Detection Of The R-wave In Ecg Signals

2005

Sasanka Valluri

University of Central Florida

Find similar works at: http://stars.library.ucf.edu/etd

University of Central Florida Libraries http://library.ucf.edu

Part of the Electrical and Electronics Commons

STARS Citation

http://stars.library.ucf.edu/etd/407

This Masters Thesis (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of STARS. For more information, please contact lee.dotson@ucf.edu.
DETECTION OF THE R-WAVE IN ECG SIGNALS

by

SASANKA VALLURI

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science in Electrical Engineering
in the Department of Electrical and Computer Engineering
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Spring Term
2005
ABSTRACT

This thesis aims at providing a new approach for detecting R-waves in the ECG signal and generating the corresponding R-wave impulses with the delay between the original R-waves and the R-wave impulses being lesser than 100 ms. The algorithm was implemented in Matlab and tested with good results against 90 different ECG recordings from the MIT-BIH database [1].

The Discrete Wavelet Transform (DWT) forms the heart of the algorithm providing a multi-resolution analysis of the ECG signal. The wavelet transform decomposes the ECG signal into frequency scales where the ECG characteristic waveforms are indicated by zero crossings. The adaptive threshold algorithms discussed in this thesis search for valid zero crossings which characterize the R-waves and also remove the Pre-ventricular Contractions (PVC's). The adaptive threshold algorithms allow the decision thresholds to adjust for signal quality changes and eliminate the need for manual adjustments when changing from patient to patient.

The delay between the R-waves in the original ECG signal and the R-wave impulses obtained from the algorithm was found to be less than 100 ms.
ACKNOWLEDGMENTS

I would like to express my sincere appreciation to Dr. Arthur R. Weeks for his valuable assistance and ideas. I would also like to thank my parents and Sujatha Chiluvuri for their support.
# TABLE OF CONTENTS

ABSTRACT ........................................................................................................................................ iii

LIST OF FIGURES ................................................................................................................................... viii

LIST OF TABLES ................................................................................................................................... xi

CHAPTER 1: INTRODUCTION .................................................................................................................. 1

1.1. Overview (Heart Disease and Stroke) ......................................................................................... 1

1.2. Introduction to the Heart and its Functioning ........................................................................... 2

1.2.1. Structure of the Heart ............................................................................................................. 3

1.3. Electrical System of the Heart ................................................................................................... 8

1.4. ECG Waveform and its Components ......................................................................................... 10

1.4.1. ECG Signal Characteristics ................................................................................................. 12

1.5. Arrhythmia ..................................................................................................................................... 12

1.6. Previous Research On the Detection of the ECG Signal Characteristics ......................... 15

1.7. Work Done In This Thesis ......................................................................................................... 17

CHAPTER 2: WAVELET THEORY ........................................................................................................... 18

2.1. Fundamental Concepts ............................................................................................................... 18

2.1.1. Fourier Transforms ............................................................................................................... 19

2.1.2. Discrete Fourier Transform ................................................................................................ 21

2.1.3. Fast Fourier Transform (FFT) ............................................................................................ 22

2.1.4. Short Time Fourier Transform (STFT) .............................................................................. 25

2.2. Wavelet Analysis ......................................................................................................................... 26

2.2.1. The Continuous Wavelet Transform and Wavelet Series ................................................. 27

2.2.2. The Discrete Wavelet Transform ......................................................................................... 28
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Properties and Types of Wavelets</td>
<td>29</td>
</tr>
<tr>
<td>Sub-band Coding and Multi-resolution Analysis</td>
<td>30</td>
</tr>
<tr>
<td>2.3. Equivalent Wavelet Filter</td>
<td>33</td>
</tr>
<tr>
<td>2.4. Wavelet Transform Maxima</td>
<td>34</td>
</tr>
<tr>
<td>2.5. Quadratic Spline Wavelet</td>
<td>35</td>
</tr>
<tr>
<td><strong>CHAPTER 3: PROCESSING OF ECG SIGNALS</strong></td>
<td>44</td>
</tr>
<tr>
<td>3.1. ECG Analysis</td>
<td>44</td>
</tr>
<tr>
<td>3.1.1. Discussions and Goals</td>
<td>44</td>
</tr>
<tr>
<td>ECG signal and all the wavelet plots</td>
<td>45</td>
</tr>
<tr>
<td>3.2. Detection of the R-waves</td>
<td>53</td>
</tr>
<tr>
<td>3.3. Detection of the PVC Output</td>
<td>56</td>
</tr>
<tr>
<td>3.4. The Adaptive Threshold Algorithms</td>
<td>57</td>
</tr>
<tr>
<td>3.4.1. QRS Thresholds of the 1\textsuperscript{st} Wavelet Signal</td>
<td>58</td>
</tr>
<tr>
<td>Example</td>
<td>61</td>
</tr>
<tr>
<td>3.4.2. PVC Detection Based on the 1\textsuperscript{st} Wavelet Thresholds</td>
<td>64</td>
</tr>
<tr>
<td><strong>CHAPTER 4: STATISTICS</strong></td>
<td>73</td>
</tr>
<tr>
<td>4.1. Validation database</td>
<td>73</td>
</tr>
<tr>
<td>4.2. Example</td>
<td>73</td>
</tr>
<tr>
<td>4.3 Statistical Results</td>
<td>74</td>
</tr>
<tr>
<td>4.4 Problems Faced During R-Wave and PVC Detection</td>
<td>77</td>
</tr>
<tr>
<td><strong>CHAPTER 5: CONCLUSIONS</strong></td>
<td>82</td>
</tr>
<tr>
<td>5.1. Problems in ECG Processing</td>
<td>82</td>
</tr>
<tr>
<td>5.2. Future Work</td>
<td>82</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>5.3. Summary Of What Has Been Done In This Thesis</td>
<td>83</td>
</tr>
<tr>
<td>APPENDIX A</td>
<td>85</td>
</tr>
<tr>
<td>APPENDIX B</td>
<td>88</td>
</tr>
<tr>
<td>MATLAB CODE</td>
<td>88</td>
</tr>
<tr>
<td>LIST OF REFERENCES</td>
<td>104</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1 Location of the Heart ........................................................................................................... 2
Figure 2 Structure of the Heart ........................................................................................................... 4
Figure 3 Valves and Chambers of the Heart ...................................................................................... 5
Figure 4 Blood Flow .......................................................................................................................... 6
Figure 5 Blood Flow .......................................................................................................................... 7
Figure 6 Electrical System of the Heart ............................................................................................. 9
Figure 7 ECG Waveform ................................................................................................................... 10
Figure 8 Arrhythmia .......................................................................................................................... 13
Figure 9 Time Domain Representation of x (t) .................................................................................. 20
Figure 10 Frequency Domain Representation of x (t) .................................................................... 21
Figure 11 Chirp Signal ...................................................................................................................... 23
Figure 12 Time Domain Representation of a Signal ........................................................................ 24
Figure 13 Frequency Domain Representation of the Signal ............................................................ 25
Figure 14 Wave ................................................................................................................................. 27
Figure 15 Wavelet .............................................................................................................................. 27
Figure 16 Different families of Wavelets ......................................................................................... 30
Figure 17 Three level Wavelet decomposition tree ....................................................................... 32
Figure 18 Three Level Wavelet Reconstruction tree .................................................................... 32
Figure 19 Filter Banks, Modified Mallat's algorithm ..................................................................... 34
Figure 20 Mother Wavelet of the Quadratic Spline Wavelet ........................................ 36
Figure 21 Frequency Spectrum of the Q Filters............................................................ 37
Figure 22 ECG Signal from the MIT-BIH Database [1] ................................................ 38
Figure 23 First Wavelet of the ECG Signal from the MIT-BIH Database [1] ............... 39
Figure 24 Second Wavelet of the ECG Signal from the MIT-BIH Database [1] ........... 40
Figure 25 Third Wavelet of the ECG Signal from the MIT-BIH Database [1] .......... 41
Figure 26 Fourth Wavelet of the ECG Signal from the MIT-BIH Database [1] ............ 42
Figure 27 Fifth Wavelet of the ECG Signal from the MIT-BIH Database [1] ............... 43
Figure 28 Quadratic Spline Mother Wavelet................................................................. 46
Figure 29 Frequency Response of the Q filters ............................................................ 47
Figure 30 ECG signal from the MIT-BIH Database [1] ................................................ 48
Figure 31 First Wavelet of the ECG Signal from the MIT-BIH Database [1] ............. 49
Figure 32 Second Wavelet of the ECG Signal from the MIT-BIH Database [1] .......... 50
Figure 33 Third Wavelet of the ECG Signal from the MIT-BIH Database [1] ............ 51
Figure 34 Fourth Wavelet of the ECG Signal from the MIT-BIH Database [1] ............ 52
Figure 35 Fifth Wavelet of the ECG Signal from the MIT-BIH Database [1] ............... 53
Figure 36 First Wavelet of the ECG Signal before the QRS threshold algorithm ...... 54
Figure 37 First Wavelet of the ECG Signal after the QRS threshold algorithm ......... 55
Figure 38 ECG signal from the MIT-BIH Database [1] ................................................ 56
Figure 39 First Wavelet of the ECG Signal from the MIT-BIH Database ................. 57
Figure 40 Flow Chart for the QRS threshold algorithm ............................................. 59
Figure 41 ECG Signal from the MIT-BIH Database [1] ................................................ 61
Figure 42 First Wavelet of the ECG Signal from the MIT-BIH Database ................. 62
Figure 43 First Wavelet Peaks before the QRS detection .......................................... 63
Figure 44 First Wavelet Peaks after the QRS detection .................................................. 64
Figure 46 ECG Signal from the MIT-BIH Database [1] .................................................. 67
Figure 47 Finding the Max Values using sliding windows ............................................. 68
Figure 48 Example of ECG signal from the MIT-BIH Database [1]. ......................... 73
Figure 49 ECG signal from the MIT-BIH Database [1] ............................................... 78
Figure 50 ECG signal from the MIT-BIH Database [1] ............................................... 78
Figure 51 ECG signal from the MIT-BIH Database [1] ............................................... 79
Figure 52 ECG signal from the MIT-BIH Database [1] ............................................... 80
Figure 53 ECG signal from the MIT-BIH Database [1] ............................................... 80
Figure 54 ECG signal from the MIT-BIH Database [1] ............................................... 81
LIST OF TABLES

Table 1 Time characteristics of a normal ECG signal .................................................. 12

Table 2 Example of the PVC detection in a Maxn array .................................................. 70

Table 3 Example of the PVC detection in a Maxn array .................................................. 71

Table 4 Test Results from R-Wave and PVC Detection for 2000 samples ....................... 75
CHAPTER 1

INTRODUCTION

1.1. Overview (Heart Disease and Stroke)

Cardiovascular disease (CVD), including heart disease and stroke, accounts for around 17 million deaths each year. It is the leading cause of death in the US with almost 2,000 Americans dying each day i.e. 1 death every 43 seconds. Heart attacks occur when the blood flow is blocked owing to the presence of a blood clot while strokes are a result of blocked or burst blood vessels in the brain. Congenital heart defects and a range of other conditions, which occur due to improper pumping of blood cause long term problems, and even death for sufferers.

Diagnosis of a possible heart disease can be performed by several tests. The choice of the tests is based on the patient's risk factors, history of heart problems and current symptoms. One of the basic noninvasive tests is the Electrocardiogram (ECG or EKG) which is used to assess the heart rate and rhythm.

The ECG can be used to detect heart disease, heart attack, abnormal heart rhythms and an enlarged heart condition that may cause heart failure. The electrical activity of the heart is monitored by placing electrical wires with adhesive ends on the arms, chest and legs of the patient and these readings are simultaneously recorded on graph paper.

The ECG signal is analyzed in a standardized sequence of steps to avoid missing the subtle abnormalities in the ECG tracing. After the analysis, the ECG is either interpreted as "Normal" or "Abnormal". Occasionally, the term "Borderline" is used in
case the significance of certain findings is not clear enough. Based on the ECG report, further medication or medical procedures are followed.

1.2. Introduction to the Heart and its Functioning

The heart is a hollow, four chambered, cone shaped muscular organ (as shown in Fig.1) found behind the sternum (breastbone) and between the lungs. It is located such that 2/3 of it is to the left of the midline of the body while 1/3 is to the right. The heart is roughly the size of a human fist, 5 inches (12 cm) long, 3.5 inches (8-9 cm) wide, 2.5 inches (6 cm) from front to back and weighs less than 0.5% of the total body weight.

![Figure 1 Location of the Heart](image)

The wall of the heart is made up of three layers, pericardium, myocardium and endocardium. The pericardium and endocardium are the thin protective outer and inner layers respectively while the myocardium is the thick muscular layer that provides the heart with the strength to function as a pump.
1.2.1. Structure of the Heart

The heart is divided into four chambers or compartments as shown in Fig 2, where each upper chamber is called an atrium and each lower chamber is called a ventricle. The atria are thin walled structures serving as the collecting points for the blood returning from the rest of the body. The ventricles are thick walled due to the presence of a great number of muscles that are required to pump blood to the lungs and the rest of the body.

