Construction of Effective HMMs for Classification between Normal and Abnormal Respiration

Masaru Yamashita^{*} ^{*}Nagasaki University, Nagasaki, Japan E-mail: masaru@cis.nagasaki-u.ac.jp Tel: +81-95-8192580

Abstract- In many situations, abnormal sounds termed adventitious sounds are included as the lung sound of a subject suffering from a pulmonary disease. Thus, we aimed to detect abnormal sounds from auscultatory sound automatically. For this purpose, we expressed the acoustic features of normal lung sound for healthy subjects and abnormal lung sound for patients by using HMMs (Hidden Markov Models) and distinguished between normal and abnormal lung sounds. Furthermore, we detected abnormal sounds under a noisy environment including heart sounds by using a heart sound model. However, the duration time and the property for segments of respiratory, heart, and adventitious sounds varied. In our previous method, we constructed the HMMs with the same number of states and mixtures (topology) for all kinds of segments. Since we did not consider an appropriate topology, the classification rate between normal and abnormal respiration was low (88.96 %). In this paper, we proposed to construct the appropriate HMMs for each segment. By selecting a suitable topology for each segment, the classification rate was increased (91.35 %). The result showed the effectiveness of the proposed method considering the topology of HMMs.

I. INTRODUCTION

Auscultation of the lungs is a means of detecting patients suffering from pulmonary diseases. Despite other noninvasive inexpensive methods, the auscultation using a stethoscope can obtain valuable information regarding health status. In many cases, abnormal sounds (called adventitious sounds [1]) are included in the lung sound of a subject suffering from pulmonary disease, and even today the auscultation is an effective method to diagnose a pulmonary However, it requires expert knowledge and disease. experience. Therefore, perceiving the difference between healthy and ill subjects is difficult for non-medical personnel. This may be the reason auscultation does not penetrate common households. Furthermore, it is difficult for the elderly or persons in depopulated areas to visit the hospital. If we can discriminate between healthy and ill subjects at home, early detection of pulmonary disease can be expected.

Several studies have been conducted with the aim of automatically detecting adventitious sounds from lung sound [2-4]. In these studies, a specific adventitious sound was detected by either using a wavelet transform or a frame of adventitious sound was discriminated by using the short-time spectrum. However, the time of occurrence and duration of adventitious sounds vary. Therefore, it is desirable to discriminate the sound using the features of the whole



Fig. 1 Example of respiratory sounds including adventitious sound, heart sounds and the other noises [7].

respiration and its inflection. Furthermore, the features of adventitious and respiratory sounds depend on the individual and the degree of progress of the disease. Therefore, we consider that the features should be expressed statistically. In the previous study, we expressed the time-series of acoustic features of the lung sound by constructing the HMMs and discriminated between normal and abnormal respiratory In auscultation, noises hinder detecting sounds [5]. adventitious with high accuracy. The auscultatory sounds often include noises from the body and rustle of the stethoscope. A typical noise from the body is the sound of the heart. Fig. 1 shows examples of respiratory sounds including adventitious, heart sounds, and other noises. The appearance frequency of heart sounds auscultated near the heart is high. The database used in our study includes many heart sounds; consequently, many normal respiratory sounds were identified as abnormal respiratory sounds.

To distinguish adventitious sounds from heart sounds, we constructed a heart sound model by using heart sounds for model learning. As a result, normal respiratory sounds were identified correctly. However in the case of abnormal respiratory sounds, the accuracy decreased.



Fig. 2 Architecture of classification system between normal and abnormal respiration.

We assumed that these models did not befitting. Therefore, we focus on analyzing the topology of the acoustic models. In this paper, we propose to construct HMMs for heart sounds with high accuracy by selecting a suitable number of states and mixtures. Furthermore, we select a suitable number of states and mixtures of HMMs for adventitious sounds.

