DISCRIMINATION BETWEEN HEALTHY SUBJECTS AND PATIENTS USING LUNG SOUNDS FROM MULTIPLE AUSCULTATION POINTS

Shohei Matsutake, Masaru Yamashita, and *Shoichi Matsunaga*

Department of Computer and Information Science, Nagasaki University, JAPAN

ABSTRACT

In this paper, we propose a robust classification method to distinguish between a healthy subject and a patient with pulmonary emphysema using lung sound samples recorded from multiple auscultation points. Although the symptom of pulmonary emphysema can be determined from lung sounds that frequently include abnormal (i.e., adventitious) sounds, these are not observed in every auscultation point. Furthermore, noise pollution during auscultation makes high-accuracy detection difficult. To overcome these difficulties, our proposed method took into account lung sound samples from multiple auscultation points in diagnosing a patient. After the calculation of the acoustic likelihood for each respiratory phase based on the maximum likelihood approach using hidden Markov models and a segmental bigram, patient diagnosis was carried out based on the comparison of the average likelihood of all auscultation points between a patient and a healthy subject. Our classification method significantly increased the classification performance to 90.5% (using samples from four auscultation points) from the 82.7% classification performance of the conventional method (using a sample from one auscultation point), validating the usefulness of our proposed method.

Index Terms— lung sound, patient classification, adventitious sound, auscultation point

1. INTRODUCTION

Diagnosis of pulmonary emphysema using a stethoscope is beneficial because auscultation of lung sounds is one of the most popular and cost-effective medical examination methods in identifying respiratory illnesses. Auscultation is based on the common knowledge that abnormal respiratory sounds are frequently observed in patients with pulmonary emphysema. Typical sounds such as wheezes are caused by abnormalities in the lungs and bronchial tubes; they are termed as *adventitious sounds* [1]. Several types of adventitious sounds exist, depending on the condition of the abnormal internal organs and the degree of inflammation. Physicians usually perform auscultation at several points of the patients' chest and back to make an accurate diagnosis. If the auscultation point is far from the abnormal parts, detecting adventitious sounds could be very difficult.

Several acoustic analyses of respiratory sounds for the detection of specific adventitious sounds have been conducted [2-5]. These analyses were performed to assist doctors in making diagnoses. The objective of our study is to develop a home-use device to identify respiratory illness by detecting abnormal respiratory sounds. We developed a classification procedure to distinguish between healthy subjects and patients by the detection of the adventitious sounds based on the maximum likelihood approach using hidden Markov models (HMMs) [6-8]. This procedure demonstrates the usefulness of the stochastic approach in the detection of abnormal respiratory sounds. However, we inferred two main problems that hinder achieving a relatively high level of classification. One problem is noise pollution in the lung sounds during auscultation because many respiration sounds include some noise from the stethoscope or the internal organs, and the spectral features of several noises are very much similar to those of some types of abnormal respiratory sounds. In our previous work [9], we proposed a classification method using not only the spectral features but also the duration of noise sounds and adventitious sounds. The other problem is the number of auscultation points. We used to have lung sound samples only from one auscultation point. However, patient diagnosis using samples only from one auscultation point is insufficient. Furthermore, a useful detection method using samples from multiple auscultation points has not yet been developed.

To address these problems, we propose a robust classification method between healthy subjects and patients with pulmonary emphysema using lung sounds recorded from multiple auscultation points. In this method, after the calculation of the acoustic likelihood of each respiratory phase for normal and abnormal respirations based on the maximum likelihood approach, the average total likelihood of the multiple auscultation points was robustly used to diagnose a patient. The validity of the proposed method is confirmed by a classification experiment using lung sound samples recorded at four auscultation points.

