

RESEARCH

Open Access



An automatic segmentation method for heart sounds

Qingshu Liu¹, Xiaomei Wu^{1,2*}  and Xiaojing Ma³

*Correspondence:

xiaomeiwu@fudan.edu.cn

¹ Department of Electronic Engineering, Fudan University, Room 522 B, Science Building, 220 Handan Rd., Shanghai, China

Full list of author information is available at the end of the article

Abstract

Background: There are two major challenges in automated heart sound analysis: segmentation and classification. An efficient segmentation is capable of providing valuable diagnostic information of patients. In addition, it is crucial for some feature-extraction based classification methods. Therefore, the segmentation of heart sound is of significant value.

Methods: This paper presents an automatic heart sound segmentation method that combines the time-domain analysis, frequency-domain analysis and time–frequency-domain analysis. Employing this method, the boundaries of heart sound components are first located, and the components are then recognized. Finally, the heart sounds are divided into several segments on the basis of the results of boundary localization and component identification.

Results: In order to evaluate the performance of the proposed method, quantitative experiments are performed on an authoritative heart sound database. The experimental results show that the boundary localization has a sensitivity (Se) of 100%, a positive predictive value (PPV) of 99.3% and an accuracy (Acc) of 99.93%. Moreover, the Se, PPV and Acc of component identification reach 98.63, 99.86 and 98.49%, respectively.

Conclusion: The proposed method shows reliable performance on the segmentation of heart sounds. Compared with previous works, this method can be applied to not only normal heart sounds, but also the sounds with S3, S4 and murmurs, thus greatly increasing the applied range.

Keywords: Heart sound segmentation, Murmur elimination, Boundary detection, Component identification

Background

Cardiovascular disease (CAD) remains the leading cause of death worldwide. In China, CAD accounted for more than 40% of deaths in 2014 [1]. Heart sounds, which are generated by the beating of heart, are considered as an important signal for detecting cardiovascular problems. In general, a normal heart sound comprises two components, namely the first heart sound (S1) and the second heart sound (S2). In some special cases, three additional components, namely the third heart sound (S3), fourth heart sound (S4), and murmurs, may appear together or separately [2]. Heart sounds have been used to diagnose cardiovascular problems for hundreds of years. Even currently, auscultation remains a crucial approach for physicians to



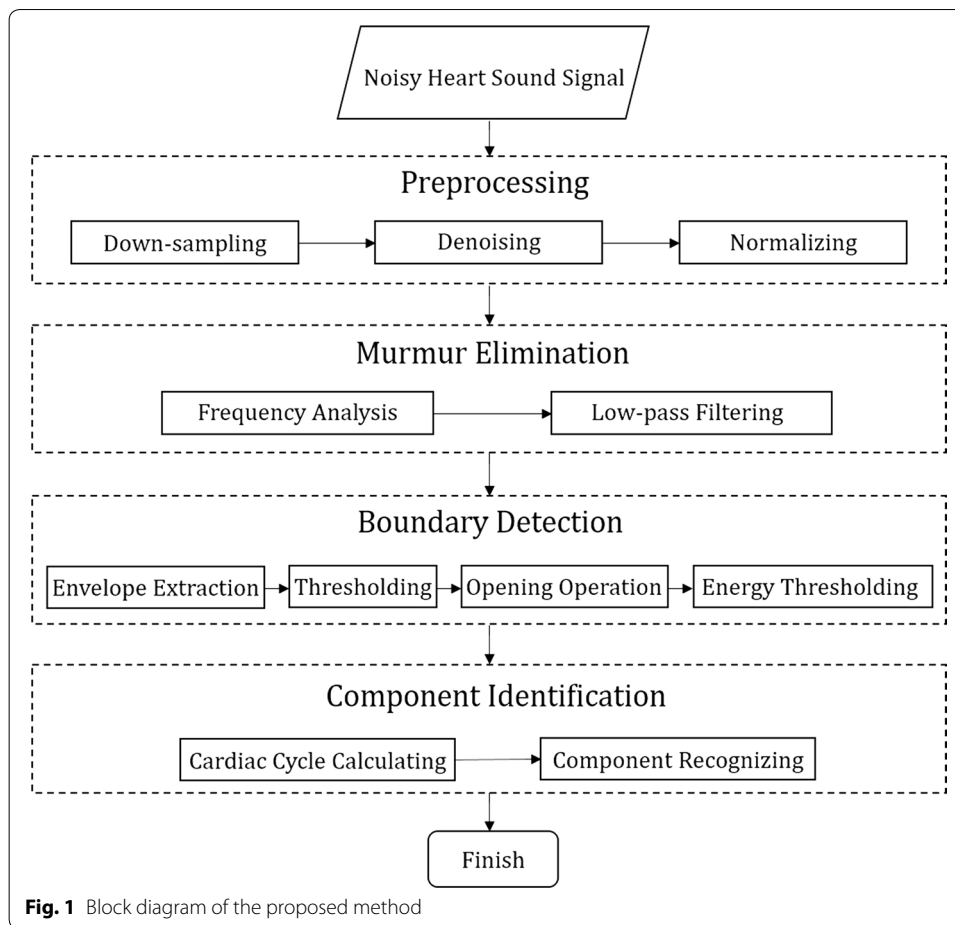
learn about patients' heart health. However, this conventional diagnostic approach is highly subjective and relies largely on physicians' experience [3]. Thus, an accurate and efficient automated heart sound analysis system is required.

There are two major challenges in developing an automated heart sound analysis tool: segmentation and classification. The primary operations of segmentation are positioning the boundaries of heart sound components and identifying the component types, namely S1, S2, S3, and S4. Several segmentation methods have been reported. However, most of the methods focus on either boundary detection or component identification. In addition, many methods only apply on normal heart sounds, thus severely limiting the application of these methods.

Liang et al. [4] proposed a Shannon-entropy-based heart sound segmentation method to recognize S1 and S2, but they did not study S3, S4, and murmurs. Kumar et al. [5] presented an S3 detection algorithm based on wavelet transform; however, this method could not detect the boundaries of the heart sound components. Moukadem et al. [6] developed an S-transform-based heart sound segmentation method; however, this method can only be applied to normal heart sounds. Springer et al. [7] successfully addressed the segmentation problem of noisy S1 and S2 by utilizing a logistic regression–hidden semi-Markov model. Nevertheless, this study, as most other studies, did not address the segmentation of heart sounds containing S3 and S4. In addition to the aforementioned studies, several other researchers have also contributed to the automated segmentation of heart sounds. For example, Naseri et al. [8] proposed a heart sound segmentation method based on the frequency-energy method, which had excellent performance for various heart sounds. Boutana et al. [9] presented a time–frequency-analysis–based heart sound segmentation method and validated the efficiency of the method through some pathological heart sounds. Tang et al. [10] reported a dynamic-clustering-based segmentation method, which was effective for both normal and abnormal heart sounds. However, these methods all have some of the aforementioned limitations.

Based on the advantages and drawbacks of the methods mentioned above, the present study is focused on developing an automated heart sound segmentation algorithm that can position the boundaries and recognize the components for both normal heart sounds and heart sounds with S3, S4 and murmurs.

In this study, a method that combines time-domain analysis, frequency-domain analysis, and time–frequency-domain analysis was developed, and some novel strategies were proposed. The complete method comprises four parts, namely preprocessing, murmur elimination, boundary detection, and component identification. The overall process is presented in Fig. 1. First, the noisy raw heart sound is standardized in the preprocessing phase. Then, the signal goes through murmur elimination algorithm, where any possible murmurs are considerably eliminated without affecting the characteristics of the heart sound signal. In the following boundary detection phase, the onset and offset of each heart sound component are positioned and marked automatically. Finally, in the component identification phase, the cardiac cycle is calculated first, and the heart sound components are then recognized in each heartbeat on the basis of cardiac cycle.



Methods

Materials

To validate the proposed method, two quantitative experiments were performed, and the heart sounds from the University of Michigan's Heart Sound & Murmur Library were employed [11]. In the first experiment, 12 types of heart sounds with variable murmurs were used to verify the effect of murmur elimination. The total length of these sounds was 759 s, and the cardiac cycle number reached 888. The second experiment was performed to quantify the effectiveness of boundary detection and component identification. For this purpose, 16 types of sounds, including 2 normal and 14 abnormal types, were used, constituting a total of 1039 s (1251 cycles). The heart sound signals were recorded at 44.1 kHz, and each one lasted for at least 56 s. For the convenience of displaying and testing, each sound was divided into three-heart-cycle segments. Moreover, every two adjacent segments had an overlap of one cardiac cycle.

Preprocessing

Down-sampling

Because the maximal frequency of the heart sounds did not exceed 1 kHz [12], the heart sound signals were down-sampled to 2 kHz according to the Nyquist-Shannon sampling theorem.

