

REVIEW

Open Access



Localization and classification of heart beats in phonocardiography signals —a comprehensive review

Shahid Ismail^{1,3*} , Imran Siddiqi¹ and Usman Akram²

Abstract

Phonocardiogram (PCG) signal represents recording of sounds and murmurs resulting from heart auscultation. Analysis of these PCG signals is critical in diagnosis of different heart diseases. Over the years, a variety of methods have been proposed for automatic analysis of PCG signals in time, frequency, and time-frequency domains. This paper presents a comprehensive survey of different methods proposed for automatic analysis of PCG signals with the objective to evaluate the current state-of-the-art and to determine the potential domains of effective analysis. An important aspect of our contribution is that the review is carried out as a function of domains of analysis rather than simply discussing different methods. Our method further splits analysis into pre-processing, localization, and classification, and details are presented in terms of techniques and classifiers used during these phases. Finally, results are summarized for normal heart beat, noisy heart beat, and different pathologies using metrics like accuracy and detection rate. In addition to time and frequency domain, time-frequency based methods including wavelet, empirical mode decomposition (EMD) and time-frequency representation (TFR) are selected for detailed analysis. The review concludes that the time-frequency representations like EMD and wavelets represent areas of exploration in future along with perspective of using different time-frequency techniques together.

Keywords: PCG signals, Classification, Localization, EMD, Wavelets, TFR

1 Introduction

Human heart is the most important organ in the body that provides blood to all parts of the body using a pump-like action. During the pumping action, electrical and mechanical activities are carried out resulting in the flow of blood. Healthy heart is very important for the normal day to day working of human body as blood carries important nutrients to the organs. Heart-related diseases, known as cardiovascular diseases (CVD), are responsible for a major proportion of deaths all around the world. According to surveys conducted by the World Health Organization (WHO), 33% of all deaths are the result of CVDs [1]. Different modalities are known to exist to monitor the health of heart. The most popular of these is the electrocardiogram or ECG. The electrical activity of heart

is recorded in the form of electrocardiogram which is composed of three main waves, 'P' wave, 'QRS' complex, and 'T' wave. The middle wave of ECG is known as 'QRS' complex that, in general, comprises three deflections 'Q' (the first negative deflection) 'R' (the first positive deflection), and 'S' (the negative deflection following the 'R' wave). Another important modality is photoplethysmography (PPG) that employs light-based sensors to estimate the rate of flow of blood by measuring the changes in the reflected/transmitted light. Like ECG, PPG can also be used to monitor various cardiac conditions. Figure 1 illustrates an example each of ECG and PPG signals [2]. A wide range of computerized systems have been proposed over the years for automatic analysis of ECG [3, 4] and PPG [5] signals as well as the combination of these and other modalities [6].

In addition to ECG and PPG, phonocardiogram (PCG), the recording of the sounds and murmurs made by heart during a cardiac cycle, can be effectively employed to study and monitor the activities of heart. Such sounds

*Correspondence: shahid.ismail@gmail.com

¹ Bahria University, Shangrilla Road, Islamabad, Pakistan

³ APCOMS, UET, Taxila, Rawalpindi, Pakistan

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

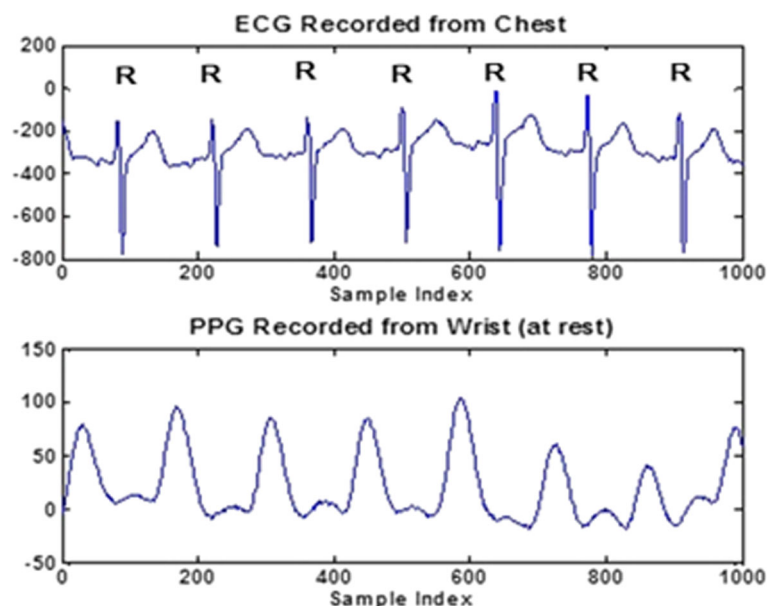


Fig. 1 Examples of ECG and PPG signals

are typically recorded using a device called phonocardiograph. The mechanical activity of heart due to its physical movement produces four distinct sounds or beats which are named as first normal heart sound (S1), second normal heart sound (S2), third abnormal heart sound (S3), and fourth abnormal heart sound (S4). S1 and S2 are normal sounds while S3, S4, murmurs, and certain other sounds usually refer to some disease or anomaly. Murmur is a noisy heart sound that can be heard when the heart valve is shut but the blood continues to flow. S1 is usually low frequency and high-amplitude signal while S2 has high frequency and low amplitude. In some cases, it is also possible to encounter S1 having low amplitude than S2 as elaborated in [7].

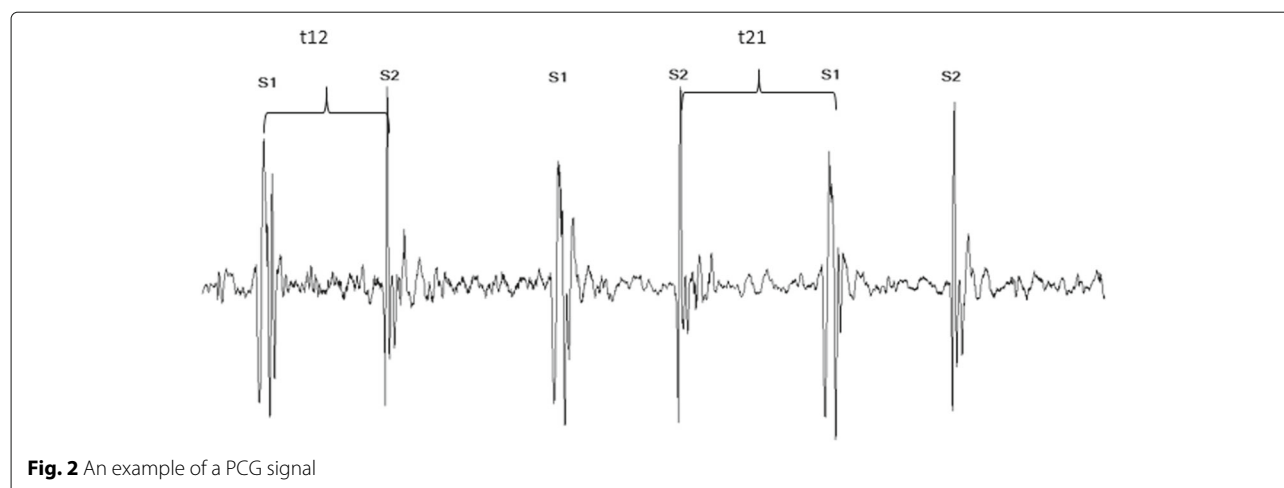
ECG, PPG, and PCG, all represent a cyclostationary signal, i.e., a signal in which the statistics of the signal vary but are repetitive with a period. ECG and PCG are highly correlated signals [8] and are known to contain more information than the PPG signal. PCG, however, enjoys a distinct advantage over ECG and PPG signals as it records the acoustic properties of the signal. These properties are better suited for murmur detection which represents abnormal heart sound [9]. Furthermore, PCG signal also has an excellent starting trigger in the form of S1 wave [10, 11].

The mechanical activity of the heart can be heard using a traditional or an electronic stethoscope. Auscultation or listening to heart is an old but very effective method of diagnosis for a number cardiovascular diseases. The recording of these sounds (PCG signal) has the same spectrum as that of audio signals. The recording process,

however, also picks noisy sound signals from the environment which distort the periodicity of the PCG signal making its analysis a challenging task. A sample PCG signal is shown in Fig. 2 where it can be seen that the first heart sound S1 has lower amplitude than the second heart sound S2. Moreover, the amplitudes of S1 and S2 are both varying. The time duration from S1 to S2 is known as systole and the one between S2 and S1 is known as diastole. Systole and diastole shown as t_{12} and t_{21} , respectively, in Fig. 2 serve as important beat classification features. The time durations t_{12} and t_{21} are also varying.

PCG, though a very complex signal, divides the normal heart sound signals into two beats (S1 and S2) and offers important details about a number of heart-related diseases [12–14]. With the advancements in technology as well as different areas of pattern classification and machine learning, efforts have been made to automate the analysis of PCG signals. The automatic analysis of these signals involves two primary challenges, localization and classification. In localization, the objective is to correctly detect the positioning of beats in the signal while classification deals with the categorizing the beats into S1 and S2 in case of normal heart sounds and into S3, S4, and murmurs etc. in case of abnormal sounds.

Localization algorithms are generally peak-based algorithms where candidate peaks are localized in a signal and windows are constructed around these potential peaks. Features are then calculated from each window, and the peaks are classified. Another approach is not to precisely localize the peaks for classification. Springer et al. [15] for instance, divides the signal into a sequence of segments



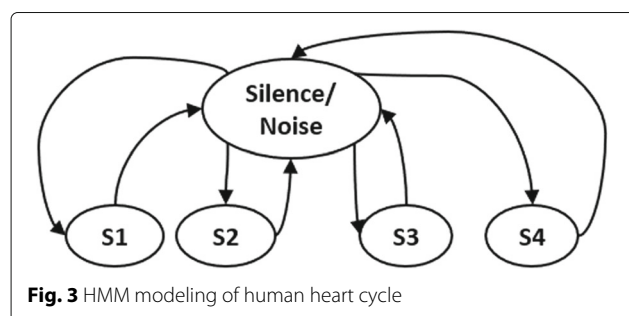
which are classified and later processed for a precise localization. Heart beat is a repetitive or cyclic process (Fig. 2) and a normal PCG signal can be divided into four segments. These include the time window containing S1, the silence window, the segment containing S2, and again the silence window. A normal heart beat can hence be modeled using these states. However, since the sequence and duration of each state within the cycle can vary, probabilistic modeling is better suited for such scenarios as opposed to deterministic approaches. Hidden Markov model (HMM), for instance, has been employed for such modeling in [15–17]. The same idea can be extended to include the abnormal beats S3 and S4 (and different other pathologies) as illustrated in Fig. 3. Such techniques allow pre-classifying these segments into one of the states while the exact location of beats within these segments are localized subsequently.

The initial developments and findings in digital processing of PCG signals have been summarized in works by [18, 19]. In a relatively recent survey, Meziani et al. [20] discussed computerized analysis of PCG signals but limit their discussion to wavelet transforms-based methods only. In another short review [21], authors present a comparative study of EMD and wavelet-based methods for analysis of PCG signals. The study concludes that EMD is better suited for PCG as compared to wavelet,

specially, when dealing with noisy signals. Authors highlighted the use of machine learning tools for better classification and suggested that the use of hybrid classifiers serves to enhance the performance. Nabih et al. [22] review the denoising, segmentation, and classification techniques for automated analysis of PCG signals. The key focus of their study is the comparison of different classifiers (artificial neural network (ANN), support vector machine (SVM), self-organizing map (SOM), and hybrid classifiers with the conclusion that while each of the processing steps is important, the choice of classifier is the most critical parameter in enhancing the overall system performance.

As discussed earlier, the present study is organized as function of time, frequency, and time-frequency methods. In time domain, the signal is sampled on time axis (Fig. 4a) and features (like amplitude, mean, and energy) are directly computed from signal. In spectral or frequency domain, signal spectra is divided into various spectral bands (Fig. 4b) and features (for example, spectral flux) are calculated from these bands. Unlike time and frequency methods, the time-frequency methods rely on simultaneous sampling on time and frequency axis. The general representations of time-frequency domain for fixed and variable time/frequency windows are illustrated in Fig. 4c, d, respectively. In our discussion, we considered time-frequency representations (like Wigner-Ville and its variants, short-time frequency transform (STFT)) and techniques based on EMD and wavelets under the category of time-frequency methods.

We first introduce the standard datasets commonly employed by researchers (to evaluate the localization and classification algorithms) in Section 2. Section 3 details the time, frequency, and time-frequency methods for analysis of PCG signals along with a critical discussion on the presented methods. Finally, Section 5 concludes the paper.



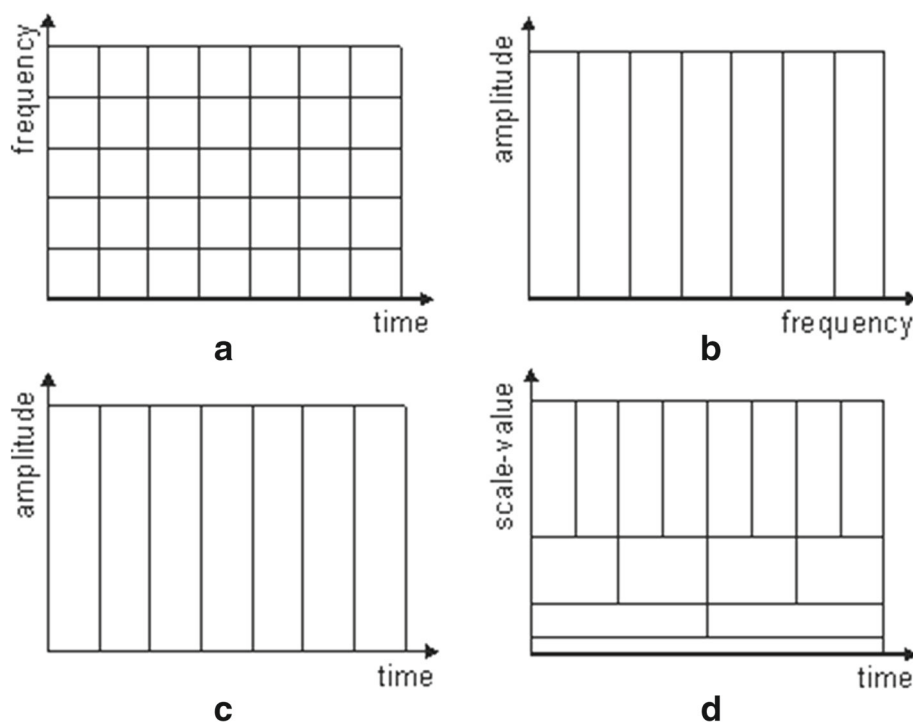


Fig. 4 Signal analysis in time, frequency, and time-frequency domain. **a** Signal/time domain. **b** Frequency/spectral domain. **c** Time and spectral domain (Gabor transform). **d** Time and spectral domain (wavelet transform)

2 Data sets

While many of the initial PCG localization and classification techniques reported results on private datasets, a number of publicly available datasets have also been developed allowing researchers meaningfully compare their developed systems using similar experimental protocols.

These public datasets include E-General Medical [23], PASCAL Heart Sound Challenge (HSC) dataset [24], the PhysioNet CinC Challenge dataset [25], and the Heart Sound and Murmur Library [26]. In the following, we present the key characteristics of these datasets for a better interpretation of the quantitative results reported by different studies (presented later in paper).

