Dinesh Kumar

AUTOMATIC HEART SOUND ANALYSIS FOR CARDIOVASCULAR DISEASE ASSESSMENT

UNIVERSIDADE DE COIMBRA

Dinesh Kumar

AUTOMATIC HEART SOUND ANALYSIS FOR CARDIOVASCULAR DISEASE ASSESSMENT

Doctoral thesis submitted to the Doctoral Program in Information Science and Technology, supervised by Prof. Dr. Paulo Fernando Pereira de Carvalho and Prof. Dr. Manuel de Jesus Antunes, and presented to the Department of Informatics Engineering of the Faculty of Sciences and Technology of the University of Coimbra.

September 2014

UNIVERSIDADE DE COIMBRA

FCTUC FACULDADE DE CIÊNCIAS
E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

September 2014

Thesis submitted to the
University of Coimbra
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Information Science and Technology

This work was carried out under the supervision of

Professor Paulo Fernando Pereira de Carvalho

Professor Associado do
Departamento de Engenharia Informática da
Faculdade de Ciências e Tecnologia da
Universidade de Coimbra

and

Professor Doutor Manuel J Antunes

Professor Catedrático da
Faculdade de Medicina da
Universidade de Coimbra

ABSTRACT

Cardiovascular diseases (CVDs) are the most deadly diseases worldwide leaving behind diabetes and cancer. Being connected to ageing population above 65 years is prone to CVDs; hence a new trend of healthcare is emerging focusing on preventive health care in order to reduce the number of hospital visits and to enable home care. Auscultation has been open of the oldest and cheapest techniques to examine the heart. Furthermore, the recent advancement in digital technology stethoscopes is renewing the interest in auscultation as a diagnosis tool, namely for applications for the homecare context. A computer-based auscultation opens new possibilities in health management by enabling assessment of the mechanical status of the heart by using inexpensive and non-invasive methods. Computer based heart sound analysis techniques facilitate physicians to diagnose many cardiac disorders, such as valvular dysfunction and congestive heart failure, as well as to measure several cardiovascular parameters such as pulmonary arterial pressure, systolic time intervals, contractility, stroke volume, etc.

In this research work, we address the problem of extracting a diagnosis using analysis of the heart sounds. Heart sound analysis consists of three main tasks: i) identification of non-cardiac sounds which are unavoidably mixed with the heart sound during auscultation; ii) segmentation of the heart sound in order to localize the main sound component; and finally, iii) classification of the abnormal heart sounds from the normal heart sounds and in case of abnormal sounds classification is performed further to identify the type of the abnormal sound.

In most previous works on heart sound analysis these three problems were tackled jointly. Two main classes of approaches are found in past years: first is to perform analysis using an auxiliary signal, such as ECG, sound signal for ambient noise, carotid pulse, etc which is acquired during auscultation; and the second is without the support of any auxiliary signal. The first group of methods can be ruled out for the sake of keeping analysis cost effective and convenient to the users, however, the second group of approaches is dependent on the use of regularity of heart sound components which may not be found in various cardiac disorders. Hence, our intention aims to develop algorithms for the each problem related to heart sound analysis with the following main requirements: i) not to use any auxiliary; ii) applicable to any age, weight, sex, body proportion subject; iii) robustness.

Noise detection technique uses the template matching based approach. We propose a new method applicable in real time to detect ambient and internal body nois-

es manifested in heart sound during acquisition. The algorithm is designed on the basis of the periodic nature of heart sounds and physiologically inspired criteria. A small segment of uncontaminated heart sound exhibiting periodicity in time as well as in the time-frequency domain is first detected and applied as a reference signal in discriminating noise from the sound.

For the segmentation problem, the heart is considered as a nonlinear dynamical system. In a first processing stage, Lyapunov exponents are computed from the phase space in order to identify the presence of murmur in the heart sound. Based on this information, the method enters one of two segmentation methods: for heart sound with murmur, an algorithm based on adaptive wavelet-nonlinear feature analysis is applied; for heart sounds without murmur, a less complex approach is followed based on the signal's envelope analysis. Recognition of S1 and S2 sounds is achieved by inspecting high-frequency signatures. This marker is physiologically motivated by the accentuated pressure differences found across heart valves, both in native and prosthetic valves, which leads to distinct high-frequency signatures of the valve closing sounds.

Since heart murmurs originate from numerous anomalies in the heart, these show different characteristics, temporal, frequency and complexity. These features are extracted in order to classify heart murmur using a supervised classifier. A new set of 17 features extracted in the time, frequency and in the state space domain is suggested. The features applied for murmur classification are selected using the floating sequential forward method (SFFS). Using this approach, the original set of 17 features is reduced to 10 features. The classification results achieved using the proposed method are compared on a common database with the classification results obtained using the feature sets proposed in two well-known state of the art methods for murmur classification.

These algorithms have been tested on the database prepared with the help of the University Hospital in Coimbra. Three classes of the heart sound databases were prepared, i) normal heart sounds from the native valves; ii) abnormal heart sounds from the native valves; iii) heart sounds from artificial valve implants. Experimental results suggest that the algorithms can be applicable for cardiac application.

Keywords: Heart sound, valvular disease, prosthetic valve, segmentation, heart murmur, classification, noise detection, time-frequency analysis, wavelet decomposition, nonlinear dynamics.

SUMÁRIO

As doenças cardiovasculares são a maior causa de morte em todo o mundo, ultrapassando de forma significativa a mortalidade devida aos diabetes e ao cancro. Dado o envelhecimento acentuado da população mundial e atendendo a que a incidência das doenças crónicas, em particular das doenças cardiovasculares, está fortemente correlacionada com a idade, observa-se uma nova tendência de cuidados de saúde focada nos cuidados de saúde preventivos, com vista a reduzir o número de episódios agudos numa tentativa de reduzir custos e de melhor a qualidade de vida dos pacientes. A implementação destas estratégias requer sistemas fiáveis de telemonitorização, porém simples e baratos por forma a permitir o seu uso prolongado por populações pouco motivadas e muitas vezes com iliteracia tecnológica.

A auscultação é a técnica mais antiga e menos onerosa para examinar o coração. Os recentes avanços em estetoscópios com tecnologia digital está a renovar o interesse no uso da auscultação como ferramenta de diagnóstico, nomeadamente para aplicações no contexto da tele-monitorização e da gestão da doença. Uma auscultação assistida por computador abre novas possibilidades na gestão da saúde ao permitir a avaliação do estado mecânico do coração, usando-se métodos não invasivos e pouco onerosos. As técnicas de análise de som cardíaco assistidas por computador facilitam o diagnóstico médico de muitos problemas cardíacos, tais como a disfunção valvular e a insuficiência cardíaca congestiva, permitindo também medir vários parâmetros cardiovasculares, como por exemplo a pressão arterial pulmonar, intervalos de tempo sistólicos, contractilidade, volume sistólico, entre outros.

Neste trabalho de investigação abordamos o problema de produzir um diagnóstico usando a análise de sons cardíacos. De uma forma geral, a análise de sons cardíacos consiste em três tarefas principais: i) na identificação de contaminações por sons não cardíacos, que se misturam inevitavelmente com o som do coração durante a auscultação; ii) na segmentação do som cardíaco por forma a localizar as sua principais componentes; e finalmente iii) na classificação dos sons cardíacos anormais e distinção dos sons cardíacos normais. No caso da existência de sons anormais, procede-se a uma classificação, por forma a identificar o tipo de som irregular.

Na literatura sobre análise de som cardíaco, estes três problemas têm sido abordados conjuntamente. De uma forma geral, podemos afirmar que emergiram duas classes de abordagens principais: a primeira classe consiste na condução da análise usando um sinal auxiliar, tal como o ECG, um sinal sonoro para o ruído ambiente, o

pulso da carótida, etc., sinal esse que é adquirido durante a auscultação; na segunda classe de métodos a tarefa é realizada sem o suporte a qualquer sinal auxiliar. O primeiro grupo de métodos não é muito prático para aplicações clínicas, dado o seu custo superior e o facto de exigir um número de sensores mais elevado o que, necessariamente, obriga a uma complexidade logística e operacional superior. No entanto, o segundo grupo de abordagens depende do uso da regularidade de componentes do som cardíaco, que poderá não ser detectada em vários problemas cardíacos. Deste modo, o objectivo deste trabalho é o de contribuir para a solução de cada um dos problemas principais relacionados com a análise do som cardíaco, dando resposta a alguns dos requisitos principais para uso clínico, ou seja: i) simplicidade pela ausência de uso de sinais auxiliares; ii) capacidade de generalização, isto é, aplicabilidade em populações heterogéneas; iii) robustez.

Na metodologia proposta para a detecção de ruído usa-se uma abordagem baseada em correspondência de modelos (*template matching*). Propomos um novo método aplicável em tempo real para detectar contaminações por fontes de ruído ambiente e fontes de ruído fisiológicos que se manifestam no som cardíaco durante a sua captura. O algoritmo está concebido com base na natureza periódica dos sons cardíacos e em critérios fisiológicos. Um pequeno segmento de um som cardíaco não contaminado é primeiramente detectado e aplicado como um sinal de referência na distinção de ruído no som cardíaco.

Para o problema da segmentação, o coração é considerado como um sistema dinâmico não linear. Numa primeira fase de processamento, os expoentes de Lyapunov são computados a partir do espaço de fase, por forma a avaliar a presença de murmúrios no som cardíaco. Com base nesta informação, o método usa um de dois algoritmos de segmentação: para o som cardíaco com murmúrio, aplica-se um algoritmo baseado numa análise de características de onduletas não lineares adaptativas; para sons cardíacos sem murmúrio, segue-se uma abordagem menos complexa, baseada na análise de envelope do sinal. O reconhecimento dos sons S1 e S2 é conseguida através do exame de uma assinatura obtida com base na informação de alta frequência. Este marcador é motivado fisiologicamente pelas acentuadas diferenças de pressão encontradas nas válvulas cardíacas, sejam estas naturais ou prostéticas, o que origina assinaturas de alta frequência distintas nos sons de fecho destes elementos.

Uma vez que os murmúrios cardíacos derivam de várias anomalias no coração, estes últimos mostram características diferentes a nível temporal, de frequência e de complexidade. Estas características são extraídas por forma a classificar-se o murmúrio cardíaco usando um classificador supervisionado. Sugerimos um novo conjunto de 17 características extraídas no tempo, frequência e no domínio do espaço de fase. As características usadas para a classificação de murmúrios são seleccionadas usando o método SFFS (*floating sequential forward method*). Ao usarmos esta abordagem, o conjunto original de 17 características é reduzido para 10. Os resultados da classifi-

cação conseguidos usando o método proposto são comparados com dois métodos do estado da arte usando uma base de dados comum.

Palavras-chave: Som cardíaco, doença valvular, válvula prostética, segmentação, sopro/murmúrio cardiac, classificação, análise tempo-frequência, decomposição de onduletas, dinâmica não linear.

ACKNOWLEDGMENTS

"Saying thank you is more than good manners. It is good spirituality."

Alfred Agache (29 Aug 1843- 15 Sep 1915)

Biomedical signal processing is the field of research that I opted to pursue. My interest in this subject comes way back from my college days during my graduation. Having graduated in electrical engineering, I wished to investigate its principles and methodologies that were taught to me during my graduation and apply then directly in human healthcare. Therefore, my first words of thanks go to Professor Bernardete Ribeiro and Professor Paulo de Carvalho who gave me access to start work on project, *MyHeart*, which had such objectives. Without whom, I could not have started this work.

I would like to express my sincere gratitude to my supervisors, Professor Paulo de Carvalho and Professor Doutor Manuel Antunes, for accepting me as one of the candidates to be guided with their experience and excellence. Their scientific support, discussion, motivation were fundamental to the conclusion of my work.

This work has been carried out under the project financed by the Science and Technology Foundation of Portugal (FCT) named *SoundForLife*. However, part of the work was initiated under the project *MyHeart* sponsored by the European Union. Therefore, I would like to thank my super supervisor, Professor Paulo de Carvalho for keeping me engaged in the project from where I have consistently received financial support. My gratitude goes to the Foundation for providing me financial assistance throughout the tenure of the projects and other allowances for personal family visits to my homeland once in year.

My gratitude is also extended to Professor Jorge Henriques and Professor Rui Pedro Paiva for their critical reviews of my articles providing me valuable suggestions and support to finalize them.

Thanks go also to CISUC, Center for Informatics and Systems of the University of Coimbra, where most of this work was developed, for the logistic and computer facilities and for the financial support regarding acquisition of material and participation in several scientific conferences in the area. I would like to express my gratitude to all the staff, researchers and collaborators for the pleasant environment and conditions offered.

The cardiothoracic surgery centre of the University Hospital of Coimbra (HUC)

has been the main partner of the project under which this work has been carried out. Therefore, I would like to extend my gratitude to Doctor Luis Eugénio, Doctor Maria Maldonado and Doctor Antonio Sá e Melo of the department of cardiology, for their medical insights and supervision in the auscultation in process and preparation of the heart sound database.

I would like to thank my friends in Portugal, specially Pedro Martins and Jose Bernardo Monteiro, who have given me their kindness, moral support, and a lot of funny moments. I have had several moments of great fun with them. Those moments were very important to me because sometimes they played as stress buster and clam refreshment to the mind, fundamental in order to sustain mental balance. Friends are indeed an integral element of one's life. My thanks also to their families for welcoming me with open hearts and letting me feel connected with love. I also thank to my old college friends, Kallol Roy and Vivek Rastogi, who inspired me to take a plunge into the research world after my graduation which eventually ended up with me enrolling in the PhD program.

I am also grateful to Joao Pedro Ramos, Tiago Sapata, Liandro do Vale, that they tested my algorithms and highlighted bugs, allowing them to be rectified in time.

I am extremely grateful to my parents, Mr. Ram Nayan and Mrs. Badama Devi, my brothers, Neeraj and Sharad, and wife Amita, for their love, affection and faith in my abilities. They were among the major driving forces that led me to keep going ahead in my research work. I would like to apologies to my family for not being able to visit them frequently. I submit my special gratitude to my mother for her patience and for the understanding that she showed in letting me choose my own time of convenience in visiting her despite missing me badly. Some words of gratitude go to my wife as living with her I was continuously reminded that my thesis must be finished.

Finally, I thank to those whose names have not been mentioned here but in one way or another they contributed to my research work.

Coimbra, September, 2014

Dinesh Kumar

The research work presented in this thesis was partially supported by the Portuguese Ministry of Science and Technology under the PhD grant SFRH/BD/48456/2008, project SoundForLife (FCOMP-01-0124-FEDER-007243) of program COMPETE/QREN, and project HeartSafe (PTDC/EEI-PRO/2857/2012).

CONTENTS

Abstract ..... **iii**
Sumário ..... **v**
Acknowledgments ..... **ix**
Contents ..... **xi**
List of Figures ..... **xv**
List of Tables ..... **xxi**
Main Abbreviations ..... **xxiii**
Chapter 1 Introduction **1**
1.1. Motivation 3
1.1.1. Diagnostic Values of Auscultation 5
1.1.2. Socioeconomic Impacts 6
1.1.3. Personalized Medical Care (pHealth) 7
1.1.4. Application for Training Medical Professionals 7
1.2. Objectives 8
1.3. Thesis Outline 11
1.4. Main Contributions 13
1.4.1. Noise detection 14
1.4.2. Segmentation 14
1.4.3. Heart Murmur Classification 15
1.4.4. List of Publications 15
Chapter 2 Origin of Heart Sound and Auscultation **18**
2.1. Introduction 19
2.2. Heart's Anatomy and Physiology 20
2.2.1. Heart Valves 22
2.2.2. Conduction System of the Heart 24
2.2.3. Cardiac Cycle with Pressure Profile 24
2.3. Heart Sound and Auscultation 26

2.3.1. Heart Sounds 28
2.3.2. Abnormal Heart Sound or Heart Murmur 30
2.3.3. Prosthetic Valve Sounds 35
2.4. Heart Sounds' Diagnostic and Prognostic Values 36
2.4.1. Cardiac Diseases 37
2.4.2. Physiological Measurements 39
2.5. Conclusions 41
Chapter 3 Heart Sound Analysis: Literature Review 43
3.1. Noise Detection in Heart Sound 44
3.1.1. Multi-channel Signal Approaches 45
3.1.2. Single Channel Signal Approaches Approaches 48
3.2. Heart Sound Segmentation 53
3.2.1. Approaches with an Auxiliary Signal 53
3.2.2. Approaches with Only Heart Sound Signal 54
3.3. Cardiac Murmur Classification 60
3.4. Limitations 65
3.5. Conclusions 67
Chapter 4 Noise Detection 70
4.1. Overview 71
4.2. Phase I: Reference Sound Detection 73
4.2.1. Periodicity in Time Domain 73
4.2.2. Periodicity in Time-Frequency Domain 78
4.2.3. A Heart Cycle Template Selection 85
4.3. Phase II: Non-cardiac Sound Detection 85
4.3.1. Spectral Energy 85
4.3.2. Temporal Energy 86
4.4. Test Database 89
4.5. Performance Assessment 90
4.6. Experiment Results and Discussions 91
4.7. Conclusions 96
Chapter 5 Heart Sound Segmentation 98
5.1. Introduction 101

5.2. Heart Sound Segmentation Scheme 103
5.3. Heart Sound Classification Using Chaos 104
5.3.1. Heart Sound Embedding 104
5.3.2. Largest Lyapunov Exponents for Heart Sound Type Detection 108
5.4. Preprocessing for Segmentation Algorithms 110
5.5. Segmentation of Heart Sounds without Murmur: Energy Envelope Method 111
5.5.1. Delimitation of Audible Sounds 112
5.5.2. Recognition of S1 and S2 Sound Components 113
5.6. Segmentation of Heart Sound with Murmur: Wavelet Decomposition and Non-linear based Segmentation (WD-NLF) 122
5.6.1. Nonlinear Features (Simplicity and Strength) 124
5.6.2. Phase I: Sound Delimitation 126
5.6.3. Phase II: Recognition of S1 and S2 130
5.7. Test Database 132
5.8. Performance Measurements 133
5.9. Experimental Results and Discussions 134
5.10. Conclusions 154
Chapter 6 Heart Murmur Classification 155
6.1. Cardiac Lesions Classification through Murmurs 156
6.2. Preprocessing and Beat Detection 158
6.3. Feature Extraction 158
6.3.1. Temporal Domain Features 159
6.3.2. Frequency Domain Features 162
6.3.3. Statistical Domain 166
6.3.4. Feature Set -II (Olmez *et al.*) 171
6.3.5. Feature Set -III (Ahlstrom *et al.*) 172
6.3.6. Feature Set- IV 176
6.4. Feature Selection and Classifier 176
6.5. Results and Discussions 178
6.6. Conclusions 180
Chapter 7 Conclusions and Perspectives 184

7.1. Summary and Conclusions 184
7.2. Perspectives for Future Research 187
Bibliography 191

LIST OF FIGURES

Figure 1.1. One heart beat or one heart cycle. Upper curve corresponds to the ECG (electrocardiogram) and the lower curve shows the phonocardiogram trace or heart sound. 4

Figure 1.2. An example of a heart sound corrupted with noise, (upper curve) Indication of clean and noisy cardiac sound, (lower curve) cardiac sound with segments of non-cardiac sounds (sound of speech and movement in stethoscope's probe). 9

Figure 1.3. An example of segmentation performed on normal heart sound containing mainly S1 and S2 sounds. Duration of the heart sound components are marked by the length of the bar shown above the sounds. 10

Figure 1.4. Blocks in identification of normal heart sound (heart sound produced from the native valves) and heart murmur (abnormal sound), and further if heart sound is identified as heart murmur then further classification into the type of murmur. 10

Figure 2.1. (Upper curve) Electrocardiogram ECG, (Lower curve) Heart sound acquired from native heart valves. The heart sound shows three main heart sound components, i.e., S1, S2 and S3. The third heart sound is a pathological sound in the case of adults. 20

Figure 2.2. (left) Anatomy of the heart and physiology (image is taken from (Malmivno and Plonsey 2009)), (right) blood flow through the chambers of the heart. (Image was taken from (Iaizzo 2005)) 21

Figure 2.3. (Upper) Native valves; (Lower left) Mechanical valves; (Lower right) Bio-prosthetic valves (figures are from (Yoganathan 2000)). 22

Figure 2.4. (Right) Cardiac electrical conduction system morphology and timing of action potentials from different of the heart. (Left) related ECG signal as measured on the body surface (figure is modified from (Malmivno and Plonsey 2009)) 23

Figure 2.5. Various cardiac events correlated in time in various curves: (from upper to lower) ECG, ventricular, aortic pressure curve, left ventricular pressure curve, left atrial pressure curve, heart sound. (Wiggers diagram was adapted from (Berne and Levy 1997)) 25

Figure 2.6. Estimated equal loudness level contours derived from Fletcher and Munson. The dashed line shows the threshold of hearing. The contours at 100 phons are drawn by a dotted line because data from only one institute are available at 100 phons. Each curve represents a sound which is perceived to have equal loudness for all frequencies. ..27

Figure 2.7. Heart sound samples along with their frequency concentration over time in presence of S3 and S4, respectively, are shown in (a) and (b). Figures (c) to (f) are magnified view of typical each sound components, i.e. S1, S2, S3, and S4. Lowest row of figures (g) to (j) are the time-frequency transform of the each components show the frequency distribution over time. In the images of time-frequency representation darker regions are exhibiting higher frequencies. (Color map: Darker space represents larger intensity of corresponding frequencies) 28

Figure 2.8. Heart murmurs (one heart cycle demonstration), systolic in (a)-(f) and diastolic in (g)-(h). Systolic murmurs are: Aortic Regurgitation in (a), Aortic Stenosis (b), Mitral regurgitation (c), Pulmonary Stenosis (d), Systolic Ejection (e), and Ventricular Septal Defect in (f). And the diastolic murmurs are: Tricuspid Regurgitation in (g) and Mitral Stenosis in (h).....33

Figure 2.9. Left figure is the time-frequency representation of a heart sound in mechanical valve (single tilted disk) at aortic position; and the right figure is time-frequency plot of the heart sound from biological valve at mitral position. 35

Figure 2.10. The main cardiac intervals in the process of cardiac functioning (figure is regenerated from (Ahlstorm 2008)). 40

Figure 3.1. Steps required in adaptive noise cancellation process in multi-channel signal approach. 47

Figure 3.2. Taped delay line adaptive line enhancement (ALE) filter model.....49

Figure 3.3. Steps required in segmentation for the sounds' components. 54

Figure 3.4. Segmentation of a heart sound using Liang's envelogram based method. The parameters c1 and c2 are the tolerance levels which are used to indentify diastolic and systolic intervals. 56

Figure 3.5. Segmentation of the murmur portion in a heart sound with aortic stenosis. Heart murmurs are segmented based on the simplicity that is below the threshold. In the example shown simplicity is computed using embedding dimension m=4 and delay = 19 parameters (see Chapter 5 for the details of these parameters). 59

Figure 3.6. Steps usually involved in cardiac murmur classification. 61

Figure 4.1. Flow chart of the proposed approach. 72

Figure 4.2. (a) Heart sound energy. (b) Heart sound envelope. (c) Autocorrelation function $y(\tau)$ of the envelope with peak identification (showing similar heart cycles in terms of vectors $y1\tau$ and $y2(\tau)$. 75

Figure 4.3. (a) Noisy heart sound energy. (b) Heart sound envelope. (c) Autocorrelation function $y(\tau)$ of the envelope with peak identification (showing dissimilar heart cycles in terms of vectors $y1(\tau)$ and $y2\tau$. ..77

Figure 4.4. (a) Normal heart sound from a single-tilted disc mechanical prosthetic valve in the mitral in the mitral position. (b) Spectrogram of (a). (c) Autocorrelation function (a) Noisy heart sound energy. (b) Heart sound envelope. (c) ACFs as $ASk(\tau)$ of spectral bands. 79

Figure 4.5. (a) Normal heart sound from native valve. (b) Spectrogram of (a). (c) Autocorrelation function. 82

Figure 4.6. (a) Noisy heart sound from valve. (b) Spectrogram of (a). (c) Autocorrelation function. 83

Figure 4.7.ROC curve for 13 heart sound samples to identify optimum cosine similarity threshold and number of aligned peaks (circled kink represents optimum value). 84

Figure 4.8. (a) Heart sound contaminated with several sources of noises. First 6 seconds correspond to reference template search window. The upper part of the graph shows noise classification result. (b) Spectral energy correlation between heart sound segments and the reference sound. (c) Relative temporal energy of the heart sound segments. 87

Figure 4.9. ROC curve for 13 heart sound samples to identify optimum spectral energy threshold as well as temporal energy (circled kink represents optimum value). 89

Figure 4.10. Sensitivity for noise detection in varying intensities of noise of the (a) heart sound from native valves, (b) prosthetic valve, (c) murmur, and (d) arrhythmia. 95

Figure 5.1. Main constituent blocks of heart sound segmentation. *Delimitation* determines onsets and offsets of the sounds which are addressed as segments or gates, and *Recognition* identifies the types of sounds. 101

Figure 5.2. Heart sound assessments and segmentation scheme. 103

Figure 5.3. Reconstruction parameters computation : (a) τ =195 first minimum in mutual information curve; (b) m=4 finding first kick in the prediction error using Cao's method (Cao 1997). 106

Figure 5.4. (a) Heart sound (normal). (b) Phase space reconstructed using delay

parameter (τ) 185 and dimension (m) 4. 107

Figure 5.5. Reconstructed phase (m=5, τ=64) of heart sound with murmur (mitral regurgitation), and magnified section is the display of diverging trajectories. 109

Figure 5.6. Largest Lyapunov exponent (LLE) estimation. 110

Figure 5.7. Illustration of the energy envelop method. 112

Figure 5.8. High frequency energy above the threshold: (a) Mechanical valve in aortic position, (b) native valve, (c) Mechanical valve in mitral position, (d) Prosthetic valve. 114

Figure 5.9. Step involved in valid heart cycles detection. 115

Figure 5.10. (a) 2-classes clustering of squared-energy peaks, (b) heart sound with high frequency signature in S2 sounds detected at 6th level of details coefficients. 116

Figure 5.11. Correlation between heart rate and the length of systole, diastole, and R-R interval. The length systole has a decreasing trend but it is stable whereas diastole decreases significantly. Therefore S2 locate, in higher heart rates, relatively later within heart cycle. ms= millisecond ; bmp= beat per minute. (figure is redrawn from (El-Segaier et al. 2005)). 118

Figure 5.12. Cases of (a) missing HFS, (b) adjacent HFS. 120

Figure 5.13. Arrhythmic sounds S1 and S2 detection using for HFS and LFS. (a) systolic interval estimation in one cycle as in HFS-LFS-HFS sequence, (b) correct LFS for being S1sound candidate. 121

Figure 5.14. Illustration of the method based on wavelet decomposition and nonlinear features (WD-NLF) 124

Figure 5.15. (top signal) Heart sound with mitral regurgitation, (middle signal) Strength profile, (bottom signal) Simplicity profile. 126

Figure 5.16. Wavelet decomposition-nonlinear features segmentation; (a) Heart sound energy in a mitral regurgitation situation (grade V), (b) Strength curve and candidate segments, (c) Simplicity curve and candidate segments, (d) Region for S1 and S2 sounds, (e) Region with murmurs and other sounds, (f) segmented S1 and S2 sounds. 129

Figure 5.17. (top) Original sound wave and gated S1 (LFS) and S2 (HFS) sounds, (bottom) High frequency energy curve with demarcation of S1 and S2 sounds. 131

Figure 5.18. (a) Box-Wishker plot of LLE, (b) ROC plot of the th_L and th_H thresholds 134

Figure 5.19. Heart sound segmentation using the WD-NLF method: In the first

column (a-d) normal heart sound from native valve is segmented. In the second column (e-h) Heart sound from prosthetic valve is segmented. 143

Figure 5.20 Heart sound with murmur (Aortic stenosis) segmented by energy envelop method..... 144

Figure 5.21. Scheme for Liang et al. and Nigam and Priemer segmentation methods' results with the energy envelop and WD-NLF. 145

Figure 5.22 (a) and (b) are the Box-Whisker plot of the sensitivities and specificities obtained from Energy envelop (EE), Liand et al. and Nigam and Priemer methods on databaseHS and databaseProsValve. Similarly, (c) and (d) are the Box-Whisker plot of sensitivities and specificities from WD-NLF, Liang et al., and Nigam and Priemer method on databaseMurmur. 151

Figure 5.23 (Upper) Box-Whisker plots of the errors in start and stop times of S1 and S2 sounds obtained from EE approach, Liang et al., and Nigam and Priemer's simplicity approach. S1_start/stop and S2_start/stop are the instants of S1 and S2, respectively, from our methods, and S1_start/stop_Nigam are from Nigam and Priemer 's approach. (Lower) Box-Whisker plots of WD-NLF approach, Liang et al., and Nigam and Priemer's approach 152

Figure 6.1 Steps blocks for heart murmur classification. 157

Figure 6.2 (upper) One heart cycle or heart beat of the heart sound from systolic murmur patients is shown. (lower) Autocorrelation function of the murmur region; three consecutive periods (for k=2) that are used to compute jitter. 161

Figure 6.3 An example of systolic ejection. Energy of the heart sound with indication of the ascending and descending time durations. 162

Figure 6.4 (upper) One heart cycle of the heart sound in mitral regurgitation case. (lower) frequency peaks in power spectrum which are considered as fundamental frequencies in our work. 166

Figure 6.5 Box-Whisker plot of the feature in the 7 classes of the heart murmurs in our prepared database..... 170

Figure 6.6 (left) The d2 wavelet coefficients of one heart beat of the all heart sound with murmurs, (right) Power of the heart beats; (a) aortic regurgitation, (b) aortic stenosis, (c) mitral regurgitation, (d) pulmonary regurgitation, (e) pulmonary stenosis, (f) sub-aortic stenosis + ventricular septal defect, (g) systolic ejection. 172

Figure 6.7 (a) One heart beat of the heart sound from aortic regurgitation patient,

(b) Shannon energy envelop and the 9 picked values, (c) wavelet decomposed signal, and 11 equally partitioned segments to compute entropy. 173

Figure 6.8 Heart sound with murmur from a systolic ejection patient: (a) one heart cycle, (b) *eigenvalues 1-8* from the SVD of S-tranfrom, (c) S-tranfrom (whiter space represente higher intensities), (d) discritisized S-transform into a 4×4 map of features (numbers represent the features denoted *STmap1-16*). The histograms (f, j, h, and l) of the respective eigenfrequencies (e, g) and eigentimes (i, j) are shown on the right. 174

Figure 6.9 Example of bispectrum of the systolic ejection murmur shown **Figure 6.8**(a). In (a) bispectrum and bold triangle is non-redundant region, and in (b) the region of interest. The smaller triangles, 16, in (b) are the features *HOS 1-16*. 175

Figure 6.10 The recurrence plot of the heart sound from the systolic ejection case shown in **Figure 6.8**(a). From this recurrence plot 10 recurrence statistics are calculated. 176

Figure 6.11 Left Box-Wishker plot is of Sensitivities and the right one is of Specificities achieved. 180

LIST OF TABLES

Table 2.1 Pitch and quality of the different kind of cardiac abnormalities. (Table is borrowed from (Erickson 2003)) 31

Table 2.2 Clincal clues to heart disease from cardiac auscultation (this table is formed based on the given information in (Chizner 2008)). 32

Table 3.1 Features and classification for heart murmur classification by the researches in recent past years. In the result column, Se is sensitivity, Sp is specificity, and Au is accuracy. 62

Table 4.1 Induced noise sources. 90

Table 4.2 Results of noise detection in terms of sensitivity (*SE* in %). (*NA = Not Available) 93

Table 4.3 Results of noise detection in terms of specificity (SP in %), computed based upon false positive (FP). 94

Table 5.1 S1 and S2 recognition performance in terms of Sensitivity (SE) and Specificity (SP): *databaseHS* and *databaseProsValve* are segmented by energy envelop method and *databaseMurmur* by WD-NLF. 137

Table 5.2 S1 and S2 delimitation errors (average value with standard deviation) *databaseHS* and *databaseProsValve* are segmented by energy envelop method and *databaseMurmur* by WD-NLF. 138

Table 5.3 S1 and S2 recognition performance in terms of Sensitivity (SE) and Specificity (SP): *databaseMurmur* is segmented using the energy envelop method and *databaseHS* and *databaseProsValve* by WD-NLF algorithm 140

Table 5.4 S1 and S2 delimitation errors (average value with standard deviation): *databaseMurmur* is segmented using the energy envelop method and *databaseHS* and *databaseProsValve* by WD-NLF algorithm 141

Table 5.5 S1 and S2 recognition performance in terms of Sensitivity (SE) and Specificity (SP) by the Liang et al method of envelogram. 146

Table 5.6 Delimitation error obtained from Liang et al method of envelogram. ... 147

Table 5.7 S1 and S2 sounds recognition performance in terms of Sensitivity (SE) and Specificity (SP) by the Nigam and Primer method. 149

Table 5.8 Delimitation error from the approach of similarity measure Nigam and Priemer approach. 150

Table 6.1 Summary of the introduced features (Feature set I). And (*) star

represents the features which has been used for the heart murmur classification problem. 168

Table 6.2 Selected features in each feature set. 177

Table 6.3 Murmur classification results with three set of selected features. 179

MAIN ABBREVIATIONS

CVD	Cardiovascular Diseases
ECG	Electrocardiogram
PCG	Phonocardiogram
CAA	Computer-Aided Auscultation
ACF	Autocorrelation Function
SNR	Singal-to-Noise Ratio
BSS	Blind source separation
STFT	Short Time Fourier Transform
LLE	Largest Lyapunov Exponent
HFS	High Frequency Segments
LFS	Low Frequency Segments
WD	Wavelet Decomposition
NLF	Non-linear Features
PPA	Peak Peeling Algorithm
EE	Energy Enevelop
WD-NLF	Wavelet Decomposition- Nonlinear Features

Chapter 1

INTRODUCTION

Cardiovascular disease is the major cause of disability and premature death throughout the world (UN 2001). In the year of 2008 nearly 17.3 million people died from CVD worldwide, of which 80% the recorded deaths happened in middle and low income countries (WHO 2011). Furthermore, according to the recent report of the WHO (World Health Organization) the projected number of possible deaths worldwide due to CVD by 2030 in the range of 23.6 million, almost 23% by leaving behind other chronic diseases such as cancer, diabetes, perinatal condition, etc., if corrective measures are not taken (Mendis et al. 2011). In Europe, CVD is the main cause of death. The prevalence of chronic cardiovascular diseases was between 20% and 45% of all deaths until 1991 (Mintz et al. 1991). However, as recent as 2005 data suggests mortality rate due to CVD in Europe is above 35%. It varies from 35% in western European region to 60% in the east European countries (Reimer et al. 2006). The known risk factors for causing chronic CVD, namely heart attack and stroke, are smoking, alcoholism, obesity, hypertension, etc. Furthermore, besides these factors CVD is found to be directly related with ageing. A report released in 2002 by the United Nation suggests that world's population is ageing because of changed social aspects of human lives, for instance, late marriages, a lowering fertility rate, a low birth rate, etc., as well as advancements in the health care (UN 2001). Another socio-economic indicator shows an ageing population that is contributing less for the country's labor force. Averaged working young force per 65 years or over has declined by 40% till 2000 from 1950. Developed countries, namely European countries, are hitting the most with the age rising of the population. If one observes closely the ageing phenomenon it can be concluded that increasing life expectancy in the developed countries is contributing significantly to the rise in the percentage CVD patients. Worldwide life expectancy at birth is projected to be 76 years old by 2050, while in developed countries it is expected to be even as high as high as 81 years (UN 2001). Aged population is relatively more prone to cardiac disorders in comparison to younger population (Mladovsky et al. 2009). Ageing causes degradation in the functionality of the cardiovascular system that may eventually lead to CVD.

In order to deal effectively with this problem many health care system is going through a paradigm shift by adopting preventative care rather than relying only on curative care strategies. Hence, the solution to this health problem is believed to be preventive healthcare, i.e., controlling costs (social and economical) by reducing preventable health conditions. This is commonly believed to be achievable by adopting preventive lifestyle as well as early diagnosis. In this sense long term telemonitoring is a promising tool to achieve early diagnosis and, therefore, avoiding potentially life critical situations as well as aggressive and expensive treatments. In the same line of reasoning there is an increasing demand for monitoring and diagnosis support solutions applicable in hazard environments (e.g. in emergency units) which are able to effectively assist doctors in their mission, while avoiding to degrade the patient's quality of life. An increasingly important key issue is the reduction of intrusiveness and invasiveness of these systems in order to minimize their interference in daily living habits and routines. Therefore, current research trends in this direction are to integrate system solutions into ordinary daily objects, such as functional clothes with integrated textile or hard sensors. In order to be cost effective and usable for long time periods, these tools require intelligent data analysis algorithms to be able to autonomously perform diagnostic functions and to support users in managing their health conditions, hence requiring low computational algorithms that could be run in real-time using low power processing devices. This combination of functional clothes and integrated electronics with on-body processing is usually defined as intelligent biomedical clothes. Therefore, in the perspective of social, economic and medical implication of CVD, a new significant research trend has emerged which aims to provide an efficient personal health system (pHealth) for CVD managements (Habetha 2006a). Another purpose of these systems is to ease physician and diagnosticians in preparing their diagnoses on regular basis of both inpatients as well as outpatients.

Heart sound is one of the biosignals that can be used in development of personalized health system. Heart sound is a valuable bio-signal to perform timely diagnosis or even to measure non-invasively important bio-variables. It is a least explored biosignals which has been given focus in the last decades compared to other heart related biosignals such as Electrocardiogram or Echocardiogram. This doctoral research aims to provide an analysis tool for heart sound which is envisaged to be applicable in controlling certain cardiac disorders and measuring cardiac parameters.

Section 1.2 Motivation

First and foremost, the motivations behind carrying out this heart sound analysis task are introduced. Preliminaries of heart sounds and cardiac disorders pertaining to their components are described. The computer based heart sound analysis' social implication and its impact on heart care system are then discussed.

Section 1.2 Objectives

This section is devoted to the objectives set out in this research.

Section 1.3 Thesis Outline

The section contains the given roadmap to the thesis. The structure of the documents is presented the contents are outlined.

Section 1.4 Main Contributions

The chapter ends with the section on main contributions. The proposed methods with regarding the three modules in heart sound analysis are summarized. The papers published in journal and conference proceedings are listed, and their topics are briefly described.

1.1. Motivation

Heart sound is a valuable bio-signal to perform timely diagnosis or even to measure non-invasively important bio-variables (Zhang and Zhang 2005). It also provides instant information of the mechanical status of the heart that physicians can usually know by hearing it through a stethoscope. The act of listening to the sound produced by heart function via stethoscope is known as auscultation (Geddes 2005; Tavel 1967). However, over these years diagnoses of severe heart diseases rely on advanced medical technologies such as echocardiography[1] or MRI (Magnetic resonance imaging)[2]. Furthermore, even to characterize the heart's function cardiac parameters such as heart rate (HR), cardiac output (CO), blood pressure (BP), etc., are used which are essentially determined using mostly Electrocardiogram (ECG) (Couceiro et al. 2008; El-Asir and Mamlook 2002; Henriques et al. 2008), Photoplethysmogrm (PPG) (Liu et al. 2009), ICG (Impedance cardiogram) (Carvalho et al. 2010; Paiva et al. 2009b), etc. The use of these techniques has increased in the last few years due to continuous improvement in their potential to provide reliable tools and assess the mechanical status of the cardiac functions. In spite of the advance-

[1] Echocardiography, also known as cardiac ultrasound, is in fact a sonogram of the heart that utilizes ultrasound based imaging technique to visualize aiming to determine a possible pathology from subcutaneous body structures, such as tendons, muscles, joints, vessels, and other internal organs.

[2] MRI (Magnetic Resonance Imaging) is a medical imaging technique that is used in radiology to visualize detailed internal structures in the body. It uses a powerful magnetic field, radio frequency pulses and a computer to produce detailed pictures of organs, soft tissues, bone, and many others internal body structures.

ments in techniques for knowing the heart condition, stethoscope is still valuable equipment used by the international medical community. Several studies have proven that heart sound can be a potential informative signal which is comparable to other bio-signals emanating from the heart; it carries information regarding rhythm, time and frequency of the sounds based on which mechanical function of the heart can be assessed (Kudriavtsev et al. 2007; Mehta and Khan 2004; Tavel 1967; Zuckerwar et al. 1993). Abnormal sounds are produced when cardiac functions improperly which are usually manifested between the main components of the heart sounds. More specifically, auscultation of the heart is still the preferred method for the heart valve disorders and heart failure diagnosis. For instance, in comparison to echocardiography and cine-fluoroscopy[3] (Geise 2001) it exhibits a much higher sensitivity, as high as about 92%, than these methods for heart valve dysfunction diagnosis (Mintz et al. 1991). Automatic processing enables evaluation of timings between events: these were used in the past as a technique called phonocardiography[4] and have very high diagnostic value.

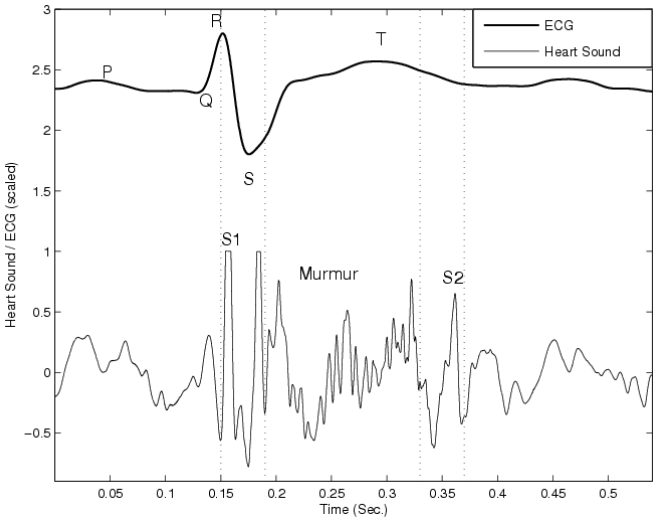


Figure 1.1. One heart beat or one heart cycle. Upper curve corresponds to the ECG (electrocardiogram) and the lower curve shows the phonocardiogram trace or heart sound.

[3] Cine-fluoroscopy is a method to record a motion picture of the body's internal organ movements; it has been extensively used to determine thrombus during prosthetic valve movements.

[4] A graphical representation of the waveform of cardiac sounds is named as phonocardiogram (PCG). Phonocardiography technique is acquiring the PCG signal.

1.1.1. Diagnostic Values of Auscultation

Experienced cardiologists are able to detect subtle heart disorders just by listening to the timbre and the sequence of its beats and murmurs. In fact, many heart disorders can be effectively diagnosed using auscultation techniques. In automatic heart sound processing, diagnostic information might be extracted from two main sets of features: morphological characteristics and existence of certain heart sound components and timing information between heart sound events.

Construction of heart sound (i.e. heart sound's components) and its origin from the heart will be explained in detail in the next chapter; therefore, only a preliminary introduction is presented here in order to highlight the relationship of some component of the heart sound and their diagnostic value. Digitized heart sound or phonocardiogram (PCG) is the graphical interpretation of the acoustic waves emanated from the functioning heart. Auscultation of the heart is performed namely for the heart valve disorders, heart failure and congenital heart lesions (Karnath and Thornton 2002). As it can be observed in **Figure 1.1**, there are two major heart sounds: S1 and S2, which are reasonably aligned with the R and T waves of the ECG, respectively. The S1 heart sound is heard at the onset of systole when the Atrioventricular valves close (Abbas 2009; Hebden and Torry 1996). This sound is transmitted to the chest wall through the ventricles. The S1 sound is a low pitched dull sound that precedes the systole. It is formed mainly by the sounds produced in closure of the Atrioventricular valves (Mitral and Tricuspid) at the start of the ventricular contraction. The Mitral and Tricuspid valves closures may be fused incompletely so that splitting may be audible. The second heart sound S2 is sharper in quality than S1 and can be heard all over the precordium. It is caused by the closure of the semilunar valves (Aortic and Pulmonary). Besides S1 and S2 sounds, heart sound might contain two other sounds, i.e. S3 and S4. The third heart sound (S3) is thought to be originated from the sudden halt in the movement of the ventricle in response to filling in early diastole after the Atrioventricular valves close (Mehta and Khan 2004). It is normally heard in children and young adults. Finally, the fourth heart sound (S4) is caused by the sudden halt of the ventricle in response to filling in presystole as a result of atrial contraction.

Heart murmur is the most common abnormal audible component in heart sound. It has to be decided whether it is pathological or physiological by physicians. Murmur may be observed between two consecutive S1 and S2, or S2 and S1 sounds. There are mainly two types of murmurs: (i) systolic murmurs, which occur between systole and diastole, and (ii) diastolic murmurs, which are located between diastole and systole. Murmurs usually exhibit lower energy when compared to S1 and S2. Actual heart murmur is created by valve damage and may also be associated with fusion of the valve leaflets (Sokolow et al. 1989). On the other hand, the S3 (ven-

tricular gallop) and S4 (atrial gallop) components of heart sounds have shown to be highly specific in predicting heart failure (HF) decompensation (Haji and Movahed 2000; Joshi 1999; Narain et al. 2005) and are highly correlated to the decrease of the ejection fraction (EF) (Mehta and Khan 2004; Narain et al. 2005). The latter may help to evaluate the risk of sudden cardiac death (SCD), since SCD is proportional to the EF decrease and HF severity (Batsford et al. 1995), as well as to monitor recovery after an ischemic event.

Myocardial relaxation and contraction are governed by intracellular recycling of calcium ions. Therefore, the timings of the basic cardiac functions are directly related to the health of the cardiac cells (Oh and Tajik 2003), which determine its ability to suppress blood delivery according to the metabolic requirements of the organs. Other very important aspect is the timings of the left ventricle, since it is this ventricle function that controls control the blood flow in the systemic circulation. Heart sounds can be applied to measure the main systolic time intervals, i.e. the pre-ejection period the iso-volumetric contraction time and the left ventricle ejection time (Carvalho et al. 2009). Heart sounds early in systole (S1 and S4) are sensitive to changes in blood volume in the heart chambers, thus changes in the heart sounds monitored during anesthesia can be used to detect hypovolemia (Fan et al. 1995). The splitting interval between the A2 (aortic) and P2 (pulmonary) components is proportional to the pulmonary artery pressure (PAP) (Popov et al. 2004). An extensive review of applications of heart sounds for clinical diagnosis can be found in (Tavel 1967).

1.1.2. Socioeconomic Impacts

In light of the paragraphs above it is evident that heart sound is an important bio-signal to expose various cardiac functions to support the examination of the human heart. Therefore, a computer aided auscultation (CAA) system for the heart, enables identification of some of the above mentioned CVDs and helps to estimate crucial cardiac measurements. This is a valuable support to clinical decision support that includes adequate methods of heart sound analysis. Heart sound recordings can be imported from the electronic stethoscope and analyzed in order to assess the heart's condition, namely abnormal sounds murmur or third and fourth heart sounds. In addition to this, given the means of transmitting the heart sound recordings patients can be monitored from their home aiming to reduce their frequent visits to the hospital. This methodology can be used in remote health care system that is widely known as tele-health or tele-medicine (Koekemoer and Scheffer 2008; Zenk et al. 2003). Tele-health contributes to bring improvements in the quality of life by offering socio-economic advantages. Socio.-economic aspects are observed in the form of the following indicators: accessibility, social isolation, cost effectiveness,

decreased health service use, education, support, health outcomes, and quality of care. Therefore, from the viewpoint of providing better quality of life in terms of socio-economic to cardiac patients a substantive research on heart sound is justified (Jennett et al. 2002; Stachura and Khashanshina 2004).

1.1.3. Personalized Medical Care (pHealth)

Another scope of the heart sound analysis can be in personal health system (pHealth). This a new trend of health care that is provided to the subjects to tailor diagnosis and therapy including prevention, home care, elderly and life style services (Blobel 2011). In this health care system, information regarding body organs' function is monitored using vital signal's regular analysis of patient at home. The vital signals originating from the respective organs encode information of their functionality, such as heart sound for the heart, breathing sound for the respiratory system, etc. To acquire these signals via an adequate portable device a new concept is adopted which is to embed sensors in the wearable garments (Cook and Song 2009). It enables patients to wear it and register their vital signals while performing formal for daily routine work. With the microsystems, telemedicine, smart implants and sensor-controlled medical devices pathological condition of both inpatients and outpatients can be monitored and treated.

Because of CVDs, being deadly and their connection with ageing they are crucial for which pHealth system can developed in order to support and (Habetha 2006b). Traditionally, the electrocardiogram (ECG) and heart sound auscultation are among the most used signals for CVD diagnosis. Other signals, such as the impedance cardiogram (ICG) as well as the photoplethysmogram (PPG), are less common in daily practice or are still used mainly in research scenarios. ECG and respiration were used to devise a personal system Heart Failure (HF) and other cardiovascular risk stratification (Atoui et al. 2006; Villalba et al. 2007). Heart sounds have the potential to enable the implementation of third generation pHealth systems, i.e., systems that are able to provide continuous feedback to the user and to the clinician. For instance, in (Couceiro et al. 2011) heart sounds are being used to evaluate cardiac output which enables hemodynamic tailoring of the patient through automated medication adaptations.

1.1.4. Application for Training Medical Professionals

The importance of auscultation has been resurrected by digital technology. Medical professionals are less and less interested in practicing auscultation based diagnosis, as suggested by one of the studies done by Dr. Salvatore Mangione of the Alleghe-

ny University of Health Sciences in Philadelphia (Salvatore 1997). He says:

"Many new doctors don't know how to use a stethoscope properly. Finding is based on a study involving more than 450 first-, second- and third-year internal medicine and family practice residents. They were asked to detect 12 different heart problems by listening with a stethoscope to cardiac rhythms played on a high-fidelity tape. On average, they were wrong four out of five times. There was very little or no improvement over the three years of training. And, in addition, there was basically no difference between residents and a group of medical students."

In fact nowadays, traditional auscultation is not considered to be a needful skill to excel when several advanced machines and techniques, such as the aforementioned MRI, Echocardiography, Pulmonary arteriography, etc., are available (Lam et al. 2005). Given the major advantage in terms of cost-savings over these advanced methods, it is stressed that medical practitioners need to renew and strengthen assessments based upon auscultation for proficient patient diagnosis. Therefore, CAA can play a crucial role in training and teaching to medical students with the new methods that include: electronic stethoscope for listening the body sounds as well as its recording, graphical users interface (GUI) for allowing users to interact with the computer, and contents for auscultation instruction and sound analysis techniques. Furthermore, this may also help in training nurses who have relatively more frequent contact with the patients compare to other medical staff (Hou et al. 2008). They pay a large role in evaluating, detection, and monitoring the patient's problems.

The advantage of the medical students' auscultation training using CAA based system can also help to reduce the auscultation proficiency gap between students with distinct skill levels. Cardiac diagnostic errors can be significantly reduced by providing training based on a vast range of heart sounds from various cardiac conditions using CAA system.

1.2. Objectives

There is a vital need for exploring heart sound in order to enhance and promote noninvasive techniques for medical examination. It is important to dissect heart sound in order to find valuable information using techniques of data processing. By keeping in mind the intricacies of producing a useful record of heart sound even with the most advanced stethoscope available in the market, certain goals are set in the direction to perform a thorough analysis of the heart sound. Typically, heart

sound analysis contains several tasks ranging from denoising, sounds localization, to its application for diagnosis and prognosis purposes such as abnormal heart sound identification/classification.

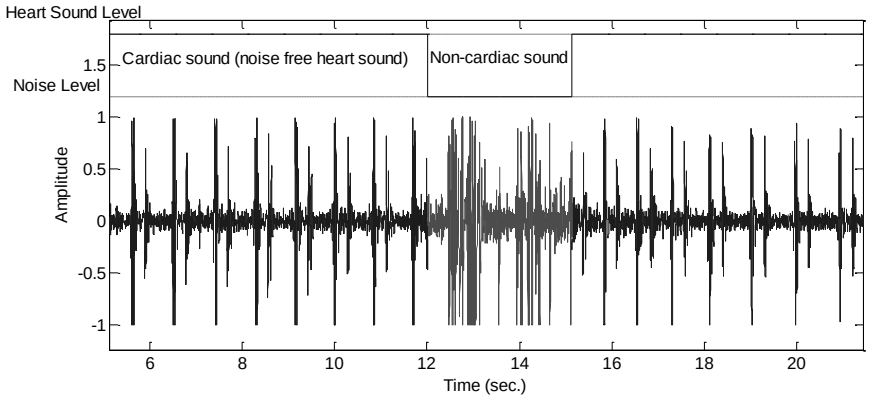


Figure 1.2. An example of a heart sound corrupted with noise, (upper curve) Indication of clean and noisy cardiac sound, (lower curve) cardiac sound with segments of non-cardiac sounds (sound of speech and movement in stethoscope's probe).

Being an acoustic signal, heart sound acquisition is prone to interference with several noise sources that may be difficult to avoid despite being collected in well-controlled room or clinical environments. These noise sources include both external ambient noise sources (e.g. ambient music, noises induced by bystanders, etc) and internal body noise sources (e.g. coughing, physiological noises, etc), which interfere with heart sound in highly complex (e.g. additive or convolutively mixture) and unpredictable ways. As shown in **Figure 1.2** non-cardiac sounds or noises are marked. These noise sources exhibit a very broad range of spectral characteristics overlapping the typical heart sound frequency range and a broad range of durations and loudness, which might alter prominent diagnosis features in heart sounds. In this context, detection of inadequate signal acquisition is of primary importance in order to timely alert and guide the user in solving the signal acquisition problems and to simplify and enhance the reliability of the diagnosis algorithms.

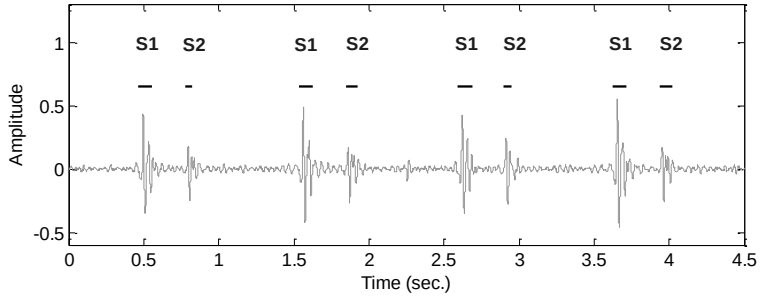


Figure 1.3. An example of segmentation performed on normal heart sound containing mainly S1 and S2 sounds. Duration of the heart sound components are marked by the length of the bar shown above the sounds.

After the noise (i.e. non-cardiac sound) elimination process from the heart sound, usually the next task is the segmentation of heart sound in order to locate its components, i.e. S1, S2, S3, S4, etc. It has been found from experience and literature that in presence of an abnormal sound, which is known as murmur sound, localization of these sounds is quite a difficult task to accomplish. On the contrary, for a normal heart sound segmentation, the process is simpler compared to the former one. In segmentation the following tasks are required to be performed: (i) finding the instants of the onset and offset of the audible sounds in the heart sound records, in other words localizing the sounds by identifying their start and end instants; (ii) find out clinically meaningful sounds; and (iii) recognition of the localized sounds as S1, S2, S3 or S4 sounds. In this thesis the S3 and S4 sounds recognition problem is not addressed. As a result of this, systolic and diastolic region in the heart cycles are also demarcated. In other words, segmentation divides the heart cycles into S1-systole-S2-diastole. Obtaining information of onset and offset instants of the S1 and S2 through the segmentation task enables to further extract valuable properties; these properties are compared with the baseline in course to identify any change in the cardiac status (see the example presented in **Figure 1.3**).

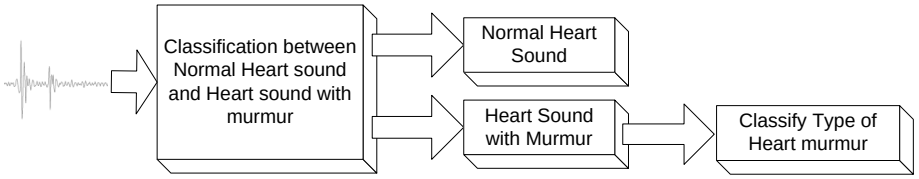


Figure 1.4. Blocks in identification of normal heart sound (heart sound produced from the native valves) and heart murmur (abnormal sound), and further if heart sound is identified as heart murmur then further classification

into the type of murmur.

As the clinical application we address the problem of murmur recognition. In this context, the main purpose of the heart sound analysis is to provide a computer assisted technique to identify abnormal sounds which often are caused by cardiac anomalies. Some cardiac anomalies cause murmur during the systolic or diastolic phase of the in a cardiac cycle (heart beat). Therefore, the next task after the segmentation of the heart sound analysis comprises the following subtasks: (i) classifying normal heart sound and abnormal heat sound; (ii) recognition of murmur category (see **Figure 1.4**).

In summary, the goal of this thesis is to research and develop effective methods that are able to solve the following problems related to automatic diagnosis from heart sound auscultation

- Recognition of uncontaminated cardiac auscultations with clinical diagnosis value.
- Segmentation of heart sounds produced under complex conditions and originated by hearts suffering from different pathologies into their main components, i.e. S1 and S2 sounds as well as systole and diastole components.
- Recognition and categorization of murmur.

1.3. Thesis Outline

In order to maintain coherence throughout the thesis, descriptions of the issues regarding heart sound analysis are presented sequentially and organized in seven chapters. The first three chapters are devoted to establishing the background of the problem. It contains explicit statements regarding the main problems, physiology of the heart and origin of the heart sound, and a summary of methods that have been implemented and applied to develop computer based heart sound analysis to provide a an efficient tool for diagnosis in the past recent years. In remaining four chapters, our proposed methods to solve the main problems, stated in section 1.2, in a scope of heart sound analysis are introduced and analyzed.

Introductory chapters: chapter 1- chapter 3

Chapter 1

This is the current chapter. In here we present the motivation of our work, the

objectives, and the main contributions of the thesis.

Chapter 2

This chapter starts with the brief chronological overview of the auscultation techniques. Large part of the chapter is devoted to an overview of the functionality of heart, and the origins of heart sounds. Besides regular components of heart sounds, the situations under which abnormal sounds are produced are introduced as well. Further, the heart sounds main components' as well as murmurs' characteristics are highlighted.

Chapter 3

A state of the art overview regarding technical solutions to handle the problems outlined in section 1.2 is presented in the chapter. We overview chronologically the main methods developed in the past few years regarding each problem.

Technical chapters: chapter 4 - chapter 7

The main scientific contributions of the thesis are presented in these four chapters. Each of these chapters, excluding chapter 7, starts with a presentation of the problem then a technical overview of the main approaches. In each chapter, the central part is devoted to the presentation of the proposed solution for the problem. A thorough result analysis and discussion is presented at the end of each chapter.

Chapter 4

The algorithm for non-cardiac sound detection is explained in chapter 4. The method explores the periodic nature of the heart sound signal in order to discriminate among contaminated and uncontaminated heart sounds. The method comprises two main steps: the first step is concerned with the detection of an uncontaminated that will serve as a patient/auscultation site specific template; the second step corresponds to a template match technique in order to detect the noisy parts in the heart sound.

Chapter 5

This chapter describes our proposed strategy for heart sound segmentation. The strategy first recognizes heart sounds with and without murmur in order to select the most appropriate segmentation algorithm. The classification is performed based on the degree of chaos of the heart sounds' samples. Chaos is measured as Lyapunov exponents from the phase space constructed from a time series or a dis-

crete signal using non-linear dynamical theory. We present two methods for heart sound segmentation: one is based on the energy envelop; whereas the other one explores wavelet decomposition- nonlinear features in order to segment more complex heart sounds. The last section of the chapter is devoted to performance assessment and discussion.

Chapter 6

Heart murmur classification is tackled in this chapter. We propose several new features to classify murmur. A feature selection strategy is followed coupled to a SVM (support vector machine) in order to identify the set of optimal features. A comparative analysis of results is introduced at the end of the chapter followed by a thorough discussion.

Chapter 7

This chapter outlines the main conclusions of the proposed methods, and some additional challenges under the scope of heart sound analysis are stated as well. Hence, this chapter expands the scope of heart sound analysis by listing a set of tasks that may lead to further research, in the spirit of developing complete heart sound based diagnosis system.

1.4. Main Contributions

In this section we list the main contributions of this thesis in relation to the problems on heart sound analysis. Three different problems have been tackled based on different techniques of signal processing. The first problem was related to noise/non-cardiac sound detection found to be the most challenging one as this problem has not yet been explored thoroughly by researchers. Since our aim is to implement the developed algorithm in real time, it requires an algorithm that demands a less complex and inexpensive computation. The other two problems tackled in this thesis, i.e., segmentation and murmur classification have been investigated extensively in the past years. Our contributions in solving these three problems are about exploring signal processing principals, where some of them have not been used until now in these problems, to develop simple algorithms aiming to reduce the time of execution, improve robustness and provide higher accuracy.

