Classification of Phonocardiogram Signals Based on Envelope Optimization Model and Support Vector Machine

Lijun Yang[a,*], Shuang Li[a], Zhi Zhang[b], Xiaohui Yang[a]

[a] *School of Mathematics and Statistics, Henan University, Kaifeng 475004, China.*
[b] *Department of Computer Science, Georgia Institute of Technology, Atlanta, GA, 30332, USA.*

Abstract

The prevention and diagnosis of cardiovascular diseases have become one of the primary problems in the medical community since the mortality of this kind of diseases accounts for 31% of all global deaths in 2016. Heart sounds, which is an important physiological signal of human body, mainly come from the pulsing of cardiac structures and blood turbulence. The analysis of heart sound signal plays an irreplaceable role in early diagnosis of heart disease since they contain a large amount of pathological information about each part of human heart. Heart sounds can be detected and recorded by Phonocardiogram (PCG). As a noninvasive method to detect and diagnose heart disease, PCG signals have been paid more and more attentions by researchers. In this paper, a novel envelope extraction model is proposed and used to estimate the cardiac cycle of each PCG signal. We give a strategy combing empirical mode decomposition (EMD) technique and the proposed envelope model to extract the time-domain features. After applying EMD process to each PCG signal, the second intrinsic mode function is chosen for further analysis. Based on the proposed envelope model, the cardiac cycles of PCG signals can be estimated and then the time-domain features can be extracted. Combining with the frequency-domain features and wavelet-domain features, the feature vectors are obtained. Finally, the support vector machine (SVM) classifier is used to detect the normal and abnormal PCG signals. Two public data sets are used to test our framework in this paper. And classification accuracies of more than 96% on both data sets show the effectiveness of the proposed model.

Keywords: Heart Sound, Phonocardiogram, Empirical Mode Decomposition, Envelope Extraction, PCG Segmentation, Support Vector Machine

1. Introduction

According to the report from World Health Organization, there are about 17.9 million people lost their lives because of cardiovascular diseases (CVDs) in 2016, which represents 31 percent of all world's deaths [26]. In China, CVD prevalence is on the rise and its mortality ranks first, higher than that of tumors and other diseases [9]. Therefore, the prevention and diagnosis of CVDs have become one of the primary problems in the medical community. Some researchers have tried to use mathematical methods to model and analyze the operation mechanism of human blood vessels [15, 16, 17, 18]. A more direct approach, however, is through cardiac auscultation. Heart sounds, which is an important physiological signal of human body, mainly come from the pulsing of cardiac structures and blood turbulence. A large amount of pathological information about each part of human heart are contained in heart sounds, such as atrial and ventricular, large blood vessels, cardiovascular and valvular function state [19]. Heart sounds reflect many pathological conditions of the cardiovascular system in human body. Therefore, it is possible to diagnose heart diseases through heart sound analysis. Because heart auscultation is non-invasive, low-cost and easily accepted by patients, it is the common way to diagnose heart sound. However, the duration of heart sounds or murmurs is so short that it is often difficult to accurately capture their intensity, frequency, duration, splitting, and their relative time relationships. Therefore, cardiac auscultation requires long-term, professional training and the accumulation of clinical experiences. All these limit the physicians from making accurate diagnoses in the early stages of cardiac diseases through the stethoscope for patients. Therefore, it is necessary to develop a computer-aided heart sound analysis system since it can help the physicians save time and give a more reliable preliminary diagnosis [46].

*Corresponding author
Email addresses: yanglijun@henu.edu.cn (Lijun Yang), lswsrain@163.com (Shuang Li), zhizhang@gatech.edu (Zhi Zhang), xhyang@henu.edu.cn (Xiaohui Yang)

Heart sounds can be recorded and displayed on the computer through phonocardiogram (PCG) [29]. PCG provides a visual representation of the heart sounds and the physicians can understand and describe a patient's heart problems by looking at the patient's PCG. Moreover, it is because of the digital form of PCG signals that some technical means can be applied to PCG analysis so that we can dig the deep pathological information. Therefore, compared with auscultation, PCG automatic analysis can monitor not only the condition of human heart but also the effect of cardiac medication on heart diseases. In the automatic analysis of PCG signals, correct localization and classification are the two challenging tasks since cardiac cycles are always inconsistent, while the accurate estimation of cardiac cycle is the premise of the two tasks. Various techniques are used by researchers for PCG signal automatic analysis, such as using filtering-based techniques for noise removal [14, 48], transform-based algorithms for features extraction [3, 4], envelope-based methods for cardiac cycle estimation [43, 28, 42], etc. PCG classification is one of the research hotspots in PCG analysis field. By using different methods to deal with PCG signals and consequently to extract the discriminative features, one can design or choose a classifier to realize automatical classification of PCG signals as normal or abnormal. To design an automatical classification system for PCG signals is of great help for the realization of concerned about the early stage of the disease and non-invasive screening.

In the last decade, lots of research work on PCG signal analysis has been carried out, which is mainly divided into the following aspects:

- Noise removal

  Heart sound recordings are always disturbed by various factors and high frequency noise in the process of collection, including lung sounds, muscle contraction, breath sounds, and background noise. All these noise components can make subsequent PCG signal analysis difficult. Therefore, it is important to use appropriate denoising algorithm on PCG signals before the further processing. The substance of denoising process is to find a suitable filter to separate the PCG signal from the noise. In reference [14], the authors used the least mean square algorithm to develop an adaptive filter for noise removal. In [33], the denoised signals were obtained by using median and low pass filter (Butterworth) on normalized PCG signals. Some researchers considered the decomposition methods for noise removal. The authors in [1] applied wavelet transform on PCG denoising based on the fact that the wavelet coefficients of true signal components always are much larger than those due to noise components. Empirical Mode Decomposition (EMD) method was used to PCG signal denoising in [38]. After EMD processing, a PCG signal can be decomposed into several subcomponents named as intrinsic mode functions (IMFs). Then, the energy density of the IMFs were compared to determinate which IMF can be used for reconstruction. In another study [48], the authors proposed a noise removal method which combined multi-level SVD and compressed sensing. Moreover, in the latest literature [34], the authors use the techniques that combine nonnegative matrix factorization and adaptive contour representation for PCG noise removal.

