A NOVEL TIME-DOMAIN DIAGNOSTIC METHOD FOR ECG SIGNAL SYSTEM BASED ON A SMART-PHONE

by

Shijie Zhou

Submitted in partial fulfillment of the
requirements for the degree of
Master of Applied Science

at

Dalhousie University
Halifax, Nova Scotia
July 2012

DALHOUSIE UNIVERSITY

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled "A NOVEL TIME-DOMAIN DIAGNOSTIC METHOD FOR ECG SIGNAL SYSTEM BASED ON A SMART-PHONE" by Shijie Zhou in partial fulfillment of the requirements for the degree of Master of Applied Science.

Dated: July 30, 2012

Co-Supervisors: _____

_____

Readers: _____

_____

DALHOUSIE UNIVERSITY

DATE: July 30, 2012

AUTHOR: Shijie Zhou

TITLE: A NOVEL TIME-DOMAIN DIAGNOSTIC METHOD FOR ECG SIGNAL SYSTEM BASED ON A SMART-PHONE

DEPARTMENT OR SCHOOL: Department of Electrical and Computer Engineering

DEGREE: M.A.Sc. CONVOCATION: October YEAR: 2012

Signature of Author

Table of Contents

List of Tables

List of Figures

Abstract

This thesis presents a novel method on a Smart-phone for ECG tele-monitoring signal analysis. The proposed system focuses on QRS complex detection, beat classification and arrhythmias classification. In the regular process, the QRS complex is detected by the Pan-Tompkins algorithm and classified as normal sinus rhythms (SRs) or premature ventricular contractions (PVCs) by existing classification methods. Subsequently, the Lempel-Ziv (LZ) complexity measure, including the K-Means clustering algorithm and the LZ complexity analysis, is utilized to further separate the high risk arrhythmias, ventricular tachycardia (VT) or ventricular fibrillation (VF). In this procedure of the high risk arrhythmias, three consecutive PVC beats in a row are considered to be an indication of the beginning of VT rhythms, at which point the following data points will be saved until up to a certain window length long are reached. The window length long ECG signal will be further classified as VT or VF by several new decision rules with heart rate detection. Furthermore, the proposed system successfully implemented on a Smart-phone adopts the time frames to indicate the analysis report for improving the reliability and error detection of arrhythmias. The new analysis method presents fairly good performance results when applied to testing records chosen from the MIT-BIH database.

List of Abbreviations Used

3G	Third-generation
AF	Atrial Fibrillation
AFL	Atrial Flutter
AN	Atrionodal
AV	Atrioventricular Node
bpm	Beats Per Minute
CVD	Cardiovascular disease
dB	Decibel
ECG	Electrocardiogram
EEG	Electroencephalogram
EMBS	Engineering in Medicine and Biology Society
FFT	Fast Fourier Transform
GA	Genetic Algorithm
Hz	Hertz
IDE	Integrated Development Environment
LA	Left Arm
LL	Left Leg
LZ	Lempel-Ziv

mm	Millimeter
mm/sec	Millimeter Per Second
mps	Meters Per Second
ms	Millisecond
mv	Millivolt
N	Nodal
NH	Nodal-His
OS	Operation System
PAC	Premature Atrial Contractions
PDA	Personal Digital Assistant
PDF	Probability Density Function
PPG	Photoplethysmograph
PVC	Premature Ventricular Contraction
RA	Right Arm
RL	Right Leg
SA	Sinoatrial Node
SNR	Signal-to-Noise Ratio
SR	Sinus Rhythm
SVT	Supraventricular Tachycardia
V	Chest
VF	Ventricular Fibrillation
VT	Ventricular Tachycardia

WL Window Length

Acknowledgements

It is greatly grateful to lots of people who have made contributions to this thesis by kinds of ways.

Foremost, I would like to owe my most sincere gratitude to my supervisor, Dr. Jason Gu for the constant support of M.A.Sc.Program, and guidance on this research. With his patience, enthusiasm, and abundant engineering knowledge to make me to understand clearly and simply, he guided and helped me in all the time of research and accomplishment of this thesis. During this Master program, he provided lots of sound advice, good research planning, and research experiences. He is an excellent advisor and mentor for my Master program study.

I would like to appreciate my co-supervisor Dr. Adel Merabet, and thesis committee: Dr. M. El-Hawary and Dr. Yajun Pan, for their dedication, insightful comments, and invaluable direction to improve the quality of this thesis.

I am grateful to my labmates in Dr. Gu's Robotic Group for their help and concern. Especially, I would like to appreciate Zichen Zhang for the research and life discussions in the three years, and Dr. Kun Zhan, Dr. Yuanlong Yu for their sharing, helpful advice and resourcefulness. As well, I would like to thank my friend Xiaoou Mao for his friendship and "Three People Coffee Break", and Mr. Roland Alexander for his kindness and help, some of whom have already been named. In addition, I am thankful to Mr. Chris Hill, Mr. Mark LeBlanc and Mr. Alex Pudsey for their technical support and secretaries Ms. Selina Cajolais and Ms. Nicole T. Smith for helping the department to run well and for assisting me in many aspects.

I wish to give my special thanks to my entire family, my grandmother, my parents, and my parent-in-law for their help, encouragement and attention.

Lastly, and most importantly, I wish to express my sincere appreciation to my loving wife Annie Zhou for continued support of all my endeavors, encouragement and love.

Chapter 1

Introduction

This chapter presents a simple research background, motivations, objectives as well as contributions of this study. A brief outline would be proposed at the end of this chapter.

1.1 Research Background

Cardiovascular disease (CVD) is the major threatening disease as well as one of the three leading factors of mortality in the world [10]. Generally speaking, patients with heart problems need to undergo a cardiac test at the hospital by using the electrocardiographic devices or instruments. An electrocardiogram (ECG) produced by an electrocardiographic device for tracking cardiac activity is widely adopted as an important indicator to record abnormal heart function and morphology, which usually measures in a non-invasive way via skin electrodes. After testing, a cardiologist makes a diagnosis that combines the ECG with the clinical symptoms to take into consideration whether there has been an abnormality in the patient's heart condition. However, the use of single ECG makes the cardiologist cause the personal subjective opinion, leading to the undetected issues of two aspects: congenital heart disease and coronary heart disease. For example, some patients with coronary artery disease manifest the normal signs of heart rhythm at most of the time until their hearts are suddenly attacked by the potential arrhythmias. As well, some patients with abnormal heart rhythms are difficult to be detected and may need to be monitored or measured over an extended period of time for accurate diagnosis by a cardiologist [11, 12]. Normally, it may not be necessary to keep the patients hospitalized for a few days of observation if there are no immediate life-threatening cardiac arrhythmias. Therefore, ECG tele-monitoring analysis systems have become a developing trend, to allow the patient to continue regular daily living while providing real-time diagnosis and keeping the Emergency Health Center updated on its user. Figure 1.1 shows

an ECG tele-monitoring analysis system. Currently, ECG tele-monitoring analysis systems have various types that can be divided into two modes in terms of operation: real-time mode and store-and-forward mode [13]. In the real-time mode, the patients' data are analyzed after acquisition, and then an intermediary platform (e.g.. Personal Computer, Personal Digital Assistant (PDA) , etc.) is utilized to transmit or analyze the ECG signal in real-time. In the store-and-forward mode, the patients' data are acquired and stored, accessing or downloading the data at a later time for offline analysis.

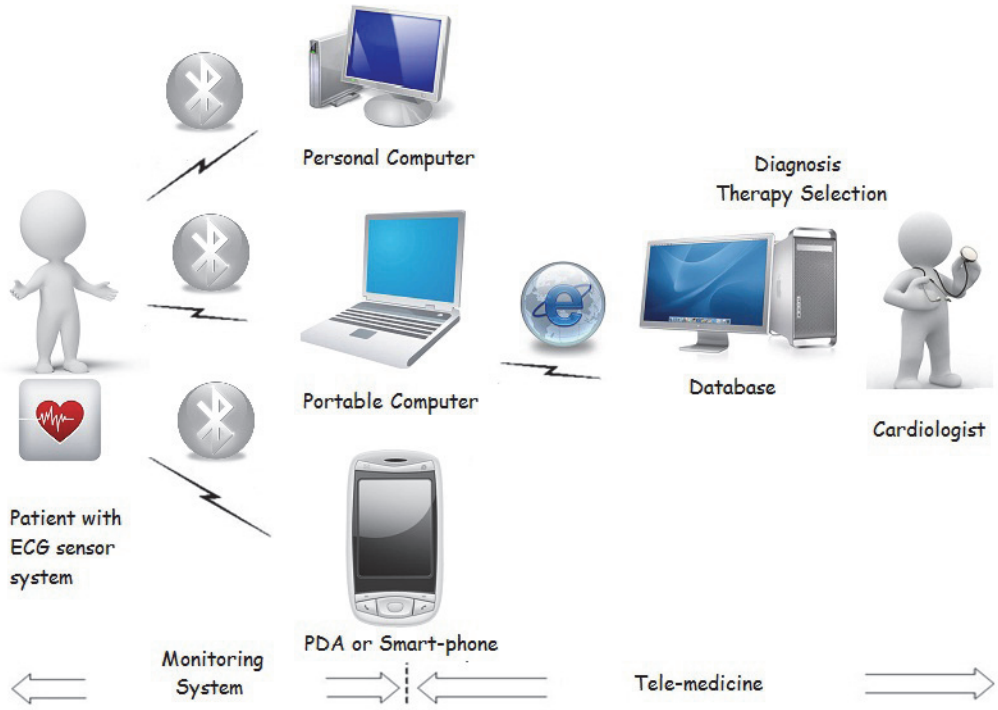


Figure 1.1: ECG tele-monitoring analysis system

In recent years, studies have been mainly focusing on ECG tele-monitoring analysis systems [14, 15, 16]. Both flexible and user-friendly, it connects the remote Emergency Health Center to the patients, therefore making the system more practical, and allowing patients to have a normal lifestyle. There are three types of ECG tele-monitoring analysis systems [17].

1. Systems that record ECG signal and perform offline classification;
2. Systems that perform remote real-time classification;
3. Systems that provide local classification in real-time.

The first one , Holter monitor, is generally worn for 24-48 hours during normal activity. After this period, it is returned to a cardiologist who then examines the records and diagnoses whether there has been any arrhythmia. Holter presented to the representative is adopted to continuously monitor and record the electrical conduction system of the heart [18]. Figure 1.2 illustrates that a Holter is attached to monitor the user; the ECG strip maps the data from the Holter detection device. The drawback of using a Holter is that it cannot perform a real-time analysis on the ECG tele-monitoring analysis system.

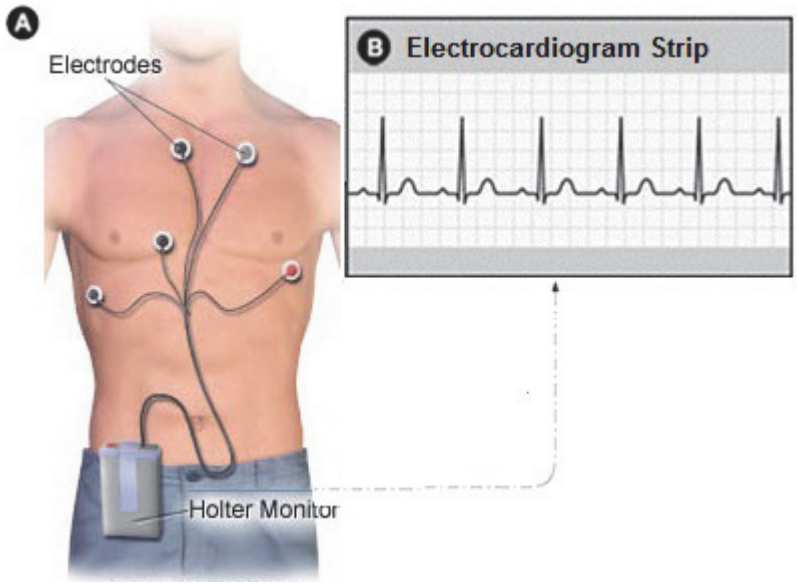


Figure 1.2: (A) Holter is attached to a user (B) Electrocardiogram strip

In order to overcome the Holter's shortcomings, people began to work on improving the heart monitor devices combining majority of wireless mobile devices with the Holters for adding an online diagnosis function. Holter was developed and expanded, becoming a part of ECG tele-monitoring analysis system in real-time to monitor people's health and wellbeing in home health care. In the second type, researchers utilize a mobile phone or a PDA as an intermediary platform, sending the ECG signal from the heart monitor device to a remote Health Service Station, where the cardiologist is available to analyze the ECG signal and give a diagnostic feedback to the user in the real-time [19, 20]. Still, the diagnostic period may be too long for the patients who need an immediate diagnosis and medical attention. There are several applications, such as Alive technology [21], Vitaphone [22]. To overcome these restrictions,

a most popular solution where an intermediary platform (e.g., Smart-phone, PDA, Tablets with 3G (Third-generation) , etc.) can perform local ECG signal analysis [17, 23, 24, 25, 26], and then send the abnormal data to a remote Health Service Station for further consultation with a doctor, or for measurement with historical information of the patient. Some examples are @home [27], or OSIRIS-SE [28]. However, this method still needs a cardiologist to give online instructions and suggestions, missing the timely rescue, increasing the human resource cost and restricting the application range.

1.2 Motivations

Heart disease causes death in one out of every three people [10]. Heart disease has been being the healthy hot issue because of the most significant cause of mortality in the world. According to the statistics [29], somebody dies of heart disease or stroke, every 7 minutes in Canada. As well, one of the important social problems in Canada we are facing now is the dramatically increasing percentage of the elderly population. The risk of CVD increases with your age, which means that there is a greater risk of CVD over age of 45 for men or 55 for women [29]. Considering this data as it relates to the aging population situation, Health Canada costs more than $22.2 billion every year for heart disease and stroke [29]. The escalating numbers of aged persons in Canada will continue to be a huge challenge in nursing homes and hospitals with regards to the financial and staffing costs. In addition, due to the accelerated pace of life, people are under more and more intensive stress than before. They ignore their health, so sudden death occurs without any medical symptom. Therefore, the focus of modern medicine has been shifting from disease diagnosis to disease prevention and control services. Since people are more concerned about their health, they strive to achieve a healthier lifestyle. There is a great demand on cost-effective devices which can monitor the physical fitness and improve the user quality of life, especially, cardiac health. With the development of wireless communication, tele-health and tele-medicine systems considered a new and emerging area have been playing an crucial role in creating versatile and cost effective alternatives to assist health care.

Over the years, most research has been dedicated to hardware issues of ECG tele-monitoring analysis systems, such as ECG signal acquisition devices [30], different telecommunication intermediary platforms [17, 27], transmission modes of communication [31], transmission protocols [32] and wireless sensor networks [33]. Along with related development of new generation mobile technologies, the hardware issues of ECG tele-monitoring analysis systems have been improved remarkably. However, little research has been done to take the performance of ECG signal analysis into consideration before in the tele-monitoring analysis system. For example, some authors have not explicitly pointed out the kind of ECG signal analysis performed in these presented systems [34], particularly the classification of the life-threatening cardiac arrhythmias, such as Ventricular Tachycardia (VT) and Ventricular Fibrillation (VF) . Besides, some researches proposed an architecture of the system and did not fully describe in details the performance of the algorithm they used [35]. Most of the algorithms were tested against their own database rather than a standard database, making it difficult to compare and evaluate their performance [36]. One could argue that these algorithms implemented on simulation platforms are too complex, and would add too much of a computation overhead to the hardware in practical application.

1.3 Objectives

The main objective of the thesis is to address an ECG signal analysis system with a novel method to improve analysis accuracy of the ECG tele-monitoring analysis system and tackle existing gaps of the current study. The proposed system implemented on a Smart-phone has a reliable and robust capacity to independently detect, and systematically classify the ECG signal based on time-domain, particularly, the arrhythmias classification of VT and VF by several new decision rules with heart rate detection. Therefore, the abnormal situation doesn't need to be transmitted to the remote Health Service Station for further analysis by a cardiologist when detecting and classifying VT and VF after a period of time. In other words, if the life-threatening arrhythmias can be efficiently detected and classified, the Smart-phone will automatically send an alert message to the remote Emergency Health Center for the timely rescue. Also, the analysis report adopts the time frames to improve the reliability

and error detection of arrhythmias in the proposed system. There are three aspects to achieve my objectives as follows.

To begin with, an algorithm architecture would be proposed, which consists of combining mature algorithms (the Pan-Tompkins algorithm [37], and the Lempel-Ziv (LZ) complexity measure algorithm [38]) separately represented by previous research in ECG signal analysis. The algorithm architecture is fast and computationally effective to be implemented on a Smart-phone for detection and classification of the ECG signal.

Secondly, the LZ complexity measure algorithm has two parts that includes coarse-graining process and complexity analysis for farther separation of the high risk arrhythmias, VT or VF. The coarse-graining process as a main part of LZ complexity measure determines how much inherent information can be retained and will consequently impact the VT and VF distinction. The question of different coarse-graining approaches interpretability in ECG signal analysis and their influence on the performance of ECG classification have not yet been previously addressed in the literature. There are four methods (K-Means clustering algorithm, Mean-value algorithm, Midpoint algorithm and Median algorithm) to be utilized in the coarse-graining process aiming at gaining a better understanding of their impact on the classification of ECG signal.

In addition, VT has two different features, monomorphic VT and polymorphic VT. Some authors only discussed the classification of monomorphic VT and VF, however, the ECG signal may contain mixed arrhythmias for a practical application, e.g. VT with monomorphic and polymorphic, 2 seconds VT followed by 2 seconds VF. Few works have taken this into consideration before. The thesis presents several new classification rules to recognize VT and VF from continuous and mixed ECG signals in this study. By combining these rules with heart rate detection, a new analysis method is described in the ECG analysis system. Then, the proposed system is implemented on a Smart-phone by using the MIT-BIH database [9] for testing.

1.4 Contributions

In this section, the scientific contributions were presented along with the corresponding or in pressing publications. The contributions of this thesis are three-fold as

follows.

1.4.1 Contributions to a novel analysis method

An algorithm architecture was presented, which consists of the Pan-Tompkins algorithm [37] and the Lempel-Ziv (LZ) complexity measure algorithm [38] in ECG signal analysis system. As well, the LZ complexity measure that includes the K-Means clustering algorithm and the LZ complexity analysis was utilized to further separate the high risk arrhythmias, VT or VF. The K-Means clustering algorithm was firstly addressed to refine the raw ECG signal in a coarse-graining process which yields much better performance of classification in the LZ complexity analysis. In addition, a novel analysis method was presented to recognize VT and VF in this study. The proposed system successfully implemented on a Smart-phone adopts the time frames to create the analysis report for improving the reliability and error detection of arrhythmias. Finally, the new analysis method indicates fairly good performance results when applied to the MIT-BIH database [9].

One paper was published in the Engineering in Medicine and Biology Society (EMBS) , 2011, Proceedings of the 33rd Annual International Conference of the IEEE [39].

1.4.2 Contributions to coarse-graining process analysis

The Lempel-Ziv (LZ) complexity measure [38] has been applied to classify VT and VF. The coarse-graining process plays a crucial role in the LZ complexity measure analysis, which directly affects the separating performance of VT and VF in ECG signal analysis. The coarse-graining approach based on the K-Means clustering algorithm on the performance of ECG arrhythmias classification has not yet been previously addressed in the literature.

A paper[40] published in 2011 EMBS presents four coarse-graining process approaches (the K-Means clustering algorithm, the Mean-value algorithm, the Mid-point algorithm and the Median algorithm). The results show that K-Means algorithm is superior to the other three approaches in VT and VF separation.

1.4.3 Contributions to arrhythmias distinction

A novel, and computationally fast method was used to classify monomorphic VT and VF, which utilized a histogram and average absolute deviation. The novelty of this method is that ECG signal statistics, morphological analysis, the histogram of signal (density estimation) and average absolute deviation altogether have been used to achieve a higher classification rate. The Student's t-test was used to analyze and reveal the significance in the histogram approach for monomorphic VT and VF arrhythmias distinction.

1.5 Organization

Chapter 1 introduces research background, motivations, objectives, contributions and outline of the thesis.

Chapter 2 starts by describing the basic concepts of ECG signal analysis, types of abnormal rhythm and the proposed ECG signal analysis system.

Chapter 3 details feature extraction. The primary part is to describe diagnostic and morphologic feature vector with time series for ECG signal analysis. There are three parts to refine some special properties within the ECG signal, which makes it to accurately detect and classify kinds of arrhythmias.

Chapter 4 investigates the algorithm for classification arrhythmia that includes the coarse-graining process and the LZ complexity analysis. In the coarse-graining process, there are four methods (K-Means clustering, Mean-value, Mid-point and Median-value) to be addressed for performance estimation. Besides, there are several new classification rules with heart rate detection to be used for separation of arrhythmias. Third part is to analyze the coarse-graining process, and makes a comparison between the K-Means clustering and the Mean-value algorithm. In addition, there is a novel method based on histogram and average absolute deviation, to be used for the arrhythmias distinction, monomorphic VT and VF.

Chapter 5 focuses on the implementation of the proposed system and evaluates the results by using the MIT-BIH database.

Chapter 6 shows the conclusion and describes the future work of the ECG analysis system.

Chapter 2

ECG Signal Analysis System

This chapter is to introduce ECG interpretation, various types of heart arrhythmias and an architecture of ECG analysis system. Firstly, there are some basic ECG terminologies to be interpreted for better understanding this proposed system. Then, abnormal heart rhythms are briefly described, and classified by the position where the arrhythmias are caused. Last but not least, this proposed system employed two mature algorithms (Pan-Tompkins algorithm [37], and Lempel and Ziv (LZ) complexity measure algorithm [38]) to detect and systematically classify the ECG signal.

2.1 Introduction

In clinical testing, the ECG has already been widely adopted to record many types of abnormal heart function and heart morphology. More specifically, an electrocardiographic device as a clinical testing tool monitors and tracks the electrical manifestation of the contractile activity of the heart related to the impulse that transmits through the heart [41]. The ECG tracing generated by pacemaker cells from the Sinoatrial node (SA) is recorded and displayed on a standard ECG graph paper which consists of small and large squares when the heart is beating in an ECG test [41]. A standard ECG tracing of a heartbeat cycle includes a P wave, a PR interval, a PR segment, a QRS complex, a ST segment, a ST interval, a QT interval and a T wave [42]. Normally, there are ten of electrodes to refer to the tracing of the voltage, yielding twelves of this type of lead recording on the ECG graph paper [1]. According to the ECG with the clinical symptoms, a cardiologist would make an effective diagnosis and evaluation, taking into consideration whether the heart has been suffering from arrhythmias.

2.1.1 The Conduction System of the Heart [1, 2, 3]

The conduction system of the heart describes the rhythmic contractile activity of the heart, which is made up of three main parts, Sinoatrial node (SA), Atrioventricular node (AV) and His-Purkinje system [41]. Figure 2.1 illustrates the whole conduction system of the heart [7]. The SA node located in the superior right corner of the right atrium is considered as the "Heart's pacemaker", which controls electrical propagation at a regular rate (60 -100 beats per minute(bpm)). The cardiac electrical impulses originate from the SA node, triggering cardiac contraction and setting the rate of contraction of the heart. Some pacemaker cells in the heart can generate periodic impulses which create the heart rate, and to rapidly conduct the stimuli to the whole heart. In other words, these pacemaker cells make up the cardiac conduction system that is a specialized pathway for the impulse spreading through the heart.

