

Phonocardiogram Signal Processing Module for Auto-Diagnosis and Telemedicine Applications

Ali Moukadem, Alain Dieterlen and Christian Brandt

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/48447>

1. Introduction

The advancement of technology has paved the way for signal processing methods to be implemented and applied in many simple tools useful in everyday life. This is most notable in the medical technology field where contributions involving the intelligent applications have boosted the quality of diagnosis. Proposing an objective signal processing methods able to extract relevant information from biosignals is a great challenge in telemedicine and auto-diagnosis fields.

For the cardiac system, many signals can be treated and monitored; ElectroCardioGram (ECG), PhonoCardioGram (PCG), Echo/Doppler and pressure monitor, see Figure 1.

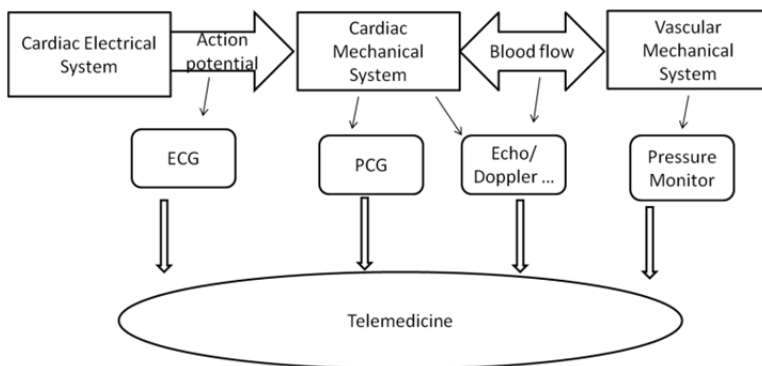


Figure 1. The cardiac activity with different measurable signals [1].

The interest of this book chapter is the PCG signal. PCG and auscultation are noninvasive, low-cost and accurate for diagnosing some heart diseases.

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.

The analysis of the cardiac sounds, solely based on the human ear, remains insufficient for a reliable diagnosis of cardiac pathologies, and for a clinician to obtain all the qualitative and quantitative information about cardiac activity especially in the field of time intervals.

Information, such as the temporal localization of the heart sounds, the number of their internal components, their frequency content, and the significance of diastolic and systolic murmurs, could all 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, different approaches could be considered for improve the electronic stethoscope:

Tool with embedded autonomous analysis, simple for home use by the general public for the purpose of auto-diagnosis, monitoring and warning in case of necessity.

Tool with sophisticated analysis (coupled to a PC, Bluetooth link) for the use of professionals in order to make an in-depth medical diagnosis and to train the medical students.

Whatever the approach, one of the first and most important phases in the analysis of heart sounds, is the segmentation of heart sounds. Heart sound segmentation partitions the PCG signals into cardiac cycles and further into S1 (first heart sound), systole, S2 (second heart sound) and diastole.

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: Shannon energy [2], Hilbert Transform [3], high order statistics [1], hidden Markov model [4] ...

In this chapter we present a new module for heart sounds segmentation based on time-frequency analysis (S-Transform). The goal of this study is to develop a generic tool, suitable for clinical and home monitoring use, robust to noise, and applicable to diverse pathological and normal heart sound signals without the necessity of any previous information about the subject. 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 are evaluated based on a database of 80 subjects (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 chapter 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 Sections 5 and 6 give the future research and 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. 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. In parallel, 2 prototypes have dedicated to the MARS500 project promoted by ESA, in order to collect signals form 6 volunteers (astronauts). The use of prototype electronic stethoscopes by different cardiologists makes the database rich in terms of qualitative diversity of collected sounds, which in turn makes the heart sounds localization more realistic.

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 subjects, including 40 cardiac pathologies sounds which contain different systolic murmurs. Each subject corresponds to one recording sound. The length of each sound is 8 seconds.

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:

$$x_{norm}(t) = \frac{x(t)}{|\max(x(t))|} \quad (1)$$

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 [1]. 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.

3.3. SRBF localization method

We proposed the RBF method as a transformation to emphasize heart sounds and it was shown to have a good performance on low level noise signals [5]. 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, we proposed a method for heart sounds localization named SRBF [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. Compared with other existing methods for heart sounds localization, SRBF was shown to have a significant enhancement in term of sensitivity and positive predictive value and the robustness of this method was shown against additive white Gaussian noise.

