THE JOURNAL OF THE ACOUSTICAL SOCIETY OF KOREA Vol.28, No.2E 2009. 6. pp. 72~78

Performance Comparison Between the Envelope Peak Detection Method and the HMM Based Method for Heart Sound Segmentation

Jang Hyun-Baek*, Chung Young-Joo*
*Department of Electronics Engineering, Keimyung University
(Received April 27 2009; revised June 2 2009; accepted June 25 2009)

Abstract

Heart sound segmentation into its components, S1, systole, S2 and diastole is the first step of analysis and the most important part in the automatic diagnosis of heart sounds. Conventionally, the Shannon energy envelope peak detection method has been popularly used due to its superior performance in locating S1 and S2. Recently, the HMM has been shown to be quite suitable in modeling the heart sound signal and its use in segmenting the heart sound signal has been suggested with some success. In this paper, we compared the two methods for heart sound segmentation using a common database. Experimental tests carried out on the 4 different types of heart sound signals showed that the segmentation accuracy relative to the manual segmentation was 97.4% in the HMM based method which was larger than 91.5% in the peak detection method.

Keywords: *Heart Sound Segmentation, Shannon Energy, Peak Detection, Hidden Markov Model*

I. Introduction

Heart sound results from the activity of the heart in the human body and clinicians can make an effective diagnosis of the patient through the auscultation in early stages of the heart disease with small expenses. It is known that, in the auscultation process, clinicians first try to separate the 4 components of the heart sound signal and then analyze each component independently.

As the heart sound analysis by auscultation requires a lot of experience and skills, a computer-aided diagnosis system which can automatically analyze the heart sound signal would be very useful for assisting the clinicians to make better diagnosis of the heart disease. In the system, the heart sound must be segmented into its components before any analysis can be done. It is well known that the more accurate the segmentation becomes, the better performance in the analysis and classification of the heart sound signal we can obtain [1]. From this viewpoint, we can see that the segmentation of the heart sound signal is very crucial in the automatic diagnosis system.

One of the traditional methods of the heart sound segmentation is to use the ECG (Electrocardigram) signal. While the segmentation information obtained from the ECG signal can be used to segment the heart sound signal, it requires the simultaneous recording of the ECG signal with the heart sound signal. To avoid such a burden, some attempts have been made to segment the heart sound signal without a reference to the ECG signal [1][4]. Liang et al. proposed an algorithm called the peak detection method based on the envelope of the normalized

Corresponding author: Chung Young-Joo (yjjung@kmu.ac.kr)
Department of Electronics Engineering, Keimyung University
Daegu, 704-701, Korea

average Shannon energy which emphasizes the medium intensity signal and attenuates the effect of low intensity signal [1]. As the envelope of Shannon energy has peaks near the centers of S1 and S2 of the heart sound signal, the segmentation of S1 and S2 can be done by finding the peaks of the envelope. However, sometimes the envelope of Shannon energy has peaks at the outside of S1 and S2 and no peaks are found in S1 and S2. To cope with such cases, some heuristics are used in the method to eliminate the extra peaks and search for the lost peaks. But the heuristic approach often lack of consistency as it heavily depends on the signal characteristics which are quite volatile. Recently, the hidden Markov model (HMM) has been used success-fully in the heart sound signal classification [2]. Also, the HMM has been shown to be very efficient in segmenting the hear sound signal by using Shannon energy as the feature vectors [3]. However, much is not known about the structure of the HMM and the training methods for the heart sound segmentation. In addition, only clean heart sound signal is used in the experiments not considering more complex signals and no comparison with the conventional peak de-tection method was done.

The purpose of this study is to contrast the con-ventional peak detection method and the emerging HMM based method in terms of how much they are accurate in segmenting heart sounds. The experimental test was carried out on the heart sound signals used for training physicians. The segmentation accuracy was measured by comparing to manual segmentation. Clinical significances were discussed in relation to the difference in the segmentation accuracy between the 2 methods.

