Physiological Signals and Classification for Happiness, Neural and Surprise Emotions

Three emotion classification

Byoung-Jun Park, Eun-Hye Jang, Sang-Hyeob Kim, Chul Huh
IT Convergence Technology Research Laboratory
Electronics and Telecommunications Research Institute
Daejeon, Republic of Korea
{bj_park, cleta4u, shk1028, chuh}@etri.re.kr

Myoung-Ae Chung
IT Convergence Services Future Technology Research Dept.
Electronics and Telecommunications Research Institute
Daejeon, Republic of Korea
machung@etri.re.kr

***Abstract*—In this study, we discuss the comparative results of emotion classification by several algorithms, which classify three different emotional states (happiness, neutral, and surprise) using physiological features. 300 students participated in this experiment. While three kinds of emotional stimuli are presented to participants, physiological signal responses (EDA, SKT, ECG, RESP, and PPG) were measured. Participants rated their own feelings to emotional stimuli on emotional assessment scale after presentation of emotional stimuli. The emotional stimuli had 96% validity and 5.8 point efficiency on average. There were significant differences of autonomic nervous system responses among three emotions by statistical analysis. The classification of three differential emotions was carried out by Fisher's linear discriminant (FLD), Support Vector Machine (SVM), and Neural Networks (NN) using difference value, which subtracts baseline from emotional state. The result of FLD showed that the accuracy of classification in three different emotions was 77.3%. 72.3% and 42.3% have obtained as the accuracy of classification by SVM and NN, respectively. This study confirmed that the three emotions can be better classified by FLD using various physiological features than SVM and NN. Further study may need to get those results to obtain more stability and reliability, as comparing with the accuracy of emotions classification by using other algorithms.**

Keywords-emotion classification; physiological signals; machine learning

I. INTRODUCTION

Recent emotion recognition studies have tried to detect a human emotion by using physiological signals. It is important to detect emotions for applying on human-computer interaction system. Various physiological signals have been widely used to recognize a human's emotion and feeling because signal acquisition using non-invasive sensors is relatively simple and physiological responses are less sensitive in social and cultural difference [1]. Physiological signal may happen to artifact due to motion, and have difficulty mapping emotion-specific responses pattern, but this has some advantages, which are less affected by environment than any other modalities as well as possible to observe user's state in real time. In addition, those can be acquired spontaneous emotional responses and not caused by responses to social masking or factitious emotion expressions. Finally, measurement of emotional responses by multi-channel physiological signals offers more information for emotion recognition, because physiological responses are related to emotional state [2]. Various physiological signals offer a great potential for the recognition of emotions in computer systems. In order to fully exploit the advantages of physiological measures, standardizations of experimental methods have to be established on the emotional model, stimulus used for the identification of physiological patterns, physiological measures, parameters for analysis, and model for pattern recognition and classification [3].

The objective of this study is to achieve emotion dataset including physiological signals for three emotions (happiness, neutral, and surprise) induced by emotional stimuli and to classify three differential emotions using physiological signals. Electrodermal activity, skin temperature, electrocardiac activity, respiration and photoplethysmography are recorded from 300 peaple; during they are exposed to visual-acoustic emotional stimuli. And participants classified their present emotion and assessed its intensity on the emotion assessment scale. As the results of emotional stimulus evaluation, emotional stimuli were shown to mean 93.3% of appropriateness and 5.71 of effectiveness; this means that each emotional stimulus caused its own emotion quite effectively. In addition that, we used Fisher's linear discriminant, which is one of the linear models as a statistical method and Support Vector Machine and Neural Networks as a machine learning algorithm [4]-[7] for classifying three emotions. The comparative analysis results by those algorithms showed that the FLD exhibits higher accuracy.

This paper is organized as follows. Section II describes emotion stimuli and physiological signals of emotions. Section III deals with the classification of emotions and presents the results of classification. Finally, some conclusions are contained in Section IV.

