NATIONAL INSTITUTE OF TECHNOLOGY, ROURKELA

Biometric Identification and Verification based on time-frequency analysis of Phonocardiogram Signal

by

Arunava Karmakar

A thesis submitted in partial fulfillment for the
degree of Master of Technology

in the
Electronics and Instrumentation Engineering
Department of Electronics and Communication

June 2012

Certificate

This is to certify that the thesis titled **"Biometric Identification and Verification based on time-frequency analysis of Phonocardiogram Signal"** submitted by **Mr. Arunava Karmakar** in partial fulfillment of the requirements for the award of **Master of Technology degree** in **Electronics and Communication Engineering** with specialization in **"Electronics and Instrumentation"** during the session 2010-2012 at **National Institute of Technology, Rourkela** is an authentic work by him under my supervision and guidance.

To the best of my knowledge, the matter embodied in the thesis has not been submitted to any other university/ institute for the award of any Degree or Diploma.

Place: NIT Rourkela
Date: 1 June 2012

Dr. Samit Ari
Asst. Professor, ECE Department
NIT Rourkela, Odisha

NATIONAL INSTITUTE OF TECHNOLOGY, ROURKELA

Abstract

Electronics and Instrumentation Engineering
Department of Electronics and Communication

Master of Technology

by Arunava Karmakar

Heart sound is generally used to determine the human heart condition. Recent reported research proved that cardiac auscultation technique which uses the characteristics of phonocardiogram (PCG) signal, can be used as biometric authentication system. An automatic method for person identification and verification from PCG using wavelet based feature set and Back Propagation Multilayer Perceptron Artificial Neural Network (BP-MLP-ANN) classifier is presented in this paper. The work proposes a time frequency domain novel feature set based on Daubechies wavelet with second level decomposition. Time-frequency domain information is obtained from wavelet transform which in turn is reflected in wavelet based feature set which carries important information for biometric identification. Database is collected from 10 volunteers (between 20-40 age groups) during three months period using a digital stethoscope manufactured by HDfono Doc. The proposed algorithm is tested on 4946 PCG samples of duration 20 seconds and yields 96.178% of identification accuracy and Equal Error Rate (EER) of 17.98%. The preprocessing before feature extraction involves selection of heart cycle, low pass filtering, extraction of heart cycle, aligning and segmentation of S1 and S2. The identification is performed over the score generated output from the ANN. The experimental result shows that the performance of the proposed method is better than the earlier reported technique, which used Linear Band Frequency Cepstral coefficient (LBFCC) feature set. Verification method is implemented based on the Mean square error (MSE) of the cumulative sum of normalized extracted feature set.

Acknowledgements

I would like to express my gratitude to my supervisor Prof. Samit Ari for his guidance, advice and constant support throughout my thesis work. I would like to thank him for being my advisor here at National Institute of Technology, Rourkela. Next, I want to express my gratitude to Prof. S. Meher, Prof. S.K. Patra, Prof. K. K. Mahapatra, Prof. S. K. Behera, Prof. Poonam Singh, Prof. T. K. Dan, Prof. U. C. Pati, Prof. A. K. Sahoo, Prof. D. P. Acharya, and Prof. S. K. Das for their teachings and extended help. I would like to thank all faculty members and staff of the Department of Electronics and Communication Engineering, N.I.T. Rourkela for their generous help which played a vital role in various ways for the successful completion of this thesis. I would also like to mention the names of Manab Das, Dipak Ghosh, Gautam, Anil, Jamal, Shravan, Brajesh, Sankata, Prasant, Santosh and Gopi for their help and support during the heart sound data collection. I would like to thank all my friends and especially my classmates for all the thoughtful and motivating discussions we had, which encouraged me to think beyond the observable. I am especially grateful to my parents for their love and support and would like to thank my parents for raising me in a way to believe that I can achieve anything in life with hard work and dedication.

Contents

Certificate i

Abstract ii

Acknowledgements iii

List of Figures vi

List of Tables vii

1 Introduction 1
- 1.1 Biometric System . . . 1
- 1.2 Positive features of a Biometric system . . . 2
- 1.3 Phonocardiogram as a physiological trait for Biometric Systems . . . 3
- 1.4 Mechanism of Heart Sound Production . . . 3
- 1.5 Auscultation Process of Heart Sound . . . 6
- 1.6 Data collection . . . 7
- 1.7 Previous Works . . . 7
  - 1.7.1 Literature review . . . 7
  - 1.7.2 Implementation of Linear Frequency Band Cepstrum coefficient technique . . . 8
    - 1.7.2.1 Feature Extraction . . . 8
    - 1.7.2.2 Classification schemes . . . 10
    - 1.7.2.3 Results . . . 11
    - 1.7.2.4 Conclusion . . . 11
- 1.8 Theoretical background to wavelet transform . . . 12
- 1.9 Proposal of Wavelet based approach . . . 14
- 1.10 Brief overview of BP-MLP-ANN . . . 14
  - 1.10.1 Structure of a single node . . . 14
  - 1.10.2 Multilayer Network . . . 15
  - 1.10.3 Training . . . 15
  - 1.10.4 Back Propagation Algorithm . . . 16
  - 1.10.5 Perceptron . . . 16

1.11 Proposal of Back Propagation Multi Layer Perceptron Artificial Neural Network as a classifier . . . 17
1.12 Motivation . . . 18
1.13 Thesis outline . . . 18

2 Preprocessing **19**
2.1 Normalization . . . 19
2.2 Low pass filtering . . . 20
2.3 Extraction of Heart cycle . . . 20
2.4 Aligning . . . 20
2.5 Segmentation . . . 22

3 Feature Extraction and Classification **25**
3.1 Feature Extraction . . . 25
3.2 Classification . . . 29
3.3 Feature fitness and Imposter Detection . . . 31
3.4 Identification . . . 32
3.5 Verification . . . 33

4 Conclusion and Future Work **37**
4.1 Conclusion . . . 37
4.2 Future works . . . 37

Bibliography **39**

List of Figures

1.1 Two phases of Biometric authentication . . . 2
1.2 Frequency representation of PCG signals . . . 4
1.3 Cross section of a human heart . . . 5
1.4 Normal phonocardiogram signal . . . 5
1.5 Auscultation sites in the chest region . . . 6
1.6 Data Collection from a volunteer using Digital Stethoscope . . . 6
1.7 Block diagram of the LBFCC feature extraction process . . . 9
1.8 Wavelet . . . 14
1.9 Structure of a single neuron. . . . 15
1.10 Structure of a Multi layer Back Propagation NN. . . . 17

2.1 Preprocessing. . . . 19
2.2 Extraction of Heart cycle. . . . 21
2.3 Heart cycle aligning process. . . . 22
2.4 Heart cycle before and after aligning. . . . 23
2.5 Process of heart cycle segmentation. . . . 24

3.1 Block diagram of feature extraction process. . . . 25
3.2 daubechies wavelet . . . 26
3.3 Second level wavelet decomposition . . . 26
3.4 Plot of the extracted feature from class 1 to 4. . . . 27
3.5 Plot of the extracted feature from class 5 to 8. . . . 28
3.6 Plot of the extracted feature from class 9 and 10. . . . 29
3.7 ANN parameters. . . . 30
3.8 Block diagram of identification process.. . . . 32
3.9 Process of Verification. . . . 34
3.10 ROC1-6. . . . 35
3.11 ROC6-10. . . . 36

List of Tables

1.1 Result of LFBCC feature set 12

3.1 NN Outputs 30
3.2 MSE chart for imposter testing for randomly choosen class 1,3,7 31
3.3 Process of Identification from Neural Network Output 32
3.4 Confusion Matrix 33
3.5 Comparison with LBFCC 33
3.6 Equal Error Rate obtained from the Identity Verification testing 36

Dedicated to
My parents

Chapter 1

Introduction

1.1 Biometric System

Security plays a vital role in present day scenario where identity fraud and terrorism possesses great threat. Recognizing humans using computers not only provide some best security solutions but also help to efficiently deliver human services. These computer based human identification systems are known as biometric systems. Present day we encounter biometric systems in offices, laptops, cars, lockers and many everyday items. Biometric authentication system associate behavioural or physiological attributes to identify or verify a person. Physiological traits are based on bodily features, like fingerprint, iris, facial structure, skin tone etc. Behavioural traits are based upon unique behavioural characteristics of an individual, like voice, signature etc. Heart sound comes under physiological traits because it is a natural sound created by the opening and closure of valves present in the heart.

Biometric authentication processes are composed of two phases. In the first phase the database is created where the feature sets which describes each individual is stored. In the second phase the extracted feature sets are compared with the feature templates stored in the database to find a match. Figure 1.1 shows the block diagram of this system comprised of two phases.

Biometric authentications can be of two types, Identification and Verification[1][2].

Identification :
The process of identification can be understood from how a crime suspect is identified by

a witness. Many persons are brought before the witness as probable suspects. Witness's job is to identify the criminal among the suspects on the basis of the physical attributes that is stored in his memory. In biometric system the classifier is first trained with the features of the different classes. Feature extracted from the quiry sample is matched among the stored classes and finally a decision is made about which class the quiry sample belongs to. How accurately the system is able to find a match is the parameter to judge how good the feature set and the classifier are.

Verification :
The process of verifying is somewhat different. Here the system has to judge between true or false. Here the system analyse the feature and take decision if the feature is good enough to be labelled as the identity it claims to be.

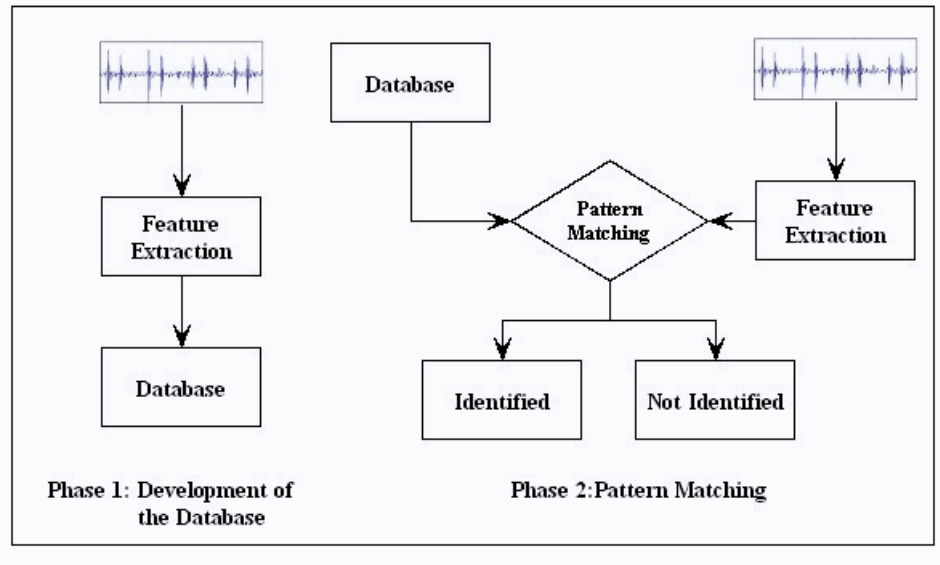

FIGURE 1.1: Two phases of Biometric authentication.

1.2 Positive features of a Biometric system

Attributes which are considered as the most important features of a biometric system are[3]:
Accuracy : Nothing is more important for a biometric system than to be accurate with minimum false accept rate (FAR) and high true accept rate (TAR).
Speed : A quick response is very desirable for a biometric system. This includes quick data acquisition, quick enrollment of inputs and processing.
Reliability : The system should be resistant to forgery and the detection should be trustworthy in a changing environment.

Universality : The biometric feature trait should be present in every living individual.
Easyaccessibility : The physiological trait that the system uses should be easily accessible.
Invariability : The system performance should remain same over a long duration of time.
Apart from these, there are many other selling points for a biometric system, like data storage requirements, cost, sensor quality etc. One of the important matter to consider is whether heart sound or phonocardiogram based biometric system will be a good option or not.

1.3 Phonocardiogram as a physiological trait for Biometric Systems

First and second heart sounds (S1 and S2) are created by sudden closure of atrioventricular and semilunar valves and are completely natural physiological signals. These are impossible to reproduce exactly similar by artificial means. Every living human being has a beating heart and the sound can easily be accessed by holding stethoscope in the auscultation sites in the chest region. By the advent of improved data acquisition systems these sounds can further be stored in a computer using a digital stethoscope and can be subjected for further processing. All these features make heart sounds suitable for the use in Biometric system. First and second heart sounds are produced from the valves present in the heart. Therefore the sound produced depends on the valves. As the valve's nature differs from person to person. Therefore the sound created by the same also differs. These are evident in the frequency components present in the heart sound. Figure 1.2 shows four PCG signal segments. Two on the left belong to person1 and two in the right belong to person 2. The frequency domain representations of these signals are also given just below to the time domain representations [4]. It shows that for an individual the frequency components present in a segment is somewhat related. Though because of the presence of murmur it cannot be easily be differentiated in only frequency domain information. Later we will see a time frequency transformation generates a better representation for biometric purposes.

