Wavelet-based fundamental heart sound recognition method using morphological and interval features

V. Nivitha Varghees, K.I. Ramachandran, K.P. Soman

Center for Computational Engineering and Networking (CEN), Amrita School of Engineering, Amrita University, Amrita Vishwa Vidyapeetham, Coimbatore, India

✉ E-mail: v_nivitha@cb.amrita.edu

Published in Healthcare Technology Letters; Received on 21st December 2016; Revised on 21st June 2017; Accepted on 28th June 2017

Accurate and reliable recognition of fundamental heart sounds (FHSs) plays a significant role in automated analysis of heart sound (HS) patterns. This Letter presents an automated wavelet-based FHS recognition (WFHSR) method using morphological and interval features. The proposed method first performs the decomposition of phonocardiogram (PCG) signal using a synchrosqueezing wavelet transform to extract the HSs and suppresses the murmurs, low-frequency and high-frequency noises. The HS delineation (HSD) is presented using Shannon energy envelope and amplitude-dependent thresholding rule. The HFS recognition (HFSR) is presented using interval, HS duration and envelope area features with a decision-rule algorithm. The performance of the method is evaluated on PASCAL HSs Challenge, PhysioNet/CinC HS Challenge, eGeneralMedical databases and real-time recorded PCG signals. Results show that the HSD approach achieves an average sensitivity (Se) of 98.87%, positive predictivity (Pp) of 97.50% with detection error rate of 3.67% for PCG signals with signal-to-noise ratio of 10 dB, and outperforms the existing HSD methods. The proposed WFHSR method achieves a Se of 99.00%, Sp of 99.08% and overall accuracy of 99.04% on both normal and abnormal PCG signals. Evaluation results show that the proposed WFHSR method is able to accurately recognise the S1/S2 HSs in noisy real-world PCG recordings with murmurs and other abnormal sounds.

1. Introduction: Accurate and reliable recognition of the fundamental heart sounds (FHSs) in a cardiac cycle of the phonocardiogram (PCG) signal plays a vital role in valvular split, cardiac stress, pulmonary artery pressure analysis, left ventricular pressure rise (LV dP/dt) measure, non-invasive blood pressure estimation and human identification [1–5]. In the past years, many S1/S2 sound recognition methods were presented based on variety of preprocessing and feature extraction techniques and classifiers such as logistic regression-hidden semi-Markov model (HSMM) and wavelet transform (WT) [4], mel-frequency cepstral coefficients (MFCCs) and deep neural networks (DNNs) [6], multifractal decomposition [7], deep convolutional neural network (CNN) and support vector machine (SVM) [8], ensemble empirical mode decomposition (EEMD) and kurtosis features [9], total variation and Shannon energy (SE) [10], S-transform [11], empirical mode decomposition (EMD) WT algorithm with SE [12], Hilbert transform and adaptive thresholding [13], expert frequency-energy based metric [14], matching pursuit [15], duration-dependent hidden Markov model (HMM) [16], sound energy [17], high-frequency signatures (HFSs) [18], homomorphic envelopogram and self-organising probabilistic model [19], adaptive sub-level tracking and Shannon-energy tracking [20], probabilistic models [21], time-delay neural network (NN) [22], NN and conventional classifiers [23], and spectral tracking [24].

