BIOMETRIC IDENTIFICATION USING PHONOCARDIOGRAM

A THESIS SUBMITTED FOR PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

BACHELOR OF TECHNOLOGY
IN
ELECTRONICS AND COMMUNICATION ENGINEERING

BY
AJAY SINGH
ROLL. NO. 107EC009
&
ASHISH KUMAR SINGH
ROLL. NO. 107EC021

UNDER THE GUIDANCE OF

Dr. SAMIT ARI
ASSISTANT PROFESSOR

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING
NATIONAL INSTITUTE OF TECHNOLOGY ROURKELA
ORISSA – 769008
INDIA

National Institute of Technology, Rourkela

CERTIFICATE

This is to certify that the thesis entitled "**Biometric Identification using Phonocardiogram**" submitted by **Mr. Ajay Singh** and **Mr. Ashish Kumar Singh** for partial fulfilment for the requirement of Bachelor in Technology degree in **Electronics and Communication Engineering** at National Institute of Technology, Rourkela is an authentic piece of work carried out by them under my guidance and supervision.

To the best of my knowledge, the matter in this thesis has not been submitted to any other University / Institute for the award of any Degree.

Date: 13th May, 2011

Dr. Samit Ari
Assistant Professor
Department of Electronics and Communication
National Institute of Technology, Rourkela
Orissa -769008
INDIA

ACKNOWLEDGEMENT

We take this opportunity to thank all the individuals whose guidance, help and timely support made us to complete the project within the stipulated time. We deeply honour and value their grateness for their help and encouragement.

First and foremost we would like to express our sincere gratitude to our project supervisor Prof. Samit Ari for his invaluable support, motivation, guidance and constant encouragement which made this project a rich and insightful experience. His timely inputs, valuable feedbacks and constructive criticism in many stages of this work were instrumental in understanding and execution of this project.

We also express our gratitude towards Prof. Sarat Kumar Patra, HOD, Department of Electronics and Communication, for his selfless support and help offered whenever needed, especially for the usage of laboratories. Also, our deep gratitute towards the M.Tech students and Reasearch scholars who when asked provided with lots of project materials selflessly and patiently helped us in database recordings. We specially thank Mr. Bhubaneswar Das of Basic Electronics lab for his help and support.

We would also like to thank all professors and lecturers, and members of the department for their generous help in various ways for the completion of this thesis. We extend our thanks to our dear friends for their support and cooperation.

Date: 13th May 2011
Place: NIT Rourkela

Ajay Singh
Deptartment of ECE
NIT Rourkela

Ashish Kumar Singh
Deptartment of ECE
NIT Rourkela

CONTENTS

ABSTRACT

Phonocardiogram (PCG) signals as a biometric is a new and novel method for user identification. Use of PCG signals for user recognition is a highly reliable method because heart sounds are produced by internal organs and cannot be forged easily as compared to other recognition systems such as fingerprint, iris, DNA etc. PCG signals have been recorded using an electronic stethoscope. Database of heart sound is made using the electronic stethoscope. In the beginning, heart sounds for different classes is observed in time as well as frequency for their uniqueness for each class. The first step performed is to extract features from the recorded heart signals. We have implemented LFBC algorithm as a feature extraction algorithm to get the cepstral component of heart sound. The next objective is to classify these feature vectors to recognize a person. A classification algorithm is first trained using a training sequence for each user to generate unique features for each user. During the testing period, the classifier uses the stored training attributes for each user and uses them to match or identify the testing sequence. We have used LBG-VQ and GMM for the classification of user classes. Both the algorithms are iterative, robust and well established methods for user identification. We have implemented the normalization at two places; first, before feature extraction; then just after the feature extraction in case of GMM classifier which is not proposed in earlier literature.

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
Introduction

1.1. Biometric Systems

In recent years, it has become very important to identify a user in applications such as personnel security, defence, finance, airport, hospital and many other important areas [1]. So, it has become mandatory to use a reliable and robust authentication and identification system to identify a user. Earlier the methods for user identification were mainly knowledge-based such as user password or possession-based such as a user key; but due to vulnerability of these methods it was easy for people to forge the information. Hence, the performance-based biometric systems for identification, where a user is recognized using his own biometrics. Biometrics uses the methods for recognizing users based upon one or more physical and behavioural traits. Hence, conventional biometric identification systems such as iris, fingerprint, face and speech have become popular for user identification and verification. However, all these identification methods have weaknesses that they can be forged as shown in Table 1.1 [2] [3] [21].

1.2. PCG Signals as a Biometric

In this project, we use the heart sounds (PCG signals) as a biometric for user identification. Use of phonocardiogram signals has many advantages over other biometrics based on the following properties of heart sounds [4]:

1. *Universal:* Every living human being has a pumping heart.
2. *Measurable:* PCG signals can be recorded using an electronic stethoscope

3. *Vulnerability:* Heart sounds cannot be copied or reproduced easily as it is based on intrinsic signals acquired from the body. Heart sounds cannot be taken without the consent of the person. Moreover, to reproduce the heart sounds, an anatomy of heart as well its surroundings has to be created as heart sounds depends on the anatomy of the body.
4. *Uniqueness:* Heart sounds depend on the physical state of an individual's health, age, size, weight, height, structure of the heart as well as the genetic factors. The heart sounds of two persons having the same type of heart diseases also vary.
5. *Simplicity:* Moreover, heart sounds are easy to obtain, by placing a stethoscope on the chest.

Table 1.1 Drawbacks of various biometric systems.

Identification trait	**weaknesses**
DNA	Easy to steal a piece of DNA
Speech	Speech can be recorded and played
Signature	Can be reproduced easily
Fingerprint	Can be recreated in latex using an object touched by the person
Face and iris	Can be recorded by a camera

Human heart sounds are natural signals, which have been applied in the doctor's auscultation for health monitoring and diagnosis for thousands of years. In the past, study of heart sounds focus mainly on the heart rate variability [5]. However, we conjecture that since the heart sounds also contain information about an individual's physiology, such signals have the potential to provide a unique identity for each person. Like ECG, these signals are

difficult to disguise and therefore reduces falsification. Moreover, heart sounds are relatively easy to obtain, by placing a conventional stethoscope on the chest.

1.3 Mechanism for heart sound production

The human heart has four chambers, two upper chambers called the atria and two lower chambers called ventricles, as shown in Fig. 1.1. There are valves located between the atria and ventricles, and between the ventricles and the major arteries from the heart [6]. These valves close and open periodically to permit blood flow in only one direction.

