Research on emotion recognition based on ECG signal

Zhongze Zhang¹, Xiaofeng Wang¹, Pengfei Li³*, Xi Chen¹, Liwei Shao²

¹State Key Laboratory of Mechatronics Engineering and Control, Beijing Institute of Technology, Beijing, China 100081
²ZhongShan Research Institute, Beijing Institute of Technology, Zhongshan, Guangdong, China 528400
³Corresponding author’s e-mail: pfli@bit.edu.cn.

Abstract: In this paper, we mainly study the emotion recognition algorithm based on ECG signals, extract the correlation feature and time-frequency domain statistical feature of ECG signals, and introduce SVW, CART and KNN three classification algorithms commonly used in emotion recognition. By comparing the accuracy of emotion recognition in the application of three classification algorithms between the correlation features of ECG signals and traditional time-frequency domain features, we found that the use of correlation features of ECG signals can get a higher recognition rate, which is 16.7%~19.7% higher than that of the traditional feature. In addition, among the three classification algorithms, KNN algorithm can get the highest emotion recognition accuracy. In order to further improve the accuracy of emotion recognition, Max-Min Ant System is combined with KNN classification algorithm in this paper to optimize the feature combination. The overall recognition rate reaches 92%, which is 16.9% higher than the accuracy of emotion recognition directly using KNN classification algorithm.

1. Introduction

Emotion is a complex psychological and physiological process. It affects every aspect of our daily life. The design of a system that can automatically identify emotions can narrow the communication gap between highly emotional humans and computers. It can improve the quality of human-computer interaction, and has great application prospects in medicine, education, entertainment, commerce and other fields[1-5].

At present, emotion recognition methods can be divided into two categories: emotion recognition based on human behavior pattern and emotion recognition based on human physiological signal[6]. The former classifies emotions according to people's different facial expressions, voice and body movements. The latter classifies emotions according to the physiological signal changes caused by different emotions. Although the emotion recognition method based on the behavior pattern is intuitive, people can control their behavior intentionally and hide the real emotion. However, the physiological signal of human body will not change due to subjective consciousness. So the emotion recognition method based on the physiological signal of human body can accurately obtain the expression of emotion.

The nerve endings of the autonomic nervous system (ANS) in the heart play a major regulatory role in the heart's activity because they affect the rhythm in which cardiac myocytes pump blood. The autonomic nervous system is divided into sympathetic nervous system and parasympathetic nervous system, which are antagonistic to each other. The nerve fibers of the cardiac sympathetic nervous
system are distributed along the atria and ventricles. When they are activated, they can stimulate the heart muscle cells to increase the heart rate. On the other hand, when the parasympathetic nervous system is excited, it reduces the load on the heart. In the presence of external mental stimulation, the sympathetic nervous system will dominate the regulation of cardiomyocytes\cite{7}. Therefore, the ECG signal is a kind of physiological signal closely related to emotion, which is widely used in emotion recognition system.

In the current research on emotion recognition based on ECG signals, researchers mainly classify them based on the characteristics of time domain and frequency domain of ECG signals. But the correlation of ECG signals is rarely used. The ECG signals affected by emotion are non-linear and non-stationary signals. The correlation is an important feature of non-stationary signals, which can reflect the correlation between the present behavior and the past behavior of time series. Therefore, the correlation characteristics of ECG signal can reflect its dynamic changes under the different emotions. In this paper, the time-domain and frequency-domain statistical features and correlation features of ECG signals are applied to emotion recognition, and a comparative study is carried out.

2. Methods

2.1. Data

In this paper, a non-contact emotional ECG signal acquisition system is used. The system structure is shown in Figure 1. The system uses a flexible active shielded electrode for non-contact detection of ECG signals, and transmits the signals to the ECG signal processing module through the shielded wire. After signal processing and A/D conversion, the data is transmitted to the PC terminal via bluetooth module.

![Figure 1. Non-contact ECG signal acquisition system: (a) Structure diagram; (b) Assembly drawing; (c) Picture of real products](image)

The acquisition system has a small size and low power consumption. It only needs two AAA dry batteries to power it. It collects data in a contactless manner and transmits data through Bluetooth.
Therefore, the acquisition system is convenient to be used in different application environments. It can allow volunteers to collect ECG signals without sensing the acquisition system, which is conducive to the expression of subjects’ emotions.

In this paper, we choose the method of video induced emotion. In order to induce the accurate emotion of the subjects, before the formal signal collection, we first investigate the subjects’ preferences for the film and television works, such as the most appreciated actors and favorite movies. Then we select the corresponding video clips as the emotional materials according to the feedback information of each subject. In order to let the subjects have enough time to immerse themselves in the video material and induce stable emotions, the induction time of each emotion is guaranteed to be more than 12 minutes, and the time of video material varies from 12 to 20 minutes.