1.4 Mechanism of Heart Sound Production

The human heart has four chambers, left and right Atrium, left and right Ventricles. Inside the heart, blood flows from Atrium to Ventricles and from Ventricles it is pumped

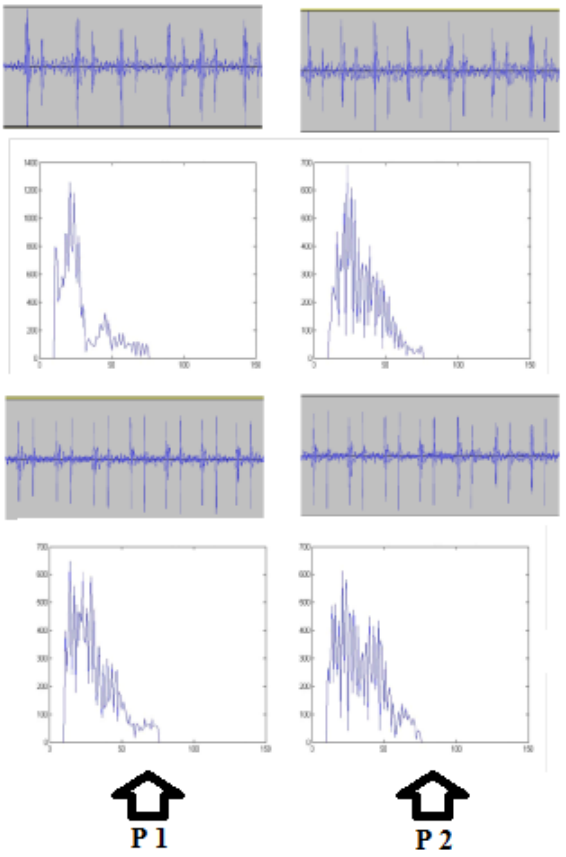

FIGURE 1.2: Frequency representation of PCG signals.

out from the heart through Pulmonary Artery and Aorta[5]. The process of pumping the blood happens in two stages, Diastole and Systole.

DiastolicPeriod :

During this period the ventricle relaxes and thus the pressure inside ventricle drops. When the pressure drops below the pressure in the atrium, the Mitral and Tricuspid valves open and blood flows into ventricles from the atrium.

SystolicPeriod :

During this period the ventricles contract. The sudden increase in pressure closes the Mitral and Tricuspid valves and opens the Aortic and Pulmonary valves and pumps out the blood.

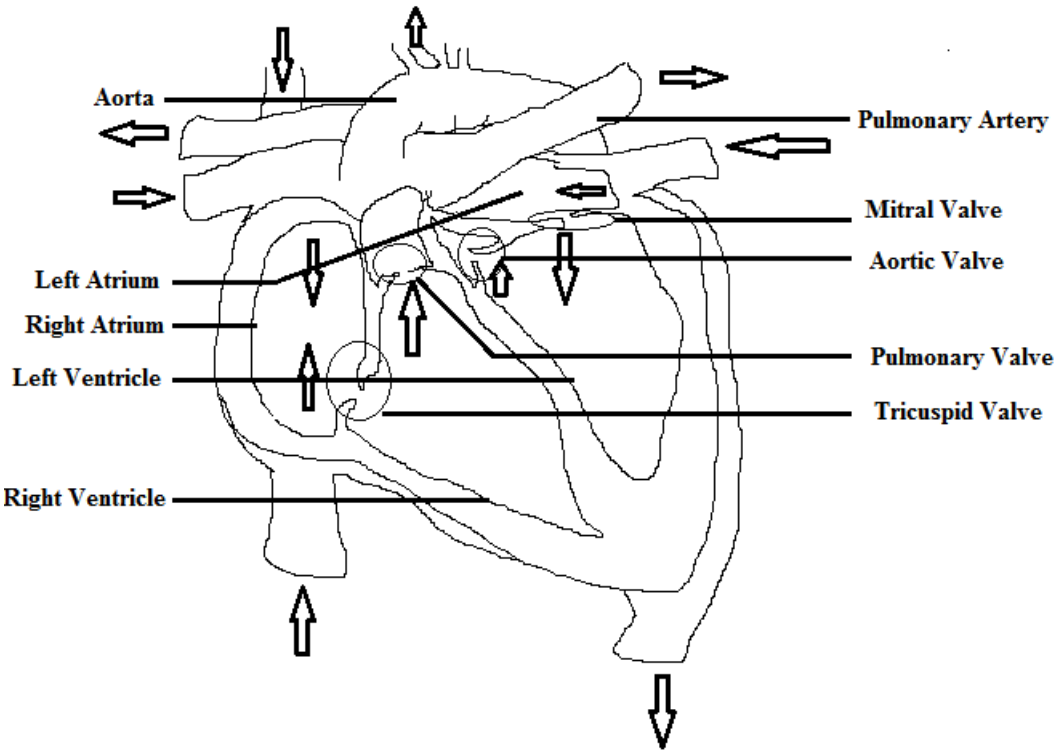

FIGURE 1.3: Cross section of a human heart.

Heart sound is in the range of 15-200 Hz. Two major heart sounds are the S1 or lubb and S2 or dubb. S1 sound is caused by the sudden blockage of reverse blood flow due to closure of the Atrioventricular valves, i.e. Tricuspid and Mitral (Bicuspid), at the beginning of ventricular contraction, or Systole. S2 sound is caused by the sudden blockage of reversing blood flow due to closure of the Semilunar valves (the Aortic valve and Pulmonary valve) at the end of ventricular systole, i.e. beginning of ventricular diastole. Figure 1.4 shows a picture of heart sounds in time domain and shows the locations of S1, S2 and Murmur.

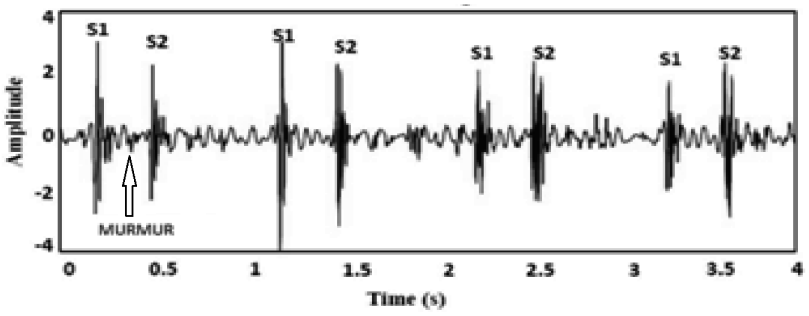

FIGURE 1.4: Normal phonocardiogram signal.

1.5 Auscultation Process of Heart Sound

Auscultation is the process of hearing the sounds produced in the body, usually with a stethoscope. There are four auscultation sites in the chest region as depicted in the Figure 1.5 below. These are Aortic, Pulmonary, Lower Sternal Border and Mitral.

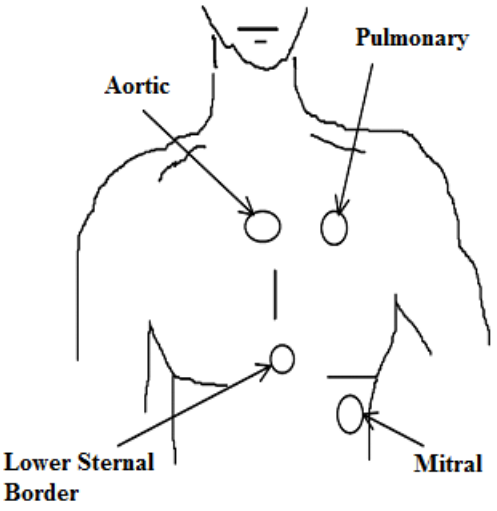

FIGURE 1.5: Auscultation sites in the chest region.

For the project it was required to store heart sounds for further processing. The digital Stethoscope manufactured by HDFono Doc has the USB connectivity and the heart sound can be stored in Personal Computer directly.

FIGURE 1.6: Data Collection from a volunteer using Digital Stethoscope.

1.6 Data collection

Unlike ECG, EEG or Voice signals, PCG signals are not freely available on a large scale for research purposes. So we had to create our own database of PCG signals. In this work ten volunteers contributed to the construction of the database of heart sounds. A digital stethoscope manufactured by HD Medical Services (India) Pvt. Ltd, was used. Through the USB connectivity the instrument can directly store the heart sound into the PC in Waveform Audio File Format, also known as WAV format. The volunteers were in the age group 20-35. For a particular volunteer each sample (of 20 second duration) was collected with a minimum time gap of one hour. This process of data collection continued for three months. All the volunteers were having normal heart sounds with no prior knowledge of heart related diseases.

1.7 Previous Works

The proposition of phonocardiogram signal for identification has been studied in many literatures so far. In this section we discuss some most important published works, their merits and demerits. We also discuss the implementation of most commonly used linear frequency band cepstral coefficient based feature and the identification result that we got with the created database.

1.7.1 Literature review

The option of Phonocardiogram signal in biometric identification was first introduced in [6]. The proposed method required localization and delineation of S1 and S2 which is followed by Z-Chirp (CZT) transforms. Feature set was classified using the Euclidean distance measure. [3] propose a Short Time DFT (STDFT) method to get Cepstral feature set followed by classification using Gaussian Mixture Model (GMM). [7] further improves the method using wavelet based noise reduction technique. But the GMM based technique is slow which creates a disadvantage for biometric usage. A fusion technique of Mel Frequency Cepstral Coefficient (MFCC) and First to Second Ratio (FSR) was introduced in [8]. An EER of 8.70 was claimed in the paper though the testing was performed upon manually chosen cleanest signals, making it a non realistic hypothesis as mentioned in [9]. [10] proposed a method for selection of clean heart sound adding practicality to the method proposed in [8]. But the fusion and selection of clean heart sound increases computational cost and time both. [9] used filter bank based Cepstral analysis. Parameters of the filter banks were varied to get the optimum values.

1.7.2 Implementation of Linear Frequency Band Cepstrum coefficient technique

1.7.2.1 Feature Extraction

The process of feature extraction starts with Discrete Short Time Fourier Transform (DSTFT) of the stored PCG signals [11][3]. DSTFT is a special class DFT technique where we use a shifting window. If x[n] is the discretize signal then DSTFT is given by:

$$DSTFT\{x[n]\} \equiv X\left(m,\omega\right) = \sum_{n=-\infty}^{\infty} x\left[n\right]\omega[n-m]e^{-i\omega n} \tag{1.1}$$

With window w[n] and m is the shift. For heart sounds we use a window of 0.5 Sec with a shift of 250 milliseconds. After performing the transform only the magnitude part is selected as it is less noisy prone. Next the signal is band pass filtered 20-150 Hz. The logarithm of the coefficients is taken as it helps to neutralize the effect of the air column channel of the Stethoscope. And then we perform Discrete Cosine Transform (DCT) to get the desired Cepstral coefficients. The overall process can be represented mathematically as:

$$x_n\left(m,\omega\right) = \log[X(m,\omega)] \tag{1.2}$$

$$X_k = \sum_{n=0}^{N-1} x_n \cos[\frac{\pi}{N}(n+\frac{1}{2})k] \tag{1.3}$$

So, X_k is our desired cepstrum coefficients. K denotes the total number of frequency bins present between 20 to 150 Hz. We reduce the dimension of feature vector to 24 by selecting only first 24 cepstrum coefficients. It is done so because it has been seen that the higher order cepstrum contain very little information.

After the dimension compression the next step is the peak removal, where we intend to remove the artifacts due to hand movement and other spurious signals collected unintentionally. To implement this we take the energy of each segment E[n], where n denotes the segment index. We take a threshold value of 15dB.

$$10lgE\left[n\right] - min_n\left(10lgE\left[n\right]\right) \geq \mu \tag{1.4}$$

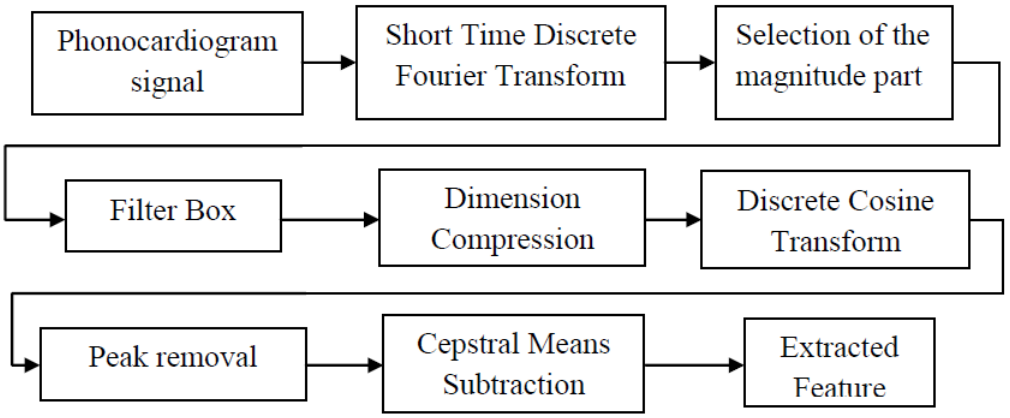

FIGURE 1.7: Block diagram of the LBFCC feature extraction process.

Next, we try to nullify the effect of the channel into the signal. The challenge is a channel transfer function depends upon the place from where data is collected which cannot be kept exact same point all the time. So it is impossible to nullify the effect completely. So we have collected the mean over a range of data and subtracted the mean of the data. If the dimension of the data is k (Cepstral coefficient/window) total transfer function of the process can be denoted as Z[n,k] and the process transfer function as X[n,k] and the average channel transfer function Y[k]. So we can write:

$$Z(n,k) = X[n,k] \times Y[k] \tag{1.5}$$

As the log Cepstral coefficient is taken, the process has become additive.

$$\log\left(|Z\left[n,k\right]|\right) \ = \log\left(|X[n,k]|\right) \ + \log(|Y[k]|) \tag{1.6}$$

So the Cepstral coefficients $C_Z[n,k]$ can be written as:

$$C_z(n,k) = C_x(n,k) + C_y(k) \tag{1.7}$$

Now after subtracting the mean $\langle C_{Z,k}[n]\rangle$over n data we can theoretically cancel the effect of the channel.

$$C_{z,k}\left[n\right] - \langle C_{z,k}\left[n\right]\rangle = C_{x,k}[n] - \langle C_{x,k}[n]\rangle \tag{1.8}$$

But in reality the effect is never fully cancelled.

