Proceedings of 20th International Congress on Acoustics, ICA 2010

23-27 August 2010, Sydney, Australia

Classification between normal and abnormal respiratory sounds based on stochastic approach

Hitoshi Yamamoto, Shoichi Matsunaga, Masaru Yamashita, Katsuya Yamauchi and Sueharu Miyahara

Department of Computer and Information Sciences, Nagasaki University, Nagasaki, JAPAN

PACS: 43.80.QF

ABSTRACT

In this paper, we propose a novel classification procedure for distinguishing between normal and abnormal respiratory sounds on the basis of stochastic approach. The main characteristic of our procedure is that two stochastic models are used to detect abnormal respiratory sounds precisely: (1) hidden Markov models (HMMs) for acoustic spectral features and (2) bigram models for the occurrence of acoustic segments in each inspiratory/expiratory period. The classification procedure comprises a training process and a test process. In the training process, acoustic models for normal and abnormal respiratory sounds are trained using a transcribed database. In the test process, the classification procedure detects the segment sequence with the highest total likelihood and yields the classification results. Our procedure achieved a classification rate of 84.2% between normal and abnormal respiratory sounds. Experimental results revealed that for the classification, use of the segment bigram led to a 4.8% reduction of error rate in comparison with the classification rate of a conventional method that uses deterministic rules to describe segment sequences instead of the segment bigram.

INTRODUCTION

The auscultation of lung sounds is one of the most popular medical examination methods used to identify respiratory illnesses. The auscultation of lung sounds is also useful because patients are able to avoid the adverse effects of radiation and the physical strain and pain associated with computed tomography (CT), magnetic resonance imaging (MRI) or endoscopic inspection. Abnormal respiratory sounds usually appear in lung sounds obtained from patients. These sounds, such as wheezes, are caused by abnormalities of the lungs and bronchial tubes; they are called “adventitious sounds.” Figure 1 shows an example of a spectrogram of abnormal respirations. In these respirations, two periods of wheeze sounds are shown. The intensity and distinctness of adventitious sounds are very low. These sounds look like environmental noises, which are frequently mixed with lung sounds. To detect the adventitious sounds correctly, then, in-depth experience and knowledge that doctors possess are required.

Children are sometimes reluctant to visit the hospitals when sick. There are also a number of people who find it difficult to visit hospitals frequently due to their living conditions. These individuals eventually visit the hospital after developing a serious disease such as heavy pneumonia. In these cases, the automated detection of abnormal respiratory sounds using a stethoscope at home could alleviate the unpleasant conditions, and appropriate medical treatment could be administered to these patients at an early stage.

Several 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 databases were prepared to obtain reliable experimental results. These studies were not, however, aimed at developing devices for the detection of abnormal respiratory sounds at home; instead, they were aimed at assisting doctors in hospitals to make diagnoses.

The objective of our study purpose was to develop a home-use technological device to detect abnormal respiratory sounds. For this purpose, we collected lung sound data from patients and healthy subjects and then developed a classification procedure for distinguishing between normal and abnormal respiratory sounds on the basis of a maximum likelihood approach using hidden Markov models (HMMs) [5,6]. To calculate the likelihood of normal/abnormal respiration, 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. Preliminary classification results indicated that the stochastic method related to acoustic HMMs is promising, but we still used deterministic rules to express the occurrence of acoustic segments in abnormal respiratory sounds. The use of deterministic rules might not be useful in order to achieve a higher classification performance.

To clarify this ambiguity, we attempt to devise a bigram model of acoustic segments instead of using deterministic connection rules among acoustic segments. For calculating the total likelihood of each inspiratory period, we add the acoustic likelihood derived from acoustic HMMs and the occurrence likelihood of acoustic segments derived from the segment bigrams.

The classification procedure comprises a training process and a test process. In the training process, acoustic models for normal and abnormal respiratory sounds are trained using a transcribed database. Furthermore, the segment bigrams in abnormal respirations are trained using the transcription. In the test process, the classification procedure detects the segment sequence with the highest likelihood and yields the classification results. To enable precise acoustic modeling in this procedure, each acoustic model for adventitious sounds and breath sounds is used to express abnormal respiratory sounds. Preliminary experimental results revealed that the segment bigram demonstrated an increase in the classification rate in comparison with that of a baseline method that uses deterministic connection rules for acoustic segments.

