

A Classification Approach for the Heart Sound Signals Using Hidden Markov Models*

Yong-Joo Chung

Department of Electronics, Keimyung University,
Daegu, S. Korea
yjjung@kmu.ac.kr

Abstract. Stethoscopic auscultation is still one of the primary tools for the diagnosis of heart diseases due to its easy accessibility and relatively low cost. Recently, many research efforts have been done on the automatic classification of heart sound signals for supporting clinicians to make better heart sound diagnosis. Conventionally, automatic classification methods of the heart sound signals have been usually based on artificial neural networks (ANNs). But, in this paper, we propose to use hidden Markov models (HMMs) as the classification tool for the heart sound signal. In the experiments classifying 10 different kinds of heart sound signals, the proposed method has shown quite successful results compared with ANNs achieving average classification rate about 99%.

1 Introduction

Heart sound auscultation is still a very important method for the diagnosis of heart diseases with its easy accessibility and the relatively low cost. However, detecting symptoms of various heart diseases by auscultation requires a skill that takes years of experiences in the field. As the skill is not easy to acquire for junior clinicians, an automatic classification system of the heart sound signal would be very useful for assisting the clinicians to make better diagnosis of the heart disease [1][2].

The dynamic spectral characteristic and non-stationary nature of the heart sound signal makes the automatic classification difficult. As the heart sound signal comes from the human body, it may be highly variable from cycle to cycle and even according to the patient's conditions. To overcome these problems, a classifier which can take into account the variability should be used in the automatic diagnosis of the heart sound signal.

Artificial neural networks (ANNs) based approaches have been widely used in the classification of heart sound signals with some success [1],[2]. ANNs are known to be efficient in classifying complex patterns and have been traditionally used for problems in bio-medical and image pattern classification [3]. However, neural networks are not designed to be suitable for time sequential input patterns like heart sound signals. In fact, nearly all ANNs used for heart sound signals are just static pattern

* This work has been supported by The Advanced Medical Technology Cluster for Diagnosis and Prediction at KNU, which carries out one of the R&D Projects sponsored by the Korea Ministry Of Commerce, Industry and Energy.

classifiers. The heart sound signal should be segmented before use in training and the ANN can only recognize the segmented data. In real situation, it may not easy to segment the relevant parts of the heart sound signal because the automatic classification system should accept the continuous input signal. Even if we can successfully segment one cycle of the heart sound signal, the non-stationary characteristics within the cycle makes it inappropriate to consider the whole samples in the cycle as an input to the ANNs, because we may fail to focus on the time-varying characteristics within the cycle. For the efficient signal classification, methods which can analyze separately the stationary parts in a cycle and later combine them would be desirable. But, neural networks have difficulty in integrating the time-varying statistical characters of the heart sound signal.

In contrary, hidden Markov models (HMMs) have shown quite successful results in classifying time sequential patterns like speech signals [4]. HMMs with its Markov chain structure can inherently incorporate the time sequential character of the signal. By using the Gaussian mixture densities, the HMMs are also expected to faithfully represent the various spectral characteristics of the heart sound signal. The non-stationary nature of the heart sound signal may be well represented by the state transitions in HMMs. As will be shown later in this paper, we could find from the classification experiments that HMMs are very efficient for modeling the dynamic and non-stationary nature of the heart sound signal.

In the next section, we will explain in detail methods how to construct a classifier using HMMs and, in section 3, we show experimental results which demonstrate the feasibility of using HMMs in classifying the heart sound signal. And finally, we make conclusion in section 4.

2 Methods

2.1 Hidden Markov Models

The basic theory of HMMs has been published in a series of papers by Baum [5] in the late 1960s and early 1970s. The underlying assumption in the use of the HMM for signal modeling is that the signal can be well characterized as a parametric random process and the parameters of the random process can be estimated in a well-defined manner. The HMM has shown to provide a highly reliable way of recognizing time sequential patterns like speech. A simple three-state HMM is shown in Fig.1 [6].