1. Right Atrium (RA): The right atrium receives venous blood from the rest of the body through the superior and inferior vena cava. The right atrium is highly distensible in order to accommodate the venous return and hence, maintains a low pressure (0-3 mmHg). The actual pressure within the right atrium depends upon the volume of blood within the atrium and the compliance of the atrium. Blood from the right atrium flows into the right ventricle.

2. Right Ventricle (RV): The right ventricle pumps out deoxygenated blood from the right atrium to the lungs through the pulmonary artery. The waste products such as carbon dioxide are carried by the blood from the right ventricle to the lungs for oxygenation (refreshment with oxygen). The refreshed blood returns from the lungs through the pulmonary veins into the left atrium of the heart.
3. **Left Atrium (LA):** Oxygenated blood enters the heart from the lungs into the Left atrium. Although the left atrium is smaller in size, it has thicker walls when compared to the right atrium. The pulmonary veins, which serve as a passage for the blood from the lungs into the heart, are the only veins that carry oxygenated blood in the whole body. The left atrium is less compliant when compared to the right atrium resulting in a higher atrial pressure (6-10 mmHg compared to 0-3 mmHg). The blood from the left atrium flows into the left ventricle.

4. **Left Ventricle (LV):** The left ventricle pumps out the oxygenated blood it receives from the left atrium into the body. It is smaller in size and has thicker walls when
compared to the right ventricle. The aorta, the largest artery in the body, passes the refreshed blood from the left ventricle into the rest of the body.

Each chamber has a one-way valve at its exit as shown in the Fig 3, so as to prevent the back flow of blood. As each chamber contracts the valve at its exit opens so as to allow the flow of blood and closes after the completion of the contraction.

Figure 3 Valves and Chambers of the Heart

1. Tricuspid Valve: Connects the Right Atrium to the Right Ventricle.
2. Pulmonary Valve: Connects the Right Ventricle to the Pulmonary artery.
3. Mitral Valve: Connects the Left Atrium to the Left Ventricle.
4. Aortic Valve: Connects the Left Ventricle to the Aorta.
Blood is pumped out of the heart when the heart muscles contract or the heart beats (called the systole). This contraction takes place in two stages as illustrated in Fig 4.

**Figure 4 Blood Flow**

- The right and left atria contract simultaneously pumping blood into the right and left ventricles respectively.
- The right and left ventricles contract concurrently pumping blood into the lungs and the body respectively.

The flow of blood is indicated by the flow graph in Fig 5.
The heart is filled with blood when the heart muscles relax in between heart beats and this process of relaxation is called the diastole. Thus, it can be summarized that the right side of the heart collects the deoxygenated blood from the body and pumps it to the lungs for oxygenation and release of waste products like carbon dioxide. The left side of the heart collects the oxygenated blood from the lungs and pumps it to the body such that all the cells in the body receive the oxygen supply they require for proper functioning.
1.3. Electrical System of the Heart

The heart contains a special group of cells called the 'Pacemaker cells' which have the ability to generate electrical activity on their own. Electricity is produced when these cells change their electric charge from positive to negative and back. The first electric wave is initiated at the top of the heart in a heart beat and due to the inherent property of the heart muscle cells to propagate electric charge to adjacent muscle cells; the initial electric wave is enough to trigger a chain reaction. Specialized fibers in the heart conduct the electrical impulse from the pacemaker to the rest of the heart.

The three important parts of the heart's electrical system as shown in Fig 6 are:

1. The SA node (Sinoatrial Node) - It is the hearts natural pacemaker which initiates each heart beat.
2. The AV node (Atrioventricular Node) - It acts as a bridge between the atria and the ventricles, allowing electrical signals from the atria into the ventricles.
3. His-Purkinje System - It carries the electrical signal throughout the ventricles and consists of the following essential parts:
   - His Bundle
   - Right Bundle Branch
   - Left Bundle Branch
   - Purkinje Fibers
Figure 6 Electrical System of the Heart

The electrical impulse from the Sinoatrial (SA) node travels to the right and left atria, resulting in a contraction. There is a delay introduced due to the resistance offered by the muscle cells. During this delay, the atria contract and the ventricles fill up with blood. Having traveled to the Atrioventricular Node, the impulse now reaches the His Bundle and then divides into the Right and Left Bundle Branches. It then spreads to the Purkinje Fibers and later to the muscles of the Right and Left Ventricle causing them to contract at the same time.
1.4. ECG Waveform and its Components

An Electrocardiogram is a recording of the electrical activity on the body surface generated by the heart. Currents flowing in the tissues around the heart cause the Electrocardiogram signals. The ECG waveform has several hills and valleys namely P, Q, R, S, T, U as shown in Fig 7.

The electrical cycle of the heart (cardiac cycle) starts with the 'resting phase', the period of time for which the heart is devoid of any electrical activity. The second phase of the cardiac cycle is the depolarization in the heart is initiated by the pacemaker cells found below the opening of the Superior Vena Cava. These cells collectively form the Sinoatrial Node (SA).

![Figure 7 ECG Waveform](image-url)
The electrical impulse propagates through the specialized cells and the cardiac cells (muscle tissue). Although the electrical discharge propagates faster through the specialized nerve tissue than the muscle tissue, these specialized cells possess the ability to reduce the speed of the electrical transmission.

The electrical impulse travels in the form of a conduction wave moving down towards the left, through both the atria and in effect depolarizing each cell. This propagation of charge is indicated by the **P wave** on the Electrocardiograph (ECG). This wave of conduction traveling through the atria meets the Atrioventricular Node located near the centre of the heart, above the interventricular septum. The Atrioventricular Node primarily delays the conduction of the electrical impulse from the atria to the ventricles. The absence of depolarization voltage (due to the undersized Atrioventricular Node) results in an isoelectric **PR segment** on the Electrocardiograph.

Depolarization in the form of a conduction wave propagates through the ventricles, down the septum reaching the His Bundle. Splitting into the right and left bundles, the electrical impulse travels to the Purkinje fibers and thus depolarizing the myocardial cells of the ventricles. Depolarization traveling from left to right results in a small negative deflection in the Electrocardiograph called the **Q wave**.

As the wave moves down into the ventricles, depolarization takes place from the endocardium to the epicardium indicated by the **R wave** on the Electrocardiograph. The direction of polarization of the ventricular muscle below the atrioventricular groove results in an **S wave**.

An **ST segment** now results corresponding to the action potential level of all the fibers. The **T wave** signals the return of the membrane potential to its baseline i.e.
polarization. The polarization of the His Bundle causes a small positive deflection called the U wave.

1.4.1. ECG Signal Characteristics

The regular Sinus rhythm has certain characteristics which can be observed in the ECG waveform. The small rounded P wave is followed by the large QRS complex made up of straight lines forming sharp waveforms and the T wave. This sequence repeats itself for every heart beat.

The analysis of the ECG signal involves observing the waveform for any abnormalities such as changes in the QRS complexes or the length of the ST segment or any discrepancies such as missing P waves or T waves. The time characteristics of the ECG waveform are also to be observed carefully. Listed below are the time characteristics of a normal ECG signal.

<table>
<thead>
<tr>
<th>Table 1 Time characteristics of a normal ECG signal</th>
</tr>
</thead>
<tbody>
<tr>
<td>P-R Interval</td>
</tr>
<tr>
<td>QRS complex Duration</td>
</tr>
<tr>
<td>Q-T Interval</td>
</tr>
<tr>
<td>Heart Rate</td>
</tr>
</tbody>
</table>

1.5. Arrhythmia

Arrhythmias or dysrhythmias are abnormal rhythms of the heart i.e. the heart may seem to miss a beat or beat irregularly or beat very fast or very slowly, thus causing the heart to pump less effectively. Normally the heart contracts 60 to 100 times per minute with each contraction representing a heart beat. Arrhythmias occur when
◆ The heart's natural pacemaker develops an abnormal rate (rhythm).

◆ The normal flow of conduction is blocked.

◆ Another part of the heart acts as the pacemaker.

It could also be caused by stress, caffeine, tobacco, alcohol and cough and cold medicines.

![Arrhythmia Diagram]

**Figure 8 Arrhythmia**

Arrhythmias are classified based on their origin in the heart (shown in Fig 8) and the effect they have on the heart's rhythm.

◆ Atrial or Supraventricular Arrhythmia: Abnormal rhythm arises in the atria. Some of the atrial arrhythmias are listed below.
  - Sinus Arrhythmia: Cyclic changes in the heart rate during breathing.
  - Sinus Tachycardia: The sinus node sends out electrical impulses faster than the usual, thus increasing the heart rate.
- **Sick Sinus Syndrome**: Improper firing of the electrical impulse leads to oscillation between a slow rate (bradycardia) and a fast rate (tachycardia).

- **Atrial Flutter**: The muscles contract quickly due to the rapidly fired signals leading to a fast heart beat.

- **Atrial Fibrillation**: Electrical signals in the atria are fired in a fast uncontrolled manner causing an irregular heart beat.

- **Ventricular Arrhythmia**: Abnormal heart rhythm originates in the ventricles and these arrhythmias are the most serious.

  - **Premature Ventricular Complexes (PVC)**: An electrical signal from the ventricles causes an early heart beat and the heart seems to pause for the next beat of the ventricle.

  - **Ventricular tachycardia**: The heart beats due to electrical signals arising from the ventricles rather than the atria.

Arrhythmias are detected by studying the Electrocardiogram for any changes in the normal rhythm. Arrhythmias are treated through

- **Drugs**: Several drugs are used based on the recommendations of the doctor and the patient's condition.

- **Cardioversion**: Electrical shock is applied to the chest wall to restore the heart rhythm.

- **Automatic implantable defibrillators**: It is surgically placed in the patient's heart, where it monitors the heart rhythm for any arrhythmia and corrects it with an electric shock.
Artificial pacemaker: It acts as the pacemaker of the heart if the natural pacemaker of the heart is dysfunctional.

Surgery: If all the above mentioned corrective procedures have no effect, surgery is performed to remove or alter the heart tissue causing arrhythmia.

1.6. Previous Research On the Detection of the ECG Signal Characteristics

Most of the research in ECG detection has been aimed at detecting the QRS complex. Gray M. Friesen [4] compares the noise sensitivity of nine QRS detector algorithms.

Many of the simple algorithms are based on the first order or second order derivative of the ECG signal. Most of the algorithms use a pre-filter to remove power line interference, and a band pass filter to remove the high frequency noise, unwanted muscular noise and other low frequency disturbances.

The algorithm which was developed based on the first derivative of the ECG signal. The first derivative (Equation 1.1) is calculated at each point of the ECG, using a formula specified by Menrad [5].

$$\begin{align*}
y(n) &= -2x(n-2) - x(n-1) + x(n+1) + 2x(n+2) \\
n &> 2
\end{align*}$$

The slope threshold is given by Equation 1.2

$$\text{Slope}_{TH} = 0.7 \text{Max}(y(n))$$

The first derivative array is searched and if the condition $y(i) > \text{Slope}_{TH}$
The QRS detector developed by Moriet Mahoudeaux [6] takes into consideration the amplitude of the ECG signal in the QRS detection. The algorithm first computes the amplitude threshold of the ECG signal using Equation 1.3

$$\text{Amplitude}_{TH} = 0.3\text{Max}(x(n)) \quad n > 0$$  \hspace{1cm} (1.3)

The first derivative \(y(n)\) is then computed as shown in Equation 1.4

$$y(n) = x(n + 1) - x(n - 1) \quad n > 1$$  \hspace{1cm} (1.4)

The three criteria that are to be met for QRS detection are as follows:

Three consecutive points in the derivative array exceed a threshold value of 0.5 (positive slope criterion) and then followed by two consecutive points with a negative slope set to -0.3.

The criterion for amplitude is as in Equation 1.5

$$x(i), x(i + 1), x(i + 2), \ldots, x(j + 1) \geq \text{Amplitude}_{TH}$$  \hspace{1cm} (1.5)

The other technique used was to pass the ECG signal through two filters. One linear filter usually a derivative operator is used to separate the QRS complex form from the P wave and T wave followed by a nonlinear squaring operator to enhance the QRS complexes. The major disadvantage of this detector is its sensitivity to noise.

The matched filter combined with the threshold detector, Antti Ruha [7], is used to cleanup noisy ECG signals. Values greater than the threshold limit are classified as QRS candidates. The impulse response of a digital matched filter is the time reversed replica of the signal to be detected.

Wavelet Transforms are now widely used to analyze the ECG characteristics. They are employed for the time and frequency analysis of the ECG signal. Decomposing
the signal into elementary building block that are localized in both time and frequency, wavelets distinguish the ECG signal from noise, artifacts and baseline drift.

1.7. Work Done In This Thesis

An algorithm is developed to detect R-waves as well as abnormal waveforms such as PVC waves (Pre ventricular contraction) in the ECG signal. The filter used to detect the QRS complexes, PVCs is based on the wavelet transform which is implemented as a filter bank with 5 outputs. The wavelet transform decomposes the ECG signal into a set of frequency bands.

The detection of PVC waves is based on wavelets. The PVC waves usually are wide and bizarre in shape when compared to the QRS complexes.