PASCAL HSC dataset was first made available in 2011 for two challenges, segmentation (localization), and classification. The database is divided into two subsets, dataset 'A' and dataset 'B'. Dataset 'A' has been collected using the iStethoscope Pro iPhone app while dataset 'B' has been gathered in a clinical setting using the digital stethoscope, DigiScope. The two datasets have 176 and 656 total auscultations, respectively. Files in both the datasets contain normal heart beats, murmurs, and extra systoles. In addition to these, dataset 'B' also contains files with artifacts and extra sounds and is more challenging.

The 2016 PhysioNet CinC database was also developed as a part of a challenge. Heart sounds in the dataset are gathered from clinical as well as non-clinical

sources. Healthy individuals as well as those with different pathologies contributed to data collection. The sound classes in the challenge include 'normal,' 'uncertain,' and 'abnormal.' The training set labeled from 'A' to 'E' has a total of 3126 files while the validation set comprises 300 recordings. The durations of the files vary from 5 to 120 s. The test set has not been made publicly available and was only used to score the participants of challenge.

E-General Medical is a well-known vendor of medical equipment and datasets. Among other datasets, a part of the dataset with heart sounds has been freely made available by E-General Medical. The dataset comprises 64 recordings with normal beats, S3, S4, and different pathologies. Majority of pathologies are systolic and diastolic in nature. In many cases, only one sample per signal type is available. Examples of such one sample per signal type include late systolic, early systolic, normal split, open snap, pan systolic, late systolic aortic stenosis, severe systolic aortic stenosis, critical systolic aortic stenosis, and systolic aortic valve replacement. Other sounds like diastolic-fixed S2 split, diastolic wide S2 split, systolic mitral regurgitation, systolic mitral prolapse, systolic split, and diastolic-fixed S2 split have more than one sample. Complete details on the dataset can be found in [23].

Heart sound and Murmur Library, University of Michigan Health Systems, is another public dataset that comprises normal heart sounds, normal split S1, systolic click, S4 gallop, S3 gallop, single S2, split S2 persistent, split S2

transient, and various type of murmurs [26]. A summary of the databases presented above can be found in Table 1.

3 Methods

As mentioned previously, the techniques proposed for automatic analysis of PCG signals can be categorized into time domain, frequency domain, and time-frequency domain. In the following, we discuss the time and frequency domains followed by the time-frequency domain.

3.1 Time and frequency analysis

The most general method of analysis is based on the time domain. In time domain, analysis is carried out on the signal itself. Statistics like peak value, mean, mode, median, and peak to peak duration of signal are typically employed for analysis. As indicated in Fig. 5, P1 and P2 are the peak values of the signal, A1 and A2 are peak to peak values, and $\Delta 1$ and $\Delta 2$ are time periods which may represent features of S1 and S2. These and similar features are extracted either globally from the complete signal or locally from different segments of the signal. The general representation of a signal in time domain is presented in Eq. 1; $f(t)$ represents the signal which is modeled by amplitude A , frequency ω , and phase θ . Each of these parameters as well as different statistics derived from these parameters have been used as features for localization and classification.

$$f(t) = A \sin(\omega * t + \theta); \quad (1)$$

For frequency domain analysis, the signal is first converted from time to frequency domain using fast Fourier transform (FFT) elaborated in Eq. 2.

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt; f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{j\omega t} d\omega \quad (2)$$

Where $f(t)$ is the signal and $F(\omega)$ is the fast Fourier transform of $f(t)$. After analysis of the frequency spectrum, any appropriate spectrum operation (for instance

filtering) is applied, and the signal is converted back to time domain using the inverse fast Fourier transform (IFFT). Bandpass filtering is usually employed in heart beat analysis to remove the low and high frequency noise from stethoscope signals. Figure 6a illustrates a sample signal before and after filtering. Figure 6b shows the spectrum of the signal, filter response, and the frequency contents after filtering. Since the frequency contents of various heart signals are different, the frequency spectra can be exploited for classification of heart beats. Finding the optimal set of filters (along with cutoff frequencies), however, can be challenging especially for noisy signals. Consequently, a number of researchers investigated Mel-Frequency Cepstral Coefficients (MFCCs) [27] for frequency domain analysis of PCG signals. In MFCC calculation, pre-emphasis is first applied on the signal to enhance the high-frequency components of signal. The signal is then divided into windows (generally rectangular), and FFT is applied to each window. The Mel filter bank is then applied and logarithm of all filter bank energies is computed. The acoustic response of human ear is not linear to all frequencies. Mel scale uses this non-linearity and maps frequencies (f) linearly below 1000 Hz and logarithmically above 1000 Hz. Conversion from frequency to Mel scale is given in Eq. 3.

$$\text{mel}(f) = 1127 \ln \left(1 + \frac{f}{700} \right) \quad (3)$$

Finally, discrete cosine transform (DCT) is applied to get the MFCCs [28]. An overview of the steps involved in MFCC calculation is presented in Fig. 7. The pre-emphases, windowing, and FFT calculation in MFCC computation act as pre-processing steps; hence, a separate pre-processing is generally not required.

Time and frequency analysis have been widely investigated to localize and classify normal as well as abnormal heart beats like third and fourth heart sound, clicks and murmurs. With a few exceptions [29, 30], analysis in the

Table 1 A summary of PCG datasets

Dataset	Reference	Total auscultations	Sound categories
PASCAL heart sound challenge dataset 2011	[24]	Dataset 'A', iStethoscope, samples: 176 dataset 'B', Digiscope, samples: 656	Dataset 'A': normal beats, murmurs, and extra systoles dataset 'B': normal beats, murmurs, extra systoles, extra heart sounds, and artifacts
PhysioNet CinC challenge dataset 2016	[25]	3126	Normal, uncertain, and abnormal
E-General Medical	[23]	64	Normal beats and different pathologies
Heart sound and Murmur Library, University of Michigan Health systems	[26]	23	Normal and abnormal sounds

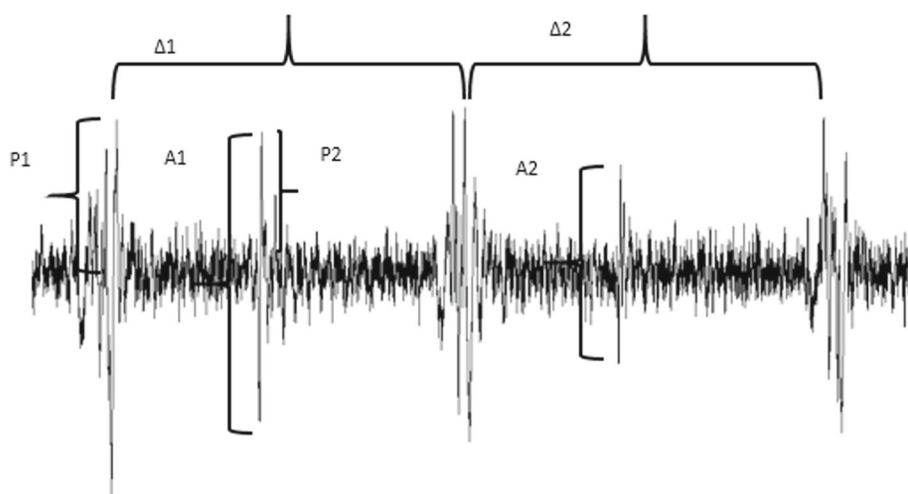


Fig. 5 Analysis of PCG signal in the time domain

time domain usually starts with pre-processing which generally comprises decimation followed by a low pass filter with a cut-off frequency of around 200 Hz. The decimation decreases the computational load on the analysis module. Low pass filtering is used as pre-processing step as heart beat has characteristics of low frequency signal. Liang et al. [31] and Potes et al. [32] first decimate the signal and pass it through low pass and band pass filters, respectively. Chauhan et al. [16] employed low pass and median filtering along with Hamming window for removal of noise, ripples, and ringing effects. Likewise, Ari et al. [33] pass the signal through a low pass filter after decimation while Hussnain et al. [34] simply used the low pass

filtering without decimation. Signal from the decimation and filtering stages is next fed to the normalization stage.

Normalization removes the amplitude variation for inter and intra signal classification. The signal is then converted to a form that is best suitable for localization and classification. Shannon energy envelop is most widely employed method investigated in most of the time-based techniques [16, 31, 34]. In short, researchers have mostly employed combinations of decimation, filtering, and Shannon energy for pre-processing in the time domain.

The signal after pre-processing has peaks at regular and non-regular intervals. Candidate peaks can be used for

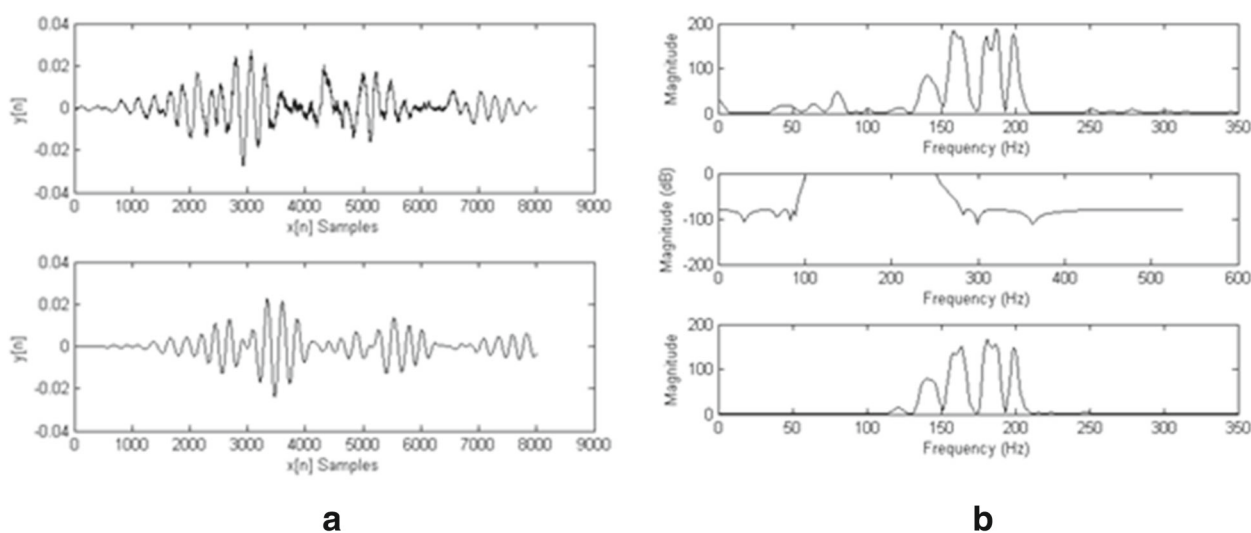


Fig. 6 Signal analysis in frequency domain. **a** Signal before and after filtering. **b** Signal spectrum before and after filtering

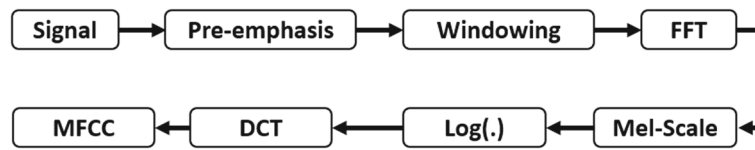


Fig. 7 Calculation of MFCC

individual beat localization which in turn can be used for heart rate detection. The time duration between S1 and S2 (systole) and the one between S2 and S1 (diastole) are usually exploited for the localization and classification of candidate peaks [16, 31, 33].

Another common approach is to employ machine learning-based techniques for localization and classification. Signals represented in an appropriate feature space are used to train a learning algorithm.

Authors in [35], for instance, use support vector machine with a modified cuckoo search (MCS) optimizer using features extracted from linear predictive coding (LPC) for classification. Similarly, Chen et al. [36] investigated deep neural networks for recognition of heart sounds S1 and S2. In another study, Potes et al. [32] applied AdaBoost-abstain (modified version of conventional AdaBoost) and convolutional neural networks (CNN) to detect abnormal heart sounds.

As discussed previously, MFCC is usually used in the frequency-based methods and is found to be effective in speech processing including speech recognition and speaker identification [36]. MFCC has been used in the feature extraction process in [36] and [32] for localization and classification. Likewise, MFCC-based features were used to train HMM classifiers in [16] and reported high classification rates.

As discussed earlier, heart beat signal is a time series data composed of different events (occurrence of S1, S2, etc.). The sequence of these events varies in case of abnormal signals. Human heart beat has therefore been modeled using HMM (Fig. 3) in a number of studies [15–17] where each state represents an event (silence, S1, S2, murmur, etc.). Although HMMs exploit this temporal dependency, the features fed to these HMMs can be extracted in the frequency domain. MFCC-based features, for instance, have been employed to train HMMs [17]. Such modeling has been applied to simple two class problem (normal and abnormal heart sound) as well as for the detection of different pathologies like ventricular septal defect (VSD), mitral regurgitation (MR), aortic stenosis (AS), aortic regurgitation (AR), patent ductus arteriosus (PDA), pulmonary regurgitation (PR), and pulmonary stenosis (PS) [17].

Ajit et al. [37] used simple FFT for localization and classification without any further processing as frequency

contents of different heart sounds have different spectrum contents. Authors in [29] used both frequency- and energy-based features for classification while Mandeep and Cheema [30] used time, frequency, cepstrum, and statistical features for localization as well as for classification. Features were evaluated using Ranker and Info Gain Attribute Evaluation algorithm, and multiple classifiers were investigated to recognize the beats. Among other notable studies, Samit et al. [33] after pre-processing, used wavelet decomposition for localization and least square SVM with least mean square (LMS) for classification. In another study, Nogueira et al. [38] exploited MFCC with time features of systole and diastole durations, and the relationship between ECG and PCG signal for heart sound classification. Ortiz et al. [39] employed dynamic time warping (DTW) along with MFCC for heart sound classification realizing a test score of 84% using SVM classifier. Tang et al. [40] proposed an ensemble of 324 features primarily composed of time, frequency, kurtosis, energy, and other features to classify a sound as normal or abnormal using ANN. Rubin et al. [41] converted PCG signal to heat map using MFCC for abnormal sound detection using convolutional neural network (CNN). It can be noticed from the discussion of the presented techniques that classifiers like ANN, SVM, and HMM have been commonly employed in time- and frequency-based methods. While HMM exploits the temporal dependencies in the signal, classifiers like ANN and SVM categories individual beats into one of the beat classes under study.

From the view point of heart sounds, S1 and S2 are considered in [30, 31, 36, 37] while [35] and [16], in addition to S1 and S2, also consider split S1, split S2, S3, S4, murmurs, clicks, and snaps. The work in [34] is focused on discrimination between normal heart sounds and murmurs. Likewise, authors in [32, 39–41] primarily distinguish between the two broad categories, normal (S1, S2) and abnormal heart sounds. Similarly, noise or spike detection is studied in [29] while the work in [33] considers normal heart conditions as well as pathological problems including aortic insufficiency, aortic stenosis, atrial septal defect, mitral regurgitation, and mitral stenosis.

