



UNIVERSITEIT•STELLENBOSCH•UNIVERSITY
jou kennisvenoot • your knowledge partner

Automated Pediatric Cardiac Auscultation

by

Jacques Pinard de Vos



*Thesis presented at the University of Stellenbosch in
partial fulfilment of the requirements for the degree of*

Master of Science in Engineering

Study leader: Dr. Mike M. Blanckenberg

April 2005

Copyright © 2005 University of Stellenbosch
All rights reserved.



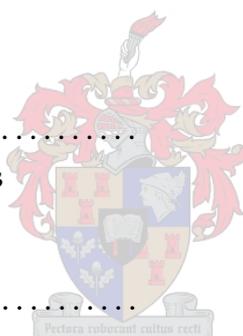
Declaration

I, the undersigned, hereby declare that the work contained in this thesis is my own original work and that I have not previously in its entirety or in part submitted it at any university for a degree.

Signature:

J.P. de Vos

Date:



Abstract

Most of the relevant and severe congenital cardiac malfunctions can be recognized in the neonatal period of a child's life. The delayed recognition of a congenital heart defect may have a serious impact on the long-term outcome of the affected child. Experienced cardiologists can usually evaluate heart murmurs with a high sensitivity and specificity, although non-specialists, with less clinical experience, may have more difficulty. Although primary care physicians frequently encounter children with heart murmurs most of these murmurs are innocent.

The aim of this project is to design an automated algorithm that can assist the primary care physician in screening and diagnosing pediatric patients with possible cardiac malfunctions. Although attempts have been made to automate screening by auscultation, no device is currently available to fulfill this function. Multiple indicators of pathology are nonetheless available from heart sounds and were elicited using several signal processing techniques. The three feature extraction algorithms (FEA's) developed respectively made use of a Direct Ratio technique, a Wavelet analysis technique and a Knowledge based neural network technique. Several implementations of each technique are evaluated to identify the best performer. To test the performance of the various algorithms, the clinical auscultation sounds and ECG-data of 163 patients, aged between 2 months and 16 years, were digitized.

Results presented show that the *De-noised Jack-Knife* neural network can classify 163 recordings with a sensitivity and specificity of 92 % and 92.9 % respectively. This study concludes that, in certain conditions, the developed automated auscultation algorithms show significant potential in their use as an alternative evaluation technique for the classification of heart sounds in normal (innocent) and pathological classes.

Opsomming

Die meeste van die relevante en ernstige aangebore hart siektes kan in die vroeë neonatale periode van 'n kind se lewe gediagnoseer word. Indien hierdie aangebore hart kondisies nie vroegtydig gediagnoseer word nie, kan dit 'n ernstige negatiewe uitwerking op die kind se langtermyn gesondheidstoestand hê. Ervare kardioloë is meestal in staat om patologiese hart kondisies met 'n hoë sensitiwiteit en spesifisiteit te identifiseer, terwyl nie-spesialiste, met minder kliniese ervaring, dit aansienlik moeiliker vind. Primêre geneesherre kom dikwels in kontak met kinders wat 'n geruis op die hart het. Baie van hierdie geruise is egter onskadelik.

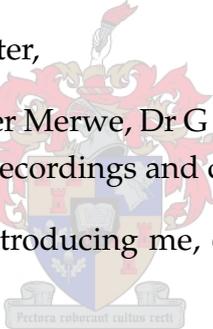
Die doel van hierdie projek is om 'n geoutomatiseerde algoritme te ontwerp, wat die primêre geneesheer kan bystaan in die ondersoek en diagnose van kinder pasiënte met moontlike hart kondisies. Ten spyte van pogings om beluistering ondersoek te outomatiseer, is geen toestel tans beskikbaar om hierdie funksie te vervul nie. Daar is egter verskeie aanwysers (tekens) van patologie teenwoordig in die hart klanke. Deur gebruik te maak van verskeie seinverwerkings tegnieke kan bg. aanwysers gebruik word om patologiese kondisies aan die lig te bring. Die drie eienskap onttrekkings algoritmes (EOA's) ontwikkel, maak onderskeidelik gebruik van - 'n Direkte Verhouding tegniek, 'n Wavelet (golfie) tegniek en 'n Kennis gebaseerde neurale netwerk tegniek. Verskeie variasies op elke tegniek is geëvalueer om die beste metode te identifiseer. Hart klanke en EKG-data van 163 pasiënte, ouderdom 2 maande tot 16 jaar, is ge-digitaliseer om die onderskeie metodes te evalueer.

Resultate, soos getoets op die 163 pasiënte, wys dat die *De-noised Jack-Knife* neurale netwerk die beste metode is om te gebruik, met 'n sensitiwiteit en spesifisiteit van 92 % en 92.9 % onderskeidelik. Die slotsom van hierdie studie is dat, in sekere omstandighede, kan die geoutomatiseerde algoritme dien as 'n alternatiewe evaluasie tegniek vir die klassifikasie van normale, onskuldige en patologiese hart klanke.

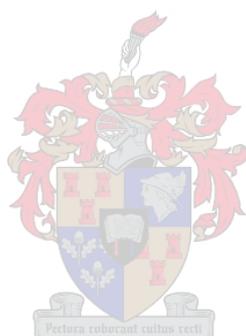
Acknowledgements

I would like to express my sincere gratitude to the following people and organization who have contributed to making this work possible:

- Dr M.M. Blanckenberg, of the University of Stellenbosch as my study leader and mentor,
- Tygerberg Children Hospital's pediatric cardiology clinic,
- Tulbagh Children Care Center,
- Dr J Hunter, Prof P.L. van der Merwe, Dr G Schoonbee and Dr A Phaff for their contribution with the data recordings and collection,
- Mr. Frank Myburgh, for introducing me, eight years ago, to the wonders of the human body.
- My wife and parents for their support, sacrifices, love and understanding,
- To our Heavenly Father, thanks for the opportunity and gifts.



Dedications



*Hierdie tesis word opgedra aan my wederhelfte,
Thia de Vos,
vir haar ondersteuning, geduld, moed inpraat en liefde.*

Contents

Declaration	ii
Abstract	iii
Opsomming	iv
Acknowledgements	v
Dedications	vi
Contents	vii
List of Figures	x
List of Tables	xvi
Nomenclature	xvii
Glossary	xx
1 Introduction	1
1.1 Context of the problem	2
1.2 Research gap	4
1.3 Objectives of this study	5
1.4 Outline of this study	6
2 Literature Review	7
2.1 Background	7
2.1.1 The Heart and the Circulatory System [1]	7
2.1.2 The fetal, transitional, and neonatal adaptations of the circulatory system	9
2.1.3 The cardiac cycle	10
2.1.4 ECG Morphology	13



2.1.5	Heart sounds	14
2.2	Heart murmurs - Innocent and pathological	15
2.2.1	Innocent murmurs	20
2.2.2	Conclusions regarding auscultation for pediatric murmur evaluation	22
2.3	Initial investigation and current theories	24
2.4	Murmur dynamics	25
2.5	Formulation of hypothesis	28
3	Methodology	29
3.1	Data collection	29
3.1.1	Subject population	29
3.1.2	Data acquisition	31
3.2	Database compilation	33
3.3	Pre-processing of heart sounds	34
3.3.1	Filtering and De-noising - ECG and Heart Sounds	34
3.3.2	Segmentation of recording into separate heart beats	43
3.3.3	Period filtering	44
3.4	Feature extraction and recognition	49
3.4.1	The Direct_Ratio method	49
3.4.2	Wavelet processing method	61
3.4.3	Artificial Knowledge Based Neural Networks	63
3.5	Statistical Analysis	73
3.5.1	Descriptive parametric statistics [2]	73
3.5.2	Sample distribution	74
3.5.3	Confidence Intervals and Hypothesis Testing	75
3.5.4	Sensitivity and specificity	77
4	Results and Findings	80
4.1	Data collection	80
4.2	Feature extraction algorithms (FEA's)	80
4.2.1	Direct Ratio results	81
4.2.2	Wavelet processing results	86
4.2.3	Artificial Knowledge Based Neural Network results	91
4.3	Simultaneous evaluation of all three methods developed	106
5	Conclusions, Limitations and Recommendations for Further Research	108
	Appendices	111

A	Information and informed consent document	112
B	Circuit schematics and board layout	115
C	Background on wavelet analysis	122
D	The Shapiro Wilk' test for normality	126
E	Matlab program code	129
E.1	Direct Ratio method	129
E.1.1	M-file used in the Direct Ratio algorithm	129
E.1.2	Code for Direct_Ratio.m	130
E.1.3	Code for Period_Filter.m	131
E.2	Wavelet analysis method	134
E.2.1	M-file used in the Wavelet analysis algorithm	134
E.2.2	Code for Wavelet.m	134
E.3	Neural network: Training data-set compilation	135
E.4	Neural network: Architecture, Initialization, Training, Testing, Validation and Performance testing	139
E.5	<i>Jack-Knife</i> neural network	144
E.5.1	<i>Jack-Knife</i> train data-set composition	144
E.5.2	<i>Jack-Knife</i> simulation and testing.(Calculation of validation recording classification)	148
	Bibliography	150

List of Figures

1.1	South African Public Health Care Statistics 2003	4
1.2	Levels and Types of Automated Systems	5
2.1	Sectional anatomy of the heart.(Courtesy of Benjamin Cummings, an im- print of Wesley Longman, Inc.)	8
2.2	Schematic diagram of the fetal circulation. The figures in the circles within the chambers and the vessels represent the oxygen saturation percent- ages for the respective parts. UV, umbilical vein; UA, umbilical artery; DV, ductus venosus; DA, ductus arteriosus; FO, foramen ovale; LV, left ventricle; LA, left atrium; RV, right ventricle; RA, right atrium; PA, pul- monary artery. <i>Draw, from a diagram illustrated by the Department of Anatomy, University of Bristol, by Thia de Vos</i>	11
2.3	The cardiac cycle. ECG section (top) and heart sound (bottom).	12
2.4	The normal heart sound (a) with three types of systolic murmurs (b, c, d). Sounds were de-noised with the fixed threshold wavelet de-noising tech- nique discussed in section 3.3.1.4, to assure the emphasis on the dynamic shape of the murmur	16
2.5	Mid-to-late systolic murmur (a) with the two types of diastolic murmurs (b, c) and an example of a continuous murmur (d). Sounds were de- noised with the fixed threshold wavelet de-noising technique discussed in section 3.3.1.4, to assure the emphasis on the dynamic shape of the murmur	17
2.6	Main auscultation areas for heart sounds	19
2.7	Levels of making a successful diagnostic differentiating, with inter-level discriminating factors (differentiators).	23
3.1	Methodology layout of the three feature extraction algorithms developed.	30

3.2	(a) Power Spectral density of stethoscope pickup in a noise-proof room, the 50 Hz mains harmonics is clearly visible. (b) & (c) show the noise differences between battery powered (red) and mains powered (blue) recordings.	36
3.3	(a) High-pass filter for ECG and (b) low pass filter for the heart sound data	37
3.4	Original ECG signal with unstable iso-electric line in blue and de-noised ECG signal in red	38
3.5	Original periodogram in green and filtered ($f_c = 650\text{Hz}$) periodogram in red	39
3.6	Daubechies Wavelet of order 5 and associated de- & recombination filter coefficients	40
3.7	This wavelet decomposition tree show approximation (V) and detail (O) spaces of 3-levels. With recombination it is shown that $s = V_3 + O_1 + O_2 + O_3$	40
3.8	Threshold values for different decomposition levels	42
3.9	(a) Normal period and (b) VSD period. Original signal is illustrated in green and the de-noised signal in red	42
3.10	Autocorrelation of the ECG waveform to calculate the heart cycle's duration	43
3.11	Flow diagram illustrating the segmentation of a heart sound into separate beats (periods). The values shown in Table 3.1 are used to classify the heart rate as normal or abnormal. The program code is available on the accompanied compact disc. The code is listed as Period_Calculator.m and all Period_Calculator's offspring files shown in Figure E.1 in Appendix E.1.	45
3.12	Result of the automatic heart cycle segmentation algorithm. All 31 cycles in this recording are extracted and copied to a ECG- and sound data matrix. The recording is represented by these two matrices in the following algorithms	46
3.13	Spectrogram of one normal heart sound period. A spectrogram is created by displaying all of the spectra computed from the heart sound period together. The lines visible on the spectrogram each represent 1 Hz along the frequency-axis, and one tenth of the total time along the time-axis. A contour plot is shown beneath the surface, on the xy-plane.	47
3.14	Mel-scale filter banks for 12 bins between 20-420 Hz	48
3.15	Flow diagram describing the automatic period filtering algorithm. The period filtering algorithm's code is listed in Appendix E.1	48
3.16	Heart cycle constituent components	50

3.17	Burke's second-order characteristic equations for the Q-T interval and the QRS complex for male and female patients.(Patient data courtesy of M.J. Burke and M. Nador, Department of Electronic Engineering, Trinity College, Dublin 2, Republic of Ireland)	51
3.18	Fitted 3 rd order equations for the Q-T interval and the QRS complex for male and female patients(Patient data courtesy of M.J. Burke and M. Nador, Department of Electronic Engineering, Trinity College, Dublin 2, Republic of Ireland)	53
3.19	Flow diagram describing the automatic segmentation algorithm, with reference to the inset figure (form Figure 3.16). The program code is available on the accompanied compact disc. The code is listed as Segmentation_Ratio.m and all Segmentation_Ratio's offspring files shown in Figure E.1in Appendix E.1	55
3.20	Output of the automatic segmentation algorithm for a normal heart sound	56
3.21	Output of the automatic segmentation algorithm for a pathological heart sound (VSD)	56
3.22	Flow diagram description of the Direct Ratio feature extraction method. The Direct Ratio algorithm's code is listed in Appendix E.1	57
3.23	Algorithm description for calculating new composition of constituent S1 (B). Program code is on the accompanied compact disc, listed as Constituent_S1.m	58
3.24	Energy content of heart cycle constituents calculated with Direct Ratio - Normal heart sound ($\frac{0}{6}$ rating)	59
3.25	Energy content of heart cycle constituents calculated with Direct Ratio - Holosystolic murmur (VSD $\frac{5}{6}$ rating)	59
3.26	Energy content of heart cycle constituents calculated with Direct Ratio - Early systolic murmur (VSD & CoArc $\frac{3}{6}$ rating)	60
3.27	Absolute values of wavelet coefficients for (a) a normal heart sound; and (b) a pathological VSD (3/6) heart sound. Colour bar indicate amplitude of absolute values	62
3.28	Algorithm flow for the wavelet analysis technique. Only for one patient (recording). The Wavelet analysis algorithm's code is listed in Appendix E.2.	63
3.29	Combined symbolic neural learning. Motivation for using neural networks for classification purposes. Framework was adopted from the knowledge-based neurocomputing flowchart presented by [3]	64

3.30	Neural network development and training methodology. The program code for the Network Architecture, Network Initialization, Training, Testing, Validation and Performance evaluation is listed in Appendix E.4 . . .	65
3.31	Algorithm flow for the construction of the training and training target data-set. Program code is listed in Appendix E.3, note that the construction of the validation matrix is done parallel in the program code.	66
3.32	Notation for describing a MLP, described with L layers , a d -dimensional input and c outputs. (Courtesy of Dr. Thomas Niesler, Stellenbosch University, South Africa [4])	68
3.33	Short notation for the 2-layer feed-forward backpropagation artificial neural network used as the classifier. The functions in both layers are sigmoid activation functions.	70
3.34	A two-sided test of the null hypothesis with $\alpha = 0.05$	76
4.1	The inset graph shows distribution of the 311 recordings studied. The primary pie chart show the distribution of the 86 pathological recordings which consist of the following conditions: ventricular septal defect (VSD), atrial septal defect(ASD), mitral incompetence or regurgitation (MI or MR), barlow syndrome (BS), aortic insufficiency (AI), aortic stenosis (AS), pulmonary stenosis (PS), pulmonary insufficiency (PI), Tetralogy of Fallot, peri-cardial friction rub (PFR) and tricuspid incompetence (TI)	81
4.2	Results of the Direct Ratio method. The inset legend show the data groups associated markers. The threshold line drawn at -22,07 dB will be discussed in a later subsection	82
4.3	(a), (b) and (c) show the difference between the normal distribution and the three data-sets	83
4.4	(a), (b) and (c) show the descriptive statistics for the Direct Ratio method, and (d), (e) and (f) the histogram distribution for the respective data-sets	84
4.5	Comparison between relative energy content for different scales tested. Only the highest energy constituent is plotted for each recording. A blue circle is a no disease case and a red cross is a pathological case	87
4.6	(a)&(b) illustrate the comparison between an actual normal distribution and the distribution of the no disease and pathological population respectively (c) & (d) illustrate the histogram distribution of the populations with their accompanied Shapiro Wilk W-test results.	88
4.7	(a), (b) and (c) show the descriptive statistics for the Wavelet analysis technique with scale = 64 and wavelet db4.	89

4.8	Receiver operating characteristics curve for classification of pathological or normal systolic heart murmur. Thresholds shifted from the minimum value in population to maximum value in the population. Data points are the corresponding sensitivity and specificity for each threshold, for different scales indicated.	90
4.9	Feed-forward neural network. The average of the input value to the last <i>sigmoid</i> function of the three validation periods per patient.	92
4.10	(a)& (b) illustrate the comparison between an actual normal distribution and the distribution of the two data-sets. (c) & (d) illustrate the histogram distribution of the data-sets with their accompanied Shapiro-Wilk W-test results	93
4.11	Descriptive statistics for the input data to the final <i>sigmoid</i> function (in the output layer) of the neural network	94
4.12	Output of the neural network. For the average of the three validation periods per patient	95
4.13	Input values to the last layer's function - average of three periods per patient with de-noised validation data input	96
4.14	Output of the neural network - average of three periods per patient with de-noised validation data input	97
4.15	<i>Jack-Knife</i> training method: Input value to the final <i>sigmoid</i> function. The average of six periods per recording (patient) is plotted	98
4.16	Distribution statistics for the six-period <i>Jack-Knife</i> training method	99
4.17	Descriptive statistics for the six-period <i>Jack-Knife</i> method	100
4.18	The <i>Jack-Knife</i> method's classification results. Trained and validated with six periods per patient. Plotted prediction is the average of the six periods. The horizontal line represents the example decision threshold	101
4.19	<i>Jack-Knife</i> de-noised training method: Input value to the final <i>sigmoid</i> function. Trained and validated with six periods per patient. Plotted prediction is the average of the six periods.	102
4.20	Distribution statistics for the <i>Jack-Knife</i> de-noised training method	103
4.21	<i>Jack-Knife</i> de-noised method's data descriptive statistics	103
4.22	<i>Jack-Knife</i> de-noised method's classification results. Trained and validated with 6 periods per patient. Plotted prediction is the average of the six periods	104
4.23	ROC curves for the four best performance methods. Legend show the indicators for the four methods.	107

B.1	Schematic diagram of the portable data acquisition unit and isolated USB or serial interface to PC	115
B.2	Schematic layout of the audio circuit. Input to circuit is a 20 - 20 000 Hz microphone pickup - implemented inside a acoustic stethoscope. A 8 th order Butterworth switch-capacitor low-pass filter ($F_c = 650\text{Hz}$) is used to filter the audio signal	116
B.3	Schematic diagram of 3-lead ECG board. A low-noise differential amplifier is used to obtain the voltage difference between the two primary electrodes. The third input is used as virtual ground. The signal is filtered with a 100 Hz LPF filter before normalized for the A/D circuitry.	117
B.4	Schematic diagram of digital acquisition board. The design consists of a 12-bit dual channel A/D converter; 2 Mb of on board flash memory for data storage; a micro processor ; an 4-channel optic isolator and a USB & serial connection. Dual power supplies are used to isolate the patient from the computer.	118
B.5	Printed circuit board layout for the audio circuit.	119
B.6	Printed circuit board layout for the of 3-lead ECG circuit	120
B.7	Printed circuit board layout for the digital acquisition board	121
C.1	Windowing regions of STFT and WT analyses	123
C.2	Wavelets to illustrate pseudo frequency	124
E.1	Flow diagram of M-files used in the Direct Ratio algorithm. Code for the Direct_Ratio.m and Period_Filter.m are listed in this Appendix. The rest of the files can be viewed on the accompanied compact disc	129
E.2	Flow diagram of M-files used in the Wavelet analysis algorithm. Code for Wavelet.m and Period_Filter.m are listed in this Appendix. The rest of the files can be viewed on the accompanied compact disc	134

List of Tables

3.1	Normal values of heart rates in pediatric patients. [wk - week; y=year; bpm = beats per minute]. Data from [5]	44
3.2	Direct Ratio calculation of mid systolic (F) constituent, for different percentage values of constituent S1	60
3.3	Pseudo frequencies computed with equation C.0.4 and 6 dB passband limits for the various scales tested. Coif2's 6 dB obtained form [6]	61
3.4	Possible patient groups.	78
4.1	Descriptive statistics for the Direct Ratio method. [SD = standard deviation and CI = confidence interval]	84
4.2	Direct Ratio method's sensitivity and specificity for different threshold values	85
4.3	Descriptive statistics for the Wavelet analysis technique with scale = 64 and wavelet db4	89
4.4	Descriptive statistics for three validation period Neural network method	94
4.5	Sensitivity and specificity for the neural network 3-period validation method	95
4.6	Sensitivity and specificity of the 3-period validation method neural network with de-noised training- and validation data input	97
4.7	Descriptive statistics for the six-period <i>Jack-Knife</i> neural network method	100
4.8	The <i>Jack-Knife</i> method's sensitivity and specificity for different training and validation periods per patient.	101
4.9	Descriptive statistics for de-noised <i>Jack Knife</i> neural network method	104
4.10	<i>Jack-Knife</i> de-noised method's sensitivity and specificity for two different threshold values.	104

Nomenclature

A/D Analogue-to-Digital

AI Aortic insufficiency

ANN Artificial neural network

ASD Atrial septal defect

AV Atrioventricular

bpm Beats per minute

CHD Congenital Heart Disease

CI Confidence interval

D Diastolic

dB decibels

DSP Digital signal processing

DT Dead time

DWT Discrete wavelet transform

ECG Electro Cardiogram

EMD Early to mid diastole

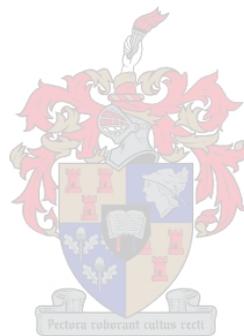
ES Early systole

FEA Feature extraction algorithms

Hz Hertz (oscillations per second)



IDWT	Inverse discrete wavelet transform
LD	Late diastolic
LPF	Low-pass filter
LS	Late systole
LUSB	Left upper sternal border
LV	Left ventricle
Mb	Mega (10^6)-bytes
MI	Mitral incompetence
MLP	Multi-layer perceptron
MR	Mitral regurgitation
MS	Mid systole
MSE	Mean square error
PC	Personal computer
PCG	Phonocardiogram
PI	Pulmonary insufficiency
PS	Pulmonary stenosis
ROC	Receiver operating characteristic
RV	Right ventricle
S	Systolic
S1	First heart sound
S2	Second heart sound
SD	Standard deviation
SNR	Signal-to-Noise Ratio



TI	Tricuspid incompetence
USB	Universal Serial Bus
VCR	Video Cassette Recorder
VSD	Ventricular septal defect
WS	Wide systole

Constants:

$$\pi = 3,141\,592\,653\,589\,793\,238\,462\,643\,383\,279\,5$$

$$e = 2,718\,281\,828\,459\,045\,235\,360\,287\,471\,352\,6$$

Symbols:

Re Reynolds number with regards to diameter

μ Mean

σ standard deviation

$\sigma_{\bar{x}}$ Standard error of the mean

p Probability

ψ Sensitivity

χ Specificity

W *Shapiro-Wilks'* W test

$s_{\bar{x}_1 - \bar{x}_2}$ Standard error of the difference between the two sample means

Z_{test} Test for hypothesis

n Number of recordings



Glossary

Artifact noise	Any man made noise.
Auscultation	To listen to the sounds made by the internal organs of the body for diagnostic purposes.
Crescendo	A gradual increase in strength or loudness
Decrescendo	With gradually diminishing force or loudness.
Differentiate	To notice or indicate differences between.
Habitus	The physique or body build.
Piezoelectric	The property of generating electric polarity in dielectric crystals subjected to mechanical stress.
Precordium	The part of the body comprising the epigastrium and anterior surface of the lower thorax.
Stenotic	A constriction or narrowing of a duct or passage.
Viscosity	The tendency of a fluid to resist flow.

Chapter 1

Introduction

Most of the relevant and severe congenital cardiac malfunctions can be recognized in the neonatal period of a child's life. Neonatal data collection gives an incidence of 8 significant congenital heart disease (CHD) out of 1000 live births [7]. With an additional 1 or 2 out of 1000 previously unknown cases presenting in school-aged children [8]. The incidence of acquired heart disease in this population is however low [8]. In contrast to the low occurrence of this disease, innocent (functional) ¹ heart murmurs are common in clinical practice and is present in at least one examination in 50 % to 90 % of children and 15 % to 44 % of young adults. Given the prevalence of innocent murmurs and the relatively low incidence of actual heart disorder, the primary physician may have difficulty differentiating which murmurs need specialist referral.

Primary physicians can readily take a history, examine the pulses, and measure the blood pressure, but sometimes lack confidence when differentiating between innocent and pathological murmurs. This, combined with the knowledge that delayed recognition of congenital heart defects may have a serious consequences, lead to the frequent unnecessary referral of patients to a pediatric cardiologist [9].

Studies show that 60 % to 80 % [8, 10] of murmurs referred for sub-specialist evaluation were found to be innocent murmurs. These statistics emphasize the need for an improvement in primary level auscultation and/or an additional screening technique. This chapter explores the current proposed solutions, their limitations and insufficiencies. The above mentioned leads to the formulation of a problem statement, and stated thereafter are the objectives of this thesis.

¹Also called functional, normal, vibratory or physiologic murmurs, this report will refer to innocent heart murmurs

1.1 Context of the problem

South Africa has a relatively youthful population with a third of the population being under 15 years of age [11]. The 1998 Demographic and Health Survey found that the infant mortality rate was 45 per 1000 live births for the preceding 10 years [12]. In 2000 congenital heart disease was the cause of death of 1238 children aged under 5 years in South Africa. According the South African Medical Research Council's report on child mortality CHD was ranked eight on the list of causes of death of children under 5 years, in South Africa for the year 2000. [13]

The analysis of biological sounds within the human body (auscultation), by the use of a stethoscope is a common practice of medical practitioners all over the world. Although auscultation is accepted as a sufficient method of diagnosing heart defects, abnormalities can easily go undetected due to the limitations in the ability of the human ear to distinguish defects from the sound of a heartbeat [14]. A pediatric patient with a heart rate of 150 bpm, has a computed interval of 0.24 seconds between the two heart sounds. To differentiate between an innocent and a pathological ejection systolic murmur in the above mentioned case, an experienced pediatric cardiologist is required. Cardiac murmurs are a common finding in children and represent the most frequent reason for referral to a pediatric cardiologist [5]. According to literature the majority of these patients can be adequately evaluated clinically, yet increasingly more extensive studies are being used in this assessment. There are many reasons for this practice, which include reduced confidence in auscultatory skills, the increased availability of diagnostic technology, the increasingly competitive nature of pediatric practice and increased medicolegal concerns [5]. Objective evidence suggests that proficiency in cardiac auscultation among physicians in training may be in decline [8, 10, 15, 16, 17], due to the availability of the above mentioned modalities.

The past several decades have seen the increased usage of the technological aspects of medicine for the making of diagnostic decisions. These methods, which include computer tomography, magnetic resonance imaging, echocardiography and cardiac catheterization have provided new modalities to assist physicians in providing quality medical care.

These techniques, via their emphasis in medical school and postgraduate training, have to a certain degree taken over as the central features of medical diagnosis at the expense of the techniques of physical examination and history taking [18]. Yet physical examination is indeed a fundamental skill in primary medicine and is of-

ten used as the tool with which to determine whether a referral to a specialist is necessary or not.

In any developing country, like South Africa, the number of available specialists are limited to such an extent, that any unnecessary referral should be minimized for the following reasons:

1. Specialists should be available for the patients needing them the most: Specialists are a very scarce and expensive resource that should be used only when required, see *Figure 1.1* for motivation;
2. For the financial benefit of the patient: *Figure 1.1* shows that the distribution of medical practitioners and specialists in South Africa are not in ratio with the regional demographic composition. It is however clear that the distribution of specialist are economically driven, leaving the poorer regions with a much larger *people-to-specialist* ratio.

Due to the statistics and the vast demographic composition of South Africa, people in poorer regions, typically living in rural areas, need to travel great distances to seek the opinion of a specialist, while the richer people typically living in urban areas, with a small *people-to-specialist* ratio only need to travel short distances to a specialist. An unnecessary referral can cost the patient a visit to a distant city or town, and the time spent at the cardiologist;

3. To minimize the anxiety of the patient and his/her family in the case of an unnecessary referral: Obtaining access (logistically and economically) to the necessary transport to-and-form the specialist can take time.

Possible solutions to the posed problem can be divided into three categories:

1. **Educational:** Better training of all primary general practitioners can lead to a decrease in the unnecessary referral rate. With reference to *Figure 1.2* the difference between a specialist and a general practitioner (in a specific field) is, however, the level of knowledge and skill required in a specific field to become a specialist. There are however a multitude of specialist fields in medicine resulting in the general practitioner having only a limited knowledge and skill in any specialist field, for example pediatric cardiology.
2. **Political:** Distribution of specialists in relation to population distribution. Given the current economic reality in South Africa, it is not feasible to attempt to let the *specialist-to-patient* ratio converge between provinces. Even if the political

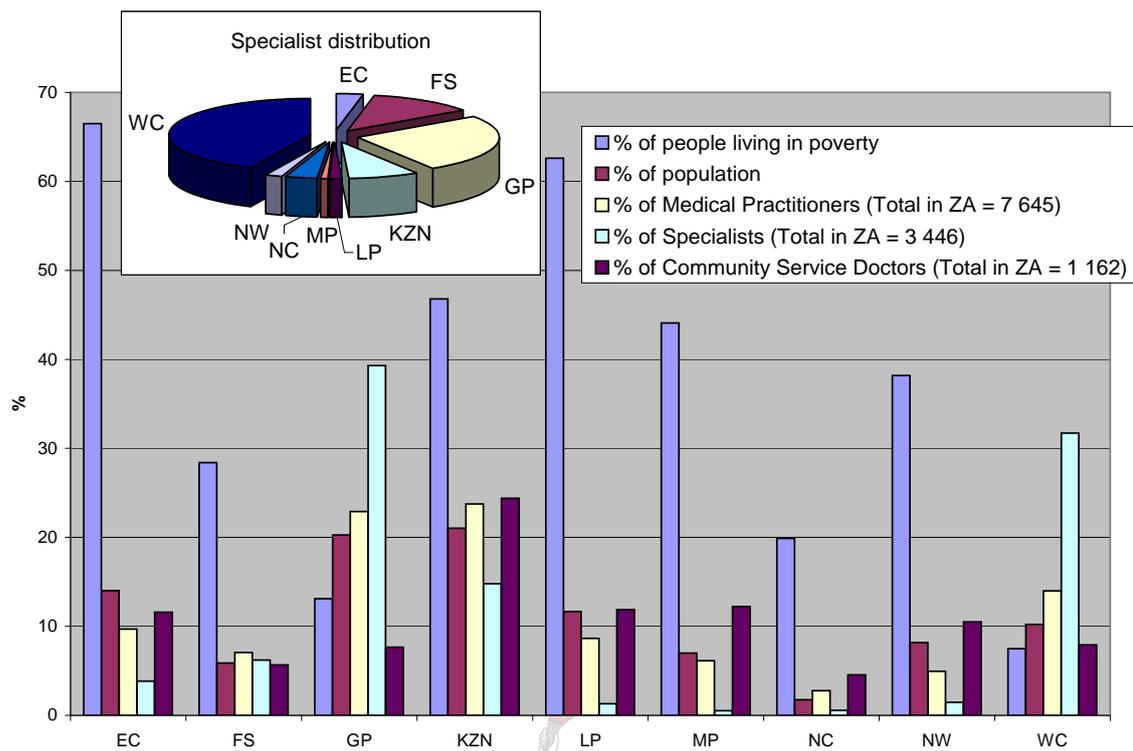


Figure 1.1: South African Public Health Care Statistics 2003. EC: Eastern Cape, FS: Free State, GP: Gauteng, KZN: KwaZulu-Natal, LP: Limpopo, MP: Mpumalanga, NC: Northern Cape, NW: North West, WC: Western Cape, ZA: South Africa (Data obtained from Statistics South Africa [19, 20])

will exists to distribute specialists more equally between provinces, it would most likely not materialize in the near future.

- 3. Research and Development:** Development of an alternative primary screening tool or method. *Figure 1.2* shows the relation between the different intellectual levels and computer levels. Developing a *Decision Support System* and/or an *Expert System* might contribute towards increasing confidence in auscultatory skills on primary healthcare level.

1.2 Research gap

There have been several studies investigating the possibility of developing an algorithm for automated diagnosis. According to literature efforts to date have been met with limited success. Only one of the studies investigated includes results obtained from an extensive clinical testing of algorithms. Most of the evaluated studies have generally focused on determining if a heart murmur exists, with a lack of em-

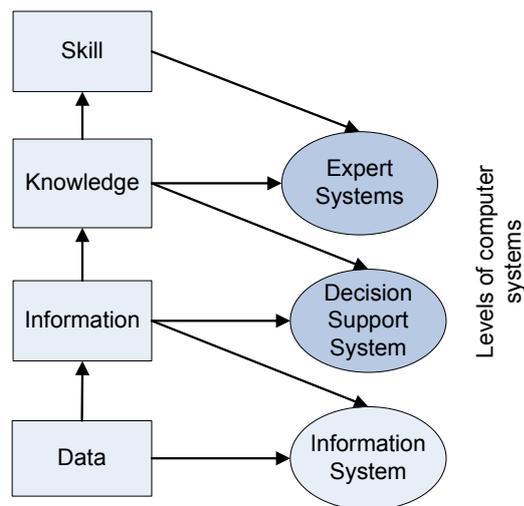


Figure 1.2: Levels and Types of Automated Systems

phasis on developing diagnostic algorithms differentiating between innocent versus pathological murmurs [?, 21, 22, 23, 24, 25].

There are, however, two parties that have published successful results. Reid Thompson from Johns Hopkins Hospital, America seems to lead the only project to have published successful results obtained from extensive clinical trails [8, 6]. This project is sponsored by the U.S. Army, resulting in a relatively strict control of information. The other research group, University of Colorado Health Science Center, examined the heart sounds using artificial neural networks, but with very limited success concerning the level of automation.

The practical implementation of an automatic primary screening device (a decision support system) for pediatric congenital heart diseases is still lacking.

1.3 Objectives of this study

The purpose of this study is to:

1. investigate various methods of extracting additional information from heart sounds;
2. look into the possibility of making a differential diagnosis;
3. design an implementable automated intelligent algorithm that can assist a primary care physician in decision making;

4. make an objective study as to whether the developed algorithm can be implemented as a primary screening device for pediatric congenital heart diseases. This, in turn, could lead to a reduction in the number of unnecessary referrals.

1.4 Outline of this study

Chapter 2 gives a background study on the cardiac anatomy and physiology and the most common congenital heart diseases, followed by a literary review on the research topic.

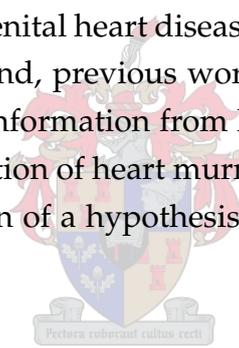
In Chapter 3 the various methodologies developed to reach the objectives of this study are discussed together with a detailed description on the statistics used to measure the performance of the various methods. Chapter 4 is an exposition of the results of the various feature extraction algorithms (FEA's) discussed in Chapter 3. The thesis is concluded with a general discussion and a summary of results, limitations and findings in Chapter 5. Recommendations for further research can also be found in Chapter 5.



Chapter 2

Literature Review

This chapter will lead a short introduction into the anatomy and physiology of the heart and the cardiovascular system, followed by a background on murmurs and the most common associated congenital heart diseases. With the structural features and composition of the heart in mind, previous work will be examined on the specific topic of extracting diagnostic information from heart sounds. Thereafter the possible explanations for the generation of heart murmurs is investigated. The chapter is concluded with the formulation of a hypothesis that will be the topic of the rest of the thesis.



2.1 Background

2.1.1 The Heart and the Circulatory System [1]

Blood vessels are subdivided into two circuits that both begin and end at the heart. The pulmonary circuit carries blood to-and-from the exchange surface of the lungs while the systemic circuit transports blood to and from the rest of the body. Arteries, or efferent vessels, carry blood away from the heart, while veins, or afferent vessels, return blood to the heart. Blood that returns to the heart in the systemic veins must complete the pulmonary circuit before re-entering the systemic arteries.

The heart consists of four muscle chambers, and of these chambers two are associated with each circuit. The right atrium receives blood from the systemic circuit, while the right ventricle discharges blood into the pulmonary (lung) circuit. The left atrium collects blood from the pulmonary circuit, while the left ventricle ejects

it into the systemic circuit. During heart beats, the two ventricles contract simultaneously to eject equal volumes of blood into the pulmonary and systemic circuits respectively.

Figure 2.1 illustrates the four internal chambers of the heart. The two atria are separated by the inter-atrial septum, while the two ventricles are divided by the inter-ventricular septum. Each atrium connects to the ventricle on the corresponding side through an atrioventricular (AV) valve. The composition of the valves ensure a one-way flow of blood from the atria into the ventricles. The right atrium receives blood from the systemic circuit via two large veins, the superior vena cava and the inferior vena cava. The superior vena cava delivers blood from the head, neck, upper limbs and chest. The inferior vena cava carries blood from the rest of the trunk, the viscera, and the lower limbs. The foramen ovale, an oval opening, permits blood flow between the two atria from the fifth week of embryonic development until birth. See section 2.1.2 for a systematic discussion on the fetal circulation.

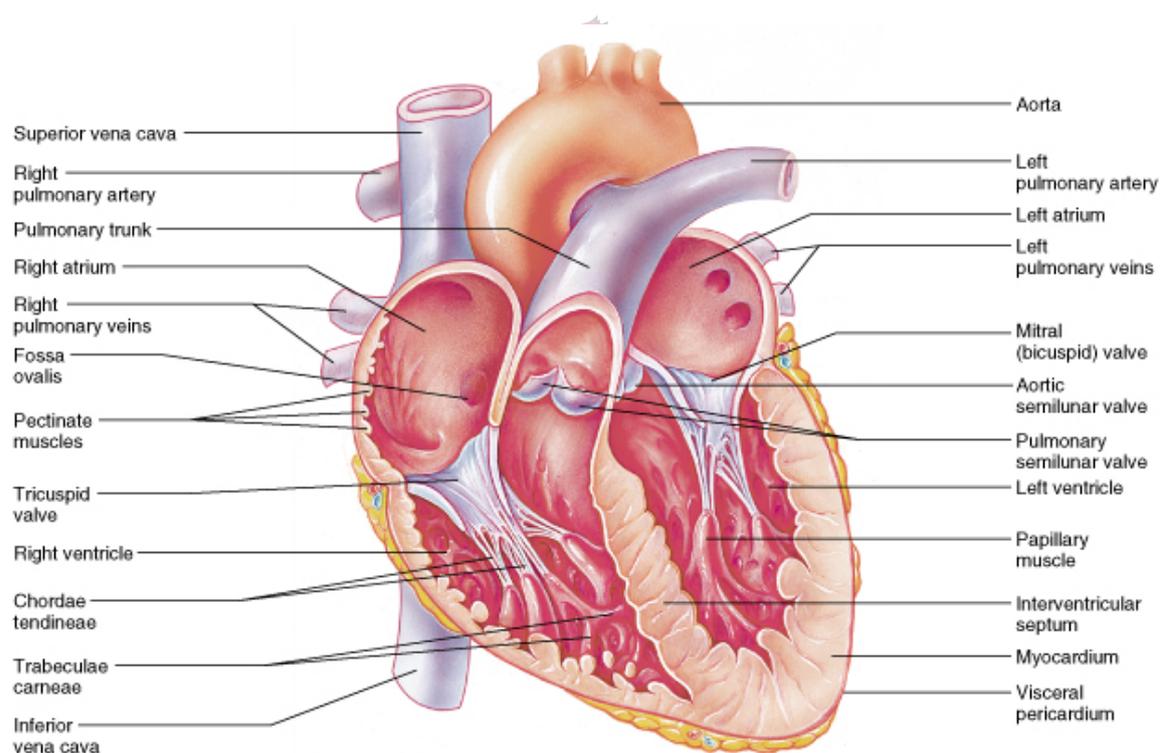


Figure 2.1: Sectional anatomy of the heart.(Courtesy of Benjamin Cummings, an imprint of Wesley Longman, Inc.)

Blood flows from the right atrium into the right ventricle through a broad opening bounded by three flaps.. These flaps, or cusps, are part of the right AV valve, also known as the tricuspid valve. Each cusp is braced by the chordae tendineae. These tendinous cords are connected to papillary muscles on the inner surface of the right

ventricle. By tensing the chordae tendineae, these muscles limit the movement of the cusps and ensure proper valve functioning.

Blood flows out of the right ventricle into the pulmonary trunk. This is the start of the pulmonary circuit. The pulmonary semilunar valve guards the entrance to this efferent trunk. Within the pulmonary trunk, blood flows into the left and right pulmonary arteries. These vessels branch out repeatedly in the lungs, supplying the capillaries where gas exchange occurs. From these respiratory capillaries, oxygenated blood collects into the left and right pulmonary veins, which deliver it to the left atrium. Similar to the right atrium, the left atrium has an external auricle and a valve, the left AV valve, or bicuspid valve. Clinicians often use the term mitral (a bishop's hat) when referring to this valve. A pair of papillary muscles braces the chordae tendineae that inserts into the mitral valve. Blood flowing out of the left ventricle passes through the aortic semilunar valve and into the aorta. This is the start of the systemic circuit.

2.1.2 The fetal, transitional, and neonatal adaptations of the circulatory system

"An understanding of the fetal, transitional, and neonatal adaptations of the circulation is important in the evaluation of the pediatric cardiovascular system because most organic heart diseases is evident in association with the circulatory changes occurring at birth." [5] The possible type of cardiac malfunction and the level of urgency that it must be act upon can be indicated by the age of the patient at the recognition of the murmur. Figure 2.2 show a schematic diagram of the fetal circulation.

In the fetus, oxygen rich blood is received from the placenta, via the umbilical vein and the ductus venosus. From the caudal vena cava, indicated on Figure 2.2, the blood flows to the right atrium from where it is directed across the foramen ovale to enter the left atrium and subsequently the left ventricle. In the fetus the de-oxygenated blood, returning from the superior vena cava and upper body segment, enters the right atrium and then moves to the right ventricle through the AV valve. From here the de-oxygenated blood moves, via the ductus arteriosus, to the descending aorta to return via the umbilical arteries to the mother's placenta. During birth, with the first breath, pulmonary arterial resistance begins to decrease and the lungs begin the process of respiration. In normal conditions pulmonary venous blood returning to the left atrium closes the flap of the foramen ovale, and the

ductus arteriosus begins to close, through mechanical and chemical mechanisms. In normal infants, this is normally accomplished 10 to 15 hours after birth. Intermittent right-to-left atrial level shunting through the foramen ovale may occur, particularly if pulmonary vascular resistance fails to decrease. Structural cardiac abnormalities requiring patency (failure of the ductus to close) of the ductus arteriosus for maintenance of either pulmonary or systemic blood flow most often present within the first few days of life. In the absence of an associated pathologic condition, hemodynamically significant ventricular septal defects are seldom present before two weeks of age, additionally atrial septal defects are seldom symptomatic in infancy.

Because the fetal heart has a circulatory system different from the one after birth, it may be days or weeks before some congenital heart defects are found. Thus, the age of the pediatric patient being evaluated influences the spectra of possible heart diseases to be considered [26].

2.1.3 The cardiac cycle

The cardiac cycle includes both a period of contraction and one of relaxation. The heart performs this combination of contraction and relaxation approximately 100 000 times a day. For any chamber in the heart, the cardiac cycle can be divided into these two phases. During contraction, or systole, the chamber pushes blood into an adjacent chamber or into an arterial trunk. And during diastole, or relaxation, the chamber fills with blood and prepares for the start of the next cardiac cycle. In the normal functioning of the circulatory system blood moves from an area of higher pressure to one of lower pressure. During the cardiac cycle, the pressure within each chamber increases during systole and decreases during diastole. An increase in pressure in one chamber will cause blood to flow to another chamber or vessel where the pressure is lower. In normal operations the AV and semilunar valves ensure that blood flows in one direction. The correct pressure relationships depend on the careful timing of contractions. In the normal heart, atrial systole and atrial diastole are out of phase with ventricular systole and diastole. *Figure 2.3* shows the duration and timing of the atrial and ventricular systole and diastole for a normal heart with a rate of 113 bpm.

The cardiac cycle starts with atrial systole. At the start of atrial systole, the ventricles are filled to approximately 70 % of capacity; atrial systole fills them completely by adding the additional 30 %. As atrial systole ends, ventricular systole begins. When the pressures in the ventricles rise above the pressure in the atria, the AV valves

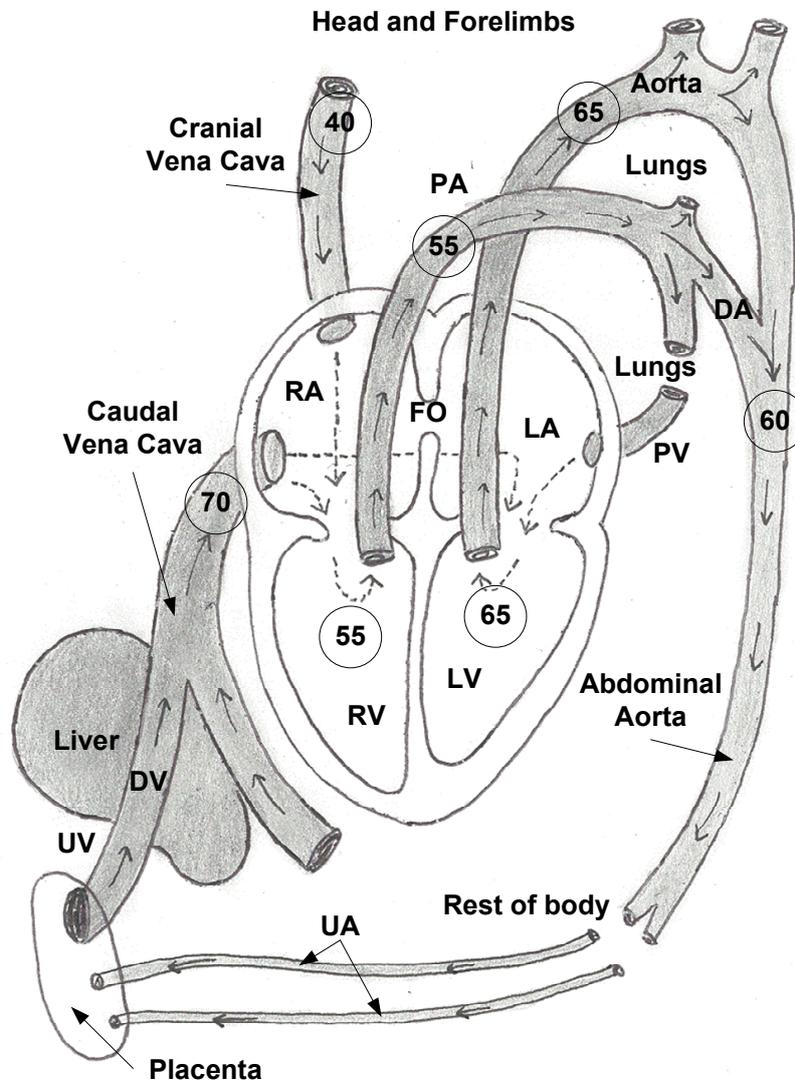


Figure 2.2: Schematic diagram of the fetal circulation. The figures in the circles within the chambers and the vessels represent the oxygen saturation percentages for the respective parts. UV, umbilical vein; UA, umbilical artery; DV, ductus venosus; DA, ductus arteriosus; FO, foramen ovale; LV, left ventricle; LA, left atrium; RV, right ventricle; RA, right atrium; PA, pulmonary artery. Draw, from a diagram illustrated by the Department of Anatomy, University of Bristol, by Thia de Vos

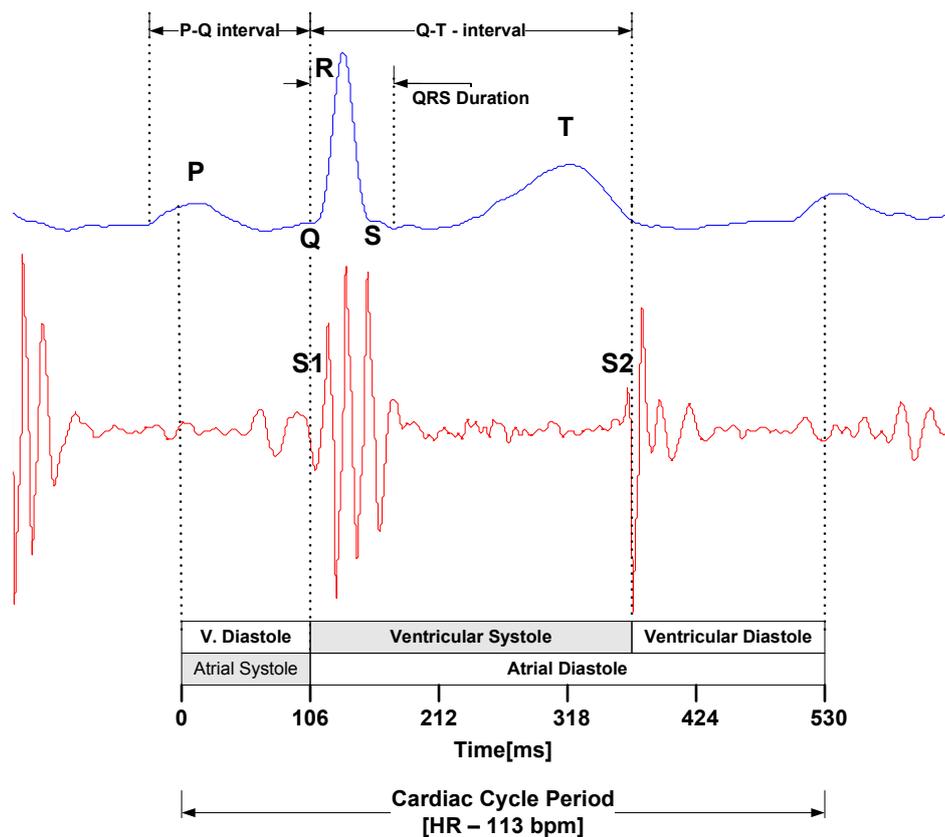


Figure 2.3: The cardiac cycle. ECG section (top) and heart sound (bottom).

swing shut. At the point where the pressure in the ventricles exceed the pressure in the aorta and pulmonary trunk, the blood pushes open the semilunar valves and flows into the aorta and pulmonary trunk.

