

Classification between Normal and Abnormal Lung Sounds Using Unsupervised Subject-Adaptation

Shunya Umeki, Masaru Yamashita and Shoichi Matsunaga

Nagasaki University, Nagasaki, Japan

E-mail: {b210018, masaru, mat} @cis.nagasaki-u.ac.jp Tel: +81-95-8192700

Abstract— In this paper, we propose an unsupervised subject-adaptation method to distinguish between normal and abnormal lung sounds. Lung sounds have varied subject-dependent acoustic characteristics. In conventional classification methods using subject-independent acoustic models, this diversity hinders the achievement of a high classification rate. To overcome this problem, we performed unsupervised subject-adaptation of acoustic lung-sound models by exclusively employing respiration periods that could be confidently considered normal/abnormal in test respiration input from an unknown subject. In our method, these confident periods were detected based on the difference between the acoustic likelihood for a normal respiratory candidate and that for an abnormal candidate. The proposed adaptation method achieved a higher classification performance of 83.7% between normal and abnormal respiration in comparison with the baseline method that did not use adaptation, which achieved a performance of 82.7%. Our method for classifying healthy subjects and patients with pulmonary emphysema achieved a higher classification rate of 84.3% relative to the baseline (83.5%).

I. INTRODUCTION

Auscultation is one of the most popular and cost-effective medical examination methods for identifying respiratory illnesses. Auscultation is based on the observation of abnormal respiratory sounds that are frequently observed in patients with pulmonary emphysema. These sounds, termed “adventitious sounds,” include wheezes and crackles, and are caused by abnormalities in the lungs and bronchial tubes [1]. Patients can produce several types of adventitious sounds depending on the condition of the abnormal internal organs and the degree of inflammation of the organs. The contamination of lung sounds by noises and the acoustical variety of lung sounds make it very difficult for an undisciplined person to detect adventitious sounds using a stethoscope. Even among healthy subjects, the acoustic characteristics of lung sounds are fairly different according to each subject’s physique, age, gender and so on.

Several studies have been conducted on the acoustic analysis of respiratory sounds for the detection of specific adventitious sounds using such as wavelet functions [2-4]. These studies were performed with the primary aim of assisting doctors in making diagnoses. We developed a classification procedure for distinguishing between normal respiration and abnormal respiration on the basis of a maximum likelihood (ML) approach, using hidden Markov models (HMMs) [5, 6]. In this procedure, the acoustic

likelihood for an abnormal respiration candidate, and that for a normal respiration candidate were compared for the classification. This procedure demonstrated the usefulness of a stochastic approach in the detection of abnormal respiratory sounds in patients. In our former study, we focused on the problem of noise contamination, because many respiratory sounds included noise from the stethoscope or internal organs. We considered the difference between the duration of adventitious sounds and noises [7] to address this problem. In the technological field of speech recognition, stochastic modeling using HMMs has been widely used and stochastic speaker-adaptation techniques have been utilized to increase the recognition performance for unknown speakers [8, 9]. The success that was achieved with speaker adaptation offered a promising way of coping with the variety of acoustical characteristics in lung sounds. We therefore anticipated that the adaptation of lung-sound models to unknown subjects might be useful for the classification between normal and abnormal lung sounds, and for the classification between healthy subjects and patients with pulmonary emphysema. However, compared with speech utterances, there are two major obstacles to perform subject-adaptation effectively: the lower intelligibility of lung sounds and the smaller amount of adaptation data for new subjects, which is usually a short series of respiratory phases within several seconds

To address these obstacles, we propose an unsupervised adaptation method for stochastic lung-sound models using the respiration input of a new subject on the basis of an ML approach. In our method, for the test input, we extracted confident respiration periods as normal/abnormal respiration, based on the difference between the acoustic likelihood for an abnormal respiratory candidate, and that for a normal respiratory candidate. Lung-sound models were adapted using only the confident respiration periods of the input and their most likely segment sequences. We assumed that this data selection approach could prevent the erroneous adaptation

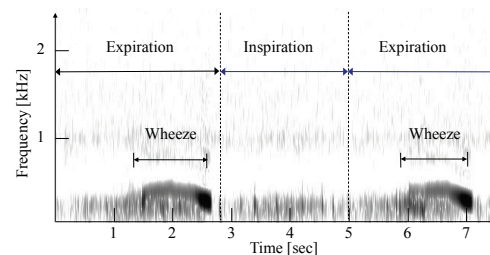


Fig. 1. Spectrogram of a typical lung sound from a patient

caused by factors such as contamination of high-level noise or vague acoustic characteristics, which may occur when inadequate periods are used. The proposed method yields improved classification performance when compared to the previous method that did not perform model adaptation. This result is also compared with the result obtained by the supervised adaptation, which relied on manual segmentation of the input.

