UNIVERSITY OF SOUTHAMPTON

FACULTY OF ENGINEERING, SCIENCE AND MATHEMATICS

INSTITUTE OF SOUND AND VIBRATION RESEARCH

HIDDEN MARKOVIAN MODELS APPLIED TO THE ANALYSIS OF HEART SOUNDS FOR DIAGNOSTIC PURPOSES

by

Eduardo Romero Vivas

Thesis submitted for the degree of Doctor of Philosophy

August, 2006

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING, SCIENCE AND MATHEMATICS
INSTITUTE OF SOUND AND VIBRATION RESEARCH

Doctor of Philosophy

HIDDEN MARKOVIAN MODELS APPLIED TO THE ANALYSIS OF HEART SOUNDS FOR DIAGNOSTIC PURPOSES

By Eduardo Romero Vivas

Auscultation has long been the primary test for initial assessment of the patient's heart conditions and early detection of anomalies. However, the reliability of this technique on subjective judgement, expertise, and individual's hearing accounts for the inconsistency in the diagnostic among experts. Despite several efforts towards the development of autonomous systems, their limited success suggests the use of new approaches for signal representation and pattern classifiers.

This thesis explores the use of hidden Markovian models (HMM) for characterisation and classification of heart sounds.

Inspired on the cardiohemic theory, Cepstral coefficients are proposed as suitable heart sound representations for analysis and classification. Their performance is evaluated by comparison with sub-band energy representations derived from a standard set of filters for clinical phonocardiography (Maass-Weber filters), and line spectral pairs, proposed as an alternative to linear prediction coefficients.

Using hidden Markov models and cepstral representations, an autonomous algorithm that requires no external signal reference is proposed for PCG signal segmentation.

Using a database of paediatric heart sounds, a screening test is devised to asses the performance of the HMM as signal classifier for systolic murmur identification. Measures of diagnostic accuracy for the algorithm are obtained using different input representations. The best results, obtained using cepstral representations, are compared to those provided by paediatric cardiologist's diagnosis based on X-ray plates and electrocardiograpy. Both results are combined in a sequential tests to assess the use of the HMM classifier as a clinical diagnosis aid.

HIDDEN MARKOVIAN MODELS APPLIED TO THE ANALYSIS OF HEART SOUNDS FOR DIAGNOSTIC PURPOSES

List of Tables and Figures

Acronyms and Abbreviations

A2	Aortic Component of the Second Sound
ANN	Artificial Neural Network
AS	Aortic Stenosis
ASD	Atrial Septal Defect
A-V	Atrio-Ventricular
CAD	Coronary Artery Disease
cm	Centimetre
dB	Decibel
DNA	Deoxyribonucleic Acid
ECG	Electrocardiogram
FFT	Fast Fourier Transform
g	Gram
GMHMM	Gaussian Mixture Hidden Markovian Model
HMM	Hidden Markovian Models
Hz	Hertz
ISVR	Institute of Sound and Vibration Research
MLP	Multi-layer Perceptron
ms	Millisecond
P2	Pulmonary Component of the Second Sound
PSP	Paralytic Shellfish Poisoning
PCG	Phonocardiogram
S1	First Heart Sound
S2	Second Heart Sound
SNR	Signal to Noise Ratio
μg	Microgram
VSD	Ventricular Septal Defect

Acknowledgements

First of all, I would like to express my gratitude to the National Council of Science and Technology (CONACYT México) for their support with special thanks to its personnel.

Thanks to Dr. Mario Martínez García from the Northwest Biological Research Centre (CIBNOR México) for his encouragement to achieve a further level in my education.

I am sincerely indebted to Professor Paul White for his kind advice and particularly grateful for his friendship and his willingness to help me during this project. Special thanks to Dr. Simpson for his help during the final stages.

I would like to thank all my professors at the ISVR and at the IPN, to all my friends in Mexico and to all the friends I have had the opportunity to meet in the United Kingdom.

Finally, I would like to express my sincere gratitude to my parents and to my brothers. This dissertation is specially dedicated to my new family: my beloved Aleyda and Itzcoatl.

Eduardo Romero Vivas

Southampton, August 2006.

1 Introduction

'I have been able to hear very plainly the beating of a Man's heart ...Who knows, I say, but it may be possible to discover the Motions of the Internal Parts of Bodies... by the sound they make'

Robert Hooke (1637-1703)

1.1 Background

Auscultation is a non-invasive technique in which a stethoscope is used to listen to the sounds in the body. The vibromyogram; a manifestation of contraction of a skeletal muscle, and the vibroarthrogram; recorded from an articulation in movement, are two good examples of the various applications of this technique to gain information about the physiology of the body. When this technique is used to listen to the heart, the sounds received provide valuable information concerning the integrity and function of the heart valves and on the hemodynamics of the system. Auscultation has a high potential for detecting various heart diseases. It has long been the primary test for initial assessment of the patient's heart conditions and early detection of anomalies.

The usefulness of auscultation for diagnosis has been recognised for a long time. During the Hippocratic period (460 to 370 BC) for example, auscultation was practised by the direct application of the ear to the patient's chest and abdomen, a process known as immediate auscultation. Hippocrates described a 'succussion splash' a noise heard when a body cavity containing air and water is shaken briskly. He provided simple descriptions of the sounds that would let one know by means of direct auscultation when the chest contained water but no pus. (Mckusick, 1958).

This form of immediate auscultation remained unchanged until Rene Laennec invented the stethoscope in 1816 (the word stethoscope derives from the Greek words *stethos* meaning chest and *skopein* that means to view or to see). Rolling a

piece of paper into a sort of cylinder and applying one end of it to the region of the heart and the other to his ear Laennec invented the first stethoscope and the use of mediate auscultation.

The invention of the stethoscope, although driven by convenience and propriety rather than for a technical need, focused the attention of the scientific community on a scientific approach to physical examination. Laennec himself carried on his research on diseases of the chest, carefully auscultating close to 3000 patients between May 1817 and May 1819, while correlating his antemortem observations with necropsy findings. His ability to play a variety of musical instruments enhanced his appreciation of the acoustic discoveries. His work '*Traité de L'auscultation Médiate*' published in 1819 provides many of the related terminology still used like the terms crepitations, bruit, rales, egophony, and rhonchi (McKusick, 1958).

Since then, many illustrious physicians have contributed to the understanding of cardiac auscultation by providing an explanation for the sounds and noises that are heard in the normal and diseased heart. Such has been the impact of this technique in medicine that nowadays the most renowned medical instrument is the stethoscope.

The information obtained by the clinician from the sounds produced by the heart is used to recognise different heart diseases; however, the use of the stethoscope is limited by several factors. The expertise of the practitioner, subjective judgement, the limitations of the terminology used, hearing constraints, and constraints of the instrument itself, can account for the inconsistency among experts (Rushmer, 1968). Referring to the instrument for example, from a technical point of view, sounds are damped and distorted by stethoscopes. Various types of stethoscopes exhibit marked differences in their efficiency at transmission of cardiovascular sounds as determined by length of transmission pathway, diameter, and stiffness of tubing, and especially, the sealing at the chest wall and in the ear canal (Groom, 1964).

Despite its limitations, cardiac auscultation remains an important tool for initial diagnosis. Many pathological conditions of the cardiovascular system cause murmurs and aberrations in heart sounds much before they are reflected as other symptoms, such as changes in the Electrocardiogram signal (Rushmer, 1970). The cost, ease of operation, and non-invasive nature of cardiac auscultation are irreplaceable by other techniques.

Nevertheless, with the advent of new technologies particularly the use of m-mode, bi-dimensional, and Doppler echocardiography, the use of auscultation as a diagnostic tool has been relegated. The complexity of this technique, the incomplete knowledge about the mechanisms generating the heart sounds, and its implicit limitations have also contributed to changing it from a diagnostic procedure to a pre-diagnostic assessment, which in the vast majority of cases needs to be supported by more complex techniques.

Nevertheless, the evolution of technology also provides the means to overcome many of the limitations associated with traditional auscultation technique:

- The use of means to record heart sounds provides a basis for comparison among experts. Phonocardiography is useful to obtain specific measurements on the timing of heart sounds and murmurs, and provides a means to archiving auscultated sounds for a better follow-up of patients. Simultaneous recording of phonocardiograms, electrocardiograms, and carotid pulse provide a better tool for diagnosis. Those recordings are also extremely useful for the training of future cardiologists.

- The use of electronic stethoscopes expands the range of frequencies that can be used in the analysis of heart sounds beyond the frequency range that can be perceived by humans. Electronic stethoscopes can amplify low-level sounds and help to minimise the sources of error due to the variability of auditory capabilities associated with ageing and differences among individuals. An electronic stethoscope can be used as a stand-alone instrument or as input stage for an electronic system for processing and displaying.

- The use of digital signal processing techniques provides different representations suitable for objective analysis, comparison, and classification. Signal processing techniques can be applied to the heart sound signals providing visual representations that could ease the interpretation of the physical phenomena involved. The use of time frequency representations helps in the detection of murmurs according to their frequency content. Changes in the spectrum of the first heart sound component are consistent with reduced ventricular elasticity associated with myocardial infarct (Adolph, 1970). Several efforts have been focused towards the development of an autonomous system in order to provide heart sound classification aiming to aid diagnosis (Iwata, 1980; Baranek, 1989; Barschdorff, 1995; Leung, 2000b; Reed, 2001; Olmez, 2003).

- Specialised interactive training software offers the opportunity to capture and share the expertise required for an accurate diagnosis. Software developed at the Institute of Sound and Vibration Research (Brown, 2002) aimed at future cardiologists allows the user to listen to a recording and either eliminate or enhance the different components until they are confident that the sounds are correctly identified. This software is used for teaching auscultation to students and junior doctors on the Wessex Cardiothoracic Unit in Southampton. The use of this training aid also helps to familiarise the user with the modified spectral composition of the heart sounds and murmurs heard with electronic stethoscopes. The difference between the sounds obtained with an electronic stethoscope and the familiar sound of the bell and diaphragm acoustic stethoscope has been stressed as one of the main difficulties in introducing the electronic stethoscope into clinical practice (Durand, 1995).

Relatively recent technical developments, such as the use of electronic stethoscopes and computers have largely contributed towards the renaissance of this ancient valuable technique leading to what is now called 'digital phonocardiology'.

1.2 Digital Phonocardiography

The use of electronic stethoscopes, the increasing processing power of digital computers and the development of signal processing techniques have renewed the interest of the scientific community in the improvement and diversification of the phonocardiogram analysis techniques. Research in this area can be classified into the following main categories: (1) Studies on the genesis, transmission and propagation of heart sounds (Fergulio, 1963; Zalter, 1963; Agress, 1964; Kingsley, 1974; Van Vollenhoven, 1971; Yoganathan, 1976; Kozmann, 1977; Stein, 1980; Durand, 1982, 1985, 1986; Baracca, 1990; Donnerstein, 1994; Wood, 1995; (2) Detection of cardiomyopathies, ventricular dysfunction and pulmonary hypertension by analysis of heart sounds (Rangayyan, 1979; Iwata, 1980; Baranek, 1989; Obaidat, 1992; Barschdorff, 1995; White, 1994; Leung, 1997, 1998a, 1998b, 2000a, 200b; Reed, 2001; Olmez, 2003); (3) Condition monitoring of prostethic heart valves (Stein, 1980; Durand, 1986, 1990); (4) Detection of coronary artery disease (Akay, 1990a, 1990b, 1991, 1993, 1994).

1.2.1 Study on the genesis transmission and propagation of heart sounds

The heart sounds are a consequence of the dynamic events associated with the contraction and relaxation of the atria and the ventricles, the valve movements, and blood flow; however, the exact mechanisms producing such sounds are still not fully understood.

There are many theories regarding the causes of heart sounds among which two main explanations are predominant. The valvular theory (McKusky, 1958) hypothesises that heart sounds are transient vibrations resulting from the abrupt tension of the valve leaflets at the end of opening and closure of the four heart valves, although opening and closure are considered to be silent. The Cardiohemic theory introduced by Rushmer (1970) attributes the heart sounds to vibrations of the whole heart structure caused by acceleration and deceleration of intracardiac

blood mass following the opening and closure of the heart valves. Although the latter is the most accepted theory there is still some controversy.

The development of intracardiac phonocardiography, signal processing, and echocardiography, has provided new approaches to investigate the basic mechanisms involved in the genesis of heart sounds and murmurs. Some of the basic controversies between theories were solved, but some others, especially those associated with the genesis of the first and third heart sounds and some functional murmurs, like the Still's murmur, are still debated (Miao, 1987).

The application of the fast Fourier transform (FFT) to the spectral analysis of heart sounds has improved the basic understanding of the phonocardiogram (PCG) and its relationship to the cardiovascular events.

Several researchers have used FFT analysis of PCG signals. For instance, Yoganathan (1976) used FFT to analyse the first (S1) and second (S2) heart sounds. They found significant peaks in the 80-400 Hz range in the aortic and pulmonary areas and concluded that those peaks are related to the elastic properties of the heart muscles and the dynamic events causing S1 and S2.

Stein (1980) investigated the frequency spectrum of S2 and related the frequency with the highest amplitude with the stiffness of the aortic valve. This work initiated studies on condition monitoring of prosthetic valves.

Application of the spectrogram in normal children showed a decrease of the frequency content of Still's innocent murmur with an increase in age and heart dimensions (Donnerstein, 1994). This study also suggested that the difficulty of detecting this murmur in older children is probably due to its decrease in frequency content, since the ear is less sensitive to lower frequencies, making this murmur difficult to detect by auscultation but not by phonocardiography.

Baracca (1990) used the spectrogram to evaluate the contributions of various heart structures to the time frequency distribution of S1 by correlating its energy in

various frequency bands with geometrical parameters of the heart structures obtained by echocardiography. The best correlation was obtained with the diameter of the left atrium, the volume of the left ventricle, the mitral valve area, and the aortic diameter.

Wood (1994, 1995) studied the local time-frequency response of S1 recorded on the epicardium of dogs during 2 hours of coronary ischemia to determine if changes in myocardial properties would affect the time frequency distribution of S1 due to a local change in the resonant properties (stifness) of the myocardium. Their results showed inconsistency with the theory supporting a resonant origin of the first heart sound. It was thus suggested that S1 appears to be a superposition of low-frequency vibrations arising from myocardial contraction and high frequency transients initiated by valvular activity.

The exact mechanisms of transmission of the heart sounds and murmurs within the heart-thorax structures and on the surface of the thorax are still unknown. The transmission characteristic of the thorax, and the various thoracic sites used to record the heart sounds and murmurs often generate major difficulties in the comparison of the results obtained from different studies.

The physical, mechanical and vibrational aspects of heart sounds and sound transmission through human tissues have been studied in detail by many researches using different methodologies (Agress, 1964; Van Vollenhoven, 1971; Kingsley, 1974; Kozmann, 1977; Durand, 1985).

Intracardiac sound generators and transducers were used to evaluate the transmission characteristics of the thorax by inside to outside transmission and vice-versa (Fergulio, 1963; Zalter, 1963). The studies indicate that the nature of the recorded external sounds depends upon cardiac action characteristics, site of sound production, and properties of the conducting tissues.

Using a multi-channel acquisition system and an array of 25 microphones Rangayyan (1987) studied the distribution of the PCG on the surface of the

thorax. This study has been used to identify optimal sites of radiation of stenotic and regurgitan murmurs by generating reference isocontour maps of the heart and murmurs.

Computer models have been developed in order to characterise the heart thorax acoustic system from the point of view of heart sound transmission (Durand, 1982, 1985, 1986).

1.2.2 Detection of cardiomyopathies, ventricular dysfunction, valvular heart disease, and pulmonary hypertension by analysis of hearts sounds

With the aim of overcoming the limitations imposed by traditional auscultation techniques, several efforts have been made to use electronic stethoscopes, computers, and advanced digital signal processing techniques to analyse heart sounds and to relate qualitatively PCG signal features with the corresponding physiological and pathological conditions.

Most of the research is focused towards the identification of murmurs, i.e. abnormal heart sounds caused by turbulent blood-flow in the heart. From the localisation of the murmurs, duration, loudness, timing, pitch, quality and shape, it is possible to identify anomalies in heart valves, ventricular dysfunction, heart defects, and pulmonary hipertension, among other pathological conditions (Guadalajara, 1998).

Initial works were orientated into finding representations that overcome human hearing constrains and could provide more information related to the physical phenomena. The search of the best feature representation was mainly based either on frequency or on time domain analysis (Rangayan, 1979; Obaidat, 1992; Leung, 1997).

Subsequent efforts have focused towards the development of autonomous systems for heart sound diagnosis in order to avoid the dependence on the skill and expertise of the listener and to extend the human hearing system capabilities

(Iwata, 1980; Baranek, 1989; Barschdorff, 1995; Leung, 2000a; Reed, 2001; Olmez, 2003). Some attempts to perform automatic classification explore the use of parametric classifiers or neural networks (Leung, 2000; DeGroff, 2001; Reed, 2001).

Time frequency methods have been applied to the analysis of signatures of paediatric heart murmurs (Leung, 1997, 1998a). The murmurs were considered as approximately cyclostationary and as a result the signal to noise ratio enhanced by averaging the time frequency distributions across cycles of heartbeats.

The use of signal processing techniques for heart sound analysis and recognition will be further explained in chapter 3.

1.2.3 Condition monitoring of prostethic heart valves

The primary function of the cardiac valves is to allow unidirectional flow of blood through the cardiac chambers. The normal valve performs this function without causing obstruction (stenosis) or reversal of flow (insufficiency), which ensures efficient transformation of energy from myocardial contraction into blood flow circulation throughout the body. When the valves malfunction sometimes compensatory mechanisms initially allow the heart to function efficiently, however, after some time these mechanisms become inadequate and the heart fails to meet the metabolic requirements of the body.

Either mechanical prosthetic valves or bioprosthetic valves extracted from pigs may replace native valves. Although the incidence of mechanical prosthetic valve fracture is low, mechanical prostheses are prone to sudden failure due to fracture of their components with major consequences and often leading to the death of the patient. Bioprosthetic valves may last from seven to twelve years but gradually fail due to tissue degeneration and calcification. The inevitable patient risk associated with both mechanical and biological prostheses requires periodical evaluations of prosthetic integrity.

Based on the theory that the opening and closure of heart valves contributes directly to heart sounds, PCG analysis offers a non-invasive and passive approach to the evaluation of prosthetic heart valves. Stein (1980) were the first to demonstrate that the dominant frequency of closure sounds produced by native and prosthetic valves was increased by stiffening of the valve leaflets due to calcification and fibrosis.

Durand (1990) studied the spectra of the first heart sound in order to determine the contribution of bio-prosthetic valves implanted in the mitral position in humans. He demonstrated that the band width of spectrum of the first heart sound broadens when the valves degenerate. Durand derived spectral parameters from the first heart sound spectra and used them to discriminate normal from degenerated bio-prosthetic valves achieving classification accuracy as high as 98%.

Durand (1986) studied the sounds of bioprosthetic valves in the aortic position. The comparison among the basic periodogram, Welch's averaged periodogram, all pole modelling, and pole zero modelling, showed that the basic periodogram provides the best compromise between spectral distribution and localisation of spectral frequency peaks of bioprosthetic valve sounds.

1.2.4 Detection of coronary artery disease

Coronary artery disease (CAD) is caused by the obstruction of coronary arteries by atheromatous plaques. Blood vessels are normally flexible, elastic, and pliant, with smooth internal surfaces. When a segment of blood vessel is hardened due to the deposition of calcium and other minerals, the segment becomes rigid. Furthermore, the development of plaque inside the vessels causes narrowing or constriction of the vessel, which impedes the flow of blood. Coronary artery disease is the leading cause of death in the western word.

The occlusion of coronary arteries is usually detected by Coronary angiography, which is an X-ray examination of the blood vessels or chambers of the heart. A very small tube (catheter) is inserted into a blood vessel in a groin or an arm. The tip of the tube is positioned either in the heart or at the beginning of the arteries supplying the heart. Then a special fluid visible by X-ray (called a contrast medium or dye) is injected. The X-ray pictures obtained (angiograms) are required in order to assess further treatment. Coronary angiography is a time consuming and expensive technique.

An alternative method consists of the development of non-invasive acoustic techniques for detection and characterisation of coronary artery disease. Akay *et al.* (1990a, 1990b, 1991, 1992, 1993, and 1994) conjectured that coronary artery disease could present high frequency sounds due to turbulence caused by the narrowing or constriction of the vessel. The study of the spectra of mid diastolic segments of the PCG over 20-30 beats of normal subjects and patients with coronary artery disease (confirmed by angiography) showed greater portions of the energy above 300 Hz when coronary artery disease was present. Subsequent studies demonstrated that relatively high power levels of resonance frequencies in the range 400-600 Hz, evident in patients with CAD, were reduced after angioplasty (Akay, 1990b).

In an attempt to improve the diagnosis ability of the technique wavelet transforms (Akay, 1994), min-max neural classifiers (Simpson, 1992, 1993), high resolution spectral estimators (Akay, 1990a, 1990b, 1991), and adaptive line enhancer techniques have been used (Akay, 1992, 1993).

1.3 Objective

Increased concern over the long term effects of radiation of various forms used in diagnosis medical imaging has reinforced the need for the development of passive

and non-invasive techniques (Verbug, 1979). This concern is particularly important in paediatrics due to the frail nature of the infants being examined. Moreover, premature births compounded by congenital defects strengthens the need for passive observation procedures.

Auscultation as a passive technique is specially suited for assessment of heart pathologies in infants. The technique has the potential to determine the presence of a heart murmur and hence, a pathological condition. Nevertheless, some limitations of this technique and its demand for high expertise from general practitioners means that often a non pathological murmur is not recognised as such and the patient needs to be referred for further analysis, bringing unnecessary concern, limitation of activity and needless antibiotic prophylaxis (Rushmer 1968).

The development of digital phonocardiology and the use of advanced signal processing techniques provide the means to overcome many of the limitations associated with classical auscultation. Autonomous systems to aid PCG diagnosis have been the focus of the latest research. However, the limited success of the different signal processing techniques applied to PCG analysis and classification has made evident the non-triviality of the problem and suggests the use of new approaches.

This work explores the use of a statistical pattern classification technique able to account for the inherent variability of the biomedical signals. Starting from a necessarily limited signal sample space available the method presented derives statistical models able to represent unseen data. The thesis also explores the use autonomous system-orientated representations to extract information based on the bio-dynamics of the system rather than the traditional features that provide visual meaningful representations. The main objective is to identify abnormal heart conditions through the analysis of the recordings of sounds produced by the heart (PCG) by means of a statistical pattern classification technique: Hidden Markovian Models (HMM).

The research is based on the classification of paediatric heart sounds corresponding to the following categories: Aortic Stenosis (AS), Atrial Septal Defects (ASDs), Ventricular Septal Defects (VSDs), Innocent Pulmonary Flow Murmurs, and normal heart sounds. These categories represent the most common conditions encountered in paediatric cardiology. The signals have been obtained from subjects in an age range between 3 months and 16 years old with the support of the Wessex Cardiac Unit of Southampton General Hospital.

1.4 Synopsis

This thesis starts with a brief introduction the traditional technique of auscultation presented in Chapter 1. The importance of the technique for diagnosis is presented in an historical context and in relation to contemporary techniques. Recent technological developments that lead to the development of digital phonocardiography have started new areas of research and diagnosis-oriented applications based on analysis of cardiovascular sounds. Those areas are briefly described before justifying the objective of this research: the identification of murmurs in paediatric PCG signals by means of statistical pattern classification techniques.

Chapter 2 introduces the biomedical background required for a good understanding of the study. Despite the controversy about the origin of heart sounds, the anatomy and physiology of the heart are presented in relation to the most generally accepted theory of the genesis of the heart sounds and murmurs: the Cardiohemic theory. The chapter also presents an overview of the characteristics described by physiologists for the classification of heart sounds and the specific signature for the systolic heart sounds studied in this research.

Chapter 3 introduces the use of signal processing techniques for PCG signal classification. Autonomous heart sound identification aimed to aid differential diagnosis is conceptualised as a signal classification task. After a review of the main approaches applied to phonocardiography, a technique of statistical signal

processing, namely the use of hidden Markovian models is proposed for this application.

Chapter 4 presents a brief introduction to hidden Markovian models. Basic concepts of the technique are explained through a practical example. The hidden Markovian model developed, although not related with the analysis of heart sounds, is useful to clearly show the relationship between the components of the model and the applications of this statistical technique. Finally, references to tutorials and details of the algorithms are provided.

Feature representations for analysis and classification of the phonocardiogram signal is the subject of Chapter 5. The most common techniques are mentioned and two of the main approaches are presented: linear prediction analysis and Mass–Weber filters. Neverthless, the use of line spectral frequencies is proposed as alternative representations for linear prediction coefficients. Inspired by the cardiohemic theory, a third approach for PCG signal representation using cepstral analysis is proposed. The viability of the latter technique for differentiation of PCGs characteristic of certain medical conditions is explored through examples.

The problem of PCG signal segmentation is addressed in Chapter 6. Since the identification of the main components of the heart sounds is considered the first step towards the automatic classification of heart sounds, a review of available techniques is presented. Segmentation is usually performed with the aid of the ECG signal. An algorithm for PCG segmentation using Cepstral analysis and HMM without an ECG reference signal is proposed. These techniques were selected in chapters 5 and 3 for signal classification and PCG representation respectively, and therefore, their use for signal segmentation is logical consequence. Finally, results obtained from the application of the algorithms on a database of hearts sound signals collected in collaboration with the cardiology unit of Southampton General Hospital are presented.

Chapter 7 explains the model selected for the classification of systolic murmurs. The algorithms are applied to real signals and the performance of the classifier is presented.

Chapter 8 presents a summary of the research and suggestions for future work.

1.5 Novel Contributions

- Use of hidden Markovian classifiers for PCG recognition.
- Cepstral representations of PCG based on the cardiohemic theory.
- Automatic PCG segmentation by hidden Markovian models without an external reference signal.
- Line spectral frequencies PCG representations for classification.
- Comparison of Mass-Weber filter PCG representations, line spectral frequencies, and cepstral coefficients for PCG analysis.

2 Biomedical Background

2.1 Overview

Basic knowledge of the related biological events is required in order to understand the problems involved in automatic phonocardiogram classification. This section describes the basic principles of the anatomy and physiology of the heart and the most accepted theory proposed for the genesis of heart sounds and murmurs. Principles of auscultation are also presented in relation to the analysis of phonocardiograms including an overall description of the general criteria used by cardiologists to identify a possible pathology from the heart sounds perceived. The main auscultation areas are briefly described before a short description of the physiological characteristics and particular signatures of the pathologies involved in this research and their related phonocardiographic signals. This section is only a brief introduction presenting the basic concepts needed, for a more detailed explanation refer to the medical literature (Rushmer, 1970; Sahver, 1985; Guadalajara, 1998; Kusumoto, 1999; Brown, 2002).

2.2 The Heart

The cardio-vascular system comprises the heart and blood vessels. The heart is the central organ of the entire system, and consists of a hollow muscle; by its contraction the blood is pumped to all parts of the body through a complicated series of tubes, termed arteries. The arteries have multiple ramifications and end in very minute vessels called arterioles, which in turn open into a mesh network of microscopic vessels called capillaries. In this passage through the capillaries of the body, the blood gives to the tissues the materials necessary for their growth and nourishment. In the capillaries the blood receives from the tissues the waste products result from metabolism and afterwards, it is collected into a series of larger vessels called veins and returned to the heart.

The heart is divided by a septum into two halves, right and left, and can be regarded as a four-chamber pump with the two upper auricle chambers receiving blood and two lower ventricles that pump blood away to organs. All the chambers are connected by valves, which allow blood to move in one direction from one chamber to the next. Figure 2.1 shows the basic anatomy of the heart.

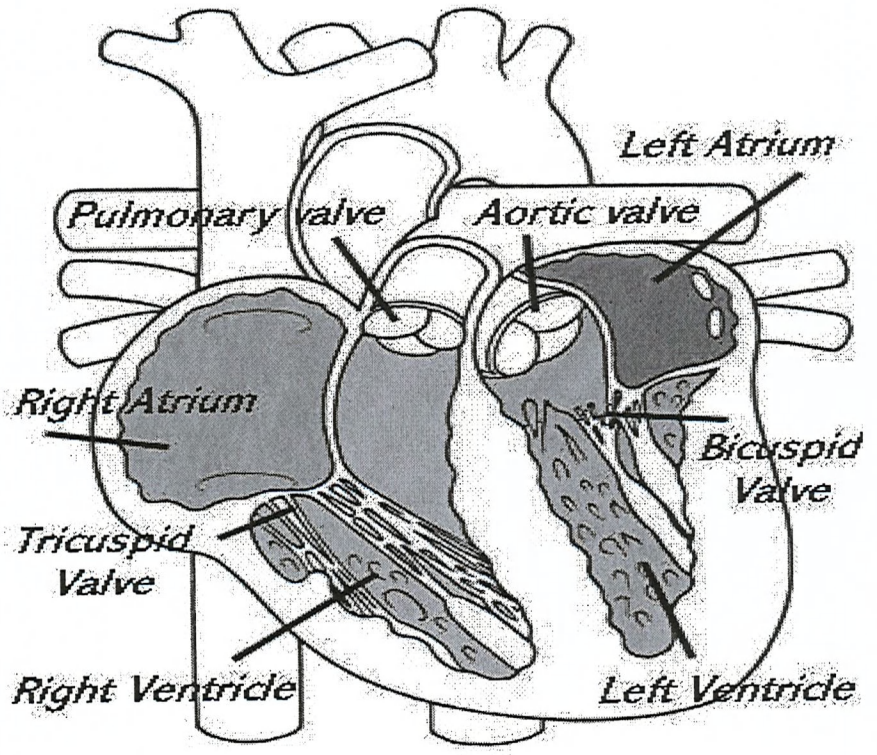


Figure 2.1. Heart Chambers. The Sourcebook of Medical Illustration, P. Cull, ed., The Parthenon Publishing Group, 1989.

In a normal cardiac cycle the right atrium receives blood from all the organs, except the lungs, through the vena cava. The tricuspid valve controls the passage of this oxygen-poor blood to the right ventricle, which in turn pumps blood to the pulmonary artery through the pulmonary valve. The pulmonary artery takes blood to the lungs where it is oxygenated, the oxygen-rich blood is then returned to the left atrium by the pulmonary veins. The mitral valve controls the passage of blood from the left atrium to the left ventricle. Finally, during ventrical contraction, the aortic valve lets the blood pass to the aorta and from there, it passes to the rest of the body.

2.3 Heart Sounds

There is a wide diversity of opinion concerning the theories attempting to explain the origin of heart sounds and murmurs. More than 40 different theories have been proposed to explain the origin of the first heart sound (Rushmer, 1970; Rangayyan, 1988). Although the controversy continues about the exact origin of the heart sounds (especially for the first and the third sounds), the concepts around the most accepted theory, which is used for cardiology teaching, are presented here (Rushmer, 1970; Kusumoto, 1999; Brown, 2002).

The cardiohemic theory proposed by Rushmer in 1970 assumes that since the chambers of the heart are filled with blood, none of these structures can vibrate independently without producing movements of blood. Similarly, vibrations in the blood must be transmitted to surrounding structures. Therefore, all structures: heart cavities, valves, and blood constitute an interdependent system vibrating at the same time. Heart sounds are regarded as vibrations or sounds due to the acceleration or deceleration of blood within the elastic system triggered by pressure gradients (Rushmer, 1970). The elasticity of a chamber completely filled with fluid is analogous to a spring whereas the fluid and supporting mass are analogous to a vibrating mass. Any sudden movement such as acceleration and deceleration throws the system into vibration.

A normal cardiac cycle contains two major sounds: the first heart sound (S1) and the second heart sound (S2). Figure 2.2 shows the main heart sounds and their relation to the electrical events of the cardiac cycle. The systolic and diastolic components are defined in relation to S1 and S2. The systolic component is usually shorter than the diastolic; a fact that is generally used during auscultation to distinguish the second heart sound from the first one.

Figure 2.3 shows the main components of an ideal heart sound. The first heart sound has four components associated with the movement of blood during the contraction of the ventricles. As the ventricles contract (systolic phase), blood shifts towards the atria, closing the atrioventricular valves (mitral and tricuspid)

with a consequential oscillation of blood generating the first component of S1. The second component of S1 begins with abrupt tension of the closed atrioventricular valves, decelerating the blood. Then, the semilunar (aortic and pulmonary) valves open and the blood is ejected out of the ventricles. A third component of the first heart sound further originates from oscillations of blood between the descending root of the aorta and ventricle, and a fourth component may arise from vibrations due to blood turbulence at the aortic and pulmonary valves. The first heart sound begins immediately after the peak of the QRS complex of the ECG.

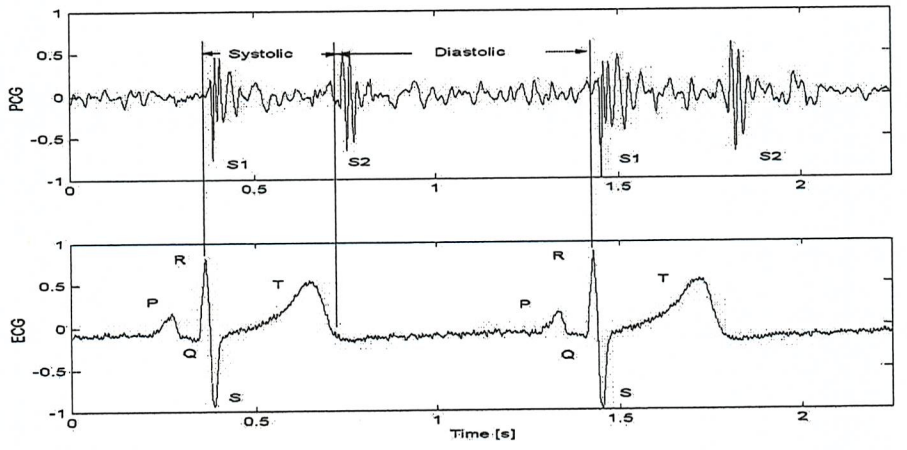


Figure 2.2. Correlation of the Components of the Heart Sounds with the Electrical Events. PCG Recorded at mitral area, ECG lead I, Fs =4 KHz.

Following the systolic pause in a normal cardiac cycle, a second heart sound occurs. S2 is coincident with the end of the T wave of the ECG since both events are related to the end of ventricular contraction. Nevertheless, the T wave is commonly referred to as the wave corresponding to ventricular relaxation through repolarisation. While this is indeed correct, relaxation through repolarization is just the final phase of contraction, but contraction and relaxation are indicated by the upstroke and downstroke of the same action potential. The T wave therefore may be said to relate to a non-specific event (Rangayyan 2002) whereas the S2 is related to the specific events of closure of the aortic and pulmonary valves. This fact and the low amplitude of the T wave (0.1 - 0.3 mV) which is almost absent in

many recordings suggest the use of another auxiliary signal: the carotid pulse for S2 identification in autonomous systems as explained in chapter 6.

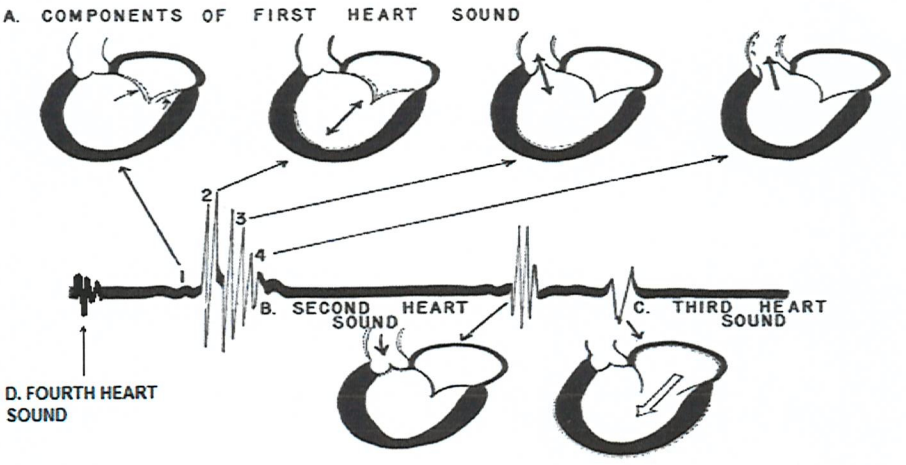


Figure 2.3. Components of an ideal heart sound. From R. F. Rushmer Cardiovascular dynamics, W. B. Saunders, Philadelphia, PA, 1976.

S2 is a low frequency vibration associated with the deceleration and reversal of flow in the aorta and pulmonary artery and with the closure of the semilunar valves (aortic and pulmonary). While the primary vibrations occur in the arteries due to deceleration of blood, the ventricles and atria also vibrate, due to transmission of vibrations through the blood, valves, and the valve rings.

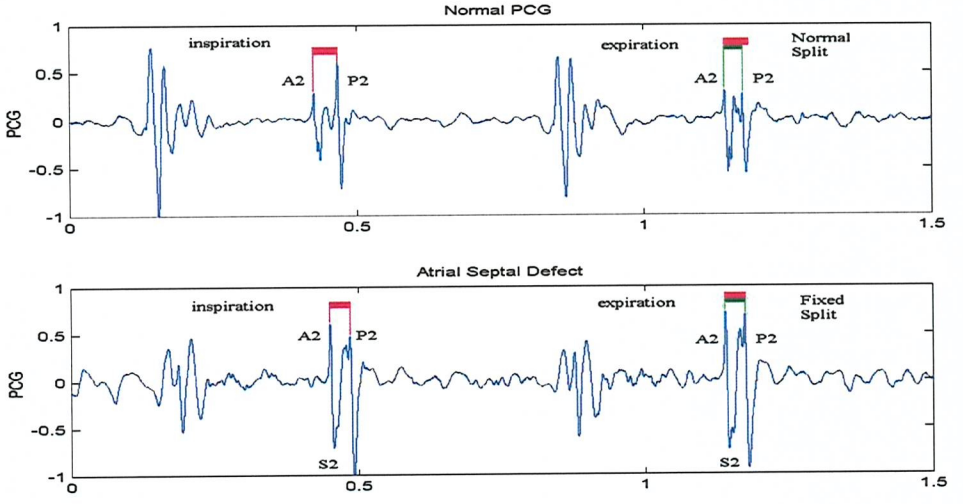


Figure 2.4. Normal and fixed split of S2. In a normal PCG the split of S2 during inspiration (pink square) last longer than during expiration (green square). In an ASD condition both last the same and the split is said to be fixed. Normal PCG Recorded at mitral area, ASD recorded at Pulmonary area, Fs =4KHz.

The second sound is composed of two components (A2 and P2) corresponding to the closure of the aortic and pulmonary valves respectively. Under normal, conditions during expiration, P2 appears from 10 to 30 ms after A2. With inspiration, overload of the right ventricle due to venous return causes a longer delay in the closure of the pulmonary valve and thus of P2. Consequently, it is easier to perceive both components. Pathological conditions could cause this gap to widen, may reverse the order of occurrence of A2 and P2, or fix the delay as in the case of atrial septal defect conditions as shown in figure 2.4 and explained in section 2.6.2.

The third sound is attributed to the sudden termination of the rapid filling phase of the ventricles from the atria and the associated vibration of the ventricular muscle walls, which are relaxed. The fourth or atrial heart sound occurs when the atria contracts and propels blood into the ventricles. These sounds are generally much quieter than the first and second ones. In addition to these sounds, valvular clicks and snaps are occasionally heard.

2.4 Murmurs

The intervals between S1 and S2 during ventricular systole, and S2 and S1 during ventricular diastole are normally silent, but murmurs may occur during those intervals. According to Rushmer heart murmurs are prolonged sounds arising from turbulent blood flow from one of four causes:

1) High rates of flow through normal valves.

- Pathological related: Murmurs may be present in pathological conditions like anemia and thyrotoxicosis, for example, where the patients have structurally and functionally normal hearts, but the high-blood flow indicates an underlying disease.

- Non pathological related: As in the case of innocent pulmonary flow murmurs, where the cardiac anatomy is normal, common in childhood and young adults and also in hyperdinamic states such as pregnancy.

2) Forward flow through a narrowed or irregular shaped valve. May occur for example when a valve does not open completely due to a calcium deposit (acquired stenosis), or in congenital malformation such as bicuspid aortic stenosis.

3) Backward flow through a leaking heart valve. Regurgitation or insufficiency occurs when valves fail to occlude completely. It can be associated with a rheumatic origin, congenital malformation, atrophy of the tendons or muscles involved in its function, or functional origins in an anatomically normal valve due to annular dilatation such as in the case of functional mitral regurgitation.

4) Flow through an abnormal cardiac or extracardiac connection. Such as in ventricular septal defect where a hole in the ventricular septum allows blood to flow from the left ventricle to right ventricle.

2.5 Auscultation

Features of heart sounds and murmurs like intensity, frequency content, and timing are affected by many physical and physiological factors such as the recording site on the thorax, intervening thoracic structures, left ventricular contractility, position of the cardiac valves at the start of the systolic cycle, the degree of the defect present, the heart rate, and blood velocity. Despite the variations introduced by all this facts, it is possible to distinguish between different murmurs by their particular morphology.

To identify a specific abnormal condition the information obtained from the heart sound is described in terms of frequency patterns (pitch and quality), intensity (loudness and patterns of sound intensity), timing (systolic and diastolic cycles),

localisation (where the murmur is best heard and spread pattern) and whether the murmur changes during respiration. Table 2.3 shows the information utilised in evaluating the significance of the heart sounds and murmurs (Rushmer, 1968).

A. Frequency patterns
   1. Pitch (high or low)
   2. Quality (presence of dominant frequencies or harmonics)

B. Intensity
   1. Loudness of heart sounds
      Grade I: very faint, difficult to hear
      Grade II: faint, but easily heard
      Grade III: moderately loud
      Grade IV: very loud
      Grade V very loud, can be heard with the edge of the stethoscope to the skin
      Grade VI: extremely loud, can be heard with the stethoscope off the skin
   2. Patterns of sound intensity of murmurs
      a. Crescendo
      b. Decrescendo
      c. Crescendo-Decrescendo ("Diamond shaped")
      d. other

C. Timing
   1. Splitting of sounds
   2. Systolic murmurs
   3. Diastolic murmurs
      a. Early diastolic
      b. Mid-diastolic
      c. Protodiastolic
   4. Intervals between sounds and murmurs
   5. Gallop rhytms and similar
   6. Identification of opening snap

D. Localisation on precordium
   1. "Mitral" area
   2. "Aortic" area
   3. "Pulmonary" area
   4. "Tricuspid" area

Table 2.3. Types of Information Employed in Auscultation Diagnosis.

The limitations of auscultation are indicated by the common tendency to classify heart murmurs according to a few general types (Table 2.4) based primarily on timing and intensity characteristics shown in Table 2.3 (Rushmer, 1968).

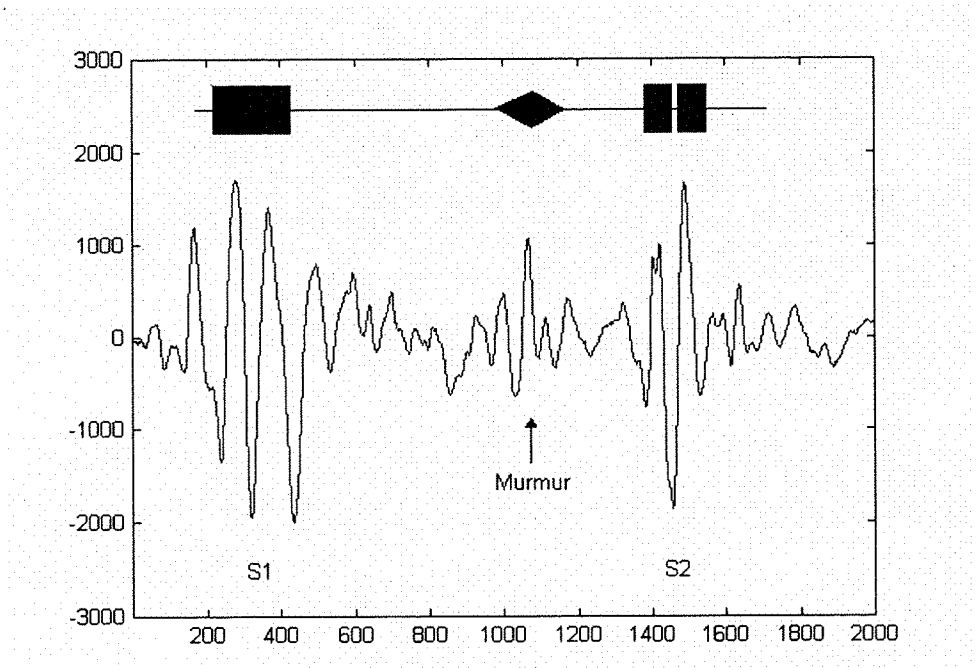


Figure 2.5. 'Diamond shaped' murmur. Named by its Crescendo – Decrescendo characteristics and well defined start and end points PCG Recorded at pulmonary area, Fs=4KHz.

Accordingly to this time and intensity classification Figure 2.5 shows a grade II crescendo – decrescendo or 'diamond' shaped mid-systolic murmur. Note that high frequency information is not evident in this time-domain plot since low frequency components have higher amplitude. Filters such as the Maass-Weber set are therefore required to extract the information provided by high frequency components.

It is important to note that murmurs vary in intensity and an increase in loudness does not necessarily indicate an increase in the severity of the disease.

If physicians were forced to rely on interpretations of heart murmurs without other clues (i.e., only by the recorded signals) they would be able to distinguish several general types of murmurs but could not specify confidently the kind of disease. To

the extent that a particular kind of murmur can be caused by a variety of pathological conditions, its diagnostic specificity is limited. The practitioner differentiates those conditions and provides a diagnosis based not only on the heart sounds, but considering the entire clinical scenario.

I. Systolic murmurs
   A. Ejection murmurs
      1. Aortic valvular stenosis
      2. Pulmonary valvular stenosis
      3. Infundibular stenosis
      4. Dilatation of aorta or pulmonary artery
      5. Increased ejection outflow rates
II. Diastolic murmurs
   A. Early diastolic decrescendo murmurs
      1. Aortic regurgitacion
      2. Pulmonary regurgitacion
   B. Atrioventricular diastolic murmur
      1. Mitral stenosis
      2. Tricuspid stenosis
III. Continuous murmurs
   A. Patent ductus arteriosus
   B. Intrathoracic arteriovenous fistulae
      1. Coronary A-V fistulae
      2. Ruptured sinus of Valsalva
      3. Intrapulmonary A-V fistulae
   C. Coarctation (systemic collaterals)
   D. Venous hum
   E. Other anomalies
IV. Musical murmurs
   A. Everted aortic cusp
   B. Chiari nets
   C. Moderator bands

Table 2.4. Categories of Murmurs.

2.5.1 Auscultation areas

Externally, particular heart sound components are best heard at certain locations on the chest, and this localisation has led to the concept of secondary sources on the chest.

Figure 2.5 shows the anatomical position of the auscultatory areas and their localisation with reference to the heart valves. Notice that the particular location of sounds on the surface probably represents the most effective transmission paths and not necessarily the shortest path.

Murmurs from the region of the pulmonary valve are most intense in the pulmonary auscultatory area, centred at the third intercostal space at the left parasternal line. The aortic area lies to the right of the sternum in the second intercostal space where the ascending aorta curves forward and most closely approaches the anterior chest wall. The tricuspid area is near the right sternal border in the fourth intercostal space, and the mitral area is near the apex of the heart where the heart is in direct contact with the anterior wall of the thorax (Rushmer, 1970).

2.6 Paediatric heart sounds

This research will be focused on the classification of paediatric heart sounds into one of five categories, two non-pathological conditions: normal and innocent

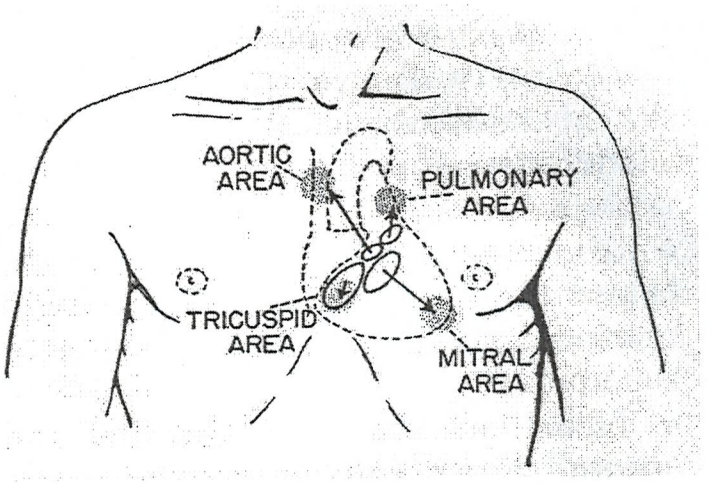


Figure 2.5. Auscultation Areas. From 'Cardiovascular Dynamics' Rushmer, 1970.

murmurs, and three pathological systolic murmurs caused by aortic stenosis, atrial septal defect, and ventricular septal defect.

2.6.1 Valvular Aortic Stenosis

The aortic valve is composed of three cusps of equal size attached symmetrically around the circumference of the valve orifice. The normal function of the aortic valve is to allow blood flow from the left ventricle to the body without obstruction or reversal of flow. The normal aortic valve closes completely and when open has a triangular orifice considerable smaller than its cross-sectional area; however, it is sufficiently large so the pressure gradient required to force blood through it is negligible. Stenosis occurs when a valve does not open completely. The causes of aortic stenosis can be divided into congenital and acquired forms.

A patient with a congenital bicuspid aortic stenosis has a valve with only two cusps. Two cusps of equal size are equally effective in completely closing but a bicuspid valve can not open widely and obstructs the flow of blood. Other forms of congenital disease include the formation of a unicuspid valve and a tricuspid valve with fused commissures. Persons born with an abnormal bicuspid valve are particular susceptible to calcification in later life.

Acquired stenosis may arise from secondary conditions such as rheumatic heart disease or idiopathic calcification of the valves. The valve may become hardened or stiff with calcium deposits or scarring, and thus, it is hard to push open. Blood has to flow through a smaller opening, and therefore less blood gets through the valve into the next chamber. The origins of aortic stenosis in infants and children are always congenital. In teenagers, this condition would be congenital when found alone, and rheumatic whenever it is accompanied with mitral stenosis although the latter patterns are characteristic of adults. Aortic stenosis in patients over 65 years old is generally due to senile, degenerative or calcific origins, resulting from mechanical wear and tear, and calcium deposits.

