EEG-BASED EMOTION RECOGNITION IN MUSIC LISTENING: A COMPARISON OF SCHEMES FOR MULTICLASS SUPPORT VECTOR MACHINE

Yuan-Pin Lin [1], Chi-Hong Wang [2], Tien-Lin Wu [1], Shyh-Kang Jeng [1] and Jyh-Horng Chen [1]

[1]Department of Electrical Engineering, National Taiwan University, Taiwan
[2]Cardinal Tien Hospital, Yung-Ho Branch, Taiwan

ABSTRACT

Currently, how to equip machines with the ability for properly recognizing users' felt-emotion during multimedia presentation is a growing issue. In this study we focused on the approach for recognizing music-induced emotional responses from brain activity. A comparative study was conducted to testify the feasibility of using hierarchical binary classifiers to improve the classification performance as compared with nonhierarchical schemes. According to our classification results, we not only found that using one-against-one scheme of hierarchical binary classifier results in an improvement to performance, but also established an alternative solution for emotion recognition by proposed model-based scheme depending on 2D emotion model.

Index Terms— Brain activity, Emotion recognition

1. INTRODUCTION

The goal of bio-inspired multimedia research is to create a new-generation interaction between human and multimedia. It allows multimedia can be directly triggered by user's simultaneous responses. Thus, this bio-interface between human and multimedia will not be relied on clicking remote controllers or series of touch buttons any more. By combining previously available techniques of multimedia and bio-signal processing, all of these corresponding applications tend to not only create the concept of the human-centered orientation, but also build a more immersive digital media environment that a user can control with his/her bio-feedback. In order to create the ultimate interface, how to equip machines for properly recognizing users' felt-emotion during multimedia presentation is a crucial issue today.

Emotion is a high level cognitive process and physiological state related to a complex of feelings, thoughts, and behaviors. It does likely contain a larger input from the brain in which limbic system charges the dominant task for generating emotion experience [1]. If capturing and interpreting the information from brain activity is possible, it would be a more natural and direct way to observe the emotion responses. Based on electroencephalography (EEG), ongoing brain activity can be noninvasive recorded. Several works have already been done for emotion recognition in EEG signals. For example, Ishino et al. [2] proposed a system for estimating the feelings of joy, anger, sorrow and relaxation by using neural network, which obtained with accuracy of 54.5% for joy, 67.7% for anger, 59% for sorrow and 62.9% for relaxation. Berkman et al. [3] used a single-layer neural network to predict three categories of emotions including positive, negative and neutral with accuracy of 43%. Chanel et al. [4] showed that arousal assessment of emotion can be obtained with a maximum accuracy of 58% for three emotion classes estimated by the Naïve Bayes classifier. Also, our previous work involved multilayer perceptron [5] and support vector machine [6] to recognize music-induced emotion response with accuracy of 69.69% and 92.73% respectively.

Considering the above research works, all of them targeted on emotional categories range from two to four. Unlike the felt-emotion categories, the expressed-emotion can be described in a total of 67 adjectives [7]. However, when the number of the categories of predicted felt-emotion increases, it would be an inevitable issue as how to maintain or even improve the performance. Accordingly, we suggest a possible solution by using a hierarchical binary classifier to handle the multi-class classification problem. There is still little discussion on the hierarchical classification in EEG-based emotion recognition in literature. Thus, in this study we will elucidate the concept of hierarchy, as well as give a comparison of performance between hierarchical and nonhierarchical classification.

This paper is structured as follows. Section 2 describes the material and method. In Section 3, we present the experimental results and corresponding discussions. Section 4 is with conclusion and future work.

2. MATERIAL AND METHOD

2.1. Data acquisition and experimental procedure

We used the same EEG dataset of 26 subjects acquired at our previous EEG study [6]. The data acquisition and experimental procedure are briefly depicted in this section. EEG data was captured by a 32-channel EEG module, where the scalp locations followed the 10-20 protocol. Sampling rate and filter bandwidth are 500Hz and 1~100 Hz respectively. With respect to experimental procedure, we

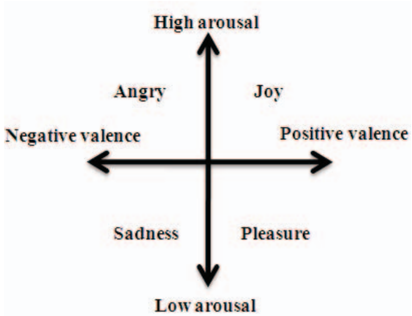

Fig. 1. The 2D emotion model and targeted four emotional states.

targeted four basic emotional states by following a simplified framework of emotion model [8] to induce emotion-specific EEG signals, including joy, anger, sadness and pleasure, as shown in Fig. 1. Emotional responses were induced by pre-labeled emotional music [9]. During our previous entire experiment, the EEG data was divided into 16 segments of 30s duration with their emotional label of self-report were used to perform the machine learning in this study.

