Detection of Heart Diseases Using PCG Signals

A Dissertation
Submitted in partial fulfillment of the requirements for the award of degree of

Master of Engineering

in

Electronic Instrumentation and Control

Submitted by:

Amandeep Cheema
(801151003)

UNDER THE GUIDANCE OF

Mr. Mandeep Singh
Assistant Professor
EIED
Thapar University, Patiala

ELECTRICAL AND INSTRUMENTATION ENGINEERING DEPARTMENT
THAPAR UNIVERSITY
(Established under the section 3 of UGC Act, 1956)
PATIALA – 147004

JULY 2013

DECLARATION

I hereby certify that the work is being presented in the thesis work entitled "**Detection of Heart Diseases Using PCG Signals**" in fulfillment of award of degree of **Master of Engineering** in **Electronic Instrumentation and Control** submitted in Electrical and Instrumentation Engineering Department, Thapar University, Patiala is an authentic record of my own work carried under the supervision of **Mr. Mandeep Singh**, Assistant Professor, Electrical and Instrumentation Engineering Department, Thapar University, Patiala, Punjab.

Date: July 15, 2013.

A.D Cheema.

Amandeep Cheema

801151003

I certify that the above statement made by the student is correct to the best of my knowledge and belief.

Date: July 15, 2013.

15/7/13

Mr. Mandeep Singh
Assistant Professor
Department of Electrical and
Instrumentation Engineering
Thapar University, Patiala
Punjab.

Countersigned By

S. Ghosh

Dr. Smarajit Ghosh
Professor & Head
Department of Electrical
Instrumentation Engineering
Thapar University, Patiala
Punjab

Dr. S.K. Mohapatra
Dean of Academic Affairs,
Thapar University, Patiala
Punjab.

ABSTRACT

In Phonocardiography (PCG), the heart sound signals are recorded using an electronic stethoscope and are displayed graphically on the PC/laptop rather than listening the heart sounds as done in traditional auscultation. More details are accessible visually, because the analysis is not limited by the human audibility range or experience of the physician while listening. The heart diseases can be detected even before the symptoms of pathology appear and this makes it a highly potential diagnostic test for the future. The aim of this study is to detect various heart diseases using the PCG signals. Many diagnostic features can be extracted using PCG which otherwise require tests like Electrocardiography (ECG) or Echocardiography. This can save a lot of money and time. Moreover, the requirement of instruments is minimum i.e. an electronic stethoscope and a PC/laptop which are available even at primary health care centres. This will increase the outreach of diagnosis to under-privileged and rural people. In order to detect murmurs efficiently, two new features have been proposed. A total of 23 features are evaluated for normal and murmur signals. Then 5 optimal features are finally selected for classification using Ranker and Info Gain Evaluation method for feature reduction. Finally, an algorithm using Naïve Bayes classifier is proposed for murmur detection and arrhythmia detection (like Tachycardia or Bradycardia).

ACKNOWLEDGEMENT

As I come to the end of my thesis, it gives me immense pleasure to express my gratitude towards the people who stood by me and have contributed to the completion of my thesis. I am extremely thankful to God who without showing his presence, was with me all the time. First of all, big thanks to my thesis guide ***Mr. Mandeep Singh***, Assistant Professor, Electrical and Instrumentation Engineering Department, Thapar University, Patiala for the immense support and guidance. Sir, I will always be thankful as you introduced me to the field of research and explained me everything with your wisdom and patience and you had an instant solution to every problem that I faced during my research work.

I am extremely thankful to ***Dr. Smarjit Ghosh***, Professor and Head, Electrical and Instrumentation Department, Thapar University, Patiala for the support provided by him. Sir, your spirit and cheerful nature always inspired me to face every problem with a smile. I knew I could always look up to you if I had some problem and this driving force and assurance was enough to work whole-heartedly.

I want to thank, ***Dr. R.K. Gorea***, Principal, Guru Nanak Mission Hospital who helped during the dataset collection phase.

I am also thankful to my friends Surabhi Gangwar, Paramjeet Shahi, Prachi Sharma, Mohit Kumar, Rajeev Kumar, Shivangi. You all have been a great support. Without you this journey and stay at Thapar would not have been as special as it is now. I am highly indebted to have friends like you. You all stood by me every time and worked as stress-busters when things were not smooth.

A very special thanks to my sister Preet Kamal Chima for the time you devoted on me. I am happy that you proofread my thesis and were patient enough to listen to me every time I needed your help. Your medical background and in-depth understanding of topics was a great help for me. I am indebted as you gave me immense support and time.

All the thanks to my mom and dad, without them nothing has been possible ever. The support given by you is incomparable. I am short of words as nothing would be sufficient enough for what you did for me in every phase.

A.D Cheema.

Place: Thapar University, Patiala

Amandeep Cheema

Date: July 15, 2013.

801151003

TABLE OF CONTENTS

Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ANFIS – Adaptive Neuro-Fuzzy Inference System

ANN – Artificial neural Network

ATP – Adenosine Triphosphate

AV Node – Atrioventricular Node

CSCW – Cardiac Sound Characteristic Waveform

DSP – Digital Subtaction Phonocardiography

DWT – Discrete Wavelet Transform

ECG – Electrocardiography

FFT – Fast Fourier Transform

FHS – Fundamental Heart Sounds

LA – Left Atrium

LV – Left Ventricle

MSE – Mean Square Error

PCG - Phonocardiography

RA – Right Atrium

RBF – Radial Basis Function

RV – Right Ventricle

SA Node – Sinoatrial Node

SRM – Structural Risk Minimization

STFT – Short Time Fourier Transform

SVM – Support Vector Machine

THD – Total Harmonic Distortion

ZCR – Zero Crossing Rate

LIST OF PUBLICATIONS

[1] Amandeep Cheema, Mandeep Singh, "Steps Involved in Heart Sound Analysis- A Review of Existing Trends", International Journal of Engineering Trends and Technology, Volume 4, Issue 7, 2013.

[2] Mandeep Singh, Amandeep Cheema, "Heart Sound Classification Using Naïve Bayesian approach", IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, Canada, 2013, (Communicated).

CHAPTER

1 INTRODUCTION

1.1 BACKGROUND

Heart disease is a major health problem and a leading cause of fatality throughout the world. The abnormalities appear in heart sounds much before the symptoms of disease start appearing. The act of analyzing sounds in the body that are produced in response to mechanical vibrations generated in the organs is called Auscultation. The potential of auscultation can be seen from the fact that whenever a patient approaches a physician, he uses a stethoscope for auscultation and if the abnormalities are observed in the heart sounds, the patient is referred for other tests like Electrocardiography (ECG) or Echocardiography. The treatment can be easier, efficient and economical if the condition is detected early. So it would be very beneficial to detect heart diseases at an early stage. The scenario could have been much better if we don't have to wait for symptoms of the heart disease to appear and then approach a cardiologist for diagnosis and treatment so that before worsening of the problem, a proper diagnosis can be done. Then other techniques like Echocardiography could be implemented so as to get a clear view of the problem. The cardiac disorders can be detected efficiently and economically using auscultation as it requires minimal equipment. Hence it is of immense use in the rural health management where other high-end equipments are not present and electronic stethoscope can be used for diagnosis in such primary health care centres. If the cardiac monitoring of patient at home is required, then due to portability stethoscope is the only available means of diagnosis and also in case of infants where other techniques like ECG are difficult to implement. Conventional auscultation using acoustic stethoscope requires extensive training and experience of the physician for proper diagnosis. Moreover, the storage of records for follow ups and future references is also not possible with conventional auscultation [1]. This is the driving force for this study in order to move towards automatic auscultation using electronic stethoscopes. In this study Phonocardiography (PCG) will be used for heart condition monitoring which find its roots in auscultation. The method proposed in this study is not a replacement of the other techniques but an early indicator of the problem in order to prevent worsening of the situation.

1.2 PHONOCARDIOGRAPHY

The word *phono* has Greek roots means -sound.

The word *cardio* is originated from Greek word kardia means – Heart.

The word *graphy* is originated from French word graphie means – process of writing or recording.

So, Phonocardiography etymologically means the process of recording of Heart Sounds. The heart sounds are primarily generated from blood turbulence. The blood turbulence occurs due to fast accelerations and retardations of the blood in the chambers and arteries caused by the contraction or closure of the heart valves, which in turn produce mechanical vibrations that propagate through the body tissues up to the surface of the thorax [2]. The heart sounds recorded by an electronic stethoscope are converted to digital signals and are plotted as in figure1-1 and termed as PCG (Phonocardiograph) signal.

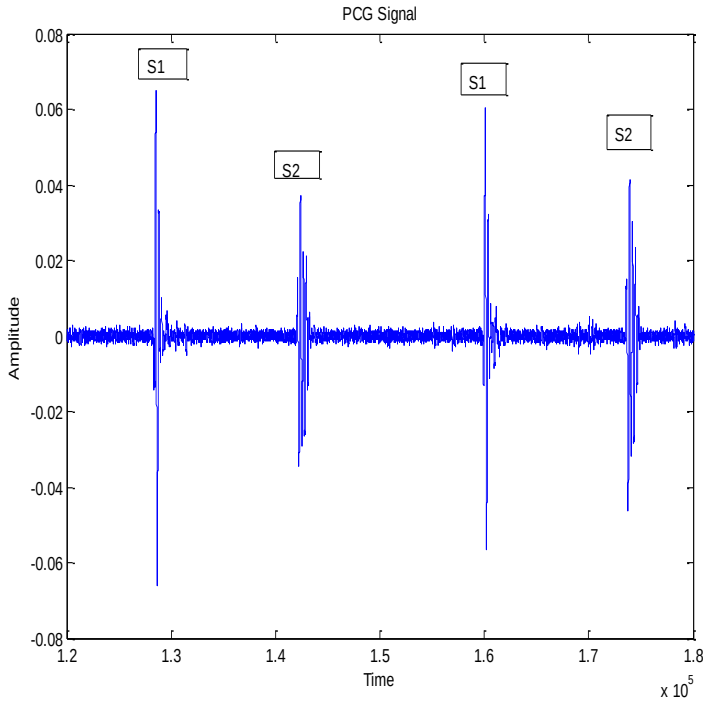


Figure 1-1 Phonocardiography Signal

Heart sounds comprise of four components among which the dominant two S1 and S2 as shown in figure 1-1 (commonly known as Fundamental Heart Sounds) are caused by closing of valves. After that S3 is caused by vibration of ventricular valves as a result of first rapid filling, S4 occurs when atria contract during the second phase of ventricular filling. The low-pitched S1 is best heard at the mitral auscultation site as it is caused by closure of mitral and tricuspid valves and shorter duration but louder sound S2 is caused by aortic and pulmonary valve closure and is best heard over the aortic auscultation area [3].

1.3 AIM OF WORK

Auscultation has immense potential in detecting heart problems at an early stage. Since auscultation requires tedious and rigorous training and auscultation skills can only be enhanced by experience, which is sometimes missing in the present scenario. This is because the diagnosis of heart diseases these days is completely dependant on modern techniques like ECG and Echocardiography which give a better view of the problem. The abnormalities appear in heart sounds much before the symptoms of pathology start appearing. So, if the science and art of auscultation are properly used then early detection of heart diseases and proper treatment can be done at an early stage and it would be a great breakthrough as only precautions and lifestyle changes could be sufficient to deal with these lifestyle related diseases at an early stage.

It is also very useful in case of infants where ECG recording and other techniques are difficult to implement.

It is also of immense use in the rural health management as sometimes stethoscope is the only available instrument for diagnosis in primary health care centres.

In this study we aim to develop a method for classification of heart sounds into normal and abnormal sounds so that state of heart could be periodically checked at home and everybody doesn't have to wait for symptoms of disease to appear and then to go to a cardiologist. Since PCG signals are an early indicator of heart problems so that before worsening of the problem, proper diagnosis can be done. Then other techniques like Echocardiography could be implemented to get a better view of the problem. So, the proposed method is not a replacement of the other techniques but an early indicator of the problem in order to prevent worsening of the situation. In this study, a classification method is proposed to separate normal and abnormal heart sound signals having murmurs without getting into the cumbersome process of segmenting fundamental heart sounds using ECG gating. In

this study we analyze Phonocardiogram (PCG) signals for normal and murmur heart sounds and established that between the fundamental heart sounds S1 and S2, where murmurs are prominent, the amplitude of the murmur signal is higher than normal signal as shown in figure 1-2. Then the segments between S1 and S2 of the same cardiac cycle, S2 of one cardiac cycle and S1 of next cardiac cycle were selected for further analysis, features were extracted in order to differentiate between normal and murmur signals. In our study we used the common assumption that systole is smaller than diastole. Another algorithm has been made which deals with the electrical disorders of heart that causes arrhythmia. The method presented in the study can detect various types of arrhythmia like Tachycardia and bradycardia efficiently.

The proposed algorithms can be easily implemented on latest electronic stethoscopes, and therefore unnecessary ECG can be avoided.

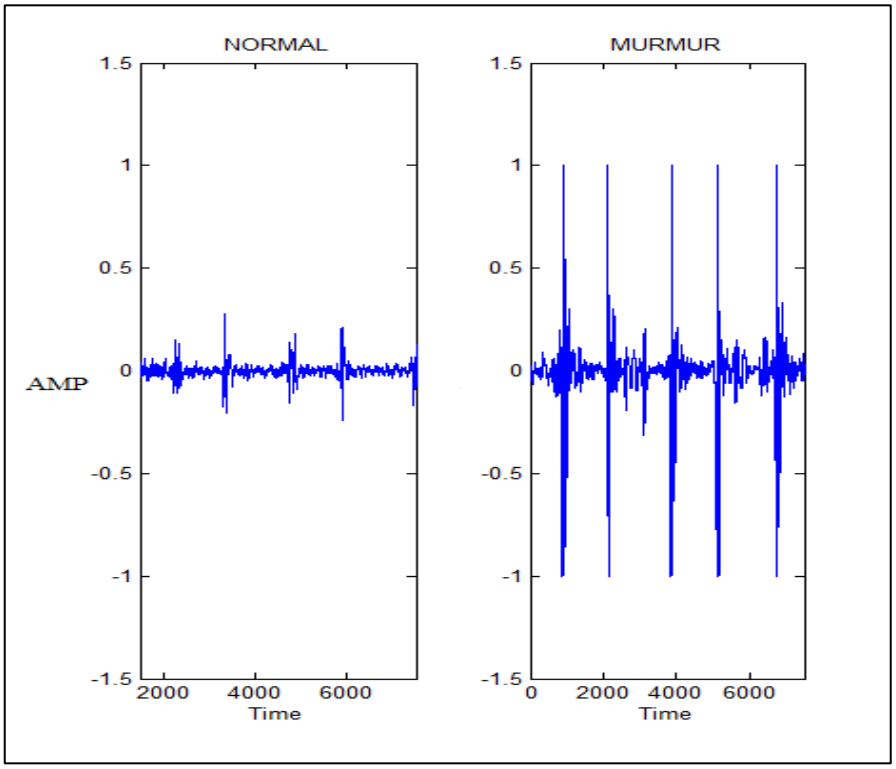


Figure 1-2 Visual amplitude difference in normal and murmur signal

1.4 OUTLINE OF THESIS

As we will be navigating through the thesis we will be proceeding with a stepwise order for better understanding.

Chapter 1 is Introduction. In this chapter the introduction to various concepts is given including the various terms used in the thesis. In this chapter firstly introduction to auscultation is given and a difference is established between conventional and automatic auscultation. Then the word Phonocardiography is introduced, which is the technique on which our study is based. Then the outline of the thesis is presented which shows how we proceed in our study and hence our thesis.

Chapter 2 is Literature Review. In this chapter a review of previous studies on this topic has been given. Here we have enlisted various works that are being done to decipher Phonocardiography signals and have been proven as the milestone in the research of this field. In this chapter we have discussed various important contributions in all important phases of analyzing and deciphering PCG signals.

Chapter 3 is The Heart. Firstly, in this chapter a discussion on the role and importance of the heart in the survival has been discussed. The most important functions of heart have been discussed. Then the anatomy i.e. the structure and major components of heart have been introduced in order to provide familiarity with the terminology. Then the next section is dedicated to the physiology i.e. the functioning of heart. A step wise discussion of physiology has been provided. The subsections include normal and abnormal physiology in order to provide a better insight to the problem. The terms systole and diastole that have been used numerous times in the text have been explained. Then comes the topic heart sounds on which this study is based. The subtopics in this section include the physics of sounds which tells how sounds originates and propagates. Then the theories behind the origin of heart sounds are discussed. Then the normal and abnormal heart sounds and the reasons behind the abnormalities are discussed. The next section is named as Auscultation to Phonocardiography. In this section we tried to establish the need to move from auscultation towards Phonocardiography. Hence we explained some drawbacks of traditional auscultation, salient features of Phonocardiography and also tried to establish the fact that PCG could be the obvious choice when other technologies fail to work. In this chapter we explained the electrical system of the heart and its abnormalities and disorders like arrhythmia.

Chapter 4 is Materials and Methodology. This is dedicated to the work done based on visual differences. First of all the materials used in this methodology are discussed. Then next subsection is

dedicated to the methodology of this approach. Various normal and murmur signals were compared visually and difference between them was established. Based on these findings an algorithm was designed to classify signals as normal and murmur. The next section describes the improvements required in this methodology and the reasons that made further study important. The next section is dedicated to the feature extraction methodology for murmur detection. This methodology finally adopted by us in order to classify heart signals. This methodology is based on feature extraction. In first section materials used for this approach are discussed. In the methodology section, we begin with signal acquisition, followed by feature extraction phase. Then we have the feature reduction phase, which reduces the computational complexities and dimensionality of the data. Then finally we discuss the classification phase which categorizes the signals into two classes namely- murmur and normal. Then the algorithm for arrhythmia detection is explained in the next section in which classification of heart problems because of electrical system disorders like Tachycardia and Bradycardia are discussed.

Chapter 5 is Results, in which the results obtained using the two algorithms are discussed. It tells about the accuracy of these two algorithms.

Chapter 6 is Conclusions and Future Scope. In this we discuss the usage and significance of the algorithms made. It also addresses the issues like benefits of using these algorithms and further improvements that could be done in this field.

