Automatic Heart Sound Analysis Module Based on Stockwell Transform

Applied on Auto-Diagnosis and Telemedicine Applications

Ali Moukadem, Alain Dieterlen
MIPS Laboratory
University of Haute Alsace
Mulhouse, FRANCE
ali.moukadem@uha.fr, alain.dieterlen@uha.fr

Chrisitian Brandt
University Hospital of Strasbourg
CIC, Inserm
christian.brandt@wanadoo.fr

***Abstract*—The aim of this paper is to present an automatic heart sound analysis method which can be used for auto-diagnosis and telemedicine applications. One of the first and most important phases in the analysis of heart sounds, is the segmentation process which partitions the sound into cardiac cycles and further into S1 (first heart sound), systole, S2 (second heart sound) and diastole. The heart sounds (S1 and S2) are localized by applying the Shannon energy of the local spectrum calculated by the S-Transform. Then, to distinguish between the first and the second heart sound, a feature extraction method based on S-Transform is also presented. The methods are evaluated on a dataset of 80 subjects, including 40 patients with cardiac pathologies sounds.**

Keywords-component; Time-Frequency, S-Transform, Heart Sounds, Auto-Diagnosis.

I. 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.

The interest of this paper 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 [1], Hilbert Transform [2], high order statistics [3], hidden Markov model [4], etc.

In this study, we present a new module for heart sounds analysis that aims to segment automatically the heart sound. 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 module can be divided into two main blocks: localization of 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

 ISBN: 978-1-61208-252-3

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 module (localization and classification). The results and discussion are presented in Section 4 and Section 5 gives the future research and the conclusion.

II. 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 is 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, two 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. The length of each sound is 8 seconds.

III. METHOD

A. 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 in the 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:

$$x_{norm}(t) = x(t) / \max(x(t)) \tag{1}$$

B. 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 [3]. The transformation try to maximize the distance between the heart sounds and the background noise, and the result is smoothed and tresholded 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. The boundaries of the heart sounds are determined by the first local minima before and after the located sound.

The results were visually inspected by a cardiologist and erroneously extracted heart sounds were excluded from the study.

1) 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:

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.

2) 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 1).

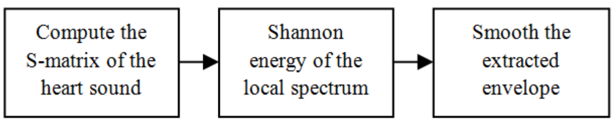

Figure 1. Block Diagram of SSE Method

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

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

Where the window function *w(τ-t)* is chosen as:

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

And *σ(f)* is a function of frequency as:

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

The proposed method calculates the Shannon energy of each the local spectrum as follows:

$$SSE(x_i) = -\int_{-\infty}^{+\infty} S(\tau, f)^2 \log\left(S(\tau, f)^2\right) df \qquad (5)$$

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 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.

C. *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) [1,9,10]. 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 2).

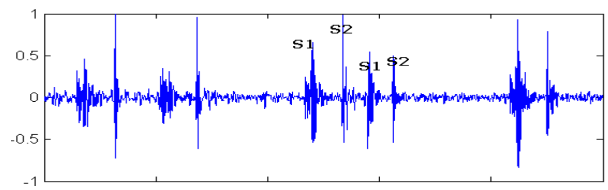

Figure 2. Example of an arrhythmic subject.

Consequently with the objective of development of a robust generic module for heart sound segmentation, we present in this paper a feature extraction methods based on the Singular Value Decomposition (SVD) technique applied on the S-matrix, to classify S1 and S2.

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 [11] authors extracted the eigenvalues of the time-frequency matrix. In [12] authors extended the method to also incorporate information from the eigenvectors to classify EEG seizures. In [13] 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 Si of the extracted heart sound Hi is decomposed by the SVD technique as follows:

$$S_i = UDV^T \qquad (6)$$

Where *U(M×M)* and *V(N×N)* are orthonormal matrices so their squared elements can be considered as density function[12], and *D(M×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.

IV. Results and Discussion

A. 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:

$$Sensitivity = TP/(TP + FN) \qquad (7)$$

And positive predictive value:

$$PPV = TP/(TP + FP) \qquad (8)$$

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).

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

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

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 3 shows the envelopes extracted by the SSE and the SRBF method that correspond to a pathologic sound with a systolic murmur. Figure 4 shows the robustness of each method against white additive noise.

