

# Automatic Heart Sound Segmentation Method Based on Cyclostationarity and Clustering

Ting Li

College of Information and Communication Engineering, Dalian Nationality University, Liaoning  
Dalian, 116600, China

email: liting@dlnu.edu.cn

**Keywords:** Heart sound signal; Segmentation; Cyclostationary; Clustering

**Abstract.** Segmenting heart sounds accurately is significant to determine the types and the severity of heart diseases. Using cyclostationary property of heart sounds and clustering method, the paper proposed an effective automatic segmentation algorithm of heart sound signals. This algorithm can segment normal heart sounds and abnormal heart sounds with medium murmur exactly. Simulation results show that this method is robust to noise, which is helpful to feature extraction, heart sound recognition and computer-aided diagnosis afterwards.

## Introduction

Phonocardiography is a noninvasive, easy, and efficient method to evaluate heart functions. Heart auscultation is widely used in medical check-up for adults and children due to its easy operation and economy. In auscultation, physicians or cardiologists try to separate heart sound signal into four phases, i.e., first heart sounds (S1), systolic phase, second heart sounds (S2), diastolic phase, and then analyze the features. However, heart sounds are transient and fast varying. Heart sounds analysis by human ears highly depends on the skills and experience of the listeners. Therefore, it is desirable to analyze the recording of heart sounds in a computerized way. The previous studies of segmentation have been reported in literature. With the reference ECG, the performance of heart sound segmentation is quite good. However, ECG is another signal source. It is maybe discommodious for medical check-up. The electronic stethoscope may become too complex if ECG circuit is embedded in. Therefore, the segmentation algorithms using heart sounds as sole source are welcome, such as [1-5]. H. Liang et al. [1] identified S1 and S2 by picking up the peaks of heart sounds envelope. Unfortunately, the peaks of envelope are often affected by noise. On the other hand, S1 and S2 may be missed if the amplitude is too small. J Vepa et al. [2] proposed a segmentation method which used energy-based and simplicity-based features computed from multi-level wavelet decomposition coefficients to identify S1 and S2. The segmentation method proposed by S Ari et al. [3] needed only the average heart rate or a rectangular window with proper width as auxiliary information, but the way to obtain the average heart rate was not mentioned in detail. An automatic segmentation is desirable.

The objective of this paper is to develop an automatic, robust algorithm for segmenting heart sound. The features of heart sounds in time domain, frequency domain and medical domain are fully exploited. The sequence of heart sound was divided cycle by cycle first. The heart sound signal of a cardiac cycle was then decomposed into a number of atoms characterized by timing delay, frequency, amplitude, time width and phase. S1 and S2 were identified by the hybrid analysis of atom clusters and medical knowledge.

## Cycle frequency spectrum and short time cycle frequency spectrum

The cardiac cycle is a sequence of events that is repeated cyclically as a heart beats. A heart sound signal is thus quasi-cyclostationary. The fundamental concept underlying the cardiac cycle detection is to exploit the repetitive behaviors. We may estimate the fundamental cycle frequency using cyclic statistics.

## Cycle Frequency Spectrum

A second-order cyclic statistics, proposed by Gardner, to assess cyclostationarity of a cyclic signals,  $x(t)$ , at cycle frequency  $\alpha$  and frequency  $f$  is cyclic spectral autocoherece function,  $\rho_x(\alpha, f)$ , which is defined as

$$\rho_x(\alpha, f) = \frac{S_x(\alpha, f)}{[S_x(0, f + \alpha/2)S_x(0, f - \alpha/2)]^{1/2}} \quad (1)$$

$S_x(\alpha, f)$  is written as

$$S_x(\alpha, f) = \int_{-\infty}^{+\infty} R_x(\alpha, \tau) e^{-j2\pi f \tau} d\tau \quad (2)$$

$$R_x(\alpha, \tau) = \int_t x(t)x(t-\tau)e^{-j2\pi\alpha t} dt \quad (3)$$

where  $R_x(\alpha, \tau)$  is cyclic correlation function.  $R_x(\alpha, \tau)$  degrades to a traditional correlation when cycle frequency  $\alpha$  is zero.

The cyclostationarity at cycle frequency  $\alpha$  is only of interest.  $\rho_x(\alpha, f)$  is transformed into a function of  $\alpha$  by an integral

$$\gamma_x(\alpha) = \int_{-\infty}^{+\infty} \rho_x(\alpha, f) df \quad (4)$$

$\gamma_x(\alpha)$  is called cycle frequency spectrum (CFS) in this paper, which shows the cyclostationarity of signal  $x(t)$  at any cycle frequency. The fundamental cycle frequency is thus located by the first peak of  $\gamma_x(\alpha)$ .