Denoising

A wavelet-based denoising method with soft threshold was utilized. The db4 wavelet was chosen as mother wavelet because its shape is similar to that of a heart sound signal, and the scale was established as 7 based on signal–noise-ratio (SNR) and normalized root mean squared error (NRMSE) [13]. The denoised heart sound was written as $HS(n)$.

Normalization

To standardize the processing of heart sound, the heart sound signals were normalized as

$$HS_N(n) = \frac{HS(n)}{\max(|HS(n)|)}, \quad n = 0, 1, \dots, N - 1 \quad (1)$$

where N is the total number of sampling points.

Unless otherwise specified, the expressions “original heart sound” and “original signal” appearing in the following refer to $HS_N(n)$.

Murmur elimination

Different from the conventional constant-cutoff-frequency-based murmur elimination method (usually 200 Hz) [8, 14], this paper presents a novel low-pass filter to remove the murmurs, namely the automatic-cutoff-frequency low pass filter (ALPF), whose cutoff frequency is calculated by analyzing the fast Fourier transform (FFT) of the heart sound.

Assuming that the modulus of the FFT sequence of $HS_N(n)$ is $FFT_H(n)$, $n = 0, 1, \dots, N - 1$, the envelope of $FFT_H(n)$ is obtained by the moving average method as follows:

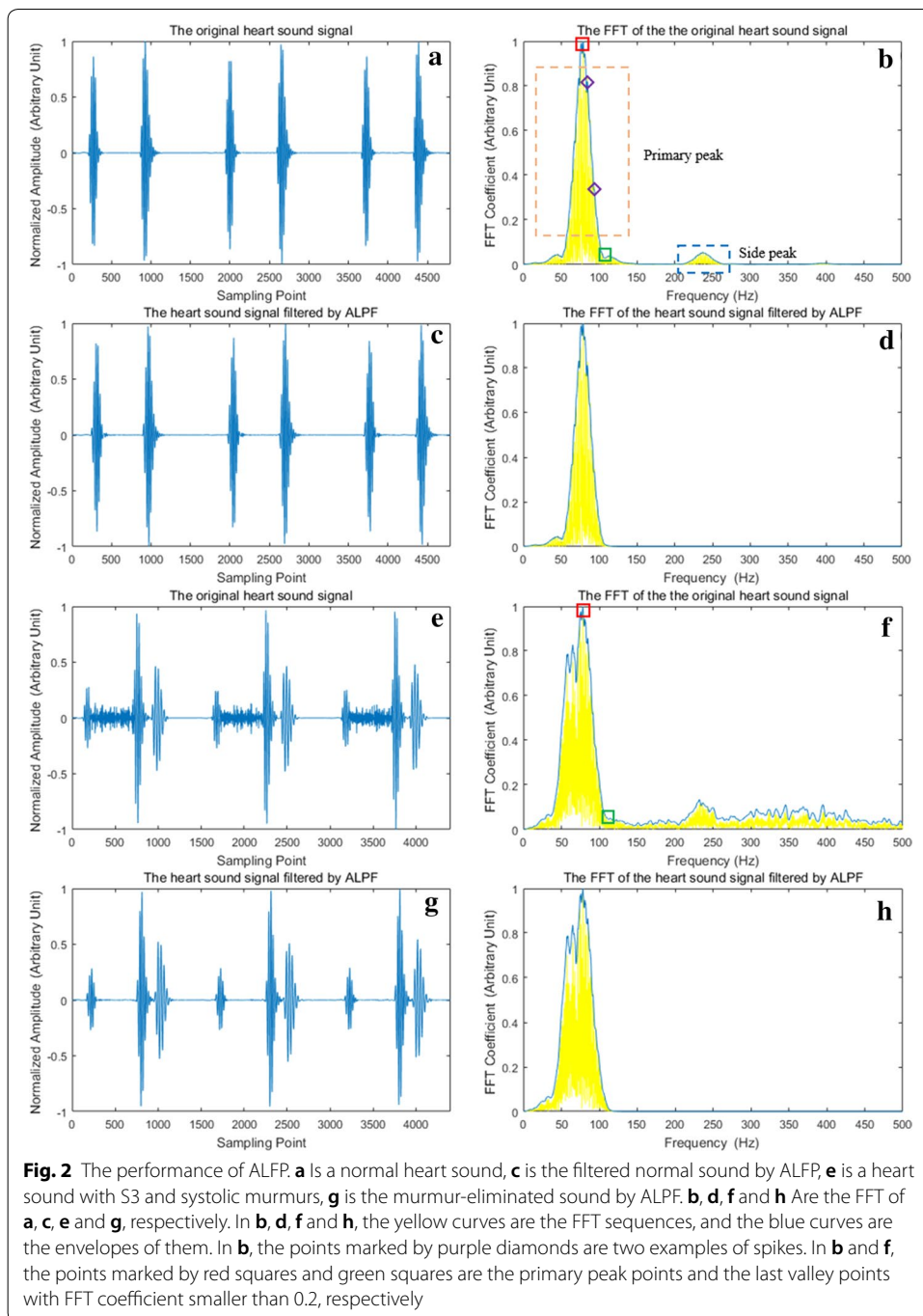
$$E_{FFT}(n) = \frac{1}{L_F + n + 1} \sum_{k=0}^{n+L_F} FFT_H(k), \quad 0 \leq n \leq L_F - 1$$

$$E_{FFT}(n) = \frac{1}{2L_F + 1} \sum_{k=n-L_F}^{n+L_F} FFT_H(k), \quad L_F \leq n \leq N - L_F - 1 \quad (2)$$

$$E_{FFT}(n) = \frac{1}{L_F + N - n} \sum_{k=n-L_F}^{n+L_F} FFT_H(k), \quad N - L_F \leq n \leq N - 1$$

where N is the sampling number of $FFT_H(n)$, and L_F is the neighborhood radius of point n , $L_F \ll N$. Note that if L_F is too large, some adjacent local peaks of $E_{FFT}(n)$ are merged together. Considering this and experimental results, L_F is set to 5. The envelope is also normalized using (1) and is written as $E_F(n)$. Unless otherwise specified, the “FFT coefficient” and “FFT envelope” appearing hereinafter refer to $E_F(n)$.

For the heart sounds without murmurs, $E_F(n)$ is generally comprised of primary peak and side peak (see in Fig. 2b). The primary peak is located in the low-frequency area with the highest amplitude, and the side peak is located in the relatively higher-frequency area with much smaller amplitude. Thus, most energy of heart sounds is concentrated within the frequencies of primary peak, and filtering out the side peak does not affect the shape of heart sounds (see in Fig. 2c, d). Therefore, the primary peak preserves the information of heart sound components (S1, S2, S3 and S4).



For the heart sounds with murmurs, the primary peak still exists, but the side peak is usually merged with the frequency band of murmurs (see in Fig. 2f). Considering that the murmurs usually have higher frequency than the primary peak, the frequency of primary peak's ending point is an appropriate estimation of the cutoff frequency to remove murmurs.

In this study, the ending point of primary peak is searched in the range of 20 to 200 Hz and is determined as the first valley point (the lowest point between two adjacent spikes)

after the primary peak point with FFT coefficient smaller than 0.2. The value 0.2 is determined based on experimental experience. If there is no such an ending point between 20 and 200 Hz, the cutoff frequency is set to 200 Hz.

$HS_N(n)$ is filtered using this frequency, and the filtered sound is then normalized; the normalized sound is denoted as HS_{NLP} .

Boundary detection

Envelope extraction

After removing the interference of murmurs, the next step is to calculate the onsets and offsets of heart sound components. Because the envelope can reduce the complexity of computing while preserving the location information of the signals, the closing operation of mathematical morphology is utilized to obtain the envelope in the proposed method and is defined as

$$[f \cdot g](n) = \{[f \oplus g] \otimes g\}(n) \quad (3)$$

where $f(n)$ refers to the input signal, and $g(n)$ is a structure element; \oplus denotes the dilation operation, and \otimes denotes the erosion operation. The lengths of $f(n)$ and $g(n)$ are P and Q respectively, and generally $P > Q$ [15, 16].

The choice of structure element $g(n)$ directly affects the shape of the envelope obtained by this operation [17]. In this study, in order to make the shape of envelope simple, $g(n)$ is designed as

$$g(n) = 0, \quad 0 \leq n \leq Q \quad (4)$$

Moreover, Q , the length of $g(n)$, is also a significant parameter. According to physiological knowledge, in general, the time interval between adjacent heart sound components (S1S2 interval, S2S3 interval, S4S1 interval, etc.) is no less than 100 ms [8, 18], which corresponds to 200 points under the sampling rate of 2 kHz. Thus, Q should be smaller than 200. In this study, Q was set as 30 on the basis of the experimental results.

