A Construction Method of Personalized ECG Template and Its Application in Premature Ventricular Contraction Recognition for ECG Mobile Phones

Chengyu Liu[1], Peng Li[1], Yatao Zhang[1], Yuan Zhang[2], Changchun Liu[1] and Shoushui Wei[1]

[1] School of Control Science and Engineering, Shandong University, Jinan, P. R. China
[2] Department of Orthopedic Surgery, Jinan Military General Hospital, Jinan, P. R. China

Abstract—**Premature ventricular contraction (PVC) is the most frequent arrhythmia encountered. PVC may occur in health subjects, which is not imminently life-threatening but may require therapy to prevent further problems. So the timely PVC recognition becomes very important for the analysis of electrocardiogram (ECG), especially for the remote ECG monitoring using mobile phones. In this paper, a construction method of personalized ECG template and a succeeding PVC recognition method based on template matching were studied. Firstly, we selected 43 ECG recordings from the MIT-BIH arrhythmia database. All recordings were divided into two datasets (DS1 for training and DS2 for testing) and each dataset approximately contains the same proportion of PVC beats. Subsequently, for each recording (30 min length) in DS1, the first 5 min recordings were used to construct the personalized ECG template and the last 25 min recordings were used for the R-wave peaks detection and PVC recognition, where the template matching method were used. The validity of the proposed methods was tested using DS2. The results showed that: 1) high beat detection accuracy was achieved for both PVC beats and non-PVC beats; 2) the sensitivity and specificity of PVC recognition were 99.11% and 99.96% for the first 5 min recordings respectively, 99.17% and 99.43% for the last 25 min recordings respectively. All the proposed methods can be real-time performed and it shows promising for the application of ECG mobile phones.**

Keywords—**Electrocardiogram (ECG), Personalized ECG template, Premature ventricular contraction (PVC), Template matching, Beat detection.**

I. INTRODUCTION

The automated recognition of arrhythmia has been a research topic for a few decades. Premature contraction is the most frequent arrhythmia encountered, which can be divided into premature supra-ventricular contraction (PSVC, originate from the atria or the atrioventricular node) and premature ventricular contraction (PVC, originate from the ventricle). These two types of premature contraction have distinct difference in electrocardiogram (ECG) morphology. Usually, PSVC has similar morphological features as a sinus beat, but PVC has a higher R peak in amplitude. PVC may occur in health subjects and it is not imminently life-threatening but may require therapy to prevent further problems [1, 2]. So the automatically recognition of PVC is very important for the clinical application and it has been well researched and successful detectors have been designed with high sensitivity and specificity [3-8]. Minami K *et al* used the Fourier-transform neural network for the real-time discrimination of ventricular tachyarrhythmia (VT) [3]. Zhang X S *et al* [4] and Sun Y *et al* [5] used nonlinear complexity measure and multiscale-based nonlinear descriptors for VT and ventricular fibrillation (VF) recognition. Shyu LY *et al* [6] and Lim J S *et al* [7] used the fuzzy neural network for PVC recognition. Tech W *et al* proposed a non-singleton fuzzy logic classifier and applied it for pattern classification of normal sinus rhythm (NSR), VT and VF [8].

The computational burden is a mainly limitation of the aforementioned methods and it limits its application on some battery-driver devices. The knowledge-based method described in [9] is the least computational expensive for its just a few relational operators. But it often cannot well identify PVC. Besides, it will give a wrong recognition result when sinus irregular occurs, which also has obvious alter in the rhythm. Another study also used the correlation coefficient based method, described in [10], constructed the standard PVC template to recognize them. This method leads to an individual problem that all PVC templates existed might mismatch some kinds of PVC beats and erroneous detection could be unavoidable. Although this problem might be solved by adding new templates to the template bank, the computational time will be another limitation for it will rise with the increase of the number of templates.

The aim of this paper is to develop a personalized ECG template construction method and a real-time PVC recognition method based on template matching. Meanwhile, we will use the MIT-BIH arrhythmia database to test the validity of those methods.

