A Robust Heart Sounds Segmentation Module Based on S-Transform

Ali Moukadem, Alain Dieterlen, Nicolas Hueber, Christian Brandt

HAL Id: hal-00984327

https://hal.science/hal-00984327

Submitted on 28 Apr 2014

A Robust Heart Sounds Segmentation Module Based on S-Transform

Ali Moukadem[1, 3], Alain Dieterlen[1], Nicolas Hueber[2], Christian Brandt[3]

[1] MIPS Laboratory, University of Haute Alsace, 68093 - MULHOUSE CEDEX FRANCE
[2] ISL: French-German Research Institute of SAINT-LOUIS, 68300 - SAINT-LOUIS FRANCE
[3] University Hospital of Strasbourg, CIC, Inserm, BP 426, 67091 STRASBOURG CEDEX FRANCE

Abstract

This paper presents a new module for heart sounds segmentation based on S-Transform. The heart sounds segmentation process segments the PhonoCardioGram (PCG) signal into four parts: S1 (first heart sound), systole, S2 (second heart sound) and diastole. It can be considered one of the most important phases in the auto-analysis of PCG signals. The proposed segmentation module can be divided into three main blocks: localization of heart sounds, boundaries detection of the localized heart sounds and classification block to distinguish between S1 and S2. An original localization method of heart sounds are proposed in this study. The method named SSE calculates the Shannon Energy of the local spectrum calculated by the S-transform for each sample of the heart sound signal. The second block contains a novel approach for the boundaries detection of S1 and S2. The energy concentrations of the S-transform of localized sounds are optimized by using a window width optimization algorithm. Then the SSE envelope is recalculated and a local adaptive threshold is applied to refine the estimated boundaries. To distinguish between S1 and S2, a feature extraction method based on the Singular Value Decomposition (SVD) of the S-matrix is applied in this study. The proposed segmentation module is evaluated at each block according to a database of 80 sounds, including 40 sounds with cardiac pathologies.

1. INTRODUCTION

Cardiac auscultation is the basis for heart examination. It provides a wealth of information about structural and functional cardiac defects, using a simple, efficient and costless medical device: the stethoscope. Invented in the nineteenth century, this acoustic instrument has proved since to be of paramount importance to the physical examination and diagnosis of cardiac pathologies. Over the course of the past two centuries, the stethoscope underwent numerous improvements to reach the development of the electronic stethoscope capable to register and optimize the quality of the acoustic signal, completed by the PhonoCardioGraphic (PCG) presentation of the auscultation signal

The PCG signal confirms, and mostly, refines the auscultation data and provides further information about the acoustic activity concerning the chronology of the pathological signs in the cardiac cycle, by locating them with respect to the normal heart sounds. The cardiac sounds are by definition non-stationary signals, and are located within the low frequency range, approximately between 10 and 750 Hz [1].

The analysis of the cardiac sounds, solely based on the human ear, is limited by the experience of the clinician for a reliable diagnosis of cardiac pathologies, and to obtain all the qualitative and quantitative information about cardiac activity. Information, such as the temporal localization of the heart sounds, the number of their internal

components, their frequential content, and the significance of diastolic and systolic murmurs, can also be studied directly on the PCG signal. In order to recognize and classify cardiovascular pathologies, advanced methods and techniques of signal processing and artificial intelligence will be used.

For that, two approaches could be considered to improve electronic stethoscopes:

• Stethoscope with embedded autonomous analysis, simple for home use by patients and paramedics, for the purpose of autodiagnosis and follow up.

• Stethoscope coupled with a hosting device or a server for sophisticated analysis (coupled to a PC with a Bluetooth link) for the use of professionals in order to improve performance of clinical medical diagnosis.

Whatever the approach, one of the first phases in the analysis of heart sounds, is the segmentation. Heart sound segmentation process segments the PhonoCardioGram (PCG) signal into four parts: S1 (first heart sound), systole, S2 (second heart sound) and diastole. First, S1 and S2 are located then systole is represented by the interval S1 to S2 and diastole by the interval S2 to S1.

Identification of the two phases of the cardiac cycle and of the heart sounds with robust differentiation between S1 and S2 even in the presence of additional heart sounds and/or murmurs is a first step in this challenge. Then there is a need to measure accurately S1 and S2 allowing the progression to automatic diagnosis of heart murmurs with the distinction of ejection and regurgitation murmurs.

This phase of autonomous detection, without the help of ECG is based on signal processing tools such as: the Shannon energy [1], which can be considered as morphological based transformation that operates on the energy of signal and its performance will certainly decrease in presence of noise or murmurs [2]. The high order statistics based methods; Shannon entropy [3], variance fractal dimension [4], recurrence time statistics [5], all these methods performed well in the presence of respiratory noise but this is not the case in the presence of murmur data [5], in addition the computing time of these methods is very high. The Radial Basis Functions (RBF) neural network method which was used to extract the envelope of the heart sound [2], it was shown to have a good performance on low level noise signals; however, in the presence of high level of noise, the performance of the RBF method decreases. This was not surprising because the method operates directly on the heart sound without any feature extraction step. To deal with this problem, a method for heart sounds localization named SRBF was proposed [6]. This method aims at extracting the envelope of the signal by applying the features extracted from the S-Transform matrix of the heart sound signal to the radial basis function (RBF) neural network (SRBF). The SRBF method was shown to have a significant enhancement in term of sensitivity comparing to other existing methods [6].

The Time-Frequency (TF) domain can lead to a more robust localization and classification methods especially for non-stationary signals like the heart sounds. S-Transform originates from two advanced signal processing tools, the short time Fourier transform (STFT) and the wavelet transform. It can be viewed as a frequency dependent STFT or a phase corrected wavelet transform [7]. The S-Transform has been proven in to perform better than other time-frequency /scale transforms for heart sounds signal analysis [8, 9].

In this paper we present a new module for heart sounds segmentation based on S-Transform. The proposed segmentation module can be divided into three main blocks: localization of heart sounds, boundaries detection of the localized heart sounds and classification block to distinguish between S1 and S2. The proposed methods in each block can be summarized as follow:

- A new localization method of heart sounds based on the S-transform is proposed in this study. The method named SSE, calculates the Shannon Energy of the local spectrum obtained by the S-transform for each sample of the heart sound signal in order to extract its envelope.

