Exploring Embedding Matrices and the Entropy Gradient for the Segmentation of Heart Sounds in Real Noisy Environments

Jorge Oliveira, Ana Castro, Miguel Coimbra, *IEEE Member*

Abstract— **In this paper we explore a novel feature for the segmentation of heart sounds: the entropy gradient. We are motivated by the fact that auscultations in real environments are highly contaminated by noise and results reinforce our suspicions that the entropy gradient is not only robust to such noise but maintains a high sensitivity to the S1 and S2 components of the signal. Our whole approach consists of three stages, out of which the last two are novel contributions to this field. The first stage consists of typical pre-processing and wavelet reconstruction to obtain the Shannon energy envelogram. On the second stage we use an embedding matrix to track the dynamics of the system, which is formed by delay vectors with higher dimension than the corresponding attractor. On the third stage, we use the eigenvalues and eigenvectors of the embedding matrix to estimate the entropy of the envelogram. Finite differences are used to estimate entropy gradients, in which standard peak picking approaches are used for heart sound segmentation. Experiments are performed on a public dataset of pediatric auscultations obtained in real environments and results show the promising potential of this novel feature for such noisy scenarios.**

I. Introduction

Cardiac auscultation is the simplest, fastest and cheapest method for heart examination. It provides information about the structural and functional characteristics of the heart using a simple medical device, the stethoscope. Electronic versions of this device are capable of registering and optimizing the quality of auscultation signals, generating what is called a phonocardiogram (PCG) signal. Data collection systems such as DigiScope [1] can obtain PCG signals such as the one depicted in Figure 1. Here we can observe the various components of a heart cycle, including S1 (first heart sound) and S2 (second heart sound). These establish the boundaries of the other two fundamental components of a heart cycle: the systole (period between S1 and S2), and the diastole (period between S2 and S1). S1 and S2 are generated by the opening and closing of the various heart valves and in some auscultations we have the presence of additional sounds such as S3, S4 or murmurs. [2]

Heart sounds segmentation is a fundamental step for extracting useful physiological information from heart sounds. It is typically preceded by a pre-processing step, and followed by feature extraction and classification steps in traditional statistical pattern recognition approaches. The PCG signal is usually pre-processed in order to extract the heart sound envelogram (HSE). Liang et al [3] presented different non-linear methods to calculate the HSE, demonstrating that the Shannon energy emphasizes the medium intensity signal and attenuates the effect of low and high intensity signal. However, in noisy environments, the HSE still keeps high noise levels, making it difficult to extract useful heart sounds. Other approaches include the use of wavelet decomposition and reconstruction [4], [5], but for higher heart rates and noisy environments, differentiating between heart sound components is still a challenge [6]. Since time-based features are not enough to differentiate the heart sound components different approaches are proposed using the frequency content of the heart sound components [7], or the entropy of the heart sound [8].

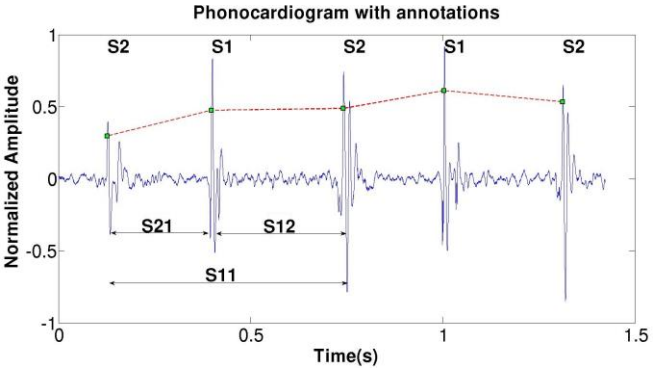

Figure1. A typical heart sound (S11) and its four main components: first and second heart sounds (S1 and S2), systole (S12), diastole (S21) and the entire heart cycle (S11)

Although PCGs are collected in clinical challenging environments, they still exhibit periodicity. Kumar et al. [6] detected heart murmurs using complexity signatures calculated directly on the PCG signal and results reported indicate that entropy is able to capture periodic signal components.