Table 2 summarizes the research in explicit time [29–31, 33–35, 38–41] and frequency [16, 17, 32, 36, 37] domains. All studies except [30, 32, 34, 38–41] consider

Table 2 Time and frequency domain

Study	Year	Pre-processing techniques	Localization techniques	Classification techniques	Beat types	Data set	Signals statistics	Evaluation metrics	Results
Nogueira et al. [38]	2017	–	Springer's segmentation algorithm [15]	Classification of heart images generated from MFCC and time features using SVM, KNN, CNN, and random forest	Normal and abnormal heart sound	PhysioNet CinC challenge, 2016	13404 heart image	Sensitivity, specificity, and overall score	96.47, 72.65, and 84.56%
Chen et al. [36]	2017	MFCC	K-means and DNN	K-means and DNN	Normal heart sound	Private dataset	920 beats	Precision, recall, and F measure	All measures > 91%
Potes et al. [32]	2016	Signal decimation and low pass filtering	Time, spectrum, MFCC-based features, CNN, AdaBoost-Abstain classifier, and ensemble classifier	Time, spectrum, MFCC-based features, CNN, AdaBoost-Abstain classifier, and ensemble classifier	Normal and abnormal heart sounds	PhysioNet CinC challenge, 2016	3240 beats and a blind dataset	specificity, sensitivity, and overall score	92.42, 77.81, and 86.02%
Ortiz et al. [39]	2016	–	Springer et al. [15]	SVM using time, MFCC, DWT, and wavelet	Normal and abnormal heart sounds	PhysioNet CinC challenge, 2016	3240 beats and a blind dataset	Test score	82.4%
Kotb et al. [17]	2016	Pre-emphasis filtering and windowing	MFCC	Heart murmur model using HMM	Normal and abnormal heart sounds	Private dataset	1069 normal and abnormal heart sound	Sensitivity, correct classification rate	98 and 96%
Tang et al. [40]	2016	Bandpass filtering	Springer et al. [15]	ANN using time, kurtosis, MFCC, energy, power spectrum cyclostationary features	Normal and abnormal heart sounds	PhysioNet CinC challenge, 2016	3240 beats and a blind dataset	Sensitivity, specificity, and over all score	80.7–81.2%, 82.9–85%, and 81.8–83.6%
Rubin et al. [41]	2016	–	Springer et al. [15]	CNN using heat maps generated using MFCC features	Normal and abnormal heart sounds	PhysioNet CinC challenge, 2016	3240 beats and a blind dataset	Specificity, sensitivity, and overall score	95, 73, and 84%
Redlarski et al. [35]	2014	–	Modified linear predictive coding(LPC)	ANN, SVM-MCS	Normal, abnormal heart sounds, and clicks	3M Poland microphone samples	72 signals	Accuracy	66–100%
Naseri et al. [29]	2013	–	Frequency and energy features	Frequency and energy features	Normal and noisy heart sound segments	Signals recorded at Rajaei Cardiovascular, Medical and Research Center, Iran	63 clean and 63 noisy segments	Sensitivity and positive predictive value	91.24 and 92.86%

Table 2 Time and frequency domain (*Continued*)

Study	Year	Pre-processing techniques	Localization techniques	Classification techniques	Beat types	Data set	Signals statistics	Evaluation metrics	Results
Singh et al. [30]	2013	–	Time domain, frequency domain, cepstrum, and statistical features	Time domain, frequency domain, cepstrum, statistical features, Bayes Net, naive Bayes, GSD, and Logit Boost	Normal heart sounds and murmurs	PASCAL CHSC2011 data set A and B	30 normal and 30 murmurs	Specificity, sensitivity, and accuracy	93.3, 93.3, and 93.3%
Ari et al. [33]	2010	Low pass filtering, normalization, and Shannon energy envelop calculation	Low pass filtering, normalization, Shannon energy envelop calculation, and wavelet decomposition	Least square SVM with LMS algorithm	Normal heart sounds, aortic insufficiency, aortic stenosis, atrial septal defect, mitral regurgitation, and mitral Stenosis	Data recorded at Maulana Azad Medical Institute, Delhi, India	512 cycles from 64 signals	Classification rate	86–100%
Chauhan et al. [16]	2008	Low pass filtering, normalization, and Shannon energy envelop calculation	Time duration and MFCC	Time duration, MFCC, and HMM	Normal heart sounds, snaps, murmurs, and clicks	Private dataset	41 Samples	Correct/ incorrect recognition rate	80–100%
Ali et al. [34]	2007	Low pass filtering	Shannon energy envelop withassification technique adaptive window and heart beat modeling	Time duration in Shannon energy envelope	Normal heart sounds, murmurs, and clicks	E-General Medical	–	–	–
Liang et al. [31]	1997	Signal decimation, low pass filtering and Shannon envelop formation	Shannon energy envelope with adaptive window, heart beat modeling, and time-based peak detection of Shannon energy envelope	Time-based peak detection of Shannon energy envelope	Normal and abnormal heart sounds	Private dataset	37 recording composed of 515 cycles	Single cardiac cycle correct/ incorrect detection	93%
Yogana- than et al. [37]	1976	–	FFT	FFT	Normal heart sounds	Private dataset	29 signals	–	–

private datasets for evaluation of the proposed techniques. These private datasets were recorded at various settings, for instance, authors in [29] and [33] used clinical settings while 3M Poland microphone samples are employed in [35]. Likewise, for public datasets, different subsets have been employed by different authors making a direct comparison of these techniques difficult.

The evaluation protocol in different studies also varies. Techniques proposed in [16, 17, 30–32, 35, 37–41] considered signals, [33] used cycles while [36] used beats. The general evaluation metrics include accuracy, specificity, sensitivity, and positive predictive value (PPV) as elaborated in the following:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (4)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (5)$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (6)$$

$$\text{PPV} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (7)$$

Where ‘TP’ represents true positives or correctly identified, ‘TN’ represents true negatives or correctly rejected, ‘FP’ for false positives or incorrectly identified, and ‘FN’ represents false negatives or incorrectly rejected. Most of the studies [17, 29, 30, 32, 35, 36, 38–41] used these measures; authors in [16, 31] employed correct and incorrect cycle detection while [33] employed classification rate for quantifying the results.

Comparing different time-based methods, it can be observed from Table 2 that a number of studies achieved an overall score of above 90% on public datasets [24, 25]. On private datasets, classification accuracies of as high as 100% have also been reported [33, 35]. Springer algorithm [15] has been employed for localization in most of studies [38, 40, 41]. For classification, peak to peak time duration has shown to be a discriminative feature [31, 34]. A similar trend is observed for frequency based methods which report, on the average, accuracies of more than 80% on the public PASCAL challenge dataset while a classification rate of 100% is claimed on a private dataset [16]. Features based on MFCC have been commonly employed in combination with variety of classifiers. In some cases, time- and frequency-based features have been combined [16, 30, 32] to further enhance the localization and classification performance. Furthermore, it is also common to employ time domain for classification frequency domain for classification [38–41].

3.2 Time-frequency analysis

Due to cyclostationary nature of PCG, a signal cannot be localized both in time and frequency due to uncertainty principle limitation. Consequently, time-frequency analysis of PCG signals is investigated in a number of studies. Such methods generally employ short-time frequency transform (STFT), time-frequency representations like Wigner-Ville, wavelet transform, and empirical mode decomposition (EMD).

3.2.1 Time-frequency representation

These methods rely on taking the Fourier transform of auto-correlation of signal as indicated in Eq. 8 [42].

$$\text{WD}(t, w) = \int_{-\infty}^{\infty} x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) e^{-jw\tau} d\tau \quad (8)$$

Where WD stands for Wigner distribution, x is the signal, x^* represents the conjugate of x while τ is the time delay and w is the frequency. For a mono component signal, Eq. 8 produces a frequency component that is representative of signal frequency. However, for multiple-component signal, this autocorrelation produces extra components known as ‘cross-terms.’ As an example, we consider a two-component signal $x(t) = x_1(t) + x_2(t)$. The autocorrelation of signal $x(t)$ will be composed of terms $x(t)*x(t) = x_1(t)*x_1(t) + x_2(t)*x_2(t) + x_1(t)*x_2(t)$. While the first two terms represent the power components of signal under study, the third term is an additional undesirable component (cross-term). Since WD computes FFT of the autocorrelation of signal, the contribution from cross-terms also appears; hence, corrupting the spectrum. To minimize the impact of cross-terms, a windowed version of above equation is normally used as presented in Eq. 9.

$$\text{PWD}(t, w) = \int_{-\infty}^{\infty} w(\tau) x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) e^{-jw\tau} d\tau \quad (9)$$

Where $w(\tau)$ represents the smoothing window. Equation 9 is known as pseudo Wigner-Ville distribution [43]. Figure 8 shows the output of the pseudo Wigner-Ville representation of heart beat signal from a sample file in the Pascal DigiScope dataset (Table 1). The peak at 1000th sample represents S1, and the one at 2000th sample represents S2. All other structures represent noise resulting from cross terms due the TFR manipulation. The noise samples result from autocorrelation of multi-spectral signals.

Vargas et al. [44] demonstrated that time-frequency methods can be used for analysis of non-stationary signal like EEG and PCG. Zhang et al. [45] used matching pursuit (MP) that works both in time and frequency domains, for PCG signal decomposition. For time scaling, the signal is

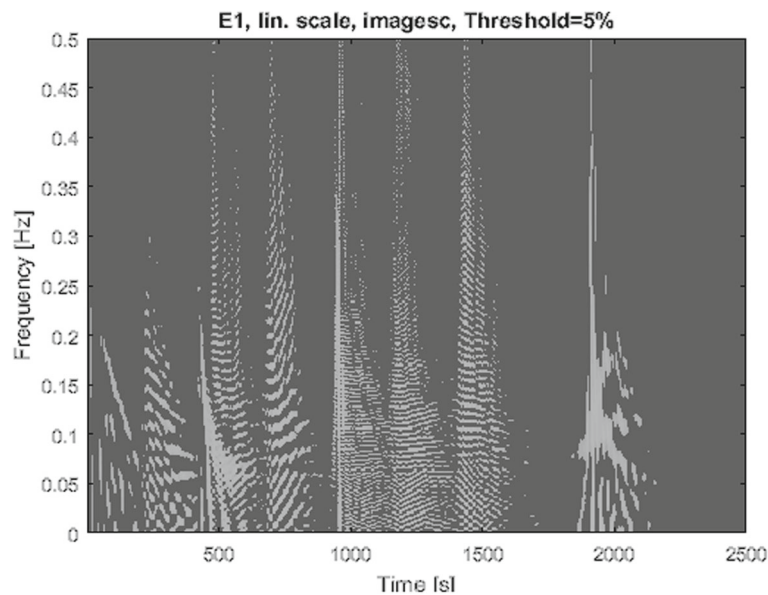


Fig. 8 Pseudo Wigner-Ville representation (time-frequency) of PCG signal

expanded in a way that the frequency components remain the same. Likewise, for frequency scaling, the frequency contents are compressed and expanded while keeping the temporal contents intact. In frequency scaling, the frequency band may be shifted towards higher frequencies or towards lower spectrum. Despite the shift, the contents in time remain the same. It is not uncommon to use both time and frequency scaling. The MP method converts the signal into time-frequency atoms with time and frequency parameters. The summation of atoms gives back the original signal with error. Making use of these parameters, the signal can be transformed. This transformation results in a modified version of the signal which cannot be directly compared with original signal. An inverse scaling is therefore used for comparison. The time-frequency representations (TFR) can be calculated from original, scaled, and inversed versions. The sum of Wigner distribution of all atoms of a signal is termed as MP-based Wigner distribution. MP-based Wigner distribution has the distinct advantage of automatic removal of cross terms over conventional Wigner distribution when applied to signal. Also, MP-based Wigner distribution gives much clear presentation than the spectrogram. The time and frequency resolution of both spectrogram and MP-WV distribution is higher than original, and components are much clear in case of normal heart sound as well as aortic regurgitation.

In another study [46], authors employed spectrogram along with Renyi entropy for localization and classification. First, TFR using spectrogram is calculated, the optimal value for window length is then calculated using Renyi Entropy Measure (RME). The RME (Eq. 10) is also used for localization of the end of normal heart sound and

start of any pathological sound based on a threshold. The method was tested for normal PCG, early aortic stenosis, late aortic stenosis, and pulmonary stenosis.

$$R_{\alpha} = \frac{1}{1-\alpha} \log_2 \int \int C_x^{\alpha}(t, f) dt df \quad (10)$$

Debbal and Bereksi-Reguig [47] compared time-frequency methods of STFT (short-time frequency transform), CWT (continuous wavelet transform), and Wigner-Ville (WV) distribution primarily for the analysis of second heart sound. Analysis was also carried out for normal and pathological cases of early stenosis aortic analysis, late aortic stenosis, and pulmonary stenosis, etc. Authors concluded that for second heart sound, WV suffers from cross-term limitation while STFT joins the split. CW outperforms the other two techniques both in terms of localization and split. Gavrovska et al. [48] used affine PWVD along with Shannon envelope, ACF, and Haar wavelet lifting for S1 and S2 localization and classification. The basic difference between the normal TFR (time frequency distribution) and affine distribution (time-scale distribution or TSD) is that TFR is covariant to time and frequency of signal while TSD is covariant with time and scale of the signal [49]. The developed technique relies on three main steps. The signal is first down sampled, and coarse detection is performed using PWVD followed by fine detection using Haar wavelet lifting scheme and normalized average Shannon energy (NASE) algorithm. The Haar wavelet-based lifting scheme emphasizes more on low frequency components of the PCG signal under consideration. Finally, identification of S1 and S2 is carried out.

Singh and Dutta [50] used time, probability, and spectrogram for automatic analysis of PCG for cardiac disorders. Barma [51] considered decomposition method called Hilbert Vibration Decomposition (HVD) along with time-frequency representation of smoothed pseudo Wigner-Ville distribution (SPWVD) for detection of third heart sound. HVD was proposed to overcome the limitations associated with Hilbert-Huang transform (HHT). HHT is unable to separate low-energy noise from heart sound which has major energy in low frequency spectrum. It also generates low amplitudes at low frequency which has no physical interpretation. In another work, Zhang et al. [52] proposed a method based on spectrogram and partial least squares regression to classify heart sound into normal, murmur, extra heart sound, extra systole, and artifacts. In an extension of this work [53], authors employed a number of machine learning algorithms for classification of heart sounds.

A summary of TFR-based localization and classification techniques is presented in Table 3. It should be noted that the evaluation datasets considered in [45, 48, 50] are private datasets comprising 11, 70, and 90 signals, respectively. The technique proposed in [46] is evaluated on the E-general Medical dataset with six signals, while no explicit details about the signals used for evaluations is given in [47]. Public datasets PASCAL CHSC 2011, 2016 PhysioNet Challenge, E-General, and Michigan Library data have been employed for evaluations in [51–53]. Gavrovska et al. [48], Zhang et al. [53], and Barma et al. [51] employed precision, efficiency, and recall to quantify the system performance. Among notable contributions, recall of more than 90% and precision of 96.39% is reported in [48]. Likewise, Barma et al. [51] realize an efficiency of 93.9% while Zhang et al. [52, 53] report a varying sensitivity of 50–100% depending upon the number and types of heart sound classes considered in the experiments.

3.2.2 Wavelet transform

Wavelet is a popular method of choice in many fields because of the flexibility it provides in terms of choice of kernel selection which in turn reflects different time and frequency resolutions. Wavelet transform divides the signal into wavelet and scale factor. The formation is given in Eq. 11 [54].

$$f(x) = \sum_k c_{j_0}(k) \varphi_{j_0,k}(x) + \sum_{j=j_0}^{\infty} \sum_k d_j(k) \psi_{j,k}(k) \quad (11)$$

c_{j_0} and d_j are coefficients while $\varphi_{j_0,k}$ and $\psi_{j,k}$ are scaling and wavelet functions, respectively. Haar, Mexican hat, Morlet, Daubechies, and Symlet wavelets are a few of the well-known functions from a very long list of wavelet functions. The different time and frequency resolutions

are usually achieved by using filter banks, an example illustrated in Fig. 9 for S1 beat selected from the PASCAL dataset. In Fig. 9a, first the beat S1 data was converted to Hamming data which was then fed to wavelet transform converting it into one scale and 12 coefficients. The average signal is presented in Fig. 9a while Fig. 9b–d present the low-band, mid-band, and high-frequency band of the beat.

