Methods for the extraction and classification of transient signals from noisy data - a case study in classifying sounds from the thorax

Thesis submitted for the degree of
Doctor of Philosophy
at the University of Leicester

by

Tuan Tran
Department of Engineering
University of Leicester
1994
It was the best of times, it was the worst of times, ...

by Charles Dickens.
Methods for the extraction and classification of transient signals from noisy data - a case study in classifying sounds from the thorax

by

Tuan Tran

Declaration of Originality

This thesis is submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in the Department of Engineering, The University of Leicester, U.K. All work recorded in this thesis is original unless otherwise acknowledged in the text or by references. No part of it has been submitted for any other degree, either to the University of Leicester or to any other University.

Tuan Tran

June 1994
Acknowledgements

There are many people whom I have encountered during the three years of my Ph.D. Before I start to acknowledge various people I would like to express my gratitude to my supervisors. A very big thank you to Professor Barrie Jones who made this project possible. A special thank you to Dr. John Fothergill for being there when there's no body around. Without their experience, guidance and support finding the right path to tread would have been a difficult task. Thanks are also due to my supervisors who have been and gone; Dr Sehmi, Michael Judge and Alastair Ruddle. I also wish to thank Dr Michael Pont and Dr Fernando Schlindwein for their expert advice.

For my friends I can only say thanks for the enlightening tea break, the nightmares, those wild nights out and, last but not least, the constant encouragement (more like aggravation) to get this thesis finished. Not mentioning your names could be a mistake, for life, so here you are guys: Salih, Gareth, Sangeet, Jian Tao, S.Q., Yeuhe, Man Ho, Patrick, Paul, Dave, Anthony, James, Ghassan, Chris, Tim, Neale, Lun, Abdul, Pete, the two Richards, Phil, Trupti and Andy.

A special thank you is reserved for my mum, brother and sister, uncles, aunts and cousins for their support and encouragement. Being so far from home, you are the only family that I have ever known. Without you those rare moments would never have happened.

Finally I would like to thank Phoebe the only person who had made a permanent impression in my life. Without whom life would have been very much different.
I would like to dedicate this thesis to my parents
Contents

Abstract

Chapter 1: Acoustic emissions from the human body
  1.1 Introduction ................................................................. 1-1
  1.2 The heart ................................................................. 1-2
  1.3 The lungs ................................................................. 1-8
  1.4 Research in chest sounds ........................................... 1-13
  1.5 Discussion ............................................................... 1-17
     References ..................................................................... 1-18

Chapter 2: Characterisation of sounds
  2.1 Introduction ................................................................. 2-1
  2.2 Characteristics of sounds ........................................... 2-1
  2.3 Speech sounds and thorax acoustics ............................ 2-9
  2.4 Methods of processing .............................................. 2-9
  2.5 Discussion ............................................................... 2-13
     References ..................................................................... 2-14

Chapter 3: Experimental & simulation techniques
  3.1 Introduction ................................................................. 3-1
  3.2 A bench top transient simulator ................................. 3-1
  3.3 A heart sound simulator ............................................. 3-12
  3.4 Discussion ............................................................... 3-30
     References ..................................................................... 3-31

Chapter 4: Extraction of chest sounds
  4.1 Introduction ................................................................. 4-1
  4.2 Extraction technique ................................................. 4-2
  4.3 Low pass filtering ..................................................... 4-8
  4.4 Signal's activity ......................................................... 4-13
  4.5 Signal's energy ......................................................... 4-14
  4.6 Short-time average zero crossing rate ....................... 4-20
Chapter 5: Classification of chest sounds

5.1 Introduction ................................................................. 5-1
5.2 Transients and their features ........................................... 5-1
5.3 Nearest neighbour clustering ......................................... 5-4
5.4 Fuzzy logic ..................................................................... 5-9
5.5 Fuzzy min-max clustering ............................................... 5-19
5.6 Discussion ..................................................................... 5-23
References ......................................................................... 5-25

Chapter 6: A chest sound analyser

6.1 Introduction .................................................................... 6-1
6.2 Overall system .............................................................. 6-1
6.3 Heart sounds ................................................................. 6-3
6.4 Lung sounds ................................................................. 6-20
6.5 Discussion ..................................................................... 6-28

Chapter 7: Conclusions and further work

Appendix A

Appendix B
Abstract

The physiological origins and physical characteristics of sounds from the thorax have been reviewed briefly. This thesis presents some signal processing algorithms and classification techniques which have been developed for the extraction and classification of those sounds.

In order to evaluate the recording equipment and signal processing algorithms, two simulators were constructed: a laboratory simulator generating lung sounds in a variable background noise environment and a heart sound simulator written such that the generated sound was defined by a set of variables.

Four conventional transformation algorithms for the transient extraction process were evaluated. Their considerable user intervention and inconsistent transformed signal led to the development of the "signal's envelope" algorithm.

The signal's envelope method was used to extract transients of interest which were then used for the classification stage. It is shown that, due to the numerical nature of the features used for the classification process, the nearest neighbour clustering algorithm could not correctly classify all the extracted transients. The numerical features were therefore converted into linguistic terms and a fuzzy logic technique was developed to classify the transients. The fuzzy inference engine was robust enough to cope with the small numerical variation in features such that the correct classification was achieved. The other classification method tried was the fuzzy "min-max" clustering algorithm. This also used numerical features for the classification process and was therefore not able to classify all of the extracted transients correctly.

A lung sound analyser was constructed using the signal's envelope and fuzzy inference engine. The system was able to extract and classify individual heart sounds, crackles and wheezes from recorded phonograms. In about 4% of cases, the heart sounds were so indistinct that only a partial classification was achieved.

It was concluded that by using simple transducers and sophisticated signal processing and classification algorithms, it was possible to construct a chest sound classifier which may be of use in a clinical environment.
Chapter 1

Acoustic emissions from the human body

1.1 Introduction

Clinical examination consists of four stages: inspection, palpation, percussion and auscultation (Macleod, 1983, Swash & Mason, 1984). A physician would use such a clinical examination in an attempt to decide upon the state of the patient's health. The inspection, palpation and percussion processes can be easily achieved without the aid of any instruments, in contrast to auscultation. Auscultation is the diagnostic technique used in medicine to listen to the various sounds in the body. These sounds arise from the chest, the heart, arteries, the thyroid gland etc (Macleod, 1983, Swash & Mason, 1984). Ever since the invention of the stethoscope by Laënnec, the role of auscultation in clinical diagnosis has increased greatly (Leatham, 1981). This is because the technique is noninvasive and does not require expensive equipment. Although the modern stethoscope is substantially different from the one used by Laënnec, the basic working principle remains the same. Its function, when examining the chest for example, is to transmit the entire range of chest sounds from the chest piece to the ear and amplifying the sounds during the transit. This amplification process makes the frequency response of the stethoscope uneven (Abella et al, 1992, Ertel et al, 1971), that is some frequencies are amplified while others are attenuated. For this reason modern stethoscopes have a selection of chest pieces; low frequencies can be best heard with the bell chest piece and high frequencies with the diaphragm chest piece. However, both the stethoscope and the
ear have their limitation as instruments for evaluating heart and respiratory sounds, the
two areas that are most often auscultated. In view of this limitation and the importance
of those sounds, this thesis presents some signal processing techniques and recording
equipment to aid the auscultation process.

In this chapter the physiology of the heart and lungs are briefly described.
Also in this chapter the theory of the generation process and characteristics of heart and
lung sounds are outlined.

1.2 The Heart

1.2.1 Physiology

The heart is the central organ in the human body, Figure 1-1. Its purpose is
to circulate blood throughout the body (Guyton, 1991). It consists of four chambers: two
atria and two ventricles. The atria act as blood reservoirs and as an entry way to the

Figure 1-1 The heart, from Guyton (1991).
ventricles. They also pump blood at low pressure to the ventricles. It is the ventricles that provide the main force for circulation. Within the heart there are four valves: two atrioventricular (A-V) valves (the tricuspid and the mitral valves) and two semilunar valves (the aortic and the pulmonary valves). The purpose of these valves is to prevent backflow of blood during systole or diastole. All of these valves operate passively, i.e. they close when there is a pressure gradient pushing the blood backward, and open when there is a pressure gradient pushing the blood forward.

The heart consists of three types of cardiac muscle: atrial muscle, ventricular muscle and special excitatory and conductive muscle fibres. The atrial and ventricular muscles contract in the same way as skeletal muscle, but the duration of the twitch contraction in these cardiac muscles is three to fifteen times longer than of skeletal muscle. The specialised excitatory and conductive muscle fibres force of contraction is much weaker than atrial and ventricular muscles, but they exhibit rhythmicity and can conduct electrical impulses to the rest of the heart. Their role in the heart is to generate excitatory impulses, which are transmitted to the whole muscle mass thus leading to heart beats.

1.2.2 Heart Sounds

The origin of heart sounds has been a controversy subject since the publication of De L’auscultation Médiate by Laënnec. Initially, it was suggested that heart sounds were caused by the vibration of the valves closing (Dock, 1933). This has been shown to contribute little if any sound due to the damping effect of the blood. Instead, the heart sounds come from the vibration of the taut valves immediately after closing along with the vibration of the adjacent blood, the heart wall and the major blood vessels around the heart (McCall & Price, 1967, Waider & Craige, 1975). The events which lead to the generation of heart sound are summarised in Figure 1-2.
Figure 1-2 Events of the cardiac cycle, from Guyton (1991).
In a healthy person there are usually only two heart sounds: the first and second heart sounds (Guyton, 1991). Referring to Figure 1-2, the first heart sound is the result of the closure of the A-V valves at the beginning of systole, as the ventricles contract. This produces a back flow of blood towards A-V valves causing them to snap shut. Due to the closing of the A-V valves the blood pushes against the valves causing them to bulge toward the atria. The chordae tendineae together with the elasticity of the valves stop the bulging and cause the back flow to bounce back into the ventricles. This effect causes the blood, the ventricular walls and valves to vibrate. This vibration then travels through the adjacent tissues to the chest wall, where it can be heard as sound using a stethoscope. The second heart sound is generated in a similar fashion to the first heart sound. It involves the closure of the semilunar valves and the bulging of the valves into the ventricles. When the blood recoils into the arteries it causes a short period of reverberation of blood back and forth between the arteries and the valves and also between the valves and the ventricular walls. These vibrations are then transmitted along the arteries before reaching the chest wall. This mechanism of heart sound generation is similar to fast flowing water along a pipe. If the flow of water is suddenly interrupted by the closure of a tap the water pushes against the valve and recoils along the pipe and causes reverberation ("hammering"). This reverberation is then transformed into sound waves.

There are some other heart sounds which can be heard in a diseased patient.

- **Third and fourth heart sounds.** These sounds are caused by abnormal filling patterns of the ventricles. In some cases the third heart sound can sometimes be heard in young healthy adults and also in pregnant women (Macleod, 1983, Swash & Mason, 1984).

- **Snapping and clicking of valves.** These are high-pitched sounds that arise from patients with abnormal valves.
- Murmurs. These are caused by turbulence in the blood at or near the valves due to a structural disorder of the valves. There are several types of murmurs and each one is associated with a particular valvular disease (Macleod, 1983, Swash & Mason, 1984).

Auscultation of the heart involves listening to heart sounds with the aid of a stethoscope (Leatham, 1975). The sounds from all the valves can be heard on the chest wall but individual valvular sounds can be located to the areas illustrated in Figure 1-3. These areas are customarily called by the name of the valves from which the sounds arise (Guyton, 1991).

![Figure 1-3 Positions for listening to heart sounds, from Guyton (1991).]
Figure 1-4 Graph of heart sounds and the audible range for the human ears. (0dB = 20\mu Pa peak pressure). Modified from Butterworth et al (1964) and Capel (1991).
When listening to heart sounds with a stethoscope the human hearing system is restricted in both the frequency and the loudness of heart sounds (Butterworth et al, 1964). Nevertheless, with the aid of a microphone enclosed in some sort of casing to reduce external noise a recording of heart sounds can be made (Leatham, 1952). These recordings are called phonocardiograms and they have been used to show that the frequency and loudness of heart sounds extend beyond the human’s audible range, Figure 1-4.

1.3 The Lungs

1.3.1 Physiology

Figure 1-5 Thoracic cage and the lung lobes, from Romanes (1986).
The lung is divided into lobes, three on the right and two on the left (Brewis, 1975). These lobes are enclosed in a thoracic cage which is formed by the vertebral column, the ribs and the intercostal spaces on either side, Figure 1-5. This thoracic cage protects the lungs and the heart and provides attachment for muscles of the thorax, abdomen, back and upper extremity. Covering the lungs is a thin layer of membrane called the visceral pleura, which passes from each lung at its roots (ie. where the main air ways and blood vessels enter the lung) to the inner surface of the chest wall, where it is called the parietal pleura, Figure 1-6. This arrangement of the membranes forms two pleural cavities, one on each side of the thorax, between the lungs and the thoracic walls. The pleural cavity is a completely closed space and contains a small quantity of serous fluid that ensures frictionless movement between the lungs and the thoracic walls when a person inhales and exhales.

Figure 1-6 Section through an intercostal space, from Snell (1981).
Within the lung lobes there is a network of pulmonary vessels and airways. The airways are reinforced with cartilage so that they do not collapse as a result of the sudden change in pressures during inspiration and expiration. These airways terminate in air filled sacs called alveoli. The walls of the alveoli are formed by a layer of epithelium which is continuous throughout the bronchial tree and it is all that separates the air in the alveoli from the pulmonary capillaries. This thin layer of cells allows rapid exchange of oxygen, from the air, and carbon dioxide, the end product of metabolism from the cells, to take place. When the alveoli are filled with freshly inhaled air, their concentration of oxygen is greater than that in the blood. On the other hand, the blood has a higher concentration of carbon dioxide than that in the alveoli. Due to this concentration gradient the carbon dioxide diffuses from the blood to the alveoli and oxygen from the alveoli diffuses to the blood stream. The carbon dioxide in the alveoli diffuses to the upper airways and is expelled into the surroundings during expiration. Clearly, any lung disorder that reduces the area for gas transfer, eg. emphysema, or which prevents the access of air to the alveoli, eg. asthma, will reduce the efficiency of the gas exchange process. Even if the gas exchange process and ventilation are unimpaired, the same outcome may arise from a decrease in the flow of blood, eg. heart failure, or the oxygen-carrying capacity of the blood, eg. anaemia.

1.3.2 Lung Sounds

Auscultation of the chest in a normal healthy person produces the so called vesicular breath sounds which are produced by the flow of air in and out of normal lung tissue and can be heard all over the chest wall (Macleod, 1983, Swash & Mason, 1984). Throughout inspiration the sound is fairly intense and of low frequency with a characteristic rustle. There is no pause before the expiratory sound, which is heard only in the early part of expiration.

In a diseased patient adventitious sounds may arise in the lung or in the pleura. At the beginning of respiratory medicine there were no accepted terminologies
for the adventitious sounds arising from the thorax. When Laënnec published his observations on auscultation he proposed that these lung sounds should be called râles.

Over the years, physicians have tried to describe what they have heard and as a result there were many terminologies to describe one particular sound (Murphy, 1981, Robertson & Coope, 1957). These terms were not to the satisfaction of many physicians, so they added their own observations and suggested amendments and the terms became more and more confusing. In 1957 Robertson and Coope proposed that the lung sounds be divided into two groups: continuous sounds eg. wheezes, and interrupted sounds eg. crackles. These two groups give a more precise acoustic description of the lung sounds and they have been adopted in recent papers on lung sounds (Benedetto et al, 1983, Chowdhury & Majumder, 1982, Fredberg & Holford 1983, Loudon & Murphy, 1984, Nath & Capel, 1974).

- **Wheeze**es are prolonged uninterrupted sounds. They arise from the bronchi and are due to the partial obstruction of the airways by mucus, by swelling of the mucosa or by the constriction of the bronchi (Forgacs, 1967 and 1978). The waveform of the sound is often irregular but it can be broken down into several superimposed sine waves, Figure 1-7. The frequency of the wheeze depends upon the diameter of the bronchi. For example, low frequency indicates the oscillation of mucus in a large airway.
Low pitched (100 Hz) 10 ms

High pitched (700 Hz) 10 ms

Figure 1-7 Two waveforms of wheezes of different frequencies, from Forgacs (1978).

Figure 1-8 Phonopneumogram of late inspiratory crackles and air flow rate, from Nath and Capel (1974).
Crackles or crepitations are discontinuous or bubbling sounds, Figure 1-7. They arise from the alveoli, the bronchi and sometimes the cavities. There are several types of crepitations and they may be heard at any time during the respiratory cycle, (Forgacs, 1967 and 1978, Macleod, 1983, Swash & Mason, 1984). Fine crepitations are thought to be due to the presence of fluid in the alveoli, ie. in patients at early stages of pneumonia. Coarse crepitations are caused by secretions in large airways such as the bronchi, ie. in patients with bronchitis.

Friction sound or pleural rub is a result of abrasion of inflamed and roughened pleural surfaces. It has a creaking or rubbing characteristic sometimes quite similar to crepitation. The friction sound may be fine or coarse. The main differences are that pleural friction sounds occur at a particular part of the inspiration phase, when the roughened surfaces are rubbing against each other, and the sounds reappear at a corresponding cycle of expiration.

There are a few similarities between lung and heart auscultation. These include the human hearing system and the equipment used during the diagnostic process. Recorded lung sounds, using microphones enclosed in a cup, are called phonopneumograms (Homma et al, 1985). These are shown in Figures 1-7 & 1-8.

1.4 Research in Chest Sounds

From the first few sections it can be seen that acoustic emissions from the thorax can be a rich source of information to aid the diagnostic process. Therefore auscultation represents a powerful noninvasive tool that can be used in combination with other information to accurately diagnose thoracic diseases. Given its effectiveness as a tool many papers have been published since its inception. With consideration to some of the numerous papers that have been published, this section will attempt to discuss some of the research areas that have been explored in the field of auscultation.
The early phases of respiratory medicine were mainly concerned with terminologies. This was because physicians not only differed in the way they interpreted the meanings of the terms, they also had different perceptions of the sounds heard. This difference in perception was at least in part due to the variation in the human hearing ability and the distortion caused by the stethoscope. An additional problem that the physicians came across, which impeded the advance of auscultatory medicine, was the lack of understanding of the mechanisms and sources of the generated sounds. Eventually the mechanisms and the sources of respiratory and heart sounds were proposed but research is still continuing up to the present time.

Most of the published papers concerning thoracic diseases describe the sounds heard during examination of diseased patients. With the development of the microphones, sounds from the chest can be recorded and studied. The following are some examples of such work.

To understand the transmission and attenuation of respiratory sounds Goncharoff et al, 1989, and Wodicka et al, 1990, investigated sound transmission in the human respiratory system. In both experiments, acoustics generated by a white noise generator were transmitted into the respiratory system by way of the mouths of the subjects. Wodicka transmitted sound waves with a white noise characteristic between 100Hz to 1000Hz and measured the acceleration on the chest wall at three different locations; trachea, and 5 cm lateral to the centre of the spine at the third and sixth thoracic vertebrae. He observed that there is a decrease in acceleration at the third and sixth thoracic vertebrae compared to the trachea position, decreasing chest wall acceleration with increasing frequency and an increase in transmission at two particular regions of frequency. The experiment showed that the lungs acted as a good absorber of high frequency sounds. Goncharoff performed a similar experiment using microphones, an acoustic frequency range from 5-20kHz and four different groups of people. He concluded that the human lungs have different sound transmission characteristics depending on their physiological status.
The recording of air flow of the subject under study could be useful when analysing respiratory sounds. Mussell et al, 1990, investigated the effect of the air flow transducer on the recorded respiratory sounds. The study was carried out on twelve normal adults and it was concluded that the air flow transducer significantly distorted tracheal breath sounds. The measurement of chest wall movement by calibrated strain gauges to give air flow rate can therefore be a useful alternative technique to an air flow transducer.

A phonopneumograph was constructed by Homma et al, 1985. The device was able to record lung sounds and respiratory flow rates on a thermal printer and oscilloscope simultaneously. These recordings were then interpreted by "trained" and "untrained" personnel. The results suggested that the phonopneumograph can be a useful device for the interpretation of lung sounds as well as a teaching aid.

Nath and Capel, 1974, observed patients with early or late inspiratory crackles. The terms early and late inspiratory crackles were defined in terms of the occurrence of the crackles with respect to the air flow rate. Early inspiratory crackles occurred at the beginning of inspiration. Late inspiratory crackles occurred in late inspiration, or mid and late inspiration, or early, mid and late inspiration. The study concluded that early inspiratory crackles are associated with severe airways obstruction, eg. asthma, and late inspiratory crackles are associated with a restrictive defect, eg. pneumonia.

Frequency analysis of lung sounds was carried out by Chowdhury and Majumder, 1982. The study was carried out on patients with respiratory diseases: wheezes, crackles and pleural friction rub. The drawback of this investigation was that frequency analysis was carried out by an array of bandpass filters instead of using the Fourier transform. Nevertheless, it was concluded that frequency analysis of respiratory sounds could prove to be an objective technique for diagnosing pulmonary diseases.
The analysis of lung sounds in the frequency domain was also carried out by Benedetto et al, 1983. The authors studied crackles occurring in patients with cryptogenic fibrosing alveolitis. The results indicated that in cryptogenic fibrosing alveolitis crackles continued until the end of the inspiration and the crackle waveforms are of broadband characteristics, i.e. energy contents of the spectrum ranged between 100-2000Hz. This result suggested that due to its broadband characteristics it may be difficult to analyse respiratory sounds in the frequency domain. This is because one cannot differentiate which part of the spectrum is background noise and which is the actual data.

As mentioned previously, heart sounds can also be heard from the chest wall. They can be recorded and studied like respiratory sounds. The early stage for heart sound research was to identify the mechanisms of generation of the various sounds heard, as described in section 1.2. When these mechanisms were identified, rapid advances were made and various papers were published on the techniques and type of sound heard for particular heart disorders. A recent monograph by Crew et al, 1988, reviewed the techniques that can be use to detect heart or respiratory sounds. The authors also discussed the various types of sound heard and their associated diseases.

When a heart valve is damaged it can be replaced by a prosthetic valve. Once in place it is important to check the working order of the valve. Prosthetic heart valves are significantly different from the normal native valves. Therefore Smith et al, 1981, assessed heart sounds from patients with a prosthetic valve implant. The authors assessed four common types of prosthetic valves (ball valve, disk valve, porcine valve and bivalve) and the mechanism of sound production and their timing were discussed.

One aspect of heart auscultation study is to find correlations between patients with similar heart sounds and their disorders. Oswald et al, 1990, reviewed the medical record of 291 patients over a period of 30 years and the average duration of follow-up was 8 years. The authors concluded that if apical systolic click(s) with no
murmur is heard then the patient does not need any valve replacement operation; if apical click(s) are accompanied by systolic murmurs or pansystolic murmurs then there is a 2-3% chance that the patient may need a mitral valve replacement operation.

Another method to find the correlation between heart sounds and their diseases is the use of computers together with some signal processing algorithms. One such study was done by Akay et al, 1991. The authors obtained spectra of diastolic heart sound data by using the modified Yule-Walker (MYW) autoregressive moving average (ARMA) method instead of using the traditional Fourier transform method. This was because the Fourier transform algorithm was more sensitive to the effects of noise on the recorded diastolic sounds than the MYW ARMA algorithm. The results showed that there is a difference in the spectral energy distribution between pre- and post-angioplasty patients and between normal and diseased patients.