Chen et al. [6] proposed S1 and S2 recognition algorithm based on the MFCCs and DNNs. Thomas et al. [7] proposed the multifractal property-based method to identify S1 and S2 heart sounds (HSs). Tschannen et al. [8] proposed a robust method for HS classification that combines a deep CNN-based feature extractor and an SVM. Papadantzil and Hadjiileontiadis [9] proposed HS segmentation and extraction using EEMD and kurtosis features. Varghees and Ramachandran [10] proposed HS activity detection using the total variation filtering. Shannon entropy and Hilbert transform. Moukadem et al. [11] presented hybrid-state system (HSS) method based on the S-transform. Sun et al. [12] presented an improved EMD-wavelet algorithm with SE envelope for extraction of the first and second HSs. The boundary estimation of S1 and S2 based on Hilbert transform and adaptive thresholding approach was presented in [13]. Naseri and Homaeinezhad [14] proposed a detection and boundary identification of PCG sounds using an expert frequency-energy based metric. Hedayioglu [15] proposed a segmentation of the second HS using matching pursuit. Schmidt et al. [16] presented a duration-dependent HMM-based HFS approach. Wang et al. [17] proposed HS energy-based detection of the S1 and S2 HSs. Kumar et al. [18] proposed HFSs-based method for detection of S1 and S2 HSs. Gill et al. [19] proposed an algorithm for detection and identification of HSs using homomorphic envelopogram and self-organising probabilistic model. Gamero and Watrous [21] proposed a S1 and S2 HS detection method using probabilistic models. Oskiper and Watrous [22] presented S1 HS detection method using a time-delay NN. The performance of NN and conventional classifiers such as linear discriminant, nearest neighbour, Naive Bayes was studied to distinguish between S1 and S2 HSs [23]. Some previous works showed that the EMD is a computationally expensive algorithm, which had computation time of 35.32 s to segment a HS of duration 1 s. Although the S1/S2 recognition is relatively simple in noise-free PCG signals, accurate and reliable recognition of the S1 and S2 HSs is still challenging task in the presence of complex heart murmurs and other high-pitched sounds and various of kinds of noise sources.

In this Letter, we present a wavelet-based FHS recognition (WFHSR) method using the morphological and interval features. This Letter addresses an accurate detection, localisation and recognition of S1 and S2 HSs from healthy and pathological PCG signals recorded in both noise-free and noisy recording conditions without the use of a reference signal. The key contributions of this Letter are: SSWST-based HS signal extraction approach; HS delineation (HSD) approach for automatically extracting interval, sound duration and envelope area features; and a decision-rule based FHS recognition (FHSR) approach. The proposed method is rigorously evaluated using the PASCAL HSs Challenge, PhysioNet/CinC HS (PCCHS) Challenge, eGeneralMedical (EGM) databases and real-time recorded PCG signals.
The remaining of this Letter is organised as follows. Section 2.1 briefly summarises the theory of synchrosqueezing WT (SSWT). Section 2.2 presents the proposed FHSR. Section 3 presents the evaluation results on both normal and abnormal PCG recordings. Finally, conclusions are drawn in Section 4.

2. Materials and method

2.1. Synchrosqueezing WT: Daubechies et al. [25] developed a new time–frequency (TF) analysis method called synchrosqueezing WT (SSWT). Synchrosqueezing adds to an existing transform by examining the local oscillations with respect to time that re-assigns each value of the original transform to a new frequency on the TF plane. Thus, the resulting TF representation is sharpened in the frequency domain. The SSWT aims to combine the advantages of the continuous WT (CWT) with the sharpening provided by the synchrosqueezing. Due to the high-resolution property of SSWT, it is becoming more and more popular for analysing multicomponent signals with oscillating modes. The SSWT provides an adaptive TF decomposition. In this Letter, we investigate the effectiveness of the SSWT for separating the HSs from heart murmurs and other low-frequency and high-frequency noises and extracting the instantaneous frequencies for recognising the FHSs. The theory of SSWT is briefly summarised as follows [25, 26].

