

## A Fusion of ECG Signal Identification Method

Huan Zhang, Lan-xiang Chong and Dan Guo

(Northwestern University, xi 'an, Shanxi)

865567067@qq.com, . 992078108@qq.com, 964896716@qq.com

**Keywords:** Ecg signal; Feature extraction; Template library; Pattern matching; Identification

**Abstract:** In order to accurately obtain the identification information, in the basis of various biometric identification technology comparative study , using a higher safety and reliability of ecg signal identification technology. This paper studies the method of ecg signal identification technology, first using wavelet transform to remove noise and interference in ecg signal, pinpoint the signal of a cycle, and then contrast the recognition rate of two feature extraction methods of harmonic fundamental wave ratio and single cycle inner time-frequency joint analysis, finally proposed the fusion feature extraction method of the ratio of harmonic fundamental wave and single cycle inner time-frequency joint analysis. Through small amount of data in the MIT - BIH ecg database experiments have been carried out to verify this method, the results show that the algorithm has a high recognition rate.

### Introduction

Along with the rapid development of modern society and the economy, various to fake and forged identity as a means of fraud and theft crime is increasing, has become an important type in the field of information security. Therefore, identity recognition technology emerge as the times require and widely used. Based biometric identification methods include fingerprint recognition, face recognition and voice recognition etc.. However, these methods are not foolproof. Such as fingerprints easily left on the grasped object, which will be used by others; face can also be extracted from pictures even with silica gel in imitation. In order to make the identity recognition technology is more safe and reliable, people on the one hand, use of multiple biometric parameters fusion technology, increase the difficulty of the parameter to be generic; on the other hand, looking for more reliable biometric identification technology, such as ecg identification technology<sup>[1,2]</sup>. Because ECG signal is bioelectricity signal acquisition in vivo produced, more difficult to counterfeit and forgery than traditional biometric signals<sup>[2-5]</sup>. Therefore, ECG identification technology has been proposed and attracted wide interest of many researchers.

In 2001, Biel proved that can realize the identification using the ECG signal, but on the further research of identity recognition based on ECG signal, have been many recognition algorithms is in recent years. Literature [6] through the principal component analysis method and the linear discrimination analysis method to realize feature extraction and dimensionality reduction, the 13 individuals to achieve the accurate recognition rate of 100%, but the sample size is too little. The literature of [7] on single lead ECG signals by the recognition and classification of distance method based on wavelet transform, the recognition rate is 89%, the recognition rate is not very high. Literature [8] using a fusion of wavelet transform and support vector machine (SVM) method for ECG identity, but as a result of this paper is based on feature points extraction, lead to the identification accuracy is influenced by feature point extraction accuracy.

In this paper, according to existing recognition algorithms with small data quantity, low recognition rate, multi feature point positioning is not accurate and a single biometric recognition accuracy is not high enough, the integration method of combined analysis of feature extraction is proposed by adopting harmonic Bobbi electrocardiographic signal feature extraction and single cycle heart electric signal time frequency, the 36 groups of MIT-BIH ECG database data experiment, the

results show that the recognition algorithm of this algorithm is better than previous ECG have higher speed and higher recognition rate.

### ECG Signal Preprocessing

Although the ECG signal acquisition technology has a great development, but in the process of collection is still affected by noise and external uncertain factors. Therefore, it is necessary before the processing and analysis of ECG signal denoising pretreatment on the.

**Ecg Signal Wavelet Pretreatment.** Db8 wavelet generating function waveform and typical ecg signal waveform is similar, and validation experiments have *db8* has better performance on ecg signal de-noising. Therefore, after the ECG by db8 mother wavelet decomposition, the coefficient signal is greater than the noise factor, he was able to find a suitable number  $THR_i$  as a threshold, when  $cd_i$  less than the threshold, then the  $cd_i$  think is mainly caused by noise and zero; when  $cd_i$  is greater than the threshold value, it is considered at this time  $cd_i$  is mainly generated by a signal, the coefficient retain  $cd_i$  this part (hard threshold method) or above the threshold value minus a threshold value (soft threshold method), and then you get denoised ECG. Given the hard threshold function at some point will have stopped, resulting in the reconstructed signal waveform discontinuity at some point, we use soft threshold de-noising function, which can effectively eliminate the signal waveform appears intermittent problems, so smoother reconstructed signal<sup>[9]</sup>. Ecg signal de-noising processing before and after contrast figure (as shown in Fig.1).

