بسم الله الرحمن الرحيم

Sudan University of science and technology
College of Graduate studies and scientific Research

Design of Heart Sound Biometric Access Control System

تصميم نظام تحكم للولوج بإستخدام صوت القلب كبصمة حيوية

Thesis submitted in partial fulfillment of the requirements for the award of degree of Master of Science in Mechatronic Engineering

By:

Haram Mohammed Alfateh Mohammed Mustafa

Supervisor:

Dr. Abuagla Babiker Mohammed Babiker

August 2015

الآيـــــة

﴿ وَقُل رَّبِّ زِدْنِي عِلْمًا ﴾

سورة طه -114

Dedication

I dedicate this research with much love and appreciation;

To the candles of my lives. My beloved mother who have always been there for me.

To my father who have always been the brick walls on whom me can learn and depend on forever.

To my brothers and sister who mean the world to me.

To my friends, family, colleagues and teachers in the Past and presents and to everyone that touch my heart.

Acknowledgement

Firstly, thanks to Allah, our creator above for being everything and for giving us the ability and strength to do anything.

*I wish to express my deepest gratitude and appreciation for my supervisor for this research **Dr. Abuagla Babiker Mohammed Babiker** for his patience and continuous guidance, advice and supervision through this work.*

*I would like to extend my gratitude to **T. Almogdad Salah Eldeen Mohammed Hassan** for his supports and encouragement for the completion of this research.*

Also thank and gratitude to all our teachers who contributed too Ur education and to everyone who helped me in this study.

Mogram

Abstract

This project reviews the extent to using of heart sound as biometric system used to identify the person and verify him. Features that have been chosen for this purpose are energy of heart sound signal and power spectrum. The heart sounds were collected from the internet. A matrix laboratory (MATLAB) algorithm was implemented to design a program that used discrete wavelet transform (DWT) for de-noising the signal, continuous wavelet transform (CWT) and Mel Frequency Cepstrum Coefficients (MFCC) for feature extraction and then a databases were build holding sounds features. Finally, Euclidean distance was used for classification sounds and display result depending on mode of operation (identification or verification mode). The program code was implemented in Graphical User Interface (GUI) to easy use of application. Then RS232 was interfaced between it and simulation circuit that consist of ATMEGA 16 (microcontroller), which simulate program via Bascom-AVR as programming language then Proteus was used to verify the code, it was displayed the name of person if the application in identification mode or allowed to access if in verification mode. The application which was programmed and simulated gave a 22.35% error in identification mode when used energy signature and gave a 15.29% error in identification mode when used power spectrum signature.

المستخلص

هذا المشروع يستعرض مدى إمكانية عمل نظام يستخدم صوت القلب كبصمة تستخدم لتمييز الشخص والتعرف عليه. الخواص التي تم إختيارها لهذا الغرض هي طاقة الصوت وقدرته, تم تجميع الأصوات من الشبكة العنكبوتية المعروفة بمسمى الإنترنت, ومن ثم تم تطبيق البرنامج بإستخدام خواريزمية معمل المصفوفات المعروف بمسمى الماتلاب, لتحليل هذه الأصوات, حيث أنه يستخدم طريقة التحويل المتقطع للمويجة لتنقية إشارة صوت القلب ومن ثم إستخدم تقنية التحويل المستمر للمويجة و تقنية ميل معاملات طيف التردد لاستخلاص الخواص لإستخدامها في تمييز الشخص والتعرف عليه, ومن ثم تم بناء قواعد البيانات من هذه الخواص المستخلصة. أخيرا تم إستخدام طريقة المسافة الإقليدية للتصنيف وبذلك تم الحصول على النتيجة وذلك إعتمادا على نمط عمل التطبيق هل نمط التعرف على الهوية أم كان نمط التحقق من الهوية. تم عرض البرنامج في شكل واجهة المستخدم الصورية لتسهيل التعامل مع التطبيق. وتم ربطها بواسطة وصلة تتابعية معروفة بإسم rs232 مع دائرة محاكاة مكونة من المتحكمة الدقيقة ATMEGA 16 حيث تم إستخدام لغة البرمجة BASCOM لإنشاء البرنامج ومن ثم أستخدم برنامج المحاكاة للدوائر الإلكترونية PROTEUS للتأكد من صلاحيه البرنامج، الدائرة تعرض اسم الشخص في حالة العمل في نمط التعرف على الهوية وأما في حالة التحقق من الهوية توضح ما إذا كان يسمح له بالمرور أم لا. حقق التطبيق نسبة خطأ مقدارها 22.35% في التعرف من الهوية في حالة كانت الميزة المستخدمة هي طاقة الصوت بينما حقق نسبة خطأ مقدارها 15.29% في التعرف من الهوية عند إستخدام تردد الصوت كميزة لعمل ذلك.

Contents

List of Figures

List of tables

List of abbreviations

FFT	Fast Fourier Transform
CZT	Z-Chirp Transform
STFT	Short Time Fourier Transform
CWT	Continuous Wavelet Transform
HS	Hilbert Spectrum
DCT	Discrete Cosine Transform
LDA	Linear Discriminant Analysis
MSA	Marginal Spectrum Analysis
SVM	Support Vector Machine
ANN	Artificial Neural Network
VQ GMM	Vector Quantization – Gaussian Mixture Modeling
MFCC	Mel Frequency Cepstral Coefficient
PCG	Phono Cardio Gram

Chapter One

INTRODUCTION

CHAPTER ONE

Introduction

1.1 General Overview

Security is the degree of resistance to, or protection from, harm. Because it is so important the ID card and password techniques was found to do these task, but because it is weak and in the absence of robust personal recognition schemes, these systems are vulnerable to the wiles of an impostor [1].

To avoid fraud the biometric techniques was found. Biometrics, are physiological or behavioral characteristics extracted from human subjects e.g. finger print, iris, face, voice and heart sounds, are used for identity and verification recognition [2, 3]. By using biometrics, it is possible to confirm or establish an individual's identity based on "who he/she is", rather than by "what he/she possesses" (e.g., an ID card) or "what he/she remembers" (e.g., a password) [2].

Biometric system is the system in which a person is identified by means of his distinguishing characteristics or quality basically one belonging to a person. Recognizing humans using computers not only provide some best security solutions but also help to efficiently deliver human services. Biometric authentication system associate behavioral or physiological attributes to identify or verify a person. Physiological characters are based on bodily features, like fingerprint, iris, and facial structure etc. Behavioral traits are based upon unique behavioral characteristics of an individual, like voice, signature etc [3].

With progress of technology, security system requires reliable personal recognition schemes to either confirm or determine the identity of an individual requesting their services. The purpose of such schemes is to ensure that the rendered services are accessed only by a legitimate user and no one else.

Likewise, Human heart sounds are natural acoustic signals conveying medical information about an individual's heart physiology. Provides a reliable biometric for human identification, Heart sound comes under physiological traits because it is a natural sound created by the opening and closure of valves present in the heart [1].

1.2 Problem Statement

A wide variety of systems, Securing personal privacy and deterring identity theft have become essential and inherently important to people and all involved parties and organizations. One of these applications is secure variety of access control mechanism. Biometric recognition refers to the automatic recognition of individuals based on their physiological and/or behavioral characteristics. However, conventional biometric systems for identity recognition suffer from Shortcomings mostly due to the possibility to falsify the in features, their performance is unsatisfactory in certain fields of application, variability with personal and environmental factors, and usually need large databases to process and consequently have high computational costs.

1.3 Proposed solution

Design of an accurate and simplified system for recognition and verification is highly required by using strong techniques for preprocessing, feature extraction and classification steps and selected accurate threshold.

1.4 Objectives

The objectives of this research are to:

- Design a classification technique to prove that the heart sound is considered one of the accurate biometric for authenticating the human into their condescending identity.

- Develop an accurate system for human identification as well as access control which is difficult to disguise, forge, or falsify. One that is easy to use and has lower computational cost.
- Validation and verification of the above mentioned proposed system

1.4 State of the art (part of related work)

There are many approaches proposed that demonstrated the possibility of utilizing PCG signal for human recognition. Like any standard recognition procedure, the proposed approaches include three major stages i.e., preprocessing, feature extraction and classification. The differences between previous studies in feature extraction and classification stages. Some previous studies extracted frequency feature by used FFT or CZT or STFT, another studies extracted energy using CWT or HS, and another used DCT, LDA, MSA, EEMD..etc, but another ones have best results which using LFBCC and MFCC this is in feature extraction stage. In classification stage different methods were used such as ANN, VQ GMM, Euclidean distance …etc. these methods were applied on data size in range between 20-30 persons. These studied had successfully percentage between 55% to 94%.

1.5 Methodology

It has two parts: software and hardware.

Software: A collection of 40 heart sounds for different persons (35 persons have two heart sounds records in different times [4], 5 persons have three heart sounds records in different times) persons was gathered, where each sound passed through three processing stages; preprocessing, feature extraction and classification. The features extracted (signatures) from the second stage are collected in the database. Then a software program using MATLAB was designed to receive a heart sounds, conduct a one-to-many comparison between the sound features and the features stored in the database and give the person identity if found in the database or give a false massage if the person is not in the database (identification), and conduct a one to one comparison between the sound features and the template features (used in access control) (verification) if same to template (do something such as unlocked door) or different (cannot access) make alarm. Hardware: After that connect the MATLAB software with micro controller circuit to display the result in LCD if he want to identify person or start do something or light the led if can access system.

1.5 Thesis Layout

- Chapter two explore the theoretical background and literature review; talk about biometrics (definition, types and uses), biometric systems, heart anatomy and physiology and heart sounds.

- Chapter three consists of design and implementation of a software program for intelligence control access using heart sounds recognition. (MATLAB program, BASCOM program and simulation using proteus). Moreover, part of the results have been shown in this chapter, which proves the practical implementation of the proposed system.

- Chapter four shows results and discussion.

- Chapter five concludes this research as well as drawing future directions for the coming researchers.

Chapter Two

THEORETICAL BACKGROUND & RELATED WORKS

CHAPTER TWO

Theoretical Background and Related Works

2.1 Security and Privacy Motivations

In the new era of technological sophistication and with the growth of the most advanced leading-edge industries, integrity of network transactions, health care, e-commerce, e-government, physical and logical access, and so on, securing personal privacy and deterring identity theft have become essential and inherently important to people and all involved parties and organizations. Some people are still reluctant to engage in ecommerce or conduct other network transactions having misgivings about well-founded systems that will protect their privacy and prevent their identity from being stolen or misused. Identity theft is the case where personal information is accessed by third parties without explicit permission from the owner. Identity fraud occurs when a criminal takes illegally-obtained personal information and uses it for financial gain. Personal information may include the social insurance number, bank or credit card account numbers, passwords, telephone calling card number, birth date, name, address and so on [2].

Identity theft methods such as phishing, hacking and spyware only constitute 12% of fraud cases; whereas, 79% of the reported cases occurred through traditional methods. These instances include stolen and lost wallets, checkbooks, credit cards, or stolen mail from unlocked mailboxes. Finally, a Friendly theft which occurs within friends, family or in-home employees who take private data for their personal gain, this motivated researchers to seek for reliable and accurate alternative for identity verification over the traditional password/ ID card based systems. Biometrics, have emerged to be a new set of technologies that promise an effective solution for this problem. On account of the fact that these credentials have the advantage of residing on the individual so that they cannot be lost, stolen, forged, or subject to failure. In the past few decades, biometrics have been extensively used for law enforcement such as criminal investigations, fatherhood determination, and forensics. Recently, they have figured prominently in a large number of civilian applications for establishing reliable person recognition. Real time recognition is performed by extracting templates and patterns from an individual and comparing them against enrolled records. Leading examples of such technologies are verification or identification systems based on the face, hand, iris, voice, and fingerprint [2].

2.2 Biometrics

The word biometric means literally a measurement of life. Many years ago, it defined as follows: A physiological or behavioral trait which may be measured, stored and thus utilized in subsequent comparison with a live sample for automated identity verification purposes [3].

For a biological trait, whether physiological or behavioral, to be qualified to be used as a biometric characteristic, according to [3]it should satisfy some requirements and criteria. Moreover, in a practical biometric system there are a number of issues that should be considered. Some of these important criteria are listed below: [2]

- Universality - This eminent criterion is connected to the genuine and natural aspect of the attribute so that it resides in all people. Although all suggested biometric traits are met in human subjects, their appearance is not guaranteed in everyone. Cases such as physical disabilities will limit the application of some biometrics to a portion of population. For instance, disabled people cannot be registered in a gait recognition system or voice recognition is not applicable to mute people.

- Uniqueness (Distinctiveness) - Refers to the distinctive ability of a biometric trait to form a unique signature so that any two individuals be sufficiently different in terms of that characteristic.

- Permanence - Is a measure of how well a biometric characteristic resists aging. The traits characteristic should be sufficiently invariant over a period of time with respect to the matching criterion. However, this is not the case with the majority of the biometrics which demand renewing the stored biometric after a period of time.

- Collectability - This criterion refers to both the ease of acquisition of a biometric trait, and its ability to be measured quantitatively.

However, in a system that employs biometrics for personal recognition, there are a number of other issues that should be considered, including:

- Performance, which refers to the achievable recognition accuracy and speed, the resources required to achieve the desired recognition accuracy and speed, as well as the operational and environmental factors that effect on the accuracy and the speed; some of the biometric traits such as DNA are not applicable for real time purposes; however, some others which are implemented in a reasonable amount of time may not have the required accuracy in recognition. Therefore, usually the choice of a biometric attribute depends on the requirements of the application and the properties of the biometric characteristic.

- Acceptability, which indicates the extent to which people are willing to accept the use of a particular biometric identifier (characteristic) in their daily lives;

- Circumvention is one of the most crucial and eminent criteria which determines the extent to which a biometric characteristic can be deployed in different applications. It reflects the robustness of the system against fraudulent methods and attacks.

Given these criteria, there is no general rule to determine the ideal biometric characteristic among all. Since each biometric has its strength and weaknesses; the choice of the optimum biometric is left upon the deciding authority based on the specifications of a biometric characteristic and the application context.[3]

2.2.1 Biometrics have two Categories

• Physiological

Also known as static biometrics: Biometrics based on data derived from the measurement of a part of a person’s anatomy. For example, fingerprints and iris patterns, as well as facial features, hand geometry and retinal blood vessels [5].

• Behavioral

Biometrics based on data derived from measurement of an action performed by a person and, distinctively, incorporating time as a metric, that is, the measured action. For example, voice (speaker verification) [5].

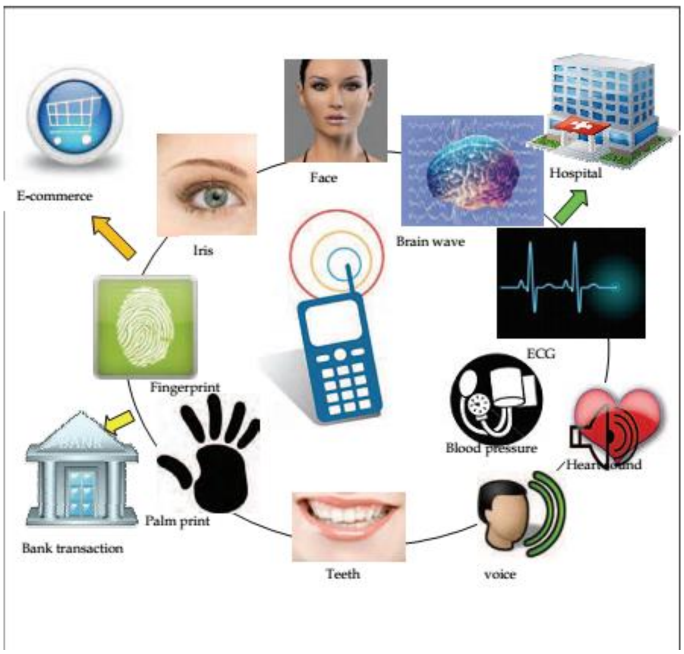


Figure 2-1 Different Types of Biometrics.

2.2.2 Biometric system

In recent years, it has become very important to identify a user in applications such as personnel security, defense, finance, airport, hospital and many other important areas [5]. 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 behavioral traits. Hence, conventional biometric identification systems such as iris, fingerprint, face and speech have become popular for user identification and verification [6].

A biometric system is essentially a pattern recognition system that operates by acquiring biometric data from an individual, extracting a feature set from the acquired data, and comparing this feature set against the template set in the database. Depending on the application context, a biometric system may operate either in verification mode or identification mode [3].

2.2.3 Verification mode

This method consists in verifying whether a person is who he or she claims to be (Figure 2-2). It is called a ”one to one” matching process, as the system has to complete a comparison between the person’s biometric and only one chosen template stored in a centralized or a distributed database, e.g. directly on a chip for an identity document. Such a method is applied when the goal is to secure and restrict specific accesses with obviously cooperative users. Identity verification is typically used for positive recognition, where the aim is to prevent multiple people from using the same identity [3].

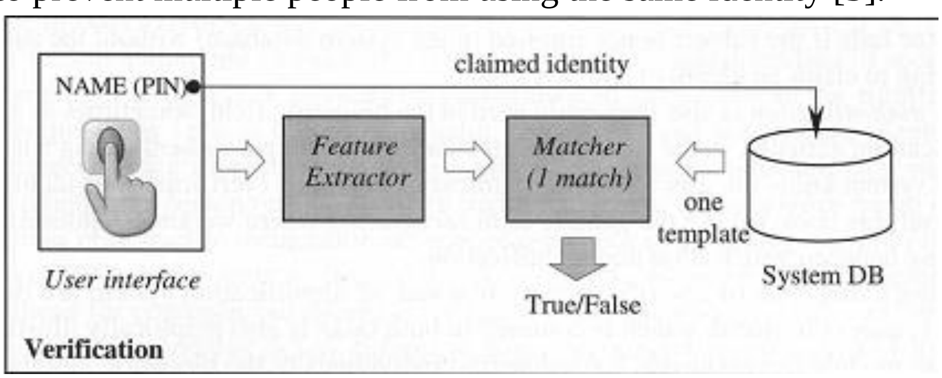


Figure 2-2 Verification Mode [3].

2.2.4 Identification mode

In this mode, the system recognizes an individual by searching the templates of all the users in the database for a match. Therefore, the system conducts a one-to-many comparison to establish an individual's identity (or fails if the subject is not enrolled in the system database) without the subject having to claim an identity (e.g., "Whose biometric data is this?"). Identification is a critical component in negative recognition applications where the system establishes whether the person is who she (implicitly or explicitly) denies to be. The purpose of negative recognition is to prevent a single person from using multiple identities [7]. Identification may also be used in positive recognition for convenience (the user is not required to claim an identity). While traditional methods of personal recognition such as passwords, PINs, keys, and tokens may work for positive recognition, negative recognition can only be established through biometrics [3].

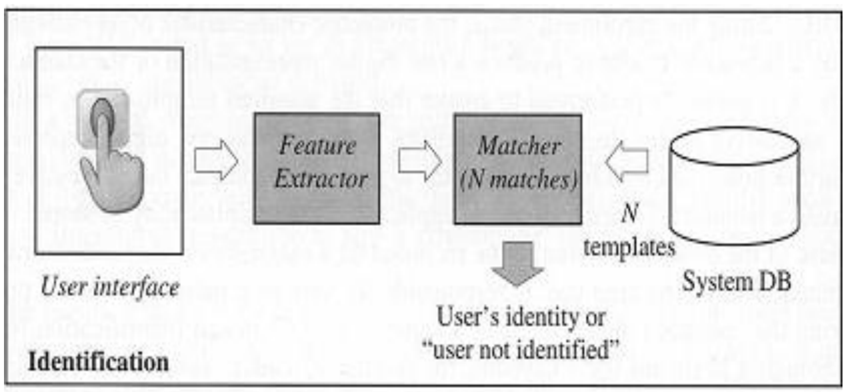


Figure 2-3 Identification Mode [3].

2.2.5 Components of a Biometric System

The actual number of discrete components in a biometric system varies with each system, and there are a variety of ways to segment a generic biometric system. However, most biometric systems consist of some variant of the following five components or subsystems: (1) sensor/data capture, (2) feature extraction, (3) data storage (also called template storage), (4) matching algorithm and (5) decision process [5].

Sensor/data capture (data collection) acquires a sample of an individual's biometric characteristics (for example, an image or signal) needed for recognition, and then converts the sample to a digital format. This process is also referred to as biometric presentation. The quality of the sensor has a significant impact on the system results. Indeed, one could argue that the sensor is the most important component of a biometric system. Biometric data collection takes place during one's registration or enrollment, and precedes verification or identification transactions. The biometric data that is captured should be high quality because poor resolution can produce false negatives and may result in re-enrollments. Biometric traits can be collected in various ways. A photograph can collect data for facial recognition, and a recorder can capture a quality voice print. The captured data is referred to as raw data. Consistency in the manner in which a sample is collected is critically important. Changes in sample collection can affect the accuracy of the template, which in turn affects the probability of a match. The sensor output is sent to the signal processor to select and to extract the distinguishing characteristics of the sample [5].

Feature extraction automatically processes the acquired information by extracting only the key data from the distinguishing features of a biometric sample, generating a digital representation referred to as the biometric template, and performing some quality control checks. This may involve a feature extraction, or it may require locating the biometric characteristics within the received sample. This is where the signal processing algorithm comes into play. Predetermined, discriminative specific features are extracted from raw data to form a new representation of only system significant information; the rest of the raw data is discarded. This new representation should be unique for each individual as well as be somewhat invariant with regard to multiple samples garnered from the same person over time. The biometric system then digitizes, compresses, and encrypts those features to produce a template. It is this template, not the raw data that will be used in the matching process. The feature extraction is executed during the enrollment process as well as during the authentication process. During enrollment, the extracted biometric data is used to create a reference template with associated personal data. In the authorization process, the feature extraction is applied to each frame of the scanned image prior to the matching process with the enrollment template. The extraction process selects the distinguishing characteristics from the raw data sample, converts it to digital format, and then processes the digitized data into a compact biometric identifier record—the biometric template. Generally, biometric templates require much less memory for storage than the biometric data samples. Raw data is simplified through the process of feature extraction. After the noise removal and adaptive algorithms such as contrast enhancement are applied, the biometric system creates a binary image of the biometric representation, and then it stores that extracted data. At the completion of the extraction process, it is virtually impossible to reconstruct the original hand scan from the digitized template. The selection of the biometric features is highly dependent on the extraction algorithm. Consequently, the features extracted by one algorithm can differ markedly from the features extracted by another algorithm even for the same body part. The implication of this reality is that a template produced by one biometric system may not be fully compatible with another biometric system of the same modality even when used on the same body part of the same person. However, the template incompatibility situation can be beneficial from a security standpoint. Extraction usually entails performing several quality control activities to ensure that the extracted features are likely to be adequately distinguishing. If the signal processing component rejects the received sample, then the sensor/data capture component collects another sample. If the signal processing component accepts the sample, it then creates a template from the extracted data. Depending on the deployed biometric system, during the enrollment process multiple sets of the individual's biometrics characteristics might be captured to create a single high-quality enrollment template, or multiple templates are created and stored to account for intraclass variations. Intraclass variations are situations in which a user's live data sample will be somewhat dissimilar to the user's template created at enrollment. The cause of the variations might stem from differing interactions of the user with the sensor (e.g., inconsistent placement or pressure), changes in the environmental conditions (e.g., lighting variation), use of different sensors during enrollment and at verification, changes in pose, or changes in the biometric trait (e.g., illness or aging). Such situations lead to a high false rejection rate. On the other end of the spectrum is interclass similarity. With interclass variation, the randomness of a data pattern among individuals is generally the norm, as measurable traits can vary significantly; and that contributes to low false acceptance rates. However, in some isolated cases, rather than variation, these large interclass similarities in the feature sets limit the discriminatory ability expected for that trait. High levels of interclass similarity lead to higher false authentication rate [5].

Template storage houses the enrollment templates to which the new biometric templates will be compared. Templates are stored within an enrollment database. Templates may be stored within a biometric capture device, on a portable medium such as a smart card, in a distributed device such as a personal computer or local server, or in a central repository. The templates can be stored in an altered format, compressed, and encrypted [5].

Matching algorithm compares each new sample template to the stored reference template and matches the similarities. The analysis is then passed as similarity or match scores to a decision processor. The similarity scores indicate the degree of fit among the templates compared. With verification applications, a single specific claim of subject enrollment results in a single similarity score, which is usually a quantitative probability estimate. With identification applications, all reference templates in the database may be compared with the subject’s sample template, and the outputs include a similarity score for each comparison. However, biometric systems only present a likelihood of a match as expressed in a probability; and despite the fact that these likelihoods are relatively accurate, they are not absolute. For larger biometric populations, response time becomes very important, and it is highly dependent on search and retrieval accuracy. Biometric data does not have a natural sorting order; and the computational overhead of pattern matching for large-scale identification systems can be extensive, requiring hours or even days of processing. Therefore, biometric practitioners have had to devise techniques to index biometric information to reduce that overhead and, consequently, to accelerating the matching process. One well-accepted method is to accomplish this is by classifying the biometric data. Thus, binning may come into play. Binning is the process of classifying biometric data to enable the presorting of very large biometric databases. This technique can significantly speed the effort in matching sample biometric data with the reference template. In binning, reference templates are partitioned into characteristics or bins. The bins selected for classifying data can be based on external characteristics (such as gender) or on internal characteristics (such as whorls, loops, and arches in the fingerprint modality). Indeed, many traditional fingerprint identification applications are based on binning. Binning is very useful in accelerating identification processing, and it enables more accurate statistical matches. Nevertheless, all such classification techniques can introduce hidden obstacles—if the reference template has been inaccurately categorized, it becomes very difficult to obtain a correct match [5].

A decision processor uses the scores of the matching component to make a system-level decision for a verification or identification transaction. Unlike a password-based system in which the submitted code is either correct or it is not, biometric matches are based on probabilities within intervals of confidence. A newly created sample template is generally not an exact match to the reference enrollment template due to an array of variables. Of course, there is usually some tolerance involved, but biometric systems can differ significantly regarding the amount of variance that can be accepted for matching. Most biometric systems establish a threshold that must be exceeded for a match to occur. Therefore, a threshold should be set at a level to accept all similarity scores that are considered matches but high enough to reject fakes. A decision is rendered based on the match score and a predetermined threshold of parameters for acceptance or rejection. Most biometric systems base their decisions on statistical methods, whereby the average of several samples is calculated, and a normal distribution of the biometric data with a mean value and standard deviation is formulated. The quantitative difference between the sample and template must lie within a certain threshold to be accepted as a match, or it is rejected. Templates are considered a match when the similarity score exceeds a preselected threshold. The individual's claim can then be verified on the basis of the decision policy. This decision process can be fully automated or it can support human intervention. The decision policy may allow or require multiple attempts before making an identification decision [5].

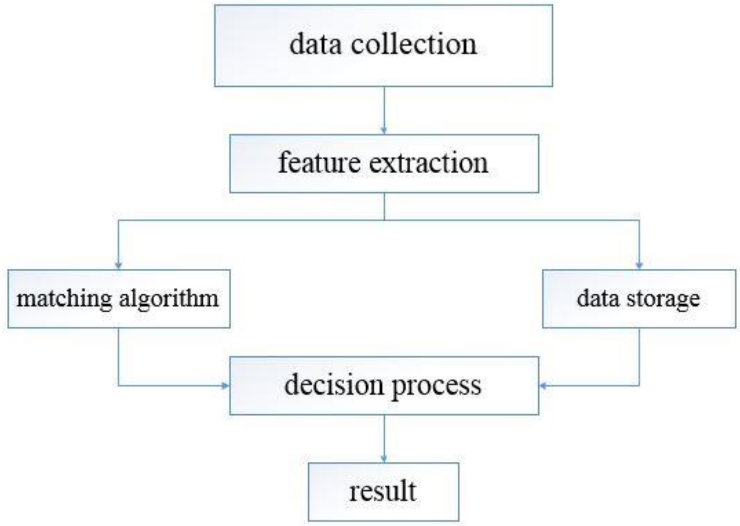


Figure 2-4 Biometric System Components.

2.2.6 Biometric operations using the processing steps

Enrollment A user is added to the biometric system. A certain number of biometric presentation of a particular user are acquired, preprocessed, transformed into features, and post processed, then used to train a user model and adapt (retrain) the world model if necessary. The user model along with impostor presentations may be used to obtain a threshold for that user. The new model is then stored, along with the threshold for that user if needed. Verification The claim to a user's identity causes the presented biometric data to be compared against the claimed user's model. Thus, the biometric data is acquired, preprocessed, transformed into features, and post processed, before being matched with the claimed user's model and the resulting score being compared with the stored threshold computed for the claimed user or a generic threshold value [3].

