EEG-Based Emotion Recognition in Listening Music by Using Support Vector Machine and Linear Dynamic System

Ruo-Nan Duan[1], Xiao-Wei Wang[1], and Bao-Liang Lu[1,2,3,4,⋆]

[1] Center for Brain-Like Computing and Machine Intelligence
Department of Computer Science and Engineering
[2] MOE-Microsoft Key Laboratory for Intelligent Computing and Intelligent Systems
[3] Shanghai Key Laboratory of Scalable Computing and Systems
[4] MOE Key Laboratory of Systems Biomedicine
Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200240, China
bllu@sjtu.edu.cn

Abstract. This paper focuses on the variation of EEG at different emotional states. We use pure music segments as stimuli to evoke the exciting or relaxing emotions of subjects. EEG power spectrum is adopted to form features, power spectrum, differential asymmetry, and rational asymmetry. A linear dynamic system approach is applied to smooth the feature sequence. Minimal-redundancy-maximal-relevance algorithm and principal component analysis are used to reduce the dimension of features. We evaluate the performance of support vector machine, k-nearest neighbor classifiers and least-squares classifiers. The accuracy of our proposed method reaches 81.03% on average. And we show that the frequency bands, beta and theta, perform better than other frequency bands in the task of emotion recognition.

Keywords: emotion recognition, electroencephalogram, power spectrum.

1 Introduction

Emotional states significantly affect the cognition and behaviors of human. Scientific findings suggest an increasingly large number of important functions of emotion [1]. Emotion recognition based on information technology is an important research topic in the field of neural engineering.

Most previous studies on emotion recognition focused on extrinsic signals, such as facial expressions [2] and voice [3]. However, human may modify their appearance deliberately, and some disabled persons can not express their emotion in these extrinsic ways. Detecting emotional states of people through their physiological signals can solve this problem in some degree. Common approaches to recognize emotion based on physiological signals include detecting emotional states

⋆ Corresponding author.

T. Huang et al. (Eds.): ICONIP 2012, Part IV, LNCS 7666, pp. 468–475, 2012.

through electroencephalography (EEG), electrocardiogram (ECG), electromyogram (EMG), skin resistance (SC), skin temperature, pulse and respiration signals. EEG signal is a typical central nervous system signal, which relates to emotion activities more closely than automatic nervous system signal, such as ECG and EMG signal.

Recently, many institutions have started to study the EEG-based emotion recognition. Aftanas *et al.* made use of Fourier transform to map the original EEG signal to theta, alpha and beta frequency bands, and calculated power spectral density of each electrodes as features to recognize emotional states [4]. Lin *et al.* classified emotional states into four categories using support vector machine, and found the 30 most relevant EEG features to the emotion recognition [5]. Nie *et al.* figured out the key brain regions and frequency bands of EEG signals according to the correlation coefficients of frequency-domain features in the emotion recognition task [6].

In this paper, we did experiments to collect EEG signals of subjects when they were listening to music with exciting or relaxing emotional elements. And we compared the performance of three kinds of frequency-domain features in five frequency bands in emotion classifying task. We adopted linear dynamic system (LDS) approach to remove the noise of the features, and employed support vector machine (SVM) as the classifier. At last, minimal-redundancy-maximal-relevance criterion (MRMR) and principal component analysis (PCA) were applied to reduce the dimension and speed up the classifying procedure.

2 Experiment

2.1 Stimuli

In order to evoke the emotion of the participants, we chose to use pure music segments in specific emotion style as stimuli, such as *Kiss The Rain* for relaxing emotion and *He's A Pirate* for exciting emotion. We used music with no lyrics for the reason that we had found that songs with lyrics in different languages impact the EEG of participants in different degrees [7]. We selected 16 pieces of pure music, each of them lasting 3 minutes.

2.2 Subjects

Five right-hand students were chosen as subjects who were all from Shanghai Jiao Tong University, aged between 21 to 24 years old, 3 males and 2 females. They had no history of brain damage and mental illness before. We informed them of the purpose and procedure of the experiment, and that the EEG recording device was harmless.

