Sensors and Materials, Vol. 35, No. 7 (2023) 2337–2354
MYU Tokyo

S & M 3326

Model for Effectively Extracting Mixed Features and Classifying Emotions from Electroencephalograms

Shijing Zhang,[1] Qunsheng Ruan,[2] Lixia Huang,[3] and Qingfeng Wu[4*]

[1]Institute of Education, Xiamen University,
Siming South Road No. 422, Siming District, Xia Men 361000, China
[2]Department of Nature Science and Computer, Ganzhou Teachers College,
Gannan Da Road No. 35, Rongjiang New District, Gan Zhou 341000, China
[3]School of Information, Mechanical and Electrical Engineering, Ningde Normal University,
College Road No. 1, Dongqiao Development District, Ning De 352100, China
[4]School of Film, Xiamen University, Siming South Road No. 422, Siming District, Xia Men 361000, China

(Received January 7, 2023; accepted June 7, 2023)

Keywords: EEG, emotion recognition, wavelet packet transform, chaos theory, mixed feature set

Emotion recognition is gaining increasing attention from researchers and health professionals, and it has become a hot research topic in recent years. However, there have been no fruitful achievements on emotion recognition owing to the low quality of features extracted from the original electroencephalogram (EEG) and low emotion recognition rate. In this study, we extract time–frequency domain and chaotic features from human electroencephalogram signals using wavelet packet transform and chaos theory. The two types of features are combined to generate a mixed feature set with low dimensions. Then, we propose a one-to-one long short support vector machine (LS-SVM) classifier based on the Gauss function for multiclass classification problems. Focusing on the indivisible classified regions in the decision model, we combine the distribution of samples with distance between samples and separating hyperplanes, and then establish a discriminant function for fuzzy classification in the one-to-one LS-SVM. Three groups of comparative experiments, namely, feature selection and dimensionality reduction, the effectiveness of optimized LS-SVM, and the classification of different types of feature set, are conducted. The experimental results for the mixed feature set demonstrate that the proposed classification model has a competitive performance.

1. Introduction

Human emotion recognition is a popular interdisciplinary research topic and has attracted attention from fields such as neuroscience, physiology, and computer science. In computer science, emotion recognition plays an important role in human–computer interaction and is useful for modeling more user-friendly systems.(1) Although recent advances in artificial intelligence (AI) have enabled computers to solve many problems as efficiently as humans, they

*Corresponding author: e-mail: qfwu@xmu.edu.cn
https://doi.org/10.18494/SAM4312

ISSN 0914-4935

are still unable to perform certain tasks. Therefore, in order to strengthen the communication between humans and machines, especially to develop a more perceptual communication between humans and computers, and to promote further development of human–computer interaction technology, it is important to understand human emotions and recognize them in the context of AI.[2,3]

Emotion classification for humans is studied using data of biological electroencephalogram (EEG) signals, which are formed by the weak current generated by the activity of neurons in the brain. EEG signal data can be collected using professional biosensor equipment. EEG-based emotion classification is considered more reliable than other approaches such as facial expression and gesture analysis, as EEG data cannot be easily faked. In this study, we mainly focus on two main tasks: extracting high-level informative features from EEG data and optimizing an existing classification model for emotion recognition. We use the SJTU Emotion EEG Dataset (SEED), for which the experiments of emotion induction, data collection, and preprocessing have been systematically completed. Our work mainly focuses on the feature extraction, feature selection and dimensionality reduction, training, and validation of classification models. In summary, the novelty of this paper is as follows.

(I) Extraction of high-level features from EEG signals. Unlike traditional methods of frequency band equalization, we use wavelet packet transform to extract five major frequency bands of EEG signals and then calculate the power spectral density (PSD) and energy characteristics of each frequency band separately on the basis of the physiological classification of EEG signals. In addition, aiming at the randomness, nonlinearity, and nonstationarity of EEG signals, we apply chaos theory to calculate chaotic parameters of the signal and use them as chaotic features. The time–frequency domain and chaotic features are combined to generate a novel set of mixed features for emotion classification.

(II) Building a classification model for emotion recognition and its optimization. We use long short support vector machine (LS-SVM) based on Gauss kernels as a classifier. The cross-validation method is applied to test the effectiveness of the model and the grid search method to optimize parameters.

(III) We divide emotions into positive, negative, and neutral categories on the basis of the EEG signal, and use a one-to-one classification approach to solve the multiclassification problem of the traditional SVM. In view of the possible nonclassifiable regions, we propose a new discriminant function to solve the problem of fuzzy classification, taking into account the distribution of samples and the distance between samples and separating hyperplanes.

The rest of this paper is as follows. The related works are discussed in Sect. 2. The process of mixed feature extraction is annotated in detail, and the design of emotion recognition classification is described in Sect. 3. Various experiments are performed, and the experimental results are showcased and analyzed in Sect. 4. Section 5 concludes our work.

2. Related Work

Recently, emotion recognition has attracted attention from researchers of behavioral sciences, medicine, and AI. In AI, a number of studies have been conducted to classify human emotions

based on physiological signals, and comprehensive research results have been emerging. Compared with other features such as facial expression, voice intonation, and body language, physiological signals have shown better results. However, data collection from these signals is considered difficult as these signals are noisy and have spontaneous and uncontrolled characteristics. Because these signals are difficult to disguise and hide, they are informationally efficient and yield highly authentic and reliable results. Schaaff and Schultz used the international affective picture system, which is intended to induce three types of emotions, namely, positive, negative, and neutral, in the subjects.(4) The EEG signals of the prefrontal region of the subjects were collected as experimental data to study the three types of emotions. Their experimental results show that the recognition accuracy is about 0.47. Lin et al. introduced the energy characteristics of EEG signals into the study of emotion recognition.(5) Considering that the energy states of the left and right brain are not completely symmetrical in different emotional states, two EEG features based on the asymmetric energy difference and asymmetric entropy were proposed for the first time. This study not only provided good results in emotion recognition, but also confirmed the relevant theoretical research of neurophysiology from the experimental viewpoint.(6,7) Koelstra et al. selected 40 music videos as motivating materials to conduct situational experiments on 32 subjects.(8) In accordance with the situation dimension model, they collected physiological signals of subjects in different situations, including 32 channels of EEG signals and eight groups of peripheral physiological signals, namely, skin electricity, skin temperature, breathing, blood pressure, and eyespot, and recorded the subjects' facial conditions with video cameras. Then, a database named DEAP for emotion recognition research was formed. The database also provides a subjective evaluation of each subject's own situation in four dimensions: effective price, aesthetic sensitivity, goodness trend, and preference in each experiment. Each group of data collected in experiments by each group of physiologists was labeled.(9,10)

The general time-domain statistical characteristics of electrical signals can also be applied to EEG signals. These measures are simple and easy to calculate and can yield good recognition results.(11,12) To remove the disadvantage that the time-domain characteristics cannot show the frequency information of the signal, researchers have added frequency-domain analysis in the preprocessing of EEG.(13) The process steps are as follows.

