

Modulation Filtering for Noise Detection in Heart Sound Signals

J. P. Ramos and P. Carvalho and R. P. Paiva and J. Henriques

Abstract—Cardiac auscultation has proven to be an excellent diagnostic tool. Heart sound processing algorithms are not completely robust in the presence of noise, requiring clean segments of heart sounds to extract reliable diagnostic features. This paper presents a new approach to detect transient noises mixed with heart sound. The algorithm explores a single channel source separation algorithm and evaluates the non-stationary separated signals. It has the potential to be applied in real-time. Using a database of heart sounds acquired in real-life scenario, the method showed a sensitivity and a specificity of 93.6% and 92.3%, respectively.

I. INTRODUCTION

Cardiovascular diseases are Europe's leading health problem, causing 42% of all deaths in the European Union (EU). Furthermore, cardiovascular diseases are estimated to cost the EU economy €192 billion a year [1]. Prevention is considered the best strategy to control such expenses. Aside from fostering a healthy lifestyle, monitoring and managing the progression of the disease helps avoiding life-critical situations as well as expensive treatments. Moreover, the diagnostic tool chosen to assess the heart condition is also of great importance. A careless selection of the tool may lead to unnecessary and inefficient use of resources and one ought to seek sources of diagnosis less expensive and equally accurate [2]. In this context, heart sounds are a promising bio-signal to deploy applications for continuous cardiac disease (tele-)management. It has been shown to be highly sensitive and specific to several major physiological events within the heart [3]. Unfortunately, the heart sound is very prone to noise interference, either internal (e.g. physiological noises) or external (e.g. surrounding noises). Noise sources exhibit very complex time-frequency characteristics. Furthermore, due to time-varying propagation routes inside the thorax, these sources tend to mix in highly complex ways with the heart sound. Hence, reducing noise contamination without affecting heart sound's main components has proven to be highly challenging and is an unsolved problem using single channel acquisition schemes. Nevertheless, these noises impair further processing modules and, in order to get a clearer diagnostic assessment, they should be identified and removed from further processing.

Several noise detection methods can be found in the literature. In [4] heart sound segments are decomposed

using a Morlet wavelet and their coefficients are fed to an artificial neural network. This method comprises a wider spectrum of classes, such as S1, S2 and murmur classification, and claims promising preliminary results. However, no actual performance results are stated. Moreover, further work from the same authors [5] only focus on heart sound main components' classification. In [6] the noise detection problem is tackled in two phases: the first phase explores the periodicity, both in time and time-frequency domain, of the heart sound signal in order to find a non-contaminated reference sound; in the second phase the reference sound is correlated with the signal using temporal and spectral similarity criteria to check for noise contamination. This method achieves not only good performance results, but also low computation times, once the reference sound has been identified. Nevertheless, the necessity to find a reference sound for further processing is a drawback in the method's applicability, since it is computational demanding and usually adds over 10 seconds to the auscultation protocol. This might be a serious impairment for daily clinical applications.

In this paper we suggest a non-cardiac sound detection method from heart sound. This approach partly explores the separation algorithm proposed by [7] in which the authors attempt to separate the heart sound - considered a quasi-stationary signal - from the lung sound (a stationary signal). Rather than completely reconstructing the signal (signal reconstruction introduces significant morphological distortions in the signal, hence affecting it's diagnostic value), we assume that any transient noise will affect the stationarity property of the non-heart signal and, therefore, we are able to detect noise disturbances in a signal. Having the features that discriminate such disturbances, one can then classify the segment using a standard classifier structure.

The paper is organised as follows. Section 2 outlines the algorithm proposed. In section 3, results are presented and discussed. Lastly, conclusions are drawn in section 4.

II. METHOD

The method found in [7] explores an alternate spectro-temporal representation, designated as modulation frequency domain. This representation is obtained by frequency decomposition of the temporal trajectories of short-term spectral components and provides information regarding the signal spectral components' rate of change. Moreover, the authors have suggested that modulation spectral content of the heart sound typically falls between approximately $2 - 20Hz$ and that lung sounds' modulation spectral contents resides at low frequencies ($< 2Hz$). The separation can thus proceed with

This work was supported in part by SoundForLife (PTDC/EIA/68620/2006; FCOMP-01-0124-FEDER-007243) financed by the Portuguese Foundation for Science and Technology.

The Authors are with the Department of Informatics Engineering of the University of Coimbra, Polo-II, Coimbra, Portugal; {jpramos@student.dei.uc.pt and {carvalho, ruipedro, jh}@dei.uc.pt

a linear phase bandpass modulation filter and its complementary bandstop filter. For a more detailed explanation the reader is referred to [7].

