Classification of Heart Sound Signals Based on AR Model

Runnan He [1], Henggui Zhang [1,2], Kuanquan Wang [1], Qince Li [1], Zhiqiang Sheng [1], Na Zhao [1]

[1] School of Computer Science and Technology, Harbin Institute of Technology, Harbin, China
[2] School of Physics and Astronomy, University of Manchester, Manchester, United Kingdom

Abstract

Heart sounds reflect information of the mechanical contraction of the heart in both of the physiological and pathological conditions. It is important to develop novel numerical algorithms to characterize the features of the heart sound as a helpful diagnostic tool of cardiovascular diseases. This study aims to develop an efficient algorithm for analyzing heart sound signals that can be used for cardiovascular disease monitoring. In the algorithms, wavelet analysis (coif5) with 5 decomposition levels was first applied to heart sound signals for noise eliminating by using a soft fixed threshold. Then, heart sound signals were decomposed by the wavelet method to reconstruct bands with different frequencies. Following this, the normalized Shannon energy of each frequency band within the same time duration was calculated to determine the position of the second heart sound (S2). Finally, the aortic valve closure (A2) of the S2 were extracted using the power spectrum analysis of Auto Regressive(AR) model, which were used to classify the normal and abnormal heart sound recordings. Results show that the modified Sensitivity (Se), Specificity (Sp) and overall score are respectively 0.87, 0.61, and 0.74.

1. Introduction

In a cardiac cycle, the electrical activity is firstly generated by the cardiac pacemaker, which then triggers atrial and ventricular contractions. This in turn pumps blood flowing between the chambers of the heart and around the body. The opening and closing behaviours of the heart valves are associated with acceleration and deceleration of blood, giving rise to vibrations of the entire cardiac structure and thus producing the heart sounds and murmurs [1]. These vibrations are audible at the chest wall, and the heart sounds can reflect the health condition of the heart. The phonocardiogram (PCG) is the graphical representation of a heart sound recording.

Fundamental heart sounds (FHSs) usually consist of two components: the first (S1) and second (S2) heart sounds. Although the FHSs are the most recognizable sounds in the heart cycle, the mechanical activity of the heart may also cause other audible sounds, such as the third heart sound (S3), the fourth heart sound (S4), systolic ejection click (EC), mid-systolic click (MC), diastolic sound or opening snap (OS), as well as heart murmurs caused by the turbulent, high-velocity flow of blood.

The segmentation of the FHSs is the first step in the automatic analysis of heart sounds. Previously several segmentation methods have been developed in the literatures [2-5].

The automated classification of pathology based on heart sound recordings has been performed for over 50 years. However, massive challenges still remain now. Gerbarg *et al.* were the first group to attempt the automatic classification of pathology in PCGs using a threshold-based method [6], motivated by the need to identify children with rheumatic heart disease (RHD). Artificial neural networks (ANNs) have been the most widely used machine learning-based approach for heart sound classification. Typical relevant studies used different signal features as the input to the ANN classifier, including wavelet features [7], time, frequency and complexity-based features [8], and time-frequency features [9]. A number of researchers have also applied support vector machines (SVM) for heart sound classification in recent years. The studies can also be divided into different groups according to the feature extraction methods, including wavelet [10], time, frequency and time-frequency feature-based classifiers [11]. Hidden Markov models (HMM) have also been employed for pathology classification in PCG recordings [5]. Clustering-based classifiers, typically the k-nearest neighbours (KNN) algorithm [12], have also been employed to classify pathology in PGCs. In addition, many other techniques have been applied, including threshold-based methods, decision trees [13] and discriminant function analysis [14].

In our study, we classified the heart sound signals based on AR model with wavelet denoising and normalized Shannon energy. Results showed that the modified Sensitivity (Se), Specificity (Sp) and overall score are respectively 0.87, 0.61, and 0.74.

2. Method

Figure 1 outlines the flow of our approach. The architecture of our proposed algorithm included signal filtering， extraction of S2 using wavelet decomposition and normalized Shannon energy, extraction of S2 features using AR model, and classification of heart sound signals. Each major step was explained in more detail in the four upcoming subsections.

