U.PORTO
FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Heart Sound Segmentation: A Stationary Wavelet Transform Based Approach

Nuno Miguel Santos Marques

Dissertação de Mestrado apresentada à
Faculdade de Ciências da Universidade do Porto

2013

MSC
2.º CICLO
FCUP
2013

U.PORTO

Heart Sound Segmentation: A Stationary Wavelet Transform Based Approach

Nuno Miguel Santos Marques

FC

Heart Sound Segmentation: A Stationary Wavelet Transform Based Approach

Nuno Miguel Santos Marques
Mestrado em Ciência de Computadores
Departamento de Ciência de Computadores
2013

Orientador
Miguel Coimbra, Professor Auxiliar, Faculdade de Ciências da Universidade do Porto

Coorientador
Rute Almeida, Investigador Auxiliar, Faculdade de Ciências da Universidade do Porto

U.PORTO

FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Todas as correções determinadas pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, _____/_____/_____

Nuno Miguel Santos Marques

Heart Sound Segmentation: A Stationary Wavelet Transform Based Approach

Tese submetida à Faculdade de Ciências da
Universidade do Porto para obtenção do grau de Mestre
em Ciência de Computadores

Departamento de Ciência de Computadores
Faculdade de Ciências da Universidade do Porto
September 2013

To my mother Odete and to my father Jorge

Acknowledgments

I would like to thank my advisors Miguel Coimbra and Rute Almeida, not only for the help, but also for their guidance and support they gave me throughout my studies; Tiago and Pedro for their friendship and support, making my time in the Computer Science Department incredibly enjoyable; my previous advisors, Luís Torgo and Vítor Costa for their patience and support in my research work; Ana Paula Tomás for helping in so many different courses and Sandra Alves for being the first Professor to support my application to my first research position; my CK friends for the camaraderie; my best friend Nelson for making me laugh so hard I forget any personal struggles i might be going through; my girlfriend and companion for putting up with me and making me happy everyday; but most of all, I would like to thank and to dedicate this work to my mother and father, for all the love and support they have given me throughout my entire life. Thank you.

Abstract

Cardiac auscultation has lost some emphasis in the cardiology practice in recent years. This is mainly due to the widespread availability of more elaborate diagnostic methods and the lack of auscultation training programmes. Auscultation, however, if done properly, remains a valuable medical procedure that allows the clinicians to make a quick diagnosis, sometimes avoiding additional and more expensive exams. The next step in the evolution of cardiac auscultation is the creation of a computer assisted cardiac assessment system that allows the detection of heart disease. Although there is a large amount of work done already in this area, there is still the need for a more reliable and accurate method.

To classify the signal extracted through a digital stethoscope, one must first divide the signal into four segments of relevance, the first heart sound(S1), the systolic period, the second heart(S2) and the diastolic period. This process is called heart sound segmentation. We can divide this process into four stages: pre-processing, where we remove the remove the signal's noisy components; representation, where the signal is transformed in a way that accentuates S1/S2 while attenuating systole/diastole segments; segmentation, where we delimit the heart sounds; and classification, where we distinguish S1 from S2. The segmentation stage can be further divided into two phases: peak and boundary detection. This thesis is structured accordingly.

We start by presenting an exploratory analysis of both datasets in terms of their spectral content. Then, we introduce the two most used types of pre-processing: filtering, where one removes the signal components that are associated with noise, and downsampling, where one shortens the length of a signal while keeping its general morphology. In the subsequent stages, we did not use any type of filtering in this stage as we wanted to show that it is possible to design a heart sound segmentation method that did not we this type of preprocessing, while achieving good results. The downsampling operation was only applied to the lengthier dataset, as its original size made the tests in the posterior stages, too computationally heavy.

The first contribution of this thesis starts by introducing, comparing and ranking different types of representations, in terms of their capabilities to detect and classify heart sounds. We found that the best representation for detection was the Shannon energy envelope, while the best representation for classification was the continuous wavelet transform.

The main contribution of this work is a novel peak detection procedure that achieves better results than the winner solution of the Classifying Heart Sounds Pascal Challenge. This challenge featured two datasets. Every test performed in this work used both datasets to assess the methods robustness facing clean and noisy signals. The novel procedure uses the inflection points of stationary wavelet transform coefficients to perform an initial segmentation followed by a hierarchical clustering procedure that picks the relevant segments. We varied the wavelet, its order, the scale and the type of coefficients to achieve maximum performance. The best performing parameter combinations achieved a total error of 56732 and 706535, while the previous best performing approaches of the challenge achieved 72242 and 1243640, for both datasets.

We also introduce two novel boundary detection methods: the longest increasing / decreasing subsequence and the difference between variations. The first is based on the assumption that the subsequence, of a given segment, with the longest contiguous increase is the beginning of an heart sound and the longest contiguous decrease is the end of an heart sound. The second proposed method maximizes the difference between a segment's variation and its neighbour's. We also obtained good results, out-performing known approaches.

In the classification stage, based on the introduced representations, we built features that described each S1/S2 segment by looking exclusively to that segment's information (individual features), and by also looking to its adjacent systole and diastole segments (neighbourhood features). Finally, we used the concatenation of both types of features to achieve the maximum accuracy. We used these features to train a machine learning algorithm, in order to predict an unseen dataset. We achieved similar results as other modern classification approaches.

Keywords. Stationary Wavelet Transform, Heart Sound Segmentation, Heart Sound Classification

Resumo

A auscultação cardíaca perdeu ênfase na prática de cardiologia nos últimos anos. Isto é maioritariamente por causa da disponibilidade de métodos de diagnóstico mais elaborados e pela sua ausência em programas de treino em hospitais e faculdades. No entanto, se for propriamente feita, continua a ser um procedimento médico que permite os profissionais da saúde fazer diagnósticos rápidos, evitando assim testes adicionais mais caros. O próximo passo na evolução da auscultação cardíaca é a criação de um sistema de apoio à decisão clínica que permita a detecção de doenças cardíacas. Embora haja uma quantidade enorme de trabalho feito nesta área, ainda existe a necessidade do desenvolvimento de métodos mais precisos e fiáveis.

Para classificar o sinal extraído através de um estetoscópio digital, devemos primeiro dividir o sinal em quatro tipos de segmento: o primeiro som cardíaco (S1), a sístole, o segundo som cardíaco (S2) e a diástole. Este processo denomina-se segmentação de som cardíaco. Podemos dividir este processo em quatro fases: pré-processamento, na qual removemos as componentes ruidosas do sinal; representação, onde transformamos o sinal de forma a que acentue os segmentos com S1/S2, atenuando os segmentos com sístoles e diástoles; segmentação, onde delimitamos os sons cardíacos; e classificação onde distinguimos os segmentos S1 de S2. Podemos ainda dividir a fase de segmentação em duas partes: detecção de picos e detecção de fronteiras. Esta tese está estruturada da maneira conforme as fases previamente mencionadas.

Começamos por apresentar uma análise exploratória dos conjuntos usados ao longo deste trabalho. Depois, introduzimos os dois métodos mais usados de pré-processamento: filtragem, onde removemos as componentes do sinal que estão associadas a ruído, e decimação, onde encurtamos o comprimento dos sinais, mantendo a sua morfologia geral. Nas fases posteriores, não usamos qualquer tipo de filtragem, dado que mostramos que é possível obter bons resultados de segmentação não usando este tipo de pré-processamento. A operação de decimação só foi aplicada ao conjunto de dados com sinais mais longos, dado que o seu tamanho original tornava os testes realizados nas

fases posteriores a esta, demasiado pesados computacionalmente.

A primeira contribuição desta tese começa por introduzir, comparar e ordenar tipos diferentes de representação, em termos da sua capacidade para detectar e classificar sons cardíacos. Concluímos dos nossos testes que a melhor representação para detecção e classificação é o envelope de energia de Shannon e a transformada wavelet contínua, respectivamente.

A contribuição principal deste trabalho é um procedimento de detecção de picos que obtém melhores resultados que a abordagem vencedora do concurso "Classifying Heart Sounds Pascal Challenge". Todas comparaçoes e testes feitos neste trabalho usam os dois conjuntos de dados apresentados neste concurso, para inferir a robustez dos métodos face a sinais limpos e ruidosos. O novo procedimento usa os pontos de inflexão da transformada wavelet estacionária seguido por um algoritmo clustering hierárquico que escolhe os segmentos relevantes(que contêm S1 ou S2). Variamos as wavelets, as ordens, as escalas e o tipo de coeficientes para atingir a máxima performance. A melhor combinação de parâmetros em obteve erros totais de 56732 e 706535, enquanto os melhores erros totais atingidos previamente no concurso foram de 72242 e 1243640, para os dois conjuntos de dados.

Também apresentamos dois novos métodos de detecção de fronteiras: a sub-sequência crescente/decrescente mais longa e a diferença entre variações. A primeira é baseada na suposição que a sub-sequência, de um dado segmento, de maior crescimento contíguo marca o início de um segmento cardíaco e que o maior decrescimento marca o seu fim. O segundo método apresentado procura comprimentos de segmento que maximize a diferença entre a sua variabilidade e a dos segmentos vizinhos. Obtivemos também bons resultados, ultrapassando outros métodos modernos.

Na fase de classificação, baseámo-nos nas representação introduzidas anteriormente, e contruímos três tipos de descriptores: descriptores que representavam a informaçao exclusivamente de um segmento (descriptores individuais), descriptores que representavam informaçao de um dado segmento e dos segmentos adjacentes, e concatenação dos dois tipos de descriptores de forma a atingir melhor precisão. Usámos estes descriptores para treinar algoritmos de aprendizagem máquina para preverem a classificação de novos segmentos. Obtivemos resultados semelhantes a outro métodos de classificação modernos.

Palavras-Chave. Transformada Wavelet Estacionária, Segmentação de Som Cardíaco, Classificação de Som Cardíaco

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Heart auscultation is a medical procedure with almost 200 years. In 1628, it was William Harvey who first concluded that the main function of the heart was to pump blood through the arteries to the body and that the pulse that created the flow, could be heard by applying their ear directly to the chest[HS02]. It was only in 1816 that Laennec, stated that *as uncomfortable for the doctor as it was for the patient, disgust in itself making it impracticable in hospitals, It was hardly suitable where most women were concerned and, with some the very size of their breasts was a physical obstacle to the employment of this method.*[HS02]. Laennec faced with a large sized woman with some symptoms of a diseased heart, rolled a piece of paper into a cylinder shape and had the impression that he could hear the heart sounds in a *"manner much more clear and distinct than I had ever been able to do by the immediate application of the ear"*[HS02] and thus invented the mediate auscultation through an instrument called the stethoscope (Greek: stethos=chest, skopein=to view or to see). Over the years, various works showed that some less frequent sounds like some types of murmurs were correlated with heart disease.

We now live in the digital era, and despite having much more sophisticated and reliable methods like the ultrasonic imaging and Doppler techniques, cardiac auscultation is still taught and used in modern cardiology, as it remains a valuable diagnostic tool[Tav06]. The common stethoscope, however, lacks some useful features like recording, and playing back sounds. It also cannot visual display or transmit the heart sounds to multiple clinicians simultaneously. These limitations have been resolved by the use of Electronic Digital Stethoscopes such as the Digiscope Prototype[Coi10] and many others, which have proved to be of great use due to its non-invasiveness and to its low cost, whether it be for analysis or for teaching young cardiologists[Tav06].

The next step in the evolution of cardiac auscultation is to create a computer assisted cardiac evaluation system that allows the detection of heart disease. Although there is a huge amount of work done already in this area, there is still the need for a more reliable method. To analyse the signal extracted through a digital stethoscope, one must first divide the signal into four segments of relevance, the first heart sound(S1), the systole period, the second heart sound(S2) and the diastole period. In some heart disease scenarios, there are some extra sounds like the S3 and S4. The aim of this work will be to divide an normal heart sound signal into four different types of segment. To give a clear explanation about these four types of segment, it is provided a brief description of the normal heart and how it produces the two main sounds: the S1 and the S2.

The heart pumps blood through the blood vessels to every part of the human body renewing its oxygen content. It has four chambers: the right and left atriums and the right and left ventricles. De-oxygenated blood from the superior and inferior vena cavae enters the heart through the right atrium which is pumped through the tricuspid valve into the right ventricle and then to the lungs where carbon dioxide is exchanged for oxygen. The left atrium receives the oxygenated blood from the lungs through the left and right pulmonary veins. The blood is then pumped into the left ventricle through the mitral valve and is sent out to the body by the aorta[SAH94].

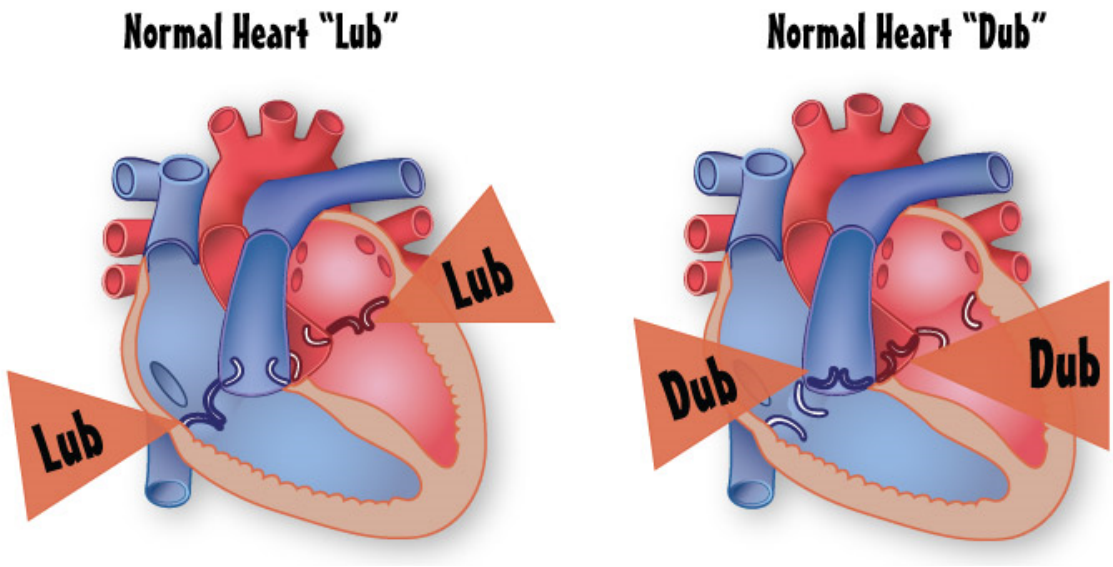


Figure 1.1: The "Lub" or S1 is caused by the closure of the mitral and tricuspid valves and marks the beginning of the systolic period. The "Dub" or S2 is caused by the closure of the aortic and pulmonary valves and it marks the beginning of the diastolic period. Adapted from [1]

As one can see in Fig.1.1 , the "Lub" or S1 is caused by the closure of the mitral and tricuspid valves and marks the beginning of the systolic period, i.e. the time in the cardiac cycle when blood is ejected from the ventricles into the great vessels. As the valves close within 100ms from each other it is heard as a single sound. The "Dub" or S2 is caused by the closure of the aortic and pulmonary valves and it marks the beginning of the diastolic period, i.e. the time when the left and right ventricules are being filled with blood. Both valves are normally heard as a single sound due to their almost simultaneous closure.

The division of the heart sound signal into S1, systole, S2 and diastole is called Heart Sound Segmentation. The most well known approach towards Heart Sound Segmentation(HSS) was presented by H.Liang's 1997 paper [LLH97]. This approach set the standard approach for heart sound segmentation dividing the method in four parts, pre-processing, representation, segmentation and classification of heart sounds. In the pre-processing stage the signal goes through filtering and downsampling operations in order to remove some artifacts through their abnormally high frequencies and by smoothing the signal. In the representation stage, the signal is transformed in order to maximize the difference between S1 and S2 from the systolic and diastolic periods. In the segmentation phase, the peaks corresponding to S1 and S2 are found and then procedure is done to detect the boundaries of the two heart sounds. In the classification, one tries to distinguish the S1 from the S2.

1.1 Evolution of HSS approaches

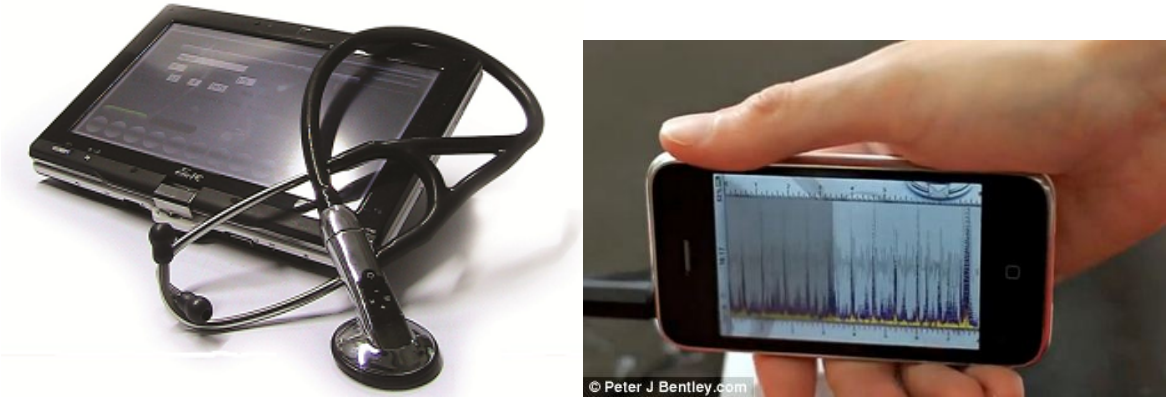
In order to give the reader an overview of the evolution of the Heart Sound Segmentation approaches, it is presented in chronological order, 3 known approaches towards HSS, starting with Liang's[LLH97]. After reducing some of the noise inherent to a heart sound signal, Liang gave his most valuable contribution in the representation stage. He proposed that a transformation to the signal called the Shannon Energy Envelope which attenuated the low-intensity parts of the signal more than the high-intensity parts and emphasized the medium intensity parts. This transformation due to its ability to accentuate the S1 and S2 heart sounds became the main reason that made this work, the most cited article in HSS. The segmentation and classification stages on the other hand, relied on some obscure thresholds leaving a path for further improvement.

[1] http://www.texasheartinstitute.org/images/ph_listen_normal-heart.jpg

In 2006, Kumar's paper [KCA+06], the segmentation and classification stages were much clearer. The authors used a simple rule to segment the signal just by using the zero-crossings of the normalized version of envelope of a transformation of the signal. The classification is done by applying a non-supervised method that bases his decision making on the fact that the duration of the diastole is longer then the systoles. This last criterion however, is not always true. When the patients are children or have an elevated heart rate the durations of the systolic and diastolic periods are sometimes indistinguishable and can lead to a high error.

In 2013, Moukadem [MDHB13] proposed a novel method that detected the boundaries of an heart sound by using the time frequency concentration information. In the classification stage, it distinguished the S1 from the S2 heart sounds using not only the durations of the systolic and diastolic period but also time frequency information features. Although this type of approach uses time frequency information for both segmentation and classification purposes, it relies on the same threshold based method of Liang for peak detection making it to sensitive to noisier auscultations.