When ventricular diastole begins, ventricular pressure declines rapidly. As ventricular pressure falls below the pressures of the atrial trunks, the semi-lunar valves close. Ventricular pressures continue to drop; as they fall below atrial pressures, the mitral and tricuspid valves open and blood flows from the atria into the ventricles. Both atria and ventricles are now in diastole; blood now flows from the major veins through the relaxed atria and into the ventricles. By the time atrial systole marks the start of another cardiac cycle, the ventricles are roughly 70 % filled [1].

Structural congenital heart diseases affect the normal cardiac cycle dynamics (timing), and will be discussed shortly.

2.1.4 ECG Morphology

The electrocardiogram (ECG) is a time-varying signal reflecting the ionic flow which causes the cardiac fibres to contract and subsequently relax [27].

The surface ECG is obtained by recording the potential difference between two electrodes placed on the body surface. A single cycle of the ECG represents the successive atrial depolarization and repolarization and ventricular depolarization and repolarization which occur during every heartbeat as described in section 2.1.3.

Each heart beat can be observed as a series of deflections away from the baseline (isoelectric line) of the ECG. These deflections reflect the time evolution of electrical activities in the heart which initiates muscle contraction. A single normal cycle of the ECG, corresponding to one heartbeat, is traditionally labeled with the letters P, Q, R, S and T on each of its turning points. The ECG may be divided into the following sections, with reference to *Figure 2.3*:

- P-wave: A small deflection away from the baseline caused by the depolarization of the atria prior to atrial contraction. The deflection appears as the activation (depolarization) wavefront propagates from the SA-node through the atria.
- PQ-interval: The time elapse between the beginning of atrial depolarization and the beginning of ventricular depolarization.
- QRS-complex: The largest amplitude section of the ECG, is caused by currents generated when the ventricles depolarize. Atrial depolarization is not visible on the ECG, because the ventricular waveform is of much greater amplitude.
- QT-interval: The time between the onset of ventricular depolarization and the end of ventricular re-polarization. The relationship between the RR-interval (heart cycle) duration and the QT-interval is discussed in detail in section 3.4.1.1.
- ST-interval: The time between the end of the S-wave and the beginning of the T-wave.
- T-wave: Ventricular re-polarization, whereby the cardiac muscle is prepared for the next cycle of the ECG.

2.1.5 Heart sounds

When listening to a normal heart sound a first and second sound can be heard. Each pair of sounds "lub-dub", "lub-dub", begin with the first sound (S1) and end with the second sound (S2). Heart sounds are of two types: high-frequency transient sounds associated with the abrupt terminal checking of valves that are closing or opening and low-frequency sounds related to early and late diastolic filling events of the ventricles. The process of listening, usually with the aid of a stethoscope, to sounds produced by the movement of gas or liquid within the body, is called auscultation. Auscultation is an aid used in diagnosis of abnormalities of the heart and other organs according to the characteristics changes in sound pattern caused by different disease processes. If reference is made to auscultation in the rest of the paper, it is in the context of cardiac auscultation and not that of other organs or processes (eg. respiratory processes).

Figure 2.3 shows the synchronous timing relationship between the ECG signal and that of the heart sound. The first heart sound (S1) arises from closure of the atrioventricular (mitral and tricuspid) valves in early isovolumic ventricular contraction and consequently is heard best in the tricuspid and mitral areas. Mitral valve closure occurs slightly in advance of tricuspid valve closure, and occasionally two components (splitting) of the S1 may be heard near the lower left sternal edge. Normally, it is heard as a single sound. The S1 is most easily heard when the heart rate is slow because the interval between the S1 and S2 is clearly shorter than the interval between the S2 and subsequent S1. The intensity of the S1 is influenced by the position of the atrioventricular valve at the onset of ventricular contraction. If the valve's leaflets are far apart, the increased excursion to accomplish valve closure increases the intensity of the S1 [5].

Shortly after the onset of ventricular contraction, the semilunar valves (aortic and pulmonary) open and permit ventricular ejection. Normally, this opening does not generate any sound. The atrioventricular valves remain tightly closed during ventricular ejection. As ventricular ejection nears completion, the pressure begins to fall within the ventricles, and the semilunar valves snap shut, closing tightly. This prevents regurgitation from the aorta and pulmonary artery back into the heart. The closure of the semilunar valves generates the S2. Normally, the second heart sound consists of a louder and earlier aortic valve closure followed by a later and quieter pulmonary valve closure sound. Normal splitting of the S2 is caused by (i) increased right heart filling during inspiration because of increased blood flow (2) diminished left heart filling because blood is retained within the small blood vessels of the lungs

when the thorax expands. During inspiration, when the right ventricle is filled more than the left, it takes slightly longer to empty. This causes the noticeable inspiratory delay in pulmonary valve closure relative to aortic valve closure. Splitting of the S2 during inspiration is thus a normal finding and should be sought in all patients [5].

2.2 Heart murmurs - Innocent and pathological

A cardiac murmur is defined as a relatively prolonged series of auditory vibrations of varying intensity (loudness), frequency (pitch) in the range of 20 Hz - 650 Hz [28], quality, configuration and duration [29]. Although the exact physical principles that govern the production of murmurs have been debated for years, it is generally agreed that turbulence is the prime factor responsible for most murmurs. See section 2.4 for discussion on murmur dynamics.

The production of murmurs can be attributed to three main factors: (1) high flow rate through normal or abnormal orifices, (2) forward flow through a constricted or irregular orifice or into a dilated vessel or chamber, (3) backward or regurgitant flow through an incompetent valve, septal defect, or *patent ductus arteriosus*. In many cases, a combination of these factors is operative [26].

Not all cardiac murmurs indicate anatomical or physiological problems. To be able to differentiate primary physicians are taught to determine and describe the following characteristics of a murmur to classify the murmur as innocent or pathological [5, 1]:

1. **Timing:** The relative position within the cardiac cycle with respect to S1 and S2, classify murmurs as either systolic, diastolic or continuous.

(i) Systolic murmurs

Systolic murmurs begin with or follow the first heart sound and end before the second heart sound. See Figure 2.4 for the following four types of systolic murmurs:

Holosystolic murmur: This murmur begins with S1 and continues with the same intensity to S2. This murmur can occur when an insufficient mitral or tricuspid valve is present or in association with the majority of ventricular septal defects [5]. Systolic insufficient(regurgitant) murmurs

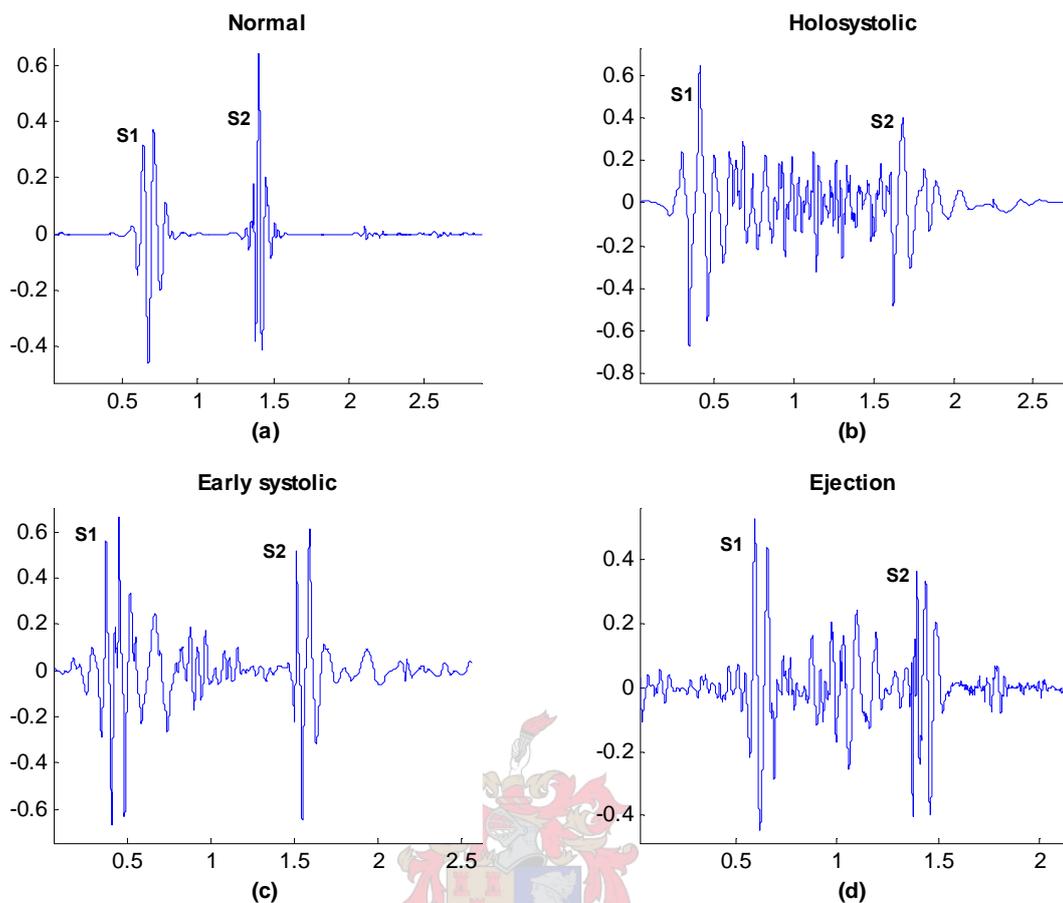


Figure 2.4: The normal heart sound (a) with three types of systolic murmurs (b, c, d). Sounds were de-noised with the fixed threshold wavelet de-noising technique discussed in section 3.3.1.4, to assure the emphasis on the dynamic shape of the murmur

are due to backwards flow from a high-pressure cardiac chamber to a low-pressure chamber [26].

Early systolic murmur: It starts abruptly with S1 but disappears before the second heart sound and is exclusively associated with small muscular VSD's.

Ejection murmur: The systolic ejection murmur begins shortly after the pressure in the left or right ventricle exceeds the aortic or pulmonary diastolic pressure sufficiently to open the aortic or pulmonary valve. Systolic ejection murmurs are due to forward flow across the left ventricular or right ventricular semilunar valves. Ejection (*Crescendo-decrescendo*) murmurs may arise from the narrowing of the semilunar valves or outflow tracts. The rising and falling nature of the murmur reflects the periods

of low-flow at the beginning and end of ventricular systole. The energy envelope of the murmur corresponds to the contour of the flow velocity.

Because of the high correlation between the shape of the murmur and its underlying flow- velocity characteristics, careful attention must be given during auscultation to the shape and the duration of the murmur as well as to its intensity.

Mid-to-late systolic murmur: This murmur begins midway through systole and is often heard in association with the midsystolic clicks and mitral insufficiency.

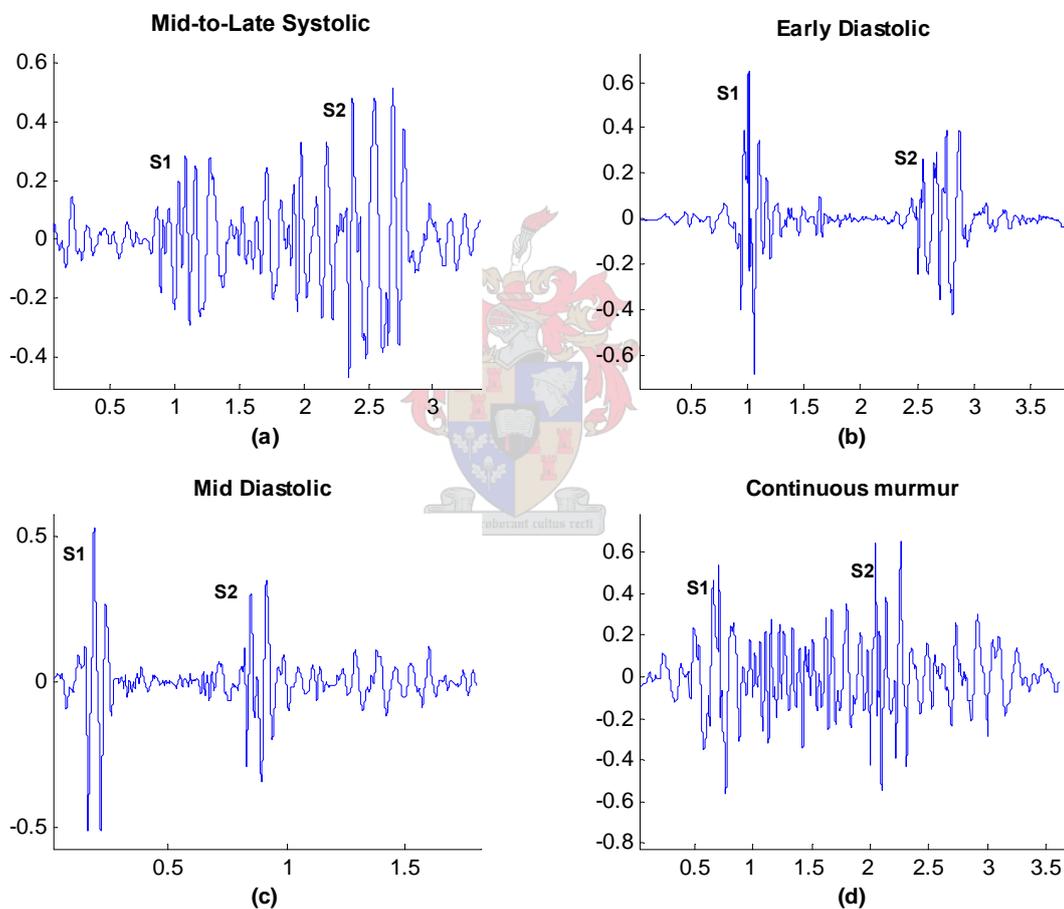


Figure 2.5: Mid-to-late systolic murmur (a) with the two types of diastolic murmurs (b, c) and an example of a continuous murmur (d). Sounds were de-noised with the fixed threshold wavelet de-noising technique discussed in section 3.3.1.4, to assure the emphasis on the dynamic shape of the murmur

(ii) Diastolic murmurs

Diastole, the period between the closure of the semilunar valves (S2) and the subsequent closure of the AV valves (S1), is normally silent. This is because the low turbulence associated with this low-pressure flow through the relatively large valve orifices. However, regurgitation of the semilunar valves, stenosis of an atrioventricular valve, or an increase flow across an atrioventricular valve can all cause turbulence and may produce diastolic heart murmurs. See Figure 2.5 (b) & (c) for the different types of diastolic murmurs.

Early diastolic murmurs are *decrecendo* in nature and arise from either aortic or pulmonary valve insufficiency. Mid-diastolic murmurs are diamond-shaped and occur because of either increased flow across a normal tricuspid or mitral valve or normal flow across an obstructed or stenotic tricuspid or mitral valve. Late diastolic or *crescendo* murmurs are created by stenotic or narrowed AV valves and occur in association with atrial contraction. "Diastolic murmurs should always be regarded as pathological" [30].

(iii) **Continuous murmurs**

Blood flow through vessels or channels distal to the aortic and pulmonary valves is not confined to systole and diastole. Thus, turbulent flow may occur throughout the cardiac cycle. The continuous murmurs extend beyond S2, see (Figure 2.5 (d)). With the exception of the venous hum (discussed later), continuous murmurs must always be considered as pathological.

2. **Intensity and loudness:** Although the intensity of the systolic murmur is not always directly proportional to the level of the hemodynamic disturbance, grading (rating) the loudness of a murmur is generally used as a differentiation indicator [26]. Murmurs are graded as follow:

Grade 1: Heard only with intense concentration.

Grade 2: Faint, but heard immediately.

Grade 3: Easily heard, of intermediate intensity.

Grade 4: Easily heard and associated with a thrill (a palpable vibration of the chest wall).

Grade 5: Very loud, thrill present, and audible with only the edge of the stethoscope on the chest wall.

Grade 6: Audible with the stethoscope off the chest wall.

The intensity of the murmur varies directly with the velocity of blood flow across the area of murmur production. See section 2.4 for detailed description on murmur dynamics. Experience has shown that systolic murmurs of grade three or more in intensity are usually pathological [26]. The intensity of the murmur, as heard at the chest wall, is also determined by the transmission characteristics of the body tissues between the source of the murmur and the stethoscope head.

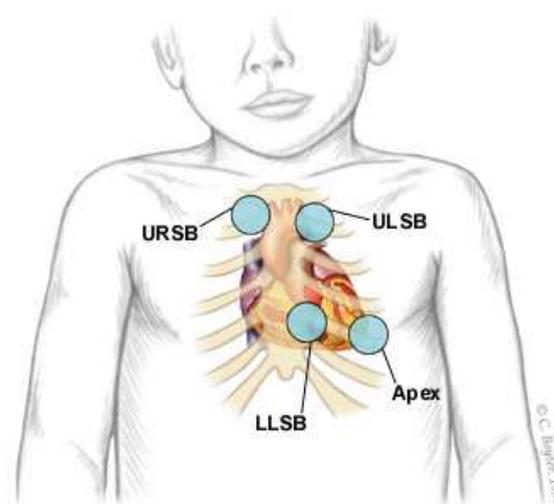


Figure 2.6: Main auscultation areas for heart sounds, murmurs and clicks: upper right sternal border (URSB) - the aortic area; upper left sternal border (ULSB)- the pulmonary area; lower left sternal border (LLSB)- the tricuspid area; apex - the mitral (bicuspid) area. (Courtesy of the American Academy of Family Physicians, August 1999)

3. Location on the chest wall with regard to:

- (a) the area where the sound is loudest (point of maximum intensity);
- (b) the area over which the sound is audible (extent of radiation).

The location and radiation of a murmur are determined by a combination of factors. Some of them are the site of origin, the intensity, and direction of blood flow, as well as the physical characteristics of the chest [5]. See Figure 2.6 for the four primary auscultation areas. These areas define the general regions where heart sounds and murmurs of the four cardiac valves are often best heard and defined. Thorough auscultation of cooperative patients, should be done with the patient in the supine, sitting, and standing positions; and

should include listening at the four indicated areas with both the bell and the diaphragm mode.

4. **Duration:** The time elapsed from beginning to end of the murmur.
5. **Configuration:** The dynamic shape (envelope) of the murmur. The duration and time intensity contour (murmur envelope) of a specific murmur are directly related to the blood flow velocity causing the murmur.
6. **Pitch:** The frequency spectra of the murmur. The frequency of the murmur bears a direct relationship to the velocity of blood flow. The low velocity flow resulting from a small pressure head across a stenotic mitral valve produces a low-pitched rumbling murmur, whereas the large diastolic pressure gradient across a regurgitant (insufficient) aortic valve causes a high-pitched murmur [31]. A recent study has demonstrated that the dominant frequencies contained in heart murmurs, due to stenosis, are directly related to the instantaneous jet velocities distal to the associated obstruction [1].
7. **Quality and associated manifestations:** The presence of harmonics and overtones, and the company that the murmur keeps can be a possible indicators of pathology. Examples include the fixed split S2 of the atrial septal defect, the decreased intensity of S2 in aortic stenosis, and the systolic click of mitral valve prolapse.



2.2.1 Innocent murmurs

Innocent heart murmurs are murmurs found in people with normal hearts and are harmless. They are common in children and may disappear and reappear throughout childhood. Their dynamics change depending on the varying acoustics with growth, and the amount of blood flow through the heart. Innocent murmurs are emphasized by fever, anemia, or increased cardiac output (such as when excited)[5]. Most innocent heart murmurs disappear, or are not heard as a child nears adulthood because of the changes in heart rate, acoustic and relative amount of blood flow through the heart.

"Innocent murmurs are almost exclusively ejection systolic in nature (never solely diastolic)" [26]. They occur, without evidence of physiological or anatomical abnormalities, when peak flow velocity in early systole exceeds the murmur threshold (see section 2.4 for further discussion) [32]. These murmurs are almost always less

than grade 3 in intensity and are subjected to considerable variation with changes in the positioning of the patient or the level of physical activity.

Considerable controversy exists as to the origin of the vibratory systolic murmur. Most authorities [5, 26] agree that innocent murmurs arise from flow across either the normal LV or RV outflow tract and always end well before semilunar valve closure. Innocent murmurs originating from the RV outflow tract have been termed innocent pulmonary systolic murmurs because of their maximal intensity in the pulmonary area. These murmurs are low to medium in pitch, with a blowing quality.

Since both innocent and pathological ejection murmurs have the same mechanism of production, it is often the company the murmur keeps that indicates the differential diagnosis of the pathological systolic ejection murmur from the innocent murmur.

The most common innocent murmurs are comprised of six systolic and two continuous types [5].

Systolic innocent murmurs:

Vibratory Still's murmur This is the most common innocent murmur in children and is most typically audible between ages 2 and 6 years. The murmur is low to medium in pitch, confined to early systole, generally grade 2 (range grade 1-3), and maximal at the lower left sternal edge. The murmur is generally loudest in the supine position and often changes in dynamics and frequency with upright positioning. The most characteristic feature of the murmur is its vibratory quality. This quality of the murmur gives it a pleasing or musical character. The Still's murmur's origin has been ascribed to vibration of the pulmonary valves during systolic ejection.

Pulmonary flow murmur An innocent pulmonary outflow tract murmur may be heard in children, adolescents and young adults. The envelope of the murmur is *crescendo-decrescendo* in nature, early to mid systolic that is confined to the second and third inter-space at the left sternal border. It is low in intensity (grade 2-3) and radiates to the pulmonary area. The pulmonary flow murmur is rough and dissonant without the vibratory musical quality of the Still's murmur.

Peripheral pulmonary arterial stenosis murmur This is a common murmur heard frequently in newborns and infants younger than one year. The audible turbu-

lence is caused by peripheral branch pulmonary arterial stenosis or narrowing. These ejection character murmurs are typically grade 1 to 2, low to moderately pitched, beginning in early to mid-systole, and extending up to and occasionally just beyond S2. This murmur is best heard with both regional and temporal variability, peripherally in the axillae and back.

Supraclavicular systolic murmur This *crescendo-decrescendo* murmur may occur in children and young adults. The murmur is low to medium in pitch, of abrupt onset, and maximal in the first half or two thirds of systole. This systolic murmur is audible maximally above the clavicles and radiates to the neck but may present to a lesser degree on the superior chest.

Aortic systolic murmur Innocent systolic flow murmurs may arise from the out-flow tract in older children and adults. The murmurs are ejection in character, confined to systole and maximally audible in the aortic area. In children, these murmurs may arise secondarily to extreme anxiety, anemia, fever, or any other condition of increased systemic cardiac output.

Continuous murmurs:

Venous hum The most common type of continuous murmur heard in children is the innocent cervical venous hum. This continuous murmur is most audible on the low anterior part of the neck but can readily extend to the infraclavicular area of the anterior chest wall. The murmur is generally more intense on the right than on the left, louder with the patient sitting than lying, and is emphasized in diastole. Intensity varies from faint to grade 6 and is quite variable in character.

Mammary arterial soufflé This murmur only occurs late in pregnancy, in lactating woman and rarely occurs in adolescence. Recording such a murmur in a pediatric patient is not possible.

A detailed description of the above cases can be found in [10, 26, 5, 7]

2.2.2 Conclusions regarding auscultation for pediatric murmur evaluation

If it is possible that the classification of the characteristics listed in section 2.2 can lead to the differentiation between a pathological and an innocent case, then in the-

ory these characteristics can be exploited to make a automated differentiation diagnosis.

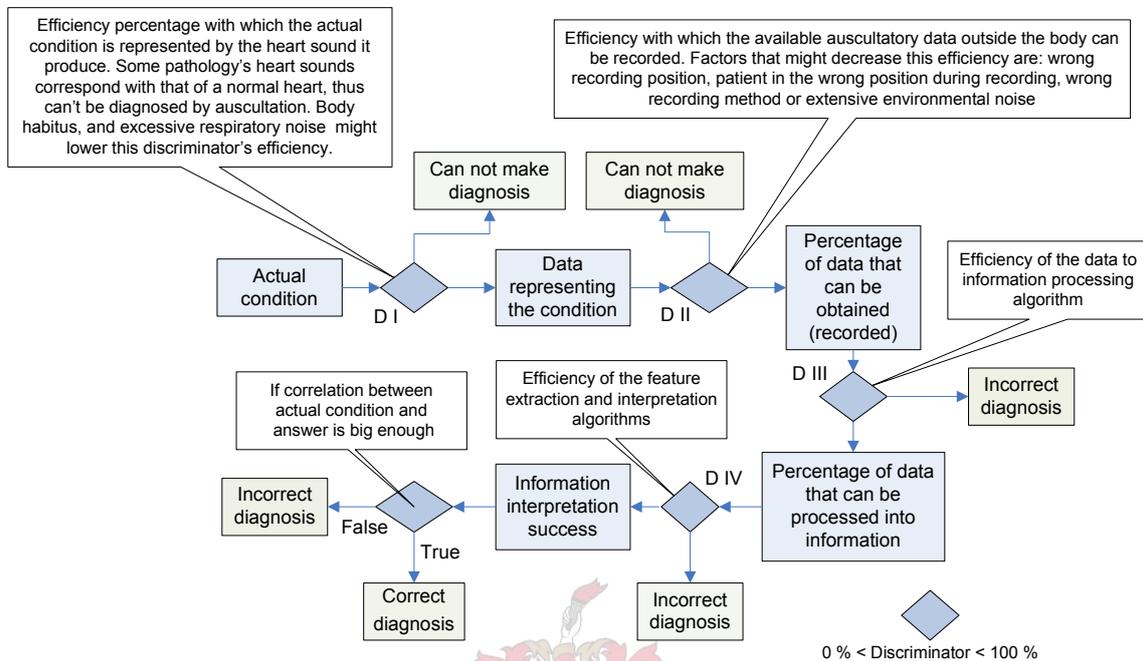


Figure 2.7: Levels of making a successful diagnostic differentiating, with inter-level discriminating factors (discriminators).

The inter-level differentiation factors of each of the discriminators shown in Figure 2.7 are important in attaining the feasibility of developing an automated diagnosis system. There is no motivation for developing an automated diagnostic system if the input data to the classification system is not a sufficient representation of the actual condition. What makes this a difficult question to answer is that each pathological condition has its own combination of differentiation factors in the above diagram (some of which are empirical; thus only learned through extensive testing)

Assigning values to the differentiation factors of the first filter stage and partly the second filter stage, poses a medical question with physical constraints and boundaries. Answering this is not within the scope of this study. The question that this study tries to address is: In the case that the data after discriminator II is a sufficient representation of the actual condition, is it possible that this data can then be sufficiently extracted (to extract something one must first be able to recognize what to extract) and exploited to obtain enough information to make a differentiation between a pathological and an innocent murmur?

In other words - can the differentiation factor in discriminators III & IV be made sufficiently high for a correct diagnosis?

In conclusion, the evaluation of cardiac murmurs represent one of the most skilled and demanding aspects of the pediatric physical assessment. In most general cases the characteristics listed in section 2.2, in combination with a physical examination, can be used as a sufficient differentiation indicator. It is however evident that in some cases not all the characteristics can be classified efficiently to make a differentiation. Recent advances in signal processing and acquisition techniques caused several research groups to investigate whether it would be possible to automatically extract some characteristics of the heart sound. The next section takes a look at the work already done.

2.3 Initial investigation and current theories

Recent literature describes the success of various time-frequency signal processing techniques in eliciting features from heart sounds to distinguish between pathology and normal heart sounds. Although automatic heart sound screening has been described as early as 1968 [33], a useful implementation of PCG signal processing techniques were only published in 1988 by Rangayyan and Lehner [34]. Their technique used Fourier transforms of the systolic and diastolic intervals to isolate energy above 200 Hz, which they linked to the presence of murmurs. Although they associated different power spectral characteristics with certain conditions they did not report any sensitivity or specificity values [6]. See section 3.5.4 for detailed explanation on sensitivity and specificity.

With the rediscovering of the application abilities of the wavelet transform in the early 1990's, McDonnell and Bentley published work using the wavelet transform in cardiovascular signal analysis [35]. They used wavelet analysis in detecting certain pathological heart conditions through auscultation. They investigated various time domain and frequency domain techniques, suggesting that looking at the time-dependent frequency and intensity of the murmur might serve as a detection mechanism for pathology. Although these studies did not include results obtained from extensive clinical testing, they pioneered a new field of investigation, making way for several successive research groups to investigate these time-frequency techniques.

In a study of 222 consecutive patients referred for evaluation of a heart murmur, Mc-

Crindle *et al* [36] found that six cardinal clinical signs on cardiac examination proved to be significant independent predictors of the presence of a confirmed cardiac lesion. The six signs were murmur intensity grade ≥ 3 , best heard at the left upper sternal border (LUSB), harsh quality, pansystolic timing, the presence of a systolic click, or the presence of an abnormal second heart sound.

In 2002 Hayek, Thompson, Tuchinda *et al* published an article describing the development of a wavelet-based time-frequency murmur diagnostic instrument [8]. They improved on previous efforts by combining a cardiologist's auscultation expertise, a large database of comprehensive heart sound files, and recent advances in signal processing techniques. The algorithm was developed to identify systolic murmurs that were indicative of heart defects and that exhibited one or more of the qualities listed by McCrindle. Their algorithm calculated the midsystolic energy (recorded at the LUSB) present in pathological murmurs using different wavelet scales. A distinction was made between healthy and pathological hearts on the basis of the calculated energy value being above or below a chosen threshold. From an extensive study done on 194 children and young adults, from which 99 had a pathological murmur present at the LUSB, their algorithm report to be 78 % sensitive and 96 % specific.

During the same period DeGroff, Bhatikar, Hertzberg *et al* [37] did research on the same topic but used artificial neural networks (ANN's) to screen for heart murmurs. Although there have been several studies on the use of ANN's on heart sounds [22, ?, 21], DeGroff *et al* published the latest results on using ANN's to distinguish between innocent and pathological murmurs. They trained a three layer feed-forward neural network with three representative consecutive heart cycles from 69 patients (37 pathological cases and 32 functional cases). The algorithm developed was not automated because the above mentioned representative heart cycles needed to be hand picked. Using the same data to validate the network, their network report to be 100 % sensitive and 100 % specific. Using new data as validation set, the ANN was able to classify 7 of 9 pathological examples and 5 of 6 innocent examples correctly. Reportedly all the misclassified pathological examples were due to a gross under-representation in the training data.

2.4 Murmur dynamics

As stated in section 2.2, turbulence is the main cause of murmur generations. Turbulence occurs when flow velocity becomes critically high due to high flow, or flow

through an irregular or narrowed area, or a combination of both [26]. Different assumptions are made during the derivation of basic fluid mechanics equations. Most models describing the flow within the human heart are, however, derived from empirical observations.

A short explanation around the transition between laminar flow and turbulent flow is given to explain the possible effect that stenotic - and/or regurgitant valves and pathology inflicted pressure differences might have in the generation of murmurs.

Poiseuille [38] found a relationship defining flow, the equation is

$$Q = \frac{(P_v - P_a) \pi r^4}{8\mu\ell} (ml/sec) \quad (2.4.1)$$

where

$\Delta P = (P_v - P_a)$ is the pressure difference between two points

ℓ is the distance between the two points

μ is the fluid's viscosity

and r is the diameter of the vessel.

The flow described in equation 2.4.1 can also be written as $Q = \frac{\Delta P}{R}$, with R equals to the flow resistance. Hence, resistance depends on 2 physical factors: (i) the effect of radius and vessel length; and (ii) the tendency of a fluid to resist flow.

From the equation Total energy = Potential energy + Kinetic energy [39], the energy captured in the flow volume is written as:

$$E = PV + 0.5\rho v^2V \quad (2.4.2)$$

where

P = pressure

V = volume

ρ = density

and v = mean velocity = $\frac{Q}{Area}$ ($cm/sec = \frac{cm^3/sec}{cm^2}$),

if the assumption is made that there are no frictional losses in this system.

Thus, in a particular vessel as you lower the diameter, velocity is increased. And in the case of laminar flow the rate of flow is increased as pressure increases. However, at some point the rate of flow trails off with the increasing pressure, and turbulence starts to occur. The Reynolds number of a particular fluid is a number that is used to

indicating the point where the flow dynamics has a good chance of changing from laminar to turbulent flow (note that R has no units).

$$R_e = \frac{vr\rho}{\mu} \quad (2.4.3)$$

where

v = mean velocity

r = vessels diameter

ρ = fluid density

μ = fluid viscosity.

According to [40] blood flow was found to be laminar when the Reynolds number was below 2,000, transitional in the range of Reynolds numbers from 2,000 to 3,000, and fully turbulent above Reynolds number 3,000. If drawing a pressure vs. flow graph the point where the flow dynamics move from laminar to turbulent will be visible.

From equation 2.4.1, 2.4.2 and 2.4.3 the following statements are derived:

- As the diameter of a orifice, vessel or chamber is narrowed the formation of turbulence is favoured.
- Increase in flow velocity because of a increase in pressure, increases chances of turbulence.
- With increase cardiac output, formation of turbulence is favoured.

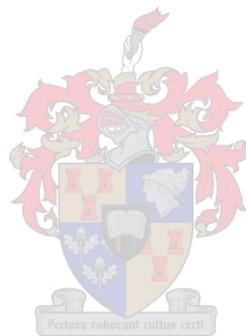
For example: In the case of a stenotic pulmonary valve, the diameter of the orifice is narrowed. During ventricular systole pressure is increased in the right ventricle causing the ventricular blood to eject through the narrowed opening. Ejection flow velocity is higher than normal, because of the decrease in orifice area. This cause the calculated Reynold number to exceed 3000, which cause turbulence on the distal side of the orifice during the early systolic contraction. This turbulence can be heard as an ejection systolic murmur. Turbulences fade away during mid-to-late systole due to the decrease in the pressure difference over the pulmonary valve. By utilizing equations 2.4.1, 2.4.2 and 2.4.3 a physical explanation can be found for all the other types of murmurs.

2.5 Formulation of hypothesis

From the above argument emerges the following hypothesis:

From listed heart murmur analysis studies and clinical observation, it is postulated that most (a high percentage of) pathological precordial murmurs possess higher frequency and intensity auscultatory sounds than innocent murmurs. In addition to the frequency and intensity of murmurs, the timing and other cardinal clinical signs accompanying the murmur also serves as a differentiation indicator. The study is also extended to analysis of the other auscultation positions, so as to more closely parallel standard auscultation examination procedures.

The following two chapters examines whether automated digital analysis of individual cardiac sound-components can differentiate between normal heart sounds, innocent murmurs and pathological murmurs.



Chapter 3

Methodology

Three feature extraction algorithms (FEA's) are developed in this chapter. These are the Direct Ratio- , Wavelet- and Neural Network algorithms. The three algorithms use the same pre-processing methods, and are evaluated using the same statistics. This chapter discusses the development of each procedure as illustrated in Figure 3.1 in detail. The procedures will be discussed in the given order. All the procedures were developed using *Matlab*. In the discussion of each procedure reference is made to the associated *Matlab* code (as completely listed in Appendix E). No results are reported during the discussion, as this is the purpose of the next chapter.



3.1 Data collection

3.1.1 Subject population

Pediatric patients, seen over a 5 month period, had their clinical auscultation sounds and ECG-data digitized. Recordings were done at three different institutes in co-operation with different medical personnel. Only patients aged between 1 month and 16 years were recorded and added to the database. The most recordings were done at the Louis Leipoldt Medi-Clinic hospital in Tygerberg, South Africa, in co-operation with Dr. Joan Hunter, a pediatric cardiologist. Heart sound- and ECG recordings were made if sufficient consulting time was available, and if the patient was co-operative enough to allow minimal background noise. Additional recordings were done at the pediatric cardiology clinic held on Mondays at the Tygerberg Children's Hospital in co-operation with Prof. PL van der Merwe, a lecturer in pediatric cardiology at the University of Stellenbosch. After a four month pe-

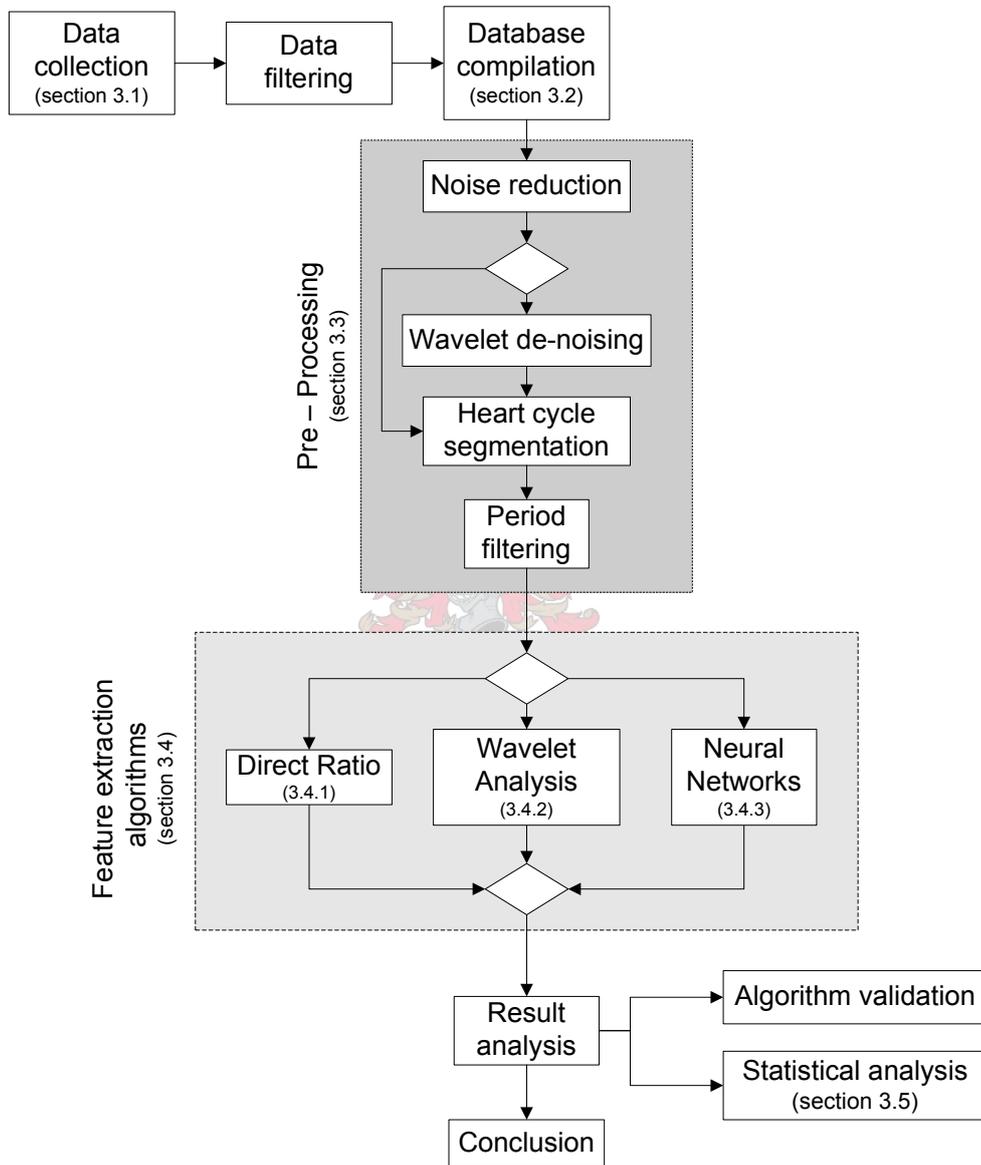


Figure 3.1: Methodology layout of the three feature extraction algorithms developed.

riod it was concluded that there were insufficient normal (no murmur) recordings in the database. This obstacle was overcome by the recordings done at the primary care center in Tulbagh in co-operation with Dr. G Schoonbee, a general practitioner, and Mrs. A Phaff, a final year medical student. Patients with and without heart disease were examined; in particular, those with innocent heart murmurs were included.

A protocol for the study was drawn up and was approved by the Human Subjects Research Committee of the University of Stellenbosch. The project's reference number is N04/04/077 and it can be consulted for further detail. Every patient's guardian or custodian completed an information and informed consent document before the child's heart sound and ECG were recorded. See Appendix A for the information and informed consent document. Patients were excluded from the database if informed consent was not obtained or if the examinations were contaminated by unacceptable noise artifacts.

3.1.2 Data acquisition

3.1.2.1 Recording equipment

Initially the proposal was to use a custom-built electronic stethoscope and ECG monitor to do the recordings. After developing the recording systems, the Human Subjects Research Committee and the medical personal performing the medical examination, stated that only medically approved recording equipment can be used during subject evaluation. The developed unit's design layout is illustrated in Appendix B, together with all the circuit schematics and board layouts.

A commercially bought *Welch & Allen Meditron Analyzer ECG* electronic stethoscope was used for the auscultation and ECG recordings. The auscultation unit used the patented *Meditron* piezoelectric contact sensor as pick-up. The *Meditron Origo Sensor System (moss)* is a new technology for registration, transmission and amplification of sound waves. The sensor which is incorporated in the *Meditron* Stethoscopes is small, but highly directional and sensitive. The sensor has an extended frequency pick-up range of (20 -20 000Hz), and amplifies sound without extraneous and disturbing electronic noise. The stethoscope has two frequency settings, one for the heart frequency range (20 - 420 Hz) and one for the lung frequency range (200-20 000 Hz). The heart frequency range was used for recordings purposes.

The heart sound and ECG-signal were recorded on separate channels using a sampling frequency of 11025 Hz per channel, with a 16 bit dynamic resolution. The data was sent via a USB connection to an *Acer Travelmate 354TEV* notebook to be collected in a database and stored in separate files as a *.wav file (wave file). A study done by B.D. Lauritz et al. [41] concluded that the *Meditron* electronic stethoscope, is adequate for the electronic referral of heart sounds.

3.1.2.2 Recording procedures

During the general examination, which included an echocardiogram study in pathological cases (if conditions apply), the specialist made a diagnosis as to whether the patient has a pathological-, innocent- or no murmur.

The 3-lead ECG's electrodes were applied to the patient's chest, with the ground electrode positioned on the left upper-abdomen, and the other two reference electrodes positioned on the left and right upper-thorax region. Electrodes were only used once to ensure good contact. If the iso-electric line was unstable, the electrodes were re-positioned until a relatively stable iso-electric line was obtained. Simultaneous to the ECG recording, auscultation recordings were done on another channel. Most recordings were done with the patient in supine position, if otherwise it was so stated in the patients examination report. Auscultation recordings were done in the areas where the murmur is most audible (at positions of maximum intensity). A minimum of two 15 second recordings were made per patient. In pathological 'textbook cases' recordings were made from each of the five usual auscultatory areas. Cases without a murmur were recorded in the LUSB area, due to its general association with maximum heart sound intensity.

After filling in the ticklist report on the auscultatory findings, a digital echocardiogram was taken and recorded on VCR to serve as the golden standard. Due to time limitations, and the limited availability of echocardiographic equipment, none of the recordings done at the Tygerberg Pediatric Cardiology Clinic is accompanied by an echocardiogram. Most of the diagnoses done here were however confirmed by an echocardiogram taken earlier, or later in the case of uncertainty. The whole recording protocol took 5 to 10 minutes per patient to complete (this includes the time it took to explain the procedure to the patient's guardian and the time it took them to complete the information and informed-consent document).

Each recording was validated by the specialist performing or assisting the recording procedure. If a recording was not a good enough representation of the specific case,

or was a too noisy recording, it was discarded from the patient's record. The patients evaluated in Tulbagh, over which there was a certain degree of uncertainty during diagnosis, were taken to the Tygerberg Pediatric Cardiology Clinic for re-assessment by a specialist.

3.2 Database compilation

After the examination and recording procedure each patient has the following data and information on his/her record:

- Patient's personal information:
 - Name and surname,
 - Age,
 - Gender,
 - Geographical information and language,
 - ID number and patient recording number;
- Informed consent document with signature or right thumb print of patient's guardian
- Clinical examination report¹;
- Ticklist with auscultation findings;
- Minimum of two auscultation and ECG recordings;
- Written comment on auscultation data;
- Diagnosis with a murmur intensity rating ((1-6) for systolic murmurs and (1-4)for diastolic murmurs) and an additional comment on the procedure;
- Digital echocardiogram with report on flow numbers ².

Before adding the data to the database, all the recordings were listened to, so as to ensure consistency in quality. Recordings with an unacceptably low quality, or excessive artifact noise, were discarded. In cases where only a small part of the recording contained unacceptable noise (normally at the beginning of the recording), only

¹Only added for patients recorded by Dr. Hunter

²If conditions apply

this part was cut from the recording. Finally before adding the data to the database, all the sound recordings were passed through a 800 Hz low-pass filter (LPF) that is part of the *Welch&Allen* recording equipment. Although the bell filter characteristics of the *Meditron* stethoscope stated a frequency passband of 20-420 Hz, recordings showed frequencies of up to 1200 Hz on the spectrogram. Investigation in ideal circumstances (no respiratory activity), showed that the heart sound's frequency spectrum reached up to a general maximum of 650 Hz [28]. The pulmonary activity frequency spectrum reached between 600 and 20 000 Hz - leaving an unwanted 150 Hz frequency overlap if a 800 Hz LPF is used for the heart sounds. Section 3.3.1.3 discusses a second low-pass filter to circumvent this problem. After the waveform was filtered it was automatically normalized with respect to the maximum amplitude equal to 0.8.

After scanning all the recordings, the patients, their personal relevant information and their associated recordings were added to the database. All recordings have their accompanied LPF critical frequency, position of recording, intensity grading of murmur, and diagnosis added to the database.

3.3 Pre-processing of heart sounds

3.3.1 Filtering and De-noising - ECG and Heart Sounds

3.3.1.1 Noise characteristics and their implication

During the recording of the ECG and heart sound data, contamination may come from any of a variety of sources. Some of this noise is of human origin for example abdominal, or respiratory noise, stethoscope diaphragm friction, movement artifact, muscle activity, speech from the surrounding environment, patient movement or the patient sucking on a dummy. A recording is also subjected to background noise originating from mains interference and the internally generated electronic noise of the recording equipment. These noises can contaminate the heart sound signal to such an extent that the signal is unsuitable for further processing for the purpose of visualization or diagnosis. A short discussion follows describing the different methods used to remove the above-mentioned noises. During the development of a noise removal technique it is important to ensure that the technique does not dispose of any information bearing data.

Before removing any noise, it must be classified as stationary or non-stationary noise. Noise from a specific source $n(t)$ is called stationary noise if the energy mean and variance of $n(t)$ are constant at all instances of time. If the energy mean and variance of $n(t)$ are not constant at all instances in time, the noise is said to be non-stationary [42]. To be technically correct the noise contaminating the heart sound signal is a combination of stationary and non-stationary noise. The first set of noise sources mentioned in the previous paragraph are classified as sources generating non-stationary noise, and the second set of mentioned sources are classified as sources generating stationary noise.

The problem with the de-noising of heart sounds is that the noise frequency range and the information bearing frequency range overlap, thus ruling out the use of bandstop or bandpass filters. A LPF was already used to do away with noise above 800 Hz.

3.3.1.2 Noise reduction through correct recording techniques

The electronic and ambient noise can be minimized by using the correct recording techniques, and by taking some characteristics of the recording equipment into account. The *Meditron ECG Analyzer* were tested with three different computers. The output when using a personal desktop computer contained an unwanted 50 Hz mains noise which was too large; this setup was accordingly not used. The second computer tested was a *Compaq* notebook with its power supply and powering unit at the same side as the USB input port. This layout architecture contributed to an excessive electronic noise level in the recordings. The third computer tested was an *Acer Travelmate* with the power supply and powering unit on the opposite side of the back panel as the USB input port. The *Acer Travelmate* performed considerably better than the first two computers tested. Accordingly the *Travelmate* was used for all the recordings.

The *Meditron* specifications state that the Signal-to-Noise (SNR) ratio will be better if the notebook is battery powered than when powered from the 220V mains. Figure 3.2, however, show the difference between the two powering methods, with the mains powered setup performing better when using the *Acer Travelmate*. Volume settings on the distribution unit do not make a measurable difference to the SNR because all the recordings are normalized. The stethoscope's head amplifier volume must be tuned to medium. Maximum volume settings contribute to a diaphragm that is over sensitive for external noise and friction.

