Cross-subject and Cross-gender Emotion Classification from EEG

Jia-Yi Zhu[1], Wei-Long Zheng[1] and Bao-Liang Lu[1,2]

[1] Center for Brain-Like Computing and Machine Intelligence, Department of Computer Science and Engineering
[2] Key Laboratory of Shanghai Education Commission for Intelligent Interaction and Cognitive Engineering,
Shanghai Jiao Tong University, Shanghai, China

Abstract— **This paper aims to explore whether different persons share similar patterns for EEG changing with emotions and examine the performance of cross-subject and cross-gender emotion classification from EEG. Movie clips are used to evoke three emotional states: positive, neutral, and negative. We adopt differential entropy (DE) as features, and apply linear dynamic system (LDS) to do feature smoothing. The average cross-subject classification accuracy is 64.82% with five frequency bands using data from 14 subjects as training set and data from the rest one subject as testing set. With the training set expanding from one subject to 14 subjects, the average accuracy will then continuously increase. Moreover, fuzzy-integral-based combination method is used to combine models across frequency bands and the average accuracy of 72.82% is obtained. The better performance of using training and testing data both from female subjects partly implies that there should be gender differences in EEG patterns when processing emotions.**

Keywords— **Emotion recognition, EEG, brain-computer interface, subject-independent, gender difference**

I. Introduction

Emotion is a general definition for subjective cognition experiences, including psychological and physiological states aroused by ones feelings, thinking, and behaviors. With the development of artificial intelligence, affective computing based on computer systems is considered to make human-machine interaction more friendly and convenient [1].

Traditional affective computing mainly focuses on recognizing emotion from facial expressions and speech. Recently, the development of brain-computer interface encourages studies on electroencephalography (EEG) based emotion recognition. Previous studies used different kinds of stimuli (pictures, music, videos) to evoke different emotions (happiness, sadness, curiosity, anger, fear, etc.). They extracted different features (power spectral density, differential entropy, etc.), and applied different models (support vector machine, k-nearest neighbor, extreme learning machine, etc.) on EEG data sets to recognize emotional states [2] [3] [4]. Although all of these studies have indicated that EEG signals do change with emotional states in specific patterns, to our best knowledge, EEG patterns and models used in the existing studies were almost subject-dependent, which means that these models were trained from EEG data of one subject, and could only be used to predict the emotional states of the same subject. In our previous work, Duan *et al.* [5] and Zhu *et al.* [4] have both confirmed that EEG patterns changing with emotions are relatively stable in different experiments of the same subject, but they still haven't reached a definite conclusion whether these patterns are universal across different subjects.

The main goal of this study is to find a subject-independent EEG model for emotion recognition, so as to explore whether different persons share similar patterns for EEG changing with emotions. Moreover, current findings have indicated that gender is a potential factor modulating emotional processing and its underlying neural mechanisms, which leads to individual differences in EEG models [6]. Therefore, the gender differences of EEG patterns are considered as one of the major factors affecting the performance of cross-subject models.

In this paper, we use movie clips as stimuli to evoke three emotional states of subjects: positive, neutral, and negative. We adopt differential entropy (DE) as EEG features, which were demonstrated to work well for classifying emotions [5]. Support vector machine (SVM) is used to train EEG models on various data sets from different subjects, and fuzzy-integral-based method is applied to combine models trained across different frequency bands and genders.

II. Experiment

A. Stimuli

The 15 Chinese movie clips lasting about 4 minutes long each were selected as stimuli to evoke three emotional states: positive, neutral, and negative. All of these movie clips came from popular Chinese movies. Feedbacks from the subjects showed that the stimuli could arouse typical emotional states successfully during the experiment.

B. Subjects

Totally 15 subjects (7 males, 8 females, aged 18 to 28) participated in this experiment. All the subjects are right-handed and healthy, with enough sleep the day before experiment. They were all told the purpose and procedure of the experiment and the harmlessness of the equipment.

D.A. Jaffray (ed.), *World Congress on Medical Physics and Biomedical Engineering, June 7-12, 2015, Toronto, Canada*,
IFMBE Proceedings 51, DOI: 10.1007/978-3-319-19387-8_288

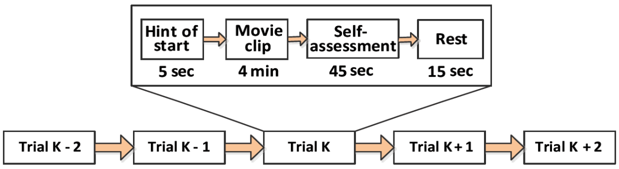

Fig. 1: Procedure of our experiment.