II. LUNG SOUND DATA

A. Training and Evaluation Data

An electronic stethoscope that incorporated a piezoelectric microphone was used when recording lung sounds. The second and fourth intercostal spaces (abbreviated as R2 and R4, respectively) on the front right of the subjects were used as the recording points. In the R2 auscultation point, 53 lung sound samples from 53 patients with pulmonary emphysema, and 53 samples from 53 healthy subjects were collected. Sixty-two samples from patients and healthy subjects were also collected in the R4 auscultation point, respectively. Each lung sound sample consisted of successive respiratory phase segments, and the average number of respiratory (inspiratory/expiratory) segments was eight. Each sample from the patients contained at least one phase segment that included adventitious sounds. We tagged the segments according to the respiratory phase (inspiratory or expiratory), diagnostic state (normal or abnormal), auscultation point, and subjects' health (healthy or patient). The subject's health was determined by a doctor on the basis of auscultation and the existence of other medical conditions.

B. Manual Segmentation of Acoustic Segments

A lung sound sample S comprised several successive respiratory phases W :

$$S = W_1 W_2 \cdots W_l \cdots W_L, \quad (l = 1, \dots, L), \quad (1)$$

where W_l is the l -th respiratory phase in which each period was manually detected. We prepared labels corresponding to the acoustic segments based on acoustic and segmental features manually [6]. In our labeling process, we assumed that an abnormal respiratory phase was composed of N successive acoustic segments and that a normal respiratory phase comprised one breath segment ($N=1$). Then, a respiratory period W is represented as

$$W_l = w_1 w_2 \cdots w_i \cdots w_N, \quad (i = 1, \dots, N), \quad (2)$$

where the i -th acoustic segment is given as w_i ($i = 1, \dots, N$). In the present study, each adventitious sound was presented using a continuous acoustic segment (e.g., rhonchus, wheezing) or a discontinuous acoustic segment (e.g., coarse crackle, fine crackle). These segments and their boundaries were used in a training process described in section III. A spectrogram of a typical lung sound from a patient is shown in Figure 1. During each expiratory phase, this lung sound contains an adventitious sound (wheeze).

III. ADAPTATION AND CLASSIFICATION STRATEGY

A. Architecture of Classification System

The architecture of our classification system between normal and abnormal respiration or between patients or healthy subjects is shown in Figure 2. The system comprises two processes: a training process and a test process.

In the training process, original acoustic HMMs θ for each kind of segment were generated for each respiratory phase (inspiration/expiration) using the segment labels and their periods detected manually ((i) in Figure 2). Segment bigrams with reference to the frequency of occurrence sequences of the segments in abnormal respiratory periods were also estimated according to the segment labels [6].

The test process comprised three steps: first, for a test respiratory input, the likelihood for a normal candidate and that for an abnormal candidate were calculated for each respiration period using original HMMs. For the abnormal candidate, the most likely segment sequence and their segment periods were obtained. Second, subject-adaptation was performed using the obtained segment labels of confident respiratory phases ((ii) in Figure 2) and the manual labels of the training data (i) on the basis of the ML approach. Third, the classification was performed using the adapted HMMs.

