CONDITIONING, DELINEATION & STATISTICAL ANALYSIS OF MULTI-CHANNEL ELECTROCARDIOGRAM

A DISSERTATION

Submitted in partial fulfilment of the
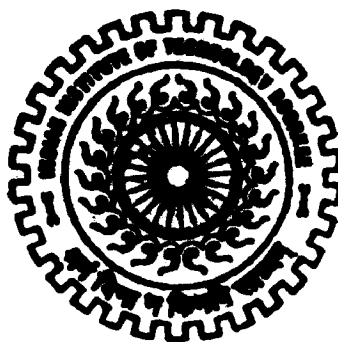*requirements for the award of the degree*
of

MASTER OF TECHNOLOGY

in

ELECTRICAL ENGINEERING

(With Specialization in Measurement & Instrumentation)

By

SEBANTI SANYAL

DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE-247 667 (INDIA)

JUNE, 2005

A Dissertation Report on

CONDITIONING, DELINEATION & STATISTICAL ANALYSIS OF MULTI-CHANNEL ELECTROCARDIOGRAM

*submitted in partial fulfillment of the
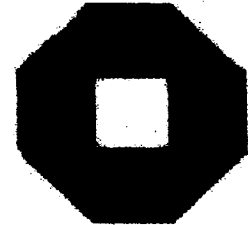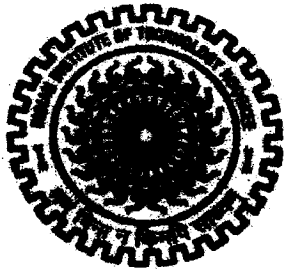requirements for the award of the degree
of*

MASTER OF TECHNOLOGY
in
ELECTRICAL ENGINEERING
(With specialization in Measurement & Instrumentation)

Submitted by

SEBANTI SANYAL
Enrollment No: 033813
EED, IITR

Supervised by

Antoun Khawaja
Institute of Biomedical Engineering,
University of Karlsruhe.

Under the esteemed guidance of

Prof. Dr. rer. nat. Olaf Dössel
Institute of Biomedical Engineering,
University of Karlsruhe, Germany.

Prof. Dr. Vinod Kumar
Electrical Engineering Department,
Indian Institute of Technology, Roorkee.

May 2005

Institut für Biomedizinische Technik
Kaiserstr. 12, (Universität)
76128 Karlsruhe, Germany
Tel: 07 21 / 608 - 26 50
Fax: 07 21 / 608 - 27 89

19.05.2005

CANDIDATE'S DECLARATION

I hereby declare that the work presented in this dissertation entitled " CONDITIONING, DELINEATION & STATISTICAL ANALYSIS OF MULTI-CHANNEL ELECTROCARDIOGRAM" submitted in partial fulfillment of the requirement for the award of the Degree of **Master of Technology** in **Electrical Engineering** with specialization in **Measurement and Instrumentation**, in the Department of Electrical Engineering, **Indian Institute of Technology, Roorkee** is an authentic record of my own work carried out from July 2004 to June 2005 under the guidance of **Dr. Vinod Kumar** (Professor, EED, IITR) and **Dr. Olaf Dössel** (Professor, Institute of Biomedical Engineering, University of Karlsruhe, Germany).

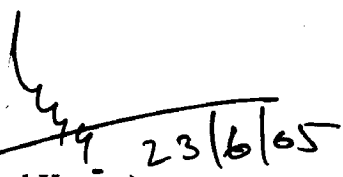I have not submitted the matter embodied in this report for the award of any other degree or diploma.

Date: 23rd June, 2004

Place: Roorkee

(SEBANTI SANYAL)

CERTIFICATE

This is to certify that the above statement made by the candidate is true to the best of my knowledge and belief.

23|6|05

(Dr. Vinod Kumar)
Professor
Department of Electrical Engineering
Indian Institute of Technology, Roorkee
Roorkee – 247667
Uttaranchal, India.

i

ACKNOWLEDGEMENTS

ABSTRACT

Multi-channel ECG Signal is recorded at the Institute of Biomedical Engineering, University of Karlsruhe on healthy volunteers. The aim of this project is to find out the pattern of variation of the QRS complex with varied respiration and heart rates.

The recorded ECG is first conditioned in order to remove noise and artifacts. A new method for baseline wander correction based on Discrete Wavelet Transform has been proposed. An algorithm using 'Haar' wavelets has been devised to automate the Multi-channel ECG delineation after filtering. The same was tested on MIT-Arrhythmia Database. The resulting sensitivity and positive predictivity was found to be highly satisfactory.

Principal component analysis is carried out on the extracted QRS waves to discover the 'basic patterns' and the 'most important deviations from mean'. Finally, a covariance matrix is calculated both in the temporal and spatio-temporal domain to represent the dependence of QRS morphology on heart rate and respiration in different individuals.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AP: Action Potential

RA: Right Atrium

RV: Right Ventricle

RBC: Red Blood Cells

LA: Left Atrium

LV: Left Ventricle

SA node: Sinoatrial node

AV node: Atrioventricular node

ECG: Electrocardiogram

EEG: Electroencephalogram

EMG: Electromyogram

EOG: Electrooculogram

WCT: Wilson's Centre Terminal

V_R: Potential measured from right arm

V_L: Potential measured from left arm

V_F: Potential measured from left foot

aVR: Augmented limb lead (right arm)

aVL: Augmented limb lead (left arm)

aVF: Augmented limb lead (left foot)

BSPM: Body Surface Potential Map

FT: Fourier Transform

STFT: Short Time Fourier Transform

WT: Wavelet Transform

CWT: Continuous Wavelet Transform

DyWT: Dyadic Wavelet Transform

DWT: Discrete Wavelet Transform

dbN: Daubechies Wavelet, Nth order

coifN: Coiflet Wavelet, Nth order

symN: Symlet Wavelet, Nth order

PCA: Principal Component Analysis

ACF: Auto-Correlation Function

CCF: Cross-Correlation Function

A.C.: Approximation Coefficients

FIR: Finite Impulse Response

LP: Low Pass

MA: Moving Average

SG: Savitzky-Golay

IIR: Infinite Impulse Response

FLDS: First Level Details Signal

DCM: Direction Change Mark

DCS: Direction Change Sharpness

WL: Window Length

IS: Incremental Step

EDCS: Extreme Direction Change Sharpness

DRM: Delineation Result Matrix

QEM: QRS Extraction Matrix

PIQM: PCA Input QRS Matrix

CHAPTER 1

INTRODUCTION

In today's competitive world, mental stress and tension have become a part and parcel in the daily life of modern people. As we are tending more and more towards the urban civilisation, we are losing touch with the nature. It is a common fact known to all that our ancestors lived a much better life although they were less privileged by science. The so-called 'fast life style' leads to the development of blood pressure and several heart diseases. Cardiac arrest is supposed to be the main cause of mortality in developed countries.

The analysis of ECG signal is extensively used as a diagnostic tool to provide information on the heart function. Multi-channel electrocardiography is carried out on healthy volunteers at the Institute for Biomedical Engineering, University of Karlsruhe. The measurement starts with the subject at rest. This is continued with measurements after physical exercise. Finally individual variations are investigated with statistical techniques for biosignal analysis.

The QRS complex, representing the contractile activity of the ventricles, is the base of analysis here. Therefore, the first concentration is on the automated delineation of ECG Signal to localize the QRS complex in different beats. Often this becomes difficult owing to the time varying morphology of the QRS and also the corruption due to noise and artefacts from various sources.

Filtering is an essential pre-processing stage before delineation. A new approach based on Discrete Wavelet Transform has been devised here to segregate the baseline wander and low-frequency disturbances from ECG. High-frequency noise and artefacts are eliminated by means of conventional Butterworth low pass or Savitzky-Golay filter.

After proper conditioning, the ECG signal is subjected to delineation. A 'Haar' wavelet based method is used in this regard. Final decision regarding the wave peaks and boundaries are made by analysing the first level details coefficients. Although this method is specifically suited for multi-channel signals, it was found to work satisfactorily on records taken from MIT-Arrhythmia Database. The methodology is extended for P and T wave delineation as well. Thereafter, QRS complexes are extracted from all the signals belonging

1

o a particular multi-channel data set. The R peaks are further aligned by means of a cross-correlation based technique.

The second stage is to carry out Principal Component Analysis on these aligned QRS complexes. The reconstruction parameters are computed from the principal components. Thereafter, a covariance matrix is calculated for each ECG dataset, which shows the correlation between the reconstruction parameters and also their relation with respiration and heart rate.

The organisation of chapters is done accordingly. The second chapter gives an overall idea of the heart function and generation of ECG waveform. The third chapter focuses on ECG measurement, the standard 12-lead system and the 32/64–channel acquisition (taken at the Institute). Fourth chapter explains the basic mathematical concepts related to this work. It describes the wavelet transform, principal component analysis and correlation technique. The fifth and sixth chapters are related to ECG signal conditioning and delineation respectively. The QRS extraction and alignment techniques are covered in the seventh chapter. The eighth chapter is the final stage, including PCA and Covariance Matrix formation. The conclusions and future directions are proposed in chapter nine. And last but not the least, the front panels of the Graphical User Interface Programs developed in Matlab 7 are given in the Appendices.

CHAPTER 2

PHYSIOLOGICAL BACKGROUND

In this chapter, we shall begin with the generation of bioelectric potentials. The heart anatomy, cardiovascular system, electrical conduction system of the heart and generation of ECG signal will be covered briefly.

2.1. SOURCES OF BIOELECTRIC POTENTIALS

In carrying out their various functions, certain systems of the body generate their own monitoring signals, which convey useful information about the functions they represent. These signals are the bioelectric potentials associated with nerve conduction, brain activity, heart beat, muscle activity, and so on. Bioelectric potentials are actually voltages produced as a result of the electrochemical activity of certain special types of cells. Through the use of transducers capable of detecting electrical voltages, these natural monitoring signals can be measured and results can be displayed in a meaningful way to aid the physician in his diagnosis and treatment of various diseases. Electrocardiogram (ECG), Electroencephalogram (EEG), Electromyogram (EMG), Electrooculogram (EOG) etc are examples of such types of signals, which represent the electrical activity of the heart, brain, muscles and eye-muscles respectively.

Certain types of cells within the body, such as nerve and muscle cells, are encased in a semi-permeable membrane that permits some substances to pass through the membrane while others are kept out. Body fluids surrounding the cells are conductive solutions containing charged atoms or ions. The principal ions are sodium ($Na+$), potassium ($K+$) and Chloride ($Cl-$). In their resting (unexcited) state, membranes of excitable cells readily permit the entry of $K+$ and $Cl-$ ions, but effectively block the entry of $Na+$ ions. The permeability for $K+$ is 50-100 times that for $Na+$ under the resting state. Since the various ions seek a balance between the inside of the cell and the outside, both according to concentration and electric charge, the inability of sodium ions to penetrate the membrane results in two conditions. First, the concentration of $Na+$ ions inside the cell becomes much lower than that in the intercellular fluid outside. Sodium ions carrying positive charge, therefore, will tend

to make the outside of the cell more positive than the inside. Secondly, in an attempt to balance the electric charge, additional K+ ions enter the cell, causing a higher concentration of potassium on the inside than on the outside. However, charge balance cannot be achieved because of the concentration imbalance of K+ ions. Equilibrium is reached with a potential difference across the membrane, negative on the inside and positive on the outside. This membrane potential is called the *resting potential* of the cell and is maintained until some kind of disturbance upsets the equilibrium. This potential ranges from –60mV to –100mV in different cells. A cell in the resting state is said to be *polarized*.

When a cell is excited by ionic current or an external stimulus, the membrane changes its characteristics and begins to allow some of the Na+ ions to enter. This movement of sodium ions into the cell constitutes an ionic current flow that further reduces the barrier of the membrane to Na+ ions. The net result is an avalanche effect in which Na+ ions literally rush into the cell to try to reach a balance with the ions outside. At the same time, potassium ions, which were in higher concentration inside the cell during the resting state, try to leave the cell but are unable to move as rapidly as the Na+ ions. As a result, the cell acquires a slightly positive potential on the inside due to the imbalance of K+ ions. This is known as the *action potential* (AP), which is about +20mV for most cells. An excited cell displaying an action potential is said to be *depolarised*. The process of changing from the resting state to the action potential is called *depolarisation*.

After some milliseconds, the rush of sodium ions through the cell membrane stops. The membrane tends to revert back to its original, selectively permeable condition. K+ ions continuously flow out of the cell leading to a repolarisation of the cell. Finally, by an active process called '*sodium pump*', the Na+ ions are transported outside the cell, and the cell is in its original state again.

Figure 2.1 shows a typical action potential waveform. The time scale depends on the type of cell. In the nerve and muscle cells, repolarisation occurs so rapidly following the depolarisation that the AP appears as a spike of as little as 1 ms duration. Heart muscles, on the other hand, repolarises much more slowly, resulting in the associated AP lasting from 150 to 300 ms.

4
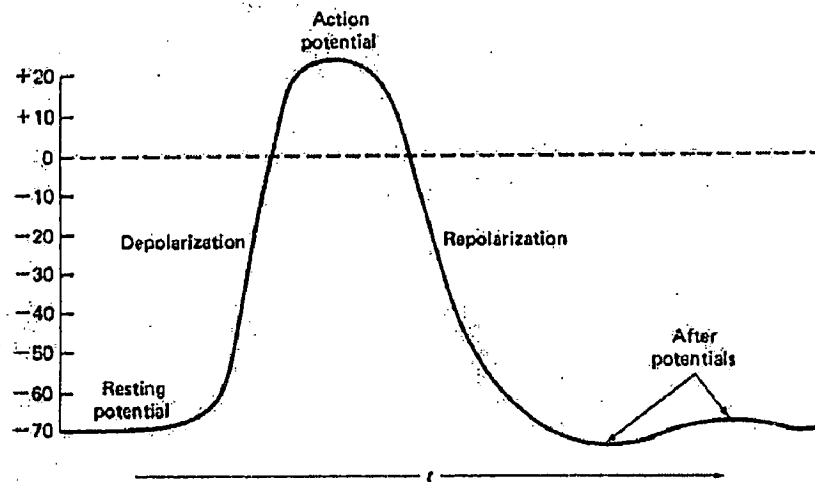
Figure 2.1: Waveform of the action potential [1]

Regardless of the method of excitation or intensity of the stimulus (provided it is sufficient to activate the cell), the AP is always the same for a given cell. This is known as the *all-or-none* law. The net height of the AP is defined as the difference between the potential of the depolarised membrane at the peak of the action potential and the resting potential.

Following the generation of the AP, there is a brief period during which the cell cannot respond to any new stimulus. This is called the *absolute refractory period* (lasting about 1 ms in nerve cells). After this, there occurs a *relative refractory period*, during which another action potential can be triggered, but a much stronger stimulation is required. In nerve cells, this period lasts for several milliseconds.

When a cell is excited and generates an AP, ionic currents begin to flow. This process can, in turn, excite neighbouring cells or adjacent areas of the same cell. Now let us see how the AP propagates in case of muscle fibre or an *unmyelinated* nerve fibre. The action potential propagates along the whole length of a fibre without decrease in amplitude by progressive depolarisation of the membrane. Current flows from a depolarised region through the intra-cellular fluid to adjacent inactive regions, thereby depolarising them. Current also flows through the extra-cellular fluids, through the depolarised membrane, and back into the intra-cellular space, completing the local circuit. The energy to maintain conduction is supplied by the fibre itself. The rate at which an action potential moves down a fibre or is propagated from cell to cell is called the *propagation rate*. The typical range for this rate is between 20-140 m/s for nerves and 0.2-0.4 m/s for heart muscles. Special time-

5

delay fibres between the atria and ventricles (to be discussed in next section) of the heart cause AP to propagate at an even slower rate, 0.03-0.05 m/s.

2.2. THE HEART ANATOMY & CARDIOVASCULAR SYSTEM

The heart is located in the chest between the lungs behind the sternum and above the diaphragm. It is surrounded by a double-layered membrane called the *pericardium*. The external layer is composed of dense fibrous tissue and the inner serous layer surrounds the heart directly. A coating layer of fluid separates the two layers of membrane, letting the heart move as it beats, still being attached to the body.

The heart is basically a four chambered organ, upper chambers called the *atria* and lower chambers known as *ventricles*. The ventricles are much larger than the atria. A wall of muscle, called the *septum* separates the left and right atria and also the left and right ventricles. Following figure shows the anatomy of heart.
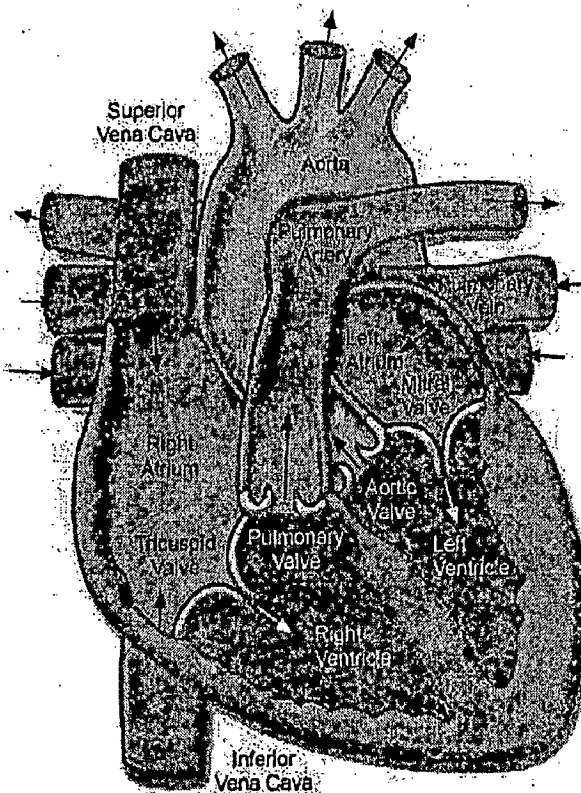
Figure 2.2: Heart Anatomy [2]

From functional point of view, the heart may be considered as a two-staged pump. The right half carries blood with carbon dioxide or CO_2 collected from different parts of the body, whereas the left half carries oxygenated blood obtained from the lungs. The

6

circulatory path for blood flow through the lungs is known as the *pulmonary circulation*. The circulatory system that supplies oxygen and nutrients to the cells of the body is called the *systemic circulation*. The cardiovascular system, as a whole, may be thought of as a complex hydraulic system comprising innumerable pipes (the *arteries* and *veins*) and a four-chambered pump (the heart).

Blood enters the heart on the right side through two main veins: the *superior vena cava*, which leads from the body's upper extremities, and the *inferior vena cava* leading from the body's organs and extremities below the heart. The incoming blood fills the storage chamber, the *right atrium* (RA). In addition to the two above-mentioned veins, the *coronary sinus* also empties into RA. The coronary sinus contains the blood that has been circulating through the heart itself via the coronary loop.

When the RA is full, it contracts and forces blood through the *tricuspid valve* into the *right ventricle* (RV). Tricuspid valve closes when the ventricular pressure exceeds the atrial pressure, thereby preventing any back flow of blood. Now RV contracts and the pressure in the ventricle forces the *pulmonary valve* to open, leading blood flow into the *pulmonary artery*. Pulmonary artery divides into two towards the two lungs. Thus, the blood from RV enters into pulmonary circulation.

The pulmonary artery bifurcates many times into smaller and smaller arteries, which become *arterioles* with extremely small cross sections. These arterioles supply blood to the alveolar capillaries, in which the exchange of oxygen and carbon dioxide takes place. The *alveoli* are the final branchings of the respiratory tree and act as the primary gas exchange units of the lung. The gas-blood barrier between the alveolar space and the pulmonary capillaries is extremely thin, allowing for rapid gas exchange. The red blood cells (RBC) are recharged with oxygen and give up their carbon dioxide. On the other side of the lung mass is a similar construction in which the capillaries feed into tiny veins, or *venules*. The latter combine to form larger veins, which in turn combine until ultimately all the oxygenated blood is returned to the heart via the pulmonary vein.

Figure 2.3 shows the schematic of cardiovascular circulation.
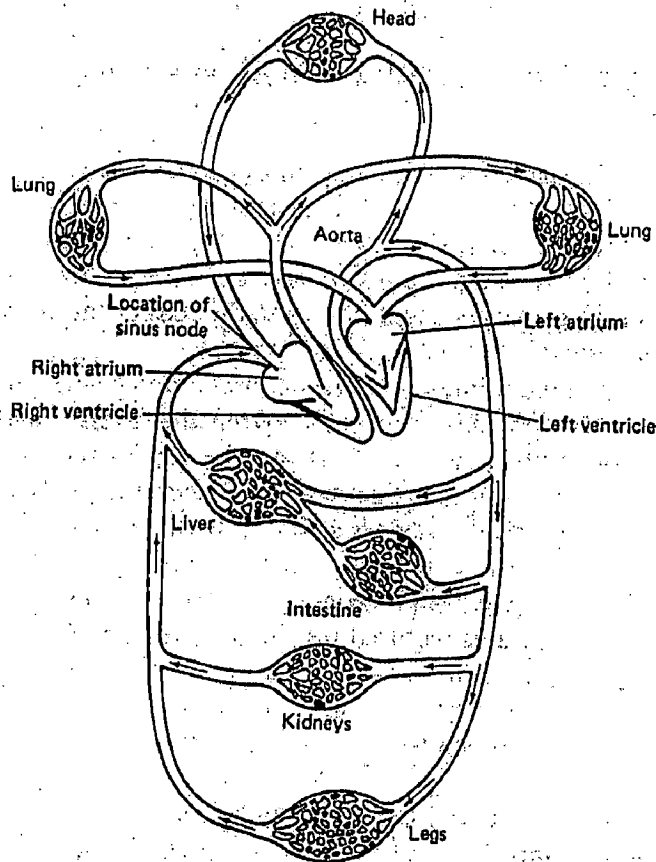
Figure 2.3: The Cardiovascular System [1]

The blood enters the *left atrium* (LA) from the *pulmonary vein*. Following the contraction of LA, blood is pumped to the *left ventricle* (LV) through *mitral* or *bicuspid valve*. After that, the ventricular muscles contract. Pressure produced by this contraction automatically closes the mitral valve and forces the *aortic valve* to open. The blood then rushes from the ventricle into the *aorta*. This action takes place synchronously with the RV as it pumps blood into the pulmonary artery.

After passing through many bifurcations of the arteries, the oxygenated blood reaches the vital organs, the brain and the extremities. A high pressure gradient exists between the arteries and veins in the systemic circulation. The left ventricle, therefore, does the most powerful pumping action in the heart.

The blood supply to the heart itself is from the aorta through the *coronary arteries* into a similar capillary system to the *cardiac veins*. This blood returns to the heart chambers through coronary sinus, as mentioned beforehand.

The heart's pumping cycle is divided into two major parts: systole and diastole. *Systole* is defined as the period of contraction of the heart muscles, specifically the ventricular muscle, during which time the blood is pumped into pulmonary artery and aorta. *Diastole* is the period of relaxation or dilation of the heart cavities as they fill with blood. Systolic (maximum) blood pressure in normal adult varies between 95 to 140 mm Hg. Normal diastolic blood pressure (lowest pressure between beats) ranges from 60 to 90 mm Hg.

The heart beats at an average rate of about 75 beats per minute (bpm) and pumps about 5 litres of blood per minute. The heart rate increases when a person stands up and decreases when he sits down, the range being from about 60 to 85 bpm. On the average, it is slightly higher in women and decreases with age. In an infant, the heart rate may be as high as 140 bpm under normal conditions. The heart rate also increases with heat exposure and other physiological and psychological factors.

2.3. ELECTRICAL CONDUCTION SYSTEM OF THE HEART & ECG

Electrical activation takes place in the heart muscle cell, or *myocyte* by means of the same mechanism as described in 2.1. Associated with the electric activation of cardiac muscle cell is its mechanical contraction, which occurs a little later.

Located in the right atrium at the superior vena cava is the *sinus node* (*sinoatrial* or SA node). This is the origin of electrical impulses in the heart and hence referred to as the 'natural pacemaker'. The SA nodal cells are self-excitatory, pacemaker cells. They generate action potential at the rate of about 70 per minute. From the sinus node, activation propagates throughout the atria, but cannot propagate directly across the boundary between atria and ventricles, because of the non-conducting barrier of fibrous tissue.

The *atrioventricular node* (AV node) is located at the boundary between the atria and ventricles. It has an intrinsic frequency of 50 pulses/min. However, if the AV node is triggered with a higher pulse frequency, it follows this higher frequency. In a normal heart, the AV node provides the only conducting path from the atria to ventricles. Thus, under normal conditions, the latter can be excited only by pulses that propagate through it.

9

Sinus node
70/min
A-V node
50/min

Bundle of His
Bundle branches
Purkinje fibers
15-30/min
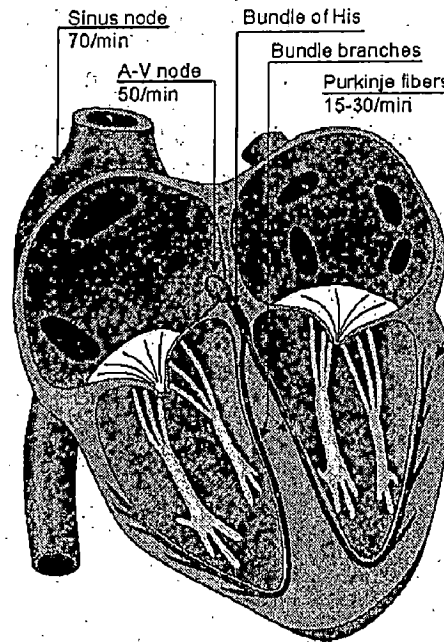
Figure 2.4: Electrical Conduction System of the Heart [3]

Propagation from the AV node to ventricles is provided by a specialized conduction system. Proximally, this system is composed of a common bundle, called the *Bundle of His*. More distally, it separates into two bundle branches propagating along each side of the septum, called the *right* and left *bundle branches*. Even more distally the bundles ramify into *Purkinje fibres* that diverge to the inner sides of the ventricular walls. Propagation along the conduction system takes place at a relatively high speed once it is within the ventricular region, but prior to this (through the AV node) the velocity is extremely slow.

From the inner side of the ventricular wall, the many activation sites cause the formation of a wavefront, which propagates through the ventricular mass towards the outer wall. This process results from cell-to-cell activation. After each ventricular muscle region has depolarized, repolarization occurs.

Because the intrinsic rate of the sinus node is the greatest, it sets the activation frequency of the whole heart. If the connection from the atria to the AV node fails, the AV node adopts its intrinsic frequency. If the conduction system fails at the bundle of His, the ventricles will beat at the rate determined by their own region that has the highest intrinsic frequency.
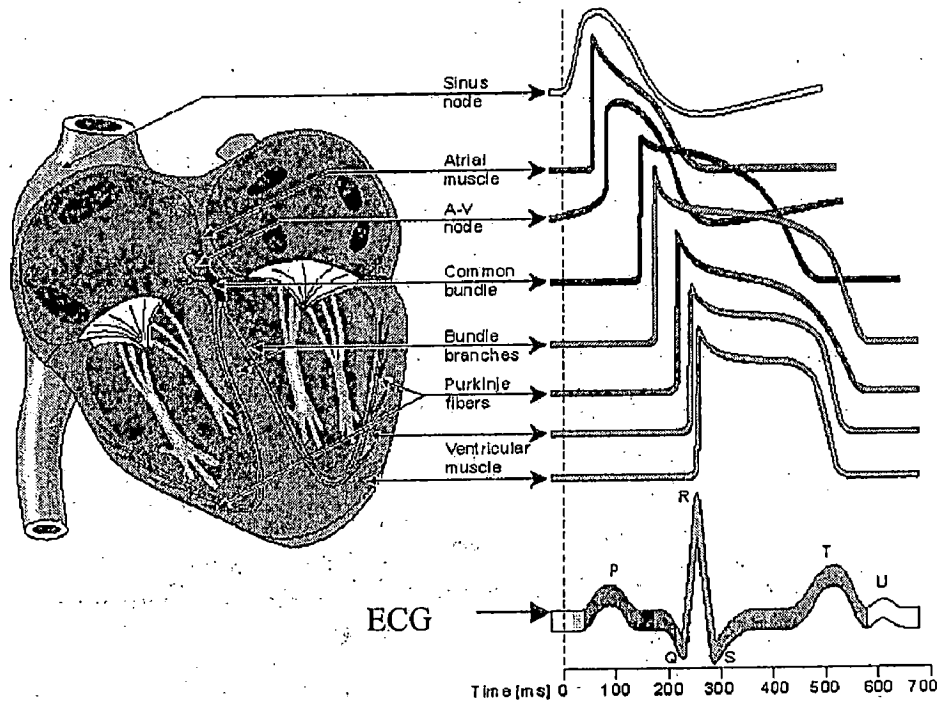
Figure 2.5: Propagation of Electrical Impulse inside the Heart [3]

The Electrocardiogram (ECG) is the electrical manifestation of the contractile activity of the heart, and can be recorded fairly easily with surface electrodes on the limbs or chest. The rhythm of the heart in terms of beats per minute may be easily estimated by counting the readily identifiable waves. ECG waveshape is altered by cardiovascular diseases and abnormalities such as myocardial ischemia and infarction, ventricular hypertrophy and conduction problems [4]. Following figure shows the ECG waves, peaks, segments and intervals more clearly.
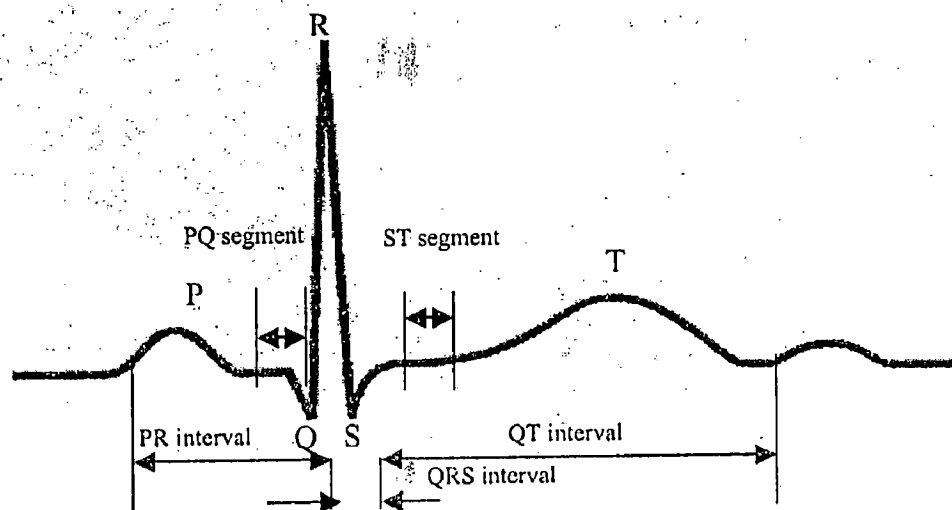
Figure 2.6: One ECG Cycle representing one Heart Beat

11

As we have already observed in figure 2.5, a cardiac cycle is reflected in a period of the repetitive ECG signal as the series of waves labelled as P,QRS, and T. If we view the cardiac cycle as a series of events, we have the following epochs in an ECG waveform [4]:

1. **The P Wave:** This wave is related to atrial depolarisation. Contraction of the atria is triggered by SA node impulse. The atria do not possess any specialized conduction nerves as the ventricles do. Contraction of the atrial muscles takes place in a slow squeezing manner, with the excitation stimulus being propagated by the muscle cells themselves. Therefore, the P wave is a slow waveform, with duration of about 80 ms. The amplitude of P wave is in the range of 0.1-0.2 mV.

2. **The PQ Segment:** The AV node provides a delay to facilitate completion of the atrial contraction and transfer of blood to the ventricles before ventricular contraction is initiated. This results in an iso-electric PQ segment of about 80 ms duration.

3. **The QRS Wave:** This is the manifestation of ventricular depolarisation. The specialised system of Purkinje fibres stimulate contraction of ventricular muscles in a rapid sequence. The almost-simultaneous contraction of the entire ventricular musculature results in a sharp and tall QRS complex of about 1mV amplitude and 80-100 ms duration. While comparing the amplitude and duration of the P and QRS waves, the smaller size of the atria in comparison to the ventricles should be taken care of.

4. **The ST Segment:** The normally flat iso-electric ST segment is related to the plateau in the action potential of the left ventricular muscle cells (referring to figure 2.1). Its duration is about 100-120 ms.

5. **The T Wave:** It is commonly referred to as the wave corresponding to ventricular repolarisation. It relates to the last phase of action potential of the ventricular muscle cells, when the potential returns from the plateau of the depolarised state to the resting potential through the process of repolarisation. The normal amplitude and duration of the T wave are 0.1-0.3 mV and 120-160 ms. This is also a slow wave.

Unlike ventricular repolarisation, atrial repolarisation does not produce any distinct waveform in the ECG as it is overshadowed by the following QRS complex.

CHAPTER 3

ECG MEASUREMENT

This chapter will give an overview of the surface electrodes used in ECG measurement and the lead systems.

3.1. BASICS

As the human body forms a volume conductor, the bioelectric phenomena occurring inside the conductor generate a signal measurable at the skin surface. The action potential propagation in the heart muscle is counted among these types of phenomena. The resulting ECG, therefore, can be measured with the help of non-invasive surface electrodes placed on the thorax or limbs.

The intervening space between the patient's skin and the electrode is filled with a coating of adhesive gel, which provides a conducting path. The most-used material for electrodes these days is *silver-silver chloride* (Ag-AgCl) since it approximates a nonpolarizable electrode. Following figure shows such an electrode. Its advantage over other types is a very small offset potential.
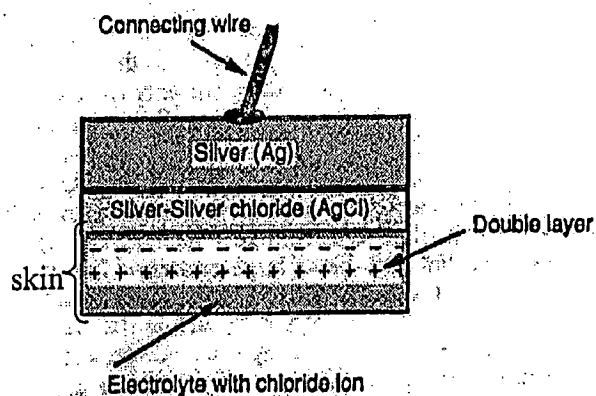
Figure 3.1: Ag-AgCl Electrode [5]

The instrument used to obtain and record the electrocardiogram is called an *electrocardiograph*. The electrocardiograph was the first electrical device to find widespread use in medical diagnostics, and it still remains the most important tool for the diagnosis of

cardiac disorders. Figure 3.2 shows the elementary building block of a typical electrocardiograph. However, the components may vary in different systems.
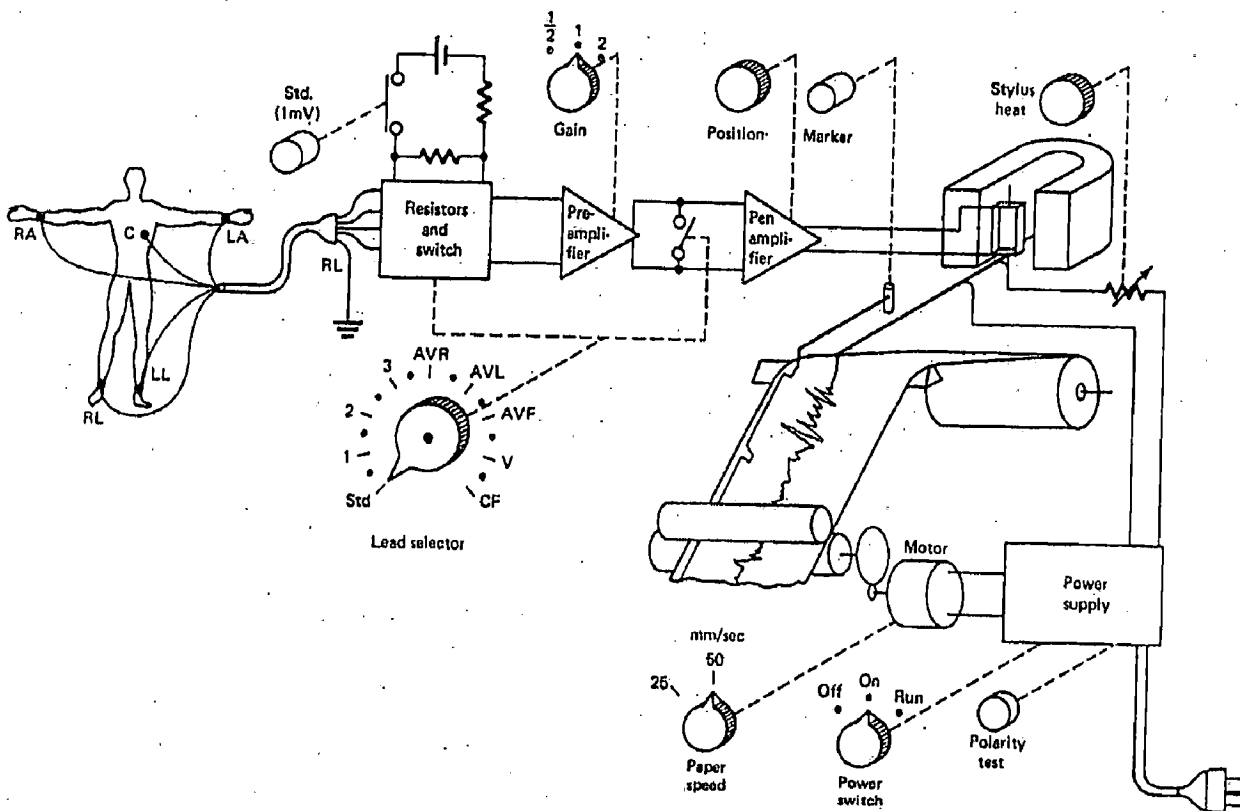
Figure 3.2: Electrocardiograph Building Blocks [1]

The electrode positions can be seen on the body of a human being on the left side of the figure 3.2. The placement of electrodes on the body is referred to as the *ECG lead system*. Here, first the standard 12-lead ECG will be discussed and then the multi-lead or multi-channel ECG, which is being recorded in the Institute.

