CLASSIFICATION BETWEEN NORMAL AND ABNORMAL RESPIRATORY SOUNDS BASED ON MAXIMUM LIKELIHOOD APPROACH

Shoichi Matsunaga, Katsuya Yamauchi, Masaru Yamashita and Sueharu Miyahara

Department of Computer and Information Sciences, Nagasaki University, JAPAN

ABSTRACT

In this paper, we have proposed a novel classification procedure for distinguishing between normal respiratory and abnormal respiratory sounds based on a maximum likelihood approach using hidden Markov models. We have assumed that each inspiratory/expiratory period consists of a time sequence of characteristic acoustic segments. The classification procedure detects the segment sequence with the highest likelihood and yields the classification result. We have proposed two elaborate acoustic modeling methods: one method is individual modeling for adventitious sound periods and for breath sound periods for the detection of abnormal respiratory sounds, and the other is a microphone-dependent modeling method for the detection of normal respiratory sounds. Classification experiments conducted using the former method revealed that this method demonstrated an increase of 19.1% in its recall rate of abnormal respiratory sounds as compared with the recall rate of a baseline method. It has also been revealed that the latter modeling method demonstrates an increase in its recall rate for the detection of not only normal respiratory sounds but also for abnormal respiratory sounds. These experimental results have confirmed the validity of our proposed classification procedure.

Index Terms— acoustic signal detection, biomedical acoustics, pattern classification, lung sounds

1. INTRODUCTION

The auscultation of lung sounds is one of the most popular medical examination methods used for diagnosing many types of disorders. The auscultation of lung sounds is useful also because it does not cause any physical strain to patients. To detect abnormalities (adventitious sounds such as wheeze) in lung sounds, however, much experience and knowledge as a doctor is required. Children occasionally hesitate to reveal their illness or to visit hospitals when ill. Further, there are a number of people who find it difficult to visit hospitals frequently due to their unsuitable living conditions. They eventually visit the hospital after developing serious diseases such as heavy pneumonia, etc. In these cases, the automated detection of abnormal respiratory sounds using a stethoscope at home can alleviate the unpleasant conditions, and appropriate medical treatment can be administered to these patients at an early stage.

A number of studies have been conducted on the acoustic analysis of breath sounds from the view point of the detection of specific adventitious lung sounds [1-4]. In these studies, large-scale lung-sound database were needed to derive reliable experimental results. The Marburug respiratory sound (MARS) database is a typical set of lung sounds collected from more than 300 patients [5]. However, these studies have not been aimed at developing devices for the detection of abnormal respiratory sounds at home but in hospitals to help doctors with performing diagnoses.

Since our purpose was to develop a technology for the detection of abnormal respiratory sounds for use at home, we acquired lung sound data from patients and able-bodied subjects, and we developed a classification procedure for distinguishing between normal respiratory sounds and abnormal respiratory sounds that included the adventitious sounds [6]. Preliminary classification results indicated that the stochastic method is promising, but precise modeling for abnormal respiratory sounds was required to achieve a higher classification performance.

To address this problem, we have proposed a new classification procedure to distinguish between normal and abnormal respiratory sounds based on a maximum likelihood approach using hidden Markov models (HMMs). For calculating the likelihood, we assumed that one section of each inspiratory/expiratory period consisted of a time series of acoustic segments that express specific acoustic features such as adventitious sounds. We hand labeled our recorded data and created a transcription corpus using segment symbols. The classification procedure comprises a training process and a test process. In the training process, acoustic models for the normal and abnormal respiratory sounds are trained using this transcribed database. In the test process, the classification procedure detects the segment sequence with the highest likelihood and yields the classification results. For the precise acoustic modeling in this procedure, each acoustic model for adventitious sounds and breath sounds are used to express abnormal respiratory sounds. Experimental results revealed that this modeling demonstrated a drastic increase in its recall rate for the detection of abnormal respiratory sounds as compared with that of a baseline method that uses a single model for the detection of abnormal respiratory sounds. Furthermore, we also developed different acoustic models depending on the type of microphone used for recording in order to express normal respiratory sounds. It was experimentally confirmed that this modeling method demonstrated an increase in the recall rate for the detection of not only normal respiratory sounds but also abnormal respiratory sounds.

