

Passive EEG-based Emotional state Cognition Using Subject-independent Feature Extraction

Sun-Hee Kim and Ngoc Anh Thi Nguyen

Abstract—Research on emotion recognition technologies, whereby emotional state is read off via using voice, facial image, bio-signal and electroencephalogram (EEG), has received considerable attention due to its prevalence in various applications including entertainment, health care, on-line education, automobile, marketing research, to robot and etc. Among the emotional state cognitive technologies utilizing the biometric data, the study on EEG is one of the most widely used. However, EEG-based emotion recognition meets challenges because of multi-channels time series data structure in which contains abnormal characteristics changing the statistical component over time. Consequently, this leads to investigate towards appropriate feature extraction approach in order to be more versatile data analysis to capture underlying information for cognitive the emotional state from EEG signals. In this paper, a subject-independent multiclass-common spatial patterns (Si-MCSP) based on passive EEG therefore is proposed to cognitive the emotional states. Si-MCSP can extract the prominent features in the terms of the subject-independence with respect to the each emotional state. Furthermore, Si-MCSP can automatically determine the numbers of the optimal features which provide an improved classification system of emotional state dramatically. Multiclass Support Vector Machine (Multi-SVM) model then is set up to classify these features for emotion recognition. The experimental results show that the proposed feature extraction method, Si-MCSP, performs higher classification accuracy than the popular feature extraction methods.

Index Terms—Emotion recognition, feature extraction, multiclass common spatial patterns, passive electroencephalogram, subject-independent.

I. INTRODUCTION

The emotion recognition technologies are to analyze the emotional state using data set such as voice signal, facial image, bio-signal, EEG and etc. [1]. In recent years, research on analyzing the emotional state of the human is progressing through qualitative and quantitative measurement evaluation, since the interest of the sentiment or emotion as mental loneliness and depression as well as the physical health has been concentrated. These technologies are applied to design

Manuscript received July 5, 2017; revised August 20. This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2015R1D1A1A01057440), this work was also supported by a Korea University Grant. This work was supported by the University of Danang, (B2017-DN03-07), Viet Nam.

S. H. Kim is with the Department of Brain and Cognitive Engineering, Korea University, Seoul 02841, South Korea (e-mail: sunheekim@korea.ac.kr).

N. A.T. Nguyen is with the Faculty of Information Technology, The University of DaNang- University of Education, 459 Ton Duc Thang Street, Da Nang, Viet Nam (e-mail: ngocanhnt@ued.udn.vn).

products for manufacturing, life environment and services in various fields such as game, military equipment, rehabilitation robot, automobile, crime prevention and etc. [1],[2].

The approach method for the emotional state cognition primarily used the physiological evaluation method based on the bio-signal. The types of the bio-signal are including: blood pressure (BP), electrocardiogram (ECG), pulse, and skin temperature, EEG, and etc.. Among them, EEG is used the most in emotion study. EEG is the flow of electricity that occurs during signal is transferred between cranial nerves on nervous system, and it differently appears depending on the state of mind and body. Therefore, EEG is the most important indicators to measure the activity condition of the brain.

EEG is mainly used to detect and to predict the brain disease such as epileptic seizure and Parkinsons in the medical field and psychology area [3], [4]. Specifically, it is applied for concentration degree measurement, depression diagnosis, stroke diagnosis, progressive bulbar palsy, spinal muscular atrophy, degenerative brain disease, and cranial nerve disease that have motor disturbance [5]. Additionally, EEG is possible the evaluation with reliability than the other bio-signal according to the research results emerged which a part to control the emotion exists on area of brain, hence the research to cognize the emotional state based on the EEG are consistently proceeding [6], [7].

The emotion is one of the important factors to express a state of mind. EEG is signal that is quite objective than other bio-signals to emotional state cognition because it is difficult to control by human deliberately. In this paper, we use the passive EEG signal to be express by suddenly external factor in order to emotional state cognition. We propose Si-MCSP which extracts the features according to the emotional state that able to raise the classification accuracy using passive EEG data.

The rest of this paper is organized as follows. Section II presents our proposed method to extract features according to emotional states. Section III shows the experimental evaluation, and Section IV provides the conclusion.