We will briefly explain the different steps of the SRBF method:

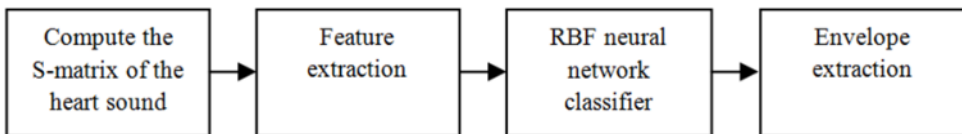


Figure 2. Block Diagram of SRBF Method

1. The S-Transform of the heart sound is calculated. A frequency range of 0-100 Hz was used to cover the main frequency band of S1 and S2 and to avoid murmurs which have in general a spectral energy above the frequency of 100 Hz [7].
2. A sliding window of 50 ms (so 100 samples) was operated on the S-matrix and an overlap of 75% was chosen. The feature extraction is done by applying some standard statistical techniques and transformations like Root Mean Square (RMS), the maximum and the average of each column of the S-matrix. Each array (100 samples) was divided into 5 segments and the mean of calculated features of each segment was calculated and

taken as input to the classifier. So for each step we have a 100 by 100 matrix which gives 15 descriptors.

3. A RBF neural network classifier is used and trained on two heart sounds samples (S1 and S2) and two no heart sound samples (systole, diastole) selected randomly from the database. The target is fixed to 1 for S1 or S2 and 0 for the other components. So the envelope of the signal is constructed by the output of the RBF neural network.

3.4. 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) 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 3).

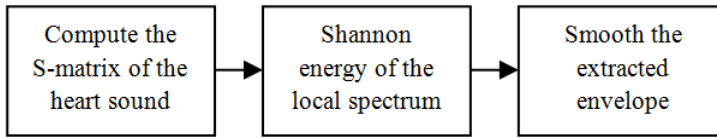


Figure 3. Block Diagram of SSE Method

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

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

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

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

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

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

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

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

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 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.

3.5. 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.

3.5.1. Overview of existing methods

Different boundaries detection algorithms exist in the literature, in [2] 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. The same authors propose another algorithm that employs the STFT (Short Time Fourier Transform) to explore the time-frequency domain of the signal [9]. 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 [10] 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.

3.5.2. The OSSE algorithm

In this chapter, 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 4).

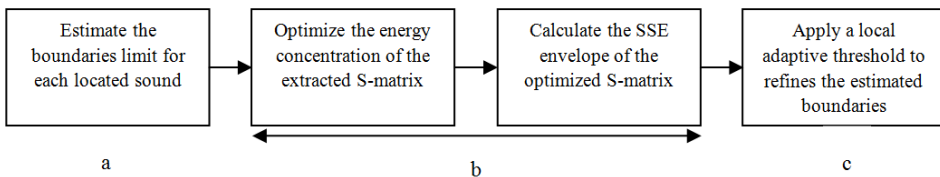


Figure 4. The block diagram of the OSSE Method

a. Estimate the boundaries limit

The boundaries limits are estimated basing on the fact that the maximum duration of S1 and S2 is 150 ms [11]. 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 TF representation of the S-transform[12-14]. The main study in the literature interested to optimize the energy concentration in the TF domain was in [14]. That is, to minimize the spread of the energy beyond the actual signal components. As it 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 [15].

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, 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 is 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 [14] and we investigate another parameter named α (see equation 6). Both of them control the Gaussian window width:

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

We note here that in this study 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 very slowly varying frequency components. The second method calculates the time-varying parameter 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 [16]. For each α (or p) from a given set, the CM (α) can be expressed by [14]:

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

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

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

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 10 % of the maximum value of the SSE envelope is applied to refine the estimated boundaries.

3.6. 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) [1,17-18]. 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 5).

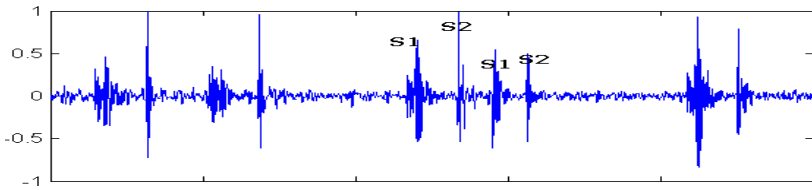


Figure 5. Example of an arrhythmic subject.

