Phonocardiogram Noise Detection in Realistic Environments

Master Thesis in Computer Science

submitted by

Nuno Miguel Tavares Costa e Silva

under the orientation of

Prof. Dr. Miguel Tavares Coimbra

Faculdade de Ciências da Universidade do Porto

Departamento de Ciência de Computadores

September 2012

Abstract

One of the main obstacles that stalls the widespread use of phonocardiograms (PCGs) in modern day medicine is the various noise components they invariably contain. Although many advances have been made towards automated heart sound segmentation and heart pathology detection and classification, an efficient method for noise handling would come as a major aid for further development in this field, especially when it comes to working with PCGs collected in realistic environments such as hospitals and clinics.

The use of a digital stethoscope to collect experimental samples of heart sounds in realistic conditions results in highly noise contaminated PCGs. Besides the fact that these environments are often crowded, the very method of data gathering (the use of a transducer for heart sound recording) contributes to the addition of background noise. Thus, only an adaptive algorithm that takes all these aspects into account will successfully increase signal-to-noise ratio.

After a process of investigating and testing commonly used PCG noise reduction and noise detection techniques, we have selected a noise detection algorithm based on periodicity measures to classify noisy and clean heart sound segments in PCGs collected in realistic environments. This thesis describes the process of selection, adaptation and testing of this algorithm. We will describe in detail the alterations performed in the previous implementation of this algorithm in order to enable processing of PCGs that contain high noise levels. The results achieved are satisfactory for the current stage of development and the conclusions taken clearly indicate a future line of work towards a robust noise detection algorithm that will aid signal processing of PCGs.

Keywords: Phonocardiography, Noise Detection, Periodicity, Signal Processing

Acknowledgements

I would like to thank my thesis advisor Prof. Miguel Coimbra for the helpful guidance and for giving me the opportunity to work in the challenging subject of phonocardiography, it has been an interesting and enriching experience.

I would also like to thank everyone within the DigiScope project, especially those working in signal processing. I am very grateful for your time and insight without which the work I present here would not have come as far as I would have hoped.

Finally I thank my family and my girlfriend for the ongoing support and understanding.

Contents

List of Figures

Abbreviations

HS	**Heart S**ound
PCG	**P**hono**C**ardio**G**ram
DTFT	**D**iscrete **T**ime **F**ourier **T**ransform
STFT	**S**hort **T**ime **F**ourier **T**ransform
FIR	**F**inite **I**mpulse **R**esponse
IIR	**I**nfinite **I**mpulse **R**esponse
ZCR	**Z**ero **C**rossing **R**ate
LTI	**L**inear **T**ime **I**nvariant
SVD	**S**ingular **V**alue **D**ecomposition
SVR	**S**ingular **V**alue **R**atio
PCA	**P**rincipal **C**omponents **A**nalysis
BPM	**B**eats **P**er **M**inute
SNR	**S**ignal to **N**oise **R**atio

Chapter 1

Introduction

This section presents the context in which the work developed in this thesis is inserted. We will address the topic of digital auscultation, which provides the input to our problem, the method by which the input has been collected, and give a brief description of the problem. Finally we will describe the motivation for this work, our main contributions and the objectives we hope to achieve.

1.1 Digital Auscultation

Cardiac auscultation is known to have been practiced during the Hippocratic period (460 to 370 BC) [1]. In that time, until the ninetieth century, auscultation was performed by pressing the hand or the hear against the patient's body, a process known as immediate auscultation. Only in 1816 did a french physician named Rene Laennec discover the stethoscope when faced with difficulties when examining a fat patient. Based on simple acoustics, he decided to roll a paper into a sort of cylinder which allowed him to better listen to heart sounds [1]. His discovery gave rise to a number of developments in cardiology and the stethoscope has since then evolved into a high precision instrument used commonly in every hospital.

With the advent of new technology in the twentieth century, which brought to light other sophisticated diagnostic modalities such as echocardiography and chest x-rays, phonocardiography, the diagnostic technique that creates and studies a graphic record

of heart sounds, became less important. The decline of auscultation teaching in medical schools and a consequent lack of confidence and accuracy in identifying heart sounds and murmurs also contributed to this fact [2]. However, with the evolution of computers and digital signal processing, phonocardiograms (PCGs) may reveal important information [3].

Traditional mechanical stethoscopes possess certain limitations in what concerns the study of the PCG: they cannot store and playback sounds, cannot offer visual display nor process the acoustic signal. In order to overcome the limitations of these mechanical tools electronic stethoscopes have been developed. Nowadays, with the advent of miniaturized and powerful technologies for computing, these limitations are rapidly receding [4] and electronic and digital stethoscopes are being used in areas such as telehealth [5], phonocardiography [6], among others.

1.2 DigiScope - Phonocardiogram Collection and Analysis

Based on the technology provided by digital stethoscopes, a tool for collecting, storing and processing acoustic auscultation signals has been created. The DigiScope collector [7] was developed with the immediate goal of creating a repository of annotated auscultation signals for biomedical signal processing and machine learning research. This repository has been the testing ground for the work developed during this thesis and will henceforth be referred to as DigiScope repository. Figure 1.1 shows the DigiScope collector with a Littmann Model 3200 digital stethoscope, used to record and transmit heart sounds.

FIGURE 1.1: The DigiScope Collector.

Using the capabilities of a digital stethoscope to record and transmit heart sounds will allow this software application to be used as a transmition tool for professionals to discuss

diagnoses together, as a teaching tool for medical students, or as a method for screening cardiac pathology by using signal processing and machine learning computational resources. The latter is the ultimate goal of this application and is better described in Figure 1.2.

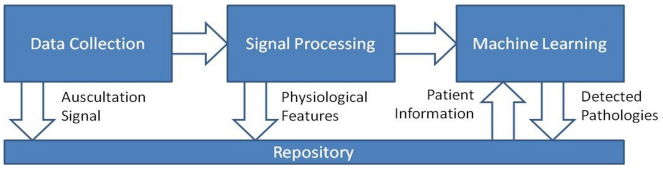

FIGURE 1.2: The DigiScope System Model. Adapted from [7].

The work described in this thesis relates directly to the signal processing phase of the DigiScope system. Currently, with the evolution of signal processing techniques, a variety of methods that rely on digital signal processing have been proposed to analyze, process and retrieve data from the PCG [8]. These methods focus mainly on heart sound denoising, segmentation and information retrieval for pathology detection and classification. However, they all rely upon the sound quality of the analyzed PCGs to achieve good results.

When PCGs are collected in controlled conditions created specifically for study purposes, background noise interference can be attenuated and the source of noise reduced to the physiological sounds produced by the human body other than heart sounds. Plus, in these cases there are no timing constraints and the person collecting the data can calmly proceed with no external factors of stress. Since DigiScope aspires to aid cardiac pathology screening in real hospital environments, the interference of external noise sources introduces a much bigger problem than in controlled environments. Our experience shows that some important noise sources which create a significant interference with PCG quality are:

- respiration sounds,
- patient sounds and movements,
- small movements of the stethoscope (friction noise),
- acoustic damping of the bones and tissue, and
- environment background noise.

Due to the significant frequency overlap and similarity to heart sounds, respiration sounds and friction noises are the biggest obstacle in collecting good quality PCGs. This obstacle constitutes the main motivation behind the work described in this thesis, which will be approached in the following section.

1.3 Noise Detection - Motivation and Objectives

The problem of dealing with external noise interference in PCGs collected in noisy environments has been approached from several angles. Noise removal has no doubt been one of the most explored methods of recovery of heart sounds from a noisy signal. Adaptive filtering ([9],[10]) and wavelet denoising [3] have been widely researched with successful results. However, a different approach has been chosen for the present work: to leave the frequency range of heart sounds intact and perform what we call noise detection. The goal, instead of separating and removing the non heart sound components of the noise contaminated PCG, is to analyze the PCG and extract the segments that can be classified as clean heart sounds and that are suitable for further signal analysis. The detection of these segments is performed based on sound quality criteria that will be explained later in Chapter 4.

This choice was motivated by the fact that noise removal alters the frequency spectrum of the signal and may introduce artifacts that damage the signal preventing any further analysis. Consequently, the signal would become unsuitable for any listening purposes due to its degradation. The method proposed by Falk et al. [11], which will be described latter, was studied and tested and the results further contributed to this conclusion.

An interesting approach towards noise detection was proposed in the Intellisensor toolbox [12], which consists in a framework implemented in Matlab that contains algorithms developed for noise detection, segmentation and murmur diagnosis of heart sounds. The noise detection algorithm implemented in this toolbox is based on periodicity measures and spectral analysis of the signal to determine which of the signal segments can be classified as clean heart sounds and therefore suitable for further analysis. This algorithm is further described in Chapter 3.

An attempt to use the noise detection algorithm's implementation to process heart sounds from the DigiScope repository was made. The results were not satisfactory due

to an important fact that motivated the work presented in Chapter 4: The algorithms developed for the Intellisensor toolbox were designed to process heart sounds collected in a controlled environment with low ambient noise. Furthermore, they were not entirely collect by health care professionals and the clinical auscultation procedure was not followed, all of which happened in DigiScope's heart sound acquisitions. Thus, an effort was made to adapt the noise detection algorithm to perform noise detection on DigiScope PCGs, which are highly contaminated with noise from different sources. The goal of this adaptation effort is simply to strengthen the DigiScope application with a robust noise detection algorithm that will filter and guarantee PCG signal quality for further signal processing and information retrieval.

1.4 Contributions

The main contributions described in this thesis are:

- Investigation and selection of an adequate methodology to assess the problem of noise contamination in PCGs.
- Adaptation of an already existing noise detection algorithm to process noise contaminated PCGs collected in clinical environments.
- Construction of a suitable data-set to assess the algorithm's performance.

1.5 Document Structure

This document is structured in the following way:

- Chapter 2 will provide background knowledge necessary to describe the subsequent chapters. The human heart will be described focusing on heart sound generation followed by a brief description of the cardiac auscultation procedure. A short introduction to the mathematical tools and definitions used will also be provided.
- Chapter 3 will provide an overview of the relevant publications studied to accomplish the proposed work.

- In Chapter 4 we will focus on describing the algorithm proposed by Kumar et al. [13], the adaptations performed for noise contaminated PCG processing and the results obtained.
- Chapter 5 will present a brief conclusion focusing on the results obtained and future work.

Chapter 2

Background

In this chapter we will provide a common basis for the concepts employed throughout this thesis. A short description of heart anatomy will be given, focusing on the way heart sound is generated, followed by an explanation on how cardiac auscultation is performed. A description of the mathematical concepts employed will follow.

2.1 The Human Heart

The goal of this section is to give a brief description of the human heart in order to understand how and why the heart sounds are generated. To achieve this we will delve superficially in its anatomy and the way the blood flow directly relates to the sounds generated. We will finish with a description of the cardiac auscultation method.

2.1.1 Anatomy and Sound Generation

The purpose of the heart is to pump blood through the blood vessels to every part of the human body renewing its oxygen content. It has four chambers separated from each other by the cardiac muscle: the right and left atriums and the right and left ventricles. As we can see in Figure 2.1 the right and left chambers of the heart have different functions. In the right side, the superior and inferior vena cava take de-oxygenated blood to the heart through the right atrium which is pumped through the tricuspid valve into the right ventricle to the lungs where carbon dioxide is exchanged for oxygen.

In the left side, the left atrium receives the oxygenated blood from the lungs through the left and right pulmonary veins. The blood is then pumped through the left ventricle through the mitral valve and is sent out to the body by the aorta.

