31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Noise Detection During Heart Sound Recording

D. Kumar, P. Carvalho, M. Antunes, J. Henriques

Abstract— **Heart sound is a valuable biosignal for early detection of a large set of cardiac diseases. Ambient and physiological noise interference is one of the most usual and high probable incidents during heart sound acquisition. It may change the prominent and crucial characteristics of heart sound which may possess important information for heart disease diagnosis. In this paper, we propose a new method to detect ambient and internal body noises in heart sounds. The algorithm utilizes physiologically inspired periodicity/semi-periodicity criteria. A small segment of clean heart sound exhibiting periodicity in the time and in the frequency domain is first detected. The sound segment is used as a template to detect uncontaminated heart sounds during recording. The technique has been tested on the heart sounds contaminated with several types of noises, recorded from 68 different subjects. Average sensitivity of 95.13% and specificity of 98.65% for non-cardiac sound detection were achieved.**

I. Introduction

Cardiovascular diseases are the leading cause of death in developed countries. In Europe it is estimated that chronic cardiovascular diseases are responsible for 20% to 45% of all deaths. Being a disease tightly connected to aging, it is observed that its incidence is on the rising due to the extended life expectancy. The solution to this health problem is believed to be changing the focus from curative healthcare to preventive healthcare. This is commonly believed to be achievable by fostering preventive lifestyles as well as early diagnosis. In this sense long term tele-monitoring is a promising tool to achieve early diagnosis which may avoid potentially life-critical situations as well as aggressive and expensive treatments. Therefore, current research trends in this direction are to integrate system solutions into ordinary daily objects, such as functional clothes with integrated textile or hard sensors. In order to be cost effective and usable for long time period, these tools require intelligent data analysis algorithms to be able to autonomously perform diagnostic functions and to support users in solving problems. Hence, low computational algorithms that can be run in real-time using low power processing devices is required. In order to design effective diagnosis algorithm using vital signals, it is observed that the noise cancelation during signal acquisition is a primary and indispensable task. This task is imperative for reliable diagnostic feature extraction.

This work was performed under the framework of the project PTDC/EIA/68620/2006 financed by FCT (Fundação para a Ciência e a Technologia).

D. Kumar, P. Carvalho and J. Henriques are with the Department of Informatics Engineering of the University of Coimbra, Polo-II, Coimbra, Portugal; E-mails: {dinesh, carvalho, jh}@dei.uc.pt

M. Antunes is with the Cardio-thoracic Surgery Center of the University Hospital of Coimbra, Coimbra, Portugal; E-mail: antunes.cct.huc@sapo.pt

Heart sound is a valuable biosignal for early detection of a large set of cardiac diseases. Unfortunately, heart sound is more prone to noise than other bio-signals, therefore, many researchers have used ECG as a reference or marker to find the non-cardiac sounds. In [1] an ECG was used to find heart beats in the heart sound. In this approach, power spectrum of each beat of heart sound is correlated with the known clean beat of heart sound to determine contaminated beats. A very well known method for speech enhancement based upon spectral domain Minimum-Mean Squared Error (MMSE) estimation was applied to reduce noise affect in heart sound [2]. This method reduces white noise from heart sound while S3 and S4 sounds are prevented using ECG gating. Another method for noise cancelation in real-time was developed using an extra acoustic sensor to capture the environmental noise. This additional signal provides a noise signal to subtract environmental noise from the contaminated heart sound signal [3]. A simple method of filtering with a certain band of frequencies was used in [4]. In order to develop cost effective, portable and practical systems, the noise removal algorithm should exclude dependence on ECG or any non-cardiac sound sensors.

Periodicity of heart sound components (namely S1 and S2) is an inspiration to detect non-cardiac sounds in the heart sound. It is observed that periodicity in heart sound tends to be violated in the presence of internal body noises or external noise sources. Therefore, its periodic nature could be a significant criterion to assess noise free heart sound. The proposed method first searches for a clean heart sound segment as a reference signal based on periodicity in the time and the frequency domain (spectrogram). Afterwards, the spectral energy of the reference signal is correlated with the rest of the heart sound in real time. Heart sound segments which exhibit low correlation coefficients with the reference heart sound signal are assessed as non-cardiac sounds. This method is able to detect numerous kinds of physiological and environmental non-cardiac sounds.