II. PHYSIOLOGICAL SIGNALS OF EMOTIONS

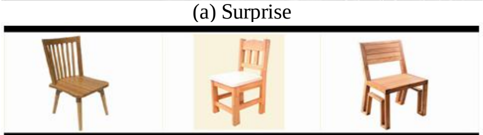
300 students (140 males, 160 females and 21.6±3.5 years) have participated in this study. They reported that they have not any history of medical illness or psychotropic

medication and any kind of medication due to heart disease, respiration disorder, or central nervous system disorder. A written consent was obtained before the beginning of the study from the participants.

The emotional stimuli, which are the 2-4 min long audio-visual film clips captured originally from movies, documentary, and TV shows, were used to successfully induce emotions (happiness, neutral, and surprise) in this study as shown in Fig. 1. Audio-visual film clips have widely used because these have the desirable properties of being readily standardized, involving no deception, and being dynamic rather than static. [8]-[11].

The audio-visual film clips that had been tested their appropriateness and effectiveness were used to provoke emotion. The appropriateness of emotional stimuli means a consistency between the intended emotion by experimenter and the participants' experienced emotion. The effectiveness is an intensity of emotions that participants rated on a 1 to 7 point Likert-type scale, that is , 1 being "least boring" and 7 being "most boring". The appropriateness and effectiveness of these stimuli were as follows; appropriateness and effectiveness of happiness were 88.0% and 5.36±1.21, 97% appropriateness and 6.06±1.05 effectiveness in surprise. In Neutral has 94.8% appropriateness.

(a) Surprise



(b) Neutral

Figure 1. Example of a stimulus for inducing an emotion.

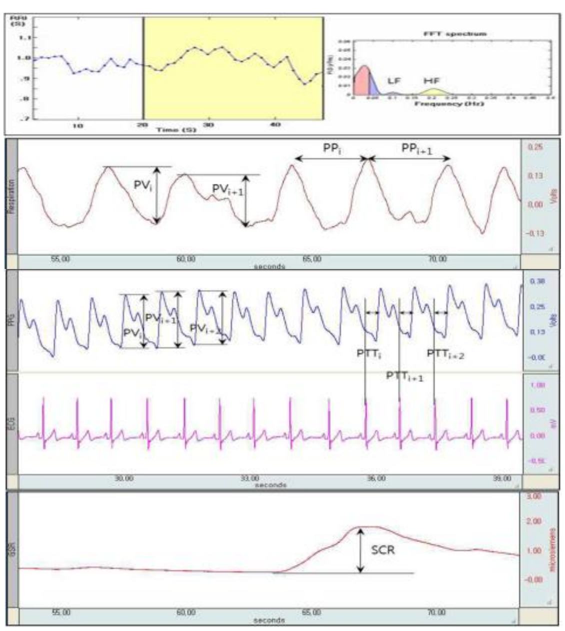


Figure 2. Analysis of physiological signals.

Prior to the experiment, we introduced detail experiment procedure to participants. They had an adaptation time to feel comfortable in the laboratory's environment and then an experimenter attached electrodes on the participants' wrist, finger, and ankle for measurement of physiological signals. Physiological signals were measured for 60 sec prior to the presentation of emotional stimulus (baseline) and for 2 to 4 min during the presentation of the stimulus (emotional state), then for 60 sec after presentation of the stimulus as recovery term. Participants rated the emotion that they experienced during presentation of the film clip on the emotion assessment scale. This procedure was repeated 3 times for elicitation of 3 differential emotions. To collect physiological signals for three emotions, the laboratory is a room with 5m x 2.5m size having a sound-proof (lower than 35dB) of the noise level where any outside noise are completely blocked.

The dataset of physiological signals such as skin temerature (SKT), electrodermal activity (EDA), photoplethysmography (PPG), respiration (RESP), and electrocardiogram (ECG) are collected by MP150 Biopac system Inc. SKT electrodes are attached on the first joint of non-dominant ring finger and on the first joint of the non-dominant thumb for PPG. EDA is measured with the use of 8 mm AgCl electrodes placed on the volar surface of the distal phalanges of the index and middle fingers of the non-dominant hand. Electrodes are filled with a 0.05 molar isotonic NaCl paste to provide a continuous connection between the electrodes and the skin. ECG electrodes are placed on both wrists and one left ankle with two kinds of electrodes, sputtered and AgCl ones. The left-ankle electrode is used as a reference.

