Pediatric Heart Sound Segmentation Using Hidden Markov Model

Pouye Sedighian[1], Andrew W. Subudhi[2], Fabien Scalzo[3], Shadnaz Asgari[1], *IEEE Member*

***Abstract*—Recent advances in technology have enabled automatic cardiac auscultation using digital stethoscopes. This in turn creates the need for development of algorithms capable of automatic segmentation of heart sounds. Pediatric heart sound segmentation is a challenging task due to various confounding factors including the significant influence of respiration on children's heart sounds. The current work investigates the application of homomorphic filtering and Hidden Markov Model for the purpose of segmenting pediatric heart sounds. The efficacy of the proposed method is evaluated on the publicly available Pascal Challenge dataset and its performance is compared with those of three other existing methods. The results show that our proposed method achieves an accuracy of 92.4%±1.1% and 93.5%±1.1% in identifying the first and second heart sound components, respectively, and is superior to three other existing methods in terms of accuracy or computational complexity.**

I. Introduction

Cardiovascular disease is the leading cause of death worldwide. An estimated 17.3 million people died from heart related problems in 2008, representing 30% of all global deaths [1]. Early detection of cardiovascular diseases may enhance the efficacy of treatments. Due its simplicity, non-invasiveness, and low-cost, cardiac auscultation remains a primary - and often only - means of cardiac examination available at basic health care clinics. In cardiac auscultation, an examiner uses a stethoscope to listen to the heart sounds generated by rhythmic contractions of the heart. Heart sounds can provide valuable information about the cardiovascular system and functionality of heart valves [2]. A normal heart cycle consists of two prominent audible components, named the first (S1) and the second (S2) heart sounds. The S1 ('lub') sound is produced when the atrioventricular (AV) valves (tricuspid and mitral) are shut to prevent back flow of blood from the ventricles to the atria during contraction. The S2 ('dub') sound occurs at the closure of semilunar (aortic and pulmonary) valves. Thus, the time interval between S1 and S2 in a single heart cycle represents systole (contraction) and the time interval between S2 and the next S1 represents diastole (relaxation). Figure 1 presents an example of heart sounds across the heart cycle.

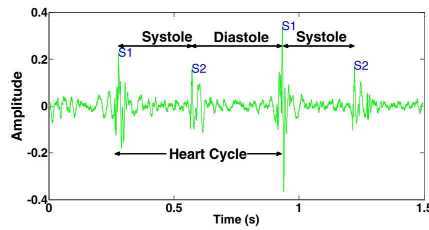

Fig 1. An example of recorded heart sound signal with identified heart cycle, and two major heart sound components (S1 and S2).

Accurate manual auscultation (heart sound interpretation) requires extensive training and experience. Therefore, the reliability of diagnosis is dependent on the level of training or experience of the clinician [3]. Automatic auscultation can address this issue by allowing for an in-depth analysis of heart sounds and identification of abnormal patterns such as murmurs[3, 4]. Using a phonocardiograph (PCG), one can automatically and quantitatively record even the sub audible components of heart sounds, in turn enabling clinicians to diagnose and treat a vast array of heart problems that cannot be detected using a regular stethoscope[5]. The first step in automatic cardiac auscultation is heart sound segmentation by identifying S1 and S2 components in each heart cycle[4]. Several segmentation techniques have been previously developed. While some techniques use information from an additional synchronized signal, such as electrocardiogram (ECG) or carotid pulse, to enhance the segmentation[6], others try to perform segmentation by merely analyzing the PCG signal [7]. Most of these techniques extract a smooth envelope of the signal using wavelet transform [4] or Hilbert transform [3]to detect the peak candidates, and then identify S1 and S2 components within each heart cycle by employing basic physiological assumptions (e.g. systolic time interval is shorter than diastolic time interval)[8]. Nonetheless, accurate heart sound segmentation is still a challenging task and requires further study, especially when the recorded signal is noisy (contaminated by environmental noises, artifacts, and pathological sources) or when the heart rate (HR) is high[9]. Pediatric heart sound segmentation can be particularly challenging. Abnormalities in ECG signals of children with hypertrophic ventricles (caused by axis deviation of the

[1]P.S and S.A. are with Computer Engineering and Computer Science Department of California State University, Long Beach CA 90840 (corresponding author's e-mail: shadnaz.asgari@csulb.edu) .