1.5 Discussion

The overwhelming amount of information, the restriction of the human auditory system and the distortion of sound by the stethoscope led to the development of the phonograms. The previous section outlined some uses of the phonograms and results obtained by various authors. Even though phonocardiographic recordings of chest sounds are available, it can be seen that the authors based their findings and conclusions on clinical experience and observation. The literature review also indicated that although some signal processing methods were used to analyse chest sounds there was little evidence that signal processing algorithms were used for the automation of chest sounds extraction and the classification of thoracic diseases. The aim of this project is to investigate the methods that can be used for the automation of chest sounds diagnosis. This automation process would contribute to the improvement of patient management. The following paragraphs will discuss the contents of the chapters in this thesis.
The introducing chapter has reviewed some characteristics of heart and lung sounds, as well as some areas of research in auscultatory medicine. Clearly, auscultation of the chest is a well-established technique and a powerful tool to aid the diagnosis of thoracic diseases. Before entering the domain of chest sound processing Chapter 2 introduces the characteristics of sound waves and the effect of sounds on some materials. This familiarisation of sound characteristics and materials can prove useful when designing equipment to record chest sounds. The chapter ends with the discussion on some methods that could be used to process chest sounds.

In Chapter 3 the methods that were used to construct two simulators are described, an experimental lung sound simulator and software heart sound simulator. These simulators were constructed to justify the materials used to construct a phonograph and to evaluate the effectiveness of the signal processing methods. The signals generated from the heart sound simulator were used to evaluate five transient extraction algorithms (Chapter 4). In this chapter the methods of calculation, advantages and disadvantages of the extraction algorithms are discussed. One of the transients extraction algorithms was then used to extract simulated heart sound transients. These heart sound transients were used for the classification process in Chapter 5. This chapter discusses advantages and disadvantages between the numerical and linguistic classification techniques. The results from Chapter 4 and 5 are used to choose the appropriate transients extraction algorithms and classification technique to construct a chest sound analyser. This chest sound analyser and its results are discussed in Chapter 6. Finally the conclusions of the methods used and works that can be done are discussed in Chapter 7.

References


analysis of respiratory crackles in cryptogenic fibrosing alveolitis. IEEE Trans on Biomedical
Engineering 1983;vol 30; p 620-3.


Stratton 1964;2nd Ed.

Ed.

1974;vol 50;p 360-5.

Eng 1982;vol 4; p 305-12.

[9] Crewe K, Crawford MH, O'Rourke RA. Differentiation of cardiac murmurs by dynamic

[10] Dock W. Mode of production of first heart sound. Archives of Internal Medicine 1933;vol 51:
p 737-46.


[14] Fredberg JJ, Holford SK. Discrete lung sounds: Crackles (rales) as stressrelaxation


J Applied Physiology 1971;vol 31; p 207-17.


Chapter 2

Characterisation of Sounds

2.1 Introduction

In order to gain familiarisation with phonograms, various aspects of sound, such as the measuring units and materials that can be used as sound insulators, are described. This chapter then briefly discusses the similarities between speech and acoustics from the thorax. It also deals with some classical digital signal processing methods that could be used to process phonograms.

2.2 Characteristics of Sounds

2.2.1 Sources & Units

Sound originates from the motion or vibration of an object in a compressible medium. This motion is transferred to the surrounding medium, usually air, as a pattern of changes in pressure. These pressure waves radiate outwards. If any components of the pressure waves lie within the range of human audibility, they are classified as sound waves. Those lying below the audible frequency range are called infrasonic, whilst those that are higher than the audible range are called ultrasonic.

There are several ways that the measurement of sound can be expressed. These include decibels (dB), phons and sones (Capel, 1991). The decibel is a function of a ratio between two values and therefore needs a reference value for its evaluation. This
reference value is usually the faintest sound that a young and healthy adult can detect. There are three variables that can be used to calculate the decibel unit, sound intensity level, sound pressure level and sound power level.

Sound intensity level is a measure of energy passing through an area of one meter square perpendicular to the propagation path (Capel, 1991). When calculating the decibel unit a reference value of $10^{-12}$ watts/m$^2$ is used. This is the faintest sound at 1kHz which can be heard by a young healthy adult. The formula for calculating the intensity level in decibels is

$$L_I = 10 \log_{10} \left( \frac{I}{I_r} \right) \quad (2.1)$$

where $I$ is the intensity in watts/m$^2$
$I_r$ is the reference intensity $10^{-12}$ watts/m$^2$

The energy transmitted by a sound wave is proportional to the square of the peak pressure exerted (Capel, 1991). As sound pressure is the easiest variable to measure it is the parameter most widely used to describe sound levels. The equation used to express sound pressure level in decibels is

$$L_P = 20 \log_{10} \left( \frac{P}{P_r} \right) \quad (2.2)$$

where $P$ is the pressure in $\mu$Pa
$P_r$ is the reference pressure = 20 $\mu$Pa

Sound power level is a measure of total energy output from a source (Capel, 1991). The measurement of sound pressure at any point does not give the true output from the source because pressure is affected by the surrounding environments and the source may be omnidirectional or unidirectional. The formula for calculating sound power level is
where \( W \) is the total radiated power
\( W_r \) is the reference power (10^{-12} \text{ watts})

The phon is a unit of loudness. It takes into account the human auditory response to various levels of loudness. 1 phon = 1 dB SPL at 1kHz, where SPL is Sound Pressure Level. Another unit used for the measurement of sound is the sone. It is also a loudness unit but it is not often encountered. 1 sone = 40 phons. Table 2-1 shows measurements of some common sounds.

<table>
<thead>
<tr>
<th>dB</th>
<th>Measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Threshold of hearing</td>
</tr>
<tr>
<td>30 - 40</td>
<td>Quiet living room</td>
</tr>
<tr>
<td>80 - 70</td>
<td>Vacuum cleaner at 1m</td>
</tr>
<tr>
<td>110</td>
<td>Pneumatic drill at 1m</td>
</tr>
<tr>
<td>120</td>
<td>Jet aircraft taking off at 60m</td>
</tr>
</tbody>
</table>


2.2.2 Propagation

Sound in air behaves in a similar way to ripples in a pond except sound radiates in three dimensions instead of two and the waves are longitudinal to the plane of propagation, i.e. they oscillate back and forth instead of up and down. An ideal point source in a homogeneous medium would radiate pressure waves like an expanding sphere. Therefore, the sound intensity diminishes as the distance from the source increases, Figure 2-1. The sound intensity at any point from the point source can be calculated using Equation 2.4 (Capel, 1991).
\[ I = \frac{W}{4\pi r^2} \]

where  
- \( I \) - Intensity in watts/m\(^2\)
- \( r \) - radius in meters
- \( W \) - source power in watts

The speed of sound is 332 m/s in air at 0 °C and at 1 bar (atmospheric pressure). The speed increases if the sound travels through a liquid medium and increases further in solids. By employing Huygen's theory of wavefront it can be shown that in a homogeneous medium sound can be reflected, refracted and diffracted similar to that of light waves, Figure 4-2 to Figure 4-4. Unlike light, sound waves are mechanical waves and have a longer wavelength. Several factors can affect its speed and propagation path. These factors include temperature, humidity, wind and the medium density. This would indicate that sound refraction can occur when a sound wave travels in the same medium because sound refraction occurs if the speed of the wave varies between different regions of the medium.
Figure 2-2 Reflection of plane waves

Figure 2-3 Reflection of spherical waves

Figure 2-4 Refraction of plane waves ($V_1 > V_2$)
As for sound diffraction, it is the spreading of sound waves through a narrow opening or round an obstacle. According to Huygen's theory, each point of a wave front is considered to be a new point source radiating in all directions. When there is no obstruction the waves are maintained in a forward direction because of the pressure of adjacent particles inhibiting side spread. If an obstruction is encountered there are no adjacent particles beyond the barrier's edge, so the waves spread behind it as if the edge was a new source. Similarly, when plane waves encounter a hole in an obstacle, they spread through the hole as if the hole was a new source and the plane waves become spherical waves. It must be noted that the statement is only true if the wavelength is large compared to the obstruction or hole size. This explains why we can hear around the corner even if we cannot see the source. This is because light waves have a smaller wavelength than sound waves.

![Figure 2-5 Diffraction of sound waves](image)

### 2.2.3 Harmonics

Pure tones consist of a single frequency and the wave is said to have a sinusoidal characteristic. In nature, only a few sources can produce a reasonably pure tone. This is because the source may have several radiating surfaces and each surface could be vibrating at a different frequency. This would result in waveforms having an
irregular shape. If these waves are analysed it can be shown that the waves consist of a fundamental sine wave with various amounts of frequencies that are simple multiples of the base tone called harmonics, Figure 2-6.

![Diagram of a pure tone together with its harmonics](image)

2.2.4 Noise Control

There are two main paths that generated noise can be transmitted to the receiver, directly and indirectly. In general, because of the characteristics of sound it is preferred to control unwanted sounds at the source, i.e. careful design of equipment to minimise noise generation. However, in the environment for recording chest sounds it is difficult to control the unwanted sounds, such as people, fluorescent lights, fans, doors and so on, and obtain a hi-fidelity recording simultaneously, unless the patient under study is placed in an anechoic room. Since an anechoic chamber is impractical for this study we will be concentrating our attention on the various methods and materials that could be used for controlling noise at the receiver, in this case the microphone.

In order to reduce the amount of unwanted sound arriving at the receiver we need to create barriers to reflect or absorb the unwanted sounds. One of the
arrangements that can be used for noise reduction is a quite room or office for the recording process. This arrangement may reduce some noises in the corridor but further noise could arise within the office. These noises include the ventilation system, recording equipment and computers. For a further reduction of the unwanted noise the microphone can be enclosed in some sort of casing. When designing and choosing materials for the enclosure of the microphone it must be noted that all materials absorb some sounds and reflect some sounds. Absorption is a function of the area of the materials and its absorption coefficients (Capel, 1991). In some enclosures sound insulation is necessary to reduce external noise, to prevent noise from escaping or to prevent reverberation. This can be achieved using the materials or combination of the materials listed in Table 2-2. The table shows the absorption coefficients for some materials; with one being 100% absorption of the sound wave and zero equal to 100% reflection. When designing a microphone's enclosure we need a rigid material to reflect noise and another material inside the enclosure to absorb the reverberations that arise from chest sounds or external noise. The materials listed in Table 2-2 could be used for the microphone's enclosure but the chosen material must be practical. The insulation of the enclosure's wall is not the only important aspect but the interface between different walls must also be taken into account when using the microphone's enclosure.

<table>
<thead>
<tr>
<th>Materials</th>
<th>125 Hz</th>
<th>250 Hz</th>
<th>500 Hz</th>
<th>1 kHz</th>
<th>2 kHz</th>
<th>4 kHz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acoustic panelling</td>
<td>0.15</td>
<td>0.3</td>
<td>0.75</td>
<td>0.85</td>
<td>0.75</td>
<td>0.4</td>
</tr>
<tr>
<td>Brick</td>
<td>0.024</td>
<td>0.025</td>
<td>0.03</td>
<td>0.04</td>
<td>0.05</td>
<td>0.07</td>
</tr>
<tr>
<td>Concrete</td>
<td>0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.03</td>
</tr>
<tr>
<td>Curtains</td>
<td>0.05</td>
<td>0.12</td>
<td>0.15</td>
<td>0.27</td>
<td>0.37</td>
<td>0.5</td>
</tr>
<tr>
<td>Fibreglass, 1 inch</td>
<td>0.07</td>
<td>0.23</td>
<td>0.42</td>
<td>0.77</td>
<td>0.73</td>
<td>0.7</td>
</tr>
<tr>
<td>Glass</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Plywood</td>
<td>0.11</td>
<td>0.12</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>-</td>
</tr>
<tr>
<td>Wood</td>
<td>0.1</td>
<td>0.11</td>
<td>0.1</td>
<td>0.08</td>
<td>0.08</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Table 2-2 Some typical absorption coefficients, taken from Capel, 1991.
2.3 Speech Sounds and Thorax Acoustics

Speech sounds can be classified into three main groups according to their mode of excitation (Rabiner, 1978). Voiced sounds are a result of oscillation of air in the vocal cords. Unvoiced sounds are generated by forcing air through a constriction thus creating turbulence. Plosive sounds are created by the built up of pressure and abruptly releasing it.

Comparing these three main speech-sounds to thorax acoustics it can be seen that they do exhibit some similarities to speech sounds. Crackles are a result of the sudden release of pressure in a closed airway, and could therefore be classified as plosive sounds. Wheezes arise from air turbulence in constricted airways and can therefore be classified as unvoiced sounds. Heart sounds can be classified in a similar way to lung sounds.

2.4 Methods of Processing

Information may be extracted from acoustic signals using various digital signal processing (DSP) algorithms. The analysis of any signals can be divided into two main categories, the time domain and the frequency domain. In this section a few known DSP techniques are briefly discussed.

Before using any DSP algorithms the analogue signal must be expressed as a sequence of numbers. From the source, the analogue signal is measured by an appropriate transducer and digitised. The digitising process must satisfy the Nyquist sampling theorem in order for a unique representation of the analogue signal as a sequence of numbers (Oppenheim, 1989).

2.4.1 Time Domain

There are several methods that can be used to extract transients of interest from the raw signal. After the extraction of individual transients there are numerous
signal processing techniques to manipulate the transients in the time domain. The section below briefly describes some of the DSP techniques.

- **Autocorrelation & Cross-correlation**

  The correlation function of a signal is a measure of the inter-dependence between the signal values at different times (Oppenheim, 1989). There are two types of correlation functions, autocorrelation and cross-correlation. The autocorrelation function of a signal $x(t)$ is defined as

  $$\phi_x(\tau) = \overline{x(t)x(t+\tau)}$$  \hspace{1cm} 2.5

  where the top bar means average in the time domain.

  and the cross-correlation function between two different signals $x(t)$ and $y(t)$ is defined as

  $$\phi_{xy}(\tau) = \overline{x(t)y(t+\tau)}$$  \hspace{1cm} 2.6

- **Signal Averaging**

  When the transients are extracted they may contain some noise. Signal averaging is one of the methods used to reduce noise and to obtain an overall average transient. There are two commonly used averaging methods, mean and median averaging. The mean signal averaging process involves the alignment of the transients, summing them and dividing by the number of realisations. Median averaging involves the alignment of the transients and then the median of each corresponding data point is obtained.

- **Feature Extraction**

  A feature of a transient is defined as a prominent or distinct part of the transient. A single transient can be uniquely described by several features. These features could be the number of peaks, number of troughs, length of the transient, amplitudes,
etc. Using these features the transients can be sorted into similar groups. The feature extraction process is a straightforward procedure but care must be taken when a noisy signal, such as an acoustic waveform, is used; for example in obtaining the number of peaks or troughs.

2.4.2 Frequency Domain

Any signals in the time domain can be transformed to the frequency domain where they can be analysed for their frequency contents. The transformation process can be carried out using the Fourier integral. The Fourier integral and its inversion formula (the Fourier transform pair) for continuous time signals are defined as (Brigham, 1988)

\[ X(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi ft} dt \quad \text{----------------------------- 2.7} \]
\[ x(t) = \int_{-\infty}^{\infty} X(f) e^{j2\pi ft} df \quad \text{----------------------------- 2.8} \]

where \( x(t) \) is the continuous time domain signal
\( X(f) \) is the Fourier transform of \( x(t) \)

It is desired to modify the Fourier transform pair, Equation (2.7) & (2.8), such that the equations are manageable for digital computation. A modification of the Fourier transform pair gives rise to the Discrete Fourier Transform (DFT). To obtain the discrete version of Equation (2.7) it is necessary to take the continuous signal, \( x(t) \), and sample it (Brigham, 1988).

The DFT algorithm is very inefficient because of redundant calculations and in 1965 Cooley and Tukey published a highly efficient algorithm to calculate the Fourier transform of a discrete time signal. This algorithm is known as the Fast Fourier Transform (FFT) and it needs \( \frac{N}{2} \log_2 N \) complex calculations instead of \( N^2 \) complex calculations used by the DFT algorithm to compute the Fourier transform of a discrete time signal of length \( N \) samples (Brigham, 1988).
When analysing the frequency spectrum of a signal one may conclude that
the signal contains noise at several different frequencies, or the important part of the
signal is composed of various frequencies. To improve the signal to noise ratio of the
original signal, a digital filter can be used. The general structure of the digital filter is
shown in Figure 2-7 and its equation below, Equation 2.9 (Oppenheim, 1989).

\[
Y(z) = X(z) \sum_{n=0}^{N} a_n z^{-n} - \sum_{n=0}^{M} b_n z^{-n}
\]

where \( z^{-1} \) is the unit delay operator
\( a_n \) & \( b_n \) are the filter's coefficients

There are two classes of digital filters that use the filter structure in Figure 2-
7, Finite Impulse Response (FIR) and Infinite Impulse Response (IIR) filters. IIR filters
have non-zero \( a \) and \( b \) coefficients whereas for FIR filters the \( b \) coefficients are equal
to zero.

Figure 2-7 General structure of a digital filter
Another type of filter that could be used to improve the signal to noise ratio is an adaptive filter. An adaptive filter consists of two distinct parts: a digital filter, FIR or IIR, with adjustable coefficients and an adaption algorithm to modify the filters coefficients. These adaption algorithms include the least mean square algorithm, the recursive least mean square algorithm and the Kalman filter algorithm (Orfanidis, 1988).

2.5 Discussion

This chapter gave an overview of some of the basic characteristics of sound and a brief introduction of some materials that can absorb or reflect sound waves. The propagation characteristics of sound waves allow them to travel via many paths, if there is an obstacle in their path, and through most materials. Therefore it is difficult to totally insulate the receiver from the background noise. For this project a sophisticated transducer could be used to record chest sounds. The transducer could be designed in a way such that it matches impedance between the chest wall and the point where sound is converted into electrical energy. This impedance matching material ensures that the generated sounds travel from the source, via the human tissues and the impedance material, and is then converted into electrical signal with little distortion. Since the speed of sound in air is approximately 332 m/s and approximately 1500 m/s in tissues, that is there is a mismatch of impedance between air and tissues, most of the background noise from the room is excluded from the recorded signal. This sophisticated transducer may be difficult to design or expensive to obtain. Therefore a simple electret microphone is used to record the chest sounds. With careful consideration of microphone placement and the recording environment, background noise could be substantially reduced. The recording environment could be in an office, where background noise is reduced by the office’s walls, or the recording could take place at night, where there are fewer activities than during the day. As for the microphone, it could be enclosed in a cup but this enclosure may not be very effective due to the way sound propagates through most materials.
The chapter also dealt with some classical techniques used for signal processing. The recorded chest sounds by a simple transducer may contain unwanted background noise, but with the application of signal processing algorithms the user may be able to compensate for the simplicity of the recording equipment.

References


Chapter 3

Experimental & Simulation Techniques

3.1 Introduction

Heart sounds and lung sounds contain both the transients of interest and unwanted background noise. The chest sounds are acoustic waveforms thus their frequency components are similar to that of the background noises. In order to have more control over the signals used to test the signal processing algorithms two simulators were constructed. This chapter describes these two acoustic simulators, the waveforms generated and their benefits for the testing of signal processing algorithms. Section 3.2 describes a simple simulator that generates lung sound and the equipment used to record real acoustic waveforms while section 3.3 describes a software simulator for heart sounds.

3.2 A Bench Top Transient Simulator

When constructing any bench top simulator there are several factors that must be taken into account to ensure that the generated signals are similar to the real signals. The purpose of the bench top simulator was to simulate lung sounds, specifically crackles. Crackles are non-musical lung sounds occurring during inspiration or expiration or both (Macleod, 1983, Swash & Mason, 1984). It was suggested that crackles are generated when a closed airway suddenly re-opens due to intra-bronchial pressure (Forgacs, 1967). Using this information it was decided that the simulator must be able to
generate aperiodic transients in a variable background noise environment. To simulate acoustics emanating from the thorax a medium which had similar characteristics to the lung was needed. Goncharoff et al, 1989, suggested that water can be used to model lung parenchyma "if we assume that the alveolar structure resembles tightly packed bubbles in a liquid-like medium". If water can be used as a medium to simulate lung parenchyma, pulsating bubbles could be used to simulate crackles.

In this section the equipments used to generate acoustic signals and the conversion of these signals into a digital format for storage is discussed. The proposed acoustic generator can be modified such that it could be used to test recording equipment to obtain the best possible signal under various conditions.

3.2.1 Experimental Set Up

- Overall Apparatus

The pulsating bubbles method was achieved by forcing air down a tube which was submerged under water. The apparatus of the bench top simulator consisted of a water tank which has a submerged tube connected to an air pump to generate the aperiodic acoustic signals. A white noise generator and speaker were also used to create various background noise environments, Figure 3-1. Any sounds produced are picked up by the transducer, digitised at 5000 Hz and stored on a computer.
**Recording Equipment**

An electret condenser microphone was used to record the generated acoustics. It was decided that the microphone should be enclosed in some type of sound insulation materials to reduce background noise.

When a sound wave impinges upon an obstructing barrier or surface, part of the wave is reflected and the remainder is transmitted into the barrier. If the barrier consists of dense, hard and rigid materials, such as glass and concrete, then most of the incident energy is reflected. However, if the barrier is non rigid and consists of soft, porous materials then a considerable portion of the incident energy is transmitted into the barrier. The transmitted sound wave will emerge from the other side of the barrier with reduced energy due to absorption by the material. Table 2-2, Chapter 2, illustrates some commonly used material’s absorption coefficients. In summary, there are two types of sound insulation materials, those that reflect sounds and those that absorb sounds.

For the microphone's enclosure there was a need to reduce external background noise by using sound insulating material and the material must electrically
insulate the patient. Although glass and concrete are ideal materials to reflect sound
waves they are not easy to use in this situation. Steel could be another option but this
material is a conductor. Given the specifications and material properties, polyethylene
was chosen as the enclosure's material. Polyethylene is a hard and rigid material but its
density is 960 kg/m^3 compared to >2400 kg/m^3 for glass and concrete. This implies
that polyethylene may not be as good sound reflector as glasses or concretes but the
material is chosen for its practicallity and because it satisfies the given specifications.
The reflected sound waves can cause reverberation within the enclosure. Therefore a
second material within the enclosure is necessary to absorb and attenuate the
reverberations. To absorb the sound energy the material must be a soft and porous, ie.
have a high absorption coefficient. For example, cellobond-k is a porous material but it is
brittle, whereas cork is a soft and porous material.

The acoustics picked up by the tie clip microphone are digitised by a Digital
Signal Processing (DSP) board and stored as a sequence of numbers on a computer.
Because the DSP board has only one input channel and more than one site on the chest
wall must be recorded simultaneously there was a need for a multiplexer system for
recording multiple data channels. The developed multiplexer system and software are
shown in Appendix A.

- Experimental Procedures

  Pulsating bubbles were created in the water tank and acoustics were
generated for different size air tubes and background noise levels. The generated
acoustics were detected using a Sony electret condenser microphone enclosed in various
enclosure arrangements. The microphone's output is amplified, digitised and stored on a
computer.

