Classification of Healthy Subjects and Patients with Pulmonary Emphysema Using Continuous Respiratory Sounds

Takanori Okubo, Naoki Nakamura, Masaru Yamashita, and Shoichi Matsunaga, *IEEE Member*

***Abstract*— In this paper, we propose a new method for classifying patients with pulmonary emphysema and healthy subjects using lung sounds. Using conventional classification methods, every boundary between inspiratory and expiratory phases in successive respiratory sounds are detected manually prior to automatic classification. However, manual segmentation must be performed accurately and has therefore created significant obstacles in achieving automatic classification. In our proposed method, adequate boundaries are detected automatically in the classification process, based on the criterion of maximizing the difference between the acoustic likelihoods for a candidate with abnormal respiration and one with normal respiration. The proposed method achieved a classification rate of 83.9% between healthy subjects and patients. The reported rate was 1.3% greater than the rate achieved using the conventional method, which required manual phase-wise segmentation. Furthermore, the resulting rate was 2.2% higher than the rate obtained by the classification in which a lung sound sample was divided into phases of equal duration, indicating the effectiveness of the proposed method.**

I. Introduction

A diagnosis of pulmonary emphysema using a stethoscope is beneficial, because the auscultation of lung sounds is one of the most popular and cost-effective medical examination methods for identifying respiratory illnesses. Auscultation is based on the heuristic that abnormal respiratory sounds usually occur in patients with pulmonary emphysema. The abnormal sounds, such as wheezes, are caused by abnormalities in the lungs and bronchial tubes, and are termed *adventitious sounds*.

Several studies have been conducted on the acoustic analysis of respiratory sounds for the detection of specific adventitious sounds using such as wavelet functions [1-5]. These studies were performed with the primary aim of assisting doctors in making diagnoses.

The objective of our study was to develop a home-use device for identifying respiratory illness by differentiating abnormal respiratory sounds from normal lung sounds. We developed a classification procedure for distinguishing between a patient and a healthy subject on the basis of a maximum likelihood approach, using hidden Markov models (HMMs) [6-8]. This procedure demonstrated the usefulness of a stochastic approach in the detection of abnormal respiratory sounds in patients. We collected lung sounds by indicating a sign for the beginning of each inspiration and expiration phase to the subjects. However, the duration of the respiration phases was somewhat different in each subject, which was similar to the distribution of respiratory phase-duration that usually occurs in auscultation. The classification procedure required respiratory phase-by-phase segmentation for the test lung sounds in advance of the classification, which was based on the calculation of the likelihood of normal and abnormal respiratory phases, and required correct performance in order to accurately calculate the exact likelihood. The delicate segmentation was difficult, and it was desired to eliminate the need for the manual segmentation step. Furthermore, noise contamination hindered the achievement of a relatively high level of automatic segmentation, because many respiration sounds included noise from the stethoscope or internal organs.

To address this problem, we propose a new classification method that uses continuous respiratory phases without performing manual segmentation between the phases. In our previous studies [8, 9], we showed that using the difference of total likelihood between candidates with normal respiration and those with abnormal respiration in a sample of lung sounds was one of the most effective classification criteria for patient detection. Based on this finding, we assumed that it would be effective to detect phase-boundaries, based on the criterion of maximizing a difference of likelihood, between the acoustic likelihood for an abnormal respiration candidate, and that for a normal respiration candidate. The proposed method does not guarantee the detection of the most adequate set of boundaries to derive the maximum difference of likelihood; however, we demonstrate experimentally that it yields improved classification performance when compared to the previous method that relied on manual segmentation.

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 PB and PC, respectively) on the front right of the subjects were used as the recording points.

In the PB 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 PC 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

The authors are with the Graduate School of Engineering, Nagasaki University, Nagasaki, 852-8521 Japan (phone: +81-95-8192700; fax: +81-95-8192575; e-mail: mat@cis.nagasaki-u.ac.jp).

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

at least one phase segment that included adventitious sounds. The segments were tagged according to the respiratory phase (inspiratory or expiratory), diagnostic state (normal or abnormal), and the subject's health state (healthy or diseased). The subject's health state was determined by a doctor, and was based on auscultation as well as several other medical conditions.

B. Manual Segmentation for Respiratory Phase Boundaries

Each lung sound sample *S* consisted of several successive respiratory phases *W*:

$$S = W_1 W_2 \cdots W_i \cdots W_N \ , \tag{1}$$

where W_i was the *i*-th respiratory phase in which the beginning and ending times were manually detected. These phase boundaries were used in a training process described in section III. A waveform, respiration phases *W*, and a spectrogram of a typical lung sound from a patient are shown in Figure 1. In our previous studies, the begging time t_0 of the first respiratory phase W_1, the ending time t_N of the last phase W_N, and the phase boundaries in a test sample were used for classification [8, 9]. On the other hand, our new method required only time t_0 and t_N in a test sample.