2.3 Procedure

Before the experiment, we recorded the information of the subjects and the environment. We played the stimuli using Stim 2 software, and ESI NeuroScan

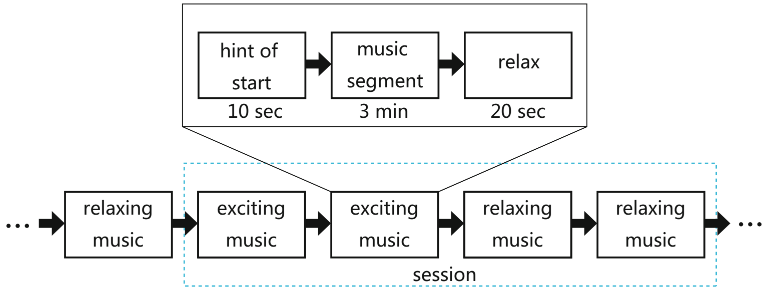

Fig. 1. Procedure of the stimuli playing

System with a 62-channel electrode cap was employed to record the EEG signals of subjects. The sampling rate was 500Hz. Those music segments were divided into 4 sessions, each session containing 2 pieces of exciting music and 2 pieces of relaxing music. We played those music segments in the order shown in Fig. 1.

3 Methods

3.1 Preprocessing

In order to speed up the computation, we down-sampled the EEG data with sample frequency 200Hz. Then we removed the electromyogram and other artifacts manually.

3.2 Feature Extraction

We used frequency-domain features and their combinations as features. At first, the 512-point short-time fourier transform (STFT) with a non-overlapped Hanning window of 1s was applied to the EEG data from each of the 62 channels to compute power spectrum. Then we averaged the spectral time series into five frequency bands (delta: 1-3Hz, theta: 4-7Hz, alpha: 8-13Hz, beta: 14-30Hz, gamma: 31-50Hz), and combined them to form 3 kinds of features named PS (power spectrum), DASM (differential asymmetry), RASM (rational asymmetry). We chose Fp1, F7, F3, FT7, FC3, T7, P7, C3, TP7, CP3, P3, O1, AF3, F5, F7, FC5, FC1, C5, C1, CP5, CP1, P5, P1, PO7, PO5, PO3, CB1 of the left hemisphere, and Fp2, F8, F4, FT8, FC4, T8, P8, C4, TP8, CP4, P4, O2, AF4, F6, F8, FC6, FC2, C6, C2, CP6, CP2, P6, P2, PO8, PO6, PO4, CB2 of the right hemisphere as the corresponding electrodes .

PS features were consisted of the average power spectrum of 62 scalp electrodes in the five frequency bands. RASM and DASM features were the average power spectrum ratios and differences of those 27 pairs of hemispheric asymmetry electrodes, respectively. Totally, we gain 580 features for one sample, and one sample per second. The number of features is shown in Table 1.

Table 1. Statistics of the feature number

Feature	Delta	Theta	Alpha	Beta	Gamma	Total
PS	62	62	62	62	62	310
RASM	27	27	27	27	27	135
DASM	27	27	27	27	27	135

3.3 Feature Smoothing

The features we extracted above contained many rapid fluctuations. As we all know, emotion usually varies smoothly, and the features relevant to emotional states should change relatively slowly. So the features that changed rapidly have nothing to do with the emotional states and we should smooth the features to remove those noise components.

We used a linear dynamic system (LDS) approach [8] to remove noise. We applied off-line LDS with the window of 20s to smooth the features and to remove noise in some degree.

3.4 Classification

Our classifying task is to use part of the EEG data of one experiment to train the model and the rest data recorded in the same experiment to test it. In practical applications, we can only use the data we have obtained in laboratory to classify the data we record later, so we used the data of the first three sessions to train the model, and the rest one session to evaluate it. Since each session contained 700 to 800 cases, we had about two thousands training samples and seven hundreds test samples.

We classified the EEG features using three kinds of algorithms, k-nearest neighbors (KNN) algorithm, least-squares (LS) classification and support vector machine (SVM). kNN model works in the way that classifying the test case to the most common category of the k nearest neighbors with known category. Least-squares classification tries to find a model that is a linear function with parameters which minimize the sum-of-squares error [10]. SVM projects the features onto another feature space using linear kernel function. Then SVM iteratively approaches the optimal hyperplane which has the maximal margins. We used LIBSVM [9] software to train the SVM classifier.

3.5 Dimension Reduction

To speed up the computing procedure and reduce the complexity of the model, we applied principal component analysis (PCA) and minimal-redundancy-maximal-relevance (MRMR) algorithm to reduce the dimension of the features. PCA projects data onto a new space with a lower dimension and maximizes the variance of the projected data. And MRMR algorithm uses a series of intuitive measures of redundancy and relevance to select optimal feature subsets [11].

4 Results

4.1 The Effect of the LDS Approach

We applied the LDS approach to smooth the features and remove the noise. A piece of feature sequence is cut to show the performance of LDS (Fig. 2).