## Short Time Cycle Frequency Spectrum

Heart beating rate varies with time. It is known as heart rate variability (HRV). This means that the fundamental cycle frequency changes with time. The equations (3) are not able to estimate the time varying cycle frequency because the overall signal is operated in one block. To track the time varying cyclic behavior, a sliding rectangle window is needed. From the nature of heartbeats, we know that a quality index does not exist for a subsequence spanning less than two cardiac cycles, where the heartbeat does not repeat. We thus conclude that the minimum length of the sliding window must be equal or greater than two cardiac cycles so that the heartbeat repeats at least once. Following the principle of joint time-frequency analysis based on short time Fourier transform, we can carry out the time cycle frequency analysis based on short time cycle frequency spectrum (STCFS). The equations (3) are modified to

$$R_x(\alpha, t) = \int_{t-\zeta}^{t+\zeta} x(\tau)x(\tau)e^{-j2\pi\alpha\tau} d\tau \quad (5)$$

where the sliding window covers  $[t-\zeta, t+\zeta]$  in time domain. We get the STCFS

$$\gamma_x(\alpha, t) = R_x(\alpha, t) / R_x(0, t) \quad (6)$$

The ICF can be found at the maximum of  $\gamma_x(\alpha, t)$  at time  $t$  in a given cycle frequency range. We define this ICF as  $\delta(t)$ . It is given by

$$\delta(t) = \arg \max_{\alpha} [\gamma_x(\alpha, t)]_{\delta(t^-)-\lambda}^{\delta(t^-)+\lambda} \quad (7)$$

where  $\lambda$  is defined as a number to limit the search range.  $\delta(t^-)$  is the cycle frequency at previous time  $t^-$ . A value of  $\lambda$  around 0.2 Hz generally yields accurate estimations. In other words, (7) is used to express the operation that the method searches for the maximum  $\gamma_x(\alpha, t)$  at each time  $t$  along the cycle frequency axis  $\alpha$  which is bounded to interval  $[\delta(t^-)-\lambda, \delta(t^-)+\lambda]$ . It can be seen that (7) is a recursive equation. The value of  $\delta(t)$  at time  $t$  is function of peak location of  $\gamma_x(\alpha, t)$  and previous  $\delta(t^-)$ . Using this recursive equation, the ICF  $\delta(t)$  at time  $t$  is forced to close to the previous one. This is reasonable because heart beating rate does not change abruptly.

## Cardiac cycles detection based on ICF

It is assumed that  $t_0$  is a starting point with which we begin to separate the heart sound signal,  $x(t)$ , to cardiac cycles. The instantaneous cardiac cycle duration at time  $t$  is the reciprocal of the

ICF, i.e.,  $1/\delta(t)$ . The cycle duration of the first cardiac cycle is estimated at time  $t_0$

$$T_1 = 1/\delta(t_0) \quad (8)$$

The first cardiac cycle is thus given by

$$c_1(t) = [x(t_0) \ x(t_0 + 1) \ \dots \ x(t_0 + T_1)] \quad (9)$$

This is reasonable because the heart beating rate does not change abruptly in short time. Similarly, the cycle duration of the  $i$ th cardiac cycle is

$$T_i = 1/\delta(t_0 + \sum_{j=1}^{i-1} T_j) \quad (10)$$

The  $i$ th cardiac cycle is

$$c_i(t) = [x(t_0 + \sum_{j=1}^{i-1} T_j) \ x(t_0 + \sum_{j=1}^{i-1} T_j + 1) \ \dots \ x(t_0 + \sum_{j=1}^i T_j)] \quad (11)$$