- The second block contains a novel approach for boundaries detection of S1 and S2. The energy concentrations of the S-transform of localized sounds are optimized by using a window width optimization algorithm then the SSE envelope is recalculated and a local adaptive threshold is applied to refine the estimated boundaries.

- A feature extraction method based on the Singular Value Decomposition (SVD) of the S-matrix is applied to classify S1 and S2.

The proposed methods are evaluated based on a database of 80 sounds (40 pathologic). This study is made under the control of an experienced cardiologist, in with the aim of validating the results of each method.

This paper is organized as follows: Section 2 describes the data base used in this study. It is followed by the Section 3 which describes the different methods proposed for the segmentation module (localization, boundaries detection and classification). The results and discussion are presented in Section 4 and Section 5 gives the conclusion.

2. Data Base

Several factors affect the quality of the acquired signal, above all, the type of the electronic stethoscope, its mode of use, the patient's position during auscultation, and the surrounding noise [5]. According to the cardiologist's experience, it's preferable that the signals remain unrefined; filtration will only be applied subsequently in the purpose of signal analysis. For this reason we used prototype stethoscopes produced by Infral Corporation, and comprising an acoustic chamber in which a sound sensor is inserted. Electronics of signal conditioning and amplification are inserted in a case along with a Bluetooth standard communication module.

Different cardiologists equipped with a prototype electronic stethoscope have contributed to a campaign of measurements in the Hospital of Strasbourg (France); recruitment was made through clinical research project (HUS-PRI 4179) with the support of the clinical investigation center (INSERM). In parallel, 2 prototypes have dedicated to the Cardio-Psy experience as a part of the MARS500 project (IBMP-Russia) promoted by ESA (European Spatial Agency – PI. A. Aubert), in order to collect signals from 6 volunteers (astronauts). The use of prototype electronic stethoscopes by different cardiologists makes the database rich in terms of qualitative diversity of collected sounds.

The sounds are recorded with 16 bits accuracy and 8000Hz sampling frequency in a wave format, using the software "Stetho" developed under Alcatel-Lucent license.

The dataset contains 80 sounds, including 40 pathological cardiac sounds. All of the participants have given a written informed consent. The normal subjects were aged from 18 to 40 years. The 6 participants of the MARS 500 experience having completed their examination as astronauts, the other have been included after examination on clinical basis.

The pathologic set of patients has been recruited during consultation of hospitalization in the University Hospital of Strasbourg. They were aged from 44 to 90 years including 20 men and 10 women. Ten of these patients have been registered twice generally before and after valvular surgery.

The auscultation was done in the conditions of clinical examination. The best auscultation focus has been registered. Duration of registration varied from 6 to 12 seconds while patients controlled their respiration.

The diagnosis of the pathologic patients was made by an experienced cardiologist as a medical file including ECG and Echocardiography-Doppler.

Patients with valvupathy (aortic stenosis, mitral insufficiency including mitral valve prolaps and mitral insufficiency and ventricural syptol defect) represented 18 registrations.

Cardiac prosthetic valve patients in mitral or aortic position represented a set of 9 patients the prothestic valves was mainly mechanical with 1 bioprothesis. 4 patients beard double prosthesis in aortic and mitral position. The other pathologic patients presented rhythm disturbances from ventricular extra systoly, AV Block (1 and 2 degree) to tachyarrhythmia in a context of ischemic cardiomyopathy.

3. Method

3.1. Preprocessing

At first the original signal is decimated by factor 4 from 8000 Hz to 2000 Hz sampling frequency and then the signal is filtered by a high-pass filter with cut-off frequency of 30 Hz, to eliminate the noise collected by the prototype stethoscope. The filtered signal is refiltered reverse direction so that there is no time delay in the resulting signal. Then, the Normalization is applied by setting the variance of the signal to a value of 1. The resulting signal is expressed by:

$$x_{norm}(t) = \frac{x(t)}{\left|\max(x(t))\right|} \qquad (1)$$

3.2. Localization of heart sounds

The localization algorithms operating on PCG data try to emphasize heart sound occurrences with an initial transformation that can be classified into three main categories: frequency based transformation, morphological transformations and complexity based transformations [5].

The transformation try to maximize the distance between the heart sounds and the background noise, and the result is smoothed and thresholded in order to apply a peak detector algorithm. We note here, that the main goal of heart sound localization is to locate the first and the second heart sounds but without distinguishing the two from each other and without detecting the boundaries of located sounds.

3.2.1. SSE localization method

A new method for the localization of heart sounds is proposed in this study (SSE). It uses the S-matrix like the SRBF method (0-100 Hz) [6] and it calculates the Shannon Energy (SE) of the local spectrum calculated by the

S-transform for each sample of the signal *x(t)*. Then, the extracted envelope is smoothed by applying an average filter (figure 1).
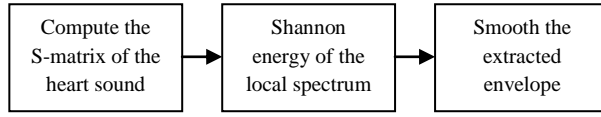

Figure 1: Block Diagram of SSE Method

The S-Transform proposed in [8], of a time series $x(t)$ is:

$$S(\tau, f) = \int_{-\infty}^{\infty} x(t)w(\tau - t)e^{-2\pi ift}dt \qquad (2)$$

Where the window function $w(\tau - t)$ is chosen as:

$$w(t, f) = \frac{1}{\sigma(f)\sqrt{2\pi}} e^{\frac{-t}{2\sigma f^2}} \qquad (3)$$

And $\sigma(f)$ is a function of frequency as:

$$\sigma(f) = \frac{1}{|f|} \qquad (4)$$

The proposed SSE method calculates the Shannon energy of each column of the extracted S-matrix as follows:

$$SSE(x_i) = \sum_{j=0}^{n} S(j,i)^2 \log(S(j,i)^2) \qquad (5)$$

Each column of the S-matrix represents the local frequency at a specific sample. The advantage of the Shannon energy transformation is its capacity to emphasize the medium intensities and to attenuate low intensities of the signal which represents the local spectrum in the case the SSE method. The main difference between the SSE and the SRBF method [6] is the training phase needed for the RBF module. The RBF neural network in the SRBF method can be considered as a non-linear filter which is replaced with a simple average filter in the SSE method.

