A Wavelet-based Approach to Electrocardiogram (ECG) and Phonocardiogram (PCG) Subject Recognition

by

Seyedeh Zahra Fatemian

A thesis submitted in conformity with the requirements
for the degree of Master of Applied Science
Graduate Department of Electrical and Computer Engineering
University of Toronto

Abstract

A Wavelet-based Approach to Electrocardiogram (ECG) and Phonocardiogram (PCG) Subject Recognition

Seyedeh Zahra Fatemian

Master of Applied Science

Graduate Department of Electrical and Computer Engineering

University of Toronto

2009

This thesis studies the applicability of two cardiac traits, the electrocardiogram (ECG) and the phonocardiogram (PCG), as biometrics. There is strong evidence that cardiac electrical activity (ECG) embeds highly distinctive characteristics, suitable for applications such as the recognition of human subjects. On the other hand, having the same origin with the ECG signal, it is believed that the PCG signal conveys distinctive information of an individual which can be deployed in biometric applications. Such recognition systems traditionally provide two modes of functionality, identification and authentication; frameworks for subject recognition are herein proposed and analyzed in both scenarios.

Moreover, the expression of the cardiac signals is subject to alternation with heart rate and noise components. Thus, the central consideration of this thesis is the design and evaluation of robust recognition approaches that can compensate for these effects. A recognition system based on each, the ECG and the PCG, is developed and evaluated. Furthermore, a fusion of the two signals in a multimodal biometric system is investigated.

Acknowledgements

Among the biometric recognition techniques, lately identity recognition based on Electrocardiogram (ECG) and heart sound have emerged to be suitable biometric modalities. As a whole new topic working in this area has been challenging, in that there have been few publicly available databases and existing software and hardware.

I would like to sincerely thank my advisor, Professor Dimitirios Hatzinakos, for all his support, guidance, and knowledge. Without his generous help this research would have not been accomplished. Special appreciation goes to all my lab mates for their help and support in collecting the databases and encouraging me. In addition, I would like to sincerely thank all participants in this research for their effort, time, and patience.

I am also grateful to the Department of Electrical and Computer Engineering for their financial support and to the Communication Group faculty members for their technical advices. Financial support of for this thesis was provided by the Natural Science and Engineering Research Council (NSERC) of Canada.

Finally, I would like to thank my family and friends for their constant support and encouragement during my studies.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 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 e-commerce 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. A publicized survey conducted by Javelin Strategy & Research [1] in 2007 concerning loss of personal privacy, fraudulent funds transfers, outright theft, and abuse of identity in network transactions, indicated that 8.1 million Americans were victimized by identity fraud or identity theft. *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

[1] http://www.privacyrights.org/ar/idtheftsurveys.htm

or credit card account numbers, passwords, telephone calling card number, birth date, name, address and so on.

According to [1], online 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, and this is reported to constitute 17% of the fraud cases.

This motivated researchers to seek for reliable and accurate alternative for identity verification over the traditional password/ ID card based systems. Biometrics, which are physiological or behavioral characteristics extracted from human subjects, 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.

1.2 Biometric Criteria and System Requirements

Biometrics employed so far fall into two main categories [2], The first group includes behavioral traits such as gait and keystroke that offer sufficient discriminatory informa-

tion to permit identity verification in low security, small scale applications. This class of biometric attributes has great potential in surveillance applications, as it is non-intrusive and can be collected from a distance; therefore, it is easy to monitor, and does not require the consent of people under surveillance. However, behavioral characteristics suffer from shortcomings mostly due to their variability with personal and environmental factors such as stress, age, injuries, and illnesses. On the other hand, they can be easily mimicked and usually need large databases to process and consequently have high computational costs.

The second group which has gained popularity and acceptance in most security applications includes physiological traits such as face, iris, fingerprint, and voice. However, each of these biometrics has its strengths and weaknesses, and the choice depends on the restrictions imposed by the application context and the properties of the biometric characteristic. For instance, the fingerprint is one of the most popular biometrics that is easy to collect applicable to real time applications. However it is vulnerable to ageing and injuries, while it can also be easily falsified using Latex or silicone gel.

Analyzing face images for recognition has drawn lots of attention since it has a non intrusive acquisition process and deploys non expensive cameras and equipments. However, this credential is affected by environmental factors such as illumination, pose, facial expression, occlusions, and even low resolution images. Moreover, it can be easily disguised or forged. Although with the improvement of processors most of face recognition algorithms are implementable in a reasonable amount of time, yet processing usually needs large databases and consequently have high computational costs.

Iris recognition techniques use pattern recognition based on high-resolution images of the iris texture of an individual. Converted into digital templates, these images provide mathematical representations of the iris that yield unambiguous verification of an individual. Depending on the camera to be used for capturing the image, the recognition efficacy is rarely impeded by glasses or contact lenses. The most important advantage

of this credential as a biometric is its stability with aging so that a single enrollment can last over a lifetime. Despite all its advantages, iris recognition suffers from some deficiencies that limit its deployment in many applications. First, recognition cannot be performed from a distance larger than a few meters from the cameral; besides, it requires the cooperation of the person to be identified by holding the head still and looking into the camera. Second, the iris scanners are expensive, the iris recognition is susceptible to poor image quality , and it has high false rejection rates.

In recent years, attention has been drawn to medical biometrics, a subset of physiological biometrics, including DNA, ECG (Electrocardiogram), EEG (Electroencephalogram), blood pressure, and heart rate. Along with benefits that come from the employment of medical traits, several restrictions and considerations are imposed that make their implementation challenging.

For a biological trait, whether physiological or behavioral, to be qualified to be used as a biometric characteristic, according to [2] 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:

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- 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 trait's 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.

When evaluating recognition systems, one should take specific parameters into consideration. Examples of which include:

Performance - Refers to the robustness of technology and achievable recognition accuracy and speed based on the operational factors which affect the accuracy and speed, in addition to the resources required to achieve this performance. Considering the technologies involved, 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 -Indicates the degree of approval for a technology by people and the extent to which they are willing to use a particular biometric identifier in their daily lives. This is highly related to the biometric characteristic employed by the system and the comfort or inhibition caused by providing that biometric trait.

Circumvention- Is one of the most crucial and eminent criteria which determines the extend to which a biometric characteristic can be deployed in different applications. It reflects the robustness of the system against fraudulent methods and attacks. A fraud can take place either with the application of falsified credentials or through mimicking someone else's behavioral attributes. Depending on the properties of the utilized bio-

metric, some systems are more robust to fraudulent methods. However, there is ongoing research on methods that compensate for this inherent weakness of some of the biometrics. For instance, in fingerprint recognition, new technologies are implemented so that a liveness indication test is performed considering the properties of live human's body tissue.

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.

1.3 Motivation for ECG

Among the biometric recognition techniques, lately electrocardiogram (ECG) has emerged to be a suitable biometric modality for medium and high security applications.

ECG reflects the cardiac electrical activity recorded by electrodes placed on the body surface. Containing a wealth of information, ECG traces have been used for medical diagnostic purposes since the early 20th century. In the last 20 years, however, researchers have been able to apply digital and numerical analysis to the data in order to perform more complex diagnostic interpretation tasks such as identifying atrial and ventricle fibrillation [3] - [4], myocardial infarction [5], identifying many cardiac diseases, and even demixing the fetal signal from the ones of the mother [6]. Demonstration of ECG's stability over a sufficiently long period of time as well as their uniqueness for individuals, motivated researchers to employ ECGs for biometric purposes [7] - [15]. ECG biometric is expected to address some deficiencies of other biometric modalities or to complement them thanks to its universality, distinctiveness, small storage requirements, speed, and most importantly the unobtrusive and inherent liveness-indicating nature.

As mentioned earlier, one of the greatest concerns in utilizing biometrics is the possibility of exposing falsified or fraudulent credentials. Automatic liveness detection has been suggested as a means to deter identifying theft and counterfeit identifiers so that the system will reassure the presence of the claimed subject. However, for the ECG signal, representation of the signal by itself indicates the presence of the alive legitimate user.

Moreover, ECG biometric can cover a greater portion of the population since physical disabilities are not an obstacle in collecting the signal; however, the application of ECG signals in the presence of cardiac disorders is still an open research. Furthermore, ECG can be easily fused with other physiological biometric modalities, such as fingerprint, in order to make a robust multiple modalities recognition system.

However, there are some concerns about application of ECG as a biometric since it conveys a wealth of information about the state of cardiac health which is considered to be a private issue. On the other hand, having an intrusive acquisition nature, this technique is best suited to verification and identification of cooperative subjects.

1.4 Motivation for Heart sound

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 [58] - [74]. 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 [56, 57].

1.5 Modes of Operation

The final goal of any biometric system is to recognize an individual based on the match established between an identifier represented by the subject to be identified and a stored sample of the claimed subject. Any biometric system can perform this recognition in two fundamental modes, namely the *identification* and the *verification*mode. The major difference between these two modes is the state of knowledge in which the system stands regarding the identity of a subject. Generally, the *identification* is required in applications such as surveillance, whereas the*verification* is performed in most authentication applications.

In the *identification* mode, the system does not have prior information about the identity of the subject. Therefore, it conducts a one-to-many comparison to find the best match between an individual's collected sample and all subjects stored in the gallery

set. In cases where the resemblance of pairs is unacceptable according to a predefined threshold, the system does not allow the identification to take place, and the subject is rejected. Utilizing such a threshold enables the system to perform a negative recognition in addition to rejection or acceptance of a claimed identity. This means that the system can determine whether a subject even exist in the system database or not. Usually the *identification* mode is time consuming due to computational efforts required to determine the best match.

On the other hand, in the *verification* mode, through an one-to-one comparison, the system validates a claimed identity by comparing someone's captured biometric characteristics with the ones corresponding to the claimed record. An individual is authenticated when there is adequate resemblance with respect to a threshold between the pair.

1.6 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:

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. 1.1 and Eq. 1.3.

$$FAR(n) = \frac{\text{Number of successful independent fraud attempts against a person n}}{\text{Number of all independent fraud attempts against a person n}} \quad (1.1)$$

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):

$$FAR = \frac{1}{N} \sum_{n}^{N} FAR(n) \quad (1.2)$$

$$FRR = \frac{\text{Number of rejected qualified attempts}}{\text{Total number of qualified attempts}} \quad (1.3)$$

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.

1.7 Ethical issues

As biometric identification technology attains a larger presence in everyday life, a range of social, legal and ethical concerns have been evoked. Concerns are based on a variety of issues, in both the social and the ethical dimension, including fears about the centralization of information, the potential for data abuses, and the development of a unique identifier. The issues of potential data misuse and high-tech surveillance have become more paramount than ever as the biometric identification technology has attained a larger presence in everyday life. However, education and appropriate policy making can diminish concerns regarding the new technology. All involved parties and organizations are making an effort to establish international policies to appropriately and globally govern the identification technologies based on the ground rules of enhanced security, respect for privacy, respect for human dignity and the technological transparency. This globalization is required to create the standards to ground electronic transactions. Moreover, there has been a lot of research about different encryption, data acquisition, and data filing methods that would guarantee confidentiality of the information and avert the privacy violations.

1.8 Research goals and Contributions

In this thesis, the analysis of two relatively new medical biometric attributes, the electrocardiogram (ECG), and the heart sound (PCG), as well as their fusion is studied. The primary objective of the current research is to address the following issues:

1. *Design of a robust recognition system based on ECG signal*: To date, two major approaches have been proposed for ECG recognition, namely fiducial-based and non fiducial-based methods, that achieve good performance in an almost ideal scenario in terms of noise and heart rate (HR) [7,15]. However, for the ECG biometric technology to mature and get embedded in different applications, challenges associated with noises encountered in practice and different HRs should be addressed; issues of which would have a profound influence on the way the whole ECG recognition process is conducted. Therefore, it is clear that there is plenty of scope for improvement in designing the ECG based recognition systems with the all-embracing goal of elegantly and accurately handling the subject recognition in real life scenarios.

   In this work, a wavelet based approach is developed and evaluated for automatic analysis of single lead ECG signals for biometric applications. So far, the wavelet transform has been applied for processing of the ECG signal in different contexts [21]- [25]; however, in this thesis an application of the DWT in the area of the ECG recognition is provided. The proposed system has two major consideration: first, utilizing a robust preprocessing stage that reduces the noise effects and allows for detection of the affected heartbeats and excluding them from further analysis. Second, delineation and standardization of the heartbeats that enables us to design a strong personalized heartbeat template for each subject regardless of the heart rate, noise, and intra-class variations.

   The design of the heartbeat template, which is the main novelty of the proposed approach, serves the system in three ways: first, it makes it capable of handling ECGs regardless of the heart rate which renders making presumptions on the individual's emotional condition unnecessary. Second, it reduces the size of the gallery set significantly since it consists of only one heartbeat template for each subject; thus, decreases the storage requirements of the system. Third, it increases the classification speed since only one template is compared against for each subject.

2. *Design of a recognition system based on PCG signal*: Human recognition based on the PCG signals is a new area only lately . Having an acoustic nature, the use of the PCG signal in recognition context is very challenging due to wave characteristic in the presence of noise components. In this work, a wavelet based approach for processing and analyzing the PCG signals for recognition purposes is developed and evaluated over the collected heart sound database. The proposed approach provides a preprocessing stage based on the discrete wavelet transform that tends to reduce the effect of noise and artifact components. Moreover, the complexity of the system is reduced by releasing the necessity for heart sound's components detection. The comparison between this method and the previous works demonstrated the robustness of the proposed system.

3. *Information fusion of both ECG and PCG*: Each biometric technique has its own limitations and advantages; therefore, in order to improve the recognition performance, multi-modal biometric fusion is introduced. The motivation for a multi-modal biometric system includes using complementary information to reduce the measurement errors as well as utilizing multiple classifier fusion technique to increase the correct classification rate. In this work, fusion of the ECG and PCG signals for establishing a robust multimodal biometric system is investigated.

4. *Propose recognition frameworks for both identification and verification modes*: Current approaches to ECG based recognition systems mainly cover the identification mode. For the heart sound recognition, no verification mode of operation has been investigated so far. This thesis will address the challenges and performance of the proposed systems under both identification and verification.

5. *Data collection*: Since there had been no publicly available datasets offering ECG recordings of subjects under varying heart rates, an experiment was conducted at the *University of Toronto* to collect the ECG signals with different heart rates.

Therefore, the performance of the proposed recognition system is evaluated under the stress conditions and compared with the previous works. Moreover, due to the absence of publicly available database of healthy heart sound signals required to evaluate the proposed PCG approach, an additional experiment was conducted in the *University of Toronto* in order to collect the required dataset.

The idea and results of this thesis is partially presented in [16] and submitted to [17].

1.9 Thesis outline

The remainder of this thesis is organized as follows: Chapter 2 provides an accurate investigation of temporal and frequency characteristics of different components of an ECG trace as well as common noise sources in recording ECGs. Moreover, a literature review of the previous works in ECG recognition area is presented in this chapter.

In chapter 3, a new wavelet based framework is developed for automatic analysis of single lead electrocardiogram (ECG) which is applicable for both identification and verification. In order to provide the basic material for implementation of the proposed algorithm, the first section of this chapter is dedicated to a brief introduction to the discrete wavelet transform and the mother wavelet used in this work.

Chapter 4 offers an introduction to the PCG signal along with the temporal and frequency characteristics of the different components of a typical PCG trace. A literature review of the previous works done in the area of PCG based human recognition is provided in the second part of this chapter. Finally, the last part of this chapter is dedicated to the introduction of a new PCG based recognition methodology.

Chapter 5 represents the experimental results of the proposed ECG and PCG based recognition systems. In addition, the fusion of these two signals in a multimodal recognition system is investigated. The proposed frameworks are evaluated for both verification

and identification, over three databases and compared with previous approaches.

Finally, in chapter 6 the conclusion and suggestions for future work are provided.

Chapter 2

Electrocardiogram (ECG): A New Biometric

The Electrocardiogram (ECG) signal reflects the recorded cardiac electrical activity of an individual over time. The first part of this chapter offers an introduction to this signal in a biometric context. An accurate investigation of the temporal and frequency characteristics of different components of an ECG trace as well as common noise sources in recording ECGs are crucial issues to be addressed prior to designing the recognition system. In this section the elements of Figure 2.1, are briefly discussed, the properties of which form the basis of the proposed wavelet preprocessing and feature extraction stages presented in the following sections. The second part of this chapter is dedicated to the analysis of the previous works in this area.

2.1 Introduction to The ECG Signal

An electrocardiogram (ECG) trace describes the electrical activity of the heart recorded by electrodes placed on the body surface. The voltage variations measured by the electrodes are caused by the action potentials of the muscle fibers of different parts of the

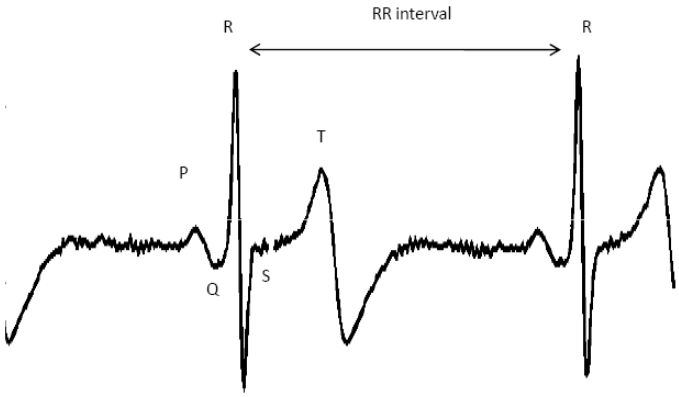

Figure 2.1: Components of an ECG signal.

heart as they make the cells contract. The resulting heartbeat in the ECG is manifested by a series of waves, whose morphology, rhythm, rate, as well as timing convey information which is used for medical diagnosis. Augustus Waller was the first person who recorded the ECG in the 1880s, followed by the Dutch physiologist Willem Einthoven who developed a recording device in the early 20th century and defined sites for electrode placement on the arms and legs. This effort was rewarded with the Nobel Prize in Medicine in 1924. Since then, ECG processing has developed dramatically and become an indispensable clinical tool in many different contexts.

The ECG is a non periodic but highly repetitive signal that may roughly be divided into the phases of depolarization and repolarization of a muscle called Myocyte. The P wave and the QRS complex of a typical ECG are generated during the depolarization of the ventricles and the atrial while the T wave corresponds to the repolarization of the ventricles. These particular elements of an ECG-complex are shown in Figure 2.1.

The number of electrodes attached to the body surface depends on the type of clinical information desired. For studying the heart rhythm it is sufficient to use only a few electrodes, whereas ten electrodes are typically used when information on waveform morphology is required (six on the chest, four on the extremities).

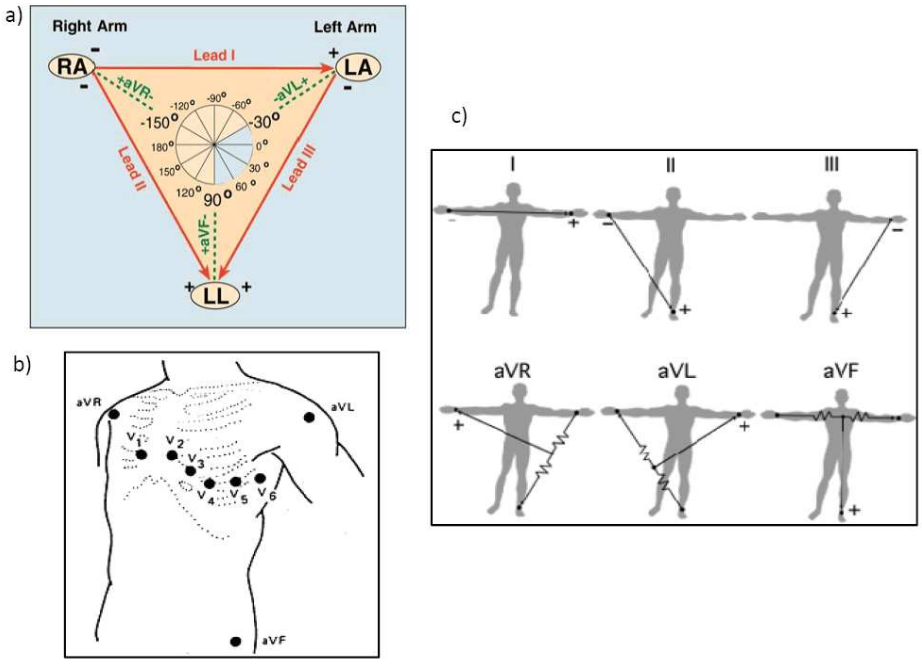

Figure 2.2: Electrode position for recording (a) the 12 directions of the bipolar limb leads, (b) the electrode positions on the chest ($V_1, V_2, V_3 V_4, V_5, and V_6$), (c) the bipolar limb leads (I,II, and III), and the augmented unipolar leads (aVR, aVL, and aVF) [18].

For regular ECG recordings, the variations in electrical potentials in 12 different directions out of the ten electrodes are measured, as depicted in Figure 2.2(a). These 12 different electrical views of the activity in the heart are normally referred to as leads. ECG is typically recorded with a multiple-lead configuration which includes monopolar or bipolar leads, or both. A monopolar lead reflects the voltage variation of a single electrode that is measured in relation to a reference electrode called the "central terminal". Four different artificial reference points that are the average of the signals seen at two or more electrodes are constructed and used to measure the potentials appearing on the left arm (aVL), the right arm (aVR), the left foot (aVF), and on the six chest electrodes (V1-V6), (Figure 2.2(b)). The three bipolar leads reflect the voltage differences between right and left arm (lead I), the right arm and left foot (lead II), and between the left arm and the right foot (lead III), (Figure 2.2(c)).

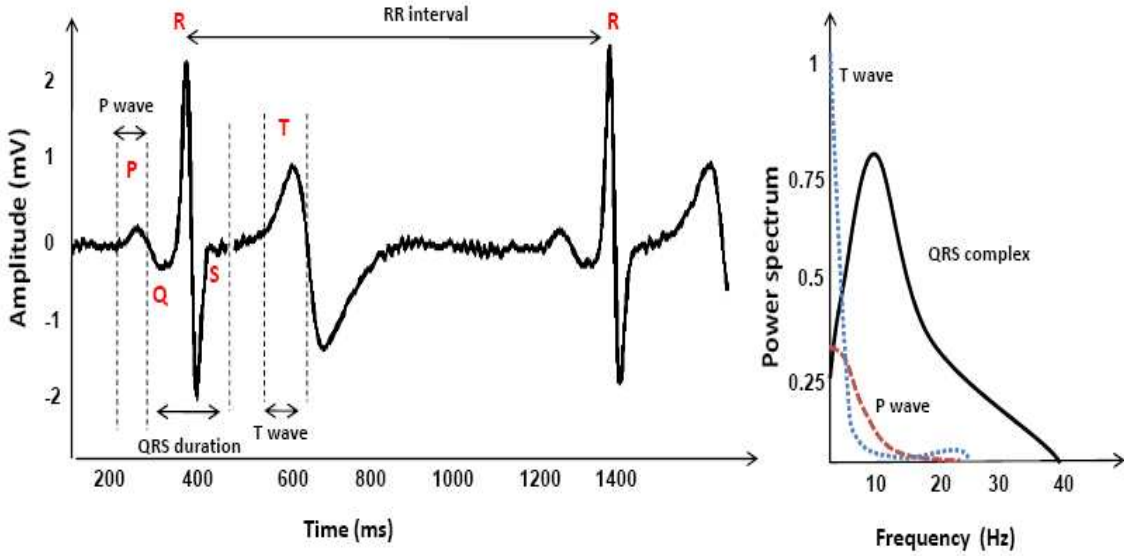

Figure 2.3: (a)Wave duration of the ECG components, (b) power spectrum of the P wave, QRS complex, and Twave.

2.1.1 ECG Waves and Time Intervals

Detection of the different elements of an ECG trace is the basis of some of the ECG recognition methods; therefore, in this section these elements and their temporal and frequency characteristics [18] are briefly discussed. The amplitude of a wave is measured with reference to the ECG baseline level and the duration of a wave is defined by two time instants at which the wave either deviates significantly from the baseline or crosses it. However, since there is no universal method to determine the onset and end of these elements [18], it is sometimes problematic to exactly localize the boundaries between the waves of an ECG trace.

- *P wave*: This wave has a smooth, monophasic morphology that reflects the sequential depolarization of the right and left atria. In most leads, the P wave has positive

polarity with amplitude that is normally less than $300\mu V$, and duration less than $120ms$. The spectral characteristic of a normal P wave is usually considered to be low-frequency, below $10-15Hz$, (Figure 2.3). However, noise-reduction techniques have helped to demonstrate the higher frequency components of the P wave that are useful for predicting the occurrence of certain arrhythmias of particular atrial origin.

- *QRS complex*: This complex reflects the depolarization of the right and left ventricles which for a healthy heart lasts for about $70ms - 110ms$. The first negative deflection of the QRS complex is denoted the Q wave, followed by R wave, while the negative deflection subsequent to the R wave is denoted as the S wave, (Figure 2.3). Since the QRS complex has the largest amplitude of the ECG waveforms, sometimes reaching $2mV-3mV$, it is first identified in any type of computer-based analysis and is used to further analyze the ECG characteristics. The frequency content of the QRS complex, due to its steep slopes, is considerably higher than that of the other ECG waves and is mostly concentrated in the interval $10Hz - 50Hz$. It is shown that the temporal duration of the QRS complex is a function of the atria-ventricle distance and is independent of the heart rate [18]. The R-R interval which is measured between two successive R waves represents the length of a ventricular cardiac cycle and indicates the ventricular rate. The RR interval is the fundamental rhythm quantity in any type of ECG interpretation and is used to characterize different arrhythmias.

- *ST segment*: This segment begins at the end of the S wave and curves into the T wave and represents the interval during which the ventricles remain in an active, depolarized state. Different changes in the ST amplitude or slop often indicate

various underlying cardiac conditions.

- *T wave*: This wave has a smooth, rounded morphology which, in most leads, is associated with a single positive peak that extends about $300ms$ after the QRS complex. The position of the T wave, which reflects ventricular repolarization, is strongly dependent on the heart rate. At rapid rates, the T wave becomes narrower and closer to the QRS complex; while on the other hand, the P wave merges with the T wave, causing the T wave end point to become fuzzy as well as the P wave onset and makes it extremely difficult to localize the wave boundaries.

- *PQ interval*: reflects the time required for the electrical impulse to propagate from the SA node to the ventricles; hence, it is the time interval from the onset of atrial depolarization to the onset of ventricular depolarization. The length of the PQ interval is weakly dependent on heart rate.

All these components tend to introduce large within-subject variability. The stability of the ECG signal over time and under different levels of tension is the most important challenge in using ECG as a biometric which is briefly discussed in next subsection.

2.1.2 ECG Stability Over Time

The ECG signals vary from person to person due to differences in position, size, and anatomy of the heart, age, gender, relative body weight, chest configuration and various other factors,yet they express cardiac features that are unique to an individual. An experiment was conducted by a working Group of *Physikalisch Technische Bundesanstalt* (PTB) to investigate the suitability of ECG signals for biometric applications [19]. Over 27000 ECG recordings were collected from a group of people each repetitively recorded

within a time period of 4 months. A simple pattern comparison technique was used to calculate a distance measure between only features from the *QRS* complex of different ECG recordings. The distribution of the test set yields a distance which indicated a clear separation among ECG recordings from different subjects and explicit similarity among recordings of each subject. Therefore, the stability of the ECG signal of a person over at least a period of four months was confirmed.

Resulting error rates for the verification and authentication processes, FAR and FRR, confirmed the potential of the ECG signals for biometric applications.

2.2 Noise and Artifacts

Raw ECG data contain some noise and artifact components that alter the expression of the ECG trace from the ideal structure described previously and render the clinical interpretation inaccurate and misleading; consequently, a preprocessing step for improving the signal quality is a necessity. It is therefore important to be familiar with the most common types of noise and artifacts in the ECG and address a method which can compensate for their presence before proceeding to the feature extraction step. Below follows some of the most common non-cardiac noise sources the first three of which are of technical origin whereas the fourth is of physiological origin [18].