2. LUNG SOUND DATABASE

Lung sounds from 109 patients with emphysema pulmonum and 53 able-bodied subjects were recorded in three hospitals. These sounds were divided into two sets, according to the type of recording instruments (stethoscope) used. In one of the sets, a condenser microphone was attached to the subjects' chest and back

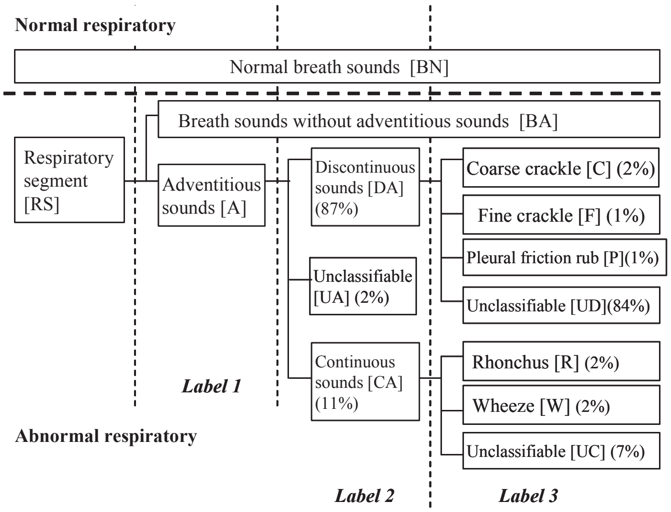

Figure 1: Hierarchical structure of the segment labels

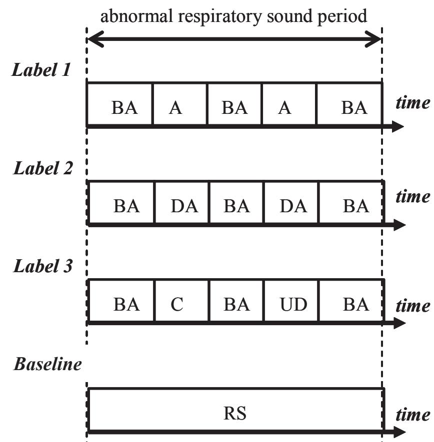

Figure 2: Segmentation of each label

by using a rubber coupler. In the other set, an electronic stethoscope incorporating a piezoelectric microphone was used. The acoustic characteristics of these two sets were different. The number of recording positions was six: two positions on the front and four positions on the back of each subject. In this paper, sounds recorded from the anterior portion on the right side of the second intercostal space were used for the experiments.

Each lung sound was divided into several respiratory phase segments, and these segments were labeled according to the respiratory phase (inspiratory or expiratory) and diagnostic state (normal or abnormal). Each segment was tested using our proposed classification procedure for distinguishing between abnormal and normal respiratory sounds.

2.1. Hand labeling of acoustic segment

We considered an abnormal inspiratory/expiratory sound to be composed of segments with acoustic characteristics. In order to recognize the diagnostic state using a statistical method, we defined the segments according to their acoustic features and assigned a symbol to each segment. The respiratory data was hand labeled using the symbols. Suppose an inspiratory/expiratory sound w comprises N segments, and let the i-th segment be s_i $(1 \le i \le N)$, then

$$w = s_1 s_2 \cdots s_i \cdots s_N , \qquad (1)$$

where the start time of segment s_{i+1} is the end time of segment s_i. In our database, one abnormal respiratory sound comprised several segments, and one normal respiratory sound comprised one normal breath segment ($N = 1$.)

In order to examine how detailed segmentation should be carried out in order to capture the acoustic features of the abnormal respiratory sounds appropriately, we prepared three types of segmentations: Labels 1, 2, and 3. The relation among these labels is indicated in Figure 1 where [_] indicates the acoustic symbols. Label 1 comprises only adventitious sound segments (A) and breath sound segments (BA) that did not include adventitious sounds. Under Label 2, the adventitious sound segments were classified into three groups: continuous sound segments (CA), discontinuous sound segments (DA), and unclassifiable sound segments (UA) that were difficult to be classified into the discontinuous or continuous sound segments. Under Label-3, the discontinuous segments were classified into four groups: coarse crackle segments, fine crackle segments (C), pleural friction rub segments, and unclassifiable sound segments (UD). The continuous sound segment under Label 3 was also classified into three subgroups: rhonchus segments, wheeze segments, and unclassifiable sound segments (UC). This hierarchy of labels was designed on the basis of the classification by the American Thoracic Society (ATS), and we introduced the three types of unclassifiable labels (UA, UD, and UC) to handle ambiguous data. The segmentation of each label is presented in Figure 2.