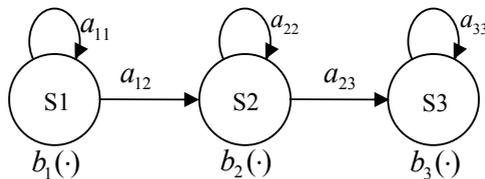


Fig. 1. A 3-state hidden Markov model

The transitions between states are controlled by the transition probability

$$a_{ij} = P(S_j | S_i). \quad (1)$$

And, the probability of generating an observation is determined by the output probability distributions in each state which are usually modeled by a mixture of multivariate Gaussian distributions. Given the observation $\mathbf{y}(t)$, the output probability distribution in state S_j is given by

$$b_j(\mathbf{y}(t)) = \sum_{m=1}^M c_{jm} N(\mathbf{y}(t); \boldsymbol{\mu}_{jm}, \boldsymbol{\Sigma}_{jm}). \quad (2)$$

where $N(\mathbf{y}(t); \boldsymbol{\mu}_{jm}, \boldsymbol{\Sigma}_{jm})$ is a multi-variate Gaussian distribution, with mean vector $\boldsymbol{\mu}_{jm}$ and covariance matrix $\boldsymbol{\Sigma}_{jm}$, each mixture component having an associated weight c_{jm} .

The HMM, as an acoustic model is required to determine the probability of the time sequential observation data $Y = \{y(1), y(2), \dots, y(T)\}$. This is done by computing the probability of the observation data $P(Y | M_i)$, $i = 1, 2, \dots, V$. Here M_i represents an HMM corresponding to a class in the classification problem and V is the total number of classes. The classification is performed by finding the class k which gives the best likelihood score.

$$k = \arg \max_{i=1,2,\dots,V} P(Y | M_i) \quad (3)$$

In practice, Viterbi decoding is employed to find the class with the best likelihood score. Meanwhile, the estimation of the parameters $(\boldsymbol{\mu}_{jm}, \boldsymbol{\Sigma}_{jm}, c_{jm})$ of the HMM is an optimization problem based on some appropriate criterion. Maximum likelihood estimation (MLE) is the most popular one and it tries to match the HMM to the training data as closely as possible. For the efficient MLE training of the HMM, the Baum-Welch algorithm based on Expectation-Maximization (EM) technique is commonly used [7].

2.2 Classification of Heart Sound Signals Using HMMs

A cycle of heart sound signals consists of four elements. The first one called S1 is heard when the mitral and tricuspid valve is closed. The second one S2 is related with the closure of the aortic and pulmonary valve. The systolic and diastolic phase refers to the intervals between S1 and S2 during which any sound is hardly heard in the normal case. We used a four-state HMM to model a cycle of the heart sound signal as shown in Fig. 2. The number of states in the HMM may be determined based on the nature of the signal being modeled. We assumed that each state of the HMM corresponds to an element of the heart sound signal because the signal characteristics in each element are thought to be homogeneous. The spectral variability in each state is modeled using multiple mixture components. By trial and error, we determined the number of mixture components in each state to be 10.

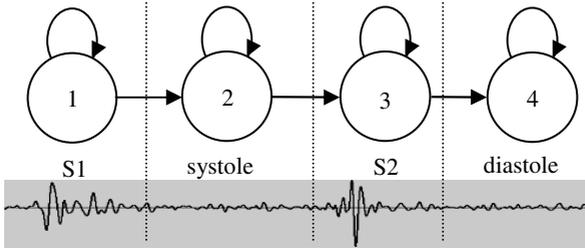


Fig. 2. An HMM for a cycle of the heart sound signal

In Fig. 3, we show the procedure of classifying the heart sound signal using the trained HMMs. The HMM parameters are estimated during the training procedure. For the initial parameter estimation, every cycle of the heart sound signal is manually segmented into 4 regions giving the statistical information corresponding to each element [7]. The feature vectors used are mel-frequency cepstral coefficients (MFCCs) and filter bank outputs both derived from the fast Fourier transform (FFT). MFCCs are popularly used for speech recognition [6] and the filter bank outputs have been usually employed for the spectral analysis of the heart sound signal.

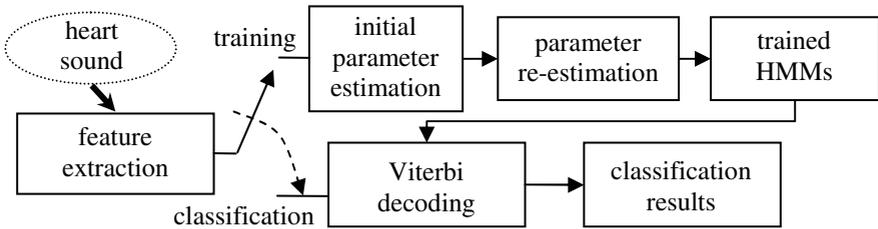


Fig. 3. The procedure of classifying the input heart sound signal using the trained HMMs