- *Baseline wander*: is extraneous noise in the ECG trace that may be caused from a variety of noise sources including perspiration, respiration, body movements, and poor electrode contact. The magnitude of this wander may exceed the amplitude of the *QRS* complex by several times, but its spectral content is usually confined to an interval below $1Hz$.

- *Electrode motion artifacts*: are manifested as large-amplitude waveforms which are mainly caused by skin stretching which alters the impedance of the skin around the electrode. They are more problematic to combat since their spectral content is mainly in the range of $1Hz - 10Hz$ that considerably overlaps within the $PQRST$ complex.

- *Power line interference (50/60 Hz)*: is high frequency noise caused by interferences from nearby devices as a result of improper grounding of the ECG equipment.

- *Electromyographic noise (EMG noise)*: is mainly caused by the electrical activity of skeletal muscles during periods of contraction or due to a sudden body movement. While the frequency component of EMG considerably overlaps with that of the QRS complex, it also extends into higher frequencies. As a result, processing the ECG trace to remove these noise effects naturally results in introducing some distortion to the signal.

- *Respiratory activity*: is a beat-to beat variation in the morphology as well as heart rate. This wander is caused during the respiratory cycle as a result of either changes in direction of the vector describing the dominant direction of the electrical wave propagation due to changes in heart position, or changes in lung conductivity.

2.3 ECG As a Biometric: Literature Review

Numerous studies have been conducted on ECG signals in different context and applications. However, regarding our application, we are interested in methodologies proposed

for ECG recognition [7] - [15], as well as ECG processing [21] - [48]. As mentioned earlier, ECG recognition approaches fall into two major categories namely fiducial-based and non fiducial-based methods. Some of the most important proposed methodologies for each of these two approaches are described in this section. Fiducial points are essentially some points of interest on a heartbeat trace such as, onset and end, peak, amplitude, and slope of each of the ECG waves and components.

2.3.1 Fiducial Based Approaches

Biel *et al.* [7] was one of the pioneers who demonstrated the possibility of utilizing ECG signal for human recognition. Using the standard 12-lead ECG signals collected from 20 subjects, they examined the effect of variation in the lead placement as well as different operators. The experimental results suggested that variations in electrode placement may lead to different traces; thus, different medical diagnosis.

Moreover, they tested the optimum number of features required for subject identification by extracting a set of features and then eliminating the correlated features not exhibiting large between-class discrimination. In order to do so, a SIEMENS ECG apparatus was used to extract 30 characteristic features, defined by physicians, which convey diagnostic information. This included the boundaries, duration, slope, and amplitude of the P and T waves, and the QRS complex. Extracting these 30 features from all 12 leads, an attribute vector with $360(12 \times 30)$ elements was obtained for each subject.

The correlation matrix exhibited a strong correlation among features from all 12 leads. Therefore, only the features extracted from the chest and limb leads were retained for further analysis. However, the similarity between the collected features from limb and chest caused the limb features to win out over the chest features owing to the fact that acquiring limb signals are more feasible. The dimensionality of the feature space of the limb lead was further reduced by eliminating the features exhibiting strong correlation.

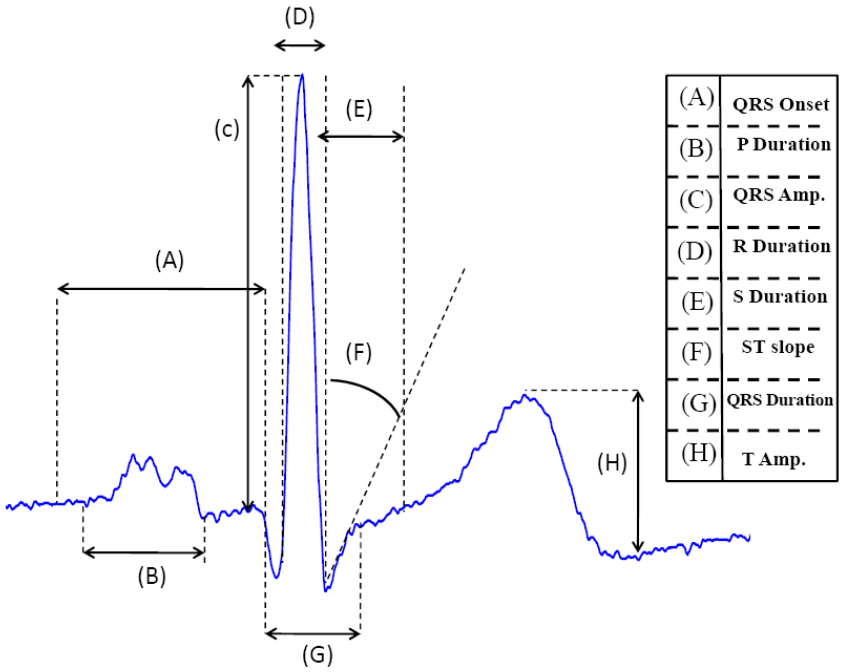

Figure 2.4: The set of 8 features extracted from each heartbeat in [7].

Eventually, a set of 12 features was retained for identification.

At the last step, a combination of multivariate analysis and Principle Component Analysis (PCA) was used to perform classification. Following this procedure, 100% recognition rate was achieved. The proposed methodology, suggested the feasibility of deploying ECG for recognition purposes. Moreover, it revealed that recordings from the standard 12 leads are highly correlated; thus, an accurate recognition can be performed using only the information of one lead. However, deploying a medical apparatus for feature extraction restricted the feature space to predefined medical features only. The application of this apparatus, on the other hand, led to the lack of an automatic feature extraction method based on signal processing techniques.

This lack of automaticity was later removed in a proposed methodology by Israel *et al.* [8] by introducing a set of 15 temporal features, demonstrated in Figure 2.5, and defining rules for their automatic extraction. This methodology consists of three major stages: prefiltering, feature extraction, and classification. For noise reduction, the input

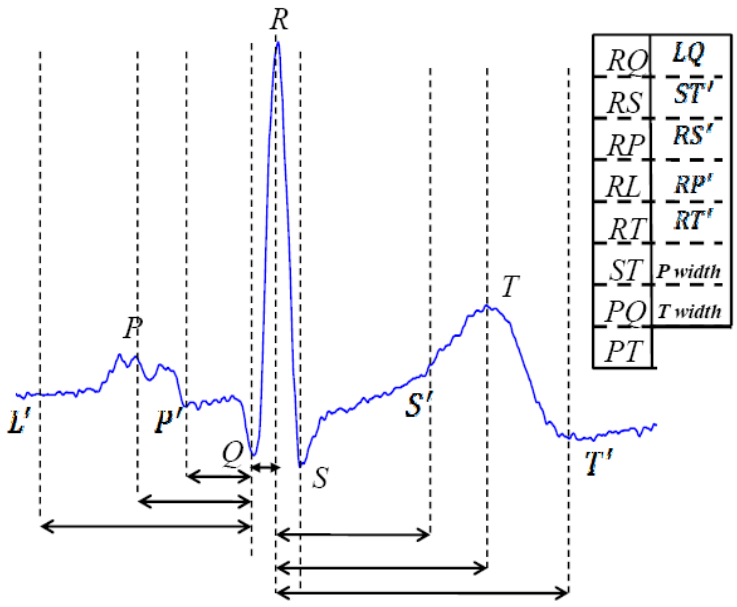

Figure 2.5: The set of 15 temporal features extracted from each heartbeat in [8].

ECG signal was first filtered by a Butterworth filter of order 4 with cutoff frequencies of $1Hz - 40Hz$. Having isolated heartbeats, the signal's peaks were detected in the time domain in the next step by finding local maxima in the regions surrounding the QRS complex, P and T waves. The boundary locations of these wave components were estimated by computation of a minimum radius of curvature.

The recognition process was completed by employing the Wilks' Lamda as the feature selector and linear discriminant analysis (LDA) for dimensionality reduction. Using 12 attributes for classification, the proposed scheme achieved 100% subject identification and 81% heartbeat identification rate for a total of 29 subjects. Although the feature extraction step was now automated, this method still encountered challenges posed by sensitivities in the localization of fiducial points of noisy or distorted heartbeats.

On the other hand, Israel *et al.* [8], conducted a series of experiments to examine the effect of electrode placement on the recognition rate. The neck and chest leads were

examined so that the system was trained with the neck lead and tested with chest lead, and vice versa. The achieved recognition rate indicated the similarity between these two leads.

In another experimental set up, the possibility of subject recognition under different stress levels was examined. Seven tasks, including reading aloud, mathematical manipulation, and virtual driving, were designed to simulate two different states of anxiety, low and high stress levels. Four possible cases were considered: training with low stress and testing with high stress, and vice versa, in addition to the inter class, low with low and high with high, testing and training. The recognition performance for intra-class cases was 78% while 100% recognition rate was achieved for inter-class experiments. The obtained result demonstrated the possibility of human recognition under different stress levels.

Although [8] offered automatic recognition, the fiducial points detection remained challenging and inaccurate due to the fact that the employed fiducial point detector was not adequately robust, especially in the presence of noise and other artifacts.

In a later work, in order to boost up the proposed system, Israel suggested a multimodal recognition system based on ECG and face attributes [9]. For the experimental set up, the ECG traces as well as the facial images were collected from 35 subjects during the enrollment process. The ECG signals were analyzed using the methodology described above [8]. On the other hand, the facial gallery set was constituted using the eigen vectors of the facial data which were selected by the principle component analysis (PCA). The face recognition system was trained on 500 images from the FERET database and then tested with the 35 collected images. Finally, the classification was performed using the Euclidean distance measure.

In order to finalize the recognition decision, three fusion methods were considered. Fusion of ECG and face features at the raw data level achieved 99% recognition rate. However, a decision based fusion obtained 94% subject recognition rate. This rate de-

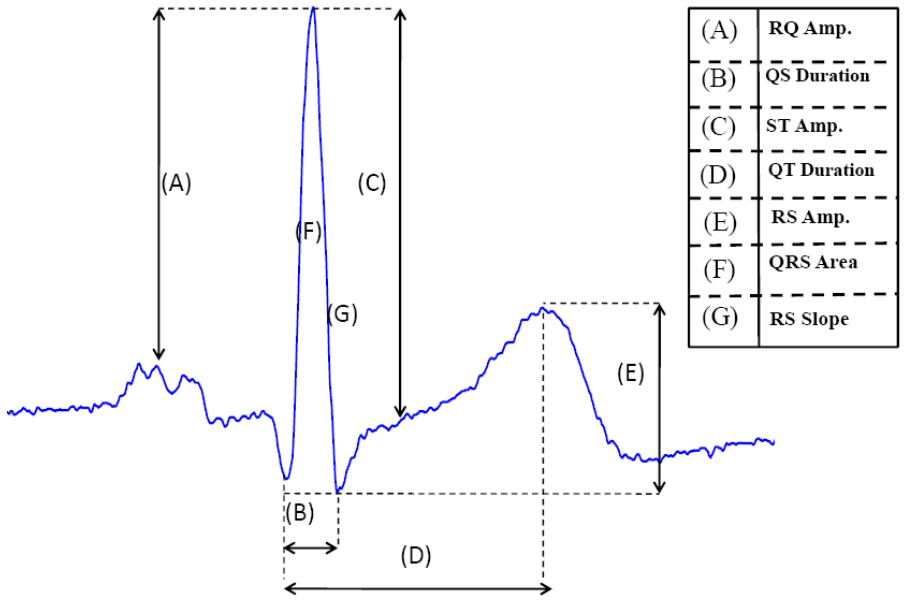

Figure 2.6: The set of 7 temporal and amplitude features extracted from each heartbeat in [10].

graded to 60% in case the voting fusion was used.

In 2002, Shen *et al.* [10], proposed a framework based on a hierarchical scheme: In the first stage, subjects were screened based on a *QRS* complex template matching. In the second stage, the Decision Based Neural Network (DBNN) was used to perform the final recognition among the possible candidates.

The first stage screening was performed on the *QRS*complex since this component of the ECG trace is considered to be invariant to heart rate. Associating the isolated *QRS* complex to a correlation coefficient measure, a similarity score was computed. A threshold value of 85% was chosen empirically to screen out the less probable subjects. Then, seven amplitude and distance fiducial features, demonstrated in Figure 2.6, were computed from the chosen candidates and used to feed the DBNN classifier.

This combined module achieved 100% recognition rate over 20 subjects. However, independent application of these techniques resulted in 95% and 80% recognition rate on

the same database.

In 2005, Shen *et al.* [11], investigated the relationship between the singe lead ECG and the body mass index (BMI) over 168 healthy subjects. The records were divided based on gender in two groups, (113 females and 55 males), in order to study the effect of age, height, weight and BMI on the normalized extracted ECG features.

A band pass filter with cut off frequencies of $1 - 50Hz$ was applied for noise reduction. The feature extraction stage was further improved so that seventeen amplitude and distance features were defined and extracted from isolated heartbeats.

Using the correlation and linear regression analysis, it was concluded that the BMI has a greater impact on the ECG attributes. Since, the selected characteristics showed 25.3% variability with respect to the BMI, but only 6.5% with respect to the age. Further analysis demonstrated that ECG variability occurs more among the male subjects due to their greater BMI.

In order to improve the identification accuracy, Wang *et al.* [12] suggested an integration of analytic and appearance features captured from heartbeats. The defined analytic features were a set of 21 temporal distances and amplitude differences among the fiducial points, whereas the appearance term refers to the preprocessed and segmented ECGs. The proposed approach included three main stages: preprocessing, feature extraction, and classification.

During the preprocessing, signals were filtered using a *4* order Butterworth filter centered at $1 - 40Hz$. Then, the R peaks were localized using a QRS detector called ECGPUWAVE and the heartbeats were segmented by a window of length $800ms$ and centered at the R peaks. After removing the outliers by a thresholding criteria, the heartbeats were aligned with respect to the position of the R peaks, as demonstrated in Figure 2.7.

In feature extraction stage, 15 temporal distance measure in addition to 6 amplitude differences were extracted based on the fiducial point detection algorithm described in [8].

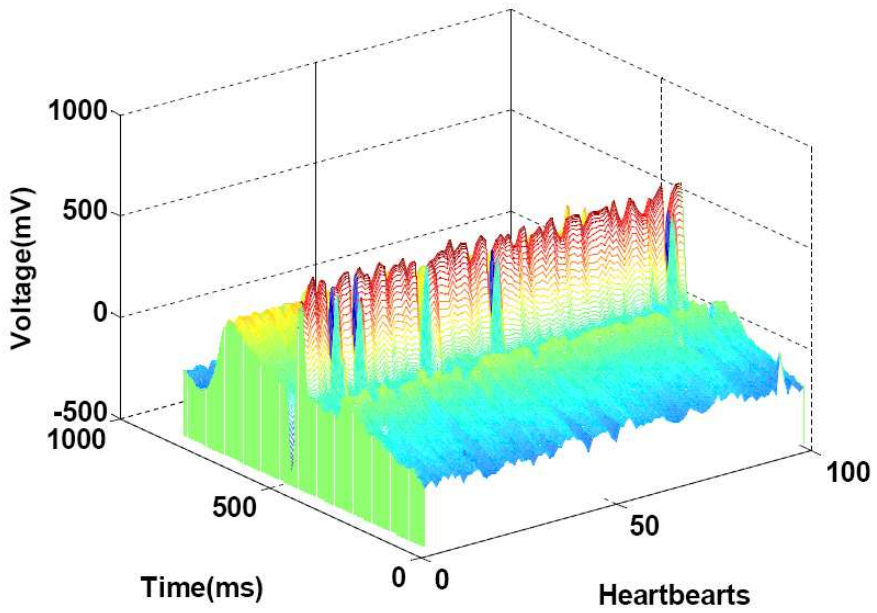

Figure 2.7: Heartbeats synchronized at the R peak.

For the appearance based features, the principle features of the heartbeats were captured using the PCA technique. Classification was performed by application of the Euclidian distance measure and the nearest neighborhood classifier.

Using the analytic and appearance features separately, 74.45% and 95.55% recognition rates were achieved respectively. With replacing dimensionality reduction technique from the PCA to the LDA (Linear Discriminent Analysis), the recognition rate degrades to 93.01%. However, Combination of discriminative information contained in the two types of features in a hierarchical scheme and deploying principal component (PCA), achieved a 100% subject and 98.9% heartbeat identification rates for 13 subjects.

Although the achieved performance in this work was considerable, accurate localization of the features was still an open problem; issue of which has a profound effect on the recognition rate in different scenarios in terms of noise and waveform anomalies.

Researches for addressing issues related to performance, as well as looking at ways to enhance the usability of ECG based biometric systems led to a different approach for ECG recognition which is called the non fiducial-based method and is described as follows.

2.3.2 Non Fiducial Based Approaches

This approach was originally proposed by Plataniotis *et al.* [13] in 2004 in order to eliminate the necessity of fiducial points' localization of the ECG signal. The proposed methodology which is referred to as *AC/DCT* method is based on a combination of autocorrelation and discrete cosine transform. The input ECG signal is first preprocessed using a filtering step similar to the previous works, the bandpass Butterworth filter of *4* order with the range $1-40Hz$. The preprocessed ECG trace is then segmented into non overlapping windows which are longer than the average heartbeat length. The underlying idea for application of windowing technique is the utilization of the ECG windows instead of the specific fiducial points extracted from each heartbeat in order to release necessity of fiducial point's detection.

In next step, the autocorrelation of windowed ECGs were computed, Figure 2.8(a). Based on the findings reported in [13], it is believed that the autocorrelation of windowed ECG signals embeds highly discriminative information that can be used for subject recognition. Owing to the energy compaction property of the DCT, discrete cosine transform (DCT) was applied as a dimensionality reduction technique on the normalized autocorrelation of each window, Figure 2.8(b). The small DCT coefficients were discarded and the retained ones were used to feed the k-nearest neighborhood classifier (KNN).

The classification performance was tested with two similarity measures, the Euclidian distance and the Gaussian log-likelihood distance, over 14 subjects. The simulation results indicated a window and subject recognition rate of 100% using Gaussian log likelihood, whereas the recognition rate using normalized Euclidian distance was 92.8%.

Later, in order to prune the search space of the AC/DCT method, Agrafioti *et al.* [14] suggested a screening procedure. In first step, all collected ECG signals were preprocessed using the bandpass Butterworth filter of *4th* order with the range $1Hz-40Hz$. Then,

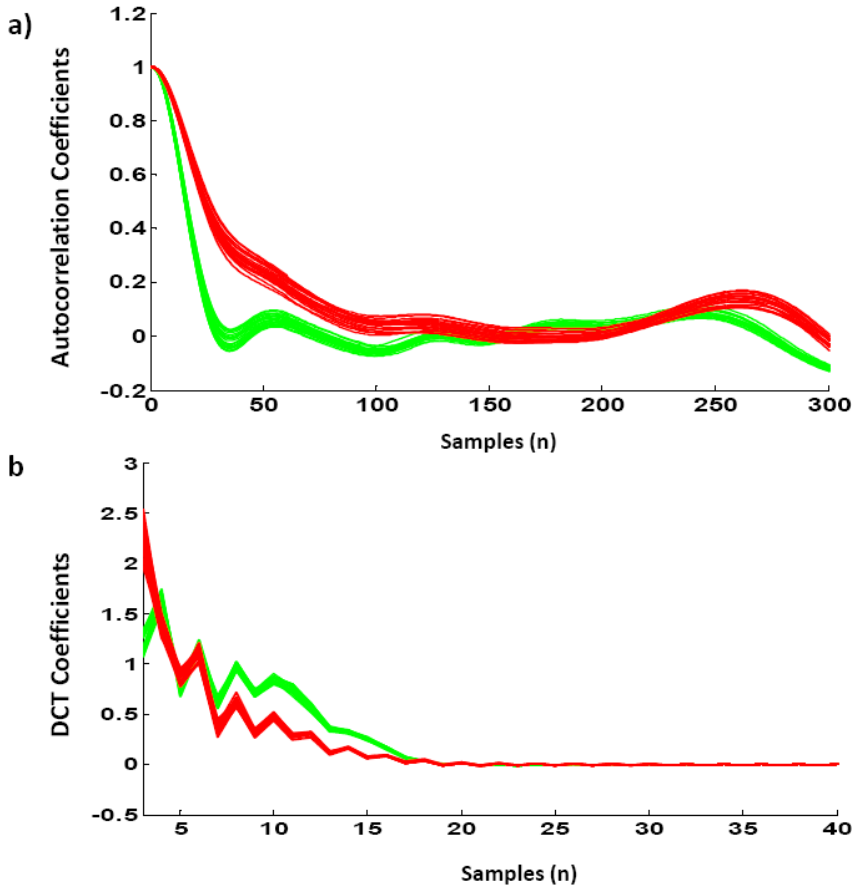

Figure 2.8: The discrimination power of the DCT and autocorrelation coefficients among different subjects (a) the autocorrelation of the ECG windows of 2 different subjects, (b) the DCT of the derived autocorrelation coefficients.

the preprocessed ECG traces were segmented into non overlapping windows with the restriction of having a size longer than the average heartbeat length so that multiple pulses were included in a window. In next step, the normalized autocorrelation of each window was computed and Template Matching was performed using the correlation coefficient (CC) of the lag L of the normalized autocorrelations. The motivation behind Template Matching is to reduce the possible number of classes and eventually carry out classification over a smaller scope.

After the elimination of the less probable subjects, the DCT and LAD were imple-

mented to reduce the dimensionality of the remaining subjects. The identification procedure was concluded by classifying the possible candidates using the Euclidean distance as the similarity measure and the nearest neighbor (NN) classifier.

Later on, Agrafioti *et al.* in [15], investigated the possibility of subject recognition under two types of cardiac irregularities commonly encountered in practice. These two types of premature heart beats were the premature ventricular contraction (PVC) and the atrial premature contraction (APC) which unlike the serious types of irregularities are not lethal but indicate some abnormalities of the heart.

A PVC results in beats whose appearance deviates substantially from healthy heart beats by inhibition of the next normal beat and introduction of a pause in the heartbeat cycle. In contrary to the PVC, the structure of the AC does not significantly differ of that of a normal heartbeat. Since the AC results in pulses that are morphologically healthy but occur earlier in time than expected.

The proposed methodology follows the same steps as was taken in the previous work [14], with the added stage of arrhythmia screening. Before extracting features for classification, various windows from the ECG records are examined and only those with a healthy structure were verified. Through a careful detection of irregular windows, the algorithm discarded those segments containing abnormal AC.

In order to detect the PVC windows, a set of criteria concerning the power distribution and complexity of ECG signals were described. After removing the PVC windows, classification was performed on the healthy AC windows like before [14]. Therefore, the application of AC heartbeats in subject recognition became technically feasible.

Experimental results over the MIT -BIH Normal Sinus Rhythm [49], MIT -BIH Arrhythmia [50], and the PTB [51] databases indicated a window recognition rate of 96.2%.

However, in the light of frequency characteristics of noise and the ECG components that were mentioned in the previous section, it is obvious that the approaches described

above suffer from some deficiencies in handling noise and different heart rates since a simple Butterworth filter with cut-off frequencies of $1Hz - 40Hz$ is not capable of removing the noise that overlaps with the spectra of the ECG signal. On the other hand, a rapid change in the heart rate usually shifts the ECG spectrum from the predefined frequency band dictated by the fix cut-off frequencies of the Butterworth filter [18]. Consequently, the recognition rate of the system degrades as a result of ECG's morphology alternation caused by discarding information above 40 Hz while inserting low frequency noise above $1Hz$ into the spectra band. This problem is dealt within a medical context using tedious and complex approaches which deploy filters specially designed to remove specific noise effects, as well as application of linear, time-variant filters in order to couple the cut-off frequencies to the prevailing HR [18].

Furthermore, in the area of ECG processing, development of accurate and robust methods for automatic ECG delineation has been under study [21] - [48]. Martinez *et al.* [21] presented a robust ECG delineation method based on discrete wavelet transform (DWT) over a wide range of morphologies that was able to handle different noises and outperformed the results of other well known algorithms, especially in determining the end of the T wave. Evaluation of the proposed approach using standard manually annotated ECG databases such as MIT-BIH Arrhythmia [50], QT [52], European ST-T [53], and CSE multilead measurement [54] databases demonstrated a robust performance over 99.66% in delineation of ECG components.

Motivated by these studies, in this thesis we propose a wavelet based approach for automatic analysis of single lead ECG for application in human recognition. The proposed system has two major consideration: first, utilizing a robust preprocessing stage that reduces the noise effects and allows for detection of the affected heartbeats and excluding them from further analysis. Second, delineation and standardization of the heartbeats that enables us to design a strong personalized heartbeat template for each

subject regardless of the heart rate, noise, and intra-class variations.

The design of the heartbeat template, which is the main novelty of the proposed approach, serves the system in three ways: first, it makes it capable of handling ECGs regardless of the heart rate which renders making presumptions on the individual's emotional condition unnecessary. Second, it reduces the size of the gallery set significantly since it consists of only one heartbeat template for each subject; thus, decreases the storage requirements of the system. Third, it increases the classification speed since only one template is compared against for each subject.

Chapter 3

ECG analysis for Human Recognition

In this chapter, a new wavelet based framework is developed for automatic analysis of single lead electrocardiogram (ECG) for application in human recognition. The proposed system is applicable to both identification and verification modes. Like any standard recognition procedure, our approach includes three major stages i.e., preprocessing, feature extraction, and classification, the respective descriptions of which are provided in the following sections. In order to provide the basic material for implementation of the proposed algorithm, the first section of this chapter is dedicated to a brief introduction to the discrete wavelet transform and the mother wavelet used in this work.

3.1 Introduction to Discrete Wavelet Transform

The wavelet transform provides a decomposition of the signal over a set of basis functions, obtained by dilation and translation of a mother wavelet by a scale factor s and a translation parameter u. The obtained family of time-frequency atoms is represented as

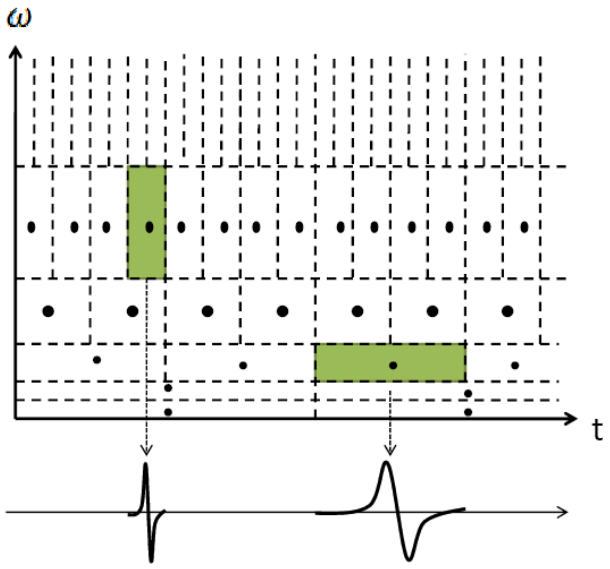

Figure 3.1: The time-frequency boxes of a wavelet basis define a tiling of the time frequency plane.

follows:

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}}\psi\Big(\frac{t-u}{s}\Big) \tag{3.1}$$

For discrete-time signals, the scale factor and/or the translation parameter are discretized. Such transform is called *dyadic wavelet* if a dyadic grid, as in Eq. 3.2, on the time-scale plane is followed. The time-frequency plane of such decomposition is shown in Figure 3.1.

$$s = 2^j, u = k2^j \qquad (j,k) \in Z^2 \tag{3.2}$$

In this way, the frequency axis is decomposed in dyadic intervals whose sizes have an exponential growth. For larger scales, the basis function is wider and the corresponding coefficient gives information about the lower frequency components of the signal. On the contrary, decreasing the scale, the higher frequency components of the signal are obtained. Therefore, the higher temporal resolution is achieved at the high frequencies, while the higher frequency resolution is achieved through the wider temporal windows.