This paper is structured as follows: the proposed methodology is presented in Section II, followed by experimental results in Section III. A discussion concludes the paper in section IV.

II. Methods

The full proposed methodology (Figure 2) can be divided into three sections: data pre-processing follows previous research and state-of-the art (2.A), embedding matrices are proposed to better capture the dynamics of the signal when facing high levels of noise (2.B), and the entropy gradient

This work was partially funded by the Fundação para a Ciência e Tecnologia (FCT, Portuguese Foundation for Science and Technology) under grant Heart Safe, PTDC/EEI-PRO/2857/2012, and and Project I-CITY - ICT for Future Health/Faculdade de Engenharia da Universidade do Porto, NORTE-07-0124-FEDER-000068.

Jorge Oliveira, and Miguel T. Coimbra are with Instituto de Telecomunicações, Faculdade de Ciências da Universidade do Porto, Departamento de Ciências dos Computadores, Rua do Campo Alegre 1021/1055, 4169-007 Porto, Portugal, {oliveira_jorge;mcoimbra}@dcc.fc.up.pt

Ana Castro is with Faculdade de Engenharia da Universidade do Porto, ana.castro@dcc.fc.up.pt

(2.C) is then calculated and used as the signal in which peak picking is applied for segmentation (2.D).

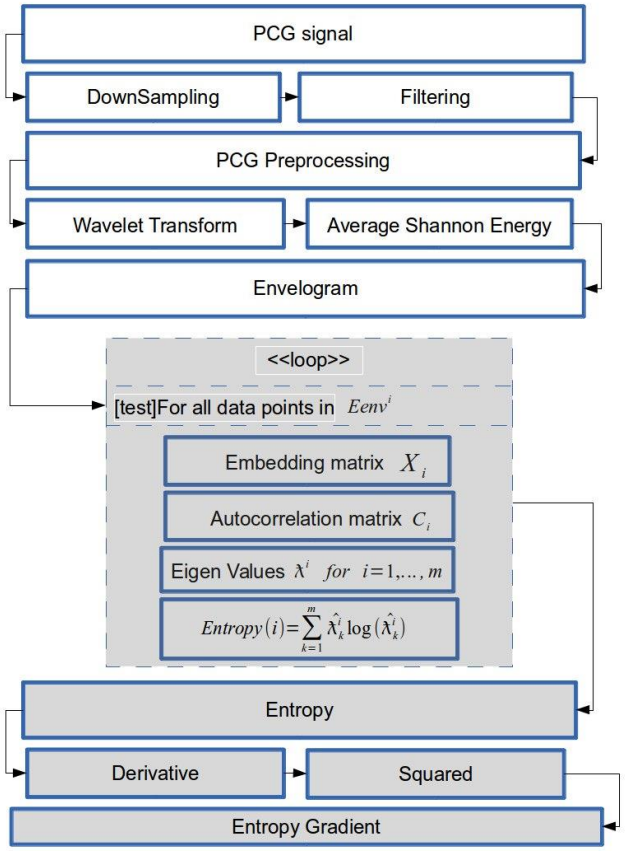

Figure 2. Structure of the proposed algorithm (novel sections in grey)

A. Signal Pre-Processing

The PCG is filtered with a band-pass, zero phase Butterworth filter order 6 (25-900Hz) to eliminate noise, decimated to 2kHz and finally normalized [9]. Afterwards, a wavelet transform using an order 6 Daubechies wavelet is applied to extract its approximation level 4 [6]. The resulting signal is normalized and then used to obtain the Shannon Energy envelope [3]. This procedure follows previous literature on this topic and due to space limitations we leave this motivation and explanation to the mentioned key references.