Identification A database of user models is searched for the most likely source of the biometric presentation. Thus, the biometric data is acquired, preprocessed, transformed into features, and post processed, before being matched with all the user models of interest. The user model that obtains the highest score with respect to the presentation is suggested to be the source of the presentation [3].

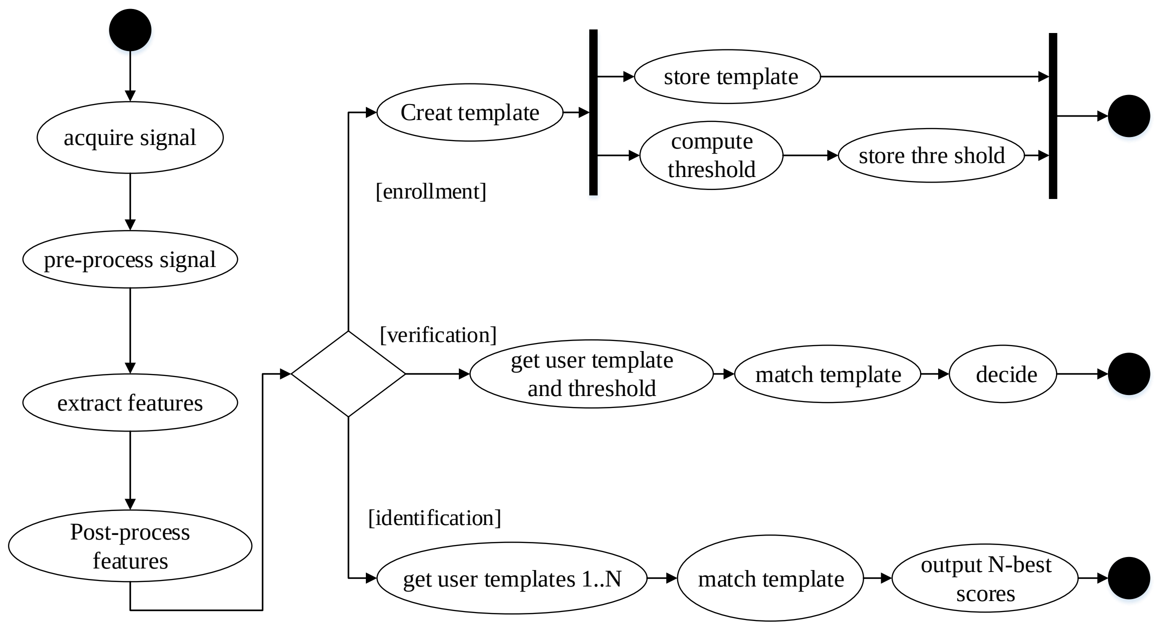


Figure 2-5 Steps of Enrollment (Make Database), Identification and Verification.

2.3 Heart sounds

Human heart sounds are natural signals, which have been thoroughly explored for health monitoring and medical diagnosis for hundreds of years. Heart auscultation, which is the interpretation of heart sounds by a physician, is a fundamental component of cardiac diagnosis. This interpretation includes the evaluation of the acoustic properties of heart sounds and murmurs such as the intensity, frequency, duration, number, and quality of the sounds. It is, however, a skill difficult to acquire. The term phonocardiography (PCG) refers to the tracing technique of heart sounds and the recording of cardiac acoustics vibration by means of a microphone-transducer. So far, the study of PCGs has focused mainly on the heart rate variability characterization of the PCG components, detection of structural abnormalities and heart defects. However, demonstration of the feasibility of applying ECG signal for human recognition and the achieved performance as well as robustness of this biometric has drawn the attentions to the PCG signals as well. Having the same origin with the ECG signal in addition to the medical information that is conveyed through the PCG signal conjectured that the PCG signal may contain information about an individual's physiology. Signals having this characteristic usually have the potential to provide a unique identity for each person. Like the ECG signals, the PCG signals are difficult to disguise, forged, or falsified. Moreover, this signal has the advent of being relatively easy to obtain, since it can be collected by placing a stethoscope on the chest. This work is concerned with utilizing the PCG signals as a physiological biometric which is a new concept only lately suggested in the literature [2].

2.3.1 Heart Anatomy

The human heart (Figure 2-6) is a muscular organ with four main chambers, two upper chambers, known as atria, and two lower chambers, known as ventricles. The four chambers of the heart, namely, the right atrium, right ventricle, left atrium, and left ventricle, supplies the force that drives blood through the circulatory system. Blood is pumped away from the heart through arteries and returns to the heart through veins [8].

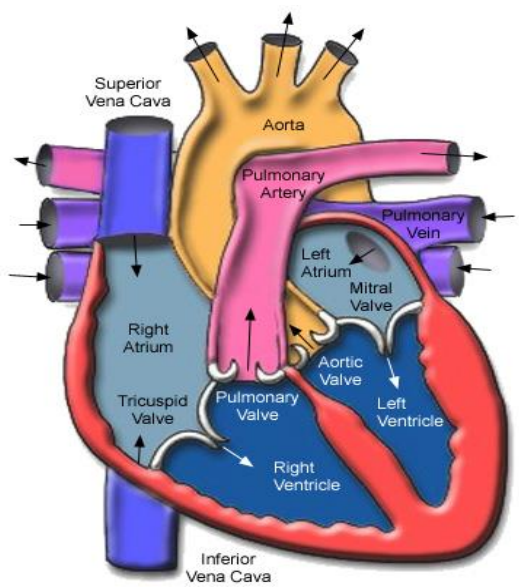


Figure 2-6 Heart Anatomy.

The human heart has four valves that help to direct flow of blood in the desired direction and prevent blood from flowing in the reverse direction. The mitral and tricuspid (atrioventricular) valves separate the left and right atria from the left and right ventricles respectively. The aortic and pulmonary (semilunar) valves separate the left and right ventricles from the aorta and pulmonary artery respectively. All four heart valves are attached to a fibrous cardiac skeleton. This skeleton is made of dense connective tissue and acts as a form of attachment for the valves and for the ventricular and atrial muscles [8].

2.3.2 Cardiac Cycle

The cardiac cycle is a synchronized sequence of contractions and relaxations of the atria and ventricles during which major events occur, such as valves opening and closing and changes in blood flow and pressure. Each contraction and relaxation is referred to as systole and diastole, respectively. Figure2 shows the events related to the cardiac cycle [8].

The diagram starts at late ventricular diastole. At this stage, the AV valves are open and the ventricles near their maximum blood volume capacity. Atrial systole will then occur, pushing the blood through the AV valves, filling the ventricles even more, increasing their pressure and volume. Next, as the ventricles begin to contract (ventricular systole), ventricular pressure (VP) rises above atrial pressure (AP), forcing the AV valves to shut. Since the semilunar valves are also closed, ventricular volume remains constant during this small period, known as Isovolumetric Contraction, causing a rapid increase in VP. When VP exceeds the pressure of the exit vessel (pulmonary artery and aorta for the right and left heart, respectively) the semilunar valves open, leading to the ejection of blood [9].

As the systole ends, the ventricular walls begin to relax (ventricular diastole) causing VP to drop drastically, falling below the exit vessel pressure, which causes the closure of semilunar valves. This period is referred to as Isovolumetric Relaxation because both semilunar and AV valves are closed, resulting in a constant ventricular volume and a further drop in VP. When VP falls below AP, AV valves open and blood flows into the ventricles, finally completing the cycle [8, 9].

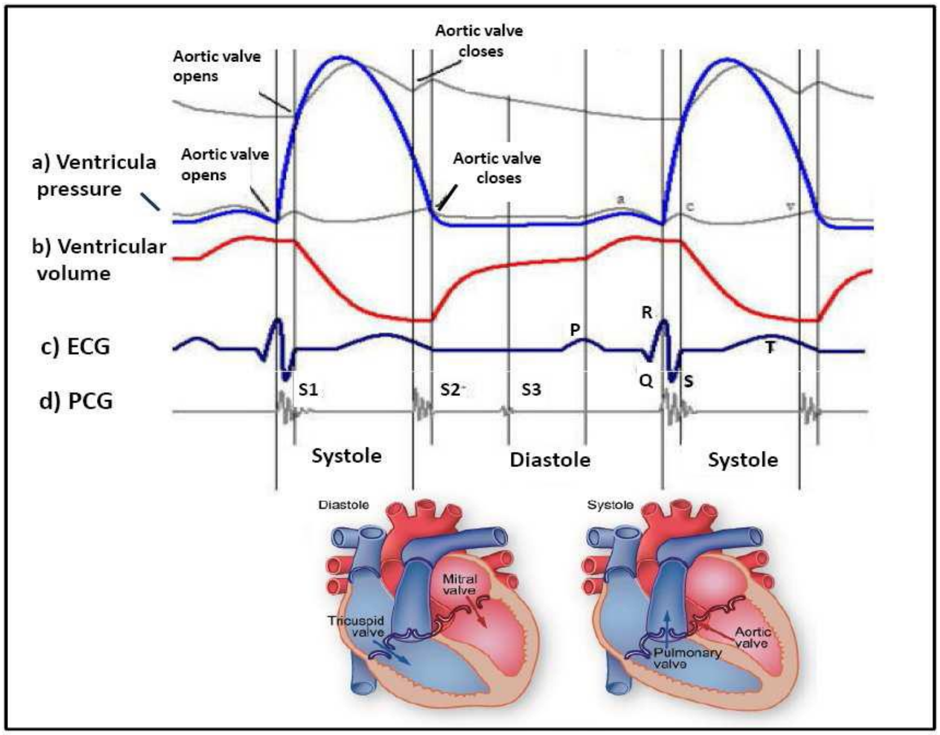


Figure 2-7 Signals of Cardiac Cycle (A) Ventricular Pressure,(B) Ventricular Volume,(C) ECG and (D) PCG(Heart Sounds)

2.3.3 Heart Sounds

In a medical context the heart sound signal is collected from four main regions on the chest wall as demonstrated in Figure 2-8 .The aortic (A), between the second and third intercostal spaces at the right sternal border; mitral (M), near the apex of the heart between the fifth and sixth intercostal spaces in the mid-clavicular line; pulmonic (P), between the second and third intercostal spaces at the left sternal border; and tricuspid (T), between the third, fourth, fifth, and sixth intercostal space at the left sternal border [10].

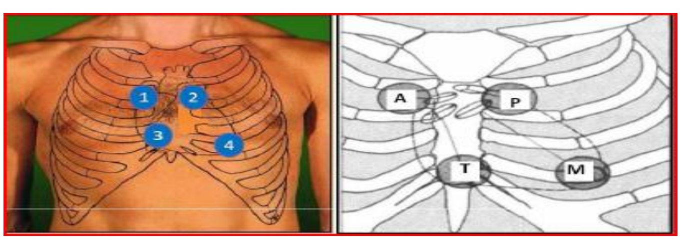


Figure 2-8 Auscultation Sites to Place Stethoscope.

2.3.4 Mechanism of Heart Sounds Production

Heart sounds can be heard throughout the heart cycle and are caused by several cardiac events such as ventricular filling, blood flow and, most of all, valve movements [8].

2.3.5 Normal Heart Sounds

There are four main heart sounds, called S1, S2, S3 and S4. Normally, only two sounds are audible, S1 and S2 sounding like the words “lub – dub”. S3 and S4 are extra heart sounds heard in both normal and abnormal situations [8].

I. First Heart Sound S1

The first heart sound forms the "lub" of "lub-dub". S1 occurs at the onset of the ventricular contraction during the closure of the AV- valves. It contains a series of low-frequency vibrations, and is usually the longest and loudest component of the PCG signal. The audible sub-components of S1 are those associated with the closure of each of the two AV-valves. S1 lasts for an average period of (100−200) ms and its frequency components lie in the range of (25−45) Hz. It is usually a single component, but may be prominently split with some pathology [2]. These components are the mitral and tricuspid components. The mitral component occurring slightly before the tricuspid one due to earlier occurrence of left ventricular contraction. S1 occurs just right after the QRS complex of the ECG [8].

II. Second Heart Sound S2

The second heart sound forms the "dub" of "lub-dub". S2 is heard at the end of the ventricular systole, during the closure of the semilunar valves. S2 lasts about 0.12s, with a frequency of 50Hz which is typically higher than S1 in terms of frequency content and shorter in terms of duration. It has aortic and pulmonary sub-components: A2 and P2 corresponding to the aortic part and pulmonary part respectively. Usually A2 and P2 are closed together, but a split S2 can occur if A2 and P2 are just far enough apart that they can be heard as two beats within S2. S2 occurs towards the end of the T wave [3, 8].

III. Third Heart Sound S3

S3 is the third low-frequency sound that may be heard at the beginning of diastole, during the rapid filling of the ventricles. Its occurrence can be normal in young people (less than 35 years of age). In ECG, S3 which has relatively lower energy occurs right after S2 [2, 8].

IV. Fourth Heart Sound S4

S4 is the fourth heart sound that may occur in late diastole during atrial contraction shortly before S1. It is always considered as an abnormality within the cardiac cycle. In ECG, S4 occurs after the P wave [2, 8].

2.3.6 Abnormal Heart Sounds

I. Murmurs

Murmurs are high-frequency, noise-like sounds that are heard between the two major heart sounds during systole or diastole. They are caused by turbulence in the blood flow through narrow cardiac valves or reflow through the atrioventricular valves due to congenital or acquired defects. They can be innocent, but can also indicate certain cardiovascular defects [2, 8].

Murmurs are described as systolic or diastolic according to their timing in the cardiac cycle. Thus, a murmur heard after the first heart sound and before the second is a systolic murmur, and which comes after the second and before the first is a diastolic murmur [2, 8].

II. Clicks and Snaps

Are associated with valves opening and indicate abnormalities and heart defects. Opening snaps of the mitral valve or ejection sound of the blood in the aorta may be heard in case of valve disease (stenosis, regurgitation). The opening snap when present, occurs shortly after S2 with the opening of the mitral and tricuspid valves [8]. Clicks are short high pitched sounds, and have three types:

1. Ejection click: is the most common click, which occurs shortly after S1 with the opening of the semilunar valves.

2. Aortic ejection clicks.

3. Pulmonic ejection clicks.[2]

2.4 Heart sounds biometric

Biometric uses the methods for recognizing users based upon one or more physical and behavioral 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 2-1 [11].

Table 2-1 weaknesses of some biometric method.

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

Biometric system using heart sounds has a number of benefits compared to the standard methods of biometrics. A lot of parts of human body had already been use in biometric technologies, and heart is not an exception. But the heart is a life support organ of the body, so that information security system will work in any human body condition and regardless of the user actions [11].

The heart sound biometric technology differs from other ones in the following properties:[11]

- Heart sonic signals cannot be lost during the life;
- heart sonic signals are difficult to be falsified;

- access control may be performed without user's actions and continuously during operation of medical devices;
- Heart sonic signals allow to diagnose the human physical and psychological condition.

The use of heart sonic signals as the authentication method is possible due to availability of a melody, which is typical of every specific man. This feature was revealed when studying the heartbeats sonic signal spectrum. The characteristics of human heart sounds may change in time depending on the human physical or psychological condition but the frequency change sequence remains unchanged during a certain period. Just the frequency change sequence produces the musical pattern of the heart sonic signal [3].

2.4.1 Salient Features of heart Sound as Biometric

Following are the factors, which differentiate heart sound from other physiological properties: [12]

i. Universality: It is universal in nature. It means it is present in every individual's heart. (Every living human being has a pumping heart.)

ii. Easy accessibility: PCG signals can be recorded using an electronic stethoscope and put it on the chest.

iii. Invariability: It is invariant in nature, because it is somehow depend upon the anatomy of the heart. (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.)

iv. Accuracy: Accuracy plays an important role in the biometric with minimum False Accept Rate (FAR) and high True Accept Rate (TAR).

v. Speed: A quick response is very desirable for a biometric system. This includes quick data acquisition, quick enrolment of inputs and processing.

vi. Reliability: There are not any chances of the forgery, because to steal the heart sound a concealed device has to be placed at chest.

vii. Usability: It is totally a time saving, a person can send a signal from several meters also.

viii. Permanent: It cannot change over a time.

2.4.2 Heart sounds biometric application

Heart sound as phonocardiogram based biometric system will be a good option or not. That is depending on the application. For universality purpose, this biometric system is best for all privacy point of you. Although for accessing BANK locker, Password protected jewellery shop, Personal Computers etc, also, using heart sounds as biometric characteristics for secure medical personal data in healthcare devices [11, 13-15].

2.5 Error Definitions

There are some challenges involved in utilizing biometric features for recognition purposes due to variations among multiple recordings of an individual. Having technical or physical origin, raw biometric data contain noise and artifact components that alter its expression from the ideal anticipated structure. Moreover, environmental and physiological factors may create variance among multiple recordings of the same subject. These factors may render interpretations inaccurate or misleading. There are four system states according to the recognition result: [2]

- Correct Acceptance: The case where the system accurately identifies the subject.
- Correct Rejection: The case where the system correctly does not validate a claimed identity.
- False Acceptance: occurs in verification mode where someone is mistakenly accepted by the system.

- False Rejection: occurs in verification mode where a genuine identity request is denied by the system.

The false acceptance rate and the false rejection rate are enumerated and computed using Eq. 1 and Eq. 3.

$$FAR(n) = \frac{Number\ of\ successful\ independent\ fraud\ attempts\ against\ a\ person\ n}{Number\ of\ all\ independent\ fraud\ attempts\ against\ a\ person\ n} \qquad \text{eq.1}$$

These values are more reliable with more independent attempts per person/characteristic. In this context, independency means that all fraud attempts have to be performed with different persons or characteristics! The overall FAR for N participants is defined as the average of all FAR (n):

$$FAR = \frac{1}{N}\sum_{n}^{N} FAR(n) \qquad \text{eq.2}$$

$$FRR = \frac{Number\ of\ rejected\ qualified\ attempts}{Total\ number\ of\ qualified\ attempts} \qquad \text{eq.3}$$

In order to reassure that impostors and intruders are not verified or validated by the system, usually a threshold is defined so that the system is not forced to either accept or reject a subject. In cases where the amount of resemblance between a pair is less than a predefined threshold, the system will refuse to validate the claimed identity as to be registered in the database. The FAR and FRR curves are plotted as a function of different thresholds and along with the equal error rate (EER) , i.e., the point at which FAR and FRR curves meet, constitute the specifications of the biometric system. Depending on the classification criteria employed by the system, the distribution of the FAR and FRR may vary. Yet, since both FAR and FRR are functions of the system's threshold, there is always a tradeoff between these two types of errors. Increasing the FAR due to lower thresholds makes the system more tolerant to input variations and noise. However, raising the threshold, in order to make the system more secure, decreases the FAR. Therefore, the operation point of the system is decided based on the requirements of the application [2].

2.6 Related Works

Human heart sounds are natural, nonlinear, non-sinusoidal and exponenttial signals; the acoustic properties of heart sounds and murmurs such as the intensity, frequency, duration, number, and quality of the sounds are the fundamental components for heart sounds analysis. These signals can be captured by placing a stethoscope on the chest or by the use of a sound recording device phonocardiograph (PCG) [2].

There are many approaches proposed that demonstrated the possibility of utilizing PCG signal for human recognition. Like any standard recognition procedure, the proposed approaches include three major stages i.e., preprocessing, feature extraction and classification.

- Phua, “Human identification using heart sound", proposed an approach for PCG recognition through the frequency analysis of the short-time discrete Fourier transform (STDFT) of the PCG traces. The PCG spectrum was processed by filtering out the frequency band outside the range of 20−150Hz and further enhanced by application of a spike removal technique. The extracted cepstral coefficients were used as the biometric features in conjunction with the discrete cosine transform (DCT) for reducing the dimensionality of the data space. Two conventional classifiers were tested for the classification stage: the Vector Quantization (VQ) and the Gaussian Mixture Modeling (GMM). However, due to the number of iterations required to train the GMM, the proposed scheme is slow and time consuming; an issue that affects the application of such a scheme in a large scale scenario. Moreover, the designed preprocessing step is incapable of reducing the inter-band noise which degrades the performance of the system for noisy data, results indicate that with well-chosen parameters, an identification rate of up to 96% is achievable for a database consisting of 10 individuals and containing 100 sounds for each person, with heart sounds collected over a period of 2 months [16].

- Another approach was proposed by Beritelli, “for human identification based on the frequency characteristics of the S1 and S2 sounds in digital PCG sequences”. A mechanism was proposed to identify the boundary of the S1 and S2 in the PCG traces. Then, the frequency analysis was performed using the Z-chirp (CZT) transform with a frequency band of 20Hz − 100Hz for obtaining an energy trend profile. The obtained signal spectrum was used as the feature vector from the PCG signal and classification was performed using the Euclidean distance measure. However, the localization and delineation of S1 and S2, which is essential for the succeeding stages, is a great challenge in the presence of noise. The complexity of this issue is brought to light by considering the fact that there is no universal definition for determining the onset and offset of these components [17].

- Another one Fatemian, “Heart id: Cardiac biometric recognition”, for human identification based on time and frequency characteristics, do not propose a pure-PCG approach, but they rather investigate the usage of both the ECG and PCG for biometric recognition. In this short summary, we will focus only on the part of their work that is related to PCG. The heart sounds are processed using the Daubechies-5 wavelet, up to the 5th scale, and retaining only coefficients from the 3rd, 4th and 5th scales. They then use two energy thresholds (low and high), to select which coefficients should be used for further stages. The remaining frames are then processed using the Short-Term Fourier Transform (STFT), the Mel-Frequency filter bank and Linear Discriminant Analysis (LDA) for dimensionality reduction. The decision is made using the Euclidean distance from the feature vector obtained in this way and the template stored in the database. They test the PCG-based system on a database of 21 people, and their combined PCG-ECG systems has better performance [2, 18].

- Another one Saad Daud, “for heart sound as a physiological biometric signature based on energy percentage in each wavelet coefficients”. An electronic stethoscope designed and implemented by the authors is used to record more than 30 heart sound and for each recording a signature is calculated. The signature is calculated during one heart cycle by calculating the total energy for each wavelet coefficient after normalization, filtering and denoising of each recording by using DWT 5th order DAUBECHIES wavelet, Feature Extraction based on CWT

MORELET wavelet and Euclidean distance is used as an identification tool between the signatures It was found that the distances for different recordings at different days for the same person are below a threshold value and above the threshold value as compared with the signatures of the other persons. Software based on MATLAB is used for recordings, denoising, signature calculations and signature comparison to prove the implemented system [19].

- Girish Gautam, first made data collection from ten volunteer of the age group 20-40 during three months period using Digital Stethoscope (100 heart samples stored in database). Then feature extraction using LFBC (linear frequency band cepstral), feature extraction method includes STDFT for converting the time domain signal into frequency domain. Then magnitude was taken and rejecting the phase part which generally include noise interference. Next the filter bank is applied, which reject the unwanted high frequency components. After that Dimension compression technique was used. Using DCT (Discrete Cosine Transform) here logarithmic first 24 coefficient was taken. Then Spike removal is done for removing the artifacts of position of hand movement while taking heart sound. At last, cepstral means subtraction is done, which removes the artifacts, here position of stethoscope is not same at all the time, after this operation is done, cepstral coefficient as our feature vector. Then Classification is done, using BP-MLP-ANN where 50 numbers of heart sound signal as Training and 50 numbers of heart sound signal as Testing are applied. The identification results show 52 % of performance accuracy [12].

- Arunava Karmakar, 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. 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). 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, 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 [20].

- Zhidong Zhao, Qinqin Shen and Fangqin Ren, heart sound biometric system based on EEMD, Hilbert Spectrum (HS), Marginal Spectrum Analysis (MSA) and Discrete Cosine Transform (DCT) which is a new feature extraction technique for identification purposes. This heart sound identification system is comprised of signal acquisition, pre-processing using fifth-order Daubechies Discrete Wavelet Transform, feature extraction, training, and identification using Euclidean Distance. Experiments on the selection of the optimal values for the system parameters are conducted. The results indicate that the new spectrum coefficients result in a significant increase in the recognition rate of 94.40% compared with that of the traditional Fourier spectrum (84.32%) based on a database of 280 heart sounds from 40 participants [21, 22].

- Ajay Singh and Ashish Kumar Singh, Database of heart sound is made using the electronic stethoscope. In the beginning, heart sounds for different classes' isobserved 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, 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 [6].

- Swati Verma and Tanuja Kashyap, "analysis of heart sound as biometric using MFCC & linear SVM classifier", 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. Mel frequency Cepstral Coefficients (MFCCs) has been used for feature extraction and then these feature vectors are classified to recognize a person, using Support Vector Machine (SVM) as classifier applied in 30 persons, The identification results show 90 % of performance accuracy [23].

- The authors of Tran et al.," Feature integration for heart sound biometrics," in Acoustics Speech and Signal Processing (ICASSP) ", one of which worked on Phua et al. (2008), take the idea of finding a good and representative feature set for heart sounds even further, exploring 7 sets of features: temporal shape, spectral shape, Cepstral coefficients, harmonic features, rhythmic features, cardiac features and the GMM super vector. They then feed all those features to a feature selection method called RFE-SVM and use two feature selection strategies (optimal and sub-optimal) to find the best set of features among the ones they considered. The tests were conducted on a database of 52 people and the results, expressed in terms of Equal Error Rate (EER), are better for the automatically selected feature sets with respect to the EERs computed over each individual feature set [24].

- In Jasper & Othman, the authors describe an experimental system where the signal is first down sampled from 11025 Hz to 2205 Hz; then it is processed using the Discrete Wavelet Transform, using the Daubechies-6 wavelet, and the D4 and D5 sub bands (34 to 138 Hz) are then selected for further processing. After a normalization and framing step, the authors then extract from the signal some energy parameters, and they find that, among the ones considered, the Shannon energy envelogram is the feature that gives the best performance on their database of 10 people [25].

- The authors of El-Bendary et al. filter the signal using the DWT; then they extract different kinds of features: auto-correlation, cross-correlation and cepstra. They then test the identities of people in their database that is composed by 40 people, using two classifiers: Mean Square Error (MSE) and k-Nearest Neighbor (kNN). On their database, the kNN classifier performs better than the MSE one [26].

- Osama and Sh-Hussain etc, “Efficient Speaker Verification System Based on Heart Sound and Speech”, the method the method selects the best fusion and normalization techniques for biometric system. The framework is developed and test the verification task. The approach in this paper is biometrics recognition, for example, providing features that can’t be easily copied, such as the Mel-Frequency Cepstral Coefficient (MFCC) as a feature vector and vector quantization (VQ) as the matching model algorithm .A simple yet highly reliable method is introduced for biometric applications. Experimental results show that the recognition rate of the Heart sound- speaker verification (HS-SV) provides an average EER of 17.8% while the average EER for the speech speaker verification model (S-SV) is 3.39%. In order to reach a higher security level an alternative to the above approach, which is based on multimodal and a fusion technique, is implemented into the system. The best performance of the work is based on simple-sum score fusion with a pricewise-linear normalization technique which provides an EER of 0.69% while the fusion type Main Rule provide an EER of 1.1% [27].

- Ekaterina Andreeva, “Alternative Biometric as Method of Information Security of Healthcare Systems”, As information technologies develop, there appear more and more medical devices for diagnosis of human condition and maintenance of life-support. Such devices as insulin pumps, implanted pacemakers and defibrillators are used for diagnosis, treatment and monitoring of the patients' condition. These devices operate using sensors interacting with the human body. The increasing complexity of such devices induced development of a Body Area Network technology which makes it possible to arrange the sensor network on the human body and perform continuous monitoring of the organism's condition. The possibility to continuously record the patient's data is useful not only in the treatment and diagnosis but also for design of the information security system for life-support medical devices. In this article discuss the problem of human authentication in BASN. Biometric technology using heart sounds applied in approach; moreover it has a number of benefits compared to the standard methods of biometrics [11].

- Francesco Beritelli and Salvatore Serrano," Human Identification Based on Segmentation and Frequency Analysis of Cardiac Sounds", The performance of traditional biometric identification systems is as yet unsatisfactory in certain fields of application. For this reason other physiological or behavioral characteristics are recently being considered, using new electrical or physical signals linked to a person's vital signs. This paper examines the biometric characteristics of PCG (Phono Cardio Gram) signals from cardiac auscultation. More specifically, the paper proposes a preliminary study related to the identification of individuals via frequency analysis of cardiac sounds. The results, obtained using a database containing several heart sound recordings from 20 different people, confirm the biometric properties of PCG signals, which can thus be included among the physiological signs used by an automatic identification system [28].