## Representation for heart sounds in time-frequency domain

To our knowledge, several models were used to decompose heart sounds in time-frequency domain, such as chirp model, damped sinusoidal model, modified Prony model, and Gaussian modulation model. T. S. Leung, et al [6] employed the Gaussian modulation model to decompose the second heart sound for diagnosis of pediatric heart disease. The Gaussian modulation model, in particular, is employed in this paper to synthesize the heart sounds of one cardiac cycle. The model is

$$h(t) = \sum_{i=1}^L a_i e^{-(t-t_i)^2/(2\sigma_i^2)} \cos(2\pi\omega_i t + \beta_i) \quad (12)$$

where  $h(t)$  is the heart sound signal of a cycle. The equation (15) means that  $h(t)$  is represented by the sum of  $L$  atoms. Every atom is characterized by five parameters.  $t_i$  is the timing delay of the  $i$ th atom with respect to the start of this cycle;  $a_i$  is the amplitude;  $\omega_i$  is the frequency;  $\sigma_i$  controls the time width to support the atom;  $\beta_i$  is the phase. Therefore, the heart sound signal of this cycle can be expressed by the set of atoms  $\{t_i, \omega_i, a_i, \sigma_i, \beta_i, 1 \leq i \leq L\}$ . The number of atoms,  $L$ , and the parameters for the atoms can be interpreted by the concept of short-time Fourier transform (STFT) analysis, as given in [7].

The definition of STFT is

$$H(t, f) = \int h(t) w(t - \tau) e^{-2\pi j f \tau} d\tau \quad (13)$$

where  $w(t)$  is a window. With comparison to (1),  $w(t)$  can be taken as a Gaussian window. From the absolute magnitude of the STFT, the parameter estimation 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,  $t_i$ , has been identified, its corresponding  $a_i$ ,  $\omega_i$  and  $\beta_i$  can be read directly from the STFT.  $\sigma_i$  is obtained by the following optimal procedure. The waveform represented by the  $i$ th atom is

$$s_i(t, \sigma_i) = a_i e^{-(t-t_i)^2/(2\sigma_i^2)} \cos(2\pi\omega_i t + \beta_i) \quad (14)$$

The residue is

$$h_i(t, \sigma_i) = h_{i-1}(t) - s_i(t, \sigma_i) \quad (15)$$

where  $h_i(t, \sigma_i)$  is the residue after the waveform represented by the  $i$ th atom is subtracted from the signal,  $h_{i-1}(t)$ .  $h_0(t)$  is the original signal of the  $m$ th cycle. The normalized residue energy is

$$\rho_i(\sigma_i) = \int |h_i(t, \sigma_i)|^2 dt / \int |h_0(t)|^2 dt \quad (16)$$

Obviously,  $\rho_i(\sigma_i)$  will be minimum if a perfect match occurs in (19).  $\rho_i(\sigma_i)$  is thus used as a criterion to optimize  $\sigma_i$ . To achieve this, we can monitor  $\rho_i(\sigma_i)$  by varying  $\sigma_i$  in predefined range. It is like a grid searching. The decomposition stops if  $\rho_i(\sigma_i)$  is small enough. The heart sounds for the cycle is reconstructed by the sum of the waveforms

$$h(t) \approx \sum_{i=1}^L s_i(t) \quad (17)$$

### Dynamic clustering based on density function

For a given heart sound signal of a cardiac cycle, two well-separated clusters can be found in the time-frequency plane if there are no murmurs, as shown in Fig. 1. Any clustering methods may be valid to cluster heart sounds. However, the clusters of heart sounds and murmurs may be close to each other. The cluster boundaries become fuzzy and difficult to determine. Among various classes of clustering methods, density-based methods are of special interest in our applications because they are concerned with statistical features of data. On the other hand, the density-based clustering has the advantages: (1) allowing arbitrary shape of cluster and (2) not requiring the number of clusters as input which is difficult to determine in our cases because we have no prior knowledge about murmur and noise.

We estimate the density of data by using Gaussian kernel density estimation, a nonparametric technique to estimate underlying density function (DF) from samples. The DF is defined to be

$$f(t, \omega) = \frac{1}{L} \sum_{i=1}^L \left\{ \frac{1}{2\pi\sqrt{(s_t)^2 + (s_f)^2}} \exp\left(-\left(\frac{t-t_i}{s_t}\right)^2 - \left(\frac{\omega-\omega_i}{s_f}\right)^2\right) \right\} \quad (18)$$

where  $s_t$  and  $s_f$  are scale parameters for time and frequency, respectively. Clusters of S1 and S2 then will be obtained according to the partition defined by the peaks of  $f(t, \omega)$ . However, there are so many peaks in Fig. 2. In order to make the peaks of S1 and S2 more outstanding, we define a new density function.