B. Embedding Matrices

Our motivation to use embedding matrices comes from the fact that heart cycles can vary in duration, not only between different persons but even for a single person. Furthermore, a sequence of samples can be used to quantify a variation but this is implicitly connected with the sampling rate used and the cycle duration, and not really to the fundamental structural pattern of the cardiac signal. Embedding matrices are a mathematical tool that can help us model non-linear data series and they can possibly be used to adequately model PCG signals. Assuming that the noise of real environments does not have relevant or repetitive structural elements, an interesting side-effect emerges, which provides us an additional layer of noise reduction.

The first step is to convert the envelope signal in a sequence of state vectors as it was demonstrated by the embedding theorem [10]. Let *x(t)* be the time series representing the heart sound envelogram. A new delay vector is formed in the i[th] state of observation $x_i = [x(i), x(i-\Gamma), \dots, x(i-(m-1)\Gamma)]$ (m is the dimension vector) by shifting Γ samples towards the right in the data series.

The difference in number of samples Γ between adjacent components of the delay vectors is equal to 2. If Γ is not close to 1 the segments constructed are formed by a sequence of data that may not correlated and loss of information is inevitable (except for Γ= 1). The number of delay vectors in a time series of N points is $P = N - \Gamma(m-1)$. All the delay vectors $x_i^j \in \mathbb{R}^m \quad j = 1,2,\cdots P$ have the same dimension independently of the observation state i. This is achieved by using a cyclic vector implementation as it is shown in Figure 3.

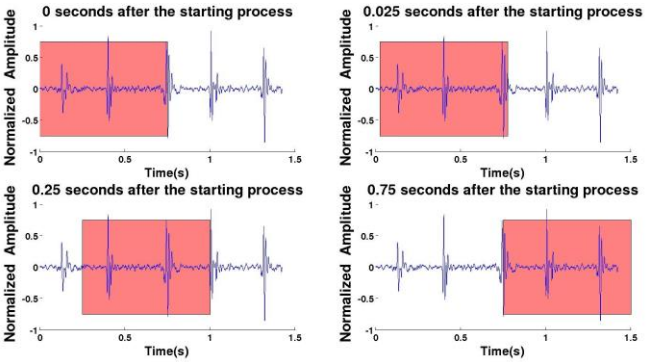

Figure 3. Using delay vectors of dimension m=(length(signal))/2+1 it is only possible to get one sample from the complete signal, therefore the rejection region *Rej* will be Rej=(length(signal))/2-1. This rejection region is the main cause of the variation observed in the signal entropy, since from one instance to another we are discarding one sample point (with some degree of predictability) and collecting a new sample point (with possible a different degree of predictability)

C. Entropy Gradient

The autocorrelation of each embedding matrix calculated in the previous section is calculated in (1).

$$C_i = X_i^T X_i \quad \text{where} \quad C_i \in \mathbb{R}^{(m \times m)} \tag{1}$$

C_i is a real symmetric matrix, its eigenvalues are real and its eigenvectors are orthogonal.

Let D_i be the diagonal matrix of the eigen values of C_i sorted in a descending order $\lambda_1 > \lambda_2 > \dots > \lambda_m$. The diagonal matrix D_i is the spectrum of the eigenvalues. The eigenvalues λ are the squared length of the semi-axis of the hyper-ellipsoid which best fits the cloud of data points and the corresponding eigenvectors give the direction of the axis in the hyper-ellipsoid [10]. The highest eigenvalues are associated with the most relevant structures of the signal. In contrast, the lowest eigenvalues are usually associated with very small variations and noise. Using the D_i vector, the *entropy* in the i^{th} state of observation is given by (2) [10].

$$H(i) = -\sum_{k=1}^{m} \lambda_k^i \log(\lambda_k^i) \tag{2}$$

The *gradient of entropy* ∇H is approximated with finite differences, using the two adjacent neighbors in each point of the function H. After differentiation ∇H is squared point by point, making all data points positive. Figure 4 shows the output of the envelogram (after 2.A), and the results obtained after the employment of the embedding matrices and the entropy gradient (2.C). Most noise appears to have been eliminated simply because noise is non-periodic and thus does not produce a significant difference in the values of entropy for different observation states.

