A Method of ECG Template Extraction for Biometrics Applications

Xiang Zhou, Yang Lu, Meng Chen, Shu-Di Bao, *IEEE Member* and Fen Miao

***Abstract*—ECG has attracted widespread attention as one of the most important non-invasive physiological signals in healthcare-system related biometrics for its characteristics like ease-of-monitoring, individual uniqueness as well as important clinical value. This study proposes a method of dynamic threshold setting to extract the most stable ECG waveform as the template for the consequent ECG identification process. With the proposed method, the accuracy of ECG biometrics using the dynamic time wraping for difference measures has been significantly improved. Analysis results with the self-built electrocardiogram database show that the deployment of the proposed method was able to reduce the half total error rate of the ECG biometric system from 3.35% to 1.45%. Its average running time on the platform of android mobile terminal was around 0.06 seconds, and thus demonstrates acceptable real-time performance.**

I. Introduction

As mobile communication technology and bio-sensor technology matures and social rejuvenation of aging and chronic disease intensifies, Internet of Things (IoT) based healthcare system has been developed rapidly [1-2]. As one of the important terminals in such systems, wearable biosensors can be easily connected to a remote server via phones or other access points for medical monitoring and diagnosis [3-4]. It has advantages of removable operation, ease-of-use, long working hours, abnormal physiological condition alarms, wireless data transmission, and so on. Among physiological signals that can be detected non-invasively, ECG (Electrocardiogram) has become a kind of must-have health information to be captured due to its ease-of-detection and importance of clinical value.

Wearable devices may be shared among family members in many application scenarios. For example, in order to reduce cost, an elderly couple is more willing to share a biosensor for health monitoring. Thus, the problem of user identification in such applications has to be addressed to avoid erroneous results of cross storage of important physiological data [5].

Biometric identification provides airtight security by identifying an individual based on his physiological and/or behavioral characteristics [6]. Each individual has unique biometric characteristics, including physiological traits such as face, fingerprint, iris, and behavioral characteristics like gait and keystroke. Although ECG signal is non-stationary, the existing studies have shown that it can be a new type of biometric traits in applications with less demanding of resolution [7-8]. The validity of using ECG for biometric identification is supported by the fact that the geometrical differences of the heart in different individuals display certain uniqueness in their ECG signals.

There are two main processes that greatly affect the identification performance of ECG biometrics system, i.e. extraction of ECG template/test waveform and identification algorithm for waveform comparison. A lot of identification algorithms have been proposed in current literatures [5, 9-10], while extraction of stable waveform has been overlooked, which is however also important because of the non-stationary characteristics of ECG signals.

This study focuses on the method of waveform extraction from a segment of ECG signals to generate both template and test waveforms, aiming at an optimized performance of identification, where the dynamic time wraping (DTW) is used as the identification method [5]. The extraction method should also be applicable to scenarios with smart mobile terminal, which have become a kind of important platforms for collecting and processing ECG signals. Once ECG signals have been collected or received by a user's smart mobile terminal, it would be great if the terminal can automatically identify the source of the signals using biometrics technology and then store the bio-data to corresponding user database, without having to ask for the account information and password. This would absolutely improve the ease-of-use of the whole healthcare system.

The remainder of this paper is organized as follows. In Section II, the identification using ECG signals is briefly introduced. In Section III, the method of ECG waveform extraction is proposed, followed by experimental analysis in Section IV. A conclusion is finally given in Section V.

II. Identification Using ECG Signal

There have been a lot of studies addressing the problem of using ECG as the biometric trait [5, 7-12]. Due to the space limitation, only the work for one-lead ECG signals by Shen *et al*. [5], which a previous work to this study, is briefly introduced here. After obtaining the ECG signal segment, the piecewise linear representation (PLR) is first used to reduce the data size while at the same time keep important information of the ECG signal segment. Then the dynamic time wraping is used for similarity measures between the signal segments of template and test.

Research supported in part by Zhejiang Nature Science Foundation (No. LY14F010005), National Nature Science Foundation of China (No. 61102087) and Ningbo Nature Science Foundation (No. 2012A610024).

