Expert Systems With Applications 84 (2017) 220–231

Contents lists available at ScienceDirect

Expert Systems With Applications

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

Heart sound classification based on scaled spectrogram and tensor decomposition

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 November 2016
Revised 10 April 2017
Accepted 5 May 2017
Available online 8 May 2017

Keywords:
Heart sound
Scaled spectrogram
Tensor decomposition

ABSTRACT

Heart sound signal analysis is an effective and convenient method for the preliminary diagnosis of heart disease. However, automatic heart sound classification is still a challenging problem which mainly reflected in heart sound segmentation and feature extraction from the corresponding segmentation results. In order to extract more discriminative features for heart sound classification, a scaled spectrogram and tensor decomposition based method was proposed in this study. In the proposed method, the spectrograms of the detected heart cycles are first scaled to a fixed size. Then a dimension reduction process of the scaled spectrograms is performed to extract the most discriminative features. During the dimension reduction process, the intrinsic structure of the scaled spectrograms, which contains important physiological and pathological information of the heart sound signals, is extracted using tensor decomposition method. As a result, the extracted features are more discriminative. Finally, the classification task is completed by support vector machine (SVM). Moreover, the proposed method is evaluated on three public datasets offered by the PASCAL classifying heart sounds challenge and 2016 PhysioNet challenge. The results show that the proposed method is competitive.

1. Introduction

Many cardiac abnormalities are reflected in heart sound signals, which makes it possible to diagnose heart disease by analysing heart sound signals. Heart sound auscultation is a frequently used method to analysis heart sound signals using a stethoscope. As auscultation is convenient to implementation, it is widely used in the clinical diagnosis of heart disease (Hanna & Silverman, 2002; Rangayyan & Lehner, 1986). However, accurate auscultation needs a long physician experience, which is not easy to obtain (Jiang & Choi, 2006). Therefore, a computer assist tool for heart sound analysis is needed to help diagnose heart disease. Phonocardiogram (PCG) signal analysis is another method of analysing 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 literatures (Herzig, Bickel, Eitan, & Intrator, 2015; Jiang & Choi, 2006). With the PCG, the automatic analysis of heart sound signals using computer technology is becoming popular. Moreover, the telemedicine is becoming available with the development of electronic stethoscopes and smart phones (Deng & Han, 2016). 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 to preliminary diagnosis.

Heart sound classification usually involves three steps. The first step is heart sound segmentation. It attempts to detect the location of the fundamental heart sounds (FHs), including the first (S1) and second (S2) heart sounds, which are the important physiological characteristics of heart sounds. According to the accurate localization of the FHs, the systolic and diastolic regions of the heart sounds are detected. In addition, the heart cycles are also identified by FHs. Many methods have been developed for heart sound segmentation, such as, the amplitude threshold based methods (Chen, Kuan, Celi, & Clifford, 2010; Liang, Lukkarinen, & Hartimo, 1997; Moukadem, Dieterlen, Hueber, & Brandt, 2013; Sun, Jiang, Wang, & Fang, 2014) and probabilistic models based methods (Schmidt, Holst-Hansen, Graff, Toft, & Struijk, 2010; Springer, Tarassenko, & Clifford, 2015). However, heart sound segmentation remains a challenging task, and it is difficult to segment the FHs accurately in a noisy environment.

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

http://dx.doi.org/10.1016/j.eswa.2017.05.014
0957-4174/

Table 1
Summary of the heart sound feature extraction methods.

Author	Features	Advantages	Disadvantages
(Ari et al., 2010)	Time	Easy to calculate and quantify the features.	Physiological and pathological information is not fully extracted.
(Safara et al., 2013)	Frequency		
(Ari et al., 2010) (Deng & Han, 2016) (Soeta & Bito, 2015) (Gokhale, 2016) (Her & Chiu, 2016) (Teo et al., 2016) (Thomae & Dominik, 2016) (Zhang et al., 2017)	Time-frequency	Discriminative features are extracted automatically according to demand, more physiological and pathological information is available.	Discriminative features are hard to extract.