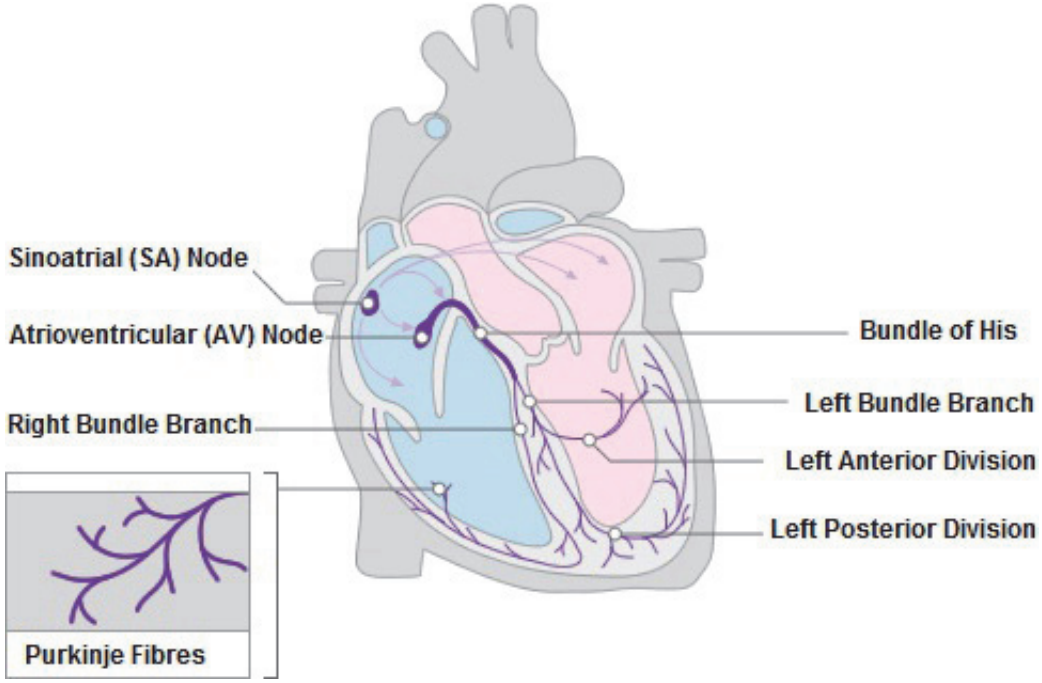


Figure 2.1: The conduction system of the heart[7]

The intrinsic electrical conduction makes the stimuli from the SA to the AV. The AV node located near the bottom of the right atrium is conducting bridge between atria and ventricle, which provide two functions: physiological conduction delay and protection of the ventricles. The AV node includes three regions: the atrionodal (AN) region, the nodal (N) region and the nodal-His (NH) region [2]. Figure 2.2

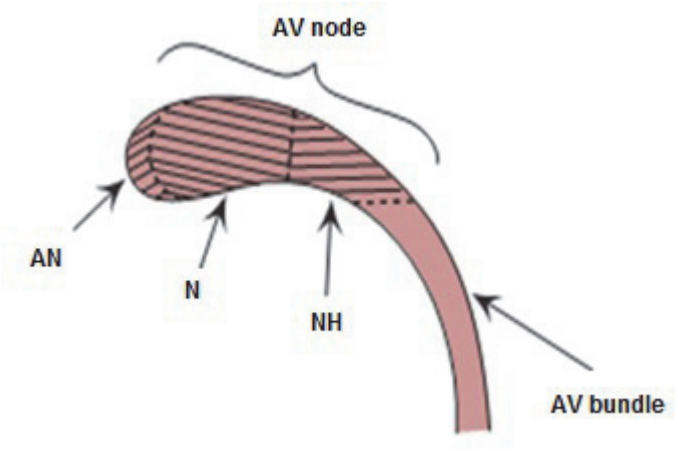


Figure 2.2: AV node conduction system [8]

illustrates the location of these regions. These stimuli are extremely slow, about 0.02 - 0.05 meters per second (mps) in the N region [2]. Therefore, it can generate an effective conduction delay that makes atrial contraction to be completed before the ventricular contraction begins. Autonomic fibers can innervate the AV junction to adjust the conduction velocity. Heart cells have the autorhythmicity in the AV junction, particularly in the AN and NH regions [2].

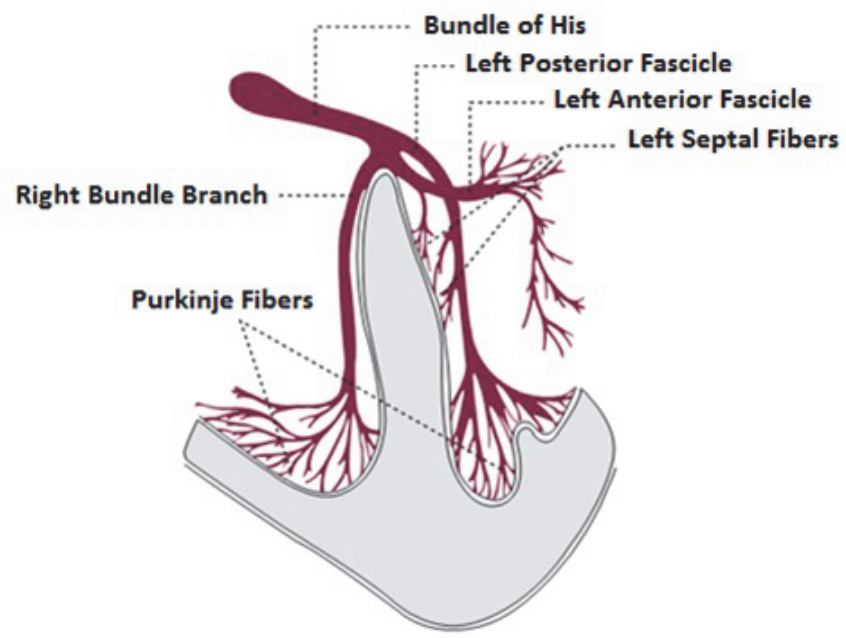


Figure 2.3: The His-Purkinje system [8]

Normal physiology further enables electrical stimuli to be relayed down from the

AV node to the ventricles or His-Purkinje system, and then be transmitted respectively to the right bundle branch and the left bundle branch. Finally, the left bundle branch divided into two fascicles/divisions, then going through the Purkinje fibers where the propagated velocity is about 1-2mps as shown in Figure 2.3.

Table 2.1: Basic Terminologies

The isoelectric line	A horizontal line with no electrical activity
Waveform	Any waveform above or below the isoelectric line in either a positive or negative direction
Segment	The region between two waveforms
Interval	One or more waveforms and a segment
Complex	Several waveforms

2.1.2 Waves, Segments and Intervals [3]

There are some basic terminologies to be addressed before introducing the waves, segments and intervals of the ECG as shown in Table 2.1. The electrical activity of the ECG tracing recording generates a heart beat that includes a P wave, a PR interval, a PR segment, a QRS complex, a ST segment, a QT interval and a T wave.

The P wave

In Figure 2.4, the depolarization of the atria means electrical activation of the atrial myocardium generating the P wave that begins at the SA node to spread throughout the atrial musculature. The duration of the P wave is less than 0.12 seconds (usually 0.08 seconds to 0.1 seconds). By observing the P wave, it is effective to distinguish various cardiac arrhythmias that cause at the atria.

The P-R Interval

As Figure 2.5 shows, there is a brief isoelectric region after the P wave, which describes that the stimuli are delayed within the AV node and the bundle of His. The duration ranges of the P-R interval are about 0.12 seconds to 0.2 seconds. The duration of the P-R Interval is from the P wave to the beginning of the QRS complex. Variations of the P-R interval result in kinds of heart disease. For example, if the P-R interval is

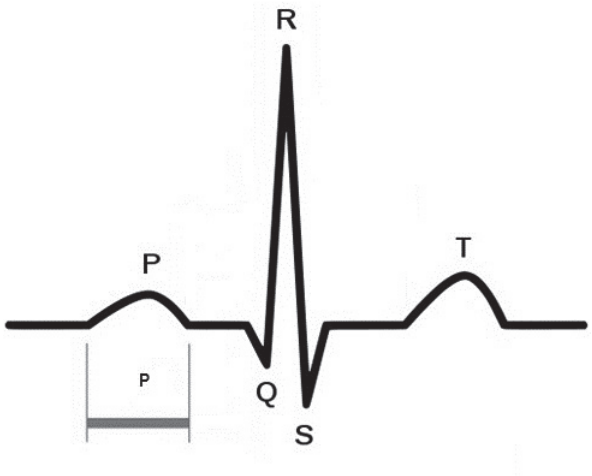


Figure 2.4: The P-wave

over 0.2 seconds, it may indicate a first degree heart block that the impulse travels slower than normal conduction time.

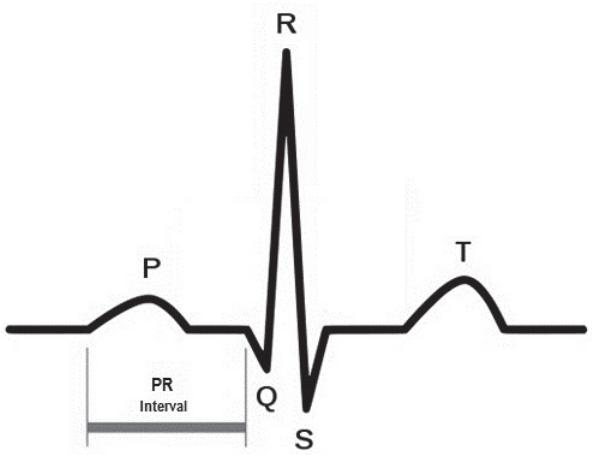


Figure 2.5: The P-R Interval

The QRS Complex

The stimuli are delivered from the Purkinje system to the ventricular muscle, giving rise to the onset of the Q wave occurrence. A fast-moving vector produces the R wave when the depolarization of the ventricular muscle. Then, most of muscle cells are depolarized, producing the peak of the R wave. Stimuli travel toward the base of ventricles when the final phase of the ventricular depolarization occurs, giving rise to

the S wave.

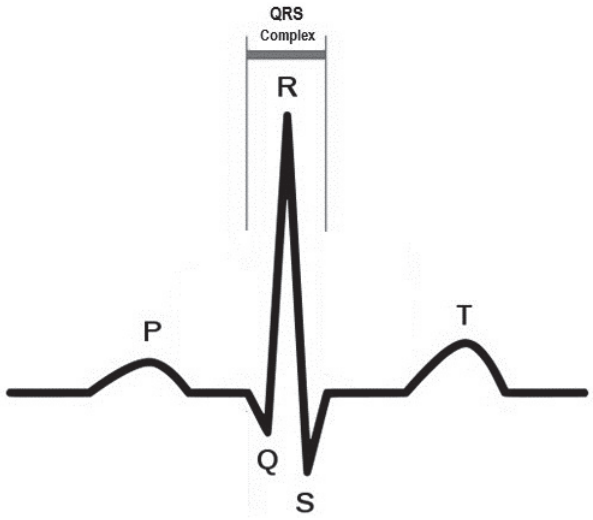


Figure 2.6: The QRS Complex

The QRS complex ranges from 0.06 seconds to 0.1 seconds in duration that consists of three waves: the Q wave, the R wave, and the S wave, which represents ventricular depolarization as shown in Figure 2.6. The duration and amplitude of the QRS complex are used to diagnose cardiac arrhythmias which generate at the ventricle.

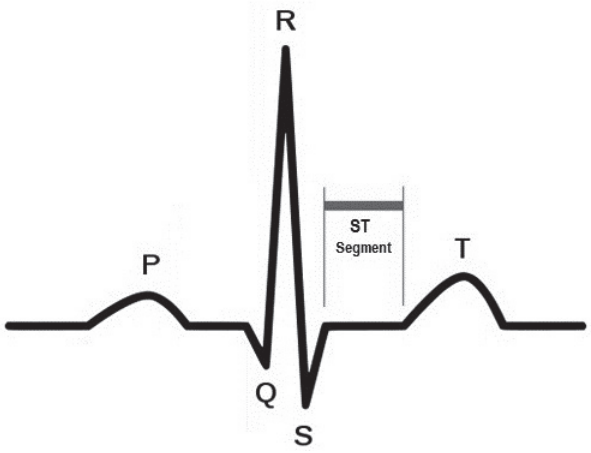


Figure 2.7: The ST Segment

The S-T Segment

The S-T segment is positioned between the end of the QRS complex and the onset of the T wave. The duration range is about 0.08 seconds to 0.12 seconds. According

to the morphology and duration, the S-T segment is analyzed and distinguished whether the heart rate is normal or abnormal. Figure 2.7 shows the S-T segment. For example, if the S-T segment is flat and depressed, it may indicate coronary ischemia. In addition, if the S-T segment is abnormally higher than the isoelectric line, and longer than 0.08 seconds, it may indicate myocardial infarction.

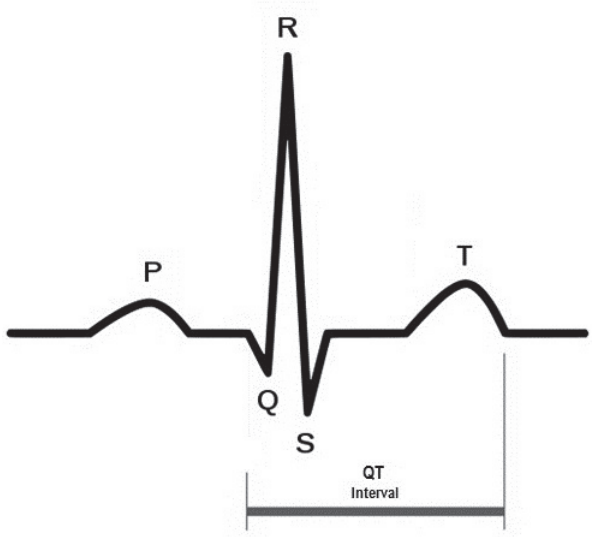


Figure 2.8: The QT Interval

The Q-T Interval

Figure 2.8 shows the Q-T interval that represents the time necessary for ventricular depolarization and repolarization. This interval can range from 0.2 seconds to 0.4 seconds depending upon the heart rate. Ventricular action potentials shorten in duration when the heart rate is high, which decreases the Q-T interval. It is a sign of sudden death when there is a prolonged Q-T interval.

The T wave

The T wave represents the restoration of the ventricles. Figure 2.9 describes a T wave. Identification of the T wave is useful in clinical significance. For example, it may generate coronary ischemia when the T wave is flat. It will be a bio-marker of myocardial infarction when the T-wave is negative in most leads except for aVR lead and sometimes in V1 lead.

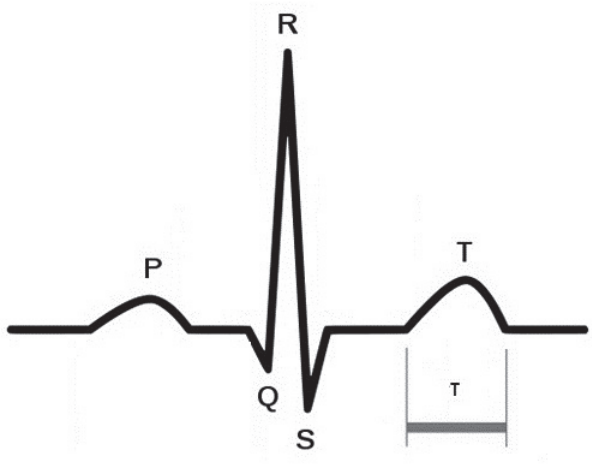


Figure 2.9: The T-wave

2.1.3 ECG Graph Paper

The ECG waveform is recorded onto ECG graph paper along the horizontal axis with time series and voltage represented on the y-axis. This graph paper based on background grid has a standard pattern, which is 1 millimeter (mm) squares representing 0.04 seconds in length with bold divisions 5mm long, representing 0.2 seconds in each larger square as shown in Figure 2.10. With standard ECG graph paper, the ECG waveform runs at 25 millimeter per second (mm/sec) in the horizontal direction, and the voltages are 1 millivolt (mv) (1mv = 10mm) on the y-axis.

According to the ECG graph paper, there are two main methods to calculate the approximate heart rate by predictable ECG paper speed 25mm/sec [43]:

- The six second Tracing Method

  Step 1.Count the number of R waves that appear within that 6 seconds period (30 large squares)

  Step 2.Multiply 10

  For example: Figure 2.11 shows 7 R-waves in the 6 seconds, so the approximate Heart Rate is equal to 70bpm (7*10).

- Large Square Method (This method only works for regular rhythms)

  Step 1.Count the number of large squares between two consecutive R-waves.

  Step 2.Divide this number into 300 for a ventricular rate.

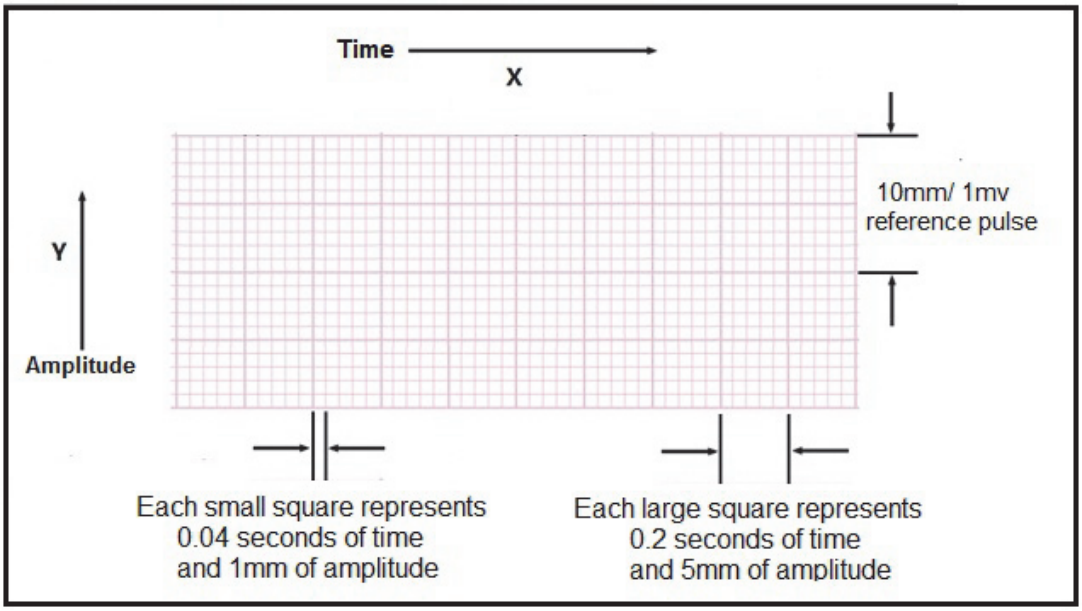


Figure 2.10: Two seconds of ECG paper

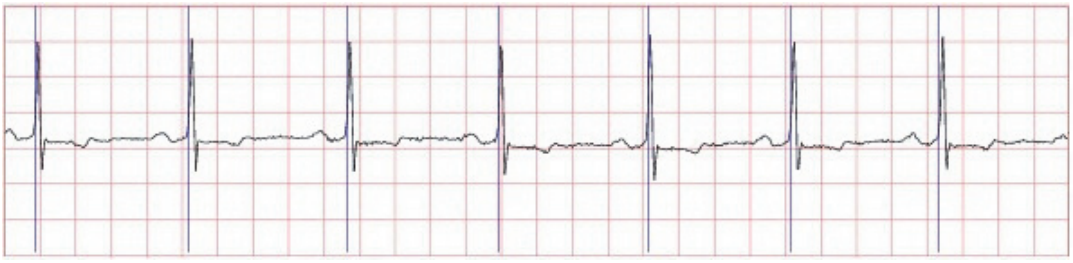


Figure 2.11: Seven R-waves in the 6 seconds

A cycle represents one beat per 0.2 seconds when only one large square is between successive R waves; therefore the heart rate is 300bpm. However, for an atrial rate, count the number of large boxes between two consecutive P waves and also divide into 300.

2.1.4 12-Lead ECG [1]

An electrocardiographic lead is a pair of electrodes that record the electrical potential activity at a specified time and location. In the ECG testing, there are five types of electrodes to be attached to the subject, such as right arm (RA) , left arm (LA) , right leg (RL) , left leg (LL) , and chest (V) showed as Figure 2.12. Normally, there are 12 leads to provide spatial information of the heart's electrical potential activity

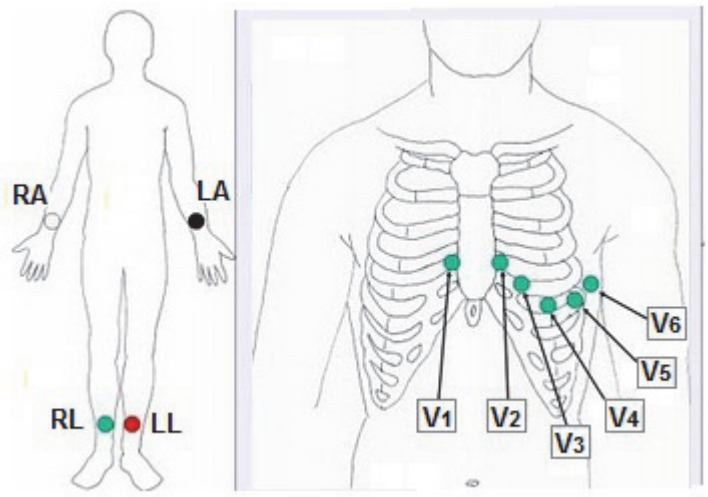
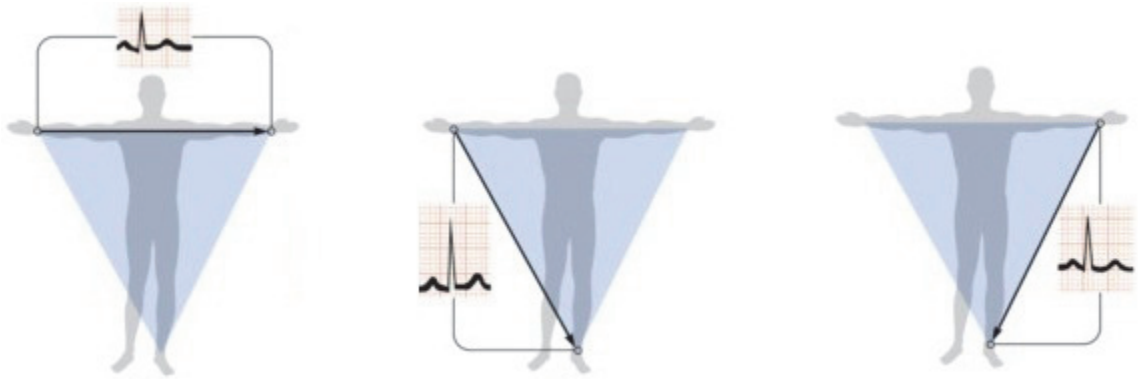


Figure 2.12: Five types of electrodes: RA, LA, RL, LL, and V

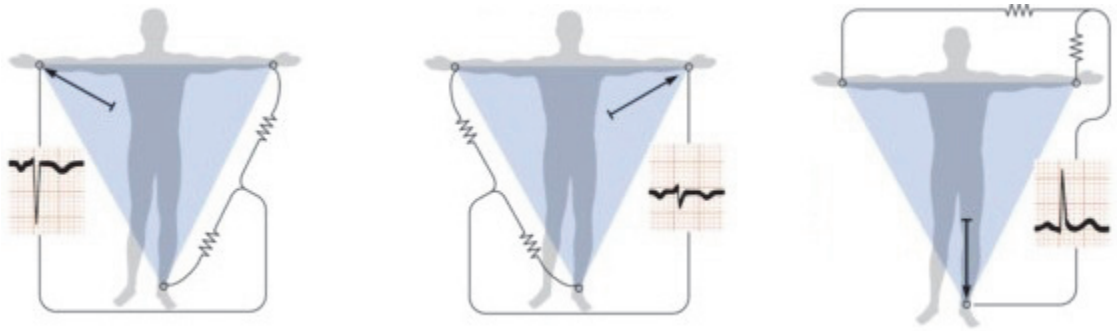
in the ECG signal diagnosis, viewing the surfaces of the heart from twelve different angles. These leads are positioned on the arms, chest and legs, especially, the right leg is used as the ground electrode.

The 12 leads include six limb leads and six chest leads that have two types: bipolar and unipolar. The bipolar lead utilizes a single positive and a single negative electrode to measure electrical potential. The standard leads (Lead I, Lead II, and Lead III) belong to the bipolar lead. The unipolar lead consists of augmented leads (Lead aVF, Lead aVL, Lead aVR) and chest leads (Lead V1 - V6). There are two electrodes in the unipolar lead that include a single positive electrode, and a composite negative electrode that is a combination of the other electrodes. Figure 2.13 illustrates the category of limb leads. The chest leads shows in Figure 2.14.

The three orthogonal planes (sagittal, frontal and horizontal) based on spherical volume conduction are adopted to view the lead vectors and position of the heart, separately as shown in Figure 2.15. Lead II, Lead III and Lead aVF located on the left foot are adjacent leads to view the inferior wall of the left ventricle. Leads I and aVL are positioned on the left arm to present the high lateral wall of the left ventricle. Leads V5 and V6 located on the left lateral chest view the lower lateral wall of the left ventricle. Leads V3 and V4 are placed on the anterior wall of the left chest, viewing the anterior wall of the left ventricle. Leads V1 and V2 are positioned on each side of the sternum to see the right ventricle and the septal wall.