1.4.1. Noise detection

Our major contribution is in non-cardiac sound detection during heart sound acquisition, exploring the periodic or semi-periodic nature of the heart function. Putting forward the argument regarding the fact that heart sound is periodic it has been observed that this periodicity is not violated in clean heart sound (noise-free heart sound) in the temporal domain as well in time-frequency domain, whereas it tends to be violated in the presence noise contaminations, both due to internal body noise sources such as breathing, muscle movements, etc., as well as due to external ambient noise sources. Using these criteria, the first few seconds of heart sound are evaluated to detect a heart beat sound template. This template is applied for the detection of non-cardiac sounds.

1.4.2. Segmentation

We propose heart sound segmentation strategy that explores specialized segmentation algorithms for distinct classes of heart sounds, i.e., heart sounds with and without murmur. In a first stage the method assesses the presence of murmur using a chaoticity analysis of the signal. Based on this analysis the strategy chooses the suitable one between the two specialized algorithms for heart sound segmentation. The first method is based on the envelogram approach introduced by Liang *et al.* (Liang et al. 1997). Herein, some novelty is brought in the second part of the segmentation process which recognizes the relevant sounds. Rather than restricting to systolic or diastolic intervals in recognizing the main heart sound components we investigated physiological properties. We explore high frequency contents in the heart sound component, as they play a kind of signature role in order to determine the second main sounds S2 (Lim et al. 1980).

The second algorithm is based on wavelet decomposition-nonlinear features (WD-NLF). This algorithm was inspired by the wavelet decomposition-fractal dimension based method which was designed to determine explosive sounds (crackles, squawks, and squeeze) and bowel sound (Hadjileontiadis 2005; Hadjileontiadis 2003) in the Lung sound. We explore nonlinear features, namely simplicity (contrary to complexity) to segment the main heart sound components even in the presence of severe heart murmur. The WD-NLF method is highly adaptive in the selection of the thresholds out of the nonlinear features. The approach for S1 and S2 recognition is borrowed from the energy envelop-based method.

1.4.3. Heart Murmur Classification

Features play a key role in training a supervised classifier that may perform classification with high accuracy. We propose a new set of features in order to compute various sound related properties, such as morphology, quality, etc. Features computed in different domains, namely temporal and frequency, and nonlinear features are employed for classification in which many of them have not been used in the past. These features are subjected to feature selection in order to identify potential uncorrelated features, and later to a SVM based classifier.

1.4.4. List of Publications

Scientific Journals

(Paper prepared for submission)

P1. D. Kumar, P. Carvalho, M. Auntunes, J. Henriques, R. P. Paiva, "*Heart murmur detection and classification using physiological features and feature selection* ", in Pattern Recognition Letters. (*To be Submited*)

P2. D. Kumar, P. Carvalho, M. Auntunes, J. Henriques, J. Habetha, "*Heart Sounds Segmentation and Recognition using Wavelet Decomposition and Signal Simplicity*", in IEEE Journal of Biomedical and Health Informatics (J-BHI). (*in Review*)

(Published papers)

P3. D. Kumar, P. Carvalho, M. Auntunes, R. P. Pavia, J. Henriques, "*Noise Detection During Heart Sound Recording Using Periodicity Signatures*", in *Physiol. Meas.* 32 (2011) 599-618, doi: 10.1088/0967-3334/32/5/008

International Conference Proceedings

P4. D. Kumar, P. Carvalho, M. Antunes, R. P. Paiva, J. Henriques, "*An Adaptive Approach to Abnormal Heart Sound Segmentation*", in Proc. Of International Conference on Acoustic and Speech Signal Processing-ICASSP, Prague, 2011.

P5. P. Carvalho, R. P. Paiva, **D. Kumar**, J. Ramos, S. Santos, J. Henriques, "*A Framework for Acoustic Cardiac Signal Analysis*", in Proc. of Bio-inspired System and Signal Processing, BIOSIGNAL 2011, Rome, Italy, 2011, pp. 151-160.

P6. **D. Kumar,** P. Carvalho, M. Antunes, R. P. Paiva, J. Henriques, *"Heart Sound Classification with Feature Selection"*, in Proc. of the 31th Annual International Conference of the IEEE Engineering in Medicine and Biology Society - EMBC, Bunes Aires, Argentina, 2010, pp. 4566-4569.

P7. **D. Kumar,** P. Carvalho, M. Antunes, R. P. Paiva, J. Henriques, *"Heart Murmur Classification using Complexity Signature"*, in Proc. of the International Conference on Pattern Recognition, Istanbul, Turkey, 2010, pp. 2564-2567.

P8. **D. Kumar,** P. Carvalho, M. Antunes, R. P. Paiva, J. Henriques, *"Noise Detection During Heart Sound Recording"*, in Proc. of the 31th Annual International Conference of the IEEE Engineering in Medicine and Biology Society- EMBC, Minneapolis, Minnesota, USA, 2009, pp. 3119-3123. *(in Special Session)*

P9. **D. Kumar,** P. Carvalho, M. Antunes, J. Henriques, A. Sá e Melo, R. Schmidt, J. Habetha, *"Discrimination of Heart Sounds Using Chaos Analysis in Various Subbands"*, in Proc. of Bio-inspired System and Signal Processing, BIOSIGNAL 2009, Porto, Portugal, 2009, pp. 369-375.

P10. **D. Kumar,** P. Carvalho, M. Antunes, J. Henriques, A. Sá e Melo, R. Schmidt, J. Habetha, *"Heart Murmur Recognition and Segmentation by Complexity Signatures"*, in Proc. of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society- EMBC, Vancouver, Canada, August 2008, pp. 2128-2132.

P11. **D. Kumar**, P. Carvalho, M. Antunes, J. Henriques, A. Sá e Melo, R. Schmidt, J. Habetha, *"Third Heart Sound Detection Using Wavelet Transform--Simplicity Filter"*, in Proc. of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology, Lyon, France, 2007, pp. 1277-1281.

P12. **D. Kumar,** P. Carvalho, M. Antunes, J. Henriques, A. Sá e Melo, R. Schmidt, J. Habetha, *"Near Real Time Noise Detection during Heart Sound Acquisition"*, in Proc. of the 15th European Signal Processing Conference, Poznan, Poland, September 2007, pp. 1387-1391.

P13. **D. Kumar,** P. Carvalho, M. Antunes, J. Henriques, A. Sá e Melo, R. Schmidt, J. Habetha, *"Detection of S1 and S2 Heart Sounds by High Frequency Signatures"*, in Proc. of the 28th Annual International Conference of the IEEE Engineering in Medicine and Biology, New York, USA, August 2006, pp. 1410-1416.

These papers are related to the three main contributions of this thesis: papers P3, P8, and P12 related to noise detection; papers P2, P8, and P13 are on heart sound segmentation; papers P1, P6, P7, and P8 concern the classification of heart sounds and murmur. Besides these papers that address three main problems stated in section 1.2, paper P11 was published on the problem of S3 sound detection. S3 sound detection is not included as a main contribution of this thesis because of the

fact that this method is yet to be tested with a large heart sound dataset. Paper P5 is a summary of the overall heart sound analysis framework developed at our research group.

Chapter 2

THE PHYSIOLOGICAL ORIGIN OF HEART SOUND AND MURMURS

Heart sounds are noises generated by the beating of the heart and resultant flow of blood through the valves and circulatory surface[5]. The sound reflects the turbulence created when the heart valves snap shut and can be listened with a stethoscope. The deep understanding of the cardiac structure and the origin of heart sounds, by cardiac auscultation can be used for assessing several mechanical and electrical heart diseases.

This chapter focuses on describing cardiac auscultation and its relashionship with heart's function.

Section 2.1 Introduction

We start with a brief introduction on the main cardiac signal that results from heart's mechanical function, describing the main components of the electrocardiogram (ECG), heart sounds and murmurs.

Section 2.2 Heart's Anatomy and Physiology

In this section, background information on the anatomy and physiology of the human heart is provided in order to understand the production of heart sounds and murmurs. The important structural components of the human heart such as cardiac chambers, vessels and valves are also discussed. The function of the heart and the variation of cardiac parameters during the heart cycle are also introduced in this section.

[5] http://en.wikipedia.org/wiki/Heart (Last viewed on 12 September 2012)

Section 2.23 Heart Sound and Murmur

This section introduces instruments for capturing heart sounds and the correlation between heart sound morphology and auscultation site. The characteristics of cardiac auscultation are related with the site where it is heard on the chest. The conventional theories regarding the generation of heart sounds and its propagation through the chest are explained in this section. This is followed by an introduction of the main heart sound components and murmurs. Important morphological characteristics of the heart sounds used for clinical purposes are highlighted.

Section 2.2 Heart Sound as Phonocardiogram

In this section, we introduce recent methods of auscultation using the electronic stethoscope. An overview on modern age stethoscope technologies is provided.

Section 2.4 The Diagnostic and Prognostic Values of Heart Sounds

The diagnostic and prognostic~~s~~ value~~s~~ of several cardiac and vascular diseases and its use to perform cardiac evaluation are described in this section.

2.1. Introduction

Heart sounds and electrical signals known as PCG and ECG, respectively, are originated due to the heart's mechanical and electrical activty (see **Figure 2.1**). This chapter starts with a brief description of the anatomy of the heart that is complemented with a brief introduction on the origin of heart sounds and the ECG. Since heart sounds are originated as a consequence of the vibration of the heart valves and shear stress of the blood on the surface of the heart vessels, it is important to highlight the physical components of the cardiac muscle that are most active. In some cases of severe valvular diseases, new heart sounds are heard, as valves are sometimes replaced with prosthetic valves that have different quality and pitch.

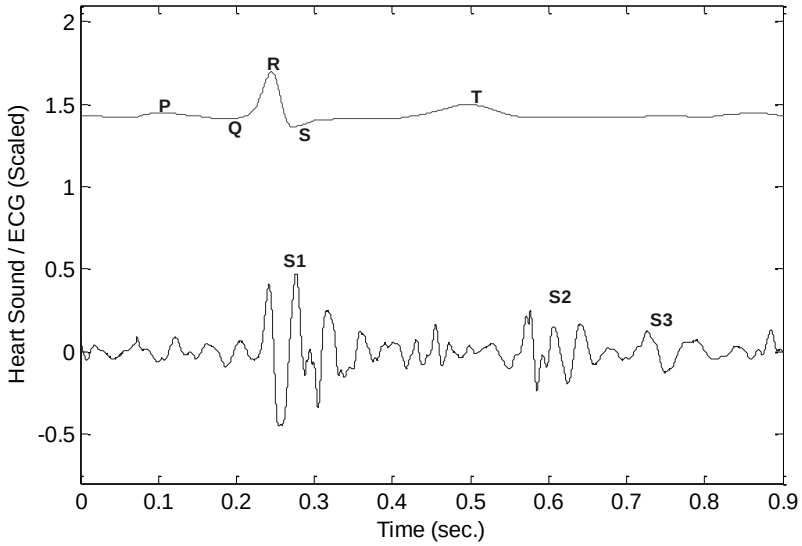


Figure 2.1. (Upper curve) Electrocardiogram ECG, (Lower curve) Heart sound acquired from native heart valves. The heart sound shows three main heart sound components, i.e., S1, S2 and S3. The third heart sound is a pathological sound in the case of adults.

Cardiac auscultation is a technique to listen to heart sound through stethoscope, usually, a stethoscope. This technique has seen consistent advancement over these years. It has reached to a stage that nowadays stethoscope is available in textiles to support patients under the homecare environment (Cho and Asada 2002).

2.2. Heart's Anatomy and Physiology

Heart is a constituent of the cardiovascular system located in the lower thorax, in the middle mediastinum slightly to the left of midsegittal plane[6] and in relation with vessels: the superior and inferior vena cava, the pulmonary artery and vein, and aorta (Williams and Warwick 1989). It weighs about 250 to 300 gram, and its wall is composed of cardiac muscle, also called myocardium; the heart has four compartments/chambers: the right and left atria and right and left ventricles. The heart is oriented such a ways that the anterior aspect is the right ventricle while the posterior aspects shows the left atrium, see **Figure 2.2**. The upper atria compartments

[6] A vertical plane through the midline of the body; divides the body into right and left halves. It is also known as median plane.

are constituted as one unit and the lower ventricles as another. The left ventricular free wall and the septum are much thicker than the right ventricular wall.

The function of the heart is to receive oxygenated blood from the lungs and to supply it to every organ of the body. In order to execute its function, the heart has four chambers and four types of valves. During pumping process, blood is forced through the valves, out of the chambers when the heart contracts, flowing from one chamber to another. (see in **Figure 2.2**).

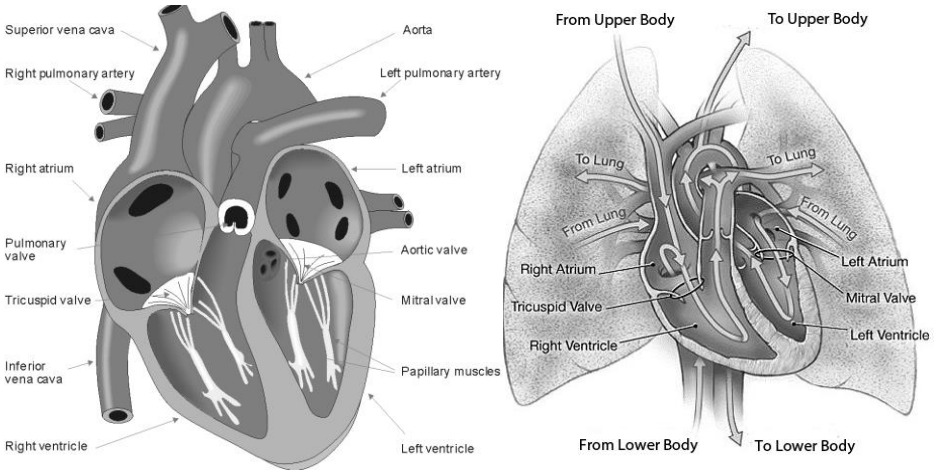


Figure 2.2. (left) Anatomy of the heart and physiology (image is taken from (Malmivno and Plonsey 2009)), (right) blood flow through the chambers of the heart. (Image was taken from (Iaizzo 2005))

The pumping action of the heart is divided into two phases: systole when the ventricles contract and eject blood from the heart, and diastole, when the ventricles are relaxed and the heart is filled with blood. Four valves prevent the blood from flowing backwards; the atrioventricular valves (mitral and tricuspid) prevent blood flowing back from the ventricles to the atria and the semilunar valves (aortic and pulmonary valves) prevent blood from flowing back towards the ventricles once being pumped into the aorta and the pulmonary artery, respectively. Deoxygenated blood of the body enters the right atrium, passes into the right ventricle and is ejected out through the pulmonary artery on its way to the lungs. Oxygenated blood from the lungs re-enters the heart in the left atrium, passes into the left ventricle and is then ejected out to throughout the body.

2.2.1. Heart Valves

The heart valves play a crucial role in the heart's function. As it has been mentioned before and can be seen in **Figure 2.2**, a normal and healthy heart encompasses four valves whose sectional views can be observed in **Figure 2.3** (upper). Mitral and tricuspid, known as atrioventricular valves, have two and three cusps, respectively. While semilunar valves, i.e. aortic and pulmonary, are shaped like a half-moon, and have three cusps, consisting of connective tissue reinforced by fibers. These valves do not have chordae tendineae instead the shape of the cusps prevents any form of prolapse. The atria are separated from the ventricles by the fibrous skeleton of the heart (Ahlstrom 2006a). The mitral and tricuspid valve leaflets are connected via the chordate tenndineae and papillary muscles to the ventricular wall.

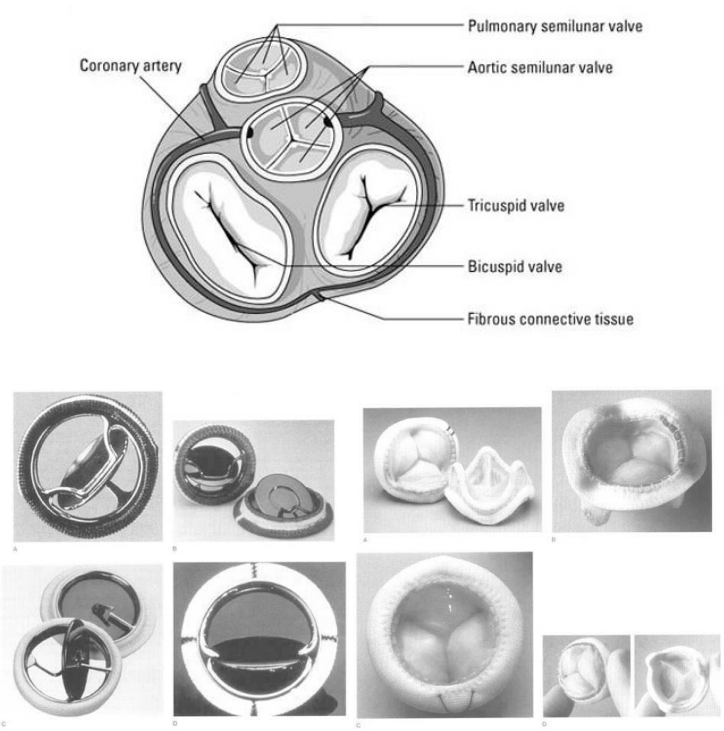


Figure 2.3. (Upper) Native valves; (Lower left) Mechanical valves; (Lower right) Bio-prosthetic valves (figures are from (Yoganathan 2000)).

Another category of heart valves is the artificial one, which is widely addressed as prosthetic valves. These valves are a replacement for faulty native valves. It was stated in section 1.1.1 under Chapter 1 that valvular degenerative dysfunction is mainly prevalent in elderly population throughout the world, especially in develoed countries. In cases of severe valve dysfunction, heart valve replacement is found to

be the only feasible treatment. The use of these valves at such magnitude caused a substantial development in the prosthetic technology since 1952 when the first mechanical valve called caged-ball, was implanted to suppress an aortic insufficiency. Some examples of prosthetic valves can be seen in **Figure 2.3** (Lower). Prosthetic valves are further split into two categories: mechanical valves and tissue based or bioprostehtic valves (see in **Figure 2.3** (lower left and right)). However, since most durable prosthetic valves are manufactured using synthetic materials, these suffer with many problems and complications, such as thrombosis, coagulation, infection, tissue growth, etc (Dayem and Raftery 1967; Pibarot et al. 2009). Regarding bioprosthetic valves, their main dysfunction is due to calcification.

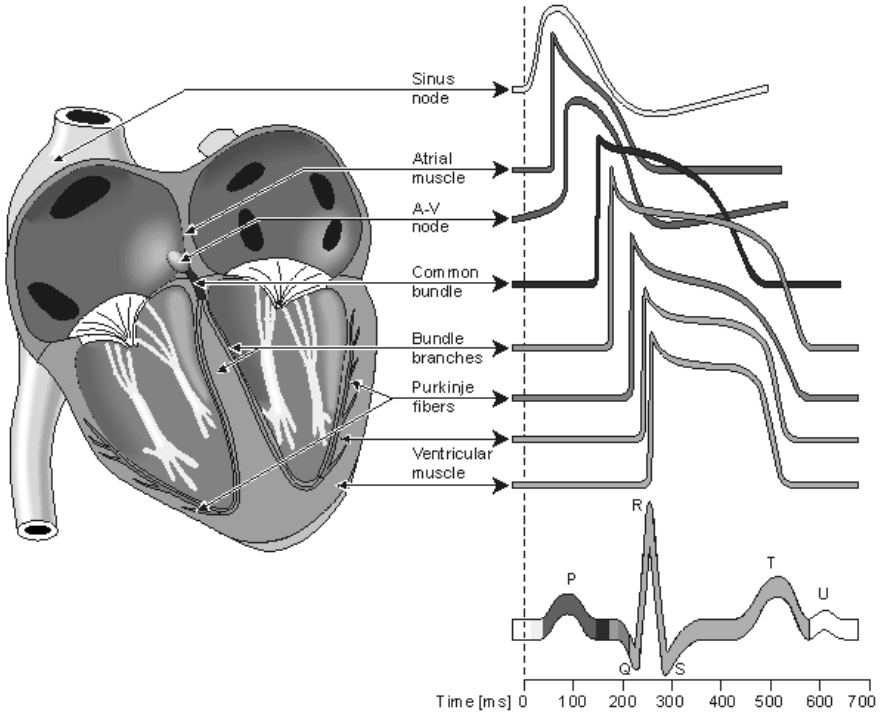


Figure 2.4. (Right) Cardiac electrical conduction system morphology and timing of action potentials from different of the heart. (Left) related ECG signal as measured on the body surface (figure is modified from (Malmivno and Plonsey 2009)).

2.2.2. Conduction System of the Heart

Normally, cardiac muscles possess four properties: automaticity (the ability to initiate an electrical impulse), conductivity (ability to conduct electrical impulses), contractility (the ability to shorten and perform work), and lusitropy (the ability to relax) (Berne and Levy 1997). In the heart muscle cell, or myocyte, electric activation takes place by means of the same mechanism as in the nerve cell, i.e. from the inflow of sodium ions across the cell membrane. It is composed by two events that are cardiac depolarization and thereafter repolarization. Depolarization and repolrization happen along with the cardiac pumping.

Under the normal cardiac electrical conduction, an impulse, or action potential, generated by the sinoatrial node (SA node), spreads through the atria and stimulates the muscles of the atria. This stimulation causes the atrium to contract. The atria are electrically isolated from the ventricles, connected via the atroventricular node (AV node). Due to the AV node, signal is delayed. This delay in the transmission allows the atria to empty before the ventricle contracts. The stimulus is conducted through the bundle of His to the bundle of branches and then to the purkinjie fibers and the endocardigram at the apex of the heart, then finally to the ventricular myocardium (see **Figure 2.4**).

Electrocardiogram (ECG) is the signal where events taking place during electrodynamics are manifested (see in **Figure 2.4**). It shows five, P-QRS-T, major waves in one heart cycle (one contraction and relaxation of the heart). The P wave is the result of an electrical impulse propagating throughout the left and right atria. The P-R time interval is the delay involved with the phase of atrial repolarization that is the AV node function. The QRS complex is produced due to depolarization through the ventricular myocardium and the T-wave is the result of re-polarization of the ventricles.

2.2.3. Cardiac Cycle with Pressure Profile

Under one heart cycle the ventricles and the atria contract and relax in two sequential phases. The phase contraction is known as systole, when the heart pumps out the blood from the ventricular chambers to pulmonary arteries and aorta, and the phase relaxation, also known as diastole, when blood flows from the atria to the ventricles. Concisely, during systole the heart chambers eject blood, and during diastole the heart chambers fill with blood. Two consecutive heart cycles are presented in **Figure 2.5**, where it can be seen that gradients in the pressure curves correspond to the events of heart sounds and ECG that is resulted from the sequential systole and diastole. The constituents of ECG were introduced in the previous subsection while heart sounds' components are described in the next subsection.

Cardiac cycle can be understood with the pressure profile which is associated with the mechanical and electrophysiological changes in the heart. The times involved in realizing contraction and relaxation are addressed as systolic and diastolic period, respectively. Thus, a heart cycle is composed by one systolic and one adjacent diastolic period.

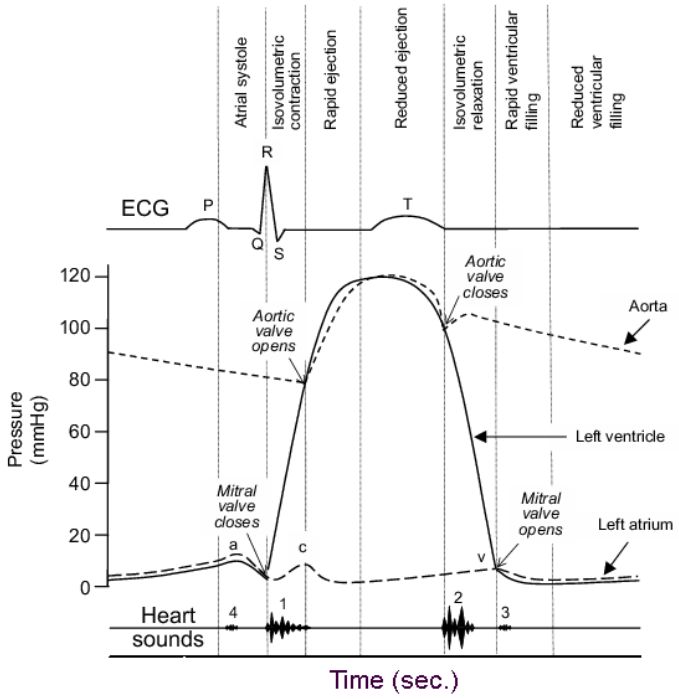


Figure 2.5. Various cardiac events correlated in time in various curves: (from upper to lower) ECG, ventricular, aortic pressure curve, left ventricular pressure curve, left atrial pressure curve, heart sound. (Wiggers diagram was adapted from (Berne and Levy 1997))

Systolic phase

Systolic phase begins after the closing of the mitral valve that is followed by closing of tricuspid valve (see A section in **Figure 2.5**). After this event, systolic period is composed by two stages: (i) the first is the stage of isovolumnic contraction in which the ventricles contract causing a rise of ventricular pressure. This incremental ventricular pressure forces the semilunar (aortic and pulmonary) to open for blood ejection from the ventricular chambers. (ii) The second stage corresponds to the ejection stage that can be subdivided into rapid ejection period and reduced

ejection period; rapid ejection is due to sharp rise in ventricular and aortic pressure leading to a high volume of ejected blood, and in reduced ejection aortic as well as ventricular pressure decline due blood flow from aorta periphery. From the atrial pressure curve it can be noticed that at the onset of the ejection there is a sharp decrease due to stretch of the atria. During the reduced ejection due to blood return to the atria there is a progressive increase in arterial pressure. The onset of the ventricular contraction tends to synchronize with the R peak of the ECG and initial vibration of the S1 heart sound.

Diastolic Phase

Like the systolic phase, the diastolic phase also comprises two stages to complete the relaxation process: (i) the first one is referred as isovolumetric relaxation that lasts from the time of closing moment of the semilunar valves to the opening instant of the atrioventricular valves. It can be noted from the pressure curves that after closing of the aortic valves, the aortic pressure progressively decreases while ventricular pressure falls steeply (by keeping ventricular volume unchanged). (ii) The second stage corresponds to the ventricle filling process which can be subdivided into rapid filling stage and reduced ventricular filling stage. The rapid filling of the ventricles starts as soon the AV valves open, hence allowing blood to flow from the atria to the ventricles. Subsequently, it produces a decrease in the atrial and in the ventricular pressures. While in the reduced filling stage, since blood returns from the periphery and from the lung to the ventricles it causes gradual rise in atrial and ventricle pressure. The onset of aortic valve closing tends to coincide with the terminating point of T-wave in ECG signal and onset of the second heart sound (S2).

2.3. Heart Sound and Auscultation

Auscultation is the technique to listen the body sound via a stethoscope. This sound is later interpreted in terms of acoustic properties, intensity, pitch and quality. Generally speaking, the stethoscope is the acoustic device that enables to capture sound produced by the movements of internal organs. However, we limit our focus to cardiac sounds and its auscultation in this section. In the first three subsections heart sound produced by native valves closure, components of the sounds, heart murmurs and heart sounds in cases of prosthetic valves implant are discussed. Later, a summary of chronological development is presented. Optimal conditions to perform auscultation as well as relating to the auscultation sites will be highlighted.

Acoustic waves are generated by vibrating objects instigated by pressure differ-

ences. The vibration sets particles in motion and if the sound is a pure tone, the individual particles move back and forth with the frequency of that tone. Each part is thus moving around its resting point, but as it pushes nearby particles they are also set in motion and this chain of events results in areas of compression and refraction. The alternating areas of compression and refraction constitute a pressure wave that moves away from the sound source. These pressure variations can be detected via the mechanical effect they exert on a membrane (the diaphragm of a stethoscope, the tympanic membrane in the ear etc.). If the sound source vibrates in a more irregular manner, the resulting sound wave will be more complex. Usually, sound is described by its intensity, duration and frequency. If the sound is non-stationary, these variables have to be analyzed as a function of time to give relevant information.

Frequency is a physical entity, and what humans perceive as frequency is called pitch (unit Mel). The two are closely related, but the relationship is not linear. Up to 1 kHz, the measured frequency and the perceived pitch are fairly the same. Above 1 kHz, a larger increase in frequency is required to create an equal perceived change in pitch. Intensity is determined by the amplitude of the vibration, the distance the wave must travel and the medium through which it travels. Similar to frequency, intensity also has a perceived correspondence, named loudness (unit phon). Intensity and loudness are not related, as it is shown in **Figure 2.6**.

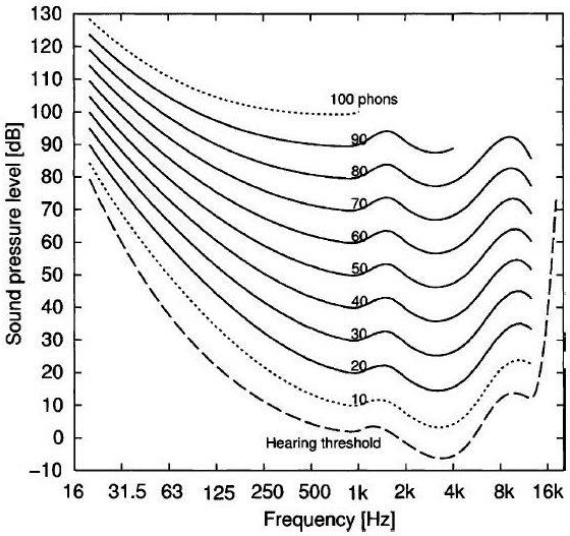


Figure 2.6. Estimated equal loudness level contours derived from Fletcher and Munson. The dashed line shows the threshold of hearing. The contours at 100 phons are drawn by a dotted line because data from only one institute are available at 100 phons. Each curve represents a sound which

is perceived to have equal loudness for all frequencies.

2.3.1. Heart Sounds

One of the basic mechanisms of the production of normal cardiac sound is the sudden acceleration or deceleration of blood, which is influenced mainly by the opening and closing of heart valves and sudden tension of intracardiac structures (such as chordae tendineae, papillary muscles, chambers walls). It results in the emanation of four heart sounds known as S1, S2, S3 and S4 sounds.

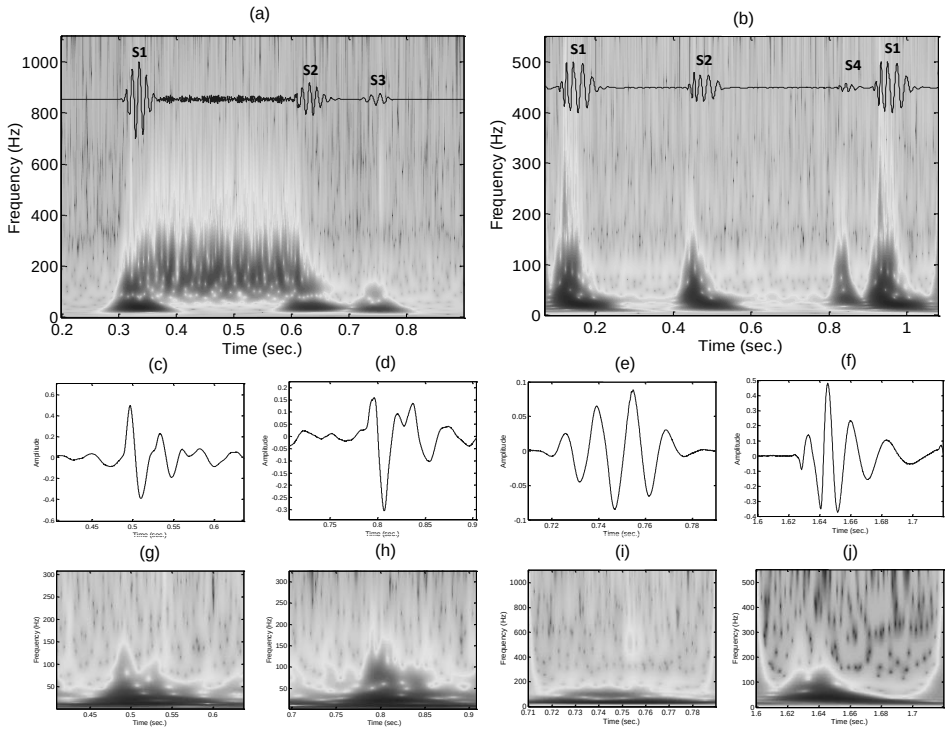


Figure 2.7. Heart sound samples along with their frequency concentration over time in presence of S3 and S4, respectively, are shown in (a) and (b). Figures (c) to (f) are magnified view of typical each sound components, i.e. S1, S2, S3, and S4. Lowest row of figures (g) to (j) are the time-frequency transform of the each components show the frequency distribution over time. In the images of time-frequency representation darker regions are exhibiting higher frequencies. (Color map: Darker space represents larger intensity of corresponding frequencies)

First Heart Sound (S1)

The most accepted theory in the production of S1 sound is the closure of the atrioventricular valves, that is, the tricuspid and mitral valves. There is essentially a change in the rate of pressure rise in the ventricle causing sudden tension of the intracardiac structure that causes the production of the S1 sound. Since the main components of S1 sound are produced by the closure of two valves (other less prominent components include the vibration induced by the opening of the aortic valve (Carvalho et al. 2009)), it is composed by two main sound components. The mitral valve closes just before the tricuspid valve resulting in consequence M1 (produced by mitral valve closure) and T1 (produced by tricuspid valve closure). Gap between the two sounds is approximately 0.02 to 0.03 second (Erickson 2003). M1 is slightly higher in intensity and frequency and requires less energy to produce. The complete S1 sounds produced by native valve closure exhibit frequency in the range of 10-150Hz. A typical S1 sound wave and its frequency contents are displayed in **Figure 2.7** (c,g).

Second Heart Sound (S2)

The second heart sound (S2) **is** a result of the closure of the aortic (A2) and pulmonary (P2) valves. Although some experimental and clinical evidence points to other intracardiac factors contributing to the formation of the S2, the first hypothesis is still the most accepted and the most cited factor is stated in the thesis. Sound A2 has more energy behind its closure and therefore is louder than P2; A2 normally precedes the P2. In normal condition, the splitting interval between these two sounds varies with the respiration cycle; it widens during and narrows during expiration. During the inspiration it is from 0.02 to 0.08 second (mean, 0.03 to 0.04 sec), while during expiration less than 0.015 second (Felner 1990; Shaver et al. 1985). The S2 sounds are higher frequency sounds than S1 sounds with the fact that Atrioventricular valves vibrate more and faster due to the greater pressure difference across them. The frequency range of typical S2 sounds is 10-400 Hz (see in **Figure 2.7**(d,h)).

Third Heart Sound (S3)

The third heart sound is the early diastolic sound also known as ventricular gallop, protodiastolic gallop, S3 gallop, and S3. The word gallop is used for the S3 sound because the sequence of heart sound with S3 or S4 resembles the canter of a horse, particularly when the heart is beating rapidly. The widely accepted theory describing the origin of S3 sound is the ventricular theory, which suggests that it is originated by ventricular compliance related with rapid deceleration of the early transmittal flow and the associated vibration of the entire cardiac-pool system (Joshi 1999; Mehta and Khan 2004). Notwithstanding the uncertainty regarding its

genesis, presence of S3 sound in the heart sound cycle of above 40 years old adults is usually related to a ventricular dysfunction. It may be a sign of ventricular distress, as in congestive heart failure. However, this sound in children is considered to be normal. Men tend to lose it in their 20s and women in their 30s. Sometimes it can also be heard in cases of patients with coronary artery disease, cardiomyopathy, incompetent valves and left-to-right shunt (ventricular septal defects or patent ductus arteriosus) and as a first clinical sign of congestive heart failure (CHF) (Gopal and Karnath 2009; Johnston et al. 2007). These sounds are usually appear after 120-180 ms from the onset of A2 sound and its frequency range is 25-70 Hz (Hult et al. 2004; Hult et al. 2005), as shown in **Figure 2.7** (e and i).

Fourth Heart Sound (S4)

The fourth heart sound is the late diastolic sound which is also addressed as atrial gallop, presystolic gallop, S4 gallop and S4. It is a low frequency sound that occurs just before the S1 sounds. This sound is a result of decreased ventricular compliance or increased volume of filling; in such situations ventricles receive additional blood from the atrium which becomes the main cause for the production of low frequency sound. It is mainly heard in athletic persons under 20 years old due to increased diastolic volume (normal in young individuals) (Collins et al. 2005; Gopal and Karnath 2009)). However, sometimes it may indicate certain abnormalities in the heart in terms of cardiac myocardial associated with decreased ventricular compliance (Gopal and Karnath 2009). Like the third heart sound it is a low frequency (10-50 Hz) heart sound (Johnston et al. 2007). Its frequency range often overlaps with the S1 sounds frequency range which makes it difficult to distinguish. A fourth heart sounds and its frequency range are shown in **Figure 2.7** (f and j).

2.3.2. Abnormal Heart Sound or Heart Murmur

Abnormal heart sounds, or heart murmurs, mainly emerge as a result of a turbulent blood flow. This is caused by the following cases: unilateral protrusion into bloodstream, circumferential narrowing, flow into a distal chamber of larger diameter than the proximal chamber, flow into a distal chamber of smaller diameter than the proximal chamber, and high flow rate (Berne and Levy 1997; Erickson 2003).

Murmurs can be heard either in the systolic phase or in the diastolic phase of the heart cycle. Based on the cardiac condition it may be of short duration pre-systolic, mid-systolic, and late systolic, in the systolic region, or pre-diastolic, mid-diastolic, and late-diastolic in the diastolic region. While another type of murmur that is usually long duration and persists throughout the heart cycles. These murmurs are referred as continuous murmurs that present during systole and the dias-

tole, and also overlap main heart sounds S1 and S2 (Michael and Chizner 2008).

Murmurs are defined by certain sound characteristics: loudness (or intensity), frequency (characterized as pitch in the description of a sound; it may be low, medium or high), quality (blowing, harsh or rough, or rumble), timing (systole and diastole), and shapes of the murmur. Loudness is judged by using six these different categories based on the ability of sound to be heard by normal hearing aid. These categories are defined in terms of grades; it is an I-through-VI grading scale and denoted as grade of the murmur/maximum grade, for instance I/VI or 1/6 (Silverman and Wooley 2008). However, the loudness of the murmur does not correlate with its significance as some insignificant abnormalities may have a loud murmur whereas an abnormality that has progressed in significance may be softer. Frequency or pitch indicates whether the sound is high, medium, or low (see in Table 2.1). Quality is tightly related to the frequencies and described as blowing (high

Pitch	***Frequency (cycle/sec)***	***Quality***	***Abnormality***
High	200-400	Blowing	Mitral regurgitation Tricuspid regurgitation Aortic regurgitation Ventricular septal defects
Medium	100-200	Harsh	Aortic stenosis Pulmonic regurgitation Atrial septal defects
Low	<100	Rumble	Mitral Stenosis

Table 2.1 Pitch and quality of the different kind of cardiac abnormalities. (Table is borrowed from (Erickson 2003))

frequency), harsh (medium frequency) or rough and rumble (low frequency). Timing is related to the position of the murmur, whether it is present in the systolic region (between S1 and S2) or in the diastolic region (between S2 and S1), see **Figure 2.8**. Murmurs are also crucially identified by their shape that is often described by crescendo/decrescendo (diamond-shaped) and holosystolic (Chizner 2008). A diamond shaped murmur starts softly, peaks in intensity, and then diminishes. A crescendo murmur starts softly and progresses to high intensity (for instance, late systolic

murmur of mitral valve prolapse). On the other hand, decrescendo begins with louder and progressively fades until the end (for instance, diastolic murmur of aortic regurgitation).. A holosystolic murmur remains unchanged throughout its length in systole or diastole (for instance, pansystolic murmur of severe mitral regurgitation) (Erickson B. 2003).

Based on the position in the heart cycle where the murmur occurs it is divided into three main classes: systolic murmur, diastolic murmurs, and a murmur that is sustained in both phases is a continuous murmur.

Categories of Murmurs	***Subcategories of Murmurs***	***Clinical Significance***
Systolic murmur	Early-mid	Innocent systolic murmur, aortic scierosis, aortic stenosis, pulmonic stenosis, hypertrophic obstructive, cardiomyopathy
	Holosystolic	Mitral regurgitation, tricuspid regurgitation, ventricular septal defect
	Late systolic	Mitral valve prolapsed, papillary muscle dysfunction, calcified mitral annulus
Diastolic Murmur	*Early (high frequency)*	Aortic regurgitation, pulmonic regugitation
	Mid-late(low frequency)	Mitral stenosis, flow rumble mitral regurgitation, aortic regurgitation
Continuous Murmur		Patent ductus arteriosus, jugular venous hum, arteriovenous fistula

Table 2.2 Clincal clues to heart disease from cardiac auscultation (this table is formed based on the given information in (Chizner 2008)).

Systolic Murmurs

Systolic murmurs are most commonly found between sounds S1 and S2. These can be innocent or pathologic. Low grade (<III/VI) early and mid-systolic mur-

murs are considered to be an innocent murmur, but in the presence of other related cardiac

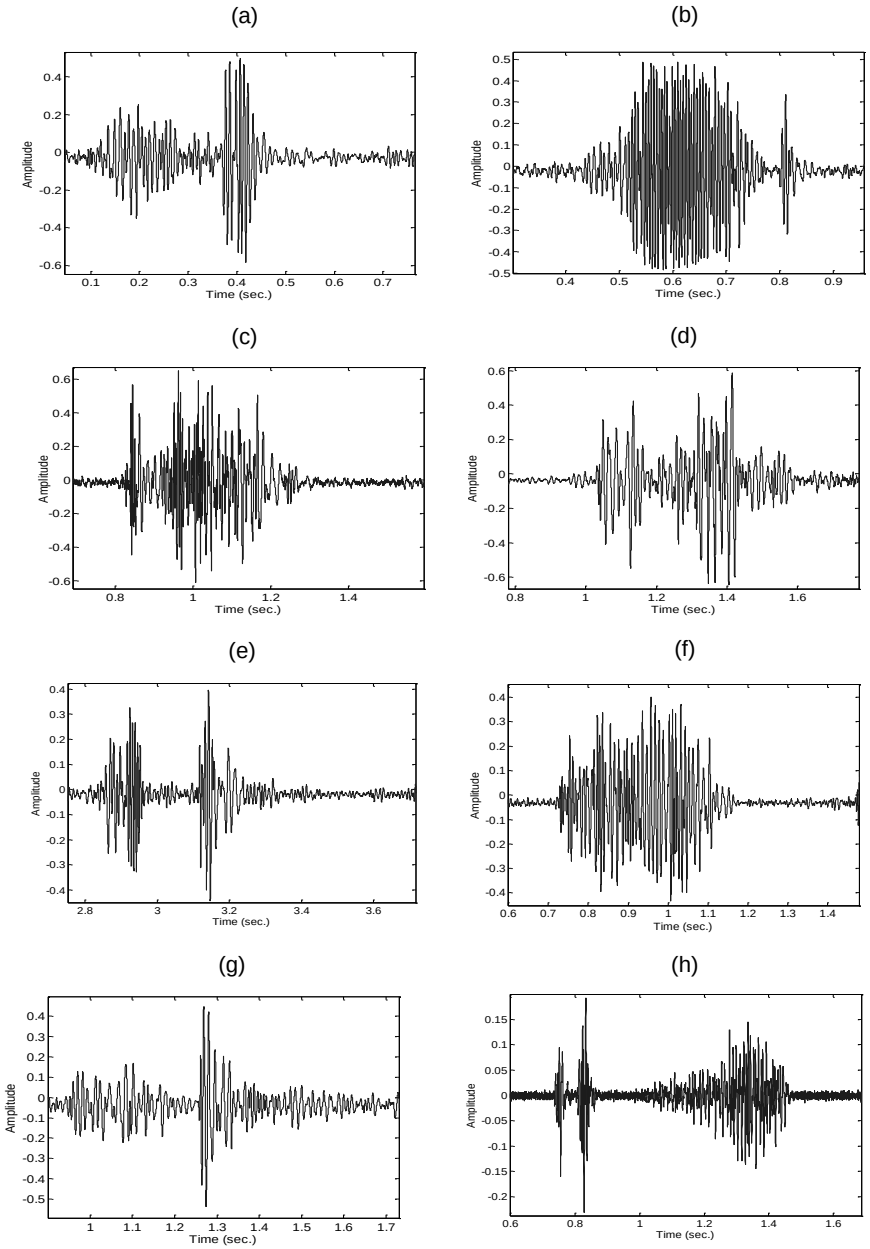


Figure 2.8. Heart murmurs (one heart cycle demonstration), systolic in (a)-(f) and diastolic in (g)-(h). Systolic murmurs are: Aortic Regurgitation in (a), Aortic Stenosis (b), Mitral regurgitation (c), Pulmonary Stenosis (d), Systolic Ejection (e), and

Ventricular Septal Defect in (f). And the diastolic murmurs are: Tricuspid Regurgitation in (g) and Mitral Stenosis in (h).

.

anomalies they can be pathologic. Loud systolic murmurs (grade >= III/VI) are more likely to be hemodynamically significant although the degree of loudness is not directly related to the severity of the anomaly. The length of a murmur is often more indicative of its severity (Chizner 2008; Erickson 2003). A murmur is considered to be a pathologic systolic murmur or innocent murmur based upon the clinical finding associated with it rather than on the basis of its characteristics. Systolic murmurs are divided into three categories based on the characteristics mentioned above in this section and mentioned morphologies as Table 2.2 reveals. The murmurs produced during the ventricular period are due to the forward flow across the aortic or pulmonary valve or the regurgitation flow from the mitral or tricuspid valve. Some prominent cardiac problems in which systolic murmurs are found to be present during auscultation are usually mitral regurgitation, aortic stenosis, and ventricular septal defect.

Diastolic Murmur

The diastolic murmur begins with or after S2 sounds or before S1 sounds. All diastolic murmurs are pathologic. Three main causes are known for diastolic murmur: aortic or pulmonary valve incompetence, mitral stenosis or tricuspid stenosis, and increased blood flow across mitral or tricuspid valves. Common pathologies related to diastolic murmurs are listed in Table 2.2.

Continuous Murmur

These murmurs are present throughout the heart cycles. It begins around the end of the S1 sound and extends in diastole. They are generated by continuous flow from a chamber with high pressure into a chamber with low pressure. The known causes of continuous murmurs are the following: obstruction in blood flow between the aorta and pulmonary arteries (Patent ductus arteriosus), abnormal connection between arteries and veins (arteriovenous fistula), abnormal flow in veins or arteries (venous hums or renal artery stenosis). Some continuous murmurs are listed in Table 2.2.

2.3.3. Prosthetic Valve Sounds

As it was mentioned in Section 2.2.1 prosthetic valves are replacements of faulty valves independently of the type of the valve. Mechanical and biological are two types of artificial valves that are usually transplanted based on the individual's characteristics. The sounds generated by prosthetic valves depend on specific details: types of valve, area of its placement and the condition health of the valve (weather the valve is functioning normally).

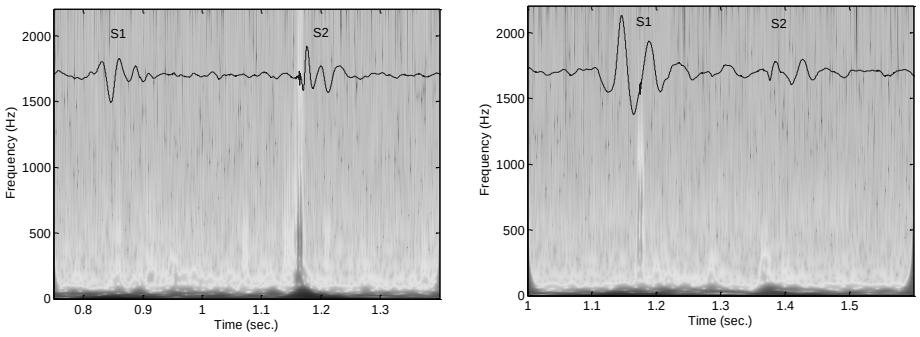


Figure 2.9. Left figure is the time-frequency representation of a heart sound in mechanical valve (single tilted disk) at aortic position; and the right figure is time-frequency plot of the heart sound from biological valve at mitral position.

Mechanical valve sound

Three commonly known categories of prosthetic heart valves are caged-ball, tilting-disk, and bileaflet. Unlike native valves they produce perceivable sound when the valves open and close rather than generating sound generation it by valves' vibration. The sounds produced by these valves closing event reveal different characteristics depending on the architecture given to the mechanical valves as shown in **Figure 2.3** (lower). For instance, the caged ball consists of a ball which moves freely within a three or four sided metallic cage mounted on a swing ring and produces short duration, high frequency, and high amplitude sound (Dayem and Raftery 1967). Its opening click sound is louder than the closing click sound which is produced just after the first heart sound if it is in mitral position or the second heart sound in aortic position (Yoganathan 2000) (see left figure in **Figure 2.9**). It can also be noticed in the figure that the closing sounds of these valves are usually very high frequency sounds compare to natural valves of bioprosthetic valves' sounds. The second category valves, known as tilting-disk valves have free–floating and lens-shaped disks mounted on a circular ring with a tilt ranging from 60 to 80 de-

grees (depending on material of the valve). They generate loud and greater frequency sounds in either implant positions (mainly aortic and mitral). In contrast to the caged ball valve, they produce a louder closing click sound with an almost inaudible opening click sound. The third category of mechanical valve consists of a bileaflet disk which is composed by two semicircular leaflets mounted on a circular sewing ring that opens in the center. Like the single tilt disk it produces a loud clicking sound when the valves close (Yoganathan 2000). Most mechanical valves suffer from a common problem of thrombosis which interfere with their free movement (Bettadapur et al. 2002). Thrombosis is the formation of a blood clot that causes immobilization of the disks and obstruction.

Biological or Bioprosthetic Valve

These valves are also known as tissue valves. They are divided into three categories based on the used materials applied in their manufacturing: the allograft type are obtained from the fascia lata of the pulmonary valve, the homologous type from a corpse or mater, and the hererologous type from the bovine aortic valves, or bovine pericardium (Bettadapur et al. 2002; Erickson B. 2003). The sounds produced by the closure of these valves exhibit characteristics similar to the sounds of natural valve; for instance, main frequency contents of the sounds from porcine tissue prosthesis are between 25 Hz to 200 Hz which are also found in several healthy natural valves (Bedi 1994; Puvimanasinghe 2004). The right figure in **Figure 2.9** demonstrates frequency spread over time in bioprosthetic heart sound.

2.4. Cardiac Auscultation and its Diagnostic and Prognostic Values

The analysis of the heart sound recorded as PCG is a basis for the preparation of valuable information for diagnosis and prognosis of cardiac related pathologies. In this section we present some major cardiac diseases and the physiological parameters mainly related to cardiac function where heart sound analysis is useful for diagnosis and/or prognosis. Each heart sound component, i.e. S1, S2, S3, S4, or murmur is prominent in order to determine certain cardiac diseases and physiological parameters. For instance, S1 and S2 sound intensity (or energy) and time between their main components are seen to diagnose several cardiac disorders (Tavel 1978). Accentuated intensity of S1 sounds are observed in cases of mitral stenosis and S2 in several cases such as dilation in aorta or pulmonary artery. Split time between the components S1 sounds is rare however S2 sound split is seen in cases of hypertension. Some examples of heart diseases diagnosed using cardiac auscultation are

given below.

2.4.1. Cardiac Diseases

There are several cardiovascular diseases that can be diagnosed with the help of heart sound. We highlight here some of the main diseases in which heart sounds can be a potential tool in their diagnosis or prognosis.

Prosthetic Valve Dysfunction

The S1 and S2 sounds are important in the early detection of prosthetic valve malfunctioning, which mainly occurs due to thrombosis, in mechanical valves and tissue degeneration, in biological valves. In the first case, long-term anticoagulant therapy is the key to reduce the thrombotic risk (Nair et al. 2003). Thrombosis in mechanical valves is mainly formed due to sheer stress created by their design and material under the influence of alteration in fluid mechanics. There is interaction between the blood plasma components and the surface of the prosthetic valves thereby thrombin enzyme converts solvable fibrinogen[7] into insolvable fibrin which leads to adhesion of platelets forming blood clot or thrombus (Furie and Furie 2008; Pibarot et al. 2009). Mechanical valves are more durable than bio-prosthetic valves therefore they are preferred valves to implant in young and middle age patients. Elderly patients may find inconveniences with the thrombolytic therapy therefore bio-prosthetic valves are preferred in the aged population (Nair et al. 2003; Pibarot et al. 2009; Vongpatanasin et al. 2003).

Heart sounds play very important role in identifying the degeneration of prosthetic valves besides other sophisticated noninvasive methods such as echocardiography. Spectral analysis of heart sound's main components S1 and S2 depending upon the implantation position was adopted to determine degradation in the functioning of valves, i.e., mechanical valves as well as bio-prosthetic valves (Baykal et al. 1995; Chen et al. 1997; Durand et al. 1990; Fritzsche et al. 2005; Kim et al. 1998; Koymen et al. 1987; Sava and McDonnell 1996; Stein et al. 1981).

Native Valve Diseases

The murmurs between S1 and S2 sounds or S2 and S1 sounds often become prominent sign of heart valve related problems. Valvular defects are usually seen be-

[7] Fibrinogen is soluble glycoprotein synthesized by liver and is part of blood coagulation.

tween systolic and diastolic regions of the heart sounds. If there is any anomaly in the opening or closing of the native valves extraneous sounds are produced during contraction and relaxation of the heart as explained in Section 2.3.2. There are two general types of cardiac valve defects: stenosis and/or insufficiency. Valvular stenosis results from a narrowing of the valve orifice that is usually caused by a thickening and increased rigidity of the valve leaflets, often accompanied by calcification. When this occurs, the valve does not open completely as blood flows from the one cardiac chamber to the other. Valvular insufficiency results from the valve leaflets not completely sealing when the valve is closed so that regurgitation of blood occurs (backward flow of blood) into the cardiac chamber.

Heart Failure

Presence of the third heart sound, S3, in adults over 40 years can be a sign heart failure (see Section 2.3.1). Congestive heart failure (CHF), also known as heart failure (HF), is the inability of the heart to pump a sufficient amount of blood to the organs or the ability to do so only with an increased filling pressure. In CHF patients the efficiency of the myocardium or heart muscle is reduced due to several conditions such as myocardial infarction (in which the heart muscle is starved of oxygen and dies), hypertension (which increases the force of contraction needed to pump blood), etc. Some of the recent studies on the quantification of the diagnostic values of S3 sound on CHF suggest the presence of S3 in acute decompensated CHF patients were as high as 93% accuracy (Collins et al. 2006; Johnston et al. 2007).

Coronary Artery Diseases

In coronary artery disease, arteriosclerotic[8] processes narrow the coronary arteries (that supply blood flow to the heart), and thereby restricting blood flow and adequate oxygenation of the myocardium. When the oxygen supply is insufficient to meet the oxygen demand the myocardium becomes ischemic, which may lead to infraction (tissue damage). The second heart sound is observed to determine features associated with the turbulent blood flow in partial occluded artery. In fact, coronary blood flow is maximum during diastole and hence sounds starting from the onset of the S2 sounds in the diastolic regions are analyzed (Akay et al. 1990; Akay et al. 1993; Semmlow et al. 1990).

Cardiomyopathy

It is a disease of the heart muscle. In most cases, cardiomyopathy causes the heart muscle to become weak as a result of this the heart becomes unable to pump efficiently. This is also one of main causes of congestive heart failure. Therefore, as

[8] Arteriosclerotic is a chronic disease in which arterial walls thicken, harden and loose elasticity. It results occlusion that becomes cause of impaired blood circulation.

in the HF cases the third heart sound (S3) is a sign of the cardiomyopathy. Furthermore, the fourth heart sound (S4) is also associated with the cardiomyopathy (Hosenpud and Greenberg 2006).

Pulmonary Hypertension

Pulmonary hypertension is caused by high blood pressure. A wide split between the two components of S2 sound, i.e. A2 and P2, leads to the presence of pulmonary hypertension. In younger subjects inspiratory splitting averages 0.04 to 0.05 second during quiet respiration. If the second sound is split by greater than 0.04 second on expiration, it is usually abnormal. Usually in normal condition, this split varies widely split into A2 and P2 sounds which can be from 30 ms to 80 ms by varies intensity of A2 of P2 from each other's (Shaver et al. 1985). Pulmonary artery pressure (PAP) can be estimated from the splitting interval of the S2 sounds (Popov et al. 2004; Sutton et al. 1968; Xu et al. 2000; Xu 2002).

2.4.2. Physiological Measurements

Besides the diagnostics and prognostics potential of the heart sounds in some major cardiac diseases as was explained in Section 2.4.1 several physiological measurements to access the cardiac status, can be performed using it in order to assess the cardiac function. Given the other measurements curves, for instance pressure profiles of the atrials and ventricles and ECG, several parameters related with cardiac outputs can be estimated with the help of heart sounds (see in **Figure 2.10**). Any alteration in these parameters from a base level (that is defined in normal cardiac functioning) is an indication of disease, for instance, hypertension, heart failure, valvular dysfunction, etc. These changes can be measured as trends over time. Some of the parameters are briefly described here.

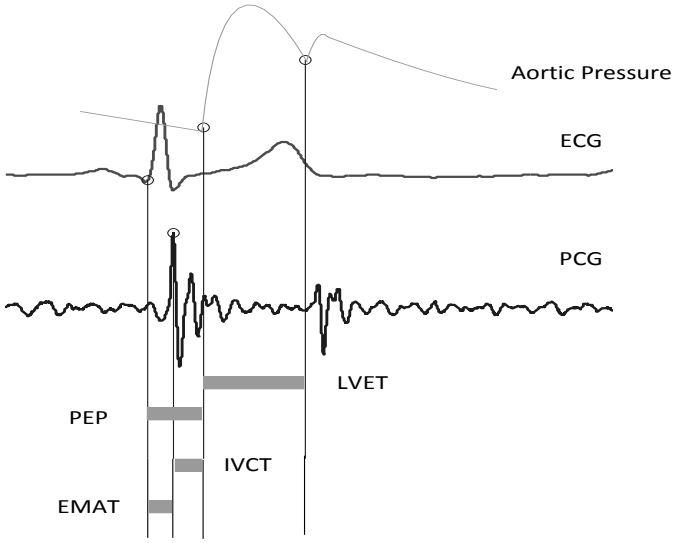


Figure 2.10. The main cardiac intervals in the process of cardiac functioning (figure is regenerated from (Ahlstorm 2008)).

Systolic Time Interval: Pre-ejection Period (PEP) and Left Ventricular Ejection Time (LVET)

Pre-ejection period is defined by the time interval between the R-peak of the ECG and the start of the opening of the aortic valve. It is composed by the two intervals: electromechanical activity time (EMAT) and isovolumic contraction time (IVCT), as **Figure 2.10** illustrates. In a heartbeat EMAT is measured from the Q-wave of the ECG to the electromechanical activity required for the left ventricle to close the mitral valve (beginning of the S1 sound). And IVCT is computed as the difference between S1 and upstroke in artiral pressure. As it can be seen in **Figure 2.10** that PEP in actual comprises EMAT and IVCT. A recent method proposed by in-house team for PEP estimation using the heart sounds can be seen in (Paiva et al. 2009b)(Paiva et al. 2009b).

PEP is defined by the time interval between the R peak of the ECG and the start of the opening of the aortic valve. Similarly, the left ventricular ejection time is defined as the time span between the start of opening and the start of the closing of the aortic valve. Using heart sound, PEP is evaluated as the time span between the R peak of the ECG and the beginning of the aortic valve sound component of the first heart sound (S1). Left ventricular ejection time (LVET) is defined as the time span between the start of opening (upstroke of the aortic pressure curve) and the

start of the closing (incisura of the same pressure curve) of the aortic valve. Using heart sound, LVET is evaluated between the beginning of the aortic valve sound component of the first heart sound (S1) and the onset of the second heart sound (S2).

Stroke Volume (SV)

It is an important cardiac parameter that measures the volume of blood pumped form one ventricle of the heart beat by beat. Ventricular SV is the difference between the ventricular end-diastolic volume (EDV) and the ventricular end-systolic volume (ESV), i.e. SV= EDV-ESV. SV is estimated using its relation with the other cardiac parameters such heart rate (HR), PEP and LVET as described in (Couceiro et al. 2011; Rietzschel et al. 2001).

Cardiac Output (CO)

The cardiac output is defined as the volume of blood ejected by the heart in a certain period of time. It is determined as product of ventricular stroke volume (SV) and heart rate (HR), i.e. CO=SV × HR, where SV can be determined from the above mentioned method (Couceiro et al. 2011; Rietzschel et al. 2001).