CHAPTER

2 LITERATURE REVIEW

Classification of heart sound is not a new topic; however it is still in a developing stage as far as embedded applications are concerned. With the use of latest technology, mobile phones can also be used as an electronic stethoscope and the digital phonocardiogram signals can be easily acquired, saved and transmitted to a cardiologist for further analysis. Most of the research in auscultation field is seen mainly till 1980s but due to other methods like Echocardiography, research trends in PCG decreased but in the past few years it has shown a boost due to improvements in personal computers and signal processing techniques [4-6]. The abrupt frequency changes, complex and highly non-stationary nature of heart sound signals makes heart sound signal analysis a tedious job. The FFT and wavelet approaches have been applied to this in various studies and the work of correlating the heart sounds with the heart defects have been established [7-12]. A lot of work has been done by Rabiner *et al.* In the field of digital signal processing, especially in speech processing area. Some features extracted by him can be applied to various other signals like heart sound signals [13]. Some related works to our study are as mentioned below.

PCG signal acquisition can be done by electronic stethoscope which responds to sound waves identically to conventional acoustic stethoscope with changes in electric field replacing the changes in air pressure. By using electronic stethoscopes the signal can be recorded, replayed and stored for future references and records. The electronic stethoscopes being very expensive and not readily available in developing countries made a very low-cost electronic stethoscope using electronic chest-piece, the amplifier circuit and a PC/laptop [14].

2.1 FEATURE EXTRACTION

Feature extraction is to calculate identifying parameters from each cardiac cycle. In this phase a careful study of the PCG signal is being done and features of the waveform are found. Any specific finding that could have a potential to discriminate between classes could be termed as a feature. Leatham *et al.* (1954) described the presence of an early systolic sound in patients with dilatation of

the pulmonary artery [4]. Since then, numerous reports have shown the systolic ejection click as a fairly constant finding in patients with minimal to moderate pulmonary stenosis.

Muruganantham, *et al.* (2003) derived various features like average power, total power, mean power frequency, median frequency, frequency variance, frequency skewness, frequency kurtosis, and jitter from frequency domain [15].

Shui *et al.* (2004) used wavelet analysis for feature extraction in order to distinguish between normal and aortic stenosis patients. Simultaneous recordings of PCG and ECG were done and ECG acts as a guide to characterize heart sounds. Variable sized windows were used, longer regions where more precise low-frequency information is desired and shorter regions where high-frequency information is analyzed. Local analysis can be performed and the time and frequency information is provided simultaneously. The time-frequency plots of the recordings showed that the frequency range and duration of the murmurs generally increase with stenosis and the second heart sound diminishes with advance degrees of stenosis [16].

Segaier *et al.* (2005) used STFT (Short Time Fourier Transform) for characterization of systolic murmur. A systolic segment between first and second heart sounds (20-70%) was selected for murmur analysis, its average frequency and mean spectral power were quantified [17].

Jiang *et al.* (2006) proposed an analytical model based on a single-DOF is proposed for extracting the cardiac sound characteristic waveforms (CSCW) from the cardiac sounds recorded by an electric stethoscope. The diagnostic parameters are [T1, T2, T11, T12] where T1 and T2 are the widths of the first sound S1 and the second sound S2, T11 is the time interval between two abutted S1, which indicates the heart beat rhythm condition and T12 is the time interval between S1 and S2, which is an indicator to express the heart valvular murmurs. To make the diagnostic parameters visually, a two-dimensional plot, scattergram and histogram of [T1, T2] and [T11, T12] is made [2].

Ahlstrom *et al.* (2005) explained basics of heart sounds [18] and Ahlstrom *et al.* (2006) used tool to be able to investigate how signal content varies over time. Stockwell's TFR formed the basis of this work. Shannon energy was used to measure intensity and a wavelet detail was used to measure intensity in a certain frequency interval. Recurrence points of the first kind, T1, are used to locate S1 and S2 after which S3 is sought in time windows 100-300 ms after the two heart sounds. Hence the tedious task of S3 location was accomplished using recurrence quantification analysis [19].

Noponen *et al.* (2007) combined spectrogram and traditional phonocardiogram to distinguish innocent murmurs from pathological murmurs. This study established that innocent murmurs have

lower frequencies (below 200 Hz) and a frequency spectrum with a more harmonic structure than pathological cases. Innocent systolic murmurs also have a shorter duration than pathological murmurs and always fade before second heart sound [20].

Amit *et al.* (2009) worked on the basis of hierarchical clustering, compact data representation in the feature space of cluster distances [21].

Maglogiannisa *et al.* (2009) proposed a diagnosis system using SVM (Support Vector Machine) to classify heart valve disease [22].

Yuenyong et al. (2011) used DWT (Discrete Wavelet Transform) for feature extraction. In this study heart sounds analysis is done by extracting an equal number of cardiac cycles from heart sounds with different heart rates using information from envelopes of autocorrelation functions without the need to label individual fundamental heart sounds (FHS) [23].

Akbari et al. (2011) used altogether new technique called Digital Subtraction Phonocardiography (DSP). It is based on the principle that murmurs are random in nature while underlying fundamental heart sounds are not (being deterministic). The DSP technique is fundamentally different from other efforts that it starts by constructing a difference signal between two time-adjacent heart cycles, which we herein call a "murmurgram". Two time-adjacent phonocardiogram cycles, PCG-1 and PCG-2, are obtained using the QRS complex of the ECG as a marker for the start of each cardiac cycle. The difference of the two PCG cycles forms a murmurgram. The Murmurgram between FHS are ought to be flat in normal cases whereas murmurgram in abnormal cases is not flat [24].

On the basis of the above discussion, a comparison table has been made as follows:

Table 2-1 Comparison of methodologies of feature extraction

Year	Methodology	Findings and Contributions	ECG Gating used	Reported Performance	Additional Remarks
2004	Wavelet analysis	Aortic Stenosis	Yes	96.4%- severe cases 70.3%- Overall	Frequency, intensity of murmurs and second heart sounds used to categorize the severity of Aortic stenosis
2005	Short Time Fourier Transform (STFT)	Systolic Murmur	Yes	100%- S1 97%- S2	20-70% of systolic murmur used for murmur analysis
2006	Cardiac Sound Characteristic Waveform (CSCW)	Diagnostic parameters T1,T2,T11,T12 Arrhythmia, Mitral Stenosis, Aortic Regurgitation	No		Adaptive THV (Threshold Value) calculated using Fuzzy C-Means (FCM)
2006	Stockwell's method for TFR, Shannon Energy, Wavelet and Recurrence Quantification analysis	S3 detection as indicator of Heart failure	No	98%- S3	S3 sought in time window 100-300 msec after two heart sounds
2007	Phonospectrography	Innocent Murmur characterization in children	No	90%- Sensitivity 91%- Specificity	Innocent murmurs have lower frequencies, shorter duration than pathological cases
2011	DWT (Discrete Wavelet Transform)	Noise robust method	No	92%- Noise-free 90%- with white noise and 10 dB SNR	Without Segmentation
2011	DSP (Digital Subtraction Phonocardiography)	Murmurgram	Yes	Visual differences	Without segmentation, Random plus deterministic PCG model

Various methodologies used for feature extraction showed various advantages and some limitations. The choice of methodology basically depends upon the application. The methodologies which considered both time and frequency domain proved to be more efficient than time and frequency domains alone.

2.2 FEATURE REDUCTON

In the feature extraction phase there could be a large number of features that are either redundant or misleading for the classification phase. This reduces the dimensionality, computation load and increases the accuracy of classification. This method ranks and selects the most important features and only the highest ranked features subset is used for classification. So, this phase comprises of two parts: ranking or feature evaluation and feature selection [25]. Various methods that can be used are as follows:

- Box Plot Method: The spacing between the different parts of the box help in indicating the degree of dispersion (spread) and skewness in the data, and identify the outliers.

- Consistency subset evaluation: This focuses on inconsistency measure according to which a feature subset is inconsistent if there exists at least two instances with same feature values but with different class labels [26].

- Info Gain Attribute evaluation: Entropy is used in information theory and forms the basis of Information gain attribute evaluation [27]. It evaluates the worth of an attribute by measuring the information gain with respect to the class [25].

- Chi Squared Attribute evaluation: It evaluates the worth of a feature by computing the value of the chi-squared statistic with respect to the class [28].

- Filtered Subset evaluation: It comprises of running an arbitrary subset evaluator on data that has been passed through an arbitrary filter [25].

- Gain Ratio Attribute evaluation: It is the non-symmetrical measure that is introduced to compensate for the bias of the Information Gain attribute evaluation [28]. It evaluates the worth of an attribute by measuring the gain ratio with respect to the class [25].

Usually more than one of the above feature reduction methodologies need to be used in order to properly reduce computation time, redundancy and dimensionality.

2.3 CLASSIFICATION

It comprises of assigning a particular class to the signals. A classification situation occurs when an object needs to be assigned in a predefined group or class based on a number of observed attributes or features related to that object. Many problems in business, science, industry, and medicine can be treated as classification problems. This includes training the classifiers and testing them.

• Back-propagation Network:

It has its roots Artificial Neural Networks which are promising alternatives to conventional classifiers. They are data-driven self-adaptive methods and their non-linearity makes them flexible for modelling complex real-world relationships [29]. It has three layers- Input layer, Hidden layer, Output layer. Training inputs are applied to the input layer of the network, and the desired outputs are then compared at the output layer. The difference between the output at output layer and the desired output is back-propagated to the previous layer(s). The back-propagated signals are usually modified by the derivative of the transfer function and the connection weights, which are usually, adjusted using the Delta Rule. The mean square error (MSE) between the output of the network and the desired output is minimized using the gradient descent algorithm [25].

• Support Vector Machines (SVM):

Support Vector Machines have been greatly used in classification and hence has been used for classification of heart sounds. SVMs are basically a set of related supervised learning methods used for classification. They belong to the family of generalized linear classification. SVM simultaneously minimizes the empirical classification error and maximize the geometric margin. So SVM is also called Maximum Margin Classifiers. SVM is based on the Structural risk Minimization (SRM) [29].

• Radial Basis Function (RBF):

Radial Basis Function (RBF) emerged as a variant of Artificial Neural Networks. Due to their non-linearity properties, RBF can model complex mappings which require multiple intermediary layers in perceptron architecture. The main advantages of RBF over feed-forward networks are its accuracy and shorter computational time. The error between the target and the desired output is minimized using a gradient descent algorithm [25].

• Adaptive Neuro-Fuzzy Inference System (ANFIS): The fuzzy logic is rule-based and is modelled on method of human thinking and decision making. On the other hand, ANN learns the problem using its ability of learning and comes through successfully on data sets it did not come across

before. ANFIS was suggested by Jang in 1993 considering the advantages of ANN and fuzzy logic methods. The membership of input/output variables are determined in ANFIS by the use of ANN's ability of learning and the conclusion is reached with the feature of reasoning and decision making of fuzzy logic method [25].

Table 2-2 Comparison of Classifiers

Classifier	**Based on**	**Performance Parameters**	**Accuracy**	**Remarks**
Back propagation	Artificial Neural Network (ANN)	Mean Square Error (MSE) minimization	90.8%	More Complexity
Support Vector Machines (SVM)	Linear Classification	Classification Error minimization and Geometric Margin maximization	95%	Long computation time
Radial Basis Function (RBF)	Artificial Neural Network (ANN)	Mean Square Error (MSE) minimization	98%	Lesser number of layers required than Back-Propagation
Adaptive Neuro-Fuzzy Inference system (ANFIS)	ANN + Fuzzy	Forward pass- Least Squares method Backward pass- Errors propagated backward	98.33%	Membership with ANN and Conclusion with Fuzzy

Usually ANFIS works well and showed highest accuracy but still the choice of classifier mainly depends on the application. The classifier that more accurately meets the intended requirements of the application has to be used.

The existing trends in all phases of signal analysis of the PCG are discussed in order to provide an overview. The methodologies that provided the best of accuracies have also been provided to have a better insight but still the performance of methodologies largely varies according to the problem. Since the PCG is an early indicator of the heart problem, a deep study of this area is required and

hence further researches could be made in this field. It would be an immense contribution to mankind if early detection of heart diseases could be done as this could decrease the magnitude of fatality to a large extent. Some simple lifestyle changes and precautions at an early stage could prevent the suffering caused by worsening of the problem. This potential of PCG, to indicate problems before symptoms appear, makes it an important area of research.

CHAPTER

3 THE HEART

3.1 ROLE AND IMPORTANCE

The heart is one of the vital organs of the body. The importance of this organ can be judged from the fact that heart beat is considered as the sign of life. If it is not beating the person is declared dead. The function of heart is to circulate the blood, oxygenated blood is circulated to the body and deoxygenated blood is circulated to the lungs. So, it acts as the pump in order to provide oxygen to whole body using blood as a medium which is necessary for survival. This oxygen is used by the body to convert carbohydrates of the food into energy which is necessary for proper functioning of the body.

Oxygen + Carbohydrates → Carbon dioxide + Water + ATP

(From atmosphere) (From food) (To atmosphere) (Energy)

$$O_2 + C_6H_{12}O_6 \rightarrow CO_2 + H_2O + ATP$$

3.2 ANATOMY

The heart is contained in the pericardium, a membranous sac consisting of an external layer of dense fibrous tissue and an inner serous layer that surrounds the heart directly [30]. It composes of four chambers and four heart valves as shown in figure 3-1.

The *four chambers* of the heart are:

- Right Atrium
- Right Ventricle
- Left Atrium
- Left Ventricle

The upper chambers are thin walled chambers called the right atrium (RA) and left atrium (LA) and the lower chambers are the right ventricle (RV) and left ventricle (LV). The walls of the left ventricle are approximately three times as thick as the right ventricle to withstand high pressures during contractions for pumping blood to the rest of the body [16].

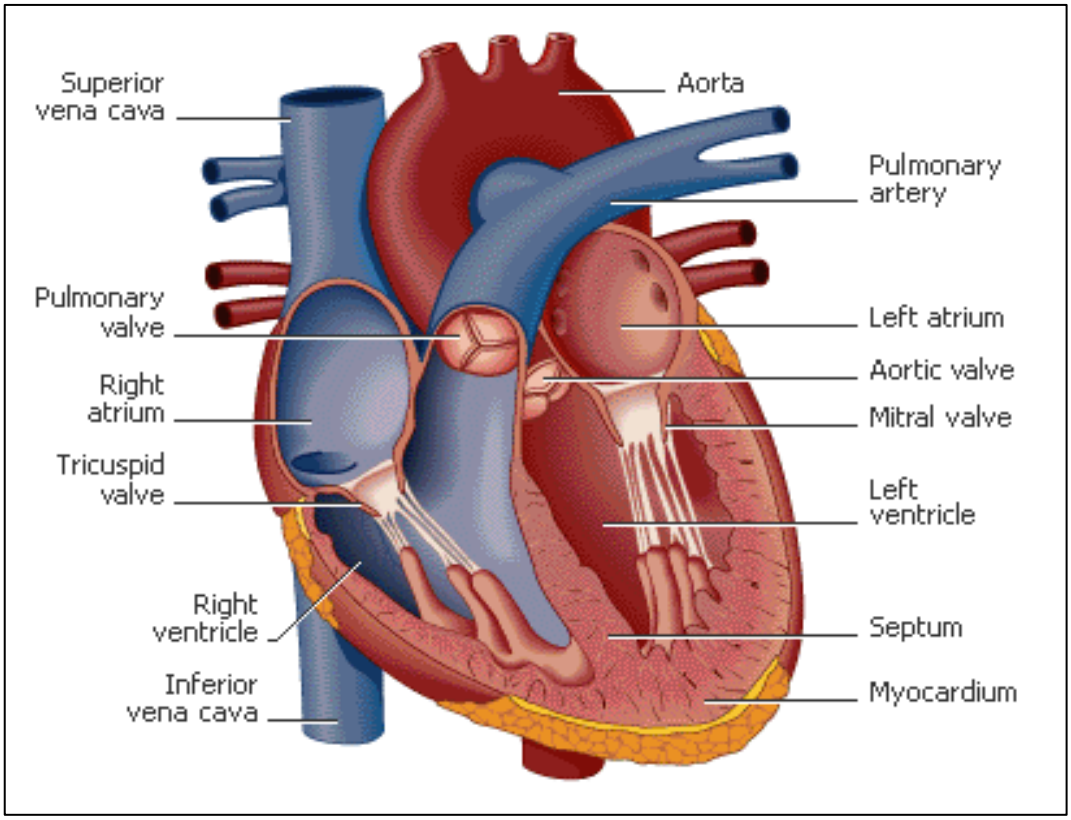


Figure 3-1 The Human Heart [31]

The left atrium is smaller than right atrium. The two sides of the heart are separated by the *septum* which is the dividing wall of tissue.

It also consists of four valves out of which two are atrioventricular valves and two are semilunar valves. These *four valves* are as follows:

- Mitral or bicuspid valve
- Tricuspid valve
- Aortic Valve
- Pulmonary valve

The left atrium is joined to the left ventricle through the *mitral valve*, which is sometimes called as *the bicuspid valve* since it consists of two cusps. Similarly, the right atrium is joined to the right ventricle through *tricuspid valve* [30]. The atrioventricular valve prevents the backward flow of blood from the ventricles to the atria. The *aortic valve* is present between the left ventricle and *aorta* for one way flow of blood i.e. from the left ventricle to aorta in order to supply oxygenated blood to the body. Similarly *pulmonary valve* is present between the right ventricle and *pulmonary artery* for unidirectional flow of blood from right ventricle to pulmonary artery which carries deoxygenated blood to lungs for gaseous exchange.

Blood enters the right atrium through two main veins: *Superior Vena Cava* (which takes blood from the body's upper extremities) and *Inferior Vena Cava* (which takes blood from extremities below the heart).

3.3 PHYSIOLOGY

This section is dedicated to the study of functioning of heart. The heart as we say, functions as a pump. It contracts and relaxes during its functioning as a pump. When the heart contract, blood is forced through the valves, from the atria to the ventricles and eventually out to the body. The two atria mainly act as collecting reservoirs for blood returning to the heart while the two ventricles act as pumps to eject the blood to the body.

