Multistage decision-based heart sound delineation method for automated analysis of heart sounds and murmurs

V. Nivitha Varghees ✉, K.I. Ramachandran

Centre for Excellence in Computational Engineering and Networking, Amrita Vishwa Vidyapeetham University, Coimbatore 641 112, Tamil Nadu, India

✉ E-mail: nivithavarghees@gmail.com

Published in Healthcare Technology Letters; Received on 17th March 2015; Revised on 4th August 2015; Accepted on 15th September 2015

A robust multistage decision-based heart sound delineation (MDHSD) method is presented for automatically determining the boundaries and peaks of heart sounds (S1, S2, S3, and S4), systolic, and diastolic murmurs (early, mid, and late) and high-pitched sounds (HPSs) of the phonocardiogram (PCG) signal. The proposed MDHSD method consists of the Gaussian kernels based signal decomposition (GSDs) and multistage decision-based delineation (MDBD). The GSD algorithm first removes the low-frequency (LF) artefacts and then decomposes the filtered signal into two subsignals: the LF sound part (S1, S2, S3, and S4) and the high-frequency sound part (murmurs and HPSs). The MDBD algorithm consists of absolute envelope extraction, adaptive thresholding, and fiducial point determination. The accuracy and robustness of the proposed method is evaluated using various types of normal and pathological PCG signals. Results show that the method achieves an average sensitivity of 98.22%, positive predictivity of 97.46%, and overall accuracy of 95.78%. The method yields maximum average delineation errors of 4.52 and 4.14 ms for determining the start-point and end-point of sounds. The proposed multistage delineation algorithm is capable of improving the delineation accuracy under time-varying amplitudes of heart sounds and various types of murmurs. The proposed method has significant potential applications in heart sounds and murmurs classification systems.

1. Introduction: Phonocardiography provides vital information for diagnosis of various cardiovascular diseases [1–8]. The normal PCG signal consists of the first heart sound (S1), the systolic pause after the sound S1, the second heart sound (S2), and the diastolic pause after the sound S2 [3–5]. The other normal and abnormal heart sounds (S3 and S4), high-pitched clicks, and different types of heart murmurs may occur in the systolic and diastolic pause intervals [5]. Heart murmurs are often characterised by the timing (early, mid, or late), intensity, duration, pitch (low, medium, or high), quality (blowing, rumbling, or musical), and configurations of crescendo, decrescendo, crescendo–decrescendo [1–8]. Thus, an automated delineation method for accurate measurements of sound parameters including amplitude, frequency content, duration, systolic, and diastolic intervals, timing, and configuration of murmurs is most important for effective diagnosis of cardiovascular diseases.

Many heart sound segmentation (HSS) methods were reported based on the reference electrocardiogram (ECG) and/or carotid pulse (CP) signals [3], empirical mode decomposition (EMD) [4, 5], hidden Markov models [7], wavelet transform (WT), and wavelet packet transform [8–10], temporal–spectral features [11–14], homomorphic envelopogram, and self-organising probabilistic model [15], Hilbert transform [16, 17], support vector machine [18], and artificial neural network [19]. On the basis of the feature extraction approaches, the HSS methods can be categorised into four major groups: (i) ECG and/or CP waveform-based methods, (ii) temporal–spectral feature-based methods, (iii) envelope-based methods, and (iv) hybrid methods. In the ECG or/and CP-based methods, segmentation performance depends heavily on the accurate detection of the R-wave and T-wave in ECG signal, and the systolic peak and dicrotic notch in CP waveforms under time-varying ECG and CP morphologies and various types of artefacts and noise [20].

In the temporal–spectral-based methods [11–14], the segmentation task was performed based on the temporal features including, short-term energy, zero-crossing rate, entropy, kurtosis, autocorrelation function, and the spectral features including sub-band energy ratio, spectral flux, spectral centroid, and harmonics are used for

detecting heart sounds S1 and S2 in the PCG signal. In the envelope-based segmentation methods, one or more feature envelope(s) are extracted by using the following operations in either time-domain or transform domain [11, 14–17]: absolute, energy, Shannon energy, Shannon entropy, instantaneous amplitude, instantaneous energy, instantaneous frequency, and autocorrelation coefficient. The decision stage includes sets of thresholds for detecting the locations of candidate peaks of the envelope. Most methods employ the searchback algorithm with medical tactics to reject or include the noise or missed peaks. Furthermore, the low-level and high-level time limits are used for deleting extra sounds and finding missed sounds, respectively.