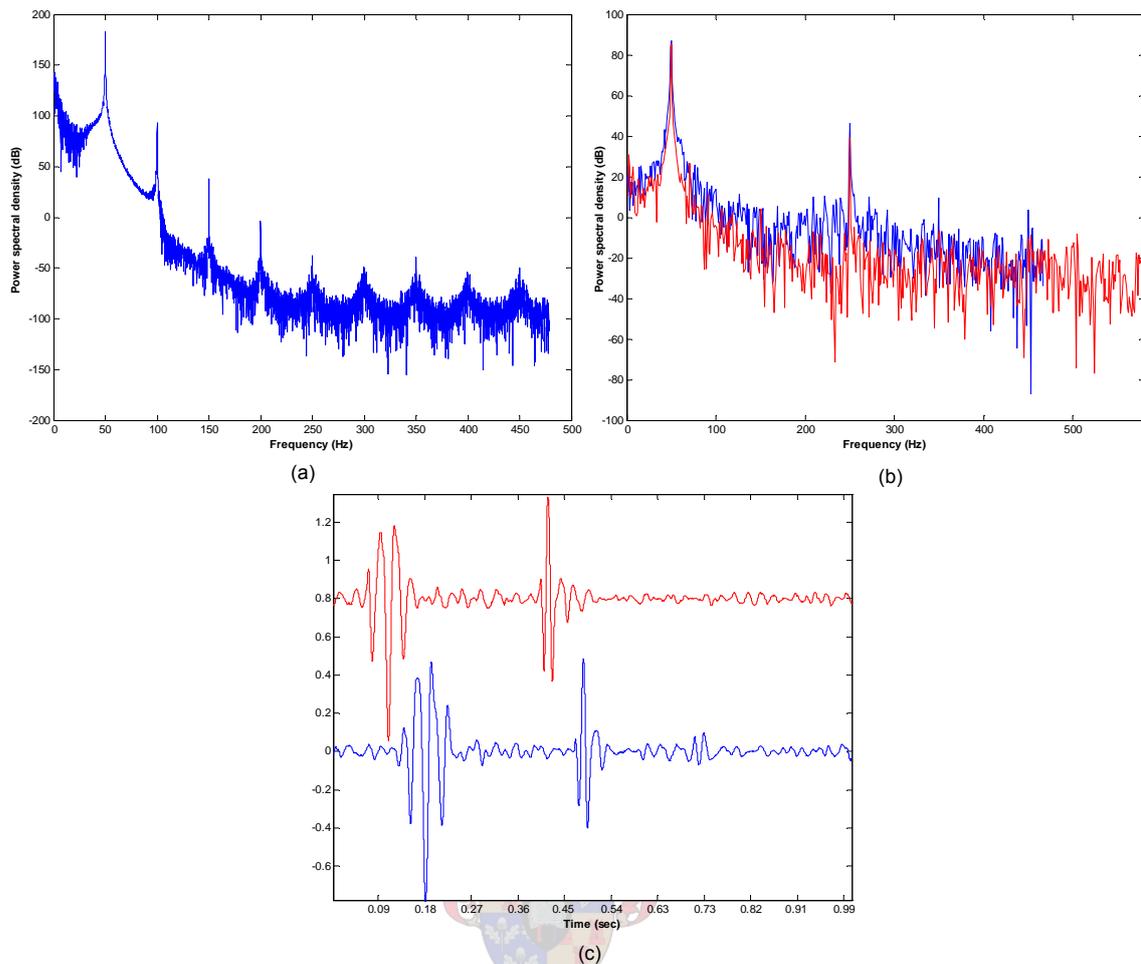


Figure 3.2: (a) Power Spectral density of stethoscope pickup in a noise-proof room, the 50 Hz mains harmonics is clearly visible. (b) & (c) show the noise differences between battery powered (red) and mains powered (blue) recordings.

End results of all the test setup recordings showed that there is no major difference between different settings (including the settings in the recording software). What is more important, is a quiet recording environment, a good firmly pressed secure recording (stable hands), correct placement of stethoscope head and a co-operative patient. [External speech (including speech in the passageway next to the evaluation room) is easily picked up during recordings]. Applying ultrasonic gel on the stethoscope head reduces the noise originating from skin contact friction remarkably which results in a better quality recording. Gel was applied if necessary.

3.3.1.3 Digital filtering of heart sound and ECG waveforms

The digital filter implemented for the ECG signal is to remove any baseline (iso-electric) drift that might influence the later discussed QRS peak-detection algorithm. According to Burk's formula for ECG constituent timing characteristics [43], the QRS-complex duration has a maximum duration of 0.12 seconds, which corresponds to 8.33 Hz in the frequency domain. The filter must thus filter out the baseline drift without affecting any data above 8.33 Hz. The developed low-pass filter has the following characteristics:

Type: Infinite impulse response 4rd order Butterworth filter

Structure: Direct Form II, 2 second-order sections

Fstop, Astop: 1 Hz, -30 dB

Fpass, Apass: 5 Hz, 1 dB

The frequency magnitude response is shown in Figure 3.3(a) and an example in Figure 3.4.

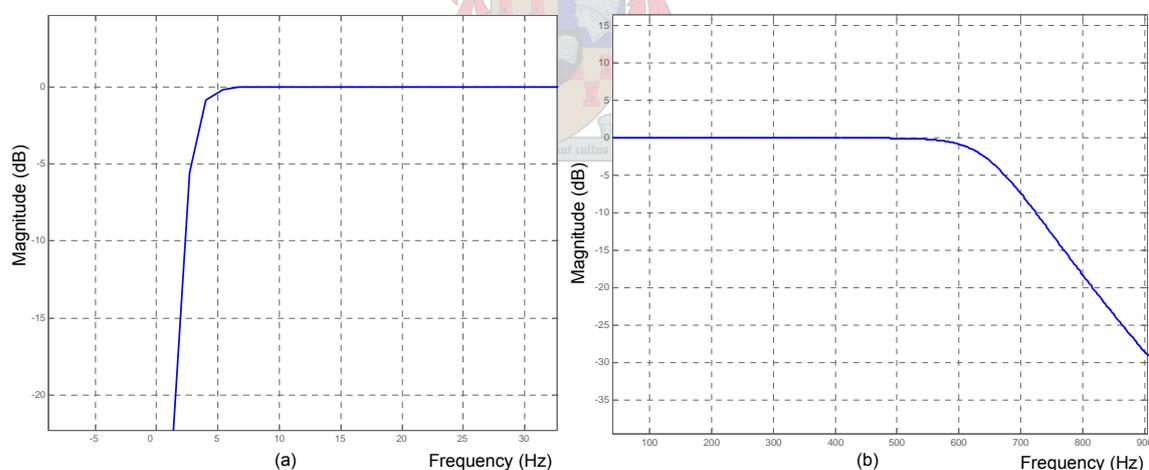


Figure 3.3: (a) High-pass filter for ECG and (b) low pass filter for the heart sound data

In section 2.2, the frequency extend of heart sounds are given as 20 to 650 Hz. The bell mode filter characteristics of the *Meditron* stethoscope are given as 20 to 420, but output signals contain frequencies of up to 1200 Hz. After completing the filtering procedure, described in section 3.2, the heart sounds are confined to the 20 to 800 Hz frequency band. Investigation of several noise free, normal and pathological, heart sound recordings showed that the average maximum frequency is in the range of

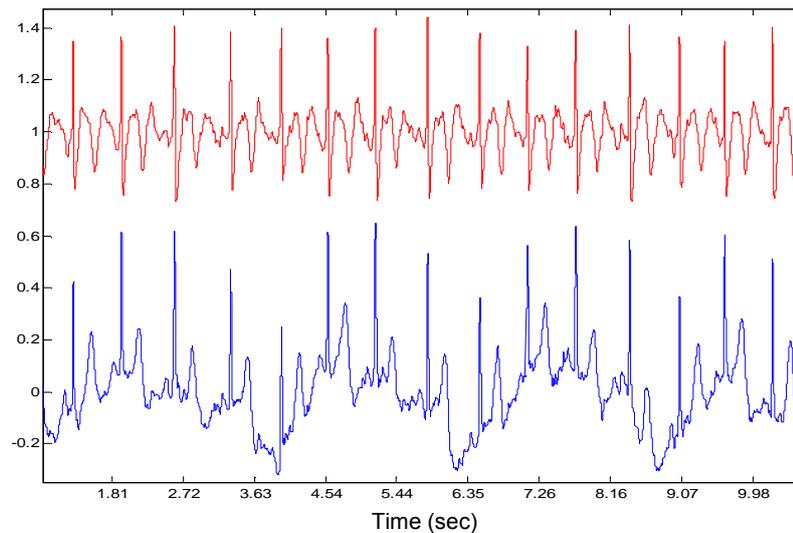


Figure 3.4: Original ECG signal with unstable iso-electric line in blue and de-noised ECG signal in red

650 Hz. To filter the 20 to 800 Hz heart sound waveform to the wanted 20 to 650 Hz frequency band, a digital LPF designed with *Matlab's* filter toolbox is used. The developed low-pass has the following characteristics:

Type: Infinite impulse response 10^{th} order Butterworth filter

Structure: Direct Form II, 5 second-order sections

Critical frequency (-3 dB): 650 Hz

The frequency magnitude response is shown in Figure 3.3(b), and an example in Figure 3.5.

3.3.1.4 Fixed threshold wavelet de-noising

Research done by L.T. Hall *et al* [44] show that the background noise in heart sound recordings can be dramatically reduced by a thresholding operation in the wavelet domain. The principle on which wavelet de-noising is based is that the background noise is confined to many small broadband wavelet coefficients that can be removed without significant degradation of the signal of interest. Appendix C contains a complete background study on the wavelet transform, the derived formulas will be used in the next explanation. Hall suggested to use the *Daubechies Wavelet* of order 5 (db5) for analyzing heart sounds. The properties of the db5 wavelet are shown in Figure 3.6. The choice is due to the heart beat signal having most of its energy distributed over a small number of db5 wavelet dimensions (scales), and therefore

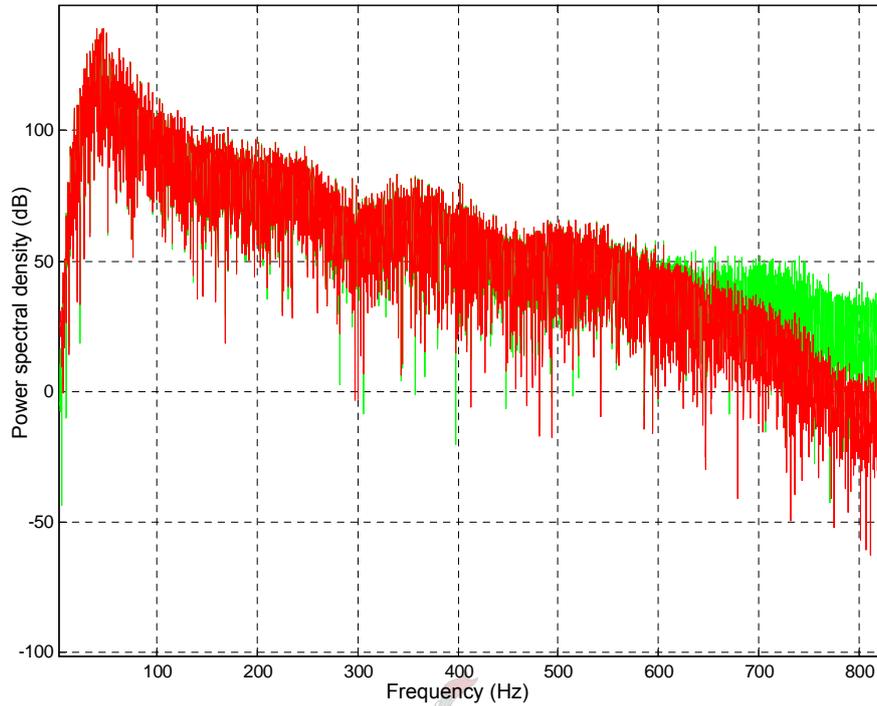


Figure 3.5: Original periodogram in green and filtered ($f_c = 650\text{Hz}$) periodogram in red

the coefficients corresponding to the heart beat signal will be large compared to any other noisy signal. The fixed threshold wavelet de-noising procedure involves three steps [45]:

(i) Decompose

Through the implementation of multi-resolution analysis, as demonstrated by Vetterli [46], the heart sound is divided into approximations and details. Where the approximations represent the slowly changing (low frequency - high scale) features of a signal and the details represent the rapidly changing (high frequency - low scale) features of the signal. Figure 3.7 shows an example decomposition, with V_x representing the approximations and O_x representing the details of the represented signal. Since the analysis process is iterative, in theory it can be continued indefinitely. In reality, the decomposition can proceed only until the individual details consist of a single sample or pixel. The sum of the final approximation and all the details, yields back the original signal, and is called recomposition. The procedure used by *MatLab* to achieve this decomposition and recomposition of a signal involves the process of applying numerous high-pass and low-pass finite impulse response (FIR) filters in succession. A decomposition level of 8 with the db5 wavelet was selected for the decomposition part of the de-noising algorithm.

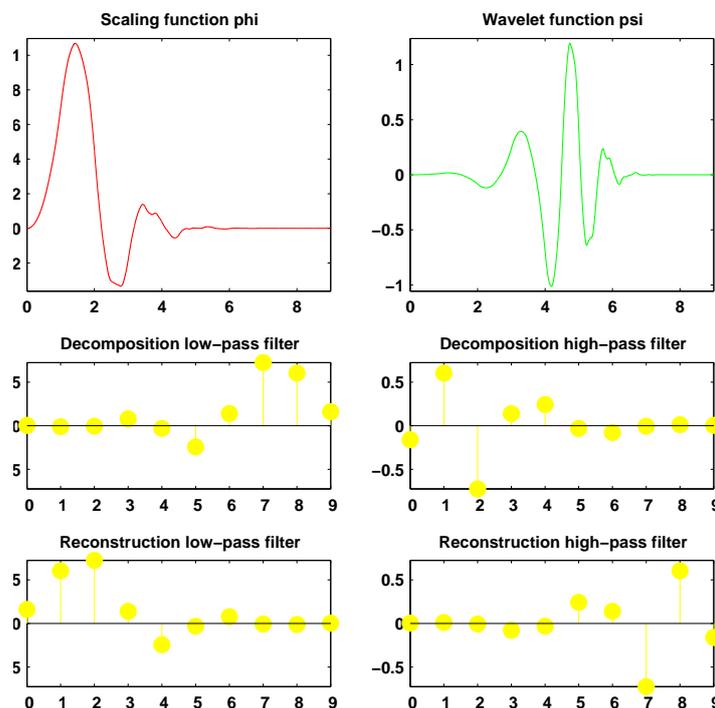


Figure 3.6: Daubechies Wavelet of order 5 and associated de- & recomposition filter coefficients

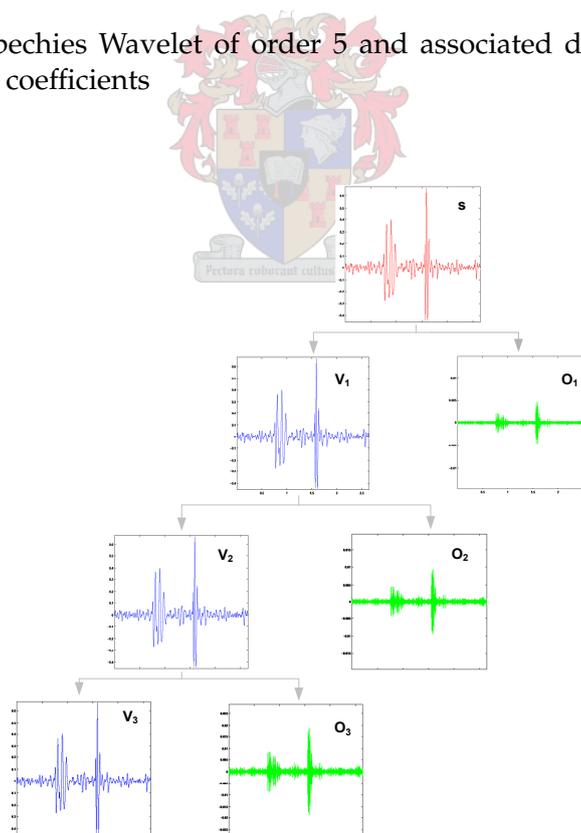


Figure 3.7: This wavelet decomposition tree show approximation (V) and detail (O) spaces of 3-levels. With recomposition it is shown that $s = V_3 + O_1 + O_2 + O_3$

(ii) Threshold detail coefficients

The threshold operation involves removing coefficients, from the various detail levels, which lie below a specified value. This method uses the property of wavelets to expose sharp discontinuities in a signal, meaning that noise can be readily revealed and hence can be removed by thresholding certain components of the wavelet decomposition. This is a very powerful concept because signals, with energy concentrated in a small number of wavelet dimensions, will have coefficients that are relatively large compared to any other signal, which has its energy spread over a larger number of wavelet dimensions, present [44]. Therefore applying the thresholding operation to the decomposed coefficients will effectively remove any unwanted signal or noise, even if the instantaneous frequency spectra of the two signals overlap.

Determining the threshold level for each decomposition level is done by trying to meet two criteria - (i) to remove as much of the noise as possible; (ii) without losing any information. This is equal to distinguish between what is noise and what is information. Because there is no practical way to determine the actual source, this vital decision is to be made on the *trial-and-error* method. Several recordings were made in a sound-proof test room, some measuring only ambient noise and others measuring a combination of ambient noise and heart sounds. Setting the threshold value according to visual interpretation of what is noise and what is information, yield a fixed setting for each recording. A comparison between all the recordings' threshold levels, showed a very high conformity between the recordings. This de-noising mechanism produces quite a measurable improvement in the signal quality. Figure 3.8 show the fixed threshold levels applied to the various decomposition levels.

In level O_8 and O_7 a low threshold is applied to retain the general shape of the signal whilst removing any unwanted low amplitude noise components. Levels O_6 to O_3 have thresholds just above the point where the heartbeat signal appears to protrude through the noise. It is at these scales that the shape of the applied wavelet corresponds very closely to the shape of the heartbeat signal, and therefore it is quite easy to distinguish between the information and the noise. All correlated levels are filtered out in levels O_2 and O_1 , arguing that this is stationary noise. The *Matlab* programming environment allow these threshold values to be programmed into a fixed threshold de-noising algorithm. This functionality was used accordingly.

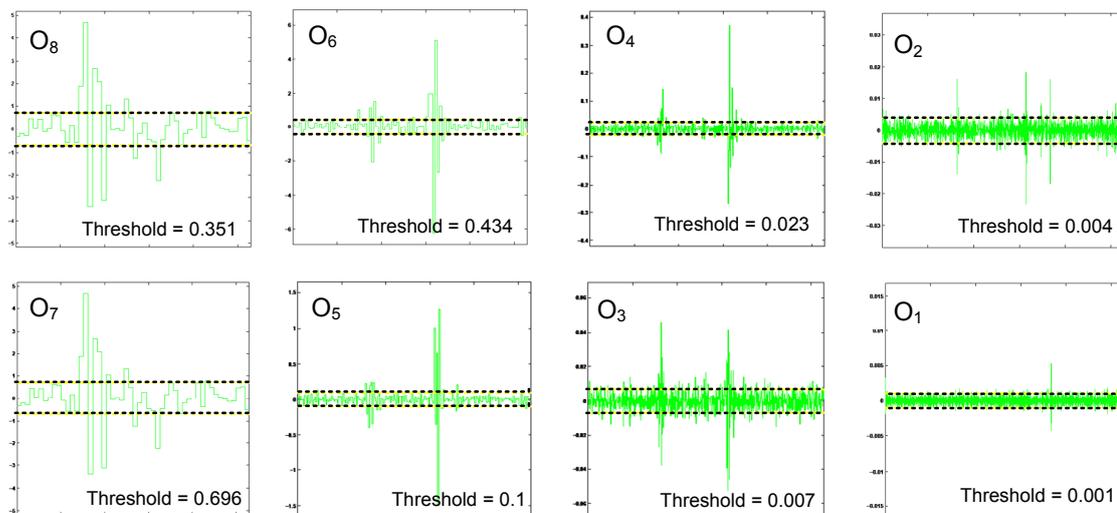


Figure 3.8: Threshold values for different decomposition levels

(iii) Reconstruct

The last step in the de-noising procedure is to compute the wavelet reconstruction through the summation of the original approximation coefficients of the last level (N) and the modified detail coefficients of levels 1 to N. Figure 3.9 shows the result of the fixed threshold wavelet de-noising algorithm on two signals.

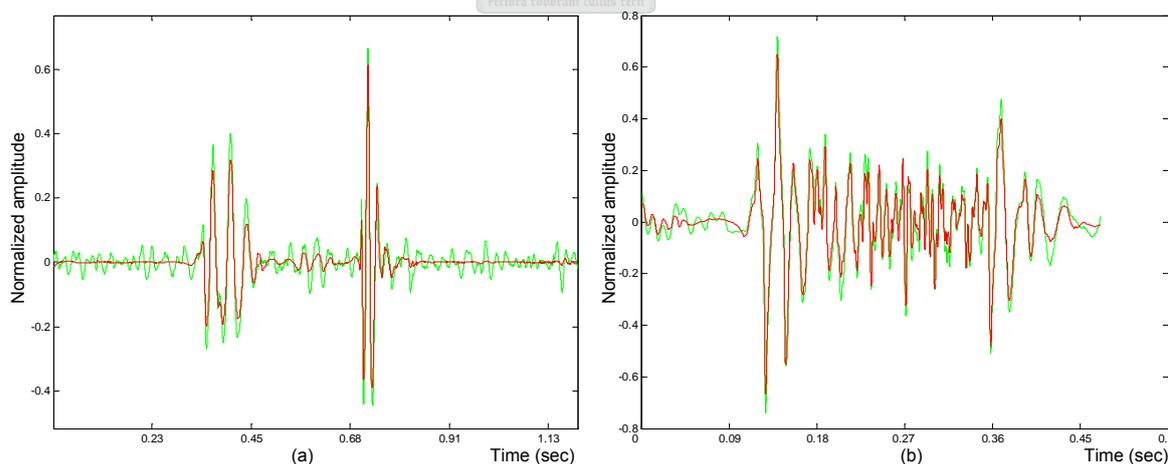


Figure 3.9: (a) Normal period and (b) VSD period. Original signal is illustrated in green and the de-noised signal in red

As illustrated in Figure 3.1 the de-noised function is not always part of the pre-processing calculations. During the discussion in Chapter 4, it will always be clearly

stated whether the de-noising block was included in the pre-processing calculations or not. From the discussion in Chapter 4 it will also be evident whether the fixed threshold de-noising algorithm discards any information bearing data or not.

3.3.2 Segmentation of recording into separate heart beats

This section describes the calculation of heart rate and the separation of heart cycles utilizing the timing relationships between the ECG and heart sound waveform described in 2.1.3. To apply feature extraction algorithms (FEA) on the recording, the FEA's need the recording to be input as separate periods to identify certain time dependent features.

The algorithm that automatically separates the heart cycles (periods) is shown in Figure 3.11. Figure 3.10 shows how the heart cycle duration is calculated using the autocorrelation property of the ECG waveform. If a recording's period length is consistent to a certain extent throughout the whole recording, the calculation of the heart rate is done automatically by the algorithm. If the difference between the inhalation- and expiration section's times, within the same recording, is too big, the heart rate calculation must be done manually, by clicking on the autocorrelation graph in the *Matlab* environment. If the heart rate of a recording is abnormally fast or slow according to Table 3.1 the user is also prompted to do the heart rate calculation manually. This is to ensure that the right heart rate is calculated in the case of a iso-electric unstable ECG recording.

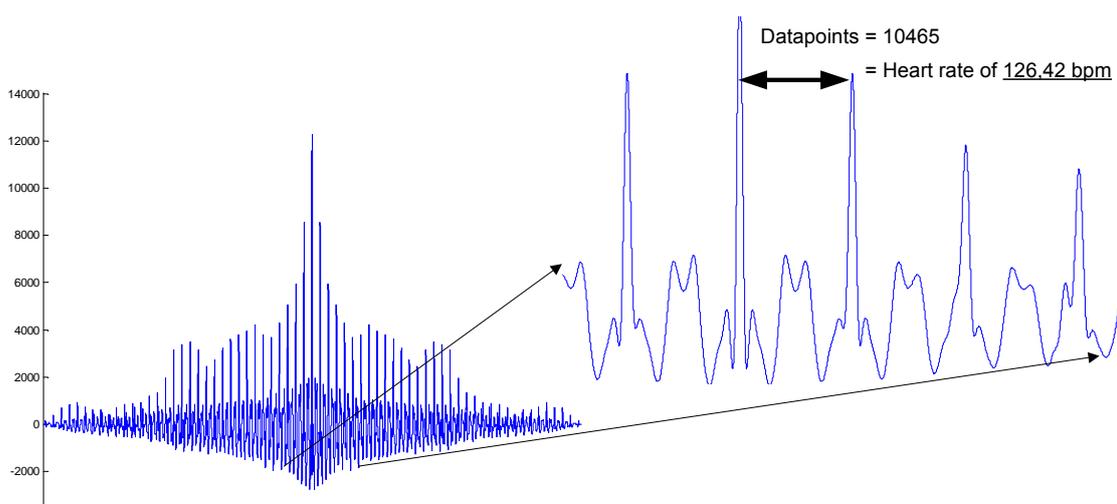


Figure 3.10: Autocorrelation of the ECG waveform to calculate the heart cycle's duration

	Birth-6 wk	6 wk-2y	2-6y	6-10y	Over 10y
Min. heart rate (bpm)	95	90	80	75	60
Max. heart rate (bpm)	155	140	120	105	100

Table 3.1: Normal values of heart rates in pediatric patients. [wk - week; y=year; bpm = beats per minute]. Data from [5]

The result of the automatic heart sound recording algorithm is shown in Figure 3.12. During the rest of the paper if referred to a heart cycle it will be in this format, with 25% of the cycle's duration preceding the peak R-wave.

3.3.3 Period filtering

After the segmentation of the recording into separate periods (heart beats), all periods are to be filtered until there is an acceptable consistency and correlation between all the periods. The process consists of identifying the unwanted (contaminated) periods and then doing away with them. Filtering unwanted periods from a recording is important to withhold unrepresentative data from the feature extraction algorithms (specifically the neural network method). When the network receives a real pathological period, but with the same features as the contaminated normal period; the feature extraction method will classify it falsely as normal. Although only recordings with a minimum artifact and high signal quality were used, anomalies do occur regularly in most recordings.

The algorithm flow description shown in Figure 3.15, illustrate the methods used to automatically remove the unwanted periods form a vector of periods.

The information used, to do the correlation calculation between the periods, is the average frequency spectrum of 12 frequency bins (between 20 – 420Hz), for 10 consecutive and equal in size time-intervals per period. The result is a 120 datapoint representation for each period of the recording. A mathematical technique called the Windowed Discrete Fourier Transform (WDFT) is applied in order to discover what frequencies are present at any given moment in the period. The result of Fourier analysis is a spectrum as shown in Figure 3.13 which contains an estimate of the short-term, time-localized frequency content of the signal. After computing the spectrum for one short section or window with a size of one tenth of the periods length, the spectrum for the adjoining window is computed: it is continued in this way until the end of the period.

After computing the spectrum new frequency bins are located using the *Mel-scale*

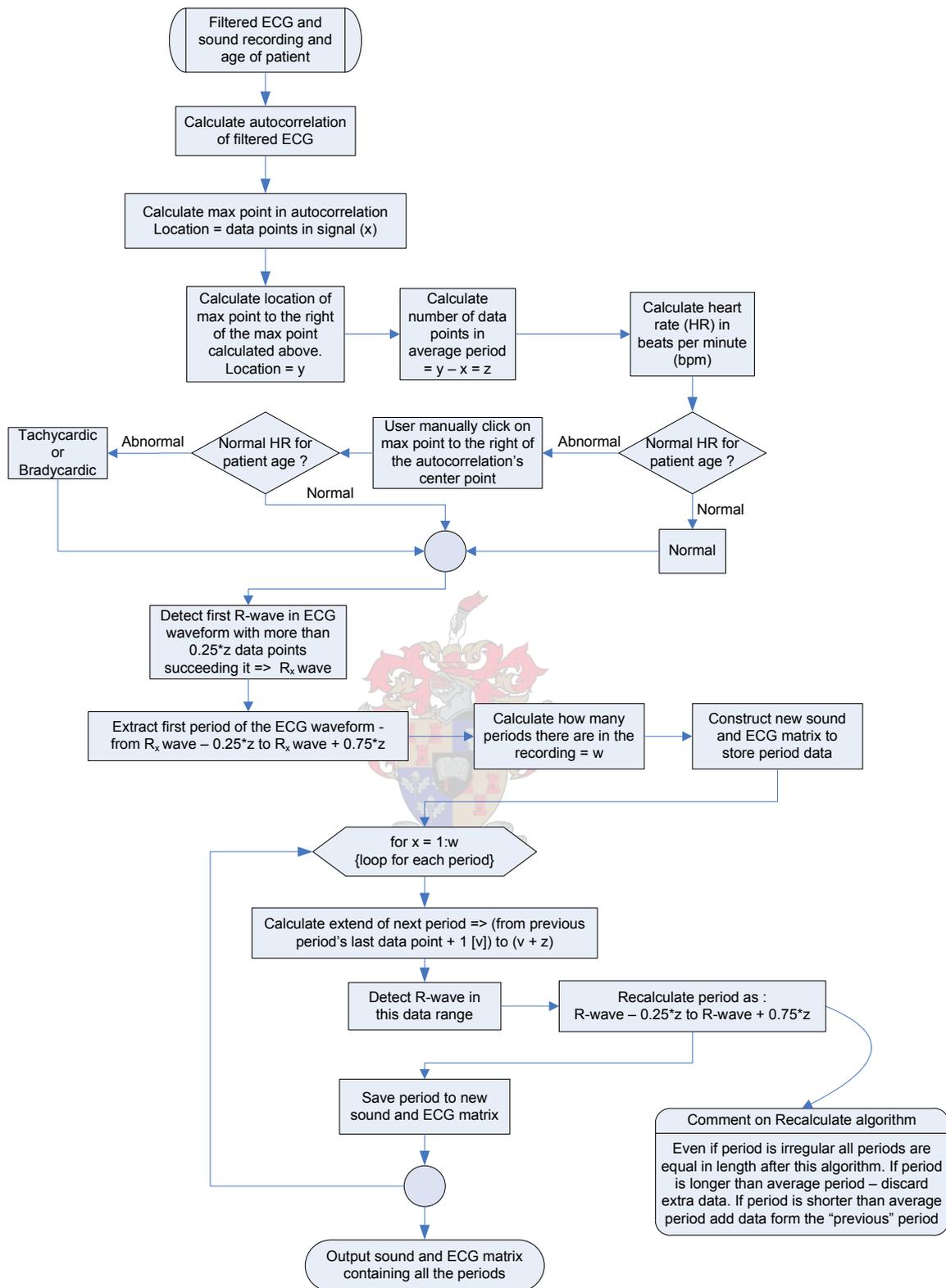


Figure 3.11: Flow diagram illustrating the segmentation of a heart sound into separate beats (periods). The values shown in Table 3.1 are used to classify the heart rate as normal or abnormal. The program code is available on the accompanied compact disc. The code is listed as `Period_Calculator.m` and all `Period_Calculator's` offspring files shown in Figure E.1 in Appendix E.1.

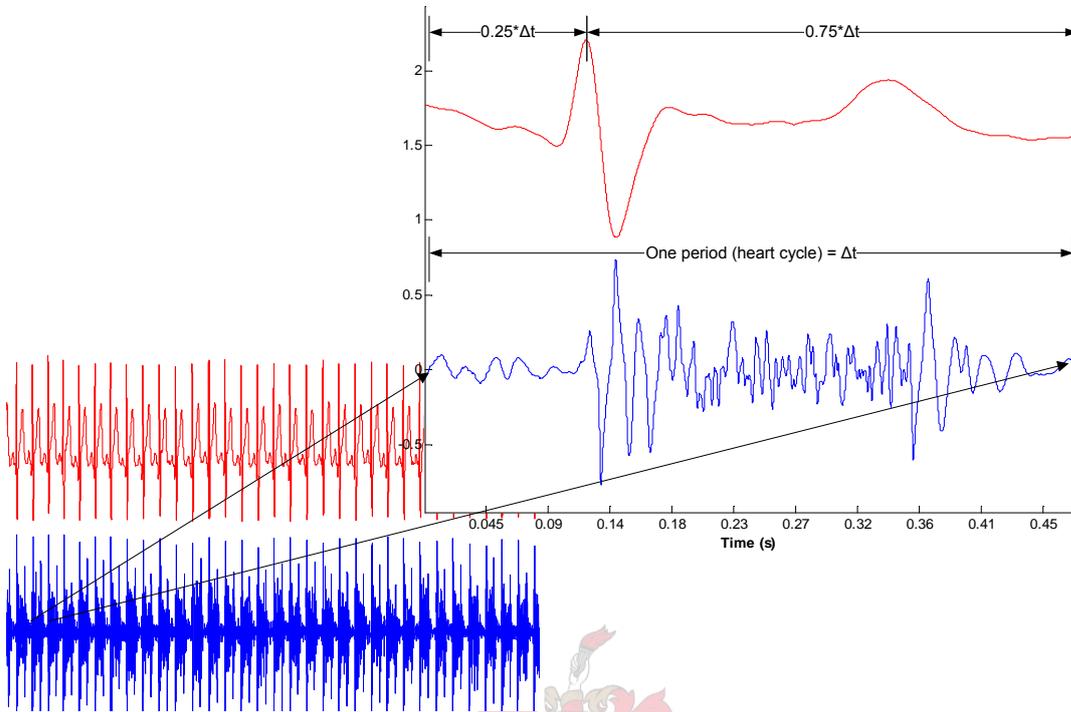


Figure 3.12: Result of the automatic heart cycle segmentation algorithm. All 31 cycles in this recording are extracted and copied to a ECG- and sound data matrix. The recording is represented by these two matrices in the following algorithms

Filter Banks which consist of triangular filters spaced on a linear-logarithm scale over the frequency-axis. The spacing of filters is inspired by critical frequency band measurements of the human auditory system, yet the bandwidth of each filter is chosen by aligning the triangle base with the center frequency of the neighboring filters [47, 4]. A person's auditory mechanism can distinguish an uncorrelated period from a number of periods. This algorithm aspires to mimic the human auditory system by using the Mel-Scale Frequency technique. The filter-bin's center frequencies are calculated in the following way [4]:

$$Mel(f) = 2595 \cdot \log\left(1 + \frac{f}{700}\right) \quad (3.3.1)$$

for 12 filter banks between 0 – 420 Hz \Rightarrow

$$\frac{Mel(420)}{12} = \frac{529.69}{12} = \underline{44.14Hz} \quad (3.3.2)$$

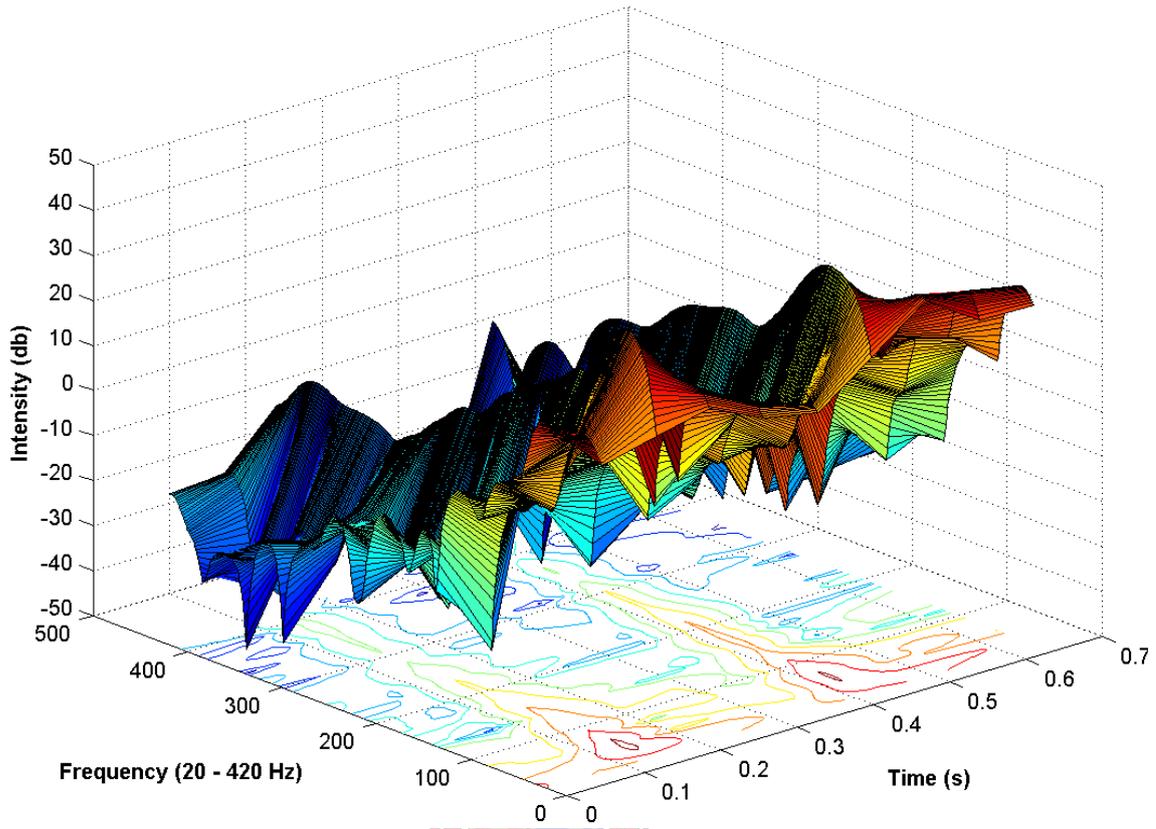


Figure 3.13: Spectrogram of one normal heart sound period. A spectrogram is created by displaying all of the spectra computed from the heart sound period together. The lines visible on the spectrogram each represent 1 Hz along the frequency-axis, and one tenth of the total time along the time-axis. A contour plot is shown beneath the surface, on the xy-plane.

from equation 3.3.1 and 3.3.2:

$$f = 700 \left[10^{\frac{n44.14}{2595}} - 1 \right] \quad (3.3.3)$$

for $n=1:12$ the 12 frequency bins centre frequencies are shown in Figure 3.3.3

Removing the unrepresentative periods from the vector require certain thresholds to be set. These thresholds were tuned using the following trial-and-error method - (1) Listen and look at the recorded signal. (2) Identify periods that does not fit into the general signal characteristics. (3) Put the signal through the above algorithm and check whether it removes the unwanted periods. (4) The threshold values were then tuned according to the outcome of this process until it gave satisfactory results with numerous recordings.

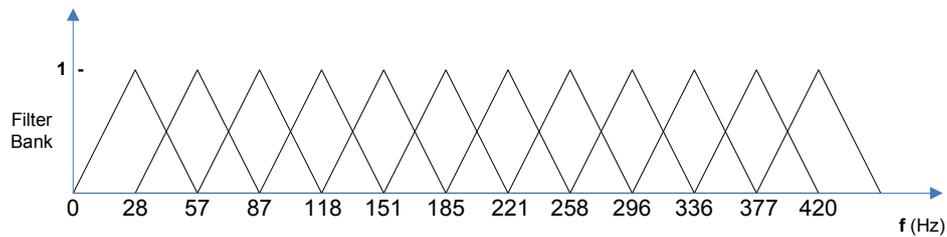


Figure 3.14: Mel-scale filter banks for 12 bins between 20-420 Hz

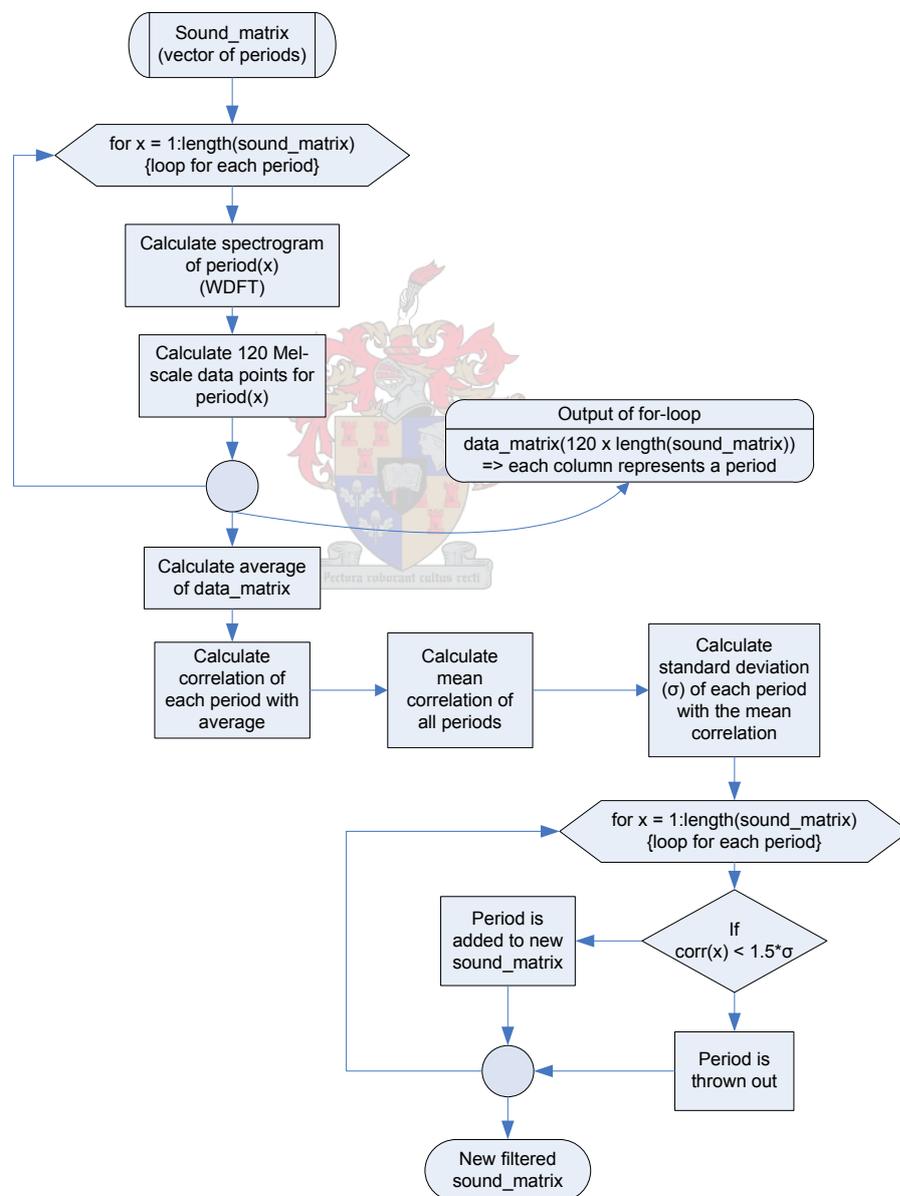


Figure 3.15: Flow diagram describing the automatic period filtering algorithm. The period filtering algorithm's code is listed in Appendix E.1

3.4 Feature extraction and recognition

After the pre-processing the recordings are ready to be analyzed and classified by the feature extraction algorithms. Each of the three methods utilize a unique signal processing technique to analyze a certain characteristic of the heart sound for classification purposes. All the employed processing techniques are discussed together with the way they are utilized to do the necessary feature extraction. All three the algorithms receive the output of the period filtering algorithm as input.

3.4.1 The Direct_Ratio method

The first method developed, extracts the time dependent energy content to serve as a indicator of pathology. This method thus examines whether the hypothesis made that pathological pre-cordial murmurs possess a higher intensity auscultatory sounds than innocent and no-murmur sounds are true. The following subsections describe the various procedures developed to automatically calculate a figure that represents the time dependent energy content of a recording.

3.4.1.1 Heart cycle constituent segmentation

To automatically calculate the intensity or loudness of a murmur we first need to calculate the timing relation of the murmur. The relative position within the cardiac cycle, in relation to S1 and S2, is important in describing and determining the characteristics of a heart murmur. The following two subsections will describe the methods developed to calculate the timing relationships of the constituent components of the heart sound.

After separating the periods of the heart sound recording using the synchronous ECG recording, the ECG can also be utilized to calculate the time relationships of the constituent components of the heart sound. The constituent components of the ECG are shown in Figure 3.16. This figure also shows the timing correlation between the ECG signal and the heart sound recording. These relations were already discussed in section 2.1.3. The described timing characteristics will be elicited to calculate the various timing components to be used to do the necessary information extraction.

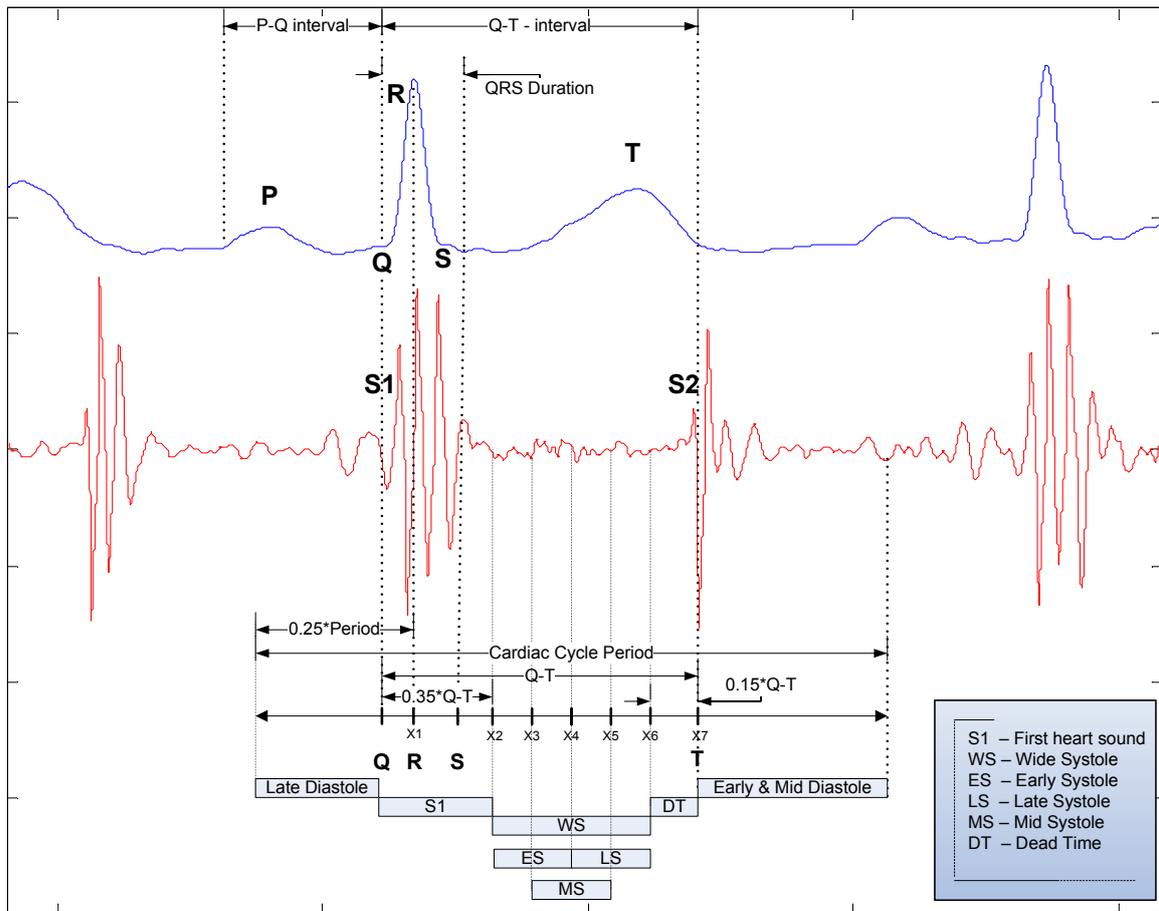


Figure 3.16: Heart cycle constituent components

3.4.1.2 The time relationships of the heart cycle constituent components

There appears to be very little information in literature concerning the time relationships of the individual components which make up the heart beat. A research group at Trinity College, Dublin, Ireland examined the timing relationships of the principle constituent components of the human ECG [43]. Using wavelet transform methods they located the positions of the onset, peak, termination and duration of the individual components of the ECG. Component times were then classified according to the heart rate associated with the cardiac cycle to which the component belonged. Second-order equations were fitted to the data to characterize its timing variation. The success of this combination is evaluated by the mean square error and the coefficients of multiple determination of the fit. The second statistic measures how successful the fit is in explaining the variation of the data. Equation 3.4.1 and 3.4.2 are used to calculate this statistic.

$$MSE = \frac{1}{P} \sum_{x=1}^P e(x)^2 = \frac{1}{P} \sum_{k=1}^P (r[x] - t[x])^2 \tag{3.4.1}$$

with $r(x)$ the actual value, $t(x)$ the estimated value, $e(x)$ the error and P the number of data points.

$$r = \frac{\sum_{x=1}^P (\hat{r} - r[x]) (\hat{t} - t[x])}{\sqrt{\sum_{x=1}^P (\hat{r} - r[x])^2 \sum_{x=1}^P (\hat{t} - t[x])^2}} \tag{3.4.2}$$

where \hat{r} and \hat{t} are the means of the actual values and the estimated values. The coefficient of determination equals r^2 .

Figure 3.17 describes the variation in the duration of the components as the cardiac cycle time altered inversely to the heart rate. It can be seen that the data obtained for the male and female subjects differs significantly, resulting in a separate equation for each sex.

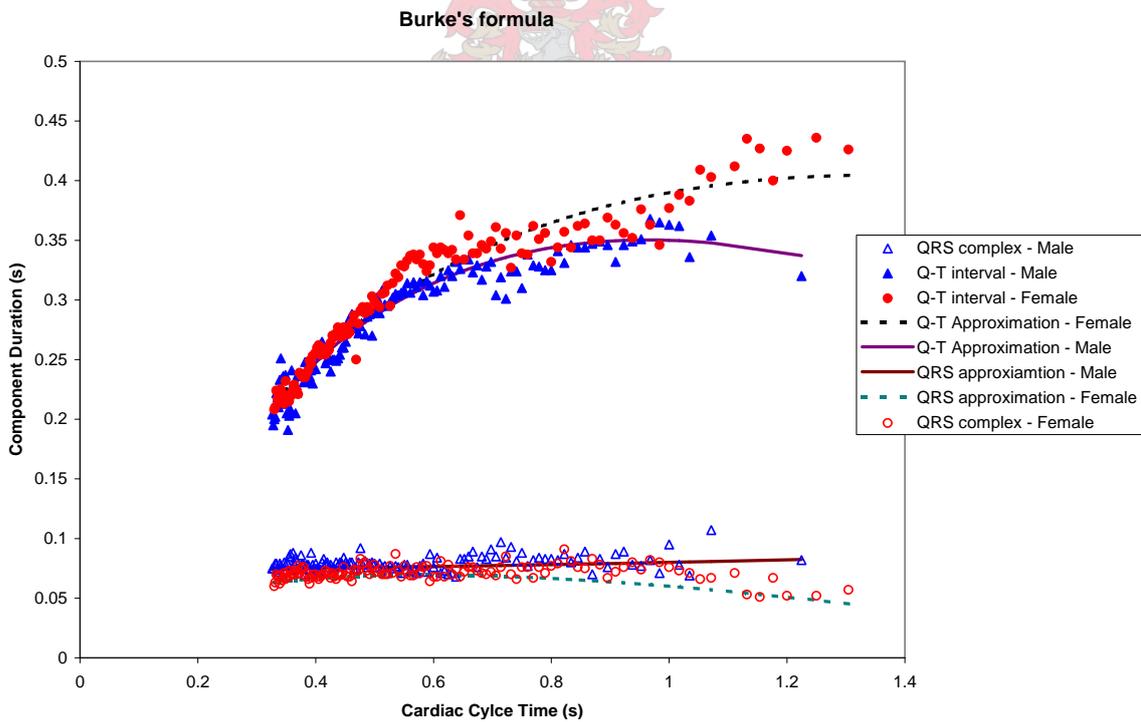


Figure 3.17: Burke's second-order characteristic equations for the Q-T interval and the QRS complex for male and female patients. (Patient data courtesy of M.J. Burke and M. Nasor, Department of Electronic Engineering, Trinity College, Dublin 2, Republic of Ireland)

The second-order equations derived for the two wanted components (the Q-T interval and the QRS complex) were as follows, with their accompanying mean square error (MSE) values and coefficients of multiple determination r^2 :

Burke formulated the equations characterizing the duration of the two wanted components as a function of the cardiac cycle.