[2]A.W.S is with the Department of Biology at University of Colorado, Colorado Springs, CO 80918 (e-mail: asubudhi@uccs.edu).

[3]F.S is with the Department of Neurology and Department of Computer Science at University of California, Los Angeles CA 90024 (e-mail: fscalzo@mednet.ucla.edu).

heart) reduce the utility of ECG in heart sound segmentation[10]. Moreover, acquiring additional synchronized signals (e.g. ECG) has its own inconveniences in children. It is also known that respiration has significant influences on cardiac assessment of children. With higher HR in children, systolic and diastolic time intervals become very short, resulting in erroneous heart sound segmentation. Given the above reasons, pediatric heart sound segmentation requires further study.

The current work investigates the application of homomorphic filtering and Hidden Markov Model (HMM) for the purpose of pediatric heart sound segmentation. The efficacy of our proposed method is evaluated on a publicly available dataset of recorded pediatric heart sound signals and its performance is compared with those of three other existing techniques.

II. METHOD

The proposed method consists of four major steps: Preprocessing; homomorphic filtering; peak detection; and peak identification using HMM. Figure 2 illustrates a block diagram of various steps of this method. In the following sections, we will describe each step in detail.

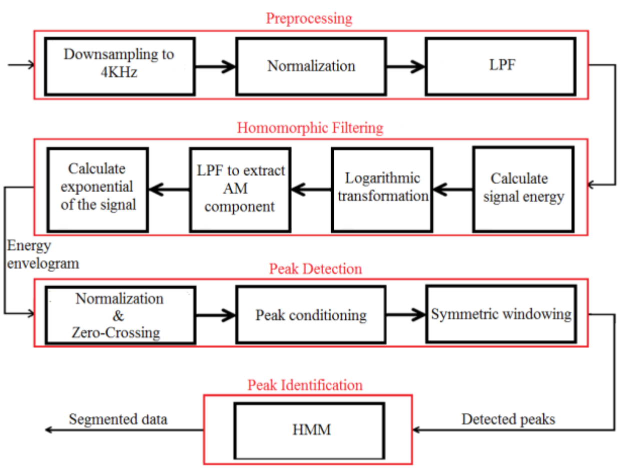

Fig 2. A block diagram of the proposed method

A. Preprocessing

The recorded heart sound signals were first down sampled to 4 KHz and normalized by dividing the signal by its maximum absolute value. Given that the possible frequency range for normal and abnormal heart sounds is between 50 and 700 Hz, a linear phase low pass filter (Chebyshev type I) with cut-off frequency of 750Hz was applied to the data [8].

B. Homomorphic Filtering

Following preprocessing of the data, a homomorphic filtering approach was taken to extract a smooth envelogram[8, 11]. This approach leverages the fact that the main heart sound components resemble amplitude modulated (AM) waveforms (*a(n)*), while abnormal heart sound patterns (murmurs) are similar to amplitude and frequency modulated (FM) waveforms (*f(n)*). Hence, by applying a logarithmic transformation to the energy of the heart sound signal, the resulting spectrum became a linear combination of a slowly varying AM component and a fast varying FM component:

$$x(n) = a(n) \cdot f(n) \rightarrow \log x(n) = \log a(n) + \log f(n) \quad (1)$$

Then, a linear phase low-pass filter (L) whose stop-band attenuates the typical high frequencies of the FM component (a Chebyshev filter with cut-off frequency of 20 Hz) was applied to extract the AM components:

$$L(\log x(n)) = L(\log a(n)) + L(\log f(n)) \approx \log a(n) \quad (2)$$

Finally, a smooth energy envelogram was obtained by calculating the exponential of the output of the low-pass filter-viz. $\exp(\log a(n)) = a(n)$. Figure 3 presents an example of a PCG signal and its extracted energy envelogram using homomorphic filtering.

