The major challenge in phonocardiography and heart sound analysis is the involvement of environments with adverse effects such as noise, distortion and multiple sources. Methods existing in literature which address the above said problem have their own limitations and lacunae. The objective of this research work is to investigate and develop noise estimation, noise reduction, segmentation and classification algorithms for Phonocardiograms, so as to improve the performance over the existing heart sound processing methods in literature.

Vishwanath Madhava Shervegar received the B.E. degree in Electronics and Communication Engineering from Manipal Academy of Higher Education, Manipal, India, the M.Tech. from Manipal Academy of Higher Education, Manipal, India and Ph.D. degree in Electrical Engineering from Visvesvaraya Technological University, Belgaum, India.

9 7 8 6 2 0 2 7 8 7 1 8 5

Shervegar

Vishwanath Madhava Shervegar

Heart Sound Analysis and Classification

Signal Processing Methods for Heart Sound Analysis

LAP LAMBERT Academic Publishing

Vishwanath Madhava Shervegar

Heart Sound Analysis and Classification

Vishwanath Madhava Shervegar

Heart Sound Analysis and Classification

Signal Processing Methods for Heart Sound Analysis

LAP LAMBERT Academic Publishing

Imprint

Cover image: www.ingimage.com

Publisher:
LAP LAMBERT Academic Publishing
is a trademark of
International Book Market Service Ltd., member of OmniScriptum Publishing Group
17 Meldrum Street, Beau Bassin 71504, Mauritius
Printed at: see last page
ISBN: 978-620-2-78718-5

Zugl. / Approved by: New Signal Processing Methods for Heart Sound Analysis

Contents

Acknowledgement

I would like to express my sincere gratitude to Dr. Ganesh V Bhat for his guidance, suggestions and encouragement during my research period. I sincerely thank him for the excellent research environment he has created to learn and pursue research work.

It is a great pleasure to express my thanks to faculty and staff members at the Canara Engineering College Mangalore for their help and cooperation during my research. I would like to express my thankfulness to the Principal, staff and Management of Moodlakatte Institute of Technology, Kundapura for providing me an opportunity to carry out my PhD on a part time basis.

Last but not the least; I would like to express my deep gratitude to my friends and colleagues, who have always been there behind me with their love, support, encouragement and prayers.

List of Figures

List of Tables

List of Symbols

$\overline{\omega_{j,k}}$	Soft threshold Coefficient
$\omega_{j,k,}$	Signal Coefficient
λ	Threshold Parameter
σ_n	Noise standard variance
N	Length of the PCG Signal
th	Threshold for log-likelihood
w_dur	Hamming Window duration
h_dur	Hop duration
num_noise	Number of initial noise frames
N_FFT	Number of FFT points
fs	Sampling Frequency
a00,a01,a10,a11	Markov Parameters for Hangover Scheme
α	Coefficient of Decision Directed SNR estimation
G_old	Hidden Markov Model based Heart Sound detection probability
γ	posteriori SNR
$\hat{\xi}$	priori SNR
$\hat{G}$	Nonlinear Gain function
I_0 , I_1	Zeroth and First order Bessel Function
L_{ML}	Log Maximum Likelihood
$\hat{\lambda}_d$	Instantaneous Noise Spectrum Estimate
$\bar{\lambda}_d$	Mean Noise Spectrum Estimate
S_{min}	Smoothed Spectrogram Estimate for 1st Iteration
$\tilde{S}$	Smoothed Spectrogram Estimate for 2nd Iteration
$S_{\min\_sw}$	First Smoothed Spectrogram running minimum
$\tilde{S}_{\min\_sw}$	Second Smoothed Spectrogram running

	minimum
$I(k,l)$	Indicator function
$\tilde{q}(k,l)$	a priori signal absence probability
$\tilde{p}(k,l)$	a priori signal presence probability
$\tilde{\alpha}_d$	time varying frequency dependent parameter
β	Fixed Bias
M_λ	3×5 λ $Matrix$
a_i	Attenuation Coefficient
$\hat{R}_k$	Risk Estimate in a Block
$R(x)$	Penalty function
σ	Noise variance
$\boldsymbol{x_{norm}(t)}$	Normalized PCG Signal
$\boldsymbol{x(n)}$	Energy of PCG Signal
$\boldsymbol{a(n)}$	Slow varying part of PCG signal
$\boldsymbol{f(n)}$	Fast varying part of PCG signal
$\boldsymbol{L}$	Low pass Filter
$I_i(dB)$	Power Spectrum of the Signal
$z(f)$	Bark Function
$L_{dB}(t)$	Loudness Function
$P\ (x_{n\|}\lambda)$	Probability Density Function
μ	Mean
∑	Standard deviation
E	Feature Vectors
J	Objective function
c_j	Cluster centers

List of Abbreviations

PCG	Phonocardiogram
RBBB	Right Bundle Branch Block
PVC	Premature Ventricular Contraction
RBBB	Left Bundle Branch Block
AS	Aortic Stenosis
PS	Pulmonic Stenosis
HOCM	hypertrophic obstructive cardiomyopathy
MVP	mitral valve prolapse
MR	mitral regurgitation
TR	tricuspid regurgitation
VSD	ventricular septal defects
PDA	patent ductus arteriosus
AVF	arterio-venous fistulae
CCCT	Cardiac Contractility Change Trend
HR	heart rate
CVD	Cardiovascular disease
ALE	Adaptive Line Enhancement Filter
AWGN	Additive White Gaussian Noise
ML	Machine Learning
HMM	Hidden Markov Model
CSCW	cardiac sound characteristic waveform
IMF	intrinsic mode functions
MLPNN	multi-layer perceptron neural network
CAD	Computer Aided Diagnosis
LSSVM	Least Squared Support Vector Machines
GTSVM	Growing Time Support Vector Machines
MFCC	Mel Scaled Frequency Cepstral Coefficients
STFT	Short Time Fourier Transform
DWT	Discrete Wavelet Transform
kNN	k-Nearest Neighbour
TFBT	Time Frequency Block Threshold method

OGS	Overlapping Group Shrinkage method
ST	Soft Threshold method
SNR	Signal to Noise Ratio
AD	Activity detection
LRT	likelihood ratio test
DD	Decision Directed method
HS	Heart Sound
IMCRA	Improved Minimum Controlled Recursive Averaging
LSA	Log Spectral Amplitude
SSNR	Segmental Signal to Noise Ratio
MSE	Mean Squared Error
CWT	Continuous Wavelet Transform
LPF	Low Pass Filter
BPF	Band Pass Filter
Se	Sensitivity
PP	Positive Predictive Value
Acc	Accuracy
TN	True Negative
FN	False Negative
FP	False Positive
TP	True Positive
ESS	Event Synchronous Segmentation
FCM	Fuzzy C Means
GMM	Gaussian Mixture Model
PDF	Probability Density Function
EM	Expectation Maximization

ABSTRACT

The major challenge in phonocardiography and heart sound analysis is the involvement of environments with adverse effects such as noise, distortion and multiple sources. Methods existing in literature which address the above said problem have their own limitations and lacunae.

The objective of this research work is to investigate and develop noise estimation, noise reduction, segmentation and classification algorithms for Phonocardiograms, so as to improve the performance over the existing heart sound processing methods in literature.

In this work the phonocardiogram available for research from MIT Heart sound database Physionet database(MITHSDB). The MITHSDB is hosted in the Physionet repository. MITHSDB is the largest open access heart sound database available till date with over 3000 heart sound recordings. The sounds in the database are corrupted with noises of various types and sources.

As the type of noise in MITHSDB is unknown, in this work the noise present is assumed as stationary/non stationary or both. Under this assumption, it is attempted to de-noise these sounds using time frequency block threshold method. Noise estimation which forms an integral part of block threshold is done using activity detection and improved minimum controlled recursive averaging.

The de-noised heart sounds are normalized and then processed for automatic segmentation event synchronous method. The event synchronous method detects the occurrence of cardiac events-the first heart sound (S1) and the second heart sound (S2) using the loudness index.

The use of loudness features obtained using loudness index is also investigated for classification of heart sounds. The loudness feature is extracted using the spectrogram of event synchronous segmentation method, this feature is then classified using Gaussian Mixture Model (GMM). The GMM utilizes clustering of sounds to classify the sounds as normal and abnormal.

The methods used in this work show good results in terms of performance metrics when compared to the current methods present in literature.

Chapter 1

Introduction

1.1 The Definition and need for Phonocardiography

A Phonocardiogram (PCG) is an audio-visual display of acoustic vibrations made in the form of cardiac sounds and murmurs by the heart. These are often obtained using an instrument called as Phonocardiograph. The study of such various types of medical recordings made during cardiac auscultation is often termed as Phonocardiography.

Heart sound analysis is a detailed step by step procedure adopted during cardiac auscultation particularly aimed at careful diagnosis of cardiac diseases, long term monitoring of patient health, as a biometric tool for patient authentication and for teaching, education and research purpose.

Phonocardiograms reveal critical health related statistics that correspond to the proper functioning of the cardiac valves. Normal heart sounds have two important and fundamental components the S1 (lub) and S2 (dub) sounds. Presence of additional sounds generally indicate abnormality. A PCG with third sound could signal critical heart failure, whereas as presence of pathological murmurs might indicate the defective valves or orifice in the septal wall.

Phonocardiogram acquisition instruments are used for long time monitoring of patients suffering from various cardiac disorders. Real time Continuous monitoring helps the doctors to study the disease on a case to case basis and take necessary and corrective action. These Phonocardiogram instruments are now equipped with both wired and/or wireless technologies that make cardiac auscultation easy and user friendly. The new systems support telemetry thus enabling the doctors to serve the patients at remote locations where access to medical facilities is quite difficult.

1.2 Overview of Heart Sound Signals

The PCG signals are non-linear and non-stationary sound vibrations picked from the chest. Sounds emanate from heart as a result of the rapid inward contraction and outward relaxation of the cardiac muscles. An acoustic stethoscope helps the doctor to identify four classes of sounds during cardiac auscultation: (i) The lub-dub heart sounds namely the S1 and the S2, (ii) The diastolic sounds namely the S3 and the S4, (iii) The murmurs and (iv) The high pitched sounds (clicks and snaps). Heart sounds show varying levels of intensity, frequency, duration,

and quality [13]- [18]. Fig. 1.2.1 shows a pictorial representation of the PCG signal with its important heart sounds.

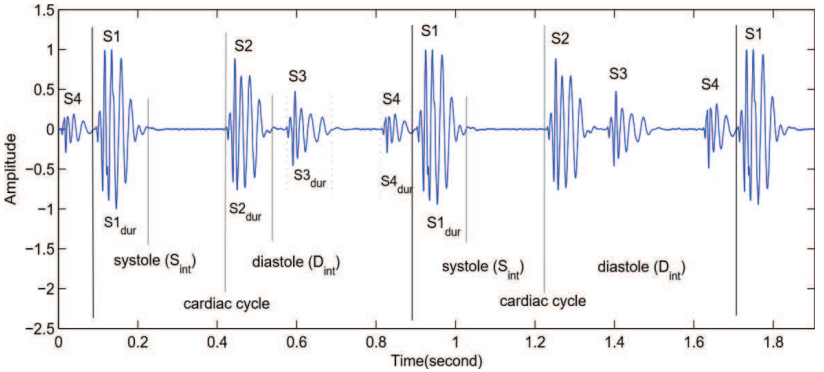

Figure 1.2.1 Time Domain Plot of PCG signal with its different components.

First Heart Sound (S1): The first component of the PCG signal (S1) starts at the onset of the systole. The mitral component (M1) begins at a time slot 20-30ms before the tricuspid component (T1) [15]. The S1 is a long duration low-pitch sound [14]. The intensity of the mitral component (M1) is much more than the tricuspid component (T1) intensity due to the sudden rise in left ventricular pressure. S1 is evaluated by noting its features such as Quality, intensity, and degree of splitting [13]- [18]. Myocardial depression, ventricular septal defect and acute aortic regurgitation are some of the diseases associated with decreased intensity of S1. Some pathologies like the Right bundle branch block (RBBB), or ventricular tachycardia or premature ventricular contraction (PVC) arises due to the increase in the splitting of S1 sound (more than 60ms) [14].

Second Heart Sound (S2): The other heart sound in the PCG signal (S2) begins at the end of the systole. Compared to S1 sound, S2 sound is usually short and is of slightly higher pitch. In the PCG signal the S2 lies in the frequency range 10-400 Hz and is of the duration 50-150ms [9]. The two S2 sounds due to aortic valve (A2) and the pulmonic valve (P2) are as short as 50ms [9]. The delay between the closure of the two valves associated with A2 and P2 causes a split S2 sound. S2 can be identified by its splitting and intensity. The amplitude and frequencies of aortic components (A2) is slightly more than the pulmonary components (P2) [3]. The splitting interval of second heart sound (S2) widens up on inspiration and narrows down on expiration. Vital parameters like aortic Blood pressure is determined by the splitting of the S2

[8]. There a number of diseases related to split S2 related pathology like Pulmonic stenosis, RBBB, left bundle branch block (LBBB), atrial septal defect and right ventricular failure. It is observed, in normal PCG signals, that the diastole is much longer than the systole [30]- [40].

Third Heart Sound (S3): The third component of the PCG signal occurs during diastolic period, usually 100-150ms after the S2. This happens when there is a sudden deceleration of blood flow within the ventricles. When compared to the S1 and S2, the S3 is of low amplitude and low frequency. The PCG signal component S3 lies in the range of 30 Hz-90 Hz and lasts for a duration of 70 ± 15ms [43], [44]. S3 is often seen in patients with acute impaired myocardial reserve [16]. The S3 provides clinical information about hemodynamic and systolic dysfunction, and is useful for diagnosis of heart failure [41]- [48]. The presence of S3 in PCG of adults has been correlated with heart failure.

Fourth Heart Sound (S4): The S4 occurs when there is contraction of atria. As a result, the blood may be forced into the distended ventricles. The S4 occurs before the first heart sound. The S4 sound signifies low-frequency vibrations with frequency band of 20-30 Hz. These are seen in the PCG of patients with diminished left ventricular compliance [13]- [18].

1.3 Heart Murmurs and Pathological Sounds

Murmurs are caused when there is turbulence in the flow of blood or vibration of the tissues. In pathological cases, different types of murmurs occur due to dysfunctions in the valves. Murmurs may be systolic, diastolic or continuous in nature. The heart sound can be adjudged using the various types of parameters such as:

1. The PCG timing parameters (early, mid, late, or pan)
2. Intensity, duration, pitch (low, medium, or high) parameters of the PCG
3. Quality parameters of the PCG (musical, blowing, harsh or rumbling)
4. Shape parameters of the PCG such as Crescendo, Decrescendo and Crescendo-decrescendo [6], [14] - [18].

The cardiac murmurs are synonymous with the pitch and intensity related to the velocity of blood. Accurate diagnosis is dependent on the timing of the murmur.

Systolic murmurs: Pathology associated with systolic murmurs are as follows:

1. Acute Mitral Regurgitation (MR) and Tricuspid Regurgitation (TR) are systolic murmur.
2. Aortic stenosis (AS), Pulmonic stenosis (PS), Hypertrophic Obstructive Cardio-Myopathy (HOCM) and Atrial Septal Defect (ASD) are Mid-Systolic murmurs.
3. Mitral Valve Prolapse (MVP) is a late-systolic murmur.
4. Mitral Regurgitation (MR), Tricuspid Regurgitation (TR), and Ventricular Septal Defects (VSD) are holosystolic or pansystolic murmurs.

Diastolic murmurs: The various pathologies associated with diastolic murmurs are as follows:

1. Aortic and Pulmonic Regurgitation are the early diastolic murmurs.
2. Mitral or Tricuspid Stenosis is a mid-late diastolic murmur.

Continuous murmurs: Patent Ductus Arteriosus (PDA) and systemic Arterio Venous Fistulae (AVF) are examples of continuous murmurs [11]- [17], [49]- [68].

Various types of heart murmurs in the PCG signal are shown in Fig 1.3.1. Defects of the semilunar valves and the mitral and tricuspid valves also results in clicks and snaps. The clicks and snaps indicate distinctive features of some type of heart defects.

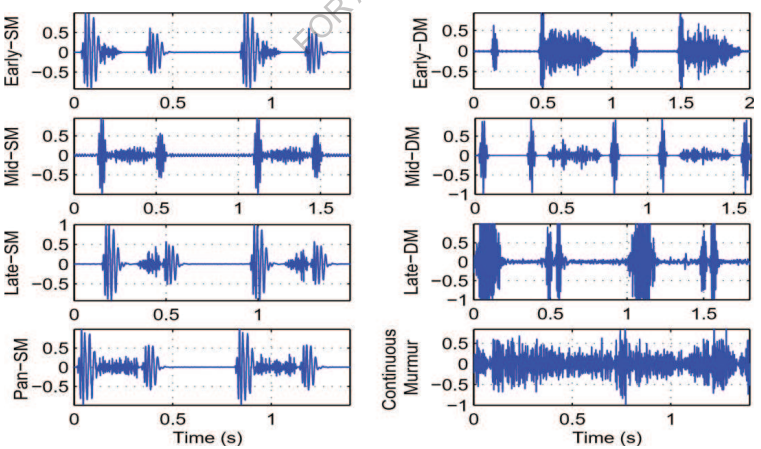

Figure 1.3.1: Various types of Murmurs that occur in PCG signal

1.4 PCG Parameter for Cardiovascular disease diagnosis

In clinical studies, specific heart sound indices are measured for evaluating heart functions of subjects, maternal, fetal and infants with various physiological and pathological conditions. The various parameters used for CVD detection are listed below.

1. Cardiac Contractility Change Trend (CCCT): Ratio of the S1 amplitude with exercising to the S1 amplitude without exercising [10].
2. Amplitude of S1 [29].
3. Ratio of S1 sound amplitude to S2 sound amplitude [10], [28].
4. Ratio of the amplitude of T1 to the amplitude of the M1.
5. Ratio of S3 amplitude to S2 amplitude.
6. Ratio of the diastole to the systole [25], [10], [27].
7. Location of S1 in the PCG [7].
8. Duration of the PCG and the Energy of Instantaneous Frequencies (EIFs).
9. A2-P2 split [1], [3].
10. Rate of heart beats (HR) [25], [27].
11. Duration and frequencies of S3 and S4 sounds; timing (location), configuration (shape), loudness (intensity), spectral content, duration of murmurs.

Modern digital electronic stethoscopes have the ability to carry out amplification, playback, display and recording of the heart sound signals in real time. But automatic and quantitative measurement of heart sound parameters helps in the diagnosis of cardiac pathology.

Chapter 2

A Review of Heart Sound Analysis

Global survey in 2012 revealed that Cardiac pathology resulted in the deaths of 17.5 million people worldwide. In low-income countries, there is shortage of high quality diagnostic equipment. In the richer economy countries, sophisticated technology have developed at a rapid pace. However, heart sound auscultation still remains an important emergency tool for the physicians and doctors. Also, access to expert diagnosis is restricted due to the high patient to doctor ratios (50,000:1) in low and mid economy countries. A solution to this problem lies in providing cure using wireless technology via the smartphone or cloud.

Methods have been developed to diagnose cardiac disorders by means of heart sound signal analysis that includes automated heart sound segmentation and classification techniques to be used in low cost clinical applications.

The literature survey also shows that the PCG provide vital information of the heart. Heart sound analysis can be broken down to three categories: i) preprocessing; ii) segmentation; and iii) classification. Preprocessing algorithms are used to identify noise and artefacts in the signal and remove them. Various machine learning algorithms have been devised to contain the redundant noise and obtain a clean signal. To screen the pathologies various segmentation and classification algorithms have been studied. The segmentation algorithms detect pathologies by identification of the cardiac events. For example, the presence of an additional S3 in adults could indicate heart failure. The classification algorithms are used to identify the pathology in the heart sounds. For example, the presence of murmurs in systole could indicate stenosis in the valves, while murmurs in the diastole could indicate regurgitation of blood in the valves.

2.1 Problem Identification

Heart sound analysis methods mentioned in Literature can be inefficient due to following reasons.

1) The method may use only a small amount of data.

2) There is severe lack of adequate separate test dataset for validation of the data.

3) The developed method may not use a different type of PCG recordings obtained under different types of environments.

4) The method may perform validation only on clean recordings.

In our work, we concentrate on more accurate techniques of cardiac sound analysis on normal

and abnormal heart sounds. That will include developed techniques for preprocessing, segmentation and classification under harsh and adverse environment. In our work we have made use of the Physionet Database (MITHSDB) for cardiac sound analysis. This database is the single largest public collection of PCG recordings/ heart sounds obtained worldwide.

The heart sound analysis can be broadly classified as i) Preprocessing ii) Segmentation iii) Classification. Preprocessing of heart sound involves identification of the type of noise underlying the signal and methods used to contain them by means of cancellation or averaging. Preprocessing is an important stage because the sound components responsible for false detection of diseases during pathology screening is virtually eliminated. The Segmentation of heart sounds follows the preprocessing stage. Segmentation of heart sounds helps us to identify the first heart sound (S1) and the second heart sound (S2). We can also identify the systole and the diastole. Segmentation gives clear idea about the start and stop instants of various heart sound components including murmurs. A single segmented PCG cycle can further be classified as normal or abnormal sound by means of the classification process. Classification is done by comparing the various features of heart sound with that of the features of the sounds in the reference database. Various supervised and unsupervised classification techniques are mentioned in Literature which have their own merits and demerits. Both methods require standard sets of features for detection of pathology in patients. Feature extraction stage is a prerequisite for accurate classification of PCG signals.

2.2 Literature Survey and Gaps present

In 1816 Laennec was the first to hear human Heart sounds. He did so by placing an end of the paper roll directly over the chest of the patient and the other end to his ears. This paved way for the phonocardiography and the actual analysis of heart sound. The initial work on heart sounds included recording of heart sound and processing of heart sound. This was followed by segmentation of heart sound for detecting the presence of S1, S2 sounds, systole-diastole duration, heart rate calculation etc. Then the last step included extra heart sound feature extraction and classification.

2.2.1 Pre-Processing Heart Sounds Using Noise Filtering.

a. Adaptive filters.

In 1996 three researchers Tinati, M.A., Bouzerdoum & A Mazumdar J [120] used LMS based

modified Adaptive Line Enhancement Filter (ALE) for containing the noise present in the Heart Sound Noise. In this method adaptive filters were used for filtering the heart sounds which consisted of a convolved noise component. The ALE method did not estimate any kind of noise but relied on the deconvolution of noise embedded in the actual signal. The deconvolution resulted in the noise free heart sound suitable for further processing. The main principle behind deconvolution in ALE is reduction of minimum mean square error to the lowest possible level by proper adjustments of the weights of the LMS filter, so as to obtain a very clean signal. The minimum mean square error is estimated taking the square root of the mean differences between the noise corrupted heart sound and the de-noised heart sound. Since then, LMS-ALE filters and their adaptations have been frequently used de-noise heart sounds. Many other works have also been reported in Literature regarding the use of the Adaptive filters for noise reduction in PCG signals [70-72]

b. Statistical filters.

To contain noise in the PCG signals the reduced-order Kalman filter was also used [121]. The first order method that uses Gauss–Markov process is utilized to convert the model into state-space. The optimal mean square of the heart sound was obtained by application of the Kalman filter. The major disadvantage of this method was, it used segmentation for identification of cardiac cycle. Identification of cardiac cycle by segmentation is quite inaccurate in the presence of high amount of background noise.

c. Time frequency techniques.

Spectral subtraction method [122] is a stationary method that is used to contain Additive White Gaussian Noise (AWGN) in PCG signals. This method failed in non-stationary noise conditions which was a major drawback in PCG de-noising since real time noise is not controlled and is usually non-stationary. In 2011, Ramos et al. [123] used modulation filtering method for heart sound noise reduction. Using a single channel of PCG signal they separated out the heart sounds and the redundant noise. The STFT was used to obtain the amplitude of the heart sound. Coherent demodulation was performed on the amplitude of the STFT to separate out the carriers and the modulators. The noises were filtered by using the filter-bank on the modulators. The filtered modulators were then combined with carriers to reconstruct the original signal. They obtained a high sensitivity of 93.6% and a specificity of 92.3% respectively.

d. Wavelet techniques.

The discrete wavelet transform is another popular method of de-noising PCG. Discrete wavelet transform has many advantages compared to its counterpart, the STFT. It uses a time varying window for de-noising and removes both the non-stationary and stationary noises with good accuracy. For heart sounds, Wavelet de-noising technique has been proposed by Gyanaparva et al. (2013) [73]. This method used a novel wavelet-based coefficient threshold technique for noise reduction. This method solves the problem of using discontinuous hard-threshold function and the problem of permanent bias in soft-threshold function by selecting a suitable adaptive threshold. This method reduced the noises effectively in fetal heart sounds [73].

e. Cycle-Frequency Domain Techniques

Cyclic stationarity is an important property of heart sounds. Based on Nonlinear Scaling in time domain, Tang et al. (2016) [92] proposed a method to reduce noise and disturbances in Heart Sounds. A piecewise linear function was used to estimate the nonlinear time scale. Clean heart sounds were thus obtained in the presence of zero mean additive noises and disturbances.

Later, Tang et al. (2016) [124] used a Fuzzy detection method for de-noising the PCG. They used PCG signal properties such as the time delay, the frequency, the amplitude, the time width, and the phase in this process. On subsequent analysis of the PCG signals it was observed that the cardiac components of the PCG clustered or congregated in the joint domains while noise did not cluster or congregate. Experiments conducted on PCG signals showed that heart sound signal components were made from lung sounds and chest motion artefacts. This was evident from the computer simulations that showed a fully reconstructed heart sound signal. The process however did not affect both the waveform and time delay of PCG signal. The similarities of the reconstructed and noise-free PCG signal were compared using correlation and normal residue parameters. Ten cardiac cycles were used for evaluation in 0dB noise environment. Using this data, they obtained a Correlation coefficient more than 0.90 and normalized residue up to 0.10.

2.2.2 Segmentation of PCG.

The PCG signal consists of two important sounds namely the first cardiac sound S1 and second cardiac sound S2 which are the first and second sounds respectively. Segmentation of the cardiac sounds help identify the systole or diastole. The closure of the AV valves triggers the

onset of the first cardiac sound S1. The closure of the Semilunar valves triggers the onset of the second heart sound.

The rule to find the start and stop of the timing of S1 and S2 is as follows.

1) The S1 begins when the mitral valves close.

2) The S2 begins when the aortic valves close.

3) The end of the S1 and S2 is triggered by some other high frequency vibrations (Moukadem et al. 2013) [51].

Methods of Cardiac segmentation are of four types. They are categorized as methods based on the shape of the envelope, type of the feature, type of Machine Learning (ML) technique and Hidden Markov Model (HMM) principle.

a. Envelope-based methods

Envelope based method uses PCG envelope extraction for segmentation of PCG signals. In a first, in 1997, Liang et al proposed an envelope based method using the normalized average Shannon energy and then he segmented the heart sounds using this envelope. This method gave emphasis to the medium-intensity heart sounds and reduced the low-intensity noise components. A total of 37 recordings contained 515 cycles. They were obtained from children with murmurs. With this method he obtained 93% accuracy. Later, Liang et al. (1997) [75] also employed wavelet decomposition for heart sound segmentation. The heart sounds were subdivided into four partitions namely S1, systole, S2 and diastole. He used a dataset of 77 noisy PCG recordings that had in it 1165 cardiac cycles of heart sounds. Using his method, he obtained a sufficient accuracy of 84% (without using the wavelet decomposition) and 93% (with using the wavelet decomposition).