- Shailesh Singh Badghare, "phonocardiogram signal for biometric application", Heart sound is distinctive in nature. Earlier work reported that, it can also contribute a lot to recognize a person by their heart sound. A novel technique is described in this thesis for the identification and verification of the person using energy based feature set and back propagation multilayer perceptron artificial neural network classifier (BP-MLP-ANN) is used in this thesis. PCG signal is invariable, unique, universal easy to accessible and unique in nature. Heart samples were collected through ten volunteers as ten data (i.e. heart sounds) per individuals. Before feature extraction, pre-processing involves extraction of cycles, alignment, and segmentation of primary heart sound S1 and S2. This Segmentation contributes to the features extraction based on energy taken 30 windows at a time. Classification was done, using BP-MLP-ANN. 69 % of total numbers of heart sound signal were used as Training and remaining 31 % of heart sound signal were used for Testing. The identification results show 63.3 % of performance accuracy [1].

- Guo et al. used a feature set of linear prediction cepstrum coefficient (LPCC), the hidden Markov model (HMM), and wavelet neural network (WNN) to acquire the heart sound classification information and to realize identity recognition [29].

- Cheng et al. presented a synthetic model of heart sounds and then used the heart sounds' linear band frequency cepstrum (HS-LBFC) as a specified configuration with similarity distance to achieve recognition and verification [30].

- Osamah and Sh-Hussain Salleh, "Multimodal Biometrics Based on Identification and Verification System", the need for an increase of reliability and security in a biometric system is motivated by the fact that there is no single technology that can realize multi-purpose scenarios. Experimental results showed that the recognition rate of Heart Sound Identification (HSI) model is 81.9%, while the rate for Speaker Identification (SI) model is 99.3% from 20 clients and 70 impostors. Heart Sound-Verification (HSV) provides an average Equal Error Rate (EER) of 13.8%, while the average EER for the Speaker Verification model (SV) is 2.1%. Electrocardiogram Identification (ECGI), on the other hand, provides an accuracy of 98.5% and ECG Verification (ECGV) EER of 4.5%. In order to reach a higher security level, an alternative multimodal and a fusion technique were implemented into the system. Through the performance analysis of the three biometric system and their combination using two multimodal biometric score level fusion The best performance of the work is based on simple-sum score fusion, with a piecewise-linear normalization technique which provides an EER of 0.7% [31].

- Andrea Spadaccini," Performance Evaluation of Heart Sounds Biometric Systems on an Open Dataset", Recently, many systems and approaches that employ heart sounds as physiological traits for biometric recognition have been investigated. However, those systems are often tested on small, diverse and closed datasets, making it difficult to compare their performance. In this paper, we present HSCT-11, an open dataset containing data collected from 206 people that can be used for performance evaluation of heart sounds biometric systems, and we use it to benchmark two such systems. The most performing one shows an Equal Error Rate of 13.66 % on this database, a result that will be the baseline for all the future evaluations made using this dataset [32].

Table below was used to explain the methods used in any step (data collection, preprocessing, feature extraction and classification), but was not represent performance in table for two reasons: first, most papers do not adopt the same performance metric, so it would be difficult to compare them; second, the database and the approach used are quite different one from another, so it would not be a fair comparison.

Chapter Three

DESIGN AND SIMULATION

CHAPTER THREE

Design and Simulation

In this chapter, an approach based on both energy and power spectrum properties of the heart sounds signal are developed for analysis of this signal for application in human recognition system. The proposed method divided in to two parts: software and hardware.

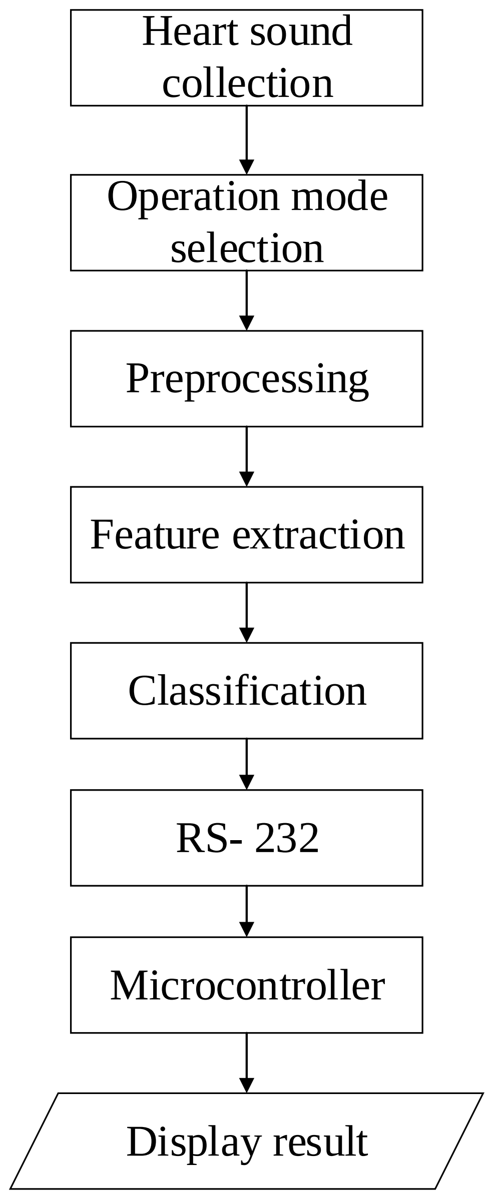


Figure 3-1 General block diagram For Heart Sound Biometric Sequence.

3.1 Software

The algorithm includes three major stages i.e., preprocessing, feature extraction, and classification, the respective descriptions of which are provided in the following sections.

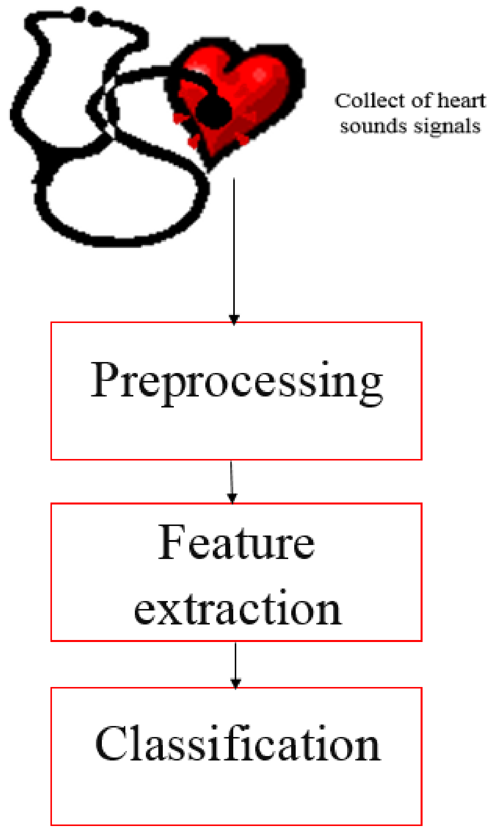


Figure 3-2 Block Diagram of Software Algorithm

Data collection

A collection of 40 heart sounds for different persons (35 persons have two heart sounds records in different times [4], 5 persons have three heart sounds records in different times)

In implementation, a collection of 40 raw heart sounds signals were gathered and entered into MATLAB GUI for processing. (1Figure (3-4))

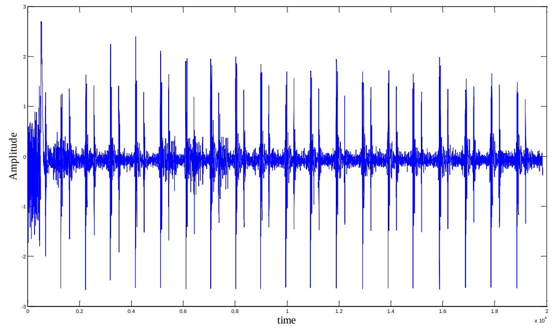


Figure 3-3 Raw Heart Sounds Signal for Person 1 (Sound 1).

Pre-processing

The purpose of this step is to eliminate noise and enhance heart sounds, making them easier to segment and identify. This stage consists of three steps:

- **De-noising**

High quality signals are essential for correct identification. Unfortunately, the presence of noise in heart sounds signals is inevitable. Even when all background noise is minimized there are always intrinsic sounds impossible to avoid: respiratory sounds, muscular movements and so on. Therefore, the de-noising stage is extremely important, ensuring elimination of noise and emphasizing relevant sounds [33].

Due to the overlapping nature of noise with the spectra of the heart sounds signal, simple band pass filtering is not effective for noise reduction. However, decomposing the signal in narrower sub-bandwidths using the wavelet transform enables us to perform the temporal noise reduction for the desired bandwidth sections. The mother wavelet implemented here is the Debauches wavelet of order 5 (db5) (figure 3-2).The choice is due to the heart beat signal having most of its energy distributed over a small number of db5 wavelet dimensions (scales), and therefore the coefficients corresponding to the heart beat signal will be large compared to any other noisy signal (figure 3-3). The de-noising procedure involves three steps [19, 33]:

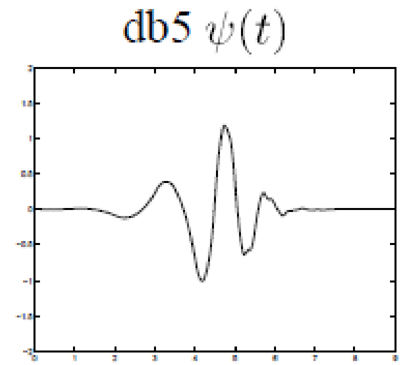


Figure 3-4 Debauches 5 (db5)

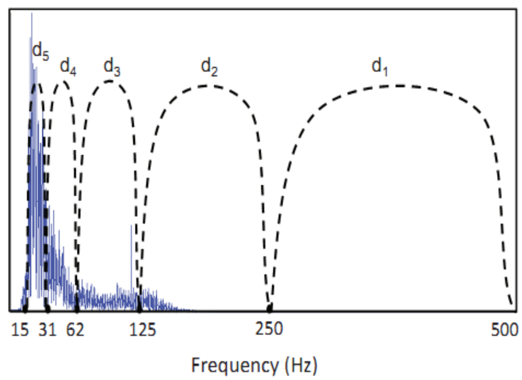


Figure 3-5 Equivalent Frequency Responses of the DWT, over the Heart Sounds Spectrum.

- **Decomposition**

The heart sound in this step is divided into approximations and details, where the approximations represent the slowly changing (low frequency- high scale) features of the signal and details represent the rapidly changing (high frequency– low scale) features of the signal.

A decomposition of level 5 with the (db5) wavelet was selected for the decomposition part of the denoising algorithm [19, 33].

- **Thresholding detailed coefficients**

This step involves thresholding the detailed coefficients of the DWT and then reconstructing the signal with the inverse discrete wavelet transform

(IDWT). There are two common methods for thresholding, soft thresholding and hard thresholding. The method chosen is the soft thresholding, where it produces better result than hard thresholding because it sets the elements whose absolute values are lower than the threshold to zero and then the nonzero coefficients remaining are shrunk and set to zero. In the other hand hard thresholding sets the elements whose absolute values are less than the threshold to zero [33].

- **Reconstruction**

The last step in the de-noising procedure is to compute the wavelet reconstruction through the summation of the original approximation coefficients of the last level (level5) and the modified detail coefficients of levels 1 to 5 [19, 33].

Each sound was passed through the preprocessing stage. It was threshold and normalized. (Figure (3-5)).

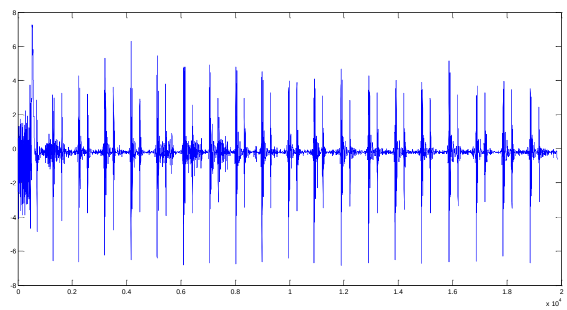


Figure 3-6 De-noised Sound1.

- **Normalization**

Normalization is used to restrict all the signals to the same range.

- **Segmentation**

The purpose of heart sounds segmentation is to separate all the cardiac cycles in a recording so that each can be analyzed individually. Here the segmentation was done manually and two cycles where selected from each sound.

In implementation, then it was manually segmented taking 2 cycles for each segment (figure 3-6, 3-7).

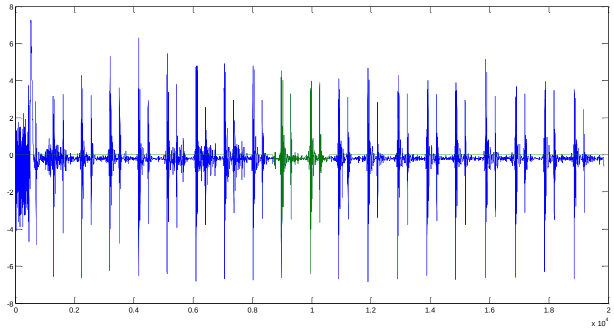


Figure 3-7 The Selected Segment (Green Colored) From Denoised Sound 1

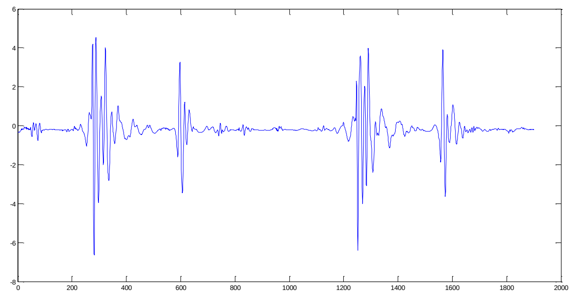


Figure 3-8 Two Cycles (One Segment) Form De-noised Sound 1.

Feature Extraction

The heart sounds is non-stationary signals and have features in both time and frequency domain, in this project extracted energy and power spectrum features from heart sounds signal.

- **Energy**

Due to the non-stationary nature of the heart sounds signal, the continuous wavelet transform was used, although a redundant transformation, but its redundancy tends to reinforce the traits and makes all information more visible, where the wavelet coefficients respond to changes in the waveforms strongly providing the required features of the signal. , This approach solved the STFT's resolution problem by analyzing the signal at different frequencies with different scales, it is designed to give good time resolution and poor frequency resolution at high frequencies and good frequency resolution and poor time resolution at low frequencies. Therefore, it made sense when the signal at hand had low frequency components for long durations [2, 34].

The mother wavelet used here is the MORLET wavelet function. It is more suitable to use because it is symmetrical, more compact (compressed) and also decays to zero fast, provides better frequency resolution, and determines accurately the frequencies of the heart sounds [34].

Here, one of parameters chosen to be extracted from the heart sounds signal as a feature is the energy of the sound. The energy is preservative for the transform. So the following equation is tenable:

$$energy = \int |x(t)|^2 dt = \frac{1}{c_\varphi} \int \frac{1}{a^2} da \int |CWT_x^\varphi(b,s)|^2 db \qquad \text{eq.4}$$

This equation represents the total energy of a domain centered at (b, a) with scale interval Δa and time interval Δb. |CWT(b ,a)|2 is defined as the wavelet scalogram. It shows how the energy of the signal varies with time and frequency, Cψ is known as admissibility constant, and is determine by:

$$c_\varphi = \int_0^\infty \frac{|\varphi(f)|^2}{f} df \qquad \text{eq.5}$$

Where ψ (f) is the Fourier transform for wavelet function and f is central frequency fc [34].

The previous formula was applied to calculate energy for all coefficients, then, using the values to calculate total energy for each scale sample. The feature extracted from this stage is collected to form a database.

In implementation, CWT was applied for each segment; energy of each coefficient was calculated using the SCALOGRAM, and then the coefficients for each point of scale where summed .The resultant matrix of the summation holds the desired feature (signature) of the heart signal.

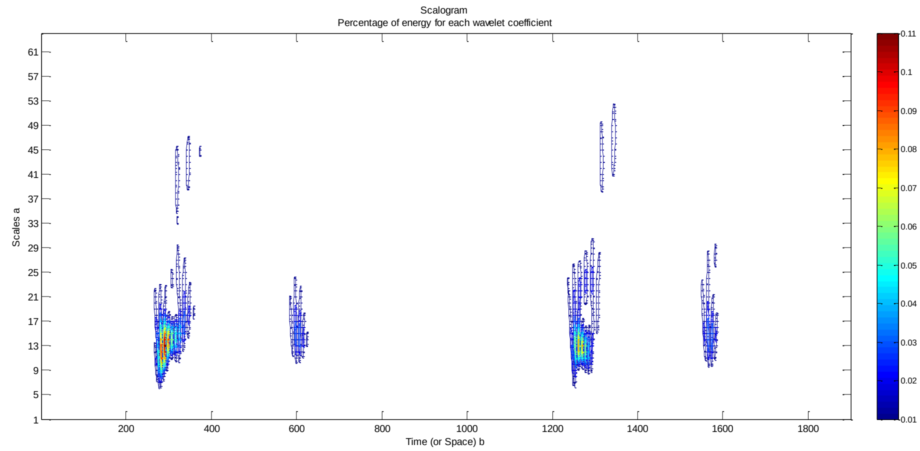


Figure 3-9 Scalogram of Selected Segment

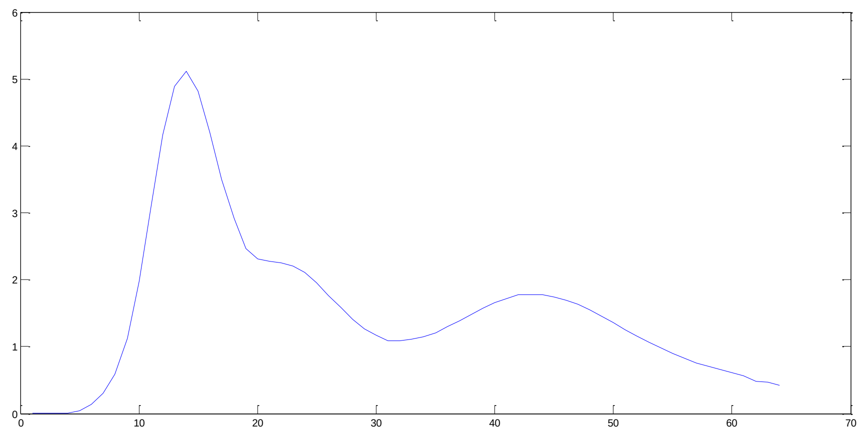


Figure 3-10 PERSON 1 Energy Signature.

All sounds were passed through the previous stages and then the features for all sounds were collected to create the desired database.

- **MFCC Feature (Power spectrum)**

Feature extraction is a special form of dimension reduction, which transforms the input data into the set features. Heart sound is an acoustic signal and many techniques used nowadays for human recognition tasks borrow speech recognition techniques. The best and popular choice for feature extraction of acoustic signals is the Mel Frequency Cepstral Coefficients (MFCC) which maps the signal onto a Mel-Scale which is non-linear and mimics the human hearing. MFCC system is still superior to Cepstral Coefficients despite linear filter-banks in the lower frequency range. The idea of using Mel Frequency Cepstral Coefficients (MFCC) as the feature set for a PCG biometric system comes from the success of MFCC for speaker identification [35] and because PCG and speech are both acoustic signals. MFCC is based on human hearing perceptions which cannot perceive frequencies over 1Khz. In other words, in MFCC is based on known variation of the human ear's critical bandwidth with frequency [23].MFCC has two types of filter which are spaced linearly at low frequency below 1000 Hz and logarithmic spacing above 1000Hz. Mel-frequency cepstrum coefficients (MFCC), which are the result of a cosine transform of the real logarithm of the short-term MFCCs are provide more efficient. It includes Mel-frequency wrap-ping and Cepstrum calculation [23].

Mel-Frequency Cepstrum Coefficients (MFCC) are one of the most widespread parametric representation of audio signals (Davis & Mermelstein (1980)). The basic idea of MFCC is the extraction of cepstrum coefficients using a non-linearly spaced filterbank; the filterbank is instead spaced according to the Mel Scale: filters are linearly spaced up to 1 kHz, and then are logarithmically spaced, decreasing detail as the frequency increases. This scale is useful because it takes into account the way we perceive sounds. The relation between the Mel frequency f˘mel and the linear frequency flin is the following: [35]

$$f_{mel} = 2595. log_{10}(\frac{1+f_{lin}}{700}) \qquad \text{eq.6}$$

Some heart-sound biometry systems use MFCC, while others use a linearly-spaced filterbank. The first step of the algorithm is to compute the FFT of the input signal; the spectrum is then feeded to the filterbank, and the *i*-th cepstrum coefficient is computed using the following formula:

$$c_i = \sum_{k=1}^{K} X_k \,.\cos\left(i.\left(k-\frac{1}{2}\right).\frac{\pi}{K}\right) \quad i = 0,\ldots,M \qquad \text{eq.7}$$

Where K is the number of filters in the filterbank, X_k is the log-energy output of the k-th filter and M is the number of coefficients that must be computed [36].

Calculation of MFCC coefficients

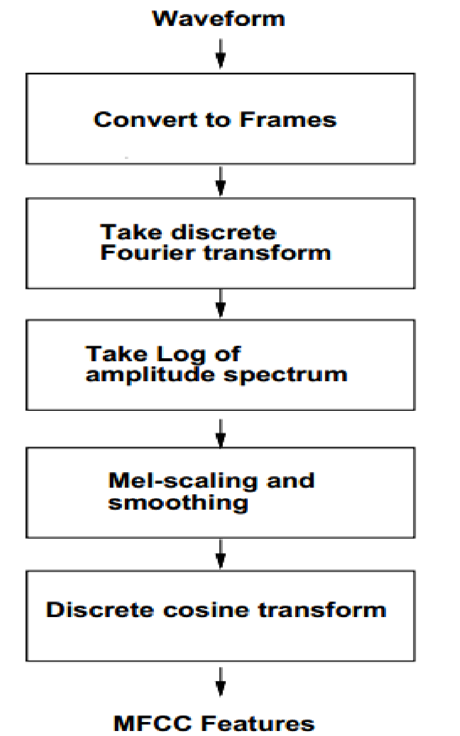


Figure 3-11 Steps of MFCC Calculation's.

From The previous formula was applied to calculate MFCC for all coefficients, the feature extracted from this stage is collected to form a database.

In implementation, MFCC was applied for each segment; power spectrum of each coefficient was calculated using this method, and then the coefficients for each point of scale where summed .The resultant matrix of the summation holds the desired feature (signature) of the heart signal.

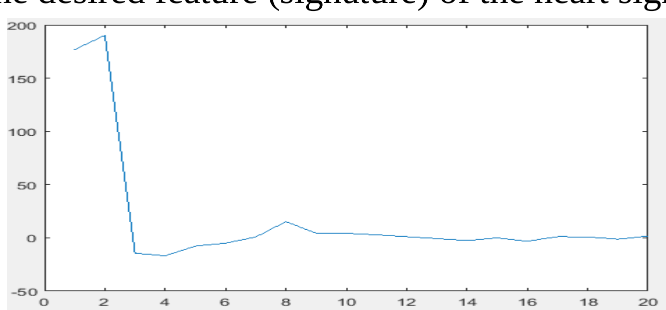


Figure 3-12 Person 1 Power Spectrum Signature.

Also, all sounds were passed through the previous stages and then the features for all sounds were collected to create another desired database.

Classification

In this stage the feature of the sound is compared with features stored in the database, and the distance (error) is calculated by using a method called Euclidean distance.

Euclidean distance is the distance between two points, determines by:

$$d(p,q) = \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2} \qquad \text{eq.8}$$

Where d is Euclidean distance, p and q are the arrays and n is the dimension of arrays [37].

Use this formula to find the result for any feature (energy and power spectrum) which use in identification mode or verification mode.

In implementation, the classification part code ,to receive a heart sound, conduct a one-to-many comparison between the sound and the collected signatures stored in the database and give the person identity if found in the database or give a false massage if the person is not in the database, that it was done if use system in identification mode, but when use it as verification mode it received heart sound , conduct a one to one comparison between the sound and the collected signatures stored in the database and give person and allow access and light the blue LED if that person who allowed to access system (if yes, in another word, if the Euclidean distance less than threshold which choose in the program) or give an alarm sounds if the person is not allowed to access to the system.

The block diagram shown in Figure below describes the block diagram of the classification step.

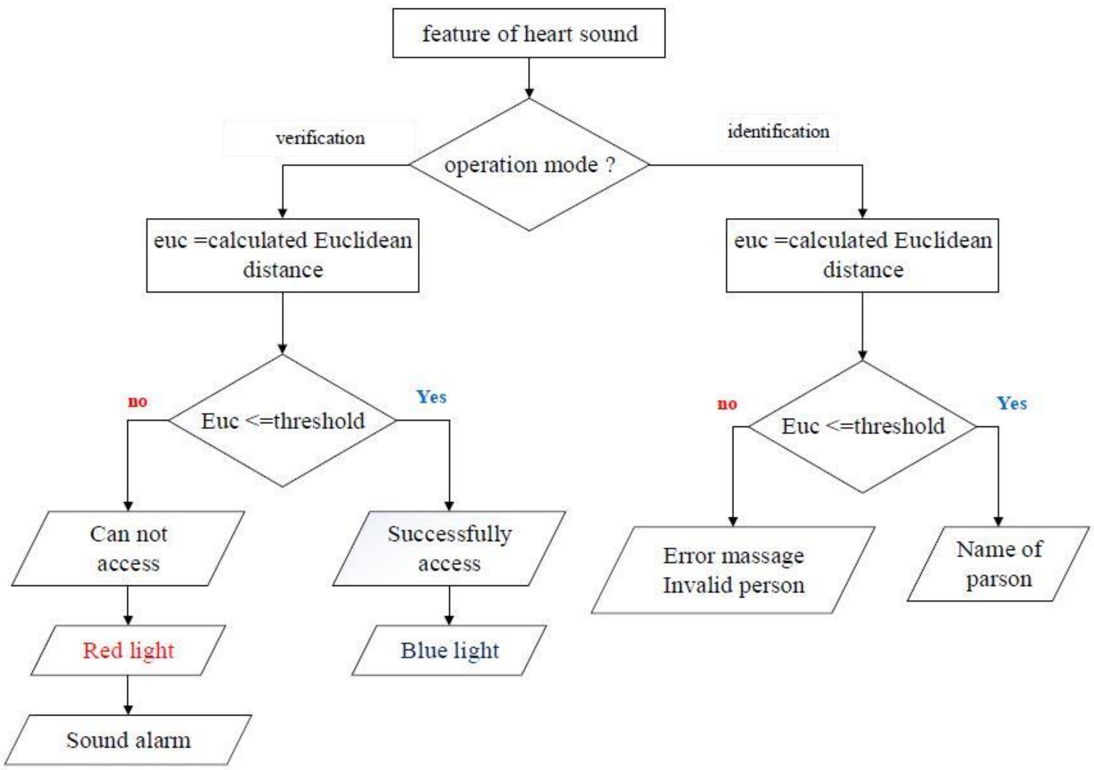


Figure 3-13 Block Diagram of Classification Step.

3.2 Hardware circuit

❖ The Figure below shown all the Circuit

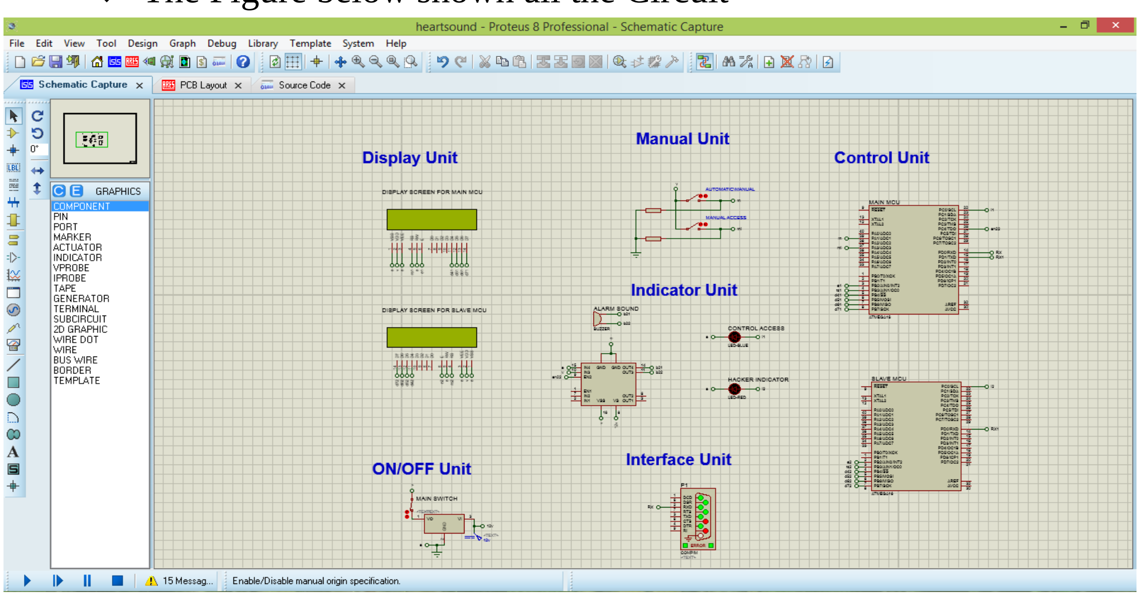


Figure 3-14 Circuit of Heart Sound Biometric System.

