Effective Phonocardiogram Segmentation Using Nonlinear Dynamic Analysis and High-Frequency Decomposition

AF Quiceno[1], E Delgado[2], M Vallverd[3], AM Matijasevic[4], G Castellanos-Domnguez[1]

[1]GC&PDS, Universidad Nacional de Colombia Sede Manizales, Colombia
[2]GINVESTAP, Research Center, Instituto Tecnolgico Metropolitano, Colombia
[3]ESAII, CREB, CIBER-BBN, Universitat Politcnica de Catalunya, Spain
[4]Telesalud, Faculty of Health Sciences, Universidad de Caldas, Colombia

Abstract

In this study, an effective methodology for segmenting the temporal trace of phonocardiographic signals (PCG) is presented. Initially, inter-beat segmentation is carried out using the DII lead of the ECG recording for locating the occurrence of the first heart sound (S1). Next, the intra-beat segmentation is achieved by using recurrence time statistics (RTS), which is sensitive to changes of the reconstructed attractor in a state space derived from nonlinear dynamic analysis. If the segmentation using RTS fails, an alternative segmentation is proposed using thresholding over the Shannon envelogram extracted from the high-frequency decomposition. The database of PCG records which was used belongs to the National University of Colombia. Inter-beat segmentation accuracy was 100% over all PCG recordings. Taking into account 360 PCG beats, where a set of 180 beats were strongly disturbed by different types of cardiac murmurs, intra-beat segmentation yielded an accuracy result of 97.7%.

1. Introduction

The segmentation process of PCG signals is a very important task in order to perform murmur detection, and diagnosis of other cardiac pathologies using computer analysis. Thus, it is essential that different components of heart cycle can be timed and separated [1]. Segmentation is performed in two main steps: inter-beat segmentation and intra-beat segmentation. The first one refers to the identification and separation of the beats with the aim of performing an individual diagnostic in each one of them. The second step includes the separation of the beat into its four principal components: first cardiac sound (S1), systole, second cardiac sound (S2) and diastole [2].

Cardiac murmurs are present in systole, diastole or both. In this way, it is crucial to precisely detect the boundaries in order to correctly locate the components of heart sounds (HS), and then, to perform an accurate diagnosis and classification of the murmur if it is present. The main problem in HS segmentation occurs when a murmur is present and produces high distortion in the temporal trace.

A large variety of algorithms to perform PCG segmentation have been proposed in the literature. In [1] the PCG signal segmentation is based on synchronized electrocardiographic (ECG) signal acquired with PCG. On the other hand, the envelope of PCG can be used to perform the segmentation, in [3] the envelope is estimated using homomorphic filtering, while in [4] it is estimated by means of Shannon normalized average energy. Moreover, Shannon energy can be used with more complex methodologies: wavelet transform in order to enhance the spectral components of S1 and S2; or with Mel-scaled filter-banks and Hidden Markov Models. Another approach to segment PCG is based on time-frequency analysis, implemented with wavelet transform or distributions belonging to the Cohen's quadratic class. That PCG signal exhibits nonlinear dynamics, this fact motivates the use of complexity measures [5], and RTS [6]. In our work, we combine the best of these methodologies with the aim of developing a reliable algorithm capable of segmenting the PCG signal with high accuracy.

In this paper, we propose a method for the boundary identification of S1 and S2, using the ECG signal as reference in order to separate each beat according to the methodology proposed in [7]; later, the end of S1, beginning of S2 and end of S2 are detected using RTS and threshold rules. The results obtained in the previous stage are validated using biomedical features; if the fiducial points are out of certain boundaries given by physiological considerations, an alternative segmentation method is used. It is based on wavelet decomposition, Shannon energy and threshold rules. The algorithms are tested in a PCG database labeled by expert cardiologists in the beginning and end of S1 and S2.

2. Methods

2.1. *R*-peak detection in ECG

The database used in this work, contains an ECG signal synchronized with PCG, thus detecting correctly the *R*-peaks in ECG, makes it possible to locate the beginning of S1.

The *R*-peaks in ECG can be detected employing algorithms for *QRS* complex detection, which have been widely reported in the literature. In this case, it is used a method proposed in [7] which consists basically in three steps: *Linear filtering*, *nonlinear transformations* and *decision rules*.