As aforementioned, our interest is not in signal separation and reconstruction, but rather in discriminating noise interference on the extracted non-cardiac signal. The algorithm proposed herein will work on 3 second frames. With such period we are sure to have more than one heart beat, per frame, which will then help discriminating non-stationary signals - noise - from quasi-stationary signals, heart sound.

The noise detection algorithm is composed by the following processing steps: a STFT is performed on the possibly contaminated signal frame $s(t)$ and the magnitude ($|\hat{S}(f)|$) for each frequency bin is computed. A Hann window of length 20 milliseconds with 75% overlap is used to compute the STFT. Afterwards, since we are only considering the non-cardiac signal, the temporal trajectory of each frequency bin is lowpass filtered with a cutoff frequency of $1Hz$. The output is the temporal trajectory ($|\hat{S}(f)|$) of the non-cardiac signal, for each frequency bin. Feature extraction is then performed along the temporal trajectories. An overview of the algorithm is depicted in Fig. 1.

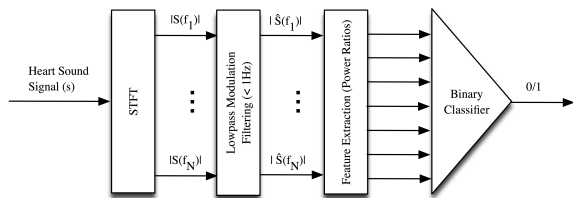


Fig. 1. Block diagram of the proposed algorithm.

A. Feature Extraction

The algorithm should be able to process the frames as they are acquired in real-time. Hence, the algorithm will resort to computationally less complex features to capture the signal's non-stationarity.

The stationarity quality of a signal is perceived both in time domain and in time-frequency domain. The features applied are thus an attempt to infer whether such stationarity has been compromised or not. In the proposed method, stationarity is assessed using power ratios of differently sized windows. Let $\hat{S}_w(f, k)$ be the windowed signal of $|\hat{S}(f, k)|$, i.e.

$$\hat{S}_w(f, k) = \begin{cases} |\hat{S}(f, k)| & , k \in \text{window} \\ 0 & , k \notin \text{window} \end{cases}$$

If L is the duration of the signal under analysis, then the power ratio for a given window w of size L is defined as in eq. 1.

$$P_w = \frac{\sum_{k \in w} |\hat{S}_w(f, k)|^2}{L} \frac{L}{\sum_{k=1}^T ||\hat{S}(f, k)||^2} \quad (1)$$

For each temporal trajectory $\hat{S}(f, k)$, five features are extracted: first the temporal trajectory is split in half and the power ratios of the left window, $P_{LW1.5}$, and the right window, $P_{RW1.5}$, are computed. Secondly, the same temporal trajectory is split in three equally sized windows and their respective power ratios, denoted by $P_{LW1.0}$, $P_{MW1.0}$ and $P_{RW1.0}$ are obtained.

The foundations of these features can be found in fig. 2. Heart rates, at rest, are between 30 to 120 beats. Hence, a 3 sec. temporal trajectory contains at least 2 complete heart beats. By splitting the temporal trajectory in half and comparing the power ratios of both windows, they should be very similar in a stationary signal. However, when a transient noise contaminates the signal, it reveals high intensity peaks that will disrupt the stability of the signal. Moreover, those peaks may be located in different positions and have different widths. Hence, the temporal trajectory is further split in three to capture possible boundary's cases, i.e., when the peak is located in the middle of the temporal trajectory.

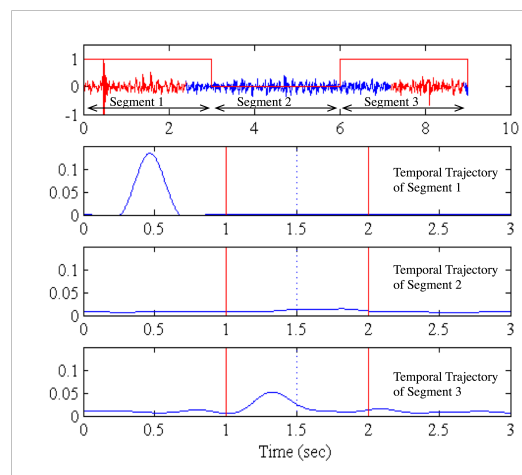


Fig. 2. Top: Contaminated heart sound and the algorithm response (blue soundwave is the clean heart sound whereas the red soundwave is noise). The other graphics represent the temporal trajectories of each segment analysed. The blue dotted line delimits the windows' split in two and the red line delimits the windows' split in three.