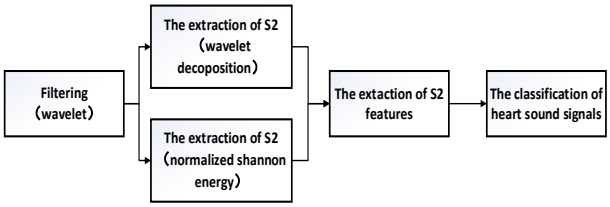

Figure 1. Flow chart of our approach.

2.1. Filtering

For signal denoising, each original heart sound signal was decomposed by muti-level discrete wavelet, which means an input PCG signal was divided into low (ai) and high frequency (di) components and then the low frequency component was input into the next layer for further decomposition. In this study, wavelet analysis (coif5) with 5 decomposition levels was first applied to heart sound signals for noise eliminating by using a soft fixed threshold. Then, heart sound signals were decomposed by the wavelet method to reconstruct bands with different frequencies. Figure 2 shows the result of denoised heart sound signal.

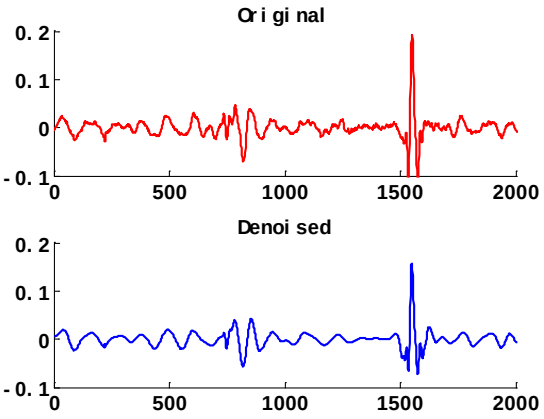

Figure 2. Denoised heart sound signal compared to original signal

2.2. The extraction of S2

2.2.1. Decompose the denoised PCG signal using wavelet decomposition and reconstruct the details and approximations

In order to identify S2s correctly, frequency band of the majority power of S2 was used. Previous studies using FFT to analyze the frequency contents of the second heart sounds have indicated that the frequency spectrum of S2 contained peaks in a low-frequency range (10 to 80Hz), a medium-frequency range (80 to 200Hz) and a high-frequency range (220 to 400Hz) [1].

Before decomposition, the denoised PCG signal was down sampled by factor 5. Some murmurs have higher frequencies than the normal sounds (up to 600Hz) [15]. Since these frequencies are still below a half of the sampling frequency 2000Hz, useful events of the heart sounds were not missed. After down sampling, a fifth-level discrete wavelet decomposition of the denoised signal was done to obtain the coefficients of all the components of the decomposition. Using these coefficients, the details and approximations in desired level were obtained by reconstruction. The details and approximations varied depending on the wavelet families and orders used in the decomposition and reconstruction. According to the characteristics of the frequency spectrum of S2 and the possible noises, details d4, d5, and approximation a4 were selected as the sources for segmentation.

2.2.2. Calculate the normalized average Shannon energy for selected approximations and details

For each selected signal, d4, d5 and a4 were segmented using a segmentation algorithm based on the envelope calculated from the normalized average Shannon energy. The average Shannon energy attenuates the effects of low value noise and makes the low intensity sounds easier to be found. The average Shannon energy was calculated in 0.02-second continuous segments throughout the normalized signal with 0.01-second overlap using the following formula:

$$E_s = -\frac{1}{N \cdot \sum_{i=1}^{N} x_{norm}^2(i) \cdot \log x_{norm}^2(i)} \quad (1)$$

where the x_{norm} is the signal sample normalized to the maximum absolute value of the studied band signal, and N is the number of samples in 0.02-second segment, here N=40.

Lastly the normalized average Shannon energy versus the whole time axis was computed as follow,

$$P_a(t) = \frac{E_s(t) - M\left(E_s(t)\right)}{S\left(E_s(t)\right)} \quad (2)$$

Where $M\left(E_s(t)\right)$ is the mean value of $E_s(t)$, $S\left(E_s(t)\right)$ is the standard deviation of $E_s(t)$.

Figure 3 shows the normalized average Shannon energy for selected approximations and details.