2.2. Amount ratio of adventitious sounds

The hand labeling of the respiratory sounds was performed by two experts and one doctor. Extremely noisy respiratory segments were excluded from our database. The number of inspiratory sounds was 740; expiratory sounds, 804; normal respiratory, 990 (64%); and abnormal respiratory sounds, 554 (36%). The number of respiratory segments recorded using the condenser microphones was 885 (57%) and that recorded using the piezoelectric microphones was 659 (43%). The amount ratio of pulmonary adventitious sounds for each acoustic label is also indicated in Figure 1. The amount ratio of continuous sounds among the adventitious sounds was 11% and that of the discontinuous sounds was 87%. However, the intelligibility of each adventitious sound was very low, and a considerable number of data belonged to the group of unclassifiable sound clusters (UD, UC) under Label 3. This indicated the difficulty in improving the performance of the classification procedure by using the detailed acoustic labels.

3. DIAGNOSTIC STATE DETECTION PROCEDURE

The architecture of the proposed classification procedure is shown in Figure 3. The system comprised the training process and the test process. The acoustic feature parameters were extracted in the feature extraction module. In the training process, acoustic models for each segment were generated for each respiratory phase. With regard to normal respiratory sounds, individual acoustic models for each type of a stethoscope (with a condenser or a piezoelectric microphone) were generated. Specifically, we developed two microphone-dependent models for the inspiratory and expiratory sounds. With regard to abnormal respiratory sounds, acoustic

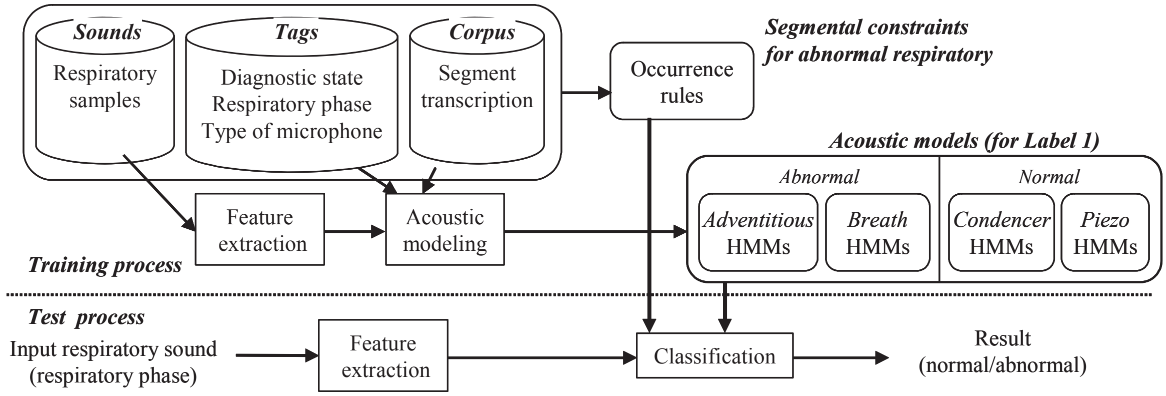

Figure 3: Architecture of classification system between normal and abnormal respiratory

models corresponding to each acoustic segment were generated for the inspiratory/expiratory sounds. Rules concerning the occurrence sequences of the acoustic segments in the abnormal respiratory sounds were also generated using the three labels. The Backus–Naur Form (BNF) was adopted to express these rules. In the test process, the acoustic likelihood of an input respiratory sound was calculated using the trained acoustic models under the constraints of the segment occurrence rules, and the diagnostic state that yielded the segment (sequence) with the highest likelihood was defined as the classification result.

