35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Heart Sound Segmentation of Pediatric Auscultations Using Wavelet Analysis

Ana Castro[1], Tiago T. V. Vinhoza[1], Sandra S. Mattos[2] and Miguel T. Coimbra[1]

Abstract— **Auscultation is widely applied in clinical activity, nonetheless sound interpretation is dependent on clinician training and experience. Heart sound features such as spatial loudness, relative amplitude, murmurs, and localization of each component may be indicative of pathology. In this study we propose a segmentation algorithm to extract heart sound components (S1 and S2) based on it's time and frequency characteristics. This algorithm takes advantage of the knowledge of the heart cycle times (systolic and diastolic periods) and of the spectral characteristics of each component, through wavelet analysis. Data collected in a clinical environment, and annotated by a clinician was used to assess algorithm's performance. Heart sound components were correctly identified in 99.5% of the annotated events. S1 and S2 detection rates were 90.9% and 93.3% respectively. The median difference between annotated and detected events was of 33.9 ms.**

I. Introduction

Stethoscopes are part of the first line of screening and diagnosis of heart pathologies, and technological developments, such as the development of the digital stethoscopes, may improve this tool. The ability of those devices to generate phonocardiograms (PCG), record and transmit the heart sound, allows further processing and analysis which led to the renaissance of the stethoscope as a first diagnosis tool [1]. Signal processing techniques applied over the phonocardiogram (PCG) allow for a more in-depth perspective on the problem of heart disease detection, by segmentation of individual heart sound, and identification of possibly abnormal sounds such as murmurs [2], [3], [4].

The heart cycle is usually described as being composed by four main components: the first sound (S1), the systolic period, the second sound (S2) and the diastolic period (Figure 1). The correct identification of heart sounds may aid in the analysis of this signal in more detail, retrieving information from each heart sound component, and murmurs presence and localization. Because this information is relevant, the correct localization of the heart sound components is imperative. Although PCG segmentation and feature extraction may aid in pathology detection, the way clinicians conduct the diagnosis process includes not only information deriving from the auscultation, but also clinical background and patient characteristics.

[1]A. Castro, Tiago T. V. Vinhoza and M. T. Coimbra are with Instituto de Telecomunic cõesFaculdade de Ciências da Universidade do Porto, Departamento de Ciência de Computadores, Rua do Campo Alegre 1021/1055, 4169-007 Porto, Portugal ana.castro, tiago.vinhoza and mcoimbra@dcc.fc.up.pt

[2]Unidade de Cardiologia e Medicina Fetal, Real Hospital Português, Recife, Brasil ssmattos@cardiol.br

Segmentation techniques have been thoroughly explored, with the use of different approaches such as the average Shannon energy [5] and wavelet decomposition and reconstruction [6], [7]. Nonetheless for higher heart rates and noisy environments, differentiating between heart sound components is still a challenge [8].

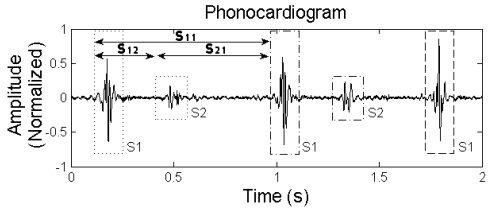


Fig. 1: Phonocardiogram (PCG) representation for one patient in the study, with correspondent heart sound components delimited: S1, S2, systolic (S_{12}), diastolic (S_{21}) and heart cycle (S_{11}) periods.

In this study we focus on the development of a segmentation algorithm, retrieving S1 and S2 localizations, for implementation in a real-time clinical decision-support tool. The proposed method provides a robust heart rate estimate prior segmentation based on singular value decomposition, and refines global heart sound localizations based on time and frequency characteristics of the heart sound through adequate wavelet sub-bands.

II. Motivation: The DigiScope Project

The goal of this work is to provide a robust segmentation algorithm for S1 and S2 identification, in realistic clinical environments. We aim to implement it in a PCG collector in clinical usage, within the DigiScope project [9], a project with the objective of providing a tool for PCG and clinical data collection, and signal processing combined with data mining techniques for clinical decision support.