1.7.2.2 Classification schemes

These features are classified using BP-MLP-ANN [12], Linde Buzo Gray Vector Quantization (LBG VQ)[13] and Gaussian Mixture Model (GMM-4).

A. BP-MLP-ANN

Multiple layers Neural Network in principle can provide an optimal solution to an arbitrary classification problem. The key power provided by such networks is that they use fairly easy algorithms where the form of the non linearity can be learned from training data. The models are thus extremely powerful and apply well to a vast array of real world application. We used Feed forward Multilayer Perceptron (MLP) Neural Network (NN) with the nonlinear sigmoid function. We have used 25 hidden nodes.

B. LBG Vector Quantization:

The VQ design problem can be stated as follows. Given a vector source with its statistical properties known, given a distortion measure, and given the number of code vectors, find a codebook and a partition which results in the smallest average distortion [14]. The LBG VQ design algorithm is an iterative algorithm. The algorithm requires an initial codebook. This initial codebook is obtained by the splitting method. In this method, an initial codevector is set as the average of the entire training sequence. The iterative algorithm is run with these vectors as the initial codebook. The process is repeated until the desired number of codevectors is obtained.

C. GMM

A Gaussian mixture density is a weighted sum of M component densities, and given by the equation:

$$p(\vec{x} \mid \lambda) = \sum_{i=1}^{M} p_i b_i(\vec{x}) \tag{1.9}$$

Where $\overrightarrow{x}$ is a D-dimensional random vector,$b_i(\overrightarrow{x})$, $i = 1,\ldots,$ M, are the component densities and p_i , $i = 1, \ldots,$ M, are the mixture weights. Each component density is a D- variate Gaussian function of the form:

$$b_i(\vec{x}) = \frac{1}{(2\pi)^{D/2}\,|\Sigma_i|^{\frac{1}{2}}} \exp^{-\frac{1}{2}(\vec{x}-\vec{\mu_1})\Sigma_i^{-1}(\vec{x}-\vec{\mu_1})} \tag{1.10}$$

With mean vector μ_i and covariance matrix $\sum_i$. The mixture weight satisfies the constraint that$\sum_{i=1}^{M} p_i = 1$. The complete Gaussian mixture density is parameterized by the mean vector, covariance matrix and mixture weights from all component densities. These parameters are collectively represented by the notation:

$$\lambda = \{p_i, \vec{\mu_i}, \Sigma_i\} \tag{1.11}$$

Each heart sound is represented by model λ. The Expectation Maximization (EM) algorithm is usually used due to its simplicity and quick convergence. We used the codebook of LBG-VQ as initial GMM fitting by EM algorithm. In the experiment four mixture densities GMM (GMM-4) is used. So in training for each person the model λ is calculated and in testing each heart sound is referred by maximum likelihood' measures. So the maximum value is taken as the possible identified person.

1.7.2.3 Results

Table 1.1 gives the identification results we obtained from the tests. The accuracy of LBG-VQ is varying significantly with changing Cepstral coefficients per window. In this experiment the accuracy achieved with 40 Cepstral coefficients is higher than 24 or 60 Cepstral coefficients.

1.7.2.4 Conclusion

This experiment conducted in our normal lab environment indicates that the LFBCC based features are not giving high identification accuracy when tested with collected database. This could very well be because of the noisy nature of the collected database. In normal condition also we have to deal with these problems. Wavelet can provide solutions to these problems. It has already been shown that a use of wavelet transforms for noise removal is very effective.

TABLE 1.1: Result of LFBCC feature set

	NEURAL NETWORK			LBG-VQ			GMM-4	
NAME (10 Samples each)	No of training samples/ person			No of Cepstral Coefficient/window			No of training samples/ person	
	150	300	450	24	40	60	300	450
Ajay	10	10	10	6	6	6	1	2
Anil	10	10	10	9	10	6	8	10
Arunava	9	10	10	8	10	9	7	10
Ashish	7	8	8	3	3	6	3	3
Brajesh	7	9	10	5	8	7	4	7
Dipak	8	9	10	2	7	8	0	2
Jayprakash	5	7	7	6	8	6	3	3
Manab	8	8	9	6	9	8	8	10
Total correct Detection	64	69	72	44	61	56	31	47
% Accuracy	80	86	92.5	55	76.2	70	39	59

1.8 Theoretical background to wavelet transform

In wavelet theory a function $f(x)$ of a continuous variable x, which belongs to a square integrable subspace $L^2(R)$, i.e. $f(x) \in L^2(R)$ can be written as [15][16]:

$$f(x) = \sum_s a_{r0,s}\phi_{r0,s}(x) + \sum_{r=r0}^{\infty}\sum_s b_{r,s}\psi_{r,s}(x), here[r \geq r0] \tag{1.12}$$

Here $\phi_{r,s}(x)$ represents a scaled and shifted (r=scaling parameter,s=shifting parameter) version of a function $\phi(x)$, called basic scaling function . It is of the form:

$$\phi_{r,s}(x) = 2^{r/2}\phi(2^r x - s) \tag{1.13}$$

$\phi_{r0,s}(x)$ is a scaling function whose scaling parameter is fixed in $r = r0$. And $\psi_{r,s}(x)$ represents a scaled and orthogonally shifted version of a function $\psi(x)$ called 'basic wavelet function'. And it is of the form:

$$\psi_{r,s}(x) = 2^{r/2}\psi(2^r x - s) \tag{1.14}$$

The $\psi(x)$ function can be represented as serirs sum of ϕ function in the following manner,

$$\psi(x) = \sum_n h_\psi(n)\sqrt{2}\phi(2x - n) \tag{1.15}$$

Provided the function $f(x)$, the scaling function $\phi(x)$, and wavelet function $\psi(x)$ is given, the coefficients and can be found by the following equations respectively,

$$a_{r0,s} = \int_{-\infty}^{\infty} f(x)\phi_{r0,s}(x)dx \tag{1.16}$$

$$b_{r0,s} = \int_{-\infty}^{\infty} f(x)\psi_{r,s}(x)dx \tag{1.17}$$

In real life however we more often than not encounter discrete signals for processing. So instead of continuous function x if we consider a function of discrete samples n=0,1,....., M-1, then in discrete domain the equation (1.12) can be represented as:

$$s(n) = \frac{1}{\sqrt{M}} \sum_{k} W_\phi(j0,k)\phi_{j0,k}(n) + \sum_{j=j0}^{\infty} \sum_{k} W_\psi(j,k)\psi_{j,k}(n) \tag{1.18}$$

j and k are the discrete scaling and shifting parameters.$\frac{1}{\sqrt{M}}$is the normalizing term. $\phi_{j0,k}(n)$ and $\psi_{j,k}(n)$ are the discrete counterpart of the coefficients $a_{r0,s}$ and $b_{r0,s}$. These coefficients can be found out from the discrete equivalent of the equations (1.16) and (1.17).

$$W_\phi(j_0,k) = \frac{1}{\sqrt{M}} \sum_{n} s(n)\phi_{j_0,k}(n) \tag{1.19}$$

$$W_\psi(j,k) = \frac{1}{\sqrt{M}} \sum_{n} s(n)\psi_{j,k}(n) \tag{1.20}$$

The discrete equivalent of equation (1.14) and equation (1.15) can be written in discrete domain respectively as,

$$\psi_{j,k} = 2^{j/2}\psi(2^j n - k) \tag{1.21}$$

$$\psi(n) = \sum_{p} h_\psi(p)\sqrt{2}\phi(2n-p) \tag{1.22}$$

Putting $n = 2^j n - k$ in equation (1.22) and substituting the $\psi(2^j n - k)$ term in equation (1.21) we get the value of $\psi_{j,k}$. This value of $\psi_{j,k}$ is substituted in equation (1.20) to determine $W_\psi(j,k)$. This value can then be represented as,

$$W_\psi(j,k) = \sum_{m} h_\psi(m-2k)W_\phi(j+1,m) \tag{1.23}$$

Here $m = p + 2k$. And by similar process it can also br shown as,

$$W_\phi(j,k) = \sum_{m} h_\phi(m-2k)W_\phi(j+1,m) \tag{1.24}$$

Equation (1.23) and (1.24) are easily realizable [15]. As shown in the Figure 1.8.

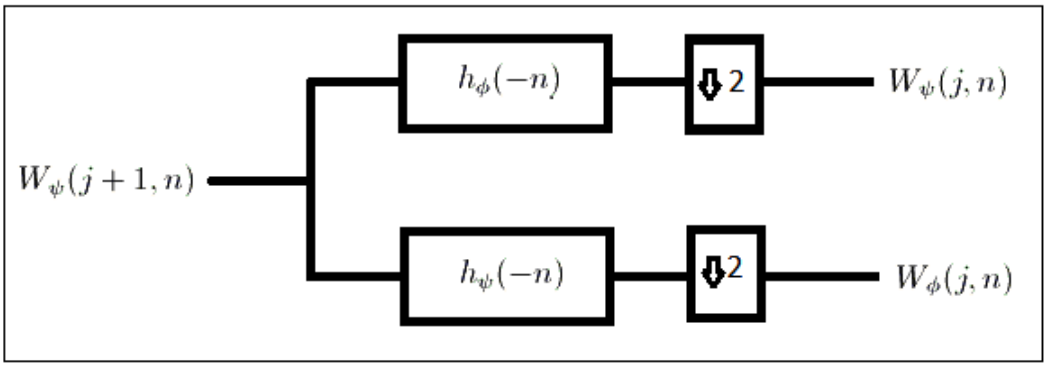

FIGURE 1.8: Realization of wavelet coefficients.

So a cascaded structure of highpass and lowpass filters followed by a down-sampler can realize the detail and approximation wavelet coefficients.

1.9 Proposal of Wavelet based approach

The peaks and valleys of S1 and S2 carry important biometric information for an individual. To entrap this information a time frequency domain analysis like wavelet transform should be suitable. Wavelet is a useful tool in assessing the local scale behaviour of functions and measures. Essential information about different irregularities of signal is carried in the corresponding maxima lines across the scales of the Wavelet transform which allow the accurate assessment of the properties of signal. Therefore, wavelet based approach which preserves the time frequency information plus gives much better performance in noisy data is a good option.

1.10 Brief overview of BP-MLP-ANN

The fundamental components of an ann is called nodes. Nodes act like artificial neurons. This concept is motivated by real neurons present in the nervous system.

1.10.1 Structure of a single node

In mathematical terms we can write:

$$v_k = \sum_{j=0}^{m} \omega_{kj} x_j \tag{1.25}$$

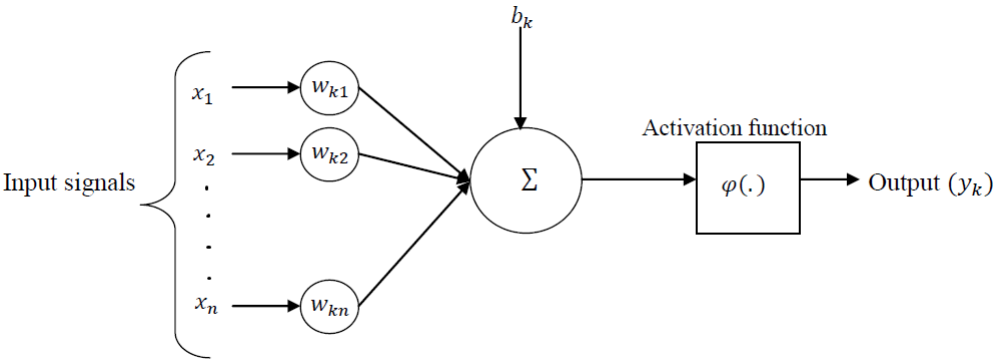

FIGURE 1.9: Structure of a single neuron.

Here $x_1, x_2, ...$ are the input signals and $\omega_1, \omega_2, ...$ are the synaptic weights of the kth neuron. υ_k is the activation potential of the neuron. Bias b_k is considered as $b_k = \omega_{0j} x_0$. Output of the neuron y_k can be represented as:

$$y_k = \varphi(\upsilon_k) \tag{1.26}$$

$\varphi()$ is called the activation function. Activation function can be deterministic or it can be probabilistic. Most popular and commonly used activation function is sigmoid function.

$$\varphi(v) = \frac{1}{1 + exp(-av)} \tag{1.27}$$

1.10.2 Multilayer Network

In general NNs are made of multiple hidden layers of nodes or artificial neurons. The overall function of this hidden layer is to make the network output follow the desired pattern. The network with multiple layer can follow higher order statistics.

1.10.3 Training

Training is the process to obtain the desired output when certain inputs are given. NN weights are adjusted according to the error signal generated. The error can be represented in many ways. In the simplest form it is the difference between the desired output and actual output.