3 Results and Discussion

The heart sound data used for the experiments were obtained from the clinical training CDs for the physicians [9]. The original data were down sampled to 16 KHz and stored in a 16 bit resolution. The heart sound signal was already diagnosed and labeled as a specific heart condition. The classification experiments were done using 159 heart sound examples corresponding to 10 different heart conditions: normal sound, innocent murmur, AR (Aortic Regurgitation), AS (Aortic Stenosis), CA (Coarctation of the Aorta), MR (Mitral Regurgitation), MS (Mitral Stenosis), MVP (Mitral Valve Prolapse), TR (Tricuspid Regurgitation) and VSD (Ventricular Septal Defect).

An HMM was constructed for each type of heart condition using the corresponding data. To overcome the problem of small amount of data collected, the classification test was done by the Jack-Knifing method. In the process, the HMM is trained with all

the available examples except the one which is used for the testing. This process is repeated so that all the examples can be used for the testing. The test results are then averaged to give the final classification rate.

The feasibility of modeling the heart sound signal using HMMs can be checked by the segmentation results which can be obtained in the Viterbi decoding. Using the training data, we could verify that each state of the HMM matches quite closely with the respective components of the heart sound signal as we expected. In Fig. 4, we show the matching relations between the HMM states and the heart sound signal waveforms as the number of mixture components varies. In the figure, the vertical lines represent the boundaries between the states.

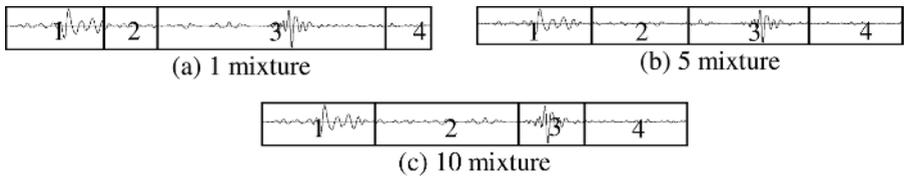


Fig. 4. The matching relations between the HMM states and the heart sound signal waveforms

Although the matching between them is poor when the number of mixtures is 1, it becomes quite accurate as we increase the number of mixture components to 10. With these segmentation results satisfying our expectation, we may conclude that the HMM is quite suitable for the stochastic modeling of the heart sound signal.

Conventional approaches for the heart sound signal classification have been usually based on ANNs [8]. ANNs are known to be able to discriminate complex patterns by generating nonlinear functions of the input. While they have proved useful in recognizing static patterns like spelled characters, they may not be tailored for time sequential input patterns. To compare the performance of the proposed HMM-based classifier with ANNs, we constructed an ANN and tested its performance in recognizing the heart sound signals. The ANN consists of 3 layers (input layer, hidden layer and output layer) connected in sequence. The nodes in each layer process the outputs from the previous layer or directly take in the input features of the heart sound signal. The number of nodes in the input layer is 210 equal to the dimension of the input feature vector and the number of nodes in the hidden and output layer is 20 and 10, respectively [8]. The node in the output layer corresponds to each class of the heart sound signals to be discriminated. In the training, we used 10 different classes of heart sound signals including the normal one.

We investigated the spectral characteristic of the heart sound signal by obtaining the normalized energy spectrum through the fast Fourier transform (FFT). In addition to the frame-level energy spectrum, the energy spectrum of the whole cycle of the heart sound signal was also obtained. The length of the cycle ranges from 500 ms to 1 sec and the frame length was fixed at 2.5 ms. The heart sound signal is processed on a frame basis in the HMM while the whole cycle is fed into the ANN as an input. So,

