Time-Frequency and Time-Scale Analysis of Phonocardiograms with Coronary Artery Disease Before and After Angioplasty

by

Mohammad Ali Tinati

B.S., M.S., Northeastern University, Boston, Massachusetts, USA

Thesis submitted for the degree of
Doctor of Philosophy
in
Department of Electrical and Electronic Engineering
The University of Adelaide

October, 1998

Statement of Originality

I hereby declare that this work contains no material which has been accepted for the award of any degree or diploma in any university or other tertiary institution and to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of my thesis, when deposited in the University Library, being available for loan and photocopying.

Mohammad Ali Tinati

Acknowledgments

This thesis is the culmination of four years of research and could not have been completed without the help and support of a number of people. I would like to take this opportunity to thank them.

First I would like to thank my supervisor Dr. Abdesselam Bouzerdoum whose constructive criticism, guidance, and enthusiasm have been a constant source of motivation to me. I wish to thank my Co-supervisor Dr. J. Mazumdar for his help and concerns.

The research work reported in this thesis was carried out in part in the Cardiology Department of the Royal Adelaide Hospital. I would like to gratefully acknowledge everyone who made available their assistance and advice. In particular, I would like to thank Dr. L. Mahar, and Mr. R. Cowie.

Last but not least of all, I would like to thank my wife Sima and my sons Farzan and Saman for their companionship, support and understanding when much of my time was spent for preparation of this thesis.

Finally I would like to acknowledge that this thesis was financially supported by the Ministry of Culture Higher Education of Iran.

Dedicated to my wife and children.

Contents

List of Figures

List of Tables

Abstract

The objective of this study is to investigate the effects that coronary artery disease (CAD) has on heart sounds. Heart sounds are acoustic phenomenon caused by acceleration and deceleration of blood within the heart and vibration of the cardiac structure. We performed this study by analyzing the heart sounds of patients with coronary artery disease before and after angioplasty operation.

The heart sounds are recorded from patients undergoing angioplasty operation in the cardiology department of the Royal Adelaide hospital. Recordings are performed both before and after the operation using an hp3964A FM recorder and a piezo contact sensor. The recorded signals are digitized and loaded into a unix workstation for processing. Each recording process is carried out for at least 30 seconds and an electrocardiogram is also recorded synchronously.

Adaptive line enhancement filtering is used for initial uncorrelated noise cancellation. Least square errors are used for adaptation process and the pole zeros system of the conventional tapped delay line implementation is altered for better performance.

An algorithm is proposed for heartbeat cycle separation. Every long recorded phonocardiogram is decomposed into its heartbeat cycles by aligning them with the synchronously recorded electrocardiograms. A template heartbeat signal is constructed and correlation coefficients are used to select the most correlated beat cycles. Frequency

domain averaging is used to calculate a single heartbeat cycle which is used for further analysis.

Our objective is to find the locations of concentration of energy in the time frequency planes of the heartbeat cycles prepared by the beat cycle separation algorithm and to compare and contrast them before and after angioplasty operation. For this purpose we have designed an algorithm to detect the local maxima in the time frequency plane. It starts by detecting the highest energy ridge in the modulus of the time frequency matrix and identifies its boundaries and removes it from the matrix.This is continued until all the ridges are detected or a point is reached where the amplitudes of the ridges stay relatively constant.

We have used three methods to investigate the time frequency behavior of the heartbeat cycles: short time Fourier transforms, Choi-Williams distributions and wavelet transforms. The short time Fourier transform has the advantage of being easily implementable, but suffers from the resolution problem. Therefore, we used Choi-Williams distributions technique as a high resolution method. Both of these techniques use trigonometric functions as their basis functions. Wavelet transforms use localized non-trigonometric dictionaries for signal decomposition so they could be more suitable for time frequency analysis. Therefore we use them as an alternative way to study the time frequency behavior of heart sound signals.

In all three methods, regions of concentration of energy in time frequency (scale) are investigated in different bands of frequencies for first, second, and third heart sounds. We found there are certain similar time frequency patterns in the modulus of the time frequency matrices of before and after angioplasty signals. Comparing the results, it is discovered that the patterns have some average frequency shifts for 85% of the patients. The frequency shift in frequency bands below 50 Hz is less than 10 Hz, in frequency bands from 50 Hz to 150 Hz, it is about 10 to 20 Hz and the frequency shift is greatest in the third heart sound region. Above 150 Hz, frequency shifts exceed 20 Hz. In this frequency band, the frequency difference between the patterns of before and after angioplasty are almost the same in the first, second, and third heart sound regions.

List of Publications

1) M. Tinati, A. Bouzerdoum, J. Mazumdar, L. Mahar, "Modified adaptive lice enhancement filter and its application to heart sound noise cancellation" *Proc. Fourth Int. Symposium on Signal Processing (ISSPA' 96)*, pp. 815-818 August 1996, Gold Cost Australia

2) M. Tinati, A. Bouzerdoum, J. Mazumdar, L. Mahar, "Time-Frequency analysis of heart sounds before and after angioplasty" *IEEE Int. Conf. on Digital Signal Processing (DSP' 97)*, pp. 75-79, July, 1997, Santori Greece.

3) M. Tinati, A. Bouzerdoum, J. Mazumdar, L. Mahar, "Short time Fourier analysis of phonocardiograms before and after angioplasty", *Proc. Third Biennial Eng. Math. and Application Conf., (EMAC' 98),* pp. 495-498, July 1998, Adelaide Australia

Chapter 1

Introduction

1.1 General Introduction

The modern medical diagnosis techniques of the heart diseases benefit a great deal from high technological advancements made in recent decades. Very sophisticated apparatuses have been developed which help physicians in saving many lives every day around the globe, but still many people lose their lives because of heart failure. Heart diseases are the most common cause of death in most countries. Among the heart diseases, *coronary artery disease* is the most common for human beings. More than one third of fatalities are caused by heart failure and most of them are due to coronary artery disease (CAD) [56].

This chapter discusses very briefly the basic facts and concepts about the heart. The primary concern is to highlight the importance of coronary arteries in the *vascular* system, introduce the coronary artery disease, and indicate the diagnostic potentials of the heart sound in regard to CAD. In section 1.2, a brief description of the human heart is given. In section 1.3, CAD is described and pathological methods of detecting CAD are expressed. In the next two sections, the heart sounds and electrical activities of the

heart are discussed. In section 1.6, some of the recent research in phonocardiography conducted in various parts of the world, is reviewed very briefly. Then the scope of the research presented in this thesis and material covered in the following chapters are summarized.

1.2 Brief Description of the Heart

The human heart is a bio-mechanical pump which is neuro-electrically actuated and circulates millions of litres of blood throughout the body every year. The oxygen requirements of the human body is maintained by its function. The heart is the most vital organ in the human body because if it stops for a few seconds, a person will lose consciousness, and if it stops longer than that, the metabolic needs of brain tissues will be cut off and the damage to them may be permanent. The cardiovascular system of the human body consists of the heart, the *arterial vessels*, the *capillaries*, and the v*enous vessels*. Figure 1.1 shows the cross section of an adult human heart[1].

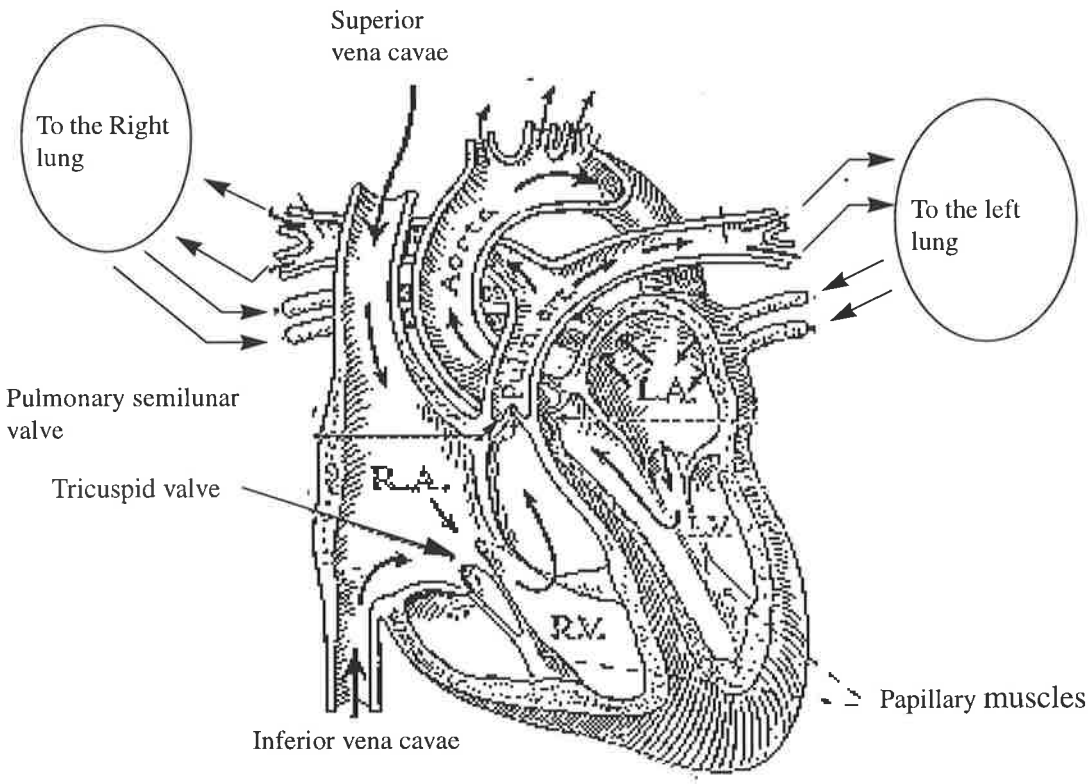


Figure 1.1 Adult human heart.

1. From A. Helfgot, "Mathematical modelling of hearts', PhD dissertaion, University of Adelaide, 1977.

There are four chambers, four valves, veins and arteries. The heart essentially functions on the basis of two basic systems that operate simultaneously. One system consists of the *left atrium* (LA), the *left ventricle* (LV), and the corresponding inlet, outlet and valves. The other system consists of the *right atrium* (RA), the *right ventricle* (RV), and the corresponding inlet, outlet and valves.

Periodical contraction followed by a relaxation of the heart muscles pumps blood out of the heart and then draws blood into the heart. The sequence of events occurring during contraction and relaxation is called the *cardiac cycle*. Contraction of *myocardiums* (muscles of the heart itself) is due to electrical activity of the heart. When an electrical impulse passes through a heart muscle, depolarization causes contraction, and when the pulse travels past the site then relaxation occurs.

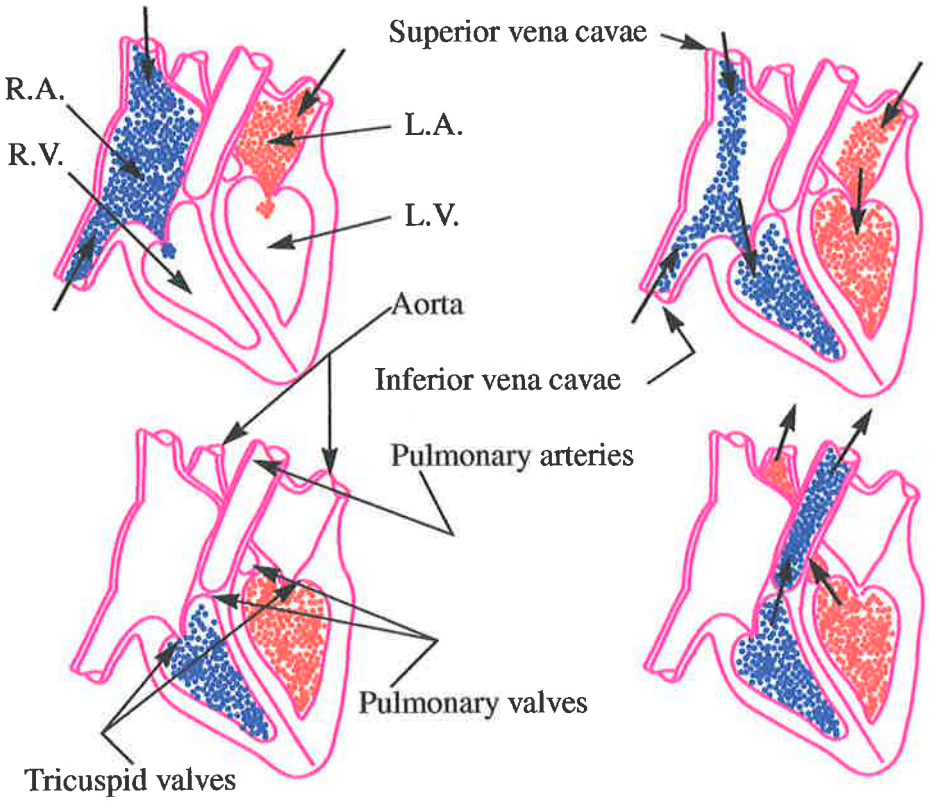


Figure 1.2 Successive stage in pumping cycle.

The oxygen-poor and carbon dioxide-rich blood from both upper and lower parts of the body enter the RA via the great veins, the superior and the inferior *vena cava*. The *tricuspid valve* (an inlet valve), which is situated between the RA and the RV chambers, acts as a one way passage and lets the blood flow from the RA to the RV.

The *pulmonary semilunar valve* (an outlet valve) facilitates unidirectional flow of the blood from the RV to the lungs via the pulmonary arteries. The oxygenized blood from the left and right lungs return to the LA by means of four pulmonary veins (two for each). The LV receives the blood from the LA through the *mitral valve*, and pumps it to the *aorta* through the *aortic semilunar valve*. These processes are shown in Figure 1.2. The filling phase of the cardiac cycle is called the *diastolic phase*, while the contraction period of the ventricular is known as the *systolic phase*, in which the heart sends out the blood.

1.3 Coronary Artery Disease

The heart needs energy in order to keep on functioning regularly. It relies almost exclusively on the oxidations of substrates for generation of energy. Normally there is a balance between the myocardial oxygen supplies and demands. The *coronary arteries* supply the metabolic needs of the heart muscles. These arteries arise from the aorta just after its origin from the left ventricle and mainly run around and down the surface of the heart. So, the coronary blood flow is an essential determinant of the myocardial performance. Blood flow into the coronary arteries takes place in both the systolic and the diastolic phases. There are two branches of the coronary arteries, left coronary artery and right coronary artery. In Figure 1.3 a picture of a human heart with coronary arteries is shown, and in Figure 1.4 the coronary blood vessels with their branches are shown in more detail [116].

In the left coronary artery, the blood flow is mainly in the diastolic phase whereas in the right coronary artery which is distributed down to the low-pressure right ventricle, the blood flow is evenly spread between the systolic and diastolic phases.

If the myocardial oxygen supplies and demands are not balanced, then that is indicative of possible occlusions in the coronary arteries. The existence of any stenoses in these veins is called coronary artery disease (CAD), which is the main cause of death in the world [56]

CAD is initiated by an abnormal reaction between *serum lipids* and *arterial intima*, causing in turn a series of changes that eventually lead to the formation of intimal plaques and other abnormalities responsible for occlusive arterial diseases [60],

[113]. There are many risk factors involved in CAD such as: diet, tobacco, absence of exercise, cholesterol level of blood, etc.

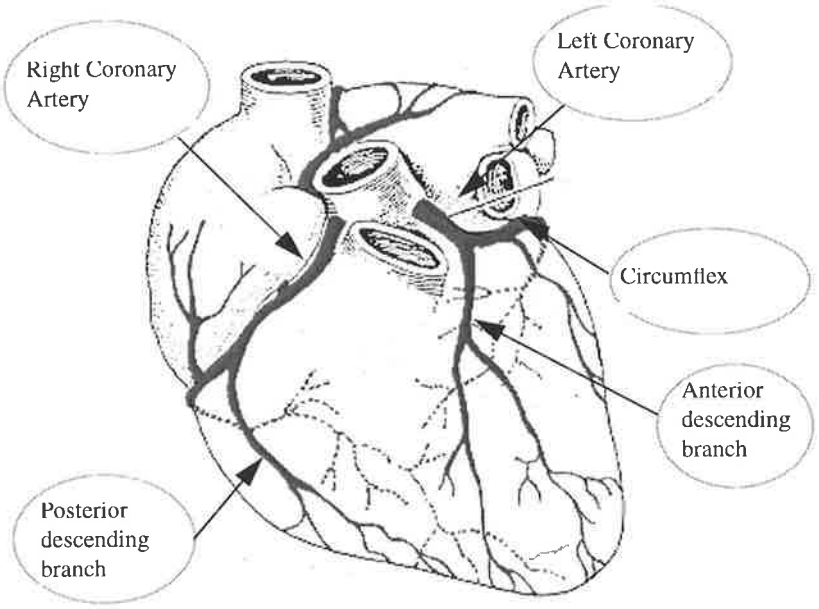


Figure 1.3 The coronary arteries.

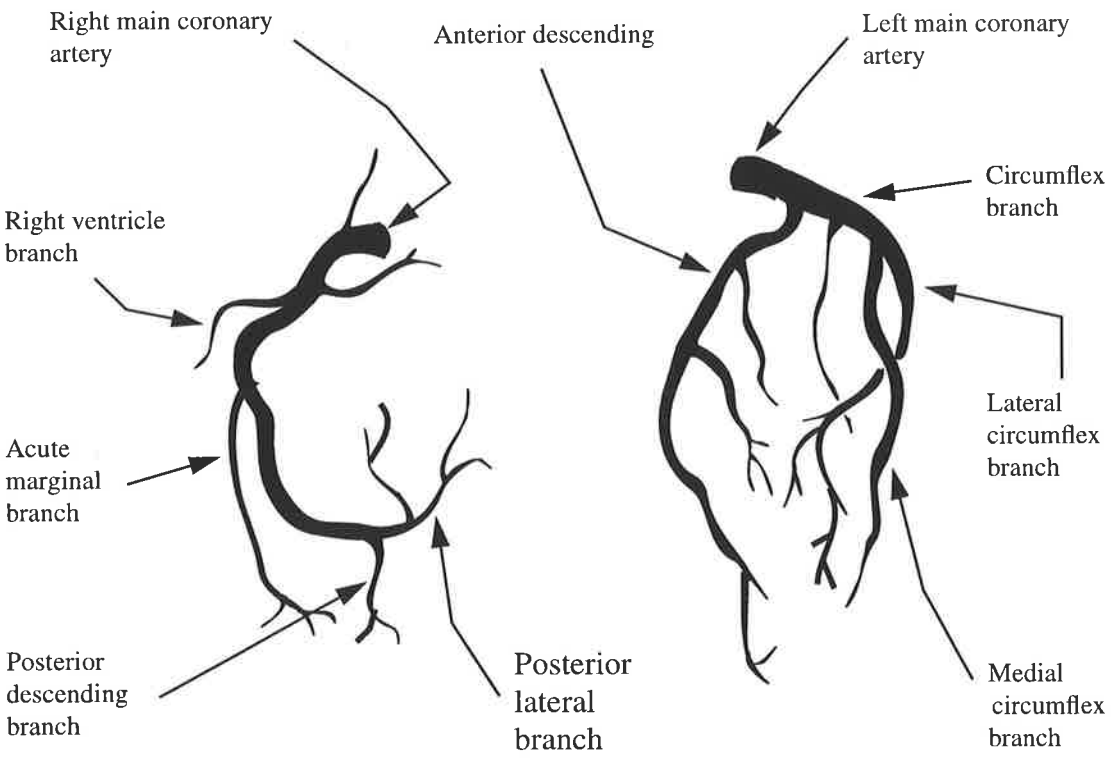


Figure 1.4 a) Right and b) Left coronary arteries.

There is no uniform presenting symptom for CAD. Chest discomfort usually is the predominant symptom. However, cases of CAD also occur in which there is not any *ischaemic* chest discomfort. The consequences are silent myocardial ischaemia, cardiac arrhythmias, congestive heart failure, and sudden death [21]. In clinical practice, there are number of techniques used by physicians to detect and diagnose CAD. These diagnostic approaches can be categorized into invasive and non-invasive methods. Some of these methods are explained very briefly in the following paragraphs.

Electrocardiography is a noninvasive method and it is probably the most frequently used technique for heart examination. Myocardial ischaemia and infarction are among the other heart diseases, besides coronary artery disease, that are assessed by this method. The abnormalities in the cardiovascular system, especially in case of coronary artery disease, reveal themselves more prominently in electrocardiograms if taken during or immediately after exercise. This is referred to as *stress test* [16], [17], [111].

Echocardiography has been one of the essential noninvasive tests for cardiovascular diseases and has become an integral part of cardiac diagnostic evaluations. Echocardiography has found many applications in the diagnosis of congenital heart diseases. The term echocardiography refers to a group of tests that use ultrasounds to assess the condition of the heart by obtaining information in the form of echoes. Coronary arteries are examined using ultrasonic transducers placed at the tip of an intracoronary catheter. The transducer then provides a cross sectional view of the coronary arteries [57].

Radionuclide scanning is a semi-invasive method for CAD diagnosis. It is based on the detection of radiation emitted from the patient following injection of a solution of a radionuclide substance such as Tallium-201 into the blood stream. An image of the heart is constructed based on emitted gamma rays, which is used for CAD detection [133].

Radiographic examination is another noninvasive method for heart imaging which involves taking chest X-rays. Since calcium is deposited early in formation of a plaque in the coronary arteries, therefore X-ray imaging can be used to monitor evolution of the CAD process. This method is rarely used for CAD diagnosis because its ability is relatively poor in detecting CAD [118].

All of these methods have their own advantages and disadvantages, but the most reliable method for CAD diagnosis is *angiography*. Angiography is an invasive method to reveal coronary artery disease and treat the stenosed coronaries. It is a painful and sometime risky operation. Angiograms are performed only for evaluation of the percentage of occlusion. In this case the operation consists of introducing a catether into the coronary arteries. Then for a very short period of time a dye is released into the blood stream and at the same time an X-ray is taken. The X-ray shows where and in what percentage occlusion has occurred.

In operations where the intention is to reopen the blocked arteries, a balloon tipped catheter is sent into the coronary arteries via the veins of the leg or the arm of the patient. The location of the stenosis is found by releasing a dye into the blood stream when the catheter has reached the heart. Monitoring the arteries visually using X-ray monitoring system reveals the exact location of the occlusion. When the balloon is placed exactly at the location of stenoses, it is inflated under high pressure. The plaque is forced to compress under the pressure where it stays compressed after the balloon is removed. This results in total elimination or considerable reduction in stenoses.

This method is very painful, patients are not given anaesthesia during the operation, it is costly, and it involves some elements of risk. There is a probability of dissection in the coronary arteries caused by pressure exerted when the balloon is inflated, which requires immediate bypass surgery. On the other hand, this method is the most reliable method for coronary artery diagnosis and treatment [21], [33], [56], [113], [116].

1.4 Heart Sounds

Acceleration and deceleration of blood within the heart and vibration of the cardiac structure give rise to a series of acoustic events called *heart sounds* and *murmurs*. The energy involved causing the acoustic events, give rise to pressure gradients in different chambers of the heart and great vessels. The vibrations of cardiac structures and acceleration of blood masses inside the vessels are the results of the pressure gradients [75]. The generated sounds are loud enough to be audible by a stethoscope applied

to the chest, or they can be recorded or registered graphically, a process referred to as *phonocardiography*. A heart sound is an acoustic event which is brief and transient with varying loudness and frequency. On the other hand, a murmur is an event which is more prolonged and has statistical characteristics, that is their characteristics may vary from beat to beat [21], [85].

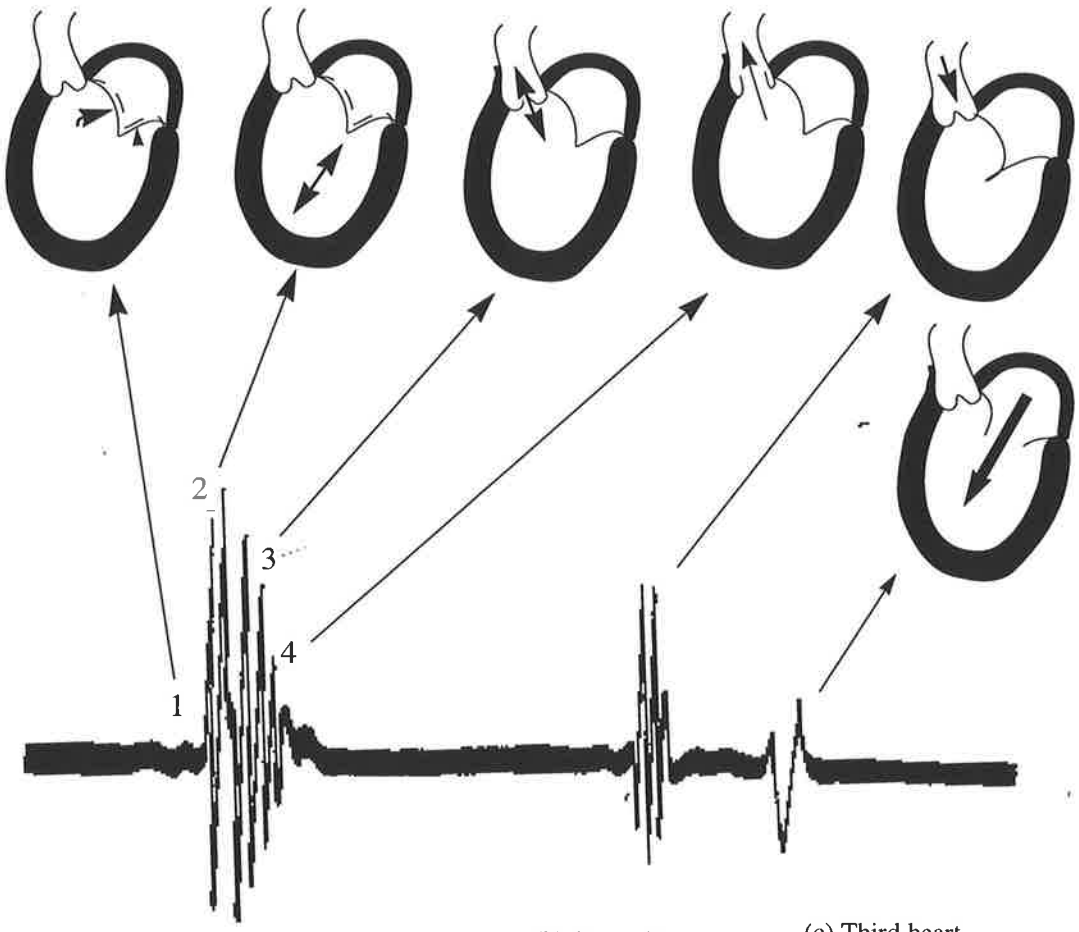
According to their occurrence in the cardiac cycle, heart sounds are categorized into four basic groups: first, second, third, and fourth heart sounds. Sometimes more descriptive terms are applied to characterize them relative to systolic and diastolic portions of the cardiac cycle. Rumser has related the contraction and relaxation events of the heart to the occurrence of the four heart sounds [107]. This is shown in Figure 1.5.

The first heart sound starts with small low frequency vibration, which are usually below 20 Hz, at the onset of rise in the left ventricular pressure. Then it is followed with a series of high frequency events, which could be up to 250 Hz, rapid rise in left ventricular pressure, change in rate of pressure rise due to opening of the aortic valve, and change of the rate of the pressure in the ascending aorta. The amplitude of the first heart sound depends on the rate of change of pressure on the left ventricle [74], [75].

The origin of the second heart sound is believed to be associated with tightening and closure of two semilunar valves; that is, the aortic and the pulmonary valves. It has been pointed out that the coaptation of leaflet is silent but rapid vibrations immediately after closure create the sound. The second heart sound has high frequency components as well [21], [74], [108], [113].

The third heart sound is an early diastolic sound and it occurs towards the end of the period of the rapid fillings of the ventricles. As for its origin, it is believed that when the ventricles are rapidly filling, there is a tendency for sudden longitudinal expansion of the ventricle walls, and it is the limitations of sudden expansion of the ventricle walls that cause the third heart sound [41], [42], [92]. The third heart sound occurs from 180 to 200 milliseconds after the second heart sound [74].

The fourth heart sound occurs in the presystolic phase of diastole. As with the third heart sound, the fourth heart sound also coincides with diastolic filling phase [130]. This sound is believed to have a similar mechanism as the third heart sound. The fourth heart sound immediately follows the first heart sound.



(a) Components of first heart sound

(b) Second heart sound

(c) Third heart sound

Figure 1.5 Schematic drawings of the causes of various components of the heart sounds based on the concept that the vibrations are induced by acceleration and deceleration of the blood within the elastic chambers.

a) The first sound can be divided into four components. The initial vibrations occur when the first myocardial contractions in the ventricle shift blood toward the artium to approximate and seal the atrioventricular valves. The second component begins with abrupt tension of closed atrioventricular valves decelerating the moving blood. It may represent oscillation of blood initiated by over distension of the atrioventricular valves, countered by recoil of the contracting ventricular myocardiums. The reaction would be similar to tapping a balloon filled with water. The third component may involve oscillations of blood between the distending root of the aorta and the ventricular walls. The fourth component probably represents vibrations due to turbulence in blood flowing rapidly through the ascending aorta and pulmonary artery.

b) The second heart sound is introduced by a few low-frequency vibrations which may accompany the deceleration and reversal of flow through the aorta and pulmonary artery prior to the closure and tensing of the semilunar valves. Although the primary vibrations occur in the arteries, they are also transmitted to the ventricles and atria by movements of the blood, valves and valve rings.

c) The third heart sound occurs at the end of the rapid filling phase. Sudden termination of the rapid-filling phase may throw the entire atrioventricular system into vibrations which have very low frequency because the walls are relaxed [107].

1.5 Electrical Activity of the Heart

The development of direct intracellular recording has provided a clear understanding of the electrophysiology of the heart [37]. A single cell is said to be polarized when there is an accumulation of positive charges outside the cell and a negative charge of equal magnitude inside it. This state is also called the resting state of the cell. When an external stimulus is applied, it changes the charge distribution around the cell membrane, and consequently the electric potential across the membrane reverses. In this condition, it is said that the cell is depolarized. When a cell depolarizes, it changes the state of the neighbouring cells and causes them to depolarise too. Thus, a chain reaction spreads out throughout the entire muscle masses of the heart.

Contraction or depolarization of the heart muscles in each cardiac cycle is initiated by an electric pulse generated at a place in the heart called *sino-atrial* (S-A) node which is located at the junction of the right atrium and the superior vena cava. As a result of such an electrical activity, an electric field is generated. The recording of the cardiac electrical cycle is called *electrocardiogram* (ECG). In Figure 1.5, an ECG and a PCG signals are shown.

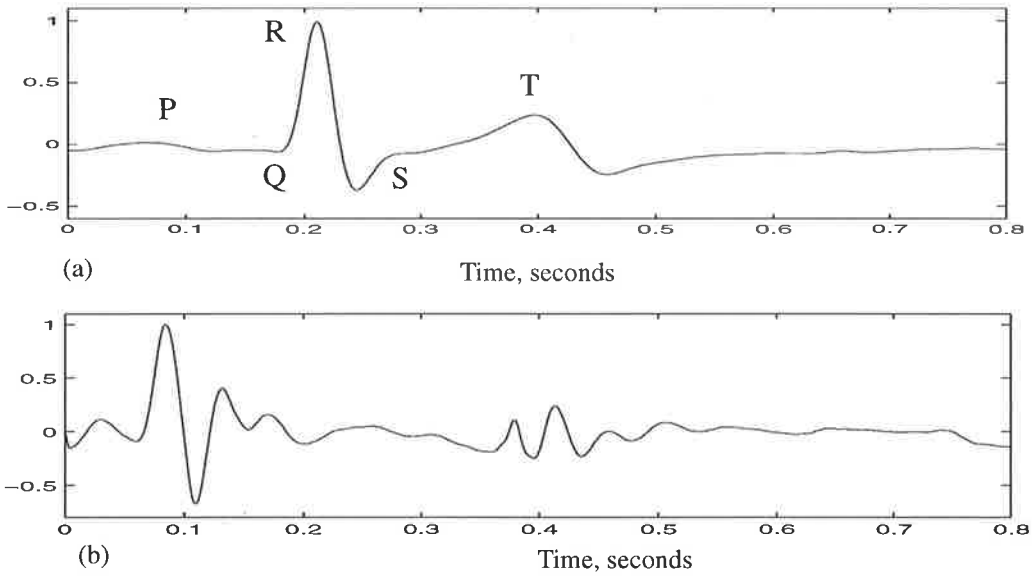


Figure 1.6 a) Typical ECG signal, showing the QRS complex and P- and T-waves, b) Typical PCG signal.

The electrical pulse at the S-A node depolarizes the atria and in turn the *atrioventricular* node. The P-wave in the ECG signal represents atrial depolarization or contraction. After this the impulse travels through fibres of the myocardium spreading

throughout the heart. Ventricular depolarization corresponds to the QRS complex and its repolarization corresponds to the T-wave in the ECG signal. Therefore the electrical cardiac cycle consists of three phases: polarization, depolarization, and relaxation.

1.6 Literature Review

Phonocardiography has been a subject of research for many years and still is. Some thirty years ago a phonocardiographic case study was conducted on a forty-nine year old man, who was believed to have coronary artery disease [40]. The patient's phonocardiogram showed an audible diastolic murmur in the frequency range of 80 to 240 Hz. It was recorded and heard easily when the patient sat up. The same results were obtained on two occasions and several days apart. The diastolic murmur was localized to the pulmonic area. The heart valves, the great vessels in chest and orifices of the coronary arteries of the patient were all normal. The position of the stenosis was recognised to be the descending branch of the left coronary artery. Study of the cross section of the thorax verified that the recognized segment for occlusion was the correct one.

Researchers have looked at the heart sounds for various purposes. Some work has been done solely for better understanding of the heart sounds, while other research objectives were to find better diagnosis techniques using phonocardiograms. Various signal processing techniques have been used to analyse phonocardiograms. A number of research works on spectral analysis, time-frequency analysis, and wavelet analysis of the heart sounds have been selected and are summarized below.

1.6.1 Spectral Analysis

In [5], series of parametric methods were applied to the isolated diastolic heart sounds to obtain spectral features of the signals. A 300 millisecond long timing window synchronized with the electrocardiogram was used to separate the diastolic portion of the heart sounds. The window was delayed from the R-wave $300-550$ milliseconds to place it just after the second heart sound. The process was carried out for over 25 heart cycles [114]. Autoregressive modelling methods using adaptive recursive lattice and gradient lattice were used to get the prediction coefficients. It was concluded

that the percentage of spectral energy at high frequencies is greater for normal subjects than patients with coronary artery disease. In another study, three eigen vector methods (MUSIC, Minimum-Norm, and Pisarenko) were used to calculate the frequency spectra of the signals. The MUSIC and Minimum-Norm algorithms produced the same estimate of the frequency peaks for all of the data, but the Pisarenko algorithm produced spurious zeros. It was found that relative power level of the second peak changed in a consistent manner after angioplastic surgery for all data [6]. In a further normal/ abnormal study, autoregressive moving average modelling was used. The autoregressive and moving average coefficients were calculated using a modified Yule-Walker technique. The analysis was carried out for 30 patients, with 20 abnormal heart sounds. The results obtained showed compliance with previous findings [7].

Linear predictive coding was used to evaluate the spectral components of the second heart sound [87], [88]. It was demonstrated that the frequency peaks of the spectrum can be enhanced by a pole enhancement technique which improves the quality of resonance in the spectrum.

In [110], spectral analysis has been performed for the first and second heart sounds. Analysis was carried out using a modified forward-backward Prony's method. Heart sounds were obtained from normal hearts and also from patients who had mechanical valves implanted. It was found that certain frequency bands were affected with the mitral valve implantation while with the aortic valves the high frequency band was affected.

1.6.2 Joint Time-Frequency Analysis

Joint time-frequency techniques such as short-time Fourier transforms, Cohen's class of distributions have been used for three dimensional representation and analysis of the heart sounds and electrocardiograms. Many analysis of the time-frequency localization of different signals have been reported [20], [19], [49], [79], [80], [83], [137]. Spectrograph and binomial time-frequency representations have been used to expose the understanding of the mechanics of the first heart sound [137], [139].

In [54], Cohen's class of distributions was used to study the time-frequency behaviour of the heart sounds. The assessment of the bioprosthetic valves was per-

formed by utilising a Bessel based kernel function. Time-frequency analysis of the heart sounds was carried out with the normal and degenerated bioprosthetic valves. It was shown that the complex frequency distribution of the heart sounds generated by the degenerated valve differs from that of the heart sound with the normal valves.

In [106], various time-frequency methods such as the short time Fourier transform, the Choi-Williams distributions and Binomial distributions are applied to the heart sounds. The heart sounds were visualised as pictures in time-frequency plane. It is concluded that the amount of information gained depended greatly on the techniques used, and higher resolution methods gave much better results.

In [139], heart sounds of eight dogs were recorded from an open chest surgery, to investigate the origin of the first heart sound. The time-frequency characteristics of the signals were studied using the binomial transform. Joint time-frequency analysis provided much better assessment of individual cardiac structures and results of the study showed that the presence of rapid frequency dynamics in the first heart sounds limits the usefulness of stationary analysis techniques for this purpose.

In [1], joint time-frequency analysis techniques were used to obtain information about the temporal variations of the spectral components of the heart sounds in a normal/abnormal study. Wavelet transforms and short time Fourier transform were utilised in identifications of normal sounds. It was concluded that both techniques provide useful information about the state of heart. The significance of some of the information was elucidated with the aid of other techniques such as echocardiography.

In [22], several generalized time-frequency distributions were used to study multi-component characteristics of intracardiac phonocardiograms. Signals were obtained from six patients using catheter-mounted transducers. Methods of study included spectrograms, binomial distributions, and Zhao-Atlas-Marks distributions. It was suggested that some of the time-frequency techniques could be utilised in the design of automated auscultation systems.

A specially designed flow simulator was used to produce sounds in the mitral valve prostheses. Time variations of power spectra of the produced sounds were studied using the short-time Fourier representation. Three types of valves were investigated

for structural deficiencies and functional abnormalities. Joint time-frequency technique was found useful in quality control purposes and identifying faulty valves [94].

In [139], the Wigner distributions and the short-time Fourier transforms were used to study the frequency behaviour of the first heart sound during the isovolumic contraction period. The heart sounds were recorded directly from the anterior left ventricular epicardium from adult dogs. Both time-frequency techniques detected a rise in frequency of the first heart sound shortly after the R-wave of electrocardiogram, but the power spectrum failed to provide and confirm that phenomenon. In [138], the same time-frequency techniques were employed to reveal the frequency dynamics of the first heart sounds of the human heart noninvasively. The multi-component behaviour of the first heart sound was observed in the joint time-frequency plane.

The short-time Fourier transform was used for 3-D representation of phonocardiograms. Filter bank summation method was used as analysis technique. The 3-D representation was used to identify the first and the second heart sounds, which had not been revealed in the recordings [115].

An automatic pattern recognition system is proposed in [85] for cardiac diagnostic purposes. The cardiac cycle was detected and its spectral features were extracted using autoregressive modelling. A data bank was generated which contained the features of the reference patterns for comparison. Arterial septal defect, cardiac valve stenoses, left bundle branch block, cardiac valve regurgitation, as well as normal heart sounds were studied.

Another automatic diagnosis system based on pattern recognition of phonocardiograms was reported in [62]. Spectral patterns of phonocardiograms were extracted using linear prediction analysis. A single beat cycle was divided into four parts and studied. It was reported that the linear prediction analysis can extract low frequency components of the first and the second heart sounds. Five different heart diseases were investigated, and it is reported that 90% correct classification was obtained by this method.

Based on statistical methods, normal and abnormal heart sounds with different cardiac diseases were processed [15]. It was assumed that heart sounds can be classified into deterministic and nondeterministic sounds. Deterministic sounds are those

events that repeat themselves in every heart-beat such as opening and closing sounds of valves. In contrast, a nondeterministic sound is one which may vary from beat to beat such as sound of randomly varying turbulent flow. The averages and variances were calculated for an ensemble of heart-beat cycles for all of the data. Variations of sounds in time was assessed for diagnostic values.

1.6.3 Wavelet Transform Analysis

In [43], heart sounds were studied using wavelet transforms. It was found that this technique is capable of distinguishing high and low frequency systolic and diastolic murmurs, and it can easily detect the temporal locations of the auscultatory gap, which is a silence that may appear when the velocity of arterial flow into the arm is reduced. Abnormal phonocardiograms were distinguished from normal. It was argued that these findings cannot be achieved using direct Fourier transform techniques.

In [3], abnormal heart sounds were studied using discrete time wavelet transforms. The intention was to detect very low frequency murmurs which fall well below audio frequency range and which cannot be identified by normal auscultatory procedures. It was found that wavelets are very useful for this application.

In [4], ways of detecting the coronary artery disease were studied using heart sounds based on neural networks and wavelet transforms. Diastolic heart sounds were extracted from phonocardiograms of patients with coronary artery disease. A feature vector is formed using the wavelet coefficients and information from physical examination of the patients. This feature vector was used as a pattern to neural networks for classifications of normal and abnormal sounds.

In [13], the phonocardiograms of patients with aortic valvular disease were analyzed. The diagnostic potential of the wavelet analysis technique was emphasized. Time-frequency characteristics of normal and abnormal cases were compared.

In [90], a comparison was performed on the analysis of the first and second heart sounds in phonocardiograms using the short-time Fourier transform, the wavelet transform, and the Wigner distribution methods. The aim was to detect the aortic valve components and pulmonary valve components of the second heart sounds and four

components of the first sound. It was shown that wavelet transforms are capable of detecting these components, but the direct Fourier transform and short-time Fourier transforms fail to do it accurately, and the Wigner distributions' results are not as clear as the wavelets.

1.7 Scope, Outline, and Contributions of the Thesis

1.7.1 Scope of the Thesis

The importance and diagnostic potentials of the heart sounds has been discussed in previous sections. It was highlighted that coronary artery disease is one of the major heart diseases. It was pointed out that the events occurring in the heart are reflected in the heart sound at different temporal locations; that is, heart-beat signal is divided into four heart sounds. We propose to apply time-frequency and time-scale techniques to phonocardiograms of patients with coronary artery disease. We intend to elucidate a mechanism by which phonocardiograms are used to investigate the changes introduced to the heart sounds by the blockage of coronary arteries. The methodology taken is to compare the features of the diseased heart sounds with its features when stenoses are removed, totally or partially.

The spectral analysis techniques such as: autoregressive moving average modelling [7], autoregressive modelling [5], linear predictive coding [87], [88], modified forward-backward Prony's method [110], used for various heart sound analysis purposes, mentioned in section 1.6, all provide only spectral information of the signal. They do not yield any temporal localization of the spectral components of the signal. Hence, in order to obtain more revealing picture of the temporal localization of the spectral components of the heart sounds, joint time-frequency techniques are chosen to study the heart sounds in this thesis. Utilization of the joint time-frequency distributions and the wavelet transforms serves two purposes: first these methods will show the temporal locations of the energy spread over the heart-beat cycle, second they will indicate how the frequency contents of all of the four heart sounds vary over time. The time-frequency techniques could aid in detection of the different sound components, which may be difficult to distinguish in the original phonocardiograms. We have selected

three techniques from many different techniques available, that is short-time Fourier transform, Choi-Williams distributions, and wavelet transforms. The short-time Fourier transform is a well-known technique and can be considered as a benchmark for comparison. The Choi-Williams distributions is an alternative method for time-frequency analysis and overcomes the shortcomings of the short-time Fourier transform. The wavelet transforms are different from other two just mentioned. In the section 1.7.2, the advantages and disadvantages of these techniques will be explained in more detail.

Time-frequency techniques were used for the analysis of phonocardiograms, but to the best of my knowledge it has not been used to analyse phonocardiograms before and after angioplasty. Understanding of the changes in time-frequency features of the heart sound of patients before and after angioplasty can help physicians to provide a better diagnosis of coronary artery disease. It may also help medical students to better understand the structure of the diseased phonocardiograms. That is, the results obtained in this study can show how the spectral components of a diseased phonocardiogram is related to temporal localization of the signal.

1.7.2 Outline of the Thesis

This thesis is divided into nine chapters. In Chapter 2, we explain the details of the recording process of phonocardiograms from patients with the coronary artery disease. The heart sound data are recorded from patients undergoing angioplasty surgery; special equipment set-up was arranged for this purpose. Recordings were carried out before and after angioplasty. Synchronised electrocardiograms were also recorded.

One of the prime concerns in a real data analysis is noise cancellation. It must be performed without distorting the signal. Obviously, if the noise spectrum does not overlap the spectrum of the signal, noise cancellation becomes a trivial job and simple filtering will accomplish the task. However, most of the time that is not the case; therefore, one of the approaches taken to reduce noise is to use adaptive filtering. In Chapter 3, the use of adaptive filters in noise cancellation from heart sounds is discussed. The kind of noise present in the recorded heart sounds is expected to be uncorrelated with the signal; therefore, an adaptive line enhancement filter is used to reduce the noise. A modified adaptive line enhancement filter is developed for this purpose and discussed in Chapter 3.

In Chapter 4, an algorithm is developed to prepare a single heart-beat sound signal to be used in the analysis carried out in the subsequent chapters. The beat cycles for each heart sound are extracted from a long data record, then they are averaged and a single heart-beat cycle is obtained. The heart-beat separation algorithm is based on correlation and uses the synchronously recorded electrocardiograms to decompose the long data record to its constituent cardiac cycles.

The time-frequency/time-scale analysis involves finding the locations of concentrations of energy in the time-frequency/time-scale planes. The moduli of time-frequency/time-scale representation of the heart sounds are considered as three dimensional images with locations of concentrations of energy as local maxima points. In Chapter 5, an algorithm is developed for detection of local maxima in time-frequency/time-scale planes.

In Chapter 6, results of applying short-time Fourier transform to heart sound analysis are presented. Although this technique has resolution restrictions, it is easy to implement and it has less computational load compared to other methods of time-frequency calculation. The most suitable analysis window has been found to be the chebychev window for this application. A method is developed for recognition of time-frequency patterns of local maxima.

In Chapter 7, the application of another time-frequency technique in heart sounds analysis is investigated. The Wigner distribution is an alternative method for time-frequency analysis. It overcomes the resolution restriction of the short-time Fourier transform, but it has the severe problem of cross terms, which are terms generated due to bilinear structure of the Wigner distribution; further explanation of cross terms can be found in Chapter 7. In some applications, such as heart sound analysis, cross terms highly restrict its usefulness. In order to reduce the effect of cross terms, the Choi-Williams distribution is used. This method has a great computational load; therefore, an algorithm is developed to calculate the time-frequency matrix for the Choi-Williams distribution with a minimum number of multiplications. The algorithm is presented in Appendix B.

In the past decade wavelets have emerged as a tool for signal processing, they provide a unified, consistent, and rigorous formulation of multi-scale data analysis. In

Chapter 8, the continuous wavelet transform is used for time-frequency analysis of the heart sounds. Morlet wavelets are chosen as basis functions. The modulus of the wavelet transform of the heart sounds is used as the time-scale matrix and concentrations of energy places are located and then time-frequency patterns before and after angioplasty are compared.

Finally in Chapter 9, a comparison is made of the results obtained from Chapters 6,7, and 8. The main conclusions to the thesis are stated and the scope of the future work is outlined.

1.7.3 Contribution

The work described in this thesis contributes to the understanding of differences in the time-frequency behaviour of the heart sounds with coronary artery disease before and after angioplasty. In Chapter 3, a modification to the adaptive line enhancement filter is made to improve its noise cancellation performance. The proposed method introduces poles into the pass-band region of the all zero tapped delay line enhancement filter. An algorithm for updating the coefficients of the new filter is derived based on the method of least-squares; this algorithm is presented in Appendix A. In Chapter 4, an algorithm to extract a single heart-beat signal is proposed. A thresholding technique is introduced to determine the beginning of the cardiac cycle in the electrocardiograms. This is used to decompose the synchronously recorded phonocardiograms to heart-beat cycles. Frequency domain averaging is used to prepare the final heart-beat signal. In Chapter 5, an algorithm is developed to identify the locations of concentrations of energy in the time-frequency plane; it is called the local maxima detector. The method is based on detecting the highest energy level of the time-frequency matrix, then determining its boundaries, removing it from the time-frequency matrix, and repeating the process until all local maxima are detected. A criterion to segregate local maxima due to signal from those due to noise is also developed. In Chapter 6, a method is proposed to compare the local maxima of the heart sounds of patients before and after angioplasty. The method is based on the energy levels, temporal locations, and frequency locations of the local maxima. In Appendix B, an efficient method for calculations of the Choi-Williams distributions is proposed.

Chapter 2

Data Acquisition

2.1 Introduction

Auscultation of the heart involves actively listening to the sounds made by the contracting heart muscles, movement of blood in the vessels, and movement of the heart valves throughout the cardiac cycle. Assessment of the function of the cardiac system is usually the purpose of auscultation. It is popular and very simple to carry it out, and it is a noninvasive method. Auscultation is performed with the aid of an instrument called a *stethoscope* which minimally consists of a flat diaphragm attached to a headset via a flexible tubing. Graphical auscultation of the heart sounds can be regarded as a phonocardiogram, where the stethoscope is replaced with a microphone, and the output is directed to a recording apparatus.

The purpose of this chapter is to illustrate the process of recording of the heart sounds and electrocardiograms for this study. In section 2.2, a historical view is made of the heart sounds as a means for cardiac assessment. Then the areas of the chest

where phonocardiograms are recorded is explained in section 2.3. The heart sounds that are used in this thesis are recorded from patients undergoing angioplasty in the cardiology department of the Royal Adelaide hospital. We have set up an equipment to carry out the recordings of both the heart sounds and electrocardiograms. In section 2.4 to 2.6, the recording process of heart sounds and utilization of the equipment are explained. In section 2.7, physical conditions, location of the stenoses of the coronary arteries, and percentage of the occlusion before and after angioplasty of the patients whose phonocardiograms were recorded are presented.

2.2 Historical review

Probably for many centuries before the modern time, physicians put their ear to the chest of patients and listened and felt impulses and sounds, although there are no records available to indicate that. It was in the 17^{th} century that Carvisart and Bayle applied their ears to the surface of the chest and described the sounds that were caused by heart beat [73].

The first stethoscope which was made of wood was invented by Laenne in 1819 [11], [73]. The main objective of the invention was to avoid direct listening with the ear to the heart beat, which was considered indecent and inconvenient. Laenne also described the first and the second heart sounds, but it was not until the 19^{th} century that the wrong interpretations of the heart sounds were corrected, the fundamental basis for the cardiac auscultation was established, and the outstanding acoustical features of the stethoscope were recognized.

The first recordings of the heart sounds were made by Huerthhle in 1895 using an elementary microphone. The first clinical phonocardiogram was recorded by Lewis by means of a string galvanometer [73], [112]. Subsequently, modifications to the recording process were made by others. Good progress was made to phonocardiographic recordings as a result of availability of electronic amplification which became possible with the development of vacuum tubes.

Today, the auscultation of the heart is performed with properly fitted stethoscopes that possess a diaphragm for high frequencies and a bell for medium and low frequencies. The examiner can readily switch from one chest piece to the other [35], [71]. The heart sounds for phonocardiographic recordings are picked up by highly sensitive piezoelectric sound sensing devices, and recorded with specially designed frequency modulated (FM) recorders. Development of digital signal processing techniques has made it possible to analyse the phonocardiograms for various clinical applications and diagnostic purposes.

2.3 Areas of chest for phonocardiographic recordings

Traditionally there are six areas on the human chest from where phonocardiographic recordings and ausculation are usually carried out. These areas have been categorized from two lines of studies: 1) by comparison of auscultation with those of catheterization and surgical findings, 2) observations of phonocardiograms and their evaluations. These areas are briefly explained here and are shown in Figure 2.1 [73].

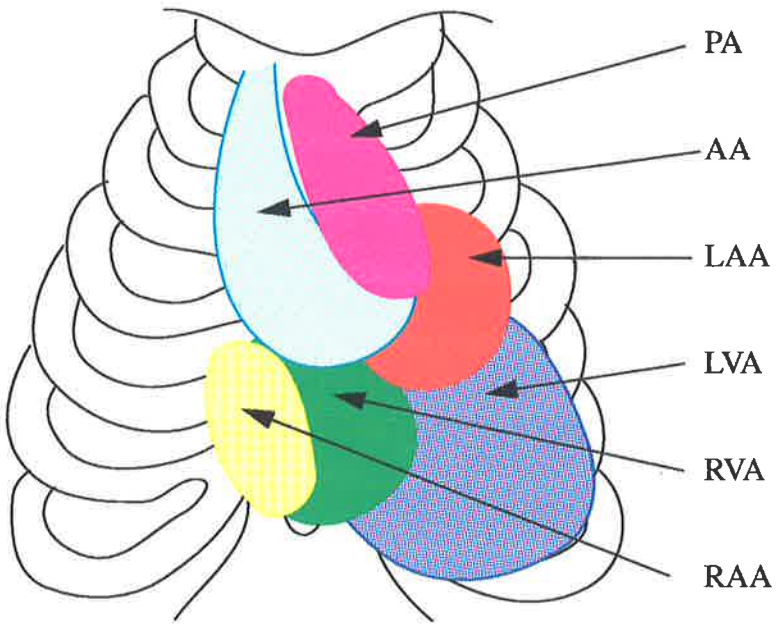


Figure 2.1 Areas of chest for phonocardiography [74].

- **Left Ventricle Area, LVA**

This area is the best location to hear any abnormality in the heart sounds caused by stenoses in the mitral valve and abnormality in the left ventricle and the left atrium. Aortic insufficiencies can be heard in this area as well. This area centres around the apex.

- **Right Ventricular Area, RVA**

This area is the best location to hear any abnormality in the heart sounds caused by stenoses in the tricuspid valve and abnormality in the right ventricle and the right atrium. Pulmonary insufficiencies can be heard in this area as well. This area includes the lower part of the sternum and the 3^{rd} and 4^{th} interspaces on both sides of the sternum.

- **Left Atrial Area, LAA**

The area to the left of the apex and above it is considered to be the left atrial area. This area is also well suited for cases with mitral insufficiencies.

- **Right Atrial Area, RAA**

Any abnormality in the tricuspid valve can be heard or recorded in this area. It extends one to two cm to the right of the sternum in the 4^{th} and 5^{th} interspaces.

- **Aortic Area, AA**

In this area vibrations caused by abnormalities in the aorta, and abnormalities in the carotid are heard or recorded. This area starts at the 3^{rd} left interspace and extends up to the 1^{st} right interspace.

- **Pulmonary Area, PA**

This is the location to hear or record abnormalities in the pulmonary arteries. This area corresponds to the 2^{nd} and 3^{rd} left interspaces close to the sternum.

2.4 Signal Acquisition Process

A schematic representation of the heart sound data acquisition system is shown in Figure 2.2. An HP model 21050A piezoelectric contact sensor was used to pick up

the heart sounds from the PA, lower part of RVA, and LVA areas of the chest of patients with coronary artery disease in the cardiology department of the Royal Adelaide Hospital. The output signal of the contact sensor was in microvolt level. The signal was amplified to volt level and then lowpass filtered with an antialiasing filter before it was digitized. The cutoff frequency of the lowpass filter was set equal to 1 *kHz*. This frequency was chosen because the sampling frequency was set equal to 2 *kHz*. The output of the lowpass filter was applied to channel one of an FM recorder. The Hewlett-Packard HP 3964A four channel FM tape recorder was used. Simultaneously, lead II electrocardiogram was recorded in channel two of the FM recorder.

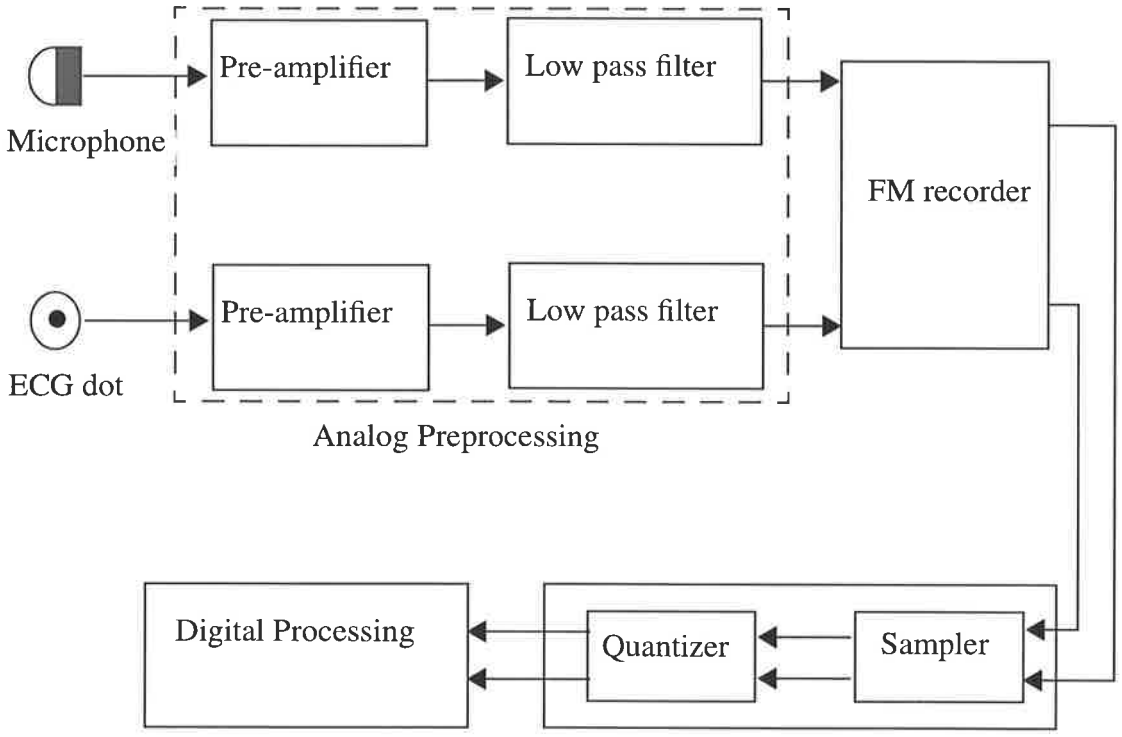


Figure 2.2 Block diagram of the data acquisition system for heart sound.

Sampling and quantizing, [91],was carried out on a Pentium PC using Labview. The sampling frequency was set to 2 *kHz*. The reason for selecting 2 *kHz* is that it is believed that the heart sounds do not contain frequency components higher than 1 *kHz*. The analog-to-digital resolution of the hardware used was 12-bit. About 25 to 45 seconds of each of the recorded data were digitized and stored for further processing. The

sampling, quantization and storing processes were carried out for both the phonocardiograms and the electrocardiograms simultaneously.

We have made 42 recordings. Each recording was carried out in two stages. First the heart sound and the electrocardiogram of the patient were recorded before the angioplasty operation. Then the recordings were repeated in the same day about an hour after angioplasty.

2.5 Pick-up device for phonocardiography

The human ear is a marvellous apparatus, but it is poorly suited for cardiac auscultation because of the characteristics of the cardiac sounds. The hearing threshold frequencies of normal people are about 40 *Hz* and 15 *kHz*. However, in heart sounds there are some low frequency components that are below the hearing threshold frequency of most people. To understand the clinical implications of the phonocardiograms and to be able to recognise certain events in heart sounds, there is need for special training.

Since heart sounds are acoustic waves, some kind of microphone is needed to pick them up. One option is to use magnetic microphones, but they are not suitable because they typically have a low cut off frequency of around 50 *Hz*, which is too high for heart sounds. Other factors such as low impedance, low sensitivity (typically about 1 *mV/Pascal*) make them useless for the heart sound recordings.

Condenser microphones are another type of microphone used for acoustical pick-ups. Their operations are based on the capacitance change caused by the movement of the diaphragm. These microphones are not suitable for phonocardiography either. They exhibit better frequency response characteristics than the magnetic microphones, that is they have lower cut off frequency (around 20 *Hz*), but this is still considered high for the heart sounds. Other features of these microphones are high impedance, higher sensitivity (typically around 10 *mV/Pascal*), and higher noise figure. These characteristics put restriction on using them for the heart sound recordings.

Another type of pick-up device is a piezoelectric sensor. Suitable wafers or slabs of crystalline or ceramic materials exhibit an electrical polarization that changes

with mechanical deformations. They generate a potential difference between opposite faces when subjected to torsional or bending stress. This phenomenon is called the *piezoelectric* effect. The usual piezoelectric microphone comprises a sandwich of oppositely polarized slabs joined to form a single unit with metal electrodes attached to them. Vibrations caused by external pressure give rise to a voltage difference at the output terminals proportional to the effective displacement. This kind of microphone is suitable for phonocardiographic recording. The Hewlett Packard HP 21050A has been used as a pick-up device for recordings in this study. This pick-up device is specially designed to pick-up heart sounds for diagnostic references in biomedical studies. Its characteristics are shown in Table 2.1. It is also supported with a perforated rubber belt and belt adapter that allow the sensor to be strapped firmly to the patient. It operates on the basis of displacement of the contact sensor tip with respect to the housing.

Table 2.1

Sensitivity	Force: $10\mu V$ / dyne Displacement: $40mV$ / micron
Frequency response, in hertz	$0.02 \leq F \leq 2000$
Capacitance, in nanofarads	$20nF$

2.6 Pre-amplification and Low-pass Filtering

Pre-amplification and low-pass filtering constitute an analog preprocessing stage which is required to amplify and filter the heart sounds and the electrocardiograms. The amplifying stage is essential in order to satisfy the amplitude requirement of the FM tape recorder and lowpass filtering is aimed for anti-aliasing purposes.

We have designed and built a dc high gain amplifier using low noise operational amplifiers. The amplifying stage is followed with a low-pass filter. A 10-pole butterworth filter with cut off frequency of 1 *kHz* has been designed and built using active analog elements. The design of the pre-amplifier and low-pass filter circuits will not be discussed here, as their designs are standard. The frequency response of the equipment (microphone, FM recorder, and filters) was checked by applying sinusoidal signals and examining the spectrum of the responses; thereby the correct operation of the equipment was verified.

2.7 Data records

Table 2.2 shows a description of stenoses for the patients used for this study before and after angioplasty. The patients were chosen to cover a vast range of age and severity of the disease as well as other pathological conditions. These conditions are shown in Table 2.2.

In Table 2.2, the second column represents the location of the stenoses. The third column is the percentage of occlusion of the stenoses before the angioplasty, while the fourth column shows the percentage of occlusion of the stenoses left after the operation. In the fifth column, blood pressures of the patients during the operation are listed. The numbers on the left are systolic pressures and those on the right are diastolic pressures. The last column shows the age of the patients.

The length of each recording was between 45 and 60 seconds. Each recording was performed in three areas of the chest; that is in the pulmonary area, right ventricle area, and left ventricle area. Recordings were carried out at patients' bedside with patients lying on the bed with their chest fully exposed.

Table 2.2

	stonoses	% before	% after	pressures S.	D.	age
1	LAD LCFX	%90 %70	%10 %0	110	60	71
2	RCA LCFX	%70 %80	%10 %10	140	80	63
3	RCA LCFX	%60 %99	%5 %20	110	65	50
4	LAD	%99	%10	100	60	38
5	LAD	%90	%30	100	50	75
6	LAD RCA	%80 %50	%20 %0	130	70	75
7	LCFX	%90	%0	150	65	78
8	LAD	%99	%0	140	80	60
9	LCFX	%90	%20	100	60	57

	stonoses	% before	% after	pressures S.	D.	age
10	RCA	%90	%0	100	60	57
11	RCA LAD	%90 %99	%10 %25	120	75	84
12	LAD	%99	%35	110	60	47
13	RCA	%80	%0	100	60	43
14	RCA	%99	%20	110	75	61
15	LAD	%80	%0	110	75	52
16	LAD	%90	%0	120	80	66
17	RCA	%60	%20	110	65	65
18	LCFX	%90	%10	130	65	60
19	RCA	%90	%10	170	100	45
20	LAD	%90	%25	120	80	70
21	LCFX RCA	%99 %70	%0 %0	110	60	70
22	LAD RCA	%90 %100	10% 25%	105	50	60
23	RCA	%85	%0	100	55	55
24	LAD (2 lesions)	%90	%0	113	65	75
25	LAD RCA	%60 %50	0% 0%	155	80	65
26	LAD	%70	%0	144	75	64
27	LAD LCFX	%50 %90	%0 %30	110	62	63
28	LCFX	%70	%20	100	55	72
29	RCA	%95	%10	130	70	75
30	LCFX	%70	%10	160	80	70
31	LAD	%100	15%	90	60	51
32	RCA LCFX	%90 %60	%10 %0	160	90	41
33	LAD	%80	%0	110	70	30
34	RCA	%90	%0	120	65	57

	stonoses	% before	% after	pressures S.	D.	age
35	RCA LAD LCFX	%100 %90 %95	10% 10% 20%	120	80	69
36	LAD	%75	%10	170	100	70
37	LAD	%90	%5	120	65	30
38	RCA	%85	%0	140	70	42
39	LAD	%90	%20	95	60	39
40	LAD LCFX	%95 %70	%10 %10	150	60	52
41	RCA	%80	%0	140	80	63
42	RCA LAD	%70 %90	%0 %10	150	60	56

RCA: right coronary artery
LCFX: lateral circumflex
LAD: left anterior descending.

In Figure 2.3 and Figure 2.4 typical recordings from Table 2.2 are shown. These figures show the phonocardiograms and the electrocardiograms of cases 1 and 10 of Table 2.2, respectively. In parts (a) of the figures, the phonocardiograms before angioplasty are shown, whereas in parts (c) the same patient's heart sound after angioplasty are plotted. In parts (b) the synchronously recorded electrocardiograms before angioplasty operation are shown and in parts (d) the synchronously recorded electrocardiograms after angioplasty are shown.

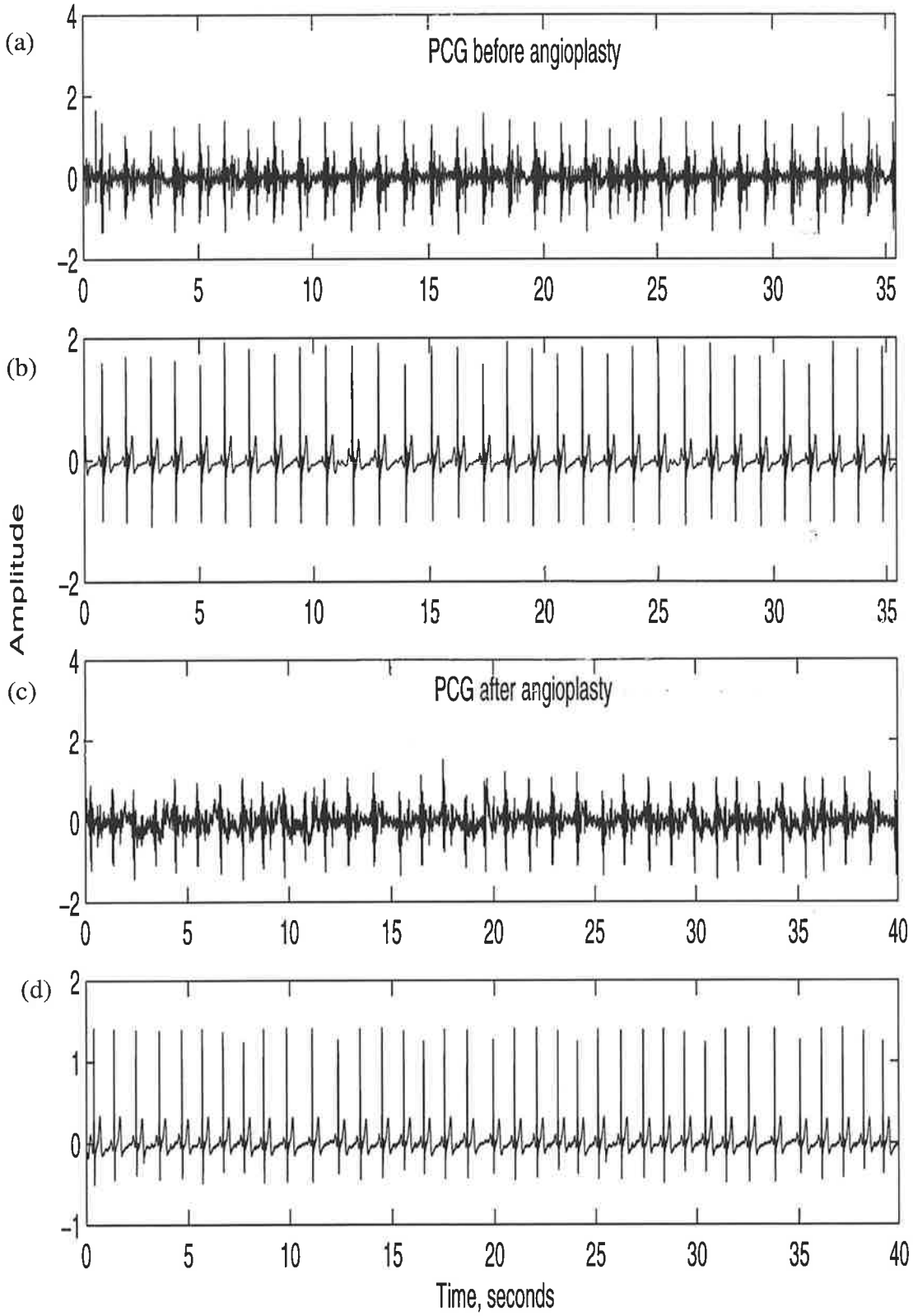


Figure 2.3 Phonocardiograms and electrocardiograms of case 1 before and after angioplasty.

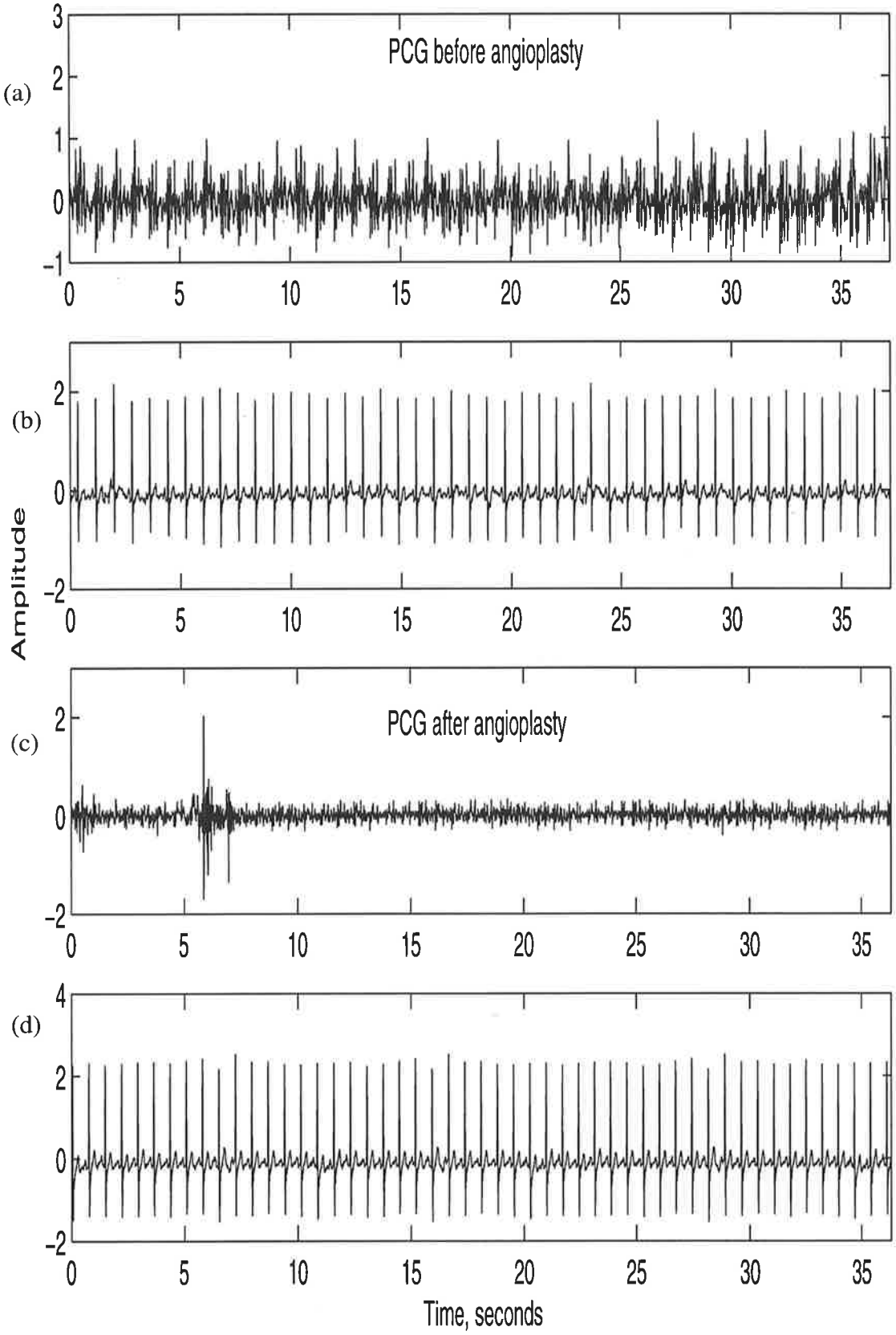


Figure 2.4 Phonocardiograms an electrocardiograms of case 10 before and after angioplasty.

Chapter 3

Adaptive Line Enhancement Filtering of Heart Sounds

3.1 Introduction

Filtering constitutes one of the major building blocks in signal processing. Filters were originally used as circuits and systems with frequency selective behaviour, but for many applications it is necessary to design more sophisticated filtering systems than collections of tuned circuits. As a result, there have been growing developments in filter design theory. Over the past four decades, adaptive filtering is one of the major developments in filter theory. It has been successfully used for many years.

Noise cancellation is one of the largest applications of filters. It is usually performed by passing a noisy signal through a fixed impulse response filter. However, this method is not a very efficient way to tackle the noise problem if the frequency spectrum of the signal and noise overlap. It is an efficient way to remove only out of band unwanted frequencies. Whenever the noise signal has a wide-band spectrum, this method is not applicable at all. In many signal processing applications such as biomedical engineering, communication networks, radar monitoring systems, control systems

and others, usually a narrow-band signal is contaminated by a wide-band noise. Therefore adaptive filters are used.

It was discovered in the 1960's that adaptive filtering is an efficient way of cancelling wide-band noise from narrow-band signals. Therefore, as an alternative method to estimation of signals from additive interferences, a new adaptive noise suppression technique was developed [134]. Since then its use has become widespread in many areas of signal processing especially in biological signal processing [8], [63], [121], [124], [135]. Adaptive filtering has found applications in many areas: sometimes the frequencies of sinusoids are of interest, sometimes an enhanced signal is desired, and sometimes a signal with time varying frequencies is tracked.