(a) Lead I, Lead II, Lead III



(b) Lead aVR, Lead aVL, Lead aVF

Figure 2.13: The category of limb leads

2.2 Abnormal Rhythm [4]

2.2.1 Definition

Arrhythmias are abnormal problem with the heart rhythm, leading to beating too fast, too slow, or with irregular rhythms. Arrhythmias are caused anywhere along the heart's electrical conduction system. Most arrhythmias are harmless, but some can be serious or even life threatening. It is called bradycardia when the heart rate is too slow under 60bpm in an adult. This means that the adequate amount of oxygen may not be pumping out to the body leading to insufficient blood supply of the body. The heart rate that exceeds the normal heart rate is called tachycardia, which is more than 100bpm. If tachycardia generates in the ventricles, it is called ventricular tachycardia. If it begins in the upper chambers, it is called supraventricular tachycardia. Generally speaking, arrhythmias can be classified into two categories in terms of the occurred location of abnormal heart rhythm, such as ventricular arrhythmias in the lower

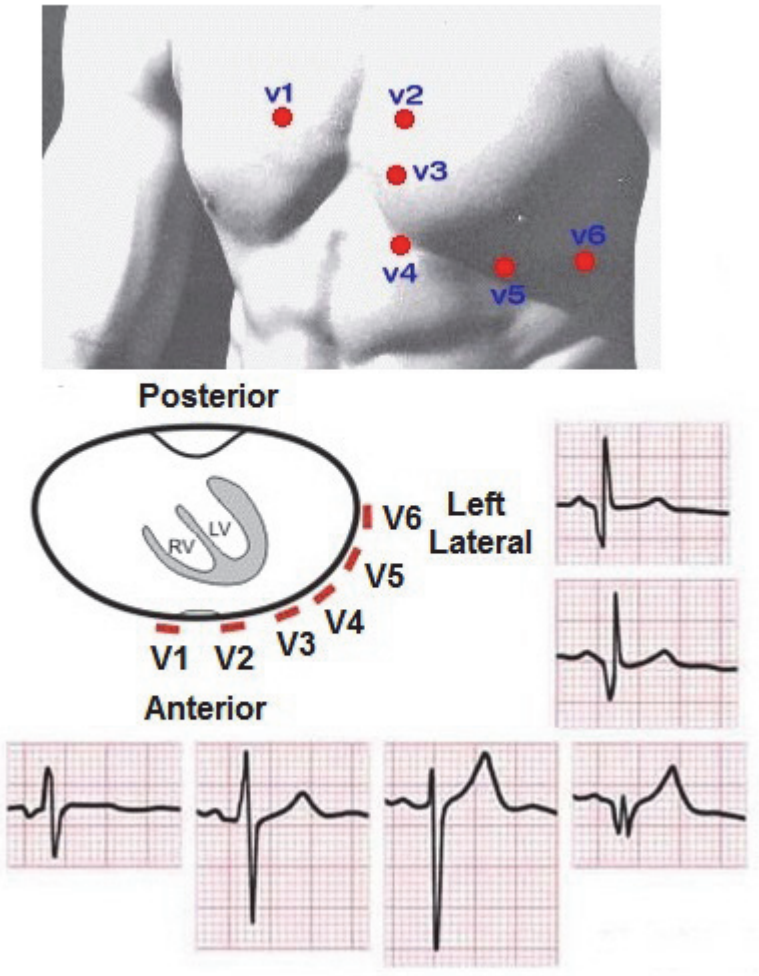


Figure 2.14: Chest Leads

chambers of the heart and supraventricular arrhythmias in the upper chambers of the heart.

2.2.2 Supraventricular Arrhythmias

Supraventricular arrhythmias called "atrial arrhythmias" originate from the upper chambers of the heart (atria) or the atrial conduction pathways. Comparing with ventricular arrhythmias, supraventricular arrhythmias are not serious and sometimes don't need timely treatment. There are several types in supraventricular arrhythmias: Supraventricular Tachycardia (SVT) , Atrial Fibrillation (AF) , Atrial Flutter (AFL) , and Premature Atrial Contractions (PAC) .

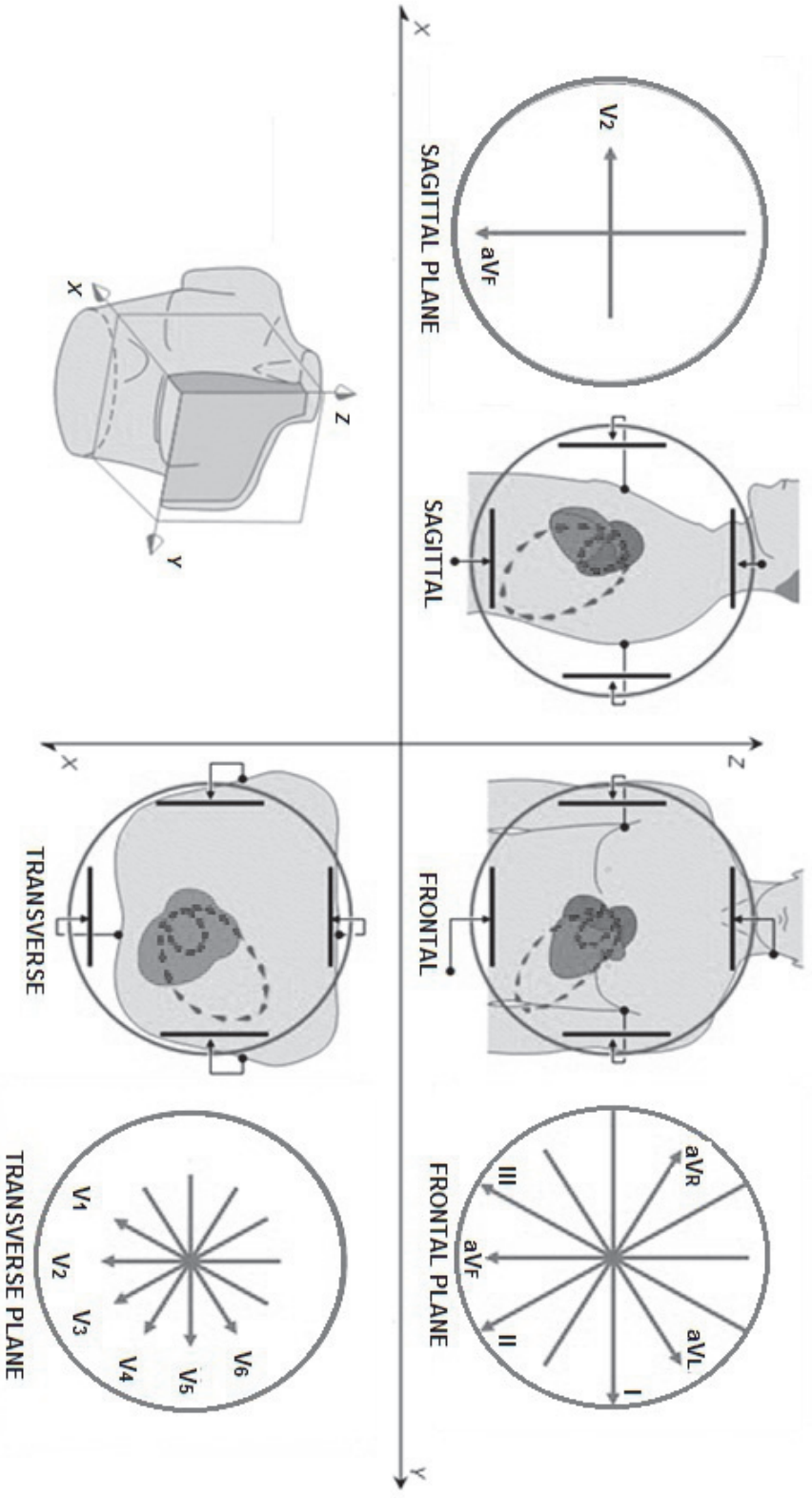


Figure 2.15: Vectors of the 12-lead ECG and heart in three orthogonal planes

SVT is a fast and regular heart rate. The heart beats of SVT are about 150-250bpm in the atria.

AF is a rapid, irregular rhythm of the heartbeat, which may cause blood clots in the heart's upper chambers.

AFL occurs in the atria of the heart. The beats over 100bpm. It happens when the atria beat very fast, causing the ventricles to beat inefficiently as well.

PACs happen when the atria contract earlier than expected, causing the atria to send an electrical impulse fast.

2.2.3 Ventricular Arrhythmias

Ventricular arrhythmias originate in one of the ventricles of the heart. Some of the arrhythmias can cause heart disease, stroke, or sudden death. Ventricular arrhythmias primarily include three types: Premature Ventricular Contraction (PVC) , Ventricular Tachycardia (VT) and Ventricular Fibrillation (VF).

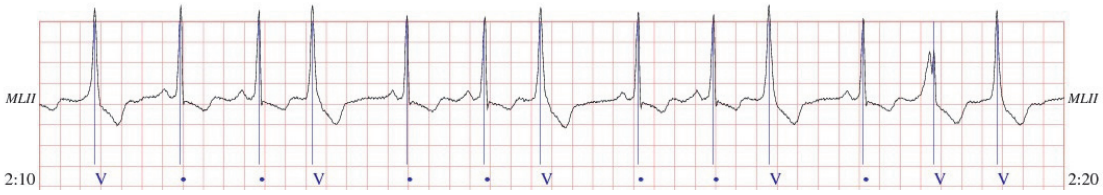
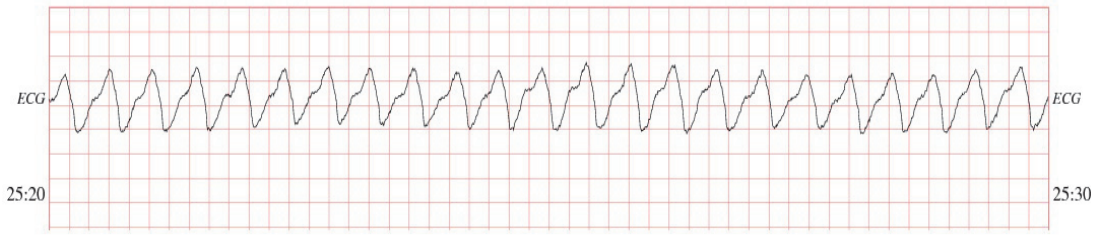


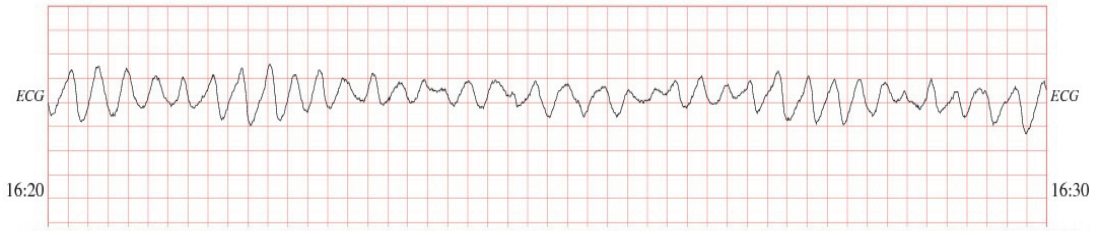
Figure 2.16: Morphology of the some PVCs Note: V means PVC [9]

PVC is extra abnormal heart rhythm generating from ventricles, which mean that blood circulation is inefficient when the ventricles contract too early and out of sequence with the normal beat. Figure 2.16 illustrates the morphology of PVC.

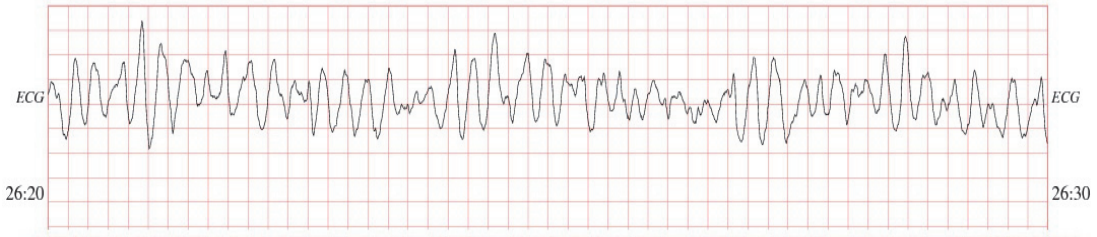
VT is a fast heart rhythm originating in the ventricles, which can be characterized by two types based on its morphologies in Figure 2.17: monomorphic VT and polymorphic VT. Each beat of monomorphic VT would look the same with a regular rhythm and fixed shape on an ECG record. Each beat of Polymorphic VT is irregular in rate and rhythm, and has variable morphologies on an ECG record, which has several combinations. For example, a fixed segment combines monomorphic VT and polymorphic VT, or a monomorphic VT may metamorphose into polymorphic VT. For the patient who has the clinical criteria for VT, the heart rate is normally over 100bpm. When VT exists in a younger healthier heart, it might still work at



(a) Monomorphic VT from Record 420



(b) Polymorphic VT form Record 427



(c) VF from Record 422

Figure 2.17: Three different types of ventricular arrhythmias [9]

a life-sustaining level. When it exists in an older frailer heart, it could be fatal [44]. Normally, people require medical assistance in less than an hour [17].

In addition, Polymorphic VT may lead to VF that is an irregular and potentially fatal beat, uncoordinated series of very rapid, ineffective contractions of the ventricles caused by many chaotic electrical impulses. Sometimes, VF is much faster, reaching 300bpm, called chaotic heartbeat, which means that blood is difficult to be pumped into the brain and body from the heart, resulting in fainting. Figure 2.17(c) shows the ECG rhythms recording VF.

PVCs can bring about serious VT when PVCs exceed two heartbeats in a row. In some cases, VT may degenerate into VF. VF is a sudden lethal arrhythmia, which has been considered a chaotic state in terms of morphologies and may lead to death in less than 3 minutes [17]. Therefore, medical assistance is required immediately for VT and VF arrhythmias. If patients with VT or VF arrhythmias can obtain timely treatment,

their life-threatening arrhythmias can be converted to normal sinus rhythms (SRs) . VT and VF manifest different morphologies that can be studied to understand the pathological changes and biological mechanisms of deadly CVD. The observed signal of the chaotic state is depicted using a nonlinear state function.

2.3 ECG Analysis Algorithms

ECG tele-monitoring analysis systems have received considerable attention over the last few years with rising interest in E-health application of cardiac arrhythmia for improving patient care and access in the public health system. In ECG signal analysis, the QRS complex represented the ventricular activity of the heart is the most remarkable waveform for cardiac event detection in ECG signal analysis, which has high potential amplitude, steep slope (R-wave) and wide duration. These features can be extracted as a characteristic quantity, and as a quantized standard in the analysis process. Currently, there are four classic types of algorithms for detection and classification of the QRS complex [45, 46]:

1. Time-domain analysis
2. Wavelet transform analysis
3. Syntax analysis
4. Neural network analysis

The time-domain analysis is a rapid detection method implemented simply, but it is noise-sensitive. The wavelet transform [47] method has very good time-frequency transformation and partial analysis ability. However, it has huge computational overhead. The neural network [48] based on the sample distinction method receives the sample adaptiveness and the restriction of training time. It is quite bad in the adaptable performance when facing the special case or special profile, and this method has expensive computational complexity. Although some of algorithms can immensely improve the detection and classification accuracy, such as Wavelet Transform [47], Neural Network Analysis [48], Syntax Analysis [49], Genetic Algorithms (GA) [50], Hibert Transform [51] , Mathematical Morphology Method [52]. These, however, generally have huge computation overhead, more resource consumption and less operation efficiency. Furthermore, among these algorithms mentioned above, most of

the presented algorithms were tested against their own database rather than a standard database, which makes the results difficult to be compared and evaluated. A reliable and robust algorithm will considerably decrease the mortality of arrhythmic patients. Meanwhile, it is necessary to accurately classify arrhythmias, especially like as the life-threatening cardiac arrhythmias, VT and VF. In other words, these arrhythmias should be notified immediately to the remote Emergency Health Center for timely rescue when detecting VT or VF. However, few works have taken this into consideration before based on Smart-phone.

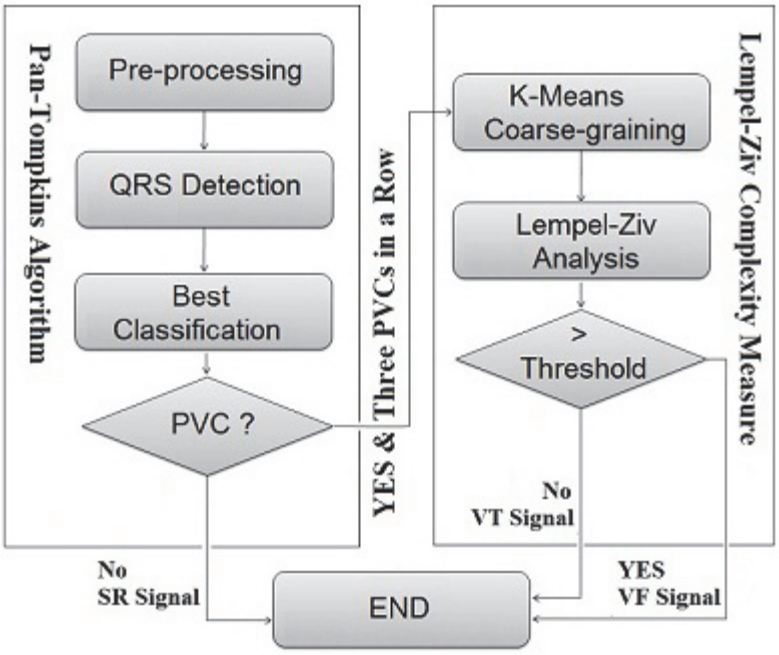


Figure 2.18: Architecture flow graph of ECG signal analysis

From what has been exposed above, it is feasible to adopt an architecture of combining time-domain algorithms on the development of a prototype that demonstrated the feasibility of using available a Smart-phone as a remote cardiac monitoring device when some weak points are overcome by learning from each other's strong points above four main types of algorithm. In this thesis, the proposed algorithm architecture is fast, low resource consumption and computational effectiveness for detection and classification of the ECG signal, which is adopted to develop smaller and more cost-efficient ECG signal analysis system based on a Smart-phone. As to the aspect of ECG signal analysis, this work applies QRS complex detection algorithm suggested

by Pan-Tompkins [37], beat classification methods introduced by Hamilton [5], and arrhythmias classification algorithm presented by Lempel and Ziv [38]. To begin with, the QRS complex is detected by the Pan-Tompkins algorithm and classified as normal SRs or PVCs by existing Hamilton classification methods. subsequently, the LZ complexity measure that includes the K-Means clustering algorithm and the LZ complexity analysis is utilized to further separate the high risk arrhythmias, VT or VF. In this procedure of the high risk arrhythmias, three consecutive PVC beats in a row are considered to be an indication of the beginning of VT rhythms, at which point the following data points will be saved until up to a certain window length long are reached. The window length long ECG signal will be further classified as VT or VF from continuous and mixed ECG signal by several new decision rules with heart rate detection. Furthermore, the proposed system successfully implemented on a Smart-phone adopts the time frames to indicate the analysis report for improving the reliability and error detection of arrhythmias. Figure 2.18 shows the proposed ECG analysis system. Therefore, the abnormal situation doesn't need to be transmitted into Health Service Station for further analysis when detecting and classifying VT and VF after a period of time. In other words, when several life-threatening arrhythmias are detected and classified, the Smart-phone will automatically send an alert message to the remote Health Service Station for the timely rescue. This ECG signal analysis architecture will be efficiently implemented on kinds of ECG tele-monitoring analysis systems, improving users' quality of life as well as save governments from expending huge amounts of money on hospital and aged care facilities.

Chapter 3

Feature Extraction

Feature extraction plays an important role to refine some special properties within the ECG signal, which would be accurately detected and systematically classified kinds of heart rhythms. In the ECG signal, P-QRS-T waves describe on a cardiac beat cycle. Figure 3.1 illustrates diagnostic and morphologic features of a normal heartbeat in ECG graph paper. The nature of the P-QRS-T waves consists of two fundamental values, interval and amplitude. The interval describes time duration; the amplitude explains the peak of the characteristic wave peak. These feature values were extracted from the ECG signal for subsequent analysis. In this chapter, a time-domain approach was addressed to represent the ECG diagnostic and morphologic features, which makes accurate and robust extraction of partial features. This chapter focuses on the feature vectors of the QRS complex analysis with time series.

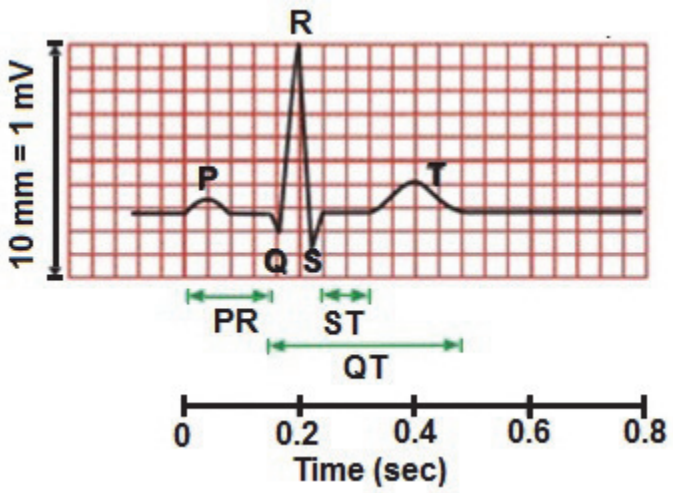


Figure 3.1: A Standard P-QRS-T wave in ECG paper

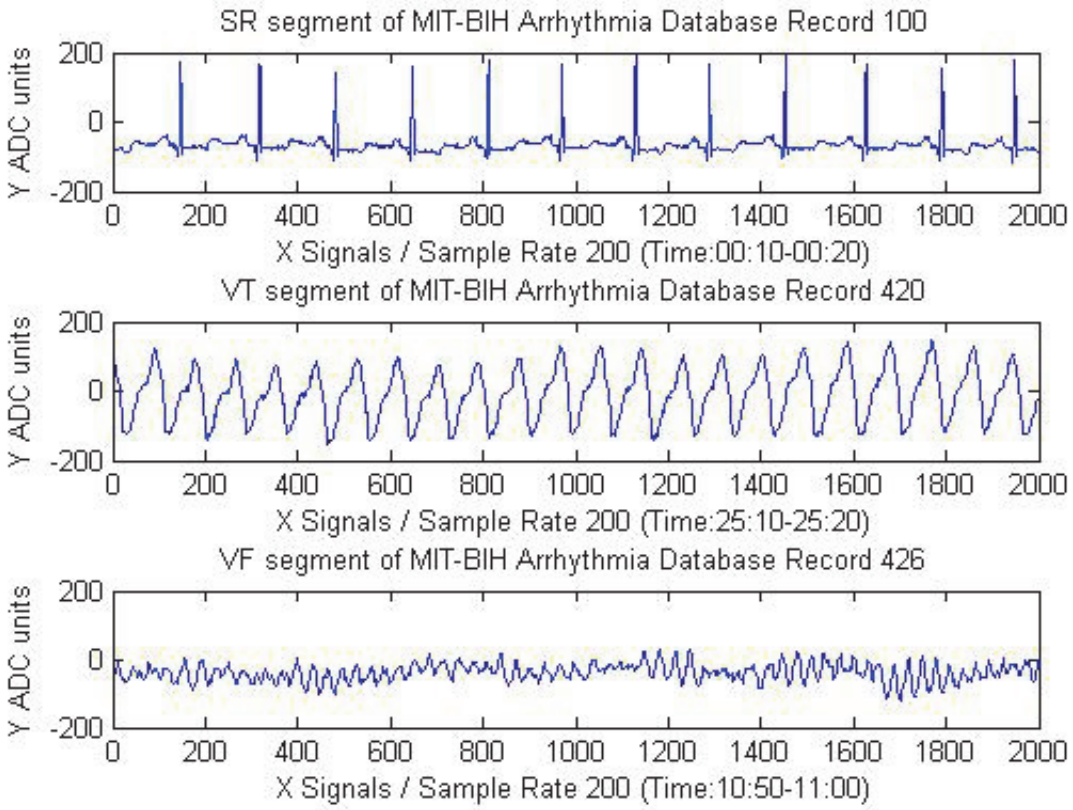


Figure 3.2: The original SR, VT and VF segments in 10 seconds

3.1 Characteristics and Analysis

In order to express the morphologies of the original ECG signal, several segments of different types are selected from MIT-BIH database [9] to analyze the spectrum by using the Fast Fourier Transform (FFT) . Figure 3.2 shows the original SR, VT and VF segments in 10 seconds. The sample rate of all segments is down-sampled from the raw sample rate to 200Hz (hertz) because of comparison and unification each other.