II. PROPOSED METHOD

CSP algorithm is a method widely known for feature extraction on EEG-based Brain-Computer Interface (BCI) technologies. It is goal to find a spatial matrix to maximize the difference in variances of each class on two classification problem [8]. The first CSP proposed by Ramoser *et al.* [8], the derived various algorithms from CSP was introduced such as Regularized-CSP (R-CSP), Filter Bank-CSP (FB-CSP), Probabilistic-CSP (P-CSP), and etc. [9]-[11]. However, these

methods have weaknesses in applying on multi-class data set because they are only designed to extract the features of data that composed of binary class. T. Yan *et al.* proposed the Multiclass-CSP extending the two-class CSP to solve the multiclass problems, and they applied MCSP to extract the feature on motor imagery EEG data with multiclass [12].

In this paper, we find the spatial principal component for emotional state by using Si-MCSP reflecting the independence of subject on EEG data set that composed of three classes with positive (Po), natural (Na), and negative (Ne) state. Given EEG signal matrix that recorded according to the emotional state of the subject, X . X consists of three class such as X_{Po} , X_{Na} , and X_{Ne} . It is a matrix of $n \times t$. n is number of channel, t means the time point. Therefore, the spatial covariance according to the each emotional state on single trial of a subject estimates as follows:

$$R_{Po} = X_{Po} X_{Po}^T, R_{Na} = X_{Na} X_{Na}^T, R_{Ne} = X_{Ne} X_{Ne}^T, \quad (1)$$

where X^T is the transpose of X . As in binary CSP, the composite covariance matrix, R computes using the spatial covariance about the each emotional state (Po, Na, Ne) as follows:

$$R = R_{Po} + R_{Na} + R_{Ne} \quad (2)$$

The composite covariance matrix R can be factored by eigenvalue decomposition as follows:

$$R = U \Lambda U^T \quad (3)$$

where U is $n \times n$ unitary matrix of principal components, Λ is the $n \times n$ diagonal matrix of eigenvalues. The whitening transformation matrix W computed using eigenvector and eigenvalue that obtained by eigenvalue decomposition of the composite covariance matrix R as follows:

$$W = \Lambda^{-\frac{1}{2}} U^T \quad (4)$$

Whitening transformation matrix, W normalizes the values of the transformed data so that the covariance matrix of the transformed signal get unit matrix and variance comes to 1 at the same time. MCSP applied one-Versus-the-Rest (OVR) method to extract the common spatial patterns depending on the each class from data with multiclass [13]. OVR assumes that multiclass segregate one class R_{Po} and the rest class

R'_{Po} , ($R'_{Po} = R_{Na} + R_{Ne}$). The transformed covariance matrix S_{Po} and S'_{Po} then computed using the whitening transformation matrix W of them as follows:

$$S_{Po} = W R_{Po} W^T, \quad S'_{Po} = W R'_{Po} W^T, \quad (5)$$

where S_{Po} and S'_{Po} share the common principal components, the sum of the corresponding eigenvalues for two matrices will be one. Therefore, S_{Po} and S'_{Po} are able to eigen-decomposition as (6) and (7):

$$S_{Po} = U \Lambda_{Po} U^T \quad (6)$$

$$S'_{Po} = U \Lambda'_{Po} U^T \quad (7)$$

where U is matrix of common principal components, spatial filter of the each class is able to comprise using commonly-shared U and whitening transformation matrix W as follows:

$$S F_{Po} = U^T W \quad (8)$$

where we select m -th common principal components, $U_m = (U_1, \dots, U_m, U_{N-m+1}, \dots, U_N)$ to extract the features of the each class, and m is $m \ll N$.

MCSP facilitates the classification by making bigger the difference in variances between the classes. Therefore, to maximize the difference in variances between one class and the remaining ones, U is arranged as descending order for the eigenvalue Λ_{Po} . Class Po produces the maximum variance if data matrix projected by the first common principal components U_1 . On the other hand, if the data matrix projected by the last common principal components U_N , Po produce the minimum variance, and Po' produce the maximum variance. MCSP produce the projection matrix of the other classes Na and Ne via above same step, and they are to be used as the features.