The signals are acquired for 1 minute long baseline state prior to presentation of emotional stimuli and 1-2 minutes long emotional states during presentation of the stimuli as emotional state. To extract features, the obtained signals are analyzed for 30 seconds from the baseline and the emotional state by AcqKnowledge (Ver. 3.8.1) software (USA) as shown in Fig. 2. 10 features, namely, SCL and SCR features from EDA signal, meanSKT from SKT, HR and RRI from ECG, RESPrate, RESTamplitude, RESP_Rsd and RESP_Asd from RESP, and BVP from PPG are extracted from the obtained emotional physiological signals.

III. CLASSIFICATION OF EMOTIONS

In this study, we have used Fisher's linear discriminant (FLD), which is one of the linear models as a statistical method and Support Vector Machine (SVM) and Neural Networks (NN) for three emotion classifications.

We have carried out the related research with the same emotions (happiness, neutral, and surprise) and the same methods (FLD, SVM and NN) as an earlier work [12]. In [12], the result of LDA (FLD) showed that the accuracy of classification was 83.4%. 75.5% and 55.6% have obtained as the accuracy of classification by SVM and MLP (NN), respectively. However, the results of the previous research were obtained by only training set. Namely, the methods used the whole emotional patterns in order to build a classifier model and measure the classification accuracy of those. This approach cannot have effectiveness because of realistic and overffiting problems.

For the effectiveness verification of the previous research, we embrace more participants, which have increased by 300 people, and 10-fold cross-validation, which is a statistical method of evaluating models. In typical cross-validation, the training and validation sets must cross-over in successive rounds such that each data point has a chance of being validated against. The three algorithms were evaluated by 10-fold cross validation and the results of this study are reported for those. These values will offer a criteria index for assessing how the results of a statistical analysis will generalize to an independent dataset.

In addition those, we reduced the number of features. We used 10 features, removed LF, HF, and HRV, whereas 13 features were used in the previous work [12].

Firstly, FLD is a method used in statistics, pattern recognition and machine learning to find a linear combination of features, which characterizes or separates classes of objects or events. In FLD, the measurement space is transformed so that the separability between the emotional states is maximized. FLD finds the direction to project data on so that between-class variance ($\mathbf{S}_B$) in maximized and within-class variance ($\mathbf{S}_W$) in minimized, and then offers a linear transformation of predictor variables, which provides a more accurate discrimination [4]. $\mathbf{S}_W$ is proportional to the sample covariance matrix for the pooled d-dimensional data. It is symmetric and positive semidefinite, and it is usually nonsingular if n>d. Likewise, $\mathbf{S}_B$ is also symmetric and positive semidefinite, but because it is the outer product of two vector, its rank is at most one.

In terms of $\mathbf{S}_B$ and $\mathbf{S}_W$, the criterion function J is written

$$J(\mathbf{w}) = (\mathbf{w}^T \mathbf{S}_B \mathbf{w})/(\mathbf{w}^T \mathbf{S}_W \mathbf{w}). \tag{1}$$

This expression is well known in mathematical physics and the generalized Rayleigh quotient. It is easy to show that a vector w that maximizes J must satisfy

$$\mathbf{S}_B \mathbf{w} = \lambda \mathbf{S}_W \mathbf{w}, \tag{2}$$

for some constant λ, which is a generalized eigenvalue.
In analysis of FLD, accuracy of all emotions was 77.3 % and accuracy of each emotion had rage of 69% to 89%. Happiness was recognized by FLD with 69.1%, neutral 88.6 % and Surprise 75.5% as shown in TABLE I.

TABLE I. RESULT OF THREE EMOTION DISCRIMINANT BY FLD

	Happiness	Neutral	Surprise	Total
Happiness	69.1	30.3	0.5	100.0
Neutral	10.3	88.6	1.1	100.0
Surprise	18.5	6.0	75.5	100.0

TABLE II. RESULT OF THREE EMOTION DISCRIMINANT BY SVM

	Happiness	Neutral	Surprise	Total
Happiness	59.0	29.8	11.2	100.0
Neutral	23.4	73.7	2.9	100.0
Surprise	12.0	6.0	82.0	100.0