II. Methods

2.1. Heart Sound Signal Segmentation Using HMMs

The heart sound signal segmentation based on the HMM is made possible through the modeling of the heart sound signal using the HMM. A four state left-to-right HMM for a cycle of the heart sound signal is shown in Fig. 1 in line with the four com-ponents of the heart sound signal, namely S1, systole, S2 and diastole. The number of states in the HMM is usually determined based on the nature of the signal being modeled. Each state of the HMM in Fig. 1 is assigned to a component of the heart sound signal because the signal characteristics in each component may be thought to be homogeneous.

Chung [2] found that the 4-state left-to-right HMM was sufficient to model a cycle of the heart sound signal. The spectral variability in each state is modeled using multiple mixtures of multivariate Gaussian distributions. Given the observation $o(t)$, the output probability distribution in the state j is given

$$b_j(o(t)) = \sum_{m=1}^{M} c_{jm} N(o(t)\ , \mu_{jm}, \Sigma_{jm}) \quad j = 1, \cdots N \tag{1}$$

where $N(o(t), \mu_{jm}, \Sigma_{jm})$ is a multivariate Gaussian distribution, with mean vector μ_{jm} and covariance matrix Σ_{jm}, each mixture component having an asso-ciated weight c_{jm}. M is the number of mixture components in each state. Also, the transition from the state i to j is controlled by the transition probability as follows.

$$a_{ij} = P(j|i) \tag{2}$$

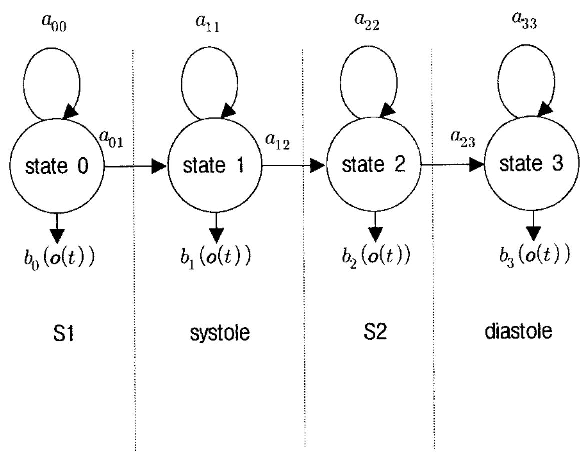


Fig. 1. A left-to-right type HMM for a cycle of the heart sound signal.

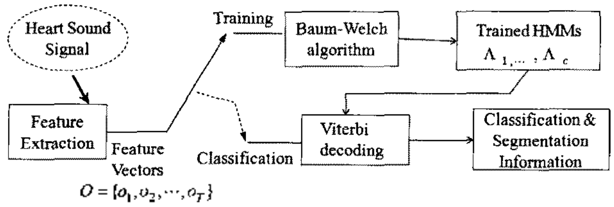


Fig. 2. The procedure of training and classifying the heart sound signal using the HMM.

Fig. 2 shows the training and classification procedure for the heart sound signal using the HMM. During the training, an HMM is constructed for each class of the heart sound signal using the training data corresponding to the class. The feature vectors extracted from the heart sound signal are input to the Baum–Welch algorithm to estimate the HMM parameters $\Lambda = \{\mu_{jm}, \Sigma_{jm}, c_{jm}, a_{ij}\}$ [5–8]. In the classification process, we calculate the likelihood score $P(O|\Lambda_i), i=1,2,3,\cdots,C$ for the sequence of input feature vectors $O=\{O_1,O_2,\cdots,O_T\}$ where C is the number of candidate classes. The class which gives the highest likelihood score is selected as the classification result. In calculating the likelihood score, we use the Viterbi decoding where the optimal state sequence S^* which gives the highest joint likelihood score is obtained as follows.

$$S^* = \underset{S}{\mathrm{argmax}}\, P(S,O|\Lambda) \tag{3}$$

Once we find S^*, the likelihood score $P(O|\Lambda)$ is approximated as follows.

$$P(O|\Lambda) \cong \max_{S} P(S,O|\Lambda) = P(S^*,O|\Lambda) \tag{4}$$

By associating the observation feature vector sequence $O=\{O_1,O_2,\cdots,O_T\}$ with the corresponding optimal state sequence $S^*=\{S_1^*,S_2^*,\cdots,S_T^*\}$, the segmentation information for the heart sound signal is obtained.