Subject-independent means that the classifier is trained from the trials conducted by a set of subjects [14]. In this paper, we apply Si-MCSP to consider independency of the subjects that can produce the projection matrix on emotional state. The projection matrix is to be used as the features to classify the emotional state. Si-MCSP extracts the features with consideration for the independency of set of subjects, not a specific subject. Therefore, the normalized covariance matrix based on the subject-independent for the each emotional state is as follows:

$$S_l^k = \sum_{k=1}^K \frac{(X_l X_l^T)}{\text{trac}(X_l X_l^T)}, \quad (k=1, \dots, K), (l=1, \dots, L) \quad (9)$$

where k is the number of subject, l denotes for the number of class for the emotional state. In this paper, we consider to multi-class classification problems. Therefore, the composited average spatial covariance matrix S computes as follows:

$$S = \frac{1}{L} \sum_{l=1}^L S_l^k \quad (10)$$

We perform the eigenvalue decomposition of the composited average spatial covariance matrix S considering the subject-independency using (3)~(7) of conventional MCSP.

In this paper, we apply the exponential forgetting factor that introduced from [15] to determine the number of optimal dimensions of projection matrix (number of the features) on the each emotional state. Generally, exponential forgetting

factor chooses the values in between 0.96 and 0.98. In this paper, we use 0.96 to find the number of the optimal common principal components.

III. EXPERIMENTAL RESULTS

In this paper, we use the SEED database in order to examine the reliable emotional state recognition [16]. SEED dataset is record the EEG of 15 subjects during the movie is playing for 4 minutes. The movie consists of fifteen-Chinese film clips with positive, neutral and negative emotions state. In order to investigate neural signatures and stable patterns across sessions and individuals, each subject performed the experiment three times with an interval of about one week. One experiment contains EEG data of 15 trials. Therefore, each subject data consist 45-trial data set that is sampling 200Hz from 62-channel (15 subjects * 45 trial * 62channel*4 min). Also, In order to filter the noise and remove the artifacts, the EEG data was processed with a bandpass filter between 0.3 to 50 Hz. This data set has three emotion classes, and each subject performs the emotional state itself among negative, neutral and positive class after subject watches movie for 4 minute.

Fig. 1 shows the EEG signal that arise during the subject sees the fifteen-Chinese film clips. Fig. 1(a) is EEG data of subject that belong to positive class, (b) is signal of the neutral class, and (c) shows the negative class. It is a signal that obtained from 32nd-channel of all subject. In this paper, we apply Si-MCSP reflecting independency of subjects to use a set of subjects, not specific subject in order to extract the features depending on emotional state from EEG. It is able to best distinguish the three emotional states.

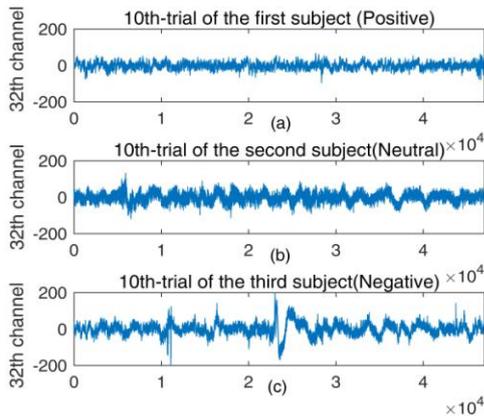


Fig. 1. Original signal for three emotional state.

TABLE I: CLASSIFICATION ACCURACY ACCORDING ON THE SUBJECT-INDEPENDENCY

Number of Subject set (5 trials per subject)	Si-MCSP	
	Number of feature	Accuracy(average)
1	3	61.1
2	4	74.5
3	4	88.6
4	3	83.4
5	4	79.8

Emotional state cognition based on the EEG can be divided into three major phases such as preprocessing step, feature

extraction step, and classification step. Preprocessing step is to filtering needless noise and artifact and etc.. Feature extraction is processing to accentuate the characteristic of data for the classification and analysis of the preprocessed data easily. Finally, classification step is series of process for verifying that obtained result through the model is correspond to the real data.

We are not considered to the preprocessing step during the major sections for emotional state cognition. SEED EEG dataset was processed with a bandpass filter between 0 to 75Hz already. In this paper, we applied Si-MCSP as feature extraction method for emotional state cognition. To find the optimal subject set reflecting the independency of subjects, we measures the classification accuracy according to number of subject set on each class.