The circuit has different units explain below:

- ON/OFF Unit works as main switch to close the circuit (make switch ON) and start to work or to OPEN the circuit (make switch OFF) and end work.

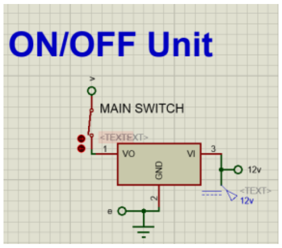


Figure 3-15 ON/ OFF Unit.

- Control unit consists of two microcontroller units, main microcontroller unit and slave micro controller unit. Main microcontroller unit was received data from GUI via interface unit and display result and send data to slave microcontroller unit if hacker was happened. Slave microcontroller unit was receive data from main microcontroller unit when hacker was happened and make red visible alarm and displayed message.

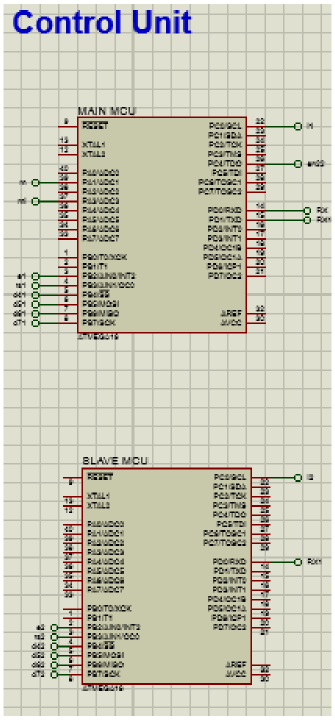


Figure 3-16 Control Unit.

3.2.1 Microcontroller

Basically a microcontroller is a computing device, and is a single integrated circuit (“Silicon chip” or IC) used to form part of a product that incorporates some software Program control. As a microcontroller is basically part of a computing system it can be used in applications requiring control, operator and user display generation, simple sequencing and many other mundane tasks [38].

A microcontroller device is not simple, but in general, a microcontroller unit may be considered as a computing device offering internal memory and a high level of input and output (I/O) device options. Ideally the use of a microcontroller device minimizes the number of external devices used in the system, and integrates as much of the external interfacing to switches, motors or other input / output devices as is practically possible [39].

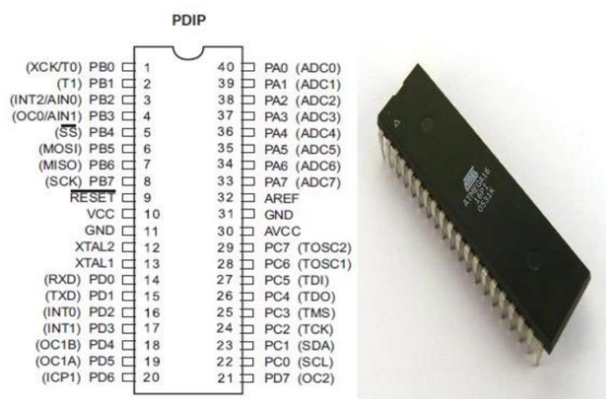


Figure 3-17 ATmega16.

- Interface unit consists of serial port (RS-232) which use as connect point between GUI and control unit, it used to send data from computer to microcontroller unit.

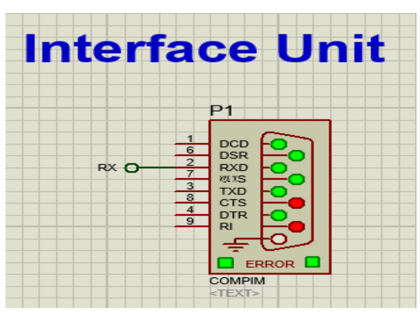


Figure 3-18 Interface Unit.

3.2.2 Rs-232 (recommended standard)

Figure 3-19 RS 232 connection.

RS-232 is a standard for serial data communication between computing equipment. This standard dates back to 1962 but has been substantially revised over the years to accommodate changes to communications technology. At a minimum, an RS-232 connection may consist of a single wire connected between two pieces of equipment. The simplest connection in common usage contains three wires: transmit (TX), receive (RX), and ground (GND). Early RS-232 connections were commonly used to connect terminal equipment to modems [44].

It needs a driver to use with microcontroller it is called max 232 and this communication called UART – universal asynchronous receiver and transmitter. The driver is 16 pin IC which is used to interface rs232 with any other devices [45].

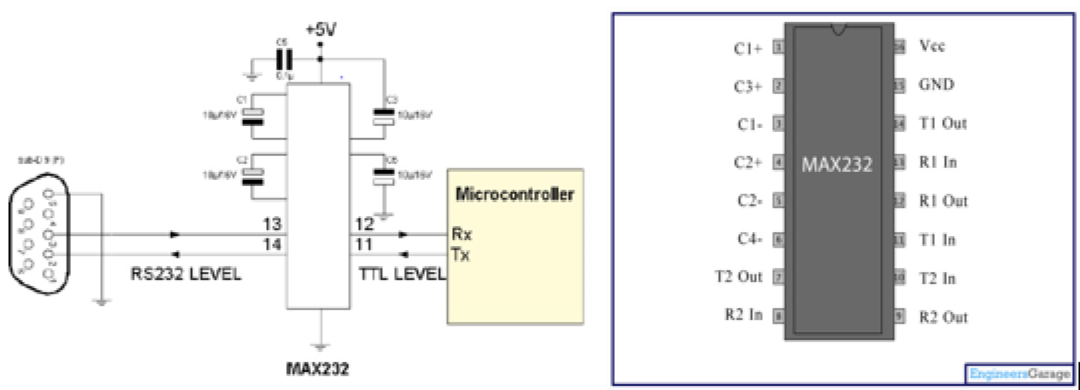


Figure 3-20 Max 232.

It uses to make interfacing between computer and hardware circuit (microcontroller) to allow send and receive data.

- Display unit consists of two LCDs, display screen for main MCU and display screen for slave MCU, it was used to display the results in different cases.

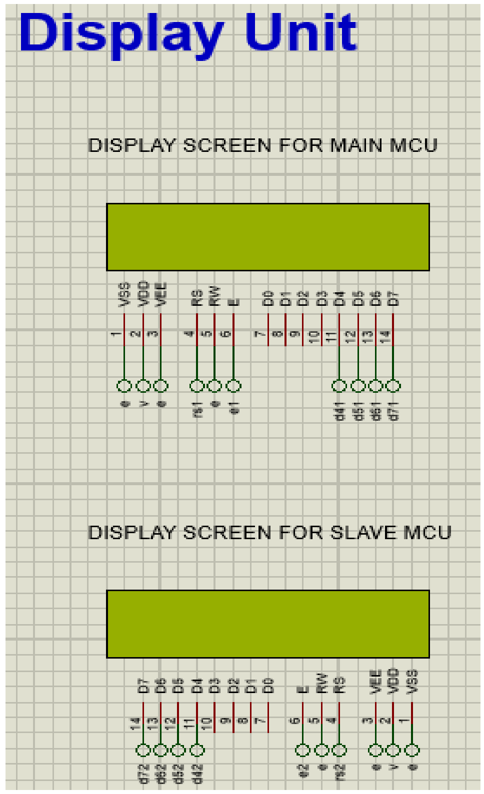


Figure 3-21 Display Unit.

3.2.3 LCD (Liquid Crystal Display)

It is a display device, Alphanumeric displays are used in a wide range of applications, including palmtop computers, word processors, photocopiers, point of sale terminals, medical instruments, cellular phones, etc. The 16 x 2 intelligent alphanumeric dot matrix display is capable of displaying 224 different characters and symbols. Use for display result [43].

Figure 3-22 LCD Display.

- Indicator unit consists of two LEDs and alarm sound, it works in verification mode. Control access LED has blue color and it lighted if person can allow to access system and connected with main MCU. Hacker indicator LED has red color and it worked if person tried to hacker and enter the system and connected with slave MCU. Alarm sound consists of buzzer which makes aloud sound if hacker was happened it connected with main MCU, it worked by supply with 12v but the voltage which out from microcontroller is 5v for this reason used l293d driver to convert 5v to 12v to work the buzzer.

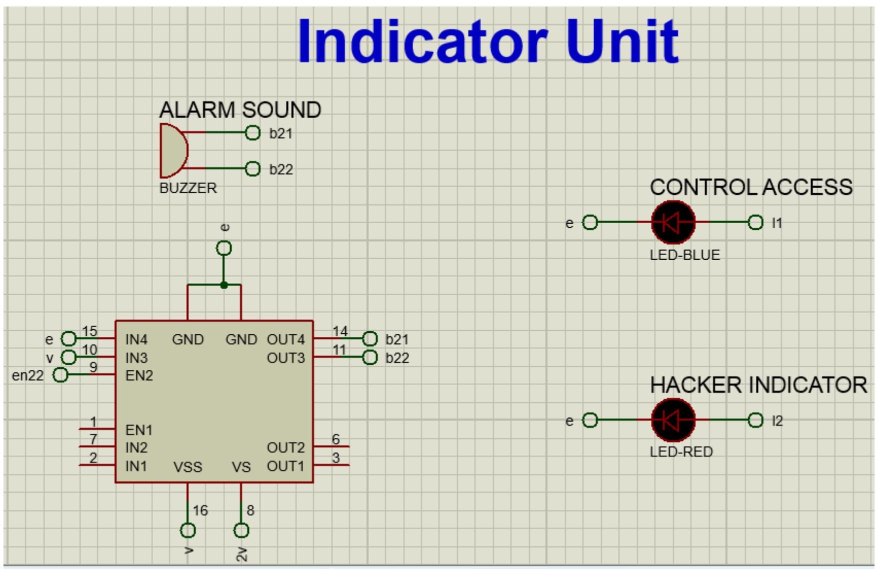


Figure 3-23 Indicator Unit.

3.2.4 Light Emitting Diodes

Commonly called LEDs, are real unsung heroes in the electronics world. They do dozens of different jobs and are found in all kinds of devices. Among other things, they form numbers on digital clocks, transmit information from remote controls, light up watches and tell you when your appliances are turned on.

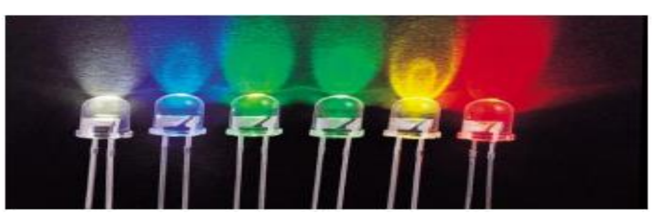


Figure 3-24 Light Emitting Diodes.

Can be color of the light (corresponding to the energy of the photon) as shown in figure 3-5 [40].

LED used to indicator when person can allowed to access to something, it makes blue color, but when the hackers tried to access will make red light and hearing alarm sounds from buzzer.

3.2.5 Buzzer

It is an electrical device, similar to a bell that makes a buzzing noise and is used for signaling. It use as alarm device when someone try to access the security system [41].

Figure 3-25 Buzzer.

- **L293D Dual H-Bridge Driver**

L293D is a dual H-Bridge driver, so with one IC we can interface two devices (here, it is buzzer). Moreover, the output supply (VCC) has a wide range from 4.5V to 36V, which has made L293D a best choice for buzzer driver [20]. The L293 comes in a standard 16-pin, dual-in line integrated circuit package as in figure 3.4. There is an L293 and an L293D part number. Pick the "D" version because it has built in freewheel diodes to minimize inductive voltage spikes [42].

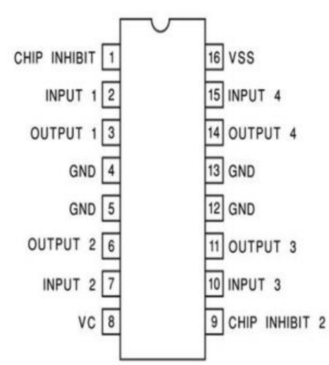


Figure 3-26 Pin of L293d.

- Manual unit consists of two switches, automatic/ manual switch that used in verification mode when emergency cases were happened (when the person which can allow to access the system died …etc), it changed mode of operation between automatic mode which depending on heart sound signature of person whose can access the system or manual mode operation which control the work of system by manual light switch. Manual light switch used to control the system manually when choose manual mode operation.

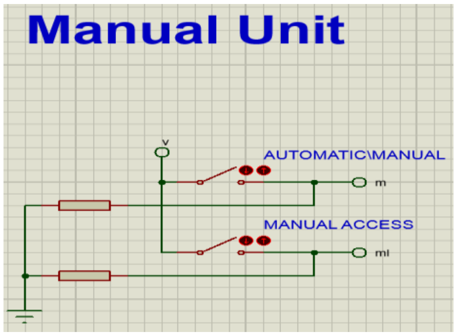


Figure 3-27 Manual Unit.

3.2.6 Switches

Several terms are used to describe switch contact.

Definition

- Pole: number of switch contact sets.

- Throw: number of conducting positions, single or double.

- Way: number of conducting positions, three or more.

- Momentary: switch returns to its normal position when released.

- Open: off position, contacts not conducting.

- Closed: on position, contacts conducting, there may be several on positions [40]

Types of switches

I. Push-to-make Single Pole, Single Throw = SPST Momentary A push-to-make switch returns to its normally open (off) position when you release the button, Shown in figure 3.8.a.

II. Push-to-make Single Pole, Double Throw = SPDT

This switch can be on in both positions, switching on a separate device in each case. It is often called a changeover switch. For example, a SPDT switch can be used to switch to do something at one position and do something else at other position as shown in figure 3.8.b.

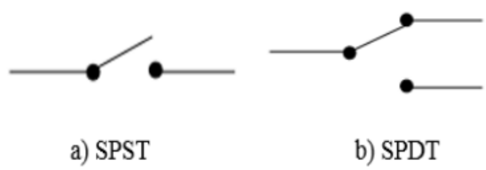


Figure 3-28 Types of Switches.

Use the switches in the circuit to control the playing of led manually or automatically by heart sounds signature and for ON/OFF of the circuit.

3.3 Simulation

3.3.1 Programs

System was simulated using Proteus 8.1 to verify the software made by Bascom 2.0.5 this will save the time and give the designer an opportunity to make any modification before test the embedded system in real life.

I. BASCOM – AVR

BASCOM-AVR is four programs in one package, it is known as an IDE (Integrated Development Environment); it includes the Program Editor, the Compiler, the Programmer and the Simulator all together [46]. For more details Appendix B.

Construction of Bascom-AVR code

i. Start by these comment

```
$regfile = "atmega16.dat" 'Bascom needs to know the-micro (Atmega16)
$crystal = 4000000 'Bascom needs to know how fast it is going 'Frequency
Config PortX = Output/Input 'make these micro pins outputs or inputs
```

ii. DO ___ LOOP

```
Do 'start of a loop

Program
Loop 'return to do and start again
```

iii. Then SUBs program:

```
Label:
Sub program

Return 'return to main program
```

II. Proteus virtual system modelling (VSM)

Proteus Virtual System Modelling (VSM) software offers the ability to co-simulate both high and low-level micro-controller code in the context mixed mode SPICE circuit simulation, animated components and microprocessor models to facilitate co-simulation of complete microcontroller based designs. For the first time ever, it is possible to develop and test such designs before a physical prototype is constructed. With this Virtual System Modelling facility, you can transform your product design cycle, reaping huge rewards in terms of reduced time to market and lower costs of development. The designer can interact with the design using on screen indicators such as LED and LCD displays and actuators such, as switches and buttons. The simulation takes place in real time. The most important feature of Proteus VSM is its ability to simulate the interaction between software running on a micro-controller and any analog or digital electronics connected to it. The micro-controller model sits on the schematic along with the other elements of pro-duct design. It simulates the execution of designer object code (machine code), just like a real chip. If the program code writes to a port, the logic levels in circuit change accordingly, and if the circuit changes the state of the processor's pins, this will be seen by the program code, just as in real systems [47].

In short, Proteus VSM improves efficiency, quality and flexibility throughout the design process [47].

III. Virtual Serial port emulator

Proteus (ISIS Design) tool is very popular platform for simulating an electronic design before making its real hardware but it's hard to get every components and its functionality in proteus. So we interface hardware components with proteus design like RFID Module, GSM Module, Bluetooth Module, Finger Print Module etc. via virtual serial port emulation (VSPE) software [48].

- **Port Emulation (VSPE)**

VSPE is use to help developers to create applications that use serial ports. It is able to create various virtual devices to transmit/receive data. One Serial port can be opened into many different applications and use their different functionality. With VSPE you are able to share physical serial port data for several applications, create virtual serial port device pairs and so on [48].

Figure 3-29 VSPE and Proteus to Simulate Serial Port.

After programming the software, make the circuit and check the simulation, the software will put in GUI form and then will be treated with the circuit from GUI.

Graphical User Interface (GUI)

A graphical user interface (GUI) is a graphical display in one or more windows containing controls, called components, which enable a user to perform interactive tasks. The user does not have to create a script or type commands at the command line to accomplish the tasks. Unlike coding programs to accomplish tasks, the user does not need to understand the details of how the tasks are performed.

3.3.2 GUI Implementation

Then the GUI program was made to be easy and usable by costumers

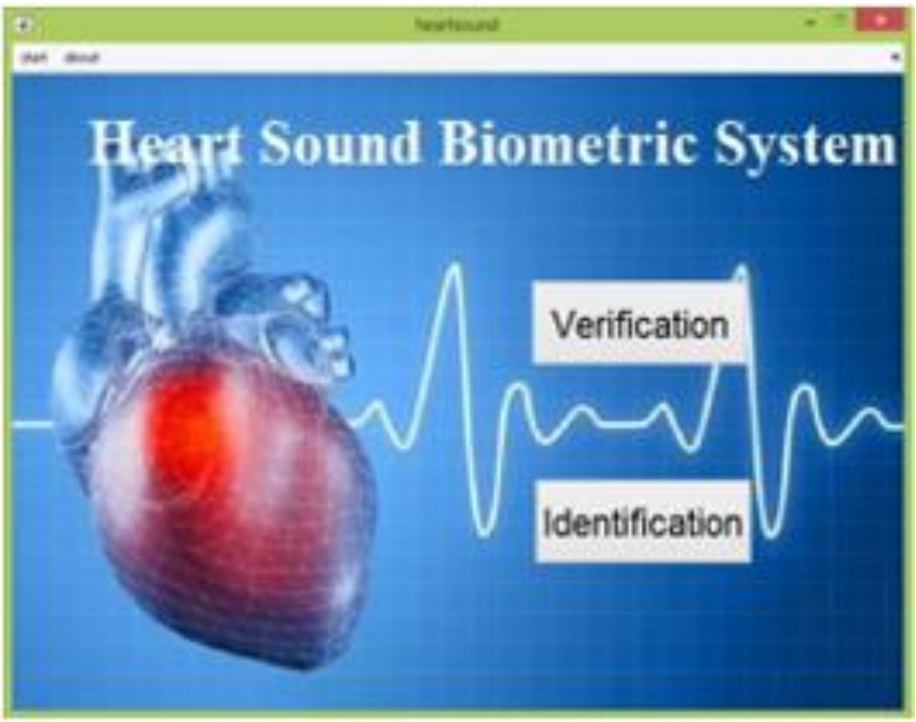


Figure 3-30 GUI of Heart Sound Biometric System.

- The Block Diagram shown in Figure below describes the block diagram of using GUI

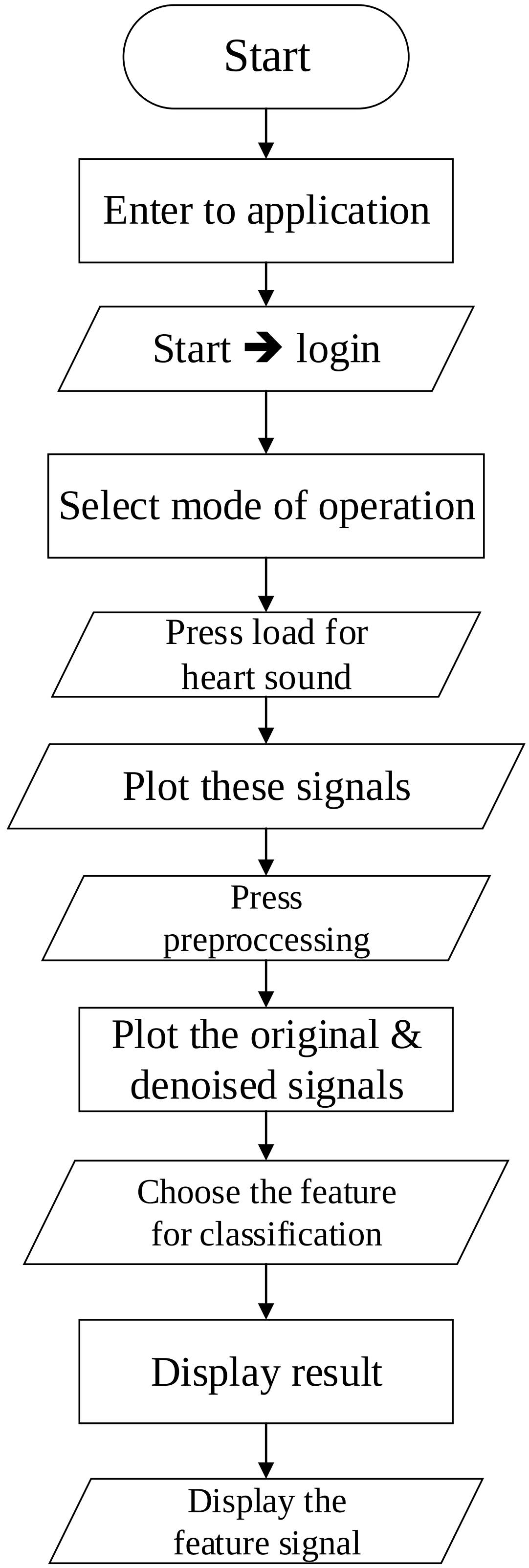


Figure 3-31 Block Diagram of GUI Sequence.

The Block Diagram shown in Figure below describes the block diagram of the identification mode.

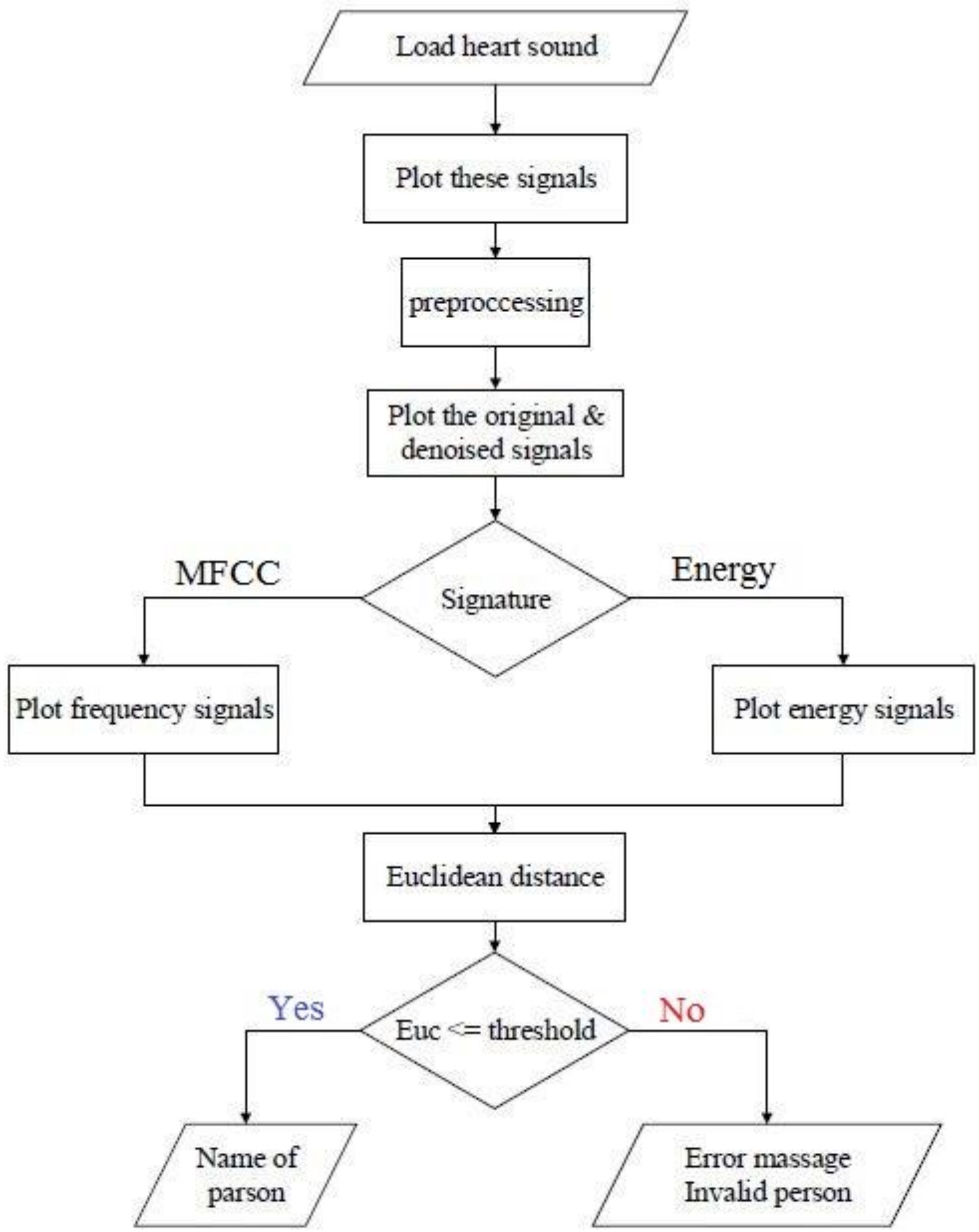


Figure 3-32 Block Diagram of Identification Mode.

The block diagram shown in Figure below describes the block diagram of the verification mode.

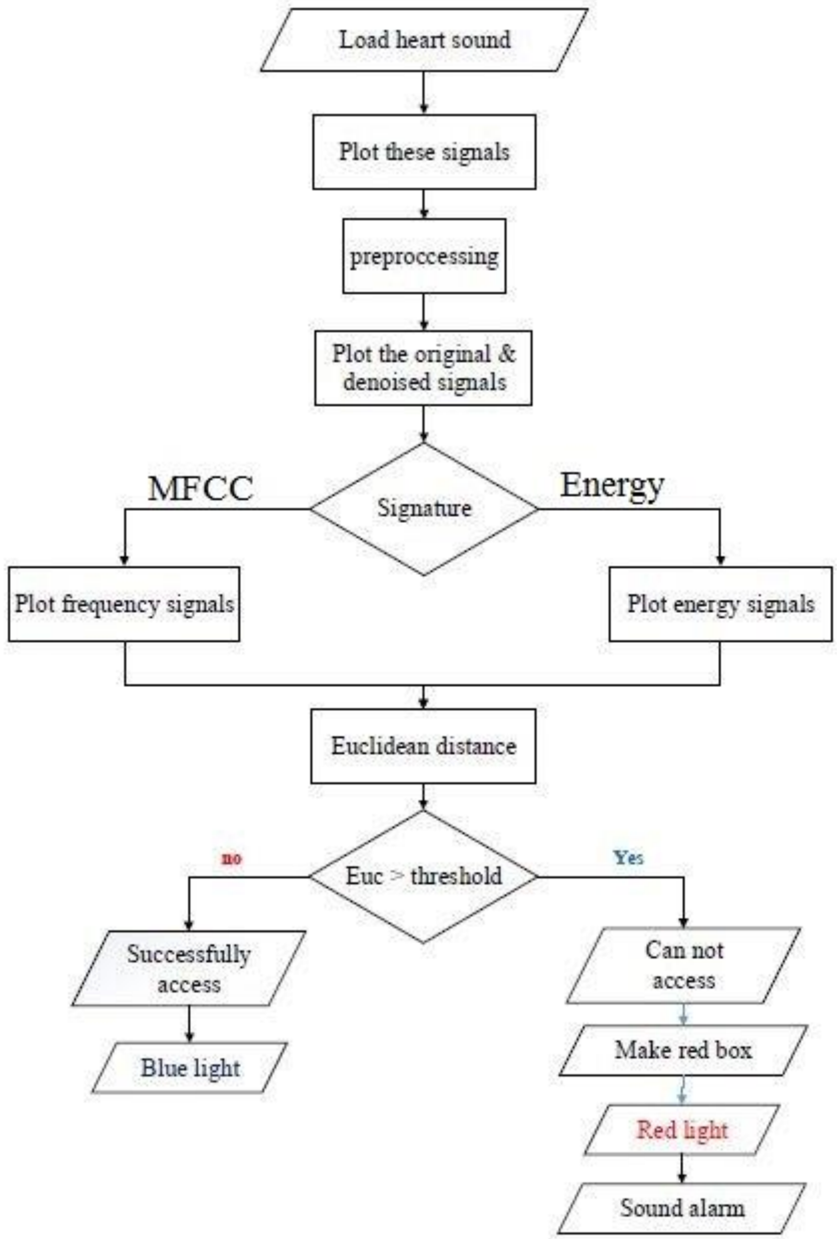


Figure 3-33 Block Diagram of Verification Mode.