2.2. Heart sound signal segmentation based on the Shannon energy envelope

Heart sound signal segmentation using the Shannon energy envelope has been proposed by Laing et al. [3] They detect the peaks of the normalized average Shannon energy to find the location of S1 and S2 in the heart sound signal. The method consists of the following 5 steps. 1) Computation of the normalized average Shannon energy 2) Detection of peaks in the normalized average Shannon energy envelope 3) Rejection of extra peaks 4) Identifying S1 and S2 5) Decision of durations of S1 and S2.

2.2.1. Computation of the normalized average Shannon energy

As the spectrum of S1 and S2 of the heart sound signal is in the frequency range of 50 ~ 100 [Hz], the heart sound signal is passed through the bandpass filter to make it easy to segment the S1 and S2 regions from the systole and diastole regions. The bandpass filtered heart sound signal $x(t)$ is normalized as follows to compensate the amplitude difference between heart sound signals.

$$x_{norm}(t) = \frac{x(t)}{\max\limits_{1 \le s \le T} |x(s)|} \tag{5}$$

Here, T represents the duration of the heart sound signal. From the normalized heart sound signal $x_{norm}(t)$, the average Shannon energy $E_{Shannon}$ can be obtained as follows.

$$E_{shannon} = \frac{-1}{L}\sum_{t=1}^{L} x_{norm}^2(t)\log x_{norm}^2(t) \tag{6}$$

L is the length of each frame where $E_{Shannon}$ is calculated. The length of a frame is 20 msec and the frames are overlapped by 10 msec. The average Shannon energy from Eq. (6) is normalized to obtain the normalized average Shannon energy E_a.

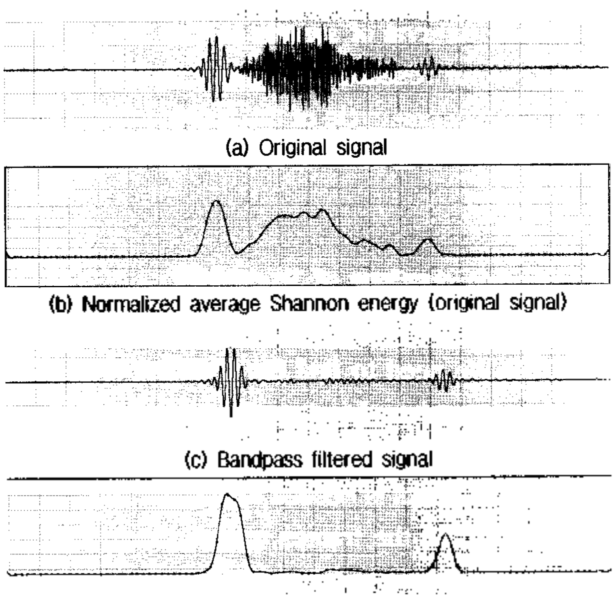
$$E_a = \frac{E_{Shannon} - Mean\{E_{Shannon}(t)\}}{Std\{E_{Shannon}(t)\}} \tag{7}$$

$Mean\{E_{Shannon}(t)\}$ and $Std\{E_{Shannon}(t)\}$ represents the average and standard deviation of $E_{Shannon}(t)$, respectively.

2.2.2. Detection of peaks in the normalized average Shannon energy

Fig. 3 shows examples of the normalized average Shannon energy to one cycle of heat sound signal with or without bandpass (50 ~ 100 Hz) filtering.