2.5. Conclusions

This chapter introduced some background information on the cardiac structure, the origin of heart sounds, their characteristics and their diagnostic and prognostic value. The latter provide the motivation for using heart sound in many medical applications and are the main motivation for this thesis. At the start of this chapter, the anatomy and physiology of heart sound was explained in which the structure of the heart and its function, i.e. pumping blood to the body, were described. Then, the origin of and the important characteristics of the main heart sounds' constituents, i.e. S1, S2, S3 and S4, were highlighted. Heart sounds produced by prosthetic heart valves were also described and analyzed. Abnormal heart sounds which are originated because of some cardiovascular anomalies, namely in valvular diseases, have been explained. The devices used to capture heart sounds in order to perform computer based automatic analysis are sensor based stethoscopes. Hence, a brief introduction of the major events in the development of stethoscope over the years has also been provided. Finally, some of the main diagnostic and prognostic applications of heart sounds in cardiology are reviewed.

Chapter 3

HEART SOUND ANALYSIS: LITERATURE REVIEW

Analysis of the heart sounds to obtain valuable information to diagnose CVDs and cardiac condition monitoring are performed using signal processing techniques and various pattern recognition methods. In the context of this thesis we are mainly interested in techniques operations related to noise contamination detection of heart sound auscultation, heart sound segmentation and recognition of its main components, i.e., the S1 and S2 sounds, and murmur recognition and classification. The review of the state of the art of these topics will be main goal of this chapter.

Signal or time series analysis has been widely dealt with the principles and the computation methods of applied mathematics, in which techniques of signal processing and statistics are mainly employed. Most analysis techniques using frequency or time-frequency transformation assume that heart sound signals are Gaussian and non-stationary. On the other hand, high order statistics and nonlinear dynamics based methods presume that heart sound signals are non-Gaussian and non-stationary. Therefore, in this chapter applied methods with both assumptions will be highlighted.

Section 3.2 Noise Detection in Heart Sound

This section reviews methods involved in detecting noise in the heart sounds. The literature is scares with respect to specific methods that target this specific problem. Two categories of approaches are reviewed: single and multi sound-channel based algorithms.The former might be further categorized based on the requirement for an additional synchronization signal, such as the ECG.

Section 3.2 Heart Sound Segmentation

We introduce recent and widely used approaches for the fundamental problem related to heart sound segmentation into its main constituent parts. This is a very rich topic in literature. Indeed, there is large number of methods available in literature that exhibit significant results and indicate diagnostic usage. As was proposed for noise detection methods, we propose to categorize existing heart sound segmenation approaches into two main categories based the requirement on an auxiliary signal and without.

Section 3.2 Cardiac Murmur Classification

This section of the chapter reviews current approaches for cardiac murmur classification. Most of the reported approaches in literature rely on the classical strategy for pattern recognion composed by a feature extraction phase followed by a classification phase. The former is the most challenging one, since it is responsible to capture the main discriminative dimensions in data and, hence, it is where most of the approaches diverge. Different sets of features for different ranges of cardiac murmurs have been suggested. We present some selected methods based on their impact in heart sound research literature.

Section 3.2 Limitations

Finally in the last section the major shortcomings of the existing methods and the main challenges in developing algorithms for the heart sound analysis are described.

3.1. Noise Detection in Heart Sound

In heart sound analysis the first and foremost task is finding a suitable mechanism to diminish intereference of ambient and other undesirable sound during heart sound recording. Sensor based stethoscopes are prone to capture all types of acoustic. To the best of our knowledge, currently no commercial stethoscope exists that is capable of dealing with noise during auscultation. Possible contamination noise sources in heart sound auscultation exhibit a broad range of spectral characteristics overlapping the typical heart sound frequency range as well as duration and loudness, and, therefore, might alter prominent diagnosis features of heart sounds. In this context, the detection of an inadequate signal acquisition is of primary importance in order to send an alert in real time and guide the user in solving the signal acquisition problems and to simplify and enhance the reliability of the diagnosis algorithms. For instance, in tele-health applications, this has a direct and significant

impact on system usability, user acceptance, long-term adherence and safety. Other non-negligible consequences are on the impact of the number of interventions by professionals induced by high rates of false positive diagnosis (Cleland et al. 2005), typically associated with tele-monitoring systems. These might interfere in ordinary clinical workflows and have significant financial and motivational impacts.

There are two main strategies that can be followed in dealing with noise contamintation during heart sound auscultation: (i) one might detect the positions where non-cardiac sounds are superimposed with heart sounds and subsequently eliminate those segments of corrupt heart sound; alternatively, (ii) one might try to suppress or cancel noise from the complete heart sound. Many algorithms exist of the later type (Hadjileontiadis and Panas 1997; Hall et al. 2000; Pourazad et al. 2003; Tang et al. 2010; Várady 2001; Warbhe et al. 2010). Usually, these rely on several filtering techniques including adaptive noise cancellation methods as well as blind source separation methods. Regarding the former, these have not been widely adopted. Suppressing or cancelling noise is involved with greater complexity than identifying noisy segments and removing then from the recording (Carvalho et al. 2005). Therefore, for pHealth applications where energy efficient algorithms are required noise detection may be a suitable way to deal with noise in heart sound auscultation considering that, usually, it suffices to alert the user or the bystander in order to eliminate the contamination sources if they persist.

The approaches in previous works on noise detection during the process of acquiring heart sound using computer assisted electronic stethoscopes can be categorized into two classes of algorithms, i.e, algorithms using an auxiliary reference signal and methods that do not require any auxiliary signal. In the former class, an auxiliary signal is recorded along with the heart sound. This signal is choosen based on its ability to capture the main cardiac phases as well as on its lower complexity to analyse. In this context, ECG is the most widely applied bio-signal to extract information on the length of heart cycles and the locations of S1 and S2 sound components (Carvalho et al. 2005; Paul et al. 2006; Wang et al. 2007). In the later class of methods specific characteristic information of the cardiac cycle is extracted from the heart sound recording.

3.1.1. Multi-channel Signal Approaches

Within this group of approaches of tackling noise detection in heart sound recording, two types of signals are evidenced that have been used along with the heart sound in order to ease noise elimination process. These signals are the ECG which is generally acquired using gel-based electrodes, and additional sound signals (namely to capture environment sounds).

The electrocardiogram (ECG) was used as a reference or marker to find noise contaminations in heart sounds by Carvalho *et al* (Carvalho et al. 2005). The basic premise of this method is the correlation of the power spectral density (PSD) of the sound of each heart cycle with respect to a reference heart sound's PSD. Three main steps were adopted in the whole noise identification procedure: (i) heart cycle identification, (ii) reference sound identification, and correlation between the PSD of the reference sound and remaining heart sound cycles. ECG is used to determine heart cycles (or heart beats) by identifying consecutive Q components of the ECG. Then, from the first 10 heart cycles a reference sound is searched by correlating PSD of each heart cycle to the PSD of the rest of heart cycles. A heart cycle whose average PSD correlation is largest among the 10 heart cycles is selected as the reference heart cycle. This method is effective in determining noises such as cough, speech, stethoscope movement, and high frequency environmental sounds.

Speech recognition methods served as inspiration therefore has been borrowed for heart sound analysis solutions. A very well-known method for speech enhancement based upon a spectral domain minimum-mean squared error (MMSE) estimation was applied to reduce noise effects in heart sound (Paul et al. 2006). Though under this method a noise spectrum is estimated using a "decision-direct" approach that does not require a separate reference signal, ECG is applied to prevent some heart sound components from being eliminated along with noisy sounds. This method reduces white noise from heart sound, while S3 and S4 sounds were prevented using ECG gating. It assumes that the clean heart sound $x(t)$ is addtiviely mixed with noises $n(t)$ such that; $y(t) = x(t) + n(t)$, where singal $y(t)$ is the signal obtained from the sensor. In order to extract the clean heart sound signal $x(t)$, noise is estimated in the spectral domain (magnitude of the Fourier transform) using (3.1).

$$X(f) = G(f)Y(f) \tag{3.1}$$

In (3.1) $G(f)$ is refered as Wiener gain which is defined as in (3.2).

$$G(f) = \frac{SNR_{priori}}{1+SNR_{priori}} \tag{3.2}$$

In (3.2) SNR_{priori} is a priori signal-to-noise ratio that is computed based on the clean signal spectrum $X(f)$ and noise spectrum $N(f)$. It uses decision-direct method to estimated a priori SNR (Ephraim and Malah 1984). The method includes updating the priori SNR using the estimated signal amplitude from the previous computational frame and the current posteriori SNR. Its details can be found in (Ephraim and Malah 1984). It is understood from the previous chapter that S3 and S4 heart sounds contain lower frequency and loudness compared to the two main sounds S1 and S2 hence there was possibility of elimintaing these in the noise elimintation

process. Hence, ECG was also used to determine S3 and S4 sound's location by identifying QRS-complex and T-wave and searching spectral peaks in the heart sound aligned with these locations in order to identify these events.

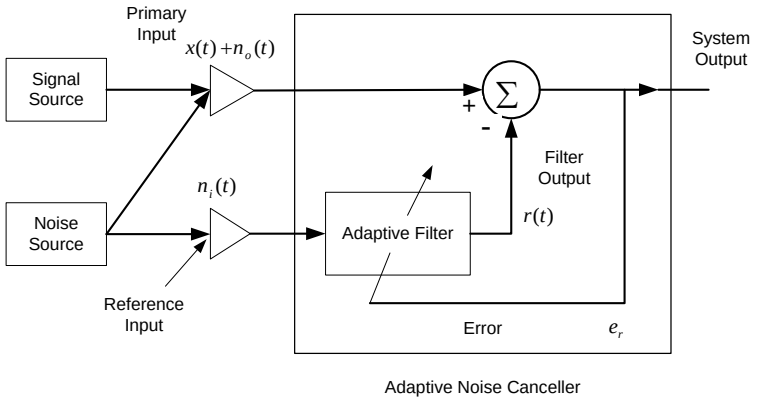


Figure 3.1. Steps required in adaptive noise cancellation process in multi-channel signal approach.

Another group of methods includes the use of an additional sensor to capture ambient noise along with the heart sound. The idea in this type of methods is to use this additional source of information to subtract ambient noise, e.g. by spectral subtraction, in order to achieve an uncontaminated heart sound. The assumption is that the stethoscope and the microphone capture environmental noises. Therefore, samples or spectral subtraction of environmental sounds from the noisy heart sound may result in noise cancellation. This approach is often used in real-time based on which follows the processing steps demonstrated in **Figure 3.1**. Two signals are collected; a heart sound signal $x(t)$ which is addtivitely mixed with $n_o(t)$, and the second signal is a reference signal which records the noise signal $n_i(t)$. Noise signals $n_o(t)$ and $n_i(t)$ are assumed to be correlated. According to **Figure 3.1** ANC receives the reference noise signal $n_i(t)$, filters it, then subtract the results from the original signal $x(t)$. The output e_r of the cancellation process is the error as defined in (3.3).

$$e_r = x(t) + n_o(t) - r(t) \tag{3.3}$$

The above equation is obtained by assuming $x(t)$ is stationary and uncorrelated with noise signal $n_o(t)$. Equation (3.3) is squared and expectation of both sides yields (3.4).

$$E(e_r{}^2) = E(x(t)^2) + E((n_o(t) - r(t))^2) \tag{3.4}$$

From (3.4) it can be deduced that the expected value of the power signal $x(t)^2$ is not affected in the process of minimizing $e_r{}^2$. Hence, in order to minimize the expected value by adjustment in filtering weights the power signal $e_r{}^2$ to obtain noise cancelled original signal $x(t)$ the signal $(n_o(t) - r(t))^2$ is minimized, as in (3.5).

$$\min E(e_r{}^2) = E(x(t)^2) + \min\ E((n_o(t) - r(t))^2) \tag{3.5}$$

The filter output $r(t)$ is then the best least squares estimate of the noise $n_o(t)$. Adjusting the adaptive filter to minimize the total output power results in the best estimate signal e_r of the input signal $x(t)$ for a given structure and adjustability of the adaptive filter as well as for a given reference signal. To obtain the signal $x(t)$ and noise reference signal, the traditional stethoscope was integrated with two microphones: the first one is attached with the stethoscope to acquire heart sound added with environmental sounds, and the second one is to acquire only environmental sound. Subsequently, noisy sounds are subtracted from the corrupted heart sound by employing the described ANC technique. This approach of noise cancellation was also adopted in (Bai and Lu 2005; Belloni et al. 2010; Giustina et al. 2011; Várady 2001).

3.1.2. Single Channel Signal Approaches Approaches

This class of methods was devised to deal with single channel recorded heart sound. In these approaches, physiological facts of the heart and therefore acoustic properties of the heart sounds contribute to significant insights to develop suitable techniques.

Typically, noise reduction in heart sound is treated as a de-noising problem whereby noise is suppressed in the recordings using suitable filtering operation. A simple filtering techniques such as low pass filtering and band pass filtering was applied in (Liang et al. 1997) and (Hurtig-Wennlöf et al. 2009), respectively. For instance, Liang et al. (Liang et al. 1997) used an eighth order Chebyshev type I low pass filter with cutoff frequency at 882Hz to filter out irrelevant sounds, whereas fifth order Butterworth band pass filter with cut-off frequencies of 5 Hz and 250 Hz was applied in (Hurtig-Wennlöf et al. 2009). The problem with these types of simple approaches is related to the fact that they are not able to cope with noise sources whose frequency range includes the pass bands of the filters. Furthermore, it implies a linear additive model of noise interference.

A simple method of samples averaging, also known as ensemble averaging, in the time domain was used in (Sörnmo and Laguna 2005). It assumes that the patterns in the signal are triggered by a measureable events and hence are devided the

signal into segments. The maximum likelihood estimator was obtained by averging samples of the derived segments. It also assumes that noise is additive, independent and Gaussian. Ensemble averaging keeps shapes of the signal unchanged, and the significant events in the signal remain evident and aligned (Sörnmo and Laguna 2005). A detail of the method includes the assumption that the pattern $y_i(t)$ is additively composed of the true pattern $x(t)$ and some noise $n_i(t)$, then $y_i(t) = x(t) + n_i(t)$. Then estimated signal can be determined as in (3.6).

$$\hat{x}(t) = \frac{1}{M}\sum_{i=1}^{M} y_i(t) \tag{3.6}$$

where M is the number of samples of each segment, and $\hat{x}(t)$ represents the estimated true signal. However, $\hat{x}(t)$ can also be calculated recursively according to (3.7).

$$\hat{x}_M(t) = \hat{t}_M(t)\frac{1}{M}(y_i(t) - \hat{x}_{M-1}(t)) \tag{3.7}$$

In (3.7), 1/M controls the updates of new information in ensemble averaging recursively. However, it can be controlled by replacing it by a weighting constant that acts like a forgetting constant, i.e., takes higher values for more recent samples and lower values for less recent ones (Gustfsson 2000).

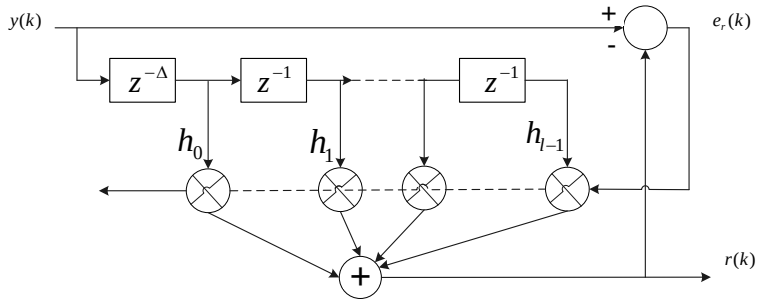


Figure 3.2. Taped delay line adaptive line enhancement (ALE) filter model.

Adaptive line enhancement (ALE) approach from the adaptive filtering techniques was utilized in (Tinati et al. 1996). ALE is configured as transversal a tapped-delay filtering block in which signal is correlated with the delayed signal in order to diminished uncorrelated noise mixed in the signal. Usually, in almost every adaptive noise cancellation method two signals are required of whom one must be a reference signal. However, in ALE the delayed signal is used as the reference signal. Therefore, according to **Figure 3.2** and equation (3.5) if the input signal is represented in discrete time *'k'* then (3.5) can be demonstrated as in (3.8).

$$\min E\left(\{{e_r}^2(k)\}\right) = E(n^2(k)) + \min\ E\{[x(k) - r(k)]^2\} \tag{3.8}$$

In (3.8), noise free heart sound signal $x(k)$ and noise signal $n(k)$ compose input signal $y(k)$ (see **Figure 3.2**), and $r(k)$ is the ouput signal. The delay parameter Δ is set long enough to decorrelate the input noise $n(k)$. The best least square estimate of the clean heart sound signal $x(k)$ can be obtained by minimizing quadratic part in (3.8).

In (Barschdorff et al. 1995), the noise detection process starts by dividing systole period of the signal into two segments, containing the first and the second heart sound, respectively. In each segment the variance is computed. All systoles with variance in either segment above a threshold value are discarded as being noisy. The threshold value used depends on the smallest variance calculated across all systoles. The method described in (Rajan et al. 1998) detects noise by passing the wavelet coefficients of half second windows of audio to single layer perceptrons that have been previously trained on clean and noisy data.

The wavelet decomposition based denoising method is an effective noise reduction approach that is used for various one dimensional as well as multidimensional signals (Strang and Ngugen 1996). In particular, when the signal is transient and its amplitude is higher than the noise amplitude (more likely to be stationary) this method is seen to be performing significantly. The basic steps are to successive decomposition of the signal in reducing frequency bands until a suitable depth is reached, and then to threshold it, and to reconstruct it back in order to obtain a noise free signal. Using this approach, it is almost possible to preserve abrupt events and remove high frequency noises. This approach has been applied for reducing noise in heart sound by (Boutouyrie et al. 1995; Xiu-min and Gui-tao 2009) and in more recent years by (Misal and Sinha 2012).

Heart sound is the natural physiological contaminating signal during auscultation of respiratory or lung sounds. Hence, separating these two sources is a challenge in order to carry out least erroneous lung sound analysis. A reasonable number of methods are available for this purpose. An adaptive filtering method based on the fourth order statistics of the recorded signal (lung sound with extraneous heart sound) was introduced in (Hadjileontiadis and Panas 1997). Reference signal for the adaptive filtering was generated using localized reference estimation algorithm of heart sound tracking, whereas the fourth order cumulant was computed in the adaptive filtering part. Methods inspired by adaptive filtering were also used with some modifications in various works related to heart sounds' separation from lung sounds (Gnitecki and Moussavi 2007; Khan and Vijayakumar 2010; Pourazad et al. 2003; Tsalaile 2008; Yip and Zhang 2001). Gnitecki and Moussavi (Gnitecki and Moussavi 2007) investigated 15 adpative filtering techniques of them 10 were linear adative

filters that require a reference signal and follow the similar computation methods as explained in 3.1.1 (see **Figure 3.1** for the filter model). The others are time-frequency filtering techniques. Linear adaptive filters assume that the noisy signals are stationary, whereas time-frequency methods assume non-stationarity of signals and, hence, due to the quasi-stationarity nature of heart sounds, it is observed that linear adaptive filters exhibit better results in heart sound cancelleation and in presevering lung sounds. It is observed that time-frequency based methods, such as the short-time Fourier transform and the wavelet transform, require much less computational time than linear adaptive filters based on least mean square (LMS) filtering, recursive least square (RLS), reduced order Kalman filtering (ROKF), block fast transversal filtering (BFTF) and liner prediction using estimating auto-regressive (AR) and moving averageing (MA) (Gnitecki and Moussavi 2007). Khan and Vijayakumar (Khan and Vijayakumar 2010) introduced multiscale product of wavelet coeficients based filtering in an attempt to find out a trade-off between loss of information and computation time in linear adaptive filtering and time- frequency based adptive filtering methods. In this approach, first heart sounds are segmented into their main components in order to remove the wavelet coeficients at every level between these main components and then estimates the gaps by using a set of linear prediction filters (Khan and Vijayakumar 2010). Pourazad et al. (Pourazad et al. 2003) introduced a novel approach of identifying heart sound in lung sound using image processing of the spectrogram of the signal to remove the stationary high frequency components. The removed samples were interpolated in time-frequency domain and finally brought denoised lung sound back to time domain using the inverse short-Fourier transform.

One might look inspiration for solving the noise detection and cancelation problem in the well-established speech recognition literature. In this area of research, it is observed that noise cancelation is typically tackled using one of two main approaches: exploring the intrinsic characteristics of vocal sound and applying blind source separation (BSS) techniques. The standard formulation of blind source separation (e.g. using ICA—independent component analysis—or the phases between different signals) requires at least as many sensors as signal sources (Comon 1992; Kristjansson et al. 2004). Extensions exist on BSS for speaker separation, which are not bounded by this independent number of observations constraint. These solutions typically rely on constraints inspired on the characteristics of voice, such as the regular harmonic structure of voiced phonemes (Zhang and Zhang 2005) *a priori* knowledge on sound models (Jang and Lee 2003; Kristjansson et al. 2004; Potamitis and Ozerov 2008), or by assuming that fundamental frequencies do not overlap (Barry et al. 2005).

The basic premise of the BSS is to observe the mixture of signals originated from independent sources and find an appropriate computation method to recover

the original signals. This process assumes that source signals are instantaneously mixed which is modeled by a linear relation between the observation $\boldsymbol{y}$ and the sources $\boldsymbol{s}$ as in (3.9).

$$\mathbf{y} = \mathbf{As},\ \ \mathbf{y} \in \mathbf{R^l}, \mathbf{s} \in \mathbf{R^m},\ \textbf{and}\ \ \mathbf{A} \in \mathbf{R^{l \times m}} \tag{3.9}$$

In (3.9) l is the number of observations and m is the number of sources and **A** is the mixture matrix. It also assumed here that the components of **s** statistically independent and non-Gaussian except at most of one component. Then, separate signal are achieved by estimating separation matrix in (3.10).

$$\mathbf{WA} = \mathbf{PD},\ \ \mathbf{W} \in \mathbf{R^{m \times l}}, \mathbf{P} \in \mathbf{R^{m \times m}},\ \textbf{and}\ \ \mathbf{A} \in \mathbf{R^{m \times m}} \tag{3.10}$$

In (3.10), **W** is the separation matilrx, **P** is a permutation matrix and **D** is a diagonal matrix. Using the sepration matrix output independent signals $\boldsymbol{r}$ are recovered according to (3.11).

$$\boldsymbol{r} = \mathbf{Wy} \tag{3.11}$$

Separation matrix $\boldsymbol{W}$ is usually estimated iteratively by optimizing some cost function of the output signal $\boldsymbol{r}$. Typical algorithms for this purpose apply high order statistics, non-linearity of neurons in a neural network, and mutual information amoung the output $\boldsymbol{x}$ components (Abbas and Bassam 2009).

Typlically, the BSS method requires multi source signals which can be inperprated as the method for multi-channel signal analysis; however, it is used in heart sound segmentation assuming that its components are produced by several cardiac sources (e.g., valves, cardiac structure, and other cardiac activites), and are independent of each other over time. Some application where this concept has been utilized to separate signal from noise, unwanted lung sound, or foetal sounds can be found in (Jimenez-Gonzalez and James 2008; Nigam and Priemer 2004; Nigam and Priemer 2005b; Nigam and Roland 2008; Tsalaile et al. 2008). In (Jimenez-Gonzalez and James 2008), foetal heart sound was extracted from noisy single channels phonograms[9]. In this approach, first the phonogram was projected into a higher dimension space using the method of delay (discussed in Chater 5, Section 5.2) and later multiple independent components were calculated by estimating separation matrix using method to search for directions of maximum kurtosis. At last, the components were projected back onto the measurement space and those associated to foetal heart sound were selected based on the signal morphology and spectra. The method

[9] Phonography is passive long term monitoring technique. It is reordcded through the varioustransducer placed on the maternal abdomen. The abdominal phonogram, carries valuable information about physiological parameters such as heart sounds, heart rate, and foetal breathing/movements.

proposed in (Nigam and Priemer 2004; Nigam and Roland 2008) utilized mutual information between the output signals as the cost function to obtain the separate signals heart sound components as well as background noise. While high orders statistics based on approximate joint diagnolization algorithm was used by exploiting quasi-periodic nature of the heart sound to extract it from the lung sound in (Tsalaile et al. 2008).

3.2. Heart Sound Segmentation

Segmentation is the next main challenge in heart sound analysis after the elimination of noisy portions in the heart sound under processing. Segmentation consists of two subtasks: delimitation, under this the onset and offset of the sounds are identified; and recognition of the sound components, i.e. systolic, diastolic sounds and murmur sounds are identified. Like in noise detection approaches, segmentation of the sound components in the heart sound recording can be decomposed into two main categories of approaches. In the category of segmenting with an auxiliary signal, a signal which requires lesser processing steps is usually acquired to extract a marker coinciding with the main heart sounds. Approaches without an auxiliary signal follows the strategy drawn in the **Figure 3.3** whereby at least one feature profile is generated that accentuates the locations of the main heart sounds using the sliding window based computation technique. Next this feature profile is thresholded to obtain the onset and offsets of the sounds.

3.2.1. Approaches with an Auxiliary Signal

As can be observed in **Figure 1.1** heart murmurs are complementary to S1 and S2 sounds. Hence, if the S1 and S2 sounds are rightly demarcated then murmurs can also be identified using appropriate processing. It must be observed that heart sounds containing murmur in systolic or diastolic phase are relatively require more processing to segment compare to the normal heart sounds due to murmurs' complexity. Therefore, in such cases using a reference signal to capture the main heart events facilitates segmentation. Most of the multichannel signals based segmentation algorithms focus more on detection of the main heart sounds rather than finding the their durations.

One of the earliest uses of ECG for computer assisted heart sound analysis was evidenced in (Wilton-Davies 1972). In (El-Segaier et al. 2005), the S1 and the S2 heart sounds are detected using the ECG signal alone. More specifically, the S1 sound is identified by searching intensity above to a defined threshold in the inter-

val 0.05-0-2R-R (time interval of R-R peak, i.e. one heart beat), and the S2 sound is detected after the T-wave within the interval of 0.6R-R. Systolic murmurs were detected between S1 and S2 by identifying the point with the highest sound intensity. To confirm it to be a murmur, this point's frequency, location, and frequency range were determined. In (Carvalho et al. 2005) a similar approach is described. In this method the limits of the S1 and the S2 sounds are determined based on self-similarity analysis based on the variance fractal dimension. The actual S1 and S2 segments are identified from a poll of candidates using the R-peak and T-wave gating strategy.

One of the first known methods using more than 2 channel signals was introduced in (Lehner and Rangayyan 1987). It used three channel signals to capture ECG, carotid pulse and heart sounds for the segmentation. The onset of the S1 sound was identified by the R-peak of the ECG and the onset of the S2 sound is identified from the timing of the Dicrotic notch in the carotid pulse. In (Barschdorff et al. 2001) a multi-sensor signals was introduced in which five synchronized signals were measured, i.e., ECG, earpulse, ultrasonic blood flow, and Doppler along with the heart sound. Later, systolic and diastolic intervals (complementary to S1 and S2 sounds' duration) are identified using information acquired from the earpulse or either of the signals.

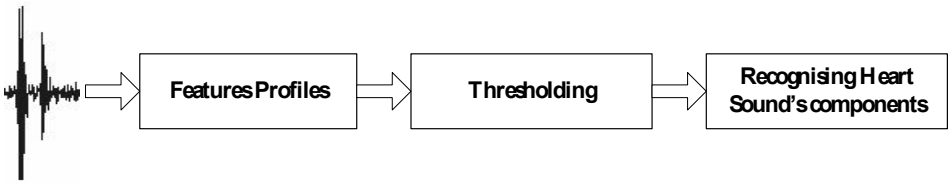


Figure 3.3. Steps required in segmentation for the sounds' components.

3.2.2. Approaches with Only Heart Sound Signal

An additional synchronized signal with the heart sound adds hardware complexity and also has a significant impact on the applicability and user friendlyness of use of the system. This has been a key motivation for the introduction of S1 and S2 sound identification schemes based on heart sound alone. Therefore, in order to avoid relying on the ECG's support in the hart sound segmentation to save hardware cost as well as processing time, there are methods that have been developed to tackle heart sound solo. In fact, this class of approaches can be further divided into unsupervised approaches and supervised approaches. The former one follows similar steps as described in **Figure 3.3**, whereas later one based on classical pattern classification techniques (feature etraction from the sound components then use a supervised

learning based classifier to identify its start and end instants).

A) Unsupervised Approaches

Segmentation using the processing steps described in the **Figure 3.3** implies adopting a feature extraction procedure that quantifies characteristics of the heart sound (such as time and frequency) components after the necessary preprocessing (background noise removal) of heart sounds. Furthermore, it must exhibit significant difference between heart sound components and extraneous sounds, which simplifies application of thresholding on the feature profile in order to note onsets and offsets of the sounds and, hence enabling their delimitation. Subsequently, the main heart sounds are identified using periodic/quasi-peridoc nature of heart sounds as well as other known physiological properties such as, heart sounds' frequency range, intentsity, etc.

Following the processing steps described in **Figure 3.3**, in (Liang et al. 1997) an envelogram based approach was introduced based on the computation of Shannon energy to construct a feature profile of the heart sound due its ability to emphasize medium frequency sounds such as S1 and S2 (see equation (5.11) for Shannon energy). An adaptive threshold based on average normalized Shannon energy of the heart sound recording is applied to the energy profile in order to identify the main peaks related to S1 and S2. In order to recognize the S1 and S2 sounds. First, the largest interval between two consecutive peaks is identified as the diastolic interval (T_{dia}) based on the fact that usually diastolic intervals are larger than systolic intervals (T_{sys}). In order to find the other S1 and S2 sounds the algorithm proceeds with the analysis to the left and the right of the largest interval, two parameters c1 and c2 are used that represent tolerance in systolic and diastolic intervals, respectively (see **Figure 3.4**). Based on this observation, the distinction between these two peaks is performed. This method of segmenting heart sound was also employed in (Altunkaya et al. 2010; Nazeran 2007; Wang et al. 2005). In (Wang et al. 2005) the Shannon energy was computed directly from the heart sound signal and followed the similar steps as in the envelogram based method whereas in (Nazeran 2007) this energy was computed from the wavelet decomposed, 6-level detail wavelet coefficients (frequency bands 62.5 – 125 Hz), signal assuming that at this level effects of external noises are minimum.

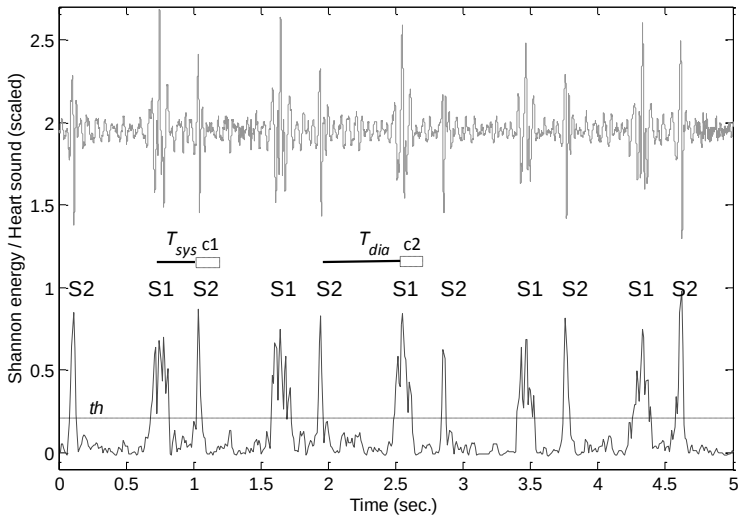


Figure 3.4. Segmentation of a heart sound using Liang's envelogram based method. The parameters c1 and c2 are the tolerance levels which are used to indentify diastolic and systolic intervals.

In another similar approach Shannon and Renyi entropy were computed for the heart sound localization in the lung sound (Yadollahi et al. 2005). In order to estimate the probability density distribution required for assessing the entropy, a kernel-based estimation approach was followed. The computational methods of Shannon and entropies are defined in (3.12) and (3.13), respectively.

$$H(p) = -\sum_{i=1}^{N} p_i \log p_i \tag{3.12}$$

where p_i describes the series of probability of the samples in a heart are sound, where $i = 1 \ldots N$, N is an integer for the total number of the events.

$$H_{\alpha}^{Re}(p) = \frac{1}{1-\alpha}\sum_{i=1}^{N} {p_i}^{\alpha} \tag{3.13}$$

In (3.13), α is a real number different from 1. Events with high probabilities exhibit significant effect on Renyi entropy when $\alpha > 1$ whereas the prominent role is transferred to low propbablities events when $\alpha < 1$. For different values of α entropy is computed and thresholded to localize heart sound components in the presence of lung sounds. After thresholding the entropy profiles, Shannon entropy demonstrated localization of the heart sound with less error (Yadollahi et al. 2005) than Renyi entropy. Shannon entropy was found to be more robust to changes in lung sound characteristics whereas the Renyi entropy based method was more sensitive to noise. In another method, heart sound's complexity was computed based on

the different fractal methods (Hasfjord 2004). The results of these show different degree of complexity of the main heart sound components, murmur, and background noise. The author investigated four methods for fractal computation: continuous box counting methods, information, variance, and Hegushi methods. The details of the computational steps and practical considerations of these methods are beyond the scope of this thesis (the reader is refered to (Hasfjord 2004) for the details). The complexity curve based on the variance and Hegushi were superior in discriminating S1 and S2 sounds as well as murmur. However, they were found to be insufficient in delimiting S1, S2 and murmur when murmurs are observed to be superimposed on S1 and S2.

As the extractions of heart sound features in temporal domain demonstrate limitations in segmenting murmur from the S1/S2 sound components, there were methods investigated in the literatures where heart sounds are transformed into other analysis domains, such as the frequency and the time-frequency domains. For instance, Fourier transform was used to extract frequency peaks in (Yoganathan et al. 1976). In (Haghighi-Mood and Torry 1995), a power spectrum based subband energy tracking method was introduced which assumes power peaks in power spectrum for S1, S2, systolic and diastolic murmurs are different. This method was also able to segment S1 and S2 in presence of certain systolic and diastolic murmur.

Temporal information of a signal corresponding to frequency intensity in power spectrum is absent which, therefore, has led to use of time-frequency based approach to extract delimitation of the sounds. For instance, in (Liang et al. 1998), a time-frequency analysis is introduced based on the spectrogram (see equation (4.5) in Chapter 4). The segmentation of S1 and S2 sounds was achieved by quantifying two high energy values (spectral energy more than 60% of a complete heart cycle) in the spectrogram, which was obtained by retaining intensity above a threshold. Besides the advantage of temporal information, this method achieved better segmentation results because of local thoresholding pertaining to each cardiac component as compare to the most methods which used selection of the sounds' envelope based on average value of the whole recording period. The short-time Fourier transform (STFT) based time-frequency analysis approach unable to exhibit high resolution frequency contents due to due to fix length of window function. Usually, non-stationary signals are comprised with several transient events which may not be demonstrated in spectrogram. Thus, wavelet transform (WT) was employed to overcome this shortcoming and to capture the transient part, namley onsets and offsets of the heart sound components, to perform heart sounds segmentation in (Boutana et al. 2007; Corona and Torry 1998; Gretzinger 1996; Kam 2003). Wavelet transform uses shifted and scaled window function in the process of signal tranoform, as shown in (3.14).

$$CWT(c,d) = \frac{1}{\sqrt{c}} \int \psi\left(\frac{t-d}{c}\right) x(t)dt \tag{3.14}$$

In (3.14), d and c are shift scale parameters, respectively, and ψ is a window function (mother wavelet). In the discrete time version of (3.14) $CWT(c,d)$ is a $l \times n$ matrix of complex coefficients, where l is the number of scales c utilised in the transform and n is the number of samples in the signal. Scaled and shifted versions of the mother wavelet are generated for different values of c and d respectively. There are different types of mother wavelets; usually one is selected based on its similarity with the waveform of the signal (Strang and Ngugen 1996). In (Corona and Torry 1998), the Morlet wavelet was applied to obtain the time-frequency representation for every heart cycle in a heart sound signal. Then, for the delimitation of the sound components the maximum value of $CWT(c,d)$ was determined in the matrix. Values below 40% values were eliminated (background noise), since they are considered as background noise. Subsequently, the remaining values of $CWT(c,d)$ sounds in the matrix were identified as the main sound components. Kam (Kam 2003) also used the Morlet wavelet to perform the segmentation of the heart sound with murmur of various intensities.

CWT is computationally demanding, therefore, a discrete wavelet transform (DWT) based approach was introduced in (Huiying et al. 1997), in which the Shannon energy was obtained from the decomposed (detail coefficients) heart sound signal. An overview of wavelet decomposition with different wavelet on range of heart sounds can be found in (Alfredson 1997).

Besides these methods other time-frequency analysis techniques, e.g. based on the matching pursuit (Zhang et al. 1998) were attempted in order to extract information from the heart sound. The matching pursuit method is based on classical Gabor wavelet or time-frequency atom which is the product of a sinusoid and a Gaussian window function (Mallat and Zhang 1993). Furthermore, it is also based on a dictionary which contains family of functions known as atoms. Therefore, a signal is decomposed into a series of time-frequency atoms using an iterative process based on selecting the largest inner product of the signal with atoms from the dictionary. In this approach, the segmentation process is performed by separating the sound components from the remaining stationary sounds by reconstructing the transient parts of the signal (Zhang et al. 1998). This method showed to be effective to remove Gaussian noise from the heart sound while separating the main heart sounds S1 and S2. To mimic the human capacity of audible discrimination of S1 and S2, one of our previous methods for heart sound segmentation utilizes Shannon energy in the wavelet domain coupled to Mel-frequency spectral coefficients (MFCC) analysis of each sound segment (Kumar et al. 2006).

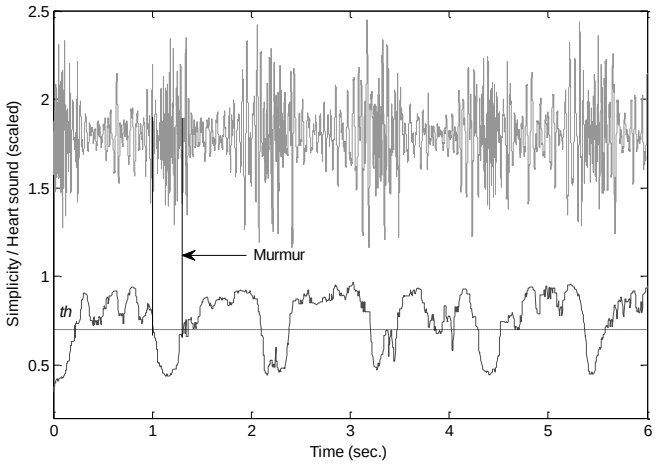


Figure 3.5. Segmentation of the murmur portion in a heart sound with aortic stenosis. Heart murmurs are segmented based on the simplicity that is below the threshold. In the example shown simplicity is computed using embedding dimension m=4 and delay = 19 parameters (see Chapter 5 for the details of these parameters).

A different analysis domain is proposed in (Nigam and Priemer 2005a). These authores attempted to delimit the main heart sound components in the presence of murmurs. Here, the heart is assumed to be a dynamic system. The segmentation is performed based on the complexity of its dynamics. The authors mainly used the heart sounds with murmurs of various grades and duration to assess their proposed method. In the original work, no results of S1 and S2 sounds' recognition in the presence of murmur were demonstrated. Simplicity (or 1/complexity) was computed (see **Figure 3.5**) using the nonlinear dynamical system theory explained in Section 5.5. The Nigam and Priemer approach computes simplicity directly from the heart sound using a fix embedding dimension parameters (see Section 5.5). It can be seen that murmurs exhibit less simplicity than other heart sound components. As a result the simplicity curve goes deep where murmurs are present.

B) Supervised Approaches

Usually, in this type of methods a typical pattern recognition approach is applied, i.e., generating features, selecting discriminative features and training a supervised classifier (e.g., ANN, SVM, GMM, etc.) with the known outcomes and inputs.

Feature generation of the sounds S1 and S2 based on wavelet decomposition and a neural network approach was proposed in (Reed et al. 2004). In this method,

the heart thorax system was modeled as a linear time-varying system with different time invariant responses.The input of the system was assumed to be the valve closure events (mainly those related to M1 and T1 for S1, A2 and P2 for S2, see Section 2.3 of Chapter 2) and the output is the sound components heard from the thorax. The sounds S1 and S2 are the output sounds producded by impulses generated by the Mitral and Tricuspid valves, and the Aortic and Pulmonary valves, respectively. The system response was then characterized using the following steps: (i) estimation of relative amplitudes and the times of input impulses, and (ii) estimation of S1 and S2 system transfer function in the frequency domain. These system transfer functions are computed from normal heart sound and heart sound with cardiac conditions. The S1 and S2 in normal and abnormal heart sounds are then used to design a neural network based classifier to perform automatic identification of heart sounds' components.

Another supervised segmentation technique was introduced based on Hidden Markov Model (HMM) (Ricke et al. 2005) which requires prior information of the heart sounds' characteristics relevant to determine S1 and S2 in a heart cycle. Therefore, this method extracts energy, Shannon energy (see equation (5.11)), and spectral features using Mel-spaced filter banks from each heart cycle in a given heart sound recording. It provides a simple method of obtaining spectral properties of an acoustic signal by creating a set of triangular filter bank across the spectrum.

3.3. Cardiac Murmur Classification

Most existing methods for cardiac murmur classification obey the sequence of the computational steps depicted in **Figure 3.6**, in which the heart cycles are identified in the segmentation stage of the heart sound analysis and the next steps are for obtaining automatic classification of various cardiac murmurs. The first challenge in murmur analysis is related with the extraction of adequate discriminative features. These are inspired by the characteristics of the different cardiac murmurs. In Table 2.8 some of the most relevant ones are listed. Most features are generated based on the computational measures of energy, spectral distribution, time-frequency transform, fractal dimension, etc. Following this, in order to identify the most discriminative set of features, a feature selection procedure is employed. At last, the key features are used as inputs to a classifier.

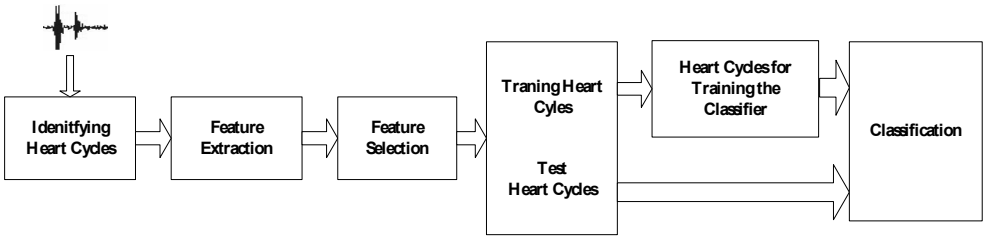


Figure 3.6. Steps usually involved in cardiac murmur classification.

Many researchers have focused their attention to find an efficient and robust solution to classify various pathological and innocent murmurs. For instance, DeGroff et al. (DeGroff et al. 2001) computed spectral energy from Fast Fourier Transform (FFT) at various spectral resolutions (1, 3, and 5 Hz) and frequency range (0 to 90, 0 to 210, 0 to 255, and 0 to 300 Hz). The spectral energy was computed at various resolutions to examine extraneous sound mixed with the heart sounds, and that can be reflected in

Previous Methods	***Number of Features***	***Murmur Classes Used in the Study***	***Reported Results (Se, Sp or Au)***
DeGroff et al.	4- Spectral energy from FFT in four frequency ranges	Innocent murmur, Pathological murmur (aortic and pulmonary stenosis, Atrial and Ventricular septal defects patent ductus arteriosus, Mitral regurgitation)	Minmum: Se=80 %, Sp=90% Maxiumum: Se= 100%, Sp=100%
Bhatikar et al.	Similar features as in *DeGroff et al.*	Innocent murmur, Pathological murmur (ventricular septal defects patent ductus artriosus, mitral regurgitation)	Se=83%, Sp=90%
Rios-Gutiérrez et al.	100 (approximation coefficients of divided in 100 summed indecies)	Mitral stenosis, Aortic stenosis, Aortic insufficiency and Mitral insufficiency	Se=73.12%, Sp=67.05%,
Gupta et al.	32 features (Power of the 2nd detail coefficients)	Normal heart sounds, systolic murmur, diastolic murmur	Averaged Au = 97.02%
Martínez-Alajarín et al.	13 (from instantaneous amplitude, energy, and frequency)	Normal heart sounds, Holosystolic murmurs, and Mid-systolic murmur	Au=100%, 92.69%, 97.57% for three classes, respectively
Strunic et al.	25 (from spectrogram matrix)	Normal sounds, Aortic stenosis and Aortic regurgitation	Au=85%

Maglogiannis et al.	100 (from sounds' envelope, energy, frequency subbands)	Normal heart sounds, Aortic stenosis, Mitral regurgitation, Aortic regurgitation, Mitral stenosis	Average, Se=91.50%, Sp=92.73%
Choi and Jiang	2 (from normalized power spectral density)	Normal heart sound, Atrial fibrillation sound, Mitral Stenosis, Aortic insufficiency	Se= 99.66%, Sp=100%

Table 3.1 Features and classification for heart murmur classification by the researches in recent past years. In the result column, Se is sensitivity, Sp is specificity, and Au is accuracy.

higher resolution spectrum. Furthermore, four frequency ranges were chosen based on the heart sound data (acquired from subjects of pathological murmurs and innocent murmur) for the study that exhibited spectral energy concentration in these. Spectral energy in four frequency ranges at different frequency resolution was input to a layered artificial neural network (ANN). A result of 100% sensitivity and specificity was reported using a database collected from 69 pediatric subjects in whom 70% of the patients had an innocent murmur. A similar approach was used in (Bhatikar et al. 2005) to discriminate innocent murmurs from pathological ones. However, the number of subjects in the database was increased to 241, of which 88 were innocent murmurs and 152 pathological murmurs (44 being murmurs of VSD). The reported results were of 83% sensitivity and 90% specificity for the complete database. These results were slightly higher as 90% sensitivity and 93% specificity in case of the discrimination between innocent murmurs and murmur of VSD.

A feedforward neural network for murmur detection and two Learning Vector Quantization (LVQ) networks were employed for murmur classification by Rios-Gutiérrez *et al.* (Rios-Gutiérrez et al. 2006). Discrete wavelet transform (DWT) with the mother wavelet Daubechies (db4) was used for the wavelet decomposition of the heart sound. Due to the ability of DWT to decompose a signal into its approximation coefficients (low pass filtered and decimated) and detail coefficients (high pass filtered and decimated), successively, it extracts the intensity of sounds in different subbands. Therefore, this technique after applying to different heart sound signals exhibited important characteristics for heart murmurs of various frequency ranges. Features were computed by normalized summation of the approximation coefficients of fourth and fifth level separately. These decomposition depths were selected based on capturing significant aspects of the heart murmurs used for classification. The murmur classification system was designed to categorize the detected murmurs into four common types: mitral stenosis, aortic stenosis, aortic insufficiency and mitral insufficiency. Two LVQ networks were trained and tested using the fourth level and fifth level of discrete wavelet transform approximations of the heart signals. The reported accuracy rate for the detection and classification subsystems

were of 73.12% and 67.05%, respectively.

In (Gupta et al. 2007), a database of 340 heart sounds was applied to classify between normal heart sounds, systolic murmur and diastolic murmur. This method also used the wavelet decomposition technique for feature extraction. However, it employed Grow and Learn (GAL) and Multilayer Perceptron - Backpropagation (MLP-BP) neural networks to perform the classification. Depending upon the heart murmur classes in the database the authors aimed to determine power of the heart sound signal in its frequency subbands. In order to compute features, the second level detail coefficients of wavelet decomposed heart sound using Daubechies-2 were selected. Each window is divided into 32 sub-windows containing each one 128 discrete data point, and finally power of each sub-window is computed (in fact, these features were originally introduced in (Olmez and Dokur 2003), which will also be explained in Chapter 6). These were applied as the input feature matrix to the Grow and Learn (GAL) and Multilayer Perceptron - Backpropagation (MLP-BP) neural networks to perform classification. According to the authors GAL outperformed the MLP-BP by producing an accuracy of 97.01% for normal heart sounds, 98.50% for systolic murmurs, and 95.55% for diastolic murmurs.

Another use of the MLP neural network with a different set of features can be found in (Martínez-Alajarín et al. 2007). Herein, the authors used 13 features extracted from temporal measurements and from the envelopes of the instantaneous amplitude, energy, and frequency as well as the duration of the murmurs. The latter was added in order to enable the classification of mid-systolic and holosystolic murmurs. A feature vector composed by 13 features extracted from the duration (one) as well as from IA, IE, and IF (four from each) were computed. To reduce redundancy in the feature matrix, principal component analysis (PCA) analysis was applied (Abbas and Bassam 2009). Using 97 heart cycles, of which 31 normal ones, 26 containing holosystolic murmur and 37 mid-systolic murmur, classification accuracies have been reported to be 100%, 92.69%, and 97.57%, respectively, for the three aforementioned classes of heart sounds.

In (Strunic et al. 2007), a feed-forward neural network classifier was employed. There, input features were extracted from the spectrogram of averaged heart cycle. In this study, instead of processing individual heart cycles of the heart sound recording for feature extraction the average of all heart cycles was used in order to obtain better representation of the completed heart sound compare to single heart cycle and also reduce the effect of anomalous heart cycles. From the spectrogram matrix the values corresponding to frequency band 0 to 195 Hz were selected because experimentally they were found to contain necessary information for ANN to identify the types of murmur used in the study. This system attained 85% accuracy in the discrimination between normal sounds, sounds with aortic stenosis and sounds with aortic regurgitation.

In (Maglogiannis et al. 2009), 100 features have been suggested to construct a multi-SVM for discriminate among normal heart sounds and four categories of murmurs (two systolic and two diastolic) murmur, i.e., aortic stenosis, mitral regurgitation, diastolic murmur and aortic regurgitation. Features were computed after demarcating four phases in the heart cycles, i.e. S1-systole-S2-diastoles, using a segmentation technique. The feature vector is constructed in order to capture frequency as well as morphological characteristics in each heart cycle phase. These features were fed to a support vector machine (SVM) based classifier which classifies in three steps: (i) healthy heart sounds from pathological heart sound (heart murmur), (ii) pathological sounds were classified into systolic murmurs and diastolic murmurs, (iii) systolic murmurs were classified into aortic stenosis (AS) and mitral regurgitation (MR), whereas diastolic murmurs were classified into aortic regurgitation (AR) and mitral stenosis (MS). The overall reported results for 160 pathological heart sound signals were 89.29% sensitivity and 93.23% specificity for systolic and diastolic murmur classification, 90.48% sensitivity and 92.86% specificity for systolic AS and MR murmur classification and 94.74% sensitivity and 92.11% specificity for AR and MS.

Recently, Choi and Jiang (Choi and Jiang 2010) developed a pattern recognition system for heart murmur classification using features extracted from a novel spectral analysis technique, i.e. the normalized autoregressive power spectral density (NAR-PSD) and classifier multi-support vector machine. The motive behind the use of NAR-PSD was to extract features based on spectral analysis of the heart murmurs. The advantage of normalized AR-PSD (computed according to (Choi and Jiang 2010)) is seen in heart murmurs where high-density frequency information distributed along a wide region compared to normal heart sounds which were important to discriminate one murmur from the other. NAR-PSD was calculated from the normalized heart sound signal by setting the variance of the signal to 1.0. Two diagnostic features were introduced for the classification of heart murmurs: frequency index of the peak value of the NAR-PSD curve and frequency elapse between the crossed points of NAR-PSD curve on a threshold (THV). Threshold value has direct connection with the value of frequency width; therefore classification results were obtained on different values THV in order to see the best possible value. The authors used this method on a large and diverse database that included 189 normal heart sounds and 293 abnormal heart sounds. Abnormal heart sounds contained 6 types of heart murmur patients: atrial fibrillation, aortic insufficiency, aortic stenosis, mitral regurgitation, mitral stenosis, and split sounds. Reported classification performance on the basis of sensitivity and specificity was found as 99.66% sensitivity and 100% specificity at the value of 10 THV.

3.4. Limitations

Here, we examine the shortcomings of the algorithms pertaining to analysis of a heart sound recording. The methods applied in the past few years related to the heart sound tasks under analysis in this chapter reported significant results. However, these methods suffer from limitations and, sometimes, very specific requirements to achieve adequate results.

The methods presented and reviewed in the earlier sections for three problems related to heart sound analysis are suffered shortcomings which may not be favorable in the aim to prepare a reliable diagnosis for certain CVD. Several methods have been developed thus far which promise significant results. Nevertheless, most of the methods among these ones highly rely on the type of heart sound database. Therefore, sometime limitations are with the type of heart sounds, that is, when a certain heart sound, mainly pathological heart sounds, are analyzed with the existing methods results may not be satisfactory for real-life use due to complex nature of heart murmurs. Here, we focus on the majors issues related to mainly the methods presented before for not considering to implement during the study.

Noise detection is one of the approaches to obtain a noise free auscultation. As it was mentioned that not many methods are known to exist which focus on determining non-cardiac sounds in the heart sounds recording (online or offline) and then remove it from the complete heart sound recording. Usually, background noise is cancelled with the help of an auxiliary signal (ECG or other acoustic signals) which requires another sensor to acquire it (Bai and Lu 2005; Belloni et al. 2010; Carvalho et al. 2005; Giustina et al. 2011; Paul et al. 2006; Várady 2001; Wang et al. 2007). Two main problems can be observed with the set of methods which require an additional signal for noise cancellation: (i) additional hardware is required; (ii) noise is not completely reduced from the input signal, although the signal distortion is less than the one obtained using conventional fixed filters. As mentioned earlier in Secton 3.1 that during auscultation of the heart sounds capture environmental and noise produced by bystanders is mixed additively or convolutley. It is observed that its spectral power distribution might spread the band of interest. Furthermore, these noise sources may not be Gaussian or stationary.

Single channel approaches for the noise reduction are also applied with certain assumptions and target the elimination of specific types of noises. For examples, a the daptive line enhancement (ALE) assumes that heart sound is corrupted with additive white noise (Tinati et al. 1996), while wavelet denoising is capable to remove particularly stationary high frequency components (Boutouyrie et al. 1995; Strang and Ngugen 1996; Xiu-min and Gui-tao 2009). Other methods assume that heart sound has only two components, i.e. S1 and S2, whereas other components, such as murmur, or low amplitude and low frequency sounds S3 and S4 are com-

pletely neglected. On the other hand, in blind source separation based methods it is assumed that heart sound components are generated by different source. These algorithms require the estimation of the mixing matrix, which tends to be computationally demanding and might not enable realtime. Furthermore, these methods can only successfully separate super-Gaussian[10] signals. BSS is an unexplored field in case of heart sound components separation; it requires much more attention for heart sound analysis as there are several algorithms based on time domain and frequency domain available which can be further investigated (Pedersen et al. 2007). An obvious disadvantage of multi-channel BSS approaches is the added complexity in hardware and software.

Hence, it is observed that in many contexts, such as in pHealth applications, the best strategy for this task might be the detection of contamination in order to alert the patient or the bystander.

Heart sound segmentation was also largely attempted with one or more auxiliary signal (El-Segaier et al. 2005; Wilton-Davies 1972). However, as it was stated before (see Chapter 1) the defined objectives in this thesis is to provide heart sound analysis without the support of any reference signal and, therefore in spite of showing high accuracy results these methods were not considered. Our methods of segmentation were inspired by the unsupervised approaches that exclude the use of an auxiliary signal (following the procedure shown in **Figure 3.3**) in order to perform components, delimitation and recognition of the sounds, of segmentation. Most single channel based methods differ mainly at the sound segment delimitation phase. For the recognition of the sound components, most approaches found in literature follow similar approaches based on physiological observations. Algorithms based on Shannon energy (Altunkaya et al. 2010; Liang et al. 1997; Nazeran 2007; Wang et al. 2005), Renyi entropy (Yadollahi et al. 2005), fractal dimension (Hasfjord 2004) were effective in identifying heart sound components S1 and S2 in absence of any additional heart sound components, such as S3, S4, or murmur. However, in the presence of murmur or sounds S3 or S4, it is observed that this simplistic approach is not sufficient to provide accurate delimitation for the all the available sounds due to overlap of features between sounds components. Hence, two main problems can be observed in such algorithms: (i) since features are mainly the reflection of amplitude (Shannon energy, Energy) or frequency contents (Entropy, Fractal dimension) of the sound samples; hence, some overlaps of the values in feature curve are expected given that the amplitude and frequency ranges of the main heart sounds, i.e., S1/S2, and murmur tend to overlap (Erickson 2003). In such situations, a threshold may not delimit all heart sound components' boundaries. (ii) Due to S3 and S4 sounds being less loud and lower frequency than the main heart sounds the possibil-

[10] Super Gaussian signals have probability distribution function more sharp and high peak than normal Gassussian signals.

ity of their to burry into background noise is high and therefore can not be clearly identified using one feature.

Time-frequency based methods, STFT (Kam 2003; Liang et al. 1998), CWT (Corona and Torry 1998), Wigner-ville distribution (Boutana et al. 2007) or matching pursuit (Zhang et al. 1998) based algorithms have been used for heart sound segmentation. However, these approaches suffer from the time-frequency resolution uncertainty, which has an impact on accuracy of heart sound components delimitation. STFT suffers with poor time-frequency resolution because of a fix size of windowing function in the process of transformation that may cause to burry low frequency sounds in the background noise. CWT changes window size depending upon signal frequency under transformation and subsequently improves resolution but since frequency representation is in fact scaling the mother wavelet it requires significant computational time. Wigner-ville distribution produces cross terms that may resemble false heart sound components, and matching pursuit is complex and computationally expensive. The DWT on the other hand is an inspiration for the algorithms introduced by us in this thesis due to its ability to decompose a signal in successive frequency bands. Another recent approach for segmentation is based on complexity measurement, which is invariant to envelope amplitude variations, and hence alleviates many of the problems of traditional segmentation.

Under the diagnosis from the heart sound, heart murmur classification is the one of the final tasks. It requires extraction of several such features which may characterize the systolic/diastolic murmurs and separate one from others. The methods shown in Table 3.1 indicate that several methods have been developed aiming to classify individual datasets. Depending upon the type of heart murmur and grades of its audibility, features were extracted from each heart cycle. It has been noticed that frequency based features in almost every method have important contribution in classification of different types of murmurs of different morphologies. Therefore, main limitation for the approaches can be the strength of dataset which must be as large and diverse as possible.

3.5. Conclusions

Heart sound is a highly valuable biosignal that has been traditionally be applied for medical diagnosis and prognosis. Due to technological advancements, namely in the field of the ultrasound and due to the high level of proficiency required, heart sound has lost importance in clinical in the daily clinical practice. However, it is observed that in the last years, heart sound processing has regained a significant attention from the research community, it is a cheap, non-invasive signal that encodes the mechanical status of the heart and hence, lead itself to be used for the deployment

automated diagnosis and prognosis medical systems. Given the high number of published approaches for heart sound analysis presented in Section 3.1 through 3.3 one might argue if technology has not reached maturity for practical clinical applications. Nowadays, it is observed that one of the most challenging application context is remote management of chronic diseases, being cardiac diseases the most meaningfull. These challenges stem from the required robustness and selfadaptation capability of analysis algorithms that have to operate adequately in non-controlled environments under non-professional guidance or support and, still, have to provide very high levels of sensitivity and specificity. In these contexts we highlight the following requirements that have a decisive impact on the adoption of the technology:

- Usability: solutions must be as simple as possible in order to provide high levels of usability. Chronic patients are mainly elderly persons with significant physical impairment. In order to avoid impacting on long term utilization, systems should be applicable without any or much reduced assistance. Single signal solutions usually lend themselves to realizations that exhibit higher degrees of usability.
- Robustness: solutions have to operate under non-controlled environements and will be operated by non-experts. Robustness and selfadaptation/selfcalibration capability will have a very direct impact on long-term adoption of pHeath solutions, since it will affect confidence in the system as well as usability. False positive rates must be as small as possible in order to avoid alarming unnacessarly the user. On the other hand false negatives must be avoided in order not to compromise the confidence in the system.
- Computational complexity: solutions must be efficient in terms of computational cost in order to provide feedback to users in real-time and for energy efficiency reasons.
- Energy efficiency: pHealth systems must be energy efficient in order to avoid frequent battery charges which would affect the usability of the system. Single channel solutions have usually a significant impact in terms of required hardware and required local processing, lending thenselfes to simpler and more energy efficient realizations.

From the above indicators, it becomes clear that existing methods do not meet these requirements adequately and therefore further research is required.

Chapter 4

NOISE DETECTION

Noise identification in the heart sound is one of the first important tasks in the heart sound analysis. Heart sound is a valuable biosignal for diagnosis of a large set of cardiac diseases. Ambient and physiological noise interference is one of the most usual and highly probable incidents during heart sound acquisition. It tends to change the morphological characteristics of heart sound that may carry important information for heart disease diagnosis. Therefore, extraneous sounds removal in the heart sounds removal must be given attention greater importance before further analysis is carried out.

In this chapter, we propose a new method that can be applicable in real time to detect ambient and internal body noises manifested in heart sound during acquisition. This method exploits the rhythmic property of the heart that is present in the phonocardiogram. Namely, it is observed that when external or internal noise is injected during heart sound acquisition the rhythmic pattern is perturbed. We were inspired by this phenomenon to design an algorithm using single heart sound recoding. In this chapter the details of the method is explained.