Later in 2013, Moukadem et al. [51] obtained the of the S-transform of the heart sound signal and then estimated its Shannon energy envelope. 40 normal cardiac sounds and 40 pathological cardiac sound were used and thus obtained a sensitivity and positive predictive value of higher than 95%.

Sun et al. 2014 [76] used Hilbert transform for heart sound segmentation. This method was divided into two parts. As a first part, he decimated the real part of the complex analytic signal

and obtained its envelope. As a second part he calculated the derivative of the imaginary part of complex analytic signal and obtained the Instantaneous frequency. For segmenting the PCG they considered the peaks of S1, the peaks of S2, the transmission points T12 from S1 to S2, and the transmission points T21 from S2 to S1. They used 7730s recordings from abnormal patients, 600s recordings from normal subjects, and 1496.8s recordings from Michigan MHSDB database. S1 and S2 without separation gave an average accuracy of 96.69%. and for sounds with separation gave an average accuracy of 97.37%.

In 2006 Jiang et.al [77] developed a new waveform based method of PCG segmentation called as the Cardiac Sound Characteristic Waveform (CSCW) method. This method was tested against Shannon energy method and Hilbert transform method. It was found that the CSCW method was more superior to both of these methods. CSCW method gave a considerably high accuracy of 100% and sensitivity of 88.2% respectively. While Shannon energy method gave an accuracy of 78.2% and sensitivity of 89.4% and Hilbert transform method gave an accuracy of 51.4% and sensitivity of 47.3%. However, the three methods were tested only on 500 selected cardiac cycles.

In 2010 Yan et al. [78] used characteristic moment waveform envelope method for PCG segmentation. With a small dataset of only 9 heart sound recordings they obtained a very high accuracy of over 99.0%.

Ari et al. 2008 [79] proposed a simple squared-energy envelope extraction method. To evaluate the heart sounds, frequency, energy and timing of PCG signal components were used. They employed a threshold-based detection method for detecting heart sound components. The method outperformed Shannon energy envelope method in terms of performance. From 71 recordings on a total of 357 cycles they obtained an accuracy of 97.47%.

b. Feature-based methods

In 2013, Naseri and Homaeinezhad, [80] developed a PCG segmentation algorithm that used synthetic decision making method. In this method they used frequency and amplitude based features of 52 PCG signals taken from patients suffering from different valve diseases. An average of 99.00% sensitivity and 98.60% positive predictive value was obtained.

In 2006, Kumar et al. [81] developed a PCG segmentation method that used features from high frequencies of PCG. They obtained fast wavelet decomposition based features from the PCG

signal and then tested them on patients with native and prosthetic heart valves. These features showed profound pressure differences in both cases. The method was later tested on two classes of patients namely one with mechanical and bio-prosthetic heart valve implants and other, the patients with native valves. Thus an average of 97.95% sensitivity and of 98.20% positive predictive value was found with this method.

In 2014, Varghees et.al, [82] developed a PCG segmentation algorithm using an instantaneous phase feature based method. They calculated the Shannon entropy and used analytic features for segmenting the PCG. No look-back steps were used. The method was tested on 701 cycles of clean and noisy PCG signals. A high sensitivity of over 99.43% and positive predictive value of over 93.56% was obtained during experiments.

In 2014 Pedrosa et al. [83] developed a PCG segmentation algorithm in which they used the analysis signal of the autocorrelation function. These periodic features were later used to segment the PCG. The method was tested on a dataset of 72 PCG recordings. 89.2% sensitivity and 98.6% positive predictive value was obtained.

Nigam and Priemer, 2005 [84] used complexity feature of the heart sound dynamics to perform PCG. Good performance was obtained on the synthetic data. They used amplitude and frequency based features for segmenting the PCG.

Vepa et al. 2008 [85] developed a PCG segmentation algorithm by using the complexity-based features. These features were obtained by combining the energy-based and simplicity-based features of the PCG signal. They were obtained from the multi-level wavelet decomposition coefficients extracted from the PCG signal. On experimental evaluation on a dataset consisting of 166 cardiac cycles they achieved an accuracy of 84.0%.

Papadaniil et.al 2014 [86] segmented the PCG signals by extracting the kurtosis-based features from the heart sounds. These features combined with different non-Gaussian intrinsic mode functions extracted from ensemble empirical mode decomposition were used for cardiac segmentation. The start positions and end positions of heart sounds were marked using the intrinsic mode functions. They obtained an accuracy of 83.05% for 11 normal patients and 32 abnormal patients.

Gharehbaghi et al. 2011 [87] introduced an ECG based cardiac segmentation method for pediatric patients. The dataset of 120 sounds of normal and pathological children, with 1976

cardiac cycles were used. They obtained an accuracy of 97% for S1 and 94% for S2.

c. Machine Learning based methods

In 2002, Oskiper and Watrous, 2002 [88] introduced a method of PCG segmentation for S1 detection by using Time Delay Neural Network. Time-delay links were used to connect to the time-frequency Morlet wavelet decomposition energy coefficients. On 30 normal subjects they obtained an accuracy as high as 96.2%.

Sepehri et al. 2010 [89], devised a PCG segmentation method by using a neural network classifier called as Multi-Layer Perceptron Neural Network (MLPNN). With 40 normal and 80 abnormal recordings of children they obtained an accuracy of 93.6%.

Chen et al., 2009 [90], developed a PCG segmentation algorithm to identify the various heart sound components. They clustered the sounds by employing K-means and segmented the PCG using a threshold based method. A high 92.1% sensitivity and a 88.4% positive predictive value was obtained for 27 recordings of healthy subjects.

Gupta et al. 2007 [91], extracted a single cardiac cycle from the PCG signal. Then they clustered the PCG by K-means and segmented using Homomorphic Filtering (HF). For a total of 340 cycles an accuracy of 90.29%. was obtained.

Tang et al. 2010, [92], segmented the PCG signal by using dynamic clustering. The PCG signal was first decomposed into cardiac cycles by frequency. Each cardiac cycle was clustered and grouped into two classes either a S1 or a S2. On 25 subjects, 94.9% accuracy was obtained for S1 and 95.9% for S2.

Rajan et al. 2006 [93], introduced SVD and Morlet wavelet decomposition based PCG segmentation. A singular value decomposition technique (SVD) was used on Morlet features to identify heart sound segments. On 42 adult patients 90.5% accuracy was obtained.

d. HMM based methods

In 2003, Gamero et.al developed a PCG segmentation algorithm by using a probabilistic finite state-machine model and applied it to PCG [94]. They used a network of two HMMs with grammar constraints on the sequence of systole and diastole. Experiments were evaluated for over 80 subjects. Promising results with 95% sensitivity and 97%. positive predictivity was

obtained.

In 2005, Ricke et al. [95], developed a PCG segmentation algorithm based on HMM method. They achieved a high 98% accuracy with eight-fold cross-validation but the method was tested only on 9 subjects.

Gill et al. 2005 [96], developed a PCG segmentation algorithm in which he incorporated the timing durations within the HMM. Homomorphic filtering followed by feature extraction was performed. The extracted features were used with GMM. The method was validated on 44 PCG recordings obtained from over 17 subjects. S1 sounds showed a high sensitivity of 98.6% and high positive predictive values of 96.9%, while S2 sounds showed high sensitivity of 98.3% and high positive predictive values 96.5%.

In 2014, Sedighian et al. [97] introduced a method to segment PCG by using Homomorphic filtering and an HMM. They used PASCAL database for testing. S1 sounds showed a high accuracy of 92.4% and while S2 sounds showed an accuracy of 93.5%.

Castro et al. 2013 [98], developed a PCG segmentation algorithm by using the wavelet analysis on the PASCAL database. S1 sounds showed a high accuracy of 90.9% and while S2 sounds showed an accuracy of 93.3%.

In 2010 Schmidt et. al [99] developed a PCG segmentation algorithm by using the hidden semi-Markov model (HSMM). The S1 and S2 sounds in 113 recordings, were hand labelled first. Gaussian distributions were derived from the average duration of the sounds and autocorrelation analysis of systole and diastole. Later, by using the homomorphic envelope and three frequency band features of 25–50, 50–100 and 100–150 Hz the Gaussian distribution-based emission probabilities for the HMM were found. The frequency features were used with Viterbi algorithm to label the states. Using this method, they obtained 98.8% sensitivity and 98.6% positive predictive value.

In 2016 Springer et al. [100] developed a PCG segmentation algorithm by using HSMM method and logistic regressions and applied it to the noisy, real-world heart sound recordings. They also developed a modified Viterbi algorithm in order to identify the sequence of states of the sounds. 10 172s recordings from 112 patients were used for testing. An average F1 score of 95.63% was obtained which was significantly higher compared to the highest score of 86.28% reported from other methods.

2.2.3. PCG feature extraction and Classification.

The heart sounds can be classified based on pathology. The normal heart sound is free of extra sounds, but abnormal heart sounds contain extra sounds such as murmurs and clicks. From the past 50 years, several developments have taken place in the automated classification of heart sounds for pathology. As early as 1963, in a first, Gerbarg et al. [101] worked on the techniques for automated heart sound classification. The work was reported on children suffering from Rheumatic Heart Disease (RHD) by using a threshold based method. The PCG classification can be grouped into four categories namely Artificial Neural Network (ANN), Support Vector Machine (SVM), Hidden Markov Model (HMM) and clustering based approach.

a. Classification based on ANN

The ANN classifiers use features of the PCG signal as inputs. Some of them are wavelet, time, frequency and complexity and time-frequency features. The most popular features used with ANN are the Wavelet-based features.

Akay et al.1994 [102], performed automated analysis of heart sounds by using wavelet based features alongside an ANN. From the diastole they computed four features namely mean, variance, skewness and kurtosis using the coefficients of wavelet. Along with these features, they fed the physical characteristics such as sex, age, weight, blood pressure into a fuzzy neural network. On 82 recordings a sensitivity of 85% and a specificity of 89% was reported.

In 1998 Liang et.al [103], developed a decomposition method based on wavelet packet for heart sound classification. They used this method to differentiate the pathological murmurs from the innocent murmurs in children. Using the information-based cost function eight nodes of the wavelet packet tree were selected. The values of the cost function were used as the feature vector. They achieved 80% sensitivity and 90% specificity on the dataset with a 65/20 train/test split.

In 2012, Uguz [104], used discrete wavelet transform features and a fuzzy logic approach along with ANN. He classified the PCG into three classes namely normal, pulmonary stenosis, and mitral stenosis. For 120 subjects a 50/50 train/test split was used. With this amount of data, he achieved 100% sensitivity, 95.24% specificity, and 98.33%.

In 2005, Bhatikar et al. [105], devised a classification algorithm for the PCG by estimating the

features of energy spectrum obtained from FFT. The features were fed as inputs to an ANN. To evaluate the performance, experiments were performed on 53 patients using a separate test set consisting of innocent and pathological murmurs. With all of these data, they reported 83% sensitivity and 90% specificity.

Sepehri et al. 2008 [106] classified the PCG by using ANN to identified the five frequency bands. The frequency bands with the greatest difference in spectral energy was used to identify normal and pathological heart sounds. 95% sensitivity and 93.33% specificity were obtained for 50 sounds.

Ahlstrom et al. 2006 [107], developed a classification algorithm for the murmurs by using the non-linear complexity-based features. They identified 14 features out of 207 available features and presented them to an ANN. 86% accuracy was obtained for normal sounds and sounds of AS and MR.

In 2007 De Vos and Blanckenberg [108], obtained a different time-frequency based features with 12 frequency bins and 10 time intervals and presented it to an ANN. The method was tested on 163 test patients. 90% sensitivity and 96.5% specificity was obtained.

In 2012, Uguz [109] introduced a classification method that used the time-frequency based features and an ANN for classification. 90.48% sensitivity, 97.44% specificity and 95% accuracy was obtained for 120 PCG signals with 50-50 test-train split.

b. Support vector machine-based classification

Ari et al. (2010) [110], classified the PCG signal by using a Least Square SVM (LSSVM). They used this method of classification to identify normal and abnormal heart sounds. 64 recordings of normal and pathological sounds were used to evaluate the proposed method. They used LSSVM on 32 patients with 50/50 split and obtained an 86.72% accuracy.

Zheng et al. 2015 [111], classified the PCG signal by clubbing heart sounds wavelet packets energy fraction and sample entropy features along with SVM. The method was tested on a dataset of 40 normal and 67 pathological patients. 97.17% accuracy, 93.48% sensitivity and 98.55% specificity was reported.

Patidar et al. 2015 [112], developed a classification algorithm for PCG by feeding the Tunable-Q Wavelet Transform features. They used these features as an input to LSSVM with varying

kernel functions. They used a dataset of 163 heart sound recordings with a total of 4628 cycles. 98.8% sensitivity and 99.3% specificity was obtained.

Maglogiannis et al. 2009 [113], developed a classification algorithm for the PCG by feeding the Shannon energy and frequency features. They were obtained from four frequency bands (50–250, 100–300, 150–350, 200–400 Hz) of the heart sound. These features are used as inputs to an SVM classifier. They used a dataset containing 38 normal and 160 heart valve disease sounds. 87.5% sensitivity, 94.74% specificity and 91.43% accuracy was reported.

Gharehbaghi et al. 2015 [114], developed a classification algorithm for the PCG by using frequency band power features during systole over varying length frames. Growing-Time SVM (GTSVM) classifier was used for classification. They used a dataset containing 50-50 train-test split data from 30 patients with Aortic Stenosis (AS), 26 with innocent murmurs and 30 normal. With all of this data and features, they reported 86.4% sensitivity and 89.3% specificity.

c. HMM-based classification

In this type of pathology classification, the posterior probability of the heart sound signals or the extracted features can be used to differentiate between healthy and pathological recordings using a given trained HMM.

Wang et al. 2007 [115], used a mixture of HMM and Mel-Frequency Cepstral Coefficients (MFCCs) for heart sound classification. The feature vector was obtained from time-domain features, short-time Fourier transforms (STFT) and MFCCs. The method was tested on 20 normal and 21 abnormal heart sounds with murmurs. With this, they obtained a sensitivity of 95.2% and a specificity of 95.3%. In this method, MFCCs is used to extract the various features as inputs to the HMM-based classifier for heart sound classification. They used a dataset containing 1381 cycles of normal and pathological heart sounds. 99.21% accuracy was obtained.

Saracoglu et al. 2012 [116], applied a HMM to the frequency spectrum extracted from entire heart cycles. They trained four HMMs, then estimated the posterior probability of the features, optimized the HMM parameters and PCA-based feature selection on the training set. They used a dataset of 60 PCG recordings. 95% sensitivity, 98.8% specificity and 97.5% accuracy was reported.

d. Clustering-based classification

In 1998, Bentley et al. [117], used DWT features alongside kNN classifier and showed that DWT features outperformed morphological based features. 100% and 87% accuracy was reported when detecting pathology in patients with heart valve and prosthetic heart valves respectively for database whose size was not specified.

Quiceno-Manrique et al 2010 [118], combined simple kNN classifier with features from various time-frequency representations. The method was tested on a subset of 16 normal and 6 pathological patients (22 recordings). With this, they reported 98% accuracy for normal and pathologic beats.

In 2010, Avendano-Valencia et al. [119], developed a classification algorithm using time-frequency features and kNN. He used this algorithm for classifying the status of the heart sound as normal and murmur. To reduce the dimension of the data they used Linear decomposition, and tiling partition of the time-frequency plane methods. 26 normal and 19 pathological recordings were used and an accuracy of 99.0% was reported using 11-fold cross-validation.

2.3 Scope of the thesis

Computer Aided Diagnosis of heart sounds includes the following steps namely, Pre-processing, Segmentation and Classification.

Pre-processing: It is used to assess the signal quality, filter out baseline changes and high frequency noises and extract relevant features. Section 2.2 lists some of the methods mentioned in literature to remove the noise from heart sound. Adaptive filter based method was able to remove only a small amount of background noise. Any further changes and modification in the filter design to remove noise resulted in loss of signal coefficients, which itself imposes some limitations for the method. Modulation filtering technique [123] and Wavelet filtering using decomposition [73] have addressed this problem. One of the drawback of these methods is that their efficiency to give a noise less filtered heart sound signal is greatly effected in cases of abnormal heart sounds where murmurs interfere with lung sound frequencies. Further Wavelet filtering using decomposition would lead to loss of important medical information related wavelet coefficients. So the effectiveness of the algorithm is reduced far too much when compared to Modulation filtering. Reduced Kalman filter [121] also failed to remove the noise under low SNR conditions. This method relied heavily on Segmentation of heart sound for

extraction of cardiac cycles under low SNR conditions. Spectral subtraction method [122] for heart sound de-noising created problems of musical noise during noise reduction. Spectral Subtraction method used activity detection to estimate the noise. Activity detection method is usually suitable under stationary noise conditions. Estimation and reduction of Non-stationary noises using [122] resulted in residual noise (musical noise). So these methods suffered miserably under Non-stationary conditions. Tang [92] in his work Non-linear time scaling in cycle frequency domain used the timings of heart sounds and murmurs to aligned them from one cycle to next cycle. Noise and disturbance were subsequently reduced by averaging. The algorithm's performance degraded, if segmentation was inaccurate, or if the assumption that heart sounds were consistent in consecutive cycles is not valid. To overcome this problem, Tang [124] proposed another method of de-noising heart sound using fuzzy detection in joint frequency-time-frequency domain. As noise level was increased the performance of this method degraded. It was noted, however, that the correlation coefficients obtained with this method were not significantly higher than those obtained with previous method [92] in higher noise. The reason behind this performance may be that atoms of heart sound signal are diluted on the joint plane due to the presence of excessive noise.

Segmentation: Segmentation plays a vital role in heart sound analysis as it is used to delineate the start and end of each phase of the heart beat (S1, systolic, S2, diastolic). Section 2.2 discusses some of prominent segmentation methods in literature. Envelope based methods used envelopes of the heart sound for segmentation. The methods relied on threshold for delineating the S1 and S2. Since it was difficult to set the threshold for various types of noises and levels, these methods failed miserably. Liang et al. [75] used Wavelet based segmentation method for delineating PCG signals. Wavelet filters used to contain noises resulted in loss of important Wavelet signal information during reconstruction. So the accuracy reduced greatly. Moukadem et al. [51] used Shannon entropy over S-Transform of the PCG signals to segment the heart sound. High accuracy was obtained for a small dataset. Sun et al [76] used Hilbert transform method to identify and delineate the heart sound signals. They used Shannon entropy to identify the transmission points of S1-S2 and S2-S1. The threshold was fixed empirically. However, the analysis was done under high signal to noise ratio conditions. Jiang and Choi [77] used a new CSCW based envelope extraction method. The method obtained good results compared to Hilbert transform method. But this method was evaluated only for 500 cardiac cycles. Ari et al [79] used simple Squared energy envelope extraction method. Though the performance of this method bettered than that of Shannon Entropy method under the presence of noise the method

showed poor performance. Features based methods used different heart sound features for segmentation. Feature extraction under the influence of noises and disturbances was difficult and the sound had to be de-noised effectively to remove the redundant noise. The number of features used also accounted for the segmentation accuracy. Naseri et.al [80] used frequency and amplitude based features, for PCG segmentation. Though the method was effective in segmenting all sounds, it was tested on a small database of 55 sounds. Kumar et al. used Wavelets based high frequency based feature extraction method. Though the method performed appropriately well, it was tested only on certain heart disorders involving bi-prosthetic valves. Varghees and Ramachandran, [82] used an instantaneous phase feature based extraction method. The method assumed that noise did not affect the phase of the PCG signal. So filtering was required to uncover the signal from noise in order to identify the correct phase of the PCG signal. Inadequate filtering resulted in false segmentation. Nigam et.al, [84] segmented the PCG by using complexity of the PCG. Using dynamic features such as amplitude and frequency they obtained fairly good accuracy. But the method suffered miserably under low SNR conditions. Vepa et al. [85] also used complexity-based features extracted from Wavelet coefficients. Though the method performed better compared to previous method, it was tested only on a small dataset. Papadaniil et.al [86] segmented the PCG by using kurtosis-based features with ensemble empirical mode decomposition. Though the method yielded high performance compared to the previous methods, it was tested only on a small dataset. Gharehbaghi et al [87] used an ECG-referred PCG segmentation method for pediatric patients. This method was validated only for a small set of infant database. Machine learning methods discussed in section 2.2 used small database for segmentation. Methods involving artificial neural networks took more computation time for training. The accuracy of the algorithms also depended on the number of sounds used. Oskiper et.al, [88], used a time-delay neural network method for PCG segmentation. Fairly good accuracy was obtained but only for a small database of sounds. Sepehri at al [89] used MLPNN to segment the heart sounds. The method was tested only on a database of infant heart sounds, but with a good accuracy. Chen et al., [90], used a K-means clustering and a threshold method for segmenting heart sound. Clustering of sounds was greatly influenced by noise presence. So heavy filtering of sounds was necessary. Inaccurate de-noising and bad threshold resulted in poor segmentation. Gupta et al. [91], segmented heart sounds by K-means clustering and homomorphic filtering method. Though the method removed murmurs and segmented S1 and S2 effectively, the evaluation of the method was done on a small dataset of sounds. Tang et al. [92], employed dynamic clustering for segmenting heart sounds. Even in this method noise

greatly influenced the detection of S1 and S2 sounds. Identification of S1 and S2 had to be preceded by suitable pre-processing techniques for optimum noise removal. Residual noise greatly affected the clustering of S1 and S2 sounds. Rajan at al. [93] used Morlet Wavelet transform method to feature extract S1 and S2 sounds and SVD method to segment them. The detection of S1 and S2 sounds boundaries were influenced by noise and led to inaccurate segmentation. This reduced the accuracy of the method. Methods using HMM also took more computational time for training for a larger database. Some of the methods required external reference signal such as ECG or carotid pulse for segmentation. Gamero et.al [94] used a HMMs with grammar constraints to segment the PCG. It was the first method using HMM and was tested only on a small dataset of heart sounds. Rickie et al [95] used HMM method for heart sound segmentation with cross validation. However, the method was tested only on 9 subjects, that too only on adults. Gill et al [96] also used HMM method along with homomorphic filtering for segmenting heart sounds. However, this method too was evaluated on a small dataset. Sedhigian et al [97] also used HMM method with homomorphic filtering on a different database. The performance however degraded with noise. Castro et al. [98] used Wavelet based features with HMM to segment heart sounds. They used same database but failed to achieve good results on a noisy dataset. Schmidt et al [99] used HSMM technique to segment the heart sound. However, the algorithm was trained to detect heart sounds taken from Michigen HSDB. The algorithm did not perform well on real time noisy heart sounds. Springer et al [100] used HSMM method along with logistic regression to segment the heart sound. However, the algorithm was trained to detect heart sounds taken from MITHSDB. The performance of algorithm was appreciable in case of real world noisy heart sounds.

Classification: Four types of classification methods are discussed in section 2.2. Classification by ANN provided lesser than 85% accuracy. Also the accuracy depended on number of features and type of features used. Akay et al [102] used ANN with wavelet based features and obtained an average accuracy of 85% which is far less compared with the current state of the art classification accuracy. Liang and Hartimo [103] used Wavelet Packet Tree along with ANN for heart sound classification and obtained an accuracy of 85% which is lower compared to the current state of the art methods. Uguz et al [104] used ANN with DWT features and Fuzzy logic approach. However, the algorithm focused only on three classes of heart sounds namely Normal, Pulmonary stenosis, Mitral stenosis. Bhatikar et al. [105], used the energy spectrum features of the Fast Fourier Transform (FFT) as inputs to an ANN. The method was tested on a small dataset of 50 patients and medium sensitivity of 83% was obtained. Sepehri et al. [106]

used ANN to identified the five frequency bands with the greatest difference in spectral energy of normal and pathological heart sounds. Their method was also tested on only 50 heart sounds for only 2 classes of heart sounds. Ahlstrom et al. [107], used non-linear complexity-based features with ANN for murmur classification. They obtained lower classification accuracy of 80% for only 2 categories of heart sound. De Vos et.al [108], used time-frequency features with ANN. With few features they obtained high accuracy, but only for lesser number of sounds. Later Uguz [109] also used time-frequency features as an input to an ANN. They obtained a high accuracy, but only for 120 sounds. SVMs are the most popular and prominent of the methods discussed in Section 2.2. However, similar to ANN the accuracy depends on the number of features and type of features used. Also of the SVM methods discussed in section 2.2 has been tested on a small database of heart sounds. And only a few SVM methods are automated. Ari et al. [110], used a Least Square SVM (LSSVM) method for PCG classification. However, their method was restricted only to 50 sounds with accuracy of 86.72%. Zheng et al. 2015 [111], used heart sounds wavelet packets energy fraction and sample entropy features for the SVM. They obtained a high accuracy of 97.17% on 107 sounds. Patidar et al. [112], used the Tunable-Q Wavelet Transform features and LSSVM for PCG classification. They obtained a high accuracy for a small dataset of only 163 sounds. Maglogiannis et al. 2009 [113], used Shannon energy and frequency features with SVM classifier. They obtained a high accuracy of 91.43% for 198 sounds. Gharehbaghi et al. [114], used frequency band power and GTSVM to classify PCG. They obtained a high accuracy for a very small dataset of 56 sounds. Classification by HMM provides good accuracy. However, classification depends on posteriori probability of Heart sound or its features. Also the accuracy depends on the number and type of features used. Wang et al. 2007 [115], classified the PCG by using a mixture of HMM and MFCCs. They obtained high accuracy for a dataset of only 41 normal and abnormal sounds. Saracoglu et al. [116], applied a HMM to the frequency spectrum extracted from entire heart cycles. They obtained a modestly high accuracy of 97.5% for 60 heart sound recordings. Clustering based methods discussed in section 2.2 is completely unsupervised type of classification. K Means algorithm is the most popular type of clustering algorithm. However most of the works related to K Means algorithm has been tested on a smaller database. Bentley et al. [117], used DWT features alongside kNN classifier and showed that DWT features outperformed morphological based features. They obtained a very high accuracy for an unknown sized dataset. Quiceno-Manrique et al [118], used a combination of simple kNN classifier with time-frequency features. They obtained a high accuracy for a small number of 22 sounds. Avendano-Valencia et al. [119], also employed time-frequency features and kNN

approach for classifying normal and murmur patients. A very high accuracy of 99% was achieved for only 45 normal and pathological sounds.

2.4 Motivation for the Present Work

Based upon the objectives and signal processing techniques of the aforementioned methods, and our preliminary results, the motivations for the present work are as follows.

• The ECG reference based segmentation methods can achieve better detection of instances of S1 and S2 components in the PCG signal under varying durations of S1 and S2 sounds, and systole and diastole intervals. But ensuring the accurate detection of the R-peak and end-T-wave is one of the major problems in the presence of time-varying PQRST morphologies, low-amplitude T waves, and various kinds of noise and artefacts.
However, these methods demand additional ECG sensing hardware requirements.
• Although identification of fundamental heart sounds (S1 and S2) is simple, it becomes more complicated when the PCG recordings include other physiological sounds and noise sources. Results showed that the segmentation performance was degraded in the presence of heart murmurs and high pitched sounds such as clicks, snaps and gallops.
• Most methods mainly focus on segmentation of PCG signal into four components: S1-systole-S2-diastole, while only very few methods have been presented for detection of S3 and S4. In pathological cases, the extra sounds including clicks, snaps, rubs and various kinds of systolic and diastolic murmurs (including, early, mid, late, and pan) may occur between the two loudest sounds S1 and S2. Most clinical studies show that different types of abnormal sounds and murmurs are often characterized by using their morphological parameters including timing, duration, intensity, frequency content, and configuration.
• Existing methods use many search-back algorithms with multiple amplitude-dependent, duration-dependent and cycle period-dependent thresholds for rejecting or including the noise segments and missing sound segments. These methods may have poor performance in detection and quantitative measurement of early, mid, late, pan systolic and diastolic murmurs and other high pitched sounds.
Various studies show that the segmentation performance should be further investigated and improved under the influence of abnormal heart sounds (S1, S2, S3, and S4) and murmurs, irregular heart rates and background noise sources. Further, determination of boundaries and peaks of low-amplitude heart sounds and murmurs are still in research. Clinical studies

highlight the need for accurate and robust generalized PCG waveform delineation and parameter extraction framework for effective screening and diagnostic of different kinds of heart defects. To the best of our knowledge, there is no systematic framework for automatically determining the peaks and boundaries of physiological sounds such as heart sounds (S1, S2, S3, and S4), heart systolic and diastolic murmurs (early, mid, late, and pan), and high-pitched sounds.