The paper is organized as follows: in the second section, the proposed method is thoroughly explained, the third section introduces results and discussion, and finally some conclusions are presented in the fourth section.

II. Method

In order to extract reliable diagnostic features from heart sound it is important to first suppress noise. During heart sound acquisition many external body noises, such as ambient noises, and internal body noises, such as heavy breathing, sounds derived by swallowing, speech etc., may be captured. These noises are linearly/nonlinearly mixed with heart sounds. Therefore, the suppression of these noises in heart sounds is not a straightforward problem. Furthermore, even a small ambiguity in suppression may lead to wrong diagnosis based on heart sound features. The strategy proposed in this paper is to detect the contaminated sound segments and to exclude them from further processing. It includes two main phases: the goal of the first phase is to find a clean segment of heart sound that will be further used as a reference sound and the second one is to correlate the remaining heart sounds with the reference in order to identify the clean cardiac sounds. The involved steps in these two components of the method can be seen in Figure 1, and are explained in the following sections.

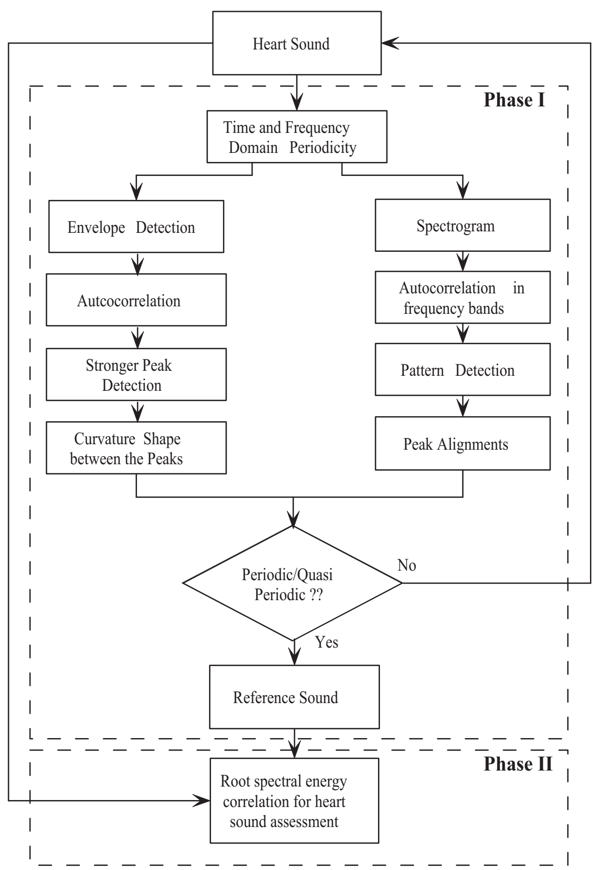

Fig. 1. Flow chart of noise detection in heart sound recording.

A. Template reference sound detection

The foremost task is to find a small segment of clean heart sounds which will be further treated as a reference sound. In this process, the period of the heart cycle is estimated in the chosen window in the time domain. Later, this estimated period is used to check about the same number of heart cycles in the frequency domain in the same analysis window. At last, one heart cycle among the most alike heart cycles is chosen to be the reference sound. The detailed method for heart cycle identification is described below:

(A) Periodicity in the time domain: The envelopes of the heart sound $x(t)$ components are extracted by applying the Hilbert transform followed by the Gammatone band-pass filter [5]. Next, the autocorrelation function of the envelope is computed. Typically, it will exhibit pronounced peaks for the main heart sound components, i.e. the S1, S2 and murmur components. The autocorrelated values are normalized by autocorrelation values of a chosen windowing function, e.g. the Hanning window. Let $x^e(t)$ be the envelope of the heart sound, $w(t)$ be a windowing function, N be number of the samples in a given segment of heart sound and τ be time lag, then autocorrelated function, can be formulated as in (1).

