Analysis of Acoustic Cardiac Signals for Heart Rate Variability and Murmur Detection Using Nonnegative Matrix Factorization-based Hierarchical Decomposition

Ghafoor Shah*‡, Peter Koch* and Constantinos B. Papadias‡
*Department of Electronic Systems
Aalborg University (AAU), 9220, Aalborg East, Denmark
Email:ghs@es.aau.dk, pk@es.aau.dk
‡Broadband Wireless and Sensor Networks Group (BWiSE)
Athens Information Technology (AIT), GR 19002, Athens, Greece
Email:ghasha@ait.edu.gr, papadias@ait.edu.gr

***Abstract*—The detection of heart rate variability (HRV) via cardiac auscultation examination can be a useful and inexpensive tool which, however, is challenging in the presence of pathological signals and murmurs. The aim of this research is to analyze acoustic cardiac signals for HRV and murmur detection. A novel method based on hierarchical decomposition of the single channel mixture using various nonnegative matrix factorization techniques is proposed, which provides unsupervised clustering of the underlying component signals. HRV is determined over the recovered normal cardiac acoustic signals. This novel decomposition technique is compared against the state-of-the-art techniques; experiments are performed using real-world clinical data, which show the potential significance of the proposed technique.**

***Keywords*-Heart rate variability; cardiac sounds; Blind source separation; Nonnegative matrix factorization**

I. INTRODUCTION

The beat-to-beat interval variation called heart rate variability (HRV) is routinely determined from electrocardiogram (ECG) signals recorded over long- or short-time periods. The HRV measured over long-time recordings can be less practical in the sense of maintaining standardized recording conditions for long periods of time in certain time critical applications. Therefore, HRV determined over short-time recordings has drawn more attention in research community [1]. However, irrespective of long- or short-time, ECG-based HRV detection is costly and time consuming, and may not be feasible in certain scenarios such as, infield or rural areas in underdeveloped countries, etc.

The feasibility of measuring HRV determined over normal cardiac sounds recorded through stethoscope was recently studied in [2]. However, cardiac conditions often produce abnormal findings in physical examinations that include pathological cardiac sounds and murmurs which, hinder the HRV detection over normal cardiac sounds. The essence of separating the normal cardiac sound from abnormal sounds and murmurs is twofold. On the one hand these abnormal findings usually provide clues to the underlaying pathophysiology, and their careful interpretation may results into successful diagnosis of the cardiac conditions and disease management [3]. On the other hand a clean normal cardiac sound facilitates the detection of HRV. However, cardiac sounds and murmurs are faint and they are separated less than $30ms$. The deciphering of these sounds through human analysis using ordinary stethoscope is hard. The ability to master the proper auscultation skills is a challenging task and as a consequence young doctors are losing interest in learning the auscultation skills [4]. According to [5], numerous studies have shown that more than 80% of the referrals (based on ordinary stethoscope examination) for further examinations such as ultrasound, which cost $300 to $1000 (in the United States) are false, causing unnecessary anxiety and cost. However, computer-aided auscultation which implies recording and analysis of cardiac sounds and murmurs, can enhance the quality of auscultation [5]. According to [6], cardiac sound analysis can be generally categorized as: 1) segmentation, whereby the cardiac cycle is determined; 2) feature extraction, whereby the distinctive characteristics are computed, and 3) classification, whereby the nature of the cardiac sounds is determined. Many methods have been developed of cardiac sounds analysis where a reference ECG signal was used (such as, see e.g., [7]).

However, in this paper, we discuss separation and analysis of acoustic cardiac signals without any reference ECG signal. A review on the cardiac sound analysis techniques developed recently was presented in [8]. A cardiac sound segmentation algorithm based on energy plot was proposed in [9]. [10] has proposed segmentation of the cardiac sound using wavelet decomposition for feature extraction. A support vector machine (SVM) was used for the classification of the normal and abnormal cardiac sounds. Localization of the cardiac sound in

978-1-4799-7502-0/14 $31.00 © 2014 IEEE
DOI 10.1109/BIBE.2014.14

respiratory signal was proposed in [11] using singular spectrum analysis (SSA), a subspace analysis technique. Separation of murmurs from the cardiac sound was proposed in [12], where, SSA technique was applied. Ensemble empirical mode decomposition (EEMD) has been applied in heart sound segmentation and extraction [13]. Matrix decomposition such as singular value decomposition (SVD) and QR decomposition (QRD) techniques were applied to the time-frequency representation produced by continuous wavelet transform (CWT), for cardiac sounds and murmurs classification [14]. Analysis of the cardiac sounds has been carried out in many research works thus far. However, most of these techniques underperform in the actual real scenarios. Therefore, recovery and analysis of the cardiac sounds in real scenarios such as under the influence of abnormal cardiac sounds, murmurs, noise, etc., needs further investigation.

