Digital Signal Processing 50 (2016) 93–102

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Effective and efficient detection of premature ventricular contractions based on variation of principal directions

Roozbeh Zarei [a,*], Jing He [b,c], Guangyan Huang [d], Yanchun Zhang [a,e]

[a] Centre for Applied Informatics, College of Engineering and Science, Victoria University, Melbourne, Australia
[b] College of Engineering and Science, Victoria University, Melbourne, Australia
[c] College of Information Engineering, Nanjing University of Finance and Economics, Nanjing, China
[d] School of Information Technology, Deakin University, Melbourne, Australia
[e] School of Computer Science, Fudan University, Shanghai, China

ARTICLE INFO

Article history:
Available online 21 December 2015

Keywords:
Electrocardiogram (ECG)
Premature Ventricular Contraction (PVC)
Principal Component Analysis (PCA)
MIT-BIH arrhythmia database

ABSTRACT

Classification of electrocardiogram (ECG) data stream is essential to diagnosis of critical heart conditions. It is vital to accurately detect abnormality in the ECG in order to prevent possible beginning of life-threatening cardiac symptoms. In this paper, we focus on identifying premature ventricular contraction (PVC) which is one of the most common heart rhythm abnormalities. We use “Replacing” strategy to check the effects of each individual heartbeat on the variation of principal directions. Based on this idea, an online PVC detection method is proposed to classify the new arriving PVC beats in the real-time and online manner. The proposed approach is tested on the MIT-BIH arrhythmia database (MIT-BIH-AR). The PVC detection accuracy was 98.77%, with the sensitivity and positive predictivity of 96.12% and 86.48%, respectively. These results are an improvement on previous reported results for PVC detection. In addition, our proposed method is effective in terms of computation time. The average execution time of our proposed method was 3.83 s for a 30 min ECG recording. It shows the capability of the classifier to detect abnormal PVCs in online manner.

1. Introduction

The electrocardiogram (ECG) is the noninvasive method used to detect heart disease. It reflects the electrical activity of the heart over time and provides information about the state of the heart. The ECG is commonly used to detect cardiac rhythm abnormalities also known as arrhythmias. Arrhythmias can be divided into two groups. The first group is life-threatening and require immediate treatment with a defibrillator such as ventricular fibrillation and tachycardia. The second group is not imminently life-threatening but may require treatment to prevent further problems such as premature ventricular contractions (PVCs).

Recently portable ambulatory ECG monitoring has attracted increasing attention due to its ability to meet the growing clinical needs for personal healthcare and remote monitoring, etc. [1]. Its aim is to help clinicians to diagnose critical cardiac conditions by processing ECG signal in real-time. Premature ventricular contractions (PVCs) are premature heartbeats originating from the ventricles of the heart. They are one of the most common heart rhythm abnormalities which can be linked to mortality associated with myocardial infarction [2]. Therefore, their immediate detection and treatment is essential to prevent possible beginning of life-threatening cardiac conditions.

Many methods for automatic classification of various arrhythmias have been developed based on different features extraction and classification methods. Features methods include ECG morphology [3,4], heartbeat interval features [5,6], frequency-based features [7,6], higher order cumulant features [8], Karhunen–Loeve expansion of ECG morphology [5], and hermite polynomials [9]. Several classifiers have been used for heartbeats classification including linear discriminants [5,10], artificial neural networks [11,6, 12], and self-organizing networks [9]. Although these methods accurately identify PVCs, they have long computational time. Therefore, due to high complexity of their algorithms, they are not suitable for portable ambulatory ECG devices where the computation resources to process ECG in real-time is limited. In addition, some of existing high detection results are based on small data sets or used overlapped training and testing data sets. For example, wavelet feature extraction in tandem with fuzzy neural network classification was used in [11] to classify PVC beat with 97.04% ac-

* Corresponding author.
E-mail addresses: roozbeh.zarei@live.vu.edu.au, roozbeh.zarei@gmail.com (R. Zarei).

http://dx.doi.org/10.1016/j.dsp.2015.12.002

curacy. However the method was tested on only seven files of the MIT/BIH arrhythmia database [13].

To address the aforementioned problems, this study focuses on developing an online PVC detection algorithm with low-complexity. We observe that replacing a non-PVC beat with a PVC beat will cause a larger effect on principal directions than replacing a non-PVC beat with another non-PVC beat. From this observation, we extract the dominant principal direction from training data set which consists of K normal beats to characterize the normal profile for the dataset. Then we apply "Replacing" procedure to check the effects of each individual new arriving heartbeat on the variation of principal directions. The beat is considered as PVC beat if this score is greater than a certain threshold. Since the proposed method needs to calculate the principal directions n times for an ECG signal with n heartbeats to classify the beats, we apply the power method [14] to alleviate this heavy loading and reduce the computational complexity. Our proposed method is evaluated on two non-overlapping data sets (DS1 and DS2) from the MIT-BIH arrhythmia database (MIT-BIH-AR) [13] in order to assess its performance over a large data set.