1.2 Datasets



(a) Digiscope Prototype. Adapted from [2]

(b) An iPhone with the Istethoscope app. Adapted from [3]

Figure 1.2: Signal Extraction Tools

To compare and validate the known and novel methods, we used two public heart sound datasets that were featured in an competition called the Classifying Heart

[2]http://digiscope.up.pt/images/articles_imgs/new_layout.png
[3]http://i.dailymail.co.uk/i/pix/2010/08/31/article-0-0AFB9BE9000005DC-123_468x286.jpg

Sounds Pascal Challenge[BNCM11], the Digiscope and Istethoscope datasets. This competition was divided in two parts. The first was to test the contestant's algorithm segmentation capabilities. The second was to test the algorithm accuracy to classify an heart sound signal into one of four different labels, Normal,Murmur,Extra Heart Sound and Artifact. For this work, it was only used the scores from the first part of the challenge in both datasets, so that it could be tested the accuracy of the novel method. Two datasets were used in order to assess the algorithm's robustness facing a clearer and a noisier heart sound signal. The clearer heart sound dataset was the Digiscope dataset.

1.2.1 Digiscope

The Digiscope Project was a Portuguese funded project that aimed to develop a digitally enhanced stethoscope that extracted clinical features from the phonocardiogram (PCG) and provided clinical auxiliary diagnostic tool regarding specific heart pathologies. The data used in this study was collected in the Real Hospital Português (Brazil), with the approval of the RHP Ethics Commitee, anonymized and shipped to Portugal. It consists in 200 auscultations were made from children using a Littmann Model 3100 electronic stethoscope with a sampling frequency of 4KHz. All of the auscultations did not have any abnormal heart sounds. An expert pointed out the correct positions of 120 auscultations, 90 of which had reference annotations available to the contestants —training data—and 30 were used for the algorithm validation—test data.

1.2.2 Istethoscope

The noisier one was the Istethoscope dataset. The iStethoscope is an iPhone app that turns the iPhone into a digital stethoscope. It uses its microphone to record ones heart sound. A person can send to an expert physician or to the University of Minnesota Duluth to be collected and analysed for scientific purposes. This dataset is overall much more noisy and has higher variability than the Digiscope dataset. The variability may be connected to patient diversity. Since we do not have information on who was auscultated, the heartbeats may present very different patterns depending on whether is a child, a regular adult or a professional athlete. One source of noise is the fact that the auscultation are not done by clinicians. Since the average person is not trained to perform heart auscultations, the heart sounds will have high variability. The

final noise factor is the lack of ambient control. The auscultation can be performed wherever and the ambient noise if its mildly present, it will be heard through the iPhone microphone. The dataset was extracted using iPhone versions 3G and 3GS with a sampling frequency of 44100Hz. Once again, we did not have information about the auscultated subjects. An expert pointed out the correct positions of 30 auscultations. 10 of these auscultations were used for the algorithm validation phase and the rest for training. Both validation datasets had the annotations hidden from the contestants in an Microsoft Excel file.

1.3 Aims and Contributions

The main contributions of this thesis are:

- Comparison of representations in terms of their capability to distinguish S1/S2 from systole/diastole and to distinguish S1 from S2.
- Propose a novel peak detection procedure
- Propose two novel boundary detection procedures
- Comparison between three types of classification features: Individual, Neighbourhood, and concatenations of the previous two.

1.4 Thesis Structure

The rest of the content of this thesis is organized as follows: In Chapter 2, it is presented some exploratory analysis of the two datasets and the differences between different methods of pre-processing; in Chapter 3, it is shown a study about the use and capabilities of different types of Representation; in Chapter 4, it is presented the main contribution of this work, the Segmentation phase. It is introduced and its results is compared with the other approaches used in the Classifying Heart Sounds Pascal Challenge; in Chapter 5 it is presented a study that compares a supervised method using different types of features vs. an unsupervised method that uses the only the systole/diastole duration criteria to identify both S1 and S2 heart sounds; Finally, in Chapter 6, an overview of the content is shown, along with some discussion.

Chapter 2

Pre-Processing

Before proceeding with the representation, segmentation and identification stages, most approaches use pre-processing techniques like filtering and downsampling. In this thesis, we did not use any type of filtering and we only downsampled the Istethoscope dataset, for complexity purposes. In this chapter, to perform an exploratory analysis of the data, we present a spectral analysis of both Digiscope and Istethoscope datasets. Although we did not use any type of filtering, we illustrate the effects of filtering methods used in other HSS approaches. We also introduce downsampling and argue about the Istethoscope's high sampling frequency causing heavy time and space complexity for the algorithms presented in posterior chapters.

To introduce spectral analysis, we present first a short introduction about Fourier transform and how it can be used to obtain a signal's spectral content.

2.1 Fourier Transform

$$X(e^{j\omega}) = \sum_{-\infty}^{\infty} x[n]e^{-j\omega n} \tag{2.1}$$

The Fourier transform(eq.2.1) of a signal $x[n]$, can be interpreted as a sequence of weighted combinations of the complex exponential sequence $e^{j\omega n}$, where ω is the real normalized frequency variable [Mit10]. In this thesis, by Fourier Transform(FT) we mean the Discrete-Time Fourier Transform(DTFT), as all digital signals have a finite discrete time domain. The magnitude of the FT can be used to compute a signal's spectral content. In the Noise removal context, this is of great use since there are spectral differences among the Systole, Diastole, S1 and S2 segments.

2.2 Spectral Analysis

Spectral Analysis aims to estimate the distribution of the power over frequency of a signal. In our particular problem, the spectral content is a mean to distinguish S1 from S2 heart sounds, as it was noted in [AT84]. We use the Periodogram (eq.2.2) as the estimator of the spectral content of the signal, i.e. the Power Spectral Density.

$$\hat{\phi}_p(\omega) = \frac{1}{N}\left|\sum_{n=1}^{N} x[n]e^{-j\omega n}\right|^2 \tag{2.2}$$

This method is a non-parametric one as it does not fit the signal to a well defined model such as the parametric approaches , e.g. the Auto-Regressive(AR). All of the different spectral estimation methods assume that the signal is stationary, i.e. the statistical properties do not change in time. The fact that the parametric methods make a stricter assumption about this property, led us to use the non-parametric approach. There are more well developed methods that aim to improve the accuracy of this estimator such as the Blackman-Tukey or the Welch methods[SM05]. However, Stoica in [SM05], shows that those methods are more or less equivalent in their properties and performance for large signal lengths.

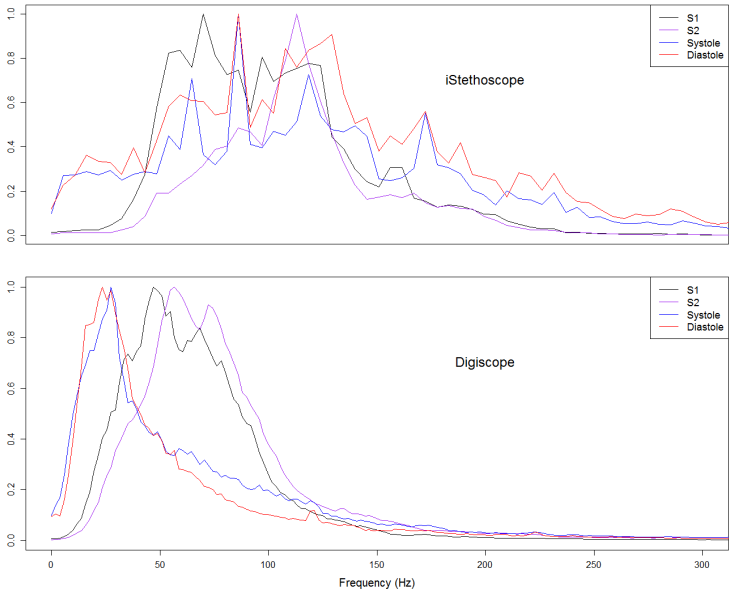


Figure 2.1: Median of periodogram of S1,S2,Systole and Diastole segments. Top Figure:Istethoscope. Bottom Figure:Digiscope.

To build this figure, we started by decomposing the Digiscope and Istethoscope into

sets of S1, systolic, S2 and diastolic segments. Then, we computed the periodogram for each segment of each type. Finally, to obtain an average spectral behaviour of all the segments of one type, we computed the median of each frequency —represented in Fig.2.1 by the x-axis —for all segments. All the median periodograms were normalized by its maximum value as it is hard to interpret the meaning of the amplitudes in the frequency domain. The periodograms had the same length, hence we represented them in the same plot as one can see in Fig.2.1. Analysing Fig.2.1, we can see that the Istethoscope segments lacks well defined spectral patterns, unlike Digiscope's. Focusing on the bottom figure(Digiscope), we see that the S2 has a spectral content around 75Hz which is slightly higher than S1(50Hz), following the claims of P.J. Arnott in [AT84]. The heart sound segments —S1,S2 —have higher frequencies than the Systolic and Diastolic segments. This difference suggests that spectral content is especially useful in a procedure where one differentiates heart sounds from the other types of segments, a procedure also known as peak detection. The results that confirm this argument are presented in Chapter 4. If we focus on the Istethoscope (Upper part of Fig.2.1) spectral content, we see that the majority of the spectral content is widely distributed around 150Hz. We can also see that the S2 spectral information completely overlaps the S1's. The difference between the spectral contents between datasets suggests that it is related to the conditions in which the dataset was extracted. Digiscope had a more controlled environment and the auscultations were performed by clinicians. The iStethoscope, on the other hand, did not share any of the Digiscope good conditions and the auscultations were made by non-experts.

2.3 Filtering

A digital filter is applied for attenuating the frequencies that not belong to the S1 or S2 heart sound frequency range. The most common filters are the convolution ones. Convolution is a mathematical operation that takes 2 signals and outputs in each points the area overlap between the input signals. In this section we analyse the effects of filtering used in other heart sound segmentation approaches. In [CVMC13], Castro applies a band pass filter with cut-off frequencies 25 and 900Hz; In [MDHB13], it is applied an high pass filter with a cut frequency of 30Hz ; in [CHJH13] the author uses a band pass filter with cut-off frequencies 50 and 200Hz. In [EMD12] the author applies a low pass filter with a cut off frequency of 800Hz.

As we can see in Fig.2.2, all filters seem to have a more or less similar effect on a heart sound signal, except the Butterworth bandpass filter with cut off frequencies 50 and

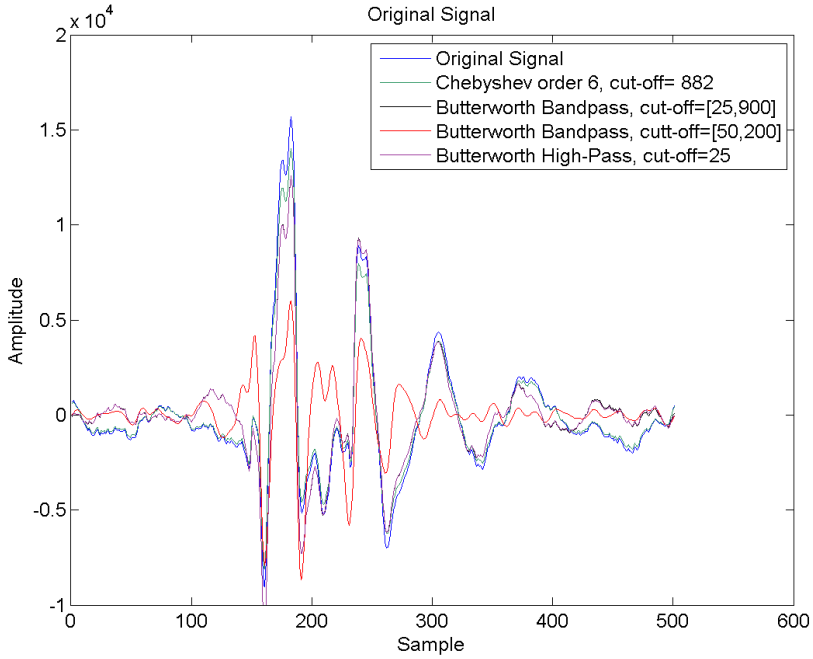


Figure 2.2: Effects of different convolutional filtering methods applied to a S1 heart sound.

200Hz, represented by the red line. Since in [CHJH13], the authors cut frequencies that are present in both S1 and S2 as it was seen on the previous section, some frequency information is lost and consequently the signal is deformed. This type of hard filtering, although it cuts frequencies that are known to be associated with heart sounds, achieves quite good results as it is shown in [CHJH13]. It is, however, a dubious method as it alters the original heart sound frequencies.

In the sub-sequent stages, we chose not to use any filtering, as we wanted to show that it is possible to achieve good heart sound segmentation results without this type of pre-processing. The results of our approach can be seen on Chapter 4, where we compare our method's performance against the winning approach —where the authors did use filtering —of the Classifying Heart Sounds Challenge.

2.4 Downsampling

Downsampling or decimation is the mathematical operation where one reduces the number of samples from a signal. in order for the signal to represent the same information with less samples, we keep every Mth sample of $x[n]$ and we remove

the in between $M-1$ samples. A downsampling of factor M is described by eq.2.3 where $x[n]$ is original signal and $y[n]$ is the downsampled signal.

$$y[n] = x[nM] \tag{2.3}$$

The Istethoscope dataset had a sampling frequency of 44100Hz, i.e. a 10 second Istethoscope signal has 441000 samples. Using such a high sampling frequency on algorithms with high time and space complexities is unfeasible. Examples of computationally heavy used approaches in this thesis are the Stationary Wavelet Transform —that keeps $2NK$ coefficients in memory, where N is the length of the signal and K is the used scale—and the Unweighted Pair Group Method with Arithmetic Mean(UPGMA)—that has a time-complexity of $O(n^2)$. To be able to perform exhaustive tests in the subsequent stages, we downsampled the Istethoscope's signals by a factor of 20 (to 2205Hz). This operation allowed us to capture frequencies up to 1102.5Hz, thus not deforming the signal.

2.5 Discussion

The need for an alternative representation of the signals especially in the iStethoscope dataset is pressing, given a distinction between S1,S2 and the diastole/systole is currently unfeasible.

The analysis of the differences of the spectral content on the Digiscope and iStethoscope, gave us an idea of the challenges of segmenting a clean and noisy heart sound. While in a clean dataset, represented by Digiscope, we can differentiate spectral content of signal and noise and consequently detect S1 and S2 heart sounds more successfully, in a noisy dataset, represented by Istethoscope, it asks for a more suitable representation for detecting and distinguishing S1 from S2.

Filtering approaches are widely used in heart sound segmentation methods. In this thesis, however, we present a heart sound segmentation method that achieves good results without using any type of filtering. In the remaining chapters, we downsampled the Istethoscope dataset to 2205Hz. This is due to the high length of the Istethoscope original signals, leading to higher computational costs in subsequent stages. We left the Digiscope dataset untouched, in terms of filtering and downsampling, since it has a computationally acceptable sampling frequency of 4000Hz.

In the next chapter, we present different representation methods and compare their

capabilities to distinguish S1 and S2 from the systole and diastole, and to distinguish S1 from S2.

Chapter 3

Representation

To correctly segment the signal and identify the heart sounds, a suitable representation of the signal is required. The representation should accentuate the difference between the phenomena we wish to detect—S1 and S2 heart sounds—and the noise—systolic and diastolic periods. The classification improvement is obtained by building features that contain different types of information about S1 and S2, allowing a better distinction from one another. In this chapter we introduce by topics, some representations that were used in other HSS approaches, highlighting its advantages and disadvantages. After the introduction we present a study that compares the representations by their ability to distinguish the different types of segments, i.e. S1, systole, S2 and diastole.

3.1 Envelopes

The envelope of a signal bounds the peaks of the signal, obtaining a low pass filtered version. This is illustrated in Fig.3.1. The envelope originally appeared from the need to decode low frequency content signals encoded in the amplitude high-frequency radio waves, a process known as Amplitude Modulation(AM)[Got66]. In AM, one transforms the original signal, which has low frequency content(audible frequencies), into an high frequency version of the original signal .

$$m(t) = M\cos(\omega_m t + \phi_m) \tag{3.1}$$

$$c(t) = A\sin(\omega_c t + \phi_c) \tag{3.2}$$

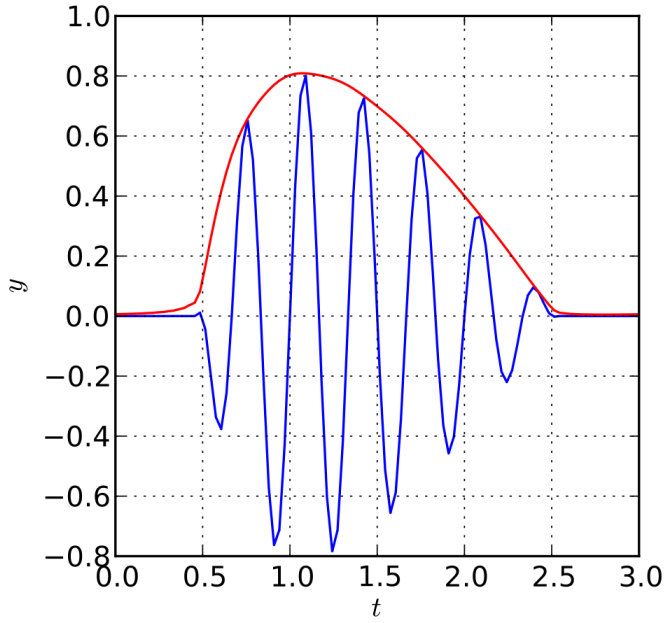


Figure 3.1: A signal's envelope. Adapted from [1]

$$y(t) = [1 + m(t)]\, c(t) = A\,[1 + M\cos(\omega_m t + \phi)\sin(\omega_c t))] \tag{3.3}$$

This transformation is called Amplitude Modulation, which consists in multiplying the modulating signal $m(t)$ with amplitude M, frequency $\frac{\omega_m}{2\pi}$ and initial phase ϕ_m(3.1), i.e. original signal, by the higher frequency signal, i.e. the carrier signal $c(t)$ (3.2). The product of the two is called the modulated signal $y(t)$(3.3). To demodulate the signal, one can use an envelope which computes the low frequency content (the original content) from the modulated signal.

In heart sound segmentation, the most known envelope is the Shannon Energy envelope. H.Liang in [LLH97], compared 4 different types of envelopes, the Shannon Energy, the Shannon Entropy, the absolute value and the Signal's energy, as we can see in Fig.3.2. Liang states that the main advantage of the Shannon Energy is that it emphasizes the medium intensities while attenuating more the low intensities. Such properties are desirable in a heart sound signal since the S1s and S2s usually are of mid-high intensity and the systole and diastole periods are low intensity, while noise might be present in high intensity peaks.