Consequently with the objective of development of a robust generic module for heart sound segmentation, we present in this chapter two feature extraction methods based on the Singular Value Decomposition (SVD) technique applied on the S-matrix, to classify S1 and S2. We investigate also, the ability of a new individual features based on the width of the optimized Gaussian window of the S-Transform, to discriminate between S1 and S2.

3.6.1. Feature extraction based on the S-Transform

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 [19] authors extracted the eigenvalues of the time-frequency matrix. In [20] authors extended the method to also incorporate information from the eigenvectors to classify EEG seizures. In [21] 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 H_i is decomposed by the SVD technique as follows:

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

Where $U(M \times M)$ and $V(N \times N)$ are orthonormal matrices so their squared elements can be considered as density function[20], 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 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.

3.6.2. Feature extraction using the EMD

In the last few years, the Empirical Mode Decomposition (EMD) has been applied in many fields one of which the biomedical signal analysis, like the emotion classification in natural speech [22], analysis of gastroesophageal information [23]. EMD has been applied to a simulated heart sounds in [24] authors show that EMD provides clear information about the components of S1 and S2 and their instantaneous frequency behaviour. In [25] authors presented a feature analysis approach of heart sound based on the improved Hilbert-Huang Transform, and applied the improved HHT by Hilbert spectrum analysis of various cases of heart sounds. In this study, a new feature extraction method based on EMD technique and Shannon energy is proposed for S1 and S2 classification.

As an alternative to the binomial TF transforms, EMD performs a multi-resolution analysis of non-stationary and nonlinear signals without the use of kernels or mother waveforms. To calculate the Intrinsic Mode Functions (IMFs), the local maxima and minima of extracted heart sound $H_i(t)$ are calculated. They are interpolated by using the cubic spline curves which generates the upper and lower envelopes, respectively. Then the mean contour $m_1(t)$ is calculated, and the first component $h_1(t)$ is given as follows:

$$h_1(t) = H_i(t) - m_1(t) \tag{10}$$

Now, h_1 has to be refined by a sifting process. In the second sifting iteration we obtain:

$$h_{11}(t) = h_1(t) - m_{11}(t) \tag{11}$$

Where m_{11} is an average contour between the upper and lower envelopes of h_1 . This operation is repeated k times until h_{1k} can be considered as zero-mean according to some stopping criterion (Rilling et al., 2003). The first intrinsic mode function $IMF_1(t)$ is given as:

$$IMF_1(t) = h_{1(k-1)}(t) - m_{1k}(t) \tag{12}$$

$IMF_1(t)$ should contain the finest scale or the shortest period component of the signal. The residue signal $r_1(t)$ is given by:

$$r_1(t) = H_i(t) - IMF_1(t) \tag{13}$$

Considering r_1 as a new signal the sifting process explained below is repeated to obtain the second $IMF_2(t)$. Similarly, a series of intrinsic mode functions are obtained and the final residue $r_n(t)$ is calculated. The stop criterion is when $r_n(t)$ becomes a monotonic function.

The initial signal $H_i(t)$ can be reconstructed as follows:

$$H_i(t) = \sum_{j=1}^n IMF_j(t) + r_n \tag{14}$$

For each IMF vector, the Shannon Energy is calculated as:

$$SE_i = -\sum_{k=1}^N IMF_i^2(k) \cdot \log(IMF_i^2(k)) \tag{15}$$

Where $i=1, \dots, 4$ and N is the number of samples of IMF_i the Shannon energy is smoothed by using a median filter, and the feature vector is obtained by applying the same SVD approach used in section 2.5.1 at each calculated IMF (Figure 6). For each extracted heart sound the first four IMF is calculated. The others IMF don't contain relevant information about S1 and S2. Five feature vectors obtained by this method are tested in the classification process; FV1 (that correspond to IMF_1 signal), FV2, FV3, FV4 and FV (that correspond to the average of calculated FVs). The length of each vector is 10.

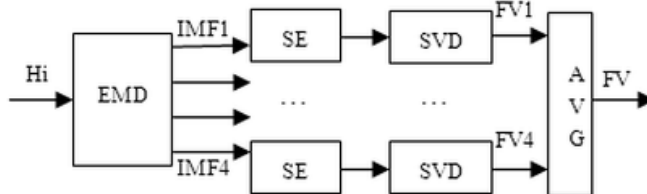


Figure 6. Feature vector (FV) of Heart Sounds (Hi) extracted using EMD and Shannon Energy (SE) before applying the SVD technique.