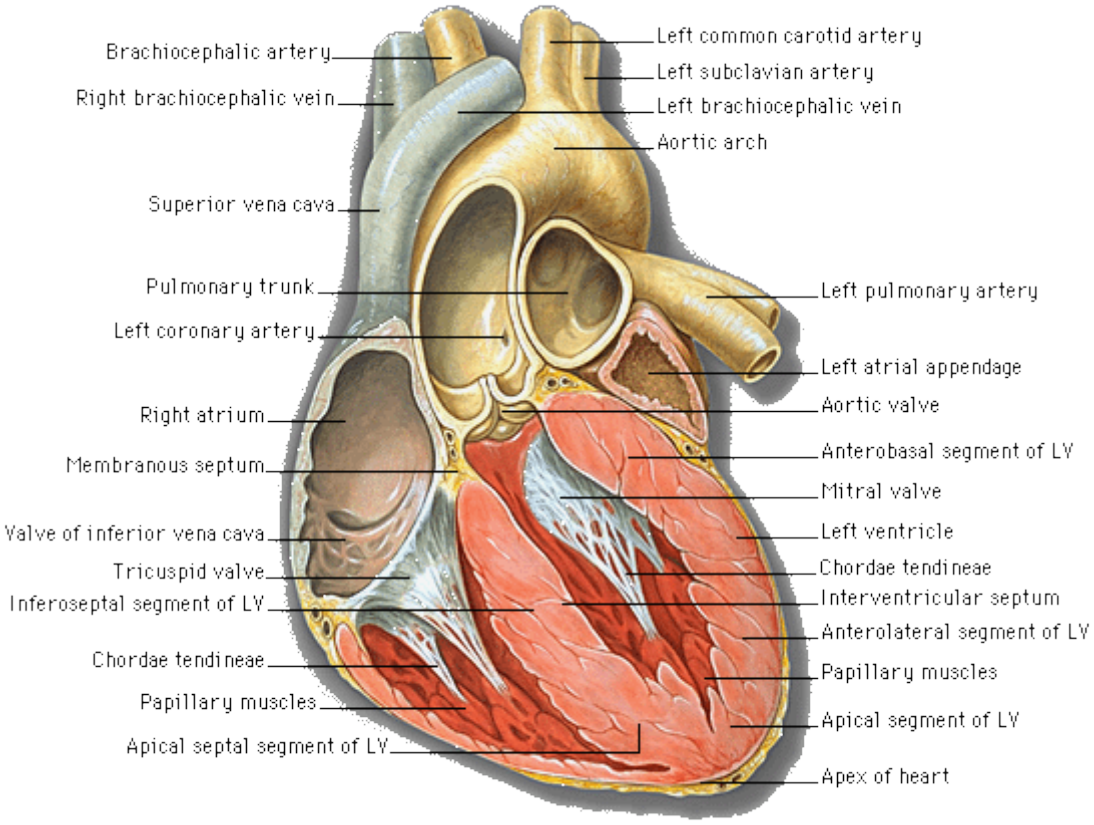

Fig. 1.1: Cross-section of a typical human heart (source: http://images.google.com).

Two sounds are normally produced as blood flows through the heart valves during each cardiac cycle as shown in Fig. 1.2. The first heart sound S1, is a low, slightly prolonged

"lubb", caused by vibrations set up by the sudden closure of the mitral and tricuspid valves as the ventricles contract and pump blood into the aorta and pulmonary artery at the start of the ventricular systole. The second sound S2 is a shorter, high-pitched "dupp", caused when the ventricles stop ejecting, relax and allow the aortic and pulmonary valves to close just after the

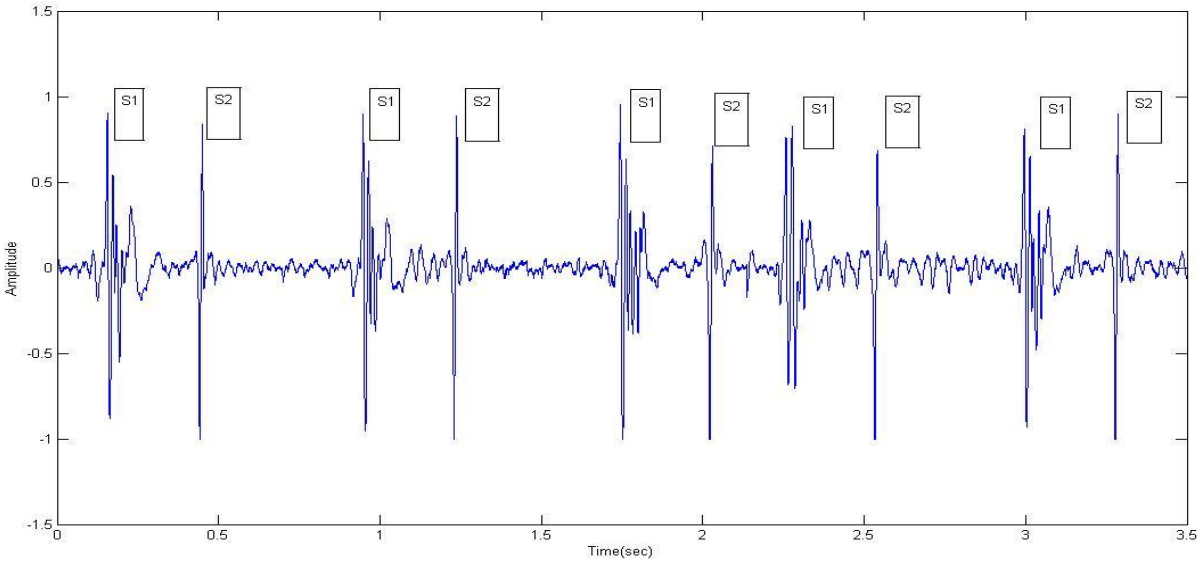

Fig. 1.2: waveform of first heart sound S1 and second heart sound S2.

end of the ventricular systole. They are the "lubb-dupp" sounds that are thought of as the heartbeat. S1 has duration of about 0.15 s and a frequency of 25–45 Hz. On the other hand, S2 lasts about 0.12 s, with a frequency of 50 Hz.

The sounds associated with the opening and closing of different values can be heard by placing a microphone directly on the various auscultation points on the chest wall as shown in Fig. 1.3 [7]. The time domain and frequency domain characteristic of heart sounds recorded from 3 persons are shown in Fig. 1.4 and Fig. 1.5 respectively. From the time and frequency domain plots, it is clear that the time and frequency characteristics are different for different persons. These observations suggest that the heart sounds are distinct for different

persons and gives us the motivation that phonocardiogram signals can be used for biometric identification of a person.

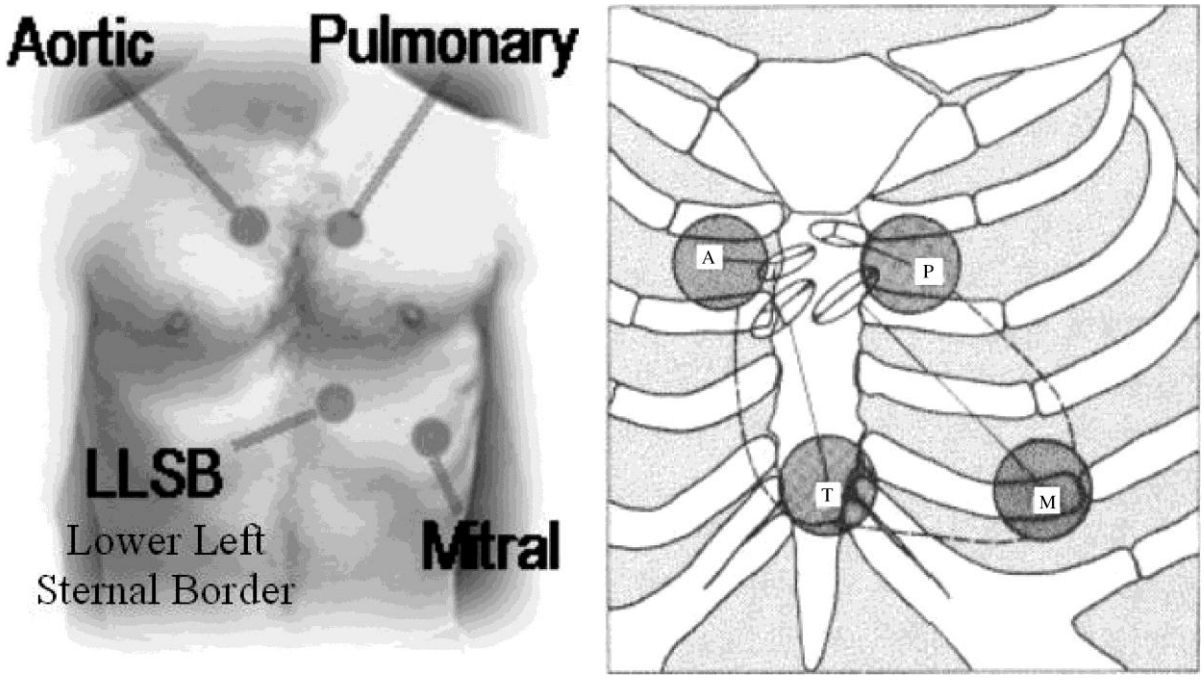

Fig. 1.3: Four sites to place stethoscope.