Lung sounds were collected by indicating a sign for the beginning of each respiration phase to the subjects on a computer monitor. However, the duration of each respiration phase collected was slightly different (Table I).

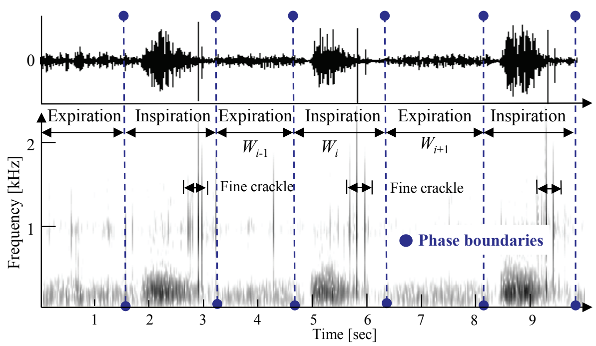

Figure 1. Phase boundaries in a typical lung sound from a patient.

TABLE I. MEAN VALUE AND STANDARD DEVIATION OF DURATION FOR EACH RESPIRATION PHASE

Respiration	Mean [sec]	S.D. [sec]
Inspiration	1.55	0.28
Expiration	1.54	0.16

C. Manual Segmentation for Acoustic Segments

We assumed that an abnormal respiratory phase *W* was composed of successive acoustic segments *w*. To model adventitious sounds stochastically, we defined the segments according to their acoustic features, and assigned a symbol to each segment. Assuming the *i*-th respiratory phase W_i comprised *M* segments: let the *j*-th segment be $w_{j,i}$ $(1 \le j \le M)$. Then,

$$W_i = w_{1,i} \cdots w_{j,i} \cdots w_{M,i} \tag{2}$$

In our data, one normal respiratory period comprised one breath segment ($M = 1$). In this study, each adventitious sound was presented using a continuous or discontinuous sound segment. Thus, the segment sequence of an abnormal respiratory period consisted of one of the two types of segments, as well as respiratory-sound segments that did not include adventitious sounds. Some typical examples of discontinuous sound segments were coarse crackle, fine crackle, and pleural friction rub. Rhonchus or wheezing sounds were examples of continuous segments [1].

III. Classification Strategy

A. Likelihood for Normal/Abnormal respiration

Our strategy for calculating the acoustic likelihood of a normal or abnormal respiratory phase was based on the maximum likelihood approach. We let the occurrence probability of the segment sequence W_i of the *i*-th respiratory period as $P(W_i)$:

$$P(W_i) = P(w_{1,i} \cdots w_{j,i} \cdots w_{M,i}) \ . \tag{3}$$

We used a segmental bigram to calculate $P(W_i)$, which was proposed in our previous work [7]. The total likelihood included the acoustic likelihood, calculated from HMMs, and the segmental sequence likelihood, calculated from the bigram. The segment (sequence) $\hat{W}_i$ with the highest likelihood $\log P(\hat{W}_i \mid X_i)$ for *i*-th respiratory X_i of an input lung sound is given below using Bayes' theorem:

$$\begin{aligned} \hat{W}_i &= \arg\max_{W_i} \log P(W_i \mid X_i) \\ &= \arg\max_{W_i} \left(\alpha \log P(W_i) + \log P(X_i \mid W_i)\right) \end{aligned} \tag{4}$$

where $\log P(X_i \mid W_i)$ was the acoustic likelihood. The weight factor α , which was set experimentally to achieve the highest rate, controlled the contribution of the bigram.

B. Criterion of Patient Detection

We calculated the likelihood $\log P(\hat{W}_i^{No} \mid X_i)$ for normal respirations, and the likelihood $\log P(\hat{W}_i^{Ab} \mid X)$ for abnormal respirations using (4). Further, the two types of likelihoods were accumulated over the entire test sample, in order to distinguish between a pulmonary emphysema patient and a healthy subject. If the total likelihood for abnormal respiration was larger than that for normal respiration, the subject of the test sample was regarded as a patient:

$$\sum_i \log P(\hat{W}_i^{Ab} \mid X_i) > \sum_i \log P(\hat{W}_i^{No} \mid X_i) \ . \tag{5}$$

The effectiveness of the described criterion was reported in previous studies [8, 9].

