Automatic Heart Sound Segmentation and Murmur Detection in Pediatric Phonocardiograms

João Pedrosa[1], Ana Castro[1], and Tiago T. V. Vinhoza[2].

Abstract— The digital analysis of heart sounds has revealed itself as an evolving field of study. In recent years, numerous approaches to create decision support systems were attempted. This paper proposes two novel algorithms: one for the segmentation of heart sounds into heart cycles and another for detecting heart murmurs. The segmentation algorithm, based on the autocorrelation function to find the periodic components of the PCG signal had a sensitivity and positive predictive value of 89.2% and 98.6%, respectively. The murmur detection algorithm is based on features collected from different domains and was evaluated in two ways: a random division between train and test set and a division according to patients. The first returned sensitivity and specificity of 98.42% and 97.21% respectively for a minimum error of 2.19%. The second division had a far worse performance with a minimum error of 33.65%. The operating point was chosen at sensitivity 69.67% and a specificity 46.91% for a total error of 38.90% by varying the percentage of segments classified as murmurs needed for a positive murmur classification.

I. Introduction

The heart, center of the cardiovascular system, is responsible for pumping blood to every vital organ of the human body and its correct functioning is therefore of paramount importance. To fulfill its function, the heart follows a series of coordinated movements known as the heart cycle, which can be separated as the systole (muscle contraction) and the diastole (muscle relaxation). A distinctive sound, lub-dub, is produced by these movements, and its graphical representation is known as the phonocardiogram (PCG). This sound can however be modified by a number of cardiopathies, producing heart murmurs. These murmurs can be either pathological or innocent and arise differently depending on their origin. 50-70% of children have a murmur (most of these are innocent) and 2-5% of adults have or will eventually have a pathological valvular murmur [1]–[3]. Several tools exist in the clinical environment to confront the problem of murmur detection. Cardiac auscultation through the use of the stethoscope is the most common technique however it is highly subjective and is dependent on a lengthy and continuous training. Echocardiography, despite its higher precision, is not available everywhere and is associated with higher costs and patient stress. Computer-aided auscultation would then allow faster and cheaper decisions by using a tool widely known by both physicians and patients: the stethoscope. The main goal of this work is to design a robust algorithm capable of segmenting a PCG signal into heart cycles and also able to detect murmurs within those cycles, independent of other sources of information.

II. Related Work

Computer-aided auscultation has been a subject of research for some time and many different methods have been applied to solve this problematic. PCG signal processing can be crudely divided into two main research areas. One is focused in the detection of events such as S1 and S2 to perform the segmentation of the PCG. The other is the detection of murmurs and, consequently, of cardiac pathologies. However, because both objectives are interconnected and are accomplished from the same base signal, they share the same basic signal processing tools. These tools range, nevertheless, a large amount of techniques from the most simple to some utterly complex [4]. On the subject of heart sound segmentation, El-Segaier developed a method based on ECG gating. This method uses the simultaneous acquisition of PCG and ECG signals. Using the temporal relations between the PCG and ECG, intervals of search for S1 and S2 were defined and the maximum in the spectrum in each of those intervals was defined as S1 or S2 [5]. The work from Oskiper and Watrous [6] also uses the ECG as auxiliary signal to train two time-delay neural networks to accomplish the segmentation task. These neural networks may then be applied without the aid of the ECG signal. The work from Liang et al. designed a segmentation algorithm dependent on the Shannon energy envelope[1]. A threshold is set to select the peaks from such envelope [7]. The works in [8] and [9] developed a methods based both on time domain and time-frequency analysis. A specific function was designed to be sensitive to high amplitudes and the specific frequencies of the main heart sounds [8] and an algorithm using self-organizing probabilistic maps was developed in [9]. Table I summarizes the results from the aforementioned works.

Regarding murmur detection, the common practice found in the literature is the following: a number of features are extracted from the interval of interest (systole or diastole) and they are introduced into a classifier for training and

---

[1]For a given signal $x(t)$, its Shannon energy envelope E is defined as $E = -x^2(t) \log x^2(t)$.