X. Zhou, Y. Lu, and M. Chen are with the School of Electronic and Information Engineering, Ningbo University of Technology , Ningbo, 315016 China (e-mail: 571995729@qq.com, 2863053347@qq.com, nbchen75@sina.com).

S. D. Bao is with the School of Electronic and Information Engineering, Ningbo University of Technology, Ningbo, 315016 China (corresponding author, phone: +86-574-87081299; e-mail: sdbao@ieee.org).

F. Miao is with Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, 518055 China (e-mail: fen.miao@siat.ac.cn).

Compared with the Euclidean distance measure, the DTW method is much more robust, allowing similar shapes to match even if time series are out of phase in the time axis. Suppose we have two time series, S and T, of length m and n, respectively, where

$$S = s_1, s_2, \ldots, s_i, \ldots, s_m \quad (1)$$

$$T = t_1, t_2, \ldots t_j, \ldots t_n \quad (2)$$

We construct an m-by-n matrix $A_{m\times n}$ according to the sort of their time positions. Each matrix element $a_{ij} = d(s_i, t_j) = \sqrt{(s_i - t_j)^2}$ corresponds to the alignment between the points s_i and t_j. A warping path W is a contiguous set of matrix elements that defines a mapping between S and T, and the k^{th} element of W is defined as $w_k = (a_{ij})_k$. The warping path is typically subject to several constraints:

(1) $\max\{m,n\} < K < m+n-1$

(2) $w_1 = a_{11}, w_k = a_{mn}$

(3) Given $w_k = a_{ij}$, $w_{k-1} = a_{i'j'}$, where $0 \le i - i' \le 1$, $0 \le j - j' \le 1$

Then, we can get the path that minimizes the warping cost:

$$DTW(S,T) = \min\left(\sum_{i=1}^{K} w_i\right) \quad (3)$$

The DTW method can be summarized as: finding the cumulative distances of the minimum cost of bending by using the dynamic time warping. $D(m,n)$ is the minimum cumulative distance of the curved path, i.e.,

$$\begin{cases} D(1,1) = a_{11}, \\ D(i,j) = a_{ij} + \min\{D(i-1, j-1), D(i, j-1), D(i-1)\} \end{cases} \quad (4)$$

which is used to find the dynamic time warping path. An example of the DTW path is illustrated in Fig. 1.

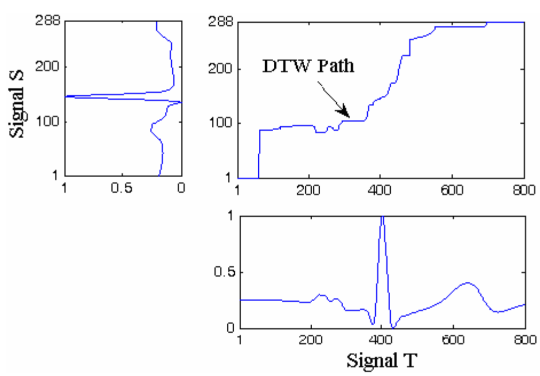


Fig.1 Example of dynamic time warping path

The performance evaluation was carried out on three ECG databases. The analysis results showed that the PLR-DTW method achieves a good accuracy rate. However, the selection of ECG template waveform has been overlooked in this previous work, which actually is very important for stable identification performance.

III. Method of ECG Extraction

A. Preprocessing

Given the raw data of ECG signals, a band-pass filter with the cutoff frequency being from 0.5 Hz to 40 Hz is first applied to denoise the ECG data, for example, removing the power-line interferences. Then, R-peaks are detected using a mathematical morphology method. It is noted that there is no strict requirement on the accuracy of the peak detection. In other words, the peaks such detected can be real R-peaks or not. This does not affect much the performance of the proposed method in this study.

Suppose that the number of detected R-peaks is N+1. Correspondingly, there are a number of N intervals in between the R-peaks. Fig. 2 depicts the R-peaks and its corresponding intervals.