2.2. Feature extraction

The raw EEG data were visually inspected to remove the artifact, where only obvious motion artifacts were blocked. Then, we use short-time Fourier transform (STFT) with Hanning window of time interval and with a 50% overlap to extract the power spectral density value over time. From previous reports about EEG classification, spectral powers of EEG components are the most common adopted features, where EEG components defined according to frequency range, including delta (1-3 Hz), theta (4-7 Hz), alpha (8-13 Hz), beta (14-30 Hz) and gamma (31-50 Hz). Therefore, after applying STFT, we can derive the power value of each EEG components across time over 32-channels.

According to our previous EEG-based emotion study[6], we found that feature type of ASM12 (power difference at a symmetric electrode pair) with temporal resolution of one second would perform higher classification accuracy. It means that using a power asymmetry index as a feature is a sensitive method for detecting brain activation related to emotional responses. By taking the advantage of sensitivity to emotion classification, we follow the same feature type ASM12 to construct the proposed classification approach. The ASM12 comprises 12 asymmetry indices derived from 12 symmetric electrode pairs, including: Fp1-Fp2, F7-F8, F3-F4, FT7-FT8, FC3-FC4, T7-T8, P7-P8, C3-C4, TP7-TP8, CP3-CP4, P3-P4, and O1-O2. These only excluded the reference electrodes of A1 and A2, as shown in Fig. 2. The feature extraction procedure repeated in each of five EEG components and then yielding 60 features. The number of sample points extracted per second from one subject was around 944 (16 30s EEG segments x around 59 points per

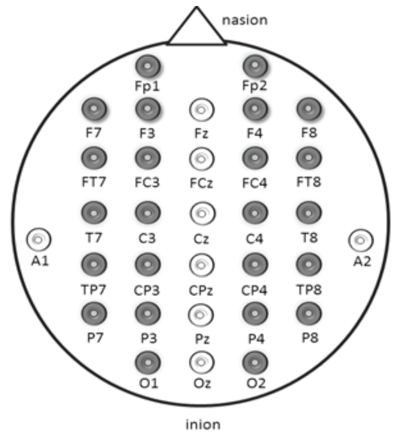

Fig. 2. The locations of 12 electrode pairs on 32-channel scalp EEG. (Electrodes in gray represent the feature extraction sites)

segment). There were 26 matrices (944x60) for conducting following classification, where rows and columns represent sample points and feature attributes respectively. The corresponding class per sample point was assigned according to the subject's self-report. It was noted that sample points extracted within one segment were regarded as the same emotional class.

2.3. Feature vector classification

We performed a comparative study to explore the feasibility for handling the multi-class emotion classification problem, using different structures of multi-class support vector machine (SVM). SVM is one of the most popular supervised learning algorithms for solving classification problems. Because SVM was originally designed for solving the binary classification problems, currently two types of approaches are developed to efficiently face the multi-class classification. One is to construct a hierarchical binary classifiers and the other is to directly obtain data optimization collectively [10]. In this study, we selected one from each type including "one-against-one" scheme for hierarchy and "all-together" scheme for optimization formulation. In addition, by exploiting the structure of 2D emotion model (valence and arousal), we further constructed a model-based hierarchical classification scheme to handle the multiclass emotion classification problem. Thus, this study included three schemes of multi-class SVM for comparison, where the detail descriptions are shown below:

2.3.1. "all-together" scheme

As compared to the performance of hierarchical binary classifier, directly optimizing all variables of multi-class in one step was included, as shown in Fig. 3(a). Here, we adopted the optimization formulation proposed by Crammer et al. [11] and then used the BSVM toolkit [10] to perform the Crammer's optimization formulation.

2.3.2. "one-against-one" scheme

In [10], it reported that one-against-one scheme is more suitable for practical use than other schemes such as one-against-all. Thus, we simply used the LIBSVM toolkit [12] which follows the structure of one-against-one for multi-class classification. This method constructs a number of k(k-1)/2 classifiers, where each one is trained on data from two classes. In the three-class case, for example, three binary classifiers, including Class 1 vs. Class 2, Class 1 vs. Class 3 and Class 2 vs. Class 3, are required to been trained. Then, the predicted class in testing stage will be decided via max wins strategy, determining data with the largest vote, as shown in Fig. 3(b). In each classifier, the radial basis function (RBF) kernel was applied to nonlinearly map data into a higher dimension space for classification.