1.10.4 Back Propagation Algorithm

In training as mentioned before, we change the synaptic weights in such a way that the overall error is reduced. The average squared error energy $\varepsilon_{av}(n)$ is obtained by adding $\varepsilon(n)$ overall n and normalizing with respect to the set size N. $\varepsilon(n)$ is the instantaneous value of the error energy obtained from adding the average individual error energy of a single node over all neurons in the output layer.

$$\varepsilon_{av}(n) = \frac{1}{N}\sum_{n=1}^{N}\varepsilon(n) \tag{1.28}$$

The change in the synaptic weight is given by:

$$\Delta\omega_{ji}(n) = -\eta\frac{\partial\varepsilon(n)}{\partial\omega_{ji}(n)} \tag{1.29}$$

η is called the learning rate parameter. The smaller its value the smaller will be the change in the weight and that means smoother trajectory in weight space. But the training time will increase if the learning rate is small. Partial derivative $\frac{\partial\varepsilon(n)}{\partial\omega_{ji}(n)}$ is called the $sensitivityfactor$. It can be shown that:

$$\frac{\partial\varepsilon(n)}{\partial\omega_{ji}(n)} = \partial_j(n)\cdot y_i(n) \tag{1.30}$$

Where $\delta_j(n) = e_j(n)\varphi'_j(\upsilon_j(n))$. So we get:

$$\Delta\omega_{ji}(n) = \eta\cdot\partial_j(n)\cdot y_i(n) \tag{1.31}$$

1.10.5 Perceptron

Suppose we have a set of learning samples consisting of an input vector x_j and a desired output y_k. For a classifcation task the y_k is usually +1 or -1 The perceptron learning rule is very simple and can be stated as follows [17]:

1. Start with random weights for the connections;

2. Select an input vector x_j from the set of training samples;

3. If $y_k \neq x_j$ the perceptron gives an incorrect response, then modify all connections w_{kj} according to:

$$\Delta w_{kj} = y_k \cdot x_j \tag{1.32}$$

4. Go back to 2.

Above, the fundamentals of Back Propagation Multilayer Perceptron Artificial Neural Network is given in the briefest way. For detail study in the above mentioned matters one is advised to read [18][19][20].

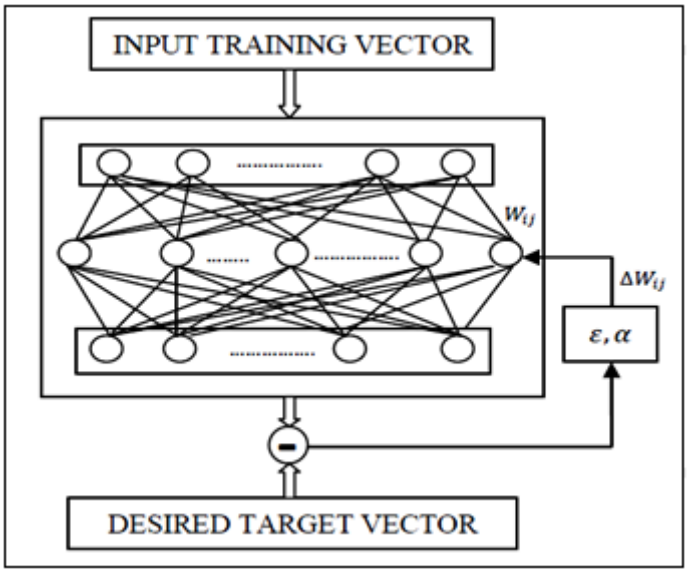

FIGURE 1.10: Structure of a Multi layer Back Propagation NN.

1.11 Proposal of Back Propagation Multi Layer Perceptron Artificial Neural Network as a classifier

Back propagation MLP ANN is very popular due to its simplicity and quick processing. MLP ANN is preferred choice of classification for speech recognition[21]. [22] uses ANN for classification of heart murmurs and a classification accuracy of 85% is achieved. [23] uses MLP ANN for wavelet based feature set for a heart murmur and the best result 86.4% is achieved. These literatures indicate the efficient performence of BP ANN in case of time-frequency domain features. The same is selected for the classification of the extracted feature in this project. For identification process the NN with 25 hidden

nodes is trained using 300 heart cycle features per class. A total of 10 classes of data is used as already mentioned before. Input training vector is provided to the input nodes. These input values flow in the forward direction from input to hidden to output layers to get the output values. The output values are compared with desired target values to generate error E signal. This error signal is then sent back to change the weight W_ij. This process continues until the error E reaches a permissible optimal value E_T. A detail of Multilayer Perceptron structure and its algorithms are given in [17],[19] and [18]. Once the training or learning is complete we sent the testing samples in the input nodes and check the output from output nodes.

1.12 Motivation

The very idea that the heart auscultations can be used to identify an individual is surprising and interests many. Plus the fact that Phonocardiogram signals of an individual cannot be reconstructed by artificial means is a big advantage as a biometric trait. Many of the previous approaches to extract Biometric information had been Cepstral coefficient based. The experiments conducted in our laboratory indicates cepstral based processes lacked the same accuracy for noisy data. Secondly, the GMM based classification approach will require too much time for practical implementation. While ANN will take time for enrollment of a person but once enrolled and the network is created it will be very fast compared to GMM based systems. Overall the field of Biometric identification based on phonocardiogram is still an active field of research. The idea to try a wavelet based novel feature set was motivating enough to go for the project.

1.13 Thesis outline

The rest of the thesis is organized in the following manner. Chapter two describes the preprocessing method. Preprocessing technique is composed of Normalization, Low pass filtering, Heart cycle extraction, Aligning and Segmentation. Chapter three describes the feature extraction process from the segmented S1 and S2. Then we describe the classification process and creation of the decision vector followed by the results. First the Imposter testing based on mean square error is described. Then the identification and verification results are shown. Chapter four describes the conclusion and future work.

Chapter 2

Preprocessing

Feature extraction is the process of extracting meaningful information from a signal. In this project we require features that can describe the individuality of a class that the PCG signal belongs to. Our information lies in the S1 and S2. That is why before feature extraction segmentation of S1 and S2 is necessary. In preprocessing a PCG signal is processed and made ready for feature extraction. It is similar to cleaning and cutting the vegetables before actual cooking begins. The preprocessing process is shown in the block diagram in Figure 2.1.

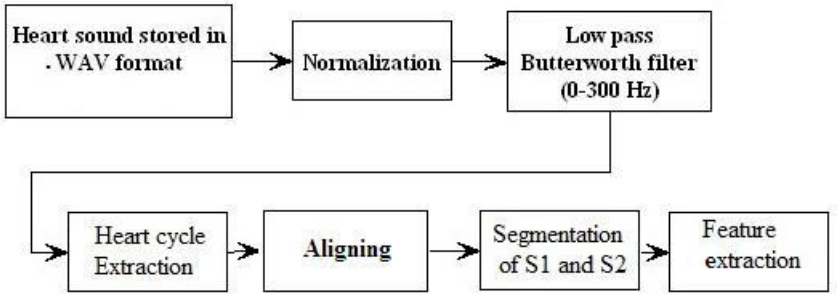

FIGURE 2.1: Preprocessing.

2.1 Normalization

Before passing through filtering we normalized the signal. This is the simplest normalization technique to make the signal vary between -1 to +1. It can be done in two steps. In the first step we find out the maximum absolute value of the signal and in the second step we divide the whole signal by that value. The resulted signal will be within +1 to -1 range.

2.2 Low pass filtering

We used a Low pass Butterworth filter of the order 10 in the range 0 to 300 Hz. This will free the signal of high frequency noises and artifacts.

2.3 Extraction of Heart cycle

Heart cycle is a quasi periodic signal whose normal period is within the range of 0.4 - 1.2 second. Some cycle extraction techniques are presented in [24], [25] and [26]. A technique described in [27] and [28] for detection of QRS complex in ECG signal can also provide a cycle detection using PCG signals because of the quasi-periodic nature of the PCG signal. Steps followed to implement the technique is described below.
Step 1: Three seconds PCG sample is extracted.
Step 2: Cardiac cycle is between 0.4 to 1.2 Sec. A lag index is created ranging from 0.4 to 1.2 second with step of 0.005 second.
Step 3: Now sum of the Autocorrelation of the sample with these lag indexes are determined.
Step 4: The position of the max peak of this sum of autocorrelation is determined.
Step 5: The corresponding lag time of the determined lag index is the determined heart cycle.
Step 6: The whole sample is divided into cycle blocks.

Process is displayed in Figure 2.2. At the top the block shows the three Sec extracted PCG signal. Just below to that is the sum of the autocorrelation output of the lag indexes. The bottom block shows the corresponding lag time of the lag indexes. This lag time is the determined heart cycle.

2.4 Aligning

Aligning process contributes to the segmentation process. Extracted heart cycle blocks were aligned using cross correlation between two different blocks for a same sample. Figure 2.3 shows two extracted cycle blocks which were correlated. Correlated function will have three regions which are labelled as A, B and C in the figure. Region A in the cross correlation is generated because $S2$ of second block coming under $S1$ of first block, B because $S1$ and $S2$ of the second block coming under the $S1$ and $S2$ of the second block, C is coming when $S1$ of second block coming under $S2$ of the first block. The

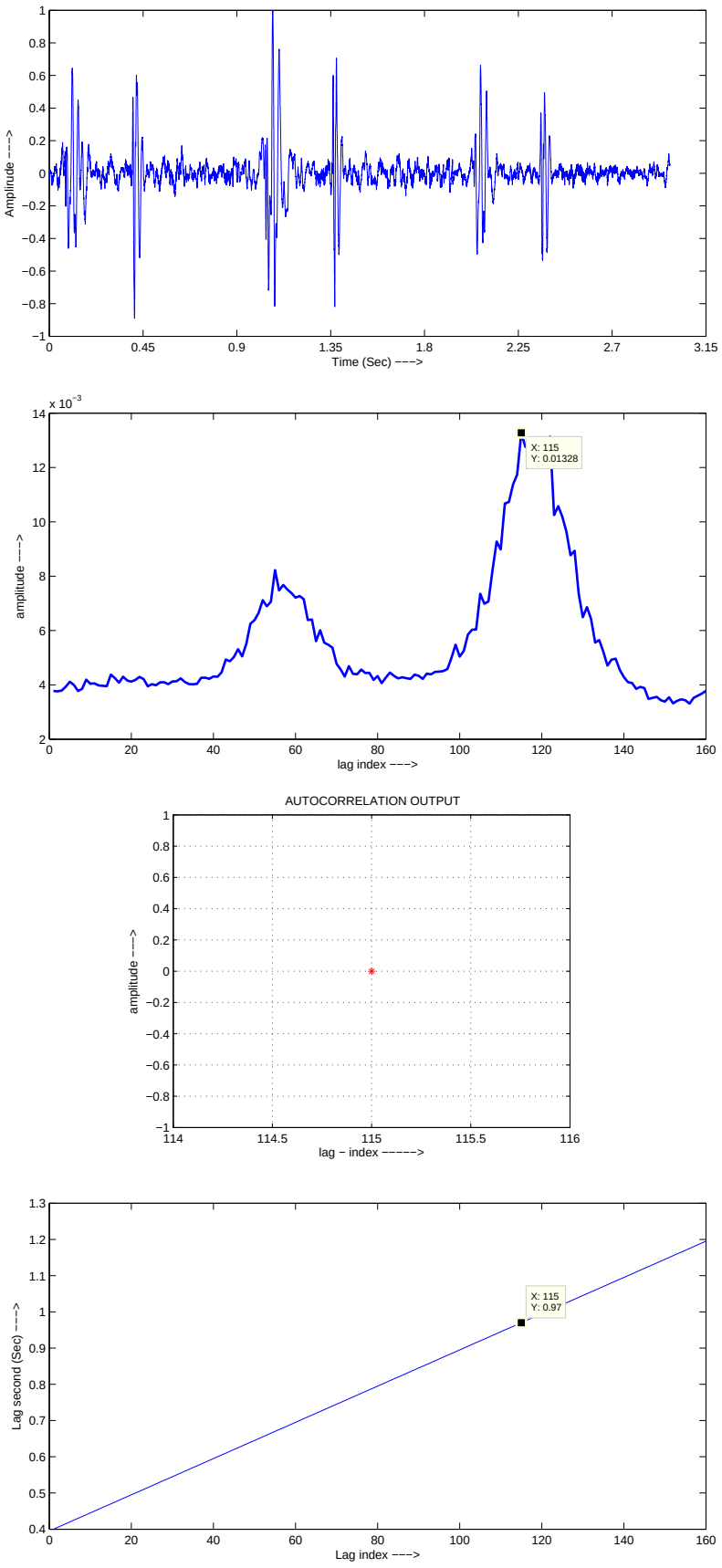

FIGURE 2.2: Extraction of Heart cycle.

peak is coming when the both blocks are perfectly aligned. The distance of the peak from the center point of the correlation window is the amount of shift required to align the two blocks properly.

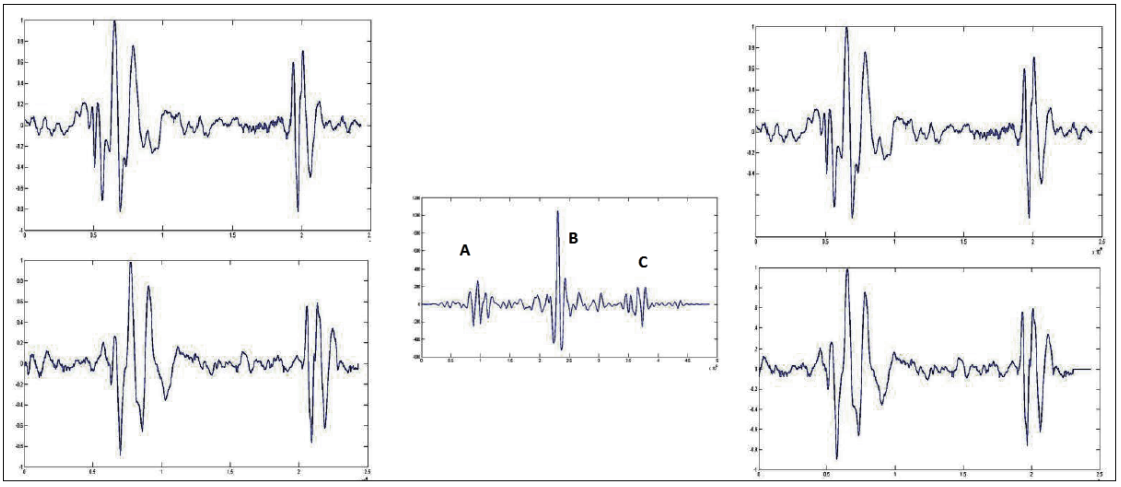

FIGURE 2.3: Heart cycle aligning process.