2. LUNG SOUND DATA

2.1. Training and evaluation data

We recorded lung sounds at four auscultation points in patients with pulmonary emphysema and in healthy subjects using an electronic stethoscope that incorporates a piezoelectric microphone. The auscultation points are shown in Figure 1. One lung sound

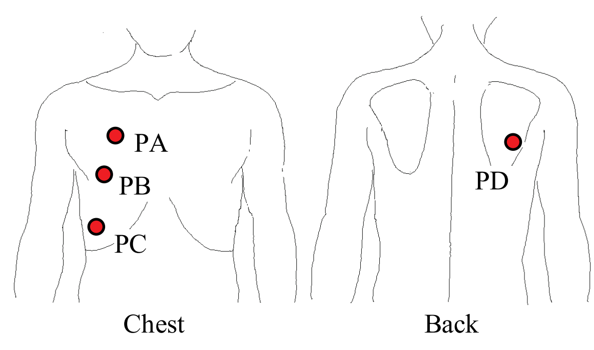

Figure 1. Multiple auscultation points in the chest and back

Table 1. Number of respiratory phase segments

Auscultation point	PA	PB	PC	PD
Patients	719	745	751	759
Healthy subjects	630	633	630	629

Table 2. Number of lung sound samples including adventitious sounds out of the 74 samples

Auscultation point	PA	PB	PC	PD
No. of samples	42	59	59	58

Table 3. Number of patients corresponding to the number of auscultation points where adventitious sounds were observed

No. of points	1	2	3	4
No. of patients	74	68	50	26

sample for each auscultation point in each subject was recorded. As a result, for each auscultation point, 74 samples from 74 patients and 63 samples from 63 healthy subjects were prepared. In our lung sound data, we prepared the patients whose four lung sound samples indicated at least one sample containing at least one adventitious sound. Each sample consisted of successive respiratory phase segments (inspiratory and expiratory), and the average number of respiratory segments was approximately 10. The number of respiratory segments for each auscultation point is shown in Table 1. The number of samples that contained adventitious sounds for each auscultation point out of the 74 patient samples is shown in Table 2. Many patient samples did not indicate adventitious sounds; in particular, in auscultation point PA (first right intercostal space), the ratio of samples containing adventitious sounds was the lowest (57%) among the four points. Table 3 shows the number of patients corresponding to the number of auscultation points where adventitious sounds were observed. The ratio of the number of patients where adventitious sounds were detected in all four auscultation points was only 35% (26/74). These investigations indicated that lung sound auscultation at multiple points is necessary to identify an unhealthy subject.

We tagged the segments according to the respiratory phase (inspiratory or expiratory), diagnostic state (normal or abnormal), auscultation point, and the subject's health states (healthy or patient). The subject's health state was identified by a doctor based on auscultation as well as on many other medical conditions. The data were divided into four groups according to the diagnostic state and the subject's health state as follows:

· Abnormal respiration from patients (AP): respiration that contained obvious adventitious sounds

· Normal respiration from healthy subjects (NH): respiration with no adventitious sounds

· Abnormal respiration from healthy subjects (AH): respiration that contained noises from internal organs

· Normal respiration from patients (NP): respiration with no obvious adventitious sounds.

In this study, the respiration data related to AP and NH were used to validate the ability to classify the abnormal respiratory phase of patients and the normal respiratory phase of healthy subjects (Section 4.2), and all respiration data were used to classify the patients and healthy subjects (Section 4.3).

2.2. Manual labeling of acoustic segments

We prepared labels corresponding to the acoustic segments based on the acoustic and segmental features. In our labeling process, we assumed that an abnormal respiratory period (phase) was composed of successive acoustic segments. To model the adventitious sounds of patients, we defined the segments according to their acoustic features and assigned a symbol w to each segment period [7].

We supposed that a respiratory phase W comprises N segments: we let the i-th acoustic segment as w_i $(1 \le i \le N)$. Then

$$W = w_1 w_2 \cdots w_i \cdots w_N . \qquad (1)$$

In our data, one abnormal respiratory period comprised several segments, and one normal respiratory period comprised one breath segment ($N = 1$). In this study, each adventitious sound was presented using a continuous or discontinuous sound segment; the segment sequence of an abnormal respiratory period thus consisted of one of the two types of segments and the respiratory-sound segments without adventitious sounds. Some typical examples of continuous sound segments are coarse crackle, fine crackle, and pleural friction rub. Rhonchus or wheezing sounds are examples of discontinuous segments. We included silent periods during the breathing segments.

In the recording process using the stethoscope, mixing of noises was inevitable; consequently, approximately 80% of all respiration phases included some noises from the stethoscope or from the internal organs.