The second step is feature extraction. It attempts to extract features for classification based on the segmentation results. Many features have been proposed in the literatures. The three main types are time (Ari, Hembram, & Saha, 2010), frequency (Safara, Doraisamy, Azman, Jantan, & Ramaiah, 2013) and time-frequency complexity-based features (Ari et al., 2010; Deng & Han, 2016; Maglogiannis, Loukis, Zafiropoulos, & Stasis, 2009). Although the time or frequency based features are easy to understand and calculate according to the physiological characteristics of the heart sound signals, such as the heart rate, the systolic and diastolic period, they may lose some important pathological information which is not easy to quantify in the time or frequency domain independently, such as the shape of S2 in the time-frequency domain. In order to improve the performance of heart sound classification, extracting more discriminative features from the time-frequency domain is becoming more and more popular in recent research (Gokhale, 2016; Her & Chiu, 2016; Teo, Yang, Feng, & Su, 2016; Thomae & Dominik, 2016; Zhang, Han, & Deng, 2017). Moreover, the PCG signals need to be represented in the transformed time-frequency domain to extract features in these methods. The commonly used transform methods for PCG signals are wavelets (Ari et al., 2010), S-transform (Livanos, Ranganathan, & Jiang, 2000; Moukadem et al., 2013) and short time Fourier transform (STFT) (Soeta & Bito, 2015). Since the spectrogram calculated by STFT is easy to implement and convenient to scale, it is used in this paper. For a better comparison, the feature extraction methods and their advantages and disadvantages are summarized in Table 1. As shown in Table 1, the heart sound feature extraction is still a challenging task which is caused by the non-stationary and diversity of the PCG signals.

The last step is classification. It aims at choosing a suitable classifier to evaluate the extracted features and complete the classification task. The commonly used methods in heart sound classification are k-Nearest Neighbors (Quiceno-Manrique, Godino-Llorente, Blanco-Velasco, & Castellanos-Dominguez, 2010), Hidden Markov Model (SaraçOğLu, 2012), Artificial Neural Network (Uğuz, 2012), decision trees (Deng & Bentley, 2012) and SVM (Zheng, Guo, & Ding, 2015). The classifier is generally chosen according to the features extracted from the heart sound signal.

The primary goal of heart sound classification is to identify different heart sound categories. It is not necessary for the detailed segmentation in some situations, especially when the heart cycles are known (Avendano-Valencia, Godino-Llorente, Blanco-Velasco, & Castellanos-Dominguez, 2010). Therefore, the heart sound signal is segmented into heart cycles instead of explicitly S1 and S2, as we did in the previous research (Zhang et al., 2017). In this way, the physiological and pathological information in the heart cycles is reserved and the segmentation is performed in a more efficient approach. Then, the spectrogram is extracted from each heart cycle after the heart cycles are detected. 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. To overcome the problem, the bilinear interpolation (Hariharan, Arbeláez, Girshick, & Malik, 2015) method is used to scale the size of the spectrogram, thus the direct comparison is enabled. 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. The commonly used method of dimension reduction is the principal component analysis (PCA) (Jolliffe, 2002) method. In order to perform the PCA, the spectrogram needs to be expanded into a vector form. Thus, during the dimension reduction process, the intrinsic structure information in the time-frequency domain is lost. However, the intrinsic structure of the scaled spectrograms, such as the relative position between the S1 and S2 in the time-frequency domain, contains the physiological and pathological information of the heart sound signal, which is important for heart sound classification. To keep the intrinsic structure of the scaled spectrograms during the dimension reduction process, Avendano-Valencia et al. (2010) proposed the 2D-PCA method. Whereas, the time or frequency information is extracted independently in the 2D-PCA method. As a result, the intrinsic structure is not fully extracted. In order to extract the relevant intrinsic structure information more efficiently during the dimension reduction process, the Tucker-2 tensor decomposition (Kolda & Bader, 2009) method is used in our study for its ability to keep the intrinsic structure information. Moreover, the extracted features are competitive compared with the features calculated from the literatures by evaluation. Finally, the classification is performed using the SVM classifier.

The main framework of this paper is shown in Fig. 1 and consists of three steps: heart cycle segmentation, feature extraction, including spectrogram scaling and tensor decomposition, and classification. The main contribution of this paper is that we proposed a more efficient method to extract the intrinsic structure of the scaled heart cycle spectrograms during the dimension reduction process using tensor decomposition. Thus, more useful physiological and pathological information is reserved during the dimension reduction process. As a result, the extracted features are more discriminative. Moreover, the extracted features are adopted for heart sound classification which can provide a primary diagnosis in the primary health center and home care. In addition, the tensor decomposition method has not been applied for heart sound spectrogram analysis in the literatures to our knowledge.