Generally, adaptive filters use a reference noise signal to adapt their impulse response. The adaptive line enhancement (ALE) filter is a special case where no reference noise signal is required for retrieval of signal from noise. The steady state behaviour of an ALE filter with an input composed of multiple sinusoids with additive white noise is analysed in [143].

In order to emphasize the applicability of the adaptive and ALE filters, we will briefly review some of their applications in biological signal processing areas. In [63], adaptive filters were used to minimize the effect of the heart sounds on lung sounds. A 300 coefficient finite impulse response filter was used. A least-mean-square algorithm was used for adaptation of the filter. It was feasible to reduce the heart sound from breath sound by 50-80 percent. In [8], an adaptive line enhancement filter was used for background noise cancellation of the diastolic heart sounds. The lattice implementations with filter order of 10 and decorrelation delay of 7 was used. The performance of the filter was assessed by spectral analysis of the signals. It was found that ALE reduced the background noise of the signals. One of the areas of applications of the ALE filters is noise cancellations of the electrocardiograms (ECG). In [124] a multi-input adaptive filter was used for noise cancellation of ambulatory ECG signal for arrhythmia detection purposes. A reference noise signal was obtained for various noises that appeared in ambulatory ECGs. In [135], a reference signal based adaptive power-line noise cancellation from ECG signal was proposed.

The objective of this chapter is to investigate the application of adaptive line enhancement filter for background noise cancellation from the heart sound signals. We intend to improve the signal-to-noise ratio (SNR) of the recorded phonocardiograms by utilizing ALE filters. The organization of this chapter is as follows. In sections 3.2 to 3.6, we briefly review adaptive filters, adaptation algorithms, and adaptive line enhancement filters. This is followed by an explanation of the pole-zero diagram of a converged tapped delay line ALE filter. Then a new implementation for transversal tapped delay line adaptive line enhancer is introduced in section 3.7. We will call it the Modified ALE filter. We should mention that our intention is not to go very rigorously in design details of adaptive filters, rather we are seeking to improve the SNR of the heart sounds and therefore we modify the ALE configuration to the extent that it can serve our purpose. In section 3.8, the performance of the modified ALE filter by a sinusoidal signal contaminated with noise is illustrated. Then in section 3.9, the adaptation algorithm for the modified ALE filter is explained and its derivation is given in appendix B. In section 3.10, heart sound signals are applied to the ALE and modified ALE filters and results are compared.

3.2 Adaptive Filters

Adaptive filters have two major differences with non-adaptive filters [61]. First, their internal parameters can be used to control their transfer function over a useful range. Second, the features of the external environment can be tracked by the transfer function of the filter. This is accomplished by an adaptation algorithm. The adaptation algorithm acts as a control element which shapes the internal parameters of the filter according to features of the external environment. Parameter variations imply that these filters are time- or shift-variant. However, the variations are very slow relative to the highest frequency component of the input signal.

A block diagram of an adaptive filter is shown in Figure 3.1. In this figure, the signal to be filtered is the *primary input*. The output of the filter is compared with another input, the *desired input*. The difference between the output of the filter and the desired input, is called the *error signal*. The objective is to make the error signal as

small as possible. A performance index is defined, which usually is in the quadratic form of the error signal. Then the performance index is minimized and used to update the impulse response parameters, $h(n)$, of the filter.

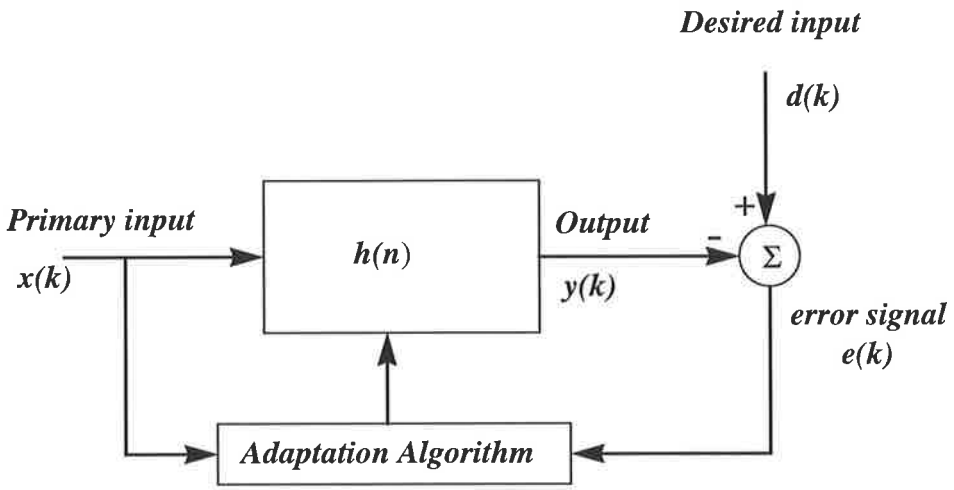


Figure 3.1 Block diagram of an adaptive filter.

3.3 Structures for Adaptive Filters

There are many ways in which an adaptive filter can be realized. The most popular ones are: transversal (also known as finite impulse response (FIR) filters), recursive, and lattice filters. Transversal filters are easily understood and analysed, and they are stable but they have a finite-duration impulse response. Recursive filters have infinite duration impulse responses and sharp spectral cutoff, but they can become unstable. Lattice filters are essentially the same as transversal filters except that their order can be updated when deemed necessary. Since we are going to use transversal filter for this application, we will not discuss any aspects of recursive and lattice filters.

An N-tap transversal tapped delay line filter is shown in Figure 3.2. The input vector to the filter is composed of the N most recent samples of input signal $\boldsymbol{x}(k)$. The input vector to the filter taps and unit sample response of the filter are denoted, herein by $\boldsymbol{s}(k)$ and $\boldsymbol{h}(k)$, respectively.

$$\mathbf{s}(k) = \left[x(k)\ x(k-1)\ x(k-2)\ \ldots\ x(k-N+1)\right]^T \quad \text{Eq 3.1}$$

$$\boldsymbol{h}(k) = [h_0, h_2, h_2, \ldots, h_{N-1}]^T \quad \text{Eq 3.2}$$

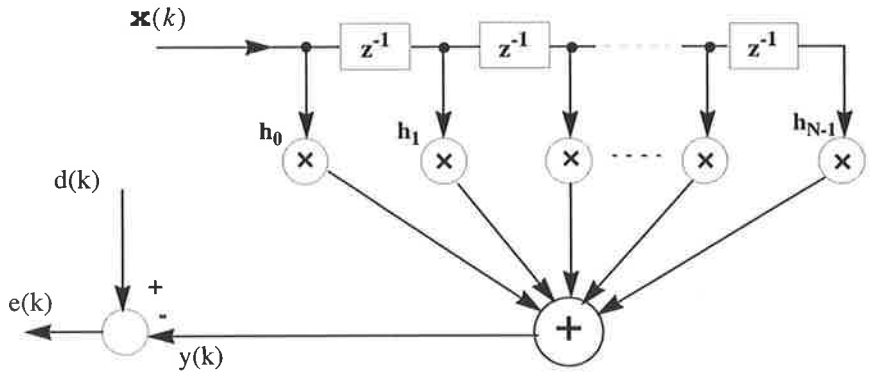


Figure 3.2 Tapped delay line implementation of FIR filter.

3.4 Adaptation Algorithms

Adaptation is accomplished by the adaptive algorithms. The objective is to generate an output, $y(k)$, which results in a minimum error, $e(k)$, with respect to a reference signal, $d(k)$. The error signal is expressed as

$$e(k) = y(k) - d(k) \quad \text{Eq 3.3}$$

and it should be as small as possible. The minimization criterion is usually in the quadratic form. The most common criteria are the mean squared error (MSE) and the least squares (LS).

3.4.1 The Mean Squared Error Criterion

The basis for mean squared error algorithms is the statistical properties of data signals. The input signals $\mathbf{x}(k)$ and $d(k)$ are assumed to be wide-sense stationary discrete random processes with known autocorrelation and power spectra. The expected value (denoted by $\mathcal{E}$) of the quadratic form of the error signal is the objective function to be minimised. Therefore the performance index is defined as

$$J = \mathcal{E}\{e^2(k)\} = \mathcal{E}\{[d(k) - y(k)]^2\} \quad \text{Eq 3.4}$$

The output of the filter is equal to the convolution of its primary input with its impulse response.

$$y(k) = \sum_{j=0}^{N-1} h_j x(k-j) \quad \text{Eq 3.5}$$

Substituting Eq 3.5 into Eq 3.4 results in

$$J = \mathcal{E}\{[d(k) - \boldsymbol{h}^T(k)\mathbf{s}(k)]^2\} \quad \text{Eq 3.6}$$

or

$$J = \mathcal{E}\{d^2(k)\} - 2\mathcal{E}\{d(k)\mathbf{s}^T(k)\}\boldsymbol{h}(k) + \boldsymbol{h}^T(k)\mathcal{E}\{\mathbf{s}(k)\mathbf{s}^T(k)\}\boldsymbol{h}(k) \quad \text{Eq 3.7}$$

The first expectation in Eq 3.7 is simply the mean squared power of the desired signal. The second expectation is the cross-correlation vector between the desired input and primary input. The last expectation is the autocorrelation matrix of the primary input. These terms are defined as

$$\sigma^2_d = \mathcal{E}\{d^2(k)\} \quad \text{Eq 3.8}$$

$$\mathbf{P}_N = \mathcal{E}\{d(k)\mathbf{s}^T(k)\} \quad \text{Eq 3.9}$$

$$\mathbf{R}_N = E\{\mathbf{s}(k)\mathbf{s}^T(k)\} \quad \text{Eq 3.10}$$

Substituting Eq 3.8 to Eq 3.10 into Eq 3.6 yields

$$J = \sigma^2_d - \mathbf{P}_N + \boldsymbol{h}^T\mathbf{R}_N\boldsymbol{h} \quad \text{Eq 3.11}$$

To get the optimum impulse response parameters of the filter, $\boldsymbol{h}$, it is essential to minimize Eq 3.11 with respect to $\boldsymbol{h}$. The matrix $\mathbf{R}$ is non-negative, symmetric, and Toeplitz [76], hence invertible. The gradient of J is given by

$$\nabla J = -2 \cdot \mathbf{P}_N + 2 \cdot \mathbf{R}_N \cdot \boldsymbol{h}$$

A null gradient yields the optimal impulse response vector as

$$\boldsymbol{h}^o = \mathbf{R}^{-1}{}_N \cdot \mathbf{P}_N \quad \text{Eq 3.12}$$

Eq 3.12 is called the *Wiener-Hopf* equation [144]. The surface of the J is depicted in Figure 3.3 for $N = 2$.

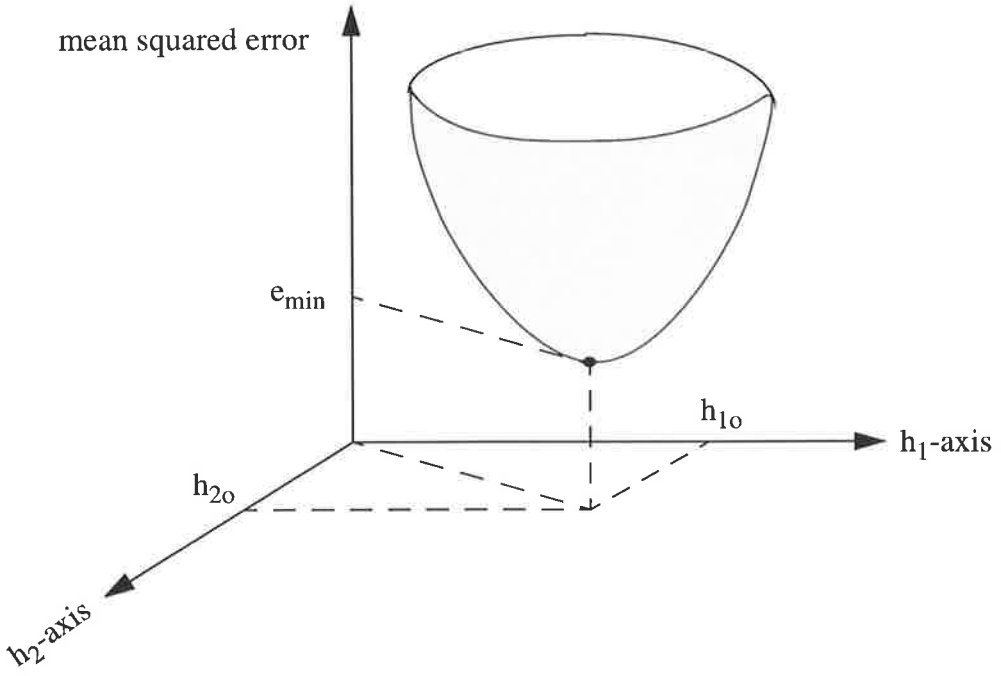


Figure 3.3 Quadratic error surface and optimum impulse response parameters.

3.4.2 The Least-Squares Algorithm

An alternative method for MSE is the least squares (LS) method which does not require a prior knowledge of the second order statistics of the signal. The performance index is defined to be the time average of the squared error. That is, the objective of the LS method is to determine the impulse response parameters of the filter which minimize the sum of squared errors. Therefore Eq 3.4 becomes

$$J = \sum_{k=1}^{L-N} e^2(k) \qquad \text{Eq 3.13}$$

where L is the number of observations in input sequence $\mathbf{x}(k)$. The autocorrelation matrix, $\mathbf{R}_N$ of Eq 3.10 changes to that of Eq 3.14 and the cross-correlation vector $\mathbf{P}_N$ of Eq 3.9 changes to that of Eq 3.15.

$$\mathbf{R} = \sum_{i=1}^{L-N} \mathbf{s}(i)\mathbf{s}^{T}(i) \qquad \text{Eq 3.14}$$

$$\mathbf{P} = \sum_{i=1}^{L-N} d(i)\mathbf{s}(i) \qquad \text{Eq 3.15}$$

where $\mathbf{s}(i)$ denotes the i^{th} input vector at the filter taps.

The minimization process is the same as MSE; that is, the gradient of J is calculated and set to zero to get the optimum solution. The optimum values of the impulse response parameters of the filter are calculated as

$$h^{o} = \mathbf{R}^{-1} \cdot \mathbf{P} \qquad \text{Eq 3.16}$$

3.5 Transversal Finite Impulse Response Filters

As mentioned before tapped delay line is the most widespread implementation structure for adaptive filters [61], [98], [135], [144]. A FIR filter has an output whose state depends on a finite number of past inputs. The direct form implementation of a FIR filter was shown in Figure 3.2, and was referred to as tapped delay line or transversal implementation. It is nonrecursive because it does not contains any feedback. Eq 3.17 represents the difference equation of the filter, and Eq 3.18 expresses its transfer function.

$$y(k) = \sum_{i=0}^{L-1} h_i x(k-i) + e(k) \qquad \text{Eq 3.17}$$

$$H(z) = \sum_{i=0}^{L-1} h_i z^{-i} \qquad \text{Eq 3.18}$$

It is clear that this structure requires *L-1* memory locations for storing the previous inputs and also *L-1* additions and *L* multiplications are performed to compute each output point. A FIR filter whose unit sample response satisfies the symmetry or asymmetry conditions of Eq 3.19 have linear phase [98]. In this case, the number of required multiplication reduces by half. Also the zeros of the filter have a symmetrical arrange-

ment; if z_k and z_k^* are a complex-conjugate pair of zeros then $1/z_k$ and $1/z_k^*$ will also be zeros of the system. This leads to a simplified system function.

$$h(k) = \pm h(L-1-k) \qquad \text{Eq 3.19}$$

3.6 The Adaptive Line Enhancement Filters

Adaptive filters are used very frequently in noise cancellation problems. It was originally proposed in [134] and the method was studied by others. No prior knowledge of signal or noise is required for its application.

Adaptive line enhancement filter (Figure 3.4) is developed for adaptive noise cancellation and has been the subject of numerous studies [2], [100], [101], [127], [142], [144]. It is used in applications where an external reference for the additive noise is not available. The input signal is delayed and applied to a tapped delay line transversal filter whose output is compared and subtracted from the original input. The error signal is used to change the impulse response parameters of the filter.

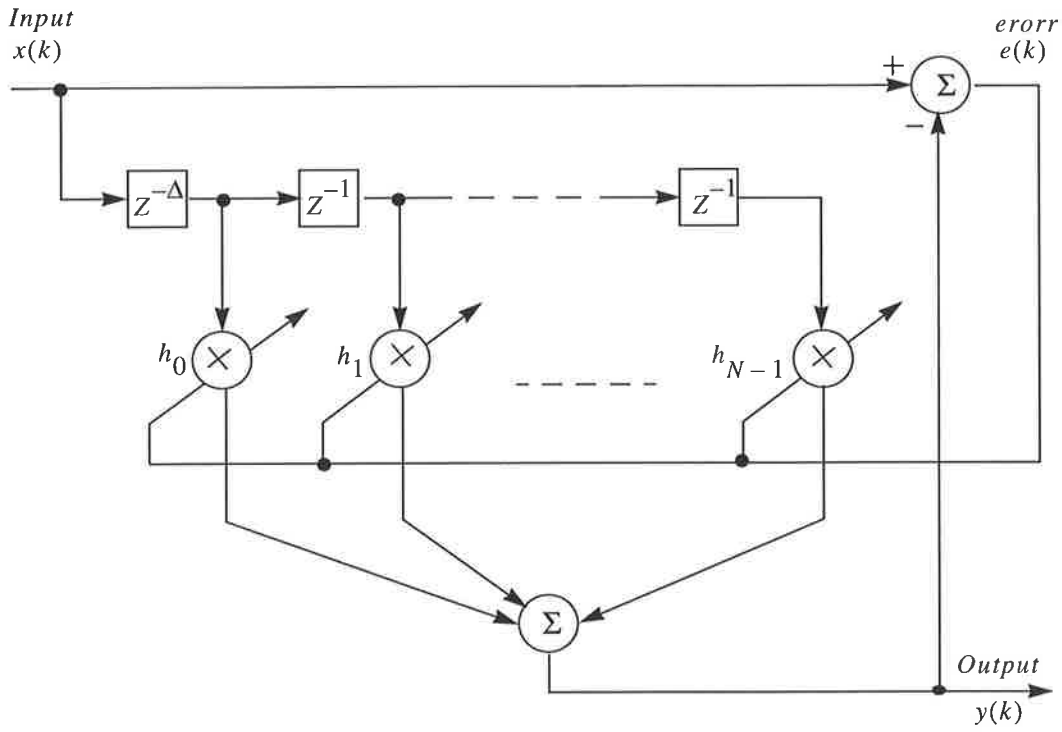


Figure 3.4 Adaptive line enhancement filter.

The operation of the adaptive line enhancer of Figure 3.4 is best understood if we consider its input to be a sinusoidal signal in white noise. The Δ delay line introduces only a simple phase shift to the sinusoidal component of the input signal while it decorrelates the noise component in the filter input with respect to those in the primary input. It should be noted that the noise and the signal must not be correlated. After a number of operations of the adaptation algorithm, the filter acts as a band-pass self-tuning filter where its pass-band covers the frequency band of the signal. The noise component of the input signal is rejected at the output. It has been argued that proper choice of Δ can significantly improve the quality of the output signal [55], [101], [142]. The input to the filter consists of two components

$$x(k) = \upsilon(k) + w(k) \qquad \text{Eq 3.20}$$

where $\upsilon(k)$ is composed of the sum of sinusoidal signals (Eq 3.21), and $w(n)$ is a broad-band additive white noise component with zero mean, independent of the sinusoidal components of the signal.

$$\upsilon(k) = \sum_{i=1}^{p} A_i \sin(\omega_i k) \qquad \text{Eq 3.21}$$

The output of the filter is calculated as

$$y(k) = \sum_{i=0}^{N-1} x(k-\Delta-i)h_i \qquad \text{Eq 3.22}$$

3.6.1 The Poles and the Zeros of the ALE Filters

The transfer function of the ALE filter can be derived from the steady state weights (impulse response parameters, $\mathbf{h}$) of Figure 3.4 as follows.

$$H(z) = h_0 + h_1 z^{-1} + \ldots + h_{N-1} z^{-(N-1)} = \sum_{i=0}^{N-1} h_i z^{-i} \qquad \text{Eq 3.23}$$

$$G(z) = z^{-\Delta} \sum_{k=0}^{N-1} h_k z^{-k} = z^{-\Delta} H(z) \qquad \text{Eq 3.24}$$

where $G(z)$ is the transfer function from the input $x(k)$ to the narrow-band output $y(k)$, and $H(z)$ is the transfer function of the tapped delay line portion of the filter.

Let the input signal consist of a single sine wave, $\nu(n)$, with an added white noise. The signal $\nu(n)$ is assumed to have an amplitude of A and normalized radian frequency of ω_i, its signal power is represented by σ_ν^2. The noise signal is assumed to have a zero mean and variance σ_w^2. It is also assumed that the signal and the noise are independent. The signal-to-noise ratio can be calculated as

$$SNR = 10 \cdot \log_{10}(\sigma_\nu^2/\sigma_w^2) \qquad \text{Eq 3.25}$$

Using Eq 3.16, the steady state values for $\boldsymbol{h}$ can be found as

$$\sum_{i=0}^{1} \phi_{ss}(l-k)h_i^o = \phi_{sx}(l+\Delta) \qquad 0 \le l \le N-1 \qquad \text{Eq 3.26}$$

where $\phi_{ss}(l-k)$ represents the elements of the autocorrelation matrix of $\mathbf{s}(k)$, and $\phi_{sx}(l+\Delta)$ represents the cross-correlation vector of $\mathbf{x}(k)$ and $\boldsymbol{s}(k)$. Eq 3.26 is solved for $\boldsymbol{h}^o$ with the method of undetermined coefficients [143]. They are found to be

$$\mathbf{h}^o{}_k = \frac{1}{\lambda}(\gamma e^{j\omega_0 k} + \gamma^* e^{-j\omega_0 k}) \qquad k = 0, 1, \ldots, N-1 \qquad \text{Eq 3.27}$$

where γ^* is the complex conjugate of γ. λ and γ are given as

$$\lambda = (N + 2\sigma_\nu^2/\sigma_w^2)^2 - (1 - \cos 2\omega_0 N)/(1 - \cos 2\omega_0) \qquad \text{Eq 3.28}$$

$$\gamma = (N + 2\sigma_\nu^2/\sigma_w^2)e^{j\omega_0\Delta} - e^{-j\omega_0\Delta}(1 - e^{-j2\omega_0 N})/(1 - e^{-j2\omega_0}) \qquad \text{Eq 3.29}$$

By substituting Eq 3.27 into Eq 3.24, location of the zeros of the transfer function of the filter can be found as

$$G(z) = \frac{z^{-\Delta}}{\lambda} \sum_{i=0}^{N-1} [\gamma e^{j\omega_0 i} + \gamma^* e^{-j\omega_0 i}] \qquad \text{Eq 3.30}$$

From the above argument we realise that the h-parameters of the ALE filter will be in sinusoidal form when the filter converges. The ALE filter is an all zero filter and when the signal-to-noise ratio of the input signal is high enough, then its zeros tend to

lie on the unit circle. The arrangement of the zeros are such that the filter acts as a band-pass filter. It is sometimes referred to as self-tuning filter. The pass-band of the filter corresponds to the locations on the z-plane where there is a lower concentration of zeros of the filter. In other words, the impulse response of the filter will be a cosine function whose frequency equals the centre frequency of the pass-band.

To illustrate the point, suppose a tapped delay line configuration has converged with a sinusoidal signal as its input signal. Based on the above discussion, the filter coefficients can be express as

$$h_i = \alpha^{(i-1)} \cdot \cos(\omega_0 i) \cdot e^{-\mu i} \qquad \text{for } i = 0, 1, \ldots, N-1 \qquad \text{Eq 3.31}$$

where ω_0 is the normalized centre frequency of the filter, α is the radius of the locations of the zeros inside the unit circle, and μ determines the out of band attenuations. Where i is the order of the filter taps. This is illustrated in Figure 3.5.

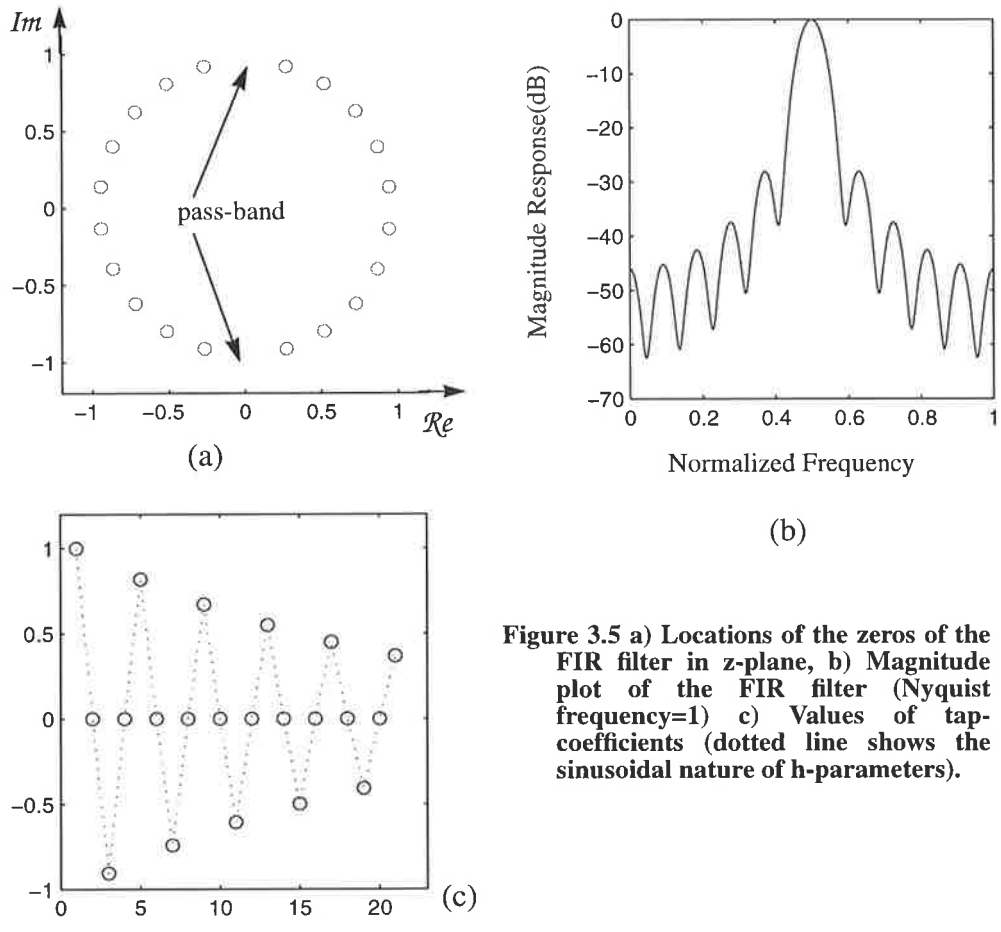


Figure 3.5 a) Locations of the zeros of the FIR filter in z-plane, b) Magnitude plot of the FIR filter (Nyquist frequency=1) c) Values of tap-coefficients (dotted line shows the sinusoidal nature of h-parameters).

In Figure 3.5 (a), the location of the zeros of the FIR filter are shown. In Figure 3.5 (b) the magnitude of the transfer function of the filter is plotted which shows that it is a band-pass filter. In Figure 3.5 (c) the coefficients of the impulse response of the filter are shown. To illustrate the point more clearly, we present a simulation example illustrated in Figure 3.6.[1]

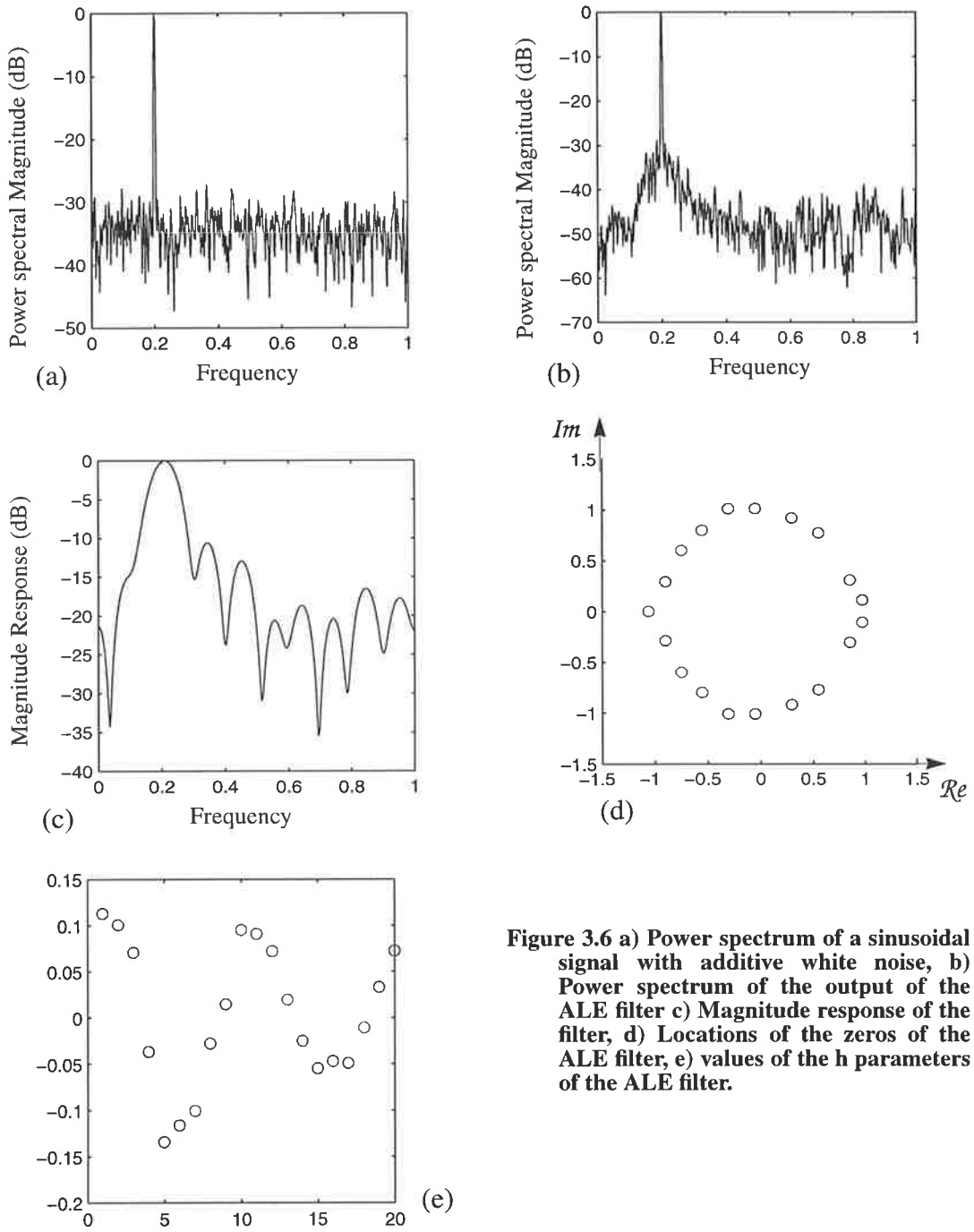


Figure 3.6 a) Power spectrum of a sinusoidal signal with additive white noise, b) Power spectrum of the output of the ALE filter c) Magnitude response of the filter, d) Locations of the zeros of the ALE filter, e) values of the h parameters of the ALE filter.

1. Note that in this figure and other figures here after Nyquist frequency =1 unless otherwise stated.

A white noise signal with zero mean was added to a sinusoidal signal. The SNR was set equal to 0 dB. The length of the ALE filter was chosen to be 20 and Δ was set equal to 15. The LS algorithm for adaptation was used and it was allowed to run for 1000 iterations. Theoretically, the larger filter length will result in better performance of the filter, but the computation will be intensive [2], [135]. The optimum value of Δ depends on the signal and noise powers [101]. Simulations were carried out for different values of filter length while keeping Δ constant. Three values of Δ were chosen, 10, 15, 20. For each value of Δ the filter length was varied from 5 to 30 in steps of 5. For each simulation, the power spectral magnitude of the output signal was plotted. Those parameter values, filter length of 20 and Δ=15, that resulted the lowest out of band mean power, were chosen. That is, for each parameter set the power spectral magnitude of the output of the filter was compared with the previous ones. Figure 3.6 shows the magnitude spectra of both input and output signals, the locations of zeros of the ALE filter, and the magnitude response of the converged filter taps. The locations of the zeros of the filter in the z-plane show that the filter has self-tuned itself to the frequency of the input signal (0.2 *Hz* with normalised Nyquist frequency). Note that for calculation of the optimum filter parameters, a program was written and the algorithm proposed in [34] was used.

3.7 Modification of the Pole-Zero Diagram of the ALE Filter

Figure 3.6 (b) shows that although the variance of the noise signal has decreased in the vicinity of the signal frequency, the noise power has not changed much. In other words the SNR of output signal has improved but not over all spectral range. In order to attenuate the noise level in frequencies that fall near the pass-band region of the filter, we propose to add poles in this region. This will make the system an IIR filter. As long as the poles are inside the unit circle the filter will be stable, and will exhibit a sharper magnitude response. This is illustrated in Figure 3.7. The pole-zero diagram of Figure 3.5 (a) is modified by adding a complex pole to it as shown in Figure 3.7 (a).

The poles are placed in the pass-band of the filter and are located inside the unit circle but very close to it. The magnitude response of the modified ALE filter is shown in Figure 3.7 (b) (solid-line), and for comparison purposes, the magnitude response of

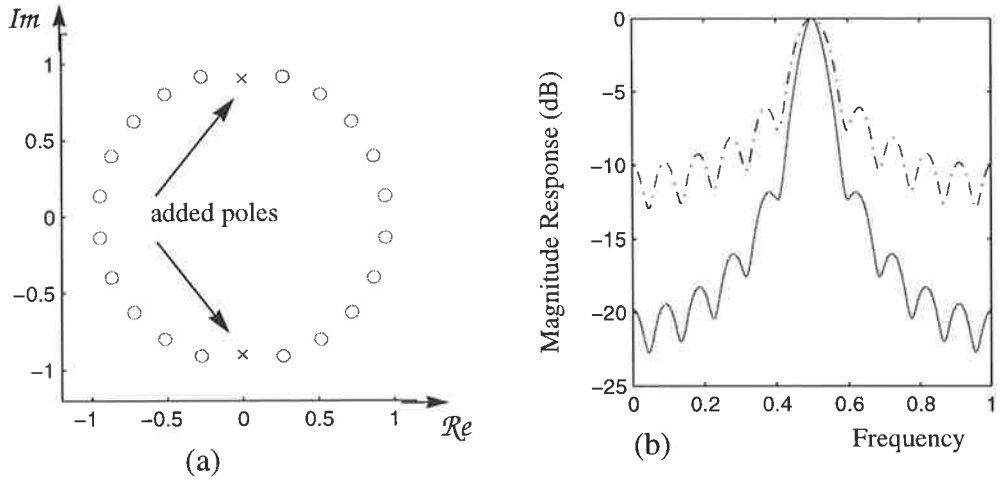


Figure 3.7 a) The modified ALE filter (a) pole-zero diagram with added poles in the pass-band, b) magnitude response of the modified filter (Solid line), magnitude response of the filter shown in Figure 3.5 (a) (dashed-line).

the original ALE filter (all zero) is shown in Figure 3.7 (b) (dashed-line). Comparison of the two plots reveals that the modified ALE filter attenuates the out-of-band frequencies more than the all-zero ALE filter. The frequencies in the vicinity of the pass-band are also attenuated more.

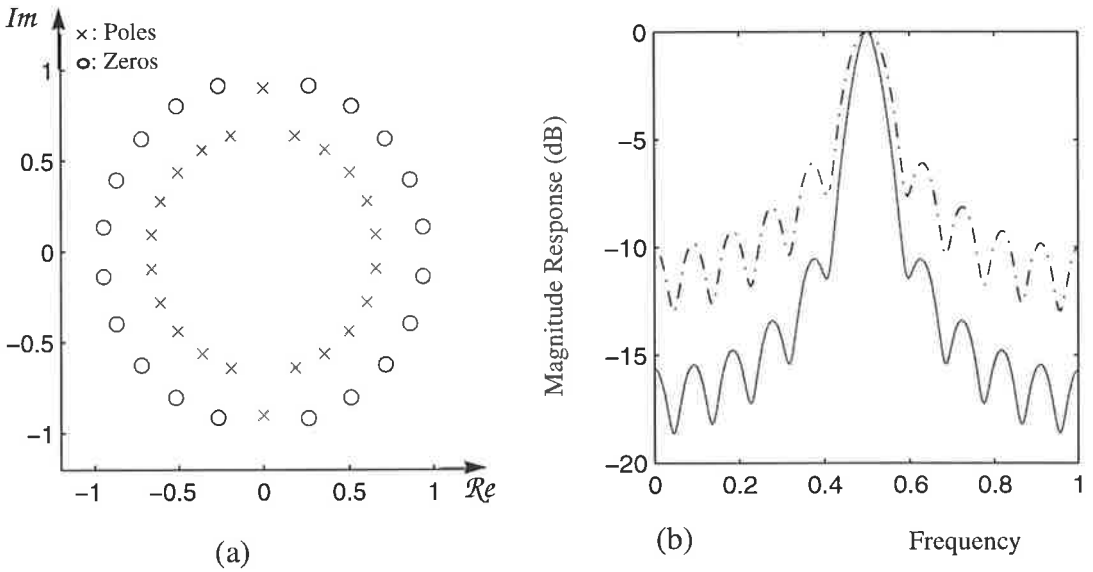


Figure 3.8 a) Modified pole-zero diagram of the filter shown in Figure 3.5 (a), b) magnitude response plot for the modified ALE filter (solid-line), and the conventional ALE filter (dashed-lined).

We have observed that adding poles in the pass-band region of the ALE filter improves the signal-to-noise ratio of the output signal. What happens if we add poles in

the stop-band region as well? The poles-zero diagram of the all-zero filter of Figure 3.5 (a) has been altered as shown in Figure 3.8 (a). That is, poles have been added in the pass-band region as well as stop-band region where they are placed on a circle with radius smaller than one located inside the unit circle. The number of the poles are equal or greater than the number of zeros. In the pass-band region we have added the same poles as were added in Figure 3.7 (a). The magnitude response of the filter is shown in Figure 3.8 (b)

Comparison of Figure 3.8 (b) and Figure 3.7 (b) shows that adding poles in the stop-band region only has slightly decreased the out of band attenuation. The fact is that as the stop-band poles get closer to the origin of the unit circle then the magnitude response curve of the modified ALE shown in Figure 3.8 (b) approaches that of the modified magnitude response plot of Figure 3.7 (b).

3.8 Modification to ALE implementation

Our main purpose for modification of the ALE filter is to get a better signal-to-noise ratio in the output of the filter. As was explained in the previous section this can be achieved by adding poles in the pass-band region of the pole-zero diagram. In order to obtain the desired modification, it is necessary to alter the implementation of the basic ALE filter of Figure 3.4. We propose the implementation shown in Figure 3.9. The output of the filter is multiplied by a scaling factor β, and delayed D samples, then processed with a scaled tapped delay line. The impulse response parameters of the filter, denoted by h_i's in the figure, are updated using the LS algorithm.

Note that multiplication of the h-parameters to α^i in Figure 3.9 changes the radius of the roots of H(z) by α.

Assume that the filter has converged and the optimum filter parameters are obtained. Then the output of the filter is given by

$$y(k) = \sum_{i=0}^{N-1} h_i x(k-\Delta-i) + \sum_{i=0}^{N-1} \alpha^i h_i y(k-D-i) \qquad \text{Eq 3.32}$$

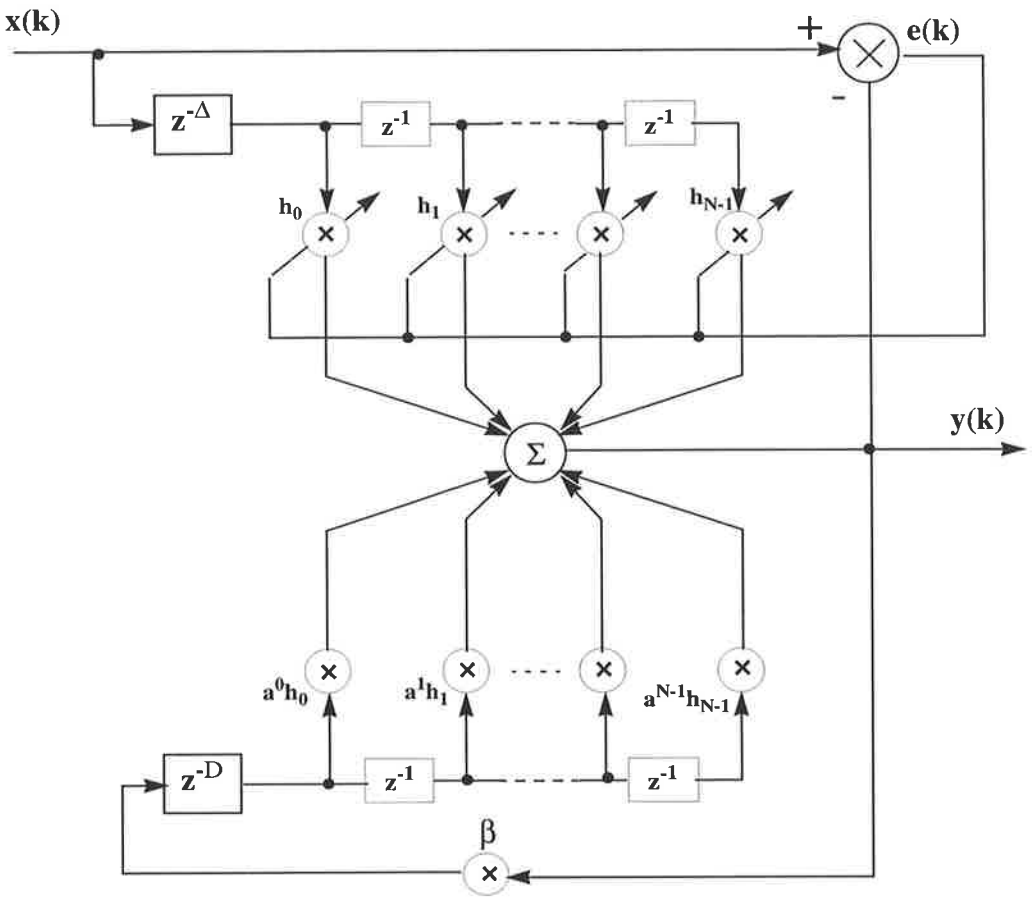


Figure 3.9 Modified adaptive line enhancement filter.

The steady state transfer function of the filter can be obtained from Eq 3.32 as

$$G(z) = \frac{z^{-\Delta}\sum_{i=0}^{N-1} h_i z^{-i}}{1-\beta z^{-D}\sum_{i=0}^{N-1} a^i h_i z^{-i}} \qquad \text{Eq 3.33}$$

or

$$G(z) = \frac{h_0 z^{N-1} + h_1 z^{N-2} + \ldots + h_{N-1}}{z^{\Delta-D}[z^{D+N-1} - \beta(h_0 z^{N-1} + \alpha^1 h_1 z^{N-2} + \ldots + \alpha^{N-1} h_{N-1})]} \qquad \text{Eq 3.34}$$

This equation shows that there are $N-1$ zeros and $D+N-1$ poles in the modified ALE filter. Of course there will be $\Delta - D$ poles (or zeros if D is greater than Δ) in the origin as well. The locations of the poles are determined by the values of α,

β, and h-parameters of the filter. The polynomial inside the brackets of the denominator of Eq 3.34 has $D+N$ coefficients and $D+N$-1 roots which constitute the poles of the filter. The first coefficient is always one and after that there are N coefficients equal to zero and the rest of the coefficients are scaled versions of h-parameters of the filter.

We used the signal shown in Figure 3.6 (a) to test the performance of the modified ALE filter. Results are shown in Figure 3.10. The following parameter values were used: $N = 20$, $\Delta = 10$, $D = 11$, $\alpha = 0.8$, and $\beta = 5$. In choosing the parameter values, we started from the initial values of N and Δ which were found with conventional ALE filter, that is we set $N = 20$ and $\Delta = 15$. The value of D was set equal to the value of Δ, and the values of α and β were set equal to one. Then the response of the filter was determined. The pole-zero diagram of the filter was examined. We also compared the power spectral density of the output with that of the previously obtained conventional ALE to reveal that if the out of band frequencies were attenuated more. This was repeated every time we changed any parameter. In this process, we investigated the effect of value of α on the locations of poles and zeros. As was expected, the value of α directly affected the location of poles. It was revealed that when α was close to one (greater than 0.9) then the stop-band poles of the filter were located very close to the unit circle and in the pass-band region they even resided outside the unit circle; therefore, the filter became unstable. When the value of α was small (less than 0.5) then all of the poles of the filter (both pass-band and stop-band poles) were close to the origin of the unit circle and therefore the filter acted like the conventional ALE. It was found that β had a great effect on the transient response of the filter. It also effected the locations of the poles of the filter. Therefore, we varied β in conjunction with α. By changing the values of Δ and D, first in steps of 5 and then in steps one, we found that the above mentioned set of values gave the best results.

In Figure 3.10 (a) the magnitude spectra of the input signal, the output of ALE filter, and the output of the modified ALE filter are shown. As the figure shows, the noise power in the vicinity of the sinusoidal component, has been reduced in the output of the modified ALE compared to the output of the ALE filter. In Figure 3.10 (b) the pole-zero diagram of the modified ALE is shown. The zeros are indicated by "o" and poles are show by "x". As the figure shows, a conjugate pair of complex poles are gen-

erated close to the unit circle in the pass-band region. All other poles are located at a radius of about 0.8 inside the unit circle. The pass-band poles are indicated as PBP in the figure.

As was mentioned earlier and we should emphasize here too that we don't intend to get involved in recursive adaptive filter design which does not comply with

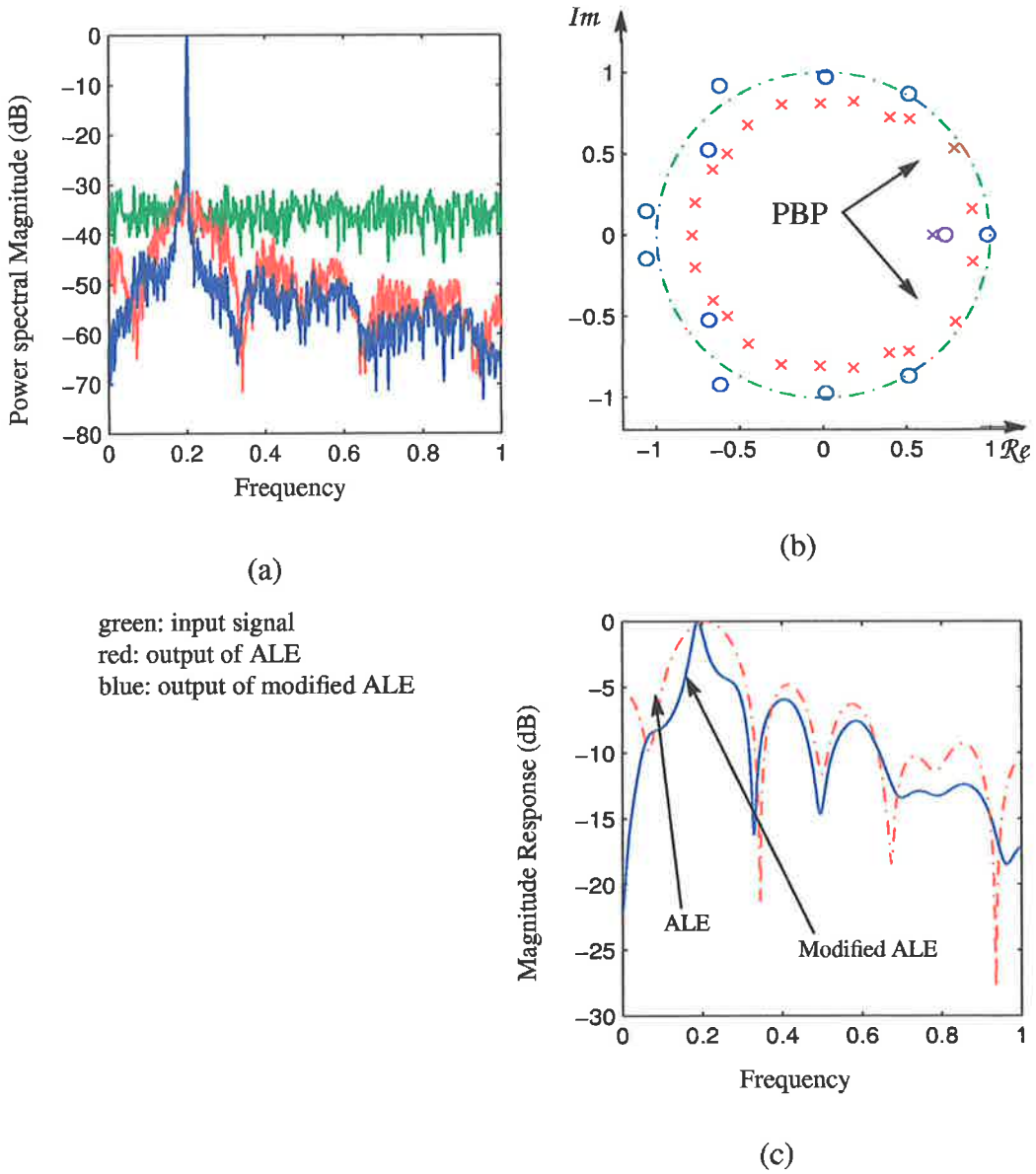


Figure 3.10 a) Magnitude spectra of the input signal, output of the ALE filter, and output of the modified ALE, b) pole-zero diagram of the modified ALE after convergence, c) magnitude response plots of the ALE and the modified ALE filters.

the aims of the thesis; rather our intention is to improve the signal-to-noise ratio of the heart sound signals by processing them with an ALE filter.

3.9 Derivation of Modified ALE Algorithm

The derivation of the modified ALE algorithm follows from a modification to the recursive least square algorithm [36]. The complete derivation is presented in appendix B; here we only outline the algorithm. Note that all the matrices and vectors mentioned in this section are defined in appendix B.

Eq 3.16 can be used to calculate the optimum values of the *h*-parameters of the filter using the LS algorithm. Assume that a new observation of the input signal is obtained. We would like to update the previously calculated values of the filter parameters instead of finding them from scratch. In appendix B we explain how to achieve this goal.

In order to start the updating process of the filter parameters, the first step is initialization of the impulse response vector of the filter and the autocorrelation matrix. The impulse response vector can be either set to zero or it can be initialized to optimum values calculated with the first set of observed data for the filter taps. The end results will be the same for both of the initialization methods. Since first method was easier to perform therefore we usually used the first method. After setting $\mathbf{h}(0)$, it is necessary to initialize the autocorrelation matrix, $\mathbf{R}(k)$. This is done by letting $\mathbf{R}(0) = \delta \cdot \mathbf{I}$, where δ is a small constant and $\mathbf{I}$ is the identity matrix, in which case

$$\mathbf{R}^{-1}(0) = \delta^{-1} \cdot \mathbf{I} \qquad \text{Eq 3.35}$$

In every step we calculate the gain parameters $\mathbf{G}(k)$. Then this vector is used for updating the vector $\mathbf{h}(k)$ and matrix $\mathbf{R}^{-1}(k)$.

3.10 Processing of phonocardiograms using the ALE and Modified ALE filters

As was mentioned earlier, the phonocardiograms were recorded in a relatively noisy environment, at the patient's bedside. Thus, the recorded signals may have been affected by the events taking place in the patient's room, such as the 50 *Hz* hum, noise

made by other people in the room. Although we took every precaution to make the recording environment as quiet as possible prior to the recordings, such as turning off all of the fluorescent lights, placing the patient's bed away from any medical equipment, asking the patients to hold their breath during recording if possible, and not performing any recording if the recovery room was too crowded with patients.

Another source of interference in the heart sounds are the lung sounds. Although we asked patients to hold their breath during the recording process, as far as they could, but sometimes this was not practical due to the patient's physical condition. Therefore, some of the recorded phonocardiograms have been contaminated with breath sounds.

ALE and modified ALE schemes were applied to the heart sounds and some typical results are presented here. In Figure 3.11 we have shown 20 seconds of a typical heart sound. We performed ALE filtering and modified ALE filtering on this signal. For comparison purposes the length of the filters and the decorrelation delays were taken equal for both filters. Based on visual inspection of the power spectra of the outputs of the filters, the following parameter values were chosen for the best results: $N = 14$, $\Delta = 9$, $D = 5$, $\alpha = 0.4$, and $\beta = 1$.

Unlike the signal of Figure 3.10 which is a sinusoidal signal where a large delay parameter will not cause decorrelation of the signal component, the delay parameter for the heart sound signal must not be very large. This will avoid decorrelation of signal components of the heart sounds.

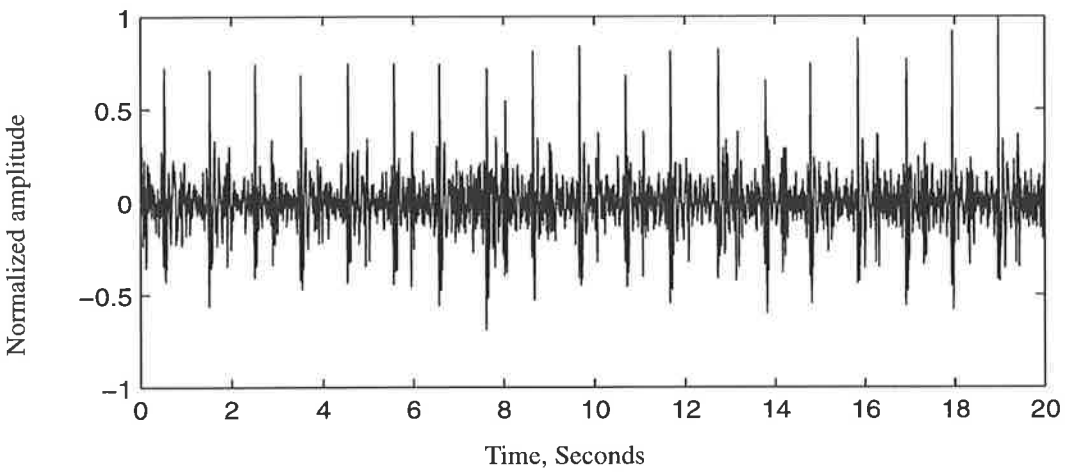


Figure 3.11 A typical heart sound used for illustration of application of modified ALE.

In Figure 3.12, we show the results obtained by ALE and modified ALE filtering of the signal of Figure 3.11. In Figure 3.12 (a), the power spectra of the signal and its filtered outputs are shown. As the figure shows, signal power has increased in the range below 400 *Hz* region after filtering with just the ALE filter. While there is not much change in the signal power beyond 400 *Hz* for this particular signal, with the modified ALE filer, not only has the signal power in the range below 400 *Hz* increased, but it has decreased in the region above 400 *Hz*. This means that the relative signal power in the low frequency region has increased. This is due to the fact that in the modified ALE filter, poles are generated in the pass-band region of the filter which seems to be in the

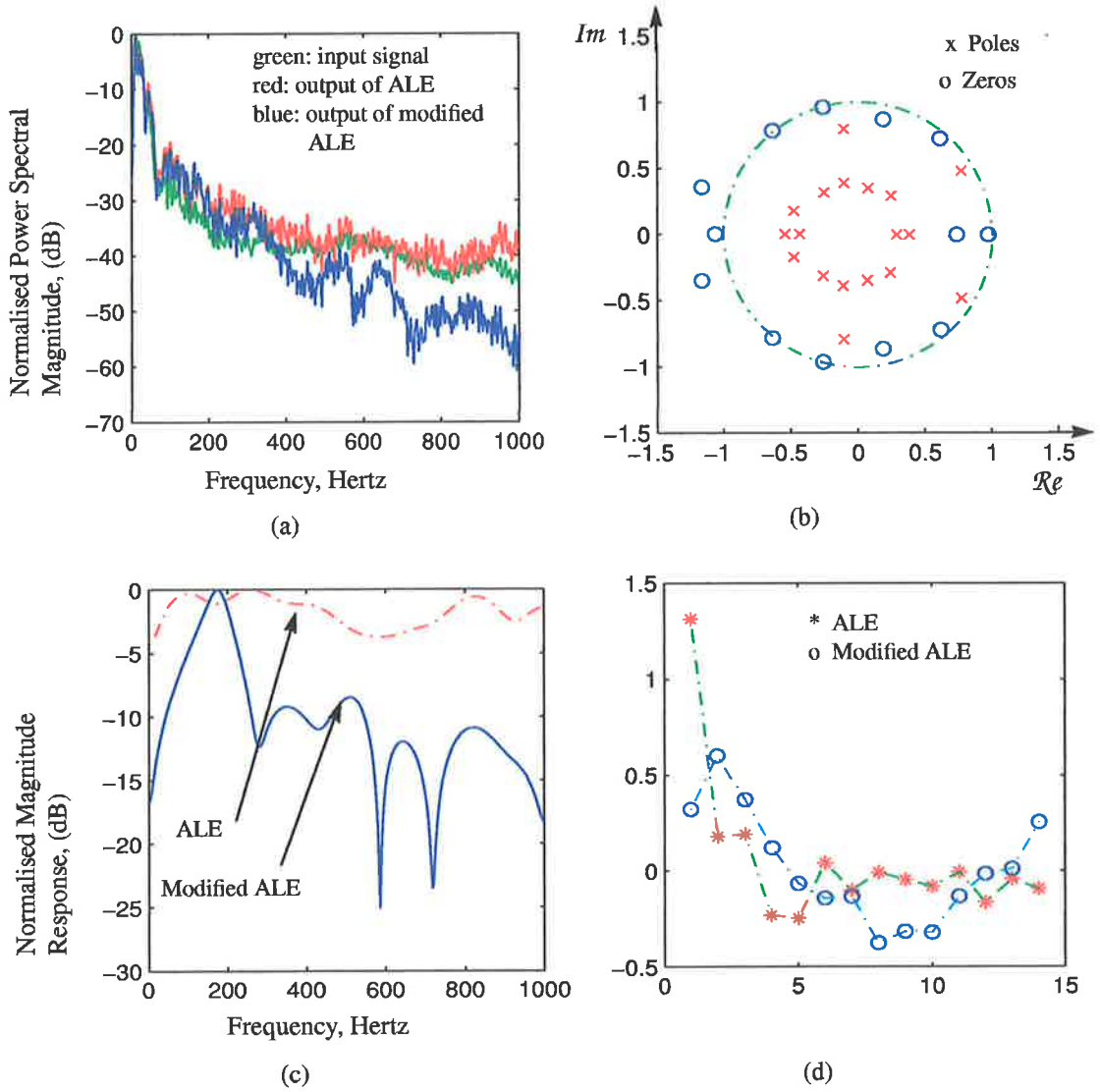


Figure 3.12 a) Magnitude spectra of the heart sound of Figure 3.11 and outputs of ALE and modified ALE, b) Pole-zero diagram of the modified ALE, c) Magnitude of the frequency responses of the ALE and modified ALE, d) h-parameters of the ALE and modified ALE (Sampling frequency=2 *kHz*).

region below 400 *Hz*, as is shown by the pair of poles closest to the unit circle in Figure 3.12 (b). Heart sounds are complicated signals, actually we can not specify a particular bandwidth for them. However, this analysis shows that most of the energy of the signal lies in a low frequency region below 300 *Hz* (Note that the Nyquist frequency is 1000 *Hz*). In Figure 3.12 (c) we show the frequency responses of the ALE and modified ALE filters. The dashed-line curve shows the magnitude response of the ALE filter. It shows that the ALE filter behaves more like an all pass filter for the heart sounds. The solid-line curve in Figure 3.12 (c) shows the magnitude response of the modified ALE filter. It shows that the modified ALE filter attenuates high-frequencies and passes the low frequencies and also has less attenuation for mid-frequencies than for high-frequencies. In Figure 3.12 (d) are shown the impulse response parameters of both of the filters.

3.11 Conclusions

In this chapter we have examined ways of improving the signal-to-noise ratio of the phonocardiograms. Adaptive filters are good tools for noise cancellation and improving the signal-to-noise ratio. They are very effective whenever a suitable reference input is available. However, in some cases such reference noise signal is not available, therefore the reference signal is obtained by delaying the process. The principle advantage of this method is its adaptive capability which allows processing of the input whose noise properties are unknown.

We have proposed a modification to ALE filters based on alteration of its pole-zero diagram. We introduce poles into the filter by adding the delayed version of the output of the filter to filter taps. This generates poles in the pass-band region of the filter and improves its performance. This was illustrated with an example in which the input signal consists of a sinusoidal signal plus an additive white noise.

Finally we have shown the performance of the filter by applying it to the heart sounds. Filter parameters were determined experimentally, that is we have observed the output of the filter with different parameters and have chosen those which gave the best results.

Chapter 4

Construction of Heart-Beat Signal from Phonocardiograms

4.1 Introduction

As mentioned in the earlier chapters, the objective of this study is to analyse heart sounds using time-frequency techniques. In Chapter 2, it was explained that long records of heart sounds were acquired from real patients before and after angioplasty operation. These signals are periodic in a sense that each beat cycle resembles its previous and following ones. This is true unless the physical conditions of the patient change. In other words, the statistical characteristics of the heart-beat cycles do not change from beat to beat. Thus, heart sounds are considered as quasi-stationary signals. This phenomenon makes it possible to treat each heart-beat as a separate data vector and construct a heart-beat signal from a recorded heart sound using an ensemble average of individual heart-beat cycles.

The purpose of this chapter is to devise a method which will enable us to construct a heart-beat signal using our recorded data set. We introduce an algorithm that decomposes a long recorded heart sound to its heart-beat cycles using simultaneously recorded electrocardiograms. The algorithm is described in two parts: Cardiac heart-beat cycle detection using ECG signals and PCG heart-beat cycle construction. The

result is a heart-beat signal which is used by the time-frequency analysis methods of the subsequent chapters. In most biological signals, the acquired data contains some artifacts, which must be identified and removed from the signal before performing any signal processing on them. In section 4.2, we will briefly explain the kind of artifacts that we may encounter in our data set. Then in section 4.3, an overview of the proposed method for construction of a heart-beat signal is presented. Description of the individual steps are given in detail for the cardiac heart-beat cycle detection. In section 4.4, details of the algorithm for construction of the PCG heart-beat cycle are described where the ECG cardiac heart-beat cycle obtained in the previous section are used. We use real data, the recorded heart sounds, to illustrate the performance of our method. In order to show the effectiveness of the scheme, we have used the same heart sound for all illustrations.

4.2 Artifacts

Phonocardiograms and electrocardiograms, like any other electrical signals, are prone to artifacts. There are three major sources of artifacts in those signals: electrical, mechanical, and physiological.

The electrical interferences are primarily caused by 50 Hz power line. Most of the time, they are easily eliminated by using a proper notch filter, or they could be eliminated by adaptive filters with a reference 50 Hz signal. The disturbing effect of the 50 Hz hum can also be minimized by isolating the recording amplifiers. This was the major source of electrical interference that we were mostly concerned with when performing our recordings. For elimination of the 50 Hz power line interferences, we used a 50 Hz notch filter, and we also placed all the circuitry inside a metal case and grounded the body of the case.

Another source of electrical noise in biological signals is radio-frequency interferences. These kinds of interferences can cause problems in the vicinity of broadcasting stations or mobile telephones. Their frequencies are usually in a far higher range than PCGs and ECGs and their effects are eliminated by using anti-aliasing filters. We used anti-aliasing filters, therefore we did not have this problem, even if the noise was present.

Mechanical noise is caused by displacement or sudden movements of the sensory devices. Displacement of the piezoelectric microphone and the ECG electrodes over the patient's skin could introduce interferences to the PCG and ECG signals, respectively. With the care taken in the recording process, we have reduced the presence of mechanically induced interferences; however, on very rare occasions we observed these kinds of noises in our signals. Most of them originated from the movements that patients made during the recording.

Physiological interferences are induced from other signals in the body. Respiration induced baseline wander is one of them. It can be minimised for the ECG signals by proper placement of the electrodes, that is the electrodes must be placed in the upper part of the chest close to the arms. Whenever it was feasible, we asked the patients to hold their breath to reduce the breath sound interferences, only in few cases it was not possible because of the physical conditions of the patients, i.e. the patient was very old with weak physical condition. Another kind of interference are muscle tremors, which are the major source of artifacts in ECG signals. These can be reduced by proper placement of the electrodes [123].

4.3 Heart-Beat Separation Algorithm

The central idea of our heart-beat separation scheme is decomposition of the phonocardiograms through detection of the heart-beats of the synchronised electrocardiograms. Beat cycle detection is easier to accomplish using the ECG signals. A brief description of an ECG wave is given in Chapter one. The QRS complex is a distinguishable component in any ECG signal. The R-waves have a relatively high amplitudes with respect to other parts in the ECG wave, therefore peak detection is applicable. We use the QRS complex as a guideline to locate the start of the cardiac cycle.

We define an ECG heart-beat cycle to be a signal starting from one P-wave and ending at the next P-wave. Therefore, the first step in our technique is automatic detection of the QRS complex. Detection of the QRS complex has been the subject of extensive research. The primary intention in most of the QRS detectors is to minimize false detection rate. Different approaches have been taken in the QRS detectors such as

amplitude measurement, amplitude squaring, peak detection, derivative, duration measurement, and others. These methods and/or combinations of them are used along with some decision making algorithm for reliable detection of the QRS complex.

The general outline of our proposed heart-beat separator is shown in Table 4.1. The first step involves detection of the QRS complex and identification of any false detection. We use a peak detector to identify the temporal positions of peaks of the R-waves. Any false detection is identified by checking the time interval between two consecutive peaks. For detection of the P-waves, first the immediate valley point before the R-wave is identified. Then the interval between two consecutive R-waves is divided into three parts. The process will be explained in details in section 4.3.3. The third section contains both the P-wave and the immediate valley point before the R-wave. Peak detection is used, in the same manner as for the R-waves detection, to find the P-waves in the interval between the beginning of the third section and the immediate valley point before the R-wave.

Next, we align the PCG and ECG signals, and use the beginning of the ECG cardiac cycles to decompose the PCG signal to its hear-beat cycles. In order to detect those heart-beat cycles which contain artifacts, a template heart-beat cycle is constructed. Then the correlation coefficient of every heart-beat cycle with respect to the template signal is calculated. A criterion is set and those heart-beat cycles that do not meet this criterion are rejected; the rest are accepted as artifact free heart-beat cycles.

Table 4.1 Outline of heartbeat cycle detection algorithm.

1) Detect the QRS peaks.

2) Identify the beginning of the cardiac cycles in the ECG signal.

3) Separate the heart-beats of the PCG signal using the cardiac cycles of the ECG signal.

4) Construct a template heart-beat signal.

5) Using the correlation coefficients, identify those beat cycles that are corrupted with artifacts.

6) Use frequency domain and calculate the PCG heart-beat signal.

The final heart-beat signal is constructed using the frequency domain mean of the artifact free beat cycles. In the following sections we describe this algorithm in greater details.

4.3.1 The QRS Detection

The ECG signal is oscillatory in nature, although it may not be exactly periodic in a strict mathematical sense. Observation of the evolution of these signals reveals that there are similar events or periods, but they may not be exactly reproducible.

One of the important tasks in ECG analysis is QRS detection; a great deal of research effort has been devoted to it [45]. Most of the QRS detectors described in the literature aimed for analysis of the ECG signal itself, but our purpose is to find the temporal location of the peaks of the QRS complexes and use them to separate the heart-beat cycles from the PCG signal. We divide our QRS detection system into two entities: the preprocessor and the decision making system, as shown in Figure 4.1. An algorithm is developed for the decision rule section of Figure 4.1 and is shown in Figure 4.2.

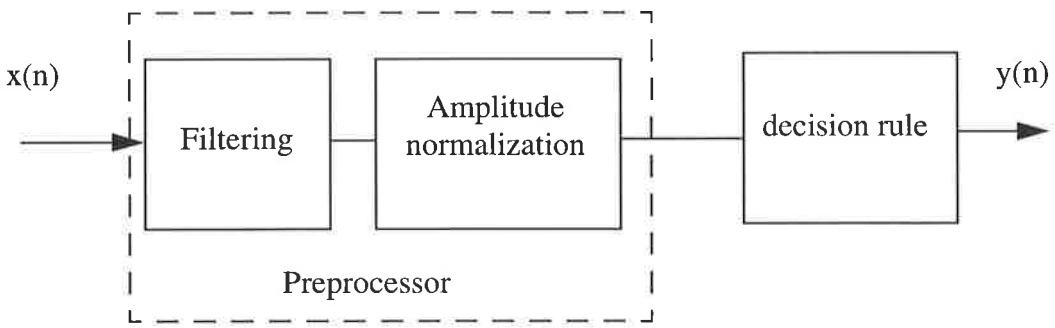


Figure 4.1 Block diagram of the QRS detector.

We observed that most of the frequency content of the ECG signal lies below $40Hz$, and since we are not interested in the spectral information of the signal, a low-pass filter with cutoff frequency of $45Hz$ is used in the filtering section of the preprocessor. As far as our application is concerned, this will not cause any distortion in the

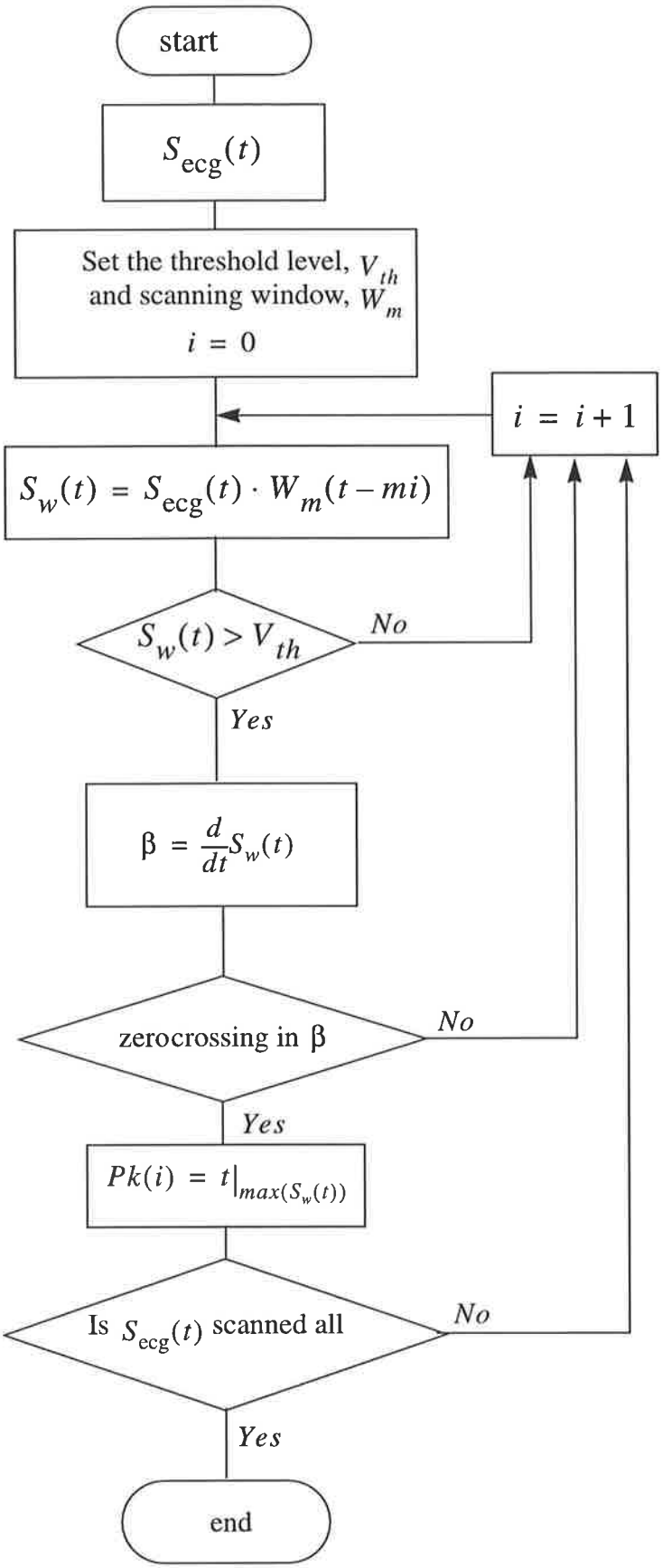


Figure 4.2 QRS detector algorithm.

signal. The high frequency noise and $50Hz$ power line interferences will be effectively removed with application of such lowpass filtering. The output of the filter is normalized in amplitude by simply dividing the signal by its maximum positive amplitude. The purpose of amplitude normalization will be explained shortly.