The $f(n)$ in (3) is replaced by HS_{NLP} , and the heart sound envelope is obtained. The envelope is then normalized. The result is denoted as E_{NLP} .

Although the ALPF can effectively reduce murmurs, the filtered heart sound may retain some low-amplitude residuals, which are also reflected in the envelope and affect the detection of boundaries. In this method, three operations are used to overcome this problem.

Thresholding processing

The first operation is thresholding.

For some murmurs, such as early systolic murmurs, their onsets partially overlap the offsets of S1s. After filtering by ALPF, the residues of murmurs may still be linked with S1s. Consequently, the murmur residues and S1s may be mixed together in E_{NLP} . Therefore, the purpose of thresholding is to separate the boundaries of murmurs from those of S1s to conduct the subsequent operations. Consequently, a small threshold can complete this task. Due to the use of small threshold, the boundary information of the heart sound components can be preserved as much as possible.

The envelope \tilde{E}_{NLP} after thresholding is calculated as

$$\tilde{E}_{NLP} = \begin{cases} E_{NLP}, & E_{NLP} \geq \theta \\ 0, & E_{NLP} < \theta \end{cases} \quad (5)$$

where $\theta = \min(\theta_A, \theta_C)$, $\theta_A = \lambda \cdot \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} [E_{NLP}(n) - \mu]^2}$, $\lambda = 0.8$, $\mu = \frac{1}{N} \sum_{m=0}^{N-1} E_{NLP}(m)$

and $\theta_C = 0.025$. The parameter θ_A enables the algorithm to determine an appropriate threshold based on the characteristics of the signal itself, and θ_C is a reference threshold, preventing error in θ_A ; the value of θ_C is determined by experimental results.

Opening operation

After thresholding, the boundaries of heart sound components and interfering residues are separated. The remaining residues are considerably smaller than the heart sound components in both duration and amplitude. Because the opening operation of mathematical morphology can reduce spikes, it is utilized as the second stage of residue removal. The opening operation is defined as [19]

$$[f \circ g](n) = \{[f \otimes g] \oplus g\}(n) \quad (6)$$

The $g(n)$ in opening operation is still defined using (4), but the length of $g(n)$ is denoted as Q' . Besides, the parameters in (6) are same as those in (3).

As in the closing operation, Q' is also a significant parameter. Because the average durations of S1 and S2 are approximately 100 ms and the durations of S3 and S4 are approximately 50 ms [8], which correspond to 200 and 100 sampling points, respectively. Thus, Q' must be smaller than 100; in this study, it was set to 50 based on experimental results. The results of the opening operation are normalized, and the normalized signal is denoted as \tilde{E}_{NO} .

Energy thresholding processing

Finally, to completely exclude the interference components from the envelope, the energy thresholding approach is used. Considering that \tilde{E}_{NO} is comprised of points of zero and non-zero, i.e., the regions located by S1, S2, S3, S4 and few murmur residues are non-zero, and the rest regions are zero. Therefore, the boundaries of each component (including the interference components escaping the opening operation) can be automatically obtained as follows:

$$\text{If } \tilde{E}_{NO}(i) == 0 \&\& \tilde{E}_{NO}(i + 1) \neq 0, \quad i = 0, 1, \dots, N - 2$$

then, the i th point is determined as a potential onset;

$$\text{If } \tilde{E}_{NO}(j - 1) \neq 0 \&\& \tilde{E}_{NO}(j) == 0, \quad j = 1, 2, \dots, N - 1$$

the j th point is determined as a potential offset. The potential onset and the potential offset always appear in pairs and these onsets and offsets are denoted as $SPT(i)$ and $EPT(j)$ respectively, where $i, j = 0, 1, \dots, M - 1$. M is the number of potential onsets/offset.

Consequently, the energy of each component can be obtained as

$$E_k = \sum_{i=SPT(k)}^{EPT(k)} \tilde{E}_{NO}^2(i), \quad k = 0, 1, \dots, M - 1 \quad (7)$$

Assuming for each E_k ,

$$\text{If } E_k < \eta$$

then the k th potential onset and offset are considered invalid and removed from SPT and EPT . Moreover, the points between these invalid onsets and offsets are determined as interference parts and are set to zero using (8)

$$\tilde{E}_{NO}(n) = \begin{cases} 0, & SPT(k) \leq n \leq EPT(k), 0 \leq k \leq M - 1, \text{ if } E_k < \eta \\ \tilde{E}_{NO}(n), & \text{else} \end{cases} \quad (8)$$

In this inequality, η is set to 0.25 based on experimental experience. The remaining potential points are determined as the final onsets and offsets of the heart sound components.

Component identification

Cardiac cycle calculation

The last step is to recognize the heart sound components. Considering the quasi-periodic nature of heart sounds, this step can be more efficiently accomplished if the cardiac cycle is calculated. In some studies, the cardiac cycle was calculated by using the partial autocorrelation function (PACF) [10, 15]. However, because of the inherent defects of PACF, the calculating results are not satisfactory. In order to overcome the shortcomings of PACF, this study proposes a cardiac cycle calculation method based on the unbiased autocorrelation function (UACF), considerably improving the applicability.

The PACF and UACF are defined as (9) and (10), respectively.

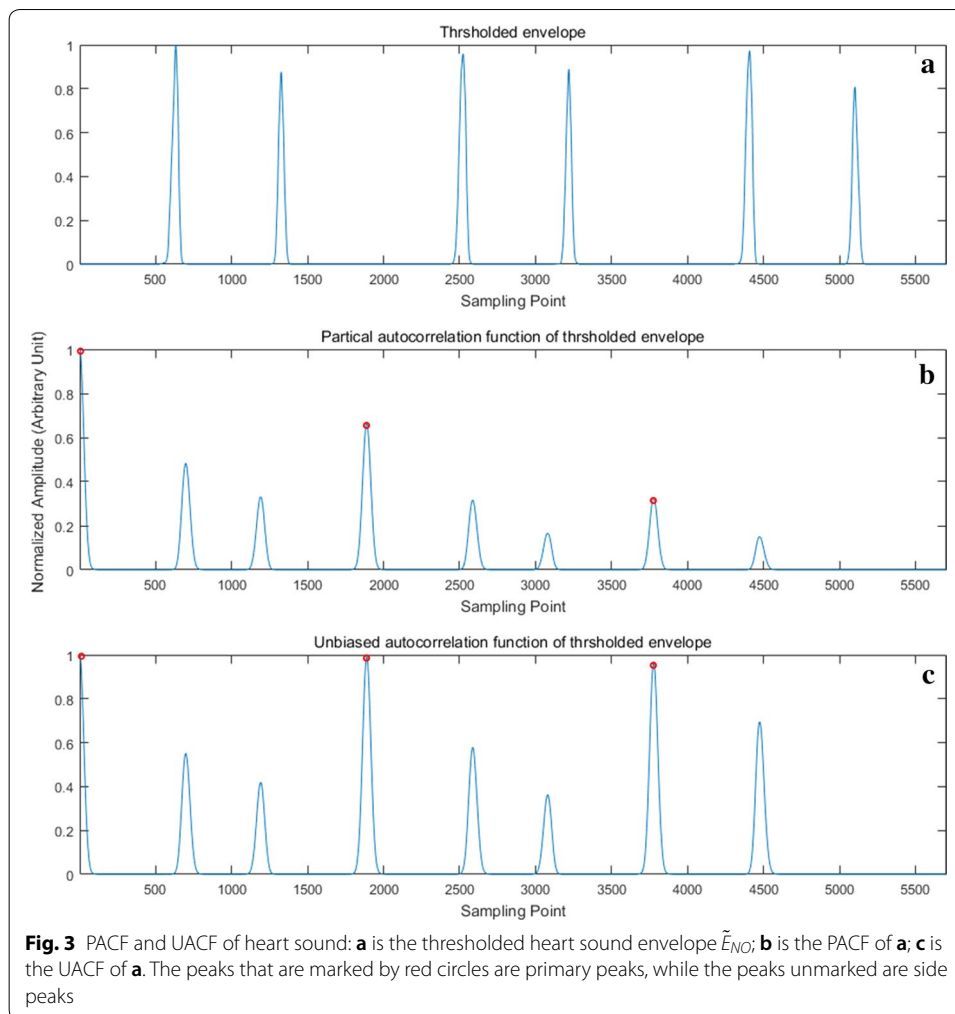
$$R(m) = \frac{1}{N} \sum_{n=0}^{N-m-1} \tilde{E}_{NO}(n) \tilde{E}_{NO}(n+m) \quad (9)$$

$$R'(m) = \frac{1}{N-m} \sum_{n=0}^{N-m-1} \tilde{E}_{NO}(n) \tilde{E}_{NO}(n+m) \quad (10)$$

where $m = 0, 1, \dots, N - 1$, N is the sampling number of \tilde{E}_{NO} . Figure 3 demonstrates the PACF and UACF.