In Fig.3.3, we can see the effect of each type of envelope presented in [LLH97] has

[1] http://upload.wikimedia.org/wikipedia/commons/thumb/d/d7/Analytic.svg/2000px-Analytic.svg.png

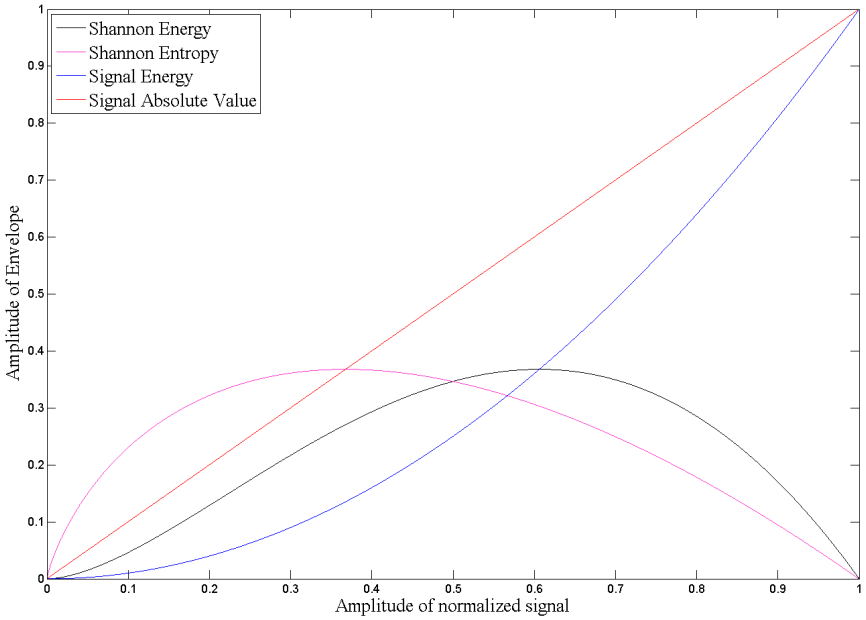


Figure 3.2: Comparison between different types of Envelopes

on a Digiscope and an Istethoscope signal. The Shannon Energy, as Liang stated, does accentuate the mid intensities while attenuating the low ones. However, one notices that while Shannon Energy does a better job by lowering the low intensities, the Shannon Entropy does a better job by setting the mid and high intensities with the approximately the same intensity. This will be particularly useful in the Segmentation stage to differentiate the S1 and S2 segments from the systole and diastole segments. The absolute value and signal energy are not considered to be good envelopes in the heart sound segmentation scenario given that while they do transform the signal, they keep the same relative intensities, consequently enhancing noise peaks.

3.2 Short Time Fourier Transform

In order to introduce the S-transform it is useful to first introduce the Short Time Fourier Transform(STFT) since one can easily derive the S-transform from the STFT.

$$STFT\left\{h(t)(\tau, f)\right\} = X(\tau, f) = \int_{-\infty}^{\infty} h(t)g(t-\tau)e^{-jft} \tag{3.4}$$

The STFT is simply an multiple application of the Fourier Transform to a signal in windows of the same size and shape. It produces a STFT matrix with 2 dimensions:

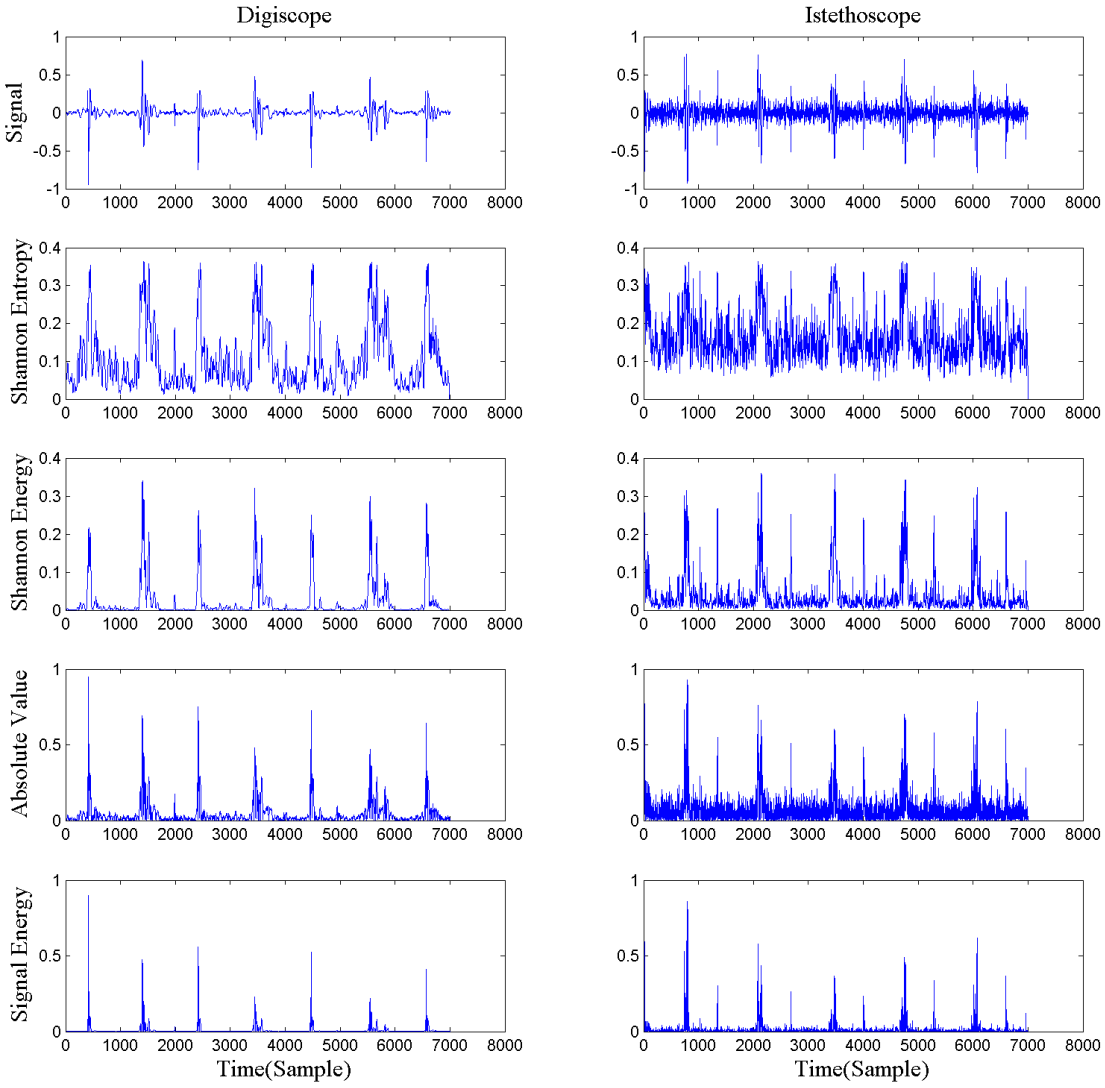


Figure 3.3: Effect of different types of Envelopes on a Digiscope and an Istethoscope Signal

the translation τ and the frequency f. The window function g can have several shapes. Known window shapes are: the Triangular(Bartlett), the Rectangular and the Hamming window. The different types of windows serve as a M size coefficient vectors that quantize the segments coefficients. The STFT has the particular application of enabling one to see how the frequencies present in a signal evolve over time. As all time-frequency representations however,it suffers from time or translation—τ—and frequency —f—resolution trade-off. Wider windows lead to a higher frequency resolution since they are able to capture both high and low frequencies in one window. They also lead to a lower time resolution because we do not know where exactly in the window were those frequencies. The narrower windows lead to a higher time resolution but a lower frequency resolution.

3.3 S-Transform

To derive the S-transform from the STFT [Sto91], we must first force the window function g to be a normalized Gaussian.

$$g(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}} \tag{3.5}$$

If we constraint the value of σ to be proportional to the inverse of the frequency we obtain the S-transform.

$$\sigma(f) = \frac{1}{|f|} \tag{3.6}$$

$$S(\tau, f) = \int_{-\infty}^{\infty} h(t) \frac{|f|}{\sqrt{2\pi}} e^{-\frac{(t-\tau)f^2}{2}} e^{-i2\pi f t} dt \tag{3.7}$$

This transform solves the trade-off by attributing windows with higher frequencies, lower widths and vice-versa.

In [MDHB13], the author uses the S-transform for two different purposes. First, he applies the S-Transform to obtain the time-frequency information in the 0-100Hz range. He then proceeds by computing the Shannon Energy Envelope of the resulting S-matrix. This step will give higher values to the time samples that contain frequency intensities between 40 and 80 Hz—the approximate frequency range of an S1 and S2.

$$\sigma(f) = \frac{\alpha}{|f|^p} \tag{3.8}$$

After obtaining the final representation, he applies the same threshold based steps to obtain the peaks that represent the S1 and S2 heart sounds. The author also modified

the S-Transform in the following way: instead of using the frequency function (3.6), he uses (3.8), where α and p are optimized parameters. These two parameters correctly tuned will give an optimized time and frequency resolution which consequently will lead to a more reliable time frequency representation, although making it very computationally expensive.

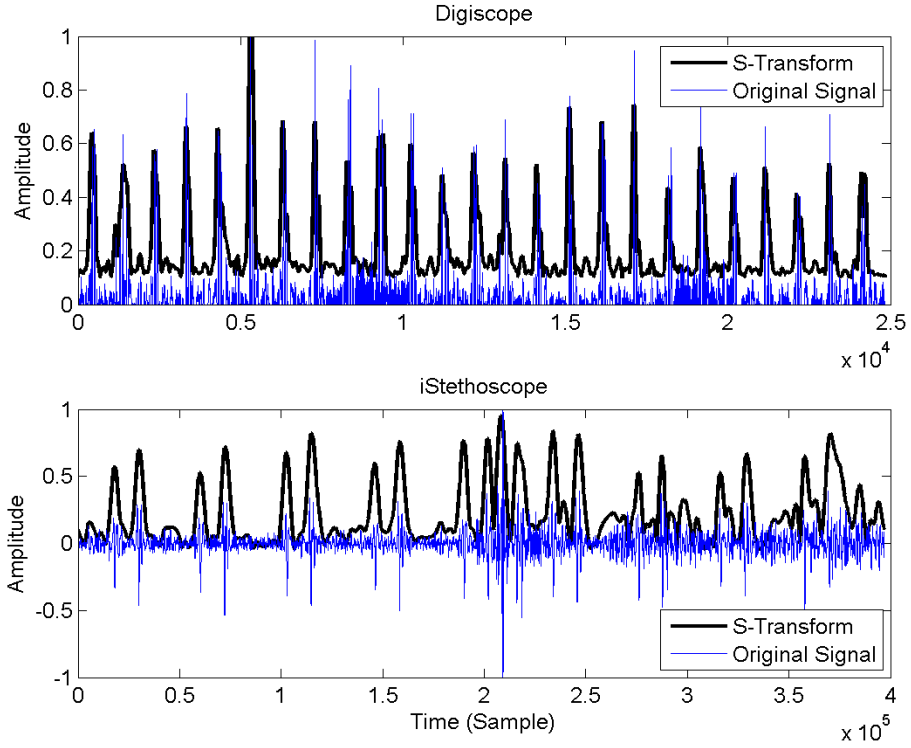


Figure 3.4: Effect of the S-Transform on a Digiscope and an Istethoscope Signal

Fig.3.4 illustrates the S-transform representation superimposed with the original signal. We can see that the S-Transform is indeed suitable for representing heart sound signals. Picking the right frequency range, by using spectral analysis, or the literature, we obtain in this case a good representation for both noisier and cleaner signal, even though the iStethoscope signal needs further processing.

3.4 Wavelets

Wavelet Transforms aim to create a matrix of coefficients that will provide information about a signals correlation with dilated,contracted and shifted versions of a mother wavelet function[BGG97]. Unlike the Short Time Fourier Transform where

the coefficients represent the correlation to complex sinusoids, the Wavelet Transform coefficients represent correlation to a small basis function called the mother wavelet. This mother wavelet is copied into scaled, shifted versions so that in the end, we end up with a multi resolution decomposition of a signal. In Fig.3.5 we can see different mother wavelet functions. The Wavelet Transform also has the property of assigning wider windows to higher scaled versions of the mother wavelet and narrower windows to lower scaled versions of the mother wavelet, the non-optimized STFT.

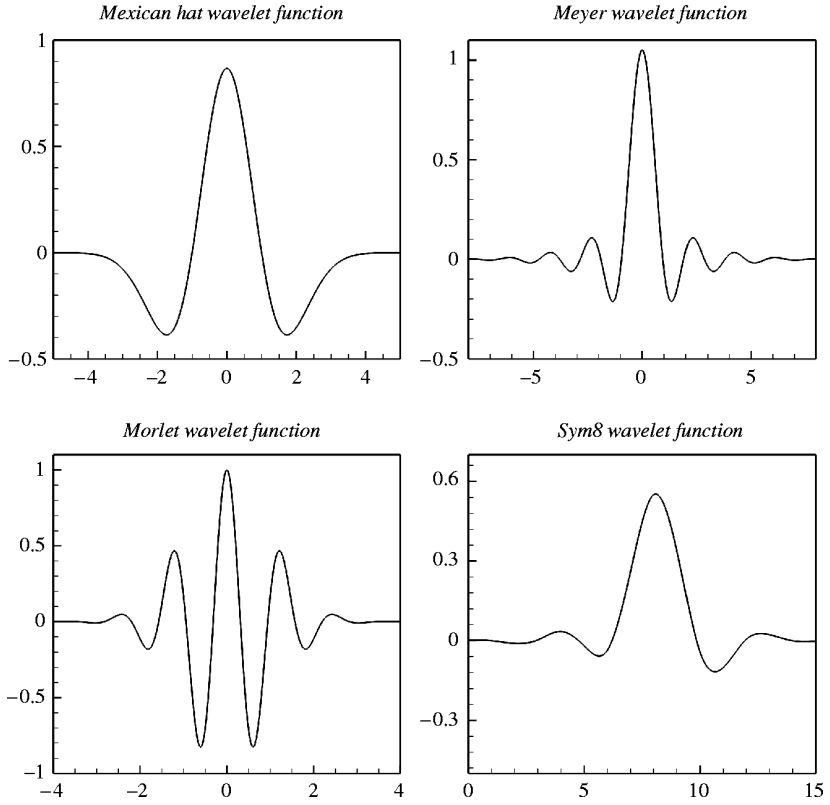


Figure 3.5: Examples of mother wavelet functions. Adapted from [2]

In this section, we show three different types of Wavelet Transform, the Continuous Wavelet Transform, the Discrete Wavelet Transform, and the Stationary Wavelet Transform. They differ between themselves in the content and redundancy they have between scales and translations.

[2]http://www.emeraldinsight.com/content_images/fig/1340180208030.png

3.4.1 Continuous Wavelet Transform

The Continuous Wavelet Transform(CWT) defined in 3.9 on a given scale computes the similarity between a window of the signal with a given mother wavelet function by the inner product. It produces a function of time n and scale s. Scaling a mother wavelet function causes its frequency domain to shrink, as it is inversely proportional to frequency. This is the property that allows the continuous wavelet transform to have a good (although redundant) resolution in both time and frequency .

$$CWT(n,s) = \frac{1}{\sqrt{|s|}} \sum_{i=n-M/2}^{n+M/2} x(i)\psi(\frac{i-n}{s}) \tag{3.9}$$

Using a linear scaling function causes the frequency to have a non linear interval in the frequency domain[ETG12]. This can be fixed applying eq.(3.10) to obtain a linear frequency function.

$$f = \frac{f_c f_s}{s} \tag{3.10}$$

The slight increase in both the shifting and the scaling of the mother wavelet function in the continuous wavelet transform, causes it to be highly redundant in both. While this property on a compression scenario is not desirable, it is useful in an event segmentation/classification scenario.

In [ETG12], the authors test several mother wavelet functions to see which is more suited to analyse a heart sound signal. The authors concluded that the Morlet wavelet fitted best in a heart sound signal, since it minimized the error in both its frequency domain(Fig.3.7), and its energy variability through time (Fig.3.6).

In Fig.3.6, we can see how the wavelet energy fits a Digiscope and an Istethoscope signal. It shows that despite fitting well in the Digiscope signal's energy, in an Istethoscope signal that is not the case since it accentuates the sidelobes of an high intensity peak present between samples 10000 and 12000. In Fig3.7, we can see the signal's frequency content calculated using a parametric approach. From this figure we can conclude that the wavelet captures the signal's frequencies. The type of representation analysis done in [ETG12] however, is not suited in a heart sound segmentation scenario.

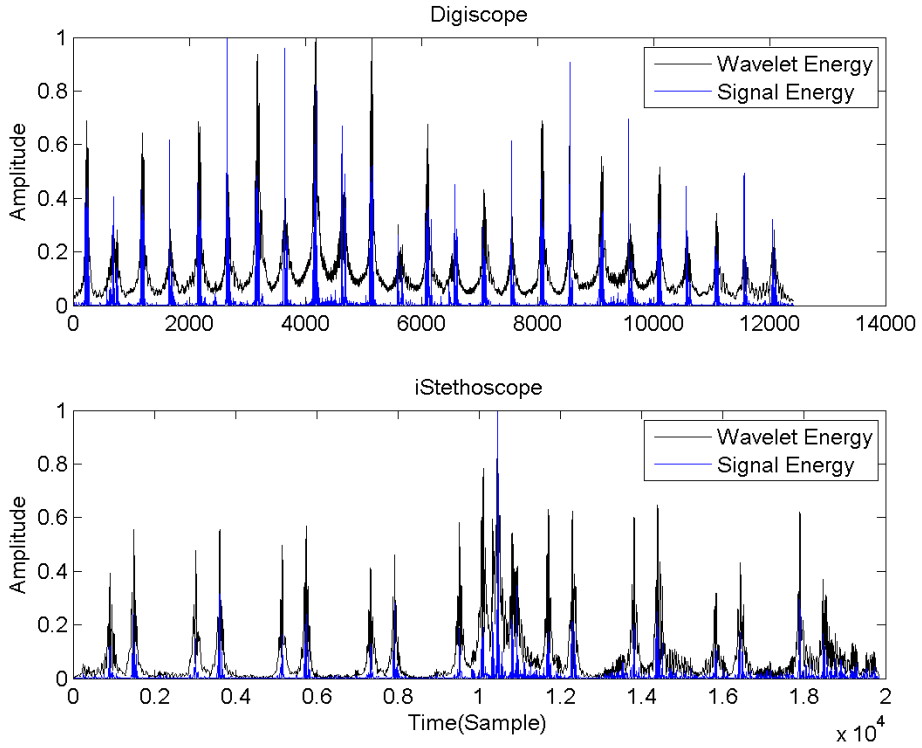


Figure 3.6: CWT Energy Adequacy in a Digiscope and an Istethoscope signal using Morlet wavelet