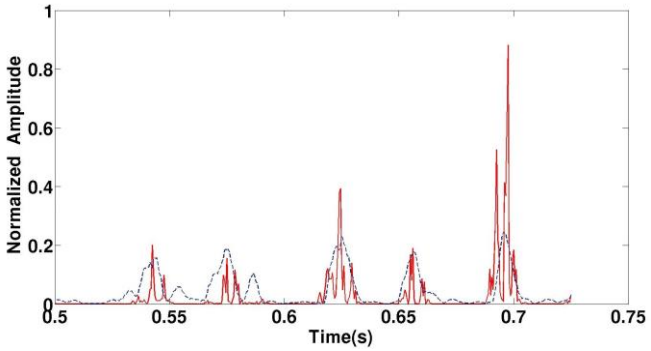

Figure 4. For a specific PCG signal we can observe its envelogram in green and its entropy gradient in red (using a time window of 25ms and a delay time equal to 2 samples)

Periodicity is mapped in the embedding matrix X. As we can observe in Figure 4, a shift in the state of observation $X_i \rightarrow X_{i+1}$ can severely change the spread distribution of the data points thus generating a large entropy gradient peak. This is mainly caused by the appearance or disappearance of S1 and S2 signals in the borders of the windows, causing severe variations in the pattern that is repeated. An example would be a window that has 3 S1 peaks and 3 S2 peaks. With a certain shift, we might lose one S1 peak and thus obtain a very different entropy measurement for a signal with 2 S1 peaks and 3 S2 peaks. Another expectable interesting observation is that the entropy gradient response to S1 and S2 are different mostly because S1 is more stable in terms of periodicity and exhibits a lower frequency content when compared to S2.

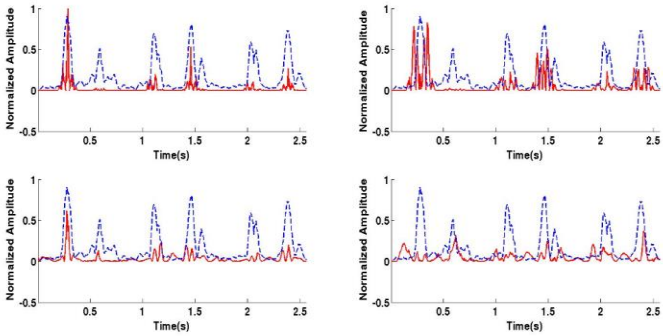

Figure 5. Output of the envelogram (blue) and entropy gradient (red) using different time windows: 25 ms (upper-left corner), 50 ms (upper-right corner), 150 ms (lower-left corner) and 250 ms (lower-right corner)

An important parameter is the dimension of the delay vector m of the embedding matrix. If m is too low then the benefit of using embedding matrices is negligible. On the other hand, if it is too high it will add redundancy and thus degrade the performance of the algorithm. An estimate of the dimension of the delay vector is calculated using the false neighbor method [10]. This method uses the idea of neighbors and two points are considered to be neighbors if their distance is less than a pre-defined threshold. Suppose that the dimension of the delay vector is m and it is projected into a lower dimensional m_0. If the number of false neighbors does not increase substantially from $m \rightarrow m_0$then it is acceptable to reduce the dimension of the delay vectors [10]. These results are plotted in the Figures 5 and 6, for a sample PCG signal.