According to *Mallat's algorithm* [20] the discrete wavelet transform, defined in Eq. 3.3,

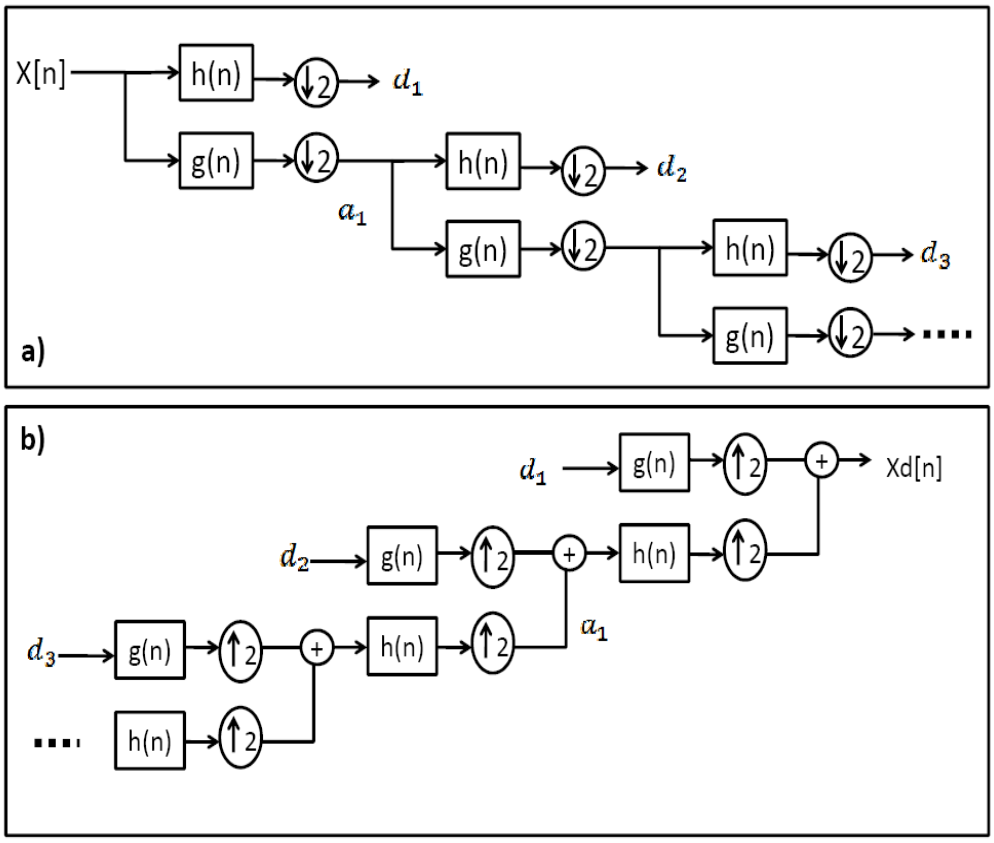

Figure 3.2: Mallat Algorithm: implementation of DWT using an octave filter bank, (a) decomposition filter bank, (b) reconstruction filter bank.

can be implemented by cascading identical low-pass and high-pass FIR filters as an octave filter bank, Figure 3.2(a). Using the wavelet coefficients and the low-pass residual, the original signal can be rebuilt using a reconstruction filter bank as illustrated in Figure 3.2(b).

$$Wx(u,s) = x(t) * \psi_s(t) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} x(t)\psi\Big(\frac{t-u}{s}\Big)dt \tag{3.3}$$

$$where: s = 2^j, b = k2^j \qquad (j,k) \in Z^2$$

The downsampler after each filter in Figure 3.2(a) removes the redundancy of the signal representation, an effect which is later compensated in the reconstruction phase, Figure 3.2(b), using an upsampler. As side effects, changing the sampling frequency of

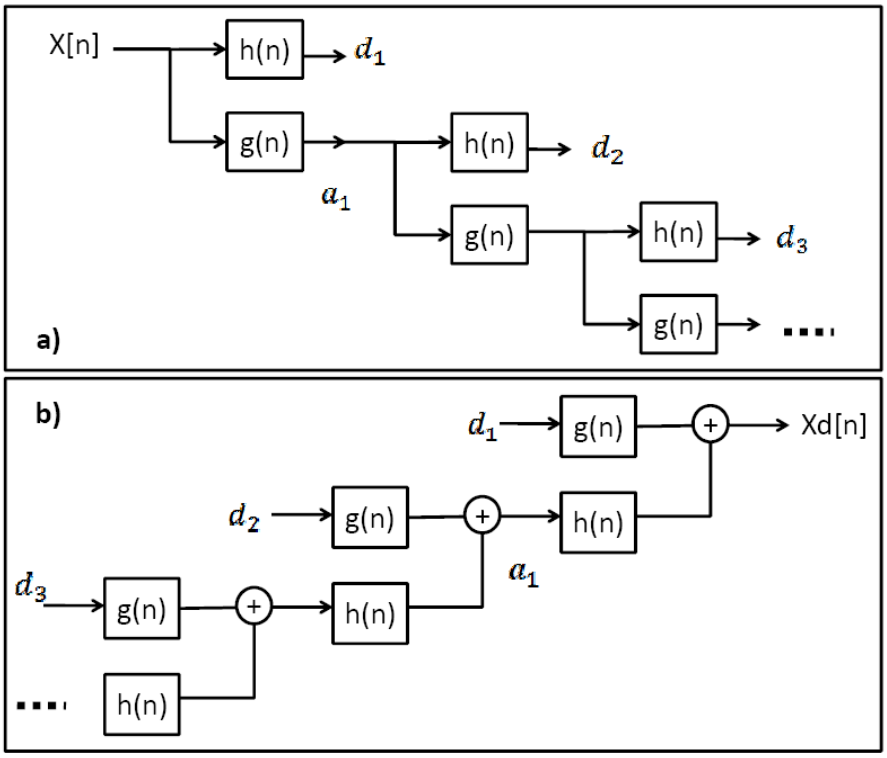

Figure 3.3: Algorithme trous: implementation of DWT without decimation, (a) decomposition filter bank, (b) reconstruction filter bank.

the signal makes its representation time-varying, and reduces the temporal resolution of the wavelet coefficients at higher scales.

In order to keep the time-invariance property and the temporal resolution at different scales, the decimation and interpolation stages are removed. Therefore, the scale s is discretized using the dyadic grid but not the translation parameter u. Fast computation of the wavelet coefficients is carried out using an algorithm called *algorithme trous* [20]; the corresponding decomposition and reconstruction filter banks of which are shown in Figure 3.3.

Providing a description of the signal in the time-frequency domain, DWT allows for the representation of the temporal features of a specific frequency band of a signal. Moreover, investigation of the nature and frequency content of different noise and artifacts affecting the ECG signal assesses their contribution to various scales. On this ground,

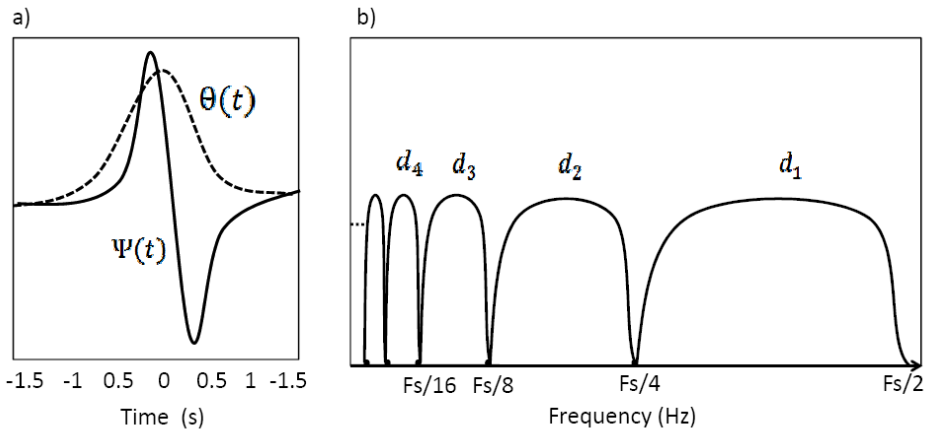

Figure 3.4: a) Spline wavelet ($\psi(t)$) and related smoothing function ($\vartheta(t)$), b) corresponding equivalent frequency responses of the DWT.

in the current implementation, DWT provides a dual functionality, on one hand, it provides noise reduction and on the other hand, using the corresponding information of the maxima lines obtained from the DWT coefficients, heartbeat selection and delineation is performed. The mother wavelet implemented in this work, depicted in Figure 3.4(a), is the *Quadratic Spline* wavelet due to its compatibility with the ECG signal [21, 22]. The selected prototype wavelet, the related Fourier transform shown in Eq. 3.4, corresponds to the following decomposition low-pass filter $h(n)$ and high-pass filter $g(n)$:

$$\psi(\Omega) = j\Omega\Big(\frac{\sin(\frac{\Omega}{4})}{\frac{\Omega}{4}}\Big)^4 \tag{3.4}$$

$$h(n) = \frac{1}{8}(\delta[n+2] + 3\delta[n+1] + 3\delta[n] + \delta[n-1]) \tag{3.5}$$

$$g(n) = 2\ (\delta[n+1] - \delta[n]) \tag{3.6}$$

It is shown in the literature [20] that the mother wavelet can be considered as the derivative of a scaling function. Therefore, the wavelet transform at scale s is proportional to the derivative of the filtered version of the signal with a scaling impulse response at that scale. As a result, the zero-crossings of the wavelet coefficients correspond to the local maxima or minima of the smoothed signal across the scales. Moreover, the maximum

slopes of the filtered signal are associated with the maximum absolute values of the wavelet coefficients. Since each heartbeat is composed of the slopes and local maxima (or minima) occurring at different time instants within the cardiac cycle, the characteristic of the wavelet coefficients across different scales can determine the expression of the ECG signal.

Based on this concept, in this work we propose a wavelet based approach for automatic analysis of single lead ECG for application in human recognition. The proposed system has two major consideration: first, utilizing a robust preprocessing stage that reduces the noise and allows for detection of the affected heartbeats and excluding them from further analysis. Second, delineation and standardization of the heartbeats that enables us to design a strong personalized heartbeat template for each subject regardless of the heart rate, noise, and intra-class variations.

The design of the heartbeat template, which is the main novelty of the proposed approach, serves the system in three ways: first, it makes it capable of handling ECGs regardless of the heart rate which renders making presumptions on the individual's emotional condition unnecessary. Second, it reduces the size of the gallery set significantly since it consists of only one heartbeat template for each subject; thus, decreases the storage requirements of the system. Third, it increases the Fclassification speed since only one template is compared against for each subject.

Like any standard recognition procedure, our approach includes the preprocessing, feature extraction, and classification stages; the respective descriptions of which are provided in the following sections.

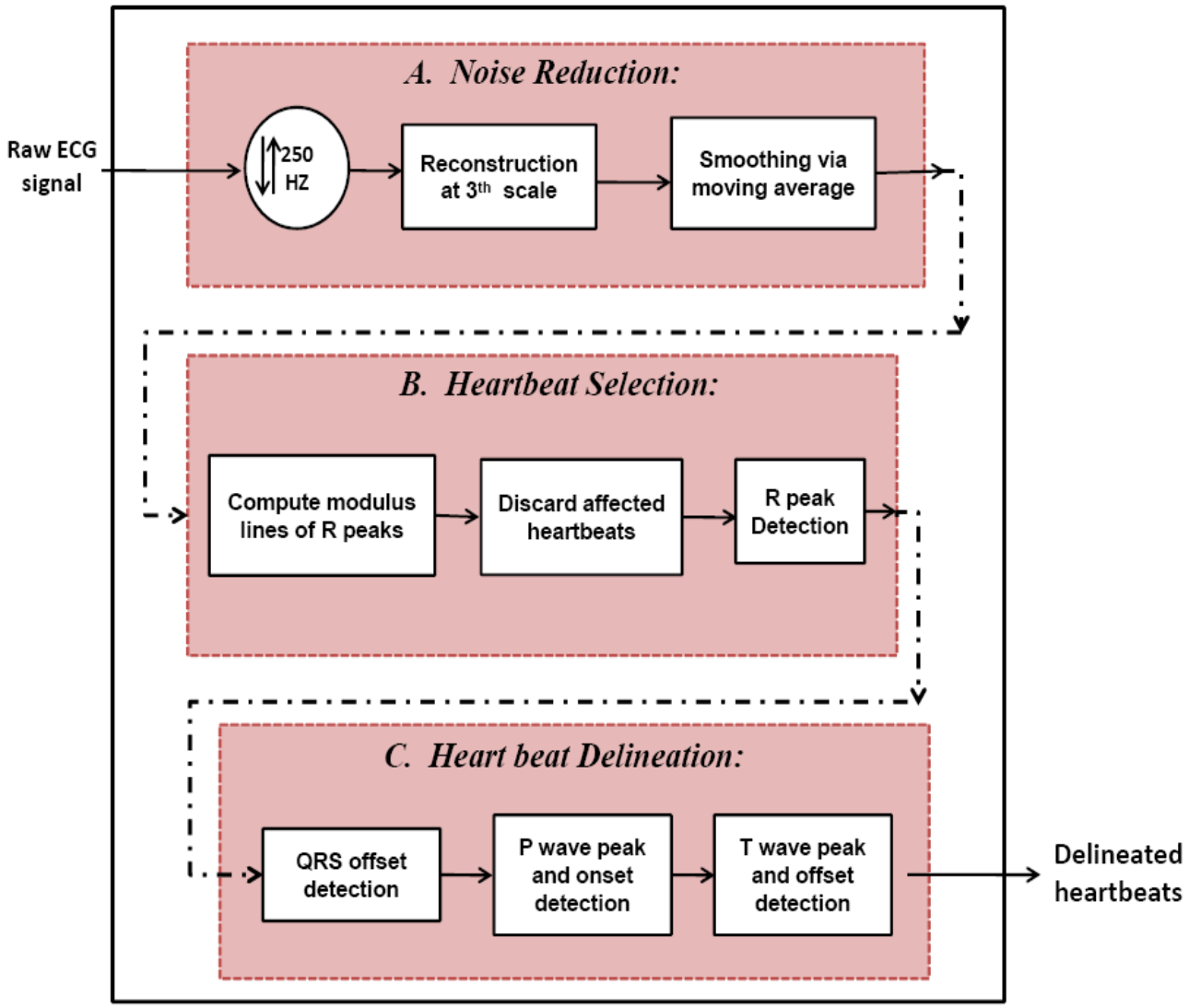

Figure 3.5: Block diagram of the preprocessing stage

3.2 Preprocessing

As mentioned before, noise and artifact components alter the expression of the ECG trace from the expected structure which leads to inaccurate interpretation and recognition. In order to have a better understanding of the raw ECG signals, some examples of the recorded ECGs affected by various noise components are depicted in Figure 3.6.

The main objective of the proposed preprocessing procedure is the detection and preparation of employable heartbeats for template extraction. In order to fulfill this aim

the preprocessing include three main steps: 1) noise reduction, 2) heartbeat selection, and finally, 3) heartbeat delineation, Figure 3.5.

The first step is designed to reduce the overall amount of noise contained in the signal by application of some time and frequency domain noise reduction techniques. However, due to the overlapping nature of noise with the ECG, some heartbeats are still affected. It is crucial to exclude such heartbeats from the template design. Therefore, in the next step the affected heartbeats as well as those heartbeats deviating from the predominant morphology detected for each subject are detected and discarded. This ensures us that only the waves without cardiac disorders are incorporated in the template design.

Finally, in the last step the remaining heartbeats are delineated based on determination of their onset and offset using DWT. The respective description of each of these steps is presented in the following subsections.

3.2.1 Noise Reduction

Noise reduction is the first step of the preprocessing stage, shown in Figure 3.5, and includes the following blocks:

- **Resampling to** $250Hz$.

  This block is needed to handle the recordings from different apparatus with various sampling rates. This sampling rate is found to be reasonable considering the frequency spectrum of the ECG signal, which is approximately in the range of $0.5Hz - 40Hz$, and the required Nyquist rate.

- **Reconstruction of the Signal using DWT**.

  According to the spectrum of the ECG signal, illustrated in Figure 3.7(a), and the corresponding equivalent responses of the DWT with *Quadratic Spline* wavelets,

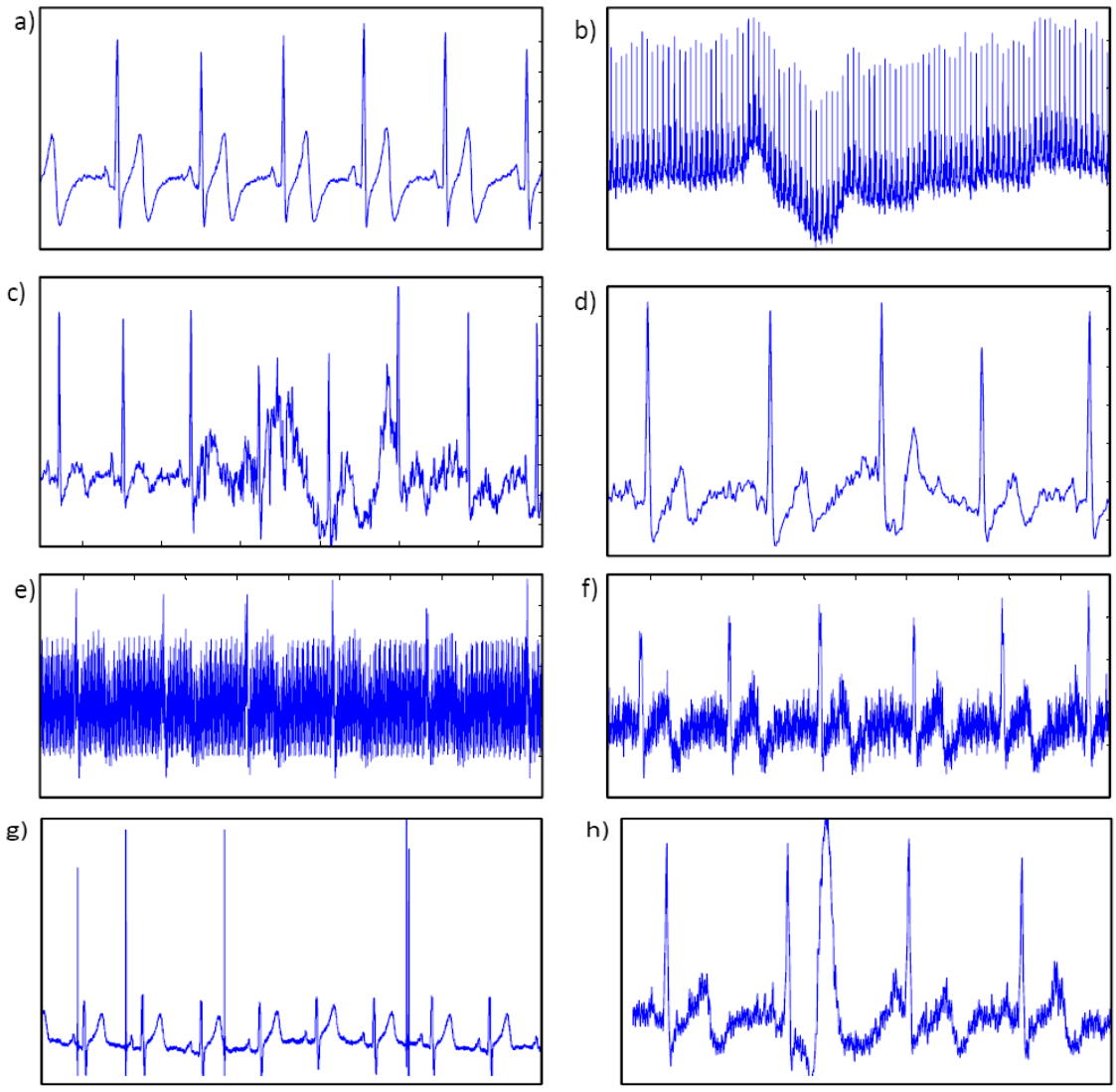

Figure 3.6: Raw ECG signals affected by different types of noise or cardiac disorders: (a) clean ECG, (b) baseline wander, (c) sudden body movement, (d) cardiac disorder, (e) high frequency noise, (f) power line inference, (g) spikes caused by motion artifacts, (h) skin scratching.

depicted in Figure 3.7(b), it is clear that most of the energy of the ECG signal is found up to the *3rd* scale. Thus, reconstruction of the signal at this scale while preserves its main frequency components, eliminates the effects of high frequency noise and power line interference. Therefore, in the next step of our implementation the ECG signal is reconstructed using the dual reconstruction filter bank at the *3rd*

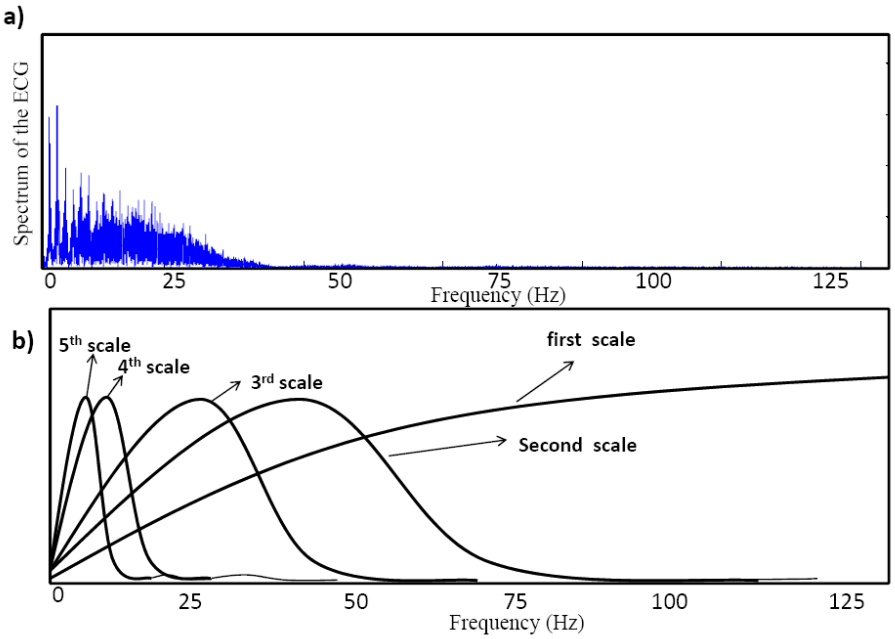

Figure 3.7: (a) Frequency spectrum of the ECG signal, (b) equivalent frequency responses of the DWT at scales 1-5 for sampling rate of 250 Hz.

scale.

- **Smoothing via Moving Average**.

  In order to eliminate the effect of high frequency ripples on the obtained signal, the ECG trace is then further smoothed down using a moving average filter which performs local regression with the weighted linear least square and the *2nd* degree polynomial model. The method assigns lower weight to outliers in the regression and zero weight to data outside six mean absolute deviations.

3.2.2 Heartbeat Selection

Due to the overlapping nature of the noise with ECG, absolutely clearing the ECG signal requires tedious and complex approaches which deploy filters specially designed to remove

specific noise components [18]. Instead, it is possible to detect the affected heartbeats and exclude them from the feature extraction. The heartbeat selection is performed using the modulus maxima lines extracted from DWT coefficients. Essential information is carried in the modulus maxima lines across the scales, analysis of which allows the detection of different irregularities of a signal [20]. A brief description of the modulus maxima lines is provided in the appendix. For further illustration, some possible irregularities of a continuous function as well as an ECG signal along with their corresponding maxima lines are depicted in Figure 3.8.

According to the shape of the *Quadratic Spline* and the wavelet coefficients, the relationship between different components of the ECG signal and the corresponding maxima lines, plotted in Figure 3.9, is empirically concluded as follow [21–25]:

1. Each sharp change in the signal is associated with a line of maxima or minima across the scales.

2. Every uni-phase wave, such as (a), (b), (c), and (d), corresponds to a positive maximum negative minimum pair of maxima lines across different scales.

3. The wave rising edge corresponds to a negative minimum line, whereas the dropping edge corresponds to a positive maximum line across the scales.

4. The peak of a uni-phase wave corresponds to the zero-crossing point of the positive maximum-negative minimum line pair.

5. The Q and S wave peaks have zero crossings associated with the maxima lines mainly at scales 2^1 and 2^2.

6. The P or T like waves have their major component across scales 2^1 to 2^5.

7. Artifacts produce isolated maximum or minimum lines.

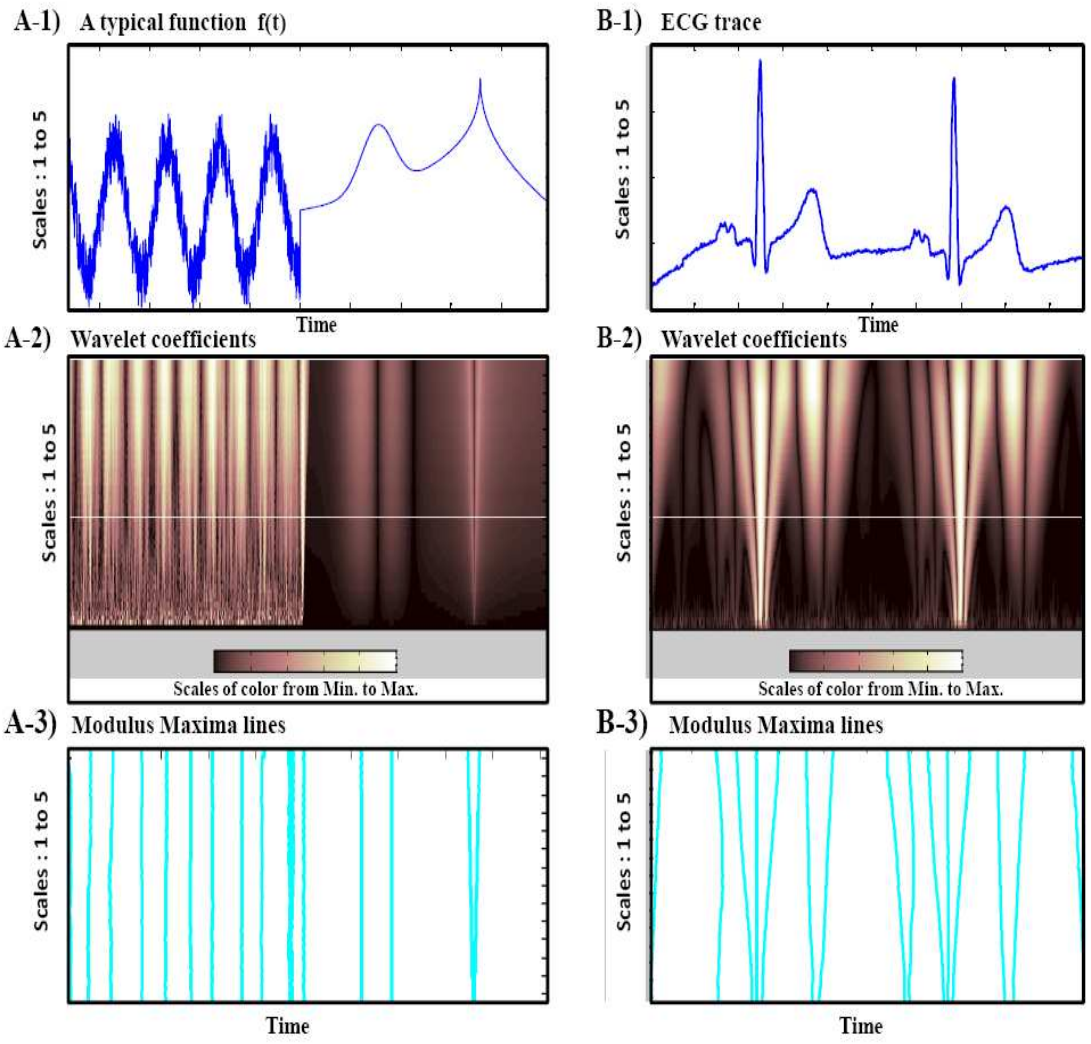

Figure 3.8: The corresponding wavelet coefficients and modulus maxima lines of: (a) a typical function, $f(t)$, (b) the ECG signal.

8. High-frequency noise, (e), affect mostly the 2^1 and 2^2 scales; the higher scales are essentially immune to this sort of noise.