We note that an automatic elimination of the extra detected peaks due to a split or a signal noise contamination is performed by applying the algorithm proposed in [1].

3.3. Boundaries detection Algorithm: an optimized S-transform approach

The boundaries detection algorithm aims at estimating the onset and the endpoint of the located heart sounds. Accurate boundaries estimation is a very important step in the heart sound segmentation module and it is essential for the extraction of meaningful features from each part of heart cycles in order to perform an auto-diagnosis process.

3.3.1 Overview of existing methods

Different boundaries detection algorithms exists in the literature, in [1] the boundaries are estimated by applying a threshold on the extracted envelope of the signal, this is not accurate for some cardiac cycles, because the

envelope threshold level is used based on the average value of the whole recordings periods. In [10] the same authors propose another algorithm that employs the STFT (Short Time Fourier Transform) to explore the time-frequency domain of the signal. Authors quantify the spectrogram at each segment to two values by applying a threshold that reserves 60% of the signal energy, however, it is not clear how the energy of the signal is calculated and the accuracy of the algorithm is not mentioned. In [11] authors use some biomedical features of heart sounds (S1 and S2) like the maximum duration of S1 and S2 to determine the limit of estimated boundaries, the disadvantage of this method is that the estimation of energy of the signal is based on the time domain only, so in the presence of high level of noise the performance of this method will decrease dramatically.

3.3.2 The OSSE Algorithm

In this paper, we propose a new algorithm to estimate the heart sounds boundaries. The proposed algorithm tries to optimize the energy concentration of the S-transform at each located sound by using a window width optimization method. The envelope of the optimized S-transform is then recalculated by using the SSE approach and an adaptive threshold is applied to determine the onset and the ending of each located heart sound. Let us assume that L is the time located sounds after applying the localization method on the heart sound and $S(M,N)$ is the S-matrix of the heart sound where M represents the frequency domain and N the time domain.

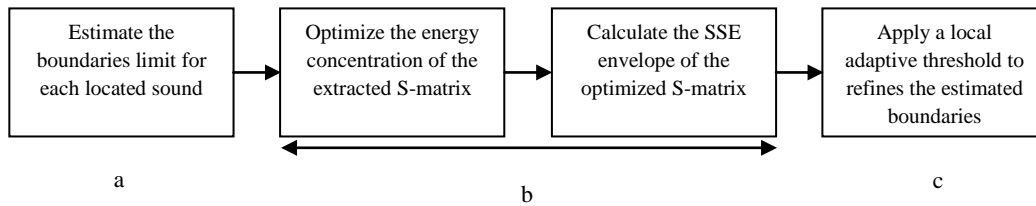The block diagram of the proposed algorithm (OSSE) is shown below (figure 2).

Figure 2: The block diagram of the OSSE Method

a. Estimate the boundaries limit

The boundaries limits are estimated based on the fact that the maximum duration of S1 and S2 is 150 ms [12]. So a 150ms window is applied in the proximity of detected S1 and S2 peaks which covers 75ms in the backward direction of the $S1$ or $S2$ peak and 75ms in the forward direction.

b. Optimized S-transform

Many studies tried to improve the Time-Frequency (TF) representation of the S-transform [13, 14]. The main study in the literature interested to optimize the energy concentration in the TF domain was in [15]. That is, to minimize the spread of the energy beyond the actual signal components. As it is well known, the ideal time-frequency transformation should only be distributed along frequencies for the duration of signal components. So the neighboring frequencies would not contain any energy and the energy contribution of each component would not exceed its duration [16].

The energy concentration in the Time-Frequency (TF) domain is a very important parameter for the algorithms that aim to detect the duration of any given events in a signal. Therefore, it should hold the same importance for the boundaries detection algorithm of heart sounds based on time-frequency features. However, in some cases,

6

the S-transform suffers from poor energy concentration in TF domain. Hence, the importance of an energy concentration optimization process to improve the boundaries estimation of the heart sounds.

The main approach used in [15] was to optimize the width of the window used in the S-transform. The width of the Gaussian window can be controlled by several ways by adding a new parameter to the window equation. We use in this study the parameter p introduced in [15] and we investigate another parameter named α (see equation 6). Both of them control the Gaussian window width:

$$\sigma(f) = \frac{\alpha}{|f|^p} \qquad (6)$$

We note here that in this paper when α vary, p is fixed to 1, and when p vary, α is fixed to 1. The optimal value can be calculated in two methods; the first method calculates one global parameter, which is recommended for signals with constant or slowly varying frequency components. The second method calculates the time-varying parameter (so an optimal parameter for each sample) which is more suitable for signals with fast varying frequency components. The disadvantage of the second approach is its high computational complexity which makes it unsuitable for applications where time is an important factor.

Based on the first approach, the optimization algorithm is applied on both parameters p and α, separately. The performance measure against each parameter is compared in section (5.2). The performance measure is based on the concentration measure (*CM*) proposed in [17]. For each α (or p) from a given set, the *CM (α)* can be expressed by:

$$CM(\alpha) = \frac{1}{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \left|\overline{S_x^\alpha(t,f)}\right| dt df} \qquad (7)$$

With $\overline{S_x^\alpha(t,f)}$ is the normalized energy of the S-transform for each α; it's given by [14]:

$$\overline{S_x^\alpha(t,f)} = \frac{S_x^\alpha(t,f)}{\sqrt{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \left|S_x^\alpha(t,f)\right|^2 dt df}} \qquad (8)$$

The *CM (α)* and *CM (p)* are calculated and compared for all existing S1 and S2 sounds in the database. We note again that the main objective is to enhance the concentration energy of the S-transform in order to detect precisely the boundaries of the located heart sounds. We consider the parameter that reaches a higher *CM* to be more appropriate for the heart sound signals.

c. The Adaptive threshold

Performing an optimized S-transform before calculating the SSE envelope makes the choice of threshold less sensitive to the variation of different heart sounds. In this study, a threshold which equals to 10 % (see figure 9) of the maximum value of the SSE envelope is applied to refine the estimated boundaries.