  The variation in background noise level was measured by using a microphone
and a ONO SOKI spectrum analyser, calibrated to measure sound pressure level in
decibels.
3.2.2 Results

<table>
<thead>
<tr>
<th></th>
<th>Tube diameter = 5mm</th>
<th>Tube diameter = 1mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polyethylene</td>
<td>Polyethylene</td>
<td>Polyethylene</td>
</tr>
<tr>
<td>&amp; cork</td>
<td>only</td>
<td>&amp; cork</td>
</tr>
<tr>
<td>~38 dB</td>
<td>Figure 3-2a</td>
<td>Figure 3-3a</td>
</tr>
<tr>
<td>~ 60 dB</td>
<td>Figure 3-2b</td>
<td>Figure 3-3b</td>
</tr>
</tbody>
</table>

Table 3-1  Recordings and corresponding figures obtained from the simulator.
(Note that the amplitudes of the figures are normalised to unity)

After the recordings were made from the above environments, Table 3-1, a recording was made by submerging the two different size diameter tubes in the water tank simultaneously. The microphone was enclosed in a polyethylene and cork enclosure with the background noise level set at ~38 dB. This recording can be viewed in Figure 3-6.

The noise level of 38 dB is the same as a quiet living room and 60 dB is approximately similar to the noise level of a normal conversation.
Figure 3-2 Bubble acoustics recorded with a polyethylene outer enclosure and cork as an inner sound insulating material. Tube diameter = 3mm
Figure 3-3 Bubble acoustics recorded using a polyethylene outer enclosure only. Tube diameter = 3mm
Figure 3-4  Bubble acoustics recorded with a polyethylene outer enclosure and cork as an inner sound insulating material. Tube diameter = 1mm
Figure 3-5 Bubble acoustics recorded using a polyethylene outer enclosure only. Tube diameter = 1mm

Figure 3-6 Bubble acoustics recorded with a polyethylene outer enclosure and cork as an inner sound insulating material. Tube diameter = 1mm and 3mm
Figure 3-7 Bubble acoustics from Figure 3-2. 20ms long transients.

Figure 3-8 Real crackle phonopneumogram. 30 ms long transients.
3.2.3 Discussion

When comparing acoustics that have been recorded with and without cork inside the polyethylene enclosure, i.e. comparison between Figure 3-2 and Figure 3-3 and between Figure 3-4 and Figure 3-5, it can be seen that the amplitudes for the acoustics which were recorded without using cork are much smaller. The comparison between these figures indicated the existence of reverberation within the polyethylene enclosure and moreover confirmed the need for a material within the cup to absorb and attenuate the unwanted acoustics. The dilemma with acoustics reverberation can be further demonstrated in Figure 3-5. Even with the background noise at approximately 38dB it is difficult to see the acoustics that were generated by the pulsating bubbles. This is because the condenser microphone not only detected the original acoustics it also recorded the reverberation, which is out of phase, as well. This resulted in the cancellation of the waveform detected.

Figure 3-7 showed the pulsating bubble acoustics for the duration of 20ms. When these transients were compared to the crackle phonograms, Figure 3-8, it can be seen that the bubble transients have a higher frequency than the crackle phonograms. This is because crackles are generated within the lung, and, as was discussed, in Chapter 1, the lungs attenuate high frequency sounds. The other factor is that crackles are caused by secretions of fluid within the lungs and the secretions are much more viscous than water. Therefore a more viscous solution should be used instead of water to simulate the lungs.

Recording was also made by using the two different size tubes simultaneously, Figure 3-6. This recording was done to demonstrate that different arrangement of tubes' size and tube locations can be used to generate different types of acoustic transients. Although different types of transients were produced the generated transients do not have the same characteristics as real crackles.
3.3 A Heart Sound Simulator

Auscultation of the heart is one of the steps in the physical examination of patients with cardiovascular diseases. With the exception of cardiologists, few physicians or medical students have the opportunity to experience the wide range of heart sounds on a regular basis. The most common methods to learn auscultatory skills are from textbooks and audio tapes. These sources do not provide the interaction needed by the users to suit their individual learning skills, e.g. when comparing similar heart sounds or changes in heart sounds caused by the patient's position.

The construction of a heart sound simulator is discussed in this section. It is divided into three main parts, Normal & Abnormal Heart Sounds, Software Design and Results. The first part describes characteristics of heart sounds that exist in normal and diseased patients. The second part outlines the methods and the structure of the heart sound simulator. The final part discusses the results obtained from the simulator and a comparison with existing heart sound simulators.

3.3.1 Normal & Abnormal Heart Sounds

Chapter 1 gave a brief review of the production of heart sounds from the human body. To construct a heart sound simulator we need to explore further the various types of heart sounds and their characteristics. The following descriptions of heart sounds are from Macleod, 1983, and Swash & Mason, 1984.

In a normal cardiac cycle the first and second heart sounds can be heard on the chest wall. In diseased patients the following sounds or deviations from the normal sounds can be heard upon auscultation: -

- The sounds may have a different intensity, either increased or decreased.
- The sounds may be abnormally split.
- Low frequency sounds occurring in diastole, i.e. third or fourth heart sounds.
• Additional sounds may be heard, e.g. clicking valves or murmurs.

**Characteristics of first and second heart sounds**

As described in chapter 1, the first and second heart sounds are due to the closure of the A-V valves and the semilunar valves respectively. The duration of these sounds is 0.14 second for the first heart sound and 0.11 second for the second heart sound. As for the frequency content of the heart sounds, the second sound has a higher frequency component compared to that of the first heart sound. This is due to tautness of the valves and the different characteristics of the blood chambers i.e. the atria and the ventricles.

**Characteristics of third and fourth heart sounds**

Besides the first and second heart sounds, lower frequency sounds can be heard in diastole. These sounds are known as third or fourth (or atrial) sounds. Both of these sounds are caused by the abnormal filling of the left ventricle. The third heart sound occurs in early diastole, about 0.15 second after the second heart sound and the fourth heart sound occurs close to the first heart sound of the next cardiac cycle. The positions of these sounds can be seen in Figure 3-9.

![Figure 3-9 Third and fourth heart sounds](image)
Characteristics of splitting heart sounds

Splitting of heart sounds are caused by the asynchronous closure of the A-V valves or the semilunar valves. Splitting is difficult to detect by auscultation because of the short interval that separates the two components of the heart sound. When heard, splitting of the first heart sound is not a sign of heart disease. It is of importance because the two components may be confused with the fourth heart sound or with the first heart sound followed by an ejection click. On the other hand, splitting of the second heart sound is associated with some heart diseases.

![Figure 3-10 Splitting of the second heart sound](image)

Characteristics of clicking and snapping valves

When either of the semilunar valves become abnormal they generate high frequency sounds called systolic clicks. There are two types of clicks, mid-systolic clicks, which occur in the middle of systole, and ejection clicks, which occur in early systole. Sometimes ejection clicks are mistaken for the splitting of the first heart sound.

Opening snaps are heard when the A-V valves are abnormal. These sounds are analogous to the systolic clicks but by tradition the sounds are known as opening snaps because they are caused by the A-V valves. Opening snaps occur in early diastole and can be mistaken for a wide splitting of the second heart sound.
Characteristics of murmurs

Murmurs are caused by the turbulence of blood flow at or near a valve or a constriction in the arteries. In the analysis of murmurs, their timing and quality are the most important features. There are three types of murmurs: systolic murmurs, diastolic murmurs and continuous murmurs.

1. Systolic murmurs can be divided into three types.

- *Ejection systolic murmurs.* This sound results from turbulent blood flow through a distorted or stenotic semilunar valve or even from turbulence caused by increased blood flow through normal semilunar valves. The murmurs increase to a crescendo about the middle of systole then diminish just before the second heart sound. Thus in the phonocardiograph of the heart sound, the murmur has a 'diamond shaped' envelope pattern.

![Figure 3-11 Ejection systolic murmur](image)

- *Pansystolic murmur* is a result of blood leaking into an area of low pressure. The murmurs start simultaneously with the first heart sound and may spill into early diastole. These murmurs are caused by the leaking of A-V valves or a ventricular septal defect.
Late systolic murmurs. Clinically these are characterised by a clear gap between them and the first heart sound. The onset of the murmur is in mid or late systole and continued up to or throughout the second heart sound.

2. Diastolic murmurs. There are three main types of diastolic murmurs: early diastolic murmurs, mid-diastolic murmurs and late diastolic murmurs.

- *Early diastolic murmurs* are caused by the leakage of the semilunar valves. The sound is loudest at the beginning of diastole and ends about mid-diastole.
• *Mid-diastolic murmurs* are caused by turbulent blood flow at the A-V valves. The sound starts later in diastole. It reaches a peak at mid-diastole and ends before late diastole. Thus in a phonograph of the heart sound, mid-diastolic murmur has a diamond shaped envelope characteristic.

![Figure 3-15 Mid-diastolic murmur](image)

• *Late diastolic murmurs* are due to the turbulence of blood flow at one of the A-V valves when the atria contract. The sound starts at late diastole and is loudest at the end of diastole.

![Figure 3-16 Late diastolic or presystolic murmur](image)

3. Continuous murmurs. As the name suggests this type of murmurs can be heard throughout systole and diastole.
3.3.2 Software Design

To construct a heart sound simulator a number of parameters must be identified. These parameters govern the characteristics of the generated heart sounds. For any oscillatory transient the main features that determine its appearance are its length, the number of turning points and the attack and decay rates. From the transient's features, seven parameters have been deduced for the generation of a heart sound transient. From here onwards referred to as feature parameters; and they include the duration of the transient, starting position, maximum amplitude, lowest and highest frequencies and the attack and decay rates. Before generating various heart sounds a number of feature parameter sets must be determined for the generation process.

**Normal Heart Sound parameters**

As described earlier in Chapter 1, the heart consists of four valves and each of these valves contributes to the genesis of normal heart sounds. The first heart sound comes from the mitral and tricuspid valves and the second heart sound comes from the aortic and pulmonary valves. Therefore, four sets of feature parameters are needed for the generation of normal heart sound.

**Abnormal Heart Sound parameters**

There are two main types of disorder that can cause abnormal heart sounds. The first is type is due to disease of the valves. The second type is due to the abnormal flow of blood through the heart caused by diseased heart muscles, atria or ventricles.
The sounds that are caused by diseased valves include splitting of heart sounds, clicking of valves and snapping of valves. To simulate the clicking and snapping of valves four extra sets of feature parameters, one for each valve, are needed. For the simulation of splitting heart sounds no extra feature parameters are required because the splitting of heart sounds is due to the asynchronous closure of the valves. This means the simulation of splitting heart sounds can be achieved by designating different relative starting positions of the valves’ transients, using the normal heart sound feature parameters. For example in the splitting of the first heart sound, the relative position parameters in time of the mitral and tricuspid valves are different but the relative positions of the aortic and pulmonary valves remain the same. Diseased valves can also give rise to murmurs. Instead of using four sets of feature parameters for the generation of murmurs (one for each individual valvular murmur) only one set of feature parameters is used for simplification.

The other heart sounds that are caused by the abnormal heart muscles are the third heart sound and fourth heart sound. To simulate these sounds two more sets of feature parameters are needed.

Other parameters

As well as the feature parameters that are needed to generate individual heart sounds, the software sound generator also needs a few more parameters to generate a complete phonocardiogram. These parameters include the length of the overall signal, number of heart beats per minute, sampling frequency, percentage of white noise and the degree of randomness of the heart sound transients.

Generation of Heart Sound

The generation of the heart sound waveform is divided into three stages.
1) The first stage involves calculating the number of samples in the desired waveform. By using the sampling frequency and the length of the signal parameters the total number of samples is evaluated, and each sample is initialised to zero, i.e. \( y_i(t) = 0 \).

3) The second stage is the generation and adding in of individual transients. These transients are added to \( y_i(t) \) according to the relative position and number of heart beats per minute parameters. To generate individual transients Equations 3.3 and 3.6 are used.

2) The final stage is the generation of white noise, \( n(t) \), to represent measurement noise and tissue noise. Using a normalised maximum value of unity the percentage of white noise can be specified. Generated white noise is scaled and then added to \( y_i(t) \), i.e. \( y_{2i}(t) = y_i(t) + n(t) \).

Each individual heart sound transient is characterised by a sinusoidal wave bounded in an envelope. The transient of the heart sound may have a broad band characteristic. Nevertheless, to make the transient generation process simple, a chirp is used. With reference to Franks (1969) a chirp, \( c(t) \), is defined as a sinusoidal wave that consists of variable instantaneous frequency, \( f_1 \), which varies linearly with time.

The chirp used in the heart sound generator is characterised by Equation 3.1.

\[
c(t) = A \sin(2\pi f_i t + 2\pi bt^2) \quad \text{3.1}
\]

and the variable instantaneous frequency is defined as

\[
f_i = 2bt + f_o \quad \text{3.2}
\]

substituting \( b \) in Equation (3.1) gives

\[
c(t) = A \sin(\pi f_o t + \pi f_i t) \quad \text{3.3}
\]
When the heart sound transient is generated using Equation 3.3, the transient has a constant amplitude, Figure 3-18. In real heart sounds this is not the case. Looking closely at a real heart sound transient it can be seen that the peaks of the transient rise to a maximum and then decay away. To create the attack and decay rate on the generated heart sound transient, the transient needs to be multiplied by an envelope that has similar characteristics. The envelope is evaluated using Equation 3.6.

\[
\text{Attack} = \left[ \left(1 - e^{-0.2T} \right) \cdot \sin\left(\frac{\pi}{2T}t\right) \right]\left(1 \cdot \frac{A}{11}\right) \quad \text{-------------- 3.4}
\]

\[
\text{Decay} = \left[ e^{-0.4T} \cdot \cos\left(\frac{\pi}{2T}t\right) \right]\left(1 \cdot \frac{D}{11}\right) \quad \text{-------------- 3.5}
\]

Envelope of the transient = Normalise(Attack * Decay) \quad \text{------- 3.6}

Where  
\begin{align*}
T & - \text{Duration of the transient} \\
A & - \text{Attack rate} \\
D & - \text{Decay rate}
\end{align*}

The constants \((0.2, 0.4 \text{ and } 11)\) in Equations 3.4 and 3.5 have been chosen after careful study of real heart sounds to give the envelopes appropriate shape. The
exponential components of the equation determine the shape of the attack and decay characteristics. The sine and cosine components ensure that the envelope starts at zero and decays to zero. The rates of attack and decay are controlled by the A and D variables respectively. Figure 3-19 demonstrates the attack, decay and envelope characteristics using different A & D rates. The result of the Attack and Decay multiplicands are normalised to a maximum value of unity.

The heart sound transient is obtained by multiplying the generated chirp and envelope together, scaled, and then added to the overall signal, $y_2(t)$.

During the generation process of the heart sound transients, a degree of randomness is introduced in the feature parameters. Thus no two transients have the same characteristics.
3.3.3 Results

The flow charts of the described procedures were composed, Appendix B, and implemented. Some examples of the menus from the program are shown in Figure 3-20. Using heart sound features, normal and abnormal heart sounds are generated, Figure 3-21a-g. The waveforms in Figure 3-21 also show real phonocardiograms taken from Butterworth et al (1964), as well as heart sounds generated by this simulator for comparison.

Not only can the users observe the generated heart sound they can also listen to it. This is done by dumping the generated sound to the D/A port on the TMS320C25 DSP board within the program environment. The D/A process is performed by the program saving the generated data in a file and then executing a DOS program. This program initialised the DSP board and outputs the contents of the file to the D/A port which is connected to a loudspeaker.
This is a heart sound simulator. Heart sounds consist of two main complexes, the first and second heart sounds. The first heart sound is a result from the vibration of blood and the heart wall due to the closure of the A-V valves. The second heart sound is generated in a similar fashion but it involves the semilunar valves. In order to generate these sounds the user needs to specify a few parameters.

1 - Sound variables
2 - First & second heart sound variables
3 - Additional sounds variables
4 - Generate/Plot/Output signal
5 - Load/Save input variables
6 - Quit

Enter an option:

Heart sound variables

- Heart rate = 75 Beats/minute
- Length of sound = 5.00 second(s)
- Sampling frequency = 10000 Hz
- % of added noise = 10.00 %

1 - Heart rate
2 - Duration of sound
3 - Sampling frequency
4 - Percentage of added noise
5 - Previous menu

Choose an option:

First & Second Heart Sound Variables

<table>
<thead>
<tr>
<th>First heart sound</th>
<th>Second heart sound</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>T</td>
</tr>
<tr>
<td>Duration (ms)</td>
<td>140.0</td>
</tr>
<tr>
<td>Rel. position (0-100)</td>
<td>1.0</td>
</tr>
<tr>
<td>Rel. intensity (0-100)</td>
<td>50.0</td>
</tr>
<tr>
<td>Lowest frequency (Hz)</td>
<td>30.0</td>
</tr>
<tr>
<td>Highest frequency (Hz)</td>
<td>30.0</td>
</tr>
<tr>
<td>Attack rate (0-10)</td>
<td>7.0</td>
</tr>
<tr>
<td>Decay rate (0-10)</td>
<td>4.0</td>
</tr>
</tbody>
</table>

1 - First heart sound variables
2 - Second heart sound variables
3 - Previous menu

Heart sound options:
Figure 3-20 Input menus for the heart sound simulator
Figure 3-21a Normal heart sound with no added noise

Figure 3-21b Normal heart sound with 10% added white noise
Figure 3-21c Splitting of the first heart sound
(Phonocardiogram taken from Butterworth et al (1964))

Figure 3-21d Splitting of the second heart sound
(Phonocardiogram taken from Butterworth et al (1964))
Figure 3-21e Third heart sound
(Phonocardiogram taken from Butterworth et al (1964))

Figure 3-21f Fourth heart sound
(Phonocardiogram taken from Butterworth et al (1964))
3.3.4 Conclusions

Comparison between the generated heart sounds and real phonocardiograms in Figure 21a-g showed some similarities. When normal simulated heart sound was played on the speaker by the DSP board the typical normal heart sound could be heard, that is the lub-dub sound emanating from the heart in a normal healthy adult which can be heard on the chest wall using a stethoscope.

There are two heart sound simulators that have been developed by Bergeron, 1988, and Takashina et al, 1990. In the simulator developed by Bergeron, real heart sound primitives are recorded. The essence of each heart sound is extracted from the recorded sounds and stored. These stored sounds are modified in amplitude, frequency, harmonic content and duration. Depending on the desired sound, each reconstituted sound is mixed in varying proportion and played on the speaker. The simulator
developed by Takashina et al, used real recorded heart sounds. The sounds are played back on the speakers which are placed inside a mannequin. This gives a realistic situation which is useful for medical students when learning the art of auscultation.

The two heart sound simulators described above, to a degree, have been validated with the help of clinical cardiologists whereas the heart sound generator developed here has not yet been thoroughly tested. Although the validation process was not completed, the heart sounds generated by constructed simulator was informally listened to and assessed by a research fellow, who is a qualified Orthopeadic surgeon. The testing and the build up of a data base of classified heart sounds could be undertaken with the aid of a cardiologist.

3.4 Discussion

There are two main uses for the bench top simulator. The first is the genesis of acoustic signals which can be used to test developed DSP algorithms. The second is testing for the immunity of the recording equipment in various background noise levels. When testing the recording equipment it is impossible to get rid of all the unwanted background noise, even if the sound insulation materials can totally reflect or absorb sound waves. This is because of the interface between the microphone’s enclosure and the subject under study. If the interface were sound proof then there is still the problem of the subject conducting background noise. The only way to record a clean acoustic signal is to put the subject under study in a sound proof room.

There are some important features that are found on the newly developed heart sound generator which are not found on the Bergeron and Takashina simulators. These are, the display of the heart sound, i.e. phonocardiograms, and the randomness of each heart sound generated. The observation of the phonocardiograms helps to establish the differences between two similar heart sounds and a better understanding on the genesis of heart sound. The randomness factors make sure that, as in nature, the
generated heart sounds are not all the same. This property makes the software simulator more flexible. The generated signals can be used to test developed signal processing algorithms to see under what conditions the algorithms are most useful and also where they break down. One disadvantage of the simulator is that breath sounds and breathing artifacts are not included in the simulated heart sounds.

References

Chapter 4

Extraction of Chest Sounds

4.1 Introduction

The automation process for locating the beginning and ending of a transient of interest in a signal is of great importance in many areas of signal processing. Such a scheme would eliminate significant amounts of computation time to extract useful information from the recorded signal. The problem of discriminating transients from background noise is not a trivial matter, except in recordings of extremely high signal to noise ratio. Such an ideal recording is not practical for many applications including chest sounds. Information from chest sounds can be processed and extracted in the time domain or in the frequency domain. The disadvantage of studying chest sounds in the frequency domain is that they have a broadband characteristic and that they lie within the frequency range of the background noise. This would make it difficult to distinguish the frequency bands that belong to normal chest sounds, abnormal chest sounds or background noise. For this reason it was decided to study chest sounds in the time domain. This chapter first discusses the general extraction technique, selection of threshold levels and the ideal transients wanted from the extraction process. The chapter then describes five transient's transformation algorithms and by testing these algorithms using the simulated data from Chapter 3 the advantages and disadvantages of the transformation techniques are discussed.
4.2 Extraction Technique

- General Method

From observations of any signals in the time domain it can be noted that they contain both the transients of interest and background noise. For most signals the transients of interest only occur in a small portion of the signal. Thus to extract information and reduce processing time it is necessary to determine the beginning and end of each transient.

In general, the detection process relies upon the transformation of the original signal such that the areas that contain the transients stand out from the rest of the signal. From this converted waveform (hereafter referred to as the transformed signal) appropriate thresholds are selected to obtain the beginning and end of the transient. A block diagram and an example of the output waveform from each block of the extraction process are shown in Figure 4-1.

![Block Diagram](image)

Figure 4-1 Block diagram of the transients extraction process and an example of the outputs from the block diagram.
• **Threshold Level**

The selection of the threshold level for the **Transients Detector** is important for the correct extraction of the wanted transients. A high threshold level could result in late estimates of the transient's starting position or the early estimates of end time, whereas a low threshold level means that the **Transients Detector** has to deal with the noise in the transformed signal. These problems indicate that the user must specify an individual threshold level for each of the signals so that the **Transients Detector** performs a reasonable estimation of the starting or ending positions. The procedure for the selection of the threshold level is time consuming and inconsistent, due to the human factor. Therefore an automatic method to obtain an appropriate threshold level is proposed.

**Figure 4-2** Block diagram for calculating the threshold level of a transformed signal.

<table>
<thead>
<tr>
<th>KEY</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSIG</td>
</tr>
<tr>
<td>OSIG</td>
</tr>
<tr>
<td>OM</td>
</tr>
<tr>
<td>NM</td>
</tr>
<tr>
<td>TL</td>
</tr>
</tbody>
</table>

- **NSIG = OSIG**, but the values of **OSIG** that are greater than OM are made equal to OM.
- | NM - OM | < 0.01
  - **YES**
  - Break
  - **NO**
  - **TL = 2 * NM**

- **NM < 0.02**
  - **YES**
  - **TL = 1.3 * NM**
  - **NO**

- **0.02 > NM > 0.15**
  - **YES**
  - **TL = 1.5 * NM**
  - **NO**
Figure 4-2 shows the block diagram for the *Threshold Level* algorithm. The algorithm starts by normalising the transformed signal such that the maximum peak is unity and then it enters a repeating process. This is termed as the "old signal". Upon entry the mean of the old signal, i.e. the normalised transformed signal, is calculated, OM. It then equates the new signal, NSIG, to the old signal, OSIG, but the values of the old signal that are greater than OM are forced to be OM. After a new signal has been obtained a new mean is calculated, NM. If the absolute difference between the new mean and the old mean is greater than 0.01 the process repeats itself. The purpose of this repeating process is to calculate the base line of the transformed signal, see Figure 4-1, which equals the new mean, NM, when the repeating process terminated. When the base line value of the transformed signal has been obtained is it scaled with a suitable constant to obtain the threshold value. These scaling constants have been chosen after studying of a number of transformed signals.