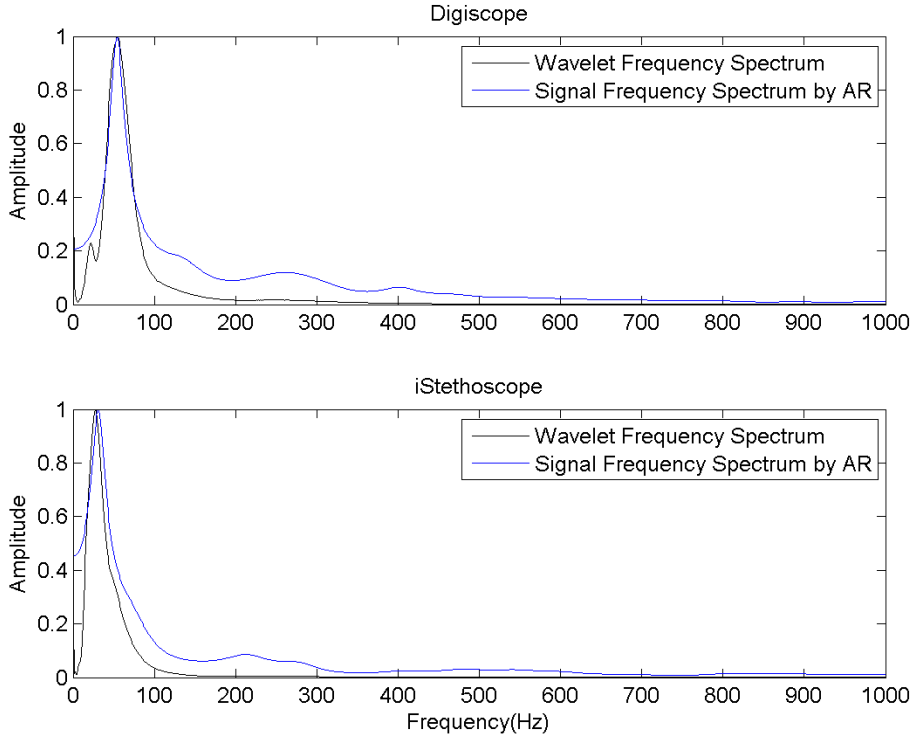


Figure 3.7: CWT Frequency Adequacy in a Digiscope and an Istethoscope signal using Morlet wavelet

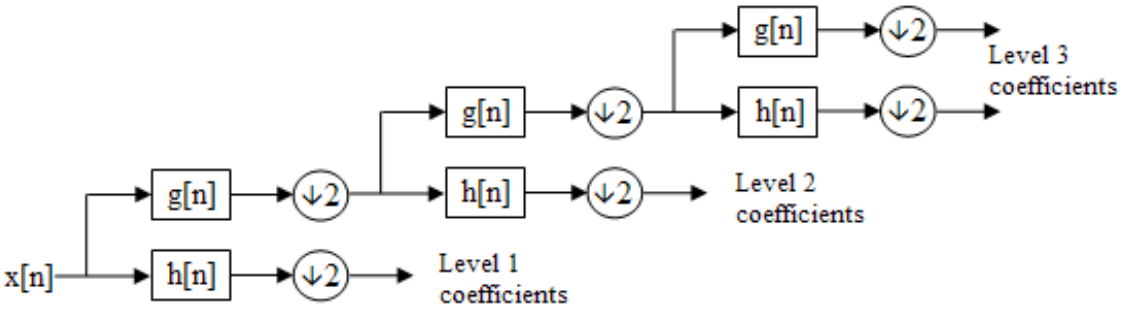


Figure 3.8: The Discrete Wavelet Transform. Adapted from [3]

3.4.2 Discrete Wavelet Transform

Perhaps the most widely used of the Wavelet Transforms, the Discrete Wavelet Transform (DWT)[BGG97] provides a compact and non-redundant representation of signals. It can be fully described by its low-pass filter $g(n)$ and by the high-pass filter $h(n)$. The filters in each scale extract the low and high frequency content of a given signal. This is especially suitable when the main frequency content of the signal lies within a specific range. It is widely used in heart sound segmentation for denoising[HSI97], and as a final representation using a specific scale[KCA+06]. This transform has the downside of the downsampling perfomed in each scale. In a event detection scenario, this means that there is not an one to one correspondence of the detail/approximation coefficients with the original signal. In heart sound segmentation, the DWT is particularly hard to use as the final representation of the signal. In a signal sampled at 4000Hz , one has to pick the fourth scale approximation coefficients to obtain the heart sound approximate frequency content[AT84] in the 0-120Hz range, thus decimating the original signal by a factor of 16.

In Fig.3.9, we can see the approximation coefficients from scales 1 to 4 and the detail coefficients in scale 4 using Daubechies wavelet of order 6. As expected, there is not much difference between the original signal and the approximation coefficients from scales 1 to 4, as the frequency range those scales feature, shown in Table.3.1, contain the natural heart sound frequencies, which are in the 0-150Hz range[AT84]. The detail coefficients, however, in the same scale show much less noise than the approximation coefficients between heart sounds, making the S1/S2 detection easier as the two types of information become more distinguishable which makes this representation of great use for the subsequent stages, i.e. the segmentation and the classification. We can also see the effects of the downsampling where in the original signal we have 4500 samples in the scale 4 we have 300 coefficients. So if we want to detect events using the approximation/detail coefficients, we have to multiply those annotations by 2^{scale}. This type of approach deeply affects the time resolution of the annotations, as we can see in Fig.3.9, making it undesirable for detailed event detection.

3.4.3 Stationary Wavelet Transform

[3]http://upload.wikimedia.org/wikipedia/commons/2/22/Wavelets_-_Filter_Bank.png
[4]http://upload.wikimedia.org/wikipedia/commons/1/16/Wavelets_-_SWT_Filter_Bank.png
[5]http://upload.wikimedia.org/wikipedia/commons/6/6b/Wavelets_-_SWT_Filters.png

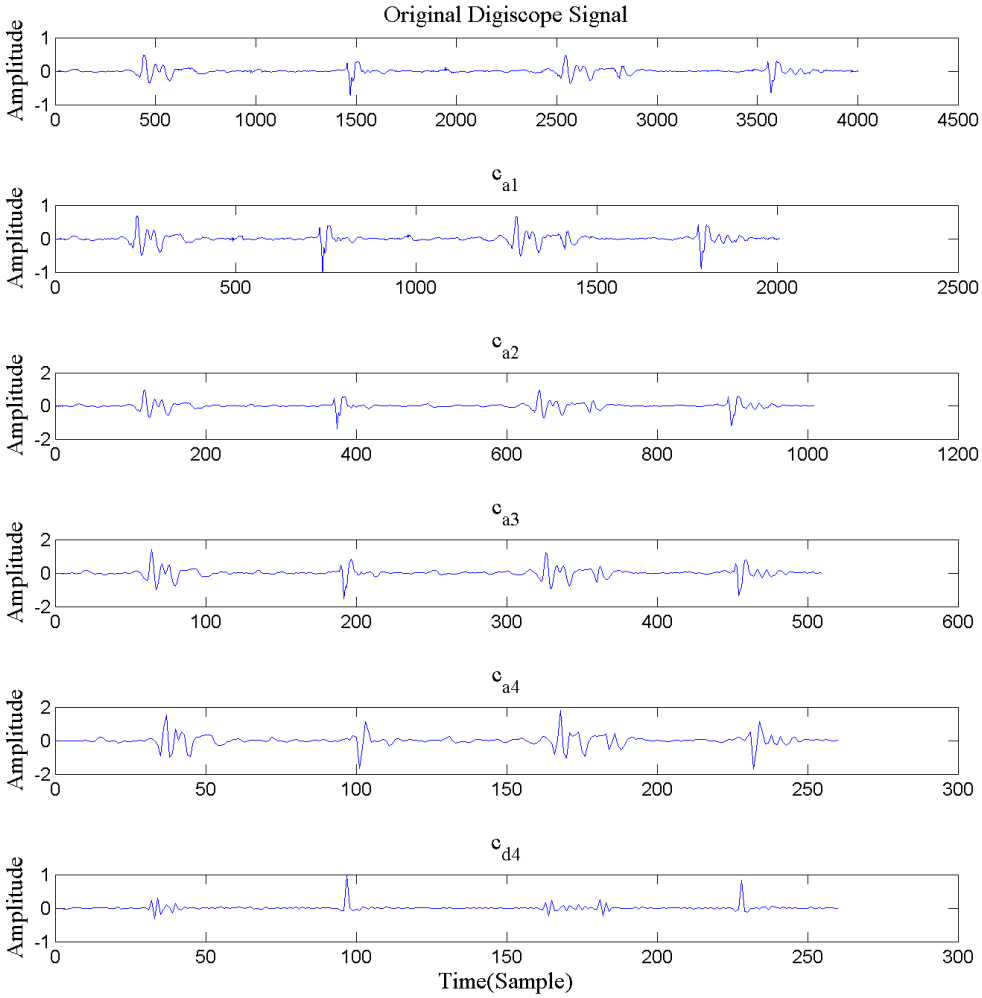
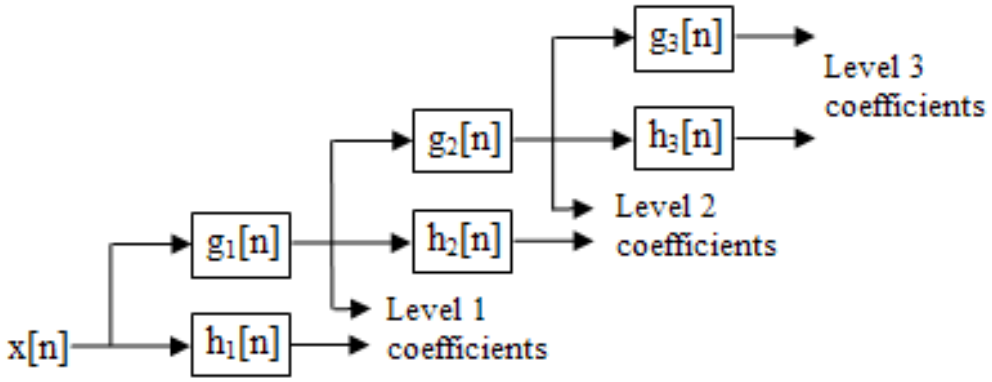


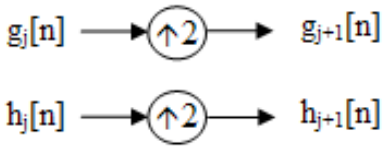
Figure 3.9: Approximation coefficients from scales 1 to 4 and Detail Coefficients in scale 4 of a Digiscope signal using Daubechies wavelet of order 6

Scale	Coefficient Type	Frequencies
1	c_a	0 to 846Hz
2	c_a	0 to 476Hz
3	c_a	0 to 239Hz
4	c_a	0 to 120Hz
4	c_d	66 to 133Hz

Table 3.1: Frequencies captured in each scale with a sampling frequency of 4000Hz, using Daubechies wavelet of order 6



(a) Stationary Wavelet Transform Diagram. Adapted from [4]



(b) Filter upsampling. Adapted from [5]

Figure 3.10: The Stationary Wavelet Transform

The Stationary Wavelet Transform centers itself around a simple idea. Unlike the DWT, where in each scale, after convolving the filter with the approximation coefficients, one downsamples the resulting signal, in the SWT the filter response is upsampled before the convolution. In practical terms, the DWT downsamples the approximation coefficients while maintaining the length of the original filter, making the relative length of the filter bigger. The SWT, on the other hand, upsamples the filter while keeping the length of the original signal, making the relative length of the filter bigger as well. It has the downside however, of being highly redundant having $2NK$ coefficients, where N is the length of the signal and K is the number of scales. In spite of that, it is particularly suitable for event detection, which is the main goal of this work. It maintains the same number of coefficients throughout all scales, thus having the desired one on one correspondence with the original signal.

Fig.3.11 shows the approximation and detail coefficients from scales 5 to 6 after applying the SWT using a Daubechies wavelet of order 10 to an heart sound signal segment. In Table.3.2, we can see to which frequency range corresponds each scale on both approximation and detail coefficients. This transform allows going further up the scales as the coefficients have the same length as the original signal, which would

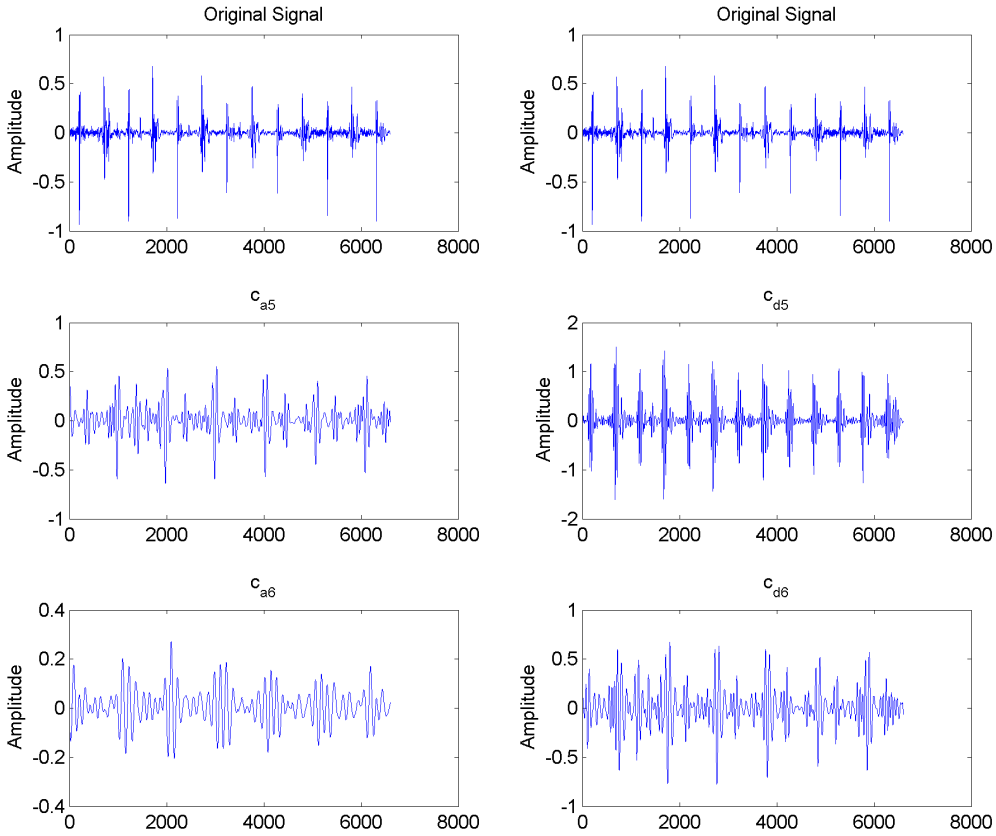


Figure 3.11: SWT Detail and Approximation Coefficients from scales 5 to 6 using Daubechies wavelet of order 10

Scale	Coefficient Type	Frequencies
5	c_a	0 to 59Hz
5	c_d	35 to 64Hz
6	c_a	0 to 30Hz
6	c_d	18 to 33Hz

Table 3.2: Frequencies in each scale with a sampling frequency of 4000Hz, using a Daubechies wavelet of order 10

be unfeasible using the DWT. Going further up the scales allows one to narrow down the frequency range, obtaining more detailed information about the signal.

3.5 Hilbert Huang Transform

Most signal analysis methods have a fixed basis function and work best when applied to stationary signals, i.e. signals that their statistical properties do not change over time. Even some of the recent methods, although designed to handle non stationary data, like the Wavelet Transform, also have a fixed basis function, which does not allow to account for morphological deformations of patterns through time. A representation which, instead of having a fixed basis function, has an adaptive basis function is therefore needed. The Hilbert Huang Transform[HS05] has this desirable property. It consist in two parts: Empirical Mode Decomposition(EMD) and Hilbert Spectral Analysis(HSA). The combination of these two methods allows an adaptive representation of a signal.

3.5.1 Empirical Mode Decomposition

Empirical Mode Decomposition(EMD), is based on the assumption that any signal consists of intrinsic modes of simple oscillations much like the rationale of a Fourier Series. The mathematical meaning of simple oscillation is that, each oscillation denoted by Huang as an Intrinsic Mode Function(IMF) and its first derivative has the same number of null points, and that the oscillation will be symmetric to the local mean. These assumptions were built based on the need to apply the Hilbert Transform, which will be introduced later on, to obtain the Hilbert spectrum, i.e. the instantaneous frequencies over time. The IMF, unlike an harmonic function, does not have the constraint of having constant amplitude and frequency. One can obtain the set of

IMFs with the following procedure:

1. identify all extrema of x(t)
2. interpolate between minima (resp. maxima),ending up with some envelope $e_{min}(t)$ (resp. $e_{max}(t)$)
3. compute the mean $m(t) = (e_{min}(t) + e_{max}(t))/2$
4. extract the detail $d(t) = x(t)m(t)$
5. if mid-stoppage criterion is not met: $x(t) = d(t)$ Go to step 1
6. if final-stoppage criterion is not met: $r(t) = x(t) - d_{final}(t)$; $x(t) = r(t)$

The two most known mid-stoppage criteria are: to compute the normalized squared error between consecutive $d(t)$ and to test if this value is below a certain threshold. This however does not guarantee that the IMF will follow the pre-defined assumptions. The second mid-stoppage criterion is to define a number S, which will define the number of consecutive times the number of zero crossing and extrema will stay the same. If they do it stops. According to Huang[HSL+98], the range of S should be between 4 and 8. The final stoppage criterion also has 2 options.: the first, is when either c_n or r_n become smaller than a predefined threshold.; the second, is when the residue r_n becomes a monotonic function. The final residue represents the trend of the data.

EMD, is used to decompose a signal into a set of intrinsic mode functions. These IMFs are known to provide insight to the physical meaning of the data[HS05]. In Fig.3.12, we can see the difference between the energy means of each type of segment throughout the IMFs. Each line represents the difference between one type and the remaining types of segment. Note that this image was done using one Digiscope illustrating signal. Analysing Fig.3.12 shows that in IMF2 one can identify a segment as a S2 through the segments IMF energy. While in IMF17 we will be able to identify an S1 segment also through the segments IMF energy. This only applies to the particular signal from which the IMF were extracted. Applying Empirical Mode Decomposition to a set of signals, although one can modify the EMD algorithm to iterate until a final number of signals, a given IMF would have different information for each signal. If a given signal is noisier then the EMD is going to have to do more steps for it to converge, resulting into a high number of IMFs. If the signal is cleaner it will result in low number of IMFs. For this reason,the choice of which IMF or IMFs to use

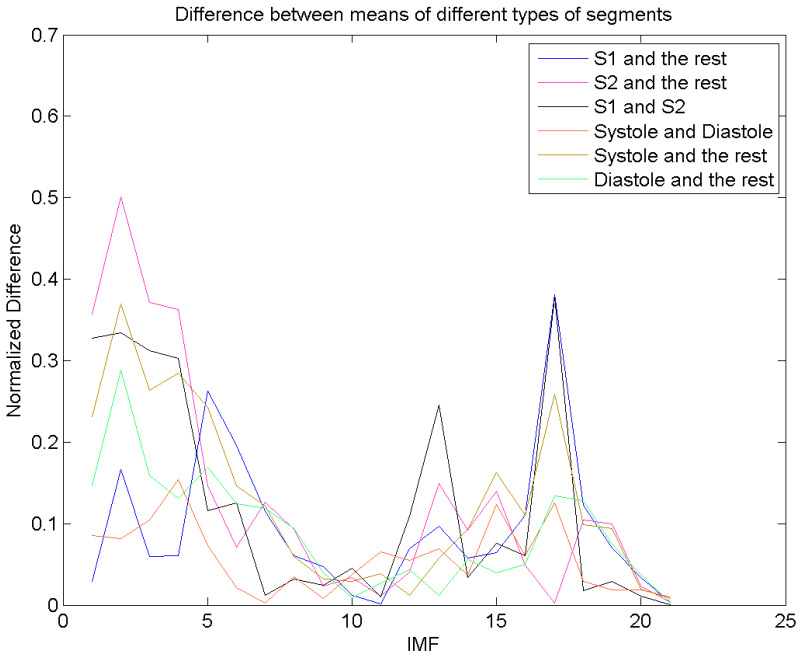


Figure 3.12: Difference between the energy means of each type of segment, i.e. S1, S2, systole and diastole, throughout the IMF's

is a difficult one and usually is done by visual inspection [EMD12, CHJH13]. This method of choosing a parameter of an algorithm does not have any either theoretical or empirical value. So either an automatic method or a quantitative study is required.

