

EEG-based Emotion Word Recognition

Weiwei Dong, Panpan Wang*, Yazhou Zhang, Tianshu Wang, Jiabin Niu and Shengnan Zhang

Department of Computer Science and Technology, Tianjin University, Tianjin 300072, China

*Corresponding author

Abstract—Emotion recognition rapidly gains interest from research community, which gives computer the ability to recognize, understand, express, and adapt to human emotions to help computer more intelligent. Traditional emotion recognition modalities (e.g. face and voice) always could not identify the totally true emotion state. Recently, with the development of electroencephalogram (EEG) equipment, utilizing EEG signals to build emotion recognition system is more and more feasible and meaningful. In this paper, an abstract emotional words classification task is designed, which requires participants judge the emotion polarity (positive or negative) for given words. Based on EEG signals, a series of study is re-explored, including data collecting, feature extraction, feature selection, and classification model. Furthermore, the Event Related Potential (ERP) components and behavioral data analysis uncover some meaningful conclusion on emotion recognition.

Keywords—affective computing; Electroencephalogram(EEG); Emotion words recognition; Feature extraction; Classification; Event Related Potential(ERP)

I. INTRODUCTION

One of the key technologies of human-computer interaction is emotion recognition technology, which is getting more and more attention in the field of artificial intelligence. Emotion recognition is a vital part of affective computing, including speech based approaches, facial expression based approaches, text based approaches, gesture based approaches and physiological signal based approaches [1]. There are numerous kinds of physiological signals, such as electroencephalogram (EEG), electromyography, electric skin resistance, skin temperature, photoelectric pulse, respiratory signal and so on [2]. The Emotion Computing Research team at the MIT lab in the United States first extracted various features from such physiological signals, and proved that it is viable that utilizing physiological signals to accomplish emotion recognition tasks. The research of EEG help researchers reveal the mechanism and cognition process of the brain for mankind. Inside the brain, the electric activities of billions of neurons on cerebral cortex form an integrated feedback, i.e. EEG [3]. All kinds of mindstates and emotion change would reflect different EEG on different positions of cerebral cortex [4]. Therefore, EEG have abundantly helpful information, and how to effectively process the EEG and extract the helpful information from EEG have great significance for the research of emotion recognition [5].

In the study of EEG based emotion recognition, the present researches only collect the emotion signals induced by film clips and music video [6]. Nevertheless, the process of human-computer interaction involves not only the interaction of visual and auditory modalities, but also the interaction of text [7]. Hence, studying the emotion change of a person will be also

interesting and meaningful, when he (she) is reading specific online novel, poem or prose. In this paper, an emotional words based experimental paradigm is designed. The target of this task is to study the difference of EEG when people watch positive or negative emotional words. Interesting phenomena were also uncovered: first, the effectiveness of EPR feature and supervised method of feature selection is proved; second, the traits of ERP waveform of the occipital electrode on this task are discovered; third, the different reaction for positive and negative words are discovered and discussed.

II. EXPERIMENTAL DESIGN

A. Experimental Material

A 160 emotional words list consisting of 80 positive words and 80 negative words, is selected randomly from China CNKI Emotion Dictionary. The selected words all contain two-characters. Each word is printed on a corresponding picture, with the same magnitude, color and typeface, as following figures:

Melancholy	Sulk	Curse
Embarrassment	Vexation	Decadent
Negative Words		
Adoration	Praise	Honor
Importance	Gratulation	Thankfulness
Positive Words		

FIGURE I. EXPERIMENTAL MATERIAL

B. Experimental Group

There are 32 experimental participants in total, including 18 males and 14 females. The age is between 22 and 25. All subjects are right-handed with normal vision or correction vision, and have had no mental or psychosocial illness, cerebral injury, or mental medication use.

C. Experiment Procedure

During the preparatory work, participants fill out a record form containing age, education, gender and other basic information, then they put on the EEG caps and relax for 1-2 minutes. After that, the experiment began. For helping participants be familiar with the experimental procedure, there is a practice. Participants are presented with 10 words (5 positive words and 5 negative words) one by one, and they need to judge the emotion polarity (positive or negative) of a word under 500ms. While if participants have not made a decision at the regular time, the time extends to 1000ms. After a judgement is given, a sign like '+' appears at screen center for a random time

about 800ms-1000ms, so as to help the participants focus on current experiment again. The words are given at random to avoid the priori speculation. In formal experiment, participants are given all the 160 words. The experimental setup is similar to practice. Different a short break after every 20 words there is a short break after every 20 words and participants decide when to restart. The procedure is shown in Figure 2.