The wavelet function \( \psi(t) \) has to be translated in time by \( u \) and scaled by \( s > 0 \) to obtain a family of wavelet functions \( \psi_{u,s}(t) \):

\[
\psi_{u,s}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right). \tag{1}
\]

The CWT of a function \( x(t) \) can be defined as

\[
W_f(u,s) = \int_{-\infty}^{\infty} x(t) \psi^*_u(t) dt, \tag{2}
\]

where \( \psi^*_u(t) \) denotes a complex conjugate of the wavelet \( \psi_{u,s}(t) \). The scale \( s \) and the angular velocity \( \omega(t) \) are related via the frequency modulation \( \eta \) as

\[
\omega(t) = \frac{\eta}{s}. \tag{3}
\]

The instantaneous frequency of \( W_f(u,s) \) for signal \( x(t) \) can be obtained by

\[
\omega(t) = -i \frac{\partial}{\partial u} W_f(u,s). \tag{4}
\]

The synchrosqueezing transform \( T_x(u,\omega) \) can be determined only at the centres \( \omega_l \) of the successive bins \([\omega_l - (1/2)\Delta \omega, \omega_l + (1/2)\Delta \omega]\) with \( \omega_l - \omega_{l-1} = \Delta \omega \) by summing different contributions

\[
T_x(u,\omega) = (\Delta \omega)^{-1} \sum_{s_l+|\omega_l - \omega| < \Delta \omega / 2} W_f(u,s_l)h^{3/2} \Delta s \tag{5}
\]

where \( \Delta \omega \) denotes the width of each frequency bin. The synchrosqueezing algorithm consists of the following three steps:

- **Step 1:** The CWT is computed for the discrete time signal with time \( u \) and scale \( s \) according to 2. The analytic wavelet is used to capture instantaneous frequency information. In this work, Morlet wavelet is used as mother wavelet which is defined as

\[
\psi(t) = \pi^{-(1/4)} e^{-t^2/2} e^{i\omega t} \tag{6}
\]

- **Step 2:** The extraction of instantaneous frequencies is performed from the CWT, \( W_f \), using a phase transform, \( \omega_f \). The phase transform is proportional to the first derivative of the CWT with respect to the translation \( u \)

\[
\omega_f(u,s) = \frac{u W_f(u,s)}{2 \pi W_f(u,s)} \tag{7}
\]

- **Step 3:** The resulting instantaneous frequency value is reassigned to a single value at the centroid of the CWT TF region. This re-assignment results in sharpened output from the synchrosqueezed transform as compared with the CWT.

2.2. Proposed fundamental heart recognition algorithm: The proposed wavelet-based FHSR (WFHHR) algorithm consists of three major stages. In the first stage, the HSs are extracted from the PCG signal by suppressing the heart murmurs and other external noises using SSWT algorithm. The second stage performs delineation of extracted HS signal using SE envelope and adaptive thresholding rule. In the third stage, recognition of FHSs is performed using morphological and interval features that are computed from the delineation parameters of extracted HSs (Fig. 1).

2.2.1. SSWT-based HS extraction: In practise, the PCG signals are often corrupted with external noises. Further, the pathological PCG signals include different patterns of heart murmurs and other pathological HSs. Hence, suppression of heart murmurs and background noises has become an essential preprocessing step. In recent years, the SSWT has shown promising results for high-resolution TF analysis of non-stationary signals. As compared with EMD and variants such as ensemble EMD and complete ensemble EMD, the SSWT-based decomposition method can reconstruct individual components from selected frequency bands. In this Letter, we investigate the application of SSWT for suppression of heart murmurs, high-pitched pathological sounds and background noises.

The SSWT provides an adaptive TF decomposition that decomposes the signal into well-separated intrinsic mode type (IMT) components \( s[n] \). The instantaneous frequencies of the different intrinsic mode components are obtained from the decomposition process. In this Letter, we use the instantaneous frequencies and maximum absolute amplitudes for selecting intrinsic mode components for extracting HS signal and suppression of heart murmurs and background noises. The HS signal extraction stage includes the following steps:

- Perform PCG signal decomposition using SSWT with Morlet wavelet function.
- Obtain the IMT components \( s[n] \) and their instantaneous frequencies.
- Construct the HS signal using the IMT components which are having the instantaneous frequencies between 5 and 150 Hz and the maximum amplitudes >0.1.