II. LUNG SOUND DATABASE

A. Hand Labeling

We recorded the lung sounds by using an electronic stethoscope. After that, we manually performed segmentation based on recorded sounds, waveform, spectrogram, and power. At first, we divided the lung sound into the inspiration and expiration sound segments (respiratory sound segment). Next, we divided the respiratory sound segment into adventitious sound segments and the other breathing sound segments. In addition, we marked the heart sound segments on the lung sounds that were recorded from auscultation points near the heart. As we can observe the first sounds (S1) and second sounds (S2) clearly, we marked S1 and S2 as heart sound. If the occurrence interval of adventitious sounds and heart sounds were shorter than 100 ms we regarded them as one segment.

B. Definition of Normal and Abnormal Respiration

Acoustic features of some noises are similar to adventitious sounds. Some respiratory sounds from healthy subjects include the adventitious sound. It is difficult for a nonmedical person to diagnose it. Conversely, some respiratory sounds from the patient do not include adventitious sounds. However, we cannot term them as normal respiratory sounds. Then, we defined normal and abnormal respiration. In our study, we grouped the respiratory sounds into four categories.

-Abnormal respirations from patients (AP): respirations include adventitious sounds from patients.

-Abnormal respirations from healthy subjects (AH): respirations include noises resembling the adventitious sounds from healthy subjects.

-Normal respirations from patients (NP): respirations do not include adventitious sounds or noises resembling the adventitious sounds from patients.

-Normal respirations from healthy subjects (NH): respirations do not include adventitious sounds or noises resembling the adventitious sounds from patients.

In our discrimination experiment, we used only NH as normal respiration and AP as abnormal respiration. That is, we do not use AH and NP.

III. DETECTION OF ABNORMAL RESPIRATION

A. Fundamental of Classification Procedure

Generally in the field of speech recognition, the acoustic models of the phoneme (as smallest unit of speech) and the occurrence probability of words are used to construct stochastic models. Then, we applied the technique to the lung sound. Fig. 2 shows the architecture of the classification system between normal and abnormal respiration [6].

The classification procedure consists of the training and test processes. In the training process, the HMMs as the acoustic model and the segment sequence model [6] that defines the occurrence probability of the divided segments are trained. In the test process, input respiration is discriminated between normal and abnormal respiration based on the maximum likelihood approach. If we assume that sample respiration W consists of N segments, it can be expressed as $W = w_1w_2 \cdots w_i \cdots w_N$ where w_i is the *i*-th segment of W.

The training process can be explained as follows. First, we extract acoustic features and train each segment. In the case of normal respiration, if we assume it does not include heart sounds it consists of one segment (N=1). Conversely, in the case of abnormal respiration including adventitious sound, it consists of at least two segments ($N \ge 2$). For example, in the case of expiration in Fig. 3, it consists of one wheeze segment and two breathing segments (N=3). In the case of inspiration in Fig. 3 that does not include adventitious sound, it consists of one breathing sound segment (N=1). The training of the segment sequence model can be explained as

follows. We calculate the occurrence probability of the segments P(W) by using segment bigram. P(W) can be written as

$$P(W) = w_1 \times \prod_{i=2}^{N} P(w_i | w_{i-1}).$$
Let $P(w_i | w_{i-1})$ be defined as
$$(1)$$

$$P(w_i|w_{i-1}) = C(w_i|w_{i-1}) = (w_{i-1}, w_i)/C(w_{i-1}),$$
(2)

where $C(w_i)$ is the count of w_i , $C(w_{i-1})$ is the count of w_{i-1} , and $C(w_i|w_{i-1})$ is the count of segment w_i after the w_{i-1} in the database for training.

The test process can be explained as follows. The maximum likelihood among the calculated likelihoods is found and the corresponding segment sequence \widehat{W} is selected to recognize the sample respiration sound. If the sequence includes at least one adventitious sound, we identify the sample respiration as abnormal sound. Conversely, we identify the sample respiration as a normal sound.

 \widehat{W} can be written as

 $\widehat{W} = \operatorname{argmax}_{W}(\log P(X|W) + \alpha \log P(W))$ (3) where X is a sample respiration and $\log P(X|W)$ is an acoustic likelihood. The weight factor was obtained experimentally.