The aim of this paper is to investigate the potential application of nonnegative matrix factorization to the single channel blind source separation (BSS) of the cardiac sounds and murmurs. Nonnegative matrix factorization (NMF) [15] technique has been widely used in the single channel BSS of audio streams, drum transcriptions and musical data [16]. NMF technique has also been recently used in the BSS of the cardiac and respiratory signals in [17], [18], [19]. NMF technique is generally applied to the magnitude spectrogram in order to produce a low dimensional approximation of the original data, in the form of two non-negative factors. One matrix having the spectral basis vectors and the second matrix containing time-variant gain information for each basis vector. Based on different NMF techniques, a multi level decomposition method is proposed in this paper. The first level decomposes the mixture into: 1) partial normal cardiac sounds, and 2) partial abnormal cardiac sounds and murmurs, by using biomedical characteristics of the cardiac sounds. In the subsequent levels, the decomposed components are further decomposed into underlying clusters by applying NMF in each stage. A novel acoustic cardiac signal (ACS)-based HRV detection method is also proposed. The main contributions of this research work are the followings:

1) Separation of normal cardiac sound from interfering murmurs and noises.
2) Cardiac sound segmentation and localization of the first heart sound (S_1) and second heart sound (S_2).
3) A novel heart rate variability (HVR) detection method based on S_1S_1 (lub-lub), S_2S_2 (dub-dub), S_1S_2 (systolic), and S_2S_1 (diastolic) intervals variations.
4) Initialization of various matrices in NMF through a new scheme.

The rest of the paper is organized as follows: The characteristics and generation mechanism of cardiac sounds and murmurs is briefly discussed in Section II. The separation and analysis method is discussed in Section III. Section IV gives the experimental details and results which are discussed in Section V. Section VI concludes the work.

II. Cardiac signals

In this Section we briefly discuss the characteristics and generation mechanism of cardiac sounds and murmurs. For more details, refer to [3]. We also show the coherence between the generated acoustics and electrical signals.

A. Cardiac cycle

In blood circulation system, the cardiac cycle is the systematic and precisely timed electrical and mechanical events which result into rhythmic atrial and ventricular contractions. A cardiac cycle consists of two phases: 1) a systole, refers to ventricular contraction, and 2) a diastole, which refers to ventricular relaxation and filling. Generally the duration of systole remains constant during a cardiac cycle whereas, the length of diastole varies with the heart rate [3].

B. Cardiac sounds

Various cardiac sounds are produced during a cardiac cycle, resulting from the mechanical vibrations of different parts of the heart. During the process, the various vibrations also produce electrical signals called ECG due to depolarization and repolarization processes [3]. The coherence between the ECG and acoustic cardiac signals is illustrated in Fig. 1. In the followings we discuss the relevant cardiac sounds.

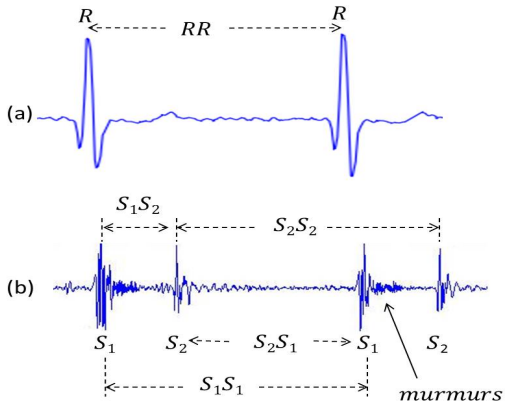


Figure 1: (a) shows ECG signal and (b) shows cardiac acoustics signals. Various intervals are also illustrated.

1) First heart sound: The first heart sound (S_1) is produced during early systole by the closure of the mitral and tricuspid values. The mitral closure which is normally preceded by tricuspid closure are separated by $10ms$. S_1 is a low-medium frequency sound.

2) Second heart sound: The second heart sound (S_2) is produced as a result of the closure of the aortic and pulmonic valves. S_2 is a high frequency sound and has two components: 1) aortic (A_2), and 2) pulmonic (P_2). A_2 and P_2 are normally heard; one sound during expiration and two different sounds during inspiration.

C. Murmurs

During hemodynamic and/or structural changes the laminar blood flow can become turbulent and produce high frequency audible noises called murmurs. Although murmurs can be nonpathological (for instance, in children), most of the times they indicate different valvular diseases depending upon their intensity, shape, pitch, location, timing, etc. [3].

III. Separation and analysis of cardiac sounds

A. Signal model

We model the various cardiac sounds and murmurs in the single channel phonocardiographic signal, i.e. the observation mixture, through the following mixing model:

$$\tilde{x}[n] = \sum_{i}^{I} a_i y_i[n] + \eta_w[n] \tag{1}$$

where $\tilde{x}[n]$ represents the discrete-time observation mixture signal and $y_i[n]$, a_i represent the i^{th} source and its amplitude, respectively. $\eta_w[n]$ and I represent white Gaussian noise (arising during signal acquisition etc.), and total number of sources, respectively. For simplicity, in our mixing model (1), we assume that cardiac sounds and murmurs combine linearly [20]. In our work, $i \in \{c, m\}$, where c, m represent the cardiac sound and murmur signal domains, respectively.