The normal aortic valve area is approximately from 3 to 4 cm^2 and starts to produce a pressure gradient when reduced to 1.5 cm^2 and below. The expulsion time of the left ventricle is prolonged accordingly to the severity of the stenosis in order to compensate for the obstruction of blood flow out of the left ventricle to the aorta. The left ventricular pressure increases due to the restriction and the ventricle compensates for this pressure overload by becoming hyperthorphic. The wall thickness increases to provide the necessary force to open the stenotic valve. Due to this compensatory mechanism patients with aortic stenosis remain asymptomatic during the long latent phase of the disease. Symptoms develop as the valve orifice narrows and the hyperthorpy is not enough to compensate. Angina (chest pain or discomfort that occurs when the heart muscle does not get enough blood), syncope (transient loss of consciousness) and congestive heart failure (a condition in which the heart can not pump enough blood to the body's other organs) are among the symptoms of advanced aortic stenosis.

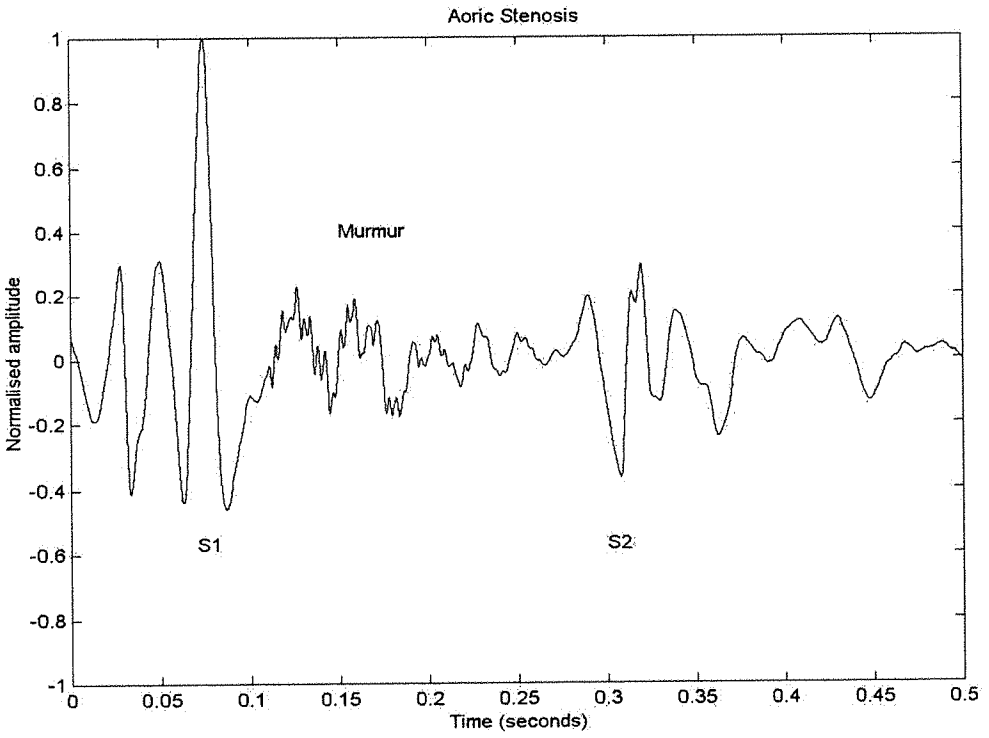


Figure 2.6. Aortic Stenosis Murmur. PCG recorded at aortic area, Fs =4 KHz.

Figure 2.6 shows a phonocardiogram (PCG) of a heart sound with a murmur due to aortic stenosis. These murmurs are typically a mid-systolic (occurring in the middle of the systolic cycle) ejection (discrete start and end points) murmurs heard best over the aortic area. The aortic stenosis murmur has a harsh quality, medium pitch, coarse, crescendo-decrescendo quality. The murmur peaks later in systole as the stenosis worsens since a larger pressure gradient with a longer ejection period are required to open the aortic valve. When the left ventricle begins to fail the murmur becomes softer and therefore, the intensity of the murmur does not necessarily correlate with the severity of stenosis.

An aortic ejection sound (ejection click) may also be present early in the natural history of aortical stenosis. The ejection click is caused by the sudden opening of a stiff aortic valve in ventricular systole and depends on valve mobility. This sound is high pitched and occurs approximately 60 ms after S1. The ejection click is heard better at the apex than in the aortic area (where the murmur is loudest but the click may be inaudible).

Because the second heart sound is largely generated by the sudden closing of the aortic valve, a poorly mobile and stenotic aortic valve may cause the second sound (S2) to become quieter or even be absent. S2 is normally created by the closure of the aortic valve originating the A2 component, followed by the pulmonary valve that originates the P2 component. In aortic stenosis, the aortic component A2 becomes progresively delayed as the degree of stenosis worsens and the left ventricular period becomes consequently simultaneously longer. If the closure of the aortic valve is delayed sufficiently, it may close after the pulmonary and therefore, two distinct sounds can be heard at expiration rather than inspiration, this condition is called abnormal paradoxical split of S2 (Isselbacher, 1984).

2.6.2 Atrial Septal Defect

An atrial septal defect (ASD) is a hole in the wall (septum) between the atria. ASD accounts for approximately 7% of all congenital heart lesions (Kusumoto, 1999).

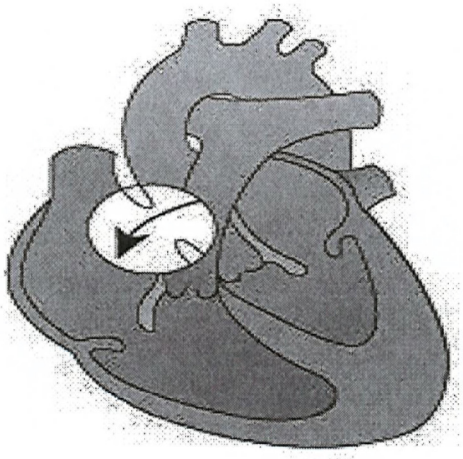


Figure 2.7. Atrial Septal Defect.

Depending on the site of the defect four categories can be identified, namely: secondum ASD, the most common type of congenital disease, endocardial cushion defect, sinus venosus type and common atrium. In all cases, due to higher compliance of the right atrium and slightly higher pressure in the left one, the shunt allows blood to seep from the left into the right atrium, as shown in Figure 2.7. Since the flow rate through the defect is slow, turbulent flow is unlikely due to this shunt, however, the higher volume of blood ejected by the right ventricle through a normal pulmonary valve produces a prominent mid-systolic flow murmur. To compensate for this increased volume, the right atrium and right ventricle hypertrophy and subsequently dilate. Additionally, pulmonary blood flow increases with subsequent enlargement of the pulmonary arteries and veins. Over time, this continued overload produces pulmonary vascular disease and pulmonary hypertension. With increased right pressure the shunt of blood left to right decreases and the shunt reverses allowing the entrance of non oxygenated blood to the systemic circulation causing the development of cyanosis.

Patients with ASD are generally asymptomatic until the third or fourth decade of life. The first symptoms are dyspnea (an unpleasant sensation of difficulty in breathing), fatigue, chest discomfort, cyanosis, and hemoptysis (coughing up of blood from the respiratory tract).

Figure 2.8 shows a typical ASD murmur. This mid-systolic ejection murmur, usually grade II, is produced by hyperflux in the pulmonary valve and consequently is best heard over the pulmonic area of the chest. With larger shunts, a mid-diastolic tricuspid murmur may also be audible.

The most characteristic feature of an ASD is the fixed split S2 as shown in figure 2.4. In a normal heart, a split S2 is caused physiologically during inspiration because the increase in venous return (due to a compliance decrease of the pulmonary arterial bed) overloads the right ventricle and delays the closure of the pulmonary valve. With an ASD, the right ventricle can be thought of as being continuously overloaded because of the left to right shunt, producing a widely split S2. Because the atria are linked via the defect, inspiration produces no net pressure change between them, and has no effect on the splitting of S2. Thus, S2 is split to the same degree during inspiration and expiration, and the split is said to be fixed.

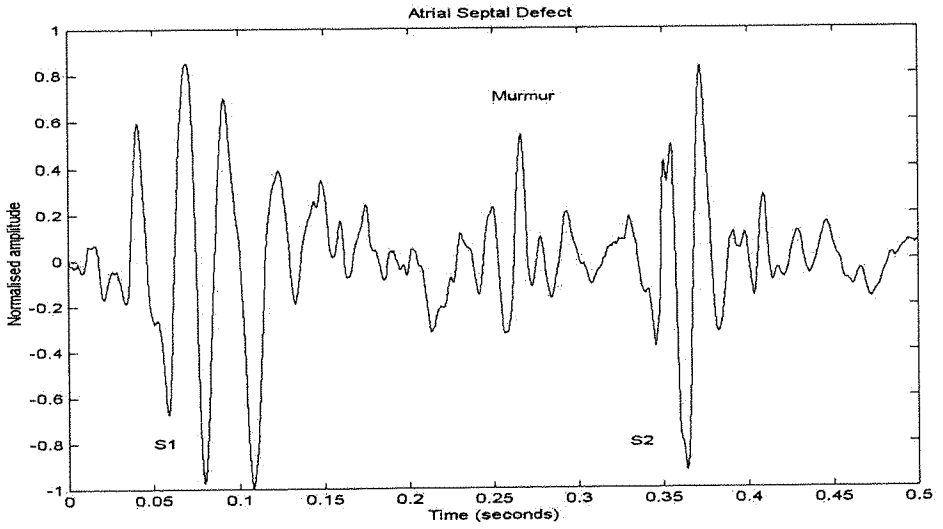


Figure 2.8. Atrial Septal Defect Murmur. PCG recorded at pulmonary area, Fs =4 KHz

2.6.3 Ventricular Septal Defect

Ventricular septal defects (VSDs) are the most common congenital malformation in infant and children, accounting for 25% of all congenital cardiac malformations. This is a type of heart defect (congenital heart disease) in which there is a hole in the septum between the ventricles. Figure 2.9 [1] shows a ventricular septal defect (VSD).

Ventricular septal defects are classified form an anatomical point of view in relation to their relative position in either the membranous, muscular or outlet portions of the septum. From a physiological point of view VSD are classified according to the relative size of the opening related to the aortic diameter and hemodynamical repercussions.

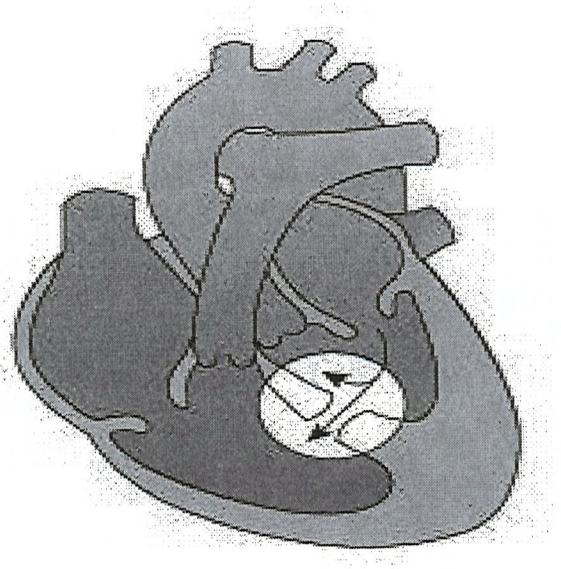


Figure 2.9. Ventricular Septal Defect.

In a foetus, during the last stages of partitioning of the heart, an aperture in the interventricular septum exists. Endocardial cushion tissue proliferates in this aperture and later forms the membranous portion of the interventricular septum. Incomplete closure of this opening lead to perimembranous VSD (that usually

extends to adjacent muscular septum) which is the most common ventricular septal defect: accounting for 65% of cases.

The least common muscular septal defects (accounting for 18% of cases) may appear anywhere in the muscular portion of the septum. The defect is probably the result of excessive resorption of the myocardial tissue during the formation of the interventricular septum.

The third type of defect is located in the outlet septum. This defect accounts for 8% to 12% of cases, and can be associated with prolapse of the aortic valve (out of its usual position) and aortic regurgitation. The fourth type relates to defects of the inlet septum accounting for approximately 7% to 10% of VSD.

The pressure in the left ventricle during the systolic cycle is significantly higher than that present in the right ventricle by a ratio of four to one and therefore an aperture between cavities propels oxygenated blood from the left ventricle into the right. The degree of flow through the shunt is related to the size of the defect and the relative resistances of the pulmonary and systemic circulation. With a small defect, there is a significant resistance to flow with a pressure gradient between ventricles. In case of large defect there is no restriction of flow between ventricles and equilibration of pressure within ventricles occurs. Large defects with significant pulmonary vascular resistance result in bi-directional shunting with a predominant right to left shunt leading to cyanosis.

Figure 2.10 shows a typical VSD murmur. These murmurs are characteristically holosystolic (completely filling the systolic phase of the cycle) since the pressure difference between the ventricles is generated almost instantly at the onset of systole, with a left to right shunt continuing throughout ventricular contraction. Small and medium shunts tend to generate harsh decrescendo murmurs. With large shunts, an intense holosystolic constant murmur is present and an apical mid-diastolic murmur may be appreciated secondary to increased flow through the mitral valve. The P2 component may be reinforced as a consequence of hypertension. VSDs murmurs are usually best heard over the tricuspid area. Since

the flow-pattern usually goes from left to right, the right ventricle suffers from volume overloads and takes longer to eject the stroke volume. This causes a slight delay in the closing of the pulmonary valve, and widely split S2 may result.

Patients with small VSD may be asymptomatic, while those with large shunts may complain of dyspnea and exercise intolerance. In some cases the defect spontaneously decreases in size or closes completely, spontaneous closure occurs in approximately 45% of patients by 3 years of age. Patients with small VSD are managed medically while large VSD often undergo surgical closure of the defect. Long term left to right shunts may lead to shunt reversal producing Eisenmenger's syndrome. These patients develop cyanosis, markedly diminished exercise tolerance, severe dyspnea, angina and dysrhythmias.

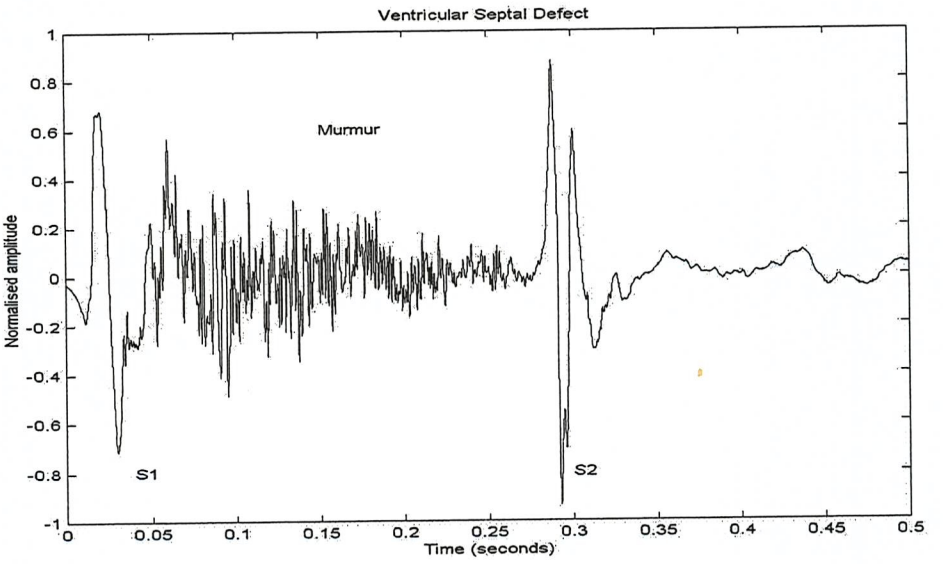


Figure 2.10. Ventricular Septal Defect Murmur. PCG recorded at middle left sternal edge area, Fs =4 KHz.

2.6.4 Innocent pulmonary flow murmur

Innocent murmurs are very common in children; 60% of all school-age children will have a murmur on random auscultation, and the majority will present an innocent murmur (Guadalajara, 1998). From the different classes of innocent murmurs innocent pulmonary flow murmurs are the most common. This kind of

murmur is actually registered in all normal subjects by intracardiac phonocardiography (Guadalajara, 1998).

Innocent pulmonary flow murmurs are non-pathological murmurs caused by the ejection of blood through a normal pulmonary valve. Since children have a more dynamic flow of blood, turbulent blood flow sometimes develops in the pulmonary artery and causes the murmur. Since the cardiac anatomy and physiology of the heart is normal there will be no cardiac symptoms associated with an innocent pulmonary flow murmur, however, children with symptoms such as chest pain which are actually non-cardiac in origin, may have co-existing innocent murmurs.

Figure 2.11 shows an innocent pulmonary flow murmur. This murmur is mid-systolic and is best heard in pulmonary auscultation area. It is seldom more than grade II (faint but easily heard, according to table 2.3), without spread and occupies approximately two thirds of systole. In children with particularly thin chest walls the murmur, and in general all the heart sounds, will be louder since a negligible amount of energy is lost in transmission to the surface.

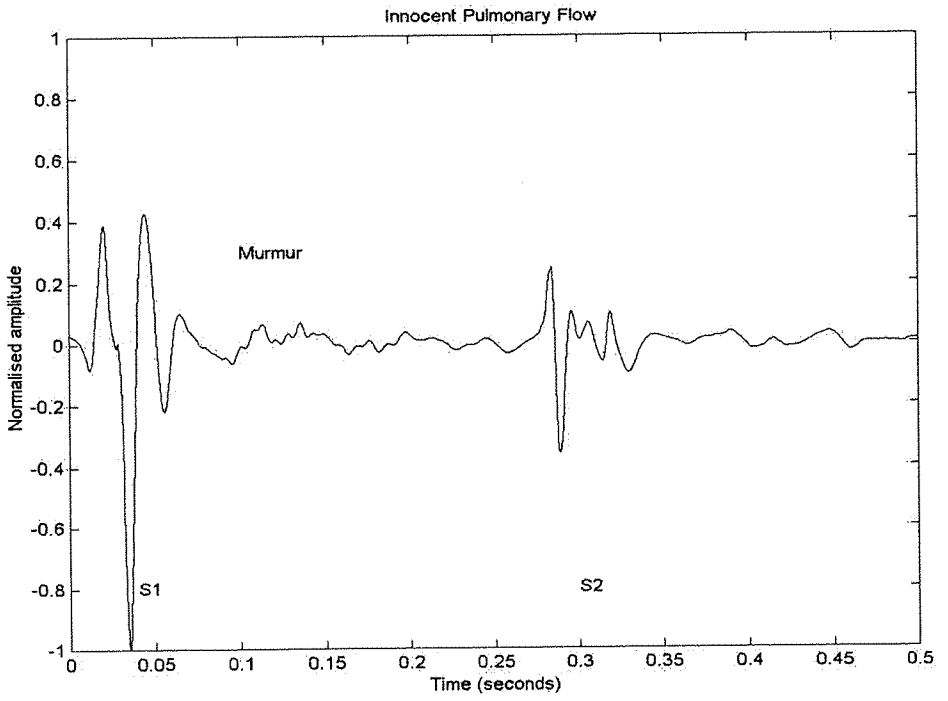


Figure 2.11. Innocent Pulmonary Flow Murmur. PCG recorded at pulmonary area, Fs =4 KHz.

The murmur has a diamond shape (a crescendo-decrescendo intensity pattern), is generally of soft quality, and has predominant mid-frequency components. The murmur tends to decrease during inspiration and also decreases when the child sits. The second heart sound splits and varies normally, and no diastolic murmur exists over the right ventricle at the lower left sternal border. Absence of all of these pathologic findings provides confidence that the pulmonic flow murmur is not associated with an atrial left to right shunt or pulmonic valve stenosis.

It should be stressed that although a description of the murmur can be provided there are no phonocardiographic findings that identify an innocent murmur with certainty. Differential diagnosis is made based on the absence of features characteristic of other possible causes and with the help of clinical manoeuvres.

2.7 Summary

A brief overview of the biomedical aspects related to the genesis of the heart sounds and murmurs has been presented as an introduction to the problem of phonocardiogram classification for the diagnosis of heart clinical conditions.

The main classification of heart murmurs, the systolic murmurs to be studied in this research and the pathological conditions related to them were presented.

This chapter also introduced the terminology employed in auscultation diagnosis to describe the frequency patterns, intensity, timing, and localisation of murmurs. These terms reflect the kind of information that the specialist looks for in order to identify abnormal heart sounds.

The next chapter presents a discussion about the signal processing techniques applied to PCG analysis and classification aimed to aid diagnosis.

3 Classification of PCG Signals to Aid Diagnosis

3.1 Overview

As described in previous chapters, the use of digital stethoscopes and advanced signal processing techniques can help to overcome some of the intrinsic limitations inherent to the traditional auscultation technique.

One of the most promising areas in the use of these techniques is the automatic identification of specific cardiopathologies that have proven to be difficult to define by differential diagnosis through traditional auscultation methods, such as the case of innocent murmurs and pathological related murmurs. Based only on the sound signal, a trained cardiologist can easily distinguish a pulmonary regurgitaton murmur from an aortic stenosis murmur. However, the distinction between, an innocent pulmonary flow murmur and an atrial septal defect murmur may be very difficult. An autonomous system that could ease the identification of similar murmurs would constitute a valuable tool to aid diagnosis.

From the signal processing point of view, the problem can be defined as a classification task. An autonomous system could assign a heart sound with a murmur to a specific class representative of a certain medical condition by objective analysis of its characteristics. Therefore, an innocent flow murmur would be correctly identified if the system were able to assign it to the innocent class rather than to the atrial septal defect class.

The identification problem can be simplified to a classification task if the signal in question belongs to one of the categories available. This approach could be helpful to differentiate murmurs with similar signatures that are known to belong to one of the available classes. However, a murmur not belonging to any of the classification groups will be wrongly assigned to one of the classes available. Consequently, this approach is only valid to confirm differential diagnosis of a specific group of pathologies with similar acoustic signatures. Other murmurs that may have similar acoustic signatures, and do not belong to the classes available,

should be previously discarded with the help of non-acoustical clues, symptoms and clinical manoeuvres. For example, the murmur of a rheumatic stenotic aortic valve of arterosclerotic nature, although similar to the murmur of a functional congenital aortic stenosis, can be differentiated simply using the age of the patient.

3.2. Signal classification

As shown in Figure 3.1, the recognition of the time sequences of the recorded PCG signals for classification is divided into four major components, namely: signal conditioning, segmentation, feature extraction, and classification of the features associated with that pattern. The performance of a signal classification procedure depends on the selection of the feature representations of the signal and the use of a convenient classifier technique for such representations. Matching these procedures eases the classification task. The work in the field has considered various combinations of these steps; although not necessarily exhaustively.

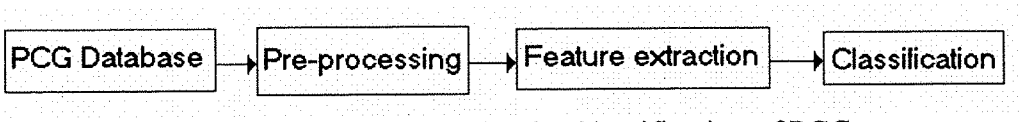


Figure 3.1. Main Stages for the Classification of PCG.

3.3 Classification approaches

The first step in any pattern-classification system is to select some representation of the input pattern. Although in some cases this is the raw input data, in general, compressing the data into a few salient features improves overall system performance.

The aim of feature extraction is to represent the signal by a choice of features that reduce variability between samples associated with the same class, whilst increasing the variability between samples belonging to different classes. In some

sense, the pattern classification can be solved trivially if the selected features are good enough (Gold, 2000, p. 105).

Given a feature vector choice, pattern classification primarily consists of the development, or training, of a system for classification of a large number of examples of the different classes. In a supervised learning process a database of patterns, labelled with the correct class, is provided for the training; whereas in an unsupervised process the data feature itself defines the classes. In the case of PCG classification, the classes are predefined by the cardiopathologies associated with the acoustical signals and therefore a training set of labelled signals, representative of each medical condition, is provided for the supervised learning process.

The computed features of each heart beat period are interpreted as a multiple dimensional feature vector within a corresponding feature space. Using this feature vector as the input to the classifier, an automatic decision is achieved generally by computing a suitable predefined distance between test and reference patterns or by assuming that separate clusters are formed in the feature space and that a decision function can be used to distinguish the boundaries. Nevertheless, decision functions are highly dependent upon the set of training samples provided. Their success when applied to new cases will depend on the accuracy of representation of the various pattern classes by the training samples.

If distinctive class distributions in the feature space are attained and the classification problem can be described algorithmically, deterministic or statistical classifiers could be utilised. The deterministic methods exploit some known specific properties of the signal. Specification of the signal model is generally straightforward; it is only required to estimate the values of the parameters of the signal model. The statistical approach, conversely, tries to characterise only the statistical properties of the signal. The underlying assumption is that the signal can be characterised as a parametric random process and that the parameters inherent to the stochastic process can be estimated (Rabiner, 1989).

Linear discriminant analysis is an approach that tries to maximise variance among classes and minimise variance within classes, given the constraint of a linear transformation of the input features. Discriminant analysis has been used in the PCG classification problem. Atrial septal defect murmurs, ventricular septal defect murmurs, and normal PCG paediatric signals were analysed by Leung (1998a) in a study conducted at the Institute of Sound and Vibration Research (ISVR) with the aim of classifying them. The systolic signals were characterised in the time-frequency domain by a representation based on the averaged spectrogram from which the spectral features were extracted to form feature vectors. Discriminant analysis was then performed to project the higher dimensional spectral features onto a two-dimensional space, in which the three groups form natural clusters that were linearly classified. Although the study shows three distinctive clusters, some samples for VSD and ASD lay near to the Normal cluster (Leung, 1998a).

An example of the use of statistical classifiers is the work conducted by Joo (1983) to study heart valve function. He proposes the use of pole-zero modelling to provide a high-resolution spectral estimate from which the frequency domain features were derived. The two maximum spectral peak locations are the features selected for the classifier. The design of the classifier was based on a statistical model for the dependence of the peak's frequencies upon the state of the valve. For PCG analysis, the probability density function is assumed to be Gaussian. A Gaussian classifier is then used to define a quadratic decision surface in the feature space. A set of twenty PCG signals from abnormal and normal aortic valves was used for training, whereas another set of the same magnitude was designed for testing. Seventeen of the twenty signals presented were correctly classified. Only one of the misclassified signals had a feature vector that was not near the decision boundary in the feature space. However, the need of a large patient population to prove the effectiveness of the method is acknowledged.

If the class related clusters in the feature space are not arranged in distinctive areas, nonlinear decision functions would be needed for classification. For simple cases of non-linear decision surfaces, quadratic surfaces can be derived.

Nevertheless, for more complex cases, artificial neural networks can be utilised to attain more general surfaces.

3.4 Neural networks

Artificial neural networks (ANN) have long been a subject of research interest for many pattern recognition tasks, including classification of phonocardiograms. Trimmed mean spectrograms and self organising maps (Leung, 1999) trimmed mean spectrograms and probability neural networks (Leung, 2000b) FFT representations and multiple layer perceptrons (DeGorf, 2001) Mass-Weber filter bands and multiple layer perceptrons (Barschoff, 1995) wavelets and multiple layer perceptrons (Reed, 2001) wavelets and grow and Learn networks (Olmez, 2002) among others have been used in this context.

ANN have proven to be powerful tools in signal classification as well as modelling numerical data. The main advantage of ANNs is their learning capability. Instead of applying time-consuming methods for analytical description of a problem, neural networks are able to represent even non-linear relationships with good accuracy by learning from examples.

The basic neural network model includes a large number of highly interconnected units, whose coupling weights can be modified by a learning process. The response of the units is a non-linear function of the combined input stimulation (by analogy with the firing threshold of neurons). Training a network involves supplying example patterns to the input units together with the desired output patterns. A learning algorithm is used to modify the connection weights in a direction that forces the model to give a closer approximation to the desired output.

The simplest ANN is the perceptron, for which the input units are connected directly to the output units. Each output unit computes a single output as a function of one or more inputs (fig 3.2.a). There is only one layer performing any

computation: the output layer, as the input layer is only used to store the input values to the system. Although this single layer can perform linear classification, this topology performs well only on very limited scenarios. To overcome such limitations, present day systems include more layers containing hidden units. This multi-layer perceptron (MLP) is capable of learning arbitrarily complex decision boundaries between different classes (fig 3.2. b).

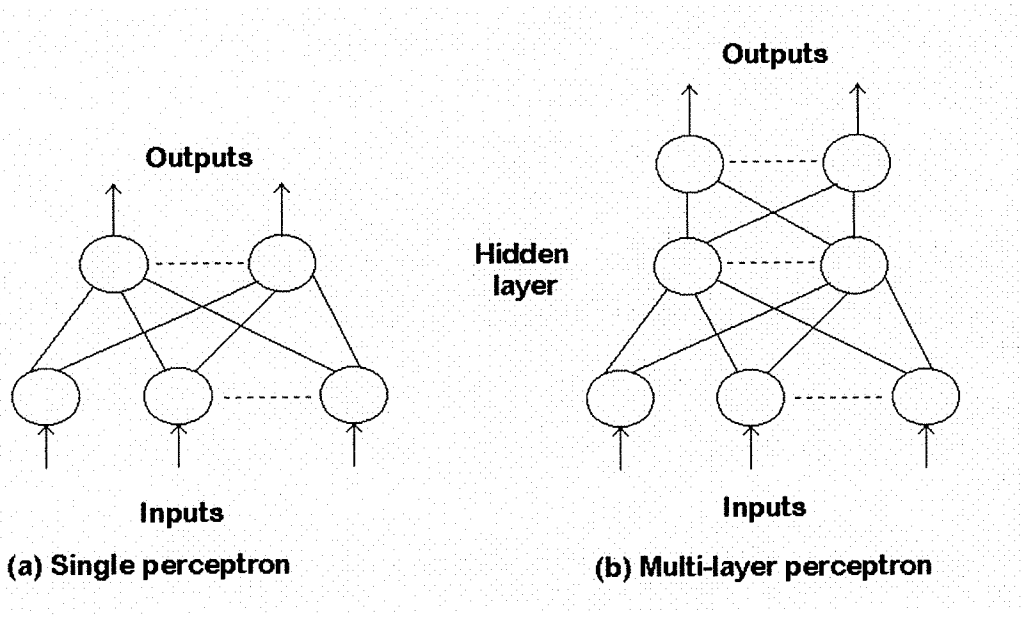


Figure 3.2 Neural Networks.

There are straightforward training procedures for this MLP. An error function (such as the mean-squared error) is defined, and this function is differentiated with respect to each of the network weights. These derivatives can then be used to find values of the weights that minimise the error function. By propagating the errors back through the network, it is possible to optimise these weights (error back propagation).

The main drawbacks of neural networks are their inability to capture temporal properties of time varying signals, their difficulty in dealing with signals of different length, and the fact that the modelling is hidden in the network and therefore it provides no insight into the classification process.

Time domain modelling: In the heart sound context before the data can be presented to the classification algorithm, individual heart cycles should be initially extracted to specify the necessary boundaries. Each cycle is then framed at regular intervals to set sub-unit boundaries and feature vectors are obtained for such frames. This approach is also used for the majority of the classification approaches. Multilayer perceptrons are feed forward networks and are unable to capture temporal properties of time varying signals directly because connections are all in one direction from input to output. There are no implicit relations established among the time sequence of input vectors, only unidirectional channels that are assumed to compute in parallel connecting the input units (features) to the outputs (classes).

Time scale variations: In order to train a connectionist network, it is necessary to specify the correspondence between the input feature vectors, from a signal pre-segmented to identify unit boundaries, and the output classes to be recognised. In the heart sound analysis context these boundaries are not usually precisely known in advance.

In some medical conditions the occurrence of murmurs is not time-fixed in the heart cycle for a specific heart condition. For example, in the case of pulmonic stenosis, during inspiration the pulmonic murmur occurs earlier in the cardiac cycle than it does during expiration. (Shaver, 1985). In aortic valvular ejection sounds the ability of the deformed valve to move plays an important role in the sound production. In patients with severe calcific fixation of the valve, when no excursion of the valve is possible, valvular ejection sounds are not recorded (Shaver, 1985). When valvular excursion is possible, the amplitude of the ejection sound correlates with valve mobility. In mobile nonstenotic bicuspid valves, the ejection sound is not only loud but is widely separated from S1, both being a function of a mobile valve with a prolonged excursion.

The use of constant length input vectors for the neural networks obtained through a linear time-length normalisation (to account for heart rate variability) increases the mismatch between sound events and linear spaced input features presented to

the network. These methods also rely on a precise segmentation algorithm and some problems may arise when signals of different heart rate have to be compared.

These time correspondence problems (segmentation and time occurrence of events within a cycle) are more important when neural networks are combined with discrete wavelets representations for their lack of time-shift invariant properties. In a recent study (Reed, 2001) a neural network was trained using ten shifted versions (over a range of one hundred time samples) of a single heart beat cycle from each type in an attempt to provide a degree of shift invariance. Such shifts are attributed in their research to variations in the heartbeat starting time found in the segmentation process. In this study the effect of noise levels in the classification process was also investigated, and although the study reveals a 100% accuracy for a SNR above 31 dB, it is worth to consider that the sample set consists of only one patient per heart condition.

Barschdorff (1995) proposes a model with no time normalisation of the input vectors for the systolic cycles, leaving the network to determine boundaries of the cycle as part of the learning process. An artificial neural network was trained using normalised power spectral densities as input features in a study to compare the performance of neural and statistical classifiers. Using this approach, the network could cope with time variations of up to 20% the length of a systolic cycle. Nevertheless, connectionist models do not cope with time scale variations very easily (Holmes, 2001). This can be partly overcome by the use of posterior neural network models, which apart from using feed forward connections, may also allow the use of recurrence by means of feed back or a set of previous activation values as part of the input.

Hidden Structure. Another practical problem in developing ANN systems is that, since all the modelling is hidden within the network, it tends to be very difficult to understand what is actually happening in that network and thus, to gain insights into how a signal recognition is performed (Holmes, 2001, p 216). As described in the biomedical background section, there are more than 40 different theories about

the heart sound genesis. A classification system that could provide some insights on the structure of the classification task may provide some clues for understanding the generation process.

3.5 Proposed technique: use of HMM for PCG classification

Several techniques have been proposed for the classification of phonocardiograms with different amounts of success. Notwithstanding, the importance of the application and the complexity of the problem demand the use of new approaches.

A technique able to account for the high variability seen in biomedical signals and with the capacity to learn, infer, and generalise from an incomplete (but representative) data set is needed. Preferably, a technique reflecting the dynamics of the system efficiently, dealing with time scale variations and with a clear structure that may provide insight in the physical phenomena.

Statistical methods seem suitable since they only require a general structure whose parameters are trained automatically using a large amount of training data. In an autonomous system the structured succession of events in time of a PCG signal could be modelled as a sequence of events with a specific pattern by means of a statistical model. Such models are a good mechanism to represent the dependencies of each sound component on the previous and posterior, providing time modelling. Furthermore, statistical distributions are a reasonable way to formally represent the variability observed in real heart sound samples.

Hidden Markov modelling is a powerful technique for modelling the temporal structure and variability of a signal, and to generalise a model through statistical distributions to include unseen data. This technique is based on a probabilistic pattern matching approach, which models a sequence as the output of a random process.

One of the main advantages of these models is their ability to deal with time scale variability (Holmes, 2001); HMMs are popular in speech because the model can cope with changes in the length of sounds. This ability to deal with time scale variability is particularly important in the case of PCGs where the systolic and diastolic cycle lengths present both inter-subject and intra-subject variability. Moreover, as described before, particular events in the cardiac cycle may exhibit relative variations within a cycle for different stages of the pathology.

The HMM methodology provides a tractable mathematical framework with straightforward algorithms for recognition and for training to create models of time sequences. The features used to represent a signal and the temporal sequence relationships are treated separately but within a consistent framework. As a consequence segmentation of particular events within a signal is possible based on their own statistical characteristics. This methodology can also be used for the segmentation of phonocardiogram signals (Romero, 2002).

The use of mathematical models to represent the signal may provide a better insight into the physical phenomena involved. The parameters of the models can be proposed based on physical phenomena or can be used as a tool for the discovery of structure in a time series. In the latter case, the significance of the states of the model is based on the parameters estimated by the model. By revealing structures in the feature representations of PCG signals, hidden Markovian models may contribute to our understanding of the genesis of the heart sounds by relating these signal defined states with the physical events. The idea of replacing priori labelling of the states of the model by a probabilistic labelling has been used to model text (Cave, 1980), phonetics (Neuburg, 1971) and speech (Poritz, 1982).

Hidden Markov Models have been successfully applied to speech signals (Gold, 2000), DNA sequences (Mount, 2004), and ECG signals for automatic pattern recognition problems (Throval, 1994). In these applications, the HMM is used as a powerful model to characterise temporal nonstationary, spatially variable, but learnable and regular patterns of the signal. One particular aspect of this model is

its sequentially arranged Markov states that allow the use of piecewise stationarity for approximating the nonstationary properties of a signal. Hidden Markovian models are supported by a strong mathematical background, and their dynamic and time variability modelling properties make them a good candidate technique for PCG classification.

3.6 Summary

The problem of PCG signal identification was discussed from the perspective of signal classification techniques. The main steps involved in the classification process were briefly explained and the principal general techniques applied to heart sound analysis were described along with some relevant examples. Particularly, the advantages and disadvantages of the use of neural networks for classification of PCG signals were evaluated. Finally after a brief discussion, hidden Markovian models were suggested as an alternative technique that may improve the classification of phonocardiograms.

The following chapter explains the theory of the hidden Markovian models.

4 Hidden Markovian Models

4.1 Overview

This section presents a brief introduction about the classification method to be used: the Hidden Markovian Models. The theory of the HMM is presented at a conceptual level by means of an example.

The problem selected is a real phenomena that can be represented by hidden Markovian models. Although it is not related with heart sound analysis, it is useful to clearly show the relation between the states of the model representing the generation mechanism and the consequences it has on the observations. In a HMM the physical significance of the states are not required to be know for it is a mathematical model, however, in practice the characteristics of the HMM are preferably chosen according to some insight regarding the generation mechanism.

In the example presented, the states are not completely hidden since they can be measured, but in practical terms they are not easily monitored in a representative manner and their complex relation to the observations make them suitable to be treated as a hidden generation mechanism.

Using this example, the relationship between the components of a HMM can be easily conceptualised and the variations between models can be illustrated intuitively. The concepts acquired through the development of this chapter are sufficient to understand the details of subsequent chapters, although the treatment of the topic is not rigorous.

A more in depth explanation of the concepts and techniques can be found in Rabiner (1989) and Jelinek (1999).

4.2 Example: Paralytic Shellfish Poisoning

Consider the example of the relationship between toxic algae blooms and the poisonous toxins in shellfish that humans consume.

Paralytic Shellfish Poisoning (PSP) is a life-threatening syndrome caused predominantly by the consumption of contaminated shellfish, with both gastrointestinal and neurologic symptoms. Cases of PSP have been reported worldwide.

Gymnodinium catenatum is a single cell alga species (dinoflagellate) that is the source of PSP marine toxins that can affect humans, birds, and fish. Gymnodinium dinoflagellates move up and down the oceanic water column, being closer to the surface in daytime and deeper during the night. Its life cycle, growing rate, and toxicity depend on several environmental factors such as temperature, salinity, and variations in nutrimental sea components. These unicellular dinoflagellates develop into an algae blooms throughout the world for unknown reasons, although a variety of factors have been studied, including change in weather, upwellings, temperature, turbulence, salinity, and transparency.

Paralytic Shellfish Poisoning is caused by a group of chemicals called the saxitoxins and gonyautoxins. These chemicals are produced by the algae and released into the shellfish when the algae are eaten. The chemicals all differ in their toxicity to humans, and change depending on the species of shellfish. After the consumption of raw or cooked contaminated shellfish (the toxins are heat stable so cooking the shellfish will not remove them) stomach acids in animals and humans also alter the chemical's toxicity.

Ingestion of molluscs contaminated results in the following clinical scenario (Kao, 1993): From five to thirty minutes after consumption, there is slight perioral tingling progressing to numbness, which spreads to face and neck when present in moderate cases. In severe cases, these symptoms spread to the extremities causing lose of co-ordination and respiratory difficulty. There are medullary disturbances

in severe cases evidenced by difficulty swallowing, sense of throat constriction, speech incoherence or complete loss of speech, as well as brain stem dysfunction. Approximately 75% of severely affected people die within two to twelve hours by complete paralysis and death from respiratory failure in the absence of ventilatory support. After twelve hours, assisted victims start to recover gradually and have no residual symptoms within a few days (Halstead 1984).

Large-scale proactive monitoring programs (assessing toxin levels in mussels, oysters, scallops, clams, etc.) prevent PSP and rapid closures to harvest of suspect or toxic areas is a common practice throughout the world.

Consider the problem of modelling the possible occurrence of PSP events (due to algae blooming) from the levels of toxins measured in shellfish.

The species that appears to take the highest amounts of toxins are green-lipped mussels although other species have been found with toxins that are in the range where illness is possible. These include tuatua, toheroa, oysters, scallops, and cockles.

In order to provide practical guidance to the public, food safety authorities can set various ranges according to the levels of toxins measured in green-lipped mussels. For instance, a safe range can be defined for levels below 30 μg/100 g saxitoxin equivalent. A caution range could be defined from 31μg/100 g to 80 μg/100 g. At this level it is expected that the toxins will not affect most people since the upper limit in this action level is around ten times lower than the lowest level associated with outbreaks. Close monitoring should be advised when toxin levels fall within this range. A warning range defined from 81 μg/100 g to 800 μg/100 g would lead to closure of commercial farms. A danger range is defined from 801 μg upwards and relates to outbreaks of PSP in humans. Previous events have reported levels as high as 4000 μg/100g.

Although the generation mechanism is the algae blooming, to obtain a close correlation between the number of *Gymnodinium catenatum* cells and shell fish

toxicity truly representative water sampling is required. On exposed coasts this is difficult to achieve. Most of the samples are collected from the shore and within the surf zone, and are therefore not truly representative of cell abundance in the adjacent water column. One approach is to model the algae process using a hidden generation mechanism. Therefore, in this example, the shellfish toxin level samples constitute the observations and an underlying hidden process, the algae toxicity, generates these observations.

If we consider for example three levels of toxicity in the algae, namely a low, medium, and high level, we can relate these algae toxin levels with the levels of shellfish toxins and therefore, to the warnings issued by the food safety authorities about the risks of PSP.

The Markov assumption presumes that the present level of algae toxicity can always be predicted solely given knowledge of the past levels; other environmental factors are not considered. In this example, and many others, such assumptions are obviously unrealistic. Nevertheless, since such simplified systems can be subjected to analysis, we often accept the assumption in the knowledge that it may generate information that is not fully accurate.

In a first order system, the actual state depends only on the previous state, and therefore, in a first order Markov model for the algae levels the current level of algae toxicity depends only on the immediate previous level.

Figure 4.1 shows all first order state transitions for the underlying Markov model of this example. Three levels of algae toxicity are defined and arrows represent the transition between them. Notice that transitions to the same state are possible. This characteristic gives the HMM the flexibility to deal with time scale variations since long lasting event can be modelled through a sequence of self transitions.

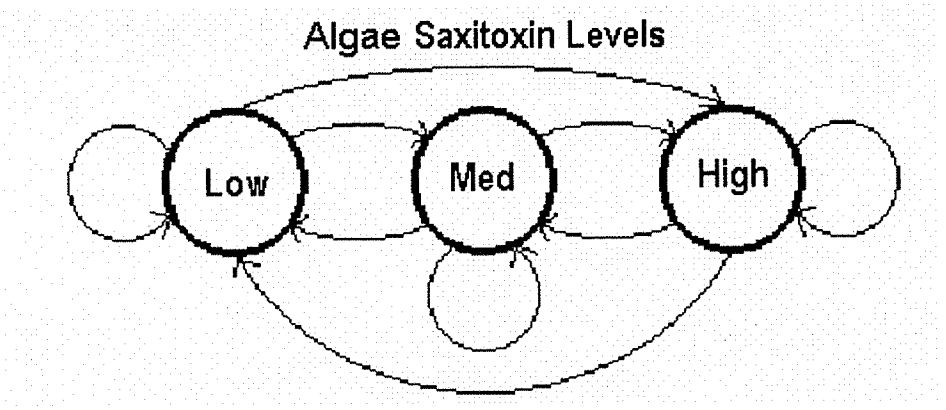


Figure 4.1. Underlying Markov Model for the Algae Toxicity.

The probability of the current level of toxicity (current state) depending on the previous level (previous state) can be expressed in a transition matrix, called the state transition matrix. Figure 4.2 shows a set of possible transition probabilities for the changes in toxicity levels in the alga *Gymonodinium catenatum*.

		Current Algae Saxitoxins Level		
		Low	Medium	High
Previous Algae Saxitoxins Level	Low	0.42	0.39	0.19
	Medium	0.12	0.43	0.45
	High	0.09	0.40	0.51

Figure 4.2. State Transition Matrix.

For the model represented by this set of probabilities, given that the previous sample of shell fish toxins level was low then there is a slightly higher probability that the levels remain low than there is that the levels to increase to a medium level. There is also a low probability it jumps straight to a high status from a low level of algae toxicity. In this way the matrix reflects the dynamics the generation process for a specific model.

Notice that since the numbers are probabilities the sum of the entries for each row is one; i.e. given a previous algae state, for example a low state, the current state can only be low again, medium or high.

To initialise such a system, we need to state what the level was (or probably was) on the first sample; we define this in a vector of initial probabilities. An example of an initial vector is shown in Figure 4.3, it corresponds to the case where the first sample is known to have had a low level of algae saxitoxins.

Until now, the model presents predetermined probability values to relate the states in the underlying process. However, given a suitable algorithm, the real values for the hidden Markov process have to be inferred from the sample sequences of observations. The observations are related to the states by a further set of probabilities, contained in the confusion matrix.

Low	Medium	High
(1.0	0	0)

Figure 4.3. Vector of Initial Probabilities.

The confusion matrix in Figure 4.4 and further ilustrated in Figure 4.5, represents the connections between the hidden states and the observable states for this example.

		Shellfish Poisoning Alert			
		Safe	Caution	Warning	Danger
Alga Saxitoxins Level	Low	0.58	0. 27	0.10	0.05
	Medium	0.17	0.34	0.33	0.16
	High	0.03	0.07	0. 28	0.62

Figure 4.4. Confusion Matrix between States and Observations.

The confusion matrix describes the probability of having a particular range of shellfish toxins and therefore, being in a particular state of alert, given a level of algae saxitoxins. In the example matrix there is a high probability that it is safe to consume shellfish for a low level of algae saxitoxins

There is also a higher probability that a danger alert level (a high level of toxins measured in the shellfish) can be associated with high levels of toxins present in Gymonodinium algae. Nevertheless, since the toxicity present in the shellfish is influenced by other factors besides the toxicity levels in the algae, there is also probability relating lower toxic levels of algae with high levels of toxicity in the shellfish population.

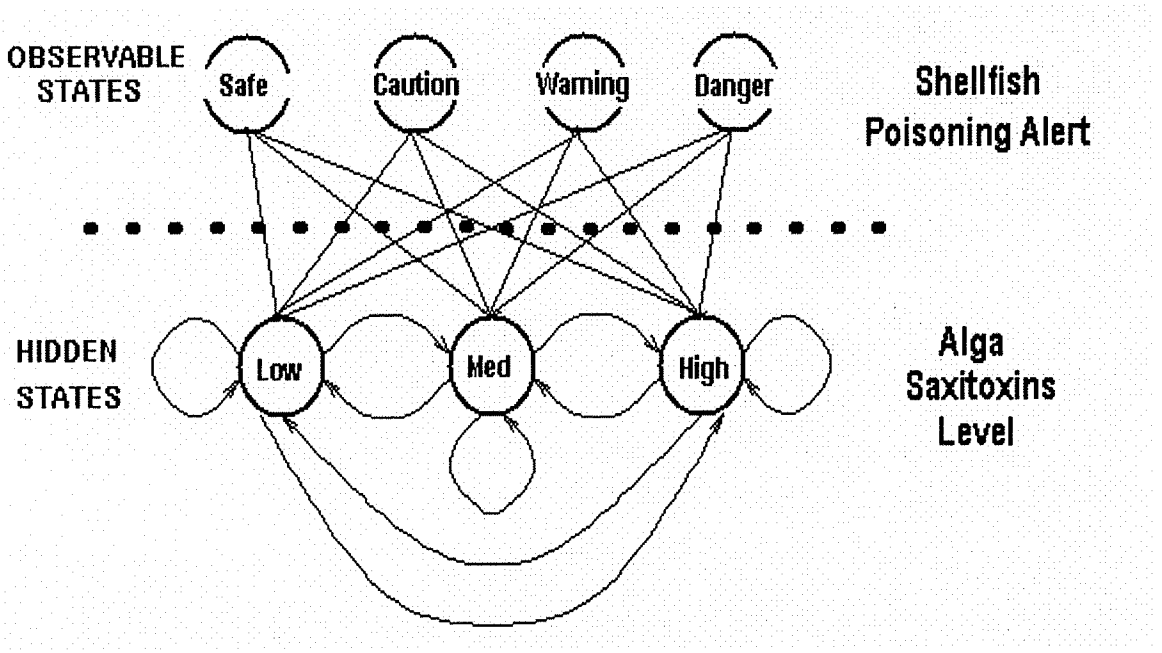


Figure 4.5. Relation between Hidden States and Observations.

4.3 Uses associated with HMMs

Learning -The sets of probability values in the matrices characterising this model are learned form sets of observations reflecting specific phenomena. In the case of PSP, this may be for example a model for summer harmful algae blooming events in a specific geographical location with observation sets obtained for events

occurring in different years. Using the same structural model of three levels of algae toxicity and four alert levels, other probability values are expected to be obtained for models of seasons without harmful algae proliferation. Similar models with characteristic probability values can be obtained for different seasons, and or geographical locations.