For PACF, as m increases, the number of sampling points involved in multiplication and accumulation decreases gradually. Consequently, the primary peaks of the PACF exhibit gradual attenuation (see in Fig. 3b), causing the side peaks between the primary peaks to be interference of the cardiac cycle calculation. By contrast, the UACF only averages the terms that are involved in multiplication and accumulation, thereby overcoming the drawback of primary peak's attenuation in the PACF.

Figure 3c shows that there is a relatively significant difference in amplitude between the primary peaks and side peaks. To further expand the amplitude difference, square energy, which is defined in (11), is utilized. After obtaining the energy signal,



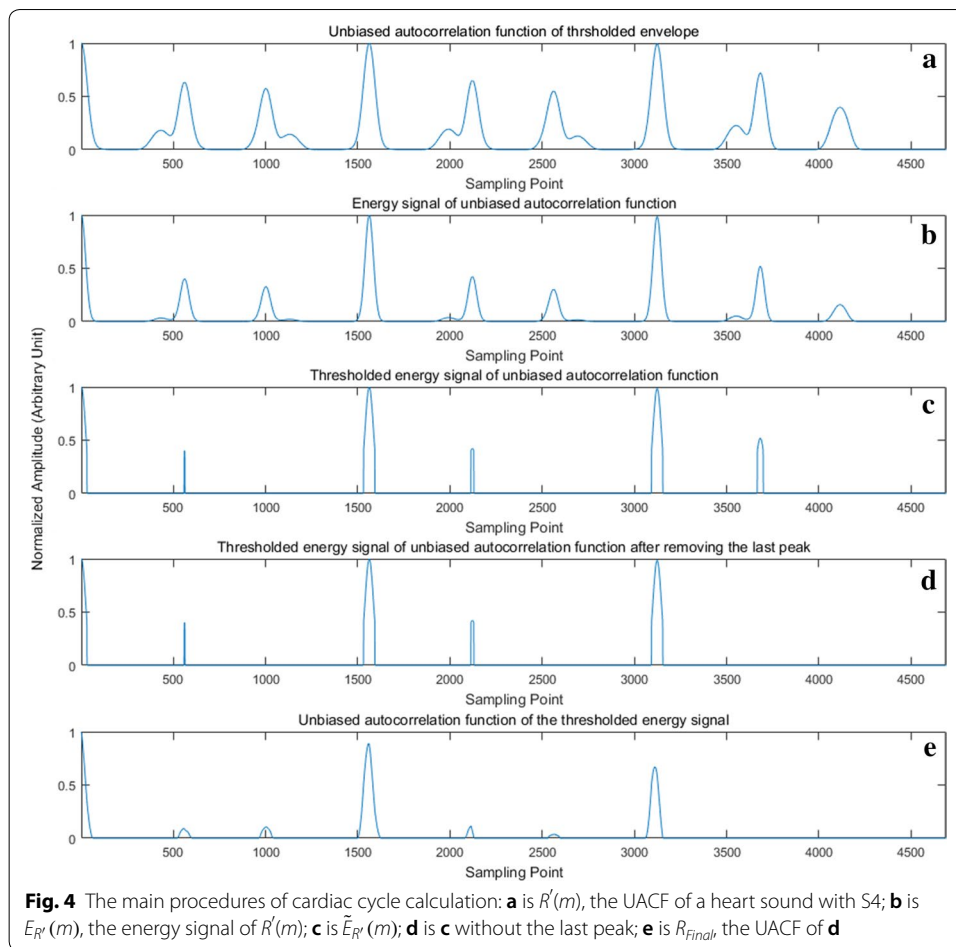
thresholding is used to remove the side peaks. The energy signal of the UACF and the energy signal after thresholding are denoted as $E_{R'}$ and $\tilde{E}_{R'}$, respectively.

$$E_{R'}(m) = R'^2(m), \quad m = 0, 2, \dots, N - 1 \quad (11)$$

$$\tilde{E}_{R'} = \begin{cases} E_{R'}, & E_{R'} \geq \sigma \\ 0, & E_{R'} < \sigma \end{cases} \quad (12)$$

where $\sigma = 0.4 \times \max(E_{R'})$.

The last peak of $\tilde{E}_{R'}$ is likely to be a side peak that escaped from thresholding because of its relatively higher amplitude (see in Fig. 4c). This kind of side peak is usually generated by the multiplication and accumulation of S1 and S2. In order to calculate the cardiac cycle accurately, the last peak of $\tilde{E}_{R'}$ is forcibly removed regardless of whether it is a side peak or a primary peak (see in Fig. 4d). Finally, the UACF of $\tilde{E}_{R'}$ with the last peak removed is calculated, and this UACF is denoted as R_{Final} (see in Fig. 4e). Then, thresholding is performed on R_{Final} with a threshold of $0.5 \times \max(R_{Final})$, leaving only primary peaks in the thresholded R_{Final} . The intervals between adjacent peaks

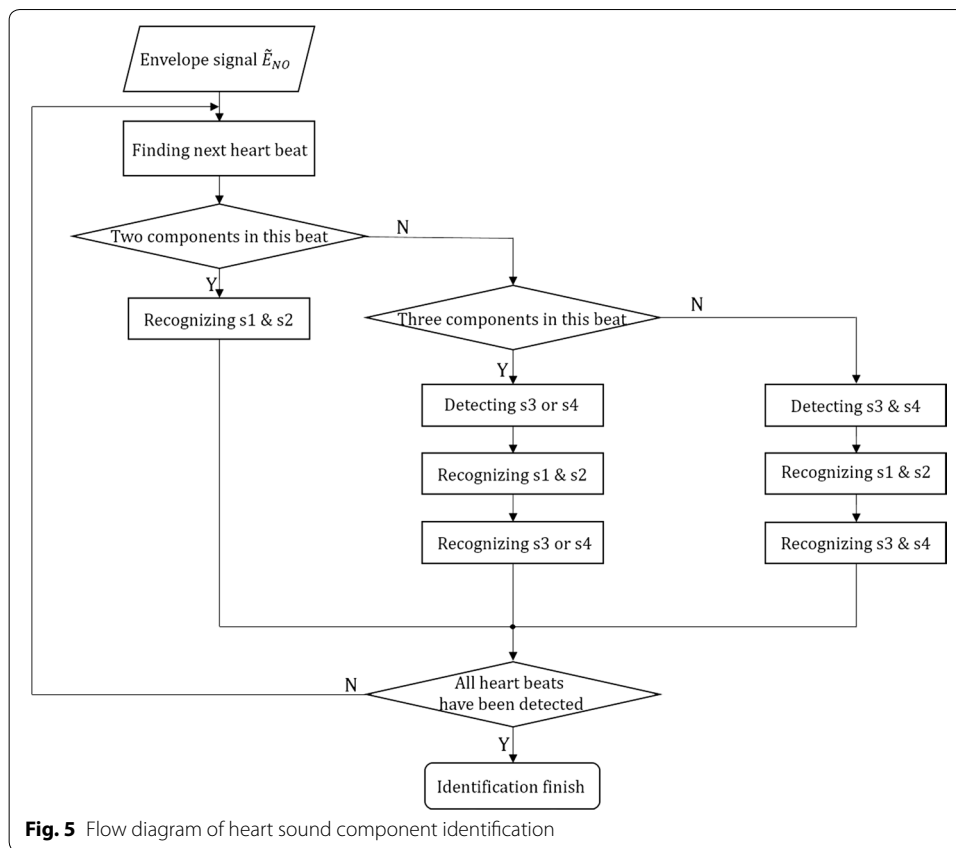


of thresholded R_{Final} are then calculated and averaged to obtain the average cardiac cycle. Figure 4 shows the main procedures of cardiac cycle calculation.

Component recognition

After determining the cardiac cycle, the heart sound components can be recognized beat by beat. Under normal circumstances, in terms of the number of components in one heartbeat, heart sounds can be divided into three categories: two components (S1 and S2), three components (S1, S2, S3, or S4) and four components (S1, S2, S3, and S4). Moreover, in the time domain, S1S2 interval < S2S1 interval, S2S3 interval < S3S1 interval, and S2S4 interval > S4S1 interval, whereas in the frequency domain, the frequencies of S1 and S2 are usually higher than those of S3 and S4 [8, 15]. These priori knowledge along with the mentioned three categories can be utilized to recognize the heart sound components.

Figure 5 presents the overall process of component recognition. The algorithm finds a heartbeat based on cardiac cycle and automatically obtains its component number. Then, the recognition processing is performed according to the component number found.