B. Preprocessing

1) Centering and normalization: To remove any direct current (DC) offset which carry no information, and to normalize the signal, we perform the following:

$$x = \frac{\tilde{x} - \mu}{\sigma}, \tag{2}$$

where $\mu = \frac{1}{N}\sum_{n=1}^{N} \tilde{x}[n]$, $\sigma = \sqrt{\frac{1}{N}\sum_{n=1}^{N} (\tilde{x}[n] - \mu)^2}$, N, and x are mean, standard deviation, length of the signal, and mean-normalized signal, respectively.

2) Time-domain filtering: To remove out-of-band noise we use a 3^{rd} order bandpass digital Butterworth filter.

3) Spectrogram formation: A simple method of spectrogram generation is short-time discrete Fourier transform (STFT) where a time-domain signal is divided into small frames using a suitable window function, and a discrete Fourier transform is performed on each frame. $X[t, f]$:=STFT{x[n]} denotes the complex spectrogram of the signal $x[n]$. t and f represent time index and frequency bin, respectively. $X = ||X[t, f]||$ represents the magnitude spectrogram of the signal $x[n]$.

4) Denoising: The time-frequency analysis can be exploited to apply the classical denoising techniques such as soft thresholding, overlapping group shrinkage (OGS), Wiener post processing, block thresholding, etc. We use OGS denoising technique. For more details we refer the reader to [21].

C. Basic non-negative matrix factorization

NMF gives parts-based decomposition and imposes the only constraint of non-negativity. Efficient algorithms for NMF computations have been developed in [15]. NMF decomposes a nonnegative matrix $X \in \mathbb{R}^{F\times T}$ into two nonnegative factors $W \in \mathbb{R}^{F\times K}$ and $H \in \mathbb{R}^{K\times T}$, where, $K < min\{F, T\}$, that is

$$X_+ = W_+ H_+ + E \tag{3}$$

where V_+ indicates that the matrix V is nonnegative and E represents reconstruction error. Different cost functions are used for minimizing the reconstruction error. We use a cost function which is the squared Euclidean distance between X and WH, being defined as

$$D_{EUD} = ||X - WH||_F^2 = \sum_{tf} (X_{tf} - (WH)_{tf})^2, \tag{4}$$

where $||\cdot||_F^2$ denotes Frobenius norm. The lower bound of the measure (4) is zero and it is optimized if $X = WH$ or $E = 0$. The corresponding multiplicative updates which converge to local minima are given as

$$W \leftarrow W \odot \frac{XH^T}{WHH^T}, \quad H \leftarrow H \odot \frac{W^T X}{W^T WH}, \tag{5}$$

where $D \odot E$ denotes element-wise multiplication, and $\frac{D}{E}$ denotes element-wise division.

D. Blind source separation method

We address the separation of normal cardiac sounds and murmurs as BSS problem. We decompose the time-frequency representation of the mixture signal into different components (clusters) using NMF technique. Supervised NMF based BSS provides good separation which requires training data of the mixing sources. However, cardiac sounds and murmurs are highly nonstationary, and it is hard to acquire training data of normal cardiac sounds in the presence of abnormal sounds and murmurs. Therefore, we aim to develop an unsupervised technique without any training data, by exploiting the biomedical characteristics of the cardiac sounds. As discussed in Section II, the heart produces low frequency sounds during normal operation, however, during different cardiac conditions higher frequency sounds (murmurs) are produced. Therefore, it can be hypothesized that: *1) The lower frequency sounds show normal cardiac operation whereas the higher frequency sounds signify abnormal cardiac activities.* We

also know from the literature [3] that normal cardiac sounds have a significant portion of their energies below 150 Hz. Therefore, we can state another hypothesis that: *2) Most of the mixture signal below* 100 Hz *partially signifies the normal cardiac operation.* We use mixture signal below 100 Hz [17], as a reference to the normal cardiac sound, in our separation method, which we denote as y_{ref}.

An overview of the proposed separation method is illustrated in Fig. 2. Motivated by the Wavelet decomposition

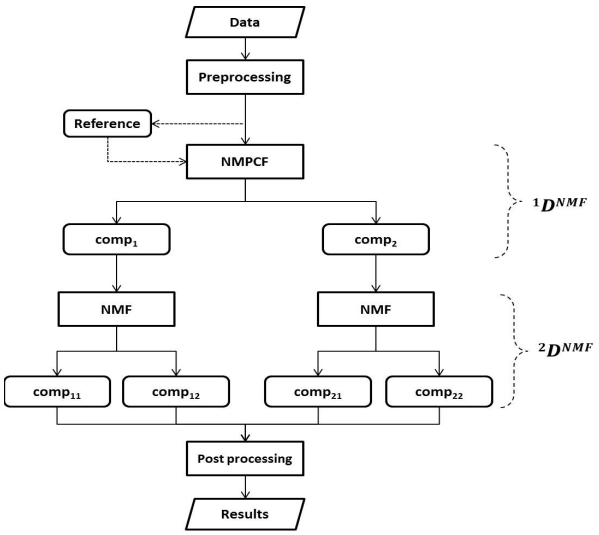


Figure 2: An overview of the proposed method, showing two levels of decomposition.