The wavelet filter bank outputs the wavelet coefficients which are used to detect the QRS complexes. To distinguish between “true” R-waves and PVCs, an adaptive threshold is implemented with a value greater than that of R-waves and less than the value of PVCs.

After identifying the PVCs, they are eliminated with other aberrations in the signal in order to produce a R-wave triggering signal. Finally, the Annotations of the R-waves and the indices of the trigger signal are compared and the delay between them is computed. The delay for most of the signals was found to be less than 10 ms.
CHAPTER 2

WAVELET THEORY

2.1. Fundamental Concepts

According to Fourier theory, any periodic waveform can be decomposed into its constituent sinusoids such that when combined they perfectly reconstruct the original waveform. Hence, a signal can be expressed as the sum of several possibly infinite series of sines and cosines (Equation 2.1).

\[ f(x) = \frac{1}{2} a_0 + \sum_{n=-\infty}^{\infty} a_n \cos(nx) + \sum_{n=-\infty}^{\infty} b_n \sin(nx) \]  

(2.1)

Fourier theory helps visualize a signal in two domains, Fourier domain and the Space domain. Space domain refers to the representation of the signal in the time domain and spatial domain while Fourier domain is the representation of the signal's amplitude, frequency components (spectral components indicating the change in the rate of something), direction and phase.

Thus, a transform is an alternate form of representing a signal in a domain providing for the better analysis of the signal. In the case of an ECG signal, any deviation form the normal arising due to a pathological condition can be better detected by analyzing the frequency domain than the time domain.
2.1.1. Fourier Transforms

A periodic wave can be represented as the sum of sines and cosines using the Fourier series representation while a non-periodic continuous signal can be expressed in terms of complex exponentials of different frequencies using the Fourier Transform (Equation 2.2).

\[
F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} \, dt
\]  

(2.2)

The continuous time domain representation of the signal can be obtained from its Fourier transform by performing an Inverse Fourier transform (Equation 2.3).

\[
f(t) = \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} \, d\omega
\]  

(2.3)
Example of a stationary signal:

Stationary signals are those signals whose frequency content does not change with respect to time.

\[ x(t) = \cos(20\pi t) + \cos(50\pi t) + \cos(100\pi t) + \cos(200\pi t) \]

Hence, the frequencies 10 Hz, 25 Hz, 50 Hz and 100 Hz are present at all instants of time.
The frequency spectrum of the time domain signal is as shown below

![Frequency domain representation of x(t)](image)

**Figure 10 Frequency Domain Representation of x(t)**

Four spectral components corresponding to frequencies 10 Hz, 25 Hz, 50 Hz and 100 Hz are shown in the frequency domain representation in Fig 10.

### 2.1.2. Discrete Fourier Transform

The Fourier Transform for the discrete signal can be defined as the Discrete Fourier Transform (DFT) (Equation 2.4). It allows the computation of the spectra from discrete-time data of N samples which satisfies the Nyquist criterion.

\[
F(k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} f(n) e^{-i2\pi kn/N} \quad \text{where } k=0, 1, 2\ldots N-1 \quad (2.4)
\]
The time domain representation of the discrete signal can be obtained from the Discrete Fourier transform using the Inverse Discrete Fourier Transform (IDFT) (Equation 2.5). The number of complex multiplication and addition operations required by the DFT and IDFT is of the order of $N^2$.

$$f(n) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} F(k) e^{-\frac{i2\pi kn}{N}} \text{ where } n=0, 1, 2\ldots N-1 \quad (2.5)$$

### 2.1.3. Fast Fourier Transform (FFT)

The Fast Fourier Transforms are fast Discrete Fourier Transform algorithms like the popular 'Radix 2' algorithms which are useful if $N$ is a regular power of 2 ($N=2^p$). The complexity involved is of the order $N \log N$.

There are two different Radix 2 algorithms, 'Decimation in Time' (DIT) and 'Decimation in Frequency' (DIF). Both these algorithms rely on the decomposition of an $N$ point transform into 2 ($N/2$) point transforms.

Although the Fourier transform provides information regarding the frequencies present in the signal, it provides no information about the occurrence of these frequencies in time. Owing to this property, Fourier transform is best suited for stationary signals wherein the frequency content of the signal does not change with time.

In the case of non-stationary signals like biological signals whose spectral content changes with time, information about the occurrence of these spectral components in time could be crucial and extremely important.
Example of a non-stationary wave

A chirp signal whose frequency changes with respect to time is plotted on the time scale in Fig 11.

![Chirp Signal](image)

**Figure 11 Chirp Signal**

A signal in the time domain (Fig 12) such that the interval 0-300ms has a 100 Hz sinusoid, the 300-600ms has a 50 Hz sinusoid, the 600-800 ms has a 25 Hz sinusoid, and the 800-1000ms has a 10 Hz sinusoid.
Figure 12 Time Domain Representation of a Signal

The frequency domain representation of the signal is shown in Fig 13. Although the frequency components are not present throughout the duration of the signal, the frequency domain shows us no such detail.
2.1.4. Short Time Fourier Transform (STFT)

The Short Time Fourier Transform (STFT) developed for both time and frequency representations analyses the signal by dividing it into several segments that can be assumed to be stationary. A window '$w$' is chosen such that it spans the segment of the signal that is assumed to be stationary.

Thus, the signal is windowed and the windowed portion of the signal is analyzed using the Fourier transform. The window is then shifted and the above mentioned process repeated for the entire signal.
The STFT suffers from a short coming whose roots go back to the Heisenberg Uncertainty Principle. According to this principle one cannot determine the exact time-frequency information of the signal, i.e. the time intervals in which a certain band of frequencies exist.

Thus, a good time and frequency resolution cannot be obtained simultaneously. A tradeoff would be to sacrifice some frequency resolution to obtain a good time resolution if the frequency components of the original signal are well separated and sacrifice some time resolution to obtain a better frequency resolution in the case of slow variation in the original signal.

2.2. Wavelet Analysis

The Wavelet transform developed as an alternative to the STFT overcomes the resolution problem. The signal is split into two parts by passing it through a high pass and low pass filter (satisfying the admissibility condition). Hence, two different versions, high pass portion and low pass portion of the same signal are obtained and these versions can further be subjected to the same operation of decomposition. The versions of the signal so obtained correspond to different frequency bands.

Unlike the STFT which provides constant resolution at all frequencies, the Wavelet Transform uses a multi-resolution technique by which different resolutions are used to analyze different frequencies. Wavelets are localized waves with their energy concentrated in time and space.
The Wavelet transform employs several wavelets of finite energy to analyze a signal. The Wavelet transform gives good time resolution and poor frequency resolution at high frequencies, while it gives good frequency resolution and poor time resolution at low frequencies.

2.2.1. The Continuous Wavelet Transform and Wavelet Series

The Continuous Wavelet Transform (CWT) is defined as follows (Equation 2.6)

\[
X_{\text{wt}}(\tau, s) = \frac{1}{\sqrt{|s|}} \int x(t) \psi^* \left( \frac{t - \tau}{s} \right) dt
\]  

(2.6)
x(t) is the signal to be analyzed. ψ(t) is the mother wavelet or the basis function. The transformed signal is a function of two variables, τ (the translation parameter) and s (the scale parameter). The wavelets used in the transformation are derived from the mother wavelet through translation (shifting) and scaling (dilation or compression).

Based on the desired characteristics, the basis functions are generated from the mother wavelet. The time information (location of the wavelet) in the Wavelet transform is provided by the translational parameter τ. The frequency information is provided by the scale parameter s (1/frequency).

Large scales corresponding to low frequencies, dilate the signal thus, providing the hidden information in the signal, while small scales corresponding to high frequencies, provides the overall information about the signal by compressing it.

It can be summarized that the Wavelet transform performs convolution between the basis functions and the signal. Wavelet series is the discretized version of the CWT obtained by sampling the time scale plane, where sampling is performed based on the Nyquist criterion.

2.2.2. The Discrete Wavelet Transform

The Wavelet Series is not a true discrete wavelet transform and is mostly redundant. The Discrete Wavelet Transform involves significantly less computational time and also provides sufficient information for both the analysis and synthesis of the original signal. The Discrete Wavelet Transform is based on sub-band coding, wherein the time scale representation of the signal is obtained by the usage of digital filtering techniques.
Properties and Types of Wavelets

The wavelet is a function with a zero average (Equation 2.7)

\[
\int_{-\infty}^{\infty} \psi(t) \, dt = 0 \quad (2.7)
\]

i.e. it has some oscillations that are both negative and positive.

Admissibility is one of the most important properties of wavelets. The square integrable function \(\psi(t)\) satisfies the admissibility condition (Equation 2.8).

\[
\int_{0}^{\infty} \left| \frac{\psi(\omega)}{\omega} \right|^2 \, d\omega < +\infty \quad (2.8)
\]

The regularity property deals with the how quickly the wavelet is decaying with decreasing scale and their smoothness. They are concentrated in time which can be proven by using the vanishing moments (Equation 2.9)

\[
M_p = \int t^p \psi(t) \, dt \quad (2.9)
\]

Wavelet transforms comprise an infinite set with different families making tradeoffs between the compact localization of the basis functions and their smoothness.
Sub-band Coding and Multi-resolution Analysis

The signal to be analyzed is passed through several filters with different cutoff frequencies at different scales. The resolution of the signal can be altered by changing the filtering operations while the scale can be changed by up-sampling and down-sampling operations. Up-sampling corresponds to increasing the sampling rate of the signal by the addition of new samples (either a zero value or interpolated values). Down-sampling refers to the reduction in the sample rate or sub sampling the signal. Sub sampling by a factor \( p \) would decrease the number samples in the signal by a factor \( p \).

The DWT analyses the signal at different frequency bands at different resolutions by decomposing the signal into a 'coarse approximation' and 'detailed information'. Two
sets of functions are employed by the DWT, the scaling functions (associated with the low pass filter) and the wavelet functions (associated with the high pass filter). The signal is filtered by passing it through successive high pass and low pass filters to obtain versions of the signal in different frequency bands.

The original signal \( x(n) \) is passed through a half band low pass and high pass filter. With the signal highest frequency being \( \pi/2 \), half of the samples are eliminated adhering to the Nyquist criterion. Thus, the signal can be sub-sampled by 2 as shown in Equation 2.10.

\[
y(n) = \sum_{n} h(k) \cdot x(2n - k) \quad (2.10)
\]

\[
y_{\text{HIGH}}(k) = \sum_{n} x(n) \cdot g(2k - n) \quad (2.11)
\]

\[
y_{\text{LOW}}(k) = \sum_{n} x(n) \cdot h(2k - n) \quad (2.12)
\]

where \( y_{\text{HIGH}}(k) \) (Equation 2.11) and \( y_{\text{LOW}}(k) \) (Equation 2.12). The decomposition performed halves the time resolution and at the same time doubles the frequency resolution. Thus, at every level, the filtering and sub-sampling will result in half the time resolution and double the frequency resolution.

The successive low pass and high pass filtering of the discrete time-domain signal as shown in the Fig 17 is called the Mallat algorithm or Mallat-tree decomposition. The sequence \( x(n) \) is passed through several levels made up of low pass (\( G_0 \)) and high pass (\( H_0 \)) filters. At each level, 'detail information' \( (d_j(n)) \) is produced by the high pass filter while the 'coarse approximations' \( (a_j(n)) \) is produced by the low pass filter.
The maximum number of levels of decomposition depends on the length of the signal. The Discrete Wavelet Transform of the original signal is obtained by concatenating all the coefficients, $a(n)$ and $d(n)$.

The reconstruction process is the reverse of decomposition, where the approximation and detail coefficients at every level are up-sampled by 2 and passed through low-pass ($G_1$) and high pass ($H_1$) synthesis filters and finally added. The same number of levels is taken as in the case of decomposition.
2.3. Equivalent Wavelet Filter

The equivalent frequency response for the jth scale is given in Equation 2.13 while Equations 2.14 and 2.15 provide the individual filter responses.

\[ Q^j(e^{i\omega}) = \begin{cases} H(e^{i\omega}) & k=1 \\ H(e^{2i\omega}) \prod_{l=0}^{k-2} G(e^{2^l\omega}) & k>2 \end{cases} \] (2.13)

where
\[ H(\omega) = \sum_{-\infty}^{\infty} h(n) e^{-ik\omega} \] and
\[ G(\omega) = \sum_{-\infty}^{\infty} g(n) e^{-ik\omega} \] (2.14)

Equivalent filters for each wavelet output can be computed from (Equation 2.16) where j represents the scale of the Wavelet Transform.

\[ Q^j(e^{i\omega}) = \begin{cases} H_0(e^{i\omega}) & k=0 \\ H_0(e^{2i\omega}) G_0(e^{i\omega}) & k=1 \\ H_0(e^{4i\omega}) G_0(e^{2i\omega}) G_0(e^{i\omega}) & k=2 \\ H_0(e^{8i\omega}) G_0(e^{4i\omega}) G_0(e^{2i\omega}) G_0(e^{i\omega}) & k=3 \\ H_0(e^{16i\omega}) G_0(e^{8i\omega}) G_0(e^{4i\omega}) G_0(e^{2i\omega}) G_0(e^{i\omega}) & k=4 \end{cases} \] (2.16)

Thus, the equivalent filter response \( q^j(n) \) (Equation 2.17) is obtained by taking the IDFT of the filters \( Q^j(\omega) \).

\[ q^j(n) = \sum_{-\infty}^{\infty} Q^j(e^{i\omega}) e^{i\omega n} \] (2.17)
The block diagram in Fig 19 shows the process of obtaining the 1\textsuperscript{st}, 2\textsuperscript{nd}, 3\textsuperscript{rd}, 4\textsuperscript{th}, 5\textsuperscript{th} Wavelets using the Mallat’s algorithm.

