Biomedical Signal Processing and Control 7 (2012) 509–516

Contents lists available at SciVerse ScienceDirect

Biomedical Signal Processing and Control

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

Segmentation of heart sounds based on dynamic clustering

Hong Tang [a,*], Ting Li [a,b], Tianshuang Qiu [a], Yongwan Park [c]

[a] *Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian 116024, China*
[b] *The College of Information and Communication Engineering, Dalian Nationalities University, Dalian 116600, China*
[c] *The Department of Information and Communication, Yeungnam University, Kyongsan-city, Kyongbuk 712-749, Republic of Korea*

ARTICLE INFO

Article history:
Received 21 April 2011
Received in revised form
20 September 2011
Accepted 20 September 2011
Available online 4 November 2011

Keywords:
Heart sound
Segmentation
Dynamic clustering
Density function
Cycle detection

ABSTRACT

The heart sound signal is first separated into cycles, where the cycle detection is based on an instantaneous cycle frequency. The heart sound data of one cardiac cycle can be decomposed into a number of atoms characterized by timing delay, frequency, amplitude, time width and phase. To segment heart sounds, we made a hypothesis that the atoms of a heart sound congregate as a cluster in time–frequency domains. We propose an atom density function to indicate clusters. To suppress clusters of murmurs and noise, weighted density function by atom energy is further proposed to improve the segmentation of heart sounds. Therefore, heart sounds are indicated by the hybrid analysis of clustering and medical knowledge. The segmentation scheme is automatic and no reference signal is needed. Twenty-six subjects, including 3 normal and 23 abnormal subjects, were tested for heart sound signals in various clinical cases. Our statistics show that the segmentation was successful for signals collected from normal subjects and patients with moderate murmurs.

1. Introduction

Phonocardiography is a noninvasive, easy, and efficient method to evaluate heart functions. It is widely used in medical check-ups for adults and children due to its easy operation and economical cost. During auscultation, cardiologists try to separate a single cycle of a heart sound signal into four phases (i.e., first heart sounds (S1), systolic phase, second heart sounds (S2), and the diastolic phase), and then analyze these features. However, heart sound signals are transient and fastly varying. Analysis of heart sound signals by human ears may be inconsistent due to their high dependence on the listener's skills and experience. Therefore, it would be desirable to analyze these recordings in an automated and computerized manner. Segmentation of heart sounds is generally an important pre-process for the automated analysis of a heart sound signal.

Previous studies of segmentation have been reported on the instantaneous energy of an ECG used to estimate the presence of S1 and S2 [1]. Heart sound signals were segmented based on their time-domain characteristics as well [2,3]. The frequency-domain-based segmentation algorithm was proposed by tracking of heart sound spectrum [4]. With the reference ECG, the performance of the heart sound segmentation is quite good. However, the ECG, which is another signal source, may not be convenient for use in a medical check-up. The electronic stethoscope may become too complex if an ECG circuit is embedded into it. Thus, segmentation algorithms using a heart sound signal as a sole source are desirable [1,7,8,18]. Liang et al. [5] segmented S1 and S2 by picking up the peaks of the heart sounds' envelope. However, the envelope is often distorted by noise. Special peak identification should be carried out to select peaks of heart sounds. On the other hand, S1 or S2 may be missed if the amplitude is too small. To segment S1 and S2, Vepa et al. [6] used energy- and simplicity-based features computed from multi-level wavelet decomposition coefficients. The segmentation method proposed by Ari et al. [7–9] needed only the average heart rate or a rectangular window with proper width as the auxiliary information; however, little information was provided on how the average heart rate was obtained. Although high frequency signatures of the heart sounds envelope can be employed to segment S1 and S2, some murmurs also have high frequency signatures [10]. The ergodic hidden Markov model for classification of a continuous heart sound signal has previously been proposed [11], wherein heart sound signals are represented by four separated states: S1, systolic phase, S2, diastolic phase. The method seemed to be advantageous; however, a time-consuming training process was needed. Further interaction between a computer and a human being was needed during the training process as well.

The objective of this paper is to develop an automated, robust algorithm to segment heart sounds. The features of heart sounds in time, frequency and medical domains are fully exploited. The heart sound signal of a cardiac cycle can be decomposed into a number of atoms. Heart sounds are segmented based on the hybrid analysis of clustering and medical knowledge. A wide range of heart sound

* Corresponding author. Fax: +86 411 84709573.
E-mail address: tanghong@dlut.edu.cn (H. Tang).

1746-8094/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bspc.2011.09.002

signals collected from 3 normal and 23 abnormal subjects were used to test the proposed algorithm.

This paper is organized as follows: the signal representation in time–frequency domains is described in Section 2. The segmentation based on hybrid analysis of clustering and medical knowledge is proposed in Section 3. Computer simulations are shown in Section 4. Discussions on cycle detection and robustness to noise are given in Section 5. To finish, Section 6 provides conclusions.