2.5 Objectives of The Work

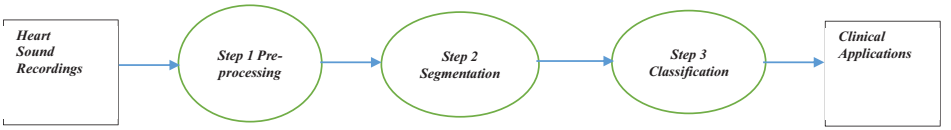

Figure 2.4.1 Block diagram of Automated analysis of heart sound.

Pre-processing of heart sound

To solve the issues of related to de-noising, an approach based on block threshold method in time frequency domain is investigated.

Section 2.3 discusses the various methods used for de-noising of PCG signals. In those methods there is no explicit description about estimation of the noise. Most of the methods rely on empirical methods to cancel or reduce noise. There are very few instances of usage of Time-Frequency methods for noise removal. Since PCG signals are non-stationary in nature, TF methods give better visualization of sounds than other methods mentioned in section 2.3. We propose to de-noise the PCG signals using time frequency block threshold method and compare the results with other time methods namely soft threshold, overlapping group shrinkage and wavelet coefficient threshold techniques. The noise can be either stationary or non-stationary. To estimate stationary noise, activity detection is used. To estimate non-stationary noise, improved minimum controlled recursive averaging method is used. The estimated noise is reduced to the lowest possible residual value using time frequency block threshold method.

Segmentation of heart sound

So a new approach involving bark spectrogram technique is proposed which will segment heart sounds with murmurs without using any form of threshold.

Section 2.3 discusses the popular segmentation methods mentioned in Literature. Most of the methods are threshold based and susceptible to noise. Since perfect noise removal is not researched yet, we devise an algorithm to segment the de-noised sound using the concept of bark spectrogram which is immune to small residual noises. We have tried to assess the performance of Segmentation after pre-processing the PCG signal corrupted with stationary and non-stationary noises. We have compared the performance evaluation of the segmentation algorithm using bark spectrogram with other popular algorithm namely Homomorphic filtering. Cardiac cycles are extracted from the segmented sounds, with each cycle consisting of a single S1, systole, S2 and diastole. The cardiac cycle is used for classification of heart sounds in the next stage of CAD.

Heart sound Feature extraction and Classification

For heart sound classification problem, we investigate the feasibility of using the Loudness features extracted from spectrogram. Loudness based features are clustered using GMM.

Section 2.3 discusses the popular Classification methods mentioned in Literature. We have created a dataset of normal and abnormal heart sounds during the process of cardiac cycle extraction. There is a collection of 3239 sounds from both datasets. The spectrogram is evaluated for each sounds, from which Loudness function is calculated empirically by row-sum of the spectral values. The mean and standard deviation this Loudness function has been evaluated to categorize heart sounds into normal and abnormal. The Loudness features have been found using these two parameters namely mean and standard deviation. GMM is used to cluster the loudness features of the heart sound. This method is compared with other clustering methods, namely K Means and Fuzzy C Means method.

2.6 Database

Heart sounds present in the Physionet MITHSDB [22] is used widely by researchers for heart sound analysis. MITHSDB is obtained from normal and pathological patients worldwide. The training set has five databases (A through E) with over 3000 PCG recordings. The duration of each PCG varies from 5 seconds to 120 seconds. The heart sound recordings were obtained from aortic area, pulmonic area, tricuspid area and mitral area. Abnormal heart sounds were

obtained from patients with a confirmed cardiac diagnosis such as mitral valve prolapse, mitral regurgitation, aortic stenosis and valvular surgery. Recordings included both children and adults. Each patient has contributed from one to six heart sound recordings. All recordings are sampled at 2,000 Hz and are in .wav format.

2.7 Popular Segmentation algorithms from Literature

The following table gives a comparison of the popular segmentation algorithms mentioned in literature.

Authors and the database used	*Performance Metrics*	*Observations*
Liang et.al (1997), 37 recordings (515 cycles) from children with murmurs (14 being pathological)	93.0% Ac	Unsupervised, optimised on entire dataset
Liang et.al (1997), 77 (1165 cycles) recordings from children with both pathological and physiological murmurs	94.6% Ac	Unsupervised, optimised on entire dataset
Sun et.al (2014), 55 recordings (7530 cycles), 51 with valve replacements	97.95% Se, 98.2% Sp	Unsupervised, optimised on entire dataset
Ari et.al (2008), 71 recordings (357 cycles), nine different pathologies	97.47% Ac	No splits between train and test
Vepa et.al (2008), 166 clean heart cycles from normal and pathological patients	84.0% Ac	Unsupervised, no stated segmentation tolerance
Gupta et.al (2007), 41 recordings (340 cycles).	90.29% Ac	Unsupervised

Mix of normal (32%), systolic (36%) and diastolic murmurs (32%)		
Chen et al (2009), 27 recordings of 30 s (997 cycles) from healthy subjects	92.1% Se, 88.4% P+	Unsupervised
Oskiper and Watrous (2002), 30 clean recordings (20 s) from healthy subjects	96.2% Ac	No split between train test set
Gharebhaghi et al (2011), 120 recordings from children, 80 with congenital heart disease (totaling 1200 s, 823 cycles in test set)	93.6% Ac on test set	50% test train set
Yan et.al (2010), Nine recordings (less than 5 s).	55% pathological 99.0% Ac on whole cycle detection	No split between train test set
Gill et.al (2011), 9426.8 s of recordings, normal (22.2%) and various pathologies (ASD, PDA, VSD, and RHD)	S1 : 98.53% Ac, S2 : 98.31% Ac, Cycles: 97.37% Ac	Unsupervised, no stated segmentation tolerance
Gamero and Watrous (2003), 80 recordings from an unknown number of patients of 6–12 s (40 healthy, 40 pathological recordings)	96% and 97% Se, 95% and 95% P+ (healthy and pathological)	No split between train test set, no stated segmentation tolerance
Tang et.al (2010), 26 clean recordings (565 cycles), 3 healthy subjects, and 23 with various pathologies	94.9% and 95.9% Ac (S1 and S2)	No split between train test set and no stated segmentation tolerance
Naseri and Homaeinezhad (2013), 50 2-min	99.0% Se and 98.6%	No split between

healthy and pathological recordings	P+	train test set and segmentation reported on 20% dataset
Varghees and Ramachandran (2014), 64 teaching quality recordings of less than 10 s (701 cycles). Various pathologies	93.06% Ac, 99.43% Se. 93.56% P+	No split between train test set and segmentation reported on partial dataset, no stated segmentation tolerance
Springer et.al (2016), 10 172 s of PCG recorded from 112 patients	F1 score of 95.63±0.85%	Large dataset, Uses ECG reference
Shervegar and Bhat (2017), 10 172 s of PCG recorded from 112 patients	**Accuracy of 96.56%**	**Large dataset, No ECG and no noise threshold**

Table 2.6.1 Popular Segmentation Algorithms from Literature

2.8 Popular Classification algorithms from Literature

The following table gives a comparison of the popular classification algorithms mentioned in literature.

Author	*Database*	*Recording length*	*Classification method*	*Features*	*Se (%)*	*Sp (%)*	*Acc (%)*	*Notes*
Uguz (2012a)	40 normal, 40 pulmonary and 40 mitral stenosis	-	ANN	Wavelet; Large number of features (>2)	100	95.24	98.33	50-50 train-test split; No cross validation; small database; No dimension reduction

Uguz (2012b)	40 normal, 40 pulmonary and 40 mitral stenosis	-	ANN	Time-Frequency; Large number of features (>2)	90.48	97.44	95	50-50 train-test splits; No cross validation; small database; No dimension reduction
Ari et.al (2010)	64 patients (normal and pathological)	Each 8 cycles	SVM	Wavelet; No of features=100	-	-	86.72	50-50 train-test splits; No cross validation; Medium accuracy; No dimension reduction
Zheng et.al (2015)	40 normal and 67 pathological	-	SVM	Wavelet; No of features=250	93.48	98.55	97.17	Cross-Validation; No effective splitting of dataset
Patidar et.al (2015)	Total 4628 heart cycles, 626 normal and 4002 pathological	-	SVM	Wavelet; No of features=180	98.8	99.3	98.9	80% training 20% test; No cross validation
Gharehbaghi et al (2015)	30 normal, 26 innocent and 30 AS	Each 10s	SVM	Frequency; No of features=25	86.4	89.3	-	50-50 train-test split; No cross validation
Sarcoghlu et.al (2012)	40 normal, 40 pulmonary and 40 mitral stenosis	-	HMM	DFT and PCA	95	98.8	97.5	50-50 train-test split
Quiceno-Manrique et al (2010)	16 normal and 6 pathological	-	kNN	Time-Frequency; No of	-	-	98	Cross-Validation

				features not specified				
Avendano-Valencia et al (2010)	26 normal and 19 pathological	-	kNN	Time-Frequency; No of features not specified	99.56	98.45	99.0	Cross-Validation
Karar et al (2017)	3 sets for normal healthy heart and 19 sets represent three cases of heart abnormalities 4, 10, and 5 datasets for AS, AI, and VSD, respectively.	-	SVM	Maximum Lyapunov Exponents	N:100 VSD: 80 AS:100 AI: 100	N:100 VSD:100 AS: 100 AI: 91.67	N: 100 VSD: 95.45 AS: 100 AI: 95.45	Small database; No train-test split specified. Cross validation
Shervegar and Bhat (2017)	**3000 sounds both normal and Pathological**	**3-50**	**GMM**	**Loudness Index; No of features=2**	**100**	**100**	**100**	**Large database; 50-50 training and test split**

Table 2.6.2 Popular Classification Algorithms from Literature

2.9 The thesis Organization

Chapter 1 gives an introduction to Phonocardiography and heart sounds. It discusses the variety of heart sounds that can be acquired from patients suffering from different heart ailments. The chapter also focusses on the clinical parameters used to identify and diagnose the diseases.

Chapter 2 gives an overview of prominent and present methodologies from literature used to for heard sound analysis. Based on the literature survey the problem identification, scope of the thesis and objectives of the work is presented. The chapter also gives details of the database used in this work.

Chapter 3 stresses the need for Pre-processing of heart sounds by de-noising. The chapter highlights the popular Discrete Wavelet Transform based de-noising procedure. The chapter also introduces the application of time frequency method of heart sound de-noising using Block threshold, Overlapping Group Shrinkage and Soft threshold methods. The chapter focusses on the performance of these algorithms and its influence on the Segmentation of the heart sounds under de-noised conditions.

Chapter 4 describes the popular and recent segmentation algorithm mentioned in Literature namely, Homomorphic Filtering. The chapter introduces the new novel low pass filtered Event Synchronous Segmentation and CWT filtered Event Synchronous Segmentation methods. The chapter mentions about the performance of these three algorithms for the given database under the influence of noises and murmurs.

Chapter 5 identifies the features present in the heart sound and classifiers for classification of these features. The chapter focusses on obtaining Loudness features from heart sound using Spectrogram. The chapter also discusses the various classifiers available from Literature for Clustering the heart sounds namely, K Means Classifier, Fuzzy C Means Classifier and Gaussian Mixture Model Classifier. The chapter also discusses the performance of these three classifiers for Loudness features extracted from de-noised heart sound.

Finally, Chapter 6 gives the concluding remarks and the future directions.

Chapter 3

De-noising of Phonocardiograms using Time Frequency methods

The heart sounds contain various types of frequency components like the S1, S2 (in normal PCG) and S3, S4, murmurs (in abnormal PCG). In spite of the advancements in modern medical equipment such as digital stethoscopes contamination of the acquired heart sounds with noise in realistic environment cannot be ruled out. Thus it becomes important to take care of this aspect during the design of heart sound analysis techniques. To study the nature of the de-noising algorithms, the noisy heart sounds present in the database are preprocessed using Wavelet threshold method. This makes the PCG more robust, where robustness refers to high quality of the PCG in terms of high Signal to Noise Ratio and audibility of the PCG in presence of artefacts and disturbing noises in the raw signal, which often hinder the process of cardiac auscultation.

To evaluate the de-noising performance noises are infused purposefully into the heart sounds which were earlier pre-processed using Wavelet threshold method. Additive White Gaussian Noise (stationary noise) and Exponential Noise (non-stationary noise) estimation and reduction is the focus of our study.

Section 2.2.1 gives details about different methods that have been evaluated in literature for removal of noise in PCG of which Wavelet and Time Frequency methods are the most popular. In this chapter adaptation and evaluation of PCG de-noising using time-frequency block threshold (TFBT) method is investigated, further the results obtained from the proposed TFBT method is compared with Time Frequency methods namely, Overlapping Group Shrinkage (OGS) [136] and Soft Threshold method [136]. These PCG de-noising methods are explained in the subsequent sub sections and their performance under the influence of stationary noise and non-stationary noise have been studied.

Noise Estimation plays an important role in PCG enhancement. A low noise estimate results in annoying residual noise in the signal and high noise estimate will result in distorted or loss in intelligibility. The focus of this part of the work lies in correct estimate of the noise in the underlying signal and further reduction of these noises using various noise reduction methods. We have also implemented the wavelet filtering of PCG using the raw sounds, followed by Block threshold method, on the de-noised dataset to obtain the new de-noised dataset. Block threshold algorithm has been used for its better performance compared to other methods such as OGS and ST in terms of SNR and SSNR.

3.1 De-noising of Phonocardiograms using Wavelets

3.1.1 Wavelet Threshold De-noising

For wavelet de-noising noise signals are assumed to be high frequency signals. Information bearing signals like Phonocardiograms appear as either low frequency or more smooth signals. To de-noise noisy Phonocardiogram signals, usually, one dimensional signal de-noising procedure is followed. One dimensional signal de-noising process includes the following: the one dimensional signal is decomposed by wavelet decomposition, a fixed threshold is selected and threshold function is then used to remove the noises which are in high frequency and the one dimensional signal is then reconstructed. Various factors such as the de-noising threshold and the selection of threshold function influence the quality of de-noising.

3.1.2 Wavelet Threshold Function

In classical threshold methods, the decomposition coefficients with values lesser than the threshold are zeroed and those decomposition coefficients with values greater than the threshold are retained [137]. The local properties of the signal are not changed. The inherent discontinuity in the signal creates certain fluctuations during the reconstruction of the original signal.

Consider

$$\overline{\omega_{j,k}} = \begin{cases} \omega_{j,k} & |\omega_{j,k}| \geq \lambda \\ 0, & |\omega_{j,k}| < \lambda \end{cases} \tag{3.1.1}$$

Above certain threshold, the soft threshold function zeros the decomposition coefficient. However, a part of the high frequency coefficients above the threshold is lost [137].

Consider

$$\overline{\omega_{j,k}} = \begin{cases} \omega_{j,k} - \frac{\lambda}{\omega_{j,k}}, & |\omega_{j,k}| \geq \lambda \\ 0, & |\omega_{j,k}| < \lambda \end{cases} \tag{3.1.2}$$

parameter $\omega_{j,k}$ represents the estimated wavelet coefficients. Parameter $\omega_{j,k}$ represents wavelet coefficients after decomposition. λ represent threshold.

3.2 Experimental Evaluation of Wavelet threshold on PCG.

The noisy heart sound data in Physionet database contains either/both stationary and non-stationary noises. To eliminate the presence of noise the heart sounds were subjected to wavelet coefficient threshold method. On the noisy data the wavelet coefficients were calculated using

different wavelets for different decomposition levels. First the sounds were corrupted with stationary AWGN. Wavelet threshold was applied to the noise. The threshold was estimated as follows. There are fixed and variable thresholds. [137]. The fixed threshold is most commonly used. Its expression is $\lambda = \sigma_n\sqrt{2lnN}$, σ_n is the noise standard variance, and N is the length of the signal. DWT performed poorly under low SNR but improves steadily as noise level drops. Soft threshold provides the best de-noising performance compared to hard threshold for an 8^{th} level decomposition using db10 mother wavelet. Appropriate wavelet and decomposition level was chosen considering the high SNR and morphological similarity of the wavelet with heart sound signal. Optimum de-noising requires selection of appropriate decomposition levels and appropriate wavelets. Since noise itself is variable and cannot be controlled, tuning the number of decomposition levels and type of wavelets becomes difficult for all noise levels. As a second step, de-noising of heart sound using time frequency method such as Block threshold method requires tuning only the window duration. Even though Block Threshold methods require noise estimation algorithms for estimation of noise variance, selection of time window for greater range of noise levels is easier. Figure 3.2.1 shows a 2 sec duration of a normal sound. The first pane shows the normal raw sound with many noise peaks in the systole and the diastole. The second pane shows the heart sound obtained by filtering the raw sound. The de-noised sound shows many peaks which remain even after wavelet filtering. The residual noises are shown by the dark circles. To remove such redundant noises time frequency block threshold method is applied. The effect of stationary and nonstationary noises and the significance of the time frequency Block threshold method to remove redundant noises are discussed in the next subsection.

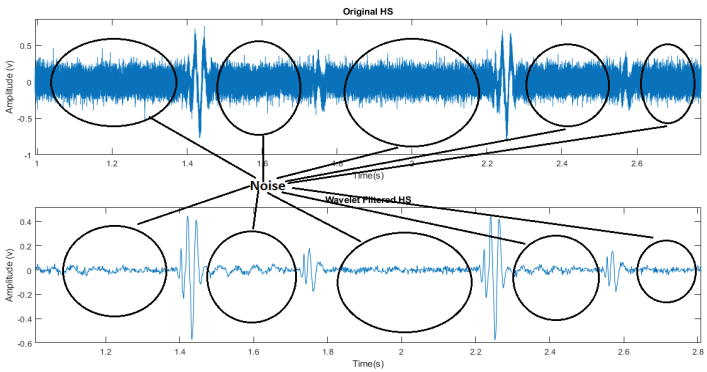

Figure 3.2.1 Normal Heart sound (pane1) and Wavelet de-noised Heart sound (pane 2), noise shown in dark circles. (X axis-time (s) and Y axis -Amplitude (V).

3.3 De-noising using Time frequency block threshold algorithm

3.3.1 Activity Detection

Activity detection (AD) is a procedure used to identify biomedical signal frames (such as heart sounds) and noise frames. In most of the conventional AD algorithms the background noises are stationary over a longer stretch than those of signal. This makes it possible to estimate the time varying noises even during the presence of signal [125]. Some decision rules are used to determine the presence or absence of biomedical signal by comparing the observed signal statistics in the current frame the estimated noise statistics. We need to minimize misdetections at weak speech components. So a hang-over scheme is used to modify this initial decision. It is difficult to optimize the relevant parameters as traditional AD algorithms are usually designed using heuristics. Recently, statistical model has been used to optimize an AD [126]. Maximum likelihood (ML) criterion is used to estimate the unknown parameters. We use these parameters and make a decision using the likelihood ratio test (LRT). The decision-directed (DD) method is used to optimize the decision rule [127]. To optimize the algorithm, we use HMM based Hang-over scheme [128]. Since HS is also a Nonlinear and non-stationary biomedical signal, we have applied DD method based and HMM based hang-over scheme for detecting stationary noises in the heart sound.

Input: A noisy heart sound.

Output: The standard deviation of the noise, sigma.

Steps:

1. Initialize the parameter for AD: threshold for log-likelihood th=0.1, hamming window duration w_dur=0.05s, hop duration h_dur=0.025s, no. of initial noise frames num_noise=20; No of FFT points N_FFT=sampling frequency, fs (=2000) *w_dur.
2. Set the Markov Parameters for Hang Over Scheme: a01=0.5, a10=0.1, a00=1-a01, a11=1-a10. Set coefficient of DD SNR estimation, $\alpha = 0.99$. Set HMM based HS detection probability, $G\_old = 1$.
3. Obtain the noisy HS by selecting a HS from the dataset and corrupting it with various

levels of random noises from -5dB to 15dB.

4. Follow the steps given below to estimate stationary noise (random) noise in the HS using AD.
5. In AD, the first few HS frames are supposed to contain noise only frames (num_noise=20). Calculate the noise variance (λ) for these frames and find the average value.
6. Divide the HS signal into signal frames n by windowing using Hamming window. Calculate the frame variance (Y) of each such signal frames.
7. Follow the steps given below to estimate the noise variance of the stationary noise in the entire signal. Repeat the following steps for 10 iterations.
8. Calculate the posteriori SNR (γ) by taking the ratio of frame variance to noise variance. First 20 frames contain only noises. So frame variance equals noise variance and $\gamma = 1$.
9. Since priori SNR ($\hat{\xi}$) is unknown its estimate is found using the following recursion.

$$\hat{\xi} = \alpha + (1-\alpha)(\gamma - 1), \text{n} = 1; \quad (3.1.3)$$

$$\hat{\xi} = \alpha \frac{\hat{A}(n-1)}{\lambda(n-1)} + (1-\alpha)(\gamma - 1), n \neq 1; \quad (3.1.4)$$

$\hat{A}(n-1)$ is the amplitude estimate of the previous (n-1) frame and $\lambda(n-1)$ is the noise variance of the previous frame.

10. Calculate the parameter $v = \frac{\hat{\xi} \times \gamma}{1+\hat{\xi}}$ and nonlinear Gain function

$$\hat{G}(n-1) = \frac{\sqrt{\pi}}{2}\left(\frac{\sqrt{v}}{\gamma}\right) \exp\left(\frac{-v}{2}\right)[(1+v)I_0\left(\frac{v}{2}\right) + vI_1\left(\frac{v}{2}\right)]. \quad (3.1.5)$$

I_0 and I_1 are modified Bessel functions of zero and first order.

11. Estimate the amplitude of the (n-1) frame by multiplying the nonlinear gain function with the spectrum of the signal frame. Also estimate noise variance

$$\lambda(n-1) = log\left(\frac{\hat{\xi}}{1+\hat{\xi}}\right) + \frac{\hat{\xi} \times \gamma}{1+\hat{\xi}} \quad (3.1.6)$$

and find the mean value by average with no. of FFT points.

12. Estimate the weights

$$w = \frac{\lambda_{mean}}{1+\lambda_{mean}} \quad (3.1.7)$$

and update

$$\lambda(n) = w \times \lambda(n-1) + (1-w)Y. \quad (3.1.8)$$

13. Find the absolute sum of the difference of $\lambda(n)$ *and* $\lambda(n-1)$. If the difference is

less than 0.000001 stop the iteration.

14. Calculate the HMM based HS detection probability

$$\boldsymbol{G} = \frac{\boldsymbol{a01}+\boldsymbol{a11}\times\boldsymbol{G\_old}}{\boldsymbol{a00}+\boldsymbol{a10}\times\boldsymbol{G\_old}}\boldsymbol{\lambda_{mean}}. \quad (3.1.9)$$

Update $G\_old = G$ and calculate log maximum likelihood

$$\boldsymbol{L_{ML}} = \frac{\mathbf{1}}{\boldsymbol{N_{FFT}}}\boldsymbol{sum}(\boldsymbol{\gamma} - \mathbf{log}(\boldsymbol{\gamma}) - \mathbf{1}). \quad (3.1.10)$$

15. Classify the signal frame as noise frame if G is less than threshold th or n is less than 20, else classify the signal frame as HS frame.

3.3.2 Improved Minimum Controlled Recursive Averaging (IMCRA)

Non Stationary and low SNR Noise spectrum estimation methods use reliable and fast tracking of variations in the noise. Activity Detector (AD) based methods depend on the heart sound absence. It is difficult to use AD for tuning weak signal components and signals with low input SNR, [129], [130], [131]. Other methods like Power spectral domain based Histograms techniques [132], [133], [134], are computationally expensive. They require more memory resources, and do not perform well in low SNR conditions. Also, the signal segments needed to build the histograms are in the order of several hundred milliseconds. To overcome this problem improved minima controlled recursive averaging (IMCRA) method was proposed by Cohen [136]. He adjusted the signal presence probability and obtained the noise estimate. He used a smoothing parameter and averaged past spectral power values. The signal presence probability depended on the minima values of the periodogram. To estimate noise using IMCRA he followed a two-step procedure that included smoothing and minimum tracking. In the first step rough activity detection was used in each frequency band. In the next step, Smoothing was used to exclude relatively strong signal components making noise estimation robust. IMCRA is used to estimate non stationary noise.

The algorithm for IMCRA is stated as follows:

Input: A noisy heart sound.

Output: The standard deviation of the noise sigma.

Steps:

1. Initialize the variables at the first frame of all frequency bins k: instantaneous noise spectrum estimate $\hat{\lambda}_d(k,0) = |Y(k,0)|^2$, mean noise spectrum estimate $\bar{\lambda}_d(k,0) = |Y(k,0)|^2$, posteriori SNR $\gamma(k,0)$=1, Smoothed Spectrogram minimum from first iteration $S_{min}(k,0) = S_f(k,0)$, Smoothed Spectrogram from first iteration $S(k,0) = S_f(k,0)$, Smoothed Spectrogram from second iteration $\tilde{S}(k,0) = S_f(k,0)$, conditional gain $G_{H1}(k,0) = 1$, Smoothed Spectrogram minimum from second iteration $\tilde{S}_{min}(k,0) = S_f(k,0)$, first Smoothed Spectrogram running minimum $S_{\min\_sw}(k,0) = S_f(k,0)$, second Smoothed Spectrogram running minimum $\tilde{S}_{\min\_sw}(k,0) = S_f(k,0)$.

2. Initialize the counter for frames, j=0.

3. For all time frames l and frequency frames k repeat the steps that follow.

4. Compute the posteriori SNR $\gamma(k,l)$ using

$$\boldsymbol{\gamma(k,l)} = \frac{|Y(k,l)|^2}{\hat{\lambda}_d(k,l)} \tag{3.1.11}$$

priori SNR

$$\hat{\xi}(k,l) = \alpha G_{H1}^2(k,l-1)\gamma(k,l-1) + (1-\alpha)max(\gamma(k,l)-1,0) \tag{3.1.12}$$

and

$$G_{H1}(k,l) = \frac{\hat{\xi}(k,l)}{1+\hat{\xi}(k,l)} exp(\frac{1}{2}\int_{v(k,l)}^{\infty}\frac{e^{-t}}{t}dt) \tag{3.1.13}$$

where α is a weighting factor (for a priori SNR estimation) that controls the tradeoff between noise reduction and speech distortion, $G_{H1}(k,l)$ is the spectral gain function of the Log-Spectral Amplitude (LSA) estimator when speech is surely present and

$$v(k,l) = \frac{\hat{\xi}(k,l)}{1+\hat{\xi}(k,l)}\gamma(k,l). \tag{3.1.14}$$

5. Compute the first iteration of the smoothed power spectrum of $S(k,l)$ using time-frequency smoothing. In frequency, we use a window function b whose

length is 2w +1:

$$S_f(k,l) = \sum_{i=-w}^{w} b(i)\,|Y(k-1,l)|^2. \tag{3.1.15}$$

In time, the smoothing is performed by a first order recursive averaging, given by

$$S(k,l) = \alpha_s S(k,l-1) + (1-\alpha_s) S_f(k,l). \tag{3.1.16}$$

The running minimum

$$S_{min}(k,l) = min(S_{min}(k,l-1), S(k,l)) \tag{3.1.17}$$

and

$$S_{min\_sw}(k,0) = min(S_{min\_sw}(k,0), S(k,l)) \tag{3.1.18}$$

are updated.