Firstly, the original time-domain signal is converted to the frequency domain to obtain the spectrum.

Secondly, the frequency band is decomposed into five sub-bands (δ, θ, α, β, and γ) that are closely related to human psychological activity.(14,15)

Finally, the five sub-bands are analyzed or counted.

There are some methods that can be used to deal with EEG. Fourier transform (FT) is usually used for time–frequency domain conversion. The signal is projected onto a fixed orthogonal function system, and a set of transformation coefficients is used to represent the time function. Each spectrum indicates many parameters such as the phase and amplitude of a certain frequency component.(16,17)

Generally, in the current EEG-based emotion recognition research results, there are deficiencies such as poor data extraction quality and low emotion prediction accuracy.

3. Design of Mixed Feature Extraction and Emotion Recognition Classification Model

3.1 General model framework

In this section, we will showcase an overall framework of mixed feature extraction and emotion recognition classification using EEGs, which mainly includes the following three steps: data preprocessing, feature selection, and emotion recognition or classification. Our aim is to obtain the time–frequency domain eigenvector and chaotic eigenvector through data preprocessing. On the basis of the two types of eigenvectors, mixed eigenvectors are formed by feature selection. Finally, we will build a classification algorithm based on the mixed eigenvectors from the original EEG. The general model framework is shown in Fig. 1.

3.2 Feature extraction in time–frequency domain

The time–frequency domain features computed in this work mainly include PSD and energy features. Since the EEG signal itself is a time series, it is necessary to map the EEG signal to a specific frequency range before we extract its frequency domain features or time–frequency domain features. In general, the FT or wavelet transform (WT) is used to divide the effective

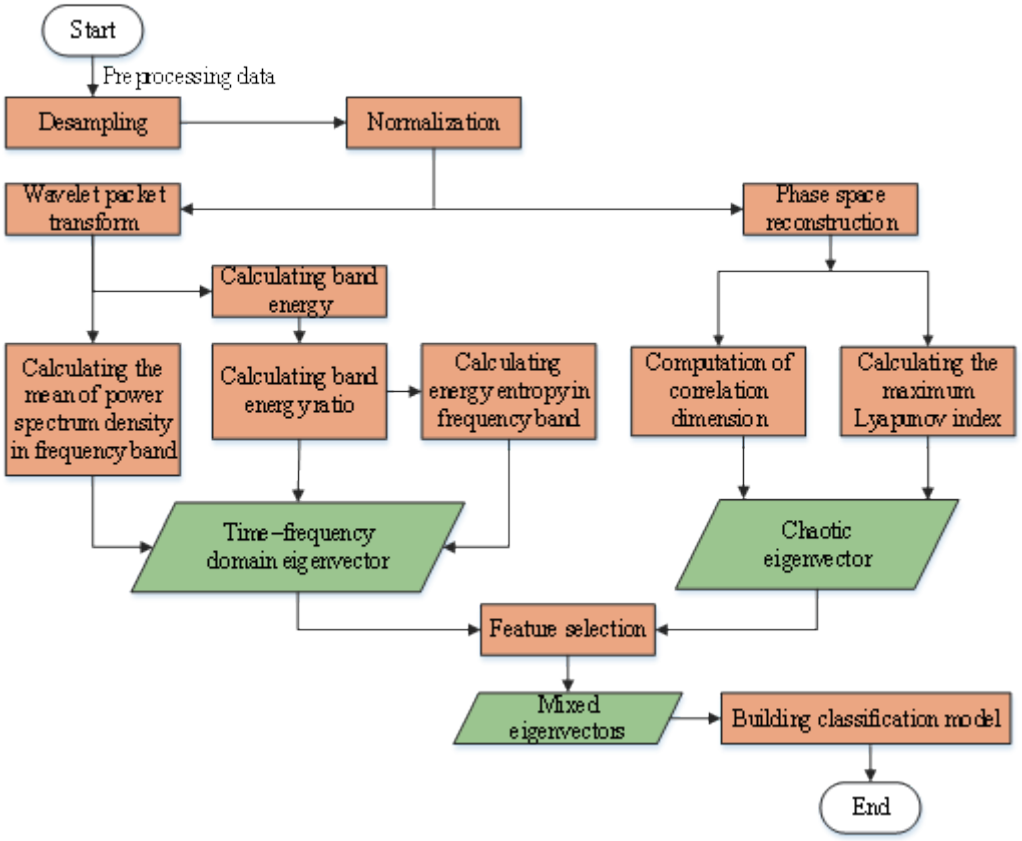

Fig. 1. (Color online) Framework of the model for mixed feature extraction and emotion recognition classification.

frequency range of EEG signals evenly, and then the required features in each frequency band are calculated. By fully considering the classification and characteristics of EEG signals, we extract the important frequency bands of EEG signals by wavelet packet transform. Then, in accordance with the classification theory of EEG signals, the relevant features are calculated in each frequency band. Specific elaborations are as follows.

3.2.1 Frequency band extraction based on wavelet packet transform

To reduce the computational complexity and computational time, we first down-sample the raw data and then extract features from SEED. Five important frequency bands of EEGs (δ, θ, α, β, and γ bands), of EEGs will be studied in the experiment. To make the frequency resolution close to the frequency range of EEG bands, the original EEG signal is decomposed into seven layers on the basis of a DB4 wavelet. The seventh layer contains a total of $2^7 = 128$ bands, and each band has a frequency range of 128/128 = 1 Hz.