Table I shows classification result. In this experiment, we use 5 trials per subject. Example, if number of subject set are two, we use total 10 trials per class to extract the feature. As a result of Table I, the estimated highest classification accuracy using Multi-SVM appeared when we use the set of three subjects. Also, we applied the 0.96% energy rate to find the number of the optimal features (dimension of feature) automatically [15]. When we apply the energy of 96% rate on projection processing, our method found the four feature dimensions. That is, our method provides the highest accuracy rate on four-dimension feature that extracted from 5 trials conducted by three subject set.

We compared the proposed method with the transitional MCSP for feature extraction on multi-class classification. Table II shows the classification result using the features that extract by Si-MCSP and the MCSP. We used Multi-SVM for multi-class in order to measure the classification accuracy. In this paper, we considered to the independency of three subjects per class in the training step based on the results of Table I. As the result, Si-MCSP shows the 5% high classification accuracy than original MCSP. Also, our method was decided the number of feature automatically.

In this paper, we measured the binary classification accuracy to verify the effects of subject-independent. That is, we applied subject-independency to extract the feature for emotional state recognition on the binary class data, and we compared the classification accuracy using the extracted feature through the existing CSP, R-CSP, P-CSP, FB-CSP and our method.

TABLE II: CLASSIFICATION ACCURACY USING FEATURES THAT WAS EXTRACT BY SI-MCSP AND MCSP

Method	Number of Subject set (random choice)	Number of trails	Number of feature	Accuracy (average)
MCSP	1	10	3	66.1
		20	3	69.6
		30	3	76.8
		40	3	82.7
Si-MCSP	3	5 (per subject)	4	87.3

The existing methods computed the covariance matrix per class for extracting the feature on binary class data using 15

trials conducted by a specific subject. Fig. 2 shows the result that the proposed method has a high accuracy of 14%, 10%, 5.8%, and 8.7% than the existing methods when we use the subject-independency in binary class data, respectively. Consequentially, we can confirm that the proposed method is more effective as the feature extraction method to classify the emotional state on EEG signal.

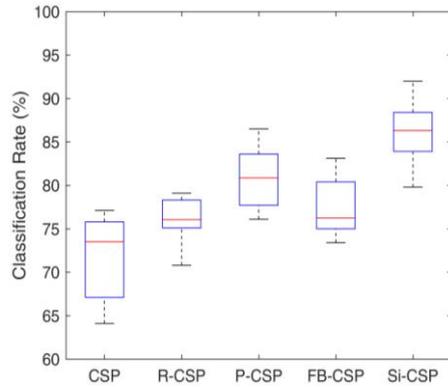


Fig. 2. Classification accuracy by feature extraction method on binary class data.

IV. CONCLUSION

In this paper, we proposed Si-MCSP that extracted the feature classable the emotional state on the passive EEG data. The proposed method extracted the significant feature that can classify the three emotional states, and it provided the 5% improved classification accuracy than the existing feature extraction method. Also, our method carried out without user intervention because the dimension of feature data was decided automatically in contrast with the existing CSP method. Si-MCSP has been reflected the subject-independency extracting the feature that is able to classify well-enough the emotional state than the existing feature extraction methods in multi-class data as well as binary class data. In future work, we will extend our method that is available for the real time processing.

REFERENCES

- [1] J. Park, "Acoustic Information-based Emotion Recognition for Human-Robot Interaction," *J. Korean Inst. Inf. Technol.*, vol. 9, pp. 39-46, Jun, 2011.
- [2] X. A. Fan, L. Z. Bi, and Z. L. Chen, "Using EEG to detect drivers' emotion with Bayesian Networks," in *Proc. 2010 ICML Conf.*, 2010, pp. 11-14.
- [3] B. Harris and I. Gath, "On time delay estimation of epileptic EEG," *IEEE Trans. Biomedical Eng.*, vol. 41, pp. 820-829, September 1994.
- [4] C. J. Stam, B. Jellesa, H. A. M. Achtereektea, S. A. R. B. Rombouts, J. P. J. Slaetsb, and R. W. M. Keunena, "Investigation of EEG non-linearity in dementia and Parkinson's disease," *Electroen. Clin. Neuro.* vol. 95, pp. 309 -317, November 1995.
- [5] D. H. Kim, and K. S. Hwang, "Development and verification of digital EEG signal transmission protocol," *J. KICS*, vol. 38C, pp. 623-629, July, 2013.
- [6] P.C. Petrantonakis, and L.J. Hadjileontiadis, "Emotion Recognition From EEG Using Higher Order Crossings," *IEEE T. Inf. Technol. B.* vol. 14, pp. 186-197, March, 2010.