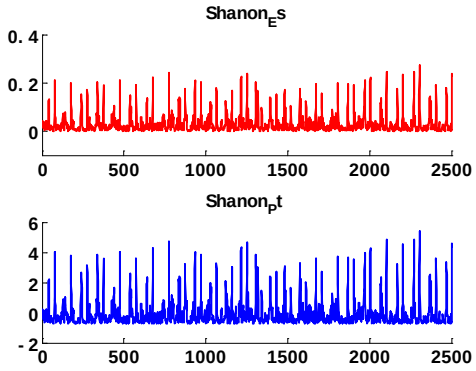

Figure 3. The normalized average Shannon energy for selected approximations and details

2.2.3. Mark the peak location of S1s and S2s

The actual heart sound recordings are very complicated and patterns of heart sounds vary significantly from recording to recording. It is difficult to pick up all the S1s and S2s using a simple threshold. There might be undesired peaks due to the second part of splitted S2 or other events. In addition, the first heart sounds usually may be too weak compared with other peaks. Moreover, artifacts similar to the real peaks both in duration and amplitude might be selected as S1s or S2s. In order to solve these problems, the threshold setting and detection rules of picking S1s and S2s were modified. Firstly, simple threshold was used to mark all the peak locations of continuous segments exceeding the threshold limit. Then time intervals between two adjacent marks were calculated. According to the mean value and standard variation of the intervals, both lower and higher time interval limits were calculated, which were used to remove extra peaks and find lost weaker peaks.

2.2.4. Decide the durations of S2s

S1s and S2s should be distinguished from each other after they are marked. Here, the identification based on the following two facts: the longest time interval between two adjacent peaks in the recording (within 20 seconds) is the diastolic period and the duration of the systolic period is relatively constant compared to the diastolic one. After the longest time interval was found, the starting and the ending marks of that interval were set as S2 and S1 respectively. Then the intervals were checked forward and backward from the longest interval on. Those marks which destroyed constancy limitations of systolic and diastolic period were discarded and the rest Sls and S2s were identified. The artifacts were discarded in this deciding procedure.

The detected positions of intensity peaks of S1s and S2s indicated the approximate locations of these sounds. The actual boundaries of these sounds were obtained by defining another lower intensity threshold value, which differed from the one for detecting S1 and S2. The boundaries of S1s and S2s were modified by confining their duration within 20ms to 150ms respectively. Finally, the systolic and diastolic periods were decided as the small transition time intervals before and after S1 and S2. Figure 4 shows the durations of S2s.

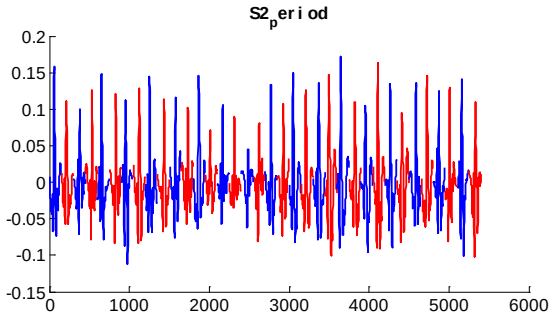

Figure 4. The durations of S2s

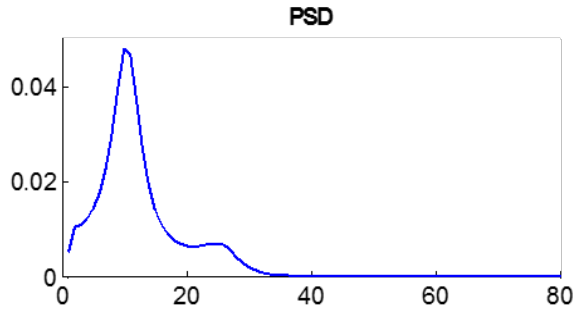

Figure 5. The PSD of S2

2.3. The extraction of S2 features

Because the recording location and physical condition of patients have great influences on the amplitude of heart sound signals, it is difficult to compare the energy differences between heart sound signals. However, the frequency characteristics of the heart sound signal is relatively stable, which can be chosen as the parameters of S2 features in according to the physical characteristics. For S2, it occurs at the beginning of diastole with the closure of the aortic and pulmonic valves, so we select the aortic valve closure (A2) as a feature of S2. The primary

peak was found using AR model to estimate the power spectral density (PSD) of heart sound signals. Figure 5 shows the PSD of S2.

2.4. The classification of heart sound signals

In this challenge, we focused on the classification of heart sound signals. In order to diagnose the heart diseases, we used A2 to classify the normal and abnormal heart sound signals, the parameters threshold setting is selected by experiment experience.