In order to avoid the influence of empathy between single group members on emotion recognition, the recruitment scope of experimental volunteers is not limited to school students, but also includes different occupations such as courier, white-collar workers and cleaners. In this study, a total of 20 volunteers were recruited to participate in ECG signal acquisition. None of them have a history of mental illness or heart disease.

In order to explore the recognition effect of different emotional dimensions, one typical emotion is selected from each of the four quadrants of the emotional dimension model. They are happy emotion, angry emotion, pleasant emotion and sad emotion. They correspond to high valence with high arousal, low valence with high arousal, high valence with low arousal, and low valence with low arousal.

In order to balance the number of samples corresponding to different emotions, a 10-minute emotional stable signal was extracted from each emotional ECG signal and divided into 5 samples, each sample for 2 minutes. Finally, 400 ECG samples of the same length were obtained, including 100 samples for each emotional tag.

2.2. Feature extraction
In this paper, two feature sets S1 and S2 are established.

| Characteristic parameter expression | Characteristic parameter name          |
|-------------------------------------|--------------------------------------|
| Max                                 | Maximum value of ECG signal          |
| Min                                 | Minimum value of ECG signal          |
| Mean                                | Mean value of ECG signal             |
| Std                                 | Standard deviation of ECG signal     |
| RRmean                              | Mean value of R-R interval           |
| RRstd                               | Standard deviation of R-R interval   |
| Energy                              | Signal band energy                   |
| Ratio                               | Proportion of signal energy          |

The time-frequency domain feature set of ECG extracted in this paper is denoted as S1 feature set, which contains the statistical characteristics of ECG signal in time domain and frequency domain. The specific characteristic parameters are shown in Table 1.

| Characteristic parameter expression | Characteristic parameter name          |
|-------------------------------------|--------------------------------------|
| α                                   | DFA scale index                      |
| α1                                 | DFA short-range scale index          |
| α2                                 | DFA long-range scale index           |
The correlation feature set of ECG signal extracted in this paper is denoted as S2 feature set, including autocorrelation feature parameter, multifractal feature parameter and cross correlation feature parameter. The specific characteristic parameters are shown in Table 2.

### 2.3. Feature selection

In this paper, we use the Max-Min Ant System (MMAS) to optimize the feature set selection. Ant Colony Optimization (ACO) is derived from ants' foraging behavior. ACO in the process of optimization is parallel, each individual can communicate with each other to achieve information transmission. So it can improve the ability of global optimization, and the algorithm has a positive feedback mechanism, which can improve the convergence speed of the algorithm. It is effective in solving the problem of multi domain combination. Moreover, the outstanding feature of the ACO is that it has strong robustness, no need for human adjustment during operation, and good scalability. However, the ACO has the disadvantages of too long time-consuming and premature phenomenon in the process of optimization. In order to solve the above problems of ACO, Dorigo et al. proposed Ant Colony System (ACS) in 1996[^7]. Compared with ACO, ACS has the following improvements: in order to speed up the convergence speed, it uses the pseudo-random proportion rule; it only updates pheromones on the best path so far; in order to increase the diversity of races and explore the possibility of other paths, the pheromone of the path is required to be updated after the ants reach the destination.

Max-Min Ant System (MMAS) was first proposed by German scholars T. Stuetzle and H. Hoos[^8]. MMAS is based on ACS and has been optimized in detail. The main optimization is reflected in three aspects. First, the concentration of pheromone is limited within $[\tau_{\text{min}}, \tau_{\text{max}}]$. When the concentration is more than $\tau_{\text{max}}$, $\tau_{\text{max}}$ is selected; when the concentration is less than $\tau_{\text{min}}$, $\tau_{\text{min}}$ is selected. In this way, the concentration of pheromone on all paths is not too large or too small. It effectively increases the probability of ants finding the optimal path and avoids the algorithm from stopping prematurely, which is conducive to finding the global optimal solution. Second, when all the ants have completed a roaming, compare the current solutions obtained by all the ants, take the shortest path between the nest and the food as the current optimal path, and update the concentration of pheromone retained on this path. In this way, the historical information during the roaming process can be retained. Third, it improves the setting of parameters at the beginning of the algorithm. For the parameter $\rho$, if it is set to a smaller value, the volatilization of pheromone will be weakened, and the ant can expand its search range, thus improving the optimization ability.