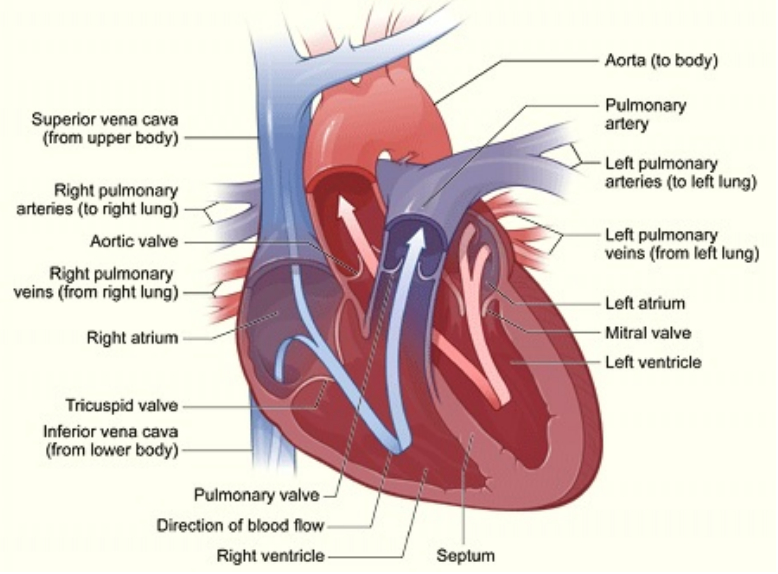

FIGURE 2.1: The interior of the human heart. Adapted from [14]

The heart valves open and close due to blood pressure differentials and create the sound a physician can hear with a stethoscope. The closure of the mitral and tricuspid valves produces S1, the heart sound that marks the start of the systole, the period when the ventricles contract and pump the blood from the heart out to the body. The second heart sound, S2, is produced by the closure of the pulmonary and aortic valves at the beginning of the diastole, the period when the ventricles are filled with blood pumped into them by the atria. These are the heart sounds produced by a normal adult person. Other heart sounds, such as murmurs, S3 and S4, often relate to a specific heart pathology. For example, the S3 gallop is frequently a sign of left ventricular failure and the S4 gallop can be encountered in patients with any condition causing decreased ventricular compliance, such as hypertension [15]. Heart murmurs can be systolic or diastolic depending on the period of the heart cycle they occur and may or may not be malignant. Innocent murmurs are commonly heard in school-aged children and result from turbulent blood flow generated by left ventricular ejection of blood [15]. Other murmurs relate directly to a specific heart failure: for example, tricuspic regurgitation is a blowing, pan-systolic murmur which is heard commonly when a right ventricular failure exists [15].

Given the previous facts we can easily conclude about the importance of the evaluation of heart sounds in any patient. As a first diagnostic tool, physicians rely on cardiac auscultation to determine the presence of any abnormal heart sounds and diagnose their causes. Next we will briefly describe this process.

2.1.2 The Auscultation Procedure

Nowadays listening to the sounds emanated from the heart has stabilized into a systematic process. Cardiac auscultation consists on an approach performed by physicians while listening to heart sounds using a stethoscope. It involves listening the heart on specific points, each of them near a cardiac valve, enabling the detection of murmurs associated with valvular abnormalities. In our study, the procedure starts with the first auscultation at the right upper sternal border (aortic area) and next at the left sternal border (pulmonic area). Next the procedure continues down the the right sternal border (tricuspid area) with the final point of auscultation being the apex (mitral area) [15]. Figure 2.2 demonstrates these points and the order of the auscultation method. The procedure can also be performed in reverse order as long as the sequence is followed.

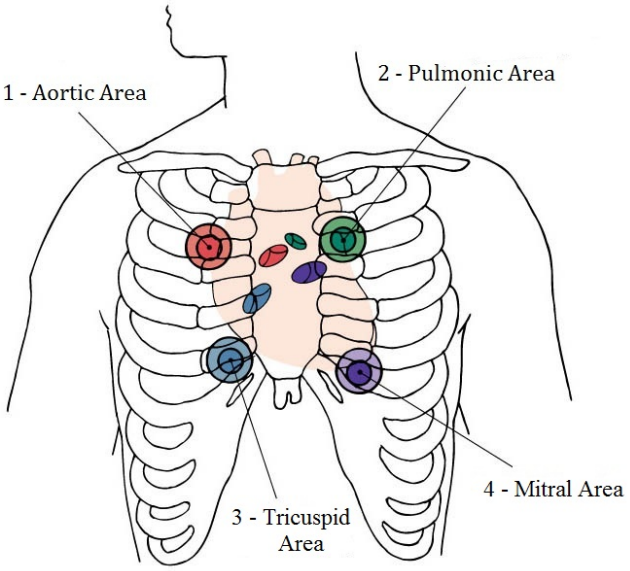

FIGURE 2.2: Points of auscultation and their order. Adapted from [16]

2.2 Mathematical Tools and Definitions

This section is dedicated to describing the underlying theory behind some of the mathematical concepts commonly applied in signal processing algorithms. Although previous knowledge of signal processing theory is a pre-requisite to fully understand these definitions and their application, we will try to explain them in a thorough way.

These concepts will be explained in the context of one-dimensional discrete signals, which is the case of a one channel (mono) audio signals such as the PCG. The independent variable here is time and any value of the signal at a specific discrete time instant is called amplitude.

2.2.1 The Fourier Transform

Inevitably we start by introducing the Fourier transform, in particular the discrete-time Fourier transform (DTFT) as the frequency-domain representation of a discrete-time sequence. The signal, originally represented in the time-domain, can be transformed into its frequency-domain equivalent, which expresses the sequence as a weighted combination of the complex exponential sequence $\{e^{j\omega n}\}$ where ω is the real normalized frequency variable [17]. The DTFT $X(e^{j\omega})$ of a discrete-time sequence $x[n]$ is defined by:

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

The result of this operation enables us to examine and if necessary process the frequency content of a certain discrete-time signal. How can this be achieved? As we can see in equation (2.1) the signal x[n] is multiplied by the complex exponential term $\{e^{-j\omega n}\}$ at frequency ω and summed over all the time instances in which the signal is defined. Euler's formula definition, given by

$$e^{ix} = \cos x + i \sin x \tag{2.2}$$

tells us that by multiplying the original signal by a complex expression with sines and cosines of frequency ω and summing this product over all time instances of the signal

will enable us to quantify the presence of frequency ω in the given input signal. This operation, performed for all angular frequencies $\omega \in [0, \pi]$, will yield the frequency content of input signal $x[n]$. To better understand this lets observe a pertinent example.

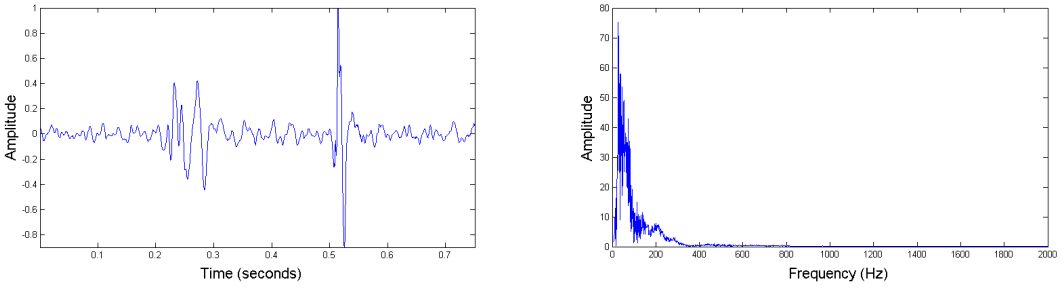

FIGURE 2.3: A discrete-time audio signal containing one heart cycle and its frequency spectrum.

Figure 2.3 shows, on the left, a discrete-time audio signal containing one normal heart cycle sampled at 4000 sampling-rate. On the right one can observe the normalized frequency content of this signal, usually referred to as the signal's frequency spectrum, after the DTFT operation. This way it is possible to observe a stronger presence of lower frequencies in this signal which derives from the fact that heart sound frequencies are located in lower bands. A number of operations can then be performed with the frequency domain representation of a signal. It is then possible to return to the signal's time-domain version using the inverse discrete-time Fourier transform:

$$x[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\omega}) e^{j\omega n} \, \mathrm{d}\omega \tag{2.3}$$

The inverse discrete-time Fourier transform can be interpreted as a linear combination of infinitesimally small complex exponential signals of the form $\frac{1}{2\pi} e^{j\omega n} \mathrm{d}\omega$, weighted by the complex constant $X(e^{j\omega})$ over the normalized angular frequency range from $-\pi$ to π [17].

2.2.2 The Short-Time Fourier Transform

Although the DTFT remains a very powerful and useful tool in analyzing the frequency spectrum of discrete signals, it possesses a strong limitation related to the analysis of non-stationary signals, that is, signals whose spectral content varies over time such as

the PCG [6]. In order to analyze the frequency components of non-stationary signals over specific time instances an improvement to the DTFT was developed, the short-time Fourier transform (STFT), which is defined by

$$X_{STFT}(e^{j\omega}, n) = \sum_{m=-\infty}^{\infty} x[n-m]w[m]e^{-j\omega m} \tag{2.4}$$

where $x[n]$ is a discrete-time signal and $w[n]$ is a specific window sequence. To circumvent the time-variant characteristics of a signal, the STFT segments it into a set of sub-sequences of smaller length and computes the DTFT separately for each one. It is of great importance to choose an adequate window length to divide the signal in order to achieve a reasonable time-frequency resolution; smaller windows will provide good time resolution and poor frequency definition and vice-versa. Therefore the window length choice should be based upon the signal characteristics and the type of information to be retrieved. Figure 2.4 demonstrates this fact with two different spectrograms which are the result of the STFT computation with six and fifty-one length windows respectively.

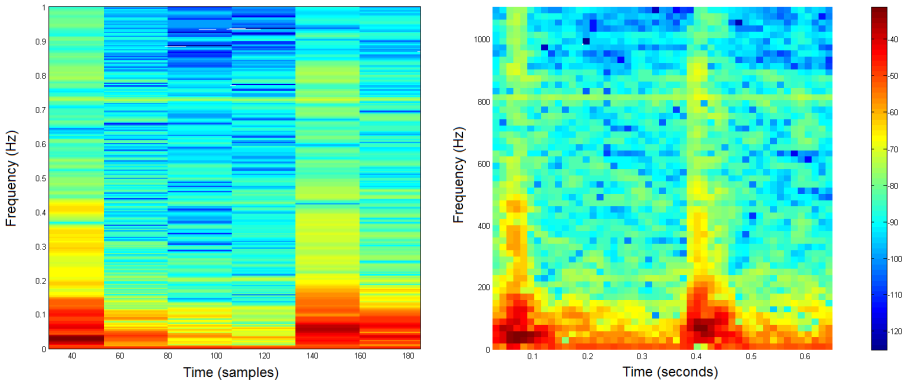

FIGURE 2.4: Spectrograms of the same heart cycle with different window lengths. Left: six windows, right: Fifty-one windows.

2.2.3 The Discrete Wavelet Transform

The wavelet transform was developed to overcome the time-frequency resolution issues of the STFT and it consists in a technique to decompose discrete time signals. Its continuous version operates in a similar way to the STFT; the signal is multiplied by a function, the wavelet, and the transform is computed for different sub-sequences of the time-domain signal. The main difference from the STFT is that, for every spectral

component, the width of the window changes, which allows for a superior time-frequency resolution. It is defined by the following equation:

$$CWT_x^{\psi}(\tau, s) = \Psi_x^{\psi} = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{\infty} x(t)\psi^* \left(\frac{t-\tau}{s} \right) \mathrm{d}t \tag{2.5}$$

As can be seen above the equation is a function of two variables, τ and s, which represent translation and scale parameters respectively. The $\psi(t)$ is the transforming function called the mother wavelet. The translation parameter corresponds to the time information in the transform domain in the same sense as in the STFT and the scale parameter corresponds to the parameter which will contract and dilate the wavelet function in order to correlate it with the signal. The transform is computed by changing the scale in each window, shifting the window in time (translation), multiplying by the signal and integrating over all times.

In its discrete case this transform requires much less computational effort in its implementations and presents sufficient information both for analysis and synthesis of the original signal. High pass and low pass filters, which will be described in the next section, are used to analyse the signal at different scales. The signal's resolution is changed by filtering operations and the scale by sub sampling operations. It is computed by repeatedly filtering the low and high bands of the signal and sub sampling these by a factor of 2, which yields the approximation and detail information respectively.

2.2.4 Digital Filters

Filtering is one of the most widely used complex signal processing operations [17]. Digital filters are commonly used for two general purposes [18] : separation of signals that have been combined or contaminated with noise or interference and restoration of signals that have been distorted in some way. They can be classified into two different sets: finite impulse response (FIR) and infinite impulse response (IIR) filters. FIR filters can be implemented by convolving a weighted sum of the input signal's samples with its impulse response, the output of a system when the input is an impulse. IIR filters, also named recursive filters, extend the FIR filters definition. Besides using input samples, they use previously calculated output samples to determine the filter's output value, and are

defined by a set of recursion coefficients. The frequency response of these filters can be obtained by calculating the discrete Fourier transform of the impulse response.