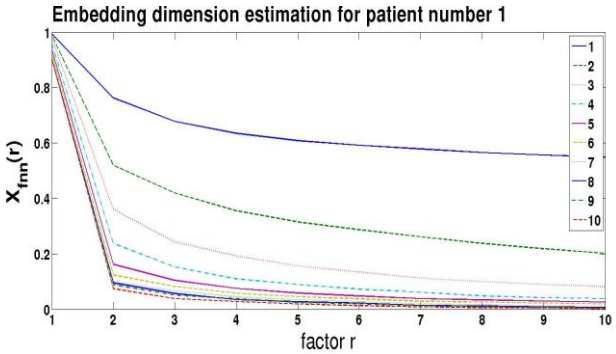

Figure 6. The relative number of false negatives plotted against the minimum distance to be considered as a possible neighbor

Observation of Figures 5 and 6 hints that a time window m of 25ms (5 samples) is enough. Although beyond the scope of this paper, a deeper inspection of how this varies when additional sounds are present (S3, S4, murmurs) is an interesting topic to be explored in the future.

D. Segmentation

A peak picking algorithm is used to search for candidates of heart sounds S1 and S2. Following previous literature [9] the algorithm uses a timing window of 50ms and only the maximum peak in it is considered a potential candidate for a peak. Peaks are only selected if their energy is superior to a 0.8 quantile of the total energy of the signal and if they are at least spaced 150ms from their neighboring peaks to ensure that heart sound splits are not erroneously detected.

III. RESULTS

A. Materials

To evaluate the segmentation algorithm performance using either entropy gradient or just the traditional envelogram signal we have used a public dataset from the Pascal 'Classifying Heart Sounds Challenge) [11]. This database includes PCGs from pediatric patients with their corresponding S1 and S2 annotated positions. This dataset was used because it has the best quality in terms of the annotation of the locations of S1 and S2.

In this study 90 heart sound segments are used corresponding to a total 1415 annotations examples of S1 and S2, within a time range between 1.2 and 14.7 seconds.

B. Performance Metric

For the evaluation of the performance of the algorithms, we have used the metric defined in the Pascal Challenge to calculate the error in S1 and S2 segmentation (3,4).

$$\Psi_k = \frac{\sum_{i=1}^{N_k/2}((\mid T_{S1,i} - \widehat{T}_{S1,i} \mid) + (\mid T_{S2,i} - \widehat{T}_{S2,i} \mid))}{N_k} \quad (3)$$

$$\Psi = \sum_{i=1}^{N_d} \Psi_k \quad (4)$$

Where $\Psi_K$is the average distance between the real observation of S1 and S2 and their automatic classification equivalents, for the k-th sound clip in the dataset. N_K is the number of S1 and S2 in the k-th sound clip. $\widehat{T}_{S1,i}$ and $\widehat{T}_{S2,i}$ indicate the real location of S1 and S2 in the i-th sound clip and $T_{S1,i}$ and $T_{S2,i}$ indicate the calculated location of S1 and S2 in their respective sound clip. N_D is the total of sound clips in the dataset. Ψ is the sum of all average distances between the calculated and the real location of S1 and S2.

The Pascal metric (PM) calculation does not take into account the number of false or true positives obtained. An algorithm with a higher rate of false positives could have a better rate in the Pascal metric than another algorithm with a lower rate of false positives. As an additional performance test we used the true positive (TP) and false positive (FP) rate to compensate these limitations.

Non-parametric paired sample test was used for the comparison of the Pascal metric (Kolmogorov-Smirnov, paired sample Wilcoxon signed ranks test), p-value<0.05 was considered significant.

C. Results

The entropy gradient performance is compared to the original envelogram signal obtained in the "Envelogram" stage of the diagram depicted in Figure 2. Both approaches use the same segmentation methodology explained in section 2.D). Table 1 presents these segmentation results, and we can observe that although the entropy gradient approach exhibits similar performance regarding the true positives detection, it reduces the number of false positives by almost half.

