Low Complexity Algorithm for Heart Sound Segmentation using the Variance Fractal Dimension

P. Carvalho†, P. Gil†, J. Henriques†, L. Eugénio‡, M. Antunes‡
†Centre for Informatics and Systems, University of Coimbra, Coimbra, Portugal
‡ Centre of Cardio-thoracic Surgery of the University Hospital of Coimbra, Coimbra, Portugal
Email: {carvalho, pgil, jh}@dei.uc.pt, antunes.cct.huc@sapo.pt

***Abstract* – This paper presents an algorithm for S1 and S2 heart sound segmentation using variance fractal dimension. Heart sound is assumed as a non-stationary signal embedding two main sounds S1 and S2, murmurs and eventually unusual ambient sound. The variance fractal dimension is applied to adaptively identify the boundaries of sound lobes. S1 components are detected using QRS synchronization while for S2 components a non-supervised classification approach is applied, based on temporal features of the lobes. This allows a 2-lead ECG signal to be used for the task. Some preliminary results are presented using recorded heart sounds taken a few days after valve replacement.**

Keywords – Heart Sound Segmentation, Variance Fractal Dimension, Clustering.

I. Introduction

Heart disorders are the primary cause of death in industrialized countries. The solution to this health problem is believed to be changing the focus from curative healthcare to preventive healthcare, i.e., controlling costs (social and economical) by reducing preventable healthcare conditions. In this sense long term tele-monitoring is a promising tool to achieve the aforementioned goal. In order to be cost effective and usable for long time periods, these tools require intelligent systems to be able to autonomously perform diagnostic functions and to support users in solving problems, hence requiring low computational algorithms that could be run in real-time using low power processing devices.

This paper introduces a low complexity algorithm for heart sound segmentation into clinically relevant lobes. This segmentation method is the basis for a tele-monitoring system, currently being developed, which aims at reducing risks associated with detection of late prosthetic heart valve dysfunction by means of a regular monitoring of timbre changes of implanted prosthetic valves as well as identification of new heart murmur development. A number of clinical disorders, namely heart disorders, can be diagnosed using auscultation techniques. Experienced cardiologists are able to detect subtle heart disorders just by listening to the timbre and the sequence of its beats and murmurs. In fact, there are several potentially deadly heart diseases, such as native and prosthetic heart valve dysfunction, where the heart sound auscultation is one of the most reliable and successful tool for early diagnosis [10]. In order to exploit the capabilities of automatic heart disorder diagnostic tools based on the heart sound, it is imperative to first carry out a segmentation for the recorded sound into a clinically meaningful sound segments or lobes. Diagnostic algorithms are then designed to identify and to assess the degree of illness based on dedicated features extracted from these lobes.

Several researchers have suggested methods for heart sound localization [8][12][13][14]. Usually in the literature, heart sound localization is addressed in order to reduce their influence in lung sound signal recordings. Nevertheless, applications can be found where the problem of heart sound component separation is addressed. Heart sounds are recognised by several means of signal processing and statistics, e.g. Wavelet decomposition methods [3], Hidden Markov models [1], decimation methods [1], linear and high order statistical methods [4], wavelet packet decomposition [5], S-transform [2], Time-frequency analysis [6] or involving artificial neural networks [7]. Even though they already might achieve significant results in identifying some particular sounds, it is recognized that these algorithms tend to be computationally very demanding and/or require supervised learning approaches, therefore not being suited for real-time heart sound segmentation in low power processing devices. Furthermore, since heart sounds, namely those produced by prosthetic valves, are highly dependent on the valve type, surgery technique, location and body morphology, supervised learning approaches tend to induce cumbersome setup procedures for each individual patient.