9. Baseline wander affects only at scales 2^4 and 2^5 .

Using this information, the heartbeat selection is performed by determination of the modulus lines corresponding to the potential R peaks

Each heartbeat is identified by the location of its R peak which has the largest amplitude of the ECG waveforms. Therefore, in the first step, the R peak location is identified;

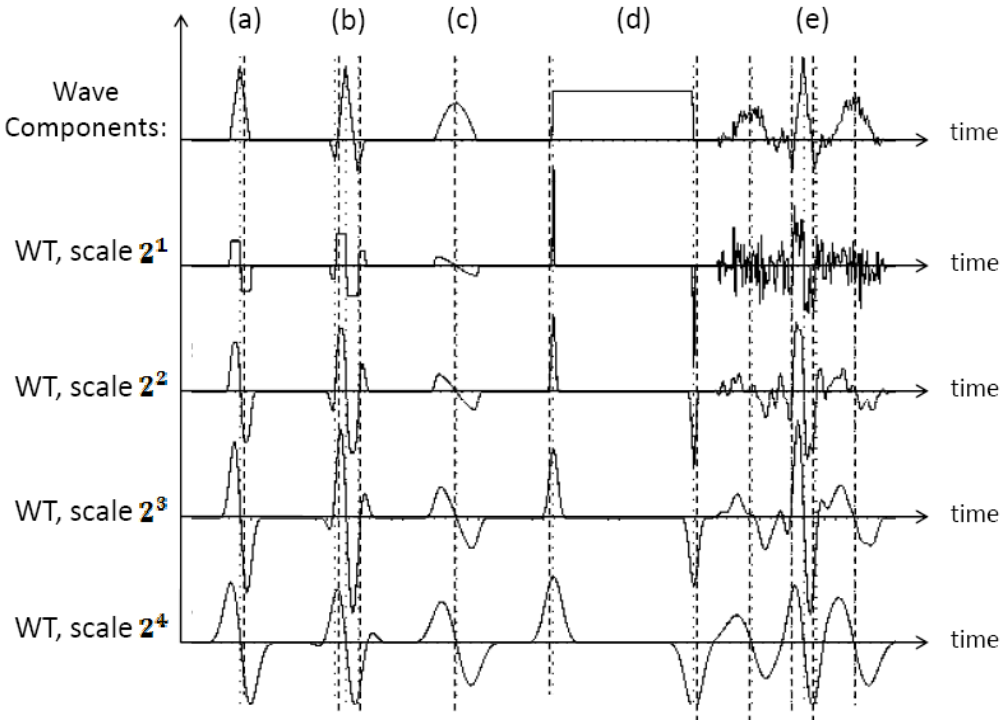

Figure 3.9: Different components of the ECG wave along with the corresponding wavelet coefficients across the scales $1:4$ [21].

based on the position of the R peaks, other heartbeat's components are detected respectively. The required algorithm for detecting the modulus maxima lines corresponding to the potential R peaks can be found in the appendix.

In the detected location set of modulus maxima lines, there may be those which are created by various noise components rather than having a physiological origin. Therefore, in the next step these maxima lines which are usually isolated or redundant are detected, so that the the heartbeats containing them are removed by discarding the corresponding R peak maxima lines. The algorithm required for detection of the isolated and redundant maxima lines is described in appendix.

Moreover, two thresholds are defined to exclude the maxima lines corresponding to the motion artifacts. One, based on the amplitude of the ECG signal, and the other based on the interval between a pair of positive maximum-negative minimum. As mentioned

in chapter 2, the R peak has the largest amplitude among the other ECG components which at most can reach to $2mV - 3mV$; however, motion artifacts are manifested as large-amplitude waveforms sometimes overriding the ECG signal. Therefore, defining a threshold on the ECG trace, the maxima lines corresponding to the points with an amplitude greater than this threshold are discarded. In this work, this threshold is empirically to $3mV$.

Furthermore, the R wave is a uni-phase wave that corresponds to a positive maximum-negative minimum pair at each characteristic scale. The interval between these two modulus maxima lines should be less than the width of the QRS complex. Not to lose the wide QRS complexes, an interval threshold is selected based on the interval of the two modulus maxima created by the widest possible QRS complex. In this work, the interval threshold is empirically defined to $120ms$.
After rejecting all maxima lines associated with noise and artifacts, the zero crossings of the $\left[W_{2^1 x(n)}\right]$ between the pairs of positive-maximum (negative-minimum) are identified as the R peaks and labeled as the set of N_r if they exceed the following thresholds, Eq. 3.7, across the scales $2^1 - 2^4$.

$$\begin{cases} \epsilon^i_{QRS} = \text{RMS}(W_{2^i} x[n]) & , i = 1, 2, 3 \\ \epsilon^4_{QRS} = 0.5 \text{ RMS}(W_{2^4} x[n]) \end{cases} \tag{3.7}$$

Both positive and negative R morphologies are allowed, but the vast majority type of the detected R morphology for each subject is considered as his/her R peak type (positive or negative). Other types of detected R peaks that deviate from the predominant morphology are removed. In this way, only the waves without cardiac disorders are retained for feature extraction. An example of the obtained ECG signal after each of these preprocessing steps is illustrated in Figure 3.10.

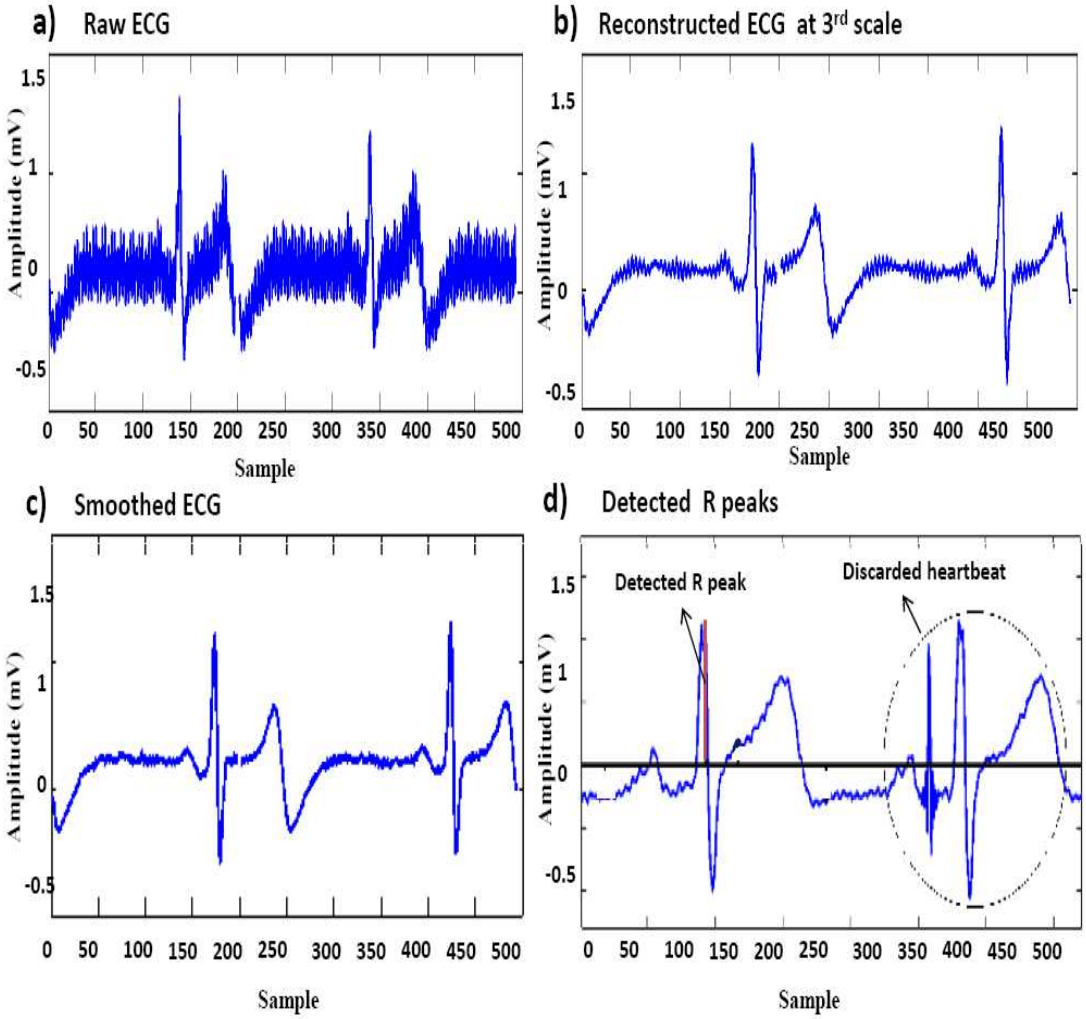

Figure 3.10: The obtained ECG signal after preprocessing steps (a) the raw ECG signal, (b) reconstructed ECG using WT coefficients up to *3th* scale, (c) smoothed ECG trace, (d) discarded heartbeats and localization of the R peaks.

3.2.3 Heartbeat Delineation

Regarding our application, the underlying idea of this stage is to delineate each heartbeat for the template design. Therefore, the onset of the P wave and the offset of the T wave of each heartbeat should be localized. Moreover, in order to design a heart rate independent template it is necessary to localize the offset of the QRS complex as well. Having obtained the R peak locations and inspired by findings in [21, 22], identifying

these boundary points is performed using a simple approach described below. The proposed algorithm, which is designed based on the characteristics of a healthy heartbeat, accurately determines the boundaries of the heartbeats without requiring the various complicated thresholds considered in [21].

- *QRS Detection and Delineation*:

  At first, around each R peak, a temporal interval surrounding the QRS complex is considered (sw_{QRS}). For considering the widest QRS complexes, the interval length is empirically defined as $120ms$. Then, the algorithm identifies the immediate pair of opposite-sign maximum modulus lines after each N_r point at scale 2^2 representing the S wave. The first sample after this pair is labeled as N_{post}. Defining a temporal search window after N_{post},(sw_{RS}), the QRS offset is determined by finding the nearest local maxima of $|\mathrm{W}_{2^2}x\,[n]\,|$ after the last peak, within the (sw_{RS}), which is followed by a local minimum of $|\mathrm{W}_{2^2}x\,[n]\,|$. In case no such point is detected the last point of the search window is considered as the offset of the QRS complex.

- *Onset of the T wave*:

  Subsequent to each QRS complex, a search window (sw_T) depending on a recursively computed RR interval is defined at scale 2^4. The area of the T wave is considered within this window where the $|\mathrm{W}_{2^4}x\,[n]\,|$ exceeds the threshold ϵ_T. The wave offset is identified as the last point of this interval.

$$\epsilon_T = 0.04\ \mathrm{RMS}(W_{2^4}x[n]) \tag{3.8}$$

- *Offset of the P wave*

The P wave delineation algorithm is similar to that of the T wave using an appropriate RR-dependent search window and adequate threshold ϵ_P.

$$\epsilon_P = 0.01 \text{ RMS}(W_{2^4} x[n]) \tag{3.9}$$

Regarding our application, a guard band of a few milliseconds is added to the detected onset of P wave as well as the end of the T wave in order to obtain a complete heartbeat trace.

3.3 Template Design

Having acquired the boundaries of each heartbeat, the next step toward recognition is feature extraction. As mentioned before, one of the challenges in utilizing ECG for recognition purposes is heart rate variability that alters the expression of the ECG morphology. Regarding our application, we are interested in designing a heartbeat template representing each subject out of his/her recorded ECG trace. Therefore, clearing off the inter-individual variability among the ECG recordings of a subject is an important issue that allows the assessment of the ECG morphology.

According to the literature [18], position of the T wave strongly depends on the heart rate, becoming closer to the QRS complex at rapid rates; however, this contraction property does not apply to the P wave or QRS complex making the temporal length of the PQ segment as well as the QRS complex weakly dependent on the heart rate.

In view of this fact, heartbeat standardization is carried out in the manner illustrated in Figure 3.11. The heart rate dependent section of each heartbeat starting from the end of the QRS complex up to the end of the T wave is resampled to $120ms$, which is the common length for normal T wave under the rest [18]. After resampling, this segment is again combined with the heart rate independent section including the P wave and the QRS complex to make the whole heartbeat. The adaptation procedure is concluded by

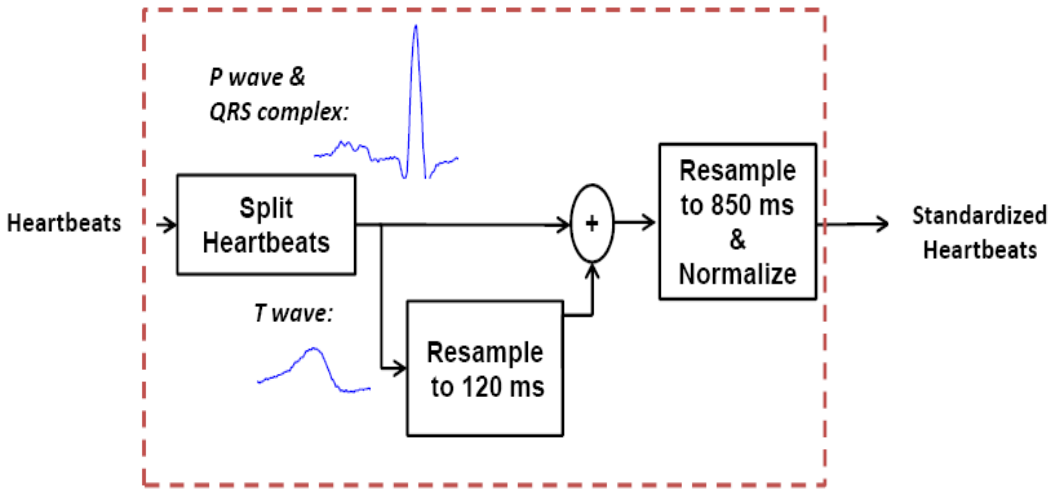

Figure 3.11: The block diagram of the template extraction stage.

resampling the whole heartbeat to 850ms, which is the normal heartbeat length under rest, and normalizing each heartbeat to have zero mean and unit variance.

At this point, the heartbeat template of each subject, representing his/her unique repetitive heartbeat morphology, can be obtained using the median of the aligned heartbeats, Figure 3.12. Considering the processing steps taken so far, the ensemble median is empirically found to be a plausible estimator since it captures the repetitive components of heartbeats while prevents insertion of the possible outliers into the template.

3.4 Classification

As in most pattern recognition problems, classification among a gallery set is the last step of the recognition process. One of the novelties of this work with respect to previous studies is the constitution of the gallery set. In the current approach, it is comprised of only *one* template for each subject rather than every heartbeat associated with it. Therefore, while retaining all discriminative information among the data captured from

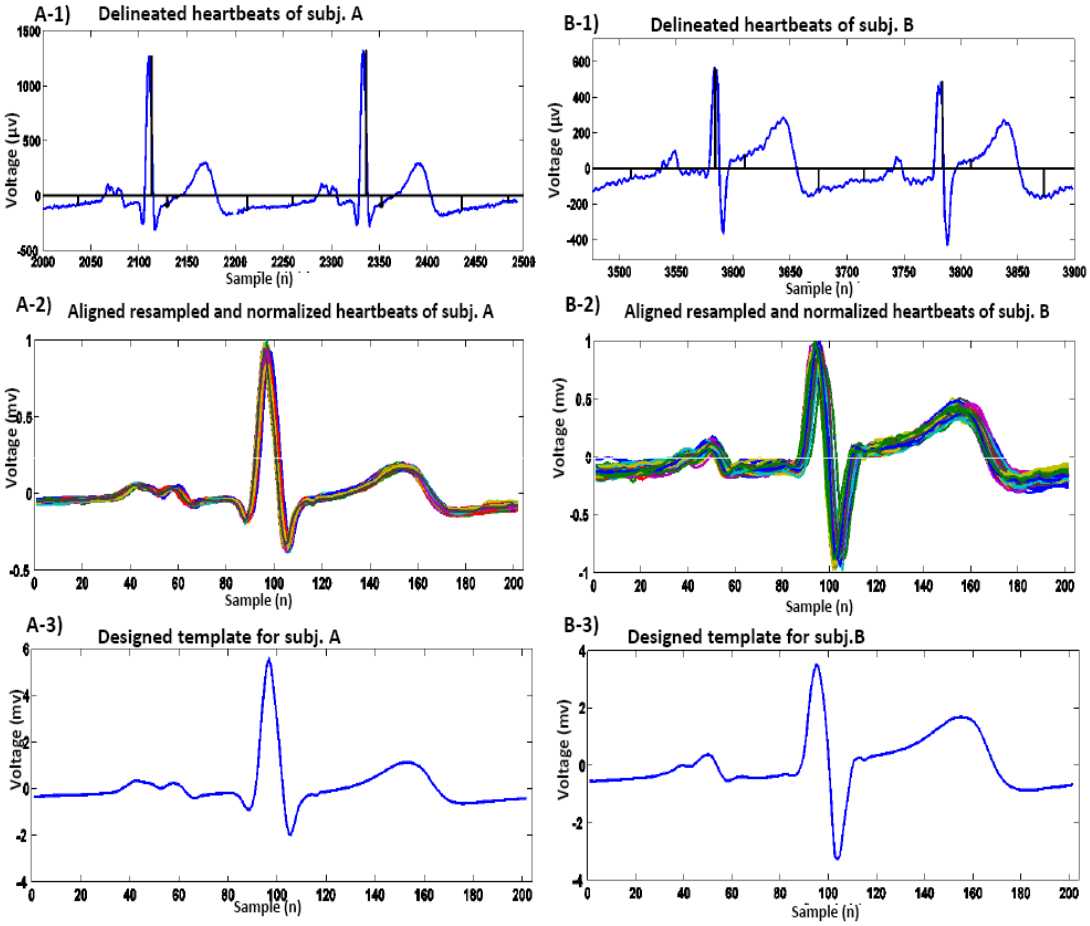

Figure 3.12: Template extraction procedure of 2 subjects (A and B), 1) Delineated heartbeats of each subject, 2) Aligned resmpled and normalized heartbeats of each subject during a recording session, 3) Corresponding extracted templates.

subjects, the size of gallery set is reduced substantially. This dramatic reduction of the gallery size serves the system in two ways: first, by decreasing the storage requirements of the system which is a major consideration in performance of any large scale biometric system; second, by speeding up the classification since the classifier needs to match only one template per subject.

All heartbeats recorded for a subject during the enrolment are incorporated in the design of a template to be stored in the gallery set. Since there is a direct relationship between the number of individuals as well as the access claims within a system and

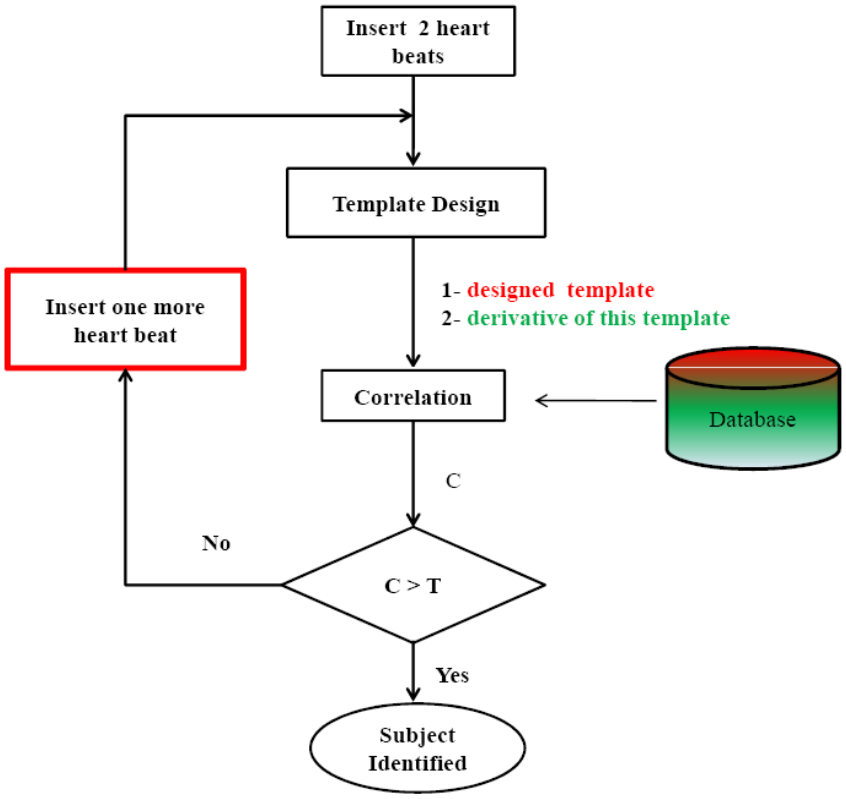

Figure 3.13: Flowchart of the classification procedure.

the number of interrogations, it is important to mitigate the data handling process by minimizing the number of descriptors for a given subject. Therefore, the templates to be used as the probe set are obtained using as small number of heartbeats as possible in an iterative manner depicted in Figure 3.13.

At first, using 2 heartbeats per subject, the heartbeat template is designed and compared with all templates stored in the gallery set. The comparison is performed using a fusion of two criteria: the maximum correlation between the templates, Eq. 3.10, and the maximum correlation between the first derivatives of the templates.

$$\hat{\rho}_{xy}[m] = \frac{\sum_{i=0}^{N-|m|-1} x[i]y[i+m]}{\sqrt{E_x \cdot E_y}} \tag{3.10}$$

where x and y represent two different heartbeat templates with the energy content of E_x and E_y respectively.

The maximum correlation metric, which exhibits the amount of similarity between two signals, achieves good classification performance due to the fact that only one template is compared against for every subject, and through all the processing and standardization steps taken for computing the robust templates. Moreover, considering the correlation coefficient between the first derivatives of the templates ensures the system that the two templates are entirely similar to each other.

The matching template in the gallery set is considered as the associated subject if the correlation coefficient between the two subjects exceeds a predefined minimum similarity threshold. If the correlation is below this threshold, one more heartbeat is added so that a new template will be obtained using all input heartbeats of the subject to be identified, compensating the effect of probable distortions by incorporation of more heartbeats in template extraction.

The number of iterations depends on the constraints imposed by the application; the more incorporated heartbeats, the higher the precision of the designed template. In our implementation, the results are shown for up to 10 iterations; if still not recognized; the system fails to recognize the subject. Investigating statistics of the average number of heartbeats required for classification of a subject, reported in chapter V, indicates that 2 heartbeats (on average) are adequate for correct recognition. This achievement greatly reduces the enrolment time and computational effort of the system compared to previous works.

3.5 Conclusion

In this chapter, we developed a wavelet-based ECG recognition system for automatic analysis of single lead electrocardiogram (ECG) for application in human recognition. The proposed system, which is applicable to both identification and verification modes,

performs subject recognition in three analysis stages: 1) ECG preprocessing and delineation using the dyadic wavelet transform, 2) Design of the personalized templates representing each subject heartbeat's morphology regardless of the heartbeat rate, 3) Classification, based on the amount of similarity among the templates.

The proposed system has two major consideration for solving the challenges of ECG recognition in real life scenarios: first, to utilize a robust preprocessing stage that reduces the noise and allows for detection of the affected heartbeats and excluding them from further analysis. Second, to delineate and standardize the heartbeats in order to design a strong personalized heartbeat template for each subject regardless of the heart rate, noise, and intra-class variations.

The design of the heartbeat template, which is the main novelty of the proposed approach, serves the system in three ways: first, it makes it capable of handling ECGs regardless of the heart rate which renders making presumptions on the individual's emotional condition unnecessary. Second, it reduces the size of the gallery set significantly since it consists of only one heartbeat template for each subject; thus, decreases the storage requirements of the system. Third, it increases the classification speed since only one template is compared against for each subject.

Evaluation of the proposed approach over different databases in both identification and verification modes is presented in chapter 5.

Chapter 4

PhonoCardioGram (PCG): A New Biometric

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 the heart sounds such as the intensity, frequency, duration, number, and quality of the sounds. It is, however, a difficult skill to acquire.

The term phonocardiography (PCG) refers to the tracing technique of the heart sounds and recording of the cardiac acoustics vibration by means of a microphone-transducer. Similar to ECG, PCG signal contains a large amount of information about the state of an individual's cardio-circulatory system. However, the traditional auscultation involves subjective judgment by the clinicians, which introduces variability in the perception and interpretation of the sounds and thereby affects the diagnostic accuracy. In recent years, considerable research has been conducted for the automatic analysis of PCG signals. The introduction of sophisticated electronic stethoscopes has provided excellent-quality digital PCG signals which are required for more objective analysis and interpretations. The idea of utilizing PCG as a physiological biometric is a new concept that is investigated in this chapter.

Understanding the nature and source of this signal is important for developing an effective tool for further analysis and processing. Therefore, the first part of this chapter offers an introduction to the PCG signal along with the temporal and frequency characteristics of different components of a typical PCG trace. A literature review of the previous works done in the area of PCG based human recognition is provided in the second part of this chapter. Finally, the last part of this chapter is dedicated to the introduction of a new PCG based recognition system.

4.1 Introduction to Phonocardiogram Signal

Heart auscultation is a fundamental tool in the diagnosis of heart diseases. Although nowadays it has been less focused due to the advent of ECG and echocardiography; still, there are some cardiac defects that are best detected by heart sounds.

The human heart is a four-chamber pump with two *atria* for the collection of blood from the veins and two *ventricles* for pumping out the blood to the arteries. The mechanical functionality of the cardiovascular system is governed by an electrical signal originated in specialized pacemaker cells in the right atrium (the *sino-atria node*), and is propagated through the atria to the AV-node (a delay junction) and to the ventricles. The periodic activity of the heart is created as a result of a complex interaction among pressure gradients, the dynamics of blood flow, and the compliance of cardiac chambers and blood vessels [55].

Two sets of valves control the flow of blood: the AV-valves (mitral and tricuspid) between the atria and the ventricles, and the semilunar valves (aortic and pulmonary) between the ventricles and the arteries. These mechanical processes produce vibrations and acoustic signals that can be recorded over the chest wall. The cardiac cycle events are demonstrated in Figure 4.1.

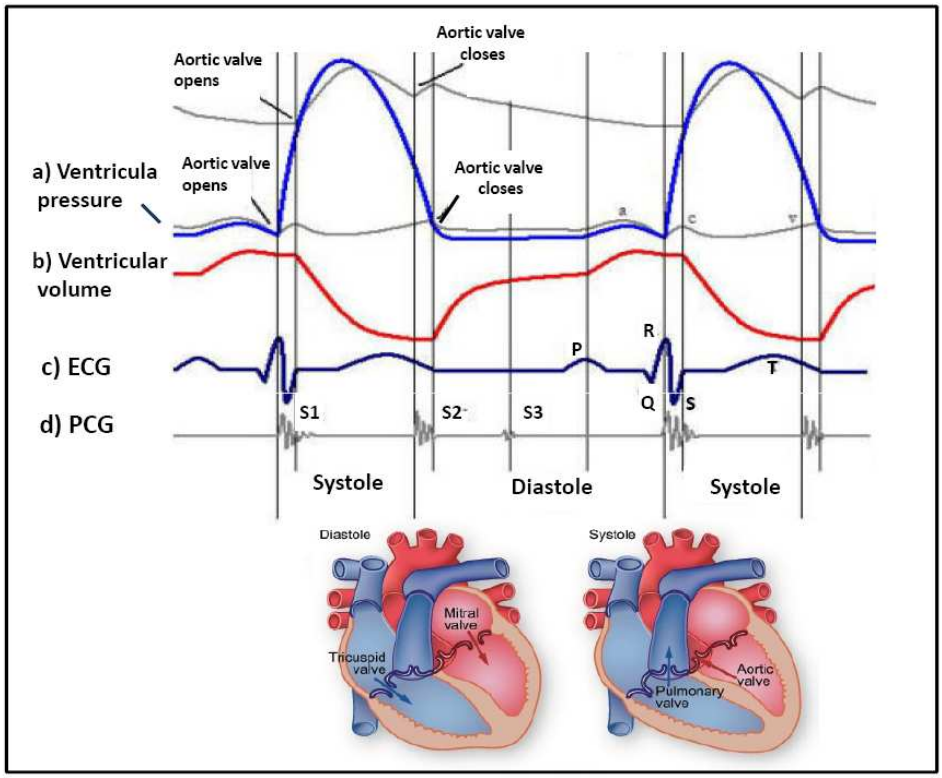

Figure 4.1: The cardiac cycle, (a) Ventricular Pressure, (b) Ventricular volume, (c) ECG trace, (d) PCG signal.