7

4. Distinguishing S1 and S2

Most of the existing methods for the segmentation of heart sounds use the feature of systole and diastole duration to classify the first heart sound (S1) and the second heart sound (S2) [1, 18, 19]. These time intervals can become problematic and useless in several clinical real life settings which are particularly represented by severe tachycardia or in tachyarrhythmia (figure 3). To deal with this problem, an unsupervised method for the discrimination of S1 and S2 using the high frequency information was proposed in [20, 21]. This method is based on a simple threshold applied on the high frequency envelope and the heart sounds which exhibit high frequency are considered as S2 sounds. However, the method is very sensitive to the selected threshold and the frequency content of heart sounds can be very close and overlap; which makes the task of detection of high frequency signatures very difficult.
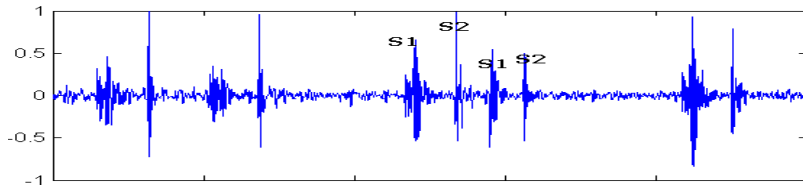

Figure 3: Example of an arrhythmic subject.

In this study, we present another method to classify S1 and S2 based on the Singular Value Decomposition (SVD) technique applied on the S-matrix and the results are analyzed and compared in section 5.3 with the high frequency signatures detection method proposed in [20].

4.1 Feature extraction based on the SVD

The SVD is a powerful tool that provides a compact matrix or compact significant information about single signal. Different ways exist in the literature aims to represent the time-frequency matrix in a compact manner by using the SVD technique. In [22], authors extracted the eigenvalues of the time-frequency matrix. In [23], authors extended the method to also incorporate information from the eigenvectors to classify EEG seizures. In [24], the last technique is applied on the S-matrix in the aim to extract features for systolic heart murmur classification. Following this approach, this study proposes a feature extraction method for S1 and S2 classification.

The time-frequency analysis is performed by the S-Transform. The S-matrix S_i of the extracted heart sound Hi is decomposed by the SVD technique as follows:

$$S_i = UDV^T \qquad (9)$$

Where $U(M \times M)$ and $V(N \times N)$ are orthonormal matrices so their squared elements can be considered as density function [21], and $D(M \times N)$ is a diagonal matrix of singular values. The columns of the orthonormal matrices U and V are called the left and right eigenvectors which contains in this case the time and frequency domain information, respectively. The eigenvectors related to the largest singular values contain more information about the structure of the signal.

Based on our experience, in this study, the first left eigenvector and the first right eigenvector that correspond to the largest singular values are used for the feature extraction process. The histogram (10 bins) for each related distribution function is calculated based on the density function. Five feature vectors obtained by this method are

8

tested in the classification process; the eigentime histogram vector U_1 (T-Features), the eigenfrequency histogram vector V_1 (F-Features), the singular values vector D_1 (SV Features) and the time-frequency vector $U_1\&V_1$ (TF Features). All vectors have a length of 10 features except the time-frequency vector that has a length of 20.

5. Results and Discussion

5.1 Localization Methods

The performance of the SBRF and the SSE methods was measured as the methods capacity to locate S1 and S2 correctly. It was measured by sensitivity and positive predictive value:

$$Sensitivity = \frac{TP}{TP + FN} \qquad (10)$$

And positive predictive value:

$$PPV = \frac{TP}{TP + FP} \qquad (11)$$

A sound is true positive (TP) or correctly located if the detected sound corresponds to a S1 or S2 sound predefined manually by the cardiologist, all other detected sounds were defined as false positive (FP) and all missed sounds are considered as false negative (FN).

Table1: Sensitivity and Positive Predictive Values for the SRBF and SSE methods applied on the clinical sounds sets (Normal and Pathological) without and with additive Gaussian noise.

Method	Normal				Pathological			
%	Sensitivity	PPV	Sensitivity (Noise)	PPV (Noise)	Sensitivity	PPV	Sensitivity (Noise)	PPV (Noise)
SRBF	91%	99%	90%	91%	94%	98%	91%	93%
SSE	96%	95%	93%	94%	97%	95%	94%	93%