Evaluation- To evaluate the risk of PSP, different HMM models can be used to characterise the relationship between the alga toxicity and the associated shellfish poisoning risk in relation to different seasonal ecosystem conditions. A set of observations of the levels of toxicity in the shellfish could then be used to identify models during initial stages of the phenomena and therefore to forecast the harmful algae blooming.

Decoding – As mentioned before, the acquisition of truly representative samples of cell abundance by water sampling in exposed coasts is not practical in most cases. This is not easily attainable since it implies identifying and counting micro organisms in discrete samples that harbour a great diversity of species where a rapid analysis of a large number of samples would be routinely desired. Nevertheless, through the models obtained it is possible to estimate the most probable sequence of algae toxicity level states from a specific set of shellfish toxicity observations, i.e. to find the hidden states that generated the observed output. Analysis of the models allows the interaction of both sets of variables to be examined.

4.4 Variations of the hidden Markovian model

The values obtained for the levels of saxitoxins measured in the shellfish are continuous, but for practical guidance the food safety authorities establish an alert system comprised of four discrete levels as described before. This set of discrete levels provides immediate guidance about the risk of shellfish consumption by contrast to providing continuous values that require interpretation. Nevertheless, from a biological perspective, continuous values may be more appropriate to track the fluctuations that characterise the process.

In this particular hidden Markov model, continuous values of shellfish saxitoxin were converted to discrete values by setting labelled ranges. The observations were characterised as discrete symbols chosen from a finite alphabet: the four alert levels, and therefore a discrete probability could be used to relate the states to the observations of this model. Such a model is called a discrete hidden Markov model.

In order to model the process using direct shellfish saxitoxin values, the relation between the states and the continuous observations can be defined by a probabilistic distribution. Assuming a parametric distribution, such as a Gaussian distribution, the necessary values can be estimated from the data by a recursive process. Many natural processes involve variable quantities that approximate reasonably well to the normal or Gaussian distribution. A weighted sum, or mixture, of Gaussian distributions is generally used to model other distributions. Provided that there is a sufficient number of Gaussian mixture components, any distribution shape can be approximated closely. The models generated are called single Gaussian hidden Markovian models and Gaussian mixture hidden Markovian models (GMHMM) respectively.

Neural networks can also be used to compute the probability for any given correspondence between states and observations generating a hybrid hidden markov neural network model (HMM/ANN model).

4.5 Solutions to the HMM Problems *(Rabiner, 1989)*

Once a system can be described as a HMM, three problems can be solved. The first two are pattern recognition problems: Finding the probability of an observed sequence given a HMM (evaluation) and finding the sequence of hidden states that most probably generated an observed sequence (decoding). The third problem is generating a HMM given a sequence of observations (learning).

4.5.1 Definitions

The example in section 4.2 shows how a Hidden Markovian Model is characterized by two model parameters (the number of states and the number of distinct observation symbols), specification of observation symbols, and the specification of three probability measures. Lets formally define these symbols.

Lets $S= \{S_1,S_2,..,S_N\}$ be the finite set of states of the model and $V= \{V_1,V_2,..,V_M\}$ the finite set representing the individual (observable) symbols. Aditionally, let q_t represent the state at time t.

A corresponds to the state transition probability matrix with entries

$$a_{ij} = P(q_{t+1} = S_j|q_t = S_i), \qquad \text{i}$$

B denotes the observation symbol probability distribution vector $B=\{b_j(k)\}$, where

$$b_j(k) = P(V_k \text{ at } t|q_t = S_j). \qquad \text{ii}$$

Π represents the initial state distribution vector $\Pi=\{\Pi_i\}$ where

$$\Pi_i = P[q_1 = S_i]. \qquad \text{iii}$$

Given appropiate values of N, M, A, B, and Π, the HMM can be used as a generator to give an observation sequence

$$O = \{o_1, o_2,\ldots,o_T\}, \qquad \text{iv}$$

where each observation O_t is one of the symbols from V, and T is the number of observations in the sequence, and

$$\lambda=(A, B, \Pi), \qquad \text{v}$$

is used to indicate the complete parameter set of the model.

The application of this model to real signals generally involves the solution of some or all of three basic problems: Evaluation, Decoding and Learning.

4.5.2 Evaluation: Finding the probability of an observed sequence given a HMM

The aim is to compute the probability of occurrence of a particular observation sequence, $O = \{o_1,\ldots,o_T\}$, given the model λ, i.e., $P(O|\lambda)$.

The most straightforward procedure to solve this problem is to enumerate every possible state of the T-length sequence. Consider the following state sequence

$Q = \{q_1,q_2,\ldots q_T\}$,

vi

where q1 is the initial state. Assuming that observations are independent, given the hidden states, the probability of the observation sequence O for the state sequence Q is

$$P(O \mid Q,\lambda) = \prod_{t=1}^{T} P(o_t \mid q_t,\lambda) = b_{q1}(o_1)b_{q2}(o_2)...b_{qT}(o_T)$$

vii

The probability of such state sequence Q is given by

$$P(Q \mid \lambda) = \pi_{q1}a_{q1q2}a_{q2q3}...a_{qT-1qT}$$

viii

The probability that O and Q occur simultaneously is the product of the above terms

$$P(O,Q \mid \lambda) = P(O \mid Q,\lambda)P(Q \mid \lambda)$$

ix

The probability of O given the model λ is obtained by adding this joint probability over all possible state sequences q is giving by

$$P(O \mid \lambda) = \sum_{q}^{N} P(O \mid q, \lambda) P(q \mid \lambda) \qquad \text{x}$$

The calculation of $P(O|\lambda)$ in this direct form involves $2TN^T$ arithmetic procedures, which make it computationally unfeasible. However, alternative and more efficient procedures (that requires only N^2T elemental mathematical operations) are available, and they are explained below:

4.5.2.1 Forward procedure

The problem of computing the probability of a particular observation sequence given a HMM can be solved by dividing the problem into computing the probability of partial observation sequences.

Consider the forward variable $\alpha_t(i)$ defined as

$$\alpha_t(i) = P(o_1 ... o_t, q_t = S_i \mid \lambda) \qquad \text{xi}$$

This is the probability of the partial sequence until a given time t and state S_i at the same time t, for a given model λ. A recursive algorithm can be used to obtain the probability of the whole sequence P(O, λ). This procedure consists of three stages:

a) Initialise

In this step the probability for the first observation (a partial sequence with one element) is computed using the set of initial probabilities of the model.

$$\alpha_1(i) = \pi_i b_i(o_1) \qquad 1 \le i \le N \qquad \text{xii}$$

b) Calculate

$$\alpha_{t+1}(j) = [\sum_{i=1}^{N} \alpha_t(i) a_{ij}] b_j(o_{t+1}) \qquad 1 \le t \le T-1, \quad 1 \le j \le N$$

xiii

Thus, the probability for the next partial observation sequence will be computed considering the values obtained in the previous one. This process continues (forward) until the last observation is reached.

c) Obtain

$$P(O \mid \lambda) = \sum_{i=1}^{N} \alpha_T(i)$$

xiv

The probability for the whole sequence is the result of adding the probabilities of all the partial observation sequences.

4.5.2.2 Backward procedure

Once again, the problem of computing the probability of a particular observation sequence given a HMM is solved by dividing the problem into computing the probability of partial observation sequences. But in this case starting from the last observation and going backwards.

In a similar manner a backward variable can be defined as

$$\beta_t(i) = P(o_{t+1} o_{t+2}, \ldots o_T \mid q_t = Si, \lambda)$$

xv

This is the probability of the partial observation sequence from t+1 to the end, given state S_i at time t and the model λ. Solving for $\beta_T(i)$

a) Initialise

$$\beta_T(i) = 1 \qquad 1 \le i \le N$$

xvi

Starting from the last observation

b) Calculate

$$\beta_t(i) = \sum_{j=1}^{N} \beta_{t+1}(j) a_{ij} b_j(o_{t+1}) \qquad t = T-1, T-2, ...,1, 1 \le i \le N$$

xvii

c) Obtain

$$p(O \mid \lambda) = \sum_{i=1}^{N} \beta_1(i) \qquad t = T-1,...,1$$

xviii

Adding all the partial sequence probabilities, the probability for the whole sequence is obtained. Again, the computation requires only N^2T calculations making this procedure more computationally feasible than the procedure described first enumerating all possibilities.

4.5.3 Decoding: Finding the sequence of hidden states that most probably generated an observed sequence

The aim is to find the single most likely state sequence, Q = {q1,q2,...q_T}, for the given observation sequence O = {o1,o2,...o_T}. This can be computed using the Viterbi algorithm (Viterbi, 1967). Defining the auxiliary variable $\delta_t(i)$

$$\delta_t(i) = \max_q P(q_1, q_2, ..., q_t = i, o_1, o_2, ...o_t \mid \lambda)$$

xix

which is the highest probability along a single path, at time t, accounting for the first t observations and ends in state S_i, therefore

$$\delta_{t+1}(j) = \max_i(\delta_t(i)a_{ij})b_j(o_{t+1}) \qquad \text{xx}$$

To retrieve the state sequence, it is necessary to keep track of the argument that maximised the previous equation for each t and j. This can be done using the array $\psi_t(j)$. The procedure can be stated as follows

a) Initialise

$$\delta_1(i) = \pi_i b_i(o_1) \qquad 1 \le i \le N \qquad \text{xxi}$$

$$\psi_1(i) = 0 \qquad \text{xxii}$$

b) Compute

$$\delta_t(j) = \max_{1\le i\le N}(\delta_{t-1}(i)a_{ij})b_j(o_t) \qquad 2 \le t \le T, 1 \le j \le N \qquad \text{xxiii}$$

$$\psi_t(j) = \arg\max_{1\le i\le N}(\delta_{t-1}(i)a_{ij}) \qquad 2 \le t \le T, 1 \le j \le N \qquad \text{xxiv}$$

c) Obtain

$$P^* = \max_{1\le i\le N}\delta_T(i) \qquad \text{xxv}$$

$$q_T^* = \arg\max_{1\le i\le N}\delta_T(i) \qquad \text{xxvi}$$

d) State sequence backtracking

$$q_t^* = \psi_{t+1}(q_{t+1}^*) \qquad t+T-1, T-2, \ldots, 1 \qquad \text{xxvii}$$

Therefore, backtracking the pointer q^*_t the single most likely state sequence can be retrieved. The Viterbi algorithm is similar (except for the backtracking) in implementation to the forward calculation previously described. The mayor

difference is the maximization (xxiii) over previous states instead of the summation (xiii).

4.5.4 Learning: Generating a HMM given a sequence of observations

The aim is to find the model parameters $\lambda = (A,B,\Pi)$ such that $P(O|\lambda)$ is locally maximised. This can be done through an iterative procedure

a) Initialise λ_0
b) Compute new model λ using λ_0 and the observations O
c) $\lambda_0 \leftarrow \lambda$
d) Repeat steps b and c until

$$\log P(O \mid \lambda) - \log P(O \mid \lambda_0) < d \qquad \text{xxviii}$$

Using the Baum-Welch procedure (Rabiner, 1989) to compute the new model parameters, we define $\xi_t(i,j)$, the probability of being in state S_i at time t and at state S_j at time t+1, given the model λ and the observation sequence O as

$$\xi(i,j) = \frac{\alpha_t(i)a_{ij}b_j(o_{t+1})\beta_{t+1}(j)}{P(O \mid \lambda)} \qquad \text{xxix}$$

$$= \frac{\alpha_t(i)a_{ij}b_j(o_{t+1})\beta_{t+1}(j)}{\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_t(i)a_{ij}b_j(o_{t+1})\beta_{t+1}(j)}$$

Let $\gamma_t(i)$ be the probability of being in state S_i at time t, given the observation sequence O and the model λ

$$\gamma_t(i) = \sum_{j=1}^{N}\xi_t(i,j) \qquad \text{xxx}$$

Summing $\gamma_t(i)$ over the time index t the number of transitions made from state Si is obtained. Similarly summation of $\xi_t(i,j)$ over t can be interpreted as the expected number of transitions from state S_i to state S_j. That is

$\sum_{t=1}^{T-1} \gamma_t(i)$ = expected number of transitions from S_i

xxxi

$\sum_{t=1}^{T-1} \xi_t(i)$ = expected number of transitions from S_i to S_j

xxxii

Consequently, the following formulas provide the new values for Π, A, and B

$$\hat{\pi} = \gamma_1(i)$$

xxxiii

$$\hat{a}_{ij} = \frac{\sum \xi_t(i,j)}{\sum \gamma_t(i)}$$

xxxiv

$$\hat{b}_j(k) = \frac{\sum_{t,o_t=k} \gamma_t(j)}{\sum \gamma_t(j)}$$

xxxv

If we define the current model λ_0 and use it along with the observations to compute the auxiliary variables $\gamma_t(i)$ and $\xi_t(i,j)$, then a new model λ can be obtained using xxxiii, xxxiv, and xxxv. Baum (1970) proved that if the new model parameters are different to the old ones, the observation sequence is more likely to have been produced by the new model.

4.6 Continuous observation densities in HMM

In order to use continuous values for the observations, some restrictions have to be imposed on the form of the probability density function (pdf) to guarantee that the parameters can be computed in a consistent way. The most general representation is a mixture of the form

$$b_j(O) = \sum_{m=1}^{M} c_{jm}\eta[O, \mu_{jm}, U_{jm}] \qquad 1 \le j \le N$$

xxxvi

Where c_{jm} is the mixture coefficient for the m^{th} mixture in state j and η is any log-concave or elliptically symmetric density (e.g., Gaussian), with mean vector μ_{jm} and covariance matrix U_{jm} for the m^{th} mixture component in state S_j. Usually, a Gaussian density is used. The mixture gains c_{jm} satisfy

$$\sum_{m=1}^{M} c_{jm} = 1 \qquad 1 \le j \le N$$

xxxvii

$$c_{jm} \ge 0 \qquad 1 \le j \le N \quad 1 \le m \le M$$

and thus, the pdf is normalised i.e.,

$$\int_{-\infty}^{\infty} b_j(x)dx = 1 \qquad 1 \le j \le N$$

xxxviii

The re-estimation formulas for the coefficients of the mixture density are

$$\overline{c}_{jk} = \frac{\sum_{t=1}^{T} \gamma_t(j,k)}{\sum_{t=1}^{T}\sum_{k=1}^{M} \gamma_t(j,k)}$$

xxxix

$$\overline{\mu}_{jk} = \frac{\sum_{t=1}^{T} \gamma_t(j,k) \cdot O_t}{\sum_{t=1}^{T} \gamma_t(j,k)}$$

xl

$$\overline{U}_{jk} = \frac{\sum_{t=1}^{T} \gamma_t(j,k) \cdot (O_t - \mu_{jk})(O_t - \mu_{jk})'}{\sum_{t=1}^{T} \gamma_t(j,k)}$$

xli

where $\gamma_t(j,k)$ is the probability of being in state j at time t with the k^{th} mixture component accounting for o_t

$$\gamma_t(j,k) = \left[\frac{\alpha_t(j)\beta_t(j)}{\sum_{j=1}^{N} \alpha_t(j)\beta_t(j)} \right] \left[\frac{c_{jk}\eta(O_t, \mu_{jk}, U_{jk})}{\sum_{m=1}^{M} c_{jm}\eta(O_t, \mu_{jk}, U_{jk})} \right]$$

xlii

The re-estimation formula for a_{ij} is identical to that used for discrete observation densities.

$$\hat{a}_{ij} = \frac{\sum \xi_t(i,j)}{\sum \gamma_t(i)}$$

xliii

For more details on the algorithms refer to the tutorial paper presented by Rabiner (1989).

4.7 Multiple Observation Sequences

To use multiple observation sequences mutual independence is assumed. The set of k observation sequences is defined as

$$O = [\, O^1, O^2, \ldots, O^k]$$

xliv

where $Ok = [\, O_1^k, O_2^k, \ldots, O_{Tk}^k]$ is the kth observation sequence. The parameters of the HMM needs to be adjusted to maximize

$$P(O \mid \lambda) = \prod_{k=1}^{k} P(O^k \mid \lambda) = \prod_{k=1}^{k} P_k$$

xlv

The formulas for the estimation of the new parameters are

$$\hat{a}_{ij} = \frac{\sum_{k=1}^{k} \frac{1}{P_k} \sum_{t=1}^{T_k-1} \alpha_t^k(i) a_{ij} b_j(O^k{}_{t+1}) \beta^k{}_{t+1}(j)}{\sum_{ik=1}^{k} \frac{1}{P_k} \sum_{t=1}^{T_k-1} \alpha^k{}_t(i) \beta^k{}_t(i)}$$

xlvi

$$\hat{b}_j(\ell) = \frac{\sum_{k=1}^{k} \frac{1}{P_k} \sum_{t=1}^{T_k-1} \alpha_t^k(i) \beta^k{}_t(j)}{\sum_{ik=1}^{k} \frac{1}{P_k} \sum_{t=1}^{T_k-1} \alpha^k{}_t(i) \beta^k{}_t(i)} \quad s.t. O_t = V_t$$

xlvii

Using the scaling factor

$$c_t = \frac{1}{\sum_{i=1}^{N} \alpha_t(i)}$$

xlvii

$$\hat{\alpha}_t(i) = \frac{\alpha_t(i)}{\sum_{i=1}^{N} \alpha_t(i)}$$

xlix

$$\hat{\beta}_t(i) = c_t \beta_t(i)$$

xlx

4.8 Summary

The statistical model assumes an underlying process (hidden) generating a set of discrete or continuous observations. The hidden process is a Markovian chain

representing the different possible states. At regularly spaced times the system undergoes a change of state (possibly returning to the same state) according to a set of probabilities associated with the current one. The model assumes that the actual state depends only on the previous ones (Markov assumption), and the probabilities associated with the states do not change with time. The observations are related to the states by a discrete probability function for discrete observations, a parametric probability function for continuous observations, or using a neural network (Rabiner, 1989).

The next chapter introduces the use of cepstral analysis in the context of phonocardiogram representation.

5 PCG Feature Extraction

5.1 Introduction

Hidden Markovian models have been proposed as an advantageous technique for the classification of PCG signals; however, the performance of a classifier also depends highly on the features selected to represent the heart sounds.

In order to perform the automatic analysis of a PCG signal, a simplified representation is necessary to ease the computation required. An overly detailed representation may also take into account characteristics of the particular training set that will not be present for independent data sets. For signal classification problems, the features selected aim to represent the signal in a manner that minimises variability within a class and retains class differences (Gold, 2000).

The PCG time domain signal is usually transformed to other domains by applying specific operations to it. Although the main purpose may be data reduction, transformation of the signal to other domains may be required to emphasise some signal patterns, due to the statistical properties of the new variables, or because of computational advantages. In some specific domains the transformation is a one to one mapping and therefore, an inverse transformation exists and can be used to recover the original time signal. When this is not the case, and the raw signal is not recoverable any longer, the transformation is still useful as long as the pattern in the new domain is still descriptive of the main signal characteristics.

The most important features of the PCG signal are the intensity and timing sequence of the components of the heart sounds and their location, frequency content, and envelope shape of murmurs, if any (Rushmer, 1970). Based on these characteristics, several methods have been employed for the analysis of PCG, among them the use of time frequency representations (Leung, 1997; Ritola, 1996) linear prediction analysis (Iwata, 1977), cepstral analysis (Rangayyan, 1978a, 1979), and wavelets (Barschdorff, 1995; Obaidat, 1992; Zhang, 1998; Olmez, 2003).

In the literature available, two different classes of feature representations can be distinguished.

- Features developed initially for analysis of the PCG. Such features aim to find an alternative representation that may emphasise certain patterns of the signal to ease visual diagnosis. Generally, these representations have a counterpart in the characteristics sought by auscultation techniques. Examples of this technique are the Maass-Weber set of filters (Holldack, 1965) and most of the time-frequency representations. Some of these representations have been later used as input features for classification algorithms.

- Feature representations specifically aimed at classification. These representations may not have a straightforward interpretation although related to a model of the physical phenomena. Nevertheless, they are selected for their computational advantages, their data reduction capabilities, and/or due to their mathematical properties. Examples of these techniques are linear prediction and pole – zero modelling.

In this research, based on a model in accordance with the most accepted theory for the genesis of heart sounds, cepstral coefficients are proposed as the preferred feature representations for the classification models. Nevertheless, there are other approaches that also provide a simplified representation of the PCG signal and are suitable to be used as input features for the signal classifier. Consequently, two other main approaches have been selected as alternative representations and for performance comparison: an energy integration method based on the use of Maass-Weber filters and linear prediction analysis.

Maass-Weber filters were selected as a simple sub band energy representation (basically a simplified spectrogram), which nevertheless, embraces the practical approach for clinical analysis of heart sounds: The use of this technique for clinical phonocardiography is considered standard practice in countries like Germany or Mexico (Holldack, 1974; Guadalajara, 1998).

Linear prediction analysis, was selected not only due to its wide success as a feature representation technique in the signal processing literature, but also because it is the main candidate to substitute the use of cepstral representations as inputs to the hidden Markovian Models.

The reminder of this chapter explores the use of these feature extraction methods for PCG representation in the context of signal classification.

5.2 Proposed technique: cepstral representations

The most accepted theory for the genesis of heart sounds (the cardiohemic theory) regards the generation of heart sounds as a consequence of a system set into vibration by the acceleration and deceleration of blood flux. These variations of blood flux are related to the sequence of the opening and closing of the heart valves and the contraction and expansion of its chambers during a cardiac cycle.

The PCG signal can be assumed to be the convolution of a system's impulse response and an excitation component. The excitation component is generated by the blood acceleration and deceleration components of S1 and S2, the noise signal caused by flow turbulence, or both, whereas the impulse response of the system would be defined by the heart structure. Accordingly, the PCG signal would be the result of the convolution of the heart structure and the vibrations induced by blood flow regardless of its generating mechanism.

Following this model, cepstral analysis can potentially be used to separate the components of the sound representing the system's response dominated by the heart structure from the components due to the blood flux excitation.

Cepstral analysis is a signal processing technique in which a measurable signal is conceived as the result of a convolution of an excitation and a system response. The aim of the technique is to transform this convolution to a linear addition so

that the high order components related to the excitation can be separated from the low order components related to the system response. Accordingly, abnormalities in the heart structure will be reflected in the lower order coefficients, whereas the contribution of the excitation due to blood flux variations will be reflected in higher order coefficients (see appendix B).

The following section explores the feasibility of using cepstral representations for the differentiation of pathological and non pathological PCGs.

5.2.1 Methodology

Cepstral analysis is a special case within a general class of methods known as 'homomorphic' signal processing. There are two main variants of cepstral analysis, namely complex cepstum and power cepstrum. The basic difference between them is that real cepstrum discards phase information about the signal while the complex cepstrum retains it. Complex cepstral is preferred in phase sensitive applications when it is required to return to a time waveform. However, since phase processing adds complexity, power cepstrum is employed more widely for practical applications. In automatic speech recognition for instance, a filter bank designed according to some model of the auditory system is used in the computation of a power cepstrum to obtain a mel-ceptrum that compensates for the unequal sensitivity of human hearing at different frequencies (Gold, 2000).

Power cepstrum is selected for this research since a simplified representation of the phonocardiograms is required. Figure 5.1 illustrates the steps involved in computing the cepstrum of the phonocardiograms. The signal is firstly divided into frames of 20 ms overlapped by 50%. The length of the window has been selected within the range proposed by Jamous (1992) for frequency analysis of heart sounds (see appendix B). Secondly, a Hamming window and a discrete Fourier transform are applied to each frame to obtain its spectrum. Afterwards, an inverse discrete Fourier transform is applied to the logarithm of each spectrum frame in order to obtain its cepstrum. The last step involves a cepstral truncation:

the higher order coefficients (assumed to represent the excitation) are discarded and only the values near to the origin (corresponding to the resonance structure) are retained since as illustrated in appendix B this low order coefficients provide a simplified spectrum representation. This set of operations is performed for each frame until the cepstrum for the whole PCG signal is computed.

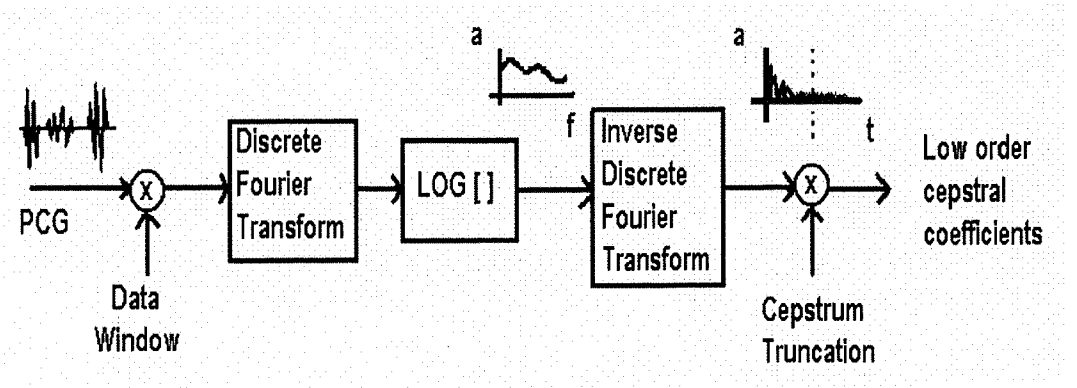


Figure 5.1. Cepstral Analysis.

Figure 5.2 shows a normal PCG signal and the evolution of the first ten cepstral coefficients. Each plot shows the evolution of one cepstral coefficient in time. The signals have been normalised to its maximum amplitude.

Considering that HMMs with more features need more training data, the limitations imposed by the size of the database lead us to set the number of input features to be small, but should still be sufficient to provide visual differentiation. Based upon the rule of thumb that the number of training samples should be five or more times the number of features used (Rangayyan, 2002, pp. 474), five coefficients were initially proposed to represent the signals. By inspection of the plots representing the samples in our database it was confirmed that these coefficients were able to provide visual differentiation. Later in the research the limitations on the number of samples available led us to use only one coefficient for heart sound segmentation.

Comparing the PCG trace and the plots corresponding to the cepstral coefficients it is possible to identify the relation between the cepstral coefficients and the mechanical events during a normal cardiac cycle. The fifth coefficient for example, shows a peak of higher magnitude for the second sound S2 that may

differentiate it from S1. The value of this representation becomes even more evident when comparing the phonocardiograms related to different pathological conditions.

Figure 5.3 shows the plots of the first five cepstral coefficients for three phonocardiograms representing a normal PCG heart sound, a ventricular septal defect (VSD) murmur and an aortic stenosis (AOS) murmur. The data in the first column of the plot represents the normal data, the second column the VSD data whilst the final column contains the AOS data. The first row shows the PCG

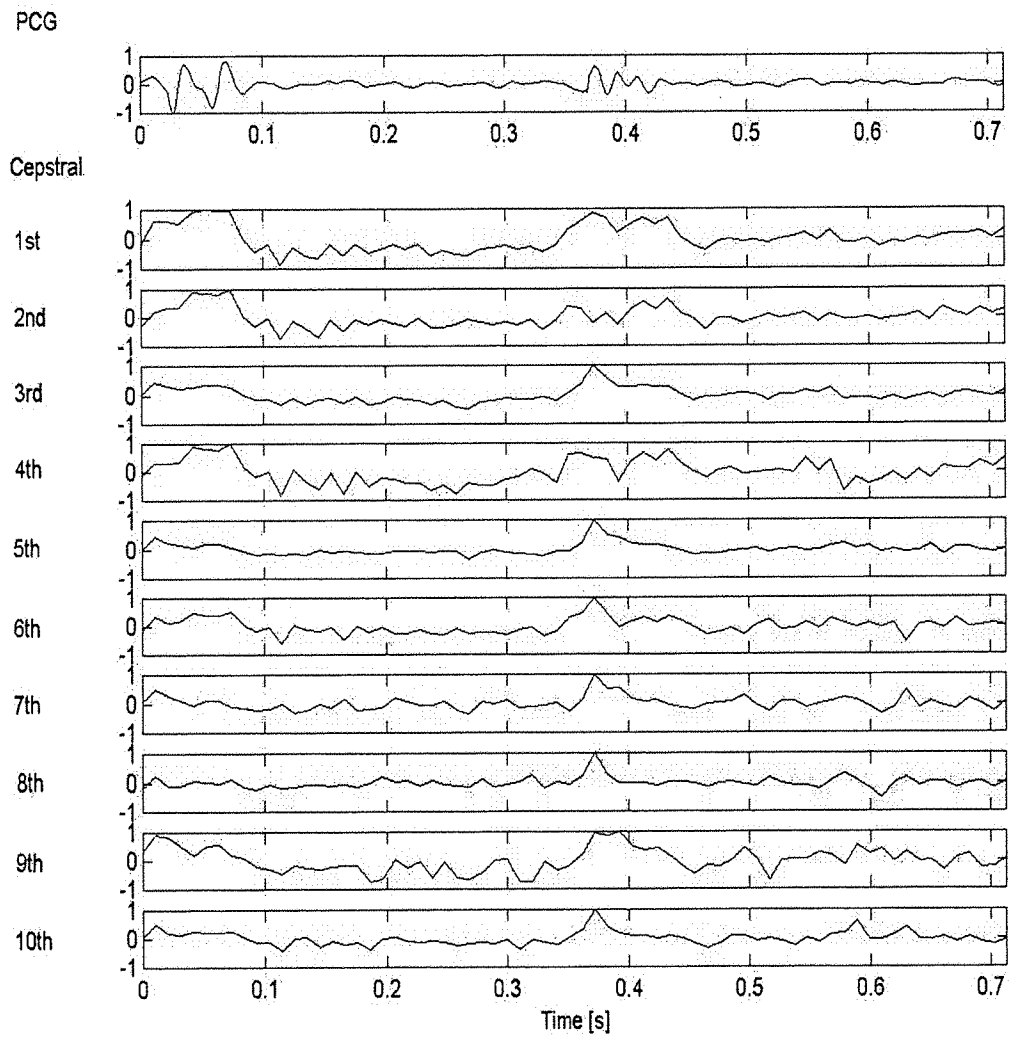


Figure 5.2. Cepstral Representation of a Normal PCG. PCG recorded at pulmonary area, Fs =4 KHz.

The first cepstral coefficient, which is essentially related to the power of the signal (see appendix B), clearly shows the difference between normal and abnormal PCGs. The first coefficient in the second column shows the distinctive pansystolic murmur characteristic of a ventricular septal defect. The murmur presents a decrescendo characteristic towards S2. Although the initial strength of this murmur is obscured by the first sound in this plot, S1 and S2 are clearly seen in the second coefficient. The murmur's characteristic contribution is more evident in the third and fourth coefficients.

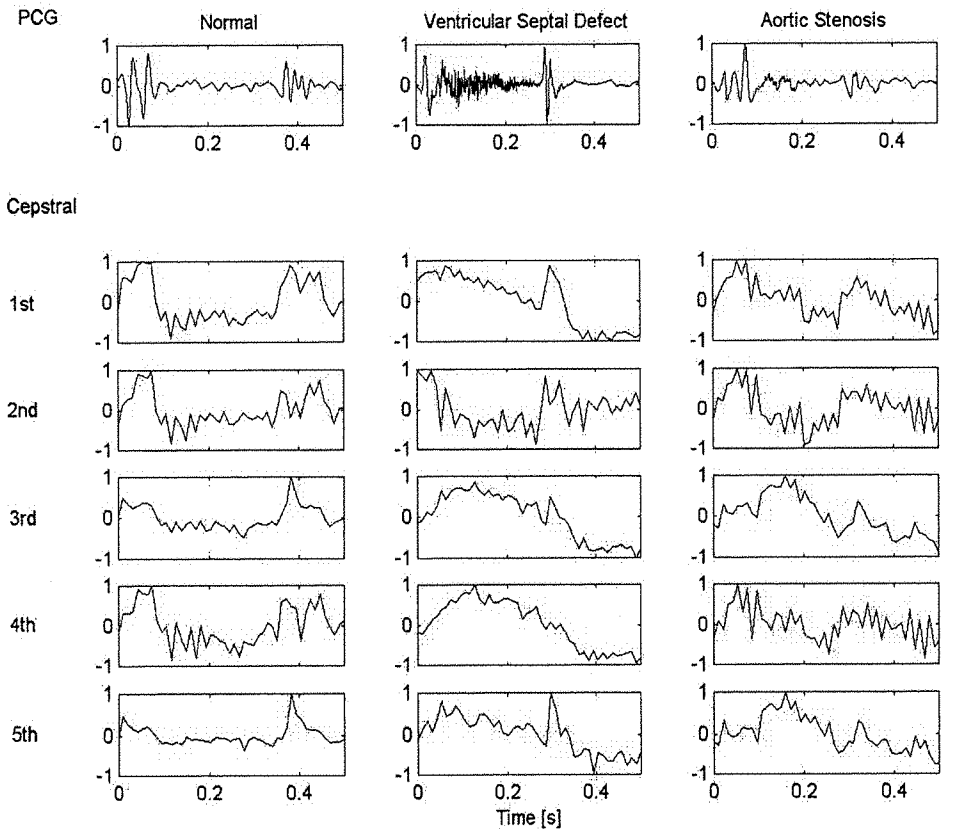


Figure 5.3. Comparison among Cepstral Representations of PCGs for Normal, Ventricular Septal Defect and Aortic Stenosis Conditions. Recorded at pulmonary, lower left sternal edge and aortic areas, Fs =4 KHz.

The third row shows the early systolic murmur caused by aortic stenosis. The plots show a pronounced increase in the third and fifth cepstral coefficients at the

time of the murmur. Note that as in the PCG, the contribution of the S2 component in the fifth coefficient appears diminished compared to the other two conditions.

The aim of figure 5.3 is to show that the cepstral coefficients obtained from the PCG display different characteristics depending on the pathologies associated with the heart sound and therefore, to show that the cepstral representations are suitable to be used in a classification algorithm.

5.2.2 Advantages of cepstral representations

The procedure to obtain the cepstral coefficient to represent PCGs was described and it was shown that this representation could be used to visually differentiate medical conditions.

Cepstral analysis has been applied before to PCG analysis by analogy to its use on speech analysis for vocal tract - excitation models (Rangayaan 1978b). In these studies, the whole cardiac cycle was used to compute the complex cepstra and after low and high liftering (filtering in the cepstral domain), the process was inverted to obtain a system response and an excitation sequence respectively. The excitation sequences were found to differ between normal and abnormal PCG (aortic stenosis and mitral stenosis); however, the system responses were discarded since it was argued that they lacked of useful information.

The concept of cepstral analysis is further expanded in this work proposing the tracking of the variations of the short-term real cepstral coefficients for classification. Instead of using the computationally demanding approach of complex cepstral deconvolution with its phase unwrapping complications, the signal is represented as pattern variations in the power cepstral domain. In the same manner as a time series is better represented in the time frequency plane to enhance frequency patterns, the PCG signal is chosen to be represented in the new plane.

Cepstral truncation is required to keep the number of input features low, and therefore, the slow varying low-quefrency coefficients related to the system response are used. However, most of the faster varying high-quefrency coefficients relating to the excitation are discarded. In an opposite approach to Rangayaan (1978a), who discards the reconstructed system response, we propose to track the variations in the low order coefficients associated with the system response. The viability of these representations to discern patterns of PCG characteristic of specific pathologies has been shown.

The use of these representations may be useful for the differentiation of turbulence induced murmurs due to abnormal communications from murmurs produced by non physiological hyperflux conditions. For example, atrial septal defect murmurs and innocent pulmonary flow murmurs present a similar grade II mid systolic ejection murmur, best heard in the pulmonary area. Both murmurs are caused by an increased volume of blood ejected by the right ventricle through a normal pulmonary valve. Nevertheless, the increased blood flow is caused by an extracardiac connection in the first case, whereas in the second case the increased blood flow results from a structural and functionally normal heart (Guadalajara, 1998). A similar excitation signal may be considered to set the system into vibration: the turbulent flow in the pulmonary valve, however, a system response from a heart with a hole in the atria, in theory, may be expected to be different from a system response of a normal heart.

The use of cepstral coefficients to represent the PCG has some computational and mathematical advantages. Since several coefficients are used to represent the PCG signal, the input features for the HMM are multidimensional and therefore, if the state observation probabilities are represented by single Gaussian distributions, they would form a multivariate Gaussian distribution (Holmes, 2001, page 143). If the features do not vary independently, a covariance matrix specifies their interdependence. The entries along the main diagonal of such a matrix represent the variance of each feature, whilst the remaining entries indicate the extent to which the separate feature distributions are correlated with each other. Assuming

this set to be of highly uncorrelated features simplifies the computation of the parameters in the hidden Markovian model since the covariance matrix can be reduced to a matrix with only diagonal elements, the variances (Jelinek, 1999).

5.3. Maass-Weber filters

During auscultation, a physician as an experienced expert, performs a subjective diagnosis of heart murmurs by an analysis of frequency and intensity patterns, relative intensity levels, timing, and localisation.

In the recording of phonocardiograms, relatively high-energy components in the low frequency regions often obscure high frequency components of relative low energy. Therefore, in clinical phonocardiography, a standard set of filters is usually required to ease the visual analysis of the PCG signals. The filters required have been empirically defined and the set comprises a low-frequency band filter t_{35}, two medium frequency filters $m1_{70}$ and $m2_{140}$, two high frequency filters $h1_{250}$ and $h2_{400}$, and a logarithmic filter m_g to mimic the response of the human ear, although this last filter is of low clinical value. This set of filters, known as the Maass-Weber filters, are standard for clinical phonocardiography analysis (Holldack, 1965, 1974; Barschdorff, 1995; Guadalajara, 1998).

Figure 5.4 shows the frequency response of the Maass-Weber set of filters. The cut-off frequencies are defined at 10% below nominal amplitude (unitary gain). The critical frequency of the low pass filter t_{35} (from the German *tiefe*) is set at 35 Hz with a 7.5 dB/octave roll-off. The first medium frequencies band pass filter $m1_{70}$ has a cut-on frequency at 70 Hz, an 18 dB/octave roll-on and a 24 dB/octave roll-off. The gain of this filter reinforces the frequencies within the band centred at 140 Hz (this reinforcement is a common characteristic of the subsequent filters). The filters $m2_{140}$, $h1_{250}$ and $h2_{400}$ have cut-on frequencies at 140 Hz, 250 Hz and 400 Hz respectively with symmetric roll-on and roll-off slopes of 24 dB/octave. Note the different maximum gains, since the filters are specified by the cut on frequencies at 0.9 gain and the central frequencies. This set was empirically

defined in the early days for PCG plotters where the gain was manually adjusted to enhance specific features on the clinical findings.

Figure 5.4. Frequency Response of the Maass-Weber Filters t_{35}, $m1_{70}$, $m2_{140}$, $h1_{250}$, $h2_{400}$.

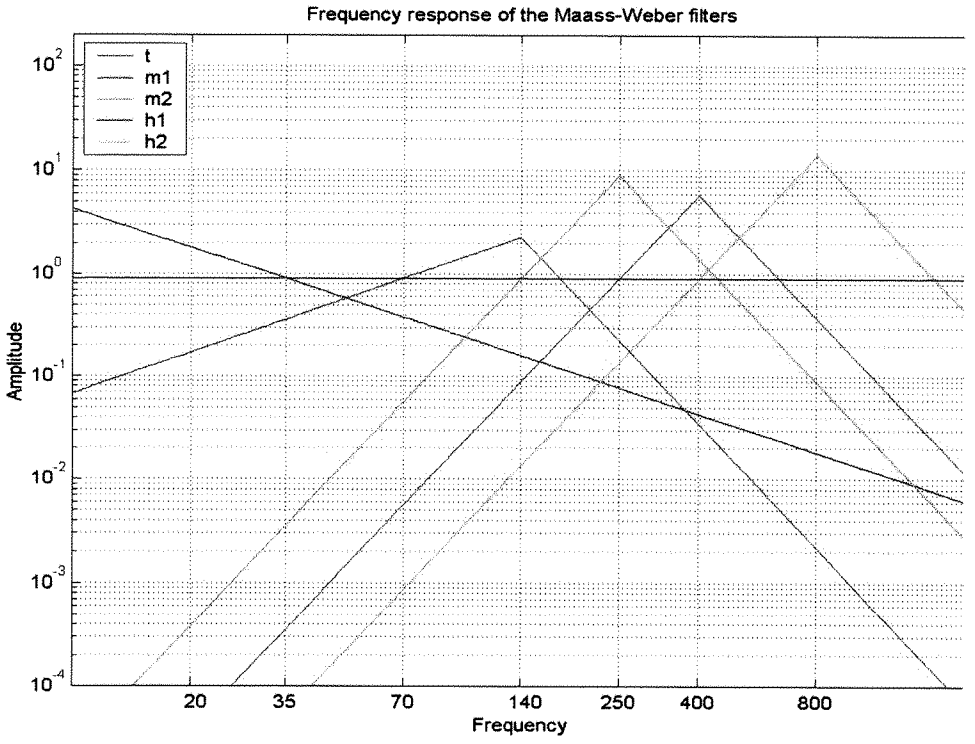


Figure 5.5 shows a normal PCG recorded at the LLSE position with this set of filters. The first and the second sound can be identified in all the frequency bands; however, the initial and final segments of S1, which are of low frequency (around 30 Hz), are better seen on t_{35} and are not visible on $m2_{140}$ or $h1_{250}$. The main group of S1 with frequencies between 100 Hz and 150 Hz is better seen on $m2_{140}$ and $h1_{250}$. The first half of the second sound (aortic) has higher amplitude and higher frequency components than the second half (pulmonic). Their respective contributions can be easily recognised on t_{35}, $m2_{140}$, and $h1_{250}$. As an example of the use of these filters to ease the identification of pathologies, Figure 5.6 shows an innocent systolic murmur recorded with the Maass-Weber set of filters. The murmur is predominant in $m1_{70}$ and $m2_{140}$ and the split in S2 is more evident in

$m1_{70}$.

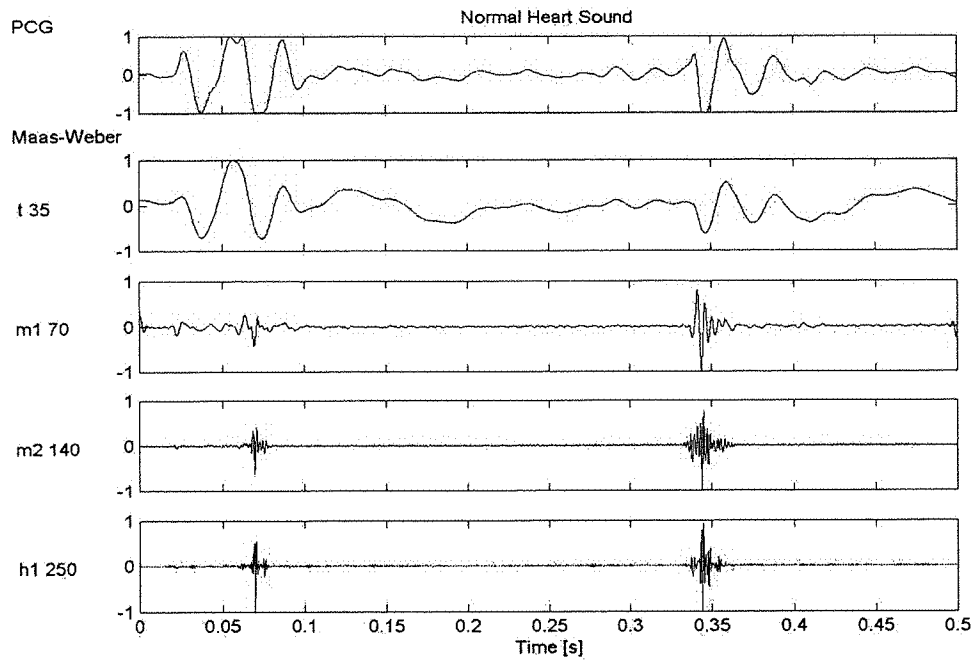


Figure 5.5. Normal Phonocardiogram Filtered with the Maass-Weber Set. PCG recorded at lower left sternal edge area, Fs =4 KHz.

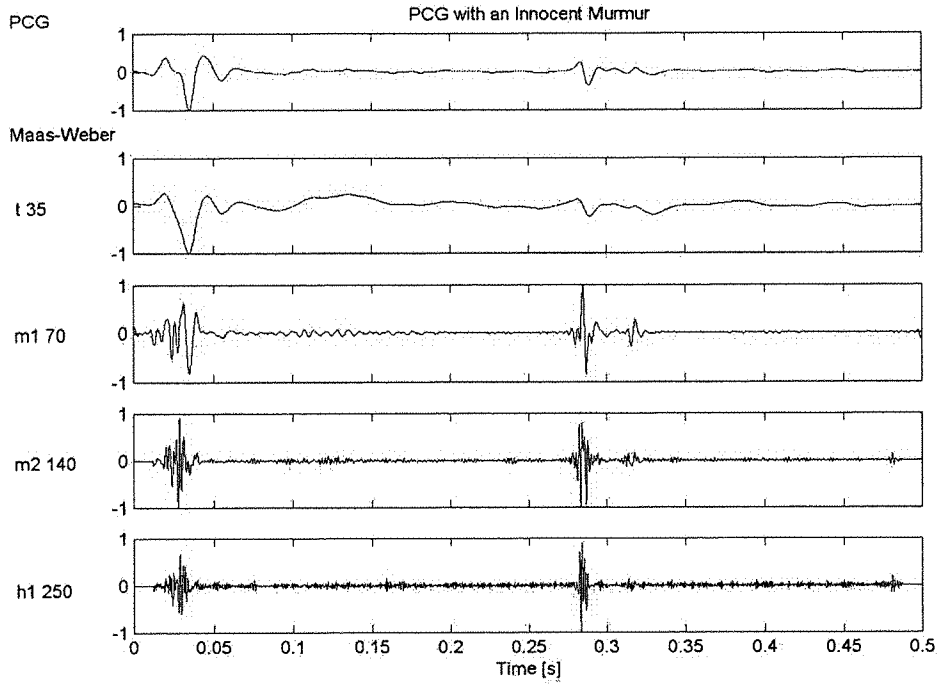


Figure 5.6. PCG with an Innocent Systolic Murmur Filtered with the Maass-Weber Set. PCG recorded at pulmonary area, Fs =4 KHz.

5.3.1. Feature representations based on Maass-Weber filters

To obtain a simple set of features, the PCG signal is segmented into frames of 20 ms overlapped by 50%. The short time Fourier spectra of each frame is computed and weighted by the five band Maass-Weber filter characteristics. A single value for each band is obtained by computing the normalised average spectral power yielding five coefficients per frame (Barschdorff, 1995).

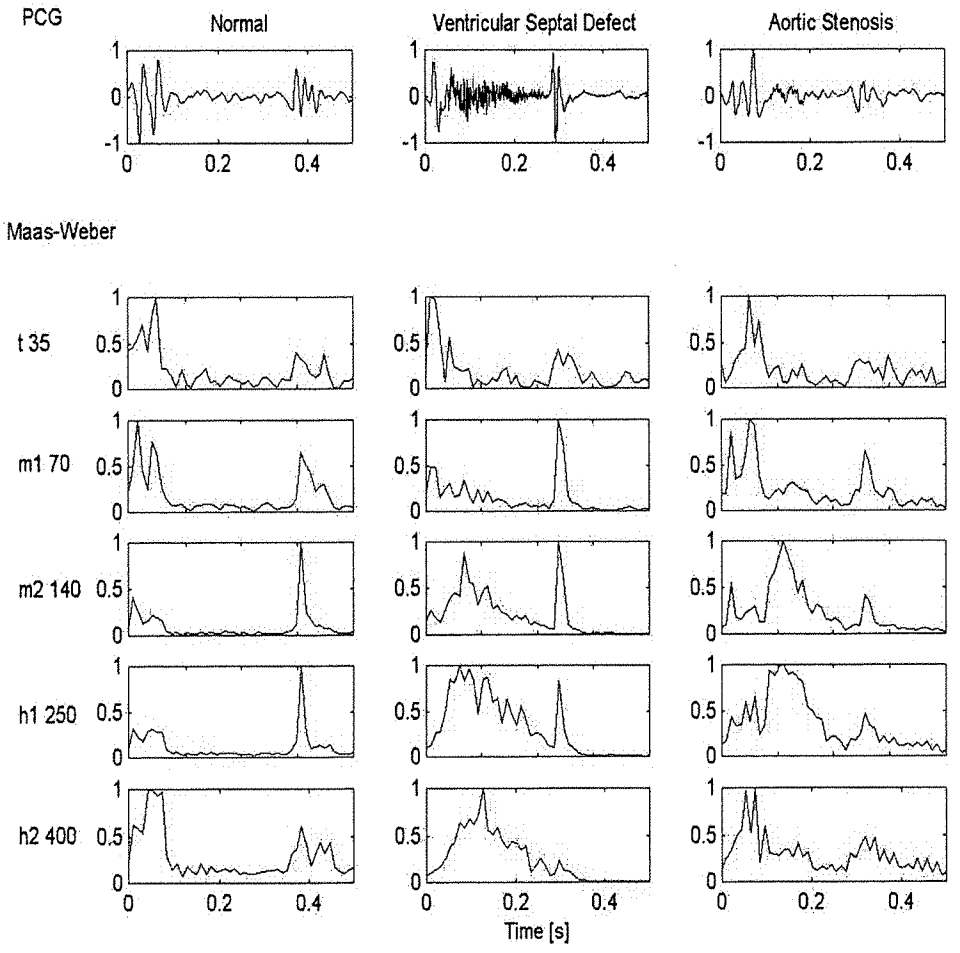


Figure 5.7. Coefficients Obtained for Each Band of the Maass-Weber Filter for Different Conditions. Recorded at pulmonary, lower left sternal edge and aortic areas, Fs =4 KHz.

Figure 5.7 compares the features obtained for signals characteristic of three clinical conditions, namely: a normal PCG, a heart sound signal with a characteristic murmur caused by a ventricular septal defect and a PCG with a

systolic murmur, typical of aortic stenosis. The first row shows the unfiltered record of the PCG, whereas the subsequent rows represent the normalised energy coefficients obtained for each frequency band in a time-frame axis. As expected, in the normal heart sound, no particular signature is present during the systolic cycle. The particular signature of a pansystolic VSD murmur is easily identified in the bands $m2_{140}$, $h1_{250}$ and $h2_{400}$. The mid-systolic ejection murmur, characteristic of aortic stenosis, is clearly seen in the bands $m2_{140}$ and $h1_{250}$. Its medium-high frequency component contributions are not as remarkable in $h2_{400}$ as in the previous bands.