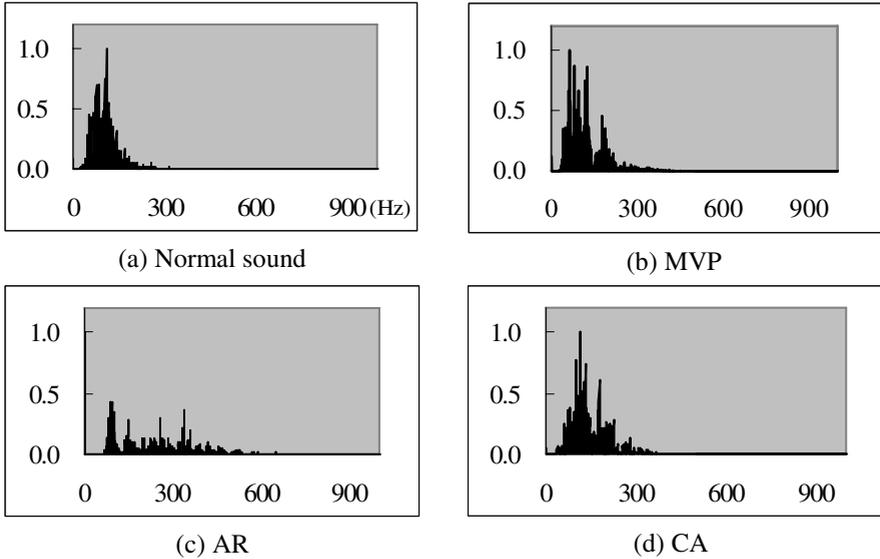


Fig. 5. Normalized energy spectrum of a cycle of the heart sound signal

we may expect that different input spectral characteristics will be modeled in the HMM and ANN, respectively. In Fig. 5, we show the normalized energy spectrum for the whole cycle of the heart sound signals. The heart sound signal seems to contain most of its energy between 0 and 300 Hz, although there are some energy for frequencies up to 600 Hz.

Contrary to the ANN, the input feature vector for the HMM is given on a frame basis. In Fig. 6, the frame-level energy spectrum is shown as time spans. In Fig. 6(a), we can find significant peaks in the energy spectrum at about 200 ms and 600 ms. They seem to correspond to S1 and S2, respectively and their frequency range is between 0 and 300Hz. Meanwhile, for the heart sound signals related with some diseases, there is considerable energy in the diastole and systole phase. In particular, in Fig. 6(c) and (d), we can see some energy peaks between S1 and S2 and their frequency range goes up to 600 Hz. For the MVP case in Fig. 6(b), although the signal waveform looks similar to the normal case, the S2 is usually split as can be seen in the spectrum.

In the experiments classifying heart sound signals, we used two different types of features. In Table 1, we show the classification results when using the MFCC and filterbank outputs. The frequency range for both types of features was relatively broad from 200 to 6400 Hz. The two kinds of features show similar results with marginal improvement obtained in the case of using filterbank outputs. Although the MFCC may be very adequate to speech signals, it was not quite successful to the heart sound signal. With these results, the filterbank outputs are used as the basic input features in the following experiments.