Wavelet transform is a powerful tool that has been employed for pre-processing, localization, and classification. Song et al. [55] denoised the signals in two phases. Environment noise in the heart sounds was removed using the LMS algorithm, while a second step of denoising was carried out using db3 wavelet. Wavelet analysis using db6 order 4 filter was performed for signal reconstruction. The authors concluded that the reconstructed signal gives better characteristics for classification between normal heart sound and murmurs. Features were computed using NASE algorithm, and fuzzy neural network with structure learning classifier was employed for classification [55].

In another study, Sepideh and Geranmayeh [56] considered modeling of the normal heart sound and three pathological disorders, namely, aortic insufficiency, the aortic stenosis, and the pulmonary stenosis sounds. Wavelet db4, order 5 analysis was then carried out on the signals producing statistical features of mean and standard deviation to train an ANN. Features were modified, and the inverse wavelet analysis using the same db4, order 5 was performed which produced the modeled sound [56].

Liang and Hartimo [57] segmented the PCG signal into four parts, i.e., systole, diastole, first heart sound, and second heart sound using wavelet decomposition and reconstruction. Feature vector is then constructed from the original PCG as well as from the third, fourth, and fifth coefficient of db6 order 5. The feature vector based on systole and diastole duration is normalized and fed to ANN classifier for classification of innocent murmur and pathological murmurs [57]. The study was later extended to consider wavelet packet decomposition for heart beat classification [58].

Tu et al. [59] denoised the signal using soft and hard thresholding method with the aim to remove the murmur and environmental noise. The signal was then reconstructed using wavelet db6 and channels 5, 6, and 7. Sound envelop was then extracted using Hilbert transform, and features mainly based on time duration of systole and diastole were used for classification. Gupta et al. [60] proposed a segmentation algorithm based on homomorphic filtering and K-means clustering. The signal was first passed through pre-processing stages comprising normalization and low pass filtering. Segmentation was performed for complete and missed cycle using db2. Features extracted from the segmented signal were then used for classification of normal, systolic, and diastolic murmur.

Table 3 Time-frequency domain

Study	Year	Pre-processing techniques	Localization techniques	Classification techniques	Beat types	Data set	Signals statistics	Evaluation metrics	Results
Zhang et al. [52]	2017	Resampling, filtering, and normalization	Time, spectrogram, and regression	SVM	Normal and abnormal heart sounds	PASCAL CHSC 2011	176 dataset A and 507 dataset B samples	Precision, sensitivity, and specificity	60–76%, 89–100%, and 39–90%
Zhang et al. [53]	2017	Resampling, filtering, and normalization	Time, tensor representation and, feature extraction	SVM	Normal and abnormal heart sounds	PASCAL CHSC 2011 and 2016 PhysioNet challenge	176 dataset A and 507 dataset B samples	Precision, sensitivity, and specificity	25–67%, 44–100%, and 39–64%
Singh et al. [50]	2017	–	Time, probability, and spectrogram-based processing	Time, probability, and spectrogram-based processing	Normal and abnormal heart sounds	National Institute of Health, India	Signal recoding from 45 healthy and 45 patients	–	–
Barna et al. [51]	2016	S3 separation and decomposition by HVD	TFR by SPWVD	Time duration-based classification	82 cycles	Data collected from Texas Heart Institute and Michigan Heart Sound and Murmur Library dataset	–	Efficiency	93.9%
Gavrovska et al. [48]	2014	Signal decimation	Pseudo-affine Wigner-Ville distribution and Haar wavelet lifting, autocorrelation, and Shannon envelopegram	Pseudo-affine Wigner-Ville distribution and Haar wavelet lifting, autocorrelation, and Shannon envelopegram	Normal heart sounds	Data acquisition at University Children Hospital in Belgrade, Serbia	Signal recoding from 35 healthy and 55 patients	Recall and precision	90.41 and 96.39%
Boutana et al. [46]	2011	–	Spectrogram with Reyni marginal entropy	Spectrogram with Reyni marginal entropy	Normal heart sounds, early aortic stenosis, late aortic stenosis, pulmonary stenosis, and mitral regurgitation	E-General Medical	6 signals	–	–
Debbal et al. [47]	2005	–	STFT, CWT, and WD	STFT, CWT, and WD	Normal heart sounds, mitral stenosis, split S2, aortic insufficiency, and stenosis	–	–	–	–
Zhang et al. [45]	1998	–	MP-based Wigner distribution	MP-based Wigner distribution	Normal heart sounds and murmurs	Private dataset	11 signals	Error reduction	–

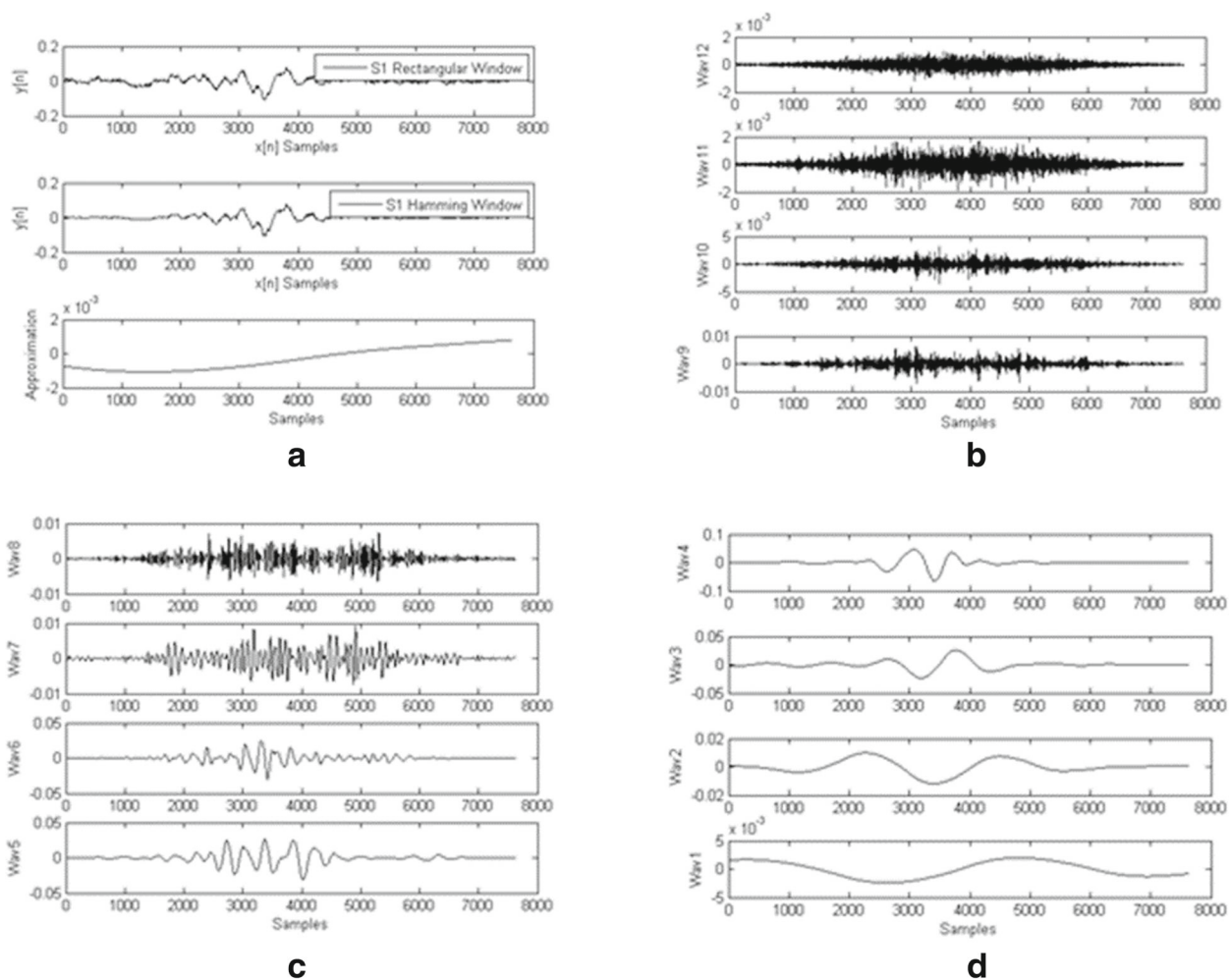


Fig. 9 Low-, mid-band, and high-frequency band extraction of S1 using wavelet transform. **a** S1 beat extracted using rectangular and hamming window with approximation. **b** S1 beat: high-frequency component extracted using wavelet transform. **c** S1 beat: mid-band frequency component extracted using wavelet transform. **d** S1 beat: low-frequency component extracted using wavelet transform

Among other wavelet based contributions, Naseri and Homaeinezhad [61] proposed a framework for classification of heart sounds S1, S2, S3, S4, murmurs, and scuffles. The authors first carry out signal normalization followed by wavelet packet decomposition for noise removal and finally bias removal using Gaussian smoothing filter. Frequency- and amplitude-based features were employed for the detection of different heart sounds.

Pedrosa et al. [62] divide their work into two parts, first is to segment the signal into periodic and non-periodic (noisy) parts, and second is to classify the segmented parts into murmurs or normal beats. Morlet-based pre-processing is used in the segmentation stage and autocorrelation function (ACF) is used for detection of periodic and noisy parts. Any periodic parts are then classified as S1 or S2 using time and time-frequency features. Murmur detection is based on features extracted from Shannon

energy, CWT, DWT, singular value decomposition (SVD), MFCC, bispectrum, variance fractal dimension, and Lyapunov exponents. Zheng et al. [63] carried out pathological signal detection based on DWT, MFCC, and dynamic time warping (DTW). First, DWT is used for envelope segmentation followed by feature extraction using MFCC. Finally, DTW is employed to measure the distance between the signal under test and signal with known pathologies. In another work, Marques et al. [64] investigated stationary wavelet transform followed by hierarchical clustering for localization S1, S2, systole, and diastole.

Deng and Bentley [65] aimed segmentation of normal heart sounds and detection of murmurs. The signal is first down sampled and db4 order 6 is used to extract the beats. Peaks are identified based on the time between the identified systole and diastole periods. The identified peaks are then classified using the time duration between peaks as

feature. Gupta et al. [66] preprocess the signal using normalization and low pass filtering. Segmentation is carried out using homomorphic filtering and K-means clustering. Once segmented, wavelet coefficients using db2 are computed to serve as features. The dimensionality of the feature vector is reduced using PCA. Classification for normal, systolic murmur, and diastolic murmur is carried out using multilayer perceptron-back propagation neural network (MLP-BP).

Abo-Zahhad et al. [67] proposed a human authentication algorithm based on discrete wavelet transform (DWT) for PCG signals. Jain and Tiwari [68] used adaptive thresholding method for denoising of PCG signals. Later, in an extension [69] to this work, adaptive algorithm was used for shrinking of wavelet coefficients for PCG denoising. Goda et al. [70] combined DWT, time features, and other features for classification of heart sounds. Abdollahpur et al. [71] used DWT with MFCC for heart sound classification while Boussaa et al. [72] compared MFCC with DWT for PCG classification. Likewise, Kay and Agarwal [73] used MFCC and continuous wavelet transform (DWT being the discrete counterpart) for heart sound classification.

There is a growing trend towards using wavelet in combination with other operators like Teager energy operator (TEO) and non-negative matrix factorization (NMF). Sattar et al. [74] used NMF for PCG segmentation. Ramovic et al. proposed a system for human authentication based on wavelet and TEO [75] while fetal heart sound detection is carried out by Koutsiana using wavelet and fractal dimensions [76].

A summarized overview of the wavelet based methods discussed in this section is presented in Table 4. Among the studies discussed, classification of normal (S1 and S2) and abnormal heart sounds is considered in [55, 61, 62, 68, 69, 71, 73, 77]. Murmurs were focused in [57, 58, 60–62, 64, 65, 67, 72], while special heart sounds like heart pathologies, scuffles, and artifacts are considered in [56, 63, 65]. Both private and public datasets have been employed for evaluations. Systems presented in [55–61, 66] have been evaluated on private datasets, while works reported in [63, 69] considered a major proportion of the E-General Medical PASCAL 11, 2016 PhysioNet challenge, Egeneral, MIT BIH, and BIOSec public datasets. Likewise, studies presented in [62, 64, 65, 67, 71, 73, 77] employed different subsets of the PHSC 2011 dataset and 2016 PhysioNet challenge. Among evaluation metrics, classical measures including accuracy, PPV, specificity, sensitivity, precision, and classification rate have been mostly used [55–58, 60–62, 65, 71, 73, 77]. Varying results have been reported in these studies ranging from as low as 50% to as high as 100%. Techniques based on wavelets report quantified results even in the earlier studies like [57, 58]

where accuracies of 74.4 and 85% are reported (respectively) on private datasets. On public datasets, accuracies of 85 and 98% are reported in [71, 73] respectively on the PhysioNet database. Among other studies, Jain et al. [68, 69] realize a 100% accuracy on E-General and a private dataset while the work presented in [72] also reports an accuracy of 100% on the MIT BIH dataset. In general, in terms of performance, wavelet-based techniques exhibit a trend similar to that of other techniques where high accuracies are reported on private datasets while the publicly available and more difficult datasets still offer a number of challenges.

3.2.3 EMD and Hilbert-Huang transform

EMD has been a popular choice for the time-frequency analysis in many fields. Unlike other time-frequency algorithms, EMD operates in time domain and operates directly on the signal.

EMD was introduced by Huang et al. [78] along with 2D graphical representation known as Hilbert spectrum for non-linear and non-stationary time series analysis. The algorithm is based on the assumption that each data series is primarily composed of a finite set of simple oscillations which are AM/FM components called intrinsic mode functions (IMF) by sifting process [79, 80]. An IMF must follow the following properties:

- The mean value of an IMF is zero.
- The difference between the number of zero crossings and the number of extrema is at most one.

The first condition implies that IMF should be a narrow band signal, and the second refers to IMF being a mono-component signal [78]. EMD comes in many flavors like bivariate EMD, complex EMD, and multivariable EMD. In our discussion, we will focus only on the standard EMD algorithm [81, 82] which is outlined in the following.