6. Based on the first iteration smoothing and minimum tracking, the rough decision about HS presence using the indicator function $I(k,l)$ for AD is obtained. It follows [13] that there exists a constant factor B_{min} , independent of the noise power spectrum, such that

$$E\left(S_{min}(k,l)\middle|\hat{\xi}(k,l)=0\right) = B_{min}^{-1}\hat{\lambda}_d(k,l). \tag{3.1.19}$$

The factor B_{min} represents the bias of a minimum noise estimate. We calculate

$$\gamma_{min}(k,l) = \frac{|Y(k,l)|^2}{B_{min}S_{min}(k,l)} \tag{3.1.20}$$

and recursive average of the a priori SNR,

$$\zeta(k,l) = \frac{S(k,l)}{B_{min}S_{min}(k,l)}. \tag{3.1.21}$$

The indicator function is given by

$$I(k,l) = \begin{cases} 1 & if\ \gamma_{min}(k,l) < \gamma(k,0)\ and\ \zeta(k,l) < \zeta(k,0) \\ 0 & otherwise \end{cases} \tag{3.1.22}$$

7. Compute the second iteration of the smoothed power spectrum of $\tilde{S}(k,l)$. In frequency, we use a window function b whose length is 2w +1:

$$\tilde{\boldsymbol{S}}_{\boldsymbol{f}}(\boldsymbol{k},\boldsymbol{l}) = \begin{cases} \dfrac{\sum_{i=-w}^{w}(\boldsymbol{b}(\boldsymbol{i})\,\boldsymbol{I}(\boldsymbol{k}-\boldsymbol{1},\boldsymbol{l})|\boldsymbol{Y}(\boldsymbol{k}-\boldsymbol{1},\boldsymbol{l})|^{2})^{2}}{\sum_{i=-w}^{w}(\boldsymbol{b}(\boldsymbol{i})\,\boldsymbol{I}(\boldsymbol{k}-\boldsymbol{1},\boldsymbol{l})} & \boldsymbol{if}\ \boldsymbol{\gamma}_{\boldsymbol{min}}(\boldsymbol{k},\boldsymbol{l}) < \boldsymbol{\gamma}(\boldsymbol{k},\boldsymbol{0})\ \boldsymbol{and}\ \boldsymbol{\zeta}(\boldsymbol{k},\boldsymbol{l}) < \boldsymbol{\zeta}(\boldsymbol{k},\boldsymbol{0}) \\ |\boldsymbol{Y}(\boldsymbol{k}-\boldsymbol{1},\boldsymbol{l})|^{2} & \boldsymbol{Otherwise} \end{cases}$$

(3.1.23)

Smoothing in time is given, as before, by a first-order recursive averaging:

$$\tilde{\boldsymbol{S}}(\boldsymbol{k},\boldsymbol{l}) = \boldsymbol{\alpha}_{\boldsymbol{s}}\tilde{\boldsymbol{S}}(\boldsymbol{k},\boldsymbol{l}-\boldsymbol{1}) + (\boldsymbol{1}-\boldsymbol{\alpha}_{\boldsymbol{s}})\tilde{\boldsymbol{S}}_{\boldsymbol{f}}(\boldsymbol{k},\boldsymbol{l}). \quad (3.1.24)$$

The running minimum

$$\tilde{\boldsymbol{S}}_{\boldsymbol{min}}(\boldsymbol{k},\boldsymbol{l}) = \boldsymbol{\min}(\tilde{\boldsymbol{S}}_{\boldsymbol{min}}(\boldsymbol{k},\boldsymbol{l}-\boldsymbol{1}),\tilde{\boldsymbol{S}}(\boldsymbol{k},\boldsymbol{l})) \quad (3.1.25)$$

And

$$\tilde{\boldsymbol{S}}_{\boldsymbol{\min\_sw}}(\boldsymbol{k},\boldsymbol{l}) = \boldsymbol{\min}(\tilde{\boldsymbol{S}}_{\boldsymbol{\min\_sw}}(\boldsymbol{k},\boldsymbol{l}-\boldsymbol{1}),\tilde{\boldsymbol{S}}(\boldsymbol{k},\boldsymbol{l}) \quad (3.1.26)$$

are updated.

8. Compute the a priori signal absence probability

$$\tilde{\boldsymbol{q}}(\boldsymbol{k},\boldsymbol{l}) = \begin{cases} \boldsymbol{1} & \boldsymbol{if}\ \tilde{\boldsymbol{\gamma}}_{\boldsymbol{min}}(\boldsymbol{k},\boldsymbol{l}) \le \boldsymbol{1}\ \boldsymbol{and}\ \tilde{\boldsymbol{\zeta}}(\boldsymbol{k},\boldsymbol{l}) < \boldsymbol{\zeta}_{\boldsymbol{0}} \\ \frac{\boldsymbol{\gamma}_{\boldsymbol{1}}-\tilde{\boldsymbol{\gamma}}_{\boldsymbol{min}}(\boldsymbol{k},\boldsymbol{l})}{\boldsymbol{\gamma}_{\boldsymbol{1}}-\boldsymbol{1}} & \boldsymbol{if}\ \boldsymbol{1} \le \tilde{\boldsymbol{\gamma}}_{\boldsymbol{min}}(\boldsymbol{k},\boldsymbol{l}) < \boldsymbol{\gamma}_{\boldsymbol{1}} \\ \boldsymbol{0} & \boldsymbol{otherwise} \end{cases}$$

(3.1.27)

$threshold\ \gamma_1 = -\log(\epsilon_1) = 3, experimentally\ \epsilon_1 (= 0.05) > \epsilon\ (= 0.01)$, a priori signal presence probability,

$$\tilde{\boldsymbol{p}}(\boldsymbol{k},\boldsymbol{l}) = \{\boldsymbol{1} + \tfrac{\boldsymbol{q}(\boldsymbol{k},\boldsymbol{l})}{\boldsymbol{1}-\boldsymbol{q}(\boldsymbol{k},\boldsymbol{l})}\left(\boldsymbol{1}+\hat{\boldsymbol{\xi}}(\boldsymbol{k},\boldsymbol{l})\right)\boldsymbol{exp}(\boldsymbol{v}(\boldsymbol{k},\boldsymbol{l}))\}^{-\boldsymbol{1}} \quad (3.1.28)$$

and time varying frequency dependent parameter:

$$\tilde{\boldsymbol{\alpha}}_{\boldsymbol{d}} = \boldsymbol{\alpha}_{\boldsymbol{d}} + (\boldsymbol{1} - \boldsymbol{\alpha}_{\boldsymbol{d}})\tilde{\boldsymbol{p}}(\boldsymbol{k},\boldsymbol{l}) \quad (3.1.29)$$

where smoothing parameter ($experimentally\ \alpha_d = 0.85$) $lies\ in \quad 0 < \alpha_d < 1$,

9. Update the noise spectrum estimate:

$$\hat{\boldsymbol{\lambda}}_{\boldsymbol{d}}(\boldsymbol{k}, \boldsymbol{l}+\mathbf{1}) = \tilde{\boldsymbol{\alpha}}_{\boldsymbol{d}}(\hat{\boldsymbol{\lambda}}_{\boldsymbol{d}}(\boldsymbol{k}, \boldsymbol{l})) + (\mathbf{1} - \tilde{\boldsymbol{\alpha}}_{\boldsymbol{d}})|\boldsymbol{Y}(\boldsymbol{k}, \mathbf{0})|^{\mathbf{2}}) \quad (3.1.30)$$

and mean noise spectrum estimate:

$$\bar{\boldsymbol{\lambda}}_{\boldsymbol{d}}(\boldsymbol{k}, \boldsymbol{l}+\mathbf{1}) = \boldsymbol{\beta} \times \hat{\boldsymbol{\lambda}}_{\boldsymbol{d}}(\boldsymbol{k}, \boldsymbol{l}+\mathbf{1}) \quad (3.1.31)$$

where β compensates the bias when signal is absent.

3.3.3 Block threshold method

This work discusses, time-frequency block threshold method for heart sound de-noising. This method utilizes a Fourier coefficient matrix filter that is used to de-noise the heart sound. Ordinary filtering using Wiener filter creates artefacts called musical noise. To prevent such artefacts, a non-diagonal processing like a block threshold method is needed. Block threshold algorithm parameters are chosen carefully to reduce the risk involved in Stein estimation. Numerical experiments demonstrate the performance and robustness of this procedure.

Input: Noisy heart sound signal with noise sigma estimated from section 3.3.1 or 3.3.2

Output: The de-noised heart sound.

Steps:

1. First the STFT ($W \times Z$) of the heart sound signal using 50ms time duration hanning window.
2. Initialize Attenuation factor matrix AttenFactorMap, Flag depth Flagdepth and thresholded coefficient matrix STFTcoefth to zero.
3. Create a 3×5 $\lambda\ Matrix$:

$$\boldsymbol{M}_{\boldsymbol{\lambda}} = \begin{pmatrix} \mathbf{1.5} & \mathbf{1.8} & \mathbf{2} & \mathbf{2.5} & \mathbf{2.5} \\ \mathbf{1.8} & \mathbf{2} & \mathbf{2.5} & \mathbf{3.5} & \mathbf{3.5} \\ \mathbf{2} & \mathbf{2.5} & \mathbf{3.5} & \mathbf{4.7} & \mathbf{4.7} \end{pmatrix}. \quad (3.1.32)$$

4. SURE Matrix of size = (3×5) is initialized to zero.
5. Start at time 1 and frequency -1.

6. **For** every block 1 to Z/8 loop over time **do**:
7. For zero frequency deal with block 1× 8:
8. Compute Attenuation coefficient,

$$\boldsymbol{a_i = 1 - \frac{1}{\xi_i + 1}} \tag{3.1.33}$$

where

$$\boldsymbol{\xi_i = \frac{\bar{Y}_i^2}{\bar{\sigma}_i^2}} \tag{3.1.34}$$

and $\bar{Y}_i^2$ is the empirical mean on the block i and place it in AttenFactorMap. Place the block 1× 8 (subdivision) in Flagdepth.

9. **For** each Macroblock in negative frequencies **do**:
10. **For** each subdivision **do**.
11. Compute the Risk:

$$\widehat{\boldsymbol{R}}_{\boldsymbol{k}} = \boldsymbol{\sigma^2}\left(\boldsymbol{B_k^{\#}} + \frac{\boldsymbol{\lambda^2 B_k^{\#} - 2\lambda\left(B_k^{\#} - 2\right)}}{\frac{\boldsymbol{Y}}{\boldsymbol{\sigma}}}\mathbf{1}_{\boldsymbol{Y \geq \lambda\sigma^2}} + \boldsymbol{B_k^{\#}}\left(\frac{\boldsymbol{Y}}{\boldsymbol{\sigma}} - \mathbf{2}\right)\mathbf{1}_{\boldsymbol{Y < \lambda\sigma^2}}\right). \tag{3.1.35}$$

This formula gives the estimation of the risk of the block i of size $B^{\#}_i$

12. Store the result in the corresponding subdivision of SURE matrix.
13. **end**
14. Find the minimum of SURE and the matched subdivision.
15. **For** each mini block of this optimal subdivision **do**
16. Compute a_i as in step 8 and store it in right place in AttenFactorMap.
17. Store the subdivision in Flagdepth.
18. **end**
19. **end**
20. **For** last block not Full in frequency **do**
21. Do the same as zero frequency.
22. **end**
23. **For** last block not Full in time and frequency **do**
24. Do hard thresholding
25. **end**
26. **end**
27. For positive frequencies conjugate the results from negative frequencies.
28. Find the thresholded coefficient matrix:

$$\boldsymbol{STFTCoefth = STFTCoef \times AttenFactorMap}. \tag{3.1.36}$$

29. Weiner filter this result.
30. Invert the STFT to obtain the reconstructed signal.

3.4 Comparison of TFBT with OGS and ST de-noising algorithm

3.4.1 Overlapping group Shrinkage algorithm

Bio-signals x are not only sparse but also exhibit a inter and intra scale clustering property. Likewise, the clustering/grouping property is also apparent in a typical heart sound spectrogram. In both cases, significant (large-amplitude) values of x tend not to be isolated. In this work simple translation-invariant shrinkage algorithm is discussed. The algorithm minimizes the cost function with the penalty where the set J defines the group. To de-noise, we discuss a simple method to set the regularization parameter lambda similar to the 'three-sigma' rule. The method allows for lambda to be selected so as to ensure that the noise variance is reduced to a specified fraction of its original value. This method does not aim to minimize the mean square error or any other measure involving the signal to be estimated, and is thus non-Bayesian. The method for setting lambda is analytically intractable due to the absence of the estimator. However, with appropriate pre-computation, the method can be implemented by table look-up. To avoid musical noise, the clustering behaviour of STFT coefficients of heart sound waveforms have been considered. The parameter lambda is used to reduce the penalty function such that the noise component of the signal is reduced to a smallest possible value. The variance of the signal is given by $\sigma_x^2(T) = 2(1+T^2)Q(T) - T\sqrt{\frac{2}{\pi}}\exp(-\frac{T^2}{2})$, where $Q(T) = \frac{1}{2\pi}\int_T^\infty e^{-t^2/2}dt = 0.5(1-\text{erf}(\frac{T}{\sqrt{2}}))$.

Input: Noisy heart sound signal with noise sigma estimated from section 3.3.1 or 3.3.2

Output: The de-noised heart sound.

Steps:

1. Initialize the vector x to zero, and assign the input y to x.
2. Index i takes values from 0 to N-1 and $x(i) \neq 0$.
3. Repeat the following steps:
4. The penalty function, $R(x)$, is chosen to promote the known behavior of x.

$$\boldsymbol{R(x) = r(i) = \sum_{j=1}^{K}[\sum_{k=1}^{K}|x(i-j+k|^2]^{\frac{1}{2}}, i = [0, \ldots . K-1]},\quad (3.1.37)$$

with K being group/block size.

5. Estimate $\boldsymbol{x(i)} = \frac{\boldsymbol{y(i)}}{\boldsymbol{1+\lambda r(i)}}$ (3.1.38)

until convergence, where λ is the regularization parameter. The parameter λ is fixed as per table 3.4.1.

	Output Standard deviation σ_x			
Group	$\mathbf{10^{-2}}$	$\mathbf{10^{-3}}$	$\mathbf{10^{-4}}$	$\mathbf{10^{-5}}$
$\mathbf{1 \times 1}$	3.36	4.38	5.24	6.00
$\mathbf{1 \times 2}$	1.69 (1.73)	2.15 (2.24)	2.38 (2.67)	2.46 (2.94)
$\mathbf{1 \times 3}$	1.16 (1.18)	1.46 (1.52)	1.60 (1.77)	1.64 (1.99)
$\mathbf{1 \times 4}$	0.89 (0.91)	1.12 (1.16)	1.23 (1.36)	1.27 (1.53)
$\mathbf{1 \times 5}$	0.73 (0.75)	0.92 (0.95)	1.01 (1.12)	1.04 (1.25)
$\mathbf{2 \times 2}$	0.86 (0.87)	1.08 (1.33)	1.19 (1.31)	1.23 (1.48)
$\mathbf{2 \times 3}$	0.59 (0.67)	0.74 (0.77)	0.80 (0.89)	0.82 (1.01)
$\mathbf{2 \times 4}$	0.46 (0.48)	0.57 (0.59)	0.62 (0.69)	0.64 (0.78)
$\mathbf{2 \times 5}$	0.38 (0.41)	0.46 (0.49)	0.51 (0.57)	0.52 (0.64)
$\mathbf{3 \times 3}$	0.41 (0.43)	0.50 (0.53)	0.55 (0.61)	0.56 (0.69)
$\mathbf{3 \times 4}$	0.33 (0.35)	0.39 (0.42)	0.43 (0.48)	0.44 (0.54)
$\mathbf{3 \times 5}$	0.29 (0.31)	0.32 (0.36)	0.35 (0.40)	0.36 (0.45)
$\mathbf{4 \times 4}$	0.27 (0.30)	0.30 (0.34)	0.33 (0.38)	0.34 (0.43)
$\mathbf{4 \times 5}$	0.24 (0.26)	0.26 (0.30)	0.27 (0.33)	0.28 (0.37)
$\mathbf{5 \times 5}$	0.21 (0.23)	0.22 (0.26)	0.23 (0.29)	0.24 (0.32)
$\mathbf{2 \times 8}$	0.28 (0.30)	0.31 (0.35)	0.33 (0.39)	0.35 (0.43)

Table 3.4.1 Regularization parameter λ to achieve specified output standard deviation when OGS is applied to a real standard normal signal: full convergence-150 iterations (25 iterations).

3.4.2 Soft Threshold Algorithm

Soft threshold algorithm is a type of threshold algorithm in which the STFT/ wavelet coefficients that are below the soft threshold T are shrinked towards zero. The noisy signal used is simulated by adding independent white Gaussian noise with standard deviation σ = 0.03. We use '3σ rule' to contain noise using Soft threshold method. The '3σ rule' states that nearly all values of a Gaussian random variable lie within three standard deviations of the mean (in fact, 99:7%). Hence, by using 3σ as a threshold with the soft threshold function, nearly all the noise will be eliminated with threshold T = 3σ = 0.09.

Input: Noisy heart sound signal with noise sigma estimated from section 3.3.1 or 3.3.2

Output: The de-noised heart sound.

Steps:

1. Initialize the vector x to zero.

2. Estimate x using the formula x = max(1 - T./abs(y), 0) .* y with T=3σ.
3. Estimate the inverse STFT to obtain the de-noised signal.

3.5 Block Threshold Algorithm-Results and Discussion

3.5.1 Performance Metrics

To evaluate the performance of de-noising in Phonocardiograms using TFBT, OGS and BT methods the following performance metrics were evaluated.

Signal to Noise Ratio: $$\boldsymbol{SNR = 10log_{10}(\frac{\sum_{n=0}^{N-1} f^2[n]}{\sum_{n=0}^{N-1}(f[n]-\tilde{f}[n])^2})} \quad (3.1.41)$$

Segmental Signal to Noise Ratio: $$\boldsymbol{SSNR = \frac{1}{H}\sum_{l=0}^{H} T(10log_{10}(\frac{\sum_{n=0}^{S-1} f^2\left[\frac{n+lS}{2}\right]}{\sum_{n=0}^{S-1}(f\left[\frac{n+lS}{2}\right]-\tilde{f}\left[\frac{n+lS}{2}\right])^2}))} \quad (3.1.42)$$

Mean Square Error: $$\boldsymbol{MSE = \frac{1}{N}\sum_{n=0}^{N-1}(f[n]-\tilde{f}[n])^2} \quad (3.1.43)$$

3.5.2 Results from the experiments

The noisy heart sounds taken from Physionet Heart sound database containing 3239 sounds are Wavelet Filtered to remove most of the noises as discussed in Section 3.2. However not all noises get removed using this method and some residual noises still remain. So the residual noises could have detrimental effects in further processing of the cardiac sound signals. To remove such noises application of Time Frequency methods is proposed as a second stage of preprocessing following the Wavelet filtering technique as the first stage. To better visualize the effect of de-noising, we compare the signal plot vis-a-vis its time frequency (STFT) representation. To study the performance of Time Frequency methods in the de-noising of heart sounds, the noises are purposefully infused. The effect of two types of noises namely stationary noise -Additive White Gaussian Noise (AWGN) and Non-Stationary noise-exponential noise have been studied and results have been reported in the form of tables 3.5.1-3.5.4. Estimation of Stationary noise is done using Activity Detection, while Non-stationary noise is estimated using IMCRA method. Figure 3.5.1 shows the time as well as STFT zoom plot of a single cycle of heart sound corrupted with exponential noise with 0.03 noise level (sigma). Figure 3.5.5 shows the time and STFT plot of heart sound corrupted with random noise with 0.03 noise level (sigma). The fundamental heart sounds S1 and S2 are visible in the signal plot as well as STFT in both of these figures. The STFT is calculated with 50% overlapping blocks of length of 512 samples for a sampling frequency of 2000Hz. Other block sizes may be suitable for other sampling frequencies. Figures 3.5.2 and 3.5.6 shows the noisy PCG signal subjected to Soft

threshold. The Noisy STFT is subjected to soft threshold with threshold parameter T selected so as to reduce the noise standard deviation down to 0.1% of its value such that noise is inaudible. The noise is sufficiently suppressed so that musical noise is not audible; however, the signal is distorted due to the relatively high threshold used. The signal is over-smoothened and spectrogram shows loss of vital information present in the signal. Figures 3.5.3 and 3.5.7 show the application of OGS algorithm to the noisy STFT. HS signals were de-noised with OGS algorithm with various group sizes. A suitable group size - 8 * 2 (i.e., eight frequency bins * two time bins) is chosen based on the better de-noising effect. Other group sizes may be more appropriate for other sampling rates and STFT block lengths. As in the soft threshold experiment, the parameter λ was selected so as to reduce the noise standard deviation down to 0.1% of its value. Regularization parameter λ was fixed as $\lambda = 0.32\sigma$ as per table 3.4.1. The signal and STFT plots indicate that OGS algorithm removes more noise compared to Soft threshold method. As a next step, Block threshold algorithm is applied to the noisy STFT. Figures 3.5.4 and 3.5.8 gives the signal and STFT plot of the noisy PCG signal subjected to TFBT method. Block size chosen was 3*5. Other block sizes are suitable for sounds with other sampling frequency. The signal and STFT plot indicate that TFBT method removes more noise compared to the other two methods. To better understand the effect of de-noising we refer to the tables 3.5.1-3.5.4.

Table 3.5.1 shows the de-noising metrics for random noise corrupted Normal PCG signal subjected to three different de-noising algorithms-ST, OGS and TFBT. The performance of de-noising was measured in terms of Signal to Noise Ratio (SNR), Segmental-Signal to Noise Ratio (SSNR) and Mean Squared Error (MSE). For better de-noising effect, the PCG signals under study should have higher SNR and SSNR and lower MSE. For a noise level of 0.03 (sigma), ST method showed 0.67 dB SNR, 0.03 dB SSNR and 0.004127 MSE. For the same noise level OGS method showed 1.48 dB SNR, 0.21 dB SSNR and 0.00329. OGS method showed better de-noising performance compared to ST method. The de-noising performance of BT algorithm was still better with 8.29 dB SNR, 3.10 dB SSNR and 0.000698 MSE. Table 3.5.2 shows the de-noising performance of random noise corrupted Abnormal PCG signal. ST method showed SNR of 0.18, SSNR of 0.01 dB and MSE of 0.005349. OGS method showed SNR of 0.33 dB, SSNR of 0.02 dB and MSE of 0.005166, while BT method showed 8.28 dB SNR, 4.76 dB SSNR and 0.000819. Table 3.5.3 shows the de-noising of Normal PCG signal corrupted with exponential noise. The noise level was kept unchanged with 0.03 (sigma). The table indicates that ST method showed 5.43 dB SNR, 0.82 dB SSNR and 0.002 MSE. At the

same time OGS algorithm showed 7.54 dB SNR, 1.23 dB SSNR and 0.000897 MSE. Analysis of normal PCG corrupted with exponential noise showed 11.11 dB SNR, 5.13 dB SSNR and 0.000427 MSE. Table 3.5.4 gives the de-noising mechanism of Abnormal PCG corrupted with exponential noise and subjected to de-noising algorithms. ST method showed 4.17 dB SNR, 1.22 dB SSNR and 0.003306 MSE. Meanwhile, OGS method showed 5.51 dB SNR, 1.49 dB SSNR and 0.001589 MSE. Similarly, TFBT method 10.81 dB SNR, 7.04 dB SSNR and 0.000468 MSE.