The original ECG signal contains many types of noise from baseline interference, motion artifact and interface, and so on. It is necessary to retain the interested signal from these noise sources. The normal ECG signal is quasi-periodic and is very difficult to identify the frequency components. Normally, the spectral analysis method is utilized to depict the frequency of the ECG signal. The power spectrum based on the

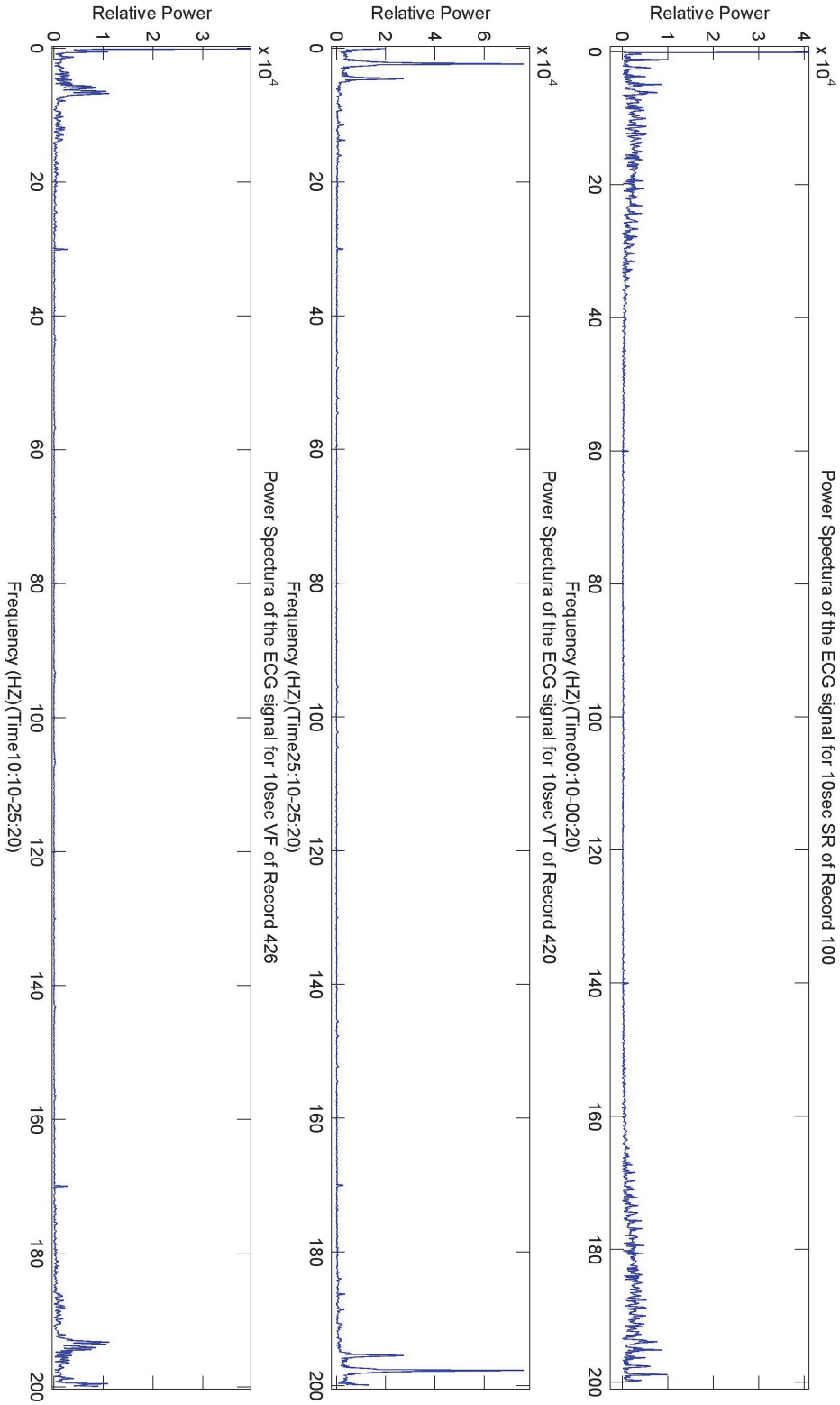


Figure 3.3: Power spectra of the ECG signal with 10 second window length

FFT can give a visual impression of different ECG signal. The ECG signal is converted into an interpretation of the frequency spectrum of the QRS complex, showing the frequency components by using the FFT. Figure 3.3 illustrates the frequencies of sampled signals from 0 to 200 with kinds of peaks at different frequencies of the sinusoid. The results show that the power spectra is symmetrical, and frequencies above 100Hz are Nyquist frequency.

In modern, multiple filters are offered in ECG monitors for ECG signal processing. There are two modes for the most devices, monitoring mode and diagnostic mode [53]. In the monitoring mode, it has a filter with a pass-band of 0.5-40Hz [53]. This pass-band filter includes low-pass filter and high-pass filter. The filter is used to narrow the bandwidth in an attempt to reduce the contaminants from baseline drift, muscle movement, power line interference, and electrode contact noise. For example, the high-pass filter can efficiently remove wandering baseline; the low-pass filter is used to reduce 50Hz or 60Hz power line interference. Diagnostic mode is used to monitor the ST-segment and the frequency range of the filter is 0.05-100Hz [53]. The typical band of frequency for VT and VF is below 12Hz [54]. The SR signal has a wider frequency band, but mostly below 20Hz [55]. The diagnostic mode is less filtered than monitoring mode in ECG display, because its pass-band is wider [56]. As there is no need to monitor the ST-segment, this system works in the monitoring mode to estimate the ECG signal. Therefore, the range of frequency is below 40Hz as shown in Figure 3.4. This study further narrows the range of the band-pass filter down to 0.6-22Hz, to maximize the energy of the signal that is of interest (SR, PVC).

3.2 ECG Signal Pre-processing

The pre-processing steps are used to improve the signal-to-noise ratio (SNR) of the ECG signal, making the signal analysis much more accurate and effective. In the original ECG signal, some types of noises overlap the ECG signal to disturb the measurement of ECG, leading to inaccurate analysis and diagnosis. For instance, body movements cause baseline wander of low-frequencies; mains interference (50Hz, 60Hz) and muscular activity lead to random noises of high-frequencies; body movements and poor electrode contact cause random shifts of the ECG signal amplitude [57]. The ECG signal pre-processing in this case refers to applying linear and nonlinear methods

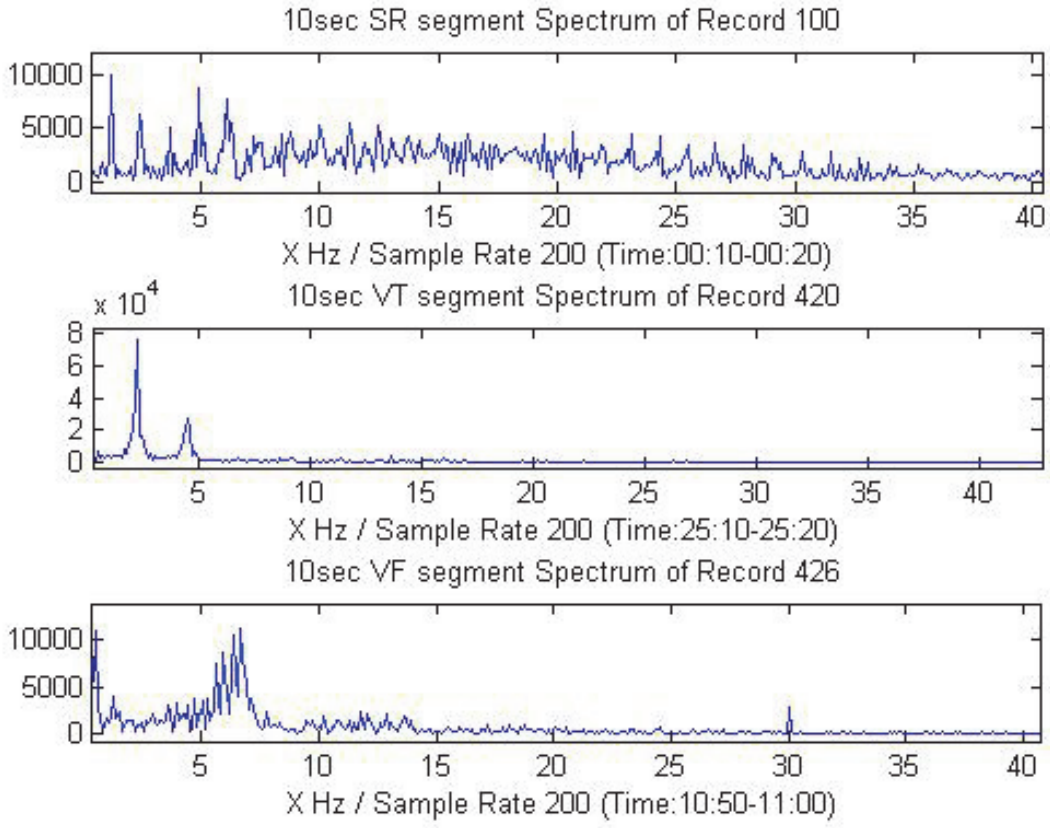


Figure 3.4: Three types of original signal in the range of 0 to 40 Hz

to improve the SNR and remove all the noise and artificial factors, which may manifest with similar morphologies as the QRS complex (the main beat morphology in the ECG waveform). Linear processes include a band-pass filter (low-pass and high-pass), a derivative, and a moving window integrator [58]. The nonlinear transformation uses the absolute value method to make the signal amplitude positive. Figure 3.5 shows the procedures of the ECG signal pre-processing. After the pre-processing, the QRS complex morphologies can be clearly manifested. The function of each step will be described in detail as follows.

3.2.1 Filtering Techniques

The filter can efficiently remove most of the noises to improve the SNR. Lots of filtering designs were unsatisfactory regarding the huge of computation time. A feasible

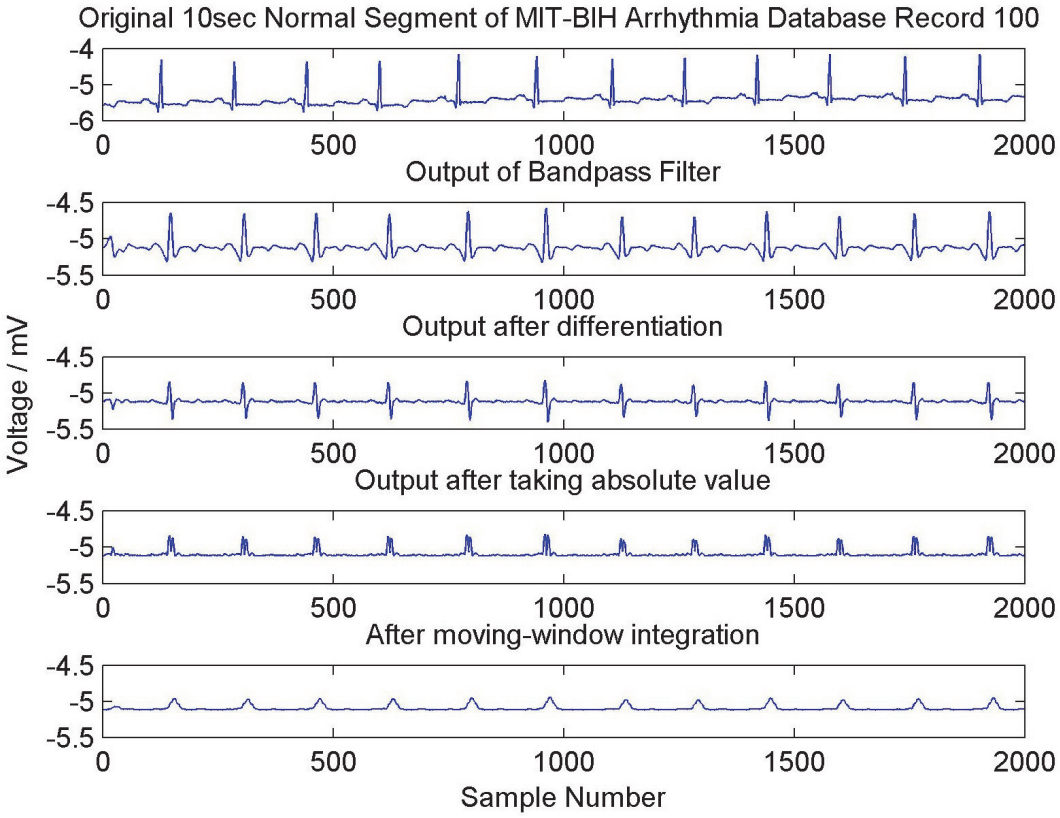


Figure 3.5: The pre-processing steps of the QRS complex

integer filter can reduce the computational time with integer-coefficients instead of floating-point coefficients in a filter equation. Lynn [59], represented the techniques of integer filter for designing band-pass filters (low-pass filter and high-pass filter) with integer coefficients.

A general formula of transfer function based on the integer-coefficients is presented as follows:

$$H(z) = \frac{(1 - z^{-m})^p}{(1 - 2 \times \cos(\theta) \times z^{-1} + z^{-2})^t} \tag{3.1}$$

The m exponent represents that the number of zero is located on the unit circle. The angle θ represents that the angular locations of the poles place around the unit circle. The powers p and t describe the order of magnitude.

Low-pass Integer Filters [57]

A low-pass filter means that the high frequencies are rejected and low frequencies can pass it. According to the general formula mentioned above, a low-pass filter can be obtained based on cutoff frequency 22Hz.

This transfer function of the second-order integer is

$$H(z) = \frac{(1 - z^{-3})^2}{(1 - z^{-1})^2} = \frac{Y(z)}{X(z)} \tag{3.2}$$

The difference equation is used to express the low-pass integer filter:

$$y(nT) = 2y(nT - T) - y(nT - 2T) + x(nT) - 2x(nT - 3T) + x(nT - 6T) \tag{3.3}$$

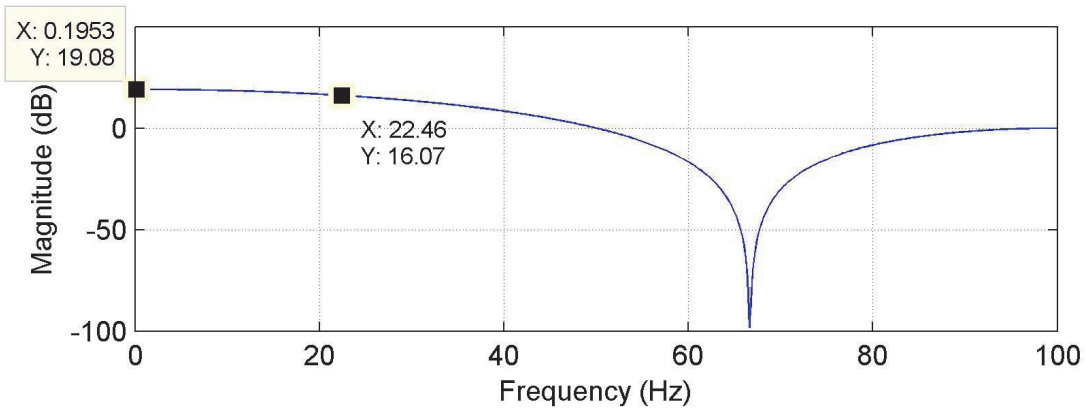


Figure 3.6: Pole-zero and Amplitude-Frequency of the low-pass filter

where the T is the delay of a sample cycle. Figure 3.6 illustrates that this cutoff frequency of low-pass filter is about 22Hz when the pass band is 3dB (Decibel) . The gain is 9 gains (19dB). The Filter has a delay of 2 samples (10ms (millisecond)).

High-pass Integer Filters [57]

In order to obtain the low cutoff frequency 0.6Hz, the high-pass filter can be implemented by two filters, an all-pass filter and a low-pass filter. This low-pass filter has a dc gain of 256 and a group delay 127.5 samples (i.e., 637.5ms).

$$H(z) = \frac{1 - z^{-256}}{1 - z^{-1}} \tag{3.4}$$

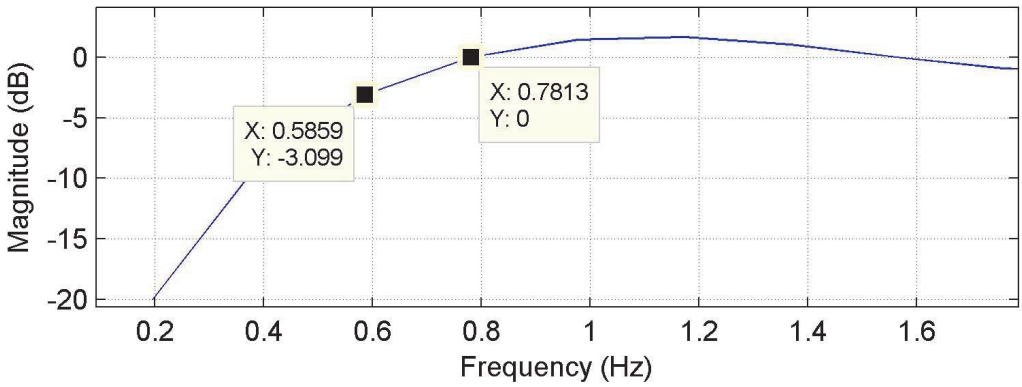


Figure 3.7: Pole-zero and Amplitude-Frequency of the high-pass filter

The all-pass filter has a delay of 128T (i.e. ,z^{-128}) for the original signal, which can be utilized to offset the delay of low-pass filter.

$$H(z) = z^{-128} \tag{3.5}$$

To produce the high-pass filter, the output of the low-pass filter is divided by its dc gain and subtracted from the delayed all-pass filter.

The transfer function of the high-pass filter is derived from:

$$H(z) = \frac{-\frac{1}{256} + z^{-128} - z^{-129} + \frac{z^{-256}}{256}}{1 - z^{-1}} \tag{3.6}$$

Figure 3.7, shows that this cutoff frequency of high-pass filter is about 0.6Hz when the pass band is 3dB. The filter has a delay of 128 samples (640ms).

$$y(nT) = y(nT - T) - \frac{x(nT)}{256} + x(nT - 128T) - x(nT - 129T) + \frac{x(nT - 256T)}{256} \tag{3.7}$$

The band-pass filter (0.6Hz-22Hz) adopted in this ECG analysis system is used to reduce the influence of muscle and electrode contact noise, 50Hz or 60Hz interference, baseline drift, and T-wave interference by the monitoring mode.

3.2.2 Derivative Technique

The QRS complex has a visual sign with high slope that can be efficiently distinguished from other ECG waves by derivative technique. After band-pass filter, a high-pass filter is used to achieve the differentiated function that can amplify the slope characteristic of the QRS complex as shown in Figure 3.8.

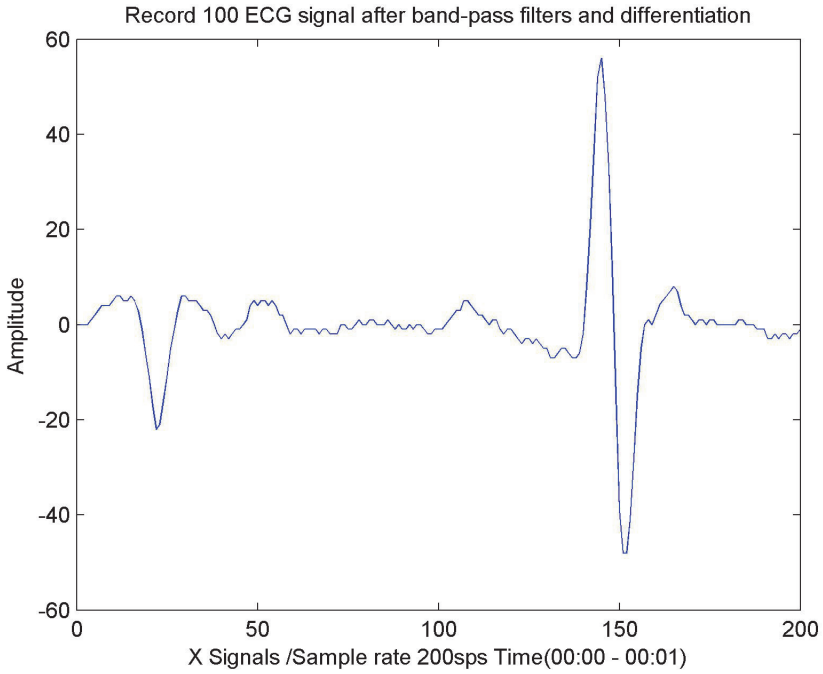


Figure 3.8: The ECG signal after differentiation.

This transfer function of the first-order integer is

$$H(z) = 1 - z^{-2} \tag{3.8}$$

The difference equation is used to express the high-pass integer filter that has a delay of 2 samples (10ms):

$$y(nT) = x(nT) - x(nT - 2T) \tag{3.9}$$

3.2.3 Absolute Function

After differentiation techniques, this part is a nonlinear operation that makes all of data points in the positive, emphasizing the higher frequencies characteristic of the

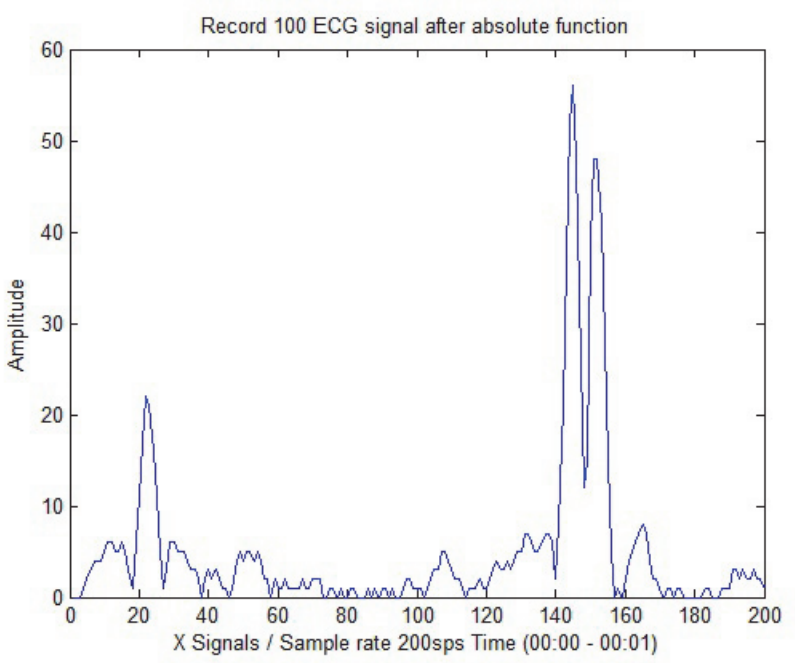


Figure 3.9: The ECG signal after absolute function.

QRS complex. Figure 3.9 illustrates the output of this stage. The operation can be implemented by:

$$y(nT) = |x(nT)| \tag{3.10}$$

3.2.4 Moving window integral

Normally, the R wave has a steep slope which is not a sole standard to guarantee the detection of the QRS complex. The arrhythmias are difficult to be detected by using the R wave only when the QRS complexes have largr amplitude and long duration. Therefore, moving window integrator is used to refine some kinds of features except the R wave. Figure 3.10 shows the ECG signal output of the moving window integration. According to the width of the moving window, the N indicates the number of samples. It is implemented by following different equation [57]:

$$y(nT) = \frac{1}{N}[x(nT - (N-1)T) + x(nT - (N-2)T) + \ldots + x(nT)] \tag{3.11}$$

Normally, the width of window should be similar with the widest QRS complex.

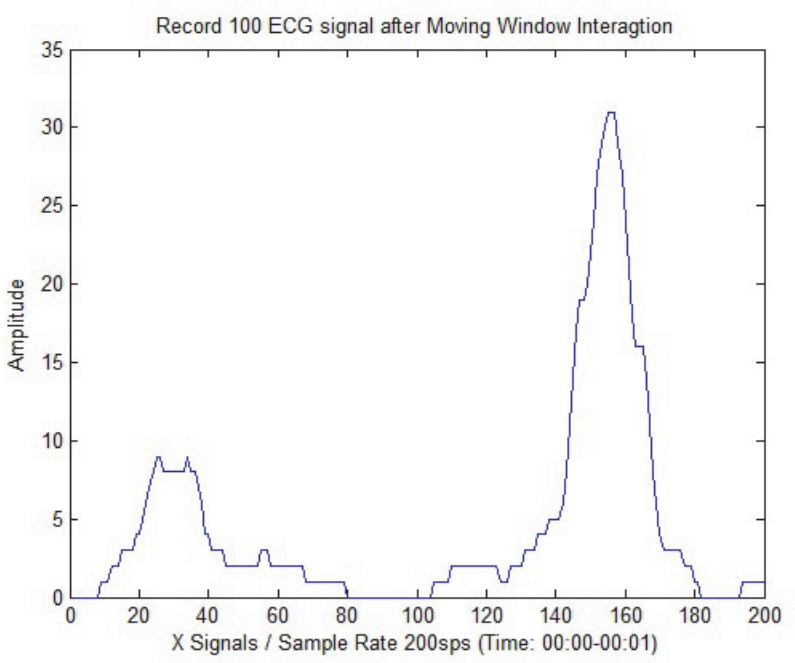


Figure 3.10: The ECG signal after moving window integration

This parameter should be chosen carefully. For example, if the width is chosen too large, it will make the waveform merging the QRS complex and T-wave together. In addition, On the other hand, if the width is too small, it will have several peaks at the output with in a QRS complex. The width of window is set to 0.08 seconds (16T) based on the QRS complex ranges presented at Chapter 2.