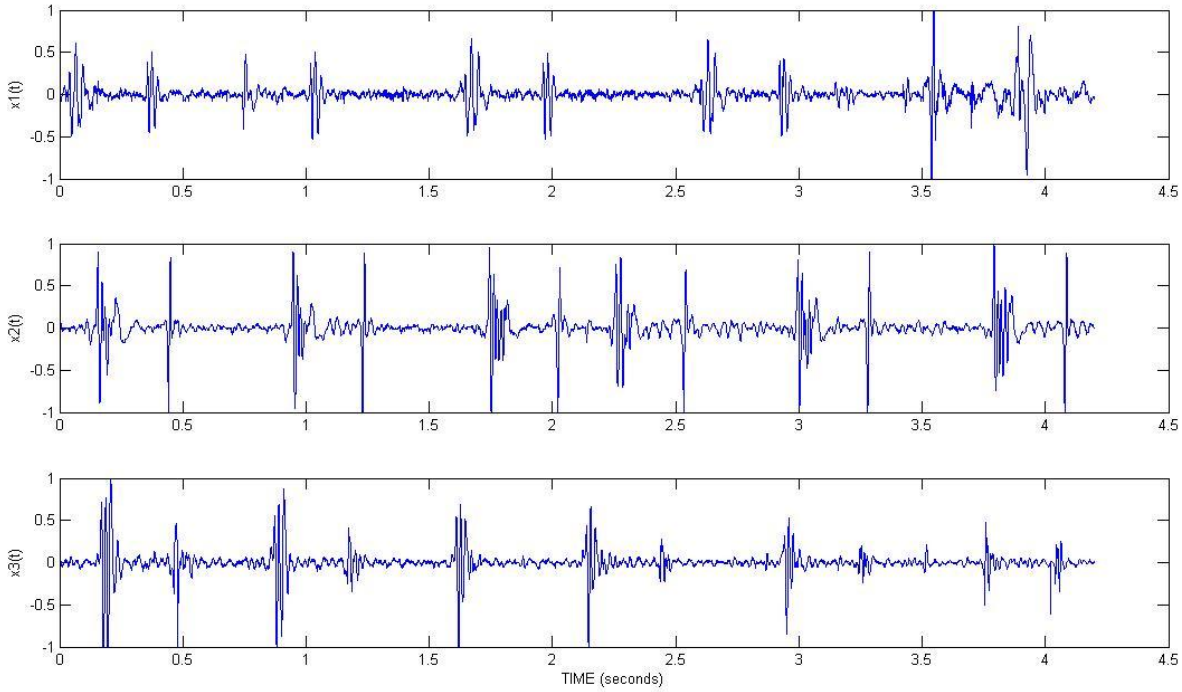

Fig. 1.4: Heart sound waveform comparison in time domain.

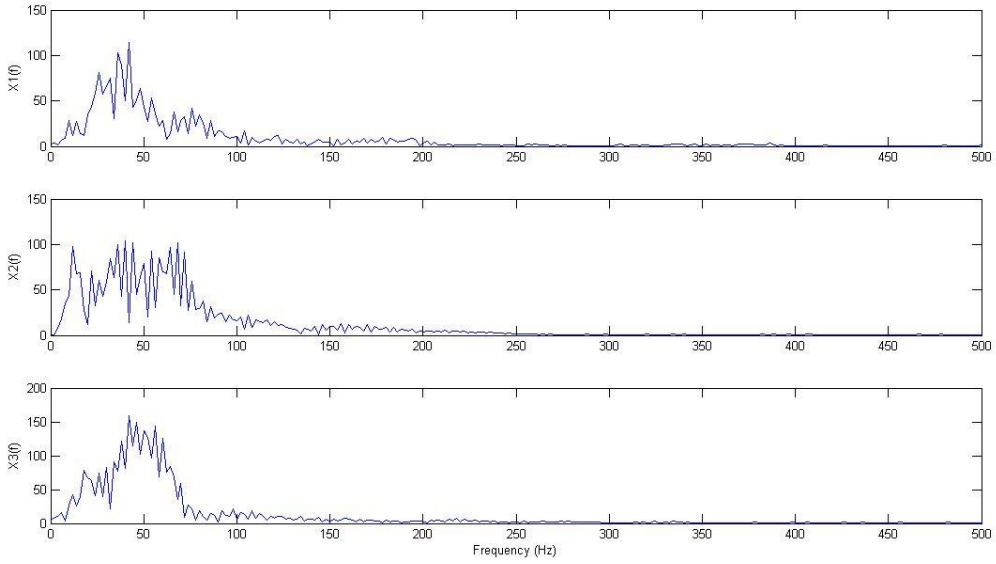

Fig. 1.5: Heart sound waveform comparison in frequency domain

1.4 OBJECTIVES

The main objective of this project is to use different classification schemes to recognize a person using phonocardiogram signals as a biometric. The first task is to record the PCG signals using an electronic stethoscope for making the database and transfer the recorded signals to a computer for further processing of the signal. The next task is to extract the features from the recorded signals and use the extracted features for classification. We propose the use of normalization of the signals at two places: first, before the feature extraction and second, after the feature extraction.

CHAPTER 2
Data Acquisition

2.1 Hardware

The electronic stethoscope used for recoding PCG signals is HD Fono / HD FonoDoc manufactured by HD Medical Services (India) Private Limited. This device allows us to adjust audio volume. It also has a visual display for observing heart sounds that represent valvular functions of the heart in real time, called Phonocardiogram (PCG). The stethoscope head is put on the user's chest for recording. It also has a USB interface with the computer for data download, review and storage. This device allows us the storage of 10 seconds signals that can be downloaded to PC/Laptop.

2.2 Software

The software used for transferring the data and storing the signals in PC is HD fono Recording and Playback Software V3.4. This software has the ability to acquire/retrieve heart sounds displayed by HD Fono device and save them along with the patient's information on the computer. It also displays waveforms for quick and correct identification of murmurs and gallops. It Visualization tools include zoom and callipers. It can be used to transfer the stored waveforms stored in device to PC for analysing and generating user's reports.

2.3 Processing Software

The PCG signals stored in the PC are processed using sound editing software AUDACITY and MATLAB. Programming for the feature extraction and classification is done in MATLAB software.

2.4 Database

Database of 8 people was made using the hardware and softwares described above. For each one of the best locations out of four locations was selected for stethoscope placement on the chest. For each user 10 samples were collected each of 1 minute duration. Out of 10 samples for each user 4 were used for training the algorithms described in the subsequent sections. Remaining 6 samples were used for matching purpose/testing the algorithms.

CHAPTER 3
Feature extraction

Feature extraction finds a transformation that converts the original heart signal into a feature space preserves the information required for the application and enables meaningful comparisons. Due to the overlap of the heart sound components and the noises and disturbances caused by other internal organs such as lungs, analysis of the heart sound in the time domain is not possible. For the biometric application, the physiological properties of the heart sounds are more important than the heart rate. Hence, we will process the heart sounds in the frequency domain.

The algorithm used for feature extraction is Linear Frequency Bands Cepstra (LFBC) [8]. It has been demonstrated that parameterization in the cepstral domain gives good discrimination and various manipulations can be performed [8]. As proposed in Ref. [8], the effect of inserting a transmission channel on the input heart sound is to multiply the heart sound spectrum by the channel transfer function. In the log cepstral domain, this multiplication becomes addition which can be removed by subtracting the cepstral mean from all input vectors. The mean is estimated over a limited amount of heart sound data so the subtraction will not be perfect. This technique is very effective because it compensates for long-term spectral effects such as those caused by different stethoscopes. As a result, we use the cepstral coefficient as the feature. The heart sound can be modelled as the outcome of a time-varying linear system as shown in Fig. 3.1, where the excitation input carries information of the heart signal, and the transfer function is time varying. The block diagram of the feature extraction module is shown in Fig. 3.2.

Fig. 3.1: Linear model of heart sound.