2.2. Recurrence time statistics (RTS)

RTS are used in order to detect abrupt changes in the signal dynamics, corresponding to S1 and S2. An arbitrary state a_{ref} is chosen on the trajectory whereupon all recurrences within a hypersphere of radious r are selected, i.e., $B_r(a_{ref}) = a : \|a - a_{ref}\| \leq r$. Ψ is then defined as the total amount of states in the set B_r, and related to the information dimension via a power law, which motivates its ability to detect weak signal transitions based on the amplitude, period, dimension and complexity of the signal.

A sliding window is used to partition the recorded PCG signal into overlapping segments, and Ψ is calculated for each segment. The r value is a very important parameter in the detection algorithm. If it is chosen too low, the hypersphere would be low on data, and if r is chosen too high, the hypersphere will contain misleading information from erroneous parts of the reconstructed state space. In this work, r is adaptive, and it becomes lower if there are not lobe detection corresponding to S1 and S2 in $\Psi(r)$ [6].

2.3. Wavelet decomposition and Shannon energy

These methods are used in the detection of S2 when it is considered that RTS did not give a good estimation of the boundaries. Basically, three steps are implemented: high-pass filtering, wavelet decomposition and Shannon energy operator applied to the detail coefficients.

From the knowledge of cardiac functionality and genesis of S1 and S2 sounds, it is known that aortic valves close with relatively large pressure difference across the valve, which originates the high frequency content in S2 sound [2]. This is the motivation to use the approximation coefficients to perform the detection of S2. To extract the signal envelope from the detail coefficients, the Shannon energy operator is applied (1).

$$E_d(d_j[n]) = -\frac{1}{N}\sum_{n=1}^{N}(d_j[n])^2 \log(d_j[n])^2 \qquad (1)$$

where d_j are the jth level detail coefficients of the wavelet transformed heart sound signal, and N is the number of samples in the selected window. This technique emphasizes the medium intensity signal components and attenuates the effect of low intensity components [2]. Using the signal envelopes provided by the Shannon energy, sound lobe boundaries are identified applying decision rules based on thresholds.

2.4. Proposed algorithm to perform PCG segmentation

According to the methods mentioned in sections 2.1 to 2.3, a segmentation algorithm was designed in order to find the points corresponding to the boundaries of S1 and S2 in PCG. This algorithm is described below:

Step 1. **Acquisition and preprocessing**:

The sounds were recorded with 16-bit accuracy using $44100Hz$ sampling frequency, in two channels: DII derivation of ECG and PCG recording. The database used in this work is made up of 22 adult subjects, who gave their informed consent, and underwent a medical examination. That, 8 recordings, of 10 seconds each one, corresponding to the four traditional focuses of auscultation were taken for each patient in the phase of post-expiratory and post-inspiratory apnea. Finally, the most representative beats (180 normal and 180 pathologic) were selected and labeled according to the criterion of cardiologists.

In this stage, PCG signals are normalized to be in the interval $[-1, 1]$, and filtered with wavelet denoising up to 8th level of decomposition using the wavelet *Coiflet 4*. Later, the PCG recording is high-pass filtered in order to enhance the frequency components of HS.

Step 2. *QRS* **and *R*-peak detection**:

The beats are segmented using the method explained in section 2.1. Henceforth, the processing is applied on the individual beats.

Step 3. **Location of the end of S1**:

RTS are obtained using a sliding window of 2000 points over the sampled signal with 90% overlap for taking well-resolution in time-domain. The dynamic attractor is reconstructed by using time-delay embedding, where the embedding parameters (i.e., embedding dimension and time lag) have to be estimated. The embedding dimension $d = 5$, is obtained using the false neighbor method and the time lag $\tau = 100$, is calculated minimizing the mutual information between each state variable and maximizing the distance from each point to the main diagonal of the state space.

The end of S1 is located in a window between 70 and $160ms$ after R-peak, hence the search is limited to this time interval. Furthermore, $\Psi(r)$ has local minima in S1 and S2, in this way $\hat{\Psi}(r) = 1 - \Psi$ is computed in order to detect positive lobes corresponding to HS. $\hat{\Psi}(r)$ is computed with an initial radius of $r_0 = 0.35$, but sometimes S1 has low intensity and there is no lobe detection using this value. Therefrom r is decreased until $\hat{\Psi}(r)$ exceeds a predefined threshold of 0.7.