Most aforementioned methods mainly focus on segmentation of PCG signal into four components: S1–systole–S2–diastole, whereas only very few methods have been presented for detections of S3 and S4 [5, 6]. In pathological cases, the extra sounds including clicks, snaps, rubs, and various kinds of systolic and diastolic murmurs (including, early, mid, and late) may occur between the two loudest sounds S1 and S2. Most clinical studies show that different types of abnormal sounds and murmurs are often characterised by using their morphological parameters including timing, duration, intensity, frequency content, and configuration [1–8]. To the best of our knowledge, there are no systematic heart sound delineation methods that can automatically determine the boundaries and peaks of both heart sounds and murmurs present in the PCG signal. Moreover, development of an accurate and reliable delineation method is still a challenging task for PCG signals composed of low-amplitude S3 and S4 sounds, high-pitched sounds (HPSs), various kinds of murmurs, irregular heart rates, and background noises.

In this Letter, we present a robust heart sound delineation method for automatically determining the boundaries and peaks of heart sounds (S1, S2, S3, and S4), the systolic and diastolic murmurs (early, mid, and late), and HPSs of the PCG signal. The proposed heart sound delineation (HSD) method consists of two major stages: (i) Gaussian kernels based signal decomposition (GSDs) and (ii) multistage decision-based delineation (MDBD). The GSD algorithm first performs the removal of low-frequency (LF) artefacts from the PCG signal and then extracts the LF component

(S1, S2, S3, and S4) and the high-frequency (HF) component including murmurs and other HPSs. The proposed MDBD algorithm consists of four major steps: the smooth absolute envelope extraction, the adaptive envelope thresholding, the residual LF component suppression, and the fiducial point determination and adjustment. Results show that the proposed method significantly reduces the influence of LF artefacts and background noises, and achieves good delineation performance in determining low-amplitude heart sounds and various types of heart murmurs. Moreover, the method elucidates the essential clinical features that are used for recognition of normal and pathological heart sounds and murmurs. The rest of this Letter is organised as follows: Section 2 describes the proposed delineation method, section 3 presents the performance evaluation results for the normal and pathological PCG signals, finally, conclusions are drawn in Section 4.

2. Proposed heart sound delineation method: Fig. 1 shows a block diagram of the proposed method, which consists of: removal of LF artefacts, decomposing signal into the LF component including heart sounds (S1, S2, S3, and S4) and the HF component including murmurs and HPSs, selection of LF and/or HF components for the delineation process, absolute envelope extraction, adaptive thresholding, and the fiducial points determination and correction.

2.1. Gaussian kernel-based PCG signal decomposition: In real-time environments, the PCG signals may be corrupted by LF artefacts which can randomly drift the local sounds and murmurs from the reference line. The presence of LF artefacts tends to increase the number of false alarm and misdetection rates. Therefore, the removal of LF artefacts from the PCG signal is most important before extracting the signal envelope for determining the fiducial points of the PCG signal. In the literature, the PCG signal enhancements were performed using moving averaging, wavelet, and EMD-based filtering techniques [4, 5, 8–11]. By considering the overall computational complexity in real-time implementation, we attempt to design an application-specific filtering technique. In this Letter, the Gaussian kernel-based signal filtering technique is designed for removal of LF artefacts, and discriminating the LF sounds (S1, S2, S3, and S4) from the heart murmurs. The first stage of the decomposition aims to remove LF artefacts from the PCG signal. In the second stage, the proposed method constructs the two subsignals: an LF signal including the heart sounds S1, S2, S3, and S4 and an HF signal including heart murmurs and HPSs.

2.1.1 Stage I – LF artefacts removal: In this Letter, the LF artefact waveform is first computed by taking the convolution of the input

PCG signal $x[n]$ and the Gaussian kernel sequence $h[k]$. The filter coefficients are directly computed from the Gaussian kernel function, which is defined as

$$h[k] = e^{-(1/2)(\alpha k/(K/2))^2} \quad (1)$$

where $-(K/2) \leq k \leq (K/2)$, α is proportional to the reciprocal of the standard deviation (σ) which denotes the width (or spread) of the Gaussian kernel, and the length of Gaussian kernel is $M = K + 1$. A smaller value of σ produces a narrower Gaussian window. The LF artefact waveform $b[n]$ is computed as

$$b[n] = \sum_{k=-\infty}^{\infty} x[k]h[n-k]. \quad (2)$$

Here, the Gaussian kernel has a length of 201 samples and α of 2. The estimated LF waveform is subtracted from the input PCG signal $x[n]$ as

$$f[n] = x[n] - b[n], \quad n = 0, 1, 2, 3, \dots, N-1, \quad (3)$$

where $f[n]$ denotes the filtered PCG signal. The results of the LF artefacts removal stage are shown in Fig. 2. It is noted that the Gaussian filter adequately removes the LF artefacts from the PCG signal as compared with filter with rectangular impulse response. Then, the normalised filtered signal is further processed at the second stage of the signal decomposition structure.