Parameter (sigma = 0.03)	**ST**	**OGS**	**BT**
SNR	0.67	1.48	**8.29**
SSNR	0.03	0.21	**3.10**
MSE	0.004127	0.00329	**0.000698**

Table 3.5.1 De-noising Normal sounds for different methods (Stationary noise)

Parameter (sigma = 0.03)	**ST**	**OGS**	**BT**
SNR	0.18	0.33	**8.28**
SSNR	0.01	0.02	**4.76**
MSE	0.005349	0.005166	**0.000819**

Table 3.5.2 De-noising Abnormal sounds for different methods (Stationary noise)

Parameter (sigma = 0.03)	**ST**	**OGS**	**BT**
SNR	5.43	7.54	**11.11**
SSNR	0.82	1.23	**5.13**
MSE	0.002	0.000897	**0.000427**

Table 3.5.3 De-noising Normal sounds for different methods (Non-stationary noise)

Parameter (sigma = 0.03)	**ST**	**OGS**	**BT**
SNR	4.17	5.51	**10.81**
SSNR	1.22	1.49	**7.04**
MSE	0.003306	0.001585	**0.000468**

Table 3.5.4 De-noising Abnormal sounds for different methods (Non Stationary noise)

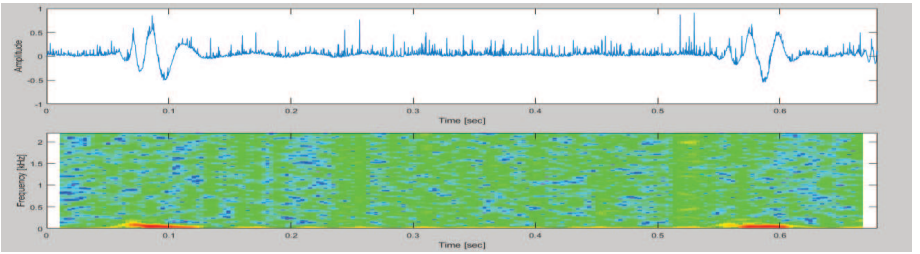

Figure 3.5.1 Non stationary exponential noise corrupted heart sound

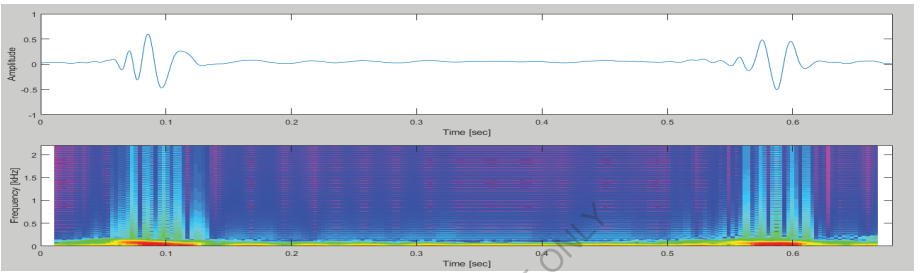

Figure 3.5.2 Soft thresholded heart sound under non stationary conditions

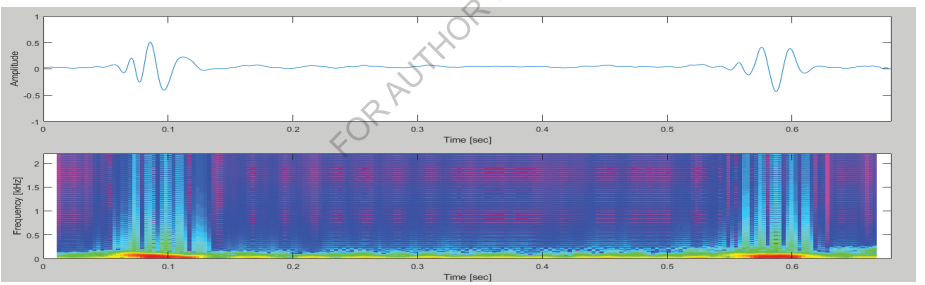

Figure 3.5.3 OGS thresholded heart sound under non stationary conditions

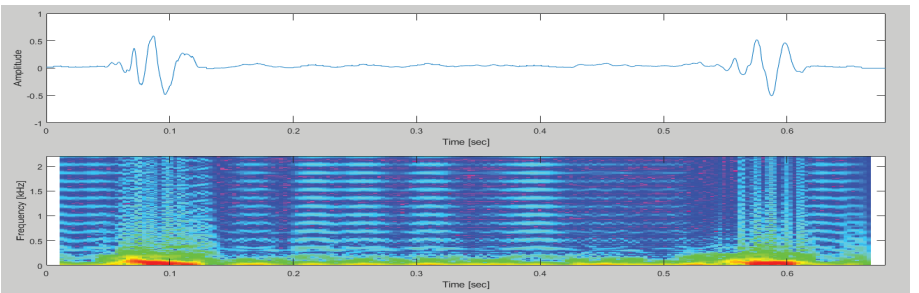

Figure 3.5.4 Block thresholded heart sound under non stationary conditions

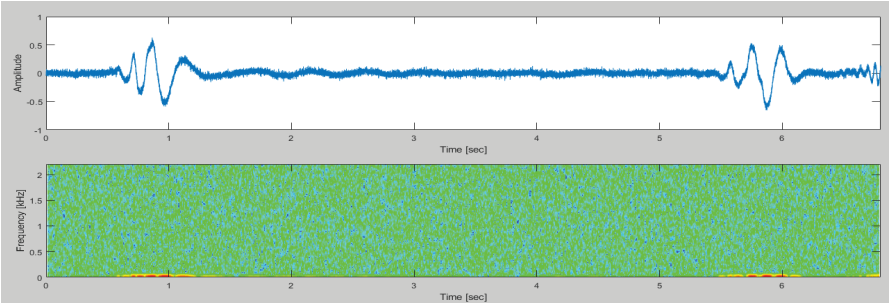

Figure 3.5.5 Noisy heart sound corrupted by random stationary noise

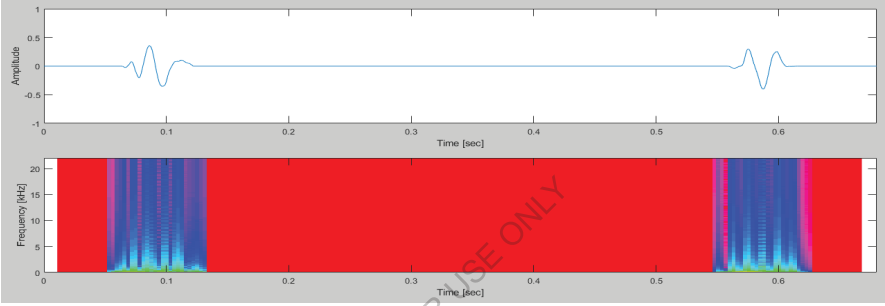

Figure 3.5.6 Soft thresholded heart sound under stationary conditions

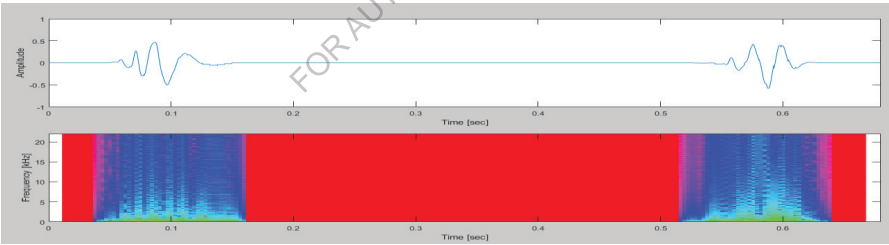

Figure 3.5.7 OGS thresholded heart sound under stationary conditions

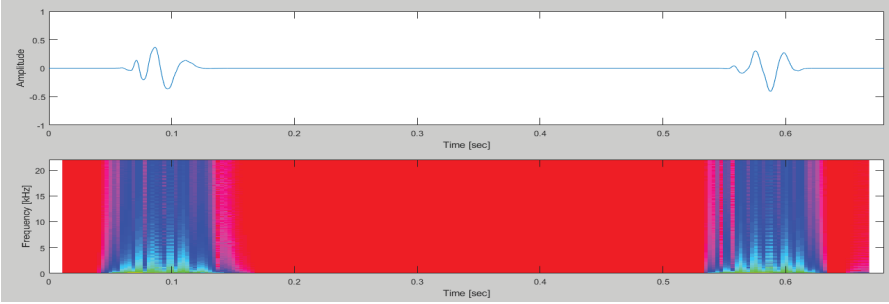

Fig 3.5.8 Block thresholded heart sound under stationary conditions

3.5.3 Observations and Discussion

From Table 3.5.1 it is clear that TFBT method outperformed the OGS and ST methods in terms of SNR, SSNR and MSE. There was good improvement in SNR and SSNR when TFBT method was used against ST method. This could be attributed to the efficient and robust noise removal in TFBT method when compared to some signal attenuation and distortion in ST method. The performance of OGS algorithm was in between that of ST method and TFBT method. This was a clear indication that OGS algorithm suffers less from distortion and attenuation as compared to ST method. However, some kind of residual noise still persists in OGS method that gets removed in TFBT method. Similar results are observed for Abnormal HS as shown in Table 3.5.2. Table 3.5.3 shows the results obtained using application of Time-Frequency methods on Normal PCG signals corrupted by Non stationary noise. Even here too, the TFBT method showed far better performance compared to its companion methods ST and OGS. One important observation that was noted in de-noising procedure adopted here, was that de-noising of non-stationary noise corrupted PCG signals yielded better results compared to stationary noise corrupted PCG signals. TFBT method removed non stationary noise more efficiently than stationary noise with respect to the companion methods. Similar results were obtained for Abnormal PCG signals corrupted with Non stationary noises as shown in Table 3.5.4.

Chapter 4

Segmentation of Phonocardiogram

Section 2.2 gives details about the different Segmentation techniques mentioned in Literature. Of the various segmentation techniques available today, the Shannon energy/ entropy envelope based techniques are the most popular and widely used. We have taken up the popular Shannon energy/entropy based Segmentation techniques namely Homomorphic Filtering (HF) and have evaluated the performance of these algorithms on a very large database of heart sounds hosted in Physionet repository.

In this chapter, we introduce a new method of Segmentation using Bark Spectrogram called the Event Synchronous Heart Sound Segmentation. Two versions of the method are discussed namely, Band pass Filtered Event Synchronous method and Continuous Wavelet Transform (CWT) Filtered Event Synchronous methods. Later in the section we describe in detail as to how cardiac events can be detected using Loudness function extracted from Bark Spectrogram of a heart sound. We compare the performance of these method with other popular method in Literature namely the one that uses Homomorphic Filtering.

4.1 Segmentation of Phonocardiograms using Homomorphic Filtering

Homomorphic filtering based Segmentation of heart sounds are divided into three important stages namely Pre-processing, peak detection using Homomorphic filtering and Heart Sound Detection using detected peaks. The block diagram with five stages are shown below in Fig 4.1.1. The first stage is the selection of Original HS from the database. The second stage is Preprocessing. Peak detection in the third stage follows the Preprocessing stage. The fourth stage is heart sound detection followed by the segmented HS in the final stage.

Figure 4.1.1 Block Diagram of HF method

Preprocessing

The recorded signal was first preprocessed before performing segmentation. PCG signals were then reduced in sample size to 4000 Hz by down-sampling and normalized as per (1).

$$x_{norm}(t) = \frac{x_{4000}(t)}{max(|x_{norm}(t)|)} \quad (4.1)$$

where $x_{4000}(t)$ is the down sampled signal. The range of frequencies of the PCG signal lies in the range from 50 Hz to 700 Hz. Higher frequencies contain noise and are of not much clinical important for analysis and diagnosis. So we have used a 700Hz cut off frequency low pass Chebyshev type I filter. To obtain zero phase distortion, the signal sequence was reversed after filtering and was then run back.

Peak detection using homomorphic filtering:

In this approach we use the structural similarity of HSs to that of the modulated components. The fundamental PCG components are similar to the AM wave. PCG murmurs are similar to the AM and FM waves. In Homomorphic filtering technique we utilize a logarithmic function. This function converts the multiplied time domain signal into added frequency domain signals. The spectrum of PCG has both slow changing part and fast changing part. The fast changing part is eliminated using a LPF.

If $v(n)$ is a PCG signal and $x(n)$, its energy, then energy is given by

$$x(n) = a(n)f(n) \quad (4.2)$$

$a(n)-$ slow changing part $f(n)$ – fast changing part. $a(n)$ consists of fundamental PCG components, while $f(n)$ consists of murmurs.

- Multiplication operation is converted to addition by taking a simple logarithmic function:

$$. z(n) = \log(x(n)) \quad (4.3)$$

which can be expressed as

$$z(n) = \log(a(n)) + \log(f(n)) \quad (4.4)$$

- The logarithms of the two signals now add together. The high frequency component shows rapid variations in time. An appropriate linear low-pass filter L is used to filter the $f(n)$ components:

$$z'(n) = L(\log(x(n))) \quad (4.5)$$

L refers to a 6th order type 1, low pass Chebyshev filter L with a passband of 10Hz to 20Hz. The formula for the Chebyshev filter is given by $|T_n(\omega)|^2 = \frac{1}{1+\varepsilon^2 {C_n}^2(\omega)}$ Where ε = ripple factor, ω = normalized cutoff frequency and ${C_n}^2(\omega) = \cos(n)\, cos^{-1}\omega$ of the 6th order.

The logarithmic function does not affect the separability of the Fourier components of $\boldsymbol{a(n)}$ and $\boldsymbol{f(n)}$. Since $\boldsymbol{L}$ is linear we have:

$$\boldsymbol{z'(n) = L(\log(a(n)) + \log(f(n)))} \tag{4.6}$$

By exponentiation we arrive at:

$$\boldsymbol{\exp(z'(n)) = exp(\log(a(n)) + \log(f(n))) = \exp(\log(a(n))) = a(n);} \tag{4.7}$$

To cut off the murmurs we used 6th order type 1, low pass Chebyshev filter $\boldsymbol{L}$ with a passband of 10Hz to 20Hz. Fig 4.1.2. shows the smooth envelope of the signal obtained using exponentiation operation. Peak of the envelope is found. Those points greater than 1.5% of the peak value were considered.

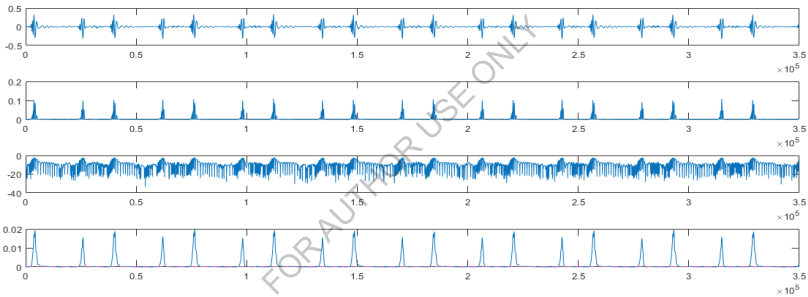

Figure 4.1.2 Segmentation Procedure Using Homomorphic Filtering (x axis –time (s) and y axis- Amplitude (v)

Heart Sound detection using detected Peaks

In this stage we detect peaks and then extract single cardiac cycle of PCG signal. Then we perform peak conditioning using homomorphic filtering which helps in cycle detection process. Peaks other than the S1 and S2 are rejected. Then we find peak width, peak start point, peak end point and distance between peaks are found.

- The mean width of detected peaks is calculated. All peaks with width lesser than 50% of mean peak width is thus rejected.
- The distance between two detected peaks, S1 and S2, cannot be less than 80ms. During inspiration the A2 and P2 are at a distance of 30-80ms from each other [9]. Peak width less than 80ms corresponds to a split S2. Such

peaks are combined into a single peak. The possible width of S1 and S2 is 80-120ms. Greater peak widths are discarded. These peaks were limited to 120ms and peak conditioning is achieved.

4.2 Event Synchronous Heart Sound Segmentation

Most of the Segmentation methods utilize threshold parameter for segmentation. Since it is difficult to set an appropriate threshold for all types of sounds these algorithms suffer from poor segmentation. So we have to look at methods where there is no use of threshold parameter. Event synchronous method utilizes calculation of Loudness function from the spectrogram. We differentiate the Loudness function to obtain the Event detection function. Event detection function helps in peak detection of S1 and S2. Two versions of this method are described in this section namely Low pass filtered Event Synchronous Method and CWT filtered Event Synchronous Method.

4.2.1 Band pass Filtered Event Synchronous Method

Figure 4.2.1 shows the block diagram of the Band Pass Filtered Event Synchronous Segmentation (BPF ESS) procedure for the event synchronous segmentation of cardiac sounds. The method is divided into five steps namely heart sound selection, Spectrogram analysis, bark spectrogram analysis, Smoothened bark spectrogram analysis, Loudness index evaluation, Cardiac event detection and smoothened cardiac events detection and Identification of the sounds S1 and S2. Each of them are discussed one after the other. Fig 4.2.2 shows the waveform plot of a normal heart sound.

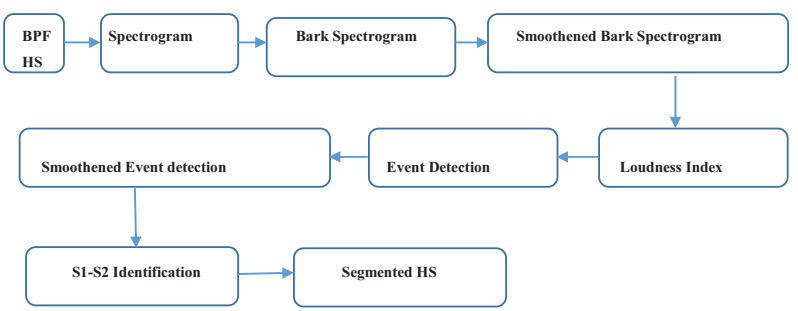

Figure 4.2.1 Block Diagram of LPF-ESS Method

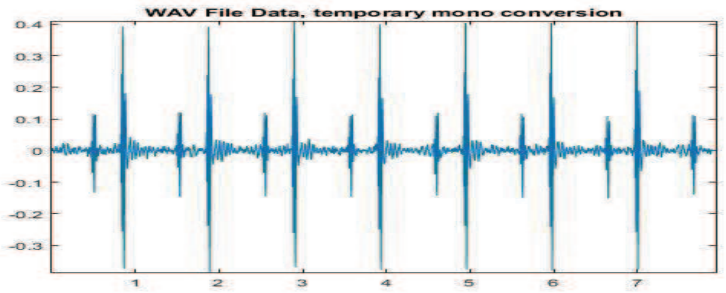

Figure 4.2.2 Signal plot of filtered normal sound (time(s)-x axis and Amplitude(v)-y axis)

Obtaining BPF heart sound signal:

Residual noises and murmurs are the main obstacles in the correct segmentation and identification of cardiac events. So a BPF is usually used for the removal of such noises and murmurs. We have chosen Chebyshev, Type 1 filter with frequency range 4 Hz to 30 Hz for removal of low frequency murmurs. For removing high frequency murmurs and noises Chebyshev, Type 1 filter with frequency range 90 Hz to 500 Hz is chosen. On using these two filters, we obtain a noise free and murmur free heart sound that is inputted to the second stage of the LPF-ESS method.

Spectrogram Analysis:

We obtain the Spectrogram of the cardiac sound signal, at roughly 3ms window. Fig 4.2.3 shows the spectrogram of the normal heart sound. However, splitting up of frequencies in the audible range by means of a spectrogram does not relate to human perception of these sounds. Humans frequency perception of sounds stretch wider with rising frequencies in the sound. Thus, the spectrogram is converted into the better representation. This scale is the Bark Scale. The power spectrum of the sound is given Eq. (4.8) [1]:

$$\boldsymbol{I_i(dB) = 20log_{10}(\frac{I_i}{I_0})} \quad \boldsymbol{, i > 0} \tag{4.8}$$

i is the instant of the power-spectrum intensity I_i. I_0 is the hearing threshold of the PCG. To obtain a reasonable trade-off between dynamic range and resolution, $I_0 = 60$ is chosen. Sound pressure levels below -60dB are completely clipped. The threshold of hearing is dependent on frequency. It is due to the outer and middle ear response of these sounds. The frequency f relates to the Bark scale z(f) [1].

$$\boldsymbol{z(f) = 13\arctan(0.00076f) + 3.5\arctan(\left(\frac{f}{7500)^2}\right))} \tag{4.9}$$

Fig 4.2.4 shows the bark spectrogram of the normal heart sound.

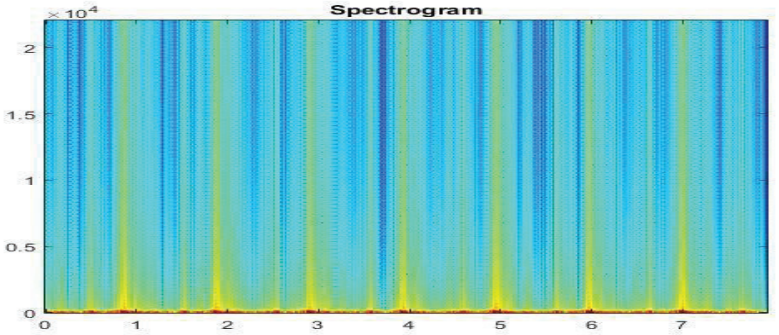

Figure 4.2.3 Spectrogram of the filtered normal sound

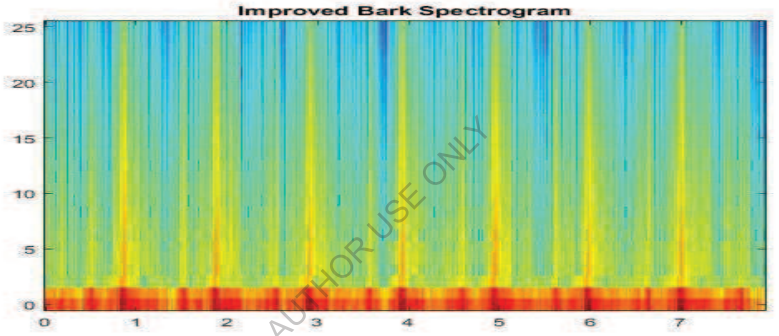

Figure 4.2.4 Bark Spectrogram of the filtered normal sound

Identification of masker and masked cardiac sounds

A Phonocardiogram has got both loud and soft sounds in it. For a very small time difference of less than 150ms it is impossible to hear both soft sounds and loud sounds of same frequency simultaneously. The loud sounds generally mask the soft sounds due to high energy content in them. This is termed as simultaneous masking. There are other two categories of masking namely: pre-masking and post-masking. Pre-masking lasts for only about 20ms duration. In this duration there are inaudible sounds softer than the masker sounds. Pre-masking is not implemented since similar effect is shown by signal-windowing artefacts during smoothing. Post-masking is a ringing type of temporal masking that lasts for about 200ms.

To smoothen the spectrogram of the cardiac sound, the envelope of each frequency band is convolved with a 200ms half-Hanning (raised cosine) window. The bark spectrogram gives audio-visual representation of each beat in those sounds that are inaudible and invisible in a normal spectrogram. Sensation of loudness index is the intensity of the cardiac sound. Fig 4.2.5

shows the bark spectrogram of the filtered heart sound.

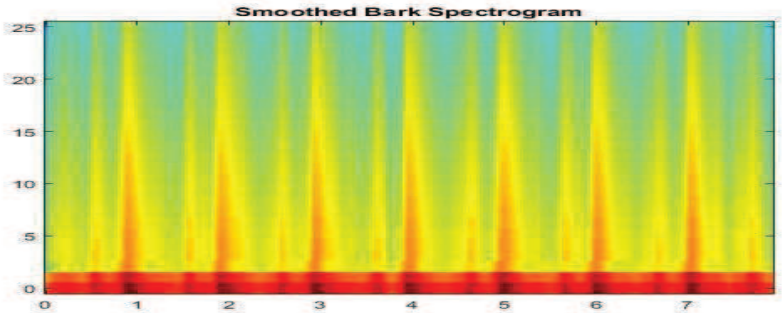

Figure 4.2.5 Smoothed Bark Spectrogram of the filtered normal sound

Valuation of loudness index

The sensation of loudness is obtained by summing the amplitudes of all frequency bands of the sound in its spectrogram:

$$\boldsymbol{L_{dB}(t) = \frac{\sum_{k=1}^{N} E_k(t)}{N}} \quad (4.10A)$$

E_k represents the magnitude of the kth frequency band present in the spectrogram. There is a total of N such bands. Loudness function is a raw and unclear plot of the energy of cardiac sound and shows the intensity of S1 and S2 loudness. Fig 3.3.6 shows the raw loudness function. To smoothen the loudness function, we convolve it with a 300ms Hanning. The smoothened loudness function is given by $\boldsymbol{L_{Sm}(t) = L_{dB}(t) * W(n)}$. (4.10B)

$$\boldsymbol{W(n) = \begin{cases} cos^2\left(\frac{\pi x}{N}\right) & x \geq N/2 \\ 0 & x < N/2 \end{cases}} \quad (4.10C)$$

represents the 300 ms Hanning window, $\boldsymbol{N}$ being the number of samples. Fig 4.2.7 shows the smoothened loudness function.

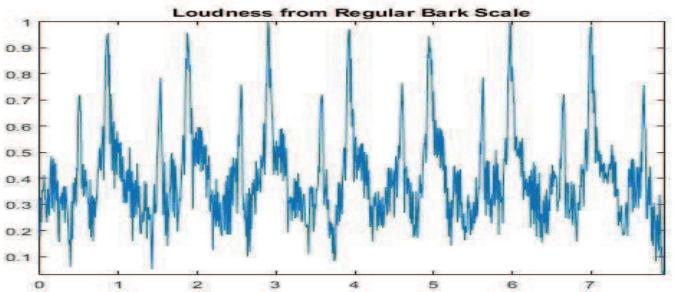

Figure 4.2.6 Loudness Index of the filtered normal sound

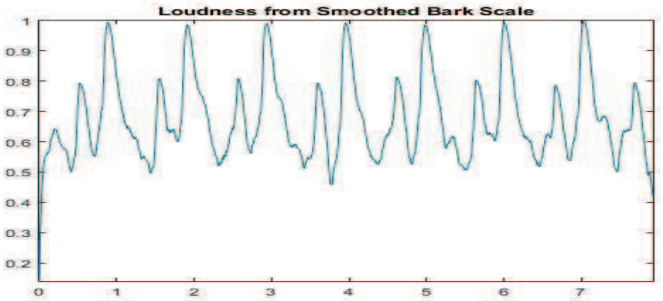

Figure 4.2.7 Smoothed Loudness Index of the filtered normal sound

Detection of Cardiac Events

We estimate the event detection function by calculating the first order difference of each spectral band, and then we find their sum. There are many onset and offset transients in this signal. A single event is represented by a transient within a 50-ms window. This event can be the corresponding cardiac sound S1 or S2 [4]. Fig 4.2.8 shows the raw event detection function. We find the smoothed events in the cardiac signal by convolving the raw event detection signal with a 400ms Hanning window. Fig 4.2.9 shows the smooth event detection function. The next step involved in the identification of cardiac events is the peak-picking stage. The segment of the sound is defined by the onset boundary and the offset boundary. An onset of the sound occurs generally with an increase in the variation of the loudness. An offset occurs with the decrease in the variation of the loudness. The local maximum is the characteristics of the softest onset moment. The local minimum is the characteristics of the softest offset. One can look at the zero crossings from the negative to the positive in the sound signal for maintaining signal continuity.

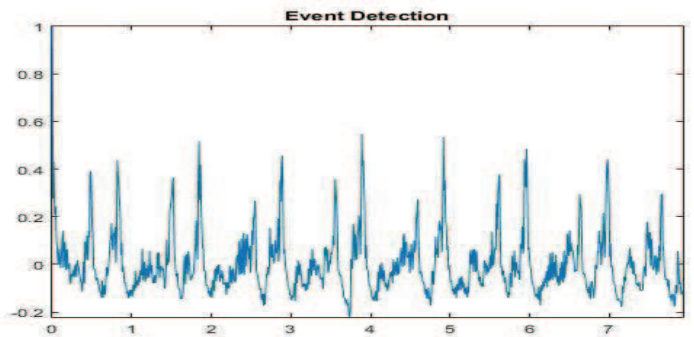

Figure 4.2.8 Event Detection of the filtered normal sound

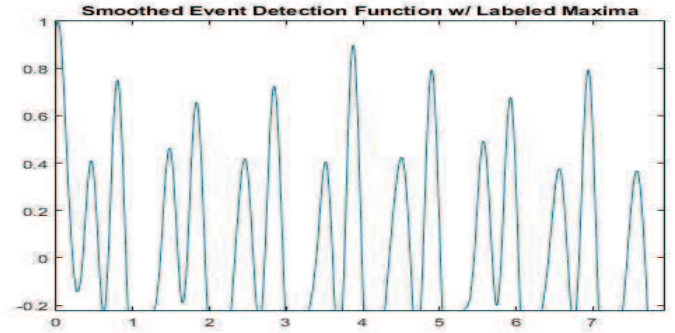

Figure 4.2.9 Smoothed Event Detection of the filtered normal sound

S1-S2 Identification

The S1 and S2 sounds are the important sounds present in the segmented cardiac cycle of the PCG. In each cardiac cycle we identify the initial S1 and S2 sounds (the first sounds) by noting the systole and diastole present in the single cardiac cycle. The diastole is the longest duration. The shortest duration between an S2/S1 and consecutive S1/S2 is the systole. Once the first sounds are identified, the other sounds are the alternating to the first sounds. Fig 4.2.10 shows the ESS method for Abnormal heart sound.

Event Synchronous Segmentation algorithm:

Input: De-noised heart sound

Output: Segmented heart sound

Steps:

1. The spectrogram of the heart sound is evaluated by taking the STFT.
2. The original spectrogram is converted into improved bark spectrogram.
3. The improved bark spectrogram is smoothened by using a hanning window of 50ms duration.
4. The loudness index is obtained by taking the row sum of the smoothened spectrogram.
5. Smoothened loudness index is obtained by windowing the original loudness index by using a hanning window.
6. Event detection function is obtained by differentiating the smoothened loudness index.
7. Smoothened Event detection function is obtained by smoothening the original event detection function by a hanning window.

8	Maxima and minima of the smoothened event detection function gives the location of first heart sound (S1) and the second heart sound (S2). S1 and S2 are identified based on the systole and diastole duration.