Male subjects:

$$T_{qt} = 1.65T_{R-R}^{\frac{1}{2}} - 0.84T_{R-R} - 0.46s \quad (3.4.3)$$

$$MSE = 1.38x10^{-4}$$

$$r^2 = 0.939$$

$$T_{qrs} = -0.02T_{R-R}^{\frac{1}{2}} + 0.02T_{R-R} + 0.08s \quad (3.4.4)$$

$$MSE = 4.198x10^{-5}$$

$$r^2 = 0.102$$

Female subjects:

$$T_{qt} = 1.28T_{R-R}^{\frac{1}{2}} - 0.55T_{R-R} - 0.34s \quad (3.4.5)$$

$$MSE = 2.124x10^{-4}$$

$$r^2 = 0.941$$

$$T_{qrs} = 0.26T_{R-R}^{\frac{1}{2}} - 0.17T_{R-R} - 0.03s \quad (3.4.6)$$

$$MSE = 6.391x10^{-5}$$

$$r^2 = 0.219$$

In order to try and improve the fit of the equations describing the data, a third order equation was fitted to the four data sets. The results are shown in Figure 3.18 and the equations obtained are:

Male subjects:

$$T_{qt} = 0.3309T_{RR}^3 - 1.084T_{RR}^2 + 1.1658T_{RR} - 0.0681 \quad (3.4.7)$$

$$MSE = 1.466x10^{-4}$$

$$r^2 = 0.936487$$

$$T_{qrs} = -0.0509T_{RR}^3 + 0.1111T_{RR}^2 - 0.0654T_{RR} + 0.0894 \quad (3.4.8)$$

$$MSE = 4.35x10^{-5}$$

$$r^2 = 0.1101$$

Female subjects:

$$T_{qt} = 0.7678T_{RR}^3 - 1.991T_{RR}^2 + 1.7956T_{RR} - 0.197 \quad (3.4.9)$$

$$MSE = 1.255 \times 10^{-4}$$

$$r^2 = 0.965$$

$$T_{qrs} = -0.0718T_{RR}^3 + 0.1023T_{RR}^2 - 0.0274T_{RR} + 0.0689 \quad (3.4.10)$$

$$MSE = 2.439 \times 10^{-5}$$

$$r^2 = 0.386$$

The third order equations for the female data sets can be seen to give an improved fit, whereas Burke's formulas show a better fit for the male data. The specific equations will be used accordingly.

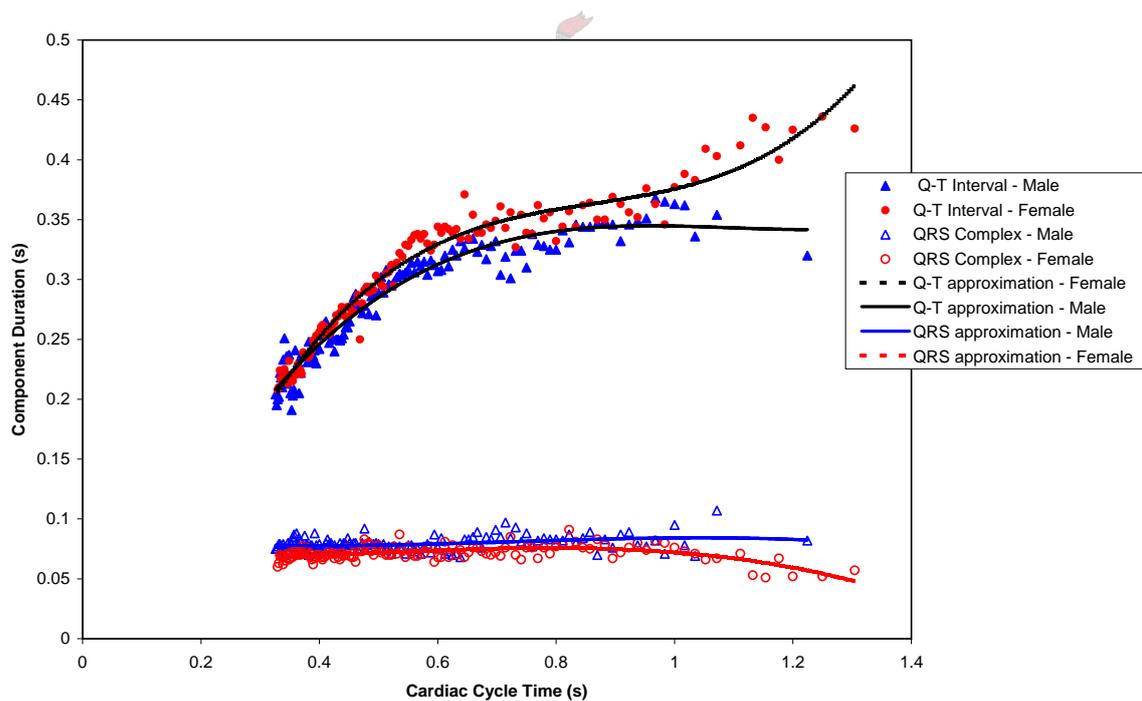


Figure 3.18: Fitted 3rd order equations for the Q-T interval and the QRS complex for male and female patients (Patient data courtesy of M.J. Burke and M. Nasor, Department of Electronic Engineering, Trinity College, Dublin 2, Republic of Ireland)

3.4.1.3 Automatic constituent segmentation

With a single heart beat as input (from the algorithm described in Figure 3.11), the algorithm shown in Figure 3.19 automatically calculates the different constituent components. The output of the algorithm is eight vectors containing the data of the following heart cycle segments of one period:

- A - Late diastolic (LD)
- B - First heart sound (S1)
- C - Wide systole (WS)
- D - Early systole (ES)
- E - Late systole (LS)
- F - Mid systole (MS)
- G - Dead time (DT)
- H - Early and mid diastole (EMD)

The eight segments' positions relatively to S1 and S2 are shown in Figure 3.16.

The different percentage values used in the above algorithm were obtained through extensive trial-and-error testing with the available data. The 35 % of T_{qt} assigned to constituent S1(B) were too long for many cases. This interval was however increased from a tested 25 % to prevent that part of S1(B) be registered as early systolic energy. The next section will describe a method to surmount this problem. Figure 3.20 and 3.21 show the output of the automatic segmentation algorithm for two different cases.

3.4.1.4 Ratio calculation

After running all the pre-processing algorithms described in section 3.3, the algorithm described in Figure 3.22 is used to calculate the ratios of the various systolic constituents' energy content with respect to the energy content of constituent S1(B). The motivation for working with relative energy values is the normalization process performed earlier on each recording. Each recording was normalized with respect to the maximum data point, which in most cases is the first heart sound. Reasoning that the rest of the the recording is normalized relative to the S1 (B) constituent, leaves it to serve as the reference constituent. Limitations of this methodology are discussed at the end of this section. The effect of extraneous noise and artifacts are minimized by using the average energy value for each constituent. This average energy value is derived from all the periods of the recording.

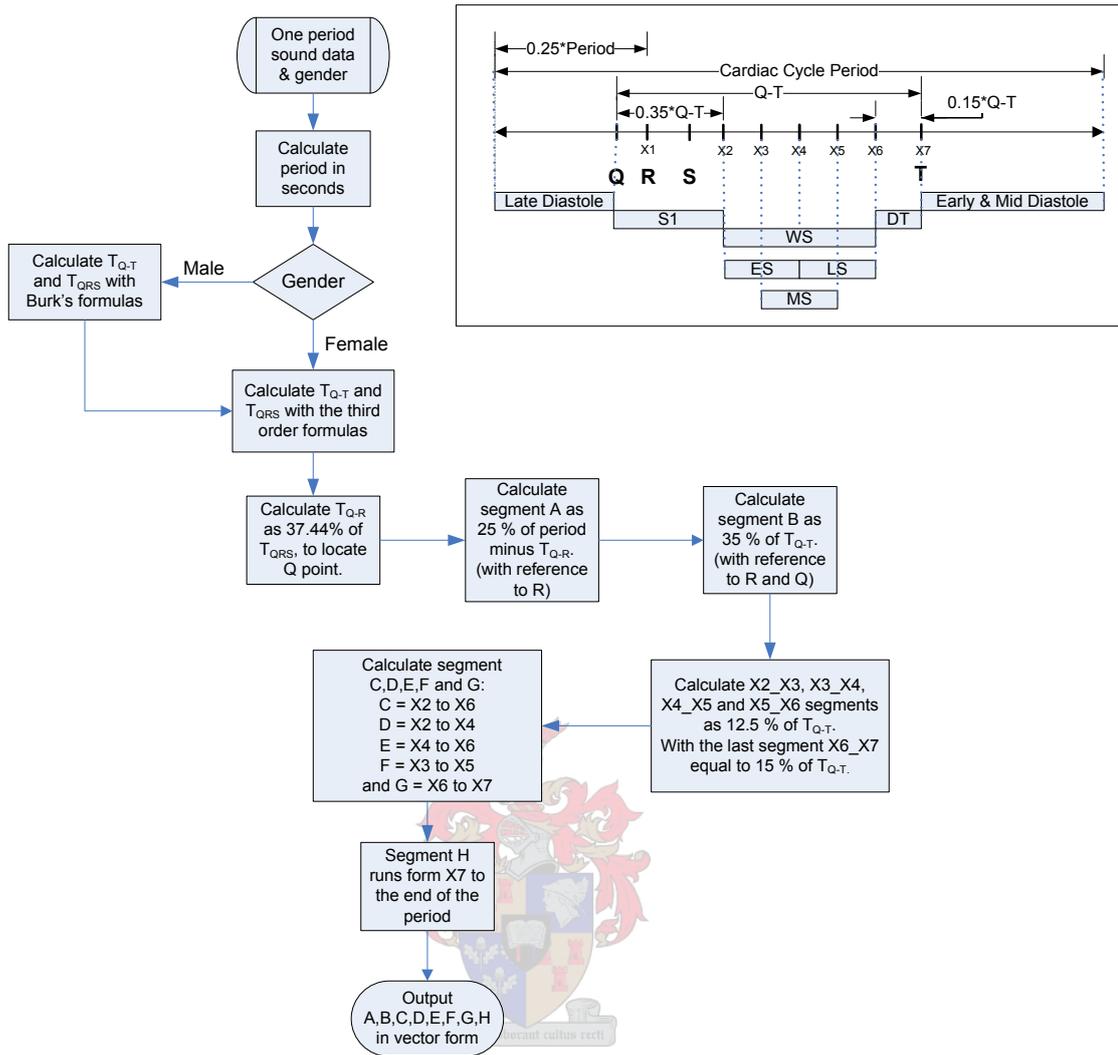


Figure 3.19: Flow diagram describing the automatic segmentation algorithm, with reference to the inset figure (from Figure 3.16). The program code is available on the accompanied compact disc. The code is listed as Segmentation_Ratio.m and all Segmentation_Ratio's offspring files shown in Figure E.1 in Appendix E.1

To overcome the problem that the 35 % of T_{qt} assigned to constituent S1 was too long for many cases, only 80 % of the energy-content of S1 (B) was used to calculate the representative energy content of S1 (B). The reasoning is as follows:

To calculate the energy ratio of one of the systolic constituents, for example mid systolic, the following calculation is made

$$F_{avg} = \frac{\sum_{x=1}^M \left(\frac{\sum_{i=1}^N F_i(x)}{N} \right)}{M} \quad (3.4.11)$$

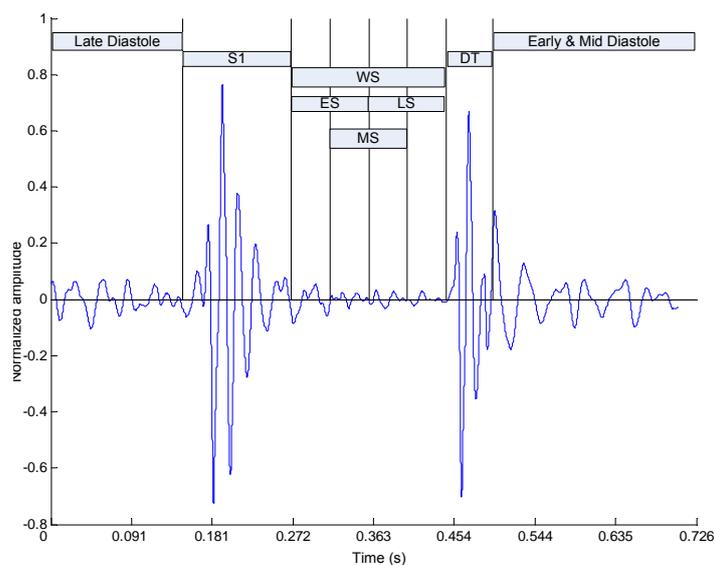


Figure 3.20: Output of the automatic segmentation algorithm for a normal heart sound

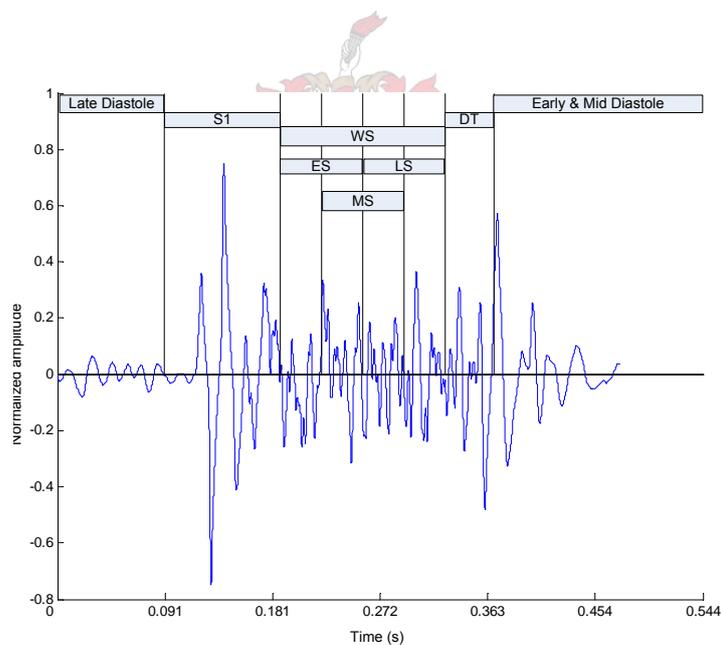


Figure 3.21: Output of the automatic segmentation algorithm for a pathological heart sound (VSD)

where M is the number of data points in constituent F and N is the number of periods in the recording.

$$B_{avg} = \frac{\sum_{x=1}^R \left(\frac{\sum_{i=1}^N B_i(x)}{N} \right)}{R} \tag{3.4.12}$$

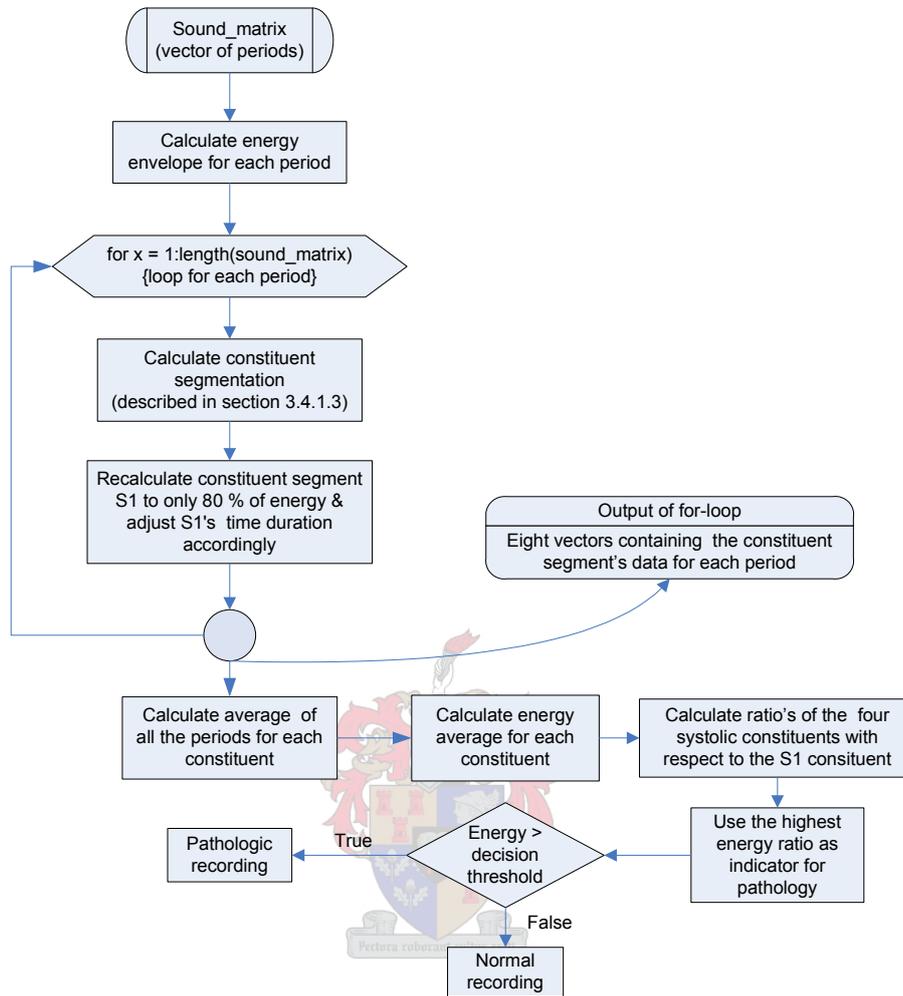


Figure 3.22: Flow diagram description of the Direct Ratio feature extraction method. The Direct Ratio algorithm's code is listed in Appendix E.1

where R is the number of data points in constituent B and N is the number of periods in the recording. Figure 3.24, 3.25 and 3.26 show calculated constituent energy results for three different recordings. The representative ratio of the mid systolic energy content is then calculated as:

$$Ratio_{FB} = \frac{F_{avg}}{B_{avg}} \quad (3.4.13)$$

From equation 3.4.11, 3.4.12 and 3.4.13 it is evident that the magnitude of $Ratio_{FB}$ is inversely proportional to the magnitude of B_{avg} , which in turn is inversely proportional to V and directly proportional to the magnitude of the energy content in constituent B .

If $S1(B)$ is taken as 35% of T_{qt} in all cases, $Ratio_{FB}$ would be unrepresentatively large for recordings with their S1 energy content distributed over a small percentage of constituent S1 (dividing total S1 energy with "dead time"). If the hypothesis is made that the S1 energy of a systolic pathological recording is more widely distributed than the S1 energy of a normal recording, then the methodology, of taking the total energy over the total elapsed time, to calculate B_{avg} is not good practice. Figure 3.24 (top) and 3.25 (top) show the difference between the S1 power distribution of a normal and a pathological recording respectively.

Choosing an energy percentage to represent both normal and pathological cases, is done by calculating the percentage of S1's time responsible for the percentage of S1's energy content. The most favourable is the percentage of energy that results in the largest difference between the calculated ratio of normal vs. pathological cases. For $x > 50\%$; the propability that $x\%$ of the energy is present in $(100 - x)\%$ of time, decreases as x increases. Calculating ratios for several recordings, at different percentages, showed that 80% is the most favourable percentage figure to use. The algorithm flow for recalculating constituent S1(B), is shown in Figure 3.23.

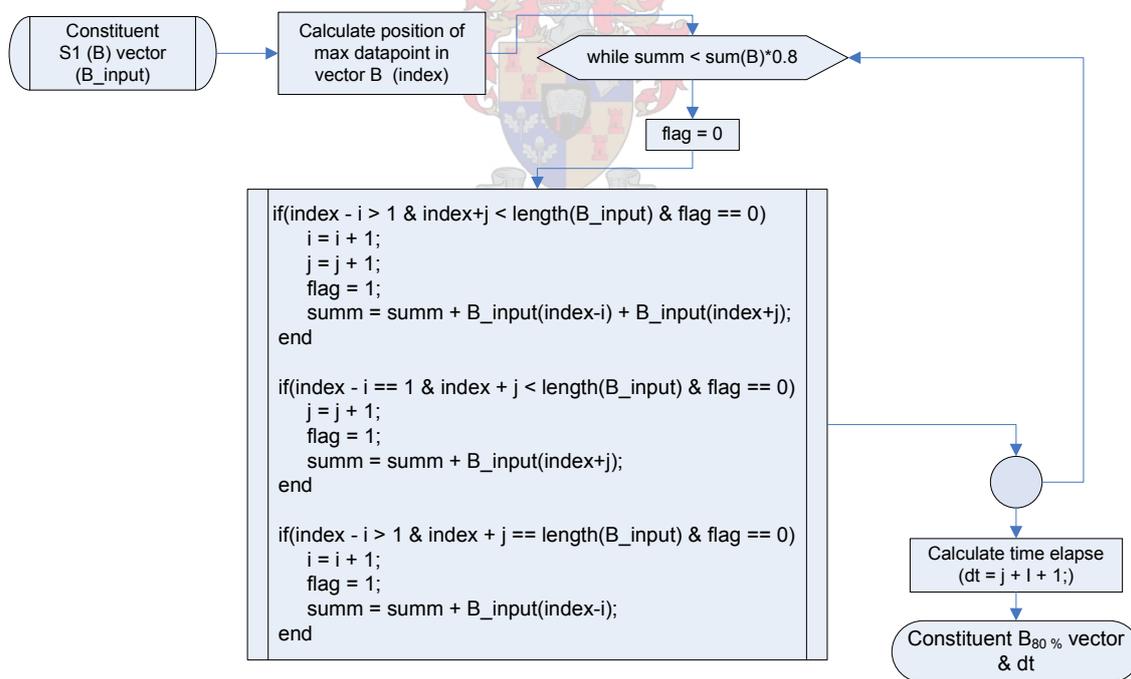


Figure 3.23: Algorithm description for calculating new composition of constituent S1 (B). Program code is on the accompanied compact disc, listed as `Constituent_S1.m`

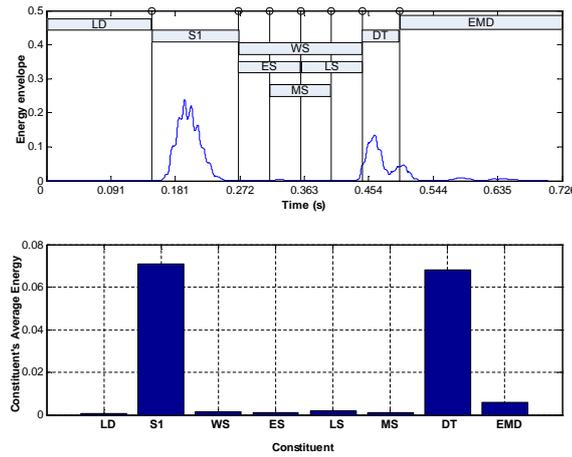


Figure 3.24: Energy content of heart cycle constituents calculated with Direct Ratio - Normal heart sound ($\frac{0}{6}$ rating)

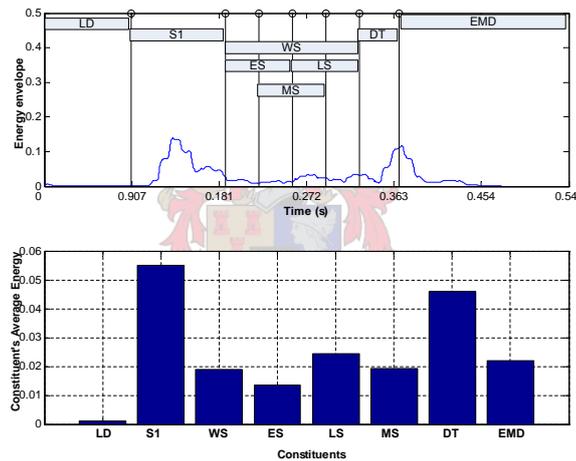


Figure 3.25: Energy content of heart cycle constituents calculated with Direct Ratio - Holosystolic murmur (VSD $\frac{5}{6}$ rating)

To illustrate the improvement made through following this method, the ratios for heart sounds shown in Figure 3.24 and 3.25 are calculated for different percentage values. See Table 3.2 for results. Although 90 % shows a better difference for this specific case, the overall best-performer for all the recordings tested was 80 % and is used accordingly.

After calculating ratios for all four systolic constituents (C, D, E and F) the highest ratio (energy value) is taken to represent the recording. Distinction is made between healthy and pathological hearts on the basis of the energy value being above or below a threshold. Chapter 4.2.1 will discuss performance (statistics) obtained across a span of threshold values.

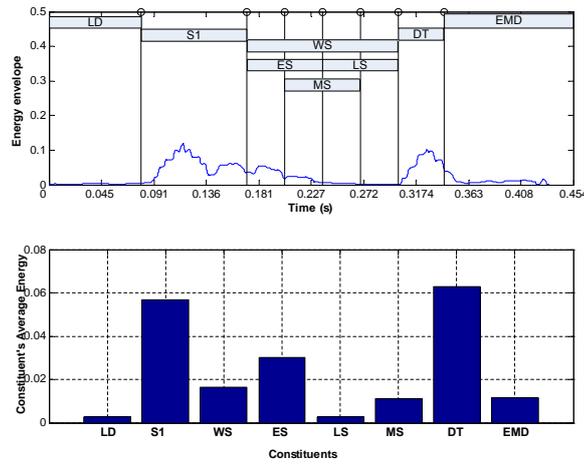


Figure 3.26: Energy content of heart cycle constituents calculated with Direct Ratio - Early systolic murmur (VSD & CoArc $\frac{3}{6}$ rating)

	100 %	90 %	80 %	70 %	60 %	50 %
Normal (dB)	-38.06	-45.14	-46.27	-46.89	-47.5	-47.85
VSD (dB)	-5.08	-7.32	-9.49	-10.88	-12.03	-12.39
Δ (dB)	32.98	37.82	36.78	36.01	35.47	35.4

Table 3.2: Direct Ratio calculation of mid systolic (F) constituent, for different percentage values of constituent S1

From the constituent bar plot of pathological cases one can, to a certain extent, make a differential diagnosis. For example from Figure 3.25 can be seen that the systolic energy is fairly evenly distributed, thus one can conclude that it is a holosystolic murmur. Whereas in Figure 3.26 most of the energy is confined to the early systolic constituent which indicates an ejection systolic type murmur. No automatic differential diagnosis was done because of the various possible conditions that can be associated with one type of murmur. Possible future developments will be discussed in Chapter 5.

The limitations of the Direct Ratio method for automatic diagnosis are the following:

- It was only developed for systolic murmur detection, due to the inability to locate the exact position of S2 for some pathological cases. In most cases diastolic murmur classification is possible, but to make the algorithm applicable to all possible pathologies this functionality was removed and not documented.
- If the indicator of pathology lies within the S1 (B) constituent it will not be picked up by the algorithm.

- In the case of a normal heart sound a wrong classification is made if the first heart sound is not proper in size. If $S1 < S2$ the normalization of the signal is made with respect to $S2$. Furthermore if $S1 \ll S2$ the ratios are calculated using a small $S1$, which result in a false classification. This can be fixed by first testing which of $S1$ or $S2$ is the normalization reference and then to calculate the ratios using the correct one. (As mentioned locating $S2$ can be problematic)

3.4.2 Wavelet processing method

Analysis done by R. Thompson *et al* [6] has shown that time-frequency analysis is a versatile technique for detecting and classifying heart problems. Of the available time-frequency techniques, they found the wavelet transformation, discussed in Appendix C, to be a useful method for representing heart sound frequency dynamics without creating cross-term artifacts.

The same constituent methodology followed in the previous method is used to develop the wavelet processing method. The only difference being that direct energy values are taken to serve as an indicator. The wavelet analysis technique is also confined to the systolic region due to the same reasons as stated in the Direct Ratio technique.

The 4th order Daubechies wavelet (db4) and the 2nd order Coiflet wavelet (coif2) were used for the wavelet processing, due to their similarities with the heart sound waveform shape. Mathematically convolution in the time domain equals multiplication in the frequency domain. Thus, according to equation C.0.6 in Appendix 3 the Fourier transform of ψ will act as a bandpass filter on the signal $f(t)$. Table 3.3 show the wavelet pseudo-frequencies at different scales together with the 6 dB passband frequency limits of the Fourier Transform of ψ (the wavelet) for each scale.

	Scales for					
	db4			coif2		
	16	32	64	16	32	64
Pseudo-frequency (Hz)	492.2	246.1	123.05	501.15	250.57	125.28
6 dB bandlimit ³ (Hz)	$\pm 200-600$	$\pm 100-300$	$\pm 50-150$	213-593	107-295	53.3-148

Table 3.3: Pseudo frequencies computed with equation C.0.4 and 6 dB passband limits for the various scales tested. Coif2's 6 dB obtained form [6]

Figure 3.27 illustrates the difference between the wavelet coefficients of a normal and a pathological heart sound. It is clear that the VSD heart cycle contains a high

frequency (low scale) energy component between S1 and S2 not present in the normal recording. This representation in the time-scale space is used as the differentiating indicator for the wavelet technique.

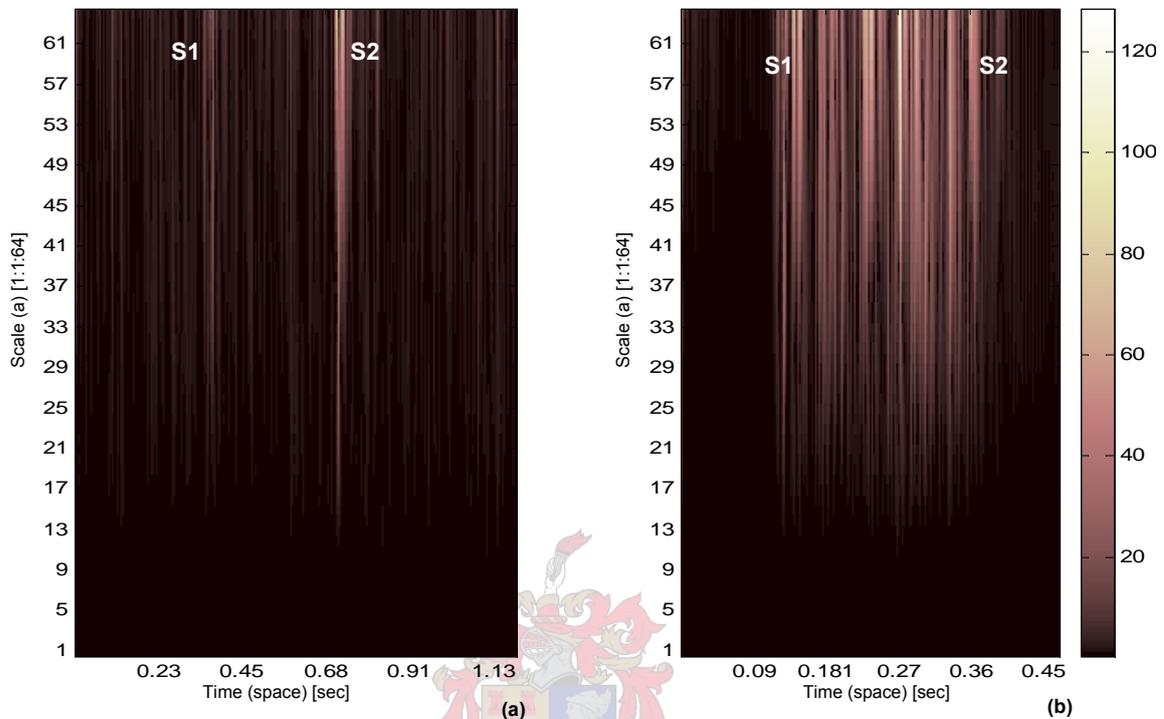


Figure 3.27: Absolute values of wavelet coefficients for (a) a normal heart sound; and (b) a pathological VSD (3/6) heart sound. Colour bar indicate amplitude of absolute values

Figure 3.28 illustrates the algorithm flow for the wavelet analysis technique. Before the constituent segmentation procedure the wavelet coefficients are calculated at the three mentioned scales using the db4 and coif2 continuous wavelet. Initially all the scales in the 20-650 Hz frequency range were tested to get the optimum performance scales. The composite energy (square of the wavelet coefficients expressed in dB) for the different scales, of all four systolic constituents, for each period of the recording were calculated for each patient.

The hypothesis is that the pathological sounds, that contains a higher frequency component than the normal sounds, will contain a higher energy content after being passed through the passband 'filters' of the different scales. Distinction between healthy and pathological hearts was made on the basis of the highest energy value of the four systolic regions being above or below a threshold value. The results are discussed in Chapter 4 section 4.2.2. The shortcomings of the wavelet analysis

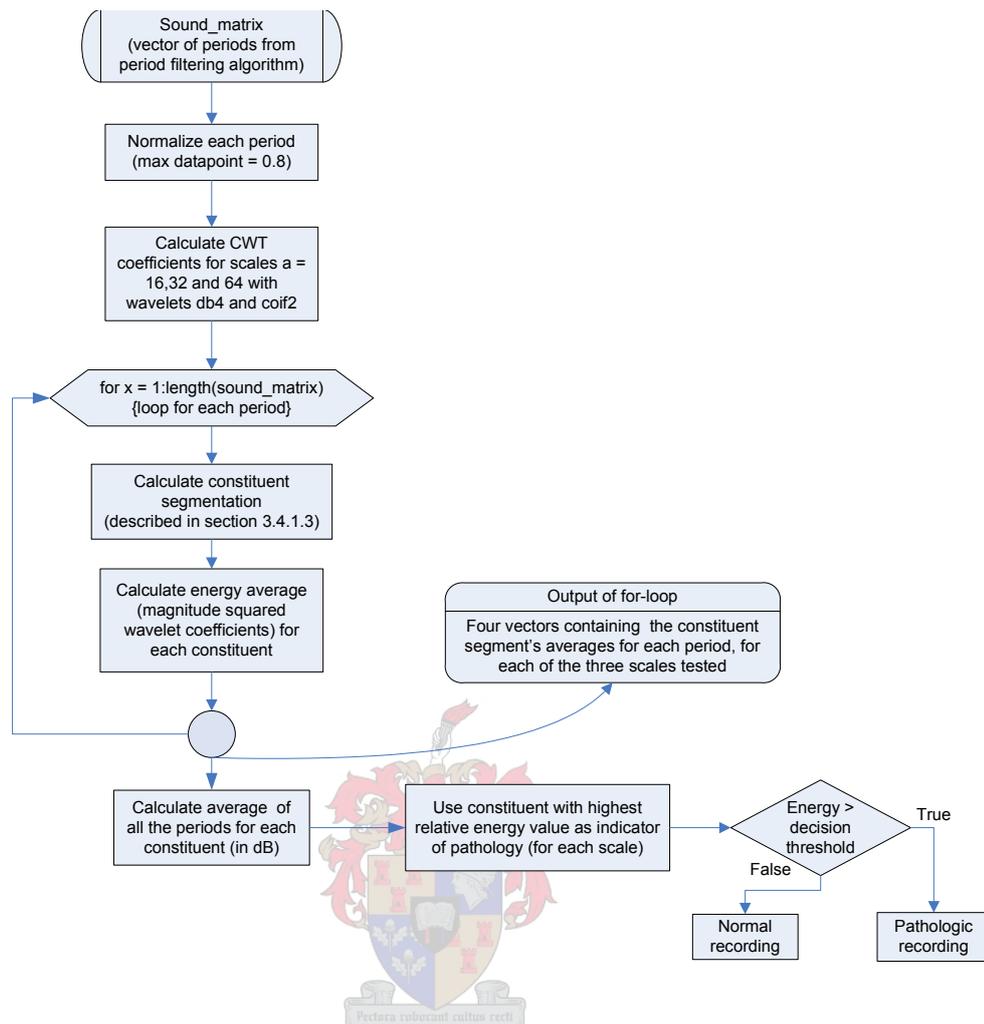


Figure 3.28: Algorithm flow for the wavelet analysis technique. Only for one patient (recording). The Wavelet analysis algorithm's code is listed in Appendix E.2.

method are limited to the first two points mentioned in the Direct Ratio method's limitation discussion.

3.4.3 Artificial Knowledge Based Neural Networks

An Artificial Neural Network (ANN) is a parallel processing system of interconnected processing elements, motivated by biological neurons and synapses. The neurons of an ANN are simple, non-linear summing nodes, and they are connected by weighted, directed interconnections. These interconnection weights are flexible parameters that are adjusted during the training of the network [48].

An artificial neural network is used as pathology classifier to test whether the clas-

sification performance of the two previous methods can be improved. The most prominent advantages of using an ANN as a classifier are:

1. The interconnection weights, that represent the solution, are found by iterative training;
2. ANN has a simple structure for physical implementation on a FPGA or equivalent;
3. A sufficiently trained ANN can easily map complex class distributions; and
4. the generalization property of the ANN produces appropriate results for the input vectors that are not presented in the training set [49]. Can classify "never seen before" inputs on the basis of correlated training data.

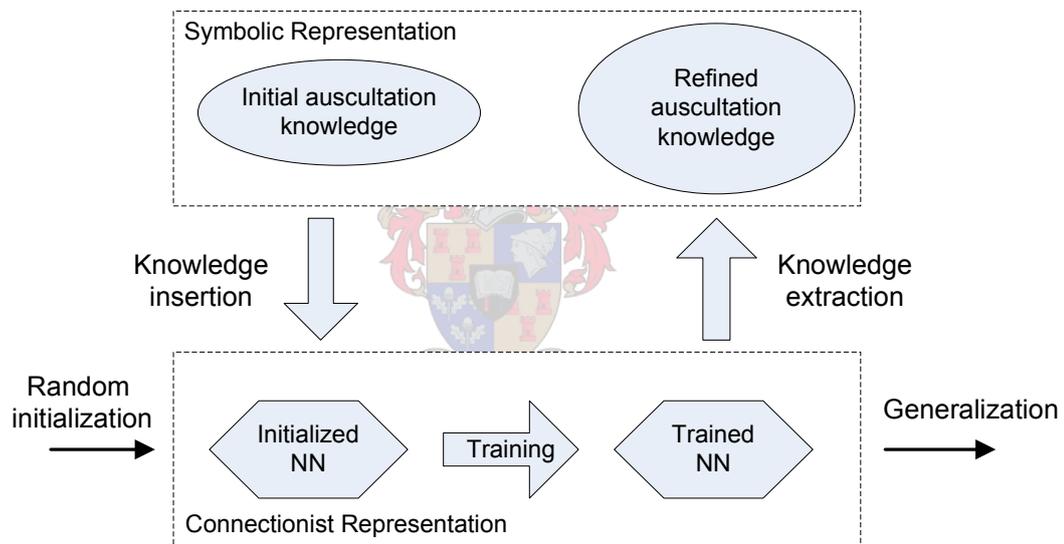


Figure 3.29: Combined symbolic neural learning. Motivation for using neural networks for classification purposes. Framework was adopted from the knowledge-based neurocomputing flowchart presented by [3]

There are seven steps in developing and training a knowledge based neural network. Figure 3.30 show the procedures to follow in order to develop an automatic heart sound classifier. *Matlab's* neural network functionalities were used to develop and test the classifier. The following subsections describe each of the procedures shown in Figure 3.30.

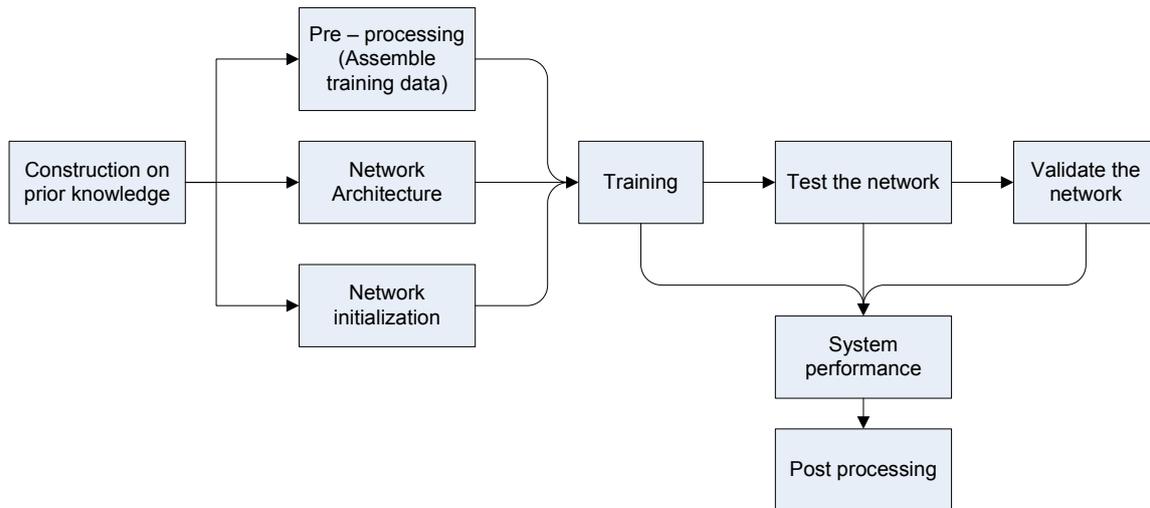


Figure 3.30: Neural network development and training methodology. The program code for the Network Architecture, Network Initialization, Training, Testing, Validation and Performance evaluation is listed in Appendix E.4

3.4.3.1 Assembling of the training data (construction of prior knowledge)

Gathering of prior information and knowledge is necessary for the successful modeling of difficult non-trivial problems [3]. We want to extract prior information that represents the heart sound recording sufficiently well to be classified by the network. Knowledge-based systems refer to systems either mainly concerned with the actual knowledge as in expert systems, or where knowledge is used advantageously with existing architecture to improve the performance of such a system. See Figure 1.2 in Chapter 1. "Knowledge-based neuro-computing (KBN) concerns methods to address the explicit representation and processing of knowledge where a neuro-computing system is involved" [50]. Thus,

**Knowledge obtained during previous methods development + Data =
Information**

Information + NN = Knowledge based neuro-computing

Different training data-sets were developed to determine optimum performance. The basic signal processing used to extract the time dependent frequency and intensity of the recording was however standard throughout the construction of the different training sets. From the information gained and lessons learned during the development of the previous methods, prior information or network inputs were constructed as shown in Figure 3.31

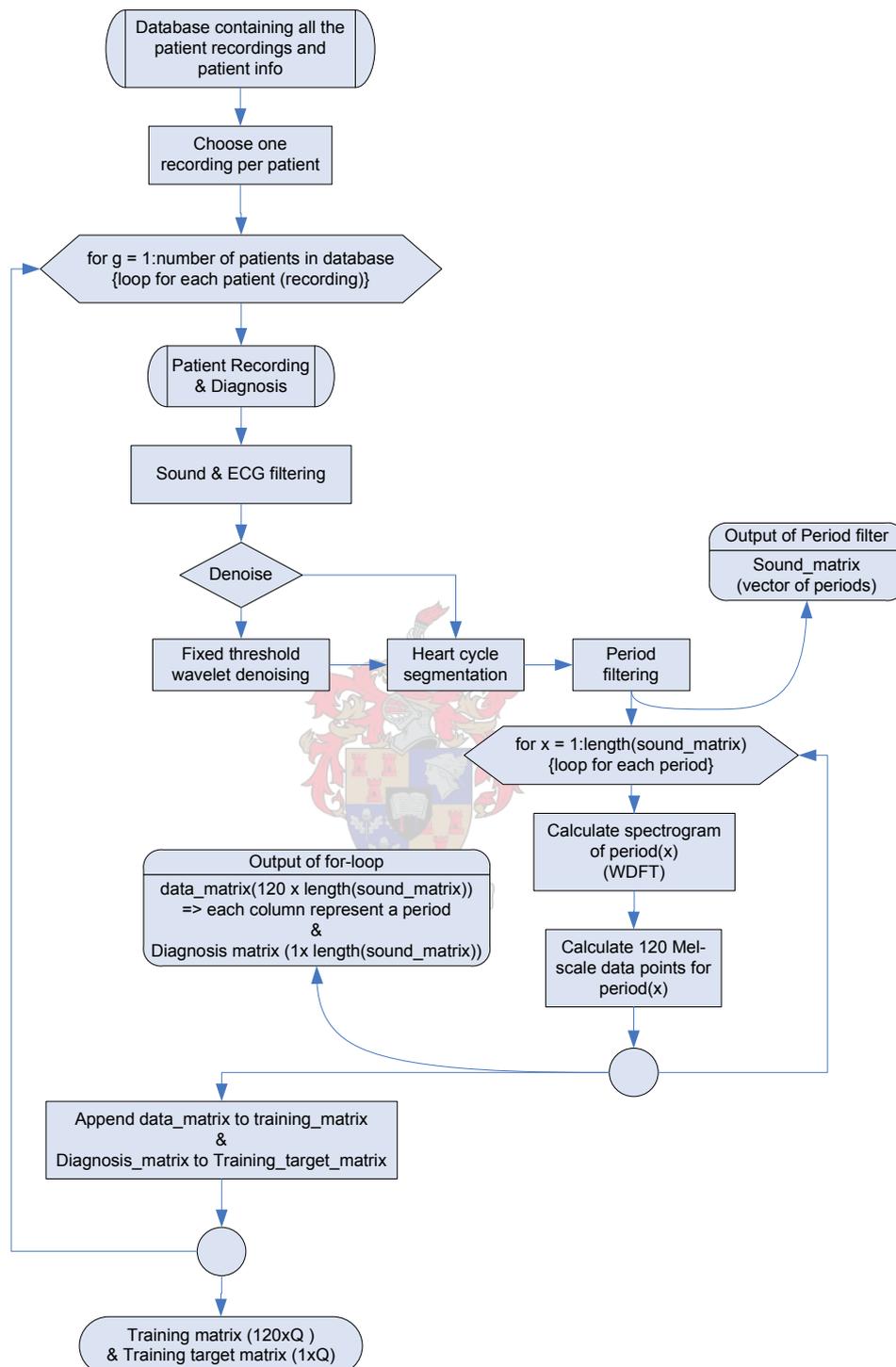


Figure 3.31: Algorithm flow for the construction of the training and training target data-set. Program code is listed in Appendix E.3, note that the construction of the validation matrix is done parallel in the program code.

The *Mel-Scale Frequency Bins*, used in the *Mel-Scale* datapoint calculation procedure, are exactly the same as those used in the period filtering algorithm described in section 3.3.3. The motivation for using this method is its effective extraction of representative frequency and intensity information. The same procedure is followed as described in section 3.3.3, to obtain the 120 representative data points per period.

The output of the algorithm described in Figure 3.31 is a training matrix and a training target matrix represented by equation 3.4.14 and 3.4.15 respectively.

$$\text{Training matrix} = \begin{bmatrix} t_{1,1} & t_{1,2} & \cdots & t_{1,Q} \\ t_{2,1} & t_{2,2} & \cdots & t_{2,Q} \\ \cdots & \cdots & \ddots & \cdots \\ t_{120,1} & t_{120,2} & \cdots & t_{120,Q} \end{bmatrix}, (-100 \leq t \leq 100 \in \mathbb{Z}) \quad (3.4.14)$$

$$\text{Training target matrix} = \begin{bmatrix} tt_1 & tt_2 & \cdots & tt_Q \end{bmatrix}, (tt \in \text{Boolean}) \quad (3.4.15)$$

where Q is the sum of all the periods from all the recordings.

Parallel to the assembling of the training matrix and training target matrix, the wanted validation matrix and validation target matrix are also assembled in exactly the same way. The validation data used for the various test sets developed will be discussed in Chapter 4.

3.4.3.2 Create the network - Network architecture

The network receives the representation of each period as a 120-element input vector. The network is then required to identify the condition by responding with a 1-element output vector. The output vector is of Boolean type, with 0 representing a normal (no-disease) case and a 1 representing a pathological case.

From previous worked done by C.G. DeGoff *et al* [37] a 2-layer log-sigmoid & log-sigmoid feed-forward multi-layer perceptron (MLP) network was chosen for the classification purpose. Figure 3.32 shows the notation for describing a feed-forward MLP. In pattern recognition it is now most common to use a 2-layer MLP with sigmoid activation functions, because it can be shown that this network can approximate any decision boundary [45, 4].