2.4. Wavelet Transform Maxima

In a practical wavelet function $\psi(t)$, the maxima of the wavelet transform corresponds to a sharp variation in the input signal. If $\Theta(t)$ is a smoothing function, and $\psi(t)$ the first order derivative of $\Theta(t)$ (Equation 2.18)

$$\psi(t) = \frac{d\{\Theta(t)\}}{dt}$$

then the smoothing function $\Theta(t)$ can be written as in Equation 2.19

$$\Theta_{mn}(t) = 2^{\frac{m-1}{2}} \Theta(2^{m} t - n)$$

---

---
where \( \Theta(t) \) is proportional to the wavelet transform \( \psi(t) \).

\[
W_{m,n}(t) = f(t)*\psi_{m,n}(t)
\]

\[
= f(t) * \frac{d\{\Theta_{m,n}(t)\}}{dt}
\]

\[
= \theta_{m,n}(t) * \frac{d\{f(t)\}}{dt}
\]

(2.20)

A wavelet satisfying the above equation would indicate the sharp variations in the signal by the zero crossing in the wavelet coefficients. This property of the wavelet transform is utilized in detecting the characteristics in the ECG signal (P wave, R wave).

### 2.5. Quadratic Spline Wavelet

The zero crossing of the Quadratic Spline Wavelet coefficients is an essential property that is utilized in detecting the characteristics of the ECG signal. The wavelet used in this thesis is a Quadratic Spline Wavelet with compact support and one vanishing moment. It is a first derivative of the smooth function.

The Fourier transform of \( \psi(\omega) \) is given by (Equation 2.21)

\[
\psi(\omega) = i\omega \left( \sin \frac{\omega}{4} \right)
\]

(2.21)
The filters \( H(\omega) \) and \( G(\omega) \) are given by Equations 2.22 and Equations 2.23 respectively.

\[
H(\omega) = e^{i\omega/2} (\cos(\frac{\omega}{2}))^3
\]  
(2.22)

\[
G(\omega) = i4e^{i\omega/2} (\sin(\frac{\omega}{2}))
\]  
(2.23)

The plot of the mother wavelet is shown in Fig 20
The frequency response of the Q filters is shown in the Fig 21.

The frequency spectrum is split up among the five filters.

The 1\textsuperscript{st} Wavelet is characterized by a high pass filter, and the rest of the wavelet filters by band pass filters. The 1\textsuperscript{st} Wavelet of an ECG signal would reflect the high frequency content in the signal (R characteristics) with large change in the amplitude.

The 2\textsuperscript{nd} Wavelet with the R wave as input would also produce a large change in the amplitude. The low frequency waves, P-waves and T-waves have less or no influence on the 1\textsuperscript{st}, 2\textsuperscript{nd} and 3\textsuperscript{rd} Wavelets. They are more or less prominent in the 4\textsuperscript{th} and 5\textsuperscript{th} wavelets.
The DC component is removed by setting the filter amplitude to zero when $\omega=0$ in order to remove its effect on the timing relationship between zero crossing in the wavelets and the maximum/minimums in the ECG signal shown in Fig 22.

Figure 22 ECG Signal from the MIT-BIH Database [1]
1st Wavelet

The high frequency content of the ECG signal is reflected in this signal as shown in Fig 23. The zero crossings of the 1st wavelet correspond to the R waves in the ECG signal. The 1st wavelet contains most of the noise energy as compared to the rest of the wavelets. This can be attributed to the high pass characteristic of the 1st Wavelet.

Figure 23 First Wavelet of the ECG Signal from the MIT-BIH Database [1]
2\textsuperscript{nd} Wavelet

The second wavelet of the ECG signal is the band passed version of the ECG signal in the high frequency range and is as shown in Fig 24.

![Second Wavelet of the ECG Signal from the MIT-BIH Database](image)

**Figure 24** Second Wavelet of the ECG Signal from the MIT-BIH Database [1]
3rd Wavelet

It is the band passed version of the ECG signal with a lower cutoff frequency as shown in Fig 25.

Figure 25 Third Wavelet of the ECG Signal from the MIT-BIH Database [1]
4\textsuperscript{th} Wavelet

The fourth wavelet of the ECG signal is as shown in Fig 26.

Figure 26 Fourth Wavelet of the ECG Signal from the MIT-BIH Database [1]
5th Wavelet

The 5th Wavelet of the ECG signal is as shown in Fig 27.

![Fifth Wavelet](image)

*Figure 27 Fifth Wavelet of the ECG Signal from the MIT-BIH Database [1]*
CHAPTER 3

PROCESSING OF ECG SIGNALS

3.1. ECG Analysis

3.1.1. Discussions and Goals

Digital signal processing of ECG signals has been very popular over the last few decades. The physiological variability of the ECG signal, various types of noise present in the signal make the detection of the characteristic waveforms of the ECG signal very difficult. Noise types such as muscular noise artifacts due to electrode motion, power line interference, and baseline wander etc are usually contained in the ECG signal.

The wavelet transform, a promising technique used in ECG processing, breaks down the ECG signal into scales and thus, makes it easier to analyze the ECG signal in different frequency ranges. A more comprehensive picture of the ECG signal compared to traditional algorithms that use derivatives is produced.

Detection of the R-waves and the elimination of the abnormalities in the ECG signal including the PVCs is the most important step in realizing the aim of this thesis, producing an R-wave triggering signal. The delay between the R-waves of the original signal as indicated by the attributes and the R-waves detected by the developed algorithm should be less than 100 ms, thus, indicating a fast detection of the ECG signal.
The algorithm is implemented in MATLAB 6.5 and tested on real patient ECG records (MIT-BIH Database) on a PC computer. The wavelet transform is used for the detection of the R-waves.

The Quadratic Spline wavelet filter banks are used for detection of the R-waves. The information that needs to be extracted from the wavelet signals is the zero crossings that have surrounding samples greater than a positive threshold and less than a negative threshold. The extreme points in the time signal such as the local minimums and the local maximums are reflected by the zero crossings in the wavelet signals.

The high frequency components of the ECG signal are contained in the 1st wavelet. The 2nd wavelet output is a band-pass filtered version of the ECG signal and has a higher center frequency than the 3rd. The 3rd, 4th, 5th wavelet outputs are also band-pass filtered versions of the ECG signal, where the 3rd wavelet has a higher center frequency than the 4th, and the 4th higher than the 5th.

**ECG signal and all the wavelet plots**

The Quadratic Spline mother wavelet given by Equation 3.1, was used in the analysis of the ECG signals and its frequency domain representation is as shown in the Fig 28.

\[
\psi(\omega) = i \omega \left( \sin \frac{\omega}{4} \right)^4
\]  

(3.1)
The equivalent frequency response of the wavelet filter banks can be obtained from Equations 3.1, 3.2, 3.3.

\[
Q' (e^{i\omega}) = \begin{cases} 
H (e^{i\omega}) & \\
H (e^{i2\omega}) G (e^{i\omega}) & \\
H (e^{i4\omega}) G (e^{i2\omega}) G (e^{i\omega}) & \\
H (e^{i8\omega}) G (e^{i4\omega}) G (e^{i2\omega}) G (e^{i\omega}) & \\
H (e^{i16\omega}) G (e^{i8\omega}) G (e^{i4\omega}) G (e^{i2\omega}) G (e^{i\omega}) & \\
\end{cases}
\]  
(3.2)

\[
H (\omega) = \sum_{-\infty}^{\infty} h(n) e^{-i k \omega}
\]  
(3.3)
\[ G(\omega) = \sum_{-\infty}^{\infty} g(n) e^{-i k \omega} \]  

(3.4)

The frequency response of the Q filters is shown in Fig 29.

![Figure 29 Frequency Response of the Q filters](image)

ECG signal under test from the MIT BIH Database is shown in Fig 30.
Figure 30 ECG signal from the MIT-BIH Database [1]
First Wavelet of the ECG signal is shown below (Fig 31).

![First Wavelet](image)

*Figure 31 First Wavelet of the ECG Signal from the MIT-BIH Database [1]*
Second Wavelet of the ECG signal is shown below (Fig 32).

Figure 32 Second Wavelet of the ECG Signal from the MIT-BIH Database [1]
Third Wavelet of the ECG signal is shown below (Fig 33).

Figure 33 Third Wavelet of the ECG Signal from the MIT-BIH Database [1]
Fourth Wavelet of the ECG signal is shown below (Fig 34).

Figure 34 Fourth Wavelet of the ECG Signal from the MIT-BIH Database [1]
Fifth Wavelet of the ECG signal is shown below (Fig 35).

Figure 35 Fifth Wavelet of the ECG Signal from the MIT-BIH Database [1]

3.2. Detection of the R-waves

Information in the 1st wavelet is used for the detection of R-waves. The zero crossing detection algorithm is based on the search for zero crossings that correspond to the R-waves, by using the adaptive threshold techniques. A valid zero crossing is found based on the criterion that it has surrounding samples greater than a positive threshold and less than a negative threshold.
The peaks corresponding to the R-waves are searched by the algorithm in the 1st wavelet. The maximums and minimums are searched within a search window set for two seconds to ensure that at least one peak that corresponds to a R-wave is within the search window. In other words, the search window should be greater than the average heart rate. The threshold value is set so that the maximums are greater than the threshold, and the minimum is less than the same threshold value with a negative sign.

Shown below in Fig 36 are the First Wavelet peaks before the QRS detection.

![Figure 36 First Wavelet of the ECG Signal before the QRS threshold algorithm](image)
After the application of the QRS algorithm, zero crossings can be clearly observed in Figure 37 but as mentioned earlier not all zero crossings correspond to R-waves. There are a few PVCs included.

Figure 37 First Wavelet of the ECG Signal after the QRS threshold algorithm

The output from the zero crossing detection algorithm corresponds to a possible R-wave. There could be PVCs along with the R-waves in the output because of the typically high amplitude of the PVCs. Explanation about how to determine whether the output was a true R-wave or PVC is given in the next section.
3.3. Detection of the PVC Output

PVCs are detected using a filter based on the amplitude of the 1st wavelet. Because of the higher amplitude of the PVC’s compared to that of the R-waves, an adaptive threshold is used to trace the R-waves and set it equal to a value greater than the wavelet amplitude of the normal R-peaks. The first wavelet is searched for peaks greater than the adaptive threshold computed. Figure 38 shows an example of an ECG signal with 12 PVC beats and 7 R-waves.

Figure 38 ECG signal from the MIT-BIH Database [1]
The First Wavelet of the ECG signal is shown in Fig 39.

![First Wavelet of the ECG Signal from the MIT-BIH Database](image)

From the above Fig 39, it is easy to locate the PVCs by looking at the 1st wavelet and the threshold (indicated by a line). Note that there are many other peaks greater than the threshold but do not correspond to the PVCs. These are the other aberrations in the ECG signal.

### 3.4. The Adaptive Threshold Algorithms

Two different types of adaptive threshold algorithms are used in this thesis.

The first one uses the first wavelet where the maximums and the minimums that correspond to the QRS complexes are tracked by the algorithm.

The second adaptive threshold algorithm also uses the first wavelet where the maximums and minimums are searched and the wavelet amplitude of the normal R-waves is estimated. PVCs can be detected when an estimate of the wavelet amplitude of normal R-waves is found.
3.4.1. QRS Thresholds of the 1st Wavelet Signal

The adaptive thresholds extracted from the 1st wavelet will be discussed in this section. The maximums and the minimums within a 2 second search window (the window should be greater than the average heart rate) are searched by the adaptive threshold algorithm as shown in Fig. 40.

This search window is limited to a size of 2 seconds in order to update the threshold more frequently. The median of five found maximums and five found minimums stored in an array is computed. The estimator can tolerate up to two peaks produced by the PVCs or small values (during missed beats) and still output a value corresponding to a R-wave. This is the reason for the median operator to be based on five peaks. An estimate of the wavelet amplitude for normal QRS complexes is given by the median operator by computing the median of the maximums and the minimums.
The purpose of the median operator is to filter out the PVCs which are illustrated in the wavelet domain as high magnitude peaks. The median filter not only filters out a maximum corresponding to a PVC but also a maximum or a minimum that are neither

---

**Figure 40 Flow Chart for the QRS threshold algorithm**

The purpose of the median operator is to filter out the PVCs which are illustrated in the wavelet domain as high magnitude peaks. The median filter not only filters out a maximum corresponding to a PVC but also a maximum or a minimum that are neither
produced by a QRS complex or a PVC wave which might occur when the ECG signal has missed beats.