TABLE III. RESULT OF THREE EMOTION DISCRIMINANT BY NN

	Happiness	Neutral	Surprise	Total
Happiness	0.5	13.8	85.6	100.0
Neutral	0.6	14.3	85.1	100.0
Surprise	0	3.0	97.0	100.0

SVM is non-linear model, which are used the well-known emotion algorithms and support vector classifier separates the emotional states with a maximal margin. The advantage of support vector classifier is that it can be extended to nonlinear boundaries by the kernel trick. SVM supervised learning models with associated learning algorithms that analyze data and recognize patterns, used for classification and regression analysis. SVM is designed for two class classification by finding the optimal hyperplane where the expected classification error of test samples is minimized and has utilized as a pattern classifier to overcome the difficulty in pattern classification due to the large amount of within-class variation of features and the overlap between classes [13].

SVM finds a hyperplane based on support vector to analyze data and recognize patterns. The complexity of the resulting classifier is characterized by the number of support vectors rather than the dimensionality of the transformed space. The goal in training SVM is to find the separating hyperplane with the largest margin [4]. We expect that the larger the margin, the better generalization of the classifier. The distance from any hyperplane to a pattern $\mathbf{y}$ is $|g(\mathbf{y})|/\|\mathbf{a}\|$, and assuming that a positive margin b exists

$$z_k\, g(\mathbf{y}_k) / \|\mathbf{a}\| \geq \quad b,\ k = 1, \ldots, n; \tag{3}$$

The goal is to find the weight vector $\mathbf{a}$ that maximizes b. Here, z_k is the class of k-th pattern, b is margin and $g(\mathbf{y})$ is a linear discriminant in an augmented $\mathbf{y}$ space,

$$g(\mathbf{y}) = \mathbf{a}^T \mathbf{y} \tag{4}$$

SVM provided accuracy of 72.3% when it classified all emotions. In happiness, accuracy of 59.0% was achieved with SVM, 73.7% in neutral, and 82.0% in surprise, refer to TABLE II.

NN is a computational intelligence model inspired by the structure and functional aspects of biological neurons [4]. NN has been widely used to deal with pattern recognition problems. The generic topology of NN consists of three layers as shown in Fig. 3. A neuron in the input layer is connected to a layer of hidden neuron, which is connected to output neuron. The activity of the input neurons represents the raw information that is fed into the network, the activity of each hidden neuron is determined by the activities of the input neuron and the weights on the connections between the input and the hidden, and the behavior of the output depends on the activity of the hidden neurons and connection weights between the hidden and the output layer.

It is realized by connecting neurons and each neuron processes information using activation function,

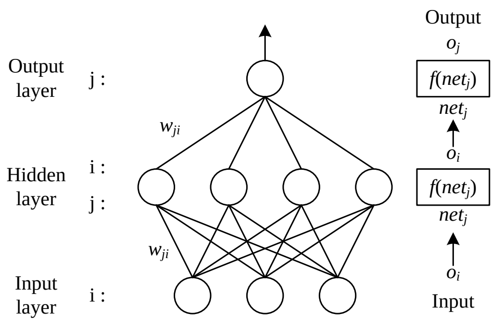


Figure 3. General structure of neural networks.

$$f_j(net_j) = 1 / (1+e^{-net_j}) \quad (5)$$

where, net_j is input of j-th neuron and computed by,

$$net_j = \sum_i w_{ji} f_i + \theta_j \quad (6)$$

where, w_{ji} is the connection weight between j-th and i-th neurons and adjusted by

$$w_{ji} \text{ (new)} = w_{ji} \text{ (old)} + \Delta w_{ji} \quad (7)$$

Δw_{ji} is calculated by back-propagation algorithm.

In TABLE III, the result of emotion recognition using NN showed that accuracy to recognize all emotions was 42.3%. NN have lower classification for happiness and neutral because of the overfitting problem for surprise. We need more research about advanced networks.

IV. CONCLUSION

This study was to identify the difference among happiness, neutral, and surprise emotions using physiological responses induced by these emotional stimuli and to find the optimal emotion recognition algorithm for classifying these three emotions. The emotion stimuli used to induce a participant's emotion were evaluated for their suitability and effectiveness. The result showed that emotional stimuli have the appropriateness of 92.5% and the effectiveness of 5.4 point (7 point Likert scale) on average.