method, we decompose the input signal in a hierarchical manner using different NMF techniques. In each decomposition level we apply NMF with $K = 2$, i.e., first level of decomposition produces two components, the second level produces four components and so on. We denote each decomposition level with ${}^{\varphi}D^{NMF}$, showing the φ_{th} decomposition level and $\varphi = \{1, 2, 3, ...\}$. As shown in Fig. 2, in ${}^1D^{NMF}$, we use nonnegative matrix partial co-factorization (NMPCF) which was used in the recovery of drum source, where a drum-only matrix *a priory* and a music matrix (mixture) were simultaneously decomposed by sharing some factors [22]. The idea of co-factorization is to use the training data of one of the two mixing sources, *a priory*, to simultaneously factorize two matrices by co-sharing some of the components. Here, we use y_{ref}, *a priori*, to separate the mixture into two components by optimizing the following cost function.

$$\min_{W_c \geq 0, W_m \geq 0, H_c \geq 0, H_m \geq 0, H_{ref} \geq 0} \frac{1}{2}||X - W_c H_c - W_m H_m||_F^2 + \frac{\gamma_c}{2}||Y_{ref} - W_c H_{ref}||_F^2 + \frac{\alpha}{2}||W_c||_2^F + \frac{\beta}{2}||W_m||_2^F \quad (6)$$

where, W_c, W_m, H_c, H_{ref}, and H_m represent the required basis spectra and time varying gain information (coefficients weights) of the cardiac sounds and murmurs, respectively. α and β are the regularization parameters whereas parameter γ_c controls the weights of the shared factor. To solve (6), multiplicative update rules are derived similar to [22]. The ${}^1D^{NMF}$ is summarized in Algorithms 1. In step (2) of Algorithm 1, the various matrices are generally initialized with nonnegative random values. However, we propose a novel scheme to initialize these matrices as the following:

$$W_c = \frac{1}{T}\sum_{t=1}^{T}||Y_{ref}[t,:]|| \quad (7)$$

$$W_m = \frac{1}{T}\sum_{t=1}^{T}||X[t,:] - Y_{ref}[t,:]||, \quad (8)$$

The proposed initialization scheme in (7) and (8) provides better results as compared to random initialization scheme in terms of reconstruction error and time of convergence.

Algorithm 1 NMF-based BSS

1: Compute magnitude spectrograms X and Y_{ref}
2: Initialize the matrices W_c, W_m, using (7) and (8)
3: Iteratively update each matrix using update rules for W_c, W_m, H_c, H_m and H_{ref} for a predefined number of iterations
4: Estimate magnitude spectrogram of the i^{th} source as $X_{(i)} = W_i H_i$

After first level of decomposition the magnitude of the mixture spectrogram X can be approximated as

$$X \approx W_c H_c + W_m H_m \quad (9)$$

where $W_c H_c$ and $W_m H_m$ are the magnitude spectrograms of the partial cardiac sound and murmurs, and can be denoted as $'X_c$ and $'X_m$, respectively. Prior to next decomposition stage, we approximate cardiac sounds and murmurs magnitude spectrograms from the original mixture spectrogram X by applying time-frequency masking discussed in Section III-G, in order to minimize the reconstruction error. The approximated spectrograms, X_c and X_m are again decomposed in the following decomposition level using basic NMF technique discussed in Section III-C. In the ${}^2D^{NMF}$ we choose $K = 2$, for both of the spectrograms.

Each decomposition level produces different features (i.e., components signal) in the given mixture signal. For instance, the ${}^1D^{NMF}$ separates normal and abnormal signals. Depending upon the classification tool, a number of decompositions can be performed. In this preliminary work, we perform only two levels of decomposition.

E. Cardiac sounds localization

Using coefficient matrix of the separated cardiac sound, i.e. H_c, a robust cardiac sound segmentation can be achieved. The coefficient matrix of the cardiac sound H_c is plotted in Fig. 5b. The peaks in Fig. 5b show the locations of S_1 and S_2. We use a simple peak detection algorithm to localize the S_1 and S_2, illustrated in Fig.5a.

F. HRV detection

The coefficient matrix H_c which capture the timing information of the cardiac sounds can be used to develop a reliable HRV detection method. However, unlike the ECG data where RR detection is trivial, H_c shows the locations of both S_1 and S_2, where differentiation between them can be challenging. The other problem in ACS-based HRV is the S_2 split into A_2 and P_2 during inspiration. Therefore, we propose the following steps for ACS-based HRV detection over the normal cardiac sounds:

1) Step 1: To mitigate the problem of erroneous peaks and S_2 splitting, we use the following moving average filtering over a specified smoothing window.

$$h_s[t] = \frac{1}{2L_s+1}(h[t+L_s]+h[t+L_s-1]+...+h[t-L_s]), \tag{10}$$

where $h[t]$, $h_s[t]$ and L_s represent, the normalized coefficient (weight) vector corresponding to normal cardiac sounds, its smoothed waveform and the smoothing window, respectively.

2) Step 2: Peak detection algorithms is used to locate the position of S_1 and S_2 within the smoothed waveform $h_s[t]$.

3) Step 3: Peak conditioning is performed to discard the erroneous peaks or relocate the missing peaks. The peak locations are used to perform segmentation which includes: 1) localizing S_1 and S_2, and 2) determining systolic and diastolic intervals.