II. MATERIALS AND METHODS

A. ECG data acquisition

ECG data were acquired from the MIT-BIH arrhythmia database, which includes recordings of many common and

M. Long (Ed.): World Congress on Medical Physics and Biomedical Engineering, IFMBE Proceedings 39, pp. 585–588, 2012.
www.springerlink.com

life-threatening arrhythmias along with examples of normal sinus rhythm [11]. The database contains 48 ECG recordings and each having two ECG signals of 30 min duration (denoted lead A and B). In 45 recordings, lead A is modified-lead II and for the other three is lead V5. Lead B is lead V1 for 40 recordings and is either lead II, V2, V4, or V5 for the other recordings. Twenty-three of the recordings are intended to serve as a representative sample of routine clinical recordings and 25 recordings contain complex PVC and PSVC arrhythmias.

In this paper, we just pay attention for PVC recognition. All the beats were divided into two classes: PVC and non-PVC beats. For all 48 recordings, 4 paced recordings (recording number: 102, 104, 107 and 217) and 1 recording with serious arrhythmia (recording number: 232) were excluded. The remaining 43 recordings were divided into two datasets and each dataset approximately contains the same proportion of PVC beats and non-PVC beats.

The first dataset (DS1) was training set and was used to evaluate the performance of different candidate ECG templates. The second dataset (DS2) was testing set and was used for a final performance evaluation of PVC recognition method based on the template matching. For each recording, the first 5 min recordings were used to construct the personalized ECG template and the last 25 min recordings were used for the performance evaluation of PVC recognition. The detailed data information can be found in Table 1. The performance evaluation results for DS1 were fed back to the process of template construction. Once the construction method for personalized ECG template was determined, DS2 was performed the similar PVC recognition operation and the final performance evaluation was performed.

Table 1 The extracted beats types of PVC and non-PVC for the dataset 1 (DS1) and dataset 2 (DS2) from the MIT-BIH arrhythmia database

Dataset		First 5 min recordings		Last 25 min recordings	
		PVC	non-PVC	PVC	non-PVC
DS1	beats number	547	7972	2755	39259
	% of total	6.4	93.6	6.6	93.4
DS2	beats number	562	7511	3020	37480
	% of total	7.0	93.0	7.5	92.5

Note: Dataset 1 comprises data from recordings 103, 108, 109, 111, 112, 116, 119, 123, 203, 208, 209, 210, 212, 213, 214, 215, 219, 222, 228, 230 and 231. Dataset 2 comprises data from recordings 100, 101, 105, 106, 113, 114, 115, 117, 118, 121, 122, 124, 200, 201, 202, 205, 207, 220, 221, 223, 233 and 234. The 4 paced recordings (102, 104, 107 and 217) and 1 serious arrhythmia recording (232) are not included in dataset 1 or 2.

B. Construction method of personalized ECG template

The first 5 min recordings were used for the construction of personalized ECG template. R-wave peaks were first detected using wavelet transform modulus maxima (WTMM) method [12]. The RR interval calculated from two consecutive R-wave peaks was then used to form the original RR sequence: $RR=\{RR_1, RR_2,\cdots, RR_N\}$. N is the number of RR intervals. Set $m=10$, then choose m consecutive RR_i values to form a RR sequence segment $X_m(i)=\{RR_i, RR_{i+1},\cdots, RR_{i+m-1}\}$, where $i=1,2,\cdots,N-m+1$. Let denote the standard deviation of $X_m(i)$ as σ_i, it forms a $N-m+1$ dimension vector $\sigma=[\sigma_1,\sigma_2,\cdots,\sigma_{N-m+1}]$. Let k denote the position of the minimal σ_i and $Rate$ denote the mean value of the k-th RR sequence segment $X_m(k)=\{RR_k, RR_{k+1},\cdots, RR_{k+m-1}\}$. In the statistical sense, $Rate$ is mostly likely to represent the real interval of cardiac cycle.