Fig. 3 Example of respiratory sounds including adventitious sound called wheeze.

B. Classification Procedure Using Heart Sound Model

To distinguish the adventitious sounds from heart sounds, we constructed a heart sound model in addition to the breathing sound model and adventitious models [7, 8]. S1 and S2 are the subjects of modeling from among four types of heart sounds (S1-S4) because they are often observed clearly.

In the training process, we trained the acoustic models. In the case of normal respiration sound, we trained the normal sound model using the breathing and heart sound segments $(N \ge 1)$. In the case of abnormal sound, we trained the model the same as the fundamental classification procedure. In the test process, the maximum likelihood among the calculated likelihoods is found and the corresponding segment sequence \widehat{W} is selected to recognize the sample respiration sound the same as the fundamental classification procedure. The difference from the fundamental classification procedure is that even if the sequence includes some heart sounds, we identify the sample respiration as normal sound.

IV. CONSTRUCTION OF APPROPRIATE HMMS FOR DETECTION OF ABNORMAL RESPIRATION

In our previous studies [5-8], we set the number of states and mixtures of HMMs for each segment as three and assumed the models were not suitable. Therefore, we focus on analyzing the topology of acoustic models. For example, the duration of the stationary sound period is significantly different between heart and adventitious sounds. Table 1 shows the mean and standard deviation (S.D.) of the duration of adventitious and heart sounds. The duration of heart sounds is shorter than the adventitious sounds. Then we focus on the model for adventitious and heart sounds.

To construct the appropriate HMMs for adventitious and heart sounds, we assume that the appropriate topologies of HMMs depend on the constancy of acoustic features and the amount of training data. Then, we construct the suitable HMMs by selecting several numbers of states and mixtures for each segment and we investigate whether the accuracy is improved or not, when the numbers are changed.

Table 1 Mean and standard deviation of duration for adventitious and heart sounds [8] (s)

for adventitious and heart sounds [8]. (s)		
Source sound	Mean	S.D.
Adventitious sound	0.53	0.31
Heart sound	0.12	0.03

V. EVALUATION EXPERIMENTS

A. Experimental Conditions

In every 10 ms, 6 Mel-Frequency Cepstrum Coefficients (MFCCs) and power were extracted as acoustic features using a 25 ms Hamming window. The lung sound data were sampled at 5 kHz. Fig. 4 shows the auscultation points. In this study, auscultated lung sounds from three points near the heart (A-C) were used for experiments. The numbers of abnormal respiratory sounds were 161, 217, and 206 for each point. As many normal respirations were selected randomly for experiments. The number of observed heart sound Since there were no significant segments was 4940. differences between acoustic features of S1 and S2, we constructed one heart sound model without distinctions We performed a leave-one-out cross between them. validation to construct the subject independent model.



Fig. 4 Auscultation points.

B. Classification Experiments

First, we decided on the number of states of HMMs for the adventitious sound segments. We found a suitable number of states from one to five where the number of the mixture was three. Fig. 5 shows the classification rate between normal and abnormal respiration for each number of states for the adventitious sound segments. The classification rate was the highest when the number of states was three. Hereafter, we set the number of states for the adventitious sound a suitable number of mixtures from one to five. Fig. 6 shows the classification rate between normal and abnormal respiration for each number of mixtures from one to five. Fig. 6 shows the classification rate between normal and abnormal respiration for each number of mixtures for the adventitious sound segments. The suitable number was two. We considered that when the number was large the amount of training data was not sufficient. That is, the model was over-fitting.



Fig. 5 Classification rate between normal and abnormal respiration for each number of states for adventitious sound segment.



Fig. 6 Classification rate between normal and abnormal respiration for each number of mixtures for adventitious sound segment.