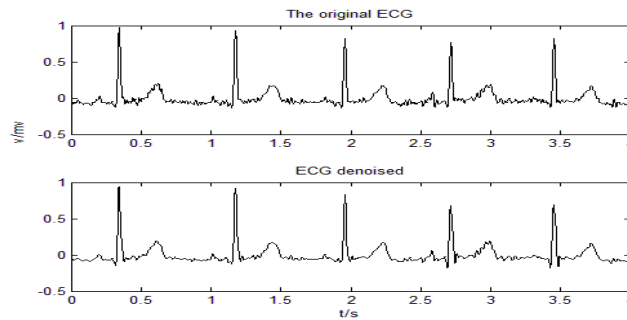


Fig. 1 ecg signal preprocessing before and after contrast

**ECG Cycle Positioning.** ECG physiologists study found that heart constantly engaged in rhythmic contraction and relaxation, cardiac rhythm determines the quasi-periodic ECG. A cycle of ECG is the electrocardiogram QRS<sup>[9]</sup> wave interval between R wave peak and the next time R wave peak.

ECG signal in the MIT-BIH database, the sampling frequency  $fs$  has 500S / s and 360S / s two kinds. The study found that each of the periodic signal is not very good start, it may be the initial signal acquisition is gatherers have not yet entered a relatively relaxed state, excluding the 400 sample data before the experiment. According to the maximum value of ECG R wave the size of  $|x|$  absolute, set a is less than 1 of the adjustment coefficient  $\alpha$ , and define the threshold:  $A0=|x| \times \alpha$ .  $\alpha$  values to ensure that the  $A0$  peak is higher than the P-wave and T-wave, so  $\alpha$  is generally from 0.6 to 0.7, the effect is shown in Fig.2. Search backward along the time axis, higher than the threshold value  $A0$  detected extreme points  $t1, t2, t3$ , etc., as shown in Fig.3. Adjacent poles point value that is the time interval for the cycle, in view of ECGQRS wave duration of not more than 0.5s, which can carry on the examination to determine the cycle. Such as the calculation of the period is less than 0.5s, it may be considered that the peak of extreme point detected than R wave, restart or revision of  $\alpha$  value.

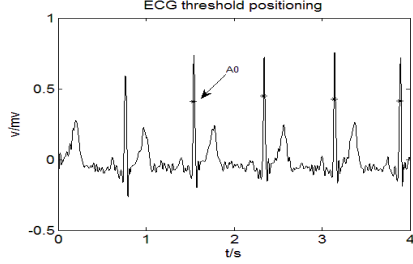


Fig. 2 Threshold of A0 in ECG waveform position

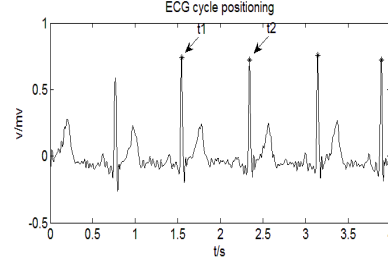


Fig. 3 ECG cycle positioning

## Spectral Feature Extraction and Experimental Analysis

**ECG Spectral Characteristics.** Known duty ratio of 1 and a period of  $T$ , an amplitude  $A$  square wave signal  $F(t)$  is the Fourier expansion as follows:

$$F(t) = \frac{4A}{\pi} \left( \sin \frac{\pi}{T} t + \frac{1}{3} \sin \frac{3\pi}{T} t + \frac{1}{5} \sin \frac{5\pi}{T} t + \dots \right) \quad (1)$$

Its ratio of 1,3,5 ... the harmonics and fundamental wave are  $1, 1/3, 1/5 \dots$ , independent of  $F(t)$  of the amplitude and frequency, i.e.  $F(T)$  of the harmonic Bobbi with amplitude frequency invariance. The ECG signals as periodic signal, should also have similar properties. ECG signal discrete Fourier series expansion<sup>[10]</sup>, as shown in Eq.2:

$$x_i = \frac{a_0}{2} + \sum_{k=1}^m \left( a_k \cos \frac{2\pi k i}{N} + b_k \sin \frac{2\pi k i}{N} \right) \quad (2)$$

Where  $N$  is a cycle of ECG signal samples. The expansion coefficients:

$$a_0 = \frac{2}{N} \sum_{i=0}^{N-1} x_i, \quad a_k = \frac{2}{N} \sum_{i=0}^{N-1} x_i \cos \frac{2\pi k i}{N} \quad (3) \quad b_k = \frac{2}{N} \sum_{i=0}^{N-1} x_i \sin \frac{2\pi k i}{N}, \quad k=1,2,\dots,m \quad (4)$$

Type  $m$  maximum  $N/2$ . Using the formula (2)~(4) can be isolated from every harmonic of ecg signals, where  $k$  harmonic Eq.5:

$$a_k \cos \left[ 2\pi \frac{k}{N} (i \square t) \right] + b_k \sin \left[ 2\pi \frac{k}{N} (i \square t) \right] \quad (5)$$

In the formula,  $T = N \Delta t$  is the ECG signal fundamental period, namely the period of ECG signal identified in section 1.2,  $k$  harmonic frequency of  $f_k = k/T$ , amplitude of  $k$  harmonic  $c_k = \sqrt{a_k^2 + b_k^2}$ .