Note the similarity between the Mass-Weber representations shown in figure 5.7 and the cepstral representations shown in figure 5.3. In the case of the ventricular septal defect murmur, for example, the signal from the filter h1 250 is remarkably similar to the signal obtained for the fifth cepstral coefficient. The fourth cepstral coefficient for the VSD murmur exhibits the strong systolic murmur and diminished S1 and S2 components, those characteristics are reflected also in the signal obtained for the filter h2 400 in figure 5.7.

If there were a restriction on the use of only one coefficient for each representation, the best choice would be the first cepstral coefficient since it better represents the distinctive shape of the murmurs and the relative amplitudes of S1, S2 and the murmur. For the Maass-Weber coefficients the h1 250 signal would be a good candidate although the relative amplitudes of the components are not maintained.

The Maass-Weber filter set has been defined through experience to aid in visual differentiation, and consequently, it is not surprising that the features obtained are a good descriptor for a visual assessment of the differences between pathologies (at least for the typical signals presented). Nevertheless, the differences that discriminate between similar murmurs are not always discernible by visual inspection of these graphs, and therefore, their performance as feature representations for an automatic classification algorithm have to be evaluated within the context of the method selected and the database available.

5.4. Linear prediction coefficients

The linear prediction model forecasts the amplitude of a signal at certain times using a linear weighted combination of a number of past samples (they are also referred to as autoregressive models). If a predictor error is defined as the difference between the actual sample value and its predicted value, a set of the best prediction coefficients can be defined by minimising a mean square error criterion.

The success with which a signal can be predicted from previous samples depends on the autocorrelation function, or equivalently, on the bandwidth and the power spectrum of the signal. A predictable signal has a smooth and correlated fluctuation in time, and the energy will be concentrated in narrowbands of frequencies, whereas the energy of an unpredictable signal is spread over a wide band of frequencies. Most signals are partially predictable and partially random. These signals can be modelled as the output of a filter excited by an uncorrelated input. The aim of linear prediction is to model the mechanism from introducing correlation in a signal; i.e. the filter characteristics, and therefore in the frequency domain LP analysis can be regarded as a method of spectral modelling. For details of the technique see Makhoul (1975), Vaseghi (2000, pp. 227 - 261).

LP analysis has been used before as input feature representations for PCG parametric classifiers (Itawa, 1977). The technique is based on a model of resonances in which the distribution of poles provides the information of interest, such as high frequency poles corresponding to murmurs and low frequency poles representing the heart sounds. A feature space formed with prominent poles of the models (with bandwidth < 80 Hz) can be used for classification (Itawa, 1980).

Figure 5.8 shows an example of the FFT spectrum of a section of the PCG signal comprising a second heart sound and the spectrum obtained by modelling with an eighth order linear predictor. A better estimate of the underlying spectrum can be

achieved by increasing the order of the filter. The signal also shows the time domain signal and the all pole diagram of the filter model.

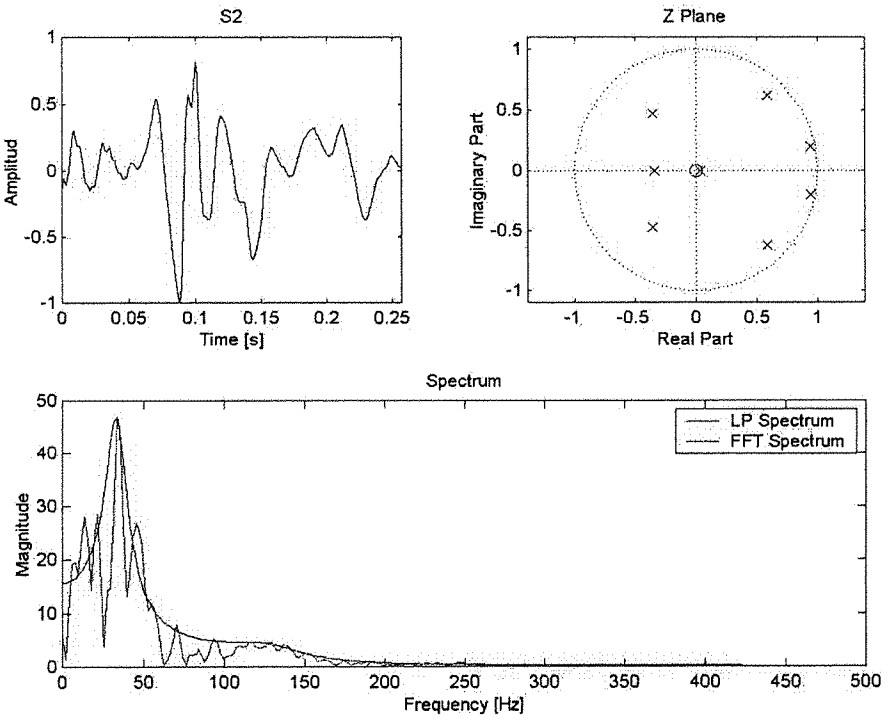


Figure 5.8. PCG Time Signal of the Second Heart Sound, Z Plane Diagram of the All-pole Filter, FFT Spectrum and LP Spectrum. PCG recorded at pulmonary area, Fs =4 KHz.

For the analysis of PCGs, Itawa (1997) suggest an eight-pole filter to model the spectral contours of 25 ms time frame windows overlapped by 50%. In practice, the prediction coefficients are not a good representation to use for most applications. In cases in which the digital word length is critical, the polynomial coefficients tend to be too sensitive to numerical precision. The coefficients are not orthogonal, which potentially creates other difficulties for classifiers that might use these features (Gold, 2000). Interpolating between parameters corresponding to two different filters will not vary the frequency response smoothly from one to the other: stability is not even guaranteed.

For all of these reasons LPC coefficients are generally transformed into one of a number of representations including pole positions, reflection coefficients, log area ratios and line spectral frequencies (Makhoul, 1975; Itakura, 1975; Deller, 2000).

The pole positions are obtained finding the roots of the polynomial defined by the set of LPC features. The polynomial roots are either real or occur in complex conjugate pairs. The main disadvantage of using the pole positions is that the frequency response is sensitive to pole positions near the unit circle.

To obtain the reflection coefficients the LPC polynomial is transformed into a set of coefficients that represent the fraction of energy reflected at each section of a non-uniform tube with as many sections as the order of the polynomial. These reflection coefficients lie between -1 and $+1$ (for totally open and totally closed boundary conditions), however, the coefficients also become very sensitive to noise when they are near to ± 1.

Log area ratios, defined as the inverse hyperbolic tangent of the reflection coefficients, are used to wrap the amplitude scale of the reflection coefficients to decrease their sensitivity to quantization errors when their magnitude is near to unit.

5.4.1 Proposed representation: line spectral pairs

Line spectral pairs are an alternative to LPC representations. In this technique, the predictor polynomial computed by the auto-correlation method of linear prediction is split into a symmetric and an anti-symmetric polynomial by extending the order of the filter (without introducing any new information). Letting the added reflection coefficient be 1 or -1 is equivalent to setting the corresponding acoustic tube model to be completely closed or completely open at the added stage (Itakura, 1975). The poles of the resulting symmetric and anti-symmetric polynomials are interlaced with each other and are on the unit circle

and therefore these pairs of poles can be uniquely represented only by their phase i.e. providing a set of line spectral frequencies.

Figure 5.9 shows the line spectral frequencies computed from the LP spectrum of figure 5.8. The vertical lines represent the line frequency pairs. The peaks in the spectra are defined by the position of the pairs and the distance between them. As shown in the figure, the first pair of LSF which are close to each other result in a sharp peak in the frequency response curve. The frequency response is relatively flat around a pair of LSF far from each other as is the case in the last pair for example. A spectrum where the signal amplitude is low in all frequency bands will be represented by equally spaced LSF.

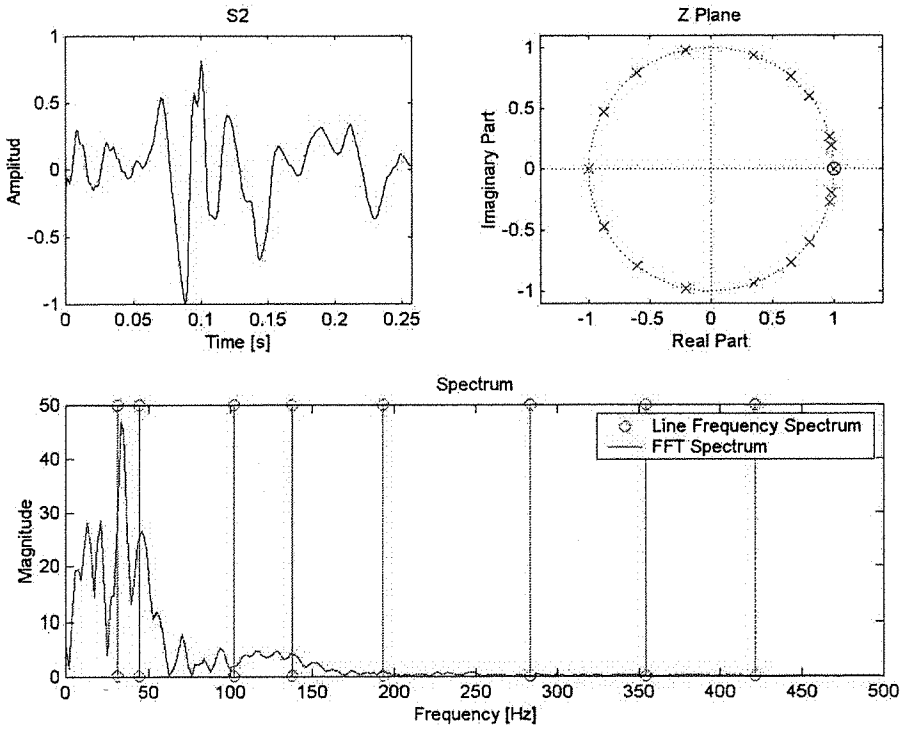


Figure 5.9. PCG Time Signal of the Second Heart sound, Z Plane Diagram.

PCG recorded at pulmonary area, Fs =4 KHz.

Figure 5.10 shows a PCG with a murmur typical of aortic stenosis and the representation of the line spectrum frequencies computed form time windows of 25 ms overlaped by 50%.

The figure shows the time evolution of the four line spectral pairs. The first line spectral pair (bottom two lines) is closely spaced and just below 50 Hz indicating a low frequency peak persisting across time. The fourth pair (uppermost lines) are general far-spaced lines characteristic of a flatter spectrum; however, within the range from 0.1 to 0.3 seconds the lines become closer with a shift of frequency showing peaks in the spectra. Although the interpretation of this line frequency

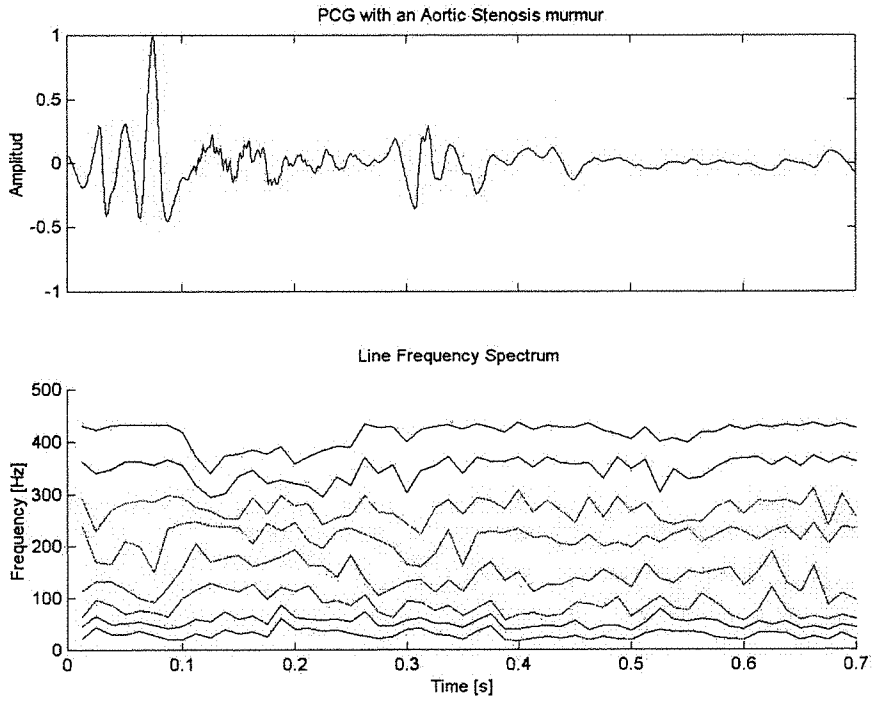


Figure 5.10. PCG with an Aortic Stenosis Murmur and its Line Frequency Spectral Representation. PCG recorded at aortic area, Fs =4 KHz.

spectrum is not straightforward, the presence of the murmur in the systolic cycle can be seen in the line frequency spectrum. To ease its interpretation, Figure 5.11 shows the spectrogram of the same signal and the line frequency spectrum superimposed.

Figures 5.11 and 5.12 show how line frequency spectral pairs can represent the spectrogram of the signal in a simplified form. Each line spectral pair represents the location of the spectral peaks by their position and the sharpness of the peak by their proximity. In Figure 5.11 closely spaced line pairs in the range from 200

to 400 Hz and 0.1 to 0.25 seconds indicate the presence of a systolic murmur. By comparison, in Figure 5.12 where no murmur is present the line pairs are far spaced.

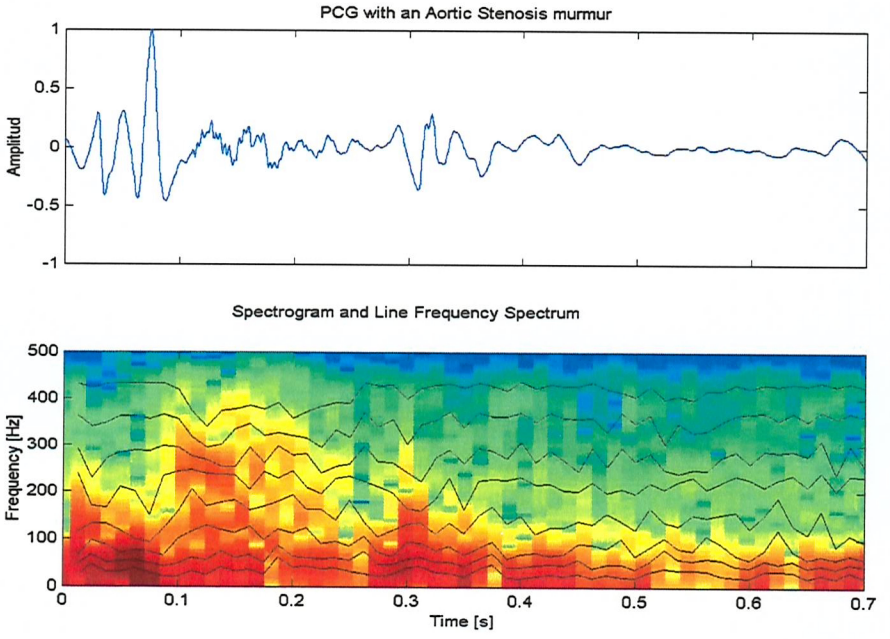


Figure 5.11. Spectrogram and Line Frequency Spectrum of a PCG with a Sytolic Murmur Typical of Aortic Stenosis. PCG recorded at aortic area, Fs =4 KHz.

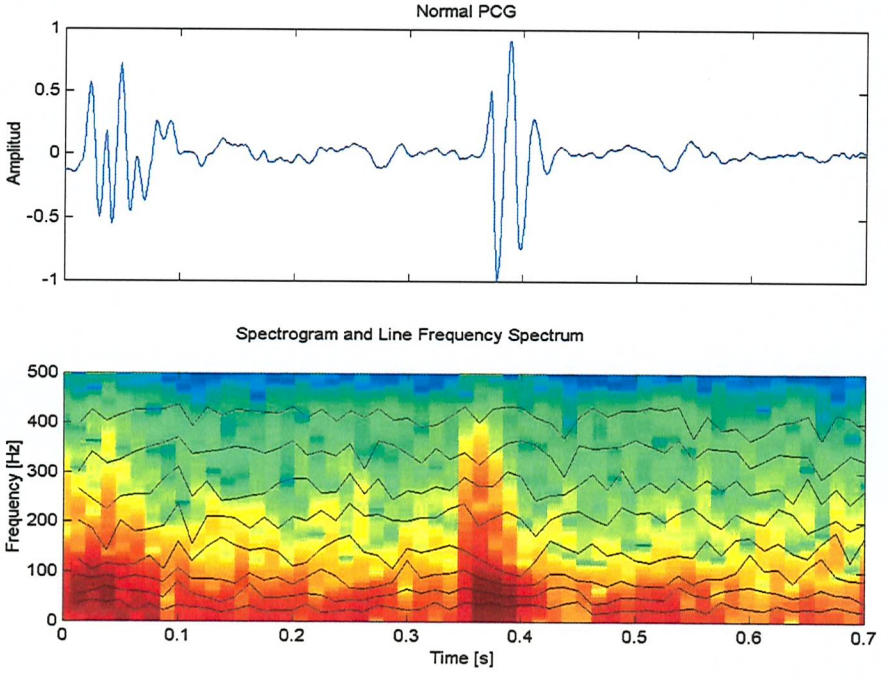


Figure 5.12. Spectrogram and Line Frequency Spectrum of a Normal PCG. PCG recorded at pulmonary area, Fs =4 KHz.

5.5 Summary

The viability of using power cepstral analysis, line spectral pairs and Maass-Weber derived energy features to provide a simplified representation of the phonocardiograms has been demonstrated. In those techniques the variations of the time signal that reflect the occurrence of pathologies lead to traceable differences in those feature representations. It has been shown that some of them may be especially suitable to ease computational demands. Nevertheless the viability of the use of those representations as inputs for the classifier needs to be assessed for the particular models selected and the database available. The next chapters will assess their performance within this context.

The following chapter shows how those representations can be used to solve a related problem: the segmentation of heart sounds.

6 Heart Sound Segmentation by Hidden Markovian Models

6.1 Introduction

During auscultation, one of the main clues to differentiate similar murmurs corresponding to different pathologies is their position with respect to the main components of the heart sound. For instance, murmurs are initially classified into systolic and diastolic murmurs depending on their timing within the heart cycle and further labelled as early, middle and late murmurs by their occurrence within the systolic and diastolic cycles. Components of the PCG are also required to be identified since they also present changes due to cardiac abnormalities, for example, the wide fixed split of S2 associated with atrial septal defects or the accentuation of the third heart sound in some ventricular septal defects.

It is therefore of prime importance for an autonomous system to be able to identify the main components of the PCG and the relative position of abnormal sound events within the heart cycle. Once the main components are identified, individual portions of the PCG signal can be extracted for a more detailed analysis and modelling. Accordingly, the segmentation of heart sounds into its components is commonly regarded as the first step towards the automatic classification of heart sounds (Iwata, 1980; Liang, 1997; Lehner, 1987).

A summary of the techniques available for the segmentation of the phonocardiogram is presented in this chapter before a new approach is proposed. Since previous techniques for heart sound signal segmentation make extensive use of other biomedical signals such as the electrocardiogram and the carotid pulse, the relation between these events is briefly explained. Using the cepstral coefficients derived from the PCG as input features, a hidden Markovian model is then proposed to identify the main components of the phonocardiogram and consequently perform the heart sound signal segmentation.

6.2 Heart sound segmentation using the electrocardiogram and carotid pulse signals

In a normal PCG (at normal heart rate), the first and the second heart sounds can be identified by visual inspection, but the presence of murmurs and high heart rates makes their localisation within the cardiac cycle difficult. Moreover, the non-specific nature of vibration signals, the various transmission paths from the heart to the chest surface, and superimposed background noises including ambient, instrumental, and extra-cardiac noise, further hinders their detection. For these reasons, several proposed systems make use of auxiliary signals like the electrocardiogram and the carotid pulse for automatic identification of heart sound components (Rangayyan, 2002).

Figure 6.1 shows the main components of the heart sound signal namely: first sound (S1), second sound (S2), the systolic and diastolic phases, and their relation to events in the electrocardiogram (ECG) and the carotid pulse (CP).

In the ECG, the sino-atrial node, a basic natural pacemaker triggering its own action potentials starts the cycle. Electrical activity is propagated through the atrial muscle causing contraction of the atria. This slow activity and the small size of the atria results in a low-amplitude slow P wave on the ECG. The excitation delayed at the atrio-ventricular node results in a normally iso-electric segment known as the PQ segment. This pause assists in the completion of blood transfer from the atria to the ventricles. The *His* bundle, the bundle branches, and the Purkinje system of specialised conduction fibers propagates the stimulus to the ventricles at high rate.

The electric stimulus is spread from the apex of the heart upwards causing rapid contraction of the ventricles, this results in the QRS wave of the ECG. As the ventricles contract, the tension in the *chordae tendineae*, and the pressure of retrograde blood flow towards the atria seals the mitral and tricuspid valves causing the initial vibrations of the first heart sound (S1) appearing immediately after the QRS complex.

Ventricular muscle cells have a relative long action potential duration, which causes the ST segment. Finally, relaxation of the ventricles causes the slow T wave and the second heart sound S2 appears slightly afterwards, due to the closure of the aortic and pulmonary valves (see chapter 2).

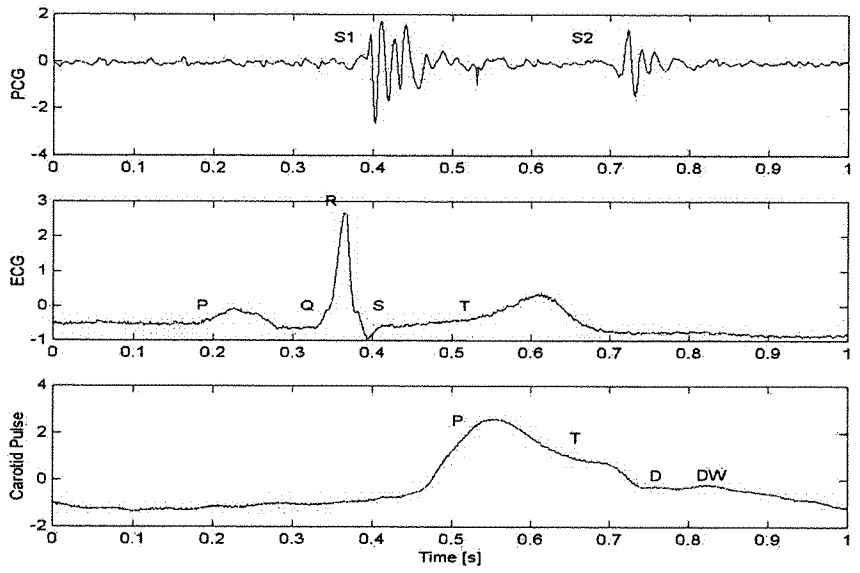


Figure 6.1. Correlation of the Components of the Heart Sounds with the Electrical Activity Represented by the ECG and the Carotid Pulse. *From Rangayyan (2002).*

The QRS complex is a reliable indicator of the beginning of S1, and since its detection is fairly easy, PCG segmentation into cardiac cycles is usually accomplished aided by the ECG trace. There are several signal processing techniques for QRS complex detection, among them the Pan-Tompkins algorithm (Pan, 1985), and several studies related to PCG analysis where the signal is first segmented using the ECG as an auxiliary signal (Barschdorff, 1995; Haghigi-Mood, 1995).

The time relationship between the T wave of the electrocardiogram and the second heart sound suggests the possibility of using the former for identification of the latter. However, as explained in chapter 2, the T wave can not be referred to a specific event *per se*, is often of low amplitude, and almost negligible in many

ECG records. For this reason, the identification of S2 is usually performed with the aid of another signal reference: the carotid pulse.

The carotid pulse is a signal recorded at the neck over the carotid artery providing indication of the variation in arterial blood pressure and volume within each heart beat. The percussion wave (P) in the carotid pulse shows the increase of pressure due to the abrupt ejection of blood from the left ventricle to the aorta.

A secondary wave caused by a reflected pulse returning from the upper body causes the tidal wave (T). Closure of the aortic valve is accompanied by deceleration and reversal of flow in the aorta causing a sudden drop of pressure. The dicrotic notch (D) is consequence of the latter and may be followed by a dicrotic wave (DW).

The dicrotic notch, measured at the carotid artery, has a delay of 4.26 ms with a standard deviation of 5 ms with respect to the second heart sound signal (S2) (Tavel, 1978) and can be used as a reliable indicator of the end of systole (Lehener, 1987).

Summarising, in order to perform the PCG segmentation the ECG signal is generally used to detect the first heart sound (Barschdorff, 1995; Haghigi-Mood, 1995; Iwata, 1980), whereas the carotid pulse is generally used as a reference to detect the second heart sound (Lehner, 1987). Nevertheless, the use of these external references involves the use of extra equipment, which may not be desirable or feasible. Furthermore, the timing between electrical and mechanical activities in a cardiac cycle will not be exactly constant among patients in some pathological conditions like Bundle branch block, severe aortic or mitral disease, and hypertropic left ventricle (Haghigi-Mood, 1995). Under these circumstances, an alternative approach is the use of the inherent characteristics of the PCG signal to perform the segmentation.

6.3 Signal processing techniques for PCG segmentation without an external reference signal

The need for an autonomous system able to identify the PCG signal component without use of an external reference has been recognised and acknowledged in the literature (Itawa, 1980, Haghigi-Mood, 1995, Rangayyan, 2002). Typical examples of methods that rely only on the PCG signal are spectral tracking at a single frequency (Itawa, 1980), spectral tracking over a frequency band (Haghigi-Mood, 1995), methods using a measure of the signal energy (Liang, 1997) and methods using matching pursuit techniques (Sava, 1998).

Itawa (1980) uses linear prediction analysis techniques as described in Chapter 5 to represent the spectra of a high pass filtered PCG signal recorded at the apex. The filter defined by standards of the Japanese Society of Medical Electronics and Biomedical Engineers has a 26 dB per octave roll-off with a cut off frequency of 250 Hz. This filter band was selected as the most representative for murmur detection. A group of sixty nine subjects was analysed to establish statistically the peak frequencies of the first and second heart sound, and consequently two tracking frequencies were chosen: 100 Hz for detection of S1 and 150 Hz for S2. Peaks in the spectral level in these frequency bands represents the occurrence of heart sounds. No carotid pulse is needed for S2 detection, and only one PCG filter sub-band (defined in a similar manner as the Maass-Weber filters described in chapter 5) is required.

Haghighi-Mood (1995) presents a similar algorithm in which instead of tracking a single predefined frequency, he tracks the energy in a certain PCG dependent frequency band and uses a length weighting function to distinguish the relatively short S1 and S2 peaks from longer peaks corresponding to murmurs. The frequency band used to track is defined for each PCG signal using the first heart cycle which is individually evaluated to obtain the first peak in the spectra considered to be either S1 or S2. Consequently, a –3 dB threshold around this peak defines the bandwidth. The method is less sensitive to recording quality and

the overall frequency response of the system since the frequency band selection depends on each PCG. No ECG signal is required to segment the PCG signal.

Liang (1997) proposes a segmentation method based on the envelope of the PCG signal calculated using the normalised average Shannon energy. The squared time signal is multiplied by its logarithm and averaged; therefore, medium intensity signals are emphasised whilst low intensity ones are more attenuated than the high intensity ones. The Shannon energy shortens the difference between the envelope intensity of the low intensity sounds and the high intensity ones, making low intensity signals easier to detect. A threshold level of this envelogram is firstly used to select the peaks. A time interval set of rules is then defined to exclude additional peaks. After examining time intervals, the threshold is adjusted to include possible missing peaks. Finally, considering that the diastolic interval is longer, and that the systolic interval remains relatively constant, S1 and S2 peaks are identified by their relative time occurrence. For PCG signals that include interfering signals like ambient noise or artifacts the method provides incorrect identification, and although these can be avoided by improving the recording techniques, cases of murmurs that overlap either S1 or S2 make the identification impossible.

Sava (1998) proposes to decompose the PCG signal in basic units arranged in an energy decreasing order. The coherent structure of the first and second sounds is expected to be represented in the first unit components since the heart sound signals convey more energy than the random murmurs, consequently the number of units is restricted to extract the main components S1 and S2. The length of the units is also limited in order to detect short transit events like the first and second sounds and to exclude long random murmurs. Once the most energetic units have been determined, they are ordered according to their time position creating a template. Subsequent events are found cross-correlating this template with the PCG signal. Considering the time characteristics of the systolic and diastolic cycles, S1 and S2 are identified. The main deficiency of this matching pursuit method is the setting of a threshold level for the cross-correlation since the

performance depends not only on the accuracy with which the template is represented, but also on the similarities of consecutive S1 and S2.

6.4 Proposed technique: PCG segmentation by hidden Markovian models and cepstral representations

Chapter three proposes that the identification of specific cardiopathologies through the automatic analysis of heart sounds can be simplified to a classification task to support differential diagnosis. Consequently, hidden Markovian models were proposed as a suitable technique to perform PCG signal classification. As part of the main stages of the process, simplified representations of the input signals are required, and therefore, Chapter four explored some techniques for feature extraction. Accordingly with the cardiohemic theory, cepstral coefficients were proposed and may provide suitable representations for the PCG signals.

To perform the automatic segmentation of the heart sound signal, a simplified representation of the PCG is also required to ease the computations. Using such representations and a sequence recognition algorithm, the main components of the PCG can be identified. Cepstral analysis and hidden Markovian models satisfy these requirements and since they will be used in a posterior stage for signal classification, it is worth exploring their use at this point for PCG segmentation.

6.4.1 Cepstral representations for S1 and S2 identification

In the previous chapter, it has been shown how the spectral envelope of the signal can be represented by the first cepstral coefficients. Figure 6.2 shows the PCG signal of a subject with normal cardiac function and the first-five cepstral coefficients computed as described in Chapter 5, see for example Figure 5.3. From the picture, it is easy to identify the portions related to the first and second sound in the first representations. Figure 5.3 shows that S1 and S2 heart sound components can also be identified in cepstral representations of PCG reflecting pathological conditions. Therefore, for the aim of signal segmentation of the PCG,

the first cepstral coefficients may provide suitable simplified representations of the signal in which the main components can be identified.

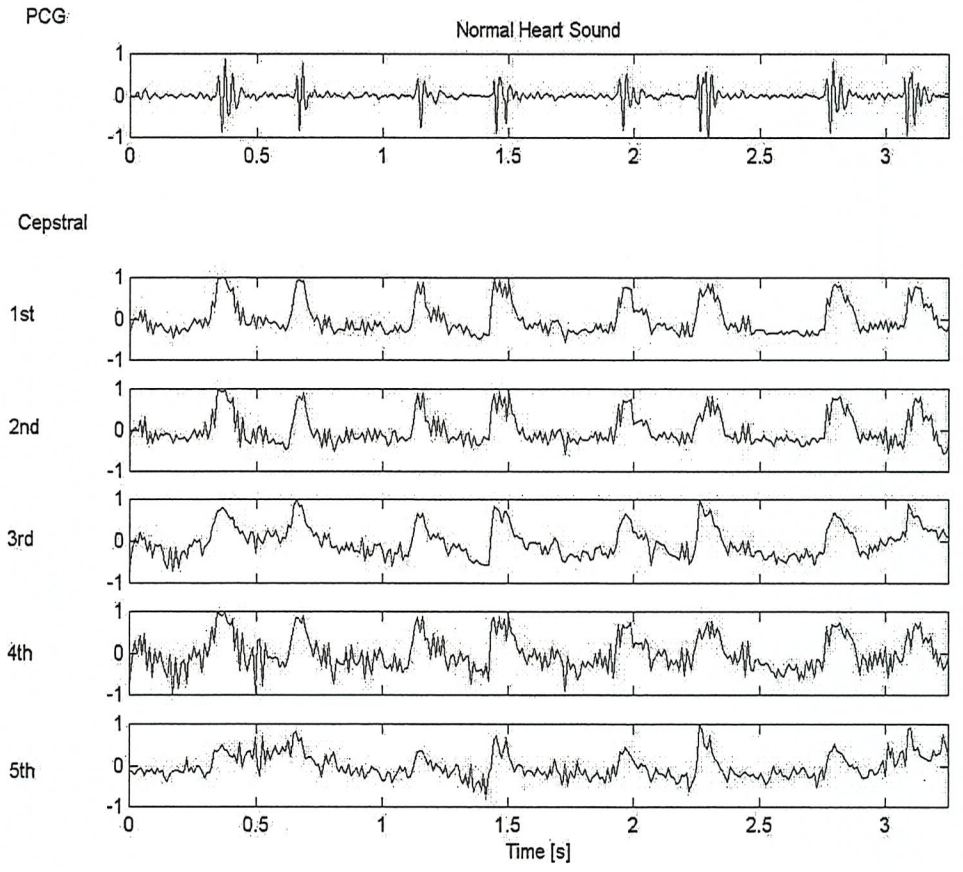


Figure 6.2. Cepstral Representation of a Normal PCG. PCG recorded at pulmonary area, Fs =4 KHz.

6.4.2 Hidden Markovian models

The cepstral representations described above can ease the visual identification of the main PCG components. Nevertheless, Cepstral coefficients and a simple technique, such as a threshold function, may not be enough to perform the segmentation considering the variability of the PCG signal, since abnormal heart conditions are reflected not only as extra sounds but also on alterations of the first and second heart sound (see figures 6.6 to 6.14). The cepstral representation of the first heart sound in figures 6.12 and 6.13 provide a good example in which a simple technique such as a threshold function may not work.

More over, for a multiple set of representations such as the Maass-Weber representations, a simple threshold would not suffice since the main contribution of different components of the signal may appear in different bands depending on its frequency content. An automatic algorithm able to distinguish these components is therefore still required.

According to the cardiohemic theory, the phonocardiogram signal recorded on the chest is the acoustical manifestation of a succession of mechanical events including variations of blood flow related to a sequence of opening and closing of heart valves and contractions of the heart chambers.

This succession of events can be modelled as a Markov process on the assumption that each of the events only depends on the previous one, and the sequence can be considered as hidden since it is not directly observable. Nevertheless, the phonocardiogram, generated by this underlying process, is available and can be modelled as the observation signal following the structure of a hidden Markov model.

The mechanical events generating each of the main PCG signal components can be represented by a single state, and consequently, each of the main components of the PCG signal, namely S1, S2, systolic and diastolic cycles can be modelled as acoustic observations related to a specific hidden state.

Using this structure and a set observation sequences (heart sounds) a model relating the hidden sequence of mechanical events to the PCG observation sequences can be trained. The parameters of the HMM namely: the state transition probabilities, the state observation probabilities and the initial conditions will be adjusted in a training stage via a recursive algorithm to represent the most probable set of values generating the training sequences.

The identification of the main components of a new PCG signal can be regarded as a decoding stage where the aim is to find the most probably sequence of hidden stages that generated the sequence of observations. The problem can be addressed

by using the Viterbi algorithm (Rabiner, 1989). Once the main components S1 and S2 are identified, the segmentation of the PCG signal can be performed accordingly.

6.4.3 The model

Due to the fact that a complex model requires more realisations of the signal in order to learn the parameters involved, it is desirable to find the simplest model that could achieve the segmentation. It is likely that the most basic model of a normal PCG is constructed by assuming that there are only two main conditions: either, a sound that could be S1 or S2, or there is a relatively silent period such in the case of systolic and diastolic events. Therefore, for simplicity, a two state HMM is initially proposed for the segmentation of normal sounds. The first state is assigned to the S1 and S2 components, whereas the second state accounts for either the systolic or diastolic cycles.

Figure 6.3 shows the two-state model for the segmentation process and an example of how the normal PCG signal can be represented as a succession of these states. The arrows indicate the possible transitions. Notice that transitions are permitted from each state to itself and to the other, this permits variations in the duration of the events. Since there are only two states there are no restrictions in the possible state transitions, nevertheless, a left-right structure HMM is shown in the diagram to represent the time succession of events embedded in the heart signal.

At this point, the model does not discriminate S1 from S2. However, the timing between events is a key parameter in the identification of the onset of the first and second sound, and the differentiation of systole from diastole can be realised by noting that the former period is shorter and its duration is relatively constant compared with the latter (Liang, 1997). Using this information, a time interval set of rules can be implemented to identify the main events of the PCG signal, which also helps to avoid false event detection.

Figure 6.3 shows the states of the model in relation to the PCG. From Figure 6.2 it can be seen that for normal PCG signals, the use of exclusively the first cepstral coefficient seems to provide a good representation for the identification of S1, S2, systolic and diastolic cycles.

The PCG signal is divided in overlapping frames of constant duration and from each frame the cepstral coefficients are computed using the methods described in Chapter 5. The continuous valued cepstral representations will be related to the states by a single Gaussian probability distribution and their characteristic parameters will be learnt from the training set of PCG signals.

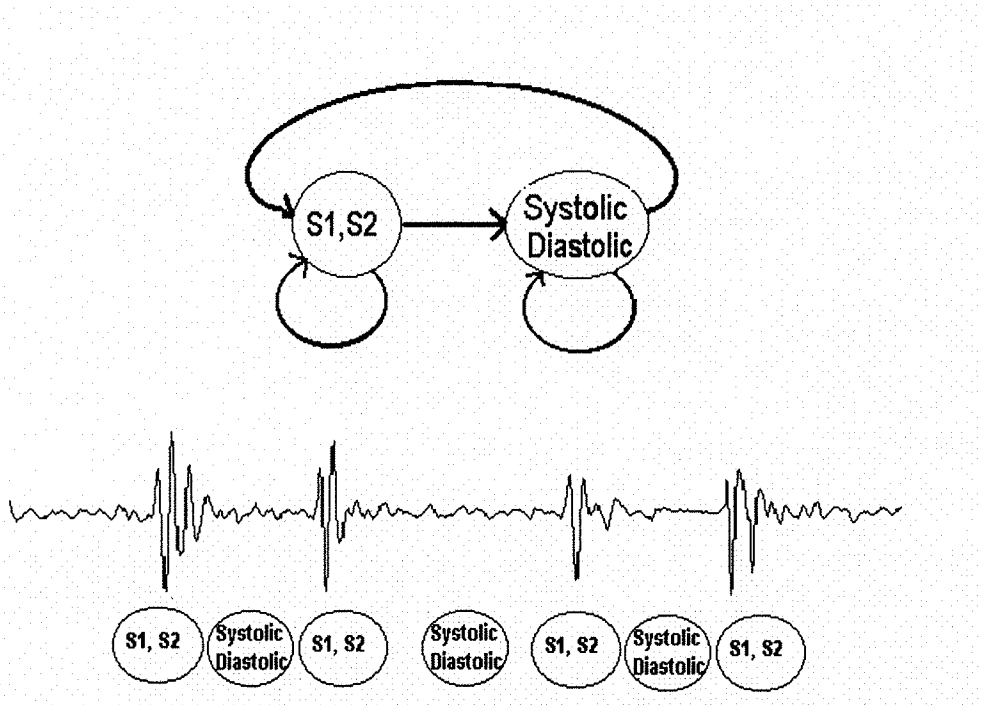


Figure 6.3. State Model for the PCG Segmentation.

6.4.4 Database

The research involves the use of an initial database of identified heart sounds collected in a previous study (Leung, 1998a) and a second set of signals acquired at the cardiology unit of Southampton General Hospital. Both studies were

approved by the Southampton and South West Health Commission Joint Ethical Committee, UK.

In the first set of heart sounds and electrocardiogram signals from children attending a paediatric outpatient clinic were digitised using a 16-bit analogue to digital converter at a sampling frequency of 4 kHz. The electronic stethoscope (e-steth range 20 Hz to 20 kHz) was positioned by a cardiologist in the standard auscultation positions, recording each heart sounds in the areas of the chest were it was best heard. The heart sounds recorded are classified according to the clinical conditions confirmed by a specialist. The second set only includes heart sound signals collected with a modern electronic stethoscope (Welchallyn 5079-405 range 20 to 20 KHz) providing higher quality signals (better signal to noise ratio).

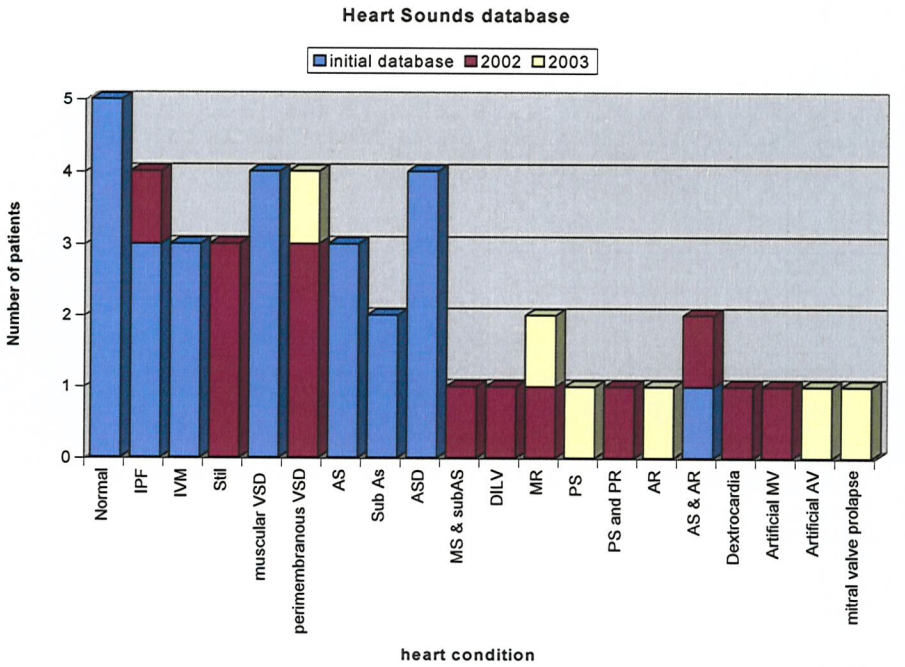


Figure 6.4. Database of Heart sounds. Normal, Innocent Pulmonary Flow Murmur, Innocent Vibartory Murmur, Innocent Still's Murmur, Muscular Ventricular Septal Defect, Perimembranous Ventricular Septal Defect, Aortic Stenosis, Sub Aortic Stenosis, Mitral Stenosis and Sub Aortic Stenosis, Double Inlet Left Ventricle, Mitral Regurgitation, Pulmonary Stenosis, Pulmonary Stenosis and Pulmonary Regurgitation, Aortic Regurgitation, Aortic Stenosis and Aortic Regurgitation, Dextrocardia, Artificial Mitral Valve, Artificial Aortic Valve, Mitral Valve Prolapse.

Figure 6.4 summarises the content of the database, the abscissa shows the categories corresponding to the medical conditions, whereas the ordinate defines the number of patients for each condition. The number of patients rather than the number of recordings is considered since a higher variability is expected to be found between subjects than between recordings of the same subject.

6.4.5 Training the HMM

Initially, to evaluate the proposed HMM based segmentation scheme only the set of normal PCG signals, the simplest case, was considered. The initial database, which contains PCG signals and simultaneous ECG recordings from five paediatric patients with normal structural and functional heart conditions, was used since an external reference is available to evaluate the performance of the segmentation algorithm. Table 6.5 presents a summary of the signals. In three cases PCG signals were recorded from two auscultation locations namely: the pulmonary and the mitral area. The first heart sound is best heard near to the apex whilst in this position the second sound appears diminished, the mitral auscultation area, located at the apex, is preferred for S1 recordings. Conversely, the pulmonary area is recommended for recordings of the second heart sound, where S1 is less intense and the splits of S2 are more noticeable (MSD, 2004). In all patients only the signals from the pulmonary auscultation area were used for the test.

Considering the size of the database, the leave-one-out method was selected to define the training and testing sets (Rangayyan, 2002). Therefore, trials were run using four out of the five PCG signals for the training stage and one for testing the segmentation algorithm until all the available signals were tested.

Each PCG time signal was normalised to have unit amplitude at its maximum point, and for each trial, five heart cycles of the four PCG training signals were

manually segmented, their first cepstral coefficient was computed and presented to the training algorithm.

Patient	Age	Condition	HR	Auscultation	Length
N1	8 yrs	Normal	74 bpm	Pulmonary	30 s
				Mitral	30 s
N2	12 yrs	Normal	78 bpm	Pulmonary	30 s
				Mitral	30 s
N3	7 yrs	Normal	86 bpm	Pulmonary	30 s
N4	2 yrs	Normal	102 bpm	Pulmonary	30 s
N5	11 yrs	Normal	60 bpm	Pulmonary	30 s
				Mitral	30 s

Table 6.5. Normal PCG Signals. Fs=4 KHz

For the two states to be trained, five cardiac cycles for each patient were considered representative since a low intra-subject variability is expected compared to a higher inter-subject variability. Nevertheless, first cepstral coefficient of the whole length of the time signal to be tested (30 seconds) was presented to the hidden Markovian model to assess the performance of the segmentation process[1]. The evaluation was based on the identification of the main components of the PCG cycles.

Notice that since only five normal patient's signals are available, if as suggested in the literature (Rangayyan, 2002), the number of samples required for a classification method is generally expected to be five times or more than the number of features used, only one cepstral coefficient was to be used. If more data were available it would be recommended to assess the performance of the

[1] Consequently, for the first trial 30 seconds of N1 (at 74 ppm) were used for testing (37 heart cycles), and 5 heart cycles of N2, N3, N4, and N5 (20 heart cycles in total) were used for training. For the second trial 30 seconds of N2 (at 78 ppm) were used for testing (39 heart cycles), and 5 heart cycles of N1, N3, N4 and N5 were used for training, and so on.

algorithms using more coefficients, but since for a fixed number of samples the addition of more features will eventually lead to a poorer performance, a balance has to be found. Nevertheless, keeping the number of coefficients low is advised to avoid an overly detailed representation.

Table 6.6 shows the values obtained for the HMM parameters on different trials. The model shows a higher probability to start in state 2 (corresponding to either the first or the second heart sound) since it was trained with segmented PCG starting with the first heart sound. The state transition matrix shows that self-transitions have a higher probability that change of states. A Gaussian probability density distribution was selected to represent the state-observations probability distributions. This distribution is completely defined by the first and second order statistics, and consequently, the values for the mean and the variance are shown in the table.

Test Signal	Training Signals	Initial Prob. $\prod_1$ $\prod_2$	Sate transitions S_{11} S_{12} S_{21} S_{22}		Mean Mu_1 Mu_2	Variance $Sigma_1$ $Sigma_2$
N1	N2,N3, N4,N5	0.0847 0.9153	0.9813 0.1545	0.0187 0.8455	-0.0889 0.4602	0.0351 0.1230
N2	N1,N3, N4,N5	0.1411 0.8549	0.9080 0.2216	0.0920 0.7784	-0.1112 0.2173	0.0319 0.1492
N3	N1,N2, N4,N5	0.2424 0.7579	0.9234 0.2019	0.0766 0.7981	-0.1036 0.4932	0.0430 0.1030
N4	N1,N2, N3,N5	0.1740 0.8260	0.9133 0.2246	0.0867 0.7754	-0.0928 0.5025	0.0381 0.1036
N5	N1,N2, N3,N4	0.1104 0.8896	0.9065 0.2242	0.0935 0.7758	-0.0931 0.4848	0.0401 0.1101

Table 6.6. Parameters for the HMM obtained during the leave-one-out training process. $\prod_x$ Initial state distribution vector, S_{xx} state transition probabilities, Mu_x and $Sigma_x$ are the mean and variance that characterize the Gaussian observation symbol distribution. 1 S1 or S2, 2 systolic or diastolic cycle

Given a training set of cepstral coefficients derived from normal phonocardiograms the parameters of the HMM were computed for each trial by means of a learning algorithm.

For instance, signals N2, N3, N4, and N5 are used in the first trial to iteratively adjust the two probability functions of the HMM (training) so as to maximize the likelihood that the sequences of values for the first cepstral coefficient could be produced by that model.

Once the model parameters are obtained, the test PCG signal N1 was used to compute the most likely sequence of states generating the sequence of observations (decoding), and therefore to segment the signal, since each state defines different portions of the PCG.

This sequence of states is obtained using the Viterbi algorithm (Rabiner, 1989) by mean of the forward probabilities, i.e. the joint probability of obtained a certain value for the cepstral coefficient, and being in a certain state (considering the transition from the previous one).

The first two values for the forward variable (one for each state) are obtained using the initial probabilities and the state observation probabilities computed from both Gaussian probability density functions. The highest forward probability value also defines the most probably state.

For the next cepstral value, the two forward probabilities are computed using the previous forward probabilities, the states transitions, and the state observation probabilities. Again, the highest forward probability will define the most probably state for this point.

Each new computation of the forward variable accounts for the previous ones. Consequently, when the last point of the cepstrum values sequence is reached, the

maximum forward probability represents the Viterbi probability for this particular HMM of producing the whole observation sequence; i.e. of generating this cepstral pattern. Back-tracking the maximum forward variables it is possible to find the most likely state sequence emitting this pattern.

The hidden states can be related to the PCG time signal through the cepstral coefficients, and since the first state represents the occurrence of either the first or the second heart sound, using this HMM it is possible to identify those events in the PCG.

The intervals between occurrences of the first state will differentiate a systolic period from a diastolic one. The former is considered relatively constant whilst the later is considered larger and with more variability. Starting from the largest, the intervals are examined forwards and backwards to maintain consistency with systolic and diastolic cycles.

Details of the algorithms for training and decoding can be found in Chapter 4 and appendix A.

6.4.6 Results and discussion

Figure 6.7 shows the graphical result of the segmentation of a normal PCG using the proposed hidden Markovian model.

The first graph shows a normalised PCG signal from an eight year old patient recorded in the pulmonary auscultation area. Marks for the first and second heart sound components for a single cardiac cycle are included for reference. The most probable sequence of state changes as obtained form the hidden Markovian model is depicted in the second graph. The first state (labelled 1 in the y-axis) represents either the systolic or diastolic cycle, whilst the second state (labelled 2 in the y-axis) represents either the S1 or S2.

The input feature for the model namely, the evolution of the first cepstral coefficient, is depicted in the third graph. Although the algorithms does not require external signals for the segmentation process, the fourth graph shows the ECG signal only as a reference to validate the identification of S1 and S2.