- Segmentation and location

  PCG segmentation is always a difficult problem in the field of PCG analysis. The key step in PCG signal analysis is to estimate the independent cardiac cycle, which is based on PCG segmentation. In a cardiac cycle, the heart can produce four heart sounds. In general, only the first (S1) and the second (S2) heart sounds can be heard and be called as the fundamental heart sounds. Accurately detecting the locations of S1 and S2 is the purpose of PCG segmentation [6]. Afterwards, the cardiac cycle can be estimated according to the locations of S1 and S2. Most of segmentation techniques use Electrocardiogram (ECG) [40, 30]. However, PCG and ECG recordings are not available simultaneously. The authors in [7] gave a complete review of the techniques and trends on PCG segmentation. Envelope-based technique is one of the popular approaches in the problem of PCG segmentation [25, 10, 43]. In [28], the authors calculated the Shannon energy envelope after applying the S-transform on PCG signal for segmentation, while [42] used the same Shannon envelope but based on the empirical wavelet transform. In [44], the authors calculated the envelope of a signal by solving an optimization problem. Inspired by this, we give a novel envelope model which derives from the modified model of [44]. And then the cardiac cycle can be estimated based on this envelope model.

- Feature extraction and classification

  Feature extraction is a key step in PCG analysis. Through various techniques, original PCG signals can be transferred to feature vectors with low dimension and meanwhile information remained. The features on PCG signals mainly can be extracted in three ways: time domain (time-domain features) [13], frequency domain (frequency-domain features) [37], and time-frequency domain (time-frequency features) [27, 2].

Although some important information of PCG signals can be found in time domain, there also has discriminative information in frequency domain that can not be ignored. Therefore, fusion features extracted from multiple techniques are often used. These features are commonly extracted by using transform ways, such as short time Fourier transform [39], S transform [28] and wavelet transform[2, 4]. EMD is another signal decomposition technique which is suitable for the non-linear and non-stationary signals. Using EMD, a PCG signal can be decomposed adaptively into several IMFs which contain different frequency components of the original signal, and then one can extract the discriminative features from these IMFs [3, 33]. After extracting features, the next step is to select an appropriate classifier to complete the classification. Various machine learning algorithms have been used by researchers to perform the classification, such as ANN [41], SVM [11, 20], deep learning [9, 31, 36] and some fusion classification methods [23, 32]. In this paper, we choose SVM classifier for PCG classification.

As mentioned above, envelope-based technique is popular in the problem of PCG segmentation, which does not require the aid of other auxiliary signals (such as ECG). Shannon energy envelope and Hilbert envelope are two kinds of envelopes commonly used. However, Shannon energy envelope often struggles to capture the details of the PCG signals, while Hilbert envelope has too many burrs and is not as smooth. All these drawbacks will affect the recognition accuracy for PCG classification. In this paper, we propose a novel envelope extraction model which is based on a constrained optimization problem. The new envelope accords with the physical intuition and is more smooth. Combing EMD method and the proposed envelope model, the cardiac cycles of PCG signals can be estimated and then the time-domain features can be extracted. Finally, the feature vectors (containing the features of time-domain, frequency-domain and wavelet-domain) are fed into SVM classifier for PCG classification. Experimental results show the good performance of the proposed envelope model for extracting the envelopes of signals and PCG classification.

This paper is organized as follows: In Section 2, we describe the proposed method in detail. It can be seen that a novel envelope extraction model is provided. Based on the proposed envelope model, cardiac cycles of heart sounds can be estimated and then the time-domain features of PCG signals are extracted. The features in frequency domain and wavelet domain are also obtained in this section. Experimental results are presented in Section 3. Finally, Section 4 is a brief conclusion of this paper.

2. Method

In this section, the proposed model for PCG classification is described in detail. Our method can be boiled down to the following steps:

Step 1 Preprocessing. The PCG recordings are always contaminated with various high frequency noise in the collection process. While most information in heart sounds is contained in the low frequency components. Therefore, the filter process is needed in preprocessing. Moreover, interception and normalization are also necessary in order to get the signals with the same length and intensity.

Step 2 Cardiac cycle estimation based on envelope extraction. Cardiac cycle is important in PCG classification, which is commonly estimated by envelope extraction. A novel envelope model based on constrained optimization will be proposed in this paper. We do not extract the envelope from PCG signal directly, instead we extract the envelope of the chosen IMF of each PCG signal after using EMD method. According to the envelope, S1 and S2, the most important components of PCG recordings, are identified to divide the heart cycle, which is used for further feature extraction.

Step 3 Feature extraction. In this paper, the features of PCG signals derive from three aspects: time-domain features based on envelope which contain the heart rate (hr), the ratio of S1 duration to heart rate (S1h/hr), the ratio of S2 duration to heart rate (s2h/hr), the ratio of diastole to systole (D/S), wavelet-domain features based on wavelet transform which contain wavelet coefficient energies (e_1 e_5), frequency-domain features based on Fourier transform which contain the energy proportions of three frequency bands (Eh, Em*and*El), the maximum power spectral density (Maxf) and the corresponding point (Maxp).

Step 4 Classification. Feature extraction is followed by the classification of PCG signals using SVM, which has a linear kernel.

A brief illustration of our method is given in Fig. 1.