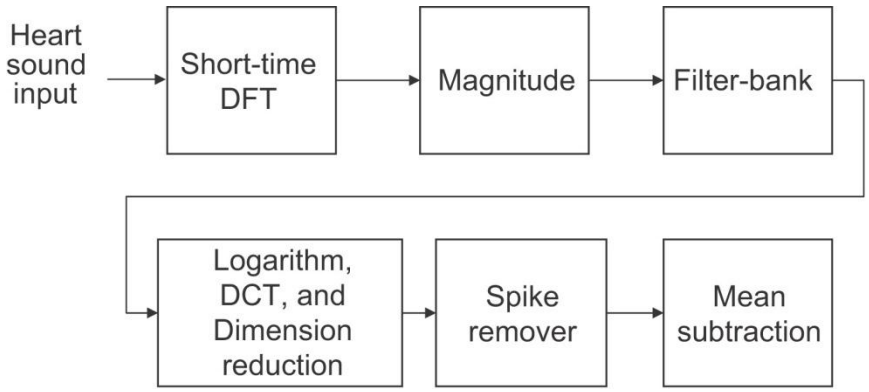

Fig. 3.2: Block diagram of feature extraction process.

3.2 Normalization

Normalization [20] is used to make all the signals to be restricted to the same range. In this technique, we divide the whole signal vector from the element of the signal whose absolute value is maximum. Thus, we limit the range of all the signal vectors to [-1, 1].

3.3 Short-time discrete Fourier Transform (STDFT):

Heart sounds are pseudo-random signals and hence STDFT [17] [18] is used to find the frequency components. The STDFT signal is given by:

$$X[n,k] = \sum_{m=0}^{N-1} x * w[m + (n-1)S] \left(-j\frac{2\pi}{N} km\right)$$

Where,

n is the frame index,

k is the frequency index,

N is the frame length,

S is the frame shift and,

w denotes the window.

Unlike speech signals, where the speech signals changes after each 20–25ms, heart sounds are more stationary and therefore the window length should be larger. The optimal window length was found to be about 500ms. For GMM as the classification method, independence between samples is required to estimate the probability distributions. Consequently, the non-overlap windowing is the most optimal. We use only the information magnitude of the STDFT signal and ignore the phase components because phase components are sensitive to the noise

3.4 Filter Bank

In this block, we pass passes the signal spectrum through a filter-bank [8]. Filter-banks are used because the sound spectrum has some special shapes and are distributed by a non-linear scale in frequency domain. Using the filter-banks with spectral characteristics which are well matched to those of the desired signal, the contribution of noise components in the frequency domain can be reduced. Mel-frequency filter-banks are best in the speech recognition and speaker identification. However, the heart sound spectrum has the range of 20–150 Hz. Full resolution is required in such a narrow band to capture more information of the signal spectrum. Hence, we simply filter out the frequency bins outside the range of 20–150 Hz.

3.5 Dimension Compression

This block is very similar to the standard MFCC feature extraction used in speaker recognition [8]. The spectral magnitude is compressed in the logarithmic domain, followed by the Discrete Cosine Transform (DCT) [18] which gives the cepstral components. The cepstral component c[n,k] can be written as

$$c[n,k] = \sum_{m=0}^{K-1} log(|X[n,m]|)\, cos\left(\frac{km\pi}{K}\right)$$

k= (1, 2… K),

Where, K is the number of bins in the frequency band of 20–150 Hz.

The first 24 coefficients for each frame are selected for dimension compression. The higher coefficients are less informative. To distinguish heart-sound's feature from the standard MFCC, this feature set is called the Linear Frequency Bands Cepstra (LFBC) [8].

3.6 Spike Removal

There is always interference in the heart sounds caused because of the movement of the stethoscope. Conventional filtering technique is ineffective because the spectra of these interferences and heart sounds overlap. We set an energy threshold to remove the high energy segments that contains the burst [8].

$$10\ log(E[n]) - min_n(10\ log(E[n]) \geq \mu,$$

Where, n is the segment index and,

μ=15dB is an energy threshold.

3.7 Cepstral Means Subtraction

There is always a fluctuation on the "relative transfer function because the positions of stethoscope cannot be fixed at all times. This acts as a "relative transfer function" as characterized by the propagation of heart sounds to the recorder. We apply the cepstral mean subtraction to remove this effect. In the frequency domain the output is equal to the multiplication of the signal measured in a fixed position, *X[n,k]*, and the relative transfer function in the frequency domain, *H[k]*. This can be expressed in the logarithmic domain by the equation [8],

$$log(|Y[n,k]|)=log(|X[n,k]|)+log(|H[k]|)$$

the cepstra of the recorded signals can be represented as follows:

$$c_Y[n,k] = c_X[n,k] + c_H[n,k]$$

$c_H[n,k]$ can be removed by taking the long term averaging in each dimension k:

$$c_Y[n,k] - \hat{c}_Y[n,k] = c_X[n,k] - \hat{c}_X[n,k]$$

3.8 Simulation Results

Frequency components for each user are obtained using STDFT. The signals after STDFT for three users are shown in Fig. 3.3. It is demonstrated that the frequency spectrum for each user is different.

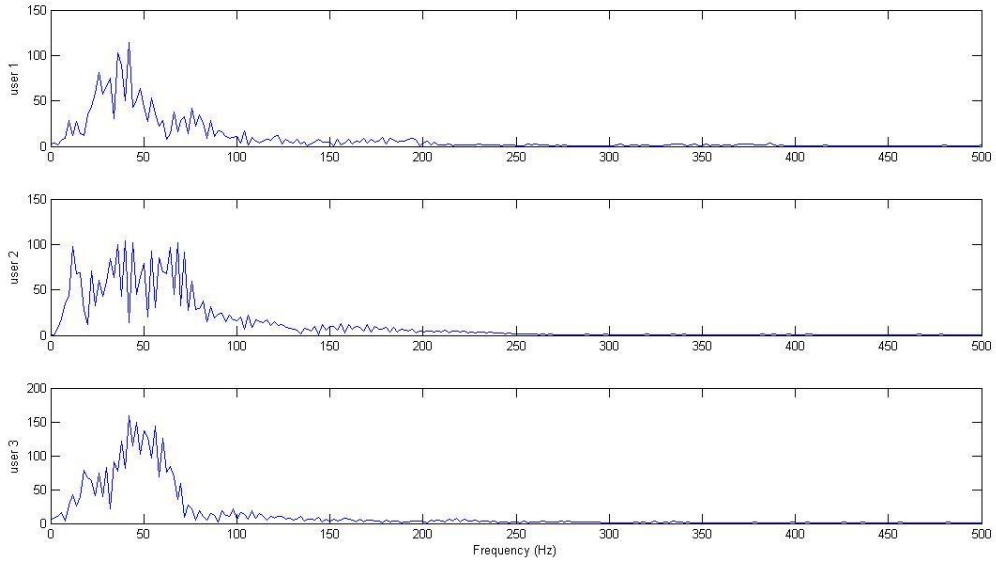

Fig. 3.3: Signals after STDFT operation

It is observed that the higher frequency components don't give much information and the information is contained within the frequency range 20-150 Hz. Hence, the frequency components outside this range are discarded using filter bank. The filtered output for three users is shown in Fig. 3.4.

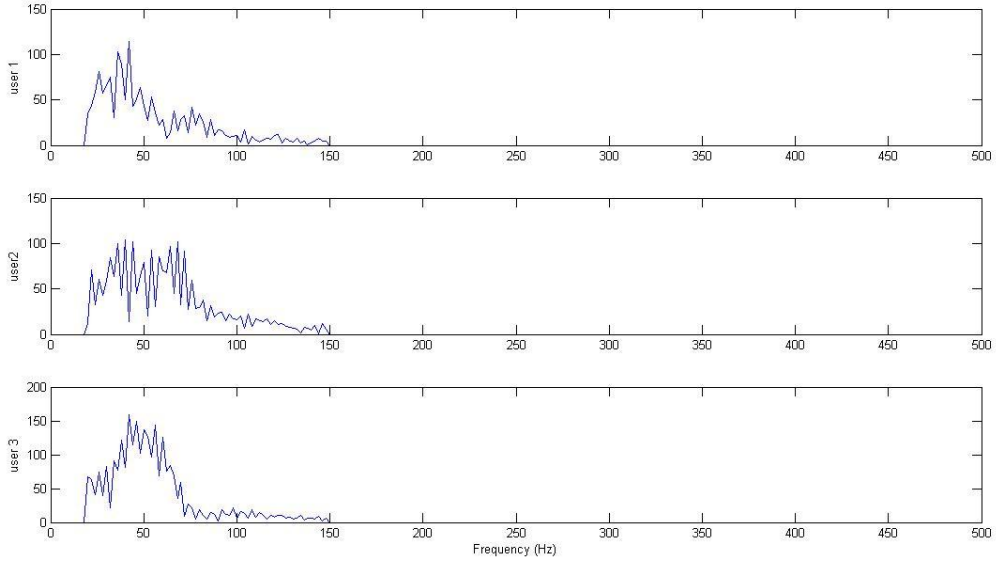

Fig. 3.4: Signals after filter operation

Finally, the cepstral components are obtained for each user by the LFBC algorithm described in the previous section. It was observed that the higher cepstral components don't have much signal information. Hence only first 24 cepstral coefficients are used and rest are discarded. The cepstra for three users are shown in Fig. 3.5.