Effectiveness of the SSWT is studied using different kinds of normal and pathological PCG signals. The decomposition results of this Letter are shown in Fig. 2. From the decomposition results, it is noted that the proposed HS signal extraction approach

![Proposed FHSR method](image-url)
adequately preserves the HS components and suppresses the heart murmurs and background noises.

A. SE envelope extraction: In practise, the dynamic amplitude of the PCG signals may vary for different subjects under normal and pathological environments. In this Letter, the absolute signal amplitude is normalised from 0 to 1. The normalised absolute signal $a[n]$ is computed as

$$a[n] = \frac{|x_{\text{norm}}[n]|}{\max_{n=1}^{N}(|x_{\text{norm}}[n]|)},$$  \hspace{1cm} (8)

where $N$ denotes the number of samples. For PCG signals with murmurs, the absolute HS signal is shown in Fig. 3c. From the results, it is noted that the absolute operation provides a linear amplification to amplitude of the HSs and other noise components. In such scenarios, the accumulation of absolute amplitudes of noise components can lead to produce a noisy HS envelope and it may produce more number of false positives. Thus, in this Letter, we employ SE-based non-linear peak amplification approach which can magnify medium HS amplitudes and diminishes low-amplitude noise components. Consequently, the SE operation results in small peak deviations between consecutive local sound envelopes by reducing magnitudes of large amplitude HSs that makes detection process much easier using a amplitude-dependent thresholding rule. The SE of the signal $e[n]$ is computed as

$$e[n] = -e[n] \log(e[n]),$$  \hspace{1cm} (9)

where $e[n]$ is the energy of the normalised absolute amplitude signal $a[n]$. The energy $e[n]$ is computed as

$$e[n] = a^2[n].$$  \hspace{1cm} (10)

From the SE computation results as shown in Fig. 3d, it is noted that the multiple peaks may increase the number of false positive detections. Thus, smoothing process is performed to obtain a smooth SE envelope by smoothing out multiple peaks and spikes. The smoothing process is implemented by using a linear zero-phase filtering with a rectangular impulse response, $h(k)$ of length $L$, which is designed to provide smoothed peaked local waves around major acoustic sounds and improve accurate determination of endpoints of the local HSs. The smoothness depends on the filter length $L$, which is set based on sound duration of 25 ms. The results of smoothing process are shown in Fig. 3e. Finally, smoothed SE envelope $g[n]$ is processed at waveform endpoint determination stage.

B. HS endpoint determination: In this Letter, we employ a simple amplitude-dependent thresholding rule for determining endpoints of HSs because the SE operation produces the local envelopes with smaller peak deviations. The endpoint determination is performed using envelope thresholding with a predefined threshold of 0.1

$$g[n] = \begin{cases} 
0 & s[n] < 0.1, \\
1 & \text{otherwise,}
\end{cases}$$  \hspace{1cm} (11)

where $g[n]$ denotes the gate signal used for determining endpoints of the local acoustic sounds. Then, the gate signal is processed for determining endpoints. The detected endpoints are further processed to reject the shorter and longer duration noise segments as compared with duration ranges of normal and abnormal FHSs. The duration of each detected sound segment is compared with a predefined minimum and maximum duration thresholds of 2 and 250 ms. The endpoint determination results are shown in Fig. 3f. Experimental results show that the proposed endpoint determination approach can accurately determine endpoints even in the presence of heart murmurs and background noises. The delineation parameters are further processed for recognition of HSs.

2.2.2. S1/S2 HS recognition: This Letter mainly focuses on recognition of FHSs such as S1 and S2 HSs. The HS parameters such as duration, interval and envelope area are used to distinguish between the first and second HSs that are described as [27]:

- **Interval (I):** The $k$th interval $(I_k)$ is computed as the difference between the start time of the $(k+1)$th segment and the start time of the $k$th segment.
- **Duration (D):** The $k$th segment duration $(D_k)$ is defined as the difference between the start-time and end-time of the $k$th.
- **Envelope area (A):** The $k$th segment envelope area $(A_k)$ is computed as the integral of absolute amplitudes over the segment duration of detected HS.
In this Letter, the sound duration and envelope area features are employed in addition with interval feature to improve the S1/S2 recognition accuracy under normal and pathological PCG signals as shown in Fig. 5.