3.3 QRS Complex Classification [5]

There are two characteristics in the feature vectors: peak height and location of peak occurrence. In order to accurately recognize the feature resulted from a QRS complex or noise, both vectors are used to evaluate and analyze the QRS complex. Therefore, lots of decision rules should be established to decide the feature vectors, and then a peak can be efficiently classified as either a QRS complex or noise, or saved for later classification.

After the moving window integrator, the filtered signal can be obtained to detect a peak. In [5], Hamilton presented several rules to accurately detect the peak for recognizing a QRS complex of the ECG.

Chapter 4

Arrhythmias Classification

In ECG signal analysis, the classification of abnormal heart rhythms plays an important role in diagnosis and measurement of cardiovascular disease (CVD). Specifically, there are two high-risk arrhythmias, ventricular tachycardia (VT) and ventricular fibrillation (VF) that can be studied to understand the pathological changes and biological mechanisms of deadly CVD. Monomorphic VT is a periodic motion that manifests a rapid heart rate, giving rise to a diminished cardiac output when it occurs. Polymorphic VT is irregular and variable morphologies, which looks somewhere betwixt and between the monomorphic VT and the VF. VF has been considered a chaotic state (a random and irregular process) in terms of its morphologies. The observed signal of the chaotic state is depicted by using a nonlinear state function. In nonlinear signal processing, Lempel and Ziv (LZ) [38] presented a practicable approach based on a coarse-graining process to estimate the randomness of finite symbolic sequences for information data processing [60], which cope with a dynamical system entering a chaotic state by quantifying the rate of new pattern occurrences along given finite symbolic sequences. Consequently, the LZ complexity measure can efficiently separate the VT and VF based on their chaotic or periodic characteristics.

In the past decade, the LZ complexity measure (which offers relevant mathematical definitions and detailed derivations) has been extensively applied in biomedical signal analysis to evaluate the complexity of physiologic signals in discrete-time [61]. For example, Electroencephalogram (EEG) complexity measure in the depth of anesthesia [62, 63] and classification of ECG signal [35, 64]. As the features of arrhythmias mentioned above, the LZ complexity measure with Mean-value coarse-graining process has been addressed by previous research that the distributions of VT and VF could be exactly separated in their own database [35, 64].

The LZ complexity measure includes several steps for separating the arrhythmias. Firstly, the original ECG signal is collected from an ECG monitor device. Then, the

ECG signal of an appropriate window length should be converted to a finite symbolic sequence with the coarse-graining technique. Lastly, the LZ complexity analysis is used to analyze the complexity by counting the number of new patterns in the given finite symbolic sequence scanned from left to right. In this procedure, the coarse-graining process determines how much inherent information can be retained and will consequently impact the separation of VT and VF. Although the original ECG signal loses a large amount of information in the process, the inherent features of the heart's dynamics are reserved, and robustness to noise is guaranteed. In this chapter, there are four methods (K-Means clustering algorithm, Mean-value algorithm, Mid-point algorithm and Median algorithm) to be utilized in coarse-graining process aiming at gaining a better understanding of their impact on the classification of ECG signal.

4.1 The Coarse-graining Process

The coarse-graining process has been used successfully to quantify the original ECG signal in a given finite symbolic sequence, where partitioning of the time series into a symbolic sequence can be done with various methods of information extraction in the information theoretical problems. With a selected window length, the finite original ECG signal is transformed into a symbolic sequence corresponding to consecutive non-overlapping time intervals of size δ. Normally, there are several types based on symbolic dynamics to transform data into symbolic characterization in nonlinear dynamics [65].

1. Static Transformation

Each data point x_i with time series is transformed into the binary symbols by means of a certain threshold X as follows (Figure 4.1(a)). In other words, as to this static transformation, assume there is a threshold value X, each data point value will be compared with X, and then to be replaced with a specific binary value accordingly.

$$s(i) = \begin{cases} 1, if\ x_i > X \\ 0, otherwise \end{cases} \tag{4.1}$$

2. Dynamic Transformation

Comparing the current data point with previous data point, original signal is extracted into a binary value. In brief, adjacent data are considered to construct a

binary sequence (Figure 4.1(b)).

$$s(i) = \begin{cases} 0, if\ x_i > x_{i-1} \\ 1, otherwise \end{cases} \tag{4.2}$$

3. Mixing of transformation

It is particular important to choose the types of transformations when the original signal is transformed into a symbolic sequence. Generally speaking, due to different conditions, the original signal is refined by using different transformation method. Therefore, the mixed transformation method can efficiently extract interior information based on different requirements.

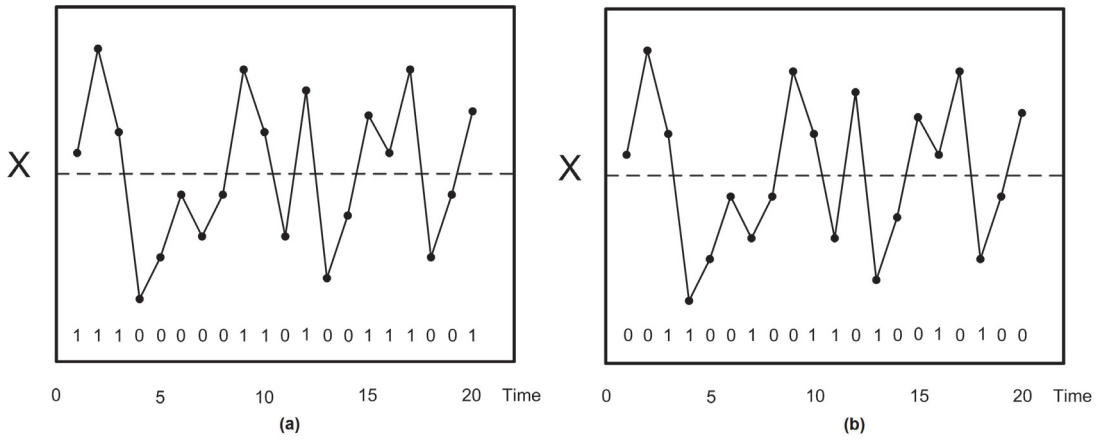


Figure 4.1: (a) Static transformation (b) Dynamical transformation

The coarse-graining process as a crucial part in the LZ complexity measure directly affects the separating performance of VT and VF in ECG signal analysis. In this following section, there are four coarse-graining process approaches to be presented, K-Means clustering, Mean-value, Median and Mid-point. Different coarse-graining process methods are used to fully understand how they affect the ECG signal accurate classification.

4.1.1 The K-Means Clustering

The K-Means clustering algorithm is a conventional pattern recognition algorithm presented in [66]. This algorithm has been successfully used to analyze the EEG signal [67]. According to the state characteristics of arrhythmias, the algorithm is

used to refine the inherent features from the ECG original signal, which is flexible and useful to classify using a certain number of clusters. In other words, the algorithm can divide the given data into several clusters based on their features. There are just as many centoids to be defined for a certain number of clusters (K clusters). The core idea is to define the number of centroids. The algorithm is presented as following procedures:

Step1. To choose K, which means that how many the initial cluster centers are represented by $z_1(1)$, $z_2(1)$,...,$z_k(1)$. Usually the first K samples of the given sample set are selected as the initial cluster centers.

Step2. The samples x are distributed among the K cluster domains using the relation as follow when the iterative step is at the kth.

$$X \in S_j(k) \; if \; \|X - z_j(k)\| < \|X - z_i(k)\| \tag{4.3}$$

for all $i = 1, 2, 3, \ldots, K, i \neq j$, where $S_j(k)$ represents the set of samples whose cluster center is $z_j(k)$.

Step3. According to the results of Step 2, the new cluster center $z_j(k+1), j = 1, 2, \ldots, K$, can be computed based on the different set of samples, minimizing the sum of the squared distances from all points in $S_j(k)$ to the new cluster center. In other words, compute the new cluster center $z_j(k+1), j = 1, 2, \ldots, K$, getting the performance index minimized. The performance index can be shown as follow:

$$J(i) = \sum_{X \in S_j(k)} \|X - z_j(k+1)\|^2, j = 1, 2, \ldots, K \tag{4.4}$$

The $z_j(k+1)$ is the sample mean of $S_j(k)$. Therefore, the new cluster center can be defined by

$$z_j(k+1) = (\frac{1}{N_j}) \sum_{X \in S_j(k)} X, j = 1, 2, \ldots, K \tag{4.5}$$

where N_j is the number of samples in $S_j(k)$.

Step4. The procedure is terminated when $z_j(k+1) = z_j(k)$ for $j = 1, 2, 3, \ldots, K$,which means that the algorithm has converged. Otherwise go to Step 2.

From what has been presented above, the ECG signal can be transformed into a binary sequence in terms of the K-Means algorithm that can be described as follows:

1. Select a window length (WL) in seconds. This WL is the set of samples. For example 5, 6, 7, 8 ..., or 10 (seconds).

2. Transmit the ECG signals into a binary sequence, let $K = 2$. The initial cluster centers $z_1(1)$, $z_2(1)$ are adjusted tinily, which is produced by a quantity of center (assume $\varepsilon = 0.01$): (x_m+($\varepsilon \times x_m$) and x_m-($\varepsilon \times x_m$)), where x_m is the mean value of the data points $\{x_i \mid i = 1, 2, 3, \ldots, n\}$ within the selected window length and $n = WL \times SampleRate$. It can be estimated as

$$x_m = (\frac{1}{n}) \sum_{i=1}^{n} x_i \tag{4.6}$$

3. The data points $\{x_i \mid i = 1, 2, 3, \ldots, n\}$ in the samples x are distributed among the K cluster domains using the relation as follow when the iterative step is at the kth.

$$\begin{aligned} x_i \in S_2(k) \text{ and } D_{Xi} = 1 \quad if \quad \|x_i - z_2(k)\| < \|x_i - z_1(k)\| \\ x_i \in S_1(k) \text{ and } D_{Xi} = 0 \quad if \quad \|x_i - z_2(k)\| < \|x_i - z_1(k)\| \end{aligned} \tag{4.7}$$

where $S_1(k)$ represents the set of samples whose cluster center is $z_1(k)$, $S_2(k)$ denotes the set of samples whose cluster center is $z_2(k)$ and D_{Xi} represents the relevant data point transformed into binary symbol.

4. Compute two new cluster centers: knowing the two sets of samples $S_1(k)$ and $S_2(k)$. Group 1, the center $z_1(k+1)$ is the average coordinate among all of members in the $S_1(k)$. Group 2, the center $z_2(k+1)$is the average coordinate among all of members in this $S_2(k)$.

5. Go to step 3 and repeat the above procedure until $z_1(k+1) = z_2(k)$ for $k = 1, 2, 3, \ldots,$

4.1.2 The Mean-value

Zhang et al. [36] utilized a method based on Mean-value to extract the original ECG signal for coarse-graining process.

1. Select a WL in seconds. This WL is the set of samples. For example 5, 6, 7, 8 ..., or 10 (seconds).

2. The mean value of the data points $\{x_i \mid i = 1, 2, 3, \ldots, n\}$ within the selected window length can be estimated as

$$x_m = (\frac{1}{n}) \sum_{i=1}^{n} x_i \tag{4.8}$$

where $n = SampleRate * WL$.

3. The average x_m is subtracted from every data point $(x_i - x_m)$, and then obtain the positive peak value V_p and negative peak value V_n.

4. Count the P_c and N_c, where P_c represents the number of data x_i $(0.0 < x_i < 10\%V_p$, and N_c represents the number of data x_i $10\%V_n < x_i < 0.0)$.

5. Select the threshold T_d, when $P_c + N_c < 40\%n$, the $T_d = 0$, if $P_c < N_c$, the $T_d = 20\%V_p$, else the $T_d = 20\%V_n$.

6. Finally, compared with the threshold T_d, the data are turned into a binary sequence $S_1, S_2, S_3, \ldots, S_r$. If$x_i < T_d$,$S_i = 0$, otherwise the $S_i = 1$.

These given parameters "10%", "40%" and "20%" were derived by an empirical study of ECG data [36].

4.1.3 The Median

The median value is determined by the entire data set, the finite sequence should be sorted and then select the one in the middle.

1. Select a WL in seconds. This WL is the set of samples. For example 5, 6, 7, 8 ..., or 10 (seconds).

2. Sort the data in ascending order, select the one in the middle

3. To subtract each data point with the median value $(x_i - x_{median})$, and then obtain the positive peak value V_p and negative peak value V_n.

4. Count the P_c and N_c, where P_c represents the number of data x_i $(0.0 < x_i < 10\%V_p$, and N_c represents the number of data x_i $10\%V_n < x_i < 0.0)$.

5. Select the threshold T_d, when $P_c + N_c < 40\%n$, the $T_d = 0$, if $P_c < N_c$, the $T_d = 20\%V_p$, else the $T_d = 20\%V_n$.

6. Finally, compared with the threshold T_d, the data are turned into a binary sequence $S_1, S_2, S_3, \ldots, S_r$. If$x_i < T_d$,$S_i = 0$, otherwise the $S_i = 1$.

As in [36], these given parameters "10%", "40%" and "20%" were used in this method for convenient comparison.

4.1.4 The Mid-point

The Mid-point is used to transform each data point into the binary sequence as follows, which is determined by only two data.

1. Select a WL in seconds. This WL is the set of samples. For example 5, 6, 7, 8 ..., or 10 (seconds).

2. The mid-point value of the data points $\{x_i \mid i = 1, 2, 3, \ldots, n\}$ within the selected window length can be estimated as

$$x_{m-point} = (x_{min} + x_{max})/2 \tag{4.9}$$

where $n = SampleRate * WL$, the x_{min} is the smallest value, and the x_{max} is the largest value.

3. To subtract each data point with the mid-point value $(x_i - x_{m-point})$, and then obtain the positive peak value V_p and negative peak value V_n.

4. Count the P_c and N_c, where P_c represents the number of data x_i $(0.0 < x_i < 10\%V_p$, and N_c represents the number of data x_i $10\%V_n < x_i < 0.0)$.

5. Select the threshold T_d, when $P_c + N_c < 40\%n$, the $T_d = 0$, if $P_c < N_c$, the $T_d = 20\%V_p$, else the $T_d = 20\%V_n$.

6. Finally, compared with the threshold T_d, the data are turned into a binary sequence $S_1, S_2, S_3, \ldots, S_r$. If $x_i < T_d$,$S_i = 0$, otherwise the $S_i = 1$.

As in [36], these given parameters "10%", "40%" and "20%" were used in this method for convenient comparison.

4.2 Lempel-Zip complexity analysis

In [38], Lempel and Zip presented that the signal must firstly be transformed into a finite symbol sequence known as a coarse-graining process, and then the complexity measure can be gained by counting the number of the distinct patterns in the given binary sequence with the discrete time series. In other words, the LZ complexity analysis is utilized to calculate the new pattern generation rate in the given finite sequence. In the context of ECG signal analysis, the signal $x(n)$ should be converted into a binary sequence which will then be scanned from left to right. Throughout the whole sequence, the complexity counter $c(n)$ is increased by one when a new

subsequence of consecutive binary sequence is encountered in the scanning process. Finally, the $C(n)$ denoted the normalized output of the LZ complexity analysis instead of $c(n)$ is used to describe the results.

The finite length sequences P whose elements are a few symbols as follows. $P = S_1, S_2, S_3, \ldots, S_n$. The $l(P)$ is denoted the length of symbol sequence. S and Q are denoted two subsequences of P and let the SQ be the concatenation of S and Q. The sequence $SQ\pi$ originates from SQ when the last character is deleted (the π is denoted an operation, which can delete the last character in the symbol sequence). Let $v(SQ\pi)$ denote the set of all different subsequences of $SQ\pi$.

Assume $S = S_1, S_2, S_3, \ldots, S_r$ and $Q = S_{r+1}$, then $SQ\pi = S_1, S_2, S_3, \ldots, S_r$; if $Q \in v(SQ\pi)$, the Q belongs to the subsequence of $SQ\pi$, not a new sequence. And S doesn't need to be changed.

Subsequently, the Q is renewed to $Q = S_{r+1}, S_{r+2}$, and judge if Q is a subsequence of $SQ\pi$ or not, and continue until $Q \notin v(SQ\pi)$.

Now $Q = S_{r+1}, S_{r+2}, S_{r+3}, \ldots, S_{r+i}$ doesn't belong to $SQ\pi = S_1, S_2, S_3, \ldots, S_{r+i-1}$, so the $c(n)$ complexity counter is increased by one.

Then,$S = S_1, S_2, S_3, \ldots, S_{r+i}$, and $Q = S_{r+i+1}$. Repeat the above procedure until Q is the last character.

Finally, the $c(n)$ value can be got by the number of different subsequences.

From what has been mentioned above, it is extremely easy to implement the LZ complexity measure algorithm. For the purpose of the complexity measure which is independent of the sequence length, $c(n)$ is normalized to $C(n)$. When the length of the sequence is n and the number of different symbols in the symbol set is α. In [38], the upper bound of $c(n)$ proved is given by

$$c(n) < \frac{n}{(1 - \epsilon_n \log_\alpha(n))} \tag{4.10}$$

Where, ϵ_n is a small quantity and $\epsilon_n \to 0$ $(n \to \infty)$. Thereby, $\frac{n}{\log_\alpha(n)}$ is the upper bound of $c(n)$. However, using the coarse-graining technique, the finite symbol sequence is denoted binary sequence (0/1 sequence), the $\alpha = 2$. The function is as follows:

$$\lim_{n\to\infty} c(n) = b(n) = \frac{n}{\log_2(n)} \tag{4.11}$$

Hence the $c(n)$ can be normalized by $b(n)$

$$C(n) < \frac{c(n)}{b(n)} \tag{4.12}$$

where $0 \leq C(n) \leq 1$. In order to obtain the result $C(n)$ independent of n [68], we need the length $n > 1000$.

Thereby, the time required at least 5 seconds window length from the raw data, for example for 1000 data points (200 sample rate * 5 seconds). In the given finite sequence, the $C(n)$ reflects the arising rate of new patterns in the sequence.

4.3 Arrhythmias Classification Rule

In this combinatorial algorithm architecture, the signal analysis is performed in the following steps. Firstly, the ECG signal is classified into different rhythms by the Pan-Tompkins algorithm. More specifically, when a beat is detected it is classified as a normal beat, or a PVC. The PVC represents any beat of ventricular origin, including VT, VF and some other rhythms. In this project, the PVC mainly consists of VT and VF. If a beat is classified as a normal beat, the system proceeds to analyze the next beat; if classified as a PVC, this beat will be further analyzed. The ECG may contain different heart rhythms and those rhythms may mix together, which makes the classification more difficult, since the LZ complexity analysis needs data of at least 1 window length long, which is 8 seconds in this case. Some decision rules are added to the classification step. Three consecutive PVC beats in a row are considered to be an indication of the beginning of VT rhythms (suggested in paper [69]), at which point the following samples will be saved until up to 1600 samples are reached. To increase the reliability of the system, the heart rate is also monitored, which may help improve the accuracy of the analysis algorithm. The number of beats in the saved 1600 samples (8 seconds) can be obtained subsequently. If no more than 80% of the beats belong to the SR, the 1600 samples will be classified as a segment of VT or VF signal, depending on the value of $C(n)$; otherwise, the samples are considered as a SR segment. The given parameter 80% was derived by the study of MIT-BIH database. The $C(n)$ denoted the normalized output of LZ complexity analysis instead of $c(n)$, where the $c(n)$ is increased by one when a new subsequence of consecutive binary sequence is encountered in the scanning process.

4.4 Coarse-graining Process Analysis

The coarse-graining process has been classified two type based on their features. K-Means cluster algorithm was used to divide the given data into two clusters based on the ECG signals' feature. However, the other three methods based on different database need to train different database for obtaining the given parameter that is used to determine the threshold. Therefore, the K-Means cluster algorithm outperforms the other three method.

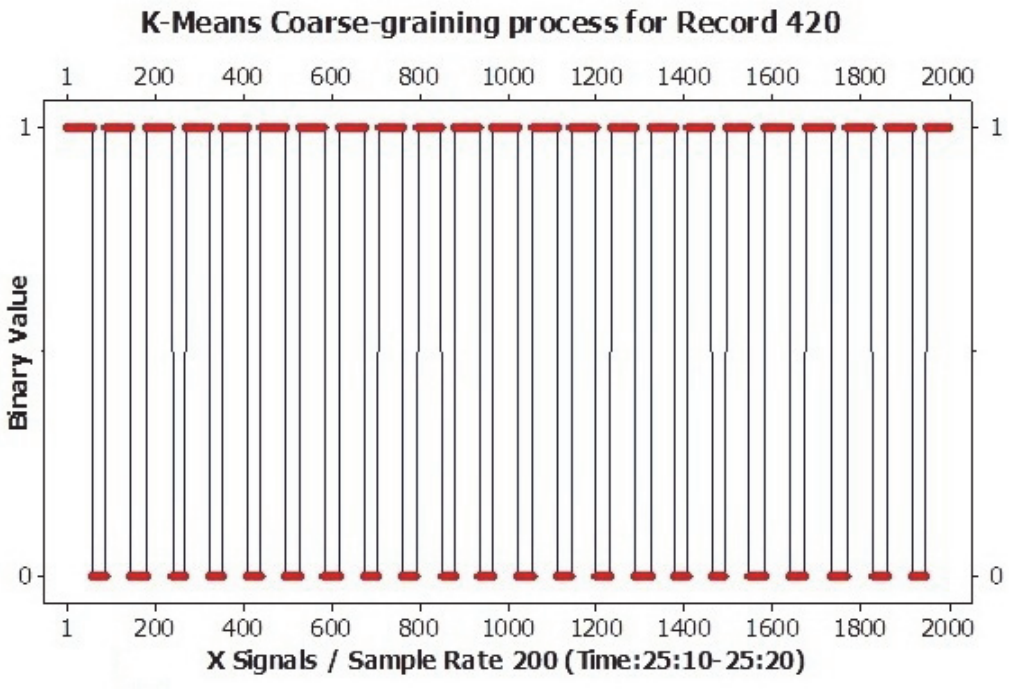
4.4.1 The K-Means clustering Algorithm Analysis

According to the features of arrhythmias, the inherent nature of the original ECG signal is refined by the K-Means clustering algorithm, which is flexible and useful to classify regarding a certain number of clusters. In other words, the algorithm is used to divide the given data into several clusters based on their features. There are just as many centroids to be defined for a certain number of clusters (K clusters). The core idea is to define the number of centroids. Therefore, the monomorphic VT signal shown in Figure 2.17(a) can be separated into two clusters regarding their periodic morphology.