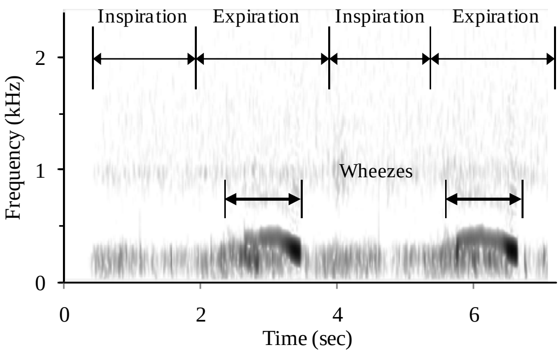


Figure 1. Spectrogram example of abnormal respirations

LUNG SOUND DATABASE

Recording

Lung sounds from 109 patients with pulmonary emphysema and 53 healthy subjects were recorded in three hospitals. These sounds were divided into two sets according to the type of recording instrument (stethoscope) used. In one of the sets recorded in each hospital, a condenser microphone was attached to the subjects' chest and back using a rubber coupler. We refer to this data set as "Set A." In the other set, an electronic stethoscope incorporating a piezoelectric microphone was used. This set is referred to as "Set B." These two sets were recorded in different hospitals and the quantity of ambient noise in Set A was much larger than that in Set B. The acoustic characteristics of these two sets were thus significantly different. There were six recording points on the body: two points on the front and four points on the back. In this study, the recording sounds from the second intercostal space on the subjects' front right were used for the experiments, as shown in Figure 2.

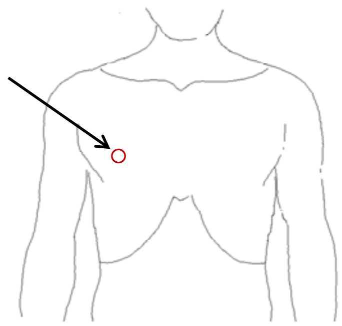


Figure 1. Recording point of the second intercostal space

Each lung sound was divided into several respiratory phase segments. These segments were labeled according to the respiratory phase (inspiratory or expiratory) and diagnostic state (normal or abnormal). Each respiration was tested using our proposed procedure to classify abnormal and normal respiration. The number of normal/abnormal respiratory periods in these sets is listed in Table 1, where Set A+B consists of all data in Set A and Set B.

Table 1. Number of respiratory periods for experiments

Respiratory	Set A	Set B	Set A+B
Normal	206	348	554 (36%)
Abnormal	679	311	990 (64%)

Hand Labeling

We consider an abnormal inspiratory/expiratory period to be composed of segments with acoustic characteristics. In order to determine 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 period w comprises N segments: let the i-th segment be w_i $(1 \le i \le N)$. Then, we have

$$W = w_1 w_2 \cdots w_i \cdots w_N \,, \tag{1}$$

where the beginning time of segment w_{i+1} is the end time of segment w_i. In our database, one abnormal respiratory period comprises several segments, and one normal respiratory period comprises one breath segment ($N = 1$).

In order to examine how detailed segmentation should be carried out in order to effectively capture the acoustic features of abnormal respiration, we prepared three types of segmentations: Labels 1, 2, and 3. The relations among these labels are shown in Figure 3 where [_] indicates the acoustic symbols. Label 1 indicates only adventitious sounds segments (A) and breath sound segments (BA) that do not contain adventitious sounds. In Label 2, the adventitious segments were classified into three kinds: continuous sound segments (CA), discontinuous sound segments (DA), and unclassifiable segments (UA) that are difficult to classify them to discontinuous or continuous segments. In Label 3, the discontinuous segment is classified into four types of sound segments: coarse crackle segments, fine crackle segments (C), pleural friction rub segments, and unclassifiable segments (UD). The continuous segment in Label 3 is also classified into three types of sound segments: rhonchus segments, wheeze segments, and unclassifiable segments (UC). This hierarchical construction of these labels was designed on the basis of the classification used by the American Thoracic Society (ATS). We introduced three kinds of unclassifiable labels (UA, UD and UC) to handle ambiguous data.