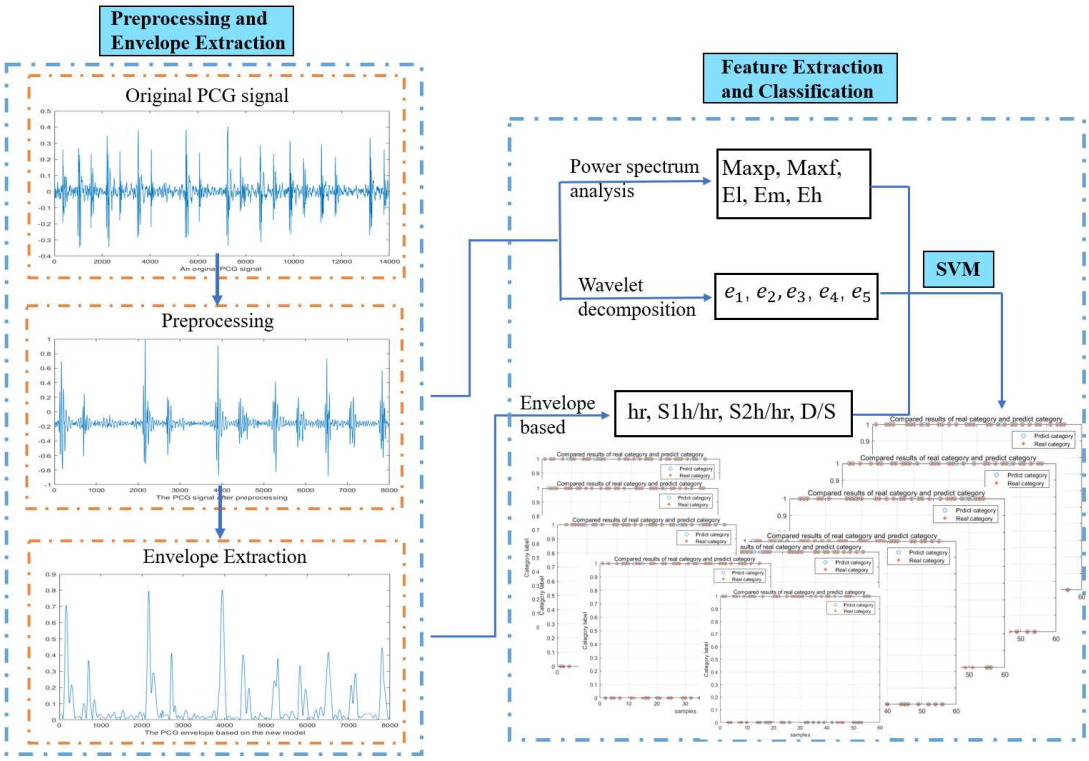

Figure 1: Flow diagram of the proposed scheme.

2.1 Preprocessing

PCG signals are preprocessed for further noise reduction. The main information in PCG recordings is distributed in the low frequency, whereas noises are in the high frequency part. Hence, we resample all the PCG signals to 2000 Hz in this step. To realize the resampling, we use the Butterworth band pass filter which order is 5 and the cutoff frequencies from 30 HZ to 200 Hz. Additionally, wavelet threshold denoising method is used for removing the noise which spectrum is overlapped with those of PCG signal. In this paper, we choose sym8 as the wavelet basis function to perform 3 layer decomposition by using the Rigrsure soft threshold. After that, the following formula is used for the normalization of PCG signals [45]:

$$y[n] = \frac{x[n]}{\max_n(|x[n]|)} \tag{1}$$

where $x[n]$ is the original PCG signal and $y[n]$ is the corresponding PCG signal after normalized.

Due to the synchronization with the biological process of repeating the heart cycle, heart sound signal is approximately periodic. Therefore, we can apply the interception method, which can not only reduce the complexity of data processing, but also obtain a relatively complete cardiac cycle. The detailed process of interception is as follows:

- A suitable threshold α is selected, and the normalized signal is binarized according to this threshold. Suppose $z[n]$ is the signal after binarized, then we have:

$$z(n) = \begin{cases} 1, & |y[n]| \geq \alpha \max_n(y[n]) \\ 0, & |y[n]| < \alpha \max_n(y[n]). \end{cases} \tag{2}$$

- Since the sampling frequency in this paper is 2000 Hz, we first find out where the first two thousand points have the largest interval, and then use the position corresponding to the half of the interval as the starting point for interception. From the starting point, we intercept 8000 sample points backwards, and use these 8000 sample points as the last intercepted signal.

The heart sound signal before and after interception processing is shown in Fig. 2.

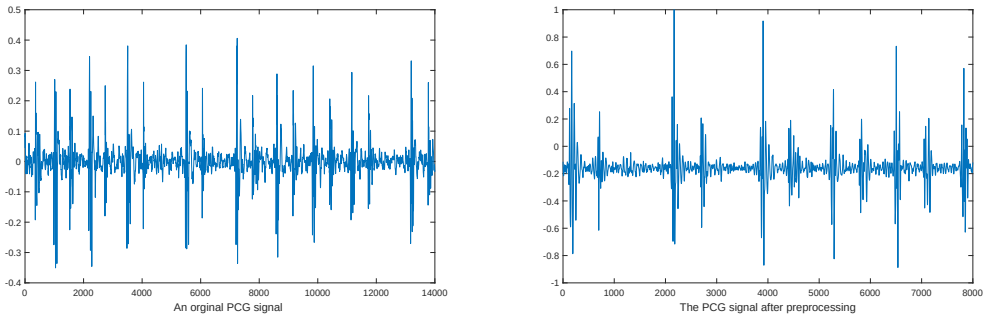

Figure 2: Left: A PCG signal; Right: the PCG signal after preprocessing.

2.2 Envelope extraction model for PCG signals

Cardiac cycle is important in heart sound classification. One of the common estimation methods is based on envelope extraction for PCG signals. In this subsection, we propose a novel envelope extraction model which is inspired from [44]. We do not extract the envelope from the original PCG signal directly, instead we extract the envelope of the chosen IMF after using EMD to the PCG signal. The IMFs of each PCG signal are separated by the iterative sifting process of EMD which is arranged by frequency from high to low. The frequency of the main components of the heart sound signal is concentrated in the range of 50-200 Hz, which is in the middle of the entire signal frequency range. Therefore, in most cases, the second or third IMF is selected as the best IMF component that represents the corresponding PCG signal. In this paper, the second IMF of each PCG signal is chosen for further analysis. For detailed description of EMD, the interested readers are referred to [24].