Some systolic murmurs give rise to extra lobes at the end of the analysis window, to solve this problem, if a lobe is detected in this zone, and if it exceeds certain threshold (80 % of the maximum value), it is removed. Finally, a forward search is performed from the maximum of the lobe corresponding to S1, until the signal has an amplitude or slope below established thresholds.

Step 4. **Location of beginning and end of S2**:

The estimation of $\hat{\Psi}(r)$ is computed in the same way aforementioned for S1, but in a different analysis window $(250 - 610ms)$. Moreover, systolic murmurs can produce lobes in the first part of the window, so if they are present, they are taken away.

After this, it is performed a lobe search based on an ascending amplitude threshold (50 % of the maximum value), and a descending threshold (10% of the maximum). Once the lobes are detected, they are validated, removing the ones that start beyond $430ms$, which is physiologically impossible. Finally, S2 is assumed to be in the remaining lobe with major amplitude and its boundaries are located computing backward and forward search from the maximum of the lobe.

Step 5. **Alternative segmentation of S2**:

The boundaries estimated in *Step 4* are compared with some criteria in order to decide whether they are correct or wrong:

- S2 length must be more than $35ms$.
- S2 length must be less than $240ms$.
- S2 beginning must be at least $265ms$ after R-peak.
- S2 beginning cannot be more than $400ms$ after R-peak.
- There must be at least one lobe detection in *Step 4*.

If one of the aforesaid criteria is not meet, an alternative segmentation method is used, based on wavelet decomposition and Shannon energy. First, the signal is downsampled to $4000Hz$, later, the wavelet decomposition is calculated up to 6th level using the wavelet *Daubechies 6*. Then, the energy is computed in the detail coefficients, using the expression (1), with $N = 100$ and the envelope E_d of the PCG signal is obtained.

E_d is analyzed within the same window used in *Step 4*, and the lobes corresponding to possible murmurs are removed if they are present. Afterwards, a lobe search is performed in order to find the lobes that exceed certain amplitude threshold (30% of the maximum of E_d). Then,

the lobes beginning before than $265ms$ after R-peak, or beyond $400ms$ before R-peak are deleted. Finally, if there is no lobe that meets the conditions, it is assigned a lobe with physiological mean values: beginning at $265ms$, end at $610ms$ and maximum point located $440ms$ after R-peak. The last step is to detect the boundaries of the major lobe as explained in *Step 4*.

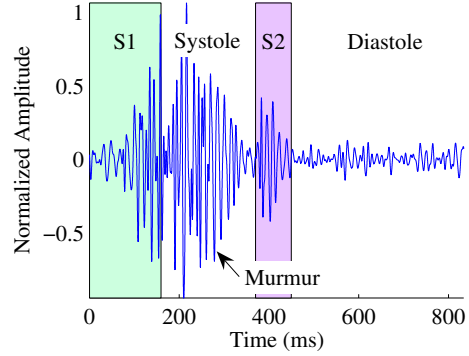Figure 1 shows a PCG signal with severe systolic murmur effectively segmented by the system.

Figure 1. Correctly segmented PCG signal with systolic murmur.

3. Results

The algorithms were tested in a PCG database labeled by expert cardiologists. The beat separation using the ECG channel (inter-beat segmentation) gave a performance of 100% in the whole database. In order to evaluate the performance of intra-beat segmentation, there are three fiducial points labeled by cardiologists: end of S1, beginning of S2 and end of S2.

Figure 2a shows the location of the three intra-beat segmentation events labeled by cardiologists. The end of S1 is located around $120ms$ after R-peak with low dispersion; whereas the average location of the beginning of S2 is $350ms$ after R-peak and the end of S2 at $430ms$. Figure 2b shows the absolute difference in time (ms) of the points calculated by the algorithm with respect to the labels; differences greater than $30ms$ are considered as wrong detections.