DIAGNOSTIC STATE DETECTION PROCEDURE

Formulation

Let the occurrence probability of the segment sequence W be $P(W)$:

$$P(W) = P(w_1 w_2 \cdots w_i \cdots w_N)\,. \tag{2}$$

In this study, we use a segmental bigram to calculate $P(W)$:

$$P(W) \approx \sum_{i=1}^{N} P(w_i \mid w_{i-1})\,. \tag{3}$$

The total likelihood, composed of the acoustic likelihood derived from HMMs and the segmental sequence likelihood derived from equation (3), is calculated using a weight factor α . The diagnostic state (normal/abnormal) that gives the segment (sequence) $\hat{W}$ with the highest likelihood is the classification result as given below.

$$\arg\max_{W} P(W \mid X) = \alpha \log P(W) + \log P(X \mid W) \cdot \quad (4)$$

where X is the unknown respiratory input and $P(X \mid W)$ is the acoustic likelihood. The weight factor α controls the contribution of the occurrence probability of the segmental sequence. If α is equal to 0, classification is carried out using acoustic HMMs only. In this paper, the value of α is experimentally acquired to achieve the best performance. Equation (4) is widely used in the speech recognition field, where $P(W)$ is usually calculated from the word n-gram model.

Classification system

The architecture of our classification system is shown in Figure 4. The system comprises a training process and a test process. Acoustic feature parameters were extracted in the feature extraction module.

In the training process, acoustic HMMs for each segment are generated in the case of each respiratory phase. With regard to normal respiration, individual acoustic models for each type of stethoscope (condenser or piezoelectric microphone) were generated. Thus, we prepared two microphone-dependent models for inspiration and expiration. With regard to abnormal respiration, acoustic models corresponding to each acoustic segment type were generated for inspiration/expiration. Segment bigrams with reference to the occurrence sequences of the acoustic segments in abnormal respiration are also estimated according to the three kinds of hand labeling: Labels 1, 2, and 3.

In the test process, the acoustic likelihood of an input respiration is calculated using the trained acoustic HMMs and segment bigrams. The diagnostic state that gives the segment (sequence) with the highest likelihood is derived as the classification result.

EVALUATION EXPERIMENTS

Experimental conditions

Our discrimination test was conducted using all the 1544 samples. We performed a leave-one-out cross validation on these samples. In addition, samples recorded from the same subject as the test sample were excluded in the training process so that our experiments were subject-independent. The respiratory data were sampled at 10 kHz. Every 10 ms a vector of power and 5 mel-warped cepstral coefficients was computed using a 25-ms Hamming window. Acoustic models for normal respiration were generated using the breath sounds

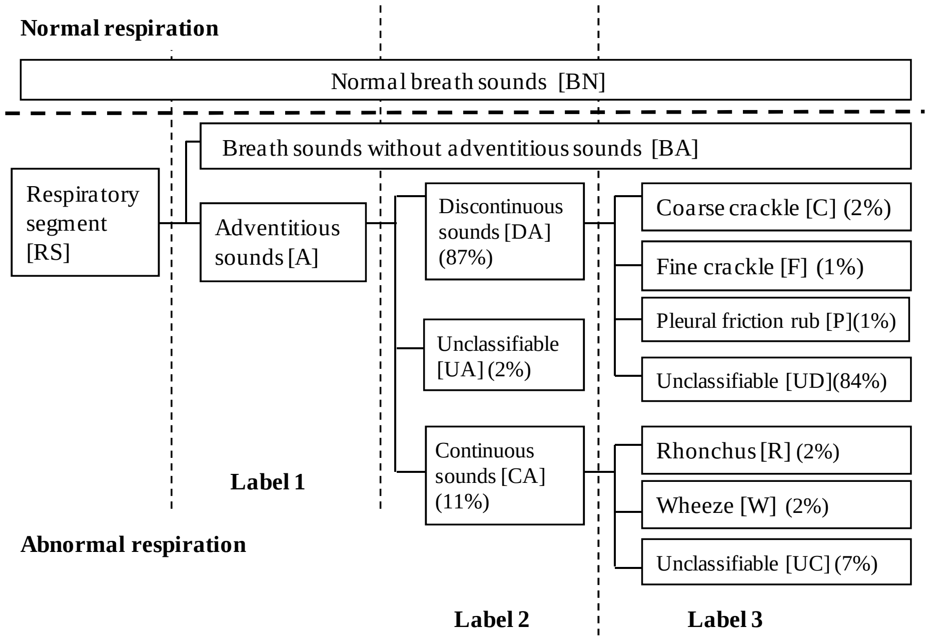


Figure 3. Hierarchical structure of the segment labels [6]