The waveform represented by the  $i$ th atom is  $s_i(t)$  as given in (14). The atom energy is defined as the waveform's energy

$$E_i = \int_{t_i-3\sigma_i}^{t_i+3\sigma_i} [s_i(t)]^2 dt \quad (19)$$

The  $i$ th atom has time width over  $[t_i-3\sigma_i, t_i+3\sigma_i]$  to support with Gaussian model. Obviously, the greater  $E_i$  is, the higher contribution the atom gives to the heart sound signal. In order to reveal those atoms who contribute more to the heart sound signal, we introduce a weighted density function by atoms energy, which is written as

$$f_w(t, \omega) = \frac{1}{L \sum_{i=1}^L E_i} \sum_{i=1}^L E_i \left\{ \frac{1}{2\pi\sqrt{(s_t)^2 + (s_f)^2}} \exp\left(-\left(\frac{t-t_i}{s_t}\right)^2 - \left(\frac{\omega-\omega_i}{s_f}\right)^2\right) \right\} \quad (20)$$

The cluster boundary is defined as the level set

$$u_{\text{boundary}}(t, \omega) := \{f_w(t, \omega) = th\} \quad (21)$$

where  $th$  is a small number.

### Simulation results and discussions

As an example, we apply the proposed segmentation algorithm to a normal heart sound. The data was recorded in the authors' laboratory. The subject was on his back on an examination bed and kept stable conditions. The sensor was put on mitral site. ECG and heart sounds were recorded synchronously. The heart sound signal of a cardiac cycle was indicated by ECG. The sampling rate was 2 KHz. The heart sounds are shown in Fig. 1, only first 6 cycles are shown in the figure. Single cardiac periods segmented by our algorithm are shown in Fig. 2. One cycle heart sound waveforms and the corresponding atoms are given in Fig. 3(a) and Fig. 3(b), respectively. The non-weighted density function and weighted density function is given in Fig. 3(c) and Fig. 3(d). It is obviously that the peaks of S1 and S2 are more outstanding in weighted DF and the peaks of murmur are greatly suppressed. The segmentation results are marked in Fig. 3(a) by the rectangles. Further, cluster 2 is for S1 and cluster 1 is for S2 because  $T_{12}$  is the time duration of diastolic phase.

The proposed segmentation was tested by various clinical heart sounds, including normal heart

sounds, mitral stenosis, mitral regurgitation, aortic stenosis, aortic regurgitation, pulmonary stenosis, summation gallop, third heart sound (S3), fourth heart sound (S4), respectively. The recordings of heart sounds were downloaded from web sites [8-10]. There are totally 565 cycles from 26 subjects. The heart sounds were firstly segmented in hand and then we compared the results with those of the proposed algorithm. The results indicated over 97 percent correctness in the identification of S1 and S2.

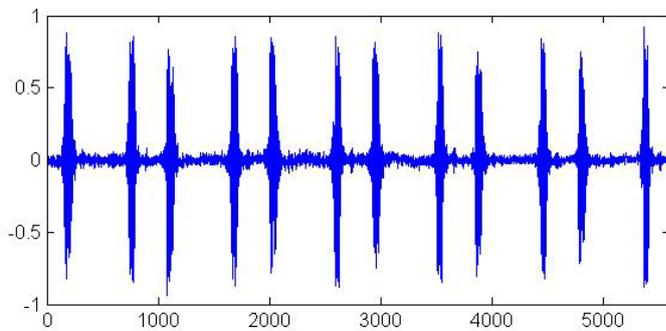


Fig. 1 A normal heart sound

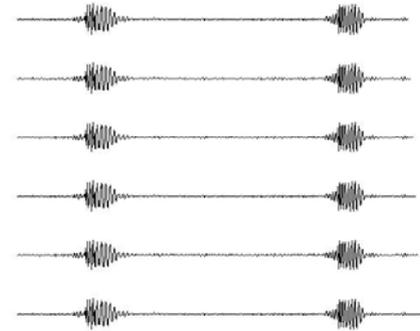
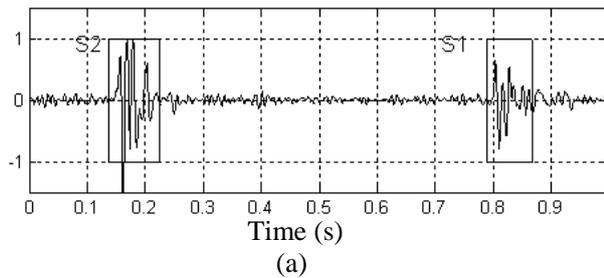
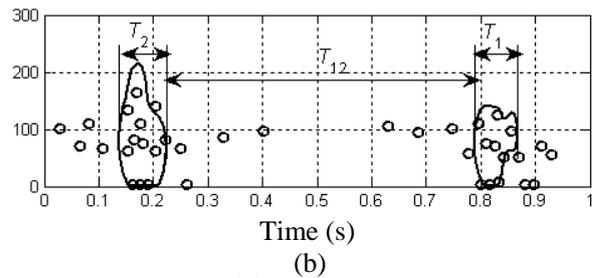