The heart sound segmentation method described herein is based on the variance fractal dimension (VFD) to achieve low-level sound segmentation. Using the ECG as a reference signal, Q components of the ECG are identified in order to classify the systole sound (S1) using a nearest neighbour approach. In order to keep system sensors and complexity to a minimum, a 2-lead ECG is used instead of higher complexity configurations. Under this setup it is not possible to identify T-wave components which mark the start of the diastole of the heart cycle. Even for a 3-lead ECG the existence of clear T-waves may not be assumed for all lead positions. In order to classify the diastole sound components (S2) a classification approach is applied based on temporal features.

The paper is structured as follows: in section 2 the proposed algorithm for heart sound segmentation into coherent sound lobes is described. In section 3 the problem of features extraction and heart sound classification are addressed and some results discussed. In section 4, obtained results are analysed. Finally, in section 5 the main conclusions are drawn and some future research directions are pointed out.

II. Segmentation of Heart Sound into Lobes

The main stages of the proposed segmentation algorithm for sound lobes identification are depicted in Figure 1. After heart sound acquisition and quality assessment, the signal is first high-passed to eliminate inaudible components, which may be induced by slow movements, such as chest and muscle movements, during sound recording. In the current implementation of the algorithm a fourth order Butterworth filter with a cut-off frequency of 25Hz is utilized for this propose. In the second stage the variance fractal dimension of the filtered signal is computed using several scale resolutions. It should be noted that significant sound lobs should exhibit persistency at fine and coarse detail scales. This observation will be one of the criteria applied during the lob validation stage. After low pass filtering, the variance fractal dimension is applied to identify sound lob boundaries. Finally, some criteria are then applied to reject false sound lobes.

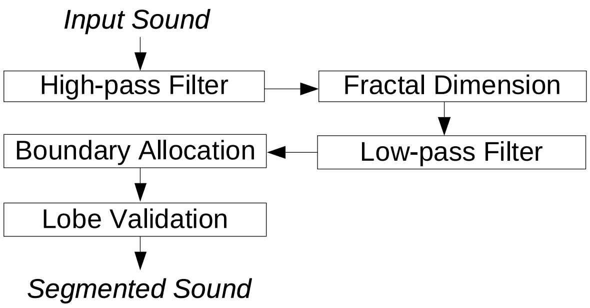

Figure 1: Block Diagram of the proposed algorithm.

Prior to segmentation the sound quality is validated. In the current implementation, this validation is performed using the correlation between the spectral power distributions (SPD) of the sound of each heart cycle with respect to a reference SPD. The heart cycle is considered to be defined between each consecutive pair of identified Q components of the ECG. A threshold approach is followed to discard noisy sound segments, which exhibit a correlation lower than 0.995 with respect to the reference SPD. The reference heart cycle is computed from the first 10 heart cycles acquired, i.e., the cycle whose SPD exhibits the largest average correlation in the set among all cycles in the set. Furthermore, if more than 2 cycles exhibit an average correlation below 0.995 the entire set is considered noisy and discarded. This simple procedure is able to detect noisy sound segments due to coughing, speaking, stethoscope movements and ambient noises.

A. Variance Fractal Dimension

Heart sounds exhibit a set of properties which suggest they are fractal in nature [8]. First, these signals do not self-cross. Second, these signals exhibit quasi-periodicity, since they emerge from natural biological processes, i.e., heart beats. Furthermore, they are self-affine, since, in order to scale them, a different scaling factor is required for each axis [16]. This suggests that fractal dimension can be utilised to properly characterise and analyse these signals.

Fractal dimension quantifies the complexity of a pattern or the information embodied in the pattern in terms of morphology, entropy, spectra or variance [15]. In this work the variance fractal dimension is utilised, since it enables real time computation.

VFD has found several applications in signal analysis. For instance, in [8] VFD is used to locate heart sounds in lung sound recordings, Lazareck and Moussavi [17] have developed an algorithm based on VFD for swallowing sound segmentation, whereas Yap and Moussavi [18] applied the same analysis tool for respiratory onset detection. In another field of application, Hall et al. [9] utilize VFD for the detection of transient in radio frequency fingerprinting.