Furthermore, different noises can have their influence in different frequency bins. To avoid increasing the classifiers' complexity, the classifier is fed with the power ratios from the temporal trajectory that exhibits the maximum unbalance in the temporal power distribution. These are captured by features $diff_{max1.5}$ and $diff_{max1.0}$ defined in eq. 2 and eq. 3.

The actual feature set fed to the classifier is composed by $diff_{max1.5}$, $diff_{max1.0}$ and the power ratios $P_{LW1.5}$, $P_{RW1.5}$, $P_{LW1.0}$, $P_{MW1.0}$ and $P_{RW1.0}$ extracted from the temporal trajectories where $diff_{max1.5}$ and $diff_{max1.0}$ are observed, respectively.

$$diff_{max1.5} = \max_f (|P_{LW1.5}(f) - P_{RW1.5}(f)|) \quad (2)$$

$$diff_{max1.0} = \max_f (|P_{LW1.0}(f) - P_{MW1.0}(f)|, |P_{LW1.0}(f) - P_{RW1.0}(f)|, |P_{MW1.0}(f) - P_{RW1.0}(f)|) \quad (3)$$

The classifier structure used to perform the binary classification was a SVM with a radial basis kernel function. The classification is done for the whole 3 sec. segment, regardless of the noise's duration.

III. RESULTS AND DISCUSSION

The database is composed by 163 heart sounds' records from 71 different patients with prosthetic valve implants (both mechanical and bioprosthetic) and healthy individuals as well. The collected dataset has a total of 97 minutes and is composed by 678 segments of contaminated heart sounds and by 1275 clean heart sound segments. All sound samples were acquired with 16-bit resolution and a $44.1KHz$ sampling rate. These were downsampled to $11025Hz$ prior to processing. Several common types of noises expected to occur during real-life auscultation were induced according to a predefined protocol during acquisition. Non-cardiac sounds like vocal sounds, breathing and coughing at different intensity levels and pitches, sensor abrasion and other ambient noises can be found along the dataset. It should be mentioned that noises induced by accident during the signal acquisition protocol have also been annotated.

A. Performance Assessment

In order to validate the classification algorithm, K-fold cross-validation was used with $K = 10$, which divides the dataset into K balanced subsets, trains the classifier with $K - 1$ sets and tests with the other one.

The validation assessment is done using two measures: sensitivity and specificity. The overall obtained results using the entire dataset are as follows:

- Sensitivity (%): 93.6 ± 3.0
- Specificity (%): 92.3 ± 2.5

The mean and standard deviation were calculated after five runs of 10-fold cross validation.

In order to get a better insight of the algorithm's performance, a simulation study was performed. In this study, clean heart sounds were additively mixed with noise contaminations with different gains. Namely, let $x(t)$ be the clean heart and $v(t)$ the noise sound to be mixed, then the resulting mixed sound would be $u(t) = x(t) + Kv(t)$, where $K \in [0, 1]$ denotes noise gain. Heart sounds with different conditions were used and were contaminated with noise of the same categories found in the dataset. The purpose of this study is to infer the algorithm's behaviour when noise intensity and heart conditions vary.

During the process, the SNR of the heart sound mixed with noise sources was varied using a linear gain and the sensitivity was calculated. The results are reported in fig. 3 and 4 for different heart conditions and for different types of noises. During each simulation, 2 heart sound signals

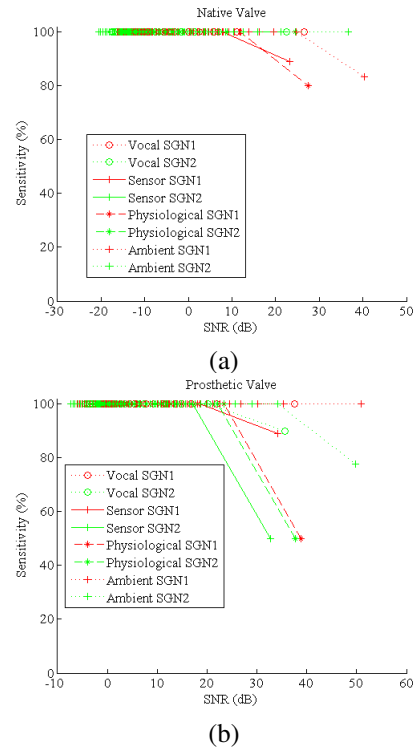


Fig. 3. Sensitivity for noise detection with different noise intensities of two heart sounds with (a) native valves and (b) prosthetic valves.

(denoted SGN1 and SGN2) were contaminated with vocal, physiological, ambient noises and sensor movement.