2.3.3. "model-based" scheme

The 2D emotion model is constructed from two axes of arousal and valence. Therefore, model-based scheme decomposes the four-class classification into two-level nested binary classifier based on arousal-specific level and valence-specific level to recognize four emotion states, as shown in Fig. 3(c). Feature data according to their four-class emotion labels was separated into categories of valence and arousal. The valence level comprises positive valence (joy and pleasure) and negative valence (angry and sadness), whereas arousal level contains high arousal (joy and angry) and low arousal (pleasure and sadness). At each node of level, we also used LIBSVM toolkit [12] to build the SVM classifier and chose RBF as kernel. The entire training feature vector and their corresponded labels were trained in the valence-specific level and arousal-specific level separately. Besides, the first level will select a higher-performance one from valence and arousal, which would not limit the final classification performance.

3. RESULT AND DISCUSSION

In this section, the results of using three types of multi-class classification will be shown and discussed. We used 10-fold cross-validation method to increase the reliability of the classification results. Data is randomly divided into 10 subsets. This procedure will repeat for 10 times such that each subset has a chance to be the testing data and remaining subsets are used to train the classifier. The criterion for evaluating the performance of three schemes of multi-class classification is their accuracy rate which is derived from the ratio of correctly classified number of samples from the total number of samples.

Table 1 provides an overall results of four emotions, valence-specific and arousal specific classification by using different schemes of multi-class SVM classifier. The cases of arousal and valence are two-class classification problems for distinguishing high/low arousal and positive/negative valence, whereas the case of four emotions is to recognize one emotion class from joy, angry, sadness and pleasure

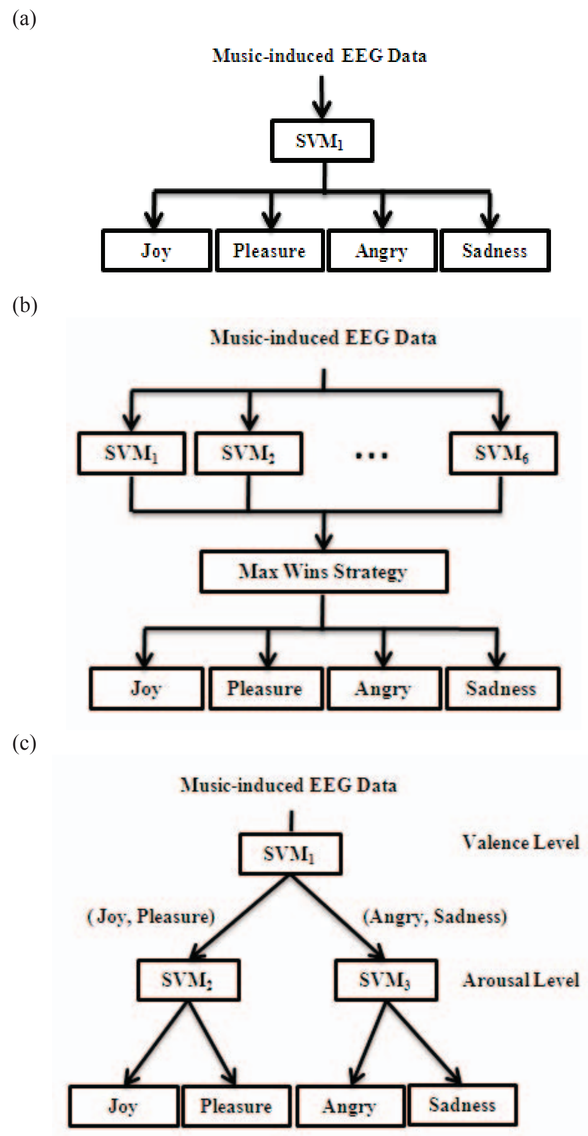

Fig. 3. Different schemes of multi-class SVM classifier, (a) all-together, (b) one-against-one, and (c) model-based.

directly. As the obtained results, using one-against-one scheme of SVM can obtain better performance than all-together scheme across three classification cases. The averaged classification accuracy for four emotions, valence and arousal were 92.57%, 94.86% and 94.43% respectively. It also found that both valence and arousal cases yield higher accuracy than that in direct four-class classification. But, it seems no apparent difference between arousal and valence classification.

Next, model-based scheme composed of two-level binary classifiers was proposed. Due to no obvious difference in performance between valence and arousal, we applied both at first level in turn for comparison. The model-based scheme obtained an average accuracy of

Table 1.
Averaged Classification Results (Standard Deviation)
Using Different Schemes of Multi-class SVM Classifiers

Schemes	Four	Valence	Arousal
All-togerther	82.37 (4.12)	86.42 (3.71)	85.86 (3.46)
One-against-one	**92.57 (2.16)**	**94.86 (1.76)**	**94.43 (2.12)**
Model-based (1st: Valence)	90.52 (2.84)	x	x
Model-based (1st: Arousal)	90.72 (2.73)	x	x

x: cannot perform two-class classification in model-based scheme.

