ROBUST CLASSIFICATION BETWEEN NORMAL AND ABNORMAL LUNG SOUNDS USING ADVENTITIOUS-SOUND AND HEART-SOUND MODELS

Masaru Yamashita, Masataka Himeshima, and *Shoichi Matsunaga*

Department of Computer and Information Sciences, Nagasaki University, JAPAN

ABSTRACT

In this paper, we propose a robust classification method to distinguish between normal lung sounds from healthy subjects and abnormal lung sounds containing adventitious sounds from patients by using lung sounds contaminated with heart sounds. Heart sounds make it difficult to perform the aforementioned classification with high accuracy. To address this problem, we propose the use of stochastic models to represent the acoustic feature of heart sounds in the classification method based on the maximum likelihood approach by using hidden Markov models (HMMs) for the calculation of a more exact acoustic likelihood of lung sounds. Our method distinguishes between periods of adventitious sounds and heart sounds by yielding the most likely acoustic segment sequence for each respiration. For the test set of lung sounds contaminated with heart sounds, our classification method achieved a higher classification rate between normal and abnormal lung sounds compared to the conventional method without heart-sound models (88.6% vs. 87.1%, respectively). The classification of healthy and patient subjects by using the proposed method also achieved a higher classification rate of 83.0% when two samples from different auscultation points were used.

Index Terms— lung sound, classification, adventitious sound, heart sound, pulmonary emphysema

1. INTRODUCTION

Auscultation of lung sounds by using a stethoscope is one of the most popular and cost-effective medical examination methods for identifying respiratory illnesses. The auscultation performed by a physician is based on the common knowledge that abnormal respiratory sounds are frequently observed in patients with pulmonary emphysema. The abnormal sounds that are caused by anomalies in the lungs and bronchial tubes are termed *adventitious sounds* [1]. Several types of adventitious sounds exist, depending on the condition of the abnormal internal organs and the degree of inflammation, and the typical sounds include wheeze, rhonchi, and crackle, among others.

Noise pollution during auscultation makes it difficult to detect adventitious sounds through a stethoscope automatically. Almost respiration sounds contain some noises from the stethoscope or the internal organs. One of the typical noises from the internal organs is the heart sound. The main heart sounds detected in recorded lung sounds are the first heart sound (S1) and the second heart sounds (S2) during a cardiac cycle. S1 and S2 are produced by the closure of valves. The frequency of heart-sound contamination depends on the auscultation point, and the frequency is high if the auscultation point is near the heart. Physicians usually perform auscultation at several points on the patient's chest and back to make an accurate diagnosis. Detecting adventitious sounds could be quite difficult if the auscultation point is far from the abnormal region. If the abnormal region is near the heart, correctly detecting the adventitious sounds in lung sounds that are contaminated with heart sounds is necessary.

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 was to develop a home-use device to identify respiratory illness by detecting abnormal respiratory sounds. We developed a classification procedure to distinguish between normal and abnormal respiratory sounds based on the maximum likelihood approach by using hidden Markov models (HMMs) [6-8]. This procedure demonstrated the usefulness of the stochastic approach in the detection of abnormal respiratory sounds. However, noise pollution in the lung sounds during auscultation hindered the achievement of a relatively high degree of correct classification. To address this problem, we previously proposed a classification method that took into account the fact that the average duration of noise sounds was shorter than that of adventitious sounds [9]. Although this method increased performance, we did not address the lung sound recorded close the heart in order to avoid heart-sound contamination. Several conventional approaches that reduce or cancel heart sounds have been used to manage heart-sound contamination [10-12], and these procedures should be performed before the process of detecting adventitious sounds is initiated. However, because the periods of adventitious sounds, noises, and heart sounds frequently overlap, it is very difficult to detect the exact boundaries of these sounds.

To address this problem, we propose a classification method based on the maximum likelihood approach by using a stochastic heart-sound model. In addition to the HMMs of breathing sounds and adventitious sounds that were used in our previous studies [6-9], we designed a heart-sound model by using the spectral feature of S1. Our method derives the most likely acoustic segment sequence, including the heart-sound segments, for each respiration. The validity of the proposed method was confirmed through an experiment in which we classified normal and abnormal respiratory sounds. Finally, we used this classification method to differentiate between healthy subjects and patients.

2. LUNG SOUND DATA

2.1. Respiratory segments for training and evaluation

We recorded lung sounds at 12 auscultation points in patients with pulmonary emphysema and in healthy subjects by using an electronic stethoscope that incorporates a piezoelectric microphone. One lung sound sample for each auscultation point was recorded in each subject. Because the objective of this study was to develop a method for robust detection of abnormal respiratory data

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

containing adventitious sounds among lung sounds contaminated with heart sounds, we used the lung sounds recorded at two auscultation points closer to the subjects' hearts: the second (L2) and forth (L4) intercostal spaces on the subjects' front left sides. We prepared our lung sound data from the patients whose lung sound samples contained at least one adventitious sound. As a result, 37 samples from 37 patients were prepared for each auscultation point, and 37 samples from 37 healthy subjects were subsequently prepared for comparison. Each sample consisted of successive respiratory phase segments (inspiratory and expiratory), and the average number of respiratory segments was approximately 10.