B. Likelihood Calculation for each Respiratory Phase

We denote the occurrence probability of the segment sequence W_l calculated using the segment bigram;

$$P(W_l) = P(w_1 w_2 \cdots w_i \cdots w_N) \approx \sum_{i=2}^N P(w_i | w_{i-1}). \quad (3)$$

The total likelihood is composed of the acoustic likelihood calculated from HMMs and the segmental sequence likelihood. The segment (sequence) $\hat{W}$ with the highest

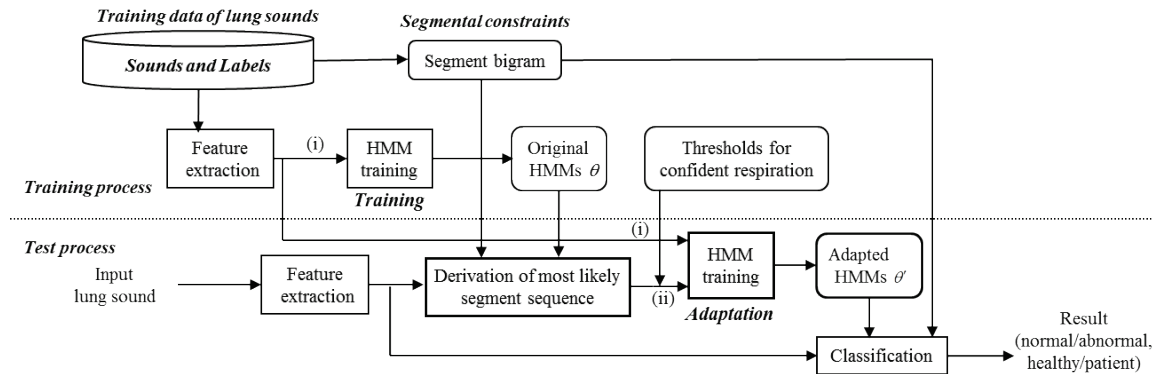


Fig. 2. Classification system between normal and abnormal lung sounds

likelihood $\log P(\hat{W} | X, \theta)$ for a respiratory phase X of a respiratory input is given below using Bayes' theorem:

$$\hat{W} = \arg \max_W \log P(W | X, \theta) \approx \arg \max_W [\alpha \log P(W) + \log P(X | W, \theta)] \quad (4)$$

The weight factor α controls the contribution of the acoustic-segment bigram and was obtained experimentally. Then, the most likely segment sequence $\hat{W}^{ab}$ for an abnormal respiratory candidate and its likelihood $\log P(\hat{W}^{ab} | X, \theta)$, and the likelihood $\log P(\hat{W}^{no} | X, \theta)$ for a normal respiration candidate $\hat{W}^{no}$ were obtained.

C. Subject Adaptation of Lung-Sound HMMs

In our adaptation method, both the obtained segment sequence of the input and all labels of the training data were used to generate adapted HMMs θ' on the basis of the ML approach. However, we did not use all of the respiratory periods in the test input for subject-adaptation. In each test lung-sound, we selected the confident respiration phases W_l considering the likelihood difference d_l :

$$d_l = \log P(\hat{W}_l^{ab} | X_l, \theta) - \log P(\hat{W}_l^{no} | X_l, \theta), \quad (5)$$

where X_l indicate the l -th respiratory phase data in the test input S . If this value d_l was larger than a predefined threshold T^{ab} , this phase X_l was used as confident adaptation data for HMMs of abnormal respiration. If this value was less than a threshold $-T^{no}$, this was used as a confident adaptation data for HMMs of normal respiration. Other phases X_l ($-T^{no} < d_l < T^{ab}$) were not used for adaptation.

D. Criterion to Detect Abnormal Phases and Patients

By using the adapted lung-sound HMMs θ' , the system calculated the likelihood for a normal respiration candidate and the likelihood for an abnormal candidate. Each target phase X_l in the test lung-sound was classified as normal or abnormal by comparing these two likelihoods.

We expanded the aforementioned criteria to classify respirations as being either normal or abnormal, and these classifications were incorporated into the criteria to identify patients by summing the difference of the likelihoods between all normal respiratory candidates and all abnormal respiratory candidates in a test sample (respiratory phase sequence):

$$D = \sum_{l=1}^L \{\log P(\hat{W}_l^{ab} | X_l, \theta') - \log P(\hat{W}_l^{no} | X_l, \theta')\} \quad (6)$$

where the adapted HMMs were used. If the score D was larger than zero, the subject of the input sample S was classified as a patient. Otherwise, the subject was classified as a healthy subject.