We can see that the murmurs in the systole region of the heart sound signal has diminished significantly after the bandpass filtering. Also, the peaks of E_a which appeared in the systole region has disappeared after the filtering, which makes it easy to segment S1 and S2 by detecting the peaks of E_a. Based on the normalized average Shannon energy, a threshold is set to eliminate the effect of noise and the very low-intensity signal. The peaks whose levels exceed the threshold are picked up and assumed temporarily to be S1 and S2. The threshold is lowered iteratively until we pick up more than 2 peaks,.



(a) Original signal

(b) Normalized average Shannon energy (original signal)

(c) Bandpass filtered signal

(d) Normalized average Shannon energy (Bandpass filtered signal)

Fig. 3. An example of the normalized average Shannon energy to one cycle of heat sound signal with or without band-pass (50~100 Hz) filtering. (a) the original heart sound signal, (b) the normalized average Shannon energy to the original signal, (c) the bandpass (50~100 Hz) filtered signal, (d) the normalized average Shannon energy to the bandpass filtered signal.

2.2.3. Eliminating extra peaks

It would be desirable if the peaks are picked up only in the regions of S1 and S2. But, in reality, due to the large variability of the heart sound signal, we can pick up extra peaks in other regions.

The time interval between two adjacent peaks is used to eliminate the extra peaks. The low-level time limit is computed based on the mean and standard deviation of the time intervals. When an interval between two adjacent peaks is less than the low-level time limit, there must be one extra peak among them which should be rejected. In that case, if the time interval exceeds 50 msec, the peak which has smaller normalized average Shannon energy is rejected. If the time interval is less than 50 msec, it is thought that a splitting of S1 has occurred and the second peak is rejected unless the normalized average Shannon energy of the first peak is significantly smaller than that of the second peak.

2.2.4. Identifying S1 and S2

If the peaks have been found, we need to identify S1 and S2. The identification is done based on the fact that the interval of systole is less than that of diastole. As our research is focused on segmenting S1 and S2 from one cycle of the heart sound signal, we identify the first peak as S1 and the second one as S2.

2.2.5. Deciding durations of S1 and S2

As the detected peaks only indicate the center of S1 and S2, we need to find the duration of S1 and S2 for the complete heart sound segmentation. A new threshold which is lower than that used for picking up the peaks is needed to find the boundaries of S1 and S2. All the frames around the peaks are considered to belong to S1 or S2 if their normalized average Shannon energy exceeds the new threshold.

III. Experimental Results and Discussion

The heart sounds for the experiments were taken from the clinical training audio CDs for the physicians [9]. For the HMM modeling, filterbank outputs from the fast Fourier transform were used as the feature vectors. Considering that the length of one cycle of the heart sound signal is in the range of 500~1000 msec, the frame length and rate for the feature vector were set to 7.5 msec and 2.5 msec, respectively. For the extraction of normalized average Shannon energy, the original heart sound signal is down sampled to 2 kHz and bandpass filtered with the passband in the range of 50~100 Hz. The normalized average Shannon energy is extracted in every frame whose length is 20 msec and the frames are overlapped by 10 msec.

For the experiments, we used 4 kinds of heart sound signals, CLEAN, AR(Aortic Regurgitation), AS (Aortic Stenosis) and TR(Tricuspid Regurgitation). The CLEAN signal has no murmurs in the systole and diastole, but the others are related with some kind of diseases and have rather complex signal shapes due to the murmurs in the systole and diastole. There are totally 67 cycles of heart sound signal and we used a cross-validation method in the experiments.

For the heart sound segmentation based on the HMM, a four state left-to-right HMM is used and each state is modeled by an output probability density function with 15 Gaussian mixtures. 13-th order filterbank outputs in addition to their regression coefficients (delta and acceleration coefficients) constitute the 39-th order feature vector in each frame.

For the heart sound segmentation based on the peak detection method, the normalized average Shannon energy is calculated in every frame.

We manually segmented the heart sound signal into S1, systole, S2 and diastole and used the manual segmentation information as a reference in calculating the accuracy of the automatic segmentation.

In Table 1, we compared the segmentation accuracy of the HMM based method and the peak detection method. The accuracy of the segmentation methods was described in terms of how much percentages of the computer-aided automatic segmentations coincide with the manual segmentation in the number of frames. As each frame of the heart sound signal can be determined to belong to one of the four components of the heart sound signal as the result of the automatic segmentation, the segmentation accuracy is easily calculated.