In deriving the variance fractal dimension, the Hurst exponent is computed based on the power law relation which exists between the variance of the signal's amplitude increments over time increments (see (1)) [8][15].

$$\operatorname{var}\left(\Delta x_{\Delta t}\right) \approx \Delta t^{2H} \tag{1}$$

Using the result from equation (1) it is observed that the Hurst exponent may be obtained from (2).

$$H = \lim_{\Delta t \to 0} \frac{1}{2} \frac{\log\left(\operatorname{var}\left(\Delta x_{\Delta t}\right)\right)}{\log\left(\Delta t\right)} \tag{2}$$

In the above equations, x and t represent, respectively the signal and time, whereas Δ stands for increment, i.e.,

$$\Delta t \equiv t_{i+1} - t_i \tag{3}$$

$$\Delta x_{\Delta t} \equiv x\left(t_{i+1}\right) - x\left(t_i\right) \tag{4}$$

From the above equations it is observed that the variance dimension is a measurement calculated by analysing the spread of the increments in the signal amplitude in the time domain. This spread is indicative of the multifractal richness in the signal, e.g., a unifractal object yields a flat line. The VFD ($D\sigma$) for a process with embedding Euclidian dimension E is computed as in equation (5),

$$D\sigma = E + 1 - H \tag{5}$$

Since it is known that the variance in heart sound lobes and channel noise differ significantly, heart sound boundaries should be clearly marked by accentuated changes in variance. Therefore, these boundaries should be captured from the variance fractal dimension. To accentuate this effect, in this work the variance fractal dimension is determined from the energy of *x* rather from *x* itself, i.e., from $x \leftarrow x^2$.

Let N_T be the size of the sliding centred analysis window. This window is chosen according to known important heart sound characteristics as will be explained later in the next section. For each window, VFD is computed for time increments $\Delta t_k \equiv t_{i+k} - t_i$. The value k represents an integer chosen such that each window contains a number of $N_k = \mathrm{int}\left(N_T/k\right)$ of Δt_k increments. In practice k is selected such that there is a reasonable separation between data samples for Δt_k and, on the other hand, N_k is sufficiently large for variance calculation. Finally, for each analysis window the variance is determined using the likelihood approach.

To obtain each VFD the previous analysis window is shifted by $p_{shift} \leq N_T$, and the VFD is calculated for this new window and set of samples. It should be noted that the effect of N_T is similar to scale in multi-scale analysis. In fact, as N_T increases it is observed that features tend to be blurred [8][15]. Therefore, N_T can be utilised as the analysis scale to devise analysis approaches common to multi-scale signal processing. For instance, significant features tend to be persistent at fine and coarse detail analysis scales. Hence, to discriminate between significant and less significant features, several N_T (small for fine analysis and large for coarse analysis) may be applied to determine their persistency.

As in multi-scale analysis it is observed that for coarser scales features tend to be shifted, leading to less accurate localization. Furthermore, for small windows VFD tends to oscillate due to redundant calculations [8]. To avoid these effects, in this work a low-pass filter (implemented by a *sinc* filter) is applied for finer scales, enabling the elimination of high frequency variations in VFD, i.e., the elimination of spurious features and VFD oscillation, while keeping localization of relevant ones.

B. Finding heart sound boundaries using VFD

To identify boundaries in the significant heart sound segments VFD is computed using two distinct scales N_T^c (coarse scale) and N_T^f (fine scale), being

$$N_T^f \equiv \frac{N_T^c}{10} \tag{6}$$

Regarding N_T^c it is defined based on the observed average duration of S1, S2 and murmur sound segments, and the average distance between the first and second heart sounds. The distance between the first and second heart sounds is typically $200ms \leq N_T^c \leq 400ms$, while the duration of S1 and S2 is usually around $100ms$. Therefore, N_T^c is chosen to be $N_T^c = 100/t_s$. This ensures the analysis window covers the main peaks, while avoiding an analysis window which includes multiple peaks. For sound signals sampled with a 44.1kHz sampling frequency it was experimentally determined that $64 \leq k \leq 128$. In each iteration the p_{shift} was selected such that the sliding window exhibits a 50 percent overlap, i.e., $p_{shift} = N_T/2$. Finally, the cut-off frequency of the *sinc* low-pass filter applied for the VFD computed with N_T^f was experimentally selected to be $0.05/ts$.