The signal $x(t)$ decomposed by EMD can be represented as follows:

$$x(t) = \sum_{i=1}^N c_i(t) + r(t) \quad (12)$$

Where $c_i(t)$ are IMF and $r(t)$ is the residue. Since $c_i(t)$ is monocomponent, by taking its Hilbert transform, it can be converted to an analytical signal. The derivation is given in the following:

$$X(t) = \sum_{i=1}^N X_i(t) + r(t) = \sum_{i=1}^N c_i(t) + jH[c_i(t)] \quad (13)$$

$X_i(t)$ can be represented with amplitude and phase as given below:

$$X_i(t) = a_i(t).e^{j\theta_i} \quad (14)$$

Table 4 Wavelet domain

Study	Year	Pre-processing techniques	Localization techniques	Classification techniques	Beat types	Data set	Signals statistics	Evaluation metrics	Results
Kay et al. [73]	2017	–	Springer's segmentation algorithm	ANN using beat duration, MFCC, CWT, and complexity features	Normal and abnormal heart sounds	2016 PhysioNet challenge	–	Accuracy	85.2%
Nabih et al. [77]	2017	Reconstruction based on DWT and IDWT	Db 10 wavelet	ANN	Normal heart sounds	PASCAL CHSC 2011 Dataset A and B	170	SNR, RMSE, and PRD	–
Jain et al. [69]	2017	Adaptive method for DWT coefficient reduction for denoising PCG	–	–	Normal heart sounds	E-General Medical and 10 signal privately recorded	–	% detection of S1 and S2	55.1–100%
Mostafa et al. [71]	2016	Downsampling, bandpass filtering, and Schmidt spike removal algorithm [104]	Springer's segmentation algorithm [15]	FDA, ANN, and voting-based classification	Normal and abnormal heart sounds	2016 PhysioNet challenge	4430 signals	Specificity, sensitivity, and over all score	87.2–97%, 81.1–96%, and 84.1–98%
Jain et al. [68]	2016	Adaptive thresholding of DWT coefficient reduction for denoising PCG	–	–	Normal heart sounds	Private dataset composed of simulated and real life data	–	% detection of S1 and S2	86.6–100%
Boussaa et al. [72]	2016	–	–	ANN-based classification using MFCC and DWT features	Normal heart sounds and murmurs	MIT BIH [103]	50 signals from BIH dataset	Correct decision rate	100%
Zahhad et al. [67]	2016	DWT-based denoising	–	Classification based on MFCC, LFCC, WPCC, and NLFCC features	Normal heart sounds and murmurs	BioSec. PCG dataset [105] and PASCAL CHSC 2011 dataset A and B	21 from BioSec. PCG dataset and 206 signals from PASCAL CHSC 2011	Correct recognition rate	–
Pedrosa et al. [62]	2014	Morlet mother-based continuous wavelet denoising	Time domain, time-frequency domain, perceptual analysis, non-linear, and chaos-based analysis	Time domain, time-frequency domain, perceptual analysis, non-linear, and chaos-based analysis	Normal heart sounds and murmurs	Data recorded at Real Hospital Portugues in Recife, Brazil and PASCAL CHSC 2011 dataset A and B	72 Digiscopes signals and 111 signals from PASCAL CHSC 2011	For segmentation: sensitivity and positive predictive, and for murmur: sensitivity and specificity	89.2 and 98.6%; and 52.38–69.67%; and 46.91–79.40%
Zheng et al. [63]	2013	Normalization, DWT and envelopegram	MFCC-based feature extraction	MFCC-based feature extraction	Normal heart sounds with pathological symptoms	E-General Medical	60 signals	Segmentation rate	–

Table 4 Wavelet domain (Continued)

Study	Year	Pre-processing techniques	Localization techniques	Classification techniques	Beat types	Data set	Signals statistics	Evaluation metrics	Results
Marques et al. [64]	2013	–	Stationary wavelet transform to segment the signal	Hierarchical clustering to distinguish S1 and S2	Normal heart sounds	PASCAL CHSC 2011	30 signals from dataset A and 120 signals from dataset B	Error reduction	–
Yiqi Deng, Peter J Bentley [65]	2012	Decompose and reconstruct using wavelet decomposition using 4 level Db6 filters and refilter the signal using spectrogram	Time duration between peaks and peak amplitude	Time duration between peaks and peak amplitude	Normal heart sounds, murmurs, extra sounds and artifact	PASCAL CHSC 2011	Dataset A and dataset B	Precision, sensitivity, and specificity	77.67 and 45.83%, 43.75–50.85%, and 44.44–58.2%
Naseri et al. [61]	2012	baseline removal, wavelet base noise removal, and normalization	Time duration, peak value, and time duration between peaks	Time duration, peak value, and time duration between peaks	Normal heart sounds, murmurs, and scouffles	Rajaei Cardiovacular, Medical and Research Center, Iran	50 signals of total duration of 52 min	Sensitivity and positive predictive value	99.00 and 98.60%
Song et al. [55]	2012	LMS and Db3 wavelet	Time durations extracted from normalized, Shannon envelop formation	Peak and duration based, fuzzy neural network with structure learning (FNNSL)	Normal and abnormal heart sounds	–	14 signals composed of private and public subsets	Classification rate	100% for normal and 50% for abnormal
Zhihai Tu et al. [59]	2010	Db6 wavelet	Heart beat time duration	Hilbert transform-based classification	Normal heart sounds	Shandong University database and real-time device	123 recordings from database and 100 from real-time device	Correct segmentation rate	94.60%
Babaei et al. [56]	2009	–	Db4 and inverse Db4	Db4 and inverse Db4, ANN and statistical classifier	Normal heart sounds, aortic insufficiency, pulmonary stenosis	Tehran Heart Center, Iran	372 cycles	Accuracy	94.42%
Gupta et al. [60]	2007	Signal decimation, low pass filtering and signal normalization	Homomorphic filtering, time duration and K-means	Wavelet Db2-based features, GAL, and MLP-BP	Normal heart sounds, systolic murmur, and diastolic murmur	Singapore General Hospital, Singapore	41 signals composed of 340 cycles	Classification rate	90.29%
Gupta et al. [66]	2005	Signal normalization	Peak detection based on homomorphic filtering and K-means	Db2 wavelet-based features and MLP-BP	Normal heart sounds, systolic murmur, and diastolic murmur	Data recorded at Singapore General Hospital, Singapore	340 cycles	Classification and segmentation accuracy	94.5 and 90.45%
Liong et al. [58]	1998	–	wavelet packet decomposition	Wavelet packet decomposition	Pathological and physiological murmurs	–	85 signal of 7–12 s	Accuracy	85%
Huiying et al. [57]	1998	Decompose and reconstruct signal using Daubechies wavelet	Shannon energy envelop-based S1 and S2 location	Feature vector based on systole time duration, diastole time duration, and ANN	Normal heart sounds and murmurs	Private dataset	78 recordings	Accuracy	74.4%

Algorithm 1: Standard EMD algorithm

-
- ```

1 function Standard EMD algorithm($x(t)$);
 Input : $x(t)$
 Output: IMF
2 Let $x(t)$ is original signal and equalize it to $x_1(t)$. Identify all
 local minima and all maxima of $x(t)$.
3 Construct a lower envelop $e_l(t)$ and upper envelope $e_u(t)$
 using interpolation.
4 Calculate the local mean $m(t) = (e_l(t) + e_u(t)) / 2$.
5 Subtract the local mean from $x_1(t)$ and obtain the IMF
 ($c_i(t) = x_1(t) - m(t)$). Where 'i' is the order of IMF.
6 If $c_i(t)$ is not an IMF, then $x_1(t) = c_i(t)$ and go to Step 2 and
 till $c_i(t)$ becomes an IMF.
7 If an IMF is obtained, then calculate residue $r(t) =$
 $x_1(t) - c_i(t)$. If $r(t)$ meets the criteria, then signal is
 decomposed into IMFs and residue else put $x_1(t) = r(t)$ and
 go to Step 2.
```
- 

Where

$$\theta_i = \arctan\left(\frac{H[c_i(t)]}{c_i(t)}\right)$$

and

$$a_i(t) = \sqrt{c_i^2(t) + H[c_i^2(t)]}$$

If phase  $\theta_i$  is differentiated with the respect to time  $t$ , the result is instantaneous frequency  $\omega_i$ . The plot of  $\theta_i$  vs  $\omega_i$  is the graphical representation known as Hilbert-Huang transform [78–80]. Figure 10 shows an actual S1 beat selected from a data file in the iStethoscope dataset. Two plots are shown, one for the actual rectangular data and other from hamming-based data of the same rectangular window. b–d present the low-band, mid-band, and high-frequency band of the beat using standard EMD algorithm.

Among EMD-based techniques, Salman et al. [83] first filter the signal using low pass filter followed by EMD-based denoising. Authors claim that EMD offers better values for a number of noise measures including signal-to-noise ratio (SNR), root mean square error (RMSE), and percent root mean square difference (PRD). Cardiac cycle was calculated using autocorrelation of normalized Hilbert transformed signal, and segmentation was carried out using Shannon energy. Finally, the time duration between peaks was used for systole and diastole identification from which different beats like S1, S2, S3, and S4 were classified [83].

In another work, Zhao et al. [84] employ two methods for instantaneous frequency calculation. One based on Hilbert transform, and the other based on TFD (EMD). The Hilbert-based approach fails for wide band signals while EMD reports satisfactory results. Authors claim to make use of EMD in its basic form to detect coronary

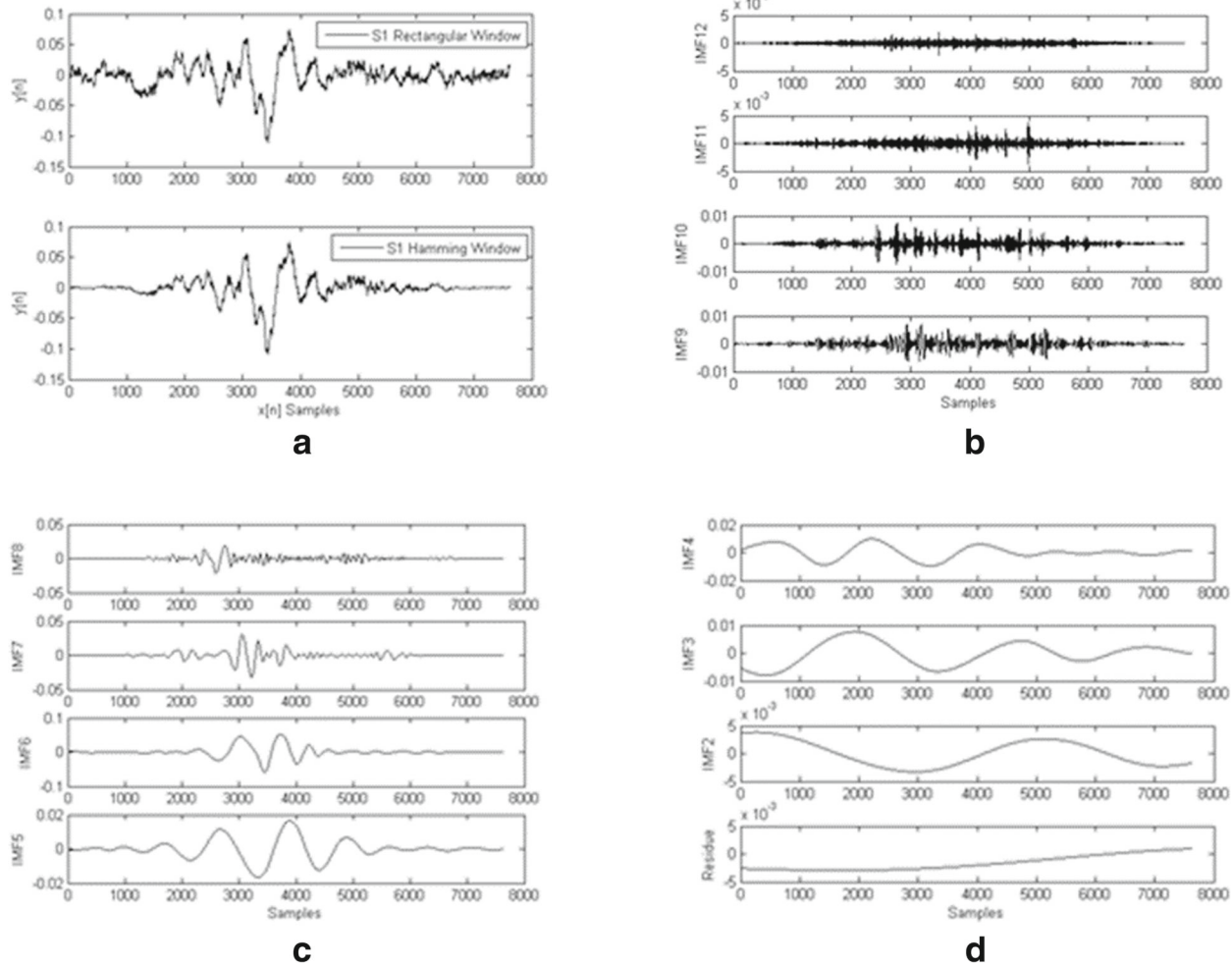
artery disease based on the instantaneous frequency calculated from diastolic murmurs using EMD and SVM. In another study [85] by the same authors, signal was denoised using db5 and, ensemble EMD (E-EMD) was then applied. EEMD removes the mode mixing problem in traditional EMD. Marginal spectrum was calculated on the Hilbert-Huang spectrum of EEMD output followed by DCT quantization and normalization of the marginal spectrum. Vector quantizer (VQ) was trained using Linde-Buzo-Gray algorithm, and classification was carried out using Euclidean distance.

One of the early attempts to analyze biomedical signals (including PCG) using EMD was made by Sun et al. [86]. The authors argued that EMD is a powerful yet little explored tool for analysis of different biomedical signals. Later, the authors investigated EMD decomposition for instantaneous frequency estimation of PCG signals and concluded that EMD being a physical decomposition is more useful than mathematical decompositions like wavelet transform [87]. Likewise, Gavrovska et al. [88] employed EMD with wavelet for PCG denoising.

Among other contributions, Moukhadem et al. [89] and Sun et al. [90] proposed algorithms for the classification of first and second heart sounds. In [90], Sun et al. used a combination of wavelet and EMD for classification. Signal is first decomposed using EMD, and the highest frequency IMF is denoised using db7 wavelet. The output signal is then reconstructed using denoised channels, and all other IMFs and NASE algorithm is applied on the resultant signal. In addition, cross correlation is calculated between the original signal and all IMFs. The IMF reporting maximum correlation is chosen and NASE is again calculated. Using separate thresholds, endpoints of S1 and S2 are detected from both denoised NASE and maximum correlation NASE signals. In [89], after pre-processing, SVD is used for feature extraction from the output of EMD and S-transform. SVD is also applied to the S-Matrix calculated from S-transform followed by k-NN for classification.

Boutana et al. [91] exploited EMD for heart sound segmentation and used it for clinical cases of late aortic stenosis, early aortic stenosis, and mitral regurgitation. The method for selection of IMFs is based on noise-only model which assumes that if the noise is additive white Gaussian noise, the logarithm-variance of each IMF varies but the variation is linear and with a parameter called the Hurst exponent. Papadaniil and Hadjileontiadis [92] employed ensemble EMD along with kurtosis for heart sound segmentation. Low pass and median filtering is first used for noise removal. Afterwards, EEMD-based IMFs are calculated. The IMFs which do not meet any of the energy criteria, instantaneous frequency criteria, and bootstrap kurtosis-based criteria are removed. S1 and S2 are then classified based on the Kurtosis measure.





**Fig. 10** EMD-based low-, mid-band, and high-frequency extraction of a sample S1 beat. **a** S1 heart beat signal extracted using rectangular and hamming window. **b** S1 heart beat: high-frequency component extracted using EMD transform. **c** S1 heart beat: mid-band frequency component extracted using EMD transform. **d** S1 heart beat: low-frequency component extracted using EMD transform

Authors in [93] suggested that translated EMD outperforms the traditional EMD. Translated EMD takes the signal, modulates it, applies EMD, and then demodulates all IMFs. The results are presented for simulated as well as real PCG data. In a relatively recent study, Jimenez et al. [94] first normalize and resample the PCG signals. Three different types of EMD analysis, namely, EMD, EEMD, and adaptive EEMD are then carried out. Afterwards,  $S\_MFCC$  obtained from the signal,  $ST\_MFCC$  calculated from energy operator in the frequency domain,  $SW\_MFCC$  from frequency bands of the spectral energy distribution, and  $SWT\_MFCC$  from combination of all previous are calculated from two signals which are sum of the even and odd IMFs. Fuzzy rough set (FRS) algorithm is then employed to reduce statistical moments obtained from HHT. At the end, classification is carried out using ergodic HMM.