The vibroacoustic heart signals, including PCGs (phonocardiogram), apical pulse (apexcardiogram) and arterial pulse (e.g. carotid pulse) convey a wealth of information mostly used in qualitative clinical assessments. The important features that are usually assessed are the rhythm, timing instances, relative intensity of the PCG components, the splitting of *S2*, the existence of murmurs or other extra sounds and their quantity. The mechanisms that generate cardiac sounds are complex and there is, so far, no general consensus as to the contribution made by various mechanical cardiac events to the formation of the single components. However, it is agreed by most investigators that the mechanical activity of the heart including the blood flow, vibrations of the chamber walls and opening and closing of the valves are the major reasons for generation of the PCGs.

The cardiac cycle consists of two periods, systole and diastole respectively. Both are periods of relatively high activity, alternating with comparatively long intervals of

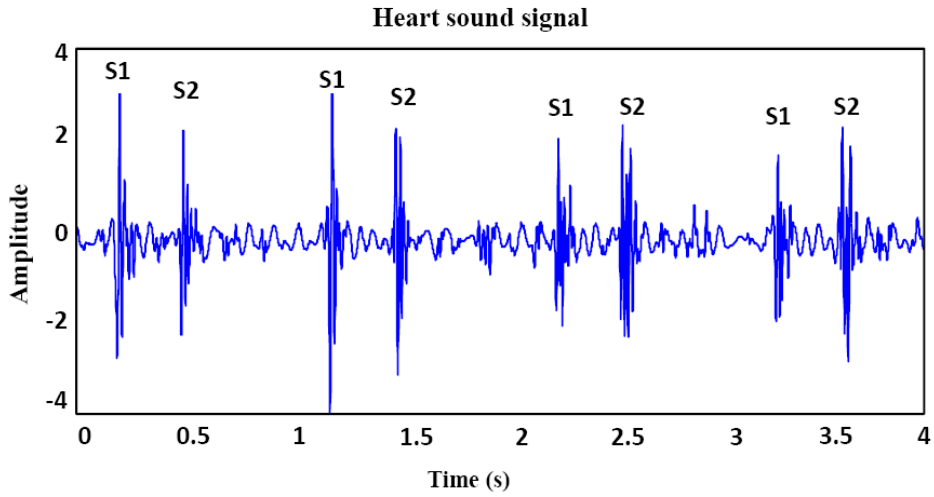

Figure 4.2: Different components of a normal PCG signal.

low activity. The major audible components of the PCG are short beats which are recognized as the primary components (*S1*, *S2*, *S3*, and *S4*). The other classes of sounds are murmurs, clicks, and snaps. However, the two major audible sounds in a normal cardiac cycle are the first and second heart sound, *S1* and *S2* as depicted in Figure 4.2.

- *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 $100ms - 200ms$ and its frequency components lie in the range of $25Hz - 45Hz$. It is usually a single component, but may be prominently split with some pathologies.

- *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*.

- *S3*: is the third low-frequency sound that may be heard at the beginning of the diastole, during the rapid filling of the ventricles. Its occurrence can be normal in young people (less than 35 years of age).

- *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.

- *Click 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 most common click is a systolic ejection click, which occurs shortly after *S1* with the opening of the semilunar valves. The opening snap when present, occurs shortly after *S2* with the opening of the mitral and tricuspid valves.

- *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.

Pathologies of the cardiovascular system occur due to different etiologies, e.g. congen-

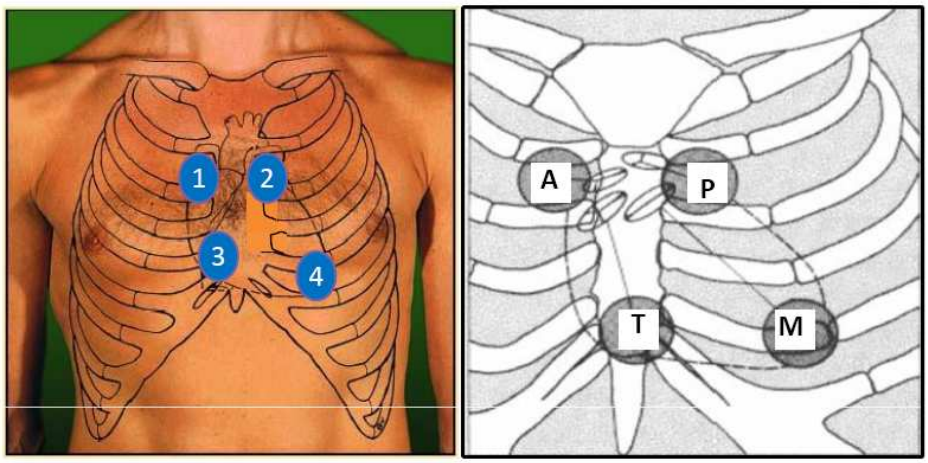

Figure 4.3: Four auscultation cites [55].

ital heart valve defects, stenotic valve, and regurgitated valve. These pathologies affect the normal sounds with respect to intensity, frequency content, and timing of components (splitting).

In a medical context the heart sound signal is collected from four main regions on the chest wall as demonstrated in Figure 4.3. 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.

However, presence of noise in recording the PCG signals is a major problem which is inevitable in practice. The noise contained in the PCG is introduced by a variety of sources such as ambient noise (e.g. background noise and stomach rumbling), respiratory sounds, thoracic muscular noise, peristaltic intestine noise, fetal breath sounds (if the subject is pregnant), non-optimal recording sites, weak sounds (obese patients), and so on.

The contribution of each source may vary significantly depending on the technical characteristics of the recording instrumentation, sensor detection bandwidth, the record-

ing environment, and the physiological status of the subject. The extraneous noises can be minimized in the vicinity of the patient during recording. Moreover, based on the type of noise and disturbance corrupting the recorded heart sound, various signal processing methods, such as notch filtering, averaging, adaptive filtering, and wavelet decomposition have been proposed in the literature for noise reduction [55].

In the next section a literature review of the approaches proposed so far for the PCG based human recognition is provided.

4.2 PCG As a New Biometric: Literature Review

Human heart sounds are natural signals, which have been thoroughly used for health monitoring and medical diagnosis for hundreds of years. So far, the study of PCG has focused mainly on the heart rate variability, characterization of the PCG components, detection of structural abnormalities, and heart defects [58] - [74]. However, demonstration of feasibility of applying ECG signal for human recognition has drawn attention of researchers to the PCG signals. Having the same origin with the ECG and considering the medical information which is conveyed through the PCG, conjectured that this 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. Similar to ECG, the PCG is difficult to disguise, forge, or falsify. Moreover, this signal has the advantage of being relatively easy to obtain, since it can be collected by placing a stethoscope on the chest.

Investigating some of the PCG's characteristics such as its ability to discriminate individuals as well as its stability over a period of time, Phua et al. [56] demonstrated the possibility of utilizing PCG signal for human recognition. They 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). The performance of the GMM and VQ methods were tested and evaluated with different parameter settings, namely the number of mixture components, the number of features, the type of feature sets (LBFC against MFCC), the frame length, and the frame shift. It was concluded that the Gaussian mixture modeling with four mixture densities (GMM-4) and a frame length of $256ms$ with no overlap can achieve a recognition rate of 96.01% over 20 subjects.

However, due to the number of iterations required to train the 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.

Another approach was proposed by Beritelli et al. [57] 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 [55].

Moreover, the position of stethoscope placement on the chest which gives the best performance in terms of the intra-class variability was investigated. Among the four typical auscultation regions, Figure 4.3, the recordings from the aortic and pulmonic regions exhibited a greater degree of separability among the classes.

In the march of the PCG recognition's progress, the proposed methodology for PCG recognition in this thesis is presented in the next section.

4.3 Recognition Methodology

Recognizing the PCG signal as a biometric, in this section, an approach based on both time and frequency properties of the PCG signal is developed for automatic analysis of this signal for application in human recognition. The proposed system is applicable to both identification and verification. Like any standard recognition procedure, the proposed approach includes three major stages i.e., preprocessing, feature extraction, and classification, the respective descriptions of which are provided in the following sections.

4.3.1 Preprocessing Using DWT

Raw PCG signal contains variety of artifacts and noise components that alter its expression from the expected structure. This renders further interpretation of the PCG signal impossible. In order to have a better perception of the raw PCG signal, some examples of the recorded PCG signals affected by noise are depicted in Figure 4.4. The main objective of the proposed wavelet based preprocessing procedure is to provide a noise reduction scheme for further analysis of the heart sound traces, Figure 4.5. The complex and highly non stationary nature of the heart sound signal is a real challenge in

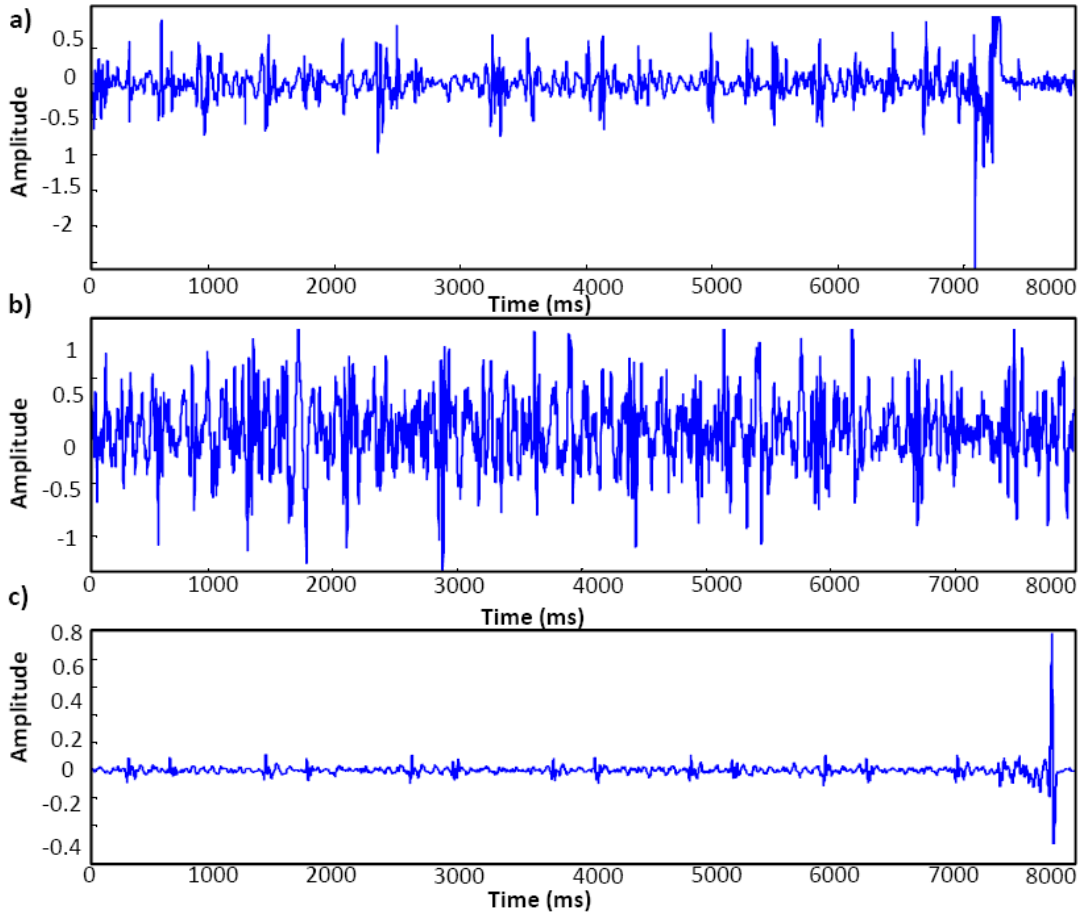

Figure 4.4: Examples of raw PCG signals, (a) usual recording, (b) contaminated with noise, (c) affected by spikes and artifacts.

automatic analysis. However, the discrete wavelet transform (DWT) provides a strong tool particularly applicable for analyzing problems of this type [20]. A large number of wavelet based approaches have been proposed in literature for detection and identification of the fundamental components of the PCG signal as well as classification and recognition of different cardiac pathologies [58] - [68]. Among various wavelets, the *Debauches* family is reported to have a good compatibility with the PCG signals [55]; the mother wavelet implemented in this work, depicted in Figure 4.6(a), is the *5th order Daubechies* wavelet. The Daubechies wavelets are a family of orthogonal wavelets that are not defined in closed form in terms of the resulting scaling and wavelet functions. The selected

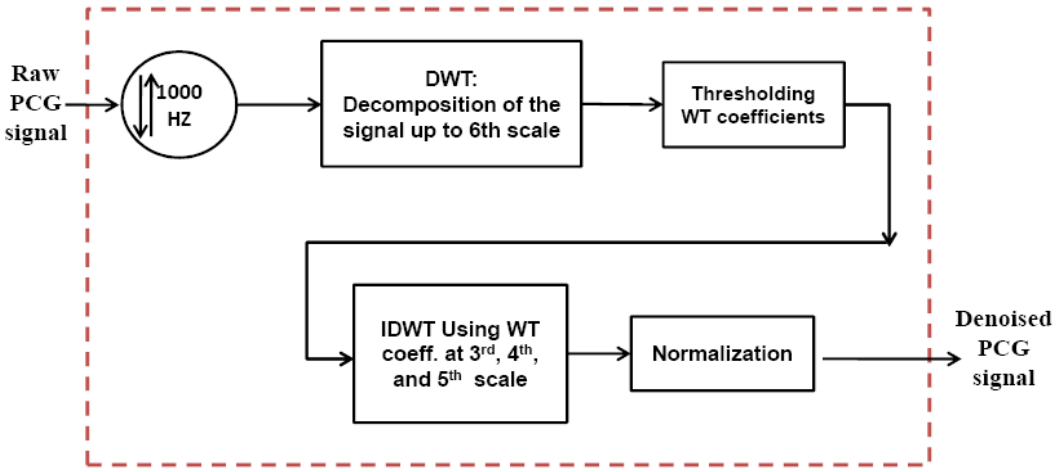

Figure 4.5: The block diagram of the preprocessing stage of the proposed PCG recognition system.

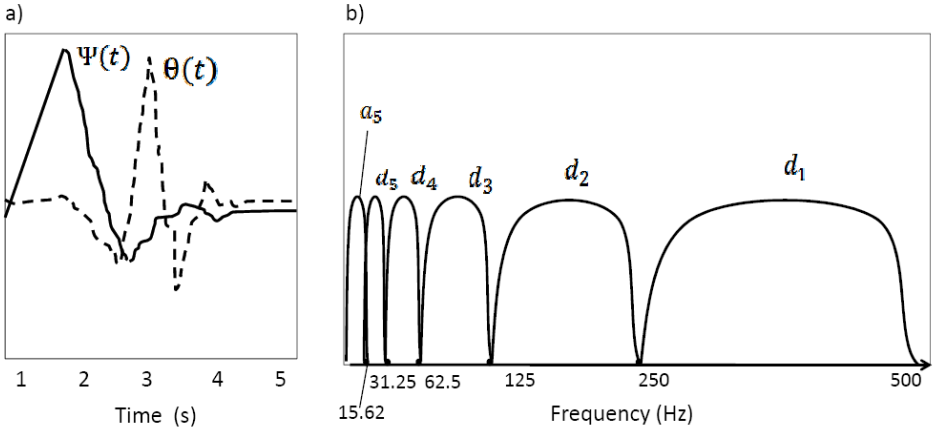

Figure 4.6: (a) 5th order Daubechies ($\psi(t)$) and related smoothing function ($\vartheta(t)$) (b) Equivalent frequency responses of the *DWT*.

prototype wavelet corresponds to the following decomposition low-pass filter $h(n)$ and high-pass filter $g(n)$:

$$h(n) = \sqrt{2}.(0.1132\delta[n] + 0.4270\delta[n-1] + 0.5122\delta[n-2] + 0.0979\delta[n-3] -$$
$$- 0.1713\delta[n-4] - 0.0228\delta[n-5] + 0.0549\delta[n-6] - 0.0044\delta[n-7] +$$
$$+ 0.0024\delta[n-8]) + \tag{4.1}$$
$$\tag{4.2}$$

$$g(n) = (-1)^{1-n}\ h[1-n] \tag{4.3}$$

The raw PCG signals from different apparatus are first resampled to $1000Hz$. This sampling rate is found to be reasonable considering the frequency spectrum of the PCG signal, which is approximately in the range of $20Hz - 150Hz$, and the required Nyquist rate. The block diagram of the preprocessing stage is represented in Figure 4.5.

Due to the overlapping nature of noise with the spectra of the PCG signal, simple bandpass 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. Comparing the frequency spectra of the PCG signal, $20Hz-150Hz$, against the corresponding equivalent frequency response of the *5th order Daubechies* wavelet, depicted in Figure 4.6(b), it is obvious that most of the energy of the PCG signal is found within the *3rd*, *4th*, and *5th* scales. Therefore, the PCG signals are decomposed up to the *5th* scale and the wavelet coefficients corresponding to the *3rd*, *4th*, and *5th* scales are retained whereas the coefficients related to other scales are set to zero. In this manner high frequency noise components (above $125Hz$) as well as the baseline wander and low frequency noise (below $15Hz$) are removed. The block diagram of the corresponding procedure is depicted in Figure 4.7.

As explained earlier, wavelet analysis is capable of exposing various discontinuities in a signal. After decomposing the signal in different scales, the energy of the main signal is concentrated in a small number of wavelet dimensions so that its corresponding coefficients will be relatively large compared to the noise components which have their energy concentrated over a larger number of wavelet dimensions. This is the underlying concept for many of the wavelet based noise reduction techniques in which the revealed noise is removed by thresholding certain components of the wavelet decomposition. The thresholding operation involves removing coefficients of the DWT which lie below a spec-

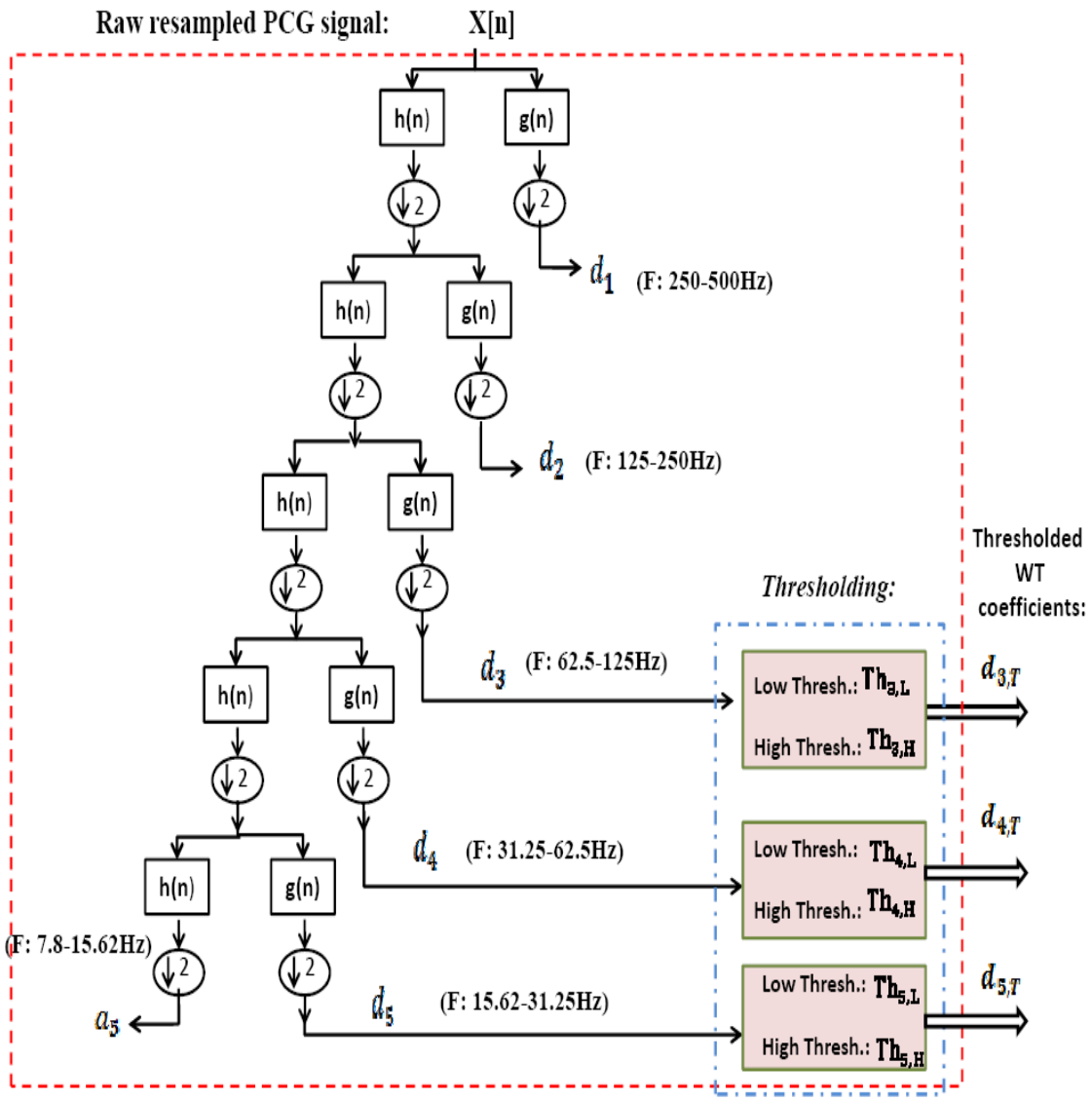

Figure 4.7: Block diagram of the Preprocessing Stage: Decomposition of the PCG signal up to *5th* scale using *5th* order Daubechies wavelets and retaining the thresholded WT coefficients at *3rd*,*4th*, and *5th* scales.

ified value, and then reconstructing the signal with the inverse discrete wavelet transform (IDWT). Therefore, applying the thresholding operation to the wavelet coefficients can remove the unwanted noise, even though the instantaneous frequency spectra of the signal and noise overlap.

Based on this ground, having the wavelet coefficients at *3rd*, *4th*, and *5th* scales, two

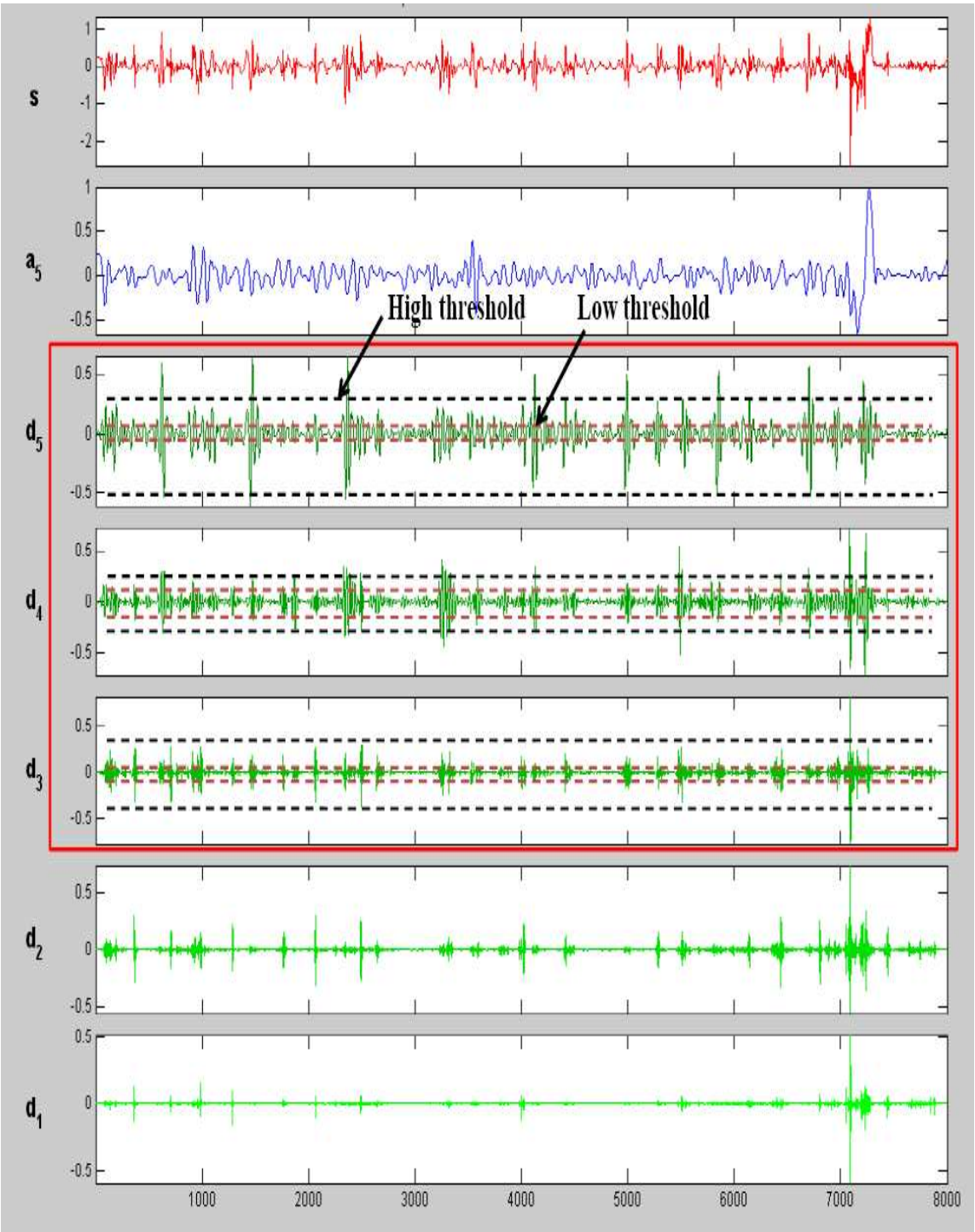

Figure 4.8: WT coefficients of a typical PCG signal (s), from scale $1:5$ ($d_1:d_5$), and the corresponding applied *Low* and *High* thresholds.

energy-based thresholds are defined regarding each scale: the *low* and the *high* threshold as defined in Eq. 4.5. The *low* threshold suppresses the noise components, whereas the *high* threshold detects the spikes. The amplitude of the wavelet coefficients at the detected spike locations is replaced by the mean of the wavelet coefficients at that scale. Finally, the PCG signals are normalized to have zero mean and unit variance. The thresholding levels applied to the wavelet coefficients are shown in Figure 4.8.

$$Th_{i,L} = T_{i,L}\ \mathrm{RMS}(W_{2^i}x[n]) \qquad , i = 3, 4, 5 \tag{4.4}$$

$$Th_{i,H} = T_{i,H}\ \mathrm{RMS}(W_{2^i}x[n]) \tag{4.5}$$

where $T_{i,L}$ and $T_{i,H}$ are the adjusting factors at scale 2^i corresponding to the low and high thresholds respectively.

The output denoised PCG signals of two subjects along with their corresponding raw PCGs are depicted in Figure 4.9.

4.3.2 Feature Extraction using STFT

In the speaker verification area, there are several techniques to extract features representing the speech signals with the objective of discrimination between the given speaker and other speakers. Having an acoustic structure, the same concepts and techniques are applicable to the PCG signal. Like other acoustic signals, in order to avoid the complex pattern of the signal in the time domain, it is preferable to process the PCGs in the frequency domain.

Due to the non stationary nature of the PCG, the discrete *Short-Time Fourier Transform (STFT)*, Eq. 4.6, is used to provide the required spectral information.

$$X[n,k] = \sum_{m=0}^{N-1} x \cdot w[m + (n-1)S] \exp\Big(- j\frac{2\pi}{N}km\Big) \tag{4.6}$$

where n is the frame index, k is the frequency index, N is the frame length, S is the frame shift, and w denotes the window.