C. Procedure

The ESI NeuroScan System with a 62-channel electrode cap was used to collect EEG signals from the subjects at a sampling rate of 1000Hz. Movie clips were played on a big screen with about 60s spare time in between. During the interval, subjects were asked to fill the feedback form. Fig. 1 shows the procedure of the whole experiment.

III. METHODS

A. Feature extraction

We firstly down sampled the EEG data to 200Hz, and did band-pass filtering from 1Hz to 75Hz to remove the EOG and EMG artifacts. In this paper, we chose differential entropy (DE) as frequency domain features for emotion recognition. It has been proved that DE is equivalent to the logarithm of energy spectrum [7]. Also, DE has been demonstrated to perform better than energy spectrum [5].

In this paper, DE was calculated by Short Time Fourier Transform (STFT) with a 1s non-overlapping hanning window. Thus, 1s long EEG signal from one channel was mapped to five common-used EEG frequency bands (Delta: 1-3Hz, Theta: 4-7Hz, Alpha: 8-13Hz, Beta: 14-30Hz, and Gamma: 31-50Hz). As the electrode cap had 62 channels, there were totally $62 \times 5 = 310$ features of one sample.

B. Feature smoothing

In order to remove rapid changes which do nothing to continuous emotional activities from EEG signals, linear dynamic system (LDS) [8] was introduced here to smooth DE features. It has also been shown that LDS is an effective way to remove noises from EEG features [5].

C. Classification and model combination

To evaluate the quality of every experiment, we first used 60% of samples from one subject to train an individual model, and used the rest 40% of samples from the same subject as testing data.

After that, we focused on choosing training and testing data from different subjects. We selected one subject and used all of the 3394 samples from this subject as testing data. Then, the training data set was continuously expanding from the remaining one subject to 14 subjects, excluding the testing subject. Consequently, we used data from 14 subjects to train models across 5 frequency bands respectively, and test these models using the same data from the rest one subject. Furthermore, cross-gender classification was done by using two models: to train models from male subjects and female subjects, respectively, and predict the data from the same or different genders. Finally, a specific model combination algorithm was applied to combine these single models across different frequency bands and genders.

In this paper, linear support vector machine (SVM) was chosen as classifier, and fuzzy-integral-based method was chosen to do model combination as well.

The concept of fuzzy integral was introduced by Sugeno [9], which is the integral of a real function with a fuzzy measure. The fuzzy measure is defined as follows:
Let X be a finite index set $X = \{1, ..., n\}$.

Definition 1 A fuzzy measure μ defined on X is a set function $\mu : P(X) \to [0,1]$, where $P(X)$ indicates the power set of X, satisfying:
(i) $\mu(\emptyset) = 0$, $\mu(X) = 1$,
(ii) $A \subseteq B \Rightarrow \mu(A) \leq \mu(B)$,

In this paper, we used one of the most representative definitions for fuzzy measure, which is the Choquet integral.

Definition 2 Let μ be a fuzzy measure on X. The discrete Choquet integral of a function $f : X \to \mathbb{R}^+$ with respect to μ is defined by

$$C_\mu(f(x_1), ..., f(x_n)) := \sum_{i=1}^{n} (f(x_i) - f(x_{i-1}))\mu(A_i), \quad (1)$$

where $\cdot_i$ indicates that the indices have been permuted so that $0 \leq f(x_1) \leq ... \leq f(x_n) \leq 1$. Also $A_i := \{x_i, ..., x_n\}$, and $f(x_0) = 0$.

Here, we chose the square error criterion to get the fuzzy measure μ, by minimizing the square error calculated from training data.

IV. RESULTS AND DISCUSSIONS

A. In-subject classification

Table 1 shows the in-subject classification results of individual models, which are trained from 60% of samples from each subject, and used to predict the rest of 40% samples from the same subject. EEG features used here are 310 DE features of the total frequency bands. The average classification accuracy is 90.97±6.68%. Therefore, we can see that there are stable patterns of EEG changing with emotions for individual subjects.

B. Cross-subject classification

Fig. 2 shows the average results of cross-subject classification using models trained on the data from different numbers of subjects. We select one subject as testing subject, and then,

Table 1: Classification accuracies using individual models

Subject	1	2	3	4	5	6
Accuracy(%)	92.99	85.12	90.90	94.87	77.60	100
Subject	7	8	9	10	11	12
Accuracy(%)	100	94.87	91.76	89.88	86.49	92.77
Subject	13	14	15	Ave.	Std.	
Accuracy(%)	84.47	82.88	100	90.97	±6.68	

sort the rest 14 training subjects randomly. Samples from one training subject are treated as one unit, added into the training data set one after another. EEG features used here are 310 DE features of total frequency bands. From Fig. 2, we can see that with more and more subjects' EEG data added into the training set, the cross-subject classification accuracies will become higher and higher generally. Therefore, we can get an observation: A universally applicable EEG model for emotion recognition can be trained using data collected from enough subjects. Nevertheless, there are still some slight ups and downs during this increasement, which are probably caused by noises from the newly-added training samples, or by the individual and gender differences of EEG patterns.