2.1.2 Stage II – obtaining LF and HF parts: In normal and pathological heart sound signals, the recorded PCG signal is composed of LF component (including heart sounds S1, S2, S3, and S4), and the HF component (including the systolic and diastolic murmurs, and other HPSs). In such scenarios, determination of boundaries and peaks of the heart sounds (S1, S2, S3, and S4) is a challenging task in the presence of different types of heart murmurs. Therefore, we implement the signal decomposition structure for the separation of the LF component from the HF component part. In the second stage of the proposed delineation method as shown in Fig. 1, the Gaussian smoothing filter (GSF) acts as a low-pass filter whereas the Gaussian derivative filter acts as a band-pass filter. The choice of alpha value plays an important role in adequately extracting the significant components of heart sounds (S1, S2, S3, and S4) and suppressing the heart murmurs which are greater than the frequency of 150 Hz and background HF noises, and also for adequately preserving the sound components of heart murmurs (150–600 Hz) and suppressing the HF noises above the

Fig. 1 Simplified block diagram of the proposed delineation method

Fig. 2 Illustrates the effectiveness of the LF artefact removal method based on the Gaussian filter and the moving average filter

frequency of 600 Hz [2–5, 7, 8]. In this Letter, a Gaussian smoothing filtering (GSF) approach is implemented using the filter coefficients derived from 25-point Gaussian kernel with α of 1. The purpose of the GSF step is to accurately determine the end-points of the heart sounds (S1, S2, S3, and S4) in the presence of heart murmurs (early, mid, late, and continuous) and HPSs. Gaussian derivative filtering (GDF) is implemented using the filter coefficients derived from 25-point Gaussian kernel with α of 5. The filtered PCG signal is obtained by taking convolution of a PCG signal with filtered coefficients derived from Gaussian derivative kernel, which acts as band-pass filters. The purpose of implementing the GDF approach is to obtain the components of heart systolic and diastolic murmurs and HPSs including clicks, snaps, and rubs so that their end-points can be accurately determined in presence of the prominent LF heart sounds (S1, S2, S3, and S4).

At this stage, the filtering capability of the proposed GSF and GDF approaches is studied for two different PCG signals including the loudest sounds S1 and S2, pan-systolic murmurs as shown in Figs. 3a1 and b1, respectively. Figs. 3a2 and b2 show the effectiveness of the GSF approach in adequately emphasising the heart sounds (S1, S2, S3, and S4) and suppressing the influence of heart murmurs, HPSs and background noises. Thus, the proposed delineation method can accurately determine the end-points of the S1, S2, S3, and S4 by processing the LF component part $x_L[n]$. Similarly, the HF component parts obtained using the GDF approach are shown in Figs. 3a3 and b3. From these results, we can observe that the GDF approach adequately emphasises the pan-systolic murmurs while significantly suppresses LF components of loudest heart sounds S1 and S2. Figs. 3a3 and b3 clearly show that the proposed GDF approach can enable accurate determination of the end-points of heart murmurs by processing the HF component part $x_H[n]$.

In some normal and pathological cases, the original PCG signal may contain only the loudest heart sounds S1 and S2 and other LF

sounds S3 and S4. In such cases, the HF component part has smaller amplitude dynamic range that may show absence of heart murmurs and other HF abnormal sounds. Furthermore, the original PCG signal may contain only the HF sounds and murmurs. In such cases, the LF component part has smaller amplitude dynamic range. Therefore, in this Letter, we implement the decision rule with relative sum of absolute amplitude ratio (RSAAR) criterion for deciding whether to perform delineation process on the extracted LF component part or/and the extracted HF component part.

