Proceedings - 19th International Conference - IEEE/EMBS Oct. 30 - Nov. 2, 1997 Chicago, IL. USA

A HEART SOUND SEGMENTATION ALGORITHM USING WAVELET DECOMPOSITION AND RECONSTRUCTION

Liang Huiying[1], Lukkarinen Sakari[2], Hartimo Iiro[1]
Helsinki University of Technology
Signal Processing and Computer Technology Laboratory[1]
Applied Electronics Laboratory[2]
Otakaari 5 A, FIN-02150 Espoo, Finland
e-mail: hliang@cc.hut.fi

ABSTRACT

A heart sound segmentation algorithm, which separates the heart sound signal into four parts: the first heart sound, the systolic period, the second heart sound, and the diastolic period has been developed. The algorithm uses discrete wavelet decomposition and reconstruction to produce intensity envelopes of approximations and details of the original phonocardiographic signal. The performance of the algorithm has been evaluated using 1165 cardiac periods from 77 digital phonocardiographic recordings including normal and abnormal heart sounds. The algorithm has shown over 93 percent correct ratio.

Key words: phonocardiography, auscultation, first heart sound, second heart sound, systolic period, diastolic period, murmurs, wavelet decomposition, wavelet reconstruction, segmentation, normalized average Shannon energy.

1. BACKGROUND

Many pathological conditions of the cardiovascular system cause murmurs and aberrations in heart sounds much before they are reflected as other symptoms, such as changes in the electrocardiogram (ECG) signal[1]. Although ECG and ultrasonic examination are widely used in cardiological diagnosis, the old-age art of the heart sound analysis by auscultation is first performed by physicians to evaluate the functional state of the heart. However, the diagnosis is seldom based on the auscultation alone due to the fact that the auscultation is subjective and depends highly on the skills of the interpreter. Still a simple initial examination should be done by the ordinary doctors using effective and objective means before the suspected patient might be sent to the cardiologist for further examinations.

In auscultation the observer tries to listen and analyze the heart sounds components separately, and then synthesize the heard features. The important components of a cardiac cycle which should be identified are: the first heart sound (S1), the systolic period, the second heart sound (S2), and the diastolic period in this sequence in time. The important features, which should be quantified are: the rhythm, the timing instants and relative intensity of the heart sound components, the splitting of S2, the existence of murmurs or other extra sounds, and the timing, intensity and quality of the murmurs and extra sounds. Although plenty of qualitative descriptions of different sounds are available, it is difficult only by listening them to quantify their properties. It might be concluded that objective and effective methods based on the quantified features of the heart sound components are needed to make reliable diagnosis.

Before any further analysis of the heart sounds can be done, segmentation of the heart sounds to different components should be performed. Some attempts to segment phonocardiographic (PCG) signals have been reported in the literature. The majority of them depend on the reference of ECG signal or/and carotid pulse[2,3,4]. M. W. Groch showed a solution where the segmentation was based on the time-domain characteristics of the signal[3]. A. Iwata used the frequency spectral tracing of high frequency (100 to 200Hz) PCG signal to segment it into components[4]. David S. Gerbarg, thirty years ago, took advantage of the time relations of the signal components to separate them without a reference to ECG[5]. The identification results showed a 92 percent correctness for the normal adult male recordings and 88 percent correct ratio for elementary-school children recordings recorded from the base area. The problem of this algorithm was on handling breathing noise, widely splitted S2, and artifacts.

The purpose of this study is to develop an automatic heart sound segmentation algorithm which would be less sensitive to ambient noises and recording locations and uses the heart sound signal as the only source.

2. MATERIALS

The sound material consists of digital heart sounds recorded on a multimedia PC equipped with an electronic stethoscope[6]. The sounds were recorded with 16-bit accuracy using 11025Hz sampling frequency. No ECG equipment was used in the mean time. The recordings included both pathological and physiological murmurs of children aged from 0.4 to 15.4 years. The samples were recorded from several auscultation locations with duration of 6 to 13 seconds. Totally 1165 cardiac cycles from 77 recordings were used to evaluate the algorithm. An experienced children cardiologist pointed out the correct positions of S1 and S2.