The block diagram shown in Figure below describes the block diagram of the automatic/ manual mode.

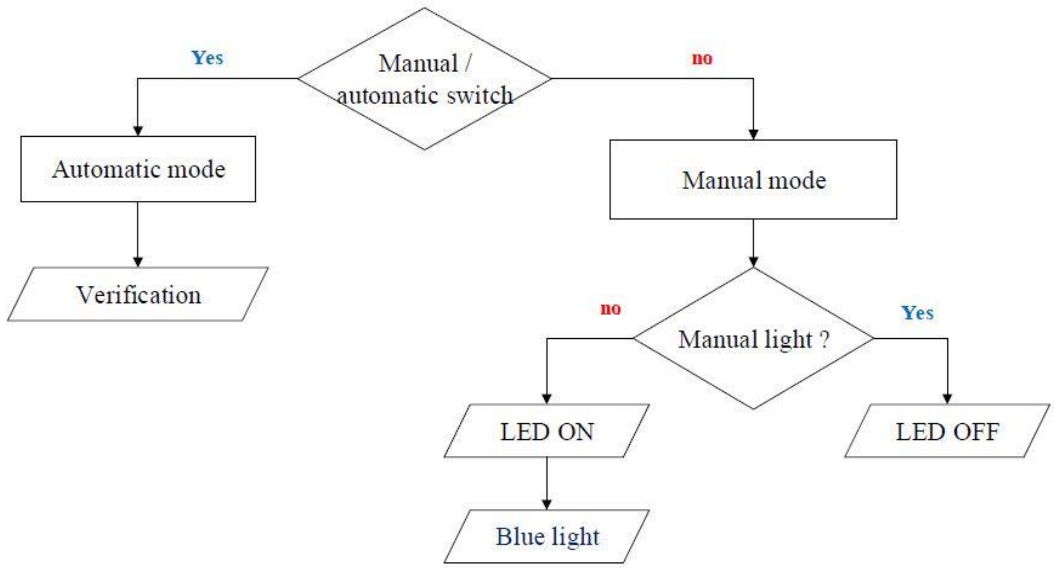


Figure 3-34 Manual / Automatic Control Mode.

Chapter Four

RESULTS AND DISCUSSION

CHAPTER FOUR

Results and Discussion

4.1 Results

Heart sound biometric system is designed and classified depending on threshold which resulted False Acceptance Rate (FAR) and False Reject Rate (FRR).

- From table4-1 and table4-2 it was noticed that the distances for different energy signatures of the same person are small when compared with different persons signature i.e; it was found that the distances for different signatures of the same person recorded at (different times) are below a threshold value and above a threshold value when compared with the signatures of other persons. From table3 it was found that the error between the same person's signatures in most cases is in the range (1-6). And this range of error is small as compared with the different person's signatures error which is above (9). Therefore, the threshold value implemented in the program was (6).

- Table 4-1 Samples of Euclidean Distance (Error) Between the Energy Signatures of The Same Person (At Different Times)

	Signature1	**Signature 2**
Person 1	0	6.178855
Person 2	0	4.804258
Person 3	0	3.418262
Person 4	5.078187	0
Person 5	4.251398	0
Person 6	0	0.868897
Person 7	3.057624	0
Person 8	0	2.522317
Person 9	0	2.109422
Person 10	0.376248	0

Table 4-2 Samples of Euclidean Distance (Error) between Energy Signatures of Different Persons

	Person 1	Person 2	Person 3	Person 4	Person 5	Person 6	Person 7	Person 8	Person 9	Person 10
Person 1	0	13.25883	20.43841	11.24915	10.7447	24.39026	17.34359	11.2471	13.11179	20.22992
Person 2	13.25883	0	25.63939	16.12119	16.54615	31.90295	26.26764	17.77539	22.15985	29.27275
Person 3	20.43841	25.63939	0	19.87729	16.31279	30.48545	26.27576	24.03225	24.64858	28.097
Person 4	11.24915	16.12119	19.87729	0	18.00674	25.08619	17.00745	16.14421	10.50641	21.18664
Person 5	10.7447	16.54615	16.31279	18.00674	0	29.43187	24.43666	20.02984	21.83759	26.63039
Person 6	24.39026	31.90295	30.48545	25.08619	29.43187	0	9.834864	24.98125	18.44474	10.316
Person 7	17.34359	26.26764	26.27576	17.00745	24.43666	9.834864	0	15.8429	8.9221	14.91084
Person 8	11.2471	17.77539	24.03225	16.14421	20.02984	24.98125	15.8429	0	10.52705	20.08713
Person 9	13.11179	22.15985	24.64858	10.50641	21.83759	18.44474	8.9221	10.52705	0	13.07577
Person 10	20.22992	29.27275	28.097	21.18664	26.63039	10.316	14.91084	20.08713	13.07577	0

- From table 4-3 and table 4-4 it was noticed that the distances for different power spectrum signatures of the same person are small when compared with different persons signature these distances were divided by 100 in classification code to make calculated and comparison easy, i.e; it was found that the distances for different signatures of the same person recorded at (different times) are below a threshold value and above a

Table 4-3 Samples of Euclidean Distance (Error) between the MFCC Feature Signatures of the Same Person (At Different Times)

	Signature1	**Signature 2**
Person 1	0	22.10924
Person 2	0	40.03968
Person 3	0	21.42619
Person 4	12.32285	0
Person 5	45.05872	0
Person 6	0	298.2824
Person 7	190.9433	0
Person 8	0	202.0956
Person 9	0	415.1519
Person 10	159.4664	0

Table 4-4 Samples of Euclidean Distance (Error) between Power Spectrum Signatures of Different Persons.

	Person 1	**Person 2**	**Person 3**	**Person 4**	**Person 5**	**Person 6**	**Person 7**	**Person 8**	**Person 9**	**Person 10**
Person 1	0	1042.4465	496.4708	667.558	900.26247	9084.79	6939.398	5822.706	9755.442	9572.443
Person 2	1042.4465	0	569.8864	606.30191	746.8988	9220.526	7076.476	5959.602	9889.407	9707.752
Person 3	496.4708	569.8864	0	857.2751	459.7874	9368.221	7224.313	6105.763	10039.28	9856.695
Person 4	667.558	606.30191	857.2751	0	1058.9289	9223.971	7081.304	5973.994	9892.772	9711.675
Person 5	900.26247	746.8988	459.7874	1058.9289	0	9121.108	6973.357	5857.232	9793.933	9609.94
Person 6	9084.79	9220.526	9368.221	9223.971	9121.108	0	2323.935	3676.224	802.2166	519.6236
Person 7	6939.398	7076.476	7224.313	7081.304	6973.357	2323.935	0	2029.642	3015.111	2795.434
Person 8	5822.706	5959.602	6105.763	5973.994	5857.232	3676.224	2029.642	0	4367.223	4130.825
Person 9	9755.442	9889.407	10039.28	9892.772	9793.933	802.2166	3015.111	4367.223	0	876.0712
Person 10	9572.443	9707.752	9856.695	9711.675	9609.94	519.6236	2795.434	4130.825	876.0712	0

threshold value when compared with the signatures of other persons. From table4-3 it was found that the error between the same person's signatures in most cases is in the range (1-3). And this range of error is small as compared with the different person's signatures error. Therefore, the threshold value implemented in the program was (4).

- The figures below shows some energy signatures

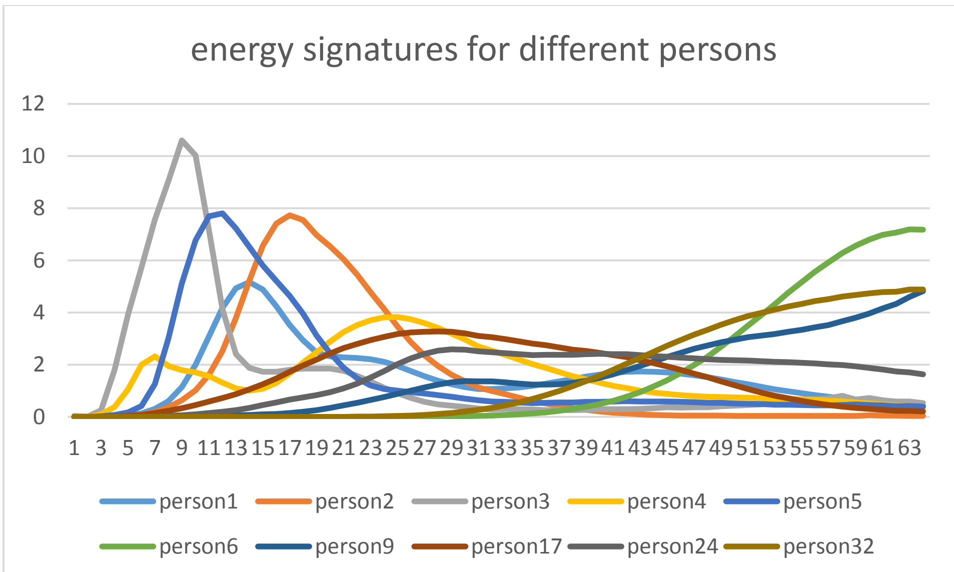


Figure 4-1 Energy Signatures for Different Persons

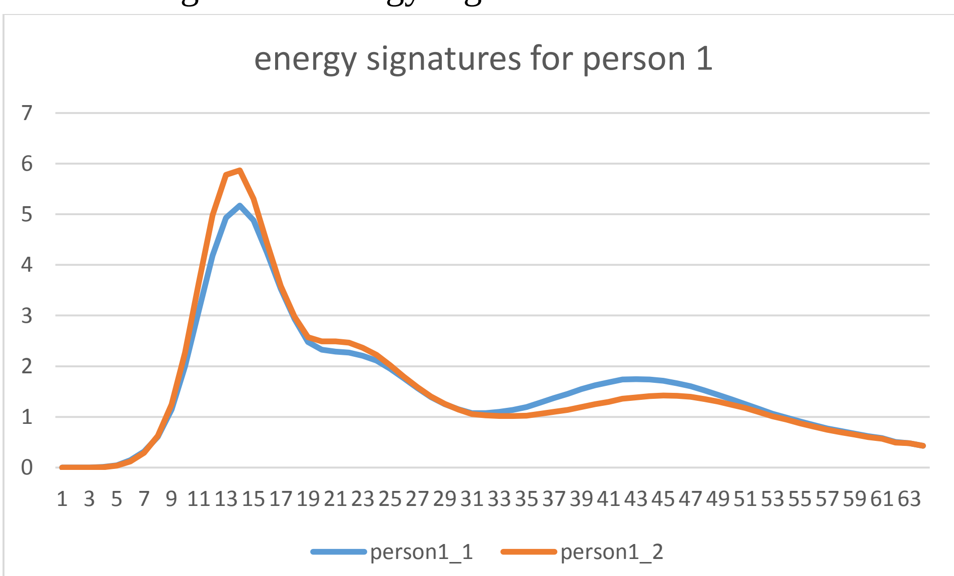


Figure 4-2 Energy Signatures for Person 1

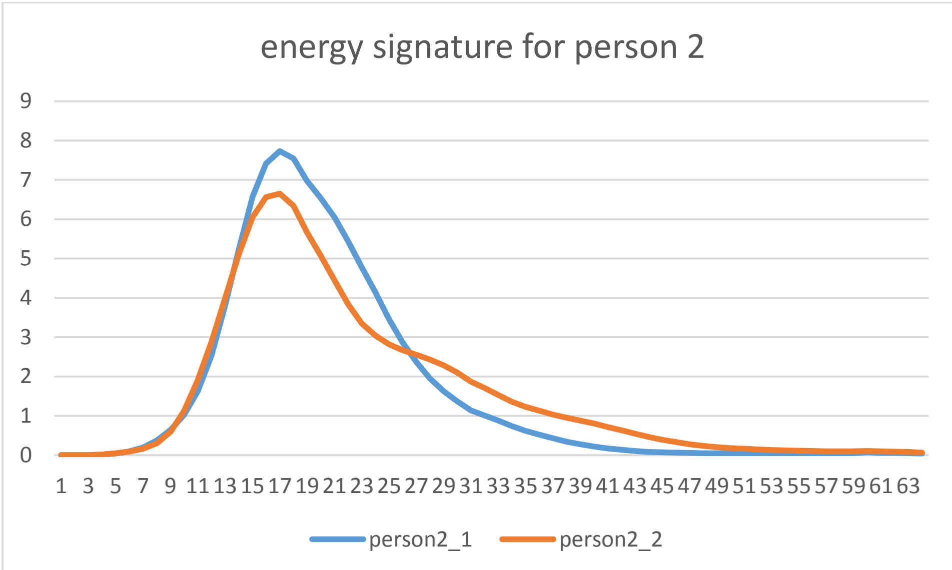


Figure 4-3 Energy Signatures for Person 2

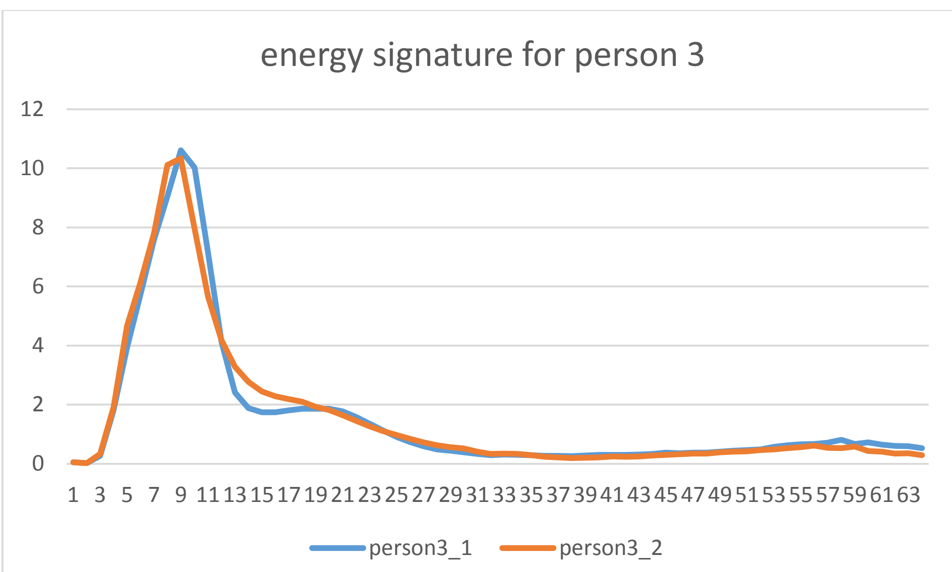


Figure 4-4 Energy Signatures for Person 3

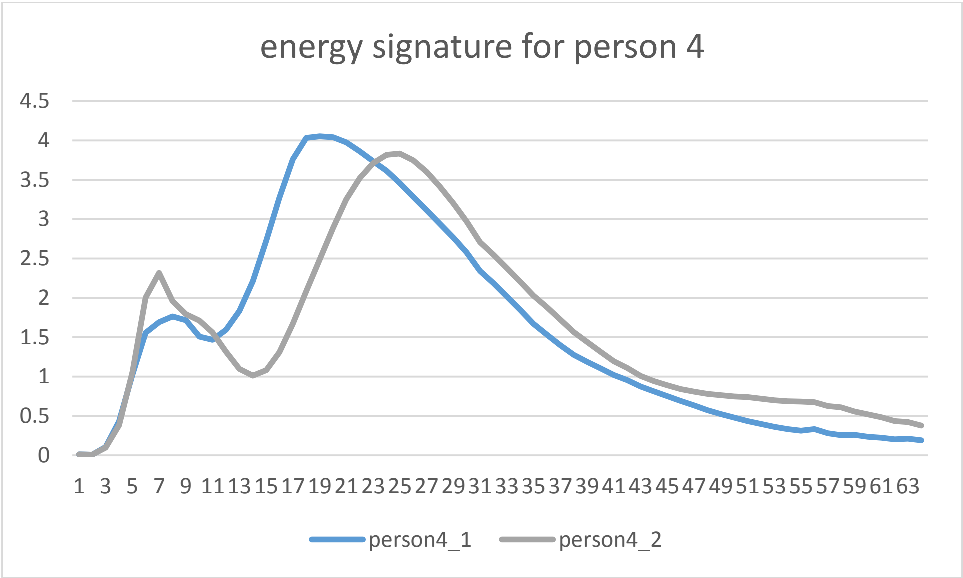


Figure 4-5 Energy Signatures for Person 4

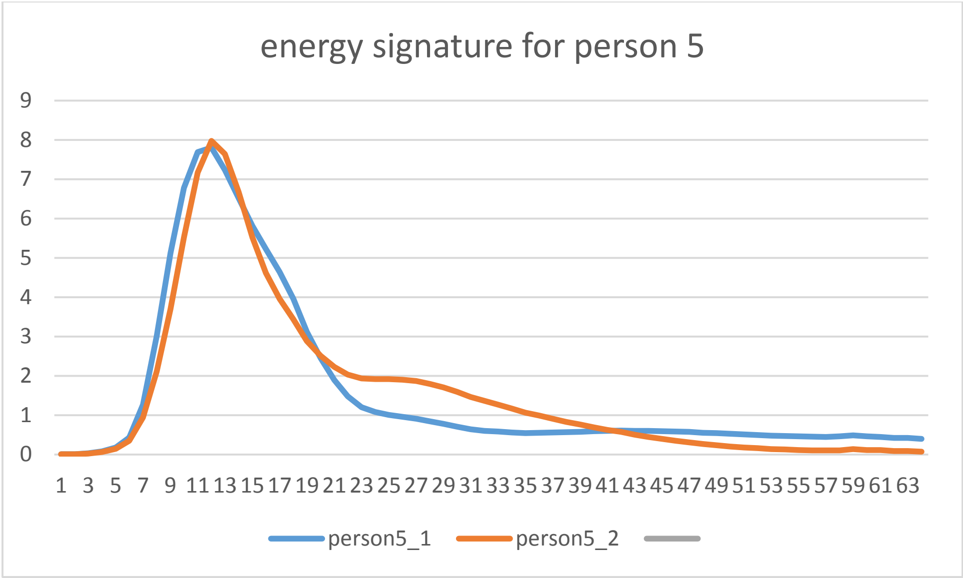


Figure 4-6 Energy Signatures for Person 5

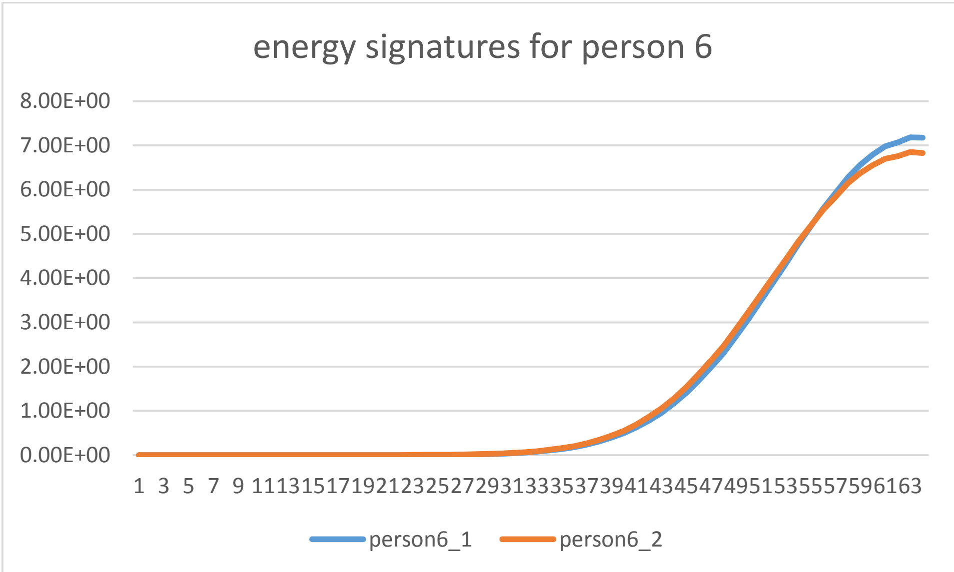


Figure 4-7 Energy Signatures for Person 6

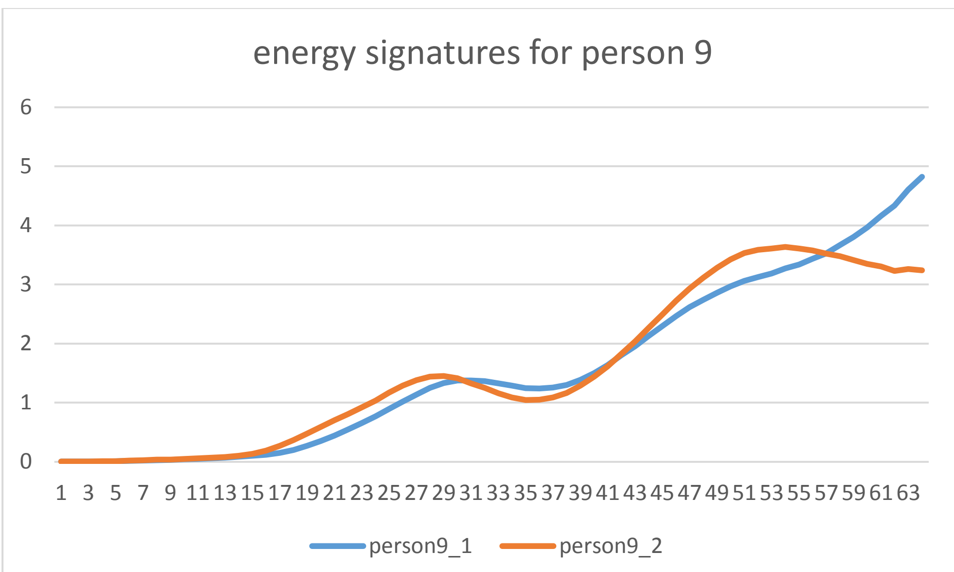


Figure 4-8 Energy Signatures for Person 9

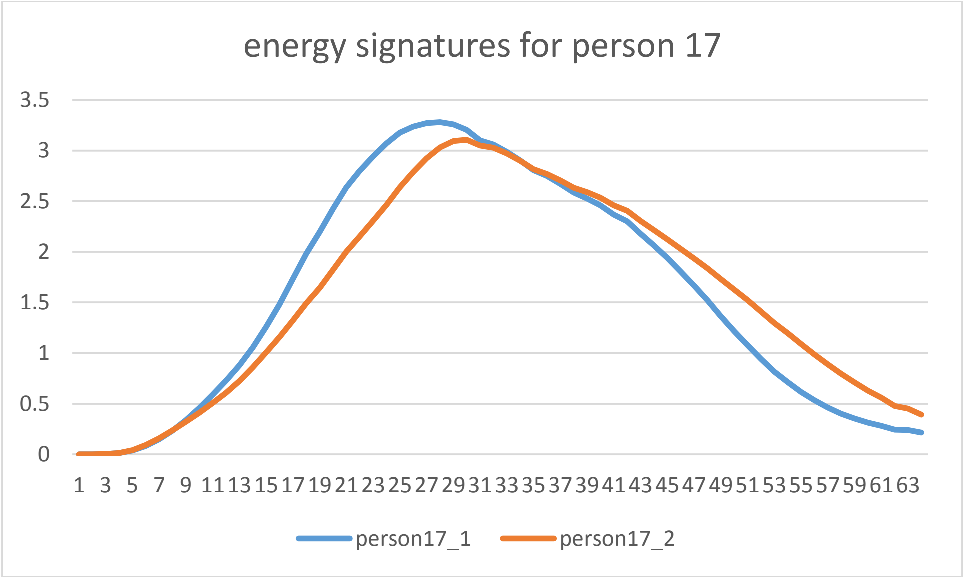


Figure 4-9 Energy Signatures for Person 17

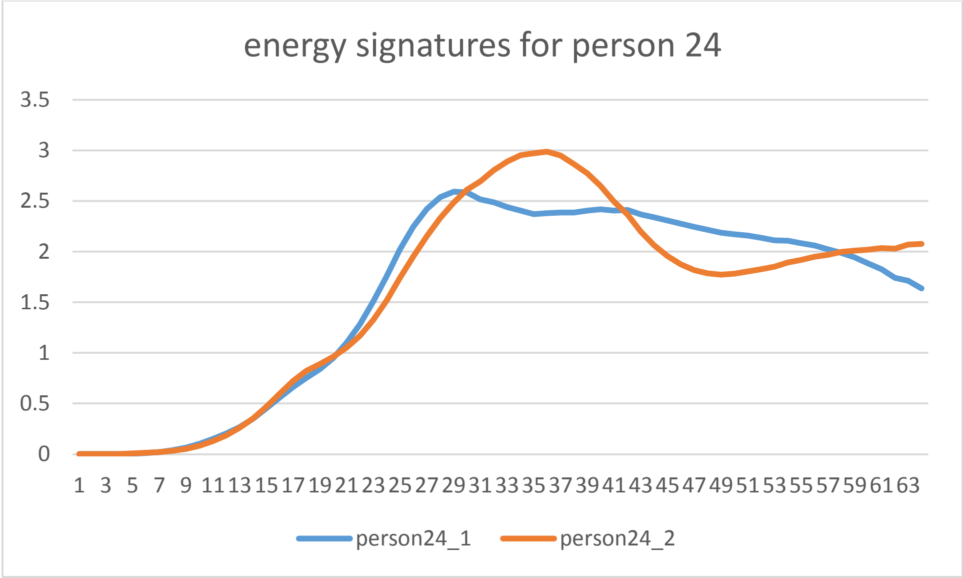


Figure 4-10 Energy Signatures for Person 24

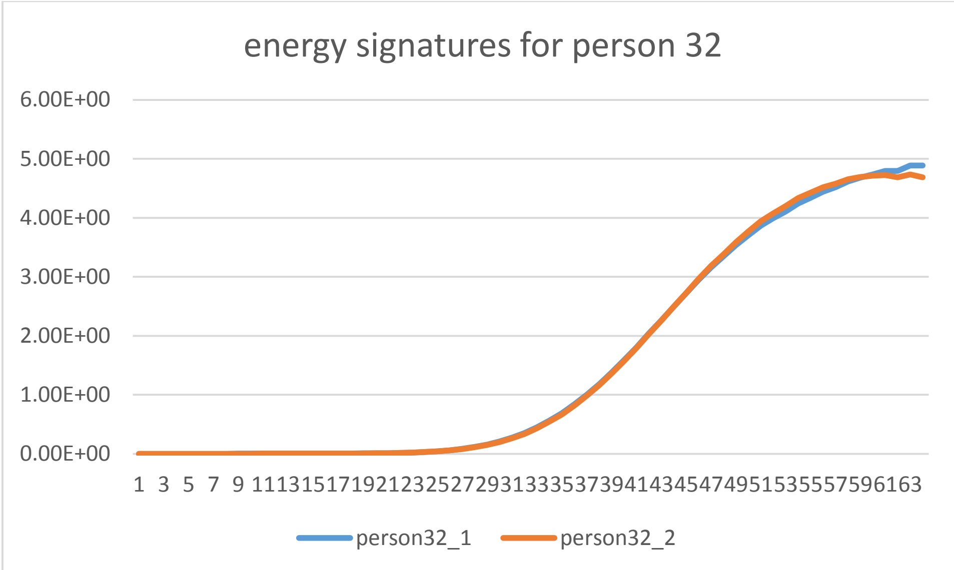


Figure 4-11 Energy Signatures for Person 32

- From the figures (4-2), (4-3), (4-4), (4-5), (4-6), (4-7),(4-8),(4-9),(4-10),(4-11), it was noticed that the shapes of the some energy signatures are similar in the same person .While in figure (4-1), there is a different in shapes between different persons.
- The figures below shows some power spectrum signatures

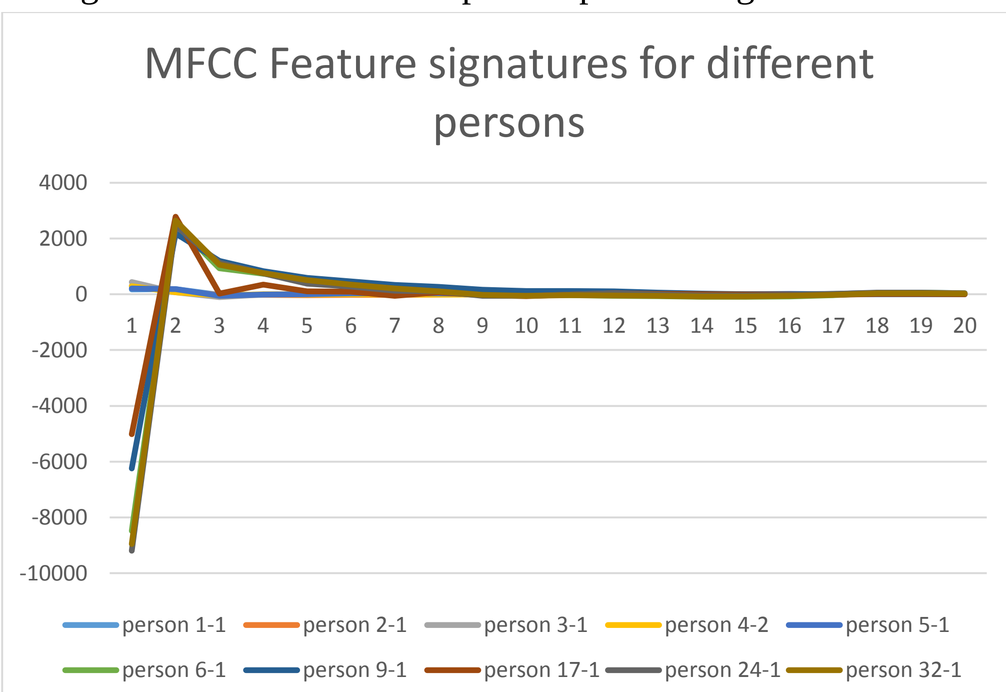


Figure 4-12 MFCC Feature Signatures for Different Persons

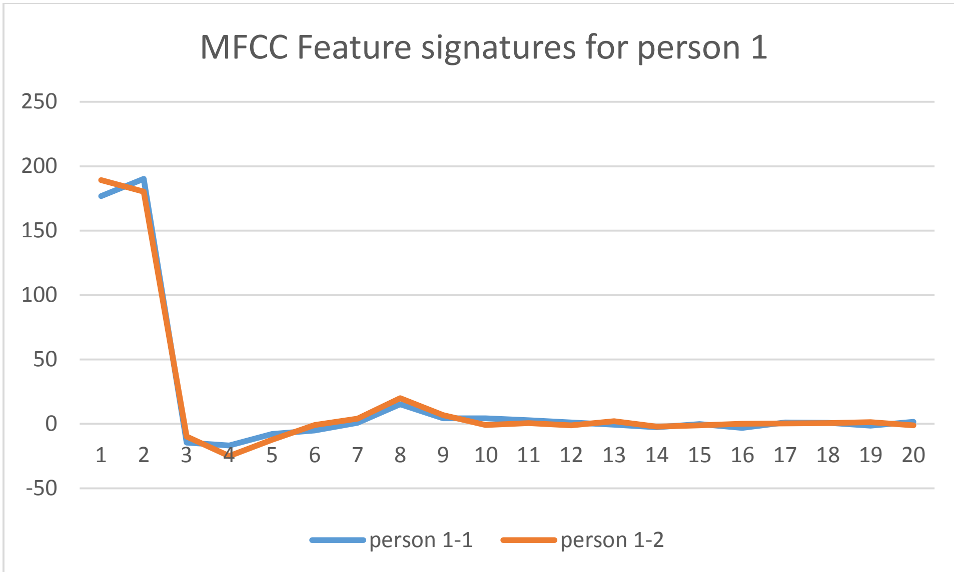


Figure 4-13 MFCC Feature Signature for Person 1 at Different Times.

