corresponding biosignals are obtained using the PPG sensor. Two features—pulse to pulse interval (PPI) and pulse amplitude (PA)—are extracted from acquired PPG data, and a nonlinear classification between arousal and relaxation is performed using SVM.

This methodology has several advantages when compared with previous similar studies. Firstly, we extracted 2 separate features from PPG, i.e., PPI and PA. Secondly, in order to improve the classification accuracy, SVM-based nonlinear classification was performed. Thirdly, to solve classification problems caused by generalized features of whole subjects, we defined each threshold according to individual features.

Experimental results showed that the average classification accuracy was 74.67%. Also, the proposed method showed the better identification performance than the single feature based methods. From this result, we confirmed that arousal and relaxation can be classified using SVM and PPG features.

***Keywords*—**Support Vector Machine, PPG, Emotion Recognition, Arousal, Relaxation

I. Introduction

RECENTLY, since emotional techniques have been applied to various industries, the study of behavioral aspects related to emotions has become a hot topic. For example, there are many published studies on emotional recognition, emotional expression, and their various products. Among them, as the classification of emotion for emotion recognition, Russell defined a two-dimensional emotion model for easily classifying complex human emotions whose 2 axes were pleasure-displeasure and arousal-relaxation [1]. In order to classify arousal and relaxation, questionnaire-answer based subjective survey methods have been widely used. However, survey-based methods have some problems caused by individual variation and variation in emotional status even within the same subject. To solve these problems, a biosignal-based objective method of emotion measurement has been adopted. In general, various emotion statuses occur with specific corresponding physiological responses. These physiological responses can be measured by acquiring biosignals. Therefore, human emotions with corresponding physiological responses can be objectively measured by analyzing features extracted from biosignals.

Among the various emotions, the classification of arousal-relaxation by measuring biosignals is possible based on analysis of either the central nervous system (CNS) or the autonomic nervous system (ANS). Examples of CNS-based measurements are electroencephalogram (EEG) and event-related potential (ERP). In contrast, ANS-based measurements include electrocardiography (ECG), photoplethysmogram (PPG), electrodermal activity (EDA), and skin temperature (SKT).

A previous study investigated the arousal-relaxation states by using the International Affective Picture System (IAPS) and classified emotional status by measuring power in alpha frequency band of EEG [2]. Additionally, Amrhein et al. classified arousal-relaxation by measuring ERP and skin conductance response (SCR) [3].

Concerning ANS changes caused by arousal and relaxation, there are cardiovascular responses and EDA. Previously, arousal level and user performance for the “stroop color-word test” were evaluated using SCR and Skin Conductance Level (SCL) [4]. Additionally, Chung et al. evaluated arousal-relaxation and pleasure-displeasure in real-time by using galvanic skin response (GSR) [5]. For arousal-relaxation classification with cardiovascular responses, blood volume pulse and ECG were used by Haag et al. [6].

These biosignal-based methods have 2 problems. Firstly, most biosignal-based studies depend on statistical significance testing by comparing mean values. Although the significance level can be calculated, the classification accuracy cannot be quantitatively measured. In addition, individual independent classification may increase the error because there are large individual variations.

The second problem with these methods is the inconvenience and intrusive nature of the research caused by the necessity of attaching sensors to the subject’s bodies. To measure biosignals, the use of sensors is inevitable. Such attachments may cause inconvenience due to many sensors and attachment sites. In particular, EEG, ECG, and EDA-based methods have the disadvantage of attaching multiple sensors onto the scalp, chest, and fingers, respectively.

To solve these problems, Support Vector Machine (SVM) based nonlinear classification and single-sensor-based PPG measurements are adopted in our proposed method.

SVM defines the optimal hyperplane by finding support vectors having maximal margins in high dimensional feature

C. J. Kim is with the Division of Fusion and Convergence of Mathematical Sciences, National Institute for Mathematical Sciences, Daejeon, Republic of Korea.

M. Whang is with the Division of Digital Media Technology, Sangmyung University, Seoul, Republic of Korea.

E. C. Lee is with the Division of Fusion and Convergence of Mathematical Sciences, National Institute for Mathematical Sciences, Daejeon, Republic of Korea. (e-mail: eclee@nims.re.kr). Dr. E. C. Lee is the corresponding author.

