Biomedical Signal Processing and Control 32 (2017) 20–28

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Heart sound classification based on scaled spectrogram and partial least squares regression

Wenjie Zhang [a], Jiqing Han [a,*], Shiwen Deng [b,*]

[a] School of Computer Science and Technology, Harbin Institute of Technology, Harbin, China
[b] School of Mathematical Sciences, Harbin Normal University, Harbin, China

ARTICLE INFO

Article history:
Received 30 May 2016
Received in revised form 9 September 2016
Accepted 13 October 2016

Keywords:
Scaled spectrogram
Heart sound
Partial least squares regression

ABSTRACT

Phonocardiogram (PCG) signal analysis is an effective and convenient method for the preliminary diagnosis of heart disease. In this study, a scaled spectrogram and partial least squares regression (PLSR) based method was proposed for the classification of PCG signals. Proposed method is mainly comprised of four stages, namely as being heart cycle estimation, spectrogram scaling, dimension reduction and classification. At the heart cycle estimation stage, the short time average magnitude difference of the Shannon energy envelope is applied. Then the spectrogram of the obtained heart cycle is calculated for feature extraction. However, the sizes of the spectrograms between different PCG signals are usually not the same. In order to overcome the difficulty of direct comparison, the bilinear interpolation is used for the spectrogram to get the scaled spectrogram with a fixed size. Nevertheless, the scaled spectrogram contains a large quantity of redundant and irrelevant information. To extract the most relevant features from the scaled spectrogram, we adopt the PLSR to reduce the dimension of the scaled spectrograms. Since PLSR has the advantage of using the category information during the dimension reduction process, the extracted features are more discriminative. Then the classification results are obtained via support vector machine (SVM). The proposed method is evaluated on two public datasets offered by the PASCAL classifying heart sounds challenge, and the results are compared to those obtained using the best methods in the challenge, thereby proving the effectiveness of our method.

1. Introduction

Many pathological conditions of the cardiovascular system are reflected in heart sound signals, which makes it possible to diagnose heart disease by analyzing heart sound signals. Heart sound auscultation is a method used to analyze heart sound signals using a stethoscope. Because of its easy implementation, auscultation is widely used in the clinical diagnosis of heart disease [1,2]. However, the accuracy of auscultation depends on the skill and subjective experience of the physician [3]. Therefore, an objective analysis of heart sound signals is necessary. PCG signal analysis is another method of analyzing heart sound signals using phonocardiograms. The physiological and pathological information has been extracted from the PCG signal using signal processing and artificial intelligence techniques in the literature [3,4]. With the PCG, the objective analysis of heart sound signals using computer technology is becoming popular. Moreover, telemedicine is becoming available with the development of electronic stethoscopes and smart phones [5]. Overall, the analysis of PCG signals has important significance for the diagnosis of heart disease. Heart sound classification aims at the automatic classification of PCG signals. It is very important for preliminary diagnosis.

Heart sound classification usually involves two steps. The first step is heart sound segmentation, which attempts to detect the location of the fundamental heart sounds (FHs). The FHs include the first (S1) and second (S2) heart sounds, which are the important physical characteristics of heart sounds. The accurate localization of the FHs shows the systolic and diastolic regions of the heart sounds. In addition, the heart cycles are identified by the FHs. Thus, the characteristics of different pathological situations in the region of one heart cycle are used to classify different heart sound categories. Many methods, such as the envelope-based method [6], the method using dynamic clustering [7] and the logistic regression-hsmm based method [8], have been developed for this task. However, heart sound segmentation remains a challenging task, and it is difficult to segment the FHs accurately in a noisy environment.

* Corresponding authors.
E-mail addresses: jqhan@hit.edu.cn (J. Han), dengswen@gmail.com (S. Deng).

http://dx.doi.org/10.1016/j.bspc.2016.10.004
1746-8094/

The second step of heart sound classification is to extract the features in one heart cycle and use the features for classification. Many features have been proposed in the literature. The three main types are time [5], frequency [9] and time-frequency complexity-based features [10,11]. Although the time-frequency-based features are more computationally complex than features based on only time or frequency, they provide more comprehensive information about the PCG signal. Thus, time-frequency-based features usually outperform other features. The commonly used time-frequency feature extraction methods for PCG signals are wavelets [10], S-transform [12] and short time Fourier transform (STFT) [13]. The magnitude of the STFT yields the spectrogram. This spectrogram is used in this paper since it is easy to implement and convenient to scale.

The primary goal of heart sound classification is to identify different heart sound categories. This is not necessary for segmentation in some situations, especially when the heart cycles are known. So the estimation of heart cycle duration and alignment methods based on the envelope are proposed to obtain the heart cycles instead of locating both S1 and S2. The calculation process is simplified in this way. Although the correct segmentation information can improve the classification performance, it requires a lot of computing. More importantly, the segmentation results are not correct in many cases which greatly affect the accuracy of the classification.