Table 2 shows the classification accuracies across 5 frequency bands, where we use samples from all 14 subjects to train models, and test samples from the rest one subject. We also use fuzzy integral to combine single models of 5 frequency bands. The "Total" here means we arrange 62 features from each band together in the sequence of Delta, Theta, Alpha, Beta, and Gamma for one sample. Considering the results of single models, the average classification accuracy for total bands is 64.82±11.10%. This implies that subject-independent patterns for EEG changing with emotions do exist in DE features. It may be noticed that the classification results are near to the chance level when using Subject 11 as testing subject. One possible reason is that besides the common stable patterns, there are still individual differences in DE features.

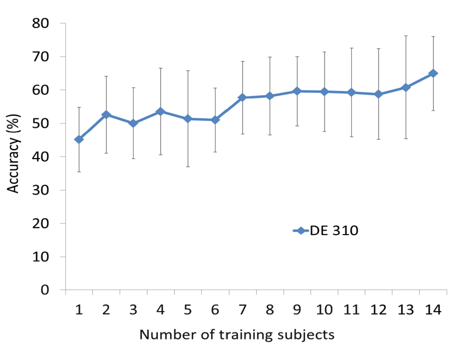

Fig. 2: Average classification accuracies using models trained from different numbers of subjects.

Table 2: Classification accuracies using cross-subject models

Test Sub.	Delta (%)	Theta (%)	Alpha (%)	Beta (%)	Gamma (%)	Total (%)	Fuzzy (%)
1	51.30	44.73	49.53	60.11	**72.57**	57.34	69.77
2	58.90	54.86	33.44	36.15	62.61	66.53	**69.89**
3	49.06	45.14	55.30	45.20	73.22	75.60	**82.70**
4	51.77	48.29	48.17	55.83	60.58	**87.15**	73.45
5	57.31	67.06	74.19	41.22	48.11	59.22	**76.52**
6	38.95	63.91	58.78	41.81	63.46	55.45	**82.29**
7	29.88	54.77	31.64	53.80	46.38	43.37	**75.69**
8	65.09	34.47	55.30	50.74	68.97	66.09	**86.59**
9	37.80	34.94	59.72	56.13	42.10	62.96	**71.89**
10	33.62	58.90	42.34	72.16	89.10	76.13	**93.99**
11	41.19	34.47	42.02	42.52	33.00	**49.18**	48.26
12	49.23	45.11	44.52	30.55	37.42	**71.42**	52.45
13	59.10	35.06	42.84	**71.92**	63.46	67.47	71.77
14	52.65	33.35	42.28	43.64	46.79	**63.52**	56.60
15	53.95	32.53	62.32	54.54	59.67	71.39	**80.41**
Ave.	48.65	45.84	49.49	50.42	57.83	64.82	**72.82**
Std.	10.24	11.78	11.46	11.99	15.31	11.10	12.55

We use fuzzy integral method to combine models trained by features from different frequency bands. We can see also from Table 2 that the average model combination result is 72.82%, which is 8% higher than the average classification result on total bands. Although the feature complexities are both 310 dimensions for the model combination results and the results on total bands, it is obvious that combining models of 5 frequency bands can get better performance in most cases. One reason is that each frequency band may weight different in emotion recognition problems.

C. *Cross-gender classification*

We firstly select a testing subject, and then, separately train two models from the rest male subjects, and the rest female subjects. These two models are both used to do cross-gender classification on the data of the testing subject (seeing Fig. 3 and Fig. 4). Fuzzy integral is also used here to combine the outputs of the male model and the female model. Here, "Same gender training" or "Different gender training" means the training and testing samples come from the same or different genders.