Normal sounds

Results in Table 1 for normal sounds show that SRBF method reaches a higher PPV (99%) than the SSE method for the clinical signals without any additive noise. However, SSE reaches a higher sensitivity (96%) than the SRBF method (91%). The supervised approach performed by the RBF block in the SRBF method makes the extracted envelope more discriminative between the different parts of the signal than the unsupervised SSE method. Therefore, it is not surprising that the number of false detected sounds in the SRBF method is lower than the SSE method, which also explains the PPV results. The same reasons can also account for the false negative alarms which are higher in the SRBF method than the SSE method and which gives a higher sensitivity to the SSE method. In the presence of an additive white Gaussian noise, the performance of the SSE method is better with 93% sensitivity and 94% PPV. The robustness of both methods against noise is very significant. This is due to the advantage of performing a time-frequency analysis which makes methods more robust against noise. Figure 4 shows the envelopes extracted by the both methods applied on normal heart sound with and without additive noise.
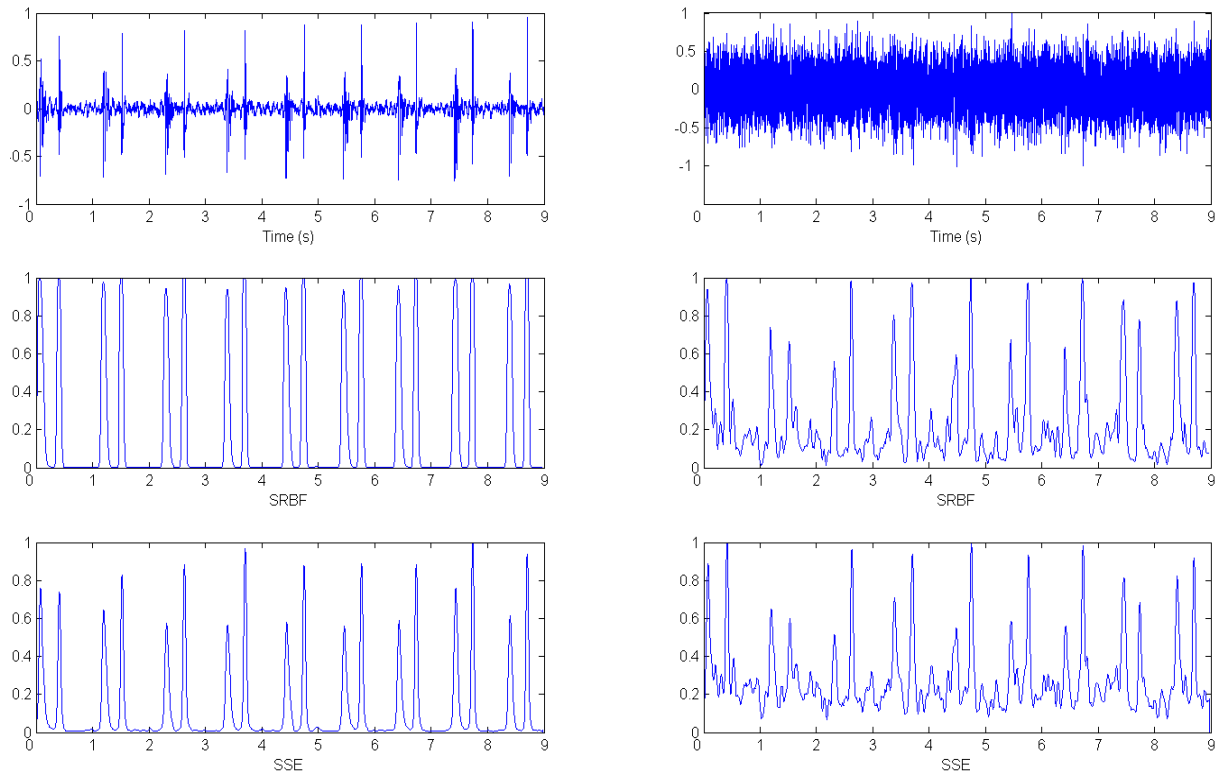
Figure 4: (top) Envelope extraction for two normal PCG signal without and with additive Gaussian noise, (middle) their SRBF envelopes, (bottom) their SSE envelopes.

Pathological sounds

The results in table 1 for pathological sounds like the normal sounds show that SRBF method reaches a higher PPV (98%) than the SSE method for the clinical signals without any additive noise. In the case of additive noise, both methods reach the same PPV (93%). The sensitivity of the SSE method (97%) still higher than the SRBF method (94%).This can be explained by the same reasons cited above. The performance of methods for pathological sounds, most notably the SSE method, is very significant. Both methods are developed to avoid the detection of murmurs so it's not surprising to reach a high performance against pathological sounds with different types of systolic murmurs.

Figure 5 shows the envelopes extracted by the SSE and the SRBF method that correspond to a pathologic sound with a heavy systolic murmur. Figure 6 shows the robustness of each method against white additive noise.
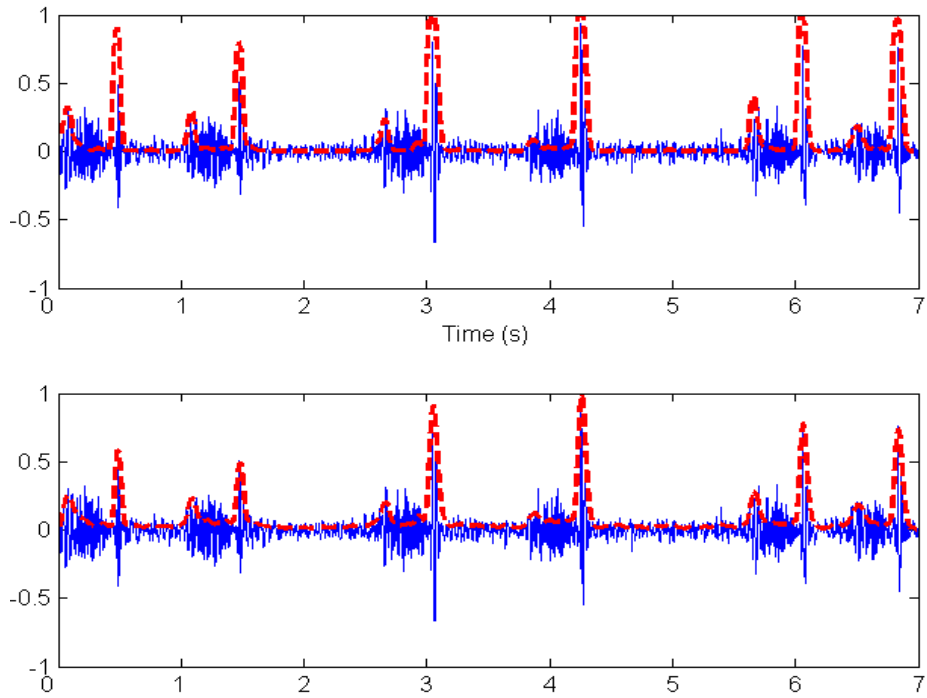
Figure 5: Envelope extraction (dashed lines) for a signal with systolic murmur, (top) SRBF envelope, (bottom) SSE envelope.
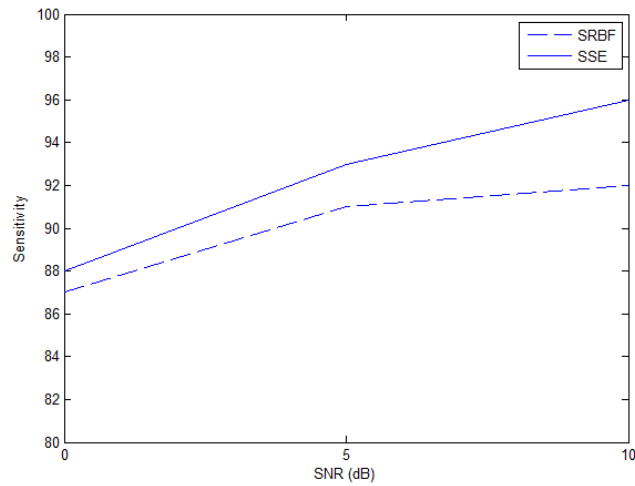

Figure 6: Sensitivity vs. SNR for the SRBF and SSE methods.