If a sub-band space is obtained using an even-order high-pass decomposition filter, further decomposition of the sub-band space will result in two sub-bands of equal bandwidth and frequencies sorted from low to high. Conversely, if a sub-band space is obtained using an odd-order high-pass decomposition filter, further decomposition of the sub-band space will result in two sub-bands of equal bandwidth and frequencies arranged from high to low. The specific permutation order is independent of the number of passes through the low-pass decomposition filter. In this paper, the Wpfrqord function in Matlab is used to rearrange the coefficients corresponding to each node obtained by wavelet packet decomposition, which further verifies the accuracy of this rule. The seven-layer wavelet packet decomposition tree structure is obtained.

A red node can be reconstructed using the wavelet packet, which indicates that a δ band can be obtained as well as a green node, a yellow node, a blue node, and an orange node corresponding to the θ, α, β, and γ bands, respectively. Table 1 further shows the numbers of nodes of the wavelet packet tree and the corresponding bandwidth ranges in these five bands.

Furthermore, according to the above information, the nodes of the wavelet packet tree contained in each frequency band are reconstructed using the wavelet packet. Five important bands of EEG signals can be extracted. For the data of the sixth channel of EEG signals collected by a subject in a certain evoking experiment in the SEED as an example,[18] the original signal graph and the signal graph of each frequency band extracted are shown in Fig. 2.

Table 1
Each band contains nodes and a frequency range.

EEG signal frequency band	Standard bandwidth range	Included nodes and frequency band range
δ band	0–3 Hz	[6,0](0–2 Hz); [7,3](2–3 Hz);
θ band	4–7 Hz	[6,3](4–6 Hz); [7,5](6–7 Hz);
α band	8–13 Hz	[5,3](8–12 Hz); [7,10](12–13 Hz);
β band	14–30 Hz	[6,4](14–16 Hz); [4,3](16–24 Hz); [5,5](24–28 Hz); [6,9](28–30 Hz);
γ band	30–80 Hz	[6,8](30–32 Hz); [2,1](32–64 Hz); [3,6](64–80 Hz)

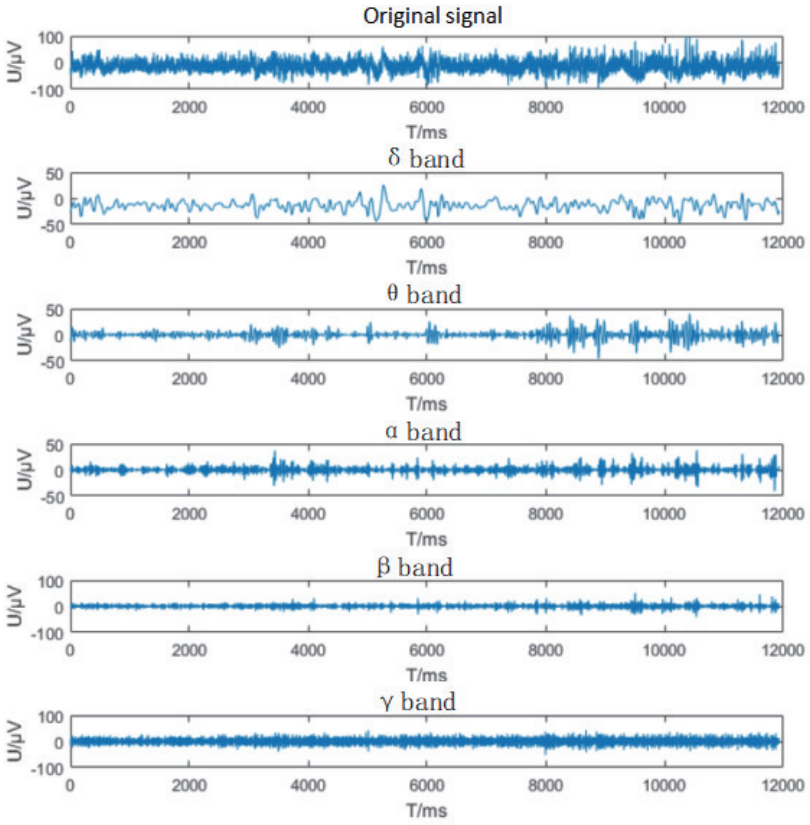

Fig. 2. (Color online) Channel 6 original signal and waveform of each frequency band of a sample.

3.2.2 PSD feature extraction

PSD is an effective index of the frequency characteristics of the time series. It can effectively characterize the distribution of power generated by random vibration using its frequency. Essentially, it describes the vibration originally existing in the time dimension from the aspect of frequency dimension.[19] Suppose that the autocorrelation function of the random signal $x(t)$ is defined as

$$r_t(k) = E\left[x(t)\overline{x(t+k)}\right], \tag{1}$$

where E represents mathematical expectations and $\overline{x(t+k)}$ represents the conjugate function of $x(t+k)$. When the autocorrelation function matches the condition of absolute integrability, its FT and corresponding inverse transform are respectively shown as follows.

$$S_t(w) = F\left[r_t(k)\right] = \int_{-\infty}^{+\infty} r_t(k) e^{-jwk} dk \tag{2}$$

$$r_t(k) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} S_t(w) e^{jwk} dw \tag{3}$$

When $k = 0$, the autocorrelation function $r_t(k)$ represents the power of the signal and its FT function $s_t(w)$ represents the power of the signal at the unit frequency, also known as PSD.

Here, we take the experimental data of a subject as an example. The histogram of the mean PSD of each band of the EEG signal in different emotional states is shown in Fig. 3.

Figure 3 shows that the average PSD of each frequency band varies significantly in different emotional states. Therefore, PSD can be used as one of the features in emotion recognition.