In our previous paper, we proposed an optimization model of signal envelope [44]. Since the second IMF of each heart sound signal is chosen to estimate the cardiac cycle, the envelope of the IMF should approximately satisfy that the upper and lower envelopes are symmetric. Different from the model in [44], we take the absolute value of the chosen IMF of given PCG signal before the envelope calculation. Assuming $z[n]$ is a PCG signal after preprocessing and $IMF_z[n]$ is the chosen IMF of $z[n]$ by using EMD. Let $u[n]$ is the extracted upper envelope from $IMF_z[n]$. According to [44], $u[n]$ can be calculated by the following L_1-minimization problem:

$$\begin{cases} \text{Minimize}_{u\in\mathbb{R}^n} & \|\Phi u\|_1; \\ \text{Subject to} & u \geq |IMF_z|,\ Au = b. \end{cases} \tag{3}$$

where Φ is 4th-order difference matrix:

$$\Phi = \begin{pmatrix} 1 & -4 & 6 & -4 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & -4 & 6 & -4 & 1 & 0 & \cdots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \cdots & 0 & 1 & -4 & 6 & -4 & 1 & 0 \\ 0 & \cdots & \cdots & 0 & 1 & -4 & 6 & -4 & 1 \end{pmatrix}.$$

It is to note that Φ also can be chosen as $2th$- or $3th$-order difference matrix. Here, A is the position matrix of the tangential points between $u[n]$ and $IMF_z[n]$. The order of A is $M \times N$ where M is the numbers of the tangential points and N is the length of the input signal. b is the extrema vector of order M which is also determined by the tangential points. The calculation of A and b can be referred to [44]. Problem (3) can be solved by Algorithm 1 which is based on the split Bregman iteration.

Algorithm 1. Upper Envelope Extraction Algorithm

1. initialize: $b^0 = c^0 = d^0 = 0, u^0 = |IMF_z|$;

2. while not converge

- $b^{k+1} = b^k + (b - Au^k), \quad c^{k+1} = c^k + (\Phi u^k - d^k)$;
- $u^{k+1} \leftarrow \min\{\frac{\lambda}{2}\|Au - b^{k+1}\|^2 + \frac{\mu}{2}\|d^k - \Phi u - c^{k+1}\|^2\}$;
- $u^{k+1} \leftarrow \text{Proj}(u^{k+1})$;
- $d^{k+1} \leftarrow \min\{\|d\|_1 + \frac{\mu}{2}\|d - \Phi u^{k+1} - c^{k+1}\|^2\}$;

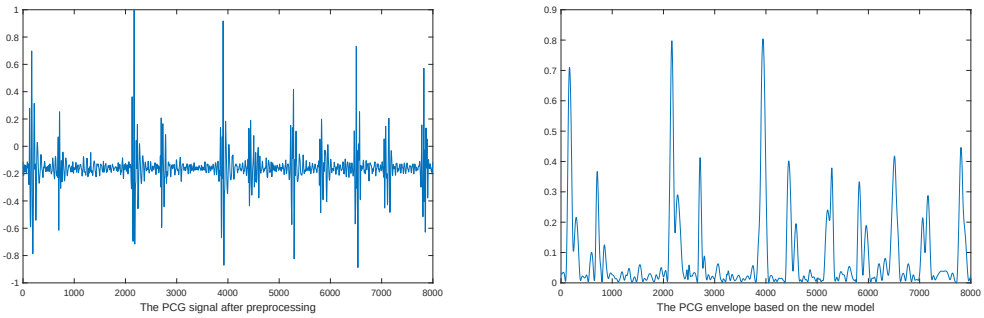

Figure 3: A normal PCG signal (left) and its envelope based on the proposed model (right).

3. end while.

Fig. 3 illustrates the envelope of one normal PCG signal.

2.3 Segmentation of Heart sound signals

Once obtaining the envelopes of PCG signals, we can estimate their cardiac cycles and segment the PCG signals consequently. In this subsection, we use the method in [47] for PCG segmentation. This method estimates the cardiac cycle in advance, meanwhile adjusts the errors during the segmentation process, which makes that errors in the segmentation process of one-stop cardiac record not in effect of the accuracy of subsequent segmentation. Different from [47], we use the proposed envelope model instead of Shannon energy envelope.

Suppose the length of preprocessed PCG signal $z[n]$ is N, and the envelope of $z[n]$ is $u[n]$, which can be calculated by the following formula

$$C[m] = \sum_{n=1}^{N-n+1} u[n] \frac{u[n+m-1]}{N-n+1}, \tag{4}$$

where $N-n+1$ is the equilibrium factor in order to avoid the autocorrelation coefficient $C[m]$ rapidly reduction with the increasing of m. Then divide $C[m]$ into several subsequences. Then binary processing is carried out for each subsequence. After that, calculate the mean and variance of the adjacent peaks of $C[m]$. Get rid of the mean which has a big variance, then the average of the remaining mean values can be used as the estimation of cardiac cycle T. With T as a reference, we can segment the PCG signals using the following steps:

(1) Set the segmentation threshold $Th = \beta * \text{mean(u)}$, where β is a parameter and in this paper $\beta = 0.9$. Based on the threshold Th, all the rising and falling segmentation points $Tr = [Tr_1, Tr_2, \cdots, Tr_m]$ and $Tn = [Tn_1, Tn_2, \cdots, Tn_m]$ of the envelope u can be extracted. The area enclosed by the envelope u between each pair of rising and falling segmentation points is calculated. If the area is too small, the corresponding segmentation points are discarded.

(2) Extract the peak points $P_i, i = 1, 2, \cdots, m-1$ between each pair of Tr_i and Tn_i. Calculate the time intervals $P_{i+2} - P_i, i = 1, 2, \cdots, m$ sequentially, if $P_{i+2} - P_i \ll T$, then discard P_i. If $P_{i+2} - P_i \gg T$, then reduce Th, repeat the step (1). After adjusting peak points, we calculate the time difference between the remaining peaks $P_{j+1} - P_j, j = 1, 2, \cdots, M-1$, and recognize the systolic or diastolic duration based on the fact that the systolic duration is smaller than diastolic duration.