5.2 Boundaries detection

Table 2: Performance measure given by the maximum values of CM (α) and CM (p) for a given parameters set of α and p, respectively.

Heart Sounds	Optimal α	CM(α)	Optimal p	CM(p)	CM(α =1, p=1)
S1	0.82±0.45	0.0185±0.0017	1.1±0.5	0.0186±0.0018	0.0177±0.0015
S2	0.55±0.3	0.0186±0.0015	1.37±0.5	0.0186±0.0018	0.0175±0.0014
Total	0.68±0.37	0.0185±0.0016	1.23±0.5	0.0186±0.0018	0.0176±0.0015

The performance measure against each parameter is compared (Table2). The values of α and p are chosen from a set; $0<α<2$, $0<p<2$, with a step of 0.1; so twenty values as total for each variable. The optimal α is reached when CM(α) is maximized, and the optimal p is reached when CM(p) is maximized. Results from Table2 show that there are no significant differences between the two parameters $α$ and p concerning the performance measure. However, results show an important difference between optimized concentration measure and standard concentration that correspond to the standard S-transform with $α=1$ and $p=1$. The maximum values of concentration measures CM(α) and CM(p), that corresponds to the optimum $α$ and p, respectively, are obtained with $α<1$ and $p>1$. This is can be explained by the fact that when $α<1$ and $p>1$, the Gaussian window of the S-transform is narrower, which improves the detection of the sudden changes in the signal, like the onset and the ending of the first and the second heart sounds. However, when a window is narrower in time domain, we loss in term of frequency resolution. The compromise is performed by the optimization process that operates on the variable that control the variance of the Gaussian window, $α$ or p for example. The criterion of the performance is the concentration energy measure. The enhancement of energy concentration in the TF domain, influence clearly on the boundaries estimation results (Table 3).

Table 3: S1 and S2 durations (ms) estimated by the SSE and OSSE methods with and without additive noise.

Method	S1(ms)	S1(Noise)	S2(ms)	S2 (Noise)
SSE	122.4±7.2	127.8±9.6	95.2±8.3	101.2±7.4
OSSE	110.7±4.32	113.6±6.5	69.1±5.4	77.9±8.2
Reference	105.8±6		74.8±5.65	

The "Reference" row in Table 3 represents the manual measures made by the cardiologists by using the software "Stetho" developed under the license of Alcatel-Lucent. Limits of heart sounds were defined with classical Phonocardiographic references; mitral closure initiating S1 with high frequency vibration, the aortic closure generating the high frequencies beginning S2 and the end of S1 and S2 are annotated by the end of the high amplitude vibrations.

Results show the efficiency of optimizing the energy concentration of the S-transform in order to estimate more realistic boundaries for S1 and S2. Measures obtained by the SSE algorithm (without optimizing the S-transform) are always higher than the results given by the OSSE algorithm where an optimization process is performed. This is not surprising since the OSSE algorithm has a better energy concentration in the TF domain, which minimizes the spread of the energy beyond the S1 and the S2. Figure 7 shows the boundaries detection results, with and without optimization of the S-transform, applied on a S2 example and figure 8 shows the OSSE results applied on the entire heart sounds (normal and pathologic). Figure 9 shows the variation of the Mean Absolute Error (MEA) in ms against the selected threshold. The MEA is calculated by the difference between the Reference results and the results obtained by the automatic method. The lower errors reached for S1 and S2 correspond to 10% and 15% threshold values, respectively. The difference between the corresponding MEA is not significant, so the selected threshold in this study is fixed to 10% of the maximum value of the SSE envelope.
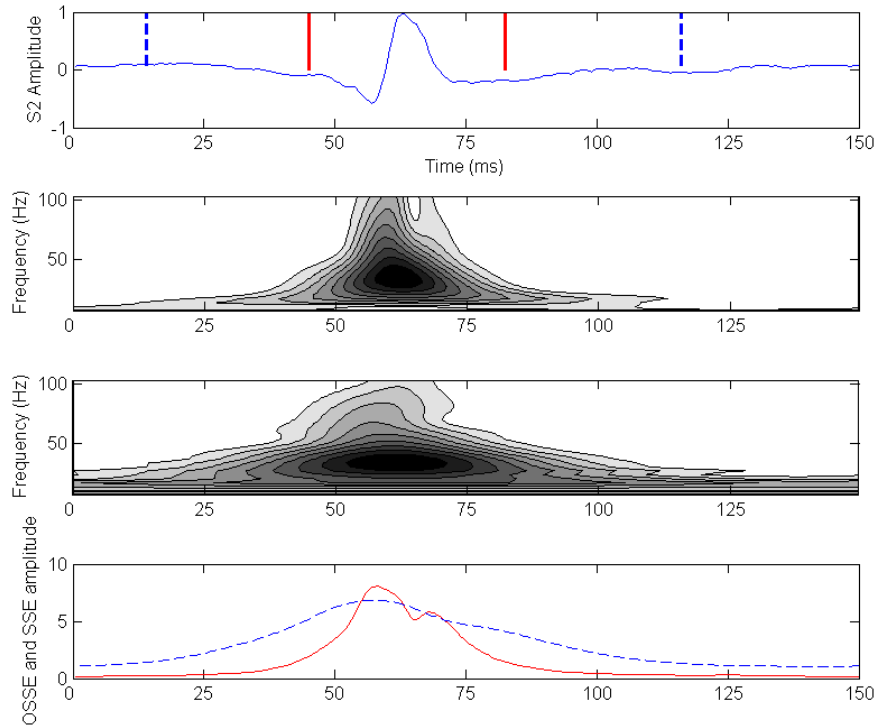
Figure 7: (top) S2 signal with two detected boundaries calculated by the optimized S-transform and the standard S-transform (dashed line), S-transform with the optimum value $\alpha=0.5$ ($p=1$), standard S-transform with $\alpha=1$ ($p=1$), (bottom) SSE envelope for the optimized S-transform and standard S-transform (dashed line).
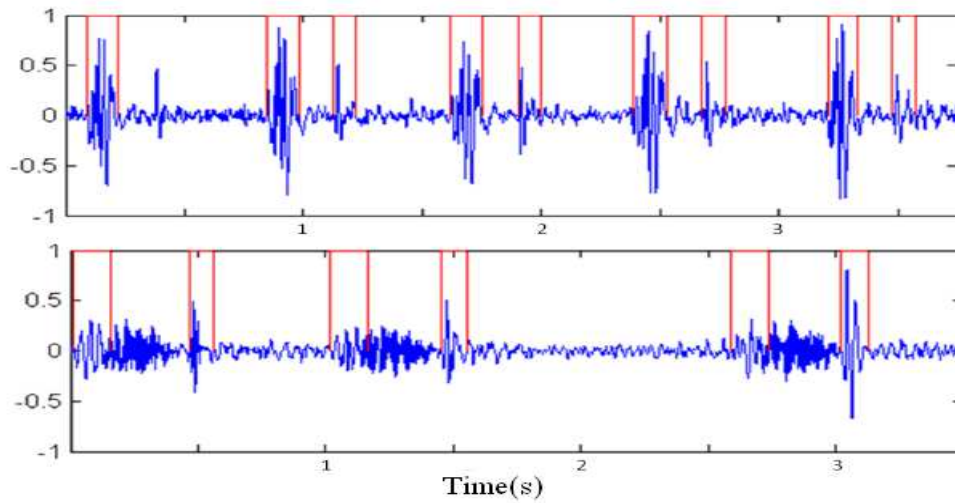