Banerjee et al. [95] used variational mode decomposition along with Shannon energy feature for heart sound localization. Variational mode decomposition was developed by Dragomiretskiy and Zosso [96] to remove noise sensitivity and sampling problems that accompany standard EMD. Salman et al. [97] used EMD for removal of white, colored, exponential, and alpha-stable noise and showed that EMD is superior when compared with wavelet and total variation denoising methods. Heart murmurs were detected by Jusak et al. [98] using complete ensemble empirical mode decomposition (CEEMD, another variant of standard EMD) and the Pearson distance metric. Authors concluded that CEEMD is computationally more complex as it extracts more modes in comparison to EMD.

Research efforts are being continuously made to enhance the EMD algorithm and proposition of new

filtering techniques for EMD. A number of variants have been proposed to handle the overshoot and undershoot end effects related to the classical EMD algorithm. These effects are demonstrated in Fig. 11. The figure shows that upper envelop is well above while the lower envelop is fairly below the original signal. Both of these introduce errors in the mean envelop. Another problem in EMD-based techniques is associated with the stopping criteria. If the algorithm stops premature, it will not reveal all the details in signal, and if it continues for too many iterations, the physically meaningless IMF would appear. The third major problem with EMD is its decomposition sequence. EMD decomposes from high frequency towards lower frequency but not from high energy towards lower energy. This effect can be minimized by using HHT transform along with EMD [99].

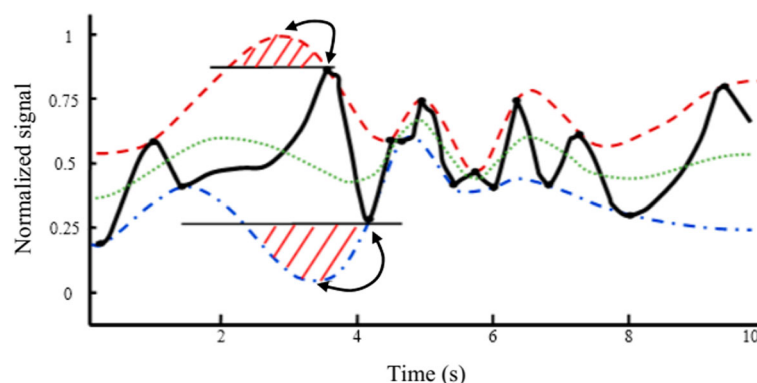
A summary of the methods discussed in this section is presented in Table 5. It should be noted that all types of heart beat have been considered in the EMD-based research. Normal heart beats of class S1 and S2 are considered in most of the studies [83–87, 89, 90, 92–95, 97, 100]. Additionally, S3 and S4 sounds are considered in [83] while S3 and S4 with gallop are considered in [101]. Studies in [86, 87] focus on abnormal heart sounds whereas various pathological states like regurgitation and stenosis are investigated in [91, 92, 100]. Likewise, murmurs have been taken into account in [94, 98, 100].

Techniques proposed in most of the aforementioned studies have been evaluated on private datasets. Only the works presented in [91, 100] are evaluated on the publicly available E-General Medical dataset. Likewise, systems reported in [83, 97, 98, 101] employed the University of Michigan dataset while PASCAL CHSC 2011 dataset is considered in [95]. Standard metrics of accuracy, specificity, sensitivity, etc. have been employed in most cases. Specific measures like mean prediction power and mean accuracy [92], correct recognition rate

(CRR) [85], correct and incorrect diagnosis [84], and SNR and ratio R [88] have also been reported. EMD algorithm is compared with other time-frequency algorithms like wavelet and DWT in [87, 88, 101]. Studies reported in [97, 98] used SNR and  $\Delta$ SNR while [95] employed the average detection rate. Gavrovska et al. [88] compared DWT, EMD, and EEMD, and concluded that EEMD reports the best performance among the three. Similarly, authors in [101] compared EMD with other denoising techniques and argued that EMD outperforms other methods. For E-General Medical public dataset, Varghees and Ramachandran [100] reported average sensitivity and positive predictivity of more than 90% while for private datasets, accuracies of as high as 99% are also reported [90].

### 3.2.4 Application of multiple time-frequency methods.

A recent trend in automated analysis of PCG signals is investigation of multiple time-frequency representations for effective classification and localization. Gavrovska et al. [48], for example, combined time-frequency representation of Wigner-Ville with Haar wavelet for normal heart beat detection in pediatric patients. Authors report a precision and recall of more 90% for 90 heart beat signals collected from 55 patients and 35 healthy subjects. The idea of multiple time-frequency methods was also exploited by Sun and Gong [90], and Gong and Nie [102] where the authors who employ EMD with wavelets for separation of S1 and S2 from noisy heart sounds. Both the studies [90, 102] report high accuracies of more than 99% and demonstrate the effectiveness of combining multiple time-frequency methods over single representation. It should however be noted that although these multiple time-frequency methods realize high accuracies, they are computationally expensive and hence, could not gain significant research attention of research community. Furthermore, the choice of different time-frequency methods and



**Fig. 11** Demonstration of overshoot and undershoot problems related with EMD algorithm

**Table 5** EMD-based research

| Study                  | Year | Pre-processing techniques                                | Localization techniques                                                                                                  | Classification techniques                                                                                                | Beat types                                                                                                                                                                                                                                                          | Data set                                                                      | Signals statistics                 | Evaluation metrics                                                                                        | Results                   |
|------------------------|------|----------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------|------------------------------------|-----------------------------------------------------------------------------------------------------------|---------------------------|
| Jusak et al. [98]      | 2016 | CEEMD and EEMD                                           | Pearson distance metric                                                                                                  | Pearson distance metric                                                                                                  | Murmur                                                                                                                                                                                                                                                              | University of Michigan Heart Sound and Murmur Library [26]                    | 10 signals                         | $\Delta$ SNR                                                                                              | –                         |
| Salman et al. [97]     | 2016 | EMD denoising                                            | –                                                                                                                        | –                                                                                                                        | Synthetic normal and abnormal heart sound corrupted by noise                                                                                                                                                                                                        | University of Michigan Heart Sound and Murmur Library [26]                    | –                                  | Signal-to-noise ratio (SNR), root mean square error (RMSE), and percent root mean square difference (PRD) | –                         |
| Banerjee et al. [95]   | 2016 | Median and low pass filtering                            | VMD, normalization, Shannon energy, and dynamic thresholding                                                             | VMD, normalization, Shannon energy, and dynamic thresholding                                                             | Normal and abnormal heart sounds                                                                                                                                                                                                                                    | PASCAL CSHC 2011 dataset A and B                                              | 10 signals                         | Average detection time error                                                                              | –                         |
| Salman et al. [83]     | 2015 | Signal decimation, low pass filtering, and EMD denoising | Time-based peak detection of Shannon energy envelope, and Heron's formula                                                | Time duration between peaks                                                                                              | Normal heart sounds                                                                                                                                                                                                                                                 | University of Michigan Heart Sound and Murmur Library                         | 8 signals of 70 s                  | –                                                                                                         | –                         |
| Salman et al. [101]    | 2015 | Wavelet transform, total variation, and EMD              | Time-based peak detection of Shannon energy envelope and Heron's formula                                                 | Wavelet transform, total variation, and EMD                                                                              | Normal heart sounds, S3, and S4 gallop                                                                                                                                                                                                                              | University of Michigan Heart Sound and Murmur library                         | –                                  | –                                                                                                         | –                         |
| Papadaniil et al. [92] | 2014 | Median and low pass filtering                            | Energy, instantaneous frequency, bootstrap, and kurtosis-based criterion                                                 | Kurtosis-based S1 and S2 detection                                                                                       | Normal heart sounds, aortic stenosis, and mitral regurgitation                                                                                                                                                                                                      | First Cardiac Clinic of Papanikolaou General Hospital in Thessaloniki, Greece | 43 signals composed of 2602 cycles | Mean prediction Power and mean accuracy                                                                   | 74–100% and 45–100%       |
| Varghees et al. [100]  | 2014 | Low pass filtering and total variational filtering       | Shannon entropy envelope computation, instantaneous phase-based boundary determination, and boundary location adjustment | Shannon entropy envelope computation, instantaneous phase-based boundary determination, and boundary location adjustment | Normal heart sounds, fixed and wide split, early systolic, late systolic and pan systolic, mitral and tricuspid stenosis, aortic and pulmonic regurgitation, mitral prolapse, ejection murmur and clicks, and different time-varying systolic and diastolic murmurs | E-General Medical                                                             | –                                  | Average sensitivity and positive predictivity                                                             | 99.43 and 93.56%          |
| Jimenez et al. [94]    | 2014 | Resampling, segmentation, and normalization              | EMD, EEMD EMD with adaptive noise                                                                                        | Features based on ergodic HMM, MFCC, and statistical moments of HHT                                                      | Normal heart sounds and murmurs                                                                                                                                                                                                                                     | Private dataset                                                               | 400 beats                          | Accuracy, specificity, and sensitivity                                                                    | For all measures above 95 |

**Table 5** EMD-based research (Continued)

| Study                 | Year | Pre-processing techniques                  | Localization techniques                                                              | Classification techniques                                                                   | Beat types                                                                    | Data set                                                                                  | Signals statistics        | Evaluation metrics                     | Results                                                           |
|-----------------------|------|--------------------------------------------|--------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------|---------------------------|----------------------------------------|-------------------------------------------------------------------|
| Gavrovska et al. [88] | 2013 | Wavelet-based denoising and lifting scheme | DWT, EMD, and EEMD                                                                   | DWT, EMD, and EEMD                                                                          | Normal heart sounds and clicks                                                | –                                                                                         | –                         | Comparison in terms of SNR and ratio R | –                                                                 |
| Boutana et al. [91]   | 2013 | –                                          | EMD                                                                                  | EMD                                                                                         | Early aortic stenosis, late aortic stenosis, mitral, and aortic regurgitation | E-General Medical                                                                         | 4 signals                 | –                                      | –                                                                 |
| Sun et al. [90]       | 2013 | EMD followed by wavelet decomposition      | Shannon envelop from denoised signal as well as from EMD channels of original signal | Shannon energy envelop from denoised signal as well as from EMD channels of original signal | Normal heart sounds                                                           | –                                                                                         | 30 signals                | Accuracy                               | 99.74%                                                            |
| Zhao et al. [85]      | 2013 | Denoising and framing                      | Hilbert-Huang transform, normalization, and DCT                                      | VQ with LBG algorithm                                                                       | Normal Heart sounds                                                           | 40 participants and 7 heart sounds for each collected by the University of Catania, Italy | 280 samples of 10 s each  | Correct recognition rate               | 94.16% in marginal spectrum system and 84.93% in Fourier spectrum |
| Moukadem et al. [89]  | 2012 | Decimation and high pass filtering         | S-transform and radial basis functions (RBF) neural network                          | Features based on SVD performed on S-transform, EMD, and KNN                                | Normal heart sounds                                                           | Data collected at Hospital of Strasbourg, France                                          | 80 signals of 8–12 s each | Sensitivity and specificity            | 94 and 97%                                                        |
| Segalen et al. [93]   | 2011 | –                                          | Modulated-EMD                                                                        | Modulated -EMD                                                                              | Normal Heart sound                                                            | Private dataset                                                                           | 20 signals                | –                                      | –                                                                 |
| Zhao Zhidong [84]     | 2005 | –                                          | Instantaneous frequency of diastolic murmurs                                         | Instantaneous frequency of diastolic murmurs and SVM                                        | Normal heart sounds and coronary artery disease                               | Private dataset                                                                           | 77 signals                | Correct and incorrect diagnosis        | 17 out 20 for abnormal and all normal correctly diagnosed         |
| Sun et al. [87]       | 2005 | –                                          | Hilbert transform on IMF given by EMD                                                | Hilbert transform on IMF given by EMD                                                       | Normal and abnormal heart sounds                                              | Private dataset                                                                           | 2 signals                 | –                                      | –                                                                 |
| Sun et al. [86]       | 2004 | –                                          | EMD                                                                                  | EMD                                                                                         | Normal and abnormal heart sounds                                              | Private dataset                                                                           | 2 signals                 | –                                      | –                                                                 |

their sequence of application also remains a challenging problem.

Comparing different time-frequency representations, it can be observed from Table 3, SPWVD outperforms other representations like spectrogram. Among wavelet-based techniques, a large number of comprehensive investigations have been carried out using a variety of wavelet filters. DWT, CWT, and Db filters have been most popular choices reporting high localization and classification accuracies. Likewise, EMD along with its various enhanced variants has remained an attractive choice for researchers for this problem. Lately, techniques based on combination of EMD and wavelets have also been explored and are known to report higher accuracies as compared to EMD or wavelet-based methods. Further discussion on comparison of different techniques is presented later in the next section.

#### 4 Review

This section presents an analysis of time, frequency, and time-frequency methods discussed in the previous sections. Time-based methods have been used extensively in the past and have still maintained their popularity because of ease with which information can be extracted. They can provide, in theory, the best localization for beats in the signal under observation. These methods need pre-processing steps to make the signal fit for subsequent steps of localization and classification. With the development of machine learning techniques, time domain methods enjoy a renewed interest which is expected to continue in the near future. This trend is very well demonstrated by the recent research worked presented in Table 2. Quantitative comparison of results also shows that time-based methods, in general, are able to localize and classify beats with more than 70% specificity.

The frequency-based methods for localization and classification usually start with MFCC which employs pre-emphasis, windowing and FFT to make the signal appropriate for classification and localization. An analysis of the quantitative results reported in Table 2 shows that on the average, the performance of frequency-based methods is more or less similar to that of time-based methods. It should however be noted that frequency-based methods are best suited for frequency band localization but fail to extract the time location of the frequency band under observation. Time-based methods, on the other hand, provide beat location in the signal but are limited in the sense that they do not provide explicit information about the frequency content of the signal.

The time-frequency methods decompose signals into different time and frequency resolutions and aim to overcome the limitations of time- and frequency-based methods. The time-frequency methods investigated in the literature include wavelets, EMD, and TFR. Although TFR

in its basic form suffers from noise due to cross terms, a number of effective techniques based on TFR are proposed for heart beat analysis. These methods attempt to suppress the cross terms while keeping as much frequency resolution as possible. The recent popularity of TFR methods is generally attributed to pseudo WV distribution and HHT which serve to reduce the noise due to cross terms and allow a better representation of the signal under study. A comparison from Table 3 shows that the results reported for TFR methods exhibit high variation, mainly as a function of the database employed for evaluation.

Among other time-frequency methods, EMD and wavelets have been mostly employed for PCG analysis. EMD is a data driven decomposition technique that decomposes a signal from high to low frequencies by generating lower and upper envelopes using spline interpolation. EMD-based techniques suffer from end-effects of undershoot and overshoot, the modeling errors inherited in the EMD algorithm. The stopping criteria for EMD are being continuously researched to extract IMFs which are of physical significance. EMD has witnessed a number of variations over the years. In addition to the standard EMD, noise-assisted EMD, ensemble EMD, multi-variable EMD, and EMD complemented by HHT have also been investigated. It can be observed from Table 5 that in general, high accuracies are reported by EMD-based techniques. The results however are not directly comparable as diverse datasets have been employed by different researchers.