TABLE I
SEGMENTATION ALGORITHMS

Ref #	Analysis Used	Auxiliary Signals	Sensitivity	PPV
[5]	T; TF	ECG	100%(S1), 97%(S2)	-
[6]	T;TF	ECG (training)	98.4%	97.8%
[7]	T	-	94.11%	98.76%
[8]	T; TF	-	99%	98.6%
[9]	T;TF	-	98.4%	96.7%

T- time domain; TF- time- frequency domain analysis

testing to find a performance value. The type of features extracted from the signals range from time domain and time-frequency domain such as in [10] to complexity ones from the works develop by Kumar et al. in [11], [12]. A wider feature analysis was conducted by Delgado-Trejos using time domain, time-frequency domain, perceptual and nonlinear features obtaining an overall accuracy of 96.11% [13]. Table II summarizes the results for these references.

TABLE II
MURMUR DETECTION ALGORITHMS

Ref #	Analysis Used	Results
[10]	T; TF	86% accuracy
[11]	NLC	91.09% sensitivity, 95.25% specificity
[12]	TF;NLC	89.1% sensitivity, 95.5% specificity
[13]	TF;P;NLC	96.11% accuracy

T- time domain; TF- time- frequency domain; P- Perceptual analysis;
NLC- nonlinear and chaos based analysis

III. METHODOLOGY

Our approach is composed by two phases: 1) the main heart sounds are identified so that the signal is separated into systolic and diastolic segments; 2) each segment is subjected to a classifier to determine the presence of a murmur.

A. Heart Sound Segmentation

As the rest of the detection algorithm is dependent of a precise segmentation, this stage needs to be as robust as possible and not only identify the correct sounds but also correctly classify them as S1 and S2. The segmentation algorithm designed is mainly based in the autocorrelation function to find the periodic components of the PCG signal.

First of all, a pre-processing routine is applied composed of continuous wavelet denoising using the Morlet mother wavelet with scales corresponding to frequencies of 15-150Hz. This mother wavelet was chosen as it has been proven to be the most effective for continuous wavelet transform for PCG analysis [14]. The remaining signal is converted to its energy envelope, normalized and decimated to obtain a smooth envelope. Secondly, the PCG is divided into 1.5s segments and a mean amplitude criterion is used to discard segments with high amplitude noise. This segment length was chosen to allow for a full heart cycle to be present at every segment. The periodic elements of the remaining segments are then detected using the autocorrelation function and the systole length is estimated according to the ACF

peaks as shown in Figure 1. Finally, the heart sounds detected using the estimated systole length are classified into S1 or S2. While this is fairly easy in most cases as the systole is normally shorter than the diastole, in children and adults with higher heart rates systoles and diastoles are approximately of the same length. Non-duration based criteria such as the differences between the frequency spectra shape of the main heart sounds [15] must then be used.
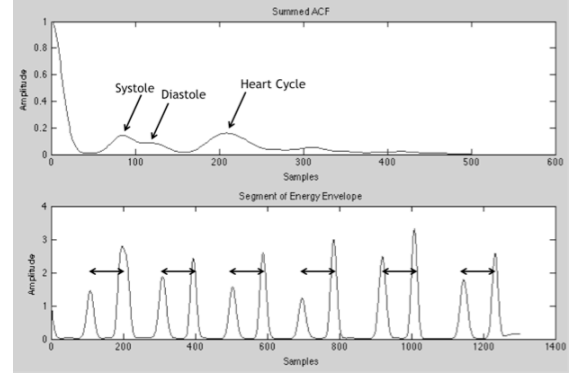

Fig. 1. Example of a summed ACF and segment with identified intervals for an estimation equal to the systole.