4) Step 4: Once we locate the position of S_1 and S_2 in separate waveforms, we can easily measure the variations in S_1S_1, S_2S_2, S_1S_2, and S_2S_1 intervals.

From the theory discussed in Section II and Fig. 1, we know that a cardiac cycle, $S_1S_1 = S_1S_2 + S_2S_1$. We also know that S_1S_2, i.e. systolic interval variation is minimal. Therefore, it can be hypothesized that: *3) The S_1S_1 variation is directly proportional to S_2S_1 variation with some constant, provided S_1S_2 does not vary significantly.* We use this hypothesis to make sure the correct localization and differentiation of S_1 and S_2 and to minimize the risk of erroneous peak detection. We provide a kind of check-and-correct mechanism by calculating the S_1S_1, S_2S_1 and S_1S_2 variations and verifying their relationships according to hypothesis *3)*. The S_1S_1 interval is analogous to RR interval which signifies the beat-to-beat variations. However, the S_1S_2 interval variation shows, intra-beat variations. The ACS-based HRV detection is illustrated through an example in Section IV-D.

G. Post processing

Once the magnitude spectrogram is approximated into original sources, the corresponding phases can also be approximated using the original spectrogram. We generate a time-frequency mask for each source and apply the corresponding mask to the original spectrogram, to recover the original sources. We construct a time-frequency mask as

$$M_i = \begin{cases} 1 \ \forall \, X_{(i)} > X_{(j)}, \ j \in \{m, c\}, j \neq i \\ 0, \ otherwise. \end{cases} \tag{11}$$

The idea of time-frequency masking is based on the assumption that cardiac sound signals are sparse [23], which means that over a small time-frequency region only one source in the mixture dominates. The time-frequency mask (11) is applied to the spectrogram of the mixture (1), to recover the time-frequency representation of original sources as

$$Y_i[t, f] = M_i \odot X[t, f]. \tag{12}$$

The inverse short-time Fourier transform (ISTFT) is used to convert the time-frequency representation back into the time-domain.

IV. Experimental results

A. Experimental setup

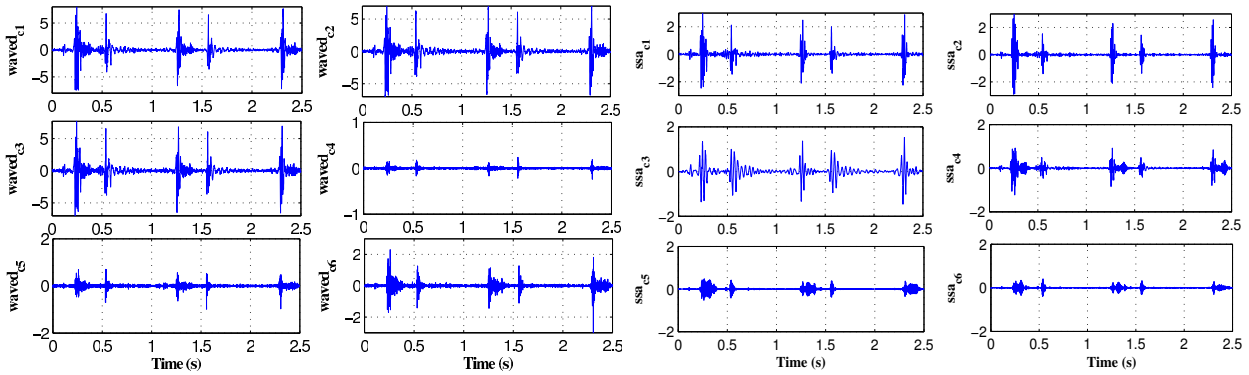
In our experiments, for STFT representation, a *Hamming* window of length 128 samples (equivalent to $32ms$), with 50% overlap was used. The parameters $\gamma_c = 1, \ \alpha = 1, \ \beta = 1$ and $K = 2$ were used in the experiments. The smoothing window $L_s = 80ms$ and a peak threshold $\gamma = 0.3$ were used. The maximum number of iterations used for NMPCF and NMF was 100 and 130, respectively. MATLAB was used for simulations.

B. Clinical Data

The data samples taken from [24] were obtained from different subjects, in noisy clinical settings, using a digital stethoscope, with a data sampling frequency of $4 \ kHz$. In the database [24], we select the samples, where each sample is generally a real mixture of normal cardiac sounds, murmurs, possibly little interference from the lung sounds and white Gaussian noise.