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). A doctor determined the subject's state of health based on auscultation in addition to many other medical conditions. The respiratory segments 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 nonpathological adventitious sounds

· 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 of the method to classify the abnormal respiratory phase of patients and the normal respiratory phase of healthy subjects (Sections 4.2 and 4.3), and all respiration data were used to classify the patients and healthy subjects (Section 4.4). The numbers of respiratory segments are listed in Table 1.

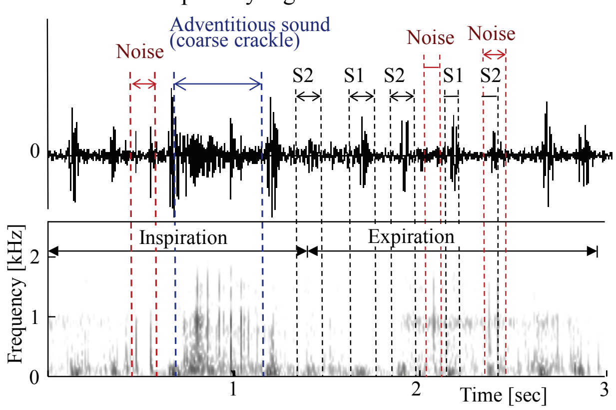

Figure 1. Typical lung sound from a patient recorded at L2

2.2. Adventitious sounds and heart sounds

A waveform and a spectrogram of a typical lung sound from a patient in our database are shown in Figure 1. During the inspiratory and expiratory periods, this lung sound contains an adventitious sound (coarse crackle), several heart sounds, and noises. Detecting the exact periods of these sounds is difficult because the spectral features of several noises are very similar to those of some adventitious sounds, and these sounds frequently appear at same time. In our database, approximately 80% of all respiration phases included some noises. The heart sounds appeared in approximately 90% of all phases in the lung sound samples recorded at L2. These numbers indicate that robust classification against heart-sound contamination is necessary to achieve higher classification performance. The sound intelligibility of S1 was greater than that of S2. In a respiratory phase, the average numbers of audible S1 and S2 sounds were 1.9 and 1.6, respectively. The spectrogram shows that the spectral energy of heart sounds are concentrated at a lower frequency compared to other types of sound. This indicated that the use of the spectral feature of the heart sound would be useful for calculating the exact acoustic likelihood for each respiration. The mean value and standard deviation of the duration of each type of sound are shown in Table 2. The duration of adventitious sounds was much longer than that of heart sounds or noise sounds, indicating that the duration of each sound would also be useful for detecting adventitious sounds.

Table 1. Number of respiratory sound samples

Respiration		Patients	Healthy subjects
L2	Normal	158 (NP)	362 (NH)
	Abnormal	214 (AP)	6 (AH)
L4	Normal	152 (NP)	362 (NH)
	Abnormal	239 (AP)	0 (AH)

Table 2. Mean value and standard deviation of duration [sec]

Sound		Mean	S.D.
Adventitious sounds		0.54	0.34
Heart sound	S1	0.096	0.021
	S2	0.090	0.017
Noise		0.13	0.20

2.3. Lung sound segmentation

The labels for the acoustic segments based on the acoustic and segmental features, such as adventitious sounds and heart sounds, were manually prepared. We assumed that a respiratory segment (period) W was composed of N successive acoustic segments as follows: we let the i-th acoustic segment be w_i $(1 \le i \le N)$. Then

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

One abnormal respiratory period comprised several breath segments, adventitious sound segments, and heart-sound segments, and one normal respiratory period comprised breath segments and heart-sound segments. Compared to our previous works [7,9], main distinguishing characteristics of this study was the use of heart-sound segments. Each heart sound was presented by using a S1 or S2 segment, and each adventitious sound was presented by using a continuous or discontinuous acoustic segment. Some typical examples of discontinuous acoustic segments are coarse crackle, fine crackle, and pleural friction rub. Rhonchus and wheezing sounds are examples of continuous acoustic segments. If an adventitious sound and a heart sound appeared at same time, we treated this acoustic period as an adventitious sound segment.

3. CLASSIFICATION METHOD

Our classification strategy comprised a training process and a test process. In the training process, three components concerning acoustic knowledge were generated by using the labels of acoustic segments: acoustic HMMs for each acoustic segment (breath segment, continuous/discontinuous adventitious segment, and heart-sound segment), an acoustic segment bigram, and a set of deterministic connection rules of acoustic segments.