2.4 Feature extraction and classification

After segmentation, time features can be extracted from the PCG signals. In this paper, the time features contain that the heart rate hr, the ratio of S1 duration to heart rate S1h/hr, the ratio of S2 duration to heart rate S2h/h, the ratio of diastole to systole D/S. Beside these time features, we also extract the features in frequency domain and wavelet domain. The 4th-order Daubechies wavelet has been used in this paper, and the energies of the 4 level detail wavelet coefficients and approximate coefficients are extracted as the wavelet domain features of PCG signals, that is e_1, e_2, e_3, e_4, e_5. Meanwhile, we can divide PCG signals into three frequency bands: high frequency, medium frequency and low frequency, and calculate the energy proportions between each frequency band and the total energy, which are Eh, Em, El. In addition, calculate the maximum power spectral density

Maxf and the corresponding point Maxp. All these features of each PCG signal make up a 14-dimensional feature vector

$$[hr, S1h/hr, S2h/hr, D/S, e_1, e_2, e_3, e_4, e_5, \text{Maxp}, \text{Maxf}, \text{Eh}, \text{Em}, \text{El}].$$

Support vector machine (SVM) has been mostly used in literature for binary classification problem via non-probabilistic supervised learning model, therefore we use the SVM classifier with linear kernel to recognize the normal and abnormal PCG signals in this paper.

3. Performance evaluation

In this section, we first give two artificial signals to test the effectiveness of the proposed envelope extraction model since the true envelopes of the two signals are known. Then this envelope model is used to segment the PCG signals to extract features and the classification results on public datasets are also given. All experiments in this paper are carried out using MATLAB R2017b on a 1.8GHz machine with 16GB RAM.

3.1 Effectiveness test of the proposed envelope model

We give two artificial signals to test our envelope model. The first signal is as follows:

$$x_1(t) = t\sin(t^2), t \in [1, 10]. \tag{5}$$

Obviously, $x_1(t)$ is an IMF whose envelope is $\rho_1(t) = t, t \in [1, 10]$. The second signal is

$$x_2(t) = (2 + \cos t)\sin(t^2), t \in [5, 15]. \tag{6}$$

$x_2(t)$ is also an IMF and its envelope is $\rho_2(t) = 2 + cost, t \in [5, 15]$. We can use the proposed envelope model to obtain the estimated envelopes of $x_1(t)$ and $x_2(t)$, which are recorded as $u_1(t)$ and $u_2(t)$, respectively. Fig. 4 plots the results of the comparison between the true envelopes and the estimated envelopes. The two subfigures in bottom of Fig. 4 give the error curve $u_i(t) - \rho_i(t), i = 1, 2$. It can be seen that the approximate accuracy reaches more than 10^{-3} which shows the performance of the proposed model. In addition, to measure quantitatively the error between two signals $u(t)$ and $\rho(t)$, we use the following index named as root mean square (RMS):

$$\text{RMS}(u, \rho) = \sqrt{\frac{1}{N}\sum_{\text{i}=1}^{N} |u(t_{\text{i}}) - \rho(t_{\text{i}})|^2}. \tag{7}$$

It is obvious that the smaller the RMS value, the closer the two signals are. The RMS values between the estimated envelopes and the true envelops of $x_1(t)$ and $x_2(t)$ are 1.3663×10^{-4} and 3.4764×10^{-4}, respectively, which also shows the performance of the proposed model.

3.2 PCG Classification

3.2.1. Data description

In this study, we use two datasets to test the effectiveness of the proposed model. Dataset I comes from the database A of the classifying heart sounds PASCAL challenge competition [5], which contains 31 normal and 34 murmurs with a 441000 Hz sampling frequency. These signals all were gathered from the general public via the iStethoscope Pro iPhone app These signals are of varying lengths, between 1 second and 30 seconds. Dataset II comes from the database A of the competition in 2016 [12], which contains 117 normal and 292 abnormal PCG recordings with 2000 Hz sampling frequency. Similar with Dataset I, the signals in Dataset II is varied from 5 seconds to 120 seconds. As the reference [21], the signals with length less than 7 seconds were excluded. As a result, Dataset I contains 26 normal and 30 murmurs, and Dataset II contains 116 normal and 290 abnormal PCG signals in this study.

3.2.2. Experimental results

In this subsection, classification experiments are performed. For evaluation, we adopt the well-known performance metrics such as sensitivity, specificity and accuracy. Many state-of-the-art methods for PCG classification also employ these metrics. The definitions of these metrics are given below:

$$S_e = \frac{TP}{TP + FN}, \tag{8}$$

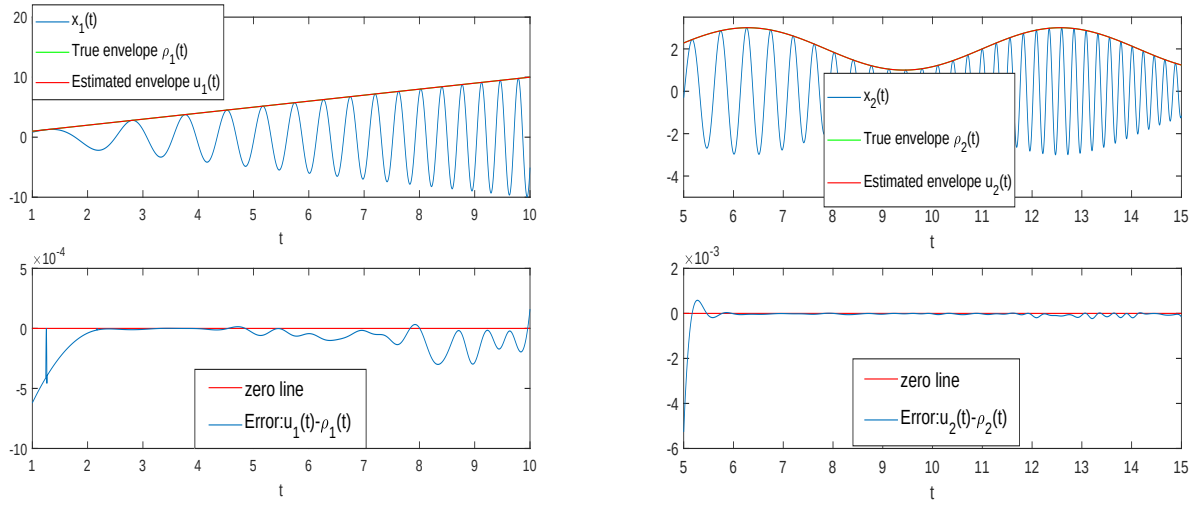

Figure 4: Left: the signal $x_1(t)$, the true envelope $\rho_1(t)$, the estimated envelope $u_1(t)$ and the approximate error $u_1(t) - \rho_1(t)$; Right: the signal $x_2(t)$, the true envelope $\rho_2(t)$, the estimated envelope $u_2(t)$ and the approximate error $u_2(t) - \rho_2(t)$.