C. Respiratory Phase Boundary Detection

Choosing the respiratory phase boundaries to include in the classification process was the main focus of the current study. In our method, the phase boundary $\hat{t}_i$, which was the ending time (to be exact, sequential number of analysis frame) of i-th expiratory phase $X_i(t_{i-1},t_i)$, corresponding to segment sequence $W_i(t_{i-1},t_i)$, was detected at the same time the acoustic likelihood of $W_i(t_{i-1},t_i)$ was calculated. According to the criterion of patient detection, we assumed that it must be effective for the detection of phase-boundaries $\{\hat{t}_i\}$ $(1 \le i \le N-1)$ based on the criterion of maximizing a difference of the two likelihoods, the acoustic likelihood for an abnormal respiration phase candidate, and that for a normal respiration phase candidate. Therefore,

$$\hat{t}_i = \underset{T_i-\Delta \le t_i \le T_i+\Delta}{\arg\max} \frac{1}{t_i - \hat{t}_{i-1}} \left| \log P(W_i^{Ab}(\hat{t}_{i-1},t_i) \mid X_i(\hat{t}_{i-1},t_i)) - \log P(W_i^{No}(\hat{t}_{i-1},t_i) \mid X_i(\hat{t}_{i-1},t_i)) \right| \quad (6)$$

$$T_i = L / N * i \quad (7)$$

where L was the duration of a lung sound sample, with the beginning time of the first phase in the sample set to zero $\hat{t}_0 = 0$. Further, T_i indicated the end of respiratory phase X_i with the lung sound sample divided equally by the number of phases N. In Eq. (6), the difference of the likelihoods was normalized by the phase duration $t_i - \hat{t}_{i-1}$ (number of analysis frames). Time width Δ was set to search an adequate phase boundary by taking into account the standard deviation of the respiratory phase duration. The search period is shown in Figure 2. The patient detection of the proposed method was performed using the detected phase boundaries $\{\hat{t}_i\}$:

$$\sum_i \log P(\hat{W}_i^{Ab}(\hat{t}_{i-1},\hat{t}_i) \mid X_i(\hat{t}_{i-1},\hat{t}_i)) > \sum_i \log P(\hat{W}_i^{No}(\hat{t}_{i-1},\hat{t}_i) \mid X_i(\hat{t}_{i-1},\hat{t}_i)) \quad (8)$$

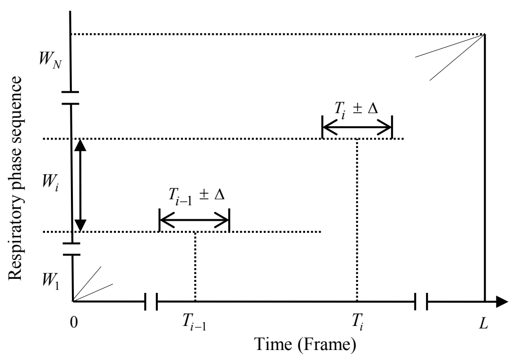

Figure 2. Detection of phase boundaries in a test sample.

C. Patient Detection Procedure

Our classification system was comprised of two processes. The first was a training process to generate acoustic models and the segment bigram, and the second was a test process for differentiating between a healthy subject and a patient using unknown lung sounds.

In the training process, segmental HMMs were adopted as an acoustic model, to describe temporal acoustic features for each kind of acoustic segment. We prepared two models for inspiration and expiration. Acoustic labels and respiratory boundaries were used in this process. Segment bigrams, which referenced the occurrence sequences of the segments in abnormal respiration, were also estimated according to the segment labels.

In the test process, input was an unknown lung sound sample composed of successive respiratory phases, where it is assumed that the number of phases, including the beginning time of the first phase, the ending time of the last phase, and the phase type (inspiratory or expiratory) of the first phase were known. The difference in the input, comparing to previous studies [8, 9], was that the temporal boundaries between the respiratory phases were not known. After acoustic feature extraction, detection of the phase boundaries and the calculation of the two likelihoods of a normal phase candidate and an abnormal candidate for each detected phase period, was carried out simultaneously. Finally, the two total likelihoods for normal and abnormal respirations were accumulated, and a classification result (healthy or diseased) was obtained.

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 PB and PC. 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 (25 ms frame-length and 10 ms frame-shift). 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 the respiratory phase, the beginning time of the first respiratory phase, and the ending time of the last phase were known. Accordingly, if the fifth 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. In addition, to ensure our experiments would be subject-independent, samples recorded from the same subject used as the test sample were excluded in the training process.