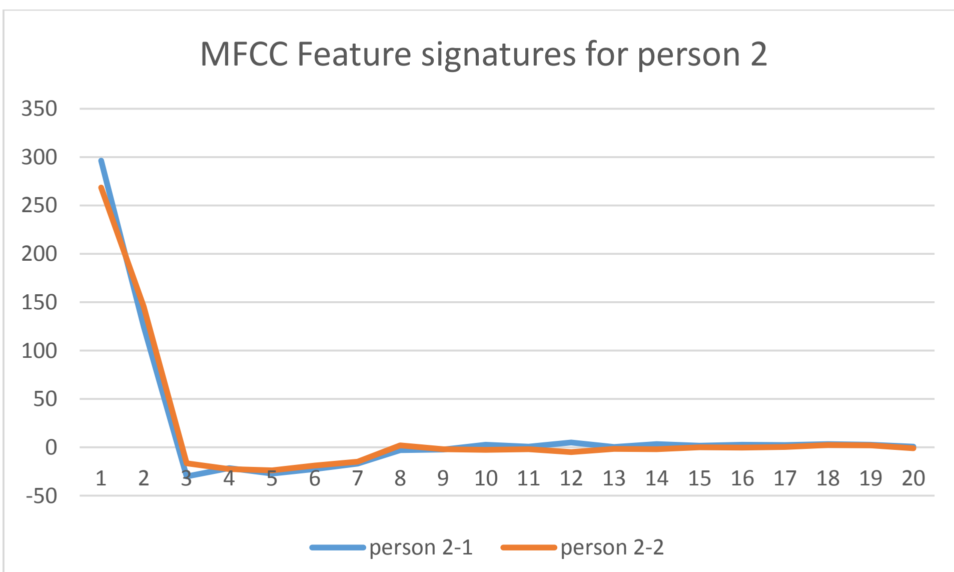


Figure 4-14 MFCC Feature Signatures for Person 2 at Different Times.

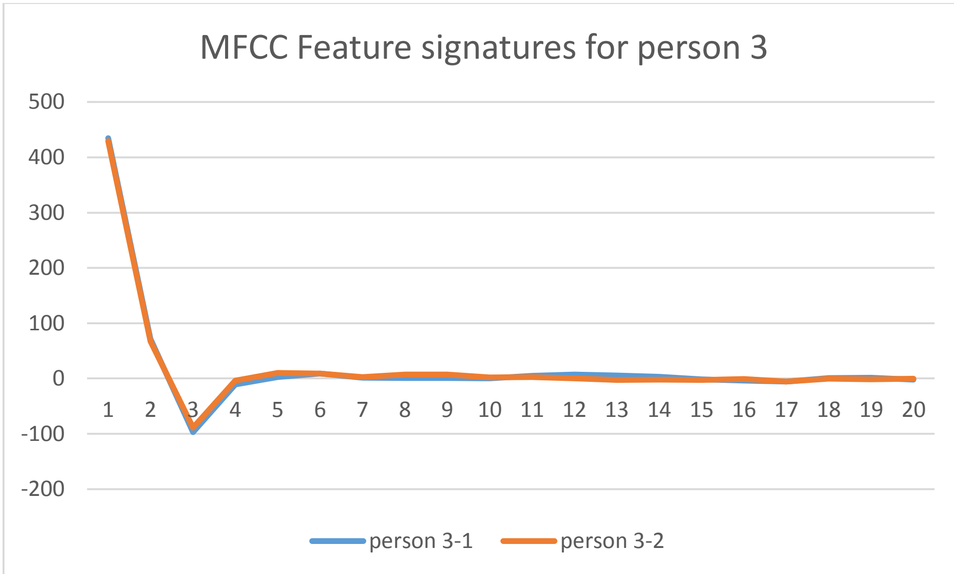


Figure 4-15 MFCC Feature Signatures for Person 3 at Different Times.

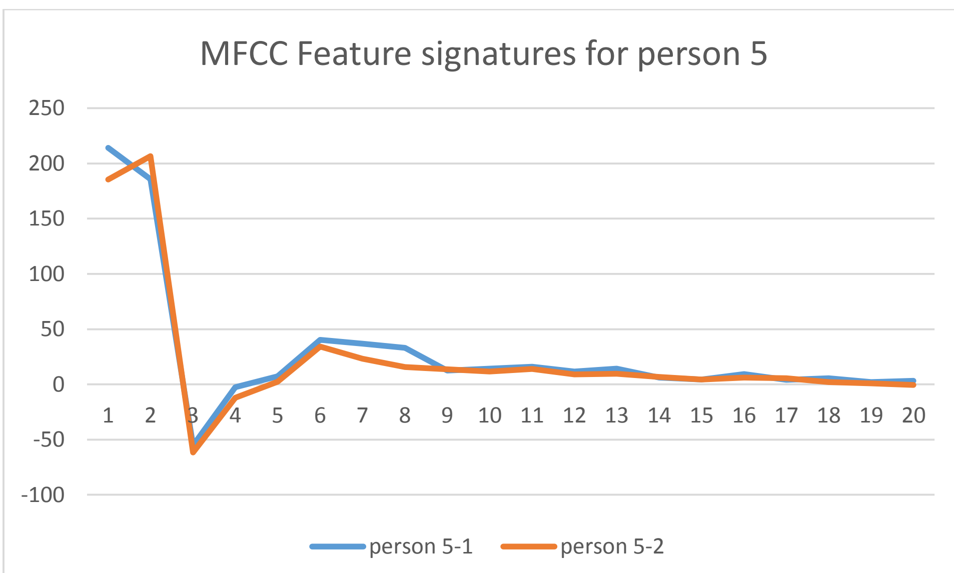


Figure 4-16 MFCC Feature Signatures for Person 5 at Different Times.

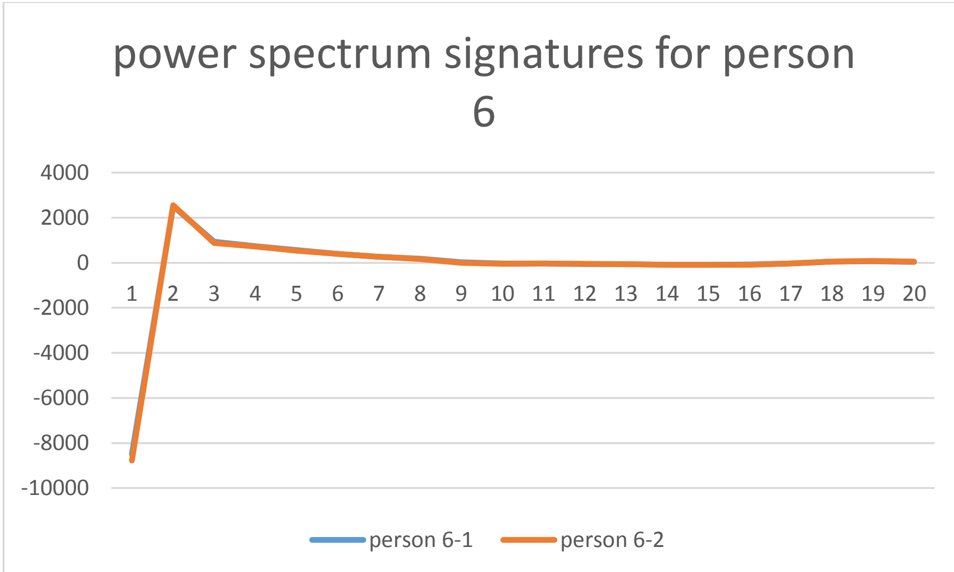


Figure 4-17 MFCC Feature Signatures for Person 6 at Different Times.

- From the figures (4-13), (4-14), (4-15), (4-16) and (4-17) it was noticed that the shapes of the some energy signatures are similar in the same person .While in figure (4-12), there is a different in shapes between different persons.

Simulation Results

- From beginning when open the GUI application this figure was seen

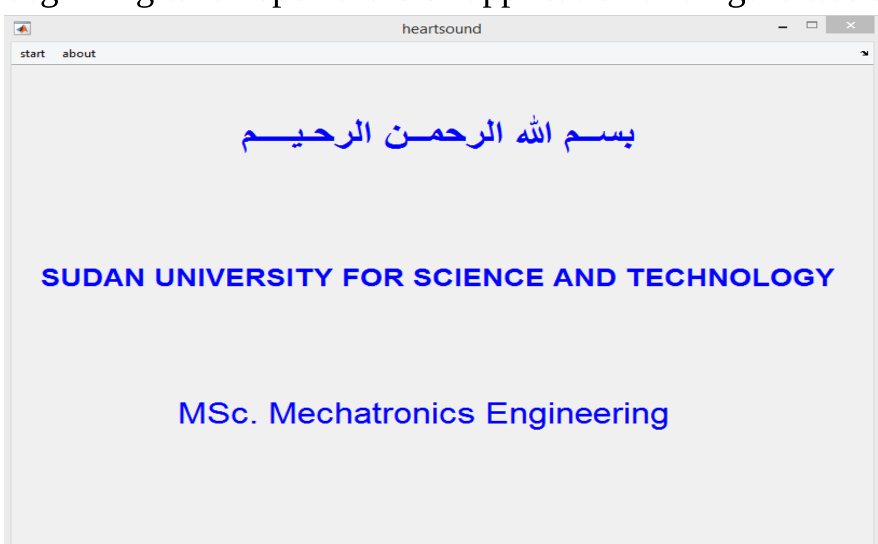


Figure 4-18 Start Running the Application.

- When the costumer want to use the application, he will be to go to start in menu bar, then he will see sub menu, choose login as see in figure below.

Figure 4-19 Login to Application

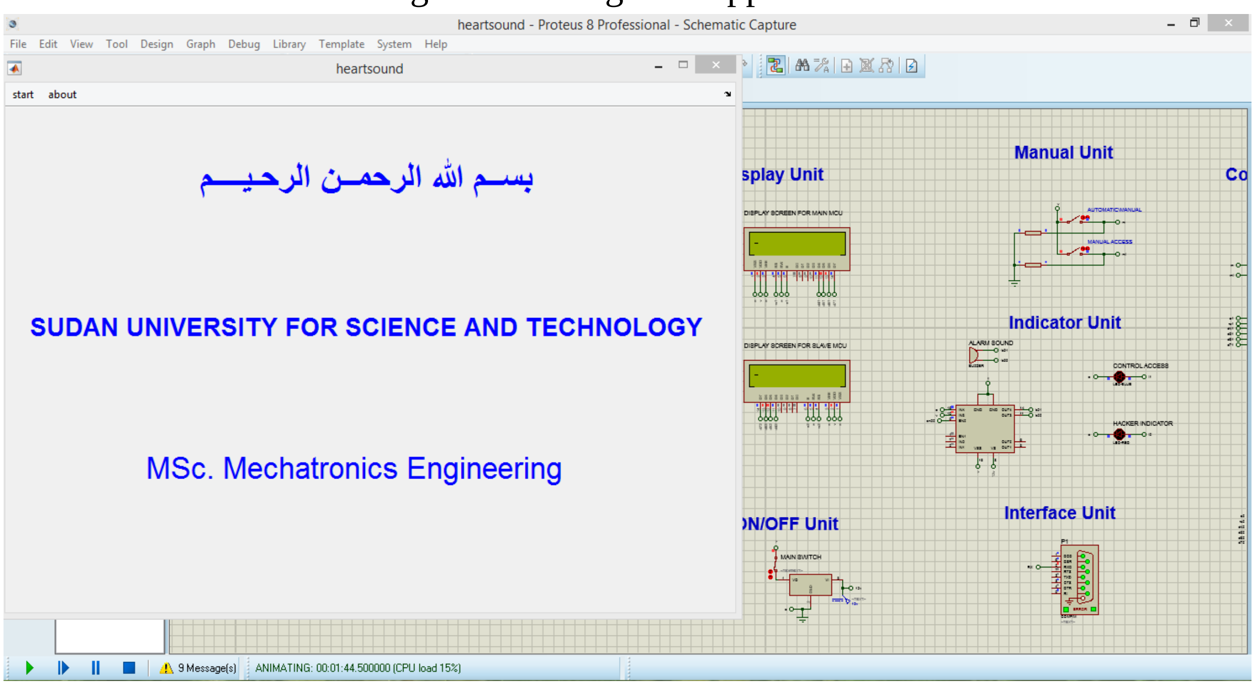


Figure 4-20 The Application and Simulation of HSBS

- Then, costumer must select the mode of operation of the system identification or verification

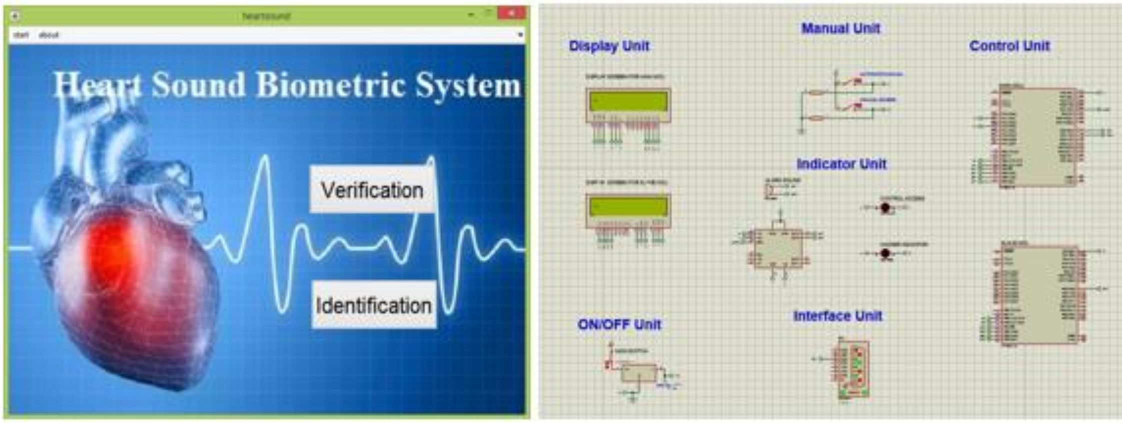


Figure 4-21 Selection of Operation Mode of the Application.

- When select the operation mode of system, this screen was appear

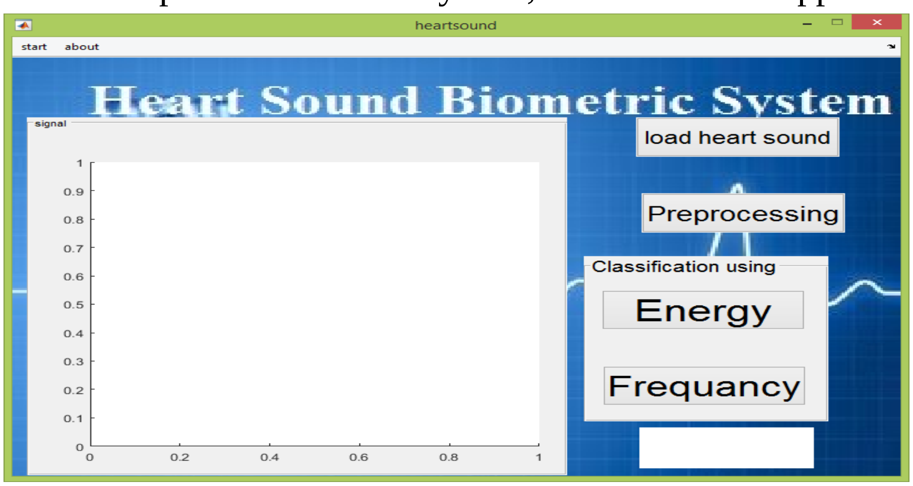


Figure 4-22 Screen of GUI after Select Operation Mode of the Application.

- Then, to load the sound, the load heart sound push button was pressed and the signal will be plot in axes.

For example, when identification mode was chosen and load heart sound was pressed, the below screens were appeared in GUI and proteus.

In proteus, LCD was displayed identification mode to obtain the mode of operation. But in GUI, when pressed load heart sound push button, windows was appeared to selected the heart sound signal and plotted this signal in axes 1.

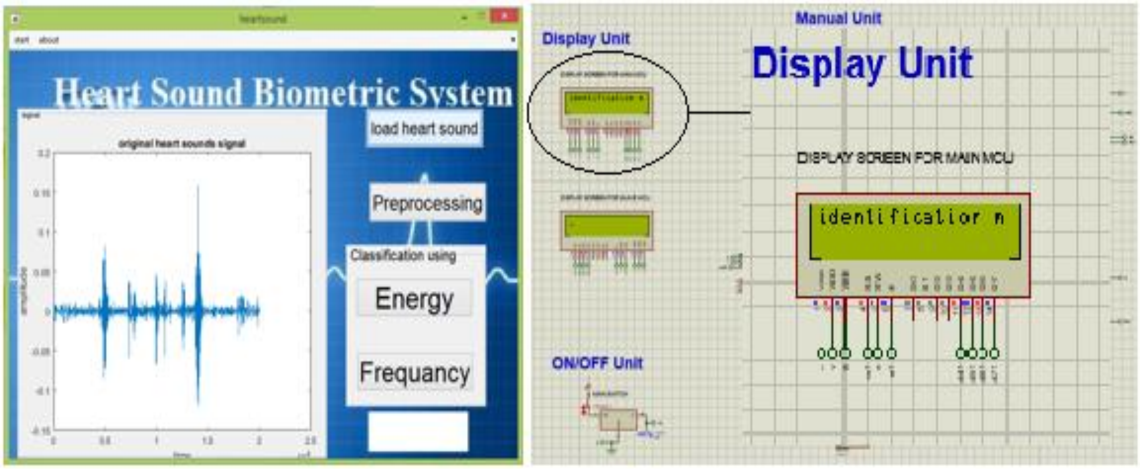


Figure 4-23 Load Heart Sound Pressed When Select Identification Mode.

But, when verification mode was chosen and load heart sound was pressed, the below screens were appeared in GUI and proteus.

In proteus, LCD was displayed verification mode to obtain the mode of operation. But in GUI, when pressed load heart sound push button, windows was appeared to selected the heart sound signal and plotted this signal in axes 1.

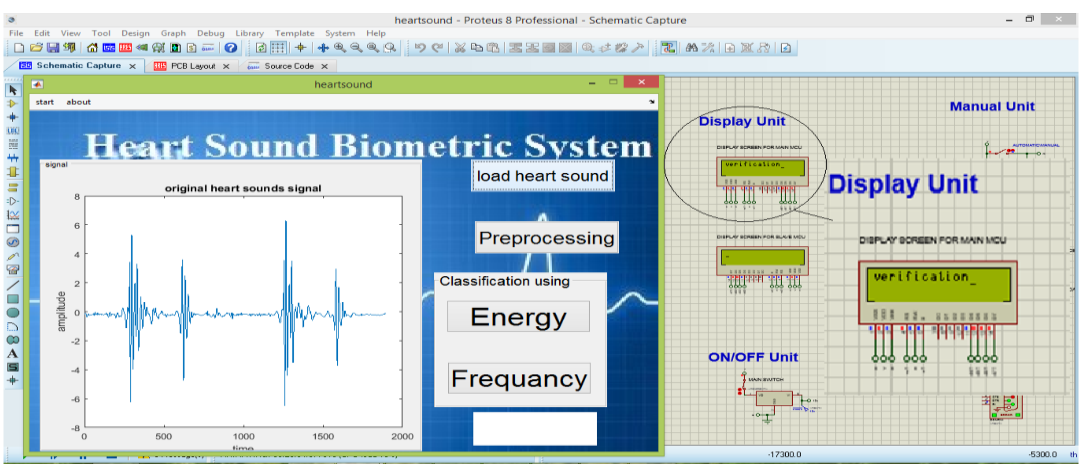


Figure 4-24 Load Heart Sound Pressed When Select Verification Mode.

- Then when preprocessing push button was pressed, the sound was passed through the preprocessing stage. It was threshold and normalized, the original heart sound signal was displayed with blue color and de-noised signal was displayed with red color, show as figure below.

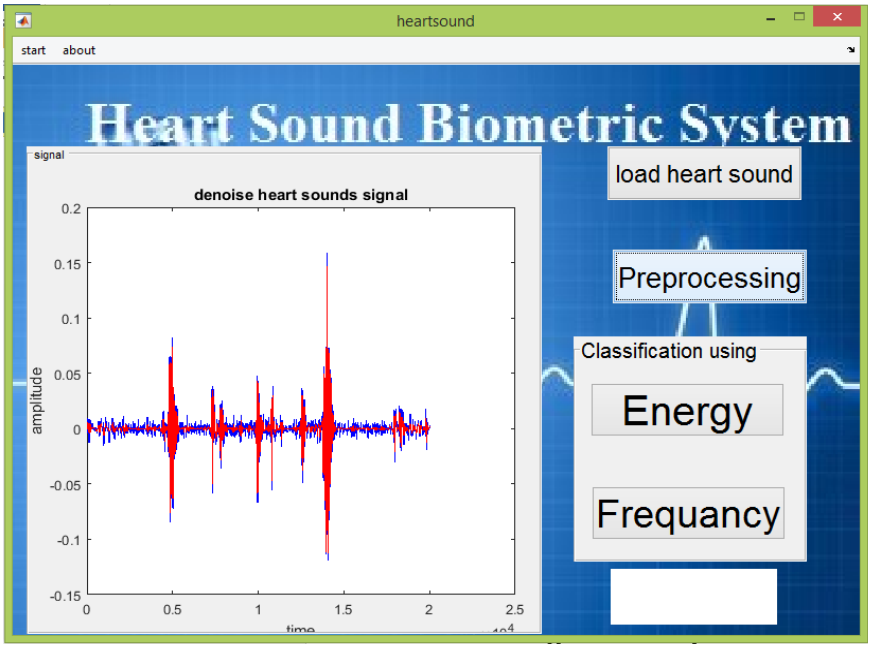


Figure 4-25 Preprocessing Pressed

- After that, the signature feature was extracted and then it was classified according of these feature, feature must select from two characteristics of signal energy or power spectrum.

For example, when identification mode was chosen and selected energy signature, the below figures was appeared in GUI and proteus.

In proteus, LCD was displayed the name of person if person find in database but if not find in database it was displayed “NO_ONE”. In GUI, axes was plotted energy signature of these heart sound signal and box was displayed the name of person if in database or invalid person if was not find in database

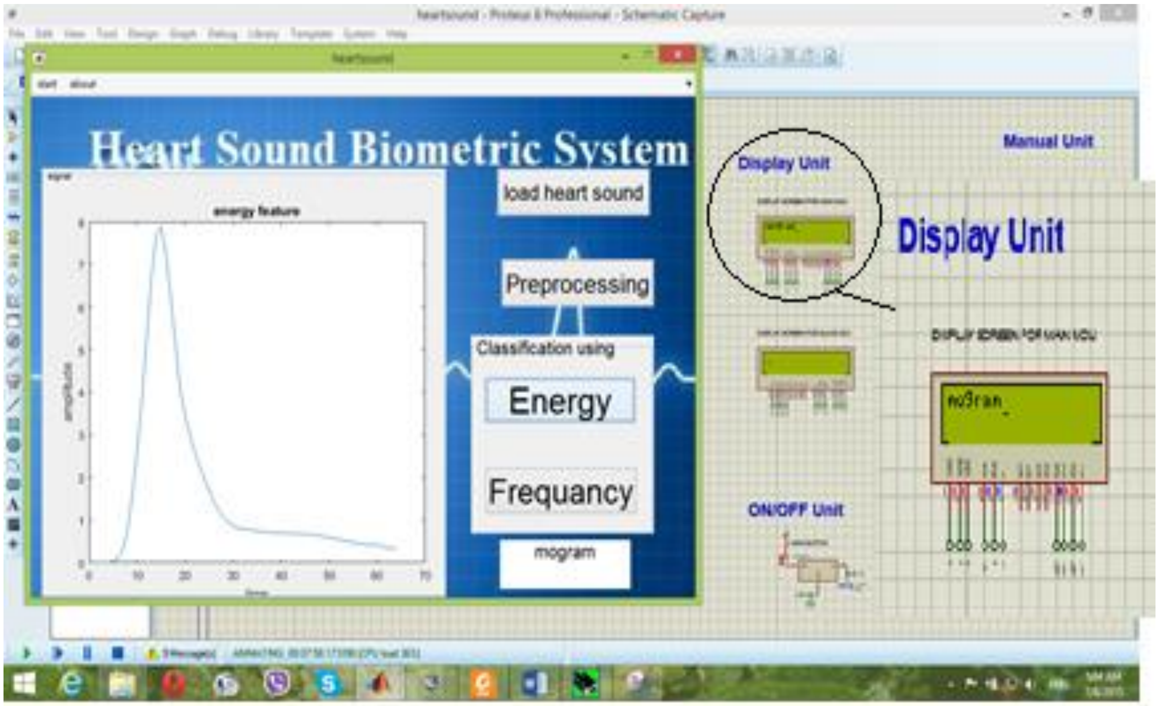


Figure 4-26 Energy Signature of Person and The System Identify Him (his name is mogram).