For the convenience of expression, in the following part, the j th component of the i th heartbeat is denoted as C_j^i , and the time interval between C_m^p and C_n^q is denoted as $T_{mr}^{p,q}$.

Assuming that the i th heartbeat is being analyzed, then the recognition processing can be expressed as follows.

Two components: In this case, $T_{12}^{i,i}$ and $T_{21}^{i,i+1}$ are calculated. If $T_{12}^{i,i} < T_{21}^{i,i+1}$, C_1^i is identified as S1, and C_2^i is identified as S2; otherwise, C_1^i is S2, and C_2^i is S1.

Three components: In this case, because S3 and S4 are less than S1 and S2 in frequency, these three components can be recognized using three steps: detecting S3/S4, recognizing S1 and S2, and identifying S3/S4.

Time–frequency analysis is required for the detection of S3 and/or S4. In this study, S-transform is employed to accomplish this task. Assuming that the signal to be analyzed is $h(n)$, $n = 0, 1, \dots, N - 1$, its Fourier transform is $H(k/NT)$, $k = 0, 1, \dots, N - 1$, where T is the sampling interval. Then, the S-transform of $h(n)$ is given by [20].

$$\begin{cases} S(jT, \frac{n}{NT}) = \sum_{m=0}^{N-1} H(\frac{m+n}{NT}) e^{-\frac{2\pi^2 m^2}{n^2}} e^{\frac{i2\pi mj}{N}}, & n \neq 0 \\ S(jT, 0) = \frac{1}{N} \sum_{m=0}^{N-1} (\frac{m}{NT}), & n = 0 \end{cases} \quad (13)$$

where $j, m, n = 0, 1, \dots, N - 1$.

To obtain sufficient time–frequency information, the S-transform is directly performed on the original heart sound signal $HS_N(n)$. The result is a complex matrix and

is written as S . Suppose that the element of the m th row in the n th column of the matrix is $S_{m,n} = a + jb$, then the modulus of $S_{m,n}$ is defined as.

$$|S|_{m,n} = |S_{m,n}| = \sqrt{a^2 + b^2} \quad (14)$$

The modulus of each element of S is calculated, and the modulus matrix $|S|$ is then obtained. Unless otherwise specified, both the expression “S transform” and “S coefficient” in the following text refer to $|S|$.

The two-dimensional $|S|$ matrix is relatively difficult to analyze; thus, the one-dimensional instantaneous frequency is utilized to detect S3 and S4, and is defined as.

$$f_H(n) = \frac{\sum_{m=0}^{N-1} F(m) \cdot |S|_{m,n}}{\sum_{m=0}^{N-1} |S|_{m,n}} \quad (15)$$

where $F(m) = \frac{m}{NT} = \frac{m \cdot f_s}{N}$ represents the frequency of the m th row of the S-transform, and f_s is the sampling rate. Because the maximal frequency of the sampled signal is half of the sampling frequency, $F(m)$ was set as $\frac{1}{2} \frac{m}{NT} = \frac{m \cdot f_s}{2N}$, $m = 0, 1, \dots, N-1$ in this study.

Furthermore, the boundaries of heart sound components are obtained in the boundary detection phase (i.e., *SPT* and *EPT*). Therefore, by extracting the parts between these onsets and offsets from $f_H(n)$, the instantaneous frequency of heart sound components is obtained (see in Fig. 6(3)).

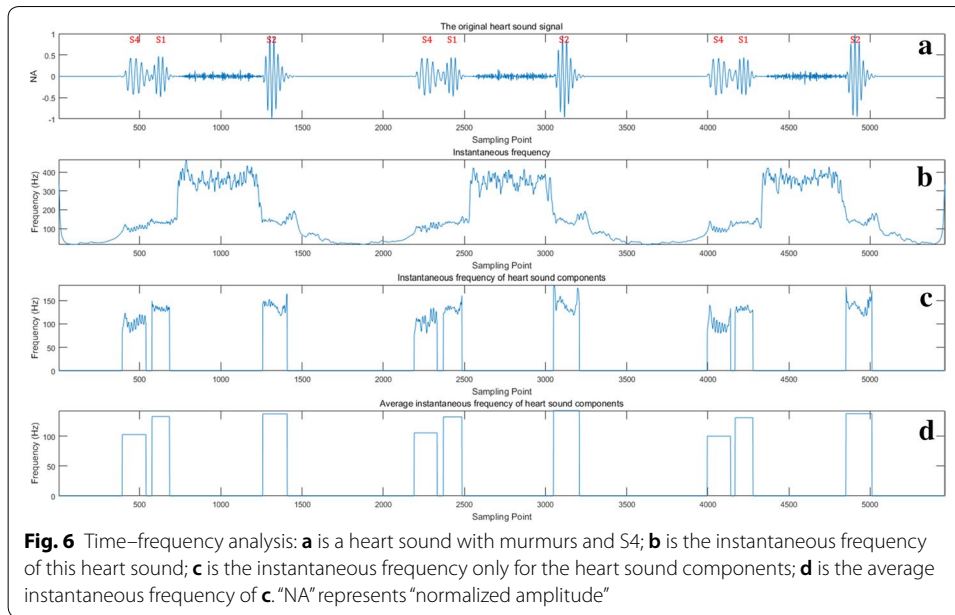
Finally, the average instantaneous frequency is calculated. Figure 6a–d present an example of these operations. It is observed that the average instantaneous frequency of S4 is significantly lower than those of S1 and S2. Similarly, the instantaneous frequency of S3 is also the lowest among S1, S2 and S3. Consequently, the component with the lowest frequency of each heartbeat is S3 or S4, which means that the previous component is S2 and the next is S1.

Without loss of generality, assuming that C_2^i is the component with the lowest frequency, C_1^i and C_3^i can be easily determined as S2 and S1, respectively. Then, $T_{12}^{i,i}$ and $T_{23}^{i,i}$ are calculated and compared. If $T_{12}^{i,i} < T_{23}^{i,i}$, C_2^i is recognized as S4; otherwise, if $T_{12}^{i,i} > T_{23}^{i,i}$, C_2^i is S3.

Four components: In this case, the four components are arranged in the order: S1 \rightarrow S2 \rightarrow S3 \rightarrow S4 \rightarrow ... S1 \rightarrow S2. Thus, the first one of the two components with the lowest instantaneous frequencies is S3, and the second is S4. Moreover, the component preceding the detected S3 and S4 is S2, and the following component is S1.

After all the components in the i th heartbeat are recognized, the algorithm finds the next beat, and performs the same recognition processing. When all the heartbeats are analyzed, the recognition process is finished.

Finally, the heart sounds are divided into segments. For the sounds containing S1 and S2 only (including the sounds with murmurs), they are segmented into S1, S1S2 interval, S2 and S2S1 interval. For the sounds containing extra S3/S4 (sounds with murmurs are also included), they are segmented into S1, S1S2 interval, S2, S2S3/S2S4 intervals, S3/S4, S3S1/S4S1 intervals.



Results

Experimental setup

Murmur elimination evaluation

For the evaluation of murmur elimination, a novel index, the signal murmur ratio (SMR), is proposed as follows:

$$SMR = 10 \cdot \lg \frac{\sum_{n \in U} X^2(n)}{\sum_{m \in V} X^2(m)} \quad (16)$$