Figure 2 depicts the scaled VFD values determined with N_T^f and N_T^c, respectively. As can be observed, sound segment boundaries are clearly identified by the slope and slope change (second derivative) of VFD and could readily be identified using a threshold procedure. However, it was observed that for heart sounds, significant segments exhibit around 50% of the overall area of VFD if a sufficiently large window is applied. Hence, sound segment boundaries are identified by the zero crossings of y, where ($\langle\cdot\rangle$ - average operator)

$$y \equiv VFD - th_{Segment} \tag{7}$$

$$th_{Segment}^k = \alpha\left\langle VFD^k\right\rangle + (1-\alpha)\, th_{Segment}^{k-1} \tag{8}$$

$$th_{Segment}^1 = \left\langle VFD^1\right\rangle \tag{9}$$

As can be observed from equations (8) and (9), a convex combination is applied in order to update the threshold for each contiguous sound window k identified during noise suppression. In the current implementation $\alpha = 0.9$.

Significant sound segments are characterised by higher VFD, since their variance is greater than the variance of the channel's noise. Hence, significant sound segments exhibit $y \geq 0$.

C. Sound segment validation

As mentioned above S1, S2 and murmur sound segments exhibit characteristic durations. Furthermore, significant sound segments should be persistently identified for different analysis scales (defined by the analysis windows). This is an useful observation to discriminate between valid and non-valid murmur segments (see Figure 2). In fact, in many situations murmur segments highly resemble channel noise. Using these principles the following criteria are applied to each of the identified sound segments for validation. Let $S_i \equiv x\left(j \cdot ts\right)$, $j = n_i^{start}, \ldots, n_i^{stop}$, be the identified sound segment. Segment S_i is considered a valid heart sound segment if the following criteria are verified:

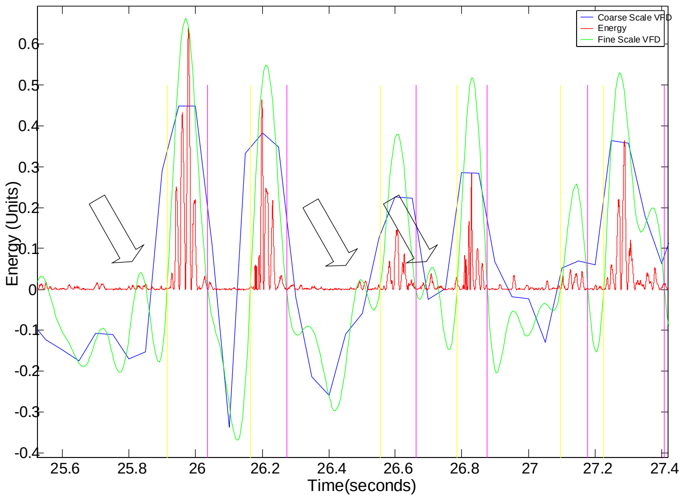

Figure 2: Identification of invalid sound lobes that are not segmented. Arrow show invalid lobes.