B. Feature Extraction for Murmur Detection

A total of 250 features were extracted from each segment to classify as presenting murmur or not. These features were extracted from different analysis domains to ensure the segments were described as thoroughly as possible. Time domain analysis was chosen as it allows the study of the amplitude in the systole, which is altered in the presence of a murmur. The time domain features extracted were obtained with the Shannon energy values for seven different locations in a PCG segment. A time-frequency analysis was also used to identify the correct frequency spectrum of the murmurs and distinguish them from noisy events. Fifty-five features were extracted from each segment using the DWT. The first five level detail coefficients were obtained with the Daubechies 1 mother wavelet. The mean and Shannon energy of each segment was obtained. Finally, the ratio between the minimum Shannon energy and the maximum between levels returned the last feature. This was done for five different sections of the segment: the whole segment, the whole segment except for the S1 and S2 and each of the thirds of the segment after removing the S1 and S2. Thirteen features were generated using the CWT. The Morlet mother wavelet was used to obtain the CWT of the PCG with scales encompassing the frequencies from 200Hz to 700Hz. Figure 2 shows how some time domain and time-frequency domain features are extracted from the PCG. This frequency range was chosen to remove any low or high frequency components and maintain the murmur bandwidth. Singular value decomposition (SVD) was then applied to each PCG segment. The eight largest eigenvalues of the TFR matrix were considered as features. The eigentimes and eigenfrequencies corresponding to the two largest eigenvalues were
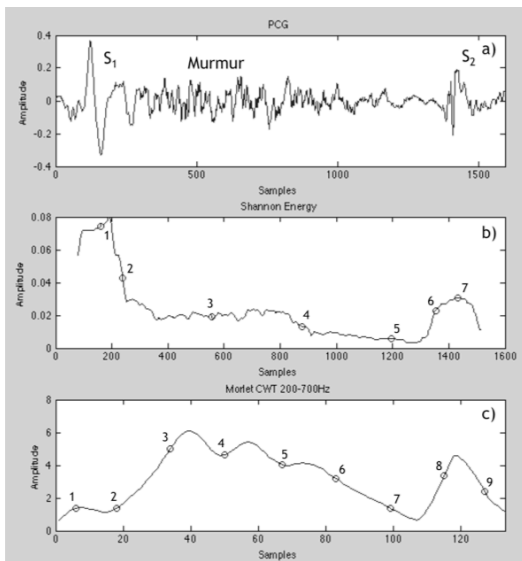
Fig. 2. a) PCG segment with heart sound components S1, S2 and systolic murmur marked; b) corresponding Shannon energy and the seven points considered as features; c) corresponding CWT obtained with scales encompassing 200Hz to 700Hz with points used for features extraction marked. Details on the feature extraction can be found in [16]

also used as features by applying a histogram (10 bins) to each of the eigenvectors probability density function. This returned ten features for each eigenvector for a total of 48 features generated by SVD from each segment [10].

Perceptual analysis using mel-frequency cepstral coefficients (MFCC) is a tool that perceives frequency in a logarithmic fashion emulating the behavior of the human ear. The PCG sections used for MFCC extraction were the same ones used to extract the DWT features. A total of 20 MFCC coefficients per section were extracted resulting in the total of 100 features from the perceptual analysis domain.

Murmurs are the highest complexity events in a PCG. This makes a nonlinear study of the PCG a promising tool for murmur detection, as the presence of a murmur will raise the overall complexity of a PCG [10], [13]. The bispectrum, a common higher order statistics, was used to extract features from the PCG regarding its nonlinear interactions. Other analysis such as Variance Fractal Dimension and Lyapunov Exponents were performed to study the complexity and chaoticity of the PCG signal.

Table III shows the features extracted as well as its amount. A k-means classifier was trained using these 250 features to classify the segments. A sequential forward feature selection (SFFS) algorithm was also used to determine the optimal feature set consisting of 167 features. As the classifier evaluates each segment, a threshold was applied to the fraction of segments classified as presenting murmur to define whether or not a patient had a murmur.