Updating the thresholds on R-wave peaks is ideal but is not always possible because of the possibility of the PVC peak updating the threshold. This problem is taken care of in the algorithm by a constant called scale_factor. Scale_factor is a constant which in most cases is set to a value of 0.35. The value of the scale_factor has been experimentally found by iterating the algorithm on several ECG signals (MIT-BIH Database). Factors involved when choosing the scale_factor normally are the influence of noise, irregular R-wave amplitude, number of false positive detections, and the number of false negative detections. Setting the scale_factor to a higher value (0.4-0.5) would be more helpful in most of the cases. The influence of high frequency noise could be decreased by increasing the scale_factor.

Disadvantages of setting the scale_factor high include missing some R-waves if the amplitude variance of the R-waves is high. The peaks from the R-waves will be greater than the final value of the threshold due to this factor. One can choose between basing the threshold on positive peaks or negative peaks, depending on the shortest distance to X-axis, by setting the threshold to scale_factor*median[max peaks] if median[max peaks] is less than –median[min peaks], else scale_factor*median[min peaks].
Example

The First Wavelet of the ECG signal is analyzed and the QRS threshold algorithm applied to the peaks of the First Wavelet. The filtered peaks of the First Wavelet are thus obtained by using the QRS threshold criterion.

Shown below in Fig 41 is the ECG signal under test.

Figure 41 ECG Signal from the MIT-BIH Database [1]
The First Wavelet obtained after using the Filter Banks based on Mallat's algorithm is shown below in Fig 42.

Figure 42 First Wavelet of the ECG Signal from the MIT-BIH Database [1]
Peaks of the First Wavelet before the usage of the ORS threshold algorithm is shown in Fig 43.

Figure 43 First Wavelet Peaks before the QRS detection
On the application of the QRS threshold algorithm, the zero crossings of the R-waves, PVCs or other aberrations, if any, are obtained as shown in Fig 44.

![Graph of First Wavelet Peaks after the QRS detection](image)

**Figure 44 First Wavelet Peaks after the QRS detection**

### 3.4.2. PVC Detection Based on the 1st Wavelet Thresholds

PVC’s are usually characterized by peaks with typically high amplitudes. Hence, the PVC detection based on the amplitude of the 1st wavelet searches for peaks (negative or positive) with abnormally high amplitudes. A pvc_threshold which is greater than R-wave peaks and less than PVC peaks is computed. The main intention of the algorithm is to save the previous peak values of R-waves and PVCs into an array and then check the
variance of the peaks. High variance normally indicates the existence of a PVC in the array.

The PVC algorithm based on the 1st wavelet uses two adaptive threshold. The first threshold called a local_threshold alienates the PVCs and R-waves from waveforms such as large peaks due to T-waves or other low frequency artifacts. The local_threshold, being used for only one purpose i.e. to find the R-waves and the PVCs, is chosen in a way that its value is less than the amplitude of the R and PVC peaks. The more important pvc_threshold is set at a value so that it is greater than the amplitude of the R-waves and lesser than the amplitude of the PVCs. A detailed discussion for computing both the thresholds is discussed in the subsequent sections.
The algorithm uses a sliding window to search for a maximum value greater than the local_threshold as shown in Fig 45. The maximum value found is stored into an array.
max_n. The local_threshold is computed based on wavelet peaks corresponding to previous R-waves.

Figure 46 ECG Signal from the MIT-BIH Database [1]
IF (present maximum > local_threshold)

\[ max_i = \text{present maximum} \]  \hspace{1cm} (3.5)

local_threshold = 0.8 * median \{ max_1, max_2 \ldots max_{10} \}

where \( max_n \) is the previous found maximum values.

Calculating the median out of 10 maximum values gives a stable estimator for the R-wave amplitude and is determined experimentally. One of the salient features of this method is that the median of the 10 maximum values will most likely be produced by a R-wave even though the maximum values are not produced by R-waves. Misleading maximum values with low amplitudes that do not correspond to the R-waves or the PVCs are removed with the help of local_threshold.
The past 10 maximum values stored in the array might or might not have PVCs. For this, two different cases are considered when the pvc_threshold is updated.

For the case when there are no PVCs found in the array, the threshold is set as

\[ pvc\_threshold = 1.33 \times (\text{greatest of max}_n \text{ array}) \]  
(3.6)

The pvc_threshold is set to a value 33% greater than the largest R-wave.

For the case when there are PVCs found in the array, the threshold is set as

\[ pvc\_threshold = \frac{\text{PVC value} + \text{R-wave with greatest amplitude}}{2} \]  
(3.7)

The pvc_threshold is set to the mean value between a PVC amplitude and the largest R-wave amplitude, a value greater than the R-wave peaks and less than the PVC peaks.

A line is drawn by the algorithm between values corresponding to R-waves and the PVCs by considering the variance within a sorted array containing the maximum values.

The figure below illustrates how PVCs and R-waves are separated. The array is first sorted from maximum to minimum. A large difference in the two sets of numbers (R-waves and PVCs) can easily be seen when the array is sorted.
Table 2 Example of the PVC detection in a Maxn array

<table>
<thead>
<tr>
<th>Unsorted</th>
<th>Sorted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max$_n$</td>
<td>Max$_n$</td>
</tr>
<tr>
<td>1.4</td>
<td>2.8</td>
</tr>
<tr>
<td>(R-wave)</td>
<td>(PVC)</td>
</tr>
<tr>
<td>1.2</td>
<td>2.3</td>
</tr>
<tr>
<td>(R-wave)</td>
<td>(PVC)</td>
</tr>
<tr>
<td>1.0</td>
<td>1.4</td>
</tr>
<tr>
<td>(R-wave)</td>
<td>(R-wave)</td>
</tr>
<tr>
<td>2.8</td>
<td>1.4</td>
</tr>
<tr>
<td>(PVC)</td>
<td>(R-wave)</td>
</tr>
<tr>
<td>1.4</td>
<td>1.3</td>
</tr>
<tr>
<td>(R-wave)</td>
<td>(R-wave)</td>
</tr>
<tr>
<td>2.3</td>
<td>1.2</td>
</tr>
<tr>
<td>(PVC)</td>
<td>(R-wave)</td>
</tr>
<tr>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>(R-wave)</td>
<td>(R-wave)</td>
</tr>
<tr>
<td>1.1</td>
<td>1.1</td>
</tr>
<tr>
<td>(R-wave)</td>
<td>(R-wave)</td>
</tr>
<tr>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>(R-wave)</td>
<td>(R-wave)</td>
</tr>
<tr>
<td>1.3</td>
<td>1.0</td>
</tr>
<tr>
<td>(R-wave)</td>
<td>(R-wave)</td>
</tr>
</tbody>
</table>

Max Variance = \((2.3-1.4)/1.4\) = 0.6428

Max Variance > 0.4 (constant) indicating the presence of the PVC in the array

The sorted array is scanned for the greatest difference between the elements.

\[
\text{diff}_{\text{max}} = \max_{n=0} \{\text{sorted}_{\text{max}_n} - \text{sorted}_{\text{max}_{n+1}}\} \tag{3.8}
\]

If \(\text{diff}_{\text{max}}\) is at a maximum when \(n=n_0\), then the condition below is true only when a PVC is detected within the 10 point array. Based upon the MIT-BIH Database, the constant was estimated to be 0.4.

\[
\frac{\text{diff}_{\text{max}}}{\text{sorted}_{\text{max}_{n_0+1}}} > \text{constant} \tag{3.9}
\]

The above equation is an indicator of at least one PVC in the 10 maximum values.
Example:

Table 3 Example of the PVC detection in a Maxn array

<table>
<thead>
<tr>
<th>Unssorted Maxn</th>
<th>Sorted Maxn</th>
<th>Max Variance=</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.4 (R-wave)</td>
<td>2.8 (PVC)</td>
<td>(2.3-1.4)/1.4</td>
</tr>
<tr>
<td>1.2 (R-wave)</td>
<td>2.3 (PVC)</td>
<td>=0.6428</td>
</tr>
<tr>
<td>1.0 (R-wave)</td>
<td>1.4 (R-wave)</td>
<td>Max Variance&gt;</td>
</tr>
<tr>
<td>2.8 (PVC)</td>
<td>1.4 (R-wave)</td>
<td>0.4 (constant)</td>
</tr>
<tr>
<td>1.4 (R-wave)</td>
<td>1.3 (R-wave)</td>
<td>indicating the</td>
</tr>
<tr>
<td>2.3 (PVC)</td>
<td>1.2 (R-wave)</td>
<td>presence of the</td>
</tr>
<tr>
<td>1.2 (R-wave)</td>
<td>1.2 (R-wave)</td>
<td>PVC in the array</td>
</tr>
<tr>
<td>1.1 (R-wave)</td>
<td>1.1 (R-wave)</td>
<td></td>
</tr>
<tr>
<td>1.0 (R-wave)</td>
<td>1.0 (R-wave)</td>
<td></td>
</tr>
<tr>
<td>1.3 (R-wave)</td>
<td>1.0 (R-wave)</td>
<td></td>
</tr>
</tbody>
</table>

The array on the left shows the R-waves and the PVC’s in the order of their occurrence in the ECG signal. The array on the right shows the result after sorting the peak values from maximum to minimum. The above condition is applied to the numbers in the arrays and a PVC occurrence is concluded.

The algorithm reads the 1st wavelet data into a search window at a time. The local_threshold is computed by taking the median of the 10 most recent maximum values. All the peak values are compared to the local_threshold and are stored into a sorted array if they are greater than the local_threshold. The sorted array is then scanned for a large difference in between the numbers. The existence of the PVCs is checked and if found, the pvc_threshold is set to the mean value of the PVC and the largest R-wave. If not found, pvc_threshold is set to a value 33% greater than the largest R-wave.

On being found, the PVCs are eliminated from the 1st wavelet. An R-wave trigger signal is generated after the elimination of other aberrations. The indices of the R-wave are read from the annotation file of the MIT-BIH database signal. These indices are
compared to the indices of the R-wave trigger signal. The difference of the indices is calculated and that value is divided by the sampling frequency. The value thus obtained gives the delay of the signal. The average delay value is found to be less than 100 ms for most of the signals.
CHAPTER 4

STATISTICS

4.1. Validation database

The MIT/BIH Database was used to evaluate the algorithm. Only channel 2 of the two-channel ECG signal in the database was used. Signal files for the MIT/BIH Database are sampled at either 250 samples per second or 360 samples per second. All the ECG recordings are annotated where each beat is described by an annotation.

4.2. Example

Figure 48 Example of ECG signal from the MIT-BIH Database [1].
4.3 Statistical Results

The algorithm presented in this thesis has been tested on the MIT/BIH Database. The performance of the algorithm was tested against the annotation files from the MIT/BIH Database.

The results of the R-wave and PVC detection in many of the ECG records were encouraging. The wavelet transforms of some ECG signals were not accurately performed because of serious high frequency noise, baseline drift, and artifacts.