The spectrogram is extracted for each heart cycle after the heart cycles are estimated. However, the sizes of the spectrograms are different since the heart rates of different PCG signals are usually not the same. This prohibits a direct comparison between the spectrograms of different PCG signals. A bilinear interpolation [14] method is used to scale the size of the spectrogram, thus enabling the direct comparison. Nevertheless, the scaled spectrogram contains a large quantity of redundant and irrelevant information. In order to extract the most relevant information, a dimension reduction process of the scaled spectrogram is adopted. In addition, the heart sound category provides valuable information to distinguish between different categories and it helps to improve the classification performance. As a result, the extracted features will be more discriminative if the category information is fully utilized during the dimension reduction process. PLSR [15] maximizes the correlation between the PCG signals and their corresponding category information during the dimension reduction process. Thus the category information is utilized. Also, PLSR is capable to robustly handle more descriptor variables than the number of samples. These are the advantage of PLSR compared with other dimensionality reduction method, such as principle component analysis (PCA) [16], linear discriminant analysis (LDA) [17]. Therefore, the discriminative features of the scaled spectrogram are extracted using PLSR in this paper. Finally, the classification is performed using the SVM classifier [11].

The main framework of this paper is shown in Fig. 1 and consists of four steps: estimation of heart cycle duration and alignment, spectrogram scaling of each heart cycle, PLSR and classification. PLSR consists of two parts, i.e., dimension reduction and regression. The contributions of this paper are threefold. First, the heart cycles are estimated and aligned instead of locating both S1 and S2 to simplify the calculation process. Second, the spectrograms of heart cycles of different lengths can be compared directly using the bilinear scaling process which has not been applied in heart sound researches to our knowledge. Third, the category information is utilized during the dimension reduction process. In this way, the extracted features are best correlated with their categories in the dimension reduction process which makes the features more discriminative.

Table 1
The number of samples in the training and testing dataset.

Dataset	Category	Training	Testing
Dataset-A	*Normal*	31	14
	Murmur	34	14
	Extra Heart Sound	19	8
	Artifact	40	16
Dataset-B	*Normal*	200	136
	Murmur	66	39
	Extrasystole	46	20

2. Method

2.1. Data collection

The datasets used in this paper, including Dataset-A and Dataset-B [18], are collected from the classifying heart sounds Pascal challenge competition. Dataset-A is collected by volunteers using iStethoscope which is an iPhone application that enables an iPhone to use its microphone as a digital stethoscope [19]. Dataset-A includes 176 records with a 44,100 Hz sampling frequency and it can be grouped into four categories: *Normal*, *Murmur*, *Extra Heart Sound* and *Artifact*. A normal heart sound has a clear lub dub, lub dub pattern, with the time from lub to dub shorter than the time from dub to the next lub [20]. In the *Murmur* category, the heart murmurs sound as though there is a whooshing, roaring, rumbling, or turbulent fluid noise in one of two temporal locations: (1) between lub and dub, or (2) between dub and lub [20]. A regular additional sound can be identified as an extra heart sound. A wide range of different sounds are contained in the *Artifact* category, including speech, music and noise. With the approval of the RHP Ethics Committee, Dataset-B was collected at the Real Hospital Portugues using a Littmann Model 3100 electronic stethoscope with a 4000 Hz sampling frequency. Dataset-B includes 507 records grouped into three categories: *Normal*, *Murmur* and *Extrasystole*. The *Extrasystole* heart sound is not the same as the extra heart sound in Dataset-A because the additional sound is not regularly occurring. Besides, no information is available on the auscultated subjects, such as gender, age, and condition. The number of samples used in the training and testing dataset is shown in Table 1.

2.2. Preprocessing

The collected PCG signals are often contaminated with high frequency noise. The main information of the PCG signals is concentrated at low frequencies. Therefore, the PCG signals are resampled to 2000 Hz before further processing. Additionally, the resampled signals are filtered with a band-pass (50–950 Hz), 6th-order Butterworth filter to further eliminate the noise. Then, the PCG signals are normalized to a fixed scale of $[-1\ 1]$:

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

where $x'[n]$ is the resampled and filtered signal and $x[n]$ is the normalized signal. A murmur PCG signal after filtering and normalization is shown in Fig. 2 (a).

2.3. Estimation of heart cycle duration and alignment

The classification task is based on each heart cycle. First, the normalized signal $x[n]$ is decomposed and reconstructed using the discrete wavelet transform to accurately estimate the heart cycle. Because the Daubechies wavelet is morphologically similar to the PCG signal, it is widely used to decompose the PCG