**Ideal Extracted Transients**

It is intended to test the five extraction algorithms using the simulated data from Chapter 3. These simulated data were sampled at 10,000Hz. The advantage of the test is that the starting and ending positions of the simulated data are known. Therefore the difference between the measured values and the expected values can be compared for the five algorithms. From Chapter 1 it can be seen that heart sounds and murmurs contained frequencies up to 1000 Hz, that is 1 ms per cycle, and approximately >100ms long. Using this information it was decided that the ideal extracted transient's positions should have the following characteristics

1. For the early detection of the transient starting positions the algorithms are allowed a 5 ms difference between measured and expected values. This is 5% of the shortest duration of a heart sound. A 0.25 ms difference was allowed for the late detection of the starting position of the transient. This is because a quarter cycle of the highest frequency heart sound may contribute little information to the classification process.
Therefore difference between the measured and expected starting positions, ie. (Measured - Expected), must lie between -5ms and 0.25 ms, Figure 4-3, where zero is the expected starting position of the transient.

2. The peaks toward the end of the transient may be much smaller than the maximum peak, and thus more difficult to detect. Therefore it was decided that a 5ms difference between measured and expected values, either way, of the transient's ending position should be accepted. This can be observed in Figure 4-3.

![Figure 4-3 Diagram showing the allowed difference between the extracted transient and the generated transient (expected transient).](image)

These conditions are used to analyse and the results are compared between five extraction algorithms and also among the simulated heart sounds that have been processed by the same algorithm. The evaluation method was achieved by calculating the percentage of measured positions that are situated within the defined criteria for each of the simulated heart sounds (normal heart sound, NHS, splitting of the first heart sound, S1, splitting of the second heart sound, S2, third heart sound, A3, and fourth heart sound, A4). This percentage can be obtained by looking at the measured values and noting down the positions that occurred within the set criteria. Using these observations the percentage of positions that are within the specified times can be calculated. This
method can be inaccurate due to the small number of simulated transients in the signal. Therefore we need a more general method of calculation to obtain the percentage of positions. The general method of calculation relies upon the frequency distribution of the measured values and the percentage of positions can be obtained by the following procedures.

1. The difference between the measured and expected starting or ending positions are calculated, ie. Difference = Measured position - Expected position.

2. The frequency distribution of the difference values can be observed by plotting the histogram of the values. The mean, \( \mu \), and standard deviation, \( \sigma \), of the differences are evaluated assuming that the differences are Normally distributed.

3. The percentage of the positions that are within the set criteria is equal to the area between the set criteria under the probability density function corresponding to the Gaussian (Normal) distribution with parameter \( \mu \) and \( \sigma \), Figure 4-4. The differences probability density function is converted into a standard normal distribution curve (with a mean of zero and a standard deviation of one) by Equation 4.1.

\[
ND = \frac{OD - \mu}{\sigma}
\]

where
\( ND \) - New differences
\( OD \) - Old differences

The function that defines a normal distribution curve is stated in Equation 4.2 and the area under the curve is the integral of Equation 4.2.

\[
PDF = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}
\]

To calculate the percentage of positions we need to integrate Equation 4.2 between the set criteria, but the set criteria must first be scaled with the factor defined in Equation 4.1.
These criteria are chosen so that a comparison can be made between the five signal transformation processes for the extraction algorithm. The disadvantages of the given criteria are:

1. A false mean or standard deviation could be calculated due to a large value in the measured differences. Therefore the scaling factor, Equation 4-1, could be grossly miscalculated which then in turn affects the position percentage.

2. The position percentages do not give an indication about the number of difference values that are positive or negative values. This is important because the wanted differences for starting positions must be $\leq 0.25$ ms, and $\geq 5$ ms for ending positions. The reason for setting a lower and upper limit for the starting and ending positions, respectively, is that the differences could have a large negative or positive value. This could result in longer extracted transients than necessary for the classification process. From the explanations, if the position percentage is small but most of the difference values are $\leq 0.25$ ms, for starting positions, and $\geq 5$ ms, for ending positions, then the transform method is suitable for the detection of starting or ending positions.
Abbreviations

The following abbreviations are used in the tables in this chapter.

- **SP** - Starting position.
- **μ** - Mean.
- **σ** - Standard deviation.
- **ULC** - Upper limit criterion.
- **LLC** - Lower limit criterion.
- **AUL** - Adjusted upper limit for integration.
- **ALL** - Adjusted lower limit for integration.
- **Q %** - Percentage of measured positions that are within the set criteria.

### 4.3 Low Pass Filtering

#### 4.3.1 Methods

A simple method that can be used to highlight the transients of interest is the low-pass filter algorithm. The transformed signal can be acquired by putting the absolute value of the original data through a low pass filter. For this method an analogue filter was chosen instead of a digital filter, even though digital filters can give a linear or zero phase response. This is because if the study proved that an analogue filter is feasible as a signal transformer it could be implemented as a piece of simple hardware.

There are a number of analogue filters that could be used for the transformation process. It is recognised that different filters have different characteristics and therefore compromises are needed to be made when choosing a filter for the transformation process.

#### 4.3.2 Results and Discussions

A Chebyshev filter was chosen to transform the simulated data. Figure 4-5 and Figure 4-6 showed one second of the transformed signals obtained by low pass filtering the absolute values of the original signals, simulated normal heart sound. The figures illustrate the increasing phase shift resulting from increasing filter order.
Therefore it was decided to set the filter order to one and a 5Hz cut off frequency was adequate to obtain the appropriate transformed signal, Figure 4-7. This filter was used as a signal transformer, Figure 4-1.

Figure 4-5  First order, 10Hz cut off frequency.

Figure 4-6  Second order, 10Hz cut off frequency.
Figure 4-7 Raw data and its transformed signal (first order, 5Hz cut off frequency).

The results of the extracted starting and ending position of the transients can be observed in Table 4-1 and 4-2 respectively. The upper half of the tables showed the difference between the measured and expected values and the bottom half showed the values calculated using the analysis procedure in section 4-2.

From Table 4-1 it can be seen that all the calculated differences are positive and the percentage of positions that lay within the given criteria is very small. This confirmed our suspicion that the phase shift characteristic of an analogue filter is not suitable for the detection of a transient's starting position. Whereas, the phase shift is advantageous for the detection of ending positions, Table 4-2. The average percentage of the difference in ending positions is ~ 21%. A closer look at the ending position differences table revealed that most of the calculated difference values are positive and some are greater than 5 ms.

These observations indicated that the low pass filter is not a suitable method for the detection of the transients' starting positions, but it is a suitable method for the detection of ending positions.
**Table 4-1 Transformed signal’s starting position differences (ms), low pass filter.**

<table>
<thead>
<tr>
<th>NHS</th>
<th>S1</th>
<th>S2</th>
<th>A3</th>
<th>A4</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.9</td>
<td>3.6</td>
<td>2.8</td>
<td>3.1</td>
<td>2.8</td>
</tr>
<tr>
<td>3.4</td>
<td>3.6</td>
<td>4.4</td>
<td>3.9</td>
<td>3.7</td>
</tr>
<tr>
<td>3.0</td>
<td>3.7</td>
<td>3.0</td>
<td>6.7</td>
<td>6.4</td>
</tr>
<tr>
<td>3.6</td>
<td>4.0</td>
<td>4.5</td>
<td>3.0</td>
<td>2.8</td>
</tr>
<tr>
<td>3.1</td>
<td>3.7</td>
<td>3.0</td>
<td>4.0</td>
<td>3.8</td>
</tr>
<tr>
<td>3.4</td>
<td>3.9</td>
<td>4.2</td>
<td>6.5</td>
<td>6.2</td>
</tr>
<tr>
<td>3.1</td>
<td>3.7</td>
<td>3.1</td>
<td>3.0</td>
<td>2.7</td>
</tr>
<tr>
<td>4.4</td>
<td>4.7</td>
<td>4.6</td>
<td>3.7</td>
<td>3.6</td>
</tr>
<tr>
<td>2.8</td>
<td>3.5</td>
<td>2.8</td>
<td>6.3</td>
<td>6.5</td>
</tr>
<tr>
<td>3.8</td>
<td>4.0</td>
<td>4.2</td>
<td>3.0</td>
<td>2.7</td>
</tr>
<tr>
<td>2.9</td>
<td>3.5</td>
<td>2.8</td>
<td>4.7</td>
<td>4.7</td>
</tr>
<tr>
<td>3.7</td>
<td>3.6</td>
<td>4.7</td>
<td>6.3</td>
<td>6.1</td>
</tr>
<tr>
<td>2.9</td>
<td>3.4</td>
<td>2.9</td>
<td>3.0</td>
<td>2.7</td>
</tr>
<tr>
<td>4.0</td>
<td>4.2</td>
<td>4.5</td>
<td>4.1</td>
<td>3.9</td>
</tr>
<tr>
<td>2.9</td>
<td>3.7</td>
<td>2.9</td>
<td>6.3</td>
<td>6.2</td>
</tr>
<tr>
<td>3.7</td>
<td>3.9</td>
<td>4.6</td>
<td>2.7</td>
<td>2.5</td>
</tr>
<tr>
<td>2.9</td>
<td>3.6</td>
<td>2.9</td>
<td>3.8</td>
<td>3.6</td>
</tr>
<tr>
<td>3.4</td>
<td>3.7</td>
<td>4.2</td>
<td>6.3</td>
<td>5.5</td>
</tr>
<tr>
<td>3.0</td>
<td>3.7</td>
<td>3.0</td>
<td>3.1</td>
<td>2.8</td>
</tr>
<tr>
<td>3.5</td>
<td>3.8</td>
<td>4.4</td>
<td>4.2</td>
<td>4.0</td>
</tr>
<tr>
<td>3.1</td>
<td>3.7</td>
<td>3.1</td>
<td>6.5</td>
<td>6.6</td>
</tr>
<tr>
<td>3.4</td>
<td>3.9</td>
<td>4.0</td>
<td>3.0</td>
<td>2.7</td>
</tr>
<tr>
<td>3.0</td>
<td>3.7</td>
<td>3.0</td>
<td>3.9</td>
<td>3.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6.6</td>
<td>6.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.0</td>
<td>2.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.7</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6.2</td>
<td>6.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.0</td>
<td>2.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.8</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6.6</td>
<td>6.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.2</td>
<td>2.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.9</td>
<td>3.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6.3</td>
<td>6.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.0</td>
<td>2.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>μ</th>
<th>3.3</th>
<th>3.7739</th>
<th>3.6391</th>
<th>4.4235</th>
<th>4.1941</th>
</tr>
</thead>
<tbody>
<tr>
<td>σ</td>
<td>0.42</td>
<td>0.2734</td>
<td>0.762</td>
<td>1.4697</td>
<td>1.4773</td>
</tr>
<tr>
<td>ULC</td>
<td>0.26</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>LLC</td>
<td>-5.00</td>
<td>-5.00</td>
<td>-5.00</td>
<td>-5.00</td>
<td>-5.00</td>
</tr>
<tr>
<td>AUL</td>
<td>-7.263</td>
<td>-12.89</td>
<td>-4.445</td>
<td>-2.84</td>
<td>-2.67</td>
</tr>
<tr>
<td>ALL</td>
<td>-19.76</td>
<td>-32.09</td>
<td>-11.34</td>
<td>-6.412</td>
<td>-6.224</td>
</tr>
<tr>
<td>Θ %</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NHS</td>
<td>S1</td>
<td>S2</td>
<td>A3</td>
<td>A4</td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>8.1</td>
<td>-0.3</td>
<td>5.3</td>
<td>-22.9</td>
<td>4.7</td>
<td></td>
</tr>
<tr>
<td>14.7</td>
<td>17.7</td>
<td>11.7</td>
<td>11.4</td>
<td>3.2</td>
<td></td>
</tr>
<tr>
<td>6.7</td>
<td>-1.6</td>
<td>6.8</td>
<td>18.1</td>
<td>13.0</td>
<td></td>
</tr>
<tr>
<td>17.0</td>
<td>18.2</td>
<td>12.7</td>
<td>-24.3</td>
<td>6.3</td>
<td></td>
</tr>
<tr>
<td>4.1</td>
<td>-0.1</td>
<td>4.2</td>
<td>14.2</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>16.0</td>
<td>17.6</td>
<td>13.3</td>
<td>18.2</td>
<td>11.9</td>
<td></td>
</tr>
<tr>
<td>4.0</td>
<td>0.4</td>
<td>4.3</td>
<td>-17.6</td>
<td>6.4</td>
<td></td>
</tr>
<tr>
<td>15.3</td>
<td>16.4</td>
<td>11.8</td>
<td>13.4</td>
<td>4.6</td>
<td></td>
</tr>
<tr>
<td>4.1</td>
<td>2.0</td>
<td>4.1</td>
<td>17.1</td>
<td>12.0</td>
<td></td>
</tr>
<tr>
<td>17.3</td>
<td>17.5</td>
<td>12.6</td>
<td>-16.2</td>
<td>6.1</td>
<td></td>
</tr>
<tr>
<td>4.3</td>
<td>-0.8</td>
<td>4.4</td>
<td>11.7</td>
<td>6.9</td>
<td></td>
</tr>
<tr>
<td>16.1</td>
<td>17.0</td>
<td>11.9</td>
<td>18.4</td>
<td>12.9</td>
<td></td>
</tr>
<tr>
<td>8.7</td>
<td>0.5</td>
<td>5.7</td>
<td>-22.6</td>
<td>5.9</td>
<td></td>
</tr>
<tr>
<td>14.4</td>
<td>16.8</td>
<td>13.3</td>
<td>14.6</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td>5.6</td>
<td>-0.4</td>
<td>6.6</td>
<td>18.5</td>
<td>12.0</td>
<td></td>
</tr>
<tr>
<td>14.7</td>
<td>17.6</td>
<td>12.5</td>
<td>-15.7</td>
<td>5.7</td>
<td></td>
</tr>
<tr>
<td>2.1</td>
<td>-0.7</td>
<td>2.1</td>
<td>13.8</td>
<td>5.0</td>
<td></td>
</tr>
<tr>
<td>14.7</td>
<td>18.0</td>
<td>13.3</td>
<td>19.4</td>
<td>13.2</td>
<td></td>
</tr>
<tr>
<td>2.3</td>
<td>0.0</td>
<td>2.3</td>
<td>-22.6</td>
<td>6.5</td>
<td></td>
</tr>
<tr>
<td>16.7</td>
<td>19.6</td>
<td>13.3</td>
<td>11.2</td>
<td>4.6</td>
<td></td>
</tr>
<tr>
<td>14.3</td>
<td>17.4</td>
<td>13.2</td>
<td>-0.05</td>
<td>7.3</td>
<td></td>
</tr>
<tr>
<td>4.8</td>
<td>0.3</td>
<td>4.8</td>
<td>12.3</td>
<td>5.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>17.6</td>
<td>12.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-21.3</td>
<td>5.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>12.4</td>
<td>3.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>17.4</td>
<td>13.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-22.6</td>
<td>4.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>12.8</td>
<td>3.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>18.7</td>
<td>13.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-22.1</td>
<td>6.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>12.3</td>
<td>3.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>17.7</td>
<td>13.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-23.2</td>
<td>6.5</td>
<td></td>
</tr>
</tbody>
</table>

| £ | 9.687 | 8.2696 | 8.4478 | 2.5147 | 7.45 |
| µ = | 5.6947 | 9.1078 | 4.2998 | 17.625 | 3.9876 |
| σ = | 5.0 | 5.0 | 5.0 | 5.0 | 5.0 |
| ULC = | -5.0 | -5.0 | -5.0 | -5.0 | -5.0 |
| LLC = | -0.823 | -0.359 | -0.802 | 0.1394 | -0.614 |
| AUL = | -2.579 | -1.457 | -3.128 | -0.422 | -3.122 |
| % ≤ | 12 | 29 | 13 | 30 | 19 |

Table 4-2 Transformed signal's ending position differences (ms), low pass filter.
4.4 Signal's Activity

4.4.1 Methods

From the study of electromyography, EMG, Gerber et al (1984) define the level of activity of a signal as the sum of the absolute change in gradients within a specified window, Equation 4-3.

\[ Ac_n = \sum_{i=n}^{n+N-1} |x[i+m] - x[i]| \]  \hspace{1cm} 4-3

where \( N \) is the window size
\( m = 1 \)

A transformed signal is obtained by using Equation 4-3. The characteristic of the transformed signal is that it highlights the sections of the signal which are most active, i.e. the sections that have large changes in gradients. The disadvantages of this method are the tedious procedure to select a window size, \( N \), and the level of noise in the recorded signal. The window size must be obtained by trial and error to acquire an appropriate transformed signal such that the transients can be extracted correctly, thus \( N \) can vary from one signal to the next.

4.4.2 Results and Discussion

![Figure 4-8 Calculated signal's activity with a window size of 20ms.](image-url)
This algorithm was designed to highlight the most active part of the signal using the absolute gradients of the signal but Figure 4-8 demonstrated otherwise. This is because the simulated heart sound contained white noise to simulate background and tissue noise. The power spectral density of white noise is constant over the entire frequency range. It is the high frequency part of the generated noise that yielded large absolute gradients. These large absolute gradients are much more significant than the absolute gradient of the transient. Therefore summing the absolute gradients within a window resulted in a large value which in turn gave rise to a meaningless transformed signal, Figure 4-8. However it would probably have been possible to have yielded a more useful transformed signal by optimising the value of m. An appropriate value of m may be the number of samples in a half cycle of the oscillatory part of the transient. Nevertheless it was felt that the process of both having to optimise the window size, N, and the delay unit, m, was too cumbersome for an operator to do each time a new recording was to be analysed. Therefore the signal's activity algorithm is not a suitable method to automatically obtain a transformed signal for a noisy original signal.

4.5 Signal's Energy

4.5.1 Methods

Since chest sounds and speech sounds have similar characteristics, it was found that the detection processes used for speech can also be used for chest sounds. One such process is the calculation of the energy of an acoustic signal. From Parseval's theorem the energy of a discrete-time signal is defined as

$$E = \sum_{n=0}^{N-1} x[n]^2 = \frac{1}{N} \sum_{k=0}^{N-1} |X[k]|^2$$  \hspace{1cm} 4-4$$

The right-hand side of Equation 4-4 is the mean square spectral amplitude, while the left-hand side is the sum of the squared magnitudes of the time series. Such a value obtained from Equation 4-4 gives little information about the properties of the
signal. Therefore it is proposed that short-time energy may be used instead (Rabiner, 1978). Short-time analysis is a technique in which short segments of the signal are isolated and processed. These short segments are called analysis frames and they can overlap one another. When the analysis process is repeated periodically a new time-dependent signal is obtained from the original signal. The transformed signal can then be used to extract the transients of interest.

Short time energy is calculated by summing the squared amplitudes of the spectral or time series within a chosen analysis frame size $N$. Thus Equation 4-4 becomes

$$E_{sn} = \sum_{n=0}^{uN}(x[m]w[m])^2 = \frac{1}{N}\sum_{k=0}^{wN}(X[k]W[k])^2 \quad \text{--------- 4-5}$$

where \( w(m) \) or \( W(k) = \begin{cases} 1 & \text{for } n \leq (mk) \leq n + N \\ 0 & \text{for otherwise} \end{cases} \)

Equation 4-5 shows that there are two ways in which short-time energy can be calculated. To optimise the number of calculations involved and thus reduce processing time, the time series is used for the transformation process.