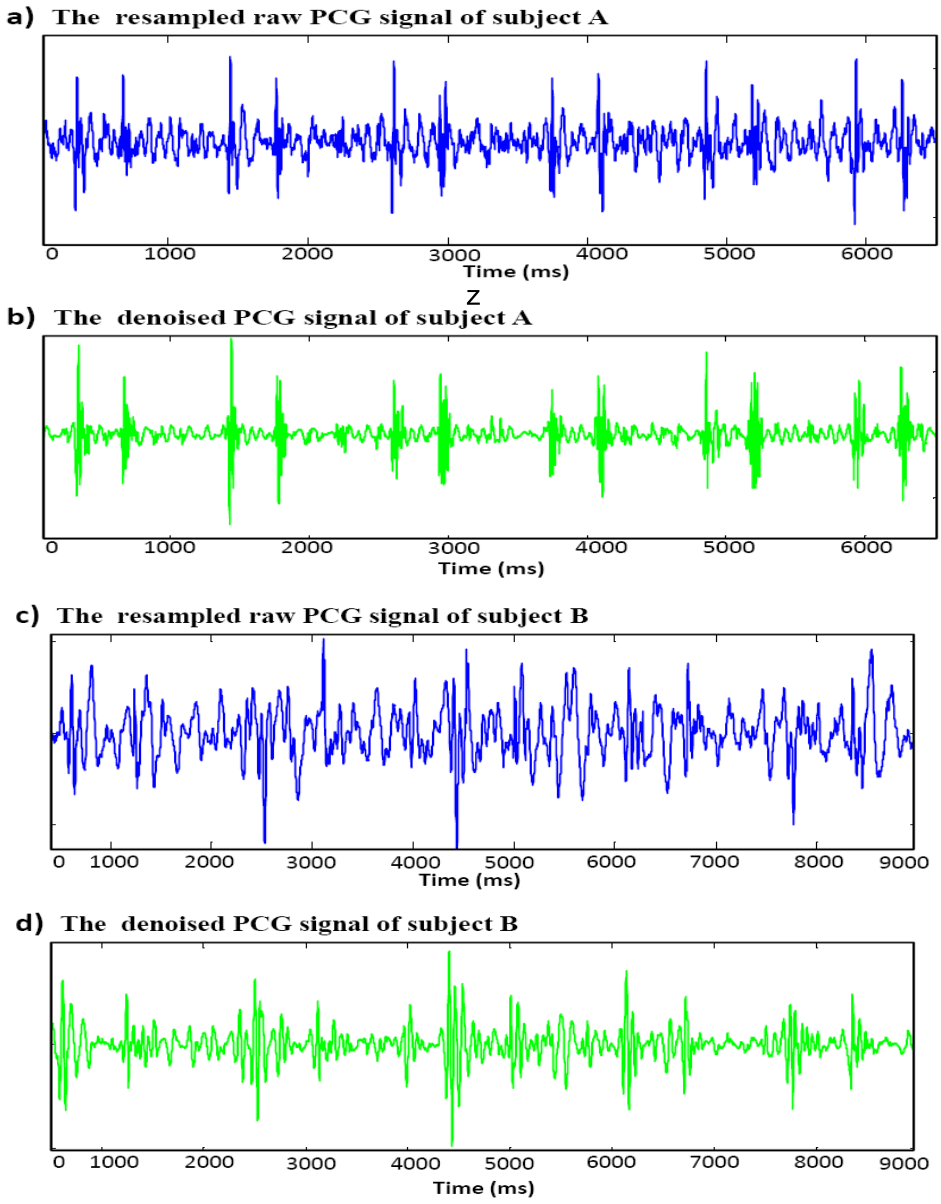

Figure 4.9: Raw PCG signals and the related denoised signals of two subjects (A and B).

In STFT, a moving window is applied to the signal and the Fourier transform is computed within this window. Reducing the length of the window leads to a decrease in the frequency resolution and an increase in the time resolution. In order to estimate the power spectral density of the PCG signal, it is necessary to divide the signal into

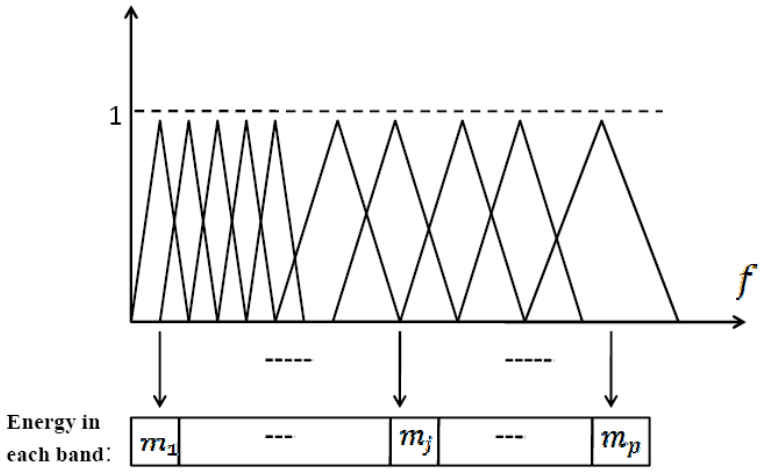

Figure 4.10: Mel frequency filter bank.

frames in which the sound remains sufficiently stationary. For the speech signals, the vocal track is changing after each $20ms - 25ms$; therefore, the optimal window length is chosen with respect to this duration. However, unlike the speech signals, the cardiac signals have a repetitive nature due to the repetitive process of the heart and circulatory system. Therefore, it is possible to define the window length based on the length of one heartbeat. The effect of different window lengths on the recognition rate is investigated in chapter 5. In this work, the windowing process is performed blindly without delineation of the heart beats or detection of the heart rate. Moreover, the correlation among the frames is reduced by applying the non-overlap windowing.

In the next step, the magnitude information of the obtained spectral of each frame is retained for the feature extraction and the spectral phase component, which is typically more sensitive to noise, is discarded. In speech recognition, it is shown that the human ear resolves frequencies non-linearly across the audio spectrum; thus, it has been empirically suggested that designing a front-end to operate in a similar non-linear manner improves the recognition performance. The filter bank analysis provides the desired non-linear frequency resolution by passing the signal spectrum through a filter-bank set which are well matched to those of the desired signal. Moreover, in this way the contribution of

the noise components in the frequency domain are reduced. Based on this fact, the PCG signals are then passed through the Mel-frequency filter-bank, which is shown to be the best in the speaker identification [75]. As illustrated in Figure 4.10, the Mel-frequency filter-bank is decomposed of triangular filters which are equally spaced along the mel-scale defined by the following equation:

$$Mel(f) = 2595log(1 + \frac{f}{700})] \tag{4.7}$$

The PCG's spectral magnitude coefficients are *binned* by correlating them with each triangular filter. The *binning* refers to the multiplication of the FFT magnitude coefficient by the corresponding filter gain and finally accumulating the results. Thus, each bin holds a weighted sum representing the spectral magnitude in that filter bank channel. The number of the filter bank channels is set empirically with respect to the spectrum of the signal. In this work, a Mel filter-bank with 10 channels is utilized. Finally, the logarithm of the obtained frames are used as the extracted features representing one subject.

4.3.3 Classification

Similar to other pattern recognition problems, classification among a gallery set is the last step of the recognition process. The gallery set is composed of the extracted frames associated with each subject. In the proposed approach, classification is performed using the majority voting rule among the frames using the nearest neighborhood classifier with the Euclidean distance measure. In order to reduce the dimensionality of the data space as well as providing more discrimination among different classes the Linear Discriminant Analysis (LDA) is applied over the feature frames.

LDA, which is a dimensionality reduction technique based on the Fisher discriminator, provides a linear projection of the data into a significantly lower dimensional space, in which the different classes are discriminated. The projection directions are chosen

nearly orthogonal to the within-class scatter so that the variations among samples of one subject are projected away while the discriminative information among different classes is retained; a set of weights are extracted to maximize the ratio of the inter-class variance to the intra-class variance [77]. The feature space is obtained in such a way that the following ratio be maximized:

$$W_{opt} = [w_1, w_2, \cdots w_m] = \text{argmax}\frac{W^T S_B W}{W^T S_W W} \tag{4.8}$$

where S_B and S_W are the between class-scatter and the within-class scatter matrices respectively, defined in Eq. 4.9 and Eq. 4.10, and W is the set of eigenvectors of $(S^{-1}_{Wprojected} S_{Bprojected})$.

$$\mathrm{S_B} = \sum_{\mathrm{i=1}}^{\mathrm{c}} \mathrm{N_i}(\mathrm{m_i} - \mathrm{m})(\mathrm{m_i} - \mathrm{m})^{\mathrm{t}} \tag{4.9}$$

$$\mathrm{S_W} = \sum_{\mathrm{i=1}}^{\mathrm{c}} \mathrm{N_i} \sum_{\mathrm{x_k} \in \mathrm{X_i}} \mathrm{N_i}(\mathrm{x_i} - \mathrm{m})(\mathrm{x_i} - \mathrm{m})^{\mathrm{t}} \tag{4.10}$$

where $\mathrm{x_i}$ are samples of the class $\mathrm{c_i}$, $\mathrm{m_i}$ is the mean of class $\mathrm{c_i}$, $\mathrm{N_i}$ the number of samples in that class, and m is the overall mean of the training set.

Since the mean and the covariance information of all subjects in the training set should be available in order to make the feature space, this method is categorized as a supervised learning technique. In order to avoid the singularity problem of the scatter matrices, the dimensionality of the training space is reduced by first applying the PCA. In this work, computational complexity of matrices is reduced by computing the within-class and between-class scatter matrices of the projected set instead of projecting the scatter matrices computed from the original training set.

$$S_{Wprojected} = W^T_{pca} S_W W_{pca} \tag{4.11}$$

$$S_{Wprojected} = W^T_{pca} S_W W_{pca} \tag{4.12}$$

$$W_{opt} = \text{argmax}\frac{W^T S_{Bprojected} W}{W^T S_{Wprojected} W} \tag{4.13}$$

Finlay, the Fisher's criteria is applied to reduce the dimension to $C-1$, i.e.

$$W_{opt}^T = W_{flda}^T \, W_{pca}^T \tag{4.14}$$

After applying the LDA over the feature frames, the subjects are classified using the majority voting rule based on the nearest neighborhood classifier with the Euclidean distance measure.

4.4 Conclusion

Demonstration of the possibility of utilizing the ECG signal for human recognition has drawn the attention of researchers to the PCG signals. These two signals have the same cardiac origin; in addition, important medical information is conveyed through the PCG signal. It has been shown in the literature that the PCG signal has the potential to describe an identity uniquely [56,57]. Like ECG, the PCG signals are difficult to disguise, forge, or falsify.

In this chapter, a wavelet-based approach for human recognition based on this new physiological biometric was developed, evaluation of which in both verification and identification modes over a heart sound database collected at University of Toronto is presented in chapter 5.

Chapter 5

Performance Evaluation

This chapter represents the experimental results of the proposed methodologies for the ECG and also PCG recognition systems. The first part of this chapter is dedicated to the evaluation of the performance of the proposed ECG and PCG frameworks for the identification and verification over some databases. The details of the databases are presented below. The fusion of the ECG and PCG signals in a multimodal biometric recognition system is investigated later in this chapter. Finally, this chapter concludes by a comparison of the proposed approaches with some of the previous works in the literature.

5.1 Databases

In this section, the databases to be used for the system evaluations are briefly introduced. For comparison purposes, two publicly available databases were used for testing, namely the MIT-BIH Normal Sinus Rhythm [49] and the PTB database [51]. As mentioned earlier, one of the challenges of exploiting ECG for recognition is its vulnerability to various stress levels which is a major issue. Since there was no publicly available dataset offering ECG recordings of subjects with different heart rates, it was impossible to evaluate this aspect of different approaches. In order to have a precise evaluation, an

experiment was conducted in the University of Toronto to collect ECGs with different heart rates. Moreover, another experiment was conducted in this university to provide the heart sound database required to evaluate the PCG recognition system. The details of these databases are described below:

5.1.1 ECG Databases

1. *The MIT-BIH Normal Sinus Rhythm Database*: This database was collected at Laboratory of Boston's Beth Israel Hospital. It contains ECG recordings of 18 people not having significant arrhythmias with the sampling rate of $128Hz$. Some of the recordings in this database were severely affected by artifacts such that no heartbeat was recognizable. For our experimental set up, a subset of 13 subjects of this database containing enough valid heartbeats was chosen. In order to form the gallery and probe sets for our experiments, the signals were partitioned into two sections, the first to build the gallery set and the remaining half to test the performance of the system.

2. *The PTB Database*: This dataset was provided by the National Metrology Institute of Germany and contains 594 ECG recordings collected of 294 subjects, both healthy and diagnosed with a variety of clinical conditions. Every record includes the 15-lead ECG, the conventional 12-lead in addition to the 3 Frank leads, with the sampling frequency of $1KHz$. A subset of 14 healthy subjects that had at least 2 recordings was chosen to form the gallery and the probe sets to conduct the experiments.

3. *U of T Rest and Exercise Databases*: Since there had been no publicly available

datasets offering ECG recordings of subjects with different heart rates, an experiment was conducted in University of Toronto in 2008 to collect such database. In this study, one-lead ECG signals were collected over a period of two months using a Vernier device [1] capturing signals at $200Hz$. Two main levels for recording the ECG signals were considered, the rest and the stress conditions. In this experimental setup the heart rate change required for the stress level was achieved by a short time duration exercise.

Subjects between 18-40 years old were eligible to take part in this experiment if they had no history of cardiac problems. A database of 43 subjects (28 males and 15 females) was composed; for 28 subjects who were willing to take part in the exercise the stress database was collected as well.

Every visit included the same procedure and steps for collecting the one-lead ECG signal of the subjects. Each visit began by asking the participant to rest for a while. After achieving a normal heart rate, the one-lead ECG signal was collected for 2 minutes using two sensors that were attached to the participant's wrists and one sensor attached to his/her elbow. Afterwards, the electrodes were removed and some physical activities were performed for 2 minutes to reach a higher heart rate for volunteers who agreed to participate in the exercise part. Then, the ECG recording was repeated in the same manner as the rest condition for 2 minutes.

Finally, the gallery and probe sets were constituted using the earlier recordings under the rest for the gallery set, whereas the later ones under the rest and exercise were used as the probe sets.

[1] http://www.vernier.com/probes/ekg-bta.html

5.1.2 Heart Sound Database

Due to the absence of publicly available database of the healthy heart sound signals required to evaluate the proposed approach, another experiment was conducted in University of Toronto in order to collect the required database. The PCG signals were collected using the *Littmann Electronic Stethoscope* Model 4100WS [2] with the sampling rate of $8KHz$ and 16 bits. The focus of the PCG experiment was to evaluate the proposed approach for healthy subjects under the rest condition. Therefore, following the same protocol of the ECG collection experiment, subjects between 18-40 years old were eligible to take part in this experiment if they had no history of cardiac problems. A database of 21 subjects was composed, each having 6 recording of length $8seconds$ under the rest condition. The location of the stethoscope on the chest was decided based on the strength of the heard heart sound. The recording position was specified for each subject, either Aotric (A) or Pulmonary (P), so that it was not changed for multiple recordings.

Moreover, for investigating the performance of a multimodal biometric system based on the ECG and PCG, the ECG signals were simultaneously recorded with the PCG signals for 3 minutes from these 21 subjects under the rest condition.

5.2 ECG Based Human Recognition System

In order to enroll one subject for the first time in the system database, his/her ECG signal with length n was recorded; the longer the recorded signal, the more precise the corresponding template is. After enrollment, every time the user arrives to the system to be identified, the user's ECG will be recorded and interrogated by the system as the probe sample. Probe set consists of the derived probe templates obtained from the probe samples corresponding to each subject. The minimum required temporal length of an

[2] http://www.allheart.com/littmann4000.html

ECG recording to be used as the probe sample for a precise recognition is discussed later on in this section. The overall procedure for designing a heartbeat template out of an ECG recording, either to be used in the gallery or the probe set, is the same as discussed in chapter 3.

The gallery sets were formed for the PTB and MIT-BIH databases using the first set of the recordings. The PTB gallery set is constituted of 14 templates derived from the total number of 1684 heartbeats recorded from all subjects (120 heartbeats per subject on average), whereas the MIT-BIH gallery set includes 13 templates derived from the total number of 479 heartbeats recorded from all subjects (37 heartbeats per subject on average). For the *U of T* rest and exercise databases, the gallery set was formed using the first recording set under the rest condition. Therefore, 43 basic templates were designed using the total number of 5779 heartbeats recoded from all subjects (134 heartbeats per subject on average).

The remaining heartbeats were used to form the probe sets. Therefore, four probe sets were constituted: the PTB probe set with 1666 heartbeats for 14 subjects, the MIT-BIH probe set with 497 heartbeats for 13 subjects, the *U of T* rest probe set with 5832 heartbeats for 43 subjects, and finally the *U of T* exercise probe set with 4702 heartbeats for 28 subjects.

The system decision was regulated by defining a threshold so that the pairs of biometric samples generating scores higher than or equal to this threshold were inferred as belonging to the same subject; whereas, those having a score below the threshold were not validated. The recognition performance of the system for identification and verification modes are presented as follow.

5.2.1 ECG: Identification Mode

The identification performance of the system as a function of different thresholds for the correlation metric is demonstrated in Table 5.2, Table 5.3, and Table 5.4 for the

Database	No. of HB	No. of subj	avrg. HB per subj
MIT-BIH	497	13	38
PTB	1666	14	119
U of T rest	5832	43	136
U of T exercise	4702	28	167

Table 5.1: Number of extracted heartbeats for MIT-BIH, PTB, *U of T* resting, and *U of T* exercise probe sets.

combination of the PTB and MIT-BIH databases, the collected *U of T* resting database, and the *U of T* exercise database respectively. For all databases, the average number of the heartbeats (HB) required to design robust templates for a precise recognition is reported as well.

The simulation results indicated 100% subject recognition rate utilizing only 2 heartbeat (on average) for the similarity thresholds up to 92% for the combination of the PTB and MIT databases, and 96.81% for the collected resting database. An identification rate of 90.51% was achieved for the exercise database for the similarity thresholds up to 92% which demonstrates the robustness of the proposed approach in handling different heart rates. As anticipated, raising the threshold, which is equivalent to increasing the level of similarity between the templates, to more than 98%, decreases the recognition rate. However, the average number of heartbeats required to extract precise templates increases only one heartbeat at higher thresholds (from 2 to 2.91) which indicates the outstanding ability of the system in compensating for the intra-class variations caused by noise or heart rate variability.

5.2.2 ECG: Verification Mode

The downside of any biometric system in the authentication mode is the mistaken acceptance or rejection, which are measured in false acceptance rate (FAR) and false rejection

Database	PTB+MIT (27 subj.)		
Thresholds	Avg. HB per subj.	No. of HB not recognized	subj. recognition rate
0.75	2.08	0/2163	100%
0.8	2.08	0/2163	100%
0.85	2.08	0/2163	100%
0.9	2.09	0/2163	100%
0.92	2.11	1/2163	99.96%
0.95	2.21	40/2163	98.15%
0.98	2.44	217/2163	89.97%

Table 5.2: The identification performance of the proposed ECG based system as a function of different similarity thresholds for the combination of PTB and MIT-BIH databases.

Database	U of T Resting (43 subj.)		
Thresholds	Avg. HB per subj.	No. of HB not recognized	subj. recognition rate
0.80	2.40	186/5832	96.81%
0.85	2.40	186/5832	96.81%
0.9	2.40	186/5832	96.81%
0.92	2.40	186/5832	96.81%
0.95	2.43	191/5832	96.72%
0.98	2.91	626/5832	89.26%

Table 5.3: The identification performance as a function of different similarity thresholds for the *U of T* rest database.

rate (FRR) respectively. False acceptance takes place when the system mistakenly verifies the claimed identity of an imposter; whereas, false rejection refers to cases where

Database	U of T Exercise (28 subj.)		
Thresholds	Avg. HB per subj.	No. of HB not recognized	subj. recognition rate
0.80	2.83	416/4702	91.15%
0.85	2.83	418/4702	91.11%
0.9	2.86	425/4702	90.96%
0.92	2.91	446/4702	90.51%
0.95	3.23	940/4702	80.01%
0.98	3.46	2623/4702	44.21%

Table 5.4: Identification performance as a function of different similarity thresholds for the *U of T* exercise dataset.

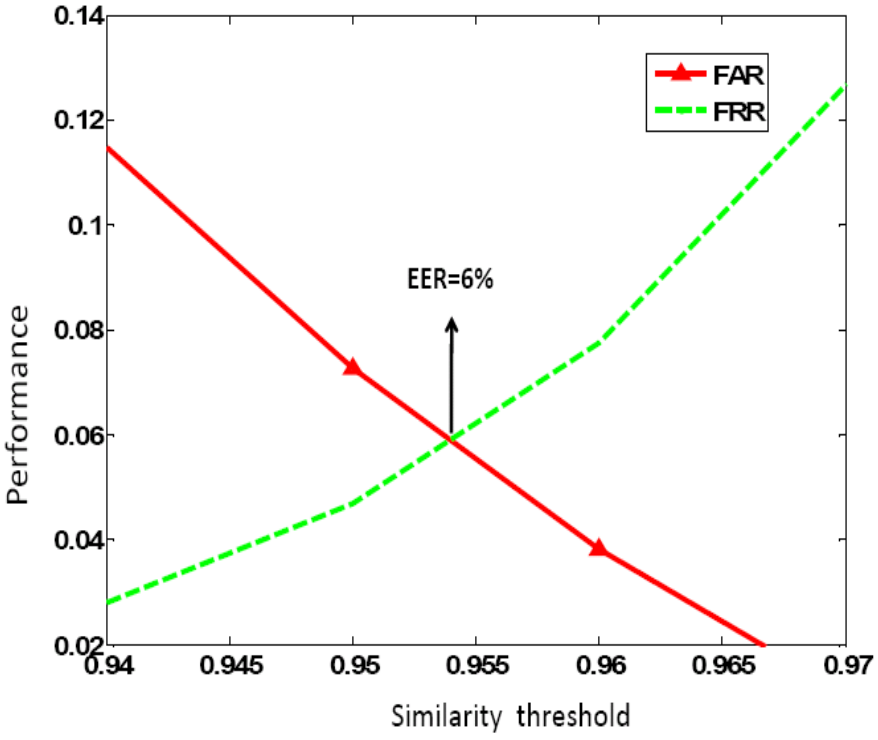

Figure 5.1: Plots of false acceptance rate (FAR) and false rejection rate (FRR) as a function of different similarity thresholds for PTB+MIT database.

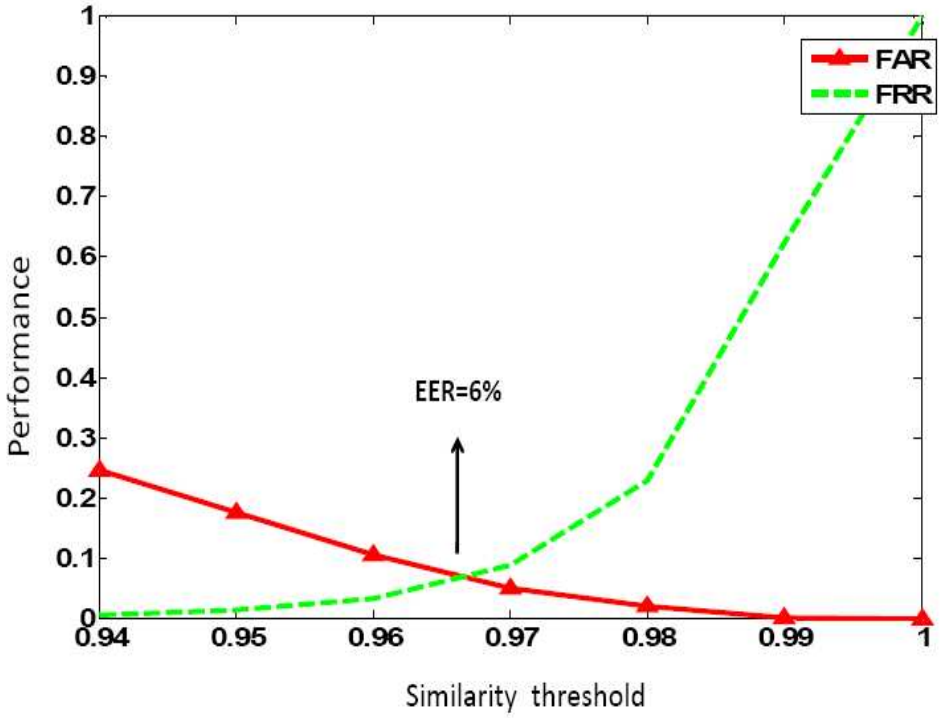

Figure 5.2: Plots of false acceptance rate (FAR) and false rejection rate (FRR) as a function of different similarity thresholds for resting database.

system denies a genuine identity claim.

Since both FAR and FRR are functions of the system threshold, there is always a tradeoff between these two types of errors. Increasing the FAR as a result of lower thresholds makes the system more tolerant to input variations and noise. However, raising the threshold in order to make the system more secure, increases the FRR. The system performance as a plot of FAR and FRR for various threshold values are depicted in Figures 5.1, 5.2, and 5.3 for the combination of the PTB and MIT-BIH databases, the collected resting database, and the collected exercise database respectively. The achieved equal error rate (EER) for the three databases were 6%,6%, and 18% respectively.

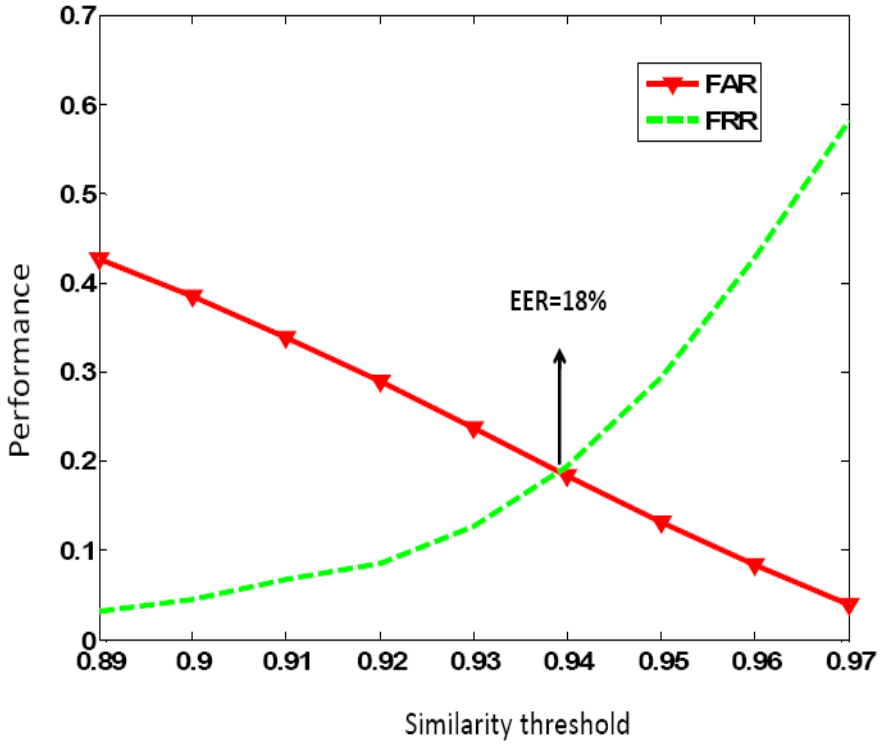

Figure 5.3: Plots of false acceptance rate (FAR) and false rejection rate (FRR) as a function of different similarity thresholds for exercise database.

5.3 PCG Based Recognition System

Being a new concept, the PCG recognition is not as developed as the ECG recognition. There were three major considerations for the PCG based human recognition in this work: first, improvement of the methodologies proposed so far in terms of accuracy and the computational efforts; second, extending the application of the PCG signal to the verification mode; and third, collecting the required heart sound database for evaluating the proposed approach. Various stress conditions and/or pathologies are not considered in the experimental set up.

- **PCG: Identification Mode**

  Two experiments were performed in order to evaluate the performance of the proposed PCG methodology in the identification mode. In the first experiment, the

Distance Threshold	No. of samples per subject in the gallery set				
	2	4	8	10	15
4%	26.9%	50%	72.3%	74.5%	80.6%
5%	46%	77%	88.5%	95.1%	97%
6%	62.5%	88.7%	98%	98.4%	99.6%
8%	78%	94.5%	98.4%	100%	100%
10%	87.9%	97.6%	99%	100%	100%
15%	95.3%	99.4%	100%	100%	100%

Table 5.5: The average identification rate as a function of different thresholds and different number of samples per subject in the gallery set, for 50 trials.

identification performance of the system was evaluated as a function of the number of samples per gallery subject. Therefore, the gallery and the probe sets were formed so that N frames out of total 48 frames obtained for each subject were chosen randomly in order to constitute the gallery set, whereas the remaining frames, $(48 - N)$, were used as the probe set. This procedure was iterated 50 times. The average identification performance as a function of different thresholds for the normalized minimum Euclidian distance is tabulated in Table 5.5. As it is apparent from this table, the proposed method achieves 100% recognition rate utilizing 10 frames (heartbeats) per subject (on average) in the gallery set for a maximum distance threshold of 8%. Increasing the gallery size leads to a more accurate training of the LDA; thus, a better identification rate is achieved for all thresholds.