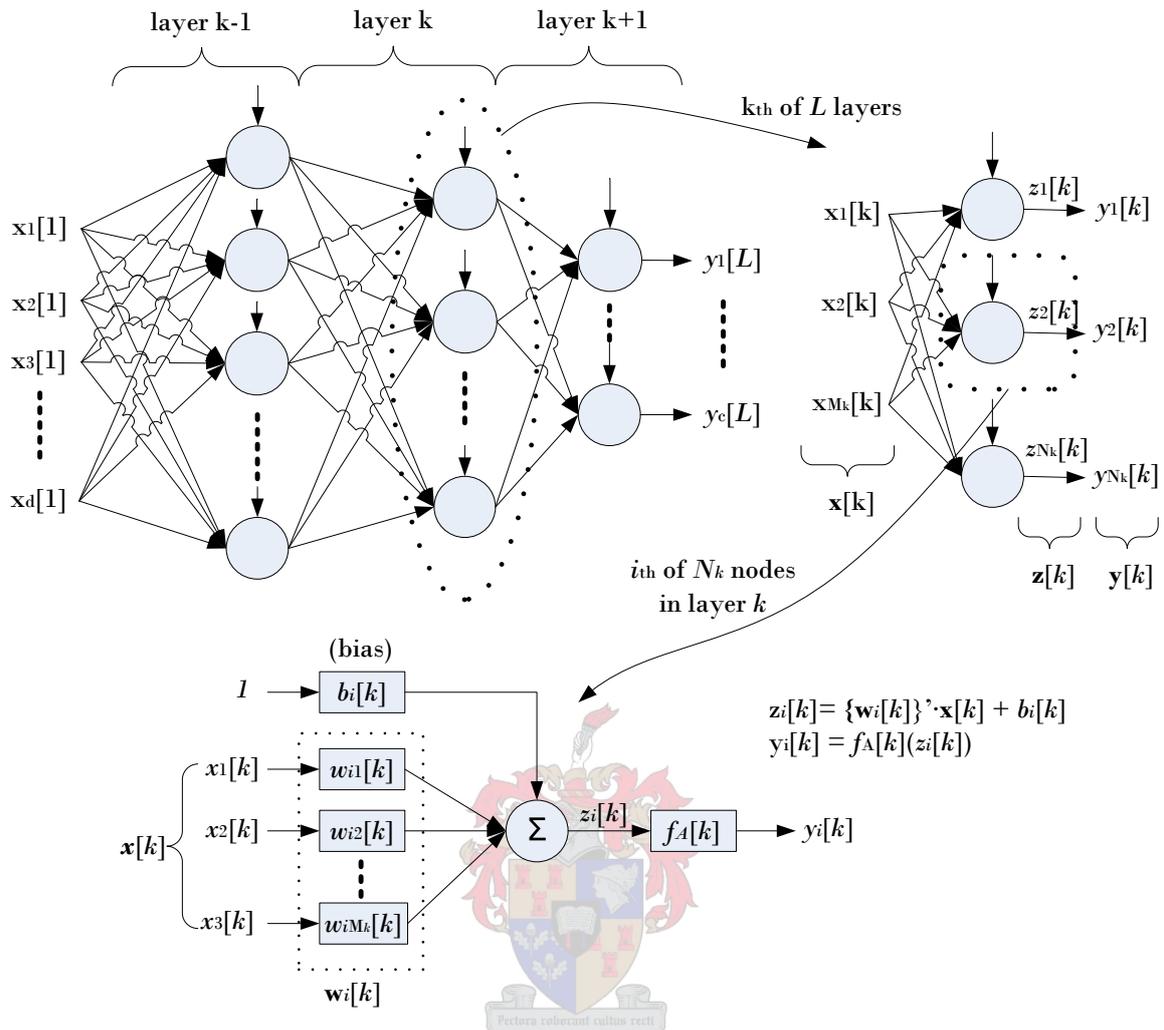


Figure 3.32: Notation for describing a MLP, described with L layers , a d -dimensional input and c outputs. (Courtesy of Dr. Thomas Niesler, Stellenbosch University, South Africa [4])

Consider layer k which has M_k inputs and N_k nodes

$$\mathbf{x}[k] = \begin{bmatrix} x_1[k] \\ x_2[k] \\ \vdots \\ x_{M_k}[k] \end{bmatrix} \text{ is the input vector to layer } k \quad (3.4.16)$$

$$\mathbf{z}[k] = \begin{bmatrix} z_1[k] \\ z_2[k] \\ \vdots \\ z_{N_k}[k] \end{bmatrix} \text{ is the vector of summation output for layer } k \quad (3.4.17)$$

$$\mathbf{y}[k] = \begin{bmatrix} y_1[k] \\ y_2[k] \\ \vdots \\ y_{N_k} \end{bmatrix} \text{ is the vector of outputs for layer } k \quad (3.4.18)$$

$$\mathbf{w}_i^T[k] = \left[w_{i1}[k] \quad w_{i2}[k] \quad \cdots \quad w_{iM_k}[k] \right] \text{ is the weight vector of node } i \text{ in layer } k \quad (3.4.19)$$

$$\mathbf{W}[k] = \begin{bmatrix} \mathbf{w}_1^T[k] \\ \mathbf{w}_2^T[k] \\ \vdots \\ \mathbf{w}_{N_k}^T[k] \end{bmatrix} = \begin{bmatrix} w_{11}[k] & w_{12}[k] & \cdots & w_{1M_k}[k] \\ w_{21}[k] & w_{22}[k] & \cdots & w_{2M_k}[k] \\ \cdots & \cdots & \ddots & \cdots \\ w_{N_k1}[k] & w_{N_k2}[k] & \cdots & w_{N_kM_k}[k] \end{bmatrix} \text{ is the weight matrix for layer } k \quad (3.4.20)$$

The MLP feed-forward ANN is equivalent to the non-linear functional

$$Y_i^k = F \left(\sum_{j=1}^{M_{k-1}} w_{ij}^k Y_j^{k-1} + b_i^k \right) \quad (3.4.21)$$

where Y_i^k is the output of the i^{th} neuron in the k^{th} layer, w_{ij}^k is the weight of the connection from the j^{th} neuron in the $(k-1)^{th}$ layer to the i^{th} neuron in the k^{th} layer, b_i^k is the bias connected to the i^{th} neuron in the k^{th} layer and M_{k-1} is the number of neurons in the $(k-1)^{th}$ layer. F is the activation function, which may be thought of as providing a non-linear gain for the neuron. From the work done by [37], the log-sigmoid function shown in equation (3.4.22) was chosen as the activation function.

$$F = \frac{1}{1 + e^{-u}} \quad (3.4.22)$$

The sigmoid function shown above takes the input, which may have any value between plus and minus infinity, and squashes the output into the range 0 to 1. This output range is perfect for learning to output boolean values. This function is also suitable because it is differentiable, which is a precondition when using the back-propagation training algorithm described in 3.4.3.4.

The network needs 120 inputs, to receive the 120-element input vector, and one neuron in its output layer to identify a condition. The hidden layer (first layer) has 30 perceptrons. This number was obtained from the network developed by [37]. If the network has trouble learning, then neurons can be added to this layer.

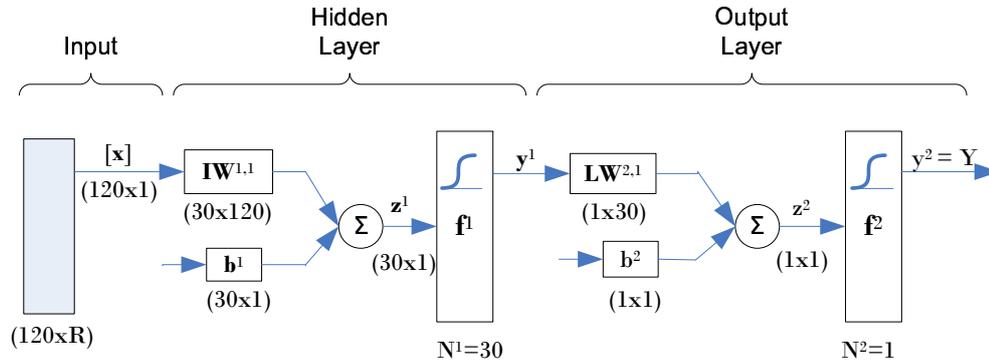


Figure 3.33: Short notation for the 2-layer feed-forward backpropagation artificial neural network used as the classifier. The functions in both layers are sigmoid activation functions.

3.4.3.3 Initialization of the network weights

Before training the network the weights and biases of the network must be initialized. Initialization of the layer's weights and biases is done with the *Nguyen-Widrow* layer initialization function [51]. The *Nguyen-Widrow* method generates initial weight and bias values for a layer, so that the active regions of the layer's neurons will be distributed roughly evenly over the input space.

Advantages over purely random weights and biases are:

1. Few neurons are wasted (since all the neurons are in the input space).
2. Training works faster (since each area of the input space has neurons).

3.4.3.4 Training the network: Background and application

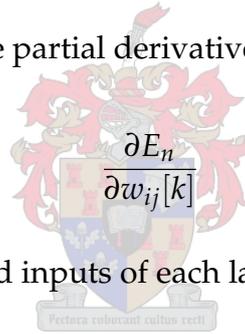
In the training mode, the network is presented data on the basis of which it learns the underlying mapping. The first step of training is the forward pass, which consists of calculating the output vector by running the input vector through the ANN. This is followed by a backward pass, where the error derivatives are calculated for each weight. The error derivatives for a weight are summed until all the data points

have been run through the network once. This constitutes an epoch. The weights are updated after each epoch so that the ANN error decreases. The error backpropagation learning algorithm is deployed for this purpose. The backpropagation algorithm is a method of calculating the partial derivatives of a certain cost function with respect to the weights of a multi-layer perceptron [45, 4]. The cost function used is the least mean square error function shown in equation (3.4.23). As each input is applied to the network, the network output is compared to the target output. The training and learning algorithm aims to minimize the average of the sum of these errors to a goal of 10^5 , within a 4000 epoch.

$$E = \frac{1}{R} \sum_{k=1}^R e(k)^2 = \frac{1}{R} \sum_{k=1}^R (t(k) - a(k))^2 \quad (3.4.23)$$

with $t(k)$ the target output, $a(k)$ the actual output, $e(k)$ the error and R the number of training data vectors.

Backpropagation will allow the partial derivatives to be calculated.



$$\frac{\partial E_n}{\partial w_{ij}[k]} \quad (3.4.24)$$

with $i = 1, 2, \dots, N_k$ (all nodes and inputs of each layer)

$j = 0, 1, 2, \dots, M_k$

$k = 1, 2, \dots, L$

These partial derivatives can then be used to batch update the network's weights and bias values by steepest descent adaptive learning optimization schemes.

$$w_{ij}^{(t+1)}[k] = w_{ij}^{(t)}[k] + \Delta w_{ij}^{(t)}[k] \quad (3.4.25)$$

where

$$\Delta w_{ij}^{(t)}[k] = -\eta \cdot \frac{\partial E}{\partial w_{ij}^{(t)}[k]} \quad (3.4.26)$$

where

$$\frac{\partial E}{\partial w_{ij}^{(t)}[k]} = \sum_{n=1}^R \frac{\partial E_n}{\partial w_{ij}^{(t)}[k]} \quad (3.4.27)$$

Each batch weight update occurs only after all R training data vectors have been processed (one such training iteration is called an epoch).

The error surface of the MLP has many local minima. The steepest-descent update equation (3.4.25) is a local optimization method, thus it detects only the local minimum which might generally not be the global minimum. The momentum parameter and term are introduced to equation (3.4.26) to avoid the training algorithm to get caught in a local minimum. The result, equation (3.4.25) with $\Delta w_{ij}^{(t)} [k]$ equal to equation 3.4.28, is a network training function that updates weight and bias values according to gradient descent momentum and an adaptive learning rate.

$$\Delta w_{ij}^{(t)} [k] = -\eta \cdot \frac{\partial E}{\partial w_{ij}^{(t)} [k]} + \beta \cdot w_{ij}^{(t-1)} [k] \quad (3.4.28)$$

with step size η and momentum parameter β , where $0 < \beta < 1$.

Choosing the correct step-size is important in that the convergence heavily depends on the step size. The step-size might overshoot the local minimum if too big and get stuck in a local minimum and result in a very slow training algorithm if too small. The general rule of thumb is to increase the step size when the gradient maintains the same sign for many iterations, and to decrease the step size when the sign of the gradient begins to oscillate.

If the network is capable of detecting conditions that was represented in the training data-set, it may not detect the conditions it has not seen before. Properly trained backpropagation networks tend to give reasonable answers when presented with inputs that they have never seen. This is the generalization property of the network. Improved generalization could be obtained by training the network on more data-sets. Two different training methods are used to evaluate the performance of the network in the respective situations. These methods will be discussed in Chapter 4 in combination with their results.

3.4.3.5 Test the network

After training the network, the network is simulated to test whether each period's output corresponds with its target training value. Each recording is represented by the average output of all its periods. Testing (simulating) the network is, however, not a figure of merit for the classification performance of the network because all the test data was included in the training data set. Testing a network is merely done to verify whether the network was able to find a correlation between the different representatives within a given condition (e.g. normal or pathological). To really test

the network's performance, a separate data set must be used. This is called network validation.

3.4.3.6 Validate the network (Algorithm validation)

The validation process is exactly the same as the simulation process described above, the only difference being that the validation of the network is done on a data set never seen before by the network. The validation process thus tests the generalization property of the network. As mentioned in section 3.4.3.4, the network was trained on two different training methods. For each training method the validation method is unique, testing different performance characteristics of the network. The different validation methods used are discussed in Chapter 4.

3.4.3.7 Post-processing (Statistical analysis)

With the networks prediction plotted on the y-axis and the corresponding validation record number on the x-axis, a threshold line is drawn at a specific y-value to distinguish between normal and pathological cases. With the threshold at different values, the erroneous classifications are calculated to serve as a performance indicator. The different statistics used to analyze the performance of the various networks tested are discussed in the next section. The results obtained with the different classifiers, combined with a complete analysis of the results, follow in the next chapter.

3.5 Statistical Analysis

The results obtained from the three methods were analyzed using various statistical approaches so as to show that the research results were not by chance, but were statistically significant results. The different statistics used to analyze the performance of the various algorithms developed are described in this section.

3.5.1 Descriptive parametric statistics [2]

The patients examined can be divided into three sample (population) groups: pathological, functional (innocent murmur) and normal (no-murmur) - each with n number of recordings.

The equation for determining the arithmetic mean of n observations is

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} \quad (3.5.1)$$

with sample standard deviation (SD) (an estimate of the population standard deviation) of

$$s = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}} \quad (3.5.2)$$

where

\bar{X} = the sample mean,

n = the number of samples,

X_i = the i^{th} individual values .

3.5.2 Sample distribution

If random samples were selected from a population that was known to be normally distributed with a finite mean μ and a standard deviation σ , then the sampling distribution of sample means \bar{X} will be normally distributed. The population we are working with is however not known to be normally distributed.

The central Limit Theorem: The central limit theorem states that, if all samples of size n are selected from a population with a finite mean μ and a standard deviation σ , then as n increases in size, the distribution of sample means (\bar{X}) will tend towards a normal distribution, with a mean of μ (the population mean), and a standard deviation equal to σ / \sqrt{n} (called the standard error of the mean or $\sigma_{\bar{x}}$).

Large samples ($n > 30$) [52] are necessary for the central limit theorem to be used as the basis of normality for a sampling distribution. The need for a large sample arises because both the population distribution, and the population standard deviation are unknown. When small samples ($n < 31$) are selected the use of the central limit theorem, to imply normality of the sampling distribution, is not valid.

A more significant and defined test for testing normality is the *Shapiro-Wilks' W test*. The *Shapiro-Wilks' W test* is the preferred test of normality because of its good power properties as compared to a wide range of alternative tests (Shapiro, Wilk, & Chen, 1968). The *Shapiro-Wilks' W test* is also recommended for small and medium samples of up to $n = 2000$.

The W value may be thought of as the correlation between the given data and their corresponding normal scores, with $W = 1$ when the given data is perfectly normal in distribution. When W is significantly smaller than 1, the assumption of normality is not satisfied. If the W statistic is significant ($p < 0.05$), then the hypothesis that the respective distribution is normal should be rejected. The function used and shown in Appendix D was obtained from the StatLib Library [53]. It seems to agree with the Shapiro Wilks' tests performed by other programs. As an alternative to the central limiting theorem the *Shapiro-Wilks'* W test will be used as the decisive normality test.

If \bar{X} and s^2 are the mean and variance of a sample of size n from a normal distribution $N(\mu, \sigma^2)$, where μ and σ^2 are unknown, then

$$\bar{X} \pm t_{n-1; \alpha/2} \frac{s}{\sqrt{(n)}} \quad (3.5.3)$$

is a $100(1 - \alpha)\%$ confidence interval for μ , with $t_{n-1; \alpha/2}$ the Student's t-distribution for $n-1$ degrees of freedom ⁴ [2].

If the population, from which the sample is drawn, has unknown mean and unknown variance, but is not quite normal equation 3.5.3 is an approximate $100(1 - \alpha)\%$ confidence interval, adequate for most practical purposes [2].

If n is large (greater than about 30), $t_{n-1; \alpha/2}$ is approximately equal to the area under the normal curve measured in standard deviation units $Z_{\alpha/2}$. Equation 3.5.3 is then rewritten as

$$\bar{X} \pm Z_{\alpha/2} \frac{s}{\sqrt{(n)}} \quad (3.5.4)$$

for the $100(1 - \alpha)\%$ confidence interval of μ , with $Z_{(\alpha/2=0.025)} = 1.96$ for a 95% confidence interval [2].

3.5.3 Confidence Intervals and Hypothesis Testing

When a hypothesis about the difference between the pathological and the no-disease population means is tested, the following null hypothesis is tested [52]:

$$H_0 : \mu_1 - \mu_2 = 0 \text{ or } \mu_1 = \mu_2 \quad (3.5.5)$$

⁴Read value from table originally published by [54]

against the alternate hypotheses,

$$H_A : \mu_1 - \mu_2 \neq 0 \quad (3.5.6)$$

Figure 3.34 shows the two-sided test, for the difference between the pathological and the no-disease sample means, hypothesis testing.

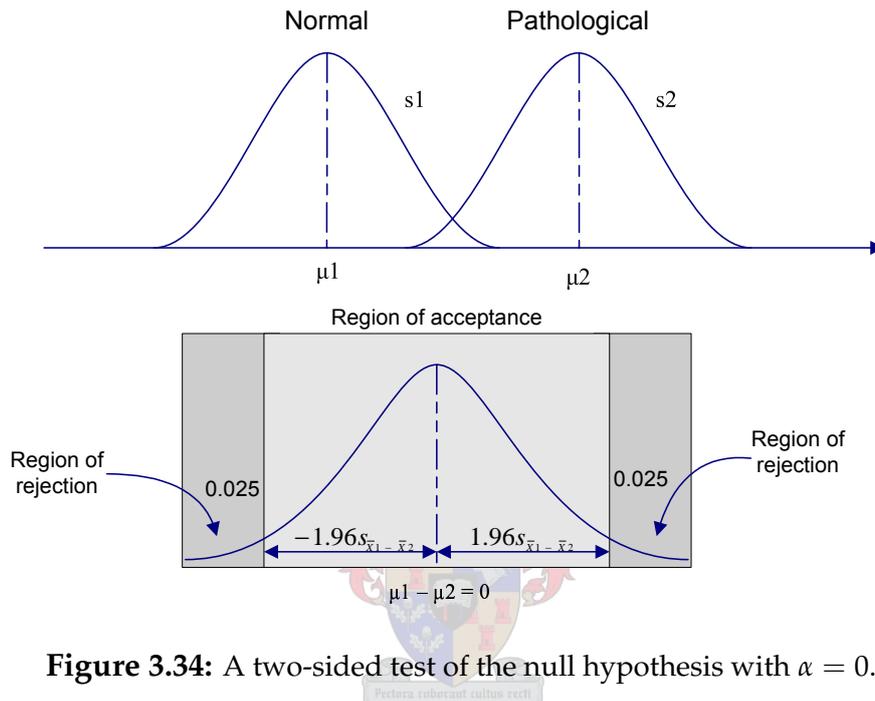


Figure 3.34: A two-sided test of the null hypothesis with $\alpha = 0.05$

To test the hypothesis the following six steps are executed:

- Step 1: Calculate sample size of both populations;
- Step 2: Calculate mean value of both samples μ_1 and μ_2 using equation 3.5.1;
- Step 3: Calculate the sample standard deviation of both samples s_1 and s_2 using equation 3.5.2;
- Step 4: Calculate the standard error of the difference between the two sample means using equation 3.5.7 when both samples are large (more than 30).

$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \quad (3.5.7)$$

or equation 3.5.8 when both or one of the samples are small (30 or less).

$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{n_1 + n_2 - 2}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \quad (3.5.8)$$

Step 5: Construct the confidence interval estimate of the difference between two population means ($\mu_1 - \mu_2$). If the population variances are unknown, and the shapes of the population distribution are not necessarily normal, then equation 3.5.9 can be used provided both sample sizes are large (n_1 and $n_2 > 30$) so that the central limiting theorem is effective.

$$\mu_1 - \mu_2 = (\bar{X}_1 - \bar{X}_2) \pm Z_{\alpha/2} s_{\bar{x}_1 - \bar{x}_2} \quad (3.5.9)$$

where

$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \quad (3.5.10)$$

and with $Z_{(\alpha/2=0.025)} = 1.96$ for a 95% confidence interval [2].

Step 6: With all the variables calculated the hypothesis can now be tested using equation 3.5.11

$$Z_{test} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{s_{\bar{x}_1 - \bar{x}_2}} \quad (3.5.11)$$

Reject the null hypothesis if $Z_{test} < -1.96$ or $Z_{test} > 1.96$ for a level 0.05 significance. This means that there is a statistical significant difference between the mean of the normal population and the pathological population. The difference is real, not chance, with a probability of being correct equal to 0.95. Thus the hypothesis can be adopted that there is a difference between the energy content in a normal patient recording and that of a pathological recording which can serve as a measure for diagnosis. If the null hypothesis is adopted there is no significant difference between the mean of the normal population to that of the pathological population, if this is the case this figure cannot be used as a measure for diagnosis.

3.5.4 Sensitivity and specificity

The following section describes a general approach to calculate the relative sensitivity and specificity for the purpose of estimating the expected proportion of erroneous decisions. The sensitivity and specificity are calculated relative to clearly defined, relevant reference standards 'golden standards'. The diagnosis made with an ultra sound unit (if conditions apply) was used as the golden standard. If an

ultra sound unit was not available the diagnosis of the specialist was considered as the golden standard. The sensitivity and specificity are calculated by choosing a threshold value on the classification axis, and then calculating the number of false negatives and false positives.

Sensitivity refers to the proportion of people with disease who have a positive test result. The higher the sensitivity the fewer false negatives (sick people lying beneath the threshold line).

Specificity refers to the proportion of people without disease who have a negative test result. The higher the specificity the fewer the false positives (healthy people lying above the threshold line).

To calculate the sensitivity and the specificity we can assume that there are four possible groups of patients, as indicated by a, b, c and d in Table 3.4:

	Golden Standard	
	Positive	Negative
Diagnostic Test +	a True Positive	b False Positive
Diagnostic Test -	c False Negative	d True Negative

Table 3.4: Possible patient groups.

From Table 3.4 we determine the sensitivity and specificity as follows:

$$\text{Sensitivity}(\psi) = \frac{a}{a + c} \quad (3.5.12)$$

$$\text{Specificity}(\chi) = \frac{d}{b + d} \quad (3.5.13)$$

Statistical confidence intervals are placed around estimates, indicating the level of confidence that the true measure of sensitivity and specificity is included in the interval. For the 95 % confidence interval of the sensitivity for a specific threshold, equation (3.5.14) is used

$$\psi \pm 1.96 \sqrt{\frac{\psi(1 - \psi)}{nP}} \quad (3.5.14)$$

and equation (3.5.15) is used to calculate the 95 % confidence interval of the specificity for a specific threshold;

$$\chi \pm 1.96 \sqrt{\frac{\chi(1-\chi)}{n(1-P)}} \quad (3.5.15)$$

with

$P = \frac{a+b}{a+b+c+d}$ the prevalence
and $n = a + b + c + d$ the total sample size.

The developed algorithms' performances for a given energy threshold value, was measured by the ratio of called positives to true positives, with a universe of positives (the sensitivity), paired with the ratio of called negatives to true negatives, with a universe of negatives (the specificity). A plot of sensitivity % vs. [100-specificity]% for a variety of threshold values is called a receiver operation characteristic (ROC) curve. No particular fixed threshold value was selected for any of the algorithms developed; the objective was to obtain the superior ROC curve. Once this is done, the user is free to choose his or her own threshold point on the ROC curve on the basis of personal criteria for the trade off of sensitivity for specificity. Example optimum threshold values was selected in the next chapter to illustrate the trade off between sensitivity and specificity.

Choosing a fixed threshold value to make the decision whether to refer a patient or not, is not an automated statistical decision, but rather an ethical and economical one. This decision is up to the practitioner making the diagnosis, and will be influenced by the patients specific conditions. If a practitioner uses this method to make a more confident diagnosis, the calculated sensitivity and specificity can serve to indicate with how much confidence he/she can interpret the selected/specific algorithm's output.

Chapter 4

Results and Findings

4.1 Data collection

Over a 5 month period the clinical auscultation sounds and ECG-data of 171 patients were digitized. A total of 411 recordings were made, giving an average of 2 to 3 recordings per patient. Eight of the patients' recordings were discarded due to their age being above 16 years. After filtering the recordings of the remaining 163 patients, 311 recordings were found to be a good enough representation of the respective conditions, and suitable for further investigation. Figure 4.1 shows the composition of the database containing the 311 recordings of the 163 pediatric patients. Only one recording per patient was used for the performance testing of the various algorithms developed. If more than one recording of a patient was included in the database, the most representative recording was chosen for this purpose. Consequently only 163 recordings of the available 311 recordings were used for algorithm testing. The average age of all the patients in the database is 5 years and 10 months, and range is between 2 months and 16 years.

4.2 Feature extraction algorithms (FEA's)

This section is an exposition of the results of the various feature extraction algorithms discussed in Chapter 3. Results are presented in the same order as their associated methodology description in the previous chapter. Results are presented in graphical format with associated tables containing the statistic data. A short discussion concerning the performance of each method follows each of the results, to-

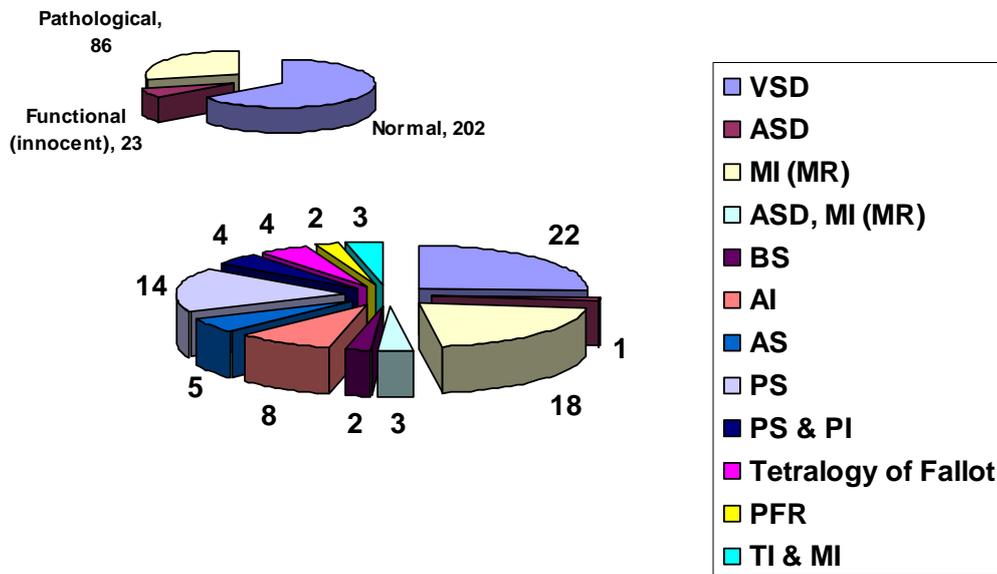


Figure 4.1: The inset graph shows distribution of the 311 recordings studied. The primary pie chart show the distribution of the 86 pathological recordings which consist of the following conditions: ventricular septal defect (VSD), atrial septal defect(ASD), mitral incompetence or regurgitation (MI or MR), barlow syndrome (BS), aortic insufficiency (AI), aortic stenosis (AS), pulmonary stenosis (PS), pulmonary insufficiency (PI), Tetralogy of Fallot, peri-cardial friction rub (PFR) and tricuspid incompetence (TI)

gether with the lessons learnt during development and testing and the contribution it made to the ensuing method.

4.2.1 Direct Ratio results

Figure 4.2 shows the output of the Direct Ratio method, with the energy value in dB plotted on the y-axis versus the corresponding patient on the x-axis. Only the highest energy value (dB) of the four systolic constituents of each patient was used for classification. All three data groups are shown with separate indicators. Cases with a diastolic murmur as indicator of pathology were left out of this analysis. Statistics are calculated for three different comparisons: no murmur vs. pathological; innocent murmur vs. pathological; and no disease (no murmur and innocent murmur) vs. pathological.

Figure 4.3, and 4.4(d),(e) illustrates the *Shapiro-Wilks' W* result for the test of normality on all three the data sets. With a p -value of $p < 0.05$ considered significant

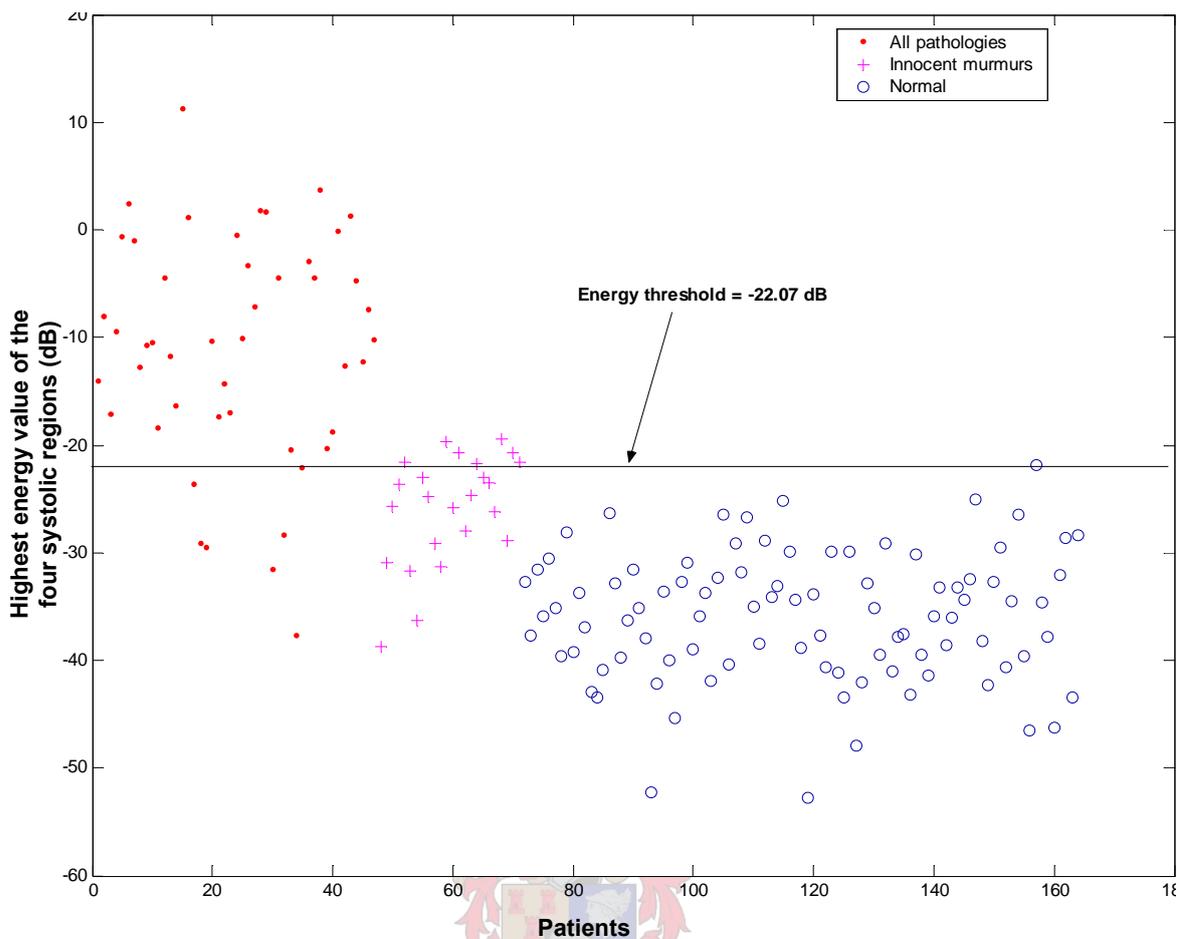


Figure 4.2: Results of the Direct Ratio method. The inset legend show the data groups associated markers. The threshold line drawn at $-22,07$ dB will be discussed in a later subsection

- the pathological and normal heart sound data-set are classified as normal with a normality of $W = 0.977$ ($p = 0.472$) and $W = 0.98951$ ($p = 0.67392$) respectively. According to the W -test, the innocent murmur data-set is also classified as normal with a normality of $W = 0.917$ ($p = 0.05052$) but is categorized as a *Student t* distribution in further calculations, due to its size ($n=24$) and its almost significant p -value of 0.05052 .

The calculated descriptive statistics for the data in Figure 4.2 are shown in Table 4.1

Testing whether the null hypothesis can be rejected for the mean difference between all three data-sets with respect to the pathological data set can be done by following the six steps described in section 3.5.3.

Test the hypothesis for mean difference between the population mean of the no murmur (normal) data-set and the pathological data-set using equation 3.5.7, 3.5.9 and

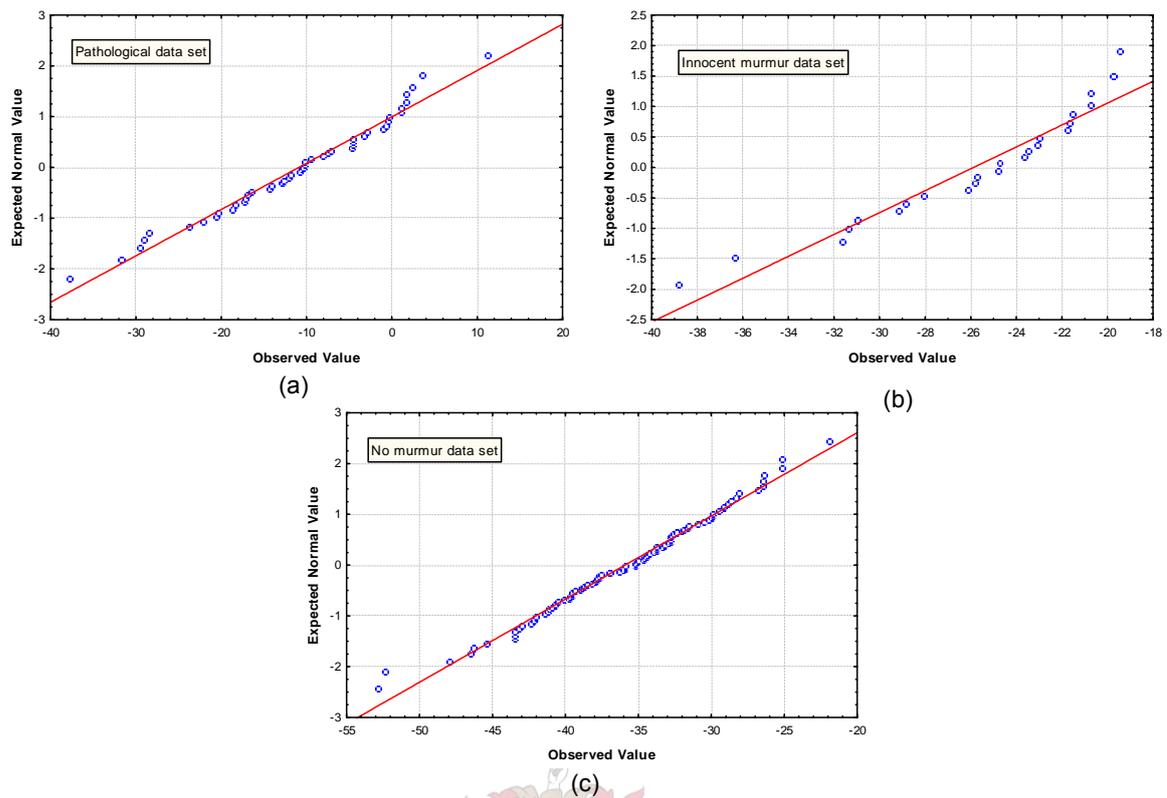


Figure 4.3: (a), (b) and (c) show the difference between the normal distribution and the three data-sets

3.5.11:

$$\begin{aligned} \Rightarrow s_{\bar{x}_1 - \bar{x}_2} &= 1.66; \\ \Rightarrow 21.75 &< \mu_1 - \mu_2 < 28.26 \\ \Rightarrow Z_{test1} &= 15.05 \end{aligned}$$

Test the hypothesis for mean difference between the population mean of the functional (innocent) murmur data-set and the pathological data set using equation 3.5.8, 3.5.9 and 3.5.11:

$$\begin{aligned} \Rightarrow s_{\bar{x}_1 - \bar{x}_2} &= 2.29 \\ \Rightarrow 10.28 &< \mu_1 - \mu_2 < 19.63 \\ \Rightarrow Z_{test2} &= 6.54 \end{aligned}$$

Test the hypothesis for mean difference between the population mean of the no-disease data-set and the pathological data-set using equation 3.5.7, 3.5.9 and 3.5.11:

$$\begin{aligned} \Rightarrow s_{\bar{x}_1 - \bar{x}_2} &= 1.67 \\ \Rightarrow 19.66 &< \mu_1 - \mu_2 < 26.23 \\ \Rightarrow Z_{test3} &= 13.7 \end{aligned}$$

The null hypothesis can be rejected in all three cases because $Z_{test1,2, \text{and} 3} > 1.96$. This means that there is a statistically significant difference between the mean of the three

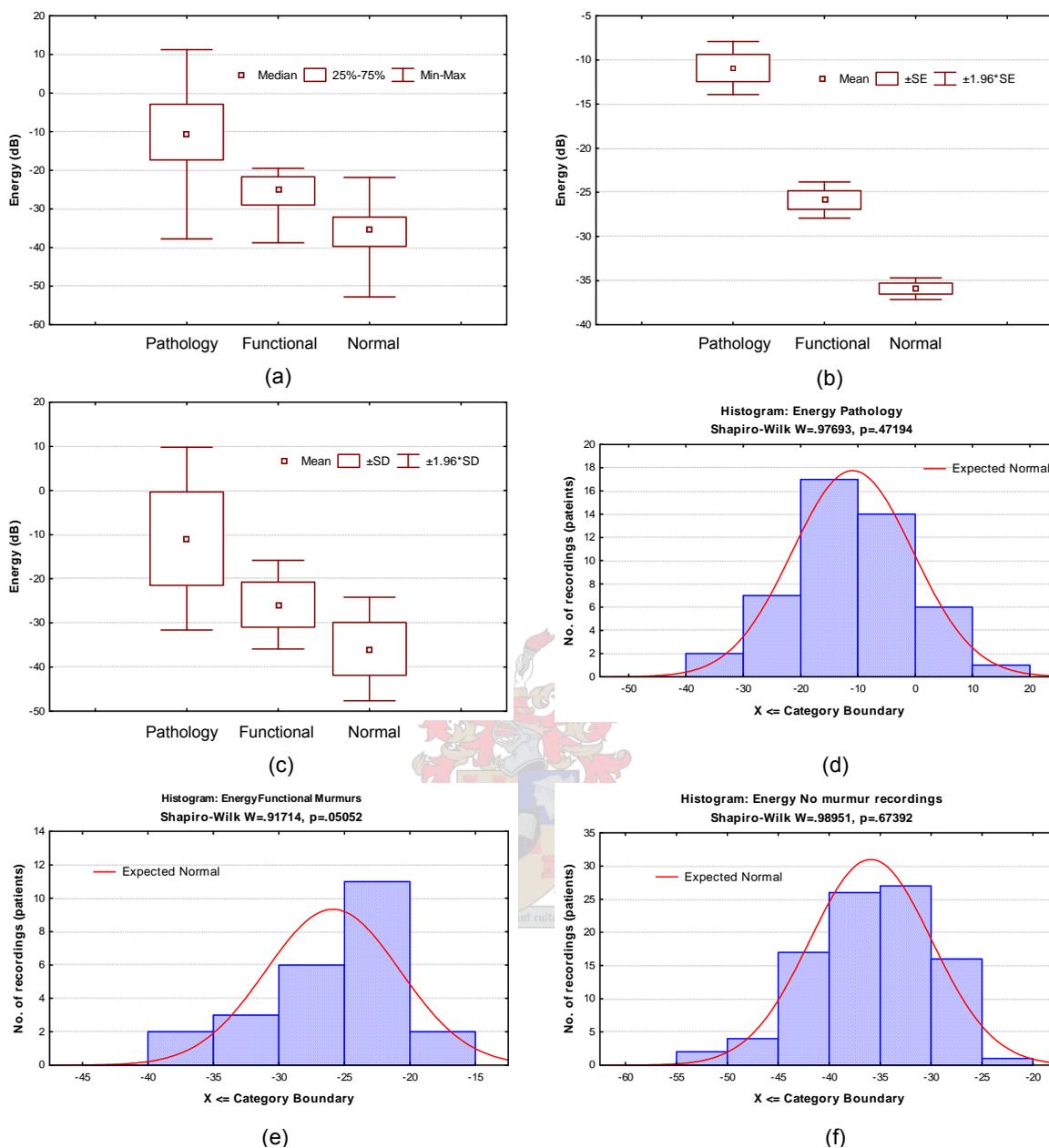


Figure 4.4: (a), (b) and (c) show the descriptive statistics for the Direct Ratio method, and (d), (e) and (f) the histogram distribution for the respective data-sets

	n	Mean	Minimum	Maximum	SD	-95% CI	95% CI
No murmur	93	-35.916	-52.805	-21.8537	5.986	-34.7	-37.13
Innocent murmur	23	-25.872	-38.774	-19.473	5.1281	-23.71	-28.03
No disease	116	-33.856	-52.805	-19.473	7.087	-32.58	-35.14
Pathological	47	-10.91	-37.748	11.244	10.564	-7.89	-13.93

Table 4.1: Descriptive statistics for the Direct Ratio method. [SD = standard deviation and CI = confidence interval]

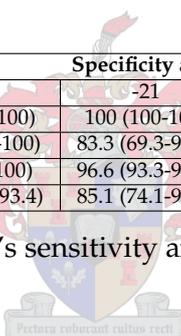
populations and the pathological population. The difference is real, not chance, with a probability of being correct equal to 0.95. Although it is evident from viewing Figure 4.2, the hypothesis can now be adapted, that there is a statistically significant difference between the energy content in a normal patient recording, and that of a pathological recording, which can serve as a measure for diagnosis.

4.2.1.1 Sensitivity and Specificity

With the mean variance hypothesis accepted sensitivity and specificity for the various data groups can now be calculated. The sensitivity and specificity for five different threshold values, together with their 95 % confidence interval, are shown in Table 4.2. If a threshold value of -22.07 dB is chosen, as an example, the sensitivity and specificity of the algorithm is 87.2 (77.9-96.7)% and 93.2 (88.5-96.7)% respectively. Refer to section 3.5.4, in the previous chapter, for a specific energy threshold choice.

	n	Specificity at energy threshold cutpoint (db)				
		-20	-21	-22	-22.07	-23
No murmur (95 % CI)	93	100 (100-100)	100 (100-100)	98.9 (97-100)	98.9 (96.9-100)	98.9 (97-100)
Innocent murmur (95 % CI)	24	91.7 (81.9-100)	83.3 (69.3-97.4)	70.8 (52.6-89)	70.8 (52.3-89.4)	66.7 (47-86.4)
No disease (95 % CI)	117	98.3 (96-100)	96.6 (93.3-99.8)	93.2 (88.6-97.8)	93.2 (88.5-96.7)	92.3 (87.4-97.2)
Sensitivity (95 % CI)	47	80.9 (68.3-93.4)	85.1 (74.1-96.1)	85.1 (74.2-96)	87.2 (77.9-96.7)	87.2 (77.1-97.3)

Table 4.2: Direct Ratio method's sensitivity and specificity for different threshold values



The misclassified pathological cases include the following:¹

- A 15 year old girl with pulmonary stenosis (Rating = $\frac{3}{6}$ (S)).*
- A 10 year old girl with mitral incompetence (Rating = $\frac{4}{6}$ (S)).*
- A 12 year old boy with mitral incompetence (Rating = $\frac{3}{6}$ (S)).*
- A 2 year and six month old boy with tricuspid hypertrophy (*Glen shunt*). Post evaluation show that recording is 40 % irregular with in- and exhalation. (Rating = $\frac{2}{6}$ (S)).
- A 10 year old girl who had a pulmonary valve replaced. (*Glen shunt*) (Rating = $\frac{2}{6}$ (S) & $\frac{2}{4}$ (D)).
- A 12 year old girl with mitral incompetence (Rating = $\frac{1.5}{6}$ (S)).*

¹If -22.07 dB is used as the threshold value. A * indicates that the recording was also misclassified by one of the other two methods.

- A 3 year old boy with mitral incompetence (Rating = $\frac{2}{6}$ (S)).

The misclassified normal and innocent cases were due to (i) a too large energy content in the early- or mid-systolic region; (ii) an insufficient S1 sound; or (iii) a too irregular heart rhythm.

4.2.1.2 Discussion

In this initial algorithm, only the time-dependent intensity of the murmur was used as an indicator of whether the recording is pathological or not. Four of the seven misclassified patients were cases of mitral incompetence (MI). A possible explanation for this is that the intensity and duration of the systolic murmur because of MI is absolutely no indication as to the severity of the regurgitation [55]. If the left ventricle is larger than normal the murmur is a very short late systolic murmur, thus confined to a short time period and averaged to a low energy value.

The overall results obtained using this limited method are however encouraging, and show that systolic energy present in pathological murmurs can be automatically detected. It also re-affirms the assumption that murmur intensity correlates with the likelihood of pathology. Individual murmur timing (location) differentiation was done by consulting the envelope of the four systolic regions. This was useful in differentiating between the different types of systolic murmurs. Diagnosing possible differentiation in pathology was however not performed, due to limited clinical knowledge and experience. Potentially this might serve as an additional tool to assist the primary physician in differentiating between certain pathologies.

The obtained results call for the other indicators of pathology, mentioned by McCrindle *et al* [36], to be investigated. Combining the extraction of additional characteristics with the methodology used in the Direct Ratio algorithm might increase the sensitivity and specificity of an automated algorithm. The next subsection analyzes whether this is the case.

4.2.2 Wavelet processing results

The result of separating normal heart sounds from pathological heart sounds by frequency band limited energy values are shown in Figure 4.5. Thompson [6] suggested the use of the 2nd order coiflet wavelet. Analysis, however, has shown that the db4 wavelet produce better sensitivity and specificity. Results for the three opti-

num filters (scales) tested are shown, for analysis with the db4 wavelet. Although the energy values for normal and pathological cases overlap considerably for the lower scales tested, the overlap decreases for scale = 64.

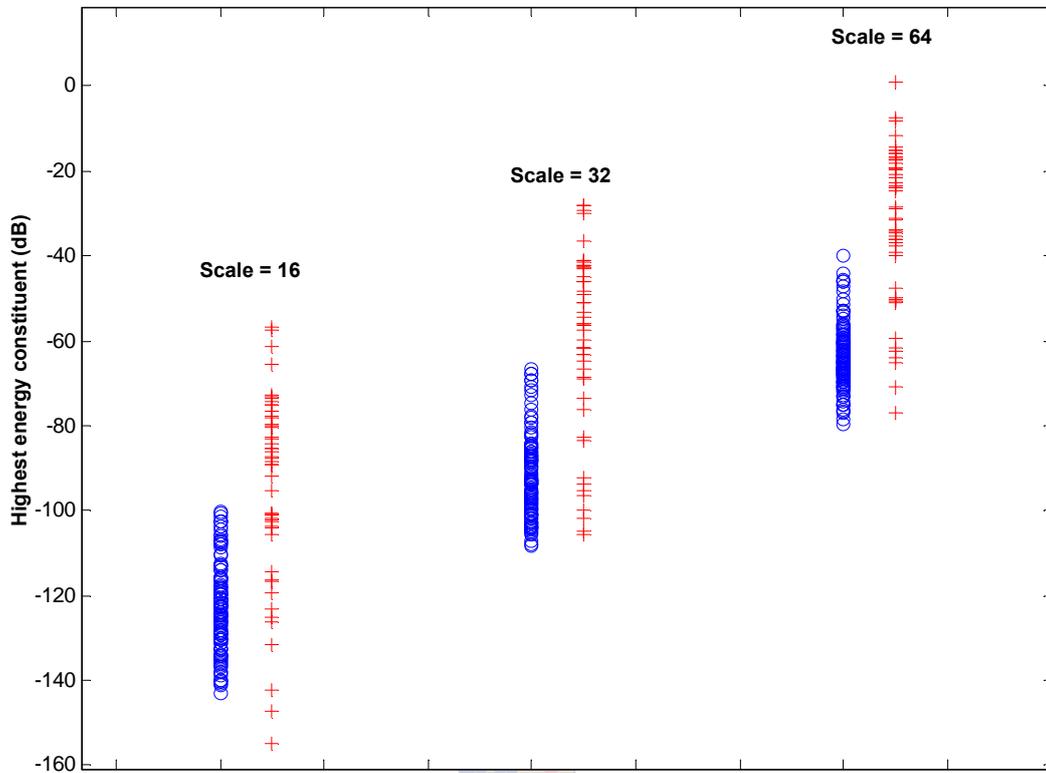


Figure 4.5: Comparison between relative energy content for different scales tested. Only the highest energy constituent is plotted for each recording. A blue circle is a no disease case and a red cross is a pathological case

Figure 4.6 illustrates the *Shapiro-Wilks'* W test result for the test of normality on the two populations. With a p -value of $p < 0.05$ considered significant - the no disease population is classified as normal with a normality of $W = 0.9745 (p = 0.029)$. With a calculated normality of $W = 0.916 (p = 0.00173)$ the pathological population distribution is classified as a *Student t* distribution.

Testing whether the null hypothesis can be rejected, for the mean difference between the two populations (scale = 64), can be done by following the six steps described in section 3.5.3.