The main problem concerning short-time calculations is the size of the analysis frame used for the calculations. If the frame size, $N$, were to be very long, $E_s$ would change very little and so the transformed signal would not give good discrimination between sound and silence intervals. On the other hand, if $N$ were too short the transformed signal would not provide a smooth representation of the energy function.
4.5.2 Results and Discussions

Figure 4-9 to Figure 4-11 illustrates the transformed signals obtained using a window size of 50 ms, 20 ms and 5 ms from a simulated normal heart sound. It can be seen that the energy algorithm introduced very little noise, compared to the other algorithms, in the transformed signal. This is because the sum of the squared values for a transient of interest within the analysis window is very large compared to the background noise. The figures also show the effect of the window size. From these observations it was decided that a 20 ms window size is adequate for the signal transformation process.
The average percentage for the starting positions that are within the specified criteria is ~19%. When looking at the upper half of the Table 4-3 it was noted that most of the differences between the measured and expected values are negative and the mean of the simulated heart sounds differences are less than -5 ms, which is one of the criteria for starting positions in section 4-2. These observations confirmed that the algorithm is suitable for extracting the starting positions of the transients. Whereas the results of differences for the ending positions, Table 4-4, showed that some of the detected ending positions are much earlier than the expected. This is because toward the end of the transients the magnitude if the values are small and became much smaller when squared, compared to the maximum peak of the transient.
<table>
<thead>
<tr>
<th></th>
<th>NHS</th>
<th>S1</th>
<th>S2</th>
<th>A3</th>
<th>A4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-7.5</td>
<td>-5.0</td>
<td>-6.5</td>
<td>-6.7</td>
<td>-7.2</td>
</tr>
<tr>
<td>2</td>
<td>-4.5</td>
<td>-5.2</td>
<td>-4.8</td>
<td>-5.9</td>
<td>-6.2</td>
</tr>
<tr>
<td>3</td>
<td>-6.9</td>
<td>-5.0</td>
<td>-6.5</td>
<td>-2.7</td>
<td>-3.1</td>
</tr>
<tr>
<td>4</td>
<td>-5.2</td>
<td>-5.6</td>
<td>-6.6</td>
<td>-6.8</td>
<td>-7.1</td>
</tr>
<tr>
<td>5</td>
<td>-6.9</td>
<td>-5.9</td>
<td>-6.5</td>
<td>-6.8</td>
<td>-6.0</td>
</tr>
<tr>
<td>6</td>
<td>-5.5</td>
<td>-5.9</td>
<td>-4.8</td>
<td>-2.7</td>
<td>-3.4</td>
</tr>
<tr>
<td>7</td>
<td>-6.8</td>
<td>-6.1</td>
<td>-5.5</td>
<td>-6.8</td>
<td>-7.2</td>
</tr>
<tr>
<td>8</td>
<td>-5.5</td>
<td>-4.9</td>
<td>-4.8</td>
<td>-4.0</td>
<td>-6.3</td>
</tr>
<tr>
<td>9</td>
<td>-7.0</td>
<td>-6.2</td>
<td>-5.7</td>
<td>-2.8</td>
<td>-2.9</td>
</tr>
<tr>
<td>10</td>
<td>-6.1</td>
<td>-5.6</td>
<td>-4.9</td>
<td>-6.7</td>
<td>-7.1</td>
</tr>
<tr>
<td>11</td>
<td>-7.1</td>
<td>-6.2</td>
<td>-5.7</td>
<td>-5.0</td>
<td>-5.1</td>
</tr>
<tr>
<td>12</td>
<td>-6.3</td>
<td>-6.1</td>
<td>-4.5</td>
<td>-2.9</td>
<td>-3.5</td>
</tr>
<tr>
<td>13</td>
<td>-7.0</td>
<td>-6.1</td>
<td>-6.6</td>
<td>-6.8</td>
<td>-7.2</td>
</tr>
<tr>
<td>14</td>
<td>-6.9</td>
<td>-5.6</td>
<td>-4.8</td>
<td>-6.6</td>
<td>-6.9</td>
</tr>
<tr>
<td>15</td>
<td>-7.0</td>
<td>-5.9</td>
<td>-6.6</td>
<td>-2.7</td>
<td>-3.0</td>
</tr>
<tr>
<td>16</td>
<td>-6.2</td>
<td>-5.8</td>
<td>-4.4</td>
<td>-7.0</td>
<td>-7.4</td>
</tr>
<tr>
<td>17</td>
<td>-7.1</td>
<td>-6.1</td>
<td>-5.6</td>
<td>-6.8</td>
<td>-6.1</td>
</tr>
<tr>
<td>18</td>
<td>-6.5</td>
<td>-6.1</td>
<td>-4.9</td>
<td>-2.9</td>
<td>-3.6</td>
</tr>
<tr>
<td>19</td>
<td>-6.9</td>
<td>-6.0</td>
<td>-5.5</td>
<td>-6.6</td>
<td>-7.1</td>
</tr>
<tr>
<td>20</td>
<td>-6.4</td>
<td>-5.9</td>
<td>-4.7</td>
<td>-6.6</td>
<td>-6.9</td>
</tr>
<tr>
<td>21</td>
<td>-6.8</td>
<td>-6.0</td>
<td>-4.4</td>
<td>-2.8</td>
<td>-3.0</td>
</tr>
<tr>
<td>22</td>
<td>-6.5</td>
<td>-5.8</td>
<td>-4.9</td>
<td>-2.7</td>
<td>-3.5</td>
</tr>
<tr>
<td>23</td>
<td>-6.9</td>
<td>-6.0</td>
<td>-5.6</td>
<td>-6.9</td>
<td>-6.1</td>
</tr>
<tr>
<td>24</td>
<td>2.7</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
</tr>
<tr>
<td>25</td>
<td>-6.8</td>
<td>-7.2</td>
<td>-6.9</td>
<td>-7.1</td>
<td>-6.1</td>
</tr>
<tr>
<td>26</td>
<td>-2.9</td>
<td>-3.1</td>
<td>-2.8</td>
<td>-3.1</td>
<td>-3.1</td>
</tr>
<tr>
<td>27</td>
<td>-6.8</td>
<td>-7.1</td>
<td>-6.8</td>
<td>-7.1</td>
<td>-6.1</td>
</tr>
<tr>
<td>28</td>
<td>-5.7</td>
<td>-5.9</td>
<td>-5.8</td>
<td>-5.8</td>
<td>-5.8</td>
</tr>
<tr>
<td>29</td>
<td>-2.9</td>
<td>-3.3</td>
<td>-2.9</td>
<td>-3.3</td>
<td>-2.9</td>
</tr>
<tr>
<td>30</td>
<td>-6.8</td>
<td>-7.2</td>
<td>-6.8</td>
<td>-7.2</td>
<td>-6.8</td>
</tr>
</tbody>
</table>

Table 4.3 Transformed signal's starting position differences (ms), signal's energy.
\[ EP = Measured - Expected \]

<table>
<thead>
<tr>
<th>NHS</th>
<th>S1</th>
<th>S2</th>
<th>A3</th>
<th>A4</th>
</tr>
</thead>
<tbody>
<tr>
<td>-4.4</td>
<td>-3.5</td>
<td>-6.8</td>
<td>-4.2</td>
<td>-8.2</td>
</tr>
<tr>
<td>-2.9</td>
<td>-0.9</td>
<td>-3.6</td>
<td>-4.1</td>
<td>-16.6</td>
</tr>
<tr>
<td>-4.8</td>
<td>-14.0</td>
<td>-7.0</td>
<td>-0.9</td>
<td>-3.5</td>
</tr>
<tr>
<td>-1.9</td>
<td>0.5</td>
<td>-4.8</td>
<td>-42.0</td>
<td>-11.8</td>
</tr>
<tr>
<td>-5.6</td>
<td>-12.5</td>
<td>-8.0</td>
<td>-3.4</td>
<td>-19.2</td>
</tr>
<tr>
<td>-2.6</td>
<td>0.9</td>
<td>-2.8</td>
<td>-1.8</td>
<td>-4.3</td>
</tr>
<tr>
<td>-6.9</td>
<td>-13.0</td>
<td>-8.2</td>
<td>-42.6</td>
<td>-8.8</td>
</tr>
<tr>
<td>-2.2</td>
<td>-2.5</td>
<td>-3.9</td>
<td>-3.8</td>
<td>-16.0</td>
</tr>
<tr>
<td>-6.0</td>
<td>-12.4</td>
<td>-34.3</td>
<td>-2.1</td>
<td>-4.8</td>
</tr>
<tr>
<td>-2.7</td>
<td>0.4</td>
<td>-3.4</td>
<td>-41.9</td>
<td>-9.4</td>
</tr>
<tr>
<td>-5.7</td>
<td>-13.3</td>
<td>-34.5</td>
<td>-4.1</td>
<td>-8.4</td>
</tr>
<tr>
<td>-3.1</td>
<td>1.0</td>
<td>-3.5</td>
<td>-1.5</td>
<td>-4.4</td>
</tr>
<tr>
<td>-5.0</td>
<td>-13.3</td>
<td>-5.9</td>
<td>-42.2</td>
<td>-13</td>
</tr>
<tr>
<td>-2.9</td>
<td>-0.1</td>
<td>-3.4</td>
<td>-3.6</td>
<td>-20.9</td>
</tr>
<tr>
<td>-4.7</td>
<td>-12.6</td>
<td>-33.8</td>
<td>-1.5</td>
<td>-4.4</td>
</tr>
<tr>
<td>-3</td>
<td>0.6</td>
<td>-3.8</td>
<td>-41.3</td>
<td>-11.4</td>
</tr>
<tr>
<td>-6.5</td>
<td>-13.2</td>
<td>-34.5</td>
<td>-3.1</td>
<td>-9.9</td>
</tr>
<tr>
<td>-2.0</td>
<td>1.3</td>
<td>-2.9</td>
<td>-1.6</td>
<td>-4.1</td>
</tr>
<tr>
<td>-6.3</td>
<td>-13.3</td>
<td>-3.9</td>
<td>-42.2</td>
<td>-8.5</td>
</tr>
<tr>
<td>-2.3</td>
<td>1.4</td>
<td>-2.8</td>
<td>-4.8</td>
<td>-16.4</td>
</tr>
<tr>
<td>-6.2</td>
<td>-14.5</td>
<td>-7.4</td>
<td>-1.9</td>
<td>-4.1</td>
</tr>
<tr>
<td>-2.8</td>
<td>0.6</td>
<td>-3.4</td>
<td>-41.0</td>
<td>-7.0</td>
</tr>
<tr>
<td>-5.1</td>
<td>-13.4</td>
<td>-7.1</td>
<td>-4.4</td>
<td>-10.4</td>
</tr>
</tbody>
</table>

Table 4-4 Transformed signal's ending position differences (ms), signal's energy.
4.6 Short-Time Average Zero Crossing Rate

4.6.1 Methods

Another algorithm that can be used for the detection of speech sound is the short-time average zero crossing rate. Zero crossing is said to occur if successive samples have different algebraic signs. The rate at which zero crossing occurs is approximately proportional to the highest frequency component of a signal (Rabiner, 1978). An example of the average zero crossing rate can be demonstrated using a sinusoidal wave of frequency $F_0$ which is sampled at $F_s$. The sinusoidal wave has $F_s/F_0$ samples per cycle and each cycle has two zero crossings thus the average rate of zero crossing is

$$Z = \frac{2F_0}{F_s} \text{ crossings/sample} \quad \text{(4-6)}$$

This technique of zero crossing rate gives a reasonable estimate of the frequency component of narrow band signals but it becomes less precise for broad-band signals (Rabiner, 1978). However, estimates of the signal spectral properties can be obtained by using the short time average zero crossing. The short time average zero crossing is defined as

$$Z_{s_k} = \sum_{m=n}^{m+n+N} \left| \text{sgn}(x[m]) - \text{sgn}(x[m-1]) \right| w[m] \quad \text{(4-7)}$$

where

$$\text{sgn}(x[m]) = \begin{cases} 1 & \text{for } x[m] \geq 0 \\ -1 & \text{for } x[m] < 0 \end{cases}$$

and

$$w[m] = \begin{cases} \frac{1}{2N} & \text{for } n \leq m \leq n+N \\ 0 & \text{for otherwise} \end{cases}$$

Equation 4-7 seems to be a complicated operation but the calculation is quite simple. It involves checking samples in pairs for a change of algebraic signs, in order to
obtain the average of the zero crossing rate over a specific frame size, \( N \). Using this algorithm a transformed signal is obtained and then used to extract transients of interest.

In addition to having the short-time process disadvantages the short-time average zero crossing rate algorithm can also be affected by noise and DC offset. The noise and DC offset can be minimised by careful design of equipment used. Due to the sensitivity of the algorithm to noise a variety of algorithms have been proposed (Rabiner, 1978), to overcome this problem.

### 4.6.2 Results and Discussions

Using Equation 4-7 the calculated transformed signal highlighted transients of interest by the troughs of the transformed signal. Therefore the calculated transformed signal was inverted and then normalised to unity for the extraction algorithm.

Figures 4-12 to 4-14 show the transformed signals obtained by using three different window sizes. By studying the transformed signals obtained by a number of window sizes it was decided that a window size of 10 ms can be used for the signal transformation process.

![Figure 4-12 50ms window size.](image)
The chosen window size was used to transform a number of simulated heart sounds and their positions were extracted. Table 4-5 and Table 4-6 showed the results of the extraction process. The two tables showed that the percentages of positions that are within the set criteria are ~97% for starting positions and ~68% for ending positions. These percentages seem to indicate that the average zero crossing rate is a good method for extracting transients. It must be noted that the signals used to test the algorithms are simulated and they do not possess a realistic effect of noise in the signal as real recorded data. These noise effects include base line wandering, DC shift and spurious noise. The effect of these noises can prove to be serious when using this algorithm. This is because the principal method of the algorithm is that it measured the zero crossing rate of the signal and the amount of zero crossing rate can be wrongly calculated if the signal
contained dc shift or base line wandering. The circumstances can be remedied by using various filters to get rid of these unwanted noise effects.

\[ SP = \text{Measured} - \text{Expected} \]

<table>
<thead>
<tr>
<th></th>
<th>NHS</th>
<th>S1</th>
<th>S2</th>
<th>A3</th>
<th>A4</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2.6</td>
<td>-2.9</td>
<td>-2.5</td>
<td>-1.7</td>
<td>-2.1</td>
<td></td>
</tr>
<tr>
<td>-1.4</td>
<td>-1.5</td>
<td>-0.5</td>
<td>0.6</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>-4.0</td>
<td>-3.6</td>
<td>-3.7</td>
<td>0.6</td>
<td>-1.0</td>
<td></td>
</tr>
<tr>
<td>-2.3</td>
<td>-0.2</td>
<td>-1.1</td>
<td>0.1</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>-0.6</td>
<td>-0.6</td>
<td>-0.5</td>
<td>0.6</td>
<td>-0.6</td>
<td></td>
</tr>
<tr>
<td>-2.2</td>
<td>-2.8</td>
<td>-2.2</td>
<td>0.5</td>
<td>-1.3</td>
<td></td>
</tr>
<tr>
<td>-2.9</td>
<td>-2.9</td>
<td>-2.8</td>
<td>-2.7</td>
<td>-2.7</td>
<td></td>
</tr>
<tr>
<td>-0.9</td>
<td>-1.5</td>
<td>-2.7</td>
<td>-2.4</td>
<td>-2.4</td>
<td></td>
</tr>
<tr>
<td>-2.0</td>
<td>-2.0</td>
<td>-1.9</td>
<td>-0.4</td>
<td>-0.6</td>
<td></td>
</tr>
<tr>
<td>-1.1</td>
<td>-3.6</td>
<td>-0.7</td>
<td>-2.1</td>
<td>-2.1</td>
<td></td>
</tr>
<tr>
<td>-0.5</td>
<td>-0.3</td>
<td>-0.4</td>
<td>-3.0</td>
<td>-3.0</td>
<td></td>
</tr>
<tr>
<td>-1.1</td>
<td>-1.5</td>
<td>0.1</td>
<td>-0.5</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>-3.0</td>
<td>-3.7</td>
<td>-3.7</td>
<td>-2.3</td>
<td>-2.3</td>
<td></td>
</tr>
<tr>
<td>-2.3</td>
<td>-1.4</td>
<td>-0.9</td>
<td>-0.5</td>
<td>-0.5</td>
<td></td>
</tr>
<tr>
<td>-1.0</td>
<td>-0.7</td>
<td>-0.9</td>
<td>-2.0</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td>-5.7</td>
<td>-1.6</td>
<td>-6.0</td>
<td>-1.8</td>
<td>-1.8</td>
<td></td>
</tr>
<tr>
<td>-2.9</td>
<td>-2.4</td>
<td>-2.7</td>
<td>-2.0</td>
<td>-2.3</td>
<td></td>
</tr>
<tr>
<td>-0.4</td>
<td>-0.7</td>
<td>-0.4</td>
<td>-1.1</td>
<td>-0.5</td>
<td></td>
</tr>
<tr>
<td>-0.6</td>
<td>-0.6</td>
<td>-0.5</td>
<td>-2.0</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>-1.8</td>
<td>-0.6</td>
<td>-1.2</td>
<td>-0.2</td>
<td>-0.2</td>
<td></td>
</tr>
<tr>
<td>-2.2</td>
<td>-1.7</td>
<td>-1.9</td>
<td>1.6</td>
<td>-1.5</td>
<td></td>
</tr>
<tr>
<td>-2.3</td>
<td>-2.7</td>
<td>-2.6</td>
<td>-2.0</td>
<td>-2.0</td>
<td></td>
</tr>
<tr>
<td>-2.9</td>
<td>-2.6</td>
<td>-2.8</td>
<td>-2.2</td>
<td>-2.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>0.6</td>
<td>0.5</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.2</td>
<td>3.2</td>
<td>1.7</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.4</td>
<td>0.6</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.6</td>
<td>2.8</td>
<td>2.1</td>
<td>2.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.1</td>
<td>2.1</td>
<td>0.5</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-1.9</td>
<td>-1.9</td>
<td>0.5</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-2.0</td>
<td>-2.0</td>
<td>-0.5</td>
<td>-0.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.6</td>
<td>-3.8</td>
<td>-3.2</td>
<td>-3.2</td>
<td></td>
</tr>
</tbody>
</table>

\[ \mu = -2.065 -1.826 -1.813 -1.185 -1.391 \]

\[ \sigma = 1.2967 1.1165 1.4916 1.3578 1.2894 \]

\[ \text{ULC} = 0.25 \quad 0.25 \quad 0.25 \quad 0.25 \quad 0.25 \]

\[ \text{LLC} = -5.00 \quad -5.00 \quad -5.00 \quad -5.00 \quad -5.00 \]

\[ \text{AUL} = 1.7869 1.8595 1.3831 1.057 1.2728 \]

\[ \text{ALL} = -2.265 -2.843 -2.137 -2.809 -2.799 \]

| Qt. % | 99 | 100 | 97 | 93 | 96 |

Table 4-5 Transformed signal's starting position differences (ms), average zero crossing rate.

4-23
Table 4-6 Transformed signal's ending position differences (ms), average zero crossing rate.

<table>
<thead>
<tr>
<th>EP = Measured - Expected</th>
<th>N</th>
<th>S1</th>
<th>S2</th>
<th>A3</th>
<th>A4</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.0</td>
<td>-8.1</td>
<td>-5.2</td>
<td>-42.5</td>
<td>-6.9</td>
<td></td>
</tr>
<tr>
<td>4.5</td>
<td>-1.9</td>
<td>2.1</td>
<td>-4.2</td>
<td>-10.7</td>
<td></td>
</tr>
<tr>
<td>2.8</td>
<td>-11.9</td>
<td>2.9</td>
<td>-2.3</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>2.2</td>
<td>0.2</td>
<td>-1.6</td>
<td>-4.5</td>
<td>-6.0</td>
<td></td>
</tr>
<tr>
<td>4.2</td>
<td>-9.7</td>
<td>-4.4</td>
<td>-2.4</td>
<td>-7.3</td>
<td></td>
</tr>
<tr>
<td>2.3</td>
<td>-1.2</td>
<td>-2.3</td>
<td>-2.4</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>3.4</td>
<td>-1.3</td>
<td>-3.5</td>
<td>-6.4</td>
<td>-7.5</td>
<td></td>
</tr>
<tr>
<td>3.6</td>
<td>-0.4</td>
<td>-4.0</td>
<td>-3.3</td>
<td>-1.14</td>
<td></td>
</tr>
<tr>
<td>3.6</td>
<td>-0.6</td>
<td>-3.8</td>
<td>-3.5</td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td>3.5</td>
<td>0.2</td>
<td>-2.4</td>
<td>-38.2</td>
<td>-6.2</td>
<td></td>
</tr>
<tr>
<td>3.1</td>
<td>-11.9</td>
<td>-3.3</td>
<td>-4.6</td>
<td>-6.6</td>
<td></td>
</tr>
<tr>
<td>2.9</td>
<td>2.3</td>
<td>-1.1</td>
<td>-0.8</td>
<td>3.7</td>
<td></td>
</tr>
<tr>
<td>7.0</td>
<td>-9.7</td>
<td>-7.4</td>
<td>-41.1</td>
<td>-7.1</td>
<td></td>
</tr>
<tr>
<td>2.3</td>
<td>-1.5</td>
<td>-1.9</td>
<td>-4.6</td>
<td>-0.8</td>
<td></td>
</tr>
<tr>
<td>0.6</td>
<td>-4.5</td>
<td>-0.8</td>
<td>-4.2</td>
<td>-0.7</td>
<td></td>
</tr>
<tr>
<td>2.7</td>
<td>0.7</td>
<td>-1.2</td>
<td>-22.9</td>
<td>-6.1</td>
<td></td>
</tr>
<tr>
<td>1.1</td>
<td>-0.9</td>
<td>1.0</td>
<td>-3.6</td>
<td>-0.6</td>
<td></td>
</tr>
<tr>
<td>0.7</td>
<td>20</td>
<td>-2.2</td>
<td>-3.3</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>3.8</td>
<td>-102</td>
<td>-3.9</td>
<td>-43.3</td>
<td>-6.8</td>
<td></td>
</tr>
<tr>
<td>3.8</td>
<td>0.9</td>
<td>-1.6</td>
<td>-3.3</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>4.7</td>
<td>-105</td>
<td>-4.8</td>
<td>-4.0</td>
<td>3.2</td>
<td></td>
</tr>
<tr>
<td>3.7</td>
<td>-1.1</td>
<td>-0.8</td>
<td>-42.0</td>
<td>-6.6</td>
<td></td>
</tr>
<tr>
<td>4.6</td>
<td>-108</td>
<td>-5.2</td>
<td>-2.4</td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td>-2.9</td>
<td>0.1</td>
<td>-4.8</td>
<td>3.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-2.0</td>
<td>0.6</td>
<td>-4.8</td>
<td>-6.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-2.3</td>
<td>1.3</td>
<td>-4.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.7</td>
<td>2.2</td>
<td>-3.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-3.6</td>
<td>1.4</td>
<td>-3.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.6</td>
<td>3.2</td>
<td>-42.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-3.3</td>
<td>8.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

μ = -2.917  -4.439  -2.652  -16.72  -2.194
σ = 2.0858  5.3786  2.132  18.875  4.0371
ULC = 5.00  5.00  5.00  5.00  5.00
LLC = -5.00  -5.00  -5.00  -5.00  -5.00
AUL = 3.7958  1.755  3.5892  1.1506  1.782
AUL = -0.998  -0.104  -1.101  0.6208  -0.695
O % = 92  56  94  14  83

Table 4-6 Transformed signal's ending position differences (ms), average zero crossing rate.
4.7 Signal's Envelope

4.7.1 Methods

The transformed signal discriminates the background noise and the transients of interest by having the characteristic of the raw signal's envelope. Most of the algorithms described so far have the disadvantage of the difficult task of selecting the analysis frame size. Therefore an algorithm to obtain the signal's envelope without choosing the frame size is proposed below.

The envelope of a signal can be calculated by using the Hilbert transform (Whalen, 1971). The definition of the Hilbert transform is the convolution of $x(t)$ with

$$\frac{1}{\pi t}$$

where $\hat{x}(t)$ is the Hilbert transform of $x(t)$

$$\hat{x}(t) = x(t) * \frac{1}{\pi t} = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(\tau)}{t - \tau} d\tau$$  \hspace{1cm} 4-8

There are two methods that can be used to calculate Equation 4-8. The first is the evaluation of the convolution integral. The second is the use of the time-convolution theorem. The first method involves the manipulation of the two signals and then finding the area under the product of $x(\tau)$ and $1/(t-\tau)$ (Brigham, 1988). The second theorem allows one to convolve two functions in the time domain by a simple multiplication in the frequency domain. Assume that the Fourier transform of $x(t)$ is $X(f)$ and the Fourier transform of $\frac{1}{\pi t}$ is $-j\text{sgn}(f)$. Applying the time-convolution theorem to Equation 4-8 and the Fourier transform of the Hilbert transform is defined as

$$\mathcal{F}[\hat{x}(t)] = -jX(f)\text{sgn}(f)$$  \hspace{1cm} 4-9

where $\text{sgn}(f)$ is the signum function, ie.


\[ \text{sgn}(f) = \begin{cases} 
1 & \text{for } f > 0 \\
0 & \text{for } f = 0 \\
-1 & \text{for } f < 0 
\end{cases} \]

This means that the Hilbert transform can be calculated by taking the inverse Fourier transform of the product of the signum function and Fourier transform of \( x(t) \).

One of the properties of the Hilbert transform is that it behaves as a 90° phase shifter. Using this property, the analytical (or pre-envelope) of a real signal is defined as

\[ x_a(t) = x(t) + j \hat{x}(t) \]

This analytical signal contains a real part, which is the original signal, and an imaginary part, which contains the Hilbert transform. To obtain the signal's envelope, the absolute value of the analytical envelope is evaluated (Whalen, 1971).

4.7.2 Results and Discussions

Figure 4-15 shows the transformed signal obtained by using the Hilbert transform method. It can be seen that the transformed signal highlighted the wanted transient but it also contained a certain amount of noise. This amount of noise depended upon the noise which is in the original signal. If the noise in the transformed signal is too great for the extraction algorithm it can be filtered out or by squaring the transformed signal, Figure 4-16. From the two figures it can be observed that the squaring process reduced the noise in the transformed signal, but at the same time the squared transformed signal did not highlight the end of the transients. Therefore, if the filtering or the squaring process was to take place care must be taken so that the appropriate positions are obtained.
Table 4-7 and Table 4-8 show the analysis process from section 4.2 for this technique. The average percentage of measured positions that are within the given criteria is ~40%, for staring positions, and ~4%, for ending positions. Looking closer at the table for the starting positions it can be seen that most of the measured positions are negative and thus this algorithm is suitable for the detection of transients' starting positions and not for ending positions.