$$S_p = \frac{TN}{FP + TN}, \tag{9}$$

$$AC = \frac{TP + TN}{TP + FP + FN + TN}, \tag{10}$$

where S_e is the sensitivity, S_p is the specificity, AC is the classification accuracy, TP and TN represent the total number of detected true positives and true negatives, respectively. FP is false positive while FN is false negative. Furthermore, a receiver operating characteristic (ROC) curve is a comprehensive index reflecting the continuous variables of the sensitivity and specificity. On the ROC curve, the point closest to the upper left of the coordinate graph is the critical value with high sensitivity and specificity. Therefore, we use ROC curves to compare the performance of our model with others. The area under the curve (AUC) of ROC is also suitable to binary classification problem. The larger the area under the curve, the higher the detection accuracy.

Envelope based method is one of important ways used for this task, in which Hilbert envelope and Shannon energy envelope are the most commonly used. Fig. 5 shows the comparison of different envelope based methods for classification. In Fig. 5, we give the ROC curves and AUC values based on different envelope methods. The green line and blue line plot the ROC curves by using Hilbert envelope and Shannon envelope, respectively. They both extract the time features from PCG signals directly without using EMD method first. It can be seen that the classification accuracy based on Shannon envelope is higher than that based on Hilbert envelope. Therefore, we choose Shannon envelope plus EMD process to compare with the proposed model. Firstly, we apply EMD process on the each PCG signal and choose the second IMF of it for further anlysis. Secondly, we segment these IMFs to estimate the cardiac cycles and then extract the time features. Thirdly, the time features combined with Fourier features and wavelet energy features are fed to SVM for classification. In Fig. 5, the yellow line represents the ROC curve based on Shannon envelope plus EMD process and the red one represents the ROC curve and AUC based on the proposed model. Although compared with using Shannon envelope alone, the accuracy based Shannon envelope plus EMD process is improved and reaches 89%. It still is lower than the result based on the proposed model which classification accuracy reaches 96.67%.

In addition, we give the classification results of 10-fold cross-validation on the two datasets. The details are shown in Table 1. It can be seen that the sensitivities of each fold iteration of Dateset I and the specificities of each fold interation of Dataset II reach the superior value 1. In both Dataset I and Dataset II, the average accuracies for PCG classification are more than 96%. Specifically, the average accuracy on Dataset I is 96.67% and the average accuracy on Dataset II is 99.51%.

We also compare the classification results based on the proposed model with the latest results for PCG classification, which is shown in Table 2. Time-frequency domain features are used in [47]. Meanwhile, its time-domain features contain the electrocardio recordings. The data used in [47] comes from two sources. One is the

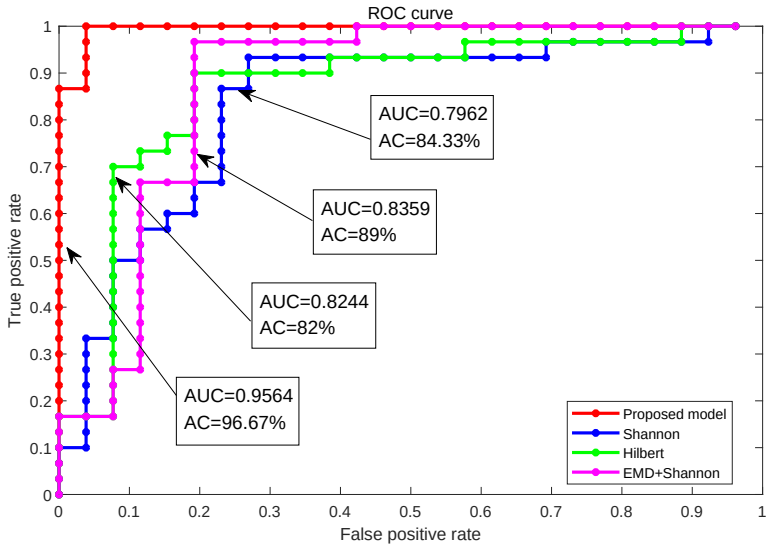

Figure 5: ROC curves comparison of different methods.

Table 1: Results of 10-fold cross validation on two datasets.

fold iteration	Dataset I			Dataset II		
	S_e	S_p	AC	S_e	S_p	AC
1	1	0.6667	0.8333	1	1	1
2	1	1	1	0.9167	1	0.9756
3	1	1	1	1	1	1
4	1	0.6667	0.8333	1	1	1
5	1	1	1	1	1	1
6	1	1	1	1	1	1
7	1	1	1	0.9167	1	0.9756
8	1	1	1	1	1	1
9	1	1	1	1	1	1
10	1	1	1	1	1	1
mean	**1**	**0.9333**	**0.9667**	**0.9833**	**1**	**0.9951**

Table 2: Comparison of the proposed method with existing methods(%).

	year	Features	Classifier	S_e	S_p	AC	
Zhao[47]	2017	Time-frequency domain	SVM	92.31	100	96.08	
Zhang[45]	2017	Based on tensor decomposition	SVM	-	-	90	
Han[22]	2018	MFCC map	CNN	98.33	84.67	-	
Noman[31]	2018	Raw (norm-dur) & MFCC	Ensemble CNN	89.94	86.35	89.22	
Our method	2019	Time-frequency domain & wavelet domain	SVM	89.83	95.63	**96.67**	(Dataset I)
				98.33	100	**99.51**	(Dataset II)

database A of [5], the other is collected by themselves. They choose the PCG signals which lengths are all more than 10 seconds to the final classification. Finally, they reached 96.08% classification accuracy. The authors in [45] proposed a new method to extract more discriminative features by using scaled spectrogram and tensor decomposition. They used the data from [12] and SVM classifier. The final accuracy for PCG classification is 90%. [22] and [31] both use Mel-frequency cepstrum coefficients (MFCC) for feature extraction and convolutional neural network (CNN) as the classifier. They also used the data from [12]. In [31], the authors combined 1D-CNN and 2D-CNN for classification and reached 89.22% accuracy. However, [31] only used 1D-CNN and reached the sensitivity of 98.33%, the specificity of 84.67%. Based on the proposed model, we achieve the classification accuracy of 96.67% on Dataset I and 99.51% on Dataset II. However, because the experiments in these studies are based on different datasets, it is difficult to directly compare the results in many cases.