Test the hypothesis for mean difference between the population mean of the no-disease data-set and the pathological data set using equation 3.5.7, 3.5.9 and 3.5.11:

$$\Rightarrow s_{\bar{x}_1 - \bar{x}_2} = 1.9861$$

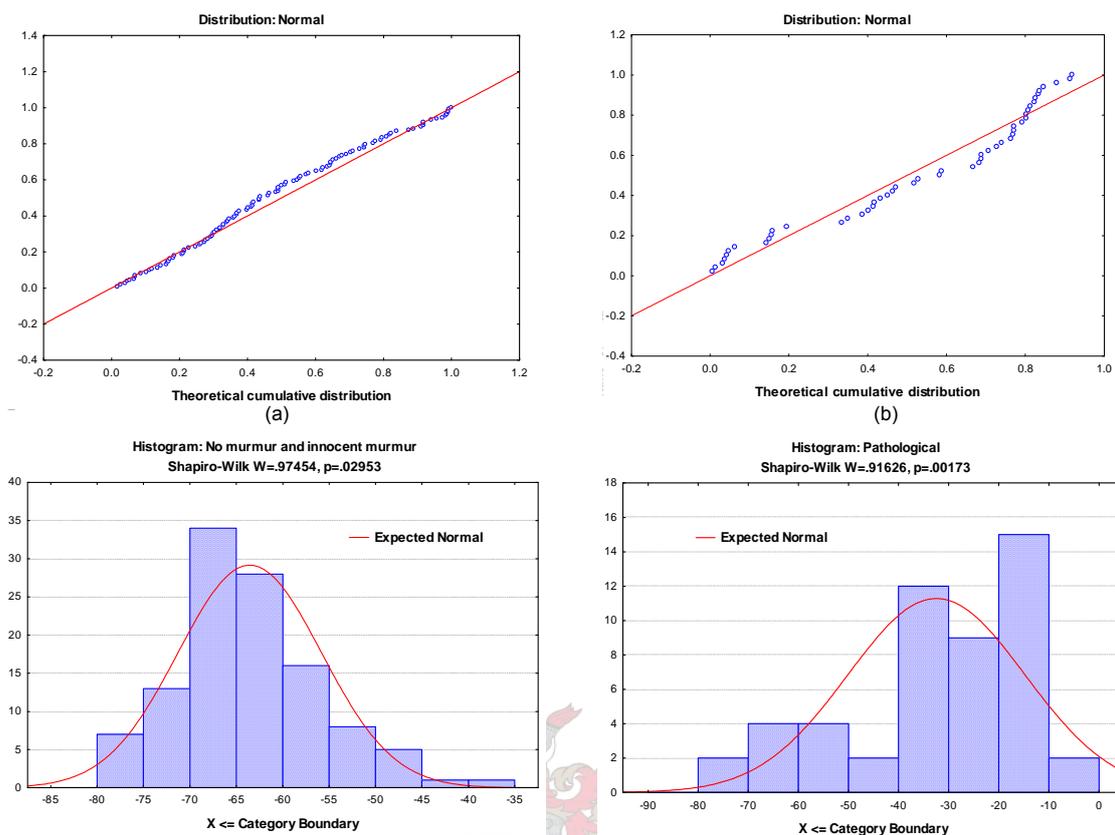


Figure 4.6: (a)&(b) illustrate the comparison between an actual normal distribution and the distribution of the no disease and pathological population respectively (c) & (d) illustrate the histogram distribution of the populations with their accompanied Shapiro Wilk W-test results.

$$\Rightarrow 27.0468 < \mu_1 - \mu_2 < 35.1579$$

$$\Rightarrow Z_{test} = 15.6603$$

The null hypothesis can be rejected because $Z_{test} > 1.96$. This means that there is a statistically significant difference between the mean of the no-disease population and the pathological population. The same conclusion is drawn as during the hypothesis testing of the Direct Ratio method. The descriptive statistics obtained with scale = 64 are shown in Figure 4.7 and Table 4.3.

4.2.2.1 Sensitivity and Specificity

To compare the performance of the different scales relative to each other a receiver operating characteristic (ROC) curve is drawn in Figure 4.8. It is evident from this analysis that scale = 64 produced the best performance, with an optimum sensitivity and specificity of 86.28 % and 92,11 % respectively at an energy threshold value of

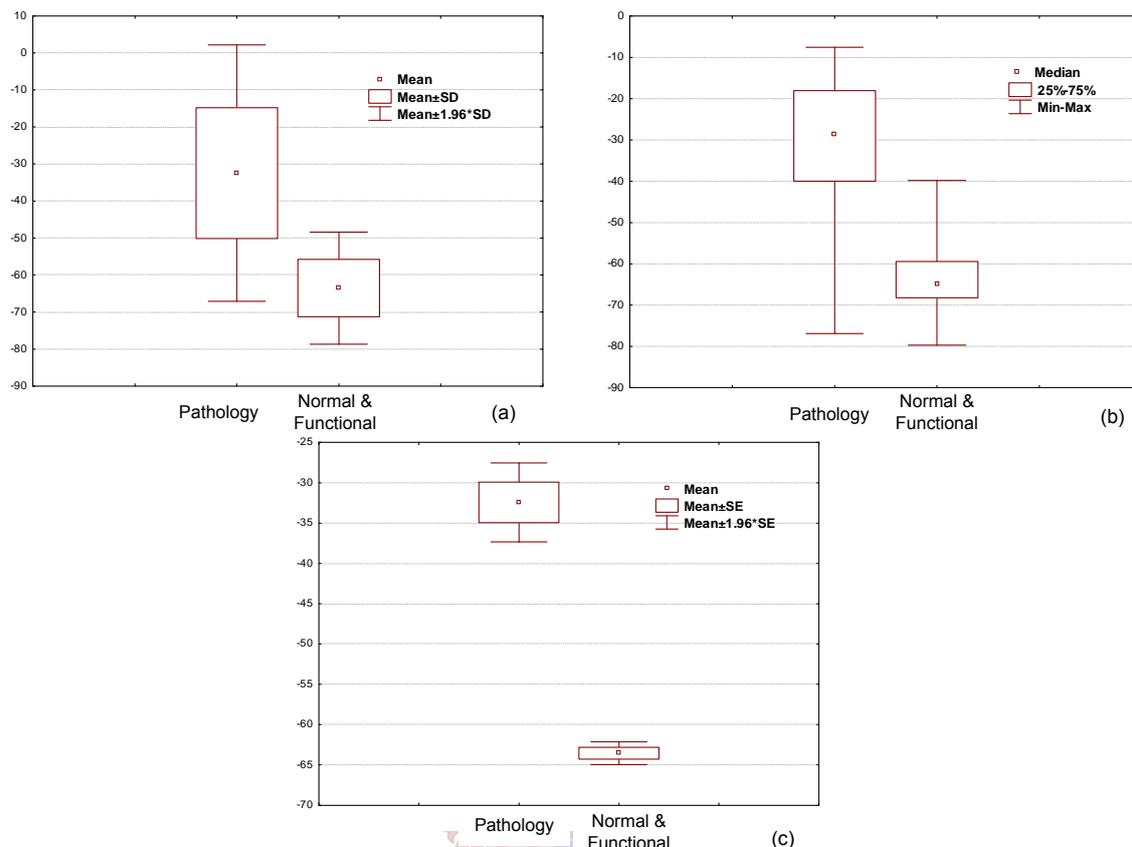


Figure 4.7: (a), (b) and (c) show the descriptive statistics for the Wavelet analysis technique with scale = 64 and wavelet db4.

	n	Mean	-95% CI	95% CI	SD	Standard Error	Minimum	Maximum
No disease	116	-63.55	-64.99	-62.11	7.73	0.727	-79.68	-39.8
Pathological	47	-32.45	-37.47	-27.42	17.68	2.5	-76.85	-7.51

Table 4.3: Descriptive statistics for the Wavelet analysis technique with scale = 64 and wavelet db4

-51.3 dB. Refer to section 3.5.4 in the previous chapter for a specific energy threshold choice.

The misclassified pathological cases include the following:²

- A 16 year old girl with mitral incompetence (Rating = $\frac{1}{6}$ (S)).
- A 10 year old girl with mitral incompetence (Rating = $\frac{4}{6}$ (S)).*
- A 12 year old boy with mitral incompetence (Rating = $\frac{3}{6}$ (S)).*

²If -51.3 dB is used as the threshold value. A * indicates that the recording was also misclassified by one of the other two methods.

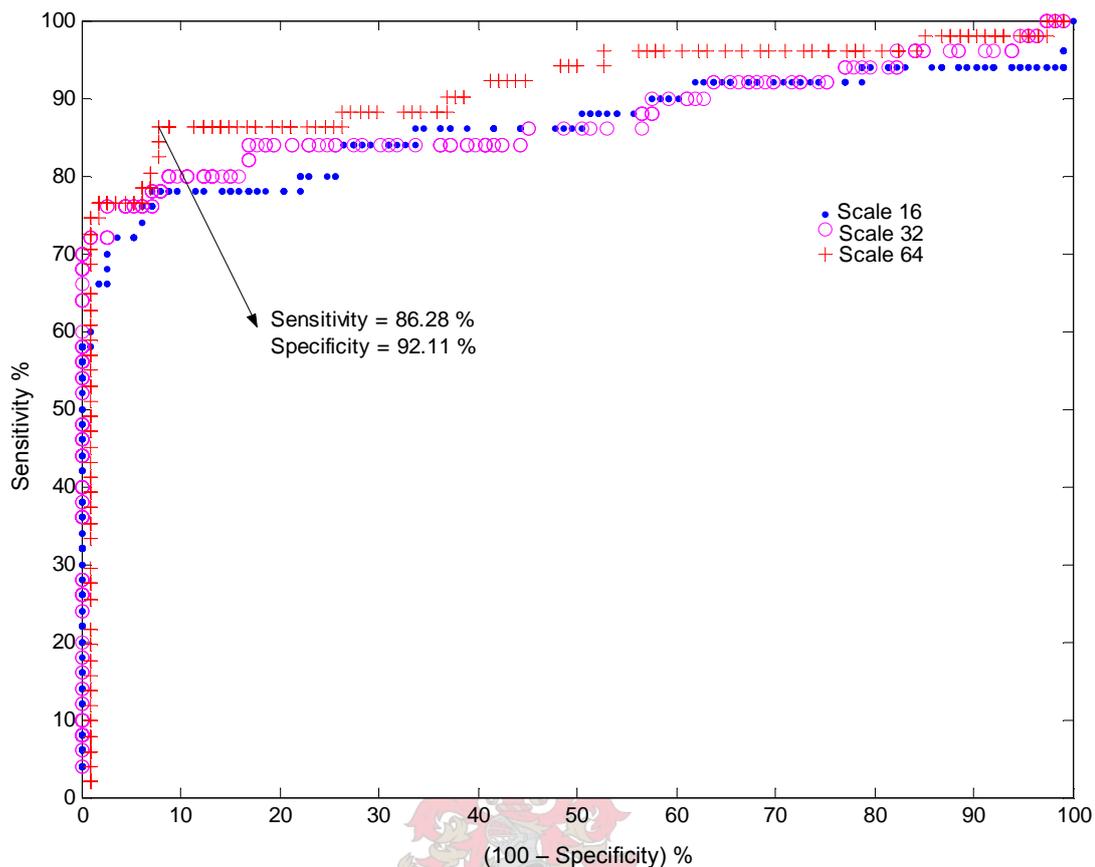


Figure 4.8: Receiver operating characteristics curve for classification of pathological or normal systolic heart murmur. Thresholds shifted from the minimum value in population to maximum value in the population. Data points are the corresponding sensitivity and specificity for each threshold, for different scales indicated.

- A 2 year and 3 month old boy with pulmonary stenosis (Rating = $\frac{1}{6}$ (S)).
- A 12 year old girl with mitral incompetence (Rating = $\frac{1.5}{6}$ (S)).*
- A 3 year old girl with a VSD (Rating = $\frac{3}{6}$ (S)).

The misclassified normal and innocent cases were due to the same reasons stated during the Direct Ratio explanation.

4.2.2.2 Discussion

In this section automated spectral analysis was evaluated to assess with the differentiation between pathological and normal digitized heart sounds. The performance of the wavelet analysis method compared well to the Direct Ratio method's sensitivity and specificity. At this stage the two algorithms, which are designed to screen

for only three of the six cardinal clinical signs of pathology listed by McCrindle *et al* [36], performs acceptably against a diverse group of pathologies.

Three of the MI cases missed by the Direct Ratio were again misclassified by the Wavelet method. This can be ascribed to the fact that both methods use the same constituent segmentation algorithm. The limitations of the Wavelet analysis method are the same as the limitations listed for the Direct Ratio method. Additional pathology indicators present in separate scales (frequency banks) were neglected. Analysis was only done on one scale at a time, meaning that if the indicator of pathology lies in the scale = 16 frequency bank, and the recording has normal frequency characteristics in the other frequency banks, then the recording will be misclassified, if one of the other frequency banks was used as the indicator. The next method seeks a manner to analyze the whole frequency spectrum at once for possible indicators of pathology. The first two methods developed was designed to detect pathological systolic murmurs and not all pathological lesions. Many examples of severe heart defects exist without a related systolic murmur. The next method aims to detect murmurs of all types.

4.2.3 Artificial Knowledge Based Neural Network results

The neural network described in section 3.4.3 was trained, tested and validated on several training- and validation data-sets to test the various recognition capabilities of the network. Both systolic and diastolic pathological murmur cases were included in the pathological data-set. This section displays the results in graphical and table format for all the different test setups.

4.2.3.1 Validation on three periods

Due to the limited number of recordings, *the validation on three periods* test was developed to evaluate the trained network's ability to recognize patterns that have been included in the training data-set. The same recordings were not used for both the training and the validation data set, instead only the last three periods of each recording were used as validation data while the rest of the periods served as training data. This method test the network's ability to classify a recording, when represented with 3 periods, as normal or pathological, if the other Σ (periods in recordings) - 3 periods of the same recording were part of the training data-set.

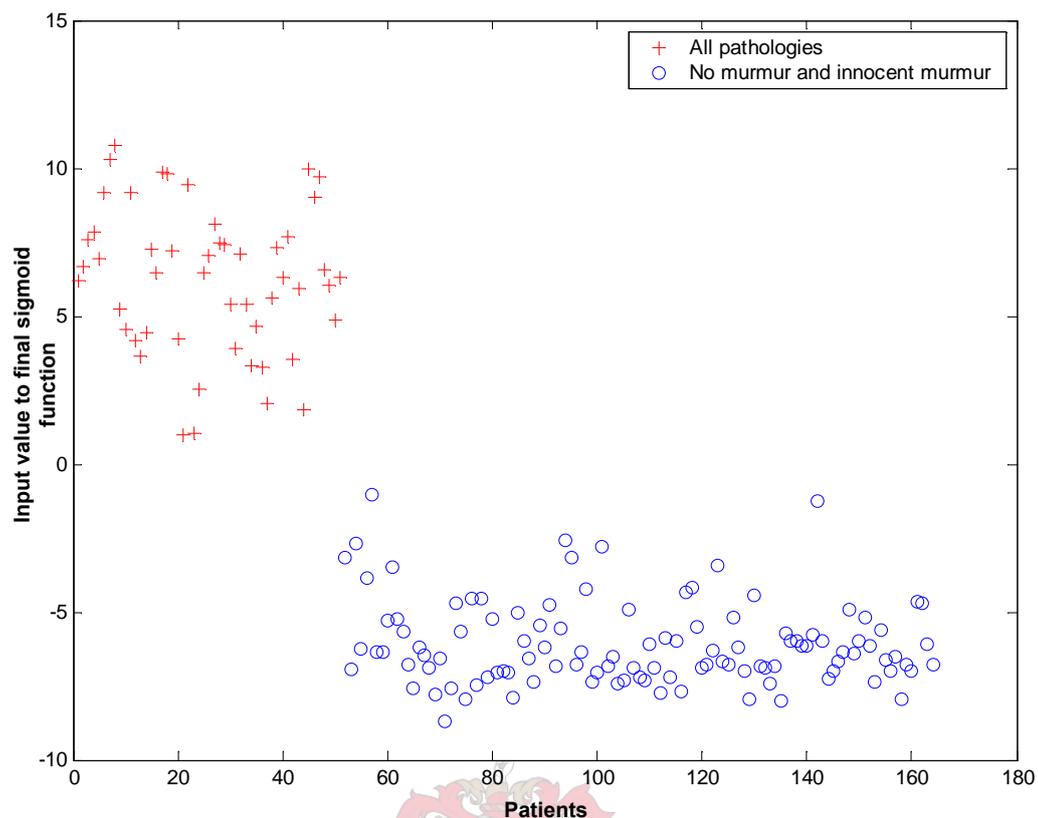


Figure 4.9: Feed-forward neural network. The average of the input value to the last *sigmoid* function of the three validation periods per patient.

The network was trained to output a 1 if the input vector represents a pathological case and a 0 if the input vector represents normal or innocent murmur case. The three validation periods' output were averaged and then plotted to represent a signal recording (patient). Figure 4.9 show the input value of the final *sigmoid* function for all 163 recordings. This graph verifies that the classification process is being performed by the first layer and the weighted summation of the second (output) layer, and that the final *sigmoid* function only needs to perform a bounding function to either of the boolean values.

Figure 4.10 illustrates the *Shapiro-Wilks'* W test result for the test of normality on the two data sets. With a p -value of $p < 0.05$ considered significant - the pathologic data-set are classified as normal with a normality of $W = 0.977(p = 0.429)$. According to the W -test the no disease (no-murmur and innocent murmur) data-set is not a normal distribution. With a calculated normality of $W = 0.91193(p = 0.00001)$ the data-sets distribution is classified as a *Student t* distribution.

Testing whether the null hypothesis can be rejected for the mean difference between the two data-sets can be done by following the six steps described in section

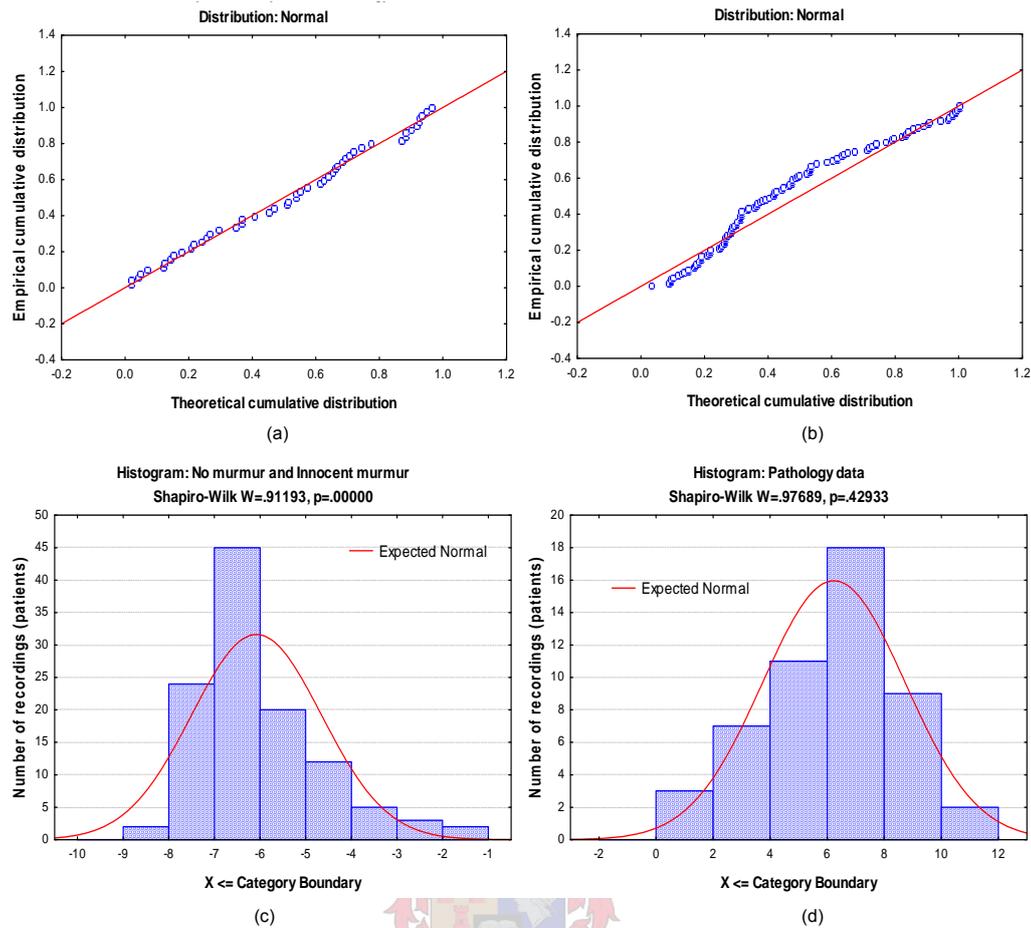


Figure 4.10: (a) & (b) illustrate the comparison between an actual normal distribution and the distribution of the two data-sets. (c) & (d) illustrate the histogram distribution of the data-sets with their accompanied Shapiro-Wilk W-test results

3.5.3.

Test the hypothesis for mean difference between the population mean of the no-disease data-set and the pathological data-set using equation 3.5.7, 3.5.9 and 3.5.11:

$$\begin{aligned} \Rightarrow s_{\bar{x}_1 - \bar{x}_2} &= 0.3715 \\ \Rightarrow 11.5819 &< \mu_1 - \mu_2 < 13.0382 \\ \Rightarrow Z_{test} &= 33.1349 \end{aligned}$$

The null hypothesis can be rejected because $Z_{test} > 1.96$. This means that there is a statistically significant difference between the mean of the no-disease population and the pathological population. The same conclusion is drawn as during the hypothesis testing of the Direct Ratio method.

Although the sensitivity and specificity of the above method is 100%, other training data-sets were tested to check whether the mean difference between the two groups

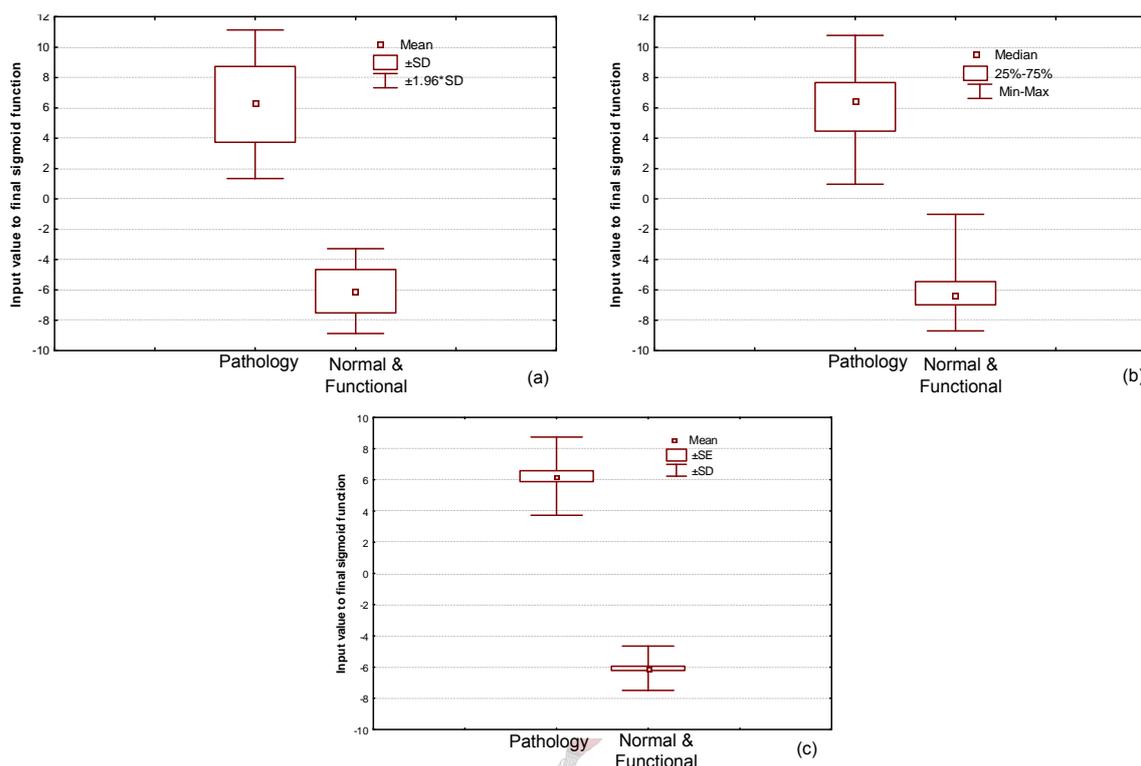


Figure 4.11: Descriptive statistics for the input data to the final *sigmoid* function (in the output layer) of the neural network

	n	Mean	-95% CI	95% CI	SD	Standard Error	Minimum	Maximum
No disease	113	-6.08	-6.34	-5.81	1.43	0.134	-8.69	-1.02
Pathological	50	6.24	5.52	6.94	2.49	0.35	0.96	10.77

Table 4.4: Descriptive statistics for three validation period Neural network method

can be improved. The training and validation data-sets used in the above method were de-noised during pre-processing using the fixed threshold wavelet de-noising algorithm discussed in section 3.3.1.4. The input values to the final *sigmoid* function, illustrated in Figure 4.13, show a decrease in mean difference between the two data-sets. Figure 4.14 and Table 4.6 illustrate the resulting decline in both the sensitivity and the specificity of the classifier.

4.2.3.2 Discussion

The three-period validation method illustrates that the neural network developed can classify three-period recordings, with a 100 % sensitivity and specificity, if the residual periods of the same recordings were included in the training data. Thus, if there is a good enough correlation between all murmurs of a certain type (because

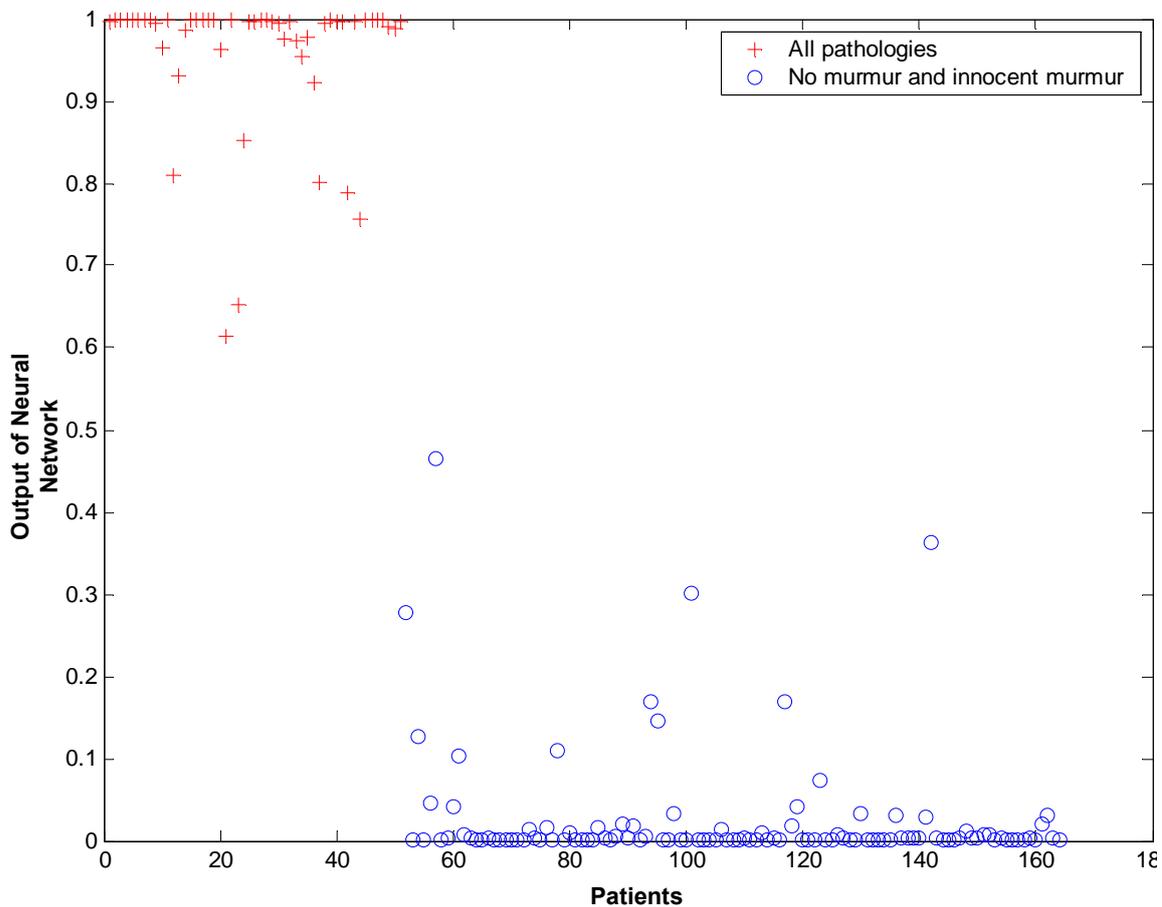


Figure 4.12: Output of the neural network. For the average of the three validation periods per patient

	n	Threshold = 0.5
Specificity (95 % CI)	113	100 (100-100)
Sensitivity (95 % CI)	50	100 (100-100)

Table 4.5: Sensitivity and specificity for the neural network 3-period validation method

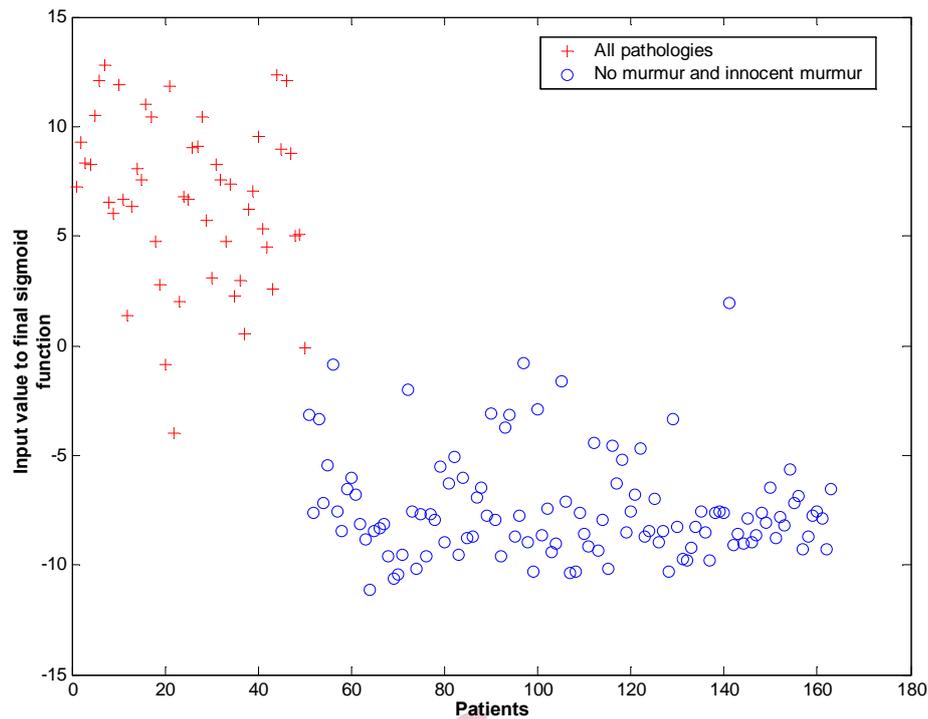


Figure 4.13: Input values to the last layer's function - average of three periods per patient with de-noised validation data input

of a certain condition), and all possible murmur types (conditions) are included in the training data set of the network, then the network will be able to classify any recording with a 100 % sensitivity and specificity. To verify the usefulness of the network, the first question to be answered is thus: *What is the correlation between a specific murmur type recorded from different patients with the same condition?* And the second question is: *How many differentiable murmur types are there?* If the clinical answer to these two questions are that there exists an adequate correlation between murmurs of the same type, and that there are only a certain amount of murmur types, then this method holds promising prospect for automated classification. If the answers to these questions, however, differ from the above-mentioned, the motivation for using a neural network as a classifier is lost. The next training method employed, aspires to verify how well the neural network performs when validated with recordings never seen before.

Another issue that surfaced during the network training and validating with both the de-noised and the impure data-sets is the decrease in performance when using the de-noised recordings. Section 3.3.1.4 states that determining the threshold level for each decomposition level is done by attempting to meet two criteria: (i) to remove as much of the noise as possible; (ii) without losing any information. By

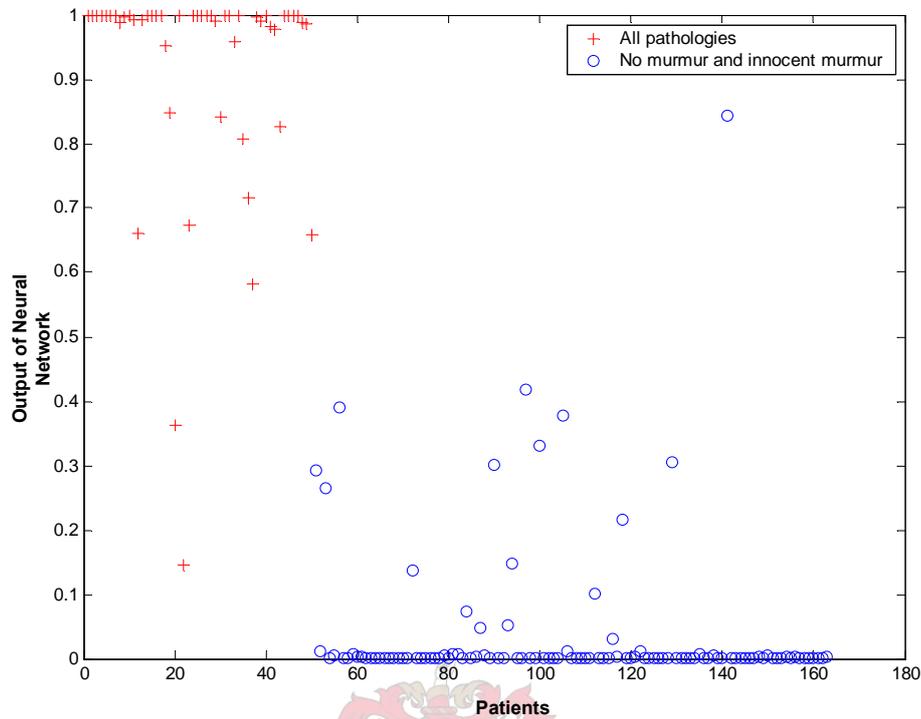


Figure 4.14: Output of the neural network - average of three periods per patient with de-noised validation data input

	n	Threshold = 0.5
Specificity (95 % CI)	113	99 (97.4-100)
Sensitivity (95 % CI)	50	96 (90.5-100)

Table 4.6: Sensitivity and specificity of the 3-period validation method neural network with de-noised training- and validation data input

comparing the network prediction variance in Figures 4.9 and 4.13, and the sensitivity and specificity in Tables 4.5 and 4.6, it was shown that information bearing data had been lost during the implementation of the fixed wavelet de-noising algorithm. Thus, according to this analysis criteria (ii) was not met during the de-noising algorithm.

4.2.3.3 *Jack-Knife* neural network

The second test used to evaluate the recognition capabilities of the developed neural network is the *Jack-Knife* method. The *Jack-Knife* method is an iterative process in which one recording is recruited for validation at a time [37, 56]. The neural network is trained using the remaining data and is validated on the single, left-out validation recording. When using the *Jack-Knife* method the classifier does not see the validation recording during its training, this ensures that the evaluation is unbiased. This approach measures the power of the generalization of the classification process rather than of one specific classifier.

Each of the 163 recordings was recruited for validation, one at a time, creating 163 separate trained networks. The network consists of exactly the same architecture, for each iteration. See Appendix E.5 for the code listing of the *Jack-Knife* training dataset composition, the *Jack-Knife* training and validation and the *Jack-Knife* simulation and testing.

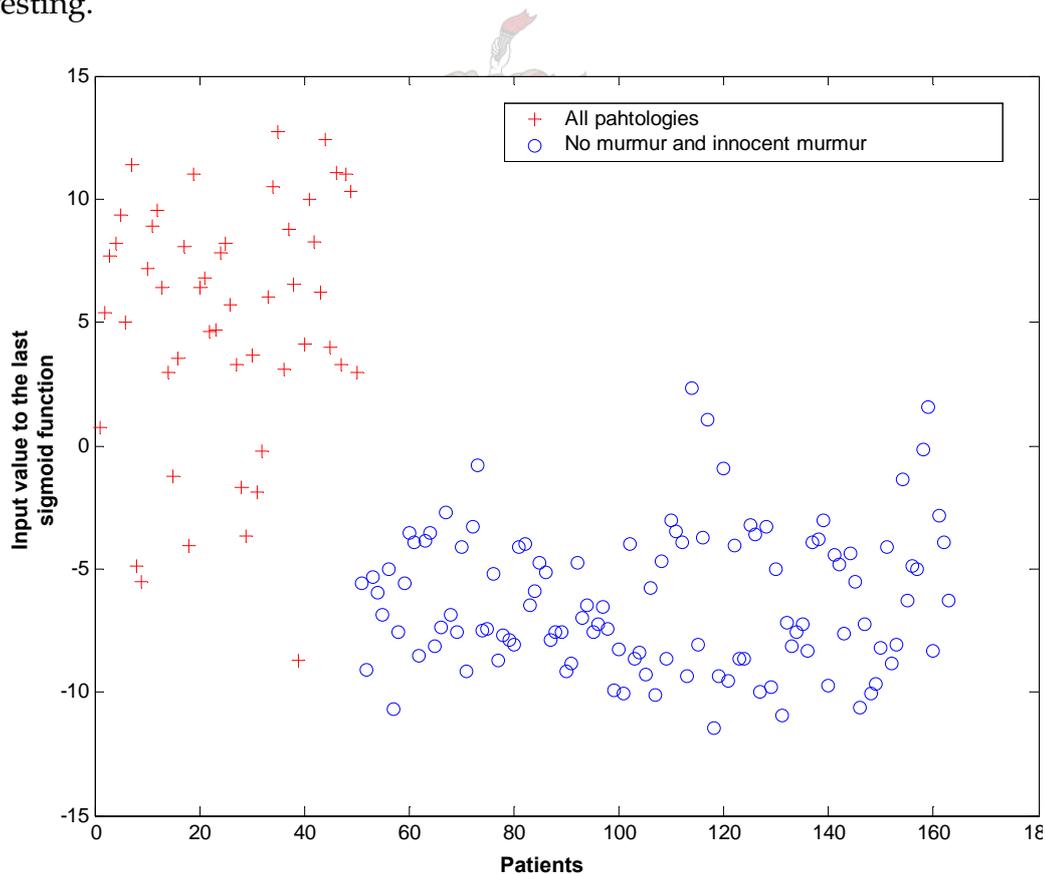


Figure 4.15: *Jack-Knife* training method: Input value to the final *sigmoid* function. The average of six periods per recording (patient) is plotted

The *Jack-Knife* method was tested with three different settings of periods-per-recording.

Four, six and eight periods-per-recording were respectively included in the training and validation data-set for each recording. The six period-per-recording performed statistically the best, if the mean and variance of the two patient groups are used as measurement. Figure 4.15 show the average input values to the final *sigmoid* function for the six-period *Jack-Knife* training method. This graph is drawn to calculate the distribution and descriptive statistics of the two patient groups.

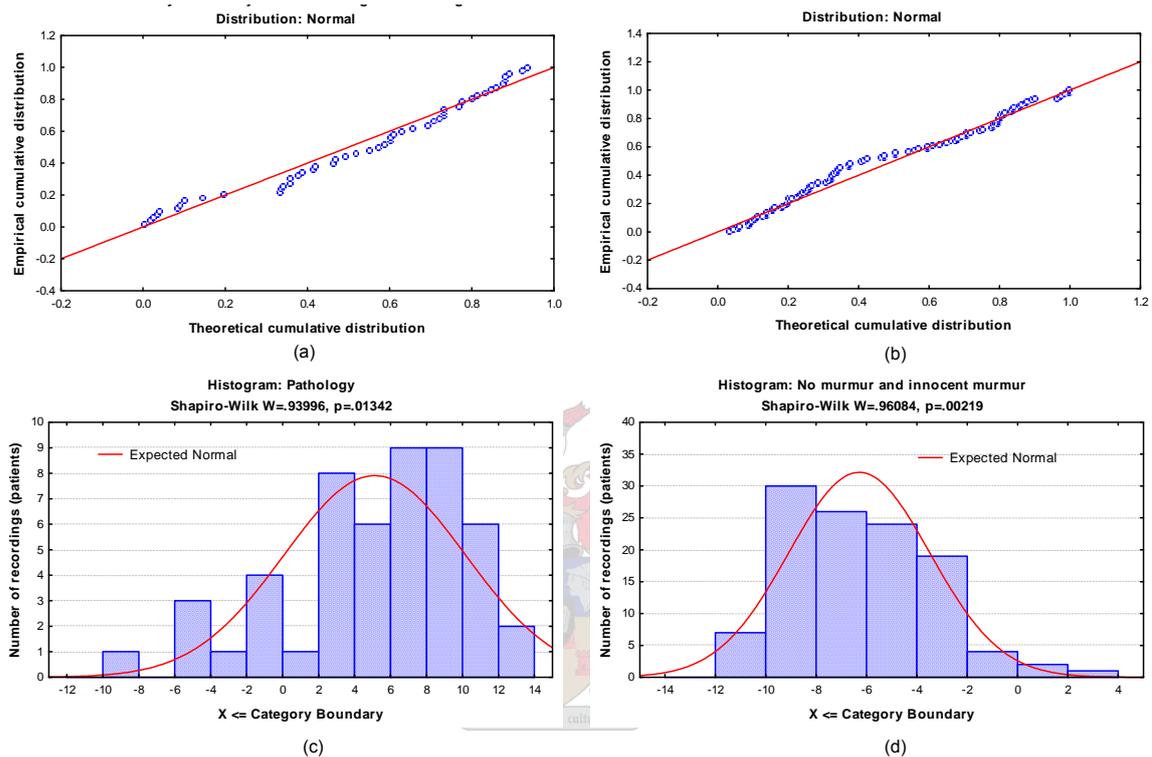


Figure 4.16: Distribution statistics for the six-period *Jack-Knife* training method

Figure 4.16 illustrates the *Shapiro-Wilks' W* test result for the test of normality on the two data sets. With a p -value of $p < 0.05$ considered significant and a calculated normality of $W = 0.94(p = 0.0134)$ and $W = 0.961(p = 0.00219)$ respectively, neither of the data-sets can be classified as a normal distribution. Accordingly both data-sets were classified as a *Student t* distribution, with descriptive statistics as shown in Table 4.7.

Testing whether the null hypothesis can be rejected for the mean difference between the two data-sets can be done by following the six steps described in section 3.5.3.

Test the hypothesis for mean difference between the population mean of the no-disease data-set and the pathological data-set, using equation 3.5.7, 3.5.9 and 3.5.11:

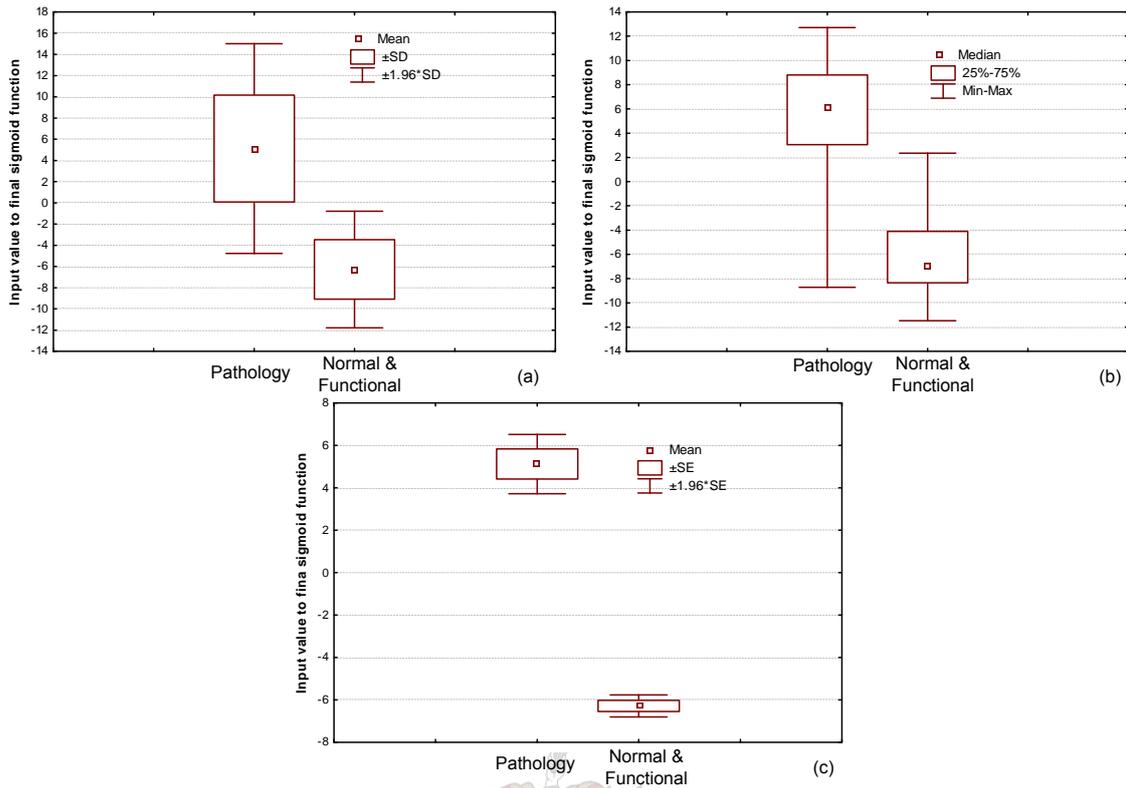


Figure 4.17: Descriptive statistics for the six-period *Jack-Knife* method

	n	Mean	-95% CI	95% CI	SD	Standard Error	Minimum	Maximum
No disease	113	-6.29	-6.81	-5.76	2.8	0.264	-11.48	2.34
Pathological	50	5.12	3.69	6.56	5.04	0.71	-8.72	12.72

Table 4.7: Descriptive statistics for the six-period *Jack-Knife* neural network method

$$\begin{aligned} \Rightarrow s_{\bar{x}_1 - \bar{x}_2} &= 0.3715 \\ \Rightarrow 11.5819 &< \mu_1 - \mu_2 < 13.0382 \\ \Rightarrow Z_{test} &= 33.1349 \end{aligned}$$

The null hypothesis can be rejected because $Z_{test} > 1.96$. Although it is evident from viewing Figure 4.15 the hypothesis can now be adapted that there is a statistically significant difference between the intensity and the frequency content in a normal recording and that of a pathological recording which can serve as a measure for diagnosis.

Table 4.8 shows the results obtained with the *Jack-Knife* method with three different periods per recording settings. The six-periods setting has the highest sensitivity of the three settings tested. Figure 4.18 illustrates the results obtained with the six-period setting, with each period-energy-spectrum represented by a 120 data-point

vector. Concerning the choice of a specific threshold value, the same argument is used as stated during selection of a threshold value for the Direct Ratio method. An example threshold value of 0.26 displayed acceptable results to both sides for the available data set used.

	n	Number of training and validations periods per recording (Threshold value = 0.26)		
		4	6	8
Specificity (95 % CI)	113	93 (88.2-97.6)	94 (89.3-98.3)	95 (90.6 - 98.8)
Sensitivity (95 % CI)	50	84 (73.8-94.2)	88 (79.1-96.9)	82 (71-93)

Table 4.8: The *Jack-Knife* method’s sensitivity and specificity for different training and validation periods per patient.

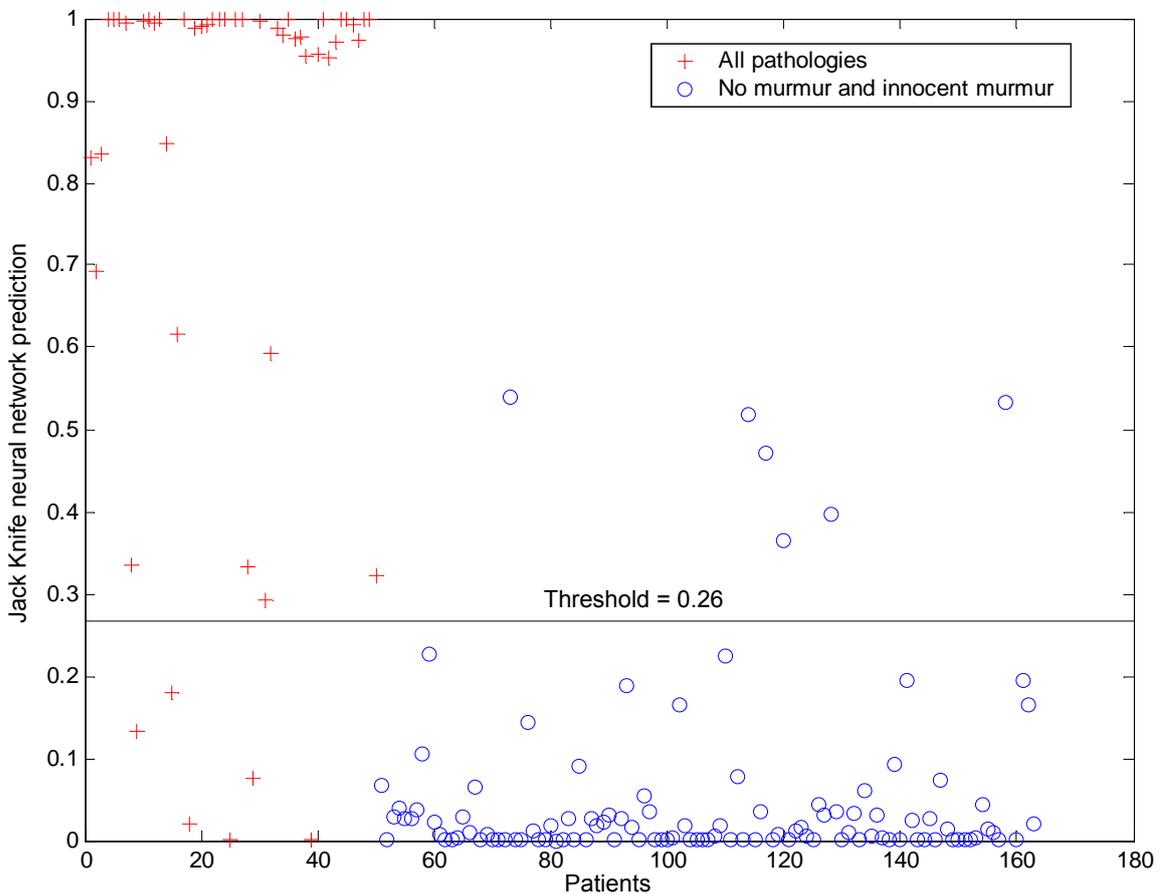


Figure 4.18: The *Jack-Knife* method’s classification results. Trained and validated with six periods per patient. Plotted prediction is the average of the six periods. The horizontal line represents the example decision threshold

Although the de-noising of the training- and validation data in the previous method had a negative effect on the performance of the network, the same procedure was

followed with the *Jack-Knife* method to check for correlated behavior. Figure 4.19 shows the average input value to the final *sigmoid* function obtained with the de-noised training and validation data sets.