$$y(t) = \frac{\int_0^N x^e(t)w(t) * x^e(t-\tau)w(t-\tau)d\tau}{\int_0^N w(t) * w(t-\tau)d\tau}, \qquad (1)$$

Two factors are important to examine periodicity from normalized autocorrelation $y(t)$: the first one is the stronger peaks and the second one is the curvature shape between the stronger peaks. Each prominent peak corresponds to one heart cycle. In the next subsections, prominent peaks detection using heart cycle (heart rate estimation) and curvature shape using radial distance between two contiguous heart cycles are explained.

(A1) Prominent peaks detection: In this task, the heart rate in a given analysis window is estimated using $y(t)$ and a physiological criterion. The prominent peaks in $y(t)$ are found using the estimated duration (or heart rate) of heart cycles in a given segment of the analysis signal. Heart rate in a given segment of heart sound is estimated using the large difference between two first singular values of the autocorrelation matrix (achieved via singular value decomposition) which is constructed by delaying and rearranging the autocorrelation function $y(t)$ of the envelopes [7]. The value of delay is set between 500 ms and 1200 ms, as heart rate at rest are usually found between 50 beats/min and 120 beats/min.

Prominent peaks detection is followed according to the algorithm described in [6]. Unlike this one, in the proposed approach the estimated heart rate facilitates to discard weak peaks around the strong peaks in $y(t)$. Strong peaks are directly related to the main components in the heart sound which occur only once per heart cycle (see in Figure 2).

(A2) Periodicity check criterion: All detected stronger peaks of $y(t)$ in the previous step enable to find the shape similarity between two consecutive heart cycles (contiguous pair of stronger peaks). The similarity is measured by the radial distance between two vectors. Let $y_1(t)$ and $y_2(t)$ be the two vectors, the radial distance is given by,

$$Cos(\theta) = \frac{< y_1(t), y_2(t) >}{| y_1(t) | | y_2(t) |}, \qquad (2)$$

where $< . >$ is the inner product operator and $| . |$ represents mean square root value of the vectors. In all situations of

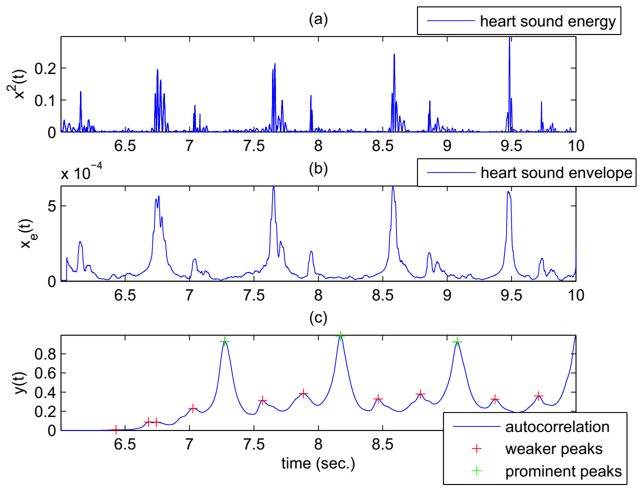

Fig. 2. (a) Heart sound energy ($x^2(t)$); (b) Heart sound envelope; (c) Autocorrelation function of the envelope with peaks identification.

similar periodic shapes, it is observed that the $Cos(\theta)$ value lies in the vicinity of 1.0, as shown in Figure 2(c).