3.6.3. New individual features

The parameters α and p used to optimize the width of the Gaussian window of the S-Transform, are tested as a new individual features to discriminate between S1 and S2. It is known from a physiological point of view, that S1 is more complicated than S2 [26]. However, S2 in general contain higher frequency than S1. These physiological differences will necessarily lead to different time-frequency content behavior which we will aim to reveal with α and p parameters. Figure 7 shows a S1 and S2 signals examples with the corresponding optimized S-transform obtained with $\alpha=0.8$ and 0.5, respectively.

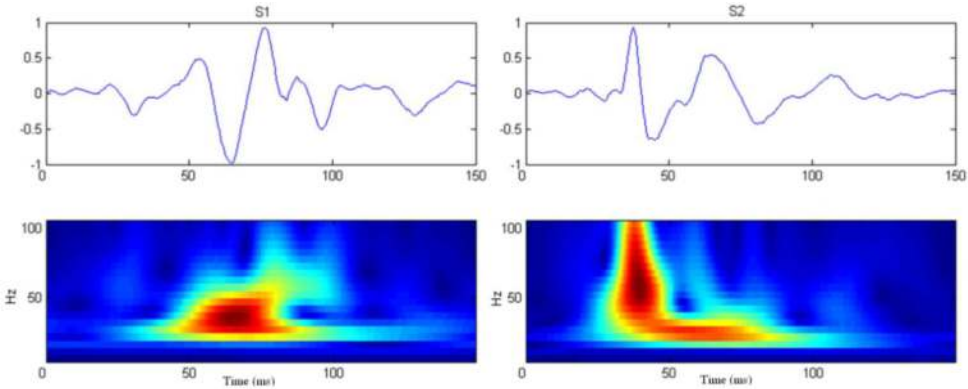


Figure 7. S1 and S2 signals (top), Optimized S-transform obtained with $\alpha=0.8$ for S1 and $\alpha=0.5$ for S2 (bottom).

4. Results and discussion

4.1. 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:

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (16)$$

And positive predictive value:

$$PPV = \frac{TP}{TP + FP} \quad (17)$$

A sound is true positive (TP) if it is correctly located, all others detected sounds are considered as false positive (FP) and all missed sounds are considered as false negative (FN).

Results in Table 1 show that SRBF method reaches a higher PPV (98%) than the SSE method for the clinical signals without any additive noise. However, SSE reaches a higher sensitivity

(96%) than the SRBF method (92%). 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 8 shows the envelopes extracted by the SSE and the SRBF method that correspond to a pathologic sound with a systolic murmur. Figure 9 shows the robustness of each method against white additive noise.

Method	Sensitivity	PPV	Sensitivity (Noise)	PPV (Noise)
SRBF	92%	98%	91%	93%
SSE	96%	95%	93%	94%

Table 1. Sensitivity and Positive Predictive Values for the SRBF and SSE methods applied on the clinical sounds set without and with additive Gaussian noise.

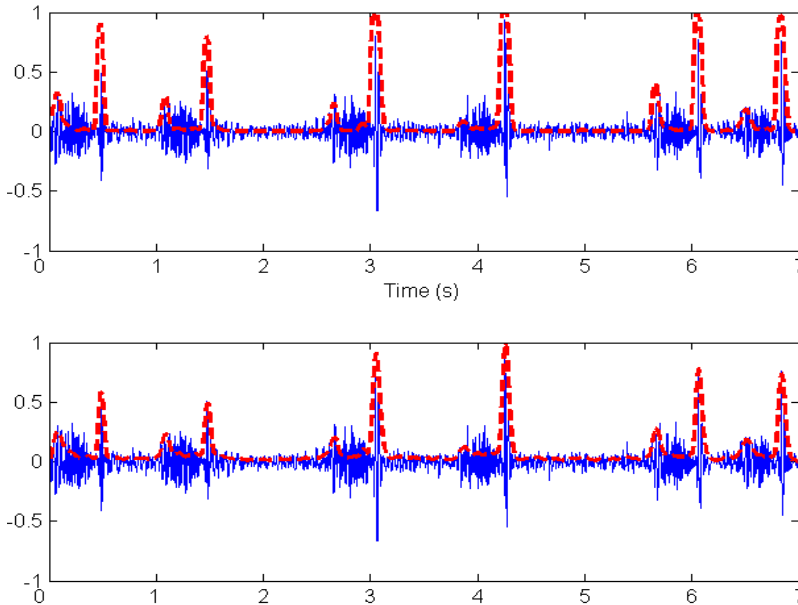


Figure 8. Envelope extraction (dashed lines) for a signal with systolic murmur, (top) SRBF envelope, (bottom) SSE envelope.