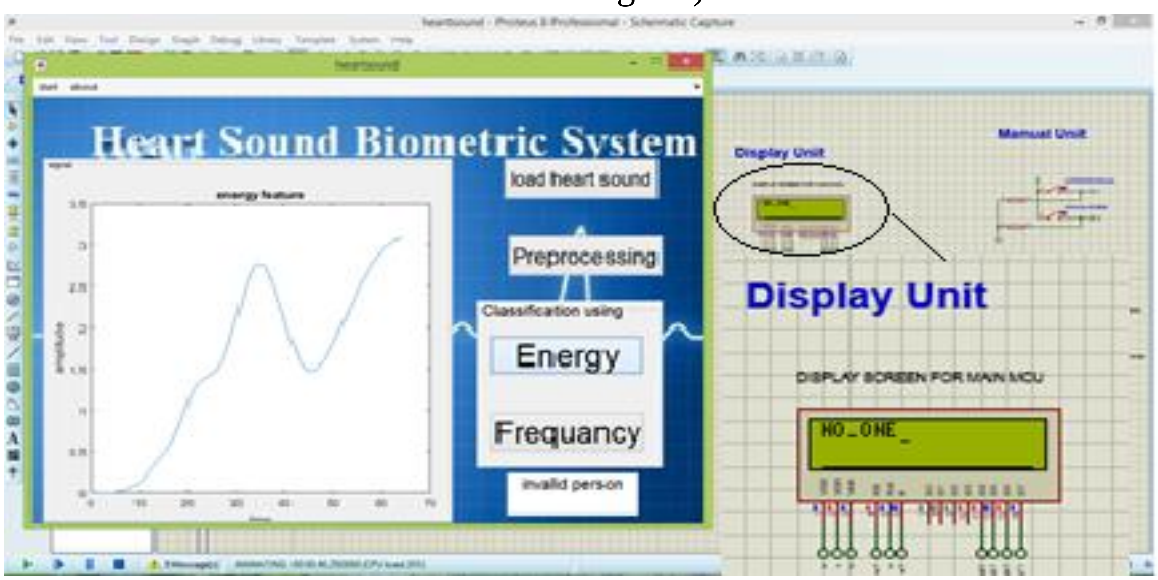


Figure 4-27 Energy Signature of Person is Not in Database of the System.

For example, when identification mode was chosen and selected power spectrum signature, the below figures was appeared in GUI and proteus.

In proteus, LCD was displayed the name of person if person find in database but if not find in database it was displayed "NO_ONE". In GUI, axes was plotted power spectrum signature of these heart sound signal and box was displayed the name of person if in database or invalid person if was not find in database.

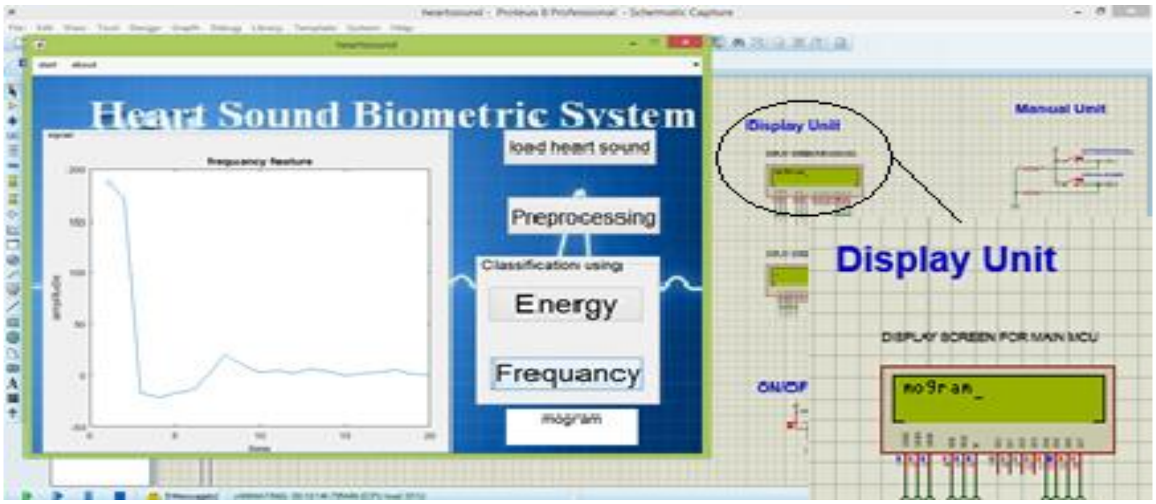


Figure 4-28 MFCC Feature Signature of Person and The System Identify Him (his name is mogram).

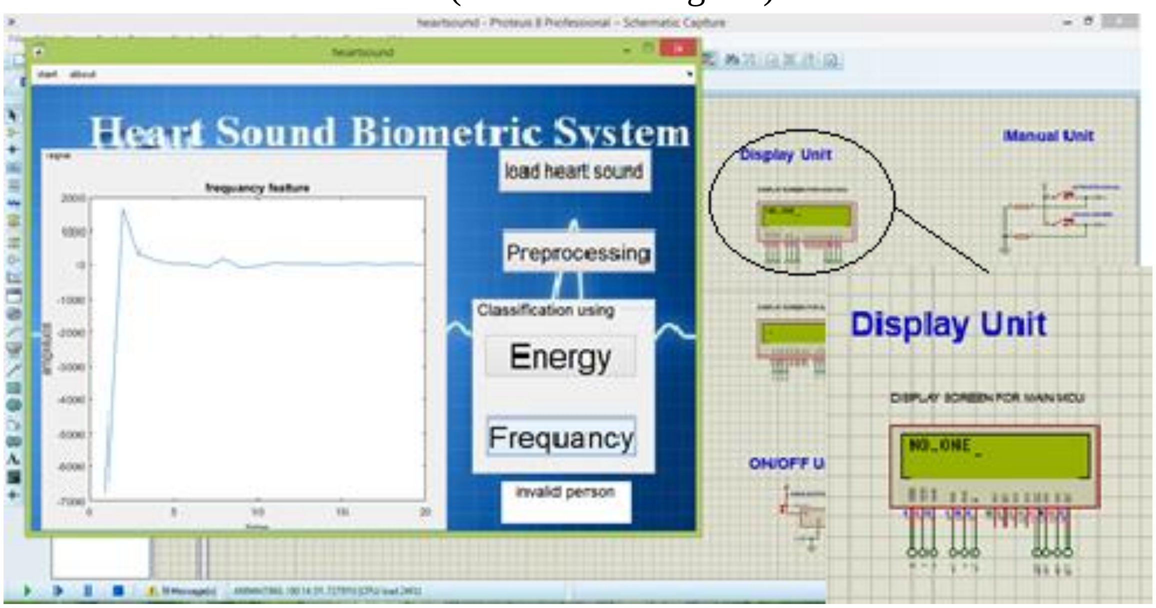


Figure 4-29 MFCC Feature Signature of Person is Not in the Database.

But, when verification mode was chosen and selected energy signature, the below figures was appeared in GUI and proteus.

In proteus, LCD was displayed “successfully access” if person can allow to access to light the LED and the LED was lighted with blue light, but if person cannot allowed to access the LCD was displayed “cannot access” and make alarm sound from buzzer and the SLAVE MCU make visible alarm by make red light from LED and their LCD was displayed “ATTENTION hacker system”. In GUI, axes was plotted energy signature of these heart sound signal and box was displayed “successfully access” if person can allow to access the system, if not box was displayed “cannot access” and colored the box with red color.

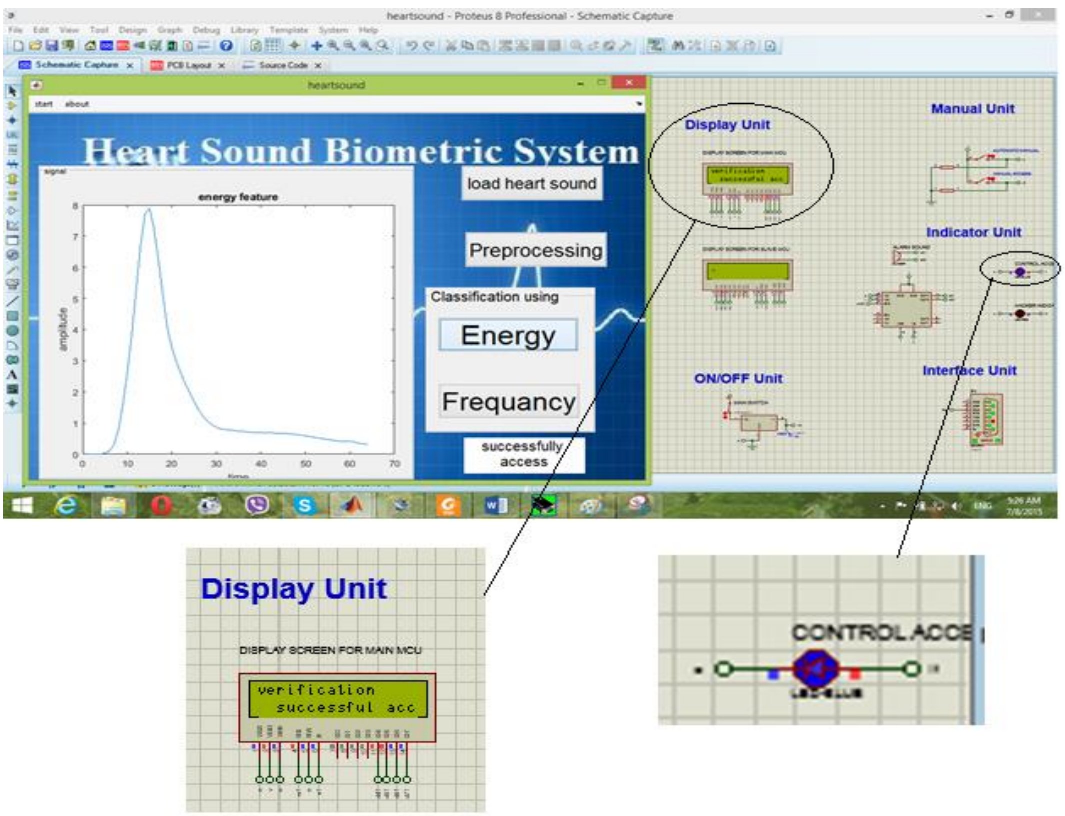


Figure 4-30 Energy Signature of Person Whose Can Allow Access in Verification Mode.

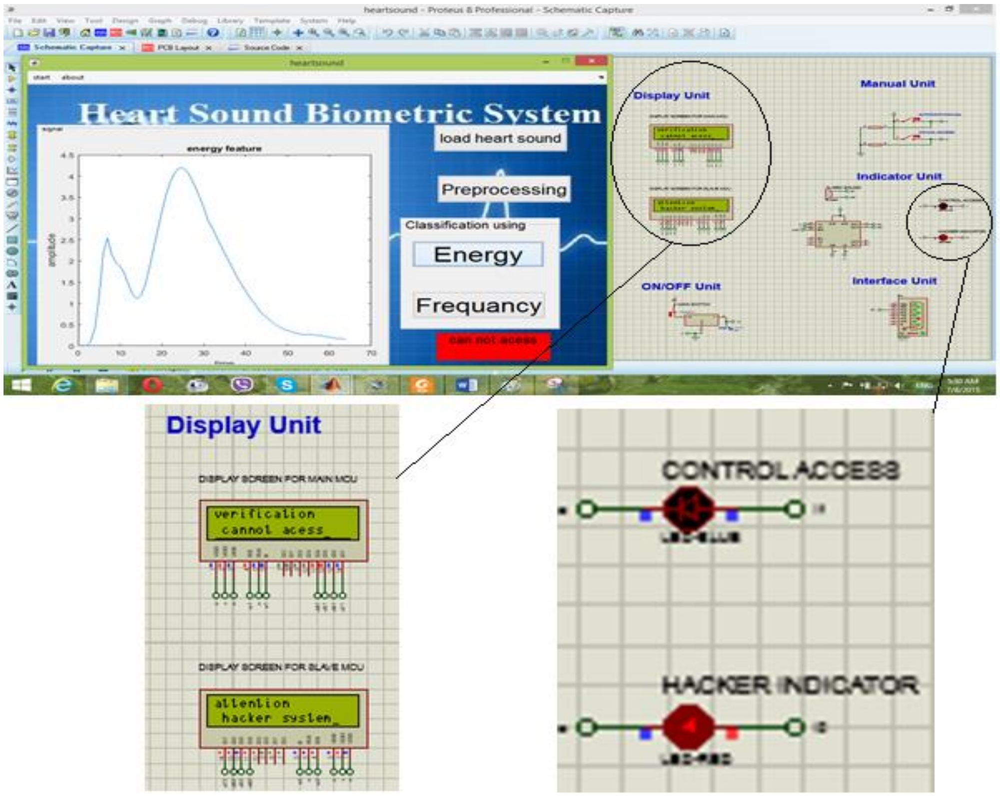


Figure 4-31 Energy Signature for Person Whose Cannot Access in Verification Mode.

But, when verification mode was chosen and selected power spectrum signature, the below figures was appeared in GUI and proteus.

In proteus, LCD was displayed “successfully access” if person can allow to access to light the LED and the LED was lighted with blue light, but if person cannot allowed to access the LCD was displayed “cannot access” and make alarm sound from buzzer and the SLAVE MCU make visible alarm by make red light from LED and their LCD was displayed “ATTENTION hacker system”. In GUI, axes was plotted power spectrum signature of these heart sound signal and box was displayed “successfully access” if person can allow to access the system, if not box was displayed “cannot access” and colored the box with red color.

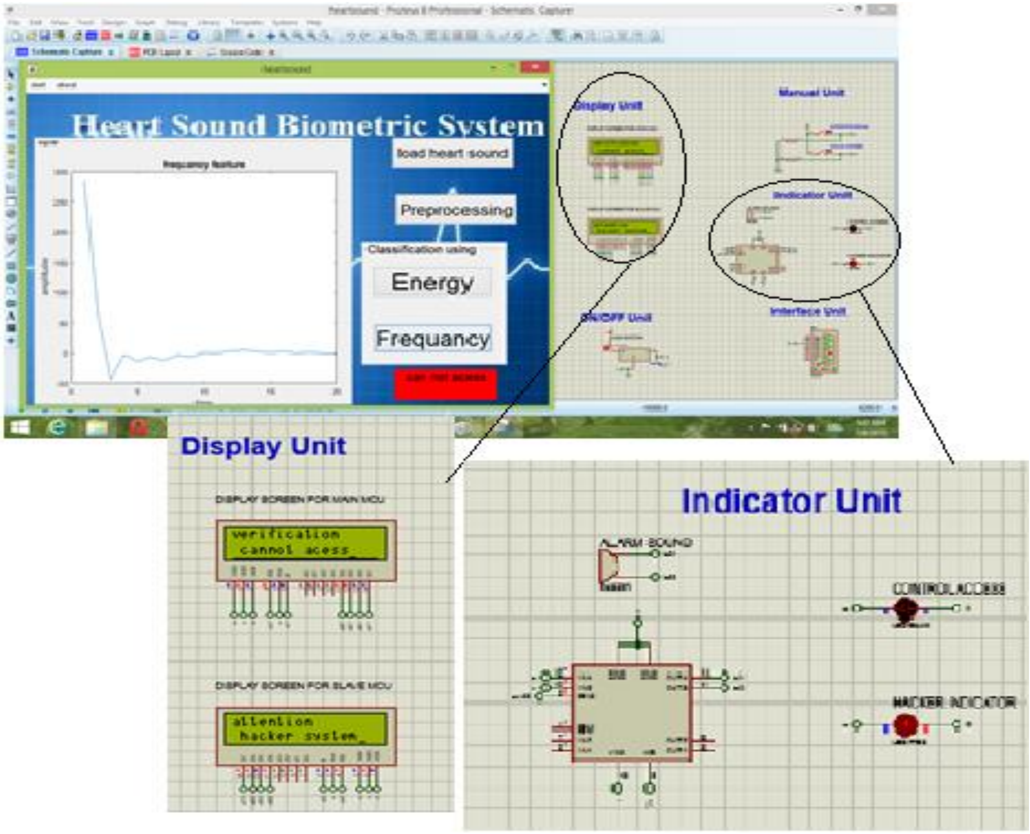


Figure 4-32 MFCC Feature Signature for Person Cannot Access in Verification Mode.

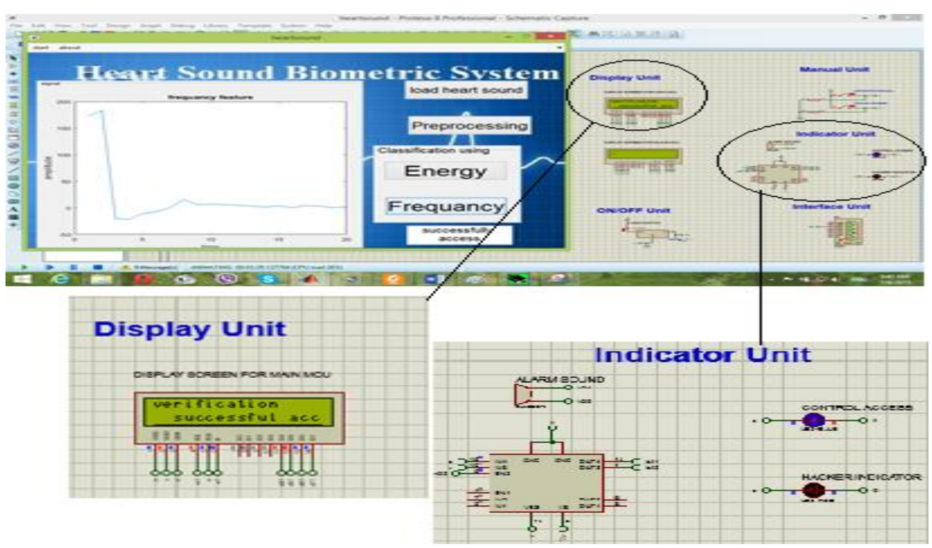


Figure 4-33 MFCC Feature Signature for Person Whose Allow to Access in Verification Mode.

In verification mode, in emergency cases, when you want to access the system the manual unit was used. The automatic / manual switch was make on, and the LCD was display manual to explain the mode of control access of the system, then control of lighting LED by manual light switch.

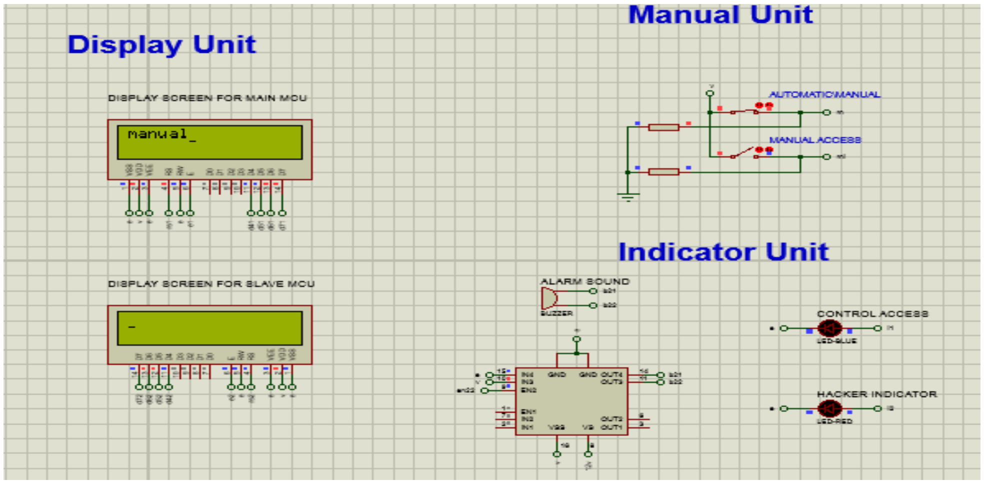


Figure 4-34 Manual Control Mode (LED is not lighted)

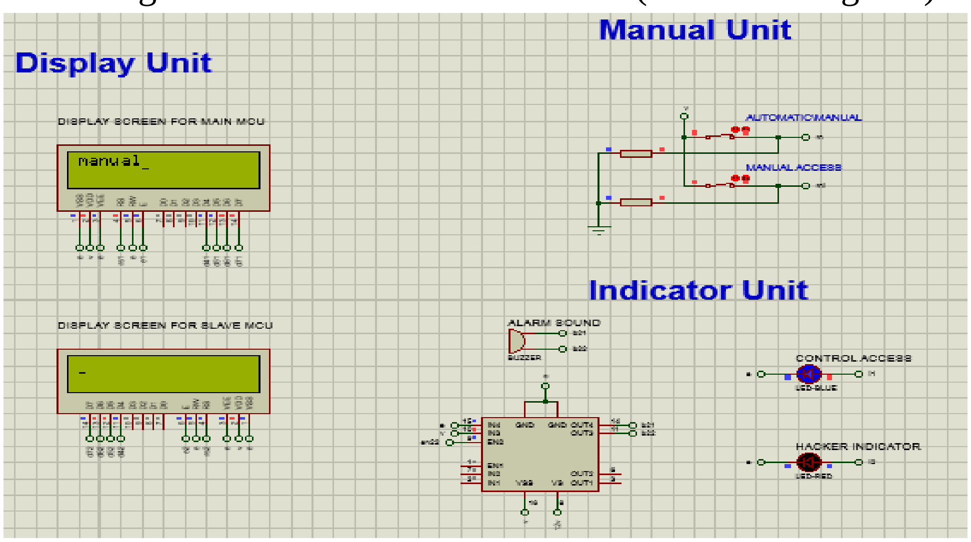


Figure 4-35 Manual Control Mode (LED is lighting)

In menu bar, about was option in it, when it was selected another windows was appeared that explain the general view of heart sound signature system as figure below

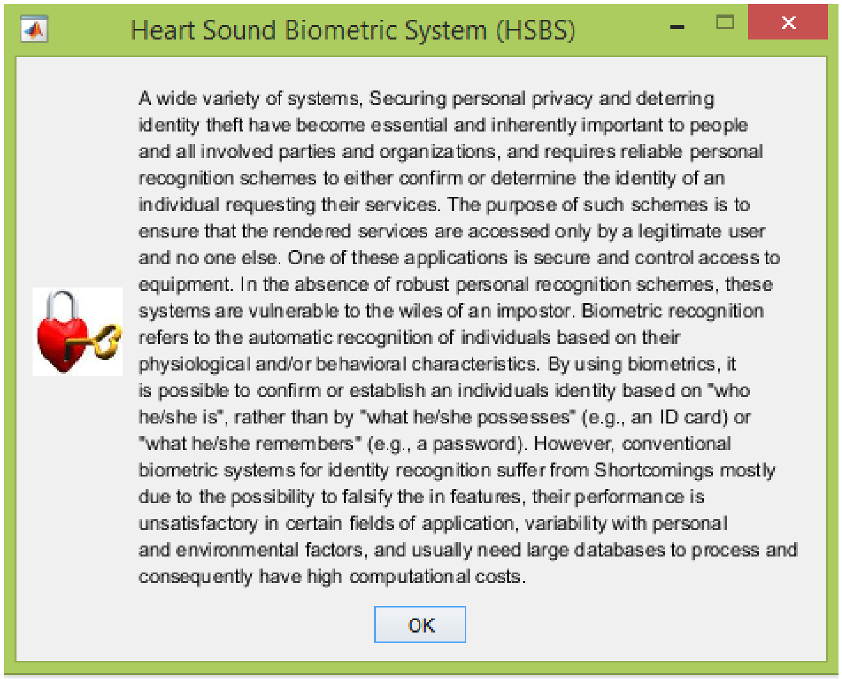


Figure 4-36 About Pressed (Information Window).

4.2 Discussion

- The algorithm performs well to identify signatures, but there were a mismatch in identification of small number of signatures, because these signatures were out of the threshold value (9) when using energy feature.

- The algorithm performs well to identify signatures, but there were a mismatch in identification of small number of signatures, because these signatures were out of the threshold value (4) when classified depend on MFCC feature.

- The algorithm performs well to verify signature, and did not have any mismatch in verification when classified depend on energy or MFCC feature.

- The percentage of error in the algorithm = (mismatch signatures/ total signatures)*100%

- Energy error= (19/85)*100%= 22.35%
- Power spectrum error=(13/85)*100%=15.29%

- From the results it was noticed that the distance between different signatures for different persons were very obvious .In reality, this is due to the variability of heart sounds as hemodynamic events such as changes in atrial pressure , ventricular pressures , ventricular volume, aortic (or pulmonary artery) pressure and ventricular volume from one person to another. And these changes depend on the person's age, gender, life style, activities, illness and psychological states.

- Also there is a slight different between the same person's signature (recorded at different times) due to the variability of heart rate which is controlled by autonomous nervous system, hormonal system, respiratory system and other mechanical or electrical factors.

Chapter Five

CONCLUSION AND RECOMMENDATIONS

CHAPTER FIVE

Conclusion and Recommendation

5.1 Conclusion

In this project, we have investigated the possibility of using heart sound as a biometric signature used for human identification and verification. The most important feature in using the heart sound as a biometric is that it cannot be easily simulated or copied, as compared to other biometrics. The main unique property of the heart sounds beyond their application for human identification is their ability to reflect information about the heart's physiology.

5.2 Recommendations

The recommendations from this project are:

1. Use a larger database of heart sounds for different persons and more heart sounds for same person in different times and after make an effort.

2. Use a reliable and trusted electronic stethoscope for sounds recording.

3. 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.

4. Use DSP processor and programming it instead of programming MATLAB and interface the circuit with pc for make these system more suitable, smaller and easy to use in their applications.

REFERENCES

1. Badghare, S.S., Analysis of. 2014, National Institute of Technology, Rourkela.

2. Fatemian, S.Z., A wavelet-based approach to electrocardiogram (ECG) and phonocardiogram (PCG) subject recognition. 2009, University of Toronto.

3. Jain, A.K., A. Ross, and S. Prabhakar, An introduction to biometric recognition. Circuits and Systems for Video Technology, IEEE Transactions on, 2004. 14(1): p. 4-20.

4. n.p. heart sounds records, http://www.diit.unict.it/hsct11,2011, [cited 4/march /2015].

5. Wilson, C., Vein pattern recognition: a privacy-enhancing biometric. 2011: CRC press.

6. Singh, A. and A.K. Singh, Biometric Identification using Phonocardio-gram. 2011.

7. Ashbourn, J., Guide to Biometrics for Large-Scale Systems. 2011: Spr-inger.

8. Sukkar, M.Y., M. Ardawi, and H. El Munshid, Concise human physiol-ogy. 2000: Wiley-Blackwell.

9. Hall, J.E. and A.C. Guyton, Textbook of medical physiology. 2011: Sa-unders London.

10. Ghali, N.I., R. Wahid, and A.E. Hassanien, Heart Sounds Human Identification and Verification Approaches using Vector Quantization and Gaussian Mixture Models. International Journal of Systems Biology and Biomedical Technologies (IJSBBT), 2012. 1(4): p. 74-87.

11. Andreeva, E., Alternative Biometric as Method of Information Securi-ty of Healthcare Systems.

12. Gautam, G., Biometric identification using analysis of cardiac sound. 2013.

13. Agrafioti, F., D. Hatzinakos, and J. Gao, Heart biometrics: Theory, methods and applications. 2011: INTECH Open Access Publisher.

14. Beritelli, F. 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. 2009. IEEE.

15. Yang, J. and N. Poh, Recent application in biometrics. 2011.

16. Phua, K., et al. Human identification using heart sound. in Second International Workshop on Multimodal User Authentication, Toulouse, France. 2006.

17. Beritelli, F. and A. Spadaccini. A statistical approach to biometric identity verification based on heart sounds. in Emerging Security Information Systems and Technologies (SECURWARE), 2010 Fourth International Conference on. 2010. IEEE.

18. Fatemian, S.Z., F. Agrafioti, and D. Hatzinakos. Heartid: Cardiac biometric recognition. in Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on. 2010. IEEE.

19. Al-Shamma, S.D. and M.C. Al-Noaemi. Heart sound as a physiological biometric signature. in Biomedical Engineering Conference (CIBEC), 2010 5th Cairo International. 2010. IEEE.