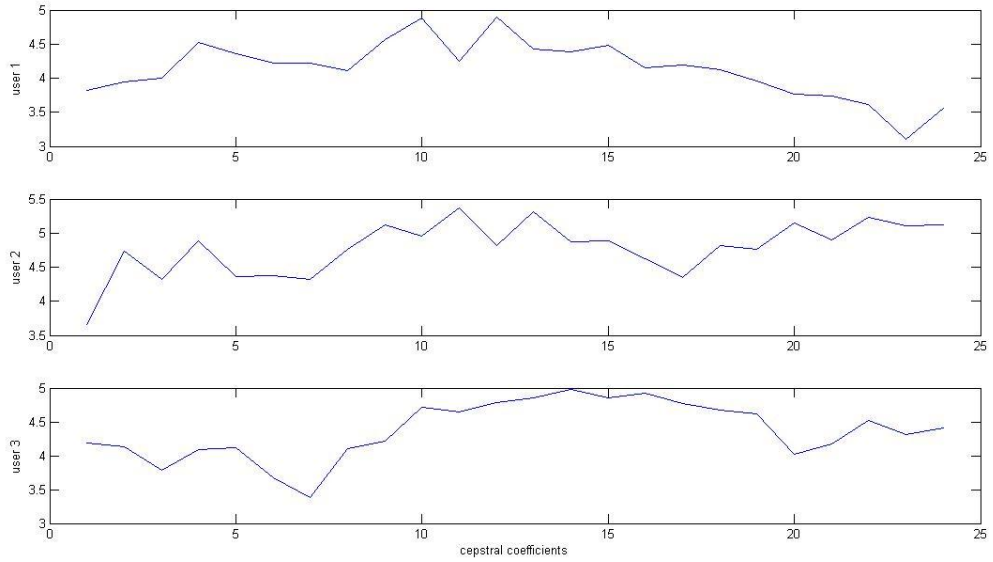

Fig. 3.5: Cepstrum for three users

CHAPTER 4
Classification methods

Vector quantization (VQ) and GMM are well known conventional and successful methods for the speaker recognition approaches [9] and [10]. So, these methods can be applied to the phonocardiogram signal recognition.

4.1 Vector Quantization

Vector quantization (VQ) [19] is a lossy-data-compression method based on the principle of block coding [11]. It is a fixed-to-fixed length algorithm. In the earlier days, the design of a vector quantization (VQ) is considered to be a challenging problem due to the need for multi-dimensional integration. In 1980, Linde, Buzo, and Gray (LBG) proposed a VQ design algorithm [11] based on a training sequence. The use of a training sequence bypasses the need for multi-dimensional integration. VQ that is designed using this algorithm are referred to in the literature as an LBG-VQ.

4.1.1 Design Problem

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 result in the smallest average distortion.

We assume that there is a training sequence consisting of *M* source vectors [11]:

$$T = \{x_1, x_2, \dots x_M\}$$

The training sequence is the feature vectors that we get after the feature extraction.

The dimension of each vector is *k*,

$$x_m = (x_{m,1}, x_{m,2}, \dots, x_{m,k}), \qquad m=1, 2,\dots, M$$

Let N be the number of code vectors and let

$$C = \{c_1, c_2, \dots c_N\},$$

represents the codebook. Each code vector is *k*-dimensional,

$$x_m = (c_{m,1}, c_{m,2}, \dots, c_{m,k}), \qquad m=1, 2,\dots,N$$

Let S_n be the encoding region associated with code vector c_n and let

$$P = \{S_1, S_2, \dots S_N\},$$

denote the partition of the space. If the source vector x_m is in the encoding region S_n, then its approximation (denoted by $Q(x_m)$ is c_n:

$$Q(x_m) = c_n, \qquad \text{if } x_m \varepsilon\ S_N$$

Assuming a squared error distortion measure [20], the average distortion is given by:

$$D_{ave} = \frac{1}{Mk} \sum_{m=1}^{M} \|x_m - Q(x_m)\|^2$$

The design problem can be stated as follows [9]: Given *T* and *N,* find *C* and *P* such that D_{ave} is minimized.

4.1.2 Optimality criteria

If *C* and *P* are a solution to the above minimization problem, then it must satisfy the following two criteria [11]:

- *Nearest Neighbor Condition*: This condition says that the encoding region S_n should consists of all vectors that are closer to c_n, than any of the other code-vectors. For those vectors lying on the boundary (blue lines), any tie-breaking procedure will do.

$$S_n = \{x: \|x - c_n\|^2 \leq x: \|x - c_n'\|^2 \text{ for all } n' = 1,2,\dots, N\}$$

- *Centroid Condition:* This condition says that the code vector c_n,should be average of all those training vectors that are in encoding region S_n. In implementation, one should ensure that at least one training vector belongs to each encoding region (so that the denominator in the above equation is never 0).

$$c_n = \frac{\sum_{x_m \varepsilon Sn} x_m}{\sum_{x_m \varepsilon Sn} 1} \qquad n=1, 2,\ldots,N$$

4.1.3 LBG Design Algorithm

The LBG VQ design algorithm [5], [9] is an iterative algorithm which alternatively solves the above two optimality criteria. The algorithm requires an initial codebook. This initial codebook is obtained by the splitting method. In this method, an initial code vector is set as the average of the entire training sequence. This code vector is then split into two. The iterative algorithm is run with these two vectors as the initial codebook. The final two code vectors are split into four and the process is repeated until the desired number of code vectors is obtained [9].

1. Given *T*. Fixed ε>0 to be a small number.
2. Let *N*=1 and

$$c_1^* = \frac{1}{M}\sum_{m=1}^{M} x_m$$

Calculate

$$D_{ave}^* = \frac{1}{Mk}\sum_{m=1}^{M}\|x_m - c_1^*\|^2$$

3. Splitting**:** For *i*=1,2,… , *N*, set

$$c_i^{(0)}=(1+\varepsilon)c_1^*$$

$$c_i^{(0)}=(1+\varepsilon)c_1^*$$

Set *N*=2*N*.

4. Iteration: Let $D_{ave}^{(0)}=D_{ave}^{*}$. Set the iteration index i=0.

    a. For *m*=1,2,…,M, find the minimum value of

    $$\|x_m - c_n^i\|^2 ,$$

    over all *n*=1,2,… ,*N*. Let n* be the index which achieves the minimum. Set

    $$Q(x_m) = c_n^{(i)},$$

    b. For *n*=1,2,… ,*N*, update the code vector

    $$c_n = \frac{\sum_{Q(x_m)=c_n^{(i)}} x_m}{\sum_{Q(x_m)=c_n^{(i)}} 1}$$

    c. Set,*i*=*i*+1

    d. Calculate

    $$D_{ave}^{(i)} = \frac{1}{Mk} \sum_{m=1}^{M} \|x_m - Q(xm)\|^2$$

    e. If ($D_{ave}^{(i-1)} = D_{ave}^{(i)}$)/ $D_{ave}^{(i-1)}$ >ε, go back to Step (i).

    f. Set $D_{ave}^{*} = D_{ave}^{(i)}$ For n-1,2,… ,N, set

    $$c_n^{*} = c_n^{(i)}$$

    as the final code vectors.