Our main concern in the QRS detector, is to avoid detection of false peaks. Visual inspection of our recorded ECG data revealed that about 85% of the recordings are artifact free. In Figure 4.3 we have shown 15 seconds of different ECG recordings. Types of disturbances that we have encountered are shown in (b), (c), (e), and (f) where

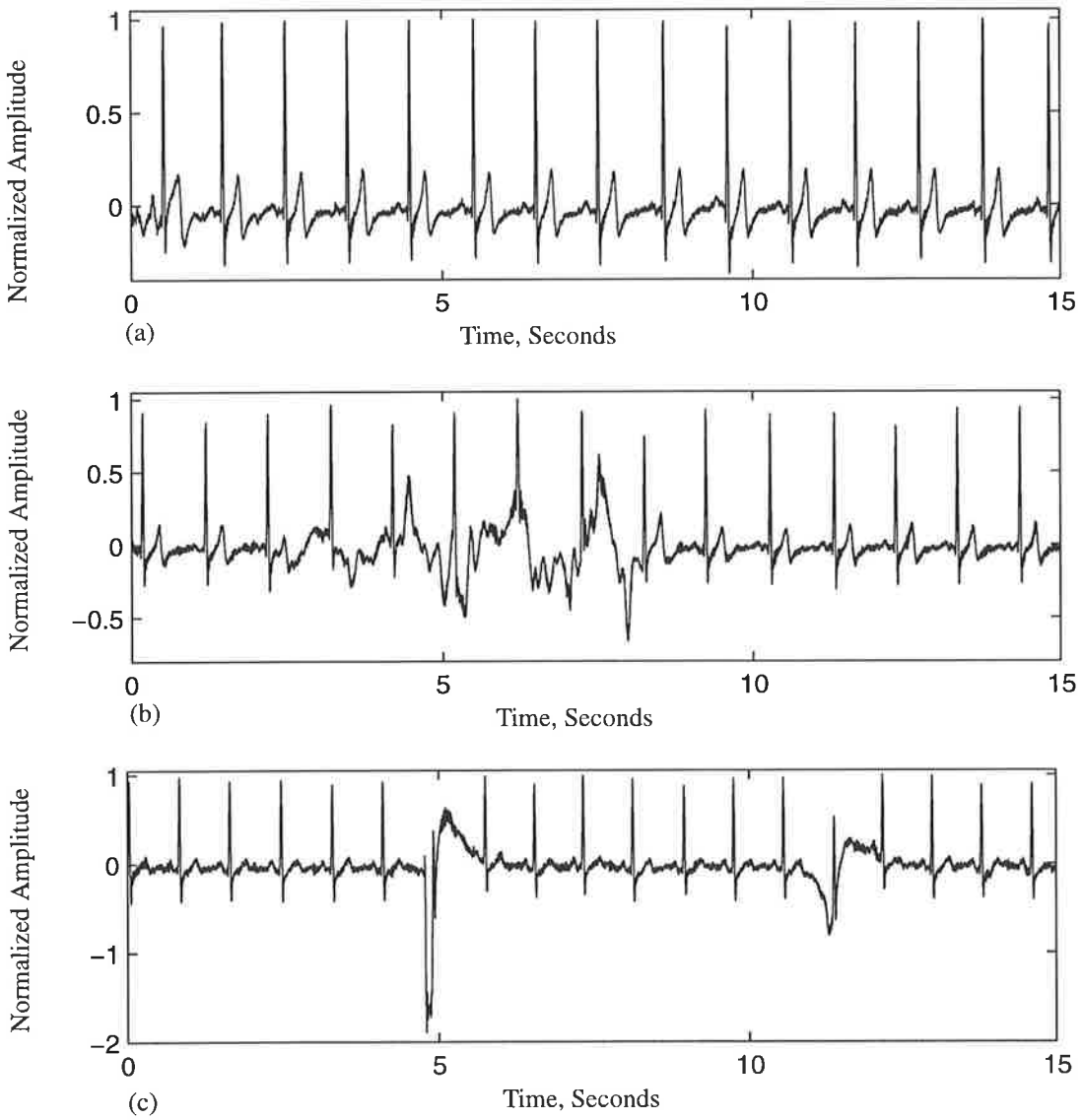


Figure 4.3 Typical ECG signals from data set a) artifact free, b) and c) contain artifacts.

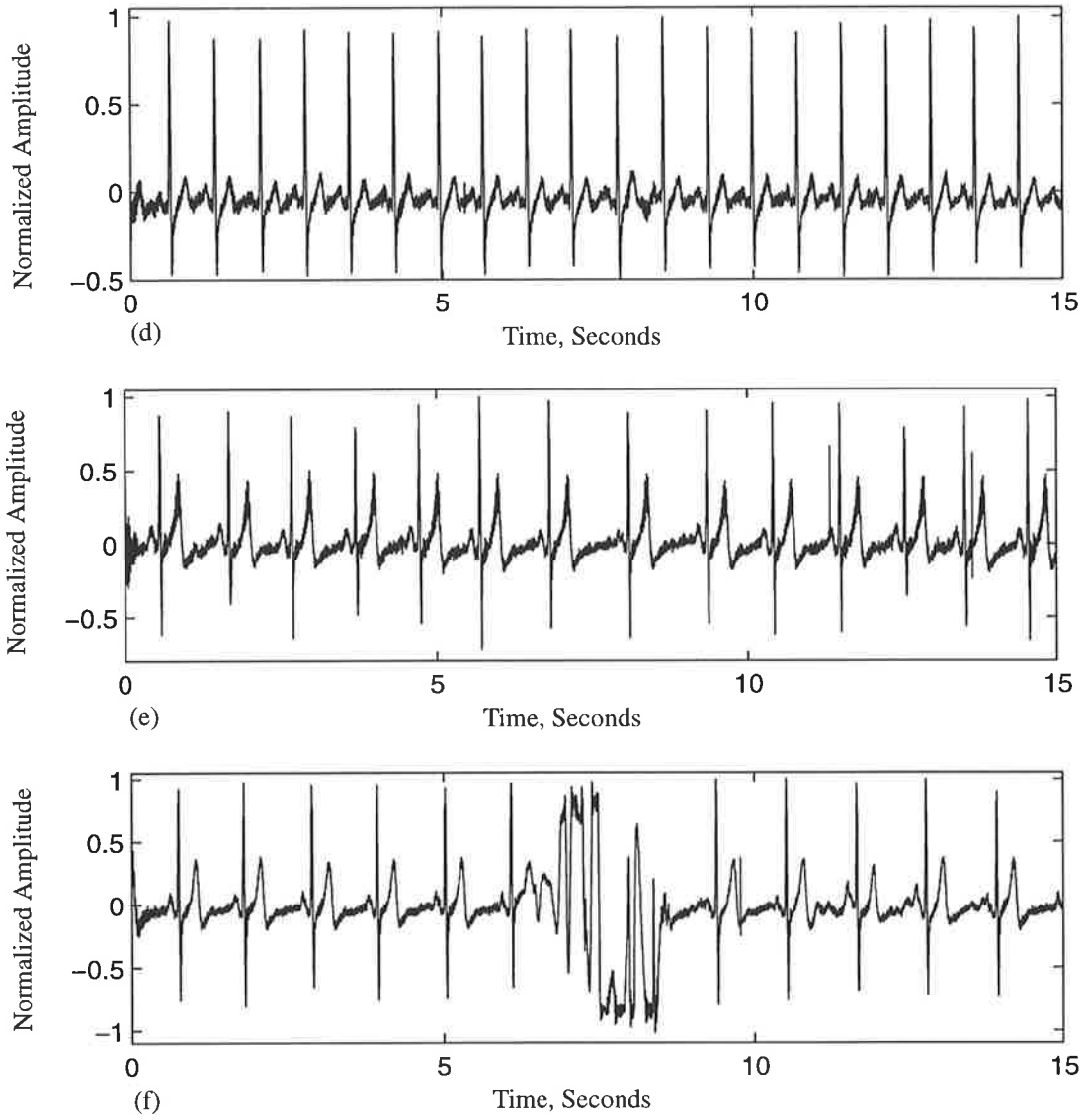


Figure 4.3 Cont d) artifact free, e) and f) contain artifacts.

signals of (a) and (d) do not contain any artifacts. The artifacts shown in Figure 4.3 (b), (c), (e), and (f) are the worst cases that we have pinpointed in our data set. Most likely those artifacts have mechanical origins, that is they are caused by improper attachment of the ECG electrodes to the patient's skin due to the movements made by the patient.

As stated above, the input signal was initially normalized in amplitude. This is essential for threshold settings. This results in a bounded threshold level, that is the threshold level can only be set between zero and one as

$$0 \leq V_{th} \leq 1 \qquad \text{Eq 4.1}$$

There are usually some amplitude fluctuations in the QRS amplitudes. If we examine the individual amplitudes of the signal of Figure 4.3 (a), as shown in Figure 4.4, it is clear that there is only 4.3% fluctuation in peaks of the QRS complexes of this signal. The fluctuation in amplitudes of 99% of our recordings falls below 5%.

We used a thresholding scheme to detect the peaks of the QRS complexes. A threshold level equal to 0.6 was selected for most signals. This is well below any amplitude variations and high enough to be above the amplitude levels of the T-waves. The amplitudes of the T-waves were mostly less than one third of the amplitudes of the QRS complexes.

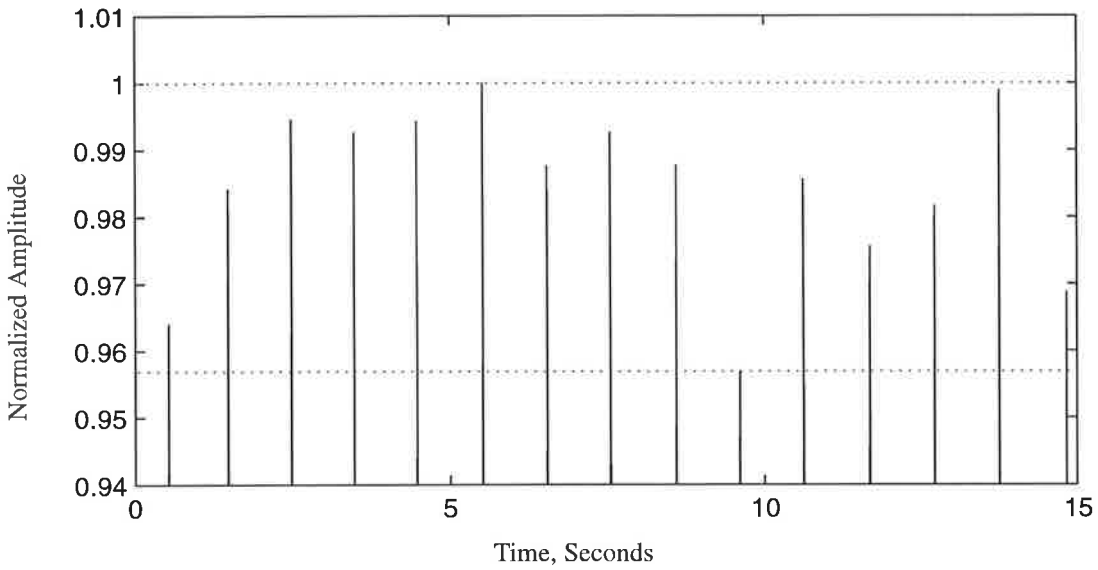


Figure 4.4 Normalized QRS peaks of the signal of Figure 4.3 (a), to show amplitude fluctuations.

As it is implied in the algorithm of the Figure 4.2, the input data is scanned with a rectangular window. The width of the window is much less than a heart-beat period, whose length could vary from about half a second to 1.2 seconds for different people. A window with length, m, equal to 10 milliseconds was used. A short window length was selected to reduce the possibility of missing a QRS complex, which may occur if there are noise spikes with height above the threshold level in the vicinity of the QRS complex. This is illustrated pictorially in Figure 4.5.

Within each window those parts of the data that are above the threshold level are selected and their maximum value is found. The thresholding ensures that only regions

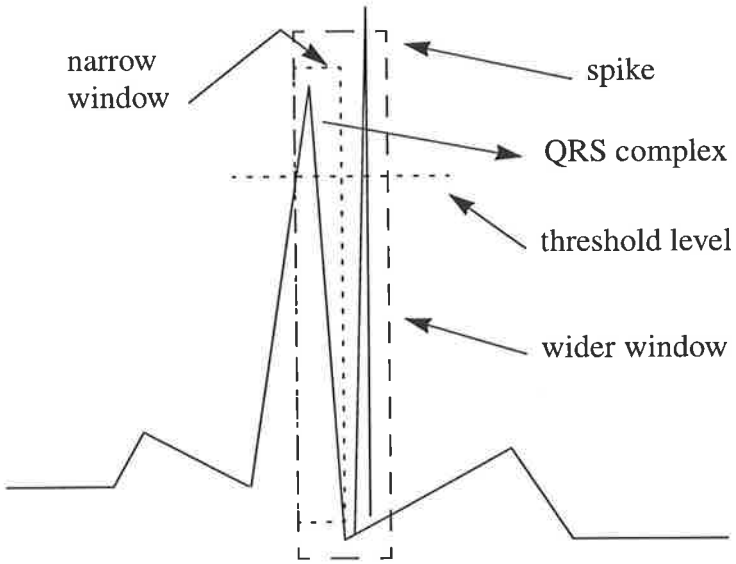


Figure 4.5 Pictorial explanation for scanning widow size selection.

of the signal that are in the vicinity of the peak of the QRS complex will be considered for peak detection. The signal shown in Figure 4.3 (a) is lowpass filtered, and its amplitude is normalized. A single beat of this signal is shown in Figure 4.6 (a). The windowing scheme is applied to it and parts of the signal above the threshold level are detected, they are indicated as being between τ_1 and τ_2 in Figure 4.6 (b).

The portion of the signal that is above the threshold level is denoted as $S_{pk}(t)$, defined by Eq 4.2, and plotted in Figure 4.6 (b) for further illustrations.

$$S_{\text{pk}}(t) = S_{\text{ecg}}(t) \qquad \text{for} \quad \tau_1 \le t \le \tau_2 \qquad \text{Eq 4.2}$$

Figure 4.6 (b) also shows one possible arrangement of three consecutive 10 milliseconds window positions. Each window may only contain the ascending portion of the signal, it may contain the actual peak of the QRS complex, or it may cover only the descending portion of the signal, as shown in Figure 4.6 (b). Any other combination is also conceivable. In order to distinguish and locate the position of the actual QRS peak, we calculate the slope of the signal within each windowed part of the signal. Let $W_m(t-mi)$ represent the i$^{\text{th}}$ window, i.e. for $i = 1, 2, 3$ as in Figure 4.6 (b).

$$\beta_i = \frac{d}{dt}[S_{ecg}(t)W_m(t-mi)] \qquad \text{Eq 4.3}$$

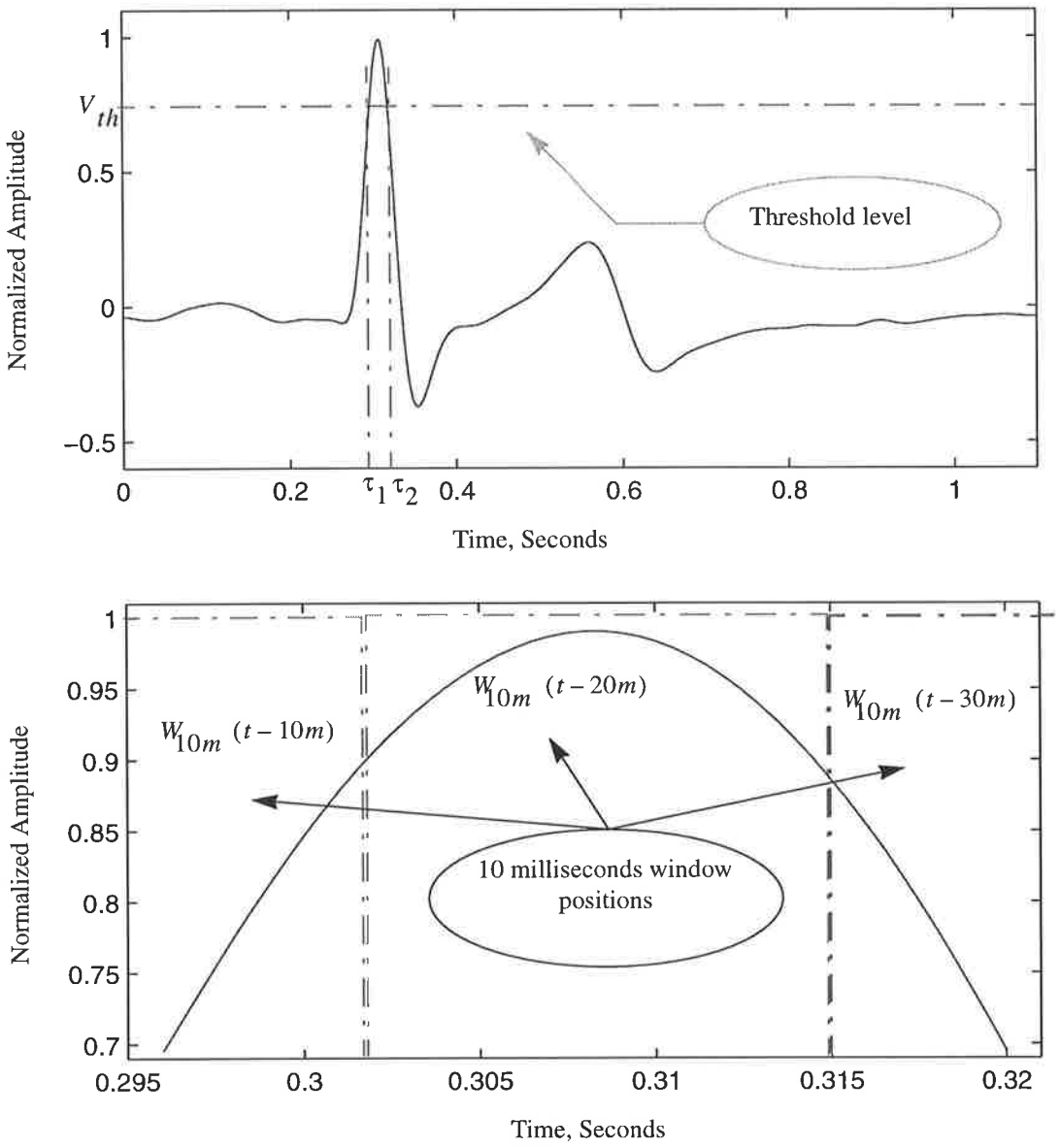


Figure 4.6 a) Single beat cycle of signal of Figure 4.3 (a), b) the part above the threshold level.

Zero crossing in any of the slope curves represent a maximum in the ECG signal in the corresponding region. The slope curve of Figure 4.6 (b) is plotted in Figure 4.7. It is clear that $W_{10m}\ (t-10m)$ and $W_{10m}\ (t-30m)$ do not contain any peaks and the signal within these windows are increasing and decreasing, respectively. In $W_{10m}\ (t-20m)$ there is a zero crossing at $t = 0.3007$ seconds which indicates that a maximum point within the current window has been detected. The scanning process is continued until all of the signal is scanned, then execution of the scanning algorithm is terminated.

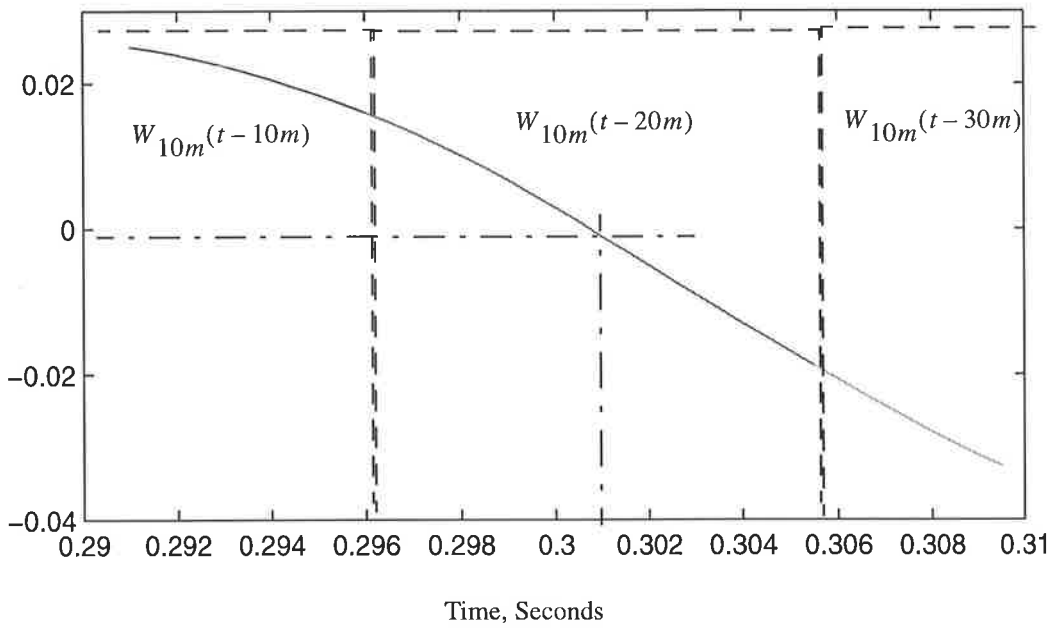


Figure 4.7 Slope of the signal of Figure 4.6b.

4.3.2 True QRS Interval Detection

The primary objective of this part is to ensure that the QRS peaks detected with the algorithm of Figure 4.2 are real QRS peaks. That is, we want to identify any false detections and remove them. There are two possibilities for detection of a wrong peak. First, there might be an artifact whose height is higher than the threshold level. The second possibility is the presence of small fluctuations (due to noise) near to the QRS peak.

Based on the outcome of the algorithm of Figure 4.2, the average of the QRS peak-to-peak intervals is calculated as

$$T = \frac{1}{M}\sum_{j=1}^{M-1} PK(j+1) - PK(j) \qquad \text{Eq 4.4}$$

where M is the number of the detected QRS peaks, and PK is a vector that holds the locations of the QRS peaks.

We define the minimum and the maximum acceptable time interval between two consecutive QRS peaks as follows:

$$\Delta T_{mn} = T - 0.1T \quad \text{Eq 4.5}$$

$$\Delta T_{mx} = T + 0.1T \quad \text{Eq 4.6}$$

The mean distance of each i^{th} QRS peak with respect to its adjacent neighbours is calculated as

$$\delta t_i = \frac{\Delta t_{i-1} + \Delta t_{i+1}}{2} \quad \text{Eq 4.7}$$

where Δt_{i-1} represents the distance of the i^{th} QRS peak from its preceding, and Δt_{i+1} represents the distance of the i^{th} QRS peak from its subsequent QRS peak.

The δt_i for every QRS peak is compared with ΔT_{mx} and ΔT_{mn}. If it is within the range, it is accepted as a true QRS peak, otherwise it is considered to be a false peak. That is for a true QRS peak, Eq 4.8 must be satisfied.

$$\Delta T_{mn} \leq \delta t_i \leq \Delta T_{mx} \quad \text{Eq 4.8}$$

If a false QRS peak occurs, then there will be three or less consecutive δt_i that will not satisfy Eq 4.8. That is, there will be a local valley centred on *i,* in the values of δt_i. In order to find the location of the false QRS peak, it is necessary to find the position of the bottom of the valley, therefore we calculate the derivative of the local δt_i and then find the position at which it becomes zero. In other words, we find the slope of the local valley and detect the position where it becomes zero.

The above QRS interval detection algorithm is shown in Figure 4.8. The performance of this algorithm is illustrated in Figure 4.9 using the signal of Figure 4.3 (b). As is shown in Figure 4.9, application of the QRS detector algorithm has resulted in the detection of 16 QRS peaks, where the threshold level was set equal to 0.6. As the figure shows, the 9^{th} QRS peak is a false detection. The mean of the peak-to-peak intervals for this signal is calculated as $T = 0.9443$ seconds. Therefore a QRS interval from 0.8499 to 1.0387 seconds will be accepted as the true QRS interval. In Table 4.2, the time intervals between the consecutive peaks of Figure 4.9 are also shown. Comparing these values with the values of ΔT_{mn} and ΔT_{mx} reveals that there is a local

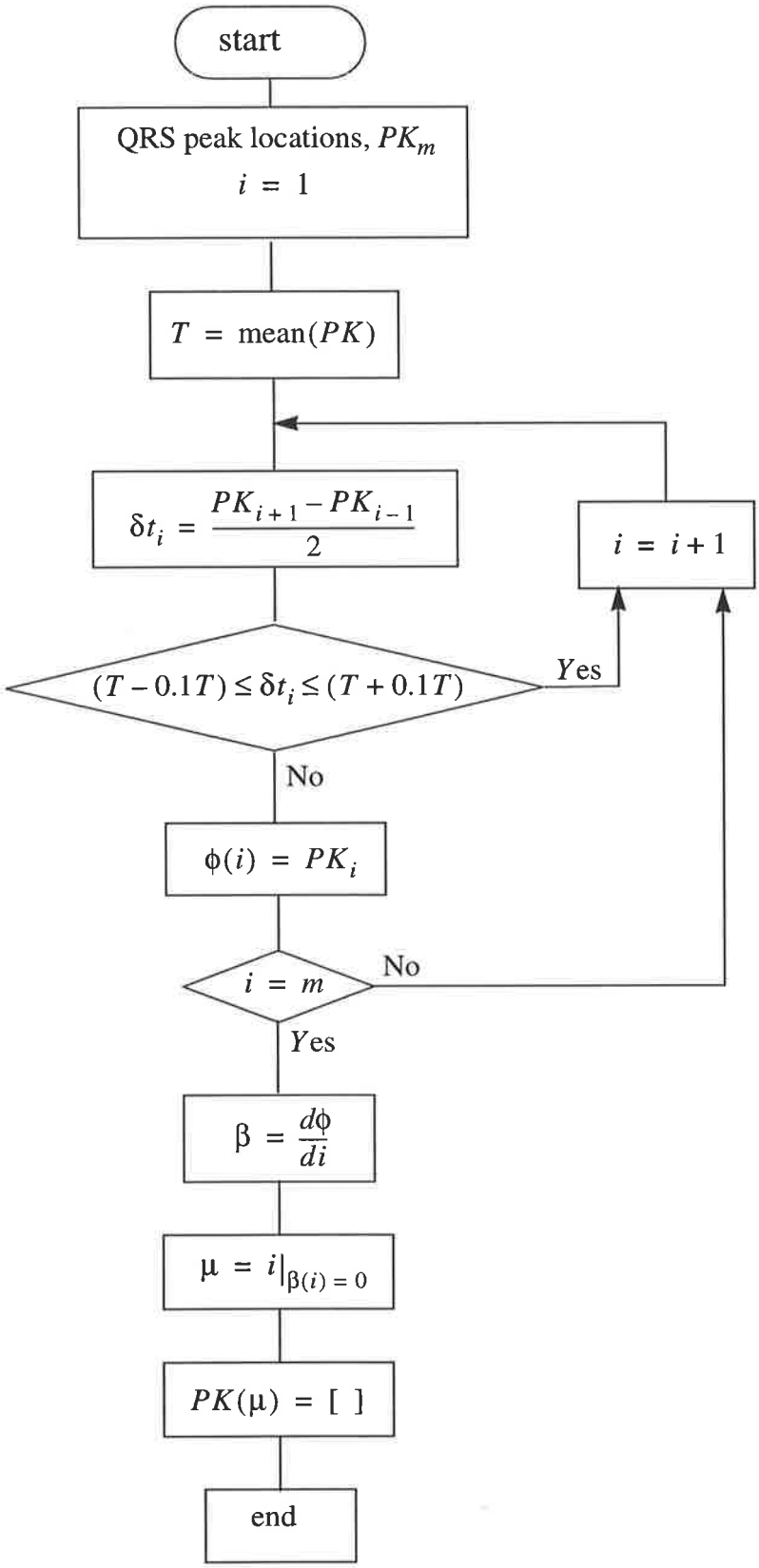


Figure 4.8 The QRS true intervals detection algorithm.

valley below the threshold level of ΔT_{mn} centred around the 9^{th} entry. This implies that the 9^{th} QRS peak is not a true QRS peak, and hence the 9^{th} peak is deleted.

Table 4.2 The time intervals between the consecutive peaks of Figure 4.9.

1.0165	1.0052	1.0028	1.0000	0.9938	1.0152	1.0390	0.6610	0.4990 *****	0.8518
1.0053	1.0375	1.0245	1.0018	1.0098					

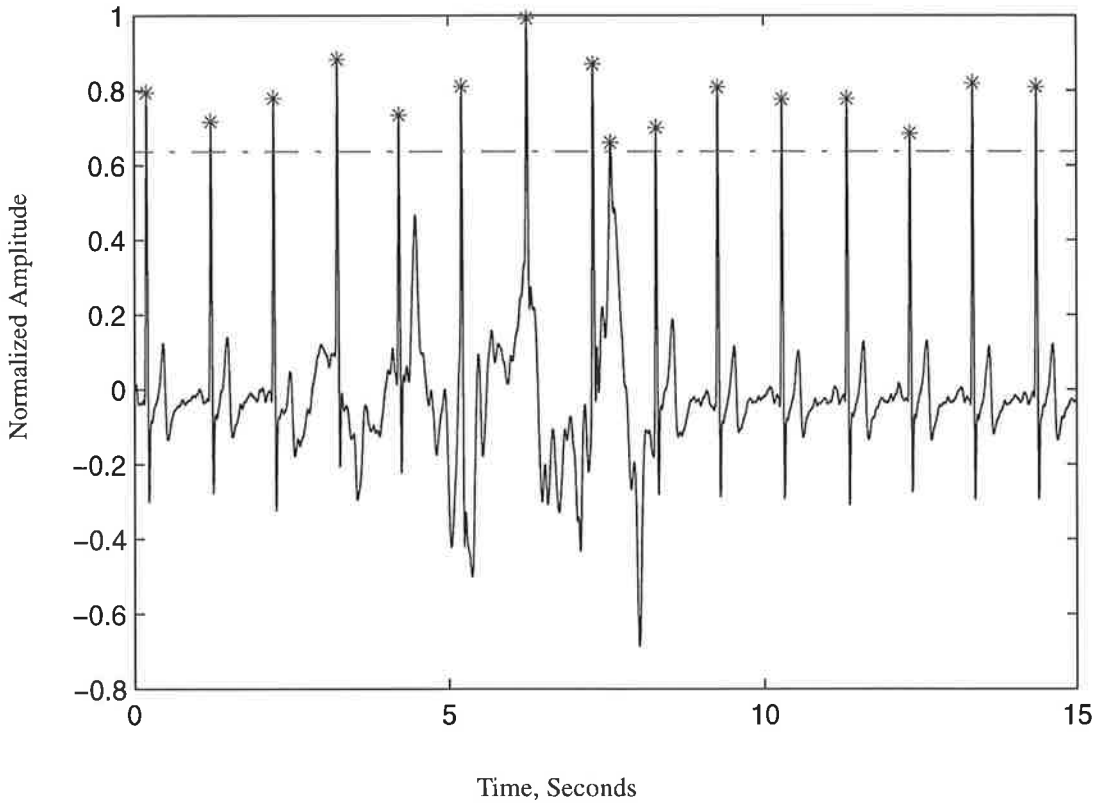


Figure 4.9 QRS peaks detector applied to signal of Figure 4.3 (b).

4.3.3 ECG Beat Detection

So far we have been able to successfully detect the QRS peaks in the ECG signal. In this section we identify the beginning of the cardiac cycle with respect to the QRS peaks. We set the P-wave as the beginning of the cardiac cycle. The P-waves

occur before the QRS complexes. If we divide the time intervals between two consecutive QRS peaks to three sections, the P-wave occurs in the third section. This is illustrated in Figure 4.10 for the ECG signal of Figure 4.3 (a).

The P-waves have relatively low amplitudes and therefore are very susceptible to noise. Our method of detecting the P-waves is similar to the QRS detector explained previously. We use an amplitude detector to locate the peaks of the P-waves, and then use its period to identify false detections. The peak detector will not accomplish the task in this case unless regions with amplitudes higher than the amplitudes of the P-waves are removed prior to its application. In Figure 4.11 an algorithm is shown for detection of the P-waves and identification of the cardiac cycle.

Initially, the time interval between two consecutive peaks is divided into three sections, Figure 4.10. The onset time for the third section of the i^{th} QRS peak is calculated as follows:

$$tr_i = PK(i) - (PK(i) - PK(i-1))/3 \qquad \text{Eq 4.9}$$

where its offset time is the i^{th} entry in the vector PK. Onset and offset times are shown in Figure 4.10.

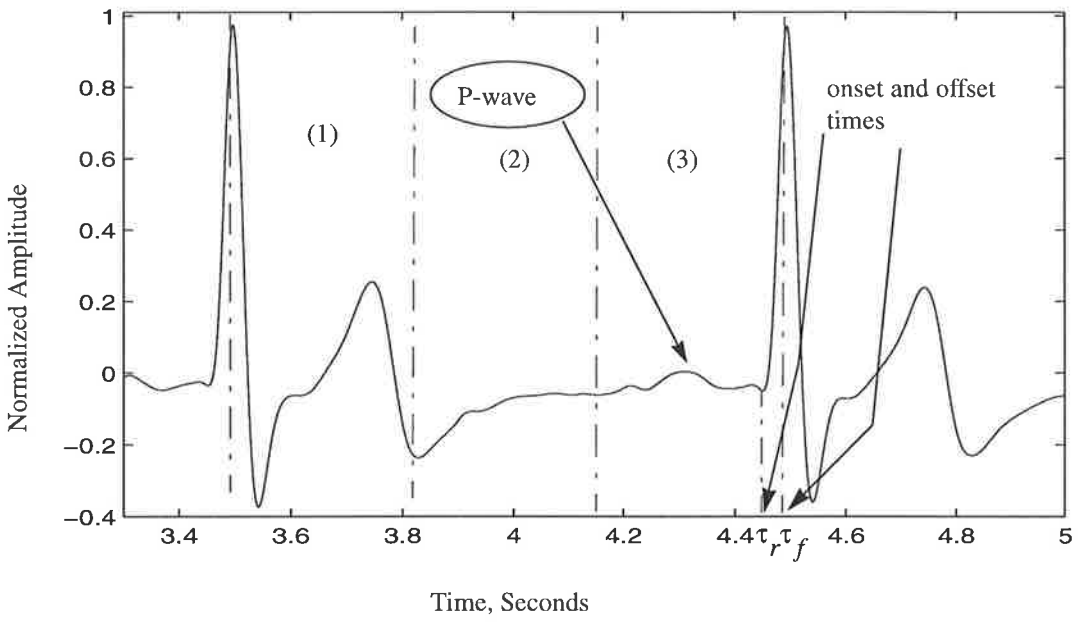


Figure 4.10 Two consecutive cycles of the ECG signal of Figure 4.3 (a).

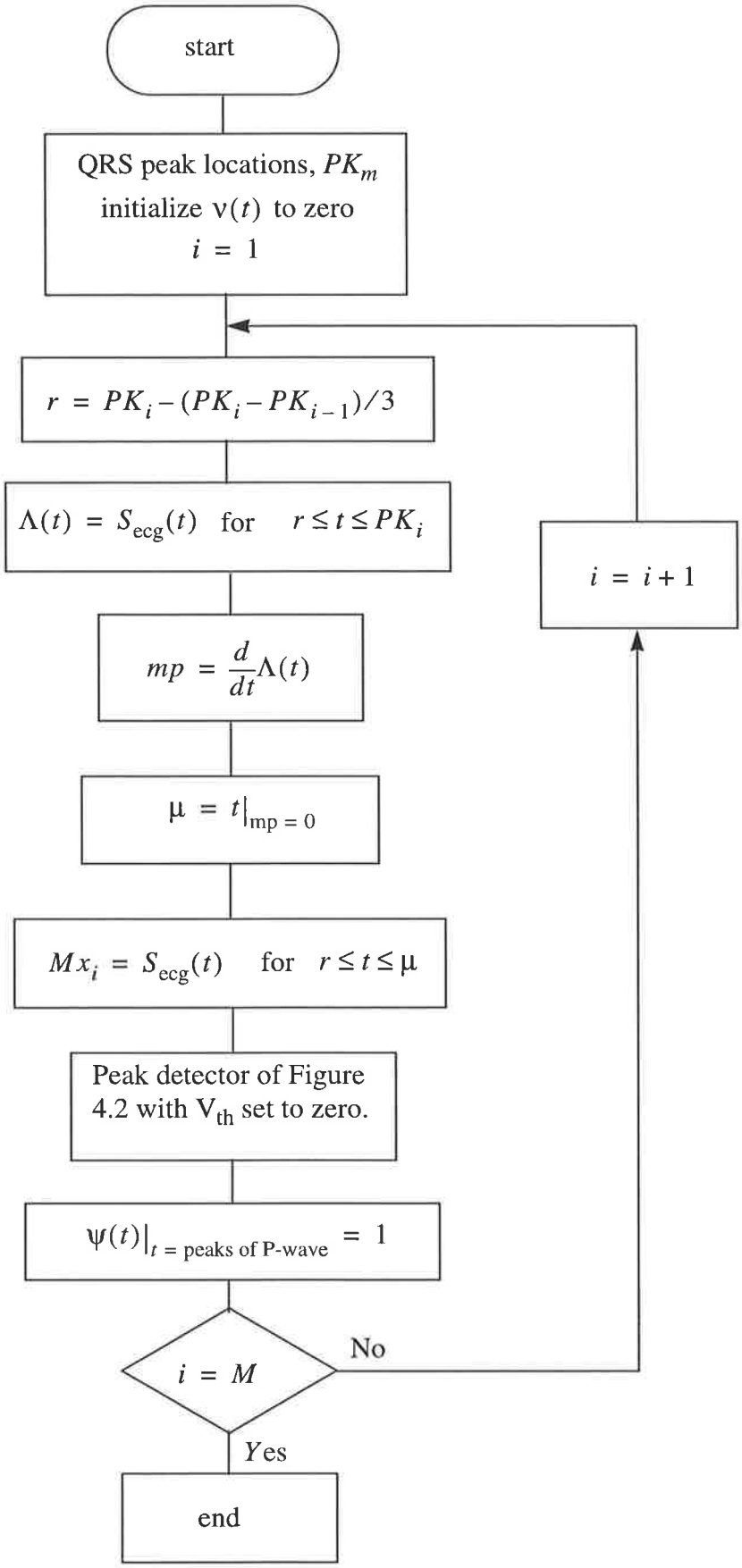


Figure 4.11 ECG beat cycle detection algorithm.

Since we are going to use the peak detector to find the peaks of the P-waves, it is essential to remove the portion of the QRS complex that falls within the section three. This is the part of the signal in between τ_r and τ_f, in Figure 4.10. This is accomplished by calculating the slope of the signal from tr_i to PK_i as

$$\mathrm{mp}_i = \frac{d}{dt} S_{\mathrm{ecg}}(t) \qquad \text{for} \qquad tr_i \le t \le PK(i) \qquad \text{Eq 4.10}$$

The signal mp_i is scanned in a bottom-up fashion; that is, scanning starts from the end and proceeds towards the beginning. When the first zero crossing is encountered, the scanning is stopped. This point corresponds to the first minimum in the signal just before the QRS complex. This is illustrated in Figure 4.12. The part of the signal to the right of zero crossing is discarded. For the remaining part of the signal, the algorithms of Figure 4.1 and Figure 4.8 are applied. This process is repeated for all of the QRS intervals. The outcome of this algorithm is a vector the same size as the ECG signal $S_{ecg}(t)$. The elements of this vector are all zeros except in those locations that the peaks of P-waves occur, where a value of one is assigned. This vector is designated as $\psi(t)$ in Figure 4.11.

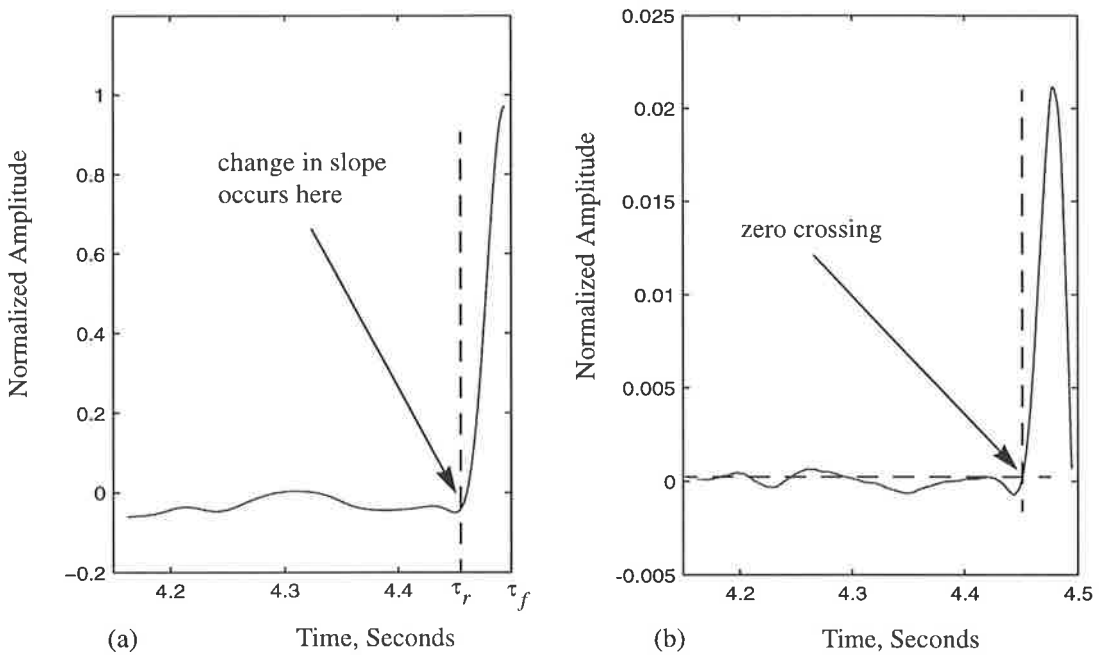


Figure 4.12 Section 3 in Figure 4.10 and its slope, (a) signal, (b) slope.

4.4 PCG Segmentation

The prime purpose for developing the QRS detection algorithm was to devise a scheme which would enable us to decompose the PCG signal to its beat cycles with minimum misdetection. As was mentioned before, the ECG and PCG signals are recorded simultaneously; therefore, the decomposition process of the PCG signal to its beat cycles can be accomplished by simply aligning it with the results obtained by the algorithm of Figure 4.11, $\psi(t)$, which is a train of pulses located at the beginning of each cardiac cycle.

Decomposition of the PCG signal is accomplished using $\psi(t)$. Each decomposed segment of the PCG signal corresponds to one period of $\psi(t)$. In Figure 4.13 (a) and (b), 30 seconds of the PCG signal and its synchronously recorded ECG signal are shown.

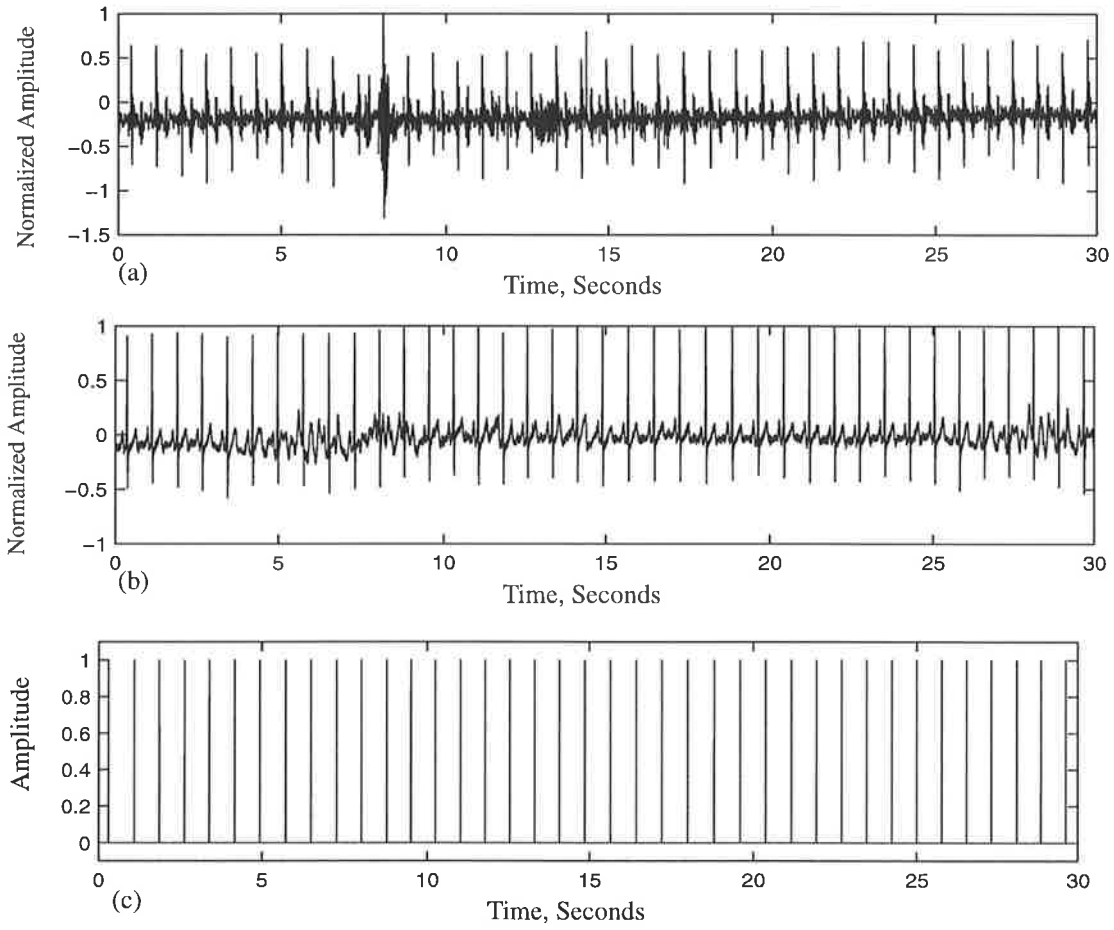


Figure 4.13 a) PCG signal, b) corresponding ECG signal, c) output of the QRS detection algorithm.

The cardiac cycle detection algorithm is applied to the signal of Figure 4.13 (a) and the resultant signal, $\psi(t)$, is shown in Figure 4.13 (c).

4.4.1 PCG Beat separation

As the PCG beats are separated, all the beat cycles are made the same length as the longest cycle in the signal. The shorter duration beat cycles were padded with the mean of the last three samples in the beat cycle. Padding with the mean of the last three samples is justifiable because: first, by doing so we avoid any sudden change in the signal amplitude, e.g. by padding with zeros; second, we are adding the last moment characteristics of the signal to its end which most likely would have occurred if the beat cycle was longer.

For the separated heart-beat cycles, it is necessary to identify those cycles that contain artifacts. Let the i^{th} PCG heart-beat be represented by $S_{\text{pcg}}^{i}(t)$. Every recorded PCG signal contains at least 30 cycles of the heart-beats, that is i equals at least 30. Visual inspections of some of the signals of our data set, revealed that most of the artifacts only occupy a small fraction of time, that is only a few beat cycles are corrupted.

A correlation scheme is devised to distinguish the artifact free heart-beat cycles within the PCG signal and is shown in Figure 4.14. The principle behind it is to compare, on a sample-by-sample basis, every beat cycle with a template signal. In the rest of the chapter, this algorithm is explained in details with illustration of its operation using the heart sound signal and the electrocardiogram shown in Figure 4.13.

The template signal is constructed by taking the ensemble average of all the beat cycles as:

$$S_{\text{tmp}}(t) = \frac{1}{m} \sum_{i=1}^{m} S_{\text{pcg}}^{i}(t) \qquad \text{Eq 4.11}$$

We examined all of the recorded signals and found out that the number of the artifact free cycles were at least ten times of the corrupted cycles; hence, the signal calculated by Eq 4.11 can serve as a template for clustering the heart-beat cycles. In Figure 4.15 (a) all of the detected beat cycles of the signal of Figure 4.13 (b) are shown. The presence of artifacts in some cycles are clearly shown in the figure. As Figure 4.15 (a) shows, the majority of the heart-beat cycles coincide while few of them show differences.

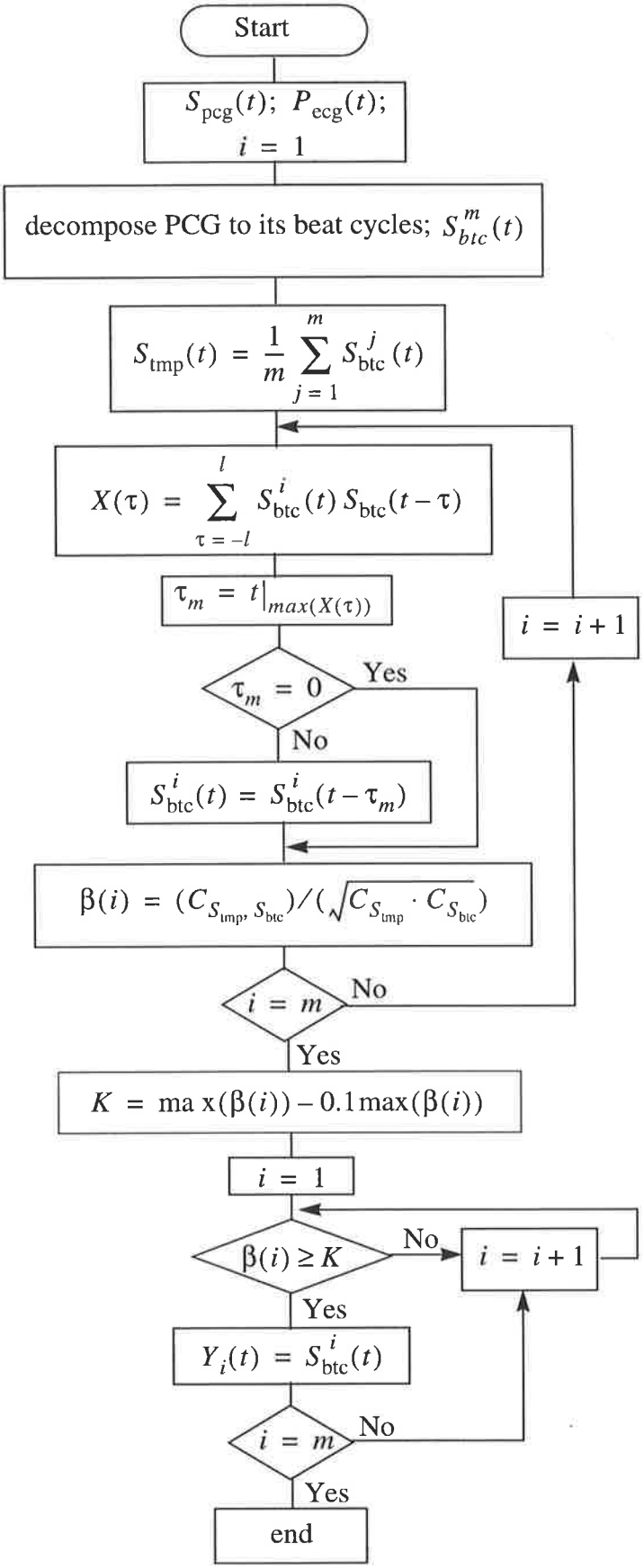


Figure 4.14 PCG beat cycle detection algorithm.

We denote beat cycles as $S_{btc}(t)$. In Figure 4.15 (b) the template signal constructed from signals of Figure 4.15 (a) is shown.

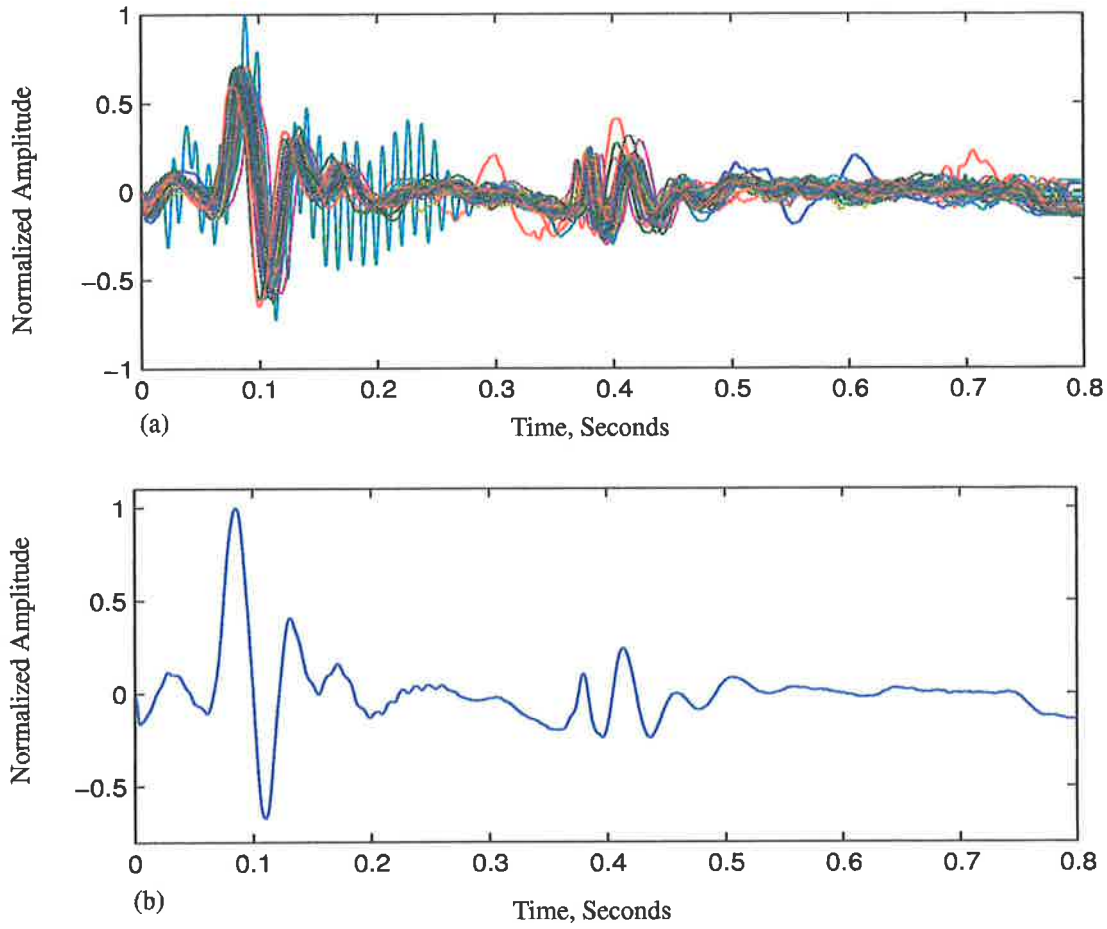


Figure 4.15 a) Beat cycles of signal of Figure 4.13(b), b) template signal constructed from signals of part (a) of this figure.

4.4.2 The Matching Process

Matching of the beat cycle waveforms with the template signal is measured by calculating the correlation coefficient between the template signal and each of the beat cycles. The correlation coefficient is defined as

$$\beta = (C_{S_{tmp}, S_{btc}})/(\sqrt{C_{S_{tmp}} \cdot C_{S_{btc}}}) \qquad \text{Eq 4.12}$$

where $C_{S_{tmp}, S_{btc}}$ represents the cross-covariance between the template signal and each of the beat cycles. $C_{S_{tmp}}$ and $C_{S_{btc}}$ are the covariances of the template signal and each of the beat cycles respectively. Close examination of the Figure 4.15 (a) indicates that

the maximum peak of the template signal may not exactly coincide with the maximum peak of some of the beat cycles. In order to maximize the correlation coefficients, it essential to align the maximum peaks of the signals involved in every correlation calculation. This is accomplished by performing a shifting and multiplying process as in Eq 4.13. The maximum point of $X(\tau)$ is found. The temporal position of this point indicates the amount of shift in time that is necessary for maximum alignment of two signals. The corresponding value of τ is used to shift the signal $S_{\text{btc}}(t)$.

$$X(\tau) = \sum_{\tau=-l}^{l} S_{\text{tmp}}(t) S_{\text{btc}}(t-\tau) \qquad \text{Eq 4.13}$$

where l is equal to the length the template signal (both signals have the same length). If the maximum occurs at $\tau = 0$, that means the two signals are aligned completely and there is no need to shift $S_{\text{btc}}(t)$, on the other hand if $\tau \neq 0$, it implies that $S_{\text{btc}}(t)$ must be shifted to $t = \tau_m$ at which the maximum of $X(\tau)$ occurs. This is illustrated in Figure 4.16. The particular beat cycle shown in Figure 4.16 (a) does not coincide with template signal which is clear by examining the peaks of the first heart sound of the signals. The small plot on the corner of the Figure 4.16 (b) shows that the peak of the signal does not occur exactly at τ=0, it is offset by 5.5 milliseconds; therefore it is necessary to shift the beat cycle by about 5.5 milliseconds. This is an acceptable phenomenon because the template signal is an average valued signal. If the detected beat cycles have had slight time shift between their maximum value point, then there would be a slight time shift between the beat cycles and template signal. The time shift could be due to beat-to-beat variations of heart sounds and/or slight error in locating the beginning of the cardiac cycles. At the final stage we use the frequency domain to calculate the heart-beat signal; therefore, this slight variation in beat cycles is not much of a concern.

The correlation coefficients are calculated and tabulated in Table 4.3, and also plotted in Figure 4.17. We set a threshold level for the correlation coefficients; that is, those beat cycles that have their correlation coefficient below 90% of the maximum correlation coefficient are rejected. In Table 4.3 the threshold level is found to be equal to 0.8702. The threshold level is also shown in Figure 4.17. As the figure shows there

are only three beat cycles that do not meet our imposed criterion for correlation coefficients. Accepted and rejected beat cycles are shown in Figure 4.18 (a) and (b) respectively.

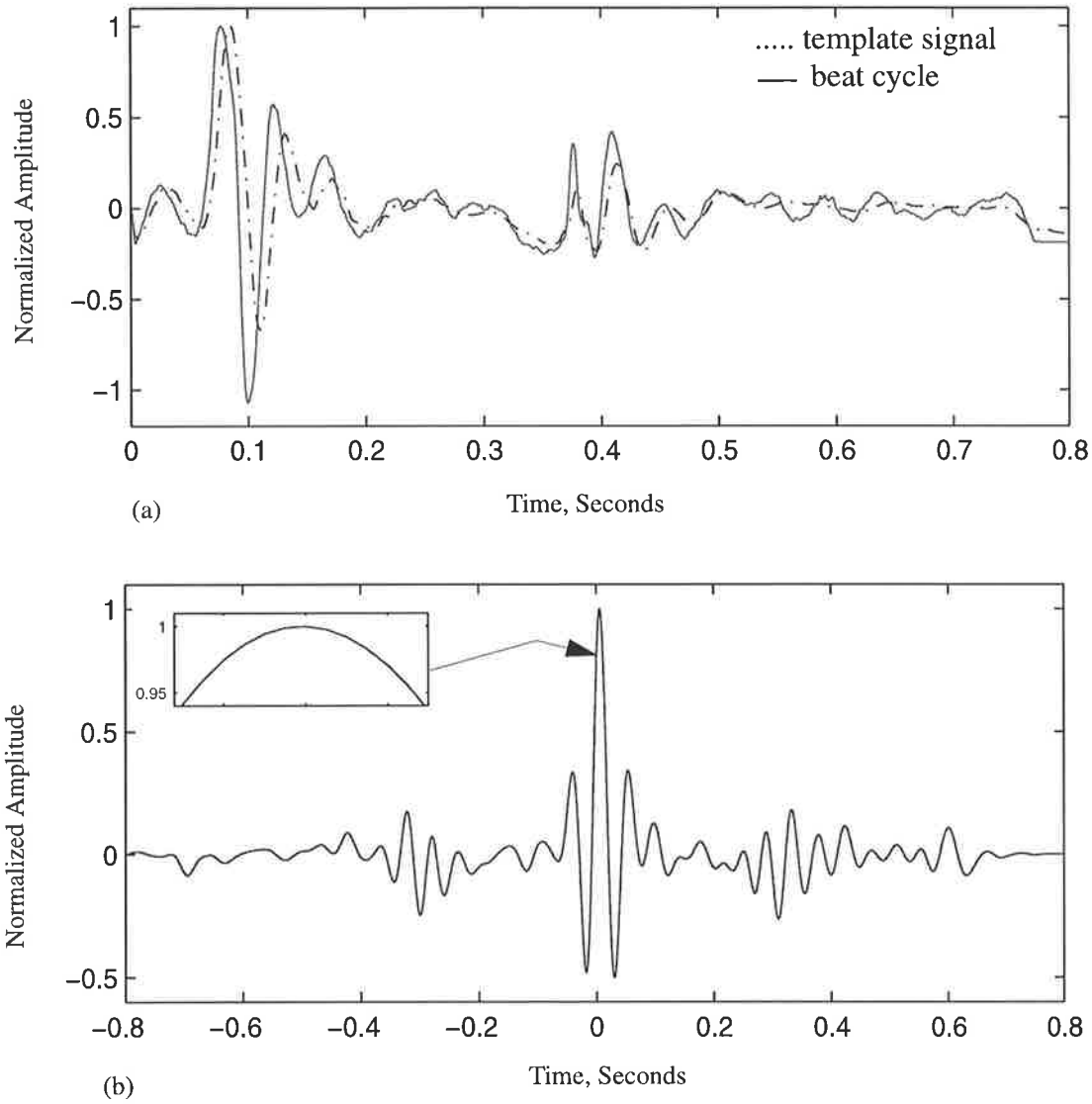


Figure 4.16 a) Template signal and a beat cycle, 2) $X(\tau)$ (see Eq 4.13) of the signals of part (a).

Table 4.3 Correlation coefficients of the signals of Figure 4.15 (a) with the signal of Figure 4.15 (b).

0.9660	0.9625	0.9395	0.9505	0.9375	0.9634	0.9482	0.9387	0.9170
0.6944	0.6568	0.9181	0.8946	0.9060	0.8681	0.9194	0.9455	0.9198
0.9322	0.9323	0.9613	0.9400	0.9172	0.9359	0.9490	0.9405	0.9400
0.9278	0.9448	0.9669	0.9645	0.9394	0.9500	0.9389	0.9658	0.9485
0.9386	0.9171							

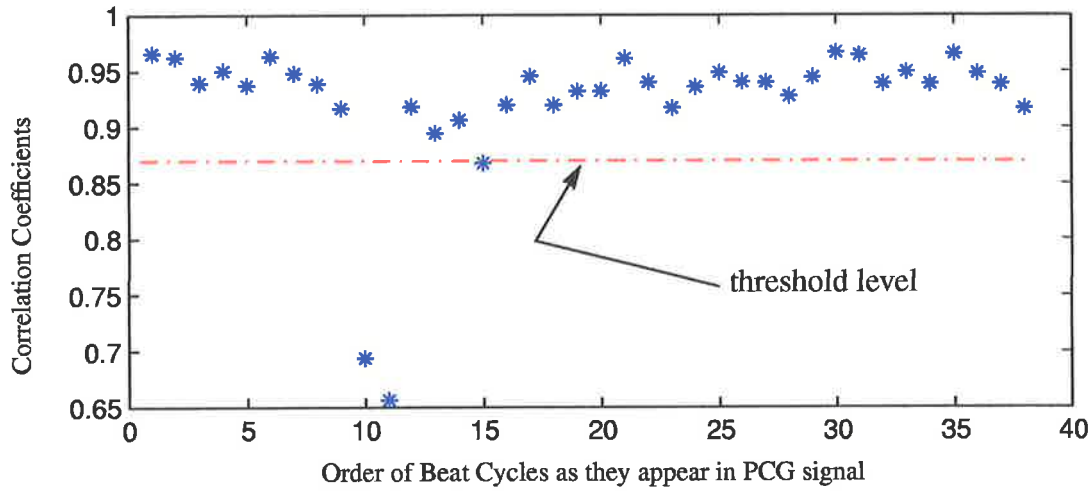


Figure 4.17 Correlation coefficients for the signals of Figure 4.15.

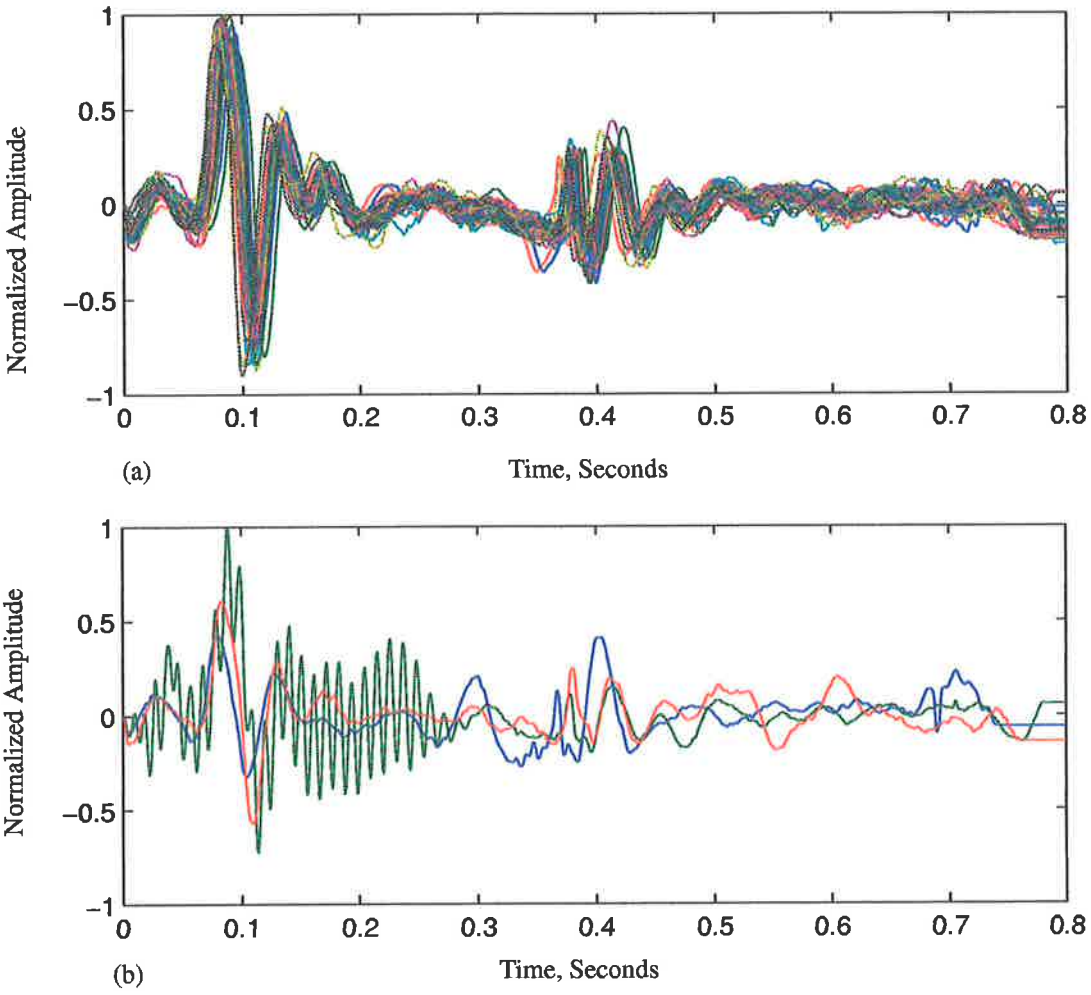


Figure 4.18 a) Accepted beat cycles, and b) rejected beat cycles for the PCG signal of Figure 4.13.

4.4.3 The Heart-beat signal

The final heart beat cycle is constructed using only the artifact free beat cycles. Note that these signals do not have the same length. Discrete Fourier transform (DFT) of the each signal is calculated. Then the mean of the DFT of beat cycles is obtained. The frequency domain is used to avoid the error that may occur in using time domain averaging because beat cycles did not have the same length in time domain. To obtain the time domain signal, the averaged frequency domain signal is converted back to time domain. The final signal is plotted in Figure 4.19.

4.5 Conclusions

We have demonstrated that the ECG signal can be used to decomposed the PCG signal into its constituent heart-beat cycles. Five algorithms have been devised to carry out the task. The beginning of the cardiac cycles are detected using the ECG signal and they are used to decompose the PCG signals. Correlation coefficients are used to identify the artifact free heart-beat cycles. Using the frequency domain, the mean of the heart-beat cycles were found as the final heart-beat signal. We have successfully tested our algorithms for all of the signals in our data set.

The ECG signal to be used must not contain very high background noise. If that is the case, which is very unlikely, then a noise removal technique must be used prior using the algorithms of this chapter.

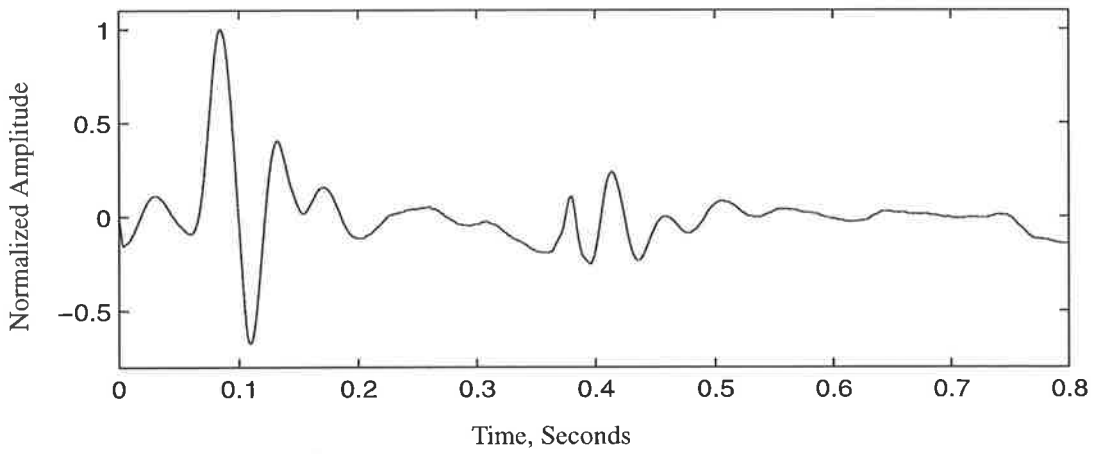


Figure 4.19 Final heartbeat signal.

Chapter 5

Local Maxima Detection

5.1 Introduction

In three-dimensional pattern recognition, edges are among the most important features of the pattern. To define an edge, consider a tiled floor. At first glance it is reasonable to decide that the contours of the floor are indeed the edges where the tiles define a pattern. We may also include the contours of each tile among the set of edges and consider the surface of each tile as a pattern texture. Discrimination of edges from textures depend upon the scale of the analysis. In image processing, points of sharp variations in intensity define the important features of the image, they are considered to be the edges in the image, and generally constitute the important image structures.

One of the bases for segmentation of an image is considered to be a discontinuity in the gray level values [46]. The principle areas of interest are detection of lines and edges in the image. In general, automatic segmentation of patterns in an image is an extremely difficult task. In processing of an image, this is the step that determines the eventual success or failure of the analysis.

In a 3-D representation of the modulus of a time-frequency matrix, locations of concentrations of energy can be viewed as local ridges. Ridges contain crucial information on the characteristics of the signal. The objective of this chapter is to develop an algorithm which can be used for detection of locations of concentration of energy in the modulus of a time-frequency plane. We call it local maxima detector. The algorithm is divided into four parts each of which is explained in a separate section. In section 5.2, the modulus of the time-frequency matrix is viewed as an image and ridges are considered to be the textures of the image. The outline of the proposed local maxima detector algorithm is presented. In section 5.3, the effect of smoothing to suppress small fluctuations on the surface of the modulus of the time-frequency matrix is explained. In section 5.4, a method is proposed for detection of edges of the locations of concentrations of energy in the time-frequency matrix of the heart sound signal. Generally, a time domain windowed sinusoidal function exhibits a cone shaped body in 3-D time-frequency representation. Although, heart sounds are very complex signals, nevertheless their locations of concentrations of energy in the time-frequency matrix have shapes very similar to cone shaped bodies. In section 5.5, two methods are proposed for detection of the boundaries of the cone shaped bodies in the time-frequency matrix. The local maxima detector is presented in complete form in section 5.6. A criterion for accepting the detected local maxima is discussed in section 5.7. In section 5.8, simulation examples are presented.

5.2 Time-Frequency Planes

In signal analysis and signal processing, characterization of signals according to regions of concentrations of energy in the time-frequency plane has been used in many applications such as radar/sonar detection, computer vision, image processing, and speech processing. These regions contain valuable information about the characteristics of the signal.

Time-frequency representations (TFR) characterize signals over time-frequency planes. In any linear time-frequency technique such as short-time Fourier transform and wavelet transforms, or bilinear methods such as Wigner distributions, a time series signal is transformed to a two dimensional time-frequency plane. The main feature of

these methods is to make use of the time-frequency localization properties of the representations. Therefore, time domain and frequency domain analyses are combined to yield a more revealing picture of the temporal localization of the spectral components of the signal.

Let $s(t) = A(t)e^{j\varphi(t)}$ be a signal as a function of time with time varying frequency. The time-frequency representation of *s(t)* reveals local properties in the regions of the time-frequency plane determined by the instantaneous frequency $\varphi'(t)$, where $\varphi'(t)$ is the time derivative of $\varphi(t)$. Examination of the modulus of the time-frequency or time-scale matrix of a time series signal reveals the regions of concentration of energy of the signal as a function of time and frequency or time and scale. In other words, the values of the surface of the time-frequency plane give an indication as to which spectral components are present at which time. Three-dimensional plots of time-frequency surfaces have been used as pictorial representations enabling a signal processor to analyse how the spectral components of the signal vary with time.

Locations of concentration of energy in the modulus of the time-frequency plane are called *local maxima* hereafter. The time-frequency plane can be considered as a generalization of the one-dimensional concepts of instantaneous energy $|s(t)|^2$ and power spectral density $|S(f)|^2$. The two-dimensional energy distributions $T(t_0, f_0)$ provides a measure of the local signal energy or information about the presence of a sinusoidal component of frequency f_0 at time t_0.

We consider the modulus of time-frequency representation of heart sounds as a three-dimensional image, and we will call it a time-frequency image (TFI). The modulus of a function is a positive and real function; therefore, the aforementioned time-frequency image is real valued. Although heart sound signals that are used for calculation of time-frequency matrices are processed for noise cancellation before being used, they may still have some noise components associated with them. In addition to that, some of the time-frequency techniques introduce unwanted components themselves, such as cross terms appearing in the Wigner distributions. Therefore, it is logical to consider TFI to be composed of two parts: signal and noise.

Patterns are defined to give quantitative or structural description of an object in an image. Image analysis is a process of discovering, identifying, and understanding patterns that are relevant to the performance of an image based task. We consider the local maxima detection as a pattern recognition task. TFI patterns are the ridges of concentration of energy in the time-frequency plane. Patterns of TFIs are partially known prior to processing operation (they are cone shaped bodies), unlike the other image processing tasks where patterns are sometimes totally unknown. This makes the processing of TFIs somehow easier. Therefore, local maxima detection is actually a process of locating patterns rather than recognising them. For clarity, we present a simple example.

Example: Let $x(t)$ be a time series signal with three sinusoidal parts located at different times, with additive white gaussian noise. Therefore, *x(t)* can be expressed as

$$x(t) = \sum_{i=1}^{3} s_i(t)w(t-\tau_i) + n(t) \qquad \text{Eq 5.1}$$

where $w(t-\tau_i)$ represents a Hamming window centred at time τ_i, $\mathcal{E}\{n(t)\} = 0$ and $\mathcal{E}\{n^2(t)\} = 0.1$. This signal is plotted in Figure 5.1. Sinusoidal parts were Hamming windowed in order to avoid step changes in the time domain which would have resulted in unwanted frequencies.