3. Results

Based on the method described above, the classification in test set of PhysioNet database were estimated.

In this study, the modified Sensitivity (Se), Specificity (Sp) and overall score are respectively 0.87, 0.61, and 0.74 in the Event-1. Although the accuracy of classification is not high, the running time is fast which can satisfy the real-time classification. The average running time for training set is 5.33% and the maximum running time is 5.65%. Meanwhile, the average running time for test set is 5.33% and the maximum running time is 5.59%.

4. Discussion

We applied wavelet analysis to heart sound signals for noise eliminating. Through the comparison and analysis, wavelet analysis (coif5) with 5 decomposition levels was first applied to heart sound signals for noise eliminations by using a soft fixed threshold.

It is notable that thresholds for classifying heart sound signals were tightly correlated with more tests on different types of data. In future studies, this method can be further improved by introducing new algorithms for the extraction of S2 features.

It should also be noted that in our approach only the main A2 frequency have been extracted, which is better than extracting more features in terms of expediting calculation speed. In the further work, new approaches like Hilbert transform should be added to improve our method for heart sound signals analysis.

Acknowledgements

This study was supported in part by the National Natural Science Foundation of China (NSFC) under Grant No. 61572152 and No. 61571165.

References

[1] Leatham A. Auscultation of the heart and phonocardiography. 1975

[2] Papadaniil CD, Hadjileontiadis LJ. Efficient heart sound segmentation and extraction using ensemble empirical mode decomposition and kurtosis features. *Biomedical and Health Informatics, IEEE Journal of*. 2014;18:1138-1152

[3] Tang H, Li T, Qiu T, Park Y. Segmentation of heart sounds based on dynamic clustering. *Biomedical Signal Processing and Control*. 2012;7:509-516

[4] Springer D, Tarassenko L, Clifford G. Logistic regression-hsmm-based heart sound segmentation. 2015

[5] Springer DB, Tarassenko L, Clifford GD. Support vector machine hidden semi-markov model-based heart sound segmentation. *Computing in Cardiology Conference (CinC), 2014*. 2014:625-628

[6] Gerbarg DS, Taranta A, Spagnuolo M, Hofler JJ. Computer analysis of phonocardiograms. *Progress in Cardiovascular Diseases*. 1963;5:393-405

[7] Liang H. A feature extraction algorithm based on wavelet packet decomposition for heart sound signals. *Time-Frequency and Time-Scale Analysis, 1998. Proceedings of the IEEE-SP International Symposium on*. 1998:93-96

[8] Schmidt SE, Græbe M, Toft E, Struijk JJ. No evidence of nonlinear or chaotic behavior of cardiovascular murmurs. *Biomedical Signal Processing and Control*. 2011;6:157-163

[9] De Vos JP, Blanckenberg MM. Automated pediatric cardiac auscultation. *Biomedical Engineering, IEEE Transactions on*. 2007;54:244-252

[10] Ari S, Hembram K, Saha G. Detection of cardiac abnormality from pcg signal using lms based least square svm classifier. *Expert Systems with Applications*. 2010;37:8019-8026

[11] Maglogiannis I, Loukis E, Zafiropoulos E, Stasis A. Support vectors machine-based identification of heart valve diseases using heart sounds. *Computer methods and programs in biomedicine*. 2009;95:47-61

[12] Bentley P, Grant P, McDonnell J. Time-frequency and time-scale techniques for the classification of native and bioprosthetic heart valve sounds. *IEEE transactions on bio-medical engineering*. 1998;45:125-128

[13] Pavlopoulos SA, Stasis AC, Loukis EN. A decision tree–based method for the differential diagnosis of aortic stenosis from mitral regurgitation using heart sounds. *Biomedical engineering online*. 2004;3:21

[14] Schmidt SE, Holst-Hansen C, Hansen J, Toft E, Struijk JJ. Acoustic features for the identification of coronary artery disease. *Biomedical Engineering, IEEE Transactions on*. 2015;62:2611-2619

[15] Rangayyan RM, Lehner RJ. Phonocardiogram signal analysis: A review. *Critical reviews in biomedical engineering*. 1986;15:211-236

Address for correspondence.

Henggui Zhang - E-mail:henggui.zhang@manchester.ac.uk
Room 3.07, Shuster building
The School of Physics and Astronomy
The University of Manchester, Oxford Road
Manchester, M13 9PL, UK