3.2. STANDARD 12-LEAD SYSTEM

In clinical practice, the standard 12-channel ECG is obtained using *four limb electrodes* and *six chest electrodes*. The variations in electrical potentials in 12 different directions out of the ten electrodes are measured. These 12 different electrical views of the activity in the heart are normally referred to as leads.

Potentials V_L, V_R and V_F are recorded from left arm, right arm and left foot respectively. The right foot is used to place the reference electrode. We get the three lead voltages I, II and III from V_L, V_R and V_F as follows:

$$I = V_L - V_R$$
$$II = V_F - V_R \quad\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(3.1)$$
$$III = V_F - V_L$$
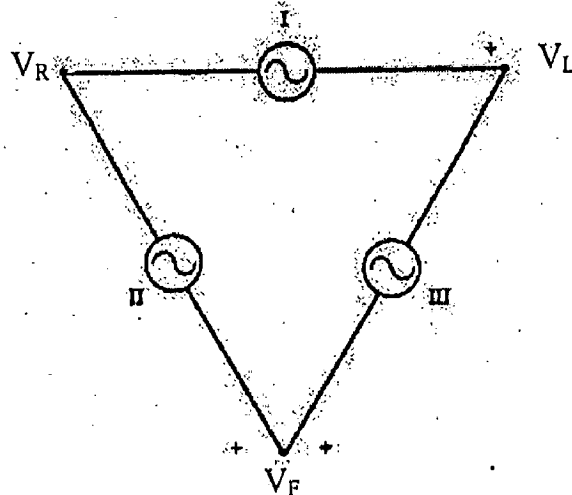
Figure 3.3: Derivation of the Three Main Leads [5]

According to Kirchoff's law,

$$II = I + III \quad\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(3.2)$$

The hypothetical equilateral triangle formed by leads I, II and III (as in Figure 3.3) is known as *Einthoven's triangle*. The centre of the triangle represents *Wilson's Central Terminal* (WCT). Schematically, the heart is assumed to be placed at the centre of this triangle.

The potential at WCT (V_{WCT}) is given as:

$$V_{WCT} = \frac{V_L + V_R + V_F}{3} \quad\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(3.3)$$

This is taken as the reference for chest leads.

There are three other limb leads known as *augmented limb leads*. These are marked as aVR, aVL and aVF (aV for the augmented lead, R for the right arm, L for the left arm, and F for the left foot). These leads are obtained by using the exploring electrode on the limb indicated by the lead name, with the reference being Wilson's central terminal without the exploring limb lead as shown in the following figure.
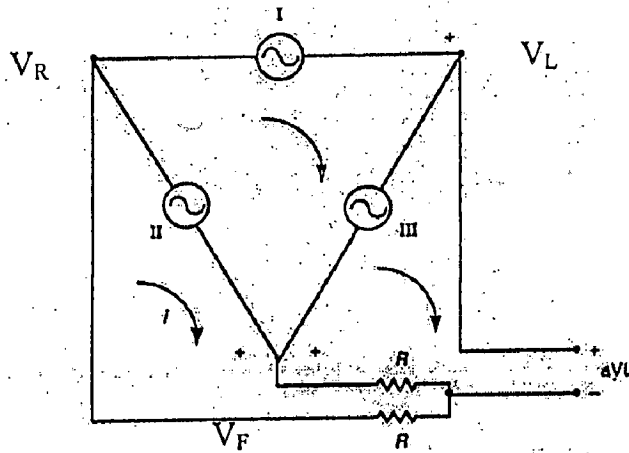
Figure 3.4: Measurement of the Augmented Limb Lead aVL [1]

We can express the augmented leads in terms of leads I, II and III as follows:

$$aVL = \frac{I - III}{2}$$

$$aVR = \frac{-I - II}{2} \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(3.4)$$

$$aVF = \frac{II + III}{2}$$

Following figures illustrate the vector relationships among the limb leads.
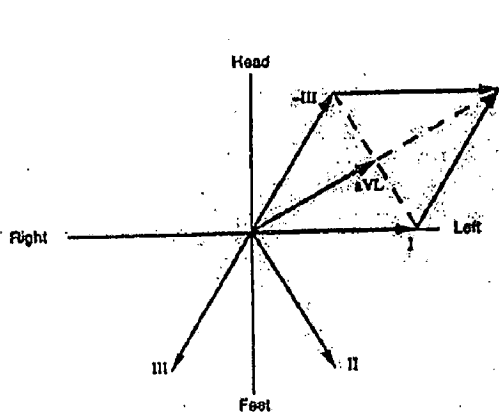
Figure 3.5: aVL Derived from I & III [5]
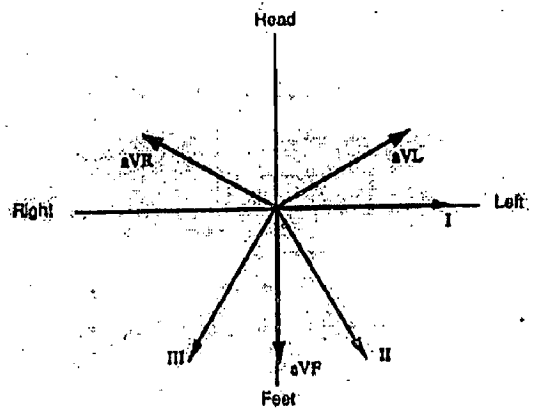
Figure 3.6: Vector Relationships [5]

We now combine all the limb leads in terms of Einthoven's triangle and WCT as in figure 3.7.

16
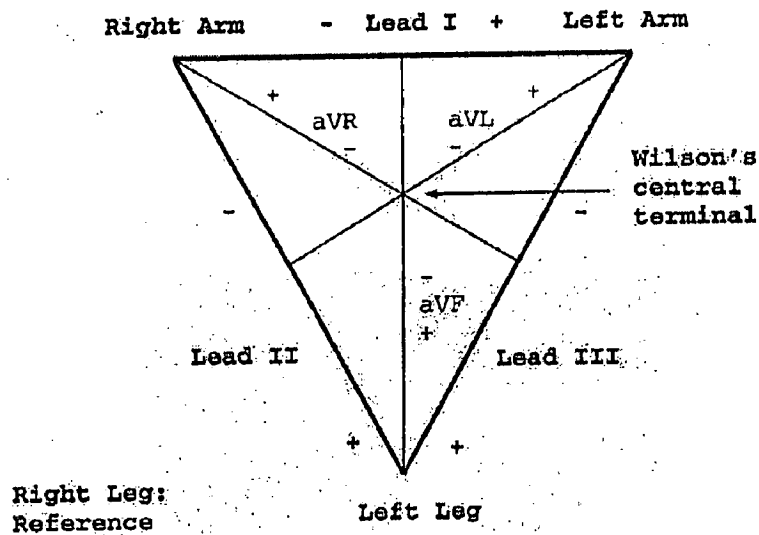
Figure 3.7: Einthoven's Triangle[4]

For measuring the potentials close to the heart, Wilson introduced the *precordial leads* (chest leads) in 1944. These leads, V1-V6 are placed over the left chest as described in figure 3.8. The points V1 and V2 are located at the fourth *intercostal space* just to the right and left of the *sternum*, respectively. V4 is located in the fifth intercostal space at the *midclavicular line*. The location of V3 is half-way between V2 and V4. V5 and V6 leads are at the same horizontal level as V4, but on the *anterior axillary line* and the *midaxillary line*, respectively.
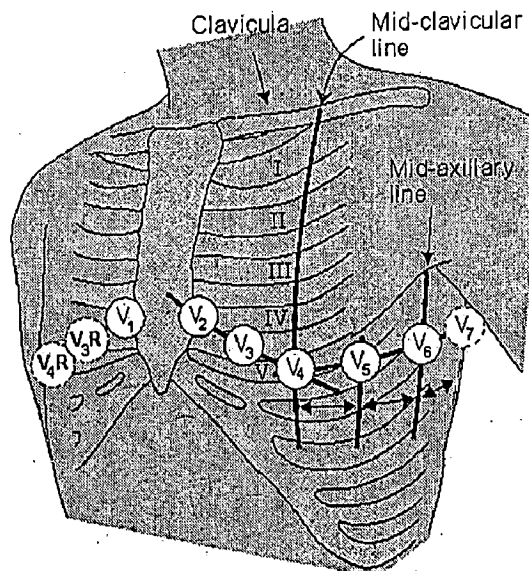
Figure 3.8: Placement of Six Precordial Electrodes [3]

In the above figure, the leads V_4R and V_3R are not used in 12-lead ECG system.

Figure 3.9 summarizes the electrode positions (both limb and chest) for the 12-lead system.

Bipolar limb leads

(Augmented) Unipolar limb leads

Unipolar chest leads

$V_1 - V_6$
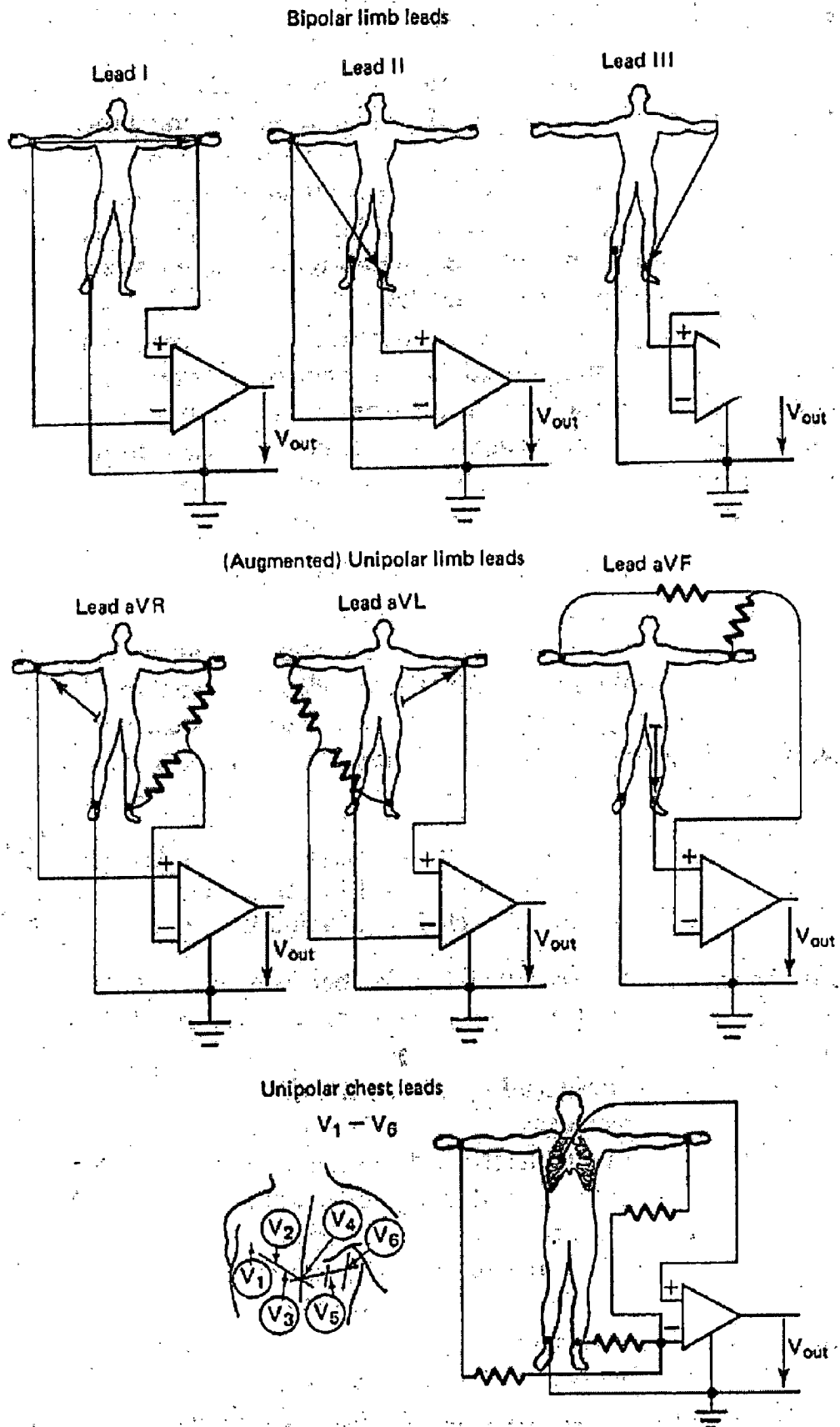
Figure 3.9: Configuration of the 12 Leads[1]

3.3. MULTICHANNEL ECG AND ITS UTILITY

Multichannel ECG is taken so as to get a volume information about the electrical activity of the heart, instead of a 2-D information taken using the famous 12-lead system. Another aim of multichannel electrocardiography is to get the Body Surface Potential Map (BSPM).

BSPM is a temporal sequence of potential distributions observed on the thorax throughout one or more electrical cardiac cycles. In practice, tens or hundreds of unipolar ECGs are recorded, either simultaneously or individually with subsequent time-alignment. For each time instant of interest, the potential measured in each lead is associated with the spatial location at which it was measured. The spatial distribution of the set of potentials at a particular instant in the cardiac cycle may then be displayed as isopotential contour maps showing lines which connect all sites which have the same potential. BSPM may give additional diagnostic information for all types of heart diseases.

64 and 32 Channel ECG records are taken at the Institute of Biomedical Engineering, University of Karlsruhe using two different acquisition systems.

Following diagram 3.10 shows the electrode positions for 32-channel ECG acquisition.
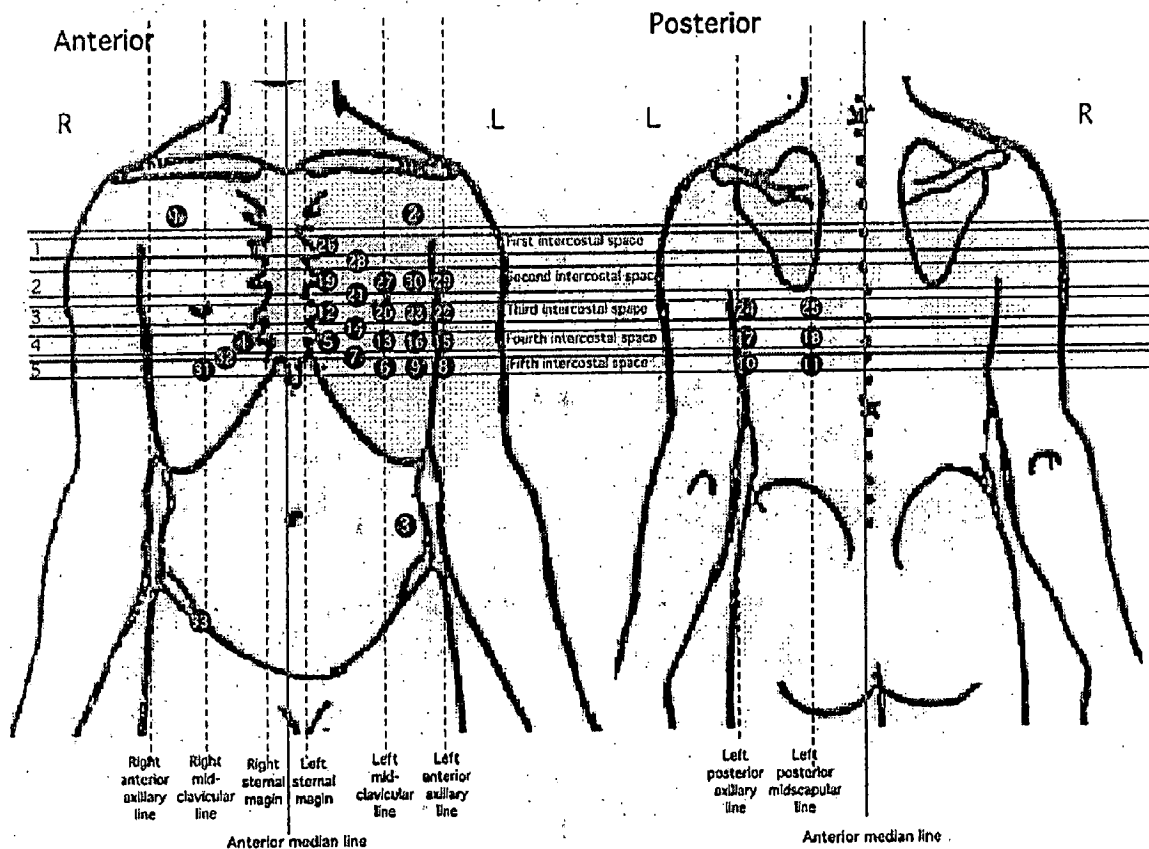
Figure 3.10: Electrode Positions for 32-Channel Recording [6]

Table 3.1 gives the description of electrode positions corresponding to Figure 3.10.

Electrode	Name	Position
1	RA	A point in the right interclavicular fossa, medial to the border of the deltoid muscle and 2 cm below the lower border of the clavicle
2	LA	A point in the left interclavicular fossa, medial to the border of the deltoid muscle and 2 cm below the lower border of the clavicle
3	LL	Halfway between the costal margin and the Iliac crest
4	V1	Right sternal margin, fourth intercostal space
5	V2	Left sternal margin, fourth intercostal space
6	V4	Left midclavicular line, fifth intercostal space
7	V3	Midway between V2-V4
8	V6	Left midaxillary line V4-V5 level
9	V5	Left anterior axillary line, V4 level
10	V7	Left posterior axillary line V4-V5 level
11	V8	Left midscapular line V4-V5 level
12	V2 1H	Left sternal margin, third intercostal space
13	V4 1H	Left midclavicular line, fourth intercostal space
14	V3 1H	Midway between V2 1H-V4 1H
15	V6 1H	Left midaxillary line V4 1H-V5 1H level
16	V5 1H	Left anterior axillary line, V4 1H level
17	V7 1H	Left posterior axillary line V4 1H-V5 1H level
18	V8 1H	Left midscapular line V4 1H-V5 1H level
19	V2 2H	Left sternal margin, second intercostal space
20	V4 2H	Left midclavicular line, third intercostal space
21	V3 2H	Midway between V2 2H-V4 2H
22	V6 2H	Left midaxillary line V4 2H-V5 2H level
23	V5 2H	Left anterior axillary line, V4 2H level
24	V7 2H	Left posterior axillary line V4 2H-V5 2H level
25	V8 2H	Left midscapular line V4 2H-V5 2H level
26	V2 3H	Left sternal margin, first intercostal space
27	V4 3H	Left midclavicular line, second intercostal space
28	V3 3H	Midway between V2 3H-V4 3H
29	V6 3H	Left midaxillary line V4 3H-V5 3H level
30	V5 3H	Left anterior axillary line, V4 3H level
31	V3 R	Right sternal margin, fourth intercostal space
32	V2 R	Right midclavicular line, fifth intercostal space
33	RL(ref)	In the region of the right Iliac fossa*

* As recommended by the American College of Cardiology

Table 3.1 : Description of Electrode Positions [6]

The limb electrodes of 12-lead system are moved onto the thorax for multi-channel ECG acquisition, as per the recommendation of Mason and Likar [7].
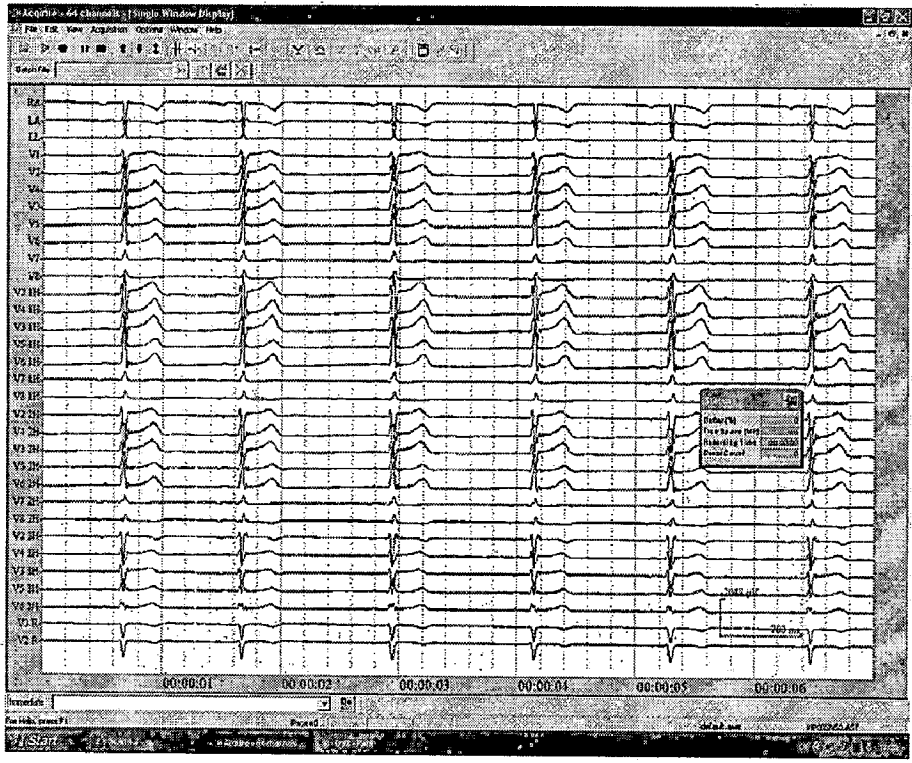
Figure 3.11: 32-Channel ECG Record taken at the Institute [6]

Figure 3.12 shows the self-explanatory electrode positions for 64-channel ECG acquisition.
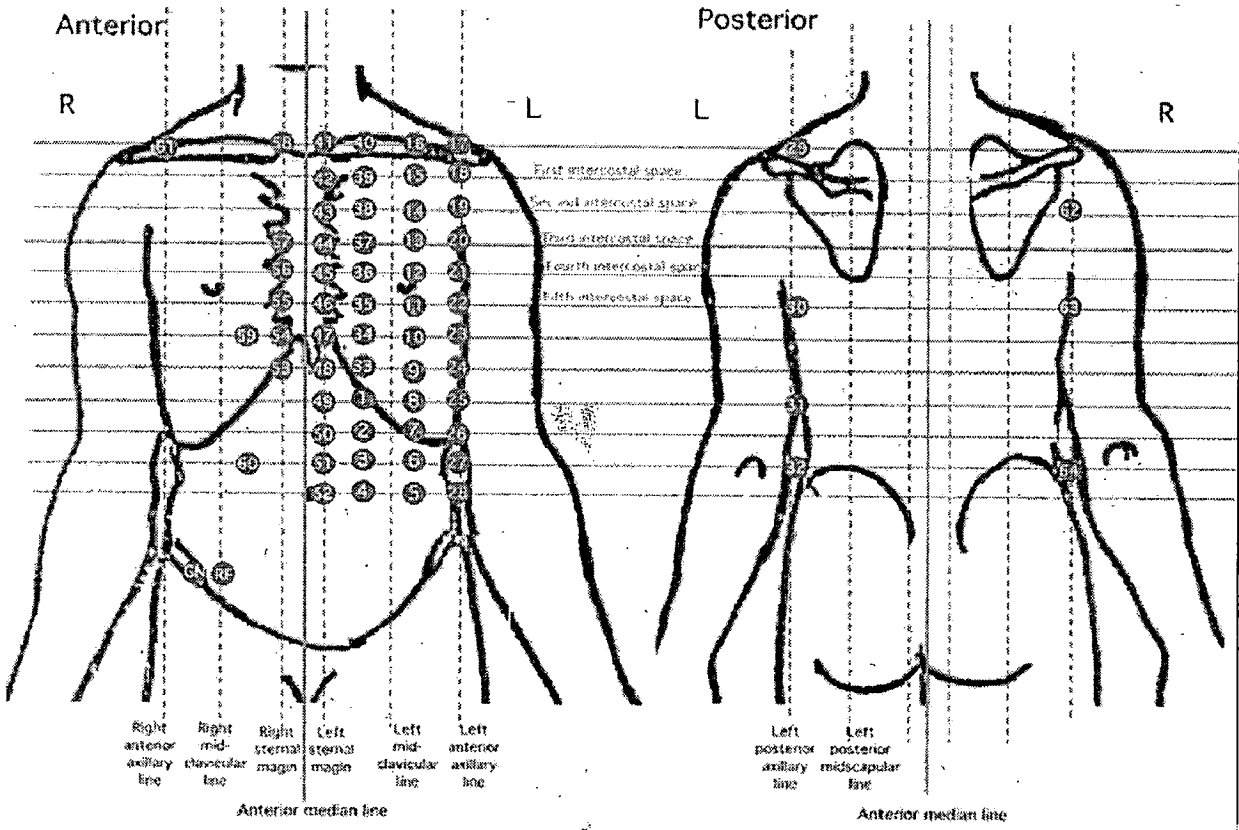
Figure 3.12: Electrode Positions for 64-Channel Recording [6]

CHAPTER 4

BACKGROUND MATHEMATICS

This chapter forms the backbone of my work. The important mathematical concepts used throughout the rest of the thesis will be described here. We shall start with Wavelet Transform, which is used for ECG delineation and filtering as well. Next we will move on to Principal Component Analysis, which finds application in Chapter 8. Correlation techniques, which are applicable for any biomedical signal analysis, will also be covered at the end.

4.1. WAVELET TRANSFORM

4.1.1. DEVELOPMENT OF WAVELET THEORY: HISTORICAL BACKGROUND

Mathematical transformations are applied to a signal to obtain further information, which is not readily available in its original time-domain form.

Fourier Transform (FT) is the oldest of all transforms used in signal processing. Historically, *Joseph Fourier* (1770-1830) first introduced the remarkable idea of expansion of a function in terms of a trigonometric series. FT decomposes a signal into complex exponential functions of different frequencies. For a continuous signal x(t), FT is defined as follows:

$$X(f) = \int_{-\infty}^{\infty} x(t).e^{-j2\pi ft} dt \quad\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(4.1)$$

The analysis coefficients (or the spectra) X(f) are computed as *inner products of the signal with sinusoidal basis functions of infinite duration*. The trigonometric kernel exp(-j2πft), used here, oscillates indefinitely, and hence, the localized information contained in the signal x(t) gets lost [8]. While the spectrum X(f) shows the overall strength with which any frequency f is contained in the signal x(t), it does not generally provide easy-to-interpret information about the time-localization of spectral components [9]. The analysis coefficients X(f) define the notion of global frequency 'f' in a signal [10].

However, time domain and frequency domain constitute two alternative ways of looking at a signal. Although FT allows a passage from one domain to the other, it does not allow a combination of the two. This method enables us to investigate problems either in the time domain or in the frequency domain, but not simultaneously in both. Fourier transform theory has been very useful for analysing harmonic signals, or signals for which there is no need for local information [8][9].

Fourier analysis is therefore an effective tool for studying stationary signals (with time-independent frequency content). However, many of the practically encountered signals (like the ECG as we shall see in the next section) are non-stationary. A complete analysis of non-stationary signals requires a joint time-frequency representation.

The basic idea of *time-frequency representations* of signals is to map a one-dimensional signal of time, x(t), into a two-dimensional function of time and frequency, $T_x(t,f)$. Thus, they combine time-domain and frequency-domain analyses to yield a

potentially more revealing picture of the temporal localization of a signal's spectral components [9].

In order to incorporate both time and frequency localization properties in FT, *Dennis Gabor* in 1946 first introduced the *windowed Fourier Transform* or *Short Time Fourier Transform*. His major idea was to use a *time-localization window function* g(t-τ) for extracting local information from the Fourier transform of a signal. The parameter τ corresponds to the position of the window in time. 'τ' is kept on varying to translate the window until the whole of time-domain is covered. The width of this window must be less or equal to the segment of the signal where stationarity is valid (i.e. frequency does not change).

$$STFT_x^{(g)}(\tau, f) = \int_t [x(t).g^*(t-\tau)].e^{-j2\pi ft} dt \quad \text{.......................(4.2)}$$

Although STFT overcomes the drawback of Fourier Transform apparently, it has got a serious problem related to the resolution in time and frequency. The root of this problem goes back to *Heisenberg's Uncertainty Principle*, according to which exact time-frequency representation of a signal is not possible. Therefore, we can never know precisely which of the spectral components exists at what instants of time. What we can know is the time interval during which a certain band of frequency exists. A broader window gives better frequency resolution and poor time resolution. On the contrary, the time resolution can be improved at the cost of frequency resolution with shorter window. Once the window is chosen for STFT, the resolution in time and frequency domain gets *fixed*. However, many signals encountered in our practical life (ECG being no exception) require a more flexible approach regarding this resolution [11].

Wavelet transform (WT) was developed to overcome this *fixed resolution problem* of STFT. The Multi-Resolutional Approach (MRA) in time and frequency domain is the heart of WT.

4.1.2. DEFINING WAVELETS

The basis of FT, as we have already seen, is sinusoidal waves of infinite duration. Fourier transform decomposes the signal of interest into sinusoids of different frequencies.

On the other hand, WT decomposes the original signal into a set of compactly supported basis functions called wavelets (small waves), obtained from a single prototype mother wavelet by means of dilation and translation (as we will see later). Wavelets are localized waves of finite energy. They have their energy concentrated in time or space and are suited to analyse transient signals.
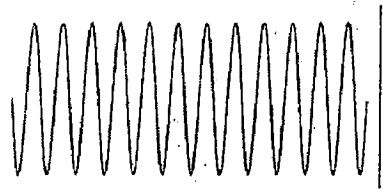
(a) Wave   (b) Wavelet

Figure 4.1: Wave and Wavelet [12]

Comparison between Wave and Wavelet:

Wave	Wavelet
1. Oscillating Function of time.	1. Oscillating Function of time.
2. Duration Infinite.	2. Duration Finite.
3. Smooth and Predictable.	3. Irregular and Asymmetric.
4. Periodic	4. Aperiodic
5. Ex: Sine & Cosine Waves	5. Ex: Daubechies, Haar,
Symlet,	Coiflet Wavelet Families.
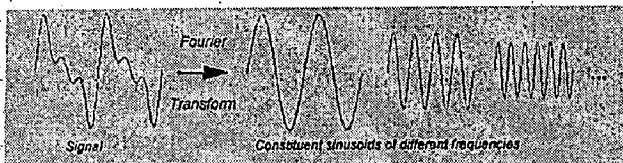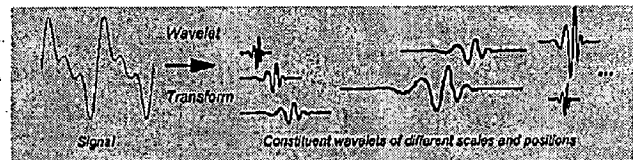
Figure 4.2:   (a) Fourier Decomposition   (b) Wavelet Decomposition [13]

4.1.3. USEFULNESS OF WT IN ECG ANALYSIS

Electrocardiogram signal, by its very nature is a non-stationary one. It is characterized by a cyclic occurrence of patterns with different frequency contents (QRS complexes, P and T waves) [7]. Each of these patterns represents a certain distinct event as discussed in 2.3. The QRS complex, representing ventricular depolarisation, occurs as a high frequency, high amplitude spike of very small duration in the ECG cycle. On the other hand, P wave (atrial

depolarisation) or T wave (ventricular repolarisation) have got relatively low frequency contents, leading to their smoother appearance.
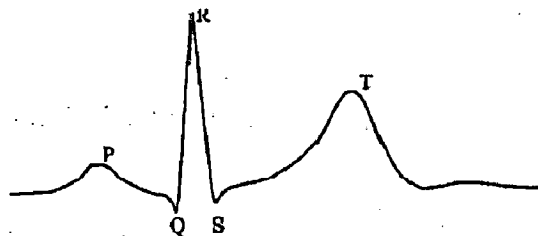
Figure 4.3: Elements of ECG waveform

The overall frequency range of ECG signal is 0.5-100 Hz with amplitude in ranges of mV. Low frequency components span for almost the entire duration, whereas the high frequency features (e.g. QRS complex) occur as transients. Moreover, the noise and artifacts affecting the ECG signal also appear at different frequency bands and time intervals. By decomposing into elementary building blocks that are well localized in time and frequency, WT can distinguish ECG waves from serious noise, rtifacts and baseline drift [15].

The mentioned characteristics of ECG signal makes wavelet transform the most suitable tool to deal with it. The wavelets have time-width adapted to their frequencies. The scale of wavelet represents its frequency range. The resolution of wavelets at different scales varies in the time and frequency domains — finer temporal resolution and coarser frequency resolution at higher scales and just the opposite at lower scales. In the application of ECG, it can be well appreciated that the R Peak needs to be marked with better temporal resolution, as compared to other waves during the cardiac cycle.

4.1.4. CONTINUOUS WAVELET TRANSFORM (CWT)

The Continuous Wavelet Transform (CWT) was developed as an alternative to STFT to overcome the fixed resolution problem in time-frequency plane. The Wavelet Transform provides a description of the signal in the time-scale domain (scale being in a sense opposite to frequency), allowing the representation of the temporal features of the signal at different resolutions. The wavelet (of a particular scale) is shifted along the signal and for every position of it, the similarity between a localised section of the original signal and the wavelet is examined by calculating CWT coefficients. Then this process is repeated many times with

a slightly compressed (or dilated) wavelet for every new cycle. At the end, the result will be a collection of time-scale representation of the signal, all with different resolutions.

The continuous wavelet transform can be defined mathematically as an *inner product of the signal and the wavelet basis functions.*

$$CWT_x^{\psi}(\tau,s) = \int x(t)\psi^*_{\tau,s}dt \dots\dots\dots(4.3)$$

The basis wavelet functions $\psi_{\tau,s}(t)$ are derived from a single prototype *Mother Wavelet* $\psi(t)$ as follows:

$$\psi_{\tau,s} = \frac{1}{\sqrt{s}}\psi(\frac{t-\tau}{s}) \dots\dots\dots\dots(4.4)$$

Therefore, the transformed signal is obtained as a function of two variables, namely, τ and s, the *translation* and *scaling coefficient*, respectively. The translation and dilation (or, scaling) operations are performed on the mother wavelet to produce the basis wavelet functions.

The term *translation* is related to the location of the wavelet, as it is shifted through the signal x(t). This term corresponds to time information in the transform domain. Translating or shifting a wavelet means hastening or delaying its onset as shown below.