The data used in this study was collected in the Real Hospital Português de Beneficência, Brazil, anonymized and shipped to Portugal with the approval of the RHP Ethics Commitee. The Ethics Committee of the University of Porto, Portugal approved this study.

Data collected with a Littmann® 3200 electronic stethoscope, and the DigiScope Collector presented in Figure 2. This was the data used in the construction of the Pascal Challenge data set later described [10].

III. Methods

The segmentation algorithm here proposed retrieves the information on heart sounds localization, and classification

Fig. 2: Prototype of the DigiScope Collector system, composed by a tablet and an electronic stethoscope Littmann® 3200.

into its sub-components, for data collected in the clinical environment.

A. Segmentation Algorithm

The proposed segmentation algorithm can be described in five sequential stages as presented in Figure 3: pre-processing of the heart sound; filtered envelope extraction; heart rate estimation; heart sound candidates identification and pre-classification; classification refinement. Each stage will be detailed below.

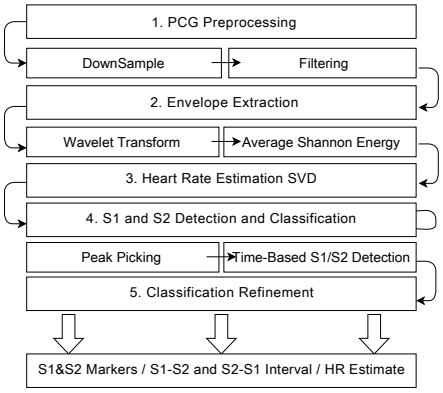


Fig. 3: Segmentation algorithm procedure in five sequential steps: pre-processing, envelope retrieval, heart rate estimation, heart sound classification, and classification refinement.

1) Data Pre-Processing: The pre-processing of the PCG includes the decimation of the original signal to 2kHz, normalization, and filtering with a band-pass, zero-phase, Butterworth filter order 6 (25-900 Hz)) to eliminate out of band noise.

2) Filtered Envelope: To obtain the signal envelope, with limited influence of impulse noise and highlighting S1 and S2, an order 6 Daubechies wavelet was used on the pre-processed PCG, to extract the approximation and detail coefficients at level 4 [8] given by $\phi_{a,4}(i)$ and $\phi_{d,4}(i)$ respectively. The approximation coefficients were normalized and then used to obtain an average Shannon Energy envelope, due to its non-linear relation between amplitude and energy [5], and given by (1).

$$E_S = -\frac{1}{N}\sum_{i=1}^{N} \overline{\phi}_{a,4}(i)^2(i) \cdot \log \overline{\phi}_{a,4}(i)^2(i) \qquad (1)$$

where $\overline{\phi}_{a,4}(i) = \phi_{a,4}(i)/\max_i |\phi_{a,4}(i)|$ are the normalized approximation coefficients.

The contour envelope of the former is calculated using only the interpolated energy peaks within a time range of 40 ms, as represented in Figure 4, providing a smooth maximum energy envelope for further processing. The proposed envelope gives a smooth version of the signal contour, allowing a more robust estimation of heart rate through singular value decomposition (SVD), as described in the next section.

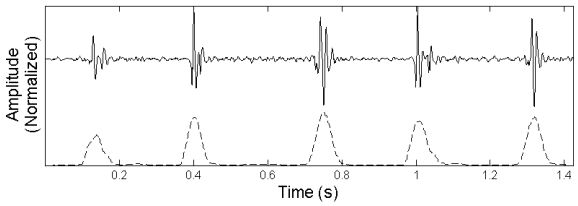


Fig. 4: Representation of the pre-processed phonocardiogram on top, and respective energy envelope below, for one of the sound segments of the study.