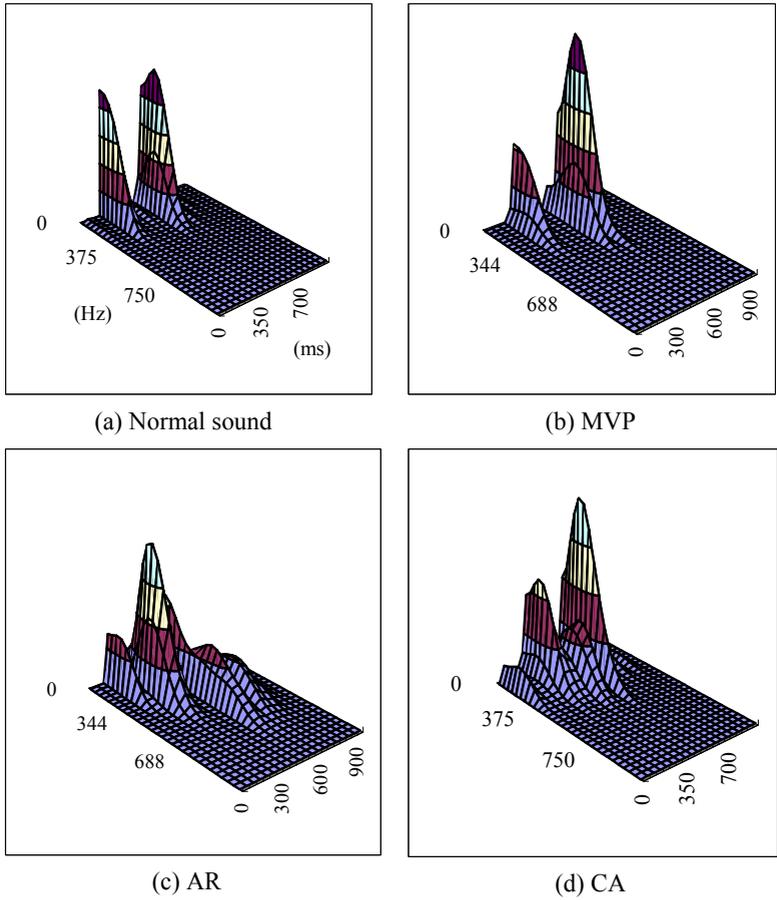


Fig. 6. Normalized frame-level energy spectrum variations with time

As mentioned earlier, the effective frequency range in the spectrum of the heart sound signal is usually well below 600 Hz. So, we experimented by varying the frequency ranges in the filterbank outputs. The frequency ranges considered are 0~900 Hz, 0~210 Hz, 200~900 Hz, 200~300 Hz and 30~900 Hz. The classification results are shown in Table 2. The 0~900 Hz and 200~900Hz ranges performed better than others although there were not significant differences in the classification results with the various frequency ranges. Also, the results in Table 2 were better than the previous results in Table1 indicating that it is important to consider only the relevant frequency ranges in obtaining the filterbank outputs.

To compare the performance of the proposed method with the conventional approaches, an ANN was trained and tested in the classification experiments. Its structure is as described in the above and two different frequency ranges were considered in the filterbank outputs. From the results in Table 3, we can see that the performance of the ANN was quite inferior to the HMM. The HMM's ability to model

efficiently the dynamic time sequential input patterns may be the reason for the superior performance over the ANN.

Table 1. The classification results depending on input features

	MFCC		Filterbank output	
	Accuracy(%)	Correct / Total	Accuracy(%)	Correct /Total
Normal sound	100	15/15	100	15/15
Innocent murmur	92.86	13/14	92.86	13/14
AR	100	14/14	100	14/14
AS	100	18/18	100	18/18
CA	100	20/20	95	19/20
MR	100	21/21	100	21/21
MS	100	14/14	100	14/14
MVP	76.92	10/13	92.31	12/13
TR	100	20/20	100	20/20
VSD	100	10/10	100	10/10
Average	97.48	155/159	98.11	156/159

Table 2. The results with various frequency ranges in the filterbank outputs

Range(Hz)	Average(%)	Correct/Total
0~900	99.37	158/159
0~210	98.74	157/159
200~900	99.37	158/159
200~300	98.11	156/159
30~900	98.74	157/159

Table 3. The classification results of the ANN

Range(Hz)	Average(%)	Correct/Total
0~210	93.08	148/159
0~420	90.56	144/159

4 Conclusion

As a preliminary study of developing an automatic diagnosis system for heart diseases, we proposed a statistical classifier using HMMs. Although the number of diseases to be classified was fairly large compared with the previous works, it achieved a

satisfactory classification rate about 99%. In particular, the proposed method showed quite superior performances compared with the ANN. This seems to come from the HMM's ability to cope with the dynamic time sequential input patterns. However, some of the heart signals were found to be difficult to discriminate. As the HMM is very flexible in modeling the signals, we may improve the discrimination between the models by careful investigation on the heart sound signal characteristics. For example, the number of states and mixture components can be varied depending on the signal types and the amount of the training data. Also features which can contribute more to discriminating between classes can be considered and various estimation criterions for the HMM parameters can be considered for the better modeling.

References

1. Cathers, I.: Neural Network Assisted Cardiac Auscultation. *Artif. Intell. Med.* 7 (1995) 53–66
2. Bhatikar, S.R., DeGroff, C., Mahajan, R. L.: A Classifier Based on Artificial Neural Network Approach for Cardiac Auscultation in Pediatrics. *Artif. Intell. Med.* 33 (2005) 251–260
3. Lippmann, R. P.: An Introduction to Computing with Neural Nets, *IEEE ASSP Magazine* April (1987) 4-22
4. Rabiner, L. R.: A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, *Proceedings of the IEEE*, 77, Feb. (1989)
5. Baum, L. E., Petrie, T., Soules, G., Weiss, N.: A Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains, *Annals of Mathematical Statistics* 41 (1970) 164-171
6. Lee, K. F.,: *Automatic Speech Recognition*, Kluwer Academic Publishers (1989)
7. Rabiner, D. R.,: A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition *IEEE Proceedings* (1998)
8. DeGroff C, Bhatikar S, Hertzberg J, Shandas R, Valdes-Cruz L, Mahajan R, "Artificial neural network-based method of screening heart murmur in children." *Circulation* 103, 2711-6, 2001.
9. Mason D.,: *Listening to the Heart*, Hahnemann University (2000)