5. Repeat Steps 3 and 4 until the desired number of code vectors are obtained.

4.1.4 Matching Algorithm

- For testing, the extracted feature vectors are mapped to the quantized space obtained during training.
- Distortion is calculated by calculating the average distance of the testing feature vector from the code vectors of each user generated during training.
- The codebook, for which the distortion is minimum, is taken as a match.

4.1.5 Parameters used

1. Threshold, ε=0.001
2. Number of training vectors for each user, M = 480 (4 samples of 1 minute each)
3. Number of code vectors, N = 8

4.1.6 Simulation Results

- Total number of users = 8
- Length of each sample = 1minute
- Total number of training samples for each user = 4
- Total number of testing samples for each user = 6
- Total number of testing samples = 6*8 = 48 samples
- Total number of samples matched correctly = 39
- Total number of mismatch = 9
- Accuracy of the algorithm=39/48*100 = 81.25%

Hence, an accuracy of 81.25% was achieved using LGB-VQ algorithm [9] [11]. Table 4.1 elaborates in detail the matching accuracy for different users which is measured using the distortion of the testing samples for each user from the samples from the database.

Table 4.1 Distortion of testing samples for each user

Distortion from different users	TESTING SAMPLES OF USER 1					
	1	2	3	4	5	6
1	123.29602	135.05503	110.91204	112.42552	140.48974	119.35754
2	139.46906	140.05533	118.03497	126.36017	156.21953	135.85946
3	133.13051	138.31153	111.3568	117.43518	145.91584	129.07704
4	131.13084	141.48886	111.5071	119.87895	145.51849	129.00216
5	135.63569	140.32307	119.43973	123.99398	149.82502	129.57042
6	129.7968	138.4468	117.51373	119.79257	145.06039	123.4855
7	130.10378	135.44991	112.68265	120.18531	147.54791	128.03929
8	130.44248	137.69459	113.26103	115.335	147.84013	124.12876
	TESTING SAMPLES OF USER 2					
	1	2	3	4	5	6
1	115.71986	116.68066	105.57611	113.10451	121.50116	99.795568
2	108.68176	108.72687	100.21558	106.85937	116.40973	94.417286
3	111.51221	113.94851	104.33573	109.26654	117.66454	97.027154
4	110.91259	113.54518	104.55136	109.24979	118.32758	99.093
5	112.72898	112.70452	103.32044	109.88826	120.67631	100.08787
6	112.82366	114.1217	103.3564	110.1797	117.57014	99.442057
7	112.99424	112.26002	103.04122	108.68368	116.95251	99.894215
8	112.11043	114.85864	104.1365	110.62567	118.73745	96.626932
	TESTING SAMPLES OF USER 3					
	1	2	3	4	5	6
1	127.49318	93.995419	147.77205	151.6504	138.09103	142.93057
2	132.00905	98.436619	147.2147	152.28172	139.92505	155.58558
3	121.88142	91.0375	141.37242	148.61088	135.84673	134.72329
4	127.88979	96.678393	144.76436	154.581	131.32526	148.85389
5	132.56635	103.07042	155.05401	155.40889	135.93063	150.30692
6	128.94207	98.714613	147.04726	150.87891	131.06824	148.24478
7	127.19962	95.948537	146.89333	153.09533	135.66425	145.78088
8	126.16144	92.194574	145.71407	151.35727	132.71427	144.6421
	TESTING SAMPLES OF USER 4					
	1	2	3	4	5	6
1	123.66165	146.18036	144.46816	128.92931	148.5529	139.72892
2	127.08876	146.48499	149.80419	131.09875	151.77042	140.52855
3	119.92198	144.78224	140.88383	127.78448	147.17763	140.36675
4	114.83553	146.33345	135.02984	125.47581	143.9754	133.88021
5	123.5244	147.08899	145.47209	126.55038	152.28035	140.96865
6	118.13347	148.14468	136.82714	118.24063	147.98747	138.30589
7	124.76335	141.45074	145.12515	127.29052	151.7787	139.90017
8	123.29121	144.50474	143.21481	119.32069	151.12861	136.87513

Distortion from different users	TESTING SAMPLES OF USER 5					
	1	2	3	4	5	6
1	120.69368	141.50944	119.19777	136.38067	127.17992	131.22901
2	116.05916	148.16	122.82356	139.84404	134.81284	138.07743
3	116.95013	139.00664	116.06076	132.23227	126.952	130.96994
4	116.53862	139.21603	115.60681	137.42037	124.00731	127.97106
5	112.13049	136.31119	112.75866	134.52144	119.84244	124.56339
6	117.75783	136.36738	115.19286	134.71796	123.49975	129.60469
7	107.88854	137.36244	118.57622	135.02698	125.8877	128.39566
8	113.84086	138.91668	115.17895	135.8984	123.27937	128.14741
	TESTING SAMPLES OF USER 6					
1	1	2	3	4	5	6
2	111.84658	96.096461	138.33305	132.64034	125.83236	146.00028
3	104.47313	90.191741	130.61441	127.46397	110.31408	133.585
4	96.204789	86.236461	128.43218	127.4774	103.232	138.80068
5	101.32804	90.493212	133.08436	132.43929	118.24336	145.41865
6	92.128324	82.563871	127.64691	125.42082	102.8333	136.91662
7	104.64624	93.461525	134.41351	133.07835	117.57059	140.75542
8	96.638031	83.865445	135.23066	130.23185	109.27844	136.52357
	TESTING SAMPLES OF USER 7					
	1	2	3	4	5	6
1	107.86363	136.94625	148.56329	118.91715	100.64625	145.17064
2	109.06051	144.92832	147.09551	118.62185	104.37783	147.73499
3	105.69257	142.98396	147.90124	117.89723	101.51801	143.08098
4	106.61763	141.59848	152.82664	116.78857	98.6814	148.12468
5	107.38338	142.16568	147.857	117.96179	99.513868	146.96557
6	106.59738	141.54126	149.47664	117.80307	100.25909	148.71874
7	102.08331	138.56263	140.13958	113.09241	95.403361	138.27615
8	105.23385	137.23456	146.80261	117.49934	100.28184	146.2447
	TESTING SAMPLES OF USER 8					
	1	2	3	4	5	6
1	120.23121	127.25893	124.24967	130.74134	125.71729	123.66842
2	134.27944	140.69237	132.79413	138.39877	134.03453	133.59355
3	127.73654	135.1348	129.20715	130.24875	127.27726	128.12592
4	130.13225	132.12627	130.36657	133.11874	127.39573	127.49284
5	127.81403	128.00302	125.77263	135.23219	121.09521	122.57889
6	126.97009	130.01245	127.90983	132.81282	123.73391	125.45549
7	127.40728	128.1145	124.68999	132.5757	122.95325	124.89971
8	124.56823	125.83394	121.76383	131.51536	119.11413	120.74858

4.2 Gaussian Mixture Modeling (GMM)

4.2.1 Introduction

The second classification method implemented for matching is the GMM [10], which can be considered as a stochastic generalization of VQ. The GMM method is a stochastic process where the matching process is based on the probabilistic calculation of the user features.

For identification system using heart sound, the GMM model is trained for each person and a model is generated for each user which contains information based on statistical processing of data. During testing, each user is referred to by his/her model and the user is identified based on maximum probability criteria.

4.2.2 Normalization

In this step, we normalize the extracted feature vectors that we got after the feature extraction method. This limits all the feature vectors to a common range of [-1, 1].

4.2.3 Algorithm description

In the first step we choose the number of component densities required to specify a user. GMM [10] algorithm is named based on number of component densities. For example, if the number of component densities is four, the GMM is called as GMM-4. Similarly, if the number of components is eight, the GMM is called as GMM-8.

Mixture density in GMM is the weighted sum of M component densities and is given by the equation,

$$p(\bar{x}|\lambda) = \sum_{i=1}^{M} p_i b_i(\bar{x})$$

Where, x is a D-dimensional feature vector,

$b_i(\bar{x})$, i=1,2,…, M, are the component densities and

p_i, i=1,2,… ,M, are the mixture weights.