Filter performance can be characterized by properties which can be observed in time and frequency domain. Important time domain properties are the step response, that defines the accuracy with which events are distinguished in a signal, the overshoot, a basic distortion of the signal's time information that should be eliminated, and linear phase occurrence, which defines whether the filter alters or not the signal's phase content.

Frequency domain properties define the alterations a filter can perform on a signal's frequency spectrum. First we define the four basic frequency responses shown in Figure 2.5:

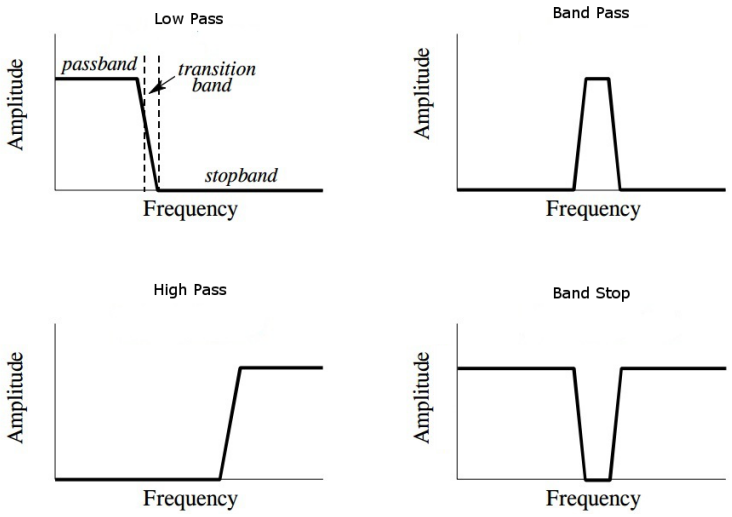

FIGURE 2.5: Common filter responses. Adapted from [18].

As can be seen in the top left graph the pass band and stop band refer to the frequencies passed and blocked respectively, with the transition band located between them. Important properties in this context are the roll-off, which is the width of the transition band, the existence of pass-band ripple, that should never occur because it alters the frequencies of the pass-band and stop-band attenuation, which refers to the amount of attenuation the filter performs on the stop-band.

It is important to mention that one cannot generalize whether a given property quantification is "good" or "bad" [18] because they are directly connected to the filter's characteristics and purposes, and these may require different filter implementations.

Filter applications can be classified by their time domain applications such as smoothing or waveform shapping, their frequency domain characteristics that alter a signal's frequency spectrum according to a specific goal, and can be custom made for a particular application in which a more elaborate design can be employed.

Figure 2.6 shows two typically used low pass filters, namely the Chebyshev type I and Butterworth filters, so that we may compare some of their main characteristics.

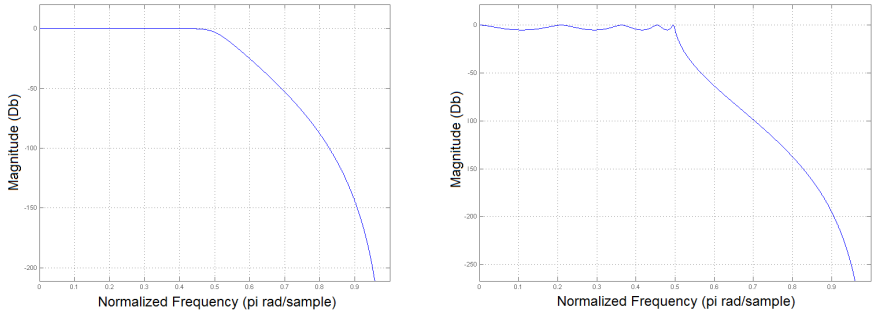

FIGURE 2.6: Butterworth (left) and Chebyshev type 1 (right) filter magnitude responses (Db).

Chebyshev type I filters, developed by the Russian mathematician Pafnuti Chebyshev, are designed to achieve the best possible balance between a faster roll-off and allowing ripple in the frequency response. A Butterworth filter is a specific case of the above mentioned filter: it is the Chebishev type I filter that has 0% ripple amplitude. On the left we have a 9 order low pass Butterworth filter with normalized cutoff frequency of 0.5. As we can see there are no ripples in the pass band but the roll-off is quite slower than in the next case. On the right we have a plot of a Chebishev type I low pass filter with the same cutoff frequency. In this case, by having allowed 5% ripple amplitude presence in the pass band, we achieve a faster roll off than on the previous case.

2.2.5 Zero Crossing Rate

As the name states, zero crossing rate (ZCR) measures the number of times the sign of a signal changes in a specified window or signal length [19]. It can be synthesized by the following equation:

$$ZCR_N = \frac{1}{N-1} \sum_{t=1}^{N-1} 0.5|sgn\{x[t]\} - sgn\{x[t-1]\}| \tag{2.6}$$

where

$$sgn\{x[t]\} = \begin{cases} 1, & x[t] \geq 0 \\ -1, & x[t] < 0 \end{cases}$$

and N corresponds to the window or signal length. ZCR can be used in a number of applications such as spectral analysis and noise estimation [20], extracting features from percussive sounds [21], speech recognition [22], among others. However its use for detailed spectral analysis can become problematic in signals that contain rich spectral content [19].

2.2.6 Cosine Similarity

Cosine similarity is a rather simple similarity test that has produced reliable results in various areas [23]. It relies on the inner product between two vectors and measures similarity based on the radial distance between these two vectors [13]. Equation 2.7 describes this operation:

$$Cos(\theta) = \frac{< x(t), y(t) >}{|x(t)||y(t)|} \tag{2.7}$$

where x and y are equal length vectors, $<,>$ is the inner product operator and $|\_|$ represents the Euclidian norm of a vector.

2.2.7 Convolution and Correlation

Convolution is a very important operation in signal processing that involves four basic operations: time-reversal, multiplication, addition and delay [17]. The convolution sum of two discrete sequences, $x[n]$ and $h[n]$ is given by

$$y[n] = \sum_{l=-\infty}^{\infty} x[l]h[n-l] \tag{2.8}$$

and is usually denoted by $y[n] = x[n] \otimes h[n]$. The output $y[n]$ can be seen as a function of the area overlap between the two input functions. In the case of linear time-invariant

(LTI) discrete time systems, such as FIR filters, that can be completely characterized by their impulse response, we can compute the system's output by convolving signal $x[n]$ with the system's impulse response $h[n]$.

Another important operation in signal processing is correlation. It can be seen as a measure of similarity between signals obtained by a sliding dot product between two functions. When the two signals are distinct it is named cross correlation and is expressed by the following equation:

$$r_{xy}[l] = \sum_{n=-\infty}^{\infty} x[n]y[n-l] \qquad l = 0, \pm 1, \pm 2, ... \tag{2.9}$$

where the l parameter indicates the lag between the function pair. In the case where $x[n]$ and $y[n]$ are the same signal this operation is called auto correlation and is typically used for finding repeating patterns within a signal.

2.2.8 Singular Value Decomposition

Last but not least we introduce a linear algebra operation named singular value decomposition (SVD). It is usually presented by the following equation:

$$A = USV^T \tag{2.10}$$

where U and V are orthonormal matrices: $U^TU = I$ and $V^TV = I$. S is a diagonal matrix that contains the square roots of eigenvalues from U and V in descending order. SVD can be informally interpreted as a set of three matrix transformations that can be easily explained for the case of A being a $3x3$ matrix and a data set D contained in a $3XN$ matrix by their geometric meaning:

1. Alignment: By multiplying V^T by matrix D, whose columns contain data points in three dimensions, we rotate the data in three dimensional space: the points shift from the standard basis to the V-basis

2. Dilation: Secondly we perform the matrix product SV^TD that stretches or contracts the points of D by the magnitude of the diagonal elements of S, the singular values, and zeroes the points of D that correspond to the zeros on the diagonal of S. This operation typically reduces the dimensions by which the data points in D are represented in the V-basis.

3. Hanging: The final step is to multiply U by SV^TD. This product will shift the data points from U-basis, which is the output basis of the last operation, back into the standard basis.

However in data sets of high dimension, which is the case of the work described in this thesis, this geometric approach is not always so evident. Therefore we shall adress SVD in a statistical context, in which it is typically named principal components analysis (PCA). This method is useful in identifying patterns in data, highlighting their similarities and differences and compressing the data by eliminating the number of dimensions with minimal loss of information.

This can be accomplished on a data set D by using the following method: first we subtract the mean for each data dimension of D and calculate its covariance matrix and then we calculate the eigenvalues and eigenvectors of the covariance matrix. By observing these results we can conclude that for the highest eigenvalue (referred to as principal component), its corresponding eigenvector will represent the data with the lowest variance possible in a new basis defined by all the eigenvectors. And also that the highest the eigenvalue is, the best its eigenvector will represent the data in the new basis.

These are the most important conclusions one can take from the output of PCA. They will be further developed latter in this document.

Chapter 3

Noise Handling in Phonocardiograms

As stated before in Chapter 1, two major alternatives arise as solutions for noise handling in PCGs: noise detection and noise removal. Each of them possesses its usefulness when dealing with particular types of noise interference.

Next we will introduce the work previously done in the areas of noise removal, noise detection and periodicity assessment, that directly relates to the work developed in this thesis. We will start by describing the noise removal approach studied and the motivation for abandoning this approach. Next we will focus on the noise detection methods explored and finally on the periodicity assessment method used will be described.

3.1 Noise Removal

3.1.1 Modulation Filtering for Heart and Lung Sound Separation from Breath Sound Recordings

Noise removal in phonocardiograms has been a subject tackled in various approaches. Different techniques include adaptive filtering [10], wavelet filtering [3], among others. Falk et al. [11] proposed a method based on an alternate spectro-temporal representation to separate heart from lung sounds. Due to the frequency overlap in these two sounds this method aims to explore the different stationary properties of the two sounds.

The method consists in decomposing the signal in the temporal trajectories of each frequency component. The resulting representation named modulation spectrum contains information about the rate of change of the signal's spectral components. The assumption investigated assumes that the spectral components of heart sound change at a different rate from the spectral components of lung sounds [11]. Heart sounds, due to their "quasi-periodic" characteristics, present modulation spectral content that falls between 2-20 Hz and lung sounds, having an important "stationary" property , are situated at low frequencies (< 2 Hz).

The implementation studied was the one developed for the Intellisensor toolbox [12]. The procedure consists of:

- Extracting the signal's temporal trajectories by taking an N-point discrete Fourier transform (DFT) of each frequency bin;
- Performing modulation filtering in each frequency bin using two FIR filters: a bandpass filter with cutoff frequencies of 2-20 Hz extracts the heart sound signal and a complementary bandstop filter extracts the lung sound signal.
- Correcting the negative power spectral values using a half-wave rectifier that clips negative values to zero and optionally reducing unwanted signal artifacts by filtering the cubic root compressed magnitude trajectories instead of the magnitude trajectories.
- After calculating the inverse discrete Fourier transform (IDFT) of each filtered signal, both signals are windowed and reconstructed using the overlap-and-add method.

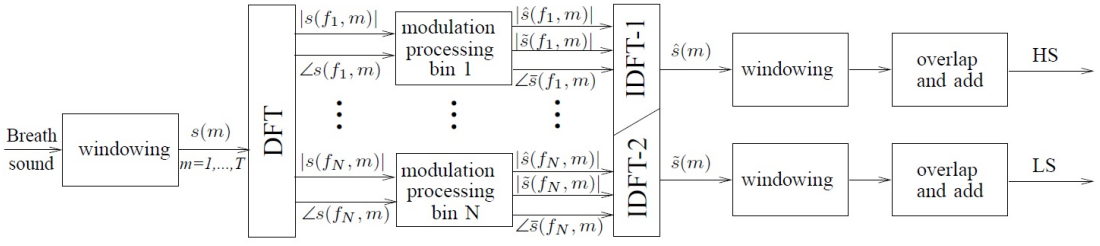

FIGURE 3.1: Block diagram of the modulation filtering method, in [11].

The block diagram of the algorithm is shown in Figure 3.1 where $s(m)$ is the analyzed input signal, $|s(f_n, m)|$ and $\angle s(f_n, m)$ are the magnitude and phase components of frequency bin n from the STFT, $\hat{s}(f, m)$ and $\tilde{s}(f, m)$ are the modulation filtered bandpass and bandstop signals of frequency bin n and $\hat{s}(m)$ and $\tilde{s}(m)$ are the inverse discrete Fourier transform (IDFT) reconstructed signals.