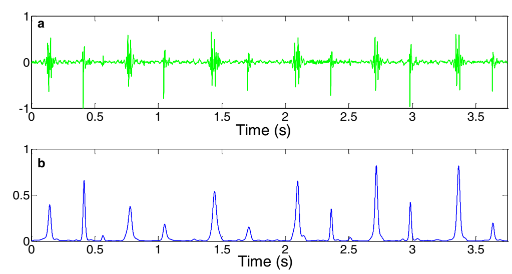

Fig 3. Result of homomorphic filtering: a) PCG signal; b) extracted energy envelogram using homomorphic filtering.

C. Peak Detection

In order to detect S1 and S2 peaks, the energy envelogram was normalized (by removing its average value) and zero crossed. Then the following two criteria were applied for peak conditioning (determining a subset of detected peak candidates which corresponds to the true S1 and S2 peaks):

- All peaks with a width less than half of the average width of detected peaks in a PCG signal were considered spurious peaks and therefore were rejected[8].
- Neighboring peaks with a space less than 150 ms were considered as split heart sounds, and therefore were combined into a single peak[9].

Finally, a symmetric window with time duration of 240 ms was centered at the location of each conditioned peak, and the time of the maximum absolute value of the signal within that window was determined to be the location of the corresponding peak[11].

D. Peak Identification

Following peak detection, a one dimensional and two-state continuous density HMM was applied to classify peaks as either S1 or S2. The two states of our HMM represented the true S1 and S2 events, whereas the observation of each state was the time interval between the current peak and the following one, modeled using a Gaussian distribution whose initial parameter values (mean and covariance matrix) were chosen randomly within some physiologically accepted range of systole and diastole[12]. Given the fact that in general the probability of transitioning from one peak to the same type of peak in a heart cycle is low, the initial transition matrix was chosen as a symmetric matrix whose diagonal elements are random small numbers (less than 0.2). For an unsupervised training of the employed HMM, model parameter estimation was conducted solely based on detected heart sound peaks with no accompanying annotations[11]. For this purpose, the maximum likelihood estimation of the HMM parameters were obtained via the expectation maximization (EM) algorithm. Finally, the Viterbi algorithm was employed to compute the most probable sequence of states given the model parameters and a sequence of observations.

III. Performance Assessment

The performance of the proposed method was evaluated using a 20-fold cross validation (100 realizations) on a publicly available dataset, Pascal Challenge "Classifying Heart Sounds Challenge" (Btraining normal Dataset)[13]. This dataset includes phonocardiograms of pediatric patients with corresponding S1 and S2 annotations. To further assess the efficacy of the proposed method, we compared its performance with those of three other existing methods:

- The Gupta method[8] relies on the assumption that systolic periods are shorter than diastolic periods. Hence, it employs a K-means clustering of extracted time intervals between consecutive peaks to identify S1 and S2 components.
- The Sepehri method[10], identifies S1 and S2 based on the assumption that variance in diastolic periods is greater than that of systolic periods in children.
- The Castro method[9], the most recently developed technique for pediatric heart sound segmentation and tested on the Pascal Challenge dataset, employs wavelet transform to extract Shannon energy envelogram of a PCG signal and detect the corresponding peaks. Then, time and frequency characteristic of heart sounds along with a HR estimate is used to identify S1 and S2 components.

IV. Results

Figure 4 presents the results of heart sound segmentation on a sample phonocardiogram of the employed dataset using the Gupta, Sepehri and our proposed methods. We observe that while the other two time-based methods fail to correctly identify heart sounds, our proposed method is able to correctly segment the entire PCG signal.