3. CLASSIFICATION METHODS

3.1. Acoustic likelihood calculation

Our strategy in calculating the acoustic likelihood for a normal/abnormal respiratory phase was based on the maximum likelihood approach. We let the occurrence probability of the segment sequence $W_{j,k,l}$ of the l-th respiratory phase in the sample from the k-th subject's j-th auscultation point as $P(W_{j,k,l})$. We used a segmental bigram to calculate $P(W_{j,k,l})$ [7], i.e.,

$$P(W_{j,k,l}) = P(w_1 w_2 \cdots w_i \cdots w_N) \approx \prod_{i=2}^{N} P(w_{i-1} \mid w_i) , \qquad (2)$$

where w_i is the i-th acoustic segment of $W_{j,k,l}$, as described in Section 2.2. The total likelihood is composed of the acoustic likelihood calculated from HMMs and the segmental sequence likelihood calculated from the bigram. The segment (sequence) $\hat{W}_{j,k,l}$ with the highest likelihood $\log P(\hat{W}_{j,k,l} \mid X_{j,k,l})$ for a unknown respiratory input $X_{j,k,l}$ is given below using the Bayes' theorem:

$$\begin{aligned}\hat{W}_{j,k,l} &= \underset{W_{j,k,l}}{\arg\max}\, P(W_{j,k,l} \mid X) \\ &\approx \underset{W_{j,k,l}}{\arg\max}\left[\alpha_j \log P(W_{j,k,l}) + \log P(X \mid W_{j,k,l})\right]\end{aligned} \qquad (3)$$

where $\log P(X \mid W)$ is the acoustic likelihood and $X_{j,k,l}$ is abbreviated as X. The weight factor α_j for each auscultation point controlled the contribution of the bigram, and this factor was obtained experimentally.

3.2. Criteria of patient detection

The classification for the healthy subjects and the patients was conducted using four lung sound samples from different auscultation points. The likelihood $\log P\left(\hat{W}^{No} \mid X\right)$ for normal and the likelihood $\log P\left(\hat{W}^{Ab} \mid X\right)$ for abnormal respirations were used in the classification. We employed four criteria in identifying a patient.

(C1) Detection of one or more abnormal respirations, i.e.,

for each subject k, $\exists j \exists l \left[\log P\left(\hat{W}^{Ab}_{j,k,l} \mid X\right) > \log P\left(\hat{W}^{No}_{j,k,l} \mid X\right)\right]$.

If at least one abnormal respiration period was detected among the four lung sound data samples from each subject k, the subject was regarded as a patient. If adventitious sounds could be ideally detected, this criterion would be sufficient to identify a patient.

(C2) Detection of one or more confident abnormal respirations for each subject: we proposed the idea of "confident abnormal respiration" in our previous paper [8]. The aim of this idea is to reduce the detection error of adventitious sounds caused by noises during auscultation. This idea is described as follows: if the difference between the likelihood for the normal respiration phase and the likelihood for the abnormal respiration phase is larger than a threshold *Th* for the respiratory input, we regard this test respiration phase to be abnormal with confidence. This threshold was determined experimentally for each auscultation point j. Then, if one or more confident abnormal respiration phases were detected in the four lung sound samples, we classified the subject as a patient. This logic is formulated as follows:

for each subject k, $\exists j \exists l \left[\log P\left(\hat{W}^{Ab}_{j,k,l} \mid X\right) - \log P\left(\hat{W}^{No}_{j,k,l} \mid X\right) > Th_j\right]$.

(C3) Detection of one or more abnormal samples for each subject: the occurrence frequency and clarity of adventitious sounds were different among data samples according to the auscultation points. This result was mainly caused by the difference in the distance from the auscultation point to the abnormalities in the lungs or bronchial tubes. Then, the aim of this criterion was to detect at least one abnormal sample, apparently recorded closest to the abnormal part, among all samples. In this criterion the total likelihood of all respiratory phases in a sample was used. If the total likelihood of the abnormal respiration was larger than that of the normal respiration, the subject was regarded as a patient, i.e.,

for each subject k, $\exists j \left[\sum_{l} \log P\left(\hat{W}^{Ab}_{j,k,l} \mid X\right) > \sum_{l} \log P\left(\hat{W}^{No}_{j,k,l} \mid X\right)\right]$.