(B) Periodicity in the frequency domain: The periodicity of heart sounds in the time domain may not be affected by many non-cardiac sounds. For instance, swallowing, breathing or high pitched voice can not be identified using the time domain periodicity detection technique. However, influence of the noise source in the heart sound periodicity can be seen in the frequency domain. The spectrogram is utilized to find the periodic patterns in different frequency bands. Let $S(f,t)$ be the short Fourier transform of $x(t)$, i.e.,

$$S(f,t) = \int_0^M x(\tau)w(t-\tau)exp(\frac{-2jf\pi\tau}{M})d\tau, \qquad (3)$$

where M is the window size. It has been observed in normal/abnormal heart sounds that most power is concentrated up to frequency of 600Hz that corresponds to the 15 frequency bands from the spectrogram matrix. Hence, spectral energies of these frequency bands are taken for periodicity verification, as depicted in Figure 3 and 4. It should be noticed that normal heart sounds exhibit regular patterns in these frequency bands which are linearly dependent. These linear dependencies may monotonically decrease/increase in heart sounds from prosthetic valve click and native valve clicks. In these situations, the peaks in power are almost aligned. These phenomenons are seen in the heart sounds which exhibit periodicity in the frequency domain. The methods for the verification of S1 and S2 periodicity in the frequency domain using 15 frequency bands and the criterion for power peaks alignment in these bands are explained next.

(B1) Pattern detection in the frequency bands: In order to extract periodic patterns from the spectrogram matrix, the rows are autocorrelated as given in (1). Let $S^k(f,t)$ be the autocorrelated function of the k^{th} frequency band in the spectrogram (see Figure 3 and 4). It is observed that the autocorrelated power in the frequency bands are in regular patterns, where the stronger peaks occur almost at the same time. Furthermore, it is also seen that the widths of these peaks decrease in higher frequency bands. On the other hand, the peak widths increase in absence of signal power in the high frequency bands; usually, strong peak widths are absent in native heart valve clicks. Therefore, these observations inspire in building the heuristic which may be applied to verify if the cardiac signal is clean. One of the observations is linear independence of the rows of $S^k(f,t)$. Singular value decomposition technique is applied for linear independency measurement. Since the peak widths in $S^k(f,t)$ increase or decrease in ascending frequency bands, each five contiguous ascending frequency bands are grouped according to (4) and the singular values are computed using the singular value decomposition (SVD) technique.

$$Sg^{(k,k+4)}(f,t) = \begin{pmatrix} S^k(f,t) \\ S^{k+1}(f,t) \\ . \\ . \\ S^{k+4}(f,t) \end{pmatrix}, k = 1,6,11. \qquad (4)$$

In (4) $Sg^{(k,k+4)}$ is the matrix formed by grouping of $S^k(f,t)$ rows for each five ascending frequency bands. The SVD of matrix $Sg^{(k,k+4)}(f,t)$ provides the singular values which reveal linear independence. The parameter $\rho = (\sigma_2/\sigma_1)^2$, where σ_1 and σ_2 are the two largest eigen values of the constructed matrix in (4), exhibits low value for linear dependent rows. Let ρ_1, ρ_2 and ρ_3 be the singular values ratios of the matrix $Sg^{(1,5)}(f,t)$, $Sg^{(5,10)}(f,t)$ and $Sg^{(10,15)}$ respectively, then the most significant observations regarding pure heart sounds are: $\rho_i > \rho_{i+1}$ or $\rho_i < \rho_{i+1}$, $i = 1,2$. In the situations of non cardiac sounds these sequences are violated.

It should be noted that the prominent peaks are due to heart valve clicks. For S1 and S2 sounds of native valves, it is observed that the width and hight of the peaks decrease in the high frequency bands while in prosthetic valve heart sound peak widths decrease but peaks heights increases due to high power availability in high frequency bands. Therefore, in prosthetic valve heart sounds linear independency monotonically decreases in the high frequency bands.

(B2) Peak alignment in the frequency bands: As it has already been explained, the peaks are due to S1 and S2 heart sounds. Therefore, they should exhibit alignment in $Sg^{(k,k+4)}(f,t)$ frequency bands, whereas in presence of noise the peak alignment is violated. In order to check the alignment, all main peaks are found using the previously mentioned peak detection technique. Afterwards, defining a time tolerance ($\pm 5\%$ of the time of first peak), all peaks alignment are inspected.

Finally, a segment of a clean heart sound is selected from the periodic heart sound segment. Length of selected reference sound corresponds to the distance between two consecutive prominent peaks in the $y(t)$.