4. EVALUATION EXPERIMENTS

Classification experiments were conducted using the all (1544) the data described in Sec. 2.2. We performed a leave-one-out cross validation on these data. In addition, data recorded from the same subject to the test sample were excluded in the training process to perform subject-independent experiments. The respiratory data were sampled at 10 kHz. Every 10 ms, a vector of 5 mel-warped cepstral coefficients and power was computed using a 25-ms Hamming window. The acoustic models for the normal respiratory sounds were generated using the normal breath sounds (BN in Figure 1). In our experiments, we presupposed that the respiratory phase is known. Subsequently, if the test data was an expiratory sound, the acoustic models generated with the expiratory sounds were used for classification. On the other hand, we presupposed that the recording condition for the test data is unknown. Subsequently, two types of respiratory models (corresponding to the condenser and the piezoelectric microphone) for normal respiratory sounds were used simultaneously.

4.1. Performance of baseline method

A preliminary classification experiment was performed to evaluate the performance of a baseline method based on a maximum likelihood approach. This method was realized using a diagnostic-state tag for each respiratory sound. The acoustic parameters for the entire abnormal inspiratory/expiratory durations (RS in Figures 2) were used to generate abnormal acoustic models. In the test process, HMMs with three states and two Gaussian probability density functions (2-mixture pdfs) were used [7]. The classification result is shown in Table 1. The recall rate of the baseline method for the detection of abnormal respiratory sounds was 72.8% and that of the normal respiratory sounds was 67.9%. The average recall rate weighted with the data amount for each diagnosis phase is indicated as "Average." This result indicated that the maximum likelihood approach using the HMMs is promising for the classification procedure.

4.2. Use of modeling for adventitious sound periods

One of the characteristics of our proposed procedure was the use of HMMs for both adventitious sounds and breath sounds in the modeling of abnormal respiratory sounds. To evaluate the proposed modeling method, a classification experiment was performed wherein the adventitious-sound models and the breath-sound models for the detection of abnormal respiratory sounds, which were generated using the transcription of Label 1, were used. This classification result is also shown in Table 1. On comparing the performance of the proposed modeling method with that of the baseline method, the recall rate of the modeling method for the detection of normal respiratory sounds was found to decrease by 1.7%. However, the recall rate of the modeling method for the detection of abnormal respiratory sounds was increased by 19.1%. This great improvement confirmed the validity of our proposed modeling method for the detection of abnormal respiratory sounds.

4.3. Effect of further detailed segmentation

We used the two sets of acoustic models based on Labels 2 and 3. The Label 2 set comprised continuous sound models, discontinuous sound models, and breath models for abnormal sound periods. The Label 3 set included six types of specific adventitious sounds, two types of unclassifiable models, and breath models, as shown in Figure 1. These classification results are also shown in Table 1.

With regard to the detection of abnormal respiratory sounds, the detailed modeling method (using Labels 2 and 3) as compared to that using Label 1 demonstrated a slight increase in its recall rates. However, the recall rates of the detailed modeling method for the detection of normal respiratory sounds decreased. The acoustic feature of the surrounding noises included in the breath sounds was similar to that of the discontinuous adventitious sounds. Consequently, we considered that detailed modeling for adventitious sounds with a small amount of data resulted in the lack of the robustness for the detection of normal respiratory sound and decreased the recall rate of the detailed modeling method for the detection of normal respiratory sounds.

Table 1: Recall rates for detection of abnormal and normal respiratory sounds for each segmentation (2-mixture) [%]

Label \ Diagnosis	Abnormal	Normal	Average
Baseline (Sec 4.1)	72.8	67.9	71.0
Label 1 (Sec 4.2)	91.9	66.2	82.7
Label 2 (Sec 4.3)	93.2	64.8	83.0
Label 3 (Sec 4.3)	93.1	60.1	81.3

4.4. Use of microphone-dependent modeling

In all the previous experiments, two types of normal breath models (using the condenser and piezoelectric microphones and referred to as "mic-dependent" models in this paper) were used to achieve a better performance with regard to the detection of normal respiratory sounds. To evaluate the performance of the mic-dependent models, we performed a classification experiment using a unified model (referred as "mic-closed" model) for breath sounds in normal respiratory periods. This model was generated using inspiratory or expiratory sounds of the normal respiratory period. The experimental results are shown in Table 2, where the mixture number of pdfs is one or two for the mic-closed models and the mic-dependent models uses a 1-mixture model. The number of acoustic parameters for the 2-mixture mic-closed models was equal to that for the two mic-dependent 1-mixture models. Table 2 indicates that the use of the mic-dependent model enabled the detection of the normal respiratory sounds (67.9%), and also increased the detection rate of the abnormal respiratory sounds by 3.2%, revealing the effectiveness of our modeling method for the detection of normal respiratory sounds.