4. Conclusion and future work

The main contribution of this paper is to propose a novel model of PCG envelope extraction. The model is based on a constraint optimization and EMD technique. After using EMD process to each PCG signal, the second IMF is chosen for further analysis as the representation of the corresponding PCG signal. Based on the proposed envelope model and EMD process, the cardiac cycles of PCG signals can be estimated and then the time-domain features can be extracted. Combining with the frequency-domain features and wavelet-domain features, the feature vectors of PCG signals are obtained. Finally, SVM classifier with a linear kernel is used to recognize the normal and abnormal PCG signals. The proposed famework is evaluated on the public datasets from the PASCAL classifying heart sounds challenge and 2016 PhysioNet challenge. The classification accuracies on two datasets are both above 96%, which show the effectiveness of the proposed model. In future work, we will study a fast iterative algorithm to improve the speed of the envelope extraction model, which make the big data classification available. Meanwhile, the real-time performance of the segmentation process of PCG signals will also be taken into consideration.

Acknowledgements

This research was partially supported by NSFC (No.11701144, 11471101), Higher Education School Young Backbone Teacher Training Program of Henan Province (No.2017GGJS020), Key Project of Science and Technology Research of Education Department of Henan Province (No.16A120002), NSF of Henan Province (No.162300410061).

References

[1] M. Ali, E. El-Dahshan, A. Yahia, Denoising of Heart Sound Signals Using Discrete Wavelet Transform, Circuits System Signal Processing 36(2017)4482-4497.

[2] S. Ari, K. Hembram, G. Saha, Detection of cardiac abnormality from PCG signal using LMS based least square SVM classifier, Expert Systems with Applications, 37(12)(2010)8019-8026.

[3] S. Ari, G. Saha, Classification of heart sounds using empirical mode decomposition based features, International Journal of Medical Engineering and Informatics, 1(2008)91-108.

[4] C. Balili, M. Sobrepena, P. Naval, Classification of heart sounds using discrete and continuous wavelet transform and random forests, 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR), DOI: 10.1109/ACPR.2015.7486584.

[5] P. Bentley, G. Nordehn, M. Coimbra, S. Mannor, The PASCAL Classifying Heart Sounds Challenge 2011 (CHSC2011) Results, [Online] http://www.peterjbentley.com/heartchallenge/index.html.

[6] F. Chakir, A. Jilbab, C. Nacir, A. Hammouch, Phonocardiogram signals processing approach for PASCAL Classifying Heart Sounds Challenge, Signal Image and Video Processing, 12(6)(2018)1149-1155.

[7] T. Chakrabarti, S. Saha, S. Roy, I. Chel, Phonocardiogram signal analysis - practices, trends and challenges: a critical review, 2015 International Conference and Workshop on Computing and Communication (IEMCON), Vancouver, BC, Canda, 2015.

[8] T. Chen, S. Yang, L. Ho, ea al., S1 and S2 heart sound recognition using deep neural networks, IEEE Transactions on Biomedical Engineering, 64(2)(2017)372-380.

[9] W. Chen, R Gao, L. Liu, M. Zhu, et. al., Summary of China cardiovascular disease report 2017, Chinese Circulation Journal, 33(1)(2018)1-8.

[10] S. Choi, Z.Jiang, Comparison of envelope extraction algorithms for cardiac sound signal segmentation, Expert Systems with Applications 34(2008)1056-1069.

[11] E. Comak, A. Arslan, $\dot{I}$. Türkoğlu, A decision support system based on support vector machines for diagnosis of the heart valve diseases, Computers in Biology and Medicine, 37(2007)21-27.

[12] Classification of Normal/Abnormal Heart Sound Recordings, the PhysioNet/Computing in Cardiology Challenge (2016), [Online] https://www.physionet.org/challenge/2016/#challenge-data.

[13] S. Deng, J. Han, Towards heart sound classification without segmentation via autocorrelation feature and diffusion maps, Future Generation Computer Systems, 60(2016)13-21.

[14] N. Dewangan, R. Potda, Noise cancellation using adaptive filter for PCG signal, International Journal of Emerging Trends of Technology in Computer Science, 3(4)(2014)38-43.

[15] Z. Gao, W. Hau, M. Lu, W. Huang, H. Zhang, W. Wu, X. Liu, Y, Zhang, An automated framework for detecting lumen and media-adventitia borders in intravascular ultrasound images, Ultrasound in Medicine & Biology, 41(7)(2015)2001-2021.

[16] Z. Gao, H. Xiong, X. Liu, H. Zhang, D. Ghista, W. Wu, S. Li, Robust estimation of carotid artery wall motion using the elasticity-based state-space approach, Medical Image Analysis, 37(2017)1-21.

[17] Z. Gao, Y. Li, Y. Sun, J. Yang, H. Xiong, H. Zhang, X. Liu, W. Wu, D. Liang, S. Li, Motion Tracking of the Carotid Artery Wall From Ultrasound Image Sequences: a Nonlinear State-Space Approach, IEEE Transactions on Medical Imaging, 37(1)(2018)273-283.

[18] Z. Gao, H. Zhang, D. Wang, M. Guo, H. Liu, L. Zhuang, P. Shi, Robust recovery of myocardial kinematics using dual $\mathcal{H}_\infty$ criteria, Multimedia Tools and Applications, 77(17)(2018)23043-23071.

[19] X. Guo, H. Lin, S. Xiao, Medical parameters extraction of heart sounds, Computer Engineering and Applications, 47(3)(2011)214-217.