After the aligning process is over the segmentation process becomes easy because now the S1 and S2 lies in the same locations of the block compared to the first cycle block. So in the next step of segmentation we only find out the locations of S1 and S2 for the first block and since all the blocks are aligned according to the first block we apply the same location values to determine the beginning and the end of S1 and S2 for all the blocks of this sample. Figure 2.4 shows the heart cycle before and after aligning and its crosscorrelation window. At the bottom the first cycle is shown which is taken as strandard for aligning all other blocks.

2.5 Segmentation

Let x(n) is the first heart cycle block. The process of segmentation is described below:
Step-1) We find $y(n) = x^2(n)$.
Step-2) We take window of 50. We place the window in the beginning of the y(n). Take the mean of the window. Place it in Z(k). We slide the widow by one sample and do the same operation. This will remove if any spike is present in the sample. If size of y(n) is m then size of Z(k) will be (m-49)
Step-3) We take a window of 400 samples. Place it at the beginning of the Z(k). Take the maximum value and place it in a(i). We slide the window 400 samples with zero overlap.
Step-4) We take a threshold of 0.07. The crossing points of the threshold is multiplied

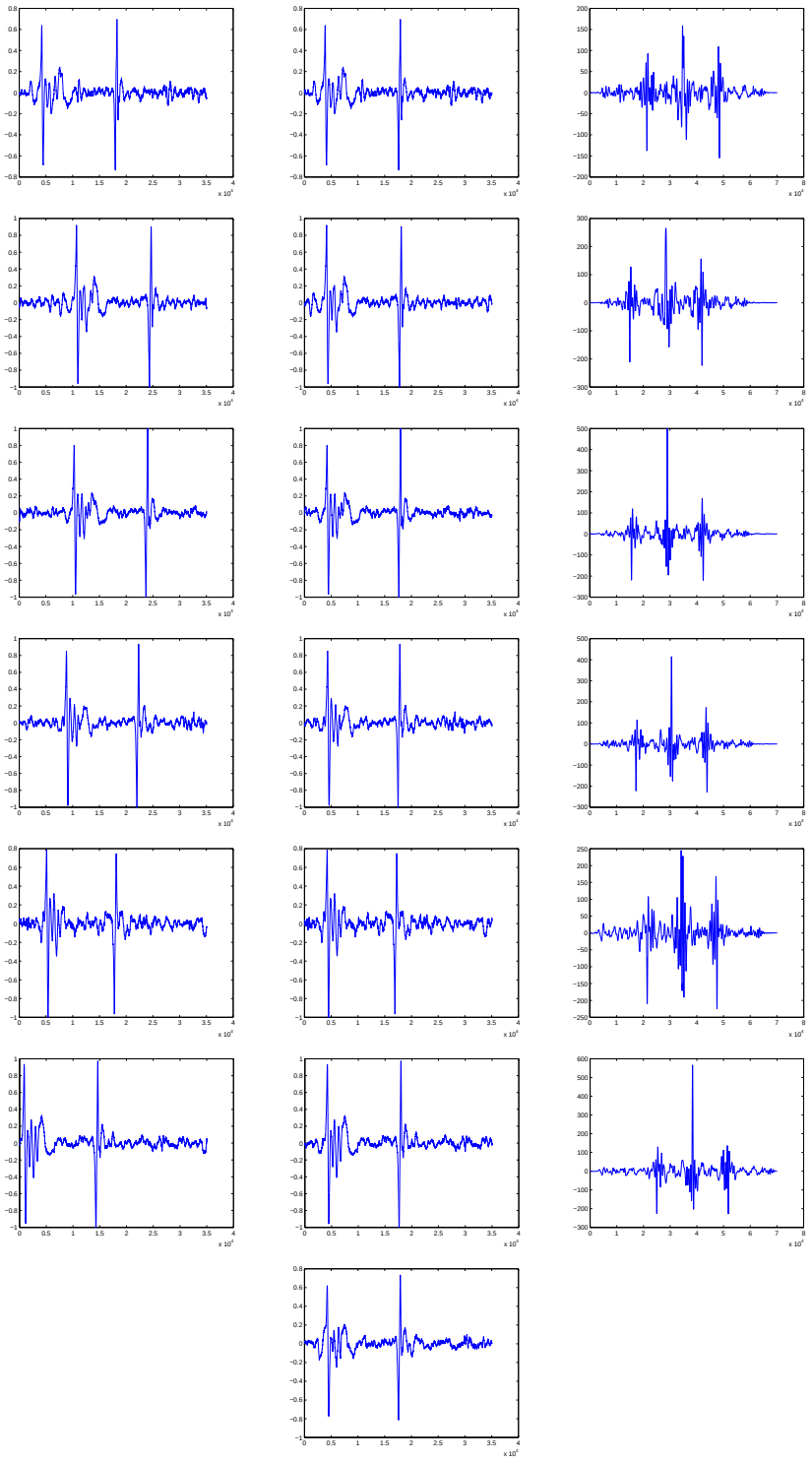

FIGURE 2.4: Heart cycle before and after aligning.

with 400 to get the beginning and end of S1 and S2.

Step-5) Since all blocks are aligned and of same size, we segment all the aligned blocks. Figure 2.5 shows the process of segmentation explained above. Second block is after squaring mean window operation is over. And last block is after max window has operated. We apply the threshold of 0.7 after this.

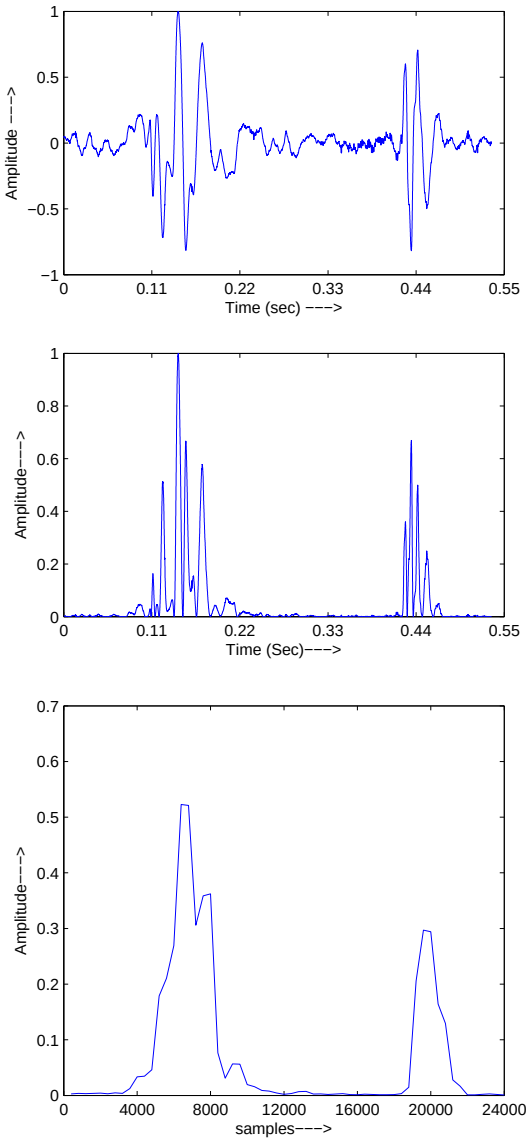

FIGURE 2.5: Process of heart cycle segmentation.

Chapter 3

Feature Extraction and Classification

3.1 Feature Extraction

The extracted S1 is sent first followed by S2. Using Dubeshie's 2nd order wavelet the signal is decomposed upto second level.

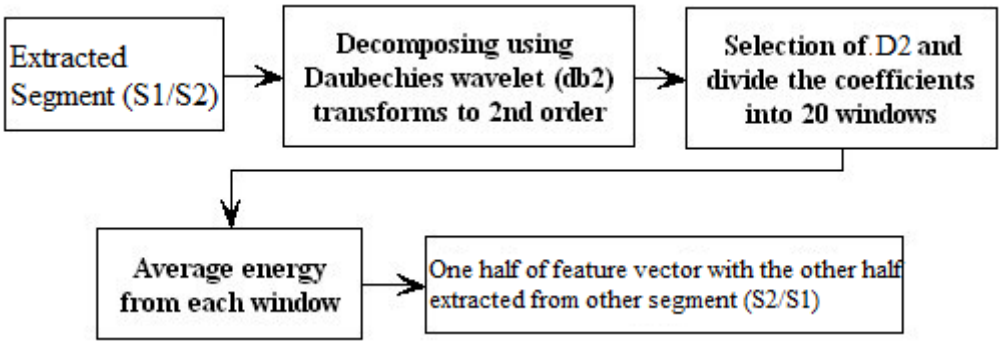

FIGURE 3.1: Block diagram of feature extraction process.

As shown in the Figure 3.1 the segmented parts are send for feature extraction first S1 is processed. Using Daubescies second order wavelet. Daubescies second order wavelet is displayed in Figure 3.2.

We decomposed the signal upto second level. Because we already have lowered the frequency range upto 300 Hz, the frequency 75 to 150 Hz is entraped in D2 of the detail coefficient. Figure 3.3 shows the frequency ranges of different coefficients.

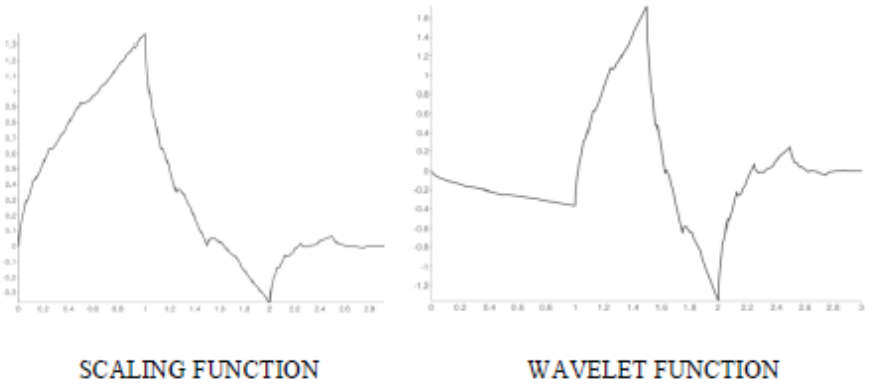

FIGURE 3.2: Second order Daubechies wavelet.

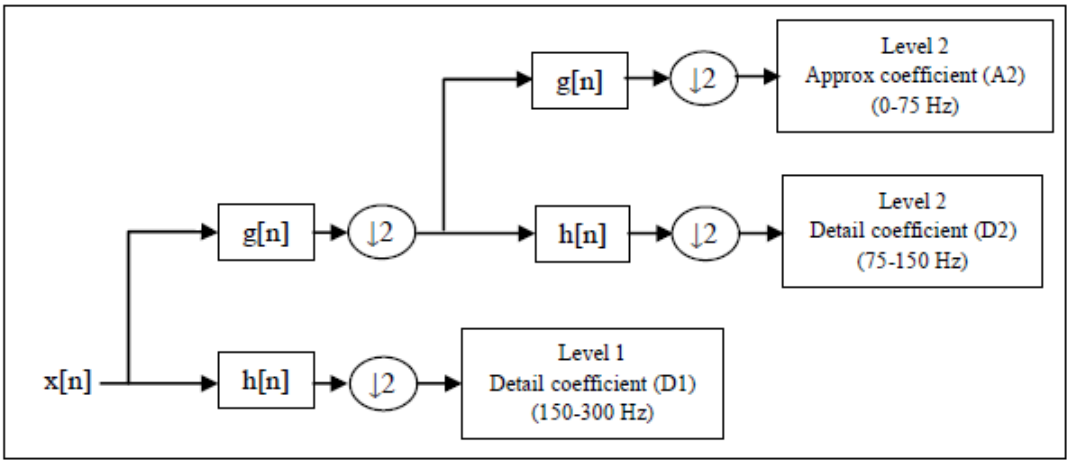

FIGURE 3.3: Second level wavelet decomposition.

Selected second level detail coefficient D2 is divided into N= 20 windows. And the average energy is determined from each window using the equation 2.1.

$$F[k] = \frac{1}{N} \sum_{n=1}^{N} X_{WT}^{2}[n] \tag{3.1}$$

After 20 coefficient is generated from S1 the same process is applied into S2. So finally we get a feature length of 40. Where first 20 is feature extracted from S1 and next 20 is feature extracted from S2. To check the nature of the extracted feature below we show the plot of features. Plot shows for same class the feature vectors extracted matches very closely, where as for different persons it varies significantly. Figures 3.4, 3.5 and 3.6 shows the plot of two samples of each classes.