(a) Wavelet Function $\psi(t)$ (b) Wavelet Function translated by τ, $\psi(t-\tau)$
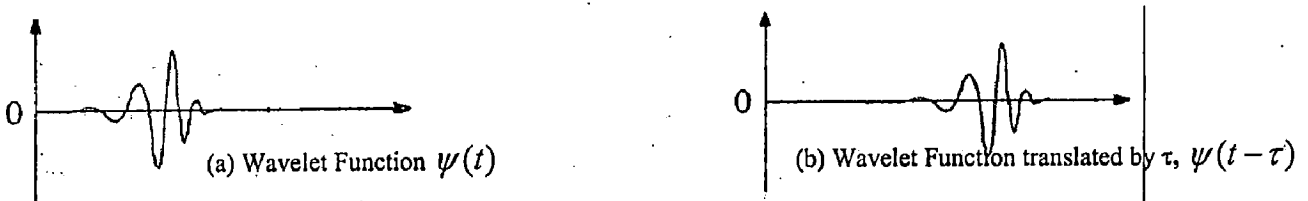
Figure 4.4: Translation of Wavelet Function [13]

The parameter *scale* in the wavelet analysis is similar to that used in maps. High scale gives a gross or global picture of the signal, whereas low scale corresponds to a detailed view. Similarly, in terms of frequency, low frequencies correspond to a global information of a signal (that usually spans the entire signal), whereas high frequencies correspond to a detailed information of a transient pattern in the signal (having relatively short duration). That is why scaling conveys a notion of something reciprocal to the frequency. Scaling, as a mathematical operation, either dilates or compresses a signal. Following figures illustrate the concept of scaling more clearly. The term '$\frac{1}{\sqrt{s}}$' in equation

(4.4) serves the purpose of *energy normalisation* of the wavelet across various scales.
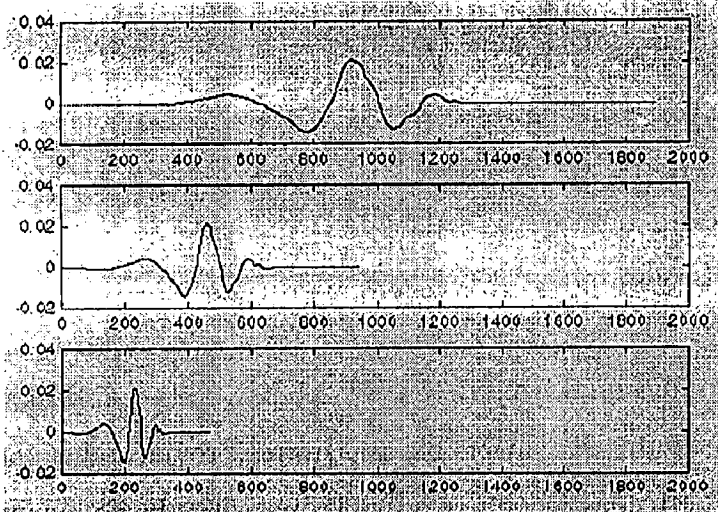
$f(t) = \psi(t)$; s =1

$f(t) = \psi(2t)$; s =1/2

$f(t) = \psi(4t)$; s =1/4

Figure 4.5: Scaling of Wavelet Function [13]

We can describe the process of CWT through the following steps and associated figure 4.6 [13].

1. Take a wavelet and compare it to a section at the start of the original signal.

2. Calculate a number, C, that represents how much similar the wavelet is with this section of the signal. The higher the value of C, more is the similarity. The value of C will depend on the shape of wavelet chosen. This C is nothing but the CWT coefficient at the initial (or smallest) scale, say $s = s_1$ and translation $\tau = 0$ (or, the value of τ corresponding to the initial position of the wavelet).

C = 0.0102
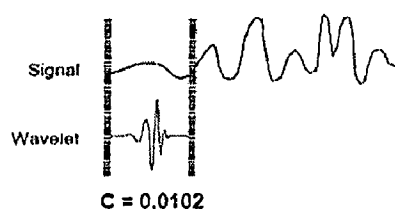
Figure 4.6 (a): Wavelet with smallest scale at initial position

3. Shift the wavelet to the right and repeat step 2 at each position, until the whole signal is covered.
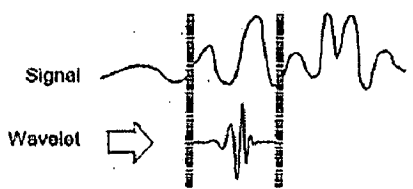
Figure 4.6 (b): Same scale Wavelet shifted to right

4. Repeat steps 1 to 3 taking a scaled (stretched) wavelet.
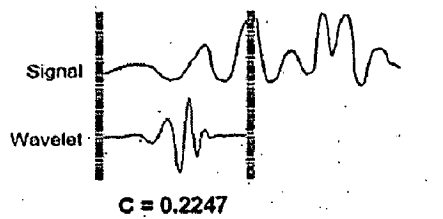
Figure 4.6 (c): Wavelet with a different scale

5. Repeat steps 1 through 4 for all scales.

Reconstruction of original time-domain signal from its CWT: Admissibility Condition

The continuous wavelet transform is a reversible transform, provided that the *admissibility condition* is satisfied. The reconstruction is possible by using the following reconstruction formula (*Inverse Wavelet Transform*):

$$x(t) = \frac{1}{c_\psi^2} \int_s \int_\tau CWT_x^\psi(\tau,s) \frac{1}{s^2} \psi\left(\frac{t-\tau}{s}\right) d\tau ds \dots\dots\dots\dots(4.5)$$

Where, C_ψ is a constant that depends on the wavelet used. The success of the reconstruction depends on this constant called, *the admissibility constant*, to satisfy the following *admissibility condition*:

$$c_\psi = \left\{ 2\pi \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\xi)|^2}{|\xi|} d\xi \right\}^{1/2} < \infty \dots\dots\dots\dots(4.6)$$

Where, $\hat{\psi}(\xi)$ is the FT of $\psi(t)$. Equation (4.6) implies that $\hat{\psi}(0) = 0$, which is

$$\int \psi(t) dt = 0 \dots\dots\dots\dots(4.7)$$

Equation (4.7) is not a very restrictive requirement since many wavelet functions can be found whose integral is zero. For equation (4.7) to be satisfied, the wavelet must be oscillatory.

4.1.5. CLARIFYING MULTI-RESOLUTIONAL APPROACH (MRA)

Unlike STFT which has a constant resolution at all times and frequencies, WT uses a Multi-Resolutional Approach (MRA), i.e. varying temporal resolution for different spectral components, which can be clarified as follows. A lower or narrower scale (higher frequencies) means lesser ambiguity in time, i.e. good time resolution. Higher scales (lower frequencies) have wider support, leading to more ambiguity in time, or in other words, poor

30

temporal resolution. The following figure compares the resolution for four different representations of the same signal.
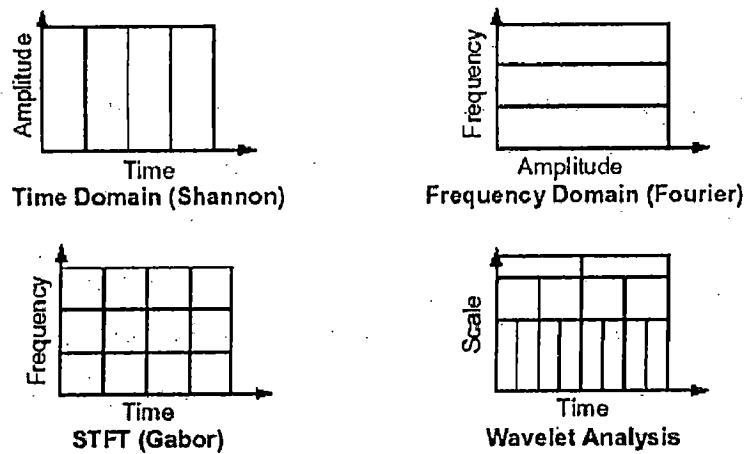
Figure 4.7: Time-Frequency Resolution at different Signal Representations [13]

The original time-domain signal has got no time resolution problem, since we know the value of the signal at every instant of time. In the Fourier transformed version, there is no resolution problem in the frequency domain, i.e. we know precisely what frequencies exist. Conversely, the frequency resolution in time domain and time resolution in Fourier domain are zero, since we have no information about them. For the two bottom diagrams, each box represents an equal area of the time-frequency plane, but different sized boxes giving different proportion to time and frequency.

All the boxes are of same size for STFT, i.e. the time and frequency resolutions are constant all over the time-frequency plane. For wavelet transform, at low frequencies (high scales), the height of the boxes are shorter (which corresponds to better frequency resolution, since there is less ambiguity regarding the value of the exact frequency), but their widths are longer (which correspond to poor time resolution, since there is more ambiguity regarding the value of the exact time). At higher frequencies (low scales), width of the boxes decreases, i.e. the time resolution gets better, and height of the boxes increases, i.e. the frequency resolution gets poorer.

4.1.6. DYADIC WAVELET TRANSFORM (DYWT)

This can be thought of as an intermediate stage between Continuous and Discrete Wavelet Transform. The CWT assigns a value to the continuum of points on the translation-scale

plane. Therefore, the computation takes a long time. Dyadic Wavelet Transform (DyWT) is based on sampling the translation-scale plane.

First the scale parameter (s) is discretized using a logarithmic rule. The base of logarithm is generally taken as 2. The scales in power of 2 are only considered (e.g. 2,4,8, and so on).

Scale discretization is given as follows,

$$s = s_0^{\,j} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(4.8)$$

where, s_0 is the base of logarithm (>1) and j=1,2,3...........,n. The number 'n', representing total number of different 'j' values, is determined considering the bandwidth of the signal x(t).

From Nyquist's Rule, we know that at higher scale (i.e. lower frequencies) the sampling rate can be reduced. In other words, if the translation-scale plane needs to be sampled with a sampling rate of N_1 at scale s_1, the same plane can be sampled with a sampling rate of N_2 at scale s_2, where $s_1 < s_2$ (corresponding to frequencies $f_1 > f_2$) and $N_1 < N_2$ [4]. The actual relationship between N_1 and N_2 is ,

$$N_2 = \frac{s_1}{s_2} N_1 \quad \text{or} \quad N_2 = \frac{f_2}{f_1} N_1 \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(4.9)$$

Therefore, at lower frequencies, the sampling rate can be reduced saving a considerable amount of computation time.

After discretizing the scale parameter (s), the translation parameter (τ) is discretized with respect to 's', i.e. a different sampling rate is used for every 's'. Following figure demonstrates this idea. For the lowest scale (s=2), we sample 32 points along the translation axis. This means, the wavelet at this scale will be compared with the signal x(t) at 32 distinct values of shift. At the next scale value (i.e. s=4), the sampling rate of translation axis is reduced by a factor of 2 and so on.
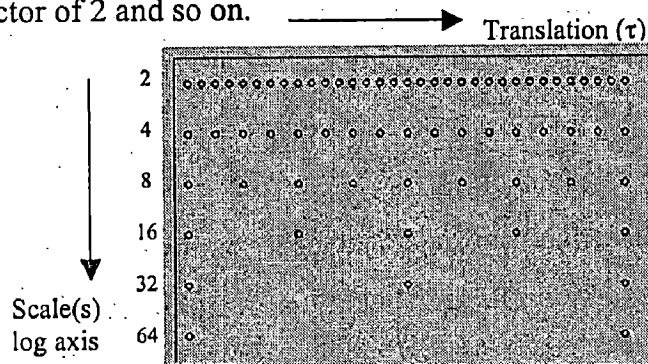
Figure 4.8: Discretization of Translation-Scale Plane [11]

In accordance with equation (4.8), the translation discretization is given as,

$$\tau = k.s_0^j.\tau_0 \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(4.10)$$

where, $\tau_0 > 0$ and k = 1, 2, 3,.................m. The number 'm', representing total number of different 'k' values, is determined considering the time span of the signal x(t).

The basis wavelet functions $\psi_{j,k}(t)$ of DyWT is derived from the mother wavelet $\psi(t)$ as follows,

$$\psi_{j,k}(t) = s_0^{-j/2}\psi(s_0^{-j}t - k\tau_0) \dots\dots\dots\dots\dots\dots\dots\dots\dots(4.11)$$

DyWT is defined mathematically as follows,

$$DyWT_x^\psi(j,k) = \int x(t)\psi_{j,k}^*(t)dt \dots\dots\dots\dots\dots\dots\dots\dots\dots(4.12)$$

Although DyWT has got computational efficiency over CWT, still it provides high degree of redundancy as far as data reconstruction is concerned. This redundancy, on the other hand consumes a significant amount of computational resources. That is why we move on to wavelet implementation based on digital filters (or the 'Discrete Wavelet Transform') for discrete time signals , which is amazingly fast in operation.

4.1.7. DISCRETE WAVELET TRANSFORM (DWT)

Basically, the discrete wavelet transform is meant to handle discrete-time signals. The DWT is considerably easier to implement when compared to the CWT. The DWT provides sufficient information both for analysis and synthesis of the original signal, with a significant reduction in the computation time.

The foundations of DWT go back to 1976 when techniques to decompose discrete time signals were devised. In the case of DWT, time-scale representation of a digital signal is obtained using *digital filtering* techniques. The signal to be analysed is passed through filters with different cut-off frequencies at different scales. Wavelets can be realized by *iteration of filters with rescaling*. The resolution of the signal, which is a measure of the amount of detail information in the signal, is changed by the filtering operations, and the scale is changed by up-sampling and down-sampling (sub-sampling) operations.

The DWT is computed by successive low-pass and high-pass filtering of the discrete time-domain signal as shown in the following figure. This is called the *Mallat*

Algorithm or Mallat-tree decomposition. Its significance is in the manner it connects the continuous time multi-resolution to discrete-time filters. The signal is denoted by the sequence x[n], integer 'n' denoting the sample number. G_0 and H_0 are the low and high pass *Analysis filters* (filters used for decomposition) respectively. At each level, the high pass filter produces *detail* information d[n], whereas, the low pass filter associated with scaling function produces coarse *approximations*, a[n].
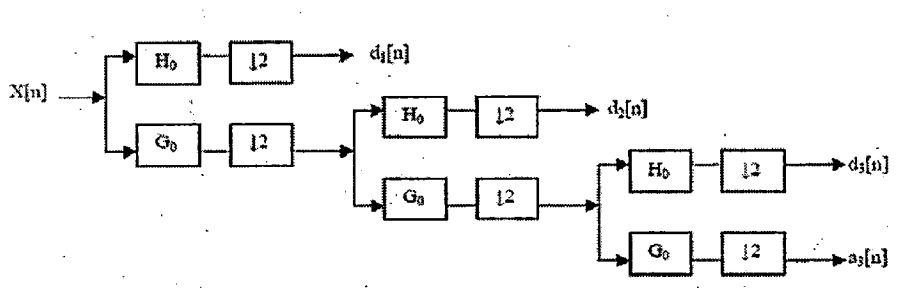
Figure 4.9: Three-level Wavelet Decomposition Tree [12]

At each decomposition level, the half band filters produce signals spanning only half the frequency band. This doubles the frequency resolution, as the uncertainty in frequency is reduced by half. According to *Nyquist's rule*, the sampling frequency of a signal needs to be at least double of its maximum frequency content in order to have a successful reconstruction. However, after each level of decomposition, the maximum frequency itself gets halved, and hence now its sampling frequency can also be reduced proportionally without any loss of information. This decimation by 2 halves the time resolution as the entire signal is now represented by only half the number of samples. This also doubles the scale.

We can see the following example to conceptualise the appearance of details and approximation coefficients across different levels.

cA, cD: Approximations and Details
time-aligned to original signal
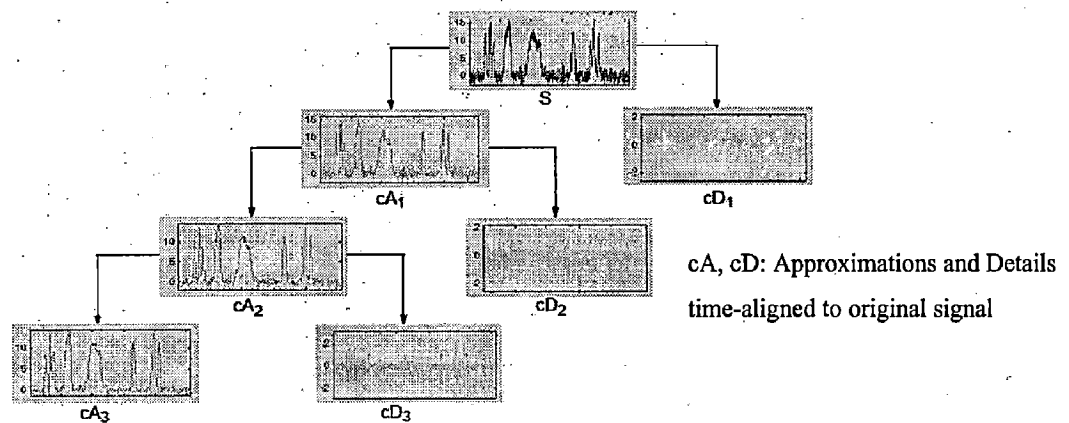
Figure 4.10: Illustration of Approximation and Detail Coefficients [13]

The filtering and decimation process is continued until the desired level is reached. The maximum number of levels depends on the length of the signal. The DWT of the original signal is then obtained by concatenating all the coefficients, a[n] and d[n], starting from the last level of decomposition.

Reconstruction of original time-domain signal from its DWT: Orthogonal Filters

The reconstruction is basically the reverse process of decomposition. The approximation and detail coefficients at every level are up-sampled by two, passed through the low pass and high pass *synthesis filters* (G_1 and H_1) and then added. This process is continued through the same number of levels as in the decomposition process to obtain the original signal. The Mallat Algorithm works equally well if the analysis filters, G_0 and H_0, are exchanged with the synthesis filters, G_1 and H_1.
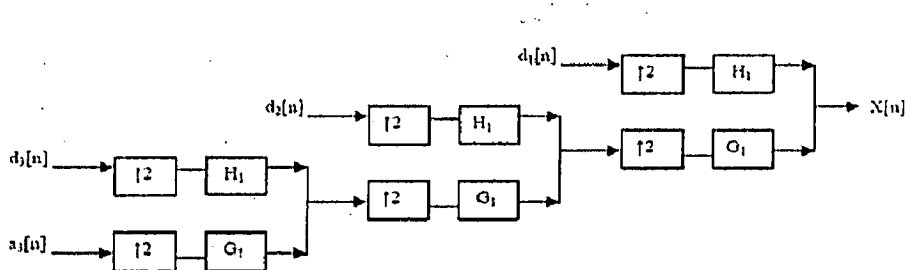
Figure 4.11: Three-level Wavelet Reconstruction Tree [12]

To achieve desired reconstruction features, we will be using only *orthogonal filters* [5] in discrete domain. Coefficients of orthogonal filters are real numbers. The filters are of the same length and not symmetric. The low pass filter, G_0 and the high pass filter, H_0 are related to each other by,

$$H_0(z) = z^{-N}G_0(-z^{-1}) \qquad\qquad (4.13)$$

Equation (4.13) means that the two filters are alternated flip of each other. These are also called *Conjugate Mirror Filters* (CMF). Also, for perfect reconstruction, the synthesis filters are identical to analysis filters except for a time reversal.

Orthogonal Wavelet Families

We will discuss the mother wavelets namely, Haar, Daubechies, Symlets and Coiflets here. These all are compactly supported orthogonal wavelets and capable of perfect reconstruction.

Scaling Function and *Wavelet Functions* are associated with Low-pass and High-pass filtering respectively. In this section, 'ϕ' denotes the Scaling Function and 'ψ', the Wavelet Function.

Daubechies Wavelets (dbN, N:filter order):

scaling function, ϕ →

← wavelet function, ψ

decomposition low-pass filter →

← decomposition high-pass filter

reconstruction low-pass filter →

← reconstruction high-pass filter
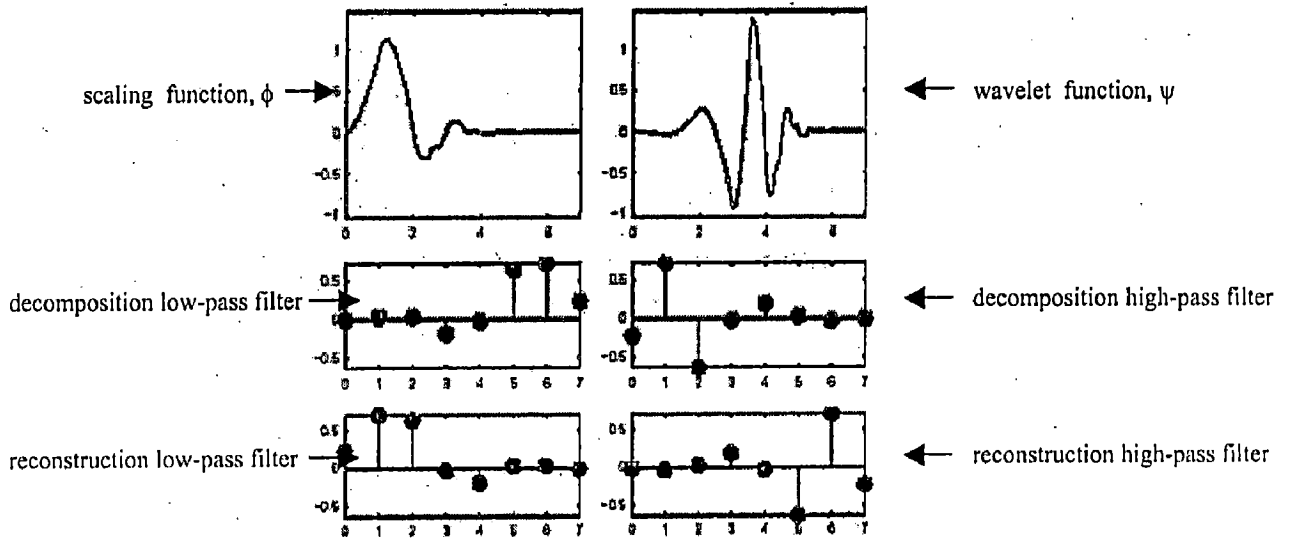
Fig 4.12: Daubechies wavelet db4 [13]

Daubechies wavelets are also called Maxflat wavelets as their frequency responses have maximum flatness at frequencies 0 and π.

Haar Wavelet:

This is nothing but the db1 wavelet, i.e. the Daubechies wavelet of order 1. This is the simplest wavelet imaginable. The wavelet function is a simple step function.

The wavelet function (ψ) and scaling function (ϕ) are defined as follows:

$$\psi(x) = 1, \quad \text{if } 0 \le x \le \frac{1}{2}$$
$$= -1, \quad \text{if } \frac{1}{2} \le x \le 1$$
$$= 0, \quad \text{if } x \notin [0,1]$$

$$\phi(x) = 1, \quad \text{if } x \in [0,1]$$
$$= 0, \quad \text{if } x \notin [0,1]$$
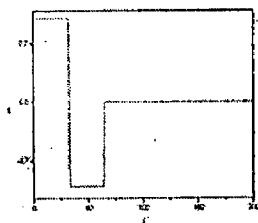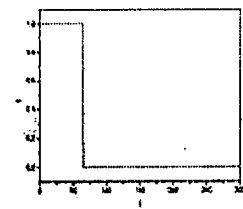
(a) wavelet function

(b) scaling function

Figure 4.13: Haar wavelet[16]

36

Symlet Wavelets (symN, N:filter order):

Symlet wavelets are modifications of Daubechies wavelets in order to improve symmetry, while retaining great simplicity. Still, Symlets are only near symmetric.
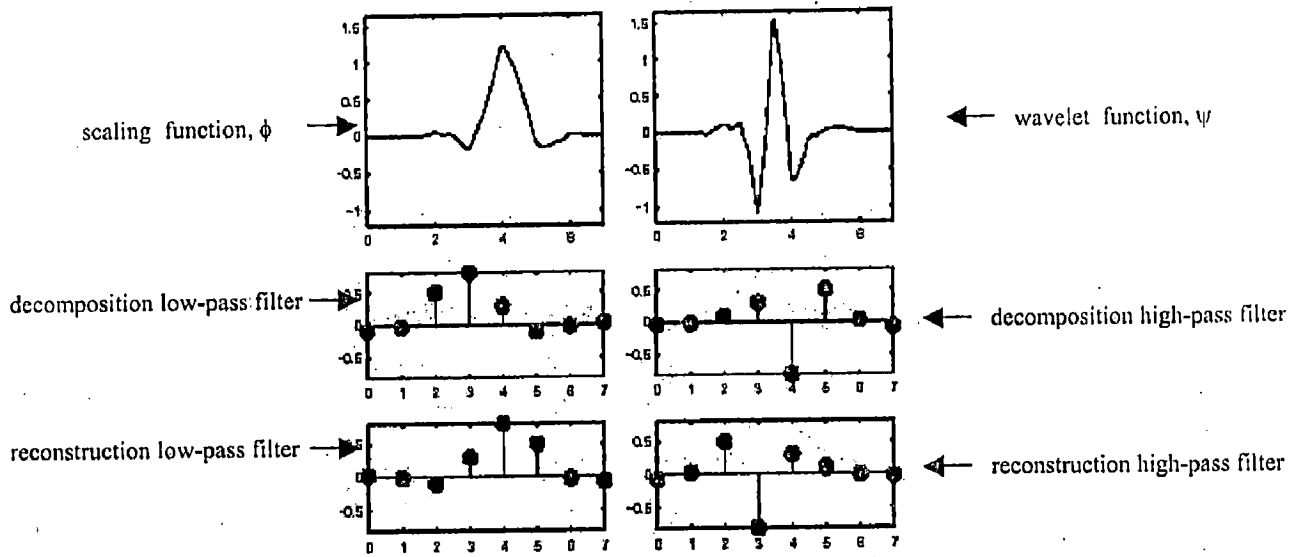
Fig 4.14: Symlet, sym4 [13]

Coiflet Wavelets (coifN, N:filter order)::

The 'coifN' ψ and ϕ functions are much more symmetrical than 'dbN'.
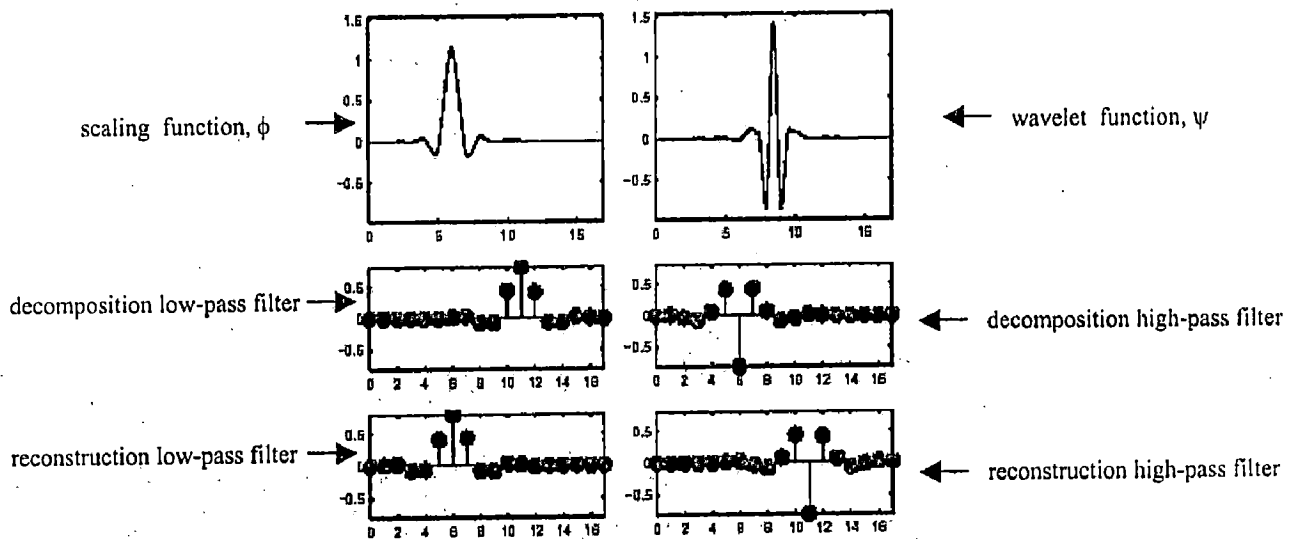
Fig 4.15: Coiflet, coif3 [13]

4.2. PRINCIPAL COMPONENT ANALYSIS

4.2.1. DEFINING PCA

Principal Component Analysis (PCA) is an exploratory multivariate statistical technique that allows the identification of key variables (or combinations of variables) in a multidimensional data set that best explains the differences between observations. Different dimensions convey the notion of different variables observed against equal span of time. So, each dimension is a time-vector and all the time-vectors are of same length. Given 'n' observations on 'm' variables, the goal of PCA is to reduce the dimensionality of the data matrix by finding 'r' ($\leq$ m) new variables. These 'r' principal components account together for as much of the variance in the original 'm' variables as possible, while remaining mutually uncorrelated and orthogonal [17].

In our application, each variable (or a 'dimension') refers to a QRS complex extracted from multi-channel ECG data after conditioning and delineation. PCA is applied both in temporal and spatio-temporal domain to find out a minimal set of QRS complexes characteristic to that data-set which conveys relevant diagnostic information .

Principal Component Analysis is designed to capture the variance in a data set in terms of principal components [19]. PCA involves a mathematical procedure that transforms a number of (possibly) correlated variables into a (smaller) number of uncorrelated variables called principal components. The first principal component accounts for as much of the variability in the data as possible, and each succeeding component accounts for as much of the remaining variability as possible [20].

4.2.2 RELEVANT MATHEMATICAL TERMS

Variance:

This is a measure of the spread of a variable from the mean. Let us consider 'n' observations of a variable A, namely, a_1, a_2, ..., a_n. The variance of A can be defined as:

$$S_A = \frac{\sum_{i=1}^{n}(a_i - \bar{a})^2}{(n-1)} \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(4.14)$$

38

Where, $\bar{a}$ denotes the mean. $[\bar{a} = (a_1 + a_2 + \ldots\ldots\ldots + a_n)/n\ ]$

The square root of variance is known as the *Standard Deviation*.

Covariance:

For two-dimensional data (two different variables having the same no. of observations), covariance provides a measure to find out how much the dimensions vary from the mean with respect to each other [5].

The covariance between two variables A = $(a_1, a_2, \ldots\ldots\ldots, a_n)$ and B = $(b_1, b_2, \ldots\ldots\ldots, b_n)$ can be mathematically defined as:

$$S_{AB} = \frac{\sum_{i=1}^{n}(a_i - \bar{a})(b_i - \bar{b})}{(n-1)} \quad\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4.15)$$

Covariance Matrix:

Covariance is always calculated between two dimensions. However, if we have a data set with more than two dimensions, there is more than one covariance measurement that can be calculated [21].

Let us consider three variables A, B and C with the same number of observations. Here, we can calculate three different covariance measures, namely, S_{AB}, S_{BC} and S_{CA}. Following is the covariance matrix for these three variables:

$$S = \begin{bmatrix} S_{AA} & S_{AB} & S_{AC} \\ S_{BA} & S_{BB} & S_{BC} \\ S_{CA} & S_{CB} & S_{CC} \end{bmatrix} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4.16)$$

The diagonal terms are nothing but the variances, e.g. $S_{AA} = S_A$ and so on. Also, the covariance matrix is a symmetric one, i.e. $S_{AB} = S_{BA}$ and so on.

For an m-dimensional data set, we have $\frac{m!}{(m-2)! \times 2}$ different covariance values and 'm' variances. Here, the covariance matrix will be a square symmetric matrix of order m×m. The formal definition can be given as,

$$S^{m \times m} = (S_{ij} : S_{ij} = \text{covariance (Dim}_i, \text{Dim}_j))\ldots\ldots\ldots\ldots(4.17)$$

Orthogonality & Orthonormality:

Two vectors are said to be *orthogonal* if their dot-product (or scalar product) is zero. This means that the vectors are mutually perpendicular.

The dot-product of two vectors $V = (w_1, w2, \ldots\ldots\ldots\ldots, wn)$ and $W = (w_1, w_2, \ldots\ldots\ldots\ldots, w_n)$ is defined as in equation 4.18.

$$V.W = v_1 w_1 + v_2 w_2 + \cdots\cdots\cdots + v_n w_n = \sum_{j=1}^{n} v_j w_j \quad\ldots\ldots\ldots\ldots\ldots\ldots(4.18)$$

For orthogonality,

$$V.W = 0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4.19)$$

A set of vectors $U_1, U_2, \ldots\ldots\ldots\ldots, Um$ (where each U_i is a vector with m components, i.e. $U_i = (u_{i1}, u_{i2}, \ldots\ldots\ldots\ldots, u_{in})$ is said to be *orthonormal* if they are pair-wise orthogonal to each other and all have a length equal to unity. Mathematically, this can be defined as,

$$U_i.U_j = 0 \text{ if } i \neq j$$
$$= 1 \text{ if } i = j \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4.20)$$

An orthonormal (or, orthogonal) matrix is defined as a matrix whose columns are orthonormal (or, orthogonal) to each other.

Corollary 1:

The inverse of an orthogonal matrix is its transpose [18].

Eigenvectors and Eigenvalues of a matrix:

Let H be an nxn matrix. The number λ is an eigenvalue of H if there exists a non-zero vector V (nx1) such that,

$$HV = \lambda V \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4.21)$$

In this case, vector V is called the eigenvector of H corresponding to the eigenvalue λ. Following is an example of an eigenvector:

$$\begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix} \times \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 12 \\ 8 \end{bmatrix} = 4 \times \begin{bmatrix} 3 \\ 2 \end{bmatrix} \ldots\ldots\ldots\ldots\ldots\ldots(4.22)$$

Here, $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ is an eigenvector of the matrix $\begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix}$ corresponding to the eigenvalue 4.

Eigenvectors can only be found for square matrices. However, not every square matrix does have eigenvectors. Given an nxn matrix that does have eigenvectors, there are n of them.

All eigenvectors of a matrix are orthogonal to each other, no matter how many eigenvectors we have. It means we can express our data in terms of these perpendicular eigenvectors taken as the axes (or basis vectors). If we restrict the length of all eigenvectors to unity, they construct a set of orthonormal vectors.

Corollary 2:

A square symmetric matrix is diagonalized by a matrix of its orthonormal eigenvectors [18]. This means an n×n symmetric matrix H can be expressed as in equation 4.23.

$$H = EDE^T \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(4.23)$$

Where, E is the matrix of the eigenvectors of H arranged as columns and E^T is the transpose of E, and D is a diagonal matrix.

4.2.3. MATHEMATICAL THEORY OF PCA

PCA is a well-established technique for dimensionality reduction without much loss of information. In other words, PCA can be thought of an analysis tool, which finds out the most meaningful basis to re-express a high-dimensional data set. This process automatically filters out undesired noise present in the data-set to certain extent, making the data more accessible for visualization and analysis. The new basis (computed by PCA) expresses the data without redundancy.

Let us go back to the example cited in section 4.2.1, where we have 'n' observations of 'm' variables. The data matrix X is defined as follows:

OBSERVATIONS ⟶

X_{11}	X_{12}	X_{13}	...	...	...	X_{1n}
X_{21}						X_{2n}
X_{31}						⋮
⋮						⋮
⋮						⋮
X_{m1}	...	...	...	...	...	X_{mn}

V
A
R
I
A
B
L
S
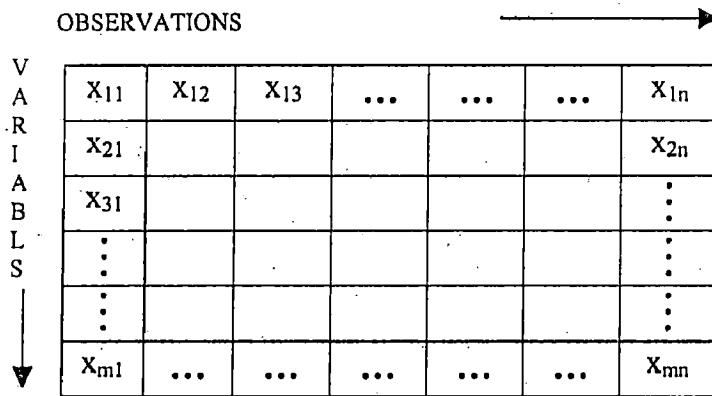
Figure 4.16 : PCA Input Data Matrix Structure

Here, any observation (along the column of X) refers to a vector having m components, as given in equation 4.24.

41

$$\overrightarrow{OBS(i)} = \begin{bmatrix} x_{1i} \\ x_{2i} \\ \vdots \\ x_{mi} \end{bmatrix} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(4.24)$$

And hence X can be re-written as:

$$X = [\ \overrightarrow{OBS(1)} \quad \overrightarrow{OBS(2)} \quad \dots\dots\dots\dots\dots \quad \overrightarrow{OBS(n)}\ ]\dots\dots\dots\dots\dots(4.25)$$

The basis of representation for the observations is an m-dimensional vector space.

$$B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = I$$

Initial basis (B): Identity Matrix (I) of order m×m.

$b_1,\dots\dots,b_m$: Initial basis vectors (each having m components)

Using PCA, we will try to find out another basis (of smaller dimension) that best re-expresses our data set [2]. The new basis vectors will be a linear combination of the original basis vectors $(b_1,\dots\dots,b_m)$.

0   In other words, we need to seek for a transformation matrix P (new basis matrix) of order r×m, which transforms an m-dimensional observation into an r-dimensional one (r≤m). This matrix P when multiplied to X will produce a new matrix Y (of order r×n), which is the re-representation of our data set.

$$PX = Y \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(4.26)$$

Equation (4.26) represents a change of basis equation. The rows of P $(p_1,\dots\dots,p_r)$ are the new basis vectors for expressing the columns of X. [Initially we had BX =X]. The rows of P are called the principal components corresponding to dataset X.

The physical interpretation of PCA can be given in terms of projection. Let us consider two-dimensional data (m=2), initially expressed in terms of conventional x and y co-ordinates. Following figures demonstrates the role of PCA.