- [7] A. T. Sohaib, S. Qureshi, J. Hagelback, O. Hilborn, and P. Jercic, "Evaluating classifiers for emotion recognition using EEG," in *Proc. 2013 AC Conf.*, 2013, pp. 492-501.
- [8] H. Ramoser, J. Müller-Gerking, and G. Pfurtscheller, "Optimal Spatial Filtering of Single Trial EEG During Imagined Hand Movement," *IEEE T. Rehabil. Eng.*, vol. 8, pp. 441-446, December 2000.
- [9] H. Lu, H. L. Eng, C. Guan, K. N. Plataniotis, and A. N. Venetsanopoulos, "Regularized Common Spatial Pattern With Aggregation for EEG Classification in Small-Sample Setting," *IEEE T. Bio-Med. Eng.*, vol. 57, pp. 2936-2946, December 2010.
- [10] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, "Filter Bank Common Spatial Pattern (FBCSP) in brain-computer interface," in *Proc. 2008 IJCNN Conf.*, 2008, pp. 2390-2397.
- [11] W. Wu, Z. Chen, X. Gao, Y. Li, E. N. Brown, and S. Gao, "Probabilistic common spatial patterns for multichannel EEG Analysis," *IEEE T. Pattern Anal.* vol.37, pp.639-653, June, 2015.
- [12] T. Yan, T. Jingtian, and G. Andong, "Multi-class EEG classification for brain computer interface based on CSP," in *Proc. 2008 BMEI Conf.*, 2008, pp. 469-472.
- [13] W. Wu, X. Gao, and S. Gao, "One-Versus-the-Rest(OVR) algorithm: An extension of Common Spatial Patterns (CSP) Algorithm to multi-class case," in *Proc. 2005 EMBC Conf.*, 2005, pp. 2387-2390.
- [14] T. T. A. Hoang, "Multivariate Features for Multi-class Brain Computer Interface Systems," Ph.D. dissertation, Dept. Inf. Sci. and Eng., Univ. of Canberra, Canberra, Australia, 2014.
- [15] S. Papadimitriou, J. Sun, C. Faloutsos, "Streaming pattern discovery in multiple time-series," in *Proc. 2005 VLDB Conf.*, 2005, pp. 697-708.
- [16] X. A. Fan, L. Z. Bi, and Z. L. Chen, "Using EEG to detect drivers' emotion with Bayesian Networks," in *Proc. 2010 ICMLC Conf.*, 2010, pp. 11-14.



Sun-Hee Kim received the B.S. in multimedia from Korean Educational Development Institute in 2004 and the M.S. degree in computer science from Dongguk University, Korea in 2006. She received the Ph.D. degrees in computer science from Chonnam National in 2011. Between 2011 and 2013, she was a visited as post-doctoral in Carnegie Mellon University, and she worked on data mining problems in bio-signal with the Database Group. She is currently a Research Professor at Department of Brain and Cognitive Engineering, Korea University, Seoul, Korea. Her current research interests are data mining, machine learning, bioinformatics, and its applications.



Ngoc Anh Thi Nguyen received the B.S. in Faculty Mathematics-Informatics from Da Nang Education University, Viet Nam and M.S. at the Dept. Electronics and Computer Engineering, Chonnam National University, Korea. She received the Ph.D. degrees in electronics and computer engineering from Chonnam National University, Korea. She is currently a lecturer at Faculty of Information Technology of The University of DaNang-University of Education, Viet Nam. Her research interests are data mining, machine learning, data analysis, signal processing, pattern recognition and mathematical modeling, mainly working on time series mining, applying on bioinformatics and neuroscience.