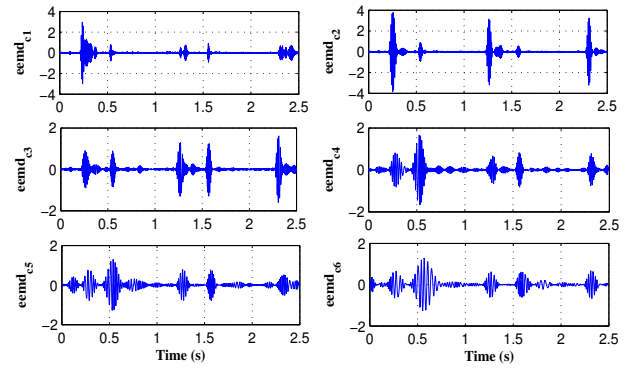
C. Decomposition comparisons

We compare our analysis and decomposition method with the three widely used state-of-the-art decomposition tools: 1) Wavelet decomposition (WD); 2) singular spectrum analysis (SSA), and empirical embedded mode decomposition (EEMD). To provided a fair comparison, we use the same data sample with all the mentioned tools. To the best of our understandings, we tried to achieve the best possible results using similar settings, i.e. decomposition levels, etc., with all

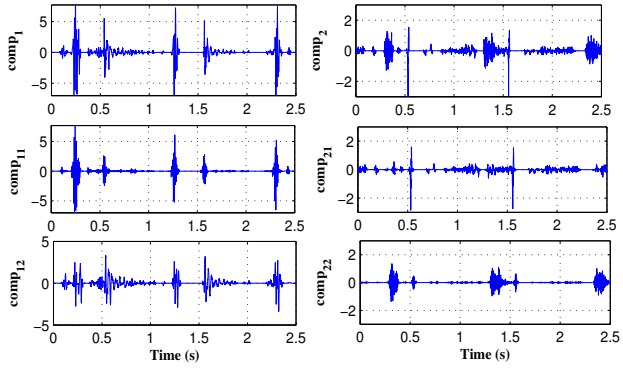


(a) Signal decomposition using Wavelet decomposition.

(b) Signal decomposition using SSA.



(c) Signal decomposition using EEMD.



(d) Signal decomposition using proposed technique.

Figure 3: Comparison of the proposed decomposition method with the state-of-the-art decomposition tools.

the mentioned tools. In the WD, we use "db" as a mother wavelet, with 3 levels of decompositions. The reconstructed 3 approximations and 3 details components which we denote as waved$_{c1}$-waved$_{c6}$, are illustrated in Fig. 3a. In the SSA, we use a window length of 50 samples. The reconstructed, first 6 principal components denoted as ssa$_{c1}$-ssa$_{c6}$, are given in Fig. 3b. The resulting 6 components based on EEMD which we denote as, eemd$_{c1}$-eemd$_{c6}$, are illustrated in Fig. 3c. The reconstructed components using the proposed NMF-based decomposition method are given in Fig. 3d. As shown in Fig. 3, our proposed method and SSA both show better performance as compared to other methods, by extracting meaningful features. The most important extracted features i.e., normal cardiac sounds, represented as comp$_{11}$ and murmurs, represented as comp$_{22}$, using the proposed method are compared against SSA, through more elaborative graphs in Fig. 4. The excellent separation of different components (using the proposed method) and their close match with the original signals, show the potential significance of the proposed method.

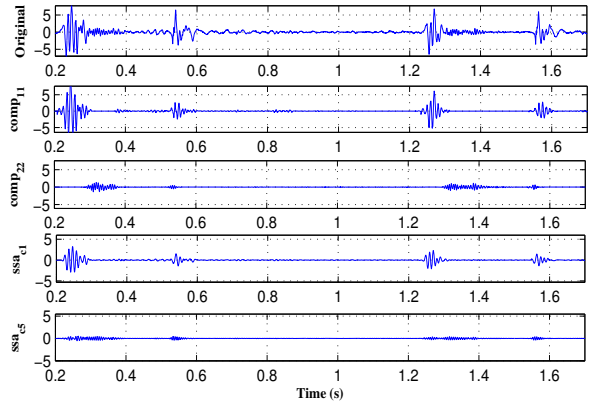
D. ACS-based HRV example

ACS-based HRV detection technique was applied to the clinical samples discussed in Section IV-B. A sample length of $25s$ was used. The samples with shorter lengths were looped to meet the length requirement. Above 90% specificity and sensitivity was achieved. An example of ACS-based HRV detection is given in Fig. 5. Fig. 5b(1)&(2) plot the coefficient weights (H_c) of the normal cardiac sounds, where the peaks indicate the locations of S_1 and S_2. The S_1 localization through peak conditioning is plotted in Fig. 5b(2). The S_1S_1, S_2S_1 and S_1S_2 variations are plotted in Fig. 5b(3). It is worth mentioning that the various interval variations are consistent with the underlying theory.

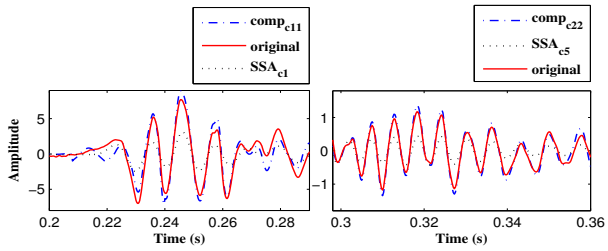
V. Discussions

NMF based decomposition of cardiac sounds provides excellent recovery of the normal cardiac sounds and murmurs, illustrated through various examples, see, e.g. Fig. 4. NMF based decomposition of cardiac sounds can add further prospectives to their analysis. For instance the coefficient matrix (H_c) which captures the timing information can be useful in the ACS-based HRV detection. The ACS-based HRV detection not only provides beat-to-beat variation but also variation within a beat which is hard to achieve with the ECG signals.