In the second experiment, the relationship between the identification performance and the window length of the STFT was investigated. As mentioned earlier, the PCG signal is a non-stationary but highly repetitive signal. The window length of the STFT should be chosen such that the signal remains sufficiently stationary within each window. The window length can be chosen with respect to the heart-

Distance Threshold	Window length		
	250 ms	500 ms	1000ms
6%	100%	100%	99%
8%	100%	100%	100%
10%	100%	100%	100%

Table 5.6: The average identification performance for different window lengths as a function of different thresholds, with 10 samples per subject in the gallery set, 50 trials.

beat duration; thus, for the experimental set up three window lengths were considered, one of the length of a normal heartbeat under the rest condition,1000ms, and the others of a length in which different components of a typical heartbeat remain stationary, 500ms and 250ms.

The windowing process was performed blindly without delineation of the heartbeats or detection of the heart rate. For the window length of 250ms a total number of 192 frames were constructed from the recorded PGS signals of each subject. The obtained number of the frames for the window length of 500ms and 1000ms were 96, and 48 respectively. In all cases, a gallery set was constituted using 10 frames randomly chosen out of the total frames associated with each subject. The average identification rate for 50 iterations as a function of distance threshold is tabulated in Table 5.6. From this table, it is apparent that decreasing the window length from 1000ms to 250ms does not have a profound influence on the identification rate. Therefore, it is suggested to choose the window length equal to 250ms in order to decrease the required length of the PCG signal and consequently to decrease the enrollment time.

- **PCG: Verification Mode**

The system performance in the verification mode is illustrated as a plot of FAR

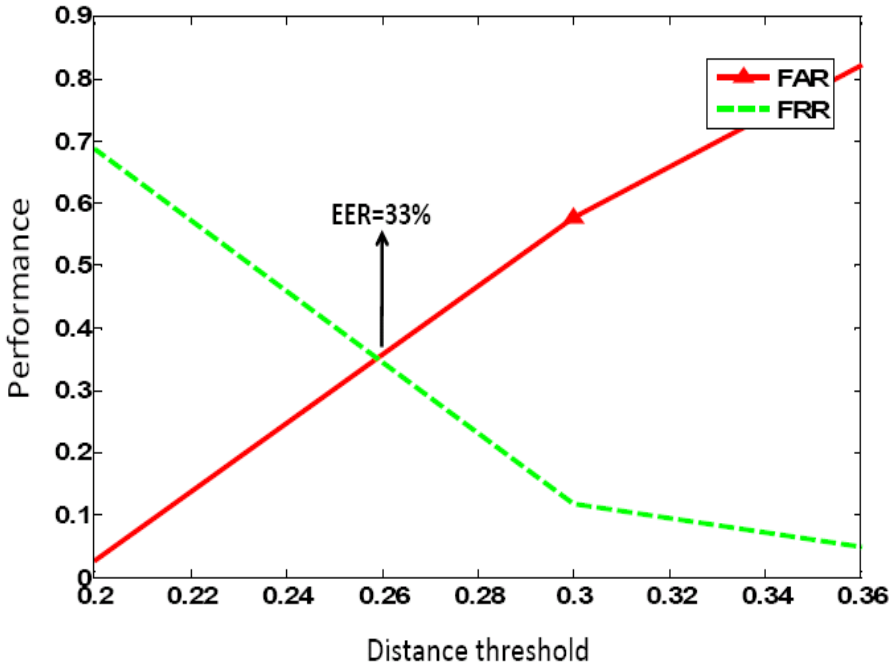

Figure 5.4: Plots of false acceptance rate (FAR) and false rejection rate (FRR) as a function of different thresholds for the *U of T* PCG dataset.

and FRR for various threshold values in Figures 5.4. As it is can be seen, similar to the ECG signal, increasing the FAR as a result of higher distance thresholds makes the system more tolerant to input variations and noise. However, decreasing the threshold in order to make the system more secure, decreases the FAR. The Achieved EER for this scheme was equal to 33%.

5.4 Fusion of the ECG and PCG Signals

Each biometric technique has its own limitations and advantages; therefore, in order to improve the recognition performance, multi-modal biometric fusion is introduced. Multi-modal biometric system is a relatively new application of the information fusion while individual biometrics has been used for a fairly long time. Information fusion is usually

considered as the combination of various sources of information, either to generate one representational format, or to reach a decision. Multi-modal biometric fusion has drawn a lot of attention recently with the motivation of utilizing complementary information to reduce the measurement errors, as well as utilizing multiple classifiers to increase the correct classification rate [78]. According to the relationship of the sources to be fused, information fusion can be performed over dependent or independent sources. Fusion under both categories can be carried out at three different levels: the raw data-level fusion, the feature-level fusion, and the decision-level fusion, the corresponding block diagrams of which are demonstrated in Figure 5.5.

In the raw data-level fusion (Figure 5.5(a)), the raw data from multiple sensors are combined directly. For the dependent sources, raw data level fusion is widely applied particularly in target-tracking, remote sensing image fusion, navigation, etc. In the biometric context, the sensor refers to the signal collectors from either different biometrics (in multimodal biometrics systems) or multiple measurements from the same biometric (in unimodal biometric configurations). However, due to the complexity of fusing raw information from independent biometrics traits, it is preferable to use the raw data level fusion in the unimodal contexts. For instance, multiple images are fused together to produce one single image by averaging the intensities of multiple images. In this manner, the random measurement noise of the fingerprint images can be reduced. After fusion of the raw data, the feature is extracted from the fused data by feature extraction module, and finally classified.

The feature-level fusion, depicted in Figure 5.5(b), is the second type of data fusion which is basically divided into two approaches: the feature concatenation and the classifier fusion. In the feature concatenation approach, the feature vectors from the multiple biometrics are concatenated to form a new feature vector with a larger dimension. This concatenated vector is then used to feed the classifier. However, in the classifier fusion approach, the individual feature vectors from each sensor are first input into their own

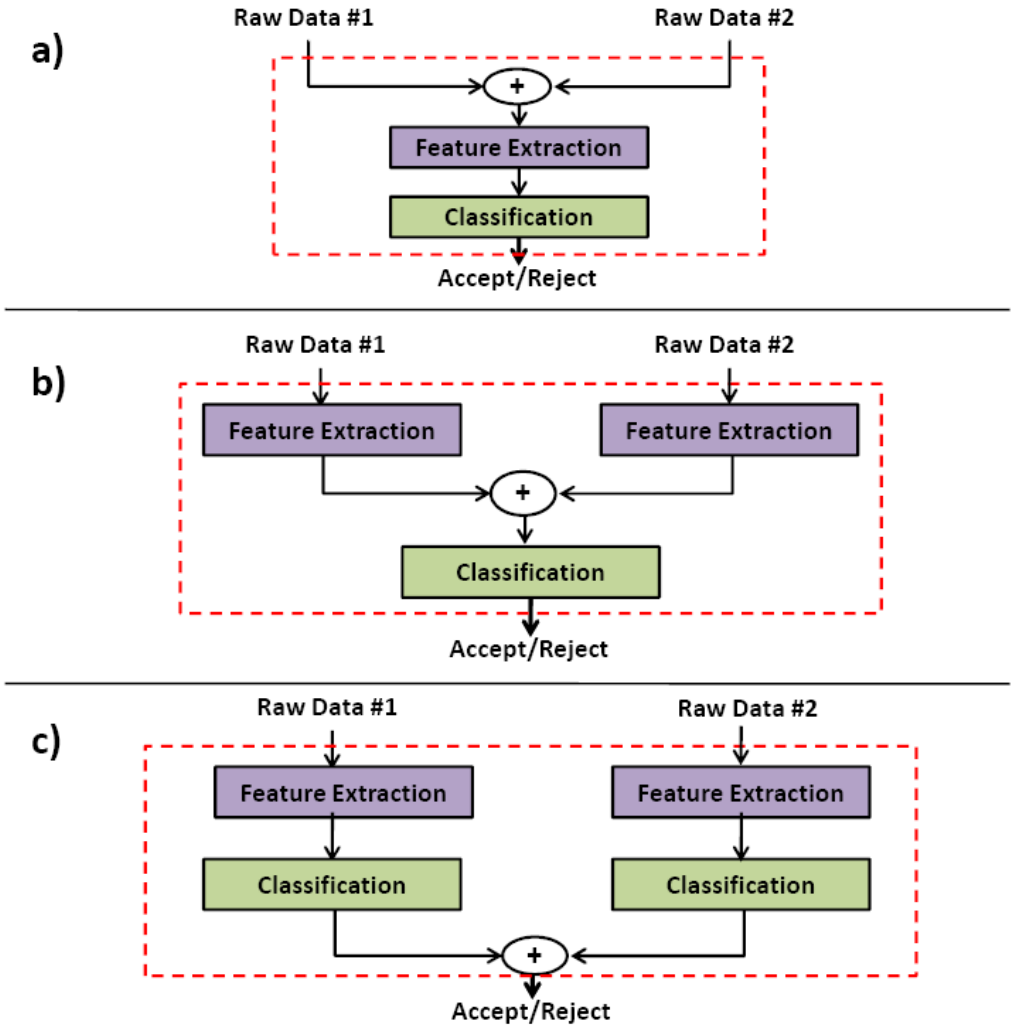

Figure 5.5: Block diagram of the biometric fusion levels, (a) raw data-level fusion, (b) feature-level fusion, (c) decision-level fusion.

feature matching stages to generate a matching score. This score indicates the proximity of the feature vector with the template vector stored in the system database. The classification is finalized based on the combination of these scores using combination method such as the majority voting scheme. Feature level fusion is applicable to both dependent and independent sources. However, since the feature vector resulted from concatenation of different sources usually have large dimensions, the application of the dimensionality reduction techniques is required before the classification.

In the case of decision level fusion (Figure 5.5(c)), individual classifiers output a hard decision whether to accept or reject the claimed identity. Then, the final decision about the vicinity of the claimed identity is made by asserting these individual decisions using a combination scheme including rule-based techniques (majority voting, AND/OR, PROD/SUM), machine learning and NN classifier, linear classifier, statistical decision theory (ML, Bayesian, DM), fuzzy k-means, logistic regression etc. Therefore, this scheme is more appropriate for the combination of the independent sources.

Sharing the same origin, the PCG and ECG signals are counted as dependent biometric sources. In this section some analysis is developed to determine whether the fusion of these two dependant traits could improve the overall recognition performance of the system.

The fusion is not feasible at the raw data level since the two signals require different noise reduction techniques. On the other hand, in the proposed ECG methodology, one template was derived using all heartbeats available for a particular subject. Therefore, each subject corresponds to only one template. However, in the proposed approach to PCG recognition, designing of such template was impossible due to the characteristics of the heart sound signal. Instead, a set of frames were developed and stored as the feature sets representing an individual. On account of this fact, the fusion of these two signals is not possible at the feature-level either. Therefore, the decision level fusion seems to be the only appropriate fusion level which is explored in this section.

The performance of the multimodal biometric system based on the ECG and PCG signals is investigated in the identification mode. The evaluation is performed using the database containing simultaneously collected ECG and PCG signals of 21 subjects at U *of* T. The block diagram of this procedure is depicted in Figure 5.6

When the subject to be identified arrives to the system his/her ECG and PCG signals are recorded for T seconds. As it was investigated in the previous sections, 3 heartbeats (on average) are enough in order to derive a precise template required for the ECG

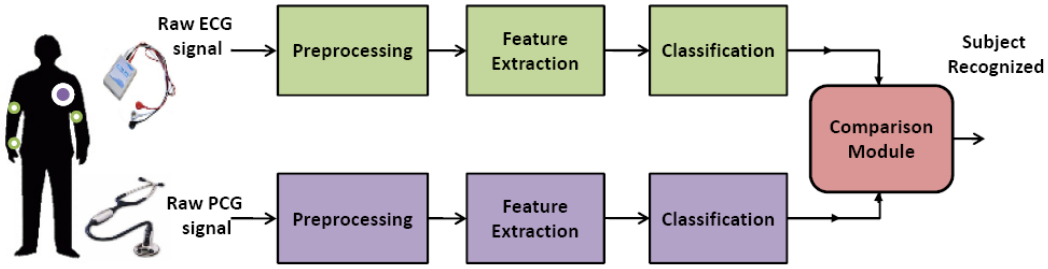

Figure 5.6: Block diagram of the fusion of the ECG and PCG signals.

recognition. Moreover, it was shown that a 100% identification rate is achieved for the PCG recognition using 8 frames of length $250ms$ for a distance threshold less than 10%. Therefore, the length T is set to 4 seconds.

At first, both the ECG and PCG signals are preprocessed and the required features are extracted using the procedures describe before. Then, classification at each module is performed according to the related criteria of that module. If the identified subject of both modules match, the system declares the subject. Otherwise, the gallery sets are pruned to the corresponding data of the most probable subjects. Comparison is again performed in each module using the pruned dataset and the subjects are classified based on the decision of the module with higher certainty. In this way, these two credentials, PCG and ECG, compensate each other weaknesses and lead the system to a more precise classification.

The experimental results over the simultaneous ECG and PCG datasets achieved a recognition rate of 98.4% for the fused scheme, while the identification rate obtained by the individual modules were 91% ,for a similarity threshold of 95%, and 85% ,with 70% certainty, respectively. This procedure is more illustrated in Figure 5.7.

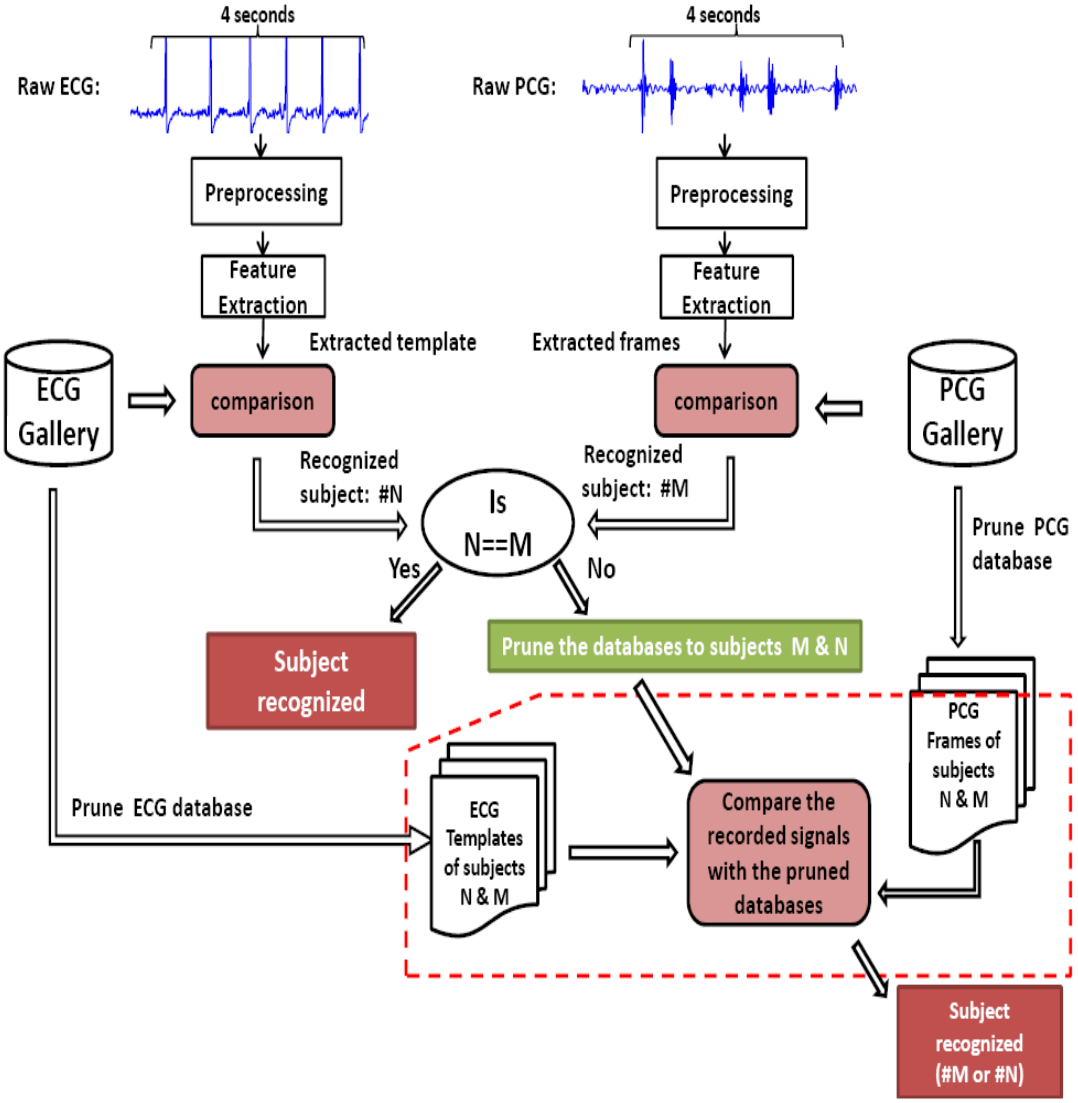

Figure 5.7: Block diagram of the fusion of the ECG and PCG signals.

5.5 Comparison of the Proposed Systems with Other Schemes

In this section a comparison of the proposed ECG and PCG methodologies against some of the previous approaches is presented.

5.5.1 Comparison of the ECG Based Recognition Systems

In this part, two of the most dominant works in the ECG recognition area, one from the fiducial based approaches and one from the non fiducial based approaches, are reproduced and evaluated over the three available databases: the combination of the PTB and MIT-BIH datasets, the *U of T* resting database, and the *U of T* exercise database. The data offered by all datasets are first down sampled to $250Hz$, so that the system be able to handle it equivalently.

1. ***Fiducial Based Approaches: Israel et al. [8]***

   As mentioned in chapter 2, one of the earliest works for applying the ECG in the identification context, was reported by Israel *et al.* [8], the work which is reproduced in this part. Following the same procedure of this work, at first the collected ECG signals were filtered using a *4th* order Butterworth filter with the cutoff frequencies of $0.5Hz - 40Hz$. The peaks of signal's components (R, T, and P waves) were detected in the time domain in the next step by finding the local maxima in the regions surrounding each wave. The boundary locations of these wave components were estimated by computation of minimum radius of curvature around each peak, and finally the heartbeats were synchronized using the R peak positions.

   Based on the detected points, 15 temporal distances were defined, [8], to form the feature space. In order to account for the heart rate variability, the values related to the P and the T wave's distances were normalized by dividing by the $\acute{L}\acute{T}$ distance, whereas the raw RQ and RS distances were used as the features. In the next step, the LDA was applied to reduce the dimensionality of the data space and the classification was performed using the majority voting with the Euclidian distance measure. The identification performance achieved for this scheme is tabulated in Table 5.7.

   As it is apparent from the simulation results, this approach achieves 96% recogni-

Database	PTB+MIT	U of T resting	U of T exercise
Israel *et al.* [8]	96%	65%	43%
Plataniotis *et al.* [13]	94%	76%	64%
Proposed method	98%	97%	91%

Table 5.7: Comparison of the identification performance of three methodologies over 3 different databases.

tion rate for the combination of the PTB and MIT-BIH datasets. This rate reduces to 65% for the *U of T* resting database. The reason for this dramatic reduction is the disability of the system in correct localization of the fiducial points in the presence of the noise components. The recognition rate degrades even more to 43% for the *U of T* exercise database. This proves that the normalization method offered in this work is not capable of compensating for the heart rate variabilities; thus, has led to the poor adaptability of the system for recognizing ECGs under stress conditions. According to Table 5.7, the proposed methodology in this thesis overrides this method for the resting dataset by 32% and for the exercise dataset by 48%.

2. ***Non Fiducial Based Approaches: Plataniotis et al. [13]***

This approach was first proposed by Plataniotis *et al.* [13] to eliminate the necessity of fiducial points' localization of the ECG signal. The proposed methodology which is referred to as AC/DCT is based on a combination of autocorrelation and discrete cosine transform. The input ECG signal is first preprocessed using the *4th* order bandpass Butterworth filter with the range of $1Hz-40Hz$. The preprocessed ECG trace is then segmented into non overlapping windows which are longer than the average heartbeat length. Therefore, the ECG windows are applied instead of the specific fiducial points extracted from each heartbeat.

Then, the autocorrelation of the windowed ECGs were computed, normalized, and

further compacted using the discrete cosine transform (DCT). The small DCT coefficients were discarded and the retained ones were used to feed the k-nearest neighborhood classifier (KNN).

The best identification rate achieved for this system, with respect to the number of DCT coefficients, is tabulated in Table 5.7. The window length was set to $5ms$ and the identification was performed using different number of DCT coefficients. As it is apparent from the simulation results, this approach achieves 92.5% recognition rate for 300 DCT coefficients over the combination of the PTB and MIT-BIH database. This rate reduced to 86.1% using 100 DCT coefficients for the *U of T* resting database. Using the same preprocessing stage as [8], the disability of the system in compensating the noise effects has led to this reduction of the recognition rate. Moreover, the recognition rate of the system drops dramatically to 60.1% using 100 DCT coefficients over the *U of T* exercise database. This indicates the poor adaptability of the system for recognizing ECGs under different stress conditions suggesting that the combination of the autocorrelation and DCT, does not contain the personalized information embedded in the ECG signal regardless of the heart rate. Therefore, some normalization and adaption procedures are required to account for the changes in the heart rate. According to table 5.7, the proposed approach in this thesis overrides this method for the *U of T* resting dataset by 21% and for the *U of T* exercise dataset base by 27%. A comparison among different ECG recognition approaches is provided in Figure 5.8.

5.5.2 Comparison of the PCG Based Systems

In this part the proposed PCG approach in this thesis is compared against the method proposed by Phua *et al.* [56], which is reproduced in this section following the same steps.

	Biel et . al	Israel et . al	Wang et . al	Plataniotis et . al	Agrafiotis et . al	Proposed method
Preprocessing method	-	Butterworth filter of order 4 (2-40 Hz)	Butterworth filtering of order 4 (1-40 Hz)	Butterworth filtering of order 4 (0.5-40 Hz)	Butterworth filtering of order 4 (1-40 Hz)	Discreet Wavelet Transform
Ability to remove noise	-	poor	poor	poor	poor	good
Extracted features	Temporal + amplitude + slopes	15 Temporal	21 Temporal and amplitude + appearance	autocorrelation of windows	autocorrelation of windows	Heartbeat templates
Extraction method	Machine based	Automatic	Automatic	Automatic	Automatic	Automatic
Heart rate adaptability procedure	-	Normalizing features	-	-	-	Design of normalized templates
Dimensionality reduction method	LDA and	PCA	PCA or LDA	DCT	LDA	-
Classification method	Majority voting	SIMCA model	Nearest centre, Nearest neighbor, LDA	Euclidean Distance+ Gaussian likelihood	Nearest neighbor on Euclidean distance	Fusion of the correlation of the templates and derivatives of the templates
Gallery constitution	Vectors of features	Vectors of features	Windows of size 0.8s	Blind Windows of size 10s	Blind Windows of size 5s	1 template per subject

Figure 5.8: Comparison of different ECG recognition approaches.

In the preprocessing stage, Short-time discrete Fourier transform (STDFT) was applied over the PCG signals with an optimal window length of $500ms$. Then, the frequency spectrum of the signal was redistributed using the Mel-frequency filter bank so that its spectral characteristic be best match to those of the desire signal. In order to provide the full resolution, then the frequency bins outside the range of $20-150Hz$ were filtered out. The output spectral magnitude was transferred to the logarithm domain, followed by the discrete cosine transform (DCT). The dimension reduction was achieved by selecting the

Overlap	0	25%	50%	75%	90%
Identification rate	65%	75%	85%	88%	86%

Table 5.8: The identification performance as a function of different window overlaps and a window length of $256ms$ and 100 iterations.

first 24 DCT coefficients of each frame.

The quality of the PCG recordings was then enhanced by application of a spike removal technique based on the energy threshold as defined in [56]. In order to compensate for the effect of the stethoscope's position, the cepstral mean subtraction was performed and the obtained cepstral coefficients were used as the biometric features to be classified. Finally, the classification was concluded using the Gaussian Mixture Modeling with four mixture densities (GMM-4) as the classifier. Using a frame length of $256ms$ and 100 iterations, the achieved identification rate as a function of different window overlaps is tabulated in 5.8. The simulation results in this table indicate that increasing the overlaps between the frames enhance the identification rate of the system, since the small data set problem is avoided by increasing the number of frames.

However, as mentioned before, due to utilization of the GMM technique, the proposed scheme is slow and time consuming since large number of iterations are required for training the classifier. The best identification rate achieved by this scheme compared to the proposed PCG methodology in this thesis is degraded 12% due to the incapability of the designed preprocessing steps for reducing the noise effects.

Chapter 6

Conclusion and Future Improvements

6.1 Research Summary

In this thesis the biometric applicability of two cardiac traits, the electrocardiogram (ECG) and the phonocardiogram (PCG), was studied. There is strong evidence in the literature that the heart's electrical activity (ECG) embeds highly distinctive characteristics, suitable for applications such as human recognition. On the other hand, having the same origin with the ECG l, the PCG signals have been investigated recently for recognition applications and it has been shown that PCGs also convey distinctive information representing an individual.

However, as in most pattern recognition problems, the presence of noise and artifacts is a major issue that affects the recognition performance. Moreover, the expression of the cardiac signals is subject to alternation with different heart rates which further complicates the problem. Thus, a central consideration of this thesis was the design and evaluation of robust preprocessing stages based on the discrete wavelet transform that can compensate for these effects.

In the first part of this thesis we developed and evaluated a wavelet-based ECG recognition system, applicable to both identification and verification, which performs subject recognition in three stages: 1) ECG preprocessing and delineation using the dyadic wavelet transform, 2) Design of personalized templates representing each subject heartbeat's morphology regardless of the heartbeat rate, 3) Classification, based on the amount of similarity among the obtained templates. The main novelties of the proposed ECG approach lie in the design of a robust preprocessing stage for noise reduction and the design of a personalized template which represents individual heartbeat's pattern regardless of the heart rate. In this way, the computational effort of the classification stage as well as the storage requirements of the system are reduced substantially, by storing only one template per subject in the gallery and probe sets.

Moreover, an experiment was conducted in University of Toronto in order to collect one-lead ECG recordings of individuals with different heart rates. Two main levels for recording the ECG signals were considered, the rest and the stress conditions; the heart rate change required for the stress level was achieved by a short time duration exercise. A database of ECG recordings of 43 subjects under the rest condition and 28 subjects under the stress condition was composed.

The proposed algorithm has been validated using three databases, the combination of two standard databases, namely the PTB and the MIT-BIH, and the two collected datasets in University of Toronto, the *U of T* resting and exercise databases.

For the combination of the PTB and MIT-BIH databases with 27 subjects, the proposed algorithm achieved a robust recognition rate of 98.2% for a similarity threshold of 95% utilizing only 2 heartbeats (on average) per subject for the template design. The identification performance of this methodology for the *U of T* resting database with 43 subjects, was reported as 96.8% for the same similarity threshold (95%) using 2.5 heartbeats (on average) for designing the templates.

Moreover, the performance of the system for the ECG signals under the stress conditions of the exercise database collected at the *U of T*, indicated an identification rate of 91% for a similarity threshold of 92%, utilizing 3 heartbeats (on average) per subject for template design. In the verification mode, the reported EER of the system was 6%,6%, and 18% for the combination of the PTB and MIT-BIH databases, the *U of T* resting database, and the *U of T* exercise database respectively.

To the best of our knowledge, the clue for this outstanding performance is the robust noise attenuation achieved by the application of the multiscale approach. Reduction of the inter-variability among the heartbeats, as a result of template design, is another reason for this improvement.