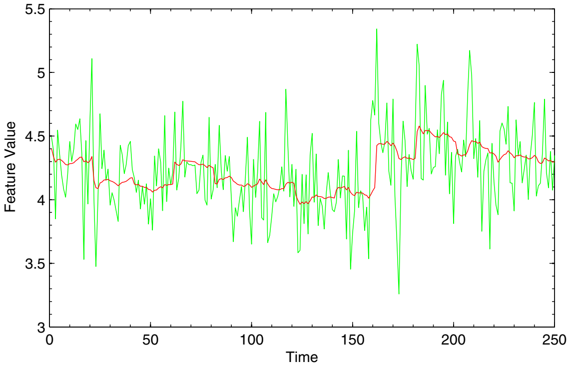

Fig. 2. Performance of LDS smoothing on feature sequence. The green curve represents the original feature values, and the red one stands for the feature values after smoothing. Feature sequence shown is part of the 24th PS feature in gamma frequency band of subject 4 in session 1.

In the experiment, we trained and tested SVM models using original features and features after LDS smoothing respectively. The features we adopted to evaluate the performance of LDS smoothing were PS in five frequency bands. The comparison of classification accuracies between original features and smoothed features is shown in Table 2. The accuracies indicate that LDS smoothing helps to increase the accuracy in some degree. And the effect of LDS smoothing becomes significant on unclean data.

Table 2. Classification accuracies of features with and without LDS smoothing

	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Average
No smoothing	74.97	70.96	71.84	**98.97**	65.01	76.35
LDS smoothing	**76.48**	**71.95**	**85.86**	98.63	**67.47**	**80.08**

4.2 Performance of Features and Classification Methods

To compare the performance of different frequency bands and different features, we adopted linear-SVM to classify the features after smoothing them. The accuracies of classification is shown in Table 3. It is obvious that the frequency

Table 3. Classification accuracies using different kinds of features

Subject	Feature Name	Delta	Theta	Alpha	Beta	Gamma	Total
1	DASM	57.59	62.22	67.79	74.86	77.75	**84.13**
	RASM	57.82	64.54	56.32	76.38	80.30	**81.34**
	PS	70.10	63.38	72.54	77.98	**82.39**	76.48
2	DASM	59.07	78.72	72.46	77.85	**81.23**	78.60
	RASM	63.95	77.97	52.82	77.35	**80.48**	74.72
	PS	61.58	74.34	70.09	70.34	**74.72**	71.96
3	DASM	78.91	59.79	60.95	65.70	72.42	**84.82**
	RASM	80.53	56.32	57.94	67.09	75.55	**84.36**
	PS	78.68	63.27	59.10	70.10	84.01	**85.86**
4	DASM	60.96	83.45	76.03	98.06	**98.06**	97.95
	RASM	56.51	72.83	59.13	98.06	**98.17**	98.17
	PS	62.67	80.25	89.38	98.40	98.40	**98.63**
5	DASM	64.95	63.93	59.82	**66.78**	57.53	66.32
	RASM	61.99	63.13	64.50	59.36	57.19	**64.61**
	PS	**78.54**	61.19	65.98	65.53	65.30	67.47
Average	DASM	64.30	69.62	67.41	76.65	77.40	**82.36**
	RASM	64.16	67.18	58.14	75.65	78.34	**80.64**
	PS	70.31	68.49	71.42	76.47	**80.96**	80.08

bands beta and gamma perform better than frequency bands delta, theta and alpha. This result is consistent with previous results that EEG signals in high frequency bands are more relevant to the emotional states of human [12]. We find that features, RASM and DASM, gain relatively high classification accuracies with lower feature dimensions. This result reflects that the differences and ratios of EEG signals between left and right hemispheres are related to the human emotion in some degree.

To evaluate the performance of different classifiers, KNN algorithm, linear-squares classification and liner-SVM were used to classify LDS-smoothed EEG features. To compare the performance of the three classifiers, we show the accuracies of classifying PS features in all frequency bands in Table 4. According to the result, KNN classifier and SVM perform better than least-squares classification. The performance of KNN classifier and SVM are similar to each other. However, SVM works much faster than KNN classifier in this task because of the high dimension of the features.

Table 4. Classification accuracies of different classifiers

Classifier	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Average
KNN	72.19	71.34	**87.72**	97.83	**69.41**	79.69
LS	**78.45**	62.08	79.49	88.47	59.02	73.50
SVM	76.48	**71.95**	85.86	**98.63**	67.47	**80.08**

4.3 Dimension Reduction

PCA and MRMR were used to reduce the dimension of the features. The original features were the combinations of PS, DASM and RASM in five frequency bands, so the original dimension of the features was 580. After the dimension was reduced to 510, 460, . . . , and 10 with the intervals of 50, the performance of SVM classifier changes as shown in Fig. 3.