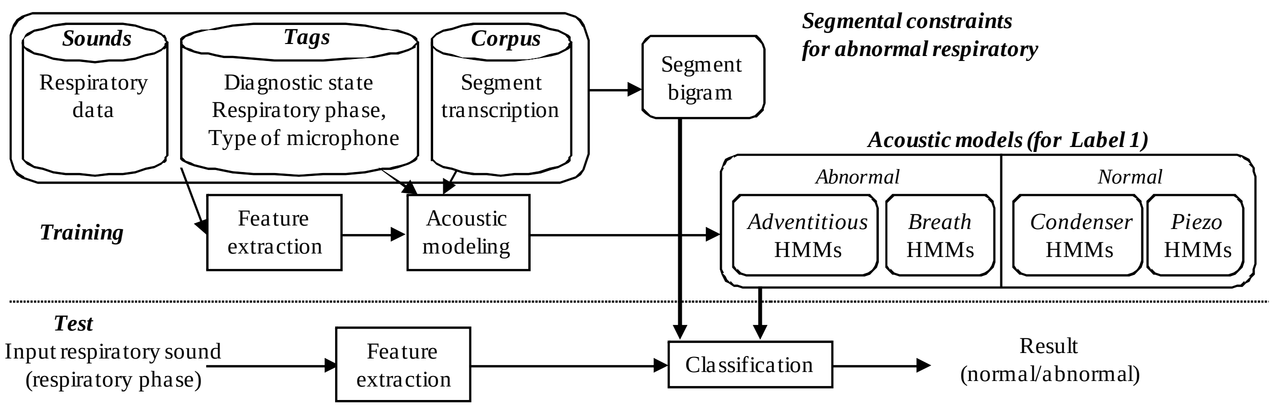


Figure 4. Architecture of classification procedure for distinguishing between normal and abnormal respiratory sounds

that are expressed as BN in Figure 1.

In our experiments, we presupposed that the respiratory phase is known. As such, if the test sample is expiratory, acoustic models generated with expiratory sounds are used for classification. On the other hand, we presupposed that the recording condition for the test sample is unknown. In this case, two kinds of respiratory models (a condenser and piezoelectric microphone) for normal respiration were used simultaneously.

Three types of data sets, A, B, and A+B, were used for our experiments. If a test sample was an element of one data set, samples of the same data set were used for acoustic modeling. HMMs with three states and two Gaussian probability density functions (two-mixture of PDFs) were used in the modeling process.

Ability of segment bigrams

We generated a segment bigram for each type of label. First, we calculated the test set perplexity *PP* for each model to evaluate the ability of the bigram. The perplexity is expressed as

$$PP = 2^H \quad . \tag{5}$$

The term *H* is the entropy of the bigram calculated using the whole test set

$$H = -\frac{1}{\sum_{j=1}^{M} N_j} \sum_{i=1}^{N_j} P\left(w_i \mid w_{i-1}\right), \tag{6}$$

where N_j is the number of segments of the *j*-th test respiratory data and *M* is the total number of abnormal respiratory test samples. The smaller value that is obtained when comparing the number of segments indicates the superiority of the stochastic segment models. Table 2 shows the test set perplexity for each label set and data set. The number of segments for each label set is also shown. This figure is also equal to the perplexity where the occurrence probabilities of all acoustic segments are equal. It is shown that all the figures of perplexity are far smaller than the number of segments in each label set, thus demonstrating the effectiveness of segment bigrams.

Table 2. Test set perplexity of each segment bigram for abnormal respiratory period

Label \ Data	Inspiration	Expiration	No. segments
Label 1	1.78	1.80	5
Label 2	1.85	1.92	6
Label 3	2.17	2.48	11

Experiments using deterministic connection rules of acoustic segments (baseline)

A preliminary classification test was conducted to confirm the performance of a conventional method (baseline) [6] using deterministic connection rules among segments instead of the proposed segment bigram. These rules expressed the sequences of acoustic segments in abnormal respiration and were generated according to three kinds of manual labeling. The Backus–Naur Form (BNF) is adopted to express these rules.

The classification results are listed in Table 2. In Label 1, the adventitious sound models and the breath sound models for abnormal respiration were generated. The Label 2 set comprises continuous sound models, discontinuous sound models, and breath models for the abnormal sound periods. The Label 3 set includes six kinds of specific adventitious sounds, two kinds of unclassifiable models, and breath models as shown in Figure 3. Each classification rate is generally high, indicating the effectiveness of stochastic acoustic modeling using HMMs. The use of detailed labels (Label 3) decreased the classification rates in comparison with the performance using Labels 1 or 2. Thus, we believe that detailed modeling for adventitious sounds using a small amount of data along with deterministic rules results in decrease in the classification performance.