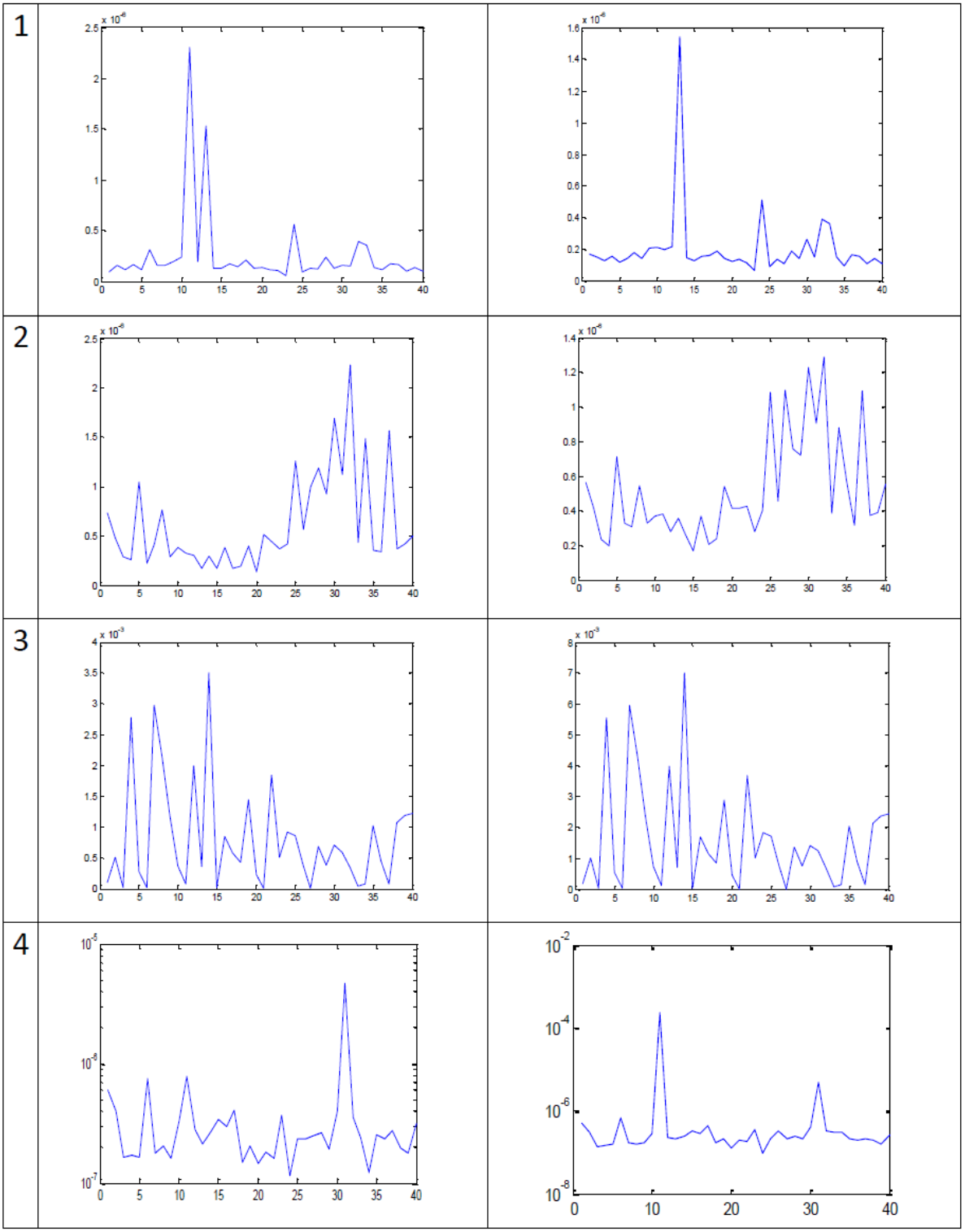

FIGURE 3.4: Plot of the extracted feature from class 1 to 4.

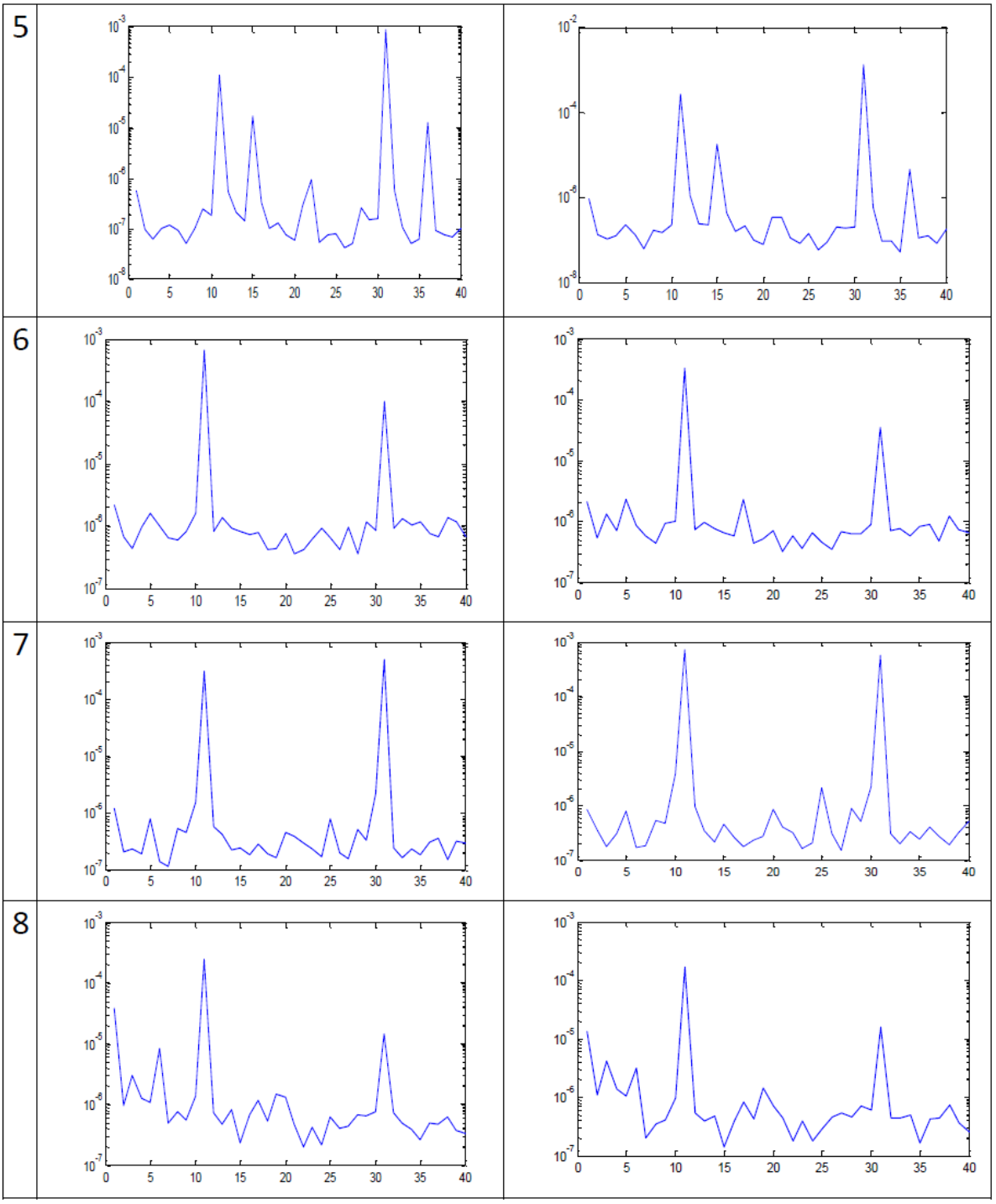

FIGURE 3.5: Plot of the extracted feature from class 5 to 8.

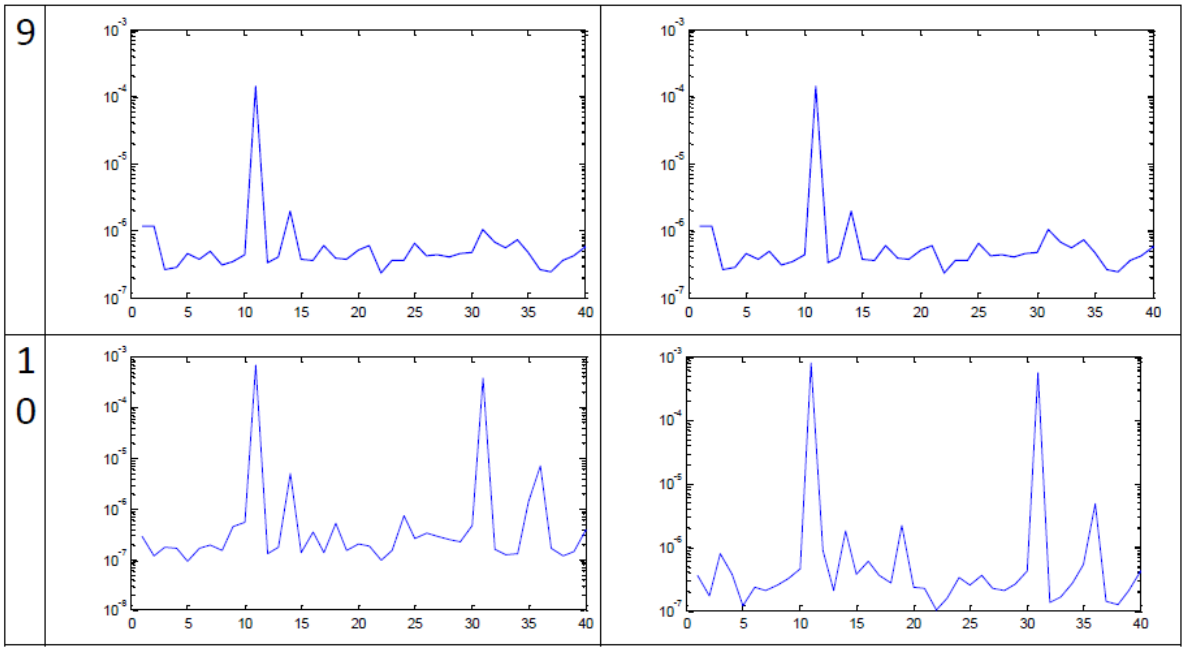

FIGURE 3.6: Plot of the extracted feature from class 9 and 10.

3.2 Classification

[22] used ANN for classification of heart murmurs and got classification accuracy of 85%. [23] shows for wavelet based feature set for a heart murmur and the best result 86.4% came for the MLP ANN classifier. So for classification the BP MLP ANN is preffered as we are also using wavelet based feature set. In the project we used score generated output from the ANN to determine the class of the testing sample. 20 Sec long heart sound testing samplea are processed. And from every sample 15 heart cycles are extracted. This generate 15 feature vectors.After normalization these are sent to the trained ANN. It was seen by trial and verify the method that neural network with 25 hidden nodes is giving the best result. The specifications of Neural Network used in Matlab is given in Figure 3.7.

15 neural network output scores for a same sample are added together to get the final decision vector of dimension 10. The highest value indicates the class identified by the system.

Table 3.1 shows the identification process. The neural network produces an output score for every feature vector extracted from one heart cycle. We add 15 such outputs of one single sample to produce our final decision vector. The highlighted part is the indicated

```
net=newff(minmax(train_data),[h s],{'tansig','tansig'},'trainlm');


% INITIALIZE NETWORK PARAMETERS
net = init(net);
net.trainParam.show =1;
net.trainParam.lr = .9;
net.trainParam.mc = 0.9;
net.trainParam.lr_inc = 1.01;
net.trainParam.lr_dec = 0.99;
net.trainParam.epochs = 1000;
net.trainParam.goal = 0.00001;
net.trainParam.max_fail = 5;
net.trainParam.min_grad = 1e-20;
net.trainParam.max_perf_inc = 1.04;
net.trainParam.time = Inf;

% TRAIN THE NETWORK WITH TRAINING DATA SET
net=train(net,train_data,target_trn);
```

FIGURE 3.7: ANN parameters.

TABLE 3.1: NN Outputs

-0.02663	0.981454	0.113668	0.574331	0.958242	0.486241	0.955438	-0.08817	0.615921	-0.04874	0.142927	0.9794	0.979398	0.022023	0.025331	6.670825
-0.00703	-0.04274	0.013481	0.073117	-0.17038	0.104555	0.232023	0.038937	0.011198	0.192994	0.048304	-0.03554	-0.0356	0.046662	0.046907	0.516883
0.054376	0.114716	0.017016	0.083482	0.00578	0.000386	-0.08472	-0.03645	0.049229	0.110458	0.042598	0.01831	0.018318	0.082215	0.082202	0.557921
0.053913	-0.10814	0.085746	0.097994	-0.26131	0.149235	0.997438	0.997709	-0.1503	0.101007	0.101183	-0.15635	-0.15638	0.089885	0.090631	1.932273
0.137382	-0.07288	0.071298	0.073494	-0.01547	0.013546	-0.06498	-0.08131	0.051063	-0.10549	0.052273	0.225751	0.225709	0.112875	0.11241	0.735667
0.068259	0.150234	0.135272	0.00774	0.99624	0.161082	-0.01571	-0.04924	0.518128	-0.06556	0.151382	-0.09712	-0.09715	0.040452	0.040025	1.944029
0.278616	0.094654	-0.02768	0.054798	-0.09454	-0.07667	0.071868	0.04369	-0.00457	-0.08659	0.001286	0.209691	0.209568	0.118181	0.117704	0.910015
0.087914	-0.55491	0.251095	-0.15884	0.189415	-0.07686	-0.68817	-0.20709	0.122875	0.663135	0.074173	0.110591	0.11096	0.265392	0.262448	0.452126
0.0239	0.07836	0.186951	0.08437	-0.36629	0.197033	-0.11094	0.052522	-0.02036	0.136953	0.222869	-0.26722	-0.26729	0.103132	0.10332	0.1573
0.293553	-0.00257	0.117387	0.049344	-0.13144	0.028144	0.055486	0.029577	-0.00064	-0.00919	0.126724	-0.10231	-0.10227	0.071001	0.071387	0.494184
NEURAL NETWORK OUTPUT FOR CLASS 1 SAMPLE TESTING															SUM

vector. In Table 3.1 the output indicates class one.