3. METHODS

The segmentation algorithm was used on the selected details and approximations of the original PCG signal. The details and approximations, which corresponded to different frequency bands, were obtained by using wavelet decomposition and reconstruction. For each frequency band signal, a similar segmentation method based on the envelope of corresponding signal was used to get the segmentation results. Then the best result of the different bands was chosen to be the final segmentation result.

The key points of the segmentation algorithm were that S1s and S2s were identified first, the intervals of S1s ant S2s were computed, and then based on these information, the intervals of the systolic and diastolic period were obtained consequently.

The whole procedure was realized in five steps.

Step 1: Decompose the original PCG signal using wavelet decomposition and reconstruct the details and approximations.

In order to identify S1s and S2s correctly, frequency band in which the majority power of S1 and S2 located should be used. Earlier studies using FFT to analyze the frequency contents of the first and second heart sounds have indicated that the frequency spectrum of S1 contains peaks in a low-frequency range (10 to 50Hz) and a medium-frequency range (50 to 140Hz). The frequency spectrum of S2 was found to contain peaks in a low-frequency range (10 to 80Hz), a medium-frequency range (80 to 200Hz) and a high-frequency range (220 to 400Hz)[1]. However, some samples have the most of the energy of S1 and/or S2 in low-frequency range meanwhile other samples have more energy in medium- or high-frequency range. In addition, ambient noises from different sources have quite different frequency contents and intensity range. It is not possible to select a fixed frequency band filter which would be suitable to eliminate noises in all samples. All these factors made us to use several frequency band signals rather than only one in order to get a good segmentation result.

Before decomposition, the original PCG signal was down sampled by factor 5. Since some murmurs having higher frequency than the normal sounds (up to 600Hz)[1] are still below half of the new sampling frequency 2205Hz, any useful events of the heart sounds are not missed. After down sampling, a fifth-level discrete wavelet decomposition of the original signal was done to obtain the coefficients of all the components of the decomposition. Using these coefficients, the details and approximations in desired level were obtained by reconstruction. The details and approximations vary depending on the wavelet families and orders used in the decomposition and reconstruction. Order six Daubechies filters, which have 11 taps, were used in our work. The frequency bands of the details and approximations in our case are:

1st-level detail (d1): 551 to 1102Hz;
2nd-level detail (d2): 275 to 551Hz;
3rd-level detail (d3): 138 to 275Hz;
4th-level detail (d4): 69 to 138Hz;
5th-level detail (d5): 34 to 69Hz;
4th-level approximation (a4): 0 to 69Hz;
5th-level approximation (a5): 0 to 34Hz.

After the reconstruction, the original PCG signal S can be expressed as:

$$\begin{aligned} S &= d1 + d2 + d3 + d4 + d5 + a5 \\ &= d1 + d2 + d3 + d4 + a4 \end{aligned} \quad (1)$$

According to the characteristics of the frequency spectrum of S1, S2 and the possible noises, details d4, d5, and approximation a4 were selected as the sources for segmentation. Fig.1 shows an example of the original heart sound recording s, its approximation a4 and details d5, d4 and d3.

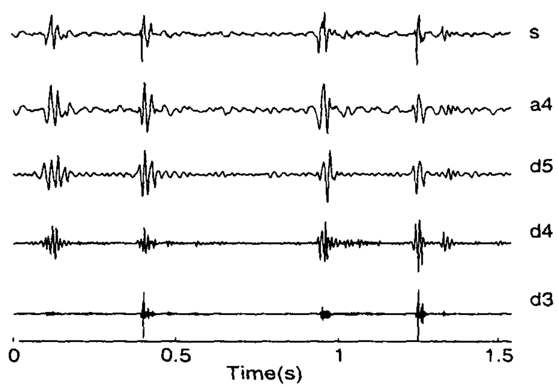

Fig 1. An original signal s, its approximation a4 and details d5, d4 and d3

Step 2: Calculate the normalized average Shannon energy for selected approximations and details.

Each selected signal d4, d5 and a4 was segmented using a segmentation algorithm based on the envelope calculated from the normalized average Shannon energy. The average Shannon energy attenuates the effects of low value noise and makes the low intensity sounds easier to be found. The average Shannon energy was calculated in continuous 0.02-second segments throughout the normalized signal with 0.01-second segment overlapping using the following formula:

$$E_S = -1/N \cdot \sum_{i=1}^{N} x_{norm}^2(i) \cdot \log x_{norm}^2(i) \quad (2)$$

where, the x_{norm} is the signal sample normalized to the maximum absolute value of the studied band signal, and N is the number of samples in 0.02-second segment, here N = 44.