3.5.2 Hilbert Transform

Practically, the Hilbert Transform is used to obtain the analytic representation of the signal.

$$\hat{x}(t) = \boldsymbol{H}\left[x(t)\right] = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{x(\tau)}{t-\tau}d\tau \tag{3.11}$$

$$x_a(t) = x(t) + j\hat{x}(t) \tag{3.12}$$

The analytic representation described in eq.3.12 can be used in signal processing for two different purposes. First, to obtain the signals envelope by computing its magnitude. Second, one can obtain the instantaneous frequencies over time, by differentiating the phase of the Hilbert Transform of a given signal. In Fig.3.13 we can see one IMF superimpose with its instantaneous frequencies. This figure suggests that the instantaneous frequencies of an IMF are too noisy to extract any physical meaning. As it was said earlier, another possible use of the Hilbert transform is the

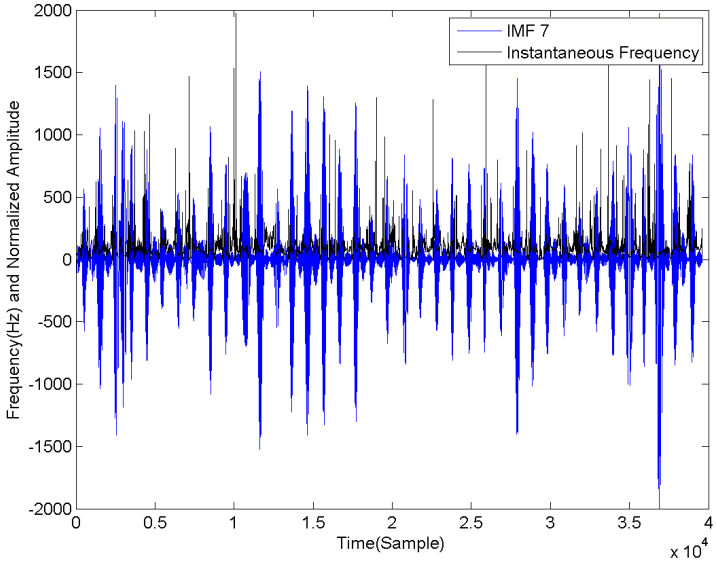


Figure 3.13: IMF7 of a Digiscope signal superimposed with its instantaneous frequencies

Hilbert envelope[ETG12]. This type of envelope is unadvisable to use as the magnitude of the Hilbert transform of a signal attenuates the low-mid intensity while accentuating every other intensities as we can see in Fig.3.14

3.6 Comparison Method

Since the segmentation and classification stages are highly dependent of the Representation, it should be tested in three different types of scenarios. First,a good heart sound representation should maximize the difference between S1 and S2 from the systole and diastole segments; Second, it should also transform the original signal in such a way that facilitates the boundary detection of S1 and S2; Finally it should contain valuable information that will help one, build features in order to correctly distinguish S1 from S2. Having the first and last guidelines in mind, in the following study it was used the maximum as a descriptive feature of a segment. The maximum was chosen given its use in a human manual annotation. Clinicians visually annotate the beginning and end of an S1/S2 by observing the regularity of certain peaks. After choosing the descriptive feature of a segment, a statistic that represented the mean behaviour of a set of one or more types of segment was needed. The median was used as a localization statistic given its robustness facing possible outliers. To complete

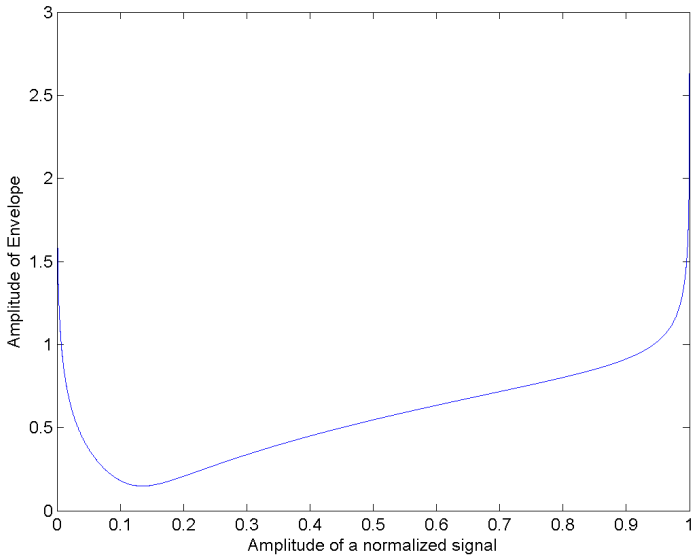


Figure 3.14: Amplitude of Hilbert envelope of a normalized signal

the comparison method design, we needed an operator that compared sets of different types of segments. For this purpose, we used the absolute difference between two sets of segments. To assess the quality of a representation, using the guidelines as our criteria, we design 2 features. The feature represented by eq.3.13, is the absolute difference between the median of the maximums of all S1's and S2's from the median of the maximums of all Systoles and Diastoles. With this method, we can determine the difference between the relevant segments, i.e. S1 and S2, from the non-relevant segments, which is going to be a vital step in the Segmentation stage. Eq.3.14, stands for the absolute difference between the median of the maximums of all S1's from the S2's. This equation allows us to distinguish S1 from S2 which is the main objective in the classification stage. The features were extracted using the annotations from the Digiscope and Istethoscope training datasets. All maximums were normalized to the [0,1] range.

$$g_1 = |median(max(w_{S1,S2}(t))) - median(max(w_{Systole,Diastole}(t)))| \tag{3.13}$$

$$g_2 = |median(max(w_{S1}(t))) - median(max(w_{S2}(t)))| \tag{3.14}$$

3.7 Results

The first and last 4 rows of tables 3.3 and 3.4 show the best Digiscope and Istethoscope results for g_1 and g_2 respectively. In these two tables are featured the best results using

the following parameter range:

- SWT,DWT and CWT Daubechies Wavelet Order=[1,...,40]
- DWT Scale=$[1, \ldots, 6]$
- CWT Frequency(*)=$[20, 40, \ldots, 500]$
- S-T Frequency(*)=$[20, 40, \ldots, 500]$
- SWT Scale=$[1, \ldots, 12]$
- SWT,DWT and CWT Coef=$[c_a, c_d]$
- EMD IMF=7
- HHT IMF=7

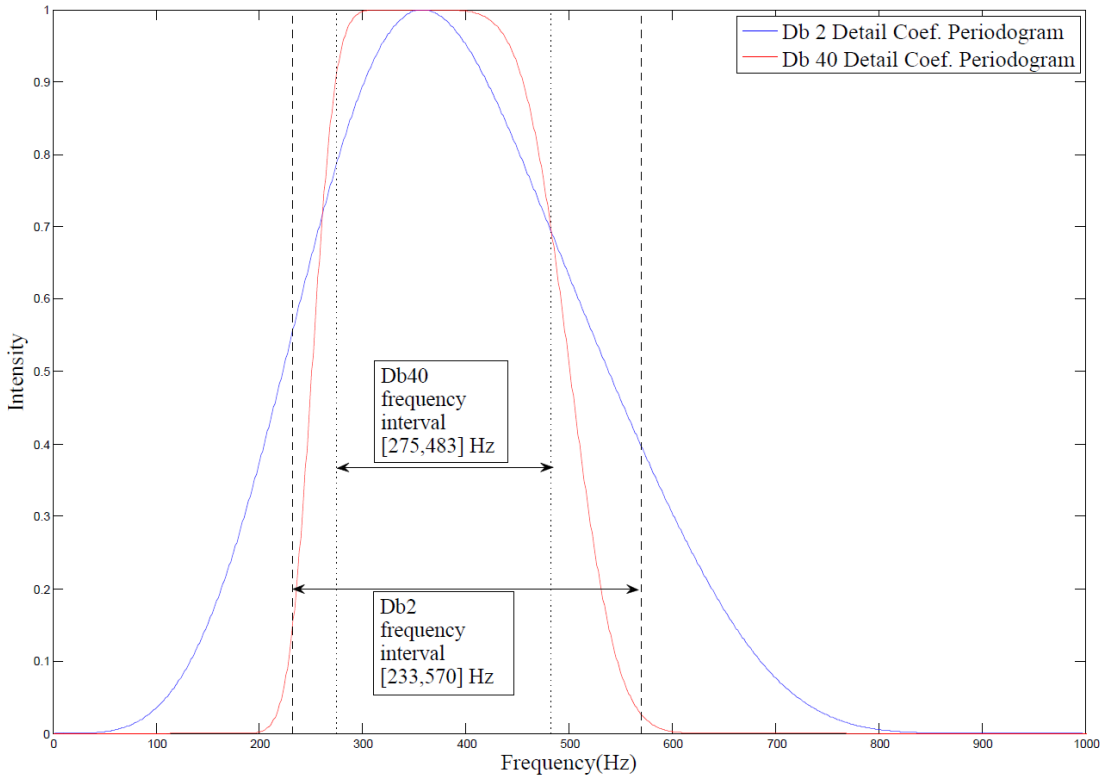


Figure 3.15: Frequency Interval of a Daubechies wavelet of scale 2 and orders 2 and 40.

To successfully analyse the implications of the SWT and DWT parameters, we computed the periodogram of the convolved filters to extract the corresponding frequency

Representation	Order	Scale	Coef	g_1	g_2
DWT	38	3	c_a	0,63	0,014
SWT	1	3	c_a	0,59	0,26
CWT	2	60(*)		0,57	0,22
S-T		380(*)		0,42	0,25
Original Signal				0,57	0,2875
Shannon Energy				**0,70**	0,18
Shannon Entropy				0,35	0,03
HHT				0,28	0,17
EMD				0,31	0,17
DWT	5	3	c_d	0,32	0,42
SWT	15	3	c_d	0,36	0,48
CWT	13	240(*)		0,40	**0,50**
S-T		500(*)		0,42	0,33

Table 3.3: Digiscope Representation Results. (*) : Frequency

interval. Fig.3.15 shows the periodograms of the detail filters of Daubechies wavelet of order 2 and 40. The left and right vertical lines represent the points that mark 0.1 and 0.9 of the area of the periodogram. We used this information as the frequency interval of a Daubechies wavelet of a given order,scale and coefficient type. As we can verify in the Fig.3.15, the larger the order, the narrower the frequency interval.

Looking at Table. 3.3 , one can see the best results for g_1 were by the DWT, SWT, CWT, Original Signal and the Shannon Energy, in lines 1, 2, 3, 5 and 6. Examining the parameters of the Wavelet transforms, we see that the best SWT and DWT representation for g_1 use the approximation coefficients in scale 3. The approximation coefficients of Daubechies wavelets of orders 1 and 38 at scale 3, captures an approximate frequency ranges of 0-314Hz and 0-227Hz, respectively. These frequency ranges cut some high frequency content related to noise usually found in the systole and diastole segments, thus attenuating their amplitude. This results in a higher difference between heart sounds and the diastolic and systolic periods. The SWT order parameter can also be explained by the fact, that the Daubechies of order 1 is the same as the Haar wavelet which in each scale the low-pass filter acts as a moving average obtaining a smoother signal which attenuates the systole and diastole segments and consequently improving the Original Signal's g_1 by 0.02. In a DWT, after 3 successive downsampling operations, the signal is smoothed, and the S1 and

S2 segments are accentuated due to resemblance between the downsampled S1 and S2 segments and the Daubechies low-pass filter at that scale/order. This explains the success of the DWT representation featured in the first row. We can see that the signal itself provides a good representation in S1 and S2 detection perspective, which was expected since, even in a noisy environment the heart sound intensities stand out. The Shannon energy is the best representation for heart sound detection purposes given its power to attenuate low intensities and accentuate the middle ones. The CWT best g_1 scores has parameters of 2 and 60 for the Daubechies order and the frequency, respectively. As it was shown in Chapter.2, the S1 and S2 have a center frequency around 60Hz which fully explains the adequacy of this parameters. The Daubechies wavelet order of the CWT cannot be fully explained due to the high translation redundancy.

Focusing on the lower part of Table 3.3, we can see the best results for the g_2 metric. We can immediately see that the classification stage is harder than the segmentation stage using this feature as comparative basis. Again the DWT and SWT have the same coefficients, which in this case, are the detail coefficients of scale 3, but with orders 5 and 15. The frequency ranges associated with the DWT and SWT are 131-269Hz and 136-249Hz. This result suggests —as the S2 heart sound has a slightly higher frequency content than the S1 —that the S1 is more attenuated than the S2 heart sound, accentuating the amplitude difference between the two. The CWT g_2 results shown in the second last line, are coherent with SWT and DWT transforms, and provide more detail to what is the frequency that best distinguishes both S1 and S2, which is 240Hz.

Table.3.4 shows the Istethoscope's representation results. The difference of the g_1 score between the Istethoscope and Digiscope datasets is due to the noisy and non-controlled auscultation conditions making the Istethoscope a harder signal to segment using only the original signal. There are however, similarities with the Digiscope results. One of them is the high order approximation coefficients in scale 3 for the DWT representation. This result confirms the robustness of this particular coefficients in the segmentation stage. In the CWT there is again a low order Daubechies wavelet and a 60Hz frequency which provides the most efficient separation between S1 and S2 from the systole and diastole type segments. The Shannon energy also provides the best results for the g_1 performance metric.

Focusing on the lower part of the table, we can see the best results for g_2. DWT uses a frequency range between 0 and 115Hz, suggesting that the S2 heart sound is attenutated, thus emphasizing S1. The CWT uses 20Hz, revealing a more detailed

Representation	Order	Scale	Coef	g_1	g_2
DWT	23	3	c_a	0,49	0,02
SWT	2	5	c_d	0,48	0,25
CWT	4	60(*)		0,49	0,29
S-T		500(*)		0,40	0,27
Original Signal				0,40	0,34
Shannon Energy				**0,61**	0,31
Shannon Entropy				0,45	0,09
HHT				0,12	0,13
EMD				0,12	0,15
DWT	23	4	c_a	0,11	**0,41**
SWT	2	5	c_a	0,41	0,39
CWT	4	20(*)		0,31	**0,41**
S-T		380(*)		0,37	0,38

Table 3.4: Istethoscope Representation Results. (*) : Frequency

frequency that characterizes S1 heart sounds. The SWT, despite having slightly worse results than the other 2, uses a narrower frequency range than the DWT, ranging from 0 to 69Hz. The frequency information that differentiates most effectively S1 from S2, is the frequency range associated with the S1 heart sound [AT84].

3.8 Discussion

In this chapter, we started by introducing different types of representations highlighting its advantages and disadvantages towards HSS. Then, in order to make a comparison between different types of segments, we chose the the segments maximum as a representative feature of each segment.

Two representation guidelines were introduced: g_1's aim was to maximize the difference between the heart sound's (S1,S2) and the systole and diastole, and g_2's, to maximize the difference between the S1 and S2 segments. The best representation in terms of g_1 was the Shannon Energy Envelope[LLH97]. This envelope method accentuates the mid intensities while attenuating the low and high ones, making it extremely suitable for heart sound detection. This result motivates our choice to use the Shannon Energy in the peak detection procedure introduced in the next Chapter.

The best representation in terms of g_2 was the CWT. This transform accentuates the high frequencies in the Digiscope dataset and the low frequencies in the Istethoscope dataset. This practical effect is the same as what we want to achieve, using g_2, is the maximum difference between S1 and S2 heart sounds. This result will be later confirmed in Chapter 5, where it is also the best individual metric that achieves higher accuracy classifying a segment as S1 or S2.

In the next chapter, we introduce a novel peak detection procedure that achieved the best segmentation result in the Classifying Heart Sounds Pascal Challenge[BNCM11].

Chapter 4

Segmentation

After representing the signal in a suitable form, the next step is to operate in this signal to detect the segments that are going to be identified in the subsequent stage. The Segmentation stage can be divided in two parts: Peak Detection and Boundary Detection. In the Peak Detection phase, the peaks that correspond to the S1 and S2 heart sounds are picked. In Boundary Detection phase, we then use these peaks to create boundaries that mark the beginning and end of these heart sounds. This chapter is divided accordingly.

4.1 Peak Detection

To get some insight about what is usually done in the Peak Detection phase, we provide an overview on some of its approaches.

The peak detection performed by Liang in [LLH97], uses exclusively thresholds to detect and discard peaks. This type of approach became a standard, that is even used in modern approaches [MDHB13]. Liang's approach is the following: First, a threshold is set to identify peaks above that threshold as S1/S2 candidates—the value of the threshold is omitted as it is a highly data-dependent parameter. This process creates several peaks belonging to the same heart sound. Then, the author uses a low level threshold—50ms—as to reject these extra peaks according to the following rule: if two adjacent peaks are within 50ms, the largest peak is selected, if two adjacent peaks exceed 50ms, the one that is more consistent with every second interval is picked. Since some of the real S1s and S2s are below the threshold, a peak recovery procedure is applied using a high level threshold. One uses the high level threshold to assume

that if the length between two adjacent peaks is above that threshold, a peak was lost. To recover that peak, Liang uses an iterated procedure that lowers the threshold until a peak that is consistent with the intervals that were computed is found, or until an iteration limit is reached.

In Castros paper [CVMC13], S1 and S2 heart sounds are picked if they are at least 150ms apart from each other and if their amplitude is above 30% of the median of the envelopes amplitude.In [EMD12], the author proposes the threshold to be 98% of the envelopes mean value. To reject extra peaks, and to recover missing peaks, the same criteria as Liang is used. In [KCA+06], the authors fully segment the heart sound signal by applying a threshold that is the average value of the envelope. In a nutshell, all of the above mentioned approaches rely on single thresholds to determine what is and what is not a S1/S2 candidate, and whether two peaks are too close/far apart from each other, in order to pick one of them.

4.1.1 SWT Inflection Point Segmentation

The key behind the success of the novel approach introduced in this chapter is the convolution operation performed by the SWT.

Fig.4.1 illustrates the convolution between two identical rectangular functions f and g. The result of this operation is a triangle denoted by the black line. The triangle represents the beginning and end of the similarity of two functions as one convolves through the other. We can verify this statement in eq.4.1, as the convolution between two functions f and g provides for each point the area overlap as g is translated through f. This idea applied to heart sound segmentation is of great use since, if there is a nearly similar mother wavelet function to the S1 and S2 heart sounds, one can immediately get the desired segmentation. This is the main idea of the novel method here introduced. We used SWT's cascade of filters to implement this idea. Although the signal is convolved with multiple low and/or high-pass filters, we can use the convolution's associative property to first convolve all the filters and then perform the final convolution with the signal. Using a convolution of filters that resembles the S1 and S2 heart sounds, in a sufficiently large scale so that the filters are approximately of the same size as the heart sounds, we obtained a series of concavities. Like the resulting triangle in Fig.4.1, the beginning and end of the concavities should delimit the corresponding heart sounds.