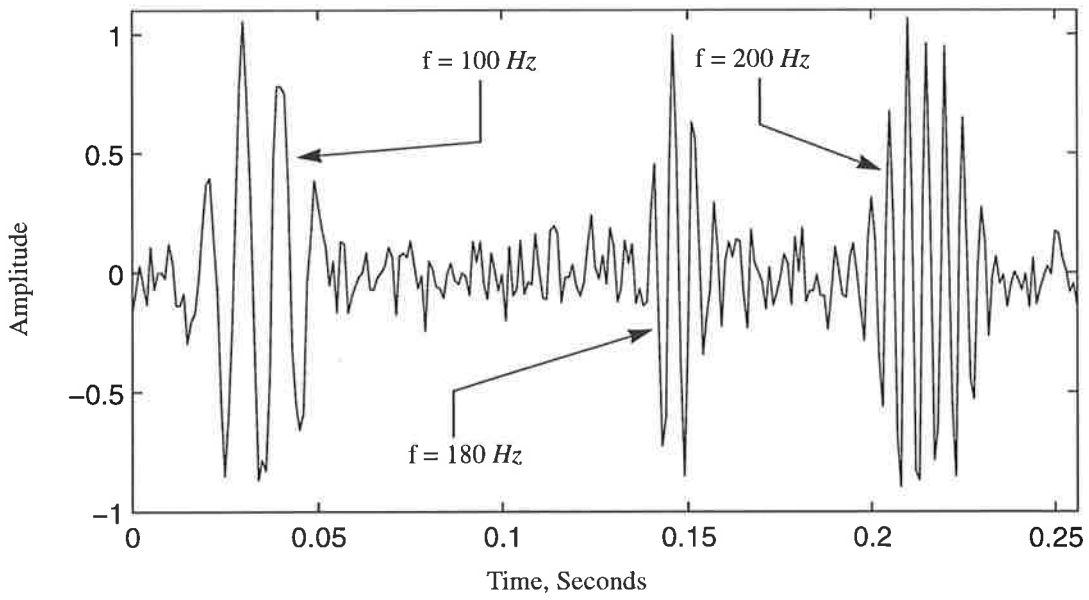


Figure 5.1 Three sinusoidal signals localized in time with white noise added.

The modulus of the Choi-Wiliams distribution, which is a time-frequency technique and is explained in details in Chapter 7, of this signal is shown in Figure 5.2. As we notice in this figure, there are almost three distinct ridges that indicate where the energies of the signal reside. These patterns are cone shaped bodies. Ideally, they should have an elliptical or circular projection on the x-y plane. There are causes that make the surface of Figure 5.2, especially areas with low energy levels, very complex surfaces. The causes are: the way the signal components are oriented in the time and frequency domains; that is, the more closer they are located to each other the more they interact and the structure of their shape changes, especially on the skirts; interferences caused by noise; and the nature of the method used for the calculation of the time frequency matrix. As we move away from the centre of the gravity of each energy concentration area, we notice that there are surface variations associated with it very far away from it.

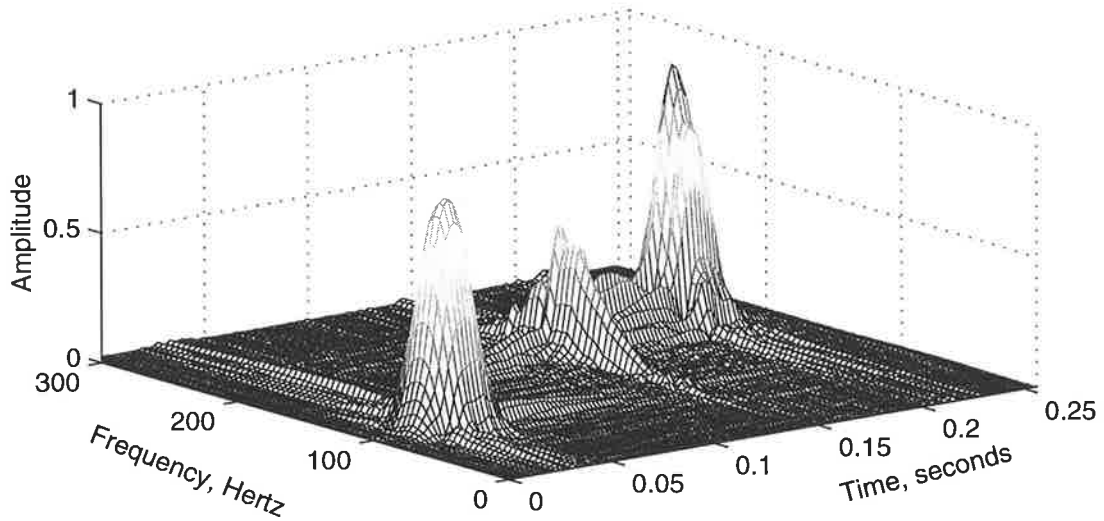


Figure 5.2 The Choi-Williams distribution of the signal shown in Figure 5.1.

Our intention is to recognize the location of these patterns in the TFIs. The strategy adopted is first to locate the centroids of the patterns and then identify their position in the time-frequency plane in order to detect all the surface areas associated with each energy concentration point. Our procedure is based on the detection and removal operation of these surface areas in the bounded domain of the time-frequency plane. Removal of the surface is necessary because our detection scheme is based on detecting the highest energy levels in the time-frequency plane. Once a detected high-

est energy level is removed, then the next highest energy level becomes dominant. Obviously the magnitudes of the ridges in the time-frequency plane for the heart sounds will not be all equal and they depend on the strength of different frequency components generated by the heart. Therefore we first detect each local maximum in TFI and then identify its contour. Next, remove the detected local maximum from the time-frequency plane. This process is repeated until all of the ridges are detected and the centroids of their locations in the time-frequency plane are identified. The outline of the local maxima detection algorithm is shown in Table 5.1.

Table 5.1 Outline of local maxima detector algorithm.

```
1. Find the row and column location of the
   maximum of the time-frequency matrix.
2. Detect its boundaries.
3. Remove it from the time-frequency matrix.
4. Repeat steps 1-3 until all local maxima are detected.
```

Two different regions of concentrations of energy in the normalized time-frequency plane of a typical heart sound are shown in Figure 5.3. Short time Fourier transform was used to obtain the time-frequency distribution. The local maximum shown in Figure 5.3 (b) has a relatively low energy and there are many surface fluctuation around it, while Figure 5.3 (a) contains a higher energy local maximum point and has a smoother surface.

In order to make the surface fluctuation in Figure 5.3 (b) become more evident, the surface of the local maximum of Figure 5.3 (b) was scanned along the time and frequency axes directions. That is, the time and frequency coordinate locations of the highest energy point of Figure 5.3 (b) was found: t=0.72 and f=180. Then at f=180 the time axis was scanned and the result is shown in Figure 5.4 (a); likewise at t=0.72 the frequency axis was scanned and the results are shown in Figure 5.4 (b). The fluctuations appearing in Figure 5.4 are due to sidelobes of the window in the short-time Fourier transform as well as the background noise in the heart sound signal. Note that a relatively low amplitude local maximum was chosen to illustrate the susceptibility of those regions to noise. This shows the necessity of applying a smoothing filter to the

time-frequency matrix, before attempting to detect the contours of the local maximum in such conditions. This is the topic of the discussion in the next section.

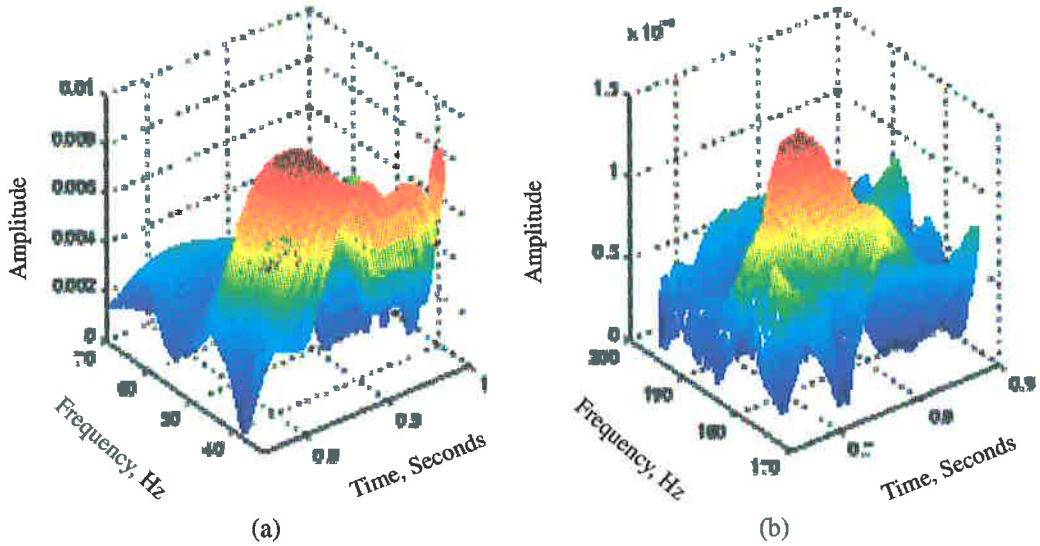


Figure 5.3 Parts of a time-frequency energy distribution of a heart sound.

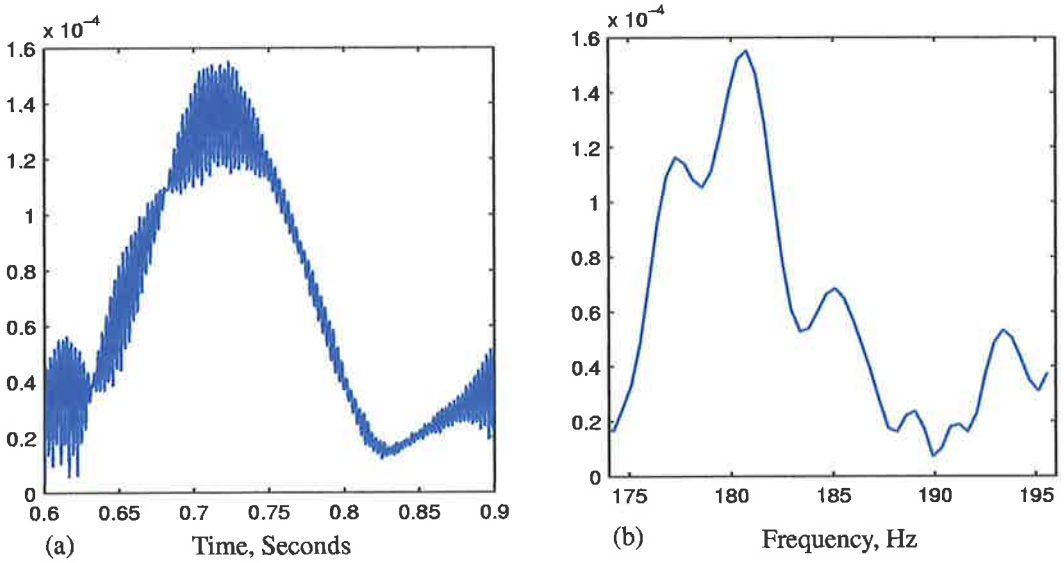


Figure 5.4 Ridges in time-frequency plane of the heart sound of Figure 5.3

5.3 Smoothing

The smoothing or filtering stage aims to suppress noise or other small fluctuations in the image surface. In image processing, smoothing has side effects such as blurring sharp edges, corners, and junctions that bear important information about the image; therefore, filtering must be done relevant to the application. This problem is not much of concern to us because we are not interested in the quality of the ridges as an

image but rather we want to smooth the surface of the ridges. However, it is necessary to choose the smoothing filter such that it would not filter out the actual ridges.

There are a number of smoothing filters that are used in image processing. They can be grouped into two large classes: linear and non-linear filters, where each class can be designed and implemented in either the frequency domain or the spatial domain. In the spatial domain, the image plane itself is the subject of manipulation, while in the frequency domain method the Fourier transform of the image is modified. The term spacial refers to all of the pixels that compose the image. Whenever image enhancement is of prime importance, then the spacial domain is used, otherwise the frequency domain is used.

The foundation of filtering techniques is the convolution theorem. Let $F(t, f)$ represent a time-frequency image, then the smoothed image can be expressed as

$$G(t, f) = h(t, f) * F(t, f) \qquad \text{Eq 5.2}$$

where $*$ is the convolution operator, $h(t, f)$ is the impulse response of the filter defined in the neighbourhood of (t, f), and $G(t, f)$ is the processed image. We define the neighbourhood of (t, f) to be a square or rectangular subimage area, centred at a local maximum, or a point where the concentration of energy is locally maximum.

One of the simplest smoothing low pass filters is the *moving average filter*. It can be designed as a finite impulse response digital filter. This filter is very effective in removing white additive Gaussian noise [96]. However, it tends to distort image details and image quality. Another type of filters that have been used for noise cancellation of images are *Wiener filters*. Their impulse responses are designed such that the mean squared error between the input image observations and the output estimation of the processor is minimized [70].

Another useful smoothing filter is *Gaussian*. In the Gaussian filter each pixel is assigned a weight whose value is inversely proportional to the distance from the central pixel. The impulse response, h(t), and the frequency response, $H(\omega)$ of the Gaussian filter are expressed as

$$h(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{x^2}{2\sigma^2}\right) \quad \text{Eq 5.3}$$

$$H(\omega) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{\sigma^2}{2}\omega^2\right) \quad \text{Eq 5.4}$$

where σ is the standard deviation of the Gaussian.

Smoothing depends on the value of σ, it should be chosen such that, the contours of the local maxima will not be smoothed too much so that we may still be able to isolate the border of the local maximum while unwanted variations in the surface still be removed. The two-dimensional Gaussian smoothing filter is

$$h(t, f) = \exp\left(-\frac{t^2 f^2}{4\sigma_t^2 \sigma_f^2}\right) \quad \text{Eq 5.5}$$

where σ_t and σ_f are the time and frequency standard deviations.

In Figure 5.5, the time domain, the frequency domain, and 3-D plots of a Gaussian filter are shown.

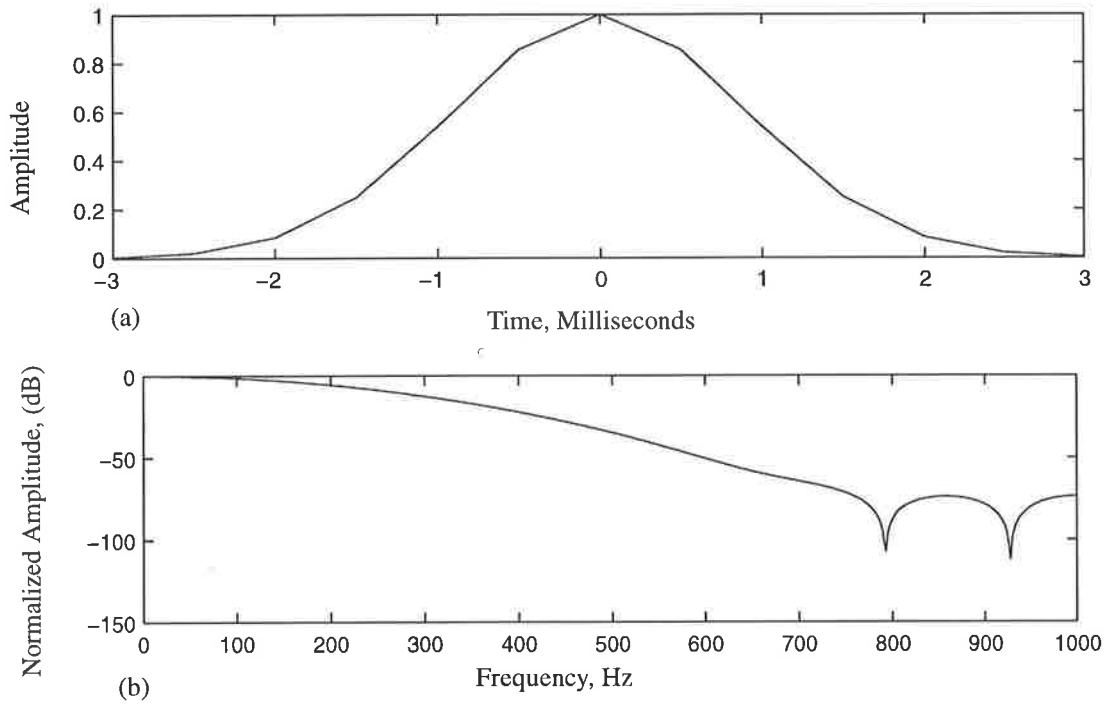
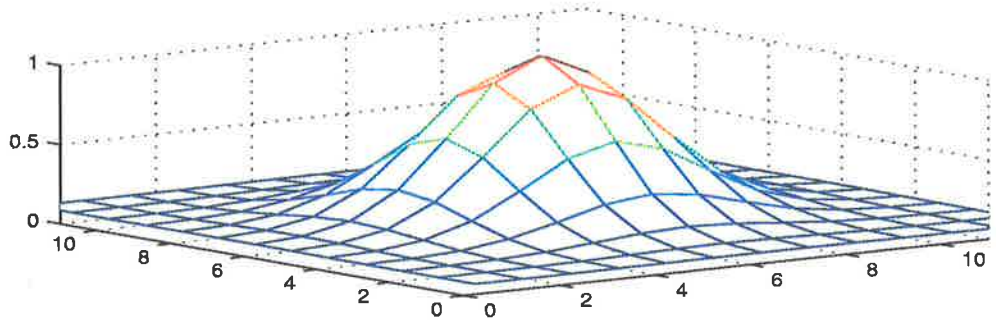
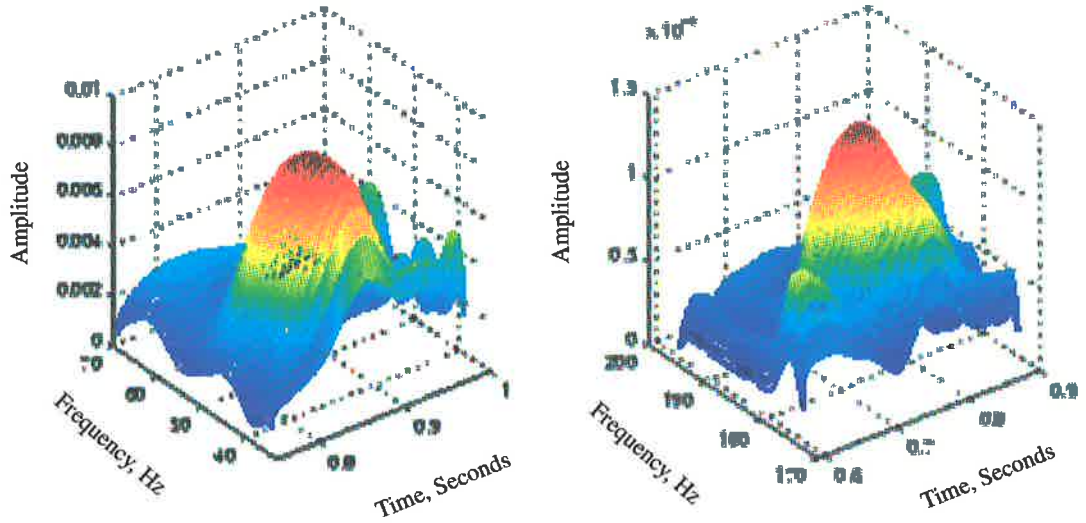


Figure 5.5 A Gaussian smoothing filter a) Time domain, b) Frequency domain.



(c)

Figure 5.5 Continued c) 3-D representation.



(a) (b)

Figure 5.6 Smoothed plots of time-frequency plots of Figure 5.3.

A 2-D Gaussian smoothing filter with σ_t and σ_f equal to 0.0316, and the time domain length of 6 milliseconds was generated. The smoothing Gaussian filter was applied to the time-frequency representations of the heart sound signals of Figure 5.3. The 2-D convolution of the filter with the signals of Figure 5.3 were obtained, results are shown in Figure 5.6. This figure clearly shows the smoothing effect of the Gaussian filter; that is, the fluctuations of the surface of the time-frequency plots have been removed. The smoothing effect becomes more evident by scanning the signals along the time and frequency axes as was performed for Figure 5.3 (b). The time-frequency representations of Figure 5.6 (b) was scanned and results are shown in Figure 5.7. By

comparing Figure 5.4 and Figure 5.7, it is clear that the surface fluctuations are removed while the shape of the ridges have not changed. This implies that the time-frequency representations have not been distorted.

5.4 Region Segmentation

It was mentioned earlier that we consider the modulus of the time-frequency plane to represent an image where the local energy concentration regions would be the textures or patterns of the image. We intend to find the location of the coordinates of each local concentration of energy and then remove it from the time-frequency plane. It was explained in the previous section that each local maximum point in the time-frequency planes of the heart sounds has a cone-shaped body. Once the centroid of a local maximum has been detected it is necessary to remove it from the time-frequency plane. This requires detecting the boundaries of the local maximum. In this section we develop a technique to accomplish that.

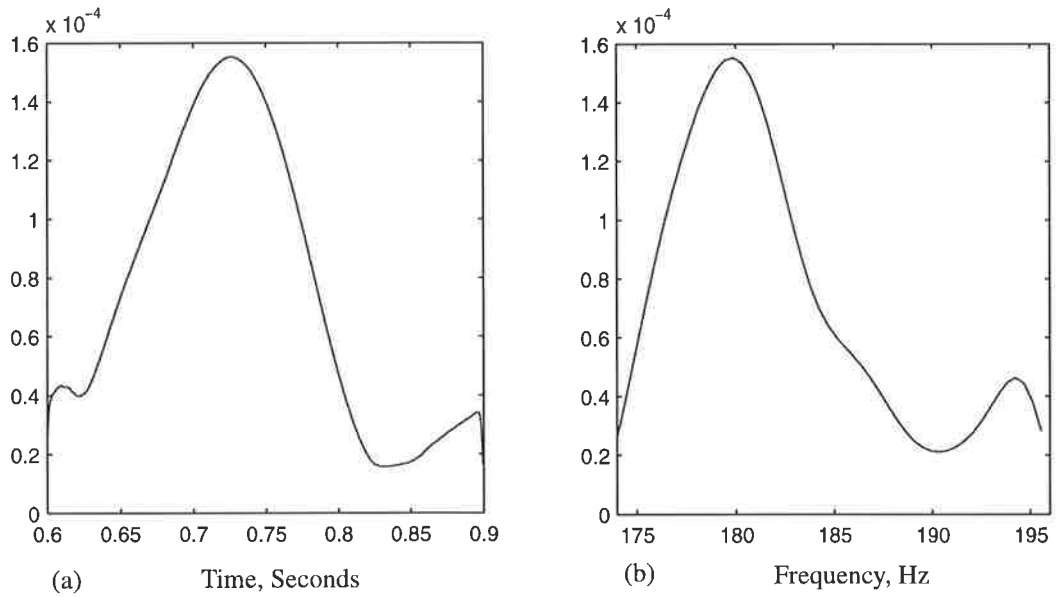


Figure 5.7 Signals obtained from scanning the time-frequency representation of Figure 5.6 (b).

Image segmentation is one of the most important subjects in any vision system. Segmentation subdivides an image into its constituent parts or objects. In order to suc-

ceed in processing of images for pattern recognition, it is necessary to achieve a good low-level feature extraction. The first step in the sequence of the operations that must be performed is to obtain a compact description of the input image. In the early stages of processing, it is essential to capture all the structural information contained in the image. Structural information refers to junctions, contours, corners, and curvature points in the image that separate patterns from each other; that is, they correspond to object boundaries or changes in surface orientations or material properties. These give important clues to the three-dimensional structure of the image. The level at which this subdivision is carried out depends on the problem being solved. The segmentation goal is achieved when the objects of interest in an application are isolated [53], [95], [102], [117].

There are many application dependent techniques used for image segmentation. Segmentation algorithms can be categorized into two broad groups. One approach is partitioning the image, based on abrupt changes in gray level. Within this category the principle areas of interest are lines and edges in the image. The other category is based on thresholding.

In the time-frequency planes, local maxima are located close to each other. The energy of local maximum is highest at its centre of gravity. As we move away from its centre of gravity the energy decreases generating a skirt around it. At the intersections of skirts a boundary is formed which we called it an *edge*.

5.4.1 Edge Detection

In the time-frequency plane, an edge is the boundary between two regions with relatively distinct energy level properties. Therefore, ideally an edge should represent an abrupt change in the energy levels of the TFIs; consequently, an ideal edge can be modelled as a step function. This is not an adequate model for edges in real TFIs because mutual effects of neighbouring patterns significantly change the energy level and shape of the edges in real TFIs. As a consequence of these effects, real TFI edges could be in the form of a ramp edge. In Figure 5.8, three different types of edges are shown.

Most of the edge detection techniques are based on some decision making stage followed by a linear filtering stage [122], [126]. The linear filtering operation is

intended for noise reduction purposes and was discussed in section 5.3 as smoothing filters. In the second stage, differentiation is used to accentuate intensity changes which will transform the image into a representation from which properties of these changes can be extracted.

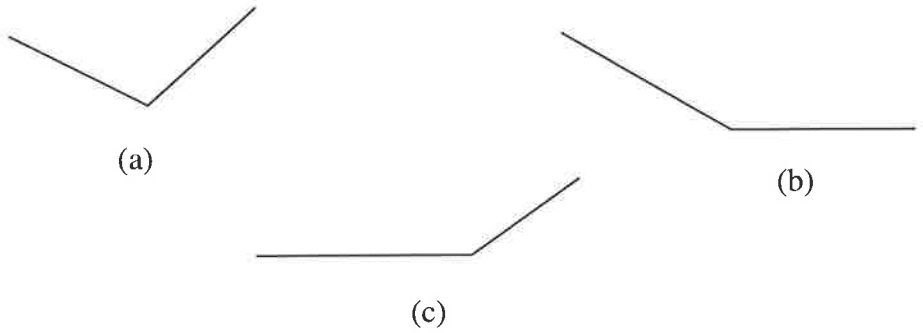


Figure 5.8 Three types of edges found in time-frequency representations of the heart sounds.

For gray level edge detection in computer image processing, two different approaches have been proposed, the extrema scheme [23], and the zero crossing scheme [83]. In the extrema scheme, edges are assumed to be sharp changes in the illuminance of the image profile. In other words, edges are step edges. In this method peaks in the first directional derivative of the Gaussian are used to find intensity change along a particular orientation in the image, this was first proposed in [23]. The filter used in this method is expressed as

$$\nabla G = \frac{1}{\sqrt{2\pi}\sigma} r \exp\left(-\frac{r^2}{2\sigma^2}\right) \qquad \text{Eq 5.6}$$

where G is the Gaussian function, ∇ represents the gradient operator, and $r^2 = x^2 + y^2$. The location of the maximum of convolution of Eq 5.6 with an image will be the place of the edge in the image.

In the second technique, the basic approach is to convolve the signal with the Laplacian-of-Gaussian (LoG) and then find the zero crossings of the convolution, this was first proposed in [83]. LoG is expressed as

$$\nabla^2 G = -\frac{1}{\sqrt{2\pi}\sigma^3}\left(1 - \frac{r^2}{\sigma^2}\right)\exp\left(-\frac{r^2}{2\sigma^2}\right) \qquad \text{Eq 5.7}$$

where G is the Gaussian function. Let $R(x, y)$ denote an image and $*$ represent the convolution operator, then the edges of the image are found by seeking the zero crossing in $F(x,y)$ by

$$F(x, y) = \nabla^2[G(r)*R(x, y)] \quad \text{Eq 5.8}$$

The accuracy of the LoG operator has been covered in details in [14].

5.4.2 Edge Detection in TFIs

It was pointed out earlier that patterns in the time-frequency planes of the heart sounds are cone shaped bodies. In order to detect the contours of these patterns, the scanned signals along the time and the frequency axes of the time-frequency plane around all of the local maxima are examined. Therefore, we restrict ourselves to one dimensional edge detection problem; that is, we examine ways of detecting edges in the scanned signals of the time and the frequency axes of the time-frequency planes. In Figure 5.9, signals obtained from scanning the TFIs of Figure 5.4 along the time axis are shown. In this figure, the edges are highlighted. Figure 5.9 (a) corresponds to Figure 5.4 (a) and Figure 5.9 (b) corresponds to Figure 5.4 (b). It is clear from Figure 5.9 that the types of edges are similar to those shown in Figure 5.8. One type of edge which is shown in Figure 5.9 (b) is formed whenever two ramps intersect. Another type of edge that we can expect in TFIs is shown in Figure 5.9 (a). This type of edge is conceivable for local maximum which may be in an isolated situation away from the other local maxima.

We can see that changes in the surface of the TFIs are not abrupt (step type) but rather they are gradual changes (ramp type). The derivative of a Gaussian edge detector technique would not be suitable here, since it is based on the assumption that the edges of concern must be step type edges.

Our edge detection algorithm is based on the zero crossing scheme. Let $TF(\mu)$ represent the signal resulting from scanning the time-frequency plane of the heart sounds along the time and frequency axes around a local maximum, where the argument μ could represent t for time axis scan or could represent f for frequency axis scan. We set the origin on the peak of the scanned signal, and examine the slope of the signal in both directions along the horizontal axis. For instance, consider the signal

shown in Figure 5.9 (b). We apply the above definitions to this signal. This signal is represented as *TF(t)*, its origin of the time axis is defined to be where the signal is maximum, and it is shown in Figure 5.10.

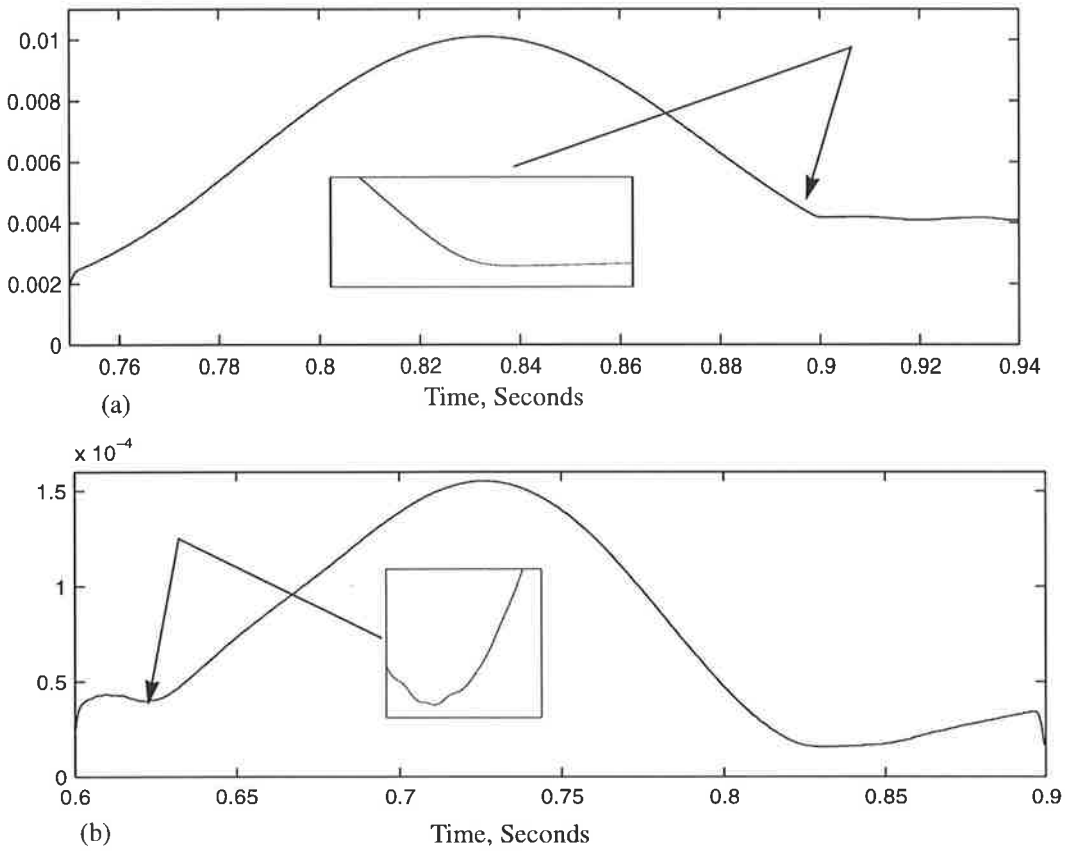


Figure 5.9 Type of edges in TFIs.

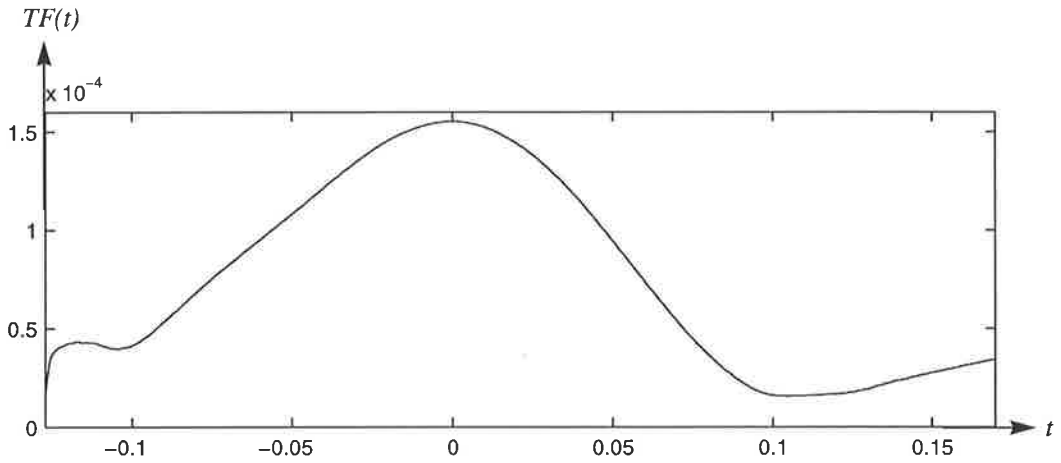


Figure 5.10 Definitions of edge detector algorithm applied to signal of Figure 5.9 (b).

It is clear that the slope is negative in both directions (left and right) of the origin until we reach the edges. Two conditions can happen as we pass the edges. If the edge is similar to one shown in Figure 5.8 (a), then the sign of the slope should become positive after the edge point is passed. However, if the edge is similar to one shown in Figure 5.8 (b) then the slope should become zero after the edge point is passed. It is clear that the edges in this figure are similar to Figure 5.8 (a); therefore, the sign of the slope changes as the edges are passed.

The above discussion implies that there is a sudden change in the sign of the slope at the edge points. Subsequently two different situations may occur here. First, the magnitude of the slope just after the edge is not zero, it could be greater than or equal to the magnitude of the slope just before the edge. Second condition is, the slope of the curve immediately after the edge becomes zero. This implies that there should be a zero crossing in the first derivative of the signal at the edge points for the first situation. For the second situation, there also will be a zero crossing but the derivative will stay zero after it reaches zero. Therefore, we focus on detecting the change in the sign of the slope of the scanned signal and finding the zero crossing points. Let the derivative of the scanned signal be

$$F(\mu) = \frac{d}{d\mu} TF(\mu) \qquad \text{Eq 5.9}$$

In Figure 5.11, $TF(\mu)$ and $F(\mu)$ corresponding to the signals shown in Figure 5.9 are shown. In Figure 5.11 (a), zero crossing corresponds to a point in $F(t)$ where $TF(t)$ becomes constant and its slope changes from negative to zero. The zero crossing of $F(t)$ in Figure 5.11 (b) corresponds to the lowest point in $TF(t)$ where its slope changes from negative to positive.

The algorithm for calculating the edge locations described above is shown in Table 5.2.

5.5 Boundary Detection Methods for TFIs

As was mentioned earlier, the proposed local maxima detection algorithm is based on detection and removal of the points of concentrations of energy of the time-

Table 5.2 Edge detection algorithm.

1) Scan the smoothed TFI along the time and frequency axes centred on the peak of the local maximum point.
2) Calculate the derivative of the scanned signal to the right and left of its peak location.
3) Find the first zero crossing of the derivative vector found in step 2 to the right and left of the peak of the scanned signal.

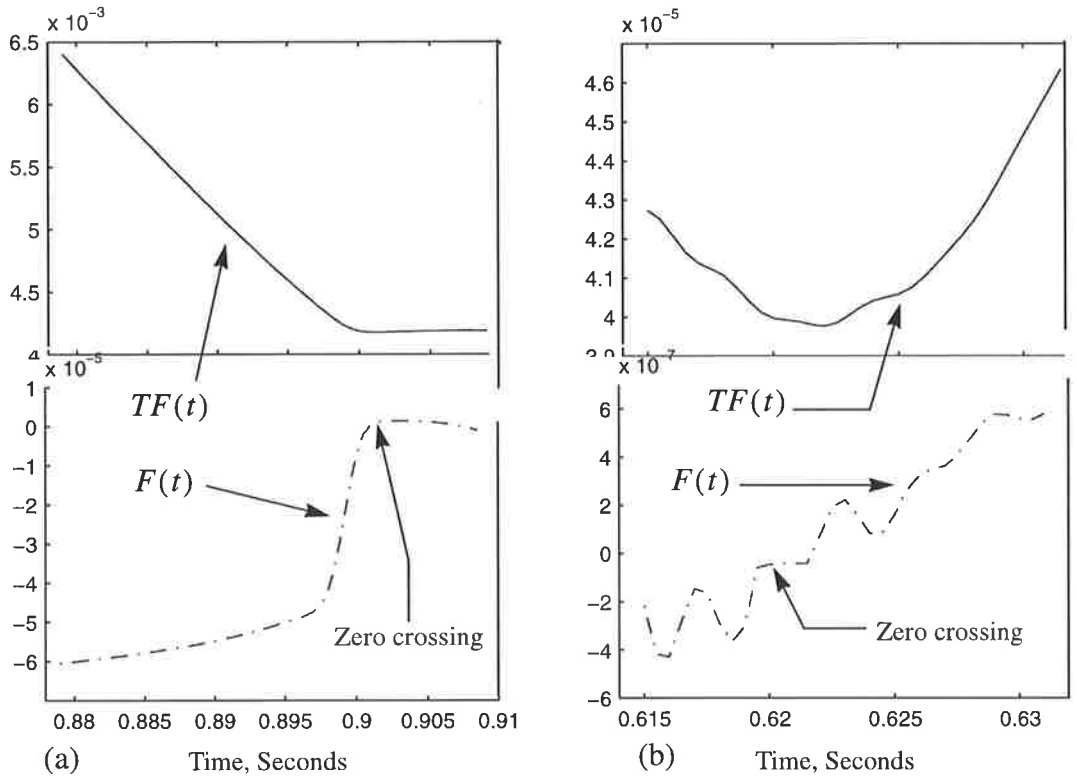


Figure 5.11 Edges of signals of Figure 5.9 and their first derivatives.

frequency plane. Therefore, it is important to detect all the surface area in TFIs that are associated with a particular local maximum, then remove it before proceeding to the next local maximum point. That is, we ought to identify the area that is confined in between the edges of every local maximum. Finding that area correctly, will reduce the probability of detecting false local maxima. In order to achieve this we propose two methods.

- ***Method 1:***

In this method, the current local maximum is scanned diagonally in all directions and edges associated with every scanned signal are found. This is shown in Figure 5.12. An initial angle value is chosen, then it is incremented by a constant value until the TFI is scanned in all directions. Let θ represent the scanning angle. Initially it is set to zero, i.e. horizontal direction, then in each step it is incremented by angle φ in each step until $\theta = 360$ is reached. This algorithm is shown in Table 5.3. In order to achieve good results with TFIs of the heart sounds, ϕ must be small, because as the distance of edge from local maximum gets larger the estimation of the connecting path between two consecutive scans becomes difficult.

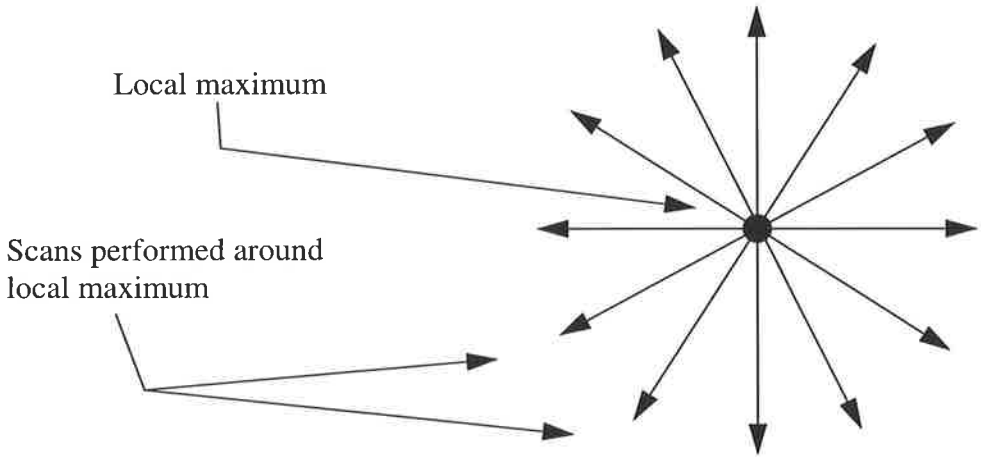


Figure 5.12 Scanning of entire pattern around a local maximum.

- ***Method 2:***

This method is based on scanning the signal only along the time axis. Initially, scanning is performed along the time axis at the frequency where the peak of the local maximum occurs and locations of the edges are found. In the next step the current scanning line (the frequency axis location) is incremented by one. Then the location of the maximum of the new scanned signal is found and the origin is set to that point, and again the location of the edges are found. In each step, the maximum value of the scanned signal is compared with the maximum value of the scanned signal in the previous step. The process is continued until a line is reached whose maximum is greater than the previous line. This is shown in Figure 5.13, and the algorithm is shown in Table 5.4.

Table 5.3 Algorithm for edge detection, method 1.

1) Initialize the angle θ. Scan the local maximum in the direction of θ.
2) Detect the edges.
3) Increment θ by φ, where φ is the incrementation angle.
4) If θ is less than 360 degrees, repeat steps 2-3 otherwise end.

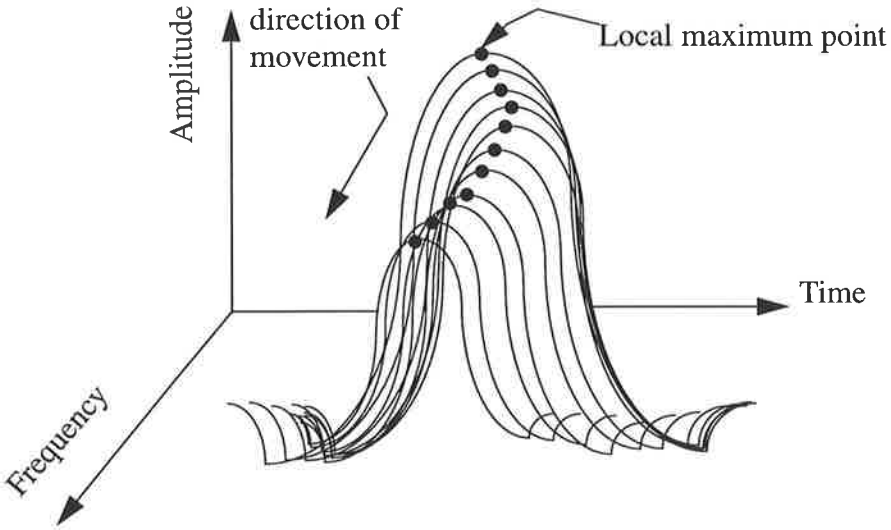


Figure 5.13 Edge detection by *method 2.*

5.6 Removing Patterns from TFIs

Once all of the edge locations are obtained for a certain pattern, they are used to construct the boundaries of the pattern for the current local maximum. The next step is to remove the detected pattern from the time-frequency plane. The removal process simply replaces all of the entries associated with the pattern of the current local maximum in the time-frequency matrix by zeros.

When all the entries within the current local maximum boundaries are set equal to zero, then the process starts again by locating another local maximum. This will continue until all local maxima are detected, or a predetermined threshold energy level is reached.

The local maxima detector algorithm explained in previous sections is shown in Figure 5.14 in block diagram form. In this figure, $X(t,f)$ represents the modulus of the time-frequency matrix.

Table 5.4 Algorithm of edge detection, method 2.

1) Scan the smoothed TFI along the time axis centred on the local maximum point.
2) Find the edges of the scanned signal.
3) Increment the frequency axis location of the current scanning. Scan and find edge locations.
4) Find the maximum of the scanned signal.
5) If maximum of the current vector is less than the maximum of the previous vector then repeat steps 3-4, otherwise end.
6) Repeat steps 1-5 with the exception that in step 3 decrement the frequency axis location.

5.7 Criteria for Accepting the Detected Local Maxima

The number of iterations in the local maxima detector algorithm is equal to the number of local maxima in the time-frequency plane. It was mentioned earlier that some of the low energy level local maxima may be due to noise. It is essential to distinguish them from the local maxima which are due to the signal components. This depends very much on the signal-to-noise ratio of the signal. For a relatively high signal-to-noise ratio, one would expect to get a higher energy local maxima for the signal components than those of noise. Therefore, if the amplitudes of local maxima are plotted, a gap in amplitude levels may be evident. That is, the local maxima of the signal components would appear in higher levels than the local maxima which are due to noise. A figure similar to Figure 5.15 may be obtained. If this is the case then a threshold level can easily be set to separate two clusters of detected local maxima. One cluster would represent the local maxima of the signal components and the other would be a cluster of unwanted local maxima.

Heart sounds are complex signals, therefore it may not be very easy to set the threshold level and separate the local maxima into two groups. Distinction between two

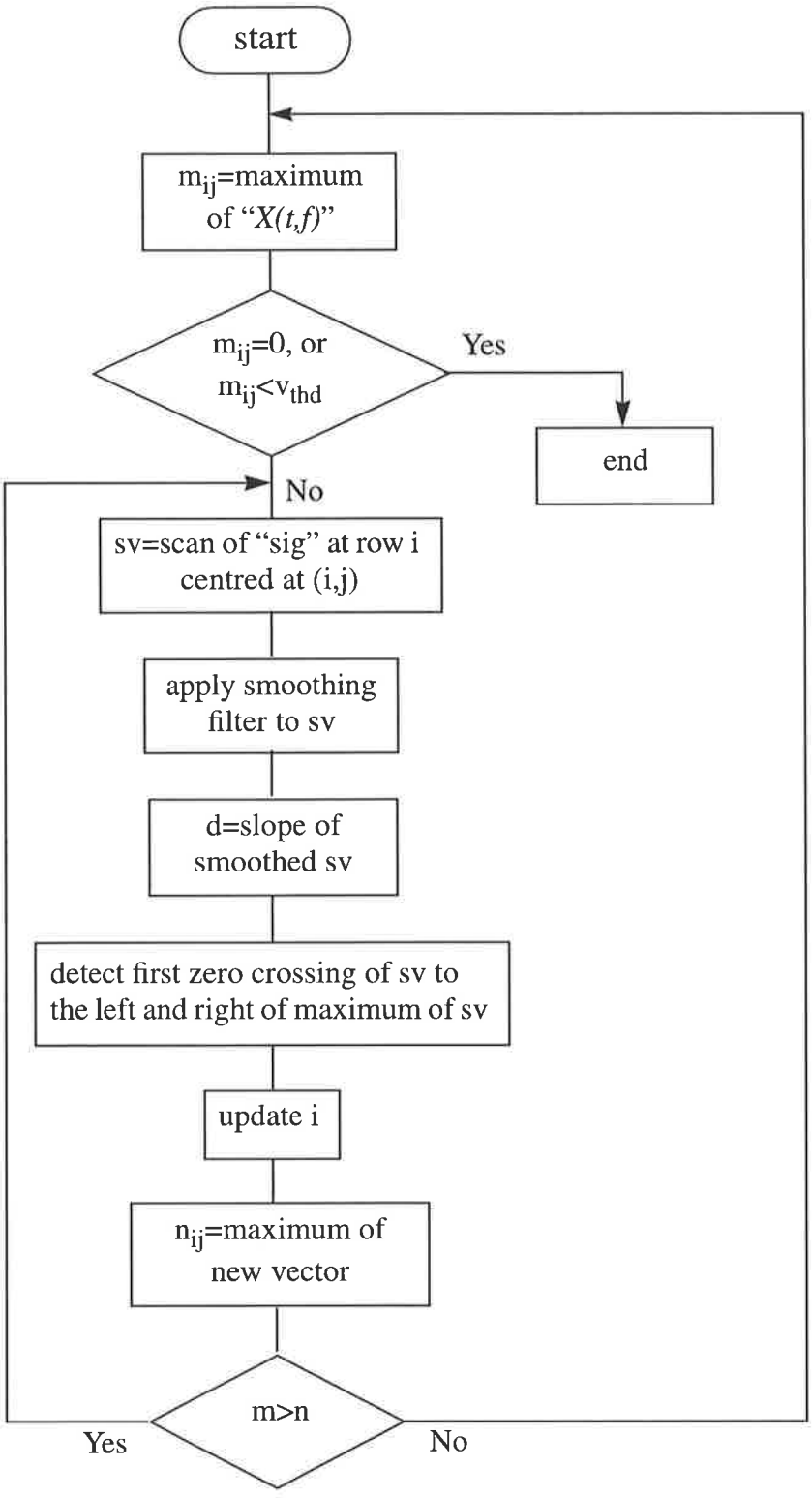


Figure 5.14 Block diagram of local maxima detector.

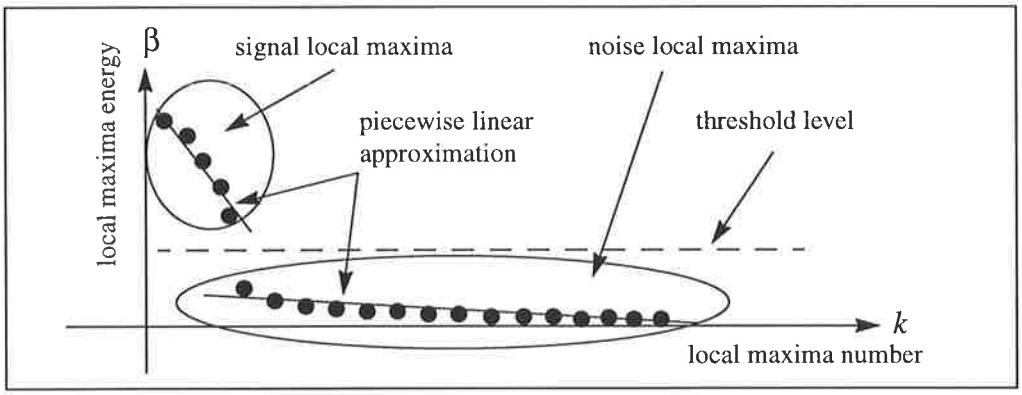


Figure 5.15 Signal and noise local maxima under high SNR condition.

clusters of local maxima for the heart sounds have been accomplished by the following method. Let $\beta(m)$ represent the energy level of the m^{th} local maximum in the TFI of the heart sound, and q be the total number of local maxima. Let the graph of $\beta(k)$ versus k be denoted as the *LM* curve, where k=1, 2, 3,..., q, and where increasing k implies decreasing energy. We found that it is possible to make a piecewise linear approximation to the *LM* curve. The piecewise linear approximation was performed in the following way. Note that, the basis for local maxima detection was detecting the highest energy level; therefore, the slope of the LM curve will be negative every where. Any large change in the values of the energy levels results in a sharp change in the slope of the curve. Therefore, if we consider the slopes of the different sections of the LM curve and compare them, then the smallest slope should be chosen. In other words, starting at the top of the LM curve, we select a part of the curve and calculate its slope. Then another portion with larger LM points is selected and its slope is calculated also. If the slope of the second part is the same as the previous one then the number of the LM points are increased, otherwise the first part is selected as a section with constant slope. This is continued until the LM curve is piecewise approximated. Therefore, there are always sections that can be approximated with a constant slope line. The number of sections varies for different heart sounds and depended on the type of time-frequency technique used. This is illustrated in Figure 5.15. Visual inspection of the TFIs of the heart sound revealed that in most of the cases the last section in piecewise linear approximated *LM* curves represents the local maxima due to noise. We used this method for clustering the local maxima of the heart sounds. Of course, it is always possible that few signal local maxima may fall into the noise local maxima cluster and vice versa.

5.8 Simulation example

In this part we will illustrate some simulations results obtained using the local maxima detector algorithm. We apply the local maxima detector to the time-frequency representation shown in Figure 5.2. The reasons for choosing Figure 5.2 are: 1) the energy concentration levels are clear in the figure; 2) it has both a high energy level local maximum as well as a low energy level one; 3) two of the local maxima are close to each other both in frequency and time, so they influence each other's boundaries a great deal; 4) the time series signal from which it is obtained has a noise component. The normalized energy level, and time-frequency locations of local maxima points for the signal of Figure 5.2 are shown in Figure 5.16 (a) and (b), respectively. The gap between the energy levels between the local maxima corresponding to signal components and the local maxima corresponding to noise, is evident. Note that there are only three signal components present in Figure 5.2, namely 100 *Hz*, 180 *Hz*, and 200 *Hz* which are located at different positions in time. These are highlighted in Figure 5.16 (b). Therefore, threshold level can be set to any value between 0.1 and 0.4 to separate the signal and noise local maxima, e.g. 0.3.

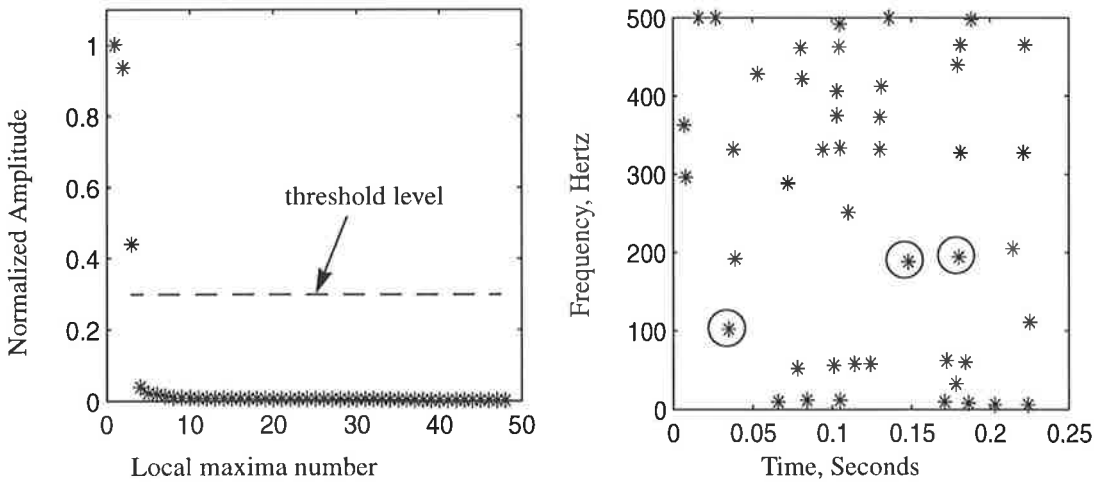


Figure 5.16 Local maxima points of Figure 5.2, a) amplitudes, b) time-frequency locations.

The local maxima detector was applied to the modulus of the time-frequency representation of a heart sound, parts of which were shown in Figure 5.3. The heart sound signal was prepared using the algorithms presented in Chapter 4. The short-time Fourier transform was used for the time-frequency calculations. The Gaussian smoothing filter of Figure 5.5 was used to suppress the surface fluctuations of the modulus of

the time-frequency matrix. The normalized energy levels of the detected local maxima are shown in Figure 5.17 (a). In order to show the energy levels more clearly, a logarithmic scale is used. As is in the Figure 5.16 (a), it is clear that there is not any distinctive gap in the energy levels. Piecewise linear approximation of this curve is shown in Figure 5.17 (b). There are five linear segments in this approximation. The fifth segment, which possesses the smallest slope, has the best approximation to the *LM* curve. This segment is considered to be due to noise, hence it is discarded. The local maxima points are shown in Figure 5.17 (c), and in Figure 5.17 (d) the accepted local maxima are shown.

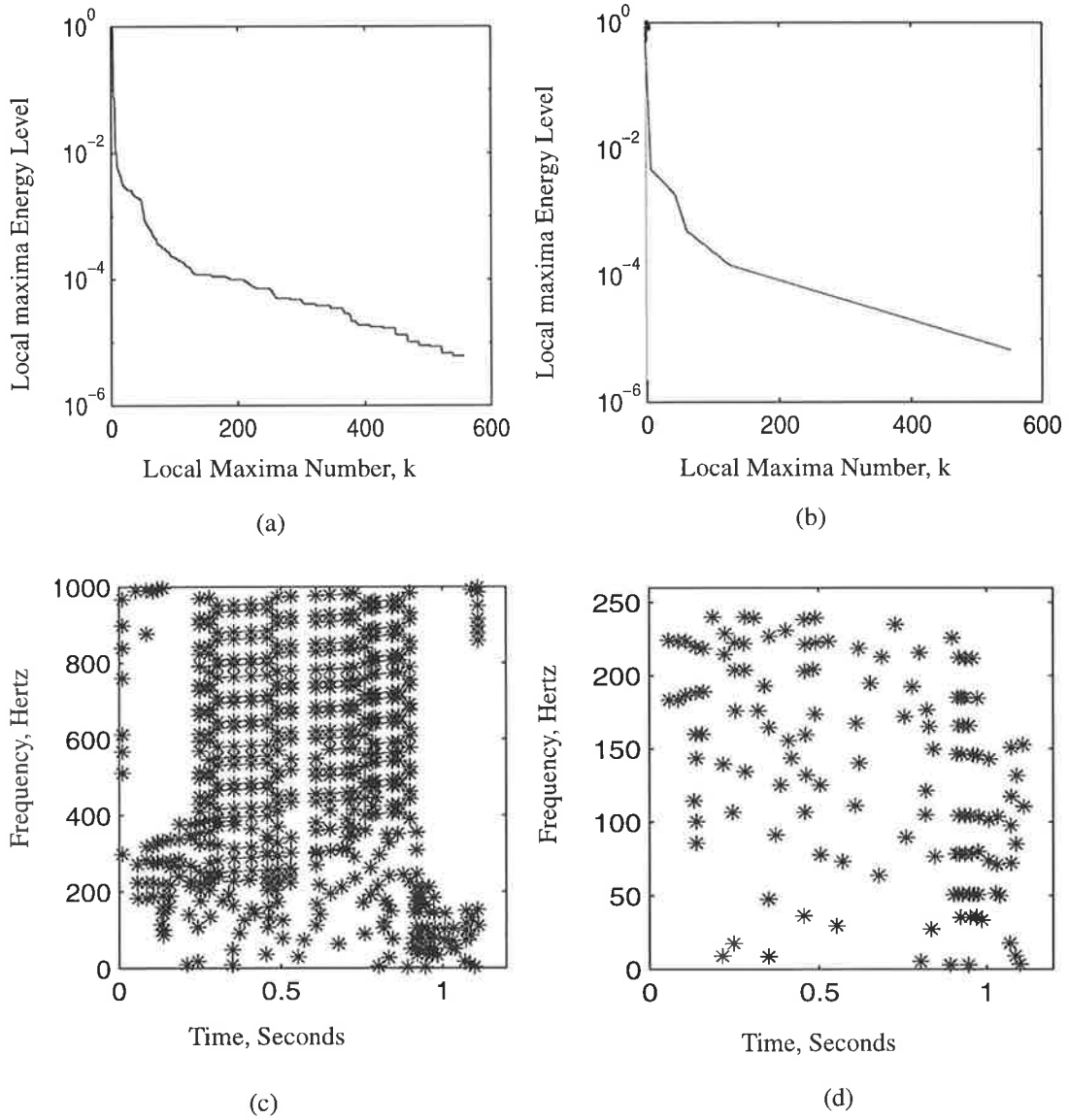


Figure 5.17 a) LM curve of the heart sound, b) piecewise linear approximation of the LM curve, c) detected local maxima, c) accepted local maxima.

5.9 Conclusions

We presented a technique to detect local maxima in the time-frequency plane with the intention of applying it to the heart sound signals. Local maximum points were defined to be the centroid of each local concentration of energy in the modulus of the time-frequency matrix. The 3-D representations of the modulus of the time-frequency matrices were perceived as 3-D images where the ridges were considered as the textures of the image. This algorithm is based on edge detection. That is, the highest energy level in the time-frequency matrix is detected, its boundaries are identified, and then it is removed from the time-frequency plane. The process is repeated until all of the local maxima are detected. This technique can be applied to the detection of ridges in the energy distribution of any signal.

We presented two methods for detection of the boundaries of the local maximum point. In 95% of the cases we were able to get the same results by both methods. We used one to confirm the results obtained by the other. The first method is more computationally intensive when φ is small.

The performance of the proposed detection technique was illustrated by a simple sinusoidal signal, with additive white noise. It was shown that the algorithm performs very well with the heart sound signals as well. We proposed a method to separate the local maxima of signal components from those of the noise components.

Chapter 6

Short-Time Fourier Transform Analysis of Heart Sounds

6.1 Introduction

The Fourier transform is among the few tools that can be considered as a keystone of modern signal processing. Development of fast Fourier transform (FFT) algorithms has greatly influenced the advancement of digital signal processing. Fourier representations are useful for two basic reasons. The first is that for linear systems it is convenient to determine the response to a superposition of sinusoids or complex exponentials. The second reason is, Fourier representations often serve to place in evidence certain properties of the signal that may be obscured or at least be less evident in the original signal. The relative intensities of the frequency components of signals are calculated using FFTs, but this does not give any information about when in time a specific frequency has occurred. For stationary signals, where signal frequencies do not change in time, FFT gives a clear understanding about the spectrum of the signals, but since in nonstationary signals the frequency content of the signals does change with time, it is necessary to analyse these signals jointly in time and frequency domains. Most of the physical signals in nature are of this kind, such as human speech, sounds made by animals, biological signals, transient signals, etc. This led to the idea of the

short-time Fourier transform (STFT) which is an extension of the Fourier transform as it was defined by Gabor [47]. Today it is the standard simple method for study of the nonstationary signals.

STFT is a windowed version of the discrete time Fourier transform (DTFT) which is calculated by a sliding window indexed by both time and frequency [9], [10], [37], [89], [97], [99]. The time-frequency description of the signal obtained by this method is called a spectrogram. Choice of the window plays an important role in the quality of the time-frequency analysis of the signal by STFT. The purpose of the window is to localize the frequency content in time. Unfortunately, temporal and frequency resolutions are inversely related to each other and they are restricted by the uncertainty principle.

Spectrograms have a long history but nevertheless, the short-time Fourier transform method only recently has been used to study the status of the heart using phonocardiograms and electrocardiograms. Some applications of STFT in heart sound analysis are pointed out in Chapter 1.

Non-heart related applications of STFT are also wide spread. For example, a spectrogram with a Gabor window has been used in [67] as a texture discriminator to identify breast masses. The complicated time-frequency characteristics of dolphin whistles were analysed in [128] using spectrogram and other time-frequency methods. In [109], a model of the bat's echolocation system using spectrogram was proposed. Underwater acoustic signals were analysed in [18].

In this chapter we first briefly review the short-time Fourier transform and then heart sounds are analysed using STFT. In section 6.2, STFT is reviewed as a windowed Fourier transform. In section 6.3, we discuss the restriction imposed on STFT by the Heisenberg inequality. In section 6.4, the application of the STFT in the analysis of the recorded heart sounds is discussed. First, the heart-beat signals obtained by the algorithms of Chapter 4 are used for STFT calculations, then the algorithms presented in Chapter 5 are used for local maxima detection of modulus of the STFT time-frequency planes. We illustrate the results obtained for phonocardiograms from both before and after angioplasty recordings. In section 6.5, a method is developed for analysis of the locations of the concentration of energy in the time-frequency matrices. Time-fre-

quency patterns in the time-frequency matrices are recognised by the proposed method. The average frequency of the patterns are calculated for first, second, and third heart sounds. Then a comparison is made and differences are observed for the results obtained from the analysis of all data records.

6.2 The Short-Time Fourier Transform

The discrete-time Fourier transform (DTFT) is used to analyse signals for their frequency contents. Let $s(n)$ be a signal defined for all n and $S(e^{j\theta})$ be its DTFT evaluated at frequency θ. The analysis and synthesis equations may then be defined as Eq 6.1 and Eq 6.2 respectively [98].

$$S(e^{j\theta}) = \sum_{n=-\infty}^{\infty} s(n)e^{-j\theta n} \qquad \text{Eq 6.1}$$

$$s(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} S(e^{j\theta})e^{j\theta n}d\theta \qquad \text{Eq 6.2}$$

The function $S(e^{j\theta})$ gives a measure of the frequency content of the signal $s(n)$ at frequency θ. These equations yield a perfect insight to signals whose frequency contents are not changing with time. However, if they are used to analyse a signal whose frequency varies as a function of time, that is if the signal is not stationary, then the results will not indicate correct spectral information of the signal. In other words, the transform $S(e^{j\theta})$ will be localised in a certain frequency, say θ_i, only if the signal $s(n)$ has infinite length in the time domain.

If $s(n)$ is time limited to a duration of N samples, then we can sample the continuous function $S(e^{j\theta})$ with uniformly spaced points in the range $(-\pi, \pi)$. This corresponds to constructing a periodic signal $\hat{s}(n)$ of infinite duration and period N by periodically extending the sequence $s(n)$ with length N, and computing its Fourier series expansion. This is expressed as

$$S(k) = \sum_{n=0}^{N-1} s(n)e^{-j\theta_k n} \qquad \text{Eq 6.3}$$

where $\theta_k = 2\pi k/N$ and $k = 0, 1, 2, \ldots\ldots, N-1$. This is the Discrete Fourier Transform (DFT) of the signal $s(n)$. Reconstruction of $s(n)$ from the samples of the its spectrum, $S(k)$, or the corresponding Inverse Discrete Fourier Transform (IDFT) is calculated as

$$s(n) = \frac{1}{N}\sum_{k=0}^{N-1} S(k)e^{j\theta_k n} \qquad \text{Eq 6.4}$$

In order to avoid time domain aliasing it is necessary to limit $s(n)$ to length N, then it can be recovered completely by taking IDFT of the N frequency samples.

In a typical DFT computational algorithm, direct calculation of N-point DFT requires $4N^2$ real multiplications. Development of computationally efficient algorithms such as the fast Fourier transform, makes direct DFT based analysis/synthesis methods very prevalent in the signal processing literature. Use of FFT algorithms allow computation of DFT to be performed in $N \log_2 N$ computations.

For the analysis of nonstationary signals it is necessary to change Eq 6.1 in such a way that the resulting transform will be in the joint time-frequency domain. Such description can be achieved by using the short-time Fourier transform [9], [47], [99]. The STFT views the DFT of the signal $s(n)$ with a moving window $w(n-m)$. The STFT of $s(n)$ is defined as

$$S_{stft}(e^{j\theta}, m) = \sum_{n=-\infty}^{\infty} w(n-m)s(n)e^{-j\theta n} \qquad \text{Eq 6.5}$$

where $w(n-m)$ is a window sequence which selects portions of $s(n)$ that are being analysed at a particular time index m and it is called the analysis window. In other words, the frequency contents of the signal $s(n)$ are examined locally around the time instant m. Therefore, the window function helps to localize the time domain data, before obtaining the frequency domain information. In Figure 6.1, a time signal which

is part of a phonocardiogram, an analysis window, and windowed signal are shown. In this figure, the window is centred at m = 1.05 seconds. As the window is moved along the time axis, the FFT of its product with the signal of Figure 6.1 (a) is calculated for each lag m, in this way the frequency content of $s(n)$ is examined only in a region around time m.

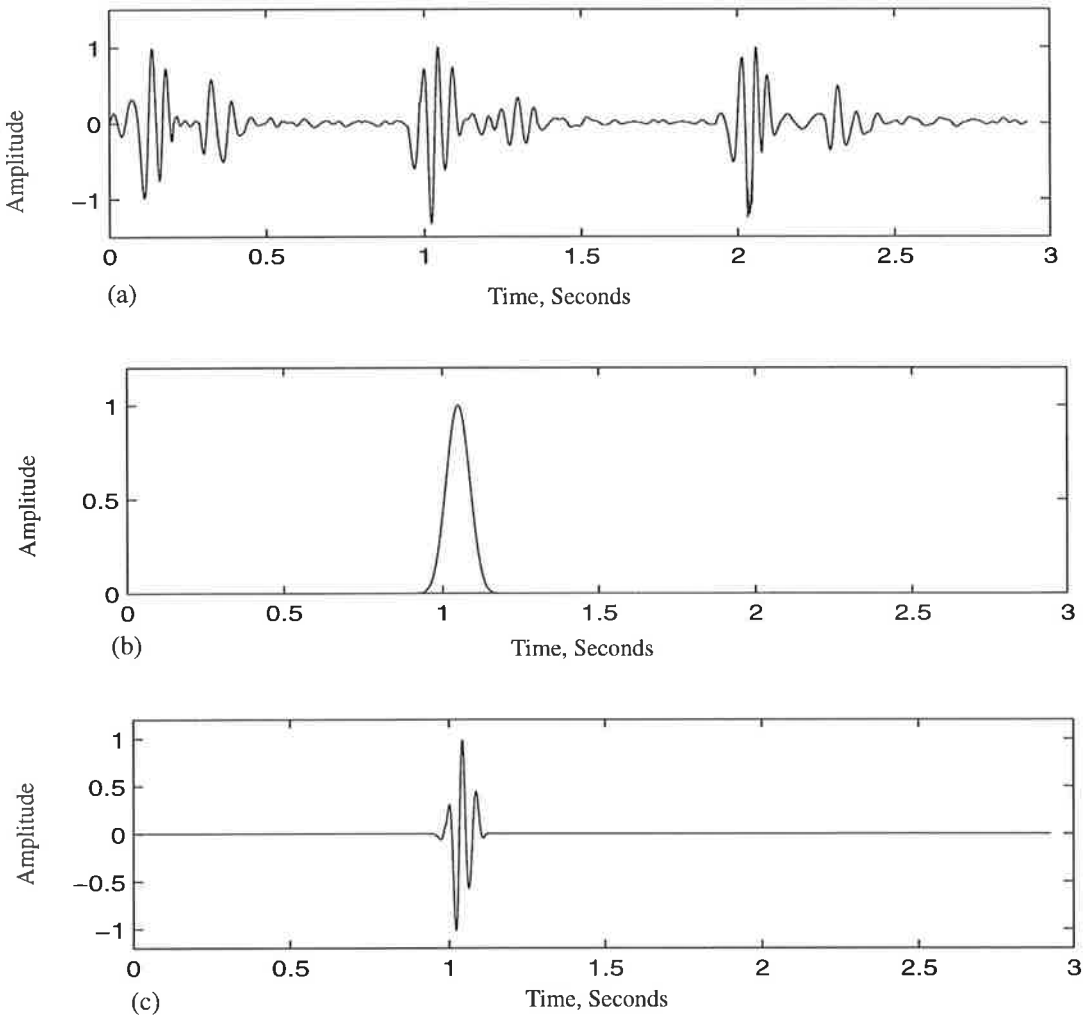


Figure 6.1 Example of a) signal, b) window, and c) windowed signal.

Generally, $S_{stft}(e^{j\theta}, m)$ expressed in Eq 6.5, is a complex valued function of two variables, θ and m, where its magnitude can be visualized as surface of the time-frequency plane with axes m and θ. In other words, the STFT maps a one-dimensional

time domain signal, into a two-dimensional function in the time and frequency plane. Evidently, it is a linear time-frequency representation.

6.3 Time-Frequency Resolution of STFT

Obviously, the choice of the window function is very important in determining the spectrogram of the input signal. If a short window is used then the resulting transform will exhibit a good resolution in time, but on the other hand a short window will only allow a small number of samples to be used in DTFT calculations and therefore, restricting the ability of the STFT to discriminate between frequencies close to each other. On the other hand, a long duration window in the time domain will discriminate close frequencies but will have a poor time domain resolution.

In order to provide a quantitative measure of the time and frequency localization of the STFT, let $w(t)$ represent a window as a function of time, and $W(\omega)$ be its Fourier transform. Therefore, $|w(t)|^2$ can be interpreted as a weight distribution of the window in time and $|W(\omega)|^2$ can be interpreted as the weight distribution of the window in frequency. The centre of the window in the time domain and the centre of the window in the frequency domain are defined by Eq 6.6 and Eq 6.7, respectively [25], [68], [144].

$$t^* = \frac{1}{\|w\|_2^2} \int_{-\infty}^{\infty} t|w(t)|^2 dt \qquad \text{Eq 6.6}$$

$$\omega^* = \frac{1}{\|W\|_2^2} \int_{-\infty}^{\infty} \omega|W(j\omega)|^2 d\omega \qquad \text{Eq 6.7}$$

where

$$\|w\|_2^2 = \int_{-\infty}^{\infty} |w(t)|^2 dt \qquad \text{Eq 6.8}$$

$$\|W\|_2^2 = \int_{-\infty}^{\infty} |W(j\omega)|^2 d\omega \qquad \text{Eq 6.9}$$

The time duration of the window, which is also known as the rms duration, is given in Eq 6.10; likewise, the rms bandwidth of $W(j\omega)$ is defined by Eq 6.11.

$$\Delta t = \frac{1}{\|w\|_2}\left\{\int_{-\infty}^{\infty}((t-t^*)^2|w(t)|^2)dt\right\}^{1/2} \quad \text{Eq 6.10}$$

$$\Delta\omega = \frac{1}{\|W\|_2}\left\{\int_{-\infty}^{\infty}((\omega-\omega^*)^2|W(j\omega)|^2)d\omega\right\}^{1/2} \quad \text{Eq 6.11}$$

where $\|w\|_2$ and $\|W\|_2$ are the Euclidean norm of $w(t)$ and $W(\omega)$, respectively. With these definitions, it can be shown that the Heisenberg uncertainty relation [84], is

$$\Delta t\Delta\omega \geq \frac{1}{2} \quad \text{Eq 6.12}$$

Eq 6.12 indicates that the uncertainty principle prohibits existence of windows with arbitrarily small duration and arbitrarily small bandwidth. This is the drawback of the STFT. Therefore, a compromise must be made between time resolution and frequency resolution.

6.4 STFT Analysis of Heart Sounds

In this section, we explain the process of the STFT analysis of the recorded heart sounds. In order to remove the uncorrelated noise from the data records, the raw data is passed through an adaptive line enhancement filter, which has been described extensively in Chapter 3. The heart-beat cycle detection algorithm of Chapter 4 is applied to the signal obtained from ALE noise canceller and the heart-beat signals are calculated for both before and after angioplasty operation. Then, the short-time Fourier transform of each heart-beat cycle (before and after angioplasty) is calculated, and the modulus of the time-frequency matrix is obtained. The local maxima of energy distribution are detected using the algorithms presented in Chapter 5. In Figure 6.2, the signal processing steps involved in time-frequency analysis of the heart sounds using STFT are shown. We first examine a single case and present all the detailed analysis for it; then, at the end of the section we present the results of the analysis of all the data records.

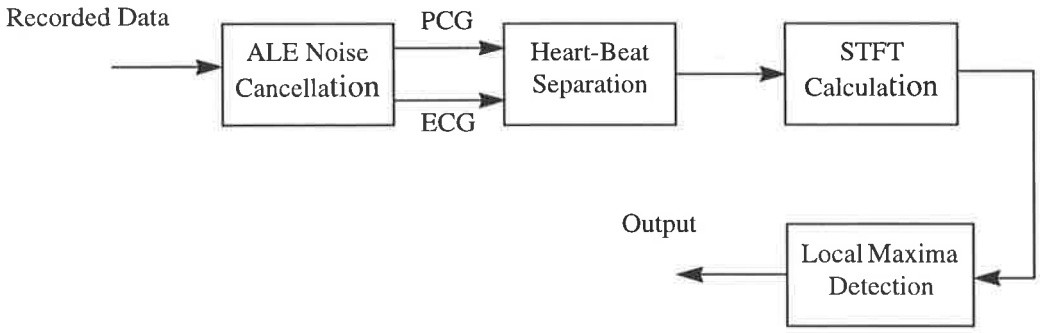


Figure 6.2 Block diagram for time-frequency analysis of heart sounds using STFT.