(C4) Comparing the two average likelihood results of all samples recorded from different auscultation points for each subject (for normal and for abnormal respirations): in this criterion, the total likelihood was averaged by the number of respiratory phases $L_{j,k}$. If the average likelihood for abnormal respiration is larger than that for normal respiration, the subject k is regarded as a patient;

for each subject k,

$$\sum_{j} \frac{1}{L_{j,k}} \sum_{l} \log P\left(\hat{W}^{Ab}_{j,k,l} \mid X\right) > \sum_{j} \frac{1}{L_{j,k}} \sum_{l} \log P\left(\hat{W}^{No}_{j,k,l} \mid X\right).$$

As described above, in C1 and C2, the decision result of each respiration was used. On the other hand, the classification results for each auscultation point were used in C3, and in C4 the latest decision for a patient was conducted using all likelihood of the four samples (Figure 2). C4 is our newly proposed criterion.

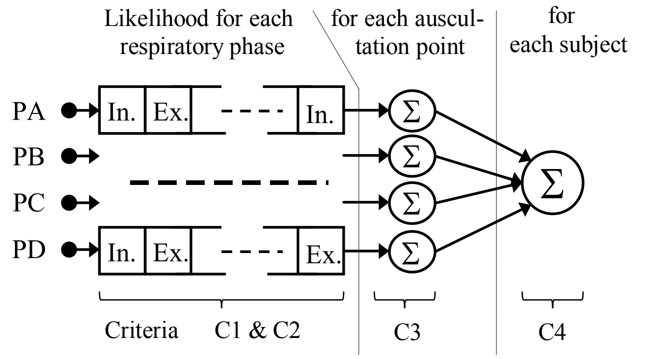

Figure 2. Likelihood calculation process and patient-detection criteria

4. EVALUATION EXPERIMENTS

4.1. Experimental conditions

We performed classification tests to evaluate the proposed method. The lung sound data were sampled at 5 kHz. For every 10 ms, a vector of 5 mel-warped cepstral coefficients and power was computed using a 25-ms Hamming window. This vector was used as an acoustic feature in modeling of the HMMs. All HMMs were generated for each auscultation point using only the lung sound samples recorded at the auscultation point (auscultation-point dependent.) The respiratory sounds from healthy subjects (NH in Section 2.1) were used for the generation of the models for normal respiration. These models were used to calculate the acoustic likelihood $\log P\left(X \mid W^{No}_{j,k,l}\right)$ for normal-respiration candidate. The models for abnormal respiration were also generated using the sounds obtained from the patients (AP). HMMs with three states and two Gaussian probability density functions were used for both models. A segment bigram for each auscultation point was also trained using the segment labels of the training samples recorded at the same auscultation point.

In our experiments, we assumed that the number of respiratory phases for each lung sound sample, the respiratory phase, and the respiratory boundaries are known. Thus, if the test sample was expiratory, acoustic models generated by the expiratory sounds were used. We performed a leave-one-out cross validation. In addition, the samples recorded from the same subject used as the test sample were excluded in the training process so that our experiments would be subject-independent.

4.2. Classification of normal and abnormal respirations

To confirm the detection performance of adventitious sounds in each auscultation point, a preliminary classification experiment to distinguish between the abnormal respiration of patients and the normal respiration of healthy subjects was performed. The evaluation samples were all abnormal respirations from patients (AP) and an almost equal number of normal respirations from healthy subjects (NH) that were randomly selected. The quantities of these samples are shown in Table 4, which also shows that the number of samples from auscultation point PA was relatively small.

Table 4. Number of respiratory sound samples at each auscultation point

Auscultation point	PA	PB	PC	PD
Healthy, normal (NH)	188	300	326	332
Patient, abnormal(AP)	188	311	320	331

Table 5. Classification performance between normal and abnormal respirations for each auscultation point[%]

Auscultation point	PA	PB	PC	PD
Healthy, normal (NH)	96.3	81.0	72.4	81.0
Patient, abnormal(AP)	84.6	85.2	90.6	82.8
Average	90.4	83.1	81.4	81.9

The obtained classification results are shown in Table 5. The average classification rate weighted with the data amount was indicated as "Average." Table 5 shows that each average classification rate is over 80%, and the performance at auscultation point PA of 90.4% was the highest rate among the four points.

4.3. Classification of healthy subjects and patients

4.3.1. Classification using single sample from an auscultation point (baseline)

We performed the classification experiment to distinguish between healthy subjects and patients using a single sample from each auscultation point. Four samples from 74 patients and 63 healthy subjects were evaluated. In this experiment, four criteria (C1, C2, C3, and C4) to identify a patient were used. When the number of auscultation point was one, C3 and C4 were the same.