As it would be expected, the sensitivity results decreases as the SNR increases. However, the degradation of those results only happens when the SNR is positive, which are the situations where the noise intensity is lower than the heart sound's. Regarding the heart condition, the algorithm seems to best behave with prosthetic valve or with no heart condition at all, having a decrease in sensitivity only above $20dB$, whilst the degradation of sensitivity with arrhythmias and murmurs starts at $10dB$. This is linked to the fact that heart sounds with murmur or arrhythmia exhibit more complex time-frequency distributions, leading to higher disturbances in stationarity. Concerning the noise class, the algorithm has a better performance when it encounters vocal sounds and a lower performance when it encounters physiological sounds, regardless of the heart condition. This is related to the fact that physiological contaminations have their frequency content very similar to the spectral content of the heart sound. Nevertheless, it is worth noticing that such degradation is only to SNR values higher than $30dB$.

Another important measure to assess is the time the algorithm takes to extract the features and classify the segment, since it is our intention to process the segments as they are acquired to provide timely feedback to the user. For each 3 second window the algorithm is able to return a classification in $.56 \pm .024$ seconds. The environment for this evaluation was a Windows XP Intel®Pentium 4, CPU $3.00GHz$, with $2.25GB$. The Matlab®version is 7.6 (R2008a).

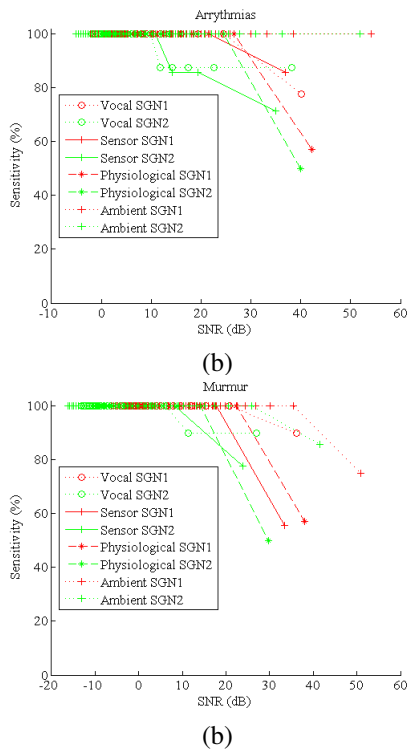


Fig. 4. Sensitivity for noise detection with different noise intensities of two different heart sounds with (a) arrhythmias and (b) murmurs.

IV. CONCLUSIONS

A novel approach for noise detection in heart sound was suggested. The extracted features rely on the assumptions that the non-cardiac signal extracted is stationary and that any interference from a transient noise will affect this quality. When compared with other state of the art algorithms, the main advantage of this approach is that it does not rely on a reference sound. Regarding its performance, the algorithm has a considerable satisfactory performance and consistency when it encounters different noise contaminations and heart conditions.

REFERENCES

- [1] S. Allender, P. Scarborough, V. Peto, and M. Rayner, "European cardiovascular disease statistics," *British Heart Foundation Health Promotion*, 2008. [Online]. Available: <http://www.heartstats.org/datapage.asp?id=7683>
- [2] C. Shub, "Echocardiography or auscultation? how to evaluate systolic murmurs." *Can Fam Physician*, vol. 49, pp. 163–7, 2003. [Online]. Available: <http://www.biomedsearch.com/nih/Echocardiography-auscultation-How-to-evaluate/12619738.html>
- [3] R. L. Watrous, "Computer-aided auscultation of the heart: from anatomy and physiology to diagnostic decision support." *Conf Proc IEEE Eng Med Biol Soc*, vol. 1, pp. 140–3, 2006.
- [4] S. Rajan, R. Doraiswami, R. Stevenson, and R. Watrous, "Wavelet based bank of correlators approach for phonocardiogram signal classification," in *Time-Frequency and Time-Scale Analysis, 1998. Proceedings of the IEEE-SP International Symposium on*, 6-9 1998, pp. 77–80.
- [5] S. Rajan, E. Budd, M. Stevenson, and R. Doraiswami, "Unsupervised and unced segmentation of the fundamental heart sounds in phonocardiograms using a time-scale representation," in *Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE*, aug. 2006, pp. 3732–3735.

- [6] D. Kumar, P. Carvalho, M. Antunes, J. Henriques, R. Schmidt, and J. Habetha, "Near real time noise detection during heart sound acquisition," in *15th European Signal Processing Conference. EURASIP, 2007*.
- [7] T. H. Falk and W.-Y. Chan, "Modulation filtering for heart and lung sound separation from breath sound recordings," in *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, 20-25 2008, pp. 1859–1862.