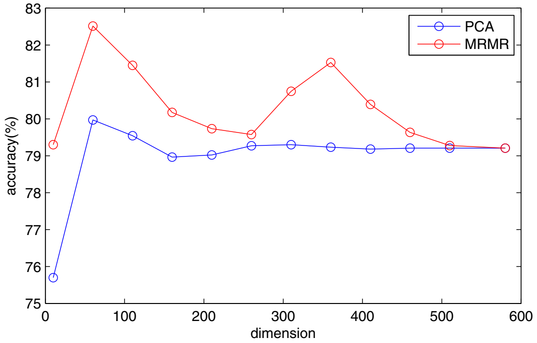

Fig. 3. Classification accuracies on features with different dimensions. The accuracies shown were the average accuracies of the five subjects by using SVM classifier and two reduction methods.

The figure indicated that the dimension reduction using PCA affected the accuracy in a small range when the dimension was larger than 100, whereas MRMR showed the possibility of increasing the accuracy. This increment was caused by the removing of emotion irrelevant features which annoyed the classifier. And this result implies that we can reduce the dimension of features to speed up the calculation.

5 Conclusion

Using the EEG data recorded from subjects when they were in some emotional state evoked by music, we obtained an efficient and stable emotion classifier. The accuracies of classification reached 81.03% on average. In our study, we showed that the frequency bands beta and gamma performed better than other frequency bands in the task of emotion recognition. In addition, PCA, MRMR and LDS can help to increase the speed and improve the stability of the classification procedure, respectively.

Acknowledgments. This research was partially supported by the National Basic Research Program of China (Grant No. 2009CB320901), and the European Union Seventh Framework Programme (Grant No. 247619).

References

1. Picard, R.W., Klein, J.: Toward Computers that Recognize and Respond to User Emotion: Theoretical and Practical Implications. Interacting with Computers 14(2), 141–169 (2002)
2. Anderson, K., McOwan, P.W.: A Real-Time Automated System for the Recognition of Human Facial Expressions. IEEE Transaction on System, Man, and Cybernetics, Part B: Cybernetics 36(1), 96–105 (2006)
3. Bruck, C., Kreifelts, B., Wildgruber, D.: Emotional Voices in Context: a Neurobiological Model of Multimodal Affective Information Processing. Physics of Life Reviews 8, 383–403 (2011)
4. Aftanas, L.I., Lotova, N.V., Koshkarov, V.I., et al.: Non-Linear Dynamic Complexity of the Human EEG During Evoked Emotions. International Journal of Psychophysiology 28, 63–76 (1998)
5. Lin, Y.P., Wang, C.H., Jung, T.P., et al.: EEG-Based Emotion Recognition in Music Listening. IEEE Transaction on Biomedical Engineering 57(7), 1798–1806 (2010)
6. Nie, D., Wang, X.W., Shi, L.C., et al.: EEG-Based Emotion Recognition During Watching Movies. In: Proceedings of IEEE EMBS Conference on Neural Engineering, pp. 667–670. IEEE Press (2011)
7. Shi, S.J., Lu, B.L.: EEG Signal Classification during Listening to Native and Foreign Languages Songs. In: Proceedings of 4th International IEEE EMBS Conference on Neural Engineering, pp. 440–443 (2009)
8. Shi, L.C., Lu, B.L.: Off-Line and On-Line Vigilance Estimation Based on Linear Dynamical System and Manifold Learning. In: Proceedings of 32nd International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 6587–6590. IEEE Press (2010)
9. Chang, C.C., Lin, C.J.: LIBSVM: A library for support vector machines, http://www.csie.ntu.edu.tw/~cjlin/libsvm
10. Bishop, C.M.: Pattern Recognition and Machine Learning. Springer, New York (2006)
11. Peng, H., Long, F., Ding, C.: Feature Selection Based on Mutual Information: Criteria of Max-Dependency, Max-Relevance, and Min-Redundancy. IEEE Transactions on Pattern Analysis and Machine Intelligence 27(8), 1226–1238 (2005)
12. Muller, M.M., Keil, A., Gruber, T., Elbert, T.: Processing of Affective Pictures Modulates Right-Hemispheric Gamma Band EEG Activity. Clinical Neurophysiology 110(11), 1913–1920 (1999)