International Journal of Medical, Medicine and Health Sciences
ISSN: 2517-9969
Vol:5, No:11, 2011

space [7]. Consequently, SVM shows a high degree of accuracy due to its nonlinear classification. For example, a previous study recognized 4 kinds of emotions (sadness, anger, stress, and surprise) by using SVM classification and multiple features extracted from ECG, EDA, and SKT [8]. In another investigation, Nocua et al. used SVM and biosignal analysis of multiple features to evaluate fall detection [9].

PPG is a biosignal that can measure cardiovascular responses such as ECG. In the case of PPG, measuring with a single sensor is possible. In addition, since the sensor is worn on the finger or earlobe, it is minimally intrusive.

As an advantage of the above-mentioned SVM and PPG, we propose a new method for classifying 2 features extracted from PPG data on arousal-relaxation status with SVM. The features are pulse-to-pulse interval (PPI) and pulse amplitude (PA), respectively. In addition, each classifier is determined for each subject by individual SVM training to improve the accuracy of classification.

II. Proposed Method

A. Dual-Feature Extraction from PPG

PPG is a signal derived from light absorption changes in pulse oximeters in contact with the skin [10]. The change in the light absorption reflects a change in the blood flow rate. Thus, the PPG signal is subordinate to pulse pressure, i.e., the difference between systolic and diastolic pressure in the arteries. Fig. 1 shows a part of 1 PPG signal. V is the systolic pressure beginning in the left ventricle, and P is the maximum systolic pressure. As shown in Fig. 1, 2 features such as PPI and PA can be extracted from 1 cycle of PPG.

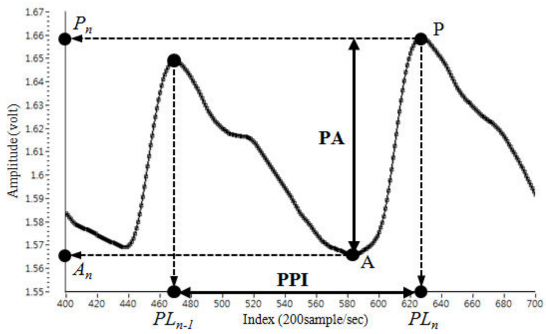


Fig. 1 Concepts of PA and PPI in PPG signal

The PPI is the time between 2 successive peaks of PPG. When the n^{th} location of the peak is PL_n, the PPI_n is calculated, as shown in (1). Additionally, the PA is a potential difference between peak and valley in the pulse. The PA_n subtracts V_n from n^{th} P_n, as in (1).

$$\begin{aligned} PPI_n &= PL_n - PL_{n-1} \\ PA_n &= P_n - V_n \end{aligned} \tag{1}$$

B. SVM Classification

SVM is a learning algorithm based classification method, which finds the optimal hyperplane to divide data into 2 categories at the maximal margin [7]. Figs. 2 (a) and (b) show the general linear classification method and SVM, respectively. The general linear classification method obtained the center between the 2 groups, which finds hyperplane from them, as shown in Fig. 2 (a). However, SVM uses the boundary data of each group. As shown in Fig. 2 (b), SVM finds “L1” and “L2” vectors from the boundary of each group. Then, a vector having a maximal margin with “L1” and “L2” is determined as a support vector. Unlike the general linear classification, SVM can determine a nonlinear classifier by using various kernels such as radial basis function (RBF), dot, analysis of variance (ANOVA), and so on. Consequently, the SVM can calculate a curve-shaped classifier in the case of a 2-dimensional feature space.