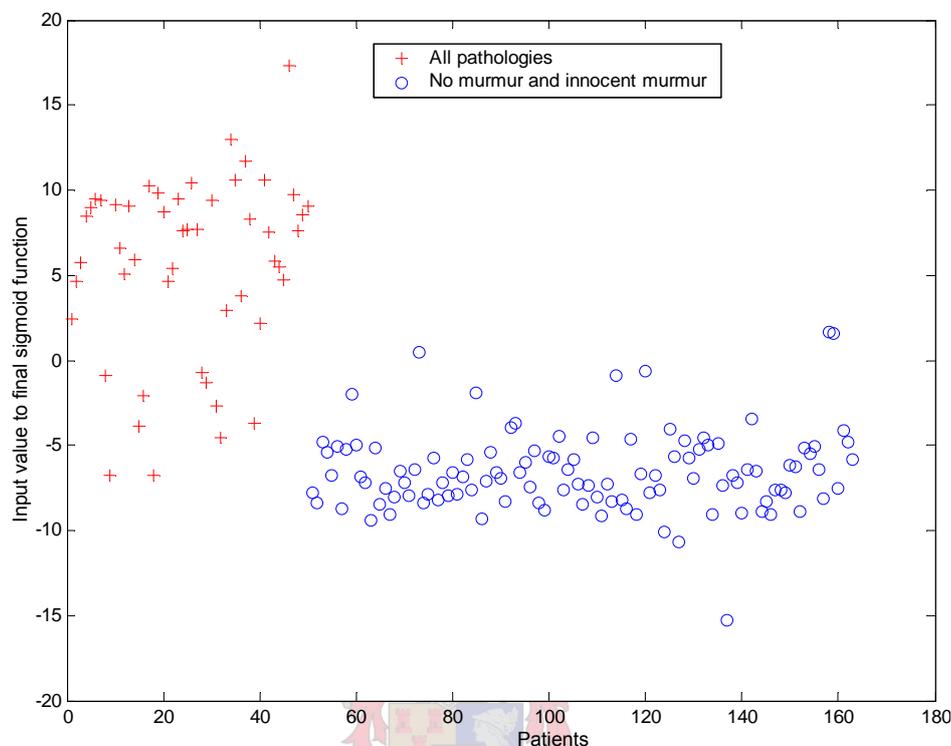


Figure 4.19: *Jack-Knife* de-noised training method: Input value to the final *sigmoid* function. Trained and validated with six periods per patient. Plotted prediction is the average of the six periods.

Figure 4.20 illustrates the *Shapiro-Wilks' W* test results. Both populations were classified as *Student t* distributions, using the same reasoning as with the normal *Jack-Knife* population. Comparisons between the descriptive statistics in Table 4.7 and 4.9, and Figure 4.17 and 4.21, show an increased performance in all the measurements for the *Jack-Knife* de-noised method. These results contradict the results obtained in the previous method; a possible explanation will be offered later.

Sensitivity and specificity for the two example threshold values drawn in Figure 4.22 are shown in Table 4.10. Concerning the choice of a specific threshold value, the same argument is used as during the selection of a threshold value for the Direct Ratio method.

The misclassified pathological cases using the de-noised *Jack-Knife* method includes the following:³

³If the more conservative threshold value of 0.26 units is used. A * indicates that the recording

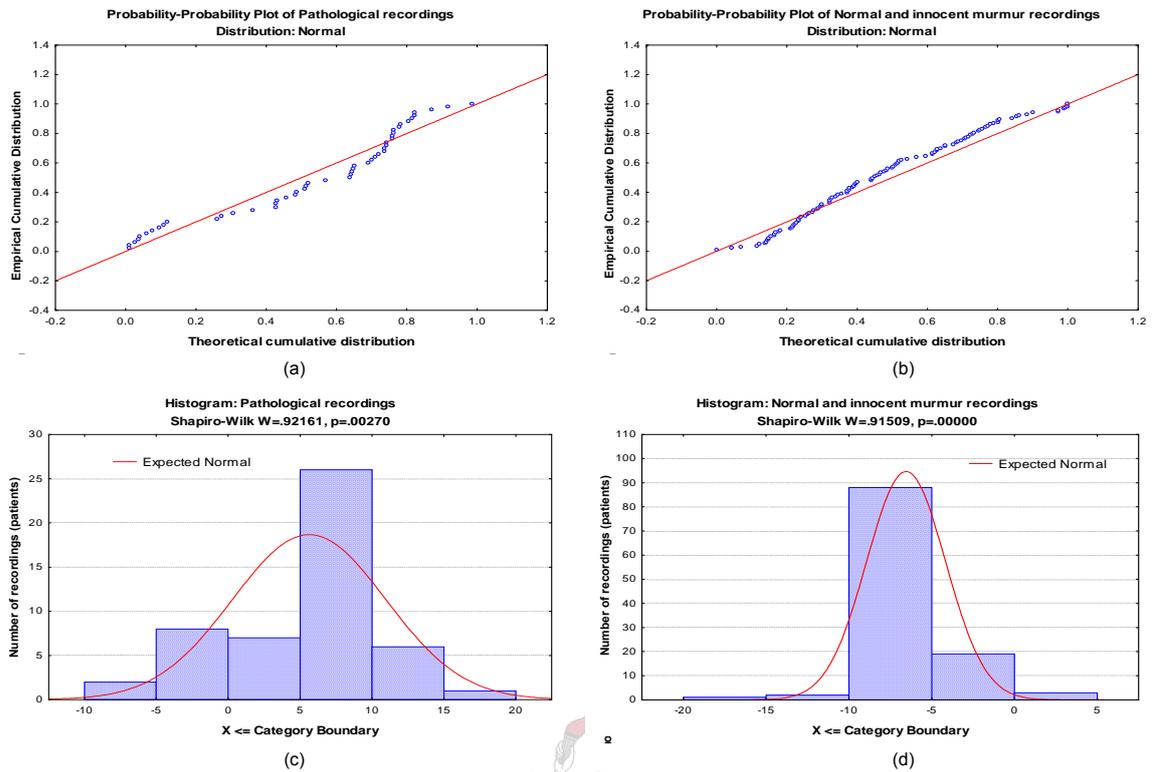


Figure 4.20: Distribution statistics for the *Jack-Knife* de-noised training method

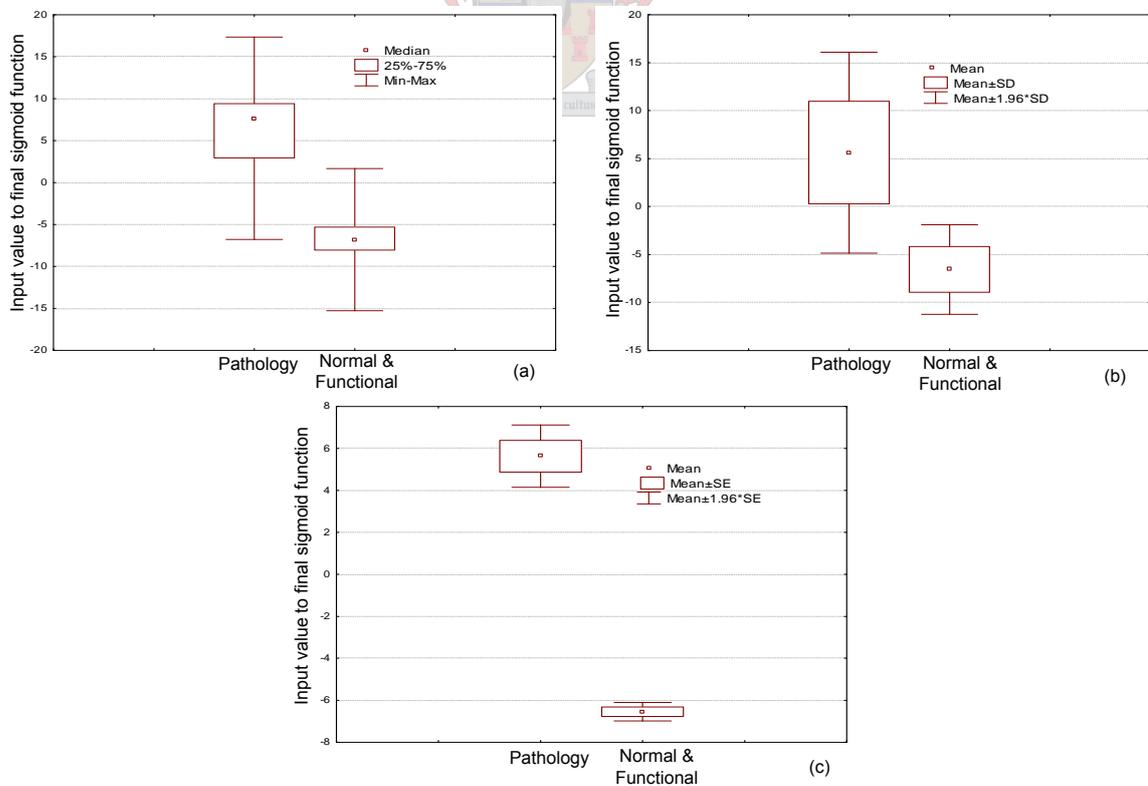


Figure 4.21: *Jack-Knife* de-noised method's data descriptive statistics

	n	Mean	-95% CI	95% CI	SD	Standard Error	Minimum	Maximum
No disease	113	-6.54	-6.99	-6.1	2.38	0.22	-15.29	1.67
Pathological	50	5.63	4.11	7.15	5.34	0.76	-6.78	17.34

Table 4.9: Descriptive statistics for de-noised *Jack Knife* neural network method

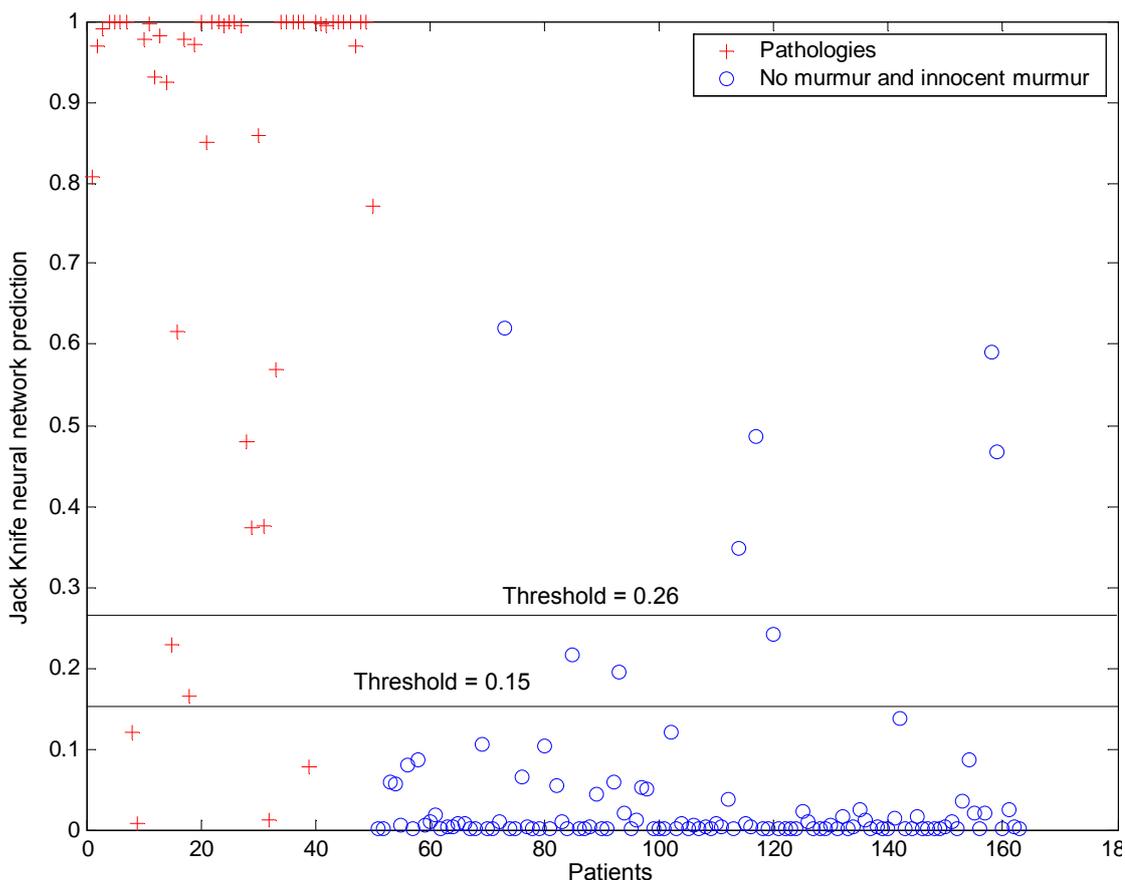


Figure 4.22: *Jack-Knife* de-noised method’s classification results. Trained and validated with 6 periods per patient. Plotted prediction is the average of the six periods

	n	(Threshold value)	
		0.26	0.15
Specificity (95 % CI)	113	95.6 (91.8-99.4)	92.9 (88.1-97.7)
Sensitivity (95 % CI)	50	88 (78.9-97.1)	92 (84.8-99.2)

Table 4.10: *Jack-Knife* de-noised method’s sensitivity and specificity for two different threshold values.

- A 3 month old baby boy with the Tetralogy of Fallot condition (Rating = $\frac{4}{6}$ (S)).
- A 10 year old boy with a stent between his pulmonary artery and aorta (Rating = $\frac{2}{6}$ (S)).
- A 2 year old girl with a primum ASD, MI and TI (Rating = $\frac{3-4}{6}$ (S)).
- A 15 year old girl with pulmonary stenosis (Rating = $\frac{3}{6}$ (S)).*
- A 5 year old boy with tricuspid and mitral incompetence (Rating = $\frac{3}{6}$ (S), $\frac{2}{4}$ (D)).
- A 1 year and 6 months old boy with pulmonary stenosis (Rating = $\frac{3}{6}$ (S)).

The misclassified normal and innocent cases were due to a large energy-content in the early- or mid systolic region (according to visual interpretation).

4.2.3.4 Discussion

According to section 3.4.3.4 the network's performance must increase with an increase in training data. Increasing the periods-per-patient from 6 to 8, however, has the opposite effect. After passing all the periods through the period-filtering algorithm the correlation between the periods are within 1.5 standard deviations of the average correlation. The decline in the network's performance can thus not be ascribed to a decrease in correlation, if the periods used are increased. (Except if the average correlation between the periods of each recording in the whole data-set is very low, which is not the case). A more acceptable hypothesis for the decline in the performance, is the increased computational load associated with an increase in training data. The training of the 163 networks with 6 periods took 14 hours and 23 minutes, while the training of the 8-period network took 20 hours and 37 minutes (this is with all the memory allocated to the *Matlab* computation process).

Misclassification is due to an under-representation of the misclassified murmur classes in the training data. Comparing the list of misclassified pathological cases with the database composition in Figure 4.1, show a definite correlation between the under-represented classes and the misclassified classes. The network's generalization would improve with better representation of all classes in the training data. The network however performed well with a wide range of murmur classes that was represented sufficiently in the training data-set. An example is the absence of the

was also misclassified by one of the other two methods.

MI cases misclassified by the other two methods. Due to a sufficient number of MI recordings in the data-set the network was able to classify all of them correctly.

Following from the discussion on the three-period validation neural network, the *Jack-Knife* method illustrates that a previously unseen murmur can be classified by a network trained on murmurs from the same class. With the limited training data-set used, an estimated answer, to the first question posed is that there is a sufficient correlation between murmurs of the same class, to justify a classification of previously unseen murmurs from a represented class. This only leaves the second question to be answered: *How many different murmur classes must be represented in the training data for a network to have a sensitivity and specificity of 100 % for all possible recordings?* The calculated specificity and sensitivity of 92.9 % and 92 % respectively show that representing all the murmur classes is not impossible, but will require more data capturing to be done from all types of congenital related pathologies. Only then can the *Jack-Knife* neural network be employed in the field as an automated classifier of heart sounds.

4.3 Simultaneous evaluation of all three methods developed

With all the statistics of the different methods developed in hand, which one is the most reliable method to use? The best performer is not indicated merely by the highest sensitivity and specificity, but will be the one with the optimum distribution and descriptive statistics, together with a good sensitivity and specificity. Comparing the distribution and descriptive statistics are however not possible because of the different measures used. An objective way of measuring performance between the different methods, is by drawing the receiver operating characteristic (ROC) curves of the various methods. The method's curve with the biggest area under the curve is the most reliable method to use. Figure 4.23 show the ROC curves of the four methods giving the best performance.

From Figure 4.23 it is evident that the de-noised *Jack-Knife* method produces the curve with the largest area under the curve. Keeping in mind that the Direct Ratio and Wavelet methods can only diagnose murmurs in the systolic region, while the *Jack-Knife* method has diagnosed the few diastolic-related pathologies present in the database successfully; the de-noised *Jack-Knife* method is the most reliable method to use.

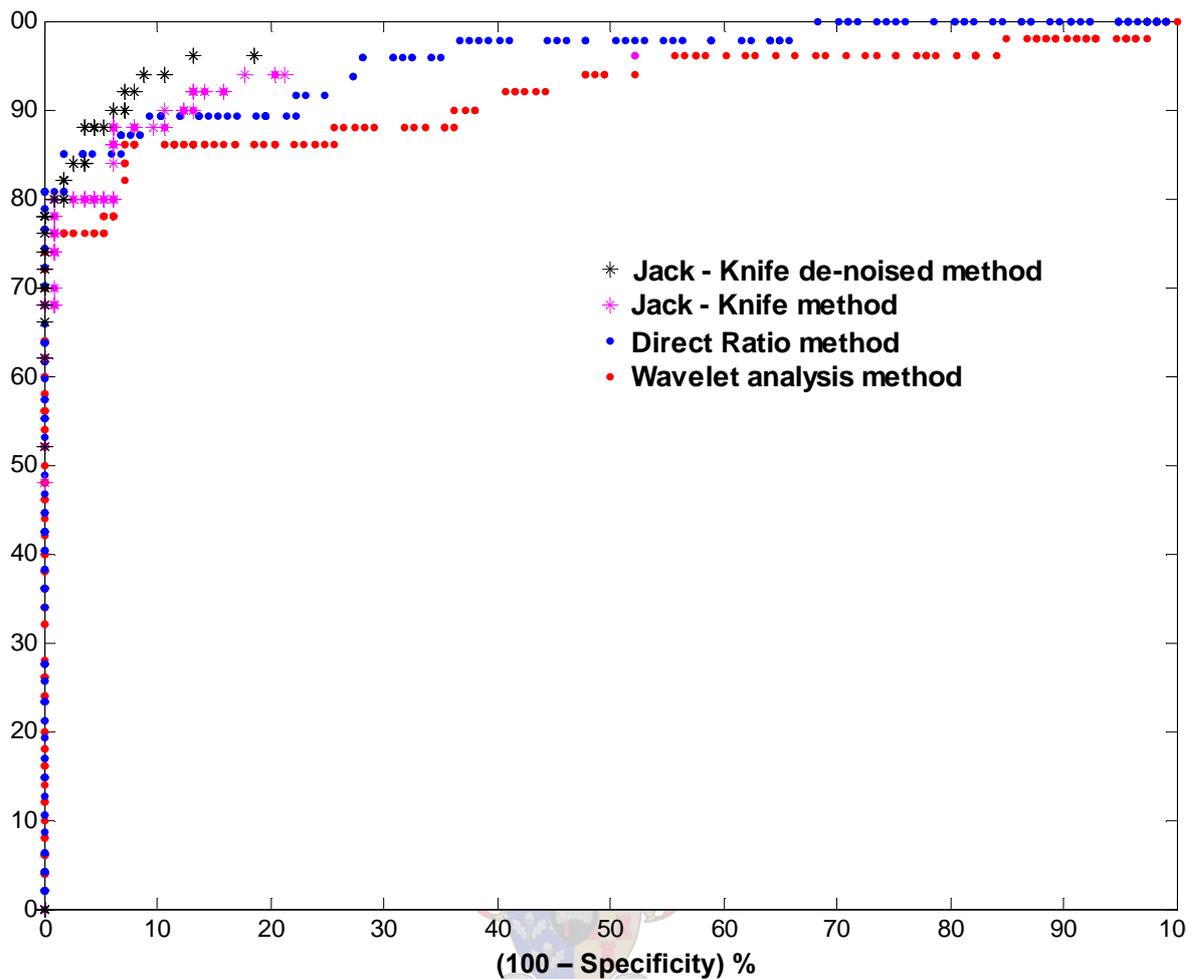


Figure 4.23: ROC curves for the four best performance methods. Legend show the indicators for the four methods.

Investigating the various misclassified cases for each method show that there is not one single recording that has been missed by all three methods evaluated. Thus to obtain a 100 % sensitivity, for the used database, one can put all three the methods in series. If one of the methods classify a recording as pathological, the recording is labeled as pathological with no influence from the other methods. This however, results in a considerable decrease in specificity, with a specificity of 85.8 %, which makes the series cascade method an unattractive option. The proposal is to gather enough patient recordings to train the de-noised neural network method sufficiently well, to represent all the murmur classes. ⁴

⁴Given that the second question posed in section 4.2.3.2 has the wanted clinical answer

Chapter 5

Conclusions, Limitations and Recommendations for Further Research

From the examination of previous attempts to develop an automated cardiac auscultation algorithm and thorough clinical observation, it was postulated that time-dependent pathological murmurs possess higher frequency and intensity auscultatory sounds than innocent murmurs. Three automated algorithms that show significant potential in their use as an alternative diagnostic tool for the classification of heart sounds into normal (innocent) and pathological classes were developed. Investigating various methods for extracting additional information from heart sounds demonstrated that the *De-noised Jack-Knife* neural network approach produces the best performance as an automated classifier, with an optimum sensitivity and specificity of 92.9 % and 92 % respectively.

Differentiating between different pathologies (types of murmurs) were only investigated with the *Direct Ratio* and *Wavelet* methods. These algorithms, however, are limited to identify only systolic murmurs. Pathology differentiation were met with limited success therefor the results were not included in this report.

The primary limitation of this project was the very restricted database of pathological and innocent murmur sounds. Results, obtained with this limited database, suggest that the generalization of the *De-noised Jack-Knife* would improve with adequate representation of all pathological classes. The most pressing need during future research will thus be to collect more data. Still lacking are validated answers to the two questions posed in the previous chapter: *What is the correlation between*

a specific murmur type recorded from different patients with the same condition? And secondly: If a murmur type is classified as a class of murmurs resulting from the same condition in different patients, and that correlate sufficiently well to represent each other, how many differentiable murmur types are there? Only after all the pathological classes are represented sufficiently, will the generalization of the network be of such a high precision that it can be field deployed.

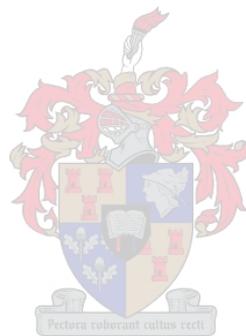
After thorough generalization of the *De-noised Jack-Knife* neural network the possibility of differentiation between conditions can be investigated. However, before adding any differentiation capabilities to the network, it is proposed to accommodate the "never seen before" case. There will always be "never seen before" murmurs, corresponding to either a certain recording position, a certain stage of the condition or a distinctive respiration frequency. To minimize the number of false negatives, the trained network must have the capability of distinguishing a murmur that correlates poorly with both populations. These murmurs must be evaluated separately to ensure that they are not misclassified.

Implementing the *De-noised Jack-Knife* algorithm as a primary screening device requires the algorithm to be implemented on a highly compact (portable) hardware device. This is possible with the newly developed DSP functionality embedded on the Altera Stratix FPGA [57]. The DSP functionality includes FFT calculations that are needed for the pre-processing of the neural network input data. After training the network on a personal computer it can be programmed onto the FPGA. In combination with a well trained network this proposed device can serve as an additional diagnostic aid that might assist the primary physician in the detection and evaluation of heart murmurs.

After the successful implementation of the trained neural network on the FPGA, the generalization of the network can be further improved by continually adding the classified (verified) "never seen before" cases, mentioned earlier, to the training dataset as they are encountered. [Where classified (verified) means - Diagnosed with a ultra-sound unit, and confirmed by a pediatric cardiologist] Further training should be administrated at a single centralized location, from where the other units in the field can then "update" their networks. Feedback, on the "never seen before" cases, from the field units can also indicate further possibilities for improvement.

In conclusion, the feasibility of using the developed algorithm as a diagnostic aid, additional to the historical and physical examination findings, depends on the general circumstances of the primary physician and the patient being evaluated. If the algorithm demonstrates a higher success rate than the current referral success rate

of the physician, then the automated auscultation algorithm might be considered as an additional evaluation technique. The proposed application is that the developed algorithm must serve as a rapid, low cost screening device that can be used by nurses to screen large numbers of children in a clinic or mobile-clinic environment. Children diagnosed as pathological can then be evaluated more carefully by a qualified physician.



Appendices



Appendix A

Information and informed consent document

INFORMATION AND INFORMED CONSENT DOCUMENT

Title of the Research Project: Digitally Automated Pediatric Cardiac Auscultation
Reference Number: N04/04/077
Principle Researcher: Jacques de Vos
Address: Department E&E Engineering
Room E245, Banghoekweg
University of Stellenbosch
Stellenbosch
7600

DECLARATION BY OR ON BEHALF OF PATIENT/*PARTICIPANT:

I, THE UNDERSIGNED, _____ (name)
[ID No: _____] the patient/*participant or* in my capacity
as _____ of the patient/*participant [ID No: _____]
of _____(address).

A. HEREBY CONFIRM AS FOLLOWS:

1. The following aspects have been explained to me/*the patient/*participant:

1.1 Aim: The aim of the project is to determine whether there is any additional information in a patient's heart sound that can be extracted to the benefit of the patient. To achieve this, heart

APPENDIX A. INFORMATION AND INFORMED CONSENT DOCUMENT 13

sound data is required for the research. Recordings done during the standard medical examination will be used for this purpose.

1.2 Procedures: There will be no deviation from the normal examination procedures during the data collection, except that the data measured by the stethoscope, EKG and echocardiogram will be recorded. The medical examination will continue as usually. The ECG recording will be done using three ECG electrodes that are placed on the patient's chest, while the heart sound recording will be done with an electronic stethoscope. Participation in this study will not result in any additional costs.

1.3 Alternatives: There are no alternatives at this moment.

1.4 Risks: There is no risk in participating. All the equipment used is medically approved.

1.5 Possible benefits: Patients involvement can lead to better and more affordable examination procedures in the future.

1.6 Confidentiality: The patient data and information extracted from the data will be treated as confidential. Patient identity will not be made public during any publication that might lead from the research.

1.7 Access to findings: Access to the research outcome can be obtained after the conclusion of the project.

1.8 Voluntary participation/refusal/discontinuation: Participation is absolutely voluntary, and participation can be refused.

2. The information above was explained to me/*the patient/*participant by _____ (name of relevant person) in Afrikaans/*English, and I am/*the participant/*patient is in command of this language. I, the participant/*patient was given the opportunity to ask questions and all these questions were answered satisfactorily.

3. No pressure was exerted on me/*the patient/*participant to consent to participation and I/*the participant/*patient understand(s) that

APPENDIX A. INFORMATION AND INFORMED CONSENT DOCUMENT 14

I/*the participant/*patient may withdraw at any stage.

4. Participation in this study will not result in any additional costs to myself/*the participant/*patient.

B. HEREBY CONSENT VOLUNTARILY TO PARTICIPATE IN THE ABOVEMENTIONED PROJECT/*THAT THE PATIENT/*POTENTIAL PARTICIPANT MAY PARTICIPATE IN THE ABOVEMENTIONED STUDY.

Signed/confirmed at _____ on _____ 20_____
(place) (date)

Signature or right thumb print of
patient/*representative of
the patient/*participant



Signature of witness

STATEMENT BY OR ON BEHALF OF INVESTIGATOR(S):

I, _____ declare that I explained the information given in this document to _____ (name of the patient/*participant) and/*or his/*her representative _____ (name of the representative); he/*she was encouraged and given ample time to ask me any questions; this conversation was conducted in Afrikaans/*English and no translator was used.

Signed at _____ on _____ 20_____
(place) (date)

Signature of investigator/*investigator's
representative

Signature of witness

Appendix B

Circuit schematics and board layout

The developed unit's design layout is illustrated Figure B.1.

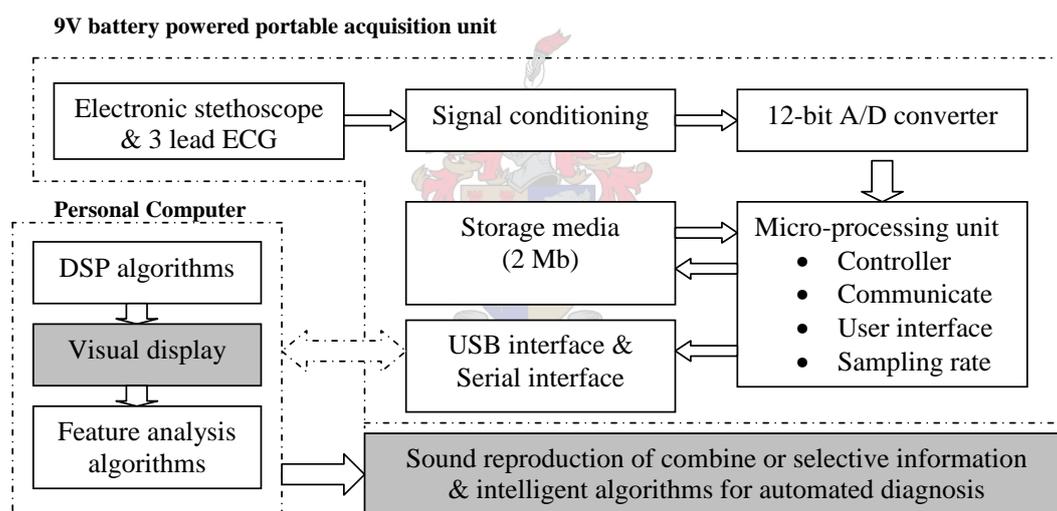


Figure B.1: Schematic diagram of the portable data acquisition unit and isolated USB or serial interface to PC

The designed unit interface with a notebook via a isolated USB connection. Here the data is captured by a communication program written in Delphi. The unit has separate filters for both the ECG signal and the heart sound signal. Each filter is designed for optimized signal conditioning. The interface board, shown in Figure B.4, has 2 Mb on-board storage capacity. (No such device is currently commercially available on the world market. The only option on the market is the *Littmann Model 4000* electronic stethoscope with an inadequate storage capacity of 6 channels of 8 seconds each).

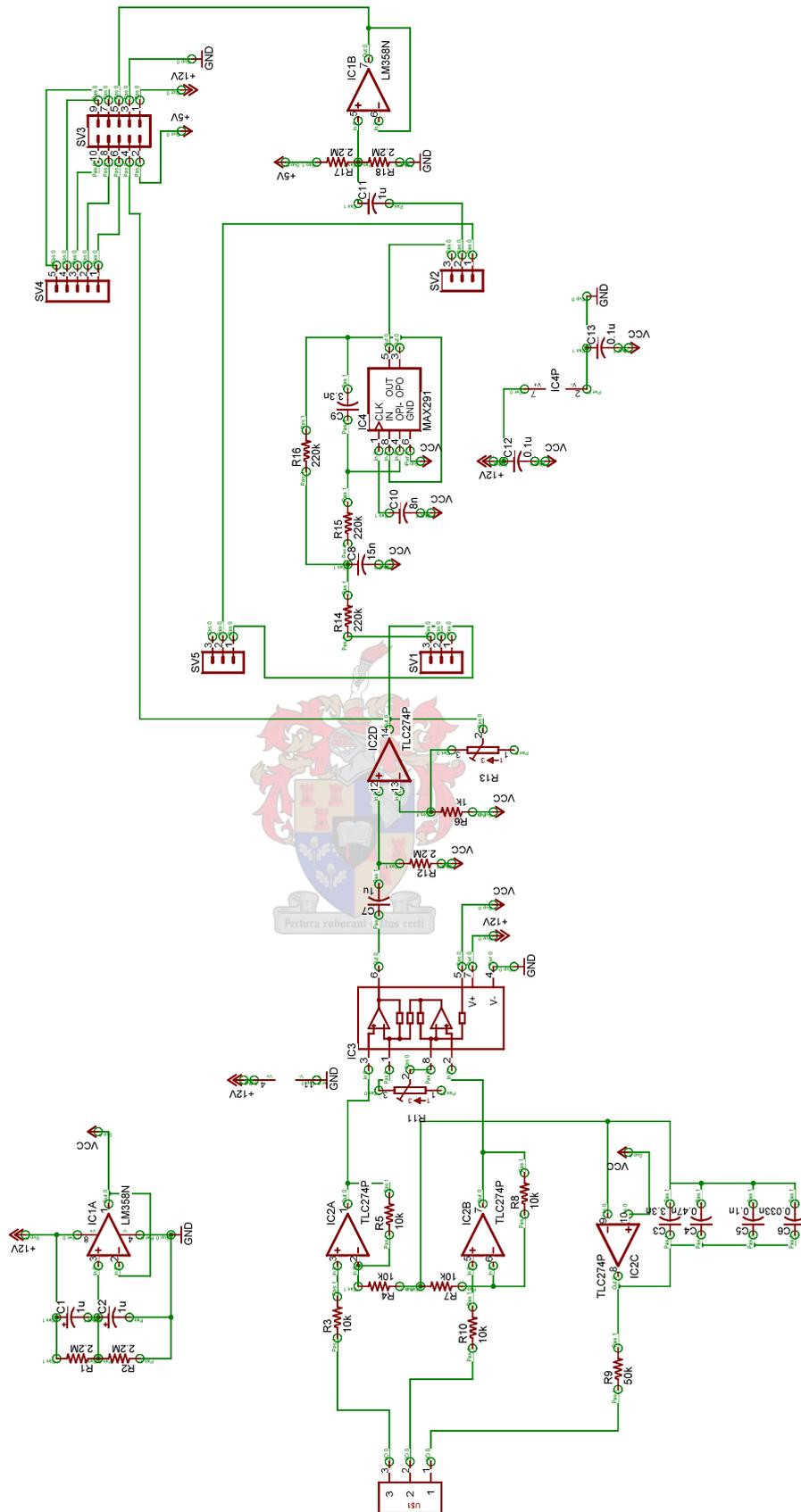


Figure B.3: Schematic diagram of 3-lead ECG board. A low-noise differential amplifier is used to obtain the voltage difference between the two primary electrodes. The third input is used as virtual ground. The signal is filtered with a 100 Hz LPF filter before normalized for the A/D circuitry.

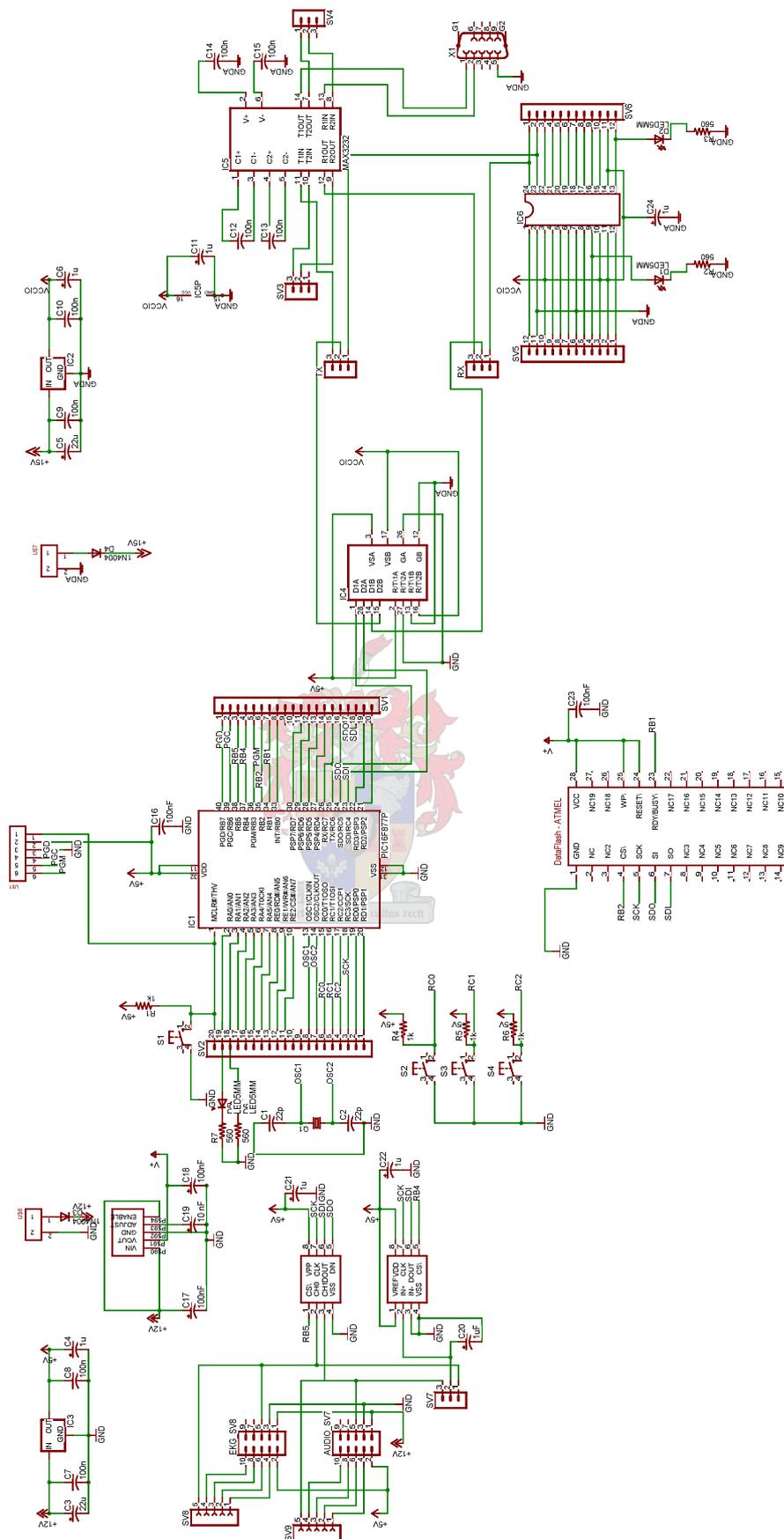


Figure B.4: Schematic diagram of digital acquisition board. The design consists of a 12-bit dual channel A/D converter; 2 Mb of on board flash memory for data storage; a micro processor ; an 4-channel optoc isolator and a USB & serial connection. Dual power supplies are used to isolate the patient from the computer.

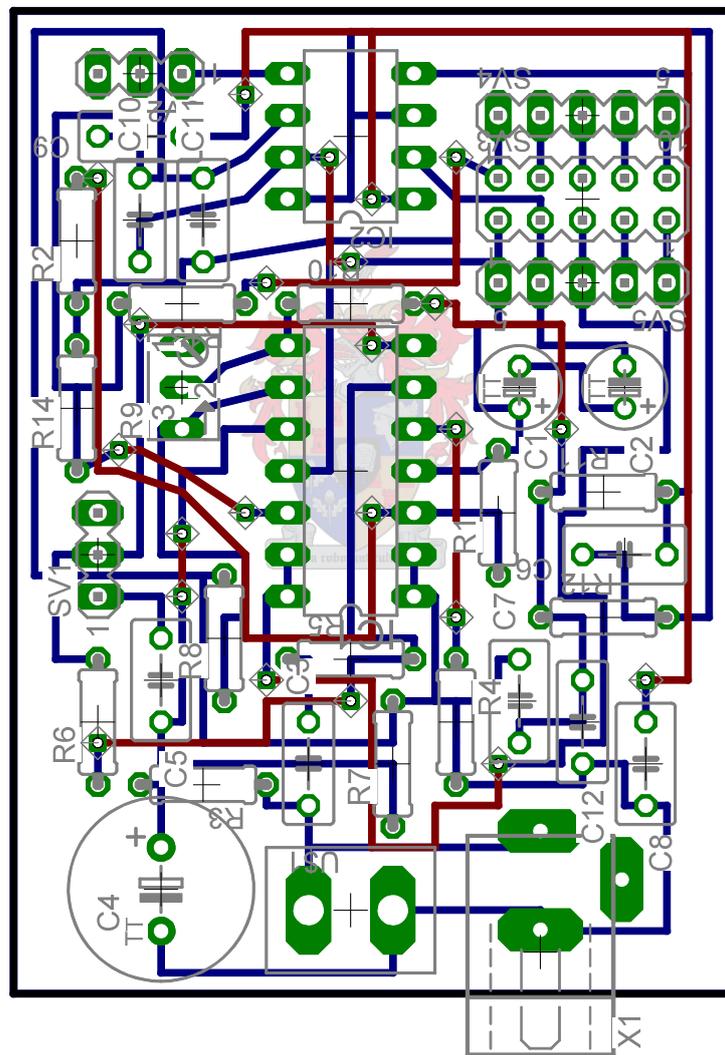


Figure B.5: Printed circuit board layout for the audio circuit.

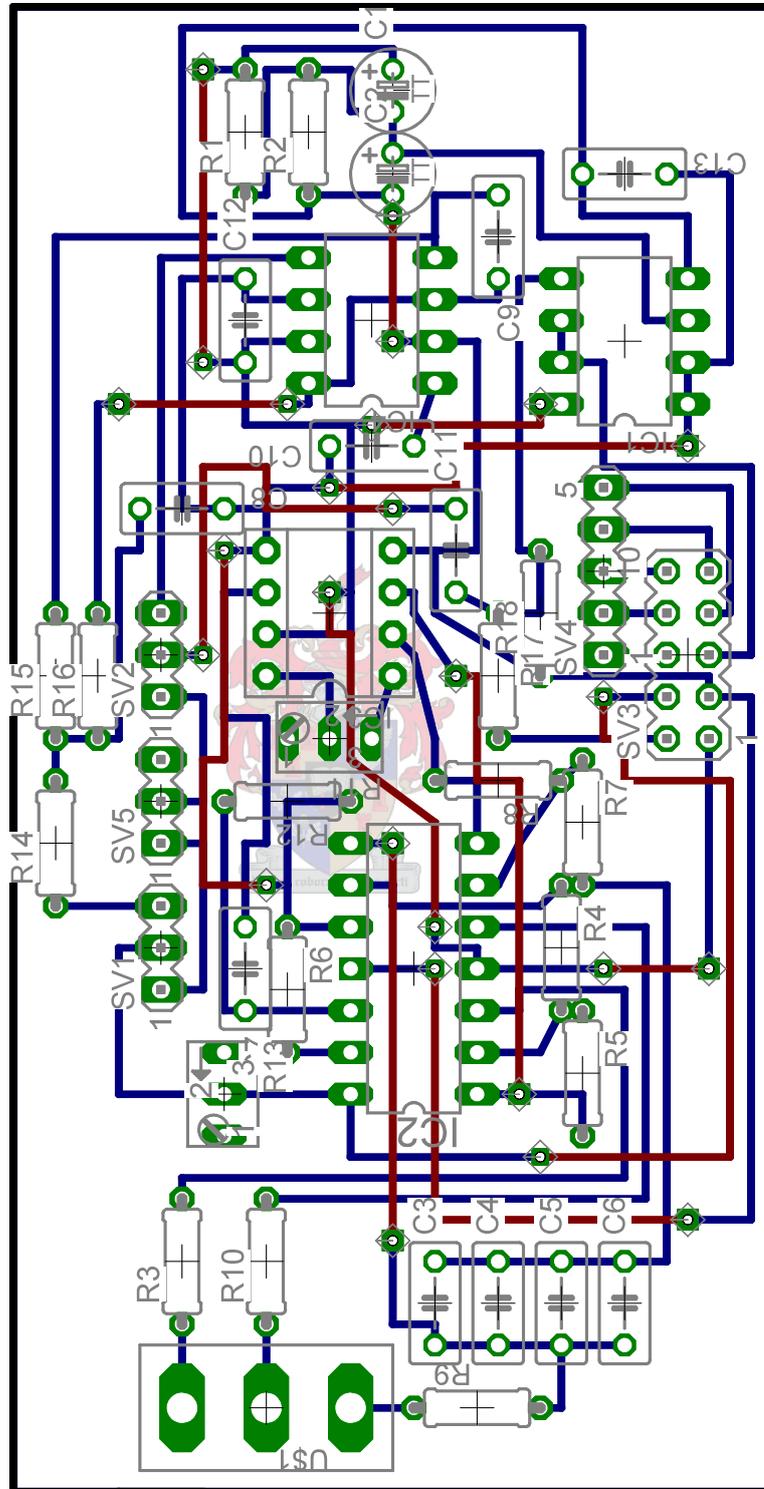


Figure B.6: Printed circuit board layout for the of 3-lead ECG circuit

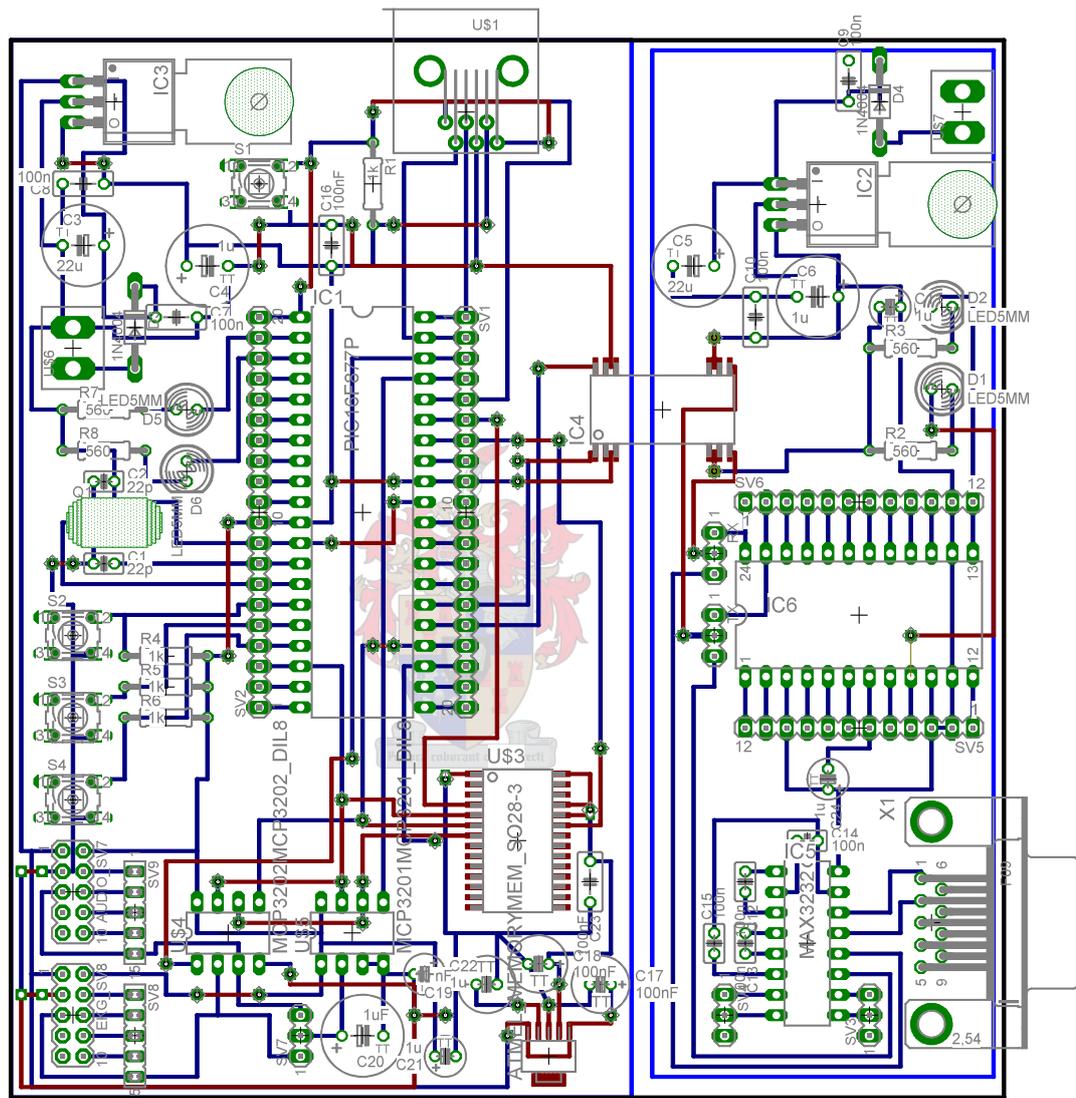


Figure B.7: Printed circuit board layout for the digital acquisition board

Appendix C

Background on wavelet analysis

The wavelet transform was derived from the Fourier transform discovered by J. Fourier. Fourier discovered that any periodic function could be expressed as an infinite sum of periodic complex exponential functions [45, 44]. This property of periodic functions was later generalized to non-periodic functions and then to (both periodic and non-periodic) discrete time functions. The Fourier Transform is normally used in the form of the Fast Fourier Transform algorithm and is mathematically described as

$$X(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi ft} dt \quad (\text{C.0.1})$$

Using the Fourier Transform, to investigate waveforms in the frequency domain, has one main disadvantage in that it does not provide sufficient information when used on non-stationary signals. The Fourier Transform only determines the frequency components of a signal integrated over all time, but no information about the instant of time at which the specific frequency component occurs. Thus for time localized frequency information extraction, the Fourier Transform is not the suitable tool to use.