For the LF component $x_L[n]$ and HF component $x_H[n]$, the RSAAR values η_L and η_H are computed as

$$\eta_L = \frac{A_L}{A_F} \quad \text{and} \quad \eta_H = \frac{A_H}{A_F}, \quad (4)$$

where A_L denotes the sum of absolute of the amplitudes of the thresholded LF component part $\bar{x}_L[n]$, A_H denotes the sum of absolute of the amplitudes of the HF component part $\bar{x}_H[n]$, and A_F denotes the sum of absolute of the amplitudes of the normalised signal $y[n]$ that is computed as $A_F = \sum_{n=0}^{N-1} |y[n]|$. To further reduce the influence of residual noise components, the LF component and HF component parts are processed with fixed thresholding with amplitude thresholds of 0.05 and 0.02, respectively. In this Letter, we assume that the amplitudes of sounds S1, S2, S3, and S4, murmurs and HPSs are larger than the predetermined thresholds within a normalised amplitude range from -1 to $+1$. Then, the A_L and A_H are computed as

$$A_L = \sum_{n=0}^{N-1} |\bar{x}_L[n]| \quad \text{and} \quad A_H = \sum_{n=0}^{N-1} |\bar{x}_H[n]|. \quad (5)$$

Finally, the computed RSAAR values η_L and η_H are compared with the predefined thresholds η_{thL} and η_{thH} , respectively. For identifying

Fig. 3 Illustrates the filtering capabilities of the GSF and GDF schemes for PCG signals with pan-systolic murmur, open snap, and background noises

the presence of heart sounds (S1, S2, S3, and S4) in the LF component, the decision rule is defined as: if $\eta_L > 0.2$, store LF component for further delineation as described in Section 2.2, otherwise discard the LF component from further processing. For identifying the presence of heart murmurs and other high-pitched heart sounds in the HF component, the decision rule is defined as: if $\eta_H > 0.1$, store HF component part for further delineation; otherwise, discard the HF component from further processing. The outputs of the proposed GSF and GDF stages are shown in Fig. 3b for abnormal PCG signal with diastolic aortic insufficiency. For the decomposition results as shown in Figs. 3b2 and b3, it is noted that the relative sum of absolute of amplitudes ratios η_L and η_H computed for the LF component and HF component parts are 0.025 and 0.7791, respectively. In such a case, the LF component is discarded from further processing since $\eta_L < 0.2$. Furthermore, the proposed method generates a PCGLF flag of 0 for LF component and a PCGHF flag of 1 for HF component. For the decomposition results as shown in Fig. 3a, the method generates a flag of 1 for both LF component and HF component parts. In this Letter, the relative amplitude thresholds are chosen based on the dynamic range of the amplitude distribution of the heart sounds and murmurs contained in the normalised PCG signal. On the basis of the PCGLF flag and PCGHF flag information, the LF component and HF component signals will be further processed at the delineation stage.

2.2. MBD algorithm: On the basis of the PCG flag information, the proposed delineation algorithm processes either only the LF component (or the HF component) or both LF and HF components. If the PCGLF flag (or the PCGHF flag) is 1, the delineation process is performed on the LF component (or the HF component). If both PCGLF flags are 1, the LF component is first processed for determining boundaries of heart sounds (S1, S2, S3, and S4). Then, the detected heart sound boundaries are used for setting those sound portions in the HF component part to zero before performing the delineation process on the HF component part. The proposed multistage decision-based PCG delineation

algorithm comprises the steps of: the smooth absolute envelope extraction, the threshold determination, the adaptive envelope thresholding, finding difference of binary output, determining boundary and peak instants, applying boundary and peak adjustment procedure, and outputting the boundary and peak parameters for further analysis of heart sounds and murmurs. Each step of the delineation algorithm is described as follows.

2.2.1 Smooth absolute envelope extraction: Unlike other segmentation methods with complex non-linear peak amplification techniques including the energy, Shannon energy, instantaneous frequency, and Hilbert envelope, the proposed method uses a simple absolute operation for extracting the signal envelope from the LF component and HF component parts. Here, the envelope smoothing process is implemented by using a linear zero-phase filtering with a rectangular impulse response, $h(k)$ of length L . The smoothing filter length is adopted based on the PCG flag information received from the previous step. The window durations of 20 and 50 ms are used for extracting the signal envelope from the LF component and HF component parts, respectively. The outputs of the envelope extraction stage are shown in Figs. 4c and g and Figs. 5c and f for different PCG signals. It shows that the envelope extraction step provides a smooth signal envelope $a[n]$ with local peaked-waves around heart sound portions.