Each component density is a Gaussian function given by [6],

$$b_i(\bar{x}) = \frac{1}{(2\pi)^{D/2}|\Sigma_i|^{1/2}} \exp\left\{-\frac{1}{2}(\bar{x} - \bar{\mu_i})' \Sigma_i^{-1} (\bar{x} - \bar{\mu_i})\right\}$$

Where, μ_i is the mean vector and Σ_i is the covariance matrix.

The mixture weights are bound by the constraint that, $\sum_{i=1}^{M} p_i$=1.

A user can be completely described by the parameters, mean vectors, covariance and mixture weights for all component densities, M. These parameters collectively represented by,

$$\lambda = \{p_i, \bar{\mu_i}, \Sigma_i\}, \quad i=1, 2,\ldots, M$$

For heart sound recognition, each speaker is represented by his/her model λ [10].

4.2.4 Maximum likelihood (ML) estimation

The goal of GMM model is to estimate the parameters λ of the GMM from the training heart signals which describes the distribution of the training feature vectors. The most popular and well-established method for estimating the parameters of GMM is maximum likelihood (ML) estimation [12].

ML estimation finds the model parameters for the given training data and maximizes the likelihood of the GMM [10], [12]. For a given T training vectors $X = \{x_1, x_2, \ldots, x_T\}$, GMM likelihood is given by

$$p(x|\lambda) = \prod_{t=1}^{T} p(\overline{x_t}|\ \lambda)$$

This is a nonlinear function of the parameter λ and hence, direct maximization is not possible. Hence we estimate ML parameters iteratively using an algorithm known as Expectation-Maximization (EM) algorithm [13].

4.2.5 Expectation Maximization (EM) Algorithm

In EM algorithm [13] we begin with an initial model for λ. After that we estimate a new model $\bar{\lambda}$ such that $p(x|\bar{\lambda}) \geq p(x|\lambda)$. This new model is taken as an initial model for the next iteration and the process is repeated until a threshold is reached. The threshold is determined such that the algorithm is convergent.

After each iteration, the parameters are calculated using following formulas which give a surety that the model's likelihood value is monotonically increasing.

- ***Mixture weight:***

$$\overline{p_i} = \frac{1}{T}\sum_{t=1}^{T} p(i\ |\overrightarrow{x_t}\ , \lambda)$$

- ***Means:***

$$\overline{\mu_i} = \frac{\sum_{t=1}^{T} p(i\ |\overrightarrow{x_t}, \lambda)\ \overrightarrow{x_t}}{\sum_{t=1}^{T} p(i\ |\overrightarrow{x_t}, \lambda)}$$

- ***Variances:***

$$\overline{\mu_i} = \frac{\sum_{t=1}^{T} p(i\ |\overrightarrow{x_t}, \lambda)\ \overrightarrow{x_t}^2}{\sum_{t=1}^{T} p(i\ |\overrightarrow{x_t}, \lambda)} - \overrightarrow{\mu_i}^2$$

The *a posteriori* probability for user class *i* is given by [10]

$$p(i\ |\overrightarrow{x_t}\ , \lambda) = \frac{p_i b_i\ (\overrightarrow{x_t})}{\sum_{k=1}^{M} p_k b_k\ (\overrightarrow{x_t})}$$

For training a Gaussian mixture model we select the order *M* of the mixture and the initial model parameters prior to the EM algorithm.

For initialization of model parameters, we take the LBQ-VQ codebook as an initial for the GMM fitting by EM algorithm.

4.2.6 Matching Algorithm

For user identification, a group of *S* users *S = (1, 2,. . . , S}* is represented by GMM's parameters $\lambda_1, \lambda_2, \ldots, \lambda_S$. The user, for which the *a posteriori* probability is maximum, is identified as match [6].

The classification rule [6] is given by,

$$\hat{S} = \arg\max_{1 \le k \le S} p(X|\lambda_k)$$

4.2.7 Simulation Results

The GMM algorithm was implemented first in the training period for generating λ for each user. The number of Gaussian mixture densities for which optimum results came was M=4. Hence this algorithm was GMM-4. During testing, the λ information stored in the training period is used and matching algorithm based on maximum probability of matching was implemented. Table 4.1 clearly demonstrates the a posteriori probability for all the users during training period.

Total number of training signals per user = 6 (of 1 minute each)

Total number of training signals for 8 users = 48

Number of users identified correctly = 35

Hence, accuracy of the GMM algorithm implemented = 35/48 *100 = 72.91%

Table 4.2 probability of matching for GMM algorithm

Probability of matching for users	TESTING SAMPLES OF USER 1					
	1	2	3	4	5	6
1	0.3052453	1.0105913	0.9180256	1.5716292	2.1221714	2.509422
2	0.1058883	0.307824	0.1434022	0.342168	0.2957833	0.2644168
3	0.1174221	0.9957861	0.1305413	0.5526605	0.2156395	1.4903397
4	0.0958147	0.2116015	0.403997	0.3274893	0.2492973	0.7835387
5	0.1257055	0.4180496	0.1265558	0.3460479	0.1159908	0.8470383
6	0.101197	0.3218849	0.4714902	0.6247784	0.1022745	0.6182716
7	0.1561092	0.8169835	0.4834926	0.6128483	0.9862494	0.2269349
8	0.1902646	0.3988331	0.2678484	1.1484947	0.2319656	0.2050847
	TESTING SAMPLES OF USER 2					
	1	2	3	4	5	6
1	0.1334416	2.5153554	0.1389285	0.4523106	1.3060099	0.0492749
2	0.517726	0.6266755	1.673393	0.5758635	0.5717287	0.7390498
3	0.3178255	3.2836101	0.5665971	1.9608275	0.8764505	0.3099441
4	0.2257807	2.5194343	0.8638951	0.516076	0.5010131	0.4643311
5	3.6372873	2.5071275	0.4191767	1.0350282	0.8851531	0.3189091
6	0.4880209	4.5388714	0.5936376	0.6056465	0.7035359	0.3459007
7	1.6538684	8.2678042	0.3270153	5.6415613	1.372504	0.2931081
8	0.8491168	0.805906	0.2712024	0.798455	0.6780509	0.4308308
	TESTING SAMPLES OF USER 3					
	1	2	3	4	5	6
1	2.4534422	1.6566185	0.9050893	0.9656017	1.6544764	0.9442298
2	0.5519181	1.3857494	0.2575687	0.2854358	2.0341275	0.7749498
3	2.7698084	6.9889462	2.6679984	1.4942529	6.9964542	3.856792
4	1.6198385	0.9910375	0.4437287	0.8841203	5.6226856	0.7520534
5	0.3288258	1.2348914	0.5228902	0.3685042	3.3126519	0.9843302
6	0.8257574	2.4734779	0.3450829	0.9187856	2.4661226	1.5249744
7	0.7018838	5.900967	0.2522972	1.4292294	1.6101185	2.8208858
8	0.8743793	1.7436654	0.1549353	0.3588307	1.8778682	0.611791
	TESTING SAMPLES OF USER 4					
	1	2	3	4	5	6
1	0.5911689	0.0391628	0.1749556	0.5137697	3.8145568	1.7664702
2	0.3543888	0.5400244	0.3347942	0.1643372	0.9365388	0.5575518
3	1.0413592	0.7399526	0.3700812	3.3878493	10.070376	0.975605
4	18.809395	0.768313	2.9119717	22.586797	303.42831	26.586224
5	0.2554774	0.4841295	0.3733521	0.6375684	0.8338768	0.2459198
6	0.3183166	0.1627882	0.8602892	0.8719282	11.461788	1.1369508
7	0.4018136	0.2004119	0.135591	1.2735881	0.9108894	2.1603475
8	0.4705731	0.3296181	0.2772492	0.5848233	1.3692882	0.8324867