Figure 8: OSSE method applied on a normal heart sound (top) and pathological heart sound (bottom).
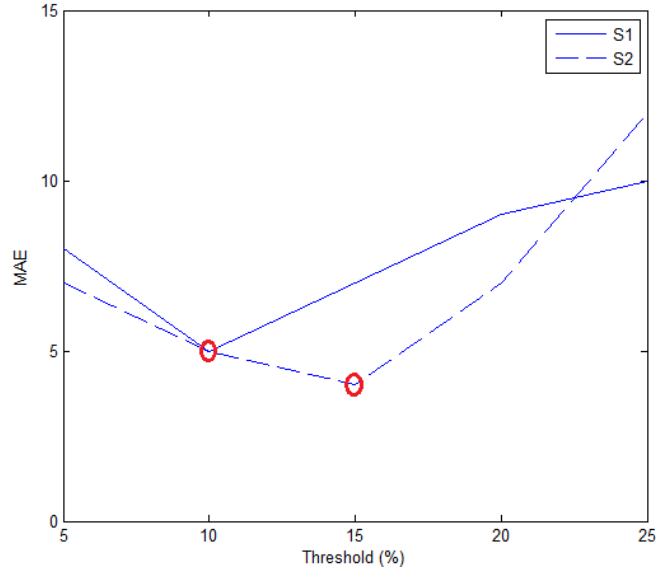
13

Figure 9: Mean Absolute Error (MEA in ms) Vs. Threshold for S1 and S2 sounds.

5.3 S1 and S2 classification

The localization of heart sounds is established by using the SSE method. The boundaries of the heart sounds are determined by the OSSE algorithm. The results were visually inspected by a cardiologist and erroneously extracted heart sounds were excluded from the study. The feature extraction process extracts a feature vector per extracted sound S_i (S1 or S2) and each of these vectors is averaged across available extracted sounds from each subject. So from each subject in the database, we obtain one S1 feature vector and one S2 feature vector to use in the training and classification process.

A 3-Neirest Neighbor (KNN) classifier is used to evaluate the performance of the four feature vectors obtained and the 5-fold approach is used for cross validation. The choice of KNN classifier was based on its simplicity of and its robustness to a noisy training data.

The method proposed in [20] based on the classification of S1 and S2 by High Frequency Signatures (HFS) is also implemented and compared with the proposed method. We note here that the HFS method tries to find the presence of high frequency information by calculating the Shannon energy of the detail coefficient obtained by the Wavelet Transform and an adaptive threshold is defined to detect the HFS, so this method doesn't use the KNN classifier.

Table 4: sensitivity and specificity for the four feature extracted vectors evaluated by a KNN classifier.

Classifier	KNN				Threshold
	T-Features	F-Features	SV Features	**TF Features**	HFS
Sensitivity	92%	81%	60%	**95%**	89%
Specificity	92%	88%	65%	**97%**	91%

The time domain feature vector reaches 92% classification rate, however, the frequency feature vector reaches 81% sensitivity and 88% specificity. The Time-Frequency vector (TF Features) reaches the higher classification

rate with 95% sensitivity and 97% specificity. The singular values are almost indistinguishable from each other and it is shown by the low classification rate for the SV features (Table4). The HFS method reaches 89% sensitivity and 91% specificity. The frequency content of heart sounds can be very close and overlap; which makes the task of detection of high frequency signatures very difficult (even in normal sounds) especially when the high frequency signature in S2 has low intensity.

In most cases seen in the medical field, S2 has a higher frequency than S1. This is due to the fact that S2 is the heart sound associated with the closure of the aortic valve in a context of high left ventricular pressure, the mitral closing occurring at low left ventricular pressure (S1). However, this criterion cannot be generalized on all real life cases because some medical conditions are characterized by S2 frequency content lower than S1 frequency content [25] like the case of mitral prosthetic valve where the mitral closure is marked by high frequencies which can easily be seen at the beginning of S1. Hence, the importance of time-frequency multi-features approach, especially in a generic module, which can explain the high performance obtained with the proposed TF features vector.

4. Conclusion

In this paper, a robust module for heart sounds segmentation has been developed. The module is divided into three blocks; localization, boundaries detection, and classification of heart sounds (S1 and S2). Several methods are proposed during this study:

- A heart sounds localization method based on the S-transform and Shannon Energy, named SSE, is proposed and evaluated against white additive Gaussian noise.