The above mentioned implementation performance was tested with heart sound signals from the DigiScope repository. Results were not satisfactory mainly due to the amount of noise that these signals contain. This caused the heart sound signal to contain a substantial quantity of sound artifacts that prevented the signal from being used for any other diagnostic purposes that involved listening.

Although not a positive result, this analysis agrees with that taken in [24] regarding the relation between the quality of the sound extracted and the amount of noise it contains.

Even though the study of one method is hardly conclusive for excluding a whole approach to a problem, we decided to tackle the problem of noise in phonocardiograms by exploring noise detection methods. This decision was motivated by the results obtained and by the notion that the amount of noise interference that the signals in the DigiScope database contain would constitute a major obstacle in removing any noise that overlaps the frequencies of the desired heart sound without compromising its quality. Furthermore, when a professional physician performs cardiac auscultation as described in section 2.1.2, the stethoscope rests in each auscultation point for a reasonably long time interval. This gives us enough margin to discard the noisy parts and extract the clean heart sound segments for further processing.

3.2 Noise Detection

3.2.1 Modulation Filtering for Noise Detection in Heart Sound Signals

Inspired by the procedure for heart and lung sound separation used by Falk et al. [11], Ramos et al. developed an adaptation of this algorithm with the purpose of noise detection in heart sound signals [25].

Based on the assumption that the modulation spectral content of the heart sound is located in the range 2-20 Hz, the modulation filtering module in this algorithm uses a

lowpass filter with a cutoff frequency of 1 Hz to extract only the non-cardiac component of the signal. Feature extraction is then performed over the non-cardiac component of each filtered frequency bin.

The algorithm processes three second windows consecutively. Each window is divided in two and three segments and the power ratios between each of these windows and the full signal are calculated according to equation 3.1.

Let $S(f,k)$ be the signal with length T and $S_w(f,k)$ the window segment under analysis. The power ratio is given by

$$P_w = \frac{\frac{\sum_{k\in w}|S_w(f,k)|^2}{L}}{\frac{\sum_{k\in T}|S(f,k)|^2}{T}} \tag{3.1}$$

where w is a window of size L.

The power ratios are the features that enable verification of the signal's stationary properties: if a transient noise contaminates the signal it reveals high intensity peaks that will disrupt the stability of the signal [25]. The power ratios that exhibit the maximum unbalance in the temporal power distribution are then fed to a classifier structure that returns a binary output.

This approach was informally tested with signals from the DigiScope database. Although the results obtained were reasonable, this method showed little flexibility regarding further adaptation to detect higher noise levels in the analyzed signals. However, due to the good sensitivity the algorithm demonstrated to transient noise, a simplified version was adapted and included in the algorithm described latter in Chapter 4.

3.2.2 Noise detection during heart sound recording using periodicity signatures

Kumar et al. [13] proposed a method for noise detection in PCGs based on the quasi-periodic properties of these signals. Having concluded that the quasi-periodicity behavior of heart sounds manifests itself in the time and time-frequency domains [13], the method proposed is divided in two phases: phase I searches the signal for an uncontaminated complete heart cycle to be used as a reference sound, by inspecting several

similarity measures based on the periodic characteristics of the signal. In phase II this reference sound is compared to the rest of the signal, using similarity criteria, to detect noisy and clean segments.

The signal is segmented into four second windows. For each window, to search for a reference sound, the algorithm uses the following features:

- Zero Crossing Rate (ZCR): Segments of 100 ms that exhibit a ZCR higher than 0.45 are considered noise.
- Periodicity in the time domain: The Hilbert transform is used to extract the heart sound envelope and its auto-correlation function is computed. Heart rate estimation is accomplished using singular value decomposition and the cosine similarity test is used to verify the degree of similarity between each heart cycle.
- Periodicity in the time-frequency domain: The spectral energy in each frequency band is verified for periodic validation by using the autocorrelation function, verifying peak alignment and singular value ratios (refer to section 3.3.1).

After the reference heart sound has been found, the algorithm relies on a set of matching procedures to perform non-cardiac sound detection. The spectral energy of the reference sound is compared to that of the segment being tested. If the correlation coefficient between the two signals is superior to a 0.98 threshold a temporal energy test is performed in order to capture short duration sound spikes. The diagram in Figure 3.2 synthesizes the two phases of the described method.

Again, the implementation studied has been that developed for the Intellisensor toolbox [12], and it was chosen for further refinement and adaptation to the DigiScope repository signals. This choice was motivated by the algorithm base assumptions: the quasi-periodicity of heart sounds and the different and complementary analysis possible in time and time-frequency domains. Although a preliminary analysis was sufficient to verify that the algorithm was not designed to process highly contaminated PCGs, its structure was flexible and a number of adaptations were instantly suggested to improve these not so satisfactory first results.

The algorithm details and the adaptations performed will be described in the following chapter along with a description of the results obtained.

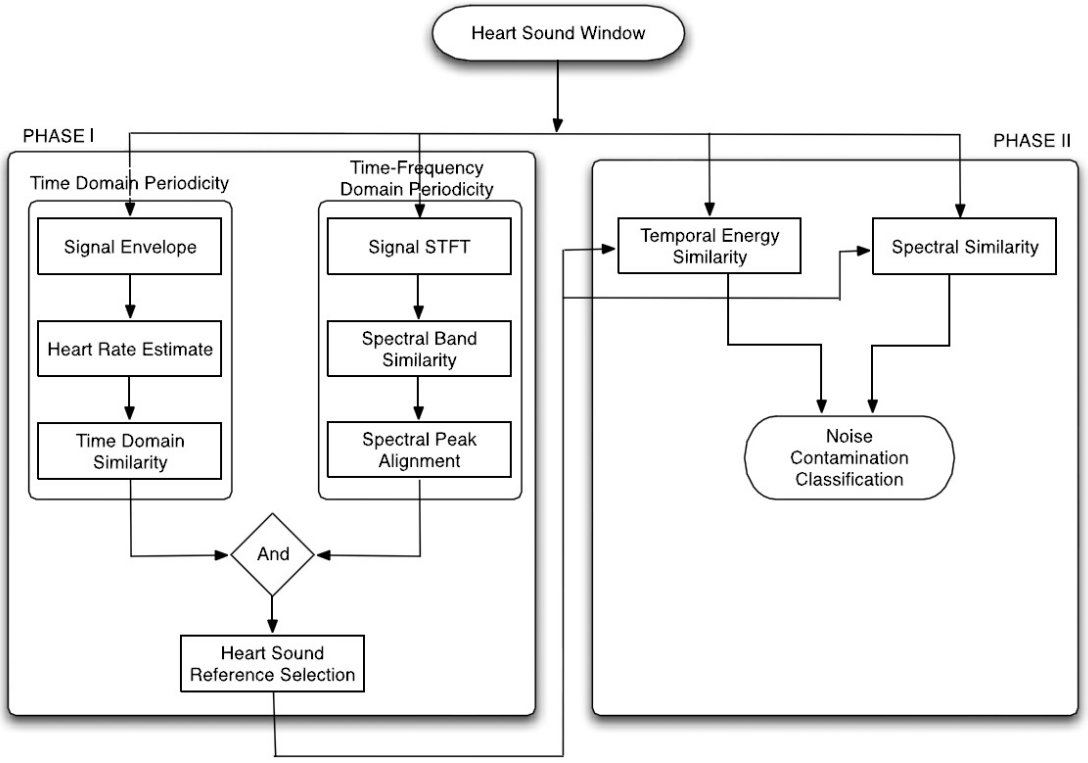

FIGURE 3.2: Flow chart of the algorithm proposed by Kumar et al., in [13].

3.3 Periodicity Assessment

3.3.1 On Multiple Pattern Extraction Using Singular Value Decomposition

The use of singular value decomposition as a method for characterizing the periodic components of a signal has been a widely explored technique in the area of biometric signal processing ([26],[27]). In [28] Kanjilal and Palit synthesize this method's application for periodicity detection and separation of the component signals.

Section 2.2.8 defines the SVD operation and introduces an interpretation for the resulting singular vectors and values. Extraction of periodic components can thus be performed by partitioning the discrete signal $y(t)$ into estimated periods and placing each period as a row of matrix $Y \in \mathbb{R}^{mxl}$

$$Y = \begin{bmatrix} y(1) & y(2) & \cdots & y(l) \\ y(l+1) & y(l+2) & \cdots & y(2l) \\ \vdots & \vdots & \ddots & \vdots \\ y[(m-1)l+1] & y[(m-1)l+2] & \cdots & y(ml) \end{bmatrix} \tag{3.2}$$

where m is the number of estimated periods and l is the length of each period.

Performing SVD on matrix Y will return a set of non zero singular values sorted in descending order. Let σ_1 and σ_2 refer to the first and second highest singular values. Calculating σ_1/σ_2, called the singular value ratio (SVR), for each estimated period value of m will allow us to observe the SVR Spectrum of the signal. As stated in [28] we can then conclude that high SVR values relate to the existence of strong periodicity in the signal for the estimated period length l.

However, this method is not as robust as one could hope for. It is prone to erroneous results in particular situations such as the occurrence of multiple peaks, generated by higher multiples of the period length, or the occurrence of high values for short period lengths. Therefore an additional interpretation of the SVR Spectrum should be performed to ensure the correctness of the conclusions drawn.

This methodology for periodicity assessment has been thoroughly studied for the purpose of heart rate estimation and heart cycle selection. The results obtained will be described in the following chapter.

Chapter 4

Periodicity Based Noise Detection

In this section we will focus on describing the various phases of the periodicity based noise detection algorithm that was introduced in Chapter 3. We will divide the explanation in four parts that refer to four different stages of the algorithm: preprocessing, reference window selection, reference heart sound selection and template matching. During this description we will give special importance to the alterations performed in order to adapt the algorithm to the substantial noise presence in PCG data from the DigiScope repository. Finally we will describe the results obtained from testing the algorithm with a small data-set contructed for that purpose, and the methodology used to collect them.

4.1 Algorithm Details and Adaptation

So that the contributions described in this chapter can be better explained and distinguished from the original algorithm implementation, we show in Figure 4.1 a similar flowchart to the one shown in Figure 3.2.

Each rectangle corresponds to a section of the algorithm and has one of three colors: Blue for sections that remain unaltered from the original implementation, red for sections that were adapted and green for sections that did not exist in the original implementation and were added for specific purposes. This part of the document will describe all of these sections focusing on the adaptations performed in each of them.

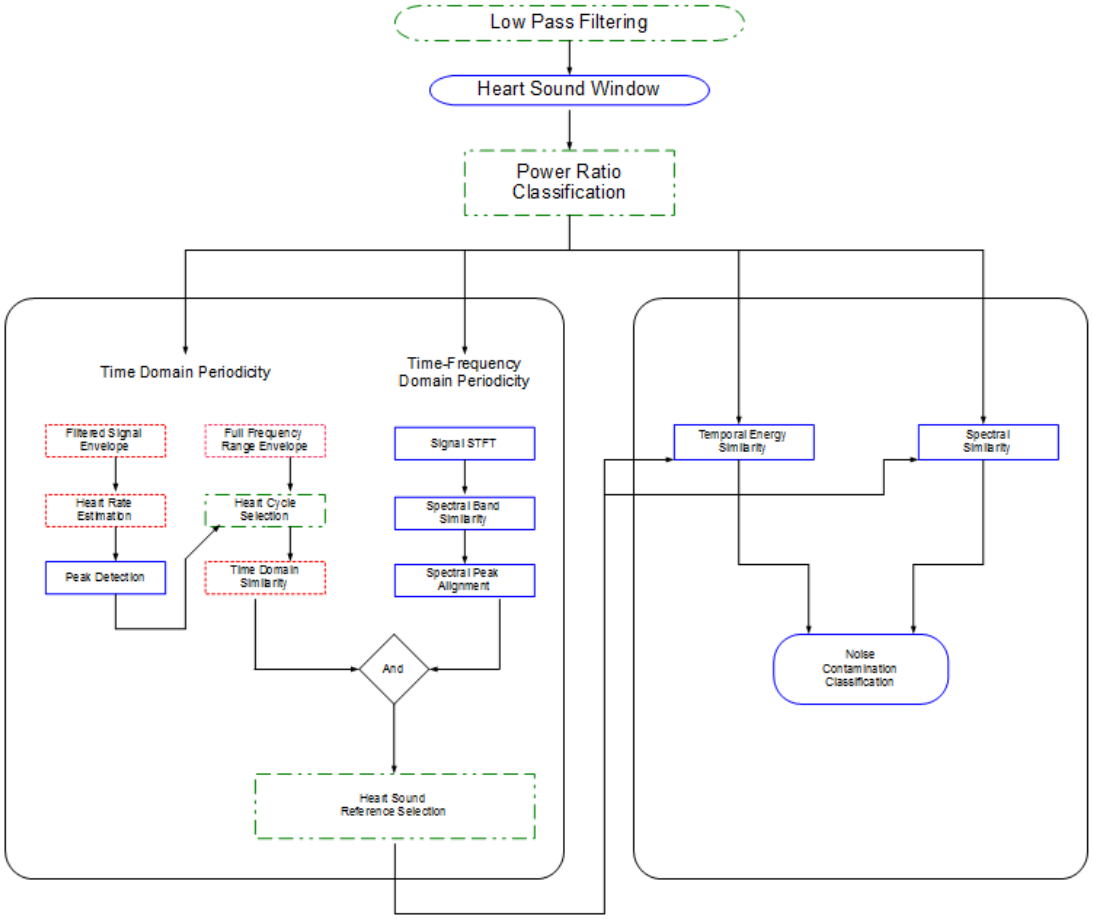

FIGURE 4.1: Flow chart of the selected noise detection algorithm with the adaptations described in this thesis. Blue rectangles correspond to unaltered sections, red rectangles to adapted sections and green rectangles to new sections.