2.2.2 Adaptive envelope thresholding: Using the absolute envelope signal $a[n]$, the end-points of the heart sounds and murmurs are determined using the adaptive envelope thresholding rule. For each signal, the amplitude threshold is adapted based on the median of the local mean estimates of the short-term absolute envelope signal $a[n]$. First, the envelope signal $a[n]$ is divided into non-overlapping segments with duration of 10 ms. The mean of the p th short-term segment $\bar{a}_p[n]$ is computed as

$$\mu_p = \frac{1}{N_s} \sum_{n=1}^{N_s} \bar{a}_p[n], \quad p = 1, 2, 3, \dots, P \quad (6)$$

Fig. 4 Illustrates the outputs of the proposed delineation method

a Original PCG signal including the heart sounds S1, S2, and pan-systolic events

b Output of the Gaussian smoothing filtering step

c Output of the absolute envelope extraction step for the LF component

d Output of the adaptive envelope thresholding step

e Output of the GDF step

f Output of the LF component portions removal step after setting the detected S1 and S2 portions to zero in the HF component

g Output of the absolute envelope extraction step

h Output of the adaptive envelope thresholding step

where P denotes the number of segments and N_s denotes the number of samples within a segment. For the estimated mean vector $\boldsymbol{\mu}$, the adaptive amplitude threshold is computed as

$$\eta_a = \eta_{\text{nf}} + \text{median}(\{\boldsymbol{\mu}_p\}_{p=1}^P) \quad (7)$$

where η_{nf} denotes the noise floor level which is set as 0.02. For the PCG signals, the estimated amplitude threshold ('Red Line') is plotted in Figs. 4c and g and Figs. 5c and f. By using the estimated

threshold η_a , the signal envelope $a[n]$ is thresholded as

$$b[n] = \begin{cases} 1, & a[n] > \eta_a \\ 0, & \text{otherwise.} \end{cases} \quad (8)$$

where $b[n]$ denotes the binary gate signal used for determining boundaries of the local sounds and murmurs present in the PCG signal. The outputs of the adaptive envelope thresholding stage are shown in Figs. 4d and h and Figs. 5d and g. From these

Fig. 5 Illustrates the effectiveness of the proposed delineation method

a Original PCG signal including the sounds S1, S2, and open snap events

b Output of the Gaussian smoothing filtering step

c Output of the absolute envelope extraction step for the LF component

d Output of the adaptive envelope thresholding step

e Output of the LF component portions removal step after setting the detected S1 and S2 portions to zero in the HF component

f Output of the absolute envelope extraction step for the signal in (e)

g Output of the adaptive envelope thresholding step

h Output of the fiducial point determination step

results, it is noted that the proposed threshold estimation procedure is capable of detecting the heart sounds (S1, S2, S3, and S4), murmurs, and other HPSs under the influence of residual background noises.

2.2.3 LF component portions removal: The PCG signal is an additive mixture of heart sounds and murmurs. Thus, the HF component part may include the residual HF components within the corresponding portions of the heart sounds (S1, S2, S3, and S4). In such a case, the proposed method first detects the sound portions in the LF component and then the detected sound portions in the HF component are set to zero if both PCGLF flag and PCGHF

flag values are 1. Otherwise, the delineation algorithm processes either LF component or HF component. The outputs of this stage are shown in Figs. 4e and 5e for two PCG signals including pansystolic murmur and high-amplitude open snap events.

2.2.4 Fiducial point determination and correction: By taking the difference of the binary gate signal $b[n]$, the start-points and end-points of the local sounds are determined by detecting the positive ones and negative ones in the difference of binary gate $b[n]$. The detected boundaries are used as guides for correcting the boundaries of the local sounds present in the PCG signal. In this Letter, two decision rules are used to reject non-heart sound segments and

merge two segments. For rejecting the falsely detected non-heart sound segments, the duration and peak amplitude values are computed for each of the detected sound segments and then compared with the predefined duration threshold of 5 ms and amplitude threshold of 0.01, respectively. The interval threshold of 10 ms is used for merging the two heart segments. If the interval between the two consecutive detected segments is less than the threshold of 10 ms, then the two successive segments are merged. The segment merging rule is implemented for the LF component, and HF component separately. The boundary adjustment procedure is implemented with the detected sound boundaries and the filtered PCG signal $f[n]$. Finally, the peak of each of the detected sound segments is computed by finding the maximum of the absolute of the original signal amplitudes within detected segment duration. The final output of the proposed delineation method is shown in Fig. 5h. The delineation results show that the method accurately determines the boundaries and peaks of the heart sounds S1 and S2 and high-pitched open snap present after the S2. On the basis of the results as shown in Figs. 4 and 5, it is noted that the proposed multistage decision algorithm is capable of accurately determining boundaries and peaks of all the normal and abnormal heart sounds and murmurs contained in the input PCG signal.