B. Classification of Healthy Subjects and Patients

1) Classification using manually segmented phases: First, to confirm the classification performance using manually segmented phases of test samples, we carried out a classification experiment (baseline). The experiment was performed by comparing the results of the two total likelihoods for all manually detected respiratory periods [Eq. (5)]. As shown in Table II, the recall rate was 73.0% for healthy subjects and 92.1% for patients. The average classification rate weighted with the data amount was 82.6%, indicated as “Average”.

2) Classification using phases divided with equal duration: According to the similar mean values and the low standard deviation of the duration for each respiratory phase (Table I), it was assumed that the simply dividing each sample would be sufficient for classification. We then divided the input sample of duration L into N phases of equal duration (L/N), where L and N were known values, and carried out experiments using these equal duration phases (equal division method). This classification method did not require manual segmentation of the respiratory phases in the same manner as the proposed method.

We obtained an average classification rate of 81.7% (Table II), which was 0.9 % lower compared to the baseline. This lower rate indicated the importance of the selection of adequate phase boundaries for the distinction between patients and healthy subjects.

3) Classification using the proposed method: A classification experiment was carried out using the proposed method [Eq. (8)]. In this experiment, adequate phase boundaries were searched around the boundary which used in the equal division method, and the search space was the period of $2\times\Delta$. In the experiment, we set $\Delta=\beta\times\sigma$, where β $(0\le\beta)$ was the constant factor defining the search space, and σ was the standard deviation of the duration for inspiratory or expiratory phase (Table I). The average classification rate between healthy subjects and patients for each β is shown in Table II and Figure 3. When β was equal to zero, the proposed method is identical to the equal division method. Figure 2 indicates that when β was larger than 0.4, the proposed method achieved a higher performance than the conventional method that used manually segmented samples, which indicated the effectiveness of our proposed method.

Finally, we investigated the difference between the boundaries detected in our proposed method and those that were manually detected. The mean and standard deviation of the boundaries (time) for the equal division method and the proposed method when β was equal to 0.5 are shown in Table III. The data in the table indicate that the mean value of the boundaries detected using the proposed method was closer to the manual boundaries than the boundaries used in the equal division method.

TABLE II. CLASSIFICATION PERFORMANCE BETWEEN HEALTHY SUBJECTS AND PATIENTS [%]

Boundary detection	β	Healthy	Diseased	Average
Manual (Baseline)	-	73.0	92.1	82.6
Equal division	0	84.3	79.1	81.7
Proposed	0.5	88.7	79.1	83.9
	1	91.3	75.7	83.5

TABLE III. MEAN VALUE AND STANDARD DEVIATION OF DIFFERENCE OF PHASE BOUNDARIES

Method	Mean [ms]	S.D. [ms]
Equal division	-120	194
Proposed $(\beta=0.5)$	-98	192

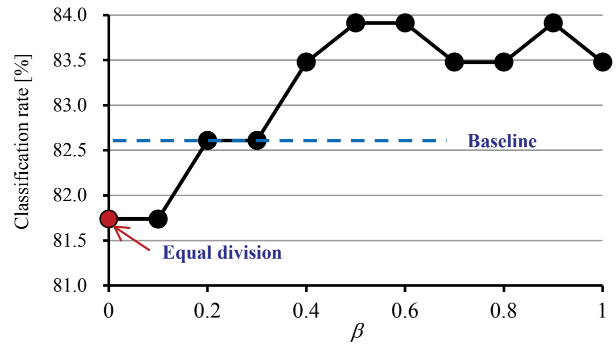

Figure 3. Classification rate of proposed method for each constant factor β.

V. CONCLUSION

This paper proposes a new method for discriminating between healthy subjects and patients with pulmonary emphysema using successive respiratory lung sound phases. Compared to previous studies, the main characteristic of the proposed method was that the method did not require difficult manual segmentation between respiratory phase boundaries. In this method, the calculation of the likelihoods for abnormal and normal respiration candidates, and the detection of phase boundaries, based on maximization of the difference between the two likelihoods, were performed simultaneously. The proposed method achieved an 83.9% classification rate, which was better than the conventional method (82.6%) that involved manual segmentation between phases, indicating the effectiveness of the proposed method.

With respect to the segmentation, however, the proposed method still required the manual detection of the beginning time of the first phase, and the ending time of the last phase in a test lung sound. Determining a solution to the manual detection problem will be the focus of our future work.

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 20th Int. Cong. 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] S. Matsutake, M. Yamashita and S. Matsunaga, "Discrimination between healthy subjects and patients using lung sounds from multiple auscultation points," *Proc. of IEEE ICASSP*, pp. 1296-1300, 2013.