Fig. 4 shows the distributions of HS features, such as interval, sound duration and envelope area, extracted from the PCG signals taken from the PASCAL HSs Challenge database and PCCHS database. The feature analysis study demonstrates the capability of recognising the first (S1) and second (S2) HSs. The results further demonstrate that the values of interval, sound duration and envelope area can vary for normal and abnormal PCG signals acquired from different subjects. Thus, this Letter employs comparison of features of consecutively detected sound segments instead of comparing the features with predefined thresholds. A decision-rule algorithm is presented using the aforementioned features as shown in Fig. 5.

In this Letter, the sound duration and envelope area features are employed in addition with interval feature to improve the S1/S2 sound recognition accuracy under normal and pathological PCG recordings.

- Step 0: Get the feature vectors such as duration (D), interval (I), envelope area (A), and the total number of sound segments (K) detected using our delineation algorithm.
- Step 1: Perform recognition of S1/S2 sounds
  - for k = 1: K
    - if (I_k < I_{k+1})
      - if (I_{k+1} - I_k) > 50 ms
        - $k^{th}$ segment is labeled as S1 sound
      - elseif ($D_k > D_{k+1}$) && ($D_k - D_{k+1}$) > 5 ms
        - $k^{th}$ segment is labeled as S1 sound
      - elseif ($A_k > 1.5 * A_{k+1}$)
        - $k^{th}$ segment is labeled as S1 sound
      - else
        - $k^{th}$ segment is labeled as S2 sound
    - end
  - end

3. Results and discussion: In this section, we evaluate the performance of the proposed HSD and FHSR algorithms using normal and abnormal PCG signals taken from standard PCG databases and real-time recorded PCG signals.

3.1. Test PCG signal databases: The test PCG databases include the PCG signals taken from the PASCAL HSs Challenge database [28], PCCHS database [29], EGM database [30] and real-time recorded PCG signals [27]. The PASCAL HSs challenge database includes PCG signals recorded using iStethoscope Pro iPhone app and clinical trials digital stethoscope DigiScope [28]. The PCCHS database includes the PCG signals of both normal and pathological subjects with different age groups [29]. These PCG databases are widely used in the literature for evaluating the performance of the heart segmentation and classification methods. The test PCG signal databases include normal sounds, normal split, early-systolic, late-systolic and pan-systolic, mitral and tricuspid stenosis, aortic and pulmonic regurgitation, fixed and wide S2 split, S1 split, mitral prolapse, ejection murmur and click and different time-varying systolic and diastolic murmurs. The PCG signals are digitised with different sampling rates and sample resolutions. The PCG signals are corrupted by various kinds of artefacts such as talking, stethoscope motion, breathing and intestinal sounds. Moreover, the PCG signals with different signal-to-noise ratios (SNRs) are created by adding additive white Gaussian noise to the input signals.

3.2. Performance of HSD: We evaluate the performance of the proposed HSD approach using the normal and abnormal PCG signals taken from the PASCAL HSs Challenge database, PCCHS database, EGM database and real-time recorded PCG signals. Manual annotation on selected PCG signals is performed to create ground-truth table including endpoints, peak amplitudes and their locations, total number of S1 and S2 sounds, and a total number of HSs. The test PCG database includes a total of 1500 fundamental HSs. The noisy PCG signals with SNR of 25–10 dB are synthetically created for validating the robustness of the proposed HSD approach.