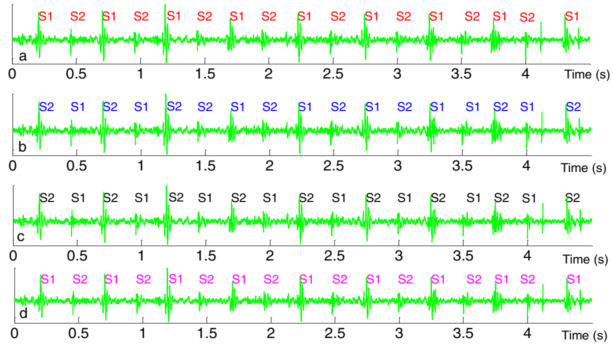

Fig 4. Comparison of the results of heart sound segmentation on a sample phonocardiogram using three different methods: a) clinician annotation; b) Gupta method; c) Sepehri Method; d) proposed method.

Table I summarizes the results of peak identification on 84 PCG records from the Pascal Challenge dataset[9] using the Gupta, Sepehri and our proposed methods. We observe that while the Gupta and Sepehri methods achieve 81.4% and 83.3% accuracy in identification of two primary heart sound components, respectively, our method is 92.9%±1.1% accurate - outperforming the other methods by at least 8%.

TABLE I. Results of peak identification on Pascal Challenge dataset using three different methods (Gupta, Sepehri, and our proposed method). Rate of identification is expressed in percentage.

		Gupta Method		Sepehri Method		Proposed Method	
	#Annotated	#Identified	Rate	#Identified	Rate	Average #Identified	Average Rate
S1	**639**	516	80.8	528	82.6	**591**	**92.4**
S2	**626**	514	82.1	526	84.0	**585**	**93.5**
Total	**1265**	1030	81.4	1054	83.3	**1176**	**92.9**

To complete our study, we compared the performance of the proposed method for pediatric heart sound segmentation with the reported performance of another technique (Castro method) recently developed and tested on the same Pascal Challenge dataset. Figure 5 presents the results of this comparison. We note that our proposed method demonstrates a comparable (slightly higher) accuracy in identification of S1 and S2 heart sounds.

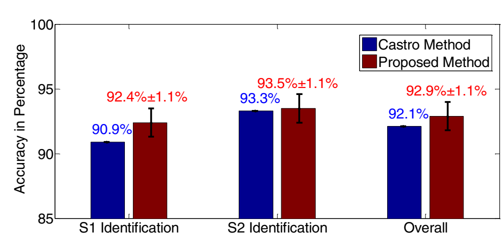

Fig 5. Comparison of heart sound segmentation results of our proposed method (mean ± standard deviation of accuracy over 100 realizations) and the Castro method using the same Pascal Challenge dataset.

V. Discussion and Conclusion

Heart sound segmentation is the first step in detection of heart pathologies through automatic cardiac auscultation. Although several segmentation techniques have been extensively explored before, only a few studies have focused on segmentation of heart sound signal in children. Pediatric heart sound segmentation poses additional challenges due to children's higher HR and significant confounding influence of respiration on auditory cardiac assessment.

The current work explored the application of homomorphic filtering and an unsupervised HMM [11] for pediatric heart sound segmentation. Evaluation of the performance of our proposed method on a publicly available dataset revealed an overall accuracy of 92.9%±1.1% in identification of the two primary heart sound components (S1 and S2). This performance was superior to the existing Gupta and Sepehri methods. The lower performances of these other methods could be explained by the fact that their underlying assumptions do not necessarily hold for children with higher HR. The Gupta method relies on the assumption that systolic periods are shorter than diastolic periods. The Sepehri method assumes the variance in diastolic periods is greater than that of systolic periods.

Comparison of our results with those reported for the Castro method demonstrated that our proposed method performs slightly better. Although the superiority of our proposed method over the Castro method in terms of accuracy of peak identification may not seem significant, the lower computational complexity of the proposed method could make it a better candidate for pediatric heart sound segmentation. In order to correctly identify S1 and S2 components in a recorded PCG signal, the Castro method relies on the estimation of HR by performing singular value decomposition on a matrix whose rows are obtained by running a moving window with varying size on the PCG signal. Although Castro et al. have tried to employ a downsampled contour envelope (by a factor of 10) to achieve a compromise between computation time and HR estimation resolution, the HR estimation is still time consuming and this may hinder the applicability of the Castro method for real-time heart sound segmentation. In fact, our calculations revealed that the average computation time for HR estimation of each PCG record in the Pascal Challenge dataset (using a varying size window from 400 to 1700 ms with incremental size of 5 ms) was 5.45 s, while the average computation time for peak identification of our proposed method was less than 1 ms (using an Intel ® core ™ CPU@ 1.9 GHz-2.4 GHz, 8GB RAM, 64 bit OS). Therefore, the proposed method is more efficient than the Castro method for pediatric heart sound segmentation.