Section 4.2 Overview

We start with a brief description of the approach. The basic steps, i.e., the first phase in which a reference signal is determined, and the second phase where template matching based criterion is adopted to assess whether a segment of heart sound is noisy, are outlined.

Section 4.2 Phase I: Reference Sound Detection

In this section, the details of the phase I introduced. The reference signal that contains noise-free heart sound, are explained. In relation, the periodicity of heart cycles extracted from the heart sound is assessed in the temporal as well as in the

time-frequency domains in order to identify an uncontaminated heart cycle of heart sound that will serve as a template.

Section 4.3 Phase II: Non-cardiac Sound Detection

After detecting a template of noise-free heart sound, the next step is to identify noisy and clean heart sound based on the matching of certain features of the template and the sound window under processing for the assessment. Hence, the features which are required to execute the matching process are introduced in this section.

Section 4.34 Test Database

This section includes description of the heart sound database prepared for the assessment of the algorithm. Our database contains different heart sounds acquired from different subjects, purposely injected with several ambient and bystanders' instigated noises. The data collection protocol is detailed here.

Section 4.3 Performance Assessment Measures

In this section, we define the performance measures that will be applied to assess the robustness of the algorithm.

Section 4.36 Experimental Results and Discussion

The achieved results by applying the measure defined in the previous section are presented here. Four categories of noises are further divided into three subcategories based on their intensity in order to examine the efficiency of the algorithm with the mild intensity range of noisy sounds.

4.1. Overview

Figure 4.1 depicts the proposed approach for the detection of heart sound segments with noise contamination. As it can be observed, the method is composed by two main phases, denoted as Phase I and Phase II. The goal of phase I is to find a segment of uncontaminated heart sound correspondent to one complete heart cycle. This heart cycle will be treated as a reference sound in phase II of the algorithm. In phase I, given a window of heart sound, first the algorithm extracts the heart rate from the envelope of the heart sound. This is performed resorting to the autocorrelation function and corresponds to step (A1) described next. The estimated heart

cycle duration is then applied to the segment's individual heart cycles, which are then assessed for noise contaminations in the time domain (see left flow of phase I in **Figure 4.1**).

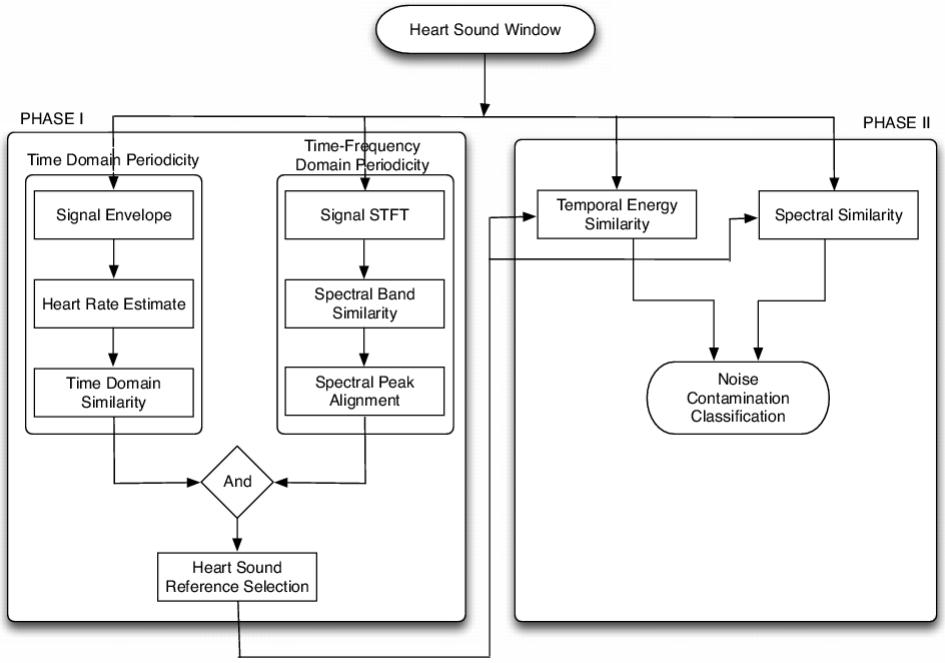


Figure 4.1. Flow chart of the proposed approach.

The later is based on the observation that heart sound exhibits a quasi-periodic behavior in the time domain, which is assessed using a similarity measure between adjacent heart beats (see step A2). This quasi-periodicity behavior of heart sound also manifests itself in the time-frequency domain. Namely, it is observed that the time distribution of the energy in adjacent frequency bands tends to be (i) highly correlated and (ii) their peaks tend to be aligned. These conditions are inspected in the processing steps on the right path of the flow chart of phase I in **Figure 4.1** and will be described in steps (B1) and (B2), respectively. The heart cycles, which meet simultaneously the temporal similarity as well as the spectral similarity criteria are taken as candidates for the reference heart sound template. The actual template is selected among this set of candidates. In phase II (see the right path of the flow chart in figure 1), each acquired heart sound window is checked for similarity against the selected heart sound reference template. Here two criteria have to be fulfilled simultaneously: (i) spectral similarity and (ii) temporal energy similarity. The later is taken into account in order to determine short duration sound spikes,

which are not captured by the spectral similarity criterion.

4.2. Phase I: Reference Sound Detection

This section introduces the algorithm steps required to identify candidates for the heart sound template.

4.2.1. Periodicity in Time Domain

The envelopes of heart sound components are extracted by applying the Hilbert transform followed by the Gammatone band-pass filter (Deshmukh et al. 2005). Next, the autocorrelation function of the envelope is computed. Typically, it will exhibit pronounced peaks for the main heart sound components, i.e. the S1, S2 and murmur components. The autocorrelated values are normalized by the autocorrelation values of a chosen windowing function, e.g. the Hanning window. Let $x^e(t)$ be the envelope of the heart sound, $x(t)$ and let $y(\tau)$ be its autocorrelated function (ACF) then ACF can be written as in (4.1).

$$y(\tau) = \frac{\int_{-\infty}^{\infty} x^e(t)w(t)x^e(t-\tau)w(t-\tau)dt}{\int_{-\infty}^{\infty} w(t)w(t-\tau)dt} \tag{4.1}$$

In (4.1), $w(t)$ is the Hanning window function, variable τ denotes lag in computation of the autocorrelation that varies from zero to the duration of the analyzed signal window (the duration of a given segment of heart sound is 4s). Herein, we also propose to use the autocorrelation function to identify the heart cycles present in the analysis window. However, this cannot be done directly by identifying the prominent peaks in $y(\tau)$, since the prominence of the autocorrelation peaks tend to be significantly disturbed in the presence of noise and murmur. Even in clear heart sounds, it is observed that usually false prominent peaks are detected. To solve this, the number of prominent peaks in $y(\tau)$ is constrained using the estimated heart rate. The actual algorithm to select the prominent peaks, which define the heart cycles, is based on an adaptation of Park's method (Park 2000), i.e. (i) first the peak with the highest amplitude is selected; (ii) the algorithm searches the adjacent prominent peak to the left using a window whose duration equals the heart cycle period; the peak with the highest amplitude inside that window is selected; (iii) the same procedure described in step (ii) is repeated with a window to the right; (iv) steps (ii) and (iii) are recursively repeated using the selected prominent peaks in the previous iteration; (v) the algorithm stops when the number of selected prominent

peaks equals the number of heart cycles inside the analysis window (these are calculated using the estimated heart rate and the duration of the window). Once all prominent peaks have been detected, the periodicity is checked using the radial distance between two contiguous heart cycles. The heart rate estimation and periodic cycle's verification procedures are as follows.

A1) Heart Rate Estimation

In the absence of an auxiliary signal, such as the ECG, it is observed that the heart rate can be estimated using the first two singular values of the rearranged autocorrelation function of the envelopes of the segment of heart sound under analysis by extending the algorithm described in (Kanjilal and Palit 1994). Let $y(\tau)$ be periodic with period T', i.e. $y(\tau + T) = y(\tau)$. For the typical population at rest, it is observed that T is between 375 and 1500 ms (heart rate is between 40 and 160 bmp (beats per minute)). Let $Y(T)$ be the data matrix which is constructed by stacking $y(\tau)$ samples after every T' ms using the following arrangement

$$Y(T') = \begin{pmatrix} y(0) \dots \dots \dots \dots \dots \dots y(T') \\ y(T' + T_s) \dots \dots \dots \dots y(2T') \\ \cdot \\ \cdot \\ \cdot \\ y\big((m-1)T' + T_s\big) \dots \dots \dots y(mT') \end{pmatrix} \tag{4.2}$$

where m is the number of periodic analysis segments in a given segment of the heart sound, and T_s is the sampling period i.e. the time interval between two consecutive lags in $y(\tau)$. If $Y(T)$ has linear dependent rows, then it will have some zero singular values. The singular values are found by singular value decomposition (SVD) of $Y(T)$. Let $\sigma_1 \geq \sigma_2 \geq \sigma_3 \dots \dots \geq \sigma_m$ be the aforementioned singular values, the periodicity can be measured using (4.3).

$$\rho = \left(\frac{\sigma_2}{\sigma_1}\right)^2 \tag{4.3}$$

A value of ρ near zero implies a strong periodicity in the signal. In the process of estimating the heart rate, $Y(T')$ is constructed by varying T'. The T' which minimizes ρ, is the estimated duration of the heart cycle. The estimated heart rate enables us to find the prominent peaks in $y(\tau)$. Since prominent peaks are directly related to the main components in the heart sound, which occur only once per heart cycle (see **Figure 4.2**), this is used as the prior information to find peaks.

Namely, inside each processing window, first the largest peak in amplitude is identified. Then the algorithm determines the second largest peak at a distance not greater than the identified heart cycle duration from the first peak. This approach is repeated for each newly identified peak until the total number of prominent peaks found equals the number of complete heart cycles in the processing window.

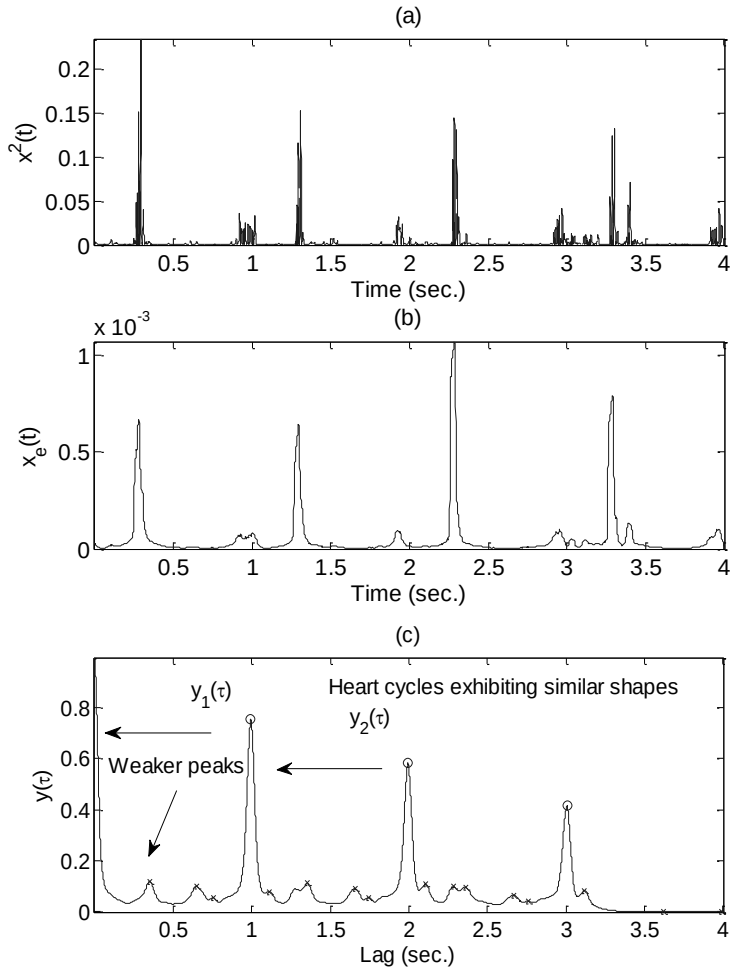


Figure 4.2. (a) Heart sound energy. (b) Heart sound envelope. (c) Autocorrelation function $y(\tau)$ of the envelope with peak identification (showing similar heart cycles in terms of vectors $y_1(\tau)$ and $y_2(\tau)$.

A2) Periodicity Check Criterion

The set of prominent peaks in $y(\tau)$, which has been detected in the previous step,

enables us to find shape similarity between two heart cycles (contiguous pair of prominent peaks). Since the prominent peak detection algorithm is constrained by the estimated heart rate, it is observed that each pair of adjacent prominent peaks defines a complete heart cycle. The similarity is measured by the radial distance between two vectors. Let $y_r(\tau)$ and $y_{r+1}(\tau)$ be the vectors representing the portions of the autocorrelation function of two consecutive heart beats; then the radial distance is given by

$$Cos(\theta) = \frac{\langle y_r(\tau), y_{r+1}(\tau) \rangle}{|y_r(\tau)||y_{r+1}(\tau)|} \tag{4.4}$$

where $<.>$ is the inner product operator and $|\cdot|$ represents Euclidean norm of the vectors.

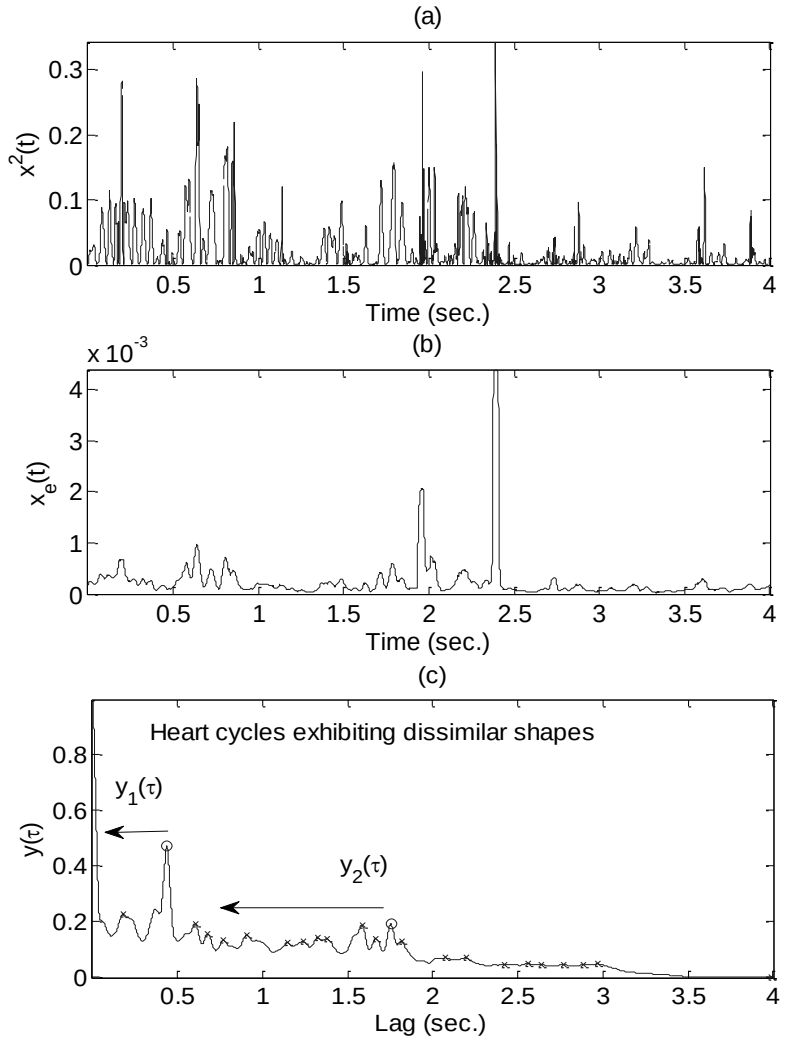


Figure 4.3. (a) Noisy heart sound energy. (b) Heart sound envelope. (c) Autocorrelation function $y(\tau)$ of the envelope with peak identification (showing dissimilar heart cycles in terms of vectors $y_1(\tau)$ and $y_2(\tau)$.

In (4.4) $y_r(\tau)$ and $y_{r+1}(\tau)$, where r is the number of prominent peaks in the analysis window, are interpolated if they do not have the same length. It is considered that two consecutive heart cycles exhibit similar shapes if their internal product is greater than 0.8 (this value is chosen using the ROC plot by varying $Cos(\theta)$ from 0.6 ro 1 along with the changing values of number of aligned peaks autocorrelated frequency band explained in the next subsection). ROC plot was drawn for 13 heart sound samples to estimate optimum similarity threshold (see **Figure 4.7**). This

threshold was tuned experimentally. Similar shaped and dissimilar shaped heart cycles (similarity indicate equal duration heart cycles whereas and dissimilarity irregular duration heart cycles) in clean and noisy heart sounds' autocorrelation function can be observed in figures 2(c) and 3(c), respectively.

4.2.2. Periodicity in Time-Frequency Domain

The periodicity check of heart sounds in the time domain is not sensitive to the presence of many non-cardiac sounds. For instance, swallowing, breathing or high pitched voice may not be identified using the time domain periodicity detection technique. However, the influence of the noise source in heart sound periodicity can be observed in the time-frequency domain. In the proposed method, the spectrogram is adopted to find periodic patterns in the time-frequency bands. Spectrogram is constructed based on the short time Fourier Transform (STFT), $S(f,t)$, of $x(t)$, which can be written as in (4.5).

$$S(f,t) = \int_{-\infty}^{\infty} x(\eta)w(\eta - t)\exp(-2jf\pi\eta)d\eta \tag{4.5}$$

In (4.5), $w(\eta)$ is the Hanning window function. The windowing function satisfies $w(\eta) = 0$ for $\eta > T/2$, where T, as mentioned in section 4.2.1, is the length of analyzed heart sound segment. In heart sounds produced by healthy subjects as well as by individuals who exhibit cardiac diseases, it is observed that most of the energy of the heart sound signal is concentrated in the 0-600 Hz frequency range. In the proposed approach, this frequency range is divided into 15 evenly divided frequency bands and the spectral energy in each of the bands is taken for periodicity validation. In **Figure 4.4**(c) the autocorrelation functions of the energy in each of these 15 frequency bands

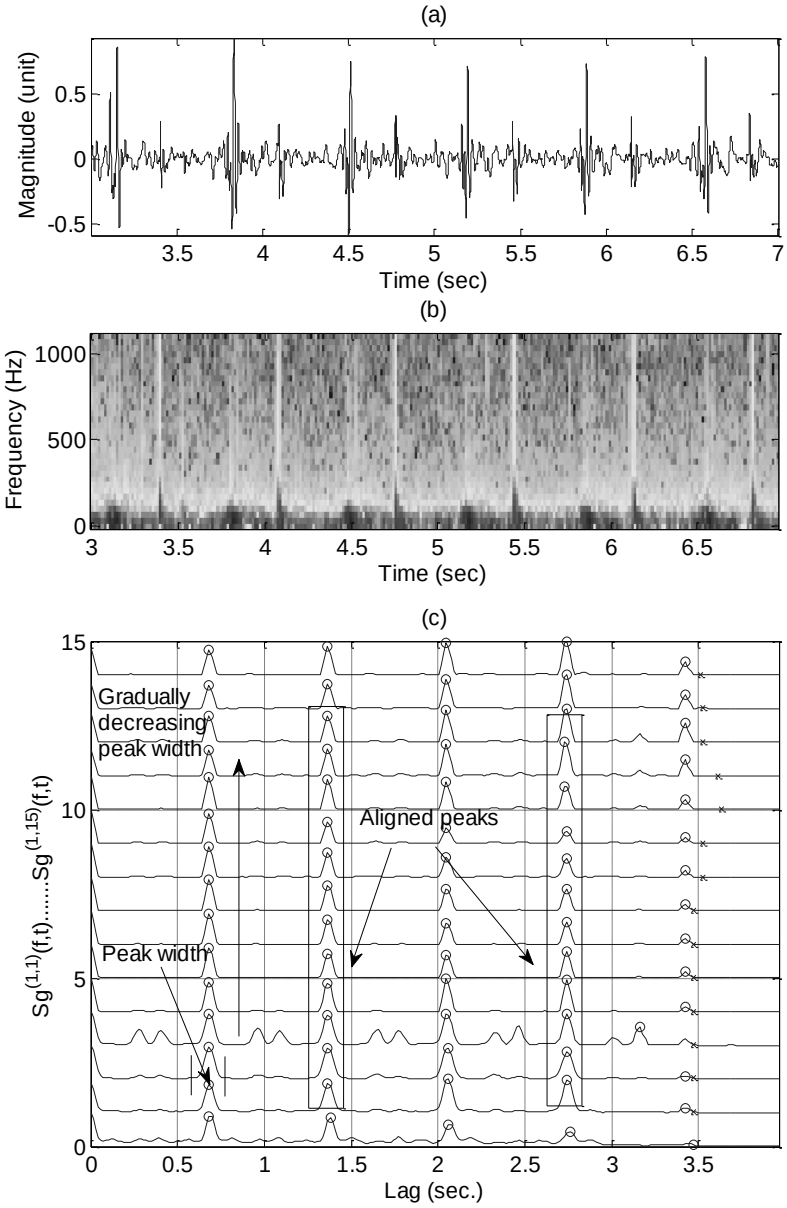


Figure 4.4. (a) Normal heart sound from a single-tilted disc mechanical prosthetic valve in the mitral in the mitral position. (b) Spectrogram of (a). (c) Autocorrelation function (a) Noisy heart sound energy. (b) Heart sound envelope. (c) ACFs as $AS^k(\tau)$ of spectral bands.

are depicted for a normal heart sound. As can be observed, the heart sounds exhibit regular patterns in the autocorrelation functions computed for each frequency band.

Furthermore, these regular patterns tend to be linearly dependent between increasing frequency bands. These linear dependencies may monotonically increase or decrease depending on the type of heart sound. Namely, it is seen that it tends to decrease for heart sounds associated to natural or to bioprosthetic heart valves, while for heart sounds produced by mechanical prosthetic heart valves it is observed that the linear dependency tends to increase with increasing frequency. The later is associated to the fact that heart sounds induced by mechanical heart valves exhibit much higher frequency content compared to natural or bioprosthetic heart valves. Furthermore, it is also seen that the peaks of the autocorrelation functions in each frequency band tend to be aligned in time. This is due to the fact that most of the signal's power is due to the S1 and the S2 heart sound components which are responsible for the main peaks in the autocorrelation functions. The methods for the verification of S1 and S2 periodicity in the time frequency domain using the 15 frequency bands and the criterion for peaks alignment in these bands are explained next.

B1) Pattern Detection in Frequency Bands

In order to extract the periodic patterns from the spectrogram, ACF of the energy of

$$S_g{}^{(m,m+4)} = \begin{pmatrix} AS^m(\tau) \\ AS^{m+1}(\tau) \\ \vdots \\ AS^{m+4}(\tau) \end{pmatrix}, m = 1, 6, and\ 11 \qquad (4.6)$$

In equation (4.6) $S_g{}^{(m,m+4)}$ is the matrix formed by grouping the $AS^k(\tau)$ rows for each five contiguous frequency bands. The linear dependence is assessed using (4.3), since ρ should be near zero for linear dependent rows. Regarding the monotony assessment of the linear dependence in increasing or decreasing frequency bands, ρ may also assist in its verification. Let ρ_1, ρ_2 and ρ_3 be the singular values ratios computed using (4.3) with the eigenvalues calculated from matrixes $Sg^{(1,5)}(\tau), Sg^{(6,10)}(\tau)$, and $Sg^{(11,15)}(\tau)$ respectively, then the most significant observations regarding pure heart sounds are: $\rho_1 > \rho_2 > \rho_3\ or\ \rho_1 < \rho_2 < \rho_3$.

In heart sounds, two distinct situations are observed for (i) native and bioprosthetic valves and (ii) for mechanical prosthetic valves. For native valves or bioprosthetic valves, it is seen that the main signal energy is concentrated in the lower frequency spectrum. Higher frequency components appear only at very short time

periods (e.g. at the onsets of the aortic valve closing in S2). Furthermore, the spectral range of heart sound produced by natural and bioprosthetic valves is lower compared to the spectral range of heart sound produced by mechanical prosthetic valves. Hence, the autocorrelation function is less regular as frequency increases (see **Figure 4.5**).

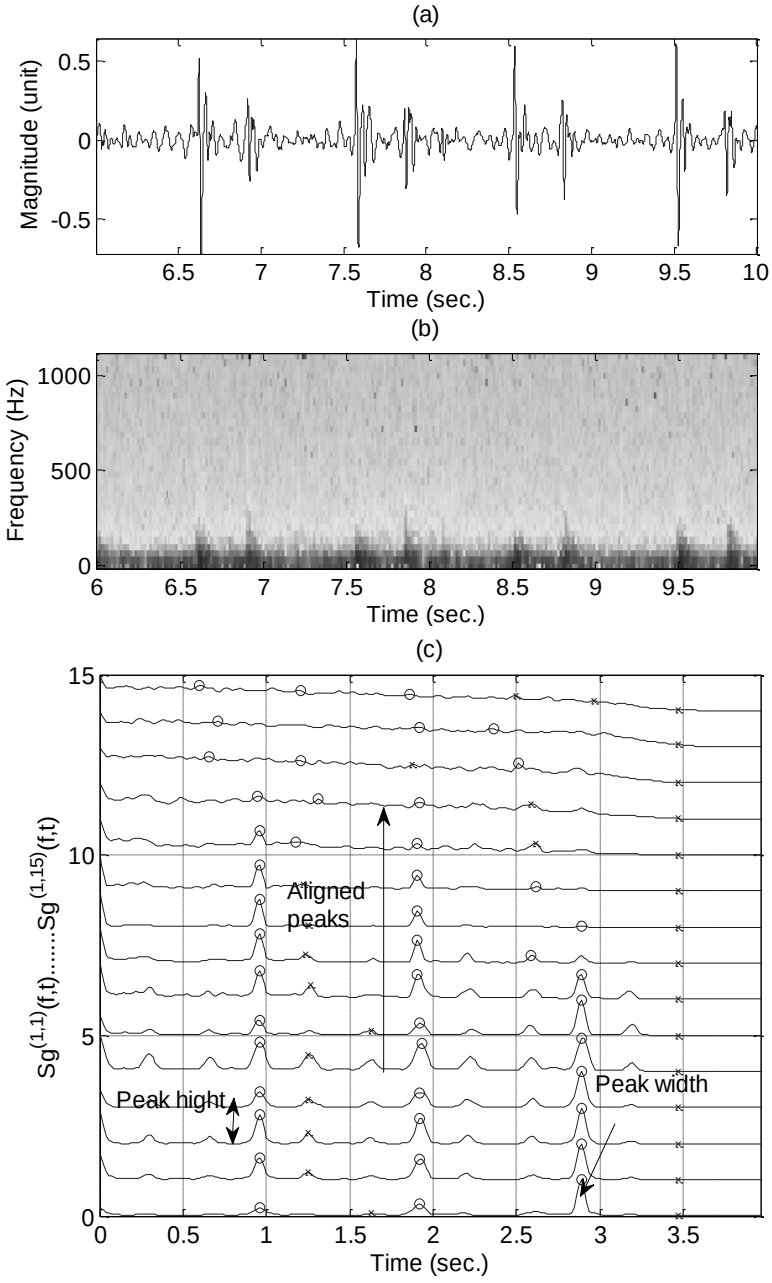


Figure 4.5. (a) Normal heart sound from native valve. (b) Spectrogram of (a). (c) Autocorrelation function.

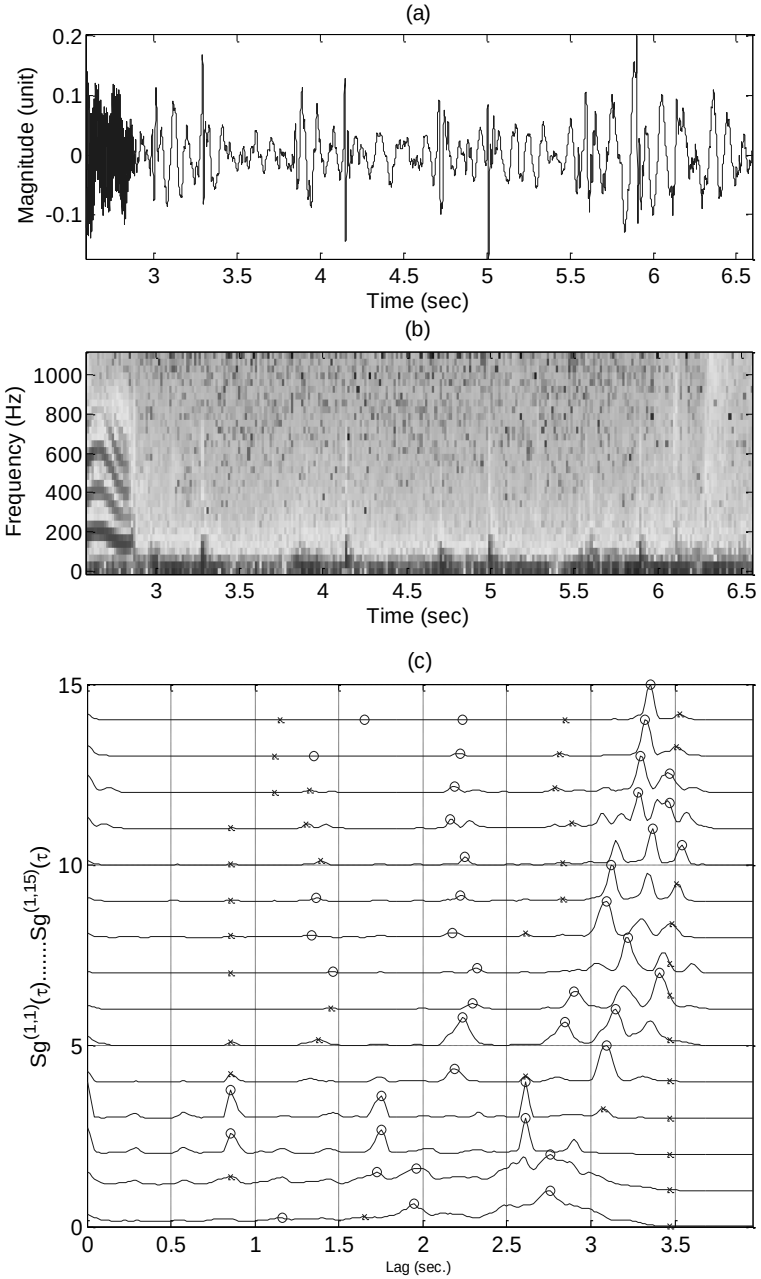


Figure 4.6. (a) Noisy heart sound from valve. (b) Spectrogram of (a). (c) Autocorrelation function.

This leads to the lower linear dependency of the autocorrelation function as frequency bands increase, i.e. $\rho_1 < \rho_2 < \rho_3$. Another situation is found in mechanical

prosthetic valves. In this situation, it is observed that the higher frequency bands have considerably more energy in a broader frequency range, leading to more pronounced and better defined peaks in time (i.e. peaks of considerable amplitude and very well localized in time) in high frequency bands. Hence, the linear dependency increases in this situation, i.e. $\rho_1 > \rho_2 > \rho_3$. On the contrary, these two conditions are not met in the presence of noise (see an example in **Figure 4.6**).

B2) Peak Alignment in frequency Domain

As has already been explained, the main autocorrelation peaks are due to S1 and S2 HSs. Therefore, they should exhibit alignment in ${S_g}^{(m,m+4)}(\tau)$ frequency bands. Otherwise, it is considered that these peaks are due to noise. In order to check the alignment, all main peaks are found using the previously described peak detection technique. Afterwards, defining a time tolerance ($\pm 5\%$ of the time of the peak in the first frequency band) the alignment of all peaks is inspected (see an example of aligned peaks in **Figure 4.4**(c)). Although these peaks do not follow regular alignment in the presence of noise, as can be seen in **Figure 4.6**(c). Finally, if 80% (obtained using ROC plot shown in **Figure 4.7**) of the total peaks are detected aligned and the monotonicity criterion is met, then the analysis window is assessed as an uncontaminated heart sound segment.

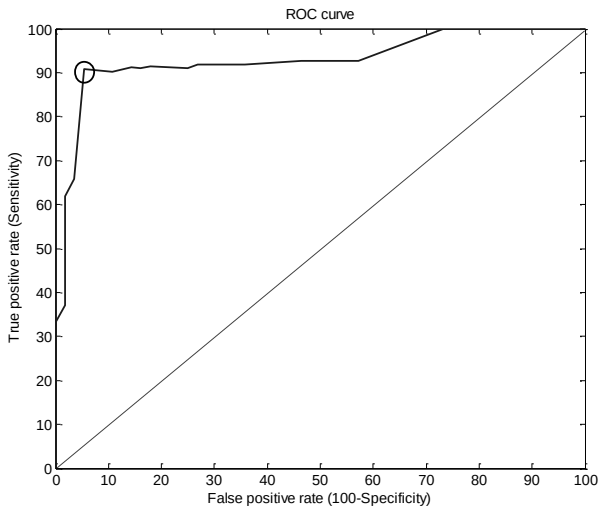


Figure 4.7. ROC curve for 13 heart sound samples to identify optimum cosine similarity threshold and number of aligned peaks (circled kink represents optimum value).

4.2.3. A Heart Cycle Template Selection

The processing steps described in the previous subsection enable the identification of uncontaminated heart sound. These steps are applied to identify a window containing clear heart sounds that will serve to extract a heart sound reference template. In order to achieve this goal, first complete heart cycles are identified inside the window using the autocorrelation function peaks. Second, the heart cycle with the highest average similarity (radial distance) with respect to all available heart cycles in the window is selected as the reference heart cycle template.

In the current implementation of the algorithm, a window of 4 seconds of heart sound is taken to estimate the reference heart sound. This analysis window was found to be sufficient to examine the signal's periodicity in the time and in the time-frequency domain. If the signal satisfies the periodicity conditions introduced in section 4.2.1 and 4.2.2, a reference heart sound is extracted using the aforementioned procedure. Otherwise, the window is shifted 1 s forward and the process is repeated until the periodicity conditions are met.

4.3. Phase II: Non-cardiac Sound Detection

The previous subsections introduced the preparation phase which is required to detect the reference heart sound. Once the reference HS has been defined a template-matching approach is applied using the following spectral and temporal features.

4.3.1. Spectral Energy

In this step, the spectral root mean square of the heart sound is calculated, i.e.

$$S_{rms}(f) = \sqrt{\int_{t}^{t+T_h} |S(f,t)|^2 dt} \tag{4.7}$$

where T_h is the length of the window of the heart which has the same duration as the reference sound (duration of the estimated heart cycle). Root mean square of the spectrogram provides an estimate of the power distribution in the frequency domain. Let $S_{rms}^{ref}(f)$ and $S_{rms}^{test}(f)$ be the spectral root mean square for the reference and the test HS signals respectively, then validation is performed using the following condition:

(4.8)

$$CorrCoef\left(S_{rms}^{ref}(f),S_{rms}^{test}(f)\right) > th$$

where *CorrCoef* is the correlation coefficient between two signals, and threshold th value was obtain using the ROC plot by varying it from 0.6 to 1 which was optimally found as 0.98 (see **Figure 4.9**). A heart sound segment is assessed as noisy if *CorrCoef* goes below th in (4.8). An example of the effects of (4.8) in assessing HS contamination by several sources is depicted in **Figure 4.8**(b).

4.3.2. Temporal Energy

Temporal energy is taken into account in order to determine short duration sound spikes. It is seen that spectral energy correlation is incapable to capture these types of noises with the defined correlation criterion in (4.8), see in **Figure 4.8**(c). However, it can be efficiently tackled with the temporal energy of the sounds.

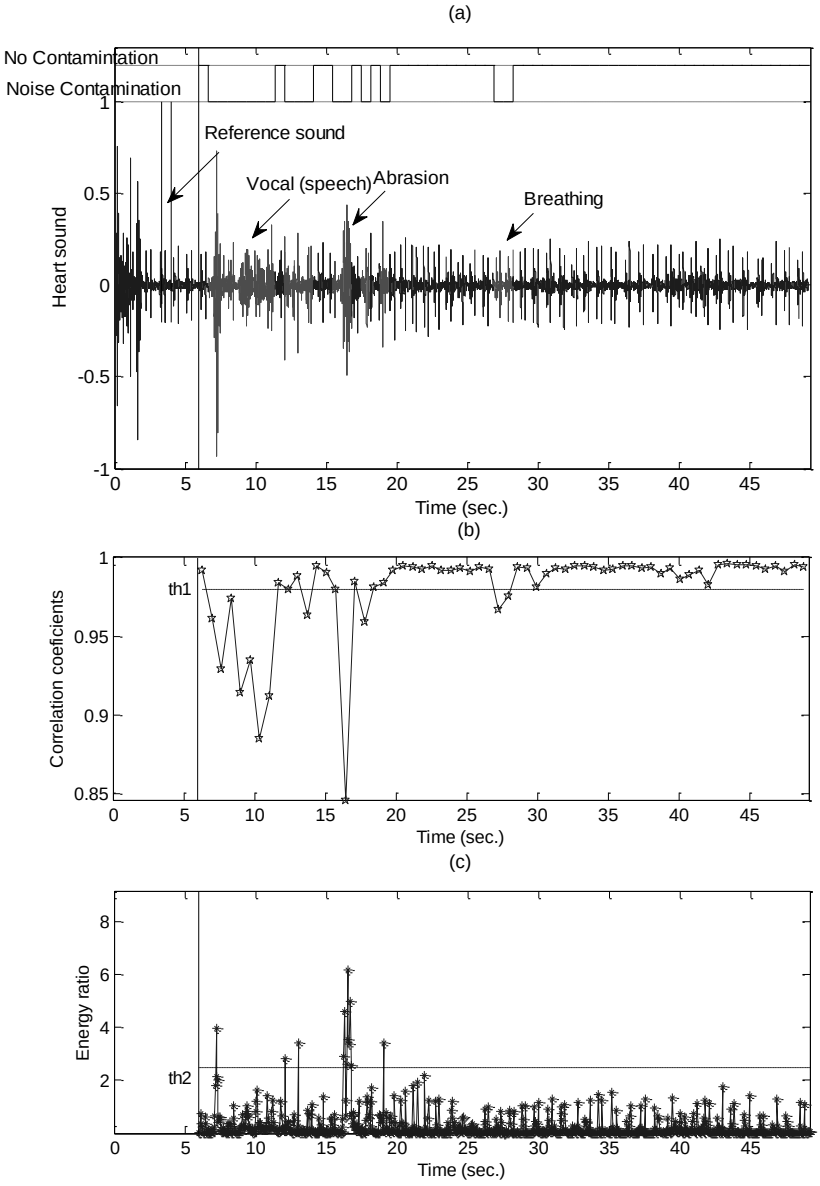


Figure 4.8. (a) Heart sound contaminated with several sources of noises. First 6 seconds correspond to reference template search window. The upper part of the graph shows noise classification result. (b) Spectral energy correlation between heart sound segments and the reference sound. (c) Relative temporal energy of the heart sound segments.

In order to obtain significant information about the instantaneous amplitude of noncardiac sounds of a short duration, it is important to compute the energy for small windows of the test signal. In our experiments we use window of duration t_w=50ms. The instantaneous energy is compared with the maximum energy of the reference sound computed in successive t_w window according to (4.9). Let $x^{ref}(t)$ be the reference signal and $x^{test}(t)$ be the sound under assessment. The relative temporal energy (RTE) of the l^{th} segment under analysis is defined in (4.10).

$$TE^{ref} = \max_{l=1,2,3.....n_p} \int_{(l-1)t_w}^{lt_w} x^{ref}(t)^2 dt \tag{4.9}$$

$$RTE_l = \frac{\int_{(l-1)t_w}^{lt_w} x^{test}(t)^2 dt}{TE^{ref}} \tag{4.10}$$

$$x^{test}(t) = \begin{cases} noise & , \exists l \in \{1,2,3,4,\ldots..n_p\} : RTE_l > th \\ heart\ sound & , otherwise \end{cases} \tag{4.11}$$

In (4.10) t_w is the window duration applied to compute the signal's energy, while $n_p = [\frac{T_h}{t_w}]$,, $l = 1,2,3,\ldots,n_p$, is the total is the total number of sound segments in the test sound. A sound segment of duration T_h is identified as contaminated with spike or impulse sounds of short duration, as formulated in (4.11), if its corresponding RTE exceeds the threshold th, where the threshold is set to 3-fold TE^{ref}.

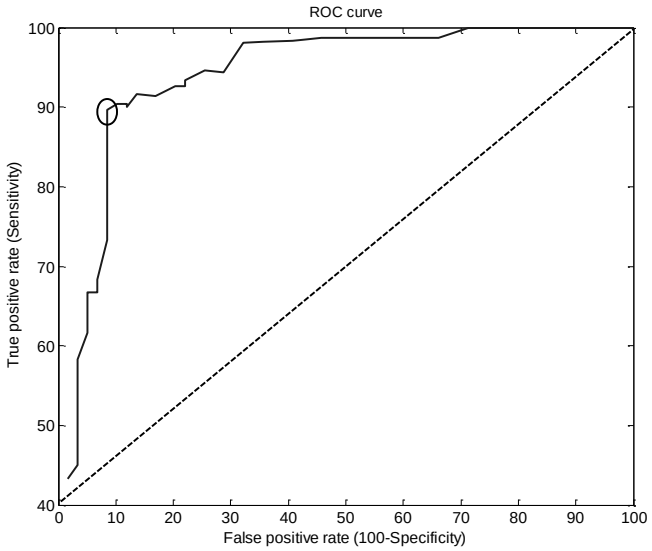


Figure 4.9. ROC curve for 13 heart sound samples to identify optimum spectral energy threshold as well as temporal energy (circled kink represents optimum value).

4.4. Test Database

The prepared data set includes 115 HS clips of recording length between 1 and 2 min. These heart sound clips have been collected from 71 different subjects at rest with the following biometric characteristics (average $\pm$ standard deviation): age = 35.26 $\pm$ 12.02 years; BMI = 25.11 $\pm$ 7.8kg/m^2; male subjects = 64; female subjects = 7. Regarding heart dysfunction, three subjects exhibited arrhythmia, 31 subjects had an artificial valve implant, and eight patients exhibited heart murmurs and the rest were healthy subjects.

Categories of Noises	**Noise Type**
Vocal	Speech at different pitches Coughing Laughing

Physiological	Breathing
	Swallowing
	Muscle movement
Sensor	Rubbing
	Movement
Ambient	Knocking door
	Noises induced by moving/falling of small objects
	Music
	Electric table fan
	Phone ringing
	Foot steps

Table 4.1 Induced noise sources.

In order to validate the performance of the algorithm, several types of noises were intentionally induced during the heart sound acquisition protocol. According to Table 4.1, in the prepared database, non-cardiac sounds were arranged into four classes: vocal sounds, physiological sounds (such as breathing, coughing, etc.), skin, rubbing or abrasion sounds via stethoscope (sensor based), and the other ambient noises. Contaminated heart sounds have been applied during the tests of both phases of the algorithm. The collection of heart sounds applied for the validation of the second phase of the method was composed by 4781 uncontaminated heart sound beats and 1659 contaminated heart beats distributed as follows: 805 vocal, 241 sensor, 161 physiological and 452 ambient noise segments. It should be mentioned that noise segments were manually annotated under the supervision of a clinical expert. Furthermore, noise contaminations induced accidently during the acquisition protocol were also annotated and included in the test database.

4.5. Performance Assessment

The performance of the algorithm is computed in the form of sensitivity (SE) and specificity (SP) measures, i.e.

$$SE\ (\%) = \frac{TP}{TP + FN}\ ;\ SE\ (\%) = \frac{TP}{TP + FN} \tag{4.12}$$

where TP denotes the number of noisy segments correctly detected as noisy segment, TN represents clean HS segments correctly detected as clean HS segments, FP stands for clean HS segments incorrectly detected as noisy segment, and FN denotes the number of noisy segments detected as clean heart sound segments.

In order to evaluate the robustness of the proposed algorithm, two types of tests have been performed: in the first test the algorithm has been applied directly to the test database; in the second test set a simulation study has been performed in order to evaluate its sensitivity with respect to noise the intensity level. In the first test, noisy segments are divided based upon their loudness into low, medium and high intensity noise classes. The loudness is defined herein as the average power of the signal in decibel, i.e. let $x(t)$ be the heart sound mixed with noise of duration T; then the loudness can be described in (4.13):

$$Loudness\ (dB) = 20\ log_{10}\sqrt{\frac{\int_0^T x^2(t)dt}{T}} \tag{4.13}$$

In the second test set performed, uncontaminated HSs have been artificially mixed additively with several types of noises with modulated intensities. Namely, let $v(t)$ be the noisy sound in the simulation the algorithm has been applied to $x(t) + Kv(t), K = 0, 0.5, 1, 1.5, \ldots, 4$. For $v(t)$ noise sources of the same categories as the ones described in section 0 have been applied. Furthermore, the signal-to-noise ratio (SNR) is computed according to (4.14):

$$SNR(dB) = 20\ log_{10}\frac{\int_t^{t+T} x(t)^2 dt}{\int_t^{t+T} K^2 v(t)^2 dt} \tag{4.14}$$

4.6. Experiment Results and Discussions

The first phase of the algorithm is related to the estimation of the reference HS. In the applied database, the proposed methodology was always able to detect a reference HS. On average, the reference HS was adequately detected during the first 11.4

s (between 4 and 25 s) of each sound clip. Regarding the processing time complexity of the algorithm, it is observed that the first phase of the algorithm is significantly more time consuming than the second one. The later one might be performed in real time in

most processing platforms. In tests performed using an implementation in Matlab 7.6 running on Windows XP using an Intel®CoreTM2Duo processor at 2.53 MHz, it was observed that the first phase took on average 1.23 s per each 4 s window of sound, while in the second phase the algorithm only took on average 0.035 s per processing window (duration of one heart cycle).

In order to examine the noise detection performance of the algorithm, noisy sound segments have been divided into three classes based on their degree of loudness, i.e. low loudness (LL), medium loudness (ML) and high loudness (HL), and into four classes of noise types based on the origin of the noise source. The considered ranges of loudness in each class were: $HL > -7.5\,dB, -14\,dB \leq ML \leq -7.5, LL \leq -14\,dB$. Table 4.2 and Table 4.3 summarize the achieved detection performance results in terms of sensitivity and specificity in each of these classes of noises for distinct types of cardiac dysfunctions. As can be observed in Table 4.2 and Table 4.3, the overall achieved detection sensitivity and specificity are 95.88% and 97.56%, respectively. The achieved detection performance for each of the considered classes of noise contaminations is relatively homogeneous. Namely, the obtained results show that the proposed method enables the detection of vocal noises at different pitches with a sensitivity of 96.62%, regardless of cardiac dysfunction. Regarding noise contaminations originated by sensor movements, i.e. noise contaminations of class sensor induced noise, the method enables its detection with 96.43% sensitivity. Another highly probable incident during HS acquisition is contamination by physiological noises. As can be observed, these types of noises are detected with 92.94% sensitivity. This lower value in sensitivity is directly related to the fact that physiological noises tend to exhibit similar spectral content as regular heart sounds and, hence, are more difficult to capture using the spectral similarity feature. Regarding ambient sound contaminations, the algorithm exhibits 94.47% sensitivity. It should be noted that these classes of noises reflect the most likely noise contaminations during heart sound acquisition. Regarding the performance with respect to the cardiac dysfunction, Table 4.2 suggests that the algorithm exhibits comparable performance irrespective of heart dysfunction and valve type (implant or native).

Noise Type	*Heart Sound/Valve type*	*Sensitivity*			*Noise Sensitivity*	*Average Sensitivity*
		HL	*ML*	*LL*		
Vocal	Native valve	98.34	94.99	89.04	96.78	96.62
	Prosthetic valve	95.45	94.32	100	94.64	
	Murmur	97.06	96	100	96.74	
	Arrhythmia	100	100	100	100	
Sensor	Native valve	100	97.95	92	97.16	96.43
	Prosthetic valve	100	90	90.91	92.68	
	Murmur	100	100	50	90.91	
	Arrhythmia	100	100	100	100	
Physiological	Native valve	100	89.66	92.08	92.20	92.94
	Prosthetic valve	100	100	91.67	92.87	
	Murmur	100	100	100	100	
	Arrhythmia	NA*	NA*	NA*	NA*	
Ambient	Native valve	98.37	92.95	94.45	95.24	94.47
	Prosthetic valve	100	93.33	90	93.18	
	Murmur	100	100	100	100	
	Arrhythmia	100	80	66.67	77.78	
Overall Sensitivity						**95.88**

Table 4.2 Results of noise detection in terms of sensitivity (*SE* in %). (*NA = Not Available)

It should be noted that the results reported for arrhythmia are not significant, since only three subjects in the database exhibited arrhythmia. As one would expect, the noise detection performance deteriorates as the noise loudness decreases. Using the test database, it is seen that the algorithm exhibits deterioration less than 10% in sensitivity when noise intensity decreases from HL to LL, irrespective of the noise type and cardiac dysfunction. As can be observed in some of the reported results that sometimes algorithm seems to perform better for low loudness noise perturbations than for medium loudness noises. This is mainly due to the way the noise intensity is assessed (4.13). The average loudness is not able to adequately capture the instantaneous intensity level of short duration noises and, therefore, it is observed that some reasonably loud short duration noises might be classified in the LL class, leading to the aforementioned effects.

Table 4.3 summarizes the specificity results achieved by the algorithm for dis-

tinct heart dysfunctions. As can be observed, the obtained specificities tend to exhibit resilience with respect to the characteristics of the HS source. Given the high values of SP experienced under the method, a very low number of false positives are expected, i.e. the method tends to exhibit a very small number of regular HS segments that are classified as noisy segments.

Heart sound	***Specificity***
Native valve	97.24
Prosthetic valve	97.61
Murmur	98.17
Arrhythmia	99.47
Overall specificity	**97.56**

Table 4.3 Results of noise detection in terms of specificity (SP in %), computed based upon false positive (FP).

In order to analyze the noise intensity implications in the algorithm's performance, a simulation study was setup using an additive noise contamination model with amplitude modulation. In this study, clean heart sounds acquired from subjects with native heart valves, prosthetic heart valves, murmur and arrhythmia were additively mixed with noise contaminations of the considered noise classes, i.e. voice, sensor, physiological and ambient noises. During this process, the SNR of the heart sound segments mixed with noise sources were varied using a linear gain. The achieved results are reported in **Figure 4.10**. As can be noticed that regardless of the type of heart dysfunction and noise contamination the sensitivity of the algorithm declines as the SNR increases. For negative SNR values, i.e. situations where the noise intensity exceeds the signal energy, the sensitivity of the method is relatively stable and very high with SE values near 100%. At 0 dB SNR, a situation where the noise exhibits the same energy as the signal, it is observed that the performance of the algorithm exhibits a SE over 85%–90%. As one would expect, since the energy of the noise decreases below the energy of the signal (positive SNR region) the number of false negatives increases rapidly, leading to a drop in sensitivity. However, it should be mentioned that even for very low energy noise contaminations, the algorithm exhibits a significant behavior. For instance, for 20 dB noise contamination, i.e. situations where the noise intensity is 10% of the energy of the signal, the simulation results suggest that the method exhibits SE values higher

than 80% for most noise contamination and heart sound types. Regarding the performance of the algorithm with respect to the contamination type, it should be observed that the detection sensitivity tends to be significantly more consistent and higher for larger ranges of SNR in noise contaminations of ambient noises and vocal noises. These types of noises tend to exhibit significant content in higher frequencies (e.g. vocal sounds are typically around 2–4 kHz bandwidth) and, therefore, exhibit spectral signatures that are easier to be captured by the spectral similarity test. Using the additive simulation study, it is seen that the performance of the algorithm tends to drop more rapidly for physiological and sensor noise contaminations compared to the other groups of noises.

This is related to different reasons; physiological noise sources tend to exhibit similar spectral content range compared to heart sound; hence, their spectral power signatures are closer to heart sound spectral power signatures. As for sensor noise sources, it is observed that these are usually characterized by short durations, leading to less well-defined spectral signatures. In this type of contaminations the temporal energy feature plays a central role.

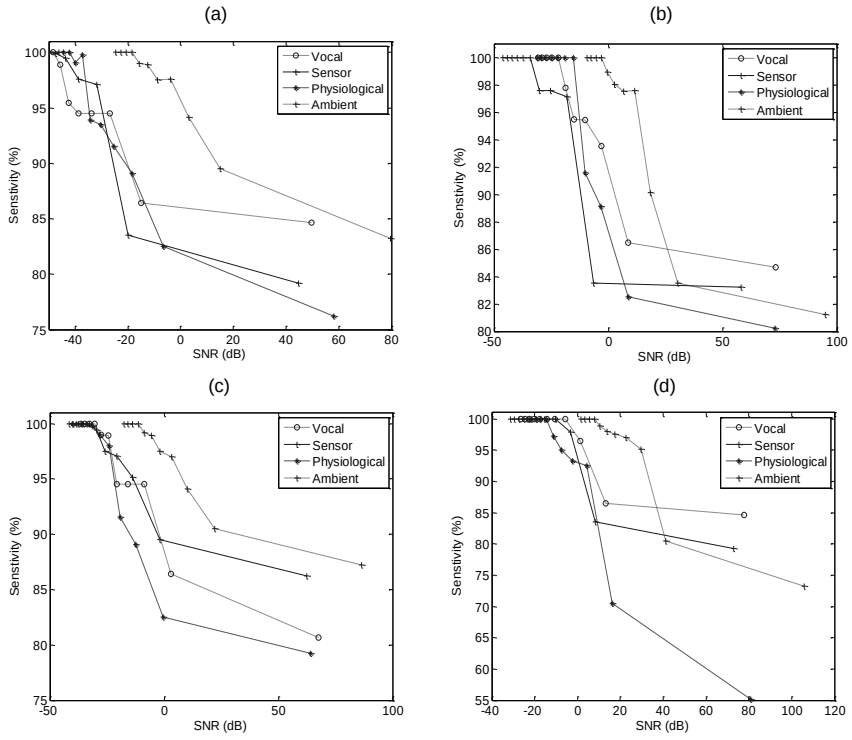


Figure 4.10. Sensitivity for noise detection in varying intensities of noise of the (a) heart sound from native valves, (b) prosthetic valve, (c) murmur, and (d) ar-

rhythmia.

4.7. Conclusions

In this chapter, an algorithm has been proposed that enables the detection of heart sound contaminations by several sources, irrespective of their spectral and intensity characteristics and interference model. The method is based on a template-based strategy composed by two main phases: in a first phase, the algorithm detects a reference heart sound with the duration of one heart cycle. This enables resilience with respect to site of collection, subject and sensor positioning as well as subject specific morphological characteristics and sound transmission paths. In a second phase, the detected reference heart sound template is applied in real time to detect deviations in the heart sound pattern. The first phase is based on two periodicity characteristics of uncontaminated heart sound: heart sound is a quasi-stationary signal exhibiting periodicity in time and it is also observed that most of its energy is concentrated in its S1 and S2 components which manifests itself in adjacent frequency bands repetitions of the same energy pattern, leading to a periodicity pattern in frequency bands. These two observations are applied to identify an uncontaminated heart sound heart cycle. It should be mentioned that this processing step is the most demanding one from the computational load perspective. Once this template has been identified, the algorithm proceeds in detecting non-cardiac sounds by checking the template against the acquired sound. Herein, two computationally simple features are extracted to analyze the similarity with respect to the template. Simulation tests using Matlab suggest that this phase of processing can be performed in real time enabling the algorithm to be used to provide feedback information to the user regarding the quality of the signal.

The method was exhaustively tested using a database of heart sound collected from 71 subjects with different cardiac dysfunctions and noise contaminations. The database has a total of 157 min of collected heart sounds with clean heart sounds as well as heart sounds that have been contaminated with some of the most probable noise sources during real life heart sound acquisitions at several intensity levels. Using this database, the algorithm achieved a sensitivity of 95.88% and a specificity of 97.56%. These high sensitivity and specificity together with its real time operation, makes this algorithm an interesting solution to deploy heart sound based biomedical systems, particularly in pHealth scenarios where feedback information has to be provided in real time to the user in order to allow us to timely alert and guide the user in solving the signal acquisition problems and to avoid false positives.

Chapter 5

HEART SOUND SEGMENTATION

The main purpose of this chapter is to investigate heart sound properties in order to develop methods to identify components in a heart sound recoding. Localizing sounds and subsequently segmenting them is the second main task in the heart sound analysis. In fact, heart sound identification consists of two tasks: *localization* and *segmentation*. The former is related to the recognition process of heart sounds, whereas the latter includes *delimitation* of the heart sound components and their classification into meaningful clinical categories, such as S1 or S2. Delimitation allows gating the sounds as segments by marking their onset and offset instants. Recognizing these segments requires physiological information with respect to heart sounds' constituents. In this chapter, we mainly focus on the segmentation part of the complete sounds' identification process, since localization of the heart sounds' components in a noisy heart sound recording has been dealt with in the previous chapter.

Generally, localization of the main heart sounds, i.e. S1 and S2, with the indication received from a supporting signal has been a popular and reliable method. However, there are many circumstances where it is advantageous to segment the heart sound directly without any auxiliary signal, as was explained in section 3.3. Because of the overlapping of murmurs with the main sounds, heart sound segmentation can be challenging without using a synchronized auxiliary signal. In order to be able to coupe with the different classes of heart sound in an effective way, we propose a segmentation strategy that utilizes specialized methods that are inspired by the characteristics of the heart sounds produced under different cardiac conditions, such as, valve insufficiency, septal defects or prosthetic valves. In the proposed approach, a selection mechanism selects the most appropriate method for a

given heart sound using a complexity analysis approach. For the most challenging heart sound segmentation problem class, we introduce a complexity-based method. Finally, for the heart sound component recognition phase, we propose a physiology-motivated feature for the recognition.

This chapter is organized in the sections outlined below.

Section 5.1 Introduction

We start this chapter by introducing the heart sound segmentation problem. Then, the previous methods are briefly explained, and in the end of the section of our approaches are outlined.

Section 5.1 Heart Sound Components and Characteristics

In this section, how heart sound's components and their properties can be used in the segmentation is explained. Although, the main segmenting sounds are S1 and S2, the properties of the third and fourth sounds are also discussed.

Section 5.23 Heart Sound Segmentation Scheme

The developed segmentation techniques are introduced in this section. A roadmap to the segmentation methods is drawn by outlining the main steps: classification of the heart sound produced from different cardiac sources, and segmentation techniques that are adequate to each class of sounds.

Section 5.24 Heart Sound Classification Using Chaos

In our scheme, two segmentation methods are proposed in which one is aimed to be applied for heart sounds without murmurs and the other one is for heart sounds with murmurs. Therefore, a mechanism to isolate one type of heart sounds from others is required. To develop this, the heart is assumed as a nonlinear dynamical system whose behavior can be known through its time series (in this case heart sound). Hence the mathematical background to show the heart's dynamics via heart sounds produced as a result of its function is described in this section.

Next, a feature corresponding to the nonlinear system, i.e. chaos, from the samples of the heart sounds is computed to determine the type of heart sounds. A rule based criterion is defined to realize this task.

Section 5.2 Preprocessing for Segmentation Algorithms

Before starting formulating the algorithms on heart sounds' segmentation,

some preprocessing steps required in the preparation of the heart sounds are presented in this section.

Section 5.5 Segmentation of Heart Sounds without Murmur: Energy Envelop Method

In this section, our approach based on the envelope of the of the heart sound generated from Shannon energy is presented. This method is suitable to segment those heart sound recordings which do not have any additional sounds (e.g., S3, S4 or murmurs) besides S1 and S2.

Under segmentation, first the approach for identifying onsets and offsets of the sounds in the heart sound recording is illustrated. Afterwards, the mechanism applied for recognition of the heart sound component based on high frequency signature of S2 sounds is explained.

Section 5.6 Segmentation of Heart Sounds with Murmur: Wavelet Decomposition-Nonlinear Features (WD-NLF) based Method

Heart sounds with additional sounds, namely murmurs, are segmented with the approach introduced in this section. The WD-NLF approach consists of four main steps: wavelet decomposition of a heart sound, nonlinear feature extraction from the wavelet coefficients using nonlinear dynamical method, thresholding, and back to the heart sound in the wavelet reconstruction process. This works in an iterative way where suitable depth of decomposition is determined.

The delimitation part is executed with the introduced approach of WD-NLF, however the sound recognition part contains major steps which involve the energy envelop method already explained in a previous section.

Section 5.68 Test Database

This section contains information regarding the heart sounds collected from three different sources.

Section 5.89 Performance Measurements

In this section, we define the measures which are applied to evaluate segmentation performance. Timing errors in detecting onsets and offsets of the heart sounds, and the classification of heart sounds' type as well as sounds' recognition using sensitivity and specificity are defined.