For evaluation of heart delineation performance, the following quantitative results such as true positive (TP) when a HS segment is correctly detected, false negative (FN) when a HS segment is not detected, and false positive (FP) when an noise segment is detected as HS, are obtained for each test signal. Then, the sensitivity (Se) = TP/(TP + FN), positive predictivity (Pp) = TP/(TP + FP) and detection error rate (DER) = (FP + FN)/TS are computed for performance evaluation. Table 1 summarises the detection performance of the proposed method and the existing methods based on the total variation filtering (TVF), Gaussian derivative filtering (GDF), and stationary WT (SWT). The delineation results are shown in Fig. 3f. Evaluation results show that the proposed HSD approach achieves an average Se of 98.87%, Pp of 97.50% with DER of 3.67% for normal and abnormal PCG signals with SNR of 10 dB. The proposed SSWT-based HSD approach outperforms the TVF- and GDF-based HSD approaches.

3.3. Performance of FHSR: The performance of the proposed FHSR method is evaluated using different kinds of HS patterns and quality of PCG recordings. The test PCG signal database also includes poor quality of FHSs. In this Letter, the performance of the FHSR method is evaluated in terms of three metrics: the Se = TN/(TN + FP), where true negative (TN) denotes the S1 sound being classified as S1 and FP denotes the S1 sound being miss-classified as S2 sound; and the Sp = TP/(TP + FN), where TP denotes the S2 sound being classified as S2 sound and FN denotes the S2 sound being miss-classified as S1 sound and the overall accuracy (OA) = (TP + TN)/(TP + TN + FP + FN). The
recognition results of this study are summarised in Table 2 for the original and noisy PCG signals. The method achieves a Se of 99.00%, Sp of 99.08% and OA of 99.04% on normal and abnormal recordings under SNR of 10 dB. The performance of the existing methods is summarised in Table 3. The EEMD with kurtosis feature method achieves an accuracy of 83.05% [9]. The homomorphic envelopogram and self-organising probabilistic model [19] had the Se of 98.6%, and Pp of 96.9% and the Se of 98.3% and Pp of 96.5% for S1 and S2 sound recognition, respectively. The probabilistic models-based method [21] had a Se of 95% and a Pp of 97%. The normalised average SE detection algorithm [17] had an accuracy of 91.2%. The DHMM method [16] identifies 739 S1 and S2 sounds out of 744, which corresponded to a Se of 99.3% and Pp of 99.1%, respectively. The EMD-DWT algorithm [12] had an accuracy rate of 100% in identifying 202 S1 sounds out of 202 and an accuracy rate of 99.48% in identifying 193 S2 sounds out of 194. The HFS method achieves Se of 89% and Sp of 91%, whereas TF vector (TF features) achieves Se of 95% and Pp of 97% [11]. The delineation results are shown in Figs. 6 and 7 for different kinds of HSs with background noises. The results show that the proposed method is able to accurately recognise the FHSs in noisy real-world PCGs with murmurs and other abnormal sounds.