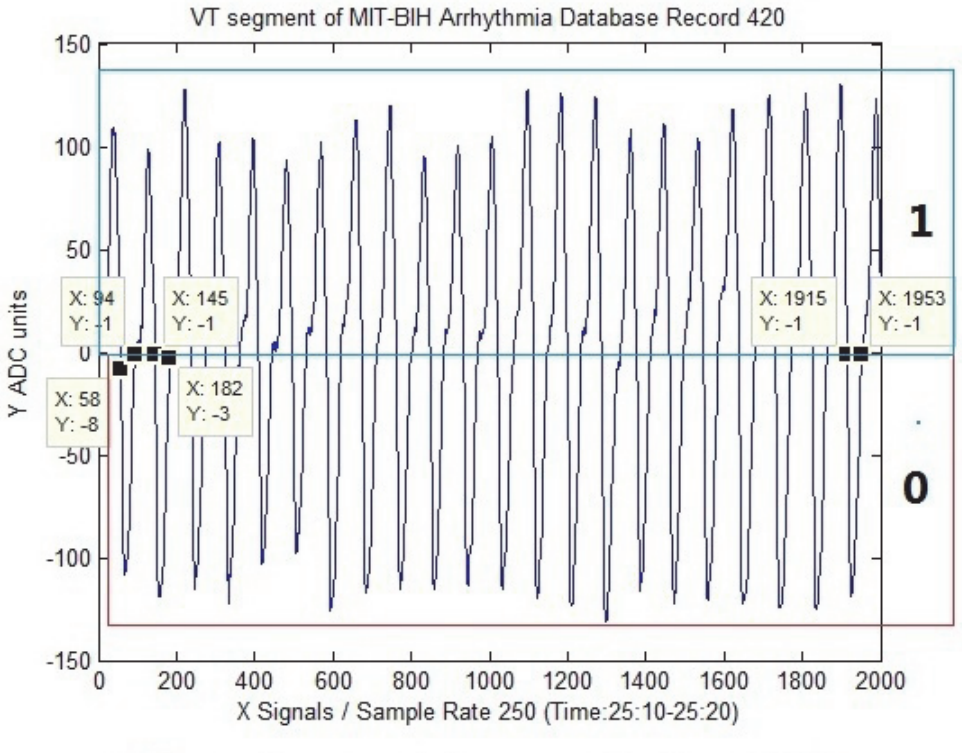
From what has been presented in the K-Means clustering algorithm above, a monomorphic VT segment with a 10 second window length collected from the MIT-BIH database can be efficiently extracted into the binary sequence in Figure4.2(b). For a sample rate of 200Hz, data points are set to 1 when exceeding the baseline. Data points are set to 0 when below the baseline. For example, Fig.4.2(a) shows that the first range ($58 < x < 94$) is set to 0; the second range ($145 < x < 182$) is set to 0; etc.; the last which ranges from 1915 to 1953 is set to 0. Throughout the whole binary sequence, the complexity $c(n)$ can be calculated when a new subsequence of consecutive binary sequence is encountered in the scanning process.

4.4.2 The Mean-value Algorithm Analysis

In [36], Zhang utilized a method based on Mean-value to extract the original ECG signal for the coarse-graining process.There are several empirical values to be used to construct this Mean-value algorithm that transforms the given original ECG signal



(a) 10 second Monomorphic VT signals



(b) Corresponding binary signals

Figure 4.2: K-Means clustering Analysis

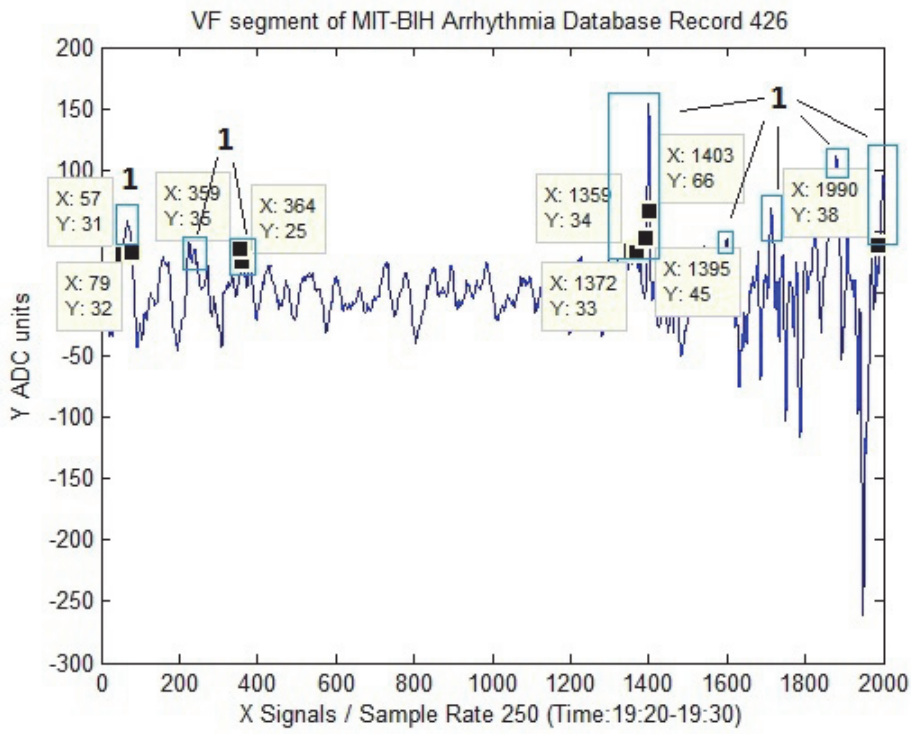


Figure 4.3: 10 seconds VF signals

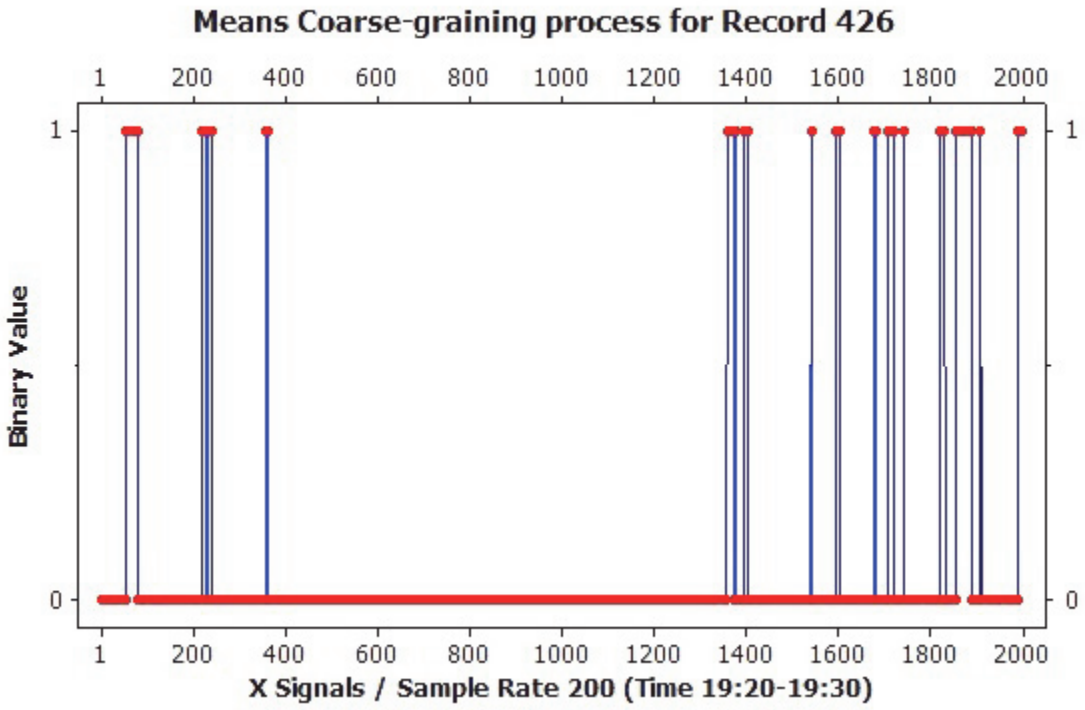


Figure 4.4: Using Mean-value algorithm

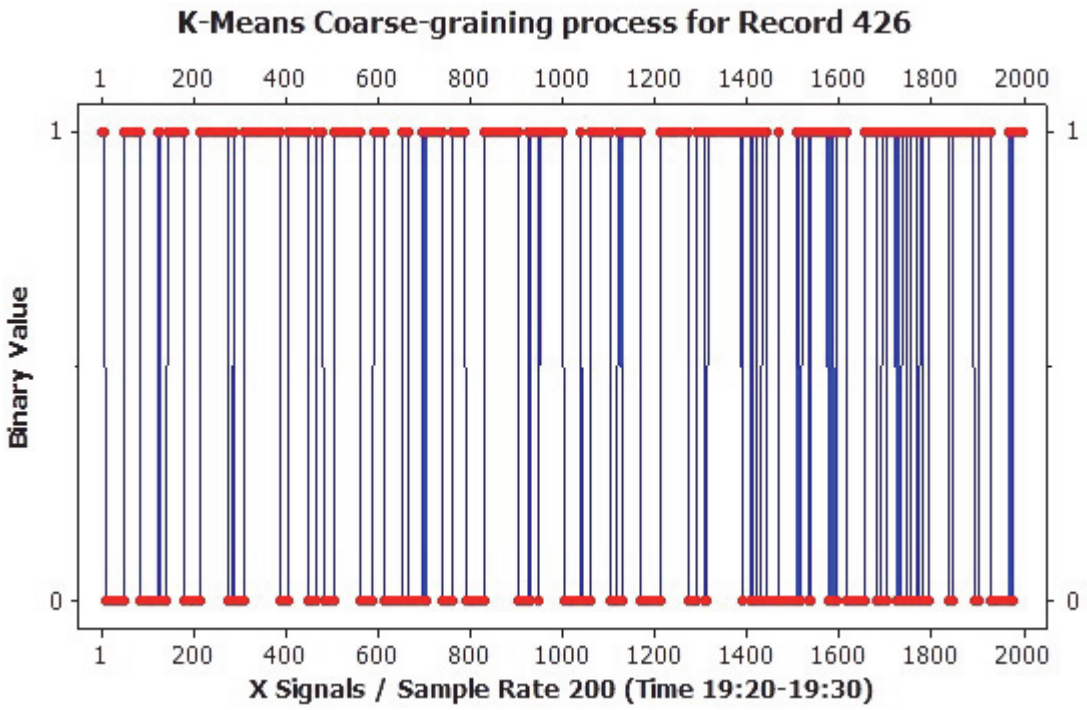


Figure 4.5: Using K-Means clustering algorithm

into a finite binary sequency. The algorithm was tested in their own database with high sensitivity to separate VT and VF. Since these empirical parameters are suitable only for their own database from where this distribution has been learned, it has a serious limitation in analyzing the public database "MIT-BIH database" using the Mean-value algorithm. An example is presented to illustrate the limitation of this algorithm as follows:

A 10 second VF segment filtered by a band-pass filter(0.6-22Hz) is analyzed by the Mean-value method shown in Figure 4.3. To begin with, the Mean value of the data points is 0.153. The Mean value is subtracted from every data point. The positive peak value V_p and the negative value V_n can be calculated separately as 153.847 and -261.153. Then, by counting the number of data points in these ranges ($0.0 < x_i < 10\%V_p$ and $10\%V_n < x_i < 0.0$), the number of data points are equal to 602, and 700 respectively. We can find that most of the original signals are evenly distributed on either side of the baseline. The threshold Td should be set to 0 regarding the distribution of the counting value and the Mean value. However, as to the given value 40% [36], the threshold is selected as Td=30.7694 for this segment. On the

basis of the selected Td, Figure 4.4 illustrates a poor refinement of the original signal. Since the baseline has an upward parallel translation of around 30.7694, most of the original signals are set as 0. Finally, the $C(n)$ is equal to 0.137072, which is close to VT's $C(n)$ value. This segment is divided into VT, making the original ECG signal to be transformed to a finite binary sequence in the wrong way. In contrast, the K-Means clustering algorithm can extract the original signal into a binary sequence very well shown in Figure 4.5. The $C(n)$ value 0.296076 is classified as VF. Therefore, the original signal is more likely to be correctly extracted into a binary sequence when the given value can be flexibly adapted regarding the range of the signal distribution. According to this segment's distribution, the given value of 40 % is changed to 70%, generating the threshold as Td=0. Finally, the $C(n)$ value can be attained (0.307042), and separated as VF. Figure 4.6 shows a good result.

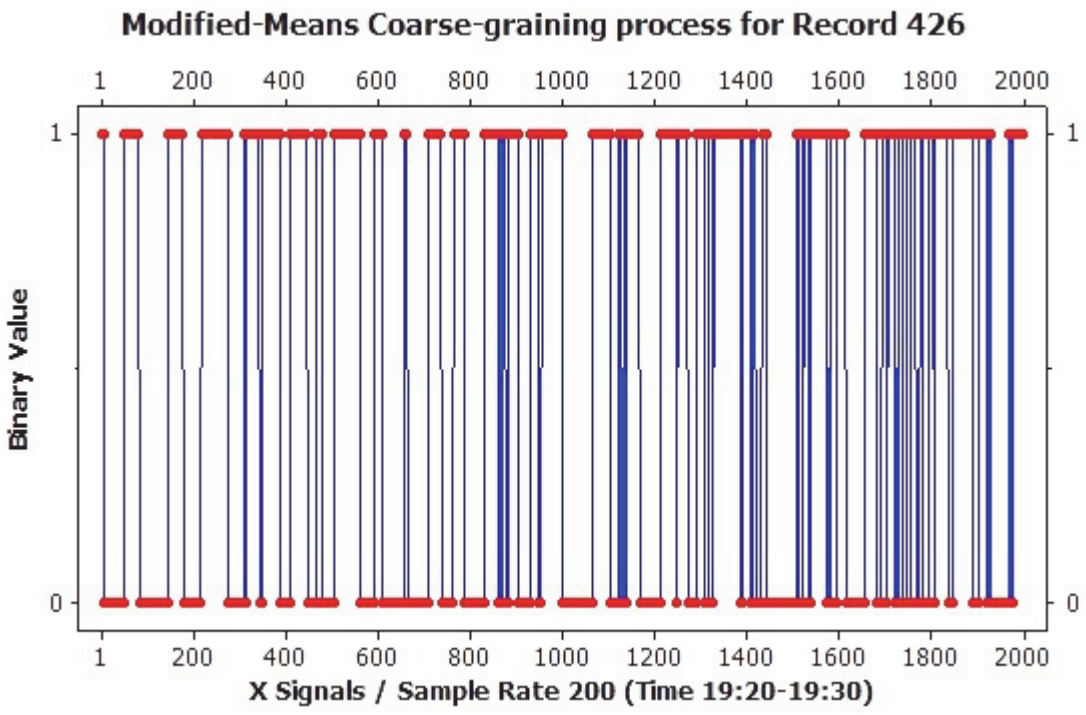


Figure 4.6: Using Modified Mean-value algorithm

4.5 Other Method For Monomorphic VT and VF distinction

4.5.1 Feature Extraction based on Histogram

According to the ECG morphology, a histogram represents a visual feature of statistical ECG signal that illustrates the density of amplitude distribution of a window length signal in successive numerical intervals of equal size. The histogram function shows the statistical nature of monomorphic VT and VF by means of their periodic or chaotic features with time series. There is a general mathematical formula to describe histogram analysis as follows:

$$H = \Sigma_{i=1}^{k} m_i \tag{4.13}$$

where the function m_i counts the numbers of data points that fall into each of the intervals. H is the total number of observed quantity and k is the total number of the intervals. The size of the intervals can reveal the different features of the whole window. The interval width W can be calculated as follows:

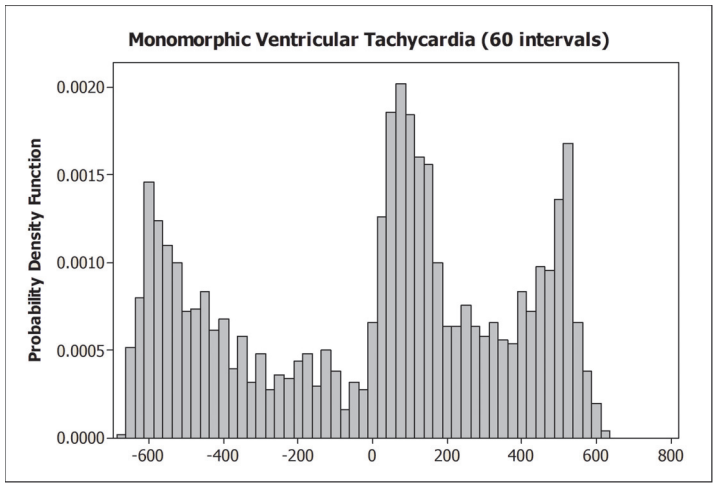
$$W = \lceil \frac{V_{max} - V_{min}}{k} \rceil \tag{4.14}$$

where the V_{max} and V_{min} are the maximum value and the minimum value among the amplitudes of a window length signal, respectively. The histogram function indicates the amplitude distribution of a window length long ECG signal across a data range from the maximum value to the minimum value. The data range can be divided into a certain number of intervals (k). The 60 intervals were chosen in order to create an efficient distribution of a window length long ECG signal as shown in Figure 4.7. It can be seen that the VF displays an approximately Gaussian distribution in Figure 4.7(b).

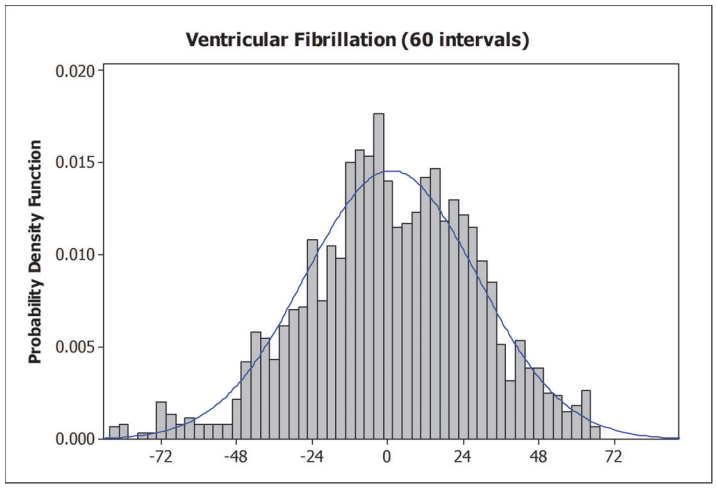
4.5.2 Average absolute deviation

The average absolute deviation can be gained by the subtraction of the reference histogram from the histogram of current data with each interval. The average absolute deviation of the window length long ECG signal containing k intervals is defined as:

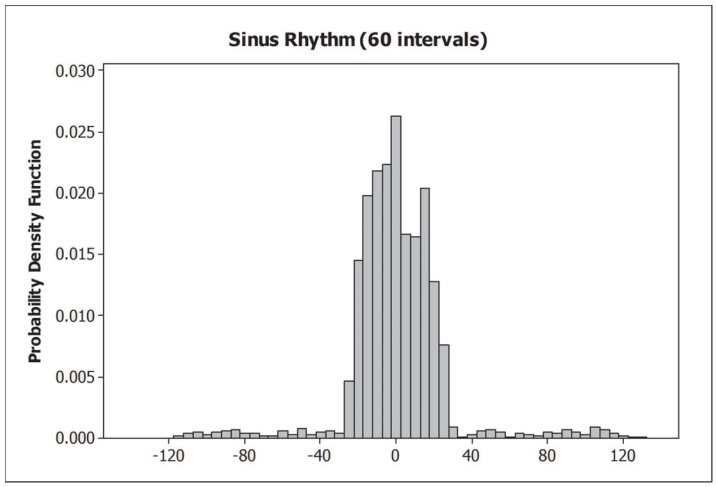
$$D(n) = \frac{\Sigma_i^k |h_i^n - Rh_i^n|}{k} \tag{4.15}$$



(a) Monomorphic VT Histogram



(b) VF Histogram



(c) SR Histogram

Figure 4.7: Histogram based on 10sec segment length signal from MIT-BIH Database

where Rh_i^n is a reference histogram that describes a selected window length long monomorphic VT based on n data points ($n = Sample\ Rate * Window\ Length$) and h_i represents the histogram of analyzed window length long signal, $i = 1, 2, 3, \cdots, k$. In other words, the method is utilized to calculate the deviation value of two different histograms in a given window length. Throughout the whole finite data, the deviation value $D(n)$ is obtained from each training segment; and then those $D(n)$s are analyzed to get a threshold value, $D(n)_T$ distinguishing between monomorphic VT and VF. The signal is considered to be monomorphic VT if its $D(n)$ is less than $D(n)_T$, otherwise, it is classified as VF.

4.5.3 Statistical analysis

The Student's t-test is used to analyze and reveal the significance in the histogram approach for the monomorphic VT and the VF arrhythmias distinction.

Table 4.1: Two-Sample T-test for Monomorphic VT Sample and VF

	N	Mean	Standard Deviation	SE Mean
Monomorphic VT	2000	1.1	71.1	1.6
VF	2000	-0.6	23.8	0.53
T-Valuse =1.02, P-Value = 0.307				

The independent t-test was performed because of the samples' characteristics for evaluating the research hypothesis that monomorphic VT of the first sample of independent variable has an exact difference average dependent variable than VF of the second sample of independent variable for each window length long ECG signal. In Table 4.1, all the values corresponding to the two types arrhythmia groups are given. In order to see the significance in different average values, the p value was considered statistically significant by looking up p-values in table 4.2 [6] when the degrees of freedom is 4000. To say further, look down the column that corresponds to the chosen value for p ($p < 0.5$), and then find the row that corresponds to the degrees of freedom (4000), and where they encounter, the critical value is found. Last thing is to compare this value with the t-value. The calculated t-value is greater than or equal to this value, then t is significant and has been found a difference. The t-test provided a p value (0.307) less than the statistical significance level (0.5), so the difference can be considered statistically significant in this situation. In the same method

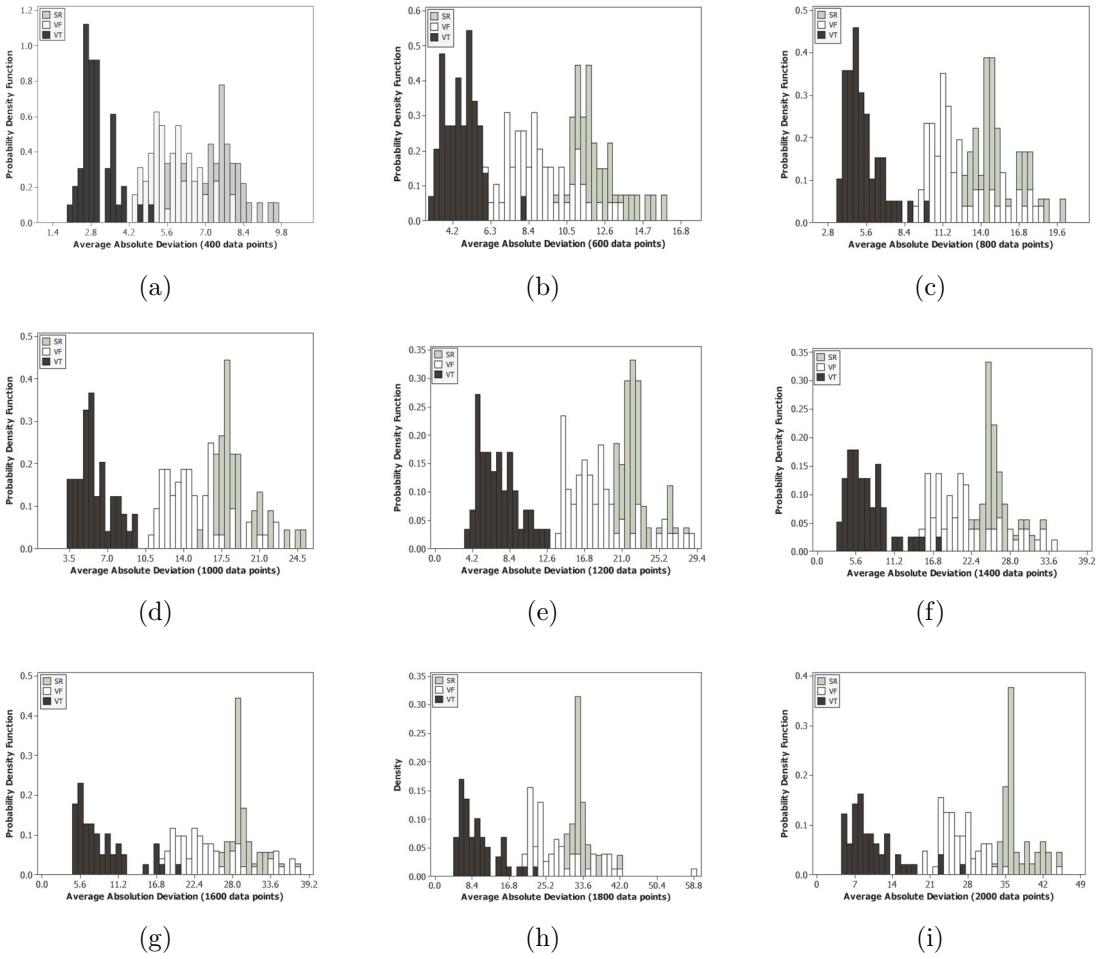


Figure 4.8: Nine representative probability density functions (PDF's) of D(n) for 158 ECG segments (45 SR, 49 VT, and 64 VF) in the training stage. There (a) Window length = 2s (400 data points), (b) Window length = 3s (600 points), and (c) window length = 4s (800 points), (d) Window length = 5s (1000 data points), (e) Window length = 6s (1200 points), (f) Window length = 7s (1400 data points), (g) Window length = 8s (1600 points), and (h) window length = 9s (1800 points), and (i) window length = 10s (2000 points).

analysis, table 4.4 illustrates that the p value (p > 0.5) was considered statistically insignificant. It indicates that the two segments belong to the same characteristic arrhythmia. It is concluded that the arrhythmias classification based on the histogram can reveal the difference of monomorphic VT and VF. The study can be concluded that the difference between the means for the two groups is different (even given the variability).