3.2.3 Frequency band energy feature extraction

As the EEG signal intensity varies in different emotional states, different emotional states of the brain can be distinguished by calculating the energy of each frequency band.[6] The experimental results illustrate that the EEG signal strength can be effectively expressed by the vibration amplitude of the EEG signal. The specific definition is

$$E_{j,i} = \sum \left[d_j^i(k) \right]^2, \tag{4}$$

where $d^i_j(k)$ represents the amplitude of the first node at the level of the wavelet packet decomposition tree. According to this definition, the energy of each frequency band can be calculated separately by obtaining each frequency band by wavelet packet transform. On this basis, the energy entropy of the wavelet packet can be calculated using Eq. (5). It shows that the

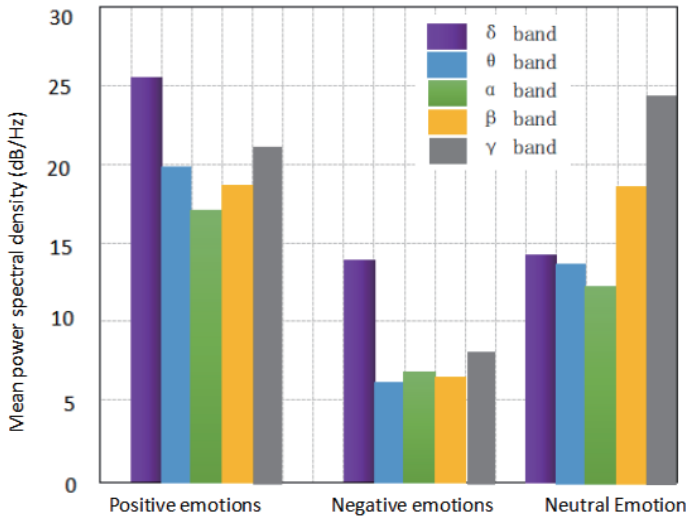

Fig. 3. (Color online) Mean PSDs of frequency bands under different emotional conditions.

higher the entropy and the more uniform the energy distribution in the corresponding frequency range, the stronger the randomness of its distribution.

$$w = -\sum \frac{E_{j,i}}{\sum E_{j,i}} \ln \frac{E_{j,i}}{\sum E_{j,i}} \tag{5}$$

Here, $\frac{E_{j,i}}{\sum E_{j,i}}$ represents the energy ratio of each frequency band, which can be used to describe the energy distribution between various emotional states in different frequency bands. The energy distribution of the channel 6 signal for the same subject under three different emotional states is shown in Fig. 4.

From Fig. 4, we can see that the energy distribution of each frequency band varies significantly in different emotional states. Particularly in the positive emotional state, the energy ratio of the δ frequency band is much higher than those of the other frequency bands, being 80% of the energy of the band. In addition to the energy ratio of each frequency band, the energy entropy of each frequency band is calculated and taken as the energy feature of the EEG signal.

3.3 Feature extraction of chaotic eigenvector

Besides the time-frequency feature, according to the randomness, nonlinearity, and nonstationarity of the EEG signal, we establish the chaotic feature vector from the EEG signal by introducing chaos theory into the EEG signal feature extraction to calculate the representative correlation dimension (CD) and the maximum Lyapunov exponent.

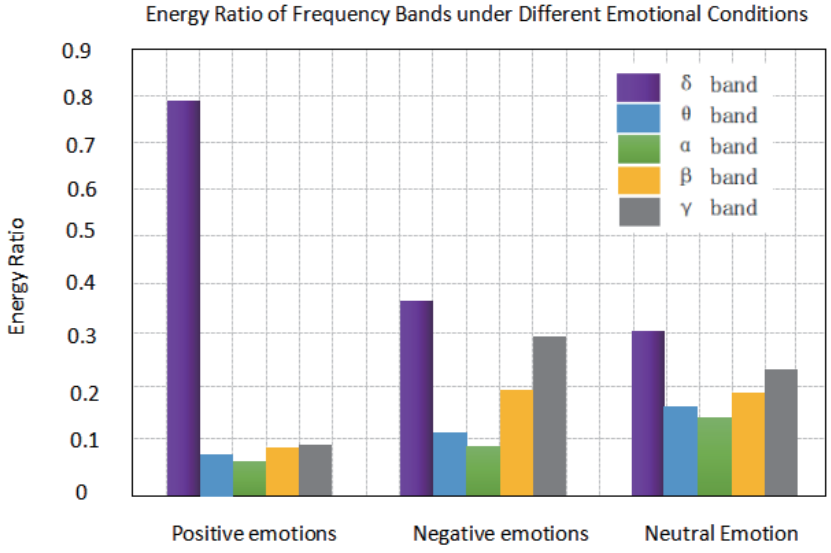

Fig. 4. (Color online) Energy distributions of frequency bands under different emotional conditions.

3.3.1 Correlation dimension

The CD is a widespread index in the nonlinear analysis of the chaotic time series. It can quantitatively describe the complexity of the internal structure of the chaotic system.[20] Firstly, the correlation integral function C(r) and D2 of the CD are defined as

$$C(r) = \frac{2}{N(N-1)} \sum_{i=1}^{N} \sum_{j=i+1}^{N} \theta(r - d_{ij}), r > 0, \tag{6}$$

$$D_2 = \lim_{r \to 0} \frac{\ln C(r)}{\ln r}, r > 0, \tag{7}$$

where $d_{ij} = \|X_i - X_j\|$ denotes the distance between two points in phase space, $\theta(x) = \begin{cases} 0, x<0 \\ 1, x \geq 0 \end{cases}$.

The CDs of the four selected channels in all samples are calculated by the above steps, and the range and mean of the CD of each channel in different emotional states are calculated and shown in Table 2.

From Table 2, it is clear that the CD of the EEG signal is not an integer, indicating that the EEG signal is indeed a nonlinear signal with chaotic characteristics. At the same time, it is ascertained that the CD values obtained are slightly different in the numerical distribution, although they overlap to some extent in different emotional states. Therefore, although the CD is not used as the main factor for emotion recognition, it contributes to improving the rate of emotion recognition.

3.3.2 Maximum Lyapunov exponents

The initial value of the chaotic system markedly affects the state of the chaotic system at each moment. For example, the separation speed between movement orbits of two original values, initially close pairs, in a chaotic system can exponentially grow with time. This phenomenon can be quantitatively described by the Lyapunov exponent.[21]

The Lyapunov exponent can represent the exponent of the rate of velocity change of a moving system in its phase space when the adjacent trajectories are close to or far from each other. We here calculate the maximum Lyapunov exponent of the four channels selected by all samples as well as the range and mean of the maximum Lyapunov exponent of each channel in different

Table 2
Ranges and means of corresponding relevance dimensions of EEG channels under different emotional conditions.