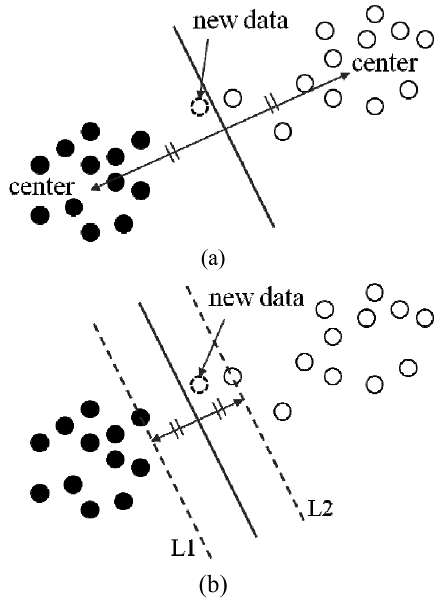


Fig. 2 Comparison of (a) general linear classification method with (b) SVM

In our research, the WinSVM software was used for the classification of arousal and relaxation on the basis of SVM [11]. The WinSVM is operated in 3 steps—optimization, learning, and prediction, consecutively. The optimization step acts to calculate the minimum mean square error (MSE) through training data. This step is for determining a kernel type and its optimal parameters. In the learning step, the data set is learned by using the previously determined kernel and parameters. That is, a maximum-margin classifier is defined by combining several support vectors in this step. Finally, the test data were verified using the previously defined classifier in the prediction step.

III. Experiments

A. Configuration

In our experiment, 5 adult subjects participated (average age: 24 ± 3 years) of whom 3 were men and 2 were women. Each subject provided informed consent and received $20 remuneration.

As the stimuli for inducing arousal and relaxation states, 4 kinds of auditory stimuli were used: 2 pieces of arousing music and 2 pieces of relaxing music. The 2 pieces of arousing music were created using electrically distorted sounds. In contrast, the 2 relaxing recordings were played on the classic guitar and Korean string instrument (Gayageum). The specifications of the arousing and relaxing music were as follows:

- Arousing Music #1
  Sound length: 2 min 37 s
  Component: uneasy electronic sounds
- Arousing Music #2
  Sound length: 3 min 5 s
  Component: taut electronic sounds
- Relaxing Music #1
  Sound length: 2 min 52 s
  Component: calm classical guitar sounds
- Relaxing Music #2
  Sound length: 2 min 59 s
  Component: quiet gayageum sounds

Before using the 4 musical stimuli, we validated whether they could be used for inducing arousal and relaxation emotions by testing on 37 subjects. Using a *t*-test, we confirmed that the arousal music evoked a different response from the relaxation music with a confidence level of 99%.

In our experiment, each subject participated in 4 trials. Each trial was carried out as shown in Fig. 3. In the "S" stage of each experiment, 1 piece of music was randomly played from the 2 arousing and 2 relaxing musical pieces.

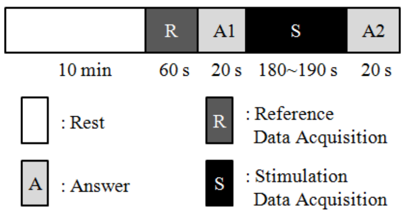


Fig. 3 Experimental sequence for 1 musical stimulus

In our experiment, the subjects sat on a chair and rested for 10 min. During the rest period, the PPG sensor was attached to the left earlobe of the subjects. After the rest period, the PPG signal was measured for 60 s, and this was used as reference data. All subjects were requested to answer a questionnaire regarding their degree of arousal (or relaxation) for 20 s before and after the measurement. During the actual music stimulation time, the PPG signal was measured. In our analysis, the results of subject's answers ("A1" and "A2" in Fig. 3) were not used because our study focused only on the PPG signal based objective measurement of emotion.

The PPG signal was recorded using the Biopac's PPG100C amplifier with TSD200A transducer of the ear-clip type. The PPG100C setting comprised a high-pass filter of 0.05 Hz and low-pass filter of 10 Hz. The National Instruments NI-DAQ 6015 A/D converter, a PPG signal measurement with the software LabVIEW 2010(National Instruments), was used to measure with a sampling rate of 200 Hz.

In our research, 2 features such as the PPI and the PA were used to classify 2 emotions. The PPI and the PA were extracted using the peak detector function of LabVIEW 2010. That function is the process of finding the locations and amplitudes of peaks. The PPI and the PA were extracted using (1) and (2) on the basis of the concept of Fig. 1.