3) Heart Rate Estimation: Regarding the time span of PCG collection and acquisition conditions, one may assume that the heart rate is almost stationary in each segment. This may be assumed since usually during an auscultation the patient is at rest, the sound segments are relatively short, and fluctuations due to the respiratory cycle may be disregarded in terms of heart cycle duration estimation. To infer the expected localization of each heart cycle, we estimate the heart rate for each segment using the contour envelope described in the previous section, down-sampled by a factor of 10, to achieve a compromise between computation time of the SVD ratio and heart rate resolution.

In pediatric patients the heart rate range is larger than in adults, therefore to estimate it time windows varying from 400 to 1700 ms were used (35-150 beats/min). For each heart cycle length, a moving window over the PCG signal is used to construct a matrix $\mathbf{A}$ containing one segment in each row, and then proceed to singular value decomposition, extracting the two singular values to produce the ratio S_R, as given by (2) and (3) [8].

$$\mathbf{A} = \mathbf{USV}^T, \mathbf{S} = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \qquad (2)$$

$$S_R = \frac{\sigma_1}{\sigma_2} \qquad (3)$$

The heart rate is estimated based on the window size with maximum S_R.

4) S1/S2 Candidates: Using the approximation Shannon energy envelope, we employ a peak-picking algorithm to search for candidates of heart sounds, with some physiological inspired constraints: peaks would only be considered if they were at least with a spacing of 150 ms from their neighboring peaks to ensure that heart sound splits were not erroneously detected (lower energy peaks are discarded). We only select as candidates peaks with energy superior to at

least 30% the median of the total energy envelope peaks. This ensures that most of the spurious peaks are not detected, and that other artifacts would not interfere on peak detection.

Based on heart physiology a simple rule to classify the detected heart sounds was applied, considering that the systolic period (S_{12}) is longer than the diastolic period (S_{21}). For higher heart rates this rule may not stand since S_{12} and S_{21} tend to become closer leading to erroneous classifications (also a problem in the presence of noise). This is the motivation for the following section, leveraging the further knowledge on the spectral characteristics of the heart sound components and heart rate to re-classify each candidate.

5) Classification Refinement: We propose a refinement to our classifier based on the spectral characteristics of the heart sound, since the S2 component tends to exhibit a higher frequency content than the S1 component [11]. This characteristic coupled with the heart rate estimate is the basis for the refined classification. For each heart cycle, the algorithm reclassifies candidates based on signal's relative energy distribution through details at levels 3 ($\phi_{d,3}(i)$) and 4 ($\phi_{d,4}(i)$) [12]. For these sub-bands the S2 component exhibits higher energy when compared to S1, allowing their separation.

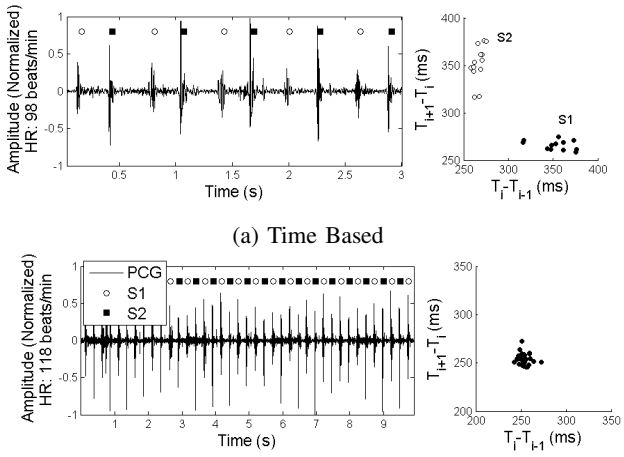
IV. Performance Assessment

To validate and assess the segmentation algorithm performance, we use annotated data from the Pascal challenge 'Classifying Heart Sounds Challenge' (Btraining_normal dataset) [10]. This database includes PCGs from pediatric patients, with corresponding annotations for the S1 and S2 components. Segmentation results were retrieved for each heart sound and compared to the annotations provided. Having data from children to test the algorithm performance poses an additional challenge due to the high heart rates. In this study 84 heart sound segments were used (639 S1 and 630 S2 annotations, 6 segments discarded), with time range between 1.2 and 14.7 s (416 s total).