90.52% for valence at first level and 90.72% for arousal at first level. As compared the model-based with all-together, we got the consistent evidence that decomposing the problem of multi-class emotion in relation to valence-specific and arousal specific levels will provide higher performance than directly multi-class classification in one step [13]. At last, when compared the performance between four schemes in four-class classification, it turned out that one-against-one scheme showed the best performance for discriminating emotional responses.

Theoretically, treating multi-class classification problem as using hierarchical binary classifiers usually yields higher accuracy than that obtained by direct multi-class classification in one step. As expected, this also proves true in our EEG data after the comparative classification. Furthermore, although the proposed model-based approach just obtained a comparable performance to one-against-one scheme, the advantage is that less number of SVMs is required. This property will help to reduce the training time, especially in the case of predicted class number increasing.

4. CONCLUSION

In this study, we have testified the feasibility of classification solution that decomposes multi-class classification into hierarchical binary classifiers for achieving better performance in EEG classification problem. We not only adopted two types of multi-class SVM schemes from nonhierarchical scheme such as all-together to hierarchical scheme such as one-against-one, but constructed a model-based scheme based on the structure of 2D emotion model. From the obtained results, we found that one-against-one scheme of multi-class SVM classifier yielded a best performance in distinguishing four emotion classes. Accordingly, the solution of combining several binary classifiers for multi-class classification actually outperforms the nonhierarchical classifier. Besides, the proposed model-based scheme achieved a comparable performance with involved less number of SVMs compared to one-against-one. We suggest that it may provide an alternative way for solving the EEG-based multi-class emotion classification problems.

5. REFERENCES

[1] T. Dalgleish, "The Emotional Brain," *Nature Reviews Neuroscience,* vol. 5, pp. 582-589, Jul 2004.

[2] K. Ishino and M. Hagiwara, "A Feeling Estimation System Using a Simple Electroencephalograph," in *IEEE International Conference on Systems, Man and Cybernetics* 2003, pp. 4204-4209 vol.5.

[3] E. Berkman, D. K. Wong, M. P. Guimaraes, E. T. Uy, J. J. Gross, and P. Suppes, "Brain Wave Recognition of Emotions in EEG," *Psychophysiology,* vol. 41, pp. S71-S71, 2004.

[4] G. Chanel, J. Kronegg, D. Grandjean, and T. Pun, "Emotion Assessment: Arousal Evaluation Using EEG's and Peripheral Physiological Signals," *Multimedia Content Representation, Classification and Security,* vol. 4105, pp. 530-537, 2006.

[5] Y. P. Lin, C. H. Wang, T. L. Wu, S. K. Jeng, and J. H. Chen, "Multilayer Perceptron for EEG Signal Classification during Listening to Emotional Music," in *IEEE International Region 10 Conference* Taipei, Taiwan, 2007.

[6] Y. P. Lin, C. H. Wang, T. L. Wu, S. K. Jeng, and J. H. Chen, "Support Vector Machine for EEG Signal Classification during Listening to Emotional Music," in *IEEE 10th Workshop on Multimedia Signal Processing*, 2008, pp. 127-130.

[7] K. Hevner, "Expression in Music - A Discussion of Experimental Studies and Theories," *Psychological Review,* vol. 42, pp. 186-204, 1935.

[8] P. Ekman, "An Argument for Basic Emotions," *Cognition & Emotion,* vol. 6, pp. 169-200, May-Jul 1992.

[9] T. L. Wu and S. K. Jeng, "Probabilistic Estimation of a Novel Music Emotion Model," in *14th International Multimedia Modeling Conference*. vol. 4903, Kyoto, Japan, 2008, pp. 487-497.

[10] C. W. Hsu and C. J. Lin, "A Comparison of Methods for Multi-class Support Vector Machines," *IEEE Transactions on Neural Networks,* vol. 13, pp. 415-425, 2002.

[11] K. Crammer and Y. Singer, "On the Learnability and Design of Output Codes for Multiclass Problems," in *Proceedings of the Thirteenth Annual Conference on Computational Learning Theory* 2000, pp. 35-46.

[12] C. C. Chang and C. J. Lin, "LIBSVM: A Library for Support Vector Machines, Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm ", 2001.

[13] J. Kim and E. Andre, "Emotion Recognition Based on Physiological Changes in Music Listening," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 30, pp. 2067-2083, 2008.