Section 5.9 Experimental Results and Discussions

Experimental results are presented and examined in this section. Besides the results achieved from the heart sounds classification, delimitation error and sounds' recognition are obtained in two additional cases: first when all heart sounds with murmurs are segmented via WD-NLF method and all heart sounds without murmur via energy envelop method, respectively, and second when the wrong segmentation algorithm is systematically applied.

Figure 5.1. Main constituent blocks of heart sound segmentation. *Delimitation* determines onsets and offsets of the sounds which are addressed as segments or gates, and *Recognition* identifies the types of sounds.

5.1. Introduction

As can be observed in **Figure 5.1** a typical and regular heart cycle is composed of S1-systole-S2-diastole. Since S3 and S4 segmentation in not in the scope of this thesis it can be stated here that segmenting the main heart sound components of the heart sounds would divide every heart cycle in the sequence of S1-systole-S2-diastole. Systolic or diastolic areas can exhibit presence of murmurs in cases of cardiac abnormalities, namely in valvular and coronary dysfunctions (Wanak 1972). As can be observed in **Figure 5.1**, heart murmurs are complementary to S1 and S2 sounds. Hence, if the delimitation of S1 and S2 sounds is known, then murmurs can also be identified.

Heart sound segmentation is a problem that has received some attention from the research community for over two decades. Ideally, heart sound segmentation can include the task of noise elimination or suppression in the heart sound before analysis is carried forward, but in this paper we only focus on the segmentation part. The problem of noise detection has been tackled in the previous chapter. Therefore, we focus only on the delimitation and recognition parts of the process of segmentation, as depicted in **Figure 5.1**. The existing methods for S1 and S2 segmentation can be categorized into two main categories of methods: i) methods re-

quiring a reference signal such as the ECG or the carotid pulse to gate the heart sound; ii) and methods that do not require any kind of reference signal (see some methods of each category in Section 3.2).

The approaches presented in Section 3.2 can exhibit some limitations with respect to the requirements laid down:

- The approaches requiring a reference signal, such as the ECG or the carotid pulse (Carvalho et al. 2005; El-Segaier et al. 2005; Lehner and Rangayyan 1987; Malarvili et al. 2003), require another sensor to acquire the reference signal which causes the use of extra hardware, power, wringing setup and cost. As a result these approaches tend to become cumbersome and expensive to use.
- In the category of direct methods, i.e., segmentation without using a reference signal, a feature curve is constructed by choosing a computation window and, subsequently, thresholding the feature curve. In this type of approaches the features are generated via these measures: envelogram (Huiying and Iiro 1998; Liang et al. 1997), entropy (Yadollahi and Moussavi 2006), instantaneous energy (Sharif et al. 2000), complexity using various methods of fractal dimension (Carvalho et al. 2005; Gnitecki and Moussavi 2003; Hasfjord 2004) , as well as some suggested methods based on the approximated entropy and spectral entropy (Rezek and Roberts 1998). The main limitation of these methods is their incapability to enhance the distinction between murmurs and low frequency high amplitude sounds, S1 and S2. Threholding is only efficient when S1 and S2 sounds show a significant difference from the murmurs.
- The second line of approaches under the direct methods is the generation of multiple features from each sound and training of a classifier (Ricke et al. 2005; Schmidt et al. 2010; Wanak 1972). Since these approaches involve the incorporation of a supervised learning based classifier which is constructed based on some known inputs and outputs, this necessarily limits the scope of adaptability of the methods.

We seek to develop a methodology for the segmentation of the heart's main sounds regardless of murmur presence, and at the same time, avoid the need for any extra reference signal as well as supervised learning. The problem with supervised lerning based approaches is that they are usually patient specific. Therefore, our proposed approach is based on chaos assessment using Lyapunov exponents. This approach enables the discrimination of heart murmur sounds from normal heart sounds. Segmentation of heart sounds without murmur is a relatively straightforward problem. We propose an algorithm based on the analysis of the signal's envel-

op. On the other hand, segmentation of heart sounds contaminated with murmur is a highly challenging task. In order to tackle this issue, we propose adaptive algorithms based on wavelet decomposition-nonlinear features (WD-NLF). These nonlinear features are simplicity and strength, extracted by applying nonlinear dynamical theory on the heart sounds. This method uses the complexity of the signal to identify candidate signal segments for S1 and S2 sounds. Discrimination of the S1 and S2 sounds is based on a physiologically-motivated high-frequency marker.

5.2. Heart Sound Segmentation Scheme

Heart sound segmentation and recognition of its main constituent parts is approached using two distinct methods: one is based on the signal's envelogram (for sounds without murmur), while the other is based on a wavelet-simplicity filter (for sounds containing murmur). The envelogram based algorithm is computationally efficient compared to the wavelet-simplicity-based approach. However, its performance degrades rapidly for HS with murmur. These abnormal sounds are segmented using an algorithm based on the wavelet-nonlinear feature filter, which is computationally more demanding. To automatically select between both methods, a selection stage has been incorporated into the segmentation module, as illustrated in **Figure 5.2**.

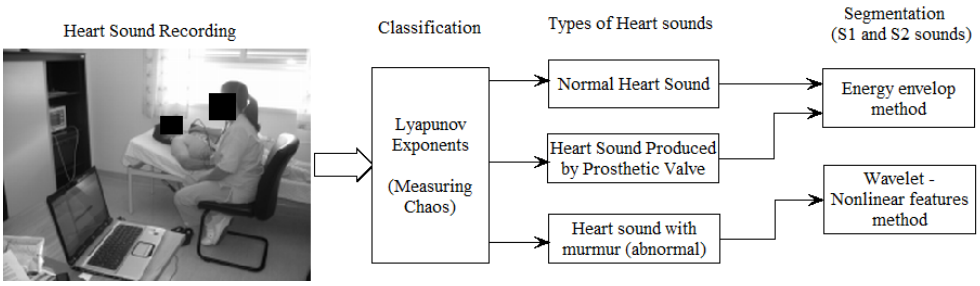


Figure 5.2. Heart sound assessments and segmentation scheme.

Heart sounds with murmur contain audible sounds in the systolic or diastolic regions of the heart cycle, making its separation from the main heart sound components difficult. Hence, sounds containing murmur should be first identified. In the current approach, this is accomplished by resorting to the degree of chaos in the signal.

In most frequency-based signal transform techniques, heart sound is assumed to be linear and Gaussian. On the other hand, in the nonlinear dynamical theory, the signal (time series) is assumed to be non-linear and non-Gaussian. In practice, heart

sound is characterized as a nonlinear, non-Gaussian, transient signal (Ahlstrom 2006b; Taplidou and Hadjileontiadis 2006). Hence, the non-linear dynamical system theory can be used to capture non-linear behavior from the heart, which could be reflected in certain nonlinear features able to discriminate between clean heart sound and heart sound contaminated with murmur. If the heart is assumed as a non-linear dynamical system that produces chaotic deterministic time series of heart sound, then the behavior of the heart can be portrayed in the state or phase spaces. From there, important attributes of heart sound, such as strength and simplicity, can be extracted. These are capable of distinguishing clean sound from contaminated sound, as it will be discussed later.

If abnormal murmurs are detected, the heart sound is segmented using the WD-NLF method. It will be seen later in the paper that heart sounds with murmurs exhibit higher chaos than normal heart sounds, but lower chaos than heart sounds induced by prosthetic valves.

5.3. Heart Sound Classification Using Chaos

5.3.1. Heart Sound Embedding

We start by assuming that the human heart is a dynamical system that produces heart sound $x(t)$. This dynamical system can be characterized as $X(t+1) = F(X(t))$, where $X(t)$ denotes the state of a system at time t. The solutions at different time instants can be regarded as points in the system phase space. F denotes a nonlinear operator acting over these points in the phase space. Hence, the closed-form solution of a dynamical system can be written as $X(t) = V(t, X(0))$, where $X(0)$ is the initial value of the state and V is a state transition function. Since the signal produced by the dynamical system is a time series, which is a function of the hidden state $X(t)$ in order to study the dynamical behavior of the system, its phase space needs to be reconstructed. In this task, it is assumed that the signal $x(t)$ is a one-dimensional (1D) projection of $X(t)$, resulting from employing the operator U, i.e., $x(t) = U[X(t)]$. A sequence of $X(t)$ over time is called a trajectory of the dynamical system. The developed trajectory reconstruction method, i.e., the reconstruction of the sequence of states $X(t)$, is based on Taken's theorem; according to that, it is possible to reconstruct a new state space, diffeomorphically (topologically) equivalent to the original state space, using scalar time series (Abarbanel 1996). Later, based on the reconstructed states of the time series, the dynamics of the system can be observed.

Let $\mathbf{X}$ be a matrix denoting the reconstructed trajectory, where each row $\mathbf{X_i}$ is the state of the system, or phase-space vector, at discrete time i, as in (5.1).

$$\mathbf{X} = [\mathbf{X_1}, \mathbf{X_2}, \mathbf{X_3}, \dots\dots\dots\dots\dots\dots, \mathbf{X_M}]^{\mathbf{T}} \tag{5.1}$$

Two methods are available in literature to reconstruct the state space: delay coordinates and derivative coordinates. Since derivatives are susceptible to noise, the method of delay coordinates is adopted. Using this method, from an N-point time series $\{x_1,\ x_2,\ x_3, \dots\dots\dots, x_N\}$ each phase-space vector $\mathbf{X_i}$, can be reconstructed as follows

$$\mathbf{X}_i = [x_i, x_{i+\tau}, x_{i+2\tau}, \dots\dots\dots\dots x_{i+(m-1)\tau}] \tag{5.2}$$

where τ is the reconstruction delay and m the embedding dimension. The trajectory, or embedded matrix $\mathbf{X}$, then an $M \times m$ matrix, where $M = N - (m-1)\tau$.

In the parameter estimation, the τ parameter is chosen before the dimension m; it is dependent on correlation between $x(t)$ and $x(t+\tau)$, a small τ is equivalent to large correlation, whereas on the other hand a large one can result in complete independence between two vectors which may lead to insignificance of embedded matrix. Hence, an optimum τ is computed based on mutual information between $x(t)$ and $x(t+\tau)$ as in (5.3)

$$I(\tau) = \sum_{n=1}^{N-\tau} p(x(t), x(t+\tau)) log_2 \frac{p(x(t),x(t+\tau))}{p(x(t))p(x(t+\tau))} \tag{5.3}$$

where $p(.)$ is probability and $I(\tau)$ is the mutual information between the samples of original time series $x(t)$ and translated time series $x(t+\tau)$. The compromised value of parameter τ time lag is calculated at the first minimum that occurs in mutual information $I(\tau)$. The optimum value for the time lag τ has to be large enough for $x(t)$ and $x(t+\tau)$ to yield minimum overlapping information without being so large that number of data points in the signals compared becomes too small. Heart sound signal is limited to N samples (data points) and for a given delay lag τ, $x(t)$ and $x(t+\tau)$ each vector contains N-τ samples. If the values of τ is too large, then the size of vectors of τ is too large, then the size of vectors $x(t)$ and $x(t+\tau)$ becomes very small and some important information may be lost. Varying the value of τ yields mutual information coefficients, as in (5.3). The optimum lag is the first

local minimum on the plot Information coefficients verses lags (τ), see **Figure 5.3**(a).

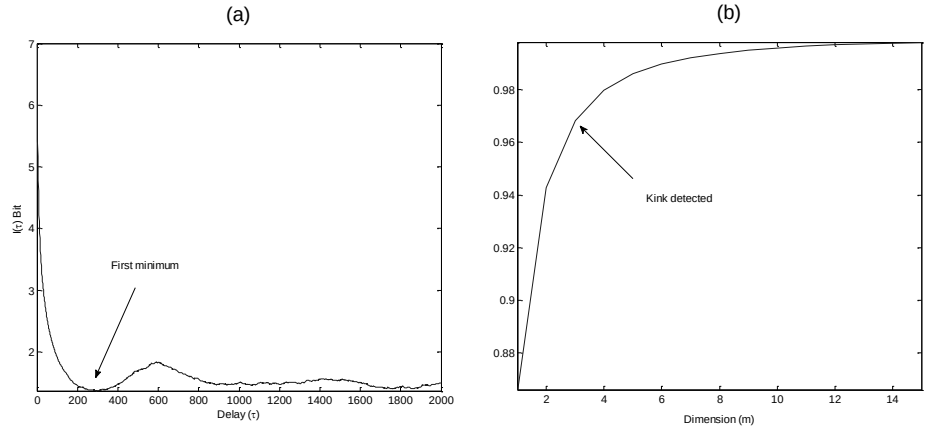


Figure 5.3. Reconstruction parameters computation: (a) τ =195 first minimum in mutual information curve; (b) m=4 finding first kick in the prediction error using Cao's method (Cao 1997).

After acquiring the delay parameter τ, the phase dimension m is estimated by the well known Cao's method (Cao 1997). The method has an advantage over the traditional false nearest neighbor method and exhibits invariance with respect to the data length; detailed mathematical elaboration can be found in (Kennel et al. 1991). According to the Cao's method if in (5.4) m is the true embedding dimension of the reconstructed vector of $\mathbf{X_i}$ heart sound $x(t)$, then the two points that are closer in the m dimensional phase space will remain closer in the m+1 dimensional phase. Hence, any two points that satisfy this condition are regarded as dimensional phase space and therefore will remain close in the true neighbors (Cao 1997). The method is applied repeatedly starting with a low value of the embedding dimension m and then increasing it until the number of false neighbor decreases to zero, or equivalently, in Cao's method embedding function defined as in (5.4).

$$E(m) = \frac{1}{N - \tau m} \sum_{i=1}^{N-m\tau} a_i(m) \tag{5.4}$$

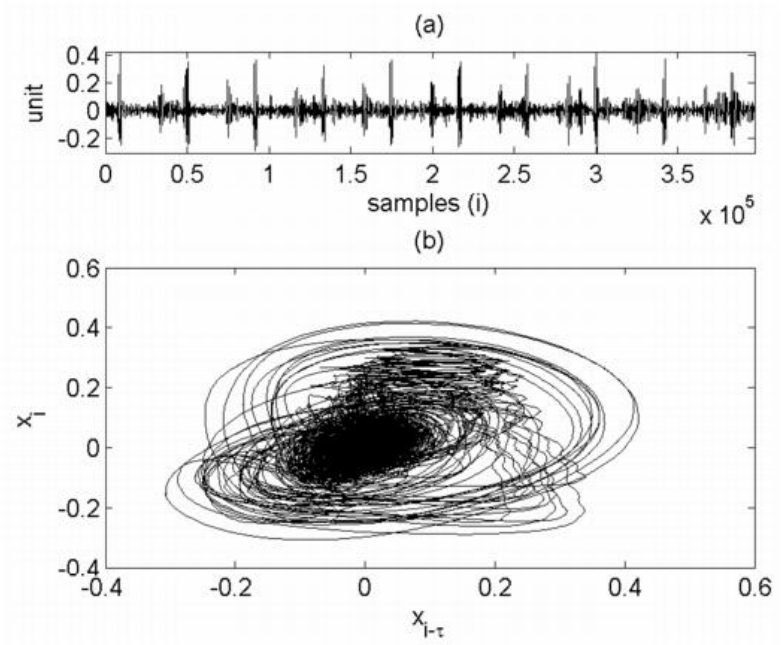


Figure 5.4. (a) Heart sound (normal). (b) Phase space reconstructed using delay parameter (τ) 185 and dimension (m) 4.

where $a_i(m)$ is the defined in (5.5)

$$a_i(m) = \frac{\left\| Y_i(m+1) - Y_{n(i,m)}(m+1) \right\|}{\left\| Y_i(m) - Y_{n(i,m)}(m) \right\|} \tag{5.5}$$

where $Y_{n(i,m)}(m)$ is the nearest neighbor of $Y_i(m)$ in the m-dimensional phase space. Regarding the proximity of two neighbors, the choice of the nearest neighbor is based on a measure of distance computed using the maximum norm function denoted by the $\|.\|$ symbol. The embedding function is then modified to model the variation from *m* to *m+1* by defining another function, $E1(m) = E(m+1)/E(m)$, which converges to 1 in the case of a finite dimensional attractor. Minimum embedding dimension is identified from the plot of $E1(m)$ verses m as the value of m at which the value of $E1(m)$ approaches 1, as can be observed Figure 4(b). However, in certain cases this may occur even with random signals. In order to distinguish deterministic data from random signals another function is defined as $E2^*(m) = E^*(m+1)E^*(m)$ is defined as in (5.6).

$$E^* = \frac{1}{N - m\tau} \sum_{i=1}^{N-m\tau} \left| x_{i+m\tau} - x_{n(i,m)+m\tau} \right| \tag{5.6}$$

In (5.6), $x_{i+m\tau} - x_{n(i,m)+m\tau}$ in the neighbor of $x_{i+m\tau}$. From an examination of the graph of $E2(m)$ versus m, a constant value of one from $E2(m)$ from different values of m indicates a random signal. The signal is found to be deterministic if the value of $E2(m)$ is not equal to 1 for at least one value of m. The embedding dimension for phase space is set to a minimum embedding dimension. An example of reconstructed phase display is shown in **Figure 5.4**.

5.3.2. Largest Lyapunov Exponents for Heart Sound Type Detection

Lyapunov exponents measure the degree of chaos in the time series acquired from dynamical systems (Abarbanel 1996). In time series produced by dynamical systems, as is the case of heart sound, according to our previous assumptions, positive characteristic exponents indicate chaos. Moreover, it is known that the calculation of the largest Lyapunov exponent (LLE) is sufficient for many applications (Rosenstein et al. 1993). In this way, our goal is to calculate the LLEs throughout time, as a basis for discriminating between heart sounds with or without murmur.

After reconstructing the system dynamics, the Lyapunov exponents can be computed from the phase space using the numerical method described in (Rosenstein et al. 1993). To calculate those exponents, the algorithm starts by locating the nearest neighbor of each point in the trajectory. Thus, for each reference point $\mathbf{X_j}$, the corresponding nearest neighbor, $\mathbf{X_j}$, is the one that minimizes the distance $d_j(0)$, according to (5.7):

$$d_j(0) = \min_{\mathbf{X_j}} \left\| \mathbf{X_j} - \mathbf{X_j} \right\| \tag{5.7}$$

In (5.7), $d_j(0)$ is the initial distance from the j^{th} point in the trajectory to its nearest neighbor, and $\|.\|$ denotes the Euclidean norm. Each pair of neighbors can thus be regarded as nearby initial conditions for different trajectories. Hence, at discrete time 1, the initial distance either expands or shrinks to $d_j(1)$ and successively over time i. The largest Lyapunov exponent is then estimated as the mean rate of separation of the nearest neighbors.

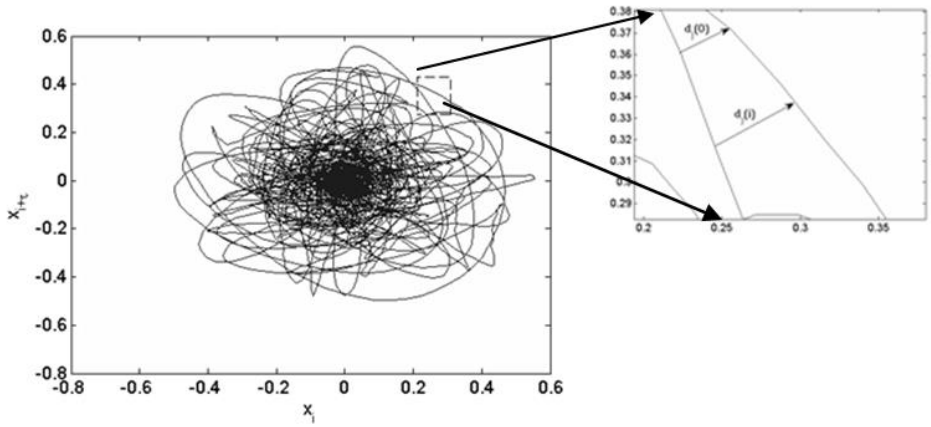


Figure 5.5. Reconstructed phase (m=5, τ=64) of heart sound with murmur (mitral regurgitation), and magnified section is the display of diverging trajectories.

To this end, it is assumed that the j^{th} pair of nearest neighbors either diverge or converge approximately at a rate given by the largest Lyapunov exponent, according to (5.8):

$$d_j(i) \approx C_j . e^{\lambda_1 (i.\Delta t)} \tag{5.8}$$

In (5.8), $C_j = d_j(0)$ is the initial separation between the j^{th} pair of nearest neighbors. By taking the logarithm of (5.8), a set of parallel lines is obtained, each with a slope roughly proportional to λ_1. Then, the largest Lyapunov exponent is calculated using a least-square fit to the "average" line defined by

$$y(i) = \frac{1}{\Delta t} \mu \left(\ln \left(d_j(i) \right) \right) \tag{5.9}$$

In (5.9), $\mu(.)$ is the mean or average operator computed over all j and y denotes the points on the divergence curve as a function of time i. The largest exponent is then estimated by calculating the slope of the divergence curve, as illustrated in **Figure 5.6**.

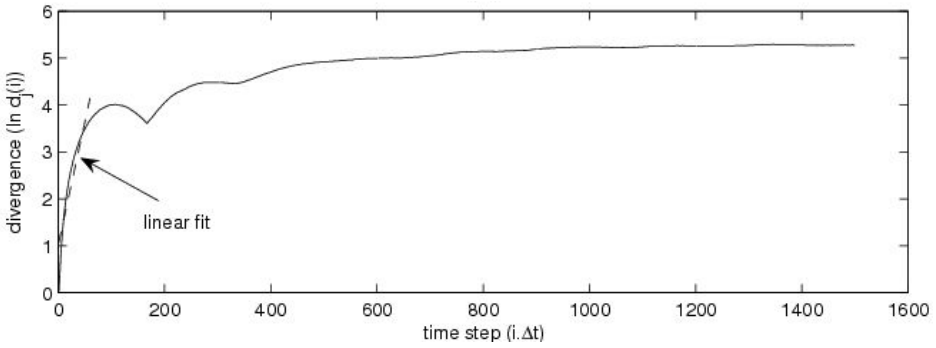


Figure 5.6. Largest Lyapunov exponent (LLE) estimation.

For a chaotic dynamical system the LLE is positive. Moreover, the larger the LLE, the more chaotic the time series is. In our heart sound database, it is observed that different types of heart sounds exhibit different ranges of LLE values. Namely, sounds from prosthetic valves are more chaotic than sounds mingled with murmur, which in turn are more chaotic than normal clean sounds. This is related to the fact that the bandwidth of prosthetic heart sounds is much higher (up to 20 kHz) compared to other heart sounds, which leads to an increased complexity. Hence, the LLE is utilized to distinguish between different classes of heart sound based on a thresholding approach, as in (5.10):

$$Heart\ Sound = \begin{cases} HS, & \lambda_1 \leq th_L \\ HSmurmur, & th_L < \lambda_1 \leq th_H \\ HSpros, & \lambda_1 > th_H \end{cases} \tag{5.10}$$

In (5.10), th_L and th_H are lower and higher thresholds, respectively. Moreover, *HS* denotes normal heart sound, $HSmurmur$ stands for heart sound with murmur and $HSpros$ represents heart sounds from prosthetic valves.

5.4. Preprocessing for Segmentation Algorithms

Heart sound segmentation mainly involves two important tasks: i) finding the start and stop instants of the sound components; ii) and identifying the main heart sound components, S1 and S2. Based on the complexity of the heart sound components, two segmentation algorithms are proposed. Normal heart sounds and heart sounds

produced by prosthetic valves are less complex than heart sounds with murmur. Therefore, as it has been previously stated, two segmentation algorithms are required, so as to find the best trade-off between computation cost and time. First, the algorithm for sounds without murmur, based on the Shannon energy envelop, will be described. Later, in the second algorithm, for sounds containing murmur, nonlinear features from the reconstructed phase space of the heart sound will be extracted. In both algorithms, wavelet decomposition is performed as a common initial step. Furthermore, a high frequency signature criterion, inspired by the heart functionality, is utilized to identify S1 and S2 sounds in both algorithms.

Wavelet decomposition (WD) provides mechanisms to analyze a signal in dyadic frequency bands, which aids to locate relevant heart sound components in a particular frequency band. Furthermore, it also suppresses the influence of stationary Gaussian noise in heart sounds, probably induced by recording equipment or ambient noise (Hossain and Mousavi 2003; Huiying and Iiro 1998). In this work, an orthonormal basis was chosen from the Daubechies family, namely *db6* basis functions. These basis functions were selected based on their morphological similarity with the temporal waveform of S1 and S2 sounds. Another advantage of *db6* is to show bandwidth efficiency with respect to samples frequencies over time, in comparison to other order of Daubechies wavelet functions and other members of the wavelet family.

In our method the approximate coefficients are used for obtaining the envelope of the sound for segmentation, whereas the detail coefficients are employed for extracting a high-frequency signature, useful in the recognition of the obtained segments, as discussed below.

5.5. Segmentation of Heart Sounds without Murmur: Energy Envelop Method (EE)

After wavelet decomposition using the *db6* basis functions, an envelope-based detector is used to detect the start and stop timings of the clinically meaningful main heart sounds. To this end, the Shannon energy is computed from the wavelet-decomposed heart sound. Since the energy envelopes of S1 and S2 heart sounds are significantly discriminative from the stationary part of the heart sound (in the systolic and diastolic regions), a simple adaptive threshold is set to identify the start and stop timings. For recognizing S1 and S2, a high frequency signature criterion is utilized. Using this criterion, one type of heart sound corresponding to a physiologically motivated high-frequency segment, usually S2, is recognized. Based on the periodic nature of heart sounds, locating one sound component along the entire

sound allows us to isolate single heart cycles. From each heart cycle, the other sound component corresponding to a low-frequency segment (usually S1), can then be recognized, according to the relation between the systolic interval and corresponding heart cycle period. The obtained low-frequency and high-frequency segments are then accurately matched to S1 or S2 sound components. The main steps of the method are illustrated in **Figure 5.7**. In the following paragraphs, the proposed method is described in detail.

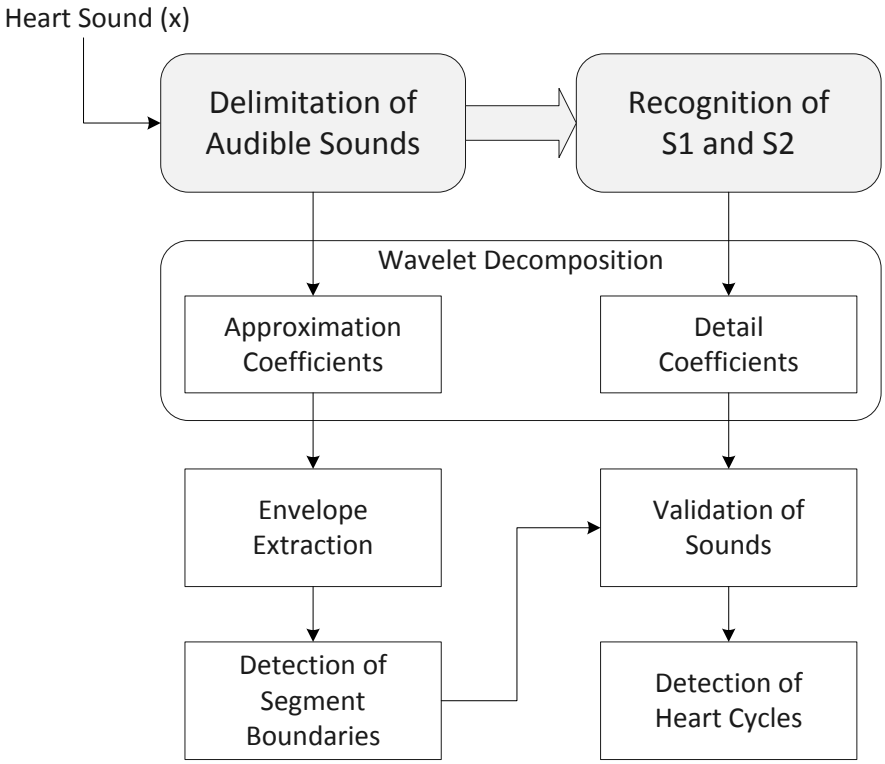


Figure 5.7. Illustration of the energy envelop method.

5.5.1. Delimitation of Audible Sounds

In order to clearly identify the boundaries of all sound segments available in the heart sound, the signal is first filtered and then its envelope is computed. The determination of segment boundaries is carried out by analyzing the zero-crossings of the normalized envelope curve.

Based on the knowledge of the frequency range of heart sounds, we concluded

that the *db6* approximation coefficients at the 5th level are the appropriate depth level for further sound delimitation (for a given sampling frequency 8820 Hz, which is obtained by downsampling the original sample frequency 44.1 kHz). These coefficients represent the transient part of the low frequency heart sound. To extract the signal envelope from the approximation coefficients, the Shannon energy operator is then applied, as in (5.11). It emphasizes the medium intensity signal and attenuates the effect of low intensity signals (Liang et al. 1997). Since the S1 and S2 heart sounds are mainly medium intensity sounds, in contrast to murmur, S3 or S4 sounds, the desired main heart sounds are therefore emphasized.

$$E_{Sh}(x_a) = \frac{1}{W}\sum_W x_a^2 \log(x_a)^2 \tag{5.11}$$

There, x_a denotes the approximation coefficients of the wavelet-transformed heart sound signal and W is the length of the analysis window. The energy is computed using a sliding and centered window (W) of 20 ms with 10 ms segment overlapped. The Shannon energy is then normalized, according to (5.12),

$$E_{Sh}^{norm}(x_a) = \frac{E_{Sh}(x_a) - \mu(x_a)}{\sigma(x_a)} \tag{5.12}$$

Where $E_{Sh}^{norm}()$ stands for the normalized Shannon energy operator, $\mu(.)$ denotes mean and $\sigma(.)$ represents standard deviation. From the zero-crossings of the normalized energy, sound segment boundaries are then identified.

The results of segment boundary detection can be seen in **Figure 5.8**. There, heart sound energy is depicted, rather than the original heart sound signal.

5.5.2. Recognition of S1 and S2 Sound Components

In the process to recognize S1 and S2 sounds from the previously obtained segments, clinically meaningful sound segments are first identified by physiological validation criteria related to heart sound production and the characteristics of heart sound components.

(A) Validation for Heart Sounds

It is observed that some irrelevant sounds and artifacts are also identified in the result of demarcating heart sounds timings. For instance, very long duration or noisy segments may not be conveniently filtered out using the wavelet decomposition and the Shannon energy approach. Therefore, based on physiological facts such as the ones below, these types of sound segments are assessed:

(i) In many situations, it is observed that two sound segments separated by an interval smaller than 50 ms may belong to the same second heart sound. Therefore, these contiguous segments of small duration but separated with less than 50 ms are merged into a single segment. In fact, the presence of a large interval between two consecutive segments is considered to correspond to splitting in the diastole, i.e., splitting of the sound induced by the closing of the aortic and the pulmonary valves.

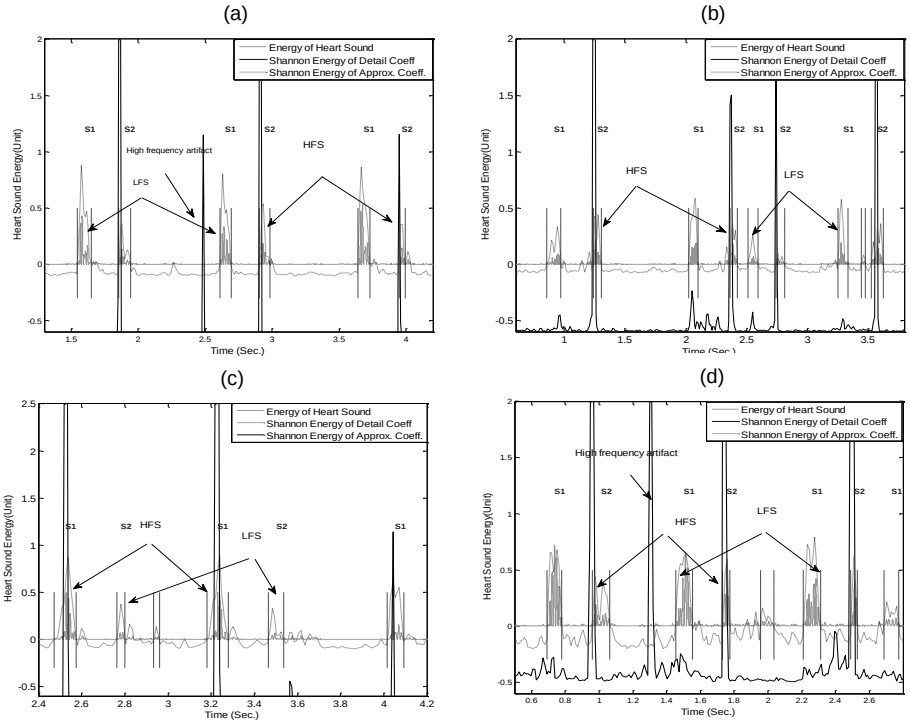


Figure 5.8. High frequency energy above the threshold: (a) Mechanical valve in aortic position, (b) native valve, (c) Mechanical valve in mitral position, (d) Prosthetic valve.

(i) The duration of S1 and S2 sounds is no more than 250 ms and no less than 30 ms in a normal population or even in cardiac patients (according to a broad range of assumed heart rates) (Tavel 1978). Hence, segments exhibiting durations outside this interval may be considered as noise and are discarded.

(B) Detection of Heart Cycles

The goal of heart cycle detection is to find at least one type of heart sound, i.e., S1 or S2. Since S2 sounds contain high frequency samples, these are identified first. Heart cycle using the high frequency signature from the details coefficients are detected by following the steps shown in **Figure 5.9**.

The major part of S1 and S2 heart sounds is the consequence of the vibration of the mitral and aortic valve leaflets, respectively. The frequency of valve vibration depends on the pressure difference across the valves. The relationship between pressure-difference is given by $f \propto Pr^{1/3}$, where f is the frequency of the valve vibration and Pr is the pressure (Lim et al. 1980). From the knowledge of cardiac functionality and genesis of S1 and S2 sounds, it is known that aortic valves close with relatively large pressure difference across the valve; this high pressure difference justifies the high frequency mingling in S2 sounds. Thus, usually S2 sounds contain higher frequency with respect to S1 sound (excluding some rare exceptions). Nevertheless, this characteristic may be used as a marker, similarly to the QRS-complex in ECG, to identify heart cycles. Some rare exceptions may occur for some models of prosthetic valves, namely single tilted disk valves. In case of mitral valve implants, these valves induce higher frequency prevalence in S1 sounds (see **Figure 5.8**(c)). From the same figure, it turns out clear that the difference of high frequency energies between S1 and S2 sounds in all categories of heart sound is significant, which assists in the discrimination of these two sounds using an adaptive threshold. Hence, the obtained sound segments can be identified as corresponding to S1, S2 or non-cardiac sounds.

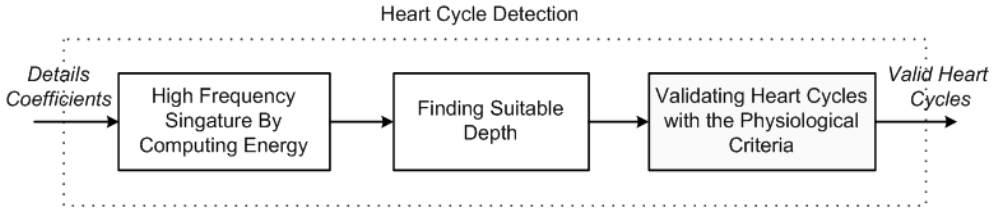


Figure 5.9. Step involved in valid heart cycles detection.

In order to find the presence of high frequency contents in at least one type of heart sound, detail coefficients of the wavelet decomposition are employed, as mentioned previously. Namely, to extract the high frequency envelopes in sound segments, the squared energy operator is applied to the detail coefficients, as in (5.13). In fact, in (Liang et al. 1997) it is shown that simple square energy is a more efficient measure to detect the presence of high intensity samples in comparison to Shannon entropy, entropy and Shannon energy.

$$E_{Sq} = \sum_{W} x_d^2 \tag{5.13}$$

In (5.13), x_d denotes the detail coefficients and W is the length of the analysis window. It is observed that normal and prosthetic valve heart sounds exhibit a significant high-frequency energy difference in both types of sound segments (S1 or S2). The goal is, then, to find the depth of detail coefficients that maximizes the difference of the energies in those classes. Since the first two levels of detail coefficients are mostly noise, the depth is determined between the 3rd to the 7th level (given the sampling frequency in our work). Next, the second goal is to find an appropriate threshold to distinguish the two classes of sounds.

From the initial segmentation using the approximation coefficients, two kinds of sound segments can be found: (i) high-frequency segments (HFS); (ii) and low-frequency segments (LFS). This is illustrated in **Figure 5.8**. In most cases (excluding the exception of single tilted disk mechanical valves), HFS correspond to S2 sounds, while LFS correspond to S1 sounds. Since heart cycles consist of both S1 and S2 sounds, heart cycles are constructed by a pair of segments from each class. Thus, one heart cycle may be defined from the beginning of one HFS to the beginning of the next HFS, i.e., one heart cycle can be determined from a HFS-LFS-HFS sequence.

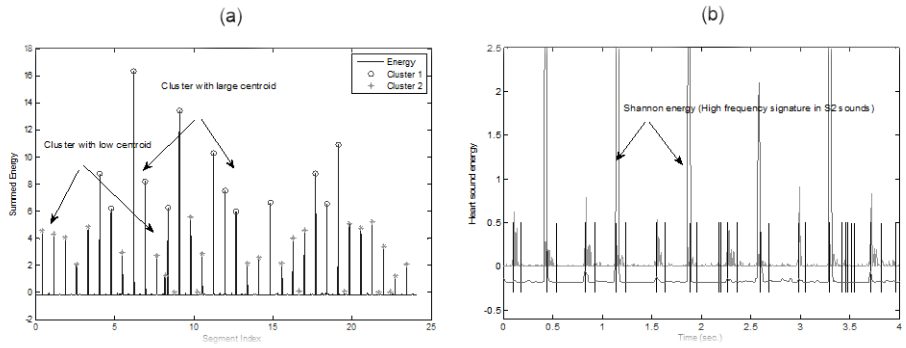


Figure 5.10. (a) 2-classes clustering of squared-energy peaks, (b) heart sound with high frequency signature in S2 sounds detected at 6th level of details coefficients.

The suitable depth for detail coefficients is between the third and seventh level. Hence, after computing the squared energy as in (5.11), the peaks in the energy profile for different levels of detail coefficients are identified. These peaks are assumed

to stem from S1 or S2 sounds. In the next step, the peaks are classified by well known K-means clustering (MacQueen 1967). It is observed that, in the presence of valid S1 and S2 sounds, the difference of the numbers of elements in both clusters is not large. The cluster with the highest frequency centroid is assumed to contain the high frequency segments (HFS). This cluster is further evaluated to check for physiologically valid heart cycles, i.e., heart cycles whose duration is between 375 ms and 1500 ms (duration of the heart cycles at rest). Furthermore, it has been observed, which is also physiologically evident, that in most heart sound the number of elements in the higher centroid cluster is nearly the same as the one in the lower centroid cluster. Based on the fact that S1 and S2 sounds are nearly equal in a given heart sound recording, and if number of elements of a smaller cluster is more than 75% [11]of the larger elements of the cluster then that depth of details coefficients are selected to determine a high frequency signature. Therefore, the depth of detail coefficients that meets the conditions described above is the one selected. In the case illustrated in **Figure 5.10**, the 6th depth proved to be the most suitable one (see **Figure 5.10**(a)).

The identified depth of details coefficients is utilized to extract a high frequency signature from heart sound segments, as in (5.13). This results as an accentuated difference of energy between HFS and LFS, as illustrated in **Figure 5.10**. Now, in order to detect individual heart cycles, an adaptive threshold is defined over the energy envelop, which can be used to mark HFS. This threshold is given by (5.14):

$$th = x_d - \alpha.\mu(x_d) \tag{5.14}$$

where α is an adjustable parameter. As stated earlier, each heart cycle generally comprises one HFS and one LFS. In other words, to identify one heart cycle, at least one LFS must be present between two consecutive HFS. Therefore, following this clue and using an iterative thresholding approach, the appropriate α value is obtained. The iterative process is started with the initial value of α as 1.0. Then, in the following iterations α is incremented with a small fraction of 0.25. Usually, in the first run most segments are identified as HFS. But as the process advances, the method eventually reaches to a stage where the number of HFS in a HFS-LFS-HFS sequence is more than 70% of the total number of HFS. **Figure 5.8** illustrates the identification of segments for a number of examples comprising normal or prosthetic valves.

[11] This number is empirically chosen with the following assumptions: (1) cluster which has greater centroid is cluster of energy of HFS, and smaller centroid is of LFS; (2) ideally, number of LFS and HFS in a heart sound recording is same.

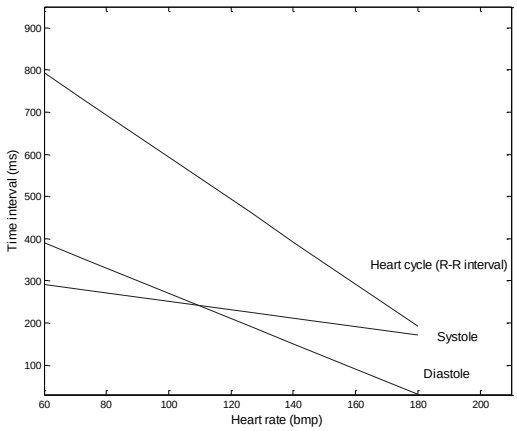


Figure 5.11. Correlation between heart rate and the length of systole, diastole, and R-R interval. The length systole has a decreasing trend but it is stable whereas diastole decreases significantly. Therefore S2 locate, in higher heart rates, relatively later within heart cycle. ms= millisecond ; bmp= beat per minute. (figure is redrawn from (El-Segaier et al. 2005)).

Due to artifacts or noisy segments in the heart sound, some segments can be inadvertently detected as HFS, leading to erroneous detection of heart cycles. Thus, all the detected heart cycles need to be validated. Furthermore, cases where some heart cycles are undetected due to missing high frequency signatures (e.g., HFS detected as LFS) must be tackled as well. In this process, the typical range of human heart rate at rest and the Shannon energy of the obtained segments are utilized.

The timing (or duration) of a heart cycle is defined as the interval between successive R-R peaks in ECG waves. In the absence of an ECG, the timing of the heart cycles can be measured from the start of one S1 sound to the start of the next S1 sound. Since systolic intervals do not change much in cycle to cycle (El-Segaier et al. 2005), therefore the timing between the start of one S2 sound to the next one can also be used to estimate heart rate. Hence, the duration of each heart cycle corresponds to the starting time of one HFS to the next HFS, as in (5.15):

$$T_{cycle}^{k} = t_{HFS}^{k+1} - t_{HFS}^{k} \tag{5.15}$$

In (5.15), T_{cycle}^{k} in the timing of the k^{th} heart cycle, while t_{HFS}^{k+1} and t_{HFS}^{k} are the starting times of the k^{th} and $(k+1)^{th}$ HFS, respectively.

In process of heart cycle detection sometimes several HFS can be found due to high frequency artefacts. However, only some of them are valid. Furthermore, there were

cases of heart cycles without containing one LFS between two HFS. **Figure 5.11** shows that heart rates in human are from 40 bpm (average of timing of heart cycles 1500 ms) to 160 bmp (average of timing of heart cycles 375 ms)[12]. Hence, if the duration of the heart cycles lies outside of this range then that heart cycle is examined; the condition for valid heart cycles can be written as in (5.16).

$$Valid\ Heart\ Cycle: 400\ ms < T_{cycle}^{k} < 1000\ ms \tag{5.16}$$

Up to this stage, a significant number of HFS are detected based on the high frequency energy marker described above. Since the Shannon energy has already been computed during the sound delimitation stage, it can be used again here to compute the Shannon energy of the HFS that are part of valid heart cycles. Let E_{Sh}^{HFS} be the Shannon energy of HFS and t^{start} and t^{stop} denote the starting and stopping time of the each initial sound segment, respectively. Then, the Shannon energy of segments that are categorized as HFS can be calculated according to (5.17) and the average energy of all detected HFS, $E_{Sh_{av}}^{HFS}$, is computed as in (5.18).

$$E_{Sh}^{HFS} = \sum_{t=t^{start}}^{t^{stop}} E_{Sh}^{norm}(t) \tag{5.17}$$

$$E_{Sh_{av}}^{HFS} = \mu(E_{Sh}^{HFS}) \tag{5.18}$$

The averaged value, $E_{Sh_{av}}^{HFS}$, is utilized to validate HFS, as well as to find the missing HFS. The following cases correspond to situations where the detected heart cycles are invalid:

- ***Case 1:*** This situation occurs when a high frequency marker goes below the threshold defined in (5.14) (see Figure 5.12 (a)) and two or more LFS are found between two contiguous HFS. In this case, the duration of the corresponding heart cycle is greater than the average of the three recent previous adjacent cycles, i.e. $T_{cycle}^{k} > \frac{1.6}{3}\sum_{i=1}^{3} T_{cycle}^{k-i}$. This criterion of the average duration of the three most recent cycles and the condition given in (5.16) are sufficient to find missing cycles even in arrhythmic heart sounds. In order to find the relevant HFS that are missed in the first step, the Shannon energy of LFS between two HFS are compared to the average Shannon energy $E_{Sh_{av}}^{HFS}$. A segment initially classified as LFS can be re-classified as HFS if its Shannon energy lies between $E_{Sh_{av}}^{HFS}$-2σ(E_{Sh}^{HFS}) and $E_{Sh_{av}}^{HFS}$+2σ(E_{Sh}^{HFS}).

[12] From http://www.heart.com/heart-rate-chart.html

Heart rates (bpm) : babies to age 1 : 100-160, children age (1-10) : 60-140, children age above 10 and adults (60-100), atheletes : (40-60)

- ***Case 2:*** This situation is met when noisy small duration segments occur (see **Figure 5.12(b)**). In this situation, the correct HFS is found using (5.17) and (5.18). Generally, in these situations Shannon energy of such HFS is found to be valid HFS candidates for S1 or S2 sounds, therefore, the criterion that was used in Case I is applied here as well, i.e. E_{Sh}^{HFS} must be within two standard deviation of averaged Shannon energy of the valid HFS $E_{Sh_{av}}^{HFS}$.

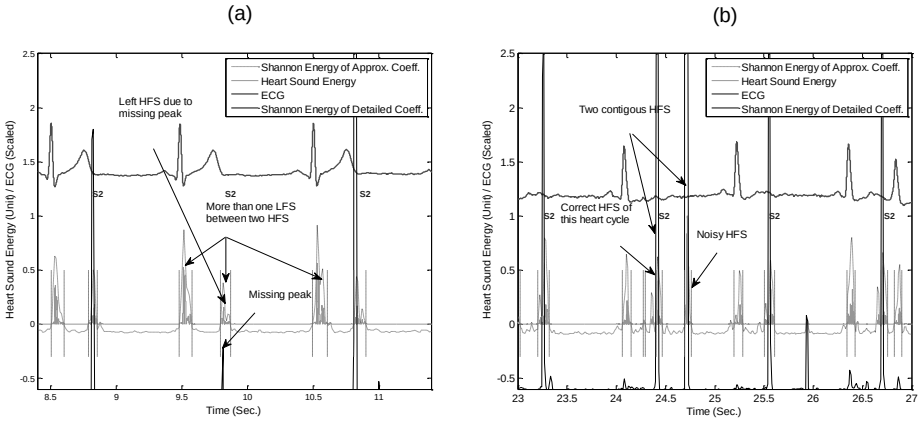


Figure 5.12. Cases of (a) missing HFS, (b) adjacent HFS.

(C) S1 and S2 Identification

In order to perform a correct detection of S1 and S2 sounds, each detected heart cycle is selected and the corresponding systolic interval, i.e. the S1-S2 interval, is estimated and compared with the duration of every HFS-LFS (or LFS-HFS) segment pair. The HFS-LFS (or LFS-HFS) pair corresponding to the smallest deviation from the estimated systolic interval of the corresponding heart cycle is assessed as an S2-S1 pair. Systolic time interval is observed to be nearly constant in a normal heart at a rest state of the subjects. However, it changes rather significantly in arrhythmic heart cases (El-Segaier et al. 2005; Hoekstra et al. 2010; Morgenstern et al. 1978). These are taken into consideration in the proposed method. The heart sounds S1 and S2 are identified and classified through the following steps:

Step1: As mentioned previously, not all HFS can be assessed as S2 sounds with complete certainty. To make the method automatic and independent from prior knowledge, the type of HFS (S1 or S2) is first identified. In this process, the first heart cycle is chosen from the set of correctly detected heart cycles, which has 2 HFS and 1 LFS (see in **Figure 5.13**(a)). The systolic interval for one heart cycle is

estimated using an empirical relationship between the duration of the heart cycle T_{cycle} and the systolic interval. It is given by,

$$T_{sys}^{est} = 0.2T_{cycle} + 160\ (ms) \tag{5.19}$$

where T_{sys}^{est} is the estimated systolic interval corresponding to a correctly detected heart cycle. This relationship provides a very close estimation of the systolic interval (El-Segaier et al. 2005). The estimated systolic interval is then evaluated to check whether it is within the allowed physiological range, i.e. between 175 ms and 380 ms, as proposed in (El-Segaier et al. 2005). Next, the durations of HFS-LFS and HFS-LFS are computed, i.e. T1 and T2, as depicted in **Figure 5.13**(a). Then, the duration, T1 or T2, having the smallest deviation with respect to T_{sys}^{est} will be evaluated as the systolic interval of the corresponding heart cycle. In the example provided in **Figure 5.13**(a), T2 has the smallest deviation with respect to T_{sys}^{est}. Therefore, LFS-HFS is a systolic pair. Once the correct systolic pair is detected, the HFS is identified as S2 for the selected cycle. Since all the HFS belong to the same class of heart sounds, the type of all HFS is determined (in **Figure 5.13**(a), all the HFS are S2 sounds). On the contrary, in heart sounds obtained from single-tilted disk in mitral position, the HFS-LFS pair corresponds to the systolic interval; consequently, the remaining HFS would be S1 sounds.

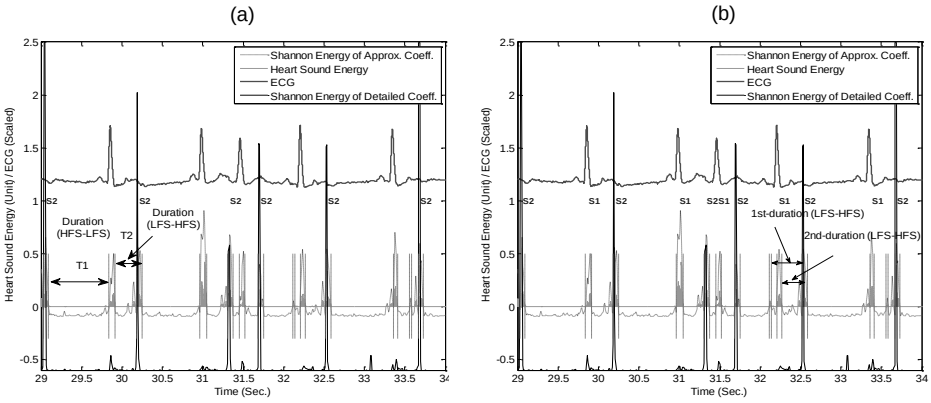


Figure 5.13. Arrhythmic sounds S1 and S2 detection using for HFS and LFS. (a) systolic interval estimation in one cycle as in HFS-LFS-HFS sequence, (b) correct LFS for being S1sound candidate.

Step2: In this step, the sound components corresponding to LFS (S1, for normal heart sounds, or S2, in case of mitral mechanical valves) are identified. Since the sound components corresponding to HFS are already identified from the previous

step, the type of LFS can be identified using each heart cycle (HFS-LFS-HFS). If there is only one LFS between two contiguous HFS, it belongs to the other class of heart sound, i.e., S1 if HFS are S2, or S2 if HFS are S1 (see **Figure 5.13**(b)).

Despite the described validation process, not all low-frequency physiologically-irrelevant segments may be eliminated. In this situation, more than one LFS between two HFS are found, in which case the correct LFS is identified by the comparison between the LFS-HFS (or HFS-LFS) duration and the estimated systolic interval. In this process, the systolic interval of the corresponding heart cycle is first estimated, as before. Then, all the LFS-HFS (or HFS-LFS) durations are computed. These durations are compared with the estimated systolic interval T_{sys}^{est}. The LFS-HFS or HFS-LFS pair which exhibits the smallest deviation from T_{sys}^{est} will be the systolic sound pair, i.e., S1-S2. In the presented example, since all the HFS are S2 sounds and LFS are suspected to be S1 sounds, the durations of LFS-HFS pairs are computed and compared with T_{sys}^{est}. The correct LFS-HFS systolic pair is then found and the type of LFS is detected. In the example illustrated in **Figure 5.13**(b), the duration of the second LFS-HFS pair comes closer to the estimated systolic interval and the corresponding LFS is identified as the S1 sound component. In exceptional cases when all HFS are S1 sounds, the durations of all the HFS-LFS are computed and compared with T_{sys}^{est}.

5.6. Segmentation of Heart Sound with Murmur: Wavelet Decomposition and Non-linear based Segmentation (WD-NLF)

In this section, the method proposed for the segmentation of heart sounds containing murmur is described. It is based upon the extraction of nonlinear features in the wavelet domain. In order to detect the presence of murmur between two main components, discriminative features able to highlight the difference between the limits of S1/S2 and murmur sounds are necessary. To this end, nonlinear features such as simplicity (complexity) and strength of the heart sound components are computed using the embedded matrix constructed in (5.1) of Section 5.3.1.

Since different kinds of murmur sounds contain a different range of frequencies, these features are extracted from wavelet-decomposed coefficients in order to capture the locations of sound components in a feature profile curve. Those features are iteratively extracted from increasing levels of approximation coefficients and, as before, the depth leading to best segmentation is selected.

This approach was inspired by an adaptive filtering method based on the wavelet transform and fractal dimension, for the enhancement and separation of explosive lung sound and bowel sound from background noise (Hadjileontiadis 2005). According to the method, first, the fractal dimension from the approximation coefficients of a wavelet-decomposed bioacoustics signal is computed, which exhibits Lung Sound and Bowling Sound peaks. Second, using an iteratively selected threshold, all peaks are picked. Similarly, our algorithm aims to highlight peaks corresponding to S1 and S2 sounds. In fact, for the separation of S1 and S2 heart sounds from murmurs, simplicity and strength are derived from the embedded matrix of the wavelet-transformed heart sound signal. This procedure highlights the desired S1 and S2 peaks.

It is observed that heart murmurs are usually characterized by having high frequency content and for being more complex than S1 and S2 sounds. In most heart sounds containing murmur, it has been observed that the S1 and S2 components exhibit high simplicity and high strength (see **Figure 5.15**). On the other hand, murmur exhibits low simplicity and high strength. As for non-audible sounds in the systolic/diastolic region, these sounds exhibit high simplicity and low strength. Hence, an adaptive iteratively threshold is used to identify the locations of S1 and S2 components, based on the common area of the profiles of these feature. To summarize, in order to automatically segment the main heart sounds, a two-stage algorithm is proposed: in the first stage, delimitation is performed using the WD-NLF method; and in the second stage, S1 and S2 sounds are recognized using a high frequency signature and physiological evidences, according to the section 5.5.2. The main steps of the algorithm are illustrated in **Figure 5.14**.

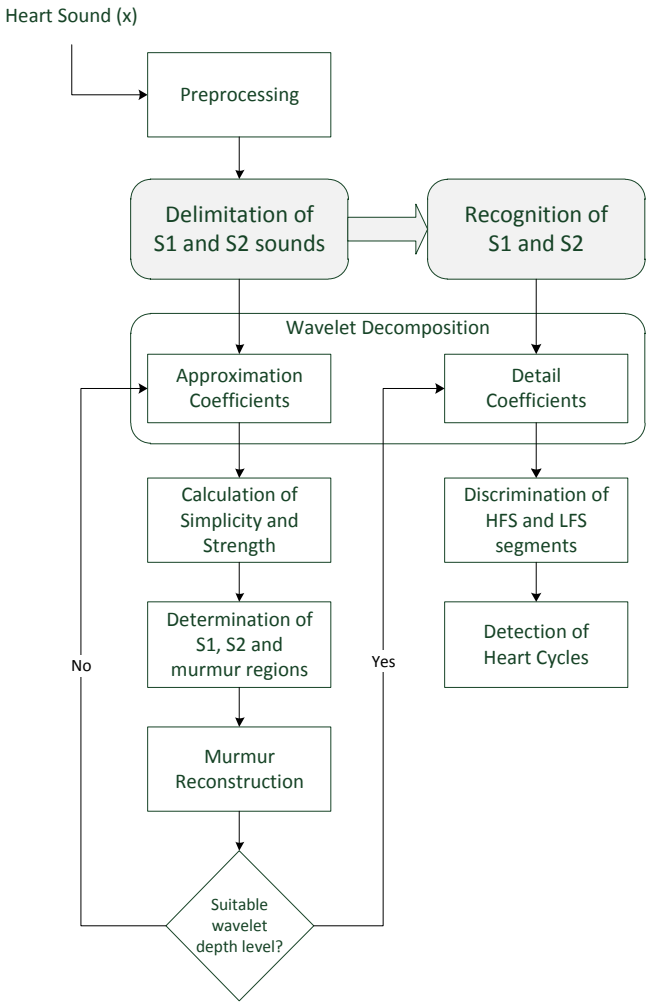


Figure 5.14. Illustration of the method based on wavelet decomposition and nonlinear features (WD-NLF).

5.6.1. Nonlinear Features (Simplicity and Strength)

The two employed features, simplicity and strength, are computed using the embedded matrix given in section 5.3.1 from (5.1). As will be described later, in this case the embedded matrix is computed from the approximation coefficients after wavelet decomposition, in a frame-wise fashion. Namely, strength is based on the loudness of the heart sound components and is calculated as in (5.20):

$$Sg = \frac{1}{M}\sqrt{\frac{\|\mathbf{X}\|}{mM}} \tag{5.20}$$

In (5.20), $\|.\|$ stands for the Euclidean norm operator and Sg denotes strength.

Simplicity is calculated based on the entropy of the normalized eigenvalues of $\mathbf{X}$. To this end, the covariance matrix is first obtained, as given by (5.21),

$$\boldsymbol{C} = \mathbf{X}^T\mathbf{X} \tag{5.21}$$

where T denotes the transpose of $\mathbf{X}$. Next, the eigenvalues of the covariance matrix $\boldsymbol{C}$ are computed. Let $\lambda_1, \lambda_2, \lambda_3, \ldots\ldots \lambda_m$ be the eigenvalues of $\boldsymbol{C}$ and let $\hat{\lambda}_1, \hat{\lambda}_2, \ldots\ldots \hat{\lambda}_m$ be the normalized eigenvalues, computed according to (5.22):

$$\widehat{\lambda}_i = \frac{\lambda_i}{\sum_{k=1}^{m}\lambda_k} \tag{5.22}$$

Now, the entropy, H, of the normalized eigenvalues is computed as in (5.23):

$$H = -\sum_{i=1}^{m}\hat{\lambda}_i \log_2 \hat{\lambda}_i \tag{5.23}$$

Finally, simplicity, Sm, is obtained from the entropy, as in (5.24).

$$Sm = \frac{1}{2^H} \tag{5.24}$$

In order to compute strength and simplicity profiles from the heart sound signal, a computational window with 20 ms duration is defined. The reason for selecting this analysis window was inspired by timing of physiologically events in the heart sounds such as duration of the constituents of S1, i.e., tricuspid and mitral valve closing sound and duration of the constituents of S2 sounds, i.e., aortic and pulmonary valve closing sounds (Kudriavtsev et al. 2004). These sounds can be captured by these measures if the split between two constituent sounds is wider (> 0.02 sec.) (Kudriavtsev et al. 2004). Moreover, to avoid abrupt transitions, a 50% window overlap is defined. **Figure 5.15** shows the strength and simplicity profiles for one sample of heart sound with murmur resulting from mitral regurgitation. As it can be seen, both the S1 and S2 sounds exhibit high strength and high simplicity. This is a crucial pattern to identify the boundaries of those components.