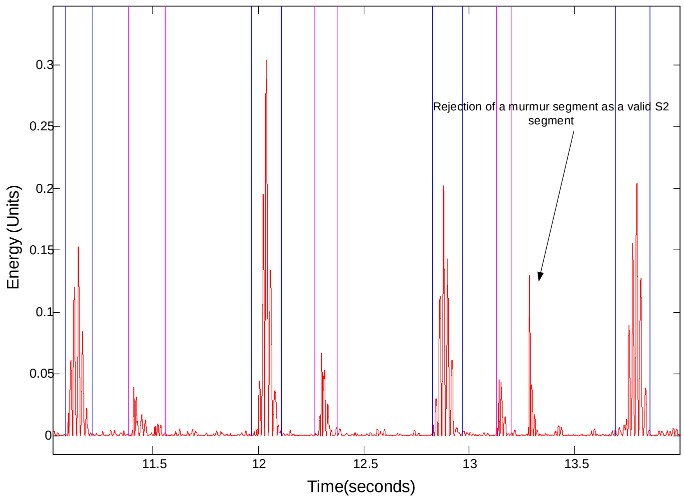

Figure 3: Example of achieved S2 identification using the defined approach. This example shows a patient with a valve prolapse. The sound energy instead of the sound wave is shown in this figure.

Duration limits:

$$t_{\min} \leq ts\left(n_i^{stop} - n_i^{start}\right) \leq t_{\max} \tag{10}$$

Fine-coarse scale support:

$$\exists j \in \left\{n_i^{start}, \ldots, n_i^{stop}\right\} : VFD\left(N_T^C, j\right) - \left\langle VFD\left(N_T^C\right)\right\rangle \geq 0 \tag{11}$$

In equation (11) $VFD\left(N_T^C, j\right)$ stands for the VFD value computed for point j using the coarse scale analysis window, while $\left\langle VFD\left(N_T^C\right)\right\rangle$ represents the average VFD value. In the current implementation of the algorithm $t_{\min} = 10ms$ and $t_{\max} = 300ms$ (average value of the S1-S2 between time). Some results of this procedure are illustrated in figure 3.

III. Classification of Heart Sound Lobes

Classification of segmented lobes is done based on their features and ECG QRS complexes. In this section the classification of the extracted sound segments is introduced. To design the classifier a two temporal related features are extracted for each identified sound segment:

S1-S2 Duration: Sound lobes of type S1 and S2 exhibit characteristic duration ranges. This is also true for the between sound duration period, namely for the interval between systolic and diastolic sounds. This time interval tends to be very regular, even for patients suffering arrhythmia.

Root Square Amplitude: RMS is a measure of the loudness of a window. This feature is unique to segmentation, since changes are important cues. Namely, S1, S2 and murmur sounds tend to exhibit a characteristic loudness for each individual and acquisition location. Since, for each data set these parameters are kept constant, loudness is a valuable cue to distinguish between these sounds as well as to identify sound segments corresponding to noise.

D. Classification

Having successfully figured out the boundaries of sound lobes of type S1 and S2 as well as major murmurs and noisy segments, the next step is to classify these segments accordingly to a predefined set of classes, i.e., S1, S2 and murmur or noise, if present.

S1 segments are identified using ECG Q component synchronization, i.e., the lobe closer to each Q component of the ECG is assumed to be a S1 lobe, systole occurs during this time period.

In order to detect S2 segments, a fuzzy C-means classifier in two classes is applied. The class that corresponds to the S2 sound component is chosen as the one which exhibits the most compact support and highest regularity with respect to the S1-S2 time interval. The motivation for this last criterion stems from the fact that, even for patients suffering from arrhythmia, S1-S2 time interval tends to be the most regular time feature. Here compactness is assessed using the class standard deviation.

After S2 class identification, outliers are removed from this class using the following criterion: let C be the centre of the identified class. For each pair of consecutive S1 sound components it is verified if only one sound lobe exists. If there exist more than one candidate segment in the heart cycle for S2 then only the one which exhibits the smallest deviation of its S1-S2 duration with respect to C is selected.

A typical result obtained using this procedure is depicted in Figure 3.

For patients exhibiting murmurs or S3/S4 heart sound splitting (these sound components usually overlap S1 and S2) it is observed that, when these murmurs occur near the S1 or S2 boundaries they might be included into these segments. In order to account for this situation the VFD of the identified segments is subsequently analysed. Should local minima in the VFD occur inside the identified segments these are used to split the segment into sub-segments. Results obtained with this procedure are depicted in Figure 4.