3. Results and discussion: In this section, we study the delineation accuracy and robustness of the proposed method using a wide variety of clean and noisy PCG signals taken from the eGeneralMedical, Michigan heart sound and murmur, and bioscience heart sound databases, which are used for performance evaluations [11, 17, 21–23]. The test PCG database includes 125 files containing different types of normal and pathological heart sounds (S1, S2, S3, and S4), systolic murmurs (early, mid, and late), diastolic murmurs (early, mid, and late), pan-systolic and pan-diastolic murmurs, continuous murmurs, and HPSs including split sounds, clicks, snaps, and rubs. The PCG signals were digitised with different sampling rates and quantisation bits. The primary purpose of using different sampling rates is to study the robustness of the optimal filtering and threshold parameters used in algorithmic components of the proposed method. It is noted that some PCG signals are corrupted with LF artefacts and HF noises. The PCG signals are corrupted with synthetically generated white Gaussian noises with signal-to-noise ratios (SNRs) ranging from 20 to 5 dB for validating the robustness of the proposed delineation method.

For performance evaluations, we created a ground-truth annotation file including the start-time and end-time instants of true sound segments, the total number of true sound segment. From the detection results, we compute three quantitative parameters: true positive (TP) denotes the number of segments which are correctly detected by the method, false negative (FN) denotes the number of segments which are missed, and false positive (FP) denotes the number of segments which are wrongly detected by the method. The performance is evaluated in terms of three metrics: (i) the sensitivity (Se) = $TP/(TP + FN)$; (ii) the positive predictivity (+P) = $TP/(TP + FP)$; and (iii) the overall accuracy (OA) = $TP/(TP + FN + FP)$. The most crucial metric for determining overall system performance is the absolute time error, which is computed between the actual value and measured value for each of the fiducial attributes. Here, the average absolute errors for each of the delineation attributes (including start-time and end-time) are computed to validate the delineation accuracy of the proposed method under different types of normal and pathological PCG signals and background noises.

In the first experiment, we study the importance of the boundary correction step including the segment merging and rejection rules. The detection results for the proposed method without (W/O) and with (W/) boundary correction procedures are summarised in Table 1 for different SNR values of 20, 15, 10, and 5 dB. From the visual inspection of detection results, it is noted that the

Table 1 Performance of the proposed delineation method W/O and W/ boundary correction procedures

SNR, dB	W/O correction				W/ correction			
	FN	FP	TP	OA, %	FN	FP	TP	OA, %
20	36	86	2821	95.85	36	64	2821	96.58
15	41	97	2816	95.33	43	68	2814	96.21
10	52	114	2805	94.41	56	76	2801	95.50
5	67	128	2790	93.47	68	84	2789	94.83

proposed method results in a large number of FP detections for the PCG signals including the diastolic patent ductus arteriosus, pericardial friction rub, systolic mitral prolapse, systolic mitral valve replacement, aortic stenosis, severe systolic aortic stenosis, and systolic aortic valve replacement events. The proposed delineation method with boundary correction procedure can significantly improve the OA, especially by reducing the FPs for different types of heart murmurs and background noises including the short-duration impulsive noises.

In the second experiment, the detection performance of the proposed method is compared with the standard wavelet-based decomposition method under different SNR values of 20, 15, 10, and 5 dB. In this Letter, the stationary WT (SWT) is implemented with 9/7 biorthogonal wavelet filters [24] and decomposition level of 8. From the decomposition results, we observed that the higher detail sub-bands D1 and D2 contain the major amplitude components of the background noise. The detail sub-bands D3, D4, and D5 include the major components of the heart murmurs and the HF portions of heart sounds S1 and S2. The detail sub-bands D6, D7, and D8 contain the LF component portions of the heart sounds (S1, S2, S3, and S4). The eighth approximation sub-band includes the LF artefact of the PCG signal. The LF component is constructed using the detail sub-bands D6, D7, and D8 and the HF component is constructed using the detail sub-bands D3, D4, and D5. Then, the proposed algorithm is employed for performing delineation process.