(a): Initial Representation of 2-D data

(b): Principal directions of variation

(c) Projecting all data points on P1
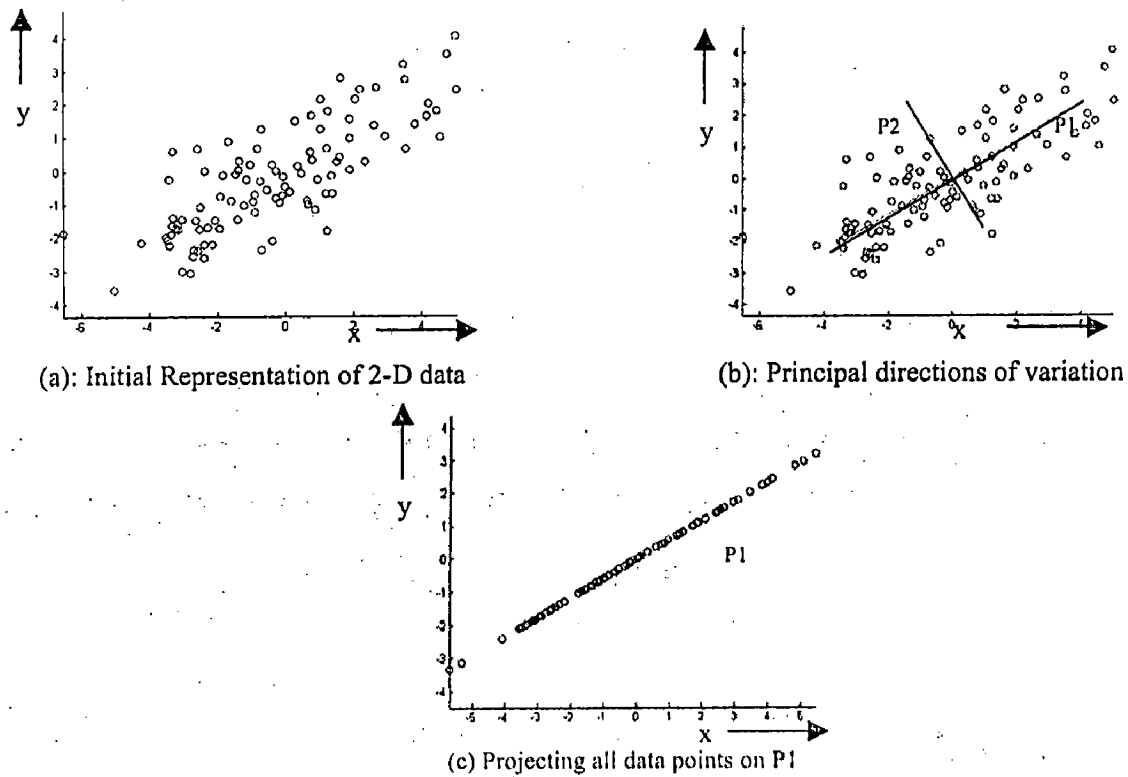
Figure 4.17: PCA Functioning

We can see that P1 in figure 4.17(b) represents the best fit line for the data. P2 represents deviation from this best fit line. The maximum variation of data occurs along P1. We can project all data points on P1 as shown in 4.17(c). Thus we come to a one-dimensional representation of the data (along a straight line). Therefore, the dimensionality is reduced (r = 1 here) while preserving the maximum variability in the data. The original basis vectors (b_1 and b_2) were nothing but the x and y axes. As apparent from 4.17(c), the new representation of data (along a straight line) can be given in terms of only one axis P1. So, P1 is the new basis vector, which is a linear combination of b_1 and b_2. The task of PCA is to find out this P1 from given b_1 and b_2.

Referring to equation (4.26), the new representation of the data (Y) should be free of noise and redundancy, as much as possible. The signal-to-noise ratio (which is a ratio of the variances of the signal and noise) can be improved by maximizing the variance along the new basis vectors. On the other hand, the redundancy of data can be measured in terms of the covariance between different variables (or dimensions). We want each variable to co-vary as little as possible with other variables. So, we need to minimize the covariance between the new set of variables (rows of Y) at the same time.

43

To quantify covariance and variance in a multi-dimensional data set, we go for the covariance matrix as discussed before (Equation 4.17). In our example, the covariance matrix for the data sets (X or Y, where each row represents a different dimension or variable) can be defined as follows (here we assume that each row of X is a zero-mean variable):

$$S_X = \frac{1}{n-1}XX^T \quad ; \quad S_Y = \frac{1}{n-1}YY^T \quad \dots\dots\dots\dots\dots\dots(4.27)$$

Computing S_X describes all relationships between pairs of measurements in our data set X. We need to determine a new basis matrix (P) which will maximize the variance inside any particular dimension and at the same time minimize the covariance between different dimensions. In other words, we need to have an optimum choice of P in order to have a diagonal covariance matrix for Y.

PCA assumes that all basis vectors $\{p_1,\dots\dots p_r$: rows of P$\}$ are orthonormal. In the language of linear algebra, PCA assumes P^T (transpose of P) to be an orthonormal matrix (columns of P^T being orthonormal to each other). Secondly, PCA assumes the directions with the largest variances are the most important ones. PCA first selects a normalized direction in m-dimensional space along which the variance of X is maximized- it saves this as p_1. Again it finds another direction along which variance is maximized. However, because of the orthonormality condition, it restricts its search to all directions perpendicular to all previous selected directions [18]. This search can theoretically continue up to m directions. However, we select only the first r directions accounting for more than say 99.9% of variance of the original data set. Next we shall see the solution for PCA using linear algebra.

First we write S_Y in terms of P as below,

$$S_Y = \frac{1}{n-1}YY^T = \frac{1}{n-1}(PX)(PX)^T = \frac{1}{n-1}PXX^TP^T = \frac{1}{n-1}P(XX^T)P^T \dots\dots\dots(4.28)$$

Let $XX^T = G$,

Then we have, $\qquad\qquad S_Y = \frac{1}{n-1}PGP^T \dots\dots\dots\dots\dots\dots\dots (4.29)$

Here, matrix G will be symmetrical by definition, as it is a product of one matrix and its transpose. The size of G would be m×m.

From 'Corollary 2' in section 4.2.2, we can write,

$$G = EDE^T \qquad\qquad\dots\dots\dots\dots\dots(4.30)$$

44

Where, D is a diagonal matrix and E is a matrix of eigenvectors of G arranged as columns.

We now select the matrix P to be a matrix where each row p_i of P is an eigenvector of XX^T. Therefore, $P = E^T$.

Substituting $P = E^T$ in Equation (4.30), we get,

$$G = P^T DP \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(4.31)$$

From 'Corollary 1' in section 4.2.2, we can see that,

$$(P^T)^{-1} = (P^T)^T$$

$$\text{or, } (P^T)^{-1} = P$$

$$\text{or, } P^T = P^{-1} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(4.32)$$

Combining equations (4.29) and (4.31), we get,

$$S_Y = \frac{1}{n-1} PGP^T = \frac{1}{n-1} P(P^T DP)P^T = \frac{1}{n-1}(PP^T)D(PP^T) = \frac{1}{n-1}(PP^{-1})D(PP^{-1})$$

Finally we arrive at,

$$S_Y = \frac{1}{n-1}D \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(4.33)$$

Therefore, we can conclude that our choice of P diagonalizes S_Y, i.e. the goal is achieved. Hence the basis vectors for the new representation or the principal components are nothing but the eigenvectors of the covariance matrix of X.

4.2.4. STEPS OF PCA IMPLEMENTATION

Step 1: Construct the data matrix X (as depicted in figure 4.16)

Step 2: Calculate the mean of each row and subtract it from the corresponding row.

Step 3: Calculate the covariance matrix S_X(mxm) of these zero-mean variables.

Step 4: Calculate the eigenvectors and eigenvalues of S_X.

Step 5: Arrange the eigenvectors in descending order of eigenvalues. Now, construct a matrix (P) with the first r eigenvectors (after being sorted) as rows. That means, the eigenvector with the highest eigenvalue will be in the first row , eigenvector with the 2nd highest eigenvalue in the 2nd row and so on.

Step 6: Now obtain the new representation of data (Y) by multiplying P with X, i.e.

Y=PX.

4.3 CORRELATION TECHNIQUES

4.3.1. AUTO AND CROSS CORRELATION FUNCTIONS

Correlation is a statistical technique, which can show whether and how strongly pairs of variables (or waveforms) are related.

A good method of measuring the similarity between two waveforms is to multiply them together, ordinate by ordinate, and to add the products over the duration of the waveforms. Let us consider the following figure.
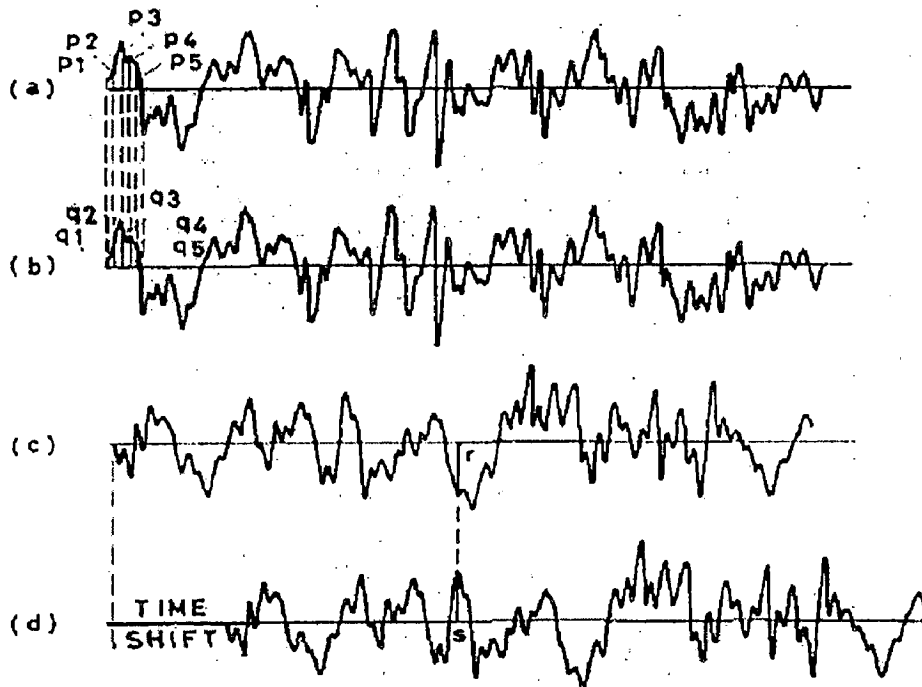
Figure 4.18: Correlation Calculation between different pairs of Waveforms [23]

In the above figure, the waveforms (a) and (b) are exactly identical. Waveforms (c) and (d) are identical in shape but having a time shift between them.

To assess the similarity between (a) and (b) in mathematical terms, we multiply ordinate p_1 by ordinate q_1, ordinate p_2 by q_2, p_3 by q_3, and so on, and finally we add all these products to obtain a single number which is a measure of the similarity. This number represents the correlation. Here, waveforms (a) and (b) are identical, so that, every ordinate, positive or negative, contributes to a positive term to the sum. Thus, we have a large value of the sum (i.e. the correlation coefficients).

If we perform the same process on waveforms (b) and (c), we find that each positive product is offset by another negative product and hence their sum becomes small. Lesser correlation is obtained between waveforms with lesser similarity.

Now we consider the correlation between waveforms (c) and (d). They are identical in shape, but one is displaced in time from the other. If we perform the same process of multiplying ordinates (of which r and s are typical), again we find that every positive product tends to be offset by a negative product, resulting in a smaller sum. Thus, if we were to plot the similarity between a waveform of figure 4.18(c) and a time-shifted version of itself, we should expect the resulting curve to assume small values for large time shifts, and to rise to a large positive maximum when the time shift is zero. This curve is called the auto-correlation function (ACF). Formal definition of ACF is given below.

The *Auto-Correlation Function* of a waveform is a graph of the similarity between the waveform and a time-shifted version of itself, as a function of this time-shift [23].

Let us consider the ACF of a sine function. The sine wave becomes exactly similar to itself whenever the time shift is an integral multiple of its time period. Hence, the ACF of it must be periodic. In reality, the ACF of a sine wave is also sinusoidal, having the same frequency, and being symmetrical about the point which represents zero time-shift, as evident from the following figure. The horizontal axis in the ACF plot represents time shift and not the absolute time.

(a) Sine Wave

(b) ACF
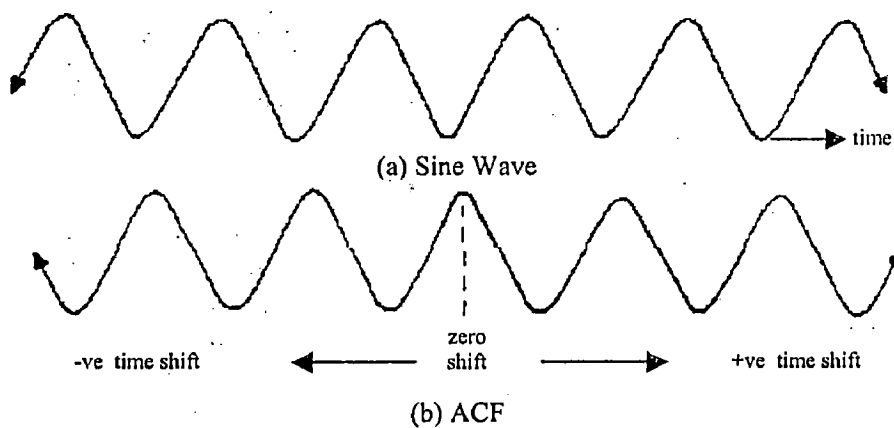
Figure 4.19: Auto-Correlation Function of a Sine Wave

Now, we consider the ACF of a broad-band noise signal in the following figure 4.20.

47

(a) Broad-Band Noise ———▶ time

-ve time shift ◀——— zero shift ———▶ +ve time shift

(b) ACF
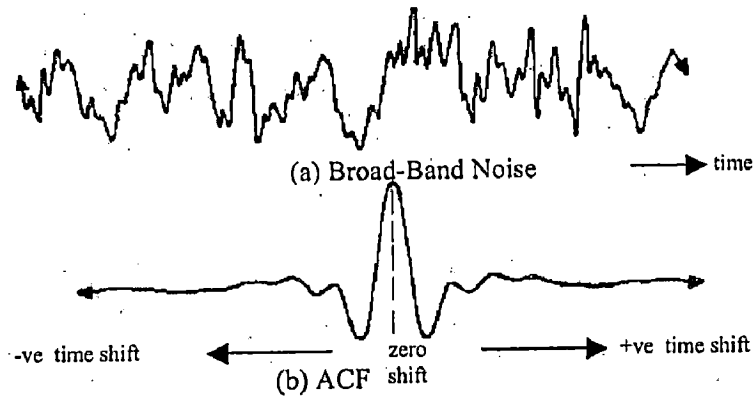
Figure 4.20: Auto-Correlation Function of a Broad-Band Noise Signal [23]

In figure 4.20, the ACF is a sharp impulse, decaying quickly from the central maximum to very low values at large time shifts. It is because the waveform in figure 4.20(a) is similar to itself only at zero time shift. A very small time shift is sufficient to destroy the similarity, and the similarity never recurs.

So far we have discussed about the ACF, which is concerned with the similarity between a waveform and a time-shifted version of itself. However, the same idea can be extended for two different non-identical waveforms, like the signals (b) and (c) in figure 4.18. This is the concept behind cross-correlation. Let us consider the figure below for a better insight into it.

Time-origin of stationary waveform

(a)

Time-origin of moving waveform

(b)

Shift between waveforms

Window of width T

(c) CCF between (a) and (b)

Zero
-ve time shift ◀——— Time-Shift ———▶ +ve time shift
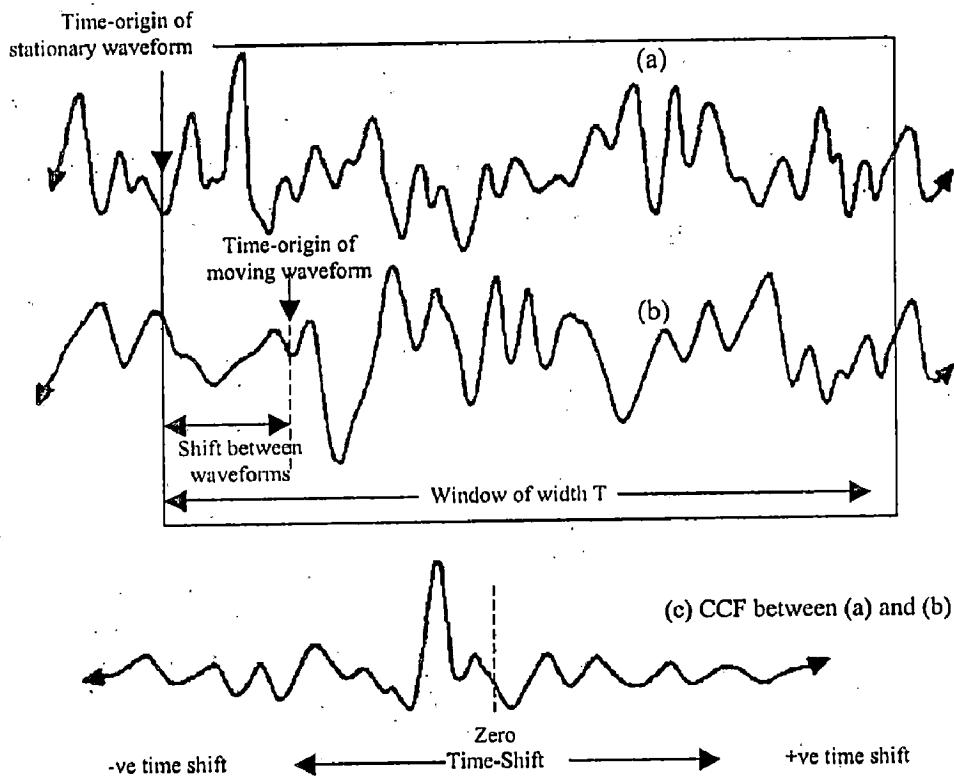
Figure 4.21: Cross-Correlation Function [23]

48

In figure 4.21, we visualize the waveform (a) as stationary and waveform (b) as sliding. We then view the two waveforms through a window of width T, and we assess the similarity between them within this time interval by our previous method of multiplying ordinates and summing products. For any particular shift of (b), this number represents the cross-correlation between (a) and (b) at that shift. The graph (c) represents the similarity between (a) and (b) as a function of the time-shift between them. (c) is the Cross-Correlation Function (CCF), whose formal definition is given below.

The *Cross-Correlation Function* of two waveforms is a graph of the similarity between the two waveforms as a function of the time shift between them [23].

In our application, we will be comparing two different signals through cross-correlation. For that, we shall deal with the cross-correlation coefficient only and not the CCF in particular. The following section provides a mathematical definition of the correlation coefficients.

4.3.2. CORRELATION COEFFICIENTS

Two signals may be compared to detect common characteristics present in them through cross correlation. The cross correlation between two digital signals x(n) and y(n) having N samples each ('n' representing sample number) is defined as their inner or scalar product as follows:

$$x.y = \sum_{n=1}^{N} x(n)y(n) \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(4.34)$$

The dot product represents the projection of one signal onto the other, with each signal being viewed as an N-dimensional vector. The dot product can be normalized by the geometric mean of the energies of the two signals [4]:

$$\gamma_{xy} = \frac{\sum_{n=1}^{N} x(n)y(n)}{\left[\sum_{n=1}^{N} x^2(n)\sum_{n=1}^{N} y^2(n)\right]^{1/2}} \dots\dots\dots\dots\dots\dots\dots\dots(4.35)$$

γ_{xy} is called the *Cross-Correlation Coefficient* between the two signals x(n) and y(n). It can have values between 0 to 1 (or 0 to 100%). Higher values of γ_{xy} signifies greater similarity between the two signals and vice versa.

The *Auto-Correlation Coefficient* ($a_{x,k}$) can be calculated in a similar manner by replacing y(n) with x(n+k) in equation (4.35). x(n+k) is nothing but the signal x(n) shifted in time by k samples. This coefficient represents mathematically the similarity between the signal x(n) with itself when shifted by k samples. The expression is as follows,

$$a_{x,k} = \frac{\sum_{n=1}^{N} x(n)x(n+k)}{\left[\sum_{n=1}^{N} x^2(n)\sum_{n=1}^{N} x^2(n+k)\right]^{1/2}} \dots\dots\dots\dots\dots\dots\dots\dots\dots(4.36)$$

$a_{x,k}$ can have values between 0 to 1 (or 0 to 100%).

50

CHAPTER 5

ECG SIGNAL CONDITIONING
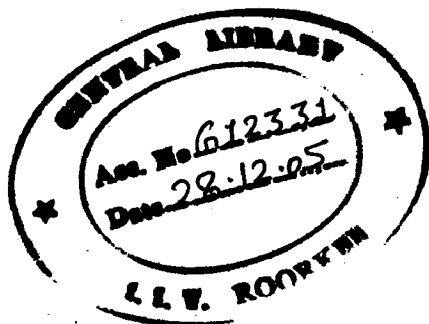
The recorded ECG signal needs to be processed or conditioned in order to remove the noise and artifacts, before delineation. ECG noise removal has always been a subject of great importance. The purpose of the de-noising filtering process is to reduce the noise level in the signal and simultaneously prevent waveform distortion. This last characteristic is of vital importance to prevent wrong diagnosis or analysis of the ECG signal [24].

Filtering of ECG is an essential pre-processing stage before delineation. Being highly sensitive to the ECG waveforms and morphologies, the delineation algorithm needs the input ECG to be free from noise as much as possible. In case of multi-channel ECG, first the reference is changed to Wilson's Centre Terminal (WCT) before any analysis. This ensures high *Common Mode Rejection Ratio (CMRR)*. Filtering is carried out in two stages, as follows:

First Stage→ Cancellation of baseline wanders

Second Stage→ Elimination of high frequency noise and artefact

5.1. CANCELLATION OF BASELINE WANDER

Normally, the baseline of ECG signal (i.e. the rest of the ECG except the peaks and waves) is supposed to be iso-electric at zero amplitude level. This is shown in the following figure.
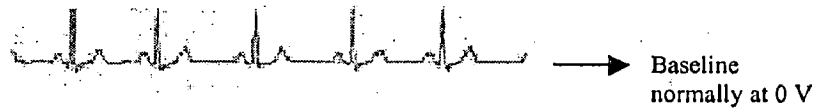
Figure 5.1: Normal ECG free of Baseline Drift
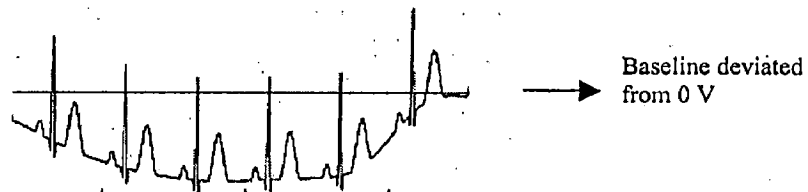
Figure 5.2 : ECG corrupted with Baseline Wander

Followings can be the sources of baseline wander [4]:

1. Coughing or breathing with large chest movement for chest-lead ECGs.

2. If an arm or leg is moved in case of limb-lead ECG acquisition.

3. Electrogastrogram (electrical activity of stomach) for chest-lead ECG.

4. Poor contact and polarization of electrodes.

5. Variations in temperature and bias in the instrumentation and amplifiers.

The frequency components of the baseline wander are usually below 0.5 Hz.

5.1.1. RELATED LITERATURE SURVEY

Several methods have been proposed in the literature to eliminate baseline wander. The first is ensemble averaging. However, this approach is not a realistic one as the ECG signal exhibits beat-to-beat variations. Among other methods, we have polynomial interpolation. Linear interpolation introduces significant distortions. A third order approximation called cubic spline [25] is proved to give better results. Interpolation techniques make use of a previous knowledge of the ECG isoelectric levels estimated from the PR intervals (knots). Therefore, the performance of this technique depends highly on the knots determination accuracy and gets degraded as the knots become more separated in time (low heart rate).

To overcome the above problem, another group proposed digital narrow-band linear-phase filtering [26]. This method can be implemented in real time, but has two major draw-backs. First, the filter needs to be a FIR filter with a long impulse response, which means a large number of coefficients. Secondly, given that ECG and baseline wander spectra usually overlap, it is not possible to remove baseline wander without distorting ECG. Another technique has been proposed in [27], which uses a time-varying linear filter that selects different cut-off frequencies as a function of the heart rate or the baseline level. This filter improves the time invariant FIR filter performance, but can yet distort the ST components of ECG and has high computational requirements.

Another group employed Short Time Fourier Transform (STFT) to get rid of baseline drift [29]. Within every window, they search for a spectral component in the range 0.0 to 1.0 Hz. Only the ECG segments containing frequency components in the specified range are high-pass filtered to cancel baseline wander. However, it is not possible to have optimal frequency and time resolution at the same time with STFT.

Adaptive filtering has also been proposed to cancel the baseline drift [28]. An adaptive transversal filter with only one weight is used, where the reference input is a constant with a value of 1 and the primary input is the ECG signal. This filter, using the Least Mean Square (LMS) Algorithm in the adaptation process, is equivalent to a linear notch filter that takes the advantage of adaptive implementation, but still modifies the ST segment.

In [30], a cascade adaptive filter has been used. The first stage of the filter is exactly similar to that mentioned above. In the second stage, the primary input is the output from the first stage and the reference input is a unit impulse sequence correlated with each QRS complex. This needs a QRS detector to generate the impulse sequence. However, in our application baseline wander cancellation is accomplished as a pre-conditioning of ECG signal before delineation.

Here, a DWT-based baseline wander removal algorithm has been employed, which is simple in implementation, yet providing promising results.

5.1.2. Motivation

The frequency range of baseline wander is below 0.5Hz. Therefore, its presence will be reflected in the higher level DWT approximation coefficients. This is the basic idea behind our approach.

First the noisy ECG signal (contaminated with baseline wander) is decomposed into a certain no. of levels (n) using Discrete Wavelet Transform (as discussed in the previous chapter, section 4.1.7). The highest level (nth level) approximation coefficients (A.C.) are supposed to represent the low frequency baseline variation signal. The nth level A.C. are made to be all zeros and then the ECG signal is reconstructed following the same procedure as mentioned in section 4.1.7.

A mother wavelet (e.g. coif4) is chosen arbitrarily and DWT decomposition is carried out on one ECG signal. It was seen that each of the 8th, 9th and 10th level approximation coefficients, when time-aligned to the original ECG, resemble the baseline wander. Following figure shows this resemblance.
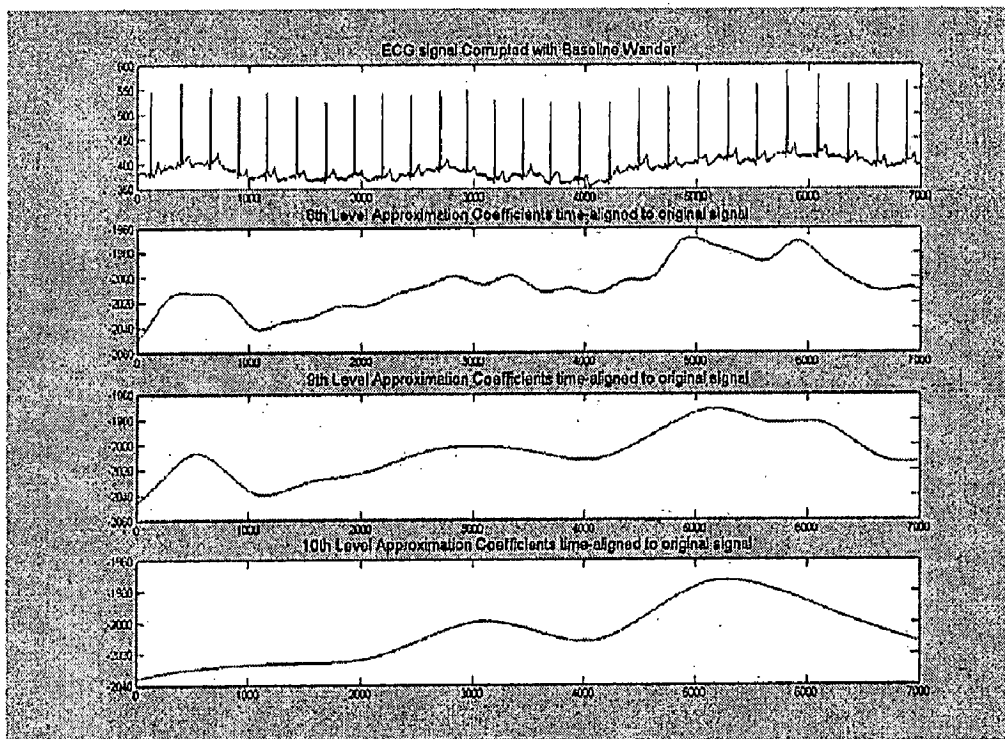

Figure 5.3: Resemblance of Baseline Wander with High Level Approximations

However, a clear idea regarding the following two things is needed.

1. Which mother wavelet should be applied for DWT analysis on the noisy ECG for best results?

2. What value of n should be chosen?

5.1.3. SIMULATION & RESULTS

Before dealing with real ECGs, artificial signals (i.e. mixture of artificial ECG and artificial baseline variation signal) were taken for experimentation. Thus, with a clear knowledge of the component signals, the performance of filtering could be judged.

Artificial ECG beats were generated in Matlab 6.5 with the help of inbuilt ECG simulator. The sampling frequency was assumed as 1 KHz, or in other words the span of 1000 samples is 1 second. The data length was taken to be 25000 samples, i.e. 25 second. Then, for 60 bpm (beats per minute) ECG, each beat should have a span of 1000 samples and hence there would be 25 beats in total. For different bpm, the span of ECG beat was varied accordingly and hence the total number of ECG beats.

To find out the suitable mother wavelet and the decomposition level (n), test was carried out on 650 artificially generated noisy ECGs. Thirteen noise-free ECG signals in the range of 60 to 180 bpm were created as discussed above. At the same time, a set of fifty sinusoidal signals with frequencies ranging from 0.01-0.5 Hz was also created, to simulate the baseline wander. Thereafter, 650 test signals in total were synthesized by mixing the artificial ECGs with artificial baseline wander signals in one to one correspondence.

Now, on each of the 650 test signals (or mixture signals), DWT analysis was carried out taking a total of 29 mother wavelets (sym1,.......,sym12, coif1,...,coif5,db1,.........,db12) under consideration.

On the same test signal, for each of the mother wavelets, the following procedure was adopted.

1. Initialise n=1; i.e. no. of decomposition levels for DWT.

2. Decompose the test signal till n levels (maximal n was 12) and get the DWT coefficients A_n, D_n, D_{n-1}, D_{n-2}, ...D_1 (referring to figure 4.9).

   A_i : ith level approximation coefficient

   D_i: ith level details coefficient

3. Perform two reconstructions.

   First Reconstruction : With A_n to be all zeros, reconstruct the signal as in Figure 4.11. The signal, reconstructed in this way is called the 'ECG reconstruction'. It should resemble the original noise-free ECG (with which the test signal is synthesized) for higher values of n.

<u>Second Reconstruction</u>: Making all the coefficients other than A_n (i.e. the details coefficients, D_n.................D_1) as zeros, reconstruct the signal as depicted in Figure 4.11. The signal, reconstructed in this way is called the 'Baseline Reconstruction'. It should resemble the original Baseline variation signal (with which the test signal is synthesized) for higher values of n.

4. Judge the resemblance between the original and reconstructed signals by means of cross-correlation. Two cross-correlation coefficients (CE & CB) were calculated.

CE= Cross-correlation between original noise free ECG (generated by in-build matlab command) and the 'ECG reconstruction'.

CB= Cross-correlation between original baseline variation signal (low frequency sinusoid) and the 'Baseline reconstruction'.

5. Repeat steps 2 to 4 for n=1,........12.

Now the whole process (from step 1 to 5) was repeated for 29 different mother wavelets applied on the same test signal.

For each test signal, two correlation matrices (one for CE and the other for CB, each of order 12×29) were constructed. The structure of each is as follows:

Mother Wavelets ─────────────────────────▶

$$
\begin{array}{c}
n \\
\downarrow
\end{array}
\quad
\begin{array}{ccc}
C(1,1). & \cdots & C(1,29) \\
\vdots & & \vdots \\
C(12,1). & \cdots & C(12,29)
\end{array}
$$
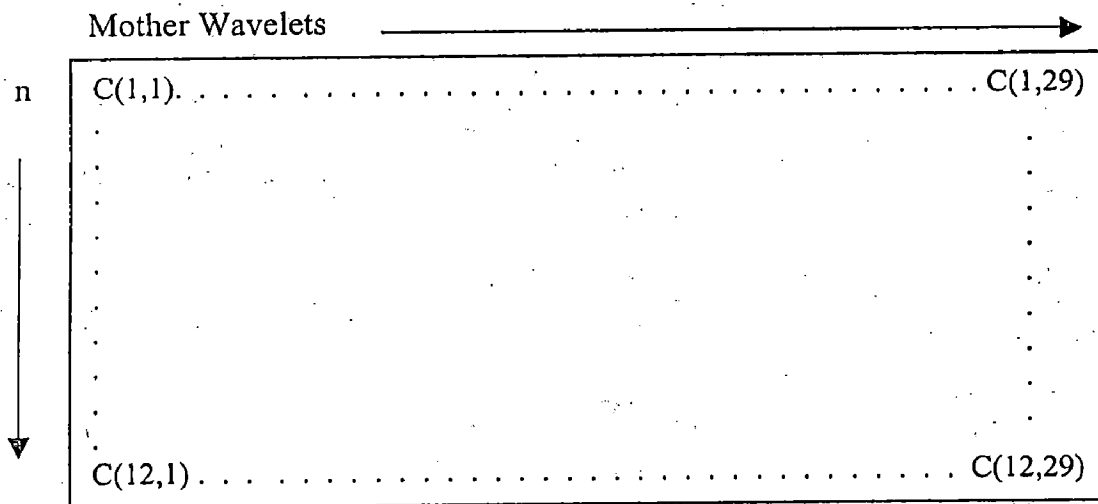
Figure 5.4: Correlation Matrix Structure

Where, C denotes either CE or CB. The 1st to 12th column represents Symlet Wavelets of 1st to 12th order (sym1,................,sym12) , the 13th to 17th column represents Coiflet Wavelets of 1st to 5th order (coif1,...........,coif5) and the remaining 18th to 12th column represents Daubechies Wavelets of 1st to 12th order (db1,................,db12).

For 650 test signals, there were 1300 correlation matrices in total, half for CEs and half for CBs. Now two more matrices were computed, the first one (CEmean) being the mean of 650 CE matrices and the second one (CBmean) being the mean of 650 CB matrices.

Now the positions of the first five greatest elements in both the matrices, CBmean and Cemean, were located. The positions are found to be exactly the same in both. This means that the highest element occurs in the same position (same row and column number) in both the matrices ---- same is true for 2nd highest and so on. All of these five highest elements are found at the row corresponding to n=9. This means that generally baseline wander signal can be located at the 9th level approximation coefficients.

The result is summarized in Table 5.1.