In addition to EMD, wavelets and their variants have been a popular as well as an effective choice of researchers for analysis of PCG signals. Wavelet is a goal-driven algorithm which decomposes a signal using dyadic filter bank. This decomposition method is limited in the sense as it does not take into account the parameters of the signal under decomposition. Despite this limitation, wavelets have been employed in all stages from pre-processing to classification.

It can be seen from Table 4 that wavelet-based techniques have been comprehensively studied and evaluated on wide variety of datasets including all three public datasets presented in this paper. These methods consistently report high localization and classification accuracies on multiple datasets.

Summarizing the key findings of our analysis, the initial attraction of employing time-frequency based methods like EMD and wavelets was their ability to represent the signal at multiple resolutions unlike time and frequency analysis. The complexity of PCG signal, however, forced the researchers to borrow techniques and features from the time and frequency domains. Features like systole and diastole time duration, beat amplitude, and beat frequency, for instance, are in common use. In general, all techniques have evolved to an extent where they are able to successfully model human heart beat under clean



environment. Nevertheless, automatic analysis of noisy heart beats still remains a challenging problem. Time and frequency methods, in most cases, are dependent on machine learning algorithms to enhance the localization and classification performance. Features extracted in the time or frequency domain are typically fed to a learning algorithm and the choice of the learning algorithm also influences the overall system performance. The time-frequency methods (especially EMD and wavelets), on the other hand, are not too sensitive to the choice of the learning algorithm. Furthermore, time and frequency methods have been mostly limited to the two-class problem of classifying the heart beats as normal or abnormal while wavelets and EMD present more sophisticated solutions detecting not only the abnormality but also classifying the pathology.

Studying the quantitative performance of time, frequency, and time-frequency methods (presented in the respective tables), it can be seen that a direct comparison of these methods is difficult due to the different types of challenges offered by each dataset. A general observation is that for small datasets and clean signals, time, frequency, and time-frequency-based methods report similar accuracies. For larger datasets and noisy signals, however, the performance of time-frequency methods, especially wavelets, remains relatively stable once compared to other methods. The PASCAL 2011 dataset A, for instance, is considered to be a very challenging set of noisy signals. Only a limited number of studies [30, 52, 53, 62, 67, 77, 95] have been evaluated on this set and among these wavelets-based techniques report the highest accuracies.

Time-frequency representations like Wigner-Ville and pseudo WV have remained relatively less explored primarily due to the problem of cross terms. Reducing the cross terms while keeping the maximum possible information in signal needs to be further investigated. EMD and wavelet-based techniques enjoy the advantage over other techniques in the sense that they decompose the signal at multiple resolutions hence removing the high-frequency noise and reducing the energy contribution from low frequencies. These noise handling capabilities make such techniques an attractive choice, especially, when dealing with noisy signals. Despite these characteristics, the problem of computerized analysis of PCG signals still remains very challenging for noisy environments. A clear evidence is the performance of different systems on the noisy signals of the PASCAL 2011 datasets which offer a great margin for enhancement. Another interesting direction could be to investigate the combination of EMD and wavelets, for instance, techniques like empirical-wavelet decomposition can be employed for analysis of noisy signals. Likewise, combination of algorithms and features from time, frequency, and time-

frequency domains can also be investigated to propose robust localization and classification techniques which can deal with different types of signals (clean, noisy, various pathologies etc.). In our present work, we are exploring such a combination where EMD or wavelets (time-frequency domain) are being employed for noise removal while localization is carried out using Springer algorithm [15] (time domain). Furthermore, a combination of features extracted from the three representations is intended to be fed to different machine learning algorithms for classification. Such rich representations are likely to enhance the localization and classification accuracies of the system.

It is also worth mentioning that the primary focus of most of the research on automatic analysis of PCG signals has been on enhancing the localization and classification accuracies. From the view point of practical applications, development of computationally efficient solutions which may work in real time is also a challenging problem that needs further exploration. In addition, the presently available PCG datasets comprise a limited number of samples and do not cover the complete range of pathologies which are likely to be encountered in clinical settings. This necessitates the development and labeling of a comprehensive dataset of PCG signals encompassing a variety of signals and covering all major pathologies. Considering the complexity of the problem, modeling the heart beats and various pathologies is likely to offer more robust solutions as opposed to the conventional techniques relying on localization and classification. To the best of authors' knowledge, very limited efforts have been made in this direction. Morlet wavelet filter, for instance, has been investigated to model the heart beat. However, modeling of such complex signals brings along its own challenges and requires significant amount of data for each pathology, a requirement that is hard to meet in the currently available datasets.

## 5 Conclusions

This paper presented an overview of the techniques proposed for computerized analysis of PCG signals which represent recordings of heart sound. Localization and classification of beats have been the key research areas with the objective to discriminate between normal and abnormal heart sounds. The variation in amplitude, frequency, and duration of beats makes PCG a very complex signal for automatic analysis. The domain has witnessed more than three decades of research, and this paper is an attempt to provide an overview of the current state-of-the-art on this subject. We organized the notable contributions to automatic analysis of PCG signals as function of domains of analysis, namely, time, frequency, and time-frequency methods. An analysis of the review techniques revealed that, in general, the performance of

time, frequency, and time-frequency is similar for small datasets and clean signals. The more challenging scenario is analysis of noisy signals where time-frequency specially EMD and wavelets has been popular choice of researchers. These methods, however, bring with them the additional computational cost and algorithmic complexity. Consequently, simple features like amplitude, energy, beat duration, and spectral flux extracted from time and frequency domain continue to sustain. These time and frequency methods have been complemented by using sophisticated machine learning to enhance the localization and classification performance.

### Abbreviations

ACF: Autocorrelation function; ANN: Artificial neural network; AR: Aortic regurgitation; AS: Aortic stenosis; CEEMD: Complete ensemble empirical mode decomposition; CNN: Convolutional neural networks; CRR: Correct recognition rate; CVD: Cardiovascular diseases; DCT: Discrete cosine transform; DTW: Dynamic time warping; DWT: Discrete wavelet transform; ECG: Electrocardiogram; EEMD: Ensemble empirical mode decomposition; EMD: Empirical mode decomposition; FFT: Fast fourier transform; HHT: Hilbert-Huang transform; HMM: Hidden Markov model; HSC: Heart sound challenge; HVD: Hilbert vibration decomposition; IFFT: Inverse fast fourier transform; IMF: Intrinsic mode functions; LMS: Least mean square; LPC: Linear predictive coding; MCS: Modified cuckoo search; MFCCS: Mel-frequency cepstral coefficients; MLP-BP: Multilayer perceptron-back propagation neural network; MP: Matching pursuit; MR: Mitral regurgitation; NASE: Normalized average Shannon energy; NMF: Non-negative matrix factorization; PCG: Phonocardiogram; PDA: Patent ductus arteriosus; PPG: Photoplethysmography; PPV: Positive predictive value; PR: Pulmonary regurgitation; PRD: Percent root mean square difference; PS: Pulmonary stenosis; RME: Renyi entropy measure; RMSE: Root mean square error; S1: First normal heart sound; S2: Second normal heart sound; S3: Third abnormal heart sound; S4: Fourth abnormal heart sound; SNR: Signal-to-noise ratio; SOM: Self organizing map; SPWVD: Smoothed Pseudo Wigner-Ville distribution; STFT: Short-time frequency transform; SVD: Singular value decomposition; SVM: Support vector machine; TEO: Teager energy operator; TFR: Time-frequency representation; TSD: Time-scale distribution; VQ: Vector quantizer; VSD: Ventricular septal defect; WD: Wigner distribution; WHO: World Health Organization

### Acknowledgements

The authors would like to thank Bahria University, Islamabad, Pakistan, for providing us with the opportunity to carry out the reported research.

### Funding

The reported research is not a part of any funded project.

### Availability of data and materials

The review referred to the following datasets which are publicly available.

- 2015 IEEE Signal processing cup. Zhilin Zhang. 2015. <https://sites.google.com/site/researchbyzhang/ieeespcup2015> <http://www.ubcieee.org/events/signal-processing-cup-2015>
- PASCAL Classifying heart sounds challenge. Peter Bentley et al. 2011. <http://www.peterjbentley.com/heartchallenge/>
- eGeneral Medical Inc.USA. eGeneralMedical.com. <http://www.egeneralmedical.com/listtohearmur.html>
- Michigan Medicine. Heart Sound and Murmur Library University of Michigan Health System. Ann Arbor. MI. [http://www.med.umich.edu/irc/psb\\_open/html/repo/primer\\_heartsound/primer\\_heartsound.html](http://www.med.umich.edu/irc/psb_open/html/repo/primer_heartsound/primer_heartsound.html)
- National Institute of General Medical Sciences and the National Institute of Biomedical Imaging and Bioengineering. PhysioNet/CinC challenge 2016. NIGMS & NIBIB. 2016. <https://www.physionet.org/physiobank/database/challenge/2016/>
- University of Toronto BioSec Lab. BioSec. PCG database. Lab BioSec. <http://www.comm.utoronto.ca/~biometrics/databases.html>

Authors referred to following implementation in the manuscript.

- ASTRES : Analysis, synthesis and transformations by reassignment, EMD and synchrosqueezing. [http://www.ens-lyon.fr/PHYSIQUE/Equipe3/ANR\\_ASTRES/](http://www.ens-lyon.fr/PHYSIQUE/Equipe3/ANR_ASTRES/)
- Massachusetts institute of technology. EMD. <http://www.mit.edu/~gari/CODE/HRV/emd.m>

### Authors' contributions

SI, PhD candidate at Bahria University, Islamabad, Pakistan, carried out a comprehensive review on localization and classification of heart beats in PCG signals as a part of his PhD research work. UA and IS, supervised the research, contributed to paper writing, paper organization, and critical analysis of the reviewed literature. All authors read and approved the final manuscript.

### Competing interests

Authors declare that they have no identifiable conflicting or competing interest.

### Publisher's Note

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

### Author details

<sup>1</sup>Bahria University, Shangrilla Road, Islamabad, Pakistan. <sup>2</sup>EME, National University of Sciences and Technology, Rawalpindi, Pakistan. <sup>3</sup>APCOMS, UET, Taxila, Rawalpindi, Pakistan.

Received: 20 November 2017 Accepted: 10 April 2018

Published online: 27 April 2018

### References

1. Mortality due Cardiovascular diseases in world. World Health Organization. <https://www.medicalnewstoday.com/articles/282929.php>
2. Z Zhang, IEEE Signal Processing Cup (2015). <http://www.ubcieee.org/events/signal-processing-cup-2015>. <https://sites.google.com/site/researchbyzhang/ieeespcup2015>
3. F Castells, et al, Principal component analysis in ECG signal processing. *EURASIP J. Adv. Signal Process.* **2007**, 074580 (2007)
4. F Minhthang Bui, D Hatzinakos, Biometric methods for secure communications in body sensor networks: resource-efficient key management and signal-level data scrambling. *EURASIP J. Adv. Signal Process.* 529879 (2008). <https://doi.org/10.1155/2008/529879>
5. M Sepulveda-Cano, et al, Selection of nonstationary dynamic features for obstructive sleep apnoea detection in children. *EURASIP J. Adv. Signal Process.* 538314 (2011). <https://doi.org/10.1155/2011/538314>
6. S Nemat, A Malhotra, GD Clifford, Data fusion for improved respiration rate estimation. *EURASIP J. Adv. Signal Process.* 2010. 926305 (2010). <http://doi.org/10.1155/2010/926305>
7. SM Debbal, F Bereksi Reguig, Computerized heart sounds analysis. *Comput. Biol. Med.* **38**, 263–280 (2008)
8. AH Osamah, et al, Multimodal biometrics based on identification and verification system. *J. Biom. Biostat.* **4**, 2 (2013)
9. W Phanphaisarn, et al, Heart detection and diagnosis based on ECG and EPCG relationships. *Med. Devices (Auckl)*. **4**, 133–144 (2011)
10. S Maderwald, et al, Is Phonocardiogram gating a reliable alternative to ECG gating in clinical routine for CINE and velocity-encoded phase contrast imaging? *Proc. Intl. Soc. Mag. Reson. Med.* **20** (2012)
11. LS Stodieck, MW Luttgies, Relationships between the electrocardiogram and phonocardiogram: potential for improved heart monitoring. *ISA Trans.* **23**(4), 59–65 (1984)
12. I Maglogiannis, et al, Support vectors machine-based identification of heart valve diseases using heart sounds. *Comput. Meth. Prog. Biomed.* **95**, 47–61 (2009)
13. H Moghaddasi, F Almasganj, A Zoroufian, Imaging of heart acoustic based on the sub-space methods using a microphone array. *Comput. Meth. Prog. Biomed.* **146**, 133–42 (2017)
14. A Sepehri Amir, et al, A novel method for pediatric heart sound segmentation without using the ECG. *Comput. Methods Prog. Biomed.* **99**, 43–48 (2010)
15. DB Springer, L Tarassenko, GD Clifford, Logistic regression-HSMM-based heart sound segmentation. *IEEE Trans. Biomed. Eng.* **63**, 822–832 (2016)