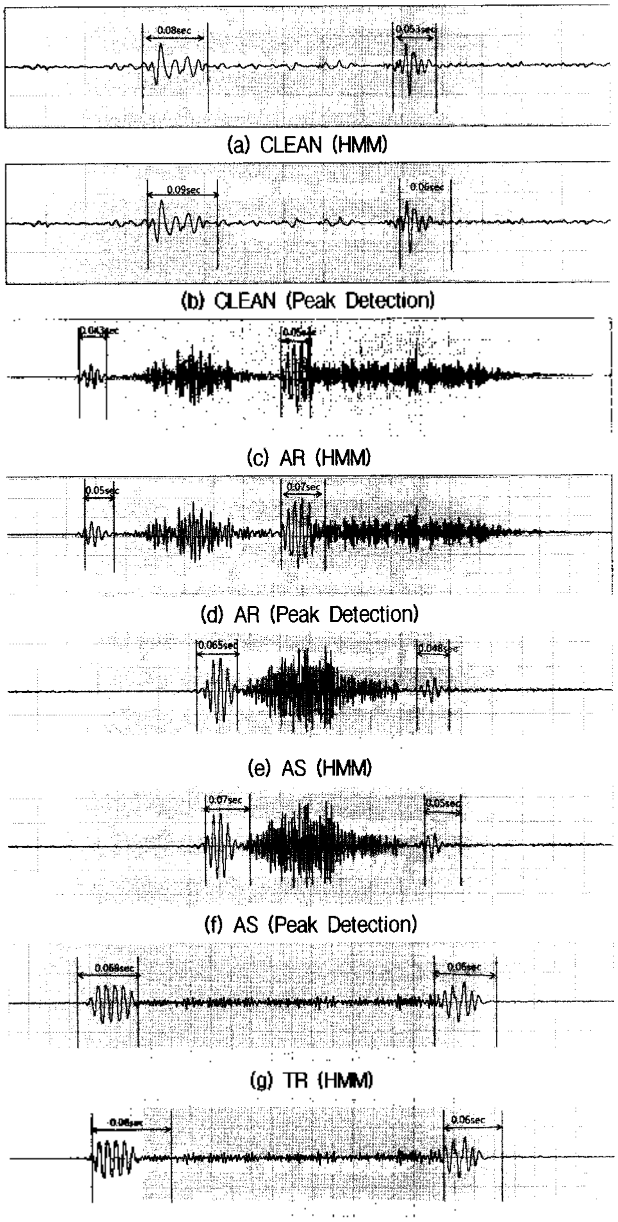
As seen in Table 1, in all of 4 kinds of the heart sound signal, the segmentation accuracy of the HMM based method was much better than that of the peak detection method. For the significance test of the segmentation accuracy, we also show the 95% confidence interval in Table 1. If there is no overlap between the confidence intervals of the two methods, we can say from the statistics that there is a significant difference in the performance of the two methods. From the confidence intervals shown in the table, we can see that the segmentation accuracy of the HMM based method is significantly better than that of the peak detection method.

From the experiments, we could see that the peak detection method successfully found the peaks in S1 and S2 without error in all of the heart sound signals as exemplified in Fig. 3. But the peak detection

Table 1. Segmentaion accuracy (%) of the peak detection method and the HMM based method.

Segmentation Method / Kinds of heart sound signal	HMM (95% confidence interval) (%)	Peak Detection (95% confidence interval) (%)	Number of Frames
CLEAN	95.7(±1.19)	89.6(±1.79)	1114
AR	98.8(±0.56)	95.2(±1.11)	1411
AS	96.8(±0.82)	92.1(±1.95)	1770
TR	98.1(±0.63)	88.9(±1.46)	1772
Average	97.4(±0.79)	91.5(±1.40)	6067

method has some drawback in accurately determining the threshold to determine the durations of S1 and S2. We determined the threshold based on the E_a values in the systole and diastole regions. But the murmurs and other noises in the systole and diastole led to the inaccurate threshold. In contrast, the HMM based method can incorporate such factors into the HMM parameters and can thus lead to more accurate segmentation result.