Order No.	Mother Wavelet	n	Mean CE (%)	Mean CB (%)
1	db11	9	99.9924	99.915
2	sym12	9	99.9913	99.9011
3	sym10	9	99.9909	99.8962
4	db10	9	99.9906	99.8925
5	coif5	9	99.9904	99.8894

Table 5.1: Results Obtained from 650 Artificial Test Signals

5.1.4. PROPOSED METHOD

The simulated ECG and baseline variation signals cover the range of practically encountered signals. Therefore, either of these five mother wavelets (listed in table 5.1) is applied on our recorded multi-channel ECG and the 9th level approximation coefficients are cancelled out to get rid of the baseline wander and at the same time to ensure least distortion in the ECG waveform. Following is the simplified strategy for baseline wander cancellation:

1. Carry out DWT decomposition of the noisy ECG at hand till the 9th level (as in Figure 4.9).

2. Make the 9th level approximation coefficients to be all zeros.

3. Now, reconstruct the ECG (as in Figure 4.11).

5.1.5. RESULT OF APPLICATION ON REAL ECG

Our algorithm was applied on Multi-Channel ECG data recorded at the Institute, as well as on signals taken from MIT- Arrhythmia Database. Following Figures demonstrate the success of our method.
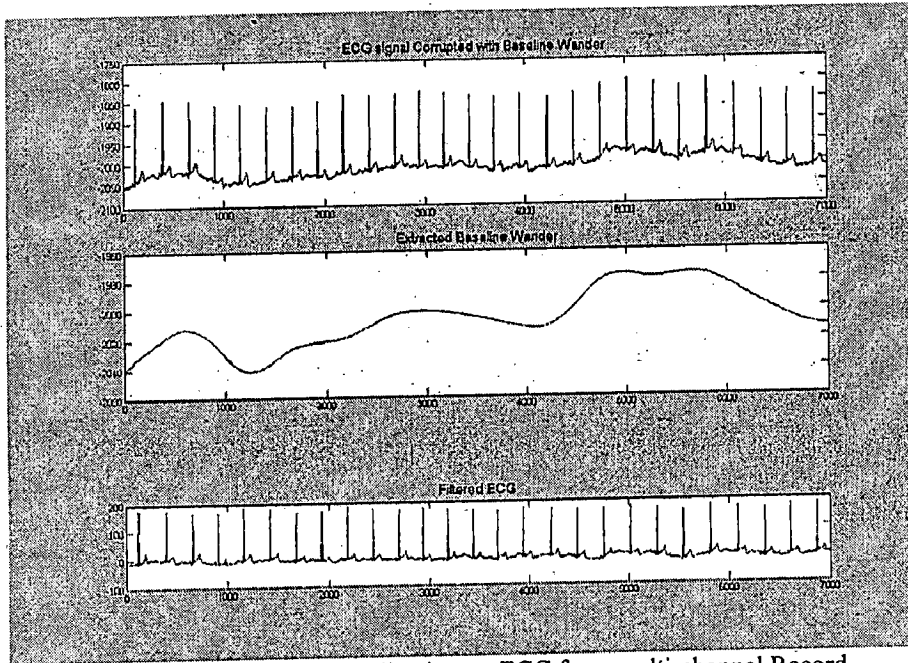
Figure 5.5: Baseline Wander Cancellation: Application on ECG from multi-channel Record
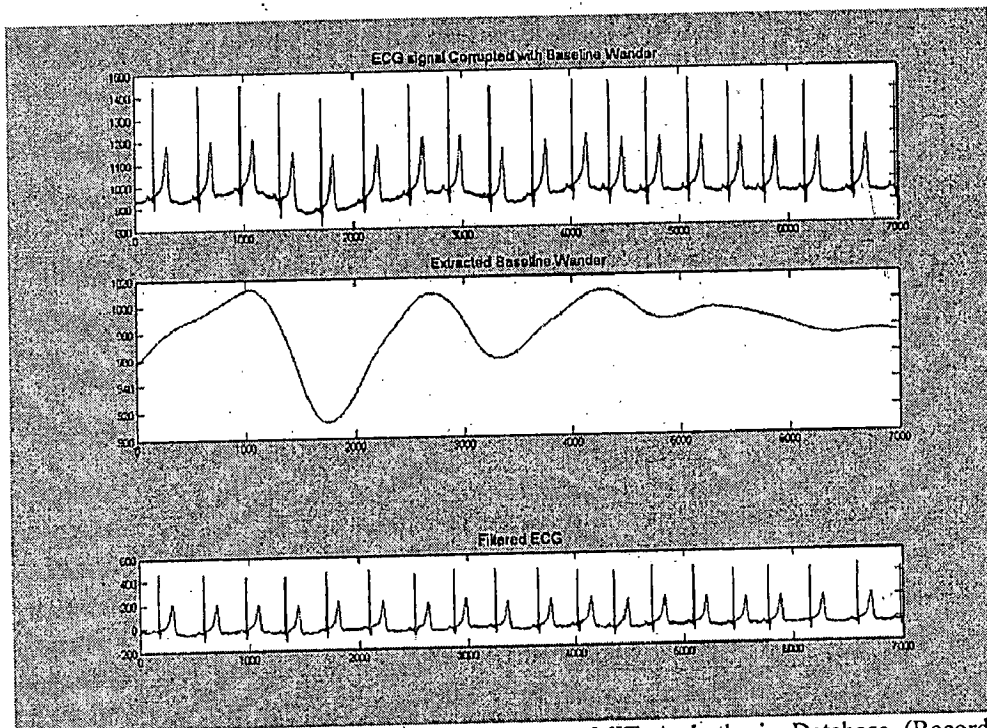
Figure 5.6: Baseline Wander Cancellation: Application on MIT Arrhythmia Database (Record no. 113, Channel 1)

After eliminating the low frequency baseline wander, the next stage is filtering of high frequency disturbance in ECG.

58

5.2. ELIMINATION OF HIGH FREQUENCY NOISE & ARTEFACT

The ECG is often corrupted with high frequency noise. This can be introduced due to instrumentation amplifiers, recording system, pick up of ambient electromagnetic signals by the cables, and so on. There may be random noises uncorrelated with the ECG signal like myoelectric, thermal etc, which can be approximated by a white noise source [24]. White noise is technically defined as a signal with uniform frequency spectrum over a wide range of frequencies (analogy with light). The harmonics of power-line interference may also be considered as a part of high-frequency noise relative to the low frequency nature of the ECG signal.
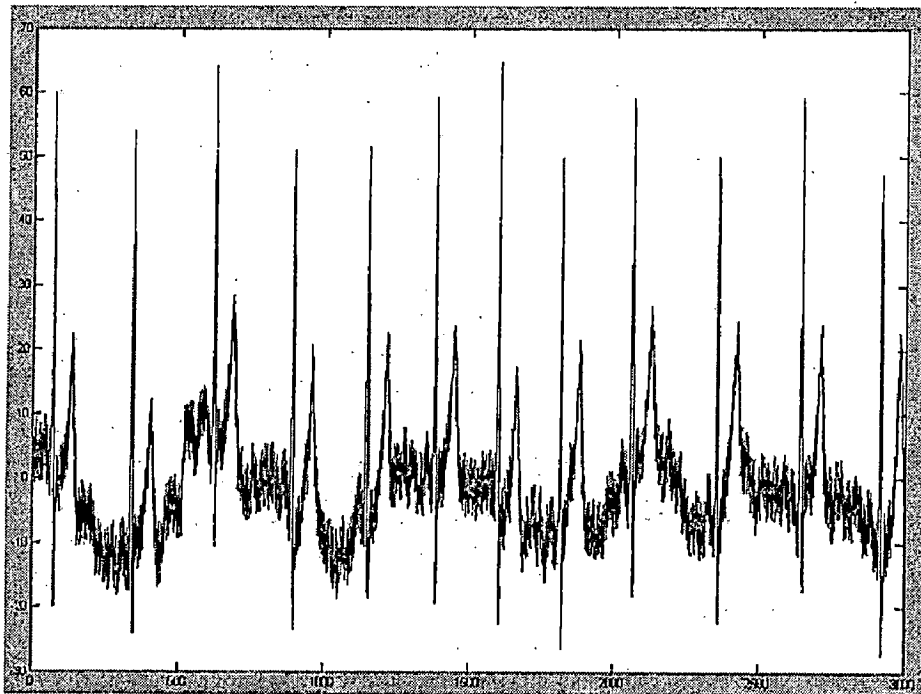
Figure 5.7: ECG Signal Corrupted with High Frequency Noise

Any of the two different approaches can be adopted to eliminate these kinds of noise, namely

1. Frequency-domain Filtering

2. Time-domain Filtering

5.2.1. FREQUENCY DOMAIN FILTERING

A filter is generally a frequency-selective device. It passes some frequency components of the signal as it is, whereas it blocks others by means of attenuation. The frequencies that are passed through the filter constitute the *pass-band*, and those that are blocked are in the *stop-*

band. Let us denote the frequency response of the filter by $H(j\omega)$ [where 'ω' is the angular frequency in radians/sec]. For frequencies in the pass-band, the magnitude $|H(j\omega)|$ is relatively high and ideally a constant. A stop-band is characterized by a small (ideally zero) $|H(j\omega)|$.

Here, we need a low pass (LP) filter to remove high frequency disturbance. At the same time, minimal loss of signal components in the pass-band needs to be ensured. Let us compare the magnitude response of four different kinds of low-pass filters.
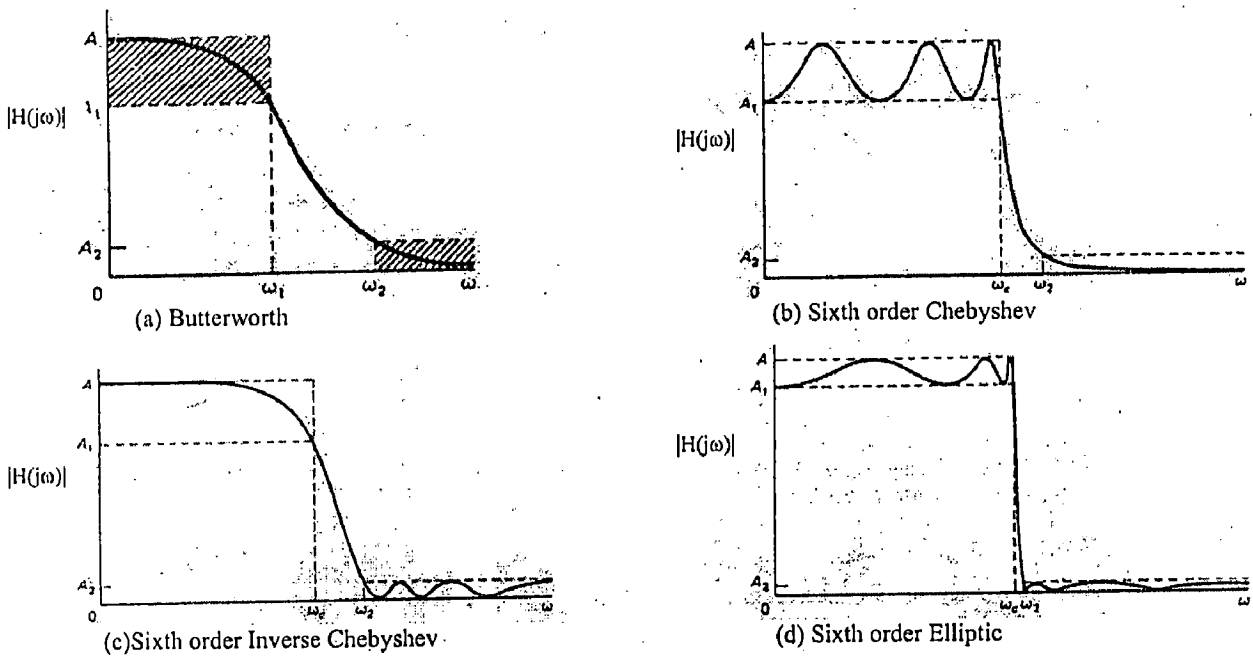
Figure 5.8: Magnitude Response of Four Different LP Filters [31]

The pass-band is characterized by $A_1 \leq |H(j\omega)| \leq A$ and the stop-band by $0 \leq |H(j\omega)| \leq A_2$. A_1 is never less than $A/2^{1/2}$. The frequency band $\omega_1 < \omega < \omega_2$, between the pass-band and stop-band is called the *transition band*. The response decreases continually within this band.

The cut-off frequency ω_c is taken same as the passband frequency ω_1 for Chebyshev and elliptic filters. For Butterworth and Inverse Chebyshev, ω_c is the frequency at which the gain falls by 3dB, i.e. where $|H(j\omega)| = A/2^{1/2}$ [31].

It is seen that although Butterworth filter does not possess a sharp transition, it offers the maximum flatness in the pass and stop bands, compared to the others. This maximally flat magnitude response makes it the most commonly used frequency domain filter.

The basic Butterworth low-pass filter function is given as,

$$| H_a(j\omega) | = \frac{1}{\left(1 + \left(\frac{j\omega}{j\omega_r} \right)^{2N} \right)^{1/2}} \qquad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(5.1)$$

Here, N denotes the order of the filter and suffix 'a' in $H_a(j\omega)$ stands for the frequency response of the analog filter. A Butterworth filter is completely specified by its order and cut-off frequency.

Following figure shows the magnitude response for different values of N.
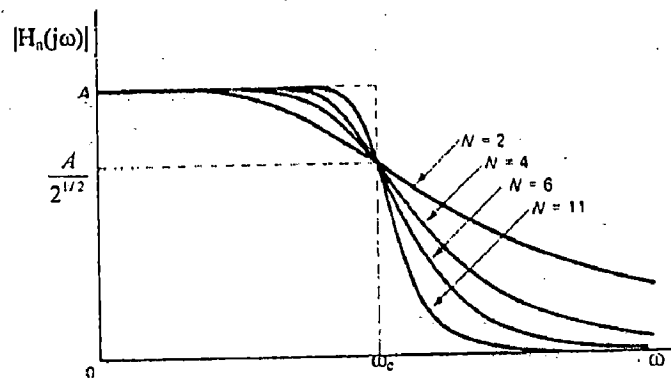
Figure 5.9: Butterworth LP Magnitude Response for Different Filter Orders [31]

The ideal low-pass filter response is shown in dotted line. The filter response is monotonic in the pass-band as well as in the stop-band. It is apparent that the Butterworth response more closely approximates the ideal response (i.e. more flatness in pass-band along with faster and sharper transition to stop-band) as the order N increases. The Butterworth filter is said to have *maximally flat magnitude response* (compared to other filters of same order) since the first (2N-1) derivatives of $| H_a(j\omega) |^2$ are zero at $\omega=0$.

However, the phase response of the Butterworth filter becomes more non-linear with increase in order as shown below.

(a) 2nd order

(b) 3rd order

(c) 5th order　　　　　　　　　　　　　　　　　(d) 6th order
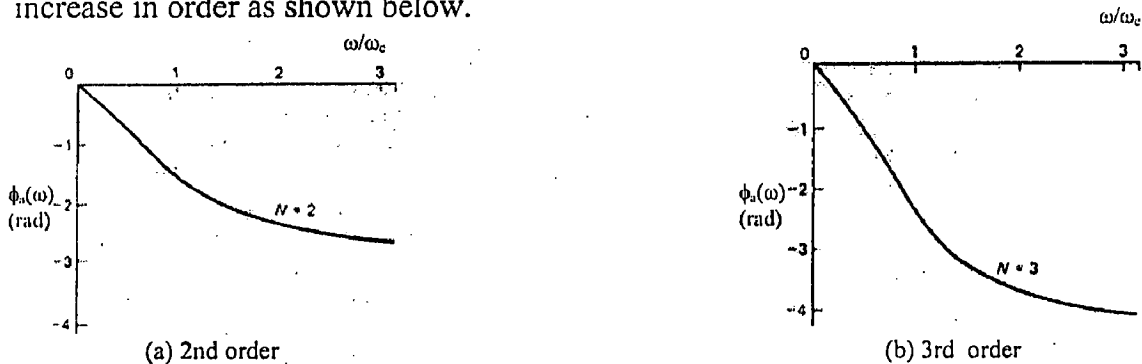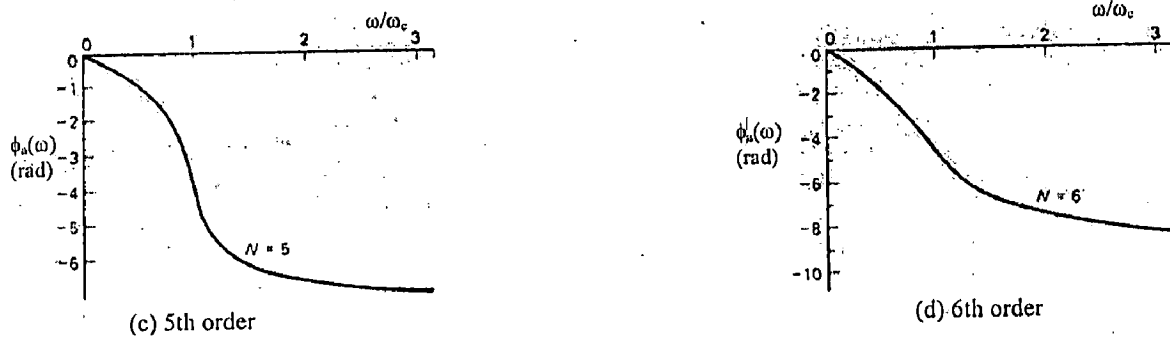
Figure 5.10: Butterworth LP Phase Response for Different Filter Orders [31]

But linear phase characteristic is essential to ensure minimum distortion of the original signal after filtering. In our application, second order Butterworth low pass filter with cut-off frequency 70 Hz seems to be the best.

However, we use the Butterworth filter in the discrete form. Let us see how the discrete transfer function H(z) can be derived from equation (5.1).

Changing equation (5.1) in terms of Laplace variable 's', we get

$$H_a(s)H_a(-s) = \frac{1}{\left(1+\left(\dfrac{s}{j\omega_c}\right)^{2N}\right)} \qquad\qquad\qquad (5.2)$$

The poles of the squared transfer function [in equation (5.2)] are located with equal spacing around a circle of radius ω_c in the s-plane, distributed symmetrically on either side of the imaginary axis s = jω.
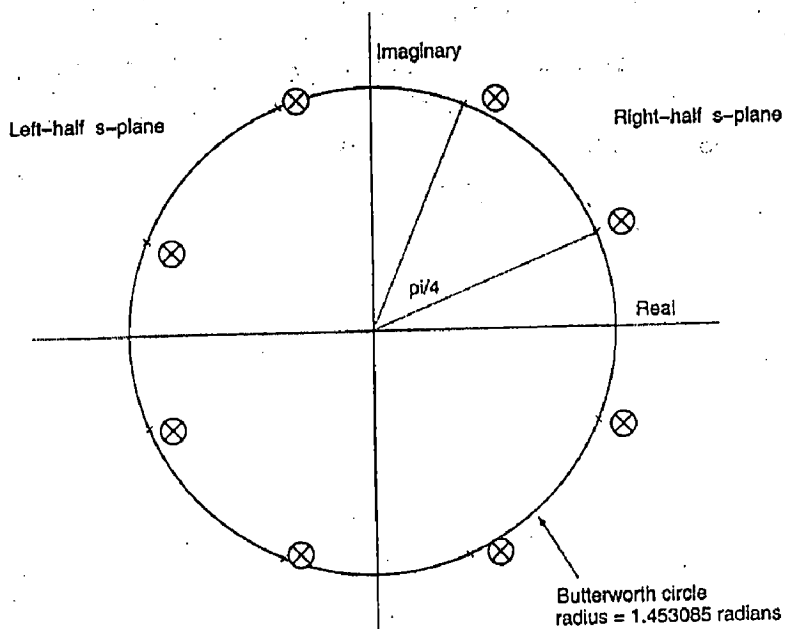
Figure 5.11: Pole Locations of $H_a(s)H_a(-s)$ [4]

62

Following observations can be made regarding the pole locations:

1. No pole will lie on the imaginary axis itself (i.e. $s = j\omega$).

2. Poles will appear on real axis for odd N.

3. The angular spacing between the poles is π/N.

4. If $H_a(s)H_a(-s)$ has a pole at s= s_p, it will have a pole at s = -s_p as well.

5. For the filter coefficients to be real, complex poles must appear in conjugate pairs.

6. The pole positions in the s-plane are given by,

$$S_i = \omega_c e^{j\pi\left(\frac{1}{2}+\frac{(2i-1)}{2N}\right)} \dots\dots\dots\dots\dots\dots\dots\dots\dots(5.3)$$

Where, i =1,2,............2N

In order to obtain a stable and causal filter, we need to form $H_a(s)$ with only the N poles on the left-hand side of the s-plane. Selecting these poles the transfer function in the analog Laplace domain becomes as given in equation 5.4.

$$H_a(s) = \frac{G}{(s-p_1)(s-p_2)(s-p_3)\dots\dots\dots(s-p_N)} \dots\dots\dots\dots\dots(5.4)$$

where, p_k, k=1,2,.......N, are the N poles of the transfer function in the left-half of the s-plane, and G is a gain factor specified as needed to normalize the gain at DC (s = 0) to be unity.

The transfer function $H_a(s)$ may be mapped to the z-domain by applying the bilinear transformation as follows:

$$s = \frac{2}{T}\left(\frac{1-z^{-1}}{1+z^{-1}}\right) \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(5.5)$$

Here, T denotes the sampling interval. If the sampling frequency of ECG recording be 1KHz, T=1 ms. The transfer function H(z) obtained after bilinear transformation can be simplified to the form,

$$H(z) = \frac{G'(1+z^{-1})^N}{\sum_{k=0}^{N}a_k z^{-k}} \dots\dots\dots\dots\dots\dots\dots\dots\dots(5.6)$$

where a_k, k=0,1,2, ...,N are the filter coefficients or tap weights (with a_0=1) and G' is the gain factor needed to obtain |H(z)|=1at DC, that is, at z =1. We can see that the filter has N zeros at z=-1.

If x(n) be the input and y(n) be the output, we have the following input-output relation of the filter:

$$y(n) = \sum_{k=0}^{N} b_k x(n-k) - \sum_{k=1}^{N} a_k y(n-k) \dots\dots\dots\dots\dots\dots(5.7)$$

Where b_k, k=0,1,2, ...,N are another set of filter coefficients or tap weights.

The filter is now in the form of an IIR filter, whose direct form realization is given in figure 5.12.
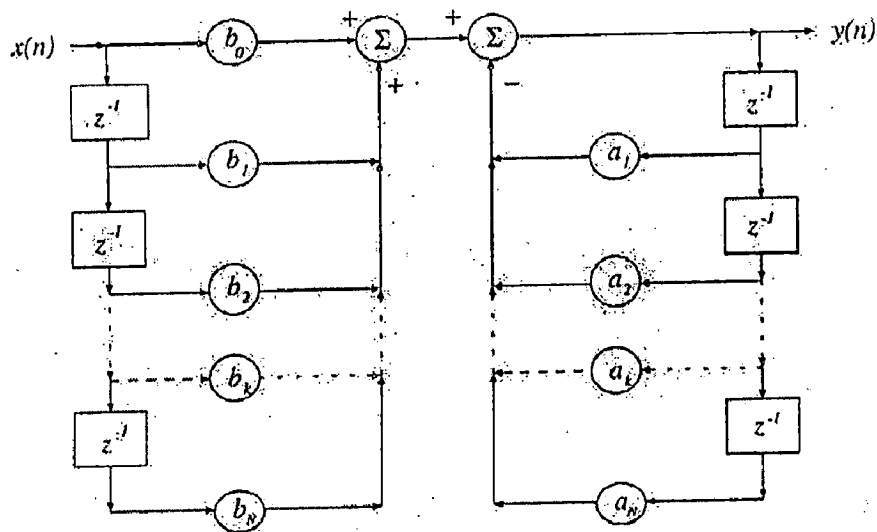
Figure 5.12: Direct Form Realization of the IIR Filter [4]

In our case, the Matlab function 'butter' is used to generate the filter coefficients with specified order (second order) and cut-off frequency (70 Hz). Following figure demonstrates the performance of this filter on ECG data. Input to this filter is the ECG obtained after adjusting the reference and cancelling the baseline wander.
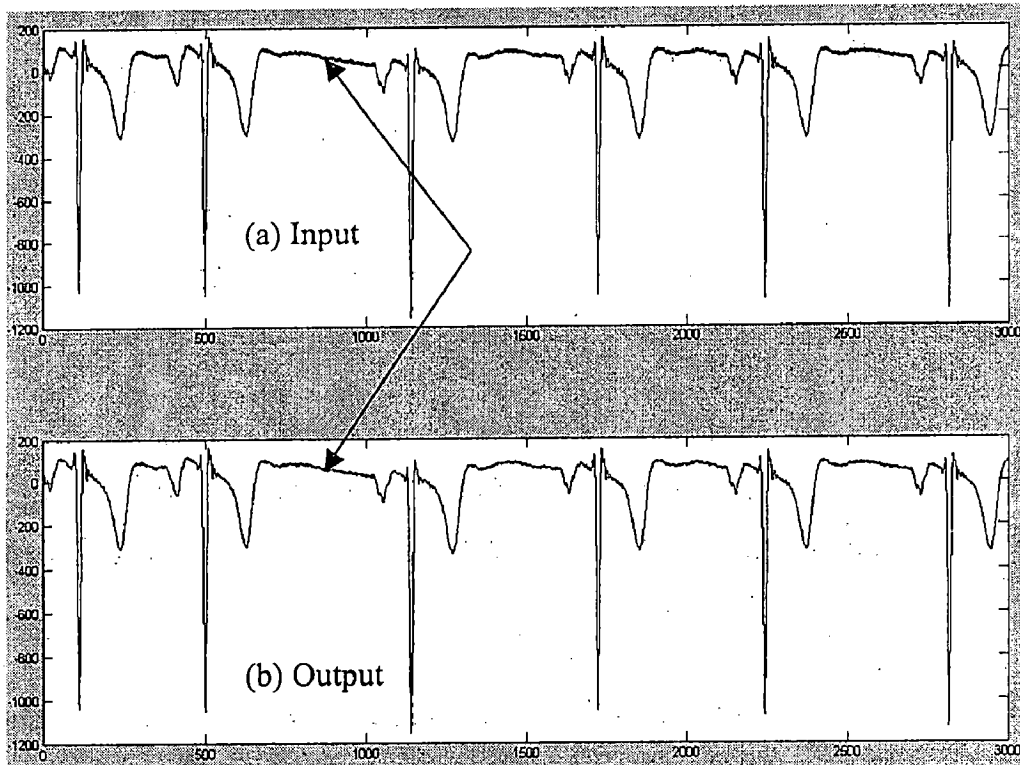
Figure 5.13: Performance of Butterworth Low Pass Filter

5.2.2. TIME DOMAIN FILTERING

Time domain filters are useful in eliminating random noise. In most cases, time-domain processing is found to be faster than frequency domain filtering [4]. For ECG signal conditioning, Savitzky-Golay (SG) Filter is sometimes used as an alternative to Butterworth LP filter. SG filter can be thought of a generalized Moving Average (MA) filtering. For better insight, we begin with MA filtering.

A *Moving Average filter* smooths data by replacing each data point with the average of the neighbouring data points defined within the span [32]. The general form is as follows,

$$y(n) = \sum_{k=0}^{N} b_k x(n-k) \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(5.8)$$

Where, x and y are the input and output of the filter, respectively. The b_k values are the filter coefficients or tap weights, k=0,1,2..........,N. 'N' is the order of the filter. The effect of division by the number of samples used (N+1) is included in the values of the filter coefficients. The signal flow diagram of a generic MA filter is given in the following figure.

65
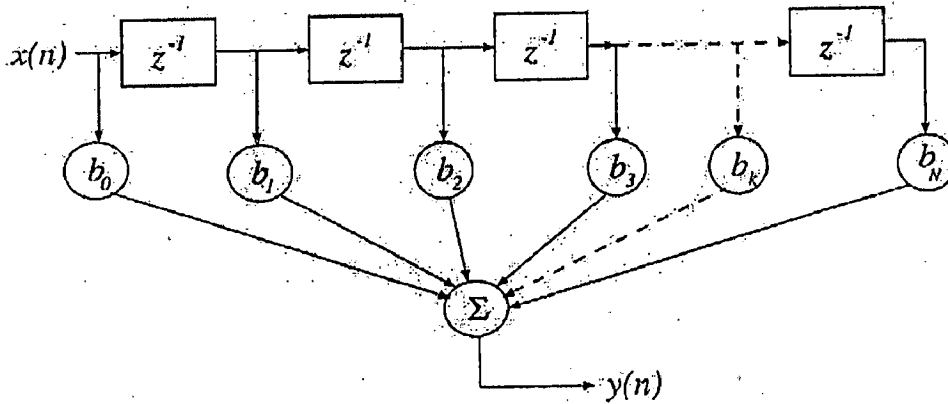
Figure 5.14: Signal Flow Diagram of an MA Filter [4]

The z-domain transfer function of the filter is given as,

$$H(z) = \frac{Y(z)}{X(z)} = \sum_{k=0}^{N} b_k z^{-k} = b_0 + b_1 z^{-1} + b_2 z^{-2} + \cdots \cdots \cdots + b_N z^{-N} \quad (5.9)$$

An MA filter is a finite impulse response (FIR) filter with the following attributes and advantages:

- The impulse response h(k) has a finite number of terms: $h(k) = b_k$, k=0,1,2.........,N.

- An FIR filter may be realized non-recursively with no feedback.

- The output depends only on the present input sample and a few past input samples.

- The filter transfer function in z-domain has no poles except at z = 0. This makes the filter inherently stable.

- The filter has linear phase if the series of tap weights is symmetric or anti-symmetric.

    Savitzky-Golay Filtering can be thought of as a generalized moving average. SG filter coefficients are derived by performing an unweighted linear least square fit using a polynomial of a given degree. For this reason, it is also called a digital smoothing polynomial filter or least squares smoothing filter.

    Matlab function 'sgolayfilt' is used to implement the SG filter with specified degree (d) of fitting polynomial and filter span (s). The following two conditions must be ensured,

    1. The span (s) must be odd. This is also known as the frame size.

    2. The polynomial degree (d) must be less than the span (s). If d = s-1, the filter will produce no smoothing.

SG filter is typically suitable to smooth out a noisy signal whose frequency span (without noise) is large, like our ECG signal. In this type of application, SG filters perform much better than standard averaging FIR filters, which tend to filter out a significant portion of the signal's high frequency content along with the noise [5]. SG filters are optimal in the sense that they minimize the least-squared error in fitting a polynomial to frames of noisy data. Following figure demonstrates the performance of SG filter (with d= 2 and s = 5) on the same signal of figure 5.13 (a).
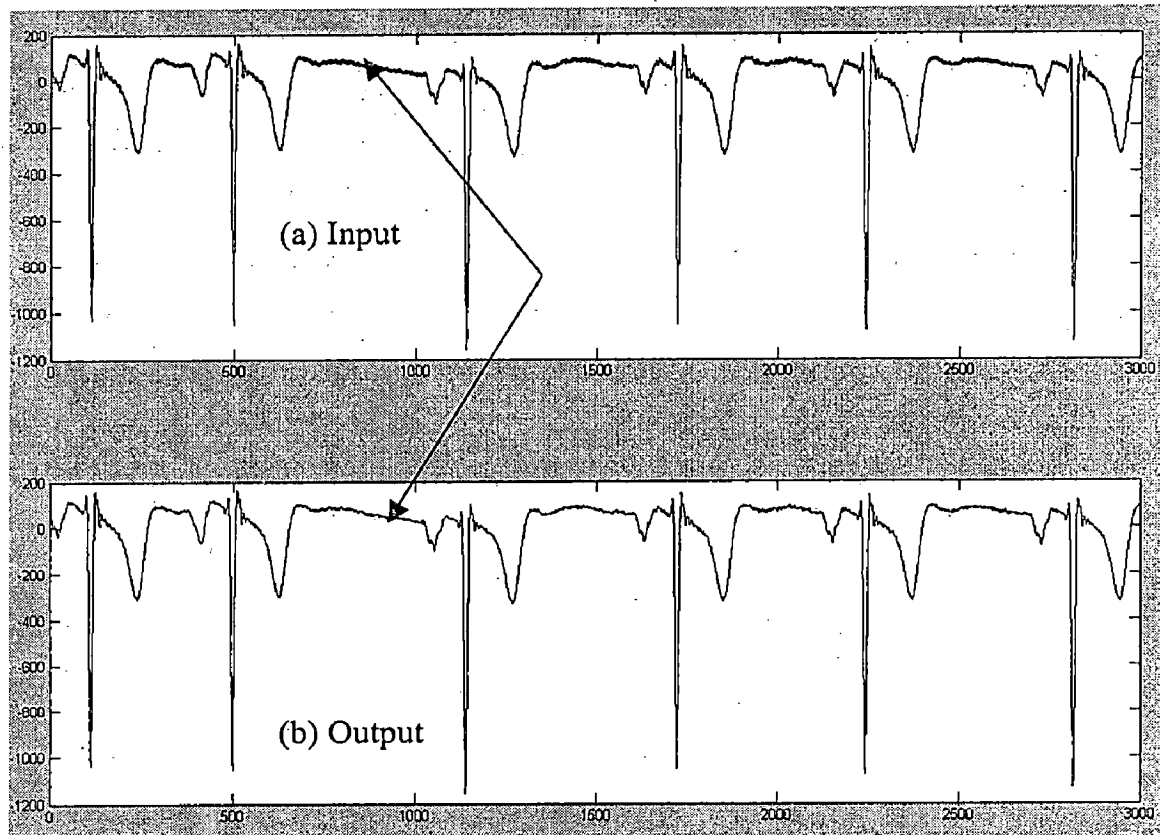
Figure 5.15: Performance of Savitzky-Golay Filter

5.2.3. DRAWBACKS OF BUTTERWORTH/ SAVITZYKY-GOLAY FILTER

The advantage of the Savitzky-Golay Filter is that it takes lesser computation time. However, it is not possible to predict which one is the best among these two. For some ECG data sets, Butterworth LP filter works better than Savitzky-Golay Filter, whereas for some other data sets the latter is found to be better. It depends on the nature of contamination.

For Butterworth low-pass filter, the cut off frequency is taken as 60-70 Hz in our application. However, ECG signal itself sometimes has frequency components more than 70

Hz (specially the QRS complex). Under such circumstances, Butterworth Filtering produces severe distortion of the signal. Savitzky-Golay filter also performs poorly on narrow peaks. This becomes evident from the following figure.
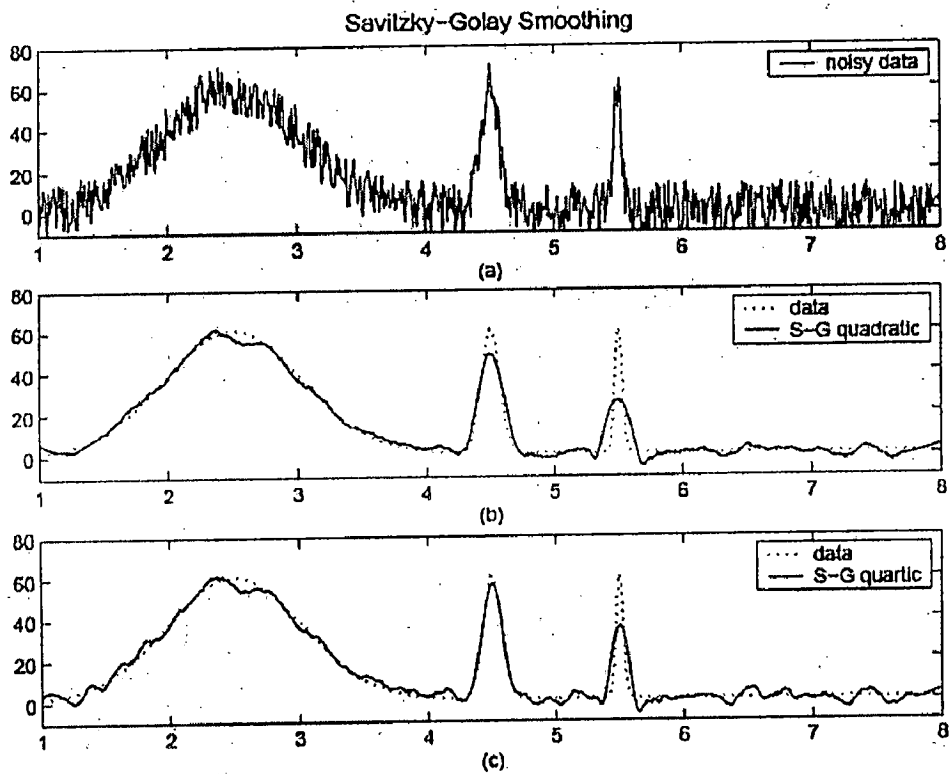
Figure 5.16: Performance of Savitzky-Golay Filter on Narrow Peaks [32]

Figure (a) shows the data added with noise. Figures (b) and (c) show the result of smoothing with a quadratic and quartic polynomial respectively. The dotted line represents the data without adding noise. Higher degree polynomials can more accurately capture the heights and widths of narrow peaks, but can do poorly at smoothing wider peaks. We use quadratic polynomial fitting in our application. This reduces the sharpness of the QRS complex sometimes.

Cancellation of frequency content greater than 70 Hz by Butterworth LP filter or smoothing by SG filter might smooth out the R peak sometimes along with the noise. Therefore, resampling (as discussed in next chapter, section 6.4) is often used during R peak detection to enhance the sharpness.

Both of the filtering schemes for high frequency noise elimination will introduce some distortion in the ECG signal, however small it may be. Still, this filtering is essential before delineation. Therefore, we use the two-stage filtered (i.e. baseline wander

cancelled and Butterworth LP/Savitzky-Golay filtered) ECG signal only for delineation purpose. For Principal Component Analysis, QRS complexes are extracted from the one-stage filtered (i.e. only Baseline wander cancelled) ECG signal based on the delineation result as discussed in Chapter 7.

Most of the clinically useful information in the ECG is found in the intervals and amplitudes defined by its significant points (characteristic wave peaks and boundaries). Delineation of ECG signal means identifying the component waves and complexes in each and every beat. In our algorithm, we will detect 11 fiducial points in each beat. P and T waves are delineated completely by locating the wave onsets, offsets and peaks for each (corresponding to six marking points per beat). For QRS complex delineation, we detect the locations of Q wave onset, Q peak, R peak , S peak and S wave offset (5 marking points per beat).

The automatic detection of ECG waves is important to cardiac disease diagnosis. A good performance of an automatic ECG analysing system depends heavily upon the accurate and reliable detection of the QRS complex, as well as the P and T waves [34].

6.1. RELATED LITERATURE SURVEY

ECG delineation, specially the QRS complex detection has always been a subject of major importance in research. An extensive review of approaches proposed in the last decade can be found in [34]. One can find in the literature many different delineation approaches based on mathematical models, the signal envelope, matched filters, ECG slope criteria, second-order derivatives, low-pass differentiation, the wavelet transform, non-linear time-scale decomposition, adaptive filtering, artificial neural networks or hidden Markov models [14].

