Research on Feature Extraction of Human ECG Identification

LIN Jia-Lun[1], LU Pan[2], WANG Wei-Quan[1,*]

1.Department of Network management center, Hainan Medical University, Hainan, 571199, China

2.School of Computer Science and Engineering, South China University of Technology, Guangzhou 510006, China

[*]email: 285995636@qq.com

Keywords: ECG; Human Identification; feature extraction; stable feature

Abstract: ECG, a kind of identity information in human identification, can be easily affected by detector's movement and the change of mood. Traditional ECG feature extraction methods require a lot of calculation which increases the computation time in the process of identification. In this paper, we mainly discuss how to seize the main, easy to extract and stable ECG feature, to complete the identification successfully and quickly. And we propose a method of feature extraction that as for intervals, we only choose QRS interval which always maintain the same. Experiments show that the identification speed has been improved when the accuracy of recognition has been little affected.

1 Introduction

Human identification based on ECG is an emerging area in biometric recognition. Electrocardiogram is the biological potential produced by human heart beat and the waveform change exist a strong regularity. What's more, it varies from people to people and different person's ECG waveforms are different. ECG signal meets the universality, stability, living and individual differences of biometric identification. Therefore, it has the feasibility to be used as a biometric identification feature.
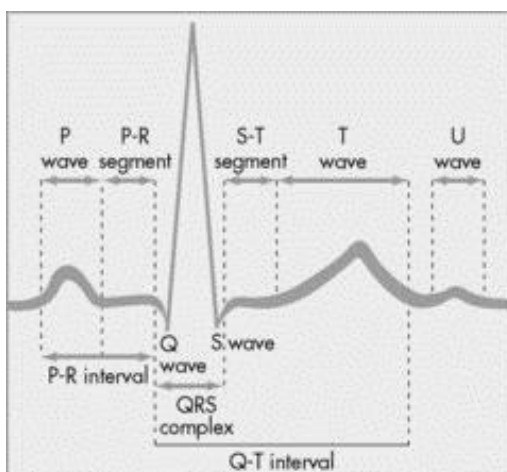

Figure.1. ECG waveform

ECG signal is composed by P wave, P-R segment, P-R interval; QRS complex, ST segment, T wave ,Q-T interval and U wave, as shown in Figure 1.There are differences on age, sex, health status, heart function between different individuals and the characteristics of ECG will be different correspondingly. Correctly extract these feature points can provide accurate basis for the final ECG identification classifier. Therefore, after accurate position each waveform, to seize the substantive

characteristics of the signal, we need extract and select the signal features.

The key research aim of ECG identification is feature selection and feature extraction which can influence the accuracy and speed of identification directly. So far, researchers at home and abroad have done a lot of research work on feature selection and achieved some stage research achievements.

For instance, Literature 2 extracts a total of 30 characteristics (interval, slope and amplitude) from 12 lead ECG signals of 20 healthy people and uses principal component analysis (PCA) to remove redundant features, selecting 10 characteristics which contribute more to classify, and at last the accurate rate reaches 100%. Israel selects 15 interval features of 29 individuals such as the width of P wave, PQ interval and so on, and using LDA achieved 81% heart beat recognition accuracy rate and 100% identification accuracy rate. Literature 4 uses the differential threshold method to extract feature points of P wave, QRS wave and T wave of 10 individuals, selects pairwise combinations of 8 amplitude and interval features and achieves 60% and 100% identification accuracy rate. Literature 5 uses feature vector composed by 15 width and interval features and 6 amplitudes features and finally achieves 100% identification accuracy rate among 29 individuals. Literature 6 extracts 13 interval features, 4 amplitude features and 3 angle features and the highest accuracy rate reaches 99% among 25 individuals. Literature 7 selected 19 feature vectors and analyzed by kernel principal component analysis. The first 10 components was used for classification and reached the highest identification rate which was 98.7342%.Literature 8 extracts 10 features including slope, harmonic number, trough number, amplitude difference, the phrase peaks and troughs corresponding frequency value, energy ratio and maximum energy of frequency, and analyze by PCA. Finally, 8 features are selected as the features of identification.

Although a lot of ECG identification researches view feature selection and feature recognition as the key point, the number and content of selected features are very different and there is no consensus on it. In this paper, in the process of the extraction of feature points, we extract the main, easy to extract and stable features under circumstance that there is little impact on the accuracy rate of identification, and then use the corresponding analysis algorithm to acquire calculation result quickly.

2   Proposed Method

2.1 Extraction of feature points

Traditionally, signal feature extraction mainly extract features such as QRS wave, P wave, T wave, P-P interval, R-R interval, Q-T interval, T-P interval and so on. The calculation is very large in the process of feature extraction and identification. Therefore, to seize the main, easy to extract and stable features is the key to complete the identification successfully and quickly.

LiYanju tests several healthy male's ECG under four status of sleep, peace, mental and exercise load. The result shows that with the increase of exercise load, T-P interval changes the most which significantly reduced while P-R interval remain unchanged. SunSuzhen selects participants with stable ECG signal from 30 rowing athletes and tests signal data before exercise, exercise, exhausted and after recovery. The result shows that P-R interval changes slightly while QRS interval is not significant changed and Q-T interval is shorted slightly. Due to the interval features can be easily affected by the change of heart rate and the cardiac cycle will also change after exercise or emotional stress. Therefore, too many interval features that are used in identification will influence the recognition rate.