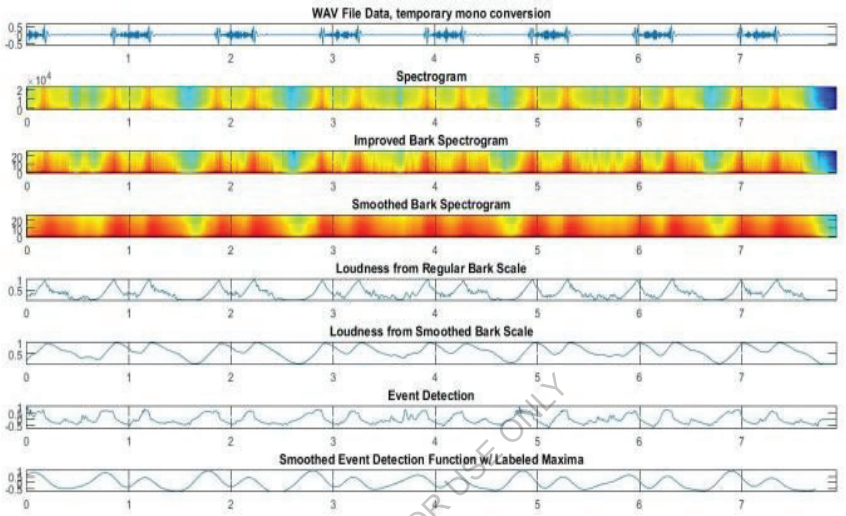

Figure 4.2.10 Event Synchronous Segmentation Procedure for Abnormal heart sound

4.2.2 CWT Filtered Event Synchronous Method

Figure 4.2.11 shows the block diagram of CWT filtered Event Synchronous Segmentation procedure. This method uses a continuous wavelet transform to obtain the noise free and murmur free heart sound. To obtain the CWT filtered heart sound, the selected heart sound is first subjected to CWT. The continuous Wavelet transform coefficients are thresholded in the frequency range 4 Hz to 30 Hz and 90 Hz to 500 Hz. This effectively curtails all the low frequency and high frequency murmurs and noises. Then the inverse CWT is taken to obtain the actual heart sound with murmurs and noises removed. The spectrogram is taken for this heart sound and then converted to bark scale. The loudness indices and event detection functions are evaluated similar to section 4.2.1. Then the smoothened event detection is found and the peaks are identified based on the systole-diastole duration.

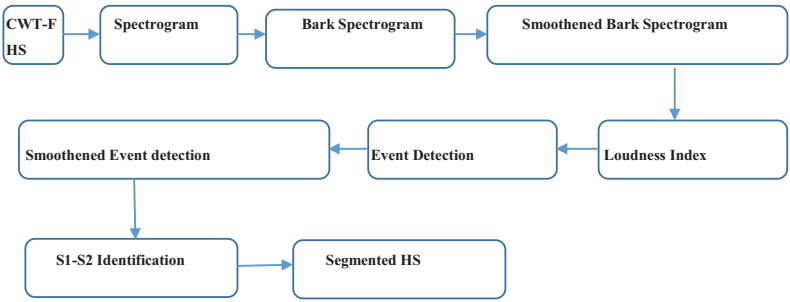

Figure 4.2.11 Block Diagram of CWT-ESS Method

4.2.3 Comparison between BPF-ESS and CWT-ESS methods

It was observed that there was residual noise (low frequency noise) in the frequency range of 4 Hz to 30 Hz even after two stages of preprocessing involving Wavelet Filtering and TFBT approach. Figures 4.2.12-4.3.13 shows how a CWT filter effectively removes the residual noises which are in the frequency range of 4 Hz to 30 Hz. Also Figures 4.2.14-4.2.15 shows how the CWT filter removes the high frequency murmurs in the frequency range 90 Hz to 500 Hz and low frequency murmurs in the frequency range 4 Hz to 30 Hz. Figures 4.2.16-4.2.19 shows a comparison of BPF-ESS and CWT-ESS methods. In Figure 4.2.16, Pane 1 shows the normal HS signal with some low frequency residual noise. Pane 2 shows BPF filtered HS with low frequency noise removed. Pane 3 shows the difference signal of the normal HS with low frequency noise and BPF filtered HS. In Figure 4.2.17, Pane 1 shows the normal HS signal with some low frequency residual noise. Pane 2 shows the CWT filtered HS with low frequency noise removed. Pane 3 shows the difference signal of the normal HS with low frequency noise and CWT filtered HS. A comparison of the larger difference signal in two figures suggests that CWT filter removes more noise than its companion BPF. In Figure 4.2.18, Pane 1 shows the heart sound with systolic murmurs lying in 90 Hz to 500 Hz frequency range. Pane 2 shows BPF filtered signal with murmurs removed. Pane 3 shows the difference signal of the two. In Figure 4.2.19, Pane 1 shows the heart sound signal with systolic murmur. Pane 2 shows the CWT filtered signal with murmurs removed. Pane 3 shows the difference signal of the two. A comparison of the larger difference signal in two figures suggests that CWT filter removes more murmurs than its companion BPF.

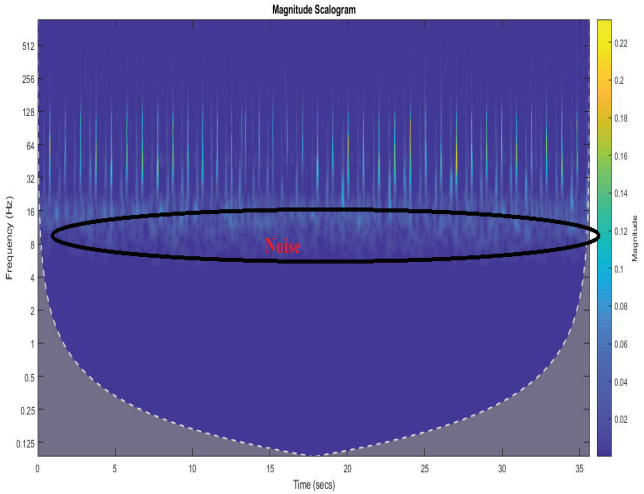

Figure 4.2.12 Original HS containing low frequency noises

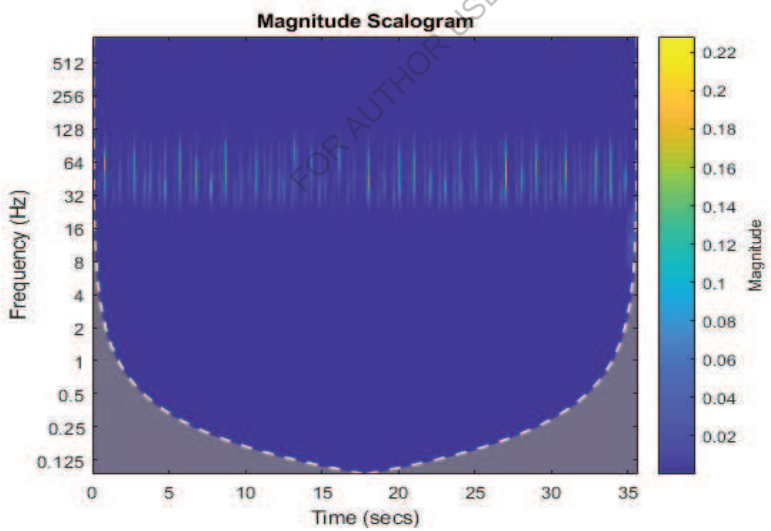

Figure 4.2.13 HS with low frequency noise removed by CWT filter

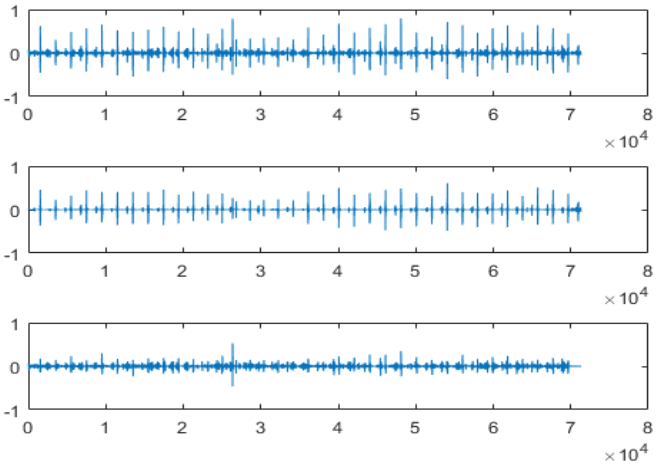

Figure 4.2.14 Original HS (Pane 1), CWT filtered HS (Pane 2), Difference Signal using CWT filter (Pane 3)

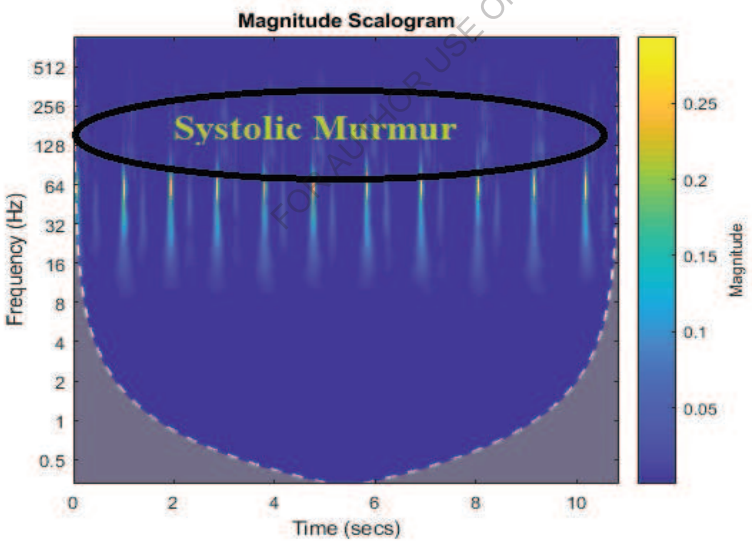

Figure 4.2.15 Original HS with Systolic Murmur

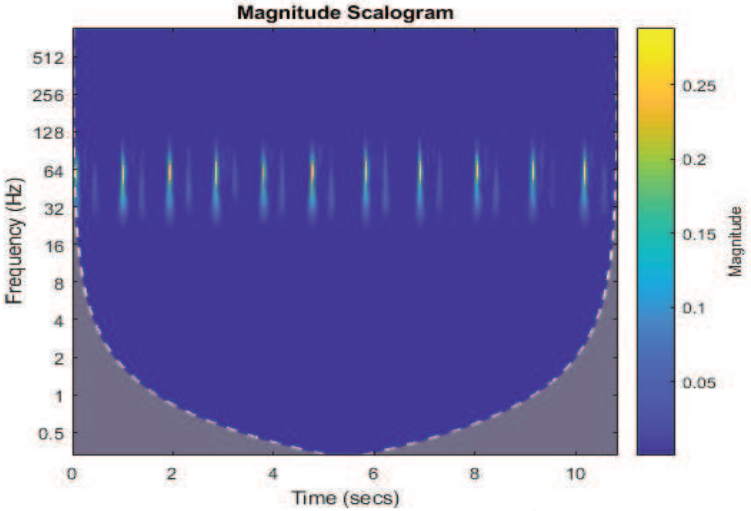

Figure 4.2.16 HS with Systolic Murmur removed using CWT Filter

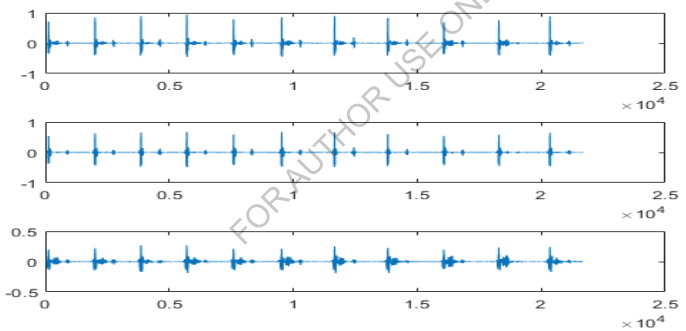

Figure 4.2.17 Original HS (Pane 1), CWT filtered HS (Pane 2), Difference Signal (Pane 3)

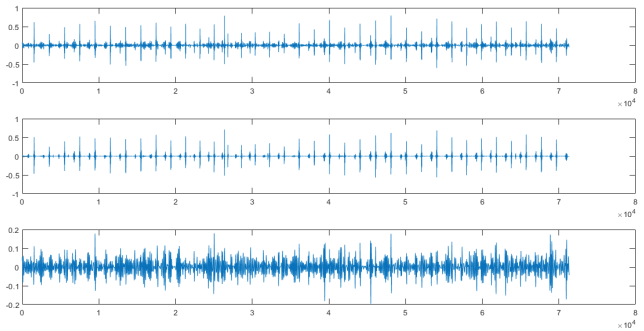

Figure 4.2.18 Original HS (Pane 1), BPF HS (Pane 2), Difference Signal using BPF

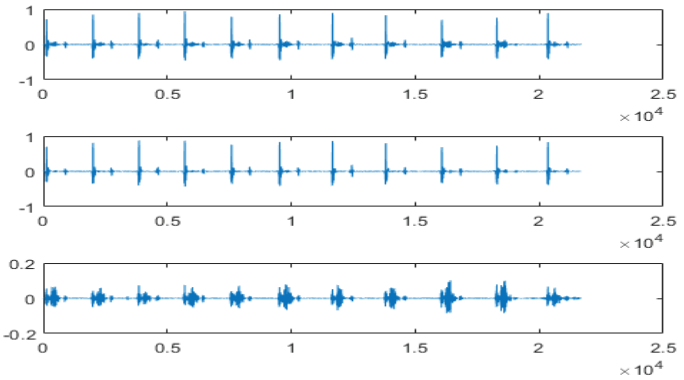

Figure 4.2.19 Original HS (Pane 1), BPF HS with murmur removed (Pane 2), Difference Signal (Pane 3)

4.3 Results of Segmentation of Phonocardiogram

4.3.1 Performance Metrics

To evaluate the performance of the Phonocardiogram Segmentation the following metrics were evaluated:

$$\textbf{Sensitivity } Se\ (\%) = \frac{TP}{TP+FN} \times 100 \quad (4.11)$$

$$\textbf{Positive Predictive Value } PP\ (\%) = \frac{TP}{TP+FP} \times 100 \quad (4.12)$$

$$\textbf{Accuracy } Acc\ (\%) = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad (4.13)$$

$$TP = S1/S2 \text{ Sound Segmented Correctly}$$

$$TN = S1/S2 \text{ Sound in Systole/Diastole}$$

$$FP = S1/S2 \text{ Incorrectly Segmented sound}$$

$$FN = \text{Noise Segmented as } S1/S2 \text{ sound}$$

4.3.2 Results from Segmentation Procedures

The heart sounds are pre-processed using two stages of filtering using wavelet transform (DWT) and Block Threshold technique (TFBT). The preprocessed heart sounds are subjected to three different segmentation techniques namely HF, BPFESS and CWTESS and analyzed. Table 4.3.1 gives the distribution of S1 and S2 sounds in the dataset. The preprocessed sounds

in the dataset were used to calculate the Shannon energy which was then filtered by an elliptical filter, using 0.1 normalized frequency, to remove the unwanted transients and obtain the smooth signal. The peaks of the resulting sounds were counted and noted as the total number of S1 and S2 sounds in the original signal. A total of 121105 S1 and S2 sounds were observed in normal PCG and 36918 S1 and S2 sounds were noted in abnormal PCG giving an overall of 158033 sounds in the dataset. The PCG signals were processed using the first method namely HF. Table 4.3.2 shows the total number of systoles and diastoles present in the HF heart sounds. These regions showed silence or no heart sounds in the PCG. A total of 65862 periods of systole and diastole were observed in the normal PCG, while abnormal PCG showed 10602 such periods resulting in 76468 periods of systole and diastole in the overall dataset. Table 4.3.3 shows the True Positive (TP) and False Negative (FN) of the normal heart sounds segmented using HF method. True positive indicate the correctly segmented heart sounds in the PCG, while the FN indicate the number of noise components wrongly segmented as heart sounds. Even after two stages of filtering there is a small chance that residual noise remains and interferes in the segmentation of PCG signals. In the HF method, for the normal PCG, 42350 TP and 0 FN were reported for S1 sounds. Similarly, 44296 TP and 0 FN were reported for S2 sounds. Thus 86646 TP and 0 FN were reported in the overall normal PCG dataset. Thus 42350 sounds were detected as S1 sounds while 44296 sounds were detected as S2 sounds amounting to a total of 86646 detected heart sounds including the noise components. Table 4.3.3 shows the break-up of S1 and S2 sounds for abnormal PCG. 9469 TP and 0 FN were reported for S1 heart sounds amounting to 9469 detected sounds. Similarly, 9785 TP and 0 FN were reported for S2 heart sounds giving total of 9785 detected sounds. All together 19254 S1 and S2 sounds were detected that included the detected noise components. Table 4.3.5 shows the breakup of abnormal and normal heart sounds that were detected using HF method. 86646 TP, 65862 TN, 34459 FP and 0 FN were reported for normal heart sounds TN corresponds to the silence periods in the PCG namely the systole and the diastole. FP corresponds to the missed sounds. FP is obtained by subtracting the TP from the number of total detected normal heart sounds that did not include the noise components. 19254 TP, 10606 TN, 17674 FP and 0 FN were reported for abnormal heart sounds. Thus an overall 105900 TP, 76468 TN, 52133 FP and 0 FN were found in the entire dataset. Table 4.2.6 lists the performance metrics of all of the cardiac sounds segmented using HF method. The normal sounds showed 100% Sensitivity, 92.06% Positive Predictive Value and 81.6% Accuracy. Sensitivity refers to the correct detection of an S1/S2 normal heart sound when noise components are present. A high value of Sensitivity indicates that there is high chance of correct Segmentation of PCG in the presence of noise. Positive Predictive Value

signifies the number of heart sounds detected when some of them are missed. A high Positive Predictive Value indicates that the rate of PCG segmentation is very high at the cost of very low number of missed sounds. Accuracy refers to the overall rate of correct segmentation of normal PCG in the presence of noise and missed heart sound components. An accurate Segmentation algorithm shows a high degree of accuracy. Abnormal PCG reported 100% Sensitivity, 52.12% Positive Predictive Value and 62.8% accuracy. It was observed that the Positive Predictive Value and Accuracy in abnormal heart sounds was slightly less when compared to that of normal sounds. The main possible cause for this was the presence of residual noise components and murmurs in the abnormal PCG that remained even after filtering. The dataset reported 100% Sensitivity, 67.01% Positive Predictive Value and 77.77% Accuracy overall.

As a second step the PCG is segmented using BPFESS method. Table 3.4.7 shows the breakup of silence period in the normal and abnormal PCG. Normal sounds reported 75760 number of systoles and diastoles. Similarly, Abnormal sounds reported 14076 number of systoles and diastoles. An overall 89836 sounds were reported as the number of systoles and diastoles in the entire dataset. Table 4.3.8 shows the segmentation of normal cardiac sounds using BPFESS method. In Normal PCG S1 sounds reported 55010 TP and 0 FN, while the S2 56486 TP and 0 FN. Thus an overall 111496 TP and 0 FN were reported for the normal dataset. Table 4.3.9 shows the segmentation of Abnormal cardiac sounds using BPFESS method. In Abnormal PCG S1 sounds showed 14176 TP and 0 FN while S2 sounds showed 14350 TP and 0 FN. This resulted in 28536 TP and 0 FN for the abnormal sounds. Table 4.3.10 shows the segmentation of all cardiac sounds. Normal sounds showed 111496 TP, 75760 TN, 9609 FP and 0 FN while abnormal sounds reported 28526 TP, 14076 TN, 8402 FP and 0 FN. Thus an effective 140022 TP, 89836 TN, 18011 FP and 0 FN was reported for the whole dataset. Table 4.3.11 shows the performance metrics of the PCG signals subjected to BPFESS method. Normal sounds showed 100 Sensitivity, 99.06% Positive Predictive Value and an Accuracy of 95.12%. While the abnormal sounds showed 100 Sensitivity, 77.25% Positive Predictive Value and 83.53% Accuracy. Thus in all, 100 Sensitivity, 88.6% Positive Predictive Value and 92.73% Accuracy was observed. It was noted that the Positive Predictive Value and Accuracy in BPFESS method was slightly higher compared to its predecessor HF method. This was because of a careful selection of a BPF to remove murmur in time domain as against the removal of murmur in the logarithmic domain using LPF in HF method.

Table 4.3.12 shows the systole-diastole or silence period in PCG signal where there is no heart

sound. 72208 number of such periods are observed in the PCG signal segmented using CWTESS method. 15477 such periods were noted in case of abnormal PCG signal. 87685 such periods were observed for the entire dataset using CWTESS method. Table 4.3.13 shows the breakup of the segmentation of normal PCG signal segmented using CWTESS method. S1 sounds showed 57060 number of TP and 0 FN. The S2 sounds showed 58410 TP and 0 FN. Thus 115470 such TP and 0 TN were observed for all of the normal sounds in the dataset. In case of Abnormal cardiac sounds shown in Table 4.3.14, 14176 TP and 0 FN were observed for S1 while S2 showed 17223 TP and 0 FN. Thus 34115 TP and 0 FN was reported in case all of the Abnormal sounds. Table 4.3.15 shows the segmentation of all cardiac sounds present at the dataset. The table indicates that Normal sounds result in 115470 TP, 72208 TN, 5635 FP and 0 FN. Also 34115 TP, 15477 TN, 2813 FP and 0 FN is found in case of Abnormal heart sounds. Altogether 149585 TP, 87685 TN, 8448 FP and 0 FN is reported for all sounds. As shown in Table 4.3.16, performance metrics evaluated for all cardiac sounds segmented using CWTESS method reveal the following information. Normal sounds showed 100% Sensitivity, 95.35% Positive Predictive Value and 97.08% Accuracy. On similar lines, Abnormal Cardiac Sounds showed 100% Sensitivity, 92.38% Positive Predictive Value and 94.63%. Thus all cardiac sounds in the dataset showed 100% Sensitivity, 94.65% Positive Predictive Value and 96.56% Accuracy. It was observed that CWTESS method outperformed their previous counterparts namely BPFESS method and HF method in terms of performance metrics.

Sounds	S1	S2	Total
Normal	59534	61571	121105
Abnormal	18484	18444	36928
Entire Dataset	78018	80015	158033

Table 4.3.1 Actual Cardiac sounds

Sounds	Total Sounds in Systole/Diastole
Normal	65862
Abnormal	10606
Entire Dataset	76468

Table 4.3.2 Actual Sounds in Systole/ Diastole Sounds (HF method)

Normal sounds	TP	FN	Total
S1	42350	0	42350
S2	44296	0	44296
Entire cycle	86646	0	86646

Table 4.3.3 Segmentation of Normal Cardiac sounds (HF method)

Abnormal sounds	TP	FN	Total
S1	9469	0	9469
S2	9785	0	9785
Entire cycle	19254	0	19254

Table 4.3.4 Segmentation of Abnormal Cardiac sounds (HF method)

All sounds	TP	TN	FP	FN	Total
Normal	86646	65862	34459	0	186967
Abnormal	19254	10606	17674	0	47534
Entire cycle	105900	76468	52133	0	234501

Table 4.3.5 Segmentation of All Cardiac sounds (HF method)

Sounds	Sensitivity(%)	Positive Predictive Value	Accuracy
Normal	100	71.5	81.6
Abnormal	100	52.14	62.8
All sounds	100	67.01	77.77

Table 4.3.6 Scoring of all cardiac sounds

Sounds	Total Sounds in Systole/Diastole
Normal	75760
Abnormal	14076
Entire Dataset	89836

Table 4.3.7 Actual Sounds in Systole/ Diastole Sounds (BPFESS method)

Normal sounds	TP	FN	Total
S1	55010	0	55010
S2	56486	0	56486
Entire cycle	111496	0	111496

Table 4.3.8 Segmentation of Normal Cardiac sounds (BPFESS method)

Abnormal sounds	TP	FN	Total
S1	14176	0	14176
S2	14350	0	14350
Entire cycle	28526	0	28526

Table 4.3.9 Segmentation of Abnormal Cardiac sounds (BPFESS method)

All sounds	TP	TN	FP	FN	Total
Normal	111496	75760	9609	0	196865
Abnormal	28526	14076	8402	0	51004
Entire cycle	140022	89836	18011	0	247869

Table 4.3.10 Segmentation of All Cardiac sounds (BPFESS method)

Sounds	Sensitivity(%)	Positive Predictive Value (%))	Accuracy(%)
Normal	100	92.06	95.12
Abnormal	100	77.25	83.53
All sounds	100	88.60	92.73

Table 4.3.11 Scoring of All Cardiac sounds

Sounds	Total Sounds in Systole/Diastole
Normal	72208
Abnormal	15477
Entire Dataset	87685

Table 4.3.12 Actual Sounds in Systole/ Diastole Sounds (CWTESS method)

Normal sounds	TP	FN	Total
S1	57060	0	57060
S2	58410	0	58410
Entire cycle	115470	0	115470

Table 4.3.13 Segmentation of Normal Cardiac sounds (CWTESS method)

Abnormal sounds	TP	FN	Total
S1	16892	0	16892
S2	17223	0	17223
Entire cycle	34115	0	34115

Table 4.3.14 Segmentation of Abnormal Cardiac sounds (CWTESS method)

All sounds	TP	TN	FP	FN	Total
Normal	115470	72208	5635	0	193313
Abnormal	34115	15477	2813	0	52405
Entire cycle	149585	87685	8448	0	245718

Table 4.3.15 Segmentation of All Cardiac sounds (CWTESS method)

Sounds	Sensitivity(%)	Positive Predictive Value (%)	Accuracy(%)
Normal	100	95.35	97.08
Abnormal	100	92.38	94.63
All sounds	100	94.65	96.56

Table 4.3.16 Scoring of all cardiac sounds

4.3.3 Observations and Discussion

From the statistics of the above mentioned tables it is clear that three aforementioned methods HF, BPFESS and CWTESS methods segmented the PCG signals, both normal and Abnormal,

with good accuracy. However, side by side comparison indicate CWTESS method showed better accuracy compared to its companion BPFESS and HF methods. It was observed that CWTESS method outperformed their previous counterparts namely BPFESS method and HF method in terms of performance metrics. It is due to the fact that CWTESS method was able to locate the cardiac events more precisely by isolating the murmurs to the maximum level. Another notable observation was that all three methods showed 100% Sensitivity. This is a clear indication of the fact that the preprocessing stage involving Wavelet Filtering followed by TFBT method removed most of the noises and the segmentation algorithms were immune to the left over bare minimal residual noises.

Chapter 5
Classification of Phonocardiograms

The automatic classification of pathology in heart sounds has been described in the literature for over 50 years. The first work on classification were reported as early as 1963 when Gerbarg et al [101] used a threshold based method to identify rheumatic heart disease in children. We have discussed the various heart sound classification methods mentioned in literature, namely, (1) Artificial Neural Network (ANN) - based classification; (2) Support Vector Machines (SVM) - based classification; (3) Hidden Markov Model (HMM) - based classification and (4) Cluster - based classification. Based on the literature survey it is seen that the results of this classification techniques are mainly dependent on the features used for classification.

In this chapter, classification techniques namely 1) FCM 2) K-Means 3) GMM are evaluated for classification of heart sounds using the mean and standard deviation of the loudness index extracted from the spectrogram of the cardiac cycle segmented using CWT-ESS method as features for classifying the segmented PCG.