IV. RESULTS

A. Databases

To evaluate the performance of the designed algorithms two different databases were used. The DigiScope database

TABLE III

FEATURES EXTRACTED FOR MURMUR DETECTION

Feature Name	Amount
Shannon Energy	7
Continuous Wavelet Transform	13
Discrete Wavelet Transform	55
Singular Value Decomposition	48
Mel-Frequency Cepstrum Coefficients	100
Bispectrum	16
Variance Fractal Dimension	9
Lyapunov Exponents	2
Total	250

was collected in the Real Hospital Português in Recife, Brasil using a Littmann 3200 stethoscope and consists on a total of 72 signals. This stethoscope was used with the DigiScope prototype developed within the homonymous project to collect, transmit and record heart sounds without interfering with clinical routine. All the sound samples were collected in the clinical environment with lengths under one minute. The selected procedure was to sequentially auscultate all four main auscultation spots (aortic, pulmonary, mitral and tricuspid). Each physician was given the freedom to decide how much time to spend on each spot depending on whether there was something particular, as they would do if examining the patient in a normal situation. Physicians were also instructed to make no additional effort to find a quiet environment for signal acquisition. The patients auscultated were of ages comprised from six months to 17 years old. The presence of murmurs and its temporal localization (systole/diastole) is annotated but the main heart sounds are not. To test and validate the heart cycle segmentation the PASCAL CHSC 2011 database [17] was used. Only a section of the database comprised of 111 signals of varying lengths between 1 and 30 seconds was used. This section is uniquely composed of signals without murmur and the signals have very little noise. The ages of the patients auscultated are not specified. Both databases are summarized in Table IV.

TABLE IV

DATABASES

Database	Number of Signals	Duration of Signals	Labelled S_1/S_2	Labelled Murmurs
Digiscope	72	≈ 1 min	No	Yes
Pascal [17]	111	1-30 s	Yes	No

To assess the performance of the heart sound segmentation algorithm three different measurements were used. The first two, the sensitivity and the positive predictive value (PPV) are widely used as was reported in Section II. The sensitivity evaluates the proportion of heart sounds that are found whereas the PPV will evaluate the proportion of the events identified that are actually heart sounds. An additional measurement δ was also used to determine the temporal precision of such detections and is described in [17].

B. Heart Sound Segmentation Results

A sensitivity of 89.2% and a positive predictive value of 98.6% were obtained and $\delta = 9.8$ ms. When comparing the

proposed segmentation algorithm performance with the ones listed in Table I it has a lower sensitivity than any other method. However, a few things must be taken into consideration. First of all, different databases were used in each study and consequently direct comparisons are always subjective. Secondly, the segmentation algorithm was designed for a different database than the one used for testing. This would not be a big problem except for the size of the PCG signals that, as mentioned earlier, hinder the heart rate estimation process. Furthermore, the goal of the designed algorithm must be taken into account. Unlike the algorithms in Table I, the ultimate goal of our segmentation algorithm is not only to identify the heart sounds but also to discard noisy regions even if the algorithm is able to pinpoint the exact locations of the heart sounds in these regions. This design was implemented because the priority was to infer in the presence of murmur and not the perfect segmentation of the signal. The average deviation of 9.8 ms is a small deviation given that the average duration of the main heart sounds is of approximately 100 ms.

C. Murmur Detection Results

The performance of the trained classifier was evaluated using two different approaches. The first approach was conducted by performing a random train-test division disregarding the patient to whom each segment originated. A minimum error of 4.65% was obtained using the whole feature set and a minimum error of 2.19% was obtained for the optimal subset of 167 features. These values are well within the range of those found in literature where the same train-test division was used. This approach does not, however, consider the interpatient variability as segments from the same patient are present in the train and test set. The second approach divides train-test sets at a patient-level. In both cases the dataset was randomly divided into train and test sets in a 60-40 ratio. As expected, the error is much larger than in the previous division: 35.5% for the whole set of features and 33.65% for the optimal subset. This is the error that should be expected when classifying a new patient due to the patient variability that exists even in patients with the same pathology. This effect is augmented even further due to the fact that the database used is relatively small and due to the noise present in the signals. The sensitivity and specificity of the trained classifier returned values at this maximum accuracy operating point of 52.38% and 79.40% respectively. By changing the patient threshold, the operating point was changed to a sensitivity of 69.67% and specificity of 46.91% to ensure a greater amount of murmur cases detected at a cost of increasing the total error to 38.9%.

