Heart Sound Segmentation Algorithm Based on Heart Sound Envelogram

H Liang, S Lukkarinen, I Hartimo

Helsinki University of Technology, Espoo, Finland

Abstract

The segmentation algorithm, which separates the heart sound signal into four parts: the first heart sound, the systole, the second heart sound and the diastole, is developed. The segmentation of phonocardiogram signal is the first step of analysis and the most important procedure in the automatic diagnosis of heart sounds. This algorithm is based on the normalized average Shannon energy of PCG signal. The performance of the algorithm has been evaluated using 515 periods of PCG signal recording from 37 objects including normal and abnormal. The algorithm has shown 93 percent correct ratio.

1. Introduction

Noninvasive study (diagnosis) methods, such as phonocardiogram (PCG) and electrocardiogram (ECG), offer useful information of functioning heart. In auscultation, the listener tries to analyze the heart sound components separately and then synthesize the heard features. Heart sound analysis by auscultation highly depends on the skills and experience of the listener [1]. Therefore the recording of heart sounds and analyzing them by a computerized and objective way would be most desirable. Before any automatic analysis can be done, the heart sound needs to be segmented into components and then analyze those components separately. The main components are the first heart sound (S1), the systolic period, the second heart sound (S2), the diastolic period in this sequence in time.

Some attempts to segment PCG signals have been reported in the literature. The majority of them depend on the reference of ECG signal or/and carotid pulse, such as [2],[3] and [4]. M. W. Groch and A. Iwata have shown a solution where the segmentation is based on the time-domain characteristics [3] and the frequency-domain characteristics [4] of the PCG signal, respectively. David S. Gerbarg [5], thirty years ago, took advantage of the time relations of the signal components to separate them based on the signal itself without a reference to ECG using a set of normal recordings.

The purpose of this study is to develop an algorithm for heart sound segmentation which uses the heart sound signal as the sole source. Based on the algorithm, every cycle of the PCG signals is separated into four parts: the first heart sound, the systolic period, the second heart sound and the diastolic period. The locations and intervals of the first heart sounds and the second heart sounds are computed first. Then based on this information, the intervals of the systolic and diastolic period are obtained consequently. Both normal and abnormal heart sound recordings are investigated.

2. Materials

The sound material consists of recordings of heart sounds recorded with a multimedia PC equipped with an electronic stethoscope. The sounds are recorded with 16-bit accuracy and 11025Hz sampling frequency. No ECG equipment has been used. Totally 37 recordings including 14 pathological murmurs and 23 physiological murmurs with total cycles of 515 have been used to evaluate the algorithm. These recordings have been made from children aged from 0.4 to 13.9 years and they are taken at several auscultation locations with duration of 7-12 seconds. The patients have different types of heart diseases. An experienced children cardiologist has pointed out the correct positions of S1 and S2.

3. Methods

The segmentation algorithm is based on the envelope calculated using the normalized average Shannon energy, which attenuates the effect of low value noise and makes the low intensity sounds easier to be found.

3.1. The normalized average Shannon energy

At first, the original signal is decimated by factor 5 from 11025Hz to 2205Hz sampling frequency using an eighth order Chebyshev type I lowpass filter with cutoff frequency at 882Hz. Here, after filtering in the forward direction, the filtered sequence is then reverses and run back through the filter. The resulting sequence has precisely zero-phase distortion, so there is no group delay.

Secondly, the signal is normalized to absolute maximum of the signal according to the equation (1)

$$x_{norm}(k) = \frac{x_{2205}(k)}{\max_i(|x_{2205}(i)|)} \quad (1)$$

where x_{2205} is the decimated signal.

Then, the envelope of normalized decimated signal is calculated. Fig.1 shows different methods of calculating the envelope of the normalized signal. Because of the symmetry of the results as we can see from the following definitions, only the positive part is shown here. The figure is drawn based on the following definitions, here x is the normalized signal, which has the real value from -1 to 1.