The detection of R-waves in some signals was very difficult due to the presence of high amplitude peaks which are neither R-waves nor PVCs. These peaks could not be eliminated because the amplitude of these waves was greater than the local_threshold and less than the pvc_threshold. These waves posed as the R-waves in the trigger signal, thus, making the delay calculations extremely difficult. Some more problems were discussed in the next section.
<table>
<thead>
<tr>
<th>ECG Signal</th>
<th>No. of R-waves</th>
<th>No. of R-waves Detected</th>
<th>No. of PVCs Detected</th>
<th>Delay in ms</th>
<th>Detection Rate of R-waves</th>
<th>Detection Rate of PVCs</th>
<th>False +ve</th>
<th>False -ve</th>
</tr>
</thead>
<tbody>
<tr>
<td>806.dat</td>
<td>13</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>84.61538</td>
<td>N/A</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>100.dat</td>
<td>9</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>101.dat</td>
<td>9</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>2.7778</td>
<td>88.88889</td>
<td>N/A</td>
<td>2</td>
</tr>
<tr>
<td>103.dat</td>
<td>19</td>
<td>7</td>
<td>0</td>
<td>5.556</td>
<td>36.84211</td>
<td>N/A</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>104.dat</td>
<td>10</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>90</td>
<td>N/A</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>106.dat</td>
<td>8</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>N/A</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>119.dat</td>
<td>13</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>N/A</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>200.dat</td>
<td>12</td>
<td>8</td>
<td>4</td>
<td>3</td>
<td>66.66667</td>
<td>N/A</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>201.dat</td>
<td>10</td>
<td>7</td>
<td>0</td>
<td>1</td>
<td>70</td>
<td>N/A</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>202.dat</td>
<td>17</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>5.556</td>
<td>11.76471</td>
<td>N/A</td>
<td>1</td>
</tr>
<tr>
<td>203.dat</td>
<td>13</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2.7778</td>
<td>7.692308</td>
<td>N/A</td>
<td>3</td>
</tr>
<tr>
<td>207.dat</td>
<td>9</td>
<td>6</td>
<td>3</td>
<td>2</td>
<td>19.44</td>
<td>66.66667</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>118.dat</td>
<td>7</td>
<td>5</td>
<td>0</td>
<td>38.889</td>
<td>71.42857</td>
<td>0.333333</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>208.dat</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>100</td>
<td>N/A</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>209.dat</td>
<td>7</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>57.14286</td>
<td>N/A</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>210.dat</td>
<td>9</td>
<td>7</td>
<td>1</td>
<td>0</td>
<td>77.77778</td>
<td>N/A</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>212.dat</td>
<td>13</td>
<td>7</td>
<td>0</td>
<td>8.33</td>
<td>53.84615</td>
<td>N/A</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>213.dat</td>
<td>12</td>
<td>9</td>
<td>0</td>
<td>13.889</td>
<td>75</td>
<td>N/A</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>214.dat</td>
<td>12</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>133.3333</td>
<td>N/A</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>215.dat</td>
<td>13</td>
<td>13</td>
<td>0</td>
<td>83.33</td>
<td>100</td>
<td>N/A</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>217.dat</td>
<td>5</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>5.556</td>
<td>160</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>219.dat</td>
<td>19</td>
<td>15</td>
<td>0</td>
<td>5.556</td>
<td>78.94737</td>
<td>N/A</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>300.dat</td>
<td>31</td>
<td>17</td>
<td>0</td>
<td>8.33</td>
<td>54.83871</td>
<td>N/A</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>301.dat</td>
<td>9</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>2.7778</td>
<td>100</td>
<td>N/A</td>
<td>3</td>
</tr>
<tr>
<td>302.dat</td>
<td>12</td>
<td>8</td>
<td>0</td>
<td>19.44</td>
<td>66.66667</td>
<td>N/A</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>303.dat</td>
<td>13</td>
<td>13</td>
<td>0</td>
<td>2.7778</td>
<td>100</td>
<td>N/A</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>304.dat</td>
<td>13</td>
<td>12</td>
<td>0</td>
<td>22.222</td>
<td>92.30769</td>
<td>N/A</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>305.dat</td>
<td>10</td>
<td>9</td>
<td>0</td>
<td>11.111</td>
<td>90</td>
<td>N/A</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>306.dat</td>
<td>9</td>
<td>10</td>
<td>0</td>
<td>11.111</td>
<td>111.1111</td>
<td>N/A</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>307.dat</td>
<td>14</td>
<td>13</td>
<td>0</td>
<td>2.7778</td>
<td>92.85714</td>
<td>N/A</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>308.dat</td>
<td>20</td>
<td>19</td>
<td>0</td>
<td>33.33</td>
<td>95</td>
<td>N/A</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>309.dat</td>
<td>12</td>
<td>12</td>
<td>0</td>
<td>8.33</td>
<td>100</td>
<td>N/A</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>310.dat</td>
<td>14</td>
<td>5</td>
<td>0</td>
<td>2.7778</td>
<td>35.71429</td>
<td>N/A</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>311.dat</td>
<td>9</td>
<td>9</td>
<td>0</td>
<td>13.88</td>
<td>100</td>
<td>1</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>312.dat</td>
<td>18</td>
<td>16</td>
<td>0</td>
<td>8.33</td>
<td>88.88889</td>
<td>N/A</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>313.dat</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8.33</td>
<td>0</td>
<td>N/A</td>
<td>3</td>
</tr>
<tr>
<td>314.dat</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2.778</td>
<td>0</td>
<td>N/A</td>
<td>2</td>
</tr>
<tr>
<td>315.dat</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2.778</td>
<td>0</td>
<td>N/A</td>
<td>2</td>
</tr>
<tr>
<td>316.dat</td>
<td>17</td>
<td>13</td>
<td>0</td>
<td>5.556</td>
<td>76.47059</td>
<td>0.4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>317.dat</td>
<td>11</td>
<td>12</td>
<td>0</td>
<td>2.778</td>
<td>109.0909</td>
<td>N/A</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>File</td>
<td>X1</td>
<td>X2</td>
<td>X3</td>
<td>X4</td>
<td>X5</td>
<td>X6</td>
<td>X7</td>
<td>X8</td>
</tr>
<tr>
<td>-----------------</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>318.dat</td>
<td>8</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>112.5</td>
<td>N/A</td>
<td>2</td>
</tr>
<tr>
<td>319.dat</td>
<td>11</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>2.778</td>
<td>63.63636</td>
<td>N/A</td>
<td>1</td>
</tr>
<tr>
<td>320.dat</td>
<td>8</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>5.556</td>
<td>100</td>
<td>N/A</td>
<td>1</td>
</tr>
<tr>
<td>321.dat</td>
<td>7</td>
<td>7</td>
<td>0</td>
<td>1</td>
<td>2.778</td>
<td>100</td>
<td>N/A</td>
<td>1</td>
</tr>
<tr>
<td>322.dat</td>
<td>13</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>53.84615</td>
<td>N/A</td>
<td>1</td>
</tr>
<tr>
<td>323.dat</td>
<td>12</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>11.11</td>
<td>100</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>324.dat</td>
<td>13</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>5.5556</td>
<td>123.0769</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>325.dat</td>
<td>9</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>8.33</td>
<td>133.3333</td>
<td>75</td>
<td>0</td>
</tr>
<tr>
<td>327.dat</td>
<td>17</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>25</td>
<td>52.94118</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>804.dat</td>
<td>7</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>7.812</td>
<td>42.85714</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel100.dat</td>
<td>6</td>
<td>7</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>116.6667</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel104.dat</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>60</td>
<td>66.6667</td>
<td>0</td>
</tr>
<tr>
<td>Sel114.dat</td>
<td>7</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>85.71429</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel117.dat</td>
<td>16</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>56.25</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>Sel123.dat</td>
<td>7</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>114.2857</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel14046.dat</td>
<td>13</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>92.30769</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sel14157.dat</td>
<td>7</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>100</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel14172.dat</td>
<td>11</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>81.81818</td>
<td>N/A</td>
<td>2</td>
</tr>
<tr>
<td>Sel15814.dat</td>
<td>8</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>50</td>
<td>N/A</td>
<td>1</td>
</tr>
<tr>
<td>Sel16265.dat</td>
<td>9</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>8</td>
<td>66.6667</td>
<td>N/A</td>
<td>2</td>
</tr>
<tr>
<td>Sel16272.dat</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>133.3333</td>
<td>N/A</td>
<td>3</td>
</tr>
<tr>
<td>Sel16273.dat</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>N/A</td>
<td>N/A</td>
<td>3</td>
</tr>
<tr>
<td>Sel16420.dat</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>N/A</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sel16483.dat</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel16539.dat</td>
<td>8</td>
<td>8</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>100</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel16773.dat</td>
<td>---</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>#VALUE!</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel16786.dat</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel16795.dat</td>
<td>16</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>16</td>
<td>100</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel17152.dat</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel17453.dat</td>
<td>11</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>100</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel213.dat</td>
<td>15</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>80</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel221.dat</td>
<td>13</td>
<td>9</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>69.23077</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel223.dat</td>
<td>19</td>
<td>11</td>
<td>0</td>
<td>1</td>
<td>16</td>
<td>57.89474</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel230.dat</td>
<td>15</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel231.dat</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>100</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel232.dat</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>16.66667</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel233.dat</td>
<td>11</td>
<td>9</td>
<td>5</td>
<td>2</td>
<td>8</td>
<td>81.81818</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel301.dat</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel302.dat</td>
<td>8</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>8</td>
<td>0</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel308.dat</td>
<td>8</td>
<td>7</td>
<td>2</td>
<td>2</td>
<td>12</td>
<td>87.5</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel808.dat</td>
<td>7</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel811.dat</td>
<td>8</td>
<td>8</td>
<td>3</td>
<td>3</td>
<td>12</td>
<td>100</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel820.dat</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>8</td>
<td>100</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel821.dat</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel840.dat</td>
<td>14</td>
<td>13</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>92.85714</td>
<td>N/A</td>
<td>0</td>
</tr>
</tbody>
</table>
### 4.4 Problems Faced During R-Wave and PVC Detection

The test results of the R-wave and PVC detection were displayed in the table.

Several figures are shown below to describe the problems faced during the detection. The reasons for the problems are noisy signals, PVC amplitude, a rhythm change initially in the ECG signals of the MIT-BIH Database. Despite the problems described above and some other problems, the algorithm works efficiently for most of the signals.

The signal quality is poor in figures (49, 50) due to high frequency noise which makes the algorithm label some of the R-waves as PVC’s. The high frequency noise hampers the detection of R-waves.

<table>
<thead>
<tr>
<th>File</th>
<th>Test 1</th>
<th>Test 2</th>
<th>Test 3</th>
<th>Test 4</th>
<th>Quality</th>
<th>Noise</th>
<th>Check 1</th>
<th>Check 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sel853.dat</td>
<td>10</td>
<td>10</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>100</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel871.dat</td>
<td>6</td>
<td>9</td>
<td>5</td>
<td>5</td>
<td>12</td>
<td>150</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel872.dat</td>
<td>6</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>100</td>
<td>N/A</td>
<td>0</td>
</tr>
<tr>
<td>Sel873.dat</td>
<td>8</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>20</td>
<td>0</td>
<td>N/A</td>
<td>1</td>
</tr>
<tr>
<td>Sel891.dat</td>
<td>9</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>100</td>
<td>1</td>
</tr>
<tr>
<td>Cu01.dat</td>
<td>11</td>
<td>11</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>100</td>
<td>33.3333</td>
<td>1</td>
</tr>
<tr>
<td>Cu03.dat</td>
<td>17</td>
<td>13</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>76.47059</td>
<td>N/A</td>
<td>1</td>
</tr>
<tr>
<td>Cu06.dat</td>
<td>7</td>
<td>7</td>
<td>3</td>
<td>3</td>
<td>12</td>
<td>100</td>
<td>N/A</td>
<td>1</td>
</tr>
</tbody>
</table>
Figure 49 ECG signal from the MIT-BIH Database [1]

Figure 50 ECG signal from the MIT-BIH Database [1]

Figure 51 illustrates the ECG signal sel14046. Here the two PVC’s have positive and negative amplitudes. This creates a problem when trying to detect the PVC’s because
the positive amplitudes are almost equal to the amplitudes of the R-waves. The tradeoff is based on choosing between a false positive R-wave and PVC in this scenario.

![Figure 51 ECG signal from the MIT-BIH Database](image)

There is a rhythm change (Figures 52, 53) associated with most of the ECG signals in the MIT-BIH Database at the beginning. Because of its high amplitude, the rhythm change is often mistaken for an R-wave or more usually PVC.
Figure 52 ECG signal from the MIT-BIH Database [1]

Figure 53 ECG signal from the MIT-BIH Database [1]
In the Figure 54, the direction of the PVC is in the opposite direction of the R-waves and other PVC’s. This creates a problem in detecting the PVC because the adaptive threshold value eliminates this PVC, its amplitude being less than the threshold value.

Figure 54 ECG signal from the MIT-BIH Database [1]
CHAPTER 5

CONCLUSIONS

5.1. Problems in ECG Processing

The ECG is a complex voltage-time relationship representation of the cardiac cycle. The physiological variability of the ECG signal and the various types of noise present in the signal can make the detection of the characteristic waveforms in an ECG signal very difficult. Noise types such as muscular noise, artifacts due to electrode motion, power line interference, and baseline wander etc could be present in an ECG signal.

The narrow and sharp shape of the R-waves makes it simple for their detection compared to P-waves and T-waves. Detection of PVC’s could be very difficult at times when noise in the ECG with very high amplitude, a characteristic of PVC, poses as PVC.

5.2. Future Work

The algorithm presented in this thesis works very well in the detection of the ECG waveforms. The high frequency noise in the ECG signal causes the algorithm to output false positive R-waves. This can be avoided by increasing the adaptive threshold that determines the valid R-waves in the wavelet signals. Estimating the noise power and then feeding back that information to the algorithm could be the possible solution for this problem because it helps in making the decisions about whether the adaptive thresholds should be increased.
In this thesis, the algorithm has been tested on the signals in the MIT-BIH Database, which had the ECG signals sampled at 360 hertz. Testing the algorithm on ECG signals that have higher sampling frequency would be an interesting follow-up. It would result in higher resolution in the wavelet signals that would be beneficial for the detection accuracy.

The delay between the R-waves of the original signal as indicated by the attributes and the R-waves detected by the developed algorithm was found to be less than 100 ms.

5.3. Summary Of What Has Been Done In This Thesis

In this thesis, an algorithm has been developed that detects the R-waves in the ECG signal including the detection of abnormal waveforms such as PVC waves (Premature Ventricular Contractions). The algorithm involves types of filters such as the wavelet transform implemented as a filter-bank, adaptive thresholds.

The wavelet filter-banks used in the algorithm decompose the ECG signal into frequency scales. The locations of the R-waves are extracted from these frequency scales. P-waves, R-waves and T-waves are easily distinguished by this approach by studying the energy in each of the frequency scale. For example, the QRS-complex has a broad frequency range and is dominant in the high frequency scales in comparing the P-wave and the T-wave, which are extracted from the lower frequency ranges.