Table 2 summarises the number of FPs, number of FNs and the number of TPs of the three HSD methods for each SNR case. Performance evaluation results show that the proposed method has 68 FNs and 84 FPs for a total of 2857 segments for the PCG signals with SNR value of 5 dB. The proposed method achieves an average Se of 97.62%, +P of 97.08%, and OA of 94.83%, whereas the wavelet decomposition-based method results in an average Se of 95.38%, +P of 93.84%, and OA of 89.76%. On the basis of our results, it is observed that the proposed method achieves an accuracy of 100% in detecting the two loudest heart sounds S1 and S2 and the different types of systolic and diastolic murmurs (early, mid, late, and pan). The method had a missed detection of 36 for a total of 250 segments of the low-amplitude S4, and the snaps and click events. The method had poor boundary detection rates for severe aortic regurgitation, pericardial friction rub, and patent ductus arteriosus events.

Table 2 Performance comparison of detection results

SNR, dB	Method	FN	FP	TP	Se, %	+P, %	OA, %
20	SWT	91	126	2766	96.81	95.64	92.73
	proposed	36	64	2821	98.74	97.78	96.58
15	SWT	104	134	2753	96.36	95.36	92.04
	proposed	43	68	2814	98.49	97.64	96.21
10	SWT	118	158	2739	95.87	94.55	90.85
	proposed	56	76	2801	98.04	97.36	95.50
5	SWT	132	179	2725	95.38	93.84	89.76
	proposed	68	84	2789	97.62	97.08	94.83

Finally, we study the delineation performance of the proposed method in accurately determining the boundary points that are widely used in measuring the clinical parameters (the amplitude S1/S2 ratio and the diastolic to systolic duration ratio) and the classification of heart sounds, different types of systolic and diastolic murmurs, and other HPSs. For each PCG signal, the delineation accuracy is computed by comparing the estimated values of delineation attributes with their actual values in the ground-truth annotation file. The delineation results are summarised in Table 3. From the results, it is noted that the most detected segments include the complete heart sounds and murmurs in addition with small pause duration, which is introduced due to the envelope smoothing process. However, the proposed method achieves acceptable delineation accuracy for heart sounds, different types of murmurs, and other HPSs.

Unlike other existing methods [14, 19], the proposed method does not employ the searchback algorithm with sets of decision thresholds determined using the past detected segment durations and peaks for identifying the missed sound segments. Thus, the proposed method is quite straightforward and does not require any learning phase for accurately determining initial threshold parameters. The WT-based method requires the eight level decomposition for removal of LF artefacts, whereas the proposed method employs only three-stage filtering structure. The proposed method has an average coding delay of 237 ms, whereas the WT-based method has coding delay of 338 ms for 2 s signal sampled with rate of 4000 Hz. On the basis of our results, we believe that the proposed delineation method is more suitable for an automated heart sound analysis, PCG signal enhancement, and heart sound biometric systems.

4. Conclusion: In this Letter, we present a robust multistage decision-based heart sound delineation method for automatically determining the boundaries and peaks of heart sounds (S1, S2, S3, and S4), murmurs (early, mid, and late) and other HPSs of the PCG signal. The proposed method consists of the GSDs and the MDBD including absolute envelope extraction, adaptive thresholding, and fiducial point determination. The accuracy and robustness of the proposed method is evaluated using various normal and pathological PCG signals under both clean and noisy environments. Evaluation results show that the method achieves an average Se of 98.22%, +P of 97.46%, and OA of 95.78%. Results further show that the method efficiently improves the delineation accuracy and also elucidates essential clinical features. Thus, the proposed method can serve as a delineation tool for analysis and diagnosis of normal and pathological heart sounds and murmurs.