Seeing the cross-gender classification results of male testing subjects in Fig. 3, the average performance of using female models is a little bit higher than using male models. However, considering the individual cases, only results of four testing subjects fit this law, and other three go against. It is probably because unknown individual differences exist among EEG patterns of male subjects. On the other hand, examining the results of female testing subjects in Fig. 4, the average performance of using female models is about 15%

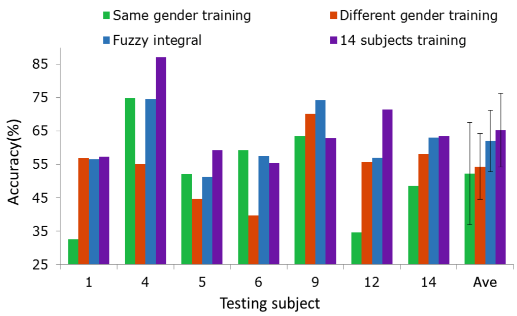

Fig. 3: Cross-gender classification accuracies using males as testing subjects

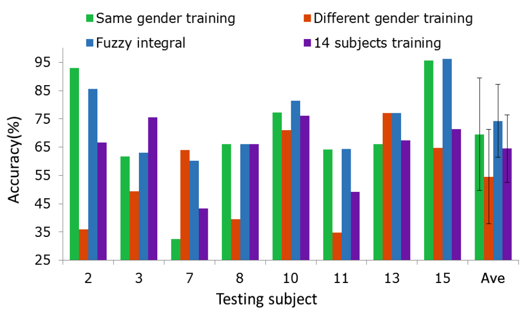

Fig. 4: Cross-gender classification accuracies using females as testing subjects

higher than using male models, and it is also about 5% higher than using models trained from 14 subjects. This difference implies that there must be gender differences between females' EEG patterns and males' EEG patterns, but whether these differences are induced by emotion changes are still not sure from this study. Also, for female testing subjects, using single female model, or combining male and female models would both improve the performance of cross-subject emotion classification from EEG.

Moreover, in both two figures, the average performance of female models is better than male models. This, to some extent, indicates that females share more similar EEG patterns when emotions are evoked, while males have more individual differences among their EEG patterns. As we have demonstrated that subjects share a stable pattern for EEG changing with emotions in the previous section, the results of cross-gender classification could illustrate that this stable pattern is likely to behave more obvious among female subjects.

V. CONCLUSION

In this paper, a series of experiments were held to collect EEG data of three emotional states (positive, neutral, negative). The 15 subjects (7 males, 8 females) participated in this experiment, and their emotions were all evoked successfully. We chose training and testing data sets from different subjects to do cross-subject and cross-gender emotion classification from EEG. According to the experimental results, we found that different persons do share similar patterns for EEG changing with emotional states, and this stable pattern is likely to behave more obviously among female subjects. Our experimental results have demonstrated that a universally applicable EEG model for emotion recognition can be trained using data collected from enough subjects. In addition, fuzzy-integral-based method works well for combining single models across different frequency bands and genders.

ACKNOWLEDGEMENTS

This work was partially supported by the grants from the National Basic Research Program of China (Grant No.2013CB329401), the National Natural Science Foundation of China (Grant No.61272248), and the Science and Technology Commission of Shanghai Municipality (Grant No.13511500200).

REFERENCES

1. Picard RW. *Affective computing*. MIT press 2000.
2. Wang XW, Nie D, Lu BL. Emotional state classification from EEG data using machine learning approach *Neurocomputing*. 2014;129:94-106.
3. Jenke R, Peer A, Buss M. Feature Extraction and Selection for Emotion Recognition from EEG *IEEE Transactions on Affective Computing*. 2014;5:327-339.
4. Zhu JY, Zheng WL, Peng Y, Duan RN, Lu BL. EEG-based emotion recognition using discriminative graph regularized extreme learning machine in *Proceedings of International Joint Conference on Neural Networks*:525-532 2014.
5. Duan RN, Zhu JY, Lu BL. Differential Entropy Feature for EEG-based Emotion Classification in *Proceedings of the 6th International IEEE/EMBS Conference on Neural Engineering*:81-84 2013.
6. Stevens JS, Hamann S. Sex differences in brain activation to emotional stimuli: A meta-analysis of neuroimaging studies *Neuropsychologia*. 2012;50:1578-1593.
7. Shi LC, Jiao YY, Lu BL. Differential entropy feature for EEG-based vigilance estimation in *Proceedings of the 35th Annu. Int. Conf. IEEE EMBS*:6627-6630 2013.
8. Shi LC, Lu BL. Off-line and on-line vigilance estimation based on linear dynamical system and manifold learning in *Proceedings of the 32nd Annu. Int. Conf. IEEE EMBS*:6587-6590 2010.
9. Sugeno M. *Theory of fuzzy integrals and its applications*. Ph.D. Thesis. Tokyo Institute of Technology 1974.

Address of the corresponding author:

Author: Bao-Liang Lu
Institute: Shanghai Jiao Tong University
Street: 800 Dong Chuan Road
City: Shanghai
Country: China
Email: bllu@sjtu.edu.cn