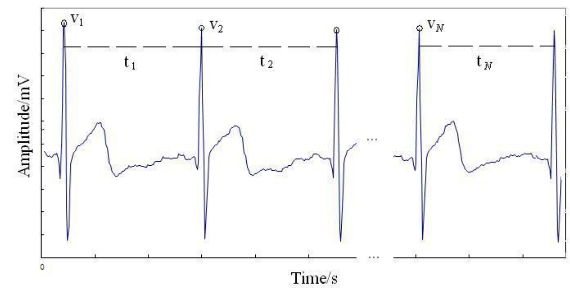


Fig. 2 Example of R-peaks and its corresponding intervals

After obtaining a sequence of R-peak related amplitude and interval, processes of data sorting, region positioning, and waveform extraction will be carried out subsequently, as depicted in Fig. 3. The data sorting is based on the amplitude and interval, respectively, followed by the process of region positioning to find the range with dense points using a cursor both along the amplitude and interval, as depicted in Fig. 4.

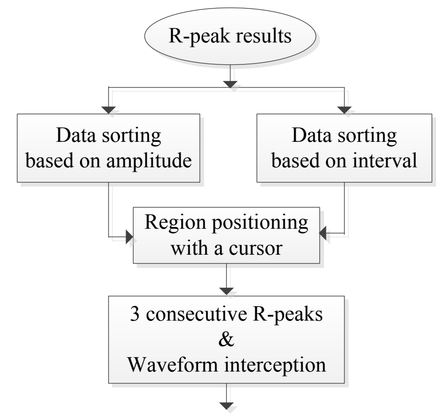


Fig. 3 Block diagram of the proposed method

B. Region positioning

After the data sorting, a 2-dimensional matrix is used to store the peak values with ascending order and its original order, denoted as A_{N*2}. Similarly, another matrix is used to store the intervals with ascending order, denoted as T_{N*2}. Thus, $A[0][0]$ and $A[0][N-1]$ represent the minimum and maximum of the peak amplitudes, respectively. $T[0][0]$ and $T[0][N-1]$

are the minimum and maximum of time intervals, respectively.

As shown in Fig. 4, a cursor is then deployed along the *x*-coordinate and *y*-coordinate, respectively, to find the region with dense points. Define the variable *w* being the width of cursor, and the variable *v* being the cursor moving speed as follows:

$$w = \begin{cases} (A[0][N-1] - A[0][0]) \times p \\ \quad or \\ (T[0][N-1] - T[0][0]) \times p \end{cases} \tag{1}$$

and

$$v = \begin{cases} (\sum_{j=0}^{N-2} A[0][j+1] - A[0][j]) / N \\ \quad or \\ (\sum_{j=0}^{N-2} T[0][j+1] - T[0][j]) / N \end{cases} \tag{2}$$

where *p* is a predefined parameter named as the width varying rate. The setting of parameter *p* is related to the value of *N*, i.e. the number of total peaks. By moving the cursor from left to right and from down to up, it is easy to find the region that has the densest points.

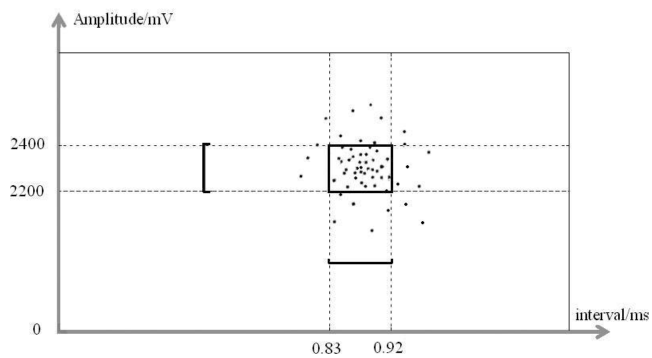


Fig. 4 Example of region positioning

C. Waveform extraction

After finding the region with densest points, 3 consecutive points that are closer to the mid-value of the region can be easily screened, which is then used to intercept the ECG waveform as the ECG template or test waveform for the subsequent identification process, as shown in Fig. 5. Such a waveform contains a complete cycle of the ECG signals, and is more suitable for consequent waveform matching using the DTW method.