Each RR_i $(i=k,k+1,\cdots,k+m-1)$ consisting in $X_m(k)$ corresponds to an ECG recording segment x_i with a whole cardiac cycle from the current R-wave peak to the succedent one. Then x_i will be enlarged or shortened to have the uniform amount of points (i.e., $Rate$). The mean value vector $\bar{x}$ of all x_i $(i=1,2,\cdots,m)$ was acquired. Two-uniform $\bar{x}$ were connected by nose-to-tail with an overlap of one point. The center segment of this connection, named ECG_{temp} with $Rate$ points, was extracted as the personalized ECG template.

C. Template matching method for PVC recognition

Template matching method is a shape-based matching of ECG morphology features and can effectively recognize the ECG morphology changes especially when PVC occurs. In this section, the PVC recognition method based on template matching concluded two steps: one was the detection for the candidates of R-wave peaks; the other one was PVC recognition.

a) Detection for the candidates of R-wave peaks

Firstly, the cross-correlation function between the ECG recordings and personalized ECG template was calculated:

$$r(n)=\sum_{i=1}^{Rate} y(n+i)ECG_{temp}(i), \tag{1}$$

where $y(n)$ is the last 25 min ECG recordings and n denotes the length of $y(n)$. Then the $r(n)$ will be actualized the normalized operation as follows:

$$\tilde{r}(n)=\left(r(n)-r_{\min}\right)/\left(r_{\max}-r_{\min}\right), \tag{2}$$

where $r_{\min}$ is the minimum of $r(n)$ and $r_{\max}$ is the maximum. The role of the cross-correlation operation is like using a comb to filter the ECG recordings. When an R-wave peak of ECG recordings just coincides with the R-wave peak of ECG template, the $\tilde{r}(n)$ is usually a big value. When an R-wave peak of ECG recordings becomes far away from the R-wave peak of ECG template, the $\tilde{r}(n)$ declines to 0 quickly. A second derivative method was used to detect the peaks of $\tilde{r}(n)$, which are assumed to be candidates of R-wave peaks. If an extra-large RR interval appeared (i.e., more than 1.9 times $Rate$), an alternative lower threshold was used to detect the peak again. At last, a decision process was performed for accurately locating of R-wave peaks in ECG recordings.

b) PVC recognition

The PVC recognition is also carried out by a template matching method. The difference to the template matching applied in the R-wave peaks detection is that this method is a beat-by-beat implementation and the template length is shortened. The normalized correlation coefficient is used as the matching measure, which is described in (3):

$$NC = \frac{\sum_{i=1}^{M}\left[z(i)-\bar{z}\right]\left[R_{temp}(i)-\bar{R}_{temp}\right]}{\sqrt{\sum_{i=1}^{M}\left[z(i)-\bar{z}\right]^2}\sqrt{\sum_{i=1}^{M}\left[R_{temp}(i)-\bar{R}_{temp}\right]^2}}, \quad (3)$$

where z is the QRS complex in the current cardiac cycle, R_{temp} is the shortened ECG template, $\bar{z}$ and $\bar{R}_{temp}$ are the mean values of z and R_{temp}, NC is the current matching result. The z and R_{temp} are acquired by cutting off the ECG recordings and ECG template respectively, making the R-wave peak as the center point. They have the same length $M = 1+2*\lfloor 0.1*Rate \rfloor$. $\lfloor * \rfloor$ means that selecting the maximum integer no more than $*$. A threshold is selected for NC and is denoted as NC_{thre}. The current beat is recognized as a non-PVC beat when $NC \geq NC_{thre}$, and a PVC beat otherwise. The choice of threshold NC_{thre} is described in the following section.