- A method for boundaries detection named OSSE is proposed. It is based on an optimization process for the energy concentration in the TF domain provided by the S-transform.

- A feature extraction method based on Singular Value Decomposition (SVD) technique to distinguish between S1 and S2 is examined.

Dividing the proposed segmentation method into three separate blocks, enable us to perform a targeted optimization at each level. This confers the feature of robustness to the proposed module, which is a more than necessary element to any auto-diagnosis module applicable in real life conditions.

The goal of this study is to develop a generic tool, suitable for clinical use, robust to noise, and applicable to diverse pathological and normal heart sound signals without any previous information about the subject. The results of the proposed methods evaluated on a database of 80 sounds in which there are 40 pathological cardiac sounds (with several systolic murmurs, prosthetic valves and tachycardia cases) are very promising. More robustness tests against noisy signals, algorithms complexity, facility of implementation and more signals, would contribute to optimize the proposed module.

5. ACKNOWLEDGEMENT

6. REFERENCES

[1] H Liang, S Lukkarinen, I Hartimo, Heart Sound Segmentation Algorithm Based on Heart Sound Envelogram, Helsinki University of Technology, Espoo, 1997, Finland.

[2] A. Moukadem, A. Dieterlen, N. Hueber, C. Brandt, Comparative study of heart sounds localization, Bioelectronics, Biomedical and Bio-inspired Systems SPIE N° 8068A-27, 2011, Prague.

[3] A. Yadollahi and Z. M. Moussavi. A robust method for heart sounds localization using lung sounds entropy. IEEE Trans Biomed Eng,53(3):497–502, 2006.

[4] F. Hasfjord. Heart sound analysis with time dependent fractal dimensions. Master's thesis, Link¨oping, Sweden, February 2004.

[5] Christer Ahlstrom, NonLinear Phonocardiographic Signal Processing, phd thesis, Link¨oping University, SE-581 85 Link¨oping, Sweden, April 2008.

[6] A. Moukadem, A. Dieterlen, N. Hueber, C. Brandt, 15TH NORDIC-BALTIC CONFERENCE ON BIOMEDICAL ENGINEERING AND MEDICAL PHYSICS (NBC 2011) IFMBE Proceedings, 2011, Volume 34, 168-171, DOI: 10.1007/978-3-642-21683-1_42.

[7] Stockwell R.G., Mansinha L., Lowe R.P., Localization of the complex spectrum: the S-transform, IEEE Trans. Sig. Proc. 44 (4) (1996) 998–1001.

[8] H Hadi et al, Classification of Heart Sound based on S-transform and neural networks, Int. Conf. on Information Science, Signal Processing and their Applications, 2010.

[9] G. Livanos, Heart Sound analysis using the S-transform, Computers in Cardiology, 2000, 27:587-590.

[10] H. Liang, S. Lukkarinen, I. Hartimo, "A boundary modification method for heart sound segmentation algorithm", Computers in Cardiology, pp.593-595, 13-16 Sept., 1998.

[11] Samit Ari, Prashant Kumar, and Goutam Saha, On An Algorithm for Boundary Estimation of Commonly Occurring Heart Valve Diseases in Time Domain, India Conference, 2006 Annual IEEE, 10.1109/INDCON.2006.302758.

[12] Robert C. Schlant and R. wayne Alexander (editors), "The Heart Arteries and veins", 8th ed., vol. 1, McGraw Hill Inc., 1994, Ch. 11.

[13] P. D.McFadden, J. G. Cook, and L. M. Forster, "Decomposition of gear vibration signals by the generalized S-transform," Mechanical Systems and Signal Processing, vol. 13, no. 5, pp.691–707, 1999.

[14] C. R. Pinnegar and L. Mansinha, "The S-transform with windows of arbitrary and varying shape," *Geophysics*, vol. 68, no. 1, pp. 381–385, 2003.

[15] Ervin Sejdi´c, Igor Djurovi´c, and Jin Jiang, A Window Width Optimized S-Transform, EURASIP Journal on Advances in Signal Processing, Volume 2008, Article ID 672941, 13 pages doi:10.1155/2008/672941.

[16] K. Gr¨ochenig, Foundations of Time-Frequency Analysis, Birkh¨auser, Boston, Mass, USA, 2001.

[17] LJ. Stankovi´c, "Measure of some time-frequency distributions concentration," *Signal Processing*, vol. 81, no. 3, pp. 621–631, 2001.

[18] Dokur Z., Ölmez T., Feature determination for heart sounds based on divergence analysis, Digital Signal Process. (2007),doi:10.1016/j.dsp. 2007.11.003.

[19] Yan Z. et al., The moment segmentation analysis of heart sound pattern, Comput. Methods Programs Biomed.(2009),doi:10.1016/j.cmppb.2009.09.008.

[20] D. Kumar et al., Detection of S1 and S2 Heart Sounds by High Frequency Signatures," in Proc.28 the Annual International Conference of the IEEE EMBS, 2006, pp. 1410-1416.

[21] Kumar, D. et al, Wavelet Transform and Simplicity based Heart Murmur Segmentation, Computers in Cardiology, vol. 33, 2006.

[22] Marinovic M. and Eichmann G., Feature extraction and pattern classification in space-spatial frequency domain. In SPIE Intelligent Robots and Computer Vision, pages 19–25, 1985.

[23] Hassanpour H., Mesbah M., Boashash B.. Time-frequency feature extraction of newborn EEG seizure using svd-based techniques. Eurasip J Appl Sig Proc, 16:2544-2554, 2004.

[24] Ahlstrom C., Hult P., Rask P., Karlsson J-E, Nylander E, Dahlstr¨om U, Ask P: Feature Extraction for Systolic Heart Murmur Classification. Annals ofBiomedical Engineering. 2006. 34(11):1666-1677.

[25] A. Moukadem, A. Dieterlen, N. Hueber, C. Brandt, Study of two feature extraction methods to distinguish between the first and the second heart sounds, International Conference on Bio-inspired Systems and Signal Processing, BIOSIGNALS 2012.