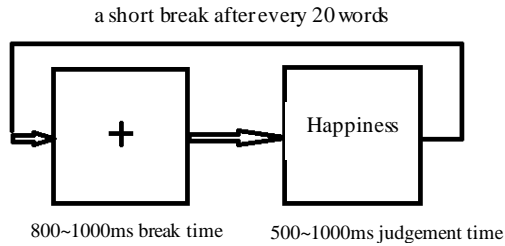


FIGURE II. EXPERIMENT PROCEDURE

D. Data Records

The EEG signals is recorded by EEG acquisition system, i.e. NeuroSCAN. The experiment flow (Figure 2) is design by Eprime-2. EEG signals is collected with 64-channel EEG caps. The electrodes are arranged in accordance with 10 ~ 20 international EEG systems. The average bilateral mastoid reference and scalp electrode resistance are all adjusted to below 5K ohms. The frequency for collecting EEG is set as DC ~ 70Hz and the sampling rate is set as 1000Hz.

III. FEATURE EXTRACTION

This experiment mainly extracts frequency band feature and event related potential (ERP) component feature based on EEG signals. The former is used to capture the signal changes in the whole time window when the participants receive a stimulus. The latter captures the signal changes in a given period of time when the participant receives a stimulus. This specific time refers to the short time window that is between the participant receives the stimulus and the participant gives a judgment.

Due to the high dimension of the data, before data processing, we remove electrodes located on the edge of the brain according to the distribution of the scalp, and select part of the electrode only according to the brain's four major divisions (forehead, temporal, apex and occipital). The number of electrode is reduced to 54 (originally 64).

A. Frequency Band Feature

EEG is divided into five frequency bands in the frequency range (1-45Hz), that is, δ wave (1-4Hz), θ wave (4-8Hz), α wave (8-13Hz), β wave (13-30Hz), and γ wave (30-45Hz). These bands have their own physiological functions. For example, the activity of α wave is related to attention, and its energy value is negatively related to the activity of the scalp. It's especially active when you're closing your eyes. The combination of α wave, θ wave and beta β can reflect the excited state. In this experiment, the time domain signals in the range of 1000ms after the stimulus appeared are extracted, and then the time domain signals are transformed into the frequency domain. First of all, the EEG signal is windowed (Hanning window), then the windowed data are processed by fast Fourier

transform, finally the power spectrum value are processed by logarithmic transformation.

B. ERP Feature

The ERP component is mainly related to the cognitive function of the brain, which occurs within a period of time after the participant has been stimulated, and reaches a peak in a certain period of time. For example, a negative potential N2 of forehead central is related to the cognitive control of uncertainty; the positive peak after the stimulus of 300ms is called P3, which is related to attention, memory processing and information retrieval in the brain. In this experiment, according to the possible range of ERP component, four kinds of ERP features are extracted in the 800ms after stimulation, namely P1 (80-150ms), N1/P2 (150-250ms), N2/P3 (250-450ms), N4 (450-800ms), and then the average voltage amplitude in these components is obtained.

IV. DATA MODELING

EEG data has the characteristics of high dimensionality, a few samples and low signal-to-noise ratio [8], so the appropriate feature selection method will be helpful to the classification task. In this section the commonly feature selection method, and the classifier used in this paper will be introduced.

A. Feature Selection

According to whether there are data labels, the methods of feature selection could divide into two types: supervised feature selection and unsupervised feature selection. The experiment results of two types of methods would be compared in the next section.

1) Supervised Feature Selection Method

a) *ReliefF Algorithm*: In ReliefF algorithm[9], the correlation between features and categories is based on the ability to distinguish between close-spaced samples. The algorithm randomly chooses a sample R from the training set D and then finds the nearest neighbor sample H from the samples R. It is called Near Hit. Finding the nearest neighbor sample M from the samples of different classes of R, which is called NearMiss. Then update the weight of each feature according to the following rules: If the distance between R and Near Miss on a feature is less than the distance between R and Near Miss, it is useful to distinguish this feature from the nearest neighbor of the same category and different category, and the weight of this feature need to be increased. Conversely, If the distance between R and Near Hit in a feature is greater than the distance between R and Near Miss, it indicates that the feature has a negative effect on the neighbors that distinguish between the same category and the different category, and the weight of this feature need to be decreased. The above process is repeated m times, and finally the average weight of each feature is obtained. The larger the weight of the feature is, the stronger the classification ability of the feature is. On the contrary, the weaker the feature classification ability is. This algorithm can be directly called through Matlab function library.

b) *MRMR Algorithm*: MRMR [10] is an abbreviation of Minimum Redundancy Maximum Relevance Algorithm, using

mutual information as a standard to control feature selection. The mutual information between two random variables x and y can be defined as:

$$I(x; y) = \iint p(x, y) \log \frac{p(x, y)}{p(x)p(y)} dx dy \quad (1)$$

where $p(x)$ and $p(y)$ are the marginal probability density functions of x and y , respectively, and $p(x, y)$ is joint probability distribution. If $I(x; y)$ is equal to 0, x and y are statistically independent.