3.3.1 NORMAL PHYSIOLOGY

Deoxygenated blood from the body enters the right atrium, passes into the right ventricle and is ejected into the pulmonary artery on the way to the lungs. Oxygenated blood from the lungs re-enter the heart in the left atrium, passes into the left ventricle and is then pumped to the body through the aorta.

The blood pressure within a chamber increases as the heart contracts, generating a flow from higher pressure areas towards lower pressure areas. This is the basis of functioning of heart. As atria contract a high pressure is created in the atria which opens the atrioventricular valves and hence blood rushes to ventricles having low-pressure. On similar lines blood is ejected out of the ventricles and during ventricular contraction atrioventricular valves prevent backflow of blood into the atria.

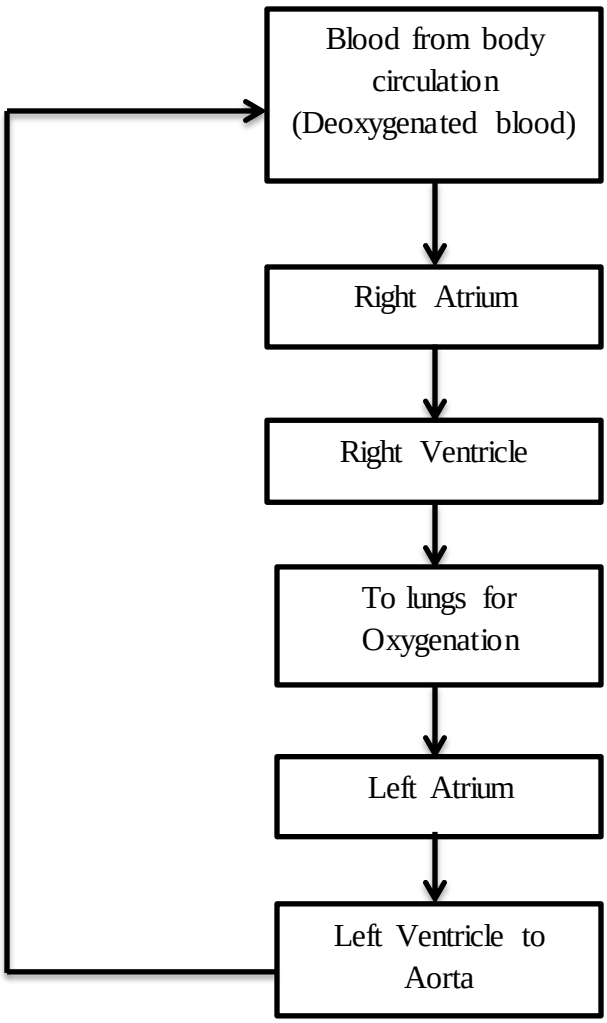


Figure 3-2 Heart Pumping Cycle

TERMINOLOGY- SYSTOLE and DIASTOLE

The heart's pumping cycle is divided into two major parts: systole and diastole.

Systole is defined as a period of contraction of heart muscles, specifically the ventricular muscle, at which time the blood is pumped into the pulmonary artery and the aorta.

Diastole is the period of dilation of the heart cavities as they fill with blood [30].

Systole is smaller than diastole.

During the rapid filling phase (atrial and ventricular diastole), venous blood from the body and from the lungs enters the atria and flows into the ventricles. As the pressure gradient between the atria and the ventricles level out (reduced filling phase), a final volume of blood is forced into the ventricles by atrial contraction (atrial systole).

At the beginning of ventricular systole, all the valves are closed resulting in an isovolumic contraction. When the pressure in the ventricles exceeds the pressure in the blood vessels, the semilunar valves open allowing blood to eject out through the aorta and the pulmonary artery. As the ventricles relax the pressure gradient reverses, the semilunar valves close and a new heat cycle begins [18-19].

3.3.2 ABNORMAL PHYSIOLOGY

The functioning of the heart can be disrupted because of many reasons. The problem could be genetic or developed with time. Deoxygenated blood from the body enters the right atrium, passes into the right ventricle and is ejected into the pulmonary artery on the way to the lungs. Oxygenated blood from the lungs re-enter the heart in the left atrium, passes into the left ventricle and is then pumped to the body through the aorta. The *genetic* problem could be that the septum or muscular partitioning of the ventricles could have a hole. So, mixing of blood from left (oxygenated) and right (deoxygenated) ventricles take place and hence causes improper functioning of the heart. This is called septal defect.

The *non-genetic* problem could be due to calcination of valves or due to cholesterol deposition and hence causes thickening of the valves. This impedes the normal blood flow and is a major cause of acquired heart diseases. Leaking valve could also be the reason of heart disease. Sometimes the valves don't close properly and blood leaks back into the atrium. This impedes the normal functioning of heart.

3.4 MECHANICAL SYSTEM OF HEART - HEART SOUNDS

3.4.1 PHYSICS OF SOUND

Before studying the heart sounds it is important to study the basic underlying acoustic phenomenon or what we say the physics of sound. A sound is generated by a vibrating object and propagates as a wave of alternating pressure. The vibrating source sets the particles in motion with the frequency of that tone. Each particle is thus moving around its resting point, but as it pushes nearby particles they are also set in motion and this chain effect result in areas of compression and rarefactions. The alternating areas of compression and rarefaction constitute a pressure wave that moves away from the sound source. These pressure variations can be detected via the mechanical effect they exert on some membrane (the diaphragm of the stethoscope) [18-19].

3.4.2 THEORIES BEHIND ORIGIN OF HEART SOUNDS

There are two theories behind the origin of heart sounds-

a.) The valvular theory states that heart sounds emanate from a point source located near the valves. This assumption is probably an oversimplification.

b.) In the cardiohemic theory heart and blood represents an interdependent system that vibrates as a whole [6].

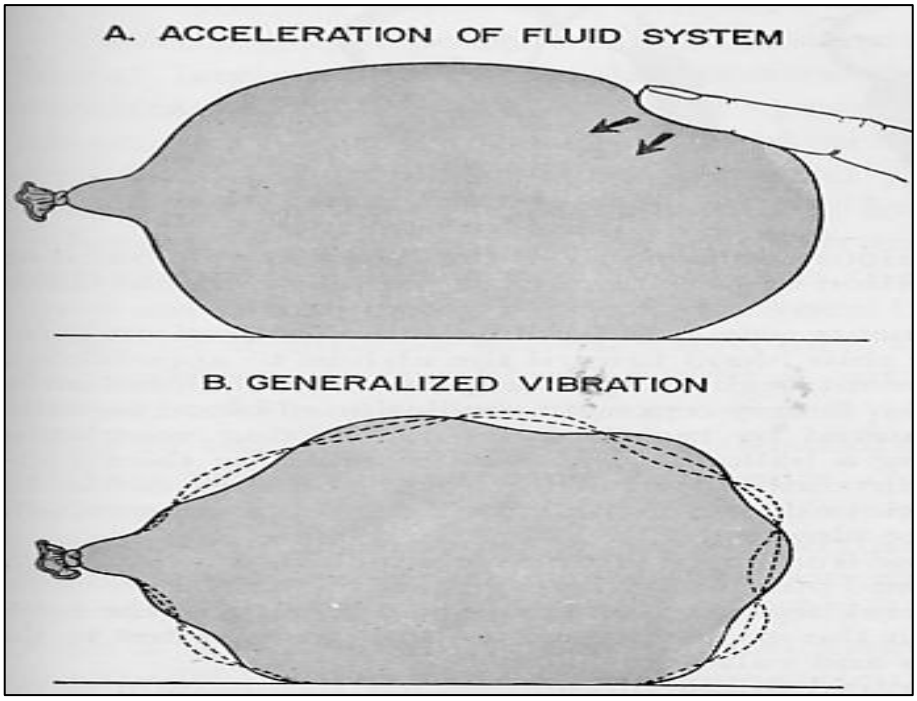


Figure 3-3 Basis of Heart Sounds [32]

3.4.3 NORMAL HEART SOUNDS

The blood flow is normally laminar in nature but when the blood passes through a small opening or has some obstruction in its path then the flow becomes turbulent. That's the reason behind the fact that when blood passes through the valves which are small openings, the blood flow becomes turbulent. Then because of this turbulence along with the vibrations caused by closure of heart valves results in the vibrations of the heart and the blood as an interdependent system. These vibrations according to cardiohemic theory are believed to be the basis of heart sounds. The normal heart sounds are:

S1: The heart sound that occurs with ventricular systole and is produced mainly by closure of the atrioventricular valves.

S2: The heart sound that signifies the beginning of diastole and is caused by closure of the semi lunar valves.

S3: The heart sound that occurs in early diastole and corresponds with the first phase of rapid ventricular filling.

S4: The heart sound occurring in late diastole, corresponding with atrial contraction [33].

The region between S1 and S2 is termed as systole. The region between S2 and next cycle's S1 is termed as diastole. Systole is smaller than diastole.

Out of these four the dominant two are S1 and S2 and are commonly known as Fundamental Heart Sounds (FHS). These two sounds are mainly audible sounds and other two are usually not audible. These two sounds are called 'lub' and 'dub' in common terminology. In medicine we call the lub sound as 'S1' and the dub sound as 'S2'. The normal heart rates at rest are 60 to 100 beats ('lub dub's) per minute.

The locations of S1 and S2 over time are as follows:

S1…….S2…….S1…….S2…….S1…….S2….

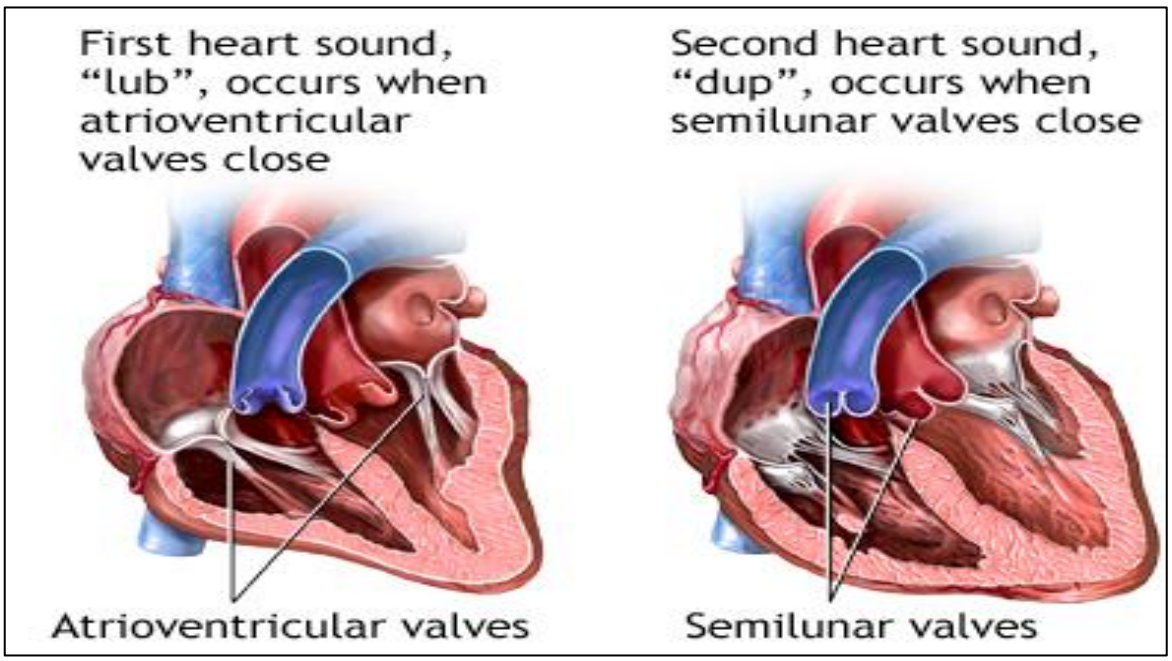


Figure 3-4 Mechanical actions producing Fundamental Heart Sounds [34]

3.4.4 ABNORMAL HEART SOUNDS

Sometimes other than the normal heart sounds some extra sounds are also heard. These may be evident because of some heart problems. Some abnormal heart sounds are as follows:

MURMURS

These are long strings of noise that can't be termed as a single sound is termed as Murmurs. Murmurs are caused by the blood turbulence which is capable of producing a sound that can be heard using a stethoscope. The murmurs can be termed as the indicators to various heart problems. Logically also if some extra blood turbulence is created that means the blood is flowing through a few extra small openings, other than just passing through the valves. The locations of murmur can be as shown:

S1*****S2……..S1*****S2……..S1*****S2

This is ***systolic murmur***. Since the region between S1 and S2 is called systole and here the murmur is appearing in this region so it is termed as systolic murmur.

S1…….S2**********S1……..S2**********S1…….S2

This is ***diastolic murmur***. Since the region between S2 and next cycle's S1 is called diastole and here the murmur is appearing in this region so it is termed as diastolic murmur.

In the description above shown the asterisks (*) denotes murmur.

Causes of murmurs

The problem causing murmurs could be congenital or developed with time. The ***congenital*** problem could be that the septum or muscular partitioning of the ventricles could have a hole. So, mixing of blood from left (oxygenated) and right (deoxygenated) ventricles take place and hence causes improper functioning of the heart. This is called septal defect. This leaking of blood from small hole present on the septum causes blood turbulence and hence the murmurs.

The ***non-genetic*** problem could be due to calcination of valves or due to cholesterol deposition and hence causes thickening of the valves. This impedes the normal blood flow and is a major cause of acquired heart diseases. Leaking valve could also be the reason of heart disease. Sometimes the valves don't close properly and blood leaks back into the atrium. This impedes the normal functioning of heart. These reasons can also cause blood turbulences and hence the murmurs.

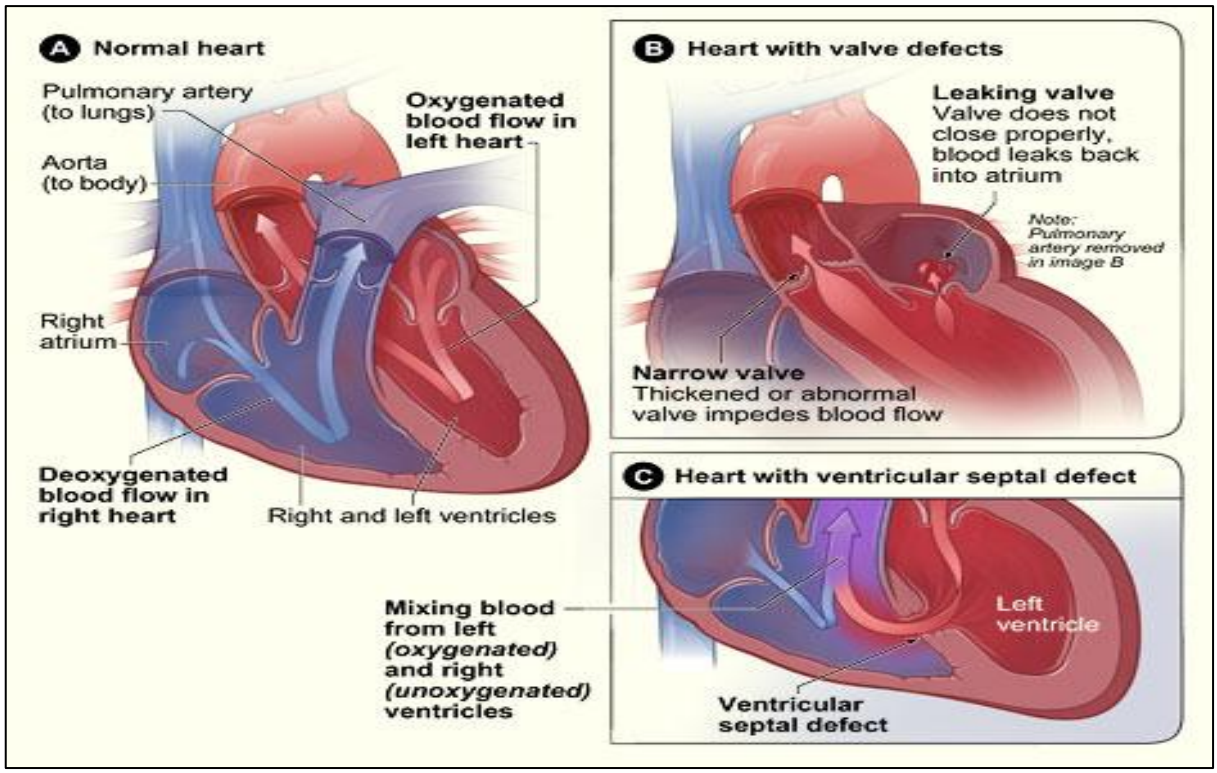


Figure 3-5 Causes of Heart Murmurs [35]

CLICKS

These appear during systole. They are called clicks because they are extremely short duration sounds. These can be differentiated from the murmurs which are longer in duration and can appear both in systole and diastole.

Types of clicks

Early systolic: Just after S1 high pitched click is heard in association with the "opening snap" of the semilunar valves which are mildly to moderately stenotic.

Mid-systolic click: A medium pitched variable sound is commonly heard in association with mitral valve prolapse.

3.5 AUSCULTATION TO PHONOCARDIOGRAPHY

3.5.1 AUSCULTATION

Auscultation is basically the act of analyzing sounds in the body that are produced in response to mechanical vibrations generated in the organs. It is an art and science of listening to the heart sounds and properly analyzing them to find the underlying problem. It requires immense experience and practice in order to properly diagnose the problem just by listening to the heart sounds. There are specific areas where the heart sounds are best heard, which are called auscultation sites.

The important auscultation sites are as follows as shown in figure 3-6:

- *Mitral area:* The cardiac apex.
- *Tricuspid area:* The fourth and fifth intercostal space along the left sternal border.
- *Aortic area:* The second intercostal space along the right sternal border.
- *Pulmonic area:* The second intercostal space along the left sternal border.