5. Funding and declaration of interests: Conflict of interest: none.

6 References

- [1] Walker H.K., Hall W.D., Hurst J.W.: 'Clinical methods: the history, physical, and laboratory examinations' (Butterworths, Boston, 1990, 3rd edn.)
- [2] Rangayyan R.M.: 'Biomedical signal analysis: a case-study approach' (Wiley-IEEE Press, 2001)
- [3] Lehner R.J., Rangayyan R.M.: 'A three-channel microcomputer system for segmentation and characterization of the phonocardiogram', *IEEE Trans. Biomed. Eng.*, 1987, **BME-34**, (6), pp. 485–489
- [4] Papadaniil C., Hadjileontiadis L.: 'Efficient heart sound segmentation and extraction using ensemble empirical mode decomposition and kurtosis features', *IEEE J. Biomed. Health Inf.*, 2014, **18**, (4), pp. 1138–1152
- [5] Tseng Y.-L., Ko P.-Y., Jaw F.-S.: 'Detection of the third and fourth heart sounds using Hilbert–Huang transform', *Biomed. Eng. Online*, 2012, **11**, (8), pp. 1–13
- [6] Ahlstrom C., Huit P., Ask P.: 'Detection of the 3rd heart sound using recurrence time statistics'. Proc. ICASSP, 2006, pp. 1040–1043
- [7] Kwak C., Kwon O.-W.: 'Cardiac disorder classification by heart sound signals using murmur likelihood and hidden Markov model state likelihood', *IET Signal Process.*, 2012, **6**, (4), pp. 326–334
- [8] Boutana D., Benidir M., Barkat B.: 'Segmentation and identification of some pathological phonocardiogram signals using time-frequency analysis', *IET Signal Process.*, 2011, **5**, (6), pp. 527–537
- [9] Wang Y., Lia W., Zhou J., *ET AL.*: 'Identification of the normal and abnormal heart sounds using wavelet-time entropy features based on OMS-WPD', *Future Gener. Comput. Syst.*, 2014, **37**, pp. 488–495
- [10] Safara F., Doraisamy S., Azman A., *ET AL.*: 'Multi-level basis selection of wavelet packet decomposition tree for heart sound classification', *Comput. Biol. Med.*, 2013, **43**, (10), pp. 1407–1414
- [11] Sabarimalai Manikandan M., Soman K.P.: 'Robust heart sound activity detection in noisy environments', *Electron. Lett.*, 2010, **46**, (16), pp. 1100–1102
- [12] Iwata A., Ishii N., Suzumura N., *ET AL.*: 'Algorithm for detecting the first and the second heart sounds by spectral tracking', *Med. Biol. Eng. Comput.*, 1980, **18**, (1), pp. 19–26
- [13] Haghighi-Mood A., Torry J.N.: 'A sub-band energy tracking algorithm for heart sound segmentation', *Comput. Cardiol.*, 1995, pp. 501–504
- [14] Liang H., Lukkarinen S., Hartimo I.: 'Heart sound segmentation algorithm based on heart sound enveloplogram', *Comput. Cardiol.*, 1997, **24**, pp. 105–108
- [15] Gill D., Gavrieli N., Intrator N.: 'Detection and identification of heart sounds using homomorphic envelopogram and self-organizing probabilistic model', *Comput. Cardiol.*, 2005, **32**, pp. 957–960
- [16] Alajarin J.M., Merino R.R.: 'Efficient method for events detection in phonocardiographic signals', *Proc. SPIE Bioengineered Bioinspired Syst. II*, 2005, **5839**, pp. 398–409
- [17] Nivitha Varghees V., Ramachandran K.I.: 'A novel heart sound activity detection framework for automated heart sound analysis', *Biomed. Signal Process. Control*, 2014, **13**, pp. 174–188
- [18] Choi S., Jiang Z.: 'Cardiac sound murmurs classification with autoregressive spectral analysis and multi-support vector machine technique', *Comput. Biol. Med.*, 2010, **40**, (1), pp. 8–20
- [19] Olmez T., Dokur Z.: 'Classification of heart sounds using an artificial neural network', *Pattern Recognit. Lett.*, 2003, **24**, pp. 617–629
- [20] Kathirvel P., Manikandan M.S., Prasanna S.R.M., *ET AL.*: 'An efficient R-peak detection based on new nonlinear transformation and first-order Gaussian differentiator', *Cardiovasc. Eng. Technol.*, 2012, **2**, (4), pp. 408–425
- [21] eGeneralMedical heart murmur database. Available at <http://www.egeneralmedical.com/listohearmur.html>
- [22] Michigan heart sound and murmur library. Available at <http://www.med.umich.edu/lrc/psb/heartounds/>
- [23] Bioscience heart sound databases. Available at http://int-prop.lf2.cuni.cz/heart_sounds/h14/sound.htm
- [24] Manikandan M.S., Soman K.P., Dandapat S.: 'Quality-driven wavelet based PCG signal coding for wireless cardiac patient monitoring'. Proc. First Int. Conf. Wireless Technologies for Humanitarian Relief, 2011, pp. 519–526