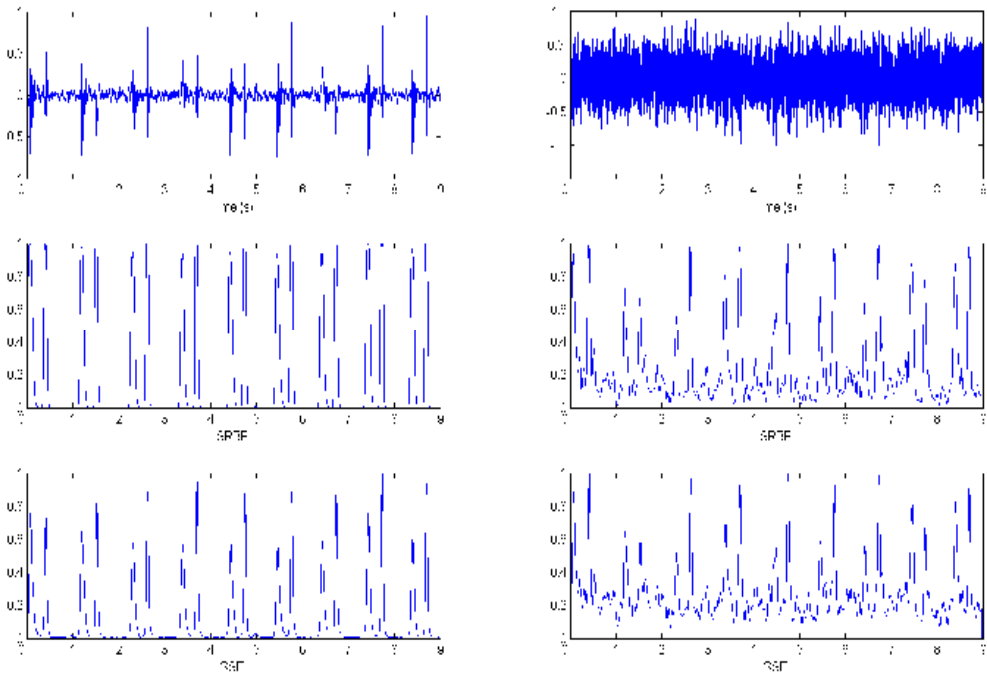


Figure 9. (top) Envelope extraction for two normal PCG signal without and with additive Gaussian noise, (middle) their SRBF envelopes, (bottom) their SSE envelopes.

4.2. Boundaries detection

The performance measure against each parameter is compared (Table 2). The values of α and p are chosen from a set; $0 < \alpha < 2$, $0 < p < 2$, with a step of 0.1; so twenty values as total for each variable.

Heart Sounds	Optimal α	CM(α)	Optimal p	CM(p)	CM($\alpha=1, p=1$)
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

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.

The optimal α is reached when CM(α) is maximized, and the optimal p is reached when CM (p) is maximized. Results from Table 2 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 $\alpha=1$ 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 $\alpha < 1$ and $p > 1$. This is can be explained by the fact that when $\alpha < 1$ and $p > 1$, the Gaussian window of the S-transform is narrower (Figure 10), 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).

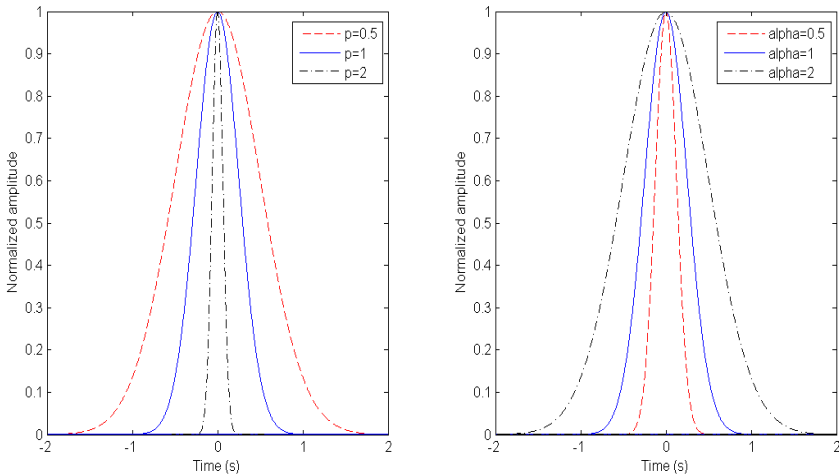


Figure 10. Normalized Gaussian window for different values of p (left) and for different values of α (right).