Lot of research has been made in the field of ECG delineation using Wavelet Transform (WT). In [15][14][35], Dyadic Wavelet Transform (DyWT) has been proposed. A spline wavelet, which is a derivative of a smoothing function, has been used as the prototype mother wavelet. The implementation is carried out by means of digital filters. The WT at a particular scale is proportional to the derivative of the filtered version of the signal with a smoothing impulse response at that scale. Therefore, the zero-crossings of the WT correspond to the local maxima or minima of the smoothed signal at different scales, and the maximum absolute values of the WT are associated with maximum slopes in the filtered signal.

In [15], first modulus maximum lines corresponding to R waves are searched across four different scales, namely, 2^1, 2^2, 2^3 and 2^4, using different threshold for different scales (based on the corresponding rms value). For a valid R wave, the *Lipschitz regularity* [34] must be greater than zero. Also, the R wave corresponds to a positive maximum-negative minimum pair at each characteristic scale. After applying certain definite criteria, the isolated and redundant modulus maximum lines were rejected. Finally the R peaks were located at the zero-crossing points between the positive maximum-negative minimum pairs at scale 2^1. On the right and left of each detected R peak, the local modulus maxima lines were taken care of for the delineation of rest of the wave peaks and boundaries.

In [14], the same procedure of [15] is extended and evaluated on several manually annotated databases. They also generalize the filter coefficients (for DyWT) for different sampling frequencies of the ECG. Moreover, they considered more morphological variations for T wave in addition to those listed in [15].

In [35], only the QRS complex detection is accomplished as they are more interested in the heart rate variability. They made use of the property that the absolute value of DyWT has localized maxima across several consecutive scales at the instant of occurrence of transients. Applying a definite threshold criterion, the peaks are located in a particular scale. Then the next higher scale is scanned in the same way. If the no. of peaks in both cases does not agree, computation is carried out for the next scale. Finally, for acceptance as QRS locations, three consecutive scales should agree on the same no. of peaks and also the corresponding peak locations in different scales must be within tolerable time deviation.

In [36], an on-line QRS detection algorithm was developed based on the Haar Wavelet and implemented as a recursive filter. They also use magnitude threshold to determine the location of R peaks.

In our algorithm, the DWT decomposition as discussed in section 4.1.7 is used with Haar function as prototype wavelet. A running window and subsequent accumulation method makes our approach threshold-independent.

6.2. MOTIVATION

The first level details coefficients obtained from the Haar wavelet based DWT decomposition of ECG signal (after conditioning) are analysed in our method. In this

section, we will try to discover the utility of 'Haar' wavelet in ECG delineation. As depicted already in figure 4.13, Haar function has a step nature. This is found to be very sensitive to any slope change in the original signal.

Before going to the real ECG signal, the Haar wavelet is applied on some signals with specific shape (e.g. a straight line with a constant slope, a triangular wave, a cosine wave etc). Every signal is decomposed into first level approximation (A1) and details coefficient (D1) as in Figure 4.9. Now the reconstruction is performed with A1 set to all zeros following the same strategy of Figure 4.11. The signal, produced this way will be nothing but the D1 coefficients time-aligned to the original signal (by means of up-sampling). This signal (reconstructed back with A1 all zeros) is referred to as the *'First Level Details Signal* (FLDS)' throughout this chapter.

First let us see the FLDS obtained from a straight line signal with constant slope.

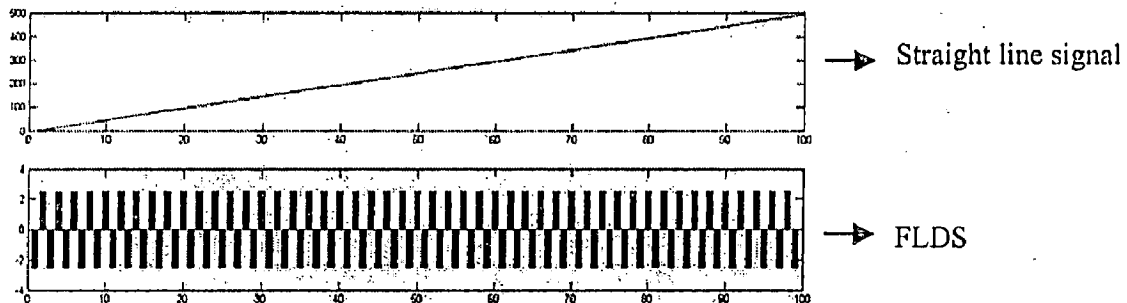Figure 6.1: First Level Details Signal Resulting from a Straight Line

The above figure shows that all the samples of FLDS are of the same amplitude and of alternating signs.

Now, the same thing is repeated on a symmetric triangular wave.
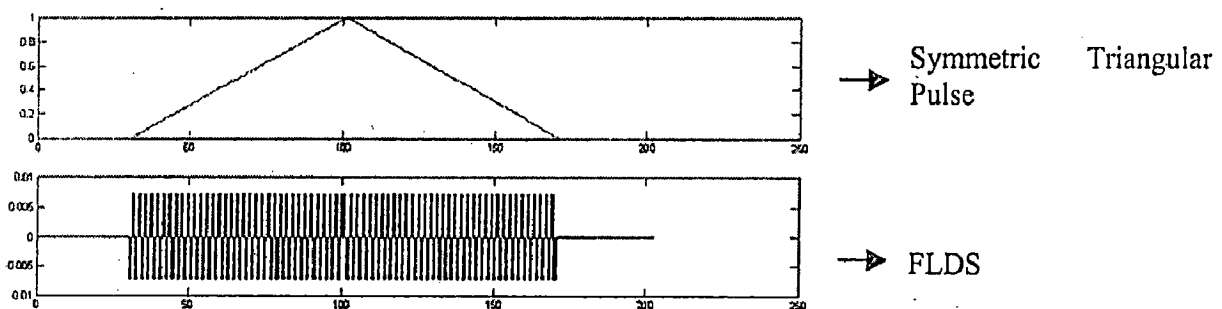

Figure 6.2: First Level Details Signal Resulting from a Triangular Pulse

The above figure illustrates the following characteristics to be discussed next.

73

(1) Samples of FLDS are of the same amplitude and alternating signs as long as the slope remains constant.

(2) When the slope of the signal is zero, FLDS samples have zero magnitude.

(3) When there is a direction change (or sign change in slope) in the original signal, e.g. at the peak of a triangle, two consecutive samples of FLDS are of same sign. Here, when the slope changes its sign from positive to negative (i.e. direction of signal changing from +ve to –ve signifying a positive peak), that reflects two consecutive positive samples in FLDS--- the first one of them marking the instant of direction change (or in other words, occurrence of peak).

(4) The slope magnitude is the same on either side of the triangle, only the sign is different. This owes to the symmetry of the triangle. So, all the samples of FLDS falling under the span of the triangle are of same absolute magnitude.

Next, let us see what happens to the FLDS if the triangle of the previous example is made just upside down.

→ Symmetric Triangular Pulse (-ve Peak)
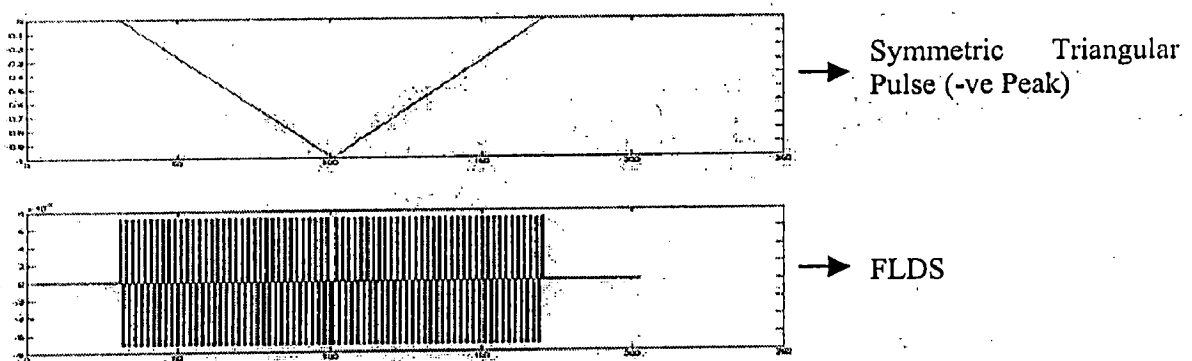
→ FLDS

Figure 6.3: FLDS Resulting from a Triangular Pulse with Negative Peak

It can be seen that when there is a negative peak (signifying direction change of the signal from –ve to +ve), there are two consecutive samples in FLDS with 'negative' sign – the first one marking the instant of the peak.

The flowing two characteristics are computed from the FLDS, on the basis of observations made so far:

1. **Direction Change Mark** (DCM): This is a time vector comprising the same no. of elements as the original signal or FLDS. All elements of this vector will have zero magnitude except at the direction changing points. Whenever there are two positive

74

consecutive samples in FLDS, the element of DCM corresponding to the first sample will be '+1'. On the other hand, two consecutive negative samples of FLDS will reflect a '–1' in the corresponding element of DCM. Therefore, a '+1' in DCM will signify a positive peak in the original signal and a '–1' will represent a negative peak.

2. *Direction Change Sharpness* (DCS): This is also a time vector having exactly the same span as DCM. All the elements of DCS will be zero except at those positions where DCM has a non-zero value.

If DCM has a '+1', the corresponding sample of FLDS is tracked. The absolute difference in magnitude between this sample and the next sample of FLDS is calculated and this value is put at the corresponding position of DCS.

If DCM has a '-1', again the corresponding sample of FLDS is tracked. The absolute difference in magnitude between this sample and the next sample of FLDS is calculated like before. Now, this difference is multiplied with –1 and the resulting negative value is put at the corresponding position of DCS.

Therefore, the samples of DCS will be a replica of those of DCM as far as the sign is concerned.

- If the DCM sample is zero, corresponding DCS sample will also be zero.
- If the DCM sample is positive (i.e. +1) , the corresponding DCS sample will also be positive (but can have any magnitude depending on FLDS).
- If the DCM sample is negative (i.e. -1) , the corresponding DCS sample will also be negative (but can have any magnitude depending on FLDS).

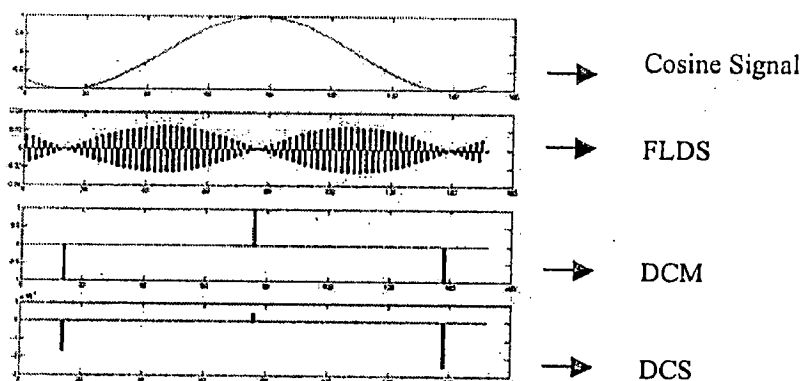Now let us see how these things , FLDS, DCM and DCS will look like on a cosine signal.

Figure 6.4: FLDS, DCM and DCS Resulting from a Cosine Signal

75

In Figure 6.4, we can see that the slope of the cosine signal varies continuously and the samples of FLDS follow the pattern of slope change closely. We have a collection of ordered pairs of consecutive samples— elements belonging to the same ordered pair will have the same magnitude but opposite sign. This leads to the symmetrical positive and negative halves of FLDS. Moreover, two consecutive samples of FLDS (belonging or not belonging to the same ordered pair) are always of opposite signs------only exceptions are found at the local extrema of the original signal (where the slope changes its sign).

However, DCM or DCS are only sensitive to the change in 'sign' of slope (i.e. a direction change in the signal) and not to the change in 'magnitude' of slope. While observing the difference in magnitude of DCS elements in Figure 6.4, the time discretization in plotting should be taken care of.

Now, let us see the application on real ECG segments (after two-stage conditioning as described in Chapter 5).
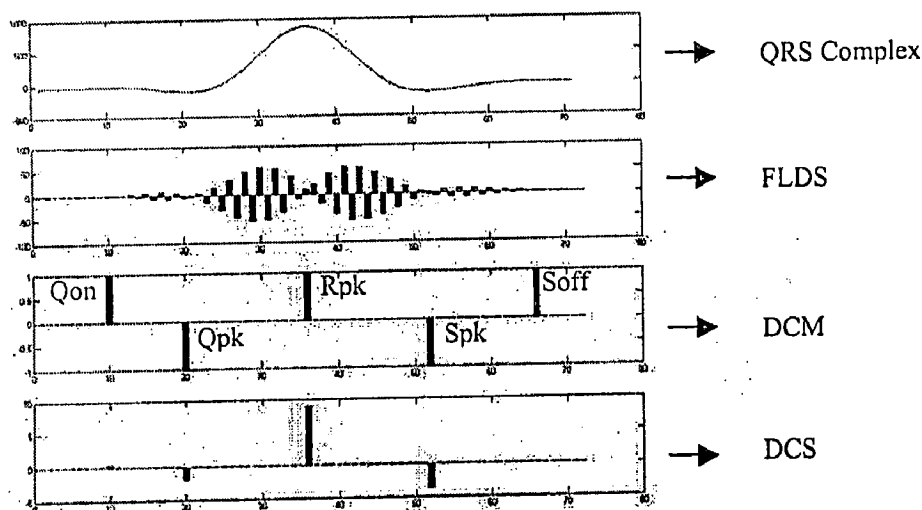
Figure 6.5: FLDS, DCM and DCS Resulting from a QRS Complex

It can be seen that five direction changing points are marked in DCM corresponding to a QRS complex, namely, Q onset, Q peak, R peak, S peak and S offset.
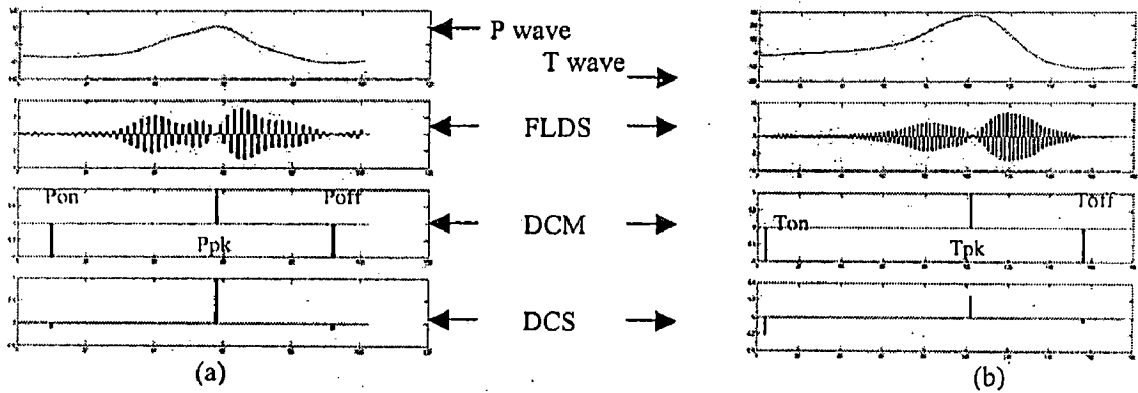
Figure 6.6: FLDS, DCM and DCS Resulting from (a) P Wave, (b) T Wave

For both of P and T waves, the onsets, offsets and peaks have been marked in DCM. Now, for each ECG beat, 11 fiducial points are marked as shown in figure 6.7.
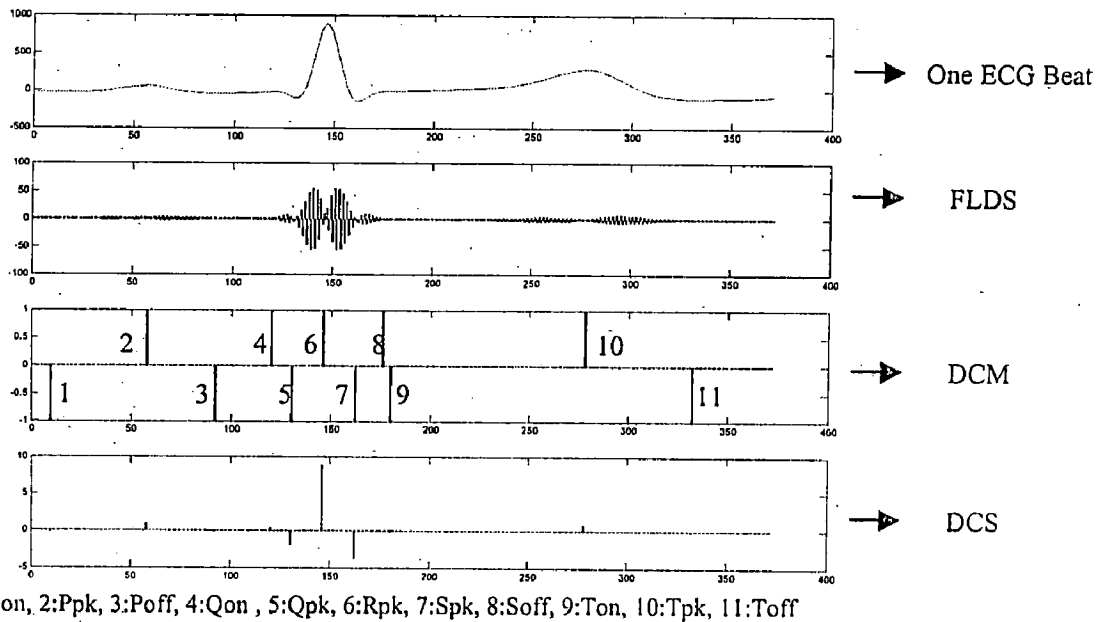
1:Pon, 2:Ppk, 3:Poff, 4:Qon , 5:Qpk, 6:Rpk, 7:Spk, 8:Soff, 9:Ton, 10:Tpk, 11:Toff

Figure 6.7: FLDS, DCM and DCS Resulting from a Whole ECG Beat

6.3. SINGLE CHANNEL DELINEATION STRATEGY

First we will start with the delineation strategy on a single channel ECG data. Then it will be modified for multi-channel delineation.

R Peak Detection

All the R peaks from the ECG data are detected first, after proper conditioning. For this, we take help of a running window of fixed length. The procedure can be described step by step as follows:

1. *Window Length Selection*:

First few seconds (2-6 seconds) of ECG containing some 5 or 6 R peaks is plotted on a graph to decide on the window size. The mean R-R interval (observed from this plot) is taken care of while deciding the window length (WL) for the whole data.

WL should only satisfy the following two conditions:

(1) It should be less than one R-R interval. This will prevent false negative (defined in section 6.5) detection.

(2) It should be more than half of the R-R interval, to prevent false positive (defined in section 6.5) detection.

2. *Extending Data Length*:

The ECG data-length is extended at the end by a set of samples of zero amplitude spanning WL. Say if the ECG length be L, the new length will be (L+WL). The last WL samples will all have zero magnitude.

This is done so that even the last data sample can be analysed properly. The window translation is carried on the signal of length (L+WL). Translation is stopped when the remaining data length is found to be less than WL.

3. *Running Fixed Length Window*:

The window is made to run over the whole data set (spanning L+WL) in steps of say 5 samples (sampling frequency of ECG being in the range of 250-1000Hz). A still smaller step will take longer computation time. However, a longer step might introduce errors in detection (reason will be clear in the next step). The incremental step (IS) of 5 samples seems optimal. IS should be chosen in such a way that WL is always divisible by it.

The instant of occurrence of the max/min (depending on R peaks being positive/negative) value of DCS inside each window is noted. If the window encompasses any R peak, then it will cause the sharpest direction change in the ECG, and hence the max/min value in the corresponding sample of DCS. If the channel is known to have positive R peaks, the time of occurrence of the highest DCS value for each and every window will be accumulated in a vector called *'Extreme Direction Change Sharpness'* (EDCS). If the ECG channel contains negative R peaks, the time of occurrence of the lowest values of DCS will be accumulated in EDCS. Following figure 6.8 illustrates the phenomenon.
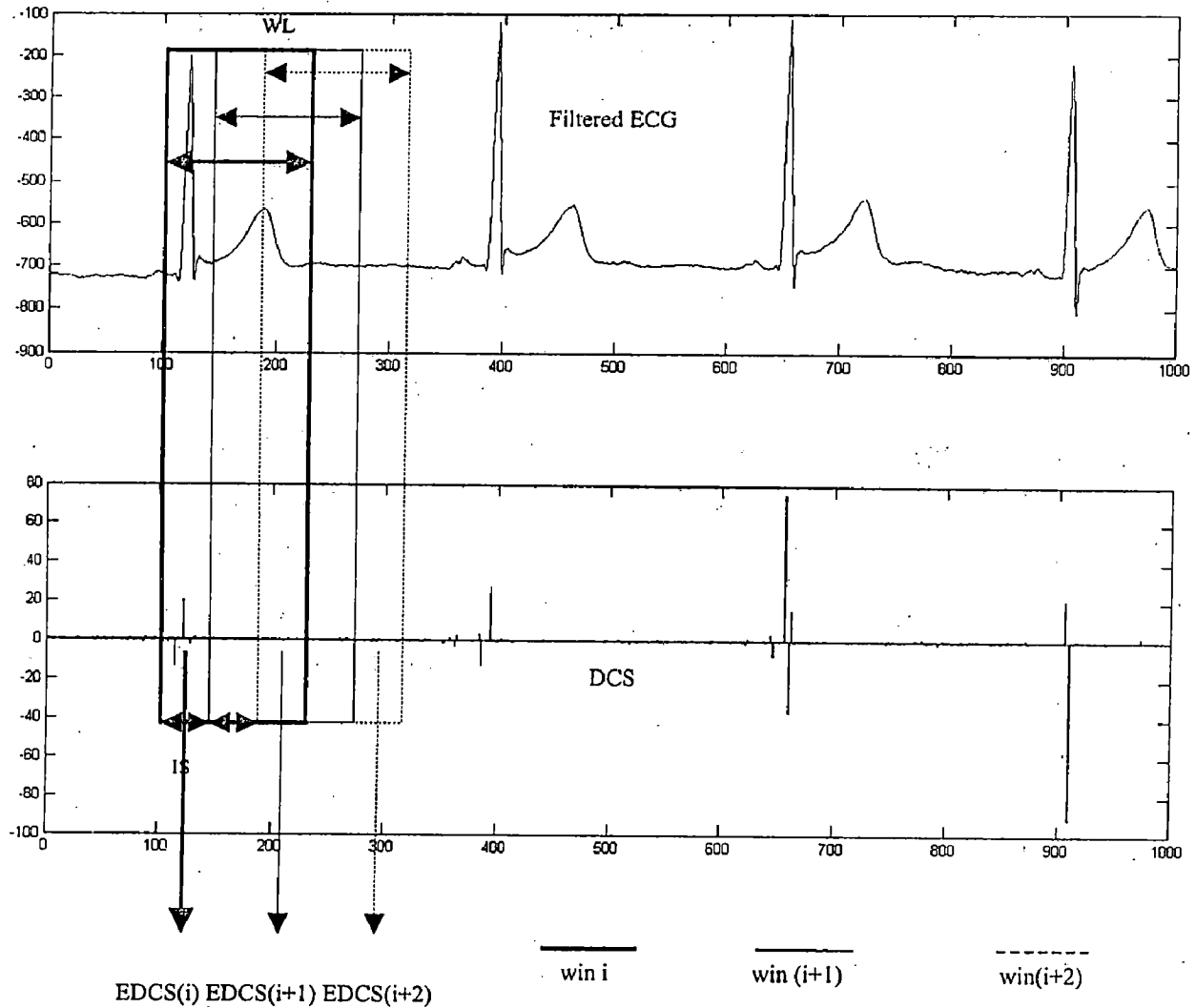
Figure 6.8: Running Window on ECG Signal

In the above figure, the ECG channel consists of positive R peaks. Hence, the instant of occurrence of the highest DCS value inside each window is saved in EDCS. Also it is seen that in each ECG beat, the DCS corresponding to the R peak is the highest.

An interesting phenomenon is noticed afterwards. If the same R peak is enveloped by 'n' consecutive windows, the same value will occur consecutively n times in the EDCS vector. Following figure makes this observation clear.
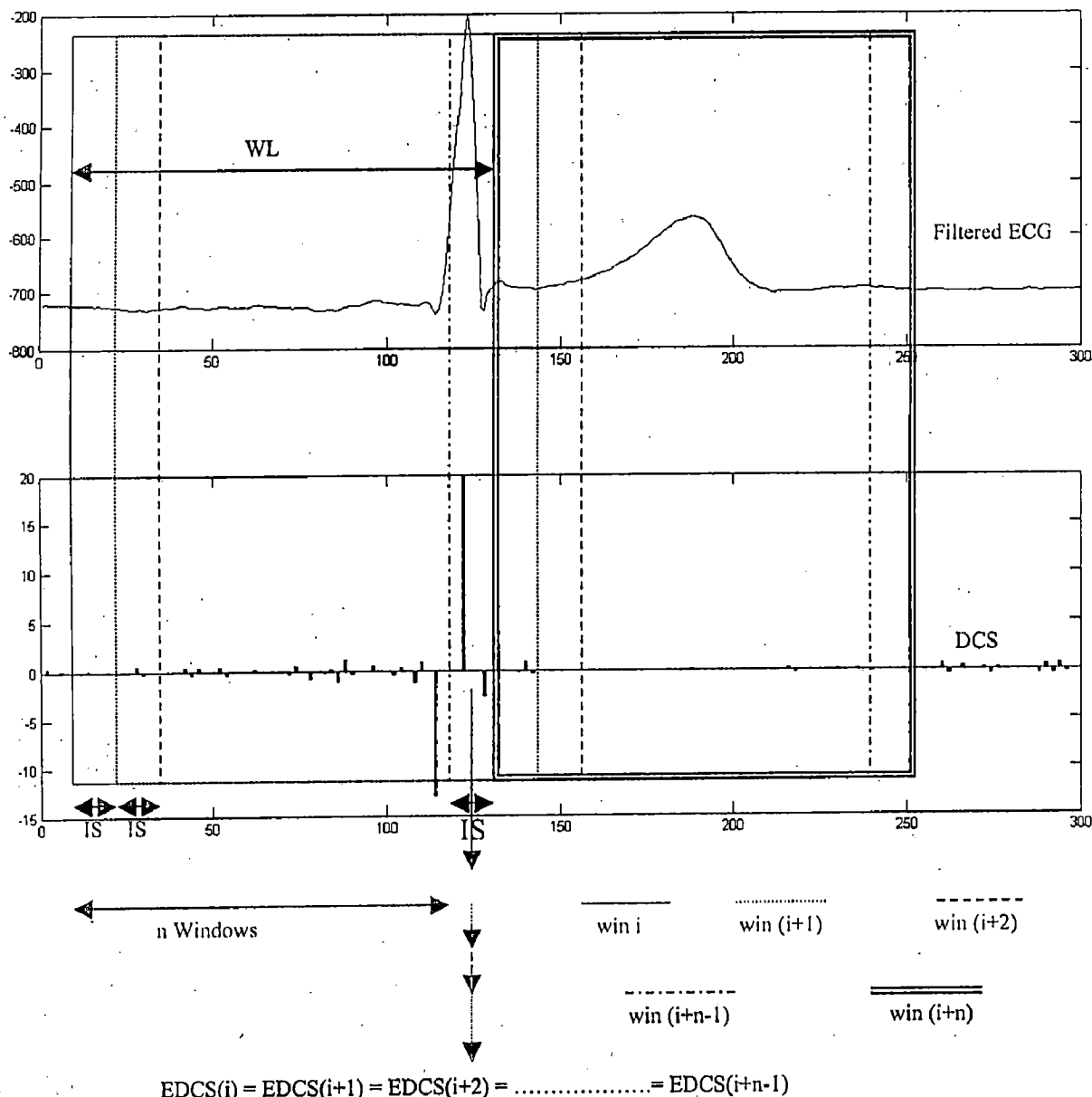
Figure 6.9: Running·Window on one ECG Beat

It is seen that all the above windows, except win (i+n) envelop the same R peak and hence they will reflect the same value in EDCS. Here the number of windows, which are tracking the same R peak, is assumed to be 'n'. Let us examine the value of this 'n'.

It can be seen that 'win i' is the first window, which starts enveloping this R peak and 'win (i+n-1)' is the last one to track this same R peak. From 'win (i+n)' onwards, this R peak will not be encompassed. The end of 'win i' and the start of the 'win (i+n)' are exactly coincident. This is ensured by the divisibility of WL by IS. Hence we can easily interpret,

$$n = WL/IS \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(6.1)$$

4. *Identification of R peaks from EDCS Values*:

After the windowing and accumulation of EDCS over the whole length of signal is accomplished, the next concentration is on localizing the R peaks. First, all the values of EDCS greater than exact ECG length (L), are discarded. Now, the elements in EDCS occurring at least 'n' times consecutively are denoted as the locations of R peak.

It can now be understood why IS should not have a large value. For larger IS, total no of window positions (in course of translation) will be lesser and that will lead to lesser accumulation of each individual R peak in EDCS. Under such circumstances, the accumulation of R peak might become comparable with that of P or T peaks. This is why IS should be kept as small as possible depending on the computation time that can be allowed and the memory resources.

5. *Enhancement of DCS*:

It is clear that the success of R peak localization solely depends on the elements of EDCS, which are in turn determined by the samples of DCS. If some mechanism can be devised, which will enhance the DCS samples corresponding to R peak locations --- that will surely be an improvement in our methodology. With this aim, three different sets of DCS are calculated:

1) *DCS_up*: ECG Signal (after extending by WL) is up-sampled by a factor of 2 and the DCS corresponding to this upsampled ECG is computed. Now, in order to achieve the same length (L+WL) as the normal ECG, this DCS is downsampled by 2.

2) *DCS_normal*: The DCS obtained from the ECG signal (after extending by WL) with normal sampling rate.

3) *DCS_down*: Original ECG Signal (after extending by WL) is down-sampled by a factor of 2 and the DCS corresponding to this down-sampled ECG is calculated. Now, in order to achieve the same length (L+WL) as the ECG, this DCS is up-sampled by 2.

Therefore, all the three, namely, DCS_up, DCS_normal and DCS_down are time aligned with the normal ECG signal (originally recorded). We calculate the modified DCS as follows,

$$DCS\_modified = DCS\_up + DCS\_normal + DCS\_down \quad \dots\dots\dots\dots(6.2)$$

The window translation is carried on DCS_modified . EDCS formation and thereafter R peak detection is now made considering the modified DCS. Up-sampling is done by means of interpolation and down-sampling, by discarding every alternate sample.

Following Figure 6.10 shows the enhancement of DCS_modified in comparision to DCS_normal corresponding to R peak locations.
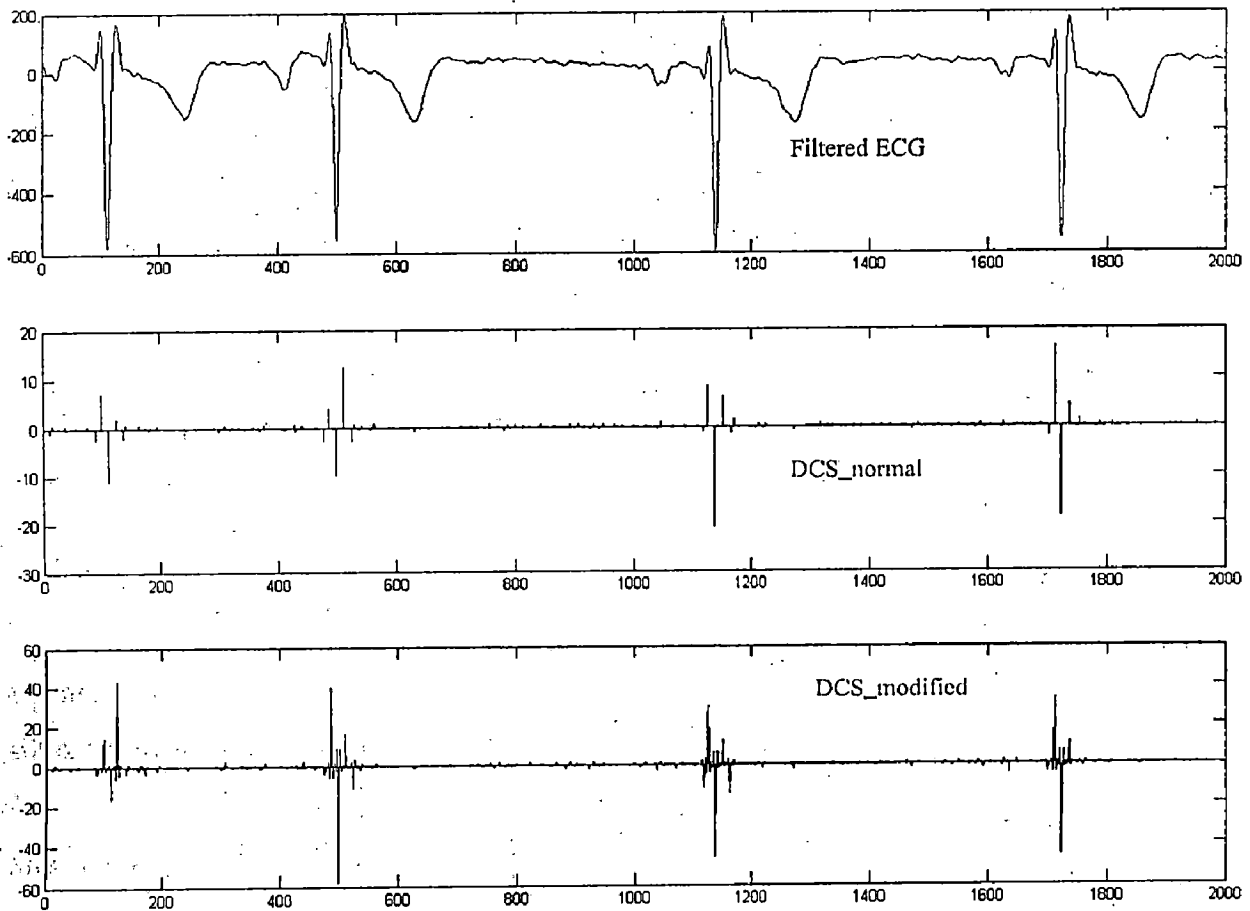
Figure 6.10: Enhancement of R Peaks: DCS_normal & DCS_modified Compared

The vertical scales should be noted while comparing DCS_normal and DCS_modified in the above figure. This modified DCS is only used for R peak detection. For delineation of P,Q,S and T waves we use only DCS_normal (refered to as DCS).

6. Modification of R peak detection method for Exercise ECG:

For exercise ECG, the bpm is not stable any more and hence the beat-to-beat variation in R-R intervals is quite high. For that reason, it is not reasonable to keep the window length (determined on the basis of initial R-R intervals) fixed for the whole data. Here, we split the ECG in segments of 5 minutes with an overlap of 2 seconds between two

consecutive segments. Now, for each segment, the R peak detection will be carried out as described before. The mean R-R interval obtained from the previous segment is taken care of to determine the window length (WL) for the next segment.

P & Q Waves Detection

After detecting all R peaks in the ECG, we now zoom into each and every beat. A search window, spanning half of the previous R-R interval is taken prior to each detected R peak as shown in **Figure 6.11**.

Figure 6.11: Detection and Delineation of P and Q waves

Here, it is an example of a positive R peak, negative Q peak and positive P peak, reflecting '+1', '-1' and '+1' in DCM respectively. The '-1' in DCM immediately prior to the R peak is the location of Q peak. The '+1' just before the Q peak is the onset of Q wave. Had the Q peak been positive, it would have caused a '+1' in DCM and the corresponding Q-onset a '-1'.

Now after detecting the Q onset, the highest positive DCS value prior to it is located. This is the P peak. The two '-1' in DCM surrounding the P peak are its onsets and offsets. Had the P wave been negative, search should have been made for the lowest negative DCS value to locate the peak and the two surrounding '+1' in DCM for the onset and offset.

83

S & T Waves Detection

In a similar method, S and T waves are delineated in each and every beat after detecting all the R peaks in the filtered ECG. Here, for each beat, a search window spanning half of the next R-R interval is takennext to each detected R peak as shown in Figure 6.12.