The multivariate MRMR algorithm is designed to implement two criterions. The first is the Maximum Relevance, which maximize the mutual information between each feature x_i and the target vector y . The second is the Minimum Redundancy, which minimize mutual information between two different features. This algorithm can obtain approximate optimal features by the forward selection method. Suppose we have chosen a set including k features, the next feature is chosen based on D-R:

$$\max_{x_j \in X - S_k} [I(x_j; y) - \frac{1}{k} \sum_{x_i \in S_k} I(x_j; x_i)] \quad (2)$$

The code of this algorithm is provided by the original author.

2) Unsupervised Feature Selection Method

PCA Algorithm: Place The main idea of Principal Component Analysis (PCA) is to retain the direction of the largest variance in data variation, and to reduce the data dimension. First, the m -dimensional mean vector μ and the covariance matrix Σ of $m \times m$ are calculated over the entire dataset. Next, we compute the eigenvectors and eigenvalues of the covariance matrix, and rearrange the matrix according to the eigenvalue. Then select the largest k eigenvalues eigenvectors to form a matrix m of $m \times k$. The final data is handled as:

$$X_{pca} = A^T (x - \mu) \quad (3)$$

The code of this algorithm is provided by the original author.

B. Classifier

This experiment directly calls the matlab built-in classification algorithm support vector machine (SVM) to classify the data after feature selection. The purpose of the SVM is to construct a hyperplane between two categories and to maximize the margin between the two categories and hyperplanes. The selected kernel function is a linear kernel function.

V. EXPERIMENT RESULT ANALYSIS

A. Classification Model

In this section, we will compare two feature selection methods on the two kinds of extracted features using SVM classifier. The first one is ERP features and the second one is frequency features. We would like to verify which features are

more effective in distinguishing between positive and negative abstract emotions by experiments.

TABLE I. CLASSIFICATION ACCURACY OF DIFFERENT SELECTION METHODS WITH DIFFERENT FEATURES

Male Participants	<i>ReliefF</i>	<i>MRMR</i>	<i>PCA</i>
ERP	0.6500	0.7125	0.6430
Frequency Band	0.6125	0.6225	0.6078

Female Participants	<i>ReliefF</i>	<i>MRMR</i>	<i>PCA</i>
ERP	0.6100	0.6511	0.5736
Frequency Band	0.5725	0.5925	0.5627

Table 1 shows the average accuracy of the SVM classifier for male and female participants respectively. From Table 1, we can figure out that the supervised feature selection method is more effective than the unsupervised algorithm. Compared with the frequency features, the ERP features, as the characteristics of EEG signal, are more effective in emotion classification tasks. The reasons are as follows:

1) The selected emotion words may not trigger the change of emotions. Therefore, there are not apparent changes of power spectrum values in each frequency band of EEG signal, resulting in the indivisibility of data.

2) The reason, why ERP features are effective, is that when making the positive and negative judgments, the ERP components have large differences. We will detail this part in the next section.

B. ERP Component Analysis

ERP is the electroencephalogram extracted from EEG signal. The latency and waveform are important features of ERP. To obtain these features, original EEG signal have to be processed offline and the steps are as follows:

1) The fusion of behavior data and EEG signal. The data can be grouped and superposed according to different behavioral response standards.

2) The removal of eye artifacts. EEG artifacts have a significant negative effect on EEG recording.

3) The segmentation of EEG signal. According to the type of event stimulus and the interval of stimulation, the EEG signal can be segmented. The processed data after the average is used for ERP analysis.

4) The baseline correction. Eliminating the deviation of EEG from baseline.

5) The removal of artifacts. Removing the higher amplitude data from the data after the segmentation.

6) Superposition averaging. After superimposing the processed segmented waveforms, the average amplitude within the stimulus interval can be found.

7) Digital filtering. Eliminating the noise components.

After investigation, the cognitive processing is related to the occipital cortex in the brain. Therefore, we analyzed the ERP waveforms of electrode Oz as shown in Figure 3.

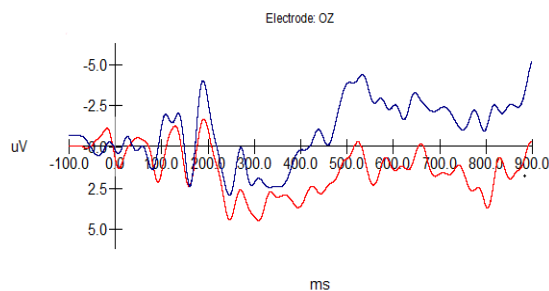


FIGURE III. ELECTRODE OZ

In this paper, the 100ms before the stimulus and the 900ms after the stimulus of EPR signal is selected waiting to be processed. According to the type of stimulus, EEG data are segmented. For every segmentation, we calculate the average values, i.e. the ERP waveform shown in Figure 3. The horizontal axis represents -100ms to 900 ms and the vertical axis represents the voltage amplitude. The red and blue lines represent the waveforms of positive and negative emotional words respectively. As shown in Figure 3, we can find that within the first 300ms, the waveform are roughly the same. However, after 300ms, the two stimuli appear a big difference. The negative words' voltage value is negative in the later period while the positive words' is positive. This shows the difference of brain cognition in judging emotional words.