Channel	Positive emotions (scope/mean)	Neutral emotions (range/mean)	Negative emotions (range/mean)
4/AF3	1.0962–1.7632/1.4981	1.2983–1.9823/1.6663	0.5242–1.6987/1.2663
6/F7	1.2916–1.6337/1.4515	0.9523–1.8473/1.5898	0.5542–1.6271/1.2885
9/F1	0.9283–1.8627/1.5681	0.8951–1.7193/1.6309	0.6582–1.1983/0.9809
12/F4	0.8175–1.7463/1.3892	1.0738–2.2324/1.7035	0.9567–1.4892/1.3023

Table 3
Range and mean of maximum Lyapunov exponent for EEG channels under different emotional conditions.

Channel	Positive emotions (scope/mean)	Neutral emotions (range/mean)	Negative emotions (range/mean)
4/AF3	0.0105–0.1174/0.0491	0.0633–0.1823/0.0931	0.0084–0.1691/0.0463
6/F7	0.0254–0.137/0.0527	0.0553–0.3133/0.0872	0.0253–0.1721/0.0586
9/F1	0.0098–0.127/0.0511	0.0951–0.2713/0.1039	0.0562–0.1283/0.0619
12/F4	0.0175–0.1433/0.0481	0.0438–0.2923/0.0715	0.0167–0.1482/0.0376

emotional states, as shown in Table 3. Similarly to the CD, the maximum Lyapunov exponent is also an important input in the classification of emotions.

3.4 Mixed feature selection

From each channel in each sample, we extract a total of 5 × 3 = 15 time–frequency domain features, including PSD mean, energy ratio, and energy entropy, which are calculated from the five frequency bands of EEG signals. We extract two chaotic features, namely, the CD and maximum Lyapunov exponent. By fusing time–frequency domain features with chaotic features, each channel contains 17 features. For each sample, there are 17 × 4 = 68 features. To remove the redundant information in these features and simplify the operation process, the principal component analysis method is used to select the fused features to realize feature dimensionality reduction.

3.5 Classification model training and parameter optimization

3.5.1 SVM classification model training

The data needed to classify emotion in this study is linear and inseparable. The least squares SVM based on the Gauss kernel function is utilized as the classification model for emotion recognition. All the parameters that must be determined in the training process include the penalty factor c and parameter α of the Gauss kernel function. These two parameters determine the performance of the SVM classification model. Thus, they are very important for the training of the classification model.

3.5.2 Decision model optimization

As a typical binary classification algorithm, SVM cannot deal with multiclass classification directly. In this work, we intend to identify three categories of emotions. Since this task involves fewer categories, we use a one-to-one method to classify the three categories of emotions and train a binary classifier between any two categories. We then use voting to decide which category the sample belongs to. However, for the three classifiers trained in this study, it is possible to get one vote for each of the three emotions by voting alone, which may result in assigning multiple labels to a single observation.

When there are an equal number of votes, the traditional processing method is to judge the samples as a category with smaller labels. Obviously, this processing method is not practical. To improve it, some researchers proposed to introduce a distance-related fuzzy discriminant function on the basis of the one-to-one SVM algorithm to determine the specific class in accordance with the reliability of voting.[22] Its essence is to calculate the distance between the sample and its separating hyperplane. Which category the sample belongs to depends on the classification result of the classifier nearest to the separating hyperplane. However, this method is sometimes unreliable because it only uses a small number of support vectors as the representative of the class to discriminate the samples without taking into account the overall distribution of the samples. If the distance between the sample and the hyperplane is taken as the basis of classification, sample A may be classified as a positive emotion. In fact, from the distribution of samples, sample A should be classified as a negative emotion.

In fact, samples belonging to the same category are apt to aggregate into a group in the spatial distribution. We propose a decision-making method that combines the distance between the sample and the class center, and the distance between the sample and the separating hyperplane in order to classify samples that may exist in nonclassifiable regions more accurately by means of a one-to-one algorithm.

Firstly, the average feature of all samples belonging to the *i*th category is taken as the class sample center o_i of this category, where o_i is denoted as

$$o_i = \frac{1}{n_i}\sum_{k=1}^{n_i} x_k, \tag{8}$$

where n_i represents the total number of samples in the *i*th category. Then, the reliability discriminant function proposed in this paper is expressed as

$$s_i\left(x_j\right) = \left[1 - \frac{d\left(x_j, o_i\right)}{\max\limits_i d\left(x_j, o_i\right)}\right] \bullet \left[1 - \frac{d_i\left(x_j\right)}{\max\limits_i d_i\left(x_j\right)}\right], \tag{9}$$

where $d(x_j, o_i)$ represents the distance between sample x_i and the center o_i of the *i*th category, and $d_i(x_j)$ represents the distance between the sample x_i and the separating hyperplane that determines the sample as the *i*th category. The decision process of one-to-one SVM based on the reliability discriminant function is shown in Fig. 5.

4. Experiment

To verify the effectiveness of the proposed method, three sets of experiments were conducted: feature selection and dimensionality reduction, decision model optimization, and different feature classification, as described in Sects. 4.2, 4.3, and 4.4, respectively.

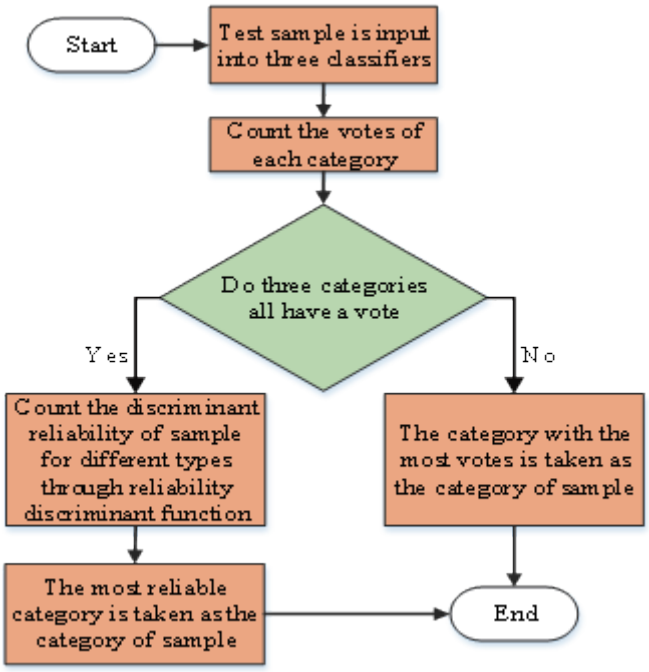

Fig. 5. (Color online) Decision flow of single sample.