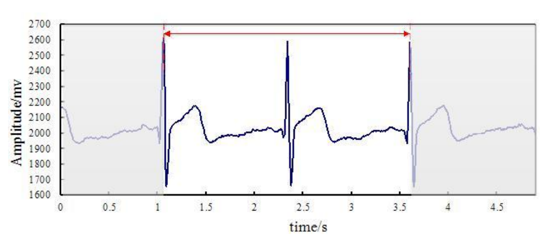


Fig. 5 Example of ECG template/test waveform

IV. Experimental Analysis

A. Experimental data

To examine the effect of the proposed method in ECG biometrics system, both analysis on the MatLab platform and real application platform were carried out. For the simulation analysis, a number of 20 healthy subjects were recruited, half of whom were females with the average age at 28. ECG signals of each subject were recorded for a period of 30 minutes under the conditions of steady and sitting status. In the real application platform, a mini-halter is used to continuously record ECG signals, which are then sent to a smart phone via Bluetooth links for user identification and data storage/transferring. The mini-halter has a sampling rate at 250 Hz and a resolution of 16 bits. The R-peaks of each ECG recording were detected using an effective algorithm.

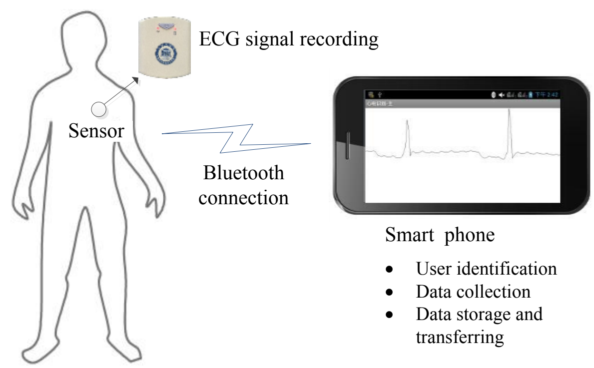


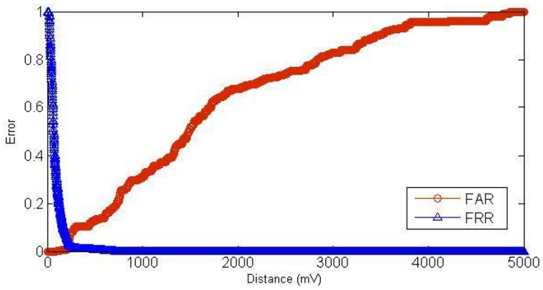
Fig. 6 Experimental environment

B. Accuracy analysis

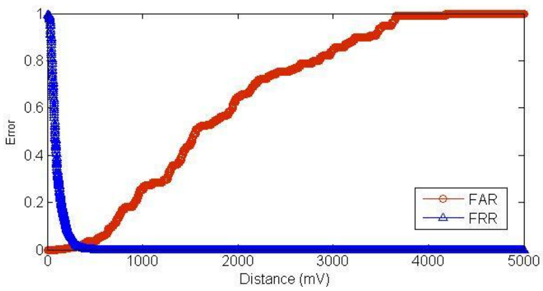
To examine the effectiveness of the proposed method, the false rejection rate (FRR) and false acceptance rate (FAR) of user identification using ECG waveforms obtained with or without the proposed extraction method are calculated, respectively. Fig. 7(a) shows the result without using the proposed extraction method for both template waveform and test waveform, Fig. 7(b) shows the result while the extraction method is used for template waveform but not for test ones, and Fig. 7(c) shows the result while the method is used for both template and test waveforms. The minimum half total error rates are given in Table I, where it can be seen that the accuracy of ECG biometrics can be significantly improved by using the proposed method for ECG waveform extraction. The result for both template and test waveform is quite similar to that for template waveform only, which indicates that proper selection of template waveform is very important while there is not necessary to do so for test waveform, which will instead affect the real-time performance of the system.

C. Time-cost analysis

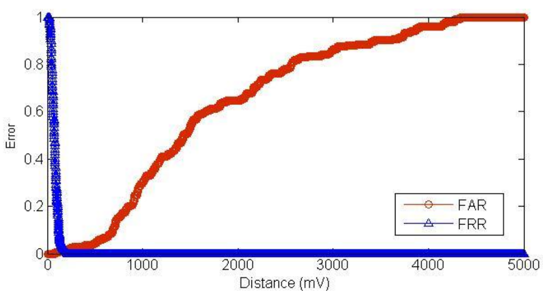
The proposed method has been implemented in the Android smartphone as shown in Fig. 6. Assume that the R-peak number needed is 100. The total running time of the region positioning process and waveform extraction was around 0.06, which is quite acceptable in such applications.