Vertical lines have been added to the graphs to relate the occurrence of the time-sample events in the PCG and ECG signals, and the time-window sequence used to represent the state changes and the progression of the cepstral coefficient. Around 0.3 s for instance, a change from the first state to the second represents the transition from either the systolic or diastolic cycle to either the first or second heart sound.

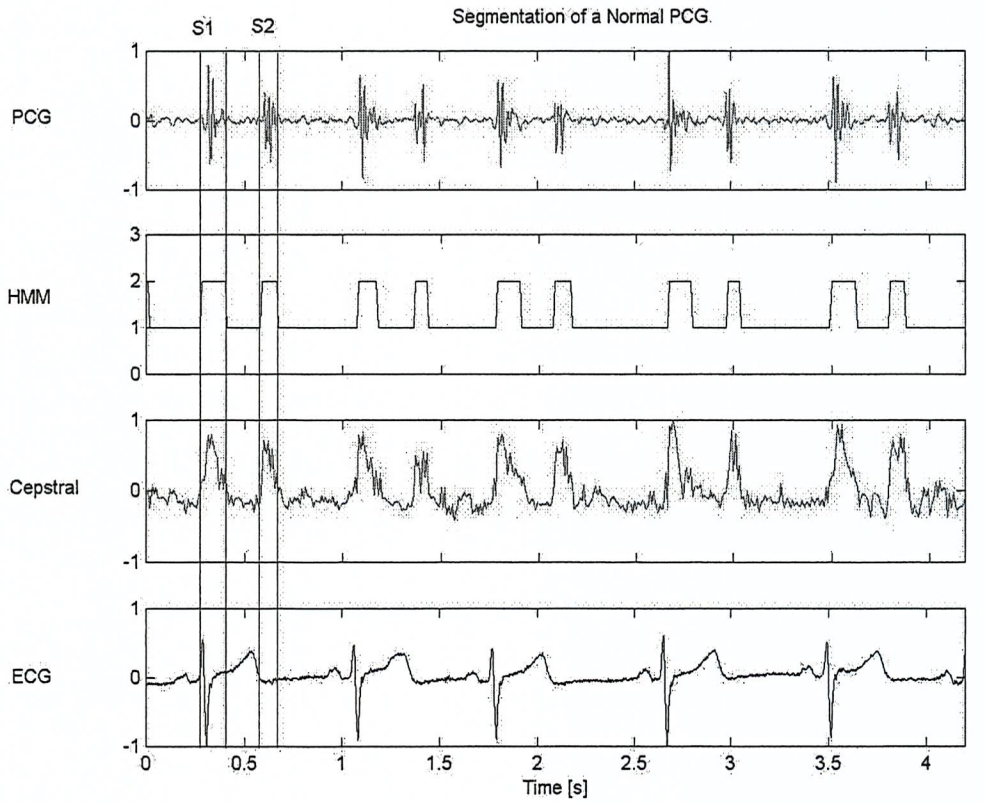


Figure 6.7. PCG Segmentation Using Hidden Markovian Models.

Considering the correct identification of the state representing either S1 or S2 specific events can be defined to assess the performance of the algorithm using standard diagnostic test (Petrie, 2000; Rangayyan, 2002).

- A true positive (TP) is obtained when S1 or S2 appear in the PCG and the HMM correctly displays the corresponding state. In Figure 6.7, the second state represents the heart sounds S1 or S2.

- A true negative (TN) represents the case when neither S1 nor S2 appear in the PCG and the HMM is in the state representing either the diastolic or the systolic cycle. The first state is used in Figure 6.7 to represent the silent periods in the PCG signal.

- A false positive (FP) is defined as the situation when either S1 or S2 is detected although they are not present in the PCG. This situation may arise from a signal imbedded in noise for example and represents the inclusion of an additional sound detected. Figure 6.8 shows an example of a false positive detected in a normal PCG. The additional detection is bounded by the vertical lines drown in the Figure. Notice that the previous S2 was correctly identified, and that the additional sound detected has a similar magnitude.

- A false negative (FN) is defined as the case where either a first or a second sound is present in the PCG but it is not detected by the HMM. In this case, the model would remain in the state representing either the diastolic or systolic cycle. Figure 6.10 shows an example of this case.

Following this convention, for the first five cycles in Figure 6.8, commencing at the first S1, the signal has the following features:

- Ten true positives: all the s1 and S2 sounds correctly identified.
- One false positive: the additional signal detected around 1.5 s.
- Nine true negatives: corresponding to the systolic and diastolic cycles correctly identified.

- Zero false negatives since there are no S1 or S2 sounds missed.

The same criterion is applied to the normal PCG signals available and the results are summarised in table 6.9. Following the leave-one-out procedure, to test the signal N1 for example, the parameters obtained for N2, N3, N4, and N5 are used.

Since each PCG is divided into S1, S2, diastolic and systolic cycles, four events were considered for each of the cardiac cycles of the five patients; and therefore, eight hundred and four samples are considered for the evaluation.

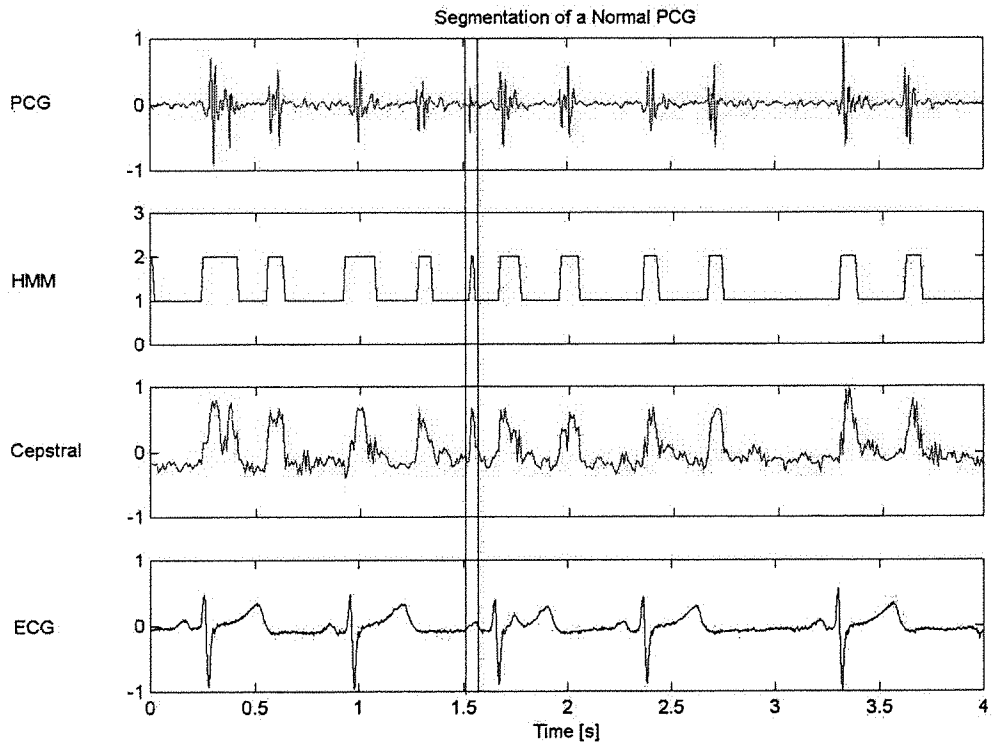


Figure 6.8. False Positive Event Detection.

Heart beats	No of samples	FP	TN	TP	FN	S+	S-
201	804	21	381	399	3	99.25%	94.77 %

Table 6.9. Evaluation Results. FP false positives, TN true negatives, TP true positives, FN false negatives, S+ sensitivity, S- Specificity.

As shown in table 6.9, the algorithm presented excellent values for both sensitivity and specificity (these terms and their significance are further explained in the next chapter). Consequently, the algorithm is highly efficient in detecting the presence or absence of either the first or the second sounds. Time rules used to differentiate systolic cycles from diastolic can be used to correct false positive detection cases.

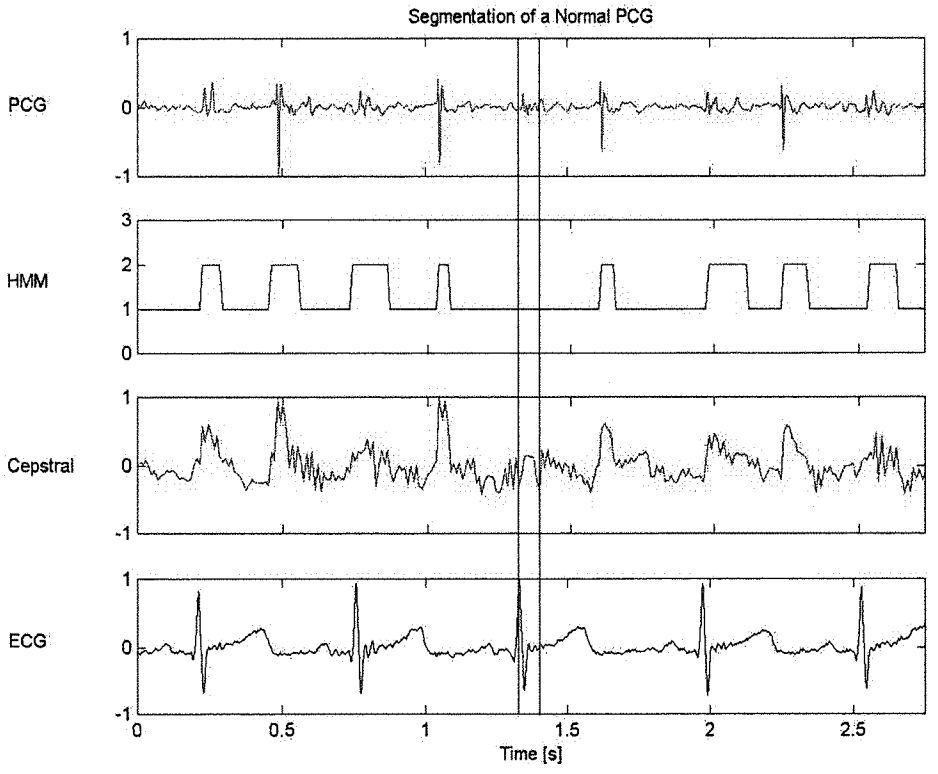


Figure 6.10. False Negative Event Detection.

6.4.6.1 Segmentation of abnormal PCGs using a HMM model of normal signals

Systolic and diastolic murmurs provide essential information in the finding and evaluation of many common heart diseases; however, abnormal cardiac conditions also introduce additional sounds and changes in other components of the PCG. The wide split of S2 in atrial septal defects or the accentuation of the third heart sound in some ventricular septal defects are good examples. These changes can be

of major diagnostic importance when differentiating certain types of murmurs through auscultation. For example, the abnormal split of S2 distinguishes aortic stenosis from innocent pulmonary flow.

Consequently, pathological conditions are not only characterised by the presence of murmurs but also reflected as changes in S1 and S2. Therefore, using a model for normal S1 and S2 states and normal systolic and diastolic states to identify their equivalents in abnormal PCG signals may present some difficulties. When a strong murmur is present in the systolic, the diastolic cycles, or both, the model for the state representing corresponding to normal silent cycles does not hold and consequently, poor performance of the algorithm is expected.

In these cases, the best approach to perform the signal segmentation would be to use hidden Markovian models trained with signals representative of similar pathological conditions, but this would require one to classify the murmur prior to performing the segmentation. This apparent contradiction implies that models of the whole cardiac cycle for different pathologies can be used to classify and segment the signal simultaneously.

Nevertheless, for conditions presenting either soft murmurs, minor S1 and S2 modifications, or both, the models trained using normal signals could be expected to perform satisfactorily. To assess the validity of this hypothesis, the hidden Markovian model obtained for segmentation of normal PCG signals was also tested using PCG signals characteristic of pathological conditions.

PCG signals representative of different medical conditions were presented to the hidden Markovian model trained with the five normal PCG signals. Only the recordings which also included simultaneous ECG signals were considered. Only five heart cycles of each of the normal signals were used for training, whilst 30 seconds of each of the abnormal signals were used for testing. Table 6.11 presents a summary of the signals.

Figures 6.12, 6.13, 6.14, and 6.15 present results for the segmentation of PCGs corresponding to atrial septal defect, innocent pulmonary flow, valvular aortic stenosis, and muscular ventricular septal defect conditions.

Figure 6.12 presents results for the segmentation of a PCG containing an atrial septal defect murmur. Notice that due to the state self transitions of the HMM, despite the wide fixed splitting of S2 characteristic of large defects and a high heart rate of the two years old patient (110 bpm), the algorithm continues to correctly identify S1 and S2.

Patient	Age	Condition	HR	Auscultation	Length
A1	8 yrs	ASD	80 bpm	Pulmonary	30 s
A2	2 yrs	ASD	110 bpm	Pulmonary	30 s
A3	16 yrs	ASD	84 bpm	MLSE	30 s
A4	2 yrs	ASD	122 bpm	Pulmonary	30 s
I1	6 yrs	IPF	88 bpm	Pulmonary	30 s
I2	17 yrs	IPF	70 bpm	Pulmonary	30 s
I3	11 yrs	IPF	52 bpm	Pulmonary	30 s
As1	14 yrs	As	74 bpm	Aortic	30 s
As2	17 yrs	As	60 bpm	Aortic	30 s
As3	8 yrs	As & Ar	74 bpm	Pulmonic	30 s
As4	11 yrs	As	84 bpm	Aortic	30 s
As5	9 yrs	Sub As	76 bpm	LLSE	30 s
As6	12 yrs	Sub As	66 bpm	Aortic	30 s
V1	8 yrs	VSD	82 bpm	Pulmonary	30 s
V2	8 yrs	VSD	102 bpm	LLSE	30 s
V3	8 dys	VSD	126 bpm	LLSE-Apex	30 s
V4	19 yrs	VSD	78 bpm	MLSE	30 s

Table 6.11. Subset of abnormal PCGs. ASD atrial septal defect, IPF innocent pulmonary flow, As aortic stenosis, Ar aortic regurgitation, VSD ventricular septal defect. Fs = 4 KHz.

The PCG signal characteristic of an innocent pulmonary flow murmur is shown in Figure 6.13. Since the hyper-flow condition occurs in an anatomically and physiologically normal heart, the model trained with normal signals correctly identifies the first and second heart sounds. As explained in Chapter 2, the signals for this condition are normally acquired in the pulmonary area. The fact that the normal signals used for the training of the HMM were recorded in the same auscultation area contributes to the excellent performance of the segmentation algorithm.

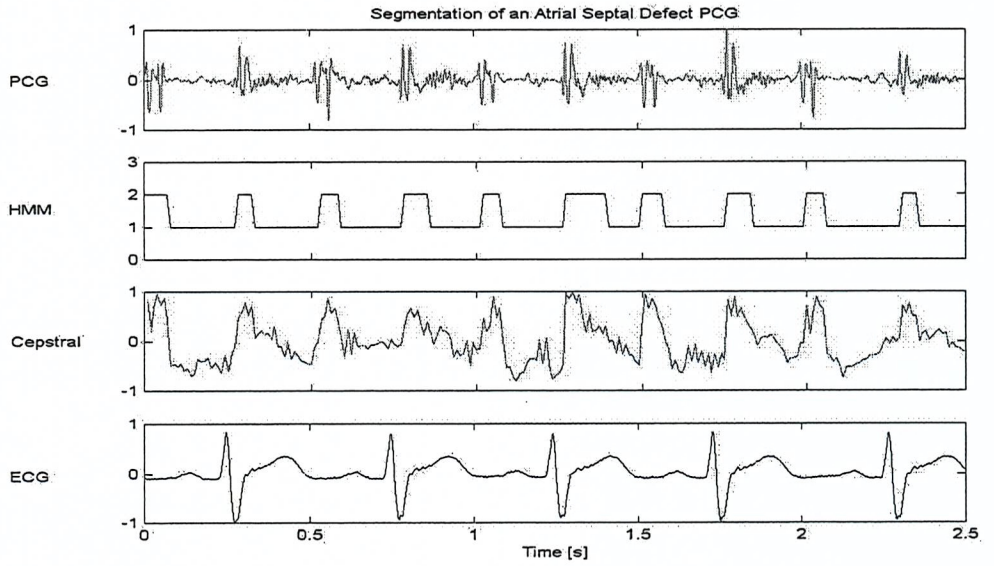


Figure 6.12. Segmentation of a PCG Characteristic of an Atrial Septal Defect Condition.

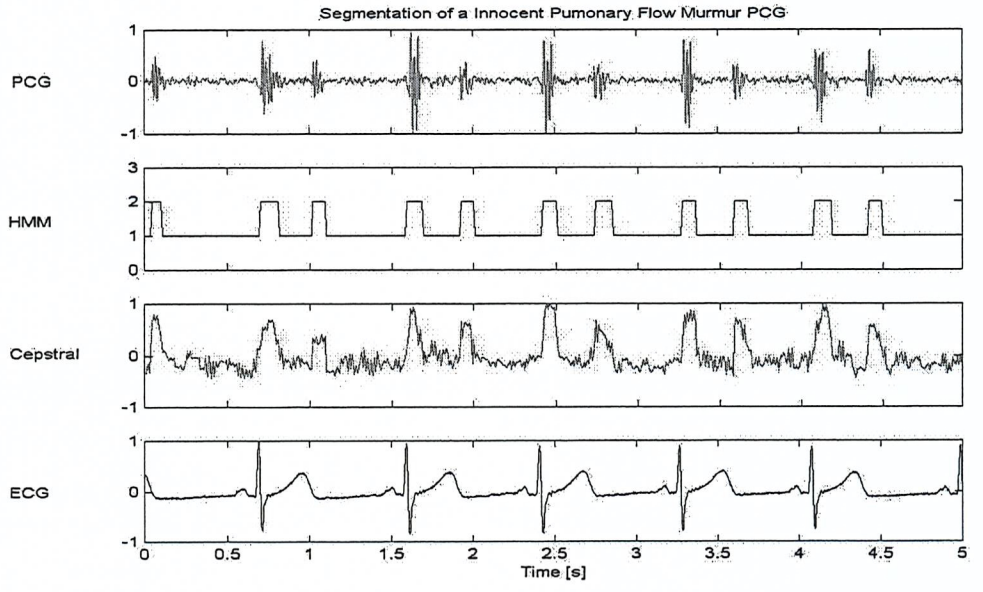


Figure 6.13. Segmentation of an Innocent Pulmonary Flow Murmur PCG.

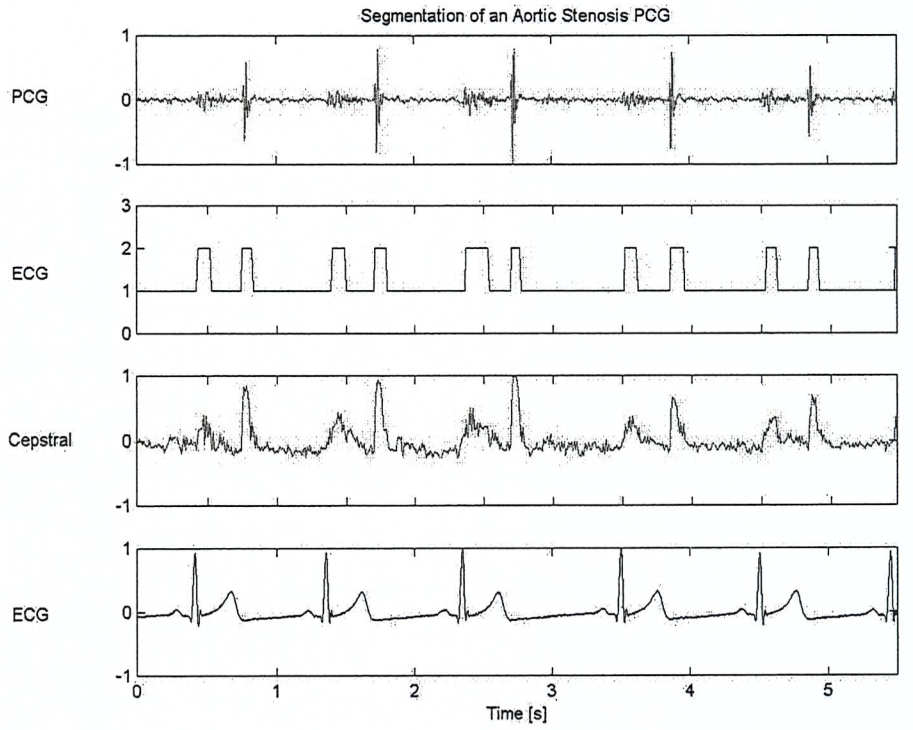


Figure 6.14. Segmentation of a PCG Characteristic of an Aortic Stenosis Condition.

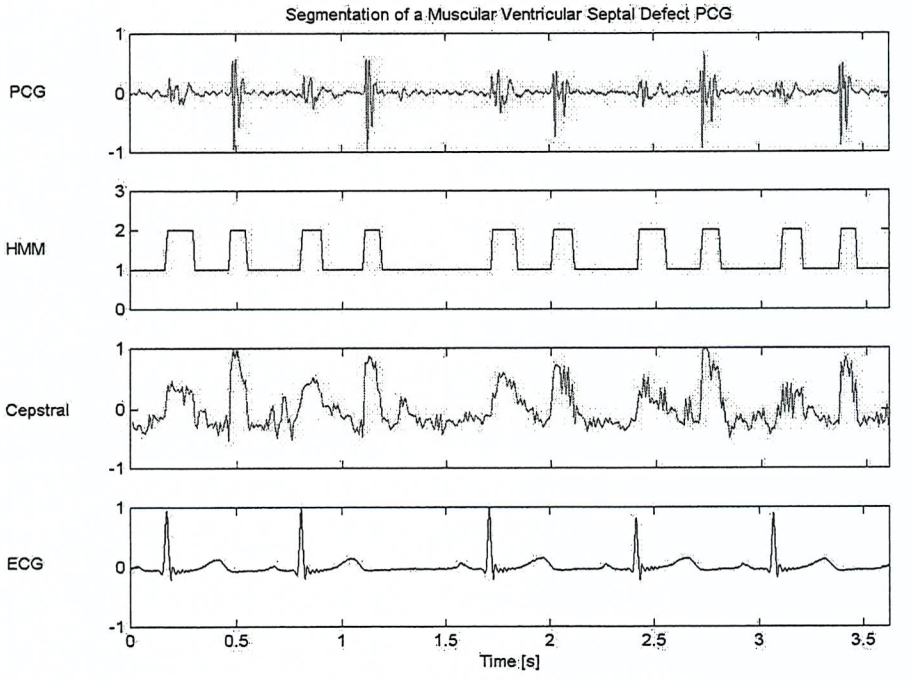


Figure 6.15. Segmentation of a PCG with a Soft Ventricular Septal Defect Murmur.

Figure 6.14 shows the segmentation of a PCG with a murmur due to an aortic stenotic valve. From the figure, it can be seen that the hidden Markovian model trained with normal heart sounds at the pulmonary area is able to detect the main components of the signal with a mild systolic murmur recorded in the aortic auscultation area.

Figures 6.15 and 6.16 show PCG signals with systolic murmurs reflecting ventricular septal defect conditions. The former shows a soft murmur recorded in the pulmonary area whilst the later, shows a strong murmur recorded in the LLSE auscultation area. The algorithm is able to identify the first and the second heart sound in the first case. However, in the second case, the strong murmur in the systolic cycle is considered as part of the first sound as shown in the first two cycles, or completely missed leading to a single state comprising the first sound, the systolic cycle, and the second sound.

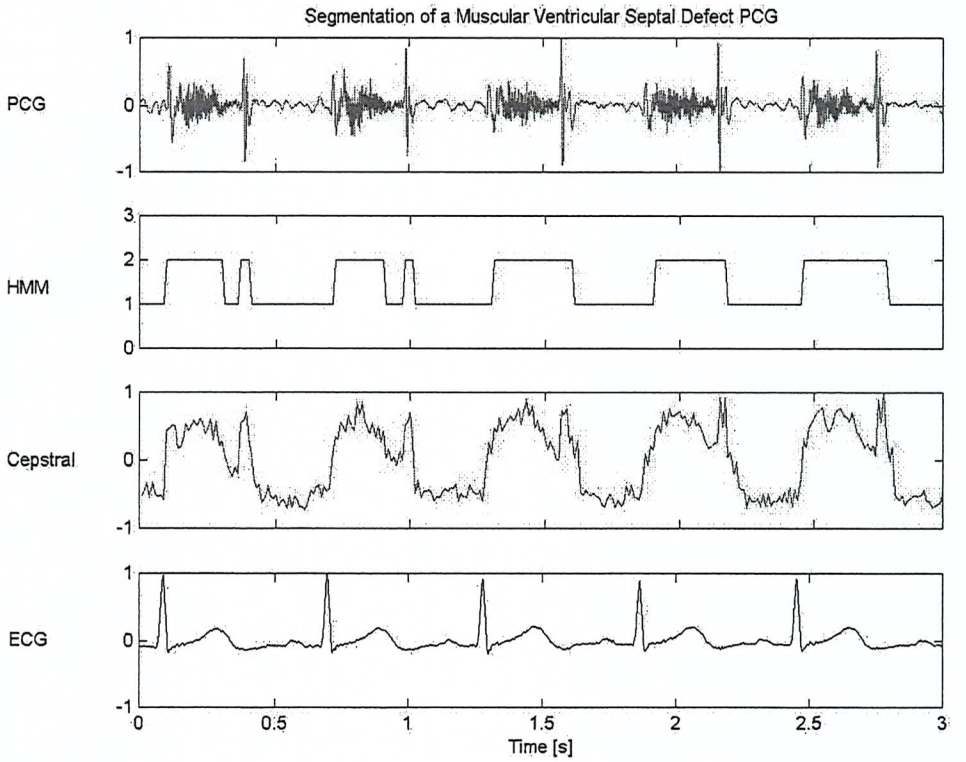


Figure 6.16. Segmentation of a PCG with a Strong Ventricular Septal Defect Murmur.

Table 6.17 summarises the performance of the segmentation algorithm for each condition and the over all performance. The model trained using normal PCG signals presents excellent results for the segmentation of PCGs representative of conditions containing mild or soft systolic murmurs and a decreased performance in the case of pathological conditions presenting very strong murmurs.

Condition	Heart beats	No of samples	FP	TN	TP	FN	S+	S-
ASD	198	792	11	385	382	14	96.46%	97.22%
IPF	105	420	2	208	207	3	98.57%	99.04%
VSD	194	776	5	383	324	64	83.50%	98.71%
AS	217	868	8	426	416	18	95.83%	98.15%
Abnormal	714	2856	26	1402	1329	99	93.06%	98.17%

Table 6.17. Evaluation Results for Abnormal PCGs. ASD atrial septal defect, IPF innocent pulmonary flow, VSD ventricular septal defect, AS aortic stenosis, FP false positives, TN true negatives, TP true positives, FN false negatives, S+ sensitivity, S- Specificity.

6.4.7 Noise in Phonocardiograms

The robustness of the algorithm to noise in the phonocardiogram signals can be shown in figure 6.18. This phonocardiogram was recorded at the apex auscultation area and is know to correspond to a patient presenting a bicuspid aortic valve condition. The murmur in the PCG can not be seen from the recording, neither the first nor the second heart sound can be clearly distinguished. More over, an ECG reference signal is not available. However, despite the noise and the clipping of the signal, the main components S1 and S2 are audible and can be recognized.

Figure 6.18 shows the noisy PCG, the state transition of the HMM, and the first cepstral coefficient representation. By comparison to all the previous figures presented in this chapter the reader may be able to detect the main components of the PCG in the cepstral representation.

An ECG auxiliary signal was not available to visually confirm the segmentation. However, each portion of the PCG was played in order verify the process. The segments seem to correspond to the events defined for S1 and S2.

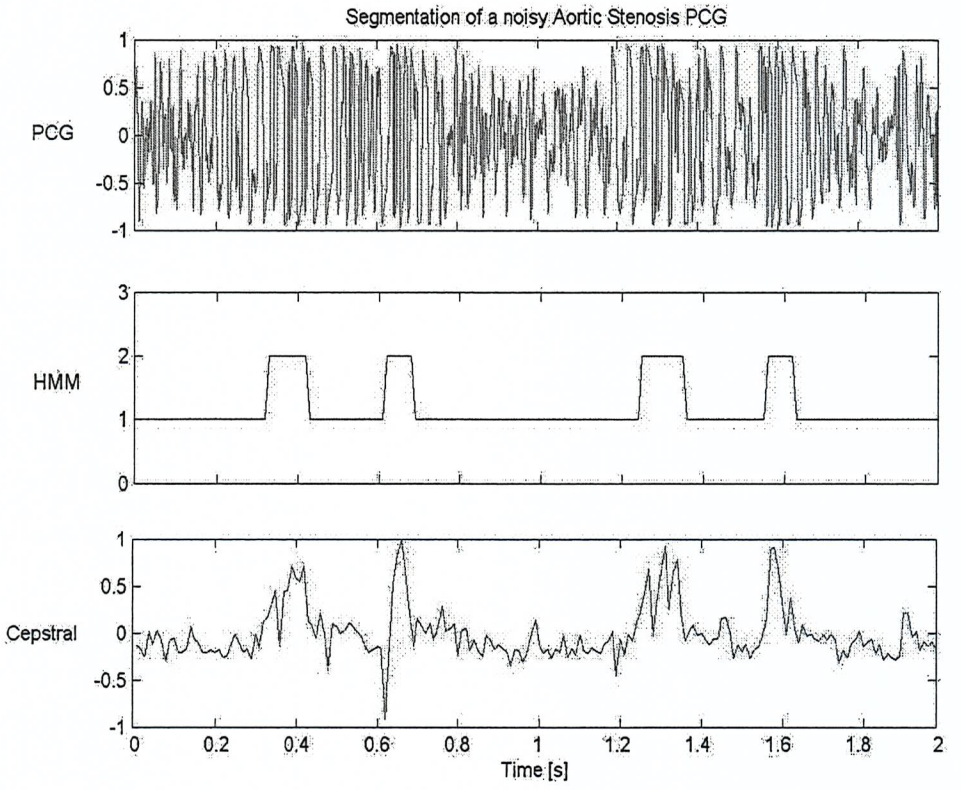


Figure 6.18. Segmentation of a noisy PCG.

In some cases low pass filtering of the signal before segmentation may help since murmurs contain relatively high frequencies, but in some others it may not be ideal. Figure 6.19 shows the same PCG, but in this case the signal was filtered using a butterworth 2nd order filtered with a cut off frequency of 70 Hz since the heart sound spectral distributions of S1 and S2 show predominant peaks below70 Hz (Rangayyan, 2001, pp 279; Yoganathan, 1976).

The figure shows in descending order the noisy PCG signal, the state sequence of the HMM, the first cepstral coefficient of the low pass filtered signal, and the low pass filtered PCG signal itself.

Despite the low pass filtering, there was a lot of noise in the signal, and the HMM state sequence obtained using this first cepstral coefficient representation was not able to identify the S1 and S2 components properly.

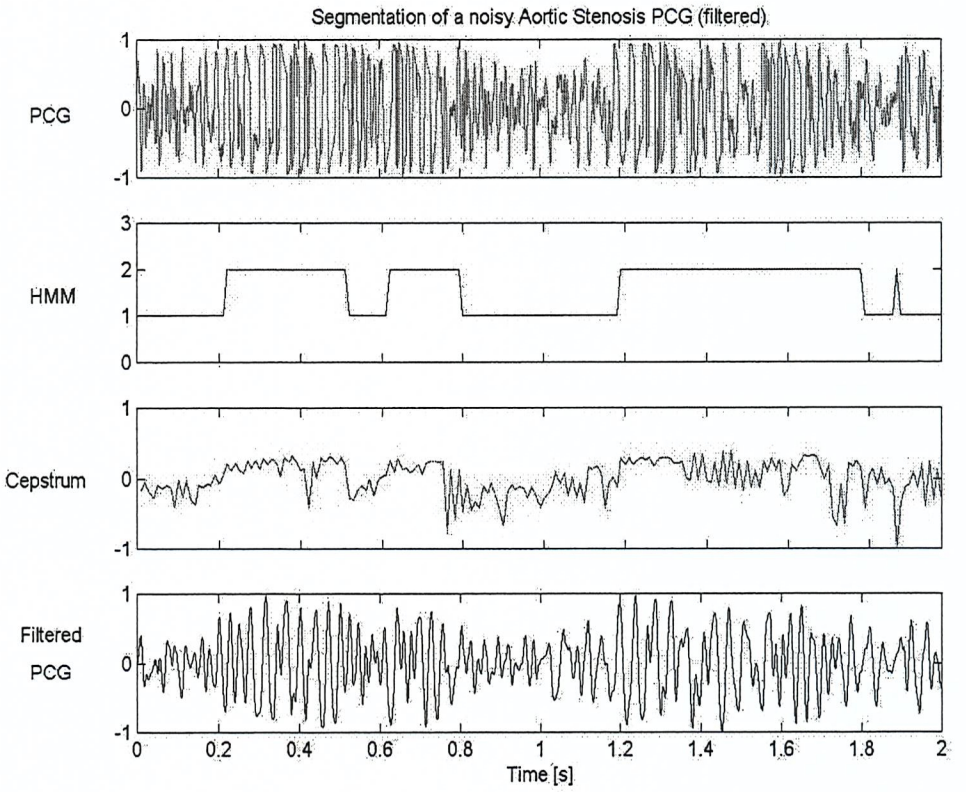


Figure 6.19. Segmentation of a noisy PCG (low pass filtered).

6.5 Summary

A systematic approach to the study of heart sounds demands the identification of the main components of the phonocardiogram signal. Cepstral coefficients and hidden Markovian models were proposed for the PCG segmentation.

This simple statistical model proves to be effective to account for the PCG intersubject variability within the normal heart sound set, and provides accurate detection even for the abnormal condition sets of PCG containing low and medium grade murmurs. In the case of high-grade murmurs the intensity of the

murmur may obscure the recognition of the first or second heart sound in the first cepstral coefficient. This can be solved either by using a higher number of cepstral coefficients in which the S1 and S2 component are clearly identified, as shown in figure 5.3c, by adding time constraints to the model, or perhaps in some cases by filtering off the high frequency murmurs from the PCG.

Considering that for the abnormal heart sounds the HMM was trained with 25 heart beats from 5 patients with normal PCGs, and tested with 714 heart beats from 17 patients representing 4 medical conditions, the high values obtained for sensitivity and specificity are promising. Nevertheless a study involving more patients and more conditions would be necessary for the evaluation to be conclusive. It is also recognised that more complex models may lead to better results although the added complexity demands more data.

The next chapter will present the use of hidden Markovian models for the classification of systolic murmurs related to the most common pathologies found in paediatric patients. The algorithms developed are aimed to implement an automatic system to aid diagnosis.

7 Systolic Murmur Classification by HMM

7.1 Introduction

Auscultation of paediatric patients using passive and non-invasive techniques is of prime importance due to the frail nature of infants. Moreover, the implications of referral for further analysis in the case of unrecognised, non-pathological heart sound murmurs in paediatric patients demand further development of passive methodologies. One of the directions to follow consists of the development of autonomous systems able to provide valuable support for medical diagnosis.

The autonomous identification of pathologies through the analysis of heart sounds can be regarded as a classification problem where the sound to be identified can be assigned to one of several known categories representing the associated pathologies. This approach is especially valid to aid differential diagnosis where signals are perceived as having similar acoustic patterns and the physiological findings are not conclusive, such as in the case of differentiation between innocent and pathological murmurs in early conditions where the patient may be otherwise asymptomatic.

The previous chapter has shown how HMM can be used to perform what is regarded as the first step towards the development of an autonomous system for classification, namely the segmentation of a PCG signal into its main components. The relative time appearance of the components provides essential information to discriminate between pathologies. Furthermore, once the main components have been identified, each segment of the signal can be analysed and modelled in more detail.

The heart sounds representative of the most common abnormal heart conditions found in paediatric patients, namely aortic stenosis, atrial septal defect, ventricular septal defect, and innocent pulmonary flow murmurs are all associated with characteristic murmurs present in the systolic cycle of the PCG. Consequently, by

analysis of the systolic segments of the PCGs, it may be possible to identify the associated pathologies and to differentiate between the pathological and non-pathological conditions.

As mentioned in Chapter 3, HMM is a suitable technique for the classification of phonocardiograms. Hidden Markovian models provide a strong mathematical basis to determine the probability with which a pattern could have been generated from a model of a specific category. Consequently, by assigning the pattern to the class providing the highest probability, a PCG signal can be identified.

Following this approach, a HMM would be trained for each one of the five medical conditions under study. To classify a new PCG, the probability of each HMM to have generated this new sequence would be computed and the signal would be classified according to the highest score. However, if the new PCG does not belong to any of these classes it would be wrongly assigned to the closest match. Nevertheless, in clinical practice, other physiological findings narrow the number of possible classes. This algorithm is designed to aid the specialist (not to substitute him/her) in differential diagnosis for specific medical conditions that present similar acoustic signals and similar physiological findings.

The following section will show how systolic murmurs, characteristic of specific heart conditions, can be identified by using HMM.

7.2 Selection of the HMM

As stated by Rabiner (1989) there is no simple, theoretically correct, way of selecting the model, or to choose the initial parameters. Nevertheless, as a general rule for a finite training set, a simple model requires fewer training sequences to obtain a good estimate of the model parameters.

The medical conditions in this study are associated with characteristic murmurs in the systolic cycle of the PCG, or by the absence of any in the case of normal

conditions. Therefore, as a first approach only the systolic segments of the PCG are going to be modelled by the HMM. A model comprising all components of the heart sound would be more appropriate for the classification process since the pathologies are reflected not only by the presence of murmurs but also by changes in the morphological characteristics of other components of the phonocardiogram signal. Nevertheless, a model of the whole PCG signal requires more utterances to capture the particularities of each pathology. Consequently, a detailed model of the systolic cycles is preferred over a more general model of the cardiac cycle to overcome limitations imposed by the size of the database.

The systolic signals are chosen to be modelled by a left-right HMM. This structure is selected for the characterisation of temporal or sequential structures since time may be visualised as having a direction from left to right. Self-loop transitions are allowed to account for variations in the duration and time scales of the signals in each state. Since a left-right model has been chosen, the initial state probability for state one is set to one, and therefore, every time a new set of observations representing a systolic cycle is evaluated the model is reset to this state.

Considering the benefits of keeping the model to the minimum number of parameters, a simple model for the systolic cycles consisting of only three hidden states is proposed. From a morphological point of view the states are selected to represent the onset, a steady state for the murmur and an exiting transition, which seems suitable to model ejection murmurs. From a time-reference point of view, the same structure may be intuitively related to the time occurrence of the murmur relative to the systolic cycle. This latter approach is useful to model pansystolic murmurs in which no marked onset and exiting transitions are exhibited. For ejection murmurs the states may be a reference for early, middle, or late ejection murmurs. It is important at this point to recall that the hidden states of the model represent an unknown generation process, and since it is a mathematical model, the duality of the state representation as a morphological or time occurrence based is only an abstraction.

A single Gaussian probability density function has been selected to represent the observation symbol probability distributions for the model in order to use the continuous valued observations (provided by cepstral coefficients, line spectral frequencies, or Maass-Weber filter representations). The univariate Gaussian functions are therefore specified by their mean and standard deviation parameters.

Figure 7.1 shows the representation of the three-state HMM proposed for the analysis of the systolic cycles. State transitions are represented by arrows. In this left-right model the sequence always start in state one, allows for self state transitions and forbids back state transitions.

Typical systolic signals for aortic stenosis and ventricular septal defect are shown in the figure along with their first cepstral coefficient representation. This simplified representation constitutes the set of observations.

The continuous values obtained from the cepstral observations are related to the hidden states by single Gaussian probability density distributions. This is graphically represented by faded shadows of values around the darkest mean value.

7.3 Binary classification: normal versus aortic stenosis classes

The implications of classification decisions made in the context of medical diagnosis go beyond statistical measures of accuracy and validity. Therefore, the effectiveness of a diagnosis technique has to be determined using standardised methods (Rangayyan, 2002). In this context, a simple *screening* test has been developed in order to assess the performance of the hidden Markovian models for murmur classification.

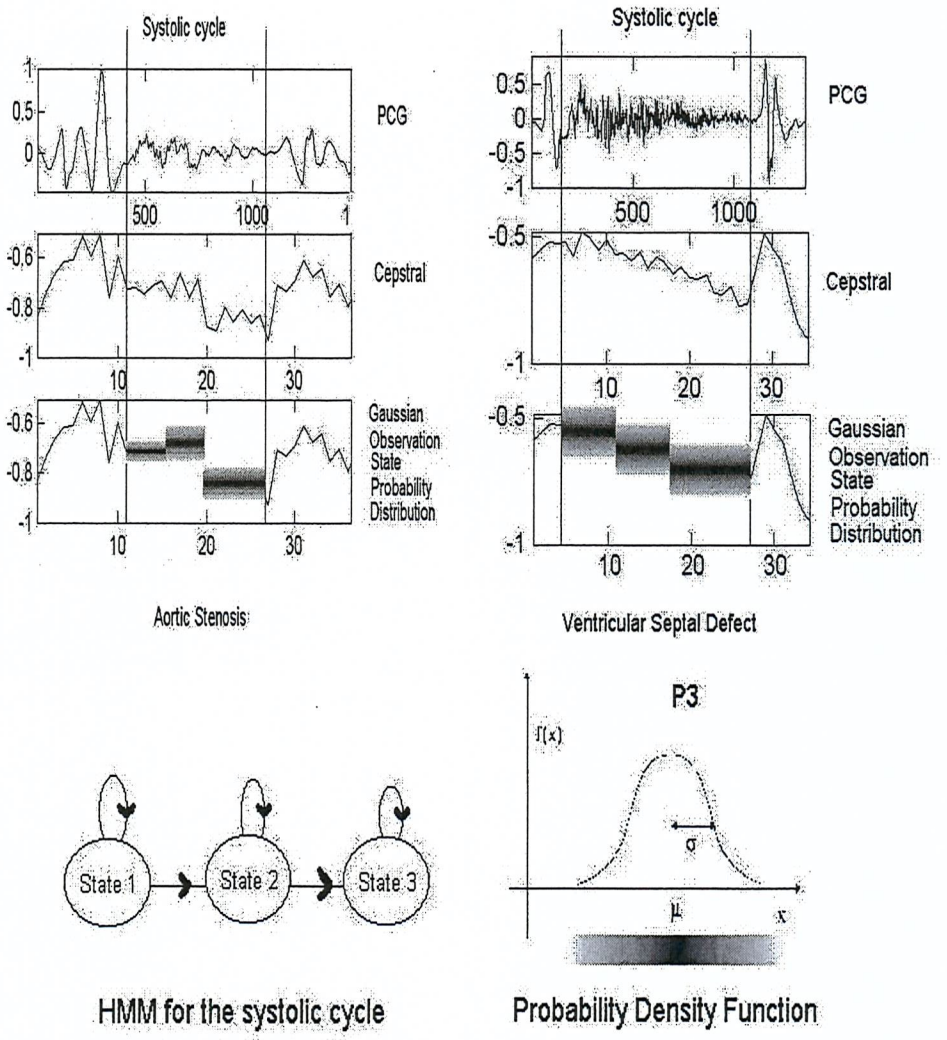


Figure 7.1. Hidden Markovian Model for Systolic Classification of PCGs.

A screening test is useful to detect the presence or absence of a specific disease in a certain study population: the decision to be made is binary. In particular, for systolic murmur classification once a signal is presented to the algorithms the decision has to be made as to whether the signal has the specific cardiopathology represented by the model, or it is more likely to be a PCG representative of a normal heart sound. Consequently, two models have to be constructed: one for

normal heart sounds and one for signals representing the medical condition on test. Considering the size of the database, the pathology represented with the higher number of samples available was selected to be modelled by the HMM, and thus the classes of aortic stenosis and normal sounds were formed.

7.3.1 Training and testing procedure

The normal class set of phonocardiogram described in Chapter 6 was used for the screening test. In the interest of optimising the use of the database, signals from the same patient recorded at different auscultation areas were considered as independent samples, and therefore, eight sample signals were used for the test.

Due to limitations of the database, the aortic stenosis set comprised not only PCG signals representing pure congenital valvar aortic stenosis but also other related pathologies were included. As in the case of normal set, signals from the same patient taken at different auscultation areas were considered as independent samples. Consequently, a total of eight sample signals, obtained from six patients, were contemplated in order to match the size of the set associated with normal class.

The aortic stenosis class was represented by the following signals:

- Two PCGs representative of sub-aortic stenosis for patients aged five and fourteen years old. This is a non-valve related pathology that, although having a similar murmur as congenital aortic stenosis, can be differentiated from the latter through analysis of the PCG (Guadalajara, 1998). One of the indications to differentiate the sub-aortic stenosis from its valvar counterpart is that the epicentre is located in the auxiliary aortic auscultation area. The signals were therefore acquired at the lower sternal edge and the third right intercostal edge instead of the second edge where the primary aortic auscultation area is defined.

- Two PCG signals from a patient aged nine years presenting a congenital aortic stenosis valve recorded at two locations. The characteristic click of this pathology was recorded first at the lower left sternal edge, and secondly the mild systolic murmur was recorded at the aortic area where it is heard best.

- A signal recorded at the pulmonary auscultation area from an eight-year old patient presenting both a systolic murmur due to valvular aortic stenosis and an early diastolic murmur due to aortic regurgitation.

- Two PCGs from a patient aged one year presenting congenital aortic stenosis. The signals were acquired at two different auscultation areas, namely the mitral area for the click and the aortic area for the murmur. This patient had had a previous cord repair.

- A signal containing a systolic murmur characteristic of valvar aortic stenosis and early diastolic murmur representative of aortic regurgitation recorded at the aortic area from a patient aged fourteen years old.

Notice that the best theoretical position to record the characteristic protosystolic click of congenital aortic stenosis is the mitral area at the apex, and that the best position to record the valvar murmur is the aortic area. Nevertheless, some recordings in the database were obtained from different locations, either the lower left sternal edge instead of the mitral area for the protosystolic click, or the pulmonary area instead of the aortic area for the systolic murmur. This represents a sub-optimal sensor placement with the exception of the recordings of sub-aortic stenosis for which the epicentre is not located in the aortic area. The decision nevertheless is based on practical considerations found during the auscultation by the consultant.

The leave-one-method, recommended when the number of available samples is small (Rangayyan, 2002), was selected for the estimation of the classification accuracy of the HMMs. Following this procedure, from the sixteen PCG signals available, eight for each medical condition, only the PCG signal of one patient

was used for each testing event. The rest of the signals, divided in two sets according to its class, were used to train the HMM for each condition.

The training sets are formed of cepstral coefficients of five systolic cycles of PCG signals representing the medical conditions under study. Although the set consisted of eight signals, these were obtained from only five or six patients. Since it is expected to find higher inter-subject than intra-subject variability in the PCG signals, five systolic cycles were considered as a reasonable number to represent the variability among systolic cycles from the same patient. Each cycle of the PCG signal to test was considered as an independent event though, and therefore, eighty systolic cycles were used for the screening test.

To train the hidden Markovian models, uniform values were provided as initial estimates for the parameters and these values were refined using an estimation–maximization algorithm (Rabiner, 1989; Murphy, 2001; Charbit, 1999) to capture the statistical parameters modelling a training set of PCGs for each of the medical conditions.

HMM	Initial Prob. $\prod_1$ $\prod_2$ $\prod_3$	Sate transitions S_{11} S_{12} S_{13} S_{21} S_{12} S_{23} S_{31} S_{32} S_{33}			Mean Mu_1 Mu_2 Mu_3	Variance $Sigma_1$ $Sigma_2$ $Sigma_3$
Normal	1	0.9165	0.0835	0	0.8887	0.0076
	0	0	0.9427	0.0573	0.7096	0.0013
	0	0	0	1	0.7845	0.0026
Aortic stenosis	1	0.8024	0.1976	0	0.9075	0.0053
	0	0	0.8971	0.1029	0.9751	0.0086
	0	0	0	1	0.7303	0.0026

Table 7.2. Values obtained for the HMM for the normal and the aortic stenosis classes during the training process. $\prod_x$ Initial state distribution vector, S_{xx} state transition probabilities, Mu_x and $Sigma_x$ are the mean and variance that characterize the Gaussian observation symbol distribution.

Testing	Training	Initial Prob. $\prod_1$ $\prod_2$ $\prod_3$	Sate transitions S_{11} S_{12} S_{13} S_{21} S_{12} S_{23} S_{31} S_{32} S_{33}			Mean Mu_1 Mu_2 Mu_3	Variance $Sigma_1$ $Sigma_2$ $Sigma_3$
As1	As2, As3,	1	0.8358	0.1642	0	0.9122	0.0054
	As4, As5,	0	0	0.8584	0.1416	0.8862	0.0058
	As6	0	0	0	1	0.6797	0.0017
As2	As1, As3,	1	0.8866	0.1134	0	0.9071	0.0056
	As4, As5,	0	0	0.8775	0.1225	0.7828	0.0031
	As6	0	0	0	1	0.6653	0.0013
As3	As1, As2,	1	0.6978	0.3022	0	0.9093	0.0055
	As4, As5,	0	0	0.8873	0.1127	0.9657	0.0074
	As6	0	0	0	1	0.7373	0.0043
As4	As1, As2,	1	0.7124	0.2878	0	0.9062	0.0053
	As3, As5,	0	0	0.8895	0.1105	0.8376	0.0056
	As6	0	0	0	1	0.7123	0.0066
As5	As1, As2,	1	0.8259	0.1741	0	0.9193	0.0043
	As3, As4,	0	0	0.8521	0.1488	0.9715	0.0087
	As6	0	0	0	1	0.7462	0.0025
As6	As1, As2,	1	0.7151	0.2849	0	0.9076	0.0052
	As3, As4,	0	0	0.8976	0.1024	0.9817	0.0086
	As5	0	0	0	1	0.7244	0.0030

Table 7.3. Values obtained for the HMM for the aortic stenosis classes for different training sets. $\prod_x$ Initial state distribution vector, S_{xx} state transition probabilities, Mu_x and $Sigma_x$ are the mean and variance that characterize the Gaussian observation symbol distribution.