4.1.1 Input

Before describing the methodology behind the algorithm it is important to describe what the input consists off. This algorithm was first designed for real time processing of PCGs recorded in controlled conditions. Since the DigiScope repository consists off PCGs collected by health care professionals in realistic environments, they were collected using the auscultation procedure described in section 2.1.2 for the duration of 60 seconds. Therefore the PCGs were collected in four auscultation points that were previously separated by an automatic process. The inputs for our algorithm will be the four segments of the PCG corresponding to four different auscultation points, which lengths may vary between 5 and 25 seconds, processed independently from each other.

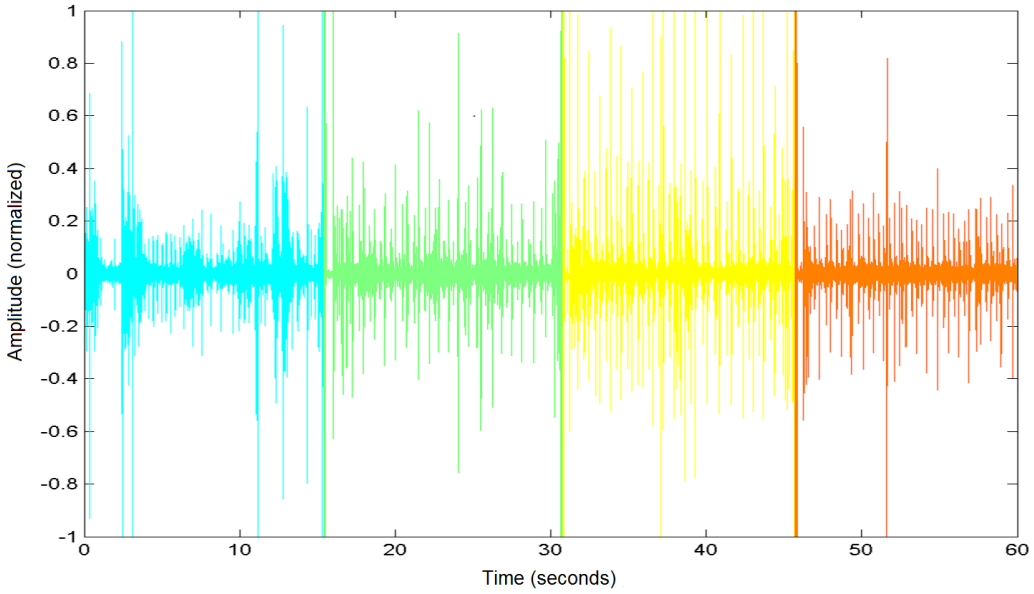

FIGURE 4.2: Example PCG divided in four segments (auscultation points): Aortic area (blue), Pulmonic area (green), Tricuspid area (yellow) and Mitral area (brown).

4.1.2 Preprocessing

The first preprocessing step added to the algorithm was a 57 order Butterworth low pass filter with a cutoff frequency of 900 Hz. The cutoff frequency value was chosen based on the fact that PCG signal energy ranges from 0 to 1000 Hz [29] and that the main part of its energy is contained between 50 and 300 Hz [6]. Thus, 900 Hz is a reasonable compromise due to the fact that, above this cutoff frequency, most of the signal's energy relates to ambient noise such a speech and background ambiance which make it very difficult to identify the higher frequencies of heart sounds. A Butterworth filter was chosen because a fast roll off was not necessary (a smooth one is preferable for listening purposes) and this way we exclude the presence of pass band ripple in the filtered signal (see section 2.2.4 for further details). As can be seen in Figure 4.3 the PCG signal contains a small amount of energy in the frequency range above the selected cutoff frequency of 900 HZ. Although no obvious conclusion can be taken by observing the filtered signals, we can clearly see that the HS energy remains present after the filtering process.

Additionally the signal is downsampled to a new sampling rate of 2205 Hz and prepared for phase two of the algorithm.

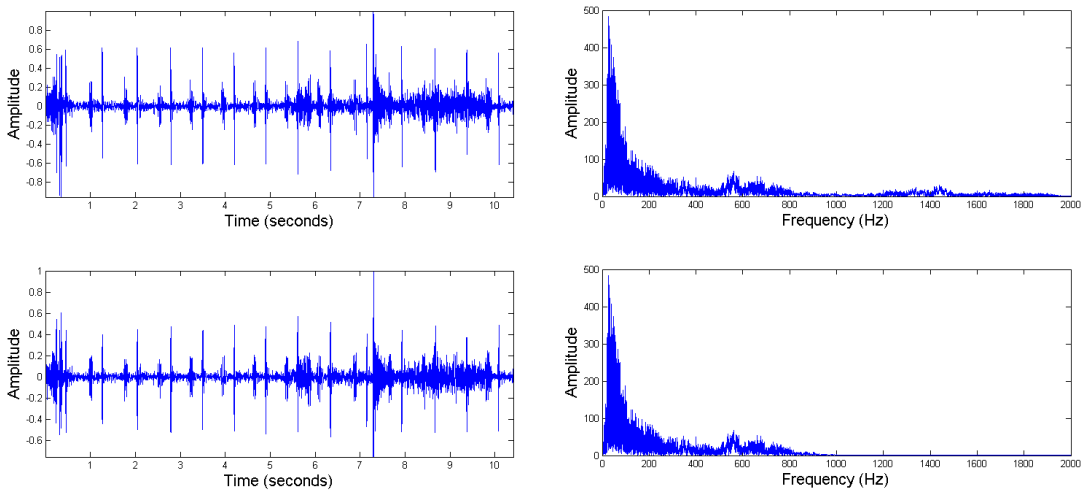

FIGURE 4.3: Butterworth lowpass filter input and output. Top right and left: input signal in time and frequency domains. Bottom right and left: output signal in time and frequency domains.

4.1.3 Reference Window Selection

Reference window selection is the first stage in finding a clean heart sound that can be used as a reference to extract all heart sounds with similar properties. This stage consists in sliding a 2,5 second window, with a step of 200 ms, and verifying that they contain a sufficiently low amount of noise that will not hinder the quality of the heart sounds contained in it.

The original algorithm would segment the signal in 4 second windows, slide the window in steps of one second and use a ZCR test to assess the presence of noise in each segment. The first window that encounters a ZCR value below a predefined threshold would be chosen as a reference window and further processing would be performed. In this version of the algorithm there are no performance constraints related to real time processing issues, since its main purpose is to process data offline. Furthermore there is a special concern in selecting the segment with better sound quality properties due to the short duration of the signals available. Therefore we chose to collect all suitable reference windows with the duration of 2,5 seconds (to encompass at least two heart cycles) and chose the best one based on criteria that will be explained latter on in this chapter. Figure 4.4 shows the segmentation of a 2,5 seconds window in a signal with 10,4 seconds length.

Noise presence verification at this stage underwent some changes. Abandoning the ZCR test was unavoidable due to the high presence of noise which hindered the test's results.

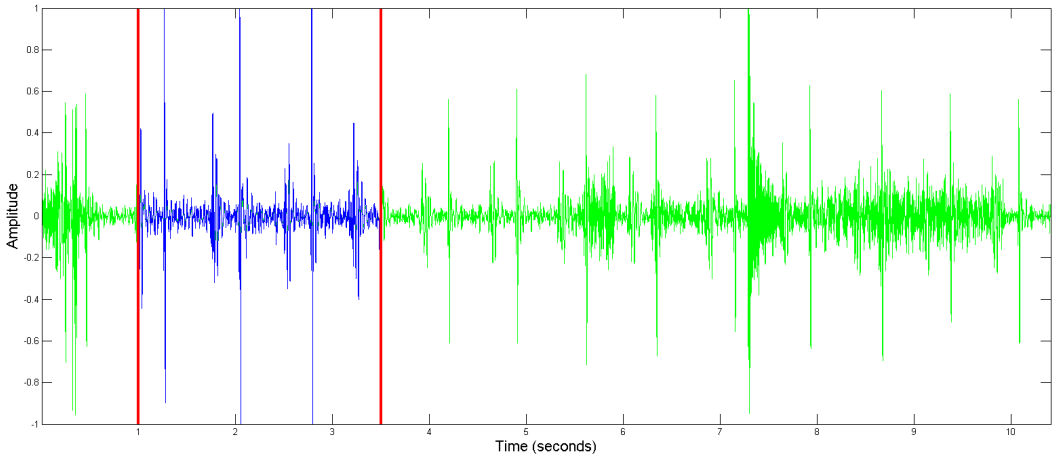

FIGURE 4.4: Segmentation of a 2,5 seconds window in a PCG signal.

After using the modulation filtering noise detection algorithm described in section 3.2.1 as a preliminary test for some time we decided that it was not necessary to undergo such heavy processing in the preprocessing window stage so we finally decided to use a simple power ratio test, inspired in the modulation filtering algorithm, to function as a selection method for the 2,5 second windows worth processing. This test divides the 2,5 windows in two and three parts, as can be observed in Figure 4.5, and calculates the power ratios P_w of each part according to equation 4.1,

$$P_w = \frac{\frac{\sum_{k \in w} |S_w(k)|^2}{L}}{\frac{\sum_{k=1}^{T} |S_w(k)|^2}{T}} \tag{4.1}$$

where T is the number of samples of signal S and w is the set that contains the samples that belong to a divided segment of size L. The features tested are the difference between the two power ratios when the 2,5 second window is divided in two parts and the maximum of the two by two differences of the three power ratios when the 2,5 second window is divided in three parts. An experimental threshold of 1 is used: if the two differences are below it the segment is suitable for further processing, otherwise the window is rejected and another segment will be chosen by sliding the window in 200 milliseconds intervals until the end of the signal. Like the modulation filtering noise detection algorithm described in section 3.2.1, this test shows good performance in detecting impulse noises with sharp peaks which are always present at the start of each signal and are generated by the placement of the stethoscope's head upon the patient's chest.

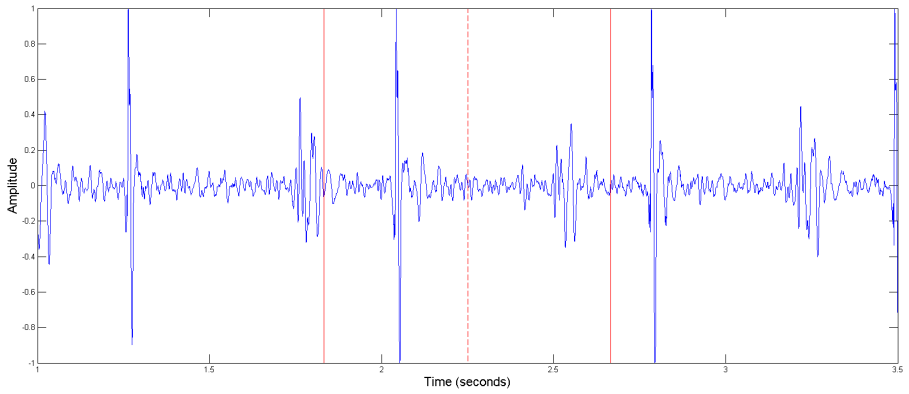

FIGURE 4.5: Division of a 2,5 second window in 2 and 3 segments.