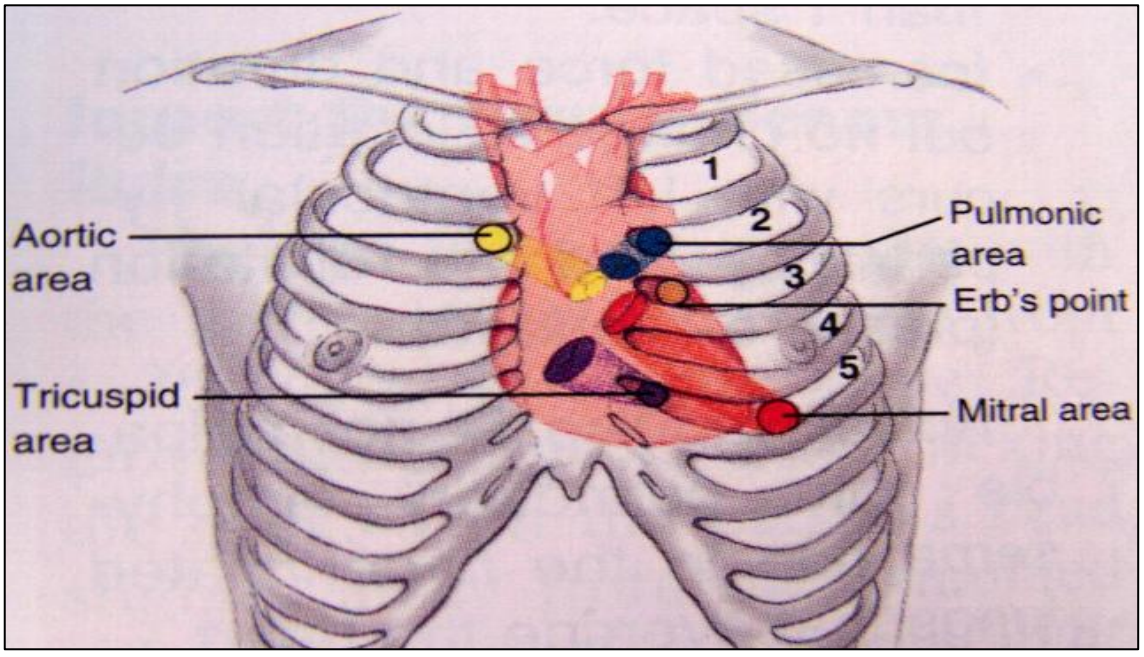


Figure 3-6 Auscultation Sites [36]

The particular heart sounds are best heard at auscultation sites which are decided on the basis of their area of origin. For example, the low-pitched S1 is best heard at the mitral auscultation site as it is

caused by closure of mitral and tricuspid valves and shorter duration but louder sound S2 is caused by aortic and pulmonary valve closure and is best heard over the aortic auscultation area.

EQUIPMENT REQUIRED FOR AUSCULTATION

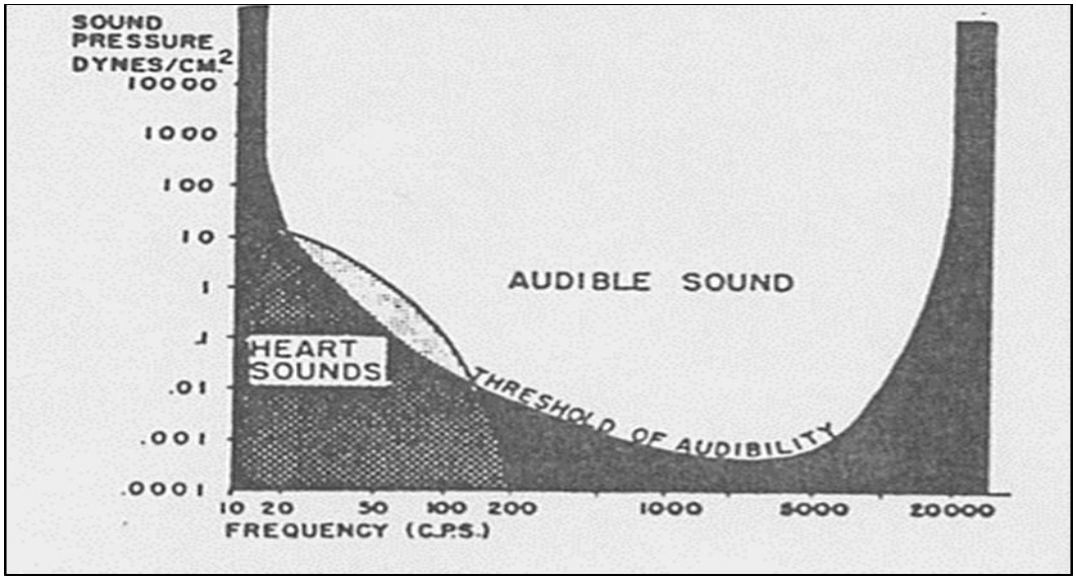


Figure 3-7 Audibility of Heart Sounds [32]

Heart sounds are basically low frequency low-pitched sounds. As shown in figure 3-7, only a small portion of heart sounds lie in the human audibility range. This means our ears are unable to properly hear these sounds and hence we will not be able to analyze the minute details prevalent in order to diagnose the patient properly. So, we need to amplify these sounds in order to hear and interpret properly and hence we use the instrument named ***stethoscope***.

Figure 3-8 Stethoscope [37]

The stethoscope is an instrument with a chest piece having a diaphragm or a membrane, which when placed at the proper auscultation site responds to heart sounds by vibrating. This creates a pressure wave which passes through binaural flexible tubing to the ears of the physician. This is acoustic stethoscope. But now-a-days ***electronic stethoscope*** is used which responds to the sound waves identically to conventional acoustic stethoscope with changes in electric field replacing changes in air pressure.

3.5.2 MOVING TOWARDS PHONOCARDIOGRAPHY

As stated earlier, Phonocardiography etymologically means the process of recording of Heart Sounds. The heart sounds recorded by an electronic stethoscope are converted to digital signals and are plotted as in figure 1-1 and termed as PCG (phonocardiograph) signal. Auscultation is an age-old science but modern day researches are based on Phonocardiography to improve its credibility in order to move towards automatic auscultation. There is a need to shift from conventional auscultation to phonocardiography because of many reasons.

DRAWBACKS OF AUSCULTATION

- As already stated that only a small portion of heart sounds lies in the human audibility range. So, there is a great possibility of missing out the minute details which may be required for proper diagnosis.
- It requires immense training and experience of the physician in order to properly diagnose the patient just on the basis of hearing the heart sounds.
- The storage of records for follow-ups and future references is not possible
- Sometimes physicians miss out some important sounds or might consider pathological murmurs as innocent murmurs.
- Noisy environments can lead to wrong diagnosis.
- Telemedicine is not possible with conventional auscultation as there is no means to transfer the heart sounds to an expert sitting miles away.

SALIENT FEATURES OF PHONOCARDIOGRAPHY

- As heart sounds and murmurs have very less overlap with human audibility range so the minute details that can be missed during auscultation can be best viewed and taken care of with the help of Phonocardiography. This is because showing heart sounds graphically have nothing to do with audibility range. This helps in better diagnosis.

- With the help of phonocardiography we are gradually moving towards automatic auscultation, which means software would be sufficient to tell state of heart. This eradicates the limits of immense training and experience of the physician to diagnose the patient properly.

- The storage of records for future references and follow-ups is possible.

- Telemedicine is possible as heart sounds recordings can be sent to expert sitting miles away in digital form via electronic media.

- Heart sounds are an early indicator of heart problems. So, PCG signal when properly analyzed can lead to cost-effective and efficient treatment at an early stage.

- Use of PCG will be extremely beneficial in the case of rural primary health care centers where stethoscope is the only available instrument for diagnosis. By using electronic stethoscope the automatic auscultation would be possible. Advice from an expert can also be taken by sending PCG signals via electronic media.

- It is sometimes also the only available option as in the case of infants where implementation of ECG and other techniques cannot be implemented.

All these above factors have been an immense driving force in order to work in the field of Phonocardiography. This study is an attempt to contribute in this field and the vision is to provide proper diagnosis to the mankind and hence improving the quality of living.

3.6 ELECTRICAL SYSTEM OF HEART

The heart beat (contraction) begins when an electrical impulse from the sinoatrial node (also called the SA node or sinus node) moves through it. The SA node is sometimes referred to as the heart's "natural pacemaker" because it initiates impulses for the heartbeat.

The normal electrical sequence begins in the right atrium and spreads throughout the atria to the atrioventricular (AV) node. From the AV node, electrical impulses travel down a group of specialized fibers Bundle of His and Punkinje fibers called the His-Purkinje system to all parts of the ventricles.

This exact route must be followed for the heart to pump properly. As long as the electrical impulse is transmitted normally, the heart pumps and beats at a regular pace [38].

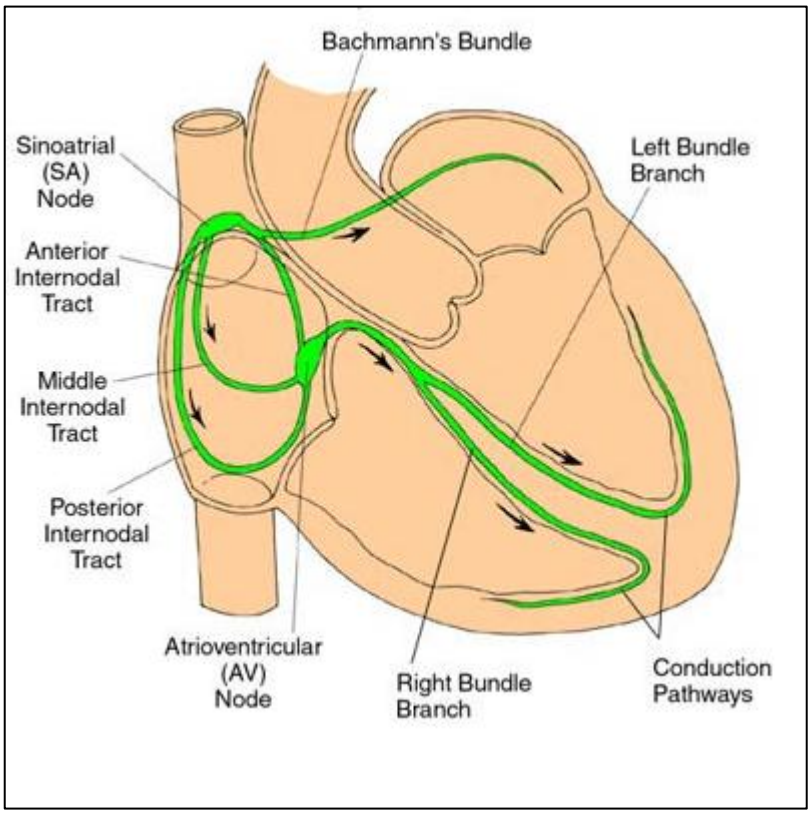


Figure 3-9 Electrical System of Heart [39]

If the functioning of the electrical system of the heart is proper, the heart rhythm or pace will be normal but if the problem persists in this system, then arrhythmia will be diagnosed in the patient.

3.6.1 ARRHYTHMIA

Arrhythmia as the name suggests, is the problem with the rhythm or rate of the heart. 'Arrhythmia' is A-Rhythmia i.e. without proper rhythm. The term "arrhythmia" refers to any change from the normal sequence of electrical impulses. The electrical impulses may happen too fast, too slowly, or erratically – causing the heart to beat too fast, too slowly, or erratically. Most arrhythmias are

harmless, but some can be serious or even life threatening. When the heart doesn't beat properly, it can't pump blood effectively. When the heart doesn't pump blood effectively, the lungs, brain and all other organs can't work properly and may shut down or be damaged. A heartbeat that is too fast is called tachycardia. A heartbeat that is too slow is called Bradycardia [2-3].

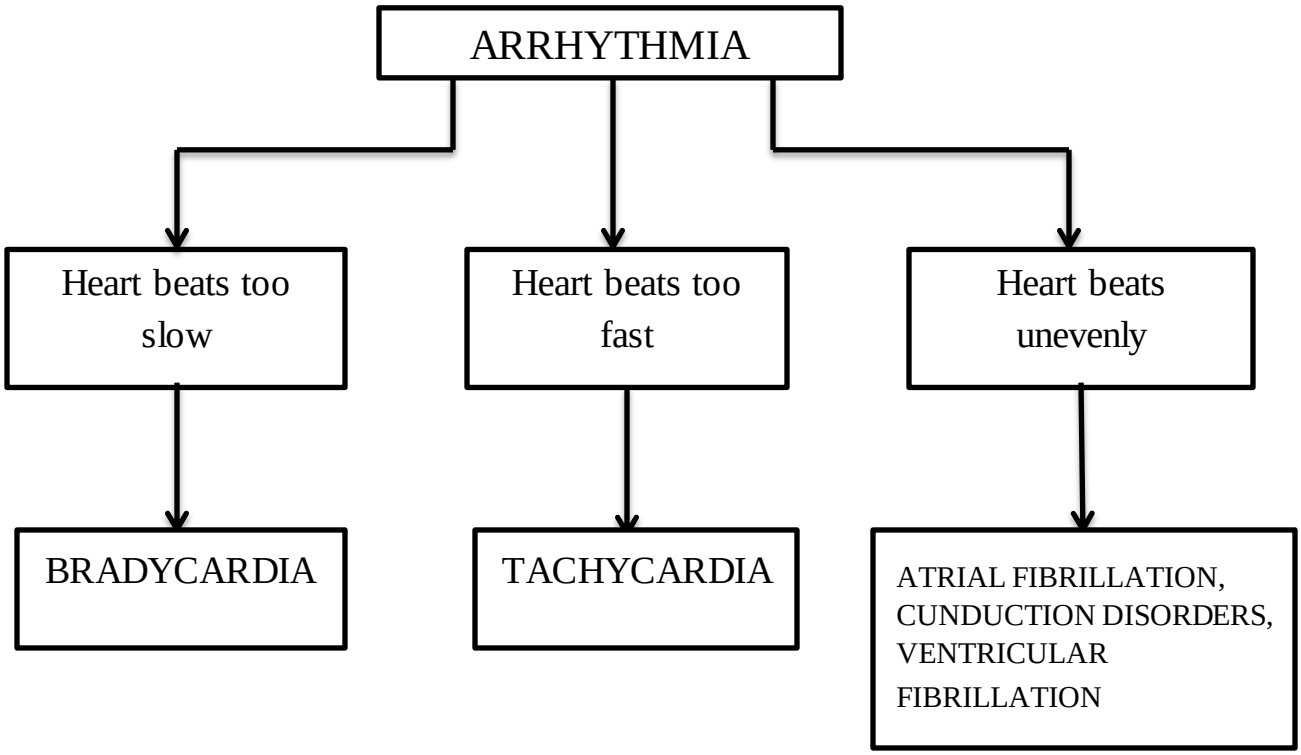


Figure 3-10 Arrhythmia and its types

The normal heart rate of an adult is 60 to 100 beats per minute i.e. 60 to 100 bpm. The heart rates vary with age. The infants and children have higher heart rates. Trained athletes tend to have lower resting heart rates so that when they exert and practice they might not feel breathlessness and proper blood circulation is maintained without much strain during training. The world record of lowest resting heart rate in a healthy human is 27 bpm.

Arrhythmia occurs due to the problems in the electrical system of the heart. Since the electrical activity of the heart is analyzed by test named as Electrocardiography (ECG), hence Arrhythmias are mainly detected using ECG.

3.6.2 IRREGULARITIES IN ELECTRICAL SYSTEM CAUSING ARRHYTHMIA

- An arrhythmia can occur if the electrical signals that control the heartbeat are delayed or blocked. This can happen if the special nerve cells that produce electrical signals don't work properly. This simply means that there is some problem in Sinoatrial Node (SA Node).
- It also can happen if the electrical signals don't travel normally through the heart. This means that there is some problem in the conduction pathway of the heart.
- An arrhythmia also can occur if another part of the heart starts to produce electrical signals. This adds to the signals from the special nerve cells and disrupts the normal heartbeat.

3.6.3 CAUSES

There can be many causes for arrhythmia. They can be broadly classified into two categories namely- Congenital and Acquired Arrhythmia.

CONGENITAL HEART DEFECTS

In some patients arrhythmia is reported due to congenital heart defects. It simply means that the defect is present from birth.

ACQUIRED ARRHYTHMIA

It is not present from birth due to a defect but acquired with time. It may be because of one or more reasons listed below:

- Smoking
- Heavy alcohol use
- Use of prohibited drugs
- Too much caffeine or nicotine
- Heart attack
- High Blood Pressure
- Heart Failure

3.6.4 TYPES OF ARRHYTHMIA

The types of arrhythmia are as shown:

- Tachycardia- very fast heart rate
- Bradycardia- slow heart rate
- Atrial Fibrillation- upper heart chambers contract irregularly
- Conduction Disorders- heart does not beat normally

- Premature contraction- early heart beat
- Ventricular Fibrillation- disorganized contraction of the lower chambers of the heart.

The two basic kinds of arrhythmia are- Tachycardia and Bradycardia.

3.6.5 TACHYCARDIA

Etymologically, Tachycardia comes from the Greek words-

Tachys means rapid or accelerated

Kardia means of the heart.

So, it simply means the rapid or accelerated state of heart.

Tachycardia is the heart rate that exceeds the normal range. A resting heart rate over 100 beats per minute is generally accepted as tachycardia. The causes of Tachycardia are often benign. Tachycardia can be dangerous depending on the speed and type of rhythm.

As the heart beats abnormally fast in tachycardia, leading in too much stress on the heart. The proper supply of blood cannot be provided to the body and also to the heart. Hence can sometimes be life threatening. If tachycardia is severe or pathologic, then it is more correctly termed as ***tachyarrhythmia***.

CAUSES OF TACHYCARDIA

- Abnormal electrical pathway
- Damage to heart tissues due to heart disease
- Smoking
- Alcohol overuse
- Due to some drugs

3.6.6 BRADYCARDIA

Etymologically, Bradycardia comes from the Greek words-

Bradys means slow or retarded

Kardia means of the heart.

So, it simply means the slowness of heart.

Bradycardia is a heart rate that is below the normal range. A resting heart rate below 60 beats per minute is termed as bradycardia. Bradycardia during sleep is considered normal and rates around 40-50 bpm is usual. Trained athletes may also have slow resting heart rates to prevent tachycardia during training.