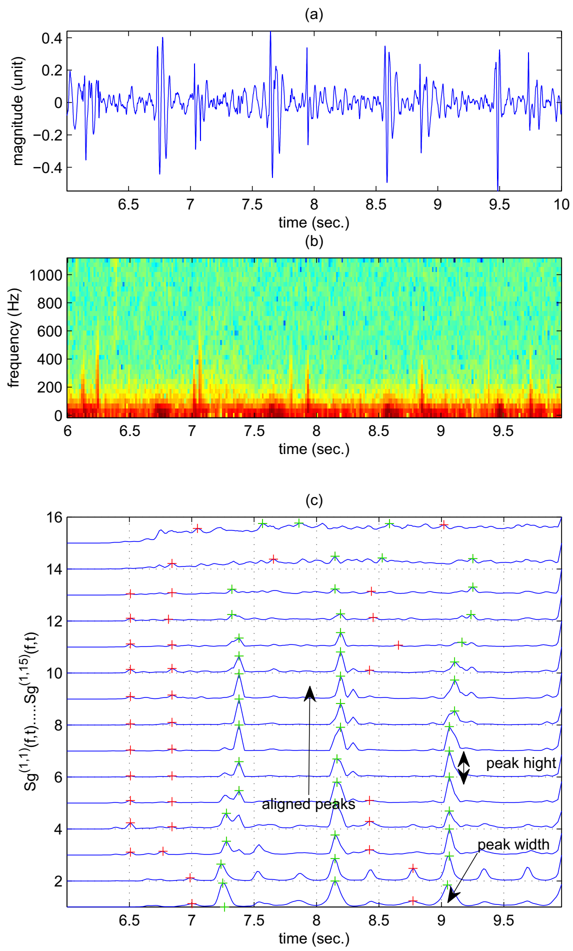

Fig. 3. (a) Normal heart sound from a native valve; (b) Spectrogram; (c) Autocorrelation function $S^k(f,t)$ of spectral power in 15 frequency bands. In this situation ρ increases in high frequency bands.

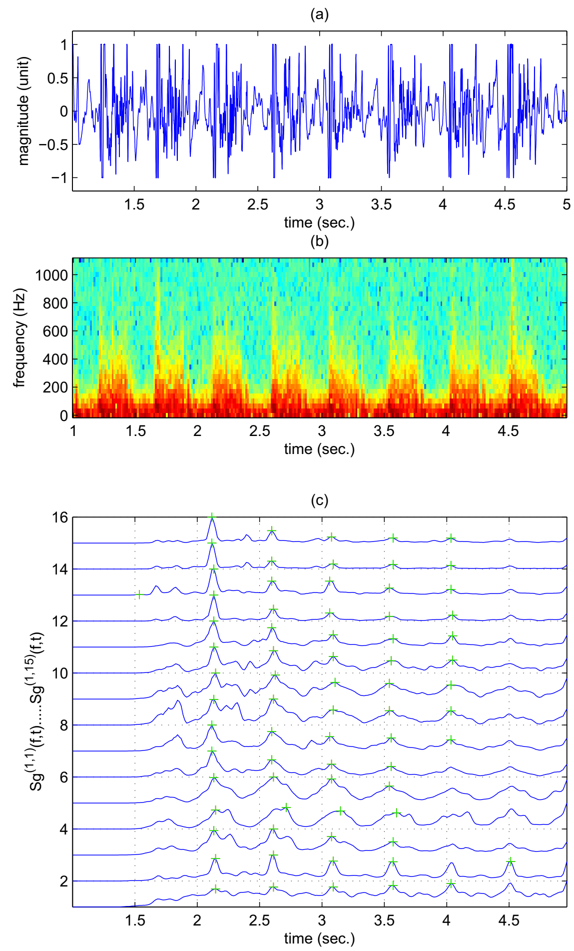

Fig. 4. (a) Heart murmur (grade 4/6) from a subject with mitral regurgitation; (b) Spectrogram; (c) Autocorrelation function $S^k(f,t)$ of spectral power in 15 frequency bands. In this situation ρ increases in high frequency bands.