Table 7.2 shows the parameters of the HMM representing the normal and aortic stenosis classes. Note how the model presented in figure 7.1 is reflected in the

parameters. The model chosen always starts in state 1, and this is reflected in the set of initial probabilities. The model is forced to end in state three, and therefore, the transition probability S_{33} is one. The left-right characteristic of the model is reflected in the fact that all backwards transitions S_{21}, S_{31} and S_{32} are cero.

The transition matrix for the normal signal shows higher self transition probabilities compared with the aortic HMM. The mean value of the normalized cepstral coefficient for the aortic stenosis class shows the same transition pattern as shown in figure 7.1. From a mean value of 0.9075 for state one, to a higher value of 0.9751 for state 2 and afterwards to the lowest value of 0.7303. A higher variance for the second state is probably related to the mid-systolic murmur, whilst the high mean and variance for state one of the normal class is probably related to inclusion of the end of S1 into the systolic cycle.

The values shown in table 7.2 were obtained using all the training signals, but in order to test the algorithm, the leave-one-out method was used. Table 7.3 shows how the parameters change depending on the set of training and testing signals.

Table 7.3 shows the strong dependence that the parameters of the HMM have on the small training set. Given the trained HMMs, the probability for each systolic cycle of being produced for a particular HMM was computed. Note that each HMM has an initial state probability, and therefore, every time a new set of cepstral observations (representing a systolic cycle) is evaluated, the most probable initial state is defined considering this probability.

Each systolic cycle (represented by its cepstral coefficients) was scored individually as successfully identified or otherwise, as misclassified. An overall score for a single patient's PCG was scored from the individual cycles.

The training and testing sets were renewed to obtain another testing event. This procedure was followed until all the sixteen PCG signals were tested.

7.4 Screening test results

In Chapter 5, cepstral coefficients and two alternative methods to extract feature representations namely Maass-Weber representations and line spectral frequencies were presented along with their viability to visually represent the explicit variations in PCG signals typical of specific diseases.

Although all the methods proved to be able to represent these variations, their ability to reflect the key parameters that would simplify the classification process needs to be evaluated within the context of a specific classifier. Some feature representations may be more suitable for certain classifiers than others. Redundancy of the information, parameter interdependence, the compromise between the length and number of parameters required for the representations versus the size of the data-base available, and implicit assumptions of the classifier are some of the factors that contribute to this interdependence between the representations and the classifier performance. Moreover, the classification performance is also highly dependent on the accuracy with which the phenomena of interest is represented in the subspace of sample signals available for training and evaluation.

A technique for phonocardiogram signal analysis and classification has to be evaluated considering both the signal representations and the classifier since they are strongly related.

Neural networks and Maass-Weber filter representations have been used before for the classification of phonocardiograms (Barschdorff, 1995) and results have been published for LPC coefficients used for heart sound analysis and classification (Itawa, 1997).

As explained in Chapter 4, the line spectral features are an alternative representation derived from LPC coefficients, although no studies have been found on the use of line spectral frequencies applied to PCG analysis.

Considering these factors, the performance of the hidden Markovian models as classifiers using cepstral cefficients, line spectral frequencies, and Maass-Weber filters as feature representations will be evaluated in the next section.

7.4.1 Hidden Markovian models and cepstral representation

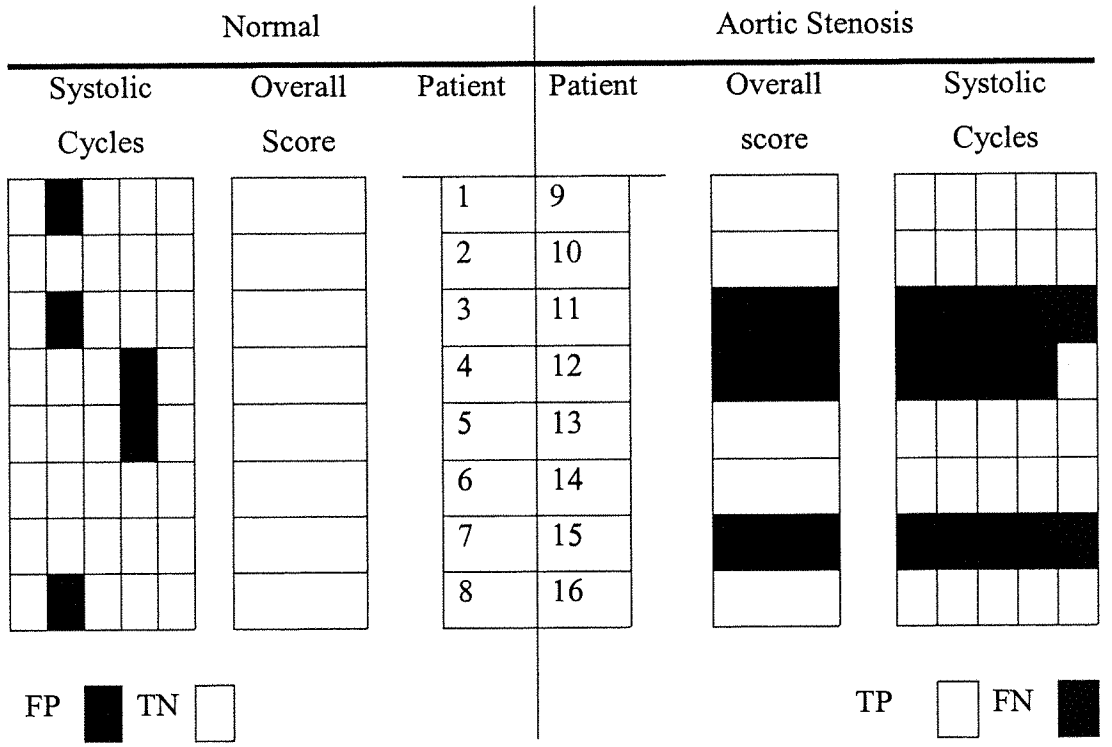


Table 7.2. Results of the Classification using Cepstral Representation. The black boxes represent incorrectly classified signals (false positives Fp for normal samples, false negatives FN for aortic systolic cycles), whereas white boxes represent signals correctly classified (true negatives TN for normal samples, true positive TP for aortic systolic cycles).

Table 7.2 presents the results obtained for the classification of the signals using cepstral representation and the HMMs for aortic stenosis and normal systolic cycles.

The black boxes in the figure represent the signals that were incorrectly classified. For patient one for example, although the signal corresponds to a normal PCG the second systolic cycle was classified as aortic stenosis. Similarly, the black boxes

for patient 11 means that all the systolic cycles were classified as normal which is incorrect.

Table 7.2 is useful to relate the performance of the classifier to the signals evaluated and therefore, it is valuable for identifying probable sources of error and possible improvements.

Analysing the performance of the HMM for aortic stenosis, from table 7.2 it can be seen that the model could not detect any murmur in the PCG of patient 15. Referring to the medical records, the signals from patient fifteen and fourteen are in fact signals from the same person, but recorded in different auscultation areas. The signals of patient fourteen correctly identified were recorded in the aortic area whereas the signals of recording fifteen were recorded in the mitral position. The mitral position is usually preferred to register the ejection click since the murmur is very soft or inaudible in this area, whereas the aortic stenosis murmur is loudest in the aortic area (Brown, 2002). Therefore, the decision of the algorithm to classify as aortic stenosis only the signals recorded in the aortic area may be correct. Nevertheless, for signals 11 and 12, both from the same person and recorded in different auscultation areas, the model could not identify the systolic murmur in the signals recorded in the aortic area.

Figure 7.3 shows PCG for the same patient recorded at different auscultation areas. The PCG signal recorded at the lower left sternal edge exhibits the characteristic click of a congenital aortic stenotic valve, whereas the systolic murmur is better seen at the recordings from the aortic auscultation area.

The results can also be represented using standard definitions for the screening test:

- A true positive (TP) is defined as the case when the systolic segment contains an aortic stenosis murmur and the algorithm acknowledges that the sequence of corresponding observations have been most probably generated by the hidden Markovian model of aortic stenosis. In Table 7.2 for example, the

systolic segments for the PCG signal of patient 9 were all correctly identified as containing aortic stenosis murmurs, and were therefore represented by a white box.

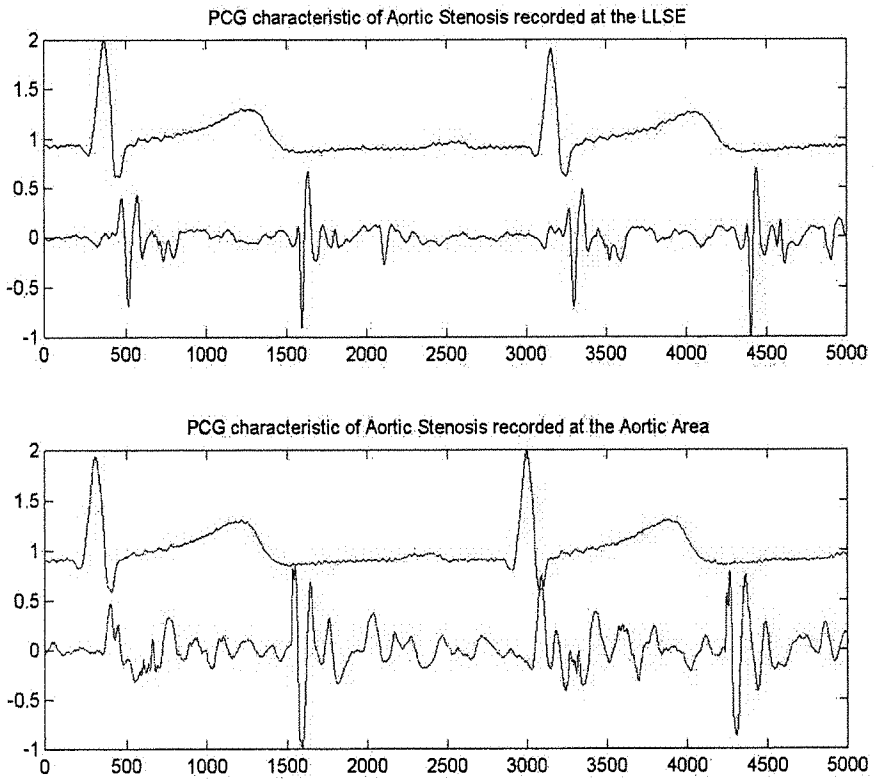


Figure 7.3. PCG Characteristic of a Congenital Aortic Stenotic Valve Recorded at the Lower Left Sternal Edge and at the Aortic Area. Fs = 4 KHZ.

- A true negative (TN) in this case represents the condition where the systolic signal segment contains no murmur and the algorithm classifies the signal as most probably generated by the hidden Markovian model of the normal PCG signals. The white boxes assigned to each systolic cycle of patient 2 exemplify this situation, i.e. the segments correspond to a normal PCG and are assigned to the HMM for normal signals.

- A false negative (FN) for this screening test is the situation when the algorithm assigns to the normal class a systolic cycle which contains a

murmur characteristic of an aortic stenosis condition. This situation is exemplified by the signal of patient 15 in Table 7.2. All the systolic cycles from a patient presenting a form of aortic stenosis were incorrectly classified as normal, and consequently they were represented by black boxes.

- A false positive (FP) occurs when a systolic signal from a normal PCG is incorrectly classified as belonging to the aortic stenosis murmur group. This false alarm is represented by black boxes in Table 7.2 as in the second systolic cycle of the signal from patient 3.

Considering the previous definitions the performance of the classification algorithm can be stated in terms of sensitivity and specificity.

- The sensitivity (S+) of a test represents its capability to detect the presence of the disease. In this case, it provides the proportion of PCG signals containing a murmur of an aortic stenosis condition that are correctly identified by the algorithm. Therefore, it can be defined as the number of true positive decisions divided by the number of signals with the disease.

$$S+ = \frac{TP}{TP+FN}.$$

- The specificity (S-) of a test indicates its accuracy in identifying the absence of the disease of concern. In this case, it provides the proportion of normal PCG signals correctly identified by the test. It can be defined as the number of true negative decisions divided by the number of subjects presenting the normal class sample.

$$S- = \frac{TN}{TN+FP}.$$

The sensitivity and specificity values provide useful indication of whether a test is useful in making diagnosis. Once the test has been performed, the sensitivity and

specificity do not indicate whether a positive result truly means the presence of a disease. That information is given by the predictive values.

- The positive predictive value (PPV) represents the percentage of the cases labelled as positive by the test that are actually positive. For this example, it represents the percentage of signals classified as aortic stenosis that are actually derived from an aortic stenosis related condition.

$$PPV=100\,\frac{TP}{(TP+FP)}.$$

- The negative predictive value (NPV) represents the percentage of cases labelled negative by the test that are actually negative. For this example it represents the percentage of signals clasiffied as normal that are actually representative of an anatomically and functionally normal heart condition.

$$NPV=100\,\frac{TN}{(TN+FN)}.$$

The predictive values are clinically useful but depend very strongly on the prevalence (Prev), that is the proportion of cases with the abnormality (Altman, 1991). A test with a constant sensitivity and specificity may have different predictive values for different groups making necessary a correction when tests conducted on different populations are compared (Chu, 1999).

$$PPV=\frac{(S+)(Prev)}{(S+)(Prev)+(1-S-)(1-Prev)}.$$

$$1-NPV=\frac{(1-S+)(Prev)}{(1-S+)(Prev)+(S-)(1-Prev)}.$$

For the screening test the prevalence is set to fifty percent since for all the medical conditions in this study the same probability of occurrence is assumed. Consequently, both classes are represented by the same number of samples.

Likelihood ratios (LR) are an alternative, newer method of judging the accuracy of a test. The likelihood ratio positive (LRP) compares the proportion of patients with the disease that have positive test results with the proportion of patients without the disease that have positive test results. The LRP is the ratio of these two proportions (Chu, 1999).

$$LRP = \frac{S+}{(1-S-)}.$$

$$LRN = \frac{1-S+}{S-}.$$

This approach may give further insight into the interpretation of diagnostic test data although it does not add new information since it uses the same quantities explained before. A likelihood ratio may demonstrate that the test is useful. However, it does not necessarily indicate that a positive test is a good indicator of the presence of disease (Altman, 1991). The likelihood ratio is especially useful to assess a sequence of tests. Table 7.4 shows the interpretation of likelihood ratios for the estimation of post-test probabilities. Values of LRP higher than 10 and values of LRN lower than 0.1 are considered to indicate that a test is of great diagnostic importance (Flores, 2002).

Following these standard diagnostic tests (Petrie, 2000; Rangayyan, 2002), Table 7.5 summarises the performance of the classifier based on the results presented in Table 7.2.

Likelihood ratio	Changes from pre-test to post-test probabilities
>10 or < 0.1	Large, often conclusive
5-10 or 0.1-0.2	Moderate
2-5 or 0.5-0.2	Small but sometimes important
1-2 or 0.5-1	Small, rarely important

Table 7.4. Interpretation of Likelihood Ratios.

No of samples	FP	TN	TP	FN	S+	S-	PPV	NPV	LRP	LRN
80	5	35	26	14	65%	87.5%	83.8%	71.4%	5.2	0.4

Table 7.5. Evaluation results using Cepstral Representation. FP false positives, TN true negatives, TP true positives, FN false negatives, S+ sensitivity, S- Specificity, PPV positive predictive value, NPV negative predictive value, LRP likelihood ratio positive, LRN likelihood ratio negative.

Table 7.5 is useful to evaluate the performance of the classifier in numerical figures suitable for comparison to other methods. In particular, the values of likelihood ratio are useful to value the classifier for clinical diagnosis.

From Table 7.5, it can be seen that the classifier is better at identifying the absence of an aortic stenosis murmur (specificity 87.5%) than at actually detecting the murmur (sensitivity 65%). From the signals classified as aortic stenosis, 83.8% are effectively aortic stenosis (positive predictive value), whereas 71.4 % of the labelled normal signals are effectively normal (negative predictive value). The likelihood positive ratio of 5.2 indicates that it is approximately five times more likely to label as "aortic stenois" a cycle with an aortic stenosis murmur, than a normal systolic cycle. In interpreting these results it has to be considered that signals recorded at sub optimal positions where the murmur is not at its loudest were included in the training process. The PCG recorded at the LLSE

shown in Figure 7.3 for example although it is from an aortic stenosis condition, it does not present the characteristic systolic murmur.

7.4.2 Hidden Markovian models and line spectral frequency representation

For the line spectral frequency representation the systolic signals are divided into frames of 25 ms overlapped by 50%. An 8-pole model provides the LP coefficients as described by Iwata (1977). However, instead of representing the signal by the 'significant' poles (bandwidth < 80 Hz) line spectral pairs and line spectral frequencies are obtained from the LP coefficients as described in Chapter 5.

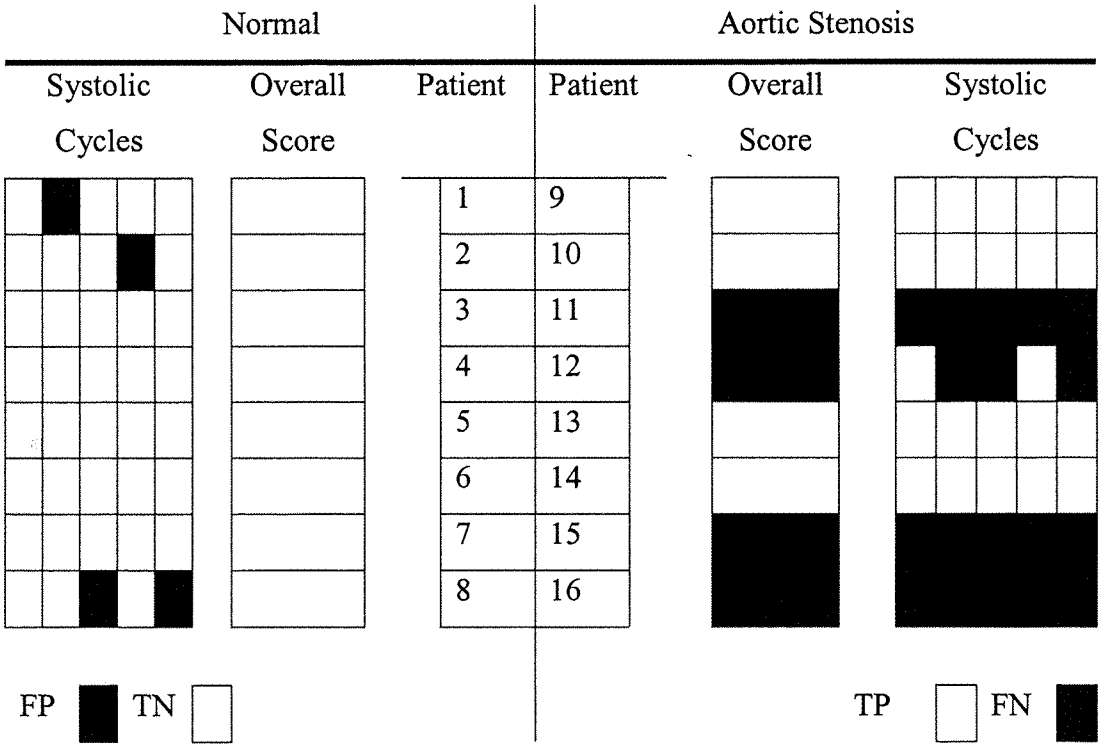


Table 7.6. Results of the Classification using Line Spectral Frequency Representation. The black boxes represent incorrectly classified signals (false positives Fp for normal samples, false negatives FN for aortic systolic cycles), whereas white boxes represent signals correctly classified (true negatives TN for normal samples, true positive TP for aortic systolic cycles).

The line spectral frequencies form a feature vector and a single Gaussian probability density distribution is assumed for each feature. Consequently the

model uses a weighted sum, or mixture, of several normal distributions and the parameters are calculated using the equations shown in section 4.6 (m= number of line spectral frequencies).

Table 7.6 and Table 7.7 present the results obtained for the screening test when line spectral frequency features are used as representations for the PCG. As in the case of the cepstral representations, no murmur is identified in signal 15 since the signals recorded in the mitral area are preferred for detection of the characteristic ejection click of congenital aortic stenosis where the systolic murmur is very soft or not present at all. By contrast, all the systolic murmurs are detected for signal 14 from the same patient but recorded in the aortic area since this position is preferred for murmur detection.

Signals 11 and 12 are recorded in different auscultation areas from the same patient presenting a mild valvular aortic stenosis. As in the case of the cepstral representations the classifier fails to detect the systolic murmurs of signal 11 recorded in the optimal aortic auscultation area. Nevertheless, murmurs are detected in two cycles of signal 12 recorded in the lower left sternal edge. Notice that this position is in between the optimal area for aortic stenosis murmur recording (aortic area) and the optimal position for the ejection click recording (mitral area).

The main difference on performance between representations is reflected on the identification of the systolic cycles of signal 16. While in the case of the cepstral representation all the cycles were correctly identified, in the case of the line spectral frequency all were wrongly classified.

The line spectral frequency representation showed a very small difference for wrong classification of normal systolic cycles. In this case only four systolic cycles were labelled abnormal compared with five cycles obtained previously using cepstral representation.

The performance of the hidden Markovian model using line spectral frequency representation is summarised in table 7.7. Using LSF the sensitivity decreases to 55% compared to 65% obtained using cepstral representations. The specificity slightly increased from 87% to 90% since only four normal cycles were wrongly classified compared to five cycles for the cepstral representation.

No of samples	FP	TN	TP	FN	S+	S-	PPV	NPV	LRP	LRN
80	4	36	22	18	55%	90%	84.6%	66.6%	5.5	0.5

Table 7.7. Evaluation Results Using Line Frequency Representation. FP false positives, TN true negatives, TP true positives, FN false negatives, S+ sensitivity, S- Specificity, PPV positive predictive value, NPV negative predictive value, LRP likelihood ratio positive, LRN likelihood ratio negative.

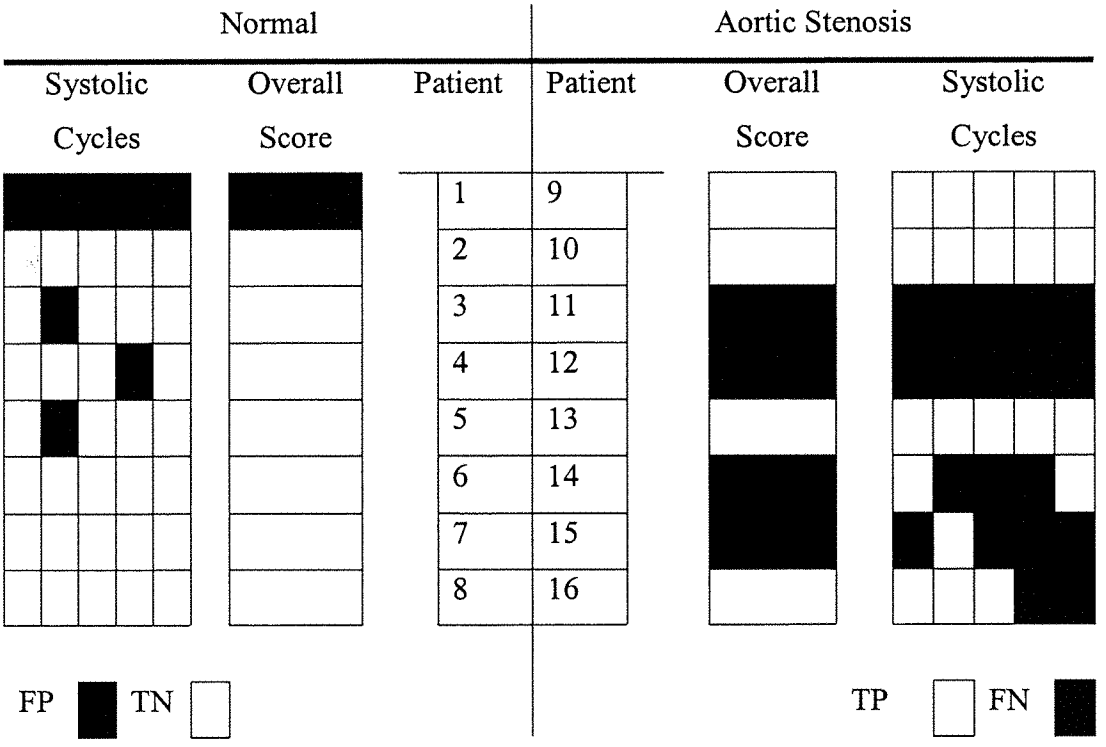


Table 7.8. Results of the Classification Using Maass-Weber Filter Representation. The black boxes represent incorrectly classified signals (false positives Fp for normal samples, false negatives FN for aortic systolic cycles), whereas white boxes represent signals correctly classified (true negatives TN for normal samples, true positive TP for aortic systolic cycles).

Positive predictive value and likelihood ratio positive are consequently slightly incremented, but negative predictive values and likelihood negative ratios reflect a decreased performance.

7.4.3 Hidden Markovian models and Maas-Weber filter representation

Table 7.8 and Table 7.9 show the results obtained for the HMM classifier using Maass-Weber filter representations as input features. The decreased performance of the algorithm is better reflected in the likelihood ratio values since the main difference in performance is given by the misclassification of the first normal signal.

No of samples	FP	TN	TP	FN	S+	S-	PPV	NPV	LRP	LRN
80	8	32	21	19	52%	80%	72.4%	62.7%	2.6	0.6

Table 7.9. Evaluation Results Using Maass-Weber Filter Representation. FP false positives, TN true negatives, TP true positives, FN false negatives, S+ sensitivity, S- Specificity, PPV positive predictive value, NPV negative predictive value, LRP likelihood ratio positive, LRN likelihood ratio negative.

7.4.4 General results

The graphical and analytical results show a better performance of the cepstral representations than that of the line spectral frequencies or Maass-Weber features as inputs for the HMM classifier. Nevertheless, conclusions have to be reached with some caution. The models for the screening test were representative of only two medical conditions. For other pathologies, other representations may surpass the performance of the cepstral representations. Moreover, it has to be taken into account that only a small set of PCG signals was available. If the three methods of signal representation have a quite similar performance, a larger database is required to truly differentiate their capabilities.

7.5 Comparison to cardiologist

Table 7.10 compares the performance of the HMM classifier and the cepstral representations with the results obtained by paediatric cardiologists. The study finds the sensibility and specificity of the clinical diagnosis of congenital heart disease, made by two paediatric cardiologists, taking as "a gold standard" an echocardiographic report (Flores, 2002). A study done in 150 children (younger than 15 years), the clinical diagnosis carried out by paediatric cardiologists with support of an electrocardiographic study and an X-ray plate of the thorax was contrasted with results obtained form the echocardiogram (continuous colour Doppler echo).

	No of samples	FP	TN	TP	FN	S+	S-	PPV	NPV	LRP	LRN
HMM	80	5	35	26	14	65%	87.5%	83.8%	71.4%	5.2	0.4
PC	150	20	38	78	14	84.8%	65.6%	71.1%	81.18%	2.46	0.23

Table 7.10. Comparative Results Between the Decisions Made by a Paediatric Cardiologist and the HMM Classifier. HMM hidden markovian model classifier, PC paediatric cardiologist, FP false positives, TN true negatives, TP true positives, FN false negatives, S+ sensitivity, S- Specificity, PPV positive predictive value, NPV negative predictive value, LRP likelihood ratio positive, LRN likelihood ratio negative.

The comparison between the performance of the HMM classifier and the performance of cardiologists is not entirely valid. The classifier was trained only for only two classes with very few training samples, whereas cardiologists were presented with ten classes, and although in both cases the decision was binary (detected or not), some conditions are harder, or easier, to be detected. Nevertheless, the comparison is made only to show reference levels to see how useful the classifier may be as a diagnosis tool.

The published predictive values for paediatric cardiologist (PPV=79.6% and NPV=73.1%) were obtained for a prevalence of 92/150. The values shown in table 7.9 (PPV=71.1% and NPV=81.18%) have been recomputed considering the prevalence of the abnormal conditions for the screening test (40/80) to ease the comparison (Chu, 1999).

As it can be seen from table 7.10 the paediatric cardiologist performs much better to detect the presence of the murmurs (higher sensitivity: S+=84.8% vs S+=65%) although the lower specificity (S-=65.6% vs S-=87.5%) may indicate a conservative approach to discarding a suspicious signal. The algorithm shows a higher certainty to an abnormal signal identified (higher PPV: 83.8% vs 71.1%) and lower certainty to a normal PCG identification (lower NPV: 71.4% vs 81.18%). Nevertheless, it should be notice that these figures were obtained for the low number of samples available.

As explained before, the likelihood positive and the likelihood negative ratios can be used as indicators of how useful is a classifier for clinical diagnosis. The results of the HMM classifier (LRP=5.2 and LRN=0.4), and the results obtained from the cardiologist with the help of ECG and X-ray plates (LRP=2.46 and LRN=0.23) are considered respectively of "moderate" and "small but sometimes important" clinical value for diagnosis (Flores, 2002).

7.6 Sequential test results

The results presented in table 7.10 compare the performance of either the cardiologist or the HMM classifier for diagnosis. Nevertheless, as an aid, it is important to assess how useful the HMM classifier is as a combined test to improve diagnosis. The results obtained by the cardiologist through the analysis of X-ray plates and electrocardiogram analysis can be used as pre-test values to asses the significance of the HMM classifier. More over, the use of these combined tests for diagnosis may prove to be useful in small communities where there is no

access to ultrasonic equipment or in cases where the access to ultrasonic services is inadequate[1].

Likelihood ratios can also be used to combine the results of multiple diagnosis tests (Chu, 1999). Lets consider a prevalence of 50% for a PCG signal containing a murmur such as in the screening test presented before. The patient's probability of a positive diagnosis after the paediatric cardiologist analysis of the X-ray plates and electrocardiogram can be obtained using the likelihood ratio positive value shown in table 7.10.

The pre-test odds for this prevalence of 50% is:

$$\text{Pre-test odds} = \frac{\text{Pre-test probability}}{\text{1- Pre-test probability}} = \frac{0.5}{1-0.5} = 1$$

Since both conditions have the same probability of occurrence. To compute the post-test odds:

$$\text{Post-test odds} = \text{Pre-test odds} \times \text{Likelihood Ratio Positive} = 1 \times 2.46 = 2.46$$

This value can be converted to post test probabilities:

$$\text{Post-test probability} = \frac{\text{Post-test odds}}{\text{1+Post test odds}} = \frac{2.46}{1+2.46} = 0.711$$

This is the positive predictive value (71.1%) for the test as presented in table 7.10. Therefore, after the diagnosis of the cardiologist the probability of disease changes from 50% to 71.1%. Notice that this value corresponds to the values computed for a prevalence of 50%.

[1] 'In the UK, increasing demand for ultrasound services and inadequate resources have led to long waiting lists with subsequent frustration of hospital clinicians, general practitioners and their patients' British Medical Ultrasound Society (Bates, 2003)

This case corresponds to a controlled test where the cardiologist is presented with equal number of abnormal and normal signals. The value for the prevalence (92/150) published by Flores (2002) differ since it reflects the particular case of the retrospective study (PPV=79.6%). This is one of the reasons explaining why the likelihood ratio, which is independent of the prevalence of the disease, is preferred over the positive predictive ratio for assessment of clinical value of a diagnosis test (Chu, 1999).

Once the paediatric cardiologist has assessed the probability of a positive diagnosis through the analysis of the ECG and X-ray plates, a sequential test using the HMM classifier would be useful if it significantly increases the pre-test to post-test probability of disease.

For the sequential test the post-test odds are given by:

$$\text{Post-test odds} = \text{previous test positive odds} \times \text{likelihood ratio positive}$$
$$= 2.4 \times 5.2 = 12.48$$

Converted to probability:

$$\text{Post-test probability} = \frac{\text{Post-test odds}}{1+\text{Post-test odds}} = \frac{12.48}{1+12.48} = 0.9258$$

From the initial assessment of the cardiologist a probability of disease of 71.1% is obtained, if the HMM classifier is then used as an aid for diagnosis the probability increases to 92.5%. This means that 92.5% of the patients with a positive diagnosis effectively present the abnormality.

The probability of the opposite situation, a patient with a negative diagnosis that actually presents the abnormality, is given by the combine negative predictive value.

For the sequential test the post-test odds are given by:

1- Post-test odds = previous test negative odds x likelihood ratio negative

= 0.23 x 0.4 = 0.092

Converted to probability:

$$1\text{- Post-test probability} = \frac{\text{Post-test odds}}{1+\text{Post-test odds}} = \frac{0.092}{1+0.092} = 0.084$$

Post test negative probability = 1- 0.084= 0.9157

Consequentially, after performing both tests, a negative diagnosis will be correct in 91.57% of cases.

Increasing the initial probability of a positive disease diagnosis from 50% to 92.58% and the probability of a negative diagnosis from 50% to 91.5% through the sequential use of both test proves the significance of the HMM classifier as an aid tool to support diagnosis.

An implicit assumption in the previous computation is that both tests are independent since different populations were screened. Nevertheless, in practice the sequential test will be applied to the same population, and although the post test probability may not be as high as the values presented, the specificity will be increased above that of any of the individual tests (Chu, 1999).

7.7 Summary

In this Chapter, the feasibility of using hidden Markovian models to classify systolic murmurs was explored. A binary screening test was designed to obtain valuable numerical figures reflecting the clinical value of the classifier.

The performance of the hidden Markovian model classifier using cepstral representations as input features was compared with the performance of a clinical cardiologist providing diagnosis based on the analysis of the electrocardiogram and x-ray plates in a similar test. Assuming similar conditions for both tests, the performance of the hidden Markovian model approach as a sequential test to aid the pre-diagnosis of the cardiologist was evaluated. Using likelihood ratios for the multiple testing, it was shown that the use of the hidden markovian model classifier is clinical useful.

The performance of the hidden Markovian model classifier using cepstral representations for PCG signals was compared against its performance using line spectral frequencies. A better performance was obtained for the former, although a larger database is required to confirm these results.

The classifier was trained only for two classes with a very limited number of samples. Nevertheless, the results obtained are promising. The validity of the method was demonstrated although the performance of the classifier is expected to improve with the addition of more training data.

8. Summary, Conclusions and Future Work

8.1 Summary

This research starts with an introduction to auscultation particularly focused on the use of the stethoscope for assessment of heart conditions and early detection of anomalies. The advantages of this non-invasive technique are highlighted in particular its value for paediatric cardiology, in which the frail nature of infants demands the use of passive observation procedures. The use of electronic stethoscopes and digital signal processing techniques to overcome inherent limitations of the traditional auscultation technique have lead to new areas of research. To exemplify this, the study of the genesis, transmission and propagation of heart sounds, the detection of cardiopathies, ventricular dysfunction and pulmonary hypertension, condition monitoring of prosthetic heart valves, and detection of coronary artery disease are briefly described in their relation to digital heart sound analysis.

The study is oriented into one of these areas of research, namely the detection of cardiopathies through the analysis of heart sounds. Centred on the paediatric population, five conditions are initially considered in this study namely: atrial septal defects, aortic stenosis, ventricular septal defect, innocent pulmonary flow murmurs, and normal heart sounds. The identification of these conditions is simplified to a classification task in which the signal is assigned to the most probable category according to its acoustic features.

After reviewing previous approaches, statistical techniques are proposed for PCG signal classification. They are considered as advantageous over parametric classifiers since inter-subject and intra-subject variations are better represented by statistical parameters and advantageous over neural networks due to their mathematical tractability. Thus, hidden Markovian models are recommended. One of the main advantages of this method is its adaptability to different signal

lengths. This feature is particularly important for the analysis of paediatric heart sound signals where a wide range of heart rates is expected.

Cepstral coefficients are proposed to provide a simplified signal representation of the phonocardiogram to serve as input for the classifier. A similar approach has been suggested before based in the similarities between speech signals and the heart sound production (Rangayyan, 1978b). In that study single cardiac cycles were analysed and the higher coefficients were proposed to reflect the differences between pathologies. In this research, cepstral analysis is proposed based on a model of the cardiohemic theory, and consequently, the phonocardiogram is modelled as a heart system response excited by either, the acceleration and deceleration of blood in the case of normal sounds, or turbulence in the case of murmurs (see appendix B).

Once a short time cepstrum is obtained the representation relies on the tracking of the low order coefficients reflecting the system response of the structure. Sensor movement artifacts and other noises commonly encountered in the PCG are mainly reflected in high order coefficients (assuming they are fast varying short time signals, see appendix B) and therefore the low order coefficients offer a certain degree of noise immunity. This cepstral truncation also provides certain degree of generalisation, since it reflects the statistical variations between signals of the same class without reflecting too much detail of a particular signal.

The graphical representations of the heart sounds, representative of the conditions under study, proved that the cepstral coefficients reflect changes in the phonocardiogram that could lead to discrimination between murmurs and ultimately to classification. Nevertheless, in order to compare the performance of the cepstral coefficients as feature representations, two other representations were applied to our experimental data.

Firstly, a sub-band energy level representation was obtained from a standard set of filters (Maass–Weber filters) used in clinical phonocardiology in other countries such as Germany and Mexico. The frequency response of these filters have been

defined through experience in clinical practice. Once the PCG signal is filtered, the energy in each band is computed. This is a sub optimal representation for the classification algorithm since features obtained using these filters are highly correlated due to the overlap of the frequency responses. Nevertheless, from a graphical perspective this representation proved to be very practical for visual identification of the pathologies under study.

A second alternative representation in which the system response is modelled as a product of resonances is proposed. In this approach, linear prediction is used to represent resonances as poles in the complex plane: high frequency poles correspond to murmurs and low frequency poles represent the heart sounds. A corresponding filter modelling the frequency response of the system is obtained.

Representations derived from linear prediction analysis have been used before as input feature representations for PCG parametric classifiers (Itawa, 1977). In practice, numerical precision and other characteristics make linear prediction coefficients an unsuitable representation for classifiers. Itawa (1980) for instance, proposes a feature space formed with prominent poles of the models (with bandwidth < 80 Hz) for classification.

The coefficients obtained through linear prediction coefficients should not be used in the statistical classifier to capture the variability among murmur samples from different subjects, since interpolation between parameters corresponding to different set of filters neither leads to a smooth frequency change nor ensures stability. For this research, line spectral frequencies were proposed as a more suitable representation to be used in the hidden Markovian model. They are more useful to represent inter subject and intra subject variability since the variations can be regarded as shifts of the position of the lines spectral frequencies on axis, and therefore, continuous and smooth transitions are possible. Nevertheless, the features are highly correlated since each of the resonances is modelled as a pair. The graphical representation of the PCG signals under study confirms the validity of using line spectral frequencies to differentiate between pathological conditions.

Since all the pathologies of interest present systolic murmurs, subtraction of the systolic cycles allows more detailed modelling. Segmentation is generally considered as the first stage towards PCG signal classification and is usually done using auxiliary signals like the ECG and carotid pulse. In paediatric cardiology, the simultaneous acquisition of these signals is not always convenient or practical. Therefore, an algorithm to segment the PCG form without the need of an external reference is required. Cepstral coefficients and hidden Markovian models were proposed for PCG segmentation.

A simple but effective model was proposed to identify the occurrence of either, the first and second heart sound, or the systolic and diastolic cycles. The time characteristics of the PCG signal provided the reference to isolate the systolic cycles. Using the leave one out method, a single Gaussian hidden Markovian model trained with the set of normal signals was proposed to perform the segmentation. The first low order cepstral coefficient conformed the input for the signal classifier. This continuous observation was related to a two-state hidden Markov process through a Gaussian distribution. The first state represented the generation mechanism for either the first or the second heart sound, the second state was related to the occurrence of either a systolic or diastolic cycle. Decoding of the state sequence provided the reference for the segmentation.

Considering the detection of the main components, the algorithm was evaluated for normal heart sounds. True positive events were defined in relation to the correct state detection and standard diagnostic tests were applied. High values of sensitivity (100%) and specificity (98%) resulted from the detection model.

The generality of the model trained using normal signals was demonstrated when used to segment PCG signals representative of other pathologies. The model resolved the main components for most of the signals using of only one cepstral coefficient. The use of cepstral truncation and a statistical model retained the main features of the signals for different auscultation areas, pathologies, and varying signal length corresponding to different heart rates.

In both cases: normal and abnormal PCGs, the identification of segments was checked visually aided by the ECG as a reference for S1. Ideally this could have been done using an autonomous system since there are several algorithms for detection of S1 using the QRS complex of the ECG, but a reliable detection of S2 requires the carotid pulse signal which was not available.

S1 and S2 were not allways correctly identified when very strong murmurs obscured the main features in the first cepstral coefficient representations as one might anticipate. The use of more coefficients is recommended in these cases since the main components are less likely to be obscured. Although a good option for the segmentation process, filtering off the high frequency murmurs from the PCG was not implemented since initially the identification of the main components (state decoding) and the classification of the murmurs (likelihood of a specific model) was conceived as a single process, and in this case the murmur needs to be kept as it is.

Once the signals were segmented, an abstract model was proposed using temporal and morphological characteristics of the systolic murmurs. A three-state Markovian model was selected as the generating mechanism related to the Cepstral observations by single Gaussian distributions.

A screening test was designed to assess the performance of the algorithm in a clinical context. For this binary test, the algorithm had to differentiate between normal and abnormal conditions. From the signals available, the class of aortic stenosis was selected since a higher number of samples represented it. Consequently, HMM models for normal and aortic stenosis classes were trained.

The leave-one out method was used during the test procedure. Standard tests were applied to assess the performance of the algorithm. Sensitivity and specificity values of 65% and 87%, positive and negative predictive values of 83% and 71%, and likelihood ratios of 5.2 and 0.4 were obtained. For the set used, the algorithm showed good performance although the algorithm proved better to detect the absence of the disease rather than its presence. The normal systolic cycles

presented less variability and therefore, the normal class was better represented with the reduced number of samples available.

To complete the database for the aortic model, signals recorded from sub-optimal areas were used for the training and testing. The inclusion of these signals diminished the performance of the model since the intensity of the murmur varies according to the auscultation area: it may be very low or even not present in some locations.

The performance of the algorithm was compared to the performance of paediatric cardiologists aided by an electrocardiographic study and X-ray plates in a similar study. The paediatric cardiologists were better able to detect the presence of a murmur (higher sensitivity: 84% vs 65%) although the HMM showed a higher certainty to an abnormal signal identified (higher PPV: 83% vs 71%). These results suggested the use of the HMM as a sequential test to follow the pre assessment of the paediatric cardiologist. Using both tests, an initial probability of a positive disease diagnosis increased from 50% to 92% whereas the probability of a negative diagnosis to be correct increased from 50% to 91%, which shows the significance of the HMM as a tool to support diagnosis.

A signal classification algorithm depends strongly on the feature representations, if the distinctive features are captured, the classification problem becomes trivial. There is also a match between feature representations and the classification algorithm, therefore the performance of the HMM and cepstral coefficients was compared to the performance of the HMM and line spectral frequencies. For the specific screening test, the performance of the HMM using cepstral coefficients was better, although it is considered that more signals are required to differentiate the performance of both representations.

8.2. Conclusions

Cepstral coefficients showed to be suitable representations for PCG analysis and classification. The features obtained using this technique are able to provide a simplified representation for the PCG and to reflect variations in the signal due to abnormalities.

Hidden Markovian models are valuable statistical methods for the analysis and classification of PCGs. Their merit as useful mathematical abstractions was demonstrated through the use of particular models proposed to undertake the practical problems of segmentation of PCG signals and classification of systolic murmurs.

Line spectral frequencies were proposed as an alternative representation for linear prediction coefficients. Although they were able to track changes in the PCG they showed poorer performance as the input for a HMM classifier when compared to the cepstral coefficients for our database. Nevertheless, in order to determine the extent of this, more signals are needed.

The use of a HMM classifier and cepstral coefficients as input features proved to be of clinical value particularly when used as a sequential test to aid diagnosis. Nevertheless, although the usefulness of the technique was demonstrated, more sample training signals are required to develop a system for clinical practice.

A considerable amount of effort has been devoted to the analysis of the influence of both training and testing sample size on the design and performance of pattern recognition systems. Although there are some recommendations for practitioners and rules-of-thumb in the literature (Rangayyan, 2002; Raudys, 1991), these are related to specific classification methods, under certain strong assumptions.

The performance of the algorithm as an aid for clinical diagnosis is ultimately compared to the performance of a human expert, and against a gold standard such as an echocardiographic report. Consequently the number of sampling signals

required to train the HMM for clinical use should be comparable to the sample size required in these studies. Flores (2002) defined a set of 150 cases for his study, and Gottdiener (2004) suggest a range from 100 to 200 subjects for the application of echochardiography for clinical trials. A set of similar size would be recommended for the training of the HMM.

8.3. Future work

8.3.1 Comparison to published techniques

The performance of a classification algorithm is strongly related to the database used for the training and testing procedures. For this reason, in order to determine the benefits of this approach, it is necessary to compare its performance with that of other classification algorithms using the same set of data.

8.3.1 Acquisition of more signals

Although the database collected contains a relevant number of PCGs, the number of signals representative of the conditions of interest for this study is very limited from the point of view of a statistical classifier. Some signals grouped in the same class, although relative to the same general pathology, have distinctive acoustical features that differentiate them according to the specialised literature in phonocardiology. In the case of aortic stenosis, for example, the study of the phonocardiogram is useful to distinguish among valvular aortic stenosis, a calcified valvular aortic stenosis, supravalvular aortic stenosis, fixed sub-valvular fibrous aortic stenosis, and dynamic aortic stenosis (Guadalajara, 1998). Consequently, more training signals of each condition are required in order to provide enough data to the models to capture the variations between sub-classes.

8.3.2. More complex models

HMM provide a structure that is broadly appropriate to represent the spectral and temporal variations of the PCG. However, it is assumed that the sound pattern is

produced in a process with instantaneous transitions between stationary states. This assumption is in direct contradiction with the fact that the heart sound signals are produced by a continuously moving physical system. However, this drawback may be diminished by a generous allocation of states to make a fair approximation of a dynamic system. More complex models with continuous state transition probabilities can also be used to represent dynamical systems. In both cases, larger sets of training data are required to account for the addition of more states or more parameters to model the dynamics of the system. In the model proposed in this study, the size of the training data limited the addition of more discrete states.

At this stage the PCG systolic signals have been modelled by HMM with a single Gaussian emission probability density. More complex models can be used to estimate the emission probability density functions, for instance, HMM with tied mixture of Gaussians, independent mixture of Gaussians, or neural networks, however, a more complex approach requires more training sequences to estimate the model parameters.

8.3.3 Full heart cycle model

Pathologies are reflected in the PCG not only with the presence of murmurs and extra heart sounds but also by changes in the components of the signal like the diminished first sound of an aortic stenosis due to left ventricular hypertrophy or the delay of A2 as the stenosis worsens. Therefore, a model that could comprise all the components of the heart sound would be highly valuable particularly in the cases of combined pathologies.

8.3.4 Alternative feature representations

The selection of cepstral analysis for feature representation was based in the cardiohemic theory for the genesis of heart sounds. Nevertheless, cepstral coefficients are not necessarily meaningful representations in the sense that the

visual features can not be directly related to the physical phenomena. Notwithstanding this, they are features possessing embedded information and are therefore useful for classification, letting the computer extract the subtle patterns characterising a specific condition. This is in contrast to the standard approach where the representations are derived to analyse the underlying physical phenomena, and the changes in the signal are visually traceable to the conditions. The second approach may provide a better visual representation but not necessarily a better feature extractor for an automatic classification system.

It is therefore recommended to try feature representations other than the classical time frequency representations, which provide easy to follow visual identification.

8.3.5 Multiple area acquisition

One of the main auscultation findings for differentiating between similar murmurs is the localisation and radiation pattern of murmurs. Simultaneous PCG acquisition for different auscultation areas will provide important information for the identification of pathologies.

8.3.6 Non-contact phonocardiology

The mass of the stethoscope used in traditional phonocardiology affects the frequency response and amplitude of the heart sound signals recorded at the chest. Moreover, handling noise and the added complexity for multiple signal acquisition suggest the use of non-contact methods for heart sounds recording. Laser phonocardiography has been used to measure foetal heart activity (Morgenstren, 1989) and to detect coronary artery disease (Furukawa, 2004). This technique, based on a Michaelson interferometer, provides higher sensitivity, wider frequency response, and may be useful for the implementation of a system for multiple point PCG acquisition.

References

Adolph R., *et al.*, 1970
'The clinical value of frequency analysis of the first heart sound in myocardial infarction'
Circulation; 41(6): 1003-1014.

Agress C., *et al.*, 1964
'Measurement of isometric contraction and ejection time by the vibrocardiogram'
American Journal of Cardiology; 13: 340-348.

Altman D., 1991
'Practical Statistics for Medical Research'
Chapman & Hall; U.K.
ISBN 0-412-27630-5

Atal B., 1974
'Effectiveness of Linear Prediction Characteristics of the Speech Wave for Automatic Speaker Identification and Verification'
Journal of the Acoustical Society of America; 55(6): 1304-1313.

Akay M., *et al.*, 1990a
'Noninvasive detection of coronary stenoses before and after angioplasty using eigenvector methods'
IEEE Transactions on Biomedical Engineering; 37(11): 1095-1104.

Akay M., *et al.*, 1990b
'Detection of coronary occlusions using autoregressive modelling of diastolic heart sounds'
IEEE Transactions on Biomedical Engineering; 37(4): 366-373.

Akay M., *et al.*, 1991
'Application of the ARMA method to acoustic detection of coronary artery disease'
Medical and Biological Engineering & Computing; 29(4): 365-372.

Akay M., *et al.*, 1992
'Application of adaptive filters to noninvasive acoustical detection of coronary occlusions before and after angioplasty'
IEEE Transactions on Biomedical Engineering; 39(2): 176-184.

Akay M., *et al.*, 1993
'Acoustical detection of coronary occlusions using neural networks'
Journal of Biomedical Engineering; 15(6): 469-473.

Akay M., *et al.*, 1994

'Dynamics of the sounds caused by partially occluded femoral arteries in dogs'
Annals of Biomedical Engineering; 22(5): 493-500.

Baracca E., *et al.*, 1990
'Dynamic spectral analysis applied to the study of the heart sounds'
Acta Cardiologica; 45(6): 505-510.

Barschdorff D., *et al.*,1995
'Automatic Phonocardiogram Signal Analysis in Infants based on Wavelet Transforms and Artificial Neural Network'
IEEE Conference Computers in Cardiology, Vienna, Austria, 10-13 September; 753-756.

Baranek R., 1989
'Automatic detection of Sounds and Murmurs in Patients with Ionescu-Shiley Aortic Bioprotheses'
Medical and Biological Engineering & Computing; 27: 449-455.