[20] G. Güraksin, H. Uğuz, Classification of heart sounds based on the least squares support vector machine, International Journal of Innovative Computing, Information and Control, 7(12)(2011)7131-7144.

[21] M. Hamidi, H. Ghassemian, M. Imani, Classification of heart sound signal using curve fitting and fractal dimension, Biomedical Signal Processing and Control, 39(2018)351-359.

[22] W. Han, Z. Yang, J. LU and S. Xie, Supervised threshold-based heart sound classification algorithm, Physiological Measurement, 39(2018)115011.

[23] Y. He, Q. Qin, V. Josef, A pedestrian detection method using SVM and CNN multistage classification, Journal of Information Hiding and Multimedia Signal Processing, 9(1)(2018)51-60.

[24] N. E. Huang, Z. Shen, S. R. Long, et al., The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, Proceedings of the Royal Society of London A 454(1971)(1998)903-995.

[25] Z. Jiang, S. Choi, A cardiac sound characteristic waveform method for in-home heart disorder monitoring with electric stethoscope, Expert Systems with Applications 31(2006)286-298.

[26] "Cardiovascular diseases (CVDs): Fact Sheets, details" World Health Organizationk, 17 May 2017. [Online]. Available: http://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds).

[27] I. Maglogiannis, E. Loukis, E. Zafiropoulos, et al., Support Vectors Machine-based indentification of heart valve diseases using heart sounds, Computer Methods and Programs in Biomedicine, 95(1)(2009)47-61.

[28] A. Moukadem et al., A robust heart sounds segmentation module based on S-transform, Biomedical Signal Processing and Control, 8(2013)273-281.

[29] Q. Mubarak, M. Akram, A. Shaukat, F. Hussain, S. Khawaja, W. Butt, Analysisi of PCG signals using quality assessment and homomorphic filters for localization and classificaiton of heart sounds, Computer Methods and Programs in Biomedicine, 164(2018)143-157.

[30] K. Nam, J. Ahn, Y. Hwang, G. Jeon, D. Jang, I. Kim, A phonocardiogram-based noise-robust real-time heart rate monitoring algorithm for outpatients during normal activities, Journal of Mechanics in Medicine and Biology, 18(5)(2018)1850044.

[31] F. Noman, C. Ting, S. Salleh and H. Ombao, Short-segment heart sound classification using an ensemble of deep convolutional neural networks, arXiv:1810.11573v1 [cs.SD] 27 Oct 2018.

[32] J. Pan, Q. Feng, L. Yan, J. Yang, Neighborhood feature line segment for image classification, IEEE Transactions on Circuits and Systems for Video Technology, 25(3)(2015)387-398.

[33] C. Papadaniil, L. Hadjileontiadis, Efficient heart sound segmentation and extraction using ensemble empirical mode decomposition and kurtosisi feature, IEEE Journal of Biomedical and Health Informatics, 18(4)(2014)1138-1152.

[34] D. Pham, S. Meignen, N. Dia, J. Fontecave-Jallon, Phonocardiogram signal denoising based on nonnegative matrix factorization and adaptive contour representation computation, IEEE Signal Processing Letters, 25(10)(2018)1475-1479.

[35] X. Quan, J. Seok, D. Bae, Detection of S1/S2 components with extraction of murmurs from phonocardiogram, IEICE Transactions on Information and Systems, E98-D(3)(2015)745-748.

[36] J. Rubin, R. Abreu, A. Ganguli, et al., Classifying heart sound recordings using deep convolutional neural networks and mel-frequency cepstral coefficients, Computing in Cardiology, 43(2016)813-816. (DOI: 10.23919/CIC.2016.7868867)

[37] F. Safara, S. Doraisamy, A. Azman, et al., Multi-level basis selection of wavelet packet decomposition tree for heart sound classification, Computers in Biology and Medicine, 43(10)(2013)1407-1414.

[38] A. Salman, N. Ahmadi, R. Mengko, A. Langi and T. Mengko, Empirical Mode Decomposition (EMD) Based Denoising Method for Heart Sound Signal and Its Performance Analysis, International Journal of Electrical and Computer Engineering, 6(5)(2016)1-8.

[39] Y. Soeta, Y. Bito, Detection of features of prosthetic cardiac valve sound by spectrogram analysis, Applied Acoustics, 89(2015)28-33.

[40] D. Springer, L. Tarassenko, G. Clifford, Logistic regression-HSMM-based Heart Sound Segmentation, IEEE Transactions on Biomedical Engineering, 63(4)(2016)822-832.

[41] I. Turkoglu, A. Arslan, E. Ilkay, An expert system for diagnosis of the heart valve diseases, Expert Systems with Applications, 23(2002)229-236.

[42] W. Varghees, K. Ramachandran, Effective heart sound segmentation and murmure classification using empirical wavelet transform and instantaneous phase of electronic stethoscope, IEEE Sensors Journal, 17(12)(2017)3861-3872.

[43] Z. Yan, Z. Jiang, A. Miyamoto, Y. Wei, The moment segmentation analysis of heart sound pattern, Computer Methods and Programs in Biomedicine 98(2010)140-150.

[44] L. Yang, Z. yang, F. Zhou, L. Yang, A novel envelope model based on convex constrained optimization, Digital Signal Processing, 29(2014)138–146.

[45] W. Zhang, J. Han, S. Deng, Heart sound classification based on scaled spectrogram and tensor decomposition, Expert Systems With Applications, 84(2017)220-231.

[46] W. Zhang, J. Han, S. Deng, Heart sound classification based on scaled spectrogram and partial least squares regression, Biomedical Signal Processing and Control, 32(2017)20-28.

[47] Z. Zhao, X. Zhang, Z. Fang, X. Chen, L. Du, T. Li, Phonocardiogram segmentation and abnormal phonocardiogram screening algorithm based on cardiac cycle estimation, Journal of Electronics & Information Technology, 39(2017)2677-2683.

[48] Y. Zheng, X. Guo, H. Jiang, B. Zhou, An innovative multi-level singular value decomposition and compressed sensing based framework for noise removal from heart sounds, Biomedical Signal Processing and Control, 38(2017)34-43.