4.1 SEED database

We use SEED from Shanghai Jiaotong University. It is an open database resource established by the Brain Computing and Machine Intelligence Center of the Computer Science and Engineering Department of Shanghai Jiaotong University of China in 2015 and can be used by researchers for emotion recognition research. The information summary of SEED is shown in Table 4.

There are 62 electrodes in the EEG signal acquisition device. The distribution of electrodes is shown in Fig. 6. The sampling rate is set to 1000 Hz during the acquisition of EEG signals. An eye-tracking device is used to record the eye electric signal, which is used to eliminate eye movement artifacts during data preprocessing.

SEED contains 15 × 15 × 3 = 675 EEG signal samples. The EEG signals collected in each evoking experiment consist of 62 channels. On the basis of the research conclusions of the database provider, the data of channels 4, 6, 9, and 12 (corresponding electrodes AF3, F7, F1, and F4, respectively) are selected in this study. Therefore, each sample used in this study includes four channels. On the premise of guaranteeing the validity of experimental data, the dimensionality of data is reduced significantly to improve the speed of subsequent data processing.

4.2 Feature selection and dimensionality reduction effect

In accordance with the method of feature selection described in Sect. 3.2, we set four thresholds of 3, 2, 1, and 0% for the dimensions of the feature subset, respectively. With the

Table 4
Information summary of SEED.

Type of information	Information quantification
Number of movie clips	15
Length of movie clips	4 min (emotional highlights)
Emotional labels	Positive, negative, and neutral
Number of subjects	15 (seven males and eight females)
Number of experiments	45 (each subject was subjected to three experiments)
Number of trials per movie clip	2–4 people
Recorded data	EEG data and video recording of facial expressions of subjects

Source: According to the content of SEED in 2015

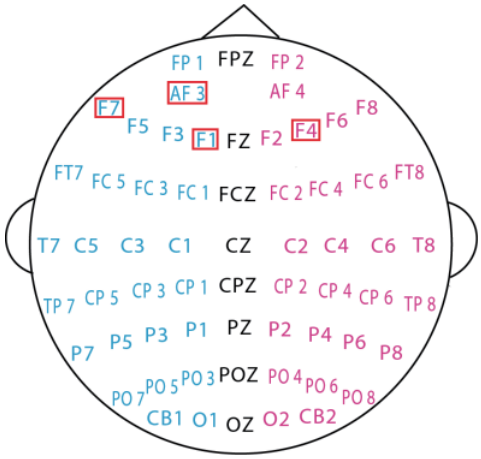

Fig. 6. (Color online) Electrode distribution of EEG signal acquisition device.

above thresholds, four feature subsets with dimensions of 18, 35, 52, and 68 are achieved. Then, we use LS-SVM to investigate the effect of feature selection. Table 5 and Fig. 7 show the accuracies of different dimension feature subsets for emotion recognition.

From Table 5 and Fig. 7, we can see that the overall trend of classification accuracy decreases with the decrease in feature dimension after feature selection and dimensionality reduction by principal component analysis. In particular, when the dimension of the feature subset is reduced from 68 to 18, the average accuracy of emotion classification decreases from 83.99 to 64.68%. However, reducing the dimensionality of features in a certain range will not decrease the accuracy of category discrimination. On the contrary, reducing the feature dimension improves the accuracy of classification for certain categories of emotions. As the dimension of the feature subset is reduced from 68 to 52, although the classification accuracy of positive and neutral emotions slightly decreases, the classification accuracy of negative emotion increases. On the whole, the average classification accuracy is slightly improved. It is shown that some redundant features exist in the 68-feature sets. Appropriate dimensionality reduction contributes not only to simplifying the operation and speeding up the classification, but also to improving the accuracy of classification in some cases.

Table 5
Classification accuracies of feature subsets from different dimensions.

	Feature weight threshold (dimension of feature subset)			
	3% (18)	2% (35)	1% (52)	0% (68)
Positive emotions	69.17%	78.28%	86.06%	**86.33%**
Negative emotions	70.23%	80.45%	**86.47%**	85.23%
Neutral emotion	54.65%	66.76%	80.93%	**81.41%**
Mean value	64.68%	75.16%	**84.49%**	84.32%

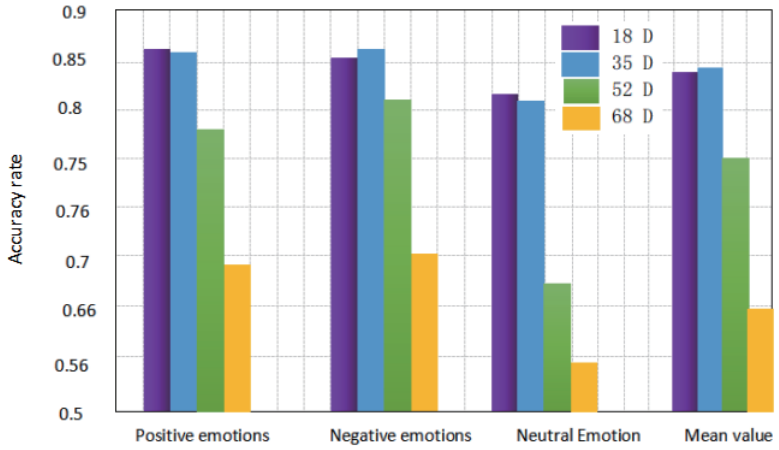

Fig. 7. (Color online) Histogram of classification accuracies of feature subsets of different dimensions.

4.3 Effect of optimizing the decision model

To tackle the three classification problems in this paper, we use the optimal combination of features described in Sect. 3.4 to construct the classification model, and we employ a one-to-one solution to train three classifiers for the three different emotions. In the decision-making stage, three classifiers are used to vote, and the winning category is selected as the category of samples. To test samples existing in the nonclassifiable region, the fuzzy and reliability discriminant functions proposed in this paper are used as the core functions of classifiers. The classification accuracies of two different discriminant functions are shown in Table 6 and Fig. 8.