Shannon energy: $E = -x^2 \cdot \log x^2$ (4)

Shannon entropy: $E = -|x| \cdot \log|x|$ (5)

absolute value: $E = |x|$ (6)

energy(square): $E = x^2$ (7)

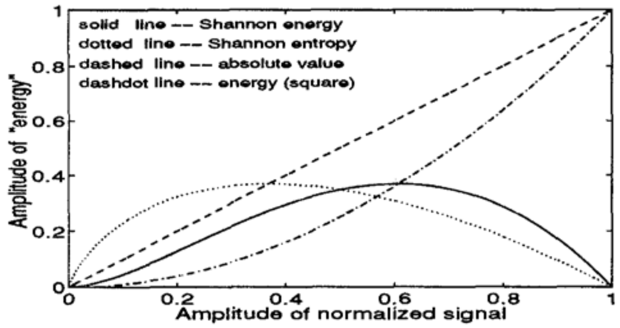

Figure 1. The comparison of different envelope methods.

The figure indicates that the energy (square) will bury the low intensity sounds under the high intensity ones by enlarging the high/low intensity ratio. The Shannon entropy accentuates the effect of low value noise that makes the envelope too noisy to read. The absolute value gives the same weight to all the signal. The Shannon energy emphasizes the medium intensity signal and attenuates the effect of low intensity signal much more than that of high intensity signal. So, the Shannon energy is better than the absolute value in shortening the difference of the envelope intensity between the low intensity sounds and the high intensity sounds. This shortening makes the finding of low intensity sounds easier.

Thirdly, the average Shannon energy is calculated in continuous 0.02-second segments throughout the signal with 0.01-second segment overlapping. The average Shannon energy is calculated as

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

where , the x_{norm} is decimated and normalized sample signal and N is signal length in 0.02-second segments, here N = 44.

Then the normalized average Shannon energy versus time axis is computed. The normalized average Shannon energy is computed as follows,

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

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

3.2. Picking up the peaks

Fig.2 Shows an original PCG signal , which is one of the recordings, and it's normalized average Shannon energy and a threshold. Based on the envelogram calculated by the normalized average Shannon energy curve, a threshold is set to eliminate the effect of noise and the very low-intensity signal. The peaks of each part whose levels exceed the threshold are picked up and assumed temporarily to be the first or the second heart sound. Here, only one peak for each overshoot is chosen even though there are more than one peaks above the threshold. The choice of the peak for each overshoot is based on the following criteria : (1) one peak is always picked up; (2)more than two peaks means the existence of splitted first or second heart sound, so the first peak is picked up in order to get the onset of each sounds.

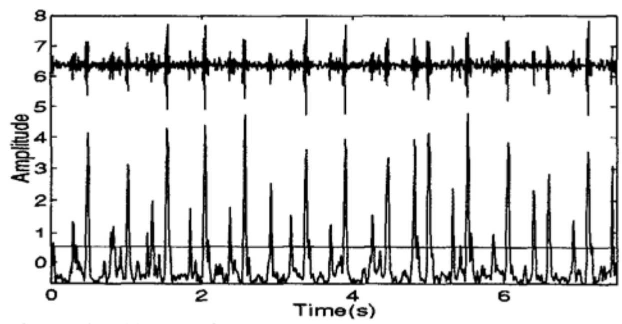

Figure 2. The original signal and it's average Shannon energy.

The actual abnormal heart sound recordings are very complicated and the patterns of heart sounds and murmurs vary largely from recording to recording even for the normal ones. To solve these problems, we made several additions in the procedure.