We also implemented the imposter checking and verification testing method using mean square error (MSE) of the normalized features. It also determines the fitness of the extracted feature. The process is developed below.

step 1: Normalize the extracted wavelet feature.

step 2: Add 10 normalized feature of same class and divide by 10 to get the mean.

step 3: Find the cumulative sum of the vector.

step 4: Three such features are created from each class.

3.3 Feature fitness and Imposter Detection

The normalized extracted feature of same class shows similarity. We developed a feature based verification cum imposter checking system. Mean of 10 normalized features of same class were determined as mentioned in the previous chapter. We compared randomly chosen three classes 1, 3 and 7, each with three such cumulative sum features as just described in the previous chapter. We determined the mean square error (MSE) between these selected features wirh all classes, i.e. total 30 samples, three from each class. The result of the mean square error is displayed in Table 3.2 below.Result shows the MSE of the same class (highlighted) is much less compared to the MSE of the other classes. A threshold value of MSE can detect an imposter and can be used for verification.

TABLE 3.2: MSE chart for imposter testing for randomly choosen class 1,3,7

		P1			P3			P7		
		1	2	3	1	2	3	1	2	3
P1	1	0	18.07679	50.46065	511.3741	42209.85	789337.3	24642.02	2246.908	127914.3
	2	18.07679	0	113.753	783.1527	41085.26	782580.9	23407.94	1959.443	125499
	3	50.46065	113.753	0	524.7715	45143.53	800818.3	26456.44	2950.077	132889.8
P2	1	805496.3	798514.5	816670.9	28783.03	32527.63	34188.71	33344.64	29086.36	33904.74
	2	924956.8	919212	938393.9	6606285	6659849	6685254	6675506	6608261	6680912
	3	6686271	6671763	6722168	765319.8	779163.1	787530.5	41352.82	36239.25	41847.47
P3	1	511.3741	783.1527	524.7715	0	121.1486	251.9691	8429.438	2399.132	8247.826
	2	42209.85	41085.26	45143.53	121.1486	0	30.41695	786291.2	764216.8	786055.1
	3	789337.3	782580.9	800818.3	251.9691	30.41695	0	131905.5	122515.3	132254.7
P4	1	133332.8	130866.7	138411.9	122603.1	129288.9	132881.2	10055.19	7803.208	10366.15
	2	10477.08	10067.5	11876.2	7636.455	9616.915	10522.85	11967893	11876710	11973502
	3	11982050	11960716	12030317	11875500	11944997	11979421	3707.634	2307.142	3864.29
P5	1	41997.35	40879.52	44922.42	2232.343	3365.424	3970.367	41142.34	36046.25	41638.7
	2	147341.8	145704.6	152485.8	35803.16	37920.23	41993.02	145743.2	136552.5	146882.6
	3	870006.1	862873.7	881962.9	135942	15520.8	147480.8	866848.1	843806.7	866518.4
P6	1	101982.7	99766.36	106400.6	845036.6	89920.34	868058.8	100756.1	92584.59	100995.8
	2	9767.112	9362.691	11127.7	92741.85	9920.1	101542.7	9358.802	7177.327	9655.62
	3	6269907	6255194	6304792	7019.553	9890.3	9808.318	6259556	6194000	6264243
P7	1	24642.02	23407.94	26456.44	8429.438	786291.2	131905.5	0	187.4656	6.449449
	2	2246.908	1959.443	2950.077	2399.132	764216.8	122515.3	187.4656	0	214.0034
	3	127914.3	125499	132889.8	8247.826	786055.1	132254.7	6.449449	214.0034	0
P8	1	557.8148	460.0017	931.1615	1063.014	5361.83	2200.272	2062.228	1060.253	2121.339
	2	3779766	3766183	3806406	117411.1	39920.83	127471.7	126516.6	117324.7	126858.1
	3	960.3259	820.6155	1439.54	3722243	3766322	3777052	3772286	3721431	3773711
P9	1	42245.7	41120.7	45180.59	36031.5	39920.83	42238.87	41388.31	36272.53	41883.22
	2	613024.4	609695.7	623448.1	589532.5	600267.4	613323.4	609760	590801.8	612104.3
	3	457319.8	452244.2	466195.8	438749.5	30920.81	456019.6	454943.6	438003.8	454887
P10	1	181915	179212.7	187898.7	169077.9	166320.55	181534.6	180191.3	169125.2	180804.3
	2	10829.25	10400.4	12260.35	7921.228	9920.80	10871.5	10398.96	8085.793	10710.46
	3	3165109	3154430	3189931	3110321	3836723	3163947	3157789	3111163	3160919

Results also confirms the fitness of the extracted feature. We can see from the highlighted parts of Table 3.2 that the MSE values are much less than others.

3.4 Identification

The process of identification as described before is the process of selecting one match among many. As shown in the Figure 3.8 the query sample is compared with the database of known samples to find a match. ANN classifier was trained using 300 normalized feature per class. After training the system was tested upon 4946 testing samples. Each sample is 20 Sec long. Each sample goes through the process of preprocessing, feature extraction and then 15 feature vectors are extracted. These vectors are fed to the trained classifier. The 15 output score generated by the ANN are added together to create the decision vector. The maximum weight of the vector is taken as the identified class by the system. The process is shown in Table 3.3. Here a testing sample of class one is identified by adding the 15 NN output scores. The highlighted portion is the decision vector. Class one is having the maximum weight in this example.

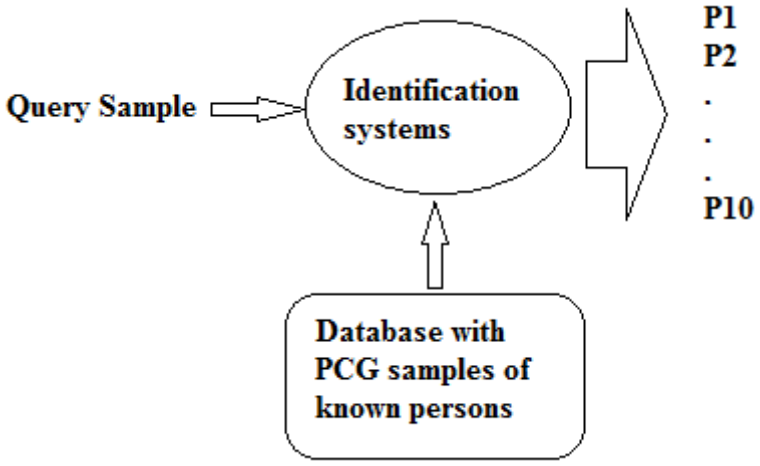

FIGURE 3.8: Block diagram of identification process.

TABLE 3.3: Process of Identification from Neural Network Output

-0.02663	0.981454	0.113668	0.574331	0.958242	0.486241	0.955438	-0.08817	0.615921	-0.04874	0.142927	0.9794	0.979398	0.022023	0.025331	6.670825
-0.00703	-0.04274	0.013481	0.073117	-0.17038	0.104555	0.232023	0.038937	0.011198	0.192994	0.048304	-0.03554	-0.0356	0.046662	0.046907	0.516883
0.054376	0.114716	0.017016	0.083482	0.00578	0.000386	-0.08472	-0.03645	0.049229	0.110458	0.042598	0.01831	0.018318	0.082215	0.082202	0.557921
0.053913	-0.10814	0.085746	0.097994	-0.26131	0.149235	0.997438	0.997709	-0.1503	0.101007	0.101183	-0.15635	-0.15638	0.089885	0.090631	1.932273
0.137382	-0.07288	0.071298	0.073494	-0.01547	0.013546	-0.06498	-0.08131	0.051063	-0.10549	0.052273	0.225751	0.225709	0.112875	0.11241	0.735667
0.068259	0.150234	0.135272	0.00774	0.99624	0.161082	-0.01571	-0.04924	0.518128	-0.06556	0.151382	-0.09712	-0.09715	0.040452	0.040025	1.944029
0.278616	0.094654	-0.02768	0.054798	-0.09454	-0.07667	0.071868	0.04369	-0.00457	-0.08659	0.001286	0.209691	0.209568	0.118181	0.117704	0.910015
0.087914	-0.55491	0.251095	-0.15884	0.189415	-0.07686	-0.68817	-0.20709	0.122875	0.663135	0.074173	0.110591	0.11096	0.265392	0.262448	0.452126
0.0239	0.07836	0.186951	0.08437	-0.36629	0.197033	-0.11094	0.052522	-0.02036	0.136953	0.222869	-0.26722	-0.26729	0.103132	0.10332	0.1573
0.293553	-0.00257	0.117387	0.049344	-0.13144	0.028144	0.055486	0.029577	-0.00064	-0.00919	0.126724	-0.10231	-0.10227	0.071001	0.071387	0.494184
NEURAL NETWORK OUTPUT FOR CLASS 1 SAMPLE TESTING															**SUM**

Table 3.4 shows the confusion matrix of the identification process. An identification accuracy of 96.178 is achieved. This result confirms that time frequency analysis of PCG signal generates useful feature set for biometric identificaion.

TABLE 3.4: Confusion Matrix

	P1	P2	P3	P4	P5	P6	P7	P7	P8	P9
P1	940	0	4	0	0	1	0	0	4	0
P2	0	249	13	0	0	0	0	0	1	3
P3	0	2	588	0	3	0	0	0	2	0
P4	0	1	11	236	0	8	0	0	3	0
P5	0	0	10	0	538	5	1	1	2	13
P6	24	0	0	0	3	660	0	0	0	0
P7	5	0	11	0	0	0	209	0	2	1
P8	10	15	0	0	6	4	0	235	0	0
P9	2	0	1	0	0	0	0	1	664	0
P10	4	0	0	0	6	0	0	1	5	438
Total sample tested: 4946 ***Correct Identification:*** 4757										
Accuracy: 96.178 %										

Comparison with LFBCC is given in 3.5. It shows our proposed system is giving better accuracy than LBFCC.

TABLE 3.5: Comparison with LBFCC

Feature	Classification	Accuracy
LBFCC	MLP ANN	92.5
Proposed Wavelet based feature	MLP ANN	96.178

3.5 Verification

The process verification is different as mentioned before. The process is shown in the Figure 3.9. There is a query sample which claims some known identity. The two samples are compared and a decision is taken if the two samples are a match or no match. That is the two belong to the same class or not.

The verification perfomence is determined from Equal Error Rate. The classification performence for verification for each class is presented in Figures 3.10 and 3.11 in terms of ROC plots. Receiver operating curve (ROC) is a graphical plot of sensitivity or true

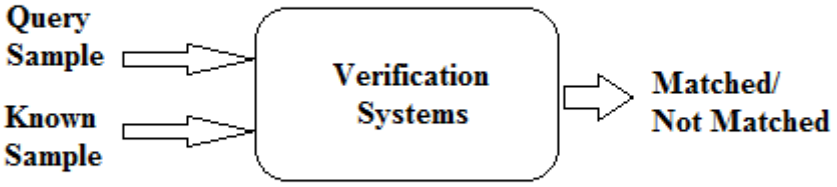

FIGURE 3.9: Process of Verification.

positive rate vs. false positive rate, as discriminative threshold is varied [29]. If we take threshold c, define binary test from continuous test result Y as, Positive if $Y \geqslant c$ and negative if $Y < c$. The corresponding true and false positive fractions at the threshold c be $TPF(c)$ and $FPF(c)$ respectively, where

$$TPF(c) = P[Y \geqslant c \mid D = 1] \tag{3.2}$$

$$FPF(c) = P[Y < c \mid D = 0] \tag{3.3}$$

The ROC curve is the entire set of possible true and false positive fraction attainable by dichotomizing Y with different thresholds. That is ROC curve is,

$$ROC(.) = \{(FPF(c), TPF(c)), c \in (-\infty, \infty)\} \tag{3.4}$$

The Equal Error Rate is given by:

$$EER = \frac{Falseacceptanceratio}{Falserejectratio} \times 100 \tag{3.5}$$

and

$$Falseacceptanceratio = 1 - Trueacceptanceratio \tag{3.6}$$

The performence of the ANN based system is given in Table 3.4

Case 1 is when the query sample is 10 heart cycle long and case 2 when it is 15 heart cycle long. We achived an EER of 17.98 which is a good EER considering the noisy nature of the testing samples.