In the next step, we found a suitable number of states and a number of mixtures for HMMs of the heart sound segment. Hereafter, we set the number of states for adventitious sounds segment to three and set the number of mixtures for adventitious sounds segment as two. This is because the values were the best number in the previously mentioned experiments. At first, we decided on the number of states of HMMs for the heart sound segment. We found a suitable number of states from one to five where the number of the mixture is three. Fig. 7 shows the classification rate between normal and abnormal respiration for each number of states for the heart sound segment. The classification rate was the highest where the number of states was two. Therefore, we considered that two states were sufficient to express the heart sound segment. This is because the heart sound segment was shorter than the adventitious sound segment and there was not much steady-state. Then we found a suitable number of mixtures from one to five where the number of states was two. Fig. 8 shows the classification rate between normal and abnormal respiration for each number of mixtures for the heart sound segment. The suitable number was two. We considered that when the number was large the amount of training data was not sufficient as the adventitious sounds segment. The above result shows the significant effectiveness (p = 0.026) of setting the suitable states and mixtures of HMMs for adventitious sounds segment and heart sound segment.



Fig. 7 Classification rate between normal and abnormal respiration for each number of states for heart sound segment.



Fig. 8 Classification rate between normal and abnormal respiration for each number of mixtures for heart sound segment.

VI. CONCLUSIONS

In this paper, we proposed to construct an appropriate HMM for heart and adventitious sounds with high accuracy by selecting a suitable number of states and mixtures to distinguish between normal and abnormal respiration sounds.

As the result of the classification experiment, we confirmed the improvement of classification rate by selecting suitable states and mixtures of HMMs for adventitious and heart sound segments. This demonstrated the effectiveness of the proposed approach. When we used a large number of mixtures of HMMs the classification rate was low, and it was considered that the amount of training data was not sufficient.

Future work includes the verification of the effect in the classification between healthy and ill subjects by using not only one respiration sample but a series of respirations. Subsequently, we should clarify the suitable number of states, mixtures and other parameters by using deep neural network, which has been proved effective in the field of speech recognition.

ACKNOWLEDGMENT

This work was supported by "Ono Charitable Trust for Acoustics" and "Kawai Foundation for Technology & Music".

REFERENCES

- [1] N. Gavriely, "Breath sounds methodology," CRC Press, 1995.
- [2] Y.P. Kahya, S. Yerer, and O. Cerid, "A wavelet-based instrument for detection of crackles in pulmonary sounds," EMBC 2001 (International Conference of the IEEE Engineering in Medicine and Biology Society), 2001.
- [3] A. Marshall, and S. Boussakta, "Signal analysis of medical acoustic sounds with applications to chest medicine," Journal of the Franklin Institute 344, pp. 230–242, 2007.
- [4] S.A. Taplidou and L.J. Hadjileontiadis, "Wheeze detection based on time-frequency analysis of breath sounds," Computers in Biology and Medicine, Vol. 37, pp.1073-1083, 2007.
- [5] S. Matsunaga, K. Yamauchi, M. Yamashita, and S. Miyahara, "Classification between normal and abnormal respiratory sounds based on maximum likelihood approach," Proc. of ICASSP 2009 (IEEE International Conference on Acoustics, Speech, and Signal Processing), pp.517-520, 2009.
- [6] H. Yamamoto, S. Matsunaga, M. Yamashita, K. Yamauchi and S. Miyahara, "Classification between normal and abnormal respiratory sounds based on stochastic approach," Proc. of ICA 2010 (International Congress on Acoustics), Paper No.671, 2010.
- [7] M. Yamashita, M. Himeshima, and S. Matsunaga, "Robust classification between normal and abnormal lung sounds using adventitious-sound and heart-sound models," Proc. of ICASSP 2014 (IEEE International Conference on Acoustics, Speech, and Signal Processing), pp. 4451-4455, 2014.
- [8] M. Taguchi, M. Yamashita, and S. Matsunaga, "Detection of abnormal lung sounds considering spectral and temporal features of heart sounds," Proc. of ICA 2016 (International Congress on Acoustics), 2016.