TABLE I. SEGMENTATION RESULTS USING ENTROPY GRADIENT AND SHANNON ENERGY ENVELOGRAM (TP – TRUE POSITIVES; FP – FALSE POSITIVES; PM – PASCAL METRIC)

$\Phi_{a,4}$	Annotated 1415			
	Detected	***TP***	***FP***	***PM***
$\nabla^2 Entropy$	1479	0.984	0.046	6
$Envelogram$	1550	0.987	0.096	9

As an additional experiment, we compared the classification performance when using different time windows for the embedding matrix construction. In Table 2 we can observe that the overall performance using *M = 25ms* is better than using a larger time window of *M = 250ms*. An 'ideal' time window should have the same dimension of the inherent attractor of the signal, and as we have explained before, this is an interesting future work topic to be explored, especially since the dataset involves a good variety of heart rates given the significant age differences of the auscultated children.

TABLE II. ENTROPY GRADIENT USING DIFFERENT TIME WINDOWS (TP – TRUE POSITIVES; FP – FALSE POSITIVES; PM – PASCAL METRIC)

SIZE	Annotated 1415			
	Detected	***TP***	***FP***	***PM***
$M = 25ms$	1479	0.984	0.046	6
$M = 250ms$	1844	0.977	0.340	20

IV. DISCUSSION

This paper introduces a new feature for heart sound segmentation that targets signals obtained in noisy environments. The results shows an increase in accuracy performance when compared to the traditional Shannon energy envelop features, maintaining the same levels of precision. This is especially motivating since all the experiments were made in real data collected in real noisy environments. We can conclude that the added adaptability of using embedding matrices, combined with the small morphological signal assumptions of a feature such as the entropy gradient are a promising avenue for future robust heart sound segmentations algorithms for real noisy environments.

REFERENCES

[1] D. Pereira, F. Hedayioglu, R. Correia, T. Silva, I. Dutra, F. Almeida, S. S. Mattos, M. Coimbra, "DigiScope – Unobtrusive Collection and Annotating of Auscultations in Real Hospital Environments", in IEEE EMBC, Boston, USA, Sep 2011.

[2] A.Guyton, J.E.Hall, eds., Textbook of Medical Physiology. Elsevier Saunders, 11th ed., Ed Hall, Jun 2006.

[3] H.Liang , S.Lukkarinen, I.Haritmo, "Heart Sound Segmentation Based on Heart Sound Envelogram", in Computers in Cardiology, vo.24, 1997.

[4] H. Liang and I. Hartimo, "A heart sound feature extraction algorithm based on wavelet decomposition and reconstruction," in Engineering in Medicine and Biology Society, 1998. Proceedings of the 20th Annual International Conference of the IEEE, vol. 3, pp. 1539 –1542 vol.3, oct-1 nov 1998.

[5] S. Omran and M. Tayel, "A heart sound segmentation and feature extraction algorithm using wavelets," in Control, Communications and Signal Processing, 2004. First International Symposium on, pp. 235 – 238, 2004.

[6] D.Kumar P.Carvalho, M.Antunes R.P.Paiva, J.Henriques, "Noise detection during heart sound recording using periodicity signatures", Physiological Measurement, vol 32 no 5, p. 599, 2011.

[7] A. Moukadem, A. Dieterlen, N. Hueber, C. Brandt, "A robust heart sounds segmentation module based on S-transform", Biomedical Signal Processing and Control, Volume 8, Issue 3, May 2013, Pages 273-281, ISSN 1746-8094,

[8] A. Yadollahi and Z. Moussavi, "A robust method for heart sounds localization using lung sounds entropy," Biomedical Engineering, IEEE Transactions on, vol. 53, pp. 497–502, March 2006.

[9] A.Castro, T.Vinhosa, S.Mattos, M.Coimbra "Heart Sound Segmentation of Pediatric Auscultation Using Wavelet Analysis", in IEEE EMBC, Osaka, JAP, Jul 2013.

[10] H. Kantz, T. Schreiber Nonlinear Time Series Analysis , 2th ed. vol .3, Ed. Cambridge University Press, Jan 2004,cp 3-9.

[11] P.Bentley, G.Nordehn, M.Coimbra, S.Mannor "The PASCAL Classifying Heart Sounds Challenge 2011 (CHSC2011) Results. " http://www.peterjbentley.com/heartchallenge/