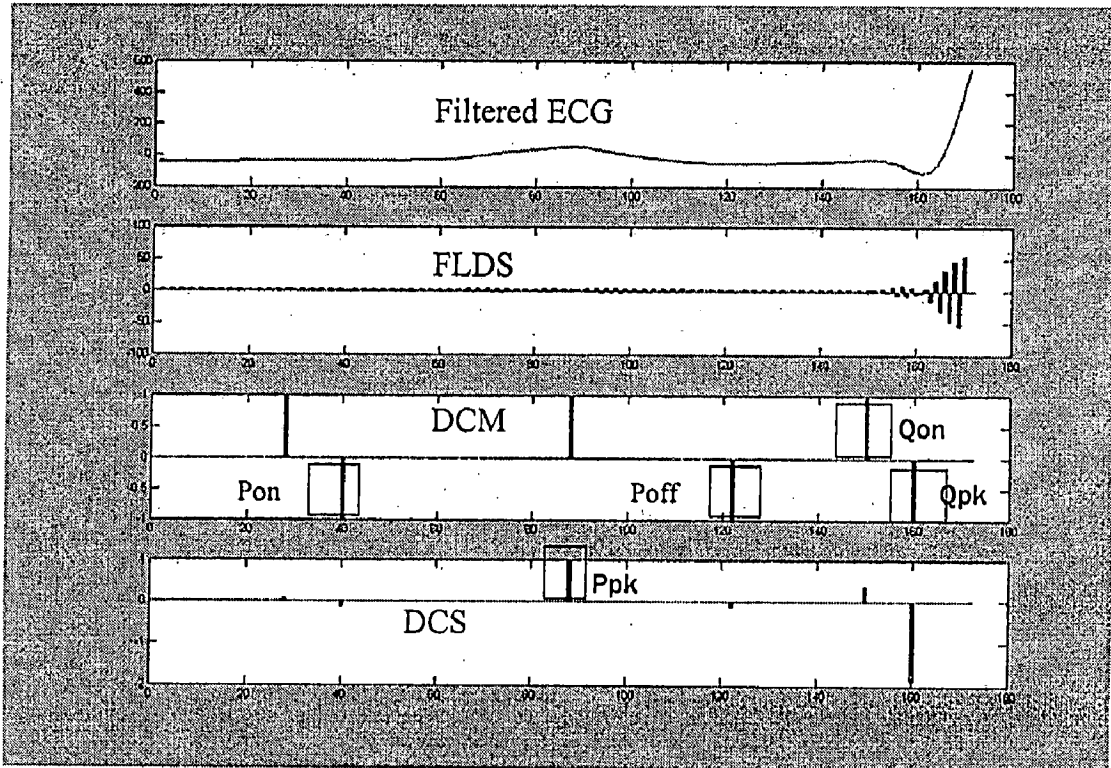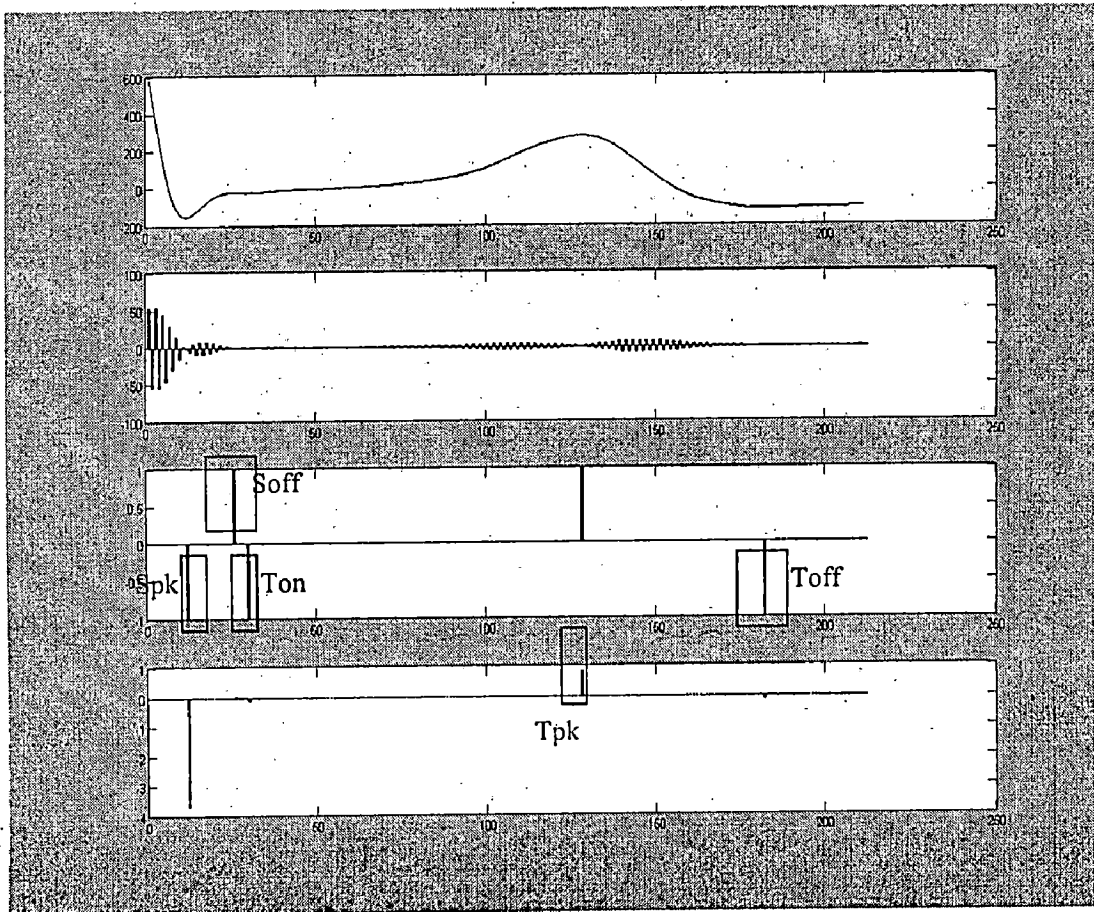

Figure 6.12: Detection and Delineation of S and T waves

This example shows positive R peak, negative S peak and positive T peak, causing '+1', ' –1' and '+1' in DCM respectively. The '–1' in DCM immediately next to the R peak is the location of S peak. The '+1' in DCM immediately after the S peak is the offset of S wave. Had the S peak been positive, it would have reflected a '+1' in DCM and the corresponding S-offset a '–1'.

On the right of S offset, the T wave peak, onset and offset are determined in the same manner as it is for P wave.

6.4. MULTI-CHANNEL DELINEATION

When there are many different channel data pertaining to the same ECG record, it is reasonable to take into consideration the delineation result obtained from a number of channels to reach the final decision.

The R peak detection is carried out simultaneously on several channels. The selected channels for this purpose should have unipolar R peaks (negative or positive) as shown in the following figure 6.13.
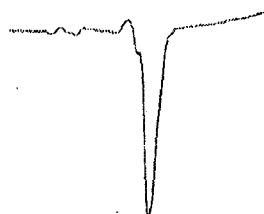
(a) Positive R Peak   (b) Negative R peak

Figure 6.13: Unipolar R peaks

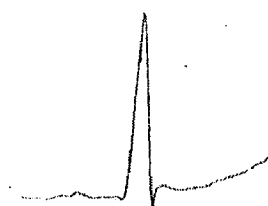The detection strategy for each individual channel will be exactly same as that described before under section 6.3. Say, we select 20 channels for R detection and the no. of beats present in the ECG is say 10. Now, each of these 10 R peaks will be detected in 20 channels. So, for each and every R peak location, there will be 20 values (or sample numbers).

Say for one particular R peak, we have the following result (denoting the sample no.) from 20 different channels:

99,99,101,99,100,120,99,99,100,101,98,100,100,94,100,99,98,100,100,99

A histogram plot of these values is given in figure 6.14.

Figure 6.14: Histogram for Same R Peak Location Detected in

85

It is seen that the values 100 and 99 are having the highest frequency of occurrence. Therefore, the final decision regarding this R peak location w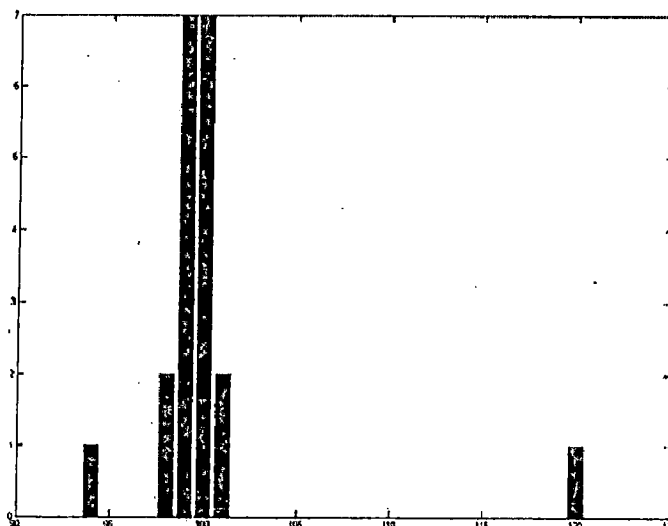ill be the sample number, obtained as the integral mean of 100 and 99 (i.e. 100). The general rule is to find out the value with highest frequency of occurrence. If more than one value is found to satisfy the same condition, the integral mean of them is taken as the final decision.

The number of false positive or false negative detections (defined in section 6.5) can be reduced considering multiple channel results. In order to be accepted as a final decision, the same R peak needs to be detected at least in half the no. of selected channels (here 10). This way, false positive detection is reduced. On the other hand, if there is any false negative detection in one channel, the same can be corrected considering delineation results from other channels. This way, our delineation algorithm becomes more robust and independent of single channel errors.

After computing the final decision regarding the R peak location in Multi-Channel ECG, next step is the delineation of P,Q,S and T waves. For P and T waves, only the peaks are detected in the first go. Again simultaneous detection is performed on selected number of channels. The channels selected for P/T wave delineation should have a prominent unipolar P/T wave (negative or positive peak). Channels selected for Q/S delineation should show prominent Q/S spikes. The same kind of histogram-based method is used to get the final delineation result for P peak, Q onset, Q peak, S peak, S offset and T peak.

Now with the final decision regarding the locations of P and T peaks, the respective onsets and offsets in different channels are identified based on the same peak locations for all channels. Again, the histogram method is taken care of to get the final delineation of onsets and offsets of P and T wave.

Therefore, the delineation is carried out in three stages:
1. R peaks delineation
2. Q peaks, Q onsets, S peaks, S offsets, T peaks and P peaks delineation
3. Delineation of the onsets and offsets of P and T waves

Each delineation stage is dependent on the previous stage. The channels for delineation should be selected judiciously so as to ensure prominent wave-shapes.

The result of delineation obtained from our algorithm is stored in a matrix form, as shown in the following figure. Each row represents information pertaining to a particular beat, whereas each column denotes one of the 11 fiducial points.

	Pon	Ppk	Poff	Qon	Qpk	Rpk	Spk	Soff	Ton	Tpk	Toff
Beat #1											
Beat #2											
⋮											
⋮											
⋮											
⋮											
⋮											
⋮											
⋮											

Figure 6.15: Structure of Delineation Result Matrix (DRM)

6.5. VALIDATION

The performance of the delineation algorithm was checked by comparing the automatic delineation result against manual delineation. Following mathematical calculations were performed to evaluate this comparison.

1. *Mean Error (M.E.):*

   Mean of the difference between manual & automatic delineation (in terms of no. of samples)

2. *Standard Deviation (S.D.):*

   Standard deviation of the difference between manual & automatic Delineation (also in terms of no. of samples)

3. *True and false positives / negatives:*

Real	Existing	Existing	Not Existing	Not Existing
Predicted	Existing	Not Existing	Existing	Not Existing
	True Positive (TP)	False Negative (FN)	False Positive (FP)	True Negative (FN)

4. *Sensitivity (Se):*

$$Se = \frac{TP}{TP + FN} \times 100\% \quad\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(6.3)$$

5. *Positive Predictivity (P+):*

$$P+ = \frac{TP}{TP + FP} \times 100\% \quad\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(6.4)$$

The algorithm of single-channel R peak detection was tested on several records taken from MIT-Arrhythmia database. The output of our algorithm was compared with the manual annotation provided with each record.

Record No.	100	101	103	113	115	122	234	100	103
Channel No.	1	1	1	1	1	1	1	2	2
Samples taken	650000	650000	650000	650000	650000	650000	650000	650000	400000
FN	0	3	0	3	1	6	20	1	0
FP	1	4	0	4	6	5	1	0	0
TP	2271	1863	2084	1791	1953	2469	2732	2271	1294
Se	100	99.84	100	99.83	99.95	99.75	99.27	99.95	100
P+	99.96	99.78	100	99.78	99.69	99.8	99.96	100	100
M.E.	-1.9	-0.48	-2.23	-1.74	-3.16	-3.95	-1.43	0.736	-2.97
S.D.	2.46	1.18	2.31	1.38	3.54	5.87	1.26	2.4	3.08

Table 6.1: Validation of Delineation Result on MIT-Arrhythmia Database

The overall Sensitivity (Se) and Predictivity (P+) obtained for MIT-Arrhythmia Database was 99.84% and 99.89% respectively.

Table 6.2 shows the validation result for a 32-Channel ECG data set with sampling frequency 1 KHz. The output of our algorithm was compared with the manual estimation of first 22 beats. Manual estimation is done by us, and hence is not a perfect cardiological annotation.

	Pon	Ppk	Poff	Qon	Qpk	Rpk	Spk	Soff	Ton	Tpk	Toff	Overall
M.E.	-0.48	-1.83	-4.52	3.52	1.17	-0.52	0.26	0.09	1.08	0.78	0.09	-0.03
S.D.	3.96	3.39	3.29	1.04	1.03	0.59	0.75	0.42	1.12	0.95	1.59	2.83
Se	100	100	100	100	100	100	100	100	100	100	100	100
P+	100	100	100	100	100	100	100	100	100	100	100	100

Table 6.2: Validation of Delineation Result on Multi-Channel ECG

The overall Sensitivity (Se) and Predictivity (P+) values for this dataset were both 100%.
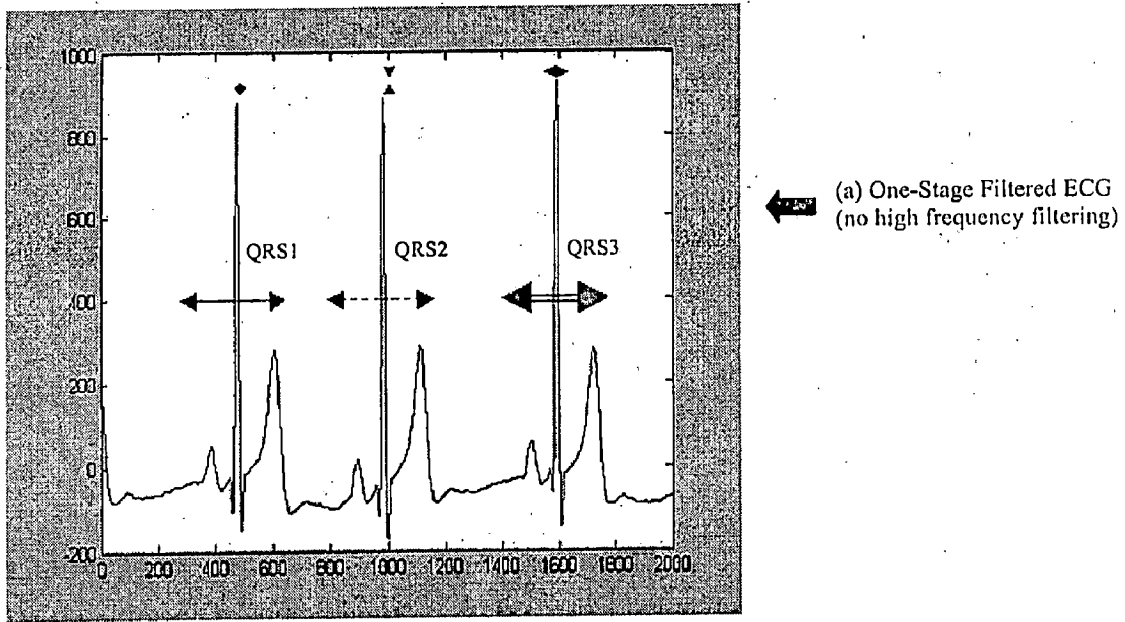
QRS COMPLEX EXTRACTION AND ALIGNMENT

The aim of QRS complex extraction and alignment is to create the input data matrix for Principal Component Analysis, as depicted in Figure 4.16. Here, different variables mean different QRS complexes. The total no. of observations should be taken so that even the longest QRS interval is taken care of. The time-alignment of QRS complexes is needed in order to avoid error in PCA output.

7.1. METHOD

First let us see the QRS extraction methodology, which is based on our delineation result.

1) *Determining the longest QRS span:* The QRS interval is defined as the time length between the Q wave onset till the S wave offset. The longest QRS span (in terms of time samples) for a particular ECG dataset is determined from the Delineation Result Matrix (DRM) (referring to figure 6.15) and a safety factor is added to it in order to take care of errors in delineation. If this comes to be an even number, we make it odd by adding 1 to it.

2) *Initialising the QRS Extraction Matrix (QEM):* This matrix should have as many no. of rows as the no. of detected beats in the corresponding ECG data set (i.e. same as the no. of rows of DRM). The number of columns should be the longest QRS span plus safety factor (odd number always).

3) *Positioning of R peak:* Although the delineation is carried on the ECG after two stages of filtering, QRS complexes are extracted from the first-stage filtered ECG (i.e. only after cancelling baseline wander but no high frequency filtering). The R peak location for a particular beat is found from DRM and then the corresponding magnitude is obtained from the first-stage filtered ECG and copied in the middle column of that particular row of QEM.

4) *Copying the QRS complex:* After positioning the R peak in the middle column of the respective row, the ECG sample magnitudes are copied from both sides of the R peak in QEM.

Figure 7.1 demonstrates the procedure of QRS extraction from a particular channel.

(a) One-Stage Filtered ECG
(no high frequency filtering)

QRS1 QRS2 QRS3

Middle
Column

(b) Extracted QRS

QRS1

QRS2

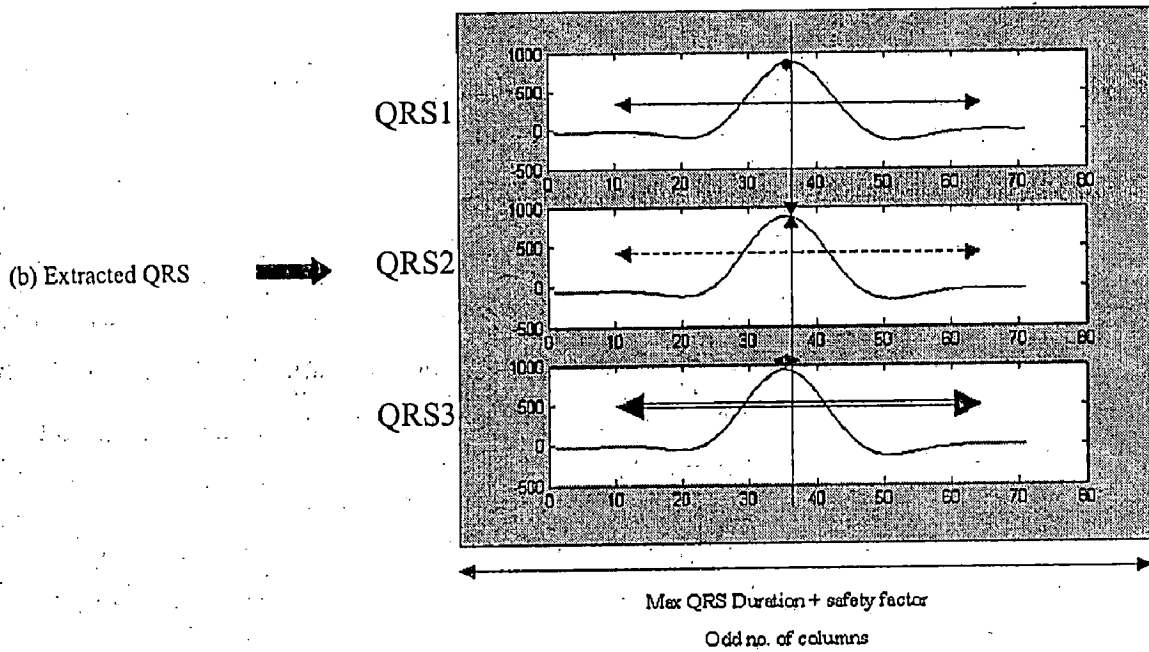QRS3

Max QRS Duration + safety factor

Odd no. of columns

Figure 7.1: QRS Extraction Procedure

If the QRS Delineation is made perfectly, then the extracted QRS will be time-aligned. We can see in figure 7.1(b) that all the R peaks lie in a vertical line. However, if there is any error that creeps in delineation result, it will lead to misalignment. Hence, improved alignment of extracted QRS complexes is performed based on cross-correlation technique. The procedure is as follows.

1) Calculate the mean QRS by taking the mean of all rows of QEM and consider this as the template for cross-correlation. The idea of taking the mean as the template for cross-correlation has been an inspiration from [37] and [38].

2) Calculate the cross-correlation coefficient between each of the extracted QRS and the template.

3) Now shift each of the QRS complexes a little bit towards left and towards right (till +3 and −3 samples) from the template and go on calculating the cross-correlation coefficient between the template and the QRS at that position.

4) Each QRS is finally aligned at the position corresponding to the highest cross-correlation coefficient. The final matrix containing aligned QRS is referred to as the PCA Input QRS Matrix (PIQM).

Before applying the above method on our extracted QRS complexes, it was tested on 100 shifted Meyer Functions. Following figure shows the result.
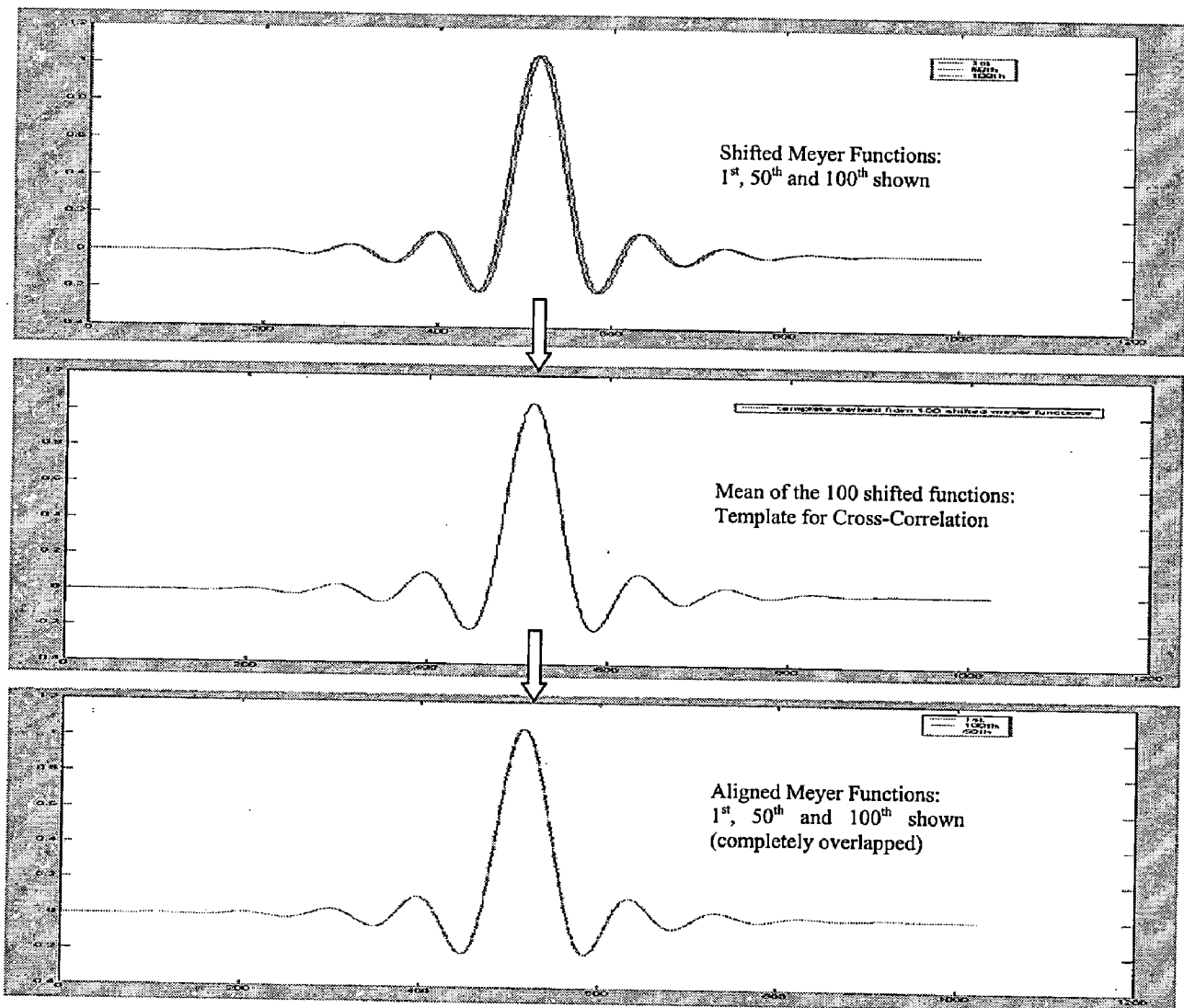

Figure 7.2: Test of Alignment Technique on Shifted Meyer Functions

93

This method will correct if there is any misalignment between the extracted QRS complexes. Following figure 7.3 shows this correction.

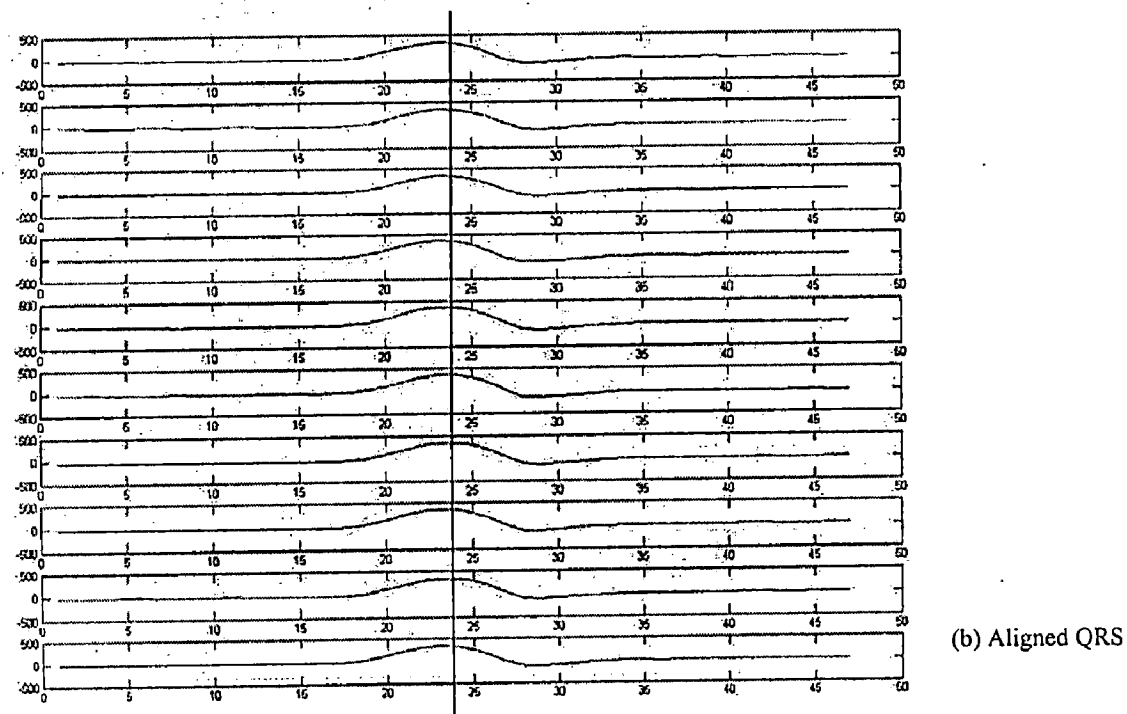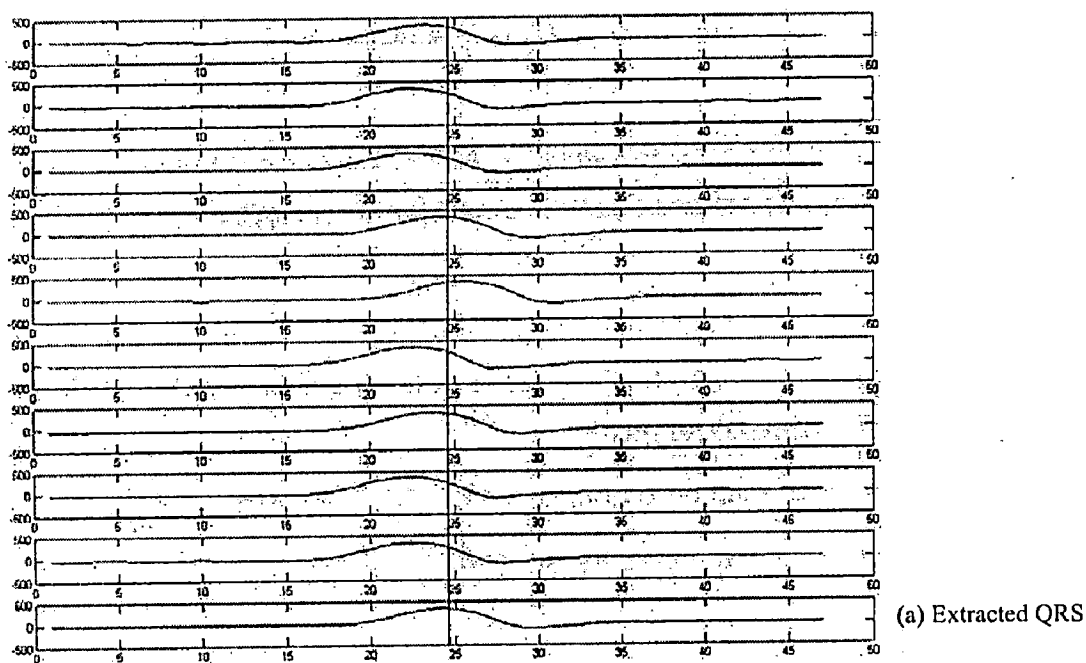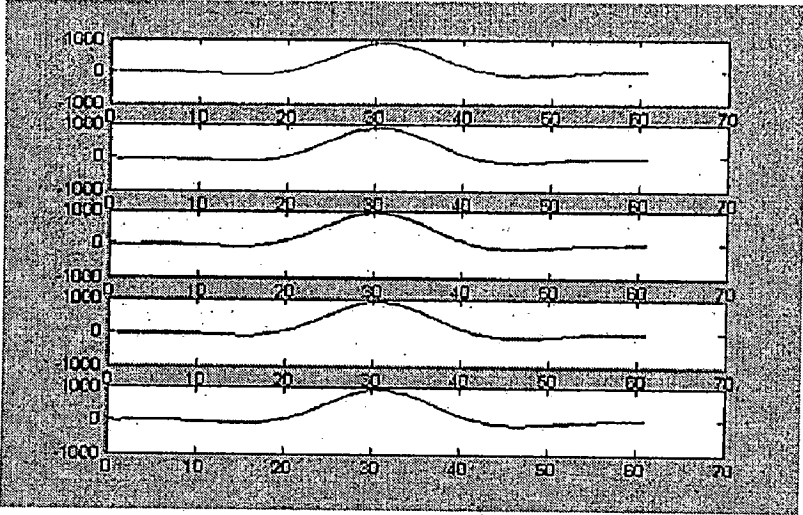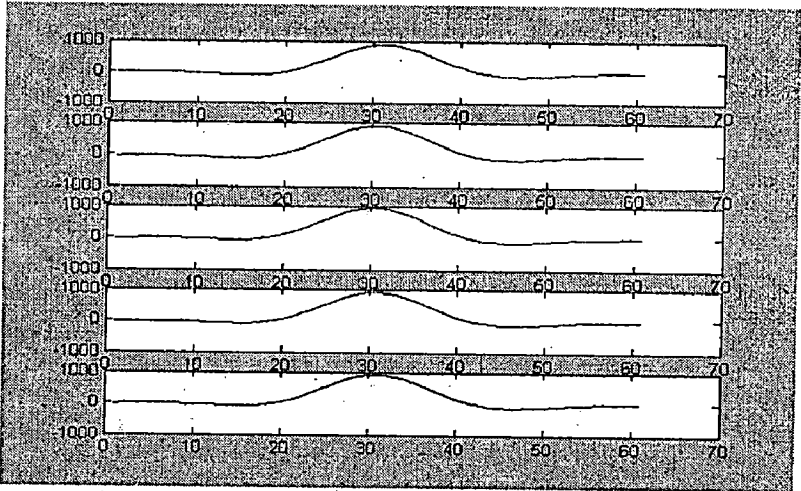(a) Extracted QRS

(b) Aligned QRS

Figure 7.3: QRS Alignment: Case 1

However, if the QRS are already aligned after extraction, this procedure will make no difference. Following figure illustrates this point.

(a) Extracted QRS: Alreday Time- aligned due to perfect Delineation

(b) Output of Alignment Algorithm: Same as input (a)

Figure 7.4: QRS Alignment: Case 2

Output of this stage, i.e. the matrix PIQM will be the input to our next PCA module.

TEMPORAL & SPATIO-TEMPORAL ANALYSIS OF QRS COMPLEX

This analysis is performed on QRS complexes extracted from Multi-Channel ECG of healthy young people, of age between 22 and 33. The respiration signal acquisition was carried out simultaneously while recording the ECG. Apart from the QRS data, two more vectors are used in the analysis of every data set, namely the heart rate vector and the respiration vector. First we shall see how these are computed.

8.1. DEFINING RESPIRATION & HEART RATE VECTORS

The *Respiration Vector* $R = [r_1, r_2 ...... r_N]$, where N represents the total number of beats in the ECG data, is computed from the acquired discrete respiration signal S(n) as follows:

$$r_p = \sum_{n=on_p}^{n=off_p} S(n) \qquad ....(8.1)$$

Where, on_p and off_p are the corresponding QRS onset and offset instants obtained from the Delineation Result Matrix (DRM).

The *Heart Rate Vector* $Hr = [h_1, h_2 ...... h_N]$ is computed from the R-R intervals (referring to DRM) as follows [39]:

$$h_p = \frac{60 \times S.F.}{DRM_{p+1,6} - DRM_{p,6}} \qquad ....(8.2)$$

S.F. denotes the sampling frequency of the ECG dataset. Here, it should be noted that the 6th column of DRM contains the R peak locations.

8.2. TEMPORAL ANALYSIS

The first step in this procedure is to carry out a Principal Component Analysis on the data matrices (D^1,.........,$D^{32/64}$: superfix denoting channel number) from different channels. The matrix D^i (i=1,2,....32/64 depending on 32/64 Channel ECG) is the transpose of PIQM (Chapter 7) for the corresponding channel. The size of D^i is M×N, where M is the QRS duration in terms of sample numbers and N is the number of heart beats, as mentioned

before. This D^i is the PCA input matrix. It should be noted that unlike the PCA data matrix definition given in Figure 4.16, here the variables (i.e. QRS) appear along the column. This modification is made to take advantage of the Matlab in-built command 'princomp'.

Similar to the PCA implementation steps mentioned under section 4.2.4, the following procedure is adopted for each and every channel (by means of 'princomp' command):

1. The mean QRS is calculated as,

$$\overline{QRS}^i = \frac{1}{N} \sum_{p=1}^{N} QRS_p^i \quad ....(8.3)$$

Here, 'i' is the index to channel no. and 'p' to the beat number, as mentioned before. Each QRS_p^i denotes a vector of length M.

2. Next the covariance matrix C^i of size N×N is computed from the zero-mean channel data:

$$C^i = \frac{1}{M-1} \sum_{p=1}^{N} \left(QRS_p^i - \overline{QRS}^i \right)^T \left(QRS_p^i - \overline{QRS}^i \right) \quad ....(8.4)$$

3. The eigenvectors and the eigenvalues of C^i are derived. The eigenvectors are then sorted in descending order of their associated eigenvalues. First k eigenvectors, namely $eig_1^i, eig_2^i, ..., eig_k^i$ are selected as the principal components. Each of eig_l^i, where $1 \le l \le k$ is a column vector of length N

4. The representation of QRS data in each of the eigenvector space is calculated as:

$$PCrep_l^i = D^i(eig_l^i) \quad ....(8.5)$$

Each $PCrep_l^i$ is a column vector of length M (i.e. the QRS vector length).