The upper part of Table 6 shows the average classification performance per auscultation point using each classification criterion. C2 (based on the detection of the confident adventitious sounds) achieved the highest performance (83.2%). Although in this study, samples from four auscultation points were used, this result showed the same trend as the experimental result described in our previous paper [8] using the samples from point PB (second right intercostal space) only. The misrecognition of noises as adventitious sounds made the recall rate of healthy subjects using C1 relatively low (45.6%).

4.3.2. Classification using samples from multiple auscultation points

Finally, the classification experiments to distinguish between healthy subjects and patients using lung sounds samples recorded from multiple auscultation points were carried out. In these experiments, we used samples from two to four auscultation points. For the experiment that used two auscultation points, six combinations of two auscultation points were examined, and the average classification rate of these six combinations was calculated. For the experiment that used three auscultation points, four combinations of three auscultation points were used.

The lower part of Table 6 shows the classification performance for each number of auscultation points using each classification criterion. The proposed classification method using the samples from four auscultation points based on the patient-detection criterion C4 achieved the highest performance of 90.5%. Furthermore, the classification performance based on C4 increased monotonically from 82.7% to 90.5% with the increase in the number of auscultation points from one to four. These results showed the usefulness of the combination of the samples from multiple auscultation points and the use of C4, which was based on the summation of the likelihood in multiple auscultation points. On the other hand, the classification performance using C1 decreased significantly with the increase in auscultation points because of the noise pollution in each sample.

To summarize the above classification results, when samples from multiple auscultation points were used for the classification, C4 (based on the latest patient decision using average likelihood of multiple samples) achieved better performance than C3 of the sample–based decision and C1 and C2 of the respiration-phase-based decision in each sample.

Table 6. Classification performance between patients and healthy subjects [%]

No. of auscultation points	Criteria	Patients	Healthy subjects	Average
1 (Baseline)	C1	97.0	45.6	74.3
	C2 [8]	86.8	79.0	83.2
	C3, C4	81.4	84.1	82.7
2	C1	99.5	20.9	63.4
	C2	97.1	66.1	82.8
	C3	93.5	73.8	84.4
	C4	85.6	86.2	85.9
3	C1	100	9.1	58.2
	C2	99.7	62.3	82.5
	C3	97.6	66.3	83.2
	C4	89.2	87.3	88.3
4	C1	100	3	55.5
	C2	100	51	77.4
	C3	99	60	82.5
	C4	91	90	90.5

5. CONCLUSIONS

This paper has proposed a new method of discriminating between healthy subjects and patients with pulmonary emphysema using lung sound samples recorded from multiple auscultation points. In this method, the likelihood of abnormal respiration containing adventitious sounds, frequently observed in the patients' lung sounds, and the likelihood for normal respiration were compared to detect the patients. The likelihood was calculated based on the maximum likelihood approach using HMMs and a segmental bigram [6,7]. The key characteristics of the proposed method in this work are as follows: it takes into account samples from multiple auscultation points, and it uses the patient-detection criterion based on the comparison of the average likelihood for all auscultation points. From the classification experiments, the proposed classification method increased the performance monotonically in response to the increase in auscultation points, showing the effectiveness of the proposed method.

We did not deal with the points close the heart to avoid heart-sound contamination, and we did not use samples from more than five auscultation points. These would be the subjects in our future work.

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

[4] 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.

[5] 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.

[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] H. Yamamoto, S. Matsunaga, M. Yamashita, K. Yamauchi, and S. Miyahara, “Classification between normal and abnormal respiratory sounds based on stochastic approach,” *Proc. of International Congress on Acoustics*, 2010.

[8] M. Yamashita, S. Matsunaga, and S. Miyahara, “Discrimination between healthy subjects and patients with pulmonary emphysema by detection of abnormal respiration,” *Proc. of IEEE ICASSP*, pp. 693-696, 2011.

[9] M. Himeshima, M.Yamashita, S. Matsunaga, and S. Miyahara, “Detection of abnormal lung sounds taking into account duration distribution for adventitious sounds,” *Proc. of EUSIPCO*, pp. 1821-1825, 2012.