The extraction of the transients' positions was also carried out for the squared transformed signal. It was found that the percentages of positions that are within the set criteria are ~71% for starting positions and ~3% for ending positions. Comparing the percentages obtained for the transformed signal and its squared showed that the
starting position's percentage of the squared signal increase but at the same time the squared signal did not highlight the end part of the transients.

<table>
<thead>
<tr>
<th>SP = Measured - Expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>NHS</td>
</tr>
<tr>
<td>-5.6</td>
</tr>
<tr>
<td>-7.3</td>
</tr>
<tr>
<td>-7.2</td>
</tr>
<tr>
<td>-8.2</td>
</tr>
<tr>
<td>-3.6</td>
</tr>
<tr>
<td>-7.5</td>
</tr>
<tr>
<td>-7.1</td>
</tr>
<tr>
<td>-5.7</td>
</tr>
<tr>
<td>-4.5</td>
</tr>
<tr>
<td>-6.3</td>
</tr>
<tr>
<td>-6.8</td>
</tr>
<tr>
<td>-8.0</td>
</tr>
<tr>
<td>-7.3</td>
</tr>
<tr>
<td>-5.3</td>
</tr>
<tr>
<td>-5.9</td>
</tr>
<tr>
<td>-6.3</td>
</tr>
<tr>
<td>-7.8</td>
</tr>
<tr>
<td>-6.7</td>
</tr>
<tr>
<td>-6.7</td>
</tr>
<tr>
<td>-6.7</td>
</tr>
<tr>
<td>-5.6</td>
</tr>
<tr>
<td>-6.1</td>
</tr>
<tr>
<td>-5.6</td>
</tr>
<tr>
<td>-7.4</td>
</tr>
<tr>
<td>-6.2</td>
</tr>
<tr>
<td>-8.1</td>
</tr>
<tr>
<td>-3.6</td>
</tr>
<tr>
<td>-7.6</td>
</tr>
<tr>
<td>-5.6</td>
</tr>
<tr>
<td>-5.8</td>
</tr>
<tr>
<td>-9.1</td>
</tr>
<tr>
<td>-7.3</td>
</tr>
<tr>
<td>-2.6</td>
</tr>
<tr>
<td>-9.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>μ</th>
<th>-6.383</th>
<th>-4.717</th>
<th>-4.743</th>
<th>-6.735</th>
<th>-2.597</th>
</tr>
</thead>
<tbody>
<tr>
<td>σ</td>
<td>1.1436</td>
<td>2.2733</td>
<td>2.6209</td>
<td>1.845</td>
<td>3.2161</td>
</tr>
<tr>
<td>ULC</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>LLC</td>
<td>-5.00</td>
<td>-5.00</td>
<td>-5.00</td>
<td>-5.00</td>
<td>-5.00</td>
</tr>
<tr>
<td>AUL</td>
<td>5.7997</td>
<td>2.1851</td>
<td>1.9083</td>
<td>3.786</td>
<td>0.8853</td>
</tr>
<tr>
<td>ALL</td>
<td>1.209</td>
<td>-0.124</td>
<td>-0.098</td>
<td>0.9405</td>
<td>-0.747</td>
</tr>
<tr>
<td>Q</td>
<td>4.57</td>
<td>56</td>
<td>9</td>
<td>75</td>
<td></td>
</tr>
</tbody>
</table>

Table 4-7 Transformed signal’s starting position differences (ms), signal’s envelope.
### Table 4-8 Transformed signal's Ending position differences (ms), signal's envelope.

<table>
<thead>
<tr>
<th>NHS S1</th>
<th>S2</th>
<th>A3</th>
<th>A4</th>
</tr>
</thead>
<tbody>
<tr>
<td>-42.3</td>
<td>-23.0</td>
<td>-42.3</td>
<td>-46.4</td>
</tr>
<tr>
<td>-13.0</td>
<td>-10.6</td>
<td>-11.3</td>
<td>-13.1</td>
</tr>
<tr>
<td>-46.3</td>
<td>-22.4</td>
<td>-46.3</td>
<td>-23.4</td>
</tr>
<tr>
<td>-10.1</td>
<td>-7.8</td>
<td>-11.6</td>
<td>-47.2</td>
</tr>
<tr>
<td>-44.1</td>
<td>-24.2</td>
<td>-44.1</td>
<td>-8.7</td>
</tr>
<tr>
<td>-12.6</td>
<td>-6.1</td>
<td>-10.2</td>
<td>-2.8</td>
</tr>
<tr>
<td>-46.7</td>
<td>-23.7</td>
<td>-46.7</td>
<td>-50.8</td>
</tr>
<tr>
<td>-15.6</td>
<td>-10.0</td>
<td>-10.3</td>
<td>-9.6</td>
</tr>
<tr>
<td>-46.1</td>
<td>-33.3</td>
<td>-45.1</td>
<td>-2.2</td>
</tr>
<tr>
<td>-10.5</td>
<td>-6.6</td>
<td>-16.5</td>
<td>-48.9</td>
</tr>
<tr>
<td>-45.0</td>
<td>-22.1</td>
<td>-45.0</td>
<td>-13.7</td>
</tr>
<tr>
<td>-10.1</td>
<td>-5.2</td>
<td>-10.6</td>
<td>-4.2</td>
</tr>
<tr>
<td>-46.7</td>
<td>-24.0</td>
<td>-44.0</td>
<td>-47.2</td>
</tr>
<tr>
<td>-11.0</td>
<td>-8.1</td>
<td>-10.6</td>
<td>-10.4</td>
</tr>
<tr>
<td>-42.6</td>
<td>-28.8</td>
<td>-42.2</td>
<td>-3.2</td>
</tr>
<tr>
<td>-10.9</td>
<td>-6.5</td>
<td>-9.7</td>
<td>-48.7</td>
</tr>
<tr>
<td>-47.1</td>
<td>-22.3</td>
<td>-47.1</td>
<td>-8.4</td>
</tr>
<tr>
<td>-10.1</td>
<td>-8.2</td>
<td>-11.2</td>
<td>-1.9</td>
</tr>
<tr>
<td>-43.7</td>
<td>-21.3</td>
<td>-43.7</td>
<td>-51.4</td>
</tr>
<tr>
<td>-10.9</td>
<td>-6.8</td>
<td>-11.4</td>
<td>-9.5</td>
</tr>
<tr>
<td>-42.6</td>
<td>-23.7</td>
<td>-42.6</td>
<td>-4.1</td>
</tr>
<tr>
<td>-12.3</td>
<td>-8.3</td>
<td>-11.9</td>
<td>-46.0</td>
</tr>
<tr>
<td>-47.5</td>
<td>-22.8</td>
<td>-47.5</td>
<td>-11.9</td>
</tr>
</tbody>
</table>

\[
\begin{align*}
\sigma &= 17.11 & 7.9551 & 17.121 & 20.352 & 7.3674 \\
ULC &= 5.00 & 5.00 & 5.00 & 5.00 & 5.00 \\
LLC &= -5.00 & -5.00 & -5.00 & -5.00 & -5.00 \\
AUL &= 1.9837 & 2.9972 & 1.9726 & 1.3049 & 2.8468 \\
ALL &= 1.3993 & 1.3417 & 1.3885 & 0.8135 & 1.4895 \\
\sigma \% &= 2 & 3 & 2 & 9 & 2 
\end{align*}
\]

The table above presents the transformed signal's ending position differences (ms) for signals' envelopes. Each cell represents the difference between the measured (NHS) and expected values for different signal types (S1, S2, A3, A4) and various parameters. The table includes mean (\(\mu\)), standard deviation (\(\sigma\)), upper limit (ULC), lower limit (LLC), and other related values along with their percentage (\(\sigma \%\)).
4.8 Discussion

This chapter described five signal transformation processes that can be used to extract transients of interest. A comparison between the method for calculation of the algorithms can be seen in Table 4-9.

The algorithms 2, 3 and 4 are commonly called short-time calculation, Table 4-9. This is because they required a specific window size to calculate the transformed signal. These short-time methods can be implemented in a fast and efficient manner but they require a certain amount of effort from the user to obtain the appropriate window size. Furthermore the recorded signal may vary considerably between recordings in terms of the amplitudes of both signals and background noise, hence it may be difficult to obtain a transformed signal for the extraction process. The difficult task of selecting an appropriate window size was overcome by algorithm 5, signal's envelope algorithm. The

<table>
<thead>
<tr>
<th>Algorithms</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Analogue Filter, low pass</td>
<td>• Can be implemented as a simple hardware</td>
<td>• Choosing filter order</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Phase distortion</td>
</tr>
<tr>
<td>2. Signal's Activity</td>
<td>• Fast to calculate</td>
<td>• Choosing window size</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Not suitable for low S/N ratio signals</td>
</tr>
<tr>
<td>3. Signal's Energy</td>
<td>• Fast to calculate</td>
<td>• Choosing window size</td>
</tr>
<tr>
<td>4. Average Zero Crossing Rate</td>
<td>• Fast to calculate</td>
<td>• Choosing window size</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Sensitive to noise such as DC offset and base line wandering</td>
</tr>
<tr>
<td>5. Signal's Envelope</td>
<td>• Correct envelope of the signal can be obtained</td>
<td>• Fourier transform calculations</td>
</tr>
</tbody>
</table>

Table 4-9 A comparison between different transformation algorithms

disadvantage of the signal's envelope algorithm is that it used the Fourier transform to calculate the transformed signal and thus longer processing time compared to algorithms
1, 2, 3 and 4. This may not be a problem given the recent advent of fast and cheap computers and the development of digital signal processors.

Table 4-10 shows the percentage of transients that lay within the given criteria for algorithms 1 to 5 using the simulated heart sounds; NHS, S1, S2, A3, and A4. In order to evaluate the quality of algorithms the average percentage for each algorithm is calculated. The result seems to indicate that the average zero crossing rate is the best algorithm out of the five tested for the detection of starting and ending positions. For this algorithm it must be noted that the signals that were used to test the algorithms were simulated signals. The simulated signals contained white noise generated by a random generator. The generated numbers could rapidly alternate about zero, Figure 4-17.

Table 4-10 Percentage of detected positions that are within the given criteria. Number 1-5 correspond to the algorithms in Table 4-9 and NHS, S1, S2, A3 and A4 are the simulated heart sounds.

<table>
<thead>
<tr>
<th></th>
<th>Starting position</th>
<th></th>
<th>Ending position</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NHS</td>
<td>S1</td>
<td>S2</td>
<td>A3</td>
</tr>
<tr>
<td>1.</td>
<td>0 0 0 0 0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>- - - - -</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>0 0 16 45 33</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>59 100 97 93 96</td>
<td>97</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>4 57 55 9 75</td>
<td>40</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4-17 An example of generated white noise.
If this happened a large number of zero crossings can be counted. This would in turn ensure the transients of interest are properly highlighted if an appropriate window size is chosen. In a recorded signal the noise contained is not white noise and the recorded signal may contain a DC shift or base line wander. This would have resulted in a decrease in the number of zero crossings and the calculated transformed signal of a recorded signal obtained using the average zero crossing rate algorithm could be completely different from the transformed signal obtained for the simulated signal.

Generated white noise or high frequency noises can also give large gradients. This study of the signal's activity algorithm showed that the method can not be used to obtain a suitable transformed signal for the transients' extraction process.

The rest of the algorithms (1, 3 and 5) demonstrated that they have the ability to detect starting positions better than ending positions, or vice versa. A closer look at the detected starting positions for signal's energy and signal's envelope revealed that the algorithms detected most of the starting positions earlier than the expected values, but some are outside the set criteria. Therefore these two algorithms are suitable for extracting the transient's starting position and not ending position. By selecting the window size for the signal's energy algorithm a better percentage can be obtained. The position's percentage for the signal's envelope algorithm can be improved by manipulating its transformed signal, ie. filtering or squaring the transformed signal.

The low pass filter method showed that it is capable of extracting the transient's ending positions. This is because extraction algorithm took advantage of a phase shift in the transformed signal due to the filter's characteristic.

In conclusion the signal's envelope technique is preferred because the transformed signal calculation is consistent for all recorded signal and the users do not need to find the appropriate window size for the transformation process. The reason for a poor performance in the detection of the transients' ending positions is due to the small amplitudes toward the end of the transients. These small amplitude values reflect upon
the transformed signal. To overcome the problem the user could further process the
signal’s envelope by simple calculations, i.e. squaring the transformed signal.

References


program for quantitative EMG signal analysis. IEEE Trans. on Biomedical Engineering
1984;vol 31; p857-63.


Chapter 5

Classification of Chest Sounds

5.1 Introduction

Large multivariate data sets can prove difficult to understand and thus methods of summarising and extracting relevant information are necessary. Before the hypothesis deduction process it is useful to divide the data sets into similar groups so that a better interpretation of the collected data can be achieved. This division process is called classification. In this chapter three classification techniques that can be used to classify extracted chest sounds are discussed. The classification techniques are then compared and tested using the simulated data from Chapter 3.

5.2 Transients and Their Features

- Transients

To test the classification methods in this chapter, simulated heart sounds from Chapter 3 were used. The starting and ending positions of the simulated heart sound transients are known. These positions could be used to obtain the transients for the classification process but a complete system for extracting and classifying transients was needed. Therefore one of the transient extraction techniques from Chapter 4 was used to extract simulated heart sounds for the classification processes. Although the results in Chapter 4 indicated that average zero crossing rate algorithm was the best method to use for the extraction of transients the Signal's Envelope method was used
instead. This is because the Signal's envelope method gave a more consistent transformed signal for different types of recording. The other reason is that the added white noise in the generated signal is beneficial to the zero crossing algorithm. This is because added noise contained a large amount of zero crossings whereas some recorded data may not have such a high noise to signal ratio, i.e. a low zero crossings rate.

Although the Signal's Envelope algorithm was an easy method to use, it was observed in Chapter 4 that the method has the following disadvantages:

1. The calculated starting positions are too early. Extracted transients may therefore contain unwanted additional data.

2. The ending positions obtained are too early, thus the end part of the transients are omitted.

From these disadvantages it was decided to process the signal's envelope further before calculating the transients' starting and ending positions. These processes are shown in the block diagram in Figure 5-1

![Figure 5-1](image)

**Figure 5-1** Block diagram showing the processing of a signal's envelope for the extraction of starting and ending positions.

Using these processed signal's envelopes the extraction algorithm was able to detect 71% of starting positions that are within the set criteria defined in Chapter 4, and 71% of the ending positions.
Features

One of the approaches for the classification process is the representation of the chest sounds as a set of features. These features could be the number of peaks or troughs, transient's length, maximum positive or negative gradient, position of transients, etc. These features are arranged to form an N-dimensional vector with each dimension representing a particular feature. Assuming that there are M chest sound transients, this would form a matrix of size MxN, Figure 5-2. This matrix is then used by the classification algorithms to group similar chest sounds.

\[
\begin{bmatrix}
1 & 2 & 3 & \cdots & N \\
1 & F_{11} & F_{12} & F_{13} & \cdots & F_{1N} \\
2 & F_{21} & F_{22} & F_{23} & \cdots & F_{2N} \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
M & F_{M1} & F_{M2} & F_{M3} & \cdots & F_{MN}
\end{bmatrix}
\]

Figure 5-2 A matrix of transients' features, M transients by N features.

For the grouping of similar heart sounds the following features were chosen for the classification process:

1. Duration of transient.
2. Separation between last and present transient.
3. Separation between next and present transient.
4. Positive area.
5. Negative area.

With the exception of feature 2 and 3 most of the above features can be easily calculated. The problems with feature 2 and 3 are that the separations for the first and last transients, respectively, are impossible to calculate. These transients could be left out of the feature matrix but they may contain important information. Therefore the separation between the first transient of the recorded data and its previous is set to zero,
and the separation between the last transient of the recorded data and its next is also set to zero.

5.3 Nearest Neighbour Clustering

5.3.1 Methods

For this method of classification there are two stages in the grouping of similar transients described by the vectors in the feature matrix, Figure 5-2. The first stage is the evaluation of the network distance; the second is obtaining clusters of similar transients using "network distance values".

The network distance (ND) between two vectors is defined as the Euclidean distance (ED) divided by the smaller of the two vector magnitudes. Assume \( \overrightarrow{x} \) and \( \overrightarrow{y} \) are two vectors of N-dimensions, the Euclidean distance between these vectors is defined as

\[
ED_\alpha = \sqrt{\sum_{n=1}^{N} (x_n - y_n)^2}
\]

and hence the network distance

\[
ND_\alpha = \frac{ED_\alpha}{vector\_length}
\]

For example, consider the feature matrix in Figure 5-2. The first transient's features are selected, i.e. the first row equals \( \overrightarrow{x} \) in equation 5-1, and the ED between it and each of the other transient's features are calculated. After the vector with minimum ED, \( \overrightarrow{y} \), has been obtained, ND is evaluated using Equation 5-2. The vector \( \overrightarrow{x} \) is removed from the feature matrix and the vector \( \overrightarrow{y} \) becomes \( \overrightarrow{x} \). The procedure is repeated until no further transients are found.

The next stage of this classification algorithm is to obtain clusters of similar transients by using the ND values. The analysis of the ND values revealed similar
transients have small ND values and groups of dissimilar transients are separated by large ND values.

For example, Figure 5-3A shows two distinctive groups of four points. If their x-y co-ordinates are used as features their NDs can be calculated. When ND\textsubscript{xy} is plotted against point numbers (x, y), Figure 5-3B, it can be observed that the two groups are separated by a sudden change in ND. Hence groups of similar points can be obtained using the ND values.

![Figure 5-3 Two clusters and their network distance](image)

Note that axes on graph A and the vertical axis on graph B are on relative scales. The first point on graph B represents the network distance between points 8 and 1.

One of the problems with using features to represent a chest sound (or most other signals) is that the features may be correlated with one another. These correlated features may limit the information to the classification process thus producing a less accurate result. *Diagonal Factor Analysis* (Loudon, 1991) is used to try to resolve this problem.

The method forms uncorrelated factors from a larger number of correlated chest sound features. The formation of the uncorrelated factors is achieved by multiplying each feature by a weighting (Loudon, 1991). These uncorrelated factors are
then used to obtain similar groups of chest sounds by using the Nearest Neighbour
Clustering algorithm described.

5.3.2 Results and Discussion

From the simulated signals (normal heart sound, splitting of the first heart
sound, splitting of the second heart sound, third heart sound and fourth heart sound)
heart sound transients were extracted, their features were calculated, labelled (from 1 to
M heart sound transients) and arranged in a matrix of M by N. The feature matrix was
factor analysed and similar transients were grouped using the nearest neighbour
clustering algorithm. Figure 5-4A-E showed transients' numbers for the five heart sounds
and their network distances calculated from the clustering procedure. The calculated
network distances were normalised to unity so that a threshold value of 0.2 was chosen
to separate groups of similar transients. The figures also showed the groups obtained by
the clustering algorithm.

The labellings of the heart sound transients are as follows:

- Normal, splitting of the first and splitting of the second heart sound transients
  are labelled such that the first heart sounds are even numbers, and odd
  numbers for the second heart sound transients.

- The third and fourth heart sound transients are labelled in multiples of three,
  ie. 1=first heart sound, 2=second heart sound, 3=third or fourth heart
  sound.

It can be observed in Figure 5-4A-E that the clustering algorithm has
classified the given features into a number of groups. For most simulated heart sound
transients, except for splitting of the first heart sound data, the correct groups were
obtained but the algorithm did not classify the first and last transients of the five
simulated heart sounds. This is because one of the features belonging to the first and last
transients is equal to zero and hence the network distances between them and the rest of
the transients are large. The clustering algorithm failed to classify the data given for the splitting of the first heart sound but it was able to sort out the clusters of heart sound transients. This is because the calculated network distance between the first and second heart sound was too small for the algorithm to differentiate between the two clusters of heart sound transients, Figure 5-4B. This problem indicated that this clustering algorithm required feature values that are approximately the same for similar transients in order for it to correctly classify similar groups of transients.

Figure 5-4 Graphs of network distances Vs the transients' number
Figure 5-4 (cont.) Graphs of network distances Vs the transients' number.
5.4 Fuzzy logic

5.4.1 Methods

In 1965 Zadeh presented a paper on fuzzy set theory from which fuzzy logic emerged. The classical logic premise has two extremes: either completely false (zero) or completely true (one). However, in the presented fuzzy set theory a premise can have a range of truth from zero to one. This would allow the premise to be partially true and partially false. This introduction of uncertainty simulates the imprecise modes of reasoning that take place when human beings attempt to make a rational decision in an imprecise environment. With this property fuzzy logic can be seen as a powerful and robust technique for problem solving.

A partition of a set of chest sounds into a certain number of groups has a property that each sound belongs to one and only one group. This statement can be an over simplification of the classification process. This is because there are some transients which definitely belong to a certain group, but some whose membership is less evident. One approach that can be used for a more accurate summary of the data set is fuzzy clustering.

There are two ways in which fuzzy logic can be used to classify chest sounds. The first is to partition the similar groups before using the nearest neighbour algorithm. The result of the partitioning process must be such that it would give the feature variables equal importance during the nearest neighbour classification process. The second is to study a large amount of data and then write precise rules and membership functions such that the fuzzy system would correctly identify a given chest sound. Figure 5-5 illustrates the flow of data through a fuzzy system. The fuzzy system consists of three stages. Each stage uses predefined input membership functions (IMF), rule base and output weightings (OW) to derive the output of the system. The first stage evaluates one or more degrees of membership (DOM) from the input variables. The degrees of membership are then used by the rule evaluation process to form a certainty
factor for each set of rules according to the calculated degree of membership. Lastly, these certainty factors are used together with the OW to evaluate the output.

Figure 5-5 Data flow through a fuzzy system. Modified from Viot, 1993.

**Membership Evaluation**

This process is responsible for calculating the degree of membership of an input in one or more quantitative sets called fuzzy sets. Figure 5-6 shows an example of some fuzzy sets for chest sound duration. These fuzzy sets can be represented as simple shapes such as triangles and trapezoids, or complicated shapes such as $S$ and $\Pi$-functions (Wang, 1991).
There are two main points to keep in mind when designing fuzzy sets. The first point is the overlapping of set boundaries. This would ensure the smooth operation of the system. The second is the number of membership functions used to describe the input range. This feature indicates the accuracy, responsiveness, stability, etc. of the system. For simplicity the trapezoidal and triangular membership functions are used.

Assume fuzzy sets A, B & C and an input of x, Figure 5-7. Membership functions such as trapezia and triangles can be described by a set of numbers. These values include the zero points, the maximum points, the centroid, the gradients, etc. If the membership function is a continuous function then a number of variables are used to define its appearance and shape. Since the proposed membership functions are composed of triangles and trapezoidals they can be described by four values. These values are the zero points (B1 and B2) and the slopes (BS1 and BS2).
To calculate the DOM of an input, for triangular and trapezoidal fuzzy set, the following procedure is adopted (Viot, 1993):

1. Obtain the membership function(s) for the input. That is to calculate the difference between $x$ & B1 and B2 & $x$. If either of the results is zero or less than zero the input is not a member of that function.

2. Using the variables that describe the membership function obtained from step 1 calculate the degree of membership of the input.