Lastly the normalized average Shannon energy versus the whole time axis was computed as follow,

$$P_a(t) = \frac{E_S(t) - M(E_S(t))}{S(E_S(t))} \qquad (3)$$

where, $M(E_S(t))$ is the mean value of $E_S(t)$;
$S(E_S(t))$ is the standard deviation of $E_S(t)$.

Step 3: Mark the peak location of each block whose level exceed the selected threshold.

The actual heart sound recordings are very complicated and patterns of heart sounds vary significantly from recording to recording. There are some problems preventing us from using a simple threshold to pick up all the S1s and S2s. There might be extra 'peaks' due to the second part of the splitted S2 or other events, like systolic or diastolic clicks caused by dysfunction of the heart. In addition, some peaks, usually the first heart sounds, can be too weak compared with other peaks to be marked. Moreover, artifacts resembling the real peaks both in duration and amplitude might be recorded and will be selected as S1s or S2s. In order to solve these problems, modifications in the threshold setting and detection rules of picking S1s and S2s has been made. Firstly, simple threshold was used to mark all the peak locations of continuous segments exceeding the threshold limit. Then time intervals between two adjacent marks were calculated. According to the mean value and standard variation of the intervals, both lower and higher time interval limits were calculated. These limits were used to remove extra peaks and find lost weaker peaks.

Step 4: Identify S1s and S2s.

After the suspected S1s and S2s has been marked, it is needed to identify which one is S1 and S2. Here, the identification has based on the following facts: (a) the longest time interval between two adjacent peaks in the recording (within 20 seconds) is the diastolic period (from the end of S2 to the beginning of S1); (b) the duration of the systolic period (from the end of S1 to the beginning of S2) is relatively constant compared to the diastolic one. After the longest time interval was found, the start and the end marks of that interval were set as S2 and S1 respectively. Then the intervals forward and backward from the longest interval on were checked. Those marks which destroyed constancy limitations of systolic and diastolic period were discarded and the rest S1s and S2s were identified. The artifacts were discarded in this identifying procedure.

Step 5: Decide the durations of S1s and S2s.

The detected positions of intensity peaks of S1s and S2s indicate the approximate locations of these sounds. The actual boundaries of these sounds were obtained by defining another lower intensity threshold value, which differed from the one for detecting S1 and S2. The boundaries of S1s and S2s were modified by confining the duration within 20ms to 150ms. Finally, the systolic and diastolic periods were decided with a small transition time before and after S1 and S2.

4. RESULTS

The segmentation results of d4, d5 and a4 were compared to choose a best one among them. In some cases all the three band signals gave correct results, in others one or two could give better results than others. The choose criterion was more identified S1s and S2s and less discarded peaks. The priority order of these three band signals is d4, d5 and a4, because d4 can give a relatively accurate boundaries of S1s and S2s. Fig 2. shows an example of segmentation result.

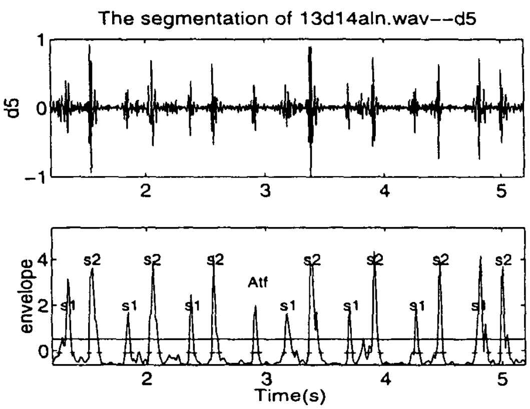

Fig 2. An example of segmentation result.