Take any of the ecg signal cycle  $N = (t_2 - t_1) / \square t, \square t = 1 / fs$ , after the above calculation, get the harmonic amplitude  $c_k$ . The ratio of  $G_k = c_k / c_1$  to define harmonic amplitude  $c_k$  and fundamental amplitude  $c_1$ , for someone ECG feature vectors. In order to get the feature vector as close as possible to the actual, in a collection of ECG waveform, continuous take 10 cycles as a group, the mean value of  $\bar{G}^1$  to the calculation of  $k=32$  harmonic Bobbi, then take second group, calculates  $\bar{G}^n$ . As the experimental study, the paper took a total of 17 groups. 17 set of curves with  $\bar{G}$  greatest error occurred in the  $k=10$ , the value is  $E_{\max} = 0.0476$ . The average of 170 cardiac cycles of harmonic Bobbi alone as a feature vector that person training base. In view of its 10 consecutive cardiac cycle harmonic Bobbi and 170 consecutive cardiac cycle harmonic Bobbi's error is less than 0.05, in order to reduce the computation, 10 consecutive cardiac cycles, with 0.05 as the threshold used for identity recognition.

**Analysis of Feature Extraction and Experimental Results.** Experiments confirmed, with  $k = 32$  harmonics can be well fitted ECG waveform. Considering the same individual in different periods of collect heart signal amplitude will have the difference, in order to facilitate the comparison and analysis of the feature vector, the maximum value of each set of harmonic amplitude was 1, and then do the following calculation.

The MIT-BIH database provides ECG data of 36 individuals, each in the two groups, each 24 minutes long. In this paper, a set of training template library, a group used to identify, and two cases experiments. The first kind of circumstance is take someone to a set of data in 170 adjacent cardiac cycle, template feature vector is obtained. Take another set of data at any place 10 piece of data, there are 10 adjacent cardiac cycle every piece of data, calculate the 10 groups of feature vector. The second case is respectively from the ECG data of 35 individuals taken 35 piece of data to do the same treatment, as a feature vector of 35 individual ecg. The results in both cases in Fig.4 (a) and (b) in Fig.4

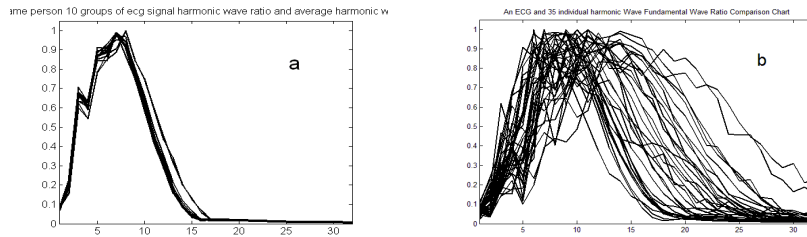


Fig. 4 Spectral features comparison chart

Figure 7 (a) difference in 10 root feature vector curve and the template feature vector of a curve and the maximum value within the range of threshold of 0.05 is 9, the correct probability of 90%. Figure 7 (b) additionally 35 individual characteristic vector curve and the template feature vector of a person in the maximum difference within the threshold value 0.05 has 11, the probability of correct 68.6%, the average accuracy in both cases 79.3%. Reduce discrimination threshold will reduce the accuracy of the first case, increase the correct rate of second kind of situations, and vice versa.

Through analyzing the error characteristic vector curve and the template feature vector curve is greater than the threshold of 0.05 of the situation analysis, found itself in the 15 harmonic later. By the period after  $T = (N \square t) / k$  the 15 harmonic is less than 0.05s. While the adult period QRS wave maximum value at 0.06s ~ 0.1s. This shows that it is possible that using Fourier series expansion of harmonic fundamental wave than QRS wave information extraction is insufficient, it is necessary to study the QRS wave.

## Feature Extraction of Time Frequency Joint Analysis

**Short-time Fourier Transform Features.** Short time Fourier transform is the signal is divided into many small time interval, and then to these small time interval signal short-time Fourier transform. That is, with divergent electrical signals from  $x(n)$  is multiplied by a center for the  $m$  window function  $w(n-m)$  and Fourier transform<sup>[11]</sup>. In other words, the time-frequency method is to use a center of symmetry sliding window function to intercept a period of observation signal, then the Fourier transform is performed on the short-time signal in different time, finally get the time-varying spectral matrix made up of paragraphs signal. The discrete Fourier transform formula:

$$X_{SIFT}(n, \omega) = \sum_m x(m)w(n-m)e^{-jm\omega} \quad (6)$$

Signal in the  $n = m$  near the time-frequency information within a short period of time.