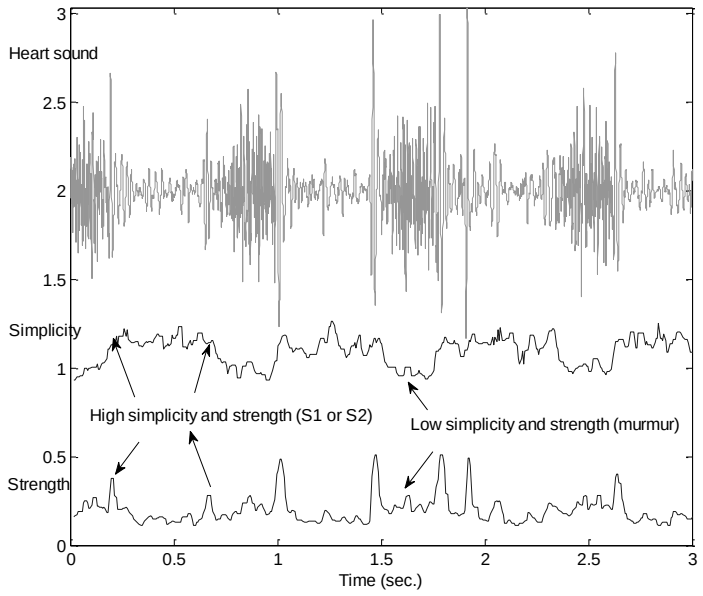


Figure 5.15. (top signal) Heart sound with mitral regurgitation, (middle signal) Strength profile, (bottom signal) Simplicity profile.

5.6.2. Phase I: Sound Delimitation

The proposed method, based on wavelet decomposition and nonlinear features (WD-NLF) uses the wavelet transform based on multi-resolution decomposition to initially decompose the heart sound into approximation and detail coefficients. Subsequently, simplicity and strength are computed from the embedded matrix, which is constructed using the approximation coefficients. As mentioned above, both S1 and S2 sounds exhibit high simplicity (*Sm*) and strength (*Sg*). Hence, the *Sm* and *Sg* peaks of S1 and S2 are identified using an iteratively adaptive threshold, which is based on mean squared error. Furthermore, the suitable depth of wavelet decomposition level is also iteratively found using a mean square error criterion. The entire algorithm involves the following steps:

Step 1: The heart sound is decomposed by the Wavelet Transform using *db6* as the mother wavelet. Let l be the l^{th} level decomposition, where $l = 1,2,3 \ldots \ldots, L$. In the following description, l starts with 1 and is iteratively incremented until final depth level, L, is reached, according to the stopping criteria defined in step 7.

Step 2: The required parameters, τ^l and m^l, for the phase space reconstruction of the approximation coefficients of the l^{th} level, are estimated.

Step 3: Simplicity (Sm^l) and Strength (Sg^l) are computed from the embedded matrix, $\mathbf{X}^{\boldsymbol{l}}$, constructed from the approximation coefficients of the l^{th} level, as described previously.

Step 4: The S1 and S2 components exhibit high strength and simplicity, hence related peaks can be observed in these curves. The peaks in both feature curves are picked using the peak picking algorithm (PPA) described in (Hadjileontiadis 2005). This algorithm finds not only peaks corresponding to S1 and S2 sounds, but also the duration of these sounds, as illustrated in **Figure 5.16**(b) and (c). This is accomplished by employing an iteratively adaptive threshold whose convergence is obtained based on a stop criterion introduced by Hadjileontiadis et al. (Hadjileontiadis 2005; Hadjileontiadis 2003). The computational steps involved in PPA are described below.

Peak Peeling Algorithm (PPA) (modified from (Hadjileontiadis 2003))

Initialization

x_a^l=$\{x_{a_i}^l\}$, $i = 1,2,3, \ldots . N$

$s^1 = NLF(x_a^l);\ \ z^0 = 0$, *where NLF is Sm or Sg*

$s = \{s_p^1\},\ p = 1,2,3 \ldots N - W$+1

$iniTh = \ 0.5(min(s_p) \ + \ (\max(s_p))$

ϵ_1=10^{-5}; $l_1 = 0$;

Step 1

do{ $l_1 = l_1 + 1$;

$\sigma^{l_1} = std(s^{l_1})$; *where, std is standard deviation*

$pNF_p^{l_1} = \{s_p^{l_1}\ if\ s_p^{l_1} > iniTh + \sigma^{l_1}; 1.0\ else\ \}$

$z^{l_1} = s^{l_1} - pNF^{l_1} + 1$

$SC^{l_1} = \left|E\{(z^{l_1})^2\} - E\{(z^{l_1-1})^2\}\right|$

$s^{l_1+1} = z^{l_1}$

} *while* $(SC_1^l < \epsilon_1)$; stopping criterion

Step 2

$$L_1 = l_1$$
$$NFDD^l = \sum_{l_1=1}^{L_1} pNF_1^l - (L_1 - 1),$$

where L_1 denotes maximum number of peeling level

In the above algorithm of PPA, x_a^l denotes the approximation coefficients of the decomposed heart sound x, N is the number of samples in x_a^l, W is the samples in the analysis window and $\mathrm{E}\{.\}$ is the expected value operator. The estimated profile curve, $NFDD$, is assigned the value 1.0 except for peaks of the low complexity and high strength part of the signal (which correspond to S1 and S2 components).

Step 5: It is observed that the correct start and stop instants of S1 and S2 sounds can be determined by common segmented time gates in both feature curves. Let $\left\{T_{Sg^l}^{start}, T_{Sg^l}^{stop}\right\}$ and $\left\{T_{Sm^l}^{start}, T_{Sm^l}^{stop}\right\}$ be the start and stop times of S1 and S2 time gates for two time-overlapping segments in the strength and simplicity curves, respectively. Thus, the correct start and stop times for the overlapping segments is set to their overlapping region, as given (5.25),

$$\left\{T_x^{start}, T_x^{stop}\right\} = \left\{T_{Sg^l}^{start}, T_{Sg^l}^{stop}\right\} \cap \left\{T_{Sm^l}^{start}, T_{Sm^l}^{stop}\right\} \tag{5.25}$$

where $\left\{T_x^{start}, T_x^{stop}\right\}$ is the start and stop time for the segments under analysis. This is repeated for all pairs of overlapping segments in both feature curves. This step is illustrated in **Figure 5.16**(b), (c) and (d).

Step 6: After estimating all segments, two delimiting square waves are constructed, one Sr^l, denoting the regions with S1 and S2 sounds and the other, Mr^l, representing the regions with murmurs and other sounds, according to (5.26) and (5.27), respectively. This is depicted in **Figure 5.16**(d) and (e).

$$Sr^l = \begin{cases} 1 & \left\{T_x^{start}, T_x^{stop}\right\} \neq 0 \\ 0 & else \end{cases} \tag{5.26}$$

$$Mr^l = 1.0 - Sr^l \tag{5.27}$$

Both square waves are then multiplied by the approximation coefficients. The out-

come of these multiplications consists of the approximation coefficients related to the S1 and S2 sound waveform and to the murmur waveform, respectively.

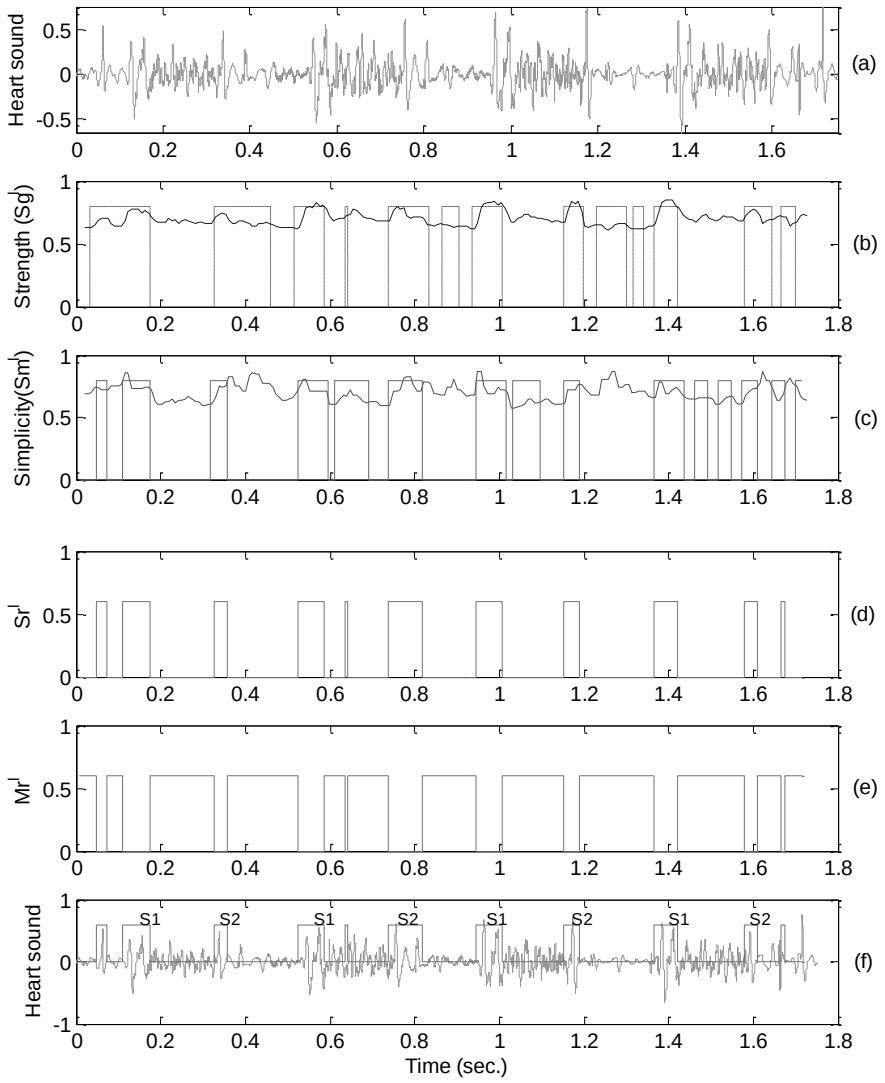


Figure 5.16. Wavelet decomposition-nonlinear features segmentation; (a) Heart sound energy in a mitral regurgitation situation (grade V), (b) Strength curve and candidate segments, (c) Simplicity curve and candidate segments, (d) Region for S1 and S2 sounds, (e) Region with murmurs and other sounds, (f) segmented S1 and S2 sounds.

Step 7: Now, the wavelet coefficients related to murmur are reconstructed in order to determine the suitable wavelet decomposition depth. Let Y_M^l be the reconstructed murmur signal then the stopping criterion is defined according to (5.28).

$$err^l = \left|E\{(Y_M^l)^2\} - E\{(Y_M^{l-1})^2\}\right| < \epsilon_2 \tag{5.28}$$

In (5.28), *err* represents error, $E\{.\}$ denotes expected value and $\epsilon_2 \in (0,1)$. In this work, $\grave{o}$ is experimentally fixed to 0.1. If the stopping criterion in (5.28) is not satisfied, then l is incremented and the algorithm is repeated from step 1 until the stopping criterion is met. In (5.28), Y_M^0 is initialized with the averaged value of the original heart sound signal.

5.6.3. Phase II: Recognition of S1 and S2

In this phase, S1 and S2 are identified from the information collected in phase I. The gated segments are assumed to be S1 and S2 sounds. From the knowledge of heart physiology, it is known that the pressures across the aortic valve are higher compared to pressures across the mitral valve. Hence, S2 sounds tend to exhibit higher frequency content than S1 sounds. High frequency signatures are obtained from the squared energy of the detail coefficients of the heart sound at the depth of decomposition determined in phase I. Based upon the squared energy of the detail coefficients, computed as in (5.13), S1 and S2 discrimination is performed according to the following steps (see **Figure 5.14**).

Step 1: Energy of the signal is computed using (5.13) from the detail coefficients at the depth level, l, used for gating in phase I.

Step 2: Since S2 is believed to exhibit higher energy than S1, a threshold is set to distinguish high frequency segments (HFS) from low frequency segments (LFS), i.e.

$$th = \frac{x_d^l - \rho\, \mu(x_d^l)}{\sigma\,(x_d^l)} \tag{5.29}$$

Where ρ is the multiplier that may vary from 1.0 to 10, μ is the mean, σ is the standard deviation, and x_d^l denotes the detail coefficients of the heart sound at the depth level, l, determined in phase I. Segments above the threshold *th* are HFS,

which are recognized as S2 sounds as they show a high frequency signature in terms of higher energy (see in **Figure 5.17**bottom)).

Step 3: Heart cycles are identified; one heart cycle contains one LFS between two HFS. Multiple LFS between two consecutive HFS are discarded using the stable systolic interval criterion defined previously. Segments caused by noise are pruned based on the fundamental physiological knowledge related to the heart sounds. For instance, heart cycles whose durations do not lie between 400 ms to 1200 ms are invalid and segments exceeding the range 40 ms to 100 ms (duration of S1 and S2) are also invalid. In the process to validate heart cycles, it is utilized the criteria defined in the previous segmentation method. Next, HFS and LFS sound segments are recognized as S2 and S1 sounds based on the method described in section 5.5.2 (B and C).

Figure 5.16 (f) and **Figure 5.17** illustrate the results of the recognition of S1 and S2 components.

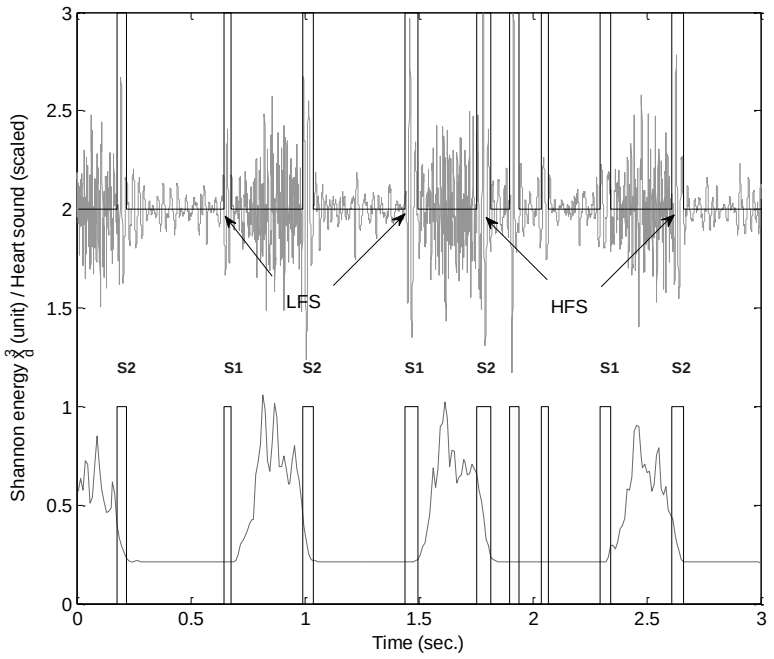


Figure 5.17. (top) Original sound wave and gated S1 (LFS) and S2 (HFS) sounds, (bottom) High frequency energy curve with demarcation of S1 and S2 sounds.

5.7. Test Database

Heart sound samples were collected in the Cardiothoracic Surgery Centre at the University Hospital of Coimbra under the guidance of an experienced cardiologist. The data collection studies were authorized by the local hospital ethical committee and informed consent was obtained from all participants. During acquisition, the patients were asked to maintain silence and to make the least possible physical movements in order to maintain the integrity of heart sound collection. As it will be explained later in this section that three classes of heart sounds were collected, in which most of the heart sounds were recorded in subject's supine position. The best site for auscultation on the chest in order to capture audible heart sound on the chest was chosen by doctors based on the clarity and loudness. Hence, no definite auscultation site was selected but second intercoastal space and the apex were found more often to be the best sites. Recording was performed with an electronic stethoscope from Meditron with an excellent signal to noise ratio and extended frequency range (20 - 20,000 Hz). Although the use of ECG is not considered in the present work, it was also recorded simultaneously to assess the segmentation efficiency of the algorithm.

Three distinct databases have been prepared in the aforementioned data collection study: i) the first database includes heart sounds acquired from volunteers with native heart valves (*databaseHS*); ii) the second database is composed by heart sounds with murmur (*databaseMurmur*); and, iii) the third database contains heart sound clips collected from patients with prosthetic heart valve implants (*databaseProsValve*).

I. databaseHS: In this database, only heart sound clips without murmur have been collected. The population was composed of 38 subjects with native heart valves. From these, 31 were healthy males, 2 were healthy females and 5 subjects were males with several types of diagnosed cardiac dysfunctions. The main biometric characteristics of the healthy sub-population were BMI = 24.48±2.46 kg/m2 and age = 29.72±8.54 years. Regarding the unhealthy sub-population, their biometrics were BMI = 25.38±3.45 kg/m2 and age = 48.2±17.15 years.

II. databaseMurmur: This database is composed of 81 heart sound clips with murmurs collected from 51 patients. The population was composed of 48 males and 3 females. The main biometric characteristics of the population were BMI = 25.41±2.16 kg/m2 and age = 64.65±8.64 years. The database contains 7 classes of abnormal heart sounds: Aortic Regurgitation (AR), Aortic Stenosis (AS), Mitral

Regurgitation (MR), Pulmonary Regurgitation (PR), Pulmonary Stenosis (PS), Subaortic Stenosis + Ventricular Septal Defect (SAS+VSD), Systolic Ejection (SYE). These are among the most prevalent murmur classes (Chizner 2008).

III. databaseProsValve: This database was collected from 55 patients with several types of prosthetic valve implants in different positions. 42 patients had the Medtronic single-leaflet mechanical valve implanted (2 patients were extremely arrhythmic) and 9 patients had the Edwards-Lifescience bio-prosthetic heart valve. From each database, 79 heart sound clips were collected, from the different subjects. The age distribution for this population was 52.40±6.23 years. Parameters for BMI, i.e. height and weight, were not collected.

5.8. Performance Measurements

In order to evaluate the proposed method, several performance measures are used. The discrimination ability between different types of heart sounds using Lyapunov exponents is evaluated using one-way variance analysis (ANOVA) with 99% confidence level. In the evaluation of this classifier, sensitivity and specificity are also employed. Sensitivity (SE) and specificity (SP) are computed according to the equation (4.12) defined in Chapter 4).

In (4.30), TP denotes True Positives, i.e. heart sounds without murmur (normal heart sounds or heart sounds produced by prosthetic heart valves) that have been correctly classified; True Negatives (TN) correspond to heart sounds with murmur correctly classified as without murmur. False Positives (FP) are murmur heart sounds classified as non-murmur, whereas False Negatives *(FN)* corresponds to non-murmur sounds classified as heart sounds containing murmur.

Regarding the S1 and S2 detection performance, two distinct measures are extracted. The first one is related to the classification accuracy. Again this is measured using SE and SP (4.30). In this context, TP denotes the number of S1 and S2 correctly detected, TN are irrelevant sound segments correctly detected as insignificant sounds, FP represents irrelevant segments detected as S1 or S2 sound and FN denotes the number of S1 and S2 are wrongly identified as irrelevant sounds.

To assess the detection accuracy of the boundaries of S1 and S2 sounds, we evaluate the detection error with respect to the annotated onsets and endings of these sounds in the databases. Thus, let $t_{reference}$ be the annotated start or stop time instants of a sound component (S1 or S2) and let $t_{detected}$ be the corresponding time instant detected using the described method. Then, the delimitation error is given by (5.30). This error is then averaged over all segments.

$$delimitation\ error = \left|t_{reference} - t_{detected}\right| \tag{5.30}$$

5.9. Experimental Results and Discussions

The first task in the introduced method is to discriminate among normal heart sounds, heart sounds with murmur and heart sounds from prosthetic valves. In order to obtain unbiased chaos in the heart sounds, LLEs were computed from the five seconds of heart sound using each one second length of sound. As a result of this each heart sound contributes with five values of largest Lyapunov exponent. A total of 395 (5×79) LLEs are obtained from each database whose values distribution are shown in the form of a spread plot in **Figure 5.18**(a). These distributions exhibit the following averages and standard deviations: 0.0651 (0.0993) for *databaseHS*, 0.2393 (0.0752) for *databaseMurmur* and 0.5269 (0.1139) for the *databaseProsValve*.

Using the one-way variance analysis (ANOVA), it is observed that the LLE of the heart sound is able to discriminate between the three groups of databases with statistical significance (F=0.0006689). In order to distinguish between the three classes of sounds, the thresholds $th_L$and th_H, defined in (5.10), are determined from the spread plots in **Figure 5.18**(a). These thresholds are estimated based on the receiver operating characteristic (ROC) curve. The ROC curve is based on the true positive rate (sensitivity) and false positive rate (1-specificity), which varies according to the defined threshold values, th_L andth_H. The optimum threshold values are determined by locating the point on the ROC curve leading to the best trade-off between sensitivity and specificity.

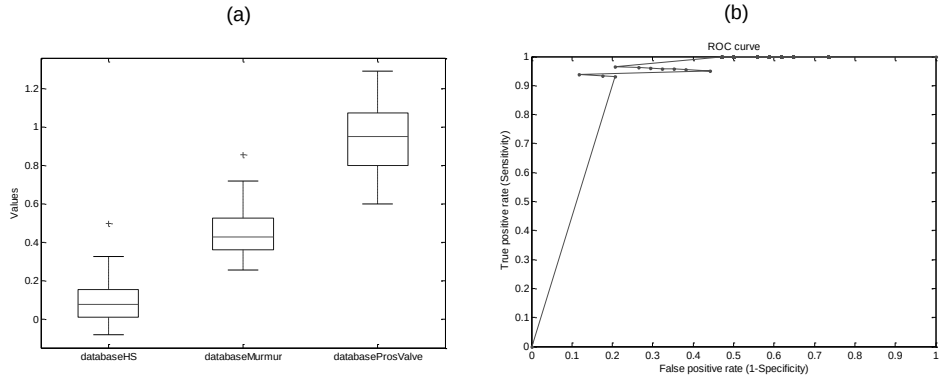


Figure 5.18. (a) Box-Wishker plot of LLE, (b) ROC plot of the th_L and th_H thresh-

olds.

To find the optimum values for th_L and th_H 40 heart sound clips from each database ($\approx$ 50% of the total of 79 heart sound clips in each group of database) is used as the training set. Then, the initial values of th_L and th_H are assigned the average values of the LLEs of normal sounds and sounds from prosthetic valves in the training set. The remaining heart sounds clips from the three databases (39 heart sound clips from each) are classified according to the rule in (5.10). The final ROC curve is plotted in **Figure 5.18**(b)). The upper triangular part of that figure represents the change in classification performance with respect to change in th_L and th_H. The best trade-off point is determined at the sensitivity of 0.96 (96.68%) and specificity of 0.92 (91.81%), corresponding to the thresholds th_L = 0.33 and th_H= 0.68.

A. Segmentation with WD-NLF and Energy envelop methods

Since segmentation is aimed to be performed after detecting the class of the heart sound using the suitable and less computational algorithm for the less complex heart to sound it is observed that some heart sounds may be misclassified, which will lead and that lead to the sub-optimal segmentation algorithm. It will be shown later in this section that heart sounds from *databaseMurmur* are not segmented with higher precision by the energy envelop when compared to the WD-NLF method.

In order to reach a significant conclusion regarding segmentation performance of the both proposed algorithms over the three databases, we have employed the energy envelop method and WD-NLF to every database separately. Therefore, segmentation performance evaluation is carried out in two ways: i) in the first, we evaluate the performance when normal heart sounds and heart sounds from prosthetic valves are segmented by the energy envelop method and heart sounds with murmur by the WD-NLF method; ii) in the second one, the segmentation of heart sounds with murmur using the energy envelop method and the segmentation of normal heart sounds/heart sounds from prosthetic valves using by the WD-NLF method are evaluated.

The achieved results for the first evaluation approach, i.e., segmentation using energy envelop algorithm for *databaseHS* and *databaseProsValve*, and WD-NLF method for *databaseMurmur*, are presented in Table 5.1 and Table 5.2. Table 5.1 presents sensitivity and specificity for S1 and S2 sound detection. Table 5.2 shows the delimitation error related to the identification of the start and end times of each S1 and S2 sound component.

From Table 5.1, we can see that sensitivity and specificity performance is lower in *databaseHS* and *databaseMurmur* than in the *databaseProsValve*. Since many weak transient waveforms are present in some of the heart sounds, namely those recorded

with subjects in supine position, the performance deteriorates due to an increase in the number of false negatives, while the number of true positives decreases. This does not happen for prosthetic valves because the sound produced by these valves tends to be significantly louder, irrespective of the patient's position. In *databaseMurmur* SAS+VSD murmurs show the best sensitivity and specificity, while PR murmurs lead to the worst performance. In fact, it is observed that stationary low grade and high frequency murmurs, such as SAS+VAD are more successfully segmented than transient loud

Heart Sound	**S1 and S2 heart sound detection**	
	Sensitivity (%)	**Specificity (%)**
databaseHS		
Healthy	92.10	93.46
Unhealthy	91.07	92.41
Average	*91.79*	*92.83*
databaseProsValve		
Mechanical valve	97.10	97.21
Bio-prosthetic valves	96.77	97.62
Average	*96.91*	*97.52*
databaseMurmur		
AR	91.77	92.88
AS	94.80	91.90
MR	92.11	90.06
PR	83.08	88.46
PS	92.55	93.02
SAS+VSD	97.00	93.78
SYE	93.23	93.04

Average	*92.98*	*91.96*
Global average	*94.07*	*93.64*

Table 5.1 S1 and S2 recognition performance in terms of Sensitivity (SE) and Specificity (SP): *databaseHS* and *databaseProsValve* are segmented by energy envelop method and *databaseMurmur* by WD-NLF.

murmurs such as PR. To summarize, the overall sensitivity and specificity performance was 94.07% and 93.64%, respectively, for the proposed S1/S2 detection scheme. These outcomes are found to be comparable with the reported results in the literature. For instance, Liang et al. (Liang et al. 1997) obtained 93% accuracy using heart sounds from healthy subjects, classified as the normal heart sound in this chapter, while El-Segaier et al. (El-Segaier et al. 2005) has obtained 100% accuracy in S1 sound detection and 97% accuracy for S2 sound in marginally large database including murmur. However, in the later one ECG signal was used as reference signal to find locations of S1 and S2 sounds with respect to QRS and T waves.

Heart Sound	**S1 error μ ($\pm\sigma$) (ms)**		**S2 error μ ($\pm$ σ) (ms)**	
	Start time	**Stop time**	**Start time**	**Stop time**
databaseHS				
Healthy	5.66 (0.93)	4.11 (0.98)	4.88 (0.83)	5.24 (0.73)
Unhealthy	5.12 (0.86)	4.98 (0.80)	6.68 (0.94)	5.44 (0.82)
Average	*5.39 (0.89)*	*4.54 (0.89)*	*5.78 (0.88)*	*5.34 (0.77)*
databaseProsValve				
Mech. valves	2.57 (0.44)	2.28 (0.40)	2.31 (0.42)	2.15 (0.5)
Biop. valves	1.55 (0.51)	2.67 (0.20)	2.43 (0.49)	3.25 (0.53)
Average	*2.06 (0.47)*	*2.47 (0.30)*	*2.37 (0.45)*	*2.70 (0.51)*
databaseMurmur				

AR	5.67 (0.88)	4.82 (0.68)	4.28 (0.66)	3.86 (0.64)
AS	3.86 (0.38)	6.05 (0.68)	5.20 (0.68)	4.28 (0.78)
MR	4.05 (0.51)	5.68 (0.81)	4.90 (0.96)	3.92 (0.84)
PR	6.20 (0.89)	4.56 (0.78)	7.80 (0.56)	3.44 (0.18)
PS	2.30 (0.64)	3.24 (0.95)	2.88 (0.38)	4.32 (1.07)
SAS+VSD	1.93 (0.86)	2.02 (0.44)	1.45 (0.89)	4.02 (0.34)
SYE	3.34(0.56)	5.10 (0.58)	7.92 (0.42)	5.11 (0.32)
Average	*3.90 (0.67)*	*4.49 (0.65)*	*4.91 (0.65)*	*4.13 (0.59)*
Global Avg.	*3.78 (0.68)*	*4.16 (0.58)*	*4.35 (0.66)*	*4.05 (0.62)*

Table 5.2 S1 and S2 delimitation errors (average value with standard deviation) *databaseHS* and *databaseProsValve* are segmented by energy envelop method and *databaseMurmur* by WD-NLF.

Regarding Table 5.2, it can be noticed that the unhealthy population in *databaseHS* exhibits larger delimitation errors (except for the onset of S1 sounds). This is especially evident in S2 sounds and the main reason for that behavior is the split sounds missing in S2, as the S2 sound is composed by aortic sound (A2) and pulmonary sound (P2), with a gap range from 5 ms to 35 ms between them. Therefore, when this gap is too large and the duration of the P2 sound is smaller than 20 ms, one component of the S2 sound might be overlooked in the segment validation process, leading to larger errors in the start and stop instants. In *databaseMurmur*, the SYE and PR murmur classes exhibit the highest error in the delimitation of S2 sounds, while SAS+VSD and PS show the lowest error for both S1 and S2 sounds. In the former, due to the overlapping of murmur with S1 and S2 sounds, simplicity and strength are not capable to demarcate the precise instants of the onset and ending of the heart sound. Finally, *databaseProsValve* shows the smallest average delimitation error, which can be easily explained based on the nature of prosthetic valve click sounds. A click sound is usually a high magnitude (loud) sound containing sharp transients at its onset and ending, which are conveniently detected by Shannon energy near the correct start and stop instants of a heart sound. The overall identification accuracy of all components was 4.08±0.64 ms, which is sufficient for most applications using heart sounds such as pulmonary artrial pressure (PAP), left ventricular ejection time (LVET), blood pressure (BP), heart rate (HR), cardiac out-

put (CO), etc. (Carvalho et al. 2009; Paiva et al. 2009a).

As for the second evaluation approach, i.e., using the incorrectly classified heart sounds, leading to heart sounds with murmur being segmented by the energy envelop method while the others are segmented by the WD-NLF approach, results are presented in Table 5.3 and Table 5.4. From those tables, it can be observed that the energy envelop method does not show comparable performance on the clips of heart sounds with murmur. On the other hand, there is no significant degradation in the accuracy of the WD-NLF method when applied on *databaseHS* and *databaseProsValve*.

Table 5.3 shows less sensitivity and specificity in nearly all high-grade murmur heart sounds. Since the intensity of murmur sounds is significantly lower than the one for S1 and S2 sounds, the threshold for murmur sounds is not conveniently set. Thus, S1 and S2 segments are not robustly detected. Similar results are seen in S1 and S2 delimitation. In Table 5.4, delimitation error for sound components is minimum in SAS+VSD murmurs while the largest error is noted in AS murmurs. As a consequence of difficulties with the definition of the threshold lead to the obtained delimitation errors.

Regarding the heart sounds from *databaseHS* and *databaseProsValve*, in this case they are segmented using the WD-NLF method. It is observed that the delimitation errors are nearly similar to the errors obtained employing the energy envelop method. Furthermore, it can also be observed that the recognition of S1 and S2 sounds yields similar sensitivity and specificity performance. Therefore, both segmentation methods are efficient to segment heart sound clips from the *databaseHS* and *databaseProsValve*. However, WD-NLF based method performs efficiently in all three databases.

Heart Sound	**S1 and S2 heart sound detection**	
	Sensitivity (%)	**Specificity (%)**
databaseHS		
Healthy	92.02	92.21
Unhealthy	88.08	90.42
Average	*91.01*	*91.33*

databaseProsValve		
Mechanical valve	95.43	94.04
Bio-prosthetic valves	93.30	95.77
Average	*94.40*	*94.66*
databaseMurmur		
AR	54.12	60.12
AS	46.98	61.04
MR	35.66	50.34
PR	60.22	64.33
PS	84.54	88.13
SAS+VSD	95.11	89.88
SYE	73.01	82.12
Average	*66.87*	73.82
Global average	*81.76*	*85.51*

Table 5.3 S1 and S2 recognition performance in terms of Sensitivity (SE) and Specificity (SP): *databaseMurmur* is segmented using the energy envelop method and *databaseHS* and *databaseProsValve* by WD-NLF algorithm.

Heart Sound	**S1 error μ ($\pm\sigma$) (msec)**		**S2 error μ ($\pm\sigma$) (ms)**	
	Start time	**Stop time**	**Start time**	**Stop time**
databaseHS				
Healthy	6.68 (1.28)	4.23 (2.30)	5.95 (1.73)	5.21 (2.12)
Unhealthy	5.68 (0.96)	5.16 (0.91)	6.22 (3.08)	6.59 (3.08)
Average	*6.18 (1.12)*	*4.69 (1.60)*	*6.08 (2.40)*	*5.91 (2.60)*
databaseProsValve				

Mech. Valve Biop. Valves	3.90 (2.71) 2.01 (0.92)	6.52 (3.51) 4.12 (1.61)	3.78 (2.28) 3.12 (2.06)	14.13 (5.13) 3.14 (1.85)
Average	*2.95 (1.81)*	*5.32 (2.56)*	*3.45 (2.17)*	*8.64 (3.49)*
databaseMurmur				
AR	15.67 (5.84)	18.98 (8.44)	21.40 (2.96)	15.12 (5.65)
AS	12.56 (6.34)	25.42 (12.14)	22.21 (3.64)	23.48 (5.46)
MR	14.15 (9.62)	23.33 (11.08)	18.47 (6.14)	16.32 (1.92)
PR	12.72 (5.32)	16.89 (5.01)	6.10 (2.56)	4.21 (2.18)
PS	7.86 (4.78)	12.11 (4.88)	6.14 (5.76)	11.22 (9.02)
SAS+VSD	4.93 (1.86)	3.72 (2.44)	2.5 (0.19)	2.12 (0.34)
SYE	1.18 (0.50)	2.14 (0.98)	2.22 (1.22)	5.08 (1.10)
Average	*9.87 (4.89)*	*14.66 (6.42)*	*11.29 (3.21)*	*11.08 (3.66)*
Global Avg.	*6.33 (2.61)*	*8.23 (3.53)*	*6.94 (2.59)*	*8.54 (3.25)*

Table 5.4 S1 and S2 delimitation errors (average value with standard deviation): *databaseMurmur* is segmented using the energy envelop method and *databaseHS* and *databaseProsValve* by WD-NLF algorithm.

The overall performance is measured in terms of the average delimitation error and recognition rate. Errors in start and stop times for S1 and S2 from Table 5.1 and Table 5.3 are averaged to compute overall delimitation error, which is found to be 6.656 ($\pm$3.721) ms. In addition to this, sensitivity and specificity are computed by averaging its values from Table 5.2 and Table 5.4. Averaged sensitivity and specificity are obtained as 84.98 ($\pm$17.98) and 86.41($\pm$13.97), respectively. As a conclusion, it becomes clear that the overall performance of the proposed method strongly depends on the initial classification stage. Particularly, misclassified murmur sounds might degrade the accuracy of the method substantially.

From the results shown in Table 5.1 to Table 5.4 it is observed that heart sounds containing mainly two components are comparably easier to segment than in case of abnormal sound present in the systolic or the diastolic region. However, it is evident from the described methods that one feature is not sufficient to segment heart sound components in presence of abnormal sounds. In the WD-NLF algo-

rithm the nonlinear features, simplicity and strength, are crucial to capture information regarding the duration of S1 and S2. In **Figure 5.19** segmentation using WD-NLF method on normal heart sound and heart sound from prosthetic valve is shown. It can be observed from **Figure 5.19**(b and c) and **Figure 5.19**(g and f) that the simplicity curve and strength curve show hikes for S1 and S2 sounds. Now, it can be noticed from the *Sm* curve that it may not be sufficient to distinguish S1 and S2 sounds in one heart cycle. Some parts of the heart sound such as non-audible sound or background noise in the systolic and diastolic regions also exhibit high simplicities. Subsequently, when the PPA is applied to apply a suitable threshold to discriminate high simpler components from lower simple ones, many non-cardiac or background noises are captured. This problem is solved using the strength feature. From **Figure 5.19**(c and g) strength can be observed as representing explicitly S1 and S2 by showing their high values. In both the curves S1 and S2 sounds exhibit high strength and simplicity which can be preceded with the PPA threshold, when finding an approach to mark the start and stop instants of S1 and S2 sounds. Sometimes, due to the presence of high frequency component heart sound induced by single tilted disk mechanical prosthetic valves, it displays extremely low simplicity (low valley in the simplicity curve). These contents exhibit such a low simplicity that it goes below the threshold found by PPA, while strength of this part is large enough to be above the threshold. Therefore, there is no common area of timings between *Sm* and *Sg* curve after thresholding such situations. As a consequence delimitation error is relatively higher than in normal heart sounds.

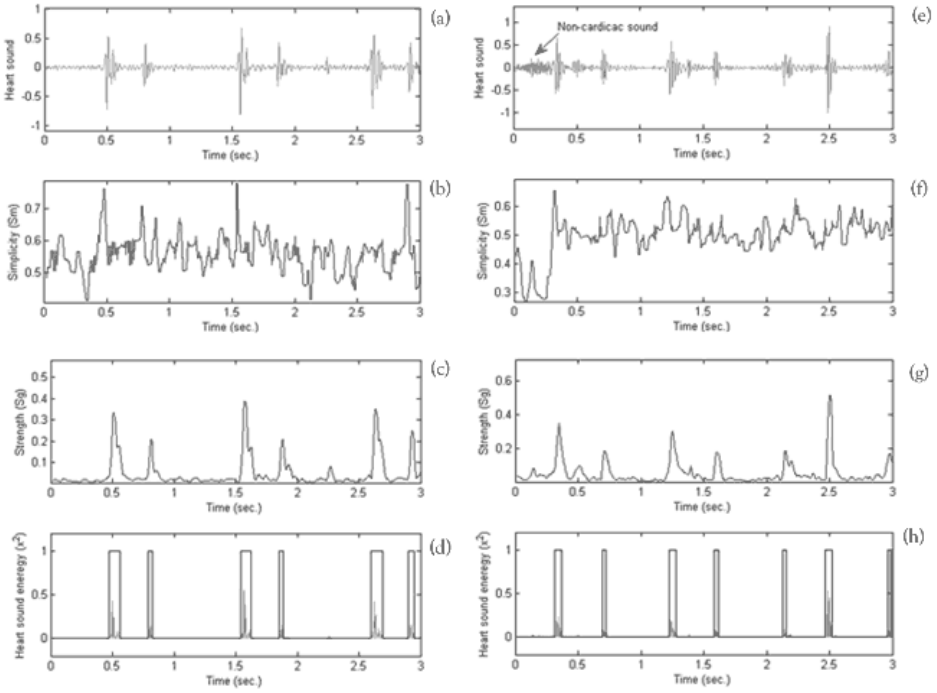


Figure 5.19. Heart sound segmentation using the WD-NLF method: In the first column (a-d) normal heart sound from native valve is segmented. In the second column (e-h) Heart sound from prosthetic valve is segmented.

Table 5.3 and Table 5.4 demonstrate that heart sound with murmur are not segmented with high accuracy using the energy envelop approach. It can also be observed that in the heart sound in case of severe aortic stenosis Shannon energy reflects high energy peaks at many places, besides the location of S1 and S2. Since Shannon energy of the samples is dependent on their amplitude, it is seen that loud murmur sounds (>= grade II/VI) exhibit large Shannon energy. As a result of this, during the thresholding operation of Shannon energy heart murmur sounds are also segmented. Many of these segments may not be discarded in the process of segment validation. An example of a heart sound with murmur segmented by the energy envelop algorithm is shown in **Figure 5.20**. As can be observed, a single feature such as the Shannon energy is incapable to delimit murmur regions from S1 and S2 sounds. Therefore, it is deduced that WD-NLF method is a significantly more efficient method to segment heart sound with murmur as well as normal heart sound and heart sound with murmur.

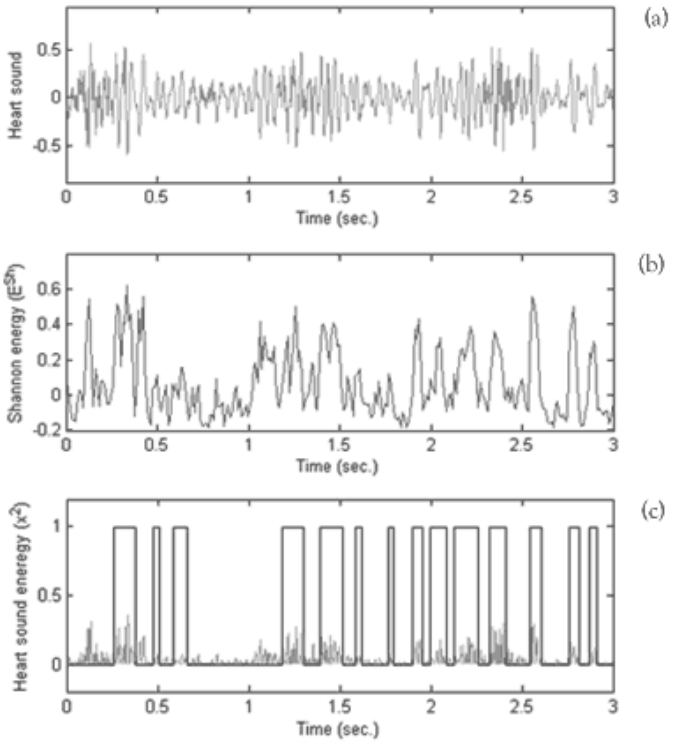


Figure 5.20 Heart sound with murmur (Aortic stenosis) segmented by energy envelop method.

B. Segmentation using Envelogram of Liang et al. and Nigam and Priemer's Simplicity based methods

In order carry out a comparative study between our methods and methods proposed by Liang et al. and Nigam and Priemer we follow the exact way theirs results' representations, i.e. Liang et al. presented the results in the form of sensitivity and specificity of the S1 and S2 sounds whereas Nigam and Priemer method showed the delimitation of the these sounds in presence of murmurs. The envelogram of Linag et al. attempted to find out S1 and S2 sounds correctly without caring much on the onsets and the offsets of these sounds while Nigam and Primer's simplicity approach focused on the identification of the S1 and S2 instants.

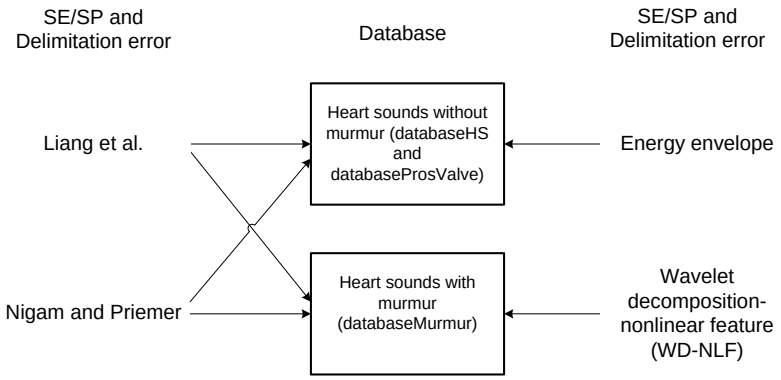


Figure 5.21. Scheme for Liang et al. and Nigam and Priemer segmentation methods' results with the energy envelop and WD-NLF.

Heart Sound	**S1 and S2 heart sound detection**	
	Sensitivity (%)	**Specificity (%)**
databaseHS		
Healthy	85.57	91.65
Unhealthy	87.22	86.73
Average	*86.19*	*88.64*
databaseProsValve		
Mechanical valve	88.60	91.22
Bio-prosthetic valves	92.10	95.70
Average	*91.14*	*93.19*
databaseMurmur		

AR	40.16	52.73
AS	49.60	56.19
MR	37.55	47.66
PR	64.36	63.65
PS	85.72	83.75
SAS+VSD	95.02	84.28
SYE	68.21	83.90
Average	*66.87*	6.92
Global average	*81.76*	*80.54*

Table 5.5 S1 and S2 recognition performance in terms of Sensitivity (SE) and Specificity (SP) by the Liang et al method of envelogram.

The results presented in the Table 5.5 and Table 5.6 are obtained using the Liang et al segmentation method. In the envelogram approach the two parameters c1 and c2 which represent tolerance in systolic and diastolic periods during the identification of S1 and S2 sounds are set to the same values as described in the original method. Though in the original work Liang et al method achieved significant results by segmenting heart sound with murmurs it does not show comparable results in our murmur database. The main reason for this behaviour is due to the murmur grades in our database that includes sounds clips with murmur of various grades (grade I-VI/VI). Therefore, Shannon energy of murmur components are significantly different from the S1 and S2 sounds' energy, which leads to omit many real peaks that correspond to S1 and S2 sounds. On the other hand, this approach is relatively efficient in segmenting a heart sound recording that contains mainly two heart sounds components such as those present in *databaseHIS* and *databaseProsValve*. As it can be noticed, in presence of two main heart sound components, Shannon energy peaks are well defined (see **Figure 5.19**). Table 5.6 presents the delimitation error achieved with Liang et al method. It is evident from the table that *databaseHIS* and *databaseProsValve* are more correctly segmented than *databaseMurmur*. Shannon energy was effective in emphasizing the S1 and S2 sound locations when heart sound has mainly two components (see **Figure 5.20**), whereas in the presence of murmur several segments in the systolic/diastolic region are identified and are often misinterpreted as S1 and S2 sounds during the recognition stage. Hence, high grade murmurs, such as AR, MR and PR in *databaseMurmur*, produce

larger errors compared to the results achieved using the sounds from the other databases as well as low grade murmurs from *databaseMurmur*.

Heart Sound	S1 error μ ($\pm\sigma$) (ms)		S2 error μ ($\pm\sigma$) (ms)	
	Start time	Stop time	Start time	Stop time
databaseHS				
Healthy	8.02 (4.9)	7.18 (4.11)	7.29 (4.13)	6.24 (3.27)
Unhealthy	8.86 (5.13)	8.06 (3.65)	6.29 (5.16)	8.15 (4.36)
Average	*8.44 (5.01)*	*7.62 (3.79)*	*6.79 (4.64)*	*7.19 (3.81)*
databaseProsValve				
Mech. valves	4.16 (1.78)	7.06 (2.81)	4.92 (2.76)	7.87 (3.76)
Biop. valves	4.55 (2.13)	4.77 (3.25)	5.39 (3.22)	5.43 (2.02)
Average	*4.35 (1.95)*	*5.91 (3.03)*	*5.15 (2.99)*	*6.65 (2.89)*
databaseMurmur				
AR	16.17 (7.13)	26.39 (9.71)	25.21 (11.13)	17.14 (9.15)
AS	13.87 (6.98)	27.71 (15.23)	24.75 (10.66)	25.38 (9.44)
MR	12.11 (7.38)	26.41 (13.11)	20.31 (9.10)	23.12 (10.81)
PR	11.03 (5.24)	18.15 (7.77)	10.18 (4.02)	9.13 (6.43)
PS	8.87 (4.99)	13.92 (6.88)	6.88 (3.34)	14.73 (8.10)
SAS+VSD	6.12 (3.34)	5.72 (3.15)	6.21 (2.01)	4.82 (1.92)
SYE	3.56 (1.31)	3.44 (1.66)	3.93 (1.92)	8.14 (3.05)
Average	*10.24 (5.19)*	*17.39 (8.21)*	*13.92 (6.03)*	*14.64 (6.70)*
Global Avg.	*7.68 (4.05)*	*9.14 (5.01)*	*8.62 (4.56)*	*9.49 (4.46)*

Table 5.6 Delimitation error obtained from Liang et al method of envelogram.

Nigam and Priemer's approach to segment heart sounds mainly focused on heart sound with various murmurs. However, the authors have not presented results of S1 and S2 detection. Instead they demonstrated the accuracy in gating of murmurs within the systole and the diastole regions. Hence, we apply our approach (high frequency signature) to recognize S1 and S2 sounds after delimiting sounds using Nigam and Priemer's approach. This method was experimented with a simulated database that is available in "http://egeneralmedical.com/list-ohearmur.html" (last seen on 10th October 2012), which contains a range of heart sounds with murmur and various other sounds originated from cardiac conditions. Following similar steps as described in the original approach, we tested the method using our *databaseMurmur*. The achieved results in terms of delimitation error and recognition of the main heart sounds are shown in Table 5.7 and Table 5.8, respectively.

Heart Sound	**S1 and S2 heart sound detection**	
	Sensitivity (%)	**Specificity (%)**
databaseHS		
Healthy	76.21	81.20
Unhealthy	73.34	85.43
Average	*74.33*	*82.71*
databaseProsValve		
Mechanical valve	84.76	87.10
Bio-prosthetic valves	77.32	79.55
Average	*79.48*	*82.32*
databaseMurmur		

AR	66.24	79.54
AS	73.43	82.32
MR	68.12	77.87
PR	77.65	84.24
PS	83.76	90.57
SAS+VSD	76.54	86.12
SYE	73.23	86.16
Average	*75.18*	83.01
Global average	*76.47*	*84.32*

Table 5.7 S1 and S2 sounds recognition performance in terms of Sensitivity (SE) and Specificity (SP) by the Nigam and Primer method.

It can be observed from Table 5.8 and from the example shown in **Figure 3.5** in Chapter 3 that the onsets of the S1 sounds as well as the offsets of S2 sounds are not identified within an error comparable to the WD-NLF algorithm in all three data-

Heart Sound	**S1 error ± (σ) (msec)**		**S2 error ± (σ) (msec)**	
	Start time	**Stop time**	**Start time**	**Stop time**
databaseHS				
Healthy	13.18 (7.10)	8.36 (6.12)	10.29 (4.13)	6.24 (3.27)
Unhealthy	14.08 (9.15)	9.46 (7.65)	15.60 (6.24)	8.15 (4.36)
Average	*13.63 (8.12)*	*8.91 (6.88)*	*12.94 (5.18)*	*7.19 (3.81)*
databaseProsValve				
Mech. valves	16.34 (8.23)	13.46 (6.81)	11.36 (5.19)	7.87 (3.76)
Biop. valves	10.55 (4.22)	12.22 (3.25)	8.48 (4.78)	5.43 (2.02)
Average	*13.44 (6.62)*	*12.84 (5.03)*	*9.92 (4.98)*	*6.65 (2.89)*
databaseMurmur				

AR	36.12 (28.36)	8.24 (5.52)	8.81 (5.31)	46.35 (28.92)
AS	43.22 (31.30)	16.25 (7.83)	7.87 (5.10)	54.36 (27.76)
MR	21.75 (15.51)	13.89 (8.87)	6.11(3.21)	33.81 (19.38)
PR	19.20 (10.9)	12.35 (7.26)	10.12 (6.81)	39.41 (19.81)
PS	8.30 (4.40)	7.24 (5.63)	7. 35 (4.22)	6.34 (3.12)
	4.98 (2.08)	5.12 (2.10)	3.45 (1.14)	4.22 (2.04)
SAS+VSD	25.34(12.78)	7.25(3.46)	11.13 (4.56)	29.64 (13.37)
SYE				
Average	*22.70 (15.04)*	*10.04 (5.81)*	*7.83 (4.33)*	*30.63 (16.62)*
Global Avg.	*16.59 (9.93)*	*10.59 (5.91)*	*10.23 (4.83)*	*14.82 (7.77)*

Table 5.8 Delimitation error from the approach of similarity measure Nigam and Priemer approach.

bases. The reason behind this is that the simplicity curve computed from the heart sound of *databaseMurmur* shows high values at places of S1/S2 sound and systolic/diastolic region where murmurs are not present. In absence of high frequency background in systolic or diastolic regions, it is observed that simplicity values tend to be similar to those observed in S1 and S2 leading to wrong threshold adjustments. Subsequently, recognition of S1 and S2 sounds is influenced by the large delimitation error (see Table 5.7).

We attempt to compare the results obtained by the methods proposed by us and the methods of Liang et al. and Nigam and Priemer separately. Since our methods aim to provide information regarding delimitation of the sounds as well as recognition of the main sounds, i.e., S1 and S2, we collectively compare our detection rate of S1 and S2 as well as delimitation error with the Liang et al.'s envelogram method and with the Nigam and Priemer's simplicity based algorithm, respectively. The comparison process follows the strategy described in **Figure 5.21**, which describes the outcomes of the segmentation using Energy envelop algorithm on *databaseHS* and *databaseProsValve* are compared with the Liang et al. and Nigam and Priemer algorithm on the same database. Similarly, the segmentation outcomes using WD-NLF on *databaseMurmur* are compared with the ones from Liang et al. and Nigam and Primer. This strategy provides relevant comparison between the Energy envelop method which was developed for heart sounds from healthy subjects (or heart sound containing mainly two components, i.e. S1 and S2) with the envelogram as well as simplicity approach.

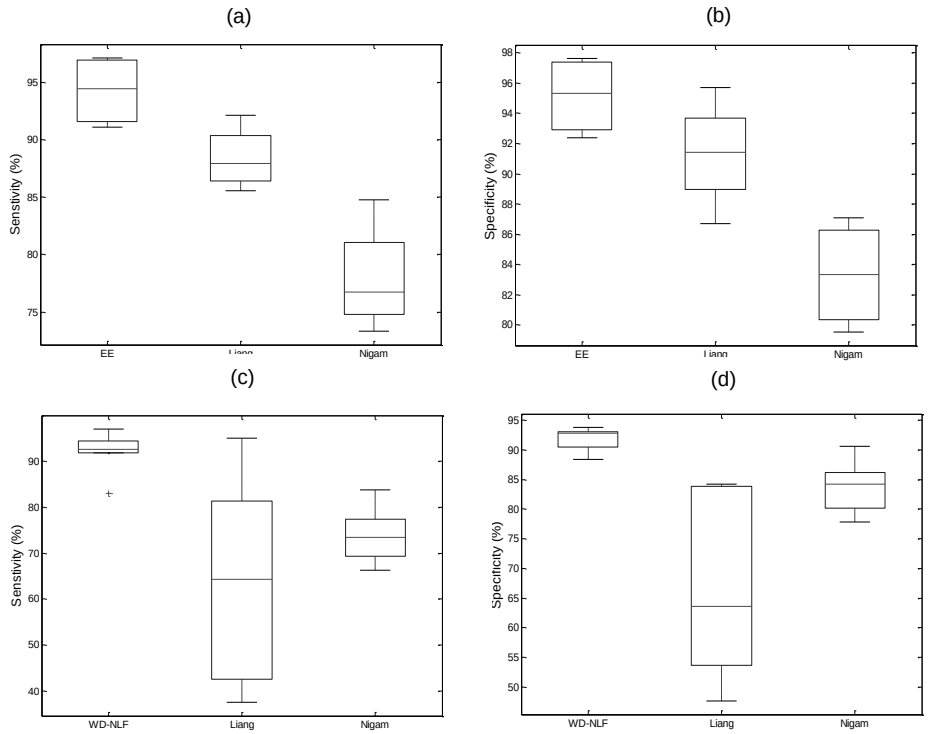


Figure 5.22 (a) and (b) are the Box-Whisker plot of the sensitivities and specificities obtained from Energy envelop (EE), Liang et al. and Nigam and Priemer methods on *databaseHS* and *databaseProsValve*. Similarly, (c) and (d) are the Box-Whisker plot of sensitivities and specificities from WD-NLF, Liang et al., and Nigam and Priemer method on *databaseMurmur*.

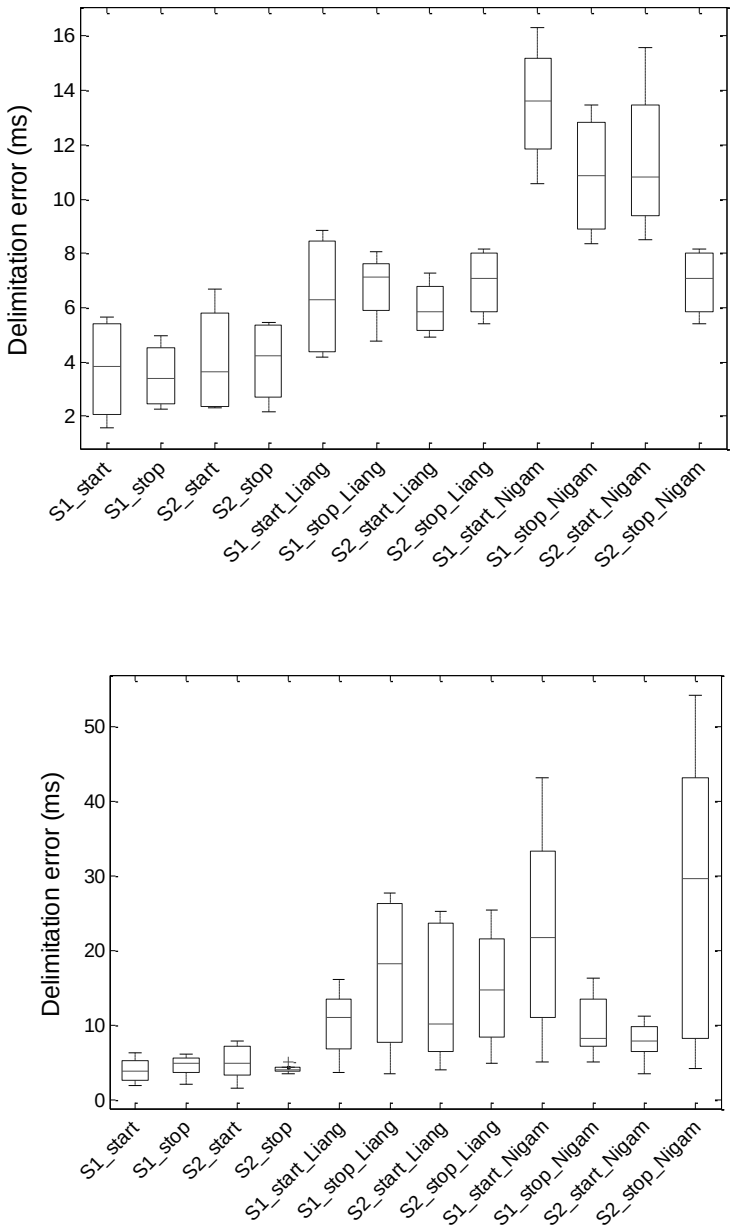


Figure 5.23 (Upper) Box-Whisker plots of the errors in start and stop times of S1 and S2 sounds obtained from EE approach, Liang et al., and Nigam and Priemer's simplicity approach. S1_start/stop and S2_start/stop are the instants of S1 and S2, respectively, from our methods, and S1_start/stop_Nigam are from Nigam and Priemer's approach. (Lower)

Box-Whisker plots of WD-NLF approach, Liang et al., and Nigam and Priemer's approach.

The detection rate of S1 and S2 sounds in terms of sensitivity and specificity by our method along with envelogram based approach of Liang et al. and simplicity based approach of Nigam and Primer are shown in **Figure 5.22**. This figure suggests that the spread in the results achieved by the envelogram approach is much more than our approach due to its low performance with the clips of heart sounds with murmur. While Nigam and Primer's approach exhibits less sensitivity and specificity when is applied on *databaseHS* and *databaseProsValve* due to large simplicity values for non-audible sounds in diastolic regions besides the main heart sound components' values. Our methods clearly outperform the both state-of-the-art methods by exhibiting better sensitivities and specificities. These sensitivities and specificities obtained were tested to find out distribution difference using suitable statistical tests. In this process, first we found out if sensitivities or specificities are Gaussian distributed, then Gaussian ones are tested with t-test otherwise rank (Mann-Whitney) test is performed (Prel et al. 2013). Therefore, sensitivities and specificities resulted from our approaches applied on *databaseHS* and *databaseProsValve* and Liang et al. and Nigam and Primer were compared with t-test. Sensitivities of EE approach showed significant difference from Liang et al. method by rejecting hypothesis of being similar distribution with p-value, p=0.0304, and from Nigam and Primer with p=0.0013 (significant level < 5%), respectively. Similarly, specificities were different with p=0.013 and p=0.0017, respectively. On the other hand, the sensitivities and specificities obtained from WD-NLF method, Liang et al. and Nigam and Primer were non-Gaussian and hence, were tested with Mann-Whitney U-test. Sensitivities from WD-NLF was different from Liang's with p=0.02 and Nigam and Primer with p=0.0001, respectively, while specificities were different with p=0.0005 and p=0.0023, respectively. Therefore, it can be deduced that EE and WD-NLF methods are significantly different, by showing better results, from the other two state-of-the-art methods.

The second comparison is drawn in terms of delimitation error which is shown in **Figure 5.23**. It is evident from the figure that there is a large difference between the average delimitation errors of and our proposed methods and the other two methods. Therefore, if the test is performed with the suitable statistical test as to measure the significant difference probably of rejecting the hypothesis will be very high ($p<0.05$). In order to examine the difference, the error obtained via each method is averaged and then test is performed. For instance, the onset and offset instants of EE method, i.e. S1_start, S1_stop, S2_start, and S2_stop are averaged, and then the averaged values are subjected for test to determine statistical difference between the delimitation errors of our methods, Liang et al. and Nigam and Priemer methods. Since averaged delimitation errors in our approaches as well as in Liang and Nigam and Priemer approaches show Gaussian distribution a student test is per-

formed. The p-value for EE method is 0.0421 with Liang et al. and 0.0004 with Nigam and Priemer. However, WD-NLF results in a p-value of 0.0056 with Liang et al. and 0.00030 with Nigam and Priemer, respectively.

5.10. Conclusions

A method composed by two sub algorithms for heart sound segmentation has been proposed based on nonlinear features and energy envelop extraction in the wavelet domain. This method aimed to segment three types of heart sound: normal heart sound produced by native valves, heart sound from prosthetic valves and heart sound with murmurs. First, the sound with murmur is discriminated based on the largest Lyapunov exponent. Then, heart sound without murmur is segmented by the energy envelop algorithm, whereas sound containing murmur is segmented by the method of wavelet decomposition-nonlinear features (WD-NLF), using simplicity and strength.