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	

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

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. 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 11 shows the boundaries detection results, with

and without optimization of the S-transform, applied on a S2 example and figure 12 shows the OSSE results applied on the entire heart sounds (normal and pathologic).

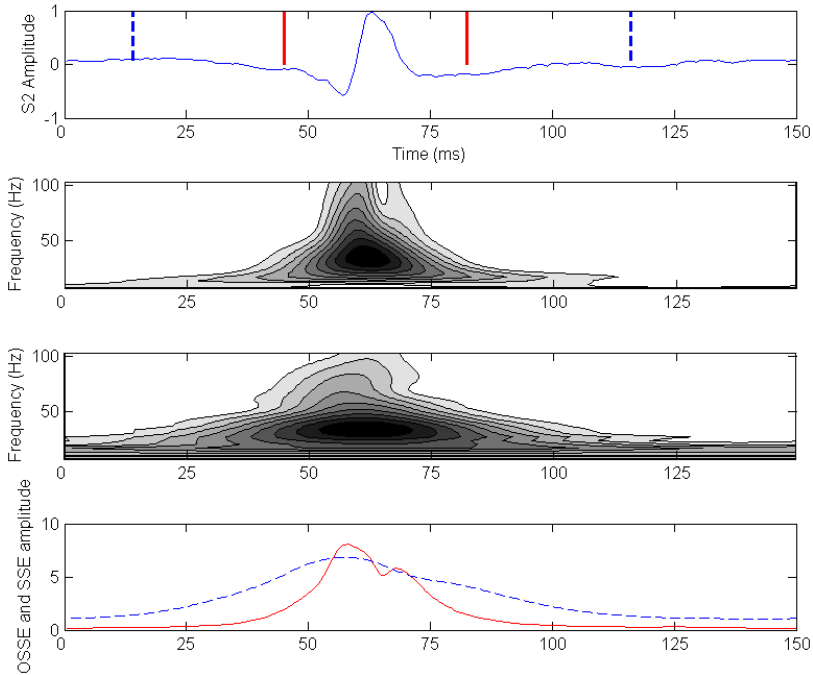


Figure 11. (top) S2 signal with two detected boundaries calculated by the optimized S-transform and the standard S-transform (dashed line), S-transform with the optimum value $\alpha=0.5$ ($p=1$), standard S-transform with $\alpha=1$ ($p=1$), (bottom) SSE envelope for the optimized S-transform and standard S-transform (dashed line).

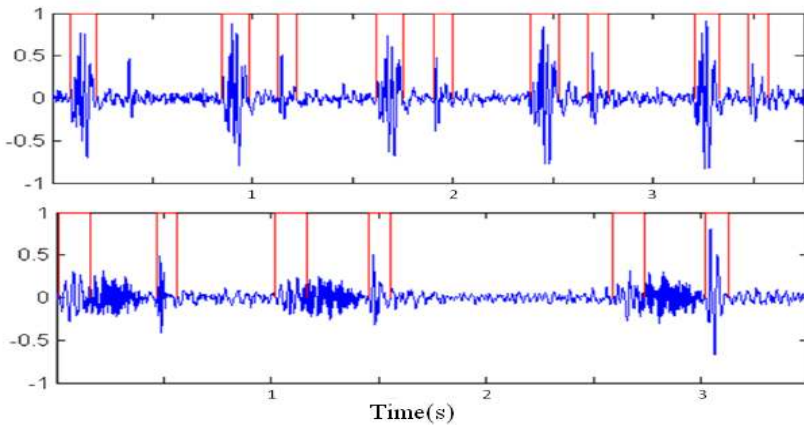


Figure 12. OSSE method applied on a normal heart sound (top) and pathologic heart sound (bottom).

4.3. Feature extraction for S1 and S2 classification

4.3.1. Evaluating the feature vectors obtained by the SVD technique

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-Nearest Neighbor (KNN) classifier is used to evaluate the performance of the four feature vectors obtained by the two methods 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 time domain feature vector reaches 92% classification rate, however, the frequency feature vector reaches 85% classification rate (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. For the EMD based method, the FV feature vector reaches a high classification rate with 94% sensitivity and 97% specificity (Table4).