[1] http://upload.wikimedia.org/wikipedia/commons/6/6a/Convolution_of_box_signal_with_itself2.gif

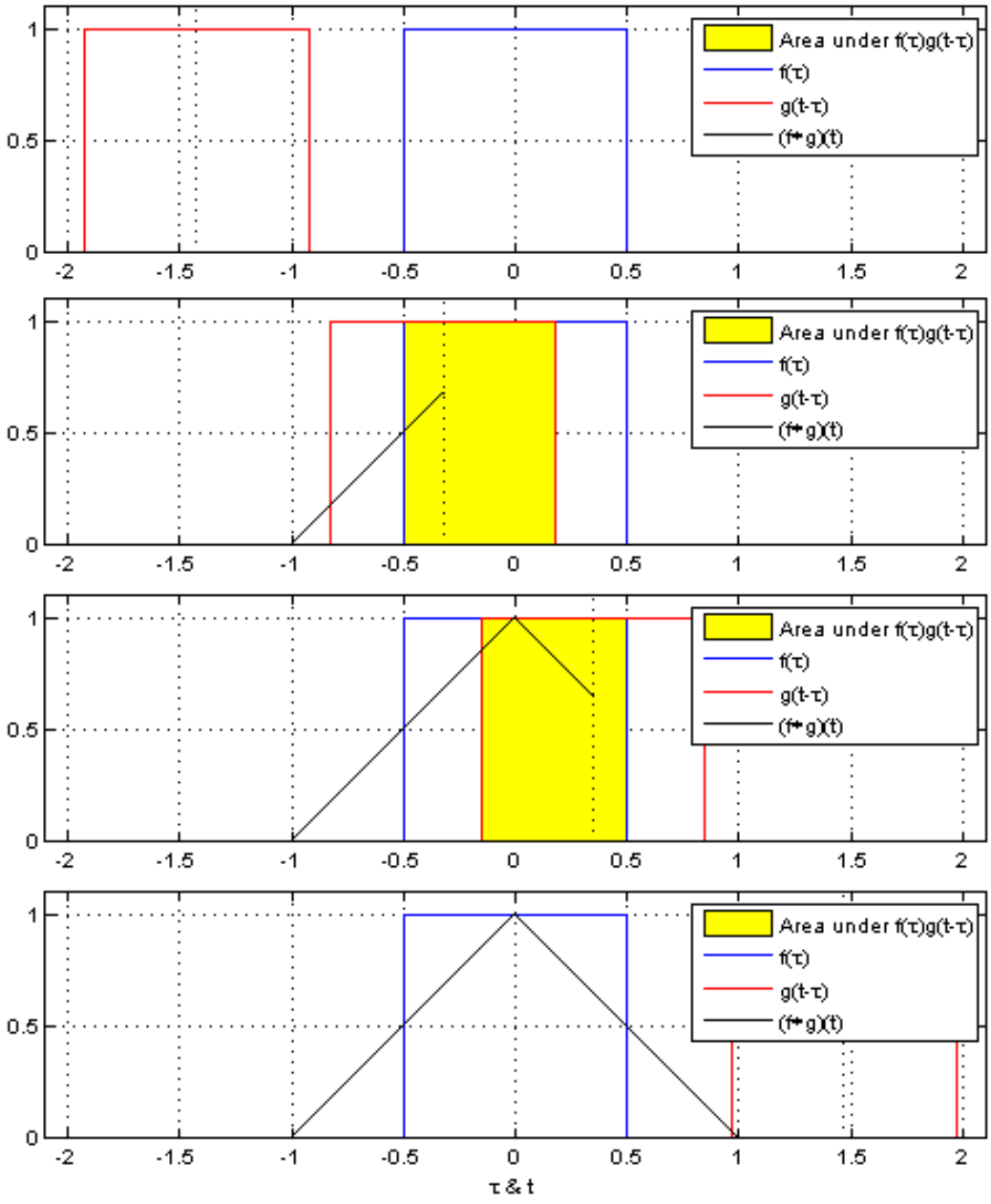


Figure 4.1: Convolution operation illustration between rectangular functions f and g. Adapted from [1]

$$(f * g)[n] = \sum_{m=-\infty}^{\infty} f[m]g[n-m] \tag{4.1}$$

Since we are not able to perfectly segment using this method, we can use it to extract the S1 and S2 representation points, i.e. their main peaks. To perform this task, we use Hierarchical Clustering.

4.1.2 Hierarchical Clustering

After computing the segments, we need to choose a representative feature that describes each segment. This feature should effectively allow a posterior classification system, to label them as S1/S2 or Noise(Systole,Diastole).The maximum absolute value of a segment was chosen as a descriptive feature of each segment. We use the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) method [Mur84], to build an hierarchical tree with the maximum of each segment. This feature resembles the one used performed visually by the clinicians. The UPGMA method is an iterative method that in each step builds one cluster from to sub-clusters which have the nearest distance according to eq. (4.2). This type of cluster aggregation allows us to obtain the S1's and S2's through the top cluster that has the highest median of maximums. This scenario, however, is not always true. The ideal scenario would happen if S1 and S2 heart sounds had the same intensity. This is not true as the intensities depend mostly on where was the auscultation spot. So to obtain the next set of S1 and S2 candidates, we pick the sub-cluster of the low intensity cluster with the highest median of intensities which can be seen in Fig.4.3. Picking only two sub-clusters can be explained by the fact that there are always two sets intensities regardless of where the auscultation spot is. If the auscultation is done in a lower part of the heart that will lead to higher S1 than S2 intensities. If it is done in a upper part of the heart, the S2 intensities will be higher than the S1's. So regardless of the auscultation spot the heart sounds will generally have two distinct patterns: the S1 and S2.

$$\frac{1}{|A|\,|B|} \sum_{x \epsilon A} \sum_{y \epsilon B} d(x, y) \tag{4.2}$$

Occasional errors in this segmentation can occur with two segments reporting the same S1 or S2. This results in redundant detections, that correspond to outliers in the time intervals between detections. To overcome this problem, the median of the intervals between consecutive peaks is computed and peaks producing intervals below half of this median are rejected, by selecting the highest of the two.

The proposed method comprises two procedures, which are a Peak Detection/Recovery phase done with Hierarchical Clustering, and a Redundant Peak Removal step. The procedure starts by computing the Shannon Energy Envelope of the original signal, as it accentuates the heart sound segments while attenuating the low intensity segments—systole and diastole. In the Peak detection, an initial segmentation is performed with the inflection points of the wavelet approximation or detail coefficients. The inflection points provide information about where the concavity changes. As mentioned earlier, these changes represent the beginning and end of similarity between the wavelets approximation/detail coefficients and the envelope of the signal. As one can see in Fig.4.2, the inflection points of the detail coefficients ($Db38c_d$) of Daubechies wavelet system of order 38 at scale 10, correspond to the boundaries of many of the S1 and S2 waves. However, due to noise and the end-effect of the convolution, this correspondence is not true for all S1 and S2 heart sounds. Nevertheless, these inflection points are able to separate segments in which exists one S1, one S2 or none of those (systole or diastole intervals), which will facilitate peak detection.

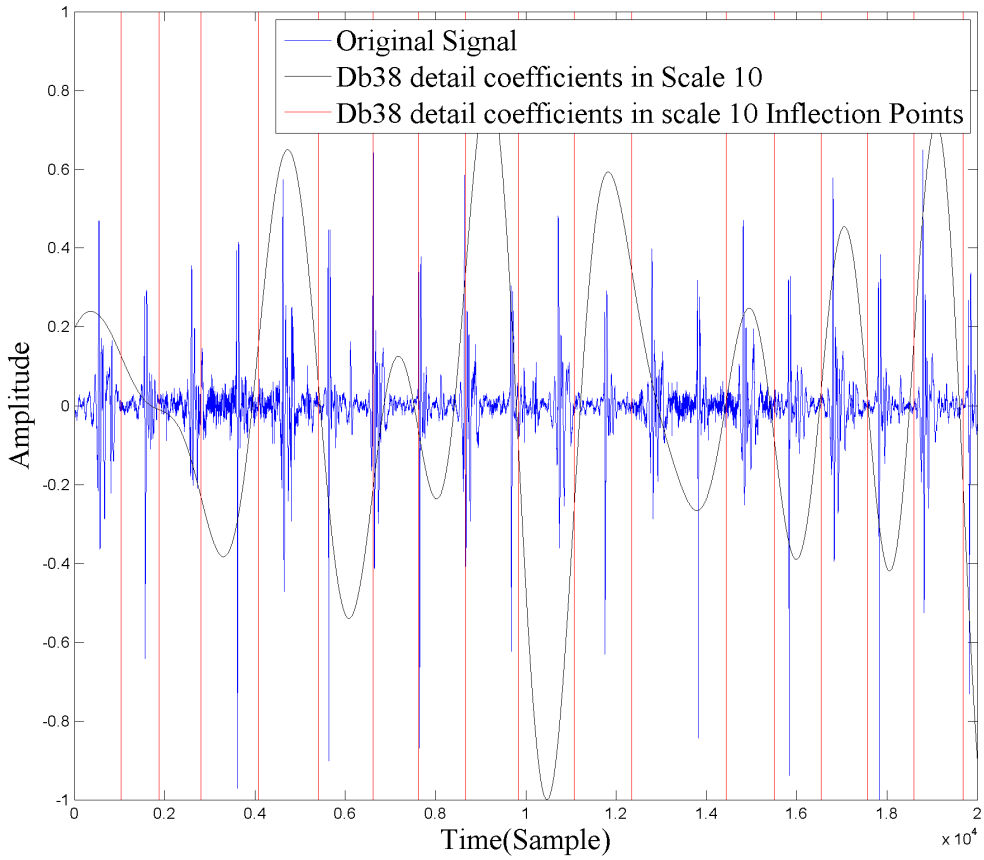


Figure 4.2: SWT Detail Coefficients in scale 10 superimposed with the a Digiscope Signal

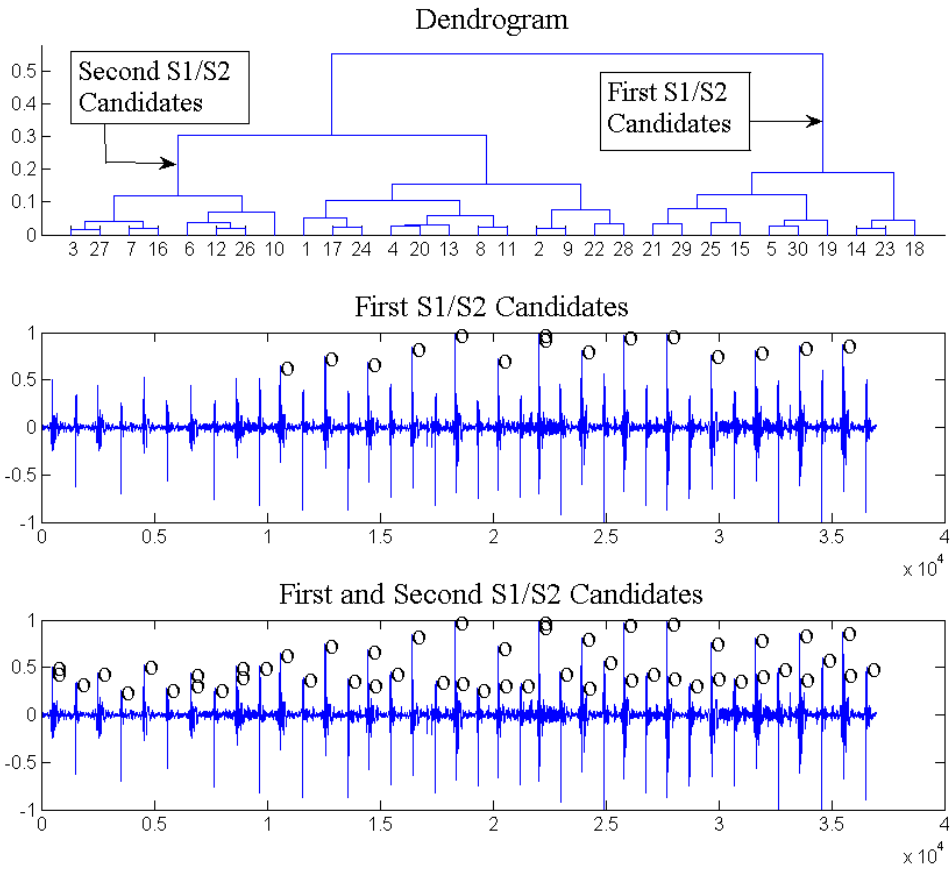


Figure 4.3: Dendrogram and the picked subclusters that represent the first and second sets of candidates (on the top). Candidates overlapped with a sample heart sound signal (two lower axis).

4.1.3 Results

To assess the performance of the method, three different types of metrics were used: Sensitivity and Positive Predictive Value (PPV), shown in eq.(4.3) and eq.(4.4), and total error show in eq.(4.5 and eq(.4.6)[BNCM11]. These first two types of metrics were computed in the training and the test set, respectively. The Sensitivity and PPV were calculated using the manual annotation of each S1/S2 boundary constructed from the provided reference marks in the training set. The criteria used to create the True Positives(TP), False Positives(FP) and False Negatives(FN) was the following:

- TP= Number of segments with detected peaks
- FN=Number of annotations without detected peaks
- FP= Number of detected peaks - TP

For the training set, the computation of total error could not be performed, as no information on how to deal with a missing segment[BNCM11] was available.

$$Sensitivity = \frac{TP}{TP + FN} \tag{4.3}$$

$$PPV = \frac{TP}{TP + FP} \tag{4.4}$$

For the test set, Sensitivity and PPV could not be obtained since the reference annotations were hidden, so only the the Total Error was calculated. The total error is defined in the following way:

$$\delta_k = \frac{\sum_{i=1}^{N_k/2} (|RS1_i - TS1_i|) + (|RS2_i - TS2_i|)}{N_k} \tag{4.5}$$

$$\delta = \sum_{i=1}^{j} \delta_k \tag{4.6}$$

where δ_k is the average distance of the k-th sound clip in the dataset and N_k is the total number of S1 and S2 in the k-th sound clip. $RS1_i$ and $RS2_i$ indicates the real location of S1 and S2 of the i-th heartbeat and $TS1_i$ and $TS2_i$ indicates the calculated location of S1 and S2 of the i-th heartbeat. j is the total of sound clips in the dataset. Finally, δ is the total error obtained after summing δ_k over all j sounds in the dataset.

Wavelet	Order	Scale	Coef
Bi-Orthogonal	$[1, 2, \ldots, 12]$	$[1, 2, \ldots, 12]$	$[c_a, c_d]$
Coiflet	$[1, 2, \ldots, 5]$	$[1, 2, \ldots, 12]$	$[c_a, c_d]$
Daubechies	$[1, 2, \ldots, 40]$	$[1, 2, \ldots, 12]$	$[c_a, c_d]$
Reverse Bi-Orthogonal	$[1, 2, \ldots, 12]$	$[1, 2, \ldots, 12]$	$[c_a, c_d]$
Symlet	$[1, 2, \ldots, 40]$	$[1, 2, \ldots, 12]$	$[c_a, c_d]$

Table 4.1: SWT's searched parameters. c_a and c_d stand for the approximation and detail coefficients.

Wavelet	Order	Scale	Coef	Total Error	Sensitivity	PPV
Bi-Orthogonal	15	10	c_a	74221	0,93	0,84
Coiflet	3	10	c_a	78663	0,93	0,82
Daubechies	9	10	c_a	56732	0,91	0,95
Reverse Bi-Orthogonal	8	10	c_d	72922	0,94	0,94
Symlet	17	10	c_d	59467	0,94	0,94

Table 4.2: Digiscope Representation Results

To tune the wavelet, its order, its scale and its coefficients(Coefs) as best as possible, we performed an exhaustive search in all the wavelets that are currently available in Matlab[MAT11]. Those and the other parameters are featured in Table.4.1.

As we tested only the procedure's segmentation performance we marked every heart sound signal first occurrence as a S1. We also marked the rest of the peaks as the following natural order —S2, S1, S2, S1, etc. These results could all be improved with a posterior identification procedure.

The Daubechies and Symlet wavelets obtained similar and better results than the others. This is due to the high number of orders both wavelets have. While the Coiflet has five, and the Bi-Orthogonal and Reverse Bi-Orthogonal, have 12 available orders in Matlab, the Daubechies and the Symlet wavelets have 40. The slight changes of wavelet's orders, allows a better fit of low and high-pass filters with S1 and S2 heart sounds, which results in a better initial segmentation. Table. 4.2 shows that scale 10 is the most suitable. Scale 10 allows the low and high-pass filters to be upsampled enough to reach a scale where the convolution between them and the original signal result in a signal with smooth waves. These waves represent the beginning and end of the similarity with S1 and S2 heart sounds.

Wavelet	Order	Scale	Coef	Total Error	Sensitivity	PPV
Bi-Orthogonal	3	8	c_a	966197	0,90	0,57
Coiflet	4	3	c_d	1272497	0,93	0,07
Daubechies	18	5	c_d	1127123	0,95	0,22
Reverse Bi-Orthogonal	8	3	c_a	1172971	0,92	0,13
Symlet	17	5	c_a	1035008	0,91	0,31
Daubechies(*)	6	10	c_d	706535	0,97	0,63

Table 4.3: Istethoscope Representation Results

Approach	Total Error	
	Digiscope	iStethoscope
Our Proposed Method	56732	706535
Stanford	76444	1243640
UCL	75569	3394378
ISEP	72242	3905581

Table 4.4: Challenge Results Comparison

Switching our focus to Table.4.3, we can see that the Bi-Orthogonal and the Daubechies wavelet, in the last row, obtained the best results. The wavelet emphasized by an asterisk was applied to the Shannon entropy envelope instead of the Shannon energy envelope of the signal. This choice was made because the Istethoscope dataset had some signals in which there were two error spikes, while the heart sound signals had intensities around $\frac{1}{20}$ of the intensities of those spikes. We suspect that this auscultation was caused by the Istethoscope user that performed that auscultation, hit the Iphone to begin and end the auscultation while placing the Istethoscope in a wrong position making the heart sounds extremely hard to detect. The Shannon Entropy was used to accentuate some low-mid intensities while attenuating the highest ones. Focusing on the Positive Predictive value, we see that lower scales have a lower PPV. This happens since the approximation/detail coefficients have higher variability at lower scales, consequently resulting in more inflection points. Even with the best approach the method still had 0.63, which is a relatively low PPV. This is due to the Istethoscope's inherent noise given the lack of a controlled environment.

Table.4.4 shows the results comparison between our proposed methods and the approaches used in the Classifying Heart Sounds Pascal Challenge. As one can see, our methods out-performed all of the other approaches in both Digicope and Istethoscope

datasets.