ACS-based HRV detection method is directly dependent

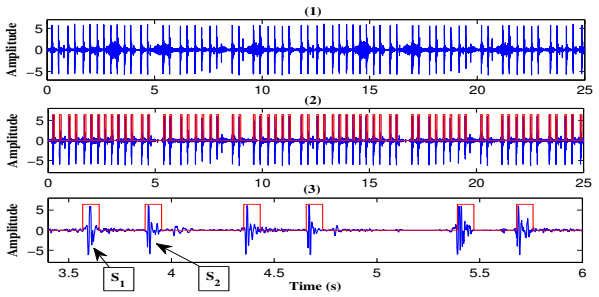


(a) The most important features, i.e. normal cardiac sounds (comp_{11}) and murmurs (comp_{22}) extracted from the mixed signal using proposed method are compared against the similar components ssa_{c1} and ssa_{c5} achieved though SSA.

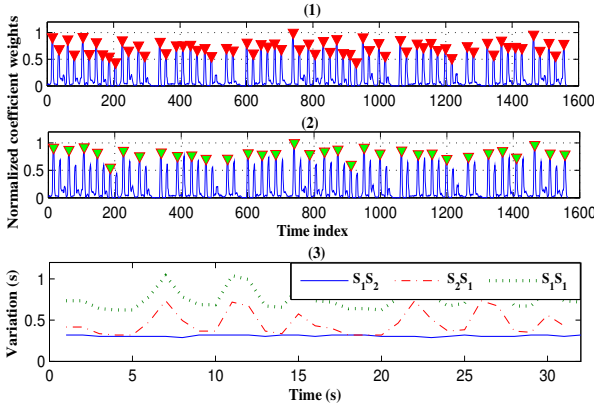


(b) Comparing the separated components with original mixture. The graph on the left side is zoomed over S_1, while the graph on the right shows a zoomed segment of the murmurs.

Figure 4: Comparison of some of the important features (components) derived using SSA and proposed method.



(a) Localization of cardiac sound components (S_1 and S_2). (1) shows original data sample, (2) shows separated normal cardiac sounds where various components are localized, and the graph in (3) is zoomed over a small segment of (2).



(b) Normalized coefficient weights of normal cardiac sounds, where peaks in (1) shows the locations of S_1 and S_2. The peaks in (2) show the location of S_1. The S_1S_1, S_2S_1 and S_1S_2 variations are illustrated in (3).

Figure 5: ACS-based HRV detection example.

on the localizations and differentiations of S_1 and S_2, which can be more robust using the recovered normal cardiac sounds shown in Fig. 5a(2)&(3) as compared to the mixture given in Fig. 5a(1). As argued in Section I, the highly nonstationary nature of the cardiac sounds which makes the localization and differentiation of S_1 and S_2 challenging, is one of the main reasons that limits the application of most of the existing automatic segmentation methods in the real scenarios. Therefore, to minimize the adverse effects in such scenarios, our proposed check-and-correct mechanism discussed in Section III-F, can be useful to verify and correct HRV detection by manually tweaking the various parameters (such as, e.g., smoothing window), if needed. For instance in Fig. 5b(3) the almost flat graph of S_1S_2 intervals and the variation of S_1S_1 which is almost directly proportional to the variation in S_2S_1, can verify the correct HRV detection.

Similar to the existing decomposition techniques which require the manual settings for various parameters, such as decomposition level, our method also needs these settings, which can be selected heuristically by analyzing large data. Nevertheless, the ultimate goal of this study is the development of a computer-aided auscultation method, allowing a physician to analyze and interpret the cardiac sounds visually. Hence the manual tweaking of different parameters can be useful for better results.

VI. CONCLUSIONS

In this paper we investigated the potential application of NMF in the recovery and analysis of cardiac sounds. A novel method for the separation of normal cardiac sounds and murmurs based on NMF is proposed. The simultaneous time-domain as well as frequency-domain features extraction via NMF can explore further analysis of the cardiac sounds. The novel HRV detection based on various cardiac sound components can incorporate additional aspects to the classical cardiac auscultation examination. The potential significance of the ACS-based HRV detection method can be twofold: 1) it can complement the existing short-time ECG signal-based HRV methods to improve HRV performance

in actual clinical settings, and 2) the solo ACS-based HRV detection method can eliminate expensive ECG setup in certain cost/time critical applications. The NMF-based cardiac sound analysis and feature extraction method is compared against the state-of-the-art analysis tools; experiments are performed on actual clinical data, which show significantly better performance. Our proposed method is particularly useful in scenarios where highly skilled professionals are not available, such as rural areas in underdeveloped countries.

ACKNOWLEDGMENTS

The authors would like to thank ELPEN (www.elpen.gr), for partially sponsoring this research work. The research work is also partially sponsored by eWALL research project funded by the European Commission under FP7 (www.ewallproject.eu). The opinions expressed in this work do not necessarily reflect those of the sponsors.