A comparison between the proposed approach and two of the previous methods, one fiducial based approach [8], and one nonfiducial based approach [13], demonstrated the robustness of the proposed method in a real life scenario. The evaluation of the scheme proposed in [13], achieved an identification rate of 86% over the collected rest data base and 61% for the collected exercise database. This reduction of the identification performance is justified as the poor adaptability of the system for recognizing ECGs under different stress conditions.

On the other hand, the performance of the fiducial based approach proposed in [8] degraded dramatically to 65% and 43% for the *U of T* resting and exercise datasets respectively. The reason for this reduction is the incapability of the system to perform correct localization of fiducial points in the presence of noise. In addition, the simple normalization step described in [8] is not capable of accounting for heart rate variability; therefore, the system cannot recognize noisy ECGs with different heart rates of the same subject.

In the second part of this thesis, the possibility of utilizing the PCG signals as a physiological biometric was investigated and a wavelet-based approach for human recognition based on this signal was developed and evaluated.

The PCG signal has same origin with the ECG; in addition, important medical information is conveyed through this signal. Therefore, it was conjectured that the PCG signal may contain information about an individual's physiology. It has been shown in the literature that the PCG signal has the potential to describe an identity uniquely [56,57]. Like ECG, the PCG signals are difficult to disguise, forge, or falsify. Moreover, this signal has the advent of being relatively easy to acquire, since it can be collected by placing a stethoscope on the chest.

There were three major consideration for the proposed PCG recognition system: first, improvement of the methodologies proposed so far in terms of accuracy and the computational efforts; second, extending the application of the PCG signal to the verification mode; and third, collecting the required heart sound database for evaluating the proposed approach.

Due to the absence of publicly available database of the healthy heart sound signals required to evaluate the proposed approach, another experiment was conducted in University of Toronto in order to collect the required database. A dataset consisting of the PCG recordings associated with 21 subjects under the rest condition was collects; various stress conditions and/or pathologies were not considered in the experimental set up.

In this thesis, a wavelet based approach based on both the time and frequency properties of PCG was developed for automatic analysis of this signal for application in human recognition. Similar to the ECG analysis, the proposed approach includes three major stages: 1) PCG preprocessing and denoising using the dyadic wavelet transform, 2) Feature extraction based on the short time Fourier transform (STFT), 3) Classification among subjects using the linear discriminant analysis (LDA) and the majority voting rule.

The proposed algorithm was evaluated on the PCG dataset collected at *U of T* in both identification and verification modes. The achieved identification performance was 100% for a normalized distance threshold of 8% and an ERR of 33% was achieved for

the verification. The comparison of the proposed methodology with the method reported in [56] indicated a 12% identification rate enhancement.

Finally, in the last part of this thesis, a fusion of the two signals in a multimodal biometric system was investigated. With regard to the specification of the PCG and ECG modules, fusion was performed at the decision level.

The proposed multimodal methodology was evaluated using a database containing simultaneously recorded ECG and PCG signals of 21 subjects at *U of T*. Experimental results suggested that these two credentials can compensate each other weaknesses and offer a more precise classification. The application of the ECG module alone obtained an identification rate of 93% for 95% similarity threshold. Also, the identification rate achieved by the PCG module with a certainty of 70% was 88.7%. However, the fusion scheme of these two modules achieved an identification rate of 96.4%.

6.2 Future Directions

The described mythologies for the proposed ECG and PCG recognition systems can be extended in three major directions. The first direction includes the design and evaluation of techniques that could perform the recognition for the cardiac signals with arrhythmias. Another extension can be performed in the area of biometric encryption, to address the issues mostly related to the application of the electrocardiogram and the heart sound for biometric applications. Moreover, the fusion of these signals with the other biometric traits in a multimodal biometric system can be investigated.

6.2.1 ECG and PCG in Arrhythmia

Arrhythmias are abnormalities of cardiac functions which are expressed in the ECG signals as distortions of a general healthy appearance and in the PCG signals as addition of other components (such as $S3$ and $S4$). Depending on the severity of the cardiac problem, the cardiac signals can be affected dramatically, rendering recognition impossible. For a commercial ECG and/ or PCG based security system, however, it is a prerequisite that the biometric recognition be possible for all subjects, whether having cardiac disorders or not. In the current work, the ECG and the PCG recognition methodologies were designated for the normal cardiac signals and tested over three databases, which did not exhibit significant cardiac disorders. The future work will extend to the investigation of different arrhythmia cases and design of an approach that could perform the recognition even for the isolated subjects.

6.2.2 Biometric Encryption

There is a range of social, legal and ethical concerns about biometric systems mainly due to the fears about the potential for data abuses and the centralization of information The issue of potential data misuse is more paramount for the medical biometrics since, by nature, they convey a wealth of information about the health state of an individual.

Being a subset of medical biometrics, there are some concerns about application of the electrocardiogram and the heart sound signals in biometric contexts. However, establishing international policies and standards to appropriately and globally govern these technologies based on the ground rules of respect for privacy and enhanced security could diminish these concerns to great level. Moreover, there has been a lot of research about different biometric encryption, data acquisition, and data filing methods that would guarantee confidentiality of the information and avert the privacy violations.

Future works can concentrate in the design of self-encryption recognition systems based on the heartbeat templates and PCG frames. In this manner, the encrypted data can be stored rather than the original data and the recognition can be performed over the encrypted data using appropriate quantization and error correction methods.

6.2.3 Multi-Modal Biometric System

Each biometric technique has its own limitations and advantages. Therefore, fusion of different biometric sources in a multi-modal biometric system can provide a more accurate recognition due to the utilization of complementary information provided by different traits. In a future work, fusion of the cardiac signals with other biometrics such as fingerprint, iris, and face images can be investigated.

Appendix A

Modulus Maxima Lines

The wavelet transform is a very useful tool in assessing the local scaling behavior of functions and measures. Essential information bout different irregularities of a signal is carried in the corresponding maxima lines across the scales of the wavelet transform, studying of which allows an accurate assessment of the properties of the signal at specific temporal points.

Modulus maximum is defined as any point (s_0, x_0) such that $|Wf(s_0, x)|$ is locally maximum at $x = x_0$. This local maximum should be a strict local maximum in either the right or the left neighborhood of x_0 to avoid having any local maxima when $|Wf(s_0, x)|$ is constant. We call modulus maxima line any connected curve in the scale-space (s, x) along which all points are modulus maxima. For the analysis of singularities it is therefore sufficient merely to consider these maxima lines, which allow detection of irregularities. Extended information about maxima lines can be found in [20].

The required algorithm for computation of the modulus maxima lines corresponding to the R waves, as well as the isolated and redundant modulus lines is provided in this section.

- *The Modulus Lines Corresponding the* R *Waves:*

  In order to detect the R wave, potential modulus maximum lines corresponding to

the R waves at characteristic scales must be determined. The detection procedure of the ECG components is inspired by findings in the [21] - [25].

The location set of modulus maxima lines is derived by searching across the largest scale to the finest scale. The algorithm searches across the scales 2^1 to 2^4 for maximum modulus lines exceeding some thresholds corresponding to the R wave. The main reason for such a search direction is that the high frequency noise does not have great compact at large scales; therefore, the modulus maximum lines created at the higher scales are barely related to the noise. In this way; since the number of modulus maxima at large scale are much fewer than those at small scale, searching the modulus maxima from large to small scale can decrease the computational effort of the algorithm. Moreover, by using an appropriate threshold for detecting the modulus maxima at large scales, the faulty detections caused by high frequency noise are reduced.

In order to eliminate the modulus maxima lines related to the P, T wave, as well as the baseline drift, some thresholds are considered at smaller scales. Based on this information, the potential R peaks are localized using the following steps:

- *Step 1* : At scale 2^4 : All of the modulus maxima larger than the threshold ϵ_R^4, Eq. A.1, are detected to obtain the location set of modulus maxima $n_k^4 \ |k = 1...N$.

$$\epsilon_R^4 = 0.5 \ \mathrm{RMS}(W_{2^4} x[n]) \tag{A.1}$$

- *Step 2* : At scale 2^3 : on the neighborhood of points found in *step 1*, n_k^4 , the modulus maxima larger than the threshold ϵ_R^3, Eq. A.2, are found.

$$\epsilon_R^3 = \mathrm{RMS}(W_{2^3} x[n]) \tag{A.2}$$

Two cases are considered:

1. If several modulus maxima exist, then the largest one is chosen. However, if the largest one is not larger than the others by a scale, (empirically found to be 1.2), the modulus maximum with its location nearest to the n_k^4 is selected .
2. If no modulus maximum exists, then all modulus points in the lower scales i.e. n_k^3, n_k^3, and n_k^1 are set to zero.

– *Step 3* : Similar to *Step 2*, the locations of the modulus maxima at scales 2^2 and 2^1 are found using the threshold ϵ_2 and ϵ_1 as follows:

$$\epsilon_i = \text{RMS}(W_{2^i} x[n]) \qquad , i = 1, 2 \tag{A.3}$$

Then, the location set of modulus maximum lines $n_k^4, n_k^3, n_k^2, n_k^1 \; | k = 1 \cdots N_1$ is obtained by searching across the detected modulus maxima lines at characteristic scales and elimination of those modulus lines with $n_k^1 = 0$.

- *Redundant and Isolated Modulus Lines*

The algorithm search across the positive maximum (negative minimum) pairs at scale 2^1, those pairs having an interval greater than $120ms$ are assigned to artifacts and the corresponding modulus maximum lines are eliminated from the set of acquired modulus maximum lines.

Next, the redundant modulus maximum lines should be eliminated. Usually, the R wave corresponds to only two modulus maximum lines. However, in the neighborhood ($120ms$) of a modulus maximum line, there are two or more modulus maximum lines having noise origins.

Analyzing the modulus lines at scale 2^3, two cases are possible, the first which is referred to as positive peak, corresponds to one positive maxima and some negative minimum; second case, which is referred to as the negative peak, is the case of one

negative minimum and some positive maxima. Imagine the case where in the neighborhood ($120ms$) of a positive maximum (negative minimum), there are two negative minimum (positive maxima). Let the two minima (maxima) be Pl and $P2$. If we define the ratio of their absolute values ($A1$ and $A2$), to the intervals between the minima and maximum ($L1$ and $L2$) as $S1$ and$S2$:

$$S1 = A1/L1 \qquad , S2 = A2/L2 \tag{A.4}$$

Then, the rules of judging redundant modulus maximum lines are as follows:

- *Rule 1*: If $S1 > 1.2\ S2$, Then : $P2$ is redundant.
- *Rule 2*: If $S2 > 1.2\ S1$, Then : $P1$ is redundant.
- *Rule 3*: Otherwise, if $P1$ and $P2$ are on the same side of the positive maximum, then the minimum farther from the maximum is redundant. If Pl and $P2$ are on different sides of the maximum, then the minimum following the maximum is redundant.

With this set up, the R wave edge with the larger slope and amplitude, or the positive R wave (corresponding to positive maximum after negative minimum in a modulus maximum pair) is selected and the redundant modulus maximum (minimum) lines are discarded.

Bibliography

[1] *Javelin Strategy and Research Report on Identity Fraud Survey Report,2007*, http://www.privacyrights.org/ar/idtheftsurveys.htm.

[2] A. K. Jain,A. Ross, S. Prabhakar,*An introduction to biometric systems*. IEEE Trans Circuit Syst Video Technol,vol. 14, no. 1,pp.4-20, 2004.

[3] E.Tatara, A.Cinar,*Interpreting ECG data by integrating statistical and artificial intelligence tools*, IEEE Engineering in Medicine and Biology Magazine, vol. 21, no. 1, pp. 36-41, Jan. 2002.

[4] J. Carlson,R. Johansson, S.B. Olsson,*Classification of electrocardiographic P-wave morphology*, IEEE Transactions on Biomedical Engineering, vol. 48, no. 4, pp. 401-405, April 2001.

[5] M. Ohlsson, H. Holst, L. Edenbrandt,*Acute Myocardial Infarction: Analysis of the ECG Using Artificial Neural Networks*, 1999.

[6] L.D. Lathauwer,B. D.Moor,J. Vandewalle,*Fetal electrocardiogram extraction by blind source subspace separation*, IEEE Transactions on Biomedical Engineering, vol. 47, no. 5, pp. 567-572, May 2000.

[7] L. Biel, O. Pettersson, L. Philipson,P. Wide,*ECG analysis:a new approach in human identification*, IEEE Trans Instrumments. vol. 50, no. 3, pp. 808-812,2001.

[8] S. A. Israel, J. M. Irvine, A. Cheng, M. D. Wiederhold,B. K. Wiederhold,*ECG to identify individuals*, Pattern Recognition vol. 38, no. 1, pp. 133-142,2005.

[9] S. A. Israel, W. T. Scruggs, W. J. Worek,J. M. Irvine,*Fusing face and ECG for personal identification*, Proceedings of 32 nd Applied Imagery Pattern Recognition workshop, pp. 226-231, 2003.

[10] T. Shen, W. J. Tompkins, Y. H. Hu,*One-lead ECG for identity verification*, Proceedings of the 2nd Conference of the IEEE Engineering in Medicine and Biology Society and the Biomedical Engineering Society, vol. 1, pp. 62-63, 2002.

[11] T. Shen, W. J. Tompkins, *Biometric statistical study of one-lead ECG features and body mass index (BMI)*, Engineering in Medicine and Biology Society, 27th Annual International Conference,pp. 1162-1165, 2005.

[12] Y. Wang, K. N. Plataniotis,D. Hatzinakos, *Integrating analytic and appearance attributes for human identification from ECG signal*, Proceedings of Biometrics Symposiums (BSYM), Baltimore, Sep. 2006.

[13] K. Platanioatis,D. Hatzinaos, and J. Lee, *ECG Biometric recognition without fiducial detection*, Proceedings of BCC'06 Baltimore, pp. 19-21, Sep. 2006.

[14] F. Agrafioti, D. Hatzinakos, *ECG based recognition using second order statistics*, Sixth Annual Conference on Communication Networks and Services Research (CNSR), May 2008, Halifax, Canada.

[15] F. Agrafioti, D. Hatzinakos, *ECG Biometric Analysis in Cardiac Irregularity Conditions*, Signal, Image and Video Processing, Springer, pp. 1863-1703.

[16] S.Z. Fatemian, D. Hatzinakos, *A new ECG feature extractor for biometric recognition*, 16th IEEE International conference on digital signal processing, Greece, July 2009.

[17] S.Z. Fatemian, D. Hatzinakos, *A Wavelet-based Aprroach to Electrocardiogram (ECG) and Phonocardiogram (PCG) Subject Recognition*, under submission to the IEEE transactions on signal processing, June 2009.

[18] L. Srnmo, P. Laguna, *Bioelectrical Signal Processing in Cardiac and Neurological Applications*,Elsevier Inc, 2005.

[19] *Human verification by heart beat signals*, http://www.berlin.ptb.de/8/84/842/BIOMETRIE/842biometriee.html.

[20] S. G. Mallat, *A Wavelet Tour of Signal Processing*,3 th edition, Elsevier Inc.,1999.

[21] J.Martinez, R. Almeida, S.Olmos, A. Rocha, P. Laguna, *A wavelet-based ECG delineator: Evaluation on standard databases* IEEE Trans Biomed Eng No.51,vol. 4, pp. 570-581, 2004.

[22] C. Li, C. Zheng, and C. Tai, *Detection of ECG characteristic points using wavelet transforms*, IEEE Trans. Biomed. Eng., vol. 42, pp.21-28, Jan. 1995.

[23] P. Laguna, R. Jan, and P. Caminal, *Automatic detection of wave boundaries in multilead ECG signals: Validation with the CSE database*, Comput. Biomed. Res., vol. 27, no. 1, pp. 45-60, Feb. 1994.

[24] J. S. Sahambi, S. Tandon, and R. K. P. Bhatt, *Using wavelet transform for ECG characterization*, IEEE Eng. Med. Biol., vol. 16, no. 1, pp.77-83, 1997.

[25] M. Bahoura, M. Hassani, and M. Hubin, *DSP implementation of wavelet transform for real time ECG wave forms detection and heart rate analysis*, Comput. Meth. Programs Biomed., no. 52, pp. 35-44,1997.

[26] B.U. Khler, C. Hennig, and R. Orglmeister, *The principles of software QRS detection*, IEEE Engineering in Medicine and Biology Society and the Biomedical Engineering Society, vol.21, pp. 42-57,Jan./Feb. 2002.

[27] O. Pahlm and L. Srnmo, *Software QRS detection in ambulatory monitoring- A review,* Med. Biol. Eng. Comp., vol. 22, pp. 289-297, 1984.

[28] F. Gritzali, *oward a generalized scheme for QRS detection in ECG waveforms*, Signal Processing, vol. 15, pp. 183-192, 1988.

[29] G. M. Friesen, *A comparison of the noise sensitivity of nine QRS detection algorithms*, IEEE Transaction on Biomedical Engineering, vol. 37, pp. 85-98,Jan. 1990.

[30] L. Srnmo, *A model-based approach to QRS delineation*, Comput. Biomed. Res., vol. 20, pp. 526-542, 1987.

[31] I. Murthy and G. D. Prasad, *Analysis of ECG from pole-zero models*, IEEE Transaction on Biomedical Engineering, vol. 39, pp. 741-751, July 1992.

[32] J. Vila, Y. Gang, J. Presedo, M. Fernandez-Delgado, and M. Malik, *A new approach for TU complex characterization*, IIEEE Transaction on Biomedical Engineering, vol. 47, pp. 764-772, June 2000.

[33] M. Nygards and L. Srnmo, *Delineation of the QRS complex using the envelope of the ECG*, Med.Biol. Eng. Comput., vol. 21, pp. 538-547,Mar. 1983.

[34] A. S. M. Koeleman, H.H. Ros, and T. J. van den Akker, *Beat-to-beat interval measurement in the electrocardiogram*, Med. Biol. Eng. Comput.,vol. 23, pp. 213-219, 1985.

[35] A. Algra and H. L. B. C. Zeelenberg, *An algorithm for computer measurement of QT intervals in the 24 hour ECG*, IEEE Computer Society Press, pp.117-119,1987.

[36] P. de Chazal and B. Celler, *Automatic measurement of the QRS onset and offset in individual ECG leads*, 18th International Conf. IEEE Eng. Med. Biol. Soc., Amsterdam, Netherlands, 1996.

[37] I. K. Daskalov, I. A. Dotsinsky, and I. I. Christov, *Developments in ECG acquisition, preprocessing, parameter measurement and recording*, IEEE Eng. Med. Biol. Mag., vol. 17, pp. 50-58, 1998.

[38] I. Daskalov and I. Christov, *Electrocardiogram signal preprocessing for automatic detection of QRS boundaries*, Med. Eng. Phys., no. 21, pp. 37-44, 1999.

[39] Med.Biol., *Automatic detection of the electrocardiogram T-wave end*, Eng. Comput., vol. 37, pp. 348-353, 1999.

[40] J. G. C. Kemmelings, A. C. Linnenbank, S. L. C. Muilwijk, A. Sippens Groenewegen, A. Peper, and C. A Grimbergen, *Automatic QRS onset and offset detection for body surface QRS integral mapping of ventricular tachycardia*, IEEE Trans. Biomed. Eng., vol. 41, pp. 830-836, Sept.1994.

[41] G. Speranza, G. Nollo, F. Ravelli, and R. Antolini, *Beat-to beat measurement and analysis of the R-T interval in 24 h ECG Holter recordings*, Med. Biol. Eng. Comput., vol. 31, no. 5, pp. 487-494, Sept. 1993.

[42] S. Meij, P. Klootwijk, J. Arends, and J. Roelandt, *An algorithm for automatic beat-to-beat measurement of the QT-interval*, in Computers in Cardiology: IEEE Computer Society Press, 1994, pp. 597-600.

[43] P. Strumillo, *Nested median filtering for detecting T-wave offset in ECGs*, Electron. Lett., vol. 38, no. 14, pp. 682-683, July 2002.

[44] E. Soria-Olivas, M. Martnez-Sober, J. Calpe-Maravilla, J. F. Guerrero-Martnez, J. Chorro-Gasc, and J. Esp-Lpez, *Application of adaptive signal processing for determining the limits of P and T waves in an ECG*, IEEE Trans. Biomed. Eng., vol. 45, pp. 1077-1080, Aug. 1998.

[45] H. Vullings, M. Verhaegen, and H. Verbruggen, *Automated ECG segmentation with dynamic time warping*, in Proc. 20th Ann. Int. Conf.IEEE Engineering in Medicine and Biology Soc., Hong Kong, 1998, pp.163-166.

[46] Z. Dokur, T. Olmez, E. Yazgan, and O. Ersoy,*Detection of ECG waveforms by neural networks*, Med. Eng. Phys., vol. 19, no. 8, pp. 738-741,1997.

[47] W. Bystricky and A. Safer, *Modelling T-end in Holter ECG's by 2-layer perceptrons*, in Proc. Computers in Cardiology. Los Alamitos, CA:IEEE Computer Society Press, 2002, vol. 29, pp. 105-108.

[48] L. Clavier, J. M. Boucher, R. Lepage, J. J. Blanc, and J. C. Cornily, *Automatic P-wave analysis of patients prone to atrial fibrillation*, Med.Biol. Eng. Comp., vol. 40, no. 1, pp. 63-71, Jan. 2002.

[49] The MIT -BIH Sinus Rhythm database, http://www.physionet.org/physiobank/database/nsrdb/

[50] The MIT -BIH Arrhythmia database, http://www.physionet.org/physiobank/database/mitdb/

[51] The PTB diagnostic ECG database, national metrology institute of germany. http://www.physionet.org/physiobank/database/ptbdb/

[52] P. Laguna, R. Mark, A. Goldberger, and G. Moody, *A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG*, in Computers in Cardiology 1997. Los Alamitos,CA: IEEE Computer Society Press, 1997, pp. 673-676.

[53] A. Taddei, G. Distante, M. Emdin, P. Pisani, G. B. Moody, C. Zeelenberg,and C. Marchesi, *The European ST -T database: Standards for evaluating systems for*

the analysis of ST-T changes in ambulatory electrocardigraphy,Eur. Heart J., vol. 13, pp. 1164-1172, 1992.

[54] J. L. Willems et al., *A reference data base for multilead electrocardiographic computer measurement programs*, J. Amer. Coll. Cardiol., vol.10, no. 6, pp. 1313-1321, 1987.

[55] A. K. Abbas and R. Bassam, *Phonocardiography Signal Processing*, Morgan and Claypool,2009.

[56] K. Phua, J. Chen, T. H. Dat, L. Shue, *Heart sound as a biometric*, Pattern Recognition, Special issues on Feature Generation and Machine Learning for Robust Multimodal Biometrics, vol. 41, no. 3, pp. 906 - 919, 2008.

[57] BF. eritelli, S. Serrano, *Biometric Identification Based on Frequency Analysis of Cardiac Sounds*, IEEE Transactions on Information Forensics and Security, vol. 2, no. 3, pp. 596-604,Sept. 2007.

[58] M. L. Jacobson, *Analysis and classification of physiological signals using wavelet transforms*, Proceedings of the 10th IEEE International Conference on Electronics, Circuits and Systems, Dec, 2003. DOI: 10.1109/ICECS.2003.1301934

[59] J. T.Willerson, J. A. Kastor, R. E. Dinsmore, E. Mundth, M. J. Buckley,W. Gerald Austen, and C. Sanders, *Non-invasive assessment of prosthetic mitral paravalvular and intravalvular regurgitation*, Brit. Heart J., vol. 34, pp. 561-568, 1972. DOI: 10.1136/hrt.34.6.561

[60] C. L.Nikias and A.P.Petropulu, *Higher-Order Spectra analysis*, PTR Prentice Hall, New Jersey, 1993.

[61] D. Braschdorff, S. Ester, T. Dorsel, and E. Most, *Neural network based multi sensor heart sound analysis*, Proc. Comput. Cardio. pp. 303-306, 1991. DOI: 10.1109/CIC.1990.144221

[62] A. K. Abbas,Rasha Bassam, *AdaptiveARMAX Cardiac Acoustic Identification*, International Conference on Modeling, Simulation and Optimization ICMSAO 2009, Al Sharjah-UAE, an, pp 209-213, 2009.

[63] P. P. Pinna, E. Piccolo, S. Bartolozzi and V. Fontana, *The FFT in the study of the fourth heart sound*, IEEE Computational Cardiology, 369-372,

[64] L. Khadra, M. Matalgah, B. El-Asir and S. Mawagdeh, *The wavelet transform and its application to phonocardiogram signal analysis*, Med. Inf. v6. 271-277,DOI: 10.3109/14639239109025301

[65] M. Matalgah, J. Knopp, S. Mawagdeh, *Interactive processing method using Gabor wavelet and the wavelet transform for the analysis of phonocardiogram signals*, Time-Frequency and Wavelets in Biomedical Signal Processing, IEEE Engineering in Medicine and Biology Society. pp. 271-304, 1997.

[66] H. P. Sava, R. Bedi, J. T. E. McDonnell, P. M. Grant, *Classification of Carpentimer-Edwards bioprosthesis heart valves using an adaptive single layer perceptro*, Proceedings of Engineering in Medicine and Biology Society, Montral, CA, Vol.1, pp. 129-130, 1995. DOI: 10.1109/IEMBS.1995.575034

[67] J. E. Hebden, J. N. Torry, *Neural network and conventional classifiers to distinguish between first and second heart sounds.* Artificial Intelligence Methods for Biomedical Data Processing, IEEE Colloquium, April, London, vol.3, pp. 1-6, 1996.

[68] M. Akay, Y. M. Akay, W. Welkowitz, *Automated noninvasive detection of coronary artery disease using wavelet-based neural networks*, Proceedings of the 16th Annual

International Conference of Engineering in Medicine and Biology Society, Baltimore, USA, vol. 1, pp. A12-A13, 1994. DOI: 10.1109/IEMBS.1994.412126

[69] P. Wang, C.S. Lim, S. Chauhan, Foo, J.Y.A., and Anantharaman, V., *Phonocardiographic Signal Analysis Using a Modified Hidden Markov Model*. Ann. Biomedical Eng., v35 i3.367-374, DOI: 10.1007/s10439-006-9232-3

[70] W. E. D. Schobben, C. W. P. Sommen, *A New Algorithm for Joint Blind Signal Separation and Acoustic Echo Canceling*, Fifth International Symposium on SignalProcessing and its Applications, ISSPA'99, Brisbane, Australia, 22-25 August, 1999.DOI: 10.1109/ISSPA.1999.815814

[71] R. J.Rangayyan,R.J.Lehner, PhonocardiogramSignal Analysis: a Review, Durand L.G.,Pibarot P.,(Digital signal processing of the phonocardiogram: Review of the most recent advances),CRC Critical Reviews in Biomedical Engineering, vol. 15, pp. 211-236, 1988.

[72] G. A. Feruglio, *A new method for producing, calibrating, and recording intracardiac sounds* J., Mar, 65:377-90, 1963. DOI: 10.1016/0002-8703(63)90013-9

[73] L.G.Gamero andR.Watrous, *Detection of the first and second heart sound using probabilistic models*, Engineering in Medicine and Biology SocietyProc. 5thAnnual International Conference of the IEEE vol.3, pp 2877-80, 2003. DOI: 10.1109/IEMBS.2003.1280519

[74] A. Haghighi-Mood and J. N. Torry, *A sub-band energy tracking algorithm for heart sound segmentation*, Computational Cardiology, pp 22-50, 1995. DOI: 10.1109/CIC.1995.482711

[75] S. Molau, M. Pitz, R. Schluter,H.Ney, *Computing Mel-frequency cepstral coefficients on the power spectrum*, IEEE International Conference on Acoustics,

Speech, and Signal Processing,ay 7-11, Utah, USA, pp. 73-76, 2001. DOI: 10.1109/ICASSP.2001.940770

[76] I.T. Jolliffe, *Principal Component Analysis, Springer Series in Statistics*, 2nd ed., Springer, NY,2002.

[77] A. M. Martinez,A. C Kak, *PCA versus LDA*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.23,No. 2, pp 228-233. DOI: 10.1109/34.908974.

[78] D. Huang, H. Leung, W. Li, *Fusion of Dependent and Independent Biometric Information Sources*, Defence R & D Canada, 2005 .