The outputs from the Membership Evaluation process are represented in linguistic, quantitative sets or fuzzy sets, and numerical format, DOM. This gives a greater expressive power and thus they can be easily interpreted. Due to these characteristics the membership function is readable, thus can be readily modified, and the rule base can be written and corrected with ease.

**Rule Evaluation**

The concept of a rule base in fuzzy logic plays an important role in many of its applications. This is because the rules determine the way the system operates, i.e. desired outputs, stability, etc. These rules have been classified in a number of ways (Zadeh, 1988). One basic class of rules are *categorical rules*, i.e. rules that do not contain fuzzy quantifiers. For the proposed system a more general class of rules known
as the dispositional rules are used. These rules may have one or more premises that may contain fuzzy quantifiers and have the form of:

\[
\text{if } \langle \text{antecedent}_1 \rangle \land \langle \text{antecedent}_2 \rangle \land \cdots \land \langle \text{antecedent}_N \rangle \\then \langle \text{result}_1 \rangle \land \langle \text{result}_2 \rangle \land \cdots \land \langle \text{result}_N \rangle
\]

where an antecedent \_N has the form:

\[
\text{Fuzzy\_set\_name(Membership\_Function, DOM, Allowable\_Min\_DOM, Allowable\_Max\_DOM)}
\]

and a result \_N has the form:

\[
\text{Name\_of\_output} = \text{Certainty\_value}
\]

The rules could be written such that they can classify heart sound transients or re-value the features such that they do not correlate with each other. Assume that the membership function for the duration of a heart sound transient consists of three fuzzy sets; short, medium and long. Below are three examples of rules that could be used as a rule base for a fuzzy system.

\textbf{Rule 1.1: } if \ DURATION('Short', DOM, 0.5, 1.0) = TRUE
\then
\begin{align*}
\text{NORMAL\_1st\_HEART\_SOUND} & = (0.0 + 0.1) = 0.1 \\
\text{NORMAL\_2nd\_HEART\_SOUND} & = (0.0 + 0.1) = 0.1 \\
\text{SPLITTING\_1st\_HEART\_SOUND} & = (0.0 + 0.0) = 0.0 \\
\text{SPLITTING\_2nd\_HEART\_SOUND} & = (0.0 + 0.0) = 0.0 \\
\text{THIRD\_HEART\_SOUND} & = (0.8 + 0.1) = 0.9 \\
\text{FOURTH\_HEART\_SOUND} & = (0.8 + 0.1) = 0.9
\end{align*}

\textbf{Rule 1.2: } if \ DURATION('Medium', DOM, 0.5, 1.0) = TRUE
\then
\begin{align*}
\text{NORMAL\_1st\_HEART\_SOUND} & = (0.7 + 0.1) = 0.8 \\
\text{NORMAL\_2nd\_HEART\_SOUND} & = (0.7 + 0.1) = 0.8 \\
\text{SPLITTING\_1st\_HEART\_SOUND} & = (0.0 + 0.1) = 0.1 \\
\text{SPLITTING\_2nd\_HEART\_SOUND} & = (0.0 + 0.1) = 0.1 \\
\text{THIRD\_HEART\_SOUND} & = (0.0 + 0.1) = 0.1 \\
\text{FOURTH\_HEART\_SOUND} & = (0.0 + 0.1) = 0.1
\end{align*}

\textbf{Rule 1.3: } if \ DURATION('Long', DOM, 0.5, 1.0) = TRUE
\then
\text{NORMAL\_1st\_HEART\_SOUND} = (0.0 + 0.1) = 0.1
Assume that there are three types of heart sound transients; third or fourth heart sound, normal heart sounds and splitting of heart sounds. The short, medium and long membership functions are for the third or fourth heart sounds, normal heart sounds and splitting of the heart sounds respectively. The procedures for assigning certainty factors for a heart sound in each rule are as follows.

- The writer must choose a maximum certainty factor for the proposed rules, eg. 1 or 100.

- The next step is the allocation of a certain proportion of the certainty value for the sound types, ie. high values for sounds that are within the membership function and the remainder of the certainty value is for sounds that lie within the overlap of the membership functions.

- This remainder is distributed according to the heart sounds and their number of occurrence within the overlap.

Take the "medium" membership function rule as an example. The duration that lies within this membership function describes the normal heart sound transients. Therefore for ease of arithmetic a certainty factor of 0.7 is assigned to the NORMAL_1st/2nd_HEART_SOUND and 0.3 is to be distributed between the heart sounds which are within the overlap of the membership functions. The overlap for the "medium" membership function consisted of the third or fourth heart sounds (short membership function), normal heart sounds (medium membership function) and splitting of the heart sounds (long membership function). This would result in
a certainty value of 0.1 for each type of heart sound. The sum of the certainty values can be observed in the written rule.

Output Evaluation

After the completion of the rule evaluation process using the result of the membership evaluation process, further processing is needed to unravel the meaning of these outputs and to resolve conflicting and meaningless values. For example, the rules for heart sound duration may conclude that NORMAL_1st_HEART_SOUND = 0.9 but the rule for positive area in a heart sound transient concluded that NORMAL_1st_HEART_SOUND = 0.2. Resolving this conflict in the output evaluation process can be just a simple weighted average equation, (Viot, 1993). If the rules for duration are considered to be more important than those for positive area, the weighting for duration could be 60 and the weighting for positive area could be 20. Evaluating this sum would give 

\[
\frac{(0.9 \times 60) + (0.2 \times 20)}{0.9 + 0.2} = 0.9 + 0.2
\]

~

certainty for the transient being a normal first heart sound. For the fuzzy system used for classifying simulated heart sounds the output evaluation procedure is as follows:

\[
(R_1 \_Weight \times R_1 \_cf) + (R_2 \_Weight \times R_2 \_cf) + \ldots + (R_n \_Weight \times R_n \_cf)
\]

where \(R_1 \_Weight\) is the weighting for rule 1 and \(R_1 \_cf\) is the certainty factor obtained for rule 1. The sum of the multiplicands were not divided by the sum of the obtained certainty factors because the certainty factors were designated such that for any one rule the sum of certainty factors equals unity.

5.4.2 Results and Discussions

The features that were used in section 5-3 were used to test the fuzzy system for classifying heart sounds. These features were studied and the following fuzzy sets were designed.
A - Spurious noise or crackle.  D - Normal 1st heart sound.
B - 3rd or 4th heart sound.  E - Splitting of 1st or 2nd heart sound.
C - Normal 2nd heart sound.  F - Lung sound, wheeze.

Figure 5-8 Fuzzy sets for durations of heart sound transients.

A - First transient, any heart sound.  E - Normal 1st to 3rd heart sound.
B - Normal 1st to 4th heart sound.  F - 4th to normal 2nd heart sound.
C - 3rd to normal 2nd heart sound.  G - Normal 1st to 2nd heart sound.
D - Normal 2nd to normal 1st heart sound.

Figure 5-9 Fuzzy sets for separations between current and previous heart sound transients.
A - Any heart sounds.              E - 3rd to normal 1st heart sound.
B - 4th to normal 1st heart sound. F - Normal 2nd to 4th heart sound.
C - Normal 2nd to 3rd heart sound. G - Normal 2nd to 1st heart sound.
D - Normal 1st to 2nd heart sound.

Figure 5-10 Fuzzy sets for separations between current and next heart sound transients.

A - Spurious noise or crackles. C - Normal 1st or 2nd heart sound.
B - 3rd or 4th heart sound. D - Splitting of 1st or 2nd heart sound.

Figure 5-11 Fuzzy sets for positive areas of heart sound transients.
A - Spurious noise or crackles.  
B - 3rd or 4th heart sound.  
C - Normal 1st or 2nd heart sound.  
D - Splitting of 1st or 2nd heart sound.

Figure 5-12 Fuzzy sets for negative areas of heart sound transients.

From the designed fuzzy sets in Figure 5-8 to Figure 5-12 rules were written for each of the membership functions in the fuzzy set. Consequently there are five sets of rules and the number of rules for each set depends upon the number of membership functions in that fuzzy set, i.e. there are six rules for the duration fuzzy set and hence there are twenty-eight rules in total. From these rules the system was able to infer the certainty factors for the various types of heart sounds. The certainty factors from the rule evaluation process were multiplied by a weighting of 20 for duration rule output, 30 for previous separation rule output, 30 for next separation rule output, 10 for positive area rule output and 10 for negative area rule output. These multiplicands were then summed and the largest result gave the type of heart sound the input belonged to.

The fuzzy inference engine was able to correctly classify the heart sound features that were used in section 5.3. An example of the fuzzy inference engine output can be observed in Figure 5-13. The heart sound simulator always generated a normal first heart sound transient first and rest of the heart sound transients followed. Therefore when the fuzzy inference engine displayed its results the user can observe that the simulated heart sound transients were correctly identified.
## 5.5 Fuzzy Min-Max Clustering

### 5.5.1 Methods

The fuzzy classification technique described above may be tedious due to the difficult task of designing membership functions and rules. What is needed is a fuzzy logic algorithm that can partition groups of feature vectors and adapt itself during the classification process. This requirement of self-teaching may bring us into the domain of neural networks. Artificial neural networks are a form of expert system that were originally designed with the thought of mimicking the human neural system (Wasserman, 1989). The network can learn from experience either by supervised or unsupervised training. Once trained, the network response can be, to a degree, insensitive to minor

---

### Table: Classification of Heart Sounds Using Fuzzy Logic

<table>
<thead>
<tr>
<th>Normal 1st</th>
<th>Normal 2nd</th>
<th>Splitting of 1st HS</th>
<th>Splitting of 2nd HS</th>
<th>Third HS</th>
<th>Fourth HS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>-</td>
<td>-</td>
<td>6</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>-</td>
<td>-</td>
<td>9</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>11</td>
<td>-</td>
<td>-</td>
<td>12</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>14</td>
<td>-</td>
<td>-</td>
<td>15</td>
<td>-</td>
</tr>
<tr>
<td>16</td>
<td>17</td>
<td>-</td>
<td>-</td>
<td>18</td>
<td>-</td>
</tr>
<tr>
<td>19</td>
<td>20</td>
<td>-</td>
<td>-</td>
<td>21</td>
<td>-</td>
</tr>
<tr>
<td>22</td>
<td>23</td>
<td>-</td>
<td>-</td>
<td>24</td>
<td>-</td>
</tr>
<tr>
<td>25</td>
<td>26</td>
<td>-</td>
<td>-</td>
<td>27</td>
<td>-</td>
</tr>
<tr>
<td>28</td>
<td>29</td>
<td>-</td>
<td>-</td>
<td>30</td>
<td>-</td>
</tr>
<tr>
<td>31</td>
<td>32</td>
<td>-</td>
<td>-</td>
<td>33</td>
<td>-</td>
</tr>
<tr>
<td>34</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Figure 5-13: An example of the fuzzy inference engine output for the simulated Third heart sound features. The heart sound data is simulated such that the first and second transient are the normal first and second heart sound respectively and third transient is the third heart sound. The simulated transients are then repeated for 10 seconds.
variations in its inputs. This ability of deriving answers from noisy inputs is a result of its structure learned from its experience and not from human intelligence or experience embedded in the form of rules and knowledge bases. It can be seen that there are similarities between fuzzy logic and neural networks and it is thought to be advantageous to combine the two systems.

The Fuzzy min-max neural network algorithm was introduced by Simpson, 1993. The advantages of this algorithm are that it can learn by itself and can utilise parallel processing. The disadvantage of the network is that the author must know in advance the number of clusters which are going to be formed so that the number of nodes in the neural network can be implemented. Another drawback, although not a disadvantage, is that the learning process of the algorithm is not a neural operation, since there is no analogy for the expansion-contraction process in the human neural system. Because of these criteria it was decided to implement the fuzzy min-max clustering technique instead of the neural network version.

The fuzzy min-max clustering algorithm requires two parameters for the clustering procedure. The first parameter is the sensitivity parameter and it regulates the degree of membership of an input. A large sensitivity parameter would result in steep slopes in the membership function and thus the boundaries of the group become more "crisp". Therefore by specifying a large sensitivity value the boundaries around groups of clusters are well defined. The second parameter is the allowable size of a group.

Figure 5-1 Example of two types of hyperboxes.
Essentially, the fuzzy min-max clustering algorithm perceives each given feature vector as a single hyperbox in the feature space. A hyperbox is a term used to describe a box that is "drawn" in n-dimensional space. Figure 5-1 illustrates two and three dimensional hyperboxes. These hyperboxes can be completely defined using their minimum and maximum points. The fuzzy min-max clustering algorithm consists of two operations; expansion and contraction. A hyperbox can be expanded to include similar transients' features and contracted to avoid overlapping of the clusters' boundaries. The following steps were used for the fuzzy min-max clustering process:

1. **Initialisation**: The transients' features are initialised such that each transient is represented as a single hyperbox.

2. **Expansion**: A hyperbox is selected and expanded to include its nearest neighbour, by calculating the degrees of membership of the other transients. This expansion process must satisfy the allowable hyperbox size, otherwise the expansion process for that hyperbox is halted. The size of a hyperbox is defined as the Euclidean distance between the minimum and maximum of the hyperbox, instead of the sum of the difference between the minimum and maximum of the hyperbox used by the author. This parameter ensures the hyperboxes do not over expand and include the other clusters.

3. **Overlap test**: Immediately after the expansion process an overlap test is performed to determine overlap between hyperboxes.

4. **Contraction**: If overlapping occurs the hyperboxes are contracted.

These steps are repeated until no more expansion could be made.

Using the fuzzy min-max clustering technique it was possible to group n-dimensional feature vectors by adjusting only two parameters.
5.5.2 Results and Discussions

<table>
<thead>
<tr>
<th>Normal heart sound</th>
<th>Splitting of 1st heart sound</th>
<th>Splitting of 2nd heart sound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>Group 2</td>
<td>Group 3</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>10</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>12</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>14</td>
<td>-</td>
</tr>
<tr>
<td>17</td>
<td>16</td>
<td>-</td>
</tr>
<tr>
<td>19</td>
<td>18</td>
<td>-</td>
</tr>
<tr>
<td>21</td>
<td>20</td>
<td>-</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Third heart sound</td>
<td>Fourth heart sound</td>
<td></td>
</tr>
<tr>
<td>Group 1</td>
<td>Group 2</td>
<td>Group 3</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>13</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>16</td>
<td>15</td>
<td>14</td>
</tr>
<tr>
<td>19</td>
<td>18</td>
<td>17</td>
</tr>
<tr>
<td>22</td>
<td>21</td>
<td>20</td>
</tr>
<tr>
<td>25</td>
<td>24</td>
<td>23</td>
</tr>
<tr>
<td>28</td>
<td>27</td>
<td>26</td>
</tr>
<tr>
<td>31</td>
<td>30</td>
<td>29</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 5-1 Groups of transients obtained by the fuzzy min-max clustering algorithm.

The fuzzy min-max clustering algorithm was tested using the features that were used in the previous two sections. The features were adjusted such that the values
lay between 0 and 1. After the features were adjusted the sensitivity value was set to 4 and the allowable Euclidean distance between the minimum and maximum of the hyperbox was set to 0.1. The results from the min-max clustering algorithm can be observed in Table 5-1. From the table it can be seen that the algorithm had mostly grouped similar heart sound transients together, i.e. the transients' numbers are in the correct groups (transient number 1 being the first heart sound). It was anticipated that the fuzzy min-max clustering algorithm could not classify the first and last heart sound transients. This is because of abnormality of the features for the first and last transients as discussed in section 5.2.

5.6 Discussion

The common aspect in the classification techniques described in this section is that they all use the transients' features for the classification process. There are a few problems associated with using features when classifying chest sound transients.

The first classification algorithm, nearest neighbour clustering, demonstrated the problem concerning the correlation between feature variables. This problem was overcome by factor analysing the feature variables (Loudon, 1991). The derived uncorrelated factors are then used by the nearest neighbour clustering algorithm to obtain groups of similar transients.

The other problem concerns the sensitivity of the feature variables to background noise; for example, the maximum positive or negative gradients. The variation in the feature variables could affect the result of the classification algorithm. This effect can be observed in the results section of nearest neighbour clustering algorithm. The features extracted for the classification process were calculated such that the separations between the current and previous, or between the current and next transient equal zero for the first and last transient respectively. This feature deviates
massively from the rest of the features and thus the nearest neighbour algorithm was not able to classify the first and last transients as one of the main groups.

Due to this problem fuzzy clustering is used. As described, fuzzy logic algorithms can cope with a certain degree of deviation from the inputs. The proposed fuzzy inference engine in section 5.4 was able to classify all the given transient's features into the appropriate heart sound. The advantage of the fuzzy inference engine is that it can classify the type of heart sound from the given features whereas the other classification procedures can only produce a list of transients that are similar to one another. One disadvantage of the proposed fuzzy inference system is that it requires a set of membership functions and rules. The membership functions are easy to write but before designing the membership functions the designer must know the characteristics of the signals involved, i.e. the designer must know the universe of discourse of the feature variables. This would mean studying a large number of chest sounds from different patients.

The design of membership functions and the rule-base must be updated when a new set of data occurs. The design process needs an expert for such an occasion and could be time consuming. Therefore the fuzzy min-max algorithm was suggested. The advantages of this algorithm were that there were only two parameters to adjust, compared to the fuzzy inference engine, and the extracted features do not need to be factorised as required by the nearest neighbour method. The algorithm was able to obtain the main clusters but it could not classify the first and last transient. The other disadvantages of this algorithm are that it can not identify the type of heart sound from the given features and the user can waste time trying to find the optimum hyperbox's size.
References


Chapter 6

Processing of Recorded Chest Sounds

6.1 Introduction

This chapter describes the procedures and results that were obtained by using recorded signals to substantiate the signal processing and classification techniques described in Chapter 4 and Chapter 5. The signal's envelope algorithm was chosen to extract the transients of interest and fuzzy logic was used as a classification method. This validation process of chest sounds would give a greater understanding of the chest sound's characteristics and accuracy of the system. The recorded heart sound's characteristics showed that they deviated from the ideal simulated heart sounds in Chapter 3. With some modifications to the extraction and classification algorithms the system was able to extract and classify the appropriate chest sounds.

6.2 Overall System

Figure 6-1 shows the overall system to process chest sounds. The system consists of three stages; the first stage is filtering of the data to get rid of DC-offset, the second stage identifies the chest sound type and finally the recorded data is processed depending on its type.
The chest sound was identified by obtaining the durations of the transients of interest. This process used a simple signal transformer, i.e. a low pass filtering algorithm [described in Chapter 4], to obtain a suitable transformed signal for the calculation of the extracted durations of the transients. It was shown in Chapter 4 that the low pass filter method was not a suitable signal transformer for the extraction of transients. Nevertheless, the algorithm was used because it was a simple and fast method for the transformation process. Due to the simplicity of the method the obtained durations may contain some unwanted transients and the transients' durations may be shorter than expected. Because the shortening of the transients is consistent it is still possible to use this technique for the identification process. Therefore these calculated durations were only used to determine the type of sound input. This was achieved by using the following observations for the extracted durations.

![Block diagram of the overall system for extracting and classifying chest sounds.](image)
• Extracted transients that contained 500ms long transients would indicate wheezes are present.

• The presence of crackles are indicated by a large number of transients that are shorter than 45ms.

• Heart sound transients have durations that are between 500ms and 45ms.

After the type of sound is determined the signal is processed accordingly.

The following sections describe the processes that are needed to extract the appropriate transients.

6.3 Heart Sounds

6.3.1 Methods

\[
\text{Signal} \quad \downarrow \\
\text{Signal's envelope} \quad \downarrow \\
\text{LP Filter} \quad \text{c.f.} = 40\text{Hz} \quad \downarrow \\
\text{Gradient} \quad \downarrow \\
\text{Extract transients} \quad \downarrow \\
\text{Extract Features} \quad \downarrow \\
\text{Classification}
\]

where c.f. is the cut off frequency

Figure 6-2 Block diagram for extracting and classifying heart sound transients.

The flow chart for processing heart sounds can be seen in Figure 6-2. It was shown in Chapter 4 that the "signal's envelope algorithm" was a simple method to calculate a transformed signal for the extraction of heart sound transients. This is because there are no parameters to adjust and the transformed signal is consistent from one signal to the next. The disadvantages of the signal's envelope algorithm are:-

• High frequency noise can occur in the signal's envelope.

• The signal's envelope highlights the transients of interest too early.
The ends of the transients of interest are sometimes not highlighted.

Due to these disadvantages it was decided to manipulate the signal's envelope in order to obtain a better transformed signal for the extraction of heart sound transients. From the study of the recorded heart sounds and their transformed signals it was found that a combination of low pass filtering and average gradient calculations, Equation 6.1, gave a satisfactory transformed signal as shown in Figure 6-2.

\[ y_n = \frac{(x_{n-1} - x_n) + (x_n - x_{n+1})}{2} = \frac{1}{2}(x_{n+1} - x_{n-1}) \]  \hspace{1cm} 6.1

where \( x \) is the filtered signal's envelope.

The reasons for using these combinations were that the filtering process would reduce high frequency noise in the signal's envelope, and the calculation of average gradients was necessary to improve the significance of the endings of the transients of interest. The sum of the filtered and average gradients of the signal's envelope was then used for the transients extraction process.

**Transient's Features**

From the extracted heart sound transients the features that were used in Chapter 5 were calculated for the classification process. These features include duration of the transient, separation between the last and present transient, separation between the next and present transient, positive area and negative area. It was discovered that the values of the features that were used for the classification of simulated heart sounds were not the same as the values for real heart sounds. This factor indicated that the simulated heart sounds were not realistic enough, i.e. the feature variables need adjusting. This resulted in the re-design of the fuzzy logic membership functions and rules. It was also noticed that recorded heart sound may differ between transients and also between patients. This may arise from the patients' state of health and the physiology of their chest, e.g. their heart rate, size of the heart or skin layers. Therefore a new set of features
was needed. The selection of these features must be made such that it will ensure robustness of the system between recordings and between patients. One feature that can be used to determine the type of heart sound (first, second, third or fourth heart sound) is the positions of the transients. These positions can vary between recordings and they can also vary within a particular phonogram. This effect was due to the patient's heart rate and therefore a ratio of separation between the transients was chosen as one of the feature for classification process, Equation 6.2.

\[
\frac{\text{separation between next and present transient}}{\text{separation between last and present transient}} \quad 6.2
\]

The separation ratio alone may be insufficient for the classification of sound type. This can be justified by observing human experiences when trying to classify a given phonogram. When presented with a phonogram the cardiologist does not look at one particular transient and try to classify it. The normal procedure is to look at the phonogram as a whole and determine the context of the transients in the phonogram. This knowledge is then used for the classification process. These methods of reasoning could be implemented using the fuzzy inference engine such that it processed all the feature vectors at the same time. The procedures for composing such a system could be too complicated and hence the fuzzy inference engine used in Chapter 4 was written to process one feature vector at a time. Nevertheless, these reasonings can be modified so that the system can use the context for the next and last transients as well as the present transient for the classification process. As a result the following features were chosen.