D. The choice of threshold NC_{thre}

The morphology feature of non-PVC beat is similar with that of R_{temp}, while the PVC beat is not. So the NC of non-PVC beat is usually close to 1, while the NC of PVC beat is usually below 0.8. By setting the range from 0.4 to 1, incrementing by 0.05, a series of threshold NC_{thre} are tested for PVC recognition. With the cardiologist's annotations, the recognition sensitivity and specificity can be calculated using the following equations:

$$\text{sensitivity} = A2/(A1+A2)\times 100\% \quad (4)$$

$$\text{specificity} = N1/(N1+N2)\times 100\% \quad (5)$$

where $A1$ is the number of PVC beats that are falsely recognized as non-PVC beats, $A2$ is the number of PVC beats that are truly recognized as PVC beats, $N1$ is the number of non-PVC beats that are truly recognized as non-PVC beats, $N2$ is the number of non-PVC beats that are falsely recognized as PVC beats. A receiver operating characteristic (ROC) curve, a graphical plot of the sensitivity and 1-specificity, will be obtained. The threshold, corresponding to the maximum sum of sensitivity and specificity, is chosen as the NC_{thre}.

III. Results and Discussion

A concise illustration of the personalized ECG template construction, the R-wave peak detection and the PVC recognition for an ECG segment from recording 119 is shown in Fig. 1. The personalized ECG template ECG_{temp} acquired from two-uniform $\bar{x}$ is shown in Fig. 1(a). The shortened ECG template R_{temp} generated using the proposed method in C of section II is shown in Fig. 1(b). These two ECG templates are acquired based on the analysis of the first 5 min recordings. An ECG segment from the last 25 min recordings is shown in Fig. 1(c). The cross-correlation function between this ECG segment and ECG_{temp} is calculated and shown in Fig. 1(d). The candidates locations of R-wave peaks are marked as a '*'. Good accuracy is also achieved even when PVC occurs. The normalized correlation coefficient NC between each QRS complex and R_{temp} is calculated and shown in Fig. 1(e). The red dot line in Fig. 1(e) is the threshold that discriminates the non-PVC beats and PVC beats and the black dot line stands for perfect matching. The beat with the correlation coefficient lower than the red dot line is recognized as a PVC beat and as a non-PVC beat inversely. The recognition results are marked as a '●' for non-PVC beat and a '■' for PVC beat in Fig. 2(c). It can be clearly seen that the normalized correlation coefficients for PVC beats are much lower than those for non-PVC beats.

PVC recognition results of the testing DS2 are shown in Table II. The recognition sensitivity and specificity calculated by (4) and (5) are also given: 99.11% and 99.96% respectively for the first 5 min recordings and 99.17% and 99.43% respectively for the last 25 min recordings. It shows very promising and efficient.

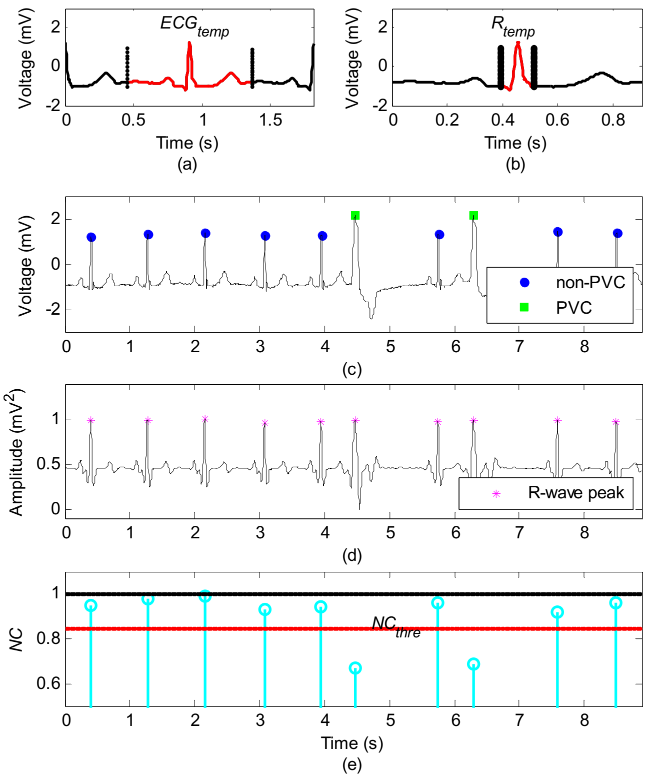

Fig. 1 The illustration of the ECG template construction and PVC recognition procedure. (a) The construction of personalized ECG template ECG_{temp}. (b) The construction of R_{temp} from ECG_{temp}. (c) A segment of ECG recording 119, with the R-wave peaks of non-PVC beats marked by '●' and PVC beats marked by '■'. (d) The result of the normalized cross-correlation function between the ECG recordings and ECG_{temp}, with the candidates of R-wave peaks marked by '*'. (e) The normalized correlation coefficient NC between each QRS complex and R_{temp}.