4.1.4 Reference Heart Sound Selection

4.1.4.1 Time Domain Periodicity

The next stage will focus on the time domain analysis of each signal contained in the 2,5 second segments. It can be divided in three important steps: information retrieval related to heart rate estimation and the number of heart cycles contained in the segment, selection of the two best quality heart cycles contained in the segment, and a similarity test to verify their resemblance as two consecutive heart cycles.

It is important to refer that two different methods for extracting the signal's envelope are used in this stage. In the first step we are concerned with getting an accurate data estimation from the PCG segment therefore we use a third-level discrete wavelet decomposition to obtain the coefficients of the third level approximation. We used order six Daubechies filters and, having previously downsampled the signal to a 2205 Hz sampling-rate, the reconstructed signal's frequency band is 0 - 138 Hz. Then we use the Hilbert transform to extract the signal's envelope. Based on the conclusions referred by Hartimo et al. [30], we can state that this filtering process will allow us to filter out most of the background and other noise types and allow our information gathering efforts to focus on the heart sounds and therefore provide a better parameter estimation.

In the next steps we are interested in evaluating the segment's sound quality and noise presence. Therefore, instead of filtering the signal, we use its whole frequency range and extract the signal's envelope applying the Hilbert transform followed by a Gammatone

band-pass filter, a filter that describes the shape of the impulse response of the auditory system (for further details see Patterson et al. [31]).

The first time domain periodicity assessment step uses a wavelet filtered envelope and is performed in the following way:

- First the segment's autocorrelation is used to detect all the local maximums of the signal: The selected segment size for maximum detection is 200 samples.
- Next heart rate estimation is performed. This step is very important mainly because if an inaccurate heart cycle period is selected the information contained in the reference heart sound will not be correct and will lead the following stages of the algorithm to erroneous results. Therefore an adaptation effort was made to correct the methodology applied here. In its previous version, the algorithm used a simplified SVD calculation to estimate heart rate frequency. However, in the presence of high noise levels, this operation needs more rigorous control of its input and output data.

  Based on the facts described in sections 2.10 and 3.3.1 we used the SVD operation to measure periodicity based on the singular values extracted. To encompass the heart rate of children we estimate heart rates between 50 and 133 beats per minute (BPM) which correspond to 1200 and 450 ms heart cycle periods respectively. To increase the accuracy of this process we test period lengths in steps of 2.5 ms which gives us a total of 280 different period length assessment tests. Furthermore we apply SVD to only two heart cycles which results in an SVD input matrix with two rows, that is, for each period length we select a window with double the period size, calculate its SVR and slide it in steps of 50 ms until the end of the segment. Before each SVD operation we normalize and interpolate the segment by a variable factor so that all the tested segments have very similar lengths. This is done to insure unbiased SVD results. The resulting SVR for each period length will be the maximum SVR calculated while sliding the two period length window over the 2,5 second segment and final SVR will be maximum of the SVR's calculated for all period lengths.

  Figure 4.6 shows the SVR values obtained for each period length: as can be observed SVR values reach a peak at 69 BPM which corresponds to an accurate

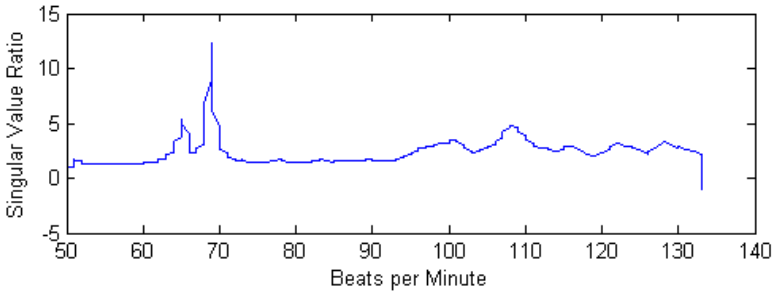

FIGURE 4.6: SVR results for a 2.5 seconds segment with an estimated cardiac frequency of 69 BPM.

heart rate frequency estimation. Figure 4.7 shows the result of the segmentation of two heart cycles based on the SVR values obtained in the previous test: the start and end of these cycles were defined by the maximum SVR values obtained while sliding the two cycle length window.

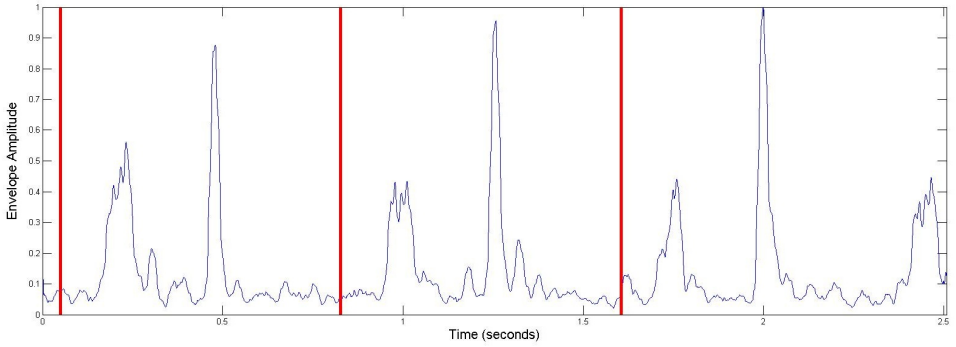

FIGURE 4.7: Heart cycle period estimation and segmentation based on SVR results.

- Based on the heart cycle period length obtained in the previous test, calculation of the strong peaks in the signal's autocorrelation is performed. Starting with the selection of the strongest peak, the local peaks previously calculated are checked in decreasing order and if they are found within the estimated period length's distance from the selected peaks, they are selected as strong peaks. This iterative process repeats itself until the estimated number of peaks is found.

The next step in time domain periodicity verification consists in selecting the segment containing the two heart cycles with better quality from the 2,5 second window. From here on we use the Hilbert transform followed by a Gammatone auditory filter to extract

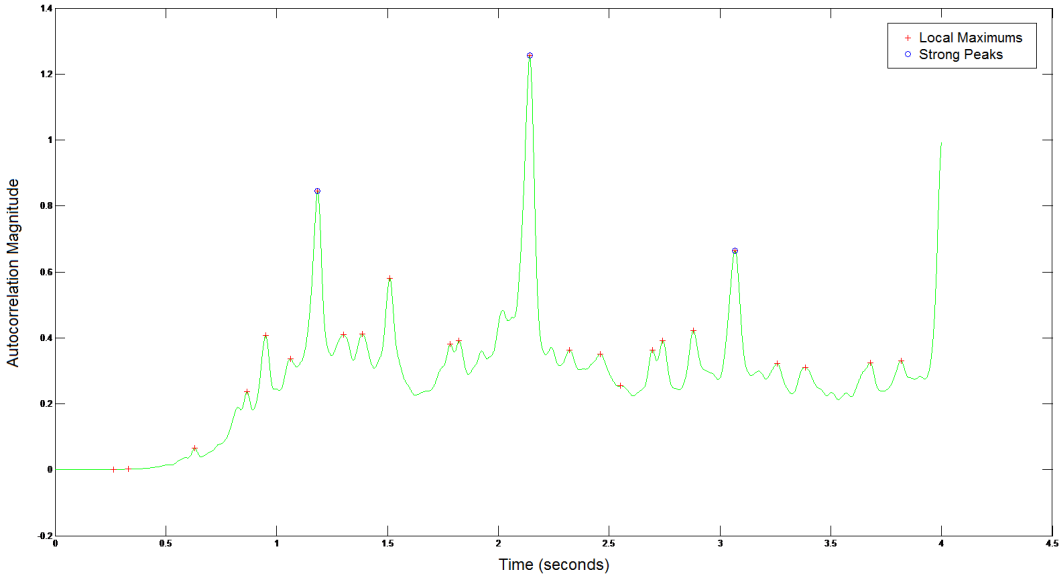

FIGURE 4.8: Peak detection of a four second HS segment.

the signal's envelope mainly because we are interested in verifying noise presence in the signal's full frequency range.

As a criterion for sound quality we use the SVD operation and base our decision in the SVR values obtained. This criterion's choice is based on the fact that high periodicity presence is directly related to low noise levels and is also used by Li et al. in [32]. This comes to us as a logical conclusion since a signal contaminated with noise, especially impulse noise, will present low periodicity presence.

To segment the two heart cycles we use a process very similar to that used for heart rate estimation. We use the estimated period length and slide the two cycle length window over the 2,5 seconds signal calculating SVR values for each step. The highest SVR value obtained will mark the limits for segmentation of the two cycles, as can be seen in Figure 4.7.

The final step in time domain periodicity verification is to use a simple test in order to verify the similarity between the two segmented heart cycles. With this purpose the cosine similarity check described in section 2.2.6 is used, with an empirically obtained threshold of 0.63 for its results. This verification is important since the use of SVD for heart rate estimation and segmentation is fallible. High noise levels will inevitably hinder the results and this additional verification will insure that that the result of heart cycle segmentation will contain two segments with similar signal envelopes.

Although time domain periodicity verification is an important step in finding a reference heart sound, it is not sufficient for assessing its quality since it is not sensitive to the presence of many non-cardiac sounds [13]. In the next section we will describe the verification performed in the time-frequency domain necessary for an accurate reference heart cycle selection.

4.1.4.2 Time-Frequency Domain Periodicity

It is important to refer that, from this point on, apart from the reference signal selection method, few changes were made in the algorithm's original implementation. As can be seen until now this algorithm consists in a sequence of tests, and if the first stages fail nothing more can be done to recover from erroneous results. Thus we chose to focus our adaptation efforts mainly on the first part of the algorithm in order to ensure reliable results for a section of the algorithm. Future work will focus on the remaining sections of the algorithm that we will proceed to describe.

Time-frequency analysis of the 2,5 second windows is performed using similar criteria to the previous algorithm steps. Transformation to the time-frequency is performed using the STFT, described in section 2.2.2, with a Hamming window, and 30 frequency bins are calculated. Based on the criteria that most of the energy of the heart sound signal is concentrated in the 0-600 Hz frequency range [13], the normalized autocorrelation for the 15 frequency bins that correspond to this frequency range is calculated for periodicity evaluation and, using the same methodology as in the previous section, the strong peaks that relate to heart cycles are calculated. Figure 4.9 shows the result of this operation on a noise contaminated segment. As can be seen, the presence of noise in the signal alters its periodic components.

To evaluate linear dependency between these frequency bins, they are grouped into three groups of five contiguous frequency bins and for each group the singular values are calculated using the SVD operation described in sections 2.10 and 3.3.1. Here the SVR is calculated using the following equation:

$$SVR = \frac{\sigma_2}{\sigma_1} 100 \tag{4.2}$$

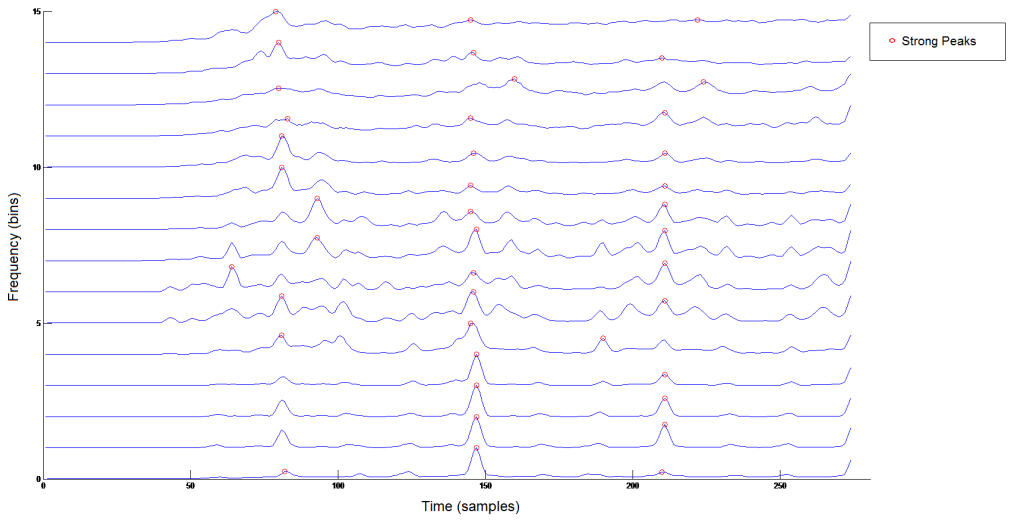

FIGURE 4.9: 15 frequency bin autocorrelation for 0 - 600 Hz frequency range with strong peaks detection results.

where σ_1 and σ_2 are the first and second highest singular values resulting from the SVD operation. Linear dependent rows, which relate to higher periodicity presence, have near zero SVR values.