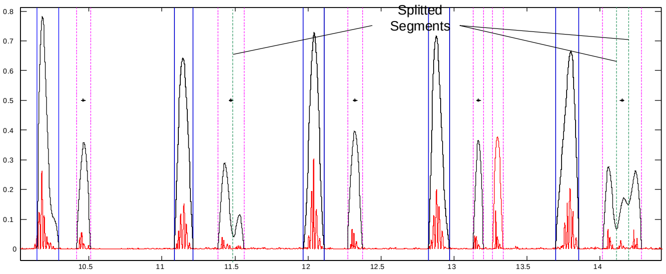

Figure 4: Identified S2 segment splitting boundaries. S2 segments are identified with the symbol '*' located inside the dashed vertical boundary lines. S1 segment boundaries are identified by continuous vertical lines. The show signal envelops correspond to the computed VFD values. The sound energy instead of the sound wave is shown in this figure.

IV. RESULTS

In Figure 5 some results obtained using the proposed segmentation methodology are shown for three distinct heart sounds taken, respectively, from a patient with a bio-prosthetic valve implant and from two patients with an aortic and a mitral mechanical valve implant. These phonocardiogram (PCG) records were taken from patients a few days after valve replacement (implant) using an electronic stethoscope presenting a flat frequency response from 20 Hz to 20 kHz (Meditron Stethoscope System) connected to a laptop computer. The sampling rate used in the acquisition was chosen as 44.1 kHz with a 16 bits resolution. Prior to the acquisition of each sound sample, the ideal position for the stethoscope was selected by an experienced cardiologist.

As can be observed from table I, for all samples correct detection results were always above 90% of the S1 and S2 components. The average detection efficiency was 96.19% for S1 and 96.23% for S2 heart sounds.

An interesting case is presented in Figure 5 middle. This patient exhibits non constant heart rhythm. Nevertheless, as can be observed, the S1-S2 interval is relatively regular. Furthermore, this patient presents some heart murmurs that, for some heart cycles, are undistinguishable in terms of loudness from the S1 component. However, as can be seen, due to the synchronization process using *Q*-wave, the algorithm is able to distinguish these segments from S1 and S2 segments. It should be noted that these conditions are quite common for patients with heart diseases, namely patients with prosthetic valve implants. This is also a function of heart sound acquisition locations. Furthermore, these patients frequently exhibit or tend to develop atrial or ventricular arrhythmias, which induce heart sounds of different characteristics. These sounds can interfere in diagnosis algorithms based on heart sound. Using the proposed algorithm, it is observed that these pathologies are easily identifiable using the ECG and, therefore, can be filtered during the sound validation stage. Solving the aforementioned situations would be a difficult task for algorithms where segment classification is not based on the ECG.

In general, it is observed that the proposed algorithm is able to detect S1 and S2 components as long as no significant regular noisy segments exist in the sound samples given as input to the segmentation algorithm. Since the classification stage relies only on temporal features, under these circumstances the algorithm would not be able to distinguish between valid S1 and S2 segments and the noisy ones, i.e., a wrong estimation of the S1-S2 interval could induce misclassification. Furthermore, this could affect the sound lobe boundary estimation accuracy in case of long duration noisy sound segments. Fortunately, these noisy sound segments (e.g. external noises, coughing, speech, etc.) tend to exhibit very different SPDs from regular heart sounds and are therefore captured using the sound validation strategy introduced in section II.

Sound	S1	S2	S1&S2
Mitral Mechanical Prosthesis	34/37 (91.9%)	35/38 (94.6%)	2/4
Bioprosthetic Valves	23/23 (100%)	23/23 (100%)	2/0
Aortic Mechanical Prosthesis	44/45 (97.8%)	44/45 (97.8%)	1/1

Table 1: Results for three sound samples. Column S1 – number of correctly detected S1 components/number of real S1 components in sound sample. Column S2 – number of correctly detected S2 components/number of real S1 components in sound sample. Column S1&S2 - number of missed S1 and S2 components/number of misclassified S1 and S2 components in sound sample.