6.4.1 Heart-Beat Signal Preparation

Here we present the STFT analysis of the phonocardiograms (before and after angioplasty) of Case 1 of Table 2.2 in details. Our aim is to demonstrate the process of the STFT analysis of the phonocardiograms by showing plots and results that were obtained. We used 40 seconds of recordings of before angioplasty and 35 seconds of recordings of after angioplasty. A ten-second segment of the PCG along with the corresponding synchronously recorded ECG signal of both before and after angioplasty of Case 1 of Table 2.2 are shown in Figure 6.3. Parts (a) and (b) of Figure 6.3 show the PCG and ECG signals for before angioplasty and parts (c) and (d) show the PCG and ECG signals for after angioplasty. The signals were processed by the ALE techniques of Chapter 3 prior using them here in order to remove any uncorrelated noise.

Heart-beat signals were prepared using the heart-beat separation algorithms of Chapter 4. The ECG signal was scanned with 10 millisecond window and the threshold level was set to 0.6. The beat cycle separation algorithm (BCSA) detected 36 most correlated beat cycles in the signal of before angioplasty and rejected one cycle. In the signal of after angioplasty, 30 most correlated beat cycles were detected and 2 beat cycles were rejected. The output of BCSA for these signals as well as the final heart-beat signals are shown in Figure 6.4. The length of the scanning window in the BCSA was set to 10 ms for both cases.

6.4.2 Analysis Window Selection

As was mentioned earlier; the results of the STFT analysis is greatly affected by the length of the window, it determines the time and frequency resolution. The type of

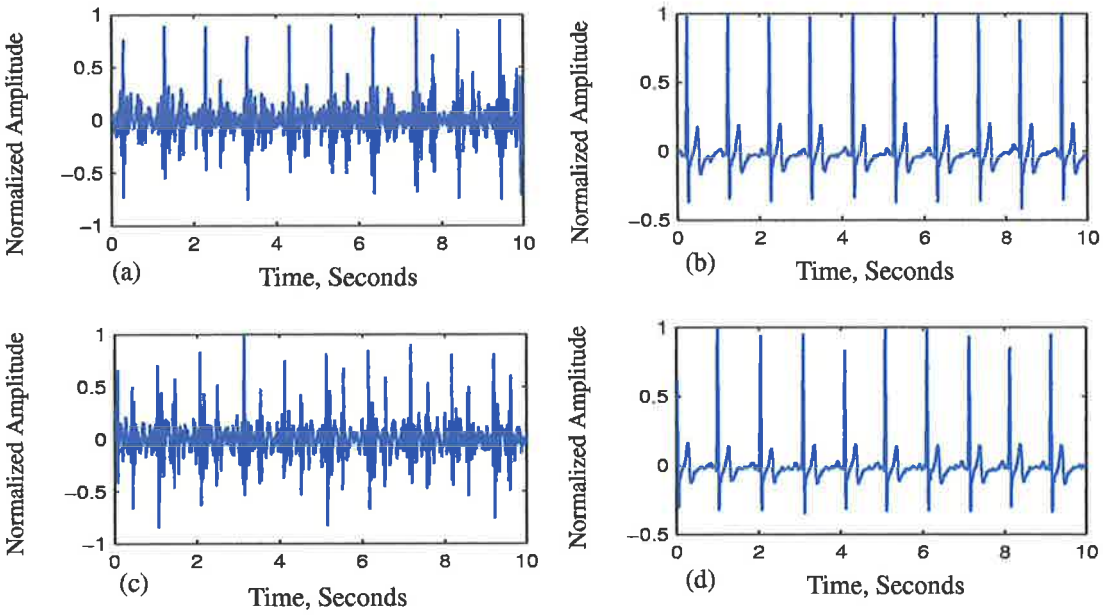
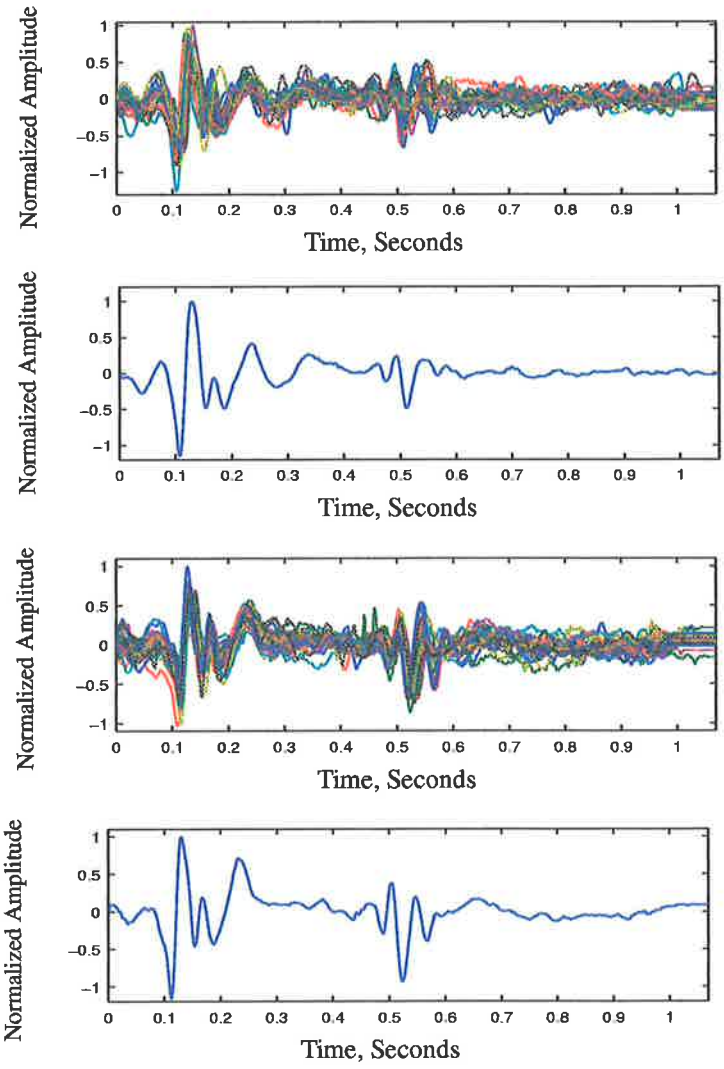


Figure 6.3 a) PCG and b) ECG before angioplasty, c) PCG and d) ECG after angioplasty.



(a) All heart-beat cycles of the recorded heart sounds before angioplasty.

(b) Heart-beat signal calculated from beat cycles of part (a).

(c) All heart-beat cycles of the recorded heart sounds after angioplasty.

(d) Heart-beat signal calculated from beat cycles of part (c).

Figure 6.4

the window is also very important. The choice of the analysis window directly affects the trade off between frequency resolution and side-lobe attenuation. The height of the side-lobe must be reduced as much as possible in order to decrease cross-term distortion, but this in turn will require wider bandwidth. On the other hand, in order to be able to resolve two closely spaced frequencies, it is necessary to have a main lobe narrow enough to distinguish them. In this application we require a small sidelobe, hence we compromise to an increased bandwidth for low sidelobe.

Four window types are considered: Chebychev, Kaiser, Hamming, and Blackman windows. The time duration of the windows were set to 250 ms. The smaller length windows were tried, but they did not have good frequency resolutions, and also higher length windows suffered from temporal resolutions. In Table 6.1 the height of the main lobe, height of the largest sidelobe, and width of the main lobe for all of the above window functions are given. The width of the main lobe is the distance (in Hertz) between first two zero crossings of the FFT of the window with respect to its peak. The magnitude plots of all four windows in the frequency domain are shown in Figure 6.5. The passbands of the windows are magnified and shown in the inset plots of Figure 6.5.

Table 6.1 Frequency response characteristics of four window types.

Window function	Height of main lobe	Height of largest sidelobe	Width of the window
Chebychev	0 dB	-80 dB	11 Hz
Kaiser	0 dB	-66.6 dB	10.5 Hz
Hamming	0 dB	-43.2 dB	9 Hz
Blackman	0 dB	-59.38 dB	10.5 Hz

Examination of Table 6.1 and Figure 6.5 reveals that the Chebychev window exhibits better sidelobe attenuation; therefore, this window was chosen in order to have least interference from sidelobe frequencies. The Chebychev has a slightly higher bandwidth than the others, but this is a secondary concern to us because we can reduce it by increasing the length of the filter. The time domain plot of the Chebychev window which was used for calculation of the time-frequency plane of the STFT of the heartbeat signals is presented in Figure 6.6.

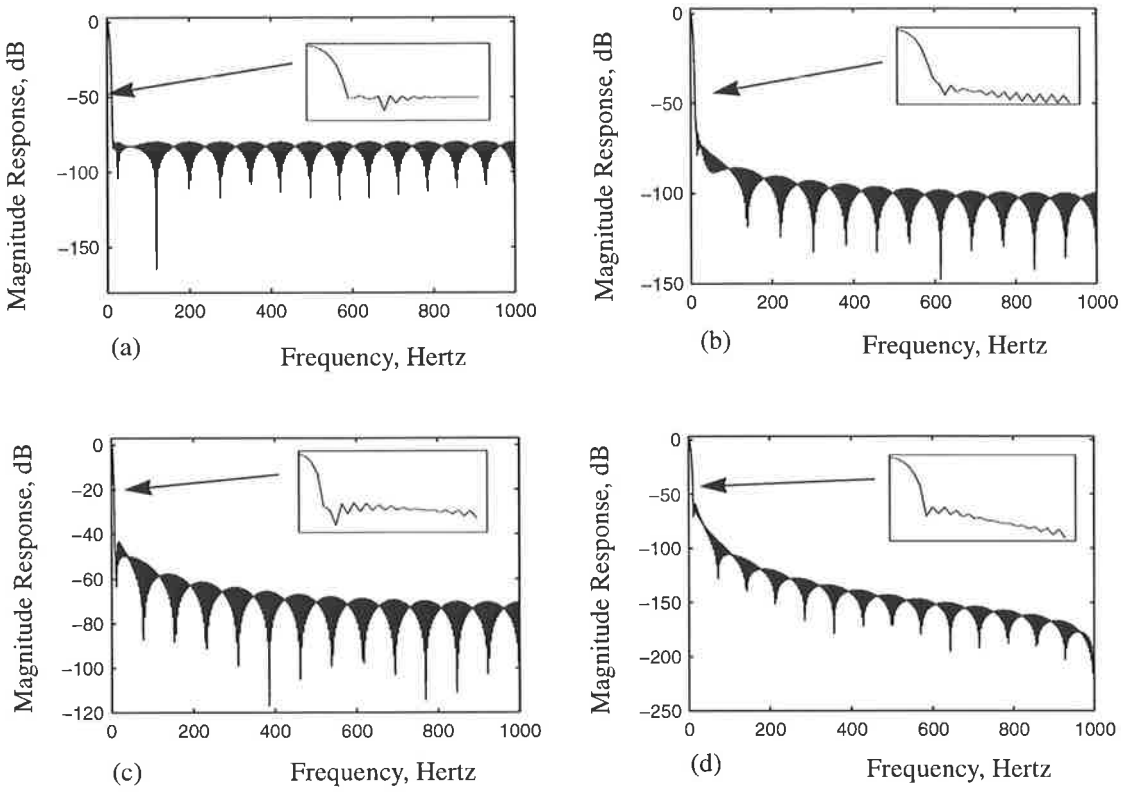


Figure 6.5 Magnitude responses of different windows are shown for comparison a) Chebychev window, b) Kaiser window with $\beta = 8$, c) Hamming window, d) Blackman window.

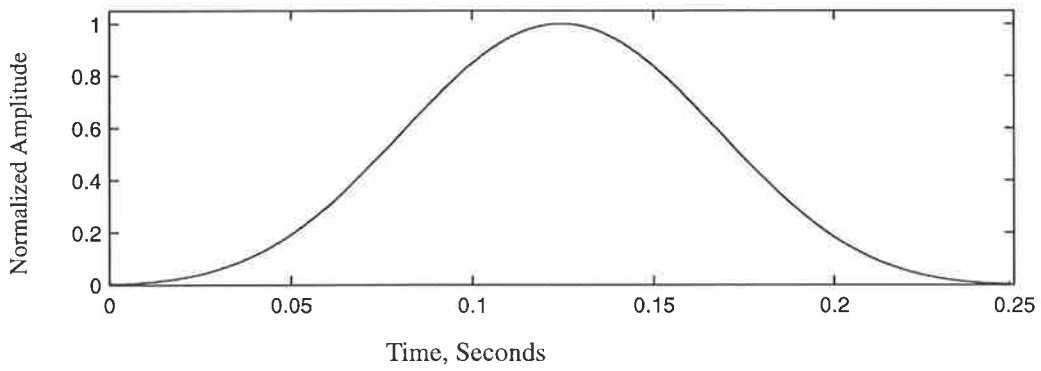


Figure 6.6 Time domain representation of the Chebychev window whose magnitude plot is shown in Figure 6.5 (a).

Once the analysis window function has been chosen, it is necessary to decide how often the window should be advanced and the windowed signal be calculated. Since our goal is to calculate the local maxima of the time frequency plane, it is important to keep the overlapping as large as possible. This will give better temporal resolu-

tion, but it will increase the computation time and the resulting STFT matrix will be large.

6.4.3 STFT Calculations

Once the window was determined, a program was written and the STFT was implemented. The STFT of the signals shown in Figure 6.4 (b) and (d) were calculated. The time shift for advancing the window was set to 0.5 milliseconds. Note that the length of the scanning window was set equal to 250 *ms;* therefore, advancing the window by 0.5 milliseconds generates large overlaps. This necessary to make the surface of the modulus of the time-frequency matrix as smooth as possible (better temporal resolution). The length of the heart-beat cycles varied from case to case, the average length was about 2000 samples.

A Gaussian smoothing mask as described in Eq 5.5 was used to suppress the small fluctuations on the surface of the STFT plane. The time domain, frequency domain, and 3-D plot of the Gaussian mask are shown in Figure 6.7.

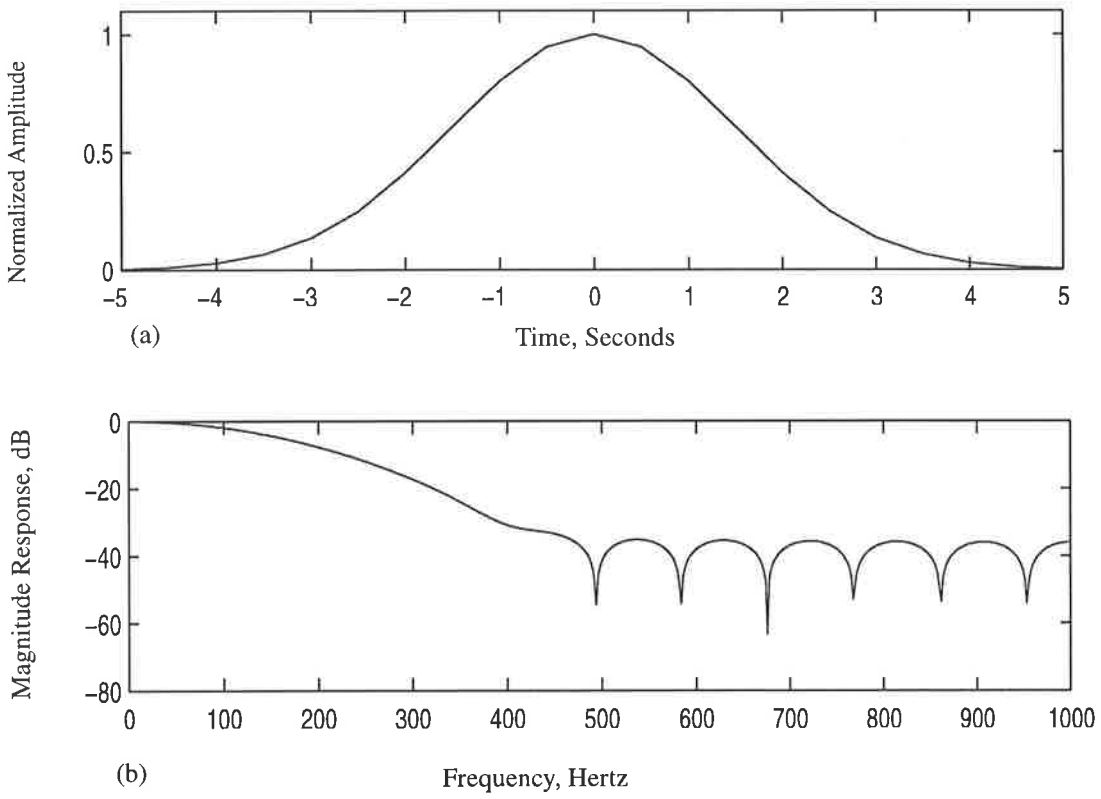
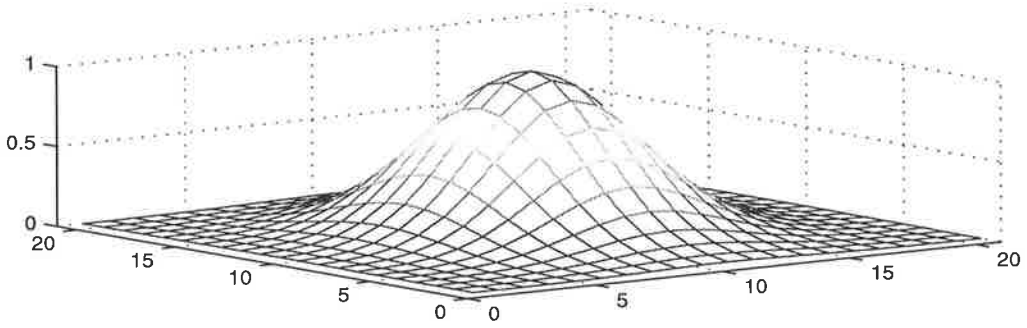


Figure 6.7 The Gaussian mask used for smoothing the STFT matrix of the signals of Figure 6.3 (b), a) time domain, b) frequency domain.



(c)

Figure 6.7 Continued c) 3_D representation

Parts of the modulus of the STFT matrix for signals of Figure 6.3 (b) and (d) are shown in Figure 6.8-1, -2, and -3. On each page eight three dimensional plots of the modulus of the STFT matrix are shown. The reason for dividing the modulus of the STFT matrix to many different sections stems from the fact that the energy of the signal varies at different locations of the time-frequency plane. Had we showed only one single plot for the entire STFT matrix, then the highest amplitude ridge would dominate and low level ridges would not be recognizable. This is clear from the amplitudes of the plots.

Figure 6.8-1 -2, -3 are shown on three pages and each page contains eight plots. Each group of four plots (two per page) divides the STFT matrix into four sections in the entire time axis and in a certain band of frequencies. The organization of Figure 6.8-1, 2, 3 is explained pictorially in Figure 6.9, which shows the time and frequency allocations of each plot of Figure 6.8-1, -2, and -3 in the STFT time-frequency matrix. For example, the term "plot #1" in the bottom left-hand corner of Figure 6.9 refers to 3-D plot shown in top left-hand corner of Figure 6.8-1.

6.5 Analysis Method for Time-Frequency Plane

The local maxima detector algorithm of Chapter 5 is applied to the STFT time-frequency matrix of signals of Figure 6.3 (b) and (d) and locations of concentration of energy are obtained. The result of the local maxima detector algorithm consists of three vectors: time, frequency, and magnitude vectors. The time vector, $\hat{T}(k)$, contains the temporal locations of the detected local maxima, the frequency vector, $\hat{F}(k)$, holds the frequencies of the detected local maxima, and the magnitude vector, $\hat{A}(k)$, contains

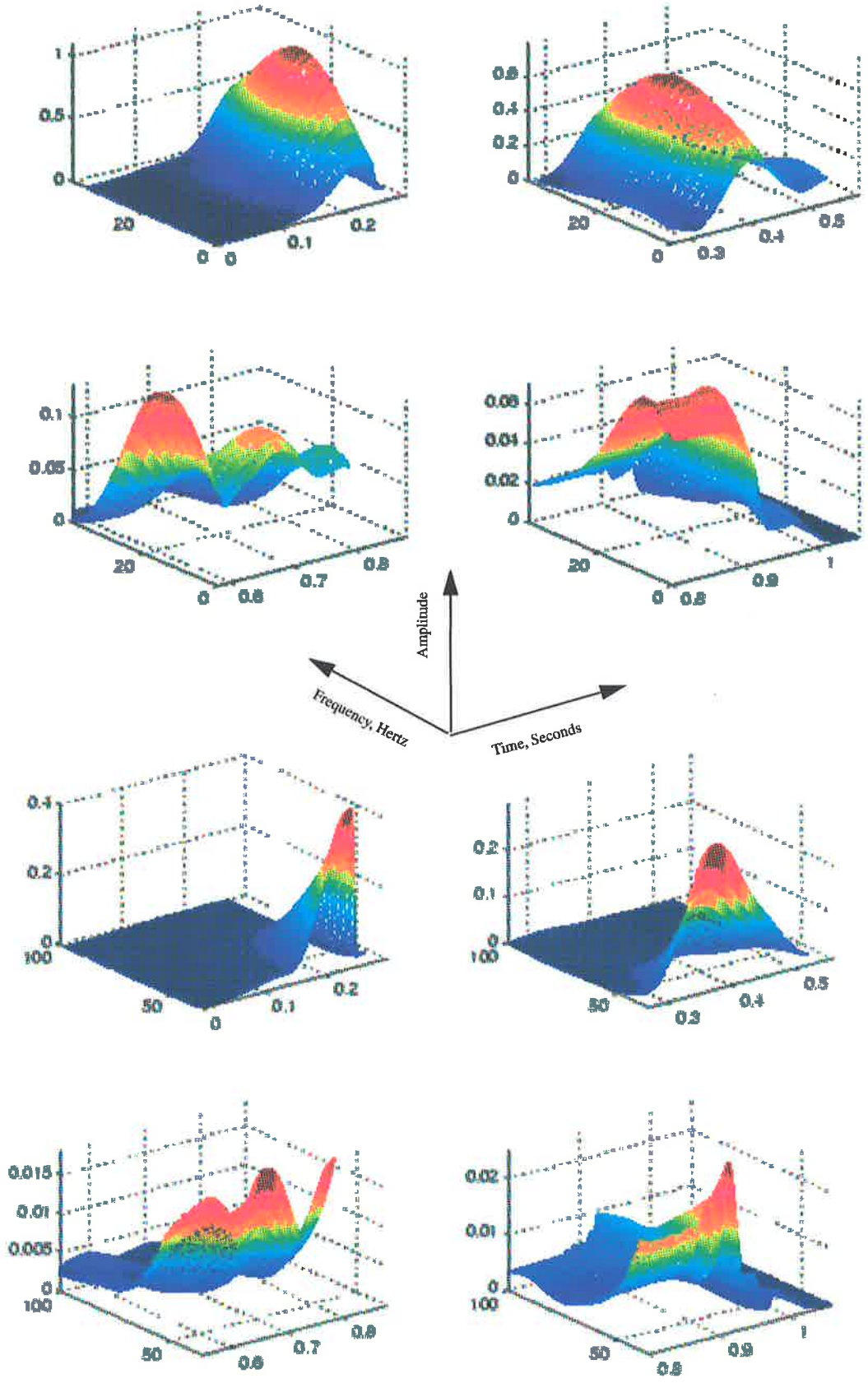


Figure 6.8-1 Time-frequency modulus of STFT of Figure 6.3 (b).

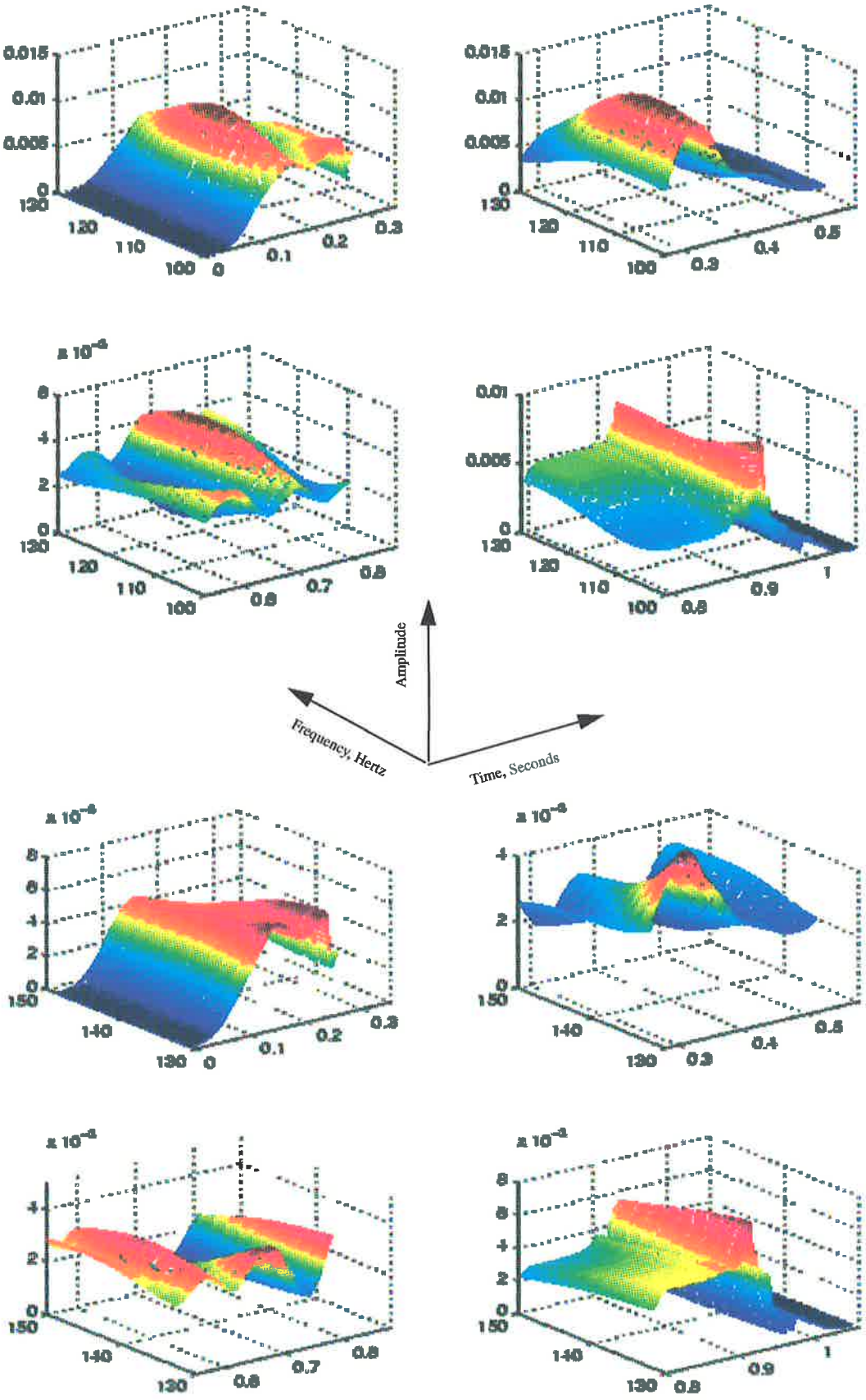


Figure 6.8-2 Continuation of Figure 6.8-1.

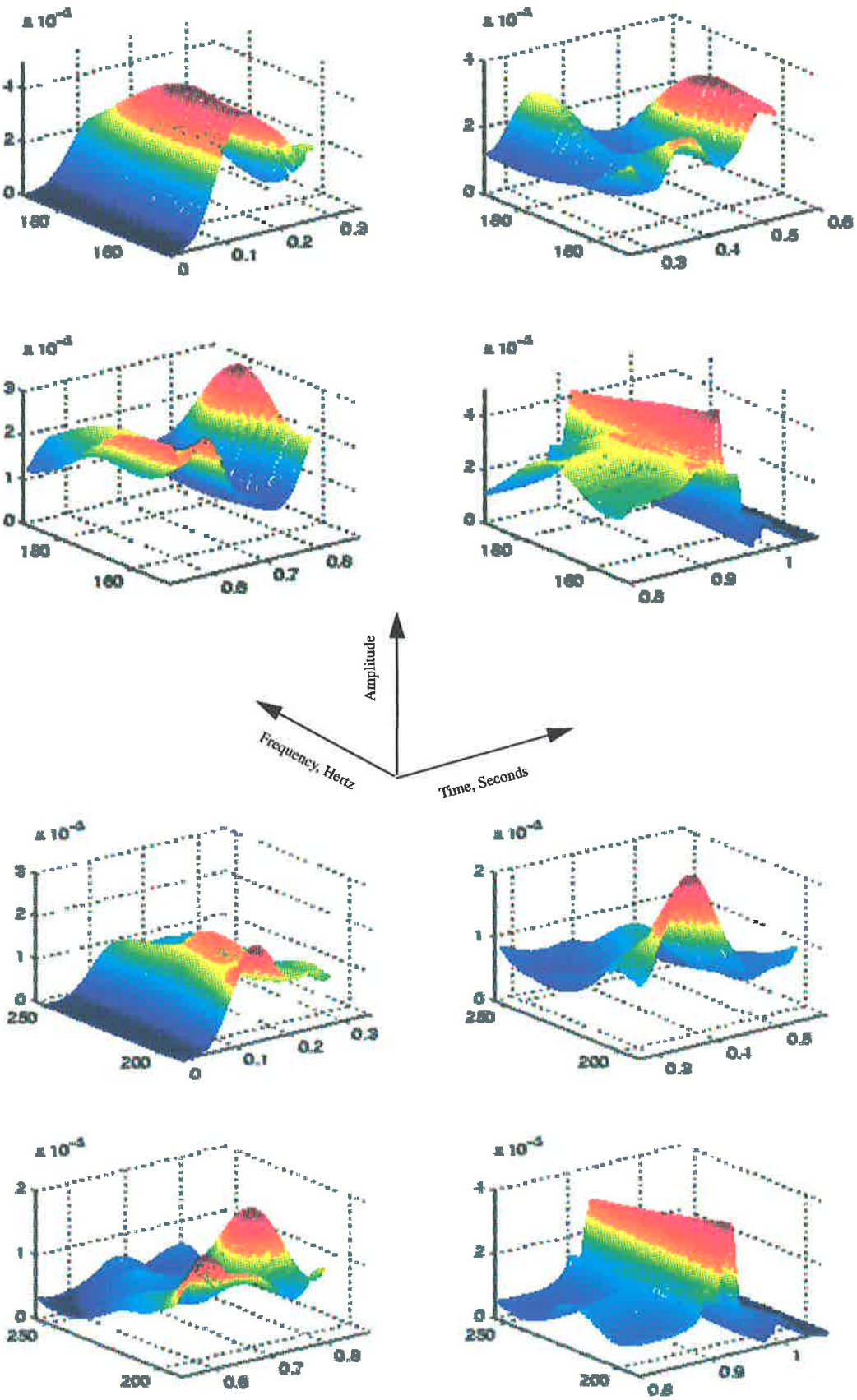


Figure 6.8-3 Continuation of Figure 6.8-2.

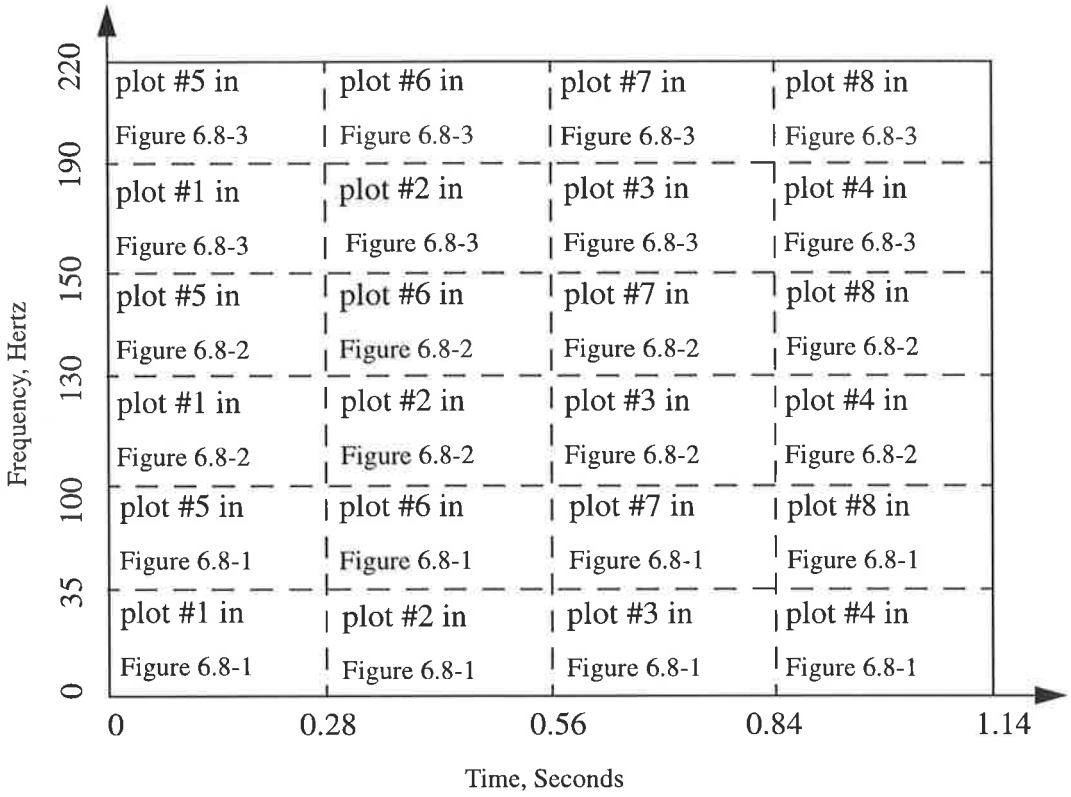


Figure 6.9 Organization of Figure 6.8-1, -2, and -3.

the magnitudes of the detected local maxima. Note that detected points represent the peaks of the local maxima in the modulus of the time-frequency plane in the STFT matrix. The results are shown in Figure 6.10 with time domain heart-beat cycles superimposed to emphasize the temporal locations of the concentration of energy relative to the heart-beat cycle. Figure 6.10 (a) shows the results for before angioplasty and Figure 6.10 (b) shows the results for after angioplasty. In these figures the vertical dotted lines split the temporal axis into first, second, third, and fourth heart sound regions. These lines are plotted with reference to the peak of the QRS complex in ECG signal. Note that the vertical axis is the frequency axis and hence means nothing to the amplitudes of the heart-beat cycles. STFT calculations are carried out with the heart-beat signals normalised in amplitude.

The prime intention here is to find any similarities and/or differences in the locations of the local maxima in the STFT planes before and after angioplasty. That is, we want to find the difference in the locations of the concentration of energy in the time-frequency plane. Time-frequency plane could reveal similarities and/or differences in the frequency contents of the first, second, third, and fourth heart sounds. As

was stated earlier there are three numbers associated with each local maximum point, that is its temporal location, spectral location, and magnitude. They completely define the spread of concentration of energy within the STFT plane. In other words, the three numbers define a local maximum as we view it as a ridge in a three-dimensional space.

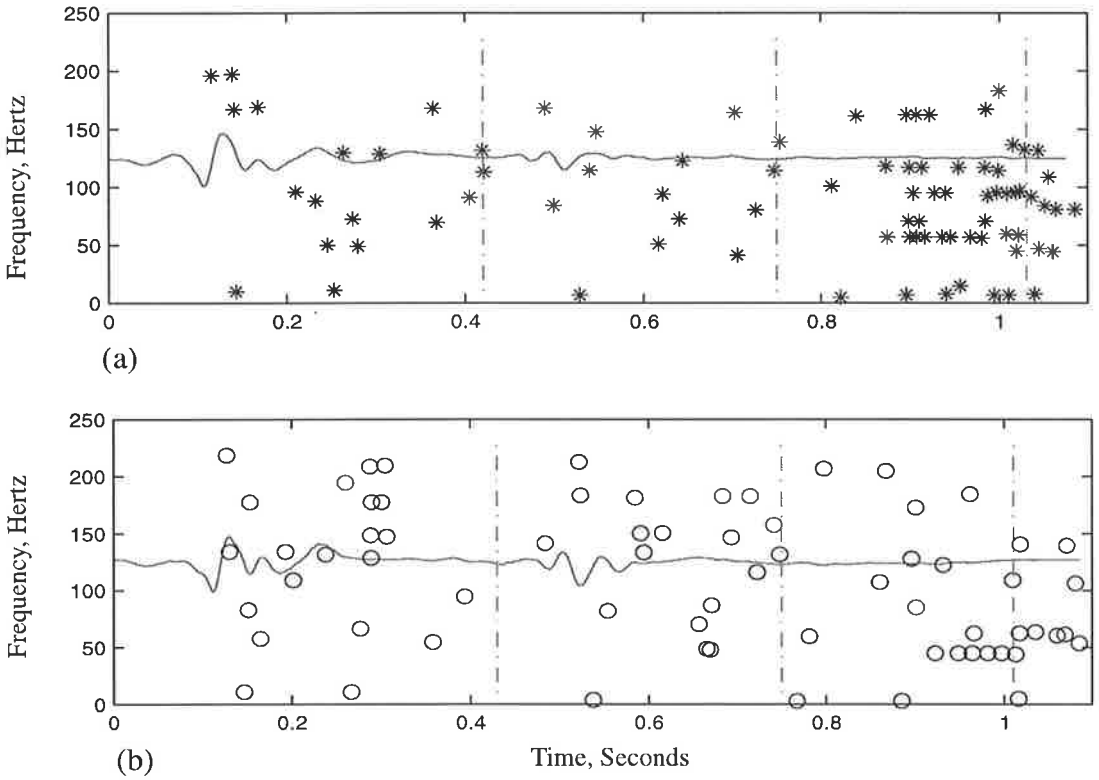


Figure 6.10 Locations of local maxima of a) signal of Figure 6.3 (b), b) signal of Figure 6.3 (d) with heart-beat cycles superimposed.

As was explained in Chapter 5, the local maxima detector algorithm works on the basis of amplitude. That is, the ridges on the time-frequency plane are detected starting from the highest amplitude ridge and moving down to lower amplitudes. Therefore; $\hat{T}, \hat{F}$, and $\hat{A}$ are sorted in descending order with respect to the amplitudes of the local maxima. That is, they are expressed as

$$\hat{A} = \left[\hat{a}_1\ \hat{a}_2\ \ldots\ \hat{a}_L\right] \qquad \text{Eq 6.13}$$

$$\hat{T} = \left[\hat{t}_1\ \hat{t}_2\ \ldots\ \hat{t}_L\right] \qquad \text{Eq 6.14}$$

$$\hat{F} = \left[\hat{f}_1\ \hat{f}_2\ \ldots\ \hat{f}_L\right] \qquad \text{Eq 6.15}$$

where $\hat{a}_1 > \hat{a}_2 > \ldots > \hat{a}_L$, and L is the number of detected local maxima. Note that the elements of vectors defined by Eq 6.14 and Eq 6.15 are not sorted in ascending or descending order. The first detected local maximum is defined by $\hat{a}_1$, $\hat{t}_1$, and $\hat{f}_1$, the second detected local maximum is defined by $\hat{a}_2$, $\hat{t}_2$, and $\hat{f}_2$, and so on.

Since we are seeking similarities in frequency content of the signals associated with first, second, and third heart sounds, we sorted the temporal, spectral, and magnitude vectors in ascending order of frequency. That is, we rearrange the vector $\hat{F}$ such that its first element is the lowest number in the vector and its last element is the highest number in the vector. That is the new frequency vector is

$$F = \left[f_1 \; f_2 \; \cdots \; f_L \right] \qquad \text{Eq 6.16}$$

where $f_1 < f_2 < \ldots < f_L$. The other two vectors, $\hat{T}$, and $\hat{A}$ are rearranged in accordance with F.

We define $\chi(k)$ as the product of temporal, spectral, and magnitude vectors of the local maxima, and call it the time-frequency-magnitude (TFM) product which is expressed as

$$\chi(k) = T(k)F(k)A(k) \qquad \text{for k=1, 2, ..., L} \qquad \text{Eq 6.17}$$

For example if the i^{th} local maximum occurs at $t = 0.5$ seconds and $f = 10$ Hz, and its magnitude $m = 20$, then $\chi(i) = (0.5)(10)(20) = 100$. The TFM product vector carries information about the temporal, spectral, and energy levels of the local maxima in the time-frequency plane. This can be regarded as a mapping of the three-dimensional space (time, frequency, and amplitude) into a one-dimensional space.

Note that the intention is to investigate the similarities of the time-frequency of the heart sounds before and after angioplasty. We do not expect to find a drastic change in the time-frequency characteristics of the signals after angioplasty operation. If there is any change at all, it is either in frequency, or may be a small change in the temporal locations of the events occurring in the heart, or it could be in the strength of the heart sound, or any combinations of those. This may result in a change in the TFM product

after angioplasty, but the general resemblance before and after angioplasty should remain. In order to show this an example is given.

- ***Example:***

Let x(t) be a signal composed of 5 sinusoids as shown in Figure 6.11 (a). We calculated the STFT matrix for this signal and its local maxima were found. The local maxima for this figure is shown by "$*$" in Figure 6.12. We change the signal of Figure 6.11 (a) in amplitude, frequency and temporal position of the its components. This is shown in Figure 6.11 (b), where the local maxima of its STFT matrix is shown by "o" in Figure 6.12. As Figure 6.12 shows, there are some differences in the locations of concentration of energy of signals of Figure 6.11 in the time-frequency plane. Note that, the amplitudes of the signal of Figure 6.11 (b) is different from that of Figure 6.11 (a), therefore the energy levels of local maxima must be different as well.

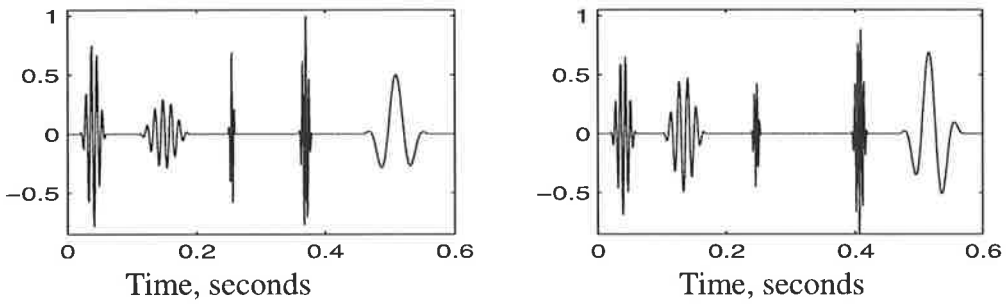


Figure 6.11 Signals for the example of TFM

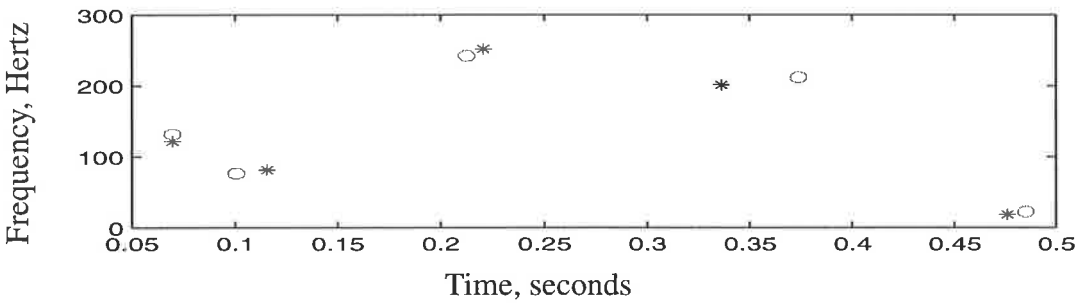


Figure 6.12 Local Maxima of signal of Figure 6.11.

The TFM product curves of the these signals are shown in Figure 6.13. As this figure shows, the TFM product curves of signals of Figure 6.11 (a) and (b) are similar although two signals have some differences.

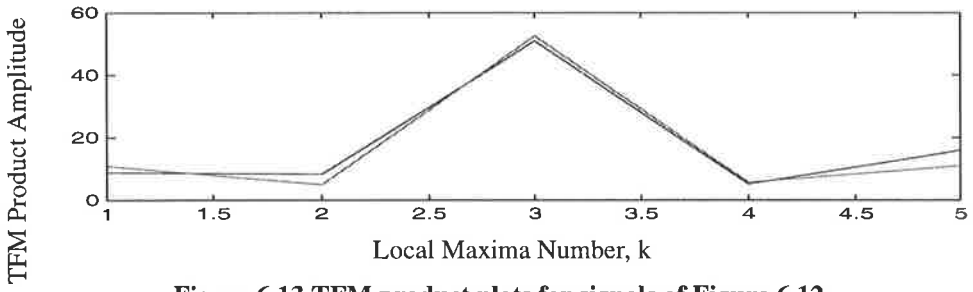


Figure 6.13 TFM product plots for signals of Figure 6.12.

We calculated the TFM product vectors for the heart sound signals before and after angioplasty for all data records. Their similarities were investigated by calculating their correlation coefficients, see Eq 4.12. The results showed more than 80% correlation for 85% of the signals. To illustrate the point, the time-frequency-magnitude product vectors of the time-frequency planes of Figure 6.10 are shown in Figure 6.14. Inspection of the graphs of this figure reveals that there are some similarities between them. In other words, the general shapes of the two waveforms seem to possess some similar variations in amplitude. The horizontal axis in this figure represents the number of local maxima in the vectors T, F, and A .

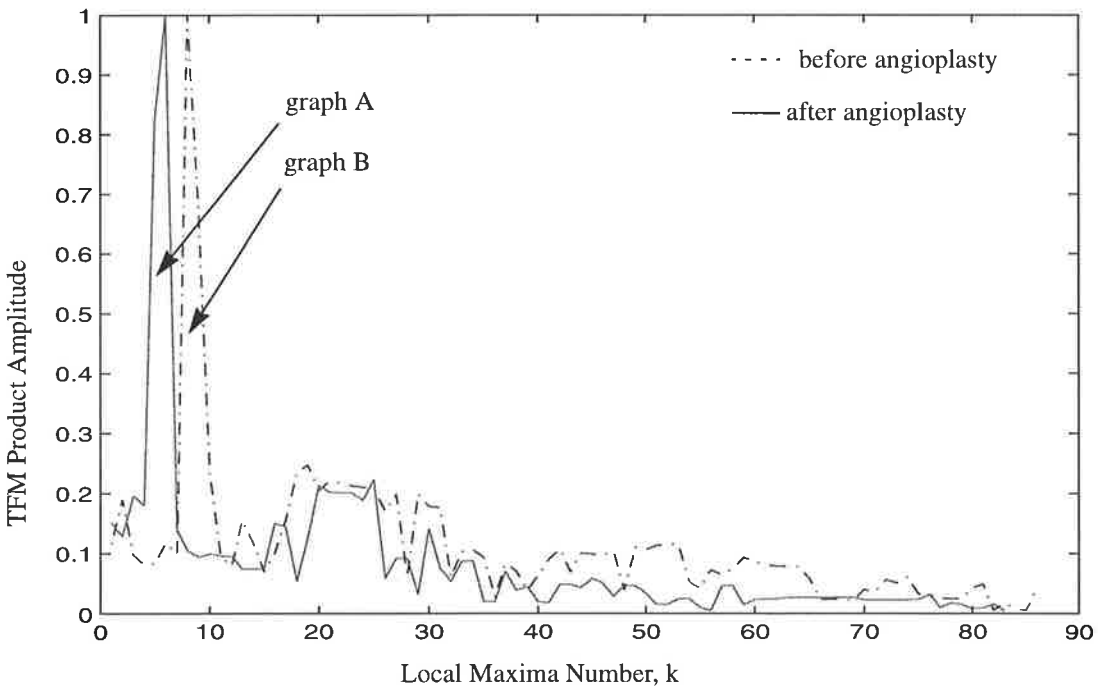


Figure 6.14 Time-Frequency-Magnitude product plots for Figure 6.10.

In Figure 6.14, the two high peaks at the beginning of the graphs indicate that there is a low frequency high energy local maxima in both signals, which corresponds to the high amplitude first heart sound signal in the heart-beat cycle. Close examination of graphs of Figure 6.14 shows that, they possess similar shapes. For example, consider part of graph A corresponding the first to the 7^{th} local maxima (from 1 to 7 in horizontal axis) with part of graph B corresponding the first to the 10^{th} local maxima (from 1 to 10 in horizontal axis). These parts are shown in Figure 6.15 (a). The general shapes of these two parts are similar, that is some similar variations in the amplitudes of the two graphs are observed, e.g. they both possess a high energy local maximum, etc. It is necessary to note that their horizontal length is not equal, in other words they represent unequal number of local maxima, therefore they cannot have exact resemblances. These pattern are not distinguishable by examining the entire local maxima plots shown in Figure 6.10. This suggests that our proposed TFM product technique may be very useful in this regard.

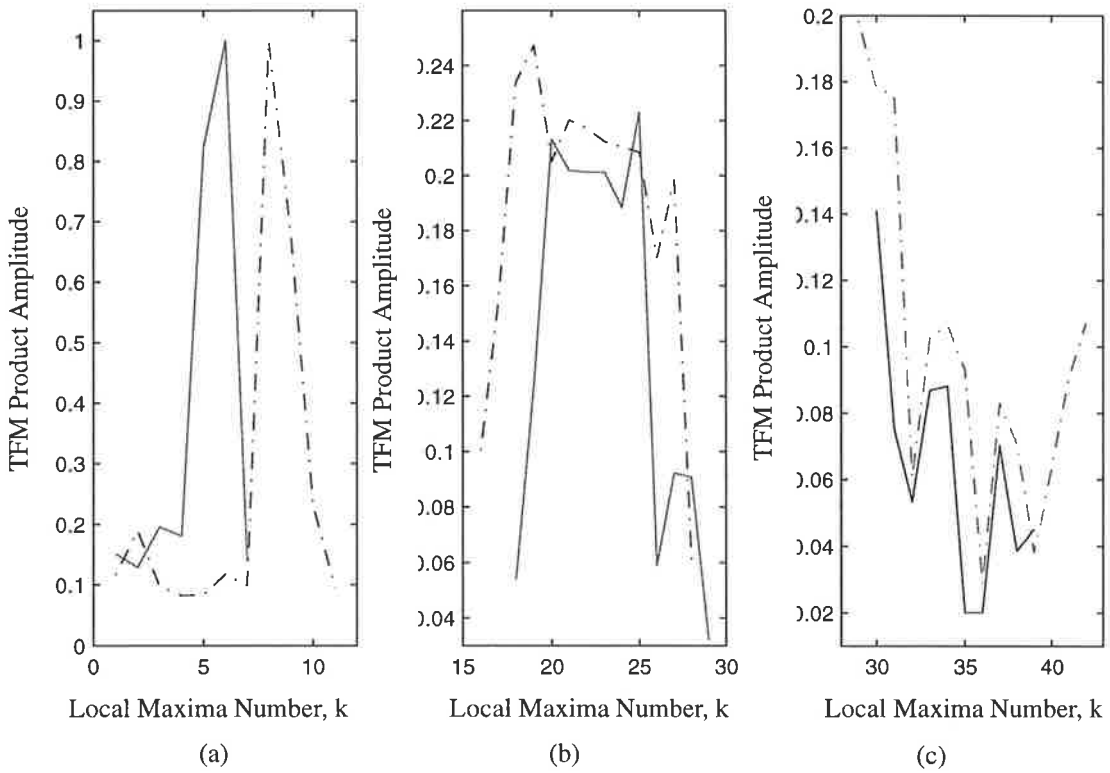


Figure 6.15 Three similar regions from Figure 6.14.

In Figure 6.15 (b) and (c), other similar parts from Figure 6.14 are shown. These similarities imply that there must be some similar local maxima patterns in the time-frequency planes. That is, the locations of the concentration of energy in the time-

frequency planes of the heart-beat signals of Figure 6.4 (a) and (b) must be somehow similar. Let $\Gamma_b(\Delta f, \Delta t)$ represent a region in the STFT plane of before angioplasty occupied by frequency band Δf and time allocation Δt, and $\Gamma_a(\Delta f, \Delta t)$ represent the similar region in the STFT plane of after angioplasty. If parts of graphs B and A of Figure 6.14 corresponding to $\Gamma_b(\Delta f, \Delta t)$ and $\Gamma_a(\Delta f, \Delta t)$ possess similar shapes then the spread of local maxima in these regions would be similar. Similarities become more evident when we examine the time-frequency matrices in those sections. For instance, plot of Figure 6.15 (a) shows that, ten lowest frequency local maxima before angioplasty must have similar time-frequency arrangement as first seven lowest frequency local maxima after angioplasty.

We call a regional arrangement of local maxima in the frequency band Δf and the time allocation Δt in time-frequency plane, a pattern. Similar patterns could be extracted from the time-frequency planes of before and after angioplasty with regards to the corresponding TFM product graphs. Three similar patterns corresponding to graphs of Figure 6.15 are shown in Figure 6.16.

Figure 6.16 shows that there are distinct time-frequency patterns which are quite similar for the before and after angioplasty signals of Figure 6.4 (a) and (b).

Using Figure 6.14, we have extracted all patterns and tabulated them in Table 6.2. The numbers in the table represent the local maxima number k. For example, the first similar patterns in the STFT matrices of signals of Figure 6.14, as was pointed out earlier, corresponds to first 10 local maxima for before angioplasty and first 7 local maxima for after angioplasty which are shown as pattern #1 in first two columns of first row in Table 6.2.

Table 6.2 Similar pattern of Figure 6.14.

#	Local maxima		#	Local maxima		#	Local maxima	
	before	**after**		**before**	**after**		**before**	**after**
1	1-10	1-7	2	11-15	8-18	3	16-28	19-29
4	29-42	30-40	5	43-57	41-56	6	58-75	57-76
7	76-86	77-83						

#: Pattern number

The patterns shown in Figure 6.16 are associated with the parts of Figure 6.14 that are shown in Figure 6.15. These patterns correspond to the patterns 1, 3, and 4 of Table 6.2. Other patterns that are listed in Table 6.2, are plotted in Figure 6.18. Figure 6.17 shows parts of the TFM plots of Figure 6.14 corresponding to the patterns of Figure 6.16. Note that any of the adjacent patterns of Figure 6.18 can be combined to make a larger pattern, and also any pattern can be divided into smaller patterns.

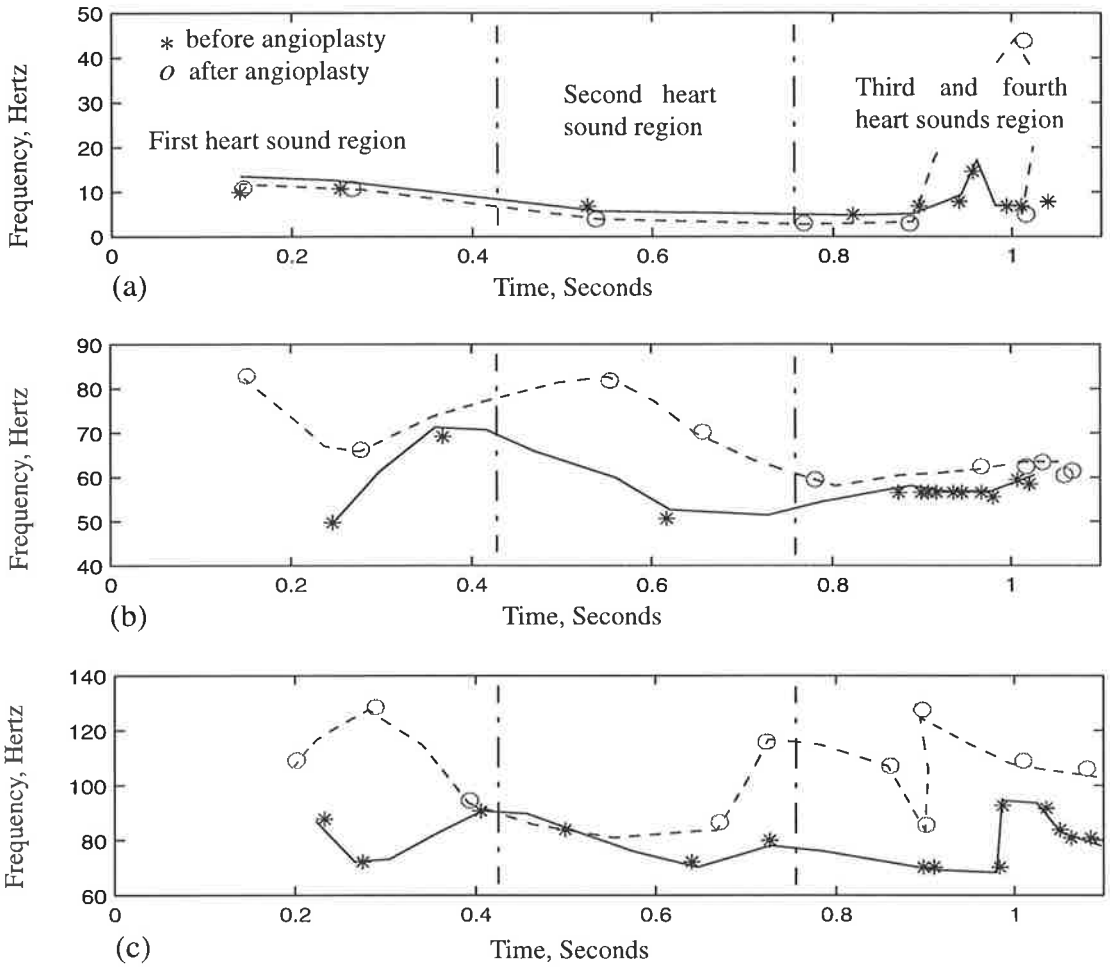


Figure 6.16 Time-frequency patterns corresponding to the TFMs of Figure 6.15.

In Figure 6.16 and Figure 6.18 "$*$" represents the local maxima before angioplasty and "o" represents the local maxima after angioplasty. In these figures, the local maxima points are connected by dotted lines to emphasize on the similarity of the patterns and they have no other meanings. As it is shown in Figure 6.15 and Figure 6.18, seven patterns were recognised up to 250 *Hz*. This was done by comparing parts of the TFM product graphs of Figure 6.14. We started by comparing parts of two graphs corresponding to local maxima 1 to 15 (numbers on the horizontal axis). It was

revealed that the most similar parts correspond to1 to 7 from graph *A* with 1 to 10 from graph B. This process was continued along the x-axis until all similar parts were recognised. For some of the cases ofTable 2.2, some patterns were detected beyond 250 *Hz* but this was not common to all signals; therefore, we restricted the analysis to the frequency range from zero to 250 *Hz*.

Examination of the time-frequency patterns for before and after angioplasty reveals that there is a frequency shift between similar patterns. That is, the average frequency in patterns of after angioplasty is higher than before angioplasty. The average frequency is calculated as the mean of the frequencies of the local maxima in a certain region. Therefore, the average frequency of the first heart sound before angioplasty in pattern #1 is the mean of the two frequencies, Figure 6.16 (a). The average frequencies of the patterns in the first, second, and third heart sound regions are calculated. The average frequency for each pattern is also calculated and the frequency difference between the average frequencies of the two patterns is found. The results are tabulated in Table 6.3.

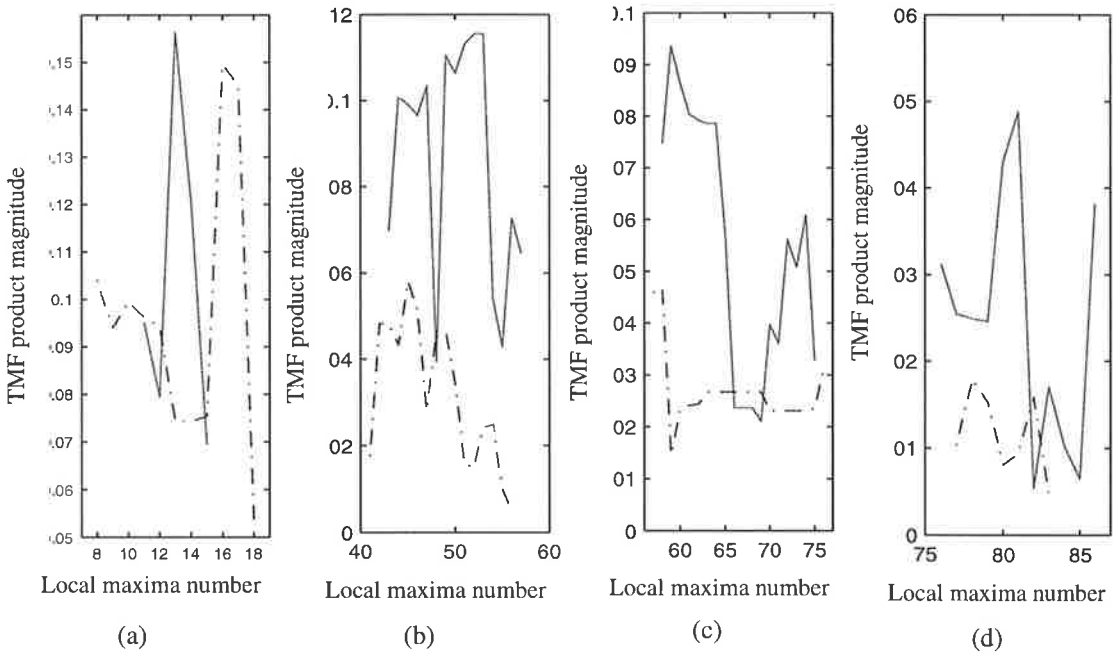


Figure 6.17 More sections of the TFMs of Figure 6.14.

In Table 6.3, the first column represents band of frequencies where patterns of signals in Figure 6.10 were identified. The 2nd and 3rd columns are average frequencies

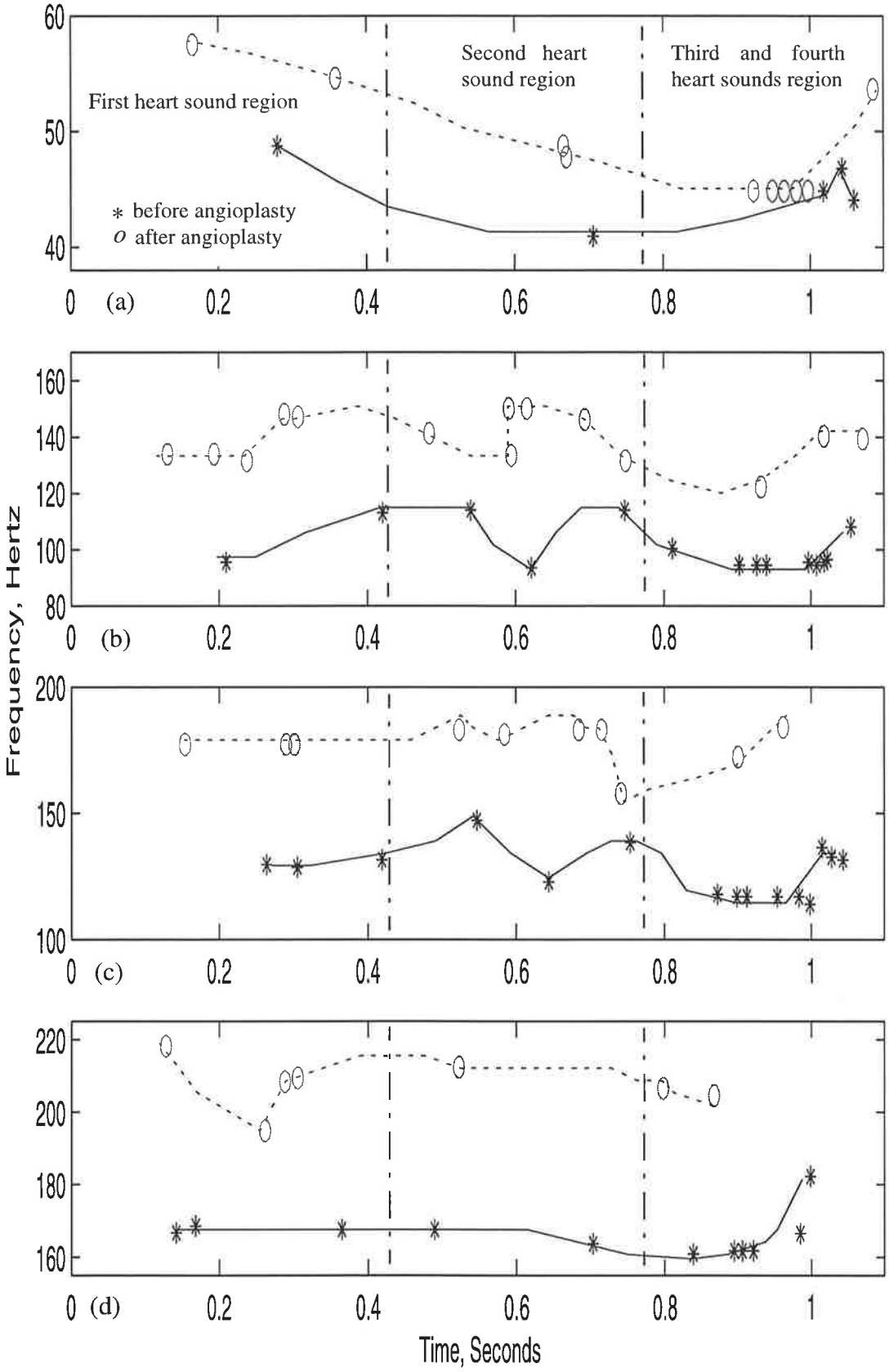


Figure 6.18 Patterns corresponding to the TFMs of Figure 6.17.

Table 6.3

Frequency Band (Hz)	1st HS MF		2nd HS MF		3rd HS MF		HS pattern MF		
	before	after	before	after	before	after	before	after	shift
0-40	11	11	7	4	7	14	8	10	2
40-60	49	56	41	48	45	46	45	50	5
40-90	59	75	51	76	57	62	56	71	15
60-140	83	110	78	101	80	107	80	106	26
80-155	104	139	104	143	99	133	102	138	36
110-185	129	177	135	181	123	178	129	179	50
160-250	168	208	166	212	166	205	167	209	42

HS: heart sounds

MF: mean frequency

of the patterns in the first heart sound regions before and after angioplasty, respectively. Likewise, the 4th and 5th columns correspond to the second heart sound regions before and after angioplasty, respectively. The 6th and 7th columns correspond to the third heart sound regions before and after angioplasty, respectively. The 8th and 9th columns show the average frequencies of the patterns before and after angioplasty, respectively. The last column shows the difference between the average frequencies of the patterns. For example, in the frequency band 0-40 *Hz* the first entry of the 1st column is 11 *Hz*, that indicates the average frequency of the first heart sound in that pattern before angioplasty. Likewise, the first entry under "HS pattern MF, before" is 8, which is the average of 11, 7, and 7. This shows that the average frequency of the whole pattern before angioplasty is 8 *Hz*. The averages were rounded off to their closest integers. The 2nd, 3rd, 4th, 5th, and 6th columns of Table 6.3 are plotted against the average frequencies of the patterns in Figure 6.19.

The frequency shift after angioplasty is clearly evident in Figure 6.19. The shift in frequency increases with the frequency. It is also interesting to note that the average frequencies of first, second, and third heart sounds are of the same order before and

after angioplasty. For example, for pattern #4 in Figure 6.19, the average frequencies before angioplasty are 83 *Hz*, 78 *Hz*, and 80 *Hz;* respectively. Likewise for pattern #4, the average frequencies after angioplasty are 110 *Hz*, 101 *Hz*, and 107 *Hz;* respectively.

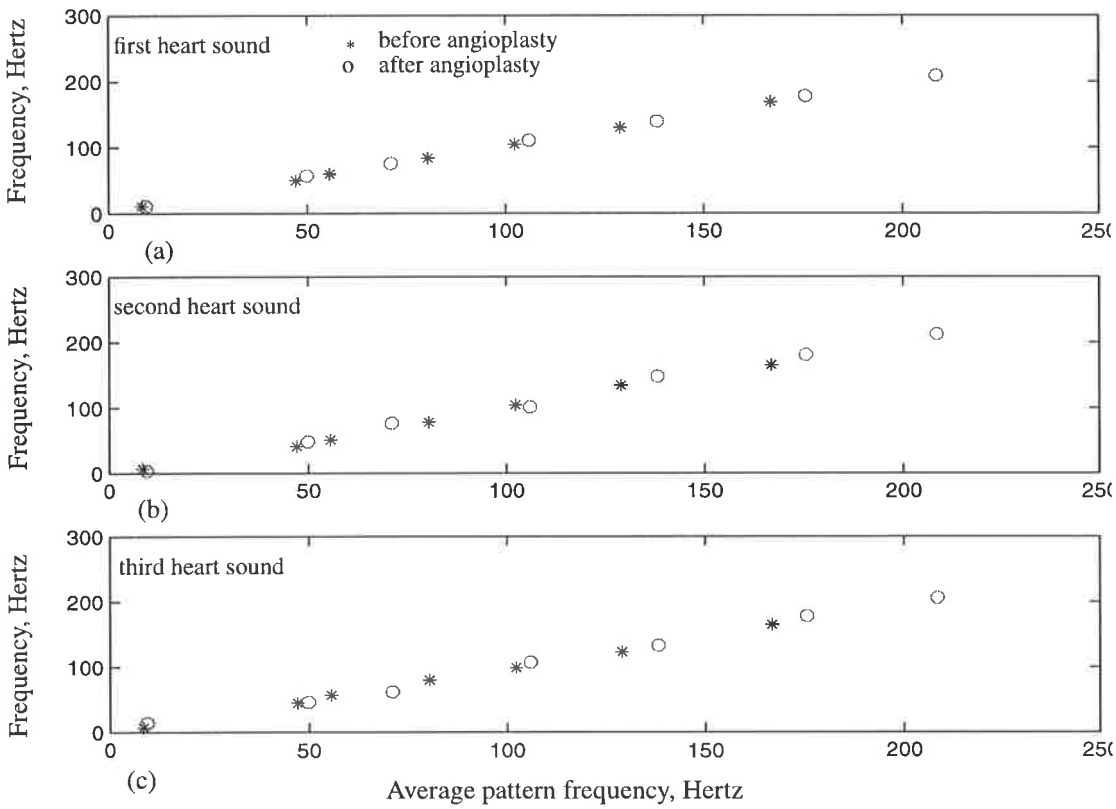


Figure 6.19 Average frequencies in a) 1[sr], b) 2nd, and c) 3rd heart sounds obtained from Table 6.2.

6.5.1 Analysis of STFT local maxima for all cases

The analysis method of section 6.5 is carried out for all of the recorded heart sounds of Table 2.2. The results for ten more cases are shown in Table 6.4. It was found that in about 85% of all cases there was a clear shift in frequency level of the time-frequency patterns after angioplasty. We have grouped the patterns of before and after angioplasty in four frequency bands according to their average pattern frequencies, as shown in Table 6.5. The frequency band designated as *below 50 Hz* includes all patterns whose average frequencies were lower than 50 *Hz*. Likewise the frequency band *50 -100 Hz* includes all patterns whose average frequencies were between 50 *Hz* and 100 *Hz*. The other two categories include patterns whose average frequencies were between 100 *Hz* and 150 *Hz*, and between150 *Hz* and 250 *Hz*. This table shows the

Table 6.4

Case	Band	1st HS MF		2nd HS MF		3rd & 4th HS MF		HS pattern MF		
		before	after	before	after	before	after	before	after	shift
# 1	0-30	11	12	5	4	7	10	8	9	1
	31-70	39	64	59	44	58	61	52	56	4
	65-95	82	78	83	82	79	80	81	80	-1
	100-130	107	113	104	107	112	110	108	110	2
	120-150	141	142	125	134	138	145	135	140	5
	155-195	172	190	161	171	160	173	164	179	15
	190-240	202	230	196	230	---	237	199	232	33
# 2	0-40	8	8	6	13	8	12	7	11	4
	30-75	48	62	44	70	51	57	48	63	15
	66-115	69	81	65	78	88	103	74	87	13
	95-155	128	135	124	130	131	149	127	138	11
	146-190	158	174	150	169	162	183	157	175	18
	180-250	201	234	211	238	218	246	210	239	29
# 3	0-35	6	5	5	5	9	11	7	7	0
	40-65	48	54	43	51	55	61	49	55	6
	55-105	69	81	64	76	75	89	69	80	11
	96-130	101	117	107	118	112	120	107	118	11
	120-160	137	152	129	135	144	154	137	147	10
	151-190	165	175	161	165	167	186	164	175	11
	190-230	198	218	195	204	201	227	198	216	18
	210-250	---	---	---	---	201	232	201	232	31
# 4	0-50	8	6	3	4	17	21	9	10	1
	35-85	60	72	47	59	46	61	51	64	13
	75-115	95	109	87	88	94	104	92	108	16
	100-170	125	138	126	127	123	139	125	135	10
	140-200	150	185	150	174	151	170	150	176	26
	160-250	199	234	193	240	199	216	197	230	33
# 5	0-45	7	10	5	6	14	25	9	14	5
	40-115	77	82	76	103	61	63	71	83	12
	90-160	116	142	97	145	113	149	109	145	36
	130-200	170	186	147	194	165	168	161	183	22
	190-235	210	228	205	220	198	213	204	220	16

Table 6.4

Case	Band	1st HS MF		2nd HS MF		3rd & 4th HS MF		HS pattern MF		
		before	after	before	after	before	after	before	after	shift
# 6	0-50	9	10	5	8	15	21	10	13	3
	20-120	67	85	55	90	61	83	61	86	25
	70-170	89	127	110	159	103	154	101	147	46
	120-250	149	195	141	206	167	223	152	208	56
# 7	0-35	7	6	5	4	10	14	7	8	1
	40-80	65	74	61	70	68	78	65	74	9
	75-120	96	109	84	98	113	113	96	108	12
	100-155	121	134	112	126	128	141	120	134	14
	140-185	159	171	161	174	171	179	164	175	11
	170-210	173	198	170	176	181	203	175	196	21
	200-250	---	---	201	230	207	235	204	233	29
# 8	0-40	6	7	4	5	9	12	6	8	2
	35-90	52	61	45	55	56	65	51	60	9
	60-100	66	82	65	90	72	92	68	88	20
	85-140	98	116	88	110	102	118	96	115	19
	130-190	149	166	148	165	155	178	147	170	23
	180-230	185	221	182	218	192	231	186	223	37
# 9	0-35	7	6	4	5	11	14	7	8	1
	30-60	52	53	57	57	47	53	52	54	2
	60-110	80	97	87	94	73	88	80	93	13
	100-145	107	130	112	141	104	128	108	133	25
	110-200	131	170	142	151	138	173	137	165	28
	140-250	172	212	168	206	166	206	169	208	39
# 10	0-45	8	9	5	7	11	16	8	11	3
	30-80	56	67	51	62	49	54	52	61	9
	70-110	81	93	76	85	88	92	82	90	8
	95-160	118	135	98	121	126	145	114	134	20
	140-190	159	178	162	174	171	182	164	178	14
	180-240	201	228	192	224	216	231	203	228	25

highest, lowest and mean average frequencies found in the patterns. For example consider the frequency band *below 50 Hz*, it is shown that the minimum average pattern frequency in the first heart sound region in all of the signals was found to be 6 *Hz*

before angioplasty and 6 *Hz* after angioplasty. Likewise, the maximum average pattern frequency in the first heart sound region was found to be 35 *Hz* and 45 *Hz* before and after angioplasty, respectively. The frequency shift for the frequency band *below 50 Hz* was found to be in the order of 5 *Hz*.