The main contributions of this paper are summarized as follows: (1) we develop a low-complexity PVC detection method based on the variation of dominant principal direction; (2) our proposed method uses a small training data set to train the classifier; (3) it achieves high accuracy detection rates over large data set and demonstrates the capability to detect abnormal PVCs in online manner.

The remainder of this paper is organized as follows. The ECG database is discussed in Section 2. Section 3 presents our proposed PVC classification method. Section 4 describes evaluation criteria which is used to evaluate the performance of proposed method. Section 5 presents the experimental results, including comparisons with previously published work. Finally, Section 6 concludes this paper.

2. ECG data

This study uses a publicly available benchmark dataset, MIT-BIH arrhythmia database [13] to test the effectiveness of the proposed method. The database consists of 48 two-lead (denoted lead A and B) recordings of approximately 30 minutes and sampled at 360 Hz. Note that only recordings of lead A (modified-lead II (MLII)) were used in this study. In accordance with the AAMI recommended practice, four recordings (#102, #104, #107 and #217) containing paced beats were discarded since these beats do not retain sufficient signal quality for reliable processing [15]. The remaining 44 recordings were divided into two datasets (DS1 and DS2) with each dataset containing ECG data from 22 recordings. For comparison purposes, We used the same dataset division scheme used in [10,16,17] for the DS1 and DS2.

The data consist of 5 different heartbeat classes. Class N contains beats originating in the sinus node (normal and bundle branch block beat types), class S contains supraventricular ectopic beats, class V contains premature ventricular contraction beats (the ventricular escape beats were also classified as the PVC beats in this study [10,16,15,17]), class F contains beats that result from fusing normal and ventricular ectopic beats, and class Q contains unknown beats. Table 1 shows the breakdown of each dataset by heartbeat classes. The first dataset (DS1) was used to evaluate the performance of different candidate classifiers. The second dataset (DS2) was used for a final performance evaluation of our proposed method.

3. Method for online PVC heartbeat detection

Fig. 1 depicts the stages of our proposed system for online detection of premature ventricular contraction beats (PVC). It consists of two stages: a processing stage and an abnormality detection stage. The ECG lead MLII is applied at the input to the processing stage. A sample ECG signal from recording #119 in the MIT-BIH arrhythmia database is shown in Fig. 2. The processing stage consists of heartbeat detection, segmentation and normalization modules. The heartbeat detection module attempts to locate all heartbeats. The segmentation module extracts single beats from the entire ECG. The normalization module is concerned with normalizing heartbeats that are processed by the classifier stage. The abnormality detection stage contains a classifier unit that classifies heartbeats into two classes: Non-PVC and PVC heartbeat. The classifier utilizes variation of dominant principal directions to determine PVC heartbeats. It contains some parameters which are set during the training phase to optimize the classification performance. The modules forming these stages are discussed in more details below.

3.1. Heartbeat detection

The "Pan and Tompkins" QRS detection algorithm [18] was used to detect the location of QRS complexes. We have chosen this method due to its computational simplicity and ease in implementation. The Pan–Tompkins algorithm identifies the R-point based on digital analyses of the slope, amplitude, and width of the ECG data. It consists of the following steps: band pass filtering of the ECG (to reduce interference present in the ECG signal), differentiating the signal (to provide the QRS complex slope information), squaring of the data samples (to emphasize peaks), moving average filtering (to smooth close-by peaks) and detecting R point. The algorithm automatically adjusts the thresholds and parameters periodically to adapt to changes in QRS morphology and heart rate.

3.2. Segmentation and normalization

After detection of R-points, the ECG signal was segmented such that each segment contain a portion of signal before and after each R-point. In this study, each segment consists of 50 samples (138.88 ms) before R-point, 99 samples (275 ms) after R-point. Each of these 150 samples (416.67 ms) segment are considered as a single beat segment. Fig. 3a depicts first segment extracted from the sample ECG signal shown in Fig. 2. Then each segment is normalized to have a mean of zero and a standard deviation of one in order to eliminate the offset effect and to decrease the effect of false classification due to personal or instrumental differences. Fig. 3b shows the effect of normalization on a segment obtained from the patient record #119.

3.3. PVC detection via variation of dominant principal directions

In this section, we first briefly review the Principal Component Analysis algorithm (PCA). Then based on the replacing strategy, we present the effects of PVC beats on the derived dominant principal direction (a direction that captures the largest variability in the data) and our proposed PVC detection method. Finally, an effective computation for estimating principal directions in replacing strategy is proposed.

3.3.1. Principal component analysis

PCA is a linear dimensionality reduction technique, which determines the principal directions of the data distribution. It involves constructing the data covariance matrix and calculating its dominant eigenvectors to obtain the principal directions. These directions contain the most important aspects of the data. Let $\mathbf{A} = [x_1^T; x_2^T; \ldots; x_n^T] \in R^{n\times p}$ where each row x_i represents a data instance in a p dimensional space, and n is the number of the