Our result showed that FLD is the best algorithm being able to classify three emotions. FLD provides extremely fast evaluations of unknown inputs performed by distance calculations between a new sample and mean of training data samples in each class weighed by their covariance matrices. The FLD shows a recognition ratio much higher chance probability, i.e. 77.3% for three emotion categories, when applied to physiological signal databases.

The values of performance (criteria index) by 10-fold cross validation are good indicators of the generalization capabilities of the constructed models. As selecting a model, if the approximation capability of a trained model is considered only, the selected model has greatly recognition accuracy; however, it has deteriorated generalization (prediction) capability and cannot apply to a real system. Especially, this is conspicuous in nonlinear problem. This important question arises, too, as to the selection of the proper structure of the emotion recognition in this study.

Although some algorithm showed lower accuracy of emotion classification for the criteria index, the results led to better chance to recognize human emotions and to identify the optimal emotion recognition algorithm by using physiological signals. For example, this will be applied to the realization of emotional interaction between man and machine and play an important role in several applications, e.g., the human-friendly personal robot or other devices. However, for more accurate and realistic applications, a novel method to identify not only basic emotions but also more various emotions such as boredom, frustration, and love, etc. must be devised before it is mentioned that emotion recognition based on physiological signals is a practicable and reliable way of enabling HCI with emotion-understanding capability.

ACKNOWLEDGMENT

This research was supported by the Converging Research Center Program funded by the Ministry of Education, Science and Technology (No. 2012K001330).

REFERENCES

[1] J. Wagner, J. Kim, and E. Andre, "From physiological signals to emotions: Implementing and comparing selected methods for feature extraction and classification," IEEE International Conference on Multimedia and Expo., vol. 7, 2005, pp. 940-943.

[2] P. D. Drummond, and S.-H. Quah, "The effect of expressing anger on cardiovascular reactivity and facial blood flow in Chinese and Caucasians," Psychophysiology, vol. 38. 2001, pp.190-196.

[3] R. W. Picard, E. Vyzas, and J. Healey, "Toward machine emotional intelligence: Analysis of affective physiological state," IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 23, 2001, pp. 1175-1191.

[4] R. O. Duda, P. E. Hart, and D. G. Stork, Pattern Classification, 2nd edn. 2000.

[5] M. Murugappan, R. Nagarajan, and S. Yaacob, "Classification of human emotion from EEG using discrete wavelet transform," Journal of Biomedical Science and Engineering, vol. 3, 2010, pp. 390-396.

[6] K. Takahashi, "Remarks on emotion recognition from bio-potential signals," 2nd International Conference on Autonomous Robots and Agents, 2004, pp. 186-191.

[7] C. Lee, and M.-G. Jang. "A Prior Model of Structural SVMs for Domain Adaptation," ETRI Journal, vol. 33, 2011, pp. 712-719.

[8] J.J. Gross, and R.W. Levenson, "Emotion elicitation using films," Cognition and Emotion, vol. 9, 1995, pp. 87-108.

[9] R.S. Lazarus, J.C. Speisman, A.M. Mordkoff, and L.A. Davidson, "A Laboratory study of psychological stress produced by an emotion picture film," Psychological Monographs, vol. 76, 1962, pp. 553.

[10] M.H. Davis, J.G. Hull, R.D. Young, and G.G. Warren, "Emotional reactions to dramatic film stimuli: the influence of cognitive and emotional empathy," Journal of personality and social psychology, vol. 52, 1987, pp. 126-133.

[11] D. Palomba, M. Sarlo, A. Angrilli, A. Mini, and L. Stegagno, "Cardiac responses associated with affective processing of unpleasant film stimuli," International Journal of Psychophysiology, vol. 36, 2000, pp. 45-57.

[12] B.-J. Park, E.-H. Jang, Sang-Hyeob Kim, C. Huh and J.-H. Sohn, "Three Differential Emotions and Classification Using Physiological Signals," IPCSIT, vol. 25, 2012, pp. 55-59.

[13] P.D. Wasserman, "Advanced Methods in Neural Computing," New York, Van Nostrand Reinhold, 1993, pp. 35–55.