### 2.4. Classifier

#### 2.4.1. KNN

K-Nearest Neighbour (KNN) is a mature pattern recognition algorithm. It is also one of the most simple and effective machine learning classification algorithms[^9]. KNN is a kind of lazy learning algorithm, which is simple in training and fast in calculation. And it does not need to use the training process of the training set. After the new samples are put in, the classification calculation is started directly. The main calculation process of this algorithm is to input a new sample into the data set.
According to the feature space of the sample set and the feature space of the new sample, the classifier will find out the k original sample subsets which are closest to the feature space of the new input sample from the existing n sample sets. Then judge the category of each sample in the k original sample subsets, and the sample category with the largest proportion is the category to which the newly input sample belongs. The reason why KNN is fast is that it can find the distance between samples directly from the existing samples, without the need for other model construction. KNN mainly uses Euclidean distance, cosine distance or similarity measure to find k initial samples which are closest to the new samples, and then predict the classification of new samples according to the classification of these initial samples.

2.4.2. SVM.
Support Vector Machine (SVM) was first proposed by Chervonenkis and Vapnik in 1995[10]. SVM maps the difficult-to-divide low-dimensional space vector set to the high-dimensional space. It has many unique advantages in dealing with the pattern recognition problem of nonlinear features or high-dimensional features. Now it has been widely used in emotion recognition, handwriting recognition and face recognition. SVM breaks through the limitation of spatial dimension and creates a hyperplane used as decision surface. The function of this hyperplane is to isolate the samples of different classifications and maximize the boundary.

2.4.3. Decision Tree.
One of the two most widely used classification models is Decision Tree, which is also a very famous prediction model in the field of machine learning[11]. Decision Tree is intuitively tree-like structures. The structure is like a tree in nature, with many branches and connections. In a Decision Tree organization, each branch's connection represents an attribute test, and a branch is an outward-extending node. Decision Tree is a very intuitive method. It uses probability analysis to illustrate the mapping relationship between object values and their attributes on the basis of all possible probabilities that have been known. Each branch is a possibility of predicting results, which are then transferred to different nodes, that is, to different categories. When the sample we are calculating completely conforms to the probability path of a certain node and branch in the Decision Tree, then the node classification that the sample finally arrives is its classification. Decision Tree based on information theory includes ID3 and Classification and Regression Trees (CART). CART is derived from ID3 and has two improvements over ID3: first, it uses information gain rate to select features without selecting more subcategories; second, it can handle continuous features.

3. Results and analysis

3.1. Effect of classifier types and feature types
In order to study the emotion recognition effect of different classification algorithms, three classification algorithms (KNN, SVW and CART) are applied to emotion recognition. In addition, in order to verify the feasibility of the application of ECG signal correlation features in emotion recognition, this paper also compares the ECG signal correlation features with traditional time-frequency domain features in emotional recognition. Among all the data, 75% of the data are used as training samples, and the remaining 25% of the test data are used for sentiment classification experiments.

The accuracy of three different classification algorithms for emotional recognition is shown in Figure 2. Each classification algorithm uses a feature set in three ways: using only the S1 feature set, using only the S2 feature set, and using both the S1 and S2 feature sets. As can be seen from Figure 2, compared with different classification algorithms, KNN algorithm has the highest emotion recognition accuracy, which is 5.9%~8.5% higher than SVW and 2%~6.5% higher than CART, but the recognition accuracy does not exceed 90%. When only using S1 feature set, the recognition accuracy is only 68.9%. By comparing the emotion recognition accuracy of different feature set combinations, it
can be seen that the accuracy rate of using only S2 feature set is the highest, which is 16.7%~19.7% higher than that of using only S1 feature set, and 9%~11.3% higher than using both S1 and S2 feature sets. It can be concluded that the correlation features of ECG signal applied to emotion recognition have better performance than the traditional time domain and frequency domain features, and the recognition rate can be increased by 20%. In addition, when S1 and S2 feature sets are used simultaneously, the number of features is the most. However, the accuracy of emotion recognition is not as good as only using S2 feature set, which shows that there are some feature parameters with low correlation with emotion in the traditional time-frequency domain features. Therefore, the accuracy of emotion recognition can be further improved by optimizing all feature parameters and selecting the best feature set.

![Figure 2. The emotion recognition accuracy of different classification algorithms](image)

3.2. Effect of feature selection

The combination of MMAS and KNN classification algorithm can solve the problem of feature combination optimization in emotion recognition. In this way, we can not only reduce the dimension of the original feature set in each emotion state, but also improve the recognition accuracy of each emotion. In this paper, firstly, MMAS algorithm combined with KNN classification algorithm is used to select the best combination of eigenvalues for emotion recognition in each emotion, and then the best combination of eigenvalues is taken as the input feature of KNN algorithm. Among all the data, 75% of the data are used as training samples, and the remaining 25% of the test data are used for sentiment classification experiments.