3.2.1. Rejecting extra peaks

The first problem is that many extra 'peaks' are picked up. This situation is shown in Fig.3. In order to eliminate the extra peaks, the intervals between each adjacent peaks are calculated. The low-level time limit and high-level time limit, which are used for deleting extra peaks and finding lost sounds respectively, are computed for each recording based on the mean value and standard deviation of these intervals. When an interval between two adjacent peaks is less than the low-level time limit, there must be one extra peak which should be rejected. The following criteria are used to decide which one should be dropped: (1) when two peaks appear within 50ms, which is the largest splitted normal sound interval[6], and the energy of the first peak is not too small compared to that of the second one, we picked up the first one. Otherwise, the second one is chosen. Here we assumed that the two peaks that meet the above conditions are actually two parts of a splitted heart sound. Otherwise, the first peak may belong to some other sound class or can be classified as noise. (2) when the interval between two adjacent peaks exceeds 50 ms, their energies are compared. If the energy of the first peak is larger than that of the second one and the last interval meets certain consistency of every second interval, we pick up the first one, otherwise the second one is picked up.

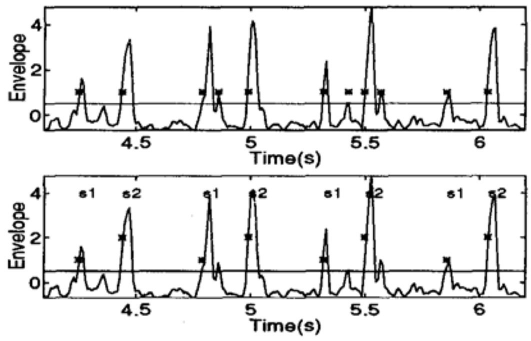

Figure 3. Rejection of extra peaks.

3.2.2. Recovering weak peaks of S1 or S2

Secondly, some peaks are so weak, usually the S1s, that their energies are less than the threshold and they are about to get lost. These weak peaks definitely need to be found, too. To do this, we still examine the intervals between the adjacent peaks. When the interval exceeds the high-level time limit, it is assumed that a peak has been lost and the threshold will be decreased by a certain amount. This reduction of the threshold will be iterated until the losing peaks are picked up or the iteration limit is reached. Here the above criteria to eliminate the extra peaks are used again to delete all "extra" peaks picked up in the procedure of finding lost ones. Fig.4 shows the "lost" peaks are found again.

The third problem is caused by the artifacts, that look like real peaks both in time interval and amplitude and can't be discarded using the methods created for extra peaks. They will be rejected in the following process.

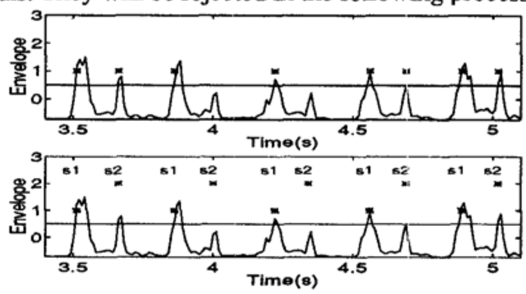

Figure 4. Recovering the "lost" peaks.

3.3 Identifying the S1s and S2s

After all the S1s and S2s have been recognised, we need to identify which one is S1 and which one is S2. Here, the identification is based on the following facts: (1) the largest interval of a recording (within 20 seconds) is the diastolic period; (2) the systolic period is relatively constant compared to the diastolic one. According to these facts, starting from the largest interval, we examine the intervals both forwards and backwards to maintain the relative consistency of systolic and diastolic intervals by different tolerances. Here, parameters c1 and c2 are used to represent the allowable tolerances in percentage of the duration of the systolic and the diastolic period respectively. After we have identified which are the systolic and which are the diastolic periods, we have also implied which are S1s and S2s, and the peaks not included to these sets are detected and discarded. Fig.5 shows the identified S1s and S2s and a discarded artifacts.