The locations of the R-waves are represented in the frequency scales as zero-crossings with surrounding samples of significant amplitude. An adaptive algorithm, to find these zero-crossings, is discussed to compute the thresholds that are used in the search for zero-crossings that are produced by R-waves.
The filter used for the detection of PVCs is discussed. PVCs typically produce peaks with great amplitude in the lower frequency scales. Only the first wavelet is used in the algorithm for the detection of the R-waves and the subsequent elimination of other abnormalities including the PVCs. PVCs (in the first wavelet) are eliminated from the first wavelet using an adaptive threshold algorithm. The delay was found to be less than 100 ms.
APPENDIX A

ANNOTATION OF ECG SIGNALS IN MIT/BIH DATABASE
<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>not-QRS (not a getann/putann code)</td>
</tr>
<tr>
<td>1</td>
<td>normal beat</td>
</tr>
<tr>
<td>2</td>
<td>left bundle branch block beat</td>
</tr>
<tr>
<td>3</td>
<td>right bundle branch block beat</td>
</tr>
<tr>
<td>4</td>
<td>aberrated atrial premature beat</td>
</tr>
<tr>
<td>5</td>
<td>premature ventricular contraction</td>
</tr>
<tr>
<td>6</td>
<td>fusion of ventricular and normal beat</td>
</tr>
<tr>
<td>7</td>
<td>nodal (junctional) premature beat</td>
</tr>
<tr>
<td>8</td>
<td>atrial premature contraction</td>
</tr>
<tr>
<td>9</td>
<td>premature or ectopic supraventricular beat</td>
</tr>
<tr>
<td>10</td>
<td>ventricular escape beat</td>
</tr>
<tr>
<td>11</td>
<td>nodal (junctional) escape beat</td>
</tr>
<tr>
<td>12</td>
<td>paced beat</td>
</tr>
<tr>
<td>13</td>
<td>unclassifiable beat</td>
</tr>
<tr>
<td>14</td>
<td>signal quality change</td>
</tr>
<tr>
<td>16</td>
<td>isolated QRS-like artifact</td>
</tr>
<tr>
<td>18</td>
<td>ST change</td>
</tr>
<tr>
<td>19</td>
<td>T-wave change</td>
</tr>
<tr>
<td>20</td>
<td>systole</td>
</tr>
<tr>
<td>21</td>
<td>diastole</td>
</tr>
<tr>
<td>22</td>
<td>comment annotation</td>
</tr>
<tr>
<td>23</td>
<td>measurement annotation</td>
</tr>
<tr>
<td>Code</td>
<td>Description</td>
</tr>
<tr>
<td>-------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>PAVE</td>
<td>/* P-wave peak */</td>
</tr>
<tr>
<td>BBB</td>
<td>/* left or right bundle branch block */</td>
</tr>
<tr>
<td>PACESP</td>
<td>/* non-conducted pacer spike */</td>
</tr>
<tr>
<td>TAVE</td>
<td>/* T-wave peak */</td>
</tr>
<tr>
<td>RHYTHM</td>
<td>/* rhythm change */</td>
</tr>
<tr>
<td>UAVE</td>
<td>/* U-wave peak */</td>
</tr>
<tr>
<td>LEARN</td>
<td>/* learning */</td>
</tr>
<tr>
<td>FLWAV</td>
<td>/* ventricular flutter wave */</td>
</tr>
<tr>
<td>VFON</td>
<td>/* start of ventricular flutter/fibrillation */</td>
</tr>
<tr>
<td>VFOFF</td>
<td>/* end of ventricular flutter/fibrillation */</td>
</tr>
<tr>
<td>AESC</td>
<td>/* atrial escape beat */</td>
</tr>
<tr>
<td>SVESC</td>
<td>/* supraventricular escape beat */</td>
</tr>
<tr>
<td>LINK</td>
<td>/* link to external data (aux contains URL) */</td>
</tr>
<tr>
<td>NAPC</td>
<td>/* non-conducted P-wave (blocked APB) */</td>
</tr>
<tr>
<td>PFUS</td>
<td>/* fusion of paced and normal beat */</td>
</tr>
<tr>
<td>WFON</td>
<td>/* waveform onset */</td>
</tr>
<tr>
<td>WFOFF</td>
<td>/* waveform end */</td>
</tr>
<tr>
<td>RONT</td>
<td>/* R-on-T premature ventricular contraction */</td>
</tr>
</tbody>
</table>
APPENDIX B

MATLAB CODE
clc; clear all;
pi=3.142;
signal=2;

%------ SPECIFY DATA -----------------------------------------------
PATH= 'C:\MATLAB6p5\work\thesis';  % path, where data are saved
HEADERFILE= '208.hea';      % header-file in text format
ATTRFILE= '208.atr';         % attributes-file in binary format
DATAFILE= '208.dat';         % data-file

SAMPLES2READ=2000;         % number of samples to be read
s1=SAMPLES2READ;                           % in case of more than one signal:%
2*SAMPLES2READ samples are read

%------ LOAD HEADER DATA ---------------------------------------------
fprintf(1,\n$> WORKING ON %s ...
', HEADERFILE);  % 1 stands for standard output i.e. display while 2 for standard error
signalh= fullfile(PATH, HEADERFILE);  % fullfile provides the full filename of the file, i.e. the path of the file
fid1=fopen(signalh,'r');  % fopen opens the file based on the permission requested and returns a 1 if opened else returns -1
z= fgetl(fid1);     % fgetl returns the next line of the file identified by fid1 and is used only in the case of text files.
A= sscanf(z, '%*s %d %d %d',[1,3]); % scanf(it is vectorized) the contents in z into the required format before outputting them. %*s is to skip the first value
nosig= A(1);  % number of signals
sfreq=A(2);  % sample rate of data
clear A;
for k=1:nosig
    z= fgetl(fid1);
    A= sscanf(z, '%*s %d %d %d %d %d',[1,5]);
    dformat(k)= A(1);
    gain(k)= A(2);              % number of integers per mV
    bitres(k)= A(3);            % bitresolution
    zerovalue(k)= A(4);         % integer value of ECG zero point
    firstvalue(k)= A(5);        % first integer value of signal (to test for errors)
end;
fclose(fid1);
clear A;

%------ LOAD BINARY DATA ---------------------------------------------
if dformat~=[212,212], error('this script does not apply binary formats different to 212.'); end;
signalld= fullfile(PATH, DATAFILE);  % data in format 212
fid2=fopen(signalld,'r');  
A= fread(fid2, [3, SAMPLES2READ], 'uint8');
Main=A;
fclose(fid2);
M2H= bitshift(A(:,2), -4);

M1H= bitand(A(:,2), 15);
PRL=bitshift(bitand(A(:,2),8),9);  % sign-bit
PRR=bitshift(bitand(A(:,2),128),5);  % sign-bit
M(:,1)= bitshift(M1H,8)+ A(:,1)-PRL;
M(:,2)= bitshift(M2H,8)+ A(:,3)-PRR;
M(1,:);
if M(1,:)~= firstvalue, error('inconsistency in the first bit values'); end;
switch nosig
  case 2
    if gain==0
      M(:,1)= (M(:,1)- zerovalue(1));
      M(:,2)= (M(:,2)- zerovalue(2));
    else
      M(:,1)= (M(:,1)- zerovalue(1))/gain(1);
      M(:,2)= (M(:,2)- zerovalue(2))/gain(2);
    end
    TIME=(0:(SAMPLES2READ-1))/sfreq;
  case 1
    M(:,1)= (M(:,1)- zerovalue(1));
    M(:,2)= (M(:,2)- zerovalue(1));
    size(M);
    M=M';
    size(M);
    M(1)=zeros(size(M(1,:))-1,1);
    size(M);
    sM=size(M);
    sM=sM(2)+1;
    M(sM)=0;
    size(M);
    M=M';
    size(M);
    M=M/gain(1);
    TIME=(0:2*(SAMPLES2READ-1))/sfreq;
  otherwise  % this case did not appear up to now!
    disp('Sorting algorithm for more than 2 signals not programmed yet!');
  end;
clear A M1H M2H PRR PRL;
fprintf(1,'\n$> LOADING DATA FINISHED 
')
%------ LOAD ATTRIBUTES DATA ---------------------------------------------------------------

atrd = fullfile(PATH, ATRFILE); % attribute file with annotation data
fid3=fopen(atrd,'r');
A = fread(fid3, [2, inf], 'uint8');
BGY = A;
close(fid3);
ATRTIME = [];
ANNOT = [];
sa = size(A);
saa = sa(1);
i = 1;
while i <= saa
    anoth = bitshift(A(i,2),-2);
    if anoth == 59
        ANNOT = [ANNOT; bitshift(A(i+3,2),-2)]; % concatenating ANNOT and the bit shift
    end;
    a = bitshift(bitand(A(i,2),3),8)+A(i,1);
    hilfe = hilfe+mod(hilfe,2);
    i = i+hilfe/2;
else
    ATRTIME = [ATRTIME; bitshift(bitand(A(i,2),3),8)+A(i,1)];
    ANNOT = [ANNOT; bitshift(A(i,2),-2)];
end;
i = i+1;
end;
ANNOT(length(ANNOT)) = [] % last line = EOF (=0)
ATRTIME(length(ATRTIME)) = [] % last line = EOF
clear A;
ATRTIME = (cumsum(ATRTIME))/sfreq;

ind = find(ATRTIME <= TIME(end));

ATRTIMED = ATRTIME(ind);
ANNOT=round(ANNOT);
ANNOTD=ANNOT(ind);

% if nosig==2
% U(:,1)=M(:,2);
% P=U';
% end
% this part check again
% U2(:,1)=M(:,1);
% P1=U2;

if nosig==1
  s1=2*SAMPLES2READ;
end
if nosig==1
  m=M(:,1);
else
  m=M(:,2);
end

fs=sfreq;
i=sqrt(-1);
hfilt=exp(2*j*pi*nn/(fs*2))*((cos(2*pi*nn/(2*fs)))^3);
gfilt=4*j*(exp((2*j*pi*nn)/(2*fs)))*(sin(2*pi*nn/(2*fs)));
for nn=1:sfreq/2
  w1(nn,1)=-eval(gfilt);
  w2(nn,1)=-eval(subs(gfilt,2*nn)*hfilt);
  w3(nn,1)=-eval(subs(gfilt,4*nn)*subs(hfilt,2*nn)*hfilt);
  w4(nn,1)=-eval(subs(gfilt,8*nn)*subs(hfilt,4*nn)*subs(hfilt,2*nn)*hfilt);
  w5(nn,1)=-eval(subs(gfilt,16*nn)*subs(hfilt,8*nn)*subs(hfilt,4*nn)*subs(hfilt,2*nn)*hfilt);
end
NX=(floor(log2(s1)))+1;
NX=9;
w1mag=abs(w1);
w2mag=abs(w2);
w3mag=abs(w3);
w4mag=abs(w4);
w5mag=abs(w5);
TIME2=[1:sfreq/2];
plot(TIME2,w1mag(1:sfreq/2), 'r');hold on;
plot(TIME2, w2mag(1:sfreq/2), 'c');hold on;
plot(TIME2, w3mag(1:sfreq/2), 'b');hold on;
plot(TIME2, w4mag(1:sfreq/2), 'm');hold on;
plot(TIME2, w5mag(1:sfreq/2), 'g');hold off;

Tdwt1=real(ifft(w1,sfreq/2));
Tdwt2=real(ifft(w2,sfreq/2));
Tdwt3=real(ifft(w3,sfreq/2));
Tdwt4=real(ifft(w4,sfreq/2));
Tdwt5=real(ifft(w5,sfreq/2));

TIME2=[1:sfreq/2];
plot(TIME2,Tdwt1(1:sfreq/2), 'r');hold on;
plot(TIME2, Tdwt2(1:sfreq/2), 'c');hold on;
plot(TIME2, Tdwt3(1:sfreq/2), 'b');hold on;
plot(TIME2, Tdwt4(1:sfreq/2), 'm');hold on;
plot(TIME2, Tdwt5(1:sfreq/2), 'g');hold off;

td_dw1=conv(Tdwt1,m);
td_dw2=conv(Tdwt2,m);
td_dw3=conv(Tdwt3,m);
td_dw4=conv(Tdwt4,m);
td_dw5=conv(Tdwt5,m);

TD_dw1=td_dw1(1:s1);
TD_dw2=td_dw2(1:s1);
TD_dw3=td_dw3(1:s1);
TD_dw4=td_dw4(1:s1);
TD_dw5=td_dw5(1:s1);
TIME2=[1:s1];
plot(TIME2,TD_dw1, 'r');
plot(TIME2, TD_dw2, 'c');hold on;
plot(TIME2, TD_dw3, 'b');hold on;
plot(TIME2,TD_dw4, 'm');hold on;
plot(TIME2,TD_dw5, 'g');
plot(TIME, m,'b');
grid on;
xlim([TIME(1), TIME(end)]);
xlabel('Time / s'); ylabel('Voltage / mV');
string=['ECG signal ',DATAFILE];
title(string); TIME2=(0:(s1-2));
TIME1=(0:2*(SAMPLES2READ)-2)/sfreq;
hold off;