What's more, by observing the waveform, we can also find that both stimuli around 100ms and 200ms induced negative peaks, which were called N1 and N2 respectively. N1 is mainly due to the stimulation of the picture and the transfer of cognition induced N2. In addition, around 300ms, there is also a positive waveform induced. This peak is called P300, which is important as the ERP component. This peak is triggered by the stimulus which produces a conscious behavior.

C. Behavioral Data Analysis

The behavioral data mainly include the reaction time and the judgment of part of speech. First, we average the judgement time. Facing the positive emotion words and negative words, the average reaction times are 612 ms and 656ms respectively. The participants react faster to positive words than negative words. Second, The average numbers of errors when judging the positive emotions and negative emotions are 6.2 and 11.5 respectively. Therefore, the participants are more confident when judging the positive words than negative words.

VI. CONCLUSION

In this paper, the building emotion recognition model and data analysis are carried out based on EEG signals. In emotion recognition model, two features, i.e. EEG frequency feature and EPR feature, are extracted. And experimental results show that EPR feature are more effective than the EEG frequency feature. Further, the supervised method of feature selection is more effective than the unsupervised algorithm for classification

modeling. According to the analysis of the ERP waveform of the occipital electrode, we found that this emotion emotional words classification task could induce N1, N2, P300 and other components. What's more, after 300ms, the judgement of positive and negative words triggered positive and negative waveforms respectively. Finally, by the analysis of behavior data, we find that the participants give faster and more accurate judgement facing positive words than negative words.

However, there are some shortcomings in this experiment. The power spectral features can not be modeled effectively and this illustrates that abstract emotion words may not trigger the change of human emotion in some situation. Therefore, more specific emotional words should be selected in the following experiment, e.g. the foods (apple, beef and so on). In addition, when choosing words, we should make an investigation in advance and grasp the properties of positive and negative of each word before the experiment.

REFERENCES

- [1] Calvo R A, D'Mello S. Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications[J]. *Affective Computing IEEE Transactions on*, 2010, 1(1):18 - 37.
- [2] Koelstra S, Muhl C, Soleymani M, et al. DEAP: A Database for Emotion Analysis; Using Physiological Signals[J]. *IEEE Transactions on Affective Computing*, 2011, 3(1):18 - 31.
- [3] Rahnuma K S, Wahab A, Kamaruddin N, et al. EEG analysis for understanding stress based on affective model basis function[J]. *Digest of Technical Papers - IEEE International Conference on Consumer Electronics*, 2011:592 - 597.
- [4] Subi D, Anuja H S. Detection of Automobile Drivers Stress from Physiological Signals[J]. *Int J Comput Sci Mobile Comput*, 2014, 3: 454-458.
- [5] Lotte F, Congedo M, Lécuyer A, et al. A review of classification algorithms for EEG-based brain-computer interfaces[J]. *Journal of Neural Engineering*, 2007, 4(2).
- [6] Khosrowabadi R, Wahab A, Kai K A, et al. Affective computation on EEG correlates of emotion from musical and vocal stimuli[C]// *Neural Networks, IEEE - INNS - ENNS International Joint Conference on*. IEEE, 2009:1590-1594.
- [7] Briesemeister B B, Kuchinke L, Jacobs A M. Emotion word recognition: Discrete information effects first, continuous later?[J]. *Brain Research*, 2014, 1564(4).
- [8] Lotte F, Congedo M, Lécuyer A, et al. TOPICAL REVIEW A Review of Classification Algorithms for EEG-based Brain-Computer Interfaces[J]. *Journal of Neural Engineering*, 2007, 4(2).
- [9] Kira K, Rendell L A. The feature selection problem: Traditional methods and a new algorithm[C]// *Aaai*. 1992, 2: 129-134.
- [10] Peng H, Ding C. Minimum redundancy and maximum relevance feature selection and recent advances in cancer classification[J]. *Feature Selection for Data Mining*, 2005, 52.
- [11] Wold S, Esbensen K, Geladi P. Principal component analysis[J]. *Chemometrics and intelligent laboratory systems*, 1987, 2(1-3): 37-52.
- [12] Bennett K P, Campbell C. Support vector machines: hype or hallelujah?[J]. *Acm Sigkdd Explorations Newsletter*, 2000, 2(2): 1-13.