<table>
<thead>
<tr>
<th>SNR</th>
<th>Method</th>
<th>FN</th>
<th>FP</th>
<th>TP</th>
<th>Se, %</th>
<th>Pp, %</th>
<th>DER, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Org.</td>
<td>TVF</td>
<td>40</td>
<td>84</td>
<td>1460</td>
<td>97.33</td>
<td>94.56</td>
<td>8.27</td>
</tr>
<tr>
<td>GDF</td>
<td>32</td>
<td>48</td>
<td>1468</td>
<td>97.87</td>
<td>96.83</td>
<td>5.33</td>
<td></td>
</tr>
<tr>
<td>SWT</td>
<td>38</td>
<td>72</td>
<td>1462</td>
<td>97.47</td>
<td>95.31</td>
<td>7.33</td>
<td></td>
</tr>
<tr>
<td>SSWT</td>
<td>18</td>
<td>25</td>
<td>1482</td>
<td>98.80</td>
<td>98.34</td>
<td>2.87</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>TVF</td>
<td>42</td>
<td>92</td>
<td>1458</td>
<td>97.20</td>
<td>94.06</td>
<td>8.93</td>
</tr>
<tr>
<td>GDF</td>
<td>32</td>
<td>52</td>
<td>1468</td>
<td>97.87</td>
<td>96.58</td>
<td>5.60</td>
<td></td>
</tr>
<tr>
<td>SWT</td>
<td>38</td>
<td>70</td>
<td>1462</td>
<td>97.47</td>
<td>95.43</td>
<td>7.20</td>
<td></td>
</tr>
<tr>
<td>SSWT</td>
<td>18</td>
<td>26</td>
<td>1482</td>
<td>98.80</td>
<td>98.28</td>
<td>2.93</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>TVF</td>
<td>42</td>
<td>90</td>
<td>1458</td>
<td>97.20</td>
<td>94.19</td>
<td>8.80</td>
</tr>
<tr>
<td>GDF</td>
<td>33</td>
<td>55</td>
<td>1467</td>
<td>97.80</td>
<td>96.39</td>
<td>5.87</td>
<td></td>
</tr>
<tr>
<td>SWT</td>
<td>35</td>
<td>82</td>
<td>1465</td>
<td>97.67</td>
<td>95.43</td>
<td>7.80</td>
<td></td>
</tr>
<tr>
<td>SSWT</td>
<td>18</td>
<td>27</td>
<td>1482</td>
<td>98.80</td>
<td>98.21</td>
<td>3.00</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>TVF</td>
<td>40</td>
<td>98</td>
<td>1460</td>
<td>97.33</td>
<td>93.71</td>
<td>9.20</td>
</tr>
<tr>
<td>GDF</td>
<td>32</td>
<td>59</td>
<td>1468</td>
<td>97.87</td>
<td>96.14</td>
<td>6.07</td>
<td></td>
</tr>
<tr>
<td>SWT</td>
<td>35</td>
<td>86</td>
<td>1465</td>
<td>97.67</td>
<td>94.46</td>
<td>8.07</td>
<td></td>
</tr>
<tr>
<td>SSWT</td>
<td>17</td>
<td>27</td>
<td>1483</td>
<td>98.87</td>
<td>98.21</td>
<td>2.93</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>TVF</td>
<td>38</td>
<td>118</td>
<td>1462</td>
<td>97.47</td>
<td>92.53</td>
<td>10.40</td>
</tr>
<tr>
<td>GDF</td>
<td>30</td>
<td>62</td>
<td>1470</td>
<td>98.00</td>
<td>95.95</td>
<td>6.13</td>
<td></td>
</tr>
<tr>
<td>SWT</td>
<td>34</td>
<td>97</td>
<td>1466</td>
<td>97.73</td>
<td>93.79</td>
<td>8.73</td>
<td></td>
</tr>
<tr>
<td>SSWT</td>
<td>17</td>
<td>38</td>
<td>1483</td>
<td>98.87</td>
<td>97.50</td>
<td>3.67</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>Signal</th>
<th>S1 recognition</th>
<th>S2 recognition</th>
<th>OA, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>TN</td>
<td>FP</td>
<td>Se, %</td>
<td>TP</td>
</tr>
<tr>
<td>SSWT</td>
<td>Org.</td>
<td>2474</td>
<td>26</td>
<td>98.96</td>
</tr>
<tr>
<td>30 dB</td>
<td>2474</td>
<td>26</td>
<td>98.96</td>
<td>2478</td>
</tr>
<tr>
<td>25 dB</td>
<td>2472</td>
<td>28</td>
<td>98.88</td>
<td>2480</td>
</tr>
<tr>
<td>20 dB</td>
<td>2468</td>
<td>32</td>
<td>98.72</td>
<td>2480</td>
</tr>
<tr>
<td>15 dB</td>
<td>2476</td>
<td>24</td>
<td>99.04</td>
<td>2476</td>