KNN	T-Features	F-Features	SV Features	TF Features	FV1	FV2	FV3	FV4	FV
Sensitivity	92%	81%	60%	95%	88%	81%	82%	65%	94%
Specificity	92%	88%	65%	97%	91%	97%	94%	95%	97%

Table 4. Sensitivity and specificity for the nine extracted feature vectors evaluated by a KNN classifier.

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. Hence, the importance of time-frequency and multi-resolution based features approach, especially in a generic module, which can explain the high performance obtained with the TF and FV features vectors.

4.3.2. Evaluating α and p to discriminate S1 and S2

The parameters used in the optimization process (section 3.3.2) to determine the boundaries of each extracted sound S_i (S1 or S2) are averaged across available extracted sounds from each subject. So from each subject in the database, we obtain one S1 feature (α or p) and one S2 feature (α or p).

The main objective is to investigate the ability of these features to discriminate between S1 and S2. The probability that the two groups (S1 and S2) comes from distributions with different medians is calculated for each feature (α and p) by the Mann-Whitney-U-test ($p < 0.005$). The receiver Operating Characteristic Curve (ROC) is also calculated for each feature and the Areas under the ROC Curve (AUC) are showed in figure 13.

The Results are presented in Table 5. Significant differences between the groups, with 95% confidence are found for both features α and p .

Feature	p-value	AUC	Sensitivity	Specificity
α	<0.0001	0.83	0.79	0.72
p	0.0047	0.64	0.609	0.671

Table 5. Significant values (U-test), AUC values, sensitivity and specificity for the parameters α and p when used to distinguished between S1 and S2.

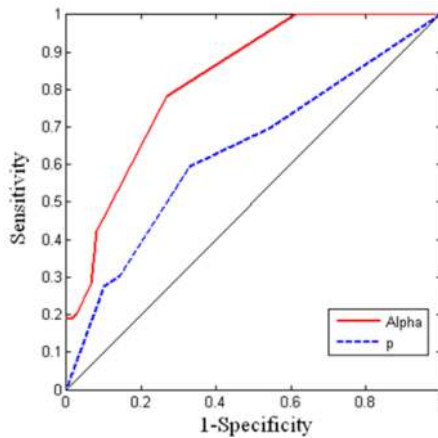


Figure 13. ROC curves for α and p parameters.

The classification results are promising for the parameter α (AUC =0.83). This is very interesting since this parameter was also used to refine the boundaries detection of S1 and S2. However, the results of the parameter p are significantly lower than the results of α (AUC =0.64). This gives a primary idea about the sensitivity of each parameter against the clinical signals. Further measures and tests should verify or deny this hypothesis.

5. Future research

5.1. Classification of heart sounds

A new time-frequency based feature is proposed and validated to distinguish with S1 and S2 (Section 4.3.2). Another parameter can be tested by applying another windows type at the S-transform like the arbitrary and varying shape window [13]. A combination of several features can also be used to classify S1 and S2 more accurately. This can be performed by

combining the α parameter with the TF_Features vector (see section 4.3.1). Then a feature selection algorithm becomes necessary to select the most accurate features.

On another hand, the classification of normal and pathological heart sounds is the final objective of any heart sounds auto-diagnosis framework. The classification rate will depend first on the segmentation results, which was the main objective of this book chapter. Then classic steps of feature extraction, feature selection, designing and testing classification systems, will be needed to complete the classification process

5.2. Real time application

One of the objectives of this study is to develop an auto diagnosis for various situations encountered in cardiology in real time. However, the S-Transform that can be considered as the heart of the proposed segmentation framework, suffers from a high computational burden. The implementation of a fast S-Transform algorithm on FPGA or GPU card will be necessary.

5.3. Sociological and psychological aspect

Introducing a smart stethoscope as a monitoring tool for home use, involves new problems related to sociological and psychological aspect of the user (patient). A smart stethoscope is a tool to facilitate the diagnosis process and to make it more objective and it will never replace the cardiologist and other advanced techniques of Cardiology. This should be taken into consideration in the deployment process in a telemedicine framework for example. The ergonomic aspect of the measuring instrument, the way to display the data and to transmit it, will be more than necessary elements to any future tool, simple for home use by the general public for the purpose of auto-diagnosis, monitoring and warning in case of necessity.

6. Conclusion

In this book chapter, a robust module for heart sounds segmentation has been proposed. 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 methods based on Singular Value Decomposition (SVD) technique to distinguish between S1 and S2 are examined. The parameters used in the time-frequency optimization process to determine the boundaries of each extracted sound are also investigated and validated as discriminative features between S1 and S2.