2. Signal representation in time–frequency domains

2.1. Signal decomposition

To our knowledge, several models have been used to decompose heart sounds into time–frequency domains, such as the chirp model [12,13], damped sinusoidal model [14,15], modified Prony model [16], and Gaussian modulation model. Leung et al. [17] employed the Gaussian modulation model to decompose the second heart sound for diagnosis of pediatric heart disease. This model, in particular, is used in this paper to synthesize the heart sound signal of one cardiac cycle. The model is

$$h(t) = \sum_{i=1}^{L} a_i e^{-(t-t_i)^2/(2\sigma_i^2)} \cos(2\pi\omega_i t + \beta_i), \quad (1)$$

where $h(t)$ is the heart sound signal of one cycle. Eq. (1) shows that $h(t)$ represents the sum of L components. Here, the components will be referred to as atoms. Each atom is characterized by five parameters. t_i is the timing delay of the ith atom with respect to the start of this cycle; a_i is the amplitude; ω_i is the frequency; σ_i controls the time width to support the atom; β_i is the phase. Therefore, the heart sound signal of this cycle can be expressed by the set of atoms $\{t_i, \omega_i, a_i, \sigma_i, \beta_i, 1 \le i \le L\}$. The number of atoms, L, and the parameters for the atoms, can be interpreted by the short-time Fourier transform (STFT) analysis [17].

The definition of STFT is

$$H(t, f) = \int h(t) w(t-\tau) e^{-2\pi f \tau} d\tau, \quad (2)$$

where $w(t)$ is a window. In comparison with (1), $w(t)$ represents a Gaussian window. From the absolute magnitude of the STFT, the estimation of the parameters begins with the identification of the atom with maximum amplitude. The atom with maximum amplitude is located by detecting the peak in the magnitude of the STFT. Once the timing delay, t_i, has been identified, its corresponding a_i, ω_i and β_i can be read directly from the STFT. σ_i is obtained by the following optimal procedure. The waveform represented by the ith atom is

$$s_i(t, \sigma_i) = a_i e^{-(t-t_i)^2/(2\sigma_i^2)} \cos(2\pi\omega_i t + \beta_i). \quad (3)$$

The residue after the waveform is subtracted from the signal, $h_{i-1}(t)$, is

$$h_i(t, \sigma_i) = h_{i-1}(t) - s_i(t, \sigma_i). \quad (4)$$

Note that $h_0(t)$ is the original signal of a cardiac cycle. The normalized residue energy is

$$\rho_i(\sigma_i) = \frac{\int |h_i(t, \sigma_i)|^2 dt}{\int |h_0(t)|^2 dt}. \quad (5)$$

Obviously, $\rho_i(\sigma_i)$ will be a minimum if a perfect match occurs in (4). $\rho_i(\sigma_i)$ is thus used as a criterion to optimize σ_i. To achieve this, we can monitor $\rho_i(\sigma_i)$ by varying σ_i in a predefined range, similar to grid searching. The decomposition stops if $\rho_i(\sigma_i)$ is small enough.

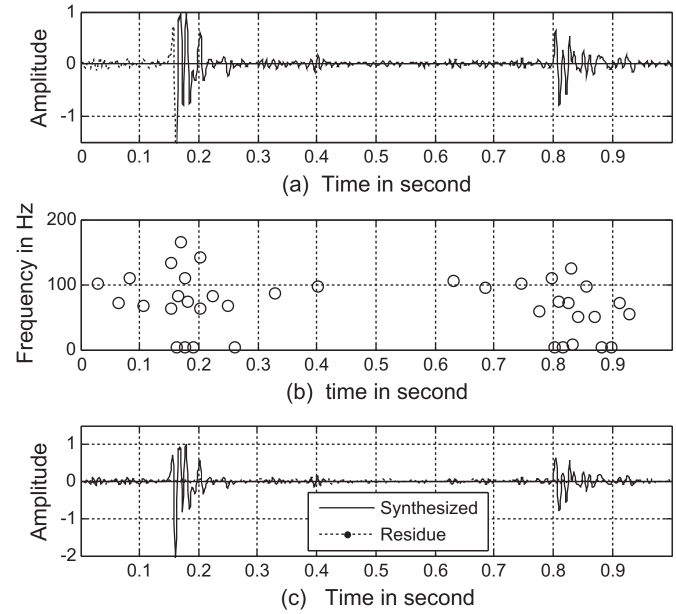

Fig. 1. An example of signal decomposition. (a) The signal of one cardiac cycle. (b) Atoms in a time–frequency plane. (c) The reconstructed signal and residue.

The heart sound signal of the cycle can be reconstructed by the sum of the waveforms

$$h(t) \approx \sum_{i=1}^{L} s_i(t). \quad (6)$$

2.2. Data acquisition

The normal heart sound signals used in this paper were collected in the authors' laboratory. Subjects were lying on his back in an examination bed and kept under stable conditions. The sensor was placed on the mitral site. The ECG and heart sound signal were synchronously recorded. It has been demonstrated that the dominant frequency band of heart sounds is generally limited to 600 Hz. The sampling rate was set as 2 kHz, which is higher than the minimum rate suggested by the sampling theory.

2.3. An example of signal decomposition

The heart sound recording is first separated into cycles by a cycle detection method previously described in Section 5.1. A representative signal of one cardiac cycle is decomposed into 38 atoms in Fig. 1. The atoms, in particular, are shown in a time–frequency plane in Fig. 1(b). The synthesized waveform is given in Fig. 1(c) where the decomposition terminates when the normalized residual energy is smaller than 0.05. We found that the features of the normal heart sound signal in time and frequency domains are perfectly represented by the atoms.

3. Segmentation of heart sounds

3.1. Features of heart sounds and murmurs in a time–frequency plane

Hypothesis. The atoms of heart sounds congregate in high density in a time–frequency plane. However, the atoms of murmurs are dispersed in the plane.

It has been previously established that S1 and S2 have a relatively high energy within a short time duration. Atoms of S1