</tr>
<tr>
<td>10 dB</td>
<td>2475</td>
<td>25</td>
<td>99.00</td>
<td>2477</td>
</tr>
<tr>
<td>SWT</td>
<td>Org.</td>
<td>2374</td>
<td>126</td>
<td>94.96</td>
</tr>
<tr>
<td>30</td>
<td>2375</td>
<td>125</td>
<td>95.00</td>
<td>2382</td>
</tr>
<tr>
<td>25</td>
<td>2368</td>
<td>132</td>
<td>94.72</td>
<td>2376</td>
</tr>
<tr>
<td>20</td>
<td>2366</td>
<td>134</td>
<td>94.64</td>
<td>2374</td>
</tr>
<tr>
<td>15</td>
<td>2365</td>
<td>135</td>
<td>94.60</td>
<td>2372</td>
</tr>
<tr>
<td>10</td>
<td>2374</td>
<td>126</td>
<td>94.96</td>
<td>2380</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reference</th>
<th>Signal processing techniques and classifiers</th>
<th>Performance, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tschannen et al. [8]</td>
<td>HSMM and Viterbi decoding, CNN, SVM</td>
<td>Se = 96.97, Pp = 99.58, DER = 3.55</td>
</tr>
<tr>
<td>Papadaniil and Hadjileontiadis [9]</td>
<td>EEMD, Kurtosis</td>
<td>OA = 94.56</td>
</tr>
<tr>
<td>Varghees and Total variation filter, Shannon entropy,</td>
<td>Hilbert transform</td>
<td>Se = 99.71, Pp = 94.20, DER = 6.44, OA = 93.95</td>
</tr>
<tr>
<td>Ramachandran [10]</td>
<td>S-Transform, SE, SVD features</td>
<td>Se = 95, Sp = 97</td>
</tr>
<tr>
<td>Sun et al. [12]</td>
<td>EMD-EN, SE, thresholding, homomorphic filtering, self-organising HMM</td>
<td>S1: Se = 98.6, Pp = 96.9, S2: Se of 98.3 and Pp of 96.5</td>
</tr>
<tr>
<td>Gill et al. [19]</td>
<td>TF, probabilistic models (HMM)</td>
<td>Se = 94.9, Pp = 96.9</td>
</tr>
<tr>
<td>Gamero and Watrous [21]</td>
<td>frequency filtering, Hilbert transform, time-domain features, NN SSWT, SE, decision rules</td>
<td>misclassification rate = 5.76</td>
</tr>
<tr>
<td>Hebden and Torry [23]</td>
<td>proposed method</td>
<td>S1: Se = 98.96, S2: Sp = 99.12, OA = 99.04</td>
</tr>
</tbody>
</table>

This is an open access article published by the IET under the Creative Commons Attribution-NoDerivs License (http://creativecommons.org/licenses/by-nd/3.0/)
4. Conclusion: This Letter presents an automatic WFHSR method based on the SSWT and morphological and interval features. In this Letter, we investigate the application of the SSWT for extracting the HSs and suppressing the heart murmurs and background noises. A simple HSD approach based on the SE envelope is presented for automatically determining endpoints and peaks of HSs. An automated FHSR approach based on the morphological and interval features. The delineation and recognition performance of the proposed approaches are evaluated on PASCAL HSs Challenge, PCCHS Challenge, EGM databases and real-time recorded PCG signals. Evaluation results show that the proposed HSD approach achieves an average Se of 98.87%, Sp of 97.50% with DER of 3.67% for PCG signals with SNR of 10 dB, and outperforms the existing approaches. The FHSR approach achieves an average a Se of 99.00%, Sp of 99.08% and OA of 99.04% on both normal and abnormal PCG signals with SNR of 10 dB. Results show that the proposed method provides promising results on both high-quality and poor-quality normal and abnormal PCG recordings.

5. Funding and declaration of interests: None declared.

6 References