IV. EVALUATION EXPERIMENTS

A. Experimental Conditions

We performed classification tests to evaluate the proposed method. The total number of test samples was 230, recorded from auscultation points R2 and R4. The lung sound data were sampled at 10 kHz. Every 10 ms, a vector of 5 mel-frequency cepstral coefficients (MFCC) and power was computed using a 25-ms Hamming window. The acoustic

models (HMMs) for normal respiration were generated using the respiratory sounds from healthy subjects, and the models for adventitious sound segments were generated using the sounds from patients. HMMs with three states and two Gaussian probability density functions were used. In our experiments, we assumed that both the respiratory phase (inspiratory or expiratory) and phase boundaries were known. Accordingly, if the target phase of the test sample was expiratory, acoustic models generated by expiratory sounds were used to calculate the acoustic likelihood of the phase. We performed a leave-one-out cross validation for each test phase or sample. The training data were other 105 samples at R2 and 123 samples at R4. In addition, the HMMs were subject-independent because the samples recorded from the same subject used as the test sample were excluded in the training process.

B. Classification of Normal and Abnormal Respiration

First, to confirm the classification performance of the proposed unsupervised adaptation, we performed experiments using three classification methods to distinguish between normal and abnormal respirations: a baseline method [6], a method using supervised models, and the proposed method using unsupervised models. In these experiments, we used 784 respiratory phases at R2 and 941 phases at R4 where each abnormal respiratory phase contained at least one adventitious sound. In the baseline method, the classification experiment was performed by using the original HMMs which was not adapted to a new subject. The supervised models were generated by using the segment labels and their periods, which were manually given, of the input sample in addition to the training dataset. Classification performance obtained by supervised adaptation is usually better than that obtained by unsupervised adaptation; however, supervised subject adaptation is not a realistic method for developing real-time processing for identifying respiratory illness. On the other hand, it is important to detect the performance of the supervised adaptation, because the classification performance by using the supervised adaptation establishes the upper bound of the performance obtained by unsupervised adaptation. Concerning the unsupervised training, we performed two types of experiment: one experiment using HMMs adapted with all respiration periods of the test input, and the other experiment using HMMs adapted with only confident periods, which is our proposed method.

The experimental results are shown in Table I. The average classification rate of 83.7% of the proposed adaptation method ($T^{ab} = 400, T^{no} = 100$) was higher than that of the baseline of 82.7%, thereby validating the unsupervised subject-adaptation. On the other hand, the classification rate (81.2%) of the unsupervised adaptation using all phases ($T^{ab} = T^{no} = 0$) was lower than that of the baseline. This result indicated that the selection of confident periods was necessary for the success of the unsupervised adaptation. Figure 3 shows the classification performances of the unsupervised adaptation for several combinations of the values of two thresholds. This figure indicates that the threshold T^{ab} for confident abnormal

TABLE I
CLASSIFICATION BETWEEN NORMAL AND ABNORMAL LUNG SOUNDS [%]

Method	Adaptation periods	Normal	Abnormal	Average
Baseline	—	79.4	89.4	82.7
Supervised ad.	All	81.7	90.9	84.7
Unsupervised adaptation	All * ¹	76.0	92.1	81.2
	Confident * ²	80.9	89.4	83.7

*1 $T^{ab} = T^{no} = 0$

*2 $T^{ab} = 400$ and $T^{no} = 100$

TABLE II
CLASSIFICATION BETWEEN HEALTHY AND DISEASED SUBJECTS [%]

Method	Adaptation periods	Healthy	Diseased	Average
Baseline	—	74.8	92.2	83.5
Supervised ad.	All	79.1	93.0	86.1
Unsupervised adaptation	All * ³	60.9	94.8	81.2
	Confident * ⁴	78.3	90.4	84.3

*3 $T^{ab} = T^{no} = 0$

*4 $T^{ab} = 400$ and $T^{no} = 100$

respirations is more important than that T^{no} for confident normal respirations. Comparing the classification rate of 84.7% of the supervised adaptation and that (83.7%) of the proposed method in Table I, it is shown that our unsupervised adaptation method achieved reasonable results, indicating the possibility of improvement. However, the actual increase in comparison with the baseline was not statistically significant (p -value: 0.24 for the proposed method and 0.054 for the supervised adaptation). We needed to prepare a large amount of lung sound data to statistically show the superiority of the proposed method. Furthermore, setting the appropriate values of the thresholds was an important problem in the proposed method. When the thresholds values were higher, the amount of available data for the adaptation was smaller. Thus, input lung sounds with more respiratory phases may yield more effective adaptation. We used the ML approach as the first trial of the subject-adaptation. Other adaptation techniques, such as maximum a posteriori (MAP) [8] or maximum likelihood linear regression (MLLR) [9], might be good alternatives.