(a)



(b)



(c)

Fig. 7 FAR-FRR performance: (a) without extraction method; (b) extraction method for template waveform; (c) extraction method for both template and test waveform

TABLE I RESULTS OF HALF TOTAL ERROR RATE IN FIG. 7

	(a)	**(b)**	**(c)**
HTER	3.35%	1.50%	1.45%

V. CONCLUSION

To address the accuracy problem of the ECG biometrics, this study has proposed a method of ECG template waveform extraction, where a moving cursor is innovatively deployed to find out the 'best' waveform given a series of ECG waveform segmentation. The analysis of false rejection rate and false acceptance rate based on self-collected experimental data has shown that the accuracy is around 2% improved while using the proposed method for ECG template waveform selection. Compared to the existing methods, the proposed one can significantly improve the accuracy with real-time performance, reducing the running time and enhancing the robust performance of identification. The implementation of the proposed method in mobile terminals has also been carried out and tested. In future, the relationship between storage space and identification accuracy will be further studied to improvement the overall performance of ECG biometrics.

REFERENCES

[1] M.Y. Suraki, M. Jahanshahi, "Internet of things and its benefits to improve service delivery in public health approach," International Conference on Application of Information and Communication Technologies, Azerbaijan, Baku, 2013, pp. 1–4.

[2] L.M.R. Tarouco, L.M. Bertholdo, L.Z. Granville, *et al.*, "Internet of Things in healthcare: interoperability and security issues," IEEE International Conference on Communications, Ottawa, Canada, 2012, pp. 6121–6125.

[3] F. Xue, Y. Li, S.D. Bao, *et al.*, "Healthcare delivery with IP-based sensor networks," IEEE-EMBS International Conference on Biomedical and Health Informatics, Hong Kong, China, 2012, pp. 792–794.

[4] E. Agu, P. Pederson, D. Strong, *et al.*, "The smartphone as a medical device: accessing enables, benefits and challenges," IEEE International Workshop of Internet-of-Things Networking and Control, New Orleans, LA, USA, 2013, pp. 48–52.

[5] J. Shen, S.D. Bao, L.C. Yang, *et al.*, "The PLR-DTW method for ECG based biometric identification," Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Boston, USA, 2011, pp. 5248–5251.

[6] S. Rane, Y. Wang, S.C. Draper, *et al*., "Secure biometrics: concepts, authentication architectures, and challenges," IEEE Signal Processing Magazine, 2013, 30(5): 51–64.

[7] I. Odinaka, P.H Lai, A.D. Kaplan, *et al.*, "ECG biometric recognition: a comparative analysis," IEEE Transactions on Information Forensics and Security, 2012, 7(6): 1812–1824.

[8] C.K. Chen, C.L. Lin, S.L. Lin, *et al*., "A chaotic theoretical approach to ECG-based identity recognition," IEEE Computational Intelligence Magazine, 2014, 9(1): 53–63.

[9] D. P. Coutinho, H. Silva, H. Gamboa, *et al*., "Novel fiducial and non-fiducial approaches to electrocardiogram-based biometric systems," IET Biometrics, 2013, 2(2): 64–75.

[10] M.S. Islam, N. Alajlan, Y. Bazi, *et al*., "HBS: a novel biometric feature based on heartbeat morphology," IEEE Transactions on Information Technology in Biomedicine, 2012, 16(3): 445–453.

[11] A. Lourenco, H. Silva, A. Fred, "ECG-based biometrics: a real time classification approach," IEEE International Workshop on Machine Learning for Signal Processing, 2012, pp. 1–6.

[12] S.I. Safie, J.J. Soragham, L. Petropoulakis, "Electrocardiogram biometric authentication using pulse active ratio," IEEE Transactions on Information Forensics and Security, 2011, 6(4): 1315–1322.