4.2 Boundary Detection

Although commonly overlooked by several Heart Sound Segmentation approaches, the boundary detection phase is paramount. It creates the final boundaries of every detected S1 and S2. These boundaries are used not only to build features that allow us to distinguish S1 from S2 heart sounds, but also in further procedures, e.g. automatic classification of heart disease procedure.

In [LMA49], the author mentions that the maximum duration of a S1 and S2 is 150ms. As the delimitation of S1 and S2 heart sounds can only be done with the use of the phonocardiogram and their maximum duration is a reasonably small quantity, some small amount of error is expected and allowed. Liang's procedure vaguely states that the duration of those heart sounds is between 20ms and 120ms without mentioning the procedure that leads to these values and how these values are used to segment the heart sounds. An alternative approach in boundary detection is the use of thresholds that automatically detect the S1/S2 boundaries. Moukadem, used this approach in [MDHB13], which shows a study that provides the Mean Absolute Error between the real S1 and S2 boundaries and different choices of threshold. It shows that the dataset's ideal threshold for S1 is 10% and the ideal threshold for S2 is 15% of the maximum amplitude of the Shannon energy envelope. To choose a single threshold that segments both S1 and S2 with near equal accuracy, the author picks 10% of the maximum value of the envelope as the final threshold. This threshold detects the S1/S2 boundaries. This type of boundary detection requires further processing given the possible presence of murmurs and artifacts with an amplitude large enough to be captured and consequently segmented by this threshold.

In this section, we present two novel boundary detection methods and compared with a baseline threshold based approach[KCA+06] approach. The first approach(Eq.4.8) is based on the difference between the variation of a main segment $s(i)$ and its neighbours $s(i-1)$ and $s(i+1)$. The aim of this approach is to vary the length of the main segment—and consequently its neighbours—to maximize the difference between the main segment's and its neighbour's standard deviations. This procedure was used on the Shannon energy envelope of the signal, as it is the best representation in terms of distinguishing S1/S2 from noise segments. The lengths of each segment varied between

60 and 150ms [LMA49].

$$s(i) = \left(\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 \right)^{\frac{1}{2}} \tag{4.7}$$

$$VBS = s(i) - s(i+1) - s(i-1) \tag{4.8}$$

The second novel approach is based on the longest increasing/decreasing sub-sequence. We apply a longest increasing and decreasing sub-sequence algorithm(LISS and LDSS) in the Shannon Energy of the signal, to extract the beginning and end point, respectively. The window length was 150ms. Fig.4.4 illustrates this method. The rationale for this particular approach is that the Longest Increasing and Decreasing Sub-Sequence in a window that contains a S1/S2, should obtain its beginning and end.

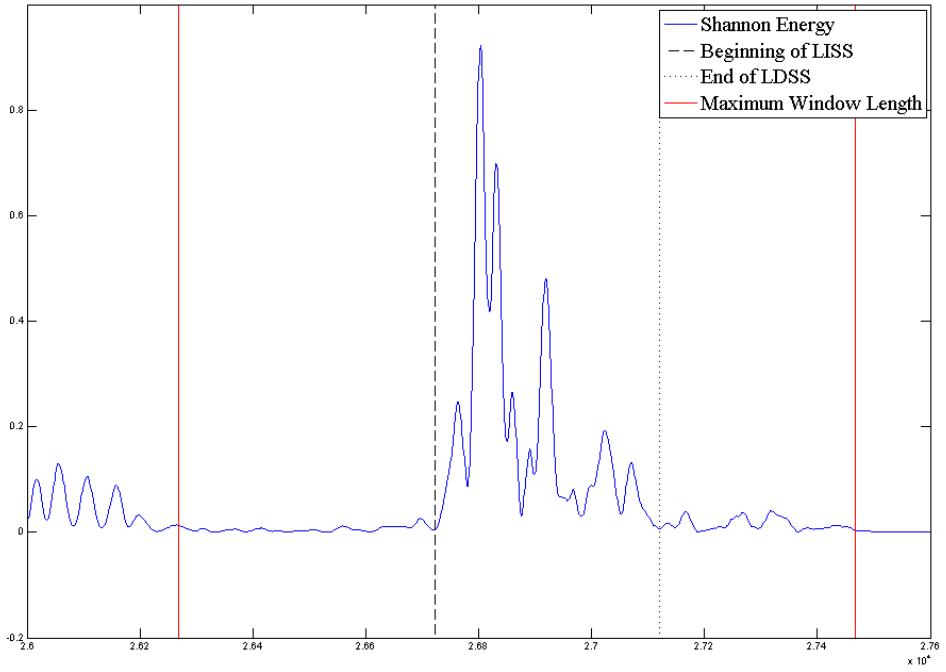


Figure 4.4: A boundary Detection perform by LISS and LDSS

The baseline approach is to detect the first and the last zero crossings in a window length of 150ms from the Normalized Shannon Energy (SE_{norm}). The following equation describes SE_{norm} :

$$SE_{norm} = SE - < SE > \tag{4.9}$$

Where <> represents the average operator.

Approach	Annotation Error	
	Digiscope	iStethoscope
a_1	29,1±14,3	37,1± 13,4
a_2	41,4± 10,8	67,1±15,2
a_3	46,8± 15,2	83,2± 20,4

Table 4.5: Boundary Detection Results in miliseconds. Mean±Standard Deviation. a_1=Difference between Variations. a_2=Longest In/Decreasing SubSequence. a_3=Threshold-based approach

The approaches were tested in the training set with the following error metric:

$$Error_{Annotation} = MAE_{leftBoundary} + MAE_{rightboundary} \tag{4.10}$$

In Table.4.5, we can see the boundary detection results for both Digiscope and Istethoscope datasets. The success of the Difference between variations approach can be explained by its iterative procedure and the metric used. A relative variation between a segment and its neighbours provides a better mathematical description of what is the most peaky subsequence of a given segment. The threshold based approach had the worse result. This result was predictable because a threshold is a parameter that can be tuned for a specific dataset, not for a wide range of datasets.

Despite achieving good results with the Difference between Variations approach, these boundaries must not be used for the Identification stage. Apart from the pre-processing stage, the success of every other stage of Heart Sound Segmentation, depends on the success of previous stages. So, every extra stage-dependency in the algorithm is ill-advised. Our suggestion is to use the maximum known duration of an S1 and a S2 (150ms) to build good classification features that allow us to distinguish an S1 from an S2. The procedures presented here should be used for the computation of a segment's duration, which is a feature of paramount importance, as we show in the next chapter.

4.3 Discussion

We start out this chapter by introducing the novel method. It uses first the Shannon Energy envelope to maximize the difference between S1/S2 heart sounds and the systole/diastole, as stated in the previous chapter. Then, we take advantage of the convolution process performed by the SWT to perform an initial segmentation by

using the inflection points of its coefficients. As this procedure produces an imperfect segmentation, we use hierarchical clustering to automatically choose the segments that contain S1's or S2's. In some cases, hierarchical clustering picks consecutive segments that represent the same peak. To overcome this problem, we use half of the median of the intervals between peaks to discard peaks that are below this threshold.

To achieve the minimum total error, we varied the wavelet types, orders, scales and coefficients types (approximation/detail). The best Digiscope parameter combination was the approximation coefficient of Daubechies 9 at scale 10. To achieve a minimum result of the Istethoscope dataset, we modified the novel procedure to use the Shannon Entropy instead of the Shannon Energy envelope. As the Shannon Entropy accentuates more the low intensities than the Shannon energy, this allowed us to detect some very low peaks that were, in fact, heart sounds. As we wanted to achieve the best possible total error in the test set we searched these parameters directly on the test-set for both the Digiscope and Istethoscope datasets. Although we did not search the training set we found that the best performing parameters—in terms of total error—in the test set matched the best parameters in the training set using the available performance metrics—sensitivity and ppv.

In the boundary detection section, we introduced two novel methods: the Longest Increasing/Decreasing Subsequence(LISS/LDSS) and the difference between variations. The LISS/LDSS is based on the assumption that within a segment the longest increase is the beginning of the heart sound and the longest decrease is the end. The Difference between Variations searches between the known lower and upper limits of a S1/S2 duration(60ms and 150ms), and maximizes the difference between the standard deviation of the main segment and its neighbours. These two novel boundary detection methods were compared against a threshold-based method. The Difference between Variations out-performed the other two methods achieving results comparable to Moukadem's approach[MDHB13]. These results are questionable as the boundary detection ground truth was annotated by the author. The ground truth was built by the author using the original annotations provided in the Classifying Heart Sounds Pascal Challenge, thus even though there may be some small translation errors, there is no major errors, i.e. absent or duplicate annotations.

In the next chapter, we focus on the classification stage. We will use the representations that were introduced and compared in the previous chapter to build representative features that allow us to distinguish S1 from S2 heart sounds.

Chapter 5

Classification

In the Classification stage, after obtaining the unlabelled segments, one classifies each segment as S1, S2, Systole or Diastole. Given the natural order of these events, we can reduce the problem of classifying systole and diastole, consequently classifying the adjacent segments as S1 or S2. This approach is mostly used as the duration of a systole/diastole is usually the key feature that allows one to distinguish S1/S2.

The usual way of classification is to apply machine learning techniques. Machine learning is a branch of artificial intelligence where its main focus is on the design of systems that can learn from data. We can successfully divide machine learning algorithms by the type of learning it uses. Two of these types are supervised and unsupervised learning. While the supervised approach requires the correct labelling , the unsupervised approach does not require the correct labelling of instances. In this scenario one usually splits the dataset into two parts: the training set and the test set.

The supervised machine learning algorithm builds rules using the training instances and predicts the labelling of the test instances. To perform classification using the unsupervised approach the algorithm discovers the structure of the training set, and applies some a-priori made rules to predict in the test set. As it was stated earlier, the key feature is the duration of the systole/diastole. This is due to the clinical fact that the diastolic duration is longer than the systolic one. This fact, however, does not apply to patients with an elevated heart rate, so other methods that describe S1 and S2 heart sounds well enough to distinguish one another, are required. Another possible solution is evaluate each segment's maximum amplitude. This approach achieves good results if the auscultation spot is always the same, however this is not the case in both Digiscope and Istethoscope datasets. Another possibility is to describe its frequency

content.

5.1 State of the Art

In [AT84], the author mentions that S2 frequency content its slightly higher than S1, as it was shown in Chapter 2 for the Digiscope dataset. However in noisy datasets such as the Istethoscope dataset this fact is harder to assess. A. Castro in [CVMC13], proposed an frequency based approach that improves the S1/S2 classification rate using the detail coefficients at levels 3 and 4 of a Daubechies Wavelet of order 6. Moukadem in [MDHB13], goes even further and uses a time frequency which obtained better results. In this chapter, we present different features using a suppervised approach. We classify each segment using the maximum absolute value of each segment and its adjacent segments using the representation that were presented in Chapter 3.

5.2 k-Nearest Neighbours

To validate and compare the Classification methods here presented, we used the k-Nearest Neighbours model[Dud76] (KNN), with k=3. We chose this model/parameter also to compare with A.Moukadem's approach [MDHB13] that used the same model setting. The KNN is a straightforward classification method where the computation of an instance distance matrix can be seen as the model training phase. To predict in a classification scenario, given an unlabelled test instance, the method searches the distance matrix for the top k most similar instances, and predicts the value of the label through majority voting. This method is illustrated in Fig.5.1, where the red star is the test instance that is going to be classified as a Class B instance, since that is the majority of the cases with $k = 3$.

5.3 Experimental Methodology

Experiments were performed using K-Fold Cross Validation[EG83], with K=5. We used this approach to compare with Moukadem's results in [MDHB13] as he also uses the same experimentation method. The K-Fold Cross Validation operates in the

[1]http://1.bp.blogspot.com/5lJFjAYvQ7Q/UIr7Qp1wvOI/AABAc/r_KtPGaat9A/s1600/knnConcept.png

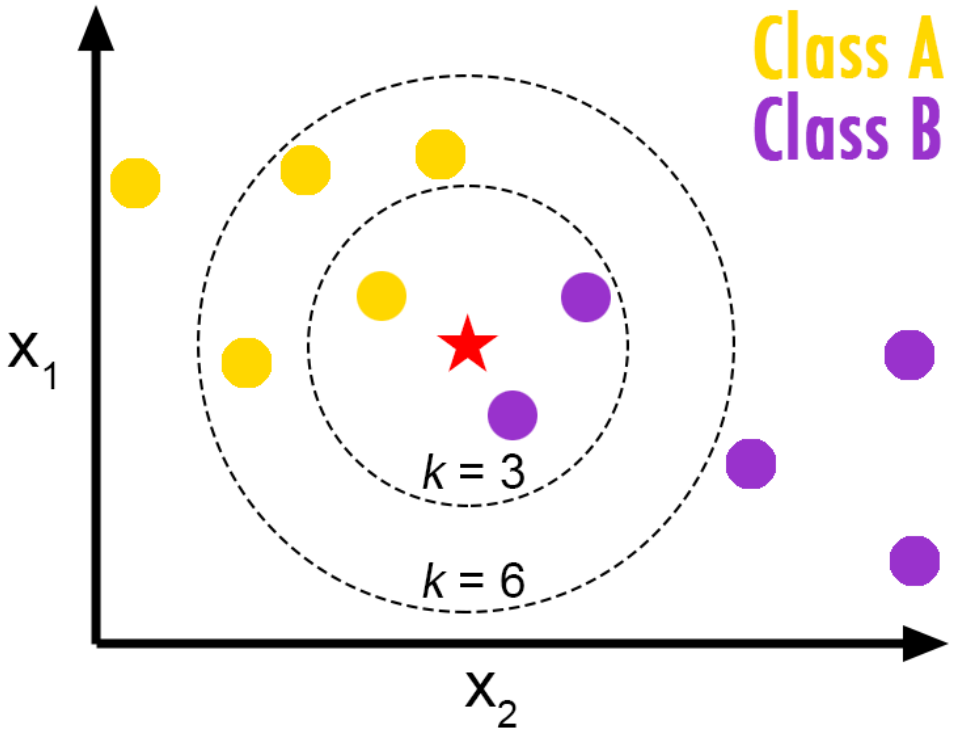


Figure 5.1: K-Nearest Neighbours Algorithm. Adapted from [1]

following way: first, the training set is divided into K sub-sets. Iteratively, using sub-set 1 to K for posterior testing, the used model is trained in $K-1$ remaining sub-sets. A performance metric is computed in each folded, and after finishing the procedure, averaged to give an overall performance metric. The 5-Fold Cross Validation is illustrated in Fig.5.2. The performance metrics used were: Accuracy, Sensitivity and Specificity. The Accuracy and Specificity are described by equations.(5.1) and (5.2). The sensitivity was already introduced in the previous chapter, but it is repeated in 5.3

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5.1}$$

$$Specificity = \frac{TN}{TN + FP} \tag{5.2}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{5.3}$$

[2]http://blog.neuroelectrics.com/Portals/181943/images/Kcrossfold-validation_machine_learning_EEG.png

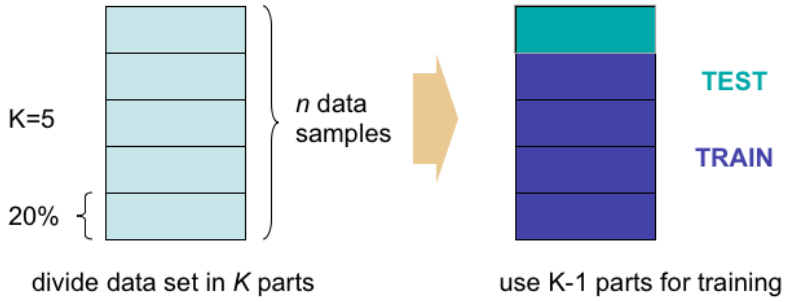


Figure 5.2: 5 fold cross validation. Adapted from [2]

5.4 Results

In order to present a more complete study, we show the results in three stages. First, only the maximum absolute value of the segment as a feature. Then, we use the maximum absolute values of the segment and its neighbouring segments (Systole,Diastole). Finally, we perform all the possible combinations of the individual and neighbourhood features to obtain the final feature set. The searched parameters were the same as the representation ones:

- SWT,DWT and CWT Wavelet=[1,...,40]
- DWT Scale=$[1, \ldots, 6]$
- CWT Frequency(*)=$[20, 40, \ldots, 500]$
- S-T Frequency(*)=$[20, 40, \ldots, 500]$
- SWT Scale=$[1, \ldots, 12]$
- SWT,DWT and CWT Coef=$[c_a, c_d]$
- EMD IMF=7
- HHT IMF=7

Table 5.1 shows the Digiscope and Istethoscope results, using individual features. Focusing on the upper part of table, we can see the the CWT obtained the best results. This was expected as the CWT also obtained the best results in Table.3.3 and Table.3.4 in Chapter 3. By capturing only frequencies(420Hz), the CWT distinguishes better S1 from S2 due to the fact that the S2 heart sound has a slightly higher frequency

Approach	Order	Scale	Coef	Accuracy	Sensitivity	Specificity
SWT	9	3	c_d	0.66	0.68	0.65
CWT	23	420(*)		**0.70**	0.71	0.69
DWT	39	4	c_a	0.67	0.67	0.68
HHT				0.52	0.57	0.47
S-T		480(*)		0.58	0.57	0.58
EMD				0.51	0.53	0.49
Original Signal				0.56	0.55	0.56
Segment Duration				0.55	0.58	0.51
Digiscope						
SWT	1	1	c_a	0.65	0.67	0.64
CWT	2	300(*)		**0.72**	0.72	0.72
DWT	1	6	c_a	**0.72**	0.73	0.72
HHT				0.53	0.55	0.50
S-T		500(*)		0.66	0.62	0.71
EMD				0.56	0.58	0.54
Original Signal				0.54	0.54	0.54
Segment Duration				0.49	0.52	0.46
Istethoscope						

Table 5.1: Individual Classification Feature results. (*) : Frequency. Above the double-line are the Digiscope Results, and below, the Istethoscope's.

content. Still, the accuracy obtained by only the maximum of the segment in the CWT representation at high frequencies is too low. We require more information about a segment in order to perform a better Classification of S1 and S2. The SWT and DWT obtain similar accuracies to the CWT. They use, however, different information about the frequency content of a segment. The SWT combination of parameters that obtain the highest accuracy uses scale 3 detail coefficients. These coefficients contain frequency information in the range of [134-256Hz]. Like the CWT, choosing a high frequency range in the SWT, allows the representation to give higher values to the coefficients that contain higher frequency information. This results in a higher difference between the coefficients that contain lower and higher frequency content. The DWT uses scale 4 approximation coefficients which feature a frequency range of [0-114Hz]. Unlike the CWT and SWT, DWT attenuates high frequency content, usually associated more with the S2 heart sounds, the practical result is the same. Accentuating the lower frequency content results in the same frequency differentiation

done in the CWT and SWT. The other and the original signal results, shown in rows 4 to 7, had the worst results. The HHT and EMD's parameters were not searched, as the information in the EMD's intrinsic mode functions (IMF), varies with the signal's length and noise. The original signal could provide good results in the Classification if the auscultation spots was always the same. In this scenario one heart sound would always have an higher intensity than the other, allowing a simple Classification using exclusively the original signal. Hence, the low accuracies obtained using the original signal, suggest that different auscultation points were used in the auscultations performed. The segment's(heart sound) duration is known as the key feature that performs a better distinction between S1 and S2 heart sounds. The bad results obtained by the Segment's Duration are due to the fact that the S1 and S2 lengths have similar lengths. What differs most from one another is the following segment's length,i.e. the systolic/diastolic period.