4.5. Effect of mixture number of Gaussian pdfs

We conducted additional experiments to evaluate the effect of the mixture number of Gaussian pdfs in HMMs. When this number was increased, further detailed modeling could be easily performed. Classification experiments were performed with the mixture number ranging from 1 to 3 using the acoustic models generated with Label 2. The experimental results are shown in Table 3. An

Table 2: Recall rates using mic-closed and mic-dependent models (Label 3) [%]

Acoustic model		Abnormal	Normal	Average
Condition (microphone)	Mixture number			
Closed	1	81.0	46.6	68.7
	2	85.7	35.6	67.7
Dependent	1	88.9	67.9	81.3

Table 3: Recall rates depending on the mixture number (Label 2) [%]

Mixture no.	Abnormal	Normal	Average
1	90.0	70.8	82.8
2	93.2	64.8	83.0
3	93.9	64.6	83.4

increase in the mixture number improved the recall rate of the modeling method for the detection of abnormal respiratory sounds. However, the recall rate in the case of normal respiratory sounds decreased. This implied the difficulty in distinguishing between normal and abnormal respiratory sounds.

5. CONCLUSIONS

In this paper, we have proposed a classification procedure for distinguishing between normal and abnormal respiratory sounds based on a maximum likelihood approach. The acoustic likelihood of an input inspiratory or expiratory lung sound phase was calculated using HMMs that were generated for each diagnostic state, and the diagnostic state that yielded the state with the highest likelihood was defined as a classification result. In our procedure, the acoustic HMMs for abnormal sounds were generated separately for adventitious sounds and breath sounds. These models demonstrated a drastic increase in their recall rate for the detection of abnormal respiratory sounds. The acoustic models for normal respiratory sounds were generated separately according to the data recorded by using microphones. These models demonstrated an increase in their recall rates for the detection of both the normal and the abnormal respiratory sounds. These results conformed the validity of the proposed procedure.

In our experiments, noises included in the normal respiratory sounds prevented the improvement in the recall rate of the proposed procedure for the detection of normal respiratory sounds. This was because the noises were acoustically similar to the adventitious sounds. Lung sounds recorded at a specific position were used in this paper. Our database consisted of lung sound data recorded at six points on the subjects' chest and back. In a future study, s robust classification method for noises will be developed by using the data recorded at different positions.

6. REFERENCES

[1] Y.P. Kahya, S. Yere and O. Cerid, "A wavelet-based instrument for detection of crackles in pulmonary sounds," *Proc. of IEEE 23rd Annual International Conference (EMBS 2003)*, pp.3175-3178, 2001.

[2] M. Bahoura and X. Lu, "Separation of crackles from vesicular sounds using wavelet packet transform," *Proc. of IEEE ICASSP*, II, pp. 1076-1079, 2006.

[3] 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.

[4] A. Marshall and S. Boussakta, "Signal analysis of medical acoustic sounds with applications to chest medicine," *Journal of the Franklin Institute*, Vol. 344, pp.230-242, 2007

[5] V. Gross, L.J. Hadjileontiadis, T. Penzel, U. Koehler, C. Vogelmeier, "Multimedia database Marburg Respiratory Sounds (MARS)," *Proc. of IEEE 25th Annual International Conference (EMBS 2001)*, Vol.1, pp.456-457, 2003

[6] K. Yamauchi, et al, "An examination of a method to detect abnormal sounds from the lung sounds recorded in different environments," *The 32nd annual conference of international lung sounds association*, pp.14, 2007

[7] S. Young, et al, *The HTK Book*, Cambridge University, 2006