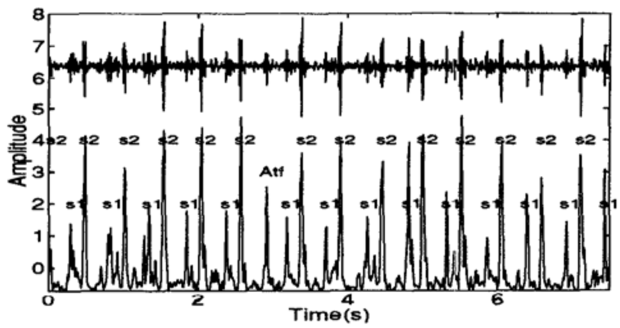

Fig.5 The identified S1 and S2 and artifacts

3.4. Decision of durations of S1 and S2

The detected S1s and S2s indicate the approximate location of these sounds. The actual duration of these sounds are obtained by defining another lower threshold, which differs from the one for detecting S1 and S2, to find the two boundaries for each sound. The detection, however, may lead to inaccurate boundaries because of noise, closed murmur, the third heart sound, the fourth heart sound, etc. We modified the deviations by confining the duration of heart sounds between 20ms and 120ms. The durations of the systole and the diastole are accordingly decided with a small transition period before and after these sounds.

4. Results

The results of the experiment indicate 93 percent correctness in the automatic identification of S1 and S2. The misinterpretations are caused by large background noise and serious murmurs. Table.1 shows the results. If we omit the missing cycles, so that they will not effect our further analysis, higher percentage of correctness will be obtained.

Table 1. The statistics of the results.

	Number of cycles	Percentage(%)
correct	479	93.01
missing	30	5.83
incorrect	6	1.17
total	515	100

The performances of this algorithm with different sets of parameters have also been compared. Table.2 shows the results. The ratios of correctness for these three cases do not differ much. It means that our algorithm is relatively robust to fluctuations of these parameters. In our research, it is the incorrect identification but not the missing sounds that will degrade the performance of the later extraction of the features of each segment. So, even though the third set of parameters can give a little better ratio of correctness, we choose the first set of parameters which gives lower ratio of incorrectness.

Table 2. Comparison of the different parameters.

parameters c1 c2 pthr	correct ratio	missing ratio	incorrect ratio
0.15 0.30 0.5	93.01%	5.83%	1.17%
0.15 0.30 0.4	89.51%	5.44%	5.05%
0.20 0.40 0.5	93.79%	1.36%	4.85%

5. Discussion

One reason of the incorrect identification is the high level of interfering signals like speech, crying, or other ambient noise. Sometimes the sudden release of the stethoscope from the patients for a short time during the recording of data has also resulted in incorrect detection. This can be helped by improving the recording techniques.

Large murmurs that overlap S1 or S2 will make the correct identification impossible. However, because the murmur is large, it can't be neglected by auscultation.

6. Conclusions

The automatic segmentation algorithm is found to be effective to segment phonocardiogram signals into four parts. The algorithm has shown 93 percent success in 37 recordings, which include 515 cycles of heart sounds. This is a good basis for further analysis of the heart sound signals.With this segmentation, we can extract the features of each segment, such as the root mean square, the peak intensity, the peak location, the duration, the splitted interval of S2, etc. We can also do other processes to each segment for the classification.

Acknowledgements

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

References

[1] Sapire DW. Understanding and diagnosing pediatric heart disease: Heart sounds and murmurs. Norwalk, Connecticut, Appleton & Lange 1992: 27-43.

[2] Lehner RJ, Rangayyan RM. A three-channel microcomputer system for segmentation and characterization of the phonocardiogram. IEEE Trans. on Biomedical Engineering 1987; 34: .

[3] Groch MW, Domnanovich JR, Erwin WD, A new heart-sounds gating devices for medical imaging. IEEE Trans. on Biomedical Engineering 1992; 39: 307-310.

[4] Iwata A, Ishii N, Suzumura N. Algorithm for detecting the first and the second heart sounds by spectral tracking. Med. & Biol. Eng. & Comput Jan 1980: 19-26.

[5] Gerbarg DS, etc. Analysis of phonocardiogram by a digital computer. Circulation Research 1962; 11: 569-576.

[6] Hurst JW. The heart arteries and veins, 7th ed. McGraw-Hill Information Services Company, New York 1990: Ch. 14:175-242.

Address for correspondence: Liang.Huiying@hut.fi

Helsinki University of Technology, Signal Processing and Computer Technology Laboratory

P.O.Box 3000, FIN-02015 HUT, FINLAND