Table 4.2: Table of Upper-Tail and Two-Tail t Critical Values [6]

one-tail p	**0.001**	**0.0025**	**0.005**	**0.01**	**0.025**	**0.05**	**0.1**	**0.25**
two-tail p	**0.002**	**0.005**	**0.01**	**0.02**	**0.05**	**0.1**	**0.2**	**0.5**
df = 1	318.3	127.3	63.66	31.82	12.71	6.314	3.078	1
2	22.33	14.09	9.925	6.965	4.303	2.92	1.886	0.816
3	10.21	7.453	5.841	4.541	3.182	2.353	1.638	0.765
4	7.173	5.598	4.604	3.747	2.776	2.132	1.533	0.741
5	5.893	4.773	4.032	3.365	2.571	2.015	1.476	0.727
6	5.208	4.317	3.707	3.143	2.447	1.943	1.44	0.718
7	4.785	4.029	3.499	2.998	2.365	1.895	1.415	0.711
8	4.501	3.833	3.355	2.896	2.306	1.86	1.397	0.706
9	4.297	3.69	3.25	2.821	2.262	1.833	1.383	0.703
10	4.144	3.581	3.169	2.764	2.228	1.812	1.372	0.7
100	3.174	2.871	2.626	2.364	1.984	1.66	1.29	0.677
200	3.131	2.839	2.601	2.345	1.972	1.653	1.286	0.676
300	3.118	2.828	2.592	2.339	1.968	1.65	1.284	0.675
400	3.111	2.823	2.588	2.336	1.966	1.649	1.284	0.675
500	3.107	2.82	2.586	2.334	1.965	1.648	1.283	0.675
600	3.104	2.817	2.584	2.333	1.964	1.647	1.283	0.675
700	3.102	2.816	2.583	2.332	1.963	1.647	1.283	0.675
800	3.1	2.815	2.582	2.331	1.963	1.647	1.283	0.675
900	3.099	2.814	2.581	2.33	1.963	1.647	1.282	0.675
1000	3.098	2.813	2.581	2.33	1.962	1.646	1.282	0.675
2000	3.094	2.81	2.578	2.328	1.961	1.646	1.282	0.675
3000	3.093	2.809	2.577	2.328	1.961	1.645	1.282	0.675
4000	3.092	2.809	2.577	2.327	1.961	1.645	1.282	0.675
5000	3.092	2.808	2.577	2.327	1.96	1.645	1.282	0.675
df = $\infty(z)$	3.09	2.807	2.576	2.326	1.96	1.645	1.282	0.674
one-tail p	**0.001**	**0.0025**	**0.005**	**0.01**	**0.025**	**0.05**	**0.1**	**0.25**
two-tail p	**0.002**	**0.005**	**0.01**	**0.02**	**0.05**	**0.1**	**0.2**	**0.5**

The $D(n)$s of 158 ECG segments (45 SR, 49 Monomorphic VT, and 64 VF) were

Table 4.4: Two-Sample T-test for VF and VF_2

	N	Mean	Standard Deviation	SE Mean
VF	2000	-0.6	23.8	0.53
VF_2	2000	-0.8	22.6	0.51
T-Valuse =0.28, P-Value = 0.777				

obtained based on different window lengths ($WL = 2s, 3s, \cdots, or 10s.$) for studying their influence on the classification results. Then, the average absolute deviation $D(n)$s for SR, Monomorphic VT, and VF are analyzed by statistical approach for training stage, and the corresponding probability density function (PDF) is described by histogram in Figure 4.8.

Table 4.5: Performance of the Histogram and Average Absolute Deviation Algorithm for Monomorphic VT and VF Classification

WL(sec)	Threshold	SENSITIVITY	
		MONOMORPHIC VT	VF
2	4.2	47/49=95.92%	61/64=95.31%
3	6.4	48/49=97.96%	51/64=79.69%
4	9	48/49=97.96%	63/64=98.44%
5	10	48/49=97.96%	63/64=98.44%
6	13	48/49=97.96%	63/64=98.44%
7	15	48/49=97.96%	63/64=98.44%
8	18	48/49=97.96%	61/64=95.31%
9	20	48/49=97.96%	59/64=92.19%
10	19.5	48/49=97.96%	64/64=100%

Note:SL= Window Length; Sensitivity%= TP/((TP+FN)) where TP=true positive,FN=false negative

All of monomorphic VT $D(n)$s are very close with the reference histogram as shown in Figure 4.8. Monomorphic VT and VF are clearly marked off in different window length. In other words, it is evident that the window length doesn't affect the distinction of monomorphic VT and VF, which means that the proposed algorithm can be used to availably classify as the monomorphic VT and VF based on ECG signal morphology. However, Figure 4.8 illustrates that there is overlap between the the distributions of VF and SR. This problem can be solved by applying the algorithm proposed by Pan-Tompkins [37] with several rules presented by Hamilton [5], which have been used in classifying normal SRs from the arrhythmias [37].

As to the statistical analysis above, we selected 49 monomorphic VT segments

and 64 VF segments in different window length from the MIT-BIH database for the testing stage. The proposed method is applied and the results are shown in Table 4.5. With different choices of window length, the method can achieve fairly high performance. The best sensitivity is achieved at the 10 second window length. The results show a 97.96% sensitivity on monomorphic VT signal, 100% sensitivity on VF.

Chapter 5

Results and Implementation

According to the above chapters, the ECG signal analysis system has been implemented both in Carbide C++ emulator and a Nokia S60 Smart-phone. In order to evaluate each component of the proposed system, the MIT-BIH database [9] is adopted for both development and evaluation stages for general and comparable results. This database offers a set of standard ECG records that are open to the public, which have been enormously helpful for developing and testing ECG analysis algorithms. Every record contains a continuous recording of ECG signal from a single subject. Most of the records have been annotated by the cardiologist and put into standard annotation files. Note that, the "MIT-BIH database" is actually a database library and is comprised of various sub-databases, for specific studies. According to the testing results, the ECG analyzer program can correctly detect the cardiac health status of the subject. Meanwhile, the algorithm architecture with the new classification rule can correctly detect and systematically classify the cardiac health status of the subject,obtaining a good performance results for detection and classification of the ECG signal based on a Smart-phone.

5.1 Testing and Results

5.1.1 Beat Detection and Classification

The QRS complex representing the ventricular activity of the heart is the most remarkable waveform for cardiac event detection in ECG signal analysis, which has high potential amplitude, steep slope (R-wave) and wide duration. These features can be extracted as characteristic quantity, and as a quantized standard in the analysis process. Beat detection is the basis of any ECG analysis algorithm. It is done by detecting the QRS complex. The algorithm proposed by Pan-Tompkins [37] has been extensively used in beat detection. Then, by combining the algorithm with several

rules presented by Hamilton [5], the beats can be classified as PVCs or normal SRs. It is computationally efficient and accurate, reported to have a sensitivity of 99.69% and positive predictivity of 99.77% [37].

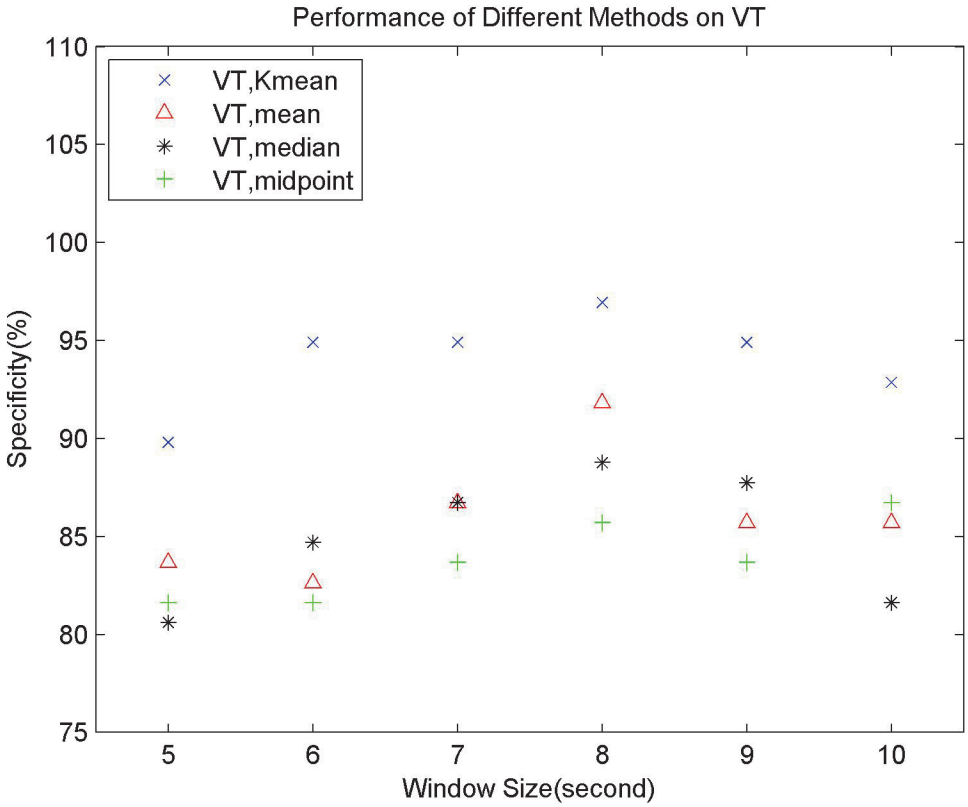


Figure 5.1: Performance of different coarse-graining process for VT

5.1.2 Data Acquisition

A set of ECG records obtained from the MIT-BIH database [9]. There are 98 monomorphic VT segments and 57 VF segments selected from the Malignant Arrhythmia subset of MIT-BIH database for development stage; the sampling frequency was set to 200Hz. In [68], in order to obtain $C(n)$ results which are independent of n, the finite string needs at least $n > 1000$. The threshold $C(n)$ value obtained from the development stage by using different coarse-graining process approaches is used to evaluate the sensitivity of VT and VF in LZ complexity measure. It is then used to interpret the impact of the coarse-graining process in ECG signal analysis. The window length was set from 5 seconds beginning. There are 6 different window

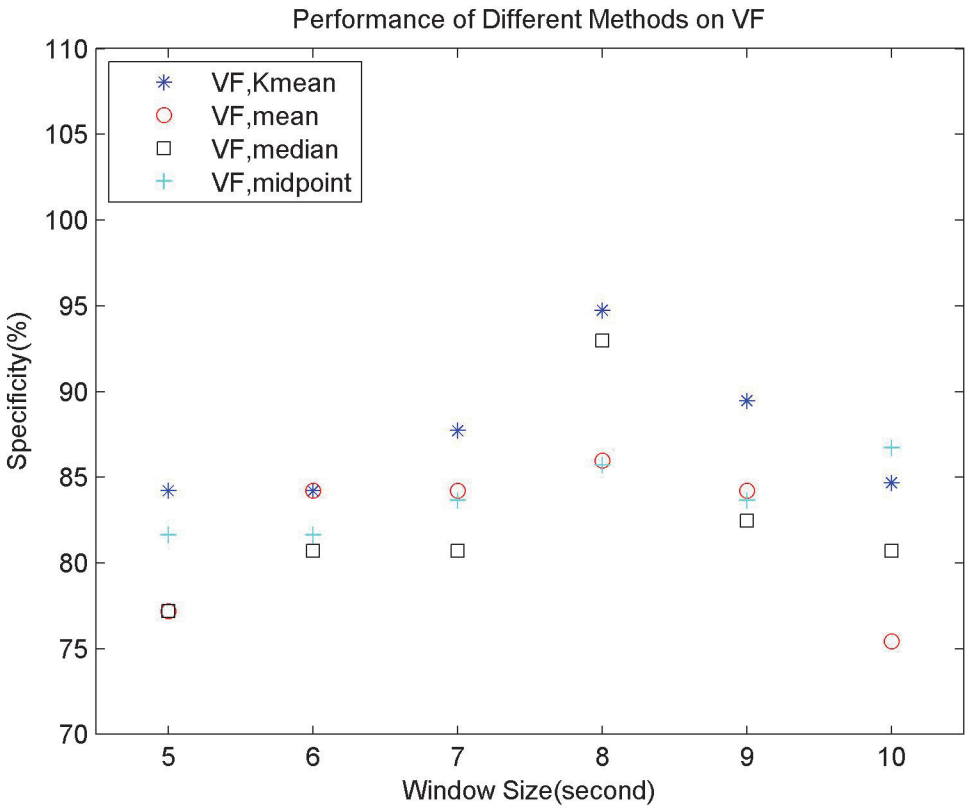


Figure 5.2: Performance of different coarse-graining process for VF

lengths to be selected for analyzing the impact of the window length on the coarse-graining process of LZ complexity measure. For evaluation stage, 100 SR segments, 98 monomorphic VT segments and 75 VF segments are selected from the MIT-BIH database at each different window length. In order to test the robustness of algorithms, 134 VT segments, including monomorphic and polymorphic VT, and 128 VF are utilized to depict the performance of combinatorial algorithms. Finally, several records are randomly collected from the MIT-BIH Arrhythmias subset and Malignant Arrhythmia subset to test the performance of the proposed system.

5.1.3 Development Stage

The VT has two types that include monomorphic VT and polymorphic VT. There are lots of difficulty to distinguish VT from VF, leading to the main error in previous studies. Thus, there are two parts to evaluate the combining algorithms. Firstly, there are monomorphic 98 VT segments and 57 VF segments selected from the Malignant

Arrhythmia subset of MIT-BIH database for development stage to get a threshold based on different window length.

According to the four coarse-graining approaches proposed above, the test results based on different window lengths (5s to 10s) are shown in the Fig.5.1 and Fig.5.2. The K-Means cluster algorithm outperforms the other three methods in terms of classification of VT and VF. This implies that the symbolic sequence constructed by the K-Means partitioning is a better representation of the original signal. The best sensitivity was obtained at the 8 second window length.

Let $C(n)$ denote the normalized output of LZ complexity algorithm. We obtained $C(n)$ for all the segments and then used statistical tools to analyze those $C(n)$s. By examining the probability density function (PDF) which is shown in Fig.5.3, a threshold for distinguishing between VT and VF is found, that is $C(n)_T = 0.234$. The signal is considered to be VT, if its C(n) is less than $C(n)_T$, otherwise, it is classified as VF.

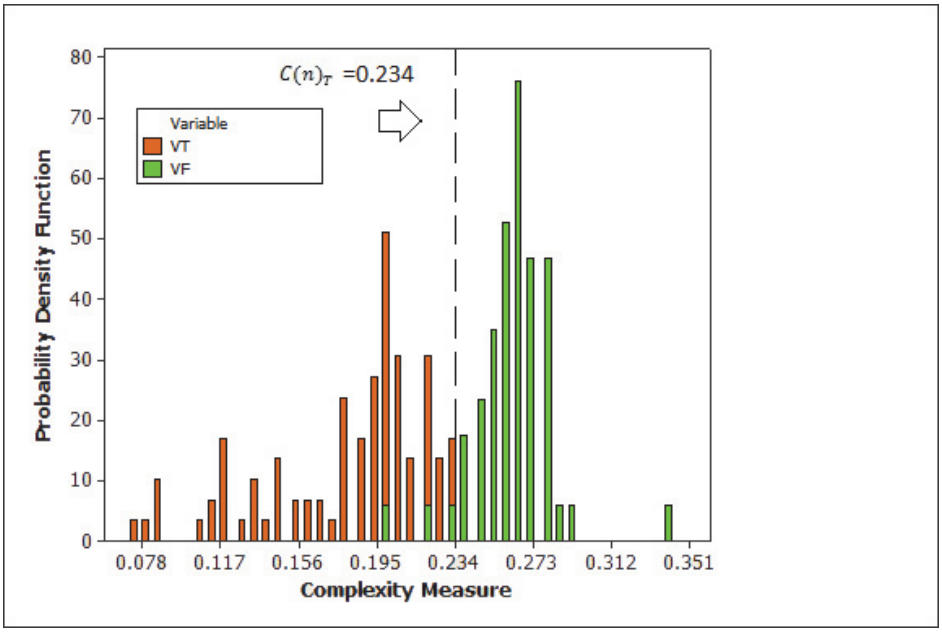


Figure 5.3: A threshold for distinguishing between VT and VF

5.1.4 Evaluation Stage

For evaluation, 100 SR segments, 98 monomorphic VT segments and 57 VF segments are selected from the MIT-BIH database at each different window length. The

$C(n)$ value obtained from the development stage is used to evaluate the sensitivity of SR, VT and VF based on different coarse-graining process approaches. The proposed method is applied and the result is shown in Table 5.1. It achieves fairly high performance when the selected window length is 8 seconds. The results show a 100% sensitivity on SR signal, 96.94% sensitivity on Monomorphic VT and 94.74% sensitivity on VF.

Table 5.1: Performance of Classification for SR, VT AND VF

WL(sec)	SENSITIVITY		
	SR	Monomorphic VT	VF
5	100/100=100%	88/98=89.80%	48/57=84.21%
6	100/100=100%	93/98=94.90%	48/57=84.21%
7	100/100=100%	93/98=94.90%	50/57=87.72%
8	100/100=100%	95/98=96.94%	54/57=94.74%
9	100/100=100%	93/98=94.90%	51/57=89.47%
10	100/100=100%	91/98=92.86%	42/57=84.69%

Fairly high performance is achieved when selecting 67 mixed monomorphic VT with polymorphic VT, and 64 VF for evaluating the ECG analysis system. In Table 5.2, the results show a 91.04% sensitivity on mixed VT and 93.75% sensitivity on VF when selecting the 10 second window length with a threshold of $C(n)_T = 0.2522$. Note:WL means window length, Sensitivity%= TP/((TP+FN)) where TP=true positive, FN=false negative

Table 5.2: Performance of Classification for VT and VF

WL(sec)	Threshold	SENSITIVITY	
		POLYMORPHIC VT	VF
5	0.2491	42/67=62.69%	53/64=82.81%
6	0.2557	56/67=83.58%	55/64=85.94%
7	0.2463	57/67=85.07%	59/64=92.19%
8	0.2461	58/67=86.57%	60/64=93.75%
9	0.2403	57/67=85.07%	58/64=90.63%
10	0.2522	61/67=91.04%	59/64=92.19%

Some records were chosen from the database to test the proposed system. According to the test results shown in Table 5.3, the ECG analysis system with the new classification rules can correctly detect the cardiac health status of the subject, generating a personalized diagnostic report.

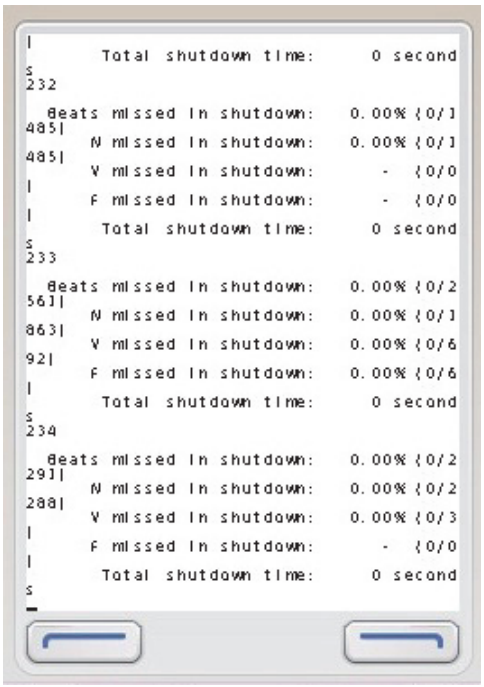
Table 5.3: Results of Testing the Proposed System

	RHYTHMS CONTAINED IN THE RECORD	
Records	Standard Annotation	Results of System
100	One VT (25m18s)	Normal Rhythm
101	No VT and VF	Normal Rhythm
103	No VT and VF	Normal Rhythm
112	No VT and VF	Normal Rhythm
113	No VT and VF	Normal Rhythm
115	No VT and VF	Normal Rhythm
220	No VT and VF	Normal Rhythm
230	One VT (29m08s)	Normal Rhythm
234	three VTs (17m06s, 21m31s, 28m29s)	Normal Rhythm
420	VT [23m50s to 28m46s]	VT [23m56s to 28m16s]
422	VT [22m12s to 22m14s]	VT [22m15s
	VT [22m15s to 25m28s]	to 25m26s]
427	VT [10m48s to 25m33s]	VT [10m53s to 25m23s]
605	VT [27m28s to 27m56s]	VT [27m36s to 27m57s]
611	VT [19m56s to 35m00s]	VT [20m00s to 35m00s]
612	VT [28m27s to 35m00s]	VT [28m30s to 35m00s]
426	VF [10m45s to 19m26s]	VF [10m48s to 17m38s]
426	VF [10m45s to 24m00s]	VF [10m48s to 23m29s]
426	VT [27m47s to 31m12s]	VT [27m35s to 31m06s]
426	VF [31m12s to 33m19s]	VF [31m16s to 33m13s]

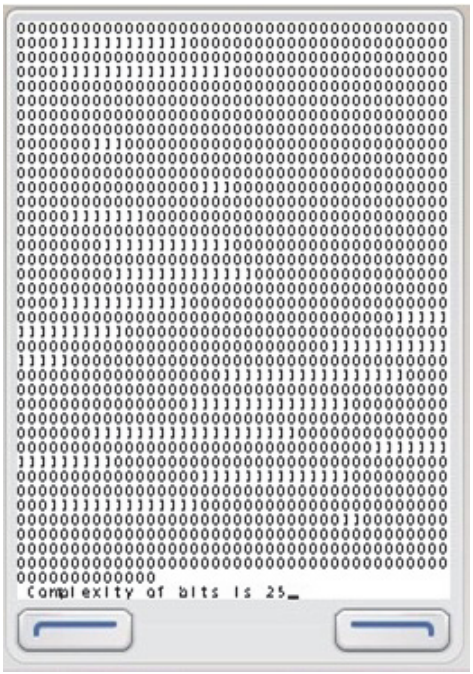
5.2 Implementation

A Nokia Smart-phone is used to test and simulate this ECG analysis system. The operation system (OS) of Smart-phone is Symbian OS. The programming is compiled in the Carbide C++ IDE (Integrated Development Environment). Symbian C++ is the native programming language of the Symbian platform. It was specifically designed for mobile devices with low power consumption and a small memory foot print. Fig.5.4 shows that the Pan-Tompkins algorithm and LZ complexity measure have been simulated in Carbide C++ emulator.

The proposed system has been implemented both in Carbide C++ emulator as shown in Fig.5.5, and a Nokia S60 Smart-phone. It runs well on the Smart-phone(Fig.5.6).

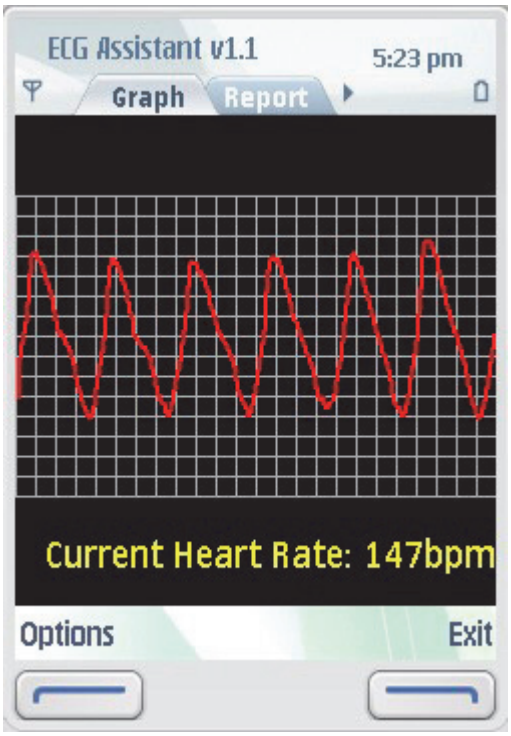


(a) Pan-tompkins algorithm runs in the emulator



(b) LZ complexity measure runs in the emulator

Figure 5.4: Algorithms run in Carbide C++ emulator.