The selection of k value in KNN algorithm will directly affect the results of the algorithm. If the value of k is small, it means that only the training samples which are close to the input test samples will have an effect on the prediction results, and overfitting is prone to occur. If the k value is large, the advantage is that the estimation error of learning can be reduced, while the disadvantage is that the approximate error of learning will increase. In this case, the training samples far away from the input test samples will also play a certain role in the prediction, making the prediction wrong. In the specific application of this paper, k value is selected as a smaller value, and the optimal k value is selected by cross validation method. The final results of emotion recognition in the MMAS-KNN model are shown in Table 3.

| Real category | Predicted category |
|---------------|-------------------|
|               | Happy | Sad  | Pleasant | Angry |
| Happy         | 91%   | 1%   | 4%       | 4%    |
| Sad           | 0     | 92%  | 0        | 8%    |
As can be seen from Table 3, the recognition rates of the MMAS-KNN model used in this paper for the four emotions of happy emotion, sad emotion, pleasant emotion and angry emotion are 91%, 90%, 88% and 97% respectively, and the overall recognition rate reaches 92%, which is 16.9% higher than that of the emotion recognition accuracy directly using KNN classification algorithm. Compared with the current research on emotion recognition based on physiological signals, the recognition rate of single channel emotion recognition using ECG signal is comparable to that of multimodal fusion emotion recognition. And this paper uses non-contact way to collect ECG signal, which is more simple and comfortable, and is suitable for more practical application scenarios.

4. Conclusion

In this paper, the algorithm of emotion recognition based on ECG signal is studied. Firstly, the correlation features and time-frequency domain statistical features of ECG are extracted, and three classification algorithms, SVW, CART and KNN, are introduced. In order to verify the feasibility of ECG signal correlation features in emotion recognition, we compare the accuracy of correlation features and traditional time-frequency domain features in the application of three classification algorithms. We find that using the correlation features of ECG signals can get a higher recognition rate, which is 16.7%~19.7% higher than the traditional features. And among the three classification algorithms, KNN algorithm has the highest accuracy rate of emotion recognition.

In order to further improve the accuracy of emotion recognition, this paper combines the MMAS algorithm with KNN classification algorithm to optimize the feature combination. Finally, the recognition rate of happy emotion, sad emotion, pleasant emotion and angry emotion are 91%, 90%, 88% and 97%, respectively. The overall recognition rate is 92%, which is 16.9% higher than that of KNN classification algorithm directly.

References

[1] Rencheng, Zheng, Shigeyuki, et al. (2015) Biosignal Analysis to Assess Mental Stress in Automatic Driving of Trucks: Palmar Perspiration and Masseter Electromyography. Sensors, 15(3), 5136-5150
[2] S. D’Mello, S. Craig, B. Gholson, S. Franklin, R. Picard, et al. (2015) Integrating Affect Sensors in an Intelligent Tutoring System. Intelligent User Interfaces, 7-13.
[3] Krithika L. B. and Lakshmi Priya G. G. (2016) Student Emotion Recognition System (SERS) for e-learning Improvement Based on Learner Concentration Metric. Procedia Computer Science, 85, 767-776.
[4] Mohammad Soleymani, et al. (2018) Affective Characterization of Movie Scenes Based on Multimedia Content Analysis and User's Physiological Emotional Responses. IEEE Transactions on Affective Computing, 3(1), 102-115.
[5] W.M. Wang, J.W. Wang and Z. Li. (2019) Multiple affective attribute classification of online customer product reviews A heuristic deep learning method for supporting Kansei engineering. Engineering Applications of Artificial Intelligence, 85, 33-45.
[6] Calvo R. A. and D'Mello S. K. (2010) Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications. IEEE Transactions on Affective Computing, 1(1), 18-37.
[7] Dorigo Marco, et al. (1996) Ant Colony System : a Cooperative Learning Approach to the Traveling Salesman. IEEE Transaction on Evolutionary Computation, 1(1), 53-66.
[8] T. Stutzle and H.H. Hoos. (2000) MAX-MIN Ant system. Future Generation Computer Systems, 16(8), 889-914.
[9] Duda R O, Hart P E, Stork D G, [Pattern Classification], Pattern classification, Wiley, 2004.
[10] Christopher J.C. Burges. (1998) A Tutorial on Support Vector Machines for Pattern Recognition. Data Mining and Knowledge Discovery, Papers 2(2), 121-167.
[11] Ross J Quinlan. (1986) Induction of Decision Trees. Machine Learning, 1(1), 81-106.