Bates J., *et al.*, 2003
'Extending the provision of Ultrasound services in the United Kingdom'
British Medical Ultrasonic Society
http://www.bmus.org/publications/strategy.htm.

Baum L., *et al.*, 1970
'A maximization technique occurring in the statistical Analysis of probabilistic functions of Markov Chains'
Annals of Mathematics and Statistics; 41(1): 164-171.

Brown E., *et al.*, 2002
'Heart Sounds Made Easy'
Elsevier Science; U.K.
ISBN 0-443-07141-1

Cave R., *et al.*, 1980
'Hidden Markov Models for English'
Proceedings Symposium Applications of HMM to Text and Speech, Princeton, USA, October; 16-56.

Charbit M., 1999
'Reconnaissance de Mots Isolés :Utilisation des Modéles HMM'
(Isolated Word Recognition: Using HMM)
http://www.tsi.enst.fr/~charbit/.

Chen D., 1997
'Time-frequency Analysis of the First Heart Sound. Part 1: Simulations and Analysis'
Medical and Biological Engineering & Computing; 25: 306-310.

Chu K., 1999
'An introduction to sensitivity, specificity, predictive values and likelihood ratios'
Emergency Medicine; 11: 175-181.

Coast D., *et al.*, 1990a
'Use of Hidden Markov Models for Electrocardiographic Signal Analysis'
Journal of Electrocardiology; 23: 184-191.

Coast D.; *et al.*, 1990b
'An Approach to Cardiac Arrhythmia Analysis using Hidden Markov Models'
IEEE Transactions on Biomedical Engineering; 37(9): 826-836.

Coast D., 1993
'Segmentation of High-Resolution ECGs using Hidden Markov Models'
Proceedings IEEE International Conference on Acoustics and Signal Processing, Minneapolis, USA., 27-30 April; 1: 67-70.

Debiais F., *et al.*, 1997a
'Time-frequency Analysis of Heart Murmurs. Part 1: Parametric Modeling and Numerical Simulations'
Medical and Biological Engineering & Computing; 35: 474-479.

Debiais F., *et al.*, 1997b
'Time-frequency Analysis of Heart Murmurs. Part 2: Optimisation of time-frequency representations and performance Evaluation'
Medical and Biological Engineering & Computing; 35: 480-485.

DeGroff C., *et al.*, 2001
'Screening Heart Murmurs in Pediatrics'
American Heart Journal; 103(22): 2711-16.

DeGroff C., *et al.*, 2001
'Artificial Neural Network – Based Method of Screening Heart Murmurs in Children'
Circulation; 103: 2711-2716.

Deller J., *et al.*, 2000
'Discrete-Time Processing of Speech Signals'
IEEE Press; U.S.A.
ISBN 0-7803-5386-2

Donnerstein R., 1989
'Continuos Spectral Analysis of Heart Murmurs for evaluating stenotic cardiac lesions'
American Journal of Cardiology; 64: 625-630.

Donnerstein R., *et al.*, 1994

'Hemodynamic and anatomic factors affecting the frequency content of Still's innocent murmur'
American Journal of Cardiology; 74(5): 508-510.

Durand L., *et al.*, 1982
'A computer model for the study of left ventricular wall contributions to the first sound'
IEEE Conference Computers in cardiology, Seattle, USA, October; 445.

Durand L., *et al.*, 1985
'Modelling the Transfer Function of The Heart Thorax Acoustic System in Dogs'
IEEE Transactions on Biomedical Engineering; 32(8): 592-599.

Durand L., *et al.*, 1986
'Evaluation of FFT-based and modern parametric methods for the spectral analysis of bioprosthetic valve sounds'
IEEE Transactions on Biomedical Engineering; 33(6): 572-578.

Durand L., *et al.*, 1990
'Comparison of pattern recognition methods for computer-assisted classification of spectra of heart sounds in patients with a porcine bioprosthetic valve implanted in the mitral position'
IEEE Transactions on Biomedical Engineering; 37(12): 1121-1129.

Durand L., *et al.*, 1995
'Digital Signal Processing of the Phonocardiogram: review of the Most Recent Advances'
Critical Reviews in Biomedical Engineering; 23(3,4): 163-219.

Ertel P., *et al.*, 1966
'Stethoscope Acoustics: The doctor and his stethoscope'
Circulation; 34: 889-908.

Fahy F., (Ed.), 1998
'Fundamentals of Noise and Vibration'
E & FN Spon. London.
ISBN 0 419 27700 8

Fergulio G., 1963
'New prospects for phonocardiography: the use of intracardiac sound generators for the study of the transmission and calibration of cardiovascular sounds'
Bollettino della Societa Italiana de Cardiologia; 8: 41-44.

Flores A., *et al.*, 2002
'Sensibilidad y especificidad del diagnóstico clínico de cardiopatías congénitas'

(Sensibility and specificity of the clinical diagnosis of congenital heart diseases)
Revista Mexicana de Pediatría; 69(3): 99-101.

Fukunaga K., *et al.*, 1989
'Effects of sample size in classifier design'
IEE Transactions on Pattern Analysis and Machine Intelligence; 11(8): 873-885.

Furukawa T., (Ed.), 2004
'Biological Imaging and Sensing'
Springer; U.S.A.
ISBN 3-540-4398-X

Gold B., *et al.*, 2000
'Speech and Audio Signal Processing'
John Wiley & Sons, Inc.; U.S.A.
ISBN 0-471-35154-7

Gottdiener J., *et al.*, 2004
'American Society of Echocardiography: Recommendations for use of Echocardiography in Clinical Trials'
Journal of the American Society of Echocardiography; 17(10): 1086-1119

Groom E., 1964
'Comparative efficiency of stethoscopes'
American Heart Journal; 68: 220-226.

Guadalajara J., 1998
'Cardiología'
(Cardiology)
Méndez Editores; México.
ISBN 968-6596-59-3

Haghighi-Mood A., *et al.*, 1995
'A Sub-Band Energy Tracking Algorithm for Heart Sound Segmentation'
IEEE Conference Computers in Cardiology, Vienna, Austria, 10-13 September ; 501-504.

Halstead B., *et al.*, 1984
'Paralytic Shellfish Poisoning'
World Health Organization; Switzerland.

Holmes J., *et al.*, 2001
'Speech Synthesis and Recognition'
Taylor & Francis; U.K.
ISBN 0-7484-0856-8

Holldack K., *et al.*, 1974

'Atlas und Lehrbuch der Phonokardiographie'
(Atlas of Phonoardiography)
Georg Thieme Verlag; Germany.

Holldack K., *et al.*, 1965
'Nociones de Fonocardiografía'
(Fundamentals of Phonocardiography)
La Prensa Médica Mexicana; México.

Isselbacher K., *et al.*, 1984
'Harrison's Principals of Internal Medicine'
McGraw-Hill, Inc.; U.S.A.

Itakura F., 1975
'Line spectrum representation of linear predictor coefficients of speech signals'
Journal of the Acoustical Society of America; 57(1): S35.

Iwata A., *et al.*, 1977
'Pattern Classification of the Phonocardiogram using Linear Prediction Analysis'
Medical and Biological Engineering & Computing; 15: 407-412.

Iwata A., *et al.*, 1980
'Algorithm for detecting the First and the Second Heart Sounds by Spectral Tracking'
Medical and Biological Engineering & Computing; 18: 19-26.

Jamous G., *et al.*, 1992
'Optimal time-window duration for computing time/frequency representations of normal phonocardiograms in dogs'
Medical and Biological Engineering & Computing; 30(5): 503-508.

Jelinek F., 1999
'Statistical Methods for Speech Recognition'
Massachusetts Institute of Technology; U.S.A.
ISBN 0-262-10066-5

Johnson G., *et al.*, 1983
'Estimation of the Severity of Aortic Stenosis by Frequency Analysis of the Murmur'
American College of Cardiology; 1(5):1315-1323.

Joo T., *et al.*, 1983
'Pole-zero modeling and classification of phonocardiograms'
IEEE Transactions on Biomedical Engineering; 30(2): 110-118.

Kang W., *et al.*, 1995

'The Application of Cepstral Coefficients and Maximum Likelihood Method in EMG Pattern Recognition'
IEEE Transactions on Biomedical Engineering; 42(8): 777-785.

Kao C., *et al.*, 1993
'Paralytic Shellfish Poisoning' In Falconer, I. A. [ed.], Algal Toxins in Seafood and Drinking Water
Academic Press; U.K.
ISBN 0-12-247990-4

Koski A., 1996
'Modelling ECG Signals with Hidden Markov Models'
Artificial Intelligence in Medicine; 8: 453-471.

Kozmann G., 1977
'An acoustic model of the heart'
Proceedings International Symposium Medicine and Informatics; 253.

Kusumoto F., 1999
'Cardiovascular Pathophysiology'
Fence Creek Publishing; U.S.A.

Lehner R., *et al.*, 1987
'A Three - Channel Microcomputer System for Segmentation and Characterisation of the Phonocardiogram'
IEEE Transactions on Biomedical Engineering; 34(6): 485-489.

Leung T., *et al.*, 1997
'Time Frequency Methods for Analysing Paediatric Heart Murmurs'
Applied Signal Processing; 4: 154-167.

Leung T., *et al.*, 1998a
'Time-Frequency Characterisation of Paediatric Heart Sounds'
PhD thesis; Southampton University; ISVR Southampton, U.K.

Leung T., *et al.*, 1998b
'Analysis of the Second Heart Sound for Diagnosis of Paediatric Heart Disease'
IEEE Proceedings Science, Measurement and Technology; 145(6): 285-290.

Leung T., *et al.*, 1999
'Characterisation of paediatric heart murmurs using self-Organising Map'
Proceedings of the joint meeting of BMES & IEEE Engineering in Medicine and Biology Society, Atlanta, USA.; 926

Leung T., *et al.*, 2000a
'Analysing Paediatric Heart Murmurs with Discriminant Analysis'

Proceedings of the 19th Annual conference of the IEEE Engineering in Medicine and Biology Society, Hong Kong; 20(3): 1628-1631.

Leung T., *et al.*, 2000b
'Classification of Heart Sounds using Time-Frequency Method and Artificial Neural Networks'
Proceedings of the World Congress on Medical Physics and Biomedical Engineering, Chicago, USA; 2: 988-991.

Liang H., *et al.*, 1997
'Heart Sound Segmentation Algorithm based on Heart Sound Envelogram'
IEEE Conference Computers in Cardiology, Lund, Sweden, 7-10 September; 24: 105-108.

Makhoul J., 1975
'Linear Prediction: A Tutorial Review'
Proceeding of the IEEE; 63(4): 561-580.

McKusick V., 1958
'The History of Cardiovascular Sound'
Cardiovascular sound in Health and disease.
Williams & Wilkins; U.K.

Messer S., *et al.*, 2001
'Optimal Wavelet Denoising for Phonocardiograms'
Microelectronics Journal; 32: 931-941.

Miao C., *et al.*, 1987
'Genesis of vibratory functional murmurs'
American Journal of Cardiology; 60(14): 1198-1199.

Morgenstern J., *et al.*, 1989
'A Laser Interferometer to Measure Fetal Heart Activity'
Clinical Physics and Physiological Measurements; 10(B): 75-78.

Mount D., 2004
'Multiple Sequence Alignment' in *Bioinformatics*
Cold Spring Harbor Laboratory Press; USA

MSD: Merck Sharp & Dohme, 2003
'Auscultación cardiaca (Cardiac Auscultation)'
CD-Rom.
http://www.msd.com/

Murphy K., 2001
'Hidden Markov Model (HMM) Toolbox'
http:// www.cs.berkeley.edu/~murphy/Bayes/hmm.html.

Neuburg E., 1971

'Markov Models for Phonetic Text'
Journal of the Acoustical Society of America; 50(A): 116.

Obaidat M., 1992
'Performance of the Short Time-Fourier Transform and Wavelet Transform to Phonocardiogram Analysis'
Proceeding of the ACM Symposium on Applied Computing, Missouri, USA; 856-862.

Olmez T., *et al.*, 2003
'Classification of Heart Sounds using an Artificial Neural Network'
Pattern Recognition Letters; 24: 617-629.

Oppenheim A., et al., 1975
'Digital Signal Processing'
Prentice Hall; USA

Owens F., 1993
'Signal Processing of Speech'
MacMillan Press Ltd.; U.K.

Pan J., *et al.*, 1985
'A real-time QRS detection algorithm'
IEEE Transactions on Biomedical Engineering; 32: 230-236.

Petrie A., *et al.*, 2000
'Medical Statistics at a Glance'
Blackwell Science; U.K.

Phua K., *et al.*, 2006
'Human Identification Using Heart Sounds'
Second Workshop on Multimodal User Authentication
Toulouse, France, May 11-12.

Portiz A., 1982
'Linear Predictive Hidden Markov Models and Speech Signal'
Proceedings IEEE International Conference on Acoustics and Signal Processing, Paris, France, May; 1: 1291-1294.

Portiz A., 1988
'Hidden Markov Models: a Guided Tour'
Proceedings IEEE International Conference on Acoustics and Signal Processing, New York, USA., 27-30 April; 1: 67-70.

Rabiner L., 1989
'A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition'
Proceedings of the IEEE; 77(2): 257-286.

Ranga A., *et al.*, 2006
'Computational Simulations of the Aortic Valve Validated by Imaging Data: Evaluation of Valve-Sparing Techniques'
Interactive CardioVascular and Thoracic Surgery; 10.1510:

Rangayyan R., *et al.*, 1978a
'Cepstral filtering of phonocardiogram signal'
Digest of International Symposium and Workshop on Biomedical Engineering; Delhi, India, Feb; (4): 118-119.

Rangayyan R., *et al.*, 1978b
'Speech Processing Techniques for the Analysis of Heart Sounds'
Proceedings of VII All India Symposium of Biomedical Engineering; Hyderabad, India, June; (8): S.4-9 - S.4-12.

Rangayyan R., *et al.*, 1979
'New Techniques for Phonocardiogram Analysis'
Medical Life Sciences Engineering; 5: 6-17.

Rangayyan R., *et al.*, 1987
'Multichanel Phonocardiography for localization of heart sound sources'
Pattern Recognition and Acoustical Imaging; 768:248.

Rangayyan R., *et al.*, 1988
'Phonocardiogram Signal Analysis: A Review'
Critical Reviews in Biomedical Engineering; 15(3): 211-236.

Rangayyan R., *et al.*, 1997
'Parametric Representation and Screening of Knee Joint Vibroarthrographic Signals'
IEEE Transactions on Biomedical Engineering; 44(11): 1068-1074.

Rangayyan R., 2002
'Biomedical Signal Analysis: A Case-study Approach'
Wiley Inter-Science; U.S.A.
ISBN 0-471-20811-6

Raudys S., *et al.*, 1991
'Small Sample Size Effects in Satistical Pattern Recognition: Recommendation for practitioners'
IEEE Transactions on Pattern Analysis and Machine Intelligence; 13(3): 252-264

Reed T., *et al.*, 2001
'Analysis of Heart Sounds for Symptom Detection and Machine – Aided Diagnosis'
Proceedings of the 4th International EUROSIM Congress, Delft, The Netherlands, June 26-29; 1-7.

Riota J. and Lukkarinen S., 1996
'Comparison of Time-Frequency Distributions in the Heart Sound Analysis'
Medical and Biological Engineering & Computing; 34(1) :89-90.

Romero E. and White P., 2002
'Heart Sound Segmentation by Hidden Markov Models'
Proceedings of the IEE seminar Medical Applications of Signal Processing; London, 7 October, 19/1-19/4.

Rushmer R., 1968
'Meaning of Murmurs'
American Journal of Cardiology; 21: 722-730.

Rushmer R., 1970
'Cardiovascular Dynamics'
WB Saunders; U.S.A.

Sava H.,1998
'Application of the matching pursuit method for structural decomposition an averaging of phonocardiographic signals'
Medical & Biological Engineering & Computing; 302-308.

Shaver J., *et al.*, 1985
'Current Problems in Cardiology'
Year Book Medical; U.S.A.

Stein P., *et al.*, 1980
'Frequency Spectrum of the Aortic Component of the Second Heart Sound in Patients with Normal Valves, Aortic Stenosis and Aortic Porcine Xenografts'
American Journal of Cardiology; 46: 48-52.

Tavel M., 1978
'Clinical Phonocardiography and external pulse recording'
Year Book Medical; U.S.A.

Thoraval L., *et al.*, 1992
'Continuously Variable Duration HMM for ECG Segmentation'
Proceedings of the IEEE Engineering in Medicine and Biology Society; 29 October - 1 November, 2: 529-530.

Thoraval L., *et al.*, 1994
'Heart Signal Recognition by Hidden Markov Models: the ECG case'
Methods Inf. Med. March; 33(1): 10-14.

University of Leeds, 1997
'Hidden Markov Models (Tutorial)'

http://www.comp.leeds.ac.uk/scs-only/teaching-materials/HiddenMarkovModels/html_dev/main.html.

Unser M. and Aldroubi A., 1996
'A Review of Wavelets in Biomedical Applications'
Proceedings of the IEEE; 84(4): 626-638.

Vaseghi S., 2000
'Advanced Digital Signal Processing and Noise Reduction'
John Wiley & Sons, Inc.; U.K. pp. 227 - 261
ISBN 0-471-62692-9

Van Vollenhoven E., 1971
'Calibration of contact microphones applied to the human chest wall'
Medical and Biological Engineering; 9(4): 365-373.

Verburg., *et al.*, 1978
'Phonocardiography Non-invasive Measurements'
Academic Press; U.K.

Viterbi A., 1967
'Error bounds for convolutional codes and an asymptotically optimal decoding algorithm'
IEEE Transactions on Information Theory, IT-13: 260-269

Webster J. ed, 1998
'Medical Instrumentation: Application and Design'
John Wiley & Sons; U.S.A.
ISBN 0-471-15368-0

Wood J., *et al.*, 1994
'Regional effects of myocardial ischemia on epicardially recorded canine first heart sounds'
Journal of Applied Physiology; 76(1): 291-302.

Wood J., *et al.*, 1995
'Time-Frequency analysis of the first heart Sound'
IEEE Engineering in Medicine and Biology Magazine; 144.

Xu J., *et al.*, 2000
'Nonlinear Transient Chirp signals Modelling of the Aortic and Pulmonary Components of the Second Heart Sound'
IEEE Transactions on Biomedical Engineering; 47(7): 1328-1335.

Yoganathan A., *et al.*, 1976
'Use of the Fast Fourier Transform for Frequency Analysis of the First Heart Sound in a Normal Man'
Medical and Biological Engineering; 14: 455.

Zalter., *et al.*, 1963
'Acoustic transmission characteristics of the thorax'
Journal of Applied Physiology; 18: 428-436.

Zhang Xuan., *et al.*, 1998
'Analysis Synthesis of the Phonocardiogram Based on the Matching Pursuit Method'
IEEE Transactions on Biomedical Engineering; 45(8): 962-971.

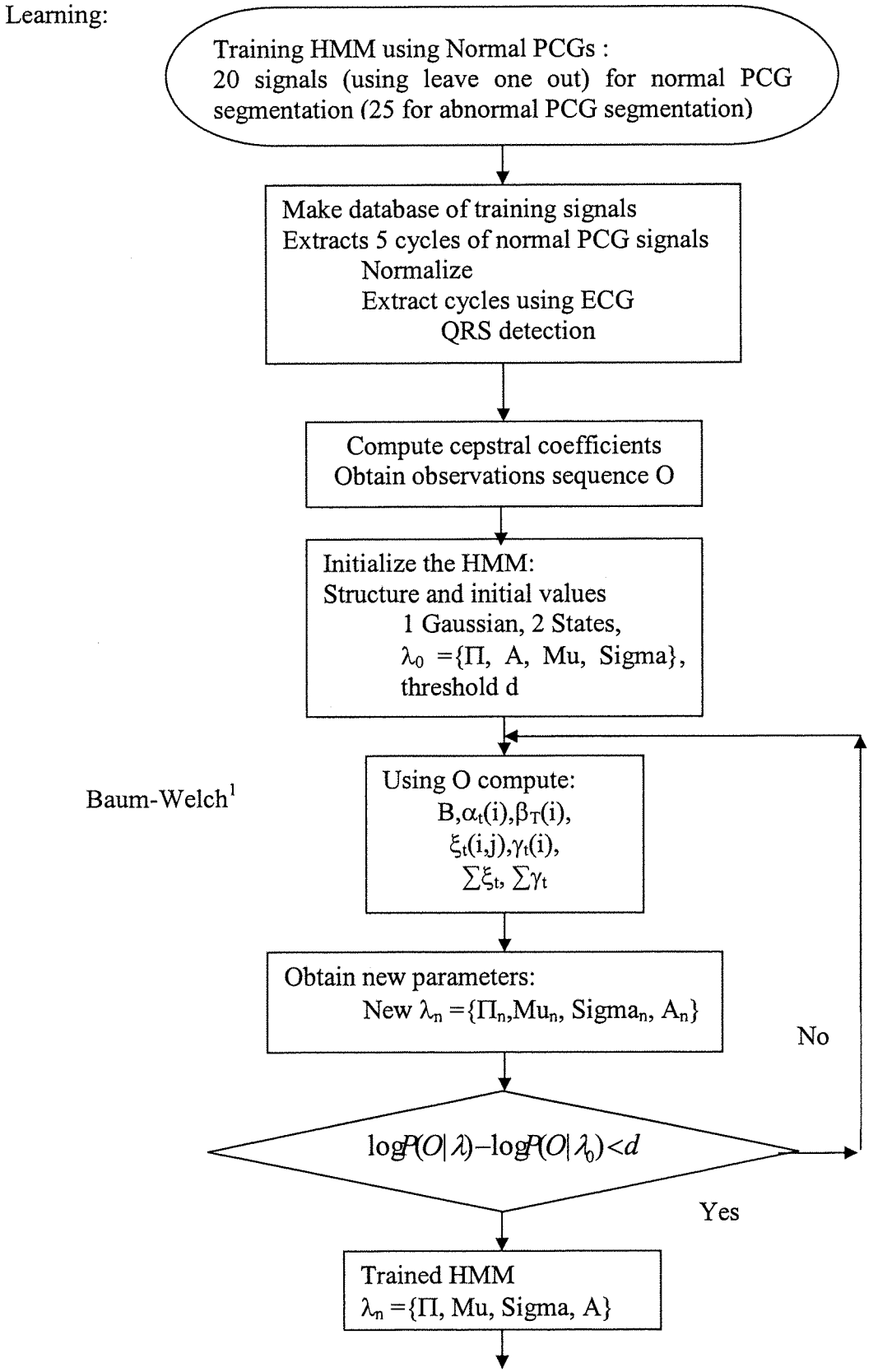
Appendix A: Flow diagrams

Implementation

The programs were initially implemented using Matlab version 5.3.1 (R11.1) running in a computer with a Pentium II processor and windows 98 system. It took 5 minutes to run the segmentation algorithms, and around 4 minutes to run the classification program for each signal under test. Considering that there were 17 signals for abnormal PCG segmentation and 16 for systolic murmur classification, this was a problem specially for debugging and features comparison. However, using the same Matlab version 5.3.1 but on a computer with an Intel Centrino M at 1.60 GHz and windows XP the segmentation and classification algorithms now run in 23 and 20 seconds per set respectively.

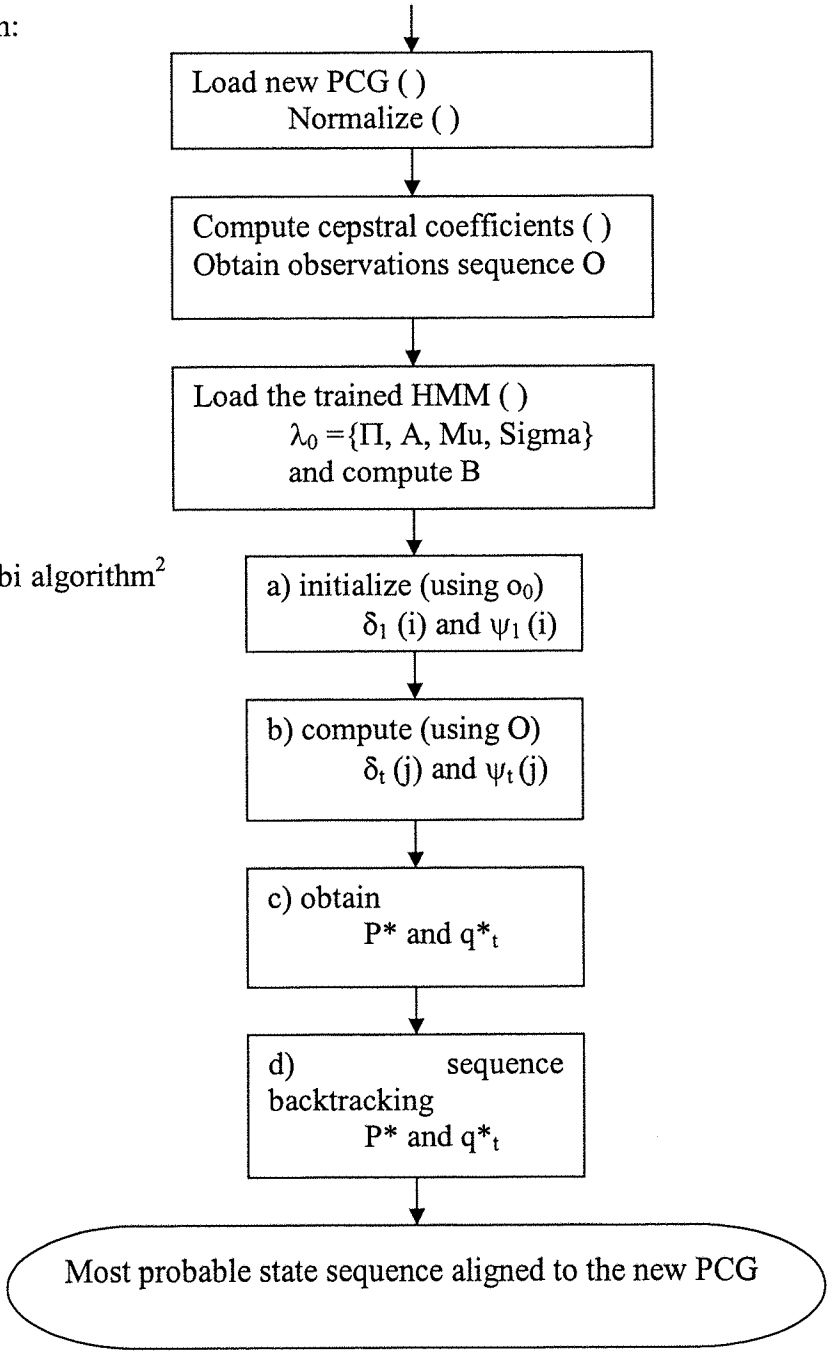
The next section shows the flow diagrams for the programs making reference to the equation defined in chapter 4.

Heart Sound Segmentation by HMM



[1] Using function 'ess_mhmm' from 'hmm_learn' Murphy's (2001) HMM toolbox

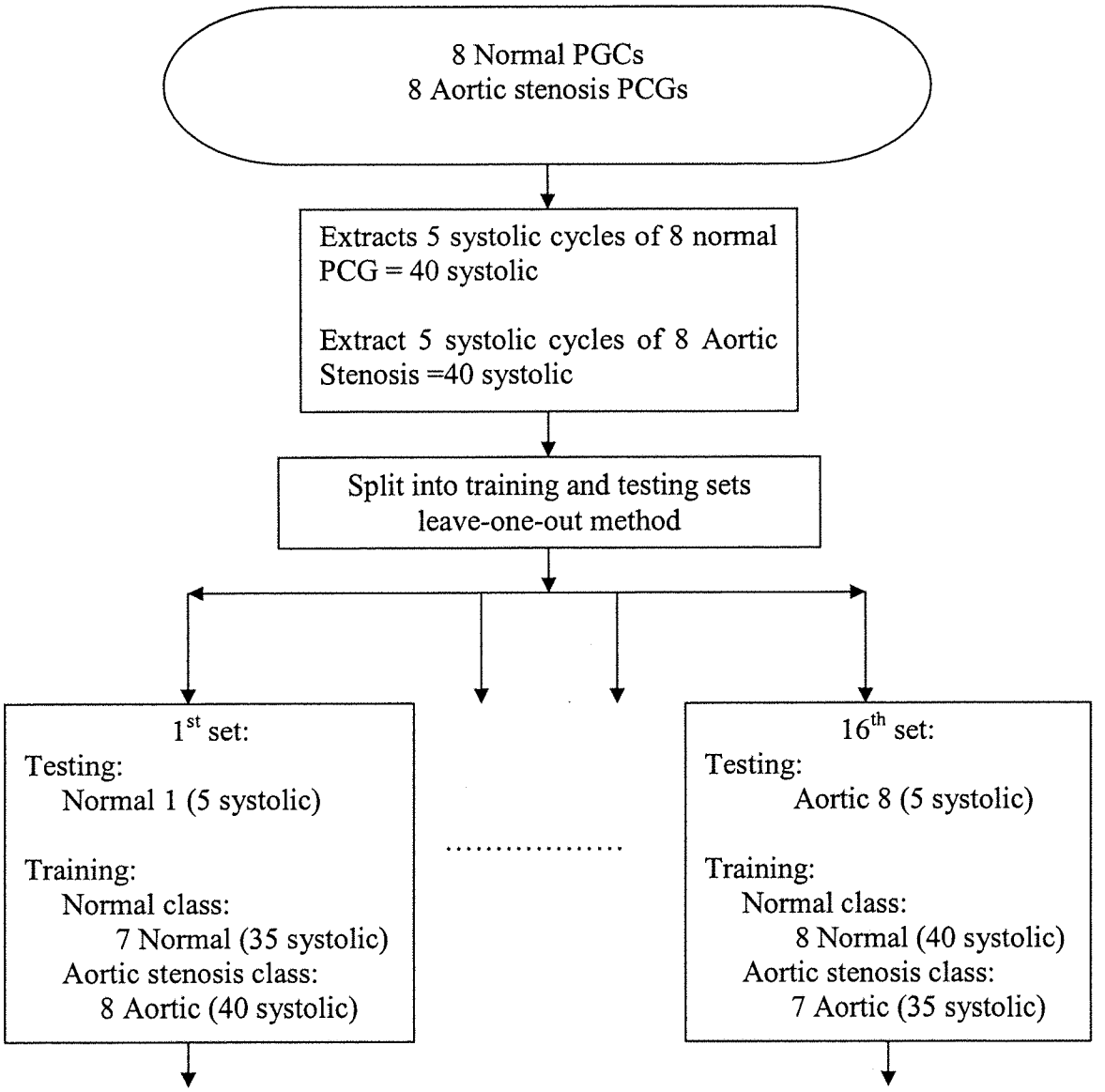
Segmentation:



[2] Using viterbi_path from Murphy's (2001) HMM toolbox

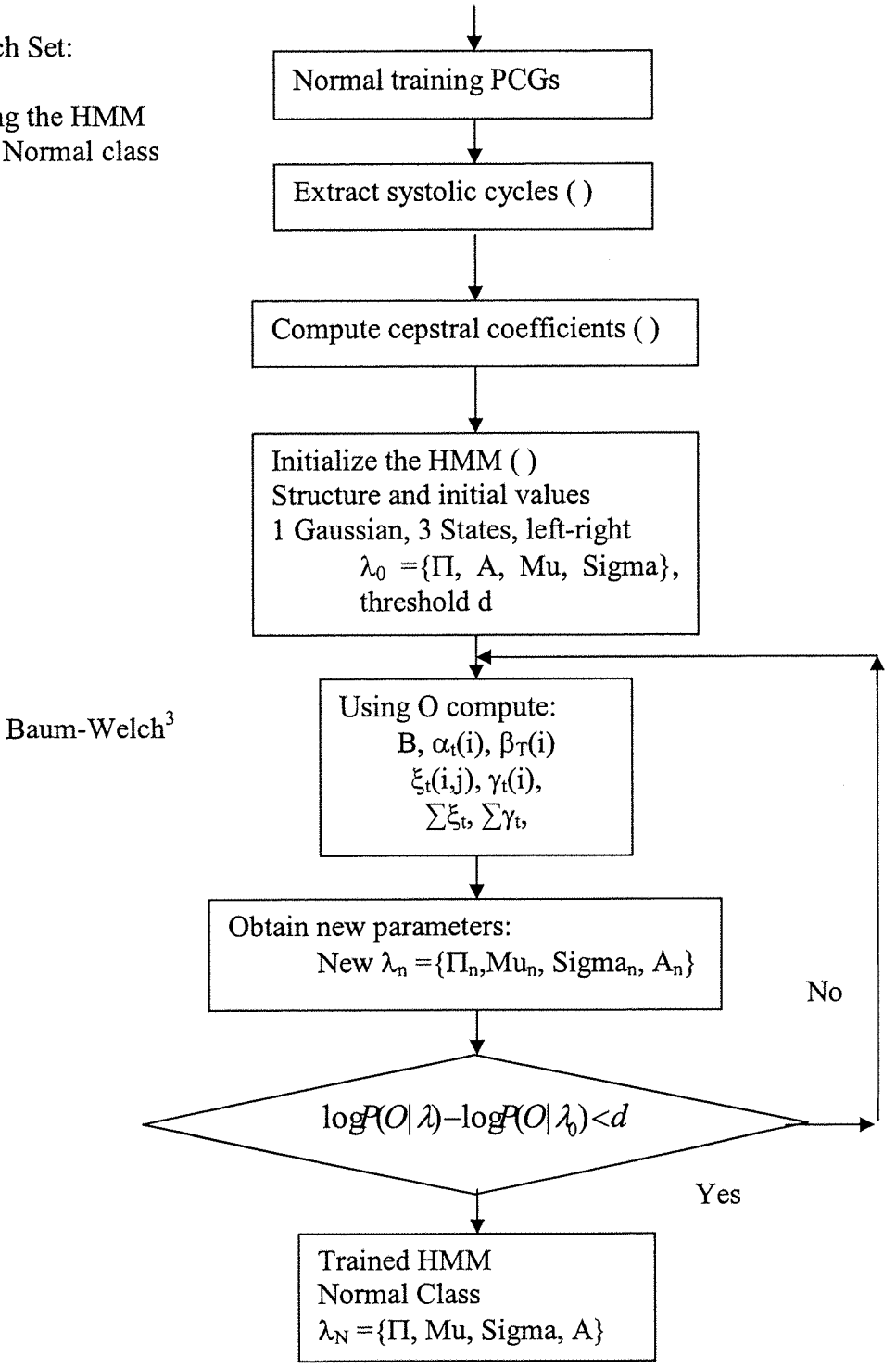
Systolic Murmur Classification by HMM

Create sets:



For each Set:

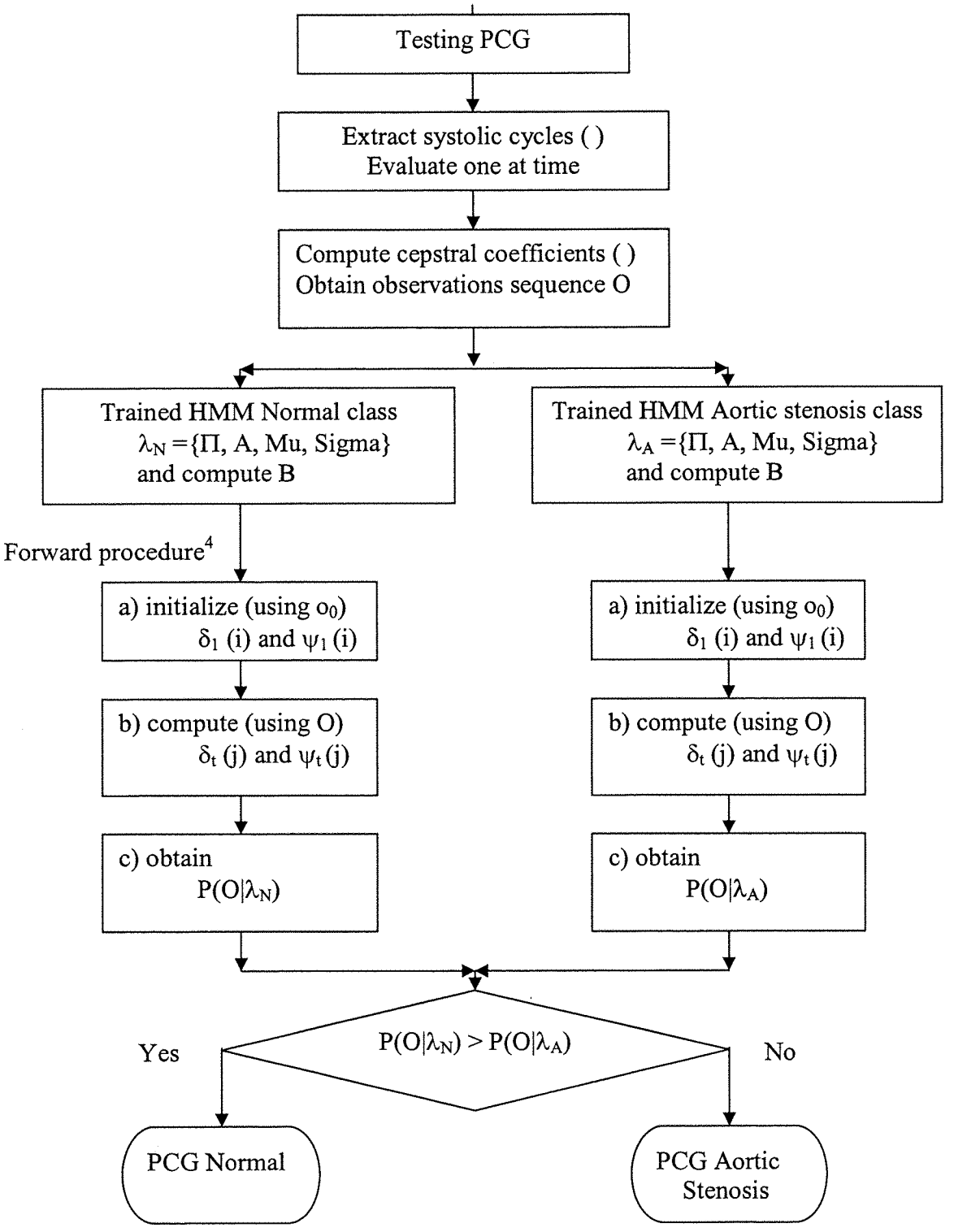
Training the HMM
for the Normal class



Same procedure for Aortic Stenosis Class.

[3] Modified function 'hmmtrain' to force left right model from Charbit's (1999) HMM toolbox.

Classification:



[4] Function 'forback' from Charbit's (1999) HMM toolbox.

Appendix B

Cepstral analysis for PCG signals

This appendix shows the steps involved to obtain the input representations for the HMM using simulated and real PCG signals. Details of the implementation, such as the window size selected for the FFT, and the choice of the number of coefficients to represent the signal, are also provided. The concept of impulse response and excitation separation through cepstral liftering is explored, and the cepstral coefficients are related to the spectrogram to explain their meaning.

PCG Signals

Synthetic and real PCG signals are going to be used in this appendix in order to explain the concepts involved in the application of cepstral analysis to phonocardiogram signals. The synthetic signals are used to explain the theoretical concepts of the technique, whilst the real PCGs are used to ilustrate the details of the implementation. The validity of these synthetic signals to represent real data may be questionable, but using simulated data it is possible to show the concepts and the details that may not be so easy to visualize using only real signals.[1]

Synthetic PCG

Simulation of the first heart sound

The synthesis of the first heart sound is based on a model proposed by Chen (1997) in which S1 is composed of two types of vibration: a valvular component and a muscular component from the myocardium. The valvular component is produced by resonance of the valve leaflets at constant frequencies, whereas the

[1] Only real PCG signals were used to develop the algorithms in the thesis.

myocardial response generates a rising frequency component with instantaneous frequency proportional to the tension of the cardiohemic system (Chen, 1997b).

The sounds produced by the valves are composed of transient signals of short duration and decaying amplitude, and therefore are modelled by a set of exponentially decaying sinusoids.

$$S_v(t) = \sum_{i=1}^{N} A_i e^{-k_i t} \sin(2\Pi f_i t + \phi_i)$$

Where A is the amplitude, k is the damping factor, f is the frequency and ϕ is the phase of the *i-th* sinusoid.

In this model, the heart sound signal is the response of a frequency selective acoustic system which consists of the structures in the chest, the heart, and the major vessels to the impulse excitation of the rapidly decelerating flow on the valve leaflets. This model is adopted for the simulation of S1 since it agrees with the theory for the genesis of the heart sounds used in this research.

A modulated frequency component associated with the myocardial tension generated during left ventricular contraction constitutes the second type of vibration. This is represented by a deterministic signal with frequency and amplitude modulation. The frequency is increased during contraction and then remains constant as the force plateau is reached. The amplitude increases rapidly, stays constant for an interval, and then fades to zero.

The myocardial component is thus represented by

$$S_m(t) = A_m(t)\sin(2\Pi(f_0 + f_m(t))t + \phi_m(t))$$

Where A_m is the amplitude modulating wave, f_0 is the carrier frequency, fm is the frequency modulating wave, and ϕ is the phase.

The first heart sound is thus represented by

$$S_t(t) = S_m(t) \qquad 0 \le t \le t_0$$

$$S_t(t) = S_m(t) + S_v(t - t_0) \qquad t \ge t_0$$

Where t_0 is the delay between the onset of the two components, since the mitral valve closes after the myocardium contracts.

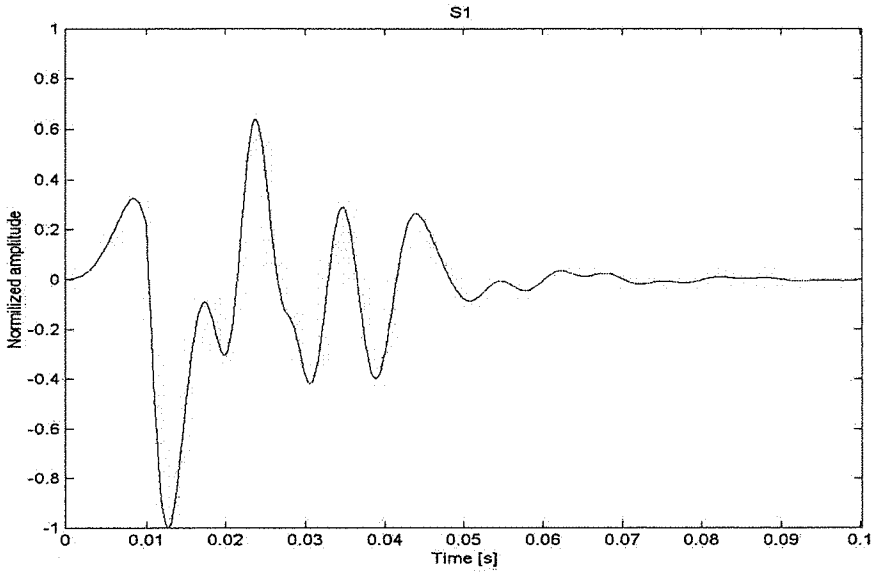


Figure B.1. Synthetic S1.

Figure B.1 show the synthetic signal representing the first heart sound. The signal has two valvular components (N=2) with frequencies of 50 and 150 Hz, amplitudes of 1 and 0.5, a damping factor of 60, and a delay of 10 ms (Chen, 1997).

The myocardial component has an amplitude modulation

$$A_m = 0.275(1.1 - 0.9\cos(83.4\pi t)) \qquad 0 \le t \le 12ms$$

$$A_m = 0.55 \qquad 12ms \le t \le 30ms$$

$$A_m = 0.275(1 - \cos(34\pi t)) \qquad 30ms \le t \le 60ms$$

$A_m = 0$ t>60ms

The frequency modulating wave is

$f_m(t) = -40\cos(34\pi t), f_0 = 60Hz, \phi_m(t) = 0$ $0 < t < 30ms$

$f_m(t) = 0, f_0 = 100Hz, \phi_m(t) = -\frac{2}{5}\pi$ $30ms < t < 60ms$

The study conducted by Chen (1997) shows that this simulated S1 signal has temporal and spectral characteristics similar to S1 recorded in humans.

Simulation of a systolic murmur

The synthesis of the systolic murmur is base on the work of Debiais (1997a). Using a database of 153 patients, the spectrogram of the PCG signals was computed using a 25 ms Hanning window to find the basic amplitude and frequency characteristics of several murmurs. Murmurs were then modelled by random signals modulated in frequency and amplitude to simulate the random vibrations of blood due to turbulence.

To simulate the crescendo-decresendo shape typical of aortic stenosis murmur, random noise is windowed in time by a Hamming window of 300 ms. The frequency pattern is obtained weighting the frequency response using a Hamming window centred in the origin in order to obtain the patterns described by Debiais (1997).

Figure B.2 shows the synthetic aortic stenosis murmur obtained by this procedure. Note the 'diamond shape' mid systolic murmur.

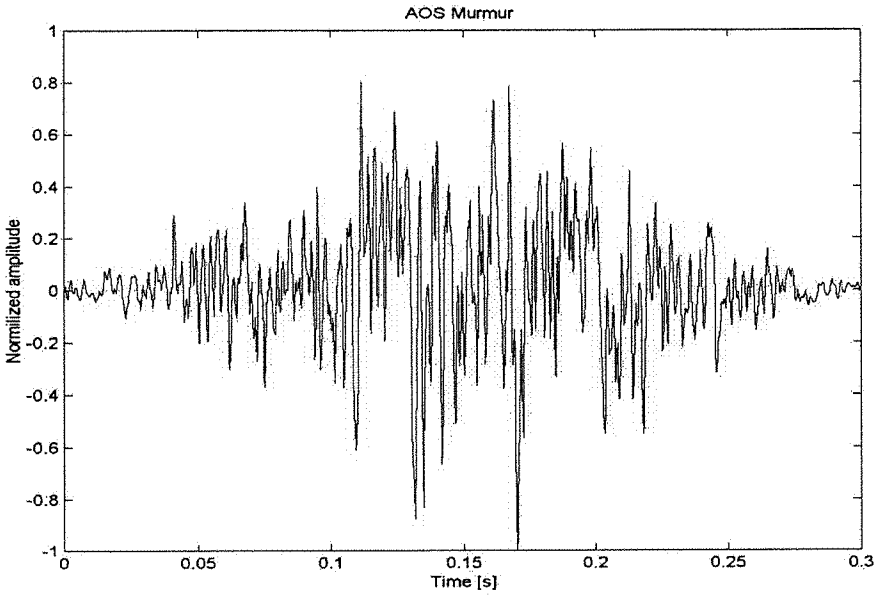


Figure B.2. Synthetic Aortic Stenosis Murmur.

Simulation of the second heart sound

Following the research of Leung (1998b) the two components of the second heart sound A2 and P2 are modelled using sinusoids and a Gaussian modulation function:

$$h(t) = \sum_{i=1}^{N} A_i e^{-(t-t_i)^2/(2\sigma_i^2)} \cos(2\pi f_i t - \phi_i)$$

Where h is the second heart sound, N is the number of components, A_i is the amplitude, σ_i is the width factor, t_i is the time delay, f_i is the frequency, and ϕ is the phase of the *i-th* sinusoid.

Figure B.3 shows the S2 generated using this mode. The components A2 and P2 are separated by 30 ms, the second component is half the amplitude of the first one, and the sinusoid frequencies are 170 and 100 Hz respectively.

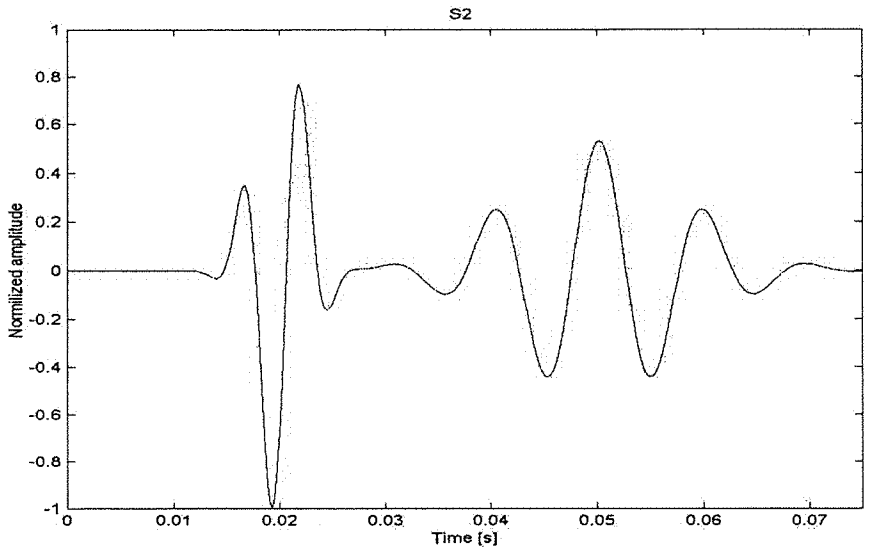


Figure B.3. Synthetic S2.

This model has been used to investigate the split in S2 for diagnosis of paediatric heart disease (Leung, 1998b).

Figure B.4. shows the synthetic PCG obtained adding all these components.

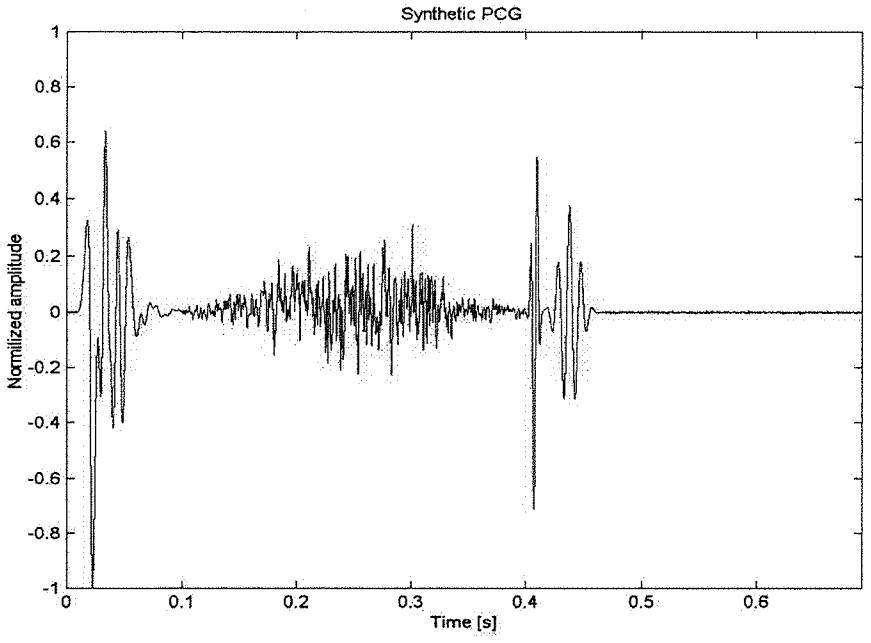


Figure B.4. Synthetic PCG.