From the results in Table 6 and Fig. 8, it is concluded that the classification accuracy of samples in the nonclassifiable regions is lower than that of total samples, which indicates that improving the recognition accuracy of samples in nonclassifiable regions is an effective means of improving the overall recognition accuracy. The fuzzy discriminant function only takes the distance between the sample in the nonclassifiable region and the separating hyperplane of each classifier as its criterion of the classification decision. This criterion is problematic. The discriminant function proposed in this paper not only considers the distance between the samples in the nonclassifiable region and the separating hyperplanes of the classifiers, but also takes the distribution of the samples into account. Experiments show that it can effectively improve the classification accuracy of the samples in the nonclassifiable region.

Table 6
Classification accuracies of different discriminant functions.

	Fuzzy discriminant function		Reliability discriminant function	
	Unclassified region	Total	Unclassified region	Total
Positive emotions	83.22%	**86.33%**	83.73%	**88.17%**
Negative emotions	84.31%	85.23%	**85.02%**	87.84%
Neutral emotion	**84.74%**	81.41%	84.61%	85.51%
Mean value	84.09%	84.32%	84.45%	87.17%

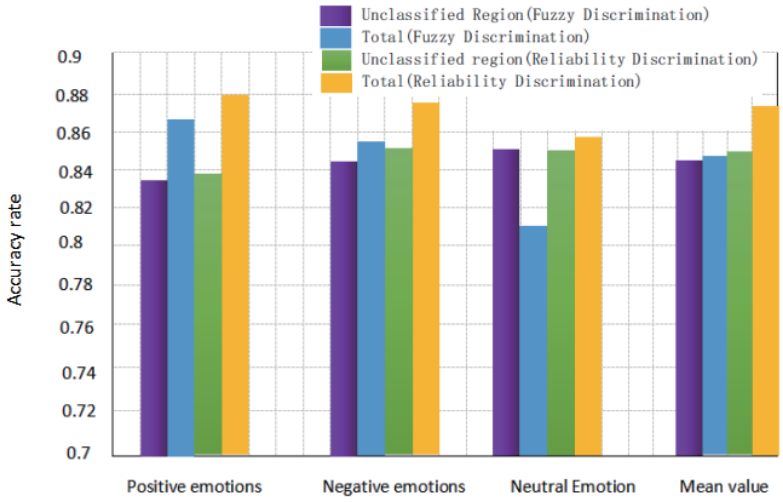

Fig. 8. (Color online) Histogram of classification accuracies of different discriminant functions.

4.4 Effect of different feature classifications

The majority of the emotion classification research based on EEG focuses on extracting and selecting the features of EEG in the time, frequency, and time–frequency domains. By fully considering the randomness, nonstationarity, and nonlinearity of EEG itself, we introduce chaos theory to extract EEG features and obtain chaotic characteristic parameters including the CD and maximum Lyapunov exponent. To compare the classification results of different features, the process described in Sect. 2 is used to train the classification model for different feature sets. In the decision-making stage, the reliability discriminant function proposed in this paper is adopted to classify the samples existing in one-to-one multiclass classification and nonclassifiable regions. The experimental results such as those in Table 7 and Fig. 9 are obtained for comparing and discussing the specific effects of emotion recognition on the following different types of features.

The time–frequency domain features with 60 dimensions include the mean PSD, energy ratio, and energy entropy of each band of EEG signals with 60 dimensions. The chaotic features with eight dimensions include the CD and maximum Lyapunov exponent. The mixed features include time–frequency domain and chaotic features. In this study, after the feature selection phase, we combine the two sets of features into the mixed feature set with 52 dimensions.

Table 7
Classification accuracies of different types of features.

Type name	Time–frequency domain features	Chaotic features	Mixed features
Positive emotions	81.46%	**83.69%**	**88.17%**
Negative emotions	82.73%	**85.83%**	**87.84%**
Neutral emotion	**80.67%**	79.31%	**85.51%**
Mean value	81.62%	**82.94%**	**87.17%**

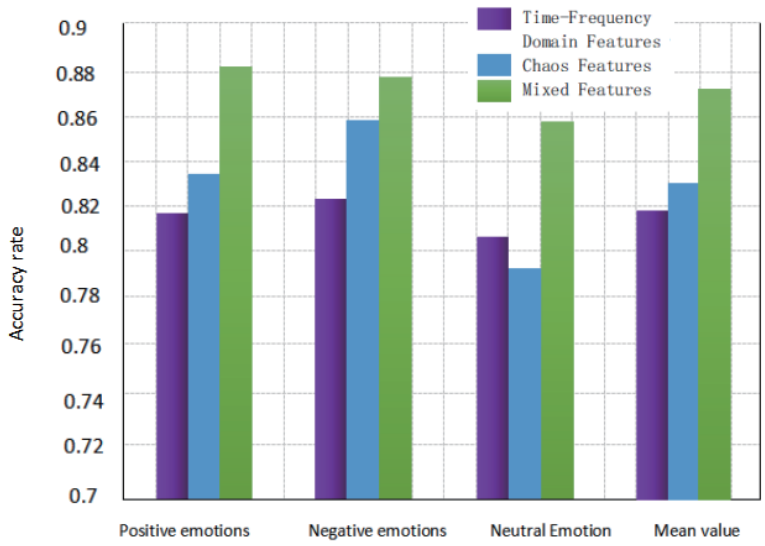

Fig. 9. (Color online) Histogram of classification accuracies of different types of features.

From the results in Table 7 and Fig. 9, it can be seen that the classification accuracy of chaotic features can be higher than that of time–frequency features in terms of the classification of positive and negative emotions. However, when distinguishing neutral emotions, the performance of time–frequency features is slightly better than that of chaotic features. Moreover, the experimental results obtained using the combination of time–frequency domain and chaotic features indicate better classification results than when using only a single type of feature.

According to the results of the experiments, the best recognition of the three types of emotions based on SEED achieved an accuracy of 87.17%, while the highest recognition accuracies of positive and negative emotions were 88.17 and 87.84%, respectively.