16. S Chauhan, et al, A computer-aided MFCC-based HMM system for automatic auscultation. *Comput. Biol. Med.* **38**, 221–33 (2008)
17. MA Kotb, H Nabih, F El Zahraa, M El Falaki, Improving the Recognition of Heart Murmur. *Int. J. Adv. Comput. Sci. Appl. (IJACSA)*. **7**, 1–10 (2016)
18. LG Durand, P Pibarot, Digital signal processing of the phonocardiogram: review of the most recent advancements. *Crit. Rev. Biomed. Eng.* **23**(3–4), 163–219 (1995)
19. RM Rangayyan, RJ Lehner, Phonocardiogram signal analysis: a review. *Crit. Rev. Biomed. Eng.* **15**(3), 211–369 (1987)
20. F Mezziani, SM Debbal, A Atbi, Analysis of phonocardiogram signals using wavelet transform. *J. Med. Eng. Technol.* **36**(6), 283–302 (2012)
21. T Chakrabarti, et al, in *2015 International conference and workshop on computing and communication (IEMCON)*. Phonocardiogram signal analysis - practices, trends and challenges: a critical review. (IEEE, New York, 2015)
22. NA Mohammed, ESA El-Dahshan, AS Yahia, A review of intelligent systems for heart sound. *J. Med. Eng. Technol.* **41**(7), 553–563 (2017)
23. eGeneral Medical Inc. USA. eGeneralMedical.com. <http://www.egeneralmedical.com/listohearmur.html>. Accessed 20 Apr 2018
24. P Bentley, et al, PASCAL Classifying heart sounds challenge (2011). <http://www.peterjbentley.com/heartchallenge/>. Accessed 20 Apr 2018
25. National Institute of General Medical Sciences and the National Institute of Biomedical Imaging and Bioengineering. PhysioNet/CinC challenge (2016). NIGMS & NIBIB. (2016). <https://www.physionet.org/physiobank/database/challenge/2016/>. Accessed 20 Apr 2018
26. Michigan Medicine. Heart Sound & Murmur Library University of Michigan Health System. Ann Arbor, MI. [http://www.med.umich.edu/lrc/psb\\_open/html/repo/primer\\_heartsound/primer\\_heartsound.html](http://www.med.umich.edu/lrc/psb_open/html/repo/primer_heartsound/primer_heartsound.html). Accessed 20 Apr 2018
27. S Davis, P Mermelstein, Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Trans. Acoust. Speech Signal Process.* **28**(4), 357–366 (1980)
28. LR Rabiner, RW Schafer, *Introduction to digital speech processing*. (Now Publishers Inc., 2007), pp. 70–73
29. H Naseri, MR Homaeinezhad, H Pourkhajeh, Noise/spike detection in phonocardiogram signal as a cyclic random process with non-stationary period interval. *Comput. Biol. Med.* **43**, 1205–1213 (2013)
30. M Singh, A Cheema, Heart sounds classification using feature extraction of phonocardiography signal. *Int. J. Comput. Appl.* **77**, 13–17 (2013)
31. H Liang, S Lukkarinen, I Hartimo, Heart sound segmentation algorithm based on heart sound envelopegram. *Comput. Cardiol.* **24**, 105–8 (1997)
32. C Potes, et al, Ensemble of feature-based and deep learning-based classifiers for detection of abnormal heart sounds. *Comput. Cardiol. Conf. (CinC)*, 621–624 (2016). <https://doi.org/10.23919/CIC.2016.7868819>
33. S Ari, K Hembram, G Saha, Detection of cardiac abnormality from PCG signal using LMS based least square SVM classifier. **37**, 8019–26 (2010)
34. H Ali, TJ Ahmad, K Shoab, Heart sound signal modeling and segmentation based on improved Shannon energy envelopegram using adaptive windows. *Asialink Int. Conf. Biomed. Eng. Tech.* (2007). Pearson Corp. London
35. G Redlarski, D Gradolewski, A Palkowski, A system for heart sounds classification. *PLoS ONE*. **9**(11), e112673 (2014). <https://doi.org/10.1371/journal.pone.0112673>
36. TE Chen, et al, S1 and S2 heart sound recognition using deep neural networks. *IEEE Trans. Biomed. Eng.* **64**, 372–80 (2017)
37. AP Yoganathan, et al, Use of the fast Fourier transform for frequency analysis of the first heart sound in normal man. *Med. Biol. Eng.* **14**, 69–73 (1976)
38. ND Marcelo, et al, *Classifying Heart Sounds Using Images of MFCC and Temporal Features*. (Springer, Cham, 2017), pp. 186–203
39. JIG Ortiz, CP Phoo, J Wiens, *Heart sound classification based on temporal alignment techniques*. (IEEE, New York, 2016), pp. 589–592
40. H Tang, H Chen, T Li, M Zhong, Classification of normal/abnormal heart sound recordings based on multi-domain features and back propagation neural network. *Comput. Cardiol. Conf. (CinC)*, 593–596 (2016). <https://doi.org/10.23919/CIC.2016.7868812>
41. J Rubin, R Abreu, A Ganguli, S Nelaturi, I Matei, K Sricharan, in *Computing in Cardiology Conference (CinC)*. Recognizing Abnormal Heart Sounds Using Deep Learning (Cornell University library, 2016), pp. 813–6
42. I Orovic, S Stankovic, Time-frequency-based speech regions characterization and eigenvalue Decomposition applied to speech watermarking. *EURASIP J. Adv. Signal Process.* 2012. **2010**, 572748 (2010)
43. YS Shin, J-J Jeon, Pseudo Wigner-Ville time-frequency distribution and its application to machinery condition monitoring. *Shock. Vib.* **1-1**, 65–76 (1993)
44. JD Martinez-Vargas, et al, Time-frequency based feature selection for discrimination of non-stationary biosignals. *EURASIP J. Adv. Signal Process.* **2012**(219), 2012 (2012)
45. X Zhang, et al, Time-frequency scaling transformation of the phonocardiogram based of the matching pursuit method. *IEEE Trans. Biomed. Eng.* **45**, 972–9 (1998)
46. D Boutana, M Benidir, B Barkat, Segmentation and identification of some pathological phonocardiogram signals using time- frequency analysis. *IET Signal Process.* **5**, 527–537 (2011)
47. SM Debbal, F Bereksi Reguig, Time-frequency analysis of the second cardiac sound in phonocardiogram signals. *Med. Phys.* **32**, 2911–7 (2011)
48. A Gavrovska, V Bogdanovic, I Reljin, B Reljin, Automatic heart sound detection in paediatric patients without electrocardiogram reference via pseudo-affine Wigner-Ville distribution and Haar wavelet lifting. *Comput. Meth. Prog. Biomed.* **113**, 515–2 (2014)
49. P Goncalves, RG Baraniuk, Pseudo affine Wigner distributions: definition and kernel formulation. *IEEE Trans. Signal Process.* **46**, 1505–1516 (1998)
50. A Sing, MK Dutta, CM Travieso, in *International Conference and Workshop on Bioinspired Intelligence (IWOB)*. Analysis of Heart Sound for Automated Diagnosis of Cardiac Disorders (IEEE, New York, 2017), pp. 1–5
51. A Barma, BW Chen, W Ji, S Rho, CH Chou, JF Wang, Detection of the third heart sound based on nonlinear signal decomposition and time-frequency localization. *IEEE Trans. Biomed. Eng.* **63**, 1718–27 (2016)
52. W Zhang, J Han, S Deng, Heart sound classification based on scaled spectrogram and least squares regression. *Biomed. Signal Process. Control.* **32**, 20–28 (2017)
53. W Zhang, J Han, S Deng, Heart sound classification based on scaled spectrogram and tensor decomposition. **84**, 220–231 (2017)
54. RC Gonzalez, RE Woods, *Digital image processing*, 3rd edn. (Pearson Corp., London, 2007), p. 508
55. D Song, L Jia, Y Lu, L Tao, in *International Conference on Systems and Informatics (ICSAI2012)*. Heart sounds monitor and analysis in noisy environments (IEEE, New York, 2012), pp. 1677–1681
56. S Babaei, A Geranmayeh, Heart sound reproduction based on neural network classification of cardiac valve disorders using wavelet transforms of PCG signals. *Comput. Biol. Med.* **39**, 8–15 (2009)
57. H Liang, I Hartimo, A heart sound feature extraction algorithm based on wavelet decomposition and reconstruction. *Proc. 20th Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.* **20**, 1539–42 (1998)
58. H Liang, I Hartimo, *A feature extraction algorithm based on wavelet packet decomposition for heart sound signals*. (IEEE, New York, 1998), pp. 93–96
59. Tu Zhihai, et al, in *Sixth International Conference on Natural Computation (ICNC 2010)*. Improved methods for detecting main components of heart sounds (IEEE, New York, 2010), pp. 3585–88
60. CN Gupta, et al, Neural network classification of homomorphic segmented heart sounds. *Appl. Soft Comput.* **7**, 286–97 (2007)
61. H Naseri, MR Homaeinezhad, Detection and boundary identification of phonocardiogram sounds using an expert frequency-energy based metric. *Ann. Biomed. Eng.* **41**, 279–92 (2012)
62. J Pedrosa, A Castro, TTV Vinhoza, in *36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Automatic heart sound segmentation and murmur detection in paediatric phonocardiograms (IEEE, New York, 2014), pp. 2294–7
63. X Zheng, M Dong, B Fu, X Fei, Automatic pathological signal detection of heart sound based on multi-resolution analysis. *Math. Comput. Contemp. Sci.* (2013)
64. N Marques, et al, Exploring the stationary wavelet transform detail coefficients for detection and identification of the S1 and S2 heart sounds. *Comp. Cardiol.* **2013**, 891–894 (2013)
65. Y Deng, PJ Bentley, A robust heart sound segmentation and classification algorithm using wavelet decomposition and spectrogram (2012)
66. CN Gupta, et al, Segmentation and classification of heart sounds. *Can. Conf. Electr. Comput. Eng.* 1674–1677 (2005). <http://dx.doi.org/10.1109/CCECE.2005.1557305>

67. M Abo-Zahhad Mohammed, et al, A comparative approach between cepstral features for human authentication using heart sounds. *Signal Image Video Process.* **10**, 843–851 (2016)
68. P Kumar Jain, A Kumar Tiwari, *An adaptive method for shrinking of wavelet coefficients for phonocardiogram denoising.* (Elsevier, New York, 2016), pp. 1–5
69. P Kumar Jain, A Kumar Tiwari, An adaptive thresholding method for the wavelet based denoising of phonocardiogram signal. *Biomed. Signal Process. Control.* **38**, 388–399 (2017)
70. MÅ Goda, P Hajas, Morphological determination of pathological PCG signals by time and frequency domain analysis. *Comput. Cardiol. Conf. (CinC)*, 1133–1136 (2016). <https://doi.org/10.23919/CIC.2016.7868947>
71. M Abdollahpur, et al, Cycle selection and neuro-voting system for classifying heart sound recordings. *Comput. Cardiol.* **43**, 176–238 (2016)
72. M Boussaa, et al, in *11th International Conference on Intelligent Systems: Theories and Applications (SITA)*. Comparison of MFCC and DWT features extractors applied to PCG classification (IEEE, New York, 2016), pp. 1–5
73. E Kay, A Agarwal, DropConnected neural networks trained on time-frequency and inter-beat features for classifying heart sounds. *Inst. Phys. Eng. Med. Physiol. Meas.* **38**, 1645–57 (2017)
74. F Sattar, et al, Time-scale-based segmentation for degraded PCG Signals Using NMF. *Non-negative Matrix Factorization Tech.*, 79–194 (2016). Springer
75. A Ramovic, et al, Wavelet and Teager Energy Operator (TEO) for Heart Sound Processing and Identification. *IFMBE Proc.* **1**(1–2), 70–73 (2017)
76. K Elisavet, et al, Fetal heart sounds detection using wavelet transform and fractal dimension. *Front. Bioeng. Biotechnol.* **5** (2017). <https://doi.org/10.3389/fbioe.2017.00049>
77. M Nabih-Ali, et al, Denoising of Heart Sound Signals Using Discrete Wavelet Transform. *Circ. Syst.* **8**(2017), 184–190 (2017)
78. NE Huang, et al, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *R. Soc.* **454**, 903–995 (1998)
79. A Zeiler, et al, in *The 2010 International Joint Conference on Neural Networks (IJCNN)*. Empirical mode decomposition - an introduction. (IEEE, New York, 2010), pp. 1–8
80. C Park, et al, Classification of motor imagery BCI using multivariate empirical mode decomposition. *IEEE Trans. Neural Syst. Rehabil. Eng.* **21**, 10–22 (2013)
81. N Rehman, DP Mandic, Multivariate empirical mode decomposition. *Proc. R. Soc. Lond. A Math. Phys. Eng. Sci.* **466**, 1291–302 (2010)
82. T Tanaka, DP Mandic, Complex empirical mode decomposition. *IEEE Signal Proc. Lett.* **14**, 101–4 (2007)
83. S Amy, et al, in *4th International Conference on Instrumentation, Communications, Information Technology and Biomedical Engineering (ICICI-BME)*. Automatic segmentation and detection of heart sound components S1, S2, S3 and S4 (IEEE, New York, 2015)
84. Z Zhao, *Noninvasive diagnosis of coronary artery disease based on instantaneous frequency of diastolic murmurs and SVM, vol 2005.* (IEEE, New York, 2005)
85. Z Zhao, Q Shen, F Ren, Heart sound biometric system based on marginal spectrum analysis. *Sensors (Basel)*. **13**(2), 2530–51 (2013)
86. L Sun, et al, in *7th International Conference on Signal Processing*. A novel method of time-frequency Representation and its application to biomedical signal processing, vol. 1 (IEEE, New York, 2004), pp. 236–9
87. L Sun, et al, in *27th Annual International Conference of the Engineering in Medicine and Biology Society*. Instantaneous frequency estimate of nonstationary phonocardiograph signals using Hilbert spectrum (IEEE, New York, 2005), p. 2005
88. A Gavrovska, M Slavkovic, I Reljin, B Reljin, Application of wavelet and EMD-based denoising to phonocardiograms. *Int. Symp. Signals Circ. Syst. ISSCS2013*, 1–4 (2013)
89. A Moukadem, A Dieterlen, C Brand, Study of two feature extraction methods to distinguish between the first and the second heart sounds. *Int. Conf. Bio-inspible Syst. Signal Process. BIOSIGNALS.* **2012**, 346–50 (2012)
90. H Sun, W Chen, J Gong, *An improved empirical mode decomposition wavelet algorithm for phonocardiogram signal denoising and its application in the first and second heart sound extraction.* pp. 187–191
91. DBB Barkat, M Benidir, Segmentation of pathological heart sound signal using empirical mode decomposition. *Int. J. Comput. Electr. Eng.*, 26–29 (2013). <https://doi.org/10.7763/IJCEE.2013.V5.655>
92. CD Papadaniil, LJ Hadjileontiadis, Efficient heart sound segmentation and extraction using ensemble empirical mode decomposition and kurtosis features. *IEEE J. Biomed. Health Inf.* **18**, 1138–52 (2014)
93. S Christophe, et al, in *4th International Symposium on Applied Sciences in Biomedical and Communication Technologies*. On the separation of heart sound components using a translated empirical mode decomposition, vol. 4. ISABEL (ACM, New York, 2011), pp. 1–5
94. JA Jimenez, MA Becerra, E Delgado-Trejos, Heart murmur detection using ensemble empirical mode decomposition and derivations of the Mel-frequency cepstral coefficients on 4-area phonocardiographic signals. *Comput. Cardiol.* 493–496 (2014)
95. S Banerjee, M Mishra, A Mukherjee, Segmentation and detection of first and second heart sounds (S1 and S2) using variational mode decomposition, 565–70 (2016). <https://doi.org/10.1109/IECBES.2016.7843513>
96. K Dragomiretskiy, D Zosso, Variational mode decomposition. *IEEE Trans. Signal Process.* **62–63**, 531–544 (2014)
97. AH Salman, et al, Empirical mode decomposition (EMD) Based Denoising Method for Heart Sound Signal and Its Performance Analysis. *Int. J. Electr. Comput. Eng. (IJECE)*. **6**, 2197–204 (2016)
98. J Jusak, I Puspasari, P Susanto, *Heart murmurs extraction using the complete ensemble empirical mode decomposition and the Pearson distance metric.* (IEEE, New York, 2016), pp. 140–145
99. A Nima, A Oveis, Short term wind power prediction based on improved kriging interpolation, empirical mode decomposition and closed-loop forecasting engine. *Sustainability*. **9**(11), 2071–1050 (2017). <http://www.mdpi.com/2071-1050/9/11/2104>. Accessed 20 Apr 2018
100. V Nivitha Varghees, KI Ramachandran, A novel heart sound activity detection framework for automated heart sound analysis. *Biomed Signal Proc. Control.* **13**, 174–188 (2014)
101. AH Salman, et al, in *2015 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*. Performance comparison of denoising methods for heart sound signal (IEEE, New York, 2015)
102. J Gong, S Nie, Y Wang, *An improved empirical mode decomposition algorithm for phonocardiogram signal de-noising and its application in S1/S2 extraction*, vol. 32. (IEEE, New York, 2015), pp. 970–4
103. GB Moody, RG Mark, The impact of the MIT-BIH arrhythmia database. *IEEE Eng. Med. Biol.* **20**, 45–50 (2001)
104. SE Schmidt, C Holst-Hansen, C Graff, E Toft, JJ Struijk, Segmentation of heart sound recordings by a duration-dependent hidden Markov model. *Physiol. Meas.* **31**(4), 513 (2010)
105. University of Toronto BioSec Lab. Biosec. PCG database. Lab BioSec. <http://www.comm.utoronto.ca/biometrics/databases.html>. Accessed 20 Apr 2018

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](http://springeropen.com)



## Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).

Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.

These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.

We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.

While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval, sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.

These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.

To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.

Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.

If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

[onlineservice@springernature.com](mailto:onlineservice@springernature.com)