Table 3. Classification performance using connection rules among acoustic segments (baseline) [%]

Label \ Data	Set A	Set B	Set A+B
Label 1	81.5	85.9	83.4
Label 2	82.0	85.3	83.4
Label 3	79.7	84.2	81.6

Experiments using segment bigrams

To evaluate the performance using segment bigrams, we carried out classification experiments for each data set and label set. Experimental results and the typical value of α for which the highest performance is achieved are listed in Table 4. A comparison of Tables 3 and 4 reveals that the performance when using the segment bigram is always superior to that when using the connection rules, thus demonstrating the effectiveness of segment bigrams. With regard to the weight α concerning the likelihood of the segment bigram, higher values are effective when detailed labels (Label 3) are used in comparison with Label 1 or 2. This means that the segment bigram is more useful than the connection rules when more detailed labels are used as acoustic segments. In our experiments, Label 2, which comprises continuous sound segments, discontinuous sound segments, and breath segments for the abnormal-sound period, achieved the highest performance among the three types of segmentation. Our procedure achieved a classification rate of 84.2% and led to a 4.8% reduction of error rate $\left((84.2-83.4)/(100-83.4)\times 100\%\right)$ in comparison with the classification rate of the baseline when Label 2 and Set A+B were used. We believe that an appropriate cluster of acoustic segments, such as continuous sound segments, is needed to achieve higher performance when the amount of training data is insufficient.

Figure 5 shows the classification performance using the Label 2 set and Set B ((a) in Figure %) or A+B (b) where the range of α is from 0 to 10. This figure also illustrates the superiority of the segment bigram and shows that it is important to select a proper value of α to achieve higher performance.

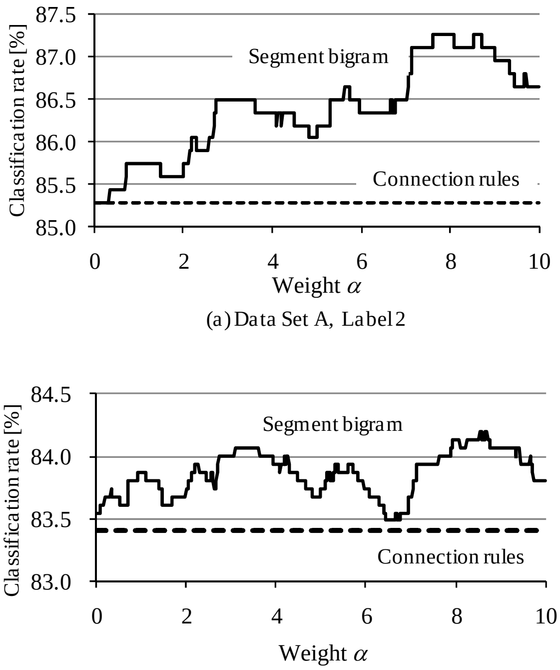
Table 4. Classification performance using acoustic segment bigram for each label set [%]

Label \ Data	Set A	Set B	Set A+B
Label 1	82.3	86.0	83.7
α	1.3	0.15	1.3
Label 2	82.5	87.3	84.2
α	0.3	8.55	8.55
Label 3	82.5	85.9	83.9
α	9.65	9.76	9.76

CONCLUSIONS

This paper proposed a classification procedure for distinguishing between normal and abnormal respiratory sounds; in this procedure, HMMs were used for acoustic spectral features and bigram models were used for the occurrence of acoustic segments in each inspiratory/expiratory period. In our approach, we assumed that each inspiratory/expiratory period consisted of a time sequence of characteristic acoustic segments. The classification procedure detected the segment sequence with the highest total likelihood using HMMs and segment bigrams, and it yielded the diagnostic state of the sequence as a classification result. Experimental results revealed that for the classification, use of the segment bigram led to reduction in the error rate in comparison with the classification rate of a conventional method that uses deterministic rules to describe segment sequences instead of the segment bigram.

In our experiments, the heuristic value, which is a weight factor of the segment bigram, is used; this weight is intentionally set to achieve the best performance. Future work will focus on automatically acquiring an appropriate weight value.



(a) Data Set A, Label2

(b) Data Set A+B, Label2

Figure 5. Classification performance where the range of the weight factor is from 0 to 10

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 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.

6 S. Matsunaga, K. Yamauchi, M. Yamashita and S. Miyahara, "Classification between normal and abnormal respiratory sounds based on maximum likelihood approach," *Proc. of IEEE ICASSP*, pp. 517-520, 2009.

7 S. Young, et al, *The HTK book*, Cambridge University, 2006.