%********************************************************** THRESHOLD ALGORITHM
**********************************************************%
windsz=2*sfreq;       % window size (2 sec)
remwind=rem(s1,windsz); % remainder if window size is not a multiple of the total
number of samples taken
no_windshift=fix(s1/windsz); % no of windows required if the s1 is completely
divisible by the window size else the number of windows required would be
no_windshift + 1
if remwind==0
    sz_remwind=remwind;
end
set=0;
finTD_dw1=zeros(s1,1);
for d=2:s1-1
    set=0;
    finTD_dw1(d)=zeros(s1,1);
    set=0;
    if abs(TD_dw1(d))>=abs(TD_dw1(d+1)) & abs(TD_dw1(d))>=abs(TD_dw1(d-1))
        set=1;
        finTD_dw1(d)=TD_dw1(d);
    end
    if set==1
        if isempty(nonzeros(finTD_dw1(1:d-1)))==0
            recent_ind=max(find(finTD_dw1(1:d-1)));
        end
    end
end
if abs(finTD_dw1(d)+finTD_dw1(recent_ind))> abs(finTD_dw1(d)-finTD_dw1(recent_ind)) &  d-recent_ind <= 30
  if abs(finTD_dw1(d))>abs(finTD_dw1(recent_ind))
    finTD_dw1(recent_ind)=0;
  else
    finTD_dw1(d)=0;
  end
end
end

Bef_finTD_dw1=finTD_dw1;

for i=1:no_windshift
  sorted_dw1=sort(finTD_dw1((i-1)*windsz+1:i*windsz));
  maxmedian_dw1(i)=median(sorted_dw1(windsz-4:windsz));
  minmedian_dw1(i)=median(sorted_dw1(1:5));
  for h=(i-1)*sfreq+1:i*windsz
    if finTD_dw1(h)>0 & finTD_dw1(h) < 0.8*maxmedian_dw1(i)
      finTD_dw1(h)=0;
    end
    if finTD_dw1(h)<0 & abs(finTD_dw1(h)) < 0.8*abs(minmedian_dw1(i))
      finTD_dw1(h)=0;
    end
  end
end

if remwind~=0
  sorted_dw1=sort(finTD_dw1((no_windshift)*windsz+1:s1));
  maxmedian_dw1(no_windshift+1)=median(sorted_dw1(remwind-4:remwind));
  minmedian_dw1(no_windshift+1)=median(sorted_dw1(1:5));
end

for h=(no_windshift*windsz)+1:s1
  if finTD_dw1(h)>0 & finTD_dw1(h) < 0.8*maxmedian_dw1(no_windshift+1)
    finTD_dw1(h)=0;
  end
  if finTD_dw1(h)<0 & abs(finTD_dw1(h)) < 0.8*abs(minmedian_dw1(no_windshift))
    finTD_dw1(h)=0;
  end
end
kv=conv(m,yo);
kv2=conv(m,y2w);
kv3=conv(m,y3w);
kv4 = conv(m, y4w);
kv5 = conv(m, y5w);
for i = 1:s1
    jk(i, 1) = kv(i, 1);
    jkfive(i, 1) = kv5(i, 1);
end
jk1 = jk;
jk5 = jk;
jk3 = jk;

%------ DISPLAY DATA -------------------------------------------------------

if nosig == 2
    grid on;
    subplot(221);
    plot(TIME, m, 'b'); hold on;
    % plot(TIME, gh, 'r'); hold off
end;
for k = 1:length(ARTRIMED)
    text(ARTRIMED(k), 0, num2str(ANNOTD(k)));
end;
grid on;
xlim([TIME(1), TIME(end)]);
xlabel('Time / s'); ylabel('Voltage / mV');
string = ['ECG signal ', DATAFILE];
title(string); TIME2 = (0:((2*s1)-2));
TIME1 = (0:2*(SAMPLES2READ)-2)/sfreq;

fprintf(1, '
$> DISPLAYING DATA FINISHED 
');
% subplot(222);
% plot(TIME1, kv, 'r'); hold on;
% xlim([TIME(1), TIME(end)]);
% xlabel('Time / s'); ylabel('Voltage / mV');
% string=['FIRST WAVELET ', DATAFILE];
% title(string); hold on;
% subplot(223);
% % plot(TIME1, -kv2, 'y'); hold on;
% xlim([TIME(1), TIME(end)]);
% xlabel('Time / s'); ylabel('Voltage / mV');
% string=['SECOND WAVELET ', DATAFILE];
% title(string); hold on;
% subplot(224);
% plot(TIME1, -kv3, 'k'); hold on;
% xlim([TIME(1), TIME(end)]);
xlabel('Time / s'); ylabel('Voltage / mV');

%string=[' THIRD WAVELET ',DATAFILE];
%title(string);

% subplot(222);
% plot(TIME1,-kv4, 'r'); hold on;
% xlim([TIME(1), TIME(end)]);
% xlabel('Time / s'); ylabel('Voltage / mV');
% string=['FOURTH WAVELET ',DATAFILE];
% title(string);

% subplot(223);
% plot(TIME1,+kv5, 'm'); hold off;
% xlim([TIME(1), TIME(end)]);
% xlabel('Time / s'); ylabel('Voltage / mV');
% string=['FIFTH WAVELET ',DATAFILE];
% title(string);


% no_of_pvcs=zeros(floor(s1/10),1);

rr=700;
uo=0; test=0;
for j=0:((floor(s1/rr))-1) % to find whether the signal is negative peaked or positive peaked..could have taken anything but taken sfreq so that i need not add any other variable.
    for p=1:rr
        map(p,1)=jk3(j*rr+p,1);
    end
    maxim(j+1,1)=max(map);
    minim(j+1,1)=min(map);
end
for i=0:5
    max_latest(i+1,1)=maxim(floor(s1/rr)-i,1);
    min_latest(i+1,1)=minim(floor(s1/rr)-i,1);
end
if median(max_latest)<=median(min_latest)
    threshold_init=median(max_latest);
    test=0;
else
    threshold_init=median(max_latest);
    test=0;
end
scale=0.5;
local_threshold=scale*threshold_init;

jk8=zeros(s1,1);
for pb=1:s1
    if (jk3(pb,1))>=local_threshold
        if (pb==1)||(pb==s1)
            if pb==1
                if jk3(pb,1)>=threshold_init
                    jk8(pb,1)=0;
            elseif pb==s1
                if jk3(pb,1)>=jk3(pb-1,1)
                    jk8(pb,1)=jk3(pb,1);
                    jk8(pb-1,1)=0;
            end
        end
    end
    elseif pb==s1
        if jk3(pb,1)>=jk3(pb-1,1)
            jk8(pb,1)=jk3(pb,1);
            jk8(pb-1,1)=0;
    end
end

jk10=zeros(s1,1);

xg=100;
for ddt=0:floor(s1/xg)-1
    for ddt1=1:xg
        jk9(ddt1,1)=jk8(ddt*xg+ddt1,1);
    end
    [mval,indv]=max(jk9);
    jk10(ddt*xg+indv,1)=mval;
jkrem=jk8;
jk1=jk8;

for j=0:((floor(s1/rr))-1)
    for i=700*j+1:700*(j+1)
        mat(i,1)=jkrem(i,1);
    end
    mat=jkrem;
yu=find(mat);
ps=size(yu);
sp=ps(1,1);
for pk=1:sp
    mat1(pk,1)=jkrem(yu(pk,1));
end

if test==1
    [ds,indx]=sort(mat);
    for k=1:rr-1
        digg(k,1)=(ds(k,1))-(ds(k+1,1));
    end
    [diffmax,inde]=min(digg);
    if abs(diffmax/ds(inde+1,1))>=0.6
        no_of_pvcs(j+1,1)=inde;
        pvc_threshold=(ds(inde,1)+ds(inde+1,1))/2;
    else
        no_of_pvcs(j+1,1)=0;
        pvc_threshold=1.33*ds(1,1);
    end
    for o=rr*(j+1):rr*(j+1)
        % if jk1(o,1)<pvc_threshold
        %     jk1(o,1)=0;
        % end
        mat(o,1)=jk1(o,1);
    end
    max2(j+1,1)=max(mat);
    if (max2(j+1,1)<=local_threshold)
        max1(j+1,1)=max2(j+1,1);
    else
        % max1(j+1,1)=0;
    end
    % matr=find(max1);
local_threshold=0.8*median(max2);
for h=rr*j+1:rr*(j+1)
    if jk1(h,1)<local_threshold
        jk1(h,1)=0;
    end
end

elseif test==0
    [ds,indx]=dsort(mat1); % sorting all the local max's
    for k=1:sp-1
        digg(k,1)=ds(k,1)-ds(k+1,1); % finding difference between consecutive elements in
        % the sorted stack
        end

    [diffmax,inde]=max(digg); % finding the index and the value of the maximum
    % difference in the stack
    sasu(j+1,1)=abs(diffmax/ds(inde+1,1));
    if abs(diffmax/ds(inde+1,1))>=0.3
        no_of_pvcs(j+1,1)=inde;
        pvc_threshold=(ds(inde,1)+ds(inde+1,1))/2;
    else
        no_of_pvcs(j+1,1)=0;
        pvc_threshold=1.33*ds(1,1);
    end
end

for o=rr*j+1:rr*(j+1)
    if jk1(o,1)>pvc_threshold
        jk1(o,1)=0;
    end
    % mat(o,1)=jkfive(o,1);
end

% max2(j+1,1)=max(mat);
% if (max2(j+1,1)>=local_threshold)
%     max1(j+1,1)=max2(j+1,1);
%     % else
%     % max1(j+1,1)=0;
% end
% matr=find(max1);

% local_threshold=0.5*median(max2);
% for h=rr*j+1:rr*(j+1)
%     if jk1(h,1)>local_threshold

%       jk1(h,1)=0;
%       end
%       end
end

pvc=0;ch=0;
rythm=0;
abber=0;
dsd=0;elf=0;
rwave=0;
szannot=size(ANNOTD);
for r=1:szannot(1)
    if ANNOTD(r,1)==5
        pvc=pvc+1;
dsd=dsd+1;
        ch=ch+1;
        pvcs(ch,1)=r;
        pvcindex(ch,1)=ATRTIMED(r,1)*sfreq;
        remov(dsd,1)=ATRTIMED(r,1)*sfreq;
    elseif ANNOTD(r,1)==28
        rythm=rythm+1;
dsd=dsd+1;
        remov(dsd,1)=ATRTIMED(r,1)*sfreq;
    elseif (ANNOTD(r,1)==1) & (ANNOTD(r,1)==5) & (ANNOTD(r,1)==28)
        abber=abber+1;
dsd=dsd+1;
        remov(dsd,1)=ATRTIMED(r,1)*sfreq;
    elseif ANNOTD(r,1)==1
        rwave=rwave+1;
        elf=elf+1;
        VERY(elf,1)=r;
        rwaveindex(elf,1)=ATRTIMED(r,1)*sfreq;
    end
 end

no_rwaves=elf;
if ANNOTD(r,1)==1
    h1b=size(rwaveindex,1)-size(find(jk1));
    fpfn=h1b(1,1);
    if fpfn>0
        F_ngtv=fpfn;
        F_pstv=0;
    else
        F_ngtv=0;
        F_pstv=0;
    end
 else
    F_ngtv=0;
    F_pstv=0;
end
F_ngtv=0;
F_pstv=fpfn;
end

for x=1:elf
    %if jk1(x,1)==0
    delay(x,1)=(rwaveindex(x,1)-jk1(x,1))/sfreq;
    %end
end
delay
uo=uo+1;
matdelay(uo,1)=mean(delay);
end
end
% [delay1,inds]=min(matdelay);
% [delay2,inds2]=max(matdelay);
% [delay3,inds3]=mean(matdelay);
% %scale=0.3+(inds*0.03);
% mindelay=delay1;
% maxdelay=delay2;
% avgdelay=delay3;
% if diffmax/ds(inde+1,1)>=0.4
totalpvcs=sum(no_of_pvcs);
%end

MaxDelay=max(delay);
MinDelay=min(delay);
AvgDelay=mean(delay);

jk2=jk1;
if dsd>0
    szremov =size(remov);
    for i=1:szremov(1)
        ttr=floor(remov(i,1));
        if jk1(ttr,1)==0
            jk2(ttr,1)=0;
        end
    end
    for ju=1:20
        if (ttr-ju)>0
            jk2(ttr-ju,1)=0;
        end
    end
end
for ju=1:20
    if(ttr+ju)<=s1
        jk2(ttr+ju,1)=0;
    end
end

rmtrigg=find(jk2);
szrmtrigg=size(rmtrigg);

suj=find(jk1);
Varnames=str2mat('Test','pvc_threshold','Threshold','MinDelay','MaxDelay','AvgDelay','# PVC','calculated PVCs','# Rythm Change','Other Abberations','False +ve','False -ve');
Data=[test pvc_threshold local_threshold MinDelay MaxDelay AvgDelay pvc totalpvcs
rythm abber F_pstv F_ngtv];
Casenames=str2mat(DATAFILE);

if signal==2
    [Data1,Varnames1,Casenames1]=tblread('Statistics_ECG.dat');
    Casenames=strvcat(Casenames1,Casenames);
    Data=vertcat(Data1,Data);
end
tblwrite(Data,Varnames,Casenames,'Statistics_ECG.dat');
type Statistics_ECG.dat;
LIST OF REFERENCES