A. PCG Segmentation Evaluation

Figure 5a presents the results for one of the sound segments studied, with correspondent estimated heart rate, and where it is evident the difference between S_{12} and S_{21}, leading to the generation of two clusters correspondent to S1 and S2 events. However, in Figure 5b we may observe a case where S_{12} and S_{21} are similar, and a simple time-based rule would not be enough to accurately classify events, and were a frequency inspired rule accurately detects the S1 and S2 components.

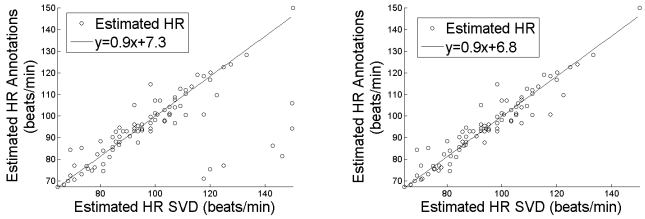
Heart rate estimation through (2) and (3) was crossed with the heart rate estimation using the annotated S1 and S2 (S_{12} and S_{21} intervals). Figure 6a shows the relation between the two, demonstrating an accurate estimation of the heart rate via SVD, although some outliers occur. These outliers occur for higher heart rates and similar envelopes, leading to an estimate that is a multiple of the real heart rate. This effect was corrected by checking the estimated heart cycles against



(a) Time Based

(b) Refined

Fig. 5: Experimental results of segmentation: on left representation of the phonocardiogram, with respective estimated heart rate (HR) and heart sounds localizations (S1 circle, S2 square); on right, representation of the time intervals between adjacent heart sound candidates retrieved clusters (T_i time of detected candidate i).



(a) Not Corrected (b) Corrected

Fig. 6: Representation of estimated heart rate (HR) through singular value decomposition (SVD) versus HR estimation based on S_{12} and S_{21} intervals (clinician annotations), and respective robust linear regression.

the number of detected peaks. If the number of candidates is similar to the estimated heart cycles than the heart rate would be corrected. Figure 6b shows the corrected SVD estimate demonstrating a strong correlation between the estimated and the annotated heart rate (ρ=0.85, P<0.001).

Table I shows the classification results considering each heart sound component, pre and post-refinement. The proposed contour envelope allowed the successful detection of 99.5% of the annotated heart sounds. We may observe that the algorithm performance was improved significantly after refinement based on spectral characteristics of each candidate.

TABLE I: Experimental results of the proposed segmentation methodology in the Pascal Challenge database [10].

		Time Based		Refined	
	# Annotated	**# Detected**	**Rate%**	**# Detected**	**Rate%**
S1	639	523	81.9	581	90.9
S2	630	517	82.1	588	93.3
Total	1269	1040	82.0	1169	92.1

We also used the metric defined in the Pascal Challenge to calculate the error in S1 and S2 segmentation,

$$\delta_k = \frac{\sum_{i=1}^{N_k/2} (|T_{S1,i} - \widehat{T}_{S1,i}|) + (|T_{S2,i} - \widehat{T}_{S2,i}|)}{N_k} \quad (4)$$

$$\delta = \sum_{i=1}^{N_D} \delta_k$$

where δ_k denotes the average distance of the k-th sound clip in the dataset. N_k is the total number of S1 and S2 in the k-th sound clip. $T_{S1,i}$ and $T_{S2,i}$ indicates the real localization of S1 and S2 of the i-th heatbeat and $\widehat{T}_{S1,i}$ and $\widehat{T}_{S2,i}$ indicates the calculated localization of S1 and S2 of the i-th heatbeat. N_D is the total of sound clips in the dataset. Finally, δ is the total error.

The proposed approach achieves an accumulated error δ of 15267.5 samples (sum of the average number of samples deviation (δ_k), at a sampling rate of 4000 Hz) [10]. Median distance between annotated and detected heart sounds was of 33.9 ms.