In severe forms of bradycardia heart pumps less. So, less pumping of heart means a lesser supply of blood and oxygen to the body to meet its needs, which can be life-threatening. Generally bradycardia is not symptomatic until the heart rate reaches below 50 bpm. If the cardiac output decreases dramatically, symptoms appear or the bradycardia reaches a severe stage then it is referred as ***bradyarrhythmia***.

CAUSES OF BRADYCARDIA

The causes of bradycardia can be:

- Prohibited drug usage can make heart pumping action slow
- SA (Sinoatrial) Node disorders
- Blockage of conduction pathway
- Damage to heart because of some disease like heart attack.

CHAPTER

4 MATERIALS AND METHODOLOGY

4.1 MURMUR DETECTION BASED ON VISUAL DIFFERENCE

Stage I of our work was to focus on the visual differences of normal and murmur signals. There were striking differences in the magnitudes of normal and murmur signals as shown in figure 4-1. The visual differences of the signals were noticed and we tried to generalize our findings.

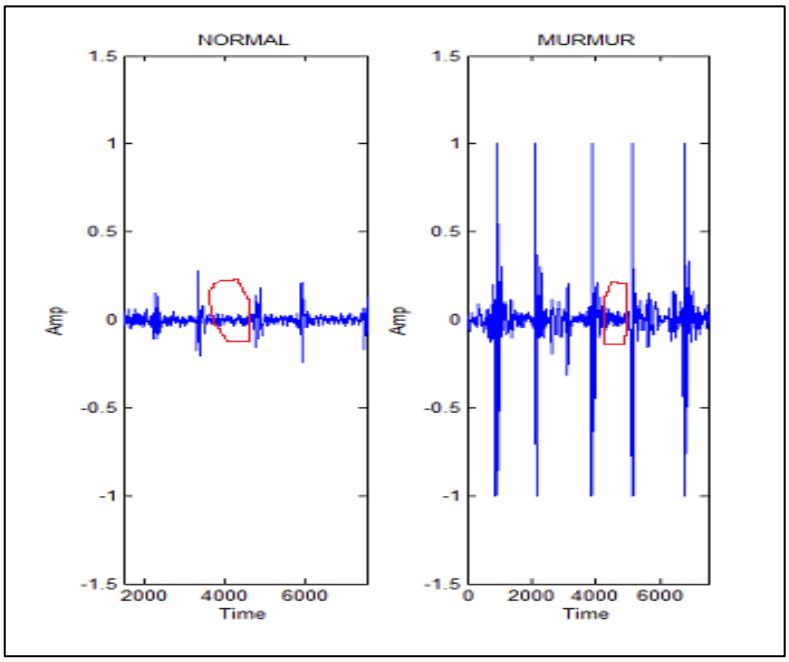


Figure 4-1 Differences in magnitude between peaks of S1 and S2

4.1.1 MATERIALS USED

The database is taken from:

www.peterjbentley.com [40]. Initially in at this preliminary stage we used 44 signals. Out of which 21 were normal and 23 were murmur signals.

- MATLAB version 7.12.635 (R2011a) was used for analyzing the signals. We developed an algorithm and code was written in MATLAB to classify signals into two classes namely- NORMAL and MURMUR.

4.1.2 METHODOLOGY

As striking differences were seen in the amplitudes of normal and murmur signals, this formed the basis in this stage. We found more profound differences in amplitude between the regions of the peaks of S1 and S2. In normal signals, the amplitude in between peaks was nearly zero but in case of murmur the amplitude was larger. So, we selected the region in between these peaks for the analysis. The methodology of selection of the in-between regions of peaks and hence the comparison of their amplitudes led to the algorithm described in the next section.

4.1.3 ALGORITHM

Step1: Peak is located on the signal by setting a threshold. Every peak lies above that value.

Step2: Then a few samples consisting of peak are skipped.

Step3: Hence we reach the in-between region.

Step 4: Then values of a few samples are taken in this in-between region and their mean is calculated.

Step4: If the mean exceeds a threshold value then it is a murmur otherwise it is a normal heart sound. This is because the values of the mean of amplitudes should be higher for murmur than for normal signals.

4.1.4 IMPROVEMENTS REQUIRED

- Accuracy needs to be improved.
- This work was based on only one parameter for classification. We aim at finding out more parameters in order to properly classify the signals.
- We aim at incorporating such a mechanism which takes care of both the cases i.e. Systolic murmur and Diastolic murmur.

4.2 MURMUR DETECTION BASED ON FEATURE EXTRACTION

4.2.1 MATERIALS USED

- The database is taken from :

  www.peterjbentley.com [40].

- Microsoft Excel 2010 was used. We used excel sheet to record the values of various features for 60 signals taken as database.

- MATLAB version 7.12.635 (R2011a) was used for analyzing the signals. We extracted the features of the signals and code for classification was written in MATLAB in order to classify the signals into two classes namely- NORMAL and MURMUR.

- Spectrum Analysers Spectra Plus SC and Sigview are used for analyzing signals and finding values of few features.

4.2.2 METHODOLOGY

For this study we proceeded in a stepwise manner as shown below.

- We began with *signal acquisition* which involved the phase of selecting the database for the study.

- Then the *feature extraction* phase. This includes selecting the features of signals which have the potential to discriminate and classify the signals into two categories namely- normal and murmur signals.

- The next phase is of *feature reduction*. To reduce the dimensionality and to remove the redundant and misleading features, this phase is very important for proper classification.

- Then comes the *classification phase*. This phase comprises of assigning a predefined class to the signals.

SIGNAL ACQUISITION

PCG signal acquisition has to be done with an electronic stethoscope rather than the traditional acoustic stethoscope. The electronic stethoscope responds to sound waves identically to conventional acoustic stethoscope with changes in electric field replaced with the changes in air pressure.

The electronic stethoscopes being very expensive is not readily available in developing countries. Shervegar *et al.* (2011) made a very low-cost electronic stethoscope using electronic chest-piece, the amplifier circuit and a PC/laptop [10]. This circuit could easily work as the expensive electronic stethoscope available in the market.

However, for this research, an electronic database of PCG signals is taken from www.peterjbentley.com [40]. The dataset used is taken from clinical trials in hospitals using the digital stethoscope DigiScope®.

FEATURE EXTRACTION

Everything in the nature is correlated and establishing this correlation is the main research motive. We believed that there will be the features that will correlate the signals belonging to the same class. The missing link was to find those features having potential to correlate signals of the same class and hence differentiating between signals of different classes.

Feature extraction is to calculate identifying parameters from each cardiac cycle. Any specific finding that could have a potential to discriminate between classes could be termed as a feature.

In this phase a careful study of the PCG signal is being done and features of the waveform are found. In this study total 23 features were extracted. These features are from various domains including time domain, frequency domain, cepstrum and statistical features that could have the potential to discriminate between normal and murmur signals. In this study we used the common assumption that systole is shorter than diastole.

Unlike other studies, in this study we have also evaluated the features in the whole signal as well as separately in systolic and diastolic regions. This is done in order to incorporate both the possibilities i.e. possibility of systolic murmur or diastolic murmur. The features that were extracted were compiled on an excel sheet in order to do further processing and analysis.

TERMINOLOGY

In the terminology used to name the features, the suffix 1 shows the feature is extracted in systole and the suffix 2 shows the feature is extracted in diastole. The various examples of this terminology are explained in this section.

The features that were extracted are as follows:

Peak Frequency: It shows the frequency at which the peak amplitude occurs. Since the murmurs and normal signals vary in amplitude and frequency, it could have been a potential feature.

Peak Frequency 1: It shows the frequency at which the peak amplitude occurs during systole. A higher peak frequency 1 is expected for murmur signals as compared to normal signals.

Peak Frequency 2: Similarly, this also shows the frequency at which the peak amplitude occurs during diastole.

Peak Amplitude: It shows the peak value of the signal. The murmur and normal signals vary in amplitude so this feature had potential.

Peak amplitude 1: It shows the peak value of the signal during systole. A higher Peak amplitude 1 values are expected for murmur signals as compared to normal signals. This is because in-between the peaks of S1 and S2 generally a higher amplitude was observed during visual analysis of signals that depicted the murmurs present in the signal.

Peak Amplitude 2: Similarly, this shows the peak value of the signal during diastole.

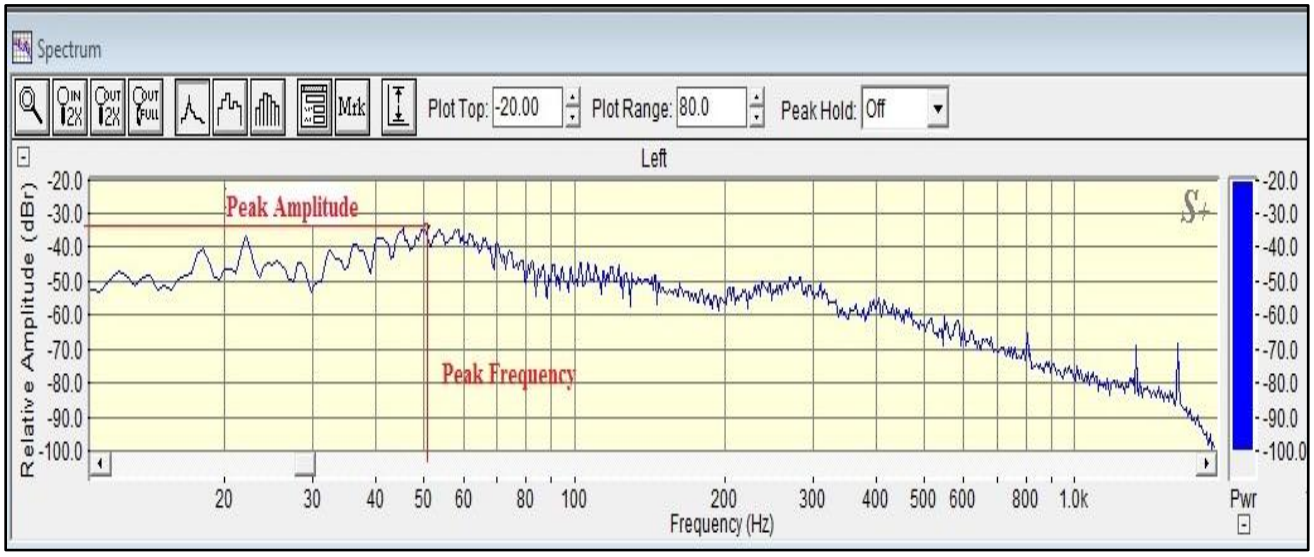


Figure 4-2 Features-Peak Amplitude and Peak Frequency

Total Power: The feature shows the total power of the signal. The murmur is a higher amplitude signals, is expected to have a higher value of this feature.

Total Power 1: It is the total power of the signal in systole region. The presence of murmur just after S1 i.e. in the region of systole yields a probability of higher value of this feature for murmur signals.

Total Power 2: Similarly, it is the total power of the signal in the region of diastole.

Total Harmonic Distortion: The total harmonic distortion of a signal is the measurement of the harmonic distortion present and is defined as the ratio of the sum of the powers of all harmonic components to the power of the fundamental frequency. A comparatively higher THD is expected for murmur signals.

Bandwidth: It is the difference between the upper and lower frequencies in a continuous set of frequencies. As murmurs are high in frequency so a higher bandwidth is expected in case of murmur signals.

Bandwidth 1: It is the bandwidth during the systole.

Bandwidth 2: It is the bandwidth during the diastole.

Q-Factor: A higher Q-Factor indicates that oscillations die out more slowly. So, a higher Q-Factor is expected in case of murmur signals.

Q-Factor1: It is the Q-factor during the duration of systole. A higher Q-Factor indicates that oscillations die out more slowly. Since murmurs can be present after S1 so a higher Q-factor can be expected during systole.

Q-Factor 2: Similarly, it is the Q-factor during the duration of diastole.

Cepstrum Peak Amplitude: The main information is concentrated in the starting of cepstrum and we observed the peak values at the start. Hence we took Cepstrum Peak Amplitude as a potential feature.

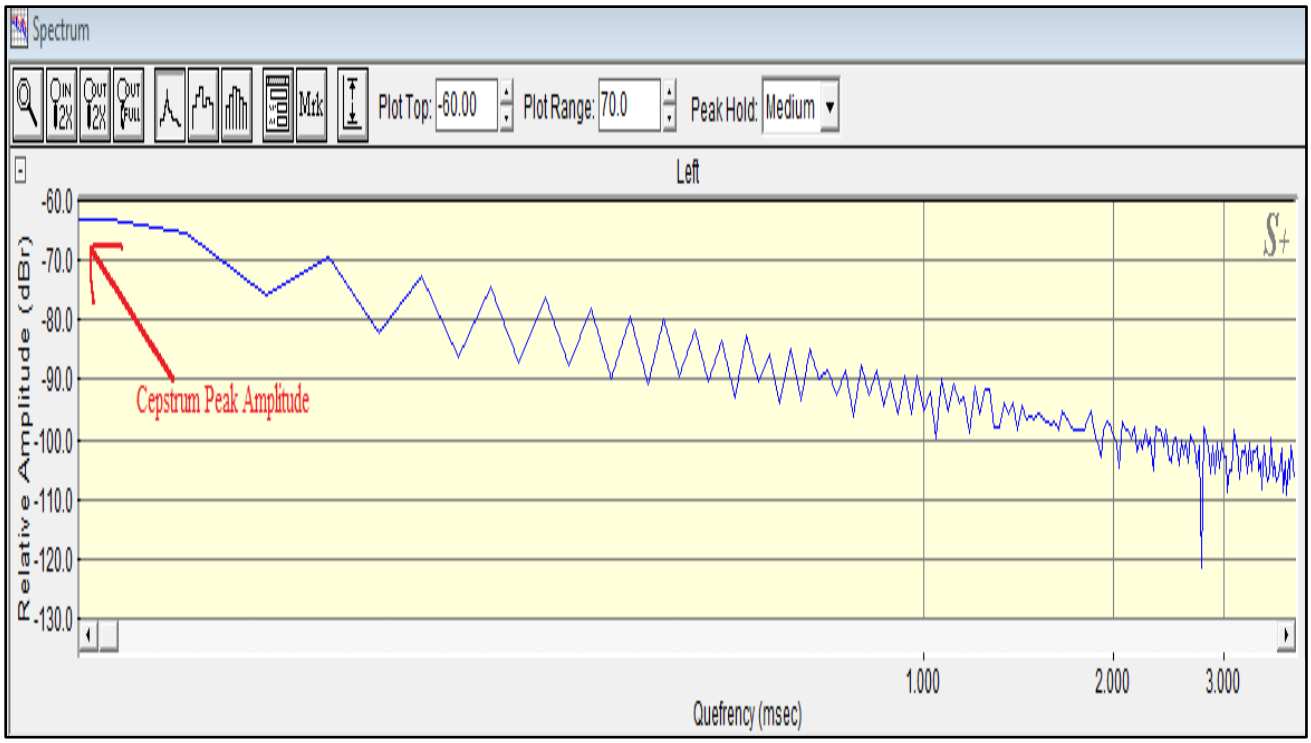


Figure 4-3 Features- Cepstrum Peak Amplitude

Mean 12: In most of the murmur signals visible differences were seen in the regions between S1 and S2 or S2 and next cycle's S1. So, the mean was found in these regions and maximum of these two values is considered as Mean12. This is done in order to explore both the possibilities i.e. systolic murmur and diastolic murmur. A higher Mean 12 is expected for murmur signals.

Zero Crossing Rate: The ZCR is the rate of sign-changes along a signal, i.e., the rate at which the signal changes from positive to negative or back. A larger value of ZCR is expected for murmur signals.

t1: It is the time duration of S1. Normally, S1 are shorter duration sounds but in case of murmurs considerably longer S1could be expected. This is because murmurs could appear just after S1 and can make S1 look longer than normal S1 sound.

t2: It is time duration of S2. It is longer than S1. Considerably a longer S2 could be expected for murmur signal than for the normal signal. This is because the murmur could appear just after S2 to make it look longer for murmur signals.

t12: It is the time duration from end of S1 to start of S2. As explained earlier that a murmur could appear just after S1 and hence the region between the end of S1 to start of S2 could become smaller. So, in case of murmurs smaller t12 is expected.

t21: It is the time duration from end of S2 to start of next cycle's S1. As the murmur could appear just after S2 hence the region between the end of S2 to start of next cycle's S1 would look smaller.

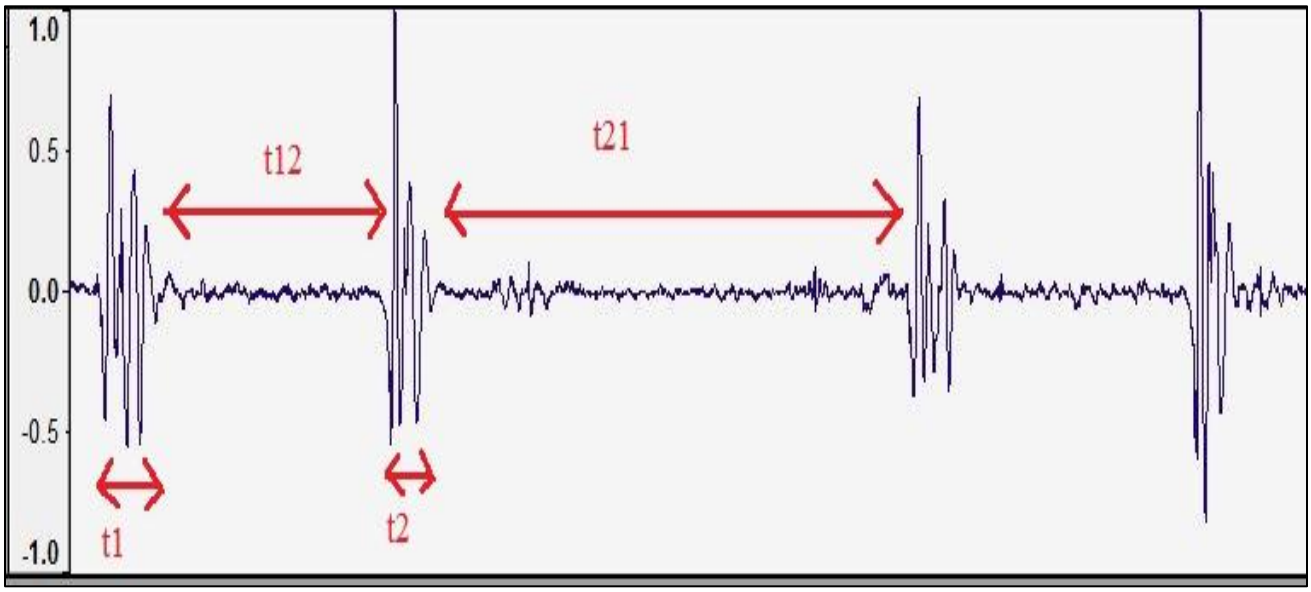


Figure 4-4 Features- t1, t2. t12, t21

The 23 features that have been evaluated for all signals are enlisted below along with the domains to which they belong and the source of that feature.