The performance of the algorithm were evaluated by two sets of samples. One set was selected by the doctor to be considered relatively clear by auscultation and included varieties of murmurs. It included 37 recordings with 515 cycles. Another set was selected by author randomly from a plenty of recordings. It included 40 recordings with 650 cycles. We had previously segmented the first clearer set and the second set recordings using only the low-frequency (10 to 90Hz) part of the original PCG signals because that frequency band has been found to contain the majority (90 percent) of the signal power of S1 and S2 in normal cases. The segmentation results of the previous algorithm[7] indicated 93 and 84 percent correctness respectively to these two sets. When using the algorithm developed in this paper, we got 97 percent correctness for the first set of data. The performance of the algorithm indicates a clear improvement for the second set of recordings: 93 percent correctness. Table 1. and Table 2. summarize the experiment results.

Table 1. Results of the segmentation of the first set recordings.

	previous algorithm		presented algorithm	
	No. of cycles	percentage (%)	No. of cycles	percentage (%)
correct	479	93.0	498	96.7
missed	30	5.8	10	1.9
incorrect	6	1.2	7	1.4
total	515	100	515	100

Table 2. Results of the segmentation of the second set recordings.

	previous algorithm		presented algorithm	
	No. of cycles	percentage (%)	No. of cycles	percentage (%)
correct	543	83.5	604	92.9
missed	55	8.5	20	3.1
incorrect	52	8.0	26	4.0
total	650	100	650	100

5. DISCUSSIONS

One reason of the incorrect identification of S1s and S2s is the high level interfering signals like speech, crying, or other ambient noises, which may overlapped with heart sounds randomly in location, intensity and frequency band and can not be deleted in these three band signals. However, the algorithm presented in this paper can eliminate the effect of respiration signal. The frequency band of which is about 30-40Hz in our case and can not be avoided in some recording procedure to children and infants. This is a benefit of the algorithm compared to the previous one used in our work.

The sudden release of the stethoscope from the patients for a short time during the recording of data had also resulted in incorrect detection. This can be avoided by improving the recording techniques.

Large intensity murmurs overlapped with S1s or S2s will make the correct segmentation and automatic analysis and identification in these three frequency bands impossible. However, because these murmurs are so intense, they can not be neglected by auscultation or by watching the PCG signal.

6. CONCLUSIONS

The presented automatic segmentation algorithm using wavelet decomposition and reconstruction to select suitable frequency band for envelope calculations has been found to be effective to segment phonocardiogram signals into four parts. The algorithm has shown over 93 percent success in 77 recordings with different kinds of murmurs, which include 1165 cycles of heart sounds. This is a good basis for further analysis of the heart sound signals.

7. ACKNOWLEDGMENTS

The authors are grateful to medical doctors Anna-Leena Noponen and Anna Angerla for recording the signals and giving clinical comments.

8. REFERENCES

[1] R. M. Rangayyan and R. J. Lehner, "Phonocardiogram Signal Analysis: A Review", CRC Critical Reviews in Biomedical Engineering, Vol.15, Issue 3, pp.211-236, 1988.

[2] R. J. Lehner and R. M. Rangayyan, "A Three-Channel Microcomputer System for Segmentation and Characterization of the Phonocardiogram", IEEE Trans. on Biomedical Engineering, Vol.BME-34, No.6, pp.485-489, June 1987.

[3] M. W. Groch, J. R. Domnanovich, and W. D. Erwin, " A New Heart-Sounds Gating Devices for Medical Imaging", IEEE Trans. on Biomedical Engineering, Vol.39, No.3, pp.307-310, March 1992.

[4] A. Iwata, N. Ishii, N. Suzumura, "Algorithm for Detecting the First and the Second Heart Sounds by spectral Tracking", Med. & Biol. Eng. & Comput., pp.19-26, Jan.1980.

[5] D. S. Gerbarg, etc., "Analysis of Phonocardiogram by a Digital Computer" , Circulation Research, Vol.11, pp.569-576, Sept.1962.

[6] S. Lukkarinen, P. Korhonen, A. Angerla, A-L. Noponen, K. Sikiö, R. Sepponen, "Multimedia Personal Computer Based Phonocardiography", 18th Annual International Conference IEEE Engineering in Medicine and Biology Society. October 31 - November 3, 1996, Amsterdam, The Netherlands.

[7] H. Liang, S. Lukkarinen, I. Hartimo, "Heart Sound Segmentation Algorithm Based on Heart Sound Envelogram", Computers in Cardiology 1997, September, Lund, Swede. (Submitted)