In 1946, Denis Gabor [45] developed a technique involving 'windowing' the signal which maps the signal into a two dimensional space of time and frequency. This technique to exhibit signals in the time-frequency domain is known as Short Time Fourier Transform (STFT), and is mathematically described below as

$$STFT_X^{(w)}(t', x) = \int_t [x(t) w^*(t - t')] e^{-j2\pi ft} dt \quad (\text{C.0.2})$$

The windowing technique involves translating (translation parameter t' in equation C.0.2) the complex conjugate of the window function $w(t)$ along the length of the signal $x(t)$, and multiplying the two functions $x(t)$ and $w^*(t)$ at the different instants of time. The exponen-

tial part converts the result of each multiplication to the frequency domain at that instant as done in the ordinary Fourier Transform.

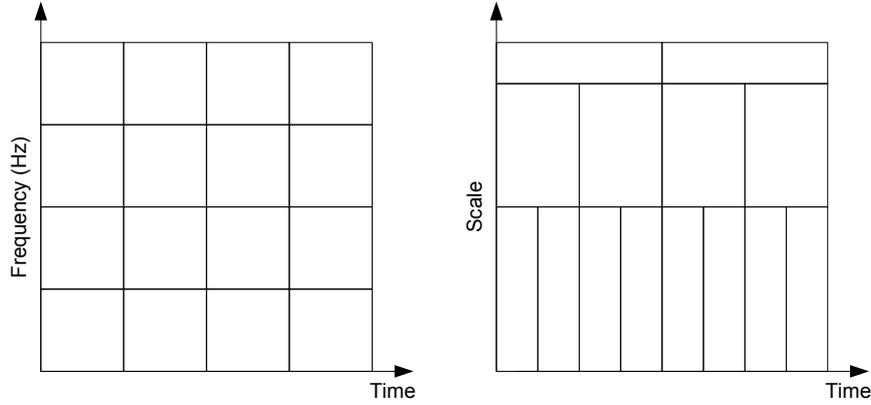


Figure C.1: Windowing regions of STFT and WT analyses

To overcome resolution problems that make it impossible to analyze the signal simultaneously in both the time and frequency domain, the wavelet transform (WT) was developed. The main difference between the STFT and the WT is that the WT uses a variable sized window region (or wavelet) to examine the signal which helps to reduce resolution problems significantly. Wavelets are families of functions $\Psi_{a,b}(t)$ generated from a single base wavelet $\Psi(t)$ called the mother wavelet, by dilation and translation [58, 45], i.e.,

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{a}} \Psi\left(\frac{t-b}{a}\right), a > 0 \quad (\text{C.0.3})$$

where a is the dilation (scale) parameter and b is the translation parameter. The theory behind the wavelet approach is that being able to dilate (stretch) or compress the wavelet, different features of the signal will be extracted. For example a narrow wavelet will show up higher frequency components, while a stretched wavelet show up lower frequency components of the signal. A comparison between the constant window regions used in STFT analysis and the variable window regions used in WT analysis is exhibited in Figure C.1. The scale of the wavelet, can also be thought as the inverse of frequency (or pseudo-frequency). The pseudo-frequency F_a in Hz corresponding to the scale a , can be calculated as follow

$$F_a = \frac{F_c}{a\Delta} \quad (\text{C.0.4})$$

where

a is a scale.

Δ is the sampling period. ($\Delta = 9.1e10^{-5}$ for $F_s = 11025\text{Hz}$)
 F_c is the center frequency of a wavelet in Hz.

The idea is to associate with a given wavelet a purely periodic signal of frequency F_c . The frequency maximizing the fft of the wavelet modulus is F_c .

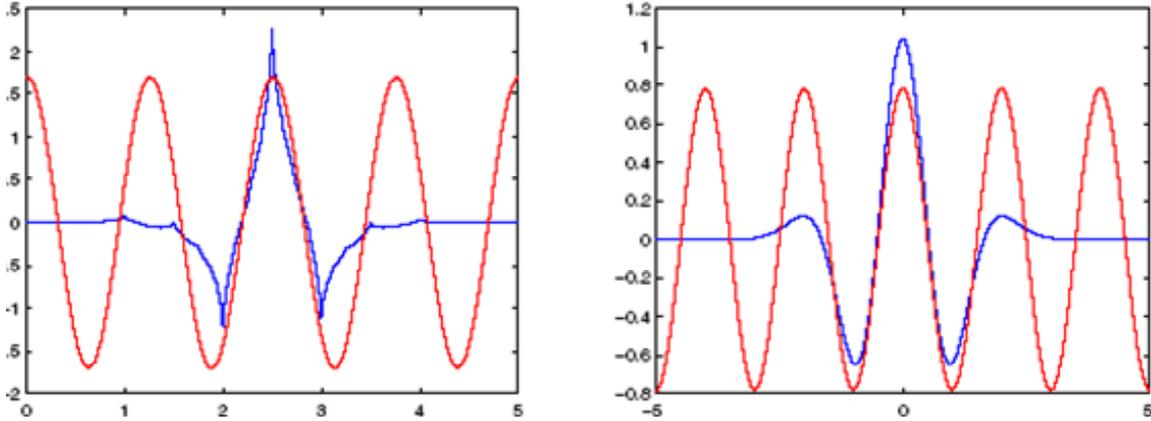


Figure C.2: Wavelets to illustrate pseudo frequency

The continuous wavelet transform of a 1-D function $f(t) \in L_2(\mathbb{R})$, where $L_2(\mathbb{R})$ denotes the vector space of measurable, square-integrable one-dimensional functions $f(t)$, is defined in a Hilbert space, as the projection of the function onto the wavelet set $\Psi_{a,b}(t)$, i.e.,

$$CWT_f^\Psi(a,b) = \Psi_f^\Psi(b,f) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) \Psi^* \left(\frac{t-b}{a} \right) dt \quad (C.0.5)$$

where * represents complex conjugate.

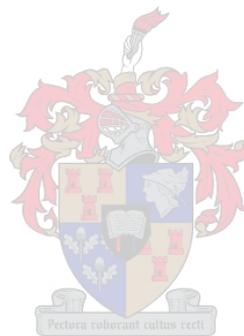
The process involved in creating the CWT is very much the same as that involved with the STFT, except that the wavelet traverses the signal many times (as indicated by the translation parameter b), with each traversal computed with a different scale. The CWT is extremely good at displaying information about the signal in great detail, but due to much needed computer power, the discrete wavelet transform (DWT) is usually used. Die DWT computes the wavelet coefficients at discrete intervals of time and scale, compared to the CWT that shifts the wavelet and changes the scale in a continuous nature in order to calculate the WT coefficients. The CWT of time series $f(t)$ can also be written as a convolution

$$CWT_f^\Psi(a,b) = f(b) * \Psi_{a,0}^*(-b) \quad (C.0.6)$$

where

$$\Psi_{a,b}(t) \equiv \frac{1}{\sqrt{a}} \Psi\left(\frac{t-b}{a}\right) \quad (\text{C.0.7})$$

Wavelets are constructed to maintain a constant ratio of center frequency to 3-dB bandwidth (Q) and have a finite duration in time. Their time-frequency resolution is inherent in their design and scale parameters. This is in contrast to Fourier decomposition, which uses the infinite time extent sine and cosine functions. Time resolution is not inherent in the Fourier Transform but is introduced by the user by windowing data. Multiple FT's using distinct windowing periods would be required to produce the constant Q decomposition offered by wavelets [44].



Appendix D

The Shapiro Wilk' test for normality

This program has a limitation and it is that only admits data sets up between 6 and 2001, because for its construction it was necessary the introduction of a program that implement the Shapiro Wilks test (<http://www.biostat.cmu.com>) and this only admits this sizes

```
shapiro.wilk.test
function(x)
{
  # "shapiro.wilk.test"
  # This function is an S version of the procedure described by
  # J. P. Royston (1982) in "An Extension of Shapiro and Wilk's
  # W Test for Normality to Large Samples", Applied Statistics,

  n <- length(x)
  index <- 1:n
  m <- qnorm((index - 0.375)/(n + 0.25))
  y <- sort(x)
  mu <- mean(y)
  SSq <- sum((y - mu)^2)
  astar <- 2 * m
  astar.p <- astar[ - ends]

  if(n <= 20)
    m <- n - 1
  else m <- n
}
```

```

if(m < 20)
  aa <- gamma(0.5 * (m + 1))/(sqrt(2)*gamma(0.5 * m + 1))
else {
  f1 <- (6 * m + 7)/(6 * m + 13)
  f2 <- exp(1)/(m + 2)
  f3 <- (m + 1)/(m + 2)
  f3 <- f3^(m - 2)
  aa <- f1 * sqrt(f2 * f3)
}

astar.1 <- (aa * sum(astar.p^2))/(1 - 2 * aa)
astar.1 <- sqrt(astar.1)
astar[1] <- - astar.1
astar[n] <- astar.1
A <- astar/sqrt(sum(astar^2))

W <- (sum(A * y)^2)/SSq
u <- log(n) - 3
lambda <- 0.118898 + 0.133414 * u + 0.327907 * u^2

logmu <- -0.37542 - 0.492145 * u - 1.124332 * u^2 - 0.199422 * u^3
logsigma <- -3.155805 + 0.729399 * u + 3.01855 * u^2 + 1.5558776 * u^3

if(n > 20)
{
  u <- log(n) - 5
  lambda <- 0.480385 + 0.318828 * u + 0.0241665 * u^3 + 0.00879701 * u^4 +
0.002989646 * u^5

  logmu <- -1.91487 - 1.37888 * u - 0.04189209 * u^2 + 0.1066339 * u^3 -
0.03513666 * u^4 - 0.01504614 * u^5

  logsigma <- -3.83538 - 1.015807 * u - 0.331885 * u^2 + 0.1773538 * u^3 -
0.01638782 * u^4 - 0.03215018 * u^5 + 0.003852646 * u^6

mu <- exp(logmu)
sigma <- exp(logsigma)

y <- (1 - W)^lambda
z <- (y - mu)/sigma

```

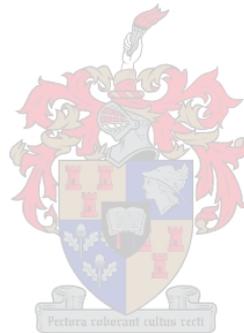
```
p <- 1 - pnorm(z)

if(n < 7) {
  warning("n is too small for this program to correctly estimate p")
  p <- NA
}

if(n > 2000) {
  warning("n is too large for this program to correctly estimate p")
  p <- NA
}

out <- list(W = W, n = n, p = p)
out
}

\endinput
```



Appendix E

Matlab program code

E.1 Direct Ratio method

E.1.1 M-file used in the Direct Ratio algorithm

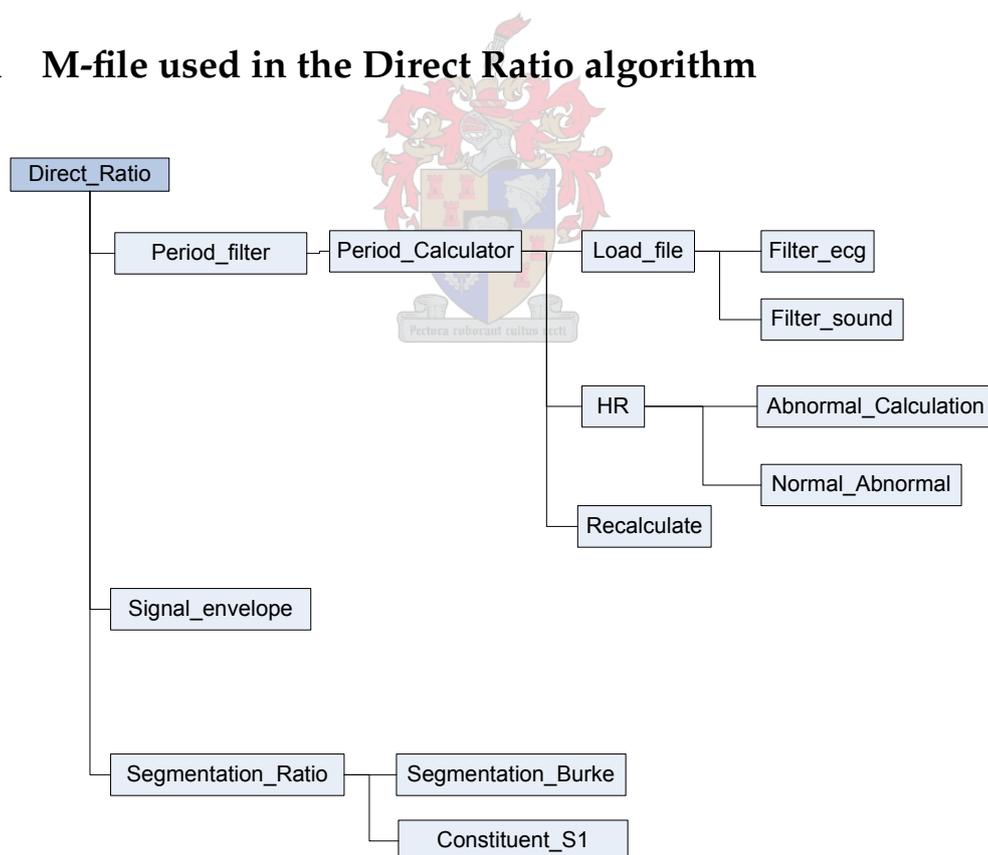


Figure E.1: Flow diagram of M-files used in the Direct Ratio algorithm. Code for the Direct_Ratio.m and Period_Filter.m are listed in this Appendix. The rest of the files can be viewed on the accompanied compact disc

E.1.2 Code for Direct_Ratio.m

```

% Program: Direct_ratio.m
% Editor: Jacques de Vos
% Date: 22-09-2004
% Version: 2.4
% -----.m files called-----
% Period_filter.m, Signal_envelope.m, Segmentation_Ratio.m
% -----

function [HR_result, C_result,D_result,E_result,F_result] =
    Direct_ratio(filename,gender,Diagnosis,age);

fs = 22050;

[HR_result, ecg_matrix, sound_matrix] = Period_filter(fs,gender,filename,age);

x = length(ecg_matrix(1,:));

C_ratio = ones(1,x);
D_ratio = ones(1,x);
E_ratio = ones(1,x);
F_ratio = ones(1,x);

for z = 1:x

figure(z);

    ecg_z = ecg_matrix(:,z);
    sound_z = sound_matrix(:,z);

    %calculate the envelope of the power to calculate ratio with this values
    [power_z] = Signal_envelope(sound_z);

    [C_ratio(z), D_ratio(z), E_ratio(z), F_ratio(z), y] =
        Segmentation_Ratio(power_z,gender,fs);

subplot(2,1,2);
bar(y);
title('[1] Late Diastole [2] S1 [3] Whole systole [4] Early Systole [5] Late Systole [6]

```



```

Mid Systole [7] Dead Time [S2 & Early & Mid Diastole');
ylabel('Constituent''s Average Energy ');
xlabel('Constituents')
end

```

```

C_result = 20*log10(sum(C_ratio)/length(C_ratio))
D_result = 20*log10(sum(D_ratio)/length(D_ratio))
E_result = 20*log10(sum(E_ratio)/length(E_ratio))
F_result = 20*log10(sum(F_ratio)/length(F_ratio))

```

E.1.3 Code for Period_Filter.m

```

% Program: Period_filter.m
% Editor: Jacques de Vos
% Date: 26-07-2004
% Version: 1.1
% -----Program description-----
% Calculate patients heart rate using cross - correlation
% Necessary for segmentation of heart sound into separate beats (periods)
% Separate recording into separate periods
% .m files called: HR.m; Recalculate.m
% Calculate various periods Mel_scale frequency components, with the
% following algorithm
% f = 20 - 420;
% n = 12; % 12 section banks
% Mel(f) = 2595*log10(1 + f_mel/700)
% f_mel = [28 57 87 119 151 185 221 258 296 336 377 420];
% Calculate correlation of individual periods with average values
% Throw out periods with an correlation below chosen threshold value
% -----

function [HR_result, ecg_filtered, sound_filtered]= Period_filter(fs,gender,filename,age);

w_avg = zeros(1,120)';
t = 1:1:120;

[HR_result, ecg_matrix, sound_matrix] = Period_calculator(fs,gender,filename,age);

```

```

x = length(sound_matrix(1,:));
diff = zeros(120,x);
corr_p = zeros(1,x);
sum_diff = zeros(1,x);
diff_fault = zeros(1,x);
reg_of_weg = zeros(1,x);
row = length(ecg_matrix(:,1)); % number of data points in data set

test_matrix_new = ones(120,x);

f_mel = [20 28 57 87 119 151 185 221 258 296 336 377 420];
mel = ones(1,12);

for w = 1:x
    window_size = floor(length(sound_matrix(:,w))/10);
    f_mat = 20:1:420;

    [B,F,T] = specgram(sound_matrix(:,w),f_mat,22050,window_size,0);

    for z = 1:10
        for q = 1:12 %forming the mel-scale frequency banks data-point
            mel(q) = sum(20*log10(abs(B(f_mel(q)-19:f_mel(q+1)-19,z))))/(f_mel(q+1)-f_mel(q));
        end

        test_matrix_new((12*z-11):(12*z),w) = mel';
    end

    %construct average representative of the spectrum bins computed
    w_avg = w_avg + test_matrix_new(:,w);
end

r =w_avg/x;

for w = 1:x
    c = test_matrix_new(:,w);
    xc=CORRCOEF(c,r);
    corr_p(1,w) = xc(1,2);
end

n = 1:1:x;

```

```
avg = sum(corr_p)/length(corr_p);

for w = 1:x
    diff_fault(1,w) = abs(corr_p(1,w) - avg);
end

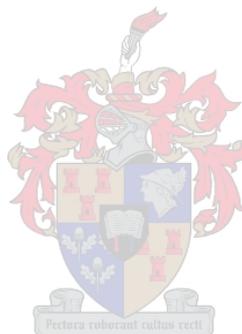
avg_fault = sum(diff_fault)/length(diff_fault);
u=0;

for w = 1:x
    if corr_p(1,w) < avg - 1.5*avg_fault
        reg_of_weg(1,w) = 1;
    else
        reg_of_weg(1,w) = 0;
        u = u + 1;
    end
end

ecg_filtered = ones(row,u);
sound_filtered = ones(row,u);

s = 1;

for m = 1:x
    if corr_p(1,m) > avg - 1.5*avg_fault
        ecg_filtered(:,s) = ecg_matrix(:,m);
        sound_filtered(:,s) = sound_matrix(:,m);
        s = s+1;
    end
end
```



E.2 Wavelet analysis method

E.2.1 M-file used in the Wavelet analysis algorithm

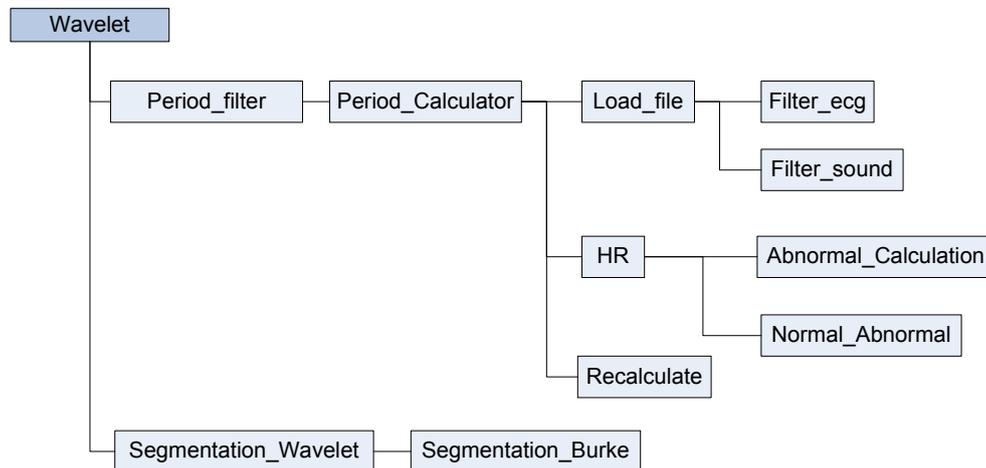


Figure E.2: Flow diagram of M-files used in the Wavelet analysis algorithm. Code for Wavelet.m and Period_Filter.m are listed in this Appendix. The rest of the files can be viewed on the accompanied compact disc

E.2.2 Code for Wavelet.m

```

% Program: Wavelet.m
% Editor: Jacques de Vos
% Date: 08-06-2004
% Version: 1.7
% -----.m files called-----
% Period_filter.m, Segmentation_Wavelet.m
% -----

function [HR_result, C_result,D_result,E_result,F_result] = Wavelet(filename,gender,age);

fs = 22050;

[HR_result, ecg_matrix, sound_matrix] = Period_filter(fs,gender,filename,age);

x = length(ecg_matrix(1,:));
  
```

```

C = ones(3,x);
D = ones(3,x);
E = ones(3,x);
F = ones(3,x);

for z = 1:x
    ecg_z = ecg_matrix(:,z);
    sound_z = sound_matrix(:,z);

    c = cwt(sound_z,1:1:64,'coif2');

    cwt16 = c(16,:);
    cwt32 = c(32,:);
    cwt64 = c(64,:);

    [C(1,z), D(1,z), E(1,z), F(1,z)] = Segmentation_Wavelet(cwt16,gender,fs);
    [C(2,z), D(2,z), E(2,z), F(2,z)] = Segmentation_Wavelet(cwt32,gender,fs);
    [C(3,z), D(3,z), E(3,z), F(3,z)] = Segmentation_Wavelet(cwt64,gender,fs);
end

for t = 1:3
    C_result(t) = 20*log10(sum(C(t,:))/x);
    D_result(t) = 20*log10(sum(D(t,:))/x);
    E_result(t) = 20*log10(sum(E(t,:))/x);
    F_result(t) = 20*log10(sum(F(t,:))/x);
end

```

E.3 Neural network: Training data-set compilation

```

% Program: Training_Mel_scale.m
% Editor: Jacques de Vos
% Date: 18-07-2004
% Version: 1.1
% -----Program description-----
% (i) Assemble the training data; using Mel scale
% frequency banks. Assemble 3-period validation matrix parallel
% f = 20 - 420;
% n = 12; % 12 section banks
% Mel(f) = 2595*log10(1 + f_mel/700)

```

```

% f_mel = [28 57 87 119 151 185 221 258 296 336 377 420];
% -----

function [training_matrix, validation_matrix, training_target, validation_target] = Training

% Get input form user for names for the different contributes of training data

[Number] = input('Number of recordings to use as training data: ');

for g = 1:Number          % number of recordings to train network with

%-----
%-----Calculate matrix for one recording-----
%-----

gender = input('Gender of patient: ','s');
[HR_result, ecg_matrix, sound_matrix] = Period_filter(fs,gender,filename,age);
[Diagnosis] = input('Normal (0) or Pathology (1): ');

x = length(sound_matrix(1,:)) - 3;

% construct matrix dimensions

training_matrix_new = ones(120,x);
validation_matrix_new = ones(120,3);

    if Diagnosis == 1
        training_target_new = ones(1,x);
        validation_target_new = ones(1,3);
    elseif Diagnosis == 0
        training_target_new = zeros(1,x);
        validation_target_new = zeros(1,3);
    end

f_mel = [20 28 57 87 119 151 185 221 258 296 336 377 420];
mel = ones(1,12);

% form matrix for training of neural network
for w = 1:x
    window_size = floor(length(sound_matrix(:,w))/10);

```

```

f_mat = 20:1:420;

[B,F,T] = specgram(sound_matrix(:,w),f_mat,22050,window_size,0);

surf(T,F,20*log10(abs(B)));
size(B)
for z = 1:10
    for q = 1:12          %forming the mel-scale frequency banks data-point
                        representative
        mel(q) =
            sum(20*log10(abs(B(f_mel(q)-19:f_mel(q+1)-19,z))))/(f_mel(q+1)-f_mel(q));
    end

    training_matrix_new((12*z-11):(12*z),w) = mel';
end
end

% form matrix for validation of trained neural network (Last three periods
% of the sound recording)
for r = 1:3

    window_size = floor(length(sound_matrix(:,w+r))/10);
    f_mat = 20:1:420;
    [B,F,T] = specgram(sound_matrix(:,w+r),f_mat,22050,window_size,0);

    surf(T,F,20*log10(abs(B)));

    for z = 1:10
        for q = 1:12
            mel(q) =
                sum(20*log10(abs(B(f_mel(q)-19:f_mel(q+1)-19,z))))/(f_mel(q+1)-f_mel(q));
        end

        validation_matrix_new((12*z-11):(12*z),r) = mel';
    end

end

end
%-----
%----- Append new recording data to old matrix-----
%-----

```

```
if g==1
    training_matrix = training_matrix_new;
    train_mel = training_matrix;
    save train_mel train_mel
    validation_matrix = validation_matrix_new;
    validate_mel = validation_matrix;
    save validate_mel validate_mel

    validation_target = validation_target_new;
    validate_target_mel = validation_target;
    save validate_target_mel validate_target_mel
    training_target = training_target_new;
    train_target_mel = training_target;
    save train_target_mel train_target_mel

else
    training_matrix = [training_matrix training_matrix_new];
    train_mel = training_matrix;
    save train_mel train_mel
    validation_matrix = [validation_matrix validation_matrix_new];
    validate_mel = validation_matrix;
    save validate_mel validate_mel

    validation_target = [validation_target validation_target_new];
    validate_target_mel = validation_target;
    save validate_target_mel validate_target_mel
    training_target = [training_target training_target_new];
    train_target_mel = training_target;
    save train_target_mel train_target_mel
end

clear training_matrix_new;
clear validation_matrix_new;
clear validation_target_new;
clear training_target_new;

end
```

E.4 Neural network: Architecture, Initialization, Training, Testing, Validation and Performance testing

```

% Program: Neural_network.m
% Editor: Jacques de Vos
% Date: 18-07-2004
% Version: 1.1
% -----Program description-----
% (i) Assemble the training data
% (ii) Create the network object
% (iii) Train the network
% (iv) Simulate the network response to new data inputs
% .m file called - Calculate_Statistics.m
% -----

function [result] = Neural_network(P,V,TT,VT);

pr = ones(120,2);
pr(:,1) = -100;
pr(:,2) = 100;

% neural network format
% constuct the neural network, with training and learning algorithms
% net = newff(PR,[S1 S2],[TF1 TF2],BTF,BLF,PF);

network = newff(pr,[30 1],{'logsig' 'logsig'}, 'traingdx', 'learngdm', 'mse');

% (iii) Train the network

network.trainParam.epochs = 3000;
network.trainParam.goal = 1e-5;

network = train(network,P,TT); % train the network with the training data

save network network % save network for later validation

Y = sim(network,P); % simulate the network

```

```

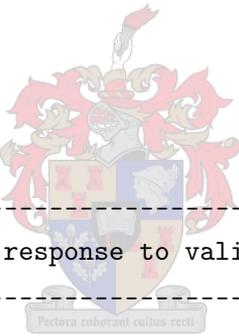
%get before final sigmoid function value for the training data
n_output = ones(1,length(P(1,:)));

for b = 1:length(P(1,:))
    n_1 = network.iw{1,1}*P(:,b) + network.b{1};
    n_output(b) = network.lw{2,1}*logsig(n_1) + network.b{2};
end

t = 1:1:length(P(1,:));

figure(1);    % plot value just before last function for training data
plot(t,n_output,'o');
hold;
plot(t,TT,'v');

figure(2);    % plot neural network output for training data
plot(t,Y,'o');
hold;
plot(t,TT,'v');

% -----
% (iv) Simulate the network response to validation (new) data inputs
% -----

Y_sim = sim(network,V);

%get before final sigmoid function value for the validation data
v_output = ones(1,length(V(1,:)));
v_result = ones(1,length(V(1,:))/3);
v_result_target = ones(1,length(V(1,:))/3);

counter = 1;

for b = 1:length(V(1,:))
    n_1 = network.iw{1,1}*V(:,b) + network.b{1};
    v_output(b) = network.lw{2,1}*logsig(n_1) + network.b{2};

    if(rem(b,3) == 0)                % calculate average of patients periods
        v_result(counter) = (v_output(b) + v_output(b-1) + v_output(b-2))/3;
        v_result_target(counter) = VT(b);
    end
end

```

```

    counter = counter + 1;
end

end

% -----
% calculate average output of three periods for validation data of each patients

v_fresult = ones(1,length(V(1,:))/3);
v_fresult_target = ones(1,length(V(1,:))/3);

counter = 1;

for b = 1:length(V(1,:))
    if(rem(b,3) == 0)                % calculate average of patient's periods
        v_fresult(counter) = (Y_sim(b) + Y_sim(b-1) + Y_sim(b-2))/3;
        v_fresult_target(counter) = VT(b);
        counter = counter + 1;
    end
end

% -----

t = 1:1:length(V(1,:));
t_2 = 1:1:length(V(1,:))/3;

figure(3); % plot value just before last function for validation data
plot(t,v_output,'o');
hold;
plot(t,VT,'^');

figure(4); % plot avg value of patient just before last function
plot(t_2,v_result,'*r');
hold;
plot(t_2,v_result_target,'ob');

figure(5); % plot neural network output for validation data
plot(t,Y_sim,'*');
hold;
plot(t,VT,'^');

figure(6); %plot validation diagnosis of each patient

```

```

plot(t_2,v_fresult,'*r');
hold;
plot(t_2,v_fresult_target,'ob');

%-----CALCULATE MEAN AND STANDARD DEVIATION FOR NORMAL
%-----AND PATHOLOGICAL POPULATION USING DATA JUST BEFORE
%-----FINAL FUNCTION

pat_size = 0;
nor_size = 0;

for w = 1:length(V(1,:))/3
    if (v_result_target(w) == 1)
        if (pat_size == 0)
            pat = v_result(w)
            pat_size = pat_size + 1;
        else
            pat = [pat v_result(w)]
            pat_size = pat_size + 1;
        end
    else
        if (nor_size == 0)
            nor = v_result(w)
            nor_size = nor_size + 1;
        else
            nor = [nor v_result(w)]
            nor_size = nor_size + 1;
        end
    end
end

% x_nor = 1:1:length(nor_size);
% x_pat = 1:1:length(pat_size);
t = 1:1:length(nor) + length(pat);

figure(7);
plot(t(1:length(pat)),pat,'or');
hold;
plot(t(length(pat) + 1:length(nor) + length(pat)),nor,'*b');

```



```

save nor nor
save pat pat

% calculate the acceptance or rejection of the null hypothesis
% using the data before the final function in the neural network

accept_reject = Calculate_Statistics(nor,pat) %calculate statistics

%Re-arrange data in pathological and normal group

fpat_size = 0;
fnor_size = 0;

for v = 1:length(V(1,:))/3
    if (v_fresult_target(v) == 1)
        if (fpat_size == 0)
            fpat = v_fresult(v)
            fpat_size = fpat_size + 1;
        else
            fpat = [fpat v_fresult(v)]
            fpat_size = fpat_size + 1;
        end
    else
        if (fnor_size == 0)
            fnor = v_fresult(v)
            fnor_size = fnor_size + 1;
        else
            fnor = [fnor v_fresult(v)]
            fnor_size = fnor_size + 1;
        end
    end
end

t = 1:1:length(fnor) + length(fpat);

figure(8)
plot(t(1:length(fpat)),fpat,'or');
hold;

```

```

plot(t(length(fpat) + 1:length(fnor) + length(fpat)),fnor,'*b');

save fnor fnor
save fpat fpat

```

E.5 *Jack-Knife* neural network

E.5.1 *Jack-Knife* train data-set composition

```

% Program: Jack_Knife_Trianing.m
% Editor: Jacques de Vos
% Date: 14-10-2004
% Version: 1.3
% -----Program description-----
% Add recordings training data to imported training matrix
% Add six periods per patient - (no validation data) only patient data
% Append new data to : train_JK and train_target_JK matrix
% -----

function [] = Jack_Knife_training(HR_result,ecg_matrix,sound_matrix,Diagnosis);

% load already existing matrix structure if there exist one
load train_JK;
load train_target_JK;

training_matrix = train_JK;
training_target = train_target_JK;

fs = 22050;

% re-use periods if recording has less than 6 periods

if length(sound_matrix(1,:)) < 6
    if length(sound_matrix(1,:)) == 4
        sound_matrix = [sound_matrix sound_matrix(:,1:2)];
    elseif length(sound_matrix(1,:)) == 5
        sound_matrix = [sound_matrix sound_matrix(:,1)];
    end
end

```

```

    end
end

x = 6;

% construct matrix dimensions

training_matrix_new = ones(120,x);

    if Diagnosis == 1
        training_target_new = ones(1,x);

    elseif Diagnosis == 0
        training_target_new = zeros(1,x);
    end

f_mel = [20 28 57 87 119 151 185 221 258 296 336 377 420];
mel = ones(1,12);

% form matrix for evaluation of Jack-Knife neural network
for w = 1:x
    window_size = floor(length(sound_matrix(:,w))/10);
    f_mat = 20:1:420;

    [B,F,T] = specgram(sound_matrix(:,w),f_mat,22050,window_size,0);

    %forming the mel-scale frequency banks data-point representative
    for z = 1:10
        for q = 1:12
            mel(q) =
                sum(20*log10(abs(B(f_mel(q)-19:f_mel(q+1)-19,z))))/(f_mel(q+1)-f_mel(q));
        end
        training_matrix_new((12*z-11):(12*z),w) = mel';
    end
end

end

%-----
%----- Append new recording data to already existing matrix-----
%-----

```

```

training_matrix = [training_matrix training_matrix_new];
train_JK = training_matrix;
save train_JK train_JK

training_target = [training_target training_target_new];
train_target_JK = training_target;
save train_target_JK train_target_JK

clear training_matrix_new;
clear training_target_new;

```

E.5.1.1 *Jack-Knife* training and validation (*Jack-Knife* Iteration)

```

% Program: Jack_Knife_NN.m
% Editor: Jacques de Vos
% Date: 14-10-2004
% Version: 1.4
% -----Program description-----
% Train the Jack-Knife neural network with 162 recordings and validate
% on the remaining recording on a ring-a-around basis. Thus, train the
% network 163 times, to obtain a classification for each patient
% -----

function [Results,wanted_results] = Jack_Knife_NN(train_matrix,target_matrix);

Results = ones(1,size(train_matrix,2)/6);
wanted_results = ones(1,size(train_matrix,2)/6);

tellerA = size(train_matrix,2) - 5;

for x = 1:size(train_matrix,2)/6

    if x == 1
        t_m = train_matrix(:,1:tellerA - 1);
        t_t_m = target_matrix(1,1:tellerA - 1);

        v_m = train_matrix(:,tellerA:tellerA + 5);
        v_t_m = target_matrix(1,tellerA:tellerA + 5);
    end
end

```

```

        [Results(x)] = Validation_NN_Jack_Knife(t_m,v_m,t_t_m,v_t_m);
        wanted_results(x) = v_t_m(1,1);

        tellerA = tellerA - 6;

elseif x == size(train_matrix,2)/6

    t_m = train_matrix(:,tellerA+6:size(train_matrix,2));
    t_t_m = target_matrix(1,tellerA+6:size(target_matrix,2));

    v_m = train_matrix(:,tellerA:tellerA + 5);
    v_t_m = target_matrix(1,tellerA:tellerA + 5);

    [Results(x)] = Validation_NN_Jack_Knife(t_m,v_m,t_t_m,v_t_m);
    wanted_results(x) = v_t_m(1,1);

    save Results Results
    save wanted_results wanted_results

else
    t_m = [train_matrix(:,1:tellerA-1)
           train_matrix(:,tellerA+6:size(train_matrix,2))];
    t_t_m = [target_matrix(:,1:tellerA-1)
             target_matrix(1,tellerA+6:size(target_matrix,2))];

    v_m = train_matrix(:,tellerA:tellerA + 5);
    v_t_m = target_matrix(1,tellerA:tellerA + 5);

    [Results(x)] = Validation_NN_Jack_Knife(t_m,v_m,t_t_m,v_t_m);
    wanted_results(x) = v_t_m(1,1);

    tellerA = tellerA - 6;

    save Results Results
    save wanted_results wanted_results
end
end % for-statement

```

E.5.2 *Jack-Knife* simulation and testing.(Calculation of validation recording classification)

```

% Program: Validation_NN_Jack_Knife.m
% Editor: Jacques de Vos
% Date: 24-08-2004
% Version: 1.6
% -----Program description-----
% (i) Assemble the training data - Jack_Knife_training.m -
% (ii) Create the network object
% (iii) Train the network
% (iv) Simulate the network response with validation vector
% (v) Send validation results back to Jack_Knife_NN
% -----

function [result] = Validation_NN_Jack_Knife(P,V,TT,VT);

pr = ones(120,2);
pr(:,1) = -100;
pr(:,2) = 100;

network = newff(pr,[30 1],{'logsig' 'logsig'}, 'traingdx', 'learngdm', 'mse');
network.trainParam.epochs = 5000;
network.trainParam.goal = 1e-5;

network = train(network,P,TT);

save network network

%Y = sim(network,P);
Y_sim = sim(network,V);

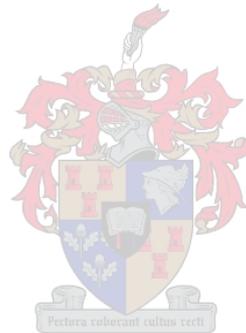
counter = 1;

% calculate average of patient's 6 periods
for b = 1:length(V(1,:))
    if(mod(b,6) == 0)
        v_fresult = (Y_sim(b) + Y_sim(b-1) + Y_sim(b-2)+ Y_sim(b-3) + Y_sim(b-4) + Y_sim(b-5))/6;
        counter = counter + 1;
    end
end

```

```
end
```

```
result = v_fresult;
```



Bibliography

- [1] E.F. Bartholomew F.H. Martini, *Essentials of Anatomy and Physiology*, Prentice Hall, Inc., second edition edition, 2000, 356-359. (Cited on pages vii, 7, 12, 15, and 20.)
- [2] S.S. Wilks I.GuttMan, *Introductory Engineering Statistics*, John Wiley & Sons, Inc., 1 st edition, 1965. (Cited on pages viii, 73, 75, and 77.)
- [3] S. Snyders, "Inductive machine learning bias in knowledge-based neurocomputing.," M.S. thesis, University of Stellenbosch, April 2003. (Cited on pages xii, 64, and 65.)
- [4] T. Niesler, "Advanced pattern processing," Cambridge University Engineering Department, SVR Group I10, 1999/00. (Cited on pages xiii, 46, 67, 68, and 71.)
- [5] A.N. Pelech, "Evaluation of the pediatric patient with a cardiac murmur," *Pediatric Cardiology*, vol. 46, no. 2, pp. 167–188, 1999. (Cited on pages xvi, 2, 9, 14, 15, 19, 20, 21, 22, and 44.)
- [6] R.A. Wojcik J.S. Lombardo W.R. Thompson C.S. Hayek, C. Tuchinda, "Wavelet processing of systolic murmurs to assist with clinical diagnosis of heart diseases," *Biomedical instrumentation and technology Association for the Advancement of Medical Instrumentation*, vol. 4, no. 37, pp. 263–270, Jul-Aug 2003. (Cited on pages xvi, 5, 24, 61, and 86.)
- [7] M. Burch, "Congenital heart disease," *Journal of Paediatrics, Obstetrics and Gynaecology*, pp. 5–14, Sep/Oct 2003. (Cited on pages 1 and 22.)
- [8] C. Tuchinda J.K. Telford W.R. Thompson, C.S. Hayek, "Automated cardiac auscultation for detection of pathological heart murmurs," *Pediatric Cardiology*, vol. 22(5), pp. 373 – 9, Sept-Oct 2001. (Cited on pages 1, 2, 5, and 25.)
- [9] F.P. Stocker J.P. Pfammatter, "Delayed recognition of haemodynamically relevant congenital heart diseases," *European Pediatric Journal*, vol. 160, pp. 231–234, 2001. (Cited on page 1.)
- [10] D.W. Hannon M.E. Mcconnell, S.B. Adkins III, "Heart murmurs in pediatric patients: When do you refer?," *American Family Physician*, Aug 1999. (Cited on pages 1, 2, and 22.)

- [11] SA Statistics, "Census 2001: Census in brief," *Statistics South Africa.*, vol. Pretoria: Statistics South Africa, 2003. (Cited on page 2.)
- [12] Macro International. Department of Health, Medical Research Council, "South africa demographic and health survey 1998. full report.," Full report, Pretoria: Department of Health., 2002. (Cited on page 2.)
- [13] N. Nannan D. Bradshaw, D. Bourne, "What are the leading causes of death among south african children?," Tech. Rep. Nr. 3, MRC Policy Brief, December 2003. (Cited on page 2.)
- [14] M. Akay, "Time frequency and wavelets in biomedical signal processing," *Series in Biomedical Engineering*, pp. 271–301, 1997. (Cited on page 2.)
- [15] P. Nutting P. Franks, C. Clancy, "Gatekeeping revisited - protecting patients for overtreatment," *New England Journal of Medicine*, vol. 327, pp. 427–429, 1992. (Cited on page 2.)
- [16] N.S. Talner P.R.A. Gaskin, S.E. Owens, "Clinical auscultation skills in pediatric residents," *Pediatrics*, vol. 105, pp. 1184–1187, 2000. (Cited on page 2.)
- [17] R.A. Waugh G.R. Correy J.R. Feussner E.W. St Clair, E.Z. Oddone, "Assessing house staff diagnostic skills using a cardiology patient simulator," *Annual Internernal Medecine*, vol. 177, pp. 751–756, 1998. (Cited on page 2.)
- [18] J.S. Alpert, "Cardiology for the primary care physician," *Current Medicine*, 1996. (Cited on page 2.)
- [19] K. Steyn D. Bradshaw, "Poverty and chronic diseases in south africa: Technical report 2001.," Tech. Rep. ISBN: 1-919809-17-1, Medical Research Council, 2002. (Cited on page 4.)
- [20] "Statistics south africa: Statistical release p0302 mid-year estimates. (various years)," Tech. Rep., Statistics South Africa, 2003. (Cited on page 4.)
- [21] D. Hudson M. Coehn, "Comparative approaches to medical reasoning," *World Scientific*, pp. 271–288, 1995. (Cited on pages 5 and 25.)
- [22] E. Trowitzsch D. Barschdorff, U. Femmer, "Automatic phonocardiogram signal analysis in infants based on wavelet transforms and artificial neural networks," *Computer Cardiology*, vol. 7, pp. 753–756, 1995. (Cited on pages 5 and 25.)
- [23] H. Kobayashi M. Okuni, T. Hasimoto, "Trial of a new cardiac mass screening system in school children.," *Japanese Circulation Journal*, vol. 42, pp. 49–52, 1978. (Cited on page 5.)
- [24] J.L. Reynolds, "Heart disease screening of preschool children," *American Journal of Disable Children*, vol. 119, pp. 488–493, 1970. (Cited on page 5.)

- [25] W.B. Collis T.S. Leung, P.R. White, "Analysing pediatric heart murmurs with discriminate analysis," *Proc IEEE.*, vol. 18, pp. 1628–1631, 1998. (Cited on page 5.)
- [26] R.A. O'Rourke V. Fuster, R.W. Alexander, *The Heart*, vol. 1, The McGraw-Hill Companies, Inc., 10th edition edition, 2001. (Cited on pages 10, 15, 16, 18, 19, 20, 21, 22, and 26.)
- [27] L. Tarassenko L.A. Smith P.E. McSharry, G.D. Clifford, "A dynamic model for generating synthetic electrocardiogram signals," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 3, pp. 289 – 294, March 2003. (Cited on page 13.)
- [28] A. Ravin, "Auscultation of the heart," *Yearbook Medical Publishers*, p. 15, 2000. (Cited on pages 15 and 34.)
- [29] A.A. Luisada A. Soffer, A. Feinstein, "Glossary of cardiologic terms related to physical diagnosis and history," *American Journal of Cardiology*, vol. 20, pp. 285–286, 1967. (Cited on page 15.)
- [30] J. Perloff, *The Clinical Recognition of Congenital Heart Disease*, Philadelphia: Saunders, 4th edition edition, 1994. (Cited on page 18.)
- [31] M.A. Tinati, *Time-frequency and time-scale analysis of phonocardiograms with coronary artery disease before and after angioplasty*, Ph.D. thesis, Department of Electrical and Electronic Engineering, The University of Adelaide, 1998. (Cited on page 20.)
- [32] E. Braunwald R.A. O'Rourke, L. Goldman, *Cardiology for the Primary Physician: Cardiac Murmurs.*, Philadelphia: Saunders, 1998, p. 155-173. (Cited on page 20.)
- [33] M.E. Gallaher R.E. Durnin, R.E. Stanton, "Heart-sound screening in children.," *Journal American Medical Association*, vol. 203, pp. 111–116, 1968. (Cited on page 24.)
- [34] R.J. Lehner R.M. Rangayyan, "Phonocardiogram signal analysis: a review," *Critical Review Biomedical Engineering*, vol. 15, pp. 211–236, 1988. (Cited on page 24.)
- [35] E.McDonnell P. Bentley, "Wavelet analysis of cardiovascular signals.," *Signal Processing VII, theories and applications. Proceedings of EUSIPCO-94 Seventh European Signal Processing Conference*, vol. 1, pp. 78–81, 1994. (Cited on page 24.)
- [36] J.S. Kan B.W. McCrindle, K.M. Shaffer, "Cardinal clinical signs in the differentiation of heart murmurs in children.," *Arch Pediatric Adolescent Medicine*, , no. 150, pp. 169–174, 1996. (Cited on pages 25, 86, and 91.)
- [37] J. Hertzberg C.G. DeGroff, S. Bhatikar, "Artificial neural network-based method of screening heart murmurs in children," *Circulation*, pp. 2712–2716, June 2001. (Cited on pages 25, 67, 69, 70, and 98.)

- [38] H. J. Gavarini, *Viscosity*, Transtronics, 3209 W.9th street Lawrence, KS 66049 USA, 2003, University unknown. (Cited on page 26.)
- [39] H. Power, *Bio-fluid Mechanics*, vol. 3 of *Advances in Fluid Mechanics*, Coputational Mechanics Publications, Wessex Institute of Technology, Southampton, UK, 1995. (Cited on page 26.)
- [40] H. Kutchal M. Gad-El-Hak, J.B. Morton, "Turbulent flow of red cells in dilute suspensions. effect on kinetics of o2 uptake.," *Biophysics Journal.*, vol. 18, no. 3, pp. 289–300, June 1977. (Cited on page 27.)
- [41] E. Arild T. Hasvold L.B. Dahl, P. Hasvold, "May heart murmurs be assessed by telemedicine?," *Tidsskr Nor Laegeforen.[PubMed]*, vol. 6, no. 123(21), pp. 3021–3, Nov 2003. (Cited on page 32.)
- [42] J.R. Peebles Z. Peyton, *Probability, Random Variables and Random Signal Principles*, Electrical Engineering. McGraw-Hill International, 4th edition, 2001. (Cited on page 35.)
- [43] M. Nasor M.J. Burke, "The time relationships of the consituent components of the human electrocardiogram," *Journal of Medical Engineering & Technology*, vol. 26, no. 1, pp. 1 –6, January/February 2002. (Cited on pages 37 and 50.)
- [44] J. Agzarian D. Abbott L.T. Hall, J.L. Maple, "Sensor system for heart sound biomonitor," *Microelectronics Journal*, vol. 31, pp. 583–592, 2001. (Cited on pages 38, 41, 122, and 125.)
- [45] G. Oppenheim J-M Poggi M. Misiti, Y. Misiti, *Wavelet Toolbox- For use with Matlab*, MatLab, 2 de edition, July 2002. (Cited on pages 39, 67, 71, 122, and 123.)
- [46] J. Kovacevic M.Vetterli, *Wavelets and Subband Coding*, Prentice-Hall, Englewood Cliffs, NJ, 1995, pp.201-298. (Cited on page 39.)
- [47] J.G. Harris M.D. Skowronski, "Increased mfcc filter bandwidth for noise-robust phoneme recognition," *International Conference on Acoustic, Speech and Signal Processing*, vol. I, pp. 801–804, 2002. (Cited on page 46.)
- [48] R.L. Mahajan S.R. Bhatikar, *Artificial Neural Network Based Diagnosis of CVD Barrel Reactor*, Center for Advanced Manufacturing and Packaging of Microwave, Optical and Digital Electronics, Department of Mechanical Engineering, University of Colorado, Boulder, CO 80309-0427, USA. (Cited on page 63.)
- [49] Z. Dokur T. Olmez, "Classification of heart sounds using an artifical neural network," *Pattern Recognition Letters*, vol. 24, pp. 617–629, 2003. (Cited on page 64.)
- [50] J.M. Zuranda I. Cloete, *Knowledge-Based Neurocomputing*, MIT Press, Cambridge, MA, 1999. (Cited on page 65.)

- [51] M. Beale H. Demuth, *Neural Network Toolbox 4.0.1 Release Notes*, MatLab, release 13 edition, 2002, Neural Networks User's Guide. (Cited on page 70.)
- [52] P.F. Rice V.E. Congelosi, P.H. Taylor, *Basic Statistics - A Real World Approach*, West Publishing Company, 3 rd. edition edition, 1983. (Cited on pages 74 and 75.)
- [53] R. L. Mason Y.M. Chou, M. Polansky, "Transforming non-normal data to normality in statistical process control," Tech. Rep., StatsLib hosted by the Department of Statistics at Carnegie Mellon University, 1998. (Cited on page 75.)
- [54] F.Yates R.A. Fisher, *Statistical Tables for Biological, Agricultural and Medical Research*, Oliver & Boyd, Edinburgh, 5 th edition, 1958. (Cited on page 75.)
- [55] R.H. Swanton, *Cardiology*, Roche CardiaCare, 4th edition, 1998. (Cited on page 86.)
- [56] R. Moore, "Jackknife error estimates," N/A, 1999. (Cited on page 98.)
- [57] Altera Inc., "Dsp blocks in stratix devices," Tech. Rep., Altera, Altera Corporation, 101 Innovation Drive, San Jose, California 95134, USA. (Cited on page 109.)
- [58] S.M. Panas L.J. Hadjileontiadis, "A wavelet-based reduction of heart sound noise form lung sounds," *International Journal of Medical Informatics*, vol. 52, pp. 183–190, 1998. (Cited on page 123.)