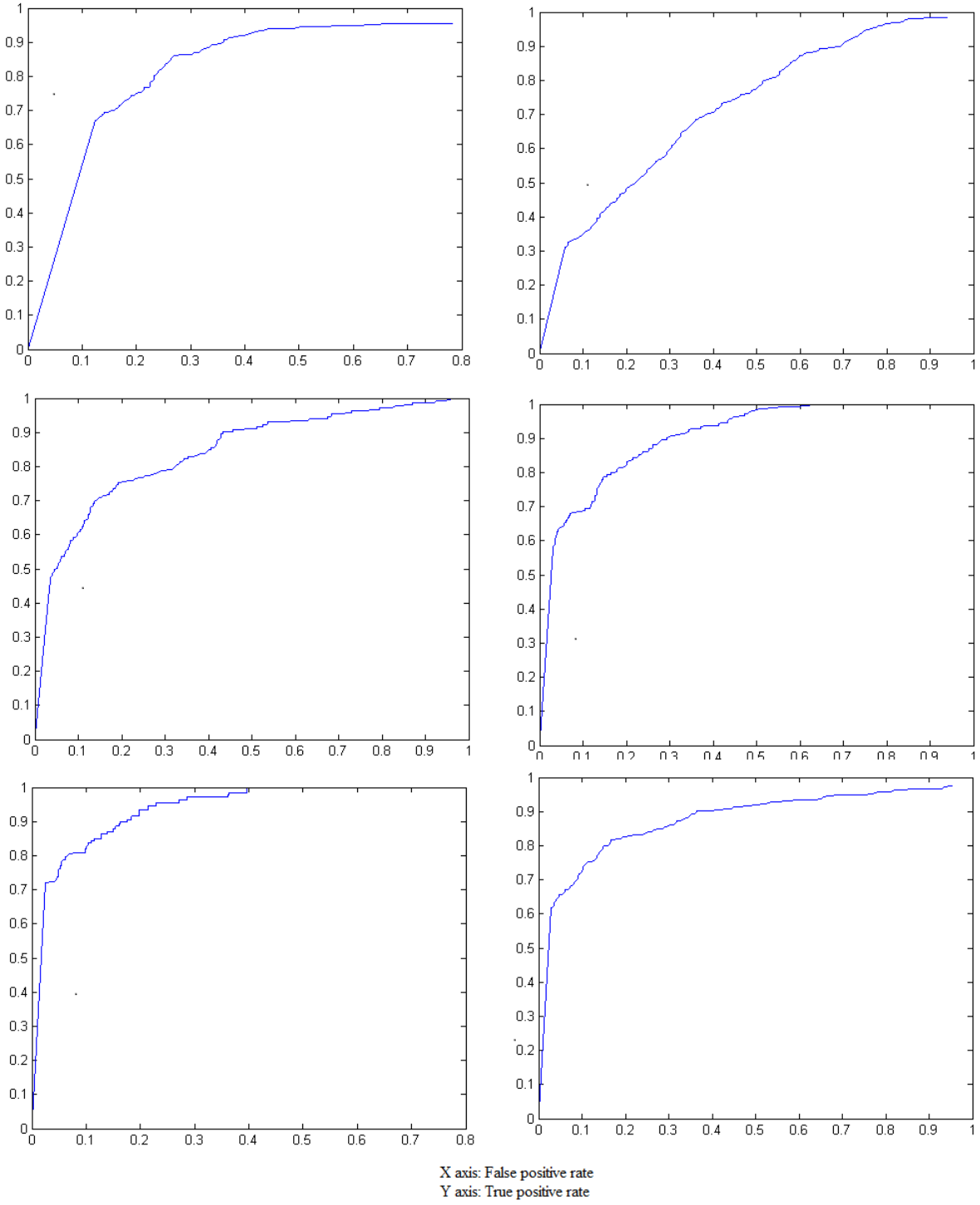

FIGURE 3.10: ROC1-6.

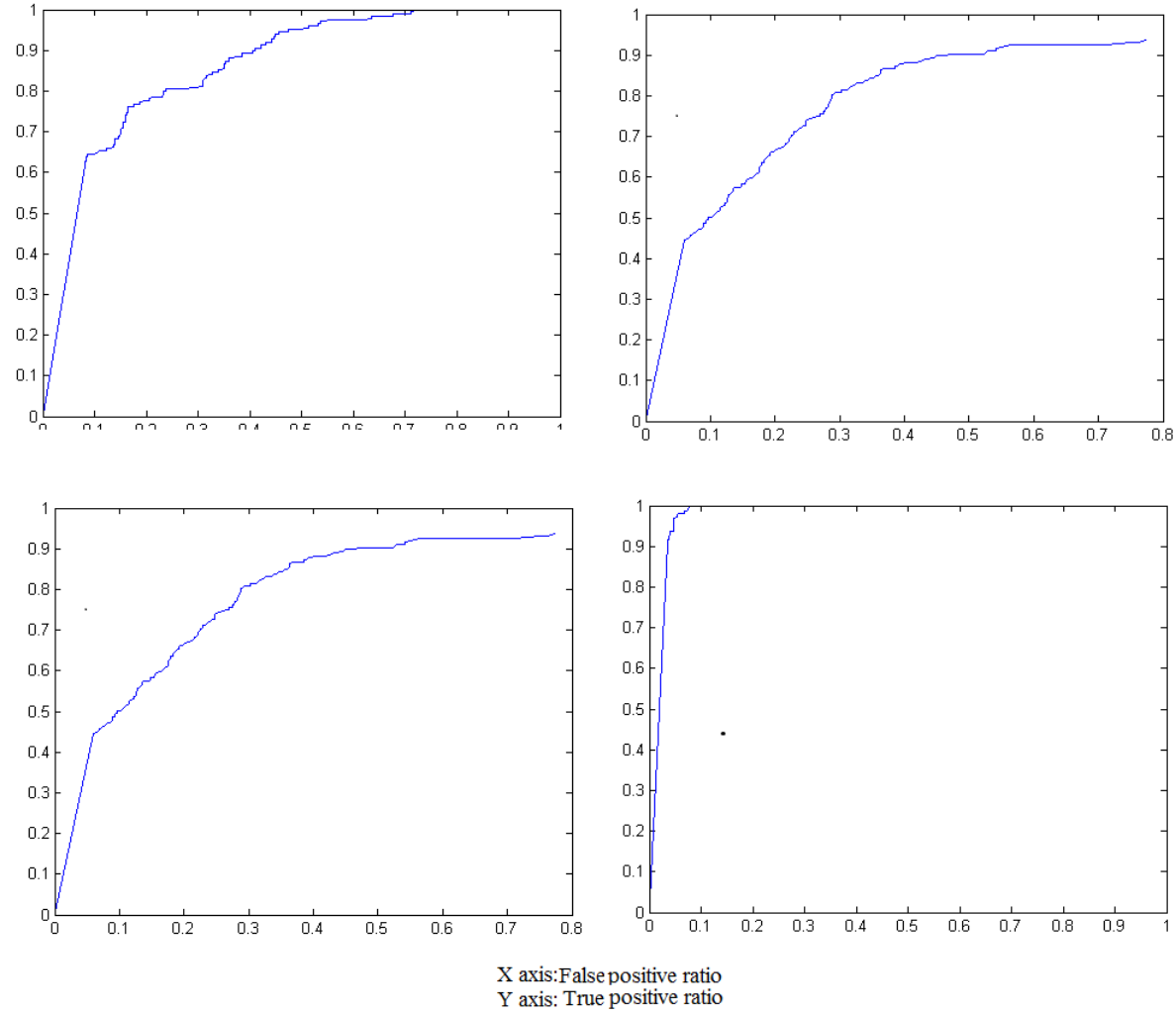

FIGURE 3.11: ROC6-10.

TABLE 3.6: Equal Error Rate obtained from the Identity Verification testing

Case	No. of feature vectors	EER
1	10	23.37
2	15	17.98

Chapter 4

Conclusion and Future Work

4.1 Conclusion

A novel technique has been proposed in this paper using wavelet based feature set for automatic identification system. The PCG signal is segmented into heart cycles, aligned, segmented into first and second heart sounds (S1 and S2) and decomposed upto second level using Daubechies second order (DB2). The feature is generated from second level coefficient energy. An imposter checking process based on MSE of Normalized mean cumulative sum of 10 features is also developed in the project. The technique is tested upon a collected heart sound data base using digital stethoscope. Comparison with Linear Frequency Cepstral Coefficient based feature set shows that the proposed method yields higher accuracy. The experimental result demonstrated that heart sound which is naturally made physiological parameter can be used in modern biometric systems.

4.2 Future works

The project has achieved some of the major objectives such as implementing a new time frequency domain feature set. It is giving satisfactory result. But some more research scopes are still left. These are

1. Trying other classifiers like a support vector machine, Gaussian Mixture Models and see their performance compared to BP MLP ANN.
2. A better segmentation technique can be applied. Because of windowing the segmentation process is not very accurate. But the problem is overcome by aligning process.
3. Integration with other biometric systems to include the benefits of other biometric

systems.

4. Test whether or not the diseased heart sounds can also be included from this system.

Bibliography

[1] Ruud Bolle and Sharath Pankanti. *Biometrics, Personal Identification in Networked Society: Personal Identification in Networked Society*. Kluwer Academic Publishers, Norwell, MA, USA, 1998. ISBN 0792383451.

[2] C. Vielhauer. *Biometric User Authentication for It Security: From Fundamentals to Handwriting*. Advances in Information Security. ISBN 9780387261942.

[3] Koksoon Phua, Jianfeng Chen, Tran Huy Dat, and Louis Shue. Heart sound as a biometric. *Pattern Recogn.*, 41(3):906–919, March 2008. ISSN 0031-3203.

[4] A. K. Abbas and R. Bassam. Phonocardiography signal processing.

[5] S. Lehrer. *Understanding Pediatric Heart Sounds*. ISBN 9780721696461.

[6] F. Beritelli and S. Serrano. Biometric identification based on frequency analysis of cardiac sounds. *Information Forensics and Security, IEEE Transactions on*, 2(3): 596 –604, sept. 2007.

[7] S.Z. Fatemian. *A wavelet-based approach to electrocardiogram (ECG) and phonocardiogram (PCG) subject recognition*. ISBN 9780494593646.

[8] F. Beritelli and A. Spadaccini. Human identity verification based on mel frequency analysis of digital heart sounds. In *Digital Signal Processing, 2009 16th International Conference on*, pages 1 –5, july 2009.

[9] F. Beritelli and A. Spadaccini. An improved biometric identification system based on heart sounds and gaussian mixture models. In *Biometric Measurements and Systems for Security and Medical Applications (BIOMS), 2010 IEEE Workshop on*, pages 31 –35, sept. 2010.

[10] F. Beritelli and A. Spadaccini. Heart sounds quality analysis for automatic cardiac biometry applications. In *Information Forensics and Security, 2009. WIFS 2009. First IEEE International Workshop on*, pages 61 –65, dec. 2009.

[11] D.A. Reynolds and R.C. Rose. Robust text-independent speaker identification using gaussian mixture speaker models. *Speech and Audio Processing, IEEE Transactions on*, 3(1):72 –83, jan 1995. ISSN 1063-6676. doi: 10.1109/89.365379.

[12] Chai Wutiwiwatchai, Sutat Sae-tang, and Chularat Tanprasert. Thai text-dependent speaker identification by ann with two different time normalization techniques.

[13] Y. Linde, A. Buzo, and R. Gray. An algorithm for vector quantizer design. *Communications, IEEE Transactions on*, 28(1):84 – 95, jan 1980. ISSN 0090-6778. doi: 10.1109/TCOM.1980.1094577.

[14] URL http://www.data-compression.com/vq.shtml#lbg.

[15] Stphane Mallat. *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*. Academic Press, 3rd edition, 2008. ISBN 0123743702, 9780123743701.

[16] Liu Chun-Lin. *A Tutorial of the Wavelet Transform*.

[17] Ben Krse, Ben Krose, Patrick van der Smagt, and Patrick Smagt. An introduction to neural networks, 1993.

[18] S.S. Haykin. *Neural networks: a comprehensive foundation*. ISBN 9780132733502.

[19] M.T. Hagan, H.B. Demuth, and M.H. Beale. *Neural Network Design*. Electrical Engineering Series. ISBN 9780534943325.

[20] E. Gelenbe. *Neural networks: advances and applications*. ISBN 9780444885333.

[21] Sutat Sae-Tang and C. Tanprasert. Feature windowing-based thai text-dependent speaker identification using mlp with backpropagation algorithm. In *Circuits and Systems, 2000. Proceedings. ISCAS 2000 Geneva. The 2000 IEEE International Symposium on*, volume 3, pages 579 –582 vol.3, 2000. doi: 10.1109/ISCAS.2000.856126.

[22] S.L. Strunic, F. Rios-Gutierrez, R. Alba-Flores, G. Nordehn, and S. Burns. Detection and classification of cardiac murmurs using segmentation techniques and artificial neural networks. In *Computational Intelligence and Data Mining, 2007. CIDM 2007. IEEE Symposium on*, pages 397 –404, 1 2007-april 5 2007. doi: 10.1109/CIDM.2007.368902.

[23] J. Vepa. Classification of heart murmurs using cepstral features and support vector machines. In *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, pages 2539 –2542, sept. 2009.

[24] Hong Tang, Ting Li, and Tianshuang Qiu. Cardiac cycle detection for heart sound signal based on instantaneous cycle frequency. In *Biomedical Engineering and Informatics (BMEI), 2011 4th International Conference on*, volume 2, pages 676 –679, oct. 2011. doi: 10.1109/BMEI.2011.6098371.

[25] S.A. Aase, S.R. Snare, O.C. Mjlstad, H. Dalen, F. Orderud, and H. Torp. Qrs detection and cardiac cycle separation without ecg. In *Ultrasonics Symposium (IUS), 2009 IEEE International*, pages 1399 –1402, sept. 2009. doi: 10.1109/ULTSYM.2009.5441469.

[26] H. Sava and L.-G. Durand. Automatic detection of cardiac cycle based on an adaptive time-frequency analysis of the phonocardiogram. In *Engineering in Medicine and Biology Society, 1997. Proceedings of the 19th Annual International Conference of the IEEE*, volume 3, pages 1316 –1319 vol.3, oct-2 nov 1997. doi: 10.1109/IEMBS.1997.756618.

[27] Jayaprakash. Sahoo, Manab.Kumar Das, Samit. Ari, and S. Behera. Autocorrelation and hilbert transform based qrs complex detection in ecg signal. *International Journal of Signal and Imaging Systems Engineering*, In Press.

[28] S. Ari J.P. Sahoo and S. Behera. A novel technique for qrs complex detection in ecg signal based on hilbert transform and autocorrelation. *International Conference on Electronic Systems, NIT, Rourkela,*, 36(5):961 –1005, Jan 7-9 2011.

[29] Thomas A. Lasko, Jui G. Bhagwat, Kelly H. Zou, and Lucila Ohno-Machado. The use of receiver operating characteristic curves in biomedical informatics. *J. of Biomedical Informatics*, 38(5):404–415, October 2005. ISSN 1532-0464.