C. Classification of Patients and Healthy Subjects

We performed experiments using the aforementioned three classification methods between patients and normal subjects. The obtained classification rates are shown in Table II. The unsupervised adaptation method using the confident phases outperformed the baseline method slightly (0.8%), showing the validity of the proposed method. However, the proposed method using confident periods decreased the detection performance of patients. This is because some patients, whose occurrences of adventitious sounds are rare, tended to be misrecognized as normal subjects.

V. CONCLUSIONS

This paper presents the results of an investigation into applying an unsupervised adaptation technique for the detection of abnormal respirations of patients. We propose an

unsupervised adaptation method using only the confident respiration periods from a test lung-sound. In our proposed method, we obtained the confident periods on the basis of the difference between the likelihood of a normal respiration candidate and that of an abnormal respiration candidate. According to the experimental results obtained for the classification between normal and abnormal respirations and those for the classification between patients and subjects, the performance of the proposed method was better than that of the baseline, which did not adopt the adaptation, and showed that it was important to prove the existence of confident abnormal respiration periods. However, we need to prepare a large amount of lung sound data to statistically show the superiority of the proposed method.

REFERENCES

- [1] Noam Gavriely, "Breath sounds methodology," CRC Press, 1995.
- [2] Y.P. Kahya, S. Yere and O. Cerid, "A wavelet-based instrument for detection of crackles in pulmonary sounds," *Proc. of IEEE EMBS*, pp.3175-3178, 2001.
- [3] S.A. Taplidou and L.J. Hadjileontiadis, "Wheeze detection based on time-frequency analysis of breath sounds," *Computers in Biology and Medicine*, 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] S. Matsunaga, et al., "Classification between normal and abnormal respiratory sounds based on maximum likelihood approach," *Proc. of IEEE ICASSP*, pp. 517-520, 2009.
- [6] H. Yamamoto, et al., "Classification between normal and abnormal respiratory sounds based on stochastic approach," *Proc. of ICA*, 2010.
- [7] M. Himeshima, et al., "Detection of abnormal lung sounds taking into account duration distribution for adventitious sounds," *Proc. of EUSIPCO*, pp. 1821-1825, 2012.
- [8] J.L. Gauvain and C.H. Lee, "Bayesian learning of Gaussian mixture densities for hidden Markov models," *Proc. of DARPA Speech and Natural Language Workshop*, pp.272-277, 1991
- [9] C.J. Leggetter and P.C. Woodland, "maximum likelihood linear regression for speaker adaptation of continuous density hidden Markov models," *Computer Speech and Language*, 9, pp.171-185, 1995

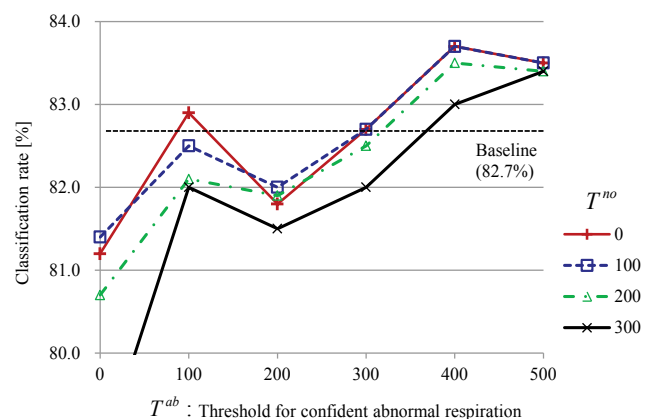


Fig. 3 Classification rate of unsupervised adaptation method for each pair of thresholds T^{ab} and T^{no} .