Let ρ_1, ρ_2 and ρ_3 denote the SVR's of frequency bin groups F^{1-5}, F^{6-10} and F^{11-15} respectively. One important observation concerning the periodicity presence in PCGs collected from patients with native heart valves can be made by relating these three values. When analyzing pure heart sounds the main signal energy is concentrated in the lower frequency spectrum, with higher frequency components appearing only at very short periods [13]. Thus the autocorrelation function becomes less regular as frequency increases. From this fact can concluded that linear dependence between frequency bins will be lower as frequency bands increase and that the relation between SVR's, $\rho_1 < \rho_2 < \rho_3$, will be verified. If this condition is checked the segment is considered for further processing.

Next, in the time-frequency domain analysis, a peak alignment verification in each of the three frequency bin groups is performed. Given that the peaks in the autocorrelation function are due to the main heart sound components energy, without noise interference these peaks should be aligned, with a time tolerance of +/- 10% of the maximal distance between two peaks in that frequency bin group. If each frequency bin group possesses three of the five frequency bins aligned, which corresponds to an 80% total alignment,

the 2,5 second window is considered an uncontaminated heart sound segment and a reference signal can then be extracted.

4.1.4.3 Reference Signal Selection

In the previous version of this algorithm, if the 2,5 segment would pass the aforementioned verification steps, the reference heart sound signal would be extracted in the following way:

- The signal is divided in equal segments of length equal to the estimated heart rate frequency.
- For each segment, the root mean square of the energy of each frequency bin is calculated in order to analyze the ratio between the decrease of energy and the decrease of frequency.
- The degree of similarity between each segment and all the others is assessed via their correlation coeficients, calculated by the Matlab function `corrcoef` according to the following equation:

$$R(i,j) = \frac{C(i,j)}{\sqrt{C(i,i)C(j,j)}} \tag{4.3}$$

  where $C(i,j)$ refers to the covariance calculation for two equally sized matrices i and j. After summing each of the resulting matrix R rows, the segment with the highest correlation coefficient sum is chosen as the reference heart sound signal.

After some experiments to improve this method, such as dividing the segments to be tested for similarity by the peaks in the signals autocorrelation function, it was decided to use a different method. Since in a previous test, presented as the last step in section 4.1.4.1, we calculated and segmented the best two heart cycles based on a quality criteria of periodicity presence, it seemed a more reliable method to chose from these two cycles the one that would serve as reference. Thus, based on the fact that abrasion noises tend to decrease with time due to the physician's settling of the stethoscope's head when good hearing spot is found, we chose to select the second segmented heart cycle as the reference heart sound to be used in the last stage of this algorithm. Figure 4.10 shows

this selection.

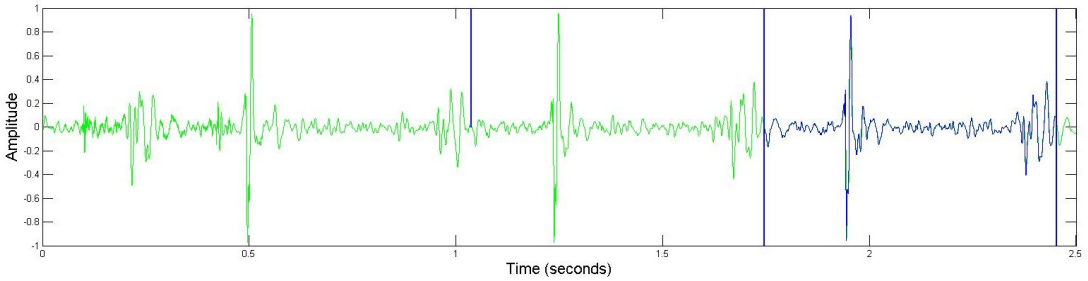

FIGURE 4.10: Selection of a reference heart sound.

4.1.5 Non Cardiac Sound Detection

This final stage of the algorithm, which has remained unaltered from its original version, consists in dividing and classifying the segments of the whole signal as clean heart sound segments or noise contaminated segments. The reference heart sound is used with a template matching approach applied using the following two features:

- *Spectral energy* consists in calculating the root mean square of the spectrogram (STFT) of the reference and tested segments to provide an estimate of the power distribution in the frequency domain. Root mean square calculation is performed according to the following equation:

$$S_{rms}(f) = \sqrt{\int_{t}^{t+T_h} |S(f,t)|^2 \, \mathrm{d}t} \tag{4.4}$$

  where S is the signal, f is the frequency bin and T_h is the length of the segments being tested, which corresponds to the length of the reference heart sound selected. Validation over all frequency bands is performed using the aforementioned Matlab function corrcoef that returns the correlation coefficient between the two signals.

$$CorrCoef(S^{ref}_{rms}(f), S^{test}_{rms}(f)) > th_1 \tag{4.5}$$

  where th_1 refers to a threshold with value 0.98. If a tested segment has a correlation coefficient below th_1 is classified as noise contaminated.

- *Temporal energy* is used to detect short duration noises. In order to gather information related to the instantaneous amplitude of both the reference and the tested signals, the energy for 50 ms windows is computed. Then the energy of each 50 ms segment of the tested signal is compared to the maximum energy of all the reference heart sound 50 ms segments. A threshold of 2.5 is used to evaluate each tested segment: if the ratio between the 50 ms segment of the tested signal's energy and the maximal energy from all the reference heart sound 50 ms segments exceeds the threshold's value the segment is classified as contaminated with an impulse sound of short duration.

If the segment successfully verifies these two tests it is classified as containing a clean heart sound.

Thus, the algorithm's output consists in an assessment matrix that contains the time intervals of each tested segment and its classification as noise contaminated or clean heart sound segment.

4.2 Experiments and Results

This section is dedicated to the preliminary results obtained with the algorithm's new revision. It will be divided in two parts: the first will explain the methodologies involved in the tests performed and the second will demonstrate the results obtained.

4.2.1 Methodology

Devising a suitable methodology for testing and quantifying the algorithm's performance was one the the main difficulties encountered when developing the work presented in this thesis. This is due to the fact that in problems where the input consists of a previously contaminated signal it is virtually impossible to rigorously separate the noisy and clean signals and quantify the segmentation's results. This is the case of the PCG signals contained in the DigiScope repository. Therefore it is important to state that the methodology encountered to test the performance of the algorithm does not allow us to infer its accuracy when confronted with data from the DigiScope repository.

In order to assess the algorithm's performance there must exist previous knowledge about which segments of the input are noise contaminated and which segments are clean heart sounds. However a 100% clean heart sound is a difficult thing to come by. Physiological noises are always present and the use of the stethoscope as a recording device introduces many artificial noises. After experimenting with a number of allegedly clean data-sets, the solution chosen for this problem was to build a small data-set with 20 clean heart sound signals built from a single segment of clean heart sound, containing a total of 405 heart cycles. The segment lengths vary between 8 and 25 seconds, which is the common time interval that a physician takes in each auscultation point. Then we manually contaminate them with small segments of different noise types gathered from the DigiScope repository. This repository provides the main motivation for the algorithm's modifications therefore an effort has been made so that the simulated signals contain similar information.

Three noise types were collected from the DigiScope repository. These are the most common noise types present in PCGs collected with digital stethoscopes and can be classified as:

- Abrasion: friction noises that occur when the physician sets and moves the stethoscope's head across the patient's chest.
- Ambient: background hospital noise, voice, music and other external noises typically present in any hospital or health care center.
- Physiological: Breathing, swallowing and other body generated sounds other than the heart sound. These noise types are very difficult to separate from the heart sound therefore only breathing sound segments, which occur more often, have been collected and added to the simulated test data.

Short noise segments with lengths between 0,5 and 4 seconds have been manually added to the test signals according to equation 4.6:

$$Y(n) = E(n) \times C + X(n) \tag{4.6}$$

where $X(n)$ is the clean signal, $E(n)$ is the noise signal, C is the coefficient intended for a specified signal to noise ratio (SNR) and $Y(n)$ is the noise contaminated signal. Three

SNR values were used when adding noise to the clean signals: 1, 5 and 10 Db. SNR was calculated according to the following equation:

$$SNR = 20 \log_{10} \frac{P_{signal}}{P_{noise}} \tag{4.7}$$

where P_{signal} denotes the clean signal's power and P_{noise} denotes the noise signal's power.

The process of contamination is mostly random in what concerns the noise type used and the time interval selected for contamination. However we chose to contaminate the initial time interval of the clean signals with an abrasion type noise to better simulate what is common in the signals from the DigiScope repository. Plus, noise segments are never overlapped. Regarding the amount of contamination we've set the time limits between 35% and 53% of the length of the clean signal. Figure 4.11 shows one artificially created sample with length of 10.4 seconds, 45.24% noise contamination, SNR = 1.0 Db and three noise types present.

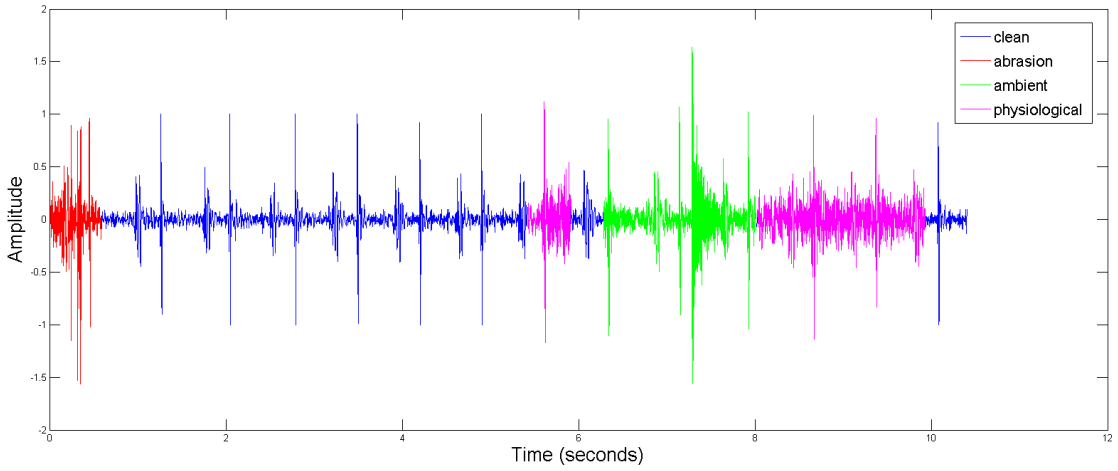

FIGURE 4.11: Artificially generated test signal containing three noise types.

Performance assessment was conducted in a similar way to that employed by Kumar et al. in [13], using sensitivity and specificity measures in the following way:

$$SE(\%) = \frac{TP}{TP + FN} \qquad SP(\%) = \frac{TN}{TN + FP} \tag{4.8}$$

where TP is the number of noisy segments correctly detected as noise contaminated, TN is the number of clean segments correctly detected as clean, FP is the number of

clean segments detected as noisy and FN is the number of noisy segments detected as clean segments.

4.2.2 Experimental Results

It is important to refer at this stage that, being a work in progress, there is yet no concern regarding performance speed issues. Therefore no remark will be made concerning this matter since the main purpose is to process data offline.

The following table shows the results obtained when processing the 20 segment data-set in terms of the sensitivity measure:

SNR (Db)	**Sensitivity (%)**
1	74.0
5	74.62
10	69.59

TABLE 4.1: Results for noise detection based on sensitivity measure (in %).