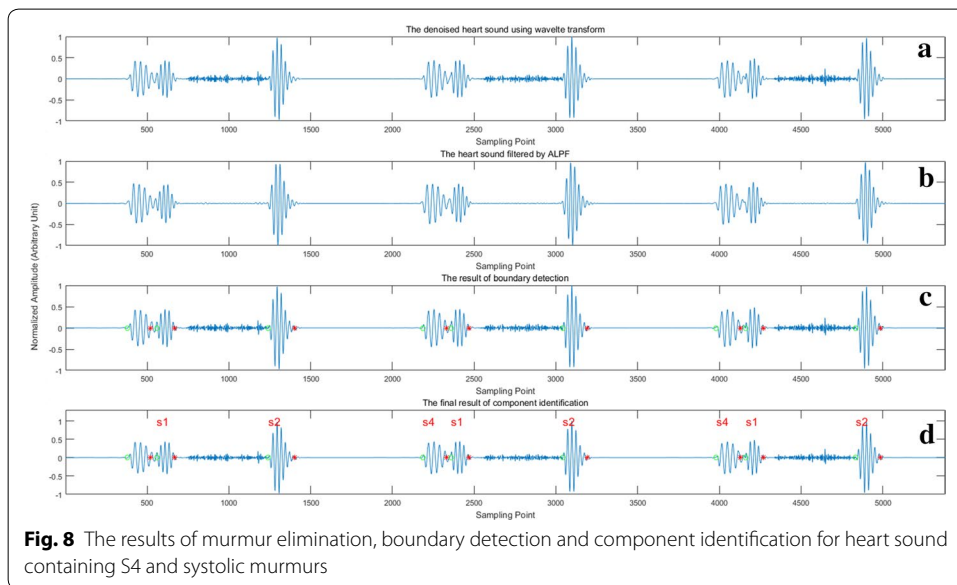
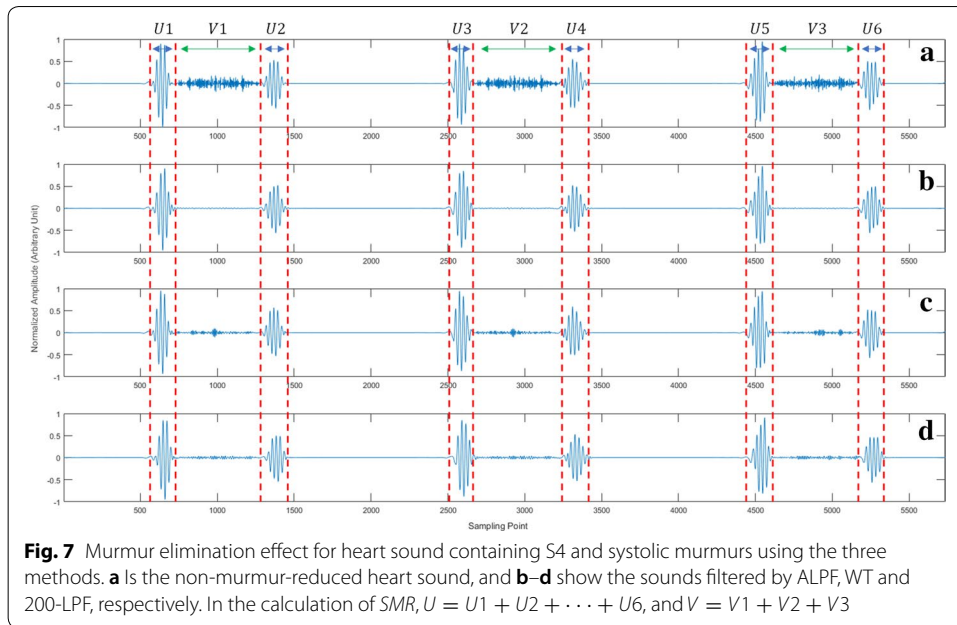
where $X(n)$ is the heart sound signal, U is the region where heart sound components (S1, S2, S3, and S4) are located, and V is the region where murmurs are located. To ensure the validity of the indicator, both U and V are manually determined from the original signal $HS_N(n)$. Thereby, for a certain heart sound, U and V are constant when calculating SMRs of the murmur-reduced heart sound and the non-murmur-reduced heart sound (see in Fig. 7). A higher SMR indicates a smaller murmur energy proportion in signal $X(n)$, which means a stronger murmur removing effect.

Boundary detection and component identification evaluation

To validate the performance of boundary detection and component identification, three classical metrics, sensitivity (Se), positive predictive value (PPV), and accuracy (Acc) are calculated. These three indices are defined as follows:

$$\text{Sensitivity(Se)} = \frac{TP}{TP + FN} \times 100\% \quad (17)$$

$$\text{Positive predictive value (PP)} = \frac{TP}{TP + FP} \times 100\% \quad (18)$$



$$\text{Accuracy (Acc)} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (19)$$

In this experiment, the onsets and offsets of heart sound components were automatically marked with green circles and red stars by the algorithm (see in Fig. 8c), respectively. In addition, the names of these components were automatically labeled above them (see in Fig. 8d). The results of boundary localization and component identification were judged by human and quantified using (17)–(19), in which the FN means false negative, FP means false positive, TN means true negative and TP means true positive.

Table 1 Statistics of SMR

UMich no.	Heart sound types	Length (s)	Cycle number	<i>SMR_A</i> /db	<i>SMR_{WPT}</i> /db	<i>SMR₂₀₀</i> /db	<i>SMR_{Ori}</i> /db
6	Early Systolic Murmur, Apex	60	66	24.24	22.91	15.52	14.60
7	Mid Systolic Murmur, Apex	67	72	28.34	25.37	22.01	12.51
8	Late Systolic Murmur, Apex	61	63	30.27	28.93	29.16	21.64
9	Holosystolic Murmur, Apex	60	63	24.26	22.20	21.96	16.07
4	Mid Systolic Click, Apex	60	72	25.12	22.25	19.50	19.14
10	Systolic Click and Late Systolic Murmur, Apex	64	72	29.03	26.35	24.18	23.23
11	S4 and Mid Systolic Murmur, Apex	65	72	34.35	28.18	28.93	17.06
12	S3 and Holosystolic Murmur, Apex	64	87	27.63	26.35	26.77	16.09
16	Early Diastolic Murmur, Aortic	61	81	32.12	31.92	31.88	31.37
21	Ejection Systolic Murmur and Transient Split S2, Pulmonic	56	75	26.41	23.68	21.33	20.47
22	Split S2 and Ejection Systolic Murmur, Pulmonic	66	78	18.07	14.46	12.52	11.26
23	Ejection Systolic Murmur and Single S2 and Ejection Click, Pulmonic	75	87	24.30	20.20	21.96	16.07
Arithmetic average \pm SD				26.95 \pm 4.30	24.42 \pm 4.54	22.61 \pm 5.86	18.35 \pm 5.41
Weighted average \pm SD				26.92 \pm 4.14	24.38 \pm 4.43	22.61 \pm 5.67	18.39 \pm 5.29

The italic values are represent the optimal results, compared with the results achieved in other methods/studies

Results of murmur elimination

The performance of ALPF was compared with two existing methods. The first method is the conventional low-pass filter with a constant cutoff frequency of 200 Hz (200-LPF), and the second method is the wavelet package transform method with a db10 mother wavelet and a scale of 5 (WPT) [14]. In the WPT method, the heart sound components were reconstructed by nodes (5, 0) to (5, 5) to remove interference components such as murmurs. The SMR of the heart sounds filtered by the ALPF, WPT and 200-LPF are denoted as SMR_A , SMR_{WPT} , and SMR_{200} , respectively. Moreover, the SMR of the original heart sounds $HS_N(n)$ was calculated and denoted as SMR_{Ori} . In addition, the SMR gains of these three methods were achieved by subtracting SMR_{Ori} from SMR_A , SMR_{WPT} and SMR_{200} , respectively. All the experimental parameters of these three methods were identical.

Tables 1 and 2 show the statistical results of SMR and SMR gains. It is observed that the proposed ALPF has the highest score of SMR and SMR gains. The overall

Table 2 Statistics of SMR gains

UMich no.	Heart sound types	Length (s)	Cycle number	SMR gains		
				ALPF	WPT	200LPF
6	Early Systolic Murmur, Apex	60	66	<i>9.64</i>	8.31	0.91
7	Mid Systolic Murmur, Apex	67	72	<i>15.84</i>	12.86	9.50
8	Late Systolic Murmur, Apex	61	63	<i>8.64</i>	7.29	7.52
9	Holosystolic Murmur, Apex	60	63	<i>8.20</i>	6.13	5.90
4	Mid Systolic Click, Apex	60	72	<i>5.99</i>	3.11	0.37
10	Systolic Click and Late Systolic Murmur, Apex	64	72	<i>5.79</i>	3.12	0.95
11	S4 and Mid Systolic Murmur, Apex	65	72	<i>17.29</i>	11.12	11.88
12	S3 and Holosystolic Murmur, Apex	64	87	<i>11.54</i>	10.25	10.68
16	Early Diastolic Murmur, Aortic	61	81	<i>0.765</i>	0.55	0.51
21	Ejection Systolic Murmur and Transient Split S2, Pulmonic	56	75	<i>5.94</i>	3.21	0.86
22	Split S2 and Ejection Systolic Murmur, Pulmonic	66	78	<i>6.81</i>	3.20	1.26
23	Ejection Systolic Murmur and Single S2 and Ejection Click, Pulmonic	75	87	<i>7.51</i>	3.70	0.79
Arithmetic Average \pm SD				<i>8.66 \pm 4.52</i>	6.07 \pm 3.88	4.26 \pm 4.51
Weighted Average \pm SD				<i>8.59 \pm 4.37</i>	5.99 \pm 3.76	4.21 \pm 4.37

The italic values are represent the optimal results, compared with the results achieved in other methods/studies

performances were evaluated by two indices, arithmetic average (AA) and weighted average (WA). AA was achieved by simply averaging the SMRs of 12 types of heart sounds and WA was affected by the weight of each type's cycle number. In terms of SMR, the AA (\pm SD) and WA (\pm SD) of SMR_A reach 26.95 ± 4.30 db and 26.92 ± 4.14 db, respectively; by contrast, the AA and WA are $24.42.07 \pm 4.54$ and 24.38 ± 4.43 for SMR_{WPT} and 22.61 ± 5.86 and 22.61 ± 5.29 for SMR_{200} . In terms of SMR gains, the heart sounds filtered by the ALPF have an arithmetic average and weighted average SMR gain (\pm SD) of 8.66 ± 4.52 db and 8.59 ± 4.37 db; by contrast, the arithmetic average and weighted average SMR gains (\pm SD) are 6.07 ± 3.88 and 5.99 ± 3.76 for WPT and 4.26 ± 4.51 and 4.21 ± 4.51 for 200-LPE.