From the PPG data, the PPI and the PA features were extracted using a sliding window concept by selecting a window size of 5 s. Since PPG data from the beginning and the end of stimulus period may be degraded by lack of concentration and boredom, we used the middle 64 s of data. Consequently, 60 PPIs and their corresponding 60 PAs were acquired from 64 s of PPG data by using the 5-s sliding window concept. Each feature was normalized on the basis of the reference data.

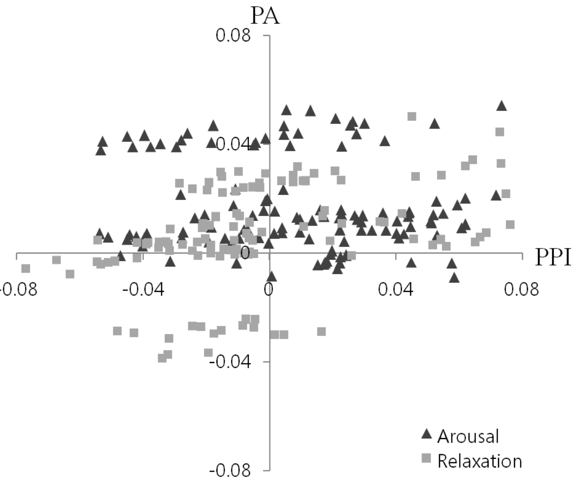
In training for SVM, 60 feature sets were divided into 2 classes consisting of 30 training data sets and 30 test data sets. However, since the amount of data was low, we used bootstrap aggregating. In detail, 30 training data points were randomly selected, and the remainder was used for testing. Using the same random selection scheme, 3 selections were performed. Consequently, 90 training data points and 90 test data points were acquired. For 1 subject's training, 180 data points for 2 arousal music stimuli and 180 data points for 2 relaxation music stimuli were used. The other 180 data points for 2 arousal music stimuli and 180 data points for 2 relaxation music stimuli were used in order to perform tests for measuring SVM classification accuracy.

B. Results

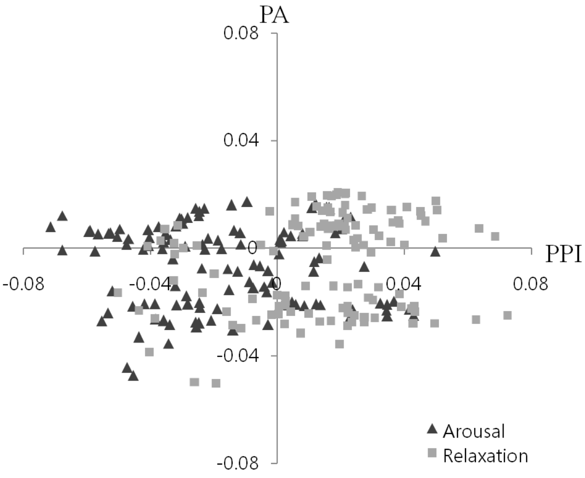
SVM for the 5 subjects were implemented individually because the patterns of physiological signals are different for each subject. As shown in Fig. 4, the patterns of PPI and the PA were different between 2 subjects. In addition, the subjects applied each optimum parameter in individual SVMs. The accuracy of prediction was calculated using output values of 3 training data sets and 3 prediction data sets per subject. The output values generated by learning were divided at arousal and relaxation by the class, and they were analyzed at each histogram. The frequency of arousal and relaxation by histogram was analyzed in terms of false arousal rate (FAR)[1] and false relaxation rate (FRR)[2]. Fig. 5 shows the receiver operating characteristic (ROC) curve for successive FAR and the FRR of the 5 subjects.

[1] The probability that the output values of the PPI and PA of relaxation is falsely classified in as arousal.