In terms of the specificity measure the following results were obtained:

SNR (Db)	**Specificity (%)**	
	Reference HS Selection	**Template Matching**
1	100.0	98.54
5	100.0	98.54
10	100.0	92.31

TABLE 4.2: Results for noise detection based on specificity measure (in %).

Specificity results for reference heart sound selection show if the segment selected for that purpose contains noise or if it was selected from a clean segment. The results are optimal which means that the initial part of the algorithm that refers to this process is functioning properly.

We can observe that with the increase of SNR the algorithm's performance is lower. This is due to the fact that noise segments with low amplitude become more difficult to detect.

As can be easily inferred, specificity results are very satisfactory due to the small presence of FP, that refer to clean heart sound segments detected as noisy segments. On the other hand, sensitivity results show that a large number of noise segments are detected as clean segments. The reason for this to happen is due to the fact that the last section of the algorithm has not been adapted to the presence of high noise levels, that is, the template matching criteria applied must still be tuned in order to improve these results.

However, quantifying results obtained when processing data directly from the DigiScope repository remains an unsolved problem. Due to the lack of knowledge concerning the amount of noise and its localization in the signal, a different methodology for verifying the results must be used. Nevertheless, we can visualize and, to some degree, infer certain conclusions from the results. To exemplify this, Figure 4.12 shows the result of processing a one minute PCG from the DigiScope repository.

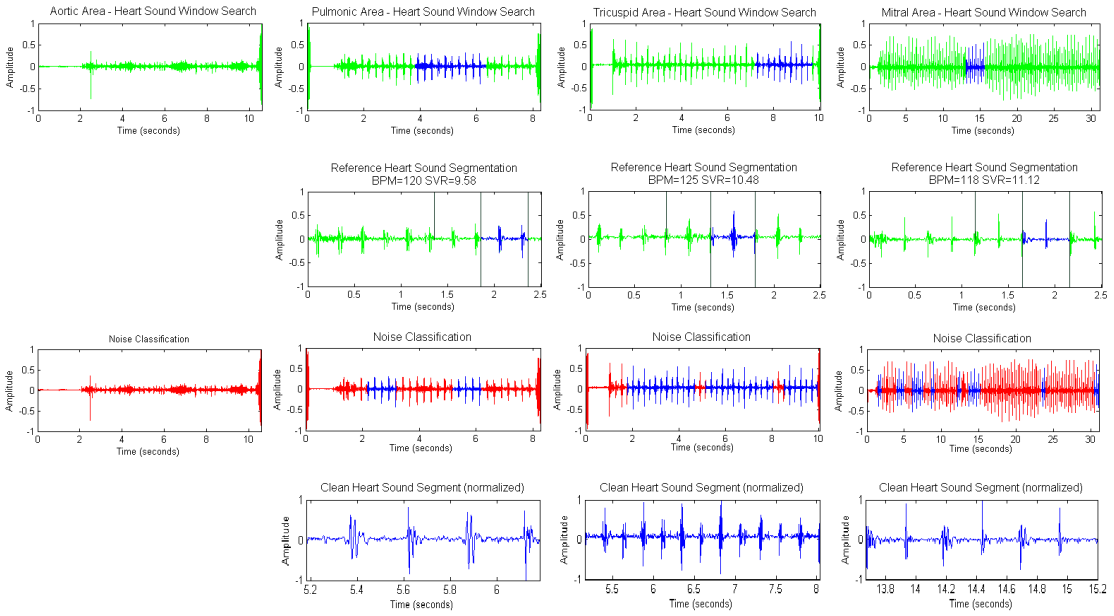

FIGURE 4.12: Noise detection results from a DigiScope repository PCG. Green signals refer to the input signal, blue signal indicate selected signals and red signals represent noise.

The first row graphics shows the segmentation of the reference HS window chosen to extract a reference cardiac cycle. As we can see, the algorithm was unable to extract such a window at the first auscultation point, the aortic area. The second row shows the segmentation of the reference heart cycle: the vertical lines denote the two heart cycles selected as having the highest SVR value, which represents the highest degree

of periodicity. The second heart cycle is chosen as reference. The third row shows the noise classification results and the forth row shows the longest segment classified as clean heart sound. This segment is the actual output of the algorithm and will be used for further signal processing and information retrieval.

While some time domain characteristics can be easily observed and certain features extracted, a more accurate and rigorous method must be devised to quantify these results in order to obtain a clear assessment of the algorithm's performance. This, as stated in the next chapter, is an important part of the future work to be developed after the presentation of this thesis.

Chapter 5

Conclusion

The work described in this thesis concerns itself with noise detection in realistic PCGs collected in hospitals and other health care facilities. The starting point was an algorithm developed for the same purpose, but with the relevant difference of aiming at processing PCGs with low noise levels collected in controlled environments. The core of our work was to adapt the algorithm to the presence of high levels of noise, which is a characteristic of realistic clinical environments, without compromising its performance. When functioning correctly, this algorithm will provide the first step in a set of algorithms developed for heart sound segmentation and pathology detection, whose ultimate goal will be to assess a patient's heart condition in a fully automatized way.

A number of important steps were taken to improve this algorithm's performance. The use of adequate filtering in the preprocessing stage and different signal envelopes for each stage of the algorithm have contributed for its better performance. Regarding the information retrieval and parameter estimation stage some relevant changes were made such as improving the autocorrelation peak estimation, the adaptation of several thresholds in periodicity and noise assessment tests, among others. But, perhaps the most important modification, was to adapt and ensure a more careful use of the SVD operation for heart rate estimation. We consider it to be a relevant contribution for a more reliable reference heart sound search since an inaccurate estimation compromises the following stages and the overall results of the algorithm.

It is important to mention that, at this stage of development, real time processing has been sacrificed in favor of better accuracy in several calculations performed, especially in

iterative processing required for reliable results. We consider this compromise inevitable when exploring new methodologies and hopefully it will be corrected when the algorithm reaches a more stable stage of development.

A number of difficulties were overcome during the development of this work. Tackling with problems that involve the chaotic behavior of noise introduces an extra difficulty which is its unpredictability. The different types of noise present in PCGs require a very robust method for detection. Thus, only with intensive testing and experimenting were we able to achieve reasonable results.

We hope, in the future, after reviewing the full extension of the algorithm, to achieve the desired result of a robust noise detection algorithm that will allow further processing of PCGs and, as an ultimate goal, to contribute to bring to reality a tool that will aid health care professionals in their daily activities.

Bibliography

[1] I. R. Hanna and M. E. Silverman. A history of cardiac auscultation and some of its contributors. *The American Journal of Cardiology*, 90(3):259 – 267, 2002. ISSN 0002-9149.

[2] R.L. Watrous. Computer-aided auscultation of the heart: From anatomy and physiology to diagnostic decision support. In *Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE*, pages 140 –143, 30 2006-sept. 3 2006.

[3] S. R. Messer, J. Agzarian, and Derek Abbott. Optimal wavelet denoising for phonocardiograms. 2001.

[4] M. E. Tavel. Cardiac auscultation. *Circulation*, 113(9):1255–1259, 2006.

[5] T. Gurney, A. Hafeez-Baig, and R. Gururajan. Exploratory study to investigate the applicability of existing digital stethoscopes in a telehealth setting. aug 2009.

[6] A. Djebbari and F. Bereksi Reguig. Short-time fourier transform analysis of the phonocardiogram signal. 2:844 –847 vol.2, 2000.

[7] D. Pereira, F. Hedayioglu, R. Correia, T. Silva, I. Dutra, F. Almeida, S.S. Mattos, and M. Coimbra. Digiscope - unobtrusive collection and annotating of auscultations in real hospital environments. In *Engineering in Medicine and Biology Society,EMBC, 2011 Annual International Conference of the IEEE*, pages 1193 –1196, 30 2011-sept. 3 2011.

[8] M. T. Coimbra F. Hedayioglu and S. S. Mattos. A survey of audio processing algorithms for digital stethoscopes. In *HEALTHINF'09*, pages 425–429, 2009.

[9] D. Song, L. Jia, Y. Lu, and L. Tao. Heart sounds monitor and analysis in noisy environments. In *Systems and Informatics (ICSAI), 2012 International Conference on*, pages 1677 –1681, may 2012.

[10] Callahan M.G. Jones J.T. Graber G.P. Foster K.S. Glifort K. Patel S.B., Callahan T.F. and Wodicka G.R. An adaptive noise reduction stethoscope for auscultation in high noise environments. may 1998.

[11] T. H. Falk and W. Chan. Modulation filtering for heart and lung sound separation from breath sound recordings. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 1859–1862, 2008.

[12] P. Carvalho, RP Paiva, D. Kumar, J. Ramos, S. Santos, and J. Henriques. A framework for acoustic cardiac signal analysis. *BioStec: BioSignals*, 2011.

[13] D. Kumar, P. Carvalho, M. Antunes, R. P. Paiva, and J. Henriques. Noise detection during heart sound recording using periodicity signatures. *Physiological Measurement*, 32(5):599, 2011.

[14] National Heart Lung and Blood Institute. How the heart works, August 2012. URL http://http://www.nhlbi.nih.gov/.

[15] B. Karnath and W. Thornton. Auscultation of the heart. *Hospital Physician*, 38 (9):39–43, sep. 2002.

[16] Indiana University Purdue University. Human anatomy website, August 2012. URL http://iupucbio2.iupui.edu.

[17] S. K. Mitra. *Digital signal processing: A computer-based approach*. MacGraw Hill, fourth edition, 2011.

[18] S. W. Smith. *The scientist and engineer's guide to digital signal processing*. California Technical Publishing, San Diego, CA, USA, 1997. ISBN 0-9660176-3-3.

[19] D. Gerhard. Audio signal classification: An overview, 2000.

[20] B. Kedem. Spectral analysis and discrimination by zero-crossings. *Proceedings of the IEEE*, 74(11):1477 – 1493, nov. 1986. ISSN 0018-9219.

[21] F. Gouyon, F. Pachet, and O. Delerue. On the use of zero-crossing rate for an application of classification of percussive sounds. In *Proceedings of the COST G-6 Conference on Digital Audio Effects (DAFX-00*, 2000.

[22] T.V. Sreenivas and R.J. Niederjohn. Zero-crossing based spectral analysis and svd spectral analysis for formant frequency estimation in noise. *Signal Processing, IEEE Transactions on*, 40(2):282 –293, feb 1992. ISSN 1053-587X.

[23] R. J. Bayardo, Y. Ma, and R. Srikant. Scaling up all pairs similarity search. In *Proceedings of the 16th international conference on World Wide Web*, WWW '07, pages 131–140, New York, NY, USA, 2007. ACM. ISBN 978-1-59593-654-7.

[24] J. Ramos. Analysis and automatic annotation of heart sounds. Master thesis, Faculdade de Cincias e Tecnologia Universidade de Coimbra.

[25] J.P. Ramos, P. Carvalho, R.P. Paiva, and J. Henriques. Modulation filtering for noise detection in heart sound signals. In *Engineering in Medicine and Biology Society,EMBC, 2011 Annual International Conference of the IEEE*, pages 6013 –6016, 30 2011-sept. 3 2011.

[26] F. Ghaderi, H. R. Mohseni, and S. Sanei. Localizing heart sounds in respiratory signals using singular spectrum analysis. *IEEE Trans. Biomed. Engineering*, 58 (12):3360–3367, 2011.

[27] P. P. Kanjilal, J. Bhattacharya, and G. Saha. Robust method for periodicity detection and characterization of irregular cyclical series in terms of embedded periodic components. *Phys. Rev. E*, 59:4013–4025, Apr 1999.

[28] P.P. Kanjilal and S. Palit. On multiple pattern extraction using singular value decomposition. *Signal Processing, IEEE Transactions on*, 43(6):1536 –1540, jun 1995. ISSN 1053-587X.

[29] F. Belloni, D. Della Giustina, M. Riva, and M. Malcangi. *A new digital stethoscope with environmental noise cancellation*, pages 169 – 174. WSEAS Press, Stevens Point, USA, 2010. ISBN 9789604742431.

[30] Liang H., 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, oct-2 nov 1997.

[31] J. Holdsworth R. Patterson, I. Nimmo-Smith and P. Rice. An efficient auditory filter bank based on the gammatone function. dec 1987.

[32] T. Li, H. Tang, T. Qiu, and Y. Park. Best subsequence selection of heart sound recording based on degree of sound periodicity. *Electronics Letters*, 47(15):841 –843, 21 2011. ISSN 0013-5194.