Following figure 8.1 shows the extracted (and aligned) QRS complexes from one channel (made overlapped on each other).
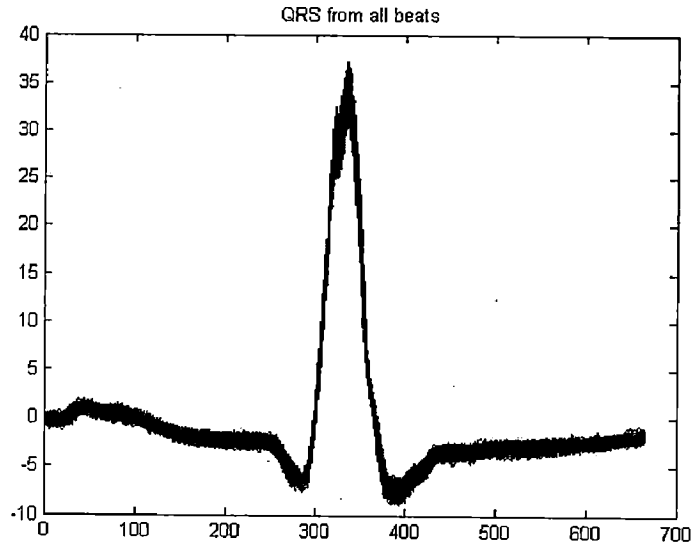
Figure 8.1: Extracted QRS complexes from one channel

Here, only the first 3 eigenvectors are selected as the principal componets, i.e. k=3. The eigenvalues associated to the rest of the eigenvectors are nearly zero. Figure 8.2 shows the representation of QRS data in each of the first three eigenvector spaces and figure 8.3 illustrates the associated eigenvalues (in terms of percentage variance).
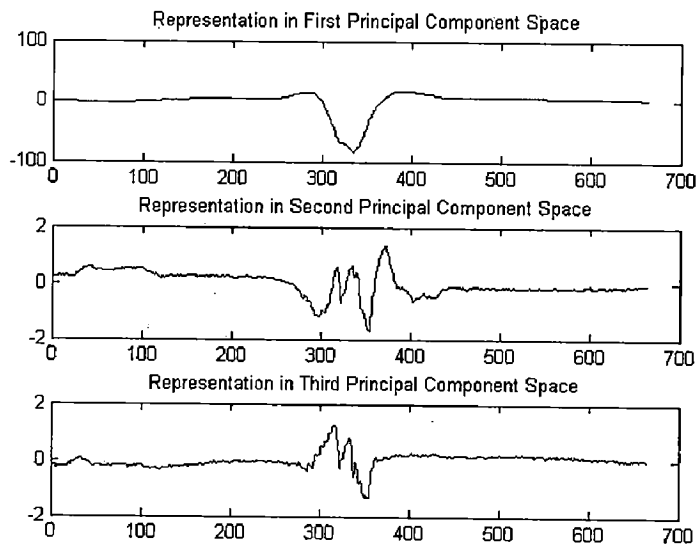

Figure 8.2: Representation of QRS data in the Principal Component Space

The change in polarity corresponding to $PCrep_i^j$ is taken care of by the reconstruction parameters to be discussed next.
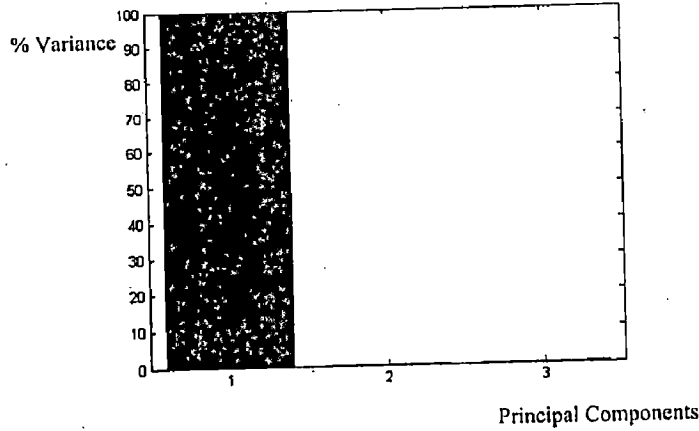
Figure 8.3: Percentage of total variance in QRS data expressed by the principal components (temporal)

5. Thereafter, the *Reconstruction Parameter Matrix* $RPM^i = \langle a_{p,l} \rangle$ for each channel is computed on the basis of selected eigenvectors as follows,

$$a_{p,l} = (PCrep_l^i)^T (QRS_p^i - \overline{QRS}^i) \quad \text{.............................(8.6)}$$

The *Reconstruction Parameter Vector* RPV^l associated with any eigenvector eig_l^i can

be given as,

$$(RPV^l)^T = [a_{1,l}, \ldots\ldots\ldots\ldots, a_{N,l}] \quad \text{.............................(8.7)}$$

RPV^l is nothing but the lth column of RPM^i.

The QRS complexes can be reconstructed in terms of the principal components as follows:

$$QRS_1^i = \overline{QRS}^i + a_{1,1} \times PCrep_1^i + a_{1,2} \times PCrep_2^i + a_{1,3} \times PCrep_3^i$$

$$QRS_2^i = \overline{QRS}^i + a_{2,1} \times PCrep_1^i + a_{2,2} \times PCrep_2^i + a_{2,3} \times PCrep_3^i$$

$$QRS_N^i = \overline{QRS}^i + a_{N,1} \times PCrep_1^i + a_{N,2} \times PCrep_2^i + a_{N,3} \times PCrep_3^i$$

The beat-to-beat variation of reconstruction parameters, heart rate and respiration for the same channel is plotted in Figure 8.4.
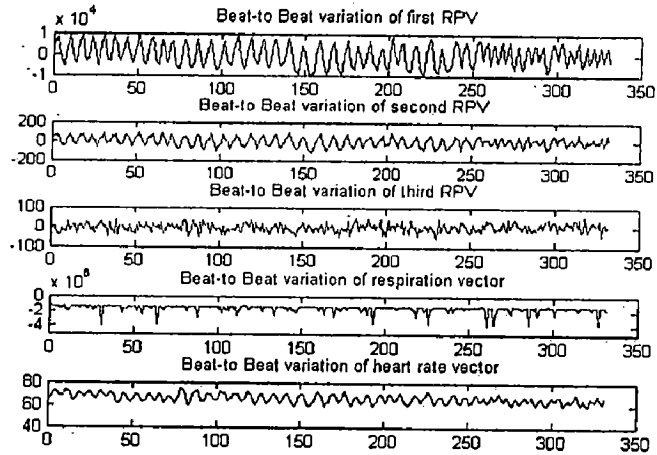
Figure 8.4: Beat to Beat Characteristics (temporal)

6. Now we calculate a covariance matrix quantifying the dependence between the reconstruction parameter, heart rate and respiration vectors as follows:

	RPV^1	RPV^2	RPV^3	R	Hr
RPV^1	2.3678×10^7	1.6059×10^5	25691	1.2586×10^8	2882.9
RPV^2	1.6059×10^5	2195.8	16.156	1.4759×10^5	84.168
RPV^3	25691	16.156	495.53	1.2467×10^5	5.4775
R	1.2586×10^8	1.4759×10^5	1.2467×10^5	1.9077×10^{11}	1.4878×10^5
Hr	2882.9	84.168	5.4775	1.4878×10^5	10.231

Table 8.1: Final Covariance Matrix (temporal)

The final covariance matrix elements, given in the table above, are neither normalised, nor zero-centred. The relative degree of covariance between each RPV and respiration and heart rate vectors are as shown in figure 8.5.
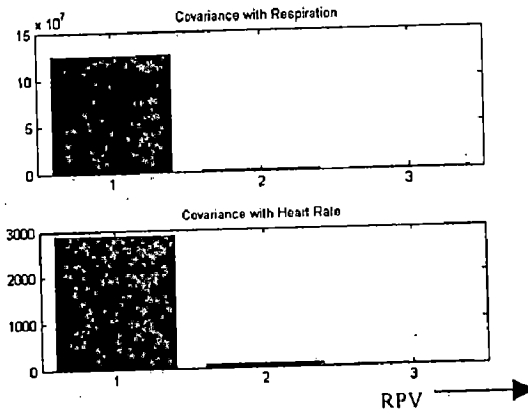
Figure 8.5: Covariance of RPVs with Respiration and Heart Rate (temporal)

8.2. SPATIO-TEMPORAL ANALYSIS

The only difference between this and the former one is the PCA input data matrix. The data matrix D for spatio-temporal analysis is a combination of all $(D^1,.........D^{32/64})$. The size of D will be M'×N, where M'(M'=M×32 or M'=M×64 depending on 32/64 channel data), is the QRS vector length over all 32/64 cannels. Schematically, the structure of D is given as. follows.

Figure 8.6: PCA Input Data Matrix for Spatio-temporal Analysis

The procedural steps (1 to 6) are exactly similar to those for temporal analysis. Spatio-temporal analysis identifies principal components in the combined spatio-temporal domain, which means considering temporal variations of all the channels pertaining to a

Multi-channel data set simultaneously. In this case, we end up with a single covariance matrix including all the channels.

Here also, the first three eigenvectors are chosen as the principal component. The percentage of total variance expressed by them is found to be similar to that for any individual channel under temporal analysis (referring to Figure 8.3).
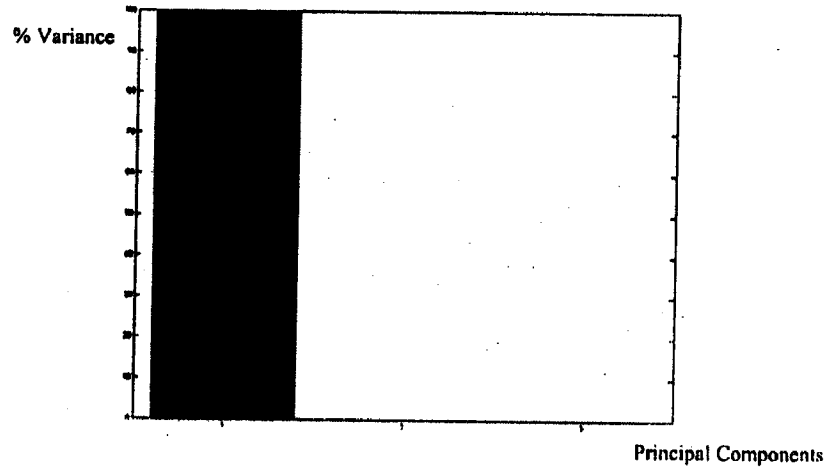

Figure 8.7: Percentage of total variance in QRS data expressed by the principal components (Spatio-temporal)

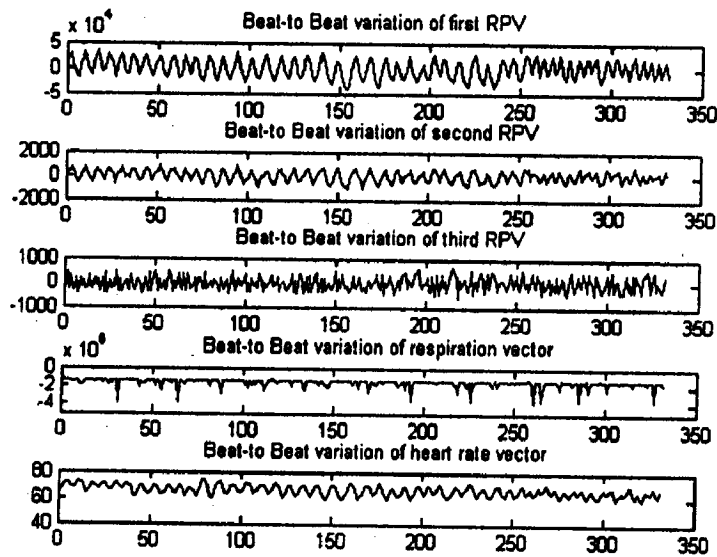The beat-to-beat characteristics for spatio-temporal analysis are shown below.

Figure 8.8: Beat to Beat Characteristics (Spatio-temporal)

The final covariance matrix for spatio-temporal analysis of the same data set is given in Table 8.2.

	RPV^1	RPV^2	RPV^3	R	Hr
RPV^1	2.373×10^{11}	3.6952×10^{10}	5.072×10^{9}	1.5987×10^{10}	4.5518×10^{10}
RPV^2	3.6952×10^{10}	2.2483×10^{10}	2.1018×10^{5}	3.0233×10^{9}	46743
RPV^3	5.072×10^{9}	2.1018×10^{5}	2.7733×10^{9}	2.1678×10^{9}	4857.8
R	1.5987×10^{10}	3.0233×10^{9}	2.1678×10^{9}	1.9077×10^{11}	1.4878×10^{5}
Hr	4.5518×10^{10}	46743	4857.8	1.4878×10^{5}	10.231

Table 8.2: Final Covariance Matrix (spatio-temporal)

Here also, the elements of final covariance matrix are not zero-centred and normalised. The relative covariance between the reconstruction parameter vectors and respiration and heart rate vectors are also found similar to those in temporal analysis (referring Figure 8.5).
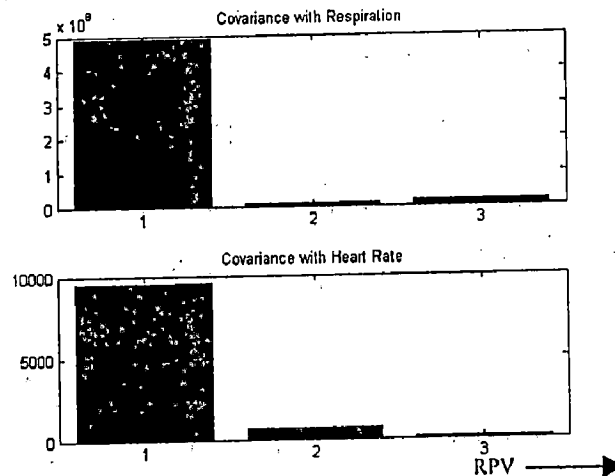

Figure 8.9: Covariance of RPVs with Respiration and Heart Rate (Spatio-temporal)

All the figures given here, for temporal and spatio-temporal analysis are pertaining to the same Multi-Channel data set.

The reconstruction parameters corresponding to the first eigenvector shows high degree of correlation (covariance) with heart rate and respiration pattern, as evident from the results of temporal and spatio-temporal analysis. Moreover, the first principal component accounts for more than 99% variability in the QRS data set. Therefore, QRS morphology is found to be highly sensitive to the heart rate variability, as well as changes in respiration pattern.

CHAPTER 9

CONCLUSIONS & FUTURE DIRECTIONS

Our analysis is based on QRS complexes. The delineation algorithm should give error-free result for the success of temporal and spatio-temporal analysis. To achieve this aim, we started with ECG conditioning. A new wavelet-based method has been proposed for the correction of baseline wander. The best prototype wavelets in this regard have been determined through simulation. For high frequency noise removal, Butterworth Low pass or Savitzky-Golay Filter was employed. However, finding out new kinds of filtering technique, which yields better results in this regard, can be an interesting area of future research. The filtering approach should be an optimum one – noise removal and ensuring minimum signal distortion at the same time.

The starting point of ECG delineation is R peak detection. Our method was tested on nine different signals (containing a total of 18,762 R peaks) from MIT Database and it yielded an overall sensitivity and positive predictivity of 99.84% and 99.89% respectively. The algorithm was applied on 30 multi-channel data sets (containing approximately 30,000 R peaks in total) recorded at the Institute. The result was found to be highly satisfactory.

However, the delineation method is based on the assumption that the R peaks cause the sharpest change in direction (highest DCS) in a ECG cycle. Whenever this condition fails, R peak detection is subjected to error. To ensure this condition, the DCS enhancement procedure using re-sampling was used sometimes. If the sampling frequency of ECG is less, this problem arises. With MIT signals sampled at the rate of 360 Hz, the use of re-sampling was essential to achieve good results in R peak detection. However, when the sampling frequency itself is more than 1KHz, no such enhancement procedure is needed. It will be interesting to find some constraints, other than re-sampling to make this delineation more robust and accurate.

In the Principal Component Analysis, we have seen that first three eigenvectors characterize the whole QRS data. The first principal component itself accounts for more than 99% variability in the QRS data set. This can be thought of as an efficient means for data compression.

The reconstruction parameters corresponding to the first three principal components, specially the first one, show high degree of correlation (covariance) with heart rate and respiration pattern. Therefore, we may conclude that the QRS morphology is highly influenced by the heart rate variability, as well as changes in respiration pattern.

For future work in this area, the same idea can be extended to discover how the ECG pattern typically changes with blood pressure. The QRS analysis, when applied to patient data (case of Ischemia, Ventricular Tachycardia etc) can be helpful to find a diagnostic pattern typical to those diseases. The correlation between QRS morphology and heart rate variability for different heart diseases may also carry useful diagnostic information.

The idea of carrying out a PCA can be extended to P and T waves as well. These represent atrial depolarisation and ventricular repolarisation respectively. For the diagnosis some diseases, P and T wave morphology plays an important role.

We hope that the slowly emerging cardiac diseases may be diagnosed in early stage with this kind of analysis.

Appendix I:

GRAPHICAL USER INTERFACE FOR ECG CONDITIONING
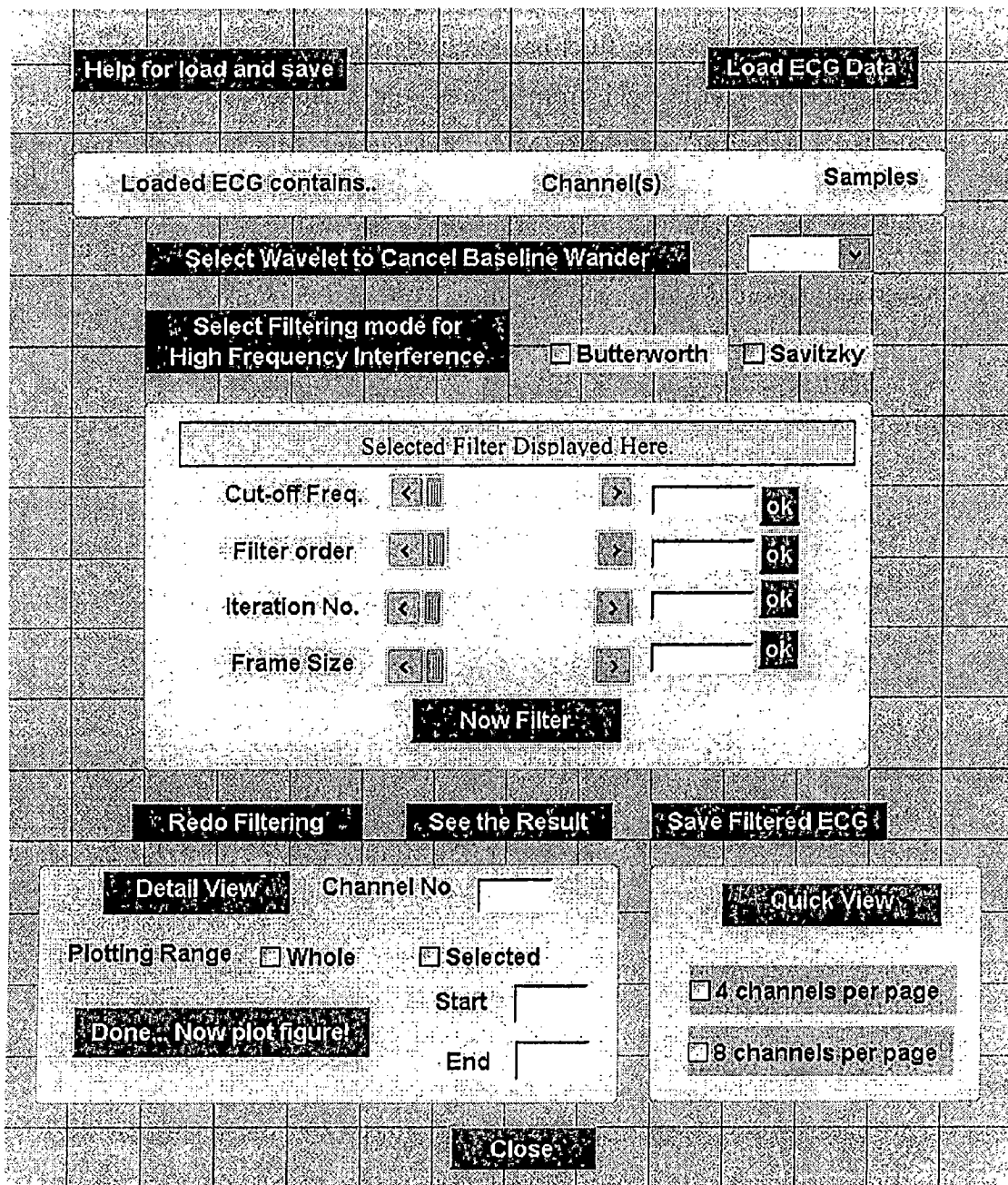
First the raw ECG data is loaded from a mat file. Immediately after that, the no. of channels and the data length (in samples) are displayed.

The first step is baseline wander cancellation. The prototype wavelet for this purpose can be selected from the pop-up menu.

The next step is high frequency interference elimination. Either Butterworth low pass or Savitzky-Golay filter can be selected. Now the selected filtering option will be displayed and the parameters for corresponding filter have to be defined by the user.

The filtering output can be seen from figures--- all channels at a time or a selected channel in selected range. There are options for redoing the filtering if the output is not satisfactory.

On pressing the save command, the filtering output (both after first stage and second stage separately) will be saved along with the filter description in a mat file.

Appendix II:

GRAPHICAL USER INTERFACE TO DECIDE CHANNELS FOR ECG DELINEATION
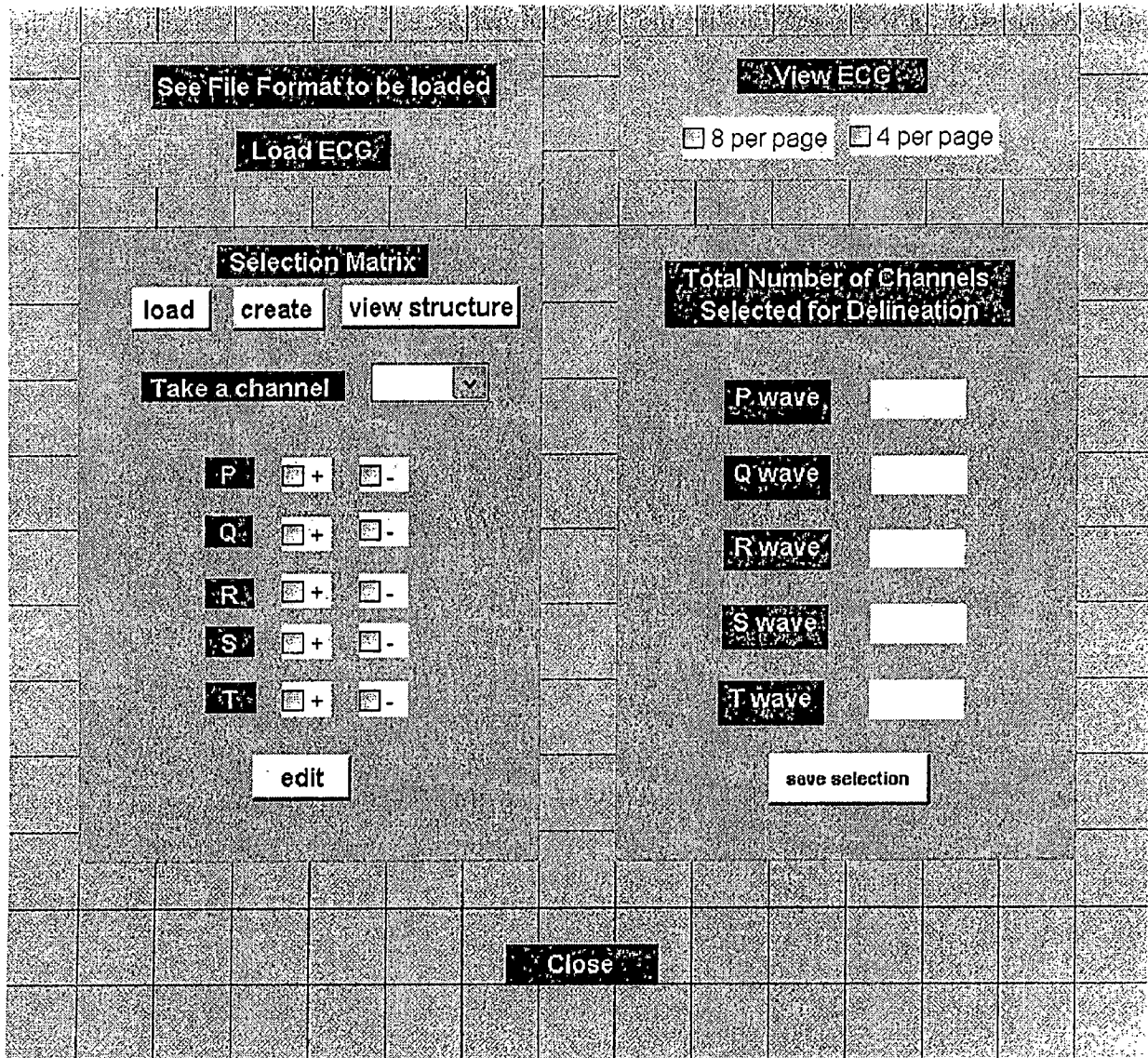
First the filtered ECG data is loaded from a mat file. The two-stage filtered ECG can be plotted on figures to aid the decision of channel selection.

First the structure of selection matrix should be viewed by pressing the corresponding button. If the selection matrix for this ECG data is already defined, the button "load" will

appear, otherwise, the button 'create' will be enabled. Now, the channels should be selected one by one and if it is found to have prominent peak (P,Q,R,S, and /or T), the corresponding one should be marked as positive or negative by the user. The "edit" pushbutton allows editing the properties for a particular channel.

The number of channels selected for delineation of any particular wave is displayed side by side to aid the user. Finally, the selection matrix can be saved in a mat file in the same folder of the loaded ECG data.

Appendix III:

GRAPHICAL USER INTERFACE FOR ECG DELINEATION
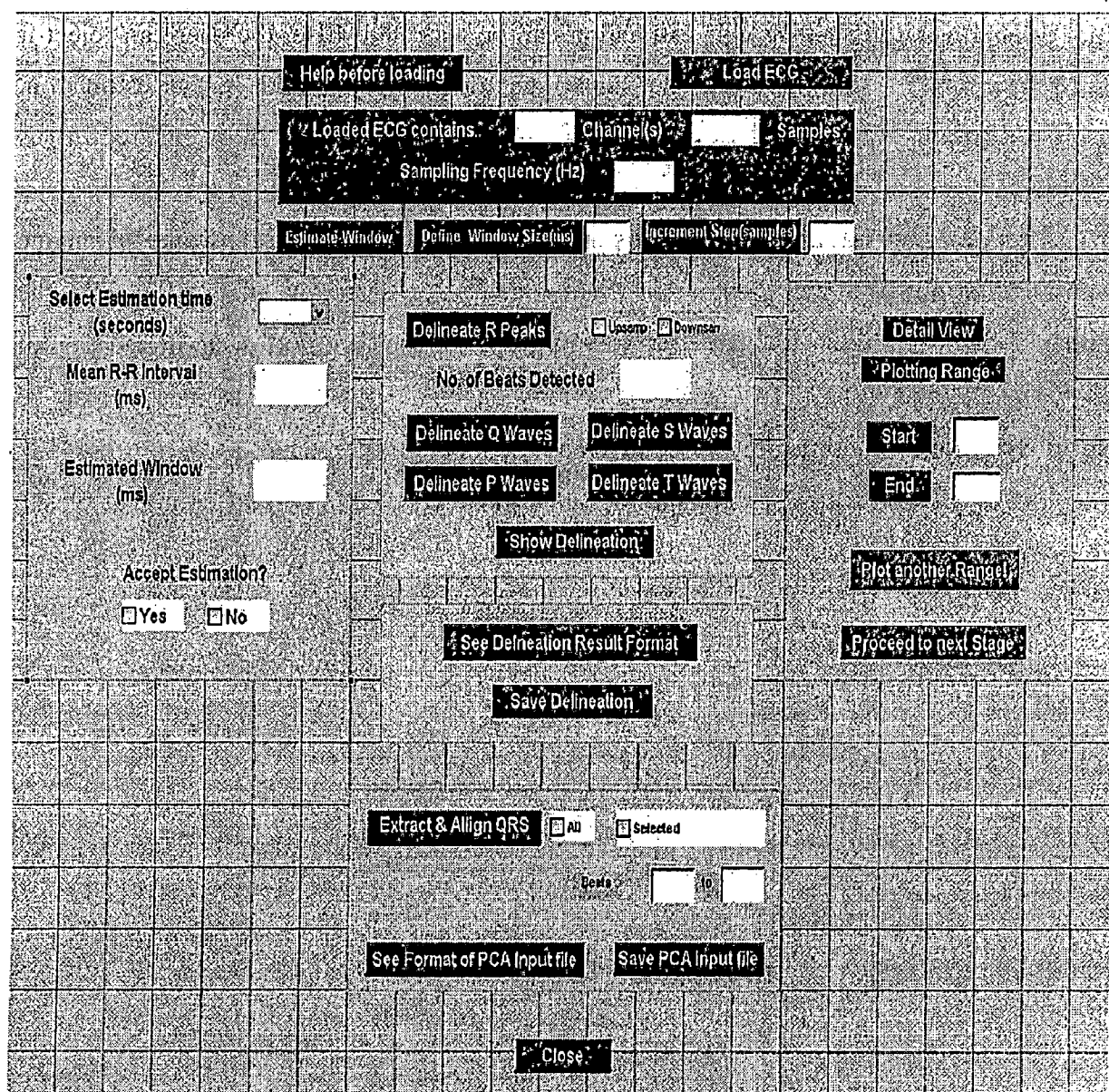
First the filtered ECG signal file is loaded. This file contains both—the first-stage filtered signal and also the two-stage filtered one. The number of channels, data length (samples) and the sampling rate is displayed after loading.

First the Window Length (WL) and Incremental Step (IS) have to be decided before proceeding to R peak delineation. IS should be chosen on the basis of displayed Sampling Frequency. There are two options to initialise WL – it can be user-defined or can be

111

estimated from initial beats of ECG. The estimation time (2-6 seconds) can be selected from the pop-up menu. 60% of the mean R-R interval resulting from the R peak detection (first few seconds, as selected by user) with defined IS and 500 ms WL is displayed as the estimated window size. Also, the detected points are plotted in a figure on the ECG. If the initial detection is satisfactory, this estimation can be accepted. Else, the user can reject it and go for estimation with over a longer time or can define WL without further estimation. During estimation, WL is always made divisible by IS. The same should be ensured if it is defined.

After deciding WL and IS, the next step is R peak detection. User may take advantage of upsampling and/or downsampling in this regard. After detecting all the R peaks, the number of beats will be displayed. Now step by step, the user can go for the delineation of Q/S wave and thereafter, P/ T wave. At every stage, the result of delineation can be displayed for the whole time range or a selected time length.

Finally, the delineation result and PCA input file are saved in two different mat files.

REFERENCES

[1]: L. Cromwell, F. J. Weibell, E. A. Pfeiffer. ,Biomedical Instrumentation and Measurements'.
Prentice-Hall, Inc.

[2]: Texas Heart Institute. 'Anatomy of the Human Heart'
http://www.tmc.edu/thi/anatomy.html

[3]: Jaakko Malmivuo, Robert Plonsey. 'Principles and Applications of Bioelectric and Biomagnetic Fields'. Oxford University Press, 1995.

[4]: R. M. Rangayyan. 'Biomedical Signal Analysis : A Case-Study Approach'. IEEE Press.

[5]: W. J. Tompkins, Editor. 'Biomedical Digital Signal Processing'. Prentice-Hall, Inc.

[6]: Antoun Khawaja. Institut für Biomedizinische Technik, Universität Karlsruhe(TH).
e-Mail: Antoun.Khawaja@ibt.uni-karlsruhe.de

[7]: Mason R.E., Likar I. 'A New System of Multiple-Lead Exercise Electrocardiography'
Am. Heart J. 1966; 71; 196-205

[8]: Lokenath Debnath. 'Wavelet Transforms and their Applications'. Birkhäuser.

[9]: F. Hlawatsch, G.F. Boudreaux-Bartels. 'Linear and Quadratic Time-Frequency Signal Representations'. IEEE SP Magazine; April 1992 ;21-67

[10]: O. Rioul, M. Vetterli. 'Wavelets and Signal Processing'. IEEE SP Magazine; October 1991; 14-38

[11]: R. Polikar. 'The Wavelet Tutorial' (four parts).
http://users.rowan.edu/~polikar/WAVELETS/WTtutorial.html

[12]: 'The Discrete Wavelet Transform'
http://etd.lib.fsu.edu/theses/available/etd-11242003-185039/unrestricted/09_ds_chapter2.pdf

[13]: Matlab Wavelet Toolbox Manual

[14]: J.P. Martinez, R. Almeida, S. Olmos, A. P. Rocha, P. Laguna. 'A Wavelet-Based ECG Delineator: Evaluation on Standard Databases'. IEEE Trans. on Biomed. Engg.; Vol. 51;
No. 4;570-581; April 2004.

[15]: C. Li, C. Zheng, C. Tai. 'Detection of ECG Characteristic Points using Wavelet Transforms'.IEEE Trans. on Biomed. Engg.; Vol. 42; No. 1; 21-28; Jan 1995.

[16]: P. Klapetek. 'Discrete Wavelet Transform'.

http://www.klapetek.cz/wdwt.html

[17]: A. Sturn. 'Cluster Analysis for Large Scale Gene Expression Studies' (48-51).
http://genome.tugraz.at/Theses/Sturn2001.pdf

[18]: J. Shlens. 'A Tutorial on Principal Component Analysis :Derivation, Discusssion and
Singular Value Decomposition'. March 25, 2003; Version 1
http://www.snl.salk.edu/~shlens/pub/notes/pca.pdf

[19]: 'Principle Component Analysis'.
http://www.ucl.ac.uk/oncology/MicroCore/HTML_resource/PCA_1.htm

[20]: 'Principal component analysis'.
http://www.fon.hum.uva.nl/praat/manual/Principal_component_analysis.html

[21]: L. I. Smith. 'A Tutorial on Principal Component Analysis'. Feb 26, 2002.
http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf

[22]: Challis R. E., Kitney R. L. 'Biomedical Signal Processing (in four parts) ; Part 1:
Time-
domain Methods'. Med. Biol. Eng. Comput., Nov 1990 ; 28(6) ; 509-24 ; PMID
2287173

[23]: N.A. Anstey. ' Correlation Techniques --- A review'.
http://www.cseg.ca/publications/journal/1966_12/1966_Anstey_N_corelation_techniques.pdf

[24] : P. M. Agante, J. P. Marques de Sa. 'ECG Noise Filtering Using Wavelets with Soft-
thresholding Methods'. Computers in Cardiology 1999 ; 26; 535-538

[25]: Meyer C. R., Keiser H. N. 'Electrocardiogram baseline noise estimation and removal
using
cubic splines and state-space computation techniques'.Computers and Biomedical
Research
; Vol 10; 1977 ; 459-470.

[26]: V. Alste J. A., Schilder T. S.'Removal of baseline wander and power-line interference
from
the ECG by an efficient FIR filter with a reduced number of taps'. IEEE Trans.
Biomed. Eng.; Vol. BME-32; No 12; 1052-1060; 1985

[27] Sörnmo L. ' Time-varying filtering for removal of baseline wander in exercise ECGs'.
Computers in Cardiology ; IEEE Computer Society Press; 145-148; 1991

[28] N. V. Thakor, Y. Zhu , 'Applications of adaptive filtering to ECG analysis: noise
cancellation and arrhythmia detection'. IEEE Trans. On Biomed. Engg.; Vol. 38; No 8;
Aug 1991

[29] S. V. Pandit. 'ECG Baseline Drift Removal through STFT'. 18th Annual Conf. of IEEE

Eng. in Med. and Biol. Society;1405-1406; 1996

[30]: P. Laguna, R. Jane, P. Caminal.' Adaptive Filtering of ECG Baseline Wander'. IEEE

[31]: Johnny R. Johnson. 'Introduction to Digital Signal Processing'.Prentice-Hall, Inc.

[32]: Matlab Curve Fitting Toolbox Manual

[33]: Matlab Signal Processing Toolbox Manual

[34]: Bert-Uwe Köhler, C. Hennig, R. Orglmeister. 'The Principles of Software QRS Detection'.
     IEEE Eng. in Med. and Biol.; 42-49; Jan/Feb 2002

[35]: S. Kadambe, R. Murray, G. Faye Boudreaux-Bartels. 'Wavelet Transform-Based QRS
     Complex Detector'. IEEE Trans. Biomed. Eng. ; Vol. 46 ; No 7; July 1999 ; 838-848

[36]: A. Gutierrez, P.R. Hernandez, M. Lara, S. J. Perez. ' A QRS Detection Algorithm Based on
     Haar Wavelet'. Computers in Cardiology 1998; Vol 25 ; 353-356

[37]: E. Laciar, R. Jane, Dana H. Brooks. 'Improved Alignment Method for Noisy High-
     Resolution ECG and Holter Records using Multiscale Cross-Correlation'. IEEE Trans.
     Biomed. Eng.; Vol. 50 ; No 3; March 2003 ; 344-353

[38]: R. Jane, H. Rix, P. Caminal. 'Alignment Methods for Averaging High-Resolution Cardiac
     Signals : A Comparative Study Performance'. IEEE Trans. Biomed. Eng. ; Vol. 38 ;
     No 6; June 1991 ; 571-579

[39]: Electrocardiogram, (ECG)-I
     http://www.cs.wright.edu/~phe/EGR199/Lab_1/