Table 4-1 List of features extracted for classification

S.No	Feature	Feature domain	Source
1	Peak Frequency	Frequency domain	[18]
2	Peak Amplitude	Time domain	[24]
3	Total Power	Time domain	[18]
4	Total Harmonic Distortion (THD)	Frequency domain	
5	Bandwidth	Frequency domain	[18]
6	Q-Factor	Frequency domain	
7	Cepstrum Peak Amplitude	Cepstrum	
8	Peak Frequency 1	Frequency domain	
9	Peak amplitude 1	Time domain	
10	Total Power 1	Time domain	
11	Bandwidth 1	Frequency domain	
12	Q-Factor 1	Frequency domain	
13	Peak Frequency 2	Frequency domain	
14	Peak Amplitude 2	Time domain	
15	Total Power 2	Time domain	
16	Bandwidth 2	Frequency domain	
17	Q-Factor 2	Frequency domain	
18	Mean 12	Statistical	
19	Zero Crossing Rate (ZCR)	Time domain	[16]
20	t1	Time domain	[5]
21	t2	Time domain	[5]
22	t12	Time domain	[5]
23	t21	Time domain	[5]

FEATURE REDUCTION

In this phase the redundant and misleading features have to be reduced and only significant features have to be retained for classification. This reduces the computational cost and makes the algorithm time efficient. The final algorithm is to have a minimum number of features and should have maximum accuracy. So, we rank the features and only the best features are used for classification. First of all Ranker and Info Gain Attribute Evaluation is applied to the feature set to find the most significant features. Ranker method ranks the attributes by their individual evaluations. Use in conjunction with attribute evaluators like Info Gain. Info Gain evaluates the worth of an attribute by measuring the information gain with respect to the class. In the feature selection phase finally 5 most significant features were selected. Only these 5 optimal features were then used for classification. This phase reduced the curse of dimensionality and when we make the final algorithm only these 5 features are to be evaluated to classify any signal as murmur signal or normal signal.

The selected features are as follows:

Table 4-2 List of selected features

S. No.	**Feature**	**Feature domain**
1.	Total Power 1	Time domain
2.	Q-Factor 1	Frequency domain
3.	t1	Time domain
4.	t12	Time domain
5.	Mean12	Statistical

The selected features have the potential to discriminate between the two classes of signals namely- normal and murmur signals. The feature, Total power 1 is the power of the signal from systole region. As expected, because of the presence of high amplitude signals in case of murmurs, this feature showed higher values for murmurs than the normal signals. This feature was ranked high enough to be used for the classification purpose.

Q-Factor 1 is also calculated in systole region. Due to the presence of murmurs the oscillations die out more slowly. Hence as expected, murmurs showed higher values of this feature and considerable differences in the values of normal and murmur signals were seen.

t1 is the time duration of S1. Due to the presence of murmur just after S1, makes S1 appear longer. So, t1 has a higher value for murmurs as compared to the normal signals.

t12 is the time duration between the end if S1 to the start of same cycle's S2. As murmurs appear just after S1, hence this duration decreases. As expected we observed larger value of this feature for normal signals as compared to murmur signals. This was also termed as an important feature but the ranker method.

Mean 12 is the maximum value amongst the two values in the regions between S1 and S2, S2 and next cycle's S1. This value is observed to be larger for murmurs than for the normal signals.

All the above selected features are considered to be very significant for classification as they are ranked higher than other features. Hence they are used for the classification purpose. By evaluating only these feature values of the signal, we will be able to achieve high classification accuracy.

CLASSIFICATION

In order to diagnose the patient, we need to classify the heart sounds of signals as normal or abnormal heart sounds. Hence, the selected features are used in classification phase. In this phase a pre-defined class is to be assigned to every signal. The pre-defined classes for classification are- normal signal and murmur signal. The classifiers according to some set of rules, classify the signals into pre-defined classes. We used the Bayes Net, Naïve Bayes and some other classifiers in MATLAB for classification. We designed the algorithm for classification. We required maximum classification accuracy. It simply means the ability of the classifier to term normal as normal signal and murmur as murmur signal. In order to automate the process of auscultation, higher classification accuracy is required so that right diagnosis can be done. For our study, the Naïve Bayes classifier was most the appropriate one, as compared to other counterparts.

4.2.3 ALGORITHM FOR MURMUR DETECTION

Step 1: Signals are taken from the database available on the internet.

Step 2: A set of 23 potential features was evaluated for these signals.

Step 3: Ranker Method and Info Gain Method were used to find out the most significant features. Then, total 5 features were selected.

Step 4: For classification we used MATLAB. We divided our data set into two parts. One was named as training set and another was named as test set. Training set was kept larger than test set.

Step 5: Training Set was loaded from excel file into a variable named training. The rows of Training corresponds to the signals, columns corresponds to features.

Step 6: Class variable stored the class of training data. Each element in Class defines which class the corresponding row of Training belongs to. It is also loaded as an excel file. The class has two values- 1 for normal and 2 for murmur signal. Training and class must have the same number of rows.

Step 7: Then the test data is loaded from excel file in a variable named test. Test and Training must have the same number of columns.

Step 8: Naïve Bayes classifier is fitted on Training and Class using the signal processing toolbox.

Step 9: Then according to the fitting of a classifier on training data, the class of Test set is predicted.

Step 10: Training and Test set were changed and step 5 to 9 are repeated. Various combinations of Training and Test sets were used and step 5 to 9 were repeated.

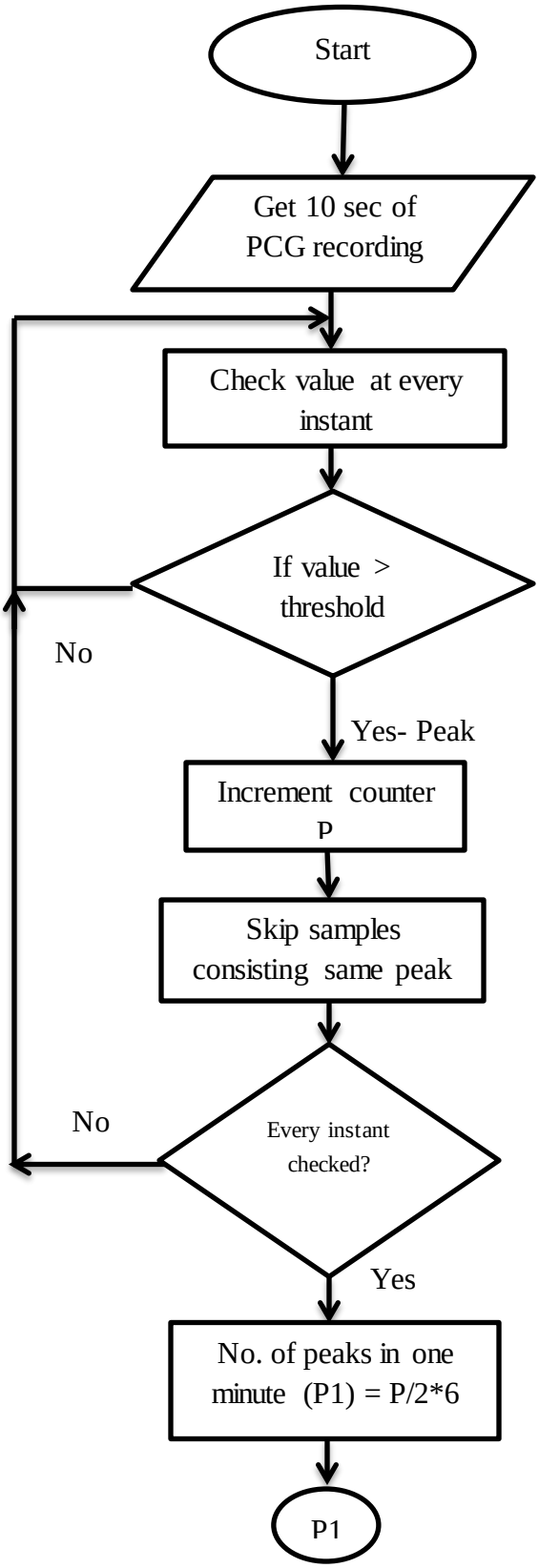
Step 11: Accuracy is calculated. It tells the ability of the methodology to term normal signal as a normal and murmur signal as a murmur.

All the steps shown above were implemented using a MATLAB code. The methodology of classification using feature extraction was implemented in order to improve the accuracy of classification.

4.3 ARRHYTHMIA DETECTION

This study also contributes in the area of arrhythmia detection using heart sounds. Arrhythmia arises due to improper functioning of the electrical system of the heart. Hence, much emphasis is laid on the detection of arrhythmia using ECG which analysis the electrical activity of heart. Echocardiography has also been used in arrhythmia detection.

In this study a method is developed to detect arrhythmia using heart sounds. We use PCG recordings to detect tachycardia and bradycardia. The correlation exists in nature and hence an abnormality due to electrical system will also reflect on the mechanical system of heart and hence in heart sounds. So, this was the driving force for an algorithm for arrhythmia detection.



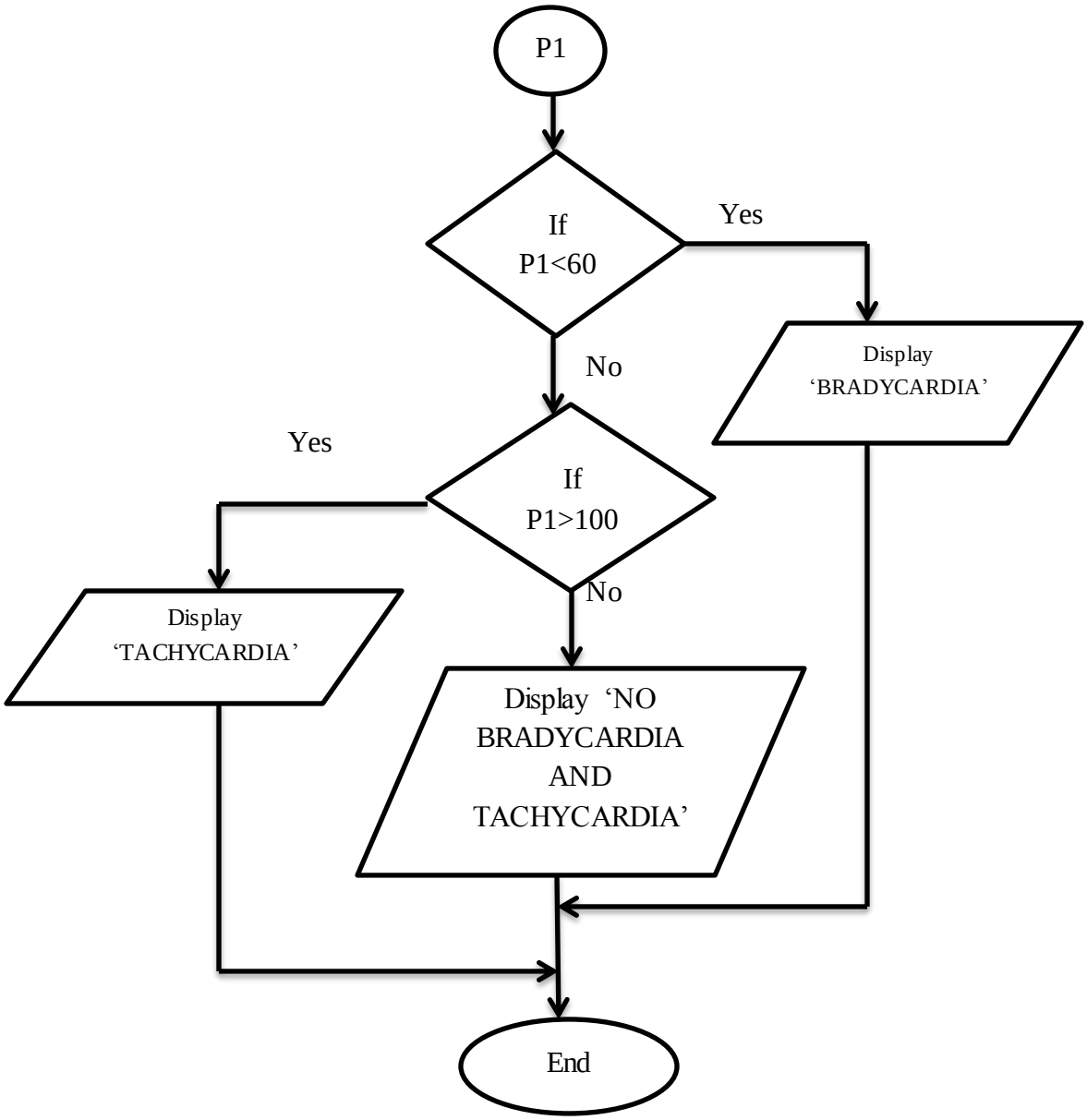


Figure 4-5 Flowchart of algorithm for arrhythmia detection

In this algorithm made for arrhythmia (Bradycardia and Tachycardia) detection, every peak is greater than the threshold. This means we are counting S1 and S2 separately. So, after peak detection process we divide the result by 2, as we count S1, S2 as one heartbeat. In the end we multiply it by 6 because we have taken 10 seconds recording but we have to convert into beats per minute and compare it with the range in order to decide bradycardia or tachycardia condition.

This algorithm is also a step in the direction of automatic auscultation. Mainly emphasis has been laid on arrhythmia detection using ECG because of its connectivity with the electrical system disorder but this method uses PCG for arrhythmia detection. The detection of heart disease using software and a laptop is the future of medicine since it rules out the probability of wrong diagnosis.

This is because there is no possibility of human error and is not bounded by human limitations. It does not require experience but the physician does. We can think of telemedicine and implement it because of these automated electronic techniques as data to larger distances could be sent only by digitization and not by conventional manual methods. The efficiency and advancements in machine learning have made these researches possible and significant.

We are heading towards the direction in which only the PCG recording would be sufficient to test the patient for various diseases. In this study we have made an algorithm which will use PCG signal to check the patient for the presence of murmurs which mainly signifies valvular defects. Another algorithm uses the same record to check arrhythmia. This seems to play a significant role in the cost-effective early detection of various diseases using the same test. This is a very promising method of diagnosis because the instruments that are required are available even at basic rural health care centres i.e. a stethoscope and a PC/laptop.

CHAPTER

5 RESULTS

The PCG signals being very potential in the heart condition detection, forms the basis of this study. We aim at differentiating between normal and abnormal signal. This study is a step forward in the direction of automatic auscultation. The automatic auscultation is the diagnosis of heart disease with the help of software. It doesn't require the experience of the physician.

5.1 MURMUR DETECTION BASED ON VISUAL DIFFERENCE

As expected, this amplitude as a parameter proved to be efficient for classification of heart sound signals as normal and murmur signals.

Based on the visual difference algorithm, the normal signal is classified as follows:

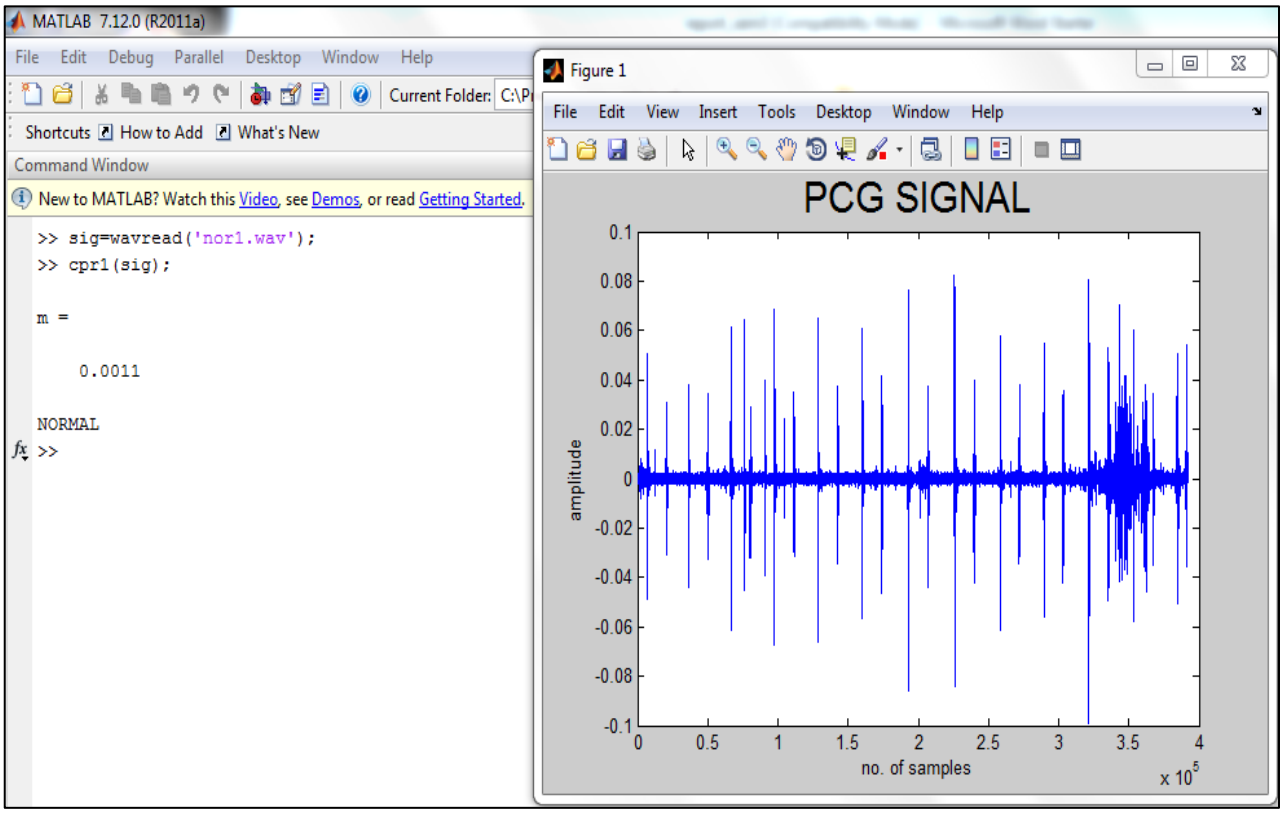


Figure 5-1 Normal Signal Classified

The murmur signal is classified as shown:

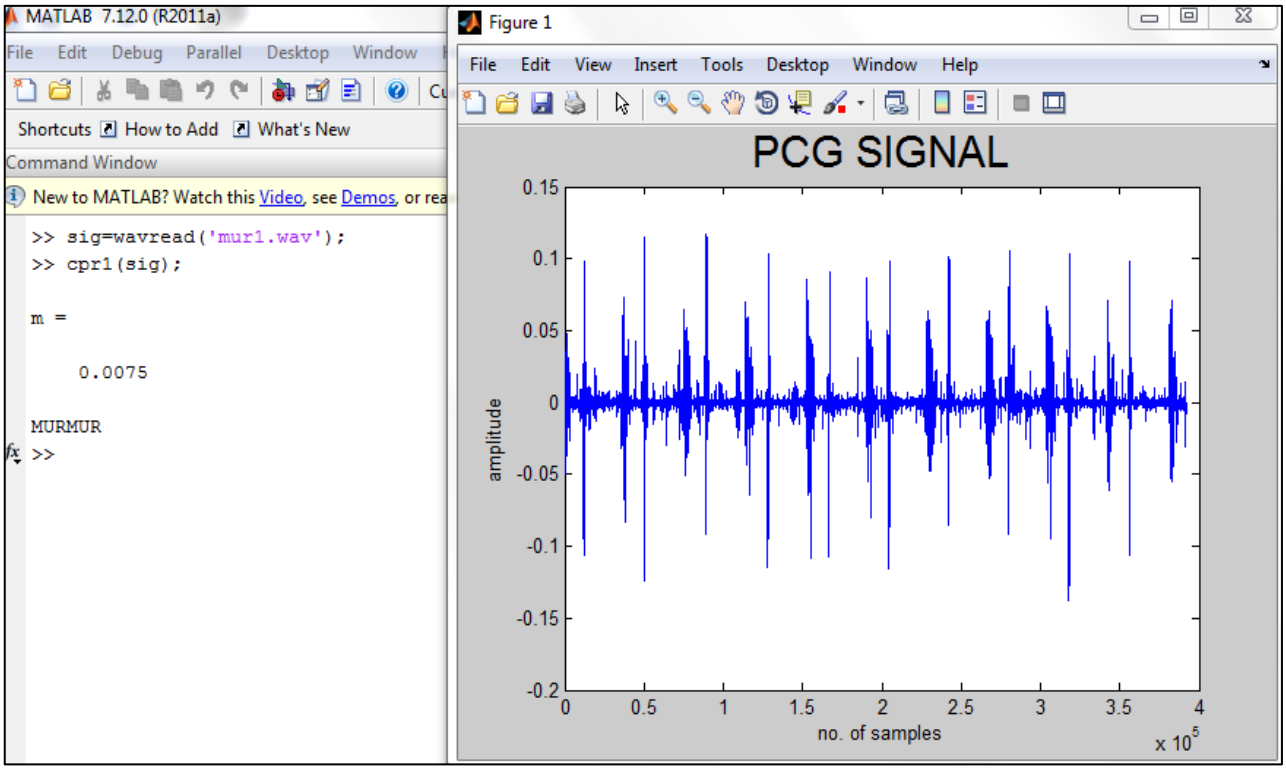


Figure 5-2 Murmur Signal Classified

CONFUSION MATRIX

A confusion matrix is a specific table layout that allows visualization of the performance of an algorithm. It reports the number of false positives, false negatives, true positives, and true negatives. This allows more detailed analysis than mere proportion of correct guesses (accuracy). Hence, by using confusion matrix we can calculate accuracy, sensitivity and specificity in order to properly analyse the efficiency of the algorithm.

Table 5-1 Formulation of Confusion Matrix

Actual Class	**Predicted Class**	
	Normal	**Murmur**
Normal	True Positive	False Negative
Murmur	False Positive	True Negative

Table 5-2 Accuracy, Senstivity and Specificity Parametrs from Confusion Matrix

Parameter	**Formula**
Accuracy	TP+TN/TP+TN+FP+FN
Sensitivity	TP/TP+FP
Specificity	TN/TN+FN

Table 5-3 Confusion Matrix for Visual Difference Methodology

Class	**Normal**	**Murmur**
Normal	18	3
Murmur	5	18

Table 5-4 Efficiency Parameters for Visual Difference Methodology

Parameter	**Value**
Accuracy	18+18/18+3+5+18*100= 81.81%
Sensitivity	18/18+5*100 = 78.2%
Specificity	18/18+3*100 = 85.7%

LIMITATIONS AND SCOPE OF IMPROVEMENT

Accuracy of this algorithm needs to be improved. This work was based on only one parameter for classification. We aim at finding out more parameters in order to properly classify the signals. We aim at incorporating such a mechanism which takes care of both the cases i.e. Systolic murmur and Diastolic murmur.

5.2 MURMUR DETECTION USING FEATURE EXTRACTION

The extracted features form the key to find the missing link in order to find the difference between the normal and murmur classes. The values of these 23 features evaluated for 30 normal signals are as shown:

Table 5-5 Extraction of Frequency domain and Cepsrum features from overall Normal Signals

S.No	Peak Freq	Peak amp	Total power	THD	BW	Q-Factor	Cep Peak Amp
1	50.293	-32.54	-19.79	37.63	1	50.294	-67.36
2	48.34	-31.83	-18.33	52.33	1.6	30.213	-66.92
3	53.711	-25.07	-14.2	25.711	1	53.712	-66.63
4	49.805	-27.3	-12.93	82.366	0.8	62.257	-68.2
5	61.035	-30.54	-15.9	41.142	0.9	67.818	-67.47
6	84.473	-29.49	-15.69	25.629	1	84.474	-67.79
7	72.266	-34.33	-19.74	28.894	1.2	60.222	-65.97
8	27.344	-39.81	-24	132.469	1.2	22.786	-67.6
9	68.848	-28.36	-13.24	51.541	0.9	76.499	-67.57
10	49.805	-29.79	-14.5	61.083	0.7	71.151	-67.71
11	69.336	-31.82	-17.35	23.791	1.1	63.034	-67.63
12	73.73	-35.54	-19.27	92.518	1	73.732	-66.2
13	83.496	-40.8	-24.76	60.281	0.8	104.372	-68
14	83.008	-31.52	-15.48	21.508	0.7	118.584	-67.62
15	29.785	-32.31	-20.29	83.079	1.6	18.616	-67.41
16	50.293	-35.35	-20.05	81.21	0.8	62.867	-67.02
17	40.527	-33.16	-20.24	-80.777	0.9	45.031	-68.05
18	53.711	-28.39	-15.91	34.321	1	53.712	-67.95
19	53.711	-30.35	-15.79	56.268	1	53.712	-66.88
20	34.18	-26.8	-13.79	59.257	1.1	31.073	-65.81
21	73.73	-27.98	-13.95	27.533	0.9	81.924	-67.77
22	44.434	-38.03	-25.69	52.059	0.9	49.371	-65.64
23	57.129	-37.11	-22.96	28.053	1.5	38.087	-65.85
24	66.895	-33.31	-18.25	39.728	1.6	41.81	-67.37
25	82.031	-33.5	-18.01	25.81	0.9	91.147	-68.47
26	42.48	-25.84	-13.57	-50.948	0.9	47.201	-66.62
27	40.527	-37.42	-24.78	53.1	1.9	21.331	-66.87
28	85.938	-28.34	-14.3	18.287	0.9	95.488	-66.55
29	24.414	-36.51	-24.09	83	1	24.414	-66.32
30	88.867	-32.86	-15.99	58.049	1.2	74.057	-67.26

Table 5-6 Extraction of features in systole and diastole regions in Normal Signals

S. No	Peak Freq1	Peak amp1	Total power1	BW1	Q-Factor1	Peak freq2	Peak amp2	Total Power2	BW2	Q-Factor2
1	18.066	-48.56	-35.6	4.8	6.5	22.949	-46.05	-33.68	4.5	5.1
2	20.508	-51.83	-37.63	22.608	0.907	28.32	-51.19	-35.87	4.3	6.586
3	41.504	-56.15	-39.92	4.7	8.831	38.086	-49.27	-34.17	5.2	7.324
4	29.297	-50.36	-33.73	3.5	8.371	17.09	-47.08	-30.33	2.6	14.649
5	29.297	-52.62	-36.26	5.5	5.327	58.105	-48.88	-32.22	3.6	16.141
6	17.578	-46.92	-31.84	4.1	7.622	30.273	-48.05	-32.54	2.3	13.162
7	25.391	-49.78	-34.66	3.3	7.694	31.25	-48.28	-36.77	3.5	8.929
8	18.066	-55.07	-38.54	24.273	0.905	33.691	-50.57	-35.65	2.9	11.618
9	56.641	-48.42	-32.65	5.4	10.489	70.313	-47.3	-31.1	3.4	20.68
10	30.273	-49.78	-33.1	6.5	4.657	16.113	-45.63	-33.22	2.8	16.741
11	67.871	-52.43	-37.5	5.9	11.504	24.414	-46.23	-32.58	2.6	9.39
12	24.414	-57.35	-40.09	28.314	0.862	108.398	-60.76	-43.1	5.1	21.255
13	6.836	-57.56	-41.98	23.496	0.894	35.645	-52.96	-40.24	5.1	6.989
14	37.598	-51.76	-35.57	8.8	4.273	19.043	-47.26	-31.72	5.1	9.287
15	27.832	-51.53	-36.17	4.3	6.7	27.344	-49.26	-36.62	2.7	10.127
16	32.227	-56.5	-39.43	5.3	6.081	48.828	-50.62	-35.62	2.6	18.78
17	18.555	-53.85	-39.09	6.5	4.282	12.695	-55.24	-41.52	8.8	3.718
18	81.543	-59.78	-41.56	6.9	11.818	19.043	-53.27	-36.92	6.8	9.766
19	23.438	-46.16	-32.3	6.3	3.7	31.738	-48.66	-34.05	4.2	7.557
20	20.02	-48.22	-33.09	23.82	0.84	22.461	-47.44	-31.86	3.1	7.245
21	16.602	-48.34	-35.16	3.6	7.731	19.531	-48.62	-34.85	5.1	7.372
22	9.766	-55.32	-39.61	13.6	2.37	14.16	-54.96	-38.93	9.3	3.883
23	20.996	-54.58	-39.83	23.596	0.89	27.344	-51.27	-37.88	4.4	6.214
24	16.602	-50.84	-36.32	4.6	6.263	10.742	-53.23	-37.45	6.3	5.348
25	62.5	-56.43	-37.88	6	10.417	34.668	-51.24	-36.11	3.8	9.123
26	42.48	-49.15	-34.42	4.1	10.361	28.32	-44.62	-31.82	1.9	14.905
27	59.082	-52.79	-35.99	5.2	11.362	9.766	-55.08	-40.48	3.1	7.56
28	27.344	-46.01	-32.51	6.1	4.483	25.291	-45.46	-32.51	3.5	7.254
29	27.344	-51.86	-40.24	4.1	6.669	25.391	-52.69	-40.02	5.2	4.883
30	6.836	-55.62	-38.59	26.302	5.718	26.367	-52.54	-37.57	5.4	4.883

Table 5-7 Extraction of Statistical Time-domain features of Normal Signals

S. No.	mean12	ZCR	t1	t2	t12	t21
1	0.0105	256.381	0.0938	0.0814	0.1565	0.7835
2	0.0112	254.892	0.1098	0.0898	0.1447	0.1497
3	0.0059	230.859	0.0955	0.0442	0.1747	0.1724
4	0.0072	236.739	0.0898	0.0586	0.2423	0.633
5	0.0133	298.207	0.07126	0.0575	0.2445	0.5416
6	0.0108	344.774	0.0753	0.0502	0.27	0.3579
7	0.0101	404.669	0.0847	0.0538	0.2308	0.3155
8	0.0091	373.709	0.0761	0.0468	0.2283	0.3219
9	0.0163	320.61	0.066	0.0812	0.2665	0.302
10	0.0052	298.51	0.0772	0.0643	0.2605	0.4149
11	0.016	210.94	0.1091	0.0935	0.1869	0.2882
12	0.006	387.637	0.0333	0.052	0.1249	0.1748
13	0.0106	269.249	0.1137	0.0455	0.182	0.2123
14	0.0045	393.588	0.116	0.0943	0.7375	0.3479
15	0.0366	309.406	0.1764	0.1764	0.1676	0.3617
16	0.0089	350.463	0.134	0.0963	0.1969	0.3602
17	0.0045	393.588	0.1139	0.0737	0.154	0.2612
18	0.0059	230.853	0.1291	0.0774	0.1239	0.1548
19	0.0233	193.866	0.1092	0.0924	0.1763	0.2351
20	0.012	157.741	0.0875	0.0876	0.1831	0.2787
21	0.009	190.255	0.0624	0.078	0.2079	0.2754
22	0.0051	318.082	0.0874	0.0874	0.1748	0.2719
23	0.0131	219.618	0.0974	0.0909	0.1884	0.2468
24	0.01	203.695	0.1082	0.0794	0.1732	0.2453
25	0.008	180.175	0.1062	0.0493	0.1745	0.2655
26	0.0113	166.172	0.0951	0.0624	0.1992	0.434
27	0.0054	223.681	0.1336	0.0927	0.1772	0.2617
28	0.0107	174.39	0.0911	0.0855	0.2138	0.2394
29	0.0048	254.803	0.1138	0.0828	0.2105	0.1966
30	0.0321	505.655	0.0922	0.0676	0.1107	0.1476

Table 5-8 Extraction of Frequency domain and Cepsrum features from overall Murmur Signals

S. No.	Peak Freq	Peak amp	Total power	THD	BW	Q-Factor	Cep Peak Amp
1	37.109	-37.29	-22.98	83.294	0.7	53.014	-66.37
2	61.035	-26.33	-12.13	21.857	1.1	55.487	-67.19
3	87.891	-30.82	-16.83	10.793	1.1	79.902	-66.89
4	25.879	-35.59	-20.8	114.241	0.9	28.754	-67.4
5	68.848	-46.96	-30.47	78.126	0.8	86.061	-66.39
6	65.43	-44.2	-27.31	70.567	1	65.431	-65.97
7	21.973	-40.43	-26.17	130.483	1.2	18.31	-66.51
8	49.316	-34.06	-20.92	25.079	1.1	44.834	-66.39
9	94.238	-27.01	-11.61	44.392	1.1	85.672	-67.03
10	76.172	-28.31	-12.04	71.25	1.1	69.248	-66.67
11	50.781	-24.17	-6.1	-134.323	0.9	56.424	-67.45
12	90.82	-34.29	-16.98	54.181	3.3	27.522	-67.49
13	58.105	-31.9	-15.74	89.804	0.9	64.563	-67.03
14	30.273	-31.81	-15.14	156.758	1.1	27.521	-67.38
15	63.477	-23.47	-10.6	51.511	1	63.478	-66.99
16	67.871	-29.17	-15.49	35.927	1.1	61.702	-66.36
17	31.738	-41.77	-23.66	201.07	0.8	39.673	-65.86
18	43.457	-37.93	-21.52	115.403	1	43.458	-66.12
19	51.27	-22.61	-6.81	113.323	0.9	56.967	-68.12
20	42.969	-24.49	-7.43	157.26	1	42.969	-68.11
21	26.855	-33.48	-19.09	119.72	1.1	24.414	-67.55
22	42.48	-34.17	-20.62	64.695	1	42.481	-67
23	70.313	-24.76	-7.32	120.34	1	70.314	-67.49
24	38.086	-26.31	-7.71	209.538	1	38.087	-67.84
25	29.297	-24.79	-7.02	205.678	1.4	20.926	-67.96
26	61.035	-38.25	-25.98	29.534	0.8	76.295	-66.5
27	30.762	-30.28	-17.55	99.348	1	30.762	-66.76
28	34.668	-26.13	-11.72	91.987	1.8	19.26	-67.65
29	53.223	-31.63	-15.41	95.924	1	53.223	-67.43
30	127.441	-24.01	-5.73	77.654	0.9	141.604	-68.34

Table 5-9 Extraction of features in systole and diastole regions in Murmur Signals