Table 6.5 Average frequencies before angioplasty and after angioplasty.

Frequency bands	Average pattern frequency in 1st heart sound		Average pattern frequency in 2nd heart sound		Average pattern frequency in 3rd heart sound		Shift in average frequency
	before	after	before	after	before	after	
Below 50 Hz	Min: 6	6	Min: 3	3	Min: 6	7	
	Max: 35	45	Max: 38	44	Max: 36	45	5
	Mn: 17	21	Mn: 18	23	Mn: 18	23	
50 Hz - 100 Hz	Min: 52	50	Min: 51	51	Min: 52	51	
	Max: 93	98	Max: 83	89	Max: 95	98	9
	Mn: 64	73	Mn: 58	68	Mn: 70	79	
100 Hz - 150 Hz	Min: 101	105	Min: 104	111	Min: 100	101	
	Max: 145	141	Max: 140	148	Max: 145	141	16
	Mn: 119	135	Mn: 122	137	Mn: 126	143	
150 Hz - 250 Hz	Min: 152	156	Min: 155	159	Min: 159	162	
	Max: 205	228	Max: 198	230	Max: 210	235	26
	Mn: 181	205	Mn: 171	197	Mn: 185	213	

Min: Minimum
Max: Maximum
Mn: Mean

6.6 Conclusions

In this chapter we briefly reviewed the short-time Fourier transform. We emphasized the importance of the window function in calculations of STFT, and highlighted the fact that it is not possible to achieve good resolution for both time and frequency. It was shown that for this application the Chebychev window is more suitable.

The heart-beat cycle of heart sounds before and after angioplasty are used and their STFT are calculated. The time-frequency matrices are smoothed by applying a two-dimensional Gaussian mask. The intention was to remove any sharp variations on the surface of the STFT plane.

The local maxima detector algorithms are used to determine the locations of concentration of energy in the STFT planes. By doing so, three vectors are obtained which define the temporal locations, spectral locations, and magnitudes of the local maxima. A method of analysis is proposed for comparison of the time-frequency behaviour of the heart-beat cycles before and after angioplasty. In this method we calculate time-frequency-magnitude product of local maxima and use it to identify time-frequency patterns in the time-frequency planes. It was shown that it could be very effective in finding the difference in the time-frequency representations of the heart sounds before and after angioplasty.

The time-frequency patterns are examined and the difference between the average frequencies of two similar patterns before and after angioplasty is calculated. It was revealed that there is a frequency shift after the angioplasty operation. Three frequency bands are considered: low, medium, and high. For the low frequency range, below 50 *Hz*, the average frequency shift is below 10 *Hz*. For the medium frequency range, between 50 *Hz* and 150 *Hz*, the average frequency shift is below 20 *Hz* and in the high frequency range, above 150 *Hz* and up to 250 *Hz*, the average frequency shift is above 20 *Hz*.

Chapter 7

Quadratic Time-Frequency Analysis of Heart Sounds

7.1 Introduction

Classical works of Wigner [136], Ville [132], and Gabor [47] led to an alternative way to look at signals with time dependent frequencies. Although, it may seem that the aim of this kind of analysis is to overcome the resolution restriction of the short-time Fourier transform, historically it is clear that the main motivation was to understand time varying spectra and their properties [31]. The resolution restriction of the short-time Fourier transform prevents its use in some applications. In 1932 Wigner [136], proposed another method of analysis for signals with time varying spectra. It was derived later by Ville in 1948 [132], and today it is called Wigner-Ville distribution (WVD). The ultimate goal in WVD analysis is to obtain a joint distribution function of time and frequency which describes the energy intensity distribution of the signal in the time-frequency plane. In other words, as with STFT, it is mapping a one-dimensional signal, which is function of time, into a two-dimensional signal, which is function of time and frequency in order to understand the time varying spectral behaviour of the signal.

WVD has a quadratic structure and it is an energetic interpretation of the signals. It does not use windowing to perform the transformation; therefore, there is no time or frequency resolution restriction associated with it. WVD is a density function which is a function of both time and frequency. Such a distribution function enables us to understand how frequencies are distributed over time, and how signal energy varies in the time-frequency plane.

The objective of this chapter is to study the energy distributions of the heart sounds in time-frequency planes using WVD before and after angioplasty. In section 7.2, the energy density function is reviewed. This is followed with an introduction to Wigner-Ville distributions in section 7.3. In section 7.4, the Cohen's class of distributions and their properties are stated very briefly. We will discuss the shortcomings of WVD for the analysis of the heart sound signals to extract the locations of concentrations of energy in the time-frequency plane. This leads us to discussion of Choi-Williams distributions (CWD) in section 7.6. We will use CWD to analyse the heart sounds. In order to reduce the computational load of the CWD, we have designed a method that uses the minimum possible number of multiplications to carry out the computation; this is explained in details in appendix B. In section 7.8, the CWD of heart sounds before and after angioplasty are calculated. Locations of the concentrations of energy in the time-frequency planes of CWD are extracted using the local maxima detector algorithm of Chapter 5. The time-frequency-magnitude product method which was proposed in Chapter 6 is used to identify the time-frequency patterns in the time-frequency planes of CWD. Patterns for a typical heart sound signal before and after angioplasty are shown and compared.

In some cases the WVD can become negative and therefore it is not considered a true distribution. In 1950's and 1960's joint time and frequency distributions were investigated from different perspectives. Some authors have derived joint time-frequency distribution functions to overcome the shortcomings of WVD [72], [82], [93], [103]. A general distribution function was introduced by Cohen [30] who introduced a kernel function into the WVD and showed that most of the joint time-frequency distribution functions actually differ from each other in the kernel function. The kernel function will be discussed later in this chapter. In fact WVD is a Cohen's class of distributions where its kernel is equal to one.

When a signal is composed of a sum of monocomponents, the inherent bilinear structure of Cohen's class of distributions causes undesirable terms which are called cross-terms [26], [58], [66]. In some applications such as pattern recognition these terms are very troublesome and sometimes might even obscure the desired auto-terms. Kernel functions determine the shape and the amount of the cross terms. Kernel functions are designed in order to reduce the effects of cross-terms considerably while keeping auto-terms unchanged. Cross-terms and auto-terms will be discussed later.

Various distributions with different kinds of kernel functions have emerged since Cohen's introduction of the kernel function. In [24], an exponential kernel function is introduced as a Cohen's class of distributions. It was shown that it satisfies all of the general properties of bilinear distributions. Its ability to suppress the cross-terms was demonstrated. For computational purposes a running window method was suggested. Reduced interference distributions were introduced in [64] that satisfy more distribution properties and associated kernel requirements.

In [54], a kernel function based on the Bessel function of the first kind of order one was proposed, and the desirable distribution properties of the Bessel kernel were shown. Its functionality as a time-frequency distribution function was tested by applying it to heart sounds and Doppler blood flow signals.

In [145], a cone-shaped kernel was introduced. The intention was to smooth the cross-term in the Wigner distributions as well as enhancing the spectral peaks and preserving the finite support property. Using frequency domain analysis, it was shown that the cone-shaped kernel has a low-pass characteristic with controlled bandwidth. It was demonstrated that this kernel can easily resolve close spectral peaks and maintain zero intervals of the signal.

Other derivations and approaches have been suggested for time-frequency distributions. Some of them have intended to arrive at a new time-frequency distribution and others have aimed to design a new kernel function [69], [81], [93], [103], [141].

7.2 Energy density function

Let $s(t)$ be a complex signal applied to a unit resistor, then the energy of the signal is

$$E = \int_{-\infty}^{\infty} |s(t)|^2 dt \qquad \text{Eq 7.1}$$

where $|s(t)|^2 = s(t)s^*(t)$, and "$*$" represents the complex conjugate. $|s(t)|^2$ is the intensity per unit time at time t, or the instantaneous power of the signal. Therefore, $|s(t)|^2 \Delta t$ is the fractional energy of the signal in the time interval Δt at time t.

Likewise in the frequency domain, the Fourier transform of the signal is

$$S(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} s(t) e^{-j\omega t} dt \qquad \text{Eq 7.2}$$

where $|S(\omega)|^2$ is the intensity per unit frequency at frequency ω, and $|S(\omega)|^2 \Delta\omega$ is the fractional energy in the frequency interval $\Delta\omega$ at frequency ω.

The quantities $|s(t)|^2$ and $|S(\omega)|^2$ are one-dimensional functions of time and frequency, respectively. The joint distribution function, $P(t, \omega)$, is defined to be the intensity of the signal at time t and frequency ω [31]. Therefore, $P(t, \omega)\Delta t \Delta\omega$ is the fractional energy in time interval Δt and frequency interval $\Delta\omega$ at time t and frequency ω. $P(t, \omega)$ must represent the instantaneous energy of the signal if the energy distributions for all frequencies are summed up at a particular time. It also must represent the energy density spectrum of the signal if the energy distributions over all times are summed up at a particular frequency. These conditions are expressed as

$$\int_{-\infty}^{\infty} P(t, \omega) d\omega = |s(t)|^2 \qquad \text{Eq 7.3}$$

$$\int_{-\infty}^{\infty} P(t, \omega) dt = |S(\omega)|^2 \qquad \text{Eq 7.4}$$

The total energy of the signal in terms of the distribution function is given as

$$E = \int_{-\infty}^{\infty} \int P(t, \omega) dt d\omega \qquad \text{Eq 7.5}$$

For Eq 7.5 to represent the total energy of the signal; in general, it is necessary that the time marginal, Eq 7.3, and the frequency marginal, Eq 7.4, be satisfied.

7.3 Wigner-Ville distribution function

The joint time-frequency distribution function given by Eq 7.6 was developed by Wigner [136] and Ville [132][1].

$$WV(t, \omega) = \frac{1}{2\pi}\int (s^*(t-\tau/2)e^{-j\omega\tau}s(t+\tau/2))d\tau \qquad \text{Eq 7.6}$$

7.3.1 Auto-terms and Cross-terms

Since energy is a quadratic function of the signal, if the signal $s(t)$ is a multi-component signal its WVD function will have two sets of distinct terms, called auto-terms and cross-terms [31], [24]. Let the signal be a multicomponent signal as

$$s(t) = \sum_{i=1}^{M} s_i(t) \qquad \text{Eq 7.7}$$

then its WVD is expressed as

$$WV(t, \omega) = \sum_{i=1}^{M} WV_i(t, \omega) + \sum_{\substack{i,k=1 \\ i \neq k}}^{M} WV_{ik}(t, \omega) \qquad \text{Eq 7.8}$$

where

$$WV_{ik}(t, \omega) = \frac{1}{4\pi^2}\iint e^{-j\varphi t - j\omega\tau} s^*_i(\varphi-\tau/2)s_k(\varphi+\tau/2)d\varphi d\tau \qquad \text{Eq 7.9}$$

Cross-terms are generated because of the bilinear nature of the distribution function; it is necessary to eliminate them in order to have a true representation of the energy distribution of the signal. Methods for suppressing the cross-terms will be discussed later. It can be shown that

1. Note that here after, in all equations integrals are taken from $-\infty$ to ∞, unless specified otherwise.

$$WV_{i,k}(t, \omega) = WV^*_{k,i}(t, \omega) \quad \text{Eq 7.10}$$

7.3.2 Marginals

Integrating Eq 7.6 with respect to time and frequency, it can be shown that the WVD satisfies the time and frequency marginals as

$$\int WV(t, \omega)d\omega = |s(t)|^2 \quad \text{Eq 7.11}$$

$$\int WV(t, \omega)dt = |S(\omega)|^2 \quad \text{Eq 7.12}$$

7.4 Properties of WVD

Some fundamental properties of the WVD are listed here; more detailed explanations can be found in [31], [27], [28], [29].

- *Reality*: Wigner-Ville distribution is real, that is

$$WV_s(t, \omega) = WV^*_s(t, \omega) \quad \text{Eq 7.13}$$

- *Uniqueness*: Wigner-Ville distribution is related to the signal up to a constant phase factor as

$$s(t) = \frac{1}{s^*(0)}\int WV_s(t/2, \omega)e^{jt\omega}d\omega \quad \text{Eq 7.14}$$

- *Instantaneous frequency*: It is defined to be the derivative of phase of the analytical signal [32]. If the signal is expressed as

$$s(t) = A(t)e^{j\varphi(t)} \quad \text{Eq 7.15}$$

then the instantaneous frequency can be calculated as

$$\Omega(t) = \frac{\partial}{\partial t}\varphi(t) \quad \text{Eq 7.16}$$

In the time-frequency analysis, at any given time the average frequency must equal the instantaneous frequency of the signal. That is

$$\Omega(t) = \frac{1}{|s(t)|^2}\int \omega WV(t, \omega)d\omega \qquad \text{Eq 7.17}$$

- *Group delay:* Group delay indicates how frequencies are delayed in time. If the spectrum of the signal is expressed as

$$S(\omega) = |S(\omega)|e^{j\psi(\omega)} \qquad \text{Eq 7.18}$$

then the group delay is defined as

$$T_G = \frac{\partial}{\partial\omega}\psi(\omega) \qquad \text{Eq 7.19}$$

In the time-frequency analysis, at any given frequency the average in the time direction (centre of gravity) must equal the group delay of the signal. That is

$$T_G = \frac{1}{|S(\omega)|^2}\int WV(t, \omega)tdt \qquad \text{Eq 7.20}$$

- *Time shift*: If the signal is translated in the time domain, then its WVD is also shifted in time. Let the shift in time domain be expressed as

$$x(t) = s(t + t_0) \qquad \text{Eq 7.21}$$

then

$$WV_x(t, \omega) = WV_s(t + t_0, \omega) \qquad \text{Eq 7.22}$$

- *Frequency shift*: Since translating a signal in the frequency domain is equivalent to multiplying it by $e^{j\Omega t}$ in the time domain; that is, if

$$X(\omega) = S(\omega + \Omega) \qquad \text{Eq 7.23}$$

then

$$x(t) = s(t)e^{j\Omega t} \qquad \text{Eq 7.24}$$

therefore, we get

$$WV_x(t, \omega) = WV_s(t, \omega - \Omega) \qquad \text{Eq 7.25}$$

- *Addition*: The additive property is actually depicted in Eq 7.10. For two signals

$$x(t) = s_1(t) + s_2(t) \quad \text{Eq 7.26}$$

$$WV_x(t, \omega) = WV_{s_1}(t, \omega) + WV_{s_2}(t, \omega) + WV_{s_1, s_2}(t, \omega) + WV_{s_2, s_1}(t, \omega) \quad \text{Eq 7.27}$$

- *Multiplication*: It is well known that multiplication in the time domain is equivalent to convolution in the frequency domain

$$x(t) = s_1(t)s_2(t) \quad \text{Eq 7.28}$$

$$X(\omega) = S_1(\omega) \otimes S_2(\omega) \quad \text{Eq 7.29}$$

This property is expressed as

$$WV_x(t, \omega) = WV_{s_1}(t, \omega) \otimes_{\omega} WV_{s_2}(t, \omega) \quad \text{Eq 7.30}$$

- *Convolution*: Convolution in the time domain is equivalent to multiplication in the frequency domain. That is

$$x(t) = s_1(t) \otimes s_2(t) \quad \text{Eq 7.31}$$

$$X(\omega) = S_1((\omega)S_2(\omega)) \quad \text{Eq 7.32}$$

This property is expressed as

$$WV(t, \omega) = WV_{s_1}(t, \omega) \otimes_{t} WV_{s_2}(t, \omega) \quad \text{Eq 7.33}$$

7.5 Cohen's class of distributions

It was realized by Cohen [30], that it is possible to generate a vast variety of distributions by modifying Eq 7.6. He proposed a general class of distribution functions by adding an additional term into Eq 7.6 as

$$P(t, \omega) = \frac{1}{4\pi^2}\iiint e^{-j(\theta t + \tau\omega - \theta u)} s^*\left(u - \frac{1}{2}\tau\right) s\left(u + \frac{1}{2}\tau\right) \Phi(\theta, \tau) du d\tau d\theta \quad \text{Eq 7.34}$$

where $\Phi(\theta, \tau)$ is an arbitrary function called the kernel function [29]. Originally it was defined to be a function of time and frequency only as in Eq 7.34, but later kernel functions with signal dependency were also developed. By choosing the kernel function, a specific distribution results. It is clear that the Wigner-Ville distribution is a

Cohen's class of distribution and its kernel function is equal to one, $\Phi(\theta, \tau) = 1$. Since the kernel function determines the properties of the distribution obtained by Eq 7.34, one can put certain constraints on the kernel function to get a particular property for the time-frequency distribution.

Many different kernel functions have been designed [24], [54], [69], [119], [145]. As stated above, the properties of the distributions are related to the properties of the kernel functions. In Table 7.1 some of the desirable properties of the kernel functions are listed [19] [31].

Table 7.1

Property	**Kernel constraints**
Frequency shift, $P_x(t, \omega) = P_y(t, \omega - \omega_0)$	Always satisfied
Time shift, $P_x(t, \omega) = P_y(t - t_0, \omega)$	Always satisfied
Time scaling, $P_x(t, \omega) = P_u\left(\alpha t, \frac{\omega}{\alpha}\right)$	$\Phi_x(\tau, \theta) = \Phi_y\left(\frac{\theta}{\alpha}, \alpha\tau\right)$
Frequency scaling, $P_x(t, \omega) = P_y\left(\frac{t}{\beta}, \beta\omega\right)$	$\Phi_x(\tau, \theta) = \Phi_y\left(\beta\theta, \frac{\tau}{\beta}\right)$
Convolution, $P_x(t, \omega) = \int P_h(t - \lambda, \omega) P_y(\lambda, \omega) d\lambda$ for $x(t) = \int h(t - \lambda) y(\lambda) d\lambda$	$\Phi_x(\tau, \theta) = e^{\tau\mu} \Phi_y(\tau, \theta)$
Modulation, $P_x(t, \omega) = \int P_h(t, \omega - \nu) P_y(t, \nu) d\nu$ for $x(t) = h(t) y(t)$	$\Phi_x(\tau, \theta) = e^{\nu\mu} \Phi_y(\tau, \theta)$
Time marginal, $\int P(t, \omega) d\omega = \lvert x(t) \rvert^2$	$\Phi(0, \theta) = 1$
Frequency marginal, $\int P(t, \omega) dt = \lvert X(\omega) \rvert^2$	$\Phi(\tau, 0) = 1$
Energy distributions, $\int P(t, \omega)(d\omega) dt = 1$	$\Phi(0, 0) = 1$

Table 7.2 shows some of the kernel functions in the Cohen's class of distributions.

Table 7.2

Time-frequency distribution	**Kernel function**
Wigner-Ville [132], [136]	1
Choi-Williams [24]	$e^{-(\tau\theta)^2/\sigma}$

Table 7.2

Time-frequency distribution	Kernel function
Cohen [31]	$\frac{\sin(\alpha\theta\tau)}{\alpha\theta\tau}$
Page [93]	$e^{j\theta\lvert\tau\rvert/2}$
Gue et al. [54]	$\frac{J_1(\alpha\theta\tau)}{(\alpha\theta\tau)/2}$, J_1 : Bessel function of order one
Zhao et al. [145]	$g(\tau)\lvert\tau\rvert\frac{\sin(\alpha\theta)}{\alpha\theta}$

7.6 Choi-Williams distributions

The Choi-Williams distribution is a member of the Cohen's class of distributions. An exponential kernel function which is proportional to the product of its time-frequency variables, referred to as the product kernel, has been used in the Choi-Williams distribution [24], Eq 7.35. It is often used as a compromise between the high resolution but cluttered WVD versus the smeared but easy to interpret spectrogram [19]. CWD preserves the signal energy marginals, and offers substantial cross-term suppression by means of the Gausian product kernel with little auto-components broadening [64]. CWD has a scaling factor that allows the user to select either a good cross-term reduction or good auto-term preservation. CWD is expressed as

$$CW(t,\omega) = \iint \frac{1}{\sqrt{4\pi\tau^2/\sigma}} \exp\left(-\frac{(\mu-t)^2}{4\tau^2/\sigma}\right) s\left(\mu+\frac{\tau}{2}\right) s^*\left(\mu-\frac{\tau}{2}\right) e^{-j\omega\tau} d\mu d\tau \qquad \text{Eq 7.35}$$

where the kernel function is expressed as in Table 7.2.

In Eq 7.35, the time-indexed autocorrelation function is multiplied to a function with weight factor in such a way that when μ is close to t, then the product $s\left(\mu+\frac{\tau}{2}\right)s^*\left(\mu-\frac{\tau}{2}\right)$ has a large weight and when μ is far from t, then it has a small weight. The parameter σ is the scaling factor, where large values of σ are used for fast changing signals while small values are used for slow varying signals. Therefore, the

kernel acts as a weighting function that reduces cross-terms by providing an exponential distributed time averaging. CWD is a real valued function and is a suitable time-frequency representation for signals that do not have time varying components like frequency modulated signals [66].

If the time series signal is lengthy, which is the case with heart-beat signals, then calculation of the CWD will take a very long time. That is, for computational purposes it is necessary to use a weighting window before evaluating CWD at each time index [24]. The windowed CWD in discrete form is calculated as

$$CW(n,\theta) = 2\sum_{\tau=-\infty}^{\infty} e^{-j2\theta\tau} \sum_{\mu=-\infty}^{\infty} W_M(\mu) \qquad \text{Eq 7.36}$$
$$\frac{1}{\sqrt{4\pi\tau^2/\sigma}} \exp\left(-\frac{\mu^2}{4\tau^2/\sigma}\right) s(n+\mu+\tau)s^*(n+\mu-\tau)$$

where $W_M(\mu)$ is a symmetrical rectangular window which has a value of 1 in the range $-M/2 \le \mu \le M/2$. The parameter M, the length of the window $W_M(\mu)$, determines the range from which the time indexed autocorrelation function is to be estimated. As long as M is large enough, the CWD calculated from Eq 7.36 will be a smoothed version of the CWD calculated from Eq 7.35.

It is possible to introduce another window for Fourier transform calculations in Eq 7.36, in order to further reduce the computational time [24]. The shape and length of such a window will determine the frequency resolution of the CWD.

In appendix B a method is developed and presented for calculations of the CWD of an analytical signal with the least possible number of multiplications. In this algorithm provision is taken so that the kernel function is calculated independently.

7.7 Simulation example

In this section, time-frequency analysis is performed on a simple signal with two sinusoidal components. Our aim is to emphasize the problem associated with cross-terms. We use analytical signals for calculations of WVD, which causes fewer

cross-term components. The most common method of generating an analytical signal is to compute its DFT (discrete Fourier transform), then zero the negative frequency components, and then compute the inverse Fourier transform of the result, which is the way that the analytical signals in this chapter are calculated. This example shows how the cross-terms might obscure the auto-terms. We will show the CWD of the same signal in order to indicate its ability in suppressing the cross-terms.

7.7.1 *Example 7.1*

A time series signal with three sinusoidal components with analog frequencies of 30 *Hz*, 115 *Hz*, and 200 *Hz*, and a sampling frequency of $1K$ *Hz*, are hamming windowed and added together. The resulting signal is shown in Figure 7.1.

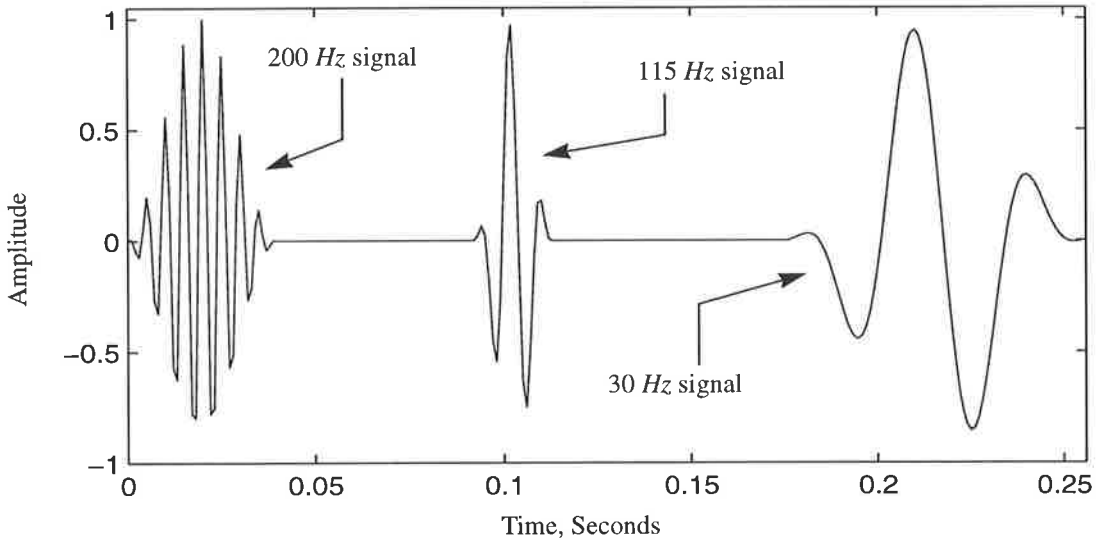


Figure 7.1 Signal for example 7.1.

In Figure 7.2 the Wigner-Ville distribution of the signal is shown. In this figure, it is clear that the auto-term corresponding to the 115 *Hz* component is obscured by cross-terms. With more complicated signals such as heart sounds, it is often extremely difficult to interpret the WVD.

In Figure 7.3, CWD of the same signal is shown. It is obtained by using the analytical signal, therefore it does not contain any cross terms due to positive and negative frequency components which would have occurred if the real signal had been used [44], [58].

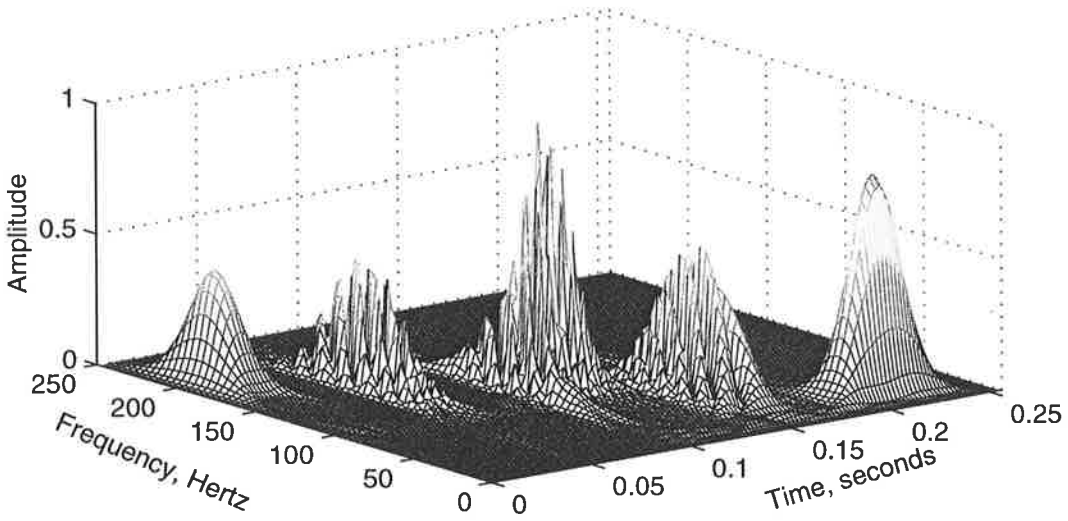


Figure 7.2 Wigner-Ville distribution of the signal of the Figure 7.1.

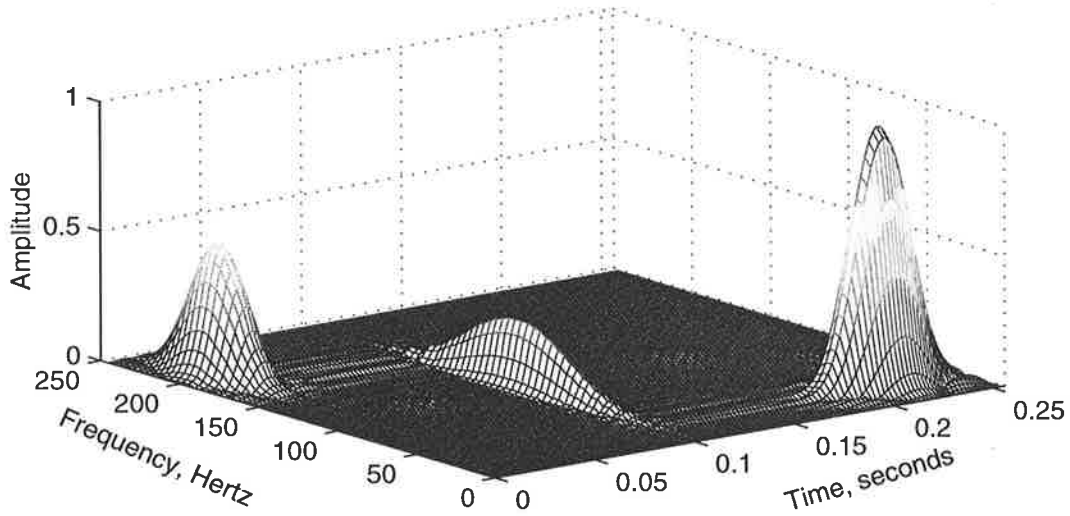


Figure 7.3 CWD of signal of Figure 7.1.

As is clear from this figure, the cross-terms of Figure 7.2 have been substantially reduced and the auto-term corresponding to the 115 Hz component is preserved. It is clear that cross-terms are a disturbing phenomena when one is dealing with pattern recognition and feature extraction from signals using the time-frequency plane.

7.8 CWD of Phonocardiograms

In this section we explain the process of analysing case #1 in Table 2.2 of Chapter 2 using CWD. For comparison purposes we have chosen the same signal that has

been used for the short-time Fourier transform analysis in Chapter 6. Again we use the algorithms presented in Chapter 4 to prepare the heart-beat signals from recorded phonocardiograms before and after angioplasty. In Figure 7.4 the heart-beat signals before and after angioplasty are shown for convenience.

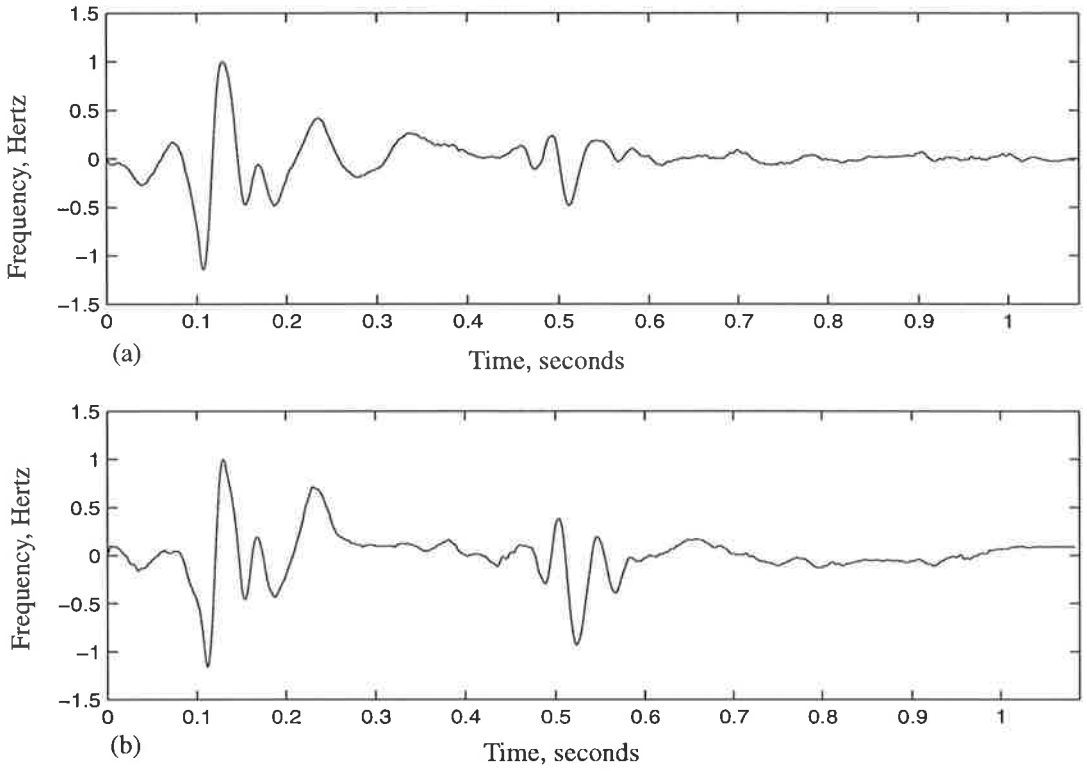


Figure 7.4 Heart-beat signals, a) before and b) after angioplasty.

As with STFT, the heart sound signals are being analysed for extraction of information using the CWD technique. Our interest in time-frequency analysis, as was mentioned earlier, is to extract time-frequency features that indicate any differences in the heart-beat sounds before and after angioplasty. Since heart sounds are multicomponent signals, WVD most probably will generate cross-terms. These cross-terms might obscure some of the low energy auto-terms; therefore, CWD is used to calculate the time-frequency matrix of the heart sounds.

The steps involved in extracting and comparing the locations of concentration of the energy in the CWD time-frequency planes of the heart-beat signals before and after angioplasty, are the same as the steps that were taken in the short-time Fourier transform analysis of Chapter 6.

7.8.1 Calculations of CWD time-frequency matrix

As was mentioned earlier, the CWD algorithm is not computationally efficient, especially when the signal is long. The length of the heart-beat signals vary from about 1500 samples to 2500 samples, depending on the heart-beat rate of the patient. Therefore, it will take a considerable amount of time to calculate the time-frequency matrix of the CWD. In order to reduce the computational time as much as possible, we used the running window scheme suggested in [24]. The window length was taken to be 200 milliseconds for a typical heart-beat signal whose length was equal to one second. That is, we set the length of the window equal to one-fifth of the total length of the heart-beat signal. For calculation of CWD, a program was written and CWD was implemented. In order to make the computation more efficient, no multiplication is performed for the bits that have value zero. This is explained in Appendix B in more detail. The value of $\sigma = 3$ was used throughout the experiment. This value was determined by trial and error from simulation results; smaller values resulted in pronounced cross-terms, while larger values smoothed out the low energy peaks.

Once the time-frequency matrix was calculated, a Gausian smoothing mask was applied to remove any sharp surface variations. The two-dimensional smoothing filter shown in Figure 6.7 was used.

The smoothed modulus of the CWD time-frequency matrix of the signal of Figure 7.4 (b) is shown in Figure 7.6-1, -2, -3. Each figure shows eight 3-D plots of the modulus of the CWD matrix, normalised in amplitude. Each group of four plots (two per page) represents the CWD matrix over the entire time axis in a certain band of frequency. For example, the top four plots of Figure 7.6-1 are the surfaces of the modulus of the CWD matrix in the 0-35 *Hz* range, where the first plot shows it from zero to 0.27 seconds, the second plot is from 0.27 to 0.54 seconds, the third plot is from 0.54 to 0.81 seconds, and the fourth plot is from 0.81 seconds to the end. The organization of Figure 7.6-1, -2, and -3 is explained pictorially in Figure 7.5, which shows the time and frequency locations of each plot of Figure 7.6-1, 2, and 3.

7.8.2 Locations of concentration of energy in CWD matrix

The local maxima detector algorithm of Chapter 5 is applied to the modulus of CWD time-frequency matrix of signals of Figure 7.4 (a) and (b) and locations of con-

centration of energy are obtained. The result for each case is three vectors: time, frequency, and magnitude vector. The number of local maxima before angioplasty were 295 and after angioplasty were 310. We performed the analysis up to 250 *Hz* in the frequency range, therefore both sets of vectors are truncated after 250 *Hz*. By doing so, the dimension of the before angioplasty vectors was reduced to 290 and that of the after angioplasty vectors was reduced to 285. It was explained in Chapter 6 that the analysis was restricted to a frequency band of 0-250 *Hz* because we did not find much information in most of the signals at frequencies higher than 250 *Hz*.

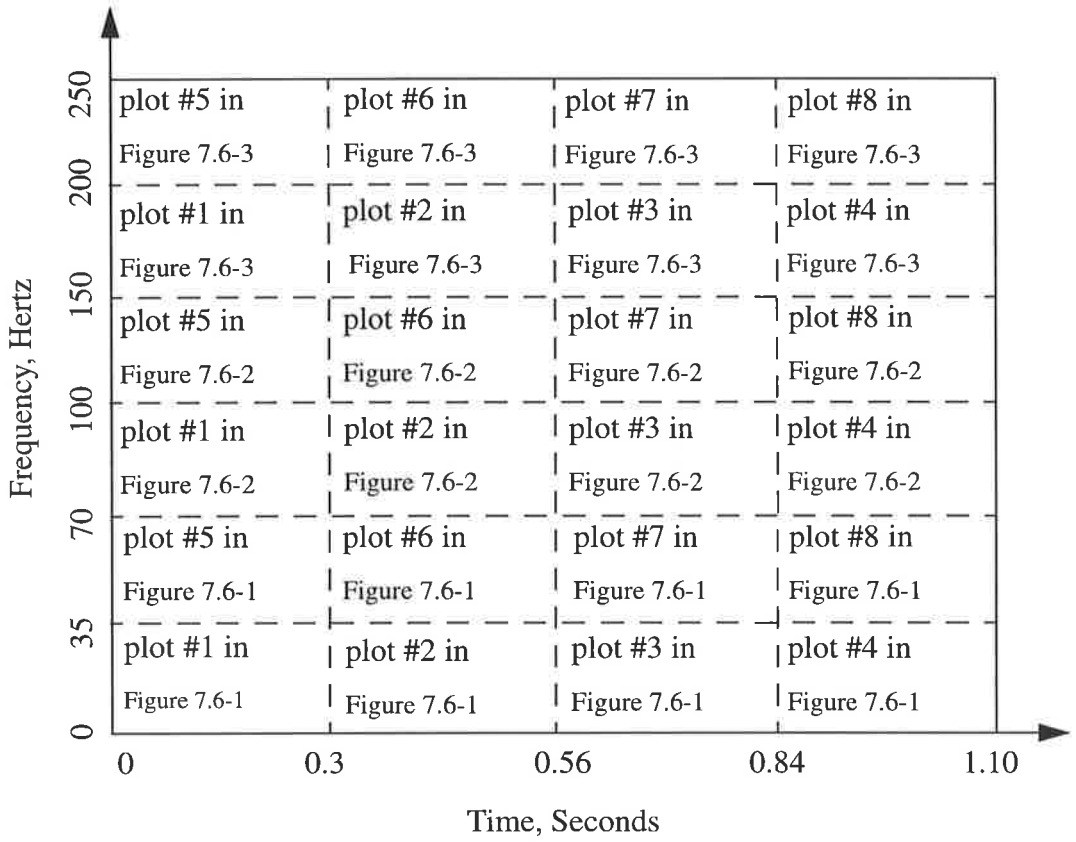


Figure 7.5 Time-frequency organizations of Figure 7.6-1, -2, -3.

7.8.3 Time-frequency patterns in CWD matrix

We now rearrange the time-frequency vectors in ascending order of frequency. Therefore, the frequency vector becomes

$$F = \left[f_1 \; f_2 \; \cdots \; f_L \right] \qquad \text{Eq 7.37}$$

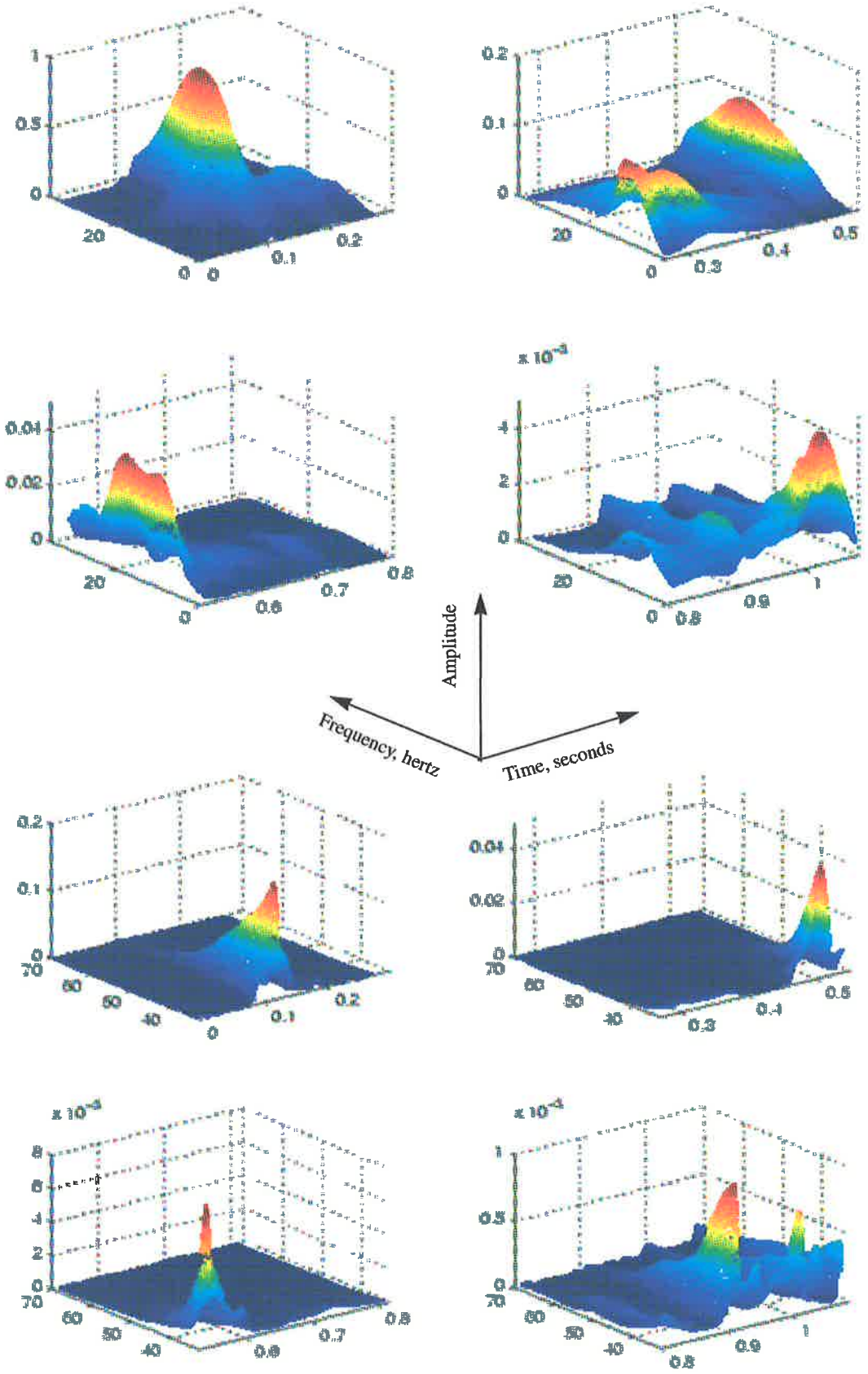


Figure 7.6-1 Parts of the modulus of the CWD of the signal of Figure 7.4 (b).

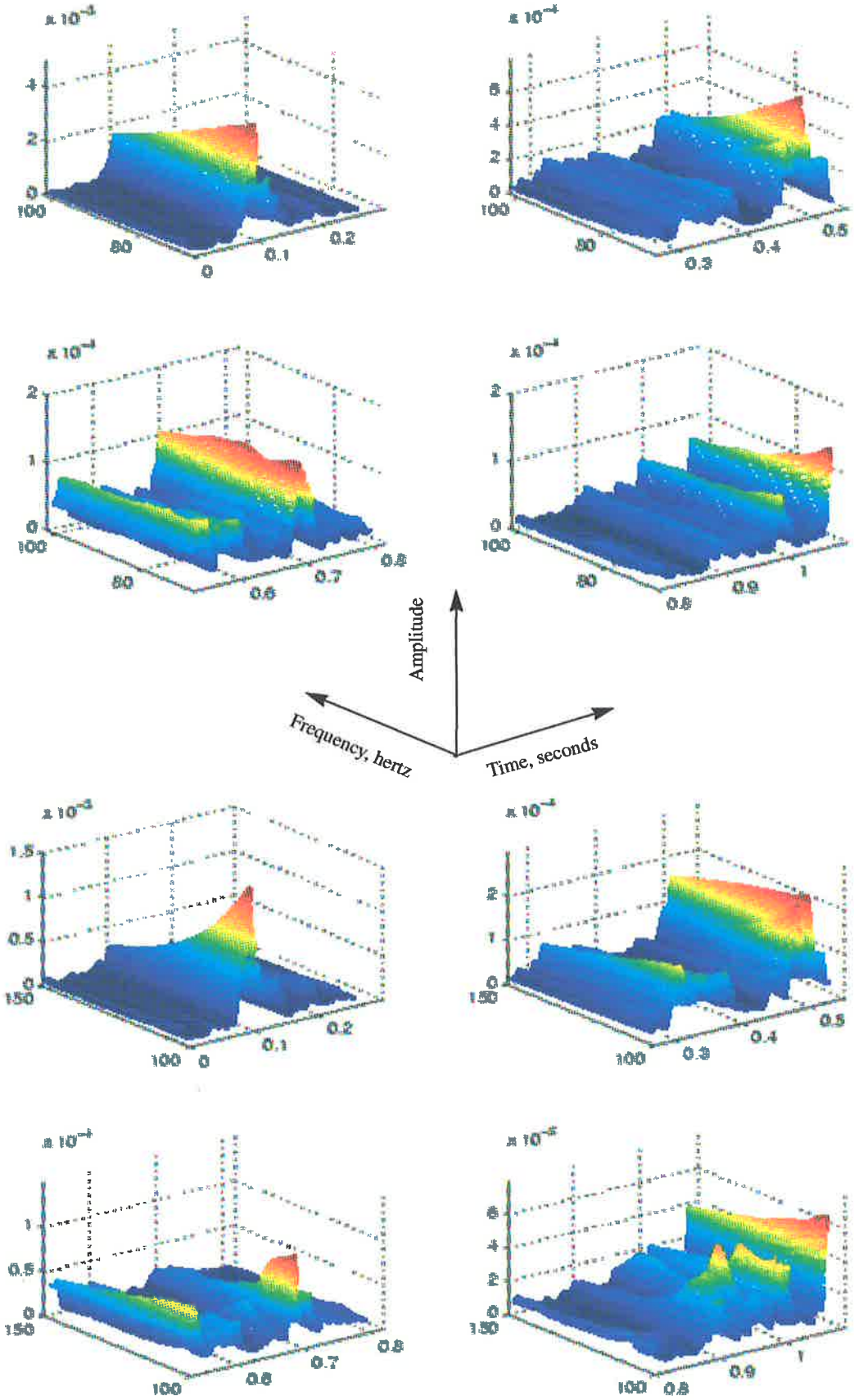


Figure 7.6-2 Parts of the modulus of the CWD of the signal of Figure 7.4 (b).

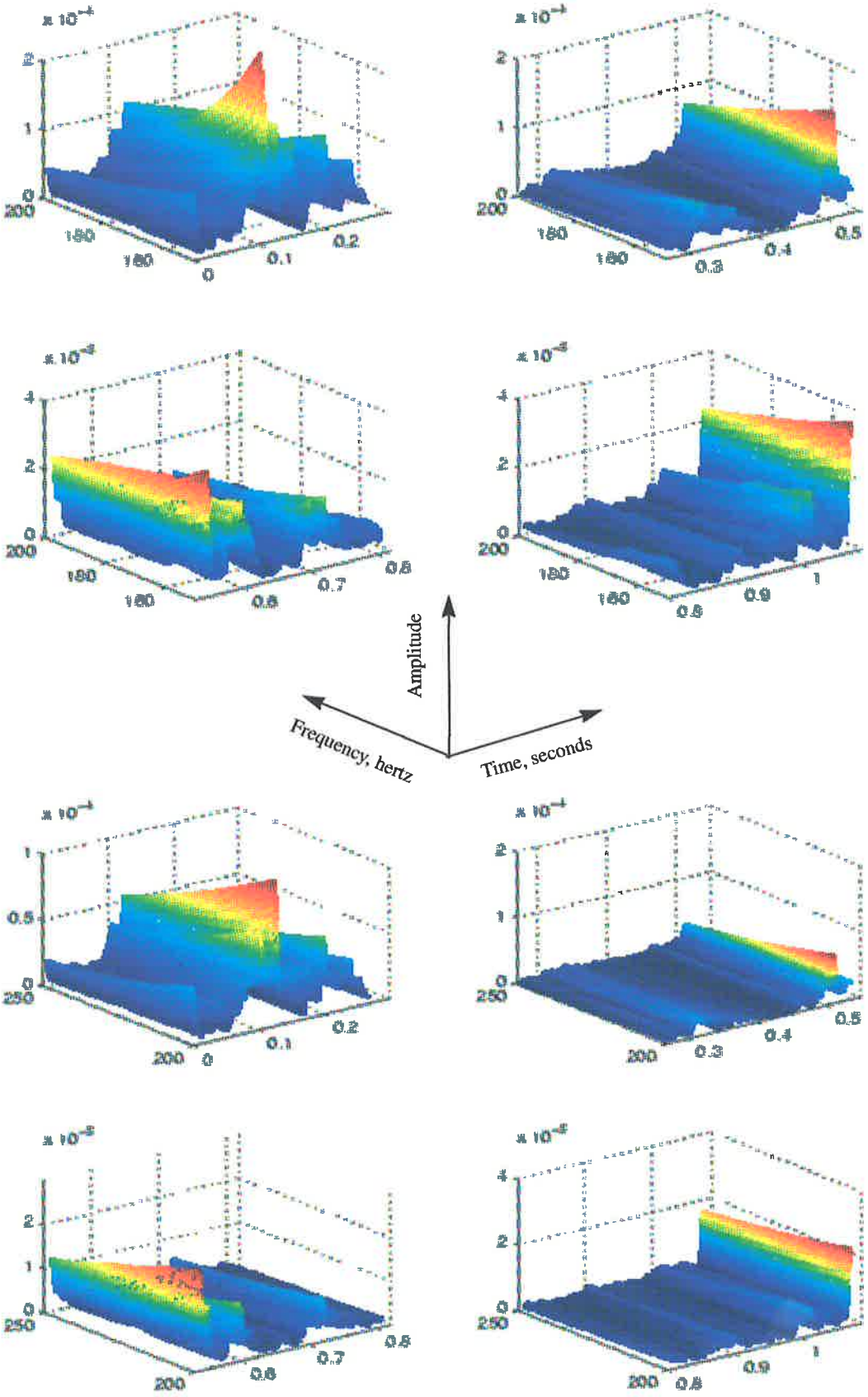


Figure 7.6-3 Parts of the modulus of the CWD of the signal of Figure 7.4 (b).

where $f_1 < f_2 < \ldots < f_L$. Note that the other two vectors T, and A (see Eq 6.13 and Eq 6.14) are rearranged with respect to the corresponding elements of F.

The time-frequency-magnitude (TFM) product vectors of before and after angioplasty planes are calculated as (see section 6.5)

$$\chi(k) = T(k)F(k)A(k) \qquad \text{for k=1, 2, ..., L} \qquad \text{Eq 7.38}$$

Detailed discussions on TFM analysis is presented in section 6.5.

The TFM vectors for signals of Figure 7.4 are plotted in Figure 7.7.

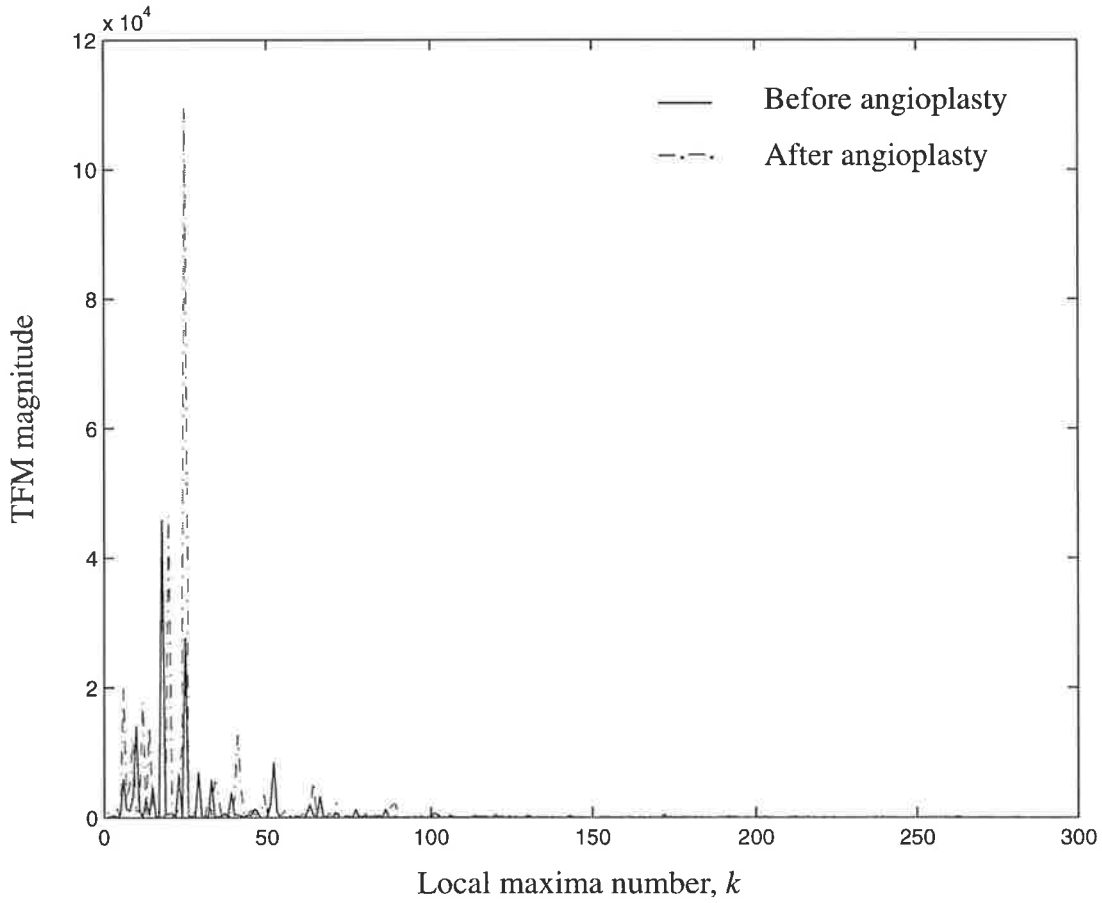


Figure 7.7 The TFM vectors for signals shown in Figure 7.4.

We calculated the average of the magnitude vectors and found that the energy of the signal after angioplasty is higher than that before angioplasty.

By comparing the two graphs of Figure 7.7, we found many similar sections. These similar sections represent similar patterns of local maxima in the modulus of the

CWD matrix. The number of detected local maxima points for CWD is larger than that detected using STFT. This is due to the fact that CWD has a better time-frequency resolution than STFT but there might be some local maxima points that are due to cross-terms. Although the Choi-Williams method suppresses the cross terms, it does not eliminate them totally, and there might be some low level local maxima points due to cross terms. When we examine the TFM graphs of Figure 7.7, we realise that in some parts the two graphs are not quite similar, which indicates that there are some local maxima generated by cross-terms. Fourteen parts were extracted and plotted in Figure 7.8. These parts were the best similar part that were recognizable. The similarity of two graphs differs from part to part. For example, in Figure 7.8 the graphs in (d) and (k) are not as similar as the graphs in (a) and (f). The number of local maxima for all of these parts are shown in Table 7.3.

Table 7.3

Local maxima			Local maxima			Local maxima	
before	**after**		**before**	**after**		**before**	**after**
1-11	1-10		12-19	11-23		20-26	22-26
27-34	27-36		35-55	37-55		56-68	56-66
69-83	67-86		84-105	86-103		106-124	104-126
125-160	127-159		161-191	160-200		192-230	197-236
231-263	237-264		264-290	265-285			

The time-frequency patterns of the local maxima corresponding to Table 7.3 are plotted in Figure 7.9-1, -2; where local maxima before and after angioplasty are represented with "$*$" and "o", respectively. It is clear that there are similar patterns in the time-frequency matrices of before and after angioplasty. As we study the patterns of Figure 7.9-1 and -2 more closely, it becomes clear that there is an average frequency shift between the patterns of before and after angioplasty; that is, the average frequency of the patterns belonging to the signal after angioplasty are higher than that the average frequency before angioplasty. It can also be noticed that the frequency shift is greater in the higher frequency range patterns. This is similar to what we found with the STFT analysis.

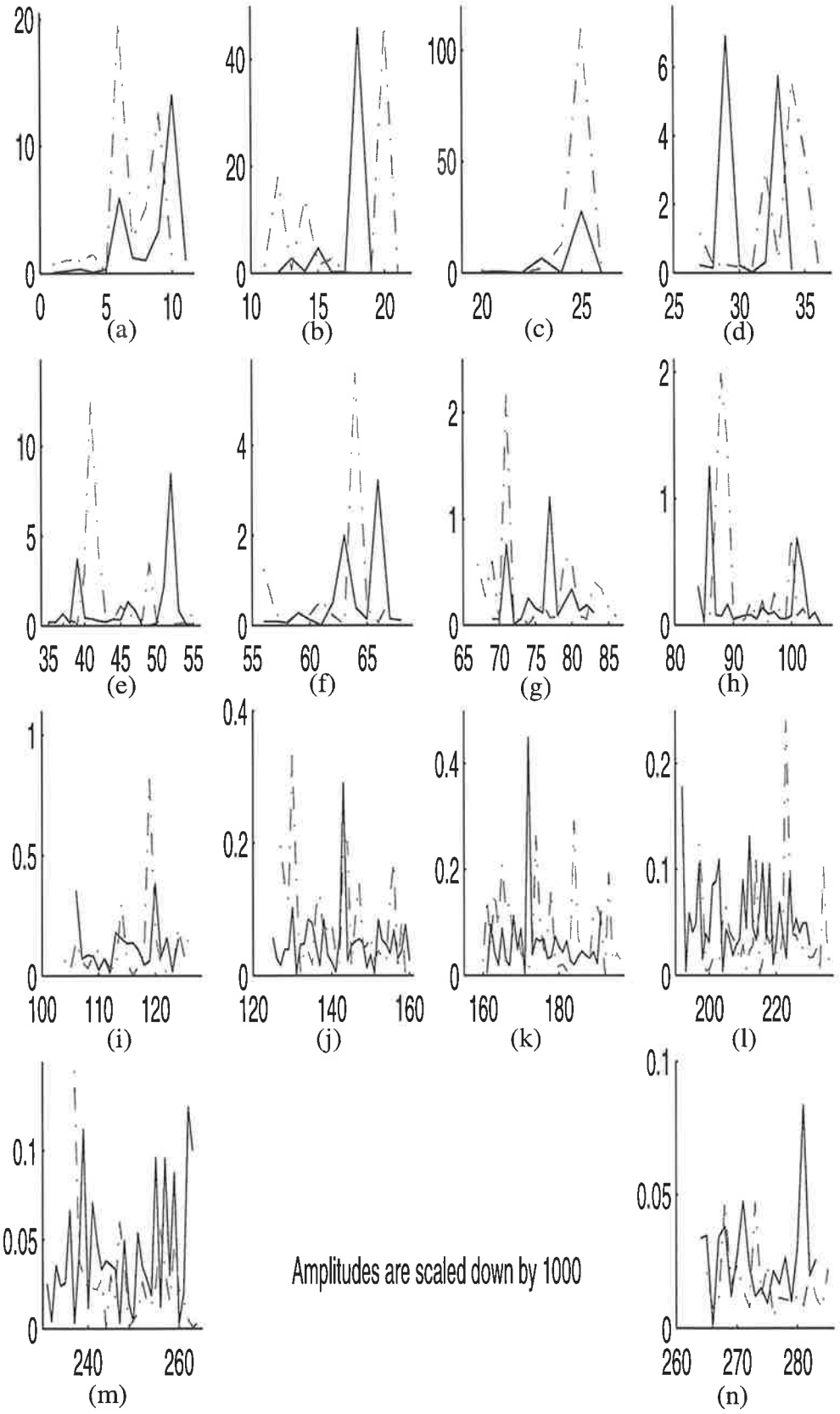


Figure 7.8 Sections of the TFM product plots of Figure 7.7.

In Table 7.4, the first column represents band of frequencies where patterns of signals in Figure 7.7 were identified. The second and third columns are average frequencies of the patterns in the first heart sound regions before and after angioplasty, respectively. Likewise, the fourth and fifth columns correspond to the second heart sound regions before and after angioplasty, respectively. The sixth and seventh columns correspond to the third heart sound regions before and after angioplasty, respec-

Table 7.4

Band	1st HS MF		2nd HS MF		2nd HS MF		HS pattern MF		
	before	after	before	after	before	after	before	after	shift
0-15	8	9	10	11	5	5	8	8	0
14-21	18	18	16	16	16	17	17	17	0
20-24	21	23	21	23	21	23	21	23	2
20-35	26	31	29	32	26	28	27	30	3
20-50	33	42	36	40	33	39	34	40	6
35-50	42	47	42	46	42	47	42	47	5
40-45	46	52	47	52	46	53	46	52	8
48-62	51	57	51	60	50	57	51	58	7
50-70	55	67	54	67	55	64	55	66	11
50-90	64	80	66	76	65	79	65	78	13
73-98	78	91	79	93	79	90	79	91	12
80-121	89	106	88	109	88	107	88	107	19
90-150	106	136	106	139	105	129	106	135	29
100-250	131	169	129	162	126	187	129	173	44

HS: Heart Sounds

MF: Mean Frequency

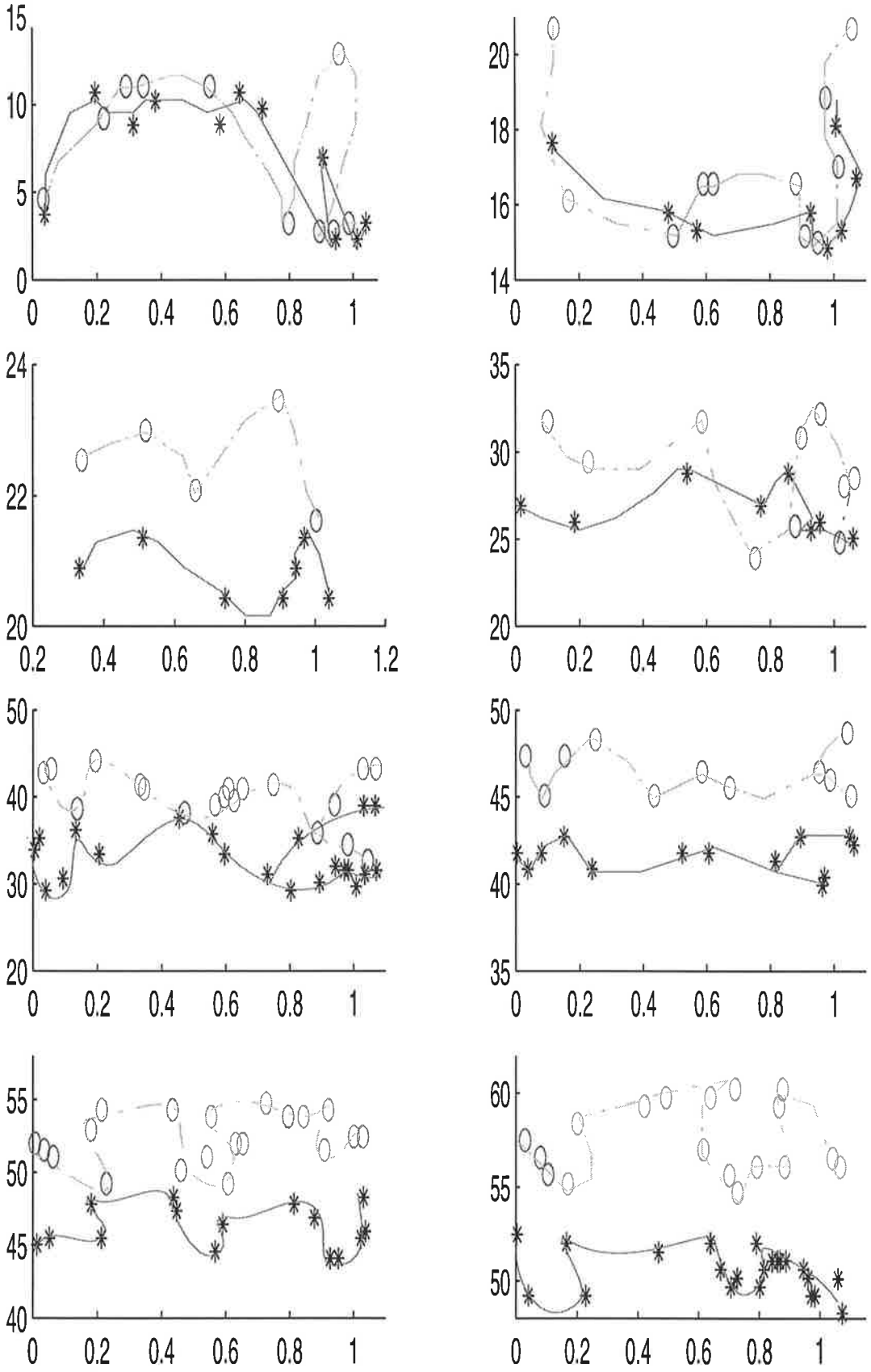


Figure 7.9-1 Time-frequency patterns corresponding to plots of Figure 7.8.

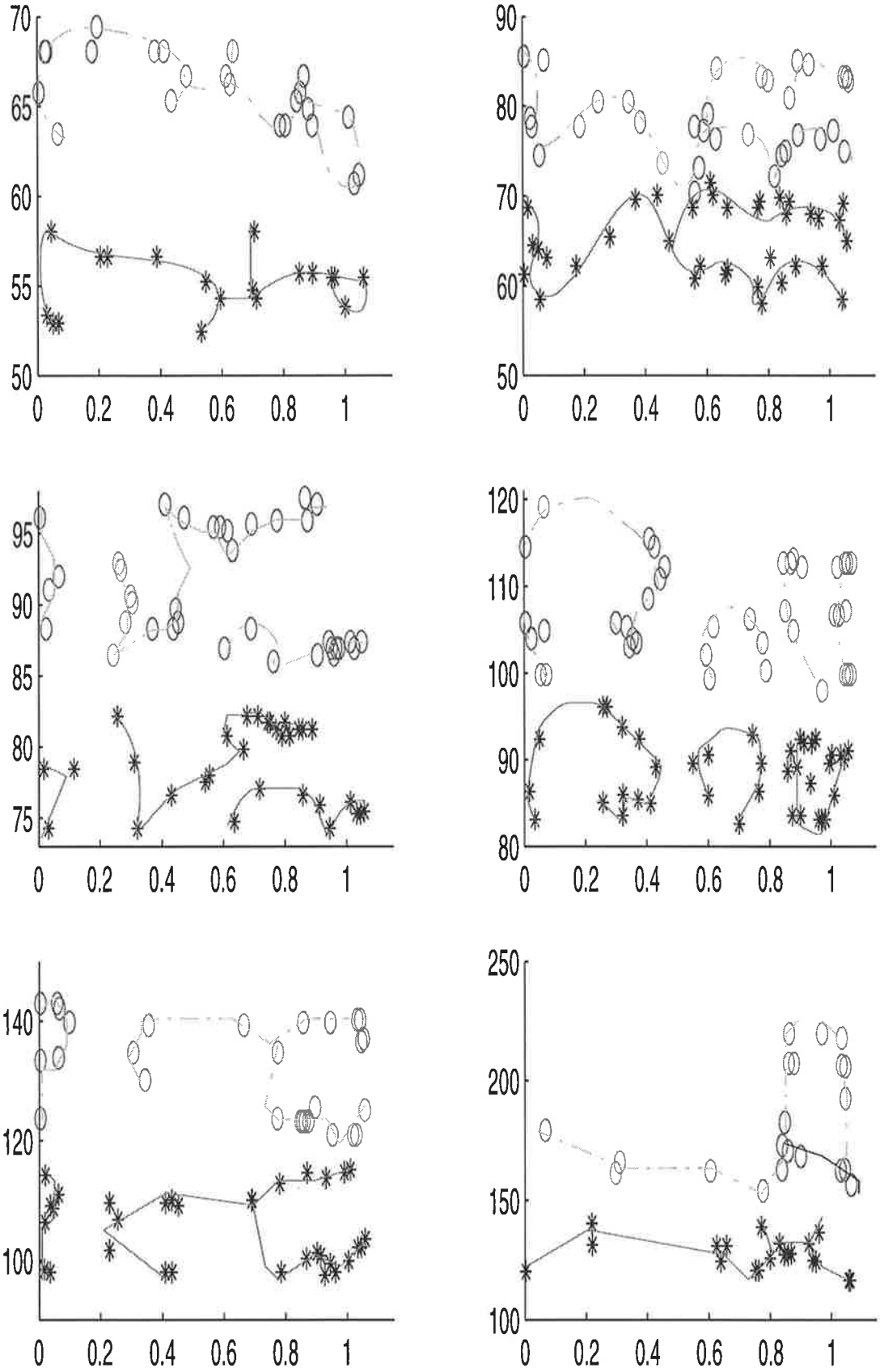


Figure 7.9-2 Time-frequency patterns corresponding to plots of Figure 7.8.

tively. The eighth and ninth columns show the average frequencies of the patterns before and after angioplasty, respectively. The last column shows the difference between the average frequencies of the patterns. For example, in the frequency band 0-15 *Hz* the first entry of the first column is 8 *Hz,* that indicates the average frequency of the first heart sound in the first pattern before angioplasty. Likewise, the first entry eighth column is 8, which is the average of 8, 10, and 5. This means that the average frequency of the whole pattern before angioplasty is 8 *Hz*. The averages were rounded off to their closest integers. The second, third, fourth, fifth, and sixth columns of Table 7.4 are plotted against the average frequencies of the patterns in Figure 7.10.

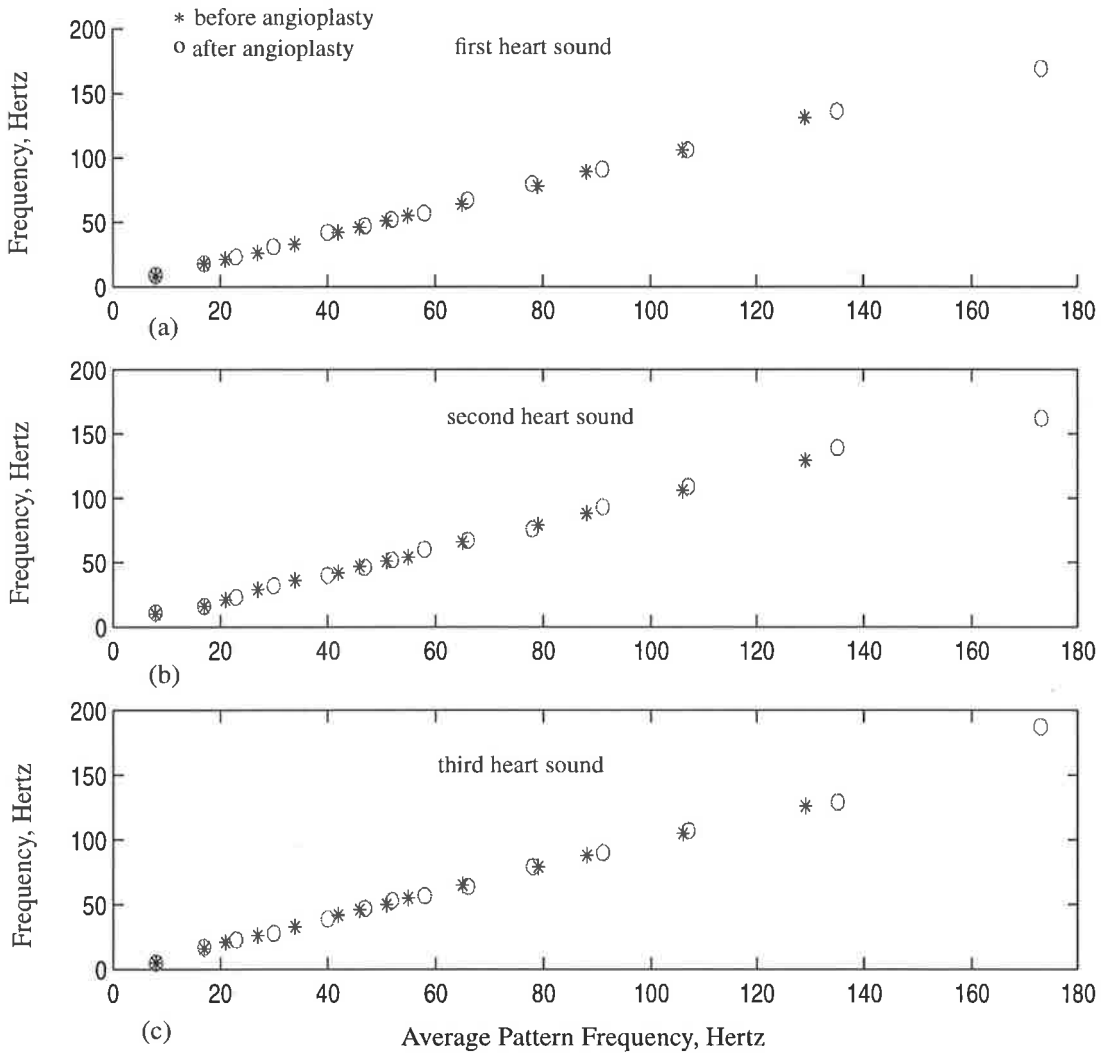


Figure 7.10 Average frequencies in a) first, b) second, and c) third heart sounds obtained from Table 7.4.

In Figure 7.10 (a), (b), and (c) the average frequencies of the first, second, and third heart sounds before and after angioplasty are shown, respectively. It is clear from the figures that there is a shift in the average frequencies after angioplasty, where the shift in the average frequencies increases with frequency.