Real signals

Real PCG signals form the database described in chapter 6 and 7 will be used to demonstrate the technique.

Cepstrum analysis

This section describes the steps involved in obtaining the cepstral representations used in the thesis. Synthetic signals and real signals are used to explain the concepts and to obtain critical parameters. In order to ease the interpretation, the analysis is described using one frame analysis first, and then the concepts are extended to short-time multiple window analysis.

One frame analysis

The PCG signal is a non-stationary signal and therefore, to obtain the cepstral representations presented in Chapter 5, short time analysis is required. This involves segmenting the signal into overlapped windows in which the signal is assumed to be stationary, and consequently the techniques developed for stationary signals can be used.

The next section shows the output at the various stages of the process of obtaining the power cepstral coefficients over a single time window. Once the basic concepts are explained, a posterior section will show the multiple window analysis.

Following the steps described in section 5.2.1 the PCG signal is first divided in frames of 20 ms and a Hamming window is applied.

The choice of window length will be justified later using real signals, but for now assume a 20 ms frame analysis as optimal. Figure B.5a shows the first 20 ms of

the synthetic PCG signal shown in figure B.4 whilst B.5.b shows the same signal but a Hamming window has been applied.

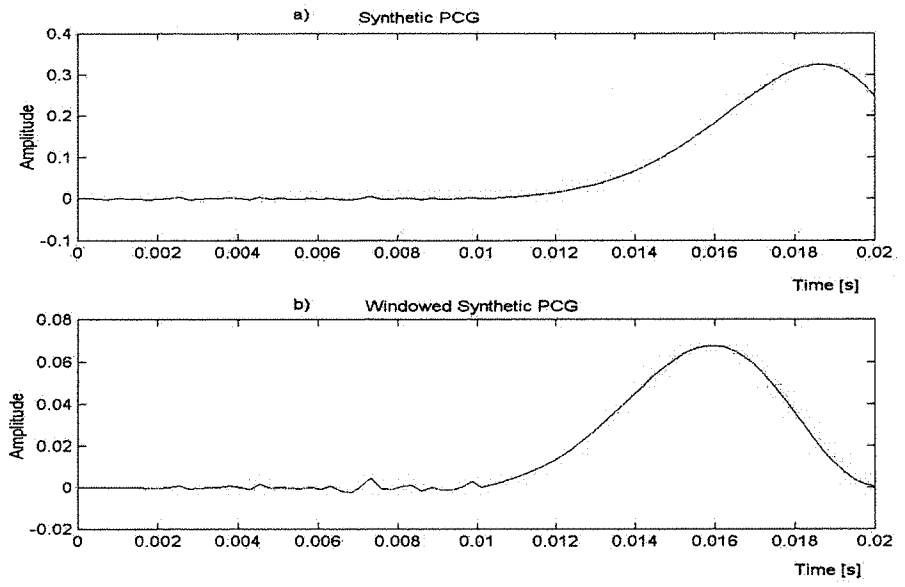


Figure B.5. 20 ms of the Synthetic PCG.
a) Raw Signal, b) Hamming Windowed Signal.

Figure B.6 show the spectrum of the signal shown in figure B.5.b, and the log spectrum computed using a discrete Fourier transform.

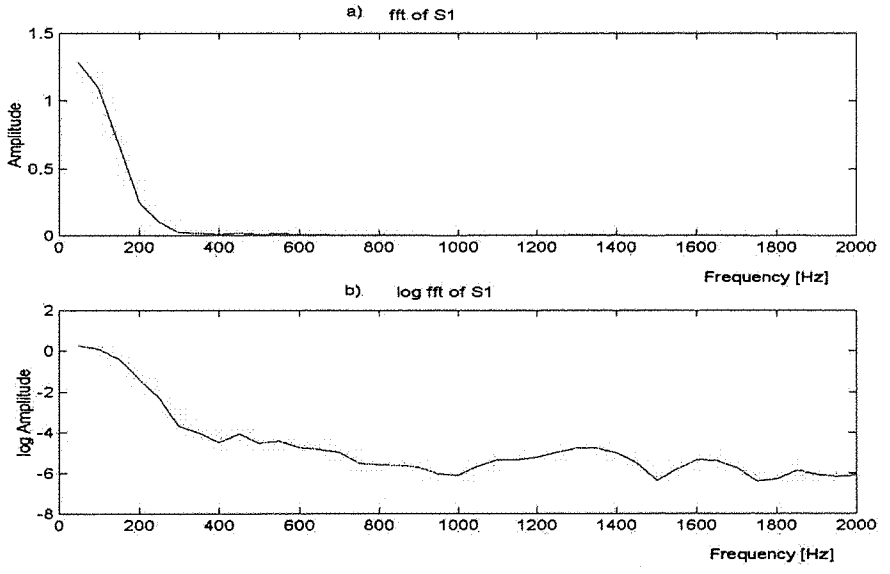


Figure B.6. Spectrum.
a) FFT, b) Log FFT.

An inverse Fourier transform applied to the signal obtained in B.5.b will provide the ceptrum shown in figure B.7.

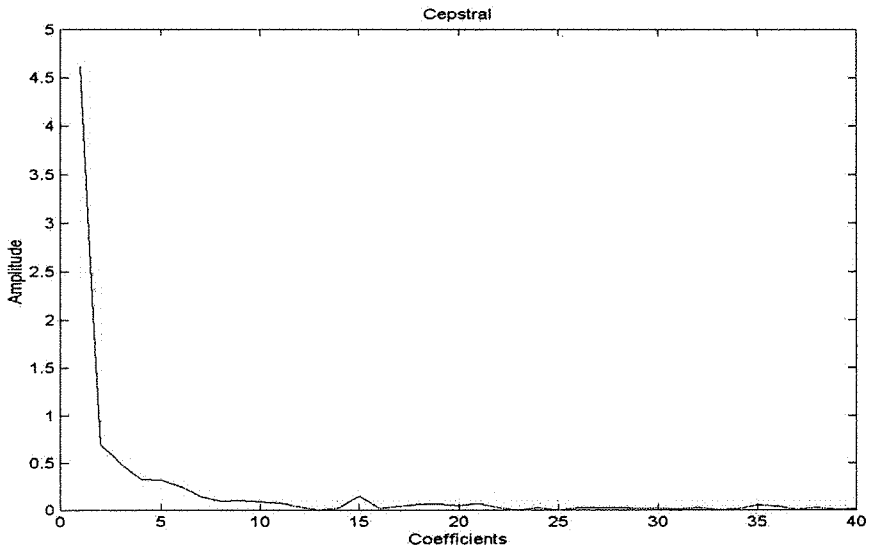


Figure B.7. Cepstral Obtained for the First 20 ms Window of the Synthetic PCG.

Figure B.7 shows the cepstral coefficients computed for the first window of 20 ms of the synthetic PCG signal.

Liftering

It is possible to go back one step in the process in order to understand how the cepstral coefficients are related to the spectrum.

Filtering the cepstrum (a process referred as liftering) can be applied to remove certain components and show the relative influence of the remaining coefficients. A simple filter is one which simply truncates the cepstral sequence by giving a weight of one to the low coefficients up to a certain index, and a weight of zero to the high order coefficients.

Figures B.8 and B.9. show the log spectrum computed from the low order coefficients.

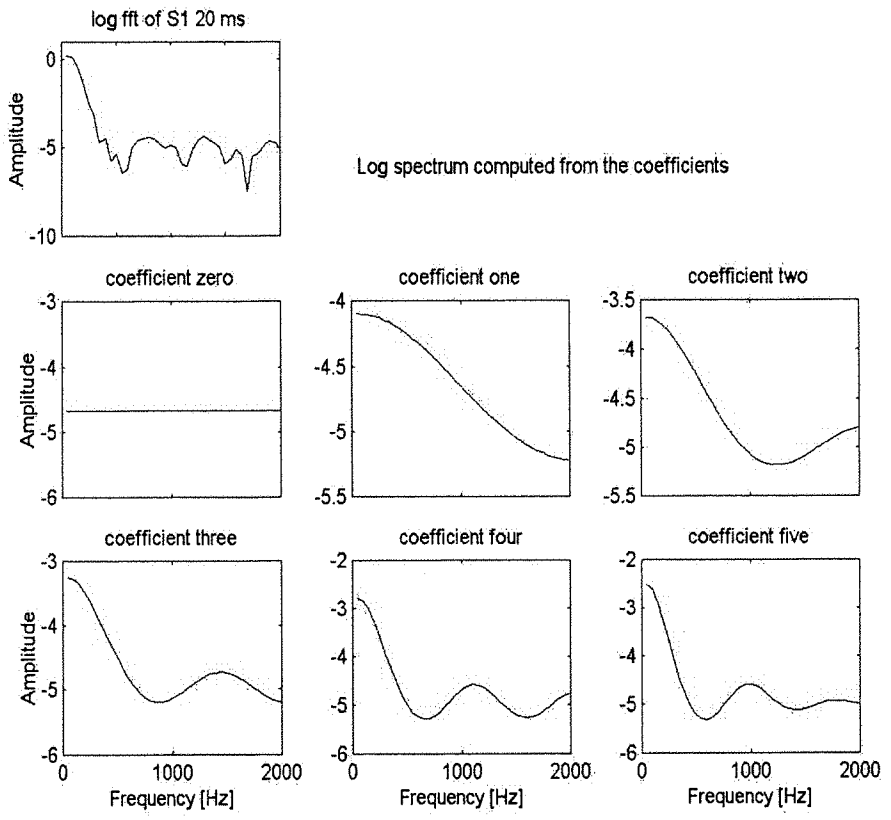


Figure B.8. Log Spectrum Computed From the Cepstral Coefficients.

The top left plot in figures B.8 and B.9 show the log spectrum obtained from the time signal. The next plots show the log spectrum obtained applying an inverse Fourier transform to the low order liftered cepstral coefficients.

Since the log magnitude spectrum is symmetrical, the FFT can be simplified to a discrete cosine transform. For a spectral representation with N channels with log magnitudes A_1 to A_N, the cepstral coefficients can be computed as follows:

$$c_j = \sqrt{\frac{2}{N}} \sum_{i+1}^{N} A_i \cos\left(\frac{\Pi j(i-0.5)}{N}\right) \qquad \text{for } 0 \le j < N$$

Figure B.8 shows that the low order cepstral coefficients provide a simplified log spectrum representation, i.e. a smoothed log spectrum. As the number of

coefficients included in the computation is increased more detail is added to the log spectrum shape. When j=0, the Equation simplifies so cepstral coefficient zero c_0 is a constant value equal to the mean of the log spectral signal computed from the time signal, and provides an indication of overall amplitude level for the frame.

Figure B.8 shows that five coefficients seem to provide a good representation of the overall shape of the log spectrum, specially comparing to figure B.9 that shows that doubling the number of coefficients does not provide much more detail. Note as a signal representation, the smoothing action on the spectrum also provides a degree of noise and artefacts immunity as shown in section 6.4.7. Depending on the application, a smoothed spectrum obtained from a few low order cepstral coefficients may be enough to represent the signal whilst in other applications a more detailed spectrum may be required.

In this thesis, initially, five cepstral coefficients were considered sufficient to represent the PCG for classification purposes, since the time evolution of these coefficients seems to provide enough detail to differentiate the pathological conditions (see figure 5.3 for example). From figure B.9 it can be see that the inclusion of more coefficients provides a more detailed representation, but the small database limited the numbers of features since a balance between sample size and feature representation has to be kept.

Multiple Frame Analysis

The procedure to obtain the power cepstral coefficients described before is applied to each window for the whole length of the PCG signals. A single value is obtained for each cepstral coefficient, per window.

The cepstral representations used in the thesis are therefore the time evolution of each of the cepstral coefficients (over all the windows).

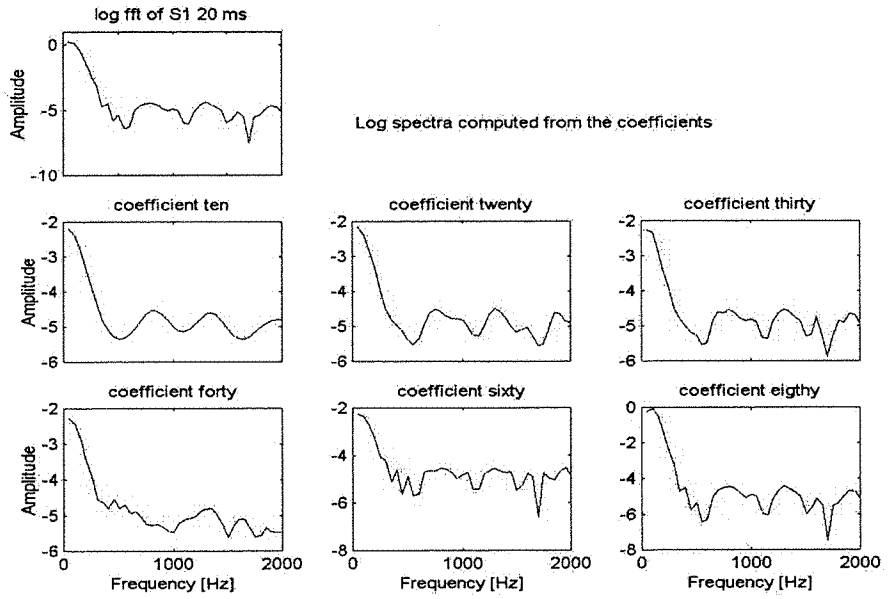


Figure B.9. Log Spectrum Computed From the Cepstral Coefficients.

Figure B.10 shows the time evolution of the cepstral coefficients (normalized) computed for the synthetic phonocardiogram. This representation is similar to figures 5.2, 5.3 or 6.2 for instance.

Window size considerations

Considering that the cepstrum is obtained from the spectrum, the literature related to the optimum window size for PCG analysis using this time-frequency representation was used as a guideline to select the window size for the cepstrurm.

Because of the non-stationarity of the PCG signal it is important to maintain an analysing time window as short as possible to guarantie the stationarity hypothesis over the analysed segments. On the other hand, a short time window will reduce the frequency resolution of the resulting spectrogram.

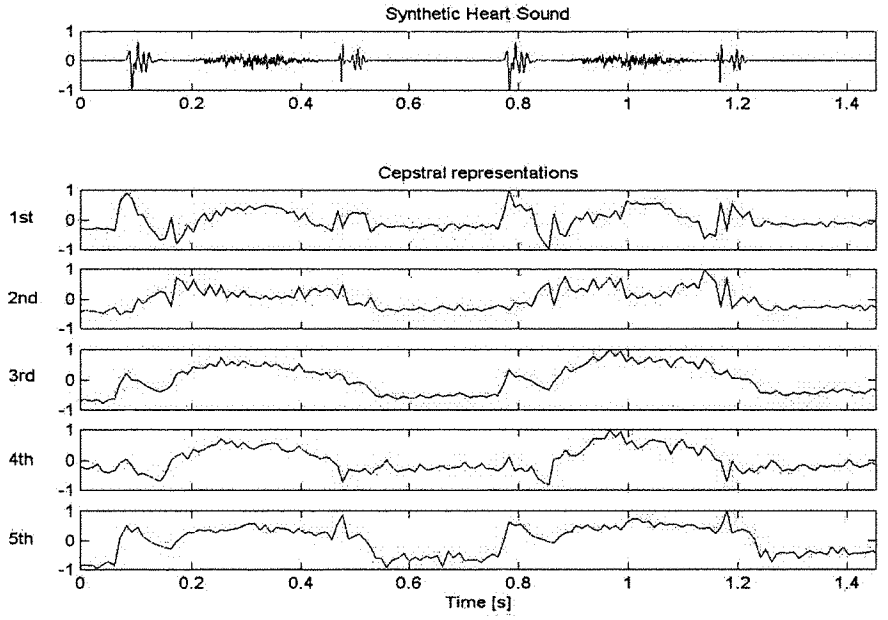


Figure B.10. Cepstral Representations of a Synthetic PCG.

The optimal duration of the time-window used to compute the time-frequency representation (spectrogram) of the phonocardiogram was studied by Jamous (1992) in four dogs by using intracardiac and thoracic measurements of the PCG. The power and cross-spectrograms of the intracardiac and thoracic PCGs were computed using a fast Fourier transform algorithm and a sine-cosine window with 10 per cent decaying time weighting functions. A coherence spectrogram was also computed for each dog to study the linear relationship between the two signals and determine the optimal time-window duration. Results show that the optimal range of the time-window duration is between 16 and 32 ms. A time-window shorter than 16 ms spreads out low-frequency components into the higher frequencies and generates a spectrographic representation with poor frequency resolution. A window larger than 32 ms increases the frequency resolution but smears the spectrographic representation of the signal in the time domain and thus cannot correctly reflect the time-varying properties of the signal.

Using synthetic murmurs, Debias (1997b) adjust the basic parameters of spectrograms, Choi-Williams, and Bessel distributions to provide the best time-frequency representations. The adjustment of the parameters is performed by

computing and minimising the relative averaged absolute error between the frequency contours at -3 dB and -10 dB of each time frequency representation of the simulated murmurs and those of the theoretical distributions found in their previous study (Debias, 1997a). They propose a 30 ms Hamming window as an acceptable compromise for the spectrogram to detect their simulated heart murmurs.

Different authors propose different window lengths for the computation of the spectrogram of PCGs (Djebari, 2000; Debias, 1979, Jamous, 1992) depending on which resolution is more important for their particular research.

Although this thesis focuses on the analysis of murmurs rather than the analysis of the second heart sound for diagnosis, the importance that the split of S2 has for the diagnosis of atrial septal defect through auscultation suggested that the PCG representations should consider the time occurrence of the A2 and P2 components of S2.

Under normal conditions, during expiration P2 appears from 10 to 30 ms after A2 (Guadalajara, 1998; Leung 1998b) and from 30 to 40 ms during inspiration. A time window of 20 ms, in the middle of this range, and within Jamous (1997) proposed range, was selected. Note that since windows are overlapped by 50 % this means that the analysis is actually performed every 10 ms.

Figure B.11 shows a PCG presents a ventricular septal defect form a 8 day old baby, recorded between the lower left external edge and the apex (HR = 126 bpm, Fs = 4kHz). Figure B.12 and B.13 show the surface plot of the spectrogram of S2 computed using a 20 ms window as proposed and a 32 ms time window.

In the spectrogram computed using a 20 ms window (Figure B.12) the two components of S2 namely A2 and P2 are clearly separated, whilst in the spectrogram of Figure B.13 computed using the 32 ms upper range value proposed by Jamous (1997) both components are smeared by the low time resolution.

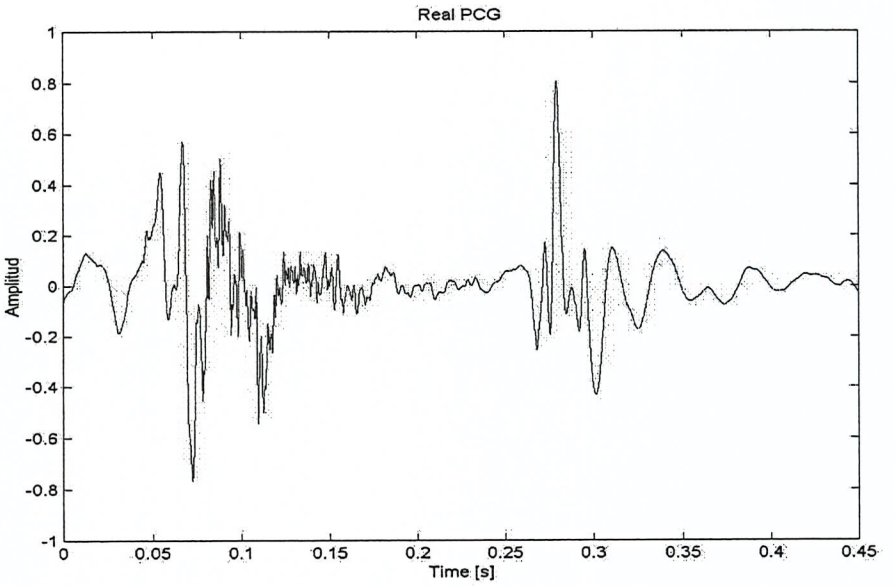


Figure B.11. Real PCG signal. Ventricular Septal Defect Murmur, HR = 126 bpm, Fs= 4 kHz, redorded at LLSE-Apex.

The 20 ms window size was used to run the algorithms of Chapter 6 and 7 for our database. The results obtained were compared to similar results obtained changing the window size around this value and within the range proposed by Jamous.

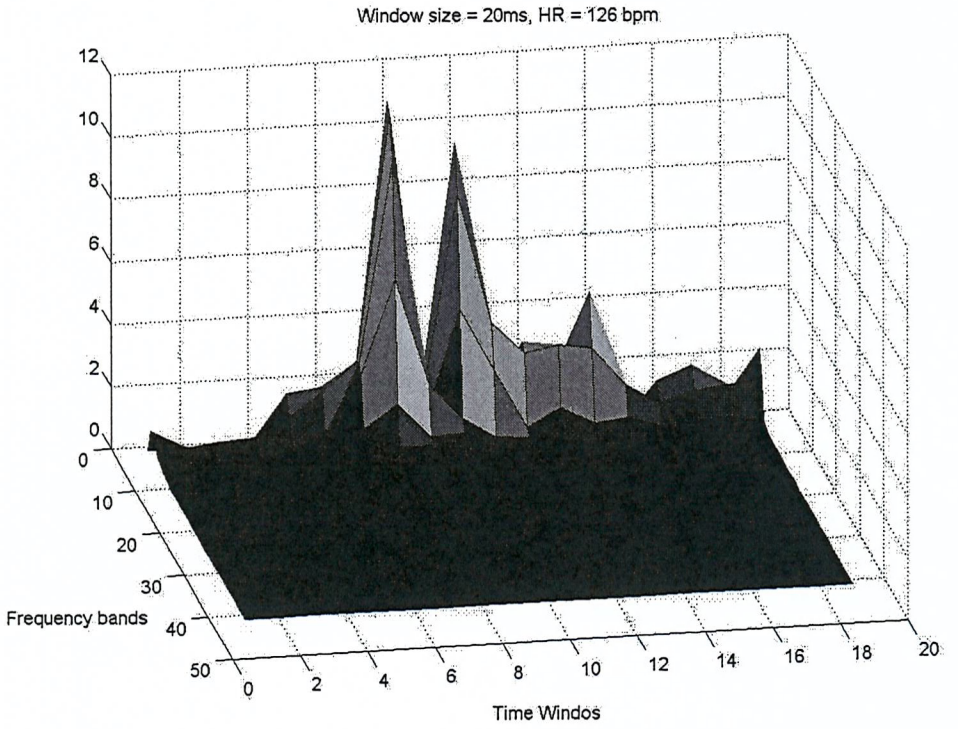


Figure B.12. Spectrogram of S2 Computed Using a 20 ms Window.

Since this value is just one of the many choices that have to be made in the implementation of the algorithm, this window size was kept because it provided good results.

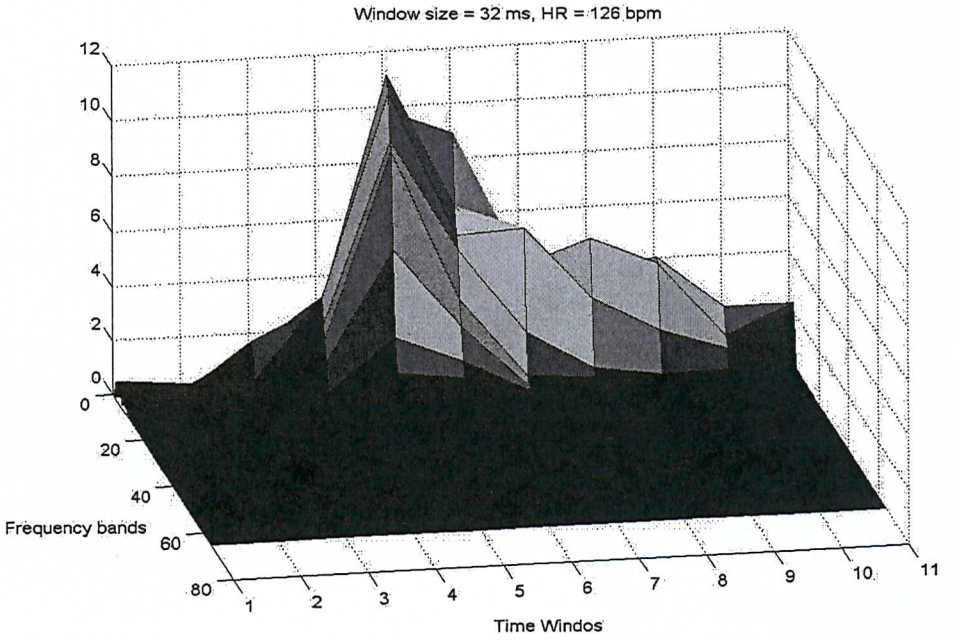


Figure B.13. Spectrogram of S2 Computed Using a 32 ms Window.

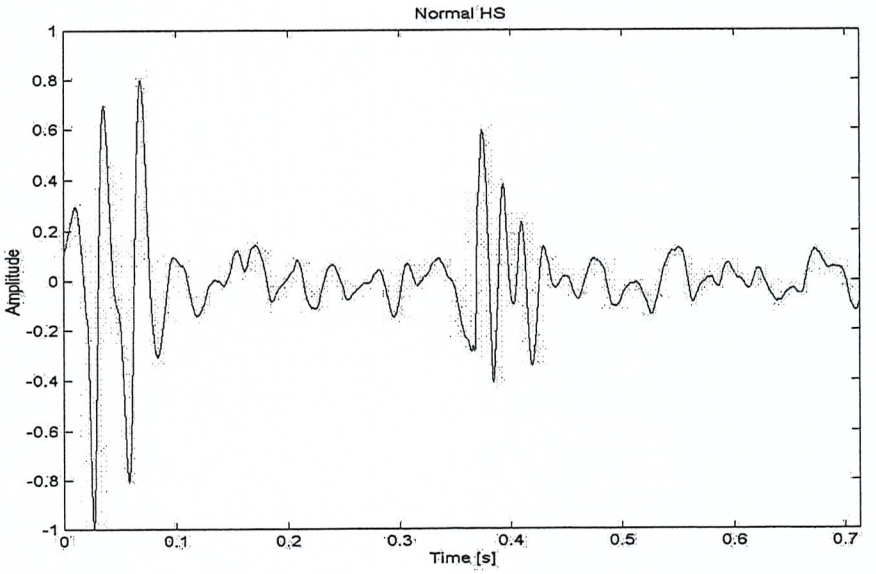


Figure B.14. Real PCG signal. Normal condition, HR = 60 bpm, Fs= 4 kHz, Recorded at Pulmonary Area.

Figure B.14 show a normal PCG recorded in the pulmonary area at a sampling frequency of 4 kHz.

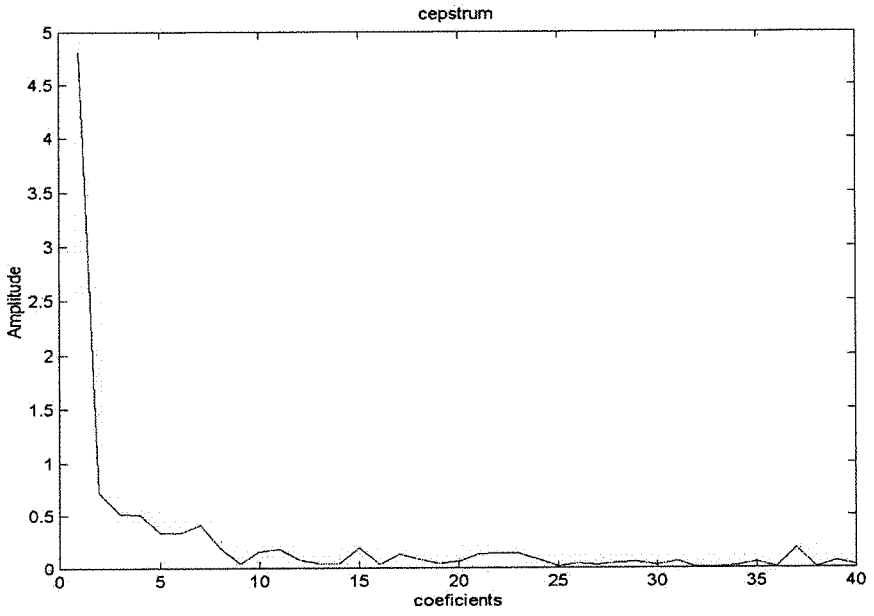


Figure B.15. Cepstral Obtained the First 20 ms Window of the Real PCG.

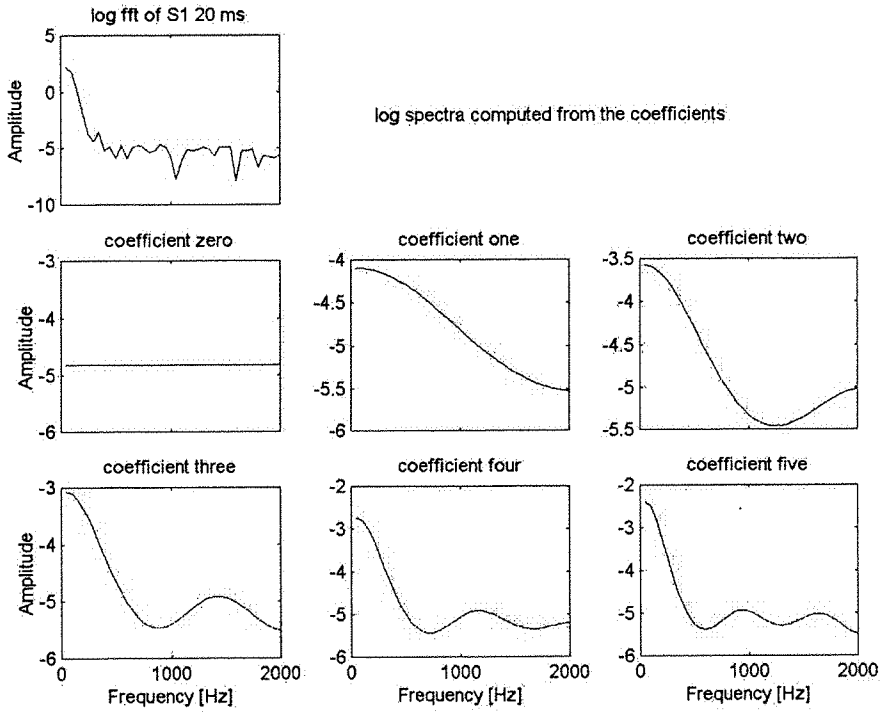


Figure B.16. Log Spectrum Computed From the Cepstral Coefficients. Real PCG.

Figures B.15 and B.16 show the cepstrum of the first 20 ms window, and the spectrum obtained from the first five coefficients. These figures are equivalent to the one frame analysis described before for synthetic signals.

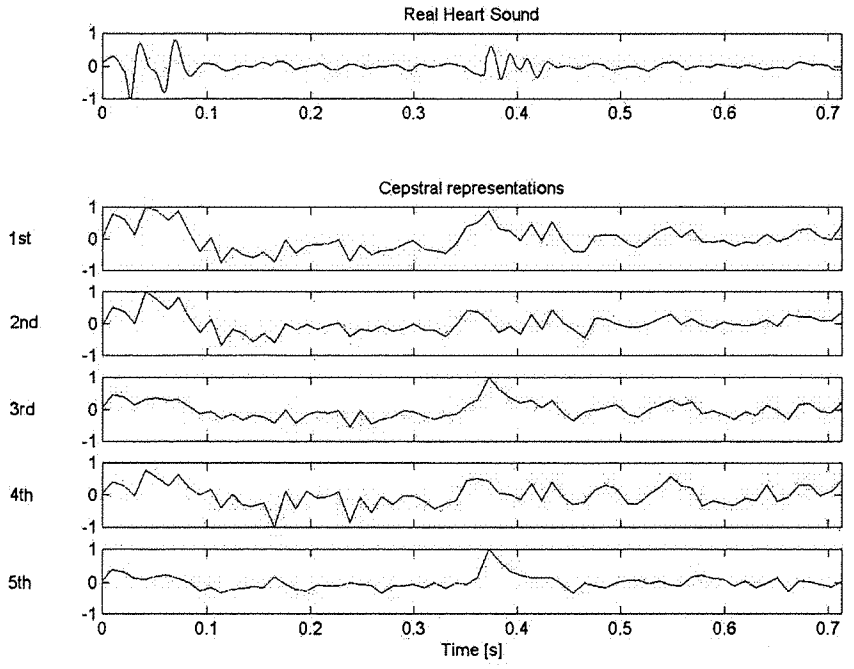


Figure B.17. Cepstral Representation of the Real PCG.

System response and excitation model

For the segmentation and classification of heart sounds a simplified representation of the PCG was required as input to the Hidden Markovian Models. Consequently, power cepstrum, that discards the phase information, was proposed over complex cepstrum since inversion to the excitation and system response time signals was not required and the phase complications of complex cepstrum made it unsuitable as a simplified representation (Rangayyan, 2002).

It is important to note that all that is required for the classifier is a signal representation that follows the changes in the PCG. There are clearly many possible signal feature that could be used, but in the current work cepstral coefficients are investigated

Representations derived from cepstral coefficients were proposed since the concept of cepstral analysis corresponds to the most accepted theory about the genesis of heart sounds. Separation of excitation from the system response is a theoretical capability of this technique (see assumptions discussed below), but the technique may also be useful when these assumptions are not justified. Nevertheless, the following discussion explores the considerations that may lead to the separation of system response and excitation in the cepstral domain.

Assumptions

The model presented makes the following assumptions:

1) The heart sound signal is assumed to be the convolution of a slowly time varying system response with a relatively fast varying excitation. Under these conditions the contribution of the system response to the cepstrum will be limited to the low order coefficients.

Oppenheim (1975) demonstrated this separation for speech signals assuming a slowly time varying vocal tract response with a relatively fast varying glottal pulse train. Following the same assumption of a slowly time varying system response Rangayyan (1978b) proposed to use this technique for PCG signals and extracts the system response sequence of the PCG using the low order coefficients (1978a). More recently Phua (2006) proposes the use of PCG cepstral representations for biometrics, the high order coefficients are also considered to be contributions from the fast time varying excitation.

2) Slowly time varying system response

The characteristics of vibrations can be described in terms of a mass supported by a spring. In an elastic chamber completely filled with fluid, the elasticity of the wall is analogous to the spring, and the fluid plus the supporting walls are

analogous to the supporting mass. The frequency at which a system vibrates depends upon the mass in motion in relation to its elasticity.

In the cardiohemic model the heart walls, blood, and valves vibrate as a whole. The character of the vibration is influenced by the nature of the specific cardiohemic system which is vibrating (Rushmer, 1970).

In the heart, the combined mass of the body and the walls of the chambers is very large in relation to the elasticity of the walls, so it tends to vibrate at low frequencies (Rushmer, 1970). When the ventricles are contracting, the elasticity of the heart walls should be greater and the vibration frequency is therefore increased.

The soft tissues of the body tend to damp the vibrations, and consequently heart sounds excited by sudden acceleration and deceleration of blood flow consist of relatively few vibrations. Nevertheless, when a turbulence flow is present the induced vibration persists as long as energy is supplied to the vibrating system (Rushmer, 1970).

Due to its inertial mass and elasticity the structure system response changes slowly to a fast varying flow excitation.

3) Fast varying excitation

The excitation in the cardiohemic model is related to sudden accelerations and decelerations of blood flow, conceptualized as short time fast varying impulses (compared to the slower time varying system response), or due to blood flow turbulence. In both cases the excitation is varying relatively fast compared to the system response.

The contribution to the spectrum from the excitation impulses can be considered flat for an ideal impulse since it is a spectrally rich signal containing all frequencies in equal amounts (Vasegi, 2000). For short impulsive forces their

spectrum falls to a low level at a frequency equal to the inverse of the pulse length (Fahy, 1998). Assuming the sudden accelerations and decelerations of blood flow to be impulses that last less than a millisecond, the amplitude spectra would spread up to a thousand Hertz. The frequencies of the normal heart sounds probably extend from below 20 Hertz to above 200 Hertz (Rushmer, 1970) and therefore the assumption of a flat spectrum over that range can be valid.

At a highly increased blood flow velocity red cells move at different velocities and directions creating blood flow turbulence (Guadlajara, 1998), this is reflected in a random excitation. A model based on a white noise random excitation force for turbulence (Debias, 1997) is assumed for the generation murmur generation.

It is important to note that the assumption of excitations with a flat spectrum is not required for the cepstrum to separate them from system response. Only a fast varying characteristic compared to a slow varying response of the system is required. Excitation signals that depart from this short time impulse and white random noise excitation assumptions (and therefore from a flat response) will have contributions on the low order ceppstrum coefficients, although the main contributions are still considered to be mainly due to the slow varying system response.

The mechanisms that lead to the genesis of the heart sounds are so complex that the basic nature and the contribution of each of its components is still unknown (Xu, 2000). Since it is not known what exactly generates heart sounds is not possible to demonstrate the validity of this model.

Representations in the frequency domain

Assuming the model presented in the previous section for the system and the excitation, in the frequency domain they will have the following characteristics.

1) System frequency response. The frequency response of this system is defined by its mass and elasticity, and can be represented as a cascade of damped resonators. The response below natural frequencies of the damped resonators is controlled by the static stiffness, and above these frequencies is dominated by the mass (Fahy, 1998). Due to its inertial mass and elasticity the structure system response does not changes abruptly, and therefore, the system response within the analysing window is smooth.

2) Excitation spectrum. The excitation spectrum is a rapidly changing flat spectra produced by white noise or short impulses

Consequently, if the system is set into vibration by an impulse or a random white noise wide band excitation, the envelope shape of the convolved signal in the frequency domain would be mainly defined by the resonances in the system's response since the excitation signals have a flat spectrum. The total response is therefore a wobbly version (due to the rapidly varying excitation spectral contributions) of the system's response spectrum.

Illustrating the Cepstrum domain and frequency domain relationship

The aim of the simulations was to show that the low order cepstral coefficients represent a smoothed spectrum, i.e. the over-all shape of the frequency response. As the high order cepstral coefficients are included, the corresponding spectrum includes finer detail.

Low order cepstral coefficients can be therefore related to the envelope shape of the spectrum and consequently related to the systems response whilst high order cepstral coefficients representing finer detail on the spectra can be related to the comparatively very short impulses generated by abrupt acceleration and deceleration of blood, or turbulence.

In order to show how the system response shapes the spectrum of the excitation, the synthetic signals proposed by Chen (1997) and Debias (1997) to simulate the system response related to the closure of the tricuspid valve and the model for the excitation due to turbulence, can be used.

Figure B.18 shows the tricuspid component of S1 as proposed by Chen (1997) and its log spectrum. Figure B.19 shows the excitation signal model for the turbulence as described by Debias (1997) and its log spectrum.

Figures B.20 shows the time signal result of convolving these system response and excitation signals and the log spectrum of the resulting signal. As shown in the figure the resulting log spectrum is mainly defined by the system response with random variations imposed by the excitation.

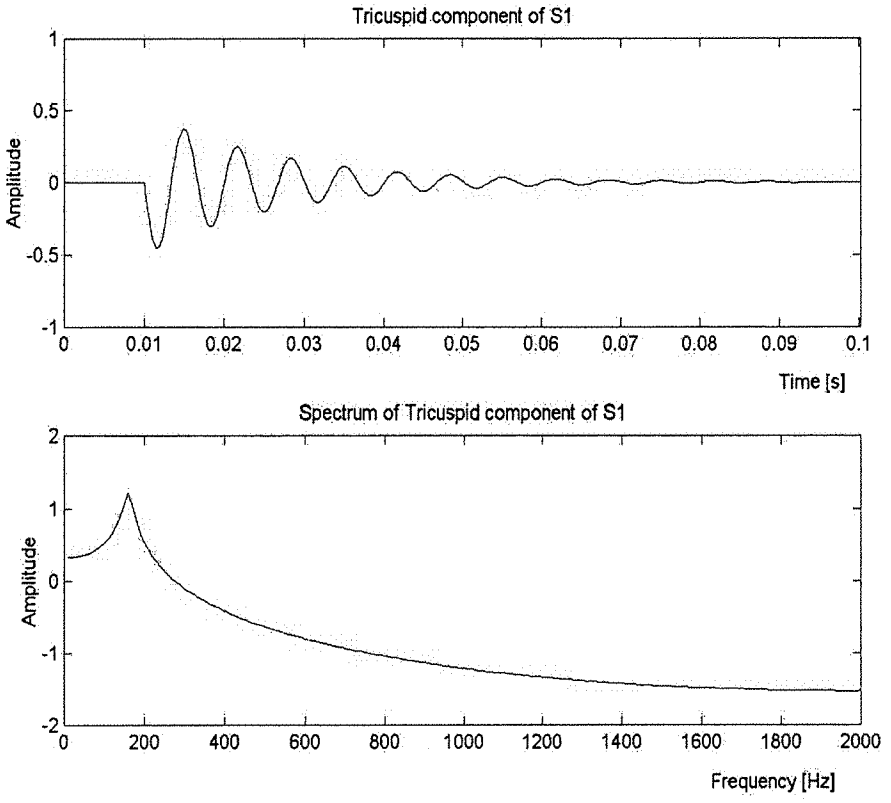


Figure B.18. Tricuspid Component of S1 and its Spectrum.

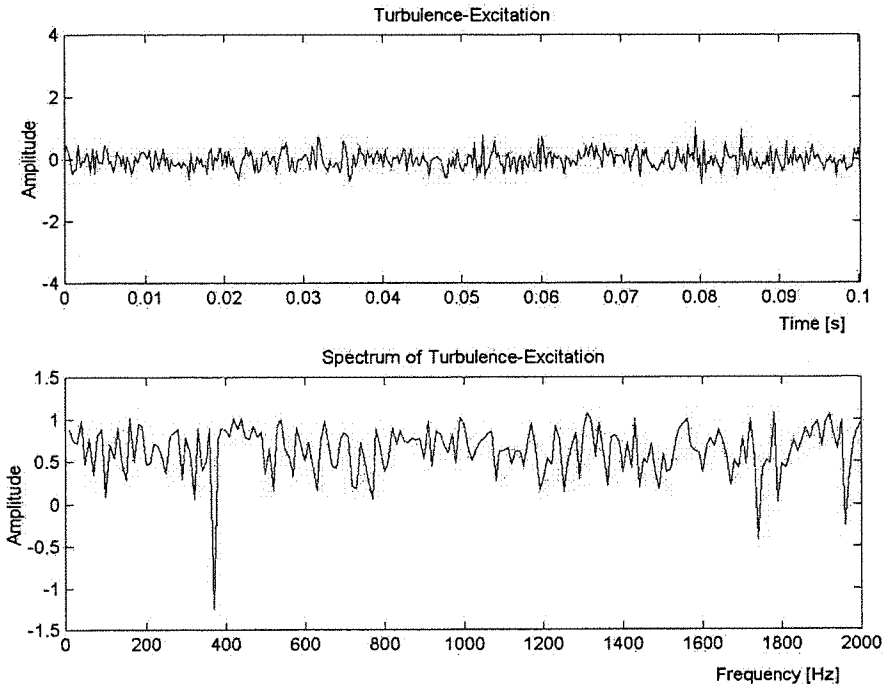


Figure B.19. Turbulence-Excitation and its Spectrum.

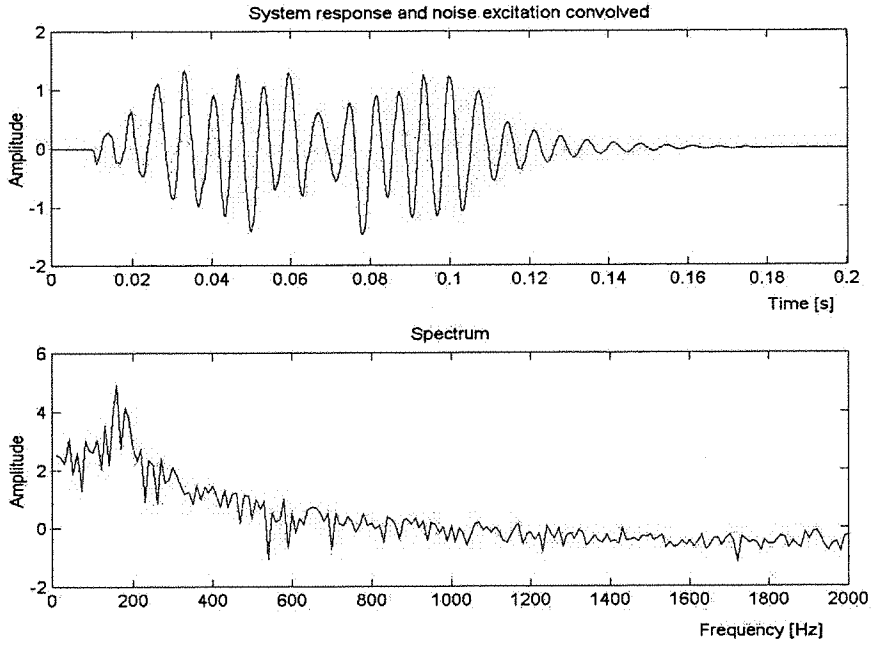


Figure B.20. System Response and Excitation Convolved and its Log Spectrum.

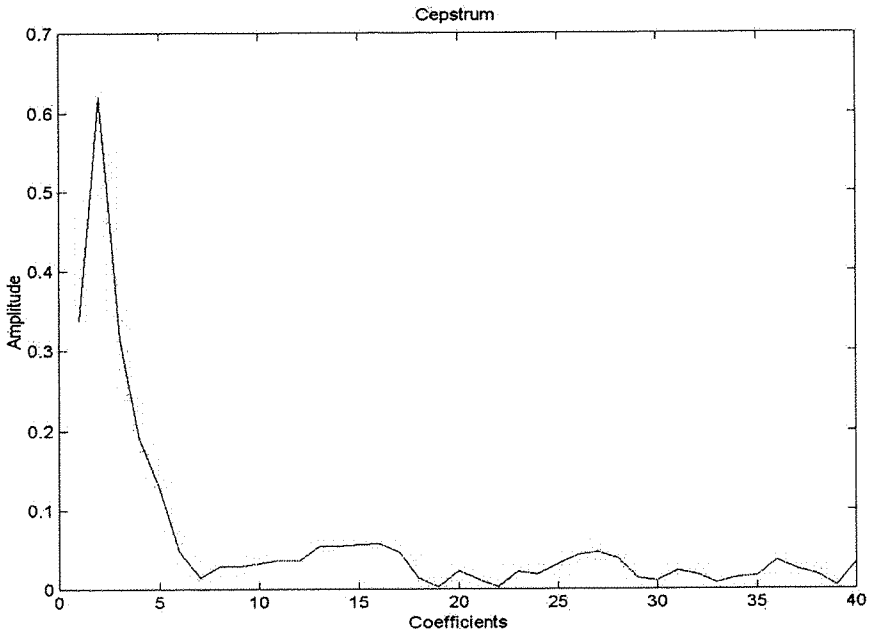


Figure B.21. Cepstrum of the Resulting Signal.

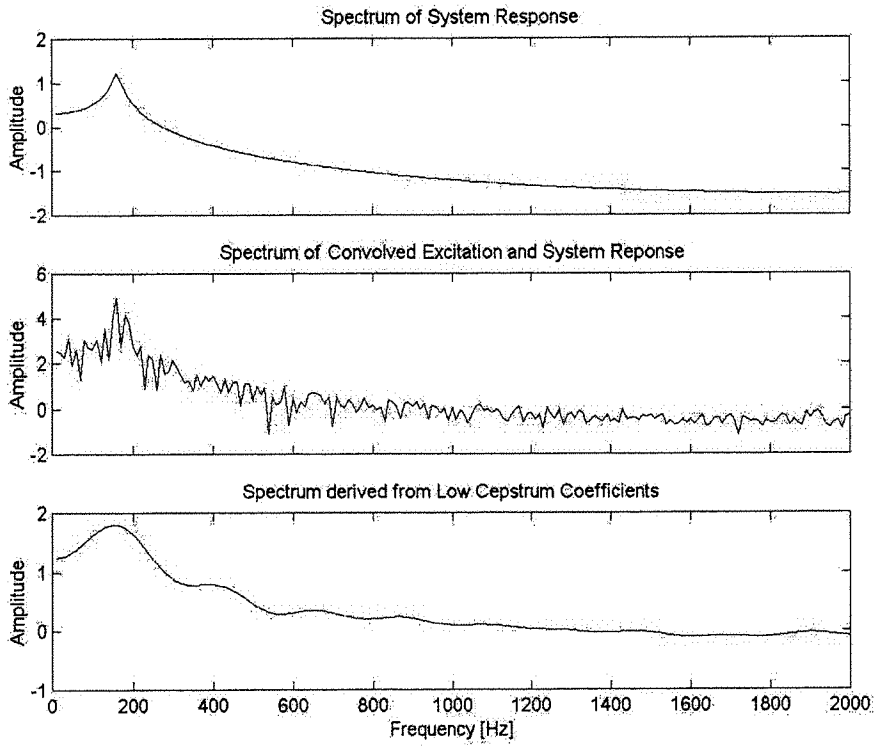


Figure B.22. Log Spectrum of the System Response, Convolved Signal and the log Spectrum Computed From the Low Order Cepstral Coefficients.

Figure B.21 shows the cepstrum of the signal result of the convolution of the system response and the excitation. Figure B.22 shows for comparison the log spectrum of the system response, the convolved signal and the log spectrum computed from the low order ceptral coefficients. The spectrum obtained using the first twenty cepstral coefficients clearly reflect the shape of the frequency response.