The performance results for each segmentation event are shown in Table 1: Mean Squared Error (MSE) with respect to the points marked by specialists, and the number of points whose difference with the label is more than $30ms$. The total MSE in all intra-beat segmentation points (360 for the end of S1 (S1-e), 360 for the beginning of S2 (S2-b), and 360 for the end of S2 (S2-e)) is $18.25ms$, and the total accuracy of the algorithm is 97.7%.
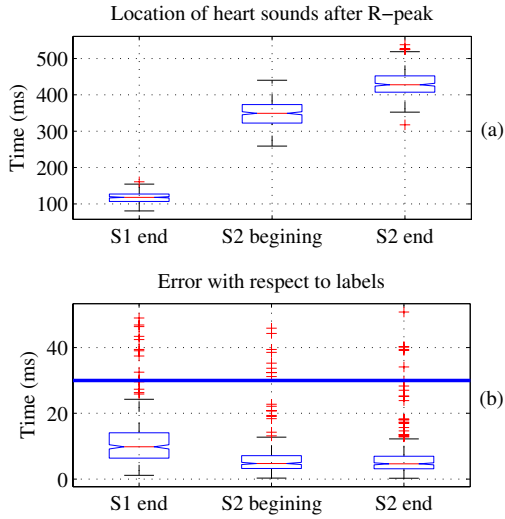
Figure 2. Boxplots corresponding to: (a) Location of S1 and S2 boundaries marked by cardiologists; (b) Error with respect to medical annotations.

Table 1. Performance results of PCG segmentation

	R-peak	S1-e	S2-b	S2-e
MSE (ms)	-	13.4	8.7	8.8
Wrong-detections	0/360	10/360	9/360	6/360

4.   Conclusions

Based on the obtained results, the proposed methodology to perform PCG segmentation works effectively in the separation of beats (inter-beat segmentation), and then, in the identification of the main intervals that compose each beat: S1, systole, S2 and diastole (intra-beat segmentation). This method has a good reliability, since results have shown high precision in event localization, including PCG signals with severe cardiac murmurs that could block out S1 and S2 sounds.

The results showed that the estimation of S2 boundaries is better than the estimation of the end of S1. This happens because S1 has low intensity in some PCG signals, and some cardiac murmurs mask S1, making the boundary localization a difficult task. In addition, as showed in Figure 2a, the location of HS is inside the physiologic standards: S1 duration between 100 and $160ms$, S2 beginning between 280 and $360ms$ and S2 duration between 80 and $140ms$.

Future work includes the application of this segmentation techniques together with feature estimation methods in order to identify cardiac pathologies, specifically, the diseases related with valvular problems, which produces cardiac murmurs whose detection and classification can be aided by the methodology proposed in this work.

Acknowledgements

References

[1] El-Segaier M, Lilja O, Lukkarinen S, Srnmo L, Sepponen R, Pesonen E. Computer-based detection and analysis of heart sound and murmur. Ann Biomed Eng 2005;33(7):937–942.

[2] Liang H, Lukkarinen S, Hartimo I. A heart sound segmentation algorithm using wavelet decomposition and reconstruction. In Proc. 19th IEEE-EMBS. 1997; 1630–1633.

[3] Gupta CN, Palaniappan R, Swaminathan S, Krishnan SM. Neural network classification of homomorphic segmented heart sounds. Applied Soft Computing 2007;(1):286–297.

[4] Liang H, Lukkarinen S, Hartimo I. Heart sound segmentation algorithm based on heart sound envelogram. In Proc. Computers in Cardiology 1997; 105–108.

[5] Nigam V, Priemer R. Accessing heart dynamics to estimate durations of heart sounds. Physiol Meas 2005;26(6):1005–1018.

[6] Ahlstrom C, Liljefeldt O, Hult P, Ask P. Heart sound cancellation from lung sound recordings using recurrence time statistics and nonlinear prediction. IEEEJSPL 2005; 12(12):812–815.

[7] Laguna P, Srnmo L. Bioelectrical signal processing in cardiac and neurological applications. Elsevier, 2005.

Address for correspondence:

Name: Andrs Felipe Quiceno-Manrique
Full postal address: Universidad Nacional de Colombia. Km. 7, Campus La Nubia. W-Mezanine 2. Manizales - Caldas. Colombia. Tel: +57 3113897397.
E-mail address: afquicenom@unal.edu.co