According to the identification and analysis of 30000 P waves, QRS waves and T waves from

12 Lead ECG of 2500 individuals, Li Shifeng concludes that P wave and T wave are both single wave shape and low waveform. Some even too flat to see clearly and they are easy to be influenced by many factors. Therefore, the P wave and T wave have large degree of variation and not suitable for identification .However, QRS integrated waves are mainly simple wave. In 30000 waves, the number of simple waves is 29347, accounting for 97.8%, while only 653 waves are complex wave, accounting for 2.2%. What's more, QRS waveform is higher than P wave and T wave. In summary, QRS wave has many advantages when used in identification such as few influence factors, small waveform change, and small variation, easy to identify .So it's very suitable for fast identification on computer. To select features not affected by heart rate and avoid the adverse effect to identification caused by heart rate change, this paper first give up the P-P interval, Q-T interval and T-P interval characteristics which are easily disturbed by heartbeat frequency, and we only choose QRS interval which always maintain the same. Specifically, the features mainly include the amplitude of R wave, amplitude of Q wave, amplitude of S wave and QRS interval.

2.2 Method of feature extraction

It is only possible to calculate heart rate, heart rate variability and analyze other details of ECG after determining QRS wave group. So the accurate extraction of the characteristic information of QRS wave group is the basis of the analysis of ECG. As the calculation and analysis of ECG and the detection of P wave and T wave are based on the localization of QRS wave group (especially R wave), so the extraction of ECG begins at QRS wave group.

In this paper, we first use the method of wavelet transform and decompose the signal by multi-scale wavelet transform, and then extract the feature point of QRS wave. Finally, extract the feature point of P wave and T wave based on the peak of R wave.

The amplitude of R wave has the greatest value in ECG signal and the peak of R wave is the maximum. If we connect it with the adjacent maxima, there will be a large slope and the feature is the most obvious. The detection and location of R wave is the primary problem in detection. Firstly, find out all maxima; then, calculate the slope between each maximum value and the maxima of both sides; select threshold to determine R peak; finally, adjust R peak based on the range of heart rate. Normal adult human heartbeat is about 75 beats per minute. The experiment has proved that 800ms data length includes P wave and T wave and can keep most of the information of heartbeat waveform. Therefore, after detecting R peak, take R as the center point and intercept ECG data with 800 time length and then we get a set of complete ECG. Moreover, a time window is so narrow that baseline drift does not cause a large disturbance to the waveform during this period. During this short time interval, there will be only one peak at most. The maximum value during this time interval is the value of R peak. And then find the minimum point at both sides of R wave in a certain range as Q wave and S wave. Using the same method to locate the second wave peak and we can get a cardiac cycle time which is equal to the interval between two peaks. Using cardiac cycle time as time window and using the previous amplitude as the threshold to find the next maximum point.

3 Experimental Results

In this paper, we take 10 samples from MIT-BIH as experiment data and select QRS feature vector in classification. Experiment results show that, by using this method, the workload of data calculation has been reduced greatly and the speed of identification has been improved obviously, and the accuracy rate can reach 90%.

4 Conclusion

As a kind of common people's identification method rather than the detection of heart disease, this feature extraction method can help to get the result successfully and quickly.

Acknowledgement

References

[1]Li Zhong-Jian, etc. Living individual ECG waveform and uniqueness [J]. Journal of Medical Forum,2008,29(4):45-46

[2]Biel L,Pettersson O,Philipson L,et al.ECG Analysis: A New Approach in Human Identification[J]. IEEE Transactions on Instrumentation and M easurement,2001,50(3): 808-812．

[3]Israel S A,Irvine J M,Cheng A,et al.ECG to Identify Individuals［J］. Pattern Recognition,2005,38(1):133-142．

[4]Wang Li，.Based on the identity of the ECG signal recognition technology research［D］. Jinan: Shandong University，2005.

[5]Wang Y，Plataniotis K N，Hatzinakos D． Integrating Analytic and Appearance Attributes for Human Identification from ECG Signals[C]／／Proceedings of 2006 Biometrics Symposium：Special Session on Research at the Biometric Consortium Conference．Washington D．C.，USA: IEEE Press，2006: 1-6．

[6]Singh Y N，Gupta P． ECG to Individual Identifica-tion［C］／／Proceedings of BTAS'08．Washington D．C.，USA: IEEE Press，2008: 1-8.

[7]Chen Xiaodan;Xu Huifang;Shen Haibin. ECG Identification Based on Fusion Features of Morphological Characteristics and KPCA [J]. Electronic Technology，2015(3):5-8.

[8]Wang Yarong. Study of Human Identification by Electrocardiography Frequency Features [D]. Tianjin：Tianjin University of Technology，2014.

[9]LI Yan-jun,YAN Hong,ZHAO Wei,YANG etc. Study on Regulation of ECG Cycle Fluctuation [J]. Space Medicine & Medical Engineering, 2009, 22(2):126-129.
WANG Zheng,YAN Qu.SpaceMedicine＆MedicalEngineering,2009,22(2):126～ 129

[10]Kong Shuzhen ,etc.The influence of different exercise intensity on the rowing athletes ECG [J]. Shaanxi Medical Journal，2011，40(10)：1340-1342.

[11]Li Shi-Feng. Electrocardiogram QRS wave in the computer identification of meaning [J]. Journal of Medical Forum, 2009 ,30(2):52-54.

[12]Wang Y, Plataniotis K N, D. Hatzinakos. Integrating analytic and appearance attributes for human identification from ECG signal [C] // Proceedings of Biometrics Symposiums (BSYM'06), Baltimore, USA, 2006.