V. CONCLUSIONS

The problem of heart sound segmentation and classification using low complexity methodologies was addressed in this paper. The variance fractal dimension, which is a measure of signal complexity, is here applied in the segmentation of heart sounds. In order to clearly detect the boundaries of segments two distinct time scales are considered, which are based on the observed average time duration of relevant heart sound segments. The classification lobes is based on ECG - QRS complex and using a fuzzy C-means classifier. Results show clearly the inherent potential of the proposed methodology.

ACKNOWLEDGEMENT

This project was partially financed by the IST FP6 project IST-2002-507816 supported by the European Union and POSI (Programa Operacional da Sociedade de Informação) of the Portuguese government supported by the European Union.

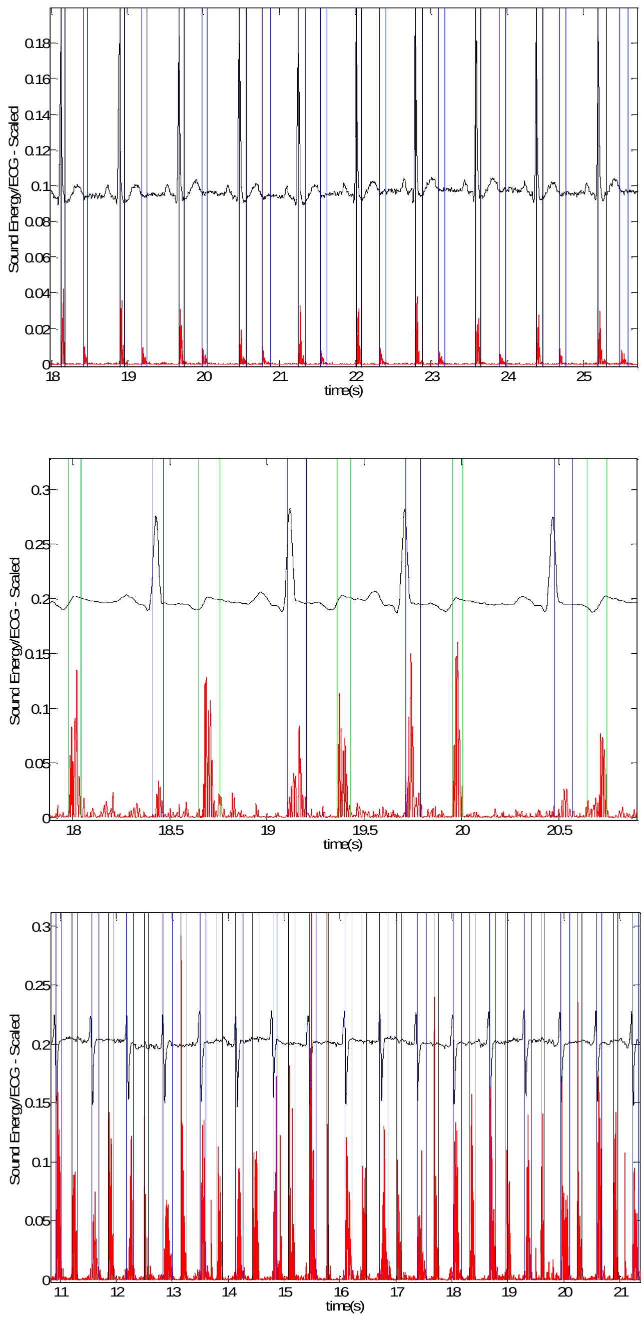

Figure 5: S1 and S2 identification results (ECG signal shown for reference). (top) Biological valves. (middle) Mechanical prosthesis in mitral position. (bottom) Mechanical prosthesis in aortic position.

REFRENCES

[1] M. El-Hanjouri, W. Alkhaldi, N. Hamdy, O. Abdel Alim, "Heart Diseases Diagnosis using HMM", IEEE MELECON 2002, Cairo, Egypt, May 7-9, 2002.