5.1 Use of Loudness features for Classification

Literature reveals that there exists silence period in the normal heart sounds namely systole and diastole. This portion of the heart sound shows minimal intensity. Cardiac auscultation reveals many high pitch sounds which might have occurred due to valve disorders such as stenosis and regurgitation. The loudness (grade 1-grade 4 sounds) of these sounds are lesser compared to the first and second heart sounds. For grade 5 and grade 6 abnormal sounds, the loudness is much more than the first and second heart sounds. This shows that more of heart sound energy is more concentrated in the systoles and diastoles of abnormal sounds than the normal ones. Thus loudness index can be used a measure to possibly identify and distinguish heart sounds. Mean and Standard deviation values of the heart sound loudness index are found and are used as the two features for the classification of these sounds. Mean and standard deviation loudness indices vary over a broader range for abnormal sounds due to the presence of murmurs while they vary over a shorter range for normal sounds.

5.2 Extraction of Loudness features

The segmented heart sounds were used to extract the single cardiac cycle that included a S1 sound, a systole, a S2 sound and a diastole. The spectrogram of this cardiac cycle was obtained

by selecting a window size of 256 samples and hop size of 128 samples (half of window size). Hanning window was selected for this purpose. The row sum of the Spectrogram coefficients was found and the loudness function was obtained. For every heart sound cycle, the mean and standard deviation of loudness indices were found from the loudness function. These two values were considered as loudness specific features that was later utilized for classification.

5.3 Classification of heart sounds using GMM

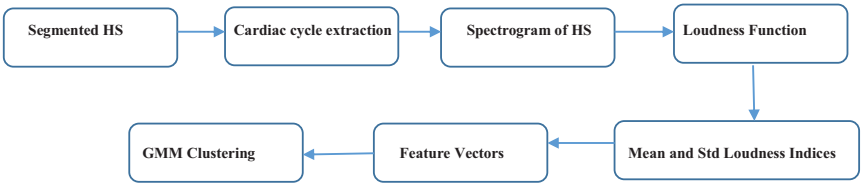

Figure 5.2.1 Block Diagram of GMM Clustering

Literature survey reveals many classification methodologies used for heart sound classification. In these methods, Euclidean distance is the most commonly used distance metric. When Euclidean metric is used the largest scaled feature gets dominated. Solution to this problem could be normalization of the continuous features to a common range of variance. Usage of Euclidean metric distorts the linear correlation that exists in the features. Euclidean distance results in hyper- spheroid clusters. This problem can be solved using Mahalanobis distance. Mahalanobis distance used in GMM results in hyper-ellipsoidal clusters. The database consists of two sound types; both of which is assumed to be generated by the two different Gaussian processes. The probability distribution function (pdf) is given by

$$P\,(x_n|\lambda) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}}\, exp(-\frac{1}{2}(x-\bar{x})^T\Sigma^{-1}(x-\bar{x}) \qquad (5.1)$$

Where d is the dimensionality. ML estimators of μ and ∑ are computed by

$$\bar{x} = \frac{1}{N}(\sum_n x_n) \qquad \sum = \frac{1}{N}(\sum_n (x-\bar{x})(x-\bar{x})^T \qquad (5.2)$$

The objective function is formed given by Eq. (5.1), by summing the class conditional density over all the classes for a feature in the feature vector; and again taking the product for all the features, assuming the features are linearly independent. Such likelihood based objective function for optimization is maximized by EM [19] algorithm, which happens to be a nonlinear optimization method. It optimizes the log likelihood over the entire feature space, including both observed data and hidden information embedded in the data. There are two steps in the EM algorithm namely Expectation step (E-step) and Maximization step (M-step). In E-step the posterior density based on conditional density using Bayes rule [20] is computed. In the M-step

the initial model is replaced with a new model which is a better one to represent the features such that the log- likelihood is more than that of the previous iteration. The iterations are repeated until the new estimate will give same model and there will not be any improvement in the model. The algorithm is briefly described here:

Input: A set of N feature vectors, $E = \{x_1, x_2, \ldots x_N\}$ model structure $\Lambda = \{\mu_k, \Sigma_k, \alpha_k\}$, $k = 1 \ldots K$, where μ's and ∑'s are parameters for the Gaussian models and αs' are prior parameters subject to $\alpha_k \geq 0, \forall k \; and \; \Sigma_k \alpha_k = 1$.

Output: Trained model parameters Λ that maximizes the data likelihood

$$(\boldsymbol{E|\Lambda}) = \prod_n \sum_k (\boldsymbol{x_n|\lambda_k}) \tag{5.3}$$

And a partition of data vectors given by the cluster identity vector $Y = \{y_{1\ldots}, y_N\}$, $y_n \in \{1\ldots, K\}$.

Steps:

1. (Initialization) Initialize the model parameters Λ.
2. (E-Step) The posterior probability of model k , given a data vector x_n and current model parameters Λ , is estimated as

$$\boldsymbol{P(k|x_n, \Lambda)} = \frac{\alpha_k p(x_k|\lambda_k)}{\sum_j \alpha_j p(x_n|\lambda_k)} \tag{5.4}$$

   where the pdf p(e|λ) is given in (5.4)
3. (M-Step) The ML re-estimation of model parameters Λ is given by

$$\boldsymbol{\mu_k^{(new)}} = \frac{\sum_n P(\boldsymbol{k|x_n, \Lambda})x_n}{\sum_n P(\boldsymbol{k|x_n, \Lambda})} \tag{5.5}$$

$$\boldsymbol{\Sigma_k^{(new)}} = \frac{\sum_n P(\boldsymbol{k|x_n, \Lambda})(x_n - \bar{x}_n)(x_n - \bar{x}_n)^T}{\sum_n P(\boldsymbol{k|x_n, \Lambda})} \tag{5.6}$$

$$\boldsymbol{\alpha_k^{(new)}} = \frac{1}{N} \sum_n \boldsymbol{P(k|x_n, \Lambda)} \tag{5.7}$$

4. (Stop) if P(E| Λ) converges; otherwise go back to Step 2;
5. For each data vector x_n, set

$$\boldsymbol{y_n} = \underset{k}{\boldsymbol{arg\,max}}(\boldsymbol{\alpha_k p(x_n|\lambda_k)}) \tag{5.8}$$

Figure 5.2.1 depicts the block diagram of a GMM clustering technique for PCG signals. As a first stage the heart sounds segmented using CWT-ESS technique is used for clustering. In the second stage the single cardiac cycle is extracted from the segmented heart sound. In the third stage spectrogram of the heart sound is obtained. Loudness function is the fourth stage in the clustering process. In the fifth stage, maximum and minimum loudness indices are extracted

from the Loudness function. As a sixth stage feature vectors are created from the loudness indices. In the final stage, these feature vectors are clustered by GMM.

5.4 K Means Clustering of heart sounds

K-means (MacQueen, 1967) is the most simplest unsupervised learning algorithms. To solve the problem, we first fixing the number of clusters k and then, k centroids one for each cluster. The centroids are placed in such a way that different location causes different result. Each point in the dataset belongs to a particular nearest centroid. As a first step, we compute the first group. Then we re-calculate k new centroids as centers of the clusters obtained from the previous step. We observe that the k centroids change their location step by step until no more changes are done.

Finally, using this algorithm we aim at minimizing an *objective function*, in this case a squared error function. The objective function

$$\boldsymbol{J} = \sum_{\boldsymbol{j=1}}^{\boldsymbol{k}} \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \left\| \boldsymbol{x}_{\boldsymbol{i}}^{(\boldsymbol{j})} - \boldsymbol{c}_{\boldsymbol{j}} \right\|^{\boldsymbol{2}} \tag{5.9}$$

Where $\left\| x_i^{(j)} - c_j \right\|^2$ is a chosen distance measure between a data point $x_i^{(j)}$ and the cluster centre c_j , is an indicator of the distance of the *n* data points from their respective cluster centers.

The algorithm is composed of the following steps:

1. *The K data-points (initial group centroids) to be placed into the space*
2. *Assigning of each of the object to the group which has the closest centroid.*
3. *Finding the positions of the K centroids after objects assignment.*
4. *Repeat the Steps 2 and 3, until the centroids are static.*

5.5 Fuzzy C Means Clustering of heart sounds

Fuzzy c-means (FCM) is a type of unsupervised clustering method in which we make the data to fit to two or more number of clusters. In 1973, Dunn first developed this method. This method was improved later by Bezdek in 1981. This method is popular with pattern recognition applications. It is based on minimizing the following objective function:

$$\boldsymbol{J_m = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}^{m}\left\|x_i - c_j\right\|^2, 1 \le m < \infty} \quad (5.10)$$

where m is any real number greater than 1, u_{ij}^{m} is the degree of membership of x_i in the cluster j, x_i is the ith of d-dimensional measured data, c_j is the d-dimension centre of the cluster, and $\|*\|$ is any norm expressing the similarity between any measured data and the centre. Fuzzy partitioning is carried out through an iterative optimization of the objective function shown above, with the update of membership u_{ij}^{m} and the cluster centres c_j by:

$$\boldsymbol{u_{ij} = \frac{1}{\sum_{k=1}^{C}\left(\frac{\left\|x_i - c_j\right\|}{\left\|x_i - c_k\right\|}\right)^{\frac{2}{m-1}}}} \quad (5.11)$$

$$\boldsymbol{c_j = \frac{\sum_{i=1}^{N} u_{ij}^{m} \cdot x_i}{\sum_{i=1}^{N} u_{ij}^{m}}} \quad (5.12)$$

This iteration will stop when

$$\boldsymbol{max_{ij}\left\{\left|u_{ij}^{(k+1)} - u_{ij}^{(k)}\right| < \varepsilon,\right.} \quad (5.13)$$

where ε is a termination criterion between 0 and 1, whereas k are the iteration steps. This procedure converges to a local minimum or a saddle point of J_m.

The algorithm is composed of the following steps:

1. *Initialize* $U = [u_{ij}]$ *matrix,* $U^{(0)}$
2. *At k-step: calculate the centres vectors* $C^{(k)} = [c_j]$ *with* $U^{(k)}$

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^{m} \cdot x_i}{\sum_{i=1}^{N} u_{ij}^{m}}$$

3. *Update* $U^{(k)}$, $U^{(k+1)}$

$$u_{ij} = \frac{1}{\sum_{k=1}^{C}\left(\frac{\left\|x_i - c_j\right\|}{\left\|x_i - c_k\right\|}\right)^{\frac{2}{m-1}}}$$

4. *If* $max_{ij}\left\{\left|u_{ij}^{(k+1)} - u_{ij}^{(k)}\right| < \varepsilon\right.$ *then STOP; otherwise return to step 2*

5.6 Comparison of Classifier performance

5.6.1 Performance Metrics

Following performance metrics were used for evaluation of heart sound Classification.

$$\textbf{\textit{Sensitivity Se}}\ (\%) = \frac{TP}{TP+FN} \times 100 \tag{5.14}$$

$$\textbf{\textit{Specificity Sp}}\ (\%) = \frac{TN}{TN+FP} \times 100 \tag{5.15}$$

$$\textbf{\textit{Accuracy Acc}}\ (\%) = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \tag{5.16}$$

$$TP = Normal\ Sound\ classified\ as\ Normal$$

$$TN = Abnormal\ Sound\ classified\ as\ Abnormal$$

$$FP = Normal\ Sound\ classified\ as\ Abnormal$$

$$FN = Abnormal\ Sound\ classified\ as\ Normal$$

5.6.2 Results from Experiments

The heart sounds from the Physionet database are initially pre-processed by filtered using Wavelet and TFBT techniques. These sounds are then processed using the CWT-ESS procedure. This results in cycles of heart beat of approximately 3s duration. Two datasets are formed consisting of abnormal and normal heart sounds apiece. In each of the datasets two groupings are done one for training set and the other for test set. 50% of the dataset sounds make up the training set while the remaining 50% sounds make up the test set. The sounds are then clustered using the GMM classifier. Fig 5.6.1 shows the segmented heart sound with segmentation boundaries. Fig 5.6.2 shows the normal heart sound with its loudness index. Fig 5.6.3 shows the abnormal heart sound with loudness features. The abnormal heart sound and normal heart sound show different types of peaks. Fig 5.6.4 shows the log likelihood function plot for the sounds. The function is a negative function and it increases steadily and then stabilizes to a constant value for a total of 101 iterations. Fig 5.6.5 shows the 3D-Scatter plot of all sounds. Two features were used the mean loudness along x axis and standard deviation of loudness along y axis. Fig 5.6.6 shows the Expectation-Maximization 3D Scatter plot of all sounds. The points in blue are the abnormal sounds with broader loudness values in terms of mean and standard deviation, while the sounds in red are the normal sounds with shorter range

loudness values. The centroids are also shown for two different clusters.

A total of 3239 sounds of type abnormal and normal were used for training and testing purpose. Three classifiers namely K Means, FCM and GMM were used.

Table 5.6.1 shows K Means Clustering in which out of 2564 normal sounds all of them were TP and 5 were FN. However, there were 660 abnormal sounds that were classified as FP and 10 sounds were TN. This led to a sensitivity of 99.61%, Specificity of 99.25% and an accuracy of 99.53%. The cause for high performance metrics for classification can be attributed to strong de-noising algorithm followed by robust segmentation to identify cardiac cycles consisting of at least one S1, one S2, a systole and a diastole.

Similarly, Table 5.6.2 shows FCM Clustering in which out of 2569 normal sounds all of them were TP and 3 were FN. At the same time, there were 662 abnormal sounds that were classified as FP and 5 sounds were TN. This led to a sensitivity of 99.80%, Specificity of 99.55% and an accuracy of 99.75%.

Table 5.6.3 shows GMM Clustering in which out of 2572 normal sounds all of them were TP and 1 were FN. There were 664 abnormal sounds that were classified as FP and 2 sounds were TN This led to a sensitivity of 99.92%, Specificity of 99.84% and an accuracy of 99.91%.

Results from Table 5.6.4 showed that with mean and standard deviation of loudness as features GMM, FCM and K-Means algorithms showed superior performance in terms of Sensitivity, Specificity and Accuracy. The residual noise was low enough such that it hardly interfered in classification process thus producing high accuracy in each of the above mentioned classifiers.

Class	**Actual Class**	
Predicted Class	2564	10
	5	660

Table 5.6.1 Confusion Matrix of all sounds using K means

Class	**Actual Class**	
Predicted Class	2569	5
	3	662

Table 5.6.2 Confusion Matrix of all sounds using FCM

Class	Actual Class	
Predicted Class	2572	2
	1	664

Table 5.6.3 Confusion Matrix of all sounds using GMM

Methods	Se(%)	Sp(%)	Acc(%)
K Means	99.61	99.25	99.53
FCM	99.80	99.55	99.75
GMM	99.92	99.84	99.91

Table 5.6.4 Scoring of all sounds using Different Methods

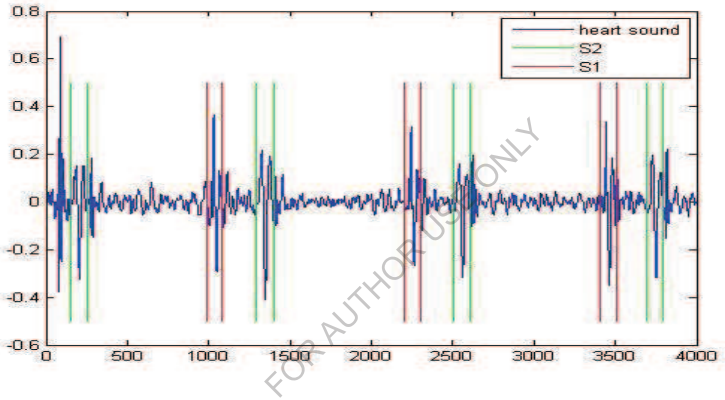

Figure 5.6.1 Segmented heart sound (blue) and Segmentation boundaries (S1 red, S2 green).

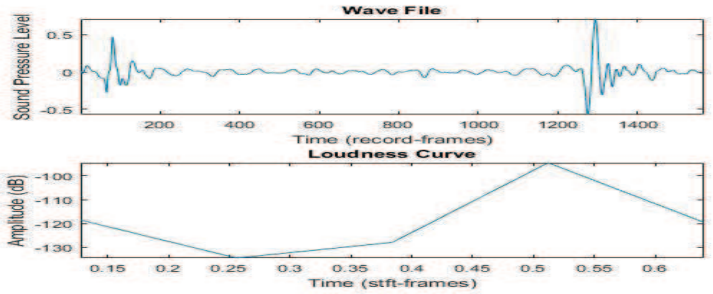

Figure 5.6.2 Normal heart sound and the loudness curve

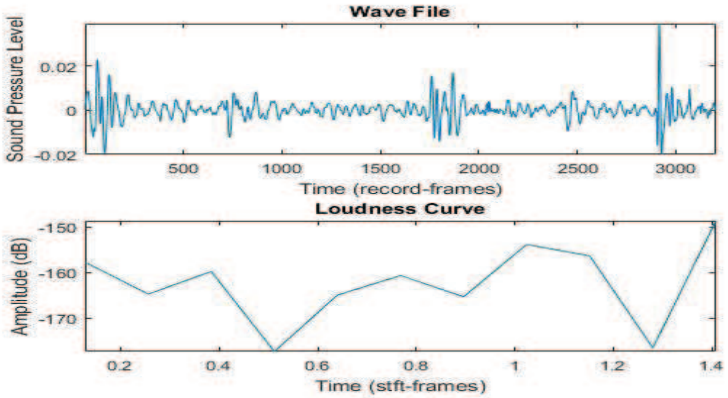

Figure 5.6.3 Abnormal Heart sound with loudness curve

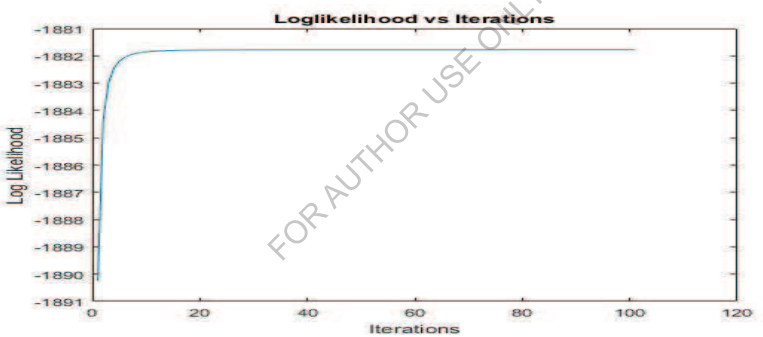

Figure 5.6.4 Log likelihood variations versus iterations

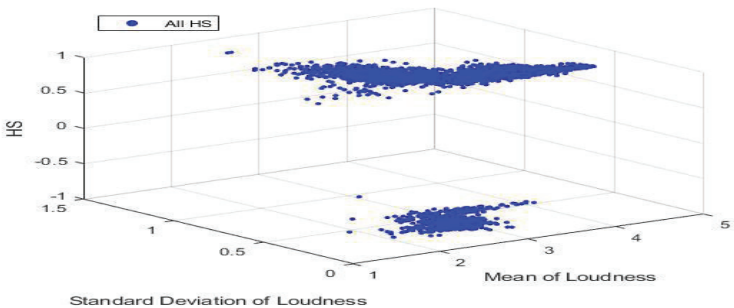

Figure 5.6.5 Scatter plot of all observations (mean loudness index-x axis and std of loudness index y axis)

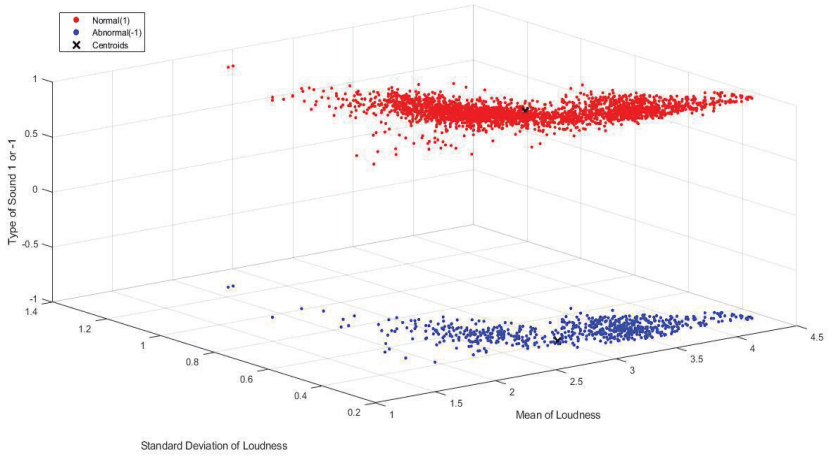

Figure 5.6.6 EM plot of all observations (mean loudness index-x axis and std of loudness index y axis)

5.6.3 Observations and Discussion

From the above tables and plot it is clear that high classification accuracy is obtained using the above mentioned classifiers. It was observed that GMM outperformed K Means and FCM classifiers in terms of Sensitivity, Specificity and Accuracy. The reason behind the high classification accuracy can be attributed to the selection of most appropriate features vectors. It was observed that selection of mean and standard deviation of loudness features as feature vectors yielded a high degree of classification and an even high degree of classification with respect to GMM in particular. Obtaining high degree of Classification has been observed in other works in literature. However, we have tried to do this with minimum number of feature vectors (=2). We have obtained a high degree of classification that is unusual with two features (minimum) and often noticed with greater number and types of feature vectors (more than two).

Chapter 6

Conclusion and Future scope

In this work Phonocardiograms available in the form of database have been utilized. The MIT heart sound database with over 3000 sounds has been used. The sounds are noisy but contain vital information regarding the status of the heart. Additional noise has been added for the purpose of analysis namely filtering. Filtering has been implemented by estimating the noise using activity detection method/ IMCRA. Activity detection estimates noise in stationary environments by detecting heart sound and non-heart sound regions in the signal. The noise is usually present in the silent non-heart sound region in the signal. For non- stationary noise estimation IMCRA is used. IMCRA method estimates the noise using Time recursive averaging technique. The estimated noise is reduced to minimal residual value using time frequency block threshold method. This method outperforms soft threshold and Overlapping group shrinkage method in terms of SNR. The de-noised sound is segmented using the CWT-ESS method. This method gives a high accuracy of 96.56% and is comparable with the state of the art segmentation techniques in terms of results. To check the robust of CWT-ESS method, it is compared with HF method and BPF-ESS methods. Cardiac cycles were extracted from the segmented heart sounds. The loudness features namely mean and standard deviation of loudness was extracted from the single heart sound cycle using Spectrogram. The features from all heart sounds are then classified using GMM by clustering the sounds into normal and pathological sounds. The classification accuracy is as high as 99.91% and is comparable to the state of the art classifiers mentioned in literature. To check the robustness of GMM classifier for loudness features, the clustering was implemented with K-Means and FCM classifier and compared. It was observed that, in terms of performance metrics namely Sensitivity, Specificity and Accuracy, with loudness as features GMM showed robust performance against the other two classifiers.

Key Contributions of This Work

Performances of existing signal processing techniques are investigated for heart sound de-noising, delineation and classification under different kinds of PCG signal patterns and noise conditions.

• Different signal processing approaches such as Time Frequency Block Thresholding (TFBT), Overlapping Group Shrinkage (OGS) and Soft thresholding (ST) are studied for suppressing the background noises.

• Performance of different Segmentation techniques like Homomorphic filtering (HF) and Band pass filtered event synchronous Segmentation (BPF-ESS) is studied using different kinds of PCG signals including large-amplitude heart sounds (S1, and S2) and small-amplitude heart sounds (S3 and S4) and other heart murmurs.

• A bark spectrogram based Continuous Wavelet transform based Event Synchronous Segmentation (CWT-ESS) waveform delineation approach is proposed for automatically determining the boundaries of heart sounds. No noise threshold was used in this procedure.

• In this study, Loudness Indices are extracted from the Spectrogram of the cardiac cycle extracted from heart sound segmented using ESS method.

• In this study, the mean and standard deviation of Loudness features obtained from Loudness Indices are clustered using GMM.

Future Directions

For real-time applications of a unified PCG signal delineation framework for both Computer Aided Diagnosis and Electronic Stethoscope, the future directions of the thesis work are summarized as follows.

• Parameter extraction accuracy of the PCG signal delineation methods namely BPF-ESS, CWT-ESS and HF will be evaluated using different kinds of pathological and non-pathological PCG signals

• Robustness of the unified PCG signal delineation framework will be studied under different kinds of physiological interferences such as lung and bowel sounds and external noise sources such as motion artefacts, speech, and instrument noise.

• Feasibility of real-time implementation of the delineation framework will be studied on the embedded processors for the development of electronic stethoscope having the capability of automatically extracting the clinical indices form the heart sound and murmur signals and time-frequency visualization of different kinds of heart sound and murmur patterns.

• In this thesis the proposed methods namely TFBT, ESS methods have been implemented using the largest available Physionet database and compared with the other time frequency methods like OGS and ST with good results. Of the latest, many signal processing methods like flexible analytic wavelet transform, Tunable-Q wavelet transform, empirical wavelet transform, variational mode decomposition and Fourier Bessel series expansion based empirical wavelet transform have been developed. These methods have used proprietary

database with less number of sounds with good results. We have compared these methods for the proprietary database with our methods with physionet database as mentioned in Table 2.6.1 and Table 2.6.2. As a future work these methods will be compared with methods under the scope of the thesis using Physionet database.

REFERENCES

[1] S. Barma, B.-W. Chen, W. Ji, F. Jiang, and J.-Fa Wang, "Measurement of duration, energy of instantaneous frequencies, and splits of subcomponents of the second heart sound," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 7, pp. 1958-1967, July 2015.

[2] B. Karnath and W. Thornton, "Auscultation of the heart," *Hospital Physician,* vol. 38, no. 9, pp. 39-43, 2002.

[3] S. Barma, B.-W. Chen, K. L. Man, and J.-F. Wang, "Quantitative measurement of split of the second heart sound (S2)," *IEEE/ACM Trans. Comput. Biology and Bioinformatics*, vol. 12, no. 4, pp. 851-860, July-August 2015.

[4] Springer et al. "Logistic regression-HSMM-based heart sound segmentation," *IEEE Trans. Biomed. Eng.* vol. 63, no. 4, pp. 822-32, April 2016.

[5] Tien-En Chen et al., "S1 and S2 heart sound recognition using deep neural networks," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 2, pp. 372-380, February 2017.

[6] C. Papadaniil and L. Hadjileontiadis, "Efficient heart sound segmentation and extraction using ensemble empirical mode decomposition and kurtosis features," *IEEE J. Biomed. and Health Informatics*, vol. 18, no.4, pp. 1138-1152, July 2014.

[7] J. Herzig, A. Bickel, A. Eitan, and N. Intrator, "Monitoring cardiac stress using features extracted from S1 heart sounds," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 4, pp. 1169-1178, April 2015.

[8] X.-Y. Zhang, E. MacPherson, and Y.-T. Zhang, "Relations between the timing of the second heart sound and aortic blood pressure," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 4, pp. 1291-1297, April 2008.

[9] J. Xu, L.-G. Durand, and P. Pibarot, "Nonlinear transient chirp signal modeling of the aortic and pulmonary components of the second heart sound," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 10, pp. 1328-1335, October 2000.

[10] S. Xiao et al., "A relative value method for measuring and evaluating cardiac reserve," *Biomed. Eng. Online*, vol. 6, pp. 1-6, December 2002.

[11] C. Kwak and O.W. Kwon, "Cardiac disorder classification by heart sound signals using murmur likelihood and hidden Markov model state likelihood," *IET Signal Processing*, vol. 6, no. 4, pp. 326-334, June 2012.

[12] S. Sun et al., "Segmentation-based heart sound feature extraction combined with classifier models for a VSD diagnosis system," *Expert Sys. Appl.*, vol. 41, no. 4, pp. 1769-1780, March 2014.