(a)

(b)

Figure 5.5: Simulation of the proposed system

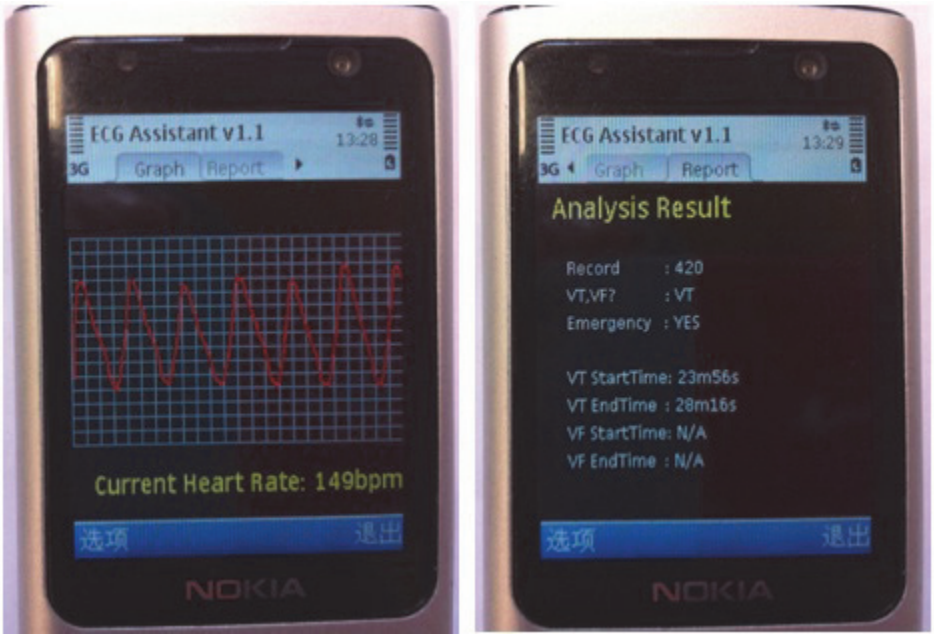


Figure 5.6: Implementation on a Nokia Smart-phone

Chapter 6

Conclusion and Future Work

6.1 Conclusion

In this thesis, it has been mainly focusing on the ECG signal analysis implemented on a Smart-phone for ECG tele-monitoring analysis system. Considering the limitations of the Smart-phone (e.g. computing capabilities, running speed), an algorithm architecture based on two time-domain algorithms are proposed for detection and classification, which prove to be reliable and robust in analyzing ECG signals. The Pan-Tompkins algorithm is applied to detect the QRS complex beats, using several classification rules to classify as SR beat or PVC beat. Having shown advantages in time-domain analysis for describing the complexity of random processes, the LZ complexity measure that consists of the K-Means clustering algorithm and the LZ complexity analysis is adopted to classify high risk arrhythmias, from PVC beat to VT beat or VF beat.

Time is a crucial issue for the life-threatening arrhythmias classification. A sensitive algorithm to detect and classify PVC to VT or VF will remarkably improve the survival probability of the patients. The LZ complexity measure shows itself possessing good advantages in time-domain analysis, which is particularly useful in describing the complexity of random processes and the information theory for analyzing the chaotic ECG signal. The coarse-graining process as a part of the LZ complex measure can efficiently quantify the signal. The core idea is to transform the original data into a finite binary sequence for refining the interior ECG signal features. Different coarse-graining methods are used to fully understand how they affect the arrhythmias classification. It is discussed that how different coarse-graining methods (the K-Means cluster, the Mean-value, the Median, the Mid-point) influence the result of LZ complexity analysis. The binary sequence obtained by using the K-Means algorithm is found to be a better representation of the original signal, and then a suitable threshold of LZ complexity can be easily obtained, with which VT and VF

can be efficiently classified. The results not only show that a proper coarse-graining technique is essential in the success of LZ complexity analysis in ECG signal classification, but also suggest the k-mean algorithm as a better method in coarse-graining compared to the other algorithms.

Besides, some additional decision rules have been addressed to accurately analyze the life-threatening arrhythmias, improving the reliability of the proposed system. The reliability and error detection of arrhythmias is improved in the proposed system by generating the analysis report in terms of the time frames. Finally, in order to obtain a comparable and generalized result, the proposed system has been tested on the MIT-BIH database showing a good performance.

In addition, the novelty of this method based on histogram and average absolute deviation is that ECG signal statistics, morphological analysis, the histogram of signal (density estimation) and average absolute deviation altogether have been used to achieve a higher performance for arrhythmias classification, monomorphic VT, and VF. The proposed method has high performance, low computational complexity, and well suited for real-time implementation for ECG Tele-monitoring analysis system.

6.2 Future Work

6.2.1 Application

The proposed system can be transplanted into many kinds of mobile platforms becoming an intermediary platform of ECG tele-monitoring analysis system, especially for use on a Smart-phone. It is possible to make the users monitor and automatically analyze their cardiac health status, from anywhere and at any time, even while moving around with an active lifestyle. More specifically, an integrated ECG analysis system is implemented to detect and systematically classify ECG signal on a Smart-phone in real-time for user monitoring and rehabilitation. When several threatening arrhythmias of high risk is detected and classified during a period of time, the Smart-phone will automatically send a message, indicating to a remote Emergency Health Center for accurate assessment and treatment. Figure 6.1 shows the ECG tele-monitoring analysis system which consists of three main components: wearable ECG sensor module, ECG signal analysis based on a Smart-phone,and a remote Emergency Health

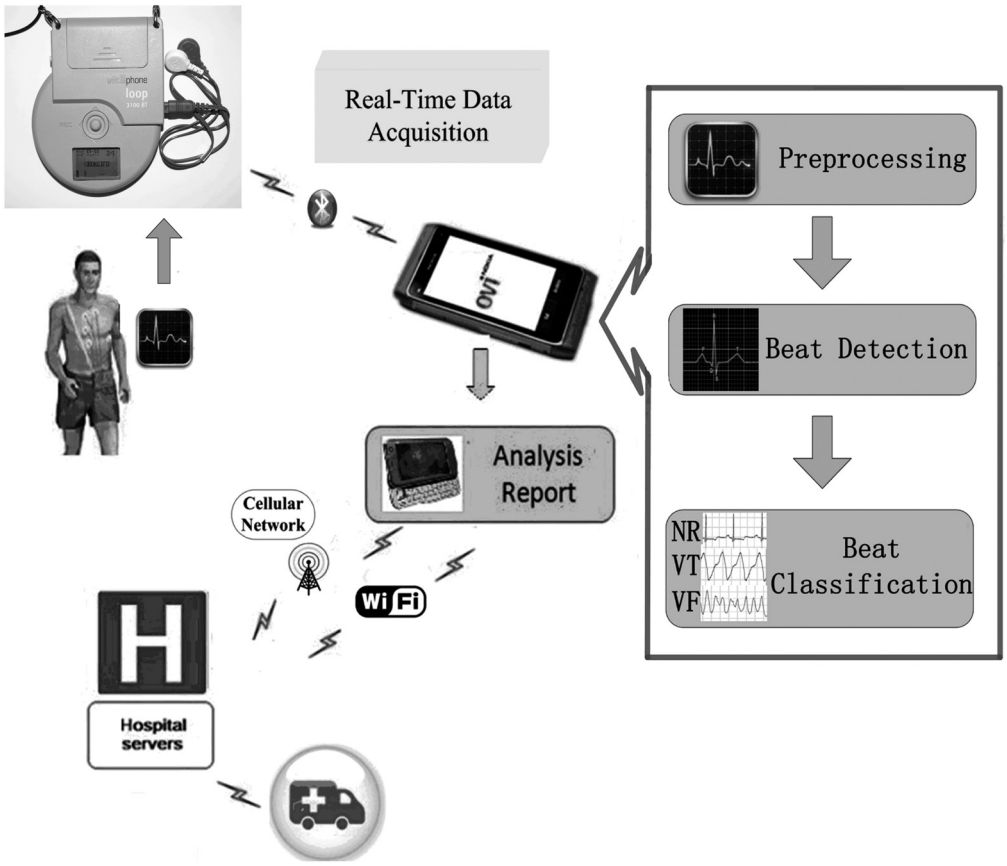


Figure 6.1: ECG tele-monitoring system based on the proposed system

Center. The ECG analysis system, carried by the users, controls ECG sensor device and performs some diagnosis and analysis in real-time, directly cause an analysis report of the user's cardiac health, and his or her subsequent heart status. In the regular process, a user is equipped with an ambulatory ECG sensor system that allows the user to continue regular daily living. A bluetooth communication network is connected between the ECG sensor system with a Smart-phone. These collecting data are transmitted through the wireless network, and then are analyzed, generating a personalized report. If there are several life-threatening arrhythmias during a period of time, the Smart-phone automatically triggers alarm messages to a remote Emergency Health Center for timely treatment.

Currently, this proposed system just was tested based on the MIT-BIH database obtaining a good performance. We will use the real patients' data for further testing.

6.2.2 Further Investigations

There will be a natural fusion of this ECG signal analysis with multi-sensor data detection. The objective is to develop a real-time ECG Healthcare System to assist health care. The system has the potential to be popularized throughout Canada and beyond. Following 4 parts will be studied and implemented:

1. Wearable multi-sensor system: ECG sensor acquires data for biometric authentication, and further detection and classification of heart beats; pulse oximeter provides detection of heart rate, and records photoplethysmograph (PPG) waves that can be combined with the ECG signal to avoid false arrhythmias detection; accelerometer has two functions: one is for activity monitoring to help plan dietary needs, and the other is for fall detection to evaluate the user's status; blood-pressure monitor is used to obtain physical blood signals as a measurable indicator for diagnosing the VF. The goal is to design and implement the multi-sensor with bluetooth communication system to collect and transmit data.

2. Wireless communication: Bluetooth technology as a communication medium is used to transmit these data. ECG signal has unique properties which can be utilized to provide wireless communication security. This work is to develop a bluetooth communication network based on ECG signal authentication. Little research has been done to take the biometric authentication based on ECG signal attributes into consideration before.

3. The assessment of quality of ECGs: This work is to develop a computer algorithm that can determine whether the quality of a signal is good or not. Moreover, when a record is unacceptable, the algorithm may also give advice as to how to correct the problems.

4. Multi-sensor data analysis: This step is to develop an effective algorithm that can fuse these multi-sensor data together to evaluate the user's physical signs, and accurately detect and classify kinds of heart rhythms step by step, generating a personalized diagnostic report for users. The multi-sensor fusion can effectively avoid undesirable alarms when detecting arrhythmias.

The system has a reliable and robust capacity to analyze kinds of physical signs, independently detect and systematically classify the types of the ECG signal as normal rhythms, atrial arrhythmias, and ventricular arrhythmias. Multi-sensor data

fusion provided by kinds of specific monitoring devices can effectively monitor users' health status, and improve the reliability and error detection of arrhythmias to avoid the false arrhythmias alarm.

Bibliography

[1] David R Ferry. *ECG in 10 Days (2nd edition)*. New York McGraw-Hill, Medical Pub, 2007.

[2] MIT open course ware. Quantitative Physiology: Organ Transport Systems. Available website(Dec.5 2011):http://ocw.mit.edu/courses/health-sciences-and-technology/hst-542j-quantitative-physiology-organ-transport-systems-spring-2004/.

[3] Adam Gacek and Witold Pedrycz. *ECG Signal Processing, Classification and Interpretation*. Springer, 2012.

[4] Texas Heart Institute. Heart Information Center – Categories of Arrhythmias. Available website (June.15 2012): http://www.texasheartinstitute.org/hic/topics/cond/arrhycat.cfm.

[5] P. Hamilton. Open source ECG analysis. In *Computers in Cardiology, 2002*, pages 101 – 104, Sept. 2002.

[6] Dawn Griffiths. *Head First Statistics*. O 'Reilly Media, 2008.

[7] University of Nottingham Division of Nursing. Available website(Dec.5 2011):http://www.nottingham.ac.uk/nursing/practice/resources/cardiology/function/conduction.php.

[8] Arnold M. *Physiology of the Heart*. New York: Raven Press Books, 1992.

[9] MIT-BIH arrhythmia database: available from: MIT-BIH Data- base Distribution, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Room E25-505A, Cambridge, MA 02139. Available Website (Dec.5 2011):http://www.physionet.org/physiobank/database/mitdb/.

[10] World Health Organization. Global Health Risks Mortality and Burden of Disease Attributable to Selected Major Risks. Available website (Jan.15 2012): http://www.who.int/healthinfo/global_burden_disease/GlobalHealthRisks_report_full.pdf.

[11] Philip de Chazal, M. O'Dwyer, and R.B. Reilly. Automatic classification of heartbeats using ECG morphology and heartbeat interval features. *Biomedical Engineering, IEEE Transactions on*, 51(7):1196 –1206, July 2004.

[12] J.J. Oresko, Zhanpeng Jin, Jun Cheng, and et al. A wearable smartphone-based platform for real-time cardiovascular disease detection via electrocardiogram processing. *Information Technology in Biomedicine, IEEE Transactions on*, 14(3):734 –740, May 2010.

[13] A.C. Norris. *Essentials of Telemedicine and Telecare*. Wiley, 2002.

[14] Sebastian Winkler, Michael Schieber, Stephanie L??cke, and et al. A new telemonitoring system intended for chronic heart failure patients using mobile telephone technology − Feasibility study. *International Journal of Cardiology*, 153(1):55 – 58, 2011.

[15] Sebastian Winkler, C. Axmann, B. Schannor, and et al. Diagnostic accuracy of a new detection algorithm for atrial fibrillation in cardiac telemonitoring with portable electrocardiogram devices. *Journal of Electrocardiology*, 44(4):460 – 464, 2011.

[16] M.C. Rodriguez-Sanchez, A. Torrado-Carvajal, J.A. Hernandez-Tamames, and et al. Novel Applications for M-Health and Free Messaging. *Pervasive Computing, IEEE*, 11(1):74 –75, Jan. 2012.

[17] J. Rodriguez, A. Goni, and A. Illarramendi. Real-time classification of ECGs on a PDA. *Information Technology in Biomedicine, IEEE Transactions on*, 9(1):23 –34, March 2005.

[18] Michael H. Crawford, Steven J. Bernstein, Prakash C. Deedwania, and et al. ACC/AHA Guidelines for Ambulatory Electrocardiography: Executive Summary and Recommendations. *ACC/AHA Practice Guidelines*, 100:886 –893, 1999.

[19] Taesoo Lee, Joohyun Hong, and Myeongchan Cho. Biomedical Digital Assistant for Ubiquitous Healthcare. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pages 1790 –1793, Aug. 2007.

[20] J.M. Cano-Garcia, E. Gonzalez-Parada, V. Alarcon-Collantes, and et al. A PDA-based portable wireless ECG monitor for medical personal area networks. In *Electrotechnical Conference, 2006. IEEE Mediterranean*, pages 713 –716, May 2006.

[21] Alive technologies. Available website(Dec.5 2011):http://www.alivetec.com/.

[22] Vitaphone. Available website(Dec.5 2011):http://www.vitaphone.de/en/company/.

[23] C. De Capua, A. Meduri, and R. Morello. A Smart ECG Measurement System Based on Web-Service-Oriented Architecture for Telemedicine Applications. *Instrumentation and Measurement, IEEE Transactions on*, 59(10):2530–2538, Oct. 2010.

[24] Lisha Zhong, Xingming Guo, and LiShan Chen. Smart-Phone Based Automatic Arrhythmia Detection and Diagnosis. In *Bioinformatics and Biomedical Engineering, (iCBBE) 2011 5th International Conference on*, pages 1 –4, May 2011.

[25] P. Leijdekkers and V. Gay. Personal Heart Monitoring System Using Smart Phones to Detect Life Threatening Arrhythmias. In *Computer-Based Medical Systems, 2006. CBMS 2006. 19th IEEE International Symposium on*, pages 157 –164, 2006.

[26] Wan-Young Chung, Chiew-Lian Yau, Kwang-Sig Shin, and et al. A Cell Phone Based Health Monitoring System with Self Analysis Processor using Wireless Sensor Network Technology. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pages 3705 –3708, Aug. 2007.

[27] Dipl. phys Ilias Sachpazidis. Home: a modular telemedicine system. In *Proc. 2 nd workshop on Mobile Computing in Medicine*, 2002.

[28] Gert Brettlecker and Heiko Schuldt. The OSIRIS-SE (stream-enabled) infrastructure for reliable data stream management on mobile devices. In *Proceedings of the 2007 ACM SIGMOD international conference on Management of data*, SIGMOD '07, pages 1097–1099, New York, NY, USA, 2007. ACM.

[29] Heart Canadian Stroke Network and Statistics Canada Stroke Foundation of Canada. 2009 tracking heart disease and stroke in canada. Available website(Dec.5 2011):http://www.who.int/healthinfo/global_burden_disease /GlobalHealthRisks_report_full.pdf.

[30] A.I. Hernandez, F. Mora, M. Villegas, and et al. Real-time ecg transmission via internet for nonclinical applications. *Information Technology in Biomedicine, IEEE Transactions on*, 5(3):253 –257, sept. 2001.

[31] Chinteng Lin, Kuancheng Chang, Chunling Lin, and et al. An Intelligent Telecardiology System Using a Wearable and Wireless ECG to Detect Atrial Fibrillation. *Information Technology in Biomedicine, IEEE Transactions on*, 14(3):726 –733, May 2010.

[32] S.K. Chen, T. Kao, C-.T. Chan, and et al. A Reliable Transmission Protocol for ZigBee-Based Wireless Patient Monitoring. *Information Technology in Biomedicine, IEEE Transactions on*, 16(1):6 –16, Jan. 2012.

[33] R.S. Dilmaghani, H. Bobarshad, M. Ghavami, and et al. Wireless Sensor Networks for Monitoring Physiological Signals of Multiple Patients. *Biomedical Circuits and Systems, IEEE Transactions on*, 5(4):347 –356, Aug. 2011.

[34] Fahim Sufi, Qiang Fang, S.S. Mahmoud, and et al. A mobile phone based intelligent telemonitoring platform. In *Medical Devices and Biosensors, 2006. 3rd IEEE/EMBS International Summer School on*, pages 101 –104, sept. 2006.

[35] U. Ayesta, L. Serrano, and I. Romero. Complexity measure revisited: a new algorithm for classifying cardiac arrhythmias. In *Engineering in Medicine and*

Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE, volume 2, pages 1589 – 1591 vol.2, 2001.

[36] Xusheng Zhang, Yisheng Zhu, N.V. Thakor, and et al. Detecting ventricular tachycardia and fibrillation by complexity measure. *Biomedical Engineering, IEEE Transactions on*, 46(5):548 –555, May 1999.

[37] Jiapu Pan and Willis J. Tompkins. A Real-Time QRS Detection Algorithm. *Biomedical Engineering, IEEE Transactions on*, BME-32(3):230 –236, March 1985.

[38] A. Lempel and J. Ziv. On the Complexity of Finite Sequences. *Information Theory, IEEE Transactions on*, 22(1):75 – 81, Jan. 1976.

[39] Shijie Zhou, Zichen Zhang, and J. Gu. Time-domain ECG signal analysis based on smart-phone. In *Engineering in Medicine and Biology Society,EMBC, 2011 Annual International Conference of the IEEE*, pages 2582 –2585, 30 2011-Sept. 3 2011.

[40] Shijie Zhou, Zichen Zhang, and J. Gu. Interpretation of coarse-graining of Lempel-Ziv complexity measure in ECG signal analysis. In *Engineering in Medicine and Biology Society,EMBC, 2011 Annual International Conference of the IEEE*, pages 2716 –2719, 30 2011-Sept. 3 2011.

[41] Daniel E. Becker. Fundamentals of Electrocardiography Interpretation. *Anesthesia Progress*, 53(2):53 – 64, 2006.

[42] P. K. Dash. Electrocardiogram Monitoring. *Indian Journal of Anaesthesia*, 46(4):251 – 260, August 2002.

[43] Malcolm S. Thaler. *The Only EKG Book You'll Ever Need*. Lippincott Williams and Wilkins, 2006.

[44] Gari D. Clifford, Francisco Azuaje, and Patrick E. McSharry. *Advanced Methods and Tools for ECG Data Analysis*.

[45] B.U. Kohler, C. Hennig, and R. Orglmeister. The principles of software QRS detection. *Engineering in Medicine and Biology Magazine, IEEE*, 21(1):42 –57, Jan.-Feb. 2002.

[46] Haiying Zhou, Kunmean Hou, and Decheng Zuo. Real-Time Automatic ECG Diagnosis Method Dedicated to Pervasive Cardiac Care. *Wireless Sensor Network*, 1:276–283, 2009. 10.1007/BF02345752.

[47] S. Mallat and W.L. Hwang. Singularity detection and processing with wavelets. *Information Theory, IEEE Transactions on*, 38(2):617 –643, mar 1992.

[48] G. Vijaya, V. Kumar, and H. K. Verma. ANN-based QRS-complex analysis of ECG. *Journal of medical engineering and technology*, 22(4):160 –167, Jul.-Aug. 1998.

[49] E. Skordalakis. Recognition of noisy peaks in ECG waveforms. *Computers and Biomedical Research*, 17(3):208 –221, Jun 1984.

[50] R. Poli, S. Cagnoni, and G. Valli. Genetic design of optimum linear and nonlinear QRS detectors. *Biomedical Engineering, IEEE Transactions on*, 42(11):1137 – 1141, Nov. 1995.

[51] M. E. Nygards and L. Sornmo. Delineation of the QRS complex using the envelope of the ECG. *Medical and biological engineering and computing*, 21(5):538 –547, Sept. 1983.

[52] P.E. Trahanias. An approach to QRS complex detection using mathematical morphology. *Biomedical Engineering, IEEE Transactions on*, 40(2):201 –205, Feb. 1993.

[53] F. A. Hensley, D. E. Martin, and G. P. Gravlee. *A practical approach to cardiac anesthesia*. Wolters Kluwer, 2008.

[54] V.K. Murthy, T.M. Grove, G.A. Harvey, and et al. Clinical Usefulness of ECG Frequency Spectrum Analysis. In *Computer Application in Medical Care, 1978. Proceedings. The Second Annual Symposium on*, pages 610 – 612, Nov. 1978.

[55] I. Jekova, A. Cansell, and I. Dotsinsky. Noise sensitivity of three surface ECG fibrillation detection algorithms. *Physiological Measurement*.

[56] Jonathan Mark. *Atlas of Cardiovascular Monitoring*. Churchill Livingstone, 1997.

[57] Rangayyan Rangaraj M. *Biomedical Signal Analysis: A Case-Study Approach*. Wiley-IEEE Press, 2001.

[58] Patrick S. Hamilton and Willis J. Tompkins. Quantitative Investigation of QRS Detection Rules Using the MIT/BIH Arrhythmia Database. *Biomedical Engineering, IEEE Transactions on*, BME-33(12):1157 –1165, Dec. 1986.

[59] P. Lynn. Online digital filters for biological signals: some fast designs for a small computer. *Medical and Biological Engineering and Computing*, 15:534–540.

[60] J. Ziv and N. Merhav. Estimating the number of states of a finite-state source. *Information Theory, IEEE Transactions on*, 38(1):61 –65, Jan. 1992.

[61] H. Zhang, Y. Zhu, and Z. Wang. Complexity measure and complexity rate information based detection of ventricular tachycardia and fibrillation. *Medical and Biological Engineering and Computing*, 38:553–557.

[62] Xusheng Zhang and R.J. Roy. Derived fuzzy knowledge model for estimating the depth of anesthesia. *Biomedical Engineering, IEEE Transactions on*, 48(3):312 –323, March 2001.

[63] X.S. Zhang, R.J. Roy, and E.W. Jensen. Eeg complexity as a measure of depth of anesthesia for patients. *Biomedical Engineering, IEEE Transactions on*, 48(12):1424 –1433, Dec. 2001.

[64] Yinlin Xu, Qianli D.Y. Ma, Schmitt, and et al. Effects of coarse-graining on the scaling behavior of long-range correlated and anti-correlated signals. In *Data Analysis, Statistics and Probability*, volume 390, pages 4057–4072, Nov. 2011.

[65] J. Kurths, U. Schwarz, A. Witt, R. Th. Krampe, and et al. Measures of Complexity in Signal Analysis, 1995.

[66] J.T. Tou and R.C. González. *Pattern Recognition Principles*. Applied mathematics and computation.

[67] Umut Orhan, Mahmut Hekim, and Mahmut Ozer. EEG signals classification using the K-means clustering and a multilayer perceptron neural network model. *Expert Systems with Applications*, 38(10):13475 – 13481, 2011.

[68] F. Kaspar and H. G. Schuster. Easily calculable measure for the complexity of spatiotemporal patterns. *Phys. Rev. A*, 36:842–848, Jul. 1987.

[69] G. D. Clifford, F. Azuaje, and P. McSharry. *Advanced Methods and Tools for ECG Data Analysis*. Artech House, 2006.