[2] The probability that the output values of the PPI and PA of arousal is falsely classified in as relaxation



(a)



(b)

Fig. 4 PPI-PA distribution of 2 subjects for arousal and relaxation (a) Subject 1 (b) Subject 2

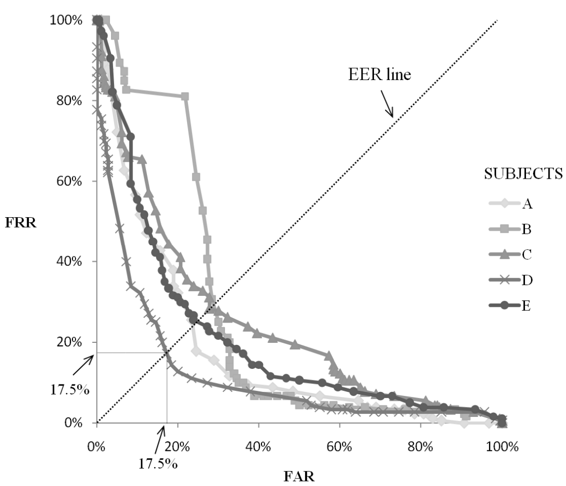


Fig. 5 ROC (Receiver Operating Characteristic) curve for the classification of 5 subjects

The output value at which the FAR became the same value as the FRR was used as a threshold to identification of arousal and relaxation. If the output value at test stage was smaller than the threshold, it was identified as an arousal state. By contrast, if the output value at test stage was greater than the threshold, it was identified as a relaxation state. The accuracy of the tests was quantified as the percentage of correctly classified values. For example, since the equal error rate (EER) of subject "D" was 17.5% as shown in Fig. 5, the overall accuracy could be determined as 82.5% (= 100% - 17.5%). Table I shows the arousal accuracy, relaxation accuracy and overall accuracy for each individual subject. The accuracies of classification for arousal sounds were obtained with an accuracy of 71.11%, 69.44%, 70.56%, 82.78%, and 72.78% from each of the 5 subjects. The classification of relaxation was acquired with an accuracy of 76.11%, 72.78%, 76.67%, 82.22%, and 72.22% for each subject. As a result, the average accuracy of 5 subjects was 74.67%.

TABLE I
ACCURACY OF INDIVIDUAL SVM

Subjects	Arousal Accuracy	Relaxation Accuracy	Overall Accuracy
A	71.11%	76.11%	73.61%
B	69.44%	72.78%	71.11%
C	70.56%	76.67%	73.61%
D	82.78%	82.22%	82.50%
E	72.78%	72.22%	72.50%

To compare our dual feature based method with single feature based ones, Table II is given. The accuracy of single feature based methods such as only using PPI or PA was analyzed by using the above used EER based scheme.

TABLE II
COMPARISON THE PROPOSED METHOD WITH SINGLE FEATURE BASED METHODS

Subjects	PPI Accuracy	PA Accuracy	SVM Accuracy
A	70.83%	44.44%	73.61%
B	66.94%	37.50%	71.11%
C	71.67%	62.22%	73.61%
D	82.22%	47.22%	82.50%
E	71.39%	51.39%	72.50%

As shown in Table II, we confirmed that the PPI based method showed better classification accuracy than the PA based one. That means that the separability of PPI is better than the one of PA. Although the separability of PA was very weak, the dual feature based method using SVM showed improved classification performance than the PPI based one. Consequently, we found that the SVM based fusion with PPI and PA was effective in terms of identification of arousal and relaxation.

IV. Conclusions

In this paper, we have proposed an identification for the arousal and relaxation states by using SVM. Two PPG features are used for the SVM. PPI and PA are extracted from the PPG signal and are fused for one feature by the SVM. In addition, we defined each threshold for the individual feature, with individual analysis for each participant. As mentioned above, this method is distinct from those used in previous similar works.

In our experiments, the accuracy was calculated by thresholds, which were selected using the output values of training data. The arousal and the relaxation music were used for the experiment. The PPG signal evoked by music was used to extract the PPI and the PA. Two features were used to distinguish between arousal and relaxation by the SVM. The SVM implemented the training and the test, and the output value obtained from training was used as the threshold for test. As a result, the proposed method was better than the single feature based methods in terms of classification accuracy.

In future work, we are planning to use biosignals other than PPG such as GSR, SKT, and so on. This is expected to raise the classification accuracy of various emotions.

Acknowledgment

This work was supported by National Institute for Mathematical Sciences (NIMS) Grant I21104 funded by the Korea Government and in part of a grant of the Korean Healthcare technology R&D Project, Ministry of Health &Welfare, Korea (A102058).