Table 2 The recognition results of PVC and non-PVC beats using the template matching method for testing dataset (DS2).

Dataset		First 5 min recordings		Last 25 min recordings	
		PVC	non-PVC	PVC	non-PVC
DS2	PVC	557	5	2995	25
	non-PVC	3	7508	214	37266
	sensitivity	99.11%		99.17%	
	specificity	99.96%		99.43%	

IV. Conclusions

Real-time construction method for personalized ECG template and R-wave peak detection method based on template matching were developed in this study. Good detection accuracy had been achieved for both non-PVC beats and ECG beats with PVC's occurrence. An efficient PVC recognizing method based on template matching between QRS complex and the shortened ECG template was also studied, which gained a promising sensitivity (more than 99%) and specificity (more than 99%) from the MIT-BIH arrhythmia database. There is a good application prospect for the real-time processing of ECG mobile phones.

Acknowledgment

We gratefully acknowledge the support of the China Postdoctoral Science Foundation funded project (No. 20110491593).

References

1. Philip C, Maria OD, Richard BR (2004) Automatic classification of heartbeats using ECG morphology and heartbeat interval features. IEEE Trans. Biomed. Eng. 51(7):1196-2006
2. Liu CY, Liu CC, Shao P et al. (2011) Comparison of different threshold values r for approximate entropy: application to investigate the heart rate variability between heart failure and healthy control groups. Physiol. Meas. 32(2):167-180
3. Minami K, Nakajima H, Toyoshima T (1999) Real-time discrimination of ventricular tachyarrhythmia with Fourier-transform neural network. IEEE Trans. Biomed. Eng. 46:179-185
4. Zhang XS, Zhu YS, Thakor NV et al. (1999) Detecting ventricular tachycardia and fibrillation by complexity measure. IEEE Trans. Biomed. Eng. 46:548-555
5. Sun Y, Chan KL, Krishnan SM (2005) Life-threatening ventricular arrhythmia recognition by nonlinear descriptor. Biomed. Eng. Online 4(1)
6. Shyu L, Wu YH, Hu W (2004) Using wavelet transform and fuzzy neural network for VPC detection from the holter ECG. IEEE Trans. Biomed. Eng. 51(7):1269-1273
7. Lim JS (2009) Finding Features for Real-Time Premature Ventricular Contraction Detection Using a Fuzzy Neural Network System. IEEE Trans. Neural. Netw. 20(3):522-527
8. Teck W, Woei W (2011) Non-singleton genetic fuzzy logic system for arrhythmias classification. Engineering Applications of Artificial Intelligence 24:251-259
9. Tsipouras MG, Fotiadis DI, Sideris D (2005) An arrhythmia classification system based on the RR interval signal. Artif. Intell. Med. 33(3):237-250
10. Chiu CC, Lin TH, Liau BY (2005) Using correlation coefficient in ECG waveform for arrhythmia detection. Biomed. Eng.-App. Bas. C. 17(3):147-152
11. Mark R, Moody G. (1997). MIT-BIH Arrhythmia Database. Available: http://ecg.mit.edu/dbinfo.html.
12. Martinez JP, Almeida R, Olmos S et al. (2004) A wavelet-based ECG delineator: evaluation on standard databases. Biomedical Engineering, IEEE Transactions on 51(4):570-581

Author: Chengyu Liu
Institute: Shandong University,
Street: Jingshi 17923
City: Jinan
Country: China
Email: bestlcy@sdu.edu.cn