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 main objective of this study is to present a robust and generic PCG segmentation method useful in real life conditions (clinical use, home care, professional use ...). The methods in the proposed framework are evaluated on a real data (80 subjects) with different noise levels and they are validated by the cardiologist.

More robustness tests against noisy signals, algorithms complexity, facility of implementation and more signals, would contribute to optimize the proposed module.

Author details

Ali Moukadem and Alain Dieterlen

MIPS Laboratory, University of Haute Alsace, Mulhouse, France

Ali Moukadem and Christian Brandt

University Hospital of Strasbourg, CIC, Inserm, Strasbourg, France

7. References

- [1] Christer Ahlstrom, NonLinear Phonocardiographic Signal Processing thesis, Linköping University, SE-581 85 Linköping, Sweden, April 2008.
- [2] H Liang, S Lukkarinen, I Hartimo, Heart Sound Segmentation Algorithm Based on Heart Sound Envelopgram, Helsinki University of Technology, Espoo, Finland.
- [3] Samjin Choi, Zhongwei Jiang, Compariason of envelope extraction algorithms for cardiac sound signal segmentation, Micro-Mechatronics Laboratory, Yamaguchi University, 2006, Japan.
- [4] Schmidt, S.E., Holst-Hansen, C., Graff, C., Toft, E., Struijk, J.J.Segmentation of heart sound recordings by a duration-dependent hidden Markov model (2010) *Physiological Measurement*, 31 (4), pp. 513-529.
- [5] A. Moukadem, A. Dieterlen, N. Hueber, C. Brandt, Comparative study of heart sounds localization, *Bioelectronics, Biomedical and Bio-inspired Systems SPIE N° 8068A-27*, Prague.
- [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] Amir A. Sepheri, et al., A nouvel method for pediatric heart sound segmentation without using the ECG, *Comput. Methods Programes Biomed.* (2009), doi:10.1016/j.cmpb.2009.10.006
- [8] Stockwell R.G., Mansinha L., Lowe R.P., Localization of the com-plex spectrum: the S-transform, *IEEE Trans. Sig. Proc.* 44 (4) (1996) 998–1001.
- [9] 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.
- [10] 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

- [11] Robert C. Schlant and R. Wayne Alexander (editors), "The Heart Arteries and Veins", 8th ed., vol. 1, McGraw Hill Inc., 1994, Ch. 11.
- [12] 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.
- [13] 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.
- [14] Ervin Sejdić, Igor Djurović, 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
- [15] K. Gröchenig, *Foundations of Time-Frequency Analysis*, Birkhäuser, Boston, Mass, USA, 2001.
- [16] Lj. Stanković, "Measure of some time-frequency distributions concentration," *Signal Processing*, vol. 81, no. 3, pp. 621–631, 2001.
- [17] 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.
- [18] Yan Z. et al., The moment segmentation analysis of heart sound pattern, *Comput. Methods Programs Biomed.* (2009), doi:10.1016/j.cmpbb.2009.09.008.
- [19] 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.
- [20] 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.
- [21] Ahlstrom C., Hult P., Rask P., Karlsson J-E, Nylander E, Dahlström U, Ask P: Feature Extraction for Systolic Heart Murmur Classification. *Annals of Biomedical Engineering*. 2006. 34(11):1666-1677.
- [22] He L., Lech M., Maddage C. N., Allen N., Study of empirical mode decomposition and spectral analysis for stress and emotion classification in natural speech, *Biomedical signal processing and control* 6 (2011) 139-146.
- [23] Liang H., Lin Z., McCallum R., Application of the Empirical Mode Decomposition to the Analysis of Esophageal Manometric Data in Gastroesophageal Reflux Disease, *IEEE T. Biomed. Eng.*, 52(10), 1692 – 1701, 2005.
- [24] Charleston-Villalobos S., Aljama-Corrales, A. T., González-Camarena R., Analysis of Simulated Heart Sounds by Intrinsic Mode Functions, *Proceedings of the 28th IEEE EMBS Annual International Conference* New York City, USA, Aug 30-Sept 3, 2006.
- [25] Liu L., Wang H., Wang Y., Tao T., Wu X., Feature Analysis of Heart Sound Based on the Improved Hilbert-Huang Transform, *3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT)*, 2010.
- [26] 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*.