B. Non-cardiac sound detection using reference signal

The previous subsections introduced the preparation phase which is required to detect the reference heart sound. This reference sound is used as a template for further noise detection. In this step, the spectral root mean square of the heart sound signal is calculated from the following equation,

$$S_{rms}(f,t) = \sqrt{\int_0^{T_w} \mid S(f,t) \mid^2 dt}, \tag{5}$$

where T_w is the size of the reference signal. Root mean square of the spectrogram provides an estimate of the power distribution in the frequency domain. Let $S_{rms}^{ref}(f,t)$ and $S_{rms}^{test}(f,t)$ be the spectral root mean square for the reference and the test heart sound signals respectively. The validation is performed using the following condition,

$$CorrCoef(S_{rms}^{ref}(f,t), S_{rms}^{test}(f,t)) > th, \tag{6}$$

where $CorrCoef$ is the correlation coefficient between two signals. The threshold th value is set to 0.99 in (6).

III. Result and Discussion

Heart sounds were recorded from different patients with prosthetic valve implants (both Mechanical and Bioprosthetic) one month after surgery and some healthy volunteers. A commercial stethoscope from Meditron is used for heart sound acquisition which has an excellent signal to noise ratio and extended frequency range (20 - 20,000 Hz). All sound samples were digitized with 16-bit resolution and 44.1 kHz sampling rate.

The prepared data set includes 95 heart sound samples of recording length of about one minute each which have been collected from 68 different subjects. In order to validate the performance of the algorithm, diverse noises were induced on purpose during the heart sound acquisition. Hence, each heart sound sample includes various non-cardiac sounds, such as speech at different pitch levels, several types of environmental sounds, human made ambient noises, and internal body sounds (heavy breathing, speech, rubbing or accidental movement of stethoscope sensor). In the data set 3 subjects exhibited of arrythmia, 31 had artificial valve, 5 patients exhibited heart murmurs and the rest were healthy.

In the prepared database, non-cardiac sounds were arranged in four classes as follows: vocal sounds, breathing sounds, skin rubbing or abrasion sounds via stethoscope, and other ambient noise such as foot steps, door knock, etc. Total number of segments of various durations (20 ms-3 sec) of non-cardiac sounds were 758 vocal, 232 abrasion, 147 breathing, and 402 others.

For the sake of computational cost each heart sound is downsampled to 2.20kHz. A window of 4 seconds was taken to process the heart sounds under this method. This length of analysis window is found to be sufficient to examine periodicity in the time and the frequency domain. If the segment satisfies the conditions to be a periodic segment, a reference sound is selected. For further processing, a window of the same length as the reference sound template is taken to select heart sound and to compare its root spectral energy to the reference sound. If no reference sound is detected in the current processing window, the window is shifted 1 second forward and the process is repeated until a periodic segment is found. Regarding the processing time complexity of the algorithm, it was observed that using Matlab version 7.6(2008a) running on Windows XP using an Intel(R)Core(TM)2Duo at 2.53 GHz, the first phase of the algorithm took an average of 1.23 seconds to process each 4 second window, while in the second phase of the algorithm only took 0.035 seconds per window.

The achieved noise detection performance of the algorithm over the given database is shown in Table I for each considered noise class. Furthermore, the overall sensitivity and specificity are 95.13% and 98.65%, respectively.

TABLE I

NOISE DETECTION (CORRECTLY DETECTED OVER TOTAL NUMBER OF NOISY SEGMENTS)

Heart Sound	Vocal	Abrasion	Breathing	others
Native valve	524/536	176/178	106/114	309/328
Prosthetic valve	115/120	41/41	24/27	36/44
Murmur	83/92	11/11	3/5	9/12
Arrythmia	9/10	2/2	1/1	15/18

It can be inferred from the Table I that the abrasion type of noise exhibits high average sensitivity of 98.29% in four considered classes of heart sounds, whereas breathing sounds show the least sensitivity of 83.75%. Long duration ($> 50ms$) vocal sounds and abrasion are distributed over large frequency range which shows significant spectral energy difference from normal heart sound for native valve subjects. On the contrary, although respiratory sounds are high frequency sounds, their spectral energy is not notably different from heart sound which requires more sophisticated method of signal processing to discriminate [8]. Finally, an example of noise detection in a heart sound mixed with some of the aforementioned noises is shown in Figure 5.