(a) CLEAN (HMM) (b) CLEAN (Peak Detection) (c) AR (HMM) (d) AR (Peak Detection) (e) AS (HMM) (f) AS (Peak Detection) (g) TR (HMM) (h) TR (Peak Detection)

Fig. 4. Examples of contrast between the segmentations made by the envelope peak detection method and the HMM based method for the 4 types of heart sound signals considered. (a) CLEAN (HMM) (b) CLEAN (Peak Detection) (c) AR (HMM) (d) AR (Peak Detection) (e) AS (HMM) (f) AS (Peak Detection) (g) TR (HMM) (h) TR (Peak Detection).

Fig. 4 contrasts between the segmentation made on the 4 types of the heart sound signals by the HMM based method and the peak detection method. For all 4 kinds of the heart sound signal, we can see that the HMM based method more correctly segments the S1 and S2 regions than the peak detection method which includes the systole and diastole parts into S1 and S2. As there are always some murmurs and noises in the systole and diastole, the exact segmentation of S1 and S2 based on the threshold determination must be a difficult task.

IV. Conclusions

The present study contrasted the conventional peak detection method for heart sound signal segmentation to an emerging method based on the HMM. Experimental tests showed that the segmentation accuracy relative to the manual segmentation was 97.4 % in the HMM based method, which was higher than 91.5% achieved in the peak detection method. As the analysis and classification of the heart sound signal for the automatic diagnosis is decisively dependent on the segmentation results, the superior performance of the HMM based method shown in this paper is quite encouraging from the perspective of clinical treatment.

Acknowledgement

This work was supported by the Korea Research Foundation(KRF) grant funded by the Korea government (MEST) (No. 2009-0075480)

References

1. H. Liang, S. Lukkarinen and L. Hartimo, "Heart Sound Segmentation

Algorithm Based on Heart Sound Envelogram," *Computers in Cardiology* pp. 105-108, 1997.
2. Y. Chung, "A Classification Approach for the Heart Sound Signals Using Hidden Markov Models," *SSPR/SPR* pp. 375-383, 2006.
3. A.D. Ricke, R.J. Povinelli and M.T. Johnson, "Automatic segmentation of heart sound signals using hidden Markov models," *Computers in Cardiology*, pp. 953-956, 2005.
4. D. Gill, N. Intrator and N. Gavriely, "A Probabilistic Model for Phonocardiograms Segmentation Based on Homomorphic Filtering," *18-th Biennial International EURASIP Conference Biosignal* pp.87-89, 2006.
5. L.E. Baum, T. Petrie, G. Soules, N. Weiss, "A Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains," *Annals of Mathematical Statistics* vol. 41 pp. 164-171, 1970.
6. K.F. Lee, *Automatic Speech Recognition*, Kluwer Academic Publishers, 1989.
7. B.H. Juang, and L.R. Rabiner, "A probabilistic distance measure for hidden Markov models," *AT&T Tech. J.*, pp. 391-408, 1984.
8. L.R. Rabiner, J.G. Wilpon, and B.H. Juang, "A segmental k-means training procedure for speech recognition," *IEEE Trans. ASSP*, pp. 2033-2045, 1990.
9. D. Mason, *Listening to the Heart*, Hahnemann University, 2000.

[Profile]

• **Jang Hyun-Baek**

He is a student at the Department of Electronics, Keimyung University. His major research interests include digital signal processing and its applications.

• **Chung Young-Joo**

He received the B.S. degree in Electronics Engineering from Seoul National University in 1988 and he also received M.S. and Ph.D. degrees in Electrical Engineering from Korea Advanced Institute of Science and Technology in 1990 and 1995, respectively. Since 1999, he has been with Keimyung University as a professor at the Dept. of Electronics Engineering.