Probability of matching for users	TESTING SAMPLES OF USER 5					
	1	2	3	4	5	6
1	0.361882	0.1334416	0.7900505	0.4839683	1.1819528	0.5601131
2	0.1665104	0.517726	0.2941798	0.1577066	0.1815039	0.2535528
3	0.265413	0.3178255	0.9056787	0.3269116	1.7068561	1.312091
4	0.0575955	0.2257807	1.6356709	0.169017	1.8280889	0.9731304
5	0.3152453	3.6372873	1.4066512	0.4627475	1.375416	0.9637991
6	0.1356935	0.4880209	1.8362612	0.5126883	4.3805771	1.6120221
7	1.3035036	1.6538684	2.2881673	0.6315786	3.6136082	3.4145925
8	1.1120009	0.8491168	1.5058038	0.1685078	1.1141782	0.5483321
	TESTING SAMPLES OF USER 6					
	1	2	3	4	5	6
1	2.6027397	1.4577717	0.0209098	0.4423808	0.1724882	0.1941023
2	1.312011	0.3143008	0.1683967	0.1650536	0.4582385	0.1615074
3	3.708837	1.1100937	0.4430339	0.9017928	0.8544069	0.2965159
4	10.285453	15.611199	0.1685961	0.2744581	1.1386429	0.3219365
5	2.2747094	1.0487854	0.2584566	0.5595458	0.7315061	0.290012
6	128.23658	7.1624753	0.5115649	2.8780724	1.196135	2.2432338
7	4.0466112	0.4177809	0.1754366	1.0385564	0.3948026	0.2858536
8	5.250256	1.0155271	0.169101	0.3504488	0.6774534	0.3020754
	TESTING SAMPLES OF USER 7					
	1	2	3	4	5	6
1	2.9280155	0.4096631	1.7465065	0.5429076	1.3949801	0.6222646
2	0.3577466	0.1449087	0.4045894	0.3551959	0.6545519	0.0919448
3	4.1168193	0.1373913	0.3090071	0.2794579	2.8447123	0.1433349
4	0.1999171	0.2551535	0.1906565	0.100501	0.920804	0.0534717
5	1.1565806	0.3336019	0.30441	0.6550373	1.048545	0.3821664
6	0.846143	0.8302776	0.1788715	0.3558203	1.3163543	0.153962
7	7.6807815	0.440093	2.840876	2.0135794	6.0861207	1.0441653
8	0.6154201	0.1402846	0.406105	0.2981293	0.8029348	0.2831999
	TESTING SAMPLES OF USER 8					
	1	2	3	4	5	6
1	0.0483223	0.5206269	0.361882	1.6993782	3.0474996	0.4893896
2	0.0393402	0.3135498	0.1665104	0.7023135	1.0992351	0.7336329
3	0.2145518	0.916437	0.265413	1.753626	2.4834402	0.4987379
4	0.1364793	0.4165047	0.0575955	1.4221358	1.9968108	0.5850998
5	0.1641518	0.6293799	0.3152453	0.8664372	1.4827072	0.7739822
6	0.0386734	0.5069832	0.1356935	1.6887445	2.2979435	0.8502223
7	0.0496738	0.6804082	1.3035036	1.0167986	1.0419757	0.2137184
8	0.2151053	1.0883502	1.1120009	2.3169957	2.5265219	2.7703224

CHAPTER 5
CONCLUSION AND FUTURE SCOPE

5.1 CONCLUSION

In this project, we used the PCG signals for user identification. After the initial study of heart signals in time domain and frequency domain we got the motivation to use PCG signals for user identification. Hence, we can conclude that heart sounds can be used as a biometric, and are reliable then other biometric identification systems because heart sounds are least susceptible to attacks from a forger. Heart sound can be itself used for identification or we can use it with other available identification system to make the overall system easy and reliable to implement.

We found that the LFBC algorithm is more suitable for heart sound feature extraction as compared to MFCC for speaker recognition. LFBC was implemented to obtain cepstral components as feature vectors. The important feature of heart sound is that it cannot be easily forged as compared to general biometrics like face, fingerprint, DNA, voice etc. It is not possible to recognize a person if he/she is not living. PCG signals are easy to capture as compared to ECG signals and enables real time identification system design.

We successfully implemented the normalization technique at two places that was not used in [8]. Here we normalized the signals before feature extraction. We also applied normalization on the feature vectors just before using them for GMM algorithm.

We studied and implemented two classification methods for matching: LBG-VQ and GMM. The LBG-VQ-4 code book was used as an initialization of the statistical means and priors for the Expectation Maximization (EM) algorithm of GMM. The accuracy in both the algorithms was found to be less than expected because of the limitations of the Data Acquisition System. The accuracy is expected to increase if the Data Acquisition System is improved to capture a long signal at a time. Moreover, the accuracy for LBG-VQ algorithm was found to be higher than GMM.

5.2 FUTURE SCOPE

- This work can be further improved to increase the robustness of the system; for that, it is highly required to make a robust Data Acquisition System which is least sensitive to noise and can record signals continuously over a long duration of time.
- Other algorithms can be implemented for feature extraction and classification. The main objective could be to find the best algorithm suitable for heart sound processing.
- Further, this work can be extended to make a real time system for user identification and verification.
- The new dimension of the work could be to use the heart sounds to find the heart diseases and other pathological cases.
- The other aim could be to analyse the heart signals over a long period of time to prove its variability or invariability.

REFERENCES

[1] J. Ortega-Garcia, J. Bigun, D. Reynolds, J. Gonzalez-Rodriguez, Authentication gets personal with biometrics, *IEEE Signal Process. Mag.*, 21 (2) (2004) 50–62.

[2] F. Dario, Biometrics: future abuses, *Comput. Fraud Secur.*, 2003 (10) (2003) 12–14.

[3] T. Matsumoto, H. Matsumoto, K.Yamada, S. Hoshino, Impact of artificial gummy fingers on fingerprint systems, *Proc. SPIE*, 4677 (2002) 275–289.

[4] A.K. Jain, A. Ross, S. Prabhakar, An introduction to biometric recognition, *IEEE Trans. Circuits Syst. Video Technol.*, 14 (1) (2004) 4–20.

[5] M. Malik, Heart rate variability: standards of measurement, *physiologicalinterpretation, and clinical use*, Europ. Heart J. Circ. 93 (1996)1043–1065.

[6] F.G. William, *Review of Medical Physiology*, Prentice-Hall, Englewood cliffs, NJ, 1997.

[7] B.N. Robert, *Noninvasive Instrumentation and Measurement in Medical Diagnosis*, CRC Press, Boca Raton, 2002.

[8] K. Phua, J. Chen, T. H. Dat, L. Shue, Heart sound as a biometric, *Pattern Recognition, The Journal of the pattern recognition society*, Pattern Recognition 41 (2008) 906 – 919.

[9] Y. Lindo, A. Buzo, R.M. Gray, An algorithm for vector quantizer design, *IEEE. Trans. Commun.* COM-28 (1980) 84–95.

[10] D.A. Reynolds, R.C. Rose, Robust text-independent speaker identification using Gaussian mixture speaker models, *IEEE Trans. Speech Audio Process*, 3 (1995) 72–83.

[11] http://www.data-compression.com/vq.shtml.

[12] ASSPG. McLachlan, *Mixture Models*, New York Marcel Dekker, 1988.

[13] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Royal Srar. Soc.*, vol. 39,pp. 1-38, 1977.

[14] R. M. Gray, ``Vector Quantization," *IEEE ASSP Magazine*, pp. 4--29, April 1984.

[15] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Kluwer Academic Publishers Norwell, 1991.

[16] H. Abut, "*Vector Quantization*", IEEE Press Piscataway, 1990.

[17] A. Djebbari and B. Reguig, "Short-time fourier transform analysis of the phonocardiogram signal*," in Proc. IEEE Int. Conf. Electron., Circuits, Syst.*, 2000, pp. 844–847.

[18] J.G. Proakis, D.G. Manolakis, *Digital Signal Processing, Principles, Algorithms, and Applications*, Fourth Edition, PHI, 2008.

[19] S. Haykin, *Neural Networks and Learning Machines*, Third Edition, PHI, 2010.

[20] http://en.wikipedia.org.

[21] M. Faundez-Zanuy, "On the vulnerability of biometric security systems", *IEEE Aerosp. Electron. Syst. Mag.*, 19 (6) (2004) 3–8.