20. Karmakar, A., Biometric identification and verification based on time-frequency analysis of phonocardiogram signal. 2012, NATIONAL INSTITUTE OF TECHNOLOGY, ROURKELA.

21. Zhao, Z. and J. Wang. Heart sound identification system. in Electronics, Communications and Control (ICECC), 2011 International Conference on. 2011. IEEE.

22. Zhao, Z., Q. Shen, and F. Ren, Heart Sound Biometric System Based on Marginal Spectrum Analysis. Sensors, 2013. 13(2): p. 2530-2551.

23. Swati Verma, T.K., Analysis of Heart Sound As Biometrics Using Mfcc & Linear Svm Classifier. IJAREEIE January 2014. Vol. 3(1).

24. Tran, H.D., Y.R. Leng, and H. Li. Feature integration for heart sound biometrics. in Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on. 2010. IEEE.

25. Jasper, J. and K.R. Othman. Feature extraction for human identification based on envelogram signal analysis of cardiac sounds in time-frequency domain. in Electronics and Information Engineering (ICEIE), 2010 International Conference On. 2010. IEEE.

26. El-Bendary, N., et al. HSAS: Heart sound authentication system. in Nature and Biologically Inspired Computing (NaBIC), 2010 Second World Congress on. 2010. IEEE.

27. Shaikh Salleh, S.H., Efficient speaker verification system based on heart sound and speech. 2012.

28. Beritelli, F. and S. Serrano, Human Identification Based on Segmentation and Frequency Analysis of Cardiac Sounds. Dipartimento di Ingegneria Informatica e delle Telecomunicazioni-Universita degli Studi di Catania-viale Andrea Doria, 2007: p. 6-95125.

29. Guo, X.-M., Y. Duan, and L.-S. Zhong, Study of human identity recognition based on HMM and WNN. Jisuanji Yingyong Yanjiu, 2010. 27(12): p. 4561-4564.

30. Cheng, X., et al., Research on heart sound identification technology. Science China Information Sciences, 2012. 55(2): p. 281-292.

31. Al-Hamdani, O., et al., Multimodal Biometrics Based on Identification and Verification System. Journal of Biometrics & Biostatistics, 2013. 4.

32. Spadaccini, A. and F. Beritelli. Performance evaluation of heart sounds biometric systems on an open dataset. in Digital Signal Processing (DSP), 2013 18th International Conference on. 2013. IEEE.

33. Simmlow, J., Biosignal and biomedical image processing. Matlab based applications.

34. Ruskai, M.B., G. Beylkin, and R. Coifman, Wavelets and their Applications. Jones and Bartlett Books in Mathematics, Boston: Jones and Bartlett, 1992, edited by Ruskai, Mary B.; Beylkin, Gregory; Coifman, Ronald, 1992. 1.

35. Martinez, J., et al. Speaker recognition using Mel frequency Cepstral Coefficients (MFCC) and Vector quantization (VQ) techniques. in CONIELECOMP 2012, 22nd International Conference on Electrical Communications and Computers. 2012.

36. Beritelli, F. and A. Spadaccini, Human Identity Verification Based on Heart Sounds: Recent Advances and Future Directions. arXiv preprint arXiv:1105.4058, 2011.

37. Danielsson, P.-E., Euclidean distance mapping. Computer Graphics and image processing, 1980. 14(3): p. 227-248.

38. Barrett, S.F. and D.J. Pack, Atmel avr microcontroller primer: Programming and interfacing. Synthesis Lectures on Digital Circuits and Systems, 2007. **2**(1): p. 1-194.

39. n.p. ATMEGA16. www.atmel.com, n.d. [cited 1/February/2015].

40. Boylestad, R.L., L. Nashelsky, and F. Monssen, Electronic devices and circuit theory. 2006: Pearson Prentice Hall.

41. Stampfli, J.-M., Electronic buzzer. 1967, Google Patents.

42. Instruments, T., L293D datasheet," Quadruple Half-H Drives. 2010.

43. LTV350QV, S.T.-L., F04 Datasheet. Samsung Electronics Co, Ltd, 2006.

44. Strangio, C.E., the RS232 standard. Online Document, 2006. 10.

45. n.p., communication ports RS232, Interfacing the Serial - RS-232 Port.htm, n.d., [cited 21/May/2015].

46. Kuhnel, C., BASCOM Programming of microcontrollers with ease: An introduction by program examples. 2001: Universal-Publishers.

47. Levin, G. and L. Nyland, An introduction to Proteus, version 0.9. UNC-CH, Technical TR95-025, August, 1993.

48. Sharma, V., Virtual Serial Port Emulator, VSPE ~2embeddedroboti-cs.htm, April/ 2014. [Cited 21/May/2015].

49. MatLab, M., The language of technical computing. The MathWorks, Inc. http://www. mathworks. Com, 2012.

50. Logan, B. Mel Frequency Cepstral Coefficients for Music Modeling. in ISMIR. 2000.

51. Davis, S.M., Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1980. Vol. 28(No. 4).

Appendices

APPENDIX A

Discrete Wavelet Transform

The Wavelet Transform is a signal processing tool which provides time-frequency representation of a signal. It was developed to overcome one of the major limitations of Fourier Transform – the lack of time resolution.

In DWT, the signal in study is processed by a sequence of filters, so it can be decomposed in several frequency bands. The process starts by simultaneously filtering the signal with two half band low-pass and high-pass filters and divided the frequency range by half for each band. The output of the high-pass filter, followed by down sampling, is known as level 1 detail coefficients (cD1), and the output of low-pass filter is called level 1 approximation coefficients (cA1) [34]. As represented in the figure A.1.

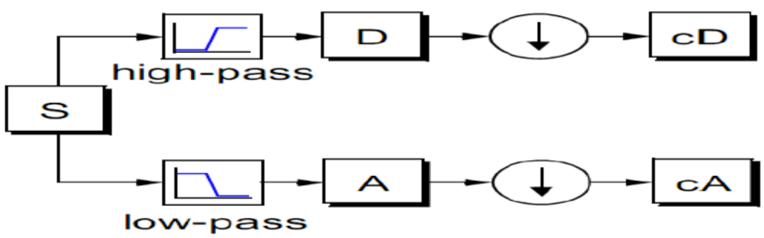


Figure 6-1 First Step of DWT

The DWT is an iterative process and subsequent levels of the decomposition are achieved by applying the same step as in figure A.1 to the approximation coefficients from the previous, upper level (Figure 6.3).

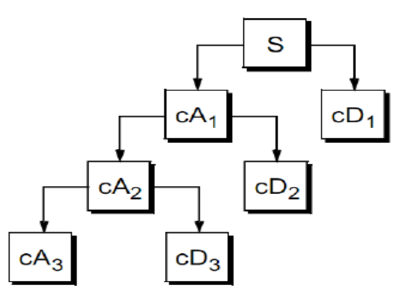


Figure 6-2 Wavelet Decomposition Tree

Inverse Discrete Wavelet Transform

After a signal is decomposed using the DWT, it is possible to reconstruct it with the Inverse Discrete Wavelet Transform (IDWT). The reconstruction is done using the DWT coefficients and carries no loss of information.

The process, similar to the DWT, consists of up sampling the DWT coefficients followed by filtering. To up sample the coefficients, zeros are inserted between each sample. The obtained signal, twice the size of the original coefficients, is then filtered using a similar filter to the one used in DWT analysis for the respective coefficients [34].

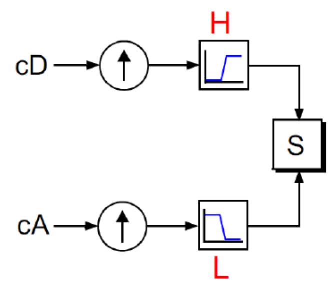


Figure 6-3 IDWT Reconstruction

The reconstruction of the detail and approximation coefficients from certain level n originates the approximation coefficients from the level n-1. By reconstructing each level consecutively, the outputs from level 1 are finally combined to generate the original signal.

The most useful feature of reconstruction is that is it possible to reconstruct the approximations and details from the respective coefficients. Reconstructed approximations and details have the same size as the original size. Consequently, by adding the desired approximations and/or details from certain levels, it is possible to obtain a signal which is a representation of the original one for the frequency bands corresponding to the coefficients used for reconstruction [34].

APPENDIX B

Fourier Transform

The Fourier Transform of a signal x(t) is defined as:

$$X(f) = \int_{-\infty}^{+\infty} x(t) \bullet e^{-2\pi jft} \qquad \text{Eq.9}$$

Where t and f are the time and frequency parameters, respectively. Fourier transform uses complex exponentials (sinusoids) as building blocks. It is unsuitable for time varying or non-stationary signals such as heart sounds .In these signals there are many non-stationary characteristics or sudden frequency changes or transients which are important and cannot be detected in time using Fourier transform [3].

Short Time Fourier Transform (STFT)

In STFT, the signal x (t) is divided into very small time segments which are assumed to be stationary. A window function 'w' is used and its width is equal to the length of the time segment assumed to be stationary. This window function is first located at the beginning of the signal and multiplied with the signal. Fourier transform is then performed on this product. The window is then shifted to the right at a later time and multiplied with the signal. Fourier transform is performed on the product again. This process is repeated until the end of the signal is reached. The STFT is of the form:

$$X(t,f) = \int_{-\infty}^{+\infty} [x(t) \bullet \omega(t - t')] \bullet e^{-2\pi jft} dt \qquad \text{Eq.10}$$

However, STFT has its limitations on the time-frequency resolution of the signal, determined by the size of the window. Once the window function is chosen, the time and frequency resolutions are fixed throughout the processing. Greater frequency resolution can only be achieved at the expense of time resolution and vice versa. Using a narrow time window gives good time resolution but poor frequency resolution. Using a wide time window gives good frequency resolution but poor time resolution [2].

Continuous Wavelet Transform (CWT)

Wavelet transform is designed to give good time resolution and poor frequency resolution at high frequencies and good frequency resolution and poor time resolution at low frequencies. This approach makes sense especially when the signal at hand has high frequency components for short durations and low frequency components for long durations, which is the case in most biological signals such as heart sounds [3].

The most significant characteristic of the wavelet transform is that it does not take the Fourier transforms of the windowed signals rather the width of the window is changed as the transform is computed foe every single spectral component [3].

The continuous wavelet transform (CWT) is defined as follows:

$$\mathrm{CWT}_x^{\psi}(b, \mathrm{a}) = \frac{1}{\sqrt{\mathrm{a}}} \int \mathrm{x(t)}\ \psi^{*}\left(\frac{\mathrm{t}-\mathrm{b}}{\mathrm{a}}\right)\mathrm{dt} \qquad \text{eq.11}$$

Where x (t) is the signal analyzed, 'a' the scale factor, 'b' the translation factor, and ψ the complex conjugate of the mother wavelet used. The mother wavelet must be of finite length or compactly supported and must be oscillatory. The wavelet must also integrate to zero. The mother wavelet is the prototype for generating the other window functions for wavelet transform of the signal. It is dilated or compressed by the scaling factor 'a' to produce a family of wavelets. The scaled wavelets include an energy normalization term $\frac{1}{\sqrt{a}}$ which keeps the energy of the scaled wavelets the same as the energy in the mother wavelet. The wavelets, each with finite energy, have their energies concentrated in time or space to enable the wavelets to analyze transient and non-stationary phenomenon [3].

The term wavelet means a small wave. The smallness refers to the condition that this (window) function is of finite length (compactly supported). The wave refers to the condition that this function is oscillatory.

The relation between scale and frequency is that low scales correspond to high frequencies and high scales to low frequencies. Regarding previous discussion we can now introduced the time-scale plane. The illustration in figure 6-5 is commonly used to explain how time and frequency resolutions should be interpreted. Every box in figure 6-4 corresponds to a value of the wavelet transform in the time- scale plane. Note that boxes have a certain non-zero area, which implies that the value of a particular point in the time- scale plane cannot be known .Regardless of the dimensions of the boxes, the areas of all boxes, both in STFT and WT, are the same and determined by Heisenberg's inequality. As a summary, the area of a box is fixed for each window function (STFT) or mother wavelet (CWT), whereas different windows or mother wavelets can result in different areas. However, all areas are lower bounded by 4 / π. That is ,we cannot reduce the areas of the boxes as much as we want due to the Heisenberg's uncertainty principle.

$$\Delta t \Delta f >= \frac{1}{4\Pi}$$

Δt and Δf at different analyzing frequencies for wavelet transform are related to the scale parameter(a= f_b / f) in the following manner:

$$\Delta t = \frac{f_b}{f} \Delta t_o \qquad \Delta f = \frac{f}{f_b} \Delta f_o$$

Where Δt_o and Δf_o are the time duration and frequency bandwidth of the mother wavelet ψ (*t*), respectively, f_b is the basic frequency of the mother wavelet, and f is the analyzing frequency.

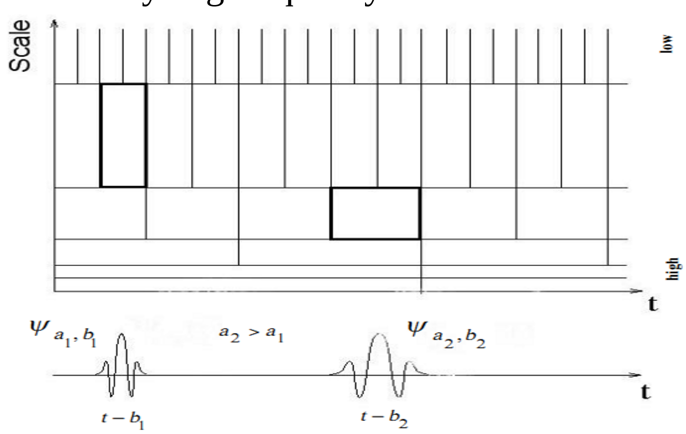


Figure 6-4 The Two Basic Wavelet Operation Scale and Translation Define a Tiling of the Scale Time Plane

❖ The algorithm of CWT can be described as the following (figure 6-5,6,7)

Select a wavelet and compare it to a segment at the beginning of the signal.

2) Compute the CWT coefficient for that segment of the signal.

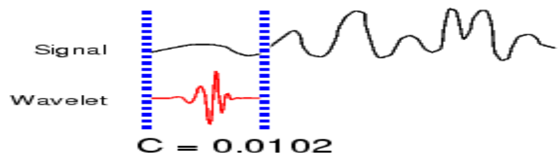


Figure 6-5 Second Step Of CWT

3) Move the wavelet to the right by the factor ‘b’, then repeat step 2, covering the entire signal.

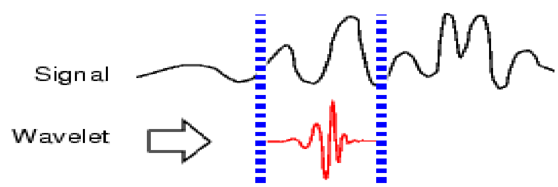


Figure 6-6 Third Step Of CWT

4) Scale the wavelet, and then repeat steps 1 to 3 for the new scale factor ‘a’.

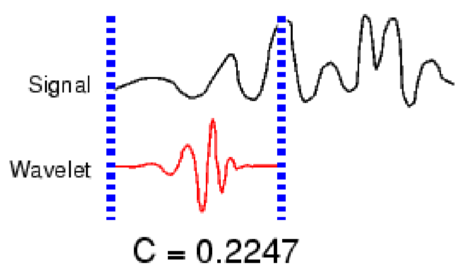


Figure 6-7 Fourth Step Of CWT

5) Repeat steps 1 to 4 for all the other scale factors ‘a’.

APENDIX C

MFCC

MFCC coefficients model the spectral energy distribution in a perceptually meaningful way.

MFCCs are the most widely used acoustic feature for speech recognition, speaker recognition, and audio classification.

MFCCs take into account certain properties of the human auditory system

- Critical band frequency resolution (approximately).
- Log power (dB) magnitudes.

Mel scale

Mel frequency scale represents subjective (perceived) pitch. It is one of the perceptually motivated frequency scales.

- Mel scale is constructed using pairwise comparisons of sinusoidal tones: a reference frequency is fixed and then a test subject (human listener) is asked to adjust the frequency of the other tone to be twice higher or lower
- Models the nonlinear perception of frequencies in the human auditory system.

For comparison, the Bark critical band scale has been constructed based on the masking properties of nearby frequency components.

- Constructed by filling the audible bandwidth with adjacent critical bands 1…26.

Note that all the scales are related and $f_{Mel}=100f_{Bark}$ (very roughly) [50].

An audio signal is constantly changing, so to simplify things we assume that on short time scales the audio signal doesn't change much (when we say it doesn't change, we mean statistically i.e. statistically stationary, obviously the samples are constantly changing on even short time scales). This is why we frame the signal into 20-40ms frames. If the frame is much shorter we don't have enough samples to get a reliable spectral estimate, if it is longer the signal changes too much throughout the frame.

The next step is to calculate the power spectrum of each frame. This is motivated by the human cochlea (an organ in the ear) which vibrates at different spots depending on the frequency of the incoming sounds. Depending on the location in the cochlea that vibrates (which wobbles small hairs), different nerves fire informing the brain that certain frequencies are present. Our periodogram estimate performs a similar job for us, identifying which frequencies are present in the frame.

The periodogram spectral estimate still contains a lot of information not required for Automatic Speech Recognition (ASR). In particular the cochlea cannot discern the difference between two closely spaced frequencies. This effect becomes more pronounced as the frequencies increase. For this reason we take clumps of periodogram bins and sum them up to get an idea of how much energy exists in various frequency regions. This is performed by our Mel filterbank: the first filter is very narrow and gives an indication of how much energy exists near 0 Hertz. As the frequencies get higher our filters get wider as we become less concerned about variations. We are only interested in roughly how much energy occurs at each spot. The Mel scale tells us exactly how to space our filterbanks and how wide to make them.

Once we have the filterbank energies, we take the logarithm of them. This is also motivated by human hearing: we don't hear loudness on a linear scale. Generally to double the percieved volume of a sound we need to put 8 times as much energy into it. This means that large variations in energy may not sound all that different if the sound is loud to begin with. This compression operation makes our features match more closely what humans actually hear. Why the logarithm and not a cube root? The logarithm allows us to use cepstral mean subtraction, which is a channel normalization technique.

The final step is to compute the DCT of the log filterbank energies. There are 2 main reasons this is performed. Because our filterbanks are all overlapping, the filterbank energies are quite correlated with each other. The DCT decorrelates the energies which means diagonal covariance matrices can be used to model the features in e.g. a HMM classifier. But notice that only 12 of the 26 DCT coefficients are kept. This is because the higher DCT coefficients represent fast changes in the filterbank energies and it turns out that these fast changes actually degrade ASR performance, so we get a small improvement by dropping them [51].

Why are MFCC coefficients successful in audio classification?

- Perceptually motivated (near log f) frequency resolution.
- Perceptually motivated decibel magnitude scale.
- Discrete cosine transform decorrelates the features (improves statistical properties by removing correlations between the features).
- Convenient control of the model order: picking only the lowest N coefficients gives lower resolution approximation of the spectral energy distribution (vocal tract etc.) [50].

APENDIX D

RS232

Communication port

A port is a point at which an external device attaches to the computer system. Ports allow data to be sent / retrieved from the external device [44].

The most common type of ports are:

1. Serial port – 1bit is transmitted at a time; COM port is a type of serial port.

It is a serial communication physical interface through which informat-ion transfers in or out one bit at a time, it represents the first interfaces to allow computers to exchange information with the outside world, the term serial refers to data sent via a single wire: the bits are sent one after other.

It also called rs-232, in it there will be two data lines: one transmission and one receive line. To send a data in this port, it has to be sent one bit after another. It generally has 9 pins explained in the below table [45]:

Table 6-1 explain bins of rs232

Abbreviation	Full Name	Function
TD	Transmit Data	Serial Data Output (TXD)
RD	Receive Data	Serial Data Input (RXD)
CS	Clear to Send	This line indicates that the Modem is ready to exchange data.
DCD	Data Carrier Detect	When the modem detects a "Carrier" from the modem at the other end of the phone line, this Line becomes active.

DSR	Data Set Ready	This tells the UART that the modem is ready to establish a link
DTR	Data Terminal Ready	This is the opposite of DSR. This tells the Modem that the UART is ready to link.
RTS	Request To Send	This line informs the Modem that the UART is ready to exchange data.
RI	Ring Indicator	Goes active when modem detects a ringing signal from the PSTN

2. Parallel port – allow multiple bits (e.g. 8 or 16) bits to be transmitted simultaneously such as printers.

Parallel data transmission involves sending data simultaneously on several wires. The parallel ports on personal computers can be used to send 8 bit simultaneously via 8 wires. It is easy to program and faster compared to the serial ports. But main disadvantage is it needs more number of transmission lines. Because of this reason parallel ports are not used in long distance communications [44].

Parallel: expensive – short distance – fast – no modulation

Serial: cheaper – long – slow

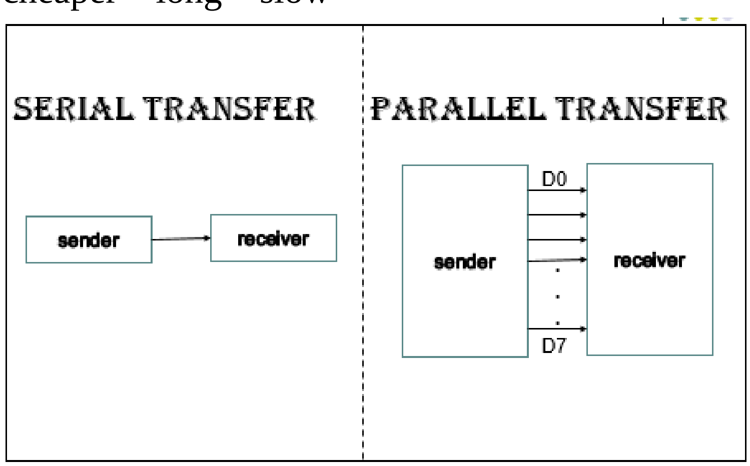


Figure 6-8 serial and parallel transfer.

APENDIX E

Graphical User Interface (GUI)

GUI components can include menus, toolbars, push buttons, radio buttons, list boxes, and sliders—just to name a few. UIs created using MATLAB tools can also perform any type of computation, read and write data files, communicate with other UIs, and display data as tables or as plots [49].

Typically, GUIs wait for a user to manipulate a control, and then respond to each user action in turn. Each control, and the UI itself, has one or more callbacks, named for the fact that they "call back" to MATLAB to ask it to do things. A particular user action, such as pressing a screen button, or passing the cursor over a component, triggers the execution of each callback. The UI then responds to these events. You, as the UI creator, write callbacks that define what the components do to handle events.

A MATLAB UI is a figure window to which you add user-operated components. You can select, size, and position these components as you like. Using callbacks you can make the components do what you want when the user clicks or manipulates the components with keystrokes. Can build MATLAB UIs in two ways: [49]

• Create the UI using GUIDE this approach starts with a figure that you populate with components from within a graphic layout editor. GUIDE creates an associated code file containing callbacks for the UI and its components. GUIDE saves both the figure (as a FIG-file) and the code file. You can launch your application from either file.

• Create the UI programmatically using this approach, you create a code file that defines all component properties and behaviors. When a user executes the file, it creates a figure, populates it with components, and handles user interactions. Typically, the figure is not saved between sessions because the code in the file creates a new one each time it runs.

When you want to build GUI by guide must do this steps: [49]

1. Start GUIDE by typing guide at the MATLAB prompt.

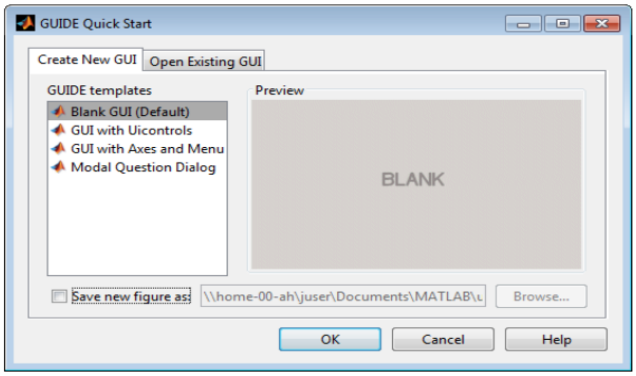


Figure 6-9 Guide Quick Start.

2. In the GUIDE Quick Start dialog box, select the Blank GUI (Default) template, and then click OK.

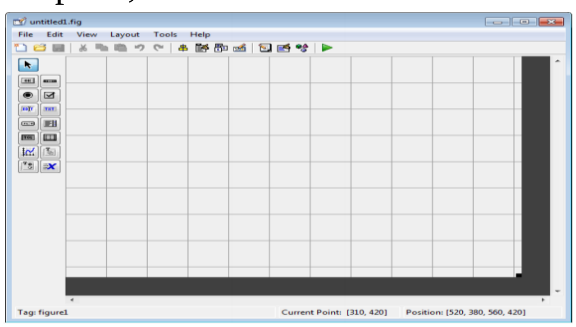


Figure 6-10 Blank GUI Template.

3. Display the names of the UI components in the component palette:
   a. Select File > Preferences > GUIDE.

b. Select Show names in component palette.

c. Click OK.

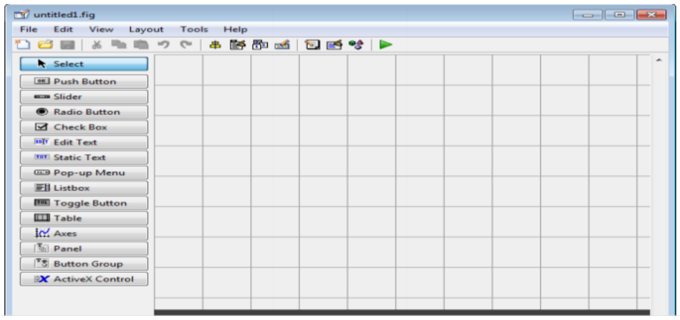
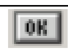
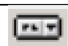
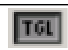
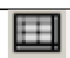
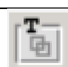


Figure 6-11 Prepare GUI Template.

And then you make the form of GUI as you like by using below components depending on applications(According to the application[49]:)

Table 6-2 GUI components.

Component	**Icon**	**Description**
Push Button		Push buttons generate an action when clicked. For example, an **OK** button might apply settings and close a dialog box. When you click a push button, it appears depressed; when you release the mouse button, the push button appears raised.
Slider		Sliders accept numeric input within a specified range by enabling the user to move a sliding bar, which is called a slider or thumb. Users move the slider by clicking the slider and dragging it, by clicking in the trough, or by clicking an arrow. The location of the slider indicates the relative location within the specified range.
Radio Button		Radio buttons are similar to check boxes, but radio buttons are typically mutually exclusive within a group of related radio buttons. That is, when you select one button the previously selected button is deselected. To activate a radio button, click the mouse button on the object. The display indicates the state of the button. Use a button group to manage mutually exclusive radio buttons.
Check Box		Check boxes can generate an action when checked and indicate their state as checked or not checked. Check boxes are useful when providing the user

		with a number of independent choices, for example, displaying a toolbar.
Edit Text		Edit text components are fields that enable users to enter or modify text strings. Use edit text when you want text as input. Users can enter numbers but you must convert them to their numeric equivalents.
Static Text		Static text controls display lines of text. Static text is typically used to label other controls, provide directions to the user, or indicate values associated with a slider. Users cannot change static text interactively.
Pop-Up Menu		Pop-up menus open to display a list of choices when users click the arrow.
List Box		List boxes display a list of items and enable users to select one or more items.
Toggle Button		Toggle buttons generate an action and indicate whether they are turned on or off. When you click a toggle button, it appears depressed, showing that it is on. When you release the mouse button, the toggle button remains depressed until you click it a second time. When you do so, the button returns to the raised state, showing that it is off. Use a button group to manage mutually exclusive toggle buttons.
Table		Use the table button to create a table component. Refer to the uitable function for more information on using this component.
Axes		Axes enable your UI to display graphics such as graphs and images. Like all graphics objects, axes have properties that you can set to control many aspects of its behavior and appearance.
Panel		Panels arrange UI components into groups. By visually grouping related controls, panels can make the user interface easier to understand. A panel can have a title and various borders. Panel children can be user interface controls and axes as well as button groups and other panels. The position of each component within a panel is interpreted relative to the panel. If you move the

		panel, its children move with it and maintain their positions on the panel.
Button Group		Button groups are like panels but are used to manage exclusive selection behavior for radio buttons and toggle buttons.
ActiveX Component		ActiveX components enable you to display ActiveX controls in your UI. They are available only on the Microsoft® Windows® platform. An ActiveX control can be the child only of a figure window. It cannot be the child of a panel or button group.