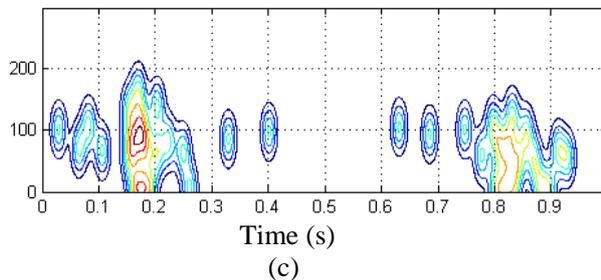
Fig. 2 Cardiac cycles detection results after compensation



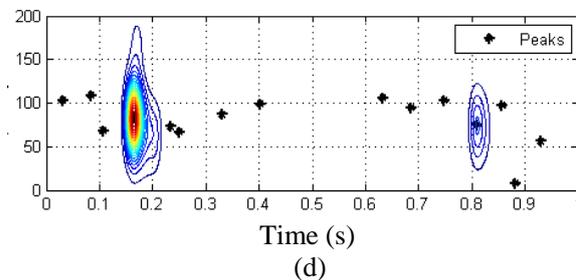
(a)



(b)



(c)



(d)

Fig. 3 Segmentation of heart sounds for a normal signal. (a) One cycle of heart sound signal. (b) The distribution of atoms on the time-frequency domain. (c) Density function. (d) Weighted density function.

## Conclusions

An algorithm to perform automatic segmentation for heart sounds was proposed in this paper. It decomposes heart sounds of a separated cycle into a number of atoms and identifies S1 and S2 based on the hybrid analysis of clustering information and medical knowledge. The algorithm was robust to additive noise (perhaps the noise is non-Gaussian and colored). Experiment results had shown over 97 percent correctness of segmentation.

## Acknowledgement

This work is supported by Fundamental Research Funds for the Central Universities (Project No. DC201501056).

## References

- [1] H. Liang, S. Lukkarinen, I. Hartimo. Heart sound segmentation algorithm based on heart sound envelopogram. *IEEE computers in cardiology*, 1997, 24, 105-108.
- [2] J. Vepa, P. Tolay, A Jain. Segmentation of heart sounds using simplicity features and timing information. *IEEE international conference on acoustics, speech, and signal processing (ICASSP)*, 2008, 469-472.
- [3] S. Ari, G. Saha. On a robust algorithm for heart sound segmentation. *Journal of mechanics in medicine and biology*, 2007, 7(2): 129-150.
- [4] D. Kumar, P. Carvalho, M. Antunes, J. henriques, L. Eugenio, R Schmidt, J. Habetha. Detection of S1 and S2 Heart sounds by high frequency signatures. *Proceedings of 28<sup>th</sup> IEEE EMBS annual international conference*, New York city, USA, Aug 30-Sept 3, 2006, 1410-1416.
- [5] Y. J. Chung. Classification of continuous heart sound signals using the ergodic hidden Markov Model. *Third Iberian conference on pattern recognition and image analysis*. Girona, Spain, June 6-8, 2007, proceedings part I, 563-570.
- [6] T. S. Leung, P. R. White, W. B. Cook, et al. Analysis of the second heart sound for diagnosis of paediatric heart disease. *IEEE proc.-Sci. Meas. Technol.*, 1998, 145(6), 285-290.
- [7] A. M. Yip, C. Ding, T. F. Chan. Dyanmic cluster formation using level set methods. *IEEE Transactions on pattern analysis and machine intelligence*, 2006, 28(6): 877-889.
- [8] Heart Sounds and Murmurs. [http://www.texasheart.org /education/cme/explore/events/eventdetail\\_5469.cfm](http://www.texasheart.org/education/cme/explore/events/eventdetail_5469.cfm). Oct 10, 2009.
- [9] Demonstrations: Heart Sounds & Murmurs. <http://depts.washington.edu/physdx/heart/demo.html>. Oct 10, 2009.
- [10] The Cardiac Exam: Auscultation. <http://filer.case.edu/dck3/heart/listen.html>. Oct 10, 2009.