The energy envelop method utilizes approximation coefficients for delimiting sounds. Categorization of the identified segments is performed using a new high-frequency signature scheme. In the WD-NLF algorithm, heart sound components are delimited using simplicity and strength measures in the wavelet domain. The presented algorithm is highly adaptive. Wavelet coefficients are adaptively selected for optimal similarity and strength profile computation. Furthermore, these profiles are adaptively thresholded in order to detect relevant sound segments. In the second phase, the method handles the detection of S1 and S2 sounds. An approach based on a high-frequency signature is utilized to identify S2; subsequently, S1 is detected with a set of knowledge-based rules defined on the basis of the heart physiology.

The evaluation of the proposed method showed significant accuracy in the recognition of S1 and S2 sounds and in the detection of their start and stop times under highly challenging conditions. This is a significant result, since performance of the most heart sound segmentation methods tend to breakdown for murmur segmentation. The obtained results were also compared with two well known methods for heart murmur delimitation and heart sounds component recognition. Our approach outperformed Liang et al.'s envelogram method as well as Nigam and Priemer's simplicity based heart murmur gating method.

Chapter 6

HEART MURMUR CLASSIFICATION

Heart murmur are originated from different types of cardiac conditions. Traditionally, classification of the murmur type is one of the main tasks in heart sound analysis. It aids physicians in assessing murmurs and to interpret its origin. In this chapter we propose a method for murmur classification. Classification of murmurs is usually approached using discriminative features followed by a supervised learning based classifier. In this chapter, some discriminative features are investigated for efficient and robust classification. The prepared set of features is compared with two sets of features used in the literature that have achieved a significant degree of correct classification in different databases.

Section 6.3 Cardiac Lesions Classification through Murmurs

In this section, we present the roadmap to classify cardiac murmurs following feature extraction from each murmur and assessing its type using a classifier.

Section 6.3 Pre-processing and Beat detection

In this section, the important steps taken to pre-process the heart sound containing murmurs before feature extraction are discussed. Murmur is present in the systolic or diastolic regions of almost every heart beat of every heart sound recoding. Therefore, the common strategy in murmur classification is to identify each heart beat (S1-systole-S2-diastole) then extract the features from each systole or diastole regions. This method also discusses the method used for heart beat (or heart cycle) detection.

Section 6.3 Feature Extraction

We then address the techniques of feature extraction from each beat of a heart sound recording. Heart murmurs exhibit different characteristics in terms of time, frequency, complexity, morphology, etc., thus methods are adopted to compute these properties. The methodologies for the features computation are introduced.

Furthermore, two previous methods of feature extraction are introduced in order to draw a comparison between the state-of-the-art results and our results by using the in-house database. The features utilized in the approaches proposed by Olmez (Olmez and Dokur 2003) and Ahlstrom are also descried.

Section 6.4 Feature Selection

After generating features from each heart beat, the following logical customary task is to select the most uncorrelated features by reducing redundant features. We used the recent sequential floating feature selection (SFFS) method to perform selection of the discriminative features. Hence, the method for feature selection and the list of the selected features are introduced in this section.

Section 6.5 Results and Discussions

The number of subjects in our test database, their specification related to age BMI, etc, and heart sound recordings from different cardiac lesions (types of murmurs) are described in this section. Finally the employed support vector machine classifier for classification between the different murmur's classes is presented.

6.1. Cardiac Lesions Classification through Murmurs

Usually heart murmurs are classified after extracting significant discriminative features. The basic steps for classifying heart murmurs constitute the processing blocks shown in **Figure 6.1**. According to this Figure heart sounds are first preprocessed aiming to remove extraneous sounds mixed with them during auscultation. In the second step heart beats are identified. This task is important to perform at this stage because auscultation time is advised to be more than the duration of a heart beat in order to assess correctly the nature of the heart sounds. It is convenient to demark heart beats in a heart sound recording and investigate murmur types beat-by-beat. Sometimes one heart sound waveform for one beat shows different

characteristic from the other that may happen due to mechanical and hemodynamic changes of heart's function. The third block includes the task related to the exploration of the properties from the sound samples of each and every beat. Properties exploration is known as feature extraction in the field of machine learning. Taking this term into consideration, samples of each beat are subjected to feature extraction methods in order to perform a dimensional reduction of the data, while preserving the most significant dimensions to discriminate to find features that highlight differences between heart sounds. These feature extraction techniques have been extensively investigated on various groups of data, besides heart sound or other biomedical related data. The fourth and final block is related to the task of classification. For this, the classifiers which can be employed are artificial neural network, Gaussian mixture model, support vector machine, etc. These ones use supervised learning methods.

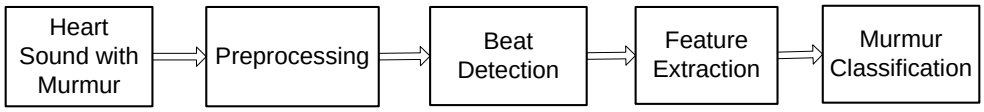


Figure 6.1 Steps blocks for heart murmur classification.

These blocks of processing have been followed by many researchers in the course to solve the heart murmur classification issue. The main challenge in the context of heart murmur classification is feature extraction. In the past recent years many researchers have attempted to solve this problem and thus far some significant results have been achieved for different databases. Some of the major works performed in murmur classification were introduced in the Section 3.3 of the Chapter 3.

Two state-of-the-art methods introduced recently are considered for a comparative study with respect the proposed method. The first algorithm considered, by Olmez *et al.* (Olmez and Dokur 2003), extracts 32 features from a wavelet decomposed heart sound and uses all the features to construct a nonlinear classifier. The second method was introduced by Ahlstrom *et al.* (Ahlstrom 2006) and resorts to 207 extracted features using several signal processing and statistical methodologies. In order to check the discriminative value of the proposed set of features, the achieved results are also compared against the classification results achieved with the best feature set selected by SFFS among all considered features in this work, i.e., the proposed feature set and the feature sets proposed by Olmez *et al.* and by Ahlstrom *et al.*.

6.2. Preprocessing and Beat Detection

This part of the heart murmur classification method is partly taken from the previous chapter, i.e., heart sound preprocessed according to the procedure adopted in the segmentation process described in chapter. It is essentially that heart sound is passed through a high pass filter FIR filter to remove low frequencies which are captured during auscultation. Heart sounds usually contain frequencies above 25 Hz. However, lower frequencies are due to muscle movements, chest skin movements, and movements of the objects. To remove these frequencies any FIR high pass filter can be used. A Butterworth with cut-off frequency of 25 Hz was used in our work.

The other part is included in this chapter in the preprocessing phase which is about the heart cycle (or heart beats) detection during processing heart sound. Since heart murmur is present in each heart cycle of the heart sound, it is required to mark the heart cycles in order to see the length and to assess heart murmur presence. In our work, each heart cycle is identified, and then the feature extracted process is started. One heart cycle consists in the sequence of S1-systole-S2-diastole. In the process of murmur identification, if locations of S1 and S2 sounds are correctly marked then locations of the murmurs can also be found using the reference of S1 and S2. The methods proposed for segmentation earlier in the previous chapter are used to demarcate the regions of S1 and S2 sound in a heartbeat.

6.3. Feature Extraction

Murmur exhibit characteristic features which correlate to their origin. A range of features is pivotal to classify heart murmurs with the maximum possible accuracy. Therefore, it is required to find the ideal set of discriminative features. As was mentioned at the beginning of this chapter, we propose the introduction of a new set of features and combine it with two sets of features from the literature in order to find suitable features that may provide adequate level of discrimination for murmur classification. Hence, in the following subsections four feature sets will be presented: the first set is the new one proposed in our work, the second one is the set of features introduced by Olmez *et al.*, the third set is the one prepared by Alhstorm *et al.*, and the fourth is the set of features that includes all the features from the previous three sets. Many properties, such as morphological properties and spectral content, of heart murmur are mostly evident in specific domains. In the proposed approach, three main computational domains are explored to construct this set of features. The extracted features are evaluated with the sequential floating feature selection

(SFFS) method (Ververidis 2008). This process reduces the number of features in the feature set by selecting discriminative features. After feature selection, the selected features are used to train a nonlinear support vector machine (SVM) classifier.

To demonstrate the methods for feature extraction the discrete representation of heart sound, $x(n)$, will be applied, where n denotes index of samples and $n = 1, 2, 3, \ldots \ldots \ldots, n_i, n_i + 1, \ldots n_e \ldots N$, where N is the number of samples in one cycle, i.e. from the starting sample of S1 sounds to the end of S2 sounds and n_i and n_e are the initial and end samples, respectively, of the murmur. Features are computed from the murmurs and then normalized to zero mean and unit standard deviation.

6.3.1. Temporal Domain Features

A straightforward way to generate discriminative features for murmur discrimination is using the time domain representation of the signal. Many of these features are inspired on the field of voice/speech recognition and music genre classification (Dargie and Denko 2010; Gajic and Paliwa 2003; Kim and Stern 2010; Pope et al. 2004; Saunders 1996; Zhou et al. 2008). These exhibit the advantage of being less computationally demanding compared to features extracted from alternative domains, while preserving significant discriminant ability for murmur classification. For instance, these features carry information regarding morphology, intensity and loudness. Below, the methods of computation of these time domain features are explained.

Loudness

Loudness is a property of the sound that is correlated with the strength of the sound and strength is directly related to the amplitude of the samples of the sound waves. The severity of the source of the heart murmur, such as stenotic or septal defects, are normally assessed based on the degree of their loudness (Desiardins et al. 1995). Most murmurs, namely for systolic murmurs, are divided into six categories depending on the sound's intensity, i.e., loudness is the criterion applied to grade audibility of the murmur. However, this does not provide significant information to make a distinction between the types of murmurs. Unlike systolic murmur, diastolic murmurs are often pathological murmurs and can not be categorized based on loudness (Erickson B. 2003; Satpathy and Mishra 2008). The loudness of a sound sample $x(n)$ is computed as in (6.1), where, n_e and n_i are indexes of the starting and ending sound samples, respectively, of the murmur.

$$loudness = \sqrt{\frac{\sum_{j=n_i}^{n_e} x(n)^2}{n_e - n_i}} \tag{6.1}$$

To avoid amplification gain dependencies, loudness applied throughout this work (denoted as *loudness 1*) is normalized by subtracting mean and deviding by standard deviation.

Jitters

This feature is widely used in speech recognition; it assesses the perturbation in period-to-period of the fundamental frequency of the voice (Li et al. 2007; Titze 1994; Vasilakis and Stylianou 2009). Fundamental frequency (F0) is the inverse of the period of a periodic or semi-periodic sound signal. Furthermore, the spectrum of the signal shows energy concentration at integer multiples of the fundamental frequency (Klapuri 2003). Unlike the main heart sound components, murmur are long duration segments which can be assume to exhibit some degree of periodicity. Therefore, measuring irregularities within the periods may render substantive information regarding the murmur's nature. To the best of our knowledge, this measure has not yet been used for heart sound analysis.

One of the most used methods for estimating the fundamental frequency is through autocorrelation of the sound signal. It is computed based on what the sound represents and its autocorrelation notion in periodicity during heart sound acquisition. High frequency long duration sounds may be mixed with heart sounds. In order to identify these corrupted sound segments with periodic disturbances, jitter is computed (Thomson and Chengalvarayan 1998). It represents the perturbation in stochastic and temporarily long sustained sounds when valves do not close normally or due to other external disturbances. As explained in **Figure 6.2**, let $T_p(k)$ be the time separation between the k^{th} and the $(k+2)^{th}$, where k=1, 2, 3, …..N_T, are local maxima of the autocorrelation of the heart sound signal then jitter in the heart sound segments is computed according to (6.2).

$$Jitter = \frac{\sum_{k=2}^{N_T-1} 2T_p(k) - T_p(k-1) - T_p(k+1)}{\sum_{k=2}^{N_T} T_p(k)} \tag{6.2}$$

Autocorrelation is computed of the whole length of the murmur (see **Figure 6.3**). For high frequency murmur, it is observed that jitters produce larger values than for low frequency murmur.

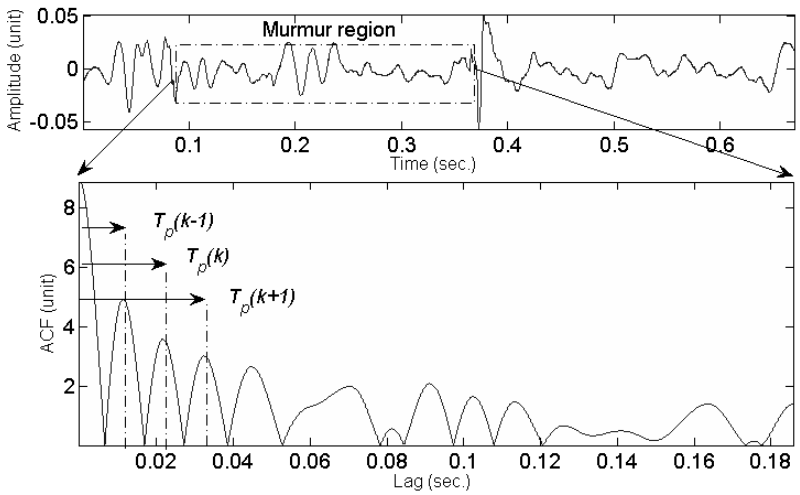


Figure 6.2 (upper) One heart cycle or heart beat of the heart sound from systolic murmur patients is shown. (lower) Autocorrelation function of the murmur region; three consecutive periods (for k=2) that are used to compute jitter.

Transition ratio

It has been observed that the morphology of murmurs in temporal domain can play a crucial role to distinguish one from the others. Transition ratio is a morphological feature that aims to compute proportional change in the times of crescendo and decrescendo in systolic or diastolic murmurs. From **Figure 6.3** it can be seen that the highest energy magnitude is identified in the systolic/diastolic region to obtain the transition ratio. Transition ratio is computed according to (6.3).

$$\mathrm{T}ransition\ ratio = \frac{T_{asc}}{T_{dsc}} \tag{6.3}$$

In (6.3) T_{asc} is the transition time taken from the first minimum of the energy curve to the maximum energy, and T_{dsc} is the time interval from the energy maximum to the last subsequent minimum energy (see **Figure 6.3**).

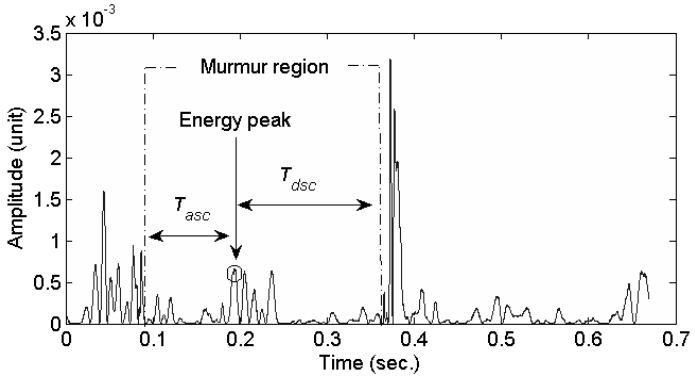


Figure 6.3 An example of systolic ejection. Energy of the heart sound with indication of the ascending and descending time durations.

Zero-crossing rate

This feature is also motivated from the speech recognition area of research where it has been used to detect voiced speech, non-voiced speech and scalene (Atal and Miner 1976; Lau and Chan 1985). Furthermore, being computationally simple and able to provide a rough estimate of dominant frequency it has also been employed to classify speech and music (Panagiotakis and Tziritas 2005), music genres (Tzanetakis and Cook 2002), singing voice and music (Zhang 2003) and various environmental sounds (Cai et al. 2006). One common characteristic of these sound signals is that they exhibit some overlapping frequency range and high values of zero-crossing rate. Zero-crossing rate also plays a significant role in sound detection and classification. Therefore, this feature can be applied to examine its ability in classifying murmurs. The computation method of zero-crossing rate is in a given window according to:

$$zcr = \frac{1}{n_e - n_i} \sum_{n=n_i}^{n_e} \left| sgn\big(x(n)\big) - sgn(x(n-1)) \right| \qquad (6.4)$$

where "**sgn**" is the *signum* function that is 1 and -1 for positive and negative values of $x(n)$ respectively, and 1 for 0. As was previously stated before in this section n_i and n_e are the beginning samples and ending samples, respectively, of the murmur sound.

6.3.2. Frequency Domain Features

Frequency based features play a crucial part in determining abnormal sound and in discriminating amoung types of murmurs. As it is shown in Table 2.1 of Chapter 2,

that frequency range of heart murmurs, in almost all types of murmur, extends from 100 Hz to 600 Hz. Frequencies of the heart murmurs can be an important indicator for identifying its origins as a result of cardiac lesions. The same table highlights that various murmurs show different pitches (blow, harsh, rumble) which enables the computation of frequency contents in order to identify the murmur type.

Three main characteristics are common in an acoustic sound based on which it can be distinguised from other sounds[13](Primentas 2003): *loudness* and *pitch* are quantitative characteristics, while *timbre* is a qualitative aspect of sound. Loudness has previously been introduced. Pitch estimation might be performed in the temporal domain based on the zero-crossing rate (Abdullah-Al-Mamun et al. 2009; Amado and Filho 2008) or in the frequency domain (Kang and Fransen 1994; Rabiner et al. 1976). Timbre or tone quality is a crucial audio feature which enables a listener to recognize the difference between two sounds of the same loudness and pitch (Mackenzie 1964). Therefore, timbre is characterized as a feature or set of features besides loudness and pitch. We attempted to use timbre with the spectrum related features which are derived from the Fourier transform. Let $\chi(f)$ be the Fourier transform of heart sound $x(n)$, once FFT coefficients are obtained, various more expressive features can be extracted, of which some are described below.

Spectral power

Spectral power is estimated as the power spectral density using several known parametric and non-parametric techniques (Porakis and Manolakis 2002). Under the nonparametric method there is a technique of periodogram for estimating power spectrum which is squared of the Fourier transform of the heart sound sample sequence $X(f)$, given as:

$$P(f) = \frac{1}{N}|\chi(f)^2| \tag{6.5}$$

Spectral power provides power carried by the heart sound signal per unit frequency. Since the main power of the main types of murmur is contained in 0-400Hz frequency range, in order to examine the dominance of spectral power at specific frequencies, it is computed in four frequency bands: 0-100Hz; 100-200Hz, 200-300Hz, and 300-400Hz. Hence, 4 features (*spectral power 1-4*), corresponding to the power in each of 4 bands, are computed by summing over frequency.

Spectral Flux

[13] http://www.animations.physics.unsw.edu.au/jw/sound-pitch-loudness-timbre.htm

Spectral flux is a measure of the amount of local spectral change and also has been seen as an important perceptual attribute in the characterization of timbre (Grey 1975). This is defined as the squared difference between the normalized magnitudes of successive spectral distributions as described in (Tzanetakis 2002) :

$$F_t = \sum_{n=1}^{N} \{N_t(n) - N_{t-1}(n)\}^2 \tag{6.6}$$

where $N_t(n)$ and $N_t(n-1)$ are the normalized magnitude of the Fourier transform, the current frame *t*, and the previous frame *t-1*, respectively. In this application, we split the total length of the murmur into two successive frames with the equal length and compute spectral change between the two. This feature is denoted as *spectral flux 1*.

Spectral Centroid

The spectral centroid is a measure of spectral shape and higher centroid values correspond to "brighter" textures with higher frequencies. The spectral centroid is defined as the center of gravity of the magnitude spectrum of the power spectrum from equation (6.5):

$$C_t = \frac{\sum_{n=1}^{N} M_t(n) \times n}{\sum_{n=1}^{N} M_t(n)} \tag{6.7}$$

where $M_t(n)$ is the magnitude of the Fourier transform at frame *t* and frequency bin *n*. This is denoted by *spectral centroid 1*.

Spectral Skewness and Kurtosis

The spectral skewness is a reflection of how the acoustic energy is distributed around the mean. This spectral distribution is sometimes referred to as the "spectral tilt" because it conveys overall slant of acoustic energy **(Nissen 2003).** This is in fact the third order moment of a distribution which is computed according to the following equation:

$$Spectral\ skewness = \sum_{j=1}^{N-1} \left(f_j - \mu\right)^3 P(f_j) \tag{6.8}$$

where f_j is the real-valued frequency corresponding to N-point FFT spectra μ is the mean of the power spectrum $P(f_j)$.

As in the spectral power curve generated from the heart sound, peaks can be noted at certain frequencies. The fourth spectral moment of kurtosis indicates the peakedness of the spectral distribution. A negative kurtic coefficient indicates a relatively flat spectral distribution, whereas a positive coefficient is characteristic of more prominent spectral peaks.

$$Spectral\ kurtosis = \sum_{j=1}^{N-1} \left(f_j - \mu\right)^4 P(f_j) \tag{6.9}$$

These two features are denoted by *spectral skewness 1*and *spectral kurtosis 1*, respectively.

Fundamental Frequency

Fundamental frequency is also referred to as the pitch of a sound. There are several methods for estimating pitch. A comparison between several existing methods can be found in (Rabiner et al. 1976). Among the simplest and fastest estimation techniques are the ones based on the power spectrum, as it is one under which the lowest frequency at which the maximum peak in the power spectrum is detected. There could be more peaks in the power spectrum which correspond to the frequencies that are integer multiples of the fundamental frequencies. Since two peaks are selected from the peaks in the power spectrum it is denoted as *fundamental frequency 1-2*.

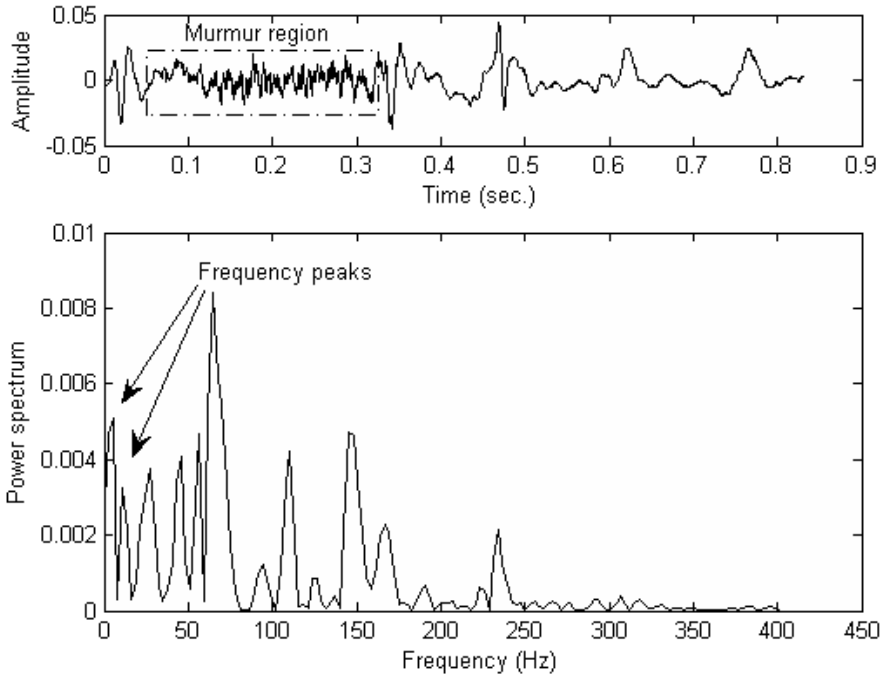


Figure 6.4 (upper) One heart cycle of the heart sound in mitral regurgitation case. (lower) frequency peaks in power spectrum which are considered as fundamental frequencies in our work.

6.3.3. Statistical Domain

The distribution and scattering of samples in the murmur is observed with recourse to histograms and phase space. These domains of features are referred as features extracted directly from the samples or features from the nonlinear dynamical modeling. In fact, some features are computed samples in temporal domain while some are obtained from the phase space which is constructed from the embedding of the signal as defined in section 5.3.1. The following three features are computed:

Statistical Skewness and Kurtosis

The motivation behind the computation of statistical skewness and kurtosis is to extract information regarding the morphology of the murmur components present in the heart sounds. By computing these features directly from the samples it can be argued that even these ones are time-domain features. However, being the

measures to probably density function (statistical inference) we categorize these features as statistical features.

As was already mentioned, skewness measures symmetry in the distribution; positive and negative values determine the distribution spreading out left or right, respectively. It is computed as in (6.10), where μ and σ are the mean and standard deviation of the samples in the given signal.

$$Skewness = \sum_{j=1}^{N} \frac{(x(j) - \mu)^3}{\sigma^3} \tag{6.10}$$

where μ and σ are the mean and standard deviation of the samples in the given signal.

The kurtosis measures the flatness of a distribution, higher values indicate higher sharper peaks and lower values indicate lower less noticeable peaks. It is computed as given as:

$$Kurtosis = \sum_{j=1}^{N} \frac{(x(j) - \mu)^4}{\sigma^4} \tag{6.11}$$

The spectral skewness and kurtosis are denoted ad *spectral skewness 1* and *spectral kurtosis 1*, respectively.

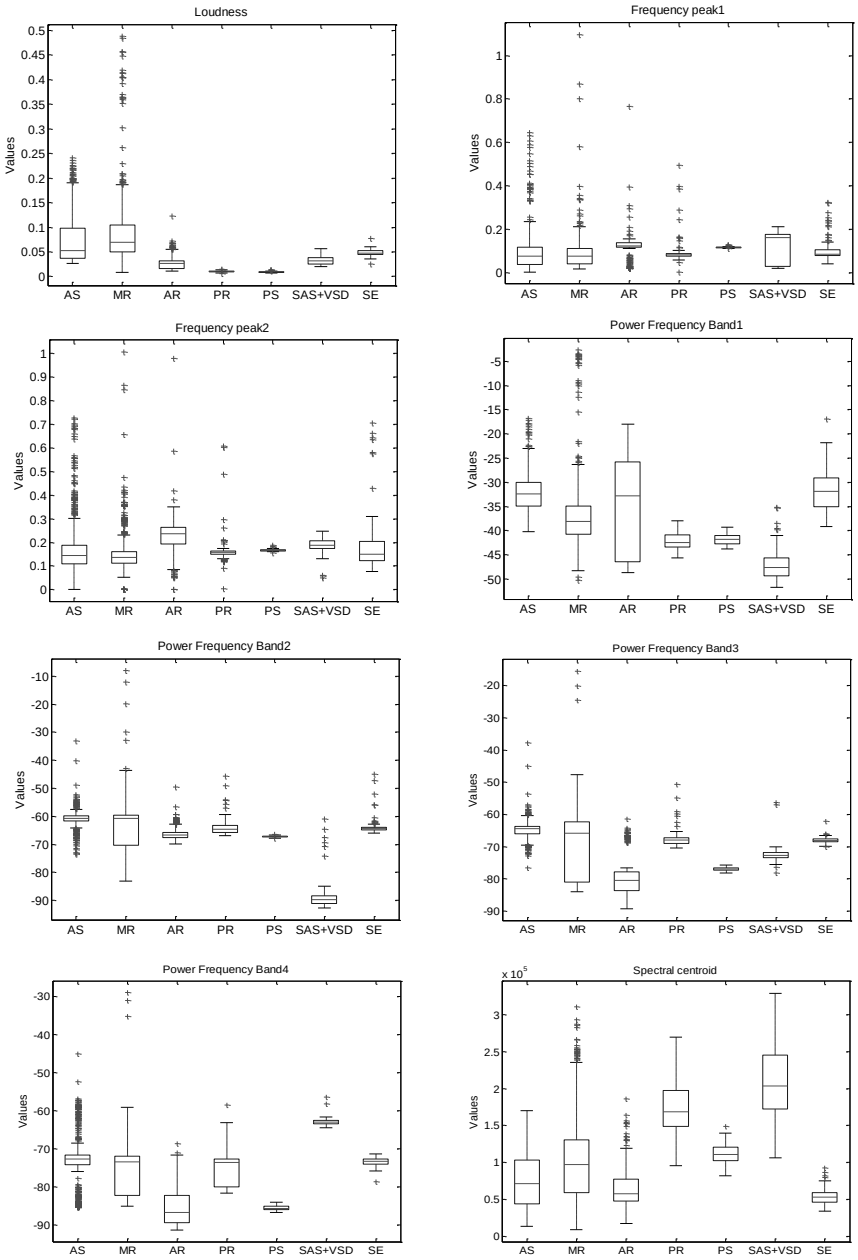
Chaos by Lyapunov Exponents

It has been observed from the chaos measurement, the heart beats with murmur show higher chaos than the heart beats which do not have murmurs in systolic or diastolic phase. Chaos is computed in terms of Lyapunov exponents using the methods explained in Section 5.3.2 of Chapter 5. The largest Lyapunov exponent is computed for each heart beat in the heart murmur database, and denoted as *Lyapunovexponent 1*.

Features	***Number of Features***	***Motivations***
Time domain features		
Loudness (loudness 1)	1	Measure loudness
Zero-crossing rate (zcr 1)	1	Rate of occurrence of the samples
Transition ratio (transition ratio1)	1*	Morphology
Jitter (jitter 1)	1*	Periodic fluctuation
Frequency domain Features		Spectral power in different frequency bands, as it is seen that in different types of murmur power concentration varies in different frequency bands.
Spectral power (spectral power 1-4)	4	
Spectral skewness	1*	
Spectral kurtosis	1*	
Spectral centroid (spectral centroid 1)	1	Frequency concentration and spectral shapes to quantify quality of the sounds
Spectral flux (spectral flux 1)	1*	
Fundamental frequency (fundamental frequency 1-2)	2*	Pitch computation
Statistical domain		
Skewness (statistical skewness 1)	1*	Samples' distribution in the sound segments
Kurtosis (statistical kurtosis1)	1*	
Chaos (Lyapunovexponent 1)	1*	Measure the degree of chaos in the samples of murmur segments

Table 6.1 Summary of the introduced features (Feature set I). And (*) star represents the features which has been used for the heart murmur classification problem.

At last, a summary of the features introduced by us for heart murmur classification is shown in Table 6.1. The table also contains the numbers of the features extracted in each domain along with the properties represented by the features. Each feature


Loudness
Frequency peak1
Frequency peak2
Power Frequency Band1
Power Frequency Band2
Power Frequency Band3
Power Frequency Band4
Spectral centroid
x 10^-5
Values
AS MR AR PR PS SAS+VSD SE

…..Contd.

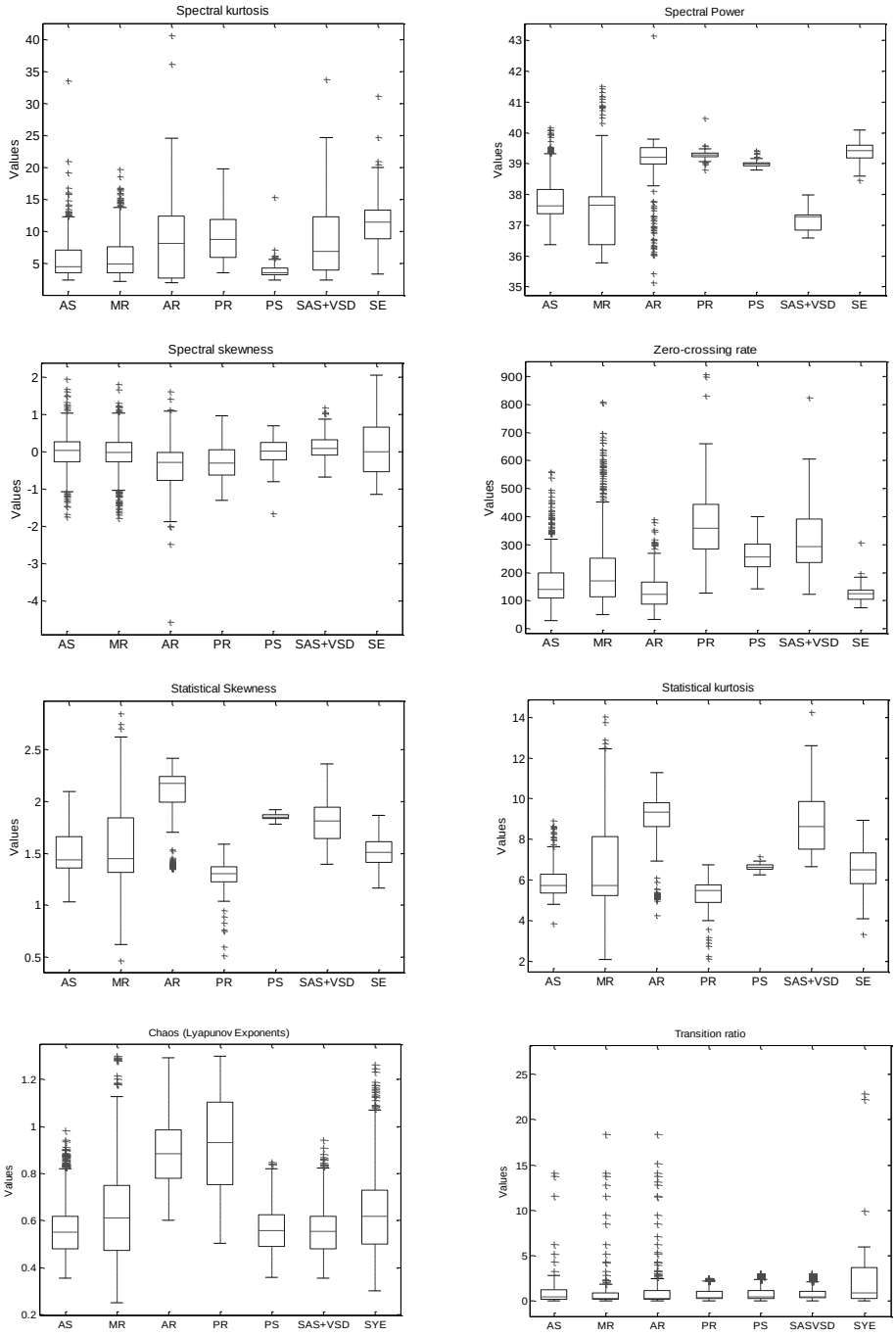


Figure 6.5 Box-Whisker plot of the feature in the 7 classes of the heart murmurs in our prepared database.

contributes a certain grade in the process to discriminate murmurs using the heart sounds. In order to view the spread of the values of the features in each murmur class included in our database (see the Section 6.4), Box and Whisker plots are shown in **Figure 6.5**.

To make it convenient in representation, these features are referred as Feature set-I which can be addressed as a set containing all the features as its elements. Follwing this notation the features extracted from the other two methods, i.e., the feature sets proposed by Olmez et al. and by Ahlstorm *et al.*, are refered as Feature set II and Feature III, repectivley.

6.3.4. Feature Set -II (Olmez *et al.*)

This set of features is generated using Olmez *et al.* (Olmez and Dokur 2003) method of feature extraction. The authors performed wavelet decomposition to represent the heart sounds in different frequency bands. The 2nd details coefficients, d2, of the Discrete Wavelet Transform (DWT) are applied to compute its power the following way: d2 is split into 32 sub-windows with 128 discrete samples each, then using FFT of each window. Hence to construct a feature matrix *(PoS 1-32)*, the power of every split signal for the 32 sub-windows is computed. The *PoS* of a heart beat from each heart sound with murmur is shown in **Figure 6.6**.

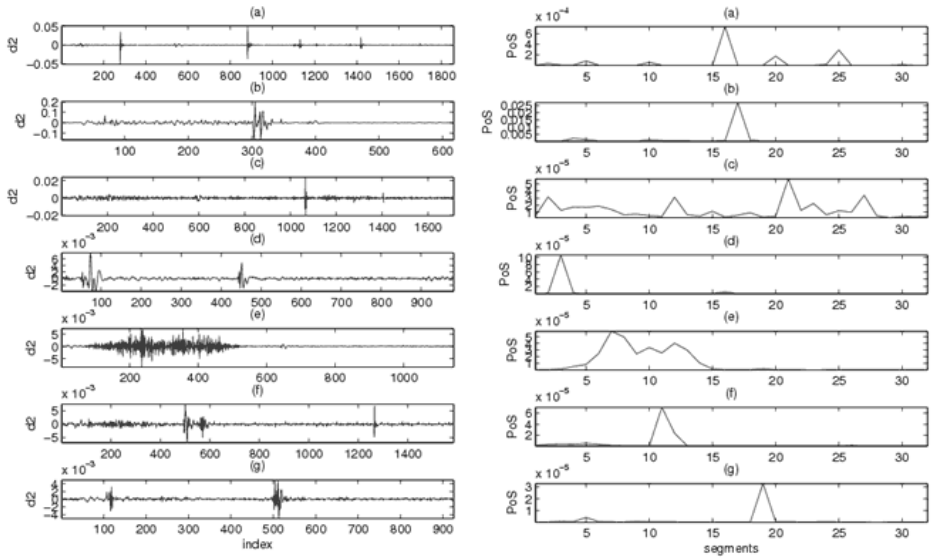


Figure 6.6 (left) The d2 wavelet coefficients of one heart beat of the all heart sound with murmurs, (right) Power of the heart beats; (a) aortic regurgitation, (b) aortic stenosis, (c) mitral regurgitation, (d) pulmonary regurgitation, (e) pulmonary stenosis, (f) sub-aortic stenosis + ventricular septal defect, (g) systolic ejection.

6.3.5. Feature Set -III (Ahlstrom *et al.*)

Feature set 3 is comprised with the feature set suggested by Ahlstorm *et al.*(Ahlstrom 2006b). The authors extracted as many as 207 features using several techniques. Though the motivations behind some of the features were not well defined, the authors intended to construct a new set of features to examine its potential in order to discriminate heart lesions based on heart murmurs. Since the complete description of the features can be found in (Ahlstrom 2006b), only a brief introduction is provided in this section.

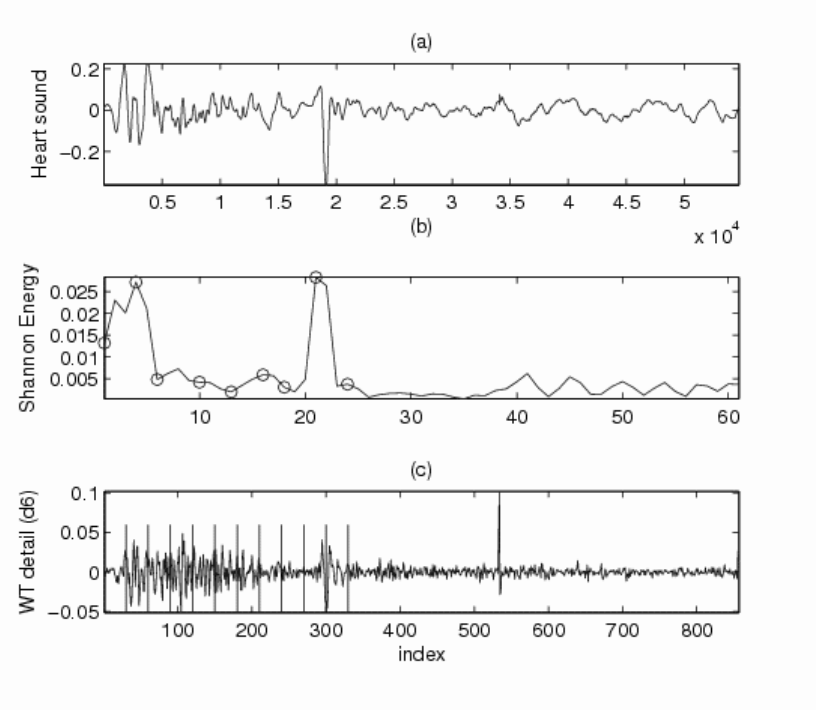


Figure 6.7 (a) One heart beat of the heart sound from aortic regurgitation patient, (b) Shannon energy envelop and the 9 picked values, (c) wavelet decomposed signal, and 11 equally partitioned segments to compute entropy.

In time domain, 9 values of Shannon energy (*Shannon energy 1-9*) are calculated. As

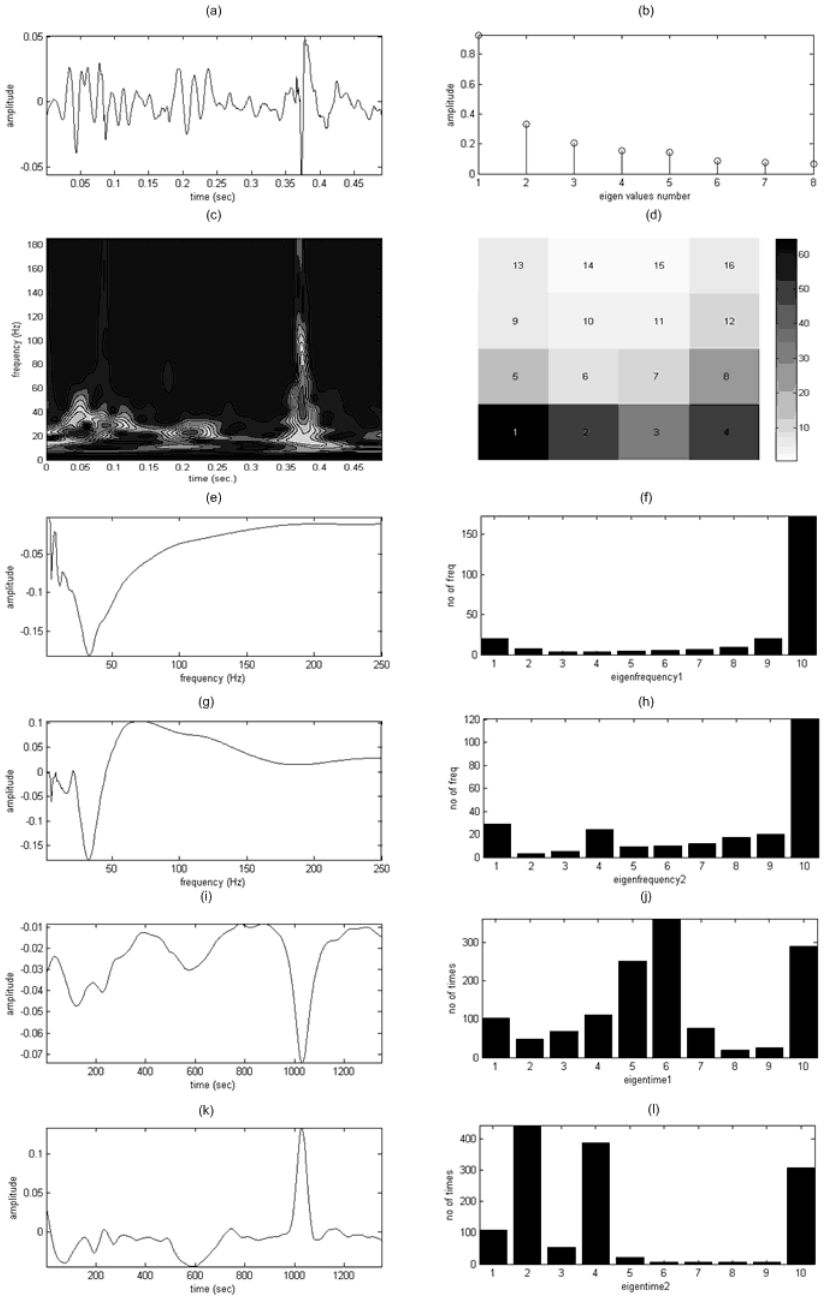


Figure 6.8 Heart sound with murmur from a systolic ejection patient: (a) one heart cycle, (b) *eigenvalues 1-8* from the SVD of S-tranfrom, (c) S-tranfrom (whiter space represente higher intensities), (d) discritisized S-transform into a 4×4 map of features (numbers represent the features denoted *STmap1-16*). The histograms (f, j, h,

and l) of the respective eigenfrequencies (e, g) and eigentimes (i, j) are shown on the right.

it can be seen in the **Figure 6.7**(a), 9 values of the Shannon energy are picked as features based on the main transition points in the heart sound's components. These parts are: times before S1, peak S1, after S1, ¼ into systole, ½ into systole, ¾ into systole, before S2, peak S2 and finally after S2.

In time-frequency domain, the first 11 features are extracted using the Shannon entropy of the 10th level wavelet detail coefficients (*WT entropy 1-11*). As depicted in **Figure 6.7**(c), 11 equal duration segments are formed to compute Shannon entropy. Energy is also computed from the high frequency components of the signal using 6th level details wavelet coefficients. It is obtained by summing of the squared samples of the 9 middle segments (leaving the first one and the last segment as shown in **Figure 6.7**(c)) and denoted as *WT detail 1-9*.

The other time-frequency domain features are obtained from the S-transform (Wang 1991). A total of 16 features, denoted as *ST map 1-16*, were achieved by limiting the frequency content of the S-transform to 150 Hz and down-sampling the result into a 4×4 matrix (see **Figure 6.8**(c-d)). Eight more features were extracted in form of eigenvalues resulted from the singular value decomposition of the S-transform (see **Figure 6.8**(b)), denoted as *eigenvalue 1-8*. The singular value decomposition S-transform is represented as $U \sum V^T$, where U and V are generally called left and right eigenvectors, or in this particular case eigentime and eigenfrequency, respectively. Next, 40 features are extracted from two left and two right eigenvectors and their corresponding histograms (*Eigenvalues 1-8*, *Eigenfrequency 1 1-10*, *Eigenfrequency 2 1-10*, *Eigentime1 1-10*, *Eigentime2*, *1-10*), see **Figure 6.8**(e-l).

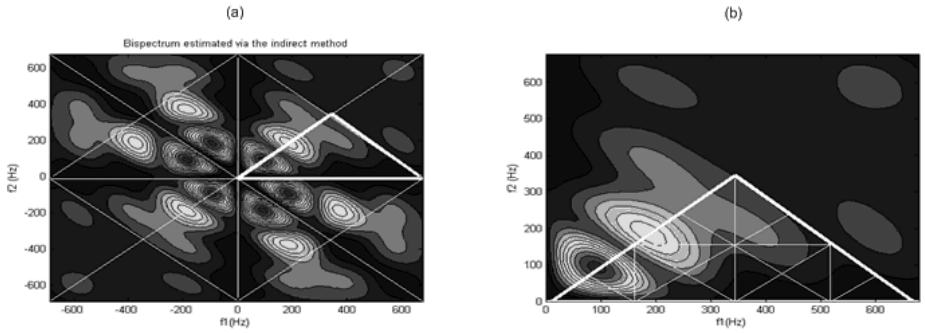


Figure 6.9 Example of bispectrum of the systolic ejection murmur shown **Figure 6.8**(a). In (a) bispectrum and bold triangle is non-redundant region, and in (b) the region of interest. The smaller triangles, 16, in (b) are the features *HOS 1-16*.

In the category of nonlinear and chaos based or state space based features, the applied feature set includes: 16 discretized equally sized triangles from non-redundant region of the bispectrum (*HOS 1-16*) (see **Figure 6.9**), 40 features from Gaussian mixture model of the reconstructed space of the heart cycle (*GMMx cycle 1-40*), 40 from murmur (*GMMx murmur 1-40*), 8 features obtained using the variance fractal dimension (*VFD 1-8*) and 10 features from recurrence quantification analysis (*RQA 1-10*) (see **Figure 6.10**).

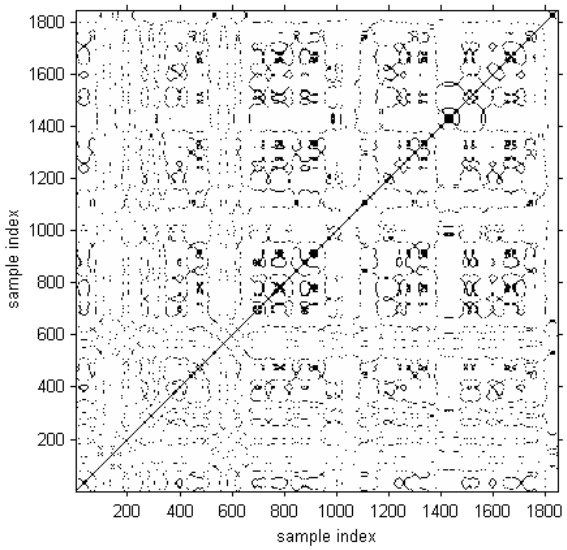


Figure 6.10 The recurrence plot of the heart sound from the systolic ejection case shown in **Figure 6.8**(a). From this recurrence plot 10 recurrence statistics are calculated.

6.3.6. Feature Set- IV

In order to find the best features among all the extracted features from the three aforementioned sets, as well as to determine the impacts of the introduced features (from set 1) a feature set 4 is created. This set of features is the superset of the previous three sets of features.

6.4. Feature Selection and Classifier

In order to reduce the classifier's complexity, feature selection might be applied to

select among the most discriminative features. In the present work, a subset of features is selected for the new feature set (feature set 1) reported in sections 6.3.1-6.3.3 using Paudil's sequential floating forward selection (SFFS) strategy. This method selects or rejects features based on the error estimate of a 1-nearest neighbor leave-one-out classifier. In our approach, the selection is performed with the toolbox developed by Ververidis (Ververidis 2008). It should be mentioned that a similar strategy is suggested by Ahlstrom *et al.* (Ahlstrom 2006) in order to reduce the original set of 207 features to a more manageable subset of 14 features. In the provided comparisons this feature set is identified by feature set 3. The features which comprise this set are the ones listed in the original work by Ahlstrom and colleagues. Regarding Olmez *et al. (Olmez and Dokur 2003)* algorithm, no feature selection has been performed because the complete set, composed by 32 features, is applied by these authors. Finally, the aforementioned feature selection strategy was applied to all the features mentioned in the previous subsection.

Feature set 1	**Feature set 2**	**Feature set 3**	**Feature set 4**
loudness 1 *zcr 1* *transition ratio1* *spectral power 2* *fundamental frequency 1* *spectral centroid 1* *spectral power 3* *spectral flux 1* *statistical skewness 1* *Lyapunovexponent 1*	*No selection* *(Kept as in original work)*	*WT detail 7 VFD 8* *Shannon energy 5* *Shannon energy 6* *Shannon energy 6* *GMM cycle 5* *Shannon energy 4* *GMM murmur 5* *Eigenfrequency1 2* *WT entropy 10* *GMM cycle 4* *Eigenfrequency1 1* *Shannon energy 8* *VFD 2* *VFD 8* *HOS 1* *(Kept as in original work)*	*loudness 1* *transition ratio1* *spectral power 2* *stat skewness 1* *Lyapunovexponent 1* *PoS 13* *PoS 16* *PoS 17* *PoS 18* *PoS 22* *Shannon energy 8* *WT details 5* *WT entropy 6* *ST map 3* *Eigenfrequencies1 4* *Eigenfrequencies2 10* *Eigentimes1 5* *HOS 6* HOS 13 *GMMx 7* *GMMx murmur 4* *VFD 3* *VFD 7*

Table 6.2 Selected features in each feature set.

Regarding the classification stage, for each of these feature sets a support vector

machine (SVM) classifier was trained and applied. A method based on the "one-against-one" (OAO) binary classification approach was chosen using a cubic polynomial kernel function (Hsu and Lin 2002).

6.5. Results and Discussions

Our four sets of features as described in the previous section were applied on the heart murmur database introduced in Section 5.7 of Chapter 5. This database is referred as *databaseMurmur*. This database is prepared with the heart sounds containing murmur from the patients admitted at the Cardiothoracic Surgery Center of the University Hospital of Coimbra. A total of 81 heart sound clips with murmurs, corresponding to 2047 beats, were collected from 51 patients. The main biometric characteristics of the population were: BMI = 25.41±2.16 Kg/m2; age = 64.65±8.64 years; number of males = 48; number of females = 3.

In the classification stage, sounds classified as murmur are categorized into the following 7 classes: Aortic Regurgitation (AR), Aortic Stenosis (AS), Mitral Regurgitation (MR), Pulmonary Regurgitation (PR), Pulmonary Stenosis (PS), Subaortic Stenosis + Ventricular Septal Defect (SAS+VSD), Systolic Ejection (SE).

The selected feature subsets for each feature set are listed in Table 6.2. The first set of features, feature set 1, is reduced to 10 features after applying SFFS. The features which comprise feature set 2 in Table 6.2 are the ones listed in the original work by Ahlstrom and colleagues. Since Olmez *et al.* the algorithm uses all 32 features, no feature selection has been performed. Finally, the SFFS feature selection strategy was applied to all the features mentioned in the previous section. The output of the SFFS on the overall 256 features is identified as feature set 4 composed by 23 features. Training was performed with 50% of the available data in the test database. The achieved results are shown in Table 6.3**Erro! A origem da referência não foi encontrada.**.

As can be observed, the best results have been achieved using feature set 4. This feature set, composed by a subset of 23 features (see table I) which comprises some of the proposed new features, exhibits a Se = 96.15% and a Sp = 96.16%. These results are very similar to the results achieved with the newly proposed feature set (set 1), which obtained Se = 95.74% and Sp = 95.01% with a significantly smaller set of 10 features, leading to a reduction in complexity. Regarding the results obtained with the feature set proposed by Ahlstrom and colleagues and Olmez and colleagues, it is observed that the former achieves similar results to the proposed strategy, although (i) a higher number (14) of features are required and (ii) a higher inter-class performance variability is observed. As for the algorithm proposed by

Olmez *et al.* the obtained results suggest that this is the less efficient method, given that it attains the lowest results, while using the highest number of features. These results seem to suggest that (i) the proposed set of features shows evidence of adequate discriminative ability to cap

Murmur Type	***Feature Set –I***	***Feature Set-II***	***Feature Set-III***	***Feature Set-IV***
Sensitivity (in %)				
AR	*94.02*	*91.02*	*92.10*	*93.02*
AS	*95.37*	*89.90*	*88.04*	*94.78*
MR	*90.85*	*86.43*	*93.21*	*95.20*
PR	*91.34*	*89.15*	*84.22*	*93.21*
PS	*100*	*100*	*94.33*	*100*
SAS+VSD	*96.17*	*95.39*	*86.71*	*90.41*
SE	*100*	*99.13*	*90.82*	*100*
Overall	*95.74*	*93.02*	*90.51*	*96.15*
Specificity (in %)				
AR	*88.25*	*96.68*	*85.33*	*91.29*
AS	*95.49*	*95.04*	*91.63*	*93.98*
MR	*93.90*	*88.86*	*88.14*	*96.33*
PR	*96.20*	*94.83*	*89.90*	*95.80*
PS	*100*	*99.15*	*97.43*	*100*
SAS+VSD	*95.07*	*95.22*	*95.32*	*95.27*
SE	*100*	*100*	*98.21*	*100*
Overall	*95.01*	*96.79*	*91.26*	*96.16*

Table 6.3 Murmur classification results with three set of selected features.

-ture the main characteristics of the most significant heart murmur classes and (ii) exhibits the overall best results in terms of trade-off between complexity and classification performance.

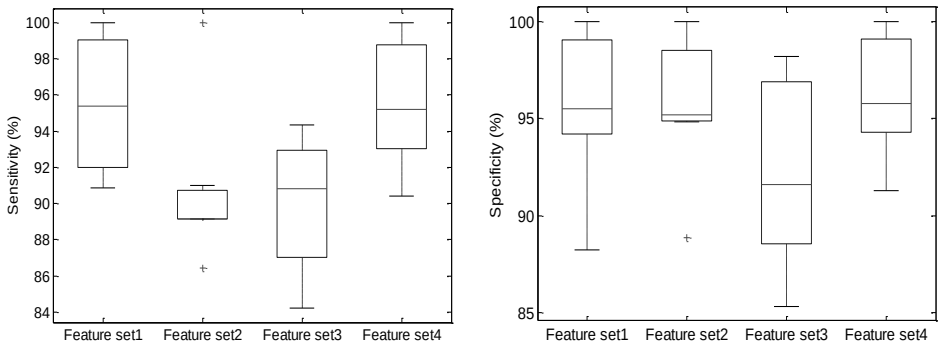


Figure 6.11 Left Box-Wishker plot is of Sensitivities and the right one is of Specificities achieved.

In order to test the impact of the different feature sets and examine the statistical difference between the results achieved using the four sets of features shown in **Figure 6.11** test is found out based on the Gaussian nature of the data. It was mentioned in section 1 of this chapter that if the distribution is Gaussian then the suitable test is the t-test, otherwise a non-parametric test such as rank based test can be applied. Sensitivities and specificities obtained from feature set I, II, III, and IV demonstrated, nearly, Gaussian distribution. Therefore, t-test was performed which to inspect the statistical significance of the observed sensitivities and specificities. It seems that the sensitivities obtained using Feature set I are different from the Feature set II, Feature set III and Feature set IV with the statistical significance of p=0.049, p= 0.017, p= 0.948 (where p is the probability of null hypothesis which assumes that the means of the two distributions are different), respectively. While specificities obtained using Feature set I with respect to the feature sets are, p=0.952, p=0.1942, and p=0.784. It can be deduced that there sensitivities in Feature I is significantly different from Feature II and Feature III while same with Feature IV. On the other hand, specificities' distribution from Feature I is same as rest of the feature sets (see in **Figure 6.11**).

6.6. Conclusions

In this chapter we tackled the problem of cardiac murmur classification using heart sounds. Seven types of cardiac murmur corresponding to different cardiac conditions were included in the database. These murmurs were of various grades and were classified by the physicians during auscultation. However, in our method we

focus only on classifying the murmur's type rather than their grades. Grade classification (or, grading murmur) is performed namely based on its audibility by the medical professional.

The proposed method for heart murmur classification required identification of the heart beats in a given heart sound recording. Within the heart beat the murmur region was also identified using the segmentation method described in Chapter 5. For classification of the heart murmur, the proposed method strategized with the three main processing blocks: feature extraction from each heart beat, selecting discriminative features based on feature selection approach, and then the SVM classifier from the automatic classification. A new feature set composed by 17 features to capture time, frequency and complexity signatures of murmurs is suggested. The most discriminant subset of 10 features has been identified using a sequential floating feature selection. This feature scheme with a SVM classifier was tested and compared with two well known sets of features, taken from the reference methods described by Ahlstrom *et al.* (Ahlstrom 2006) and by Olmez *et al.* (Olmez and Dokur 2003), coupled to SVM classifiers.

The achieved results seem to confirm that the proposed feature set has the ability to capture the discriminative morphological and spectral characteristics of heart murmurs. Furthermore, the obtained performances confirm that the results are better than the results obtained using the most cited algorithms for murmur detection from heart sounds. However, these results are reached with a smaller complexity, since only 10 features are applied.

Chapter 7

CONCLUSIONS AND PERSPECTIVES

Having reached the end of this dissertation, our first thought is that this is, at best, the end of an initial stage of heart sound analysis. Heart sound is a very informative biosignal, since it directly encodes the mechanical activity of the heart. It has been shown that heart sound has the potential to be applied in the important biomedical applications both in hospital as well as in home-care setting. Under heart sound analysis process, the three main modules are constructed: the first one contains noncardiac sound identification and elimination from the heart sounds; the second one includes, distinctively identification of normal heart sounds and abnormal heart sounds, and then segmentation by the appropriate algorithm; the third and last module is about the heart murmur classification.

In this chapter, we summarize the entire work carried out throughout this thesis, from Chapter 2 to Chapter 6, and some main conclusions are withdrawn from it. Since technical aspects of the methods were thoroughly discussed in the respective chapter therefore only main important technical points are highlighted here.

7.1. Summary and Conclusions

This document includes five main chapters, chapter 2 to 6, in which the two earlier chapters contain some background regarding of the human heart, origin of the heart sound, and some of the previous methods regarding the analysis of the heart sounds. While the later three chapters, chapter 3 to 6 contain the main contribution in solving the problems defined under heart sound analysis.

In chapter 2, some background information on the cardiac structure, the origin of heart sounds, their characteristics and their diagnostic and prognostic value were introduced. At the start of the chapter, the anatomy and physiology of heart sound was explained in which the structure of the heart and its function, i.e. pumping

blood to the body, were described. Then, the origin of and the importance of the characteristics of the main heart sounds' constituents, i.e. S1, S2, S3 and S4, were highlighted. Heart sounds produced by prosthetic heart valves were also described and analyzed. Abnormal heart sounds which are originated due to some cardiovascular anomalies, namely in valvular diseases, were explained. The devices used to capture heart sounds in order to perform computer based automatic analysis are sensor based stethoscopes. Hence, a brief introduction of the major events in the development of stethoscopes over the years has also been provided. It is followed by the small description of the main signal processing steps in the modern electronic stethoscope. Finally, some of the main diagnostic and prognostic applications of heart sounds in cardiology are reviewed. This chapter includes a small description of heart sounds and related information to understand the fundamentals for heart sound analysis.

Chapter 3 describes the most popular and up-to-date techniques applied thus far in the area of heart sound analysis. An overview of the previous approaches related to the three major problem modules: non-cardiac heart sound identification during heart sound recording; heart sound segmentation in which main heart sound components by identifying their onsets (start of the sound) and offsets (stop of the sounds) are identified; identification of abnormal heart sound from the normal heart sound, and if heart sound is classified as abnormal sounds then the determination of its type, were presented.