References

[1] J. A. Russell, "A circumplex model of affect," *Journal of Personality and Social Psychology*, vol. 39, no. 6, pp. 1161-1178, Dec. 1980.

[2] L. I. Aftanas, A. A. Varlamov, S. V. Pavlov, V. P. Makhnev, and N. V. Reva, "Time-dependent cortical asymmetries induced by emotional arousal: EEG analysis of event-related synchronization and desynchronization in individually defined frequency bands," *International Journal of Psychophysiology,* vol. 44, issue 1, pp. 67-82, Apr. 2002.

[3] C. Amrhein, A. Mühlberger, P. Pauli, and G. Wiedemann, "Modulation of event-related brain potentials during affective picture processing: a complement to startle reflex and skin conductance response?," *International Journal of Psychophysiology,* vol. 54, issue. 3, pp. 231– 240, Nov. 2004.

[4] C. Collet, C. Petit, A. Priez, and A. Dittmar, "Stroop color–word test, arousal, electrodermal activity and performance in a critical driving situation," *Biological Psychology*, vol. 69, issue. 2, pp. 195-203, May 2005.

[5] S. C. Chung *et al.*, "Development of the Real-Time Subjective Emotionality Assessment (RTSEA) system," *Behavior Research Methods*, vol. 39, no. 1, pp. 144-150, Feb. 2007.

[6] A. Haag, S. Goronzy, P. Schaich, and J. Williams, "Emotion Recognition Using Bio-sensors: First Steps towards an Automatic System," in *Affective Dialogue Systems : Lecture Notes in Computer Science* , vol. 3068, E. André, L. Dybkjær, W. Minker, and P. Heisterkamp, Ed. Heidelberg : Springer Berlin, 2004, pp. 36-48

[7] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proc. 5th annual workshop on Computational learning theory (COLT '92)*, New York, 1992, pp.144-152.

[8] K. H. Kim, S. W. Bang, and S. R. Kim, "Emotion recognition system using short-term monitoring of physiological signals," *Medical and Biological Engineering and Computing*, vol. 42, no. 3, pp. 419-427, May 2004.

[9] R. Nocua, N. Noury, C. Gehin, and A. Dittmar, "Evaluation of the autonomic nervous system for fall detection," in *Proc. 31th Annu. Int. Conf. of the IEEE Engineering in Medicine and Biology Society 2009,* Minneapolis, 2009, pp. 3225-3228.

[10] K. Shelley and S. Shelley, "Pulse Oximeter Waveform: Photoelectric Plethysmography," in *Clinical monitoring practical applications for anesthesia and critical care*, C. L. Lake, R. Hines, and C. Blitt, Ed. W.B.: Saunders Company, 2001, pp. 420-428.

[11] "winSVM," M. Sewell, Retrieved 29 June 2011, <http://www.cs.ucl.ac.uk/staff/M.Sewell/winsvm/>

Chi Jung Kim received the BS degree in digital media from Sangmyung University, Seoul, South Korea, in 2009. He received the M.S. degree in the Department of Computer Science at Sangmyung University. He is currently a Researcher in the Division of Fusion and Convergence of Mathematical Sciences at NIMS (National Institute for Mathematical Sciences) since April 2011. His research interests include psychophysiology, emotion engineering, attentive user interface, and HCI.

Mincheol Whang received the MS and Ph.D. degrees in Biomedical Engineering from Georgia Institute of Technology, Atlanta, Georgia, in Unite State, in 1990 and 1994, respectively. He was a Professor in the Division of Digital Media Engineering at Sangmyung University since March 1998. His research interests include human computer interaction, emotion engineering, human factors and bioengineering.

Eui Chul Lee received the BS degree in software from Sangmyung University, Seoul, South Korea, in 2005. He received the MS and Ph.D. degrees in Computer Science from Sangmyung University, in 2007 and 2010, respectively. He is currently a Researcher in the Division of Fusion and Convergence of Mathematical Sciences at NIMS (National Institute for Mathematical Sciences) since March 2010. His research interests include computer vision, image processing, pattern recognition, ergonomics, BCI and HCI.