7.8.4 Results from other cases

We performed the analysis for all the signals in Table 2.2. The results of five more signals are summarised in Table 7.5. The number of the patterns differed from signal to signal. The maximum number of patterns was found to be 19 while the minimum number of patterns was found to be 12.

Table 7.5 Average frequencies in time-frequency patterns in five subjects.

Band	1st HS MF		2nd HS MF		2nd HS MF		HS pattern MF		
	before	after	before	after	before	after	before	after	shift
0-16	7	8	9	10	5	6	7	8	1
13-19	16	17	15	15	16	18	16	17	1
20-25	22	24	21	24	22	24	22	24	2
21-33	25	31	28	32	26	32	26	32	6
22-52	32	43	34	43	33	45	33	45	12
29-53	38	48	41	51	42	51	40	50	10
38-60	44	52	44	54	49	59	46	55	9
42-60	47	54	49	56	50	60	49	57	8
48-62	51	57	51	60	51	58	51	58	7
52-73	56	66	54	65	57	69	56	67	11
55-90	65	82	68	83	66	87	66	84	12
68-100	72	91	78	88	75	94	75	91	16
82-118	89	110	86	107	87	109	87	109	22
90-139	102	130	108	138	101	131	104	133	29
120-178	128	164	132	169	135	175	132	169	37
150-250	172	210	181	225	174	228	176	221	45
0-22	9	10	12	12	7	8	9	10	1
12-25	14	19	15	21	14	18	14	19	5
19-26	21	24	22	25	21	23	21	24	3
20-29	22	28	23	28	23	27	23	28	5
20-35	26	32	29	33	28	34	28	33	5
30-50	32	42	37	49	33	46	34	46	8

Table 7.5 Average frequencies in time-frequency patterns in five subjects.

Band	1st HS MF before	1st HS MF after	2nd HS MF before	2nd HS MF after	2nd HS MF before	2nd HS MF after	HS pattern MF before	HS pattern MF after	shift
32-55	41	49	41	51	48	54	43	51	8
43-63	48	57	49	58	51	61	49	59	10
50-75	55	68	51	62	59	72	55	67	12
69-89	72	83	71	84	71	85	71	84	13
75-100	79	92	83	94	72	91	78	92	14
95-148	107	138	106	140	105	134	49	59	10
120-195	142	179	152	192	136	178	106	137	31
169-250	181	201	180	204	194	231	143	183	40
0-14	6	7	8	8	5	6	6	7	1
12-18	13	14	14	14	15	15	14	14	0
19-25	21	23	21	23	20	24	21	23	2
20-28	23	25	21	26	23	27	22	26	4
20-35	24	28	27	32	26	33	26	31	5
32-45	35	41	37	41	39	42	37	41	4
40-51	46	50	41	48	43	49	45	49	4
48-65	55	62	49	55	57	63	54	60	6
51-67	53	65	55	65	58	66	57	65	8
55-92	67	82	61	79	71	88	66	83	17
62-95	69	79	74	85	79	91	74	85	11
71-105	81	93	82	97	78	95	80	95	15
85-123	98	110	91	104	101	121	97	112	15
98-137	105	115	109	121	113	132	109	127	18
109-154	126	139	131	146	133	152	128	146	18
128-170	141	161	135	159	144	167	140	162	22
149-189	153	171	157	179	162	182	157	177	20
171-227	179	198	189	211	194	223	187	211	24
190-250	197	215	199	223	203	238	201	225	26
0-14	7	8	10	11	5	6	7	8	1
13-25	13	15	14	16	9	11	12	14	2
24-38	25	30	22	27	27	33	25	30	5
35-56	41	52	37	46	39	53	39	50	11
50-78	58	72	55	71	61	77	58	73	15
75-100	76	89	86	96	81	92	81	92	11
94-134	97	116	101	121	106	129	101	115	14
125-164	129	142	136	151	138	161	131	151	20
157-196	161	183	172	195	166	191	166	190	24
165-210	172	201	178	205	183	208	178	205	27
191-250	197	221	203	224	198	229	199	225	26

Table 7.5 Average frequencies in time-frequency patterns in five subjects.

Band	1st HS MF before	1st HS MF after	2nd HS MF before	2nd HS MF after	2nd HS MF before	2nd HS MF after	HS pattern MF before	HS pattern MF after	HS pattern MF shift
0-16	6	7	9	9	5	6	7	7	0
13-22	15	18	16	19	14	17	15	18	3
21-28	23	25	23	24	24	25	23	25	2
25-32	27	31	27	30	28	31	27	31	4
29-46	35	41	31	39	33	41	33	40	7
38-58	42	49	45	52	47	53	45	51	6
52-71	55	63	61	68	54	60	57	64	7
66-85	68	78	73	82	74	83	72	81	9
80-105	91	103	89	101	85	95	88	100	12
96-127	98	112	107	124	102	121	102	119	17
110-150	119	133	125	142	128	147	124	141	17
125-161	135	156	128	148	137	158	133	154	19
138-187	146	167	149	171	154	181	150	173	23
169-210	175	198	171	198	177	205	174	200	26
195-250	201	232	204	238	208	245	201	238	37

We have grouped the patterns of before and after angioplasty in four frequency bands according to their average pattern frequencies, as shown in Table 7.6. The frequency band designated as *below 50 Hz* includes all patterns whose average frequencies were lower than 50 *Hz*. Likewise the frequency band *above 50 Hz and below 100 Hz*, includes all patterns whose average frequencies were between 50 *Hz* and 100 *Hz*. The other two categories include patterns whose average frequencies were between 100 *Hz* and 150 *Hz*, and between150 *Hz* and 250 *Hz*. This table shows the highest, lowest and mean average frequencies found in the patterns.

As is shown in Table 7.6, for example in the frequency band *below 50 Hz*, the minimum average pattern frequency in first heart sound region in all of the signals was found to be *5 Hz* before angioplasty and *6 Hz* after angioplasty. Likewise the maximum average pattern frequency in the first heart sound region was found to be *38 Hz* and *45 Hz* before and after angioplasty respectively.

The third row in each category in Table 7.6 represents the mean of all of the average pattern frequencies in that category. For example in the frequency band *below 50 Hz*, the mean average pattern frequency in first heart sounds before angioplasty was found to be 18 *Hz*.

Table 7.6 Average frequencies of the patterns in first, second, and third heart sound regions.

Frequency bands	average pattern frequency in 1st heart sound before after	average pattern frequency in 2nd heart sound before after	pattern frequency in 3rd heart sound before after	Shift in average frequency
Below 50 Hz	Min.: 5 6 Max.: 38 45 Mn.: 18 22	Min.: 7 8 Max.: 40 47 Mn.: 20 25	Min.: 4 4 Max.: 39 48 Mn.: 20 24	4
50-100 *Hz*	Min.: 55 62 Max.: 79 94 Mn.: 63 69	Min.: 51 50 Max.: 84 98 Mn.: 66 74	Min.: 50 50 Max.: 79 95 Mn.: 62 71	8
100-150 *Hz*	Min.: 102 101 Max.: 143 148 Mn.: 121 138	Min.: 101 101 Max.: 144 145 Mn.: 119 134	Min.: 100 103 Max.: 142 148 Mn.: 123 139	16
150-250 *Hz*	Min.: 153 162 Max.: 197 221 Mn.: 173 192	Min.: 152 159 Max.: 203 224 Mn.: 178 201	Min.: 152 151 Max.: 206 238 Mn.: 181 208	23

Min: Minimum
Max: Maximum
Mn: Mean

The last column of Table 7.6 shows the frequency shift in mean frequency of the patterns after angioplasty; that is, it represents average of "Mn" for the first, second, and third heart sounds of the related category. For example, as is shown in the table, in the frequency band *below 50 Hz,* for the 1st heart sound before angioplasty "Mn" was found to be 18 *Hz*, for the second heart sound it was found to be 20 *Hz*, and for the third heart sound it was found to be 20 *Hz*. Therefore average "Mn" before angioplasty would be 19.33 *Hz*. Likewise, "Mn" for the first, second, and third heart sounds after angioplasty were found to be 22 *Hz*, 25 *Hz*, and 24 *Hz* respectively. Thus, average "Mn" after angioplasty would be 23.66 *Hz*. The difference between 23.66 and 19.33 is 4.33 which was rounded off to 4 which is shown in the last column of Table 7.6 in the frequency band *below 50 Hz*.

7.9 Conclusions

In this chapter we briefly reviewed Wigner-Ville and Choi-Williams distributions. We showed that the Wigner-Ville distribution is not suitable for time-frequency analysis of signals that consist of multicomponents. In order to reduce the effects of cross-terms we used the Choi-Williams distribution. We pointed out that due to the computational inefficiency of this method it is necessary to use a running window technique. We proposed a method for calculating the CWD which further reduces the computational time by performing the least possible number of multiplications only.

The heart-beat cycle of heart sounds before and after angioplasty were prepared and their CWD were calculated. The time-frequency matrices were smoothed by applying a two-dimensional Gaussian mask. This removed small and sharp variations on the surface of the modulus of the time-frequency matrices. The three dimensional plots of a typical CWD were shown.

The local maxima detector algorithms were used to determine the locations of concentration of energy in the CWD plane. The temporal, spectral, and magnitude vectors, were used to calculate the time-frequency-magnitude product vectors. The TFM product vectors for before and after angioplasty were compared and similar sections were identified. Using the similar sections, the time-frequency patterns of before and after angioplasty were plotted. It was found that two patterns in the same frequency band do have resemblances and there exists some frequency shift between them.

The average frequencies in first, second, and third heart sound regions were calculated and compared. Their plots show that the average frequencies before and after angioplasty have similar variations.

In the low frequency range, below *50 Hz*, the average frequency shift is less than *10 Hz*. For the medium frequency range, between *50 Hz* and *100 Hz*, the average frequency shift is also less than *10 Hz* and in the high frequency range, above *100 Hz* and up to *150 Hz*, the average frequency shift is between *10 Hz* and *20 Hz*, and in the frequency range above *150 Hz* and below *250 Hz* the average frequency shift is more than *20 Hz*.

Chapter 8

Wavelet Based Time-Frequency Analysis of Heart Sounds

8.1 Introduction

Just like short-time Fourier transforms, wavelets are versatile tools of harmonic analysis. Since their introduction in the1980's, considerable time has been devoted to their developments and practical applications. It is not a new theory in a sense that many of the ideas and techniques involved in wavelets, such as filter bank theory and quadrature mirror filters were developed independently in various signal processing applications and have been known for some time. On the other hand, it is a new theory since the recent development in the mathematical foundations of wavelets has provided a unified framework for the subject.

Many of the shortcomings of the short-time Fourier transform stem from the fact that its time-frequency grid is uniform for all time and frequency. That is, the analysis window has a fixed length and therefore, is unable to capture the same number of

sinusoids in both high and low frequency ranges. Thus, the spectral estimations will be good at one end and poor at the other end. Wavelet transforms provide a solution for this problem by using an analysis window which depends on both time and frequency.

The mathematical aspects of the wavelet transforms are introduced in [52] and later treated in considerable details by several authors [38], [77], [78], [120]. There are a number of books that discuss wavelets in greater details [25], [39]. Consideration of wavelets as a potential tool for signal processing followed the influential papers by Daubechies [38], and Mallat [77]: Daubechies developed a technique to generate orthonormal wavelet bases and also established a relation between the filter banks and wavelet transforms and Mallat presented multiresolution analysis of the wavelets.

Applications of wavelets as a signal processing tool are closely related to its presentation by filter banks, and hence its multiresolution analysis. Nonuniform filter bandwidths in the structure of the filter banks and nonuniform decimation ratios in the subbands are the key points of wavelet transforms. For more information on filter bank theory see [144]. Filter banks will be discussed briefly later in this chapter.

The literature on wavelet transforms is rather vast and most of it require a profound mathematical background which is perhaps unnecessary for most signal processing experts. It is beyond the scope of this chapter to go through a mathematical literature review of wavelets. From the signal processing point of view, tutorial treatments can be found in [59] and [105]. In order to highlight the applicability of wavelets in biomedical signal processing, especially in connection with the heart sounds, we have given a brief literature review on the heart sound analysis using wavelets in Chapter 1.

In this chapter we use continuous wavelet transform (CWT) for manipulation of the wavelet coefficients of the heart sounds. These coefficients are used to study the differences in time-scale characteristics of heart sounds before and after angioplasty. In section 8.2, the wavelet basis functions are discussed. In section 8.3, classification of wavelets to continuous and discrete wavelet transforms are explained. In section 8.4, desired properties of analyzing wavelet functions are highlighted. This is followed in

section 8.5, with a review of wavelets using filter banks with constant-Q filters. The time-frequency localization of wavelets are discussed in section 8.6. Then in section 8.7, we explain Morlet wavelets as analyzing wavelets, which we will use for the heart sound analysis. In section 8.8, the Chirp-z transform implementation technique for continuous wavelet transform is explained. In section 8.9, octave based time-scale representation for wavelet transforms is discussed. In section 8.10, results obtained from wavelet based analysis of a heart sound are given. The wavelet coefficients are calculated and the locations of concentration of energy in the time-scale matrix are detected. Using the time-frequency-magnitude method, described in Chapter 6, the time-frequency patterns are detected and compared before and after angioplasty. In section 8.11, results obtained from wavelet analysis of all of the recorded heart sounds are summarized.

8.2 Wavelet Basis Functions

The wavelet transform of a function is defined as a projection of this function onto a set of analyzing functions called wavelets. A wavelet is a function that should have a small concentrated burst of finite energy and must exhibit some kind of oscillations. These characteristics make it a wave like localized function in the time domain. Wavelets are generated from a single function by performing dilation (scaling) and translation (shifting). The generating function is called the *mother* wavelet.

Let $\psi(t)$ be a function of real variable t and represent a mother wavelet. Localization suggests that it should decrease rapidly to zero as $|t|$ tends to infinity. The other condition requires that $\psi(t)$ must vibrate like a wave. Let a be a scaling factor and b represent a time shift, then the mother wavelet $\psi(t)$ can be used to generate a series of wavelets as follows

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right) \qquad \text{Eq 8.1}$$

where $a>0$ and $a, b \in R$ (R represents the set of real numbers). $\psi_{a,b}(t)$ is called the wavelet basis function. Let $s(t)$ be an absolute integrable function, that is $s(t) \in L^2(R)$. The forward wavelet transform coefficients of $s(t)$ are computed using the inner products $\langle s(t), \psi_{a,b}(t) \rangle$ and defined as

$$\Gamma(a, b) = \langle s(t), \psi_{a,b}(t) \rangle = |a|^{-1/2} \int_{-\infty}^{\infty} s(t) \psi\left(\frac{t-b}{a}\right) dt \qquad \text{Eq 8.2}$$

The function $s(t)$ is reconstructed from its wavelet coefficients as

$$s(t) = C_{\psi}^{-1} \int_{a>0} \int \Gamma(a', b') \psi_{a',b'}(t) \frac{da'db'}{a'} \qquad \text{Eq 8.3}$$

where

$$C_{\psi} = \int_{-\infty}^{\infty} \frac{|\Psi(\omega)|^2}{|\omega|} d\omega \qquad \text{Eq 8.4}$$

where $\Psi(\omega)$ represents the Fourier transform of $\psi(t)$.

A condition called the admissibility condition [39], [51], is required for $\psi(t)$ to be a wavelet function. That is

$$C_{\psi} = \int_{-\infty}^{\infty} \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < \infty \qquad \text{Eq 8.5}$$

If $\Psi(\omega)$ is continuous, which is the case if $\psi(t)$ is an absolutely integrable function, then C_{ψ} can be finite only if

$$\Psi(0) = 0 \Leftrightarrow \int_{-\infty}^{\infty} \psi(t) dt = 0 \qquad \text{Eq 8.6}$$

This implies that a wavelet must be an oscillatory function with zero mean. Admissibility also suggests that $\psi(t)$ must decay with respect to time. Typical wavelet functions are shown in Figure 8.1.

Note that strictly speaking, wavelet transforms are functions of time and scale, and therefore, they are time-scale representation. Since scale is inversely related to frequency, we will use the phrases "time-frequency" and "time-scale" interchangeably throughout this chapter for interpretational convenience.

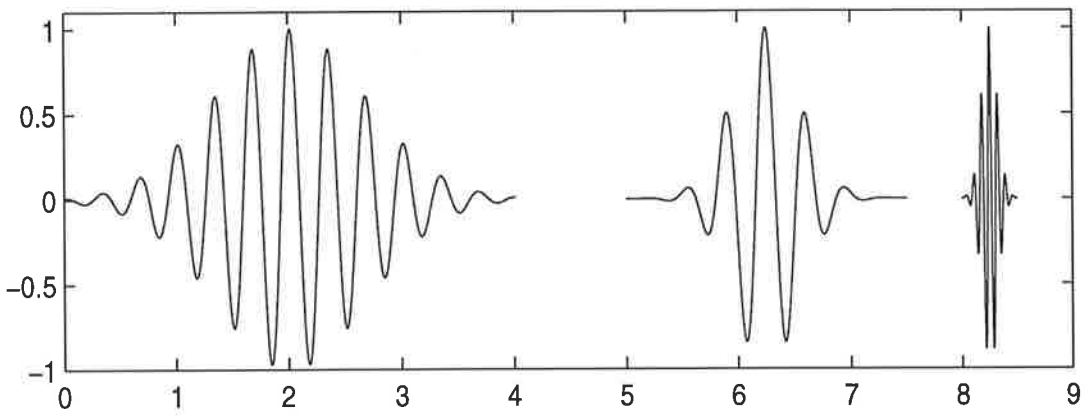


Figure 8.1 Wavelet functions.

8.3 Classification of Wavelets

Wavelets can be categorized into continuous and discrete wavelets on the basis of the values of the dilation and translation parameters, although discrete wavelet transforms are a special case of continuous wavelet transforms. That is, if the dilation and translation parameters vary continuously then continuous wavelet functions are obtained, on the other hand if these parameters are discretized then discrete wavelets are achieved [39]. The continuous wavelet transform is best suited for signal analysis [48], [129], [140], while discrete wavelet transforms are more suitable for signal coding applications.

8.3.1 The Continuous Wavelet Transforms

The continuous wavelet transforms were originally introduced in [48]. In Eq 8.1, if a and b vary continuously over R provided that $a \neq 0$, then the basis functions are continuous and the wavelet transfer of signal $s(t)$ is calculated [27], [60], [118] as

$$\Gamma(a,b) = |a|^{-1/2} \int_{-\infty}^{\infty} s(t)\psi\left(\frac{t-b}{a}\right)dt \qquad \text{Eq 8.7}$$

The CWT of a one-dimensional signal is a two dimensional function of the real variables b (time shift) and a (scale). Eq 8.7 implies that for any particular values of b and a, the wavelet transform assigns a (complex) numerical value to the signal $s(t)$ which quantitatively describes the degree of similarity between the signal and the translated and scaled version of the mother wavelet. Provided that the analyzing function is chosen with good time-frequency localization, then the CWT gives a picture of the time-frequency characteristics of the signal $s(t)$.

There are a number of alternative ways to express the CWT. One is to write it as an integral in the frequency domain as

$$\Gamma(a,b) = |a|^{-1/2}\left[\int_{-\infty}^{\infty} S(\omega)\Psi(a\omega)e^{j2\pi\omega b}d\omega\right] \qquad \text{Eq 8.8}$$

where $S(\omega)$ and $\Psi(\omega)$ are Fourier transforms of $s(t)$ and $\psi(t)$, respectively. We can also express Eq 8.7 as a convolution in the time domain as

$$\Gamma(a,b) = |a|^{-1/2}\left[s(b) * \psi\left(\frac{-b}{a}\right)\right] \qquad \text{Eq 8.9}$$

From Eq 8.7 to Eq 8.9, the wavelet transform $\Gamma(a,b)$ of the signal $s(t)$ may be interpreted as the output of an infinite bank of filters described by the impulse response $\psi(\cdot)$ over a continuous range of scale a.

8.3.2 The discrete wavelet transforms

In general, discrete wavelet transforms (DWT) are generated by sampling the CWT. That is, the dilation and translation parameters are discretized. Discretization of parameters is achieved in the following manner [39].

The scaling parameter can be discretized as

$$a = \alpha^m \qquad m \in Z \qquad \text{Eq 8.10}$$

where the dilation step $\alpha > 0$. For discretization of b one natural choice is to take multiples of a fixed value β. That is

$$b = k\beta\alpha^m \qquad \text{Eq 8.11}$$

where $\beta > 0$ and the condition is $\psi(t - k\beta)$ must cover the whole time. With the above definitions the discrete wavelets are expressed as

$$\psi_{m,k}(t) = \alpha^{-m/2}\psi(\alpha^{-m}t - k\beta) \qquad \text{Eq 8.12}$$

Typically, $\alpha = 2$ and $\beta = 1$ are used in practical applications of wavelets, though they need not be restricted to these values.

8.4 Desirable Properties of Wavelets

In addition to the conditions imposed by Eq 8.5, there are other desirable properties that if satisfied will make the analyzing function a useful wavelet.

- *Smoothness*: $\psi(t)$ must be a smooth function (n times differentiable) with continuous derivatives [104].
- *Good frequency localization*: This property is directly related to vanishing moments of $\Psi(\omega)$ [39], defined as

$$\int_{-\infty}^{\infty} t^l \psi(t)dt = 0 \qquad \text{Eq 8.13}$$

where $l = 0, 1, 2, \ldots, M-1$. Eq 8.13 implies that

$$\frac{d^l}{d\omega^l}\Psi(\omega)\Big|_{\omega=0} = 0 \qquad l = 0, 1, \ldots, M-1 \qquad \text{Eq 8.14}$$

The consequences of this condition are good frequency localization, that is $\Psi(\omega)$ would decay sufficiently fast as $\omega \to \infty$ and $\Psi(\omega)$ will be flat near $\omega = 0$. A larger number of vanishing moments results in more flatness near $\omega = 0$ when ω is small.

- *Good time localization*: This is also related to vanishing moments. In order to have a good time localization it is necessary that $\psi(t)$ together with its derivatives decay rapidly [39], that is

$$\left|\frac{d^l}{dt^l}\psi(t)\right| < \infty \qquad l = 0, 1, \ldots, M-1 \qquad \text{Eq 8.15}$$

which implies that Eq 8.13 must be satisfied.

- *Symmetry*: Symmetric wavelets do not distort the phase information of the signal [77].

If the aforementioned properties and admissibility condition are met then the wavelets behave like a band-pass filter, and the analyzing function $\psi(t)$ would be the impulse response of the filter.

8.5 Analyzing wavelets as constant-Q filters

It was concluded in the previous section that the analyzing wavelets are band-pass filters if they are designed with desired properties. As a matter of fact, wavelet analysis is very similar to the short-time Fourier transform in a sense that they are both windowed spectral analysis techniques, but there is a major difference in the kind of the window used. In the latter the bandwidth of the window function remains constant as it scans the signal, but in the wavelet analysis the bandwidth of the window varies. In other words, with short-time Fourier transform one can only see the forest or the trees but with wavelet analysis it is possible to see both the forest and the trees.

8.5.1 Constant-Q filters

It can be shown that the short-time Fourier transform can be perceived as analysis filters that are frequency-shifted versions of a base filter, where all filters in the filter bank have the same bandwidth [144]. In contrast, wavelet transforms use constant-Q filters which have a constant bandwidth to centre frequency ratio. Let $H(j\omega)$ represent a band-pass filter centered at ω_0 with a bandwidth of $\Delta\omega = \omega_h - \omega_l$ as shown in Figure 8.2 (a). Let the impulse response of the filter be *h(t)*.

The centre frequency of this filter is defined to be the centre of the band-pass region when the frequency axis is logarithmic. That is $\log(\omega_0) = \dfrac{\log(\omega_h) + \log(\omega_l)}{2}$, which implies that the centre frequency is the geometric mean of the cutoff frequencies [144], that is

$$\omega_0 = \sqrt{\omega_l \omega_h} \qquad \text{Eq 8.16}$$

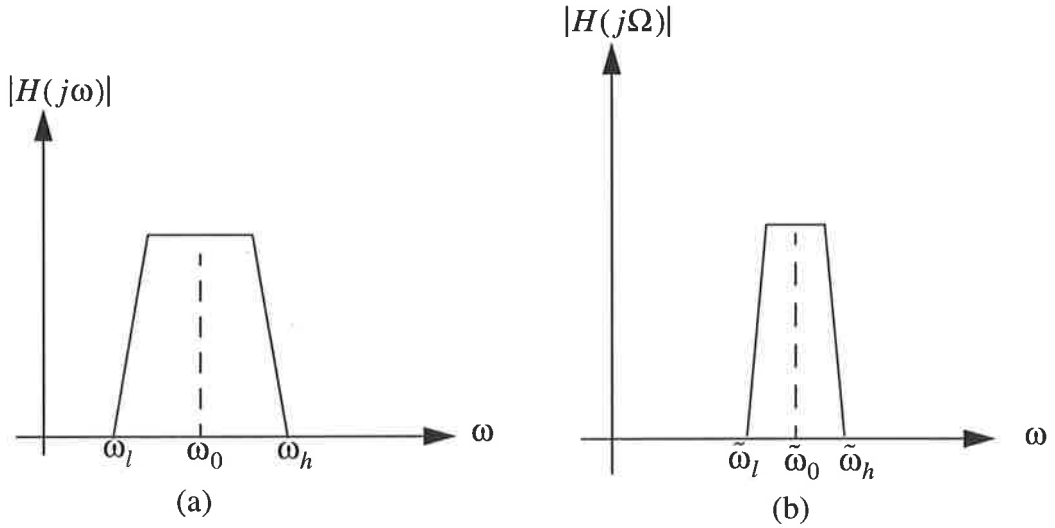


Figure 8.2 a) A bandpass filter, b) frequency-scaled version of (a).

A parameter Q is defined as the ratio of the bandwidth to centre frequency (the inverse of quality factor defined in band-pass filters) as

$$Q = \frac{\Delta\omega}{\omega_0} \quad \text{Eq 8.17}$$

The transfer function of the filter can be altered such that its centre frequency and its bandwidth change but the Q of the filter remains unchanged. Let the scaling factor be α^{-m}, then new cutoff frequencies and consequently the new centre frequency will be

$$\tilde{\omega}_l = \alpha^{-m}\omega_l \quad \text{Eq 8.18}$$

$$\tilde{\omega}_h = \alpha^{-m}\omega_h \quad \text{Eq 8.19}$$

$$\tilde{\omega}_0 = \alpha^{-m}\sqrt{\omega_l\omega_h} \quad \text{Eq 8.20}$$

If Eq 8.18 to Eq 8.20 are inserted into Eq 8.17 then it is clear that the value of Q will stay unchanged. The new transfer function becomes

$$H_m(j\omega) = H(j\alpha^m\omega) \quad \text{Eq 8.21}$$

where it is assumed that $\alpha > 1$. This is shown in Figure 8.2 (b). Eq 8.21 implies that constant-Q filters are realized from a band-pass filter by shifting them to higher/lower frequencies where it changes the bandwidth of the filter.

The impulse response of the new filter is

$$h_m(t) = \alpha^{-m}h(\alpha^{-m}t) \quad \text{Eq 8.22}$$

Constant-Q filters are used as analysis filters in wavelet analysis. As m is increased the impulse response of the filter gets longer in the time domain.

Usually it is desirable that all filters have the same energy [144]; that is, for all values of m the following expression must hold true.

$$\int_{-\infty}^{\infty} h_m^2(t)\,dt = \int_{-\infty}^{\infty} h^2(t)dt \quad \text{Eq 8.23}$$

Eq 8.23 implies that

$$h_m(t) = \alpha^{-m/2} h(\alpha^{-m} t) \qquad \text{Eq 8.24}$$

$$H_m(j\omega) = \alpha^{m/2} H(j\alpha^m \omega) \qquad \text{Eq 8.25}$$

Eq 8.25 means that as the filter is moved to higher frequencies, its height is decreased in order to maintain the same energy. This is illustrated in Figure 8.3. This figure explains how the frequency scaling can change the bandwidth of a band-pass filter.

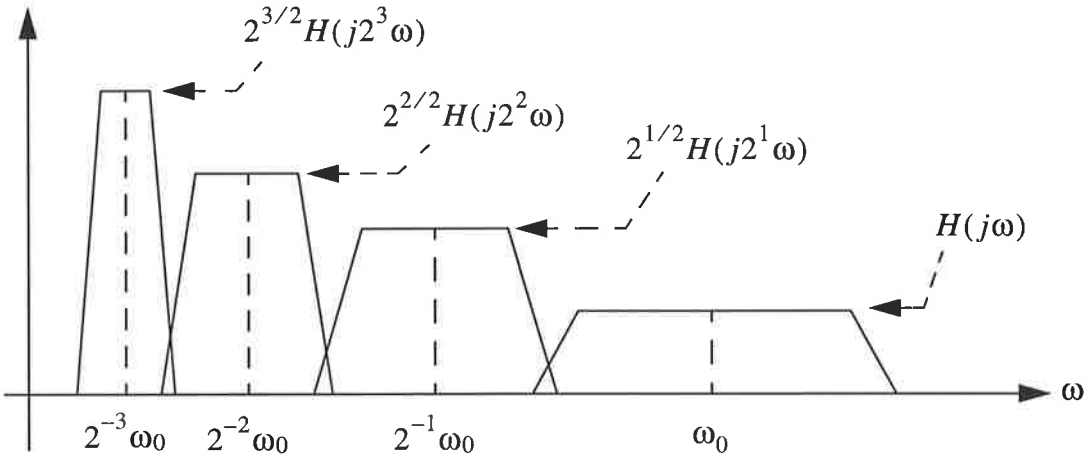


Figure 8.3 Successive constant-Q filters for m=2.

8.6 Time-Frequency Localization

In section 6.3, the time-frequency localization of STFT was explained with regard to the window function $w(t)$. A similar argument follows here. Associated with each wavelet function there is time localization and frequency localization, defined respectively as

$$\Delta t = \left\{ \int_{-\infty}^{\infty} (t - t^*)^2 |\psi(t)|^2 dt \right\}^{1/2} \qquad \text{Eq 8.26}$$

$$\Delta\omega = \left\{ \int_{-\infty}^{\infty} (\omega - \omega^*)^2 |\Psi(j\omega)|^2 d\omega \right\}^{1/2} \qquad \text{Eq 8.27}$$

where t^* and ω^* are defined as

$$t^* = \int_{-\infty}^{\infty} t|\psi(t)|^2 dt \qquad \text{Eq 8.28}$$

$$\omega^* = \int_{-\infty}^{\infty} \omega|\Psi(j\omega)|^2 d\omega \qquad \text{Eq 8.29}$$

The wavelet $\psi(t)$ is centered around (t^*, ω^*) in the time-frequency plane and contains information in the time interval $(t^* - \Delta t, t^* + \Delta t)$ and frequency interval $(\omega^* - \Delta\omega, \omega^* + \Delta\omega)$. The two intervals define a time-frequency cell whose area is restricted by the uncertainty principle, Eq 6.12. Translation of the wavelet function does not change the time and frequency localization, but dilation will change them as follows [144]

$$\Delta t_a = a\Delta t \qquad \text{Eq 8.30}$$

$$\Delta\omega_a = \frac{\Delta\omega}{a} \qquad \text{Eq 8.31}$$

Therefore, good time localization can be achieved for small values of a, which corresponds to high frequencies, or small scales. On the other hand, good frequency localization is obtained for large values of a.

8.7 Basis function selection

There are a variety of wavelet functions, each with different properties. The choice of the wavelet in an application is usually made by matching the properties of the wavelets with those required by the application. Not all wavelet functions have the desired properties of wavelets. It is not possible to construct a wavelet with an arbitrary number of desired properties.

We will not discuss the wavelet design procedures here but rather we will focus on the selection of a particular wavelet for the heart sound analysis. Three wavelets in the time and frequency domains are examined: Morlet, Meyer, and Daubechies-4 wavelets. Their time and frequency responses are shown in Figure 8.4. As was explained earlier and shown in Chapters 6 and 7, we are examining the time-frequency patterns of the heart sounds in the time-frequency plane. Analysis involves finding the frequency difference in frequency patterns in the time-frequency plane, hence it is necessary to have good time-frequency resolution.

There is a closed form expression for $h(t)$ and $H(j\omega)$ of the Morlet wavelet and its relative bandwidth is adjustable. The Morlet wavelet is explained in greater detail in section 8.7.1. In Figure 8.4 (a), the Morlet wavelet is shown in the time and frequency domains. There is no closed form expression for the Meyer wavelet. Its structure is similar to the Morlet wavelet but it has a fixed relative bandwidth and there are no free parameters for adjusting it. In Figure 8.4 (b), the time domain and the frequency domain representation of the Meyer wavelet are shown. Information on design procedures can be found in [131]. Unlike the other two wavelets, the Daubechies wavelets are not symmetric in the time domain. They are generated using finite impulse response filters proposed by Daubechies [131]. The relative bandwidth of the Daubechies wavelets are also fixed and there is no closed form expression for them. In Figure 8.4 (c), a Daubechies-4 wavelet is shown in the time and frequency domains.

In Figure 8.4, the left-hand plots are the time domain representation of the wavelets while the right-hand plots are the frequency domain representation of the corresponding left-hand wavelets. In order to make the comparison easy, we have normalized the frequency responses of the wavelets so that their maxima have unity magnitude and their centre frequencies are equal to one. The Morlet and the Meyer wavelets are generally complex valued, only their real parts are shown in Figure 8.4.

We have chosen the Morlet wavelet for the analysis of the heart sounds. The main reason behind our choice is the degree of freedom for adjusting the relative bandwidth of the analyzing wavelet. Figure 8.4 shows that the relative bandwidth of the

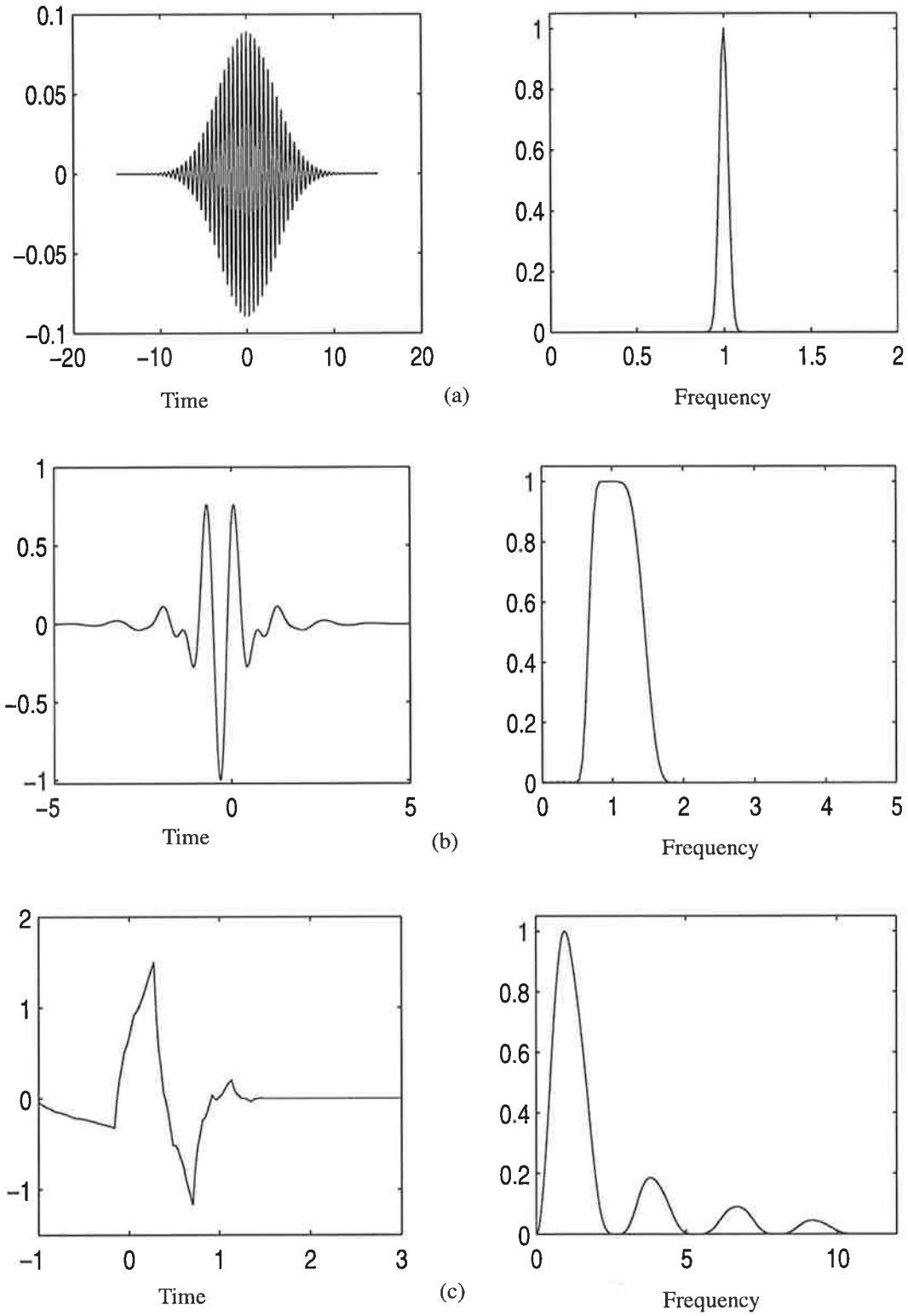


Figure 8.4 a) The Morlet wavelet, b) The Meyer wavelet c) The Daubechies wavelet.

Daubechies-4 wavelet and the Meyer wavelet are larger than the relative bandwidth of the Morlet wavelet. Furthermore, the relative bandwidth of the Morlet wavelet can be altered while the others are fixed.

8.7.1 Morlet wavelet

The Morlet wavelet is one of the wavelet functions that are often used in CWT applications. The Morlet wavelet [50], [125], is actually a modulated Gaussian expressed as

$$\psi(t) = \frac{1}{\sqrt{2\pi/\beta}} \cdot e^{j2\pi f_0 t - \beta t^2/2} \qquad \text{Eq 8.32}$$

where f_0 is the centre frequency and β is the bandwidth of the filter. The Fourier transform of the filter can be computed as

$$\Psi(f) = e^{-2\pi^2 (f - f_0)^2/\beta^2} \qquad \text{Eq 8.33}$$

This shows that the maximum value of the frequency response of the Morlet wavelet is equal to one and occurs at f_0. In other words, $\Psi(f_0) = 1$. The relative bandwidth of the filter is calculated as $RBW = \frac{2\sqrt{2\beta}}{f_0}$. The principle feature of the Morlet wavelet, of interest to us is that its relative bandwidth is easily set. This is important in obtaining a desired frequency resolution in CWT at a particular frequency. In Figure 8.4 (a), a Morlet wavelet and its Fourier transform with parameter values of $f_0 = 5$ and $\beta = 0.05$ are shown. The Morlet wavelet has two theoretical limitations which are not of much concern in some practical cases, but they are worth mentioning. First, $\Psi(0) \neq 0$ which implies that the admissibility condition is not met. However; for small values of bandwidth, $\Psi(0)$ will be small enough for practical purposes. Second, the Morlet wavelet cannot be used to form an orthonormal basis function which is of concern if perfect reconstruction of the signal using filter bank is desired.

8.8 CWT Implementation

For the calculations of the CWT, we have used the fast algorithm proposed in [65]. In this technique instead of using the time domain convolution, the frequency domain multiplication is utilized. To see how it is done, consider Eq 8.8 in which the term inside the brackets is the inverse Fourier transform of $S(\Omega)\Psi(a\Omega)$. This suggests that, the CWT can be computed by taking the Fourier transform of the signal and the wavelet function and then finding the inverse Fourier transform of the product term.

Let $s(n)$ represent N_s samples of $s(t)$, and $\psi(n)$ represent N_ψ samples of $\psi(t)$, both sampled at or above the Nyquist rate. Discretization of Eq 8.8 with uniform sampling in time, and arbitrarily spaced sampling in scale yields [65]

$$\Gamma(n, a_i) = \eta|a_i|^{1/2} \sum_{k=-M/2}^{(M/2)-1} S\left(\frac{2\pi}{M}k\right)\psi\left(\frac{2\pi}{M}a_i k\right)e^{j(2\pi/M)kn} \qquad \text{Eq 8.34}$$

where η is a constant, a_i is the discretized scale parameter chosen arbitrarily with $i = 1, \ldots, N_a$. The value of M is chosen such that to ensure frequency multiplication of $S(\Omega)$ and $\Psi(a\Omega)$ results in linear convolution in the time domain rather than circular convolution; that is, M must be greater than the sum of the lengths of $s(n)$ and the largest wavelet basis function used. Usually the next highest power of 2 greater than $N_s + N_\psi(max|a_i| - 1)$ is used.

The chirp z-transform algorithm [91], is used for computing $\psi\left(\frac{2\pi}{M}a_i k\right)$. Using the definition of the discrete Fourier transform and setting $p = k + M/2$, the chirp z-transform of the sequence $\psi(n)$ is obtained as

$$\psi\left(\frac{2\pi}{M}a_i k\right) = \psi\left(\frac{2\pi}{M}a_i k\right) = \sum_{n=0}^{M-1} \psi(n)e^{-j(2\pi/M)a_i pn}e^{j\pi a_i n} \qquad \text{Eq 8.35}$$

where $p = 0, \ldots, M-1$.

8.9 Time-Scale Representation

A program was written based on the discussion of the previous section. The parameter values in Eq 8.33 were set to $\beta = 0.05$ and $f_0 = 5$. This sets the relative bandwidth of the wavelet equal to 0.1265. The time domain and the frequency domain representation of this wavelet are shown in Figure 8.4 (a).

Calculation of the wavelet coefficients were carried out on an octave-by-octave basis [50]. That is for integer values of dilation, the frequency localization which is called an octave, varies logarithmically. Each octave is further decomposed into smaller sections called voices according to

$$\psi^{\vartheta}(t) = 2^{-(\vartheta-1)/\kappa}\psi(2^{-(\vartheta-1)/\kappa}t) \qquad \text{Eq 8.36}$$

where κ represents the number of voices per octave. One such time-scale plane is shown in Figure 8.5. In this figure four voices per octave are used. Voices are designated by "+"and octaves are represented by "o". The interval between two consecutive octaves is partitioned into four logarithmically spaced intervals along the frequency axis while the time axis intervals are left unchanged.

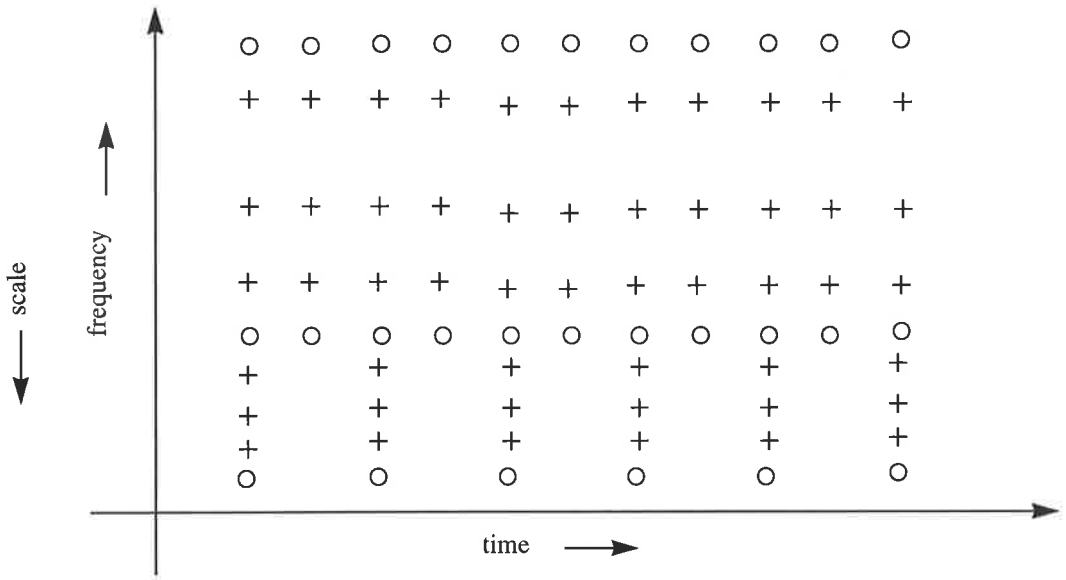


Figure 8.5 Time-frequency plane decomposition.

8.10 Wavelet analysis of heart sounds

In this section we explain the process of wavelet based analysis of recorded heart sounds. Prior to CWT calculations, the recorded phonocardiograms were passed through MALE filter for noise cancellation. Then the algorithms of Chapter 4 were used to prepare the hear-beat signals. Analytical signals were used in calculations of CWT. In order to demonstrate the analysis procedures and for the sake of comparisons, we have chosen the same signal that we used in STFT and Choi-Wiliams analyses. These signals are shown in Figure 7.4. We have performed continuous wavelet analysis using the technique explained in section 8.8. The number of octaves were calculated from the length of the signal used. For example, if the length of signal was 2000 samples, then the signal was padded with zeros to obtain the dyadic length closest to 2000, i.e. 2048. Since $2048 = 2^{11}$, the number of octaves was set equal 8. The number of octaves was set to 3 figures less than dyadic length of signal (zero padded signal if the original signal length was not dyadic). The number of voices was set to 32 in order to have less redundancy. The moduli of the CWT of the signals were calculated. In the modulus of the CWT matrix the parts that corresponded to the zero-padded section in the time axis were discarded for all frequencies. The CWT matrix is a time-scale matrix. In order to represent it in 3-D, we calculated a vector which contained centre frequencies of the filters associated with each voice level. This vector was used for time-frequency plots of the modulus of the CWT. These plots are shown in Figure 8.7-1, Figure 8.7-2, Figure 8.8-1, and Figure 8.8-2. In these figures, the x-axis represents time in seconds, the y-axis represents frequency in hertz and the z-axis is the amplitude of the modulus of the CWT matrix. The organization of these figures in the time-frequency planes are shown in Figure 8.6.

8.10.1 Locations of concentration of energy in CWT matrix

The local maxima detector algorithm of Chapter 5 is applied to the modulus of the CWT time-scale matrix of the signals of Figure 7.4 (a) and (b), and locations of concentration of energy are obtained. The time-scale matrices are mapped to time-frequency matrices using the frequency vector that was calculated in the process of CWT

calculations (explained in previous section). Then the three time, frequency, and magnitude vectors were obtained (refer to section 6.5 for more details on these vectors). These vectors show that the number of local maxima before angioplasty is 82 and after angioplasty is 79. Our preliminary findings showed that 90% of the energy of the signals were concentrated below 250 *Hz* in most cases; therefore, we have restricted this analysis to that frequency range.

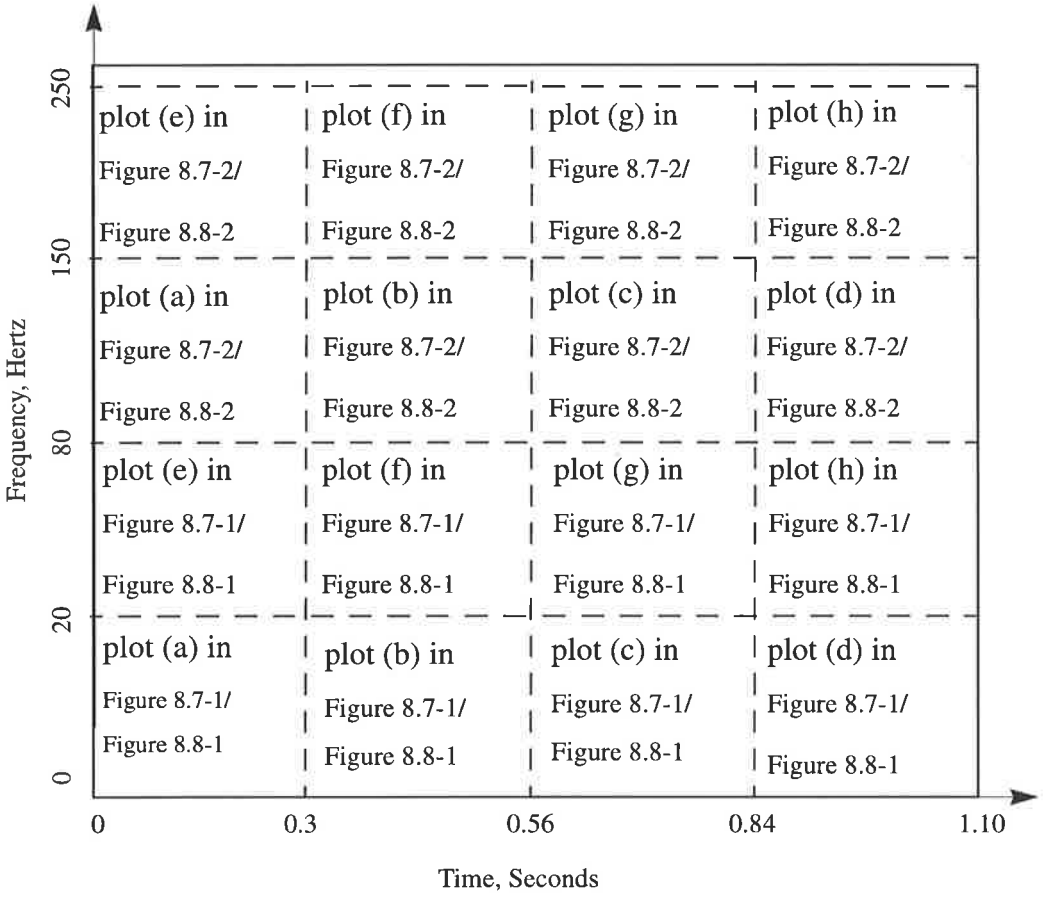


Figure 8.6 Organization of Figure 8.7-1, Figure 8.7-2, Figure 8.8-1, and Figure 8.8-2 in the time-frequency plane.

8.10.2 Time-frequency patterns in CWT matrix

In order to investigate the resemblance of the modulus of the CWT matrices, we carried out the same procedure as was previously explained in Chapter 6. That is,

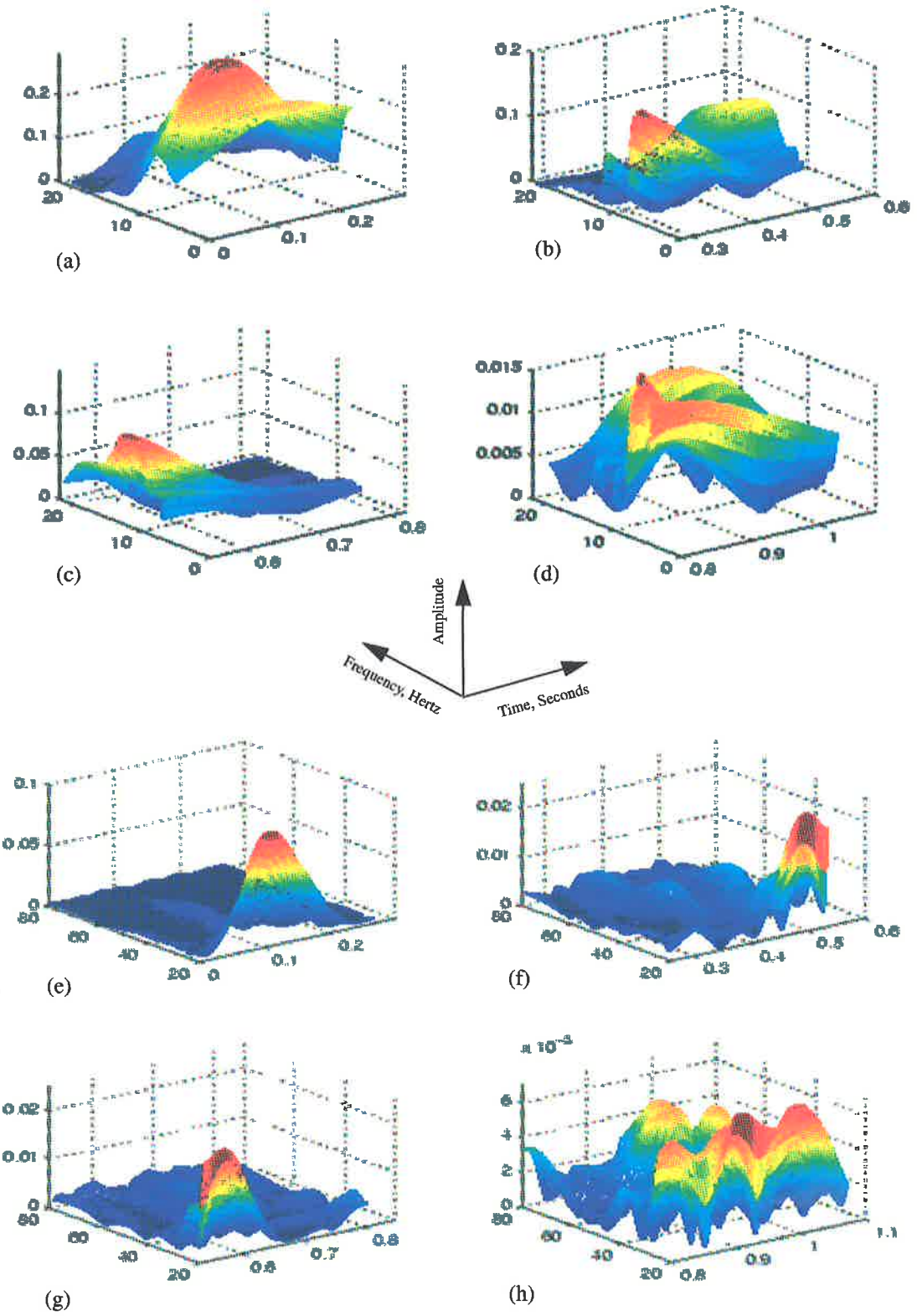


Figure 8.7-1 The modulus of the CWT of the signals of Figure 7.4 (a).

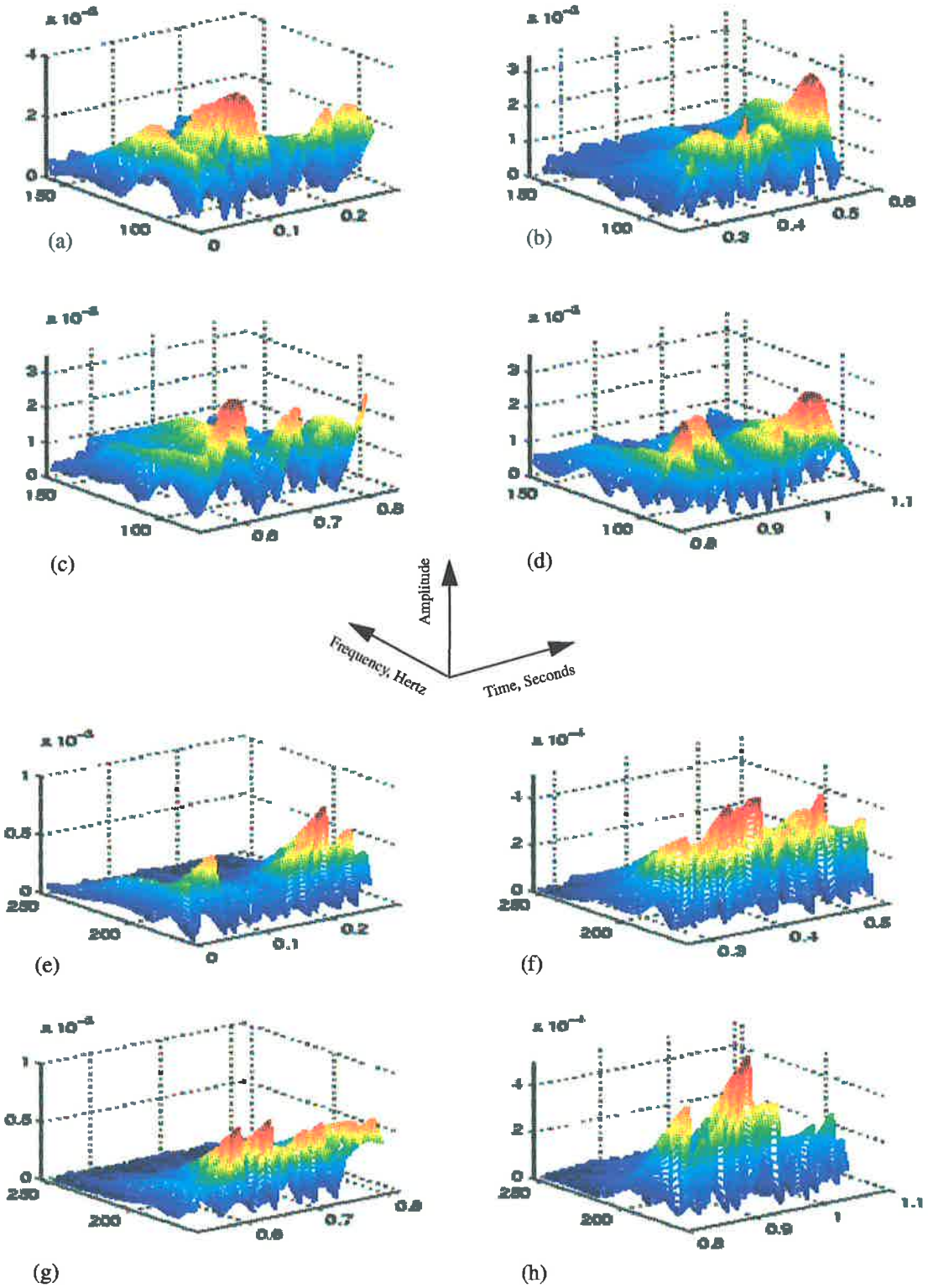


Figure 8.7-2 The modulus of the CWT of the signals of Figure 7.4 (a).

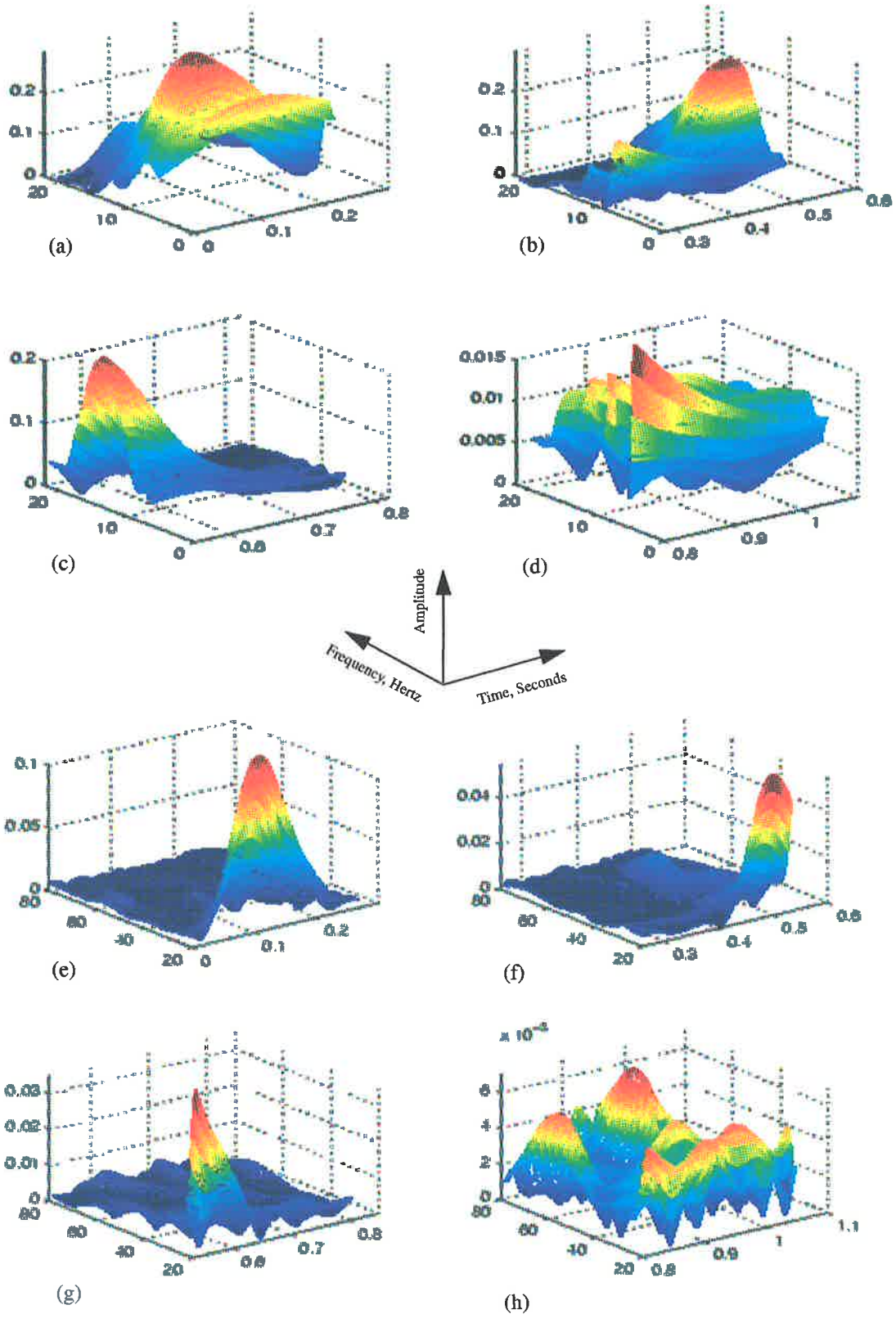


Figure 8.8-1 The modulus of the CWT of the signals of Figure 7.4 (b).

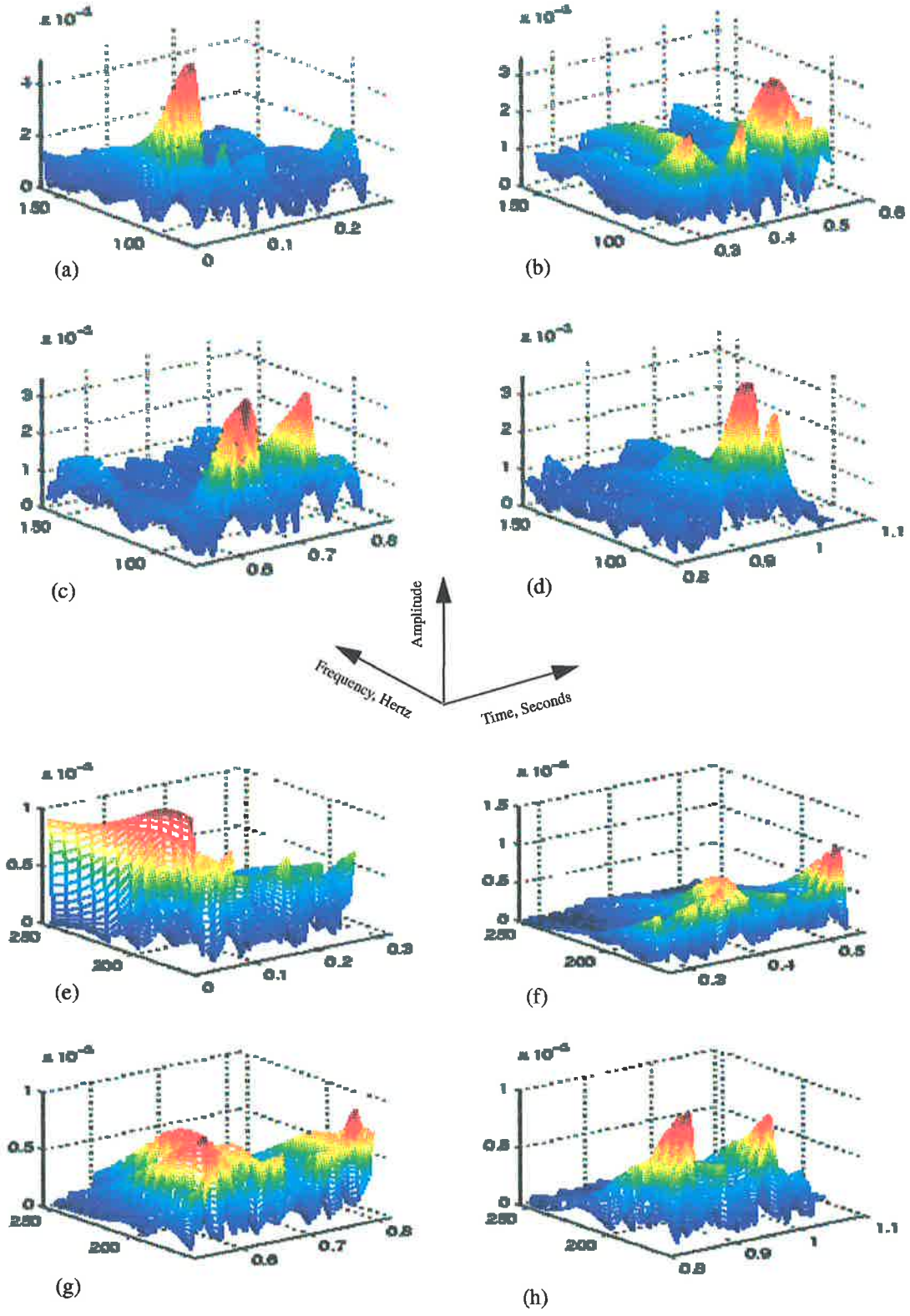


Figure 8.8-2 The modulus of the CWT of the signals of Figure 7.4 (b).

we now rearrange the time-frequency vectors in ascending order of frequency. Then the time-frequency-magnitude product vectors of before and after angioplasty are calculated.

The obtained TFM vectors are plotted in Figure 8.9. The solid line represents the TFM of the heart sound before angioplasty and the dashed-line shows the TFM of the heart sound after angioplasty. The vertical axis represents the magnitude of the TFM vector and the horizontal axis represents the local maximum number. The highest peaks in the graphs correspond to the highest energy point in the CWT matrices.

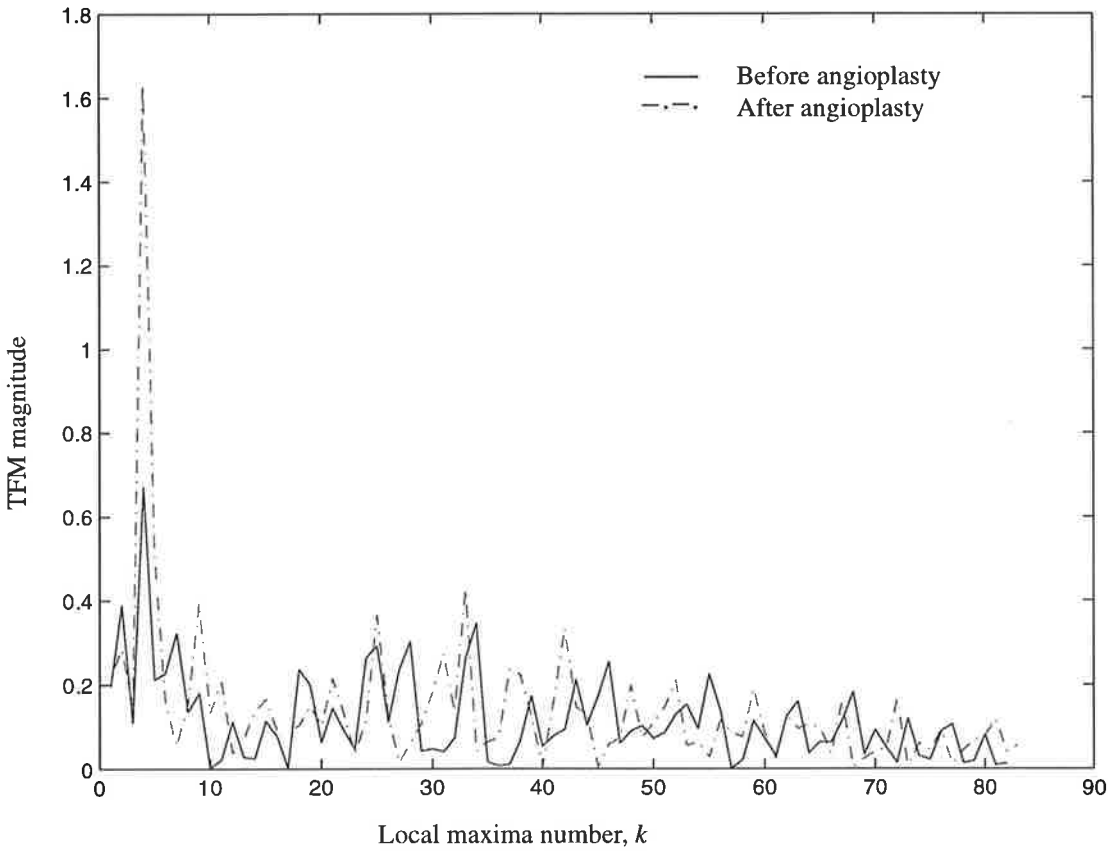


Figure 8.9 TFM graphs of the modulus of the CWT of the signals of Figure 7.4.

We calculated the average of the magnitude vectors and it was found that the energy of the signal after angioplasty is higher than before angioplasty.

Upon comparing the two graphs of Figure 8.9, many similar sections were revealed. We first compared the two signals visually and selected those sections that looked similar. These similar sections represent similar patterns of local maxima in the modulus of the CWT matrix. Then we plotted the corresponding time-frequency patterns and compared them together. This process was repeated until all the similar patterns shown in Figure 8.10 were obtained. From the total of 82 and 79 local maxima detected in the modulus of CWT matrices before and after angioplasty, seven similar sections were extracted and plotted in Figure 8.10; while the local maxima numbers for these sections are shown in Table 8.1.

Table 8.1

Local maxima			Local maxima			Local maxima	
before	**after**		**before**	**after**		**before**	**after**
1-5	1-7		6-14	8-18		15-24	19-30
25-35	31-40		36-46	41-52		47-65	53-70
66-79	71-82						

The time-frequency patterns of the local maxima corresponding to Table 8.1 are plotted in Figure 8.11. In this figure "$*$" represents the local maxima for the signal before angioplasty and "o" represents the local maxima for the signal after angioplasty. It is clear that there are similar patterns in the time-frequency matrices of before and after angioplasty. In order to show resemblances of the patterns in the time-frequency plane we have connected the local maxima points to each other. This facilitated recognition of the resemblance of the two patterns.

For this particular case, the time-frequency patterns are spread in the frequency range between zero and 220 Hz. We calculated the average frequencies of the patterns in first, second, and third heart sound regions. We found that there is a frequency shift between the average frequencies of the patterns before and after angioplasty; that is, the average frequency of the patterns belonging to the signals after angioplasty are

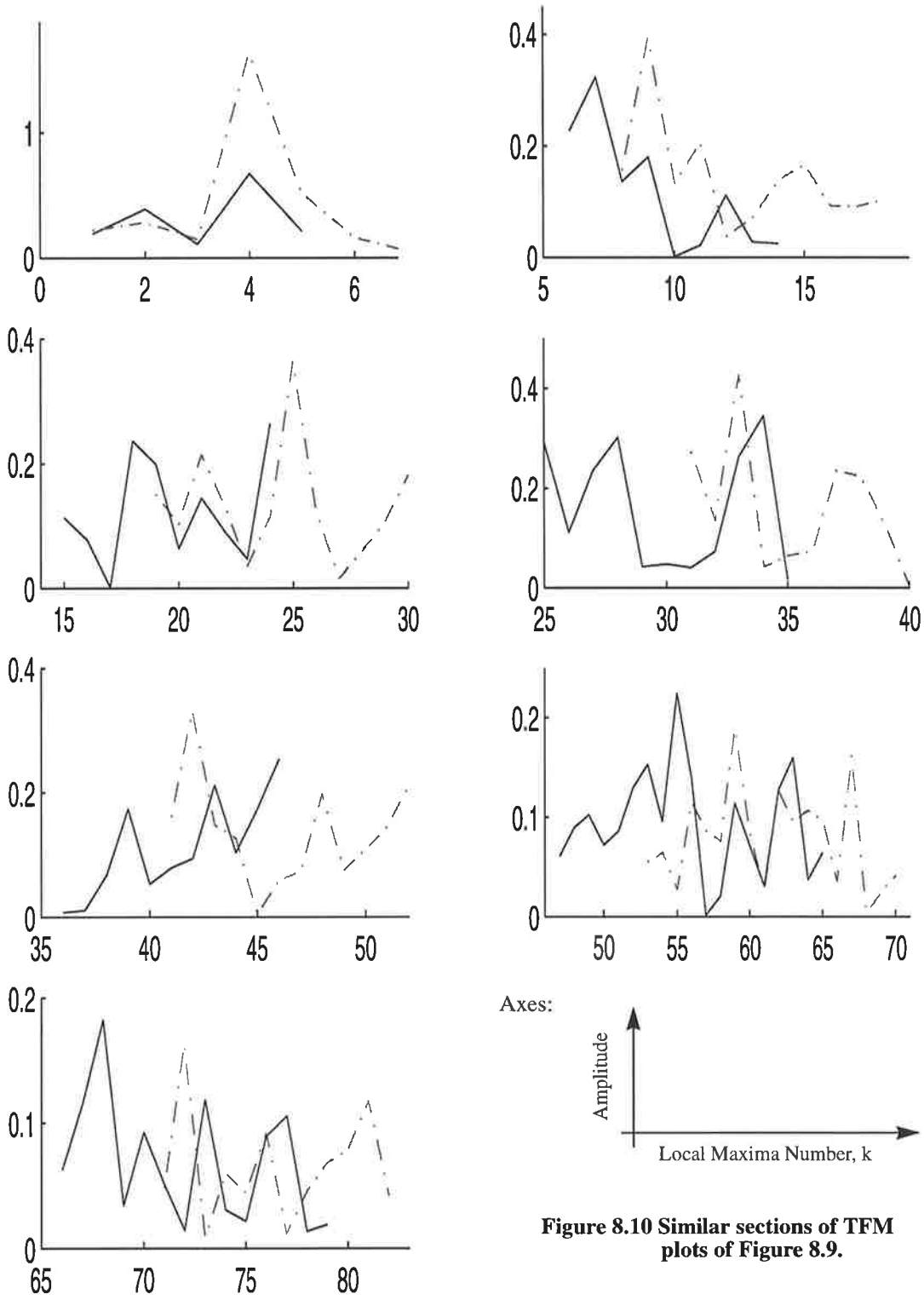


Figure 8.10 Similar sections of TFM plots of Figure 8.9.