Under normal circumstances, the width <100ms adult QRS wave, T wave or P although sometimes more than one wave, but only one R-wave, and remains unchanged. Therefore, the R-wave starting point, respectively forward, backward extraction 50ms length data as QRS template.

Selects the hamming window, window long for 1/11 of the length of the signal.

**Singular Value Feature Extraction and Results Analysis.** Short-time Fourier transform spectra matrix, the need to extract the feature vector dimensionality reduction. In view of the singular values are stability, proportion and rotational invariance, exactly in line with the requirements of pattern recognition feature. Therefore, for the same individual and different individuals were collected 10 adjacent cardiac cycles QRS wave do singular value feature vector extraction and averaged, the result shown in Fig.5.

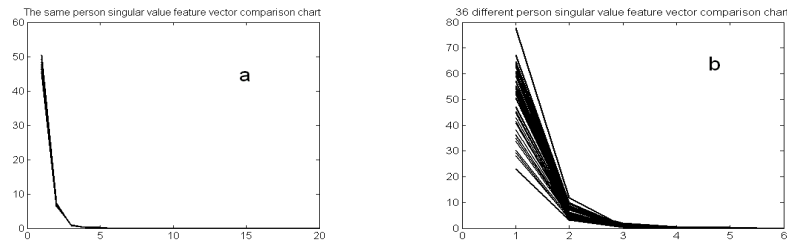


Fig .5 Singular eigenvalues comparison chart

Fig.5 (a) in the same person 10 mean curve of all feature vectors with the template feature vectors curve difference between the maximum probability of correctly within the threshold range of 0.4 to 100%. Fig.5 (b) is another 35 people, with available to joint frequency analysis method is someone's ecg feature vector, the curve with sb template feature vector curve maximum difference in threshold only within 0.4 of 2 roots, the correct probability of 94.44%. The correct rate of recognition was 97.22%.

## Conclusion

ECG signal than other biological characteristics have a higher security, not easy to be forged and has great potential of development. This study proposes a syncretic feature extraction method of ECG signal spectrum analysis and time-frequency joint analysis, then uses the MIT-BIH ECG database on two methods for experiments, found that after the feature fusion after extraction method to obtain very high recognition rate and fast recognition speed. In future research, in the hope of ECG database more abundant support, further find the right vector distance and the threshold, and to verify the robustness of the algorithm.

## References

- [1] Y.Q.Hong, Y.M.Zhou and C.L.Zheng. Based on the distance discriminant identification technology research of ecg signals[D]. Science, technology and engineering, 2014, 14(18):252-253. (In Chinese)
- [2] X.P.Zhao. Biometric identification technology development were reviewed[D]. The criminal technique, 2011, (6):44-48. (In Chinese)
- [3] Jain A K, Ross A, Prabhakar S. An introduction to biometric recognition. IEEE Transactions on Circuits and Systems for Video Technology, 2004; 14(1):4-20.
- [4] L.Yang and Z.D.Zhao. Based on the identification algorithm of ecg signal research[D]. Hangzhou university of electronic science and technology 2012. (In Chinese)
- [5] Jun Shen, Shu-Di Bao, Li-Cai Yang, The PLR-DTW method for ECG based biometric identification. Proc. 33rd Annual Conference of IEEE-EMBS, 2011. PP.5248-5251.
- [6] Yongjin Wang, Foteini Agrafioti, Dimitrios Hatzinakos, and Konstantinos N. Plataniotis. Analysis of Human Electrocardiogram for Biometric Recognition, EURASIP Journal on Advances in Signal Processing, vol. 2008, Article ID 148658.
- [7] Adrian D.C.Chan, Mohyeldin M.Hamdy, Armin Badre, and Vesal Badee, Wavelet Distance Measure for Person Identification Using Electrocardiogram. IEEE Transactions on Instrumentation and Measurement, 2008, 57(2):248-253.

- [8] G.Lv,L.Chen. ECG identification of wavelet transform and support vector machine integration[D],Computer Engineering and Applications, 2013, 49(24):195-199. (In Chinese)
- [9] Platanioatis K, Hatzinaos D, Lee J. ECG Biometric Recognition without Fiducial Detection [C]// Baltimore: Biometrics Symposium BCC'06, 2006: 19—21.
- [10] Y.G.Wan. Digital signal processing of MATLAB[M]. Version 2.Beijing: science press,2012.86-126. (In Chinese)
- [11] Y.H.Wang,G.S.Liu,X.Li and Y.D.Wang. Based on short-time Fourier transform and singular value feature extraction method for target recognition[D]. The signal processing, 1998, 14(2):123-126. (In Chinese)