Figure 7 shows an example of the murmur elimination results using the three methods.

Results of boundary detection and component identification

A detailed quantitative description of the boundary positioning result is shown in Table 3. It is observed that the proposed method achieves a sensitivity of 100% for all 16 types of heart sounds, and both the PPV and accuracy are 99.93%, indicating an outstanding performance.

Table 4 presents the quantitative data of component identification. Because of the good positioning effect, the identification performance is high, with an average Se of 98.63%, an average PPV of 99.86%, and an average accuracy of 98.49%.

Figure 8 shows the complete process of the proposed method, including murmur elimination, boundary detection, and component identification. The first component (S4) in Fig. 8d is not labeled. This is because the identification of S4 requires

Table 3 Quantitative estimation of positioning

UMich no.	Heart sound types	Length (s)	Cycle number	TP	FN	FP	TN	Se/%	PPV/%	Acc/%
1	Normal S1 and S2, Apex	69	87	174	0	0	0	100	100	100
2	Split S1, Apex	71	102	204	0	0	0	100	100	100
3	S4 Gallop, Apex	75	93	279	0	0	0	100	100	100
4	Mid Systolic Click, Apex	60	72	144	0	2	0	100	98.63	98.63
5	S3 Gallop, Apex	68	81	243	0	0	0	100	100	100
6	Early Systolic Murmur, Apex	60	66	132	0	0	0	100	100	100
7	Mid Systolic Murmur, Apex	67	72	144	0	0	0	100	100	100
8	Late Systolic Murmur, Apex	61	63	126	0	0	0	100	100	100
9	Holosystolic Murmur, Apex	60	63	126	0	0	0	100	100	100
10	Systolic Click and Late Systolic Murmur, Apex	61	72	144	0	0	0	100	100	100
11	S4 and Mid Systolic Murmur, Apex	65	87	261	0	0	0	100	100	100
12	S3 and Holosystolic Murmur, Apex	64	87	261	0	0	0	100	100	100
14	Normal S1 and S2, Aortic	61	69	138	0	0	0	100	100	100
16	Early Diastolic Murmur, Aortic	61	81	162	0	0	0	100	100	100
18	Single S2, Pulmonic	61	69	138	0	0	0	100	100	100
23	Ejection Systolic Murmur and Single S2 and Ejection Click, Pulmonic	75	87	174	0	0	0	100	100	100
Total		1039	1251	2850	0	2	0	100	99.93	99.93

Table 4 Quantitative estimation of identification

UMich no.	Heart sound types	Length (s)	Cycle number	TP	FN	FP	TN	Se/%	PPV/%	Acc/%
1	Normal S1 and S2, Apex	69	87	174	0	0	0	100	100	100
2	Split S1, Apex	71	102	204	0	0	0	100	100	100
3	S4 Gallop, Apex	75	93	279	0	0	0	100	100	100
4	Mid Systolic Click, Apex	60	72	142	3	4	0	97.73	97.26	95.30
5	S3 Gallop, Apex	68	81	243	0	0	0	100	100	100
6	Early Systolic Murmur, Apex	60	66	132	0	0	0	100	100	100
7	Mid Systolic Murmur, Apex	67	72	144	0	0	0	100	100	100
8	Late Systolic Murmur, Apex	61	63	126	0	0	0	100	100	100
9	Holosystolic Murmur, Apex	60	63	126	0	0	0	100	100	100
10	Systolic Click and Late Systolic Murmur, Apex	61	72	144	0	0	0	100	100	100
11	S4 and Mid Systolic Murmur, Apex	65	87	261	0	0	0	100	100	100
12	S3 and Holosystolic Murmur, Apex	64	87	261	0	0	0	100	100	100
14	Normal S1 and S2, Aortic	61	69	102	36	0	0	73.91	100	73.91
16	Early Diastolic Murmur, Aortic	61	81	162	0	0	0	100	100	100
18	Single S2, Pulmonic	61	69	138	0	0	0	100	100	100
23	Ejection Systolic Murmur and Single S2 and Ejection Click, Pulmonic	75	87	174	0	0	0	100	100	100
Total		1039	1251	2812	39	4	0	98.63	99.86	98.49

information obtained from the previous cardiac cycle. However, this cycle lacks the information. The unlabeled S4 component is identified in the previous segment. Therefore, there is no problem of missed detection.

Discussion

Tables 1 and 2 show the results of murmur elimination, all three methods exhibit an improvement in SMR compared with SMR_{Ori} . As a result of selecting the appropriate cutoff frequency, the proposed ALPF has the highest SMR gain among all the tested methods. The arithmetic and weighted average SMR gains of ALPF reach 8.66db and 8.59db, which are 42.67% and 43.41% higher than WPT's 6.07db and 5.99db, and 103.29% and 104.04% higher than 200-LPF's 4.26db and 4.21db, respectively, indicating optimal murmur elimination performance.

In order to verify the superiority of the proposed method for murmur elimination, in this study, the paired two-sample t-test was performed on the SMR gains calculated for 12 types of heart sounds filtered with these three methods. In the paired two-sample t-test, the significance level α was set to 0.01. Under this condition, the rejection region is $|t| \geq t_{\alpha/2}(n-1) = t_{0.005}(11) = 3.1058$, and $t_{A,WPT} = 5.5499 > 3.1058$, $t_{A,200} = 5.5039 > 3.1058$, and $t_{WPT,200} = 2.6273 < 3.1058$. Thus, the proposed murmur elimination method was verified to have significant advantages compared with WPT and 200-LPF.

For boundary detection results shown in Table 3, among all 16 types of heart sounds, the proposed method failed on one signal labeled as a mid-systolic click: two click components were not removed in boundary detection process and were erroneously positioned as true components. Time–frequency analysis was then performed to determine the reason for this.

The time–frequency analysis revealed that these click components have a relatively high number of overlaps with S1 and S2 in frequency, resulting in the failure in removing them. Compared with this sound, the other signals' interference components (clicks, murmurs, and residues of noise) show a smaller overlap with the true components; thus, the positioning score is higher. In fact, setting a higher threshold θ can completely remove the clicks, resulting in a higher performance score. However, the higher threshold causes a larger deviation between the calculated boundaries and the real values. After weighing the pros and cons, this approach was abandoned.

The identification results of Table 4 shows that the fourth signal (Umich No. 4) is consistent with the result of positioning. A difference appears in the result of the 13th signal (Umich No. 14). The FN of the signal reaches 36, resulting in a low Se and low Acc of 73.91%. Initially, this result was puzzling because the signal was labeled as normal. After a comprehensive analysis, the rhythm of some parts of this heart sound was found to be irregular, resulting in the failure of heart cycle calculation in these parts. Therefore, components of these parts could not be recognized.

For the performance evaluation, because the cycle number of each heart sound was very close and the performance of each type of signal was analyzed separately, the utilized dataset did not suffer from the problem of unbalance.