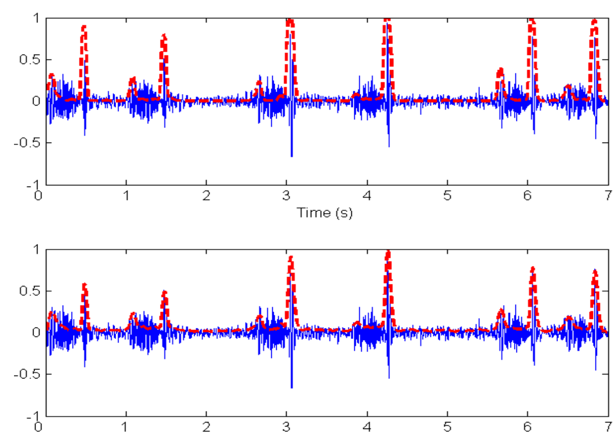

Figure 3. Envelope extraction (dashed lines) for a signal with systolic murmur, (top) SRBF envelope, (bottom) SSE envelope

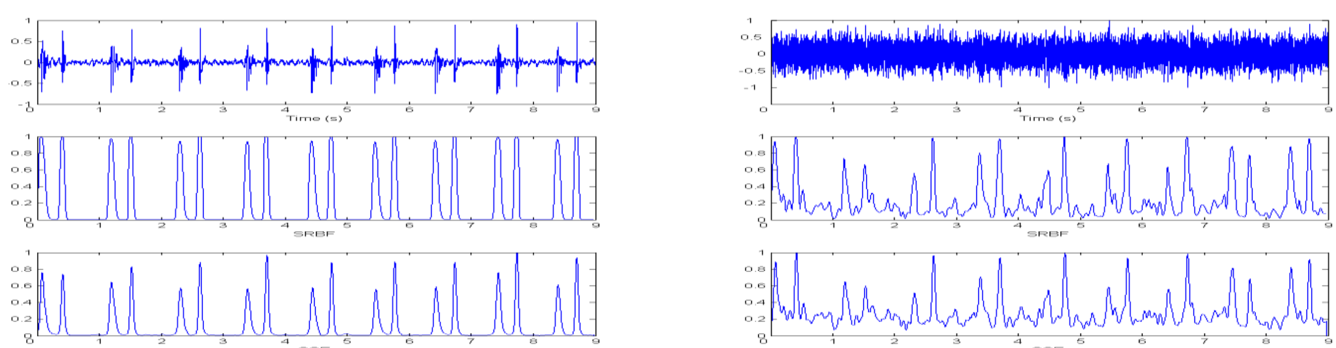

Figure 4. (top) Envelope extraction for two normal PCG signal without and with additive Gaussian noise, (middle) their SRBF envelopes, (bottom) their SSE envelopes.

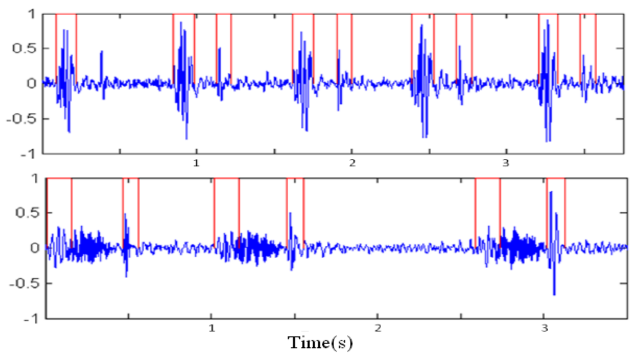

Figure 5. Segmentation module applied on normal heart sound (top) and pathological heart sound (bottom).

B. *Classification of S1and S2*

The heart sounds are segmented (Figure 5) and 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 Si (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.

TABLE II. SENSITIVITY AND SPECIFICITY FOR THE FIVE EXTRACTED FEATURE VECTORS EVALUATED BY A KNN CLASSIFIER.

KNN	**T-Features**	**F-Features**	**SV Features**	**TF Features**
Specificity	92%	81%	60%	**95%**
Sensitivity	92%	88%	65%	**97%**

A 3-Neirest 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 (Table 2).

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 features approach, especially in a generic module, which can explain the high performance obtained with the TF and FV features vectors.

V. FUTURE RESEARCH AND CONCLUSION

A. *Classification of heart sounds*

A time-frequency based features is proposed and validated to distinguish with S1 and S2. However, 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 study. Then classic steps of feature extraction, feature selection, designing and testing classification systems, will be needed to complete the classification process

B. *Real time application*

The main objective of this study was 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.

C. *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.

D. *Conclusion*

The proposed module is divided into two blocks: localization, 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 is proposed and evaluated against white additive Gaussian noise.
- A feature extraction method based on Singular Value Decomposition (SVD) technique to distinguish between S1 and S2 are examined.

The main objective of this study was 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 were evaluated on a real data (80 subjects) with different noise levels and they were validated by the cardiologist. More robustness tests against noisy signals, algorithms complexity, facility of

 ISBN: 978-1-61208-252-3

implementation and more signals, would contribute to optimize the proposed module.

ACKNOWLEDGMENT

The authors would like to thank the financial support from the French telemedicine project, E-Care.

REFERENCES

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

[2] Samjin Choi and Zhongwei Jiang, Compariason of envelope extraction algorithms for cardiac sound signal segmentation, Micro-Mechatronics Laboratory, Yamaguchi University, 2006, Japan.

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

[4] Schmidt, S.E., Holst-Hansen, C., Graff, C., Toft, E. and 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 and 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 and 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. and Lowe R.P., Localization of the com-plex spectrum: the S-transform, IEEE Trans. Sig. Proc. 44 (4) (1996) 998–1001.

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

[10] 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.

[11] 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.

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

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