The founder of SEED initially extracted the differential entropy of the EEG signal as a feature and trained and validated it using the logistic regression, K-nearest neighbor, SVM, and deep belief network. The average accuracies were 82.70, 72.60, 83.99, and 86.08% in the baseline paper,[23] respectively. Overall, the results proved that the novel feature set extracted and optimized the classification model results and achieved a higher prediction accuracy.

5. Conclusions

We developed a novel set of features from EEG signals for emotion classification. The chaos theory was introduced to calculate chaotic-parameter-like features with the aim of creating a

mixed feature set. We extracted the five key frequency bands in EEG signals by wavelet packet transform and also separated them into several frequency bands instead of simply dividing them into several frequency bands by FT and WT. Finally, we proposed a new reliability discriminant function to solve the problem of fuzzy classification and optimized the decision-making model by considering the distribution of samples and the distance between samples and separating hyperplanes. From the two aspects of data processing and the classification model, we proposed innovative measures. The experimental results proved the superiority of our method. We will actively explore the improvement of deep learning methods and their application to our experimental data in the future.

Acknowledgments

This work was supported by Key Science and Technology Projects of Jiangxi Province (No. GJJ2206003), the Department of Education Industry-University-Research Cooperation Project of Fujian Science and Technology Planning (No. 2022H6012), and the Natural Science Foundation of Fujian Province of China (Nos. 2022J011225, 2021J011169, and 2020J01435). The authors appreciate the valuable comments and suggestions from the editors and reviewers.

References

1 D. Zhang, B. K. Wan, and D. Ming: J. Biomed. Eng. **1** (2015) 44. https://doi.org/10.7507/1001-5515.20150042
2 D. Zhang, L. Q. Wu, and S. S. Li: Proc. IEEE Int. Conf. Multimedia and Expo. (2019). https://doi.org/10.1109/ICME.2019.00130
3 D. Zhang, S. Li, Q. M. Zhu, and G. Zhou: Proc. IEEE Int. Conf. Multimedia and Expo. (2019). https://doi.org/10.1109/ICME.2019.00131
4 K. Schaaff and T. Schultz: Proc. IEEE Conf. Robot and Human Interactive Communication (2009). https://doi.org/10.1109/ROMAN.2009.5326306
5 Y. P. Lin, C. H. Wang, T. P. Jung, T. L. Wu, S. K. Jeng, J. R. Duann, and J. Chen: IEEE Trans. Biomed. Eng. **57** (2010) 1798. https://doi.org/10.1109/TBME.2010.2048568
6 Z. Liu, Y. Shen, V. B. Lakshminarasimhan, P. P. Liang, A. Zadeh, and L. P. Morency: arXiv preprint arXiv:1806.00064 (2018). https://doi.org/10.48550/arXiv.1806.00064
7 D. Bertero, F. B. Siddique, C. S. Wu, Y. Wan, and P. Fung: Proc. SIGDAT Conf. Empirical Methods in Natural Language Processing (2016). https://doi.org/10.18653/v1/d16-1110
8 S. Koelstra, C. Muhl, and M. Soleymanil: IEEE Trans. Affective Comput. **3** (2012) 18. https://doi.org/10.1109/T-AFFC.2011.15
9 A. Zadeh, P. P. Liang, N. Mazumder, S. Poria, and L. P. Morency: Proc. AAAI Conf. Artificial Intelligence (2018). https://doi.org/10.1609/aaai.v32i1.12021
10 D. Hazarika, S. Poria, A. Zadeh, E. Cambria, and R. Zimmermann: Proc. SIGDAT Conf. Empirical Methods in Natural Language Processing (2018). https://doi.org/10.18653/v1/D18-1280
11 G. H. Zhang, M. J. Yu, G. Chen, Y. H. Han, D. Zhang, and G. Z. Zhao: China Sci.: Inf. Sci. **49** (2019) 097. https://en.cnki.com.cn/Article_en/CJFDTotal-PZKX201909002.htm
12 T. F. Bastos, A. Ferreira, A. C. Atencio, S. Atencio, S. Arjunan, and D. Kumarl: Proc. IEEE Conf. Intelligent Human–Computer Interaction (2012). https://doi.org/10.1109/IHCI.2012.6481860
13 R. W. Picard, E. Vyzas, and J. Healey: IEEE Trans. Pattern Anal. Mach. Intell. **23** (2001) 1175. https://doi.org/10.1109/34.954607
14 A. C. Conneau and S. Essid: Proc. IEEE Conf. Acoustics, Speech and Signal Processing (2014). https://doi.org/10.1109/ICASSP.2014.6854493
15 A. C. Conneau and S. Essid: Proc. IEEE Conf. Acoustics, Speech and Signal Processing (2014). https://doi.org/10.1109/ICASSP.2014.6854493
16 Y. Gu, K. N. Yang, S. Y. Fu, S. H. Chen, X. Y. Li, and I. Marsic: Proc. ACL Conf. Computational Linguistic Meet. (2018). https://wwwncbi. 53yu.com/pmc/articles/PMC6217979

17 S. G. Xiao, S. L. Liu, M. M. Song, N. Ang, and H. Zhang: Multibody System Dyn. **48** (2020) 451. https://doi.org/10.1007/s11044-019-09718-9
18 W. L. Zheng and B. L. Lu: IEEE Trans. Auton. Mental Develop. **7** (2015) 162. https://doi.org/10.1109/TAMD.2015.2431497
19 J. Nick and M. Max: Stat. Comput. **45** (2022). https://doi.org/10.48550/arXiv.2003.02367
20 X. M. Guo, W. Y. Zhang, and Z. H. Yuan: J. Instrum. Instrum. **35** (2018) 827. https://doi.org/10.1142/S0219519414500468
21 M. Barczy: arXiv reprint arXiv:2204.00277v2. https://doi.org/10.48550/arXiv.2204.00277
22 H. Ibrahim, S. A. Anwar, and M. I. Ahmad: Proc. IOP Conf. Emerging Electrical Energy, Electronics and Computing Technologies (2021). https://doi.org/10.1088/1742-6596/1878/1/012054
23 W. L. Zheng and B. L. Lu: IEEE Trans. Auton. Mental Develop. **7** (2015) 162. https://doi.org/10.1109/TAMD.2015.2431497