[2] G Livanos, N Ranganatha,J Jiang, "Heart Sound Analysis Using the S Transform", IEEE Computers in Cardiology 2000, vol. 27, pages 587-590, 2000.

[3] Liang H., Harmito I., "A heart sound feature extraction algorithm based on wavelet decomposition and reconstruction", 20th Annual Int. Conf. of the IEEE, vol. 3, pp.1539 – 1542, 1998.

[4] Ergen, B., Tatar, Y," The analysis of heart sounds based on linear and high order statistical methods", 23rd Annual Int. Conf. of the IEEE, vol. 3, pp. 2139 – 2141, 2001.

[5] Liang H., Nartimo I, "A feature extraction algorithm based on wavelet packet decomposition for heart sound signals" Proceedings of the IEEE-SP International Symposium, pp. 93 – 96, 1998.

[6] El-Asir B., Khadra, L., Al-Abbasi A.H., Mohammed M.M.J, "Time-frequency analysis of heart sounds", IEEE TENCON. Digital Signal Processing Applications, vol. 2, pp. 553 - 558, 1996.

[7] Curt G. DeGroff, Sanjay Bhatikar, Jean Hertzberg, Robin Shandas, L. Valdes-Cruz, R. L. Mahajan, "Artificial Neural Network–Based Method of Screening Heart Murmurs in Children", Circulation, vol. 103, pp. 2711-2716, 2001.

[8] Gnitecki J., Moussavi Z., "Variance Fractal Dimension Trajectory as a tool for Heart Sound Localization in Lung Sounds Recording", 25th Annual Int. Conf. of the IEEE, vol. 3, pp. 2420 – 2423, 2003.

[9] J. Hall, M. Barbeau, E. Kranakis, "Detection Of Transient in Radio Frequency Fingerprinting Uning Signal Phase",

[10] G. Mintz, E. Carlson, M. Kolter, "Comparison of noninvasive techniques in evaluation of the nontissue cardiac valve prothesis", J. Med. Eng. Technol., 15(6), pp.222-231, 1991.

[11] R. Charleston, M. R. Azimi-Sadjadi, R. Gonzalez-Camarena, "Interference Cancellation in Respiratory Sounds via a Multiresolution Joint Time-Delay and Signal Estimation Scheme," IEEE Trans. Biomed. Eng., vol. 44, no. 10, pp. 1006-1019, 1997.

[12] L.J. Hadjileontiadis, S.M. Panas, "Adaptive reduction of heart sounds from lung sounds using fourth-order statistics", IEEE Trans. Biomed. Eng., vol. 44, no. 7, pp. 642-648, Jul. 1997.

[13] M. Kompis, E. Russi, "Adaptive heart-noise reduction of lung sounds recorded by a single microphone," 14th Conf. IEEE Engineering in Medicine Biology Society, EMBC'92, pp. 691-692, 1992.

[14] J. Gnitecki, Z. Moussavi, H. Pasterkamp, "Recursive least squares adaptive noise cancellation filtering for heart sound reduction in lung sounds recordings," 25th Conf. IEEE Engineering in Medicine and Biology Society, EMBC'03, 2003.

[15] W. Kinsner, "Fractal and chaos engineering," Lecture Notes, University of Manitoba, 2003, ch.2 and ch. 7 part 3.

[16] H. O. Peitgen, H. Jürgens, and D. Saupe, "Chaos and Fractals: New Frontiers of Science". New York, NY: Springer-Verlag, 1992, pp. 491-496.

[17] Lazareck L, Moussavi Z., "Classification of Normal and dysphagic Swallowing Sounds by Acoustical Means", IEEE, Trans. Biomed. Eng., vol. 51, no. 12, pp. 2103-2112.

[18] Yap Y., Moussavi Z., "Acoustical Airflow Estimation From Tracheal Sound Power", Proc. IEEE Canadian Conf. Elec. Comp. Eng. (CCECE), pp. 1073-76, 2002.