On the lower part of table5.1, we can see the Istethoscope results. Again, the CWT achieved the best overall results. Despite using a lower frequency, the same principle applies in this case. The lower frequency suggests the absence of high frequencies in the Istethoscope's heart sound segments. The SWT, in this case, uses an unexpected scale. Given that the Istethoscope signals were downsampled to have a sampling frequency of 2205Hz, approximation coefficients in scale 1 contain frequency information in the range of [0-516Hz], so the previous CWT interpretation is, in this case, no longer valid. This scale suggests that applying a low pass filter with a cut-off frequency of 516Hz smooths the Original Signal in a way that it attenuates the extremely high frequency noise, allowing a more clear distinction of S1 and S2 heart sounds. The DWT, on the other hand, uses completely different frequency information while achieving the same results as the CWT. Approximation coefficients in the downsampled Istethoscope on scale 6 represent the signal's frequency content in the range of [0,28Hz]. This results in an attenuation of the rest of the frequency content of the signal, which will capture more S2's than S1's given their slightly higher frequency content. S-T achieved better results in the Istethoscope. This result increase suggests that higher frequencies are more present in this dataset than on the Digiscope. These higher frequencies allow a better distinction of S1 from S2 heart sounds. All of the other representations achieved similar results to the Digiscope's.

Table.5.2 shows the Classification results using the maximum absolute value of the heart sound and its neighbouring segments, Systole and Diastole. Using the segments information and its neighbourhood increases the SWT result, in both Digiscope and Istethoscope by 10%. This means that, despite the Systolic and Diastolic period

Approach	Order	Scale	Coef	Accuracy	Sensitivity	Specificity
SWT	30	6	c_a	**0.75**	0.77	0.73
CWT	3	340(*)		0.71	0.73	0.68
DWT	33	6	c_a	0.70	0.70	0.71
HHT				0.51	0.54	0.48
S-T		460(*)		0.61	0.62	0.61
EMD				0.50	0.54	0.46
Original Signal				0.59	0.62	0.56
Segment Duration				**0.81**	0.83	0.80
Digiscope						
SWT	38	4	c_a	**0.76**	0.83	0.69
CWT	3	260(*)		0.73	0.75	0.70
DWT	33	4	c_a	0.74	0.76	0.72
HHT				0.45	0.48	0.41
S-T		460(*)		0.68	0.71	0.64
EMD				0.51	0.53	0.50
Original Signal				0.58	0.60	0.57
Segment Duration				0.70	0.71	0.67
Istethoscope						

Table 5.2: Neighbourhood Classification Feature results. (*) : Frequency.

having usually low maximums, they form different patterns for S1 and S2 heart sounds. SWT uses scales 6 and 4 approximation coefficients which contain frequency content in the range of [0-17Hz] and [0-63Hz] in the Digiscope and Istethoscope datasets, respectively. The low frequency ranges allow us to distinguish more accurately the Systolic from the Diastolic periods. The CWT did not have a significant result increase using the neighbouring segments suggesting the absence of differentiating patterns in the neighbouring segments. DWT achieved the best results, in both the Digiscope and Istethoscope datasets, using the same scale and coefficients as the SWT. These results show the robustness of the parameters using the heart sound's neighbourhood. The Segment duration had the highest result increase in the Digiscope dataset. This was expected since the lengths of S1 and S2 are in many cases indistinguishable, while the Systolic and Diastolic periods have different lengths—the Diastolic period is usually longer. The Istethoscope Segment Duration results had almost -10% accuracy than the Digiscope's. The noise, which heavily present in this dataset, introduces some randomness in the systolic and diastolic periods maximum absolute values resulting

Type of Feature	Approach	Accuracy	Sensitivity	Specificity
Individual	CWT+ST	0.86	0.88	0.84
Neighbourhood	SWT+DWT+ST	0.83	0.86	0.80
	Digiscope			
Individual	CWT+DWT+HHT+ST+EMD	0.90	0.91	0.89
Neighbourhood	CWT+DWT+ST	0.92	0.90	0.94
	Istethoscope			

Table 5.3: Combination of the best Classification Features.

in the 10% decrease.

Table.5.3 shows the results of the best combination of Individual and Neighbourhood features, in the terms of accuracy, for the Digiscope and Istethoscope datasets. Surprisingly, the features contained information about the same frequency content, as it is the case of SWT with DWT and CWT with S-T. It was also unexpected to see the S-T be part of every top combination of features. This particular results suggest that other Transforms, such as the Wavelet or the Hilbert-Huang, are not able to capture some information that S-T can. While the other representations try to solve time/frequency trade-off by attributing longer windows to lower frequencies and vice-versa, the S-Transform uses the same window length for all frequencies. Since S-Transform achieves the best Classification result using high frequencies, we argue that its lack of time-frequency resolution results in aggregating multiple high frequencies into a single coefficient. Consequently, creating higher valued coefficients that allow a better distinction between S1 and S2 heart sounds. Comparing to Moukadem's results, we see that these combinations of features in the Digiscope dataset achieved lower results. This can be explained by the fact that our approach was tested on younger patients. Distinguishing heart sounds in children pose a much more complex problem than in older patients. As the children's heart rate is higher than other age groups, the length of the Systolic and Diastolic periods become indistinguishable, and consequently, its Time/Frequency content as well. Although we achieve good results in the Istethoscope dataset, facing its instance/patient variability and the signal's noise, their validity is arguable. Since the Istethoscope dataset has only 20 auscultations, and we perform 5-fold cross validation, this results in testing four instances in each iteration. As we have more features than instances, the over-fitting inherent to this particular model is even more present. For a more robust result in this dataset, more auscultations would be needed.

5.5 Discussion

In this chapter, we presented descriptive features based on the representations introduced in chapter 3. We picked the maximum of a segments, in a given representation, as a representative feature. This feature resembles the visual process in which a clinician annotates an heart sound signal. We started by comparing individual features —features that described only the segment itself. Then, we concatenated the neighbouring systole and diastole to a segment and used them to perform the Classification. We called this type of features, neighbourhood features. Finally, we varied the concatenation of the best individual and neighbourhood features to maximize the accuracy of the Classification procedure.

The best performing individual feature was the CWT, confirming the claim we made in chapter 3. The best performing neighbourhood feature was the Duration. This was expected as the neighbourhood duration included the adjacent diastolic duration, which is the primary feature used in a manual clinical annotation to distinguish S1 from S2. The combination of CWT and the S-T individual features achieved the best result for the Digiscope dataset while the combination of CWT, DWT and S-T neighbourhood features performed best in the Istethoscope dataset. Although we achieve good results in the Istethoscope dataset, their validity is arguable since we have more features than instances, over-fitting an already biased model to the data.

We end this thesis with the following chapter. It shows an overview of the presented heart sound segmentation stages and highlights our main findings in each of them.

Chapter 6

Conclusion

Throughout this thesis, we showed different Heart Sound Segmentation approaches divided by four main stages: pre-processing, representation, segmentation and classification. In the pre-processing chapter, we started by doing an exploratory analysis of the frequency content of both Digiscope and Istethoscope datasets. Then, we illustrated the effects of different filters on a heart sound and the effects of downsampling. We chose not to apply any type of filtering as we focused more heavily on the subsequent stages and to show that it is possible to perform a good heart sound signal segmentation without the use of this kind of preprocessing. We only downsampled the Istehoscope dataset as it was unfeasible to perform exhaustive tests in the representation, segmentation and classification stages.

In the Representation stage we presented several representations and compared their performance in terms of peak detection and classification capabilities. We found that the Shannon Energy Envelope was the best representation for detection and CWT for classification.

We proposed a Segmentation stage divided into two distinct phases: Peak and Boundary Detection. In the Peak Detection phase we introduced a novel method that is based on the Stationary Wavelet Transform inflection points. These points performed an initial segmentation followed by a Hierarchical Clustering procedure that automatically picked segments that contained S1 and S2 heart sounds. We varied the Wavelet order, scale and coefficients parameters to minimize the total error. The best performing combination in the Digiscope dataset was the approximation coefficients of the Daubechies wavelet of order 9 at scale 10. This parameter combination achieved a total error of 56732 while the winning approach of the Classifying Heart Sounds

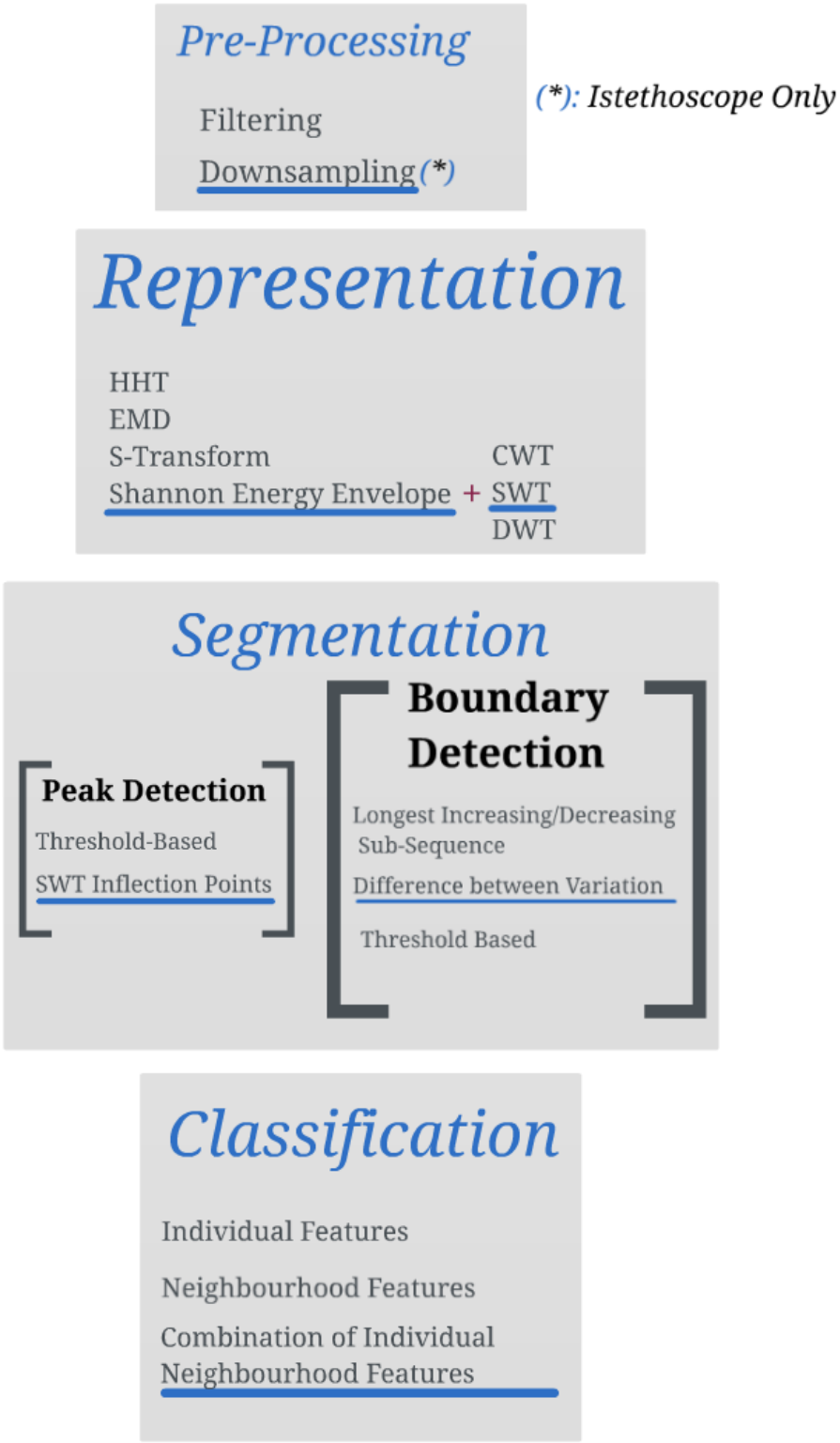


Figure 6.1: Overview of the presented work. The underlined text highlights the best performing methods.

Challenge achieved 72242. In the Istethoscope dataset as we noticed some very different patterns in its auscultation caused by the misuse of used iPhone, we resorted to the Shannon Entropy instead of the Shannon Energy Envelope. This change allowed us to detect some missing segments, achieving a total error of 706535, while the best performing approach in the challenge had 1243640. In the Boundary Detection method, we introduced two novel methods: the Longest Increasing/Decreasing Sub-Sequence and the Difference between Variations method. These two methods were compared with a baseline threshold based approach. Both methods out-performed the baseline approach, with the Difference between variation being the best of the two.

Finally in the classification stage, we presented 2 types of features: Individual and Neighbourhood Features. We found that the Neighbourhood features performed better than the individual ones. This was due to patterns available in adjacent segments. Both type of features were combined within each other to find the combination of features that maximized the Accuracy. We found that a combination of individual features performed best in the Digiscope dataset and the Neighbourhood features in the Istethoscope Dataset. We argued that the results in the Istethoscope dataset were not fully reliable as there were shortage of data in this dataset.

The main contribution of this thesis was featured in an article published in the conference "Computing in Cardiodology"[MCRA13]. An improvement could be made by finding a wavelet that detected peaks without the hierarchical clustering we used in this work. Without a doubt, the greatest improvement would be to find a specific wavelet that created the boundaries using the SWT's inflection points. In the future we hope to address these shortcomings, thus creating a better heart sound segmentation approach.

Appendix A

Acronyms

HSS - Heart Sound Segmentation
PCG - Phonocardiogram
AM - Amplitude Modulation
SWT - Stationary Wavelet Transform
DWT - Discrete Wavelet Transform
CWT - Continuous Wavelet Transform
S-T - S-Transform
HHT - Hilbert Huang Transform
EMD - Empirical Mode Decomposition
IMF - Intrinsic Mode Function
KNN - K-Nearest Neighbours
UPGMA - Unweighted Pair Group Method with Arithmetic Mean
FT - Fourier Transform
DTFT - Discrete-Time Fourier Transform
AR - Auto-Regressive model

References

[AT84] P J Arnott and M E Tavei. Spectral Analysis of Heart Sounds. 6:121–128, 1984.

[BGG97] C. Sidney Burrus, Ramesh A. Gopinath, and Haitao Guo. *Introduction to Wavelets and Wavelet Transforms: A Primer*. Prentice Hall, 1 edition, August 1997.

[BNCM11] P. Bentley, G. Nordehn, M. Coimbra, and S. Mannor. The PASCAL Classifying Heart Sounds Challenge 2011 (CHSC2011) Results. http://www.peterjbentley.com/heartchallenge/index.html, 2011.

[CHJH13] J Chen, H Hou, Chen Jie, and Hou Hailiang. Segmentation of heart sound using double-threshold. In *Proceedings - 2013 5th Conference on Measuring Technology and Mechatronics Automation, ICMTMA 2013*, pages 985–988, 2013.

[Coi10] M. Coimbra. The digiscope project - digitally enhanced stethoscope for clinical usage. http://digiscope.up.pt/, June 2010.

[CVMC13] Ana Castro, Tiago T V Vinhoza, Sandra S Mattos, and Miguel T Coimbra. Heart Sound Segmentation of Pediatric Auscultations Using Wavelet Analysis. pages 1–4, 2013.

[Dud76] Sahibsingh A. Dudani. The distance-weighted k-nearest-neighbor rule. *Systems, Man and Cybernetics, IEEE Transactions on*, SMC-6(4):325–327, 1976.

[EG83] Bradley Efron and Gail Gong. A leisurely look at the bootstrap, the jackknife, and cross-validation. *The American Statistician*, 37(1):36–48, 1983.

[EMD12] Heart sound analysis using EMD algorithm. In *Biomedical Engineering and Sciences (IECBES), 2012 IEEE EMBS Conference on*, pages 861–866, 2012.

[ETG12] Burhan Ergen, Yetkin Tatar, and Halil Ozcan Gulcur. Time-frequency analysis of phonocardiogram signals using wavelet transform: a comparative study. *Computer methods in biomechanics and biomedical engineering*, 15(4):371–81, January 2012.

[Got66] I.M. Gottlieb. *Understanding amplitude modulation*. Foulsham-Sams techn. books. H. W. Sams, 1966.

[HS02] Ibrahim R Hanna and Mark E Silverman. A history of cardiac auscultation and some of its contributors. *The American journal of cardiology*, 90(3):259–67, August 2002.

[HS05] N.E. Huang and S.S. Shen. *Hilbert-Huang Transform and Its Applications*. Interdisciplinary mathematical sciences. World Scientific, 2005.

[HSI97] Liang Huiying, L Sakari, and H Iiro. A heart sound segmentation algorithm using wavelet decomposition and reconstruction. In *Engineering in Medicine and Biology Society, 1997. Proceedings of the 19th Annual International Conference of the IEEE*, volume 4, pages 1630–1633 vol.4, 1997.

[HSL+98] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1971):903–995, March 1998.

[KCA+06] D Kumar, P Carvalho, M Antunes, P Gil, J Henriques, and L Eugenio. A New Algorithm for Detection of S1 and S2 Heart Sounds. In *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, volume 2, pages II–II, 2006.

[LLH97] H Liang, S Lukkarinen, and I Hartimo. Heart sound segmentation algorithm based on heart sound envelogram. In *Computers in Cardiology 1997*, volume 24, pages 105–108, 1997.

[LMA49] Aldo A. Luisada, F. Mendoza, and Mariano M. Alimurung. The duration of normal heart sounds. *British heart journal*, 11(1):41–7, January 1949.

[MAT11] MATLAB. *version 7.12.0 (R2011a)*. The MathWorks Inc., Natick, Massachusetts, 2011.

[MCRA13] N. Marques, M. Coimbra, Ana Paula Edward Rocha, and Rute Almeida. Exploring the stationary wavelet transform detail coefficients for the detection and identification of the s1 and s2 heart sounds. *Proc. Computing in Cardiology*, 2013.

[MDHB13] Ali Moukadem, Alain Dieterlen, Nicolas Hueber, and Christian Brandt. A robust heart sounds segmentation module based on S-transform. *Biomedical Signal Processing and Control*, 8(3):273–281, May 2013.

[Mit10] S.K. Mitra. *Digital Signal Processing: A Computer-based Approach*. McGraw-Hill Companies, 2010.

[Mur84] Fionn Murtagh. Complexities of hierarchic clustering algorithms: State of the art. *Computational Statistics Quarterly*, 1(2):101–113, 1984.

[SAH94] R.C. Schlant, R.W. Alexander, and J.W. Hurst. *Hurst's The heart, eighth edition, companion handbook*. McGraw-Hill, Health Professions Division, 1994.

[SM05] P. Stoica and R.L. Moses. *Spectral analysis of signals*. Pearson Prentice Hall, 2005.

[Sto91] R G Stockwell. Why use the S-Transform ? 00:1–31, 1991.

[Tav06] Morton E Tavel. Cardiac auscultation: a glorious past–and it does have a future! *Circulation*, 113(9):1255–9, March 2006.