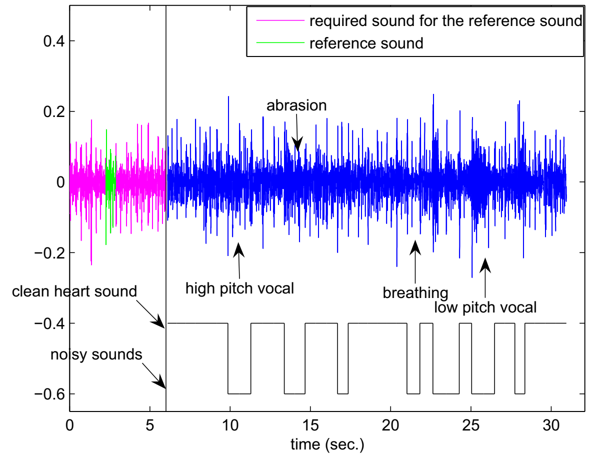

Fig. 5. An example of noise detection in a heart sound sample. The upper curve is the heart sound and the lower one is the noise detection curve.

IV. CONCLUSIONS

A new algorithm for non-cardiac sound detection in real time without ECG as a reference signal was proposed. The algorithm is composed of two main steps: first a reference signal detection based on the periodic nature of clean heart sounds. Secondly, this reference signal is compared to the subsequent recorded heart sounds. The first phase of the algorithm is slightly computationally intensive. However, once the reference signal has been detected, subsequent processing is performed quickly. At this stage, non-cardiac sounds are segregated from heart sound. This suggests that the algorithm might be applied for real time application using low power processing kits.

REFERENCES

[1] P. Carvalho, P. Gil, J. Henriques, M. Antunes and L. Eugénio, "Low Complexity Algorithm for Heart Sound Segmentation using the Variance Fractal Dimension," in *Proc. of Int. Sym. on Intelligent Signal Processing*, Algarve, Portugal, Sep 3-7. 2005, pp. 937-942.

[2] A. S. Paul, E. A. Wan, A. T. Nelson, "Noise Reduction for Heart Sounds Using a Modified Minimum- Mean Squared Error Estimator with ECG Gating," in *Proc. of 28th IEEE EMBS Annual International Conference*, New York City, USA, Aug 30-Sep 3. 2006, pp. 3385-3390.

[3] Y. W. Bai, C. L. Lu, "The embedded digital stethoscope uses the adaptive noise cancellation filter and the type I Chebyshev IIR bandpass filter to reduce the noise of the heart sound," in *Proc. of 7th International Workshop on Enterprise networking and Computing in Healthcare Industry, HEALTHCOM*, June 23-25. 2005, pp. 278-281.

[4] M. Brusco, H. Nazeran, "Development of an Intelligent PDA-based Wearble Digital Phonocardogram, " in *Proc. of 27th IEEE Engineering in Medicine and Biology*, Shanghai, China, Sept 1-4. 2005, pp. 3506-3508.

[5] O. Deshmukh, C. Y. Espy-Wilson, A. Salomon, and J. Singh, "Use of Temporal Information : Detection of Periodicity, Aperiodicity, and Pitch in Speech," *IEEE Tran. on Speech and Audio Processing*, vol. 13, pp. 776–786, Sep. 2005.

[6] T. H. Park , "Salient Feature Extraction of Musical Instrument Signal, " Masters Thesis, *Dartmouth College*, 2000.

[7] P. P. Kanjilal, S. Palit, "Extraction of multiple periodic waveforms from noisy data," in *Proc of ICASSP*, 19-22 April, 1994, pp. II361-II364.

[8] T. H. Falk, W. Chan, "Modulation Filtering for Heart and Lung Sound Separation from Breath Sound Recordings," in *Proc of IEEE-EMBC*, 20-24 August, 2008, pp. 1859-1862.