Some other operating characteristics obtained in this study and 11 previous studies are listed in Table 5 for comparison. Both the Se and PPV for positioning in literature [6] consist of two values separated by a slash. The two values are for normal and

Table 5 A comprehensive comparison with other studies

Methods	Application capability ^a				Functions ^b		Positioning ^c			Identifying		
	N	S3	S4	Mur	P	I	Se	PPV	Acc	Se	PPV	Acc
Proposed method	✓	✓	✓	✓	✓	✓	<i>100</i>	<i>99.93</i>	<i>99.93</i>	98.63	<i>99.86</i>	<i>98.49</i>
Nasari et al. [8]	✓	✓	✓	✓	✗	✓	✗	✗	✗	<i>99.00</i>	98.60	NM
Varghees et al. [21]	✓	✓	✓	✓	✓	✗	99.43	93.56	93.06	✗	✗	✗
Sepehria et al. [22]	✓	✓	✓	✓	✗	✓ [*]	✗	✗	✗	✗	✗	93.6
Moukadem et al. [6]	✓	✗	✗	✓	✓	✓	96/97	95/95	NM	95	97	NM
Moukadem et al. [26]	✓	✗	✗	✓	✓	✗	95	98	NM	✗	✗	✗
Pedrosa et al. [23]	✓	✗	✗	✗	✓	✓	89.2	98.6	NM	✗	✗	✗
Wang et al. [24]	✓	✗	✗	✓	✓	✓	NM	NM	NM	NM	NM	NM
Schmidt et al. [25]	✓	✗	✗	✗	✗	✓	✗	✗	✗	98.8	98.6	NM
Tseng et al. [2]	✓	✗	✗	✗	✗	✓	NM	NM	NM	92.4	88.1	NM
Zhong et al. [27]	✓	✗	✗	✗	✓	✓	NM	NM	NM	92.84	NM	NM
Wang et al. [28]	✓	✗	✗	✗	✓	✓	NM	NM	NM	96.8	NM	NM

The italic values are represent the optimal results, compared with the results achieved in other methods/studies

^a N: normal; Mur: murmurs

^b P: positioning; I: identifying; ✓^{*}: only available for S1 and S2

^c NM: not mentioned

pathological cases of heart sounds, respectively. As mentioned in the “**Background**”, a complete segmentation is comprised of boundary detection and component identification. However, several existing works only focus on one of them. Moreover, some studies could be applied only to the basic situation, for example, heart sounds without murmurs. Therefore, in addition to achieving excellent performance, the proposed approach achieves the functions of both boundary localization and component identification and can be applied in more conditions.

Conclusions

This paper presents an accurate heart sound segmentation algorithm that combines time-domain, frequency-domain and time–frequency-domain analysis. Compared to existing studies, this method is applicable to a wide range of heart sounds, from normal to those containing S3, S4 and various murmurs. To verify this method, quantitative experiments were performed using the University of Michigan’s Heart Sound & Murmur Library, an authoritative open database. The experimental materials incorporated two types of normal heart sounds and 14 types of abnormal heart sounds. The results show that the boundary localization has an average Se of 100%, an average PPV of 99.3% and an average Acc of 99.93%. Moreover, the Se, PPV and Acc of the component identification reach 98.63%, 99.86% and 98.49%, respectively, indicating outstanding performance of the proposed method. There are still some shortcomings of this work. For example, the component identification relies on the success of cardiac cycle calculation; therefore, this method cannot be applied to the heart sound with severe arrhythmia because of the failure to achieving accurate cardiac cycle by using UACF. The study of segmentation provides a good basis for extracting significant features of heart sounds. Therefore, the further study will focus on the classification of heart sounds.

Authors' contributions

QL developed the main algorithm. XW and XM provided professional verification of the experimental results. All authors read and approved the final manuscript.

Author details

¹ Department of Electronic Engineering, Fudan University, Room 522 B, Science Building, 220 Handan Rd., Shanghai, China. ² Key Laboratory of Medical Imaging Computing and Computer Assisted Intervention (MICCAI) of Shanghai, 138 Medical College Rd., Shanghai, China. ³ Children's Hospital of Fudan University, 399 Wanyuan Rd., Shanghai, China.

Acknowledgements

The authors would like to thank Science and Technology Commission of Shanghai Municipality for financial support.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

The datasets analyzed during the current study are available in the [Heart Sound & Murmur Library of UMich] repository, (http://www.med.umich.edu/lrc/psb_open/html/repo/primer_heartsound/primer_heartsound.html).

Consent for publication

Not applicable.

Ethics approval and consent to participate

Not applicable.

Funding

This work received financial support from Science and Technology Commission of Shanghai Municipality, Grant No.15DZ2251700 and 13DZ1941802.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 23 May 2018 Accepted: 30 July 2018

Published online: 06 August 2018

References

- Chen W, Gao R, Liu L. Report on cardiovascular diseases in China 2014. *Chin Circ J*. 2015;30:617–22.
- Tseng YL, Ko PY, Jaw FS. Detection of the third and fourth heart sounds using Hilbert-Huang transform. *Biomed Eng Online*. 2012;11:8.
- Yuenyong S, Nishihara A, Kongprawechnon W, et al. A framework for automatic heart sound analysis without segmentation. *Biomed Eng Online*. 2011;10:13.
- Liang H, Lukkarinen S, Hartimo I. Heart sound segmentation algorithm based on heart sound envelopogram. In: *Computers in Cardiology 1997*. Lund: IEEE; 1997. 7–10 September 1997, p. 105–8.
- Kumar D, et al. Third heart sound detection using wavelet transform-simplicity filter. In: *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE. IEEE, 2007*. p. 1277–81.
- Moukadem A, et al. A robust heart sounds segmentation module based on S-transform. *Biomed Signal Process Control*. 2013;8:273–81.
- Springer DB, Tarassenko L, Clifford GD. Logistic regression-HSMM-based heart sound segmentation. *IEEE Trans Biomed Eng*. 2016;63:822–32.
- Naseri H, Homaeinezhad MR. Detection and boundary identification of phonocardiogram sounds using an expert frequency-energy based metric. *Ann Biomed Eng*. 2013;41:279–92.
- Boutana D, Benidir M, Barkat B. Segmentation and identification of some pathological phonocardiogram signals using time-frequency analysis. *IET Signal Process*. 2011;5:527–37.
- Tang H, et al. Segmentation of heart sounds based on dynamic clustering. *Biomed Signal Process Control*. 2012;7:509–16.
- Heart Sound & Murmur Library. University of Michigan, US. http://www.med.umich.edu/lrc/psb_open/html/repo/primer_heartsound/primer_heartsound.html. Accessed 7 Apr 2017.
- Choi S, Jiang Z. Comparison of envelope extraction algorithms for cardiac sound signal segmentation. *Expert Syst Appl*. 2008;34:1056–69.
- Qiu S, et al. A stimulus artifact removal technique for SEMG signal processing during functional electrical stimulation. *IEEE Trans Biomed Eng*. 2015;62:8.
- Zheng Y, Guo X, Ding X. A novel hybrid energy fraction and entropy-based approach for systolic heart murmurs identification. *Expert Syst Appl*. 2015;42:2710–21.
- Chen Y, Duan H. A QRS complex detection algorithm based on mathematical morphology and envelope. In: *Engineering in Medicine and Biology Society. IEEE-EMBS 2005. 27th Annual International Conference of the IEEE, 2006*. p. 4654–57.
- Chu CHH, Delp EJ. Impulsive noise suppression and background normalization of electrocardiogram signals using morphological operators. *IEEE Trans Biomed Eng*. 1989;36:262–73.
- Trahanas PE. An approach to QRS complex detection using mathematical morphology. *IEEE Trans Biomed Eng*. 1993;40:201–5.

18. Barma S, et al. Detection of the third heart sound based on nonlinear signal decomposition and time–frequency localization. *IEEE Trans Biomed Eng.* 2016;63:1718–27.
19. Hu JL, Bao SD. An approach to QRS complex detection based on multiscale mathematical morphology. In: *Biomedical Engineering and Informatics (BMEI), 2010 3rd International Conference on.* Vol. 2. IEEE, 2010.
20. Stockwell RG, Mansinha L, Lowe RP. Localization of the complex spectrum: the S transform. *IEEE Trans Signal Process.* 1996;44:998–1001.
21. Varghees VN, Ramachandran KI. A novel heart sound activity detection framework for automated heart sound analysis. *Biomed Signal Process Control.* 2014;13:174–88.
22. Sepehri AA, et al. A novel method for pediatric heart sound segmentation without using the ECG. *Comput Methods Appl Mech Eng.* 2010;99:43–8.
23. Pedrosa J, Castro A, Vinhoza TTV. Automatic heart sound segmentation and murmur detection in pediatric phonocardiograms. In: *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE.* IEEE, 2014.
24. Wang P, et al. Phonocardiographic signal analysis method using a modified hidden markov model. *Ann Biomed Eng.* 2007;35:367–74.
25. Schmidt SE, et al. Segmentation of heart sound recordings from an electronic stethoscope by a duration dependent hidden-markov model. In: *Computers in cardiology. Bologna: IEEE; 14–17 September 2008, p. 345–48.*
26. Moukadem A, et al. Localization of heart sounds based on S-transform and radial basis function neural network. In: *15th Nordic-Baltic Conference on Biomedical Engineering and Medical Physics (NBC 2011).* Springer: Heidelberg, 2011.
27. Zhong L, et al. A robust envelope extraction algorithm for cardiac sound signal segmentation. In: *Bioinformatics and Biomedical Engineering, 2011 5th International Conference on.* IEEE, 2011.
28. Wang X, et al. Detection of the first and second heart sound using heart sound energy. In: *Biomedical Engineering and Informatics, 2009. BMEI'09. 2nd International Conference on.* IEEE, 2009.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