S. No.	Peak Freq1	Peak amp1	Total power1	BW1	Q-Factor1	Peak freq2	Peak amp2	Total Power2	BW2	Q-Factor2
1	137.695	-56.27	-38.14	8.701	15.826	22.949	-47.54	-37.09	1.9	12.078
2	27.344	-39.08	-26.22	3	9.115	27.344	-38.99	-25.88	3.8	7.196
3	293.945	-48.96	-29.91	5.4	54.431	51.27	-50.55	-36.66	2.8	18.311
4	39.063	-44.68	-29.26	5.2	7.512	25.879	-43.32	-30.69	2.7	9.585
5	19.531	-51.56	-34.83	3.4	21.255	16.602	-49.25	-35.85	2.4	31.128
6	58.594	-54.2	-35.45	8.1	7.234	23.926	-54.6	-40.17	4	5.981
7	16.602	-49.14	-35.21	25.049	0.916	20.02	-48.32	-38.13	21.42	0.935
8	17.09	-52.6	-39.28	4.9	6.378	21.973	-49.4	-37.02	3.2	6.866
9	63.965	-39.61	-21.69	5.9	10.842	63.965	-46.08	-28.72	5.9	10.842
10	99.121	-45.01	-24.73	9.8	10.115	45.41	-53.4	-34.51	8.5	5.342
11	73.73	-33.07	-12.01	6.6	11.171	23.438	-35.81	-22.57	5.5	4.261
12	124.023	-43.52	-26.59	5.7	21.759	142.09	-43.67	-25.81	1.6	88.801
13	68.848	-39.85	-21.11	9.1	7.566	41.016	-52.42	-35.48	11.2	3.662
14	48.34	-38.7	-21.46	5.9	8.193	14.648	-42.92	-30.19	4.5	6.944
15	207.52	-39.49	-18.34	24.101	8.61	21.973	-50.34	-36.97	24.373	0.902
16	199.219	-49.35	-29.89	8.1	24.593	21.484	-49.56	-32.53	23.884	0.9
17	153.32	-49.28	-28.57	15.101	10.153	15.137	-54.56	-39.55	2.1	10.696
18	30.273	-50.78	-30.02	3.4	8.904	29.297	-46.98	-35.8	3	9.766
19	432.617	-35.04	-14.03	9.101	47.537	66.406	-44.99	-28.51	8.6	7.722
20	105.957	-33.57	-12.26	5.5	19.265	39.551	-42.79	-25.82	7.2	5.493
21	120.117	-43.86	-24.42	6.1	13.048	23.926	-55.88	-39.63	15.701	8.614
22	129.883	-51.43	-31.74	20.401	6.367	15.137	-54.48	-40.75	6.4	4.272
23	144.043	-28.5	-9.31	4.4	32.735	45.41	-57.95	-37.68	53.91	0.842
24	183.105	-32.78	-11.63	7.5	24.413	55.664	-42.27	-25.58	4.9	11.36
25	114.258	-32.04	-11.57	7.1	16.093	40.527	-39.51	-39	7.2	5.629
26	22.949	-58.79	-41.45	5.8	3.957	48.828	-57.54	-39.9	4.9	9.965
27	87.402	-50.43	-33.67	5.6	15.608	25.391	-52.64	-38.12	4.1	6.193
28	113.77	-44.54	-27.83	5.3	21.466	16.113	-52.72	-36.48	2.3	14.012
29	123.047	-39.06	-23.88	10	12.305	1.465	-50.97	-35.76	1.6	19.531
30	292.48	-25.96	-8.33	6.7	43.651	33.691	-43.86	-29.64	6.9	4.883

Table 5-10 Extraction of Statistical Time-domain features of Murmur Signals

S. No.	mean12	ZCR	t1	t2	t12	t21
1	0.0077	268.967	0.1706	0.1129	0.1458	0.6021
2	0.0129	275.779	0.5374	0.3037	0	0.3272
3	0.0088	324.117	0.2157	0.1233	0.154	0.6315
4	0.0248	287.654	0.2606	0.0782	0.0695	0.2259
5	0.0248	427.819	0.4141	0.1705	0	0.2436
6	0.0248	405.428	0.3122	0.0765	0	0.344
7	0.0058	266.822	0.0842	0.0618	0.0449	0.2079
8	0.0171	185.393	0.1452	0.1307	0.1743	0.1888
9	0.0207	243.109	0.3071	0.3071	0	0.2532
10	0.0535	366.155	0.345	0.345	0	0.1785
11	0.1833	316.274	0.3932	0.3932	0	0.1368
12	0.019	326.511	0.2784	0.1218	0.1044	0.6089
13	0.0249	292.761	0.3201	0.3201	0	0.2088
14	0.0114	499.965	0.2954	0.2954	0	0.1933
15	0.0552	398.792	0.362	0.362	0	0.1164
16	0.0327	406.802	0.0767	0.0644	0.1657	0.2209
17	0.0025	453.663	0.3164	0.0716	0	0.2825
18	0.0033	448.126	0.3493	0.0649	0	0.3043
19	0.0437	550.011	0.2584	0.2584	0	0.1433
20	0.1011	515.552	0.2707	0.2707	0	0.1297
21	0.0146	331.248	0.4032	0.4032	0	0.0336
22	0.0105	349.46	0.4755	0.4755	0	0.111
23	0.0165	354.049	0.3078	0.3078	0	0.0468
24	0.1654	477.197	0.2908	0.2908	0	0.243
25	0.1686	487.583	0.2964	0.2964	0	0.1468
26	0.1686	377.907	0.0992	0.0947	0.1736	0.2074
27	0.0095	199.045	0.2032	0.1839	0.0906	0.1455
28	0.0076	338.052	0.1705	0.3922	0	0.3411
29	0.0062	520.503	0.3126	0.1316	0.0494	0.4278
30	0.021	412.515	0.3693	0.3693	0	0.1266

Out of these 23 features, the most important and significant 5 features selected during feature reduction phase enlisted below:

- Total Power 1
- Q-Factor 1
- t1
- t12
- Mean12

Selection of only a few significant features reduces the curse of dimensionality and the computation time. This means that by simply evaluating the value of signal for the above features, classification of normal and murmur signals can be done.

A few new features like Q-factor 1 and Mean 12 are introduced in this study. By using these features, the ability to differentiate between normal and murmur showed tremendous improvement as compared to the classifications without these features. Hence these features proved to be the missing link in differentiating the two types of signals.

Various classifiers were used in this study to find out the best classifier that suits our problem. The table below shows the improvement in classification accuracy of various classifiers by using these features:

Table 5-11 Accuracy achieved using different classifiers with and without using new features

Classifier	**TP Rate (N)**	**TP Rate (M)**	**Accuracy (%) *without* these new features**	**Accuracy (%) *With* these new features**
Bayes Net	0.933	0.9	86.6667	91.6667
Naïve Bayes	0.933	0.933	88.3333	93.3333
SGD	0.967	0.867	90	91.6667
Logit Boost	0.933	0.833	83.3333	88.3333

The use of these new features leads to the improvement in accuracy in every classifier listed above. The 23 features initially evaluated had the potential for classification but these selected 5 features were significant and enough for classification. As seen above these features proved to be extremely important in differentiating normal signals from the murmur signals.

Naïve Bayes proved to be the best suited classifier for this application. It gave the highest accuracy amongst all other classifiers used. Hence we used Naïve Bayes classifier because we aim to minimize the probability of murmur being classified as normal signal.

The confusion matrix using Naïve Bayes classifier is as shown:

Table 5-12 Confusion matrix using Naive Bayes classifier

Classifier	**Normal**	**Murmur**	**Class**
Naïve Bayes	28	2	**Normal**
	2	28	**Murmur**

The efficiency parameters for feature extraction methodology, evaluated from the confusion matix using Naïve Bayes classifier are as shown:

Table 5-13 Efficiency Parameters for Feature Extraction Methodology

Parameter	**Value**
Accuracy	28+28/28+2+28+2*100 = 93.33%
Sensitivity	28/28+2*100 = 93.33%
Specificity	28/28+2*100 = 93.33%

Table 5-14 Comparison of Visual Difference and Feature Extrction Methodologies

Parameter	**Visual difference Methodology**	**Feature Extraction Methodology**
Accuracy	81.81%	93.33%
Senstivity	78.2%	93.33%
Specificity	85.7%	93.33%

By using this methodology we achieved a higher accuracy. The classification accuracy showed tremendous improvement as compared to that obtained using visual differences alone. We had only one parameter in the previous methodology but here we had five parameters as deciding factors in the present feature extraction methodology.

5.3 ARRHYTHMIA DETECTION

Much emphasis has been laid on detection of arrhythmia using ECG. This is due to its connection with the electrical disorders of the heart and the ECG shows the electrical activity of heart. We aim at using PCG for heart disease diagnosis. As arrhythmias are related with heart rate or number of beats per minute, so counting the number of heart beats would be sufficient for diagnosis. We have used the same data set that we used for murmur detection, to test the algorithm made for arrhythmia detection. We took a paper [41] presented by the provider of dataset as reference. In this paper he presented the details of the dataset provided by him. In this the data set provider mentions the number of heart beats and other details in various signals of his dataset. The details of each signal including number of heart beats would be used for testing the efficiency of the algorithm. The results of testing the algorithm are shown in the table below:

Table 5-15 Results of Arrhythmia detection algorithm

S.No	Name Of Signal as in [41]	Actual number of Heart beats As in [41]	Number of Heart Beats detected (P1)	Error (In number of Heart Beats)
1	154_1306935608852_B1	4.5	4.5	0
2	159_1307018640315_B1	7	7	0
3	159_1307018640315_B2	3	3	0
4	179_1307990076841_B	16.5	16.5	0
5	184_1308073010307_D	26.5	26.5	0
6	103_1305031931979_B	12.5	12.5	0
7	106_1306776721273_C2	3	3	0
8	107_1305654946865_C1	8	7.5	-0.5
9	126_1306777102824_B	6.5	7	0.5
10	126_1306777102824_C	3.5	3.5	0
11	133_1306759619127_A	4.5	4	-0.5
12	140_1306519735121_B	13	13	0
13	146_1306778707532_B	19	19.5	0.5
14	146_1306778707532_D3	3	3	0
15	151_1306779785624_D	4.5	4.5	0

The above results show that the algorithm made for arrhythmia detection yields a very high accuracy with no or negligible error in most of the cases. It showed an error of half a heartbeat in some signals which will not have a bad impact on decision making process i.e. while deciding between the classes 'Tachycardia', 'Bradycardia' and 'No Tachycardia and Bradycardia'. This is because of the large ranges of classes which are as shown:

No Tachycardia and Bradycardia = Normal range = 60 to 100 bpm.

Tachycardia = Fast heart rate > 100 bpm.

Bradycardia = slow heart rate < 60 bpm.

So, the probability that a class will change due a small error in number of heart beats is minimal and hence the proper diagnosis of arrhythmia can be done using PCG signals.

CHAPTER

6 CONCLUSION AND FUTURE SCOPE

6.1 CONCLUSION

In this thesis an attempt is made to study and analyse the characteristic features of PCG for detection of various heart diseases. The algorithms proposed in this study are time efficient, simple, and require only PCG as input signal unlike other methods which require ECG gating. The following points are concluded during this study:

For murmur detection two new features like Q-factor 1 and Mean 12 are proposed.

Using 5 optimal features and Naïve Bayes classifier the accuracy increased from 88.33% to 93.33% and thus can lead to more reliable diagnosis.

The proposed algorithm for murmur detection is useful to detect mainly the valve-related diseases and other congenital abnormalities.

The proposed algorithm can also detect arrhythmia and can further classify it into Tachycardia or Bradycardia.

6.2 FUTURE SCOPE

The proposed method can also be implemented using the latest mobile phones with the applications which can work as electronic stethoscope or Phonocardiogram.

The method presented in this study can easily be implemented in the existing electronic stethoscope by interfacing it with the present embedded technology.

The accuracy of the presented algorithms can be further increased by incorporating Artificial Intelligence techniques or other hybrid classifiers on a larger dataset.

REFERENCES

[1] Roy D, Sargeant J, Gray J, Hoyt B, Allen M, Fleming M, "*Helping family physicians improve their cardiac auscultation skills with an interactive CD-ROM"*, Journal of Continuing Education in Health Professions 22, 2002, pp. 152-159.

[2] Jiang Z, Choi S, "*A cardiac sound characteristic waveform method for in-home heart disorder monitoring with electric stethoscope",* Expert Systems with Applications 31, 2006, pp. 286-298.

[3] Babaei S, Geranmayeh A, *"Heart Sound reproduction based on neural network classification of cardiac valve disorders using wavelet transform of PCG signals"*, Elsevier, Computers In Biology and Medicine 39, 2009, pp. 8-15.

[4] Leatham, Vogelpoel, *"Splitting of first and second Heart Sound"*, The Lancet, Volume 264, Issue 6839, 1954, pp. 607-614.

[5] Kagawa Y, Sato N, Nitta S, Saji K, Tanaka M, Shibota Y and Horiuchi T, *"Sound Spectroanalytic Diagnosis of Malfunctioning Prosthetic Valves"*, Tokuhu J. Exp. Med., 123, 1977, pp. 77-89.

[6] Durand L.C, Blanchard M, Sabbah H.N, Hamid M, Kemp S.R, Stein P.D, *"A Bayes model for automatic detection and quantification of bioprosthetic valve degeneration"*, Mathematical and Computer Modelling 11, 1988, pp. 158-163.

[7] Barschdor D, Femmer U, Trowitzsch E, *"Automatic phonocardiogram signal analysis in infants based on wavelet transforms and artificial neural networks"*, Computers in Cardiology, IEEE, Vienna, Austria, 1995, pp. 753-756.

[8] El-Asir B, Khadra L , Al-Abbasi A, Mohammed M, *"Time-frequency analysis of heart sounds"*, Proceedings of IEEE TENCON Conference on Digital Signal Processing applications, Volume 2, Perth, WA, Australia, 1996, pp. 553-558.

[9] El-Asir B, Khadra L, Al-Abbasi A, Mohammed M, *"Multiresolution analysis of Heart Sounds"*, Proceedings of Third IEEE TENCON International Conference Electronics, Circuits, and Systems, Volume 2, Rodos, Greece, 1996, pp. 1084-1087.

[10] Lee J.J, Lee S.M, Kim I.Y, Min H.K, *"Comparison between Short Time Fourier and wavelet transform for feature extraction of heart sound"*, Proceedings of IEEE TENCON 99, Volume 2, Cheju Island, South Korea, 1999, pp. 1547-1550.

[11] Shino H, Yoshida H, Mizuta H, Yana K, *"Phonocardiogram classification using time-frequency representation"*, Proceedings of 19th International Conference on Engineering in Medicine and Biology Society, Volume 4, Chicago, IL, 1997, pp. 1636-1637.

[12] Rangayyan R.M, Lehner R.J, *"Phonocardiogram Signal Analysis: A Review"*, CRC Critical Reviews in Biomedical Engineering, 15(3), 1988, pp. 211-236.

[13] Rabiner L and Shafer R, *"Digital Processing of Speech Signals"*, Prentice Hall: NJ.

[14] Vishwanath M, Shervegar, Ganesh V. Bhat, Raghavendra M, Shetty K, *"Phonocardiography- The future of cardiac auscultation"*, International Journal of Scientific & Engineering Research, Volume 2, Issue 10, 2011.

[15] Muruganantham, *"Methods for Classification of Phonocardiogram"*, IEEE TENCON, 2003.

[16] Shui L, *"AM13 Analysis of Heart Sound Thesis"*, National University of Singapore, 2004.

[17] El-Segaier M, Lilja O, Lukkarinen S, S-Ornmo L, Sepponen R, Pesonen E, *"Computer based detection and analysis of Heart Sound and Murmur"*, Annals of Biomedical Engineering, Volume 33, No. 7, 2005, pp. 937-942.

[18] Ahlstrom C, Liljefelt O, Hult P, Ask P, *"Processing of Phonocardiographic Signal- Methods for Intelligent Stethoscope"*, Studies in Science and Technology, Thesis No. 1253, 2005.

[19] Ahlstrom C, Hult P, Ask P, *"Thresholding distance plots using true recurrence points"*, Proceedings of IEEE Conference on Acoustics, Speech and Signal Processing, 2006.

[20] Noponen A.L, Lukkarinen S, Angerla A, Sepponen R, *"Phono-spectrographic analysis of heart murmur in children"*, BMC Pediatrics, 2007.

[21] Amit G, Gavriely N, Intrator N, *"Cluster analysis and Classification of Heart Sounds"*, Elsevier, Biomedical signal Processing and Control 4, 2009, pp. 26-36.

[22] Maglogiannis I, Loukis E, Zafiropoulos E, Statis A, *"Support Vector Machine- based identification of heart valve diseases using Heart Sounds"*, Elsevier, Computer methods and programs in Biomedicine 95, 2009, pp. 47-61.

[23] Yuenyong S, Nishihara A, Kongprawechnon W, Tungpimolrut K, *"A framework for automatic heart sound analysis without segmentation"*, Biomedical Engineering Online, 2011.

[24] Akbari M.A, Hassani K, Doyle J.D, Navidbakhsh M, Sangargir M, Bajelani K, Ahmadi Z.S, *"Digital Subtraction Phonocardiography (DSP) applied to detection and characterization of heart murmurs"*, Biomedical Engineering Online, 2011.

[25] Devi A, Misal A, "A survey on classifiers used in Heart Valve Disease Detection", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Volume 2, Issue 1, 2013.

[26] Dasha M, Liu H, *"Consistency-based search in feature selection"*, Elsevier, Artificial Intelligence 151, 2003, pp. 155-176.

[27] Novakovic J, *"Using Information Gain Attribute Evaluation to Classify Sonar Targets"*, 17th Telecommunication forum TELFOR, 2009.

[28] Novakovic J, Strbac P, Bulatovic D, *"Toward Optimal Feature selection using ranking methods and classification algorithms"*, Yugoslav Journal of Operations Research 21, Number 1, 2011, pp. 119-135.

[29] Zhang G.P, *"Neural Networks for classification- A Survey"*, IEEE Transactions on Systems, Man and Cybernetics- Part C: Applications and Reviews, Volume 30, Number 4, 2000.

[30] Cromwell L, Weibell F.J, Pfeiffer E.A, *"Biomedical Instrumentation and Measurements"*, Pearson Education, Second Edition.

[31] www.howstuffworks.com, Referred on May 5, 2013.

[32] Rushmer R.F, *"Cardiovascular Dynamics"*, Philadelphia, W.B. Saunders 4th edition, 1976, pp. 76-104.

[33] http://learnpediatrics.com (University of British Columbia), Referred on May 8, 2013.

[34] www.diglondon.blogspot.com, Referred on May 23, 2013.

[35] http://www.nhlbi.nih.gov (National heart, Blood and Lung Institute), U.S. Department of Health and Human Resources, Referred on May 25, 2013.

[36] www.acousticheart.com, Referred on June 3, 2013.

[37] www.medicalexpo.com, Referred on June 5, 2013.

[38] http://www.heart.org (American Heart Association), Referred on June 9, 2013.

[39] www.heart-consult.com, Referred on June 12, 2013.

[40] www.peterjbentley.com/ heartchallenge/ index.html, Referred on December 12, 2012.

[41] Bentley P.J, Nordehn G, Coimbra M, Mannor S, *"The PASCAL Classifying Heart Sounds Challenge 2011 (CHSC2011) Results"*, http//www.peterjbentley.com/heartchallenge/index.html, referred on June 20, 2013.