V. Conclusions and Future Work

This paper presented a novel algorithm for the digital analysis of heart sounds and detection of murmurs, with promising results. Given the motivation of this paper it is important to compare the obtained values to the accuracy levels of general practitioners. Lam et al. conducted a study that, among other things, evaluated the accuracy of physician trainees in the detection of murmurs obtaining accuracy values of 79.2% and 67% for pan-systolic and ejection systolic murmurs respectively [18]. By comparison to the obtained accuracy and sensitivity values of 61.1% and 69.67% respectively, one sees that it is within the range of the values obtained in Lam et al. for the physicians normal performance. To make a difference in the clinical environment, a murmur detection algorithm should however have a superior accuracy to that of trainees. To achieve this, different features should be experimented to obtain an optimal feature set. Furthermore, one must enlarge the database to diminish the effect of interpatient variability.

References

[1] J. F. Smythe, "Initial evaluation of heart murmurs: are laboratory tests necessary?," *Pediatrics*, vol. 86, no. 4, pp. 497–500, 1990.

[2] D. A. Danford, "Heart murmur in a child," *The Journal of Clinical Outcomes Management*, vol. 9, no. 3, pp. 146–158, 2002.

[3] V. T. Nkomo et al., "Burden of valvular heart diseases: a population-based study.," *The Lancet*, vol. 368, no. 9540, pp. 1005–1011, 2006.

[4] R. L. Watrous, "Computer-aided auscultation of the heart: From anatomy and physiology to diagnostic decision support," in *Conf Proc IEEE Eng Med Biol Soc*, 2006.

[5] M. El-Segaier, *Digital Analysis of Cardiac Acoustic Signals in Children*, Ph.D. thesis, University of Lund, 2007.

[6] T. Oskiper and R. Watrous, "Detection of the first heart sound using a time-delay neural network," in *Computers in Cardiology*, 2002.

[7] H. Liang, S. Lukkarinen, and I. Hartimo, "Heart sound segmentation algorithm based on heart sound envelogram," in *Computers in Cardiology*, 1997.

[8] H. Naseri and M. R. Homaeinezhad, "Detection and boundary identification of phonocardiogram sounds using an expert frequency-energy based metric," *Ann Biomed Eng*, vol. 41, no. 2, pp. 279–92, 2013.

[9] D. Gill, N. Gavrieli, and N. Intrator, "Detection and identification of heart sounds using homomorphic envelogram and self-organizing probabilistic model," in *Computers in Cardiology*, 2005.

[10] C. Ahlstrom, "Feature extraction for systolic heart murmur classification," *Ann Biomed Eng*, vol. 34, no. 11, pp. 1666–1677, 2006.

[11] D. Kumar et al., "Heart murmur recognition and segmentation by complexity signatures," in *Conf Proc IEEE Eng Med Biol Soc*, 2008, pp. 2128–2132.

[12] D. Kumar et al., "Wavelet transform and simplicity based heart murmur segmentation," in *Computers in Cardiology*, 2006.

[13] E. Delgado-Trejos et al., "Digital auscultation analysis for heart murmur classification," *Ann Biomed Eng*, vol. 37, no. 2, pp. 337–353, 2009.

[14] B. Ergen, Y. Tatar, and H. O. Gulcur, "Time-frequency analysis of phonocardiogram signals using wavelet transform: a comparative study," *Comput Methods Biomech Biomed Engin*, vol. 15, no. 4, pp. 371–381, 2012.

[15] A. Castro, T. T. V. Vinhoza, S. S. Mattos, and M. Coimbra, "Heart sound segmentation of pediatric auscultations using wavelet analysis," in *Conf Proc IEEE Eng Med Biol Soc*, 2013.

[16] J. Pedrosa, "Heart sound analysis for cardiac pathology identification: Detection and characterization of heart murmurs," *M.Sc. Thesis. Faculdade de Engenharia da Universidade do Porto*, 2013.

[17] 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.

[18] M. Z. Lam et al., "Factors influencing cardiac auscultation proficiency in physician trainees," *Singapore Medical Journal*, vol. 1, no. 46, pp. 11–14, 2005.