REFERENCES

[1] G. R. H. Sandercock, "The reliability of short-term measurements of heart rate variability," *International Journal of Cardiology*, vol. 103, pp. 238–247, Sept. 2005.

[2] H. Kamran, I. Naggar,F. Oniyuke, M. Palomeque, P. Chokshi, L. Salciccioli, M. Stewart, and J. M. Lazar, "Determination of heart rate variability with an electronic stethoscope," *Clinical automatic research*, vol. 23, pp. 41–47, Feb. 2013.

[3] L. S. Lilly, Ed., *Pathophysiology of heart disease*, 5th ed. Lippincott Williams & Wilkins, 2011.

[4] T. Rice, "learning to listen: auscultation and the transmission of auditory knowledge," *Jour. of the Royal Anth. Inst.*, pp. s41–s61, 2010.

[5] A. F. Pease, "If the heart could speak," *Pictures of the future*, pp. 60–61, 2001.

[6] S. Yuenyong, A. Nishihara, W. Kongprawechnon and K. Tungpimolrut3, "A framework for automatic heart sound analysis witwith segmentation," *Biom. Eng. Online*, p. 10(13), 2011.

[7] Y. N. Wen, A. P. Lee, F. Fang, C. N. Jin and C. M. Yu, "Beyond auscultation: Acoustic cardiography in clinical practice," *Int. Jour. of Cardiology*, pp. 548–560, 2014.

[8] B. S. Emmanuel, "A review of signal processing techniques for heart sound analysis in clinical diagnosis," *Jour.of Med. Engg. and Tech.*, vol. 36(6), pp. 303–307, 2012.

[9] H. Liang, S. Lukkarinen and I. Hartimo, "Heart sound segmentation algorithm based on heart sound envelogram," in *in Proc. of Computers in Cardiology*, Sept. 1997, pp. 105–108.

[10] S. Ari, K. Hembram and G. Saha, "Detection of cardiac abnormality from pcg signal using lms based least square svm classifier," *Int. Jour. of Expert Systems with Applications*, pp. 8019–8026, 2010.

[11] F. Ghaderi, H. R. Mohseni and S. Sanei, "Localizing heart sounds in respiratory signals using singular spectrum analysis," *IEEE Trans. on Biomed. Eng.*, vol. 58, no. 12, pp. 3360–3367, Dec. 2011.

[12] S. Sanei, M. Ghodsib and H. Hassanib, "An adaptive singular spectrum analysis approach to murmur detection from heart sounds," *Medical Engineering and Physics*, vol. 33, pp. 362–367, 2011.

[13] C. D. Papadaniil and L. J. Hadjileontiadis, "Effcient heart sound segmentation and extraction using ensemble empirical mode decomposition and kurtosis features," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, pp. 1138–1152, July 2014.

[14] Y. Chen, S. Wang, C-H. Shen and F. K. Choy, "Matrix decomposition based feature extraction for murmur classification," *Medical Engineering and Physics*, vol. 34, pp. 756–761, 2012.

[15] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, pp. 788–791, 1999.

[16] T. Virtanen, "Monaural sound source separation by non-negative matrix factorization with temporal continuity and sparseness criteria," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, no. 15, pp. 1066–1074, 2007.

[17] G. Shah and C. Papadias, "Blind recovery of cardiac and respiratory sounds using non-negative matrix factorization & time-frequency masking," in *Proc. of IEEE 13th Int. Conf. on BioInfo. and BioEng. (BIBE), Crete*, Nov 2013, pp. 1–5.

[18] C. Lin and E. Hasting, "Blind source separation of heart and lung sounds based on nonnegative matrix factorization," in *Proc. of Int. Symp. on Intel. Sign. Proces. and Comm. Syst. (ISPACS)*, Nov 2013, pp. 731–736.

[19] G. Shah, P. Koch, and C. Papadias, "On the blind recovery of cardiac and respiratory sounds," *IEEE Journal of Biomedical and Health Informatics*, to be published, early Access.

[20] L. J. Hadjileontiadis and S. M. Panas, "Adaptive reduction of heart sound from lung sounds using fourth-order statistics," *IEEE Trans. on Biomed. Eng.*, vol. 44, no. 7, pp. 642–648, July 1997.

[21] P. Y. Chen and I. W. Selesnick, "Translation-invariant shrinkage/thresholding of group sparse signals," *Signal Processing*, vol. 94, pp. 476–489, Jan 2014.

[22] M. Kim, J. Yoo, K. Kang and S. Choi, "Nonnegative matrix partial co-factorization for spectral and temporal drum source sepration," *IEEE Jour. of Sel. Topics in Sig. Pro.*, vol. 5, no. 6, pp. 1192–1204, Oct. 2011.

[23] G. Shah and C. Papadias., "Separation of cardiorespiratory sounds using time-frequency masking and sparsity," in *Proc. of 18th Int. Conf. on Digital Signal Processing (DSP), Greece*, July 2013, pp. 1–6.

[24] P. Bentley, G. Nordehn, M. Coimbra and S. Mannor, *The PASCAL Classifying Heart Sounds Challenge (CHSC2011) Results*, Std., 2011. [Online]. Available: http://www.peterjbentley.com/heartchallenge/index.html