Chapter 4 includes the algorithm that enables the detection of heart sound contaminations by several sources, irrespective of their spectral an intensity charecteristics and interference model. The method is based on a template-matching approach composed by two main phases: in a first phase, the algorithm detects a noise-free reference heart sound with the duration of one heart cycle and in the second phase, the detected reference heart sound template is applied in real time to detect deviations in the heart sound pattern. The periodic nature of the heart's dynamic is explored to determine an uncontaminated reference sound. Periodicity is observed in noise-free heart sounds. However, this property tends to vanish or completely disappear as noise corrupts the heart sound. Hence, the first phase of the algorithm applies two periodicity characteristics of uncontaminated heart sound: heart sound is a quasi-stationary signal exhibiting periodicity in the temporal domain, and it is also observed that most of its energy is concentrated in its S1 and S2 components, which manifests itself as repetitions of similar energy patterns in adjacent frequency bands, i.e., leading to a periodicity patterns in frequency bands. These two observations are applied to identify an uncontaminated heart sound heart cycle. It should be mentioned that this processing step is the most demanding one from the computational load perspective. Once this template has been identified, the algorithm proceeds in detecting non-cardiac sounds by checking the template against the acquired

sound. Two computationally simple features are extracted to analyze the similarity with respect to the template. Simulation tests using Matlab suggest that this phase of processing can be performed in real time enabling the algorithm to be used to provide feedback information to the user regarding the quality of the signal. The algorithm was tested on a reasonably large database containing range of heart sounds collected from various subjects suffering from several heterogeneous heart conditions. High sensitivity and specificity were obtained when the algorithm was applied on our database.

In chapter 5, the second module of heart sound analysis, i.e., heart sound segmentation, was presented. It started with a short introduction of the main state of the art approaches. Since heart sound segmentation amounts to be a difficult problem considering the divergence of characteristics of heart sounds in different cardiac patients, our strategy for heart sound segmentation applied different approaches according to the heart sound characteristics. In the proposed approach, first the heart sound is classified according to its characteristic, and then each class of heart sound is segmented with a different algorithm. This classification was achieved using chaos signature in the heart sounds, which has been computationally measured by estimating Lyapunov exponents. Abnormal heart sounds exhibit more chaos than normal heart sounds. This hypothesis was assessed using three heart sound databases containing, respectively, normal heart sounds from native heart valves, normal heart sounds induced by the prosthetic heart valves, and abnormal heart sounds containing murmur. A simple threshold-based classification scheme was suggested. Optimal thresholds were computed based on the ROC curve approach.

After classification of the heart sound's type, a suitable segmentation algorithm was chosen to segment the heart sound between the two proposed algorithms. Segmentation was performed in two main phases in each algorithm: the first phase included delimitation of the sounds in which the onsets (starting of the sound) and the offsets (stop of the sound) of all audible and non-audible sounds in a heart sound recording were identified, and the second phase determines physiologically relevant sounds such as S1 and S2. The first proposed method is based on energy envelop extraction of the sounds in the heart sounds' wavelet domain. It uses the approximation coefficients for delimiting the sounds, followed by sound segments' classification as S1 or S2 using the high-frequency signature. The high frequency signature is based on the hypothesis that higher pressures across the aortic valve cusps induce higher vibrations.

The second proposed algorithm is based on wavelet decomposition-nonlinear features (WD-NLF). It uses simplicity and strength as the nonlinear features. The presented algorithm is highly adaptive and suitable depth of wavelet decomposition and thresholding, being these two factors independent of the person's specific de-

tails. In the second phase, recognition of the S1 and S2 sounds using the same approach as proposed in the energy envelop method, i.e., based on the high frequency signature, is applied.

Three types of databases were used to test these two algorithms; it was observed that the WD-NLF algorithm segments heart sounds from each database with high accuracy, while the energy envelop method results with high accuracy segmentation whenever in the presence of normal heart sound. The other proposed method WD-NLF computes complexity for the segmentation of heart sounds with murmurs as well as heart sounds collected without murmur. The energy envelop method is computationally less expensive than WD-NLF; given its good performance in heart sound segmentation without murmur, it is observed that it is the adequate algorithm in these situations.

Chapter 6 includes a detailed description of heart murmur classification. It started with a short explanation of the fundamental cause of heart murmur generation. A method for heart murmur classification was presented. The strategy followed was based on feature selection followed by a SVM classifier. A new feature set composed by 17 features to capture the time, frequency and complexity signatures of murmurs is suggested. The most discriminant subset of 10 features was identified using sequential floating feature selection. This feature scheme with a SVM classifier was tested and compared with two well known sets of features, taken from the reference methods described in Ahlstorm *et al.* (Ahlstrom 2006) and Olmez *et al.* (Olmez and Dokur 2003). The achieved results seemed to confirm that the proposed feature set had the ability to capture the discriminative morphological and spectral characteristics of heart murmurs. Furthermore, the obtained performances confirmed that the results are comparable to the results obtained using the most cited algorithms for murmur detection from heart sounds. However, these results were reached with a smaller complexity, since only 10 features were applied.

7.2. Perspectives for Future Research

At this stage, we outline some of the works that can be further carried out in order to provide a complete toolbox for the heart sound analysis. In this process, we first highlight the possible improvements in our proposed methods, and then list some of the tasks besides the ones tackled in this thesis.

By agreeing with Winston Churchil's quotation, "*Now this is not the end. It is not even the beginning of the end. But it is, perhaps, the end of the beginning*", it can be argued that the proposed methods for the three defined modules of heart sound analysis can be relooked to sort out shortcomings such as reducing the required long delay during the phase I of noise detection algorithm or modifying complexity measure to

determine better delimitation in cases high grade murmurs (Mitral and Aortic regurgitation). With respect to future work concerning each module presented, there is always room for improvements. This was highlighted in the respective chapters. Furthermore, there are still several tasks that can be considered such as, for instance, third (S3) and fourth (S4) heart sound identification, split measurement between the components of S1 and S2 sounds, respectively.

In order to prove the methods under heart sound analysis module to be efficient and robust, one must have a large and diverse database. Our test database is required to be enhanced by including more heart sounds in which murmurs are present. Then, a proper annotation of heart sounds including, auscultation site, murmur's grade, and body mass index (BMI) of the patients are required to carry out a thorough study of the outcomes when these methods are applied on the database. Because of the lack of number of patients exhibiting heart sounds with murmur of various grades (I-VI) in our database, we could not present a separate set of results of segmentation and classification on heart sounds with murmur of different degrees of severity (different grades). Therefore, database enhancement and its annotation will be the foremost task if we are to extend research work on the heart sound analysis.

In pursuit to achieve computationaly simpler algorithms for each of the modules presented here, it is important to reduce the number of algebraic operations in the algorithm. Hence, it is important to keep this at the top of the priority list of requirements when an algorithm is designed for each heart sound analysis module. However, at the same time accuracy of heart sound analysis in terms of sensitivity and specificity must also be given close attention. Therefore, in the methods a trade-off must be found between computational cost and accuracy of the obtained results. For instance, one research direction for heart sound analysis is by obtaining multichannel heart sounds; it is devised to enable cardiac auscultation from more than one site on the chest via different sensors (an array of sensors) at the same time (Wood et al. 1992). This may provide better results but suffers on the count of computation load. As a synchronized multichannel heart sound recording system enables to capture the heart sound component where it is best heard most heart sound components can be segmented with high accuracy using suitable processing techniques. Such technology can be used in several other diseases (e.g. multiple sclerosis among Chronic Obstructive Pulmonary Diseases (COPD)) besides CVD.

However, some of the tasks that were not on the scope of this thesis are also important in order to develop a complete computer aided diagnosis for cardiac diseases. That is going to be one of the future challenges. The three main tasks that are important to determine some of the cardiovascular diseases (CVDs) such as heart failure (HF) decompensation, coronary heart diseases, etc. (see Section 2.4)

are: identification of S3 and S4 sounds, systolic time intervals (STI), and the timings of the components of S1 and S2. This enables us to measure the split between its components, for instance finding the duration of the tricuspid (T1) and mitral valve (M1) movements, which essentially constitutes the S1 sound, will help to determine the split between the sounds produced by the closure of these two valves. If S3 and S4 sounds are prevalent in adults then they are the key sounds, particularly the S3 sound, in order to provide prognosis for HF (Patel et al. 1993; Shah et al. 2008). In fact, we have attempted to deal with this issue by proposing a wavelet decomposition-simplicity filter which exploits lower frequency range characteristics of S3 sounds and therefore less complex when compared to the other two main heart sounds (see paper numbered as P7 in Section 1.4). Due to the lack of heart sounds with S3 sounds the methods could not be tested in order to reach a conclusive mechanism of determining the S3 sounds. Therefore, the foremost task in this regard is to accumulate a significant number of heart sounds with S3 and S4 sounds and then apply the proposed methods on them so we can evaluate if the proposed algorithm is needed to be further investigated.

The second untouched task is the identification of STIs which are determined by estimating pre-ejection period (PEP) and the left ventricle ejection time (LVET). Recently, our research group has introduced a heart sound based STI measurement algorithm based on a Bayes approach (Paiva et al. 2012), which has shown to exhibit the highest accuracy, precision and correlation with respect to the gold standard among available methods for portable STI measurement. However, the detection of the onset of the aortic valve opening event exhibits a relatively low correlation (about 0.5) regarding the gold standard and an absolute estimation error of about 20%. This approach may still be at the starting point to carry out research in order to overcome the shortcomings in this approach.

The third important task is to estimate split interval components of S1 and S2 sounds. Split in S1 sounds is often not heard due to following T1 sounds right after the M1 sounds which usually lies in the range of 0.02-0.03 seconds. However in the presence of CVDs such as the right bundle branch block, ventricular premature contraction and ventricular tachycardia split becomes wider (Lehrer 2003). Similarly, the S2 sound split widens in patients with hypertension as it was mentioned in a previous section. S2 sound splits have been reasonably explored as some of the recent methodologies were presented in (Hedayioglu et al. 2011; Nigam and Roland 2008; Santos et al. 2011; Xu et al. 2000). In (Xu et al. 2000) it was demonstrated that it is possible to model each component of S2 by a narrow-band nonlinear chirp signal, while Nigam et al. (Nigam and Roland 2008) used blind-source separation for this decomposition by assuming the mutual statistical independence of A2 and P2 in signals gathered with 4-tipped endoscope. A more recent method explored a similar idea with very interesting results using sequential auscultations instead of

simultaneous ones (Hedayioglu et al. 2011). However, S1 split interval estimation is almost an unexplored problem, which can be approached with the same strategy as for S2 sound split estimation.

BIBLIOGRAPHY

Abarbanel, H. D. (1996). *Analysis of Observed Chaotic Data*, New York: Springer-Verlag New York Inc. .

Abbas, A. K. (2009). *Phonocardiography Signal Processing.*

Abbas, A. K., and Bassam, R. (2009). *Phonocardiography Signal Processing*, Morgan and Claypool.

Abdullah-Al-Mamun, K., Sarker, F., and Muhammad, G. (2009). "A High Resolution Pitch Detection Algorithm Based on AMDF and ACF." *J. Sci. Res.* , 1(3), 508 - 515.

Ahlstorm, C. (2008). "Nonlinear Phonocardiographic Singal Processing," Linkoping University, Linkoping.

Ahlstrom, C. (2006a). "Processing for the Phonocardiographic Signal - Methods for the Intelligent Stethoscope," Linköping University, Sweden.

Ahlstrom, C., Hult, P., Rask, P., Karlsson, J., Nylander, E., Dahlstro, U., Ask, P. (2006b). "Feature Extraction for Systolic Heart Murmur Classification." *Annals of Biomedical Engineering*, 34(11), 1666–1677.

Akay, M., Semmlow, J. L., Welkowitz, W., Bauer, M. D., and Kostis, J. B. (1990). "Detection of Coronary Occlusions Using Autoregressive Modeling of Diastolic Heart Sounds." *IEEE Trans. on BioMed. Engg.*, 37(4), 366 - 373.

Akay, Y. M., Akay, M., Welkowitz, W., Semmlow, J. L., and Kostis, J. B. (1993). "Noninvasive Acoustical Detection of Coronary Artery Disease: A Comparative Study of Signal Processing Methods." *IEEE Trans. Biomed. Engg.*, 40(6), 571 - 578.

Alfredson, R. J. (1997). "Wavelets and Heart Sounds." *Acoustics Australia*, 25(1), 5 - 10.

Alpert, M. A. (1990). "Clinical Methods : The History, Physical, and Laboratory Examinations." Clinical Methods, 131 - 137.

Altunkaya, S., Kara, S., Görmüş, N., and Herdem, S. "Statistically Evaluation of Mechanical Heart Valve Thrombosis Using Heart Sounds." *World Congress on Engineering*, 1 - 5.

Amado, R. G., and Filho, J. V. "Pitch Detection Algorithms Based On Zero-Cross Rate and Autocorrelation Function For Musical Notes." *Int. Conf. on Audio, Language and Image Processing (ICALIP)* 449 - 454

Atal, B. S., and Miner, L. R. (1976). "A Pattern Recognition Approach to Voiced-Unvoiced-Silence Classification with Applications to Speech Recognition."

IEEE Tran. on Acoustics, Speech, and Signal Processing, 24(3), 201 - 212.

Atoui, H., Fayn, J., Gueyffier, F., and Rubel, P. (2006). "Cardiovascular Risk Stratification in Decision Support Systems: A Probabilistic Approach. Application to pHealth." *Computers in Cardiology*, 33, 281 - 284.

Bai, Y. W., and Lu, C. L. "The embedded digital stethoscope uses the adaptive noise cancellation filter and the type I Chebyshev IIR bandpass filter to reduce the noise of the heart sound." *Int. Workshop on Enterprise Networking and Computing in Healthcare Industry, HEALTHCOM*, 278 - 281

Barry, D., Fitzgerald, D., Coyle, E., and Lawlor, B. "Single Channel Source Separation using Short-term Independent Component Analysis." *119th Audio Engineering Society Convention*, New York, USA.

Barschdorff, D., Ester, S., Dorsel, T., and Most, E. "Neural Network based Multi Sensor Heart Sound Analysis." *IEEE Conf. on Computers in Cardiology*, 303 - 306.

Barschdorff, D., Femmer, U., and Trowitzsch, E. "Automatic Phonocardiogram Signal Analysis in Infants based on Wavelet Transforms and Artificial Neural Networks." *Computers in Cardiology*, Lund, Sweden, 753 - 756.

Batsford, W. P., Mickleborough, L. L., and Elefteriades, J. A. (1995). "Ventricular Arrhythmias in Heart Failure." *J. of Cardiol. Clin.*, 13, 87-91.

Baykal, A., Ider, Y. Z., and Koymen, H. (1995). "Distribution of Aortic Mechanical Prosthetic Valve Closure Sound Model Parameters on the Surface of the Chest." *IEEE Trans. Biomed. Engg.*, 42(4), 358 - 370.

Bedi, R. "Acoustic Analysis of Carpentier-Edwards Bioprosthetic Heart Valves." *IEEE Conf. of the Enginering in Medicine and Biology - EMBC*, 1284 - 1285.

Belloni, F., Giustina, D. D., Riva, M., and Malcangi, M. "A New Digital Stethoscope with Environmental Noise Cancellation." *Advances In Mathematical And Computational Methods*, 169 - 174.

Berne, R. M., and Levy, M. N. (1997). *Cardiovascular Physiology: Mosby's Physiology Monograph Series*, Mosby, St. Louise, Missouri, .

Bettadapur, M. S., Griffin, B. P., and Asher, C. R. (2002). "Caring for Patients with Prosthetic Heart Valves." *Cleveland Clinic Journal of Medicine*, 69(1), 75-87.

Bhatikar, S. R., DeGroff, C., and Mahajan, R. L. (2005). "A Classifier based on the Artificial Neural Network Approach for Cardiologic Auscultation in Pediatrics." *Artificial Intelligence in Medicine*, 33, 251—260.

Blobel, B. (2011). "Ontology Driven Health Information Systems Architectures Enable Phealth for Empowered Patients." *International Journal of Medical Informatics*, 80, e17 - e25.

Bonow, O. B. (2006). "ACC/AHA 2006 Guidelines for the Management of Patients With Valvular Heart Disease." *Journal of the American College of Cardiology*, 48(3), e1-148.

Boutana, D., Djeddi, M., and Benidir, M. "Identification of Aortic Stenosis and Mitral Regurgitation By Heart Sound Segmentation on Time-Frequency Domain " *Int. Sym. on image and Signal Processing and Analysis*, 1 - 6.

Boutouyrie, P., Laurent, S., Girerd, X., Benetos, A., Colley, P. L., Abergel, E., and Safar, M. (1995). "Common Carotid Artery Stiffness and Patterns of Left Ventricular Hypertrophy in Hypertensive Patients." *Hypertension*, 25, 651 - 659.

Cai, R., Lu, L., Hanjalic, A., Zhang, H., and Cai, L. (2006). "A Flexible Framework for Key Audio Effects Detection and Auditory Context Inference." *IEEE Trans. on Acoustic, Speech, and Signal Processing*, 14(3), 1026 - 1039.

Cao, L. (1997). "Practical Method for Determining the Minimum Embedding Dimension of a Scalar Time Series." *Physica*, 110, pp. 43–50.

Carvalho, P., Gil, P., Henriques, J., Antunes, M., and Eugénio, L. "Low Complexity Algorithm for Heart Sound Segmentation using the Variance Fractal Dimension." *IEEE Int. Sym. on Intelligent Signal Processing (WISP'05)*,, Algarve, Portugal, 593–595.

Carvalho, P., Paiva, R. P., Couceiro, R., Henriques, J., Antunes, M., Quintal, I., and Muehlsteff, J. "Comparison of Systolic Time Interval Measurement Modalities for Portable Devices." *Int. Conf. of the IEEE - EMBS*, Buenos Aires, Argentina.

Carvalho, P., Paiva, R. P., Couceiro, R., Henriques, J., Quintal, I., Muehlsteff, J., Aubert, X. L., and Antunes, M. "Assessing Systolic Time-Intervals from Heart Sound: a Feasibility Study." *Int. Conf. of the IEEE - EMBS*, Minneapolis, USA, 3124-3128.

Chen, D., Durand, L. G., Lee, H. C., and Weiting, D. W. (1997). "Time - frequency Analysis of the First Heart Sound: Part 3: Application to Dogs with Varying Cardiac Contractility and to Patients with Mitral Mechanical Prosthetic Heart Valves." *Med. Biol. Eng. Comput.*, 35, 455 - 461.

Chizner, M. A. (2008). "Cardiac Auscultation: Rediscovering the Lost Art." *Current Problems in Cardiology*, 33(7), 326-408.

Cho, K. J., and Asada, H. H. "Wireless, Battery-less Stethoscope for Wearable Health Monitoring " *IEEE Annual Northeast on Bioengineering Conference*, Philadelphia, PA , USA 187 - 188

Choi, S., and Jiang, Z. (2010). "Cardiac Sound Murmurs Classification with Autoregressive Spectral Analysis and Multi-Support Vector Machine Technique." *Computers in Biology and Medicine*, 40, 8 - 20.

Cleland, J. G. F., Louis, A. A., Rigby, A. S., Janssens, U., and Balk, A. H. M. M. (2005). "Noninvasive Home Telemonitoring for Patients With Heart Failure at High Risk of Recurrent Admission and Death." *J. of the American College of Cardiology*, 45(10), 1954 - 1964.

Collins, S. P., Arand, P., Lindsell, C. J., Peacock, W. F., and Storrow, A. B. (2005). "Prevalence of the Third and Fourth Heart Sound in Asymptomatic Adults." *Congest Heart Fail*, 11(5), 242 - 247.

Collins, S. P., Lindsell, C. J., Peacock, W. F. H., V. D., Askew, J., Eckert, D. C., and Storrow, A. B. (2006). "The Combined Utility of an S3 Heart Sound and B-Type Natriuretic Peptide Levels in Emergency Department Patients With Dyspnea." *Journal of Cardiac Failure*, 12(4), 286 - 292.

Comon, P. (1992). "Independent component analysis, A new concept ?" *Signal Pro-*

cessing, 36, 287 - 314.

Cook, D. J., and Song, W. (2009). "Ambient intelligence and wearable computing: Sensors on the body, in the Home, and Beyond." *Journal of Ambient Intelligence and Smart Environments*, 1, 1 - 4.

Corona, B. T., and Torry, J. N. "Time-Frequency Represetnation of Systolic Murmurs using Wavelets." *Computer in Cardiology* Cleveland, OH, USA, 601 – 604.

Couceiro, R., Carvalho, P., Aubert, X., Muehlsteff, J., Henriques, J., and Antunes, M. "Beat-to-beat Cardiac Output Inference using Heart Sounds." *Int. Conf. on IEEE - EMBS* Boston, Massachusetts, USA, 5657 - 5661.

Couceiro, R., Henriques, J., Carvalho, P., Habetha, J., Antunes, M., and Harris, M. "Detection of Atrial Fibrillation using model-based ECG analysis." *ICPR*, Florida, USA, 1 - 5

Dargie, W., and Denko, M. K. (2010). "Analysis of Error-Agnostic Time- and Frequency-Domain Features Extracted from Measurements of 3-D Accelerometer Sensors." *IEEE System Journal*, 4(1), 26 - 33.

Dayem, M. K. A., and Raftery, E. B. (1967). "Phonocardiogram of the Ball-and-Cage Aortic Valve Prosthesis." *Brit. Heart J.*, 29, 446-452.

DeGroff, C. G., Bhatikar, S., Hertzberg, J., Shandas, R., Valdes-Cruz, L., and Mahajan, R. L. (2001). "Artificial Neural Network-Based Method of Screening Heart Murmurs in Children." *Circulation, Journal of the American Heart Association* 103, 2711 - 2716.

Deshmukh, O., Espy-Wilson, C. Y., Salomon, A., and Singh, J. (2005). "Use of Temporal Information: Detection of Periodicity, Aperiodicity, and Pitch in Speech." *IEEE Trans. Speech Audio Process*, 13, 776–86.

Desiardins, V. A., Enriquez-Sarano, M., Tajik, A. J., Bailey, K. R., and Seward, J. B. (1995). "Intensity of Murmurs Correlates With Severity of Valvular Regurgitation." *The American Journal of Medicine*, 100, 149 - 156.

Durand, L. G., Blanchard, M., Cloutier, G., Sabbah, H. N., and Stein, P. D. (1990). "Comparison of Pattern Recognition Methods for Computer- Assisted Classification of Spectra of Heart Sounds in Patients with a Porcine Bioprosthetic Valve Implanted in the Mitral Position." *IEEE Trans. Biomed. Engg.*, 37(12), 1121 - 1129.

El-Asir, B., and Mamlook, R. (2002). "ECG Beat Classification by a Fuzzy Logic " *Inform. Tech. J.* , 1, 213-217.

El-Segaier, M., Lilja, O., Lukkarinen, S., Sornmo, L., Sepponen, R., and Pesonen, E. (2005). "Computer-based Detection and Analysis of Heart Sound Murmur." *Annals of Biomedical Engineering*, 33(3), 937-942.

Ephraim, Y., and Malah, D. (1984). "Speech Enhancement Using a Minimum Mean Square Error Short-Time Spectral Amplitude Estimator." *IEEE Transactions on Acoustics, Speech and Signal Processing*, 32(6), 1109 - 1121.

Erickson, B. (2003). *Heart Sound and Heart Murmurs Across the Life Span Fourth Edition*, Mosby, Inc.

Erickson B. (2003). *Heart Sound and Heart Murmurs Across the Life Span Fourth Edition*,

Mosby, Inc.

Fan, W., Brooks, D. H., Mandel, J., Calalang, I., and Philip, J. H. "Detection of Hypovolemia Using Short-Time Fourier Transform Analysis." *21st Annual Northeast IEEE Bioengineering Conference* Bar Harbor, ME , USA 116-117.

Felner, J. M. (1990). "The Second Heart Sound." Clinical Methods: The History, Physical, and Laboratory Examinations, W. H. K., W. D. Hall, and J. W. Hurst, eds., Butterworths, Boston.

Fritzsche, D., Eitz, T., Minami, K., Reber, D., Laczkovics, A., Mehlhorn, U., Horstkotte, D., and Körfer, R. (2005). "Digital Frequency Analysis of Valve Sound Phenomena in Patients after Prosthetic Valve Surgery: Its Capability as a True Home Monitoring of Valve Function." *J Heart Valve Dis.*, 14(5), 657 - 663.

Furie, B., and Furie, B. C. (2008). "Mechanisms of Thrombus Formation." *The New England Jorunal of Medicine* 359, 938 - 949.

Gajic, B., and Paliwa, K. K. (2003). "Robust Speech Recognition Using Features Based on Zero Crossings with Peak Amplitudes." *IEEE Conf. Acoustic Speech and Signal Processing (ICASSP)*, 1, 64 - 67.

Geddes, L. A. (2005). "Birth of the Stethoscope." IEEE Engineering in Medicine and Biology, 84 - 86.

Geise, R. A. (2001). "Fluoroscopy: Recording of Fluoroscopic Images and Automatic Exposure Control." *Imaging & Theapeutic Technology*, 21(1), 227 - 236.

Giustina, D. D., Riva, M., Belloni, F., and Malcangi, M. (2011). "Embedding a Multichannel Environmental Noise Cancellation Algorithm into an Electronic Stethoscope." *Int. J. of Circuits, Systems and Signal Processing*, 5(2).

Gnitecki, J., and Moussavi, Z. "Variance Fractal Dimension Trajectory as a tool for Heart Sound Localization in Lung Sounds Recording." *Int. Conf. of the IEEE - EMBS*, Germany, 2420 – 2423.

Gnitecki, J., and Moussavi, Z. M. K. (2007). "Separating Heart Sounds from Lung Sounds." IEEE Engineering in Medicine in Biology, 20 - 29.

Gopal, M., and Karnath, B. (2009). "Clinical Diagnosis of Heart Failure." *Hospital Physician*, November/December, 9 - 15.

Gretzinger, D. T. K. (1996). "Analysis of Heart Sounds and Murmurs by Digital Signal Manipulation," University of Toronto.

Grey, J. M. (1975). "An Exploration of Musical Timbre," Stanford University.

Gupta, C. N., Palaniappan, R., Swaminathan, S., and Krishnan, S. M. (2007). "Neural Network Classification of Homomorphic Segmented Heart Sounds." *Applied Soft Computing*, 7, 4251-4254.

Gustfsson, F. (2000). *Adaptive Filtering and Change Detection*, Wiley, England.

Habetha, J. "The MyHeart Project – Fighting Cardiovascular Diseases by Prevention and Early Diagnosis." *IEEE-EMBS*, New York, USA.

Habetha, J. "The MyHeart Project – Fighting Cardiovascular Diseases by Prevention and Early Diagnosis." *Int. Conf. of the IEEE - EMBS*, 1 - 4.

Hadjileontiadis, L. J. (2005). "Wavelet based Enhancement of Lung and Bowel Sounds

using Fractal Dimension Tresholding-Part I: Methodology." *IEEE Trans. on Biomedical Engineering*, 52(6), pp. 1143-1148.

Hadjileontiadis, L. J., and Panas, S. M. (1997). "Adaptive Reduction of Heart Sounds from Lung Sounds Using Fourth-Order Statistics." *IEEE Trans. Biomed. Engg.*, 44(7), 642 - 648.

Hadjileontiadis, L. J., Rekanos, I. T. (2003). "Detection of Explosive Lung and Bowel Sounds by Means of Fractal Dimesion." *IEEE Signal Processing Letters*, 10(10), pp. 311–314.

Haghighi-Mood, A., and Torry, J. N. "A Sub-band Energy Tracking Algorithm for Heart Sound Segmentation." *Int. Conf. of the IEEE - EMBS*, Vienna, Austria, 501 – 504.

Haji, S. A., and Movahed, A. (2000). "Update on Digoxin Therapy in Congestive Heart Failure." *J. of American Family Physician*, 62(5).

Hall, L. T., Maple, J. L., Agzarian, J., and Abbott, D. (2000). "Sensor System for Heart Sound Biomonitor." *Microelectronics Journal*, 31, 583–59.

Hasfjord, F. (2004). "Heart Sound Analysis with Time Dependent Fractal Dimensions," Linköpings Universitet.

Hebden, J. E., and Torry, I. N. "Neural Network And Conventional Classifiers to Distinguish Between First And Second Heart Sound." *IEEE coll. of Artificial Intelligence Methods for Biomedical Data Processing*, 1-6.

Hedayioglu, F., Jafari, M., Mattos, S., Plumbley, M., and Coimbra, M. "Separating Sources from Sequentially Acquired Mixtures of Heart signals." *Int. Conf. on Acoustic Speech and Signal Processing (ICASSP)*, Prague, Czech Republic, 653 - 656.

Henriques, J., Carvalho, P., Harris, M., Antunes, M., Couceiro, R., Brito, M., and Schmidt, R. "Assessment of Arrhythmias for Heart Failure Management." *5th pHealth Conference 2008*, Valencia, Spain.

Hoekstra, F., Habers, E., Janssen, T. W. J., Verdaasdonk, R. M., and Meijer, J. H. (2010). "Relationship Between the Initial Systolic Time Interval and RR-interval during an Exercise Stimulus Measured with Impedance Cardiography." *J. of Physics*, 224, 1-4.

Hosenpud, J. D., and Greenberg, B. H. (2006). *Congestive Heart Failure*, Lippincott Williams & Wilkins.

Hossain, I., and Mousavi, Z. "An Overview of Heart-Noise Reduction of Lung Sound Using Wavelet Transform based Filter." *Int. Conf. on the IEEE - EMBS*, Cancun, Maxico, 458 - 461.

Hou, C., Chen, Y., Hu, L., Chuang, C., Chiu, Y., and Tsai, M. "Computer-Aided Auscultation Learning System for Nursing Technique Instruction." *Int. Cont. of the IEEE - EMBS*, Vancouver, British Columbia, Canada, 1575 - 1578.

Hsu, C. W., and Lin, C. J. (2002). "A Comparison on Methods for Multi-Class Support Vector Machines." *IEEE Trans. on Neural Networks*, 13(2), 415 - 425.

Huiying, L., and Iiro, H. "A Heart Sound Feature Extraction Algorithm Based on Wavelet Decomposition and Reconstruction." *Int. Conf. of the IEEE - EMBS*,

1539 - 1542.

Huiying, L., Sakari, L., and Iiro, H. "A Heart Sound Segmentation Algorithm Using Wavelet Decomposition And Reconstruction." *Engineering in Medicine and Biology*, Chicago, 1630-1633.

Hult, P., Fjiillbrant, T., Wranne, B., and Ask, P. (2004). "Detection of The Third Heart Sound using a Tailored Wavelet Approach." *Med. Biol. Eng. Comput.*, 42, 253 - 258.

Hult, P., Fjllbrant, T., Hildn, K., Dahlstrm, U., Warnne, B., and Ask, P. (2005). "Detection of The Third Heart Sound using a Tailored Wavelet Approach: Method Verification." *J. of Medical & Biological Engg. & Comp.*, 43(8), 212 - 217.

Hurtig-Wennlöf, A., Ahlstrom, C., Egerlid, R., Resare, M., Ask, P., and Rask, P. (2009). "Heart Sounds are Altered by Open Cardiac Surgery." *Exp Clin Cardiol.* , 14(2), 18 – 20.

Iaizzo, P. A. (2005). *Handbook of Cardiac Anatomy, Physiology, and Devices*, Humana Press Inc., Totowa, New Jersey.

Isaaz, K., Bruntz, J. F., Costa, A., Winninger, D., Cerisier, A., Chillou, C., Sadoul, N., Lamaud, M., Ethevenot, G., and Aliot, E. (2003). "Noninvasive Quantitation of Blood Flow Turbulence in Patients with Aortic Valve Disease Using Online Digital Computer Analysis of Doppler Velocity Data." *Journal of the American Society of Echocardiography*, 16(9), 965 - 974.

Jang, G., and Lee, T. (2003). "A Maximum Likelihood Approach to Single-channel Source Separation." *Journal of Machine Learning Research*, 4, 1365 - 1392.

Jennett, P., Scott, R., Hailey, D., Ohinmaa, A., Thomas, R., Anderson, C., Young, B., Lorenzetti, D., Hall, L. A., Milkovich, L., Claussen, C., Perverseff, T., Brownell, S., Jadavji, A., Sanguins, J., and Yeo, S. (2002). "Socio-Economic Impact of Telehealth: Evidence Now for Health Care in the Future." Health Telematics Unit, University of Calgary.

Jimenez-Gonzalez, A., and James, C. J. "Blind Source Separation to Extract Foetal Heart Sounds from Noisy Abdominal Phonograms: A Single Channel Method." *Int. Conf. on Advances in Medical, Signal and Information Processing (MEDSIP)*, 1 - 4.

Johnson, R., and Holzer, R. (2005). "Evaluation of Asymptomatic Heart Murmurs." *Current Paediatrics*, 15, 532 – 538.

Johnston, M., Collins, S. P., and Storrow, A. B. (2007). "The Third Heart Sound for Diagnosis of Acute Heart Failure." *Curr Heart Fail Rep.*, 4(3), 164 - 180.

Joshi, N. (1999). "Third Heart Sound." *Southern Medical Journal*, 92(8), 756 - 761.

Kam, L. T. (2003). "Analysis of Heart Sound II," National University of Singapore.

Kang, G. S., and Fransen, L. J. (1994). "Speech Analysis and Synthesis Based on Pitch-Synchronous Segmentation of the Speech Waveform." Naval Research Laboratory, Washington, DC.

Kanjilal, P. P., and Palit, S. "Extraction of Multiple Periodic Waveforms from Noisy Data." *IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, II361–4.

Karnath, B., and Thornton, W. (2002). "Auscultation of the Heart." *Hospital Physician*, 39 - 43.

Kennel, M. B., Brown, R., and Abarbanel, H. D. I. (1991). "Determining Embedding Dimension for Phase-space Reconstruction Using a Geometrical Construction." *Phy. Re. A*, 45(6), 3403-3411.

Khan, T. E. A., and Vijayakumar, P. (2010). "Separating Heart Sound from Lung Sound Using LabVIEW." *Int. J. of Computer and Electrical Engineering*, 2(3), 1793 - 8163.

Kim, C., and Stern, R. M. "Feature Extraction for Robust Speech Recognition Based on Maximizing the Sharpness of the Power Distribution and on Power Flooring." *Int. Conf. on Acoustic Speech and Signal Processing (ICASSP)*, 4574 - 4577.

Kim, S. H., Lee, H. J., Jae, M. H., and Chang, B. C. (1998). "Spectral Analysis of Heart Valve Sound for Detection of Prosthetic Heart Valve Diseases." *Yonsei Medical Journal* 39(4), 302 - 308.

Klapuri, A. P. (2003). "Multiple Fundamental Frequency Estimation Based on Harmonicity and Spectral Smoothness." *IEEE Trans. On Speech and Audio Processing*, 11(6), 804 - 816.

Koekemoer, H. L., and Scheffer, C. "Heart Sound and Electrocardiogram Recording Devices for Telemedicine Environments." *Int. Conf. of the IEEE - EMBS*, Vancouver, British Columbia, Canada, 4867 - 4870.

Koymen, H., Altay, B. K., and Ider, Y. Z. (1987). "A Study of Prosthetic Heart Valve Sounds." *IEEE Trans. Biomed. Engg.*, 34(11), 853 - 863.

Kristjansson, T., Attias, H., and Hershey, J. "Single Microphone Source Separation using High Resolution Signal Reconstruction." *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Montreal, Canada, 817 - 820

Kudriavtsev, V., Polyshchuk, V., and Roy, D. L. (2007). "Heart Energy Signature Spectrogram for Cardiovascular Diagnosis." *BioMedical Engineering OnLine*, 6(16), 1 - 22.

Kudriavtsev, V., Polyshchuk, V., and Sayniina, O. "Hemodynamic Pressure Instabilities and Their Relation to Heart Auscultation." *5thInt. Symp. On Computational Technologies for Fluid/Thermal/Chemical/Stressed Systems with Industrial Applications*, 1-10.

Kumar, D., Carvalho, P., Gil, P., Henriques, J., Antunes, M., and Eugénio, L. "A New Algorithm for Detection of S1 and S2 Heart Sounds." *Int. Conf. of Acoustic and Speech Signal Processing (ICASSP)*, Toulouse, France, 105 – 108.

Lam, M. Z. C., Lee, T. J., Boey, P. Y., Ng, W. F., Hey, H. W., Ho, K. Y., and Cheong, P. Y. (2005). "Factors Influencing Cardiac Auscultation Proficiency in Physician Trainees." *Singapore Med J* 46(1), 11 - 14.

Lau, Y., and Chan, C. (1985). "Speech Recognition Based on Zero Crossing Rate and Energy." *IEEE Trans. on Acoustic, Speech, and Signal Processing*, 33(1), 320 - 323.

Lehner, R. J., and Rangayyan, R. M. (1987). "A Three-Channel Microcomputer System for Segmentation and Characterization of the Phonocardiogram " *IEEE Trans.*

Biomed. Engg., 34, 485 – 489.

Lehrer, S. (2003). *Understanding Pediatric Heart Sounds*, Saunders, St. Louis, Missouri.

Li, X., Tao, J., Johnson, M. T., Soltis, J., Savage, A., Leong, K. M., and Newman, J. D. "Stress and Emotion Classification Using Jitter and Shimmer Features." *Int. Conf. on Acoustic Speech and Signal Processing (ICASSP)*, 1081 - 1084.

Liang, H., Lukkarinen, S., and Hartimo, I. "Heart Sound Segmentation Algorithm Based on Heart Sound Envelogram." *IEEE Computers in Cardiology*, Lund, Sweden, 105–108.

Liang, H., Lukkarinen, S., and Hartimo, I. "A Boundary Modification Method for Heart Sound Segmentation Algorithm." *Computers in Cardiology*, Cleveland, OH , USA, 593 - 595

Lim, K. O., Liew, Y. C., and Oh, C. H. (1980). "Analysis of Mitral and Aortic Valve Vibrations and their Role in the Production of the First and Second Heart Sounds." *Physics in Medicine and Biology*, 25, 727-733.

Liu, S., Chang, K., and Fu, T. (2009). "Heart rate extraction from photoplethysmogram on fuzzy logic discriminator." *Engg. Applications of Artificial Intelligence*, 23(6), 968 - 977

Mackenzie, G. W. (1964). *Acoustics*, Focal Press, London.

MacQueen, J. (1967). "Some Methods for Classification and Analysis of Multivariate Observations." *Proc. of the fifth Berkeley Symp. Statistic and Probability*, 1, 281-297.

Maglogiannis, I., Loukis, E., Zafiropoulos, E., and Stasis, A. (2009). "Support Vectors Machine-based Identification of Heart Valve Diseases Using Heart Sounds." *Computer Methods and Programs in Biomedicine*, 95, 47-61.

Malarvili, M. B., Kamarulafizam, I., Hussain, S., and Helmi, D. "Heart Sound Segmentation Algorithm Based on Instantaneous Energy of Electrocardiogram." *IEEE Conf. Computers in Cardiology*, 327 - 330.

Mallat, S. G., and Zhang, Z. (1993). "Matching Pursuits With Time-Frequency Dictionaries." *IEEE Trans. on Singal Processing*, 41(12), 3397 - 3415.

Malmivno, J., and Plonsey, R. (2009). *Principles and Applications of Bioelectric and Biomagnetic Fields*, Oxford University Press, USA.

Martínez-Alajarín, J., López-Candel, J., and Ruiz-Merino, R. (2007). "Classification and Diagnosis of Heart Sounds and Murmurs Using Artificial Neural Networks " *IWINAC 2007, Part I, LNCS*, 4527/2007, 303-312.

Mehta, N. J., and Khan, I. A. (2004). "Third Heart Sound: Genesis and Clinical Importance." *J. of Cardiology*, 97(2), 183-186

Mendis, S., Puska, P., and Norrving, B. (2011). "Global Atlas on Cardiovascular Disease Prevention and Control ", World Health Organization.

Michael, A., and Chizner, M. D. (2008). "Cardiac Auscultation: Rediscovering the Lost Art." *Current Problems in Cardiology*, 33(7), 326-408.

Mintz, G., Carlson, E., and Kolter, M. (1991). "Comparison of noninvasive techniques in evalution of the nontissue cardiac valve prothesis." *The American Journal of*

Cardiology, 15(6), 222-231.

Misal, A., and Sinha, G. R. (2012). "Denoising of PCG Signal by using Wavelet Transforms." *Advances in Computational Research* 4(1), 46 - 49.

Mladovsky, P., Allin, S., Masseria, C., Hernández-Quevedo, C., McDaid, D., and Mossialos, E. (2009). "Health in the European Union." World Health Organisation, Copenhagen, Denmark.

Morgenstern, J., Czerny, H., Schmidt, H., Schulz, J., and Wernicke, K. (1978). "Systolic Time Intervals of the Fetal Cardiac Cycle." *Perinat. Med.*, 6, 173-196.

Murgo, J. P. (1998). "Systolic Ejection Murmurs in the Era of Modern Cardiology, What Do We Really Know?" *J. Am. Coll. Cardiol.*, 32, 1596 - 1602.

Nair, K., Muraleedharan, C. V., and Bhuvaneshwar, G. S. (2003). "Developments in Mechanical Heart Valve Prosthesis." *Sadhana*, 28(3 & 4), 575 – 587.

Narain, V. S., Puri, A., Gilhotra, H. S., Sadiq, P. A., Mehrotra, S., Dwivedi, S. K., Saran, R. K., and Puri, V. K. (2005). "Third Heart Sound Revisited: A Correlation with NTerminal Pro Brain Natriuretic Peptide and Echocardiography to Detect Left Ventricular Dysfunction." *Indian Heart J.*, 57, 31-34.

Nazeran, H. (2007). "Wavelet-based Segmentation and Feature Extraction of Heart Sounds for Intelligent PDA-based Phonocardiography." *Methods Inf Med*, 46, 135 – 141.

Nigam, V., and Priemer, R. "Cardiac Sound Separation." *IEEE Conf. on Computers in Cardiology*, 497 - 500.

Nigam, V., and Priemer, R. (2005a). "Accessing Heart Dynamics to Estimate Durations of Heart Sounds." *J. of Physiological Measurement*, 26, 1005 - 1018.

Nigam, V., and Priemer, R. "Source Adaptivity for Cardiac Sound Separation." *Midwest Symposium on Circuits and Systems*, 1920 - 1923.

Nigam, V., and Roland, P. (2008). "Generalized Blind Delayed Source Separation Model for Online Non-invasive Twin-fetal Sound Separation: A Phantom Study." *Journal of Medical Systems*, 2, 123 - 135.

Nissen, S. L. (2003). "An Acoustic Analysis of Voiceless Obstruents Produced by Adults and Typically Developing Children," The Ohio State University.

Oh, J. K., and Tajik, J. (2003). "The Return of Cardiac Time Intervals: The Phoenix is Rising." *J. of American colleage of Cardiology*, 42, 1471-1474.

Olmez, T., and Dokur, Z. (2003). "Classification of Heart Sounds Using an Artifical Neural Network." *Pattern Recognition Letters*, 24(1-3), 617-629.

Paiva, R., Carvalho, P., Aubert, X., Muehlsteff, J., Henriques, J., and Antunes, M. "Assessing PEP and LVET from Heart Sounds: Algorithms and Evaluation." *IEEE Engineering in Medicine and Biology Society*, Minneapolis, MN, 3129 - 3133

Paiva, R. P., Carvalho, P., Couceiro, R., Henriques, J., Antunes, M., Quintal, I., and Muehlsteff, J. (2012). "Beat-to-beat Systolic Time-Interval Measurement from Heart Sounds and ECG." *Physiol. Meas.* , 33(2), 177 - 194.

Paiva, R. P., Carvalho, P., Henriques, J., Antunes, M., Muehlsteff, J., and X., A. "As-

sessing PEP and LVET from Heart Sounds: Algorithms and Evaluation." *Int. Conf. of the IEEE - EMBS*, Minneapolis, USA.

Panagiotakis, C., and Tziritas, G. (2005). "A Speech/Music Discriminator Based on RMS and Zero-Crossings." *IEEE Trans. on Multimedia*, 7(1), 1520 - 9210

Park, T. H. (2000). "Salient Feature Extraction of Musical Instrument Signal," Dartmouth College, USA.

Patel, R., Bushnell, D. L., and Sobotka, P. A. (1993). "Implications of An Audible Third Heart Sound in Evaluating Cardiac Function." *West J Med*, 158(6), 606 - 609.

Paul, A. S., Wan, E. A., and Nelson, A. T. "Noise Reduction For Heart Sounds Using a Modified Minimum-Mean Squared Error Estimator with ECG Gating." *Int. Conf. of the IEEE - EMBS*, New York City, USA, 3385- 3390.

Pedersen, M. S., Larsen, J., Kjems, U., and Parra, L. C. (2007). "A Survey of Convolutive Blind Source Separation Methods." *Springer Handbook on Speech Processing and Speech Communication*, 1 - 34.

Pibarot, P., Dumesnil, J. G., and Magne, J. (2009). "Prosthetic Valve Dysfunction." Contemporary Cardiology: Valvular Heart Disease, Humana Press, 447 - 473.

Pope, S. T., Holm, F., and Kouznetsov, A. "Feature Extraction and Database Design for Music Software." *Int. Computer Music Conference*, Miami, USA, 1 - 8.

Popov, B., Sierra, G., Durand, L. G., Xu, J., Pibarot, P., Agarwal, R., and Lanzo, V. "Automated Extraction of Aortic and Pulmonary Components of the Second Heart Sound for the Estimation of Pulmonary Artery Pressure " *IEEE Engineering in Medicine and Biology* Quebec City, Que., Canada 921-924

Porakis, J. G., and Manolakis, D. G. (2002). *Digital Singal Processing : Principles, Algorithms, and Applications*, Prentice-Hall, Inc., N. J., USA.

Potamitis, I., and Ozerov, A. "Single Channel Source Separation using Static and Dynamic Features in the Power Domain." *European Signal Processing Conference (EUSIPCO)*, Lausanne, Switzerland, 1 - 5.

Pourazad, M. T., Mousavi, Z. K., and Thomas, G. "Heart Sound Cancellation from Lung Sound Recordings using Adaptive Threshold and 2D Interpolation in Time-Frequency Domain " *Int. Conf. of the IEEE - EMBS*, Cancun, Mexico, 2586 - 2589.

Primentas, G. (2003). "The Musical Aspects of Sound (Pitch, Dynamics, Timbre) as the Core Element For Interaction in Sound Based Computer Games.," The London Institute, London.

Puvimanasinghe, J. P. A. (2004). "Prognosis after Aortic Valve Replacement with Mechanical Valves and Bioprostheses," Erasmus University Medical Center, Rotterdam.

Rabiner, L. R., Cheng, M. J., Rosenberg, A. E., and Mcgonegal, C. A. (1976). "A comparative Performance Study of Several Pitch Detection Algorithms." *IEEE Trans. on Acoustic, Speech, and Signal Processing*, 24(5), 399 - 417.

Rajan, S., Doraiswami, R., Stevenson, R., and Watrous, R. "Wavelet Based Bank of Correlators Approach for Phonocardiogram Signal Classification." *International Symposium on IEEE-SP* Pittsburgh, PA 77 - 80.

Reed, T. R., Reed, N. E., and Fritzson, P. (2004). "Heart Sound Analysis for Symptom Detection and Computer-Aided Diagnosis." *Simulation Modelling Practice and Theory*, 12, 129 - 146.

Reimer, W. S., Simoons, M. L., Boersma, E., and Gitt, A. K. (2006). "Cardiovascular Diseases in Europe." *European Society of Cardiology*.

Reynolds, O. (1883). "An Experimental Investigation of the Circumstances Which Determine Whether the Motion of Water Shall Be Direct or Sinuous and of The Law of Resistance In Parallel Channels." *Philos Trans R Soc Lond*, 174, 935 - 87.

Rezek, I. A., and Roberts, S. J. (1998). "Stochastic Complexity Measures for Physiological Signal Analysis." *IEEE Trans. Biomed. Engg.*, 9(45), 1186 - 1191.

Ricke, A. D., Povinelli, R. J., and Johnson, M. T. "Automatic Segmentation of Heart Sound Signals Using Hidden Markov Models." *IEEE Conf. Computers in Cardiology*, 953 - 956.

Rietzschel, E., Boeykens, E., Buyzere, M., Duprez, D., and Clement, D. (2001). "A Comparision between Systolic and Diastolic Pulse Contour Analysis in the Evaluation of Arterial Stiffness." *Hypertension*, 37, 15 - 22.

Rios-Gutiérrez, F., Alba-Flores, R., Ejaz, K., Nordehn, G., Andrisevic, N., and Burns, S. "Classification of Four Types of Common Murmurs using Wavelets and a Learning Vector Quantization Network." *IEEE World Congress on Computational Intelligence*, 2207–2213.

Rosenstein, M. T., Collins, J. J., and Luca, C. J. D. (1993). "A Practical Method for Calculating Largest Lyapunov Exponents from Small Data Sets." *J. of Physica D*, 65, 117-134.

Salvatore, S. (1997). "Study: Most New Doctors Can't use Stethoscope ", <http://edition.cnn.com/HEALTH/9709/02/nfm.heart.sounds/> (14 July 2012).

Satpathy, M., and Mishra, B. R. (2008). *Clinical Diagnosis of Congenital Heart Disease*, Jaypee Brothers Medical Publishers (P) Ltd., New Delhi.

Saunders, J. "Real-time Discrimination of Broadcast Speech/Music." *IEEE International Acoustics, Speech, and Signal Processing (ICASSP)*, Atlanta, GA , USA 993 - 996.

Sava, H. P., and McDonnell, J. T. E. (1996). "Spectral Composition of Heart Sounds Before and After Mechanical Heart Valve Imdantation Using Modified Forward-Backward Prony's Method." *IEEE Trans. Biomed. Engg.*, 43(7), 734 - 742.

Schmidt, S. E., Holst-Hansen, C., Graff, C., Toft, E., and Struijk, J. J. (2010). "Segmentation of Heart Sound Recordings by a Duration-Dependent Hidden Markov Model." *Physiological Measurement*, 31, 513 – 529.

Semmlow, J. L., Akay, N., and Welkowitz, W. (1990). "Noninvasive Detection of Coronary Artery Disease Using Parametric Spectral Analysis Methods." IEEE Engineering in Medicine and Biology Magzine 33 - 35.

Shah, S. J., Marcus, G. M., Gerber, I. L., McKeown, B. H., Vessey, J. C., Jordan, M. V., Huddleston, M., Foster, E., Chatterjee, K., and Michaels, A. D. (2008). "Physiology of the Third Heart Sound: Novel Insights from Tissue Doppler Imag-

ing." *Journal of the American Society of Echocardiography*, 21(4), 395 - 400.

Sharif, Z., Mohd, Z., Ahmd, A., and Sheikh, S. "Analysis and Classification of Heart Sounds and Murmurs Based on the Instantaneous Energy and Frequency Estimators." *TENCON*, 130-134.

Shaver, J. A., Salerni, R., and Reddy, P. S. (1985). "Normal and Abnormal Heart Sounds in Cardiac Diagnosis Part I: Systolic Sounds." *Current Problems in Cardiology*, 10(3), 1 - 68.

Silverman, M. E., and Wooley, C. F. (2008). "Samuel A. Levine and the History of Grading Systolic Murmurs." *The American Journal of Cardiology*, 102(2), 1107 - 1110.

Sokolow, M., Mcllroy, M. B., and Cheitlin, M. D. (1989). *Clinical Cardiology*, Pub. Prentice Hall International Inc.

Sörnmo, L., and Laguna, P. (2005). *Bioelectrical Signal Processing in Cardiac and Neurological Applications*, Elsevier, New York.

Stachura, M. E., and Khashanshina, E. V. (2004). "Potential Impact of Telehealth on Socio-Economic Stability and Sustainability in the Process of Globalization." *Int. J. of Economic Development*, 6(3), 6 - 33.

Stein, P. D., and Sabbah, H. N. (1976). "Turbulent Blood Flow in the Ascending Aorta of Humans with Normal and Diseased Aortic Valves." *Circulation Research*, 39, 58-65.

Stein, P. D., Sabbah, H. N., Lakier, J. B., Magilligan, D. J., and Goldstein, D. (1981). "Frequency of the First Heart Sound in the Assessment of Stiffening of Mitral Bioprosthetic Valves." *Circulation J. of the American Heart Association*, 63, 200 - 203.

Strang, G., and Ngugen, T. (1996). *Wavelets and Filter Banks*, Wellesley - Cambridge Press, Wellesley MA.

Strunic, S. L., Rios-Gutiérrez, F., Alba-Flores, R., Nordehn, G., and Bums, S. "Detection and Classification of Cardiac Murmurs using Segmentation Techniques and Artificial Neural Networks." *CIDM*, 397-404.

Sutton, G., Harris, A., and Leatham, A. (1968). "Second Heart Sound in Pulmonary Hypertension." *Brit. Heart J.*, 30, 743 - 756.

Tang, H., Li, T., and Qiu, T. (2010). "Noise and Disturbance Reduction for Heart Sounds in Cycle-Frequency Domain Based on Nonlinear Time Scaling." *IEEE Trans. on BioMed. Engg.*, 57(2), 325 - 333.

Taplidou, S. A., and Hadjileontiadis, L. J. "Nonlinear Analysis of Heart Murmurs using Wavelet-Based Higherorder Spectral Parameters." *IEEE Conf. of the IEEE - EMBS*, 4502 - 4505.

Tavel, M. (1967). *Clinical Phonocardiography and External Pulse Recording*, Year Book Medical Publishers.

Tavel, M. (1978). *Clinical Phonocardiography and External Pulse Recording*, Year Book Medical Publishers.

Thomson, D. L., and Chengalvarayan, R. "Use of Periodicity and Jitter as Speech

Recognition Features." *IEEE International Conference on Acoustics, Speech and Signal Processing*, Seattle, WA , USA 21 - 24.

Tinati, M. A., Bouzerdoum, A., and Mazumdar, J. "Modified Adaptive Line Enhancement Filter and its Application to Heart Sound Noise Cancellation." *Int. Sym. on Signal Processing and its Applications*, Brisbane, Australia, 815 - 818.

Titze, I. R. (1994). "Workshop on Acoustic Voice Analysis, Summary Statement." National Center for Voice and Speech, Denver, Colorado.

Tsalaile, T. (2008). "Digital Signal Processing Algorithms and Techniques for the Enhancement of Lung Sound Measurements," Loughborough University.

Tsalaile, T., Naqvi, S. M., Nazarpour, K., Sanei, S., and Chambers, J. A. "Blind Source Extraction of Heart Sound Signals from Lung Sound Recordings Exploiting Periodicity of The Heart Sound." *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, 461 - 464

Tzanetakis, G. (2002). "Manipulation, Analysis and Retrieval Systems for Audio Signals," Princeton University.

Tzanetakis, G., and Cook, P. (2002). "Musical Genre Classification of Audio Signals." *IEEE Trans. on Speech and Audio Processing*, 10(5), 293 - 302.

UN. (2001). "World Population Ageing: 1950-2050." Population Division, DESA, United Nations, New York.

Várady, P. "Wavelet-Based Adaptive Denoising of Phonocardiographic Records." *Int. Conf. of the IEEE - EMBS*, Istanbul, Turkey, 1 - 4.

Vasilakis, M., and Stylianou, Y. (2009). "Voice Pathology Detection Based on Short-Term Jitter Estimations in Running Speech." *Folia Phoniatr Logop*, 61, 153–170.

Ververidis, D., Kotropoulos, C. (2008). "Fast and Accurate Sequential Floating Forward Feature Selection with the Bayes Classifier Applied to Speech Emotion Recognition." *Signal Processing*, 88(12), 2956-2970.

Villalba, E., Arredondo, M. T., Ottaviano, M., Salvi, D., and Hoyo-Barbolla, E. (2007). "Heart Failure Monitoring System based on Wearable and Information Technologies." *Journal of Communications*, 2(2), 10 - 21.

Vongpatanasin, W., Hillis, L. D., and Lange, R. A. (2003). "Prosthetic Heart Valves." *Medical Progress*, 335(6), 407 - 416.

Wanak, J. (1972). "Phonocardiology: Integrated Study of Heart Sounds and Murmurs: 4" Edition." Year Book Medical Publishers Inc., Chicago.

Wang, F., Syeda-Mahmood, T., and Beymer, D. "Finding Disease Similarity by Combining ECG with Heart Auscultation Sound." *Computers in Cardiology* Durham, North Carolina, USA 261–264.

Wang, P., Kim, Y., Ling, L. H., and Soh, C. B. "First Heart Sound Detection for Phonocardiogram Segmentation." *Int. Conf. of the IEEE- EMBS*, Shanghai, China, 5519 - 5522.

Wang, Y. (1991). "The Tutorial: S Transform." National Taiwan University, Taipei, Taiwan, Taipei, 1 - 23.

Warbhe, A. D., Dharaskar, R. V., and Kalambhe, B. "A Single Channel Phonocardiograph Processing using EMD, SVD, and EFICA." *Int. Conf. on Emerging Trends in Engg and Tech.*, Nagpur, India.

WHO. (2011). "Global Atlas on Cardiovascular Disease Prevention and Control ", World Health Organization.

Williams, P. L., and Warwick, R. (1989). *Gray's Anatomy*, Churchill Livingstone, Edinburgh.

Wilton-Davies, C. C. (1972). "Computer-Assisted Monitoring of ECG's And Heart Sounds." *Med. & Biol. Engng.*, 10, 153 - 162.

Wolf, A., Swift, J. B., Swinney, H. L., and Vastano, J. A. (1985). "Determining Lyapunov Exponents from a Time Series." *Physica*, 16 D, 285 - 317.

Wood, J. C., Buda, A. J., and Barry, D. T. (1992). "Time-Frequency Transforms: A New Approach to First Heart Sound Frequency Dynamics " *IEEE Trans. Biomed. Engg.*, 39(7), 730 - 740

Xiu-min, Z., and Gui-tao, C. "A Novel De-noising Method for Heart Sound Signal Using Improved Thresholding Function in Wavelet Domain." *Int. Conf. on Future BioMedical Information Engineering*, 65 - 68.

Xu, J., Durand, L., and Pibarot, P. (2000). "Nonlinear Transient Chirp Signal Modeling of the Aortic and Pulmonary Components of the Second Heart Sound." *IEEE Trans. Biomed. Engg.*, 47(7), 1328 - 1335.

Xu, J., Durand, L. G., Pibarot, P.,. (2002). "A New, Simple, and Accurate Method for Non-invasive Estimation of Pulmonary Arterial Pressure." *Heart Journal*, 76-80.

Yadollahi, A., and Moussavi, Z. (2006). "A Robust Method for Heart Sounds Localization Using Lung Sounds Entropy." *IEEE Trans. Biomed. Engg.*, 53(3), 497 - 502.

Yadollahi, A., Moussavi, Z., Shamsollahi, M. B., and Ahmadinejad, Z. (2005). "Heart Sounds Localization using Lung Sounds Entropy " *Int. J. Sci. Res.*, 15, 1 - 4.

Yip, L., and Zhang, Y. T. "Reduction of Heart Sounds from Lung Sound Recordings by Automated Gain Control and Adaptive Filtering Techniques." *Int. Conf. of the IEEE - EMBS*, 2154 - 2156

Yoganathan, A. P. (2000). "Cardiac Valve Prostheses." The Biomedical Engineering Handbook: Second Edition, Boca Raton: CRC Press LLC.

Yoganathan, A. P., Gupta, R., Udwadia, F. E., Miller, J. W., and Corcoran, W. H. (1976). "Use of the fast Fourier Transform for Frequency Analysis of the First Heart Sound in Normal Man." *J. Medical and Biological Engg.*, 69 - 73.

Zenk, B. M., Bratton, R. L., Flipse, T. R., and Page, E. E. (2003). "Accuracy of Detecting Irregular Cardiac Rhythms via Telemedicine." *Journal of Telemedicine and Telecare*, 10, 55 - 58.

Zhang, T. "System and Method for Automatic Singer Identification." *IEEE Int. Conf. on Multimedia and Expo*, Baltimore, MD, 1 - 15.

Zhang, X., Durand, L. G., Senhadji, L., Lee, H. C., and Coatrieux, J. (1998). "Time-

Frequency Scaling Transformation of the Phonocardiogram Based of the Matching Pursuit Method." *IEEE Trans. Biomed. Engg.*, 45(8), 972 - 979.

Zhang, X. Y., and Zhang, Y. T. "A Novel Method for the Noninvasive and Continuous Monitoring of Arterial Blood Pressure on an Electronic Stethoscope." *3rd European Medical & Biological Engineering Conference.*, Prague, Czech Republic, 4602-4605.

Zhou, H., Sadka, A., and Jiang, R. M. "Feature Extraction for Speech and Music Discrimination." *Int. Workshop on Content-Based Multimedia Indexing*, London, UK, 170 - 173

Zuckerwar, A., Pretlow, R., Stoughton, J., and Baker, D. (1993). "Development of a Piezopolymer Pressure Sensor for a Portable Fetal Heart Rate Monitor." *IEEE Trans. Biomed. Engg.*, 40(9), 963 – 969.