[13] Felner JM. The Second Heart Sound. In: Walker HK, Hall WD, Hurst JW, editors. Clinical Methods: The History, Physical, and Laboratory Examinations. 3rd edition. Boston: Butterworths; 1990. Chapter 23. Available from: http://www.ncbi.nlm.nih.gov/books/NBK341/

[14] Rangaraj M. Rangayyan, "Biomedical Signal Analysis: A Case-Study Approach," *Wiley-IEEE Press*, 2001.

[15] Abbas K. Abbas and Rasha Bassam, "Phonocardiography Signal Processing," *Morgan & Claypool Publishers*, 2009.

[16] A. Ravin, "Auscultation of the Heart," 3rd edition, Chicago, Year Book Medical Publishers, 1977.

[17] V. Nivitha Varghees and K. I. Ramachandran, "Heart murmur detection and classification using wavelet transform and Hilbert phase envelope," *IEEE 21st National Conference on Communications*, pp. 1-6, February 2015.

[18] R. M. Rangayyan and R. J. Lehner, "Phonocardiogram Signal Analysis: A Review", *CRC Critical Reviews in Biomedical Engineering*, vo1. 15, no. 3, pp. 211-236, February 1987.

[19] D. Gill et al., "Detection and identification of heart sounds using homomorphic envelogram and self-organizing probabilistic model," in *Proc. Comput. Cardiol.*, pp. 957-960, September 2005.

[20] P. Wang, Y. Kim, L. H. Ling, and C. B. Soh, "First heart sound detection for phonocardiogram segmentation," in *Proc. IEEE 27th Annu. Int. Conf. Engineering in Medicine and Biology*, pp. 5519-5522, January 2005.

[21] L. Gamero and R. Watrous, "Detection of the first and second heart sound using probabilistic models," in *Proc. IEEE 25th Annu. Int. Conf. IEEE Engineering in Medicine and Biology*, pp. 2877-2880, September 2003.

[22] T. Oskiper and R. Watrous, "Detection of the first heart sound using a time-delay neural network," in *Proc. Comput. Cardiol.*, pp. 537-540, September 2002.

[23] J. E. Hebden and I. N. Torry, “Neural network and conventional classifiers to distinguish between first and second heart sound,” in *Proc. IEE Coll. Artificial Intelligence Methods for Biomedical Data Processing*, vol. 3, pp. 1-6, April 1996.

[24] A. Iwata et al., “Algorithm for detecting the first and the second heart sounds by spectral tracking,” *Med. Biol. Eng. Comput.*, vol. 18, no. 1, pp. 19-26, January 1980.

[25] Yong Shao, Ying-hong Zhang, and Mamie Liu, “Using phonocardiography to investigate maternal cardiac reserve function in gestational hypertension and pre-eclampsia,” *J. Obstetrics and Gynaecology Research*, vol. 39, no. 1, pp. 53-60, January 2013.

[26] X. Yang and W. Zeng, “A relative value method for measuring and evaluating neonatal cardiac reserve,” *Indian J. Pediatr.*, vol. 77, no. 06, pp. 661-664, June 2010.

[27] Lin Yan Wei, “Study on the change of cardiac reserve function during normal vaginal delivery,” Master’s thesis, Chongqing Medical University, China, 2012.

[28] G. Xing-ming, Z. Li-sha, W. Dong, Y. Feng-zhi, and X. Shou-zhong, “Change of cardiac reserve during abnormal pregnancy and its evaluation,” *Acta Academiae Medicinae Sinicae*, vol. 33, no. 1, pp. 58-61, February 2011.

[29] B. Hansen Peter et al., “Phonocardiography as a monitor of cardiac performance during anesthesia,” *Anesthesia & Analgesia*, vol. 68, no. 3, pp. 385-392, March 1989.

[30] Isa Yildirim and Rashid Ansari, "A robust method to estimate time split in second heart sound using instantaneous frequency analysis," in *Proc. IEEE 29th Annu. Int. Conf. Engineering in Medicine and Biology*, pp. 1855-1858, August 2007.

[31] Abdelghani Djebbari and Fethi Bereksi-Reguig, "Detection of the valvular split within the second heart sound using the reassigned smoothed pseudo Wigner-Ville distribution," *Biomed. Eng. OnLine*, vol. 12, no. 37, pp. 1-21, April 2013.

[32] T. S. Leung et al., "Analysis of the second heart sound for diagnosis of paediatric heart disease," *IEE Proc. Sci. Meas. Technol.*, vol. 145, no. 6, pp. 285-290, November 1998.

[33] H. M. Mgdob et al., "Application of Morlet transform wavelet in the detection of paradoxical splitting of the second heart sound," in *Proc. Computers in Cardiology*, pp. 323-326, September 2003.

[34] S. M. Debbal and F. Bereksi-Reguig, "Automatic measure of the split in the second cardiac sound by using the wavelet transform technique," *Computers in Biology and Medicine*, vol. 37, pp. 269-276, March 2007.

[35] S. Santos, P. Carvalho, R. P. Paiva, and J. Henriques, "Detection of the S2 split using the Hilbert and wavelet transforms," in *Congresso de M´etodos Num´ericos em Engenharia*, pp. 1-11, January 2011.

[36] P. S. Vikhe, N. S. Nehe, and V. R. Thool, "Heart sound abnormality detection using short time Fourier transform and continuous wavelet transform," in *Proc. IEEE 2nd Int. Conf. Emerging Trends in Engineering and Technology*, pp. 50-54, December 2009.

[37] P. S. Vikhe, S. T. Hamde, and N. S. Nehe, "Wavelet transform based abnormality analysis of heart sound," in *Proc. Int. Conf. Advances in Computing, Control, and Telecommunication Technologies*, pp. 367-371, December 2009.

[38] Vivek Nigam and Roland Priemer, "A dynamic method to estimate the time split between the A2 and P2 components of the S2 heart sound," *Physiol. Meas.*, vol. 27, no. 7, pp. 553-567, July 2006.

[39] Hamza Cherif L and Debbal SM, "Algorithm for detection of the internal components of the heart sounds and their split using a Hilbert transform," *J. Med. Eng. Technol.*, vol. 37, no. 3, pp. 220-230, April 2013.

[40] C. S. Lima and D. Barbosa, "Automatic segmentation of the second cardiac sound by using wavelets and hidden Markov models," in *Proc. IEEE 30th Annu. Int. Conf. on Engineering in Medicine and Biology*, pp. 334-337, August 2008.

[41] N. J. Mehta and I. A. Khan, "Third heart sound: genesis and clinical importance," *Int. J. Cardiology*, vol. 97, no. 2, pp. 183-186, November 2004.

[42] Li-Sha Zhong et al., "The third heart sound after exercise in athletes: An exploratory study," *Chinese J. of Physiology* vol. 54, no. 4, pp. 219-224, August 2011.

[43] G. M. Marcus et al., "Association between phonocardiographic third and fourth heart sounds and objective measures of left ventricular function," *J. Amer. Med. Assoc.*, vol. 293, no. 18, pp. 2238-2244, May 2005.

[44] G. M. Marcus et al., “Relationship between accurate auscultation of a clinically useful third heart sound and level of experience,” *Arch. Intern. Med.,* vol. 166, no. 6, pp. 617-622, March 2006.

[45] C. Ahlstrom, P. Hult, and P. Ask, “Detection of the 3rd heart sound using recurrence time statistics,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, vol. 2, pp. 1040-10, May 2006.

[46] J. Wynne, “The clinical meaning of the third heart sound,” *Amer. J. Med.*, vol. 111, no. 2, pp. 157-158, August 2001.

[47] C. M. Tribouilloy et al., “Pathophysiologic determinants of third heart sounds: a prospective clinical and doppler echocardiographic study,” *Amer. J. Med.*, vol. 111, no. 2, pp. 96-102, August 2001.

[48] C. E. Lok, C. D. Morgan, and N. Ranganathan, “The accuracy and interobserver agreement in detecting the ’gallop sounds’ by cardiac auscultation,” *J. Chest*, vol. 114, no. 5, pp. 1283-1288, November 1998.

[49] Y. Wang et al., “Identification of the normal and abnormal heart sounds using wavelettime entropy features based on OMS-WPD,” *Future Gen. Comput. Sys.*, vol. 37, pp. 488-495, July 2014.

[50] F. Safara et al., “Multi-level basis selection of wavelet packet decomposition tree for heart sound classification,” *Comput. Biol. Med.*, vol. 43, no. 10, pp. 1407-1414, October 2013.

[51] A. Moukadem et al., “A robust heart sounds segmentation module based on S-transform,” *Biomed. Signal Process. Control*, vol. 8, no. 3, pp. 273-281, May 2013.

[52] H. Sun et al., “An improved empirical mode decomposition-wavelet algorithm for phonocardiogram signal de-noising and its application in the first and second heart sound extraction,” in *Proc. 6th Int. Conf. Biomed. Eng. and Informatics*, pp. 187-191, December 2013.

[53] Ashok Mondal et al., “Boundary estimation of cardiac events S1 and S2 based on Hilbert transform and adaptive thresholding approach,” in *Proc. Indian Conf. Medical Informatics and Telemedicine*, pp. 43-47, March 2013.

[54] H. Naseri and M. R. Homaeinezhad, “Detection and boundary identification of phonocardiogram sounds using an expert frequency-energy based metric,” *Ann. Biomed. Eng.,* vol. 41, no. 2, pp. 279-292, February 2013.

[55] F. Hedayioglu et al.,“Denoising and segmentation of the second heart sound using matching pursuit,” in *Proc. IEEE 34th Annu. Int. Conf. Engineering in Medicine and Biology*, pp. 3440-3443, 2012.

[56] Y. Chen, S. Wang, C.-H. Shen, and F. K. Choy, “Matrix decomposition based feature extraction for murmur classification,” *Med. Eng. and Physics*, vol. 34, no. 6, pp. 756-761, July 2012.

[57] D. Boutana, M. Benidir, and B. Barkat, “Segmentation and identification of some pathological phonocardiogram signals using time-frequency analysis,” *IET Signal Processing*, vol. 5, no. 6, pp. 527-537, September 2011.

[58] S. Choi, Y. Shin, and H. K. Park, “Selection of wavelet packet measures for insufficiency murmur identification,” *Expert Sys. Appl.*, vol. 38, no. 4, pp. 4264-4271, April 2011.

[59] S. Choi and Z. Jiang, "Cardiac sound murmurs classification with autoregressive spectral analysis and multi-support vector machine technique," *Comput. Bio. and Med.*, vol. 40, no. 1, pp. 8-20, January 2010.

[60] L. D. Avendano-Valencia, J. I. Godino-Llorente, M. Blanco-Velasco, and G. CastellanosDominguez, "Feature extraction from parametric time-frequency representations for heart murmur detection," *Ann. Biomed. Eng.*, vol. 38, no. 8, pp. 2716-2732, August 2010.

[61] W. -C. Kao and C. -C. Wei, "Automatic phonocardiograph signal analysis for detecting heart valve disorders," *Expert Sys. Appl.*, vol. 38, no. 6, pp. 6458-6468, June 2011.

[62] S. Yuenyong, A. Nishihara, W. Kongprawechnon, and K. Tungpimolrut, "A framework for automatic heart sound analysis without segmentation," *Biomed. Eng. Online*, vol. 10, no. 13, pp. 1-23, February 2011.

[63] S. E. Schmidt et al., "Segmentation of heart sound recordings by a duration-dependent hidden Markov model," *Physiol. Meas.*, vol. 31, no. 4, pp. 513-529, April 2010.

[64] X. Wang et al. "Detection of the first and second heart sound using heart sound energy," in *Proc. Int. Conf. Biomed. Eng. and Informatics*, pp. 1-4, October 2009.

[65] A. Castro et al., "Heart sound segmentation of pediatric auscultations using wavelet analysis," *IEEE 35th Int. Conf. Engineering in Medicine and Biology*, pp. 3909-3912, 2013.

[66] N. Atanasov and T. Ning, "Quantitative delineation of heart murmurs using features derived from autoregressive modeling," in *Proc. IEEE 33rd Annu. Northeast Conf. Bioengineering*, pp. 167-168, March 2007.

[67] D. Kumar, P. Carvalho, M. Antunes, R. P. Paiva, and J. Henriques, "Heart murmur classification with feature selection," in *Proc. IEEE 32nd Annu. Int. Conf. Engineering in Medicine and Biology*, pp. 4566-4569, August 2010.

[68] C. Ahlstrom et al., "Feature extraction for systolic heart murmur classification," *Ann. Biomed. Eng.*, vol. 34, no. 11, pp. 1666-1677, November 2006.

[69] https://physionet.org/challenge/2016/

[70] Noise Cancellation Using Adaptive Filter for PCG Signal, Naveen Dewangan, Conference: BIT, At Bhilai, Volume: Volume 3, Issue 4, July-August 2014.

[71] Implementation of Adaptive Algorithm for PCG Signal Denoising, Mr. R. M. Potdar, Dr. Mekhram Meshram, Naveen Dewangan & Dr. Ramesh Kumar, INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN ELECTRICAL, ELECTRONICS, INSTRUMENTATION AND CONTROL ENGINEERING Vol. 3, Issue 4, April 2015.

[72] Cardiac Signal Denoising Using Adaptive Filtering Techniques, Laxmi Shetty, Usha Desai, PROCEEDINGS OF NJCIET 2015, ISBN: 97-8-93-81195-82-6.

[73] DENOISING OF HEART SOUND SIGNAL USING WAVELET TRANSFORM, Gyanaprava Mishra, Kumar Biswal, Asit Kumar Mishra, International Journal of Research in Engineering and Technology, Volume: 02 Issue: 04, Apr-2013.

[74] Moukadem A, Dieterlena A, Hueberb N and Brandtc C, A Robust heart sounds segmentation module based on S-transform Biomed. Signal Process. Control Vol 8, No. 1, pp 273–81, 2013.

[75] Heart Sound Segmentation Algorithm Based on Heart Sound Envelogram, H Liang, S Lukkarinen, I Hartimo, Computing in Cardiology, 1997; IEEE: pp 105-108.

[76] Sun S, Jiang Z, Wang H and Fang Y 2014 Automatic moment segmentation and peak detection analysis of heart sound pattern via short-time modified Hilbert transform Comput. Methods Programs Biomed. 114 219–30.

[77] Jiang Z and Choi S 2006 A cardiac sound characteristic waveform method for in-home heart disorder monitoring with electric stethoscope Expert Syst. Appl. 31 286–98.

[78] Yan Z, Jiang Z, Miyamoto A and Wei Y 2010 The moment segmentation analysis of heart sound pattern Comput. Methods Programs Biomed. 98 140–50.

[79] Ari S, Kumar P and Saha G 2008 A robust heart sound segmentation algorithm for commonly occurring heart valve diseases J. Med. Eng. Technol. 32 456–65.

[80] Naseri H and Homaeinezhad M R 2013 Detection and boundary identification of phonocardiogram sounds using an expert frequency-energy based metric Ann. Biomed. Eng. 41 279–92.

[81] Kumar D, Carvalho P, Antunes M and Henriques J 2006 Detection of S1 and S2 heart sounds by high frequency signatures Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (New York: IEEE) pp 1410–6.

[82] Varghees V N and Ramachandran K 2014 A novel heart sound activity detection framework for automated heart sound analysis Biomed. Signal Process. Control 13 174–88.

[83] Pedrosa J, Castro A and Vinhoza T T V 2014 Automatic heart sound segmentation and murmur detection in pediatric phonocardiograms Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (Chicago: IEEE) pp 2294–7.

[84] Nigam V and Priemer R 2005 Accessing heart dynamics to estimate durations of heart sounds Physiol. Meas. 26 1005–18.

[85] Vepa J, Tolay P and Jain A 2008 Segmentation of heart sounds using simplicity features and timing information IEEE Int. Conf. on Acoustics, Speech and Signal Processing (Las Vegas, NV: IEEE) pp 469–72.

[86] Papadaniil C D and Hadjileontiadis L J 2014 Efficient heart sound segmentation and extraction using ensemble empirical mode decomposition and kurtosis features IEEE J. Biomed. Health Inform. 18 1138–52.

[87] Gharehbaghi A, Dutoir T, Sepehri A, Hult P and Ask P 2011 An automatic tool for pediatric heart sounds segmentation Computing in Cardiology (Hangzhou: IEEE) pp 37–40.

[88] Oskiper T and Watrous R 2002 Detection of the first heart sound using a time-delay neural network Computing in Cardiology (Memphis: IEEE) pp 537–40.

[89] Sepehri A A, Gharehbaghi A, Dutoit T, Kocharian A and Kiani A 2010 A novel method for pediatric heart sound segmentation without using the ECG Comput. Methods Programs Biomed. 99 43–8.

[90] Chen T, Kuan K, Celi L and Clifford G D 2009 Intelligent heartsound diagnostics on a cellphone using a hands-free kit AAAI Spring Symp. on Articial Intelligence for Development (Stanford University) pp 26–31.

[91] Gupta C, Palaniappan R, Swaminathan S and Krishnan S 2007 Neural network classification of homomorphic segmented heart sounds Appl. Soft. Comput. 7 286–97.

[92] Tang H, Li T and Qiu T S 2010b Noise and disturbance reduction for heart sounds in the cycle frequency domain based on non-linear time scaling IEEE Trans. Biomed. Eng. 57 325–33.

[93] Rajan S, Budd E, Stevenson M and Doraiswami R 2006 Unsupervised and uncued segmentation of the fundamental heart sounds in phonocardiograms using a time-scale representation Int. Conf. of the IEEE Engineering in Medicine and Biology Society (New York: IEEE) pp 3732–5.

[94] Gamero L G and Watrous R 2003 Detection of the first and second heart sound using probabilistic models Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (Cancun: IEEE) pp 2877–80.

[95] Ricke A D, Povinelli R J and Johnson M T 2005 Automatic segmentation of heart sound signals using hidden Markov models Computers in Cardiology (Lyon: IEEE) pp 953–6.

[96] Gill D, Gavrieli N and Intrator N 2005 Detection and identification of heart sounds using homomorphic envelogram and self-organizing probabilistic model Computers in Cardiology (Lyon: IEEE) pp 957–60.

[97] Sedighian P, Subudhi A W, Scalzo F and Asgari S 2014 Pediatric heart sound segmentation using hidden Markov model Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (Chicago: IEEE) pp 5490–3.

[98] Castro A, Vinhoza T T V, Mattos S S and Coimbra M T 2013 Heart sound segmentation of pediatric auscultations using wavelet analysis Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (Osaka: IEEE) pp 3909–12.

[99] Schmidt S E, Holst-Hansen C, Graff C, Toft E and Struijk J J 2010a Segmentation of heart sound recordings by a duration-dependent hidden Markov model Physiol. Meas. 31 513–29.

[100] Springer D B, Tarassenko L and Clifford G D 2016 Logistic regression-HSMM-based heart sound segmentation IEEE Trans. Biomed. Eng. 63 822–32.

[101] Gerbarg D S, Taranta A, Spagnuolo M and Hofler J J 1963 Computer analysis of phonocardiograms Prog. Cardiovasc. Dis. 5 393–405.

[102] Akay Y M, Akay M, Welkowitz W and Kostis J 1994 Noninvasive detection of coronary artery disease IEEE Eng. Med. Biol. 13 761–4.

[103] Liang H and Hartimo I 1998 A feature extraction algorithm based on wavelet packet decomposition for heart sound signals Proc. of the IEEE-SP Int. Symp. on Time-Frequency and Time-Scale Analysis (Pittsburgh, PA: IEEE) pp 93–6.

[104] Uguz H 2012a Adaptive neuro-fuzzy inference system for diagnosis of the heart valve diseases using wavelet transform with entropy Neural Comput. Appl. 21 1617–28.

[105] Bhatikar S R, DeGroff C and Mahajan R L 2005 A classifier based on the artificial neural network approach for cardiologic auscultation in pediatrics Artif. Intell. Med. 33 251–60.

[106] Sepehri A A, Hancq J, Dutoit T, Gharehbaghi A, Kocharian A and Kiani A 2008 Computerized screening of children congenital heart diseases Comput. Methods Programs Biomed. 92 186–92.

[107] Ahlstrom C, Hult P, Rask P, Karlsson J E, Nylander E, Dahlstrom U and Ask P 2006 Feature extraction for systolic heart murmur classification Ann. Biomed. Eng. 34 1666–77.

[108] De Vos J P and Blanckenberg M M 2007 Automated pediatric cardiac auscultation IEEE Trans. Biomed. Eng. 54 244–52.

[109] Uguz H 2012b A biomedical system based on artificial neural network and principal component analysis for diagnosis of the heart valve diseases J. Med. Syst. 36 61–72.

[110] Ari S, Hembram K and Saha G 2010 Detection of cardiac abnormality from PCG signal using LMS based least square SVM classier Expert Syst. Appl. 37 8019–26.

[111] Zheng Y N, Guo X M and Ding X R 2015 A novel hybrid energy fraction and entropy-based approach for systolic heart murmurs identification Expert Syst. Appl. 42 2710–21.

[112] Patidar S, Pachori R B and Garg N 2015 Automatic diagnosis of septal defects based on tunable-Q wavelet transform of cardiac sound signals Expert Syst. Appl. 42 3315–26.

[113] Maglogiannis I, Loukis E, Zafiropoulos E and Stasis A 2009 Support vectors machine-based identification of heart valve diseases using heart sounds Comput. Methods Programs Biomed. 95 47–61.

[114] Gharehbaghi A, Ekman I, Ask P, Nylander E and Janerot-Sjoberg B 2015 Assessment of aortic valve stenosis severity using intelligent phonocardiography Int. J. Cardiol. 198 58–60.

[115] Wang P, Lim C S, Chauhan S, Foo J Y and Anantharaman V 2007 Phonocardiographic signal analysis method using a modified hidden Markov model Ann. Biomed. Eng. 35 367–74.

[116] Saracoglu R 2012 Hidden Markov model-based classification of heart valve disease with PCA for dimension reduction Eng. Appl. Artif. Intell. 25 1523–8.

[117] Bentley P M, Nokia R D, Camberley U K, Grant P M and McDonnell J T E 1998 Time-frequency and time-scale techniques for the classification of native and bio-prosthetic heart valve sounds IEEE Trans. Biomed. Eng. 45 125–8.

[118] Quiceno-Manrique A F, Godino-Llorente J I, Blanco-Velasco M and Castellanos-Dominguez G 2010 Selection of dynamic features based on time-frequency representations for heart murmur detection from phonocardiographic signals Ann. Biomed. Eng. 38 118–37.

[119] Avendano-Valencia L D, Godino-Llorente J I, Blanco-Velasco M and Castellanos-Dominguez G 2010 Feature extraction from parametric time-frequency representations for heart murmur detection Ann. Biomed. Eng. 38 2716–32.

[120] Tinati, M.A, Bouzerdoum, A., Mazumdar, J, "Modified Adaptive Line Enhancement Filter And Its Application To Heart Sound Noise Cancellation", Signal Processing and Its Applications, 1996. ISSPA 96., Fourth International Symposium on (Volume: 2), Gold Coast, Queensland, Australia.

[121] S. Charleston ; M.R. Azimi-Sadjadi, "Reduced order Kalman filtering for the enhancement of respiratory sounds", IEEE Transactions on Biomedical

Engineering , Volume: 43 , Issue: 4 , April 1996, pp 421-424.

[122] Lu Zhang, Lei Wang, "Heart Sound Enhancement Based on Improved Spectral Subtraction", 2015 11th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS) (2015), Bangkok, Thailand, Nov. 23, 2015 to Nov. 27, 2015, ISBN: 978-1-4673-9721-6, pp: 546-553, DOI Bookmark: http://doi.ieeecomputersociety.org/10.1109/SITIS.2015.54

[123] Ramos JP, Carvalho P, Paiva RP, Henriques J," Modulation filtering for noise detection in heart sound signals", Conf Proc IEEE Eng Med Biol Soc. 2011;2011:6013-6. doi: 10.1109/IEMBS.2011.6091486.

[124] Tang H, Li T, Park Y and Qiu T S 2010a Separation of heart sound signal from noise in joint cycle frequency-time-frequency domains based on fuzzy detection IEEE Trans. Biomed. Eng. 57 2438–47.

[125] K. Srinivasan and A. Gersho, "Voice activity detection for cellular networks," in Proc. IEEE Speech Coding Workshop, Oct. 1993, pp. 85–86.

[126] J. Sohn and W. Sung, "A voice activity detector employing soft decision based noise spectrum adaptation," in Proc. Int. Conf. Acoustics, Speech, and Signal Processing, 1998, pp. 365–368.

[127] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-32, pp. 1109–1121, Dec. 1984.

[128] Jongseo Sohn, Student Member, IEEE, Nam Soo Kim, Member, IEEE, and Wonyong Sung, "A Statistical Model-Based Voice Activity Detection," IEEE Signal Processing Letters, Vol. 6, No. 1, January 1999, pp. 1-3.

[129] B. L. McKinley and G. H. Whipple, “Model based speech pause detection,”in Proc.22th IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP’97), Munich, Germany, Apr. 20–24, 1997, pp. 1179–1182.

[130]] J. Meyer, K. U. Simmer, and K. D. Kammeyer, “Comparison of one- and two-channel noise-estimation techniques,” in Proc. 5th Int. Workshop on Acoustic Echo and Noise Control (IWAENC’97), London, U.K., Sept. 11–12, 1997, pp. 137–145.

[131] A. Papoulis, Probability, Random Variables, and Stochastic Processes, third ed. New York: McGraw-Hill, 1991.

[132] H. G. Hirsch and C. Ehrlicher, “Noise estimation techniques for robust speech recognition,” in Proc. 20th IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP’95), Detroit, MI, May 8–12, 1995, pp. 153–156.

[133] R. J. McAulay and M. L. Malpass, “Speech enhancement using a soft decision noise suppression filter,” IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-28, pp. 137–145, Apr. 1980.

[134] C. Ris and S. Dupont, “Assessing local noise level estimation methods: Application to noise robust ASR,” Speech Commun., vol. 34, no. 1–2, pp. 141–158, Apr. 2001.

[135] Israel Cohen, “Noise Spectrum Estimation in Adverse Environments: Improved Minima Controlled Recursive Averaging,” IEEE Transactions On Speech and Audio Processing, Vol. 11, No. 5, September 2003.

[136] P.-Y. Chen and I. W. Selesnick, “Translation-invariant shrinkage/thresholding of group sparse signals,” Signal Processing. vol. 94, pp 476-489, January 2014.

[137] Lu Jing-yi, Lin Hong, Ye Dong, and ZhangYan-sheng, "A New Wavelet Threshold Function and Denoising Application," Mathematical Problems in Engineering, Hindawi Publishing Corporation, Volume 2016, pp 1-8.

PUBLICATIONS

1. M. Vishwanath Shervegar and Ganesh V Bhat, "Automatic Segmentation of Phonocardiogram using the occurrence of the cardiac events", Informatics in Medicine Unlocked, Vol 9C 2017, pp 6-10, ELSEVIER Publishers (Scopus Indexed Journal).

2. M. Vishwanath Shervegar and Ganesh V Bhat, "Classification of heart sounds using Gaussian Mixture Model", Porto Biomedical Journal, Vol 3, No.1, 2018, WOLTERS KLUWER Publishers (PubMed Indexed Journal).

3. M. Vishwanath Shervegar and Ganesh V Bhat, "Denoising of Phonocardiogram using Time Frequency Block Thresholding Method", International Journal of Biomedical Engineering and Technology, Vol 28, No.1, 2018, pp 1-17, INDERSCIENCE Publishers (Web of Science Indexed Journal).