1. Last transient's separation ratio.
2. Present transient's separation ratio.
3. Next transient's separation ratio.

The above features were satisfactory for the classification of first, second, third or fourth heart sound, but they were not able to distinguish between the splitting of heart sounds and normal heart sounds. Hence more features were needed for the
classification of heart sound transients. An additional feature could be the area of a heart sound transient but this also varied because of the difference in durations (due to heart rate and heart physiology) and amplitudes (due to the variation in attenuation because of the chest physiology) of the transients. These problems were resolved by interpolating the extracted transients to the same length and then normalising them to unity. This resulted in the calculated areas of the transients being approximately the same. It was found that by multiplying the area by their number of peaks and troughs the newly calculated features for different groups of transients were better partitioned. This method of calculating the new features resulted in large values for transients with a large number of peaks, and vice versa. The number of peaks and troughs in heart sound transients could also be different between people. In order to improve the robustness of the classification system a ratio of “turning point-area” was calculated, Equation 6.3, and used as one of the features. The reason for calculating the ratio in such a way was that if the phonogram only contained two types of sound then the last and next transients should belong to the same group. The average ratio of “turning point-area” would compensate for transients with abnormal “turning point-area” values.

\[
\text{if first transient then} \frac{\text{Present transient's TPA}}{\text{Next transient's TPA}} \\
\text{else if last transient then} \frac{\text{Present transient's TPA}}{\text{Last transient's TPA}} \\
\text{else} \frac{1}{2} \left( \frac{\text{Present transient's TPA}}{\text{Last transient's TPA}} + \frac{\text{Present transient's TPA}}{\text{Next transient's TPA}} \right)
\]

end

where TPA is the “turning point-area” values

Equation 6.3
Again these ratios were not sufficient to distinguish the normal and splitting of heart sounds. Working on the same principles as the separation ratios this feature was expanded to include the last, the present and the next "turning point-area" ratios. As a result the following features were chosen:

1. Last separation ratio.
2. Present separation ratio.
3. Next separation ratio.
4. Last turning point-area ratio.
5. Present turning point-area ratio.
6. Next turning point-area ratio.

-o Classification of heart sound transients-

The membership functions for the fuzzy classification system were rewritten to accommodate the newly selected features. The rule-base was also rewritten such that there were more than one antecedents in the "if statements". The results of these changes can be seen in the next section.

6.3.2 Results and Discussion

The extraction and classification abilities of the heart sound transient system were tested using five recorded heart sounds. The first two recordings, Figure 6-3 and Figure 6-5, were from a patient that had a valve replacement operation. The recordings were made using the equipment described in Chapter 3. It can be seen from these phonocardiograms, that the patient's second heart sound is split and there are other sound transients that could be 3rd or 4th heart sounds. This is because when observing transients 3 and 4 in Figure 6-3A it can be deduced that these transients are the first and second heart sound respectively. The second heart sound appears to have a double peak and therefore indicate that this heart sound has a splitting characteristic. The shape of the waveforms for the 9th and 10th transients, Figure 6-3B, do not appear to be similar to the other transients and thus their origin is not clear.
The third, fourth and fifth phonocardiograms, Figure 6-7, Figure 6-9 and Figure 6-11, were also from a patient who had had a valve replacement operation. In these phonocardiograms it can be observed that the patient has only two types of heart sound transients; splitting of the first heart sound and normal second heart sound. Observing the phonocardiogram, Figure 6-7, as a whole it can be deduced that the 3rd and 4th transients are the first and second heart sound respectively. From the phonocardiogram it can be seen that the first heart sounds' amplitudes are much smaller than the second heart sounds. This may be due to the sounds generated by two of the heart's chambers are not synchronised, splitting the heart sound.

Figures 6-3, 6-5, 6-7, 6-9 and 6-11 show the phonocardiograms used to test the system and Figures 6-4, 6-6, 6-8, 6-10 and 6-12 show the results obtained from the heart sound classification system.
Figure 6.3: CL1 phonocardiogram showing extracted transients.
The heart sound classification system was able to classify the detected heart sound transients of the GL1 phonocardiogram. Most of the extracted transients were classified as normal first heart sound and splitting of the second heart sound. It can be observed that transient number 9 and 10 do not share any common characteristics with the other transients. Therefore they were classified as the 3rd and 4th heart sound.

Looking at the overall signal it can be noticed that the patient’s heart rate is irregular and this characteristic may indicate that the 9th and 10th heart sound could be a normal first heart sound and splitting of the second heart sound, respectively. One inaccuracy of the system was not able to classify transient numbers 8 and 11, but the system was able to decide the correct categories the transients belong to (note that the system put these transients into two categories). This may be because of the unexpected appearance of transient number 9 and 10.
Figure 6-5 GL2 phonocardiogram showing extracted transients.
The classification system was able to correctly identify most of the heart sound transients with the exception of transient number 5. From the transients encountered in the GL1 phonocardiogram, transient number 9 and 10 of the GL2 phonocardiogram could be mistaken for a 3rd and 4th heart sound. Looking closer at these transients it can be observed that their characteristics are similar to those of the others but with smaller amplitudes. The system classified the 6th transient as a third heart sound and this may confirm that the 9th and 10th transients of the GL1 phonocardiogram were 3rd and 4th heart sounds.
CLASSIFICATION OF HEART SOUNDS
USING FUZZY LOGIC

There are no 3rd or 4th heart sounds in the LL1, LL3 and LL4 phonocardiograms. The phonocardiograms only contained splitting of the first heart sound and normal second heart sound. It can be seen that the splittings of the heart sounds were barely noticeable but the system was able to extract the heart sound transients and classified them correctly. The results of the classification process can be observed in Figure 6-8, 6-10 and 6-12.
Figure 6-9: LL3 phonocardiogram showing extracted transients.
### Classification of Heart Sounds Using Fuzzy Logic

File name >> LL3

<table>
<thead>
<tr>
<th>Normal 1st HS</th>
<th>Normal 2nd HS</th>
<th>Splitting of 1st HS</th>
<th>Splitting of 2nd HS</th>
<th>Third HS</th>
<th>Fourth HS</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>11</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>14</td>
<td>13</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>16</td>
<td>15</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Figure 6-10 Results of classification for phonocardiogram LL3.

(The numbers refer to the labels on the phonocardiogram)
Figure 6.11 LL4 phonocardiogram showing extracted transients.
### Classification of Heart Sounds Using Fuzzy Logic

<table>
<thead>
<tr>
<th>Normal 1st HS</th>
<th>Normal 2nd HS</th>
<th>Splitting of 1st HS</th>
<th>Splitting of 2nd HS</th>
<th>Third HS</th>
<th>Fourth HS</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>2</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-</td>
<td>4</td>
<td>3</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-</td>
<td>6</td>
<td>5</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-</td>
<td>8</td>
<td>7</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-</td>
<td>10</td>
<td>9</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-</td>
<td>12</td>
<td>11</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-</td>
<td>14</td>
<td>13</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-</td>
<td>16</td>
<td>15</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Figure 6-12 Results of classification for phonocardiogram LL4.

(The numbers refer to the labels on the phonocardiogram)

From the results of the extraction and classification process it can be seen that the system can be used to extract and classify heart sound transients from a given phonocardiogram. The phonocardiograms used were from two patients; GL's and LL's. The heart sound transients from LL's phonocardiograms were repetitive. This is because the patient's heart rate was constant and transients that belonged to the same type of heart sound were similar to each other for all of the LL's phonocardiograms. Due to these similar characteristics the system did not have problems when extracting and classifying these transients. The same system was used to extract and classify heart sound transients from GL's phonocardiograms. The system was able to extract and classify most of the heart sounds correctly. There were a few instances where the system was not able to decide whether the heart sound was normal or split. It can be seen from GL's phonocardiograms that virtually all the heart sound transients were unique. That is the
heart rate is not constant during a recording and transients belonging to the same heart sound differ in amplitude and sometimes in shape. The results from these tests demonstrated the robustness of the system as it was able to deal with variation in heart rate within the same recording or the variation in heart rate between recordings. Another feature that demonstrated the robustness of the classification system was that it could also deal with discrepancies in the shape of the heart sound transients.

The results from these phonocardiograms seem to indicate that the system can accommodate phonocardiograms from new patients. However, there are restrictions to the system. The restrictions and their reasons are as follows:

- **Extraction algorithm.** If one of the heart sound transients is too small the extraction algorithm may not be able to detect it. This would result in the wrong calculation of the feature vectors and in turn result in wrong classification of some of the heart sound transients.

- **Heart rate.** This is another important factor for the correct classification of the heart sound transients. As the heart rate increases the heart sound transients get closer and closer to each other. At a certain rate the separation ratios between the present & next transient and between the present & previous transient are similar to each other. When this happens the separation ratio features may not be able to partition the different groups clearly and this would result in the break down of the classification process.

- **Unknown transients.** When an unknown transient suddenly appears in the phonocardiogram the fuzzy inference engine cannot recognise the feature. This results in the wrong classification of the unknown transient and the one or two transients that surround it may also be wrongly classify. However, with some modifications to the membership functions and rules, the fuzzy inference engine could then correctly identify the unknown transient.
6.4 Lung Sounds

6.4.1 Methods

In Chapter 1 the two types of lung sounds; wheezes and crackles were discussed. This section describes the processes that were used to extract the lung sounds, but these processes did not attempt to group crackles that were generated from the same source, or wheezes that were generated from the same source. This is because wheezes are classified as low or high frequency, and crackles are classified as fine or coarse. Due to these imprecise descriptions of the lung sounds it was decided to extract the lung sounds and classify them as crackles or wheezes only.

![Block diagram for extracting and classifying lung sounds.](image)

where c.f. is the cut off frequency

Figure 6-13 Block diagram for extracting and classifying lung sounds.

The flow chart for extracting lung sounds can be observed in Figure 6-13. Before calculating the signal’s envelope the signal was filtered to get rid of high frequency noise. Depending on the type of sound detected the, signal’s envelope is filtered and then lung sound transients are extracted.
6.4.2 Results and Discussion

Figure 6-14 and 6-15 show ten seconds of recorded phonopneumograms. These phonopneumograms contained wheezes and by using the extraction algorithms described in Chapter 4 the extraction process was able to extract 4 wheezes each for the W1 and W2 phonopneumograms using their filtered signal's envelope. The extracted wheezes are labelled and the square brackets indicated the starting and ending positions of the wheezes. From the W2 phonopneumograms, Figure 6-15, it can be observed that a wheeze occurred between 2 and 2.6 seconds. This wheeze could be a continuation of the first extracted transient or a single wheeze. Nevertheless, the extraction algorithm did not recognise it. This may be because of the low signal to noise ratio of the W2 phonopneumogram. This factor made it difficult to distinguish between wheezes and background noise.

Figure 6-16 to Figure 6-19 show the recorded and filtered phonopneumograms that contained crackles. Some of the crackles were not distinguishable from the background noise. Therefore the recorded phonopneumograms were filtered to show where the crackles occurred. When these phonograms were put through the chest sound classification system they were recognised as phonograms that contained crackles. The crackles were extracted using the algorithms described in Chapter 4 and these extracted crackles are shown as labelled numbers in the filtered phonograms. The following numbers of crackles were extracted from the given phonograms.

- BR1 - 22 crackles were extracted.
- BR2 - 30 crackles were extracted.
- BL1 - 39 crackles were extracted.
- BL2 - 17 crackles were extracted.
It can be observed from the labelled phonopneumograms that some noise transients were extracted and some crackles were left out from the extraction process. This was because some of the crackle transients have small amplitudes due to the attenuation of the generated sounds by the airways and chest wall or the sources of the sounds were deep within the lung. Due to the crackles' small amplitudes they are lost within background noise and the signal's envelope can not highlight these small crackle transients. For these reasons the parameters for the extraction process were chosen so as to optimise the extraction of crackle transients.
Figure 6-14  W1 phonopneumogram showing wheezes.

Figure 6-15  W2 phonopneumogram showing wheezes.
Figure 6-16 BR1 phonopneumogram showing crackles; (A) recorded data, (B) filtered data.
Figure 6-17 BR2 phonopneumogram showing crackles; (A) recorded data, (B) filtered data.
6.5 Discussion

The results from the extraction and classification system indicated that it can analyse phonograms and produce the appropriate response. A low pass filter was used in the "automatic distinction algorithm" even though it was proved that the filtering process was not the most suitable method to use. Nevertheless, the "automatic distinction algorithm" was able to discriminate phonograms that contained wheezes, crackles or heart sound transients. The output from the "automatic distinction algorithm" was used to decide the processes that are needed to extract and classify the transients in the given phonogram.

Although there were only six types of heart sound transients tested from two different patients the system generally gave the correct class of the heart sound transients. It was also demonstrated that the fuzzy inference engine was flexible enough to cope with the new set of membership functions and rules for the classification of real heart sound transients. This flexibility suggested that if new heart sounds, such as snapping or clicking of valves and murmurs, were to be introduced, the membership functions and rules for the fuzzy classification system could be rewritten to accommodate the new sounds with ease. The classification process could be made more robust by the careful selection of additional features.

The other types of sounds that the chest sound classification system was able to classify and extract were wheezes and crackles. These sounds lay within the background noise frequency thus making the filtering process ineffective. Nevertheless, the system was able to extract most transients of interest with reasonable accuracy.

Wheezes are classified as low or high frequency lung sounds. Therefore classification of wheezes can be done in the frequency domain. Due to the imprecise description of wheezes it is difficult to know at what frequency ranges do physicians
classify a wheeze as a high pitch wheeze. These circumstances also apply to crackles. Due to these problems it was decided not to classify individual crackles or wheezes.
Chapter 7

Conclusions and Future Work

The literature review showed that auscultation is an established technique used to obtain information when diagnosing thoracic disorders. Due to the restriction of the human auditory system and the advancement in electronics the phonogram was developed. Using the phonogram it was possible for physicians and researchers to observe and gain a better understanding of the mechanisms that cause the various types of sounds emanating from the chest. The availability of the phonograph initiated the sudden interest in the classification of chest sounds and their associated diseases. The findings and conclusions of the investigations were mainly based on the researchers' experiences and observations. This classification by observations can be a laborious task thus a system to automate the classification of chest sound is needed.

Phonograms can be recorded by using a number of transducers but the literature review indicated that an electret condenser microphone is an acceptable transducer for most investigators. The equipment for recording chest sounds was built. It consisted of an amplifier and an electret condenser microphone enclosed in a shroud. This equipment was tested in a controlled environment.

The problem of noise contamination in the recording of chest sounds led to the development of a bench top acoustic simulator. This simulator was designed such that it
simulated crackles, lung sounds, in a controllable background noise environment. It was observed that the generated crackle transients had a higher frequency than crackles generated by the lungs. This is because the lung is a good medium for attenuating high frequency sounds and water is too fluid to simulate the lungs. A more viscous solution could be used to simulate the lungs. It was also demonstrated that, with the addition of different size tubes, different types of acoustic transients were successfully generated. Using this bench top simulator it was possible to test the recording equipment. This included the various arrangements of the microphone's shroud at different noise levels. The results indicated that due to the characteristics of sound waves it was impractical to completely eliminate background noise using the shroud. Nevertheless, a reasonably comprehensible phonogram could be obtained when the recordings took place in a quiet environment. This noise problem proved to be difficult to eliminate using conventional signal processing algorithms because chest sounds and background noise lie within the same frequency bands. For this reason the background noises had to be dealt with at the source or at the transducer level. Controlling background noise at the source is impossible so the next stage is the placement or the choice of transducers. The microphone that was used to record phonograms was an omni-directional microphone and thus the phonograms were contaminated with background noise. To reduce the background noise there are a number of transducers that can be used including: directional microphone, contact microphone or a hydrophone. Most of these transducers have a disadvantage of mismatching impedance between the chest wall and the microphone.

Another solution is to design a sealed container such that it has two flexible ends and an impedance matching fluid within, Figure 7-1. The container should have a hard and rigid outer layer to prevent background noise from entering the system, by reflection of sound waves. The inner layer of the container should also be lined with a sound absorption material to prevent sound reflection within the matching fluid and thus
prevent distortion of the chest sounds. Sounds from the chest can transmit into the fluid with little loss since acoustic impedance of the chest and the medium within the container are matched. When the transmitted chest sound reaches the end of container it encounters a flexible membrane. In order for the chest sound to convert into pressure waves the membrane is vibrated and hence sound. It is the vibration of the membrane that we want to measure. The vibration could be measured using an accelerometer, piezo electric films or strain gauges, but the mass of these transducers could alter the characteristic of the membrane's vibration. Therefore a none contact transducer could be use to measure this vibration.

The recorded sound contained both, the transients of interest and the unwanted background noise. There are a number of signal processing algorithms that could be used to extract transients of interest from the recorded signal. These algorithms have their own advantages and disadvantages. The main factor for choosing a signal processing algorithm for the extraction process is its ability to extract the wanted transients. To test the proposed signal processing algorithms a software heart sound
simulator was written. There are two heart sound simulators that have been developed by other authors. These simulators acquired heart sounds from patients and stored them. The heart sounds are played back on speakers when needed. The disadvantages with these simulators are that they require large storage space to store the recorded sounds and they are not suitable to use when testing the proposed transients' extraction algorithms. Because of these disadvantages a software heart sound simulator was designed such that it can generate heart sounds using a set of equations. Although the designed simulator has not been tested thoroughly it can be perceived that the generated signals can be used to test the proposed transients' extraction algorithms. This is because the beginnings and endings of each heart sound transient are known and the amount of noise can be controlled.

By testing the proposed signal processing algorithms using simulated heart sounds it was possible to determine the property and quality of the algorithms used to extract transients in a noisy environment. The extraction algorithm consisted of two stages; the transformation of the recorded signal and then it's use to extract transients of interest. The transients' extraction algorithm is a simple threshold selection process but this process depended upon the transformed signal obtained from the original signal. For this reason a number of signal transformers were evaluated; low pass filtering, signal's activity, signal's energy average zero crossing rate and signal's envelope. Most of the algorithms proposed by various authors were short-time analysis techniques, ie. they need a specific frame size for their computation. The advantage the first four signal transformers are that they are fast to calculate. However, the disadvantages of short-time analysis is that if a signal varies between recordings then the transients extraction process is not consistent and the algorithms are sensitive to noise. Because of these disadvantages the signal's envelope algorithm was proposed. Although the percentages of transients that lay within the given criteria were lower than some of the signal transformer algorithms this technique
was much easier to use than the other. This is because there were no parameters to adjust unlike to the other transform algorithms. Nevertheless, it was shown that the signal's envelope can be processed further, i.e. low pass filter or gradient calculations, such that the percentages of transients that lay within the given criteria increased.

The features from the extracted transients were calculated and then they were used for the classification process. If the transients are noisy then noise could be introduced in the feature vectors. It was demonstrated by the "nearest neighbour clustering algorithm" that the wrong classification can be the result of noise introduced in the feature vectors, even though diagonal factor analysis was used to uncorrelate the features. The wrong classification by the nearest neighbour clustering algorithm was due to the numerical values of the feature vectors. Therefore a fuzzy logic inference engine was used to classify similar transients. The advantage of fuzzy logic is that it describes the numerical feature vectors as a set of linguistic variables. This would make the linguistic variable not susceptible to noise. The fuzzy classification methods used were successful in determining the different type of heart sound transients.

It was discussed that if a new transient occurs the fuzzy logic system, i.e. its membership functions and rules, needed to be adjusted to accommodate for the newly detected transient. A set of rules was established for designing membership functions and rules for a fuzzy system. These rules are as follows.

- Gather the similar transient's features together.
- Select one feature and design membership functions such that the numerical values of similar transients lie within one or more membership functions. The membership function(s) that contains specific transients is noted for the designing of rules at a later stage. Repeat this procedure for all the features.
Rules consist of two sets of information. The first is the antecedents and the second is the output statements. The antecedents must consist of a membership function or a combination of membership functions for a specific output. The output statement should contain the wanted output or a number of outputs described by the antecedents. These outputs are in the form of certainty factors such that the most likely answer has the largest certain value.

From the above set of rules it is thought that an algorithm could be written to automate the designing of rules and membership functions for a fuzzy system. This algorithm could make the fuzzy classification method adaptive and self taught.

The output from the fuzzy classification system is in a form of a linguistic variable. This variable can be used by an expert system such as MIKE (Micro Interpreter for Knowledge Engineering) to diagnose thoracic disorders. The combination of transient's extraction and classification algorithms with an expert system may be a useful tool for diagnosing or studying thoracic disorders.

The other classification technique proposed by Simpson (1993) was the fuzzy min-max clustering algorithm. The algorithm arranged the feature vectors in an n-dimensional space and hyperboxes are introduced such that similar feature vectors lay within the same hyperbox. The author proposed that the size of the hyperboxes, or groups of similar transients, are controlled by the sum of the difference between the minimum and maximum of the hyperboxes. This value is thought to be meaningless and the Euclidean distance between the minimum and maximum of the hyperboxes was used instead.

The fuzzy min-max clustering algorithm was able to classify some of the given transients' features. This is due to the abnormality of the transients' features. The problem was anticipated since this classification process is somewhat similar to the nearest neighbour classification algorithm. That is they both used the feature space.
Another method that could be used to classify chest sounds is the neural network. The disadvantages of this technique are that a large number chest sounds are needed to train the network, there may be long training time and it is not possible for the network or developer to explain how the network derived the answers.

Using the signal's envelope as transients' extraction algorithm and fuzzy logic as a classification method a chest sound analyser was assembled. Signals from patients with chest disorders were used to test the chest sound analyser system. The system was able to recognise the three different types of chest sounds; crackles, wheezes and heart sounds. After the recognition process the sound was then processed according to its type. The system was able to extract the crackles and wheezes transients but it did not attempt to classify them due to the imprecise description of the sounds in numerous texts. When the system was used to process heart sound transients it was able to extract the appropriate heart sounds and classify them accordingly (normal first or second heart sound, splitting of the first or second heart sound, third heart sound and fourth heart sound). However, for some heart sounds the system was not able to decide whether the sound was normal or split. The results from the chest sound analyser test indicated that the system is robust enough to cope with variations between phonocardiograms due to anatomical and physiological differences between patients.

Since chest sounds and speech have similar properties it was thought that the techniques that were used for speech recognition could also be used the classification of chest sound. In general, a speech recognition system usually consists of two stages; the feature extractor and a classifier. The classifier is sub-divided such that the phonetic information of a signal is used for the classification process. The classification process is achieved by the usage of dynamic programming or by using hidden Markov models (Picone, 1990, and Silverman and Morgan, 1990). It was thought that continuous chest sound can be classified by dynamic programming or by using the hidden Markov model.
References


Appendix A

Recording equipment

Overall recording system

Cross-sectional drawing of the electret condenser microphone and its shroud

A.1
Flow chart for separating samples when using multi-channel recordings
Appendix B

Heart sound simulator flow chart

MAIN

WHILE option ≠ END

IF option = LS
Load / Save

IF option = NOR
Get heart sound variables

IF option = ADD
Get adventitious sound variables

IF option = GN
Generate heart sound
WHILE opt ≠ END

IF opt = LV
Load heart sound variables

IF opt = SV
Save heart sound variables

IF opt = SA
Save heart sound in an ASCII file

GET HEART SOUND VARIABLES

WHILE opt ≠ END

WHILE op = 1st Heart sound

IF op = M
Change MITRAL valve variables

IF op = T
Change TRICUSPID valve variables

WHILE op = 2nd Heart sound

IF op = A
Change AORTIC valve variables

IF op = P
Change PULMONARY valve variables
GET ADVENTITIOUS SOUND VARIABLES

WHILE opt ≠ END

IF opt = 3HS
Get third heart sound variables

IF opt = 4HS
Get fourth heart sound variables

IF opt = SC
Get snap/clicking valve variables

IF opt = MU
Get murmurs variables