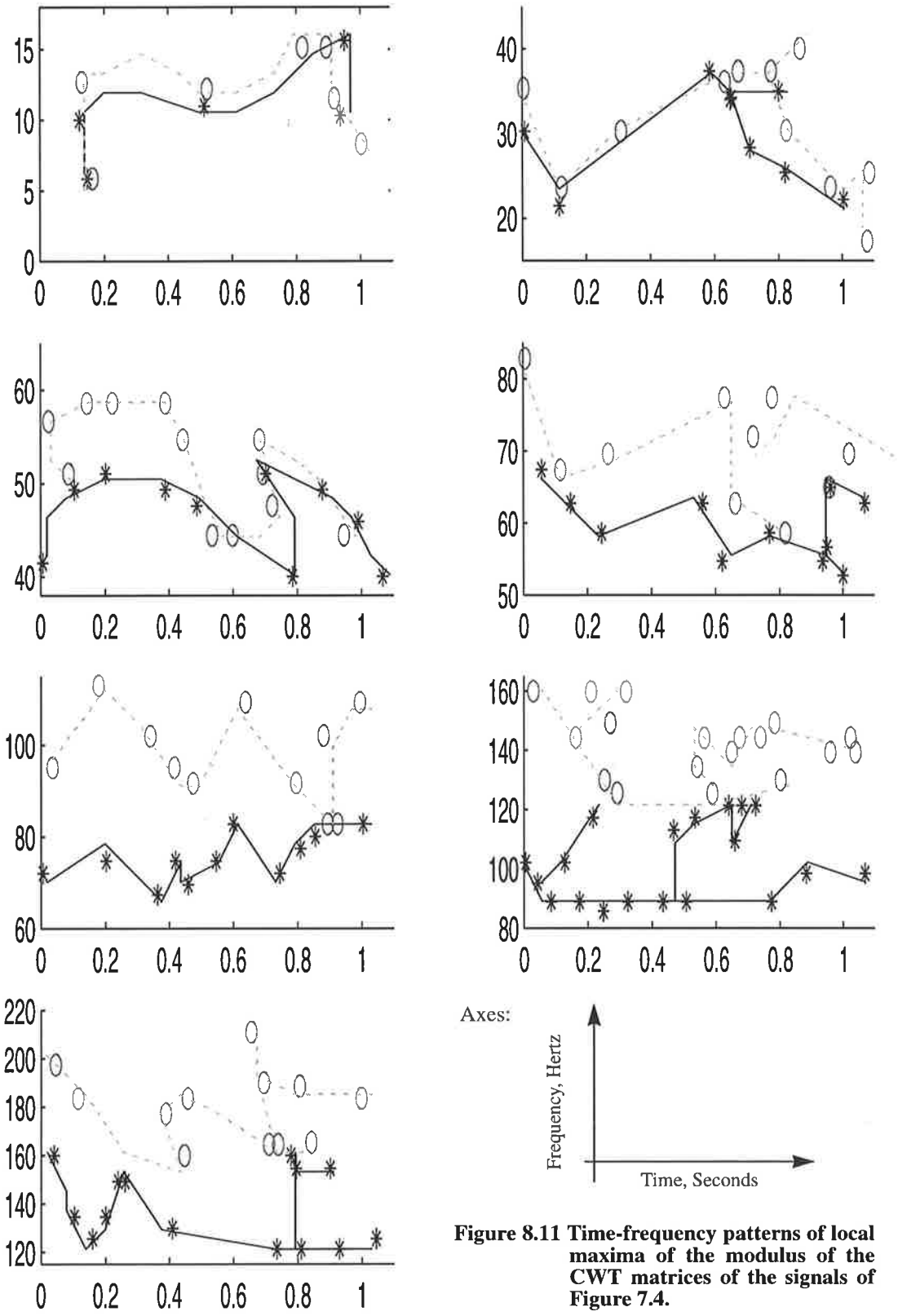


Figure 8.11 Time-frequency patterns of local maxima of the modulus of the CWT matrices of the signals of Figure 7.4.

higher than before angioplasty. It is also noticeable that the frequency shift is greater in the higher frequency range patterns. The average frequency of each pattern in first, second, and third heart sound regions are calculated and listed in Table 8.2.

Table 8.2 Average frequencies of patterns of Figure 8.11.

frequency range	1st HS MF		2nd HS MF		3rd HS MF		HS pattern MF		
	before	after	before	after	before	after	before	after	shift
0-20	8	9	10	12	13	13	10	11	1
15-45	26	29	30	29	28	29	28	29	1
40-60	45	56	42	57	40	53	42	55	13
45-90	58	73	56	75	54	76	56	75	19
60-120	71	112	72	106	68	101	70	106	36
80-165	97	151	98	145	90	145	95	147	52
115-250	139	191	124	183	128	176	130	183	53

In Table 8.2, the first column represents frequency bands in which patterns of signals in Figure 8.9 were identified. The second and third columns are average frequencies of the patterns in the first heart sound region before and after angioplasty, respectively. Likewise, the fourth and fifth columns correspond to the second heart sound region before and after angioplasty, respectively. The sixth and seventh columns correspond to the third heart sound regions before and after angioplasty, respectively. The eighth and ninth columns show the average frequencies of the patterns before and after angioplasty, respectively. The last column shows the difference between the average frequencies of the patterns. For example, in the frequency band 0-20 *Hz* the first entry of the first column is 8 Hz, which indicates the average frequency of the first heart sound in the first pattern before angioplasty. Likewise, the first entry eighth column is 10, which is the average of 8, 10, and 13. This means that the average frequency

of the whole pattern before angioplasty is 8 *Hz*. The averages were rounded off to their closest integers.

The second, third, fourth, fifth, and sixth columns of Table 8.2 are plotted against the average frequencies of the patterns in Figure 8.12. In Figure 8.12 (a), (b), and (c) the average frequencies of the first, second, and third heart sounds before and after angioplasty are shown, respectively. It is clear from the figures, that there is a shift in the average frequencies after angioplasty, where the shift in the average frequencies increases with frequency.

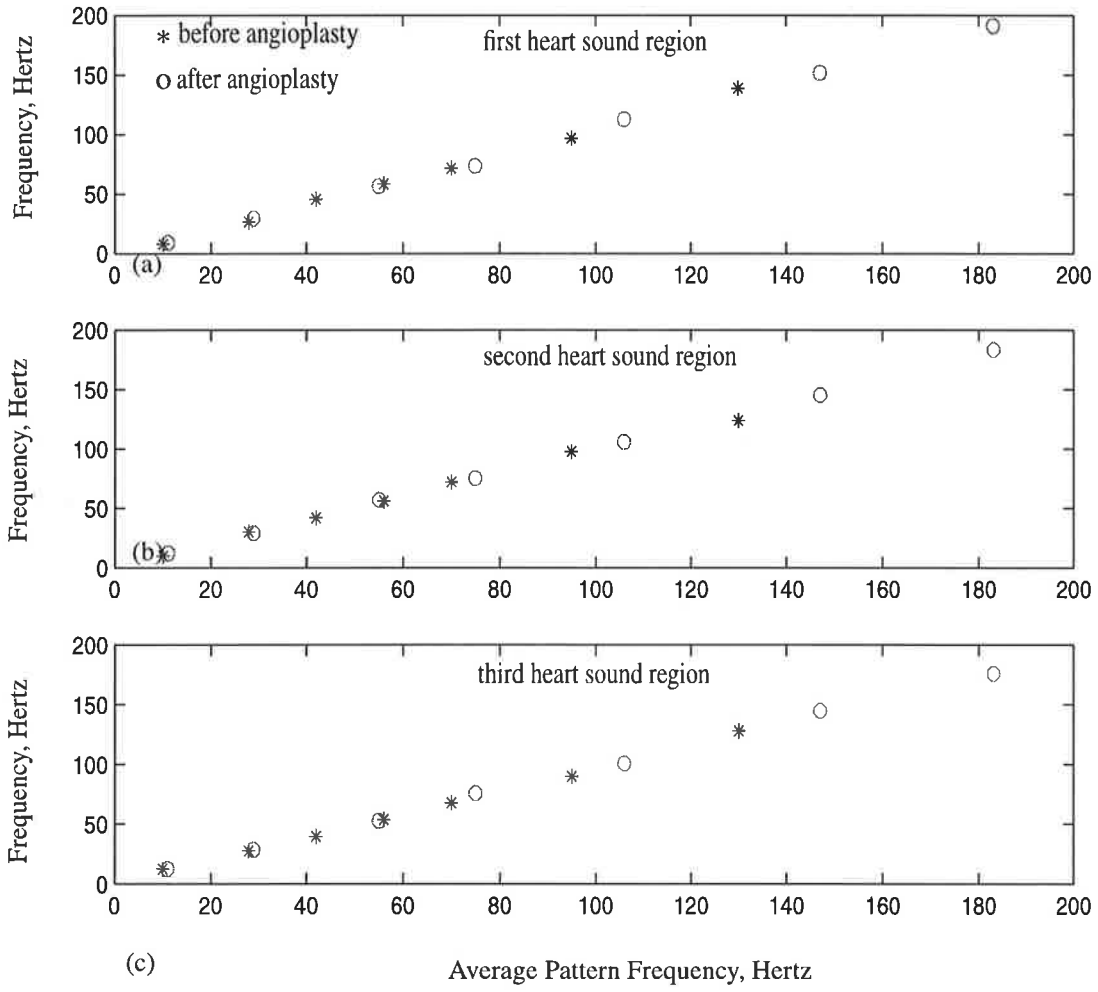


Figure 8.12 Average frequencies in a) first, b) second, and c) third heart sounds obtained from Table 8.2.

8.11 Results from other cases

We performed the analysis for all of the signals in Table 2.2. The results obtained from five heart sounds are shown in Table 8.3. The results from the whole data-set are summarized in Table 8.4. The number of patterns in the modulus of CWT matrices of the signals varies from a minimum of 5 to a maximum of 9. In Table 8.4 we have shown the average pattern frequencies of before and after angioplasty in the first, second, and third heart sound regions in the frequency band:, below 50 *Hz*, between 50 and 100 *Hz*, above 100 and below 150 *Hz*, and above 150 up to 250 *Hz*.

Table 8.3 Average frequencies in time-frequency patterns in five subjects.

Frequency Band, *Hz*	1st HS MF		2nd HS MF		3rd HS MF		HS pattern MF		
	before	after	before	after	before	after	before	after	shift
0-20	8	8	9	11	11	13	9	11	2
17-40	25	28	27	30	29	32	27	30	3
35-58	42	51	38	47	45	54	42	51	9
40-85	65	80	59	76	51	71	58	76	18
62-115	78	111	75	103	69	102	74	105	31
85-175	102	149	92	139	122	168	105	152	47
125-235	168	215	32	181	171	228	157	208	51
01-22	7	8	10	10	12	11	10	10	0
15-48	30	33	32	36	31	32	31	34	3
38-60	41	56	39	46	40	55	40	52	12
50-93	53	62	65	89	59	82	59	78	19
80-145	87	121	92	128	101	139	93	129	36
138-212	141	187	152	198	154	207	149	197	48
0-25	9	10	9	11	13	15	10	12	2
18-45	27	29	29	29	28	31	28	30	2
39-65	48	59	41	56	49	62	46	59	13
48-89	52	72	54	70	63	85	56	76	20
67-126	81	120	68	108	79	112	76	113	37
90-155	97	145	103	146	108	151	103	147	44
123-195	138	186	129	177	142	191	136	185	49
160-250	164	216	179	238	187	243	177	232	55

Table 8.3 Average frequencies in time-frequency patterns in five subjects.

Frequency Band, *Hz*	1st HS MF		2nd HS MF		3rd HS MF		HS pattern MF		
	before	after	before	after	before	after	before	after	shift
0-20	8	9	11	13	13	14	11	12	1
16-43	27	33	23	28	32	36	27	32	5
37-59	40	56	41	56	44	58	42	57	15
45-94	58	86	49	74	52	71	53	77	24
71-130	81	115	83	112	92	126	85	118	33
92-178	106	151	118	161	128	173	117	162	45
148-234	155	196	164	204	178	221	166	207	41
0-25	7	9	10	13	12	14	10	12	2
20-50	33	34	35	35	37	41	35	37	2
32-60	39	50	51	58	44	57	45	55	10
46-93	61	76	54	67	55	69	57	71	14
67-120	84	106	71	95	91	115	82	105	23
84-158	90	122	103	132	118	149	104	134	30
131-200	138	181	146	189	151	196	145	189	44
166-250	174	215	182	239	191	246	182	233	51

Table 8.4 Average frequencies before angioplasty and after angioplasty.

Frequency bands	average pattern frequency in 1st heart sound		average pattern frequency in 2nd heart sound		pattern frequency in 3rd heart sound		Shift in average frequency
	before	after	before	after	before	after	
Below 50 *Hz*	Min.: 6	7	Min.: 8	8	Min.: 10	11	
	Max.: 38	41	Max.: 40	47	Max.: 39	48	5
	Ave.: 19	20	Ave.: 20	23	Ave.: 26	29	
50-100 *Hz*	Min.: 55	62	Min.: 51	50	Min.: 50	50	
	Max.: 78	95	Max.: 82	98	Max.: 79	98	17
	Ave.: 61	74	Ave.: 62	76	Ave.: 59	82	
100-150 *Hz*	Min.: 102	101	Min.: 101	101	Min.: 100	103	
	Max.: 141	148	Max.: 140	145	Max.: 142	148	27
	Ave.: 111	138	Ave.: 118	141	Ave.: 112	143	

Table 8.4 Average frequencies before angioplasty and after angioplasty.

Frequency bands	average pattern frequency in 1st heart sound before after	average pattern frequency in 2nd heart sound before after	pattern frequency in 3rd heart sound before after	Shift in average frequency
150-250 *Hz*	Min.: 153 162 Max.: 205 225 Ave.: 173 198	Min.: 152 159 Max.: 213 234 Ave.: 178 206	Min.: 152 151 Max.: 236 247 Ave.: 171 218	34

Consider, for example, the entries corresponding to the region below 50 *Hz* in Table 8.4. The first row indicates the minimum average frequency found in the patterns representing the first heart sound was 6 *Hz* before angioplasty and 7 *Hz* after angioplasty. Likewise, the second row indicates that the maximum average frequency found in the patterns representing the first heart sound was 38 *Hz* before angioplasty and 41 *Hz* after angioplasty. The mean of all of the average frequencies of the patterns that fit into a particular category was calculated and is shown in the third row of each column in each category of Table 8.4. The last column of the table shows the frequency shift in the average pattern frequencies.

8.12 Conclusions

In this chapter we briefly reviewed the wavelet transform technique as a potential tool for time-frequency analysis. We discussed the Morlet wavelet and used it as the basis function for CWT for the analysis of the heart sounds. We pointed out that continuous wavelet transform has been useful in signal analysis in various fields. The wavelet transform is essentially convolution of the signal with a particular filter. The filter is actually the wavelet function and must have zero mean. This ensures that only the features of the signal comparable with the filter are picked up and other features that fall outside of the filter's characteristics are averaged out. We showed that the

Morlet wavelet is attractive in time-frequency analysis of the heart sounds because its relative bandwidth is adjustable.

The heart-beat cycle of heart sounds before and after angioplasty were prepared and their CWT were calculated. A program was written where a computationally efficient method in the frequency domain was used for CWT calculations. In order to map the CWT time-scale matrix to time-frequency matrix, we generated a frequency vector whose elements were the central frequencies of the wavelet functions in the relevant scale levels.

The local maxima detector algorithms were used to determine the locations of concentration of energy in the CWT plane. Three vectors, temporal, spectral, and magnitude vectors, were used to calculate the time-frequency-magnitude product vectors. The TFM product vectors for before and after angioplasty were compared and similar sections were identified. Using similar sections, the time-frequency patterns of before and after angioplasty were plotted. It was found that two patterns in the same frequency band do have resemblances and there exist some frequency shift between them.

The average frequencies in first, second, and third heart sound regions were calculated and compared. These plots show that the average frequency before and after angioplasty have similar variations.

In the low frequency range, below 50 *Hz*, the average frequency shift is less than 10 *Hz*. For the medium frequency range, between 50 *Hz* and 100 *Hz*, the average frequency shift is also below 25 *Hz* and in the high frequency range, above 100 *Hz* and up to 150 *Hz*, the average frequency shift is above 25 *Hz* and below 50 *Hz*, and in the frequency range above 150 *Hz* and below 250 *Hz* the average frequency shift is above 30 *Hz*.

Chapter 9

Conclusion and Recommendation for Future Research

9.1 Conclusions

The importance of coronary artery disease and the potentials of signal processing in time-frequency methods have given impetus to the investigations of heart sounds from people with coronary artery disease before and after angioplasty. This study was an attempt to characterize various time-frequency representations with respect to their capability to identify the difference in heart sounds with coronary artery disease before and after angioplasty.

Phonocardiograms were recorded in the cardiology department of the Royal Adelaide Hospital from patients with coronary artery disease before and after angioplasty. Recordings were carried out at the patient's bedside. Heart sounds were recorded from three positions on the chest including the apex. Each signal was

recorded for at least 45 seconds. Electrocardiograms were also recorded synchronously.

It is probable that the phonocardiograms were corrupted with noise during the recording process; therefore, adaptive line enhancement filters were used to remove any uncorrelated background noise. The tapped delay line implementation is the most widely used for adaptive filters since it is easy to implement. The least-squares algorithm is used as adaptation algorithm because it converges faster. In order to improve the performance of the adaptive line enhancement filter, a modified adaptive line enhancement filter was proposed in Chapter 3. The modification is based on adding poles into the pass-band region of the ALE filter. This was achieved by introducing a feedback to the filter taps from its output. That is, the output was scaled, delayed, and added to the input. This process generated poles inside the unit circle. The modified ALE filter was tested by applying a sinusoidal signal embedded in white noise. The results showed improvement over the original ALE filter. The better performance of the modified ALE over the original ALE was also confirmed with heart sound signals.

The parameter values of the modified ALE filter, such as the number of filter taps, decorrelation delay, feedback delay parameter, and the number of taps in the feedback loop were chosen to place poles close to the unit circle in the pass-band region. Those poles sharpened the frequency response curve of the filter at the pass-band region while attenuating the sidelobes. Therefore, an improvement in the performance of the filter was achieved.

Heart sounds are quasi-stationary signals. That is, if the physical conditions of the patient do not change during the heart sound recording process, then each heart-beat cycle in the recorded phonocardiogram will have the same temporal and spectral information as the others over a short period of time, i.e. few minutes. This led us to the development of a method to calculate the heart-beat signal from a long data record, we called it the heart-beat separation algorithm and is presented in Chapter 4.

There were two main steps in the heart-beat separation algorithm. First, the electrocardiograms were used to detect the beginning of the cardiac cycles. The PCG beat cycle calculation was only possible since the PCG and ECG signals were

recorded synchronously. The reason for using the ECG signal is that the QRS complex is easily detectable in the ECG signal. A threshold level was set and peaks of the QRS complexes were detected. Using these as pilot points, the peaks of the R-waves were detected and used as the beginning of the cardiac cycles. The ECG signals were passed through noise removing filters before processing them; however, in order to remove the artifact corrupted cycles, the durations of the beat-cycles were checked using the ECG signals.

Then the PCG signals were aligned with the ECG beat cycles and decomposed into heart-beat cycles. The second part of the algorithm consists of identifying and rejecting those heart-beat cycles that contain artifacts. Artifacts are strong disturbances in the heart sounds; therefore, they were not removed by the MALE filter of Chapter 3. The identification and rejection of the artifact corrupted heart-beat cycles was accomplished in the following manner. Detected heart-beat cycles were averaged and a template heart-beat signal was calculated. It was found that in a typical recording there were only few corrupted cycles, and hence constructing the template signal in this way is justifiable. The correlation coefficients of individual heart-beat cycles with the template signal were then calculated and a criterion was set to accept the artifact free heart-beat cycles. The lengths of the heart-beat signals were not equal; therefore, the accepted heart-beat cycles were averaged in the frequency domain to obtain the final heart-beat signal. If there were any small errors in detecting the beginnings of the cardiac cycles, using the frequency domain averaging would avoid errors occurring due to such small temporal differences. The heart-beat separation algorithm was very effective in calculating the heart-beat signals. That is, the final signal was compared to individual heart-beat cycles and was found that it had lower noise and its first, second, and third heart sounds were much easier to identify.

The modulus of the time-frequency matrix of the heart-beat signal consisted of many regions of local of concentration of energy, referred to as local maxima. The time-frequency analysis carried out in this thesis involves identifying the locations of concentrations of energy in the modulus of the time-frequency matrix. Thus, it was necessary to develop an algorithm for this purpose. We called it the local maxima detector and it was presented in Chapter 5. The local maxima detector algorithm is

based on a two step process: detection and removal of local maxima. A Gaussian smoothing filter is applied to suppress the surface fluctuations on the time-frequency plane. Then the peak of the highest energy level in the modulus of the time-frequency matrix is detected and its boundaries are identified and removed from the matrix. Removal of a detected local maximum is necessary because this will cause the next highest energy level to become the new highest energy level in the time-frequency matrix. This process is repeated until all local maxima are detected.

The boundary detection is based on finding the slope of the surface around the local maximum. The outcome of the algorithm is three vectors: a magnitude vector, a temporal vector, and a spectral vector. These vectors contain the energy levels, temporal locations, and spectral locations of the local maxima. A method was proposed to distinguish the local maxima that were due to signal from those that were due to noise. By using a piecewise linear approximation of the magnitude vector, stored in descending order, we were able to identify the local maxima due to noise as the last section, which contains the maxima of lowest energy. The local maxima detection algorithm is applicable to any time-frequency analysis and is not restricted to heart sounds.

Three time-frequency analysis methods were used to analyze the heart-beat signals: the short-time Fourier transform, the Choi-Williams distributions, and the wavelet transform. The analysis procedure was developed in Chapter 6 and was used for all three methods. The spectral vector of local maxima, obtained by the local maxima detector algorithm, was sorted in ascending order of the frequency. The magnitude and the spectral vectors were rearranged accordingly. The products of the three vectors were obtained for local maxima before and after angioplasty and were called *TFM* product vectors. The idea behind generating the *TFM* product vector was the following. If a particular local maximum before angioplasty had similar coordinates (magnitude, time, and frequency) with a local maximum after angioplasty, then their corresponding *TFM* terms would be close. Therefore, if there was any similarity in local maxima before and after angioplasty, then their *TFM* product curves would have similar shapes. This was illustrated through an example in Chapter 6.

In Chapter 6, the short-time Fourier transform was used to analyze heart-beat signals. The time-frequency resolution of STFT is restricted by the uncertainty principle. STFT is attractive to use because it is easy to implement. The type of window used in the STFT applications is very important. A trade-off must be made between the side-lobe attenuations and bandwidth of the window function, depending on the application. We found the Chebychev window to be the most suitable for heart sound analysis because it has better side lobe attenuation. The *TFM* product graphs of before and after angioplasty were compared. It was found that there are subgroups of local maxima that showed similar time-frequency arrangements before and after angioplasty; they were called time-frequency patterns.

The time-frequency patterns were one of the interesting results obtained from the STFT analysis. Patterns were extracted by comparing the TFM curves of before and after angioplasty. The average frequencies of the patterns were calculated in first, second, and third heart sound regions. It was found that the average pattern frequencies after angioplasty were higher than before angioplasty. The frequency shift was found to increase with frequency.

In order to overcome the resolution restrictions of the STFT, the Choi-Williams distribution was used for heart sound analysis in Chapter 7. The CWD is used to suppress the cross-terms, but it is computationally intensive. We used windowed CWD to decrease the computational load. The computational load was further reduced by using a minimum multiplication method, proposed in Appendix B. In this method, multiplications were performed in a triangular fashion; that is, only non-zero elements were multiplied. The symmetry conditions allowed the multiplications to be performed in a quarter of the final matrix. The CWD analysis of the heart-beat signals yielded the same general results obtained from the STFT analysis; that is, there was a frequency shift in average pattern frequencies after angioplasty. However, the number of detected time-frequency patterns was larger than the number of patterns detected using the STFT analysis. This suggests that the better resolution of CWD resulted in detection of peaks due to close frequencies, on the other hand cross-terms may have generated some of the local maxima; cross-terms do not totally disappear by using CWD. Comparison of the *TFM* product graphs of before and after angioplasty revealed that some

of the time-frequency patterns were not quite similar, which indicates that cross-terms may have been generated.

In Chapter 8, the continuous wavelet transform was used to calculate the time-scale matrix of the heart-beat signals. The continuous wavelet transform is best suited for signal analysis. The Morlet wavelet was used as the basis function because its relative bandwidth can easily be set. The frequency domain multiplication was used for calculations of the CWT matrices on the octave base, which was faster than the time domain convolution. In order to be able to compare the results with the results of previous methods, we calculated a frequency matrix whose elements were centre frequencies of the filters associated with each voice scale. It was used to map the time-scale matrix to the time-frequency matrix. The CWT analysis confirmed the frequency shift in the average frequencies of the patterns after angioplasty, and the shift increases with frequency.

A comparison of figures 6.19, 7.10, and 8.12, reveals the following:

1) At low frequencies, below 50 Hz, the results obtained from the STFT analysis show low concentration of local maxima, while the results obtained from the CWD analysis show high concentration of local maxima, and the results from CWT analysis show a medium concentration of local maxima. This indicates that the resolution restriction of STFT limited its applicability, while the cross-terms in the CWD resulted in an increase in the number of local maxima. Thus, the CWT seems to be a more suitable analysis method in this frequency range. Figures 6.19, 7.10, and 8.12 revealed that the above argument is true for first, second, and third heart sound regions.

2) Closer examination of figures 6.19, 7.10, and 8.12 at frequencies below 50 Hz range, reveals that there is a small shift in the average frequencies in both STFT and CWT analysis, while shift in average pattern frequencies in CWD analysis is not consistent. That is, for some patterns there is not any shift at all, while some patterns show frequency shift. This could be another reason for presence of some local maxima that are generated due to cross-terms.

3) In the mid-frequency range, 50-150 Hz, the number of patterns increases in STFT, decreases for CWD, and stays about the same for CWT, compared to the low frequency band.

4) In the high frequency range, 150-250 Hz, all three analysis techniques show almost similar results.

Considering the above points, we may say that wavelet transforms are more suitable for this analysis. The pattern frequency shift was found to be as follows. In the low frequency range, below 50 *Hz*, the average frequency shift was less than 10 *Hz*. For the medium frequency range, between 50 *Hz* and 100 *Hz*, the average frequency shift was below 25 *Hz* and in the high frequency range, above 100 *Hz* and up to 150 *Hz*, the average frequency shift was above 10 *Hz* and below 30 *Hz*, and in the frequency range above 150 *Hz* and below 250 *Hz* the average frequency shift was above 30 *Hz*.

9.2 Future Research

There are a couple of possibilities in pursuing this research. The ultimate goal can be the design of an automatic coronary artery disease recognition system based on heart sound signals and other parameters, such as age, blood pressure and smoking habit of a person. A neural network or a statistical pattern recognition system can be trained for detection of coronary artery disease based on heart sounds. This would require a time-frequency analysis to be conducted for normal heart sounds as well. Time-frequency patterns can be used as features for a pattern recognition system. Based on the findings of our study, it is recommended to use wavelet transform techniques to extract the time-frequency features.

Another pathway for this research to follow is to identify the locations of stenoses in the coronary arteries. That is, to find which branch of the coronary artery is blocked. This would require simultaneous multi-positional recordings of phonocardiograms.

Appendix A

Derivations of
adaptive algorithm for
modified adaptive line enhancement filter

A.1 Introduction

In this appendix we will derive the expression needed to update the *h*-parameters of the modified ALE filter shown in Figure 3.9.

A.2 Calculation of the impulse parameters of the filter

We wish to determine a weight vector **h**(k) that minimizes the performance index

$$E(k) = \sum_{i=1}^{k} e^2(i) = \sum_{i=1}^{k} [x(i) - y(i)]^2 \qquad \text{Eq A.4}$$

The output of the modified ALE filter shown in Figure 3.9 is expressed as

$$y(k) = \sum_{i=0}^{N-1} x(k-\Delta-i)h_i + \beta \sum_{i=0}^{N-1} y(k-D-i)\alpha^i h_i \qquad \text{EQ 3.1}$$

The filter parameters are defined as

$$\mathbf{h}(k) = \left[h_0 \; h_1 \; \ldots \; h_{N-1}\right]^T \qquad \text{Eq A.4}$$

At time k, the filter tap values applied from the input signal are defined as

$$\mathbf{s}(k) = \left[x(k-\Delta) \; x(k-\Delta-1) \; \ldots \; x(k-\Delta-N+1)\right]^T \qquad \text{Eq A.5}$$

The filter tap values applied from the output through the feedback loop are defined as

$$\mathbf{y}(k) = \beta\left[\alpha^0 y(k-D) \; \alpha^1 y(k-D-1) \; \ldots \; \alpha^{N-1} y(k-D-N+1)\right]^T \qquad \text{Eq A.6}$$

Therefore, E(k) can be written as

$$E(k) = \sum_{i=1}^{k} \left[x(i) - \mathbf{h}^T(k)\mathbf{s}(i) - \mathbf{h}^T(k)\mathbf{y}(i)\right]^2 \qquad \text{Eq A.7}$$

$$= \sum_{i=1}^{k} x^2(i) + \mathbf{h}^T(k)\left[\sum_{i=1}^{k} \mathbf{s}^2(i)\right]\mathbf{h}(k) + \mathbf{h}^T(k)\left[\sum_{i=1}^{k} \mathbf{y}^2(i)\right]\mathbf{h}(k)$$

$$- 2\mathbf{h}^T(k)\sum_{i=1}^{k} x(i)\mathbf{s}(i) - 2\mathbf{h}^T(k)\sum_{i=1}^{k} x(i)\mathbf{y}(i) + 2\mathbf{h}^T(k)\left[\sum_{i=1}^{k} \mathbf{s}(i)\mathbf{y}(i)\right]\mathbf{h}(k)$$

We define

$$\mathbf{R}_s(k) = \sum_{i=1}^{k} \mathbf{s}^T(k)\mathbf{s}(k) \qquad \text{Eq A.8}$$

$$\mathbf{R}_y(k) = \sum_{i=1}^{k} \mathbf{y}^T(k)\mathbf{y}(k) \qquad \text{Eq A.9}$$

$$\mathbf{R}_{sy}(k) = \sum_{i=1}^{k} \mathbf{s}^T(k)\mathbf{y}(k) \qquad \text{Eq A.10}$$

$$\mathbf{R}_{ys}(k) = \sum_{i=1}^{k} \mathbf{s}(k)\mathbf{y}^T(k) \qquad \text{Eq A.11}$$

where $\mathbf{R}_s(k)$ and $\mathbf{R}_y(k)$ are the auto-correlation matrices of $\mathbf{s}(k)$ and $\mathbf{y}(k)$; respectively, and $\mathbf{R}_{sy}(k)$ is the cross-correlation matrix of $\mathbf{s}(k)$ and $\mathbf{y}(k)$. We also define

$$\mathbf{P}_{sx}(k) = \sum_{i=1}^{k} x(i)\mathbf{s}(i) \quad \text{Eq A.12}$$

$$\mathbf{P}_{sy}(k) = \sum_{i=1}^{k} x(i)\mathbf{y}(k) \quad \text{Eq A.13}$$

where $\mathbf{P}_{sx}(k)$ and $\mathbf{P}_{sy}(k)$ are the cross-correlation vectors for x(k) with $\mathbf{s}$(k) and $\mathbf{y}$(k), respectively. Using Eq A.8 to Eq A.13 in Eq A.7 we get

$$E(k) = \sum_{i=1}^{k} x^2(i) + \mathbf{h}^T(k)\mathbf{R}_s(k)\mathbf{h}(k) + \mathbf{h}^T(k)\mathbf{R}_y(k)\mathbf{h}(k) - \quad \text{Eq A.14}$$

$$2\mathbf{h}^T(k)\mathbf{P}_{sx}(k) - 2\mathbf{h}^T(k)\mathbf{P}_{sy}(k) + \mathbf{h}^T(k)\mathbf{R}_{sy}(k)\mathbf{h}(k) +$$

$$\mathbf{h}^T(k)\mathbf{R}_{ys}(k)\mathbf{h}(k)$$

Eq A.14 can be minimized by setting its gradient equal to zero. This will result in the least-squares normal equations which can be solved for the vector $\mathbf{h}$(k).

$$\nabla E(k) = \frac{\partial}{\partial \mathbf{h}(k)} E(k) = 0 \quad \text{Eq A.15}$$

Taking the derivative of Eq A.14 we get

$$[\mathbf{R}_s(k) + \mathbf{R}_y(k) + \mathbf{R}_{sy}(k) + \mathbf{R}_{ys}(k)]\mathbf{h}(k) = \mathbf{P}_{sx}(k) + \mathbf{P}_{sy}(k) \quad \text{Eq A.16}$$

This equation is in the form of

$$\mathbf{R}(k)\mathbf{h}(k) = \mathbf{P}(k) \quad \text{Eq A.17}$$

where

$$\mathbf{R}(k) = \mathbf{R}_s(k) + \mathbf{R}_y(k) + \mathbf{R}_{sy}(k) \quad \text{Eq A.18}$$

$$\mathbf{P}(k) = \mathbf{P}_{sx}(k) + \mathbf{P}_{sy}(k) \quad \text{Eq A.19}$$

Therefore, $\mathbf{h}(k)$ can be calculated as

$$\mathbf{h}(k) = \mathbf{R}^{-1}(k)\mathbf{P}(k) \quad \text{Eq A.20}$$

A.3 Updating process

If a new data record is observed then it is not necessary to calculate the h-parameters of the filter from scratch, instead we can use the previous values of the auto-correla-

tion, and cross-correlation matrices to update the h-parameters of the filter. From Eq A.8 to Eq A.10, it is clear that

$$\mathbf{R}(k) = \mathbf{R}(k-1) + \mathbf{s}(k)\mathbf{s}^T(k) + \mathbf{y}(k)\mathbf{y}^T(k) + \mathbf{s}^T(k)\mathbf{y}(k) + \mathbf{s}(k)\mathbf{y}^T(k) \quad \text{Eq A.21}$$

and likewise from Eq A.12 and Eq A.13 we get

$$\mathbf{P}(k) = \mathbf{P}(k-1) + x(k)\mathbf{s}(k) + x(k)\mathbf{y}(k) \quad \text{Eq A.22}$$

As Eq A.20 shows, calculation of *h*-parameters require manipulation of $\mathbf{R}^{-1}(k)$. Inversion of $\mathbf{R}(k)$ is a time consuming process which we want to avoid. In order to do so, it is necessary to express $\mathbf{R}^{-1}(k)$ in term of $\mathbf{R}^{-1}(k\text{-}1)$ using the matrix inversion lemma.

A.3.1 Matrix inversion lemma

Let $\mathbf{X}$ be a matrix expressed as

$$\mathbf{X} = \mathbf{B} + \mathbf{C}\mathbf{C}^T + \mathbf{D}\mathbf{D}^T + \mathbf{C}\mathbf{D}^T + \mathbf{D}\mathbf{C}^T \quad \text{Eq A.23}$$

then the inverse of $\mathbf{X}$ is calculated as

$$\mathbf{X}^{-1} = \mathbf{B}^{-1} - [\mathbf{B}^{-1}\mathbf{C}\mathbf{C}^T\mathbf{B}^{-1} + \mathbf{B}^{-1}\mathbf{C}\mathbf{D}^T\mathbf{B}^{-1} + \mathbf{B}^{-1}\mathbf{D}\mathbf{D}^T\mathbf{B}^{-1} + \mathbf{B}^{-1}\mathbf{D}\mathbf{C}^T\mathbf{B}^{-1}]/[1 + g_1 + g_2 + g_3 + g_4] \quad \text{Eq A.24}$$

where g's are scalar quantities and are defined as

$$g_1 = \mathbf{C}^T\mathbf{B}^{-1}\mathbf{C},\ g_2 = \mathbf{C}^T\mathbf{B}^{-1}\mathbf{D},\ g_3 = \mathbf{D}^T\mathbf{B}^{-1}\mathbf{C}, \text{ and } g_4 = \mathbf{D}^T\mathbf{B}^{-1}\mathbf{D}. \quad \text{Eq A.25}$$

A.3.2 Updating $\mathbf{R}^{-1}$

Eq A.24 enables us to update the inverse of $\mathbf{R}$(k) from its previous value without having to do the inverse calculation. Comparison of Eq A.21 and Eq A.23 reveals that if $\mathbf{C} \equiv \mathbf{s}(k)$ and $\mathbf{D} \equiv \mathbf{y}(k)$, then Eq A.24 can be used to update the $\mathbf{R}^{-1}$(k) as

$$\mathbf{R}^{-1}(k) = \mathbf{R}^{-1}(k-1) - [\mathbf{R}^{-1}(k-1)\mathbf{s}(k)\mathbf{s}^T(k)\mathbf{R}^{-1}(k-1) + \mathbf{R}^{-1}(k-1)\mathbf{s}(k)\mathbf{y}^T(k)\mathbf{R}^{-1}(k-1) + \mathbf{R}^{-1}(k-1)\mathbf{y}(k)\mathbf{y}^T(k)\mathbf{R}^{-1}(k-1) + \mathbf{R}^{-1}(k-1)\mathbf{y}(k)\mathbf{s}^T(k)\mathbf{R}^{-1}(k-1)]/[1 + g_1 + g_2 + g_3 + g_4] \quad \text{Eq A.26}$$

Note that g's are calculated from $\mathbf{s}$(k), $\mathbf{y}$(k), and $\mathbf{R}^{-1}$(k-1) by using Eq A.25. We rearrange Eq A.26 and express it as

$$\mathbf{R}^{-1}(k) = \mathbf{R}^{-1}(k-1) - \mathbf{G}(k)[\mathbf{s}^T(k)\mathbf{R}^{-1}(k-1) + \mathbf{y}^T(k)\mathbf{R}^{-1}(k-1)] \quad \text{Eq A.27}$$

where $\mathbf{G}$(k) is a gain vector defined as

$$\mathbf{G}(k) = \frac{[\mathbf{R}^{-1}(k-1)\mathbf{s}(k) + \mathbf{R}^{-1}(k-1)\mathbf{y}(k)]}{1 + g_1 + g_2 + g_3 + g_4} \quad \text{Eq A.28}$$

A.3.3 Updating h

Once $\mathbf{R}^{-1}$(k) is updated we can derive an expression for updating the h-parameters of the filter. Using Eq A.27 and Eq A.22 in Eq A.20 we get

$$\begin{aligned}\mathbf{h}(k) &= \mathbf{R}^{-1}(k-1)\mathbf{P}(k-1) + \mathbf{R}^{-1}(k-1)x(k)\mathbf{s}(k) \quad \text{Eq A.29}\\ &+ \mathbf{R}^{-1}(k-1)x(k)\mathbf{y}(k) - \mathbf{G}(k)\mathbf{s}^T(k)\mathbf{R}^{-1}(k-1)\mathbf{P}(k-1)\\ &- \mathbf{G}(k)\mathbf{s}^T(k)\mathbf{R}^{-1}(k-1)x(k)\mathbf{s}(k) - \mathbf{G}(k)\mathbf{s}^T(k)\mathbf{R}^{-1}(k-1)x(k)\mathbf{y}(k)\\ &- \mathbf{G}(k)\mathbf{y}^T(k)\mathbf{R}^{-1}(k-1)\mathbf{P}(k-1) - \mathbf{G}(k)\mathbf{y}^T(k)\mathbf{R}^{-1}(k-1)x(k)\mathbf{s}(k)\\ &- \mathbf{G}(k)\mathbf{y}^T(k)\mathbf{R}^{-1}(k-1)x(k)\mathbf{y}(k)\end{aligned}$$

In order to simplify Eq A.29, it is necessary to rearrange Eq A.28 and then replace it in Eq A.29. From Eq A.28 we get

$$\mathbf{R}^{-1}(k-1)\mathbf{s}(k) + \mathbf{R}^{-1}(k-1)\mathbf{y}(k) = g\mathbf{G}(k) \quad \text{Eq A.30}$$

where

$$g = 1 + g_1 + g_2 + g_3 + g_4 \quad \text{Eq A.31}$$

From Eq A.20

$$\mathbf{h}(k-1) = \mathbf{R}^{-1}(k-1)\mathbf{P}(k-1) \quad \text{Eq A.32}$$

Now we replace Eq A.30, Eq A.31, and Eq A.32 into Eq A.29, giving

$$\mathbf{h}(k) = \mathbf{h}(k-1) - \mathbf{G}(k)[\mathbf{y}^T(k) + \mathbf{s}^T(k)]\mathbf{h}(k-1) + g\mathbf{G}(k)x(k) \quad \text{Eq A.33}$$

Eq A.33 is used for updating the h-parameters of the modified ALE filter.

Appendix B

Calculation of Choi-Williams Distribution with minimum possible multiplications

B.1 Introduction

The Choi-Williams distribution is computationally intensive when the length of the signal is large. It is necessary to use window functions to reduce the number of multiplications needed. Utilising a rectangular window of length M, we can write Eq 7.36 as

$$CW(n,\theta) = 2\sum_{\tau=-\infty}^{\infty} e^{-j2\theta\tau} \sum_{\mu=-\infty}^{\infty} W_M(\mu)\Phi(\mu,\tau)S(n,\mu,\tau) \qquad \text{Eq B.1}$$

where

$$\Phi(\mu,\tau) = \frac{1}{\sqrt{4\pi\tau^2/\sigma}}\exp\left(-\frac{\mu^2}{4\tau^2/\sigma}\right) \qquad \text{Eq B.2}$$

and

$$S(n, \mu, \tau) = s(n + \mu + \tau)s^*(n + \mu - \tau) \qquad \text{Eq B.3}$$

where $\Phi(\mu, \tau)$ is a Gausian kernel function. For fixed values of σ, the values of the kernel function, Eq B.2, can be calculated and stored in a matrix.

B.2 Calculations

Let $s(n)$ be a finite complex sequence of N samples, then the range of τ is

$$-\frac{(N-1)}{2} \leq \tau \leq \frac{(N-1)}{2} \qquad \text{Eq B.4}$$

Note that outside this range all the values of $S(n, \mu, \tau)$ are equal to zero. Therefore, the limits of the first summation in Eq B.1 are defined by Eq B.3. The calculation of $\Phi(\mu, \tau)$ for the whole range of τ with a rectangular window $W_M(\mu)$ results in a matrix of order $N + 1 \times M + 1$, with rows corresponding to values of τ and columns corresponding to values of μ. For example, for $N = 9$ and $M = 4$, the windowed $\Phi(\mu, \tau)$ matrix is shown in Table B.1.

Table B.1 Eq B.2 is calculated for N=9 and M=4.

μ \ τ	
−2	0.066 0.084 0.109 0.103 0 0.103 0.109 0.084 0.066
−1	0.069 0.091 0.132 0.219 0 0.219 0.132 0.091 0.069
0	0.070 0.094 0.141 0.282 1 0.282 0.141 0.094 0.070
1	0.069 0.091 0.132 0.219 0 0.219 0.132 0.091 0.069
2	0.066 0.084 0.109 0.103 0 0.103 0.109 0.084 0.066

The matrix of Table B.1 is symmetric about its centre in both horizontal and vertical axes. Therefore, only a quarter of the matrix is needs to be calculated; rest of it can be

generated from the symmetry property. We take advantage of this property in calculating the kernel function, which is calculated once and then used in a look-up table.

Calculation of Eq B.3 for values of τ and $\mu - n$, results in Table B.2; that is, the entries of Table B.2 are the values of S in Eq B.3 at a particular time instant i ($n = i$). The horizontal axis corresponds to values of τ and the vertical axis corresponds to values of μ calculated at time i, with the origin located in the centre of the table.

The hermetian symmetry exists in Table B.2 about the $\tau = 0$ for all values of $\mu - n$; also there are many elements in the table that are zero. Therefore, we only need to calculate half of the elements and ignore those elements that are zero. This greatly reduces the time needed for multiplication. In writing our program to calculate CWD, we have taken advantage of these properties.

Another point that has been taken into consideration is the way in which this matrix is updated with respect to time. That is, once this matrix is calculated at a time instant i, there is no need to calculate a new matrix from scratch at time instant i+1; only part of this matrix need be computed at each time instant. In other words, for each value of $\mu - n$ with

Table B.2 Eq B.3 calculated for time i

$\mu - n$ \ τ	$-\frac{(N-1)}{2}$	...	-2	-1	0	1	2	...	$\frac{(N-1)}{2}$
$-i$	0	...	0	0	$s_0 s^* 0$	0	0	...	0
$1-i$	0	...	0	$s_0 s^* 2$	$s_1 s^* 1$	$s_2 s^* 0$	0	...	0
$2-i$	0	...	$s_0 s^* 4$	$s_1 s^* 3$	$s_2 s^* 2$	$s_3 s^* 1$	$s_4 s^* 0$	...	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
-2	0	...	$s_{i-4} s^* i$	$s_{i-3} s^* i-1$	$s_{i-2} s^* i-4$	$s_{i-1} s^* i-3$	$s_i s^* i-4$	...	0
-1	0	...	$s_{i-3} s^* i+1$	$s_{i-2} s^* i$	$s_{i-1} s^* i-3$	$s_i s^* i-2$	$s_{i+1} s^* i-3$	...	0
0	$s_0 s^* N$	...	$s_{i-2} s^* i+2$	$s_{i-1} s^* i+1$	$s_i s^* i$	$s_{i+1} s^* i-1$	$s_{i+2} s^* i-2$	...	$s_N s^* 0$
1	0	...	$s_{i-1} s^* i+3$	$s_i s^* i+3$	$s_{i+1} s^* i+1$	$s_{i+3} s^* i$	$s_{i+3} s^* i-1$	...	0

$\mu-n$ \ τ	$\frac{(N-1)}{2}$	...	-2	-1	0	1	2	...	$\frac{(N-1)}{2}$
2	0	...	$s_i s^*_{i+4}$	$s_{i+1} s^*_{i+4}$	$s_{i+2} s^*_{i+2}$	$s_{i+4} s^*_{i+1}$	$s_{i+4} s^*_{i}$	...	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
$i-2$	0	...	$s_N s^*_{N+4}$	$s_{N+1} s^*_{N+3}$	$s_{N+2} s^*_{N+2}$	$s_{N+3} s^*_{N+1}$	$s_{N+4} s^*_{N}$	...	0
$i-1$	0	...	0	$s_N s^*_{N+2}$	$s_{N+1} s^*_{N+1}$	$s_{N+2} s^*_{N}$	0	...	0
i	0	...	0	0	$s_N s^*_{N}$	0	0	...	0

a window size equal to M, we only need to have the rows of the matrix from $\mu - n - M/2$ to $\mu - n + M/2$. Consequently, at each instant only $M + 1$ rows are needed. Thus, in the next instant we only have to remove the first row and add another row to the bottom. This way it is only necessary to perform multiplications for one row of the matrix of Table B.2 at every instant after the first step of the calculations.

In order to calculate the CWD according to Eq B.1, a matrix must be calculated for $\Phi(\mu, \tau)$ with the value of 1 as its centre element. The centre element must be aligned with the centre element of the Table B.2, $s_i s^*_i$, and then multiplied and summed along the columns [12]. This gives the auto-correlation of the signal at time i. Then the FFT of the result must be taken. A program was written which implements the above observations to achieve the least possible calculation time.

Bibliography

[1] R. J. Alfredson, "A comparison of the short-time Fourier transform with a discrete wavelet transform for analysing some heart sounds," *Mech. Eng. Trans.*, vol. ME21, no. 1-2, pp. 75-80, 1996.

[2] C.M Anderson, E. H. Satorius, and J. R. Zeidler, "Enhancement of finite bandwidth signals in white Gaussian noise," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. ASSP-31, pp. 17-27, 1983.

[3] R. J. Alfredson, "Wavelets and heart sounds," *Acoustics Australia*, vol. 25, no. 1, pp. 5-10, 1997

[4] M. Akay, "Automated noninvasive detection of coronary artery disease using wavelet-based neural networks," *Intelligent Eng. Sys. through Artificial Neural Networks*, Vol. 4, pp. 517-522, 1994.

[5] M. Akay, J. L. Semmlow, W. Welkowitz, M. D. Bauer, and J. B. Kostis, "Detection of coronary occlusions using autoregressive modelling of diastolic heart sounds," *IEEE Trans. Biomed. Eng.*, vol. 37, no. 4, pp. 366-370, 1990.

[6] M. Akay, J. L. Semmlow, W. Welkowitz, M. D. Bauer, and J. B. Kostis, "Noninvasive detection of coronary stenosis before and after angioplasty using eigenvector methods," *IEEE Trans. Biomed. Eng.*, vol. 37, no. 11, pp. 1095-1104, 1990.

[7] M. Akay, W. Welkowitz, J. L. Semmlow, and J. B. Kostis, "Application of ARMA method to acoustic detection of coronary artery disease," *Med. & Biol. Eng. & Comput.*, vol. 29, pp. 365-372, 1991.

[8] M. Akay, W. Welkowitz, J. L. Semmlow, Y. M. Akay, and J. B. Kostis, "Noninvasive acoustical detection of coronary artery disease using the line enhancer method," *Med. & Biol. Eng. & Comput.*, vol. 30, pp. 147-154, 1992.

[9] J. B. Allen, "Short-term spectral analysis, synthesis and modification by discrete Fourier transform," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. ASSP-25, pp. 235-238, 1977.

[10] J. B. Allen, and L. R. Rabiner, "A unified approach to short time Fourier transform analysis and synthesis," *Proc. IEEE.*, vol. 65, no. 11, pp. 1558-1564, 1977.

[11] D. Barschdorff, A. Bothe, and U. Rengshausen, "Heart sound analysis using neural network and statistical classifiers: a comparison," *Proc. Comput. Cardio.*, Jerusalam, Israel, pp. 19-22, 1989.

[12] D. T. Barry, "Fast calculation of the Choi-Williams time-frequency distribution," *IEEE Trans. Signal Processing,* vol. 40, pp. 450-455, 1992.

[13] P. M. Bentley, J. T. E. McDonnell, "Analysis of heart sounds using wavelet transforms," *Proc. 16th Intern. Conf. IEEE Eng. Medicine Biology Society,* vol. 2, pp. 1304-1305, 1994.

[14] V. Berzin, "Accuracy of Laplacian edge detectors," *Comput. Vision, Graphics, Image Processing,* vol. 27, no. 2, pp. 195-210, 1984.

[15] R. Beyar, S. Levkovitz, S. Braun, and Y. Palti, "Heart sound processing by average and variance calculation physiologic basic and clinical implications," *IEEE Trans. Biomed. Eng.*, vol. BWM-31, no. 9, pp. 591-596, 1984.

[16] T. O. Binford, "Interfering surfaces from images," *Artificial Intel.*, vol. 17, pp. 205-244, 1981. *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 38, no. 11, pp. 1829-1841, 1990.

[17] H. Blanke, M. Cohen, G. U. Schlueter, K. R. Karsch, and K. P. Pentrop "Electrocardiogarphic and coronary angiographic correlations during acute myocardial infraction," *American J. Cardiology*, vol. 54, pp. 249-255, 1984.

[18] B. Boashash, and P. O'Shea, "A methodology for detection and classification of some underwater acoustic signals using time-frequency analysis techniques," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 38, pp. 1829-1841, 1990.

[19] G. F. Boudreaux-Bartels, "Time frequency signal representations for biomedical signals," in *The biomedical engineering handbook* J. D. Bronzino, Boca Raton, Florida: CRC Press, Inc., pp. 866-885, 1995.

[20] J. M. Brady, "Computational approaches to image understanding," *Computing Surveys*, vol. 14, pp. 3-71, 1982.

[21] E. Braunwald, "*Heart disease, a textbook of cardiovascular medicine*," Philadelphia, PA: W. B. Sanders Company, 1997.

[22] J. R. Bulgrin, B. J. Rubal, T.E. Posch, J.M. Moody, "Comparison of binomial, ZAM and minimum cross-entropy time-frequency distributions of intracardiac heart sounds," *Proc. 28 Asilomar Conf. Signals, Systems Comput.*, vol. 1, pp. 383-387, 1994.

[23] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Recog. Anal. Machine Intell.*, vol. PAMI-8, pp. 679-698, 1986.

[24] H. I. Choi, and W. Willaims, "Improved time-frequency representation of multi-component signals using exponential kernels," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. ASSP-37, pp. 862-871, 1989.

[25] C. Chui, "*An introduction to wavelets,*" San Diego, Ca.: Academic Press, Inc., 1992.

[26] T. A. C. M. Claasen, W. F. G. Mecklebrauker, "Time-frequency signal analysis by means of Wigner distributions," *Proc. IEEE Int. Conf. coust. Speech Signal Processing,* vol. 3, pp. 67-72, 1981.

[27] T. A. C. M. Claasen, W. F. G. Mecklebrauker, "The Wigner distribution - a tool for time frequency analysis; part I: continuous time signals," *philips J. Res.*, vol. 35, pp. 217-250, 1980.

[28] T. A. C. M. Claasen, W. F. G. Mecklebrauker, "The Wigner distribution - a tool for time frequency analysis; part II: discrete time signals," *philips J. Res.,* vol. 35, pp. 276-300, 1980.

[29] . A. C. M. Claasen, W. F. G. Mecklebrauker, "The Wigner distribution - a tool for time frequency analysis; part III: relation with other time frequency transformations," *philips J. Res.*, vol. 35, pp. 372-389, 1980.

[30] L. Cohen., "Generalised phase-space distribution functions," *J. Math. Phys.*, vol. 7, pp. 781-786, 1966.

[31] L. Cohen, "Time-frequency distributions, a review," *Proc. IEEE,* vol. 77, no. 7, pp. 941-981, 1989.

[32] L. Cohen, and C. Lee, "Instantaneous frequency and time frequency distributions," *Proc. IEEE Int. Conf. Circuits and Systems*, vol. 2, pp. 1231-1234, 1989.

[33] P. Cohen, "*Diagnosis and therapy of coronary artery disease*," Boston, Mass.: Little, Brown and Company Inc., 1979.

[34] J. M. Cioffi, and T. Kailath, "Fast, recursive-least-square transversal filters for filtering," *IEEE Trans. Acoust. Speech Signal Processing*, vol. ASSP-32, pp. 304-337, 1984.

[35] T. B. Couniham, M. B. Rappaport, and H. B. Spague, "Physiologic and physical factors that govern the clinical appreciation of cardiac thrills," *Circulation,* vol. 4, pp. 716-720, 1951

[36] C. F. N. Cowan, and P. M., "*Filters,*" Englewood Cliffs, NJ.: Prentice Hall, 1985.

[37] R. E. Crochiere, and L. R. Rabiner, "*Multi-rate Digital Signal Processing,*" Englewood Cliffs, NJ.: Prentice Hall, 1983.

[38] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Comm. on Pure and Appl. Math.*, vol. 4, pp. 909-996, 1988.

[39] I. Daubechies, "*Ten lectures on wavelets*" SIAM, CBMS series, 1992.

[40] W. Dock, and S. Zoneraich, "A diastolic murmur arising in a stenosed coronary artery," *American J. Medicine,* vol. 42, no. 4, pp. 617-619, 1967.

[41] T. R. Downes, W. Dunson, K. Stewart, and A. Lundvisk., "Mechanics of physiologic and pathalogic S3 gallop sounds," *American Society Echocardiology*, vol. 5, pp. 211-218, 1992.

[42] D. M. Drzewiecki, M. J. Wasicko, and J. K. Li, "Diastolic mechanics and the origin of the third heart sound," *Ann. Biomed. Engin.*, vol. 19, pp. 651-667, 1991.

[43] B. El-Asir, L. Khadra, A. H. Al-Abbasi, and M. M. J. Mohammed, "Time-frequency analysis of heart sounds," *IEEE TENCON Digital Signal Processing Applications Proc.*, vol. 2, pp. 553-558, 1996.

[44] P. Flandrin, "Some feature of time frequency representation of multicomponent signals," in *Proc. IEEE 1984 int. Conf. Acoust. Speech, Signal Processing* (San Diego, CA), Mar. 1984, pp.41.B.4.1-4.

[45] G. M. Friensen, T. C. Jannett, M. A. Jadallah, S. L. Yates, S. R. Quint, and H. T. Nagle, "A comparison of the noise sensitivity of nine QRS detection algorithms," *IEEE Trans. Biomed. Eng.*, vol. BME 37, pp. 85-98, 1990.

[46] K. S. Fu, and J. K. Mui, "A survey on image segmentation," *Pattern Recognition*, vol. 13, pp. 3-16, 1981.

[47] D. Gabor, "Theory of communication," *J. IEE. (London)*, vol. 93, pp. 429-457, 1946.

[48] P. Goupillaud, A. Grossman, and J. Morlet, "Cycle-octave and related transforms in seismic signal analysis," *Geoexploration*, vol. 23, pp. 85-102, 1984/1985.

[49] D. W. Griffin, and J. S. Lim, "Signal estimation from modified short-time Fourier transform" *IEEE Trans. Acoust. Speech, Signal Processing*, vol. ASSP-32, no. 2, pp. 236-242, 1984.

[50] A. Grossman, R. Kronland-Martinet, and J. Morlet, "Reading and understanding continuous wavelet transforms," in "*Wavelets, Time-Frequency Methods and Phase Space,*" J. M. Combes, A. Grossman, and Ph. Tchamitchian (eds.), New York, NY: Springer-Verlag, pp. 2-20, 2nd edn., 1989/1990.,

[51] A. Grossman, and J. Morlet, and T. Paul, "Transforms associated to square integrable group representation-I: General results," *J. Math. Phys.*, vol. 26, pp. 2473-2479, 1985.

[52] A. Grossman, and J. Morlet, "Decomposition of Hardely functions into square integrable wavelets of constant shape," *SIAM J. Math. Anal.*, vol. 15, pp. 723-736, 1984.

[53] R. C. Gonzalez, and R. E. Woods, *Digital Image Processing*, Addison Wesley pub. Com.: USA, 1993.

[54] Z. Guo, L. G. Durand, H. C. Lee, "The time-frequency distributions of nonstationary signals based on a bessel kernel," *IEEE Trans. Signal Processing*, vol. 42, pp. 1700-1706, 1994.

[55] A. K. Gupta, "On the suppression of a sinusoidal signal in broadband noise," *IEEE Trans. Acoust. Speech, Signal Processing,* vol. ASSP-33, pp. 1024-1026, 1985.

[56] A. C. Guyton, "*Textbook of medical physiology,*" Sanders, Ph: W. B. Sanders Company, 1987.

[57] D. Hausmann, A.-J. S. Lundvisk, G. J. Friedrich, W. L. Mullen, P. S. Fitzgerald, and R. G. Yock, "Intracoronary ultrasound imaging: Intraobserver and interobsrver variability of morphometric measurements," *American Heart J.,* vol. 128, pp. 674-680, 1994.

[58] F. Hlawatsch, "Interference terms in the Wigner distributions," in *Digital Signal Processing*-84 V. Cappellini and A. G. Constantinides, Eds. New York: North-Holland, pp. 69-72, 1984.

[59] F. Hlawatsch, and G. F. Boudreaux-Bartels, "Linear and quadratic time frequency signal representation," *IEEE Signal Processing Magazine*, pp. 21-67, April 1992.

[60] B. F. Hoffman, and P. F. Cranefield, "The physiological basis of cardiac arrhythmias," *American J. Medicine*, vol. 37, pp. 670-684, 1964.

[61] M. L. Honig, and D. G. Messerschmitt, "*filters: structures, algorithms, and applications,*" Kluwer academic publishers: Boston, Mass., 1984.

[62] A. Iwata, N. Suzumura, and K. Ikegaya, "Pattern classification of phonocardiogram using linear prediction analysis," *Med. Biolog. Eng. Computing,* vol. 15, pp. 407-412, 1977.

[63] V. K. Iyer, P. A. Rammorthy, H. Fan, and Y. Ploysongsang, "Reduction of heart sounds from lung sounds by filtering," *IEEE Trans. Biomed. Eng.*, vol. BME 33, no. 12, pp. 1141-1148, 1986.

[64] J. Jeong, W. J. Williams, "Kernel design for reduced interference distributions," *IEEE Trans. Signal Processing,* vol. 40, pp. 402-412, 1992.

[65] D. L. Jones, R. G. Baraniuk, "Effective approximations of continuous wavelet transforms," *Electronics Letters*, vol. 27, pp. 748-750, 1991.

[66] D. L. Jones, T. M. Parks, "A resolution comparison of several time-frequency representations," *IEEE Trans. Signal Processing,* vol. 40, pp. 413-420, 1992.

[67] A. Kaewlium, and H. Longbotham, "Application of Gabor transform as texture discriminator of masses in digital mammograms," *Biol. Sci. Instrum.* vol. 29, 183-187, 1993.

[68] G. Kaiser, "*A friendly guide to wavelets,*" Boston Mass.: Birkhäuser, 1994

[69] S. Krishnamachari, W. J. Williams, "Kernel design in the generalized marginals domain for time-frequency analysis," *IEEE Int. Conf. Acoust. Speech Signal Processing,* vol. III, pp. 341-4, 1994.

[70] J. S. Lim, *Two dimensional signal and image processing*, Prentice Hall, 1990.

[71] D. Littmann, "An approach to the ideal stethoscope," *J. American Med. Association*, vol. 178, pp. 504-505, 1961.

[72] R. M. Loynes, "On the concept of spectrum for non-stationary processes," *J. R. Stat. Soc.,* ser. B, vol. 30, pp. 1-30, 1968.

[73] A. A. Luisada, "*From auscultation to phonocardiography,*" Saint Louis: The C. V. Mosby Company, 1965.

[74] A. A. Luisada, "Sounds and pulses as aids to cardiac diagnosis," *Med. Clinics North America*, Vol. 64, no. 1, pp. 3-32, 1980.

[75] A. A. Luisada, and F. Portaluppi, "The main heart sounds as vibrations of the cardiohemic system: old controversy and new facts," *American J. Cardiology*, vol. 52, no. 8, pp. 1133-1136, 1983.

[76] O. Macchi, "*Processing, the least mean squares approach with applications in transmission,*" John Wiley & sons: New York, N.Y., 1995.

[77] S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 674-693, 1989a.

[78] S. Mallat, "Multiresolution approximations and wavelet orthonormal bases of $L^2(R)$," *Trans. Amer. Math. Soc.*, vol. 315, pp. 69-87, 1989b.

[79] S. Mallat, and W. L. Hwang, "Singularities detection and processing with wavelets," *IEEE Trans. Info. Theory*, vol. 38, no. 2, pp. 617-643, 1992.

[80] S. Mallat, and S. Zhong, "Characterization of signals from multiscale edges," *IEEE Trans. Pattern Recog. Anal. Machine Intell.*, vol. 14, pp. 710-732, 1992.

[81] H. Margenau, R. N. Hill, "Correlation between measurements in quantum theory," *Prog. Theor. Phys.*, vol. 26, pp. 722-738, 1961.

[82] W. D. Mark, "Spectral analysis of the convolution and filtering of non-stationary statistic processes," *J. Sound Vib.,* vol. 11, pp. 19-63, 1970.

[83] D. C. Marr, E. C. Hildreth, "Theory of edge detection," *Proc. Roy. Soc. London B*, vol. 207, pp. 187-217, 1980.

[84] A. Messiah, "*Quantum mechanics,*" North-Holland, Amesterdam, 1961.

[85] A. S. A. Mohammed, and H. M. Raafat, "Recognition of heart sounds and murmurs for cardiac diagnosis," *9th Int. Conf. Pat. Recogn.*, Rome, Italy, vol. 2, pp. 1009-1011, 1988.

[86] V. S. Nalwa, and T. O. Binford, "On detecting edges," *IEEE Trans. Pattern Recog. Anal. Machine Intell.*, vol. PAMI-8, pp. 699-714, 1986.

[87] D. Nandagopal, "*Phonocardiogram frequency analysis techniques and noninvasive determination of heart valve calcification*," PhD dissertation, University of Adelaide, Adelaide South Australia, 1984.

[88] D. Nandagopal, R. E. Bogner, J. Mazumdar, "Spectral analysis of second heart sound using linear predictive coding," *Proc. 2nd Int. Conf. Mechanics Medicine Biology,* Osaka, Japan, pp. 102-103, 1980.

[89] S. N. Nawab, and T. F. Quatieri, "Short time Fourier transform," in *Advanced topics in signal processing*, J. S. Lim and A. V. Oppenheim, edt.: Prentice Hall. Englewood Cliffs. NJ, 1988.

[90] M. S. Obaidat "Phonocardiogram signal analysis: techniques and performance comparison," *J. Medical Eng. Technology*, vol. 17, no. 6, pp. 221-227, 1993.

[91] A. V. Oppenheim, and R. W. Shafer, "*Discrete time signal processing,*" Englewood Cliffs, NJ: Prentice-Hall, 1989.

[92] Y. Ozawa, D. Smith, and E. Craige, "Origin of the third heart sound, II. Studies in human subjects," *Circulation*, vol. 67, pp. 399-404, 1983.

[93] C. H. Page, "Instantaneous power spectra," *J. Appl. Phys.*, vol. 23, pp. 103-106, 1952.

[94] D. Picard, J. Charara, F. Guidoin, Y. Haggag, D. Doussart, D. Walker, and T. How, "Phonocardiogram spectral analysis simulator of mitral valve prostheses," *J. Med. Eng. Techn.*, vol. 15 pp. 222-231, 1991.

[95] I. Pitas, "*Digital Image Processing Algorithms*," Hertfordshire, UK: Prentice Hall, 1993.

[96] I. Pitas, and A. N. Venetsanopoulos, "*Nonlinear digital filters, principal and applications*," Kluwer Academic, 1990.

[97] M. R. Portnoff, "Time frequency representations of digital signals and systems based on short time Fourier analysis," *IEEE Trans. Acoust. Speech Signal Processing*, vol. 28, pp. 55-69, 1980.

[98] J. G. Proakis, D. G. Manolakis, "*Digital Signal Processing,*" Englewood Cliffs, NJ.: Prentice Hall, 1996.

[99] L. R. Rabiner, and R. W. Schafer, "*Digital Processing of Speech Signals*," Englewood Cliffs, NJ.: Prentice Hall, pp. 250-347, 1978.

[100] D. V. B. Rao, S.-Y. Kung, "Notch filtering for the retrieval of sinusoids in noise," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 32, no. 4, pp. 791-802, 1984.

[101] B. D. Rao, "Properties of FIR-ALE structure," *Digital Signal Processing*, vol. 3, pp. 29-35, 1993.

[102] T. R. Reed, and J. M. Hans Du Buf, "A review of recent texture segmentation and feature extraction techniques," *CVGIP: Image Understanding,* vol. 57, no. 3, pp. 359-372, 1993.

[103] W. Rihaczek, "signal energy distribution in time and frequency," *IEEE Trans. information Theory*, vol. IT-14, pp. 369-374, 1968.

[104] O. Rioul, "Simple regularity criteria for subdivision scheme," *SIAM J. Math. Anal.*, vol. 23, no. 6, pp. 1544-1576, 1992.

[105] O. Rioul, M. Vetterli, "Wavelet and signal processing," *IEEE Signal Processing magazine*, pp. 14-38, Oct. 1991.

[106] J. Ritola, S. Lukkarinen, "Comparison of time-frequency distributions in the heart sounds analysis," *Med. Biol. Eng. Comput.*, vol. 34, suppl. 1, p. 89-90, 1996.

[107] R. F. Rushmer, "*Cardiovascular dynamics,*" Philadelphia, PA: W. B. Sanders Company, 1970

[108] H. N. Sabbah, and P. D. Setain, "Investigation of the theory and mechanism of the origin of the second heart sound," *Circulation Research*, vol. 39, pp. 847-882, 1976.

[109] P. A. Saillant, J. A. Simmons, S. P. Dear, "A computational model of echo processing and acoustic imaging in frequency modulated echo-locating bats: The spectrogram correlation and transformation receiver," *J. ASA.* vol. 94, pp. 2691-2694, 1993.

[110] H. P. Sava, and J. T. E. McDonnell, "Spectral composition of heart sounds before and after mechanical heart valve implantation using a modified forward-backward Prony' method," *IEEE Trans. Biomed. Eng.*, vol. 43, no. 7, pp. 734-742, 1996.

[111] R. C. Schlant, C. G. Blomqvist, R. O. Brandenburg, and M. E. Josephson, "Guidelines for exercise testing: A report of the joint American College of cardiovascular procedures (subcommittee on exercise testing)," *Circulation*, vol 74 (suppl. III) pp. 653a:667a, 1986.

[112] B. L. Segal, W. Likoff, and J. H. Moyer, "*The theory and practice of auscultation,*" Philadelphia: F. A. Davis Co., 1964.

[113] A. Selzer, "*Principles & practice of clinical cardiology,*" Philadelphia, PA: W. B. Sanders Company, 1983,

[114] J. Semmilow, W. Welkowitz, J. Kostis, and J. W. Mackenzie, "Coronary artery disease correlates between diastolic auditory characteristics and coronary artery disease," *IEEE Trans. Biomed Eng.*, vol. BME-30, no. 2, pp. 136-139, 1983.

[115] X. Shilin, H. Min, Z. Diancheng, "The time-frequency representation of PCG," *Proc. IEEE inter. Conf. Industrial Techn.*, pp. 679-681, 1994.

[116] M. Sokolow, M. B. Mcllroy, "*Clinical cardiology*," Los Altos, Ca.: Lange Medical Publication, 1986.

[117] M. Sonka, V. Nlavac, and R. Boyle, "*Image Processing, Analysis and Machine Vision*," Chapman and Hall: London, 1994.

[118] W. Stanford, B. H. Thompson, and R. M. Weiss, "Coronary artery calcification: clinical significance and current methods of detection," *American J. Roentgenology*, vol. 161, pp. 1139-1146, 1993.

[119] L. Stankovic, "Auto-term representation by the reduced interference distributions: a procedure for kernel design," *IEEE Trans. Signal Processing*, vol. 44, pp. 1557-63, 1996.

[120] G. Strange, "Wavelets and dilation equations: a brief introduction," *SIAM review*, vol. 31, pp. 614-627, 1989.

[121] P. Strobach, K. Abraham-Fuchs, and W. Härer, "Event-synchronous cancellation of the heart sounds interference in biomedical signals," *IEEE Trans. Biomed. Eng.*, vol. 41, no. 4, pp. 343-350, 1994.

[122] H. D. Tagare, and R. J. P. deFigueiredo, "On the localization performance measure and optimal edge detection," *IEEE Trans. Pattern Recog. Anal. Machine Intell.*, vol. 12, pp. 1186-1190, 1990.

[123] N. V. Thakor, and J. G. Webster, "Electrode studies for the long term ambulatory ECG," *Med. Biolo. Eng. Comput.*, vol. 23, no. 2, pp. 116-121, 1985.

[124] N. V. Thakor, and Y.-S. Zho, "Applications of filtering to ECG analysis: Noise cancellation and arrhythmia detection," *IEEE. Trans. Biomed. Eng.*, vol. 38, pp. 785-793, 1991.

[125] A. Teolis, "*Computational signal processing with wavelets*," Boston, Massachusettts: Birkhäuser, 1998.

[126] V. Torre, and T. A. Poggio, "On the edge detection," *IEEE Trans. Pattern Recog. Anal. Machine Intell.*, vol. PAMI-8, no. 2, pp. 147-152, 1986.

[127] J. R. Treichler, "Transient and convergence behaviour of the line enhancer," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. ASSP_27, no. 1, pp. 53-62, 1979.

[128] P. L. Tyack, W. J. Williams, and G. Cunningham, "Time-frequency fine structure of dolphin whistles," *Proc. IEEE. Sp. Intl. Symp. T-F and T-S Anal. (Victoria, BC, Canada)*, pp.50-56, 1993.

[129] F. B. Tuteur, "Wavelet transformation in signal detection," in "*Wavelets, Time-Frequency Methods and Phase Space*," J. M. Combes, A. Grossman, and Ph. Tchamitchian (eds.), New York, NY: Springer-Verlag, pp. 132-138, 2nd edn., 1989/1990.,

[130] F. Van de Werf, J. Minten, P. Carmeliet, H. D. Geest, and H. Kesteloot, "Genesis of the third and fourth heart sounds," J. *Clinical Investigation,* vol. 73, pp. 1400-1407, 1984.

[131] M. Vetterli, and J. Kovacevic, "*Wavelets and subband coding,*" Englandwood Cliffs, NJ: Prentice-Hall, 1995.

[132] J. Ville, "Théorie et applications de la notion de signal analytique," *Cables et transmission,* vol. 2A, pp. 61-74, 1948.

[133] F. J. Th. Wackers, "Exercise myocardial perfusion imaging,", *J. Nuclear Medicine*, vol. 35, pp.726-729, 1994.

[134] B. Widrow, J. R. Glover Jr., J. M. McCool, J. Kaunitz, C. S. Williams, R. H. Hearn, J. R. Zeidler, E. Dong Jr., and R. C. Goodlin, "noise cancellation: Principal and applications," *Proc. IEEE. Trans.,* vol. 63, pp. 1692-1716, 1975.

[135] B. Widrow, S. D. Stearns, "*Adaptive signal processing*", Prentice-Hall, Inc.: New Jercy, 1985

[136] E. P. Wigner, "On the quantum correction for thermodynamic equilibrium," *Phys. Rev.*, vol. 40. pp. 749-759, 1932.

[137] J. C. Wood, D. T. Barry, "Time-frequency analysis of the first heart sound," *IEEE Eng. Med. Biol.*, vol. 14, pp. 144-151, 1995.

[138] J. C. Wood, D. T. Barry, "Quantification of first heart sound frequency dynamics across the human chest wall," *Med. Biol. Eng. Comput.*, vol. 32, pp. S71-S78, 1994.

[139] J. C. Wood, A. J. Buda, D. T. Barry, "Time-frequency transforms: A new approach to first heart sound frequency dynamics," *IEEE. Trans. Biomed. Eng.*, vol. 39, pp. 730-740, 1992.

[140] W. Yanjun, X. Jingping, Z. Yan, W. Jing, W. Bo, and C. Jingzhi.,"Time-frequency analysis of the second heart sound signals," *Proc. 17th Int. Conf. Eng. Medicine. Bio. Society.*, vol.1, pp. 131-132, 1997

[141] S. Yao, Z. He, "Analysis of multicomponent chirp signals using frequency shear representation," *J. Circuits Sys. Computers*, vol. 6, pp. 385-401, 1996.

[142] Y. Yoganandam, V. U. Reddy, and T. Kailath, "Performance analysis of the line enhancer for sinusoidal signals in broad-band noise," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 36, pp. 1749-1757, 1988.

[143] J. R. Zeidler, E. H. Satorius, D. M. Chabriery, and H. T. Wexler, "Adaptive enhancement of multiple sinusoidal in uncorrelated noise", *IEEE Trans. Acoust. Speech, Signal Processing*, vol. ASSP-26, pp. 240-254, 1978

[144] G. Zelniker, and F. Taylor, "*Advanced digital signal processing, theory and application,*" Marcel Decker Inc.: New York, 1994.

[145] Y. Zhao, L. E. Alas, R. J. Marks, "The use of cone-shaped kernels for generalized time-frequency representations of nonstationary signals," *IEEE Trans. Acoust. Speech, Signal Processing,* vol.38, pp. 1084-91, 1990.