V. DISCUSSION

The first step in the detection of heart pathologies through auscultation is to distinguish the different sub-components of the PCG, and extract clinically relevant features. In this study we propose a new methodology for heart sound segmentation, that demonstrated good results in detecting the PCG sub-components, even in pediatric patients, without the need of an external biomedical signal (electrocardiogram e.g.).

Our method showed 99.5% agreement when compared to clinically annotated heart sounds, with 90.9% and 93.3% detection rates for S1 and S2 components respectively. A robust estimation of the heart rate through SVD is proposed, demonstrating high correlation between the SVD estimation and the annotated S_{12} and S_{21} intervals. A downsampled contour envelope was used to achieve a compromise between computation time and heart rate estimation resolution with good results (ρ=0.85). This information combined with the knowledge of heart sound physiology allowed an accurate detection of each heart sound component. Using solely a time-based rule is not sufficient in the presence of higher heart rates as presented in previous studies [5], [6], and a rule based on signal's relative energy distribution over the wavelet transform sub-bands of interest demonstrated to be a good estimate of the S2 higher frequency content when compared to S1 [11].

The algorithm performance was also evaluated in terms of the accumulated average distance between clinical annotation and detected events in each sound segment, with a median distance of 33.9 ms between annotated events and the markers retrieved [10].

VI. CONCLUSIONS

A new methodology to segment the heart sound is proposed, with 92.1% accurate detection. Some problems were identified and need to be addressed in the future, such as the existence of low energy events that were not detected, and also some spurious events detected as heart sounds and misclassified in the presence of ambient noise. Performance of the method in pathological conditions also needs to be evaluated in future studies.

ACKNOWLEDGMENT

This work was partially funded by the Fundação para a Ciência e Tecnologia (FCT, Portuguese Foundation for Science and Technology) under grant PEst-OE/EEI/LA0008/2011, PTDC/EIA-CCO/100844/2008 (DigiScope), Heart Safe PTDC/EEI-PRO/2857/2012, DECA-BIO IT/LA/01075/2011.

REFERENCES

[1] S. Messer, J. Agzarian, and D. Abbott, "Optimal wavelet denoising for phonocardiograms," *Microelectronics Journal*, vol. 32, no. 12, pp. 931–941, 2001.

[2] Z. Sharif, M. Zainal, A. Sha'ameri, and S. Salleh, "Analysis and classification of heart sounds and murmurs based on the instantaneous energy and frequency estimations," in *TENCON 2000. Proceedings*, vol. 2, pp. 130 –134 vol.2, 2000.

[3] J. Ning, N. Atanasov, and T. Ning, "Quantitative analysis of heart sounds and systolic heart murmurs using wavelet transform and ar modeling," in *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, pp. 958 –961, sept. 2009.

[4] S. Schmidt, E. Toft, C. Holst-Hansen, C. Graff, and J. Struijk, "Segmentation of heart sound recordings from an electronic stethoscope by a duration dependent hidden-markov model," in *Computers in Cardiology, 2008*, pp. 345 –348, sept. 2008.

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

[6] 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.

[7] 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.

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

[9] "The digiscope project - digitally enhanced stethoscope for clinical usage." http://digiscope.up.pt/.

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

[11] P. Arnott, G. Pfeiffer, and M. Tavel, "Spectral analysis of heart sounds: Relationships between some physical characteristics and frequency spectra of first and second heart sounds in normals and hypertensives," *Journal of Biomedical Engineering*, vol. 6, no. 2, pp. 121 – 128, 1984.

[12] M. Ursino, E. Magosso, E. Gardella, G. Rubboli, and C. Tassinari, "A wavelet based analysis of energy redistribution in scalp eeg during epileptic seizures," in *Engineering in Medicine and Biology Society, 2004. IEMBS '04. 26th Annual International Conference of the IEEE*, vol. 1, pp. 255 –258, sept. 2004.