In conclusion, our application of homomorphic filtering and HMM for pediatric heart sound segmentation achieves an accuracy of 92.9%±1.1% for identification of the two primary heart sounds (S1 and S2) and is superior to three other existing methods in terms of accuracy or computational complexity.

Acknowledgment

The presented work was supported by a grant from the California State University Program for Education and Research (CSUPERB). The authors would like to thank Dr. C.N. Gupta for his assistance in implementing homomorphic filtering, and Dr. M. T. Coimbra for providing their validation study dataset.

References

[1] A. Alwan., "Global status report on noncommunicable diseases 2010.," World Health Organization2011.

[2] K. Phua, J. Chen, T. H. Dat, and L. Shue, "Heart sound as a biometric," *Pattern Recognition,* vol. 41, pp. 906-919, 3// 2008.

[3] S. Choi and Z. Jiang, "Comparison of envelope extraction algorithms for cardiac sound signal segmentation," *Expert Systems with Applications,* vol. 34, pp. 1056-1069, 2008.

[4] X. Wang, Y. Li, C. Sun, and L. Liping, "Detection of the first and second heart sound using heart sound energy," in *Biomedical Engineering and Informatics, 2009. BMEI'09. 2nd International Conference on*, 2009, pp. 1-4.

[5] H. Tang, T. Li, Y. Park, and T. Qiu, "Separation of heart sound signal from noise in joint cycle Frequency–Time–Frequency domains based on fuzzy detection," *Biomedical Engineering, IEEE Transactions on,* vol. 57, pp. 2438-2447, 2010.

[6] C. Lima and D. Barbosa, "Automatic segmentation of the second cardiac sound by using wavelets and hidden Markov models," in *Engineering in Medicine and Biology Society, 2008. 30th Annual International Conference of the IEEE*, 2008, pp. 334-337.

[7] D. Kumar, P. Carvalho, M. Antunes, P. Gil, J. Henriques, and L. Eugenio, "A new algorithm for detection of S1 and S2 heart sounds," in *Acoustics, Speech and Signal Processing, ICASSP Proceedings. IEEE International Conference on*, 2006, pp. II-II.

[8] C. N. Gupta, R. Palaniappan, S. Rajan, S. Swaminathan, and S. Krishnan, "Segmentation and classification of heart sounds," in *Electrical and Computer Engineering, 2005. Canadian Conference on*, 2005, pp. 1674-1677.

[9] A. Castro, T. T. Vinhoza, S. S. Mattos, and M. T. Coimbra, "Heart sound segmentation of pediatric auscultations using wavelet analysis," in *Engineering in Medicine and Biology Society (EMBC), 35th Annual International Conference of the IEEE*, 2013, pp. 3909-3912.

[10] A. A. Sepehri, A. Gharehbaghi, T. Dutoit, A. Kocharian, and A. Kiani, "A novel method for pediatric heart sound segmentation without using the ECG," *Computer methods and programs in biomedicine,* vol. 99, pp. 43-48, 2010.

[11] 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, pp. 957-960.

[12] A. Cantor, K. L. Wanderman, T. Karolevitch, I. Ovsyshcher, and M. Gueron, "Systolic time intervals in children: normal standards for clinical use," *Circulation,* vol. 58, pp. 1123-1129, 1978.

[13] P. Bentley, G. Nordehn, M. Coimbra, and S. Mannor, "The PASCAL Classifying Heart Sounds Challenge 2011 (CHSC2011) Results," ed, 2011.