In the test process, the likelihood of a normal/abnormal respiratory phase was calculated based on the maximum likelihood approach. We let the occurrence probability of the segment sequence W be $P(W)$. We used a segmental bigram to calculate $P(W)$ [7]. The total likelihood was composed of the acoustic likelihood calculated from the HMMs and the segmental sequence likelihood calculated from the bigram. The segment sequence $\hat{W}$ with the highest likelihood $\log P(\hat{W} \mid X)$ for an unknown respiratory input X is given below by using Bayes' theorem:

$$\hat{W} = \arg\max_{W}\left[\log P(W \mid X)\right] \approx \arg\max_{W}\left[\alpha \log P(W) + \log P(X \mid W)\right] \quad (2)$$

where $\log P(X \mid W)$ is the acoustic likelihood. The weight factor α controlled the contribution of the bigram, and this factor was obtained experimentally. Compared to our previous works [7-9], the distinguishing feature of the proposed method was that $P(W)$ and $P(X \mid W)$ in Eq. (2) were calculated by including the occurrence probability and the acoustic likelihood of heart-sound segments, respectively.

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 the power was computed by using a 25-ms Hamming window. This vector was used as an acoustic feature when modeling the HMMs. All HMMs were generated for each auscultation point by using only the lung sound samples recorded at the auscultation point. The respiratory periods from healthy subjects (NH in Section 2.1), excluding the heart-sound periods, were used to train the model for the breath segment of normal respiration. The models for heart sounds were trained by using both normal respiration and abnormal respiration, whereas the models for adventitious sounds were generated by using the sounds obtained from the patients (AP). The HMMs with three states and two Gaussian probability density functions were used. The segment bigram was also trained by using the segment labels of the training samples according to the details of segmentation.

In our experiments, we assumed that the number of respiratory periods for each lung sound sample, the respiratory phase, and the respiratory boundaries were 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, our experiments were subject-independent because the samples recorded from the same subject used as the test sample were excluded in the training process.

4.2. Preliminary classification experiments

To evaluate the validity of the use of heart-sound HMMs, three preliminary classification experiments were carried out in which different segment HMMs and deterministic connection rules were used. In the first experiment, the HMMs for the breath and adventitious segments were used, but the heart-sound model was not used. In the second experiment, the heart-sound model trained by using S1 segments (abbreviated as the S1 model) was used, and both the S1 and S2 models were used in the third experiments, adding to the models in the first experiment. The acoustic bigram was not used in each experiment. The evaluation data (NH and AP) were the lung sounds recorded at L2 in Table 1.

The classification results are shown in Table 3. These results confirmed that the use of heart-sound models was beneficial. The performance was better when using the S1 model compared to the performance when using both the S1 and S2 models. We assumed that the lower classification rate when using the two models was caused by the unintelligibility of S2. Thus, we used the S1 model as the heart-sound model in the subsequent experiments.

4.3. Classification of normal and abnormal respirations

To confirm the effectiveness of the proposed method using the S1 model, three classification methods were used to distinguish between abnormal and normal respirations: a baseline method, a duration-constrained method, and the proposed method. In the baseline method [7], the acoustic HMMs, the segment bigrams, and the connection rules were used, whereas heart sound was not taken into account. The duration-constrained method proposed in our previous study [9] took into account the duration distributions of noise and adventitious sounds to ensure the baseline was robust to noises. This was because the duration of noise sounds was remarkably shorter than that of adventitious sounds. This method decreased the misrecognition of short noises as adventitious sounds. As shown in Table 2, the durations of heart sounds and noises were shorter than that of adventitious sounds. Therefore, we presumed that this method was also effective for the detection of abnormal respiration contaminated with heart sounds.

The classification results obtained for two auscultation points are shown in Table 4, where the evaluation data were the respiration data related to AP and NH in Table 1. The result of the duration-constrained method is represented by the "Duration" and the average classification rate weighted with the data amount is represented by the "Average." The average classification rate of the proposed method (88.6%) was the highest among the three methods, validating the use of the heart-sound model. The values obtained for the duration-constrained method (87.3%) and the proposed method (88.6%) indicate that the heart-sound spectral information was more useful than the duration information. The detection rates of normal respiration for both auscultation points were increased in the proposed method, and this was because the proposed method decreased the misrecognition of heart sounds as adventitious sounds. On the other hand, the detection rates of

Table 3. Classification rate with or without heart-sound models [%]

Heart-sound model	None	S1	S1 and S2
Classification rate	86.3	89.1	87.2

Table 4. Classification performance between normal and abnormal respiration [%]

Auscul. point	Method	Healthy, normal (NH)	Patient, abnormal(AP)	Average
L2	Baseline	88.1	86.0	87.3
	Duration	90.3	82.7	87.5
	Proposed	93.6	83.2	89.8
L4	Baseline	87.6	85.8	87.1
	Duration	88.1	85.4	87.0
	Proposed	89.5	84.1	87.3
Average	Baseline	87.8	85.9	87.1
	Duration	89.2	84.1	87.3
	Proposed	91.6	83.7	88.6

Table 5. Classification performance for segments that did or did not contain heart sounds [%]

Segments		Baseline	Duration	Proposed
Heart sounds	No. phases			
Present	1007	86.0	86.4	87.6
Absent	170	93.5	92.9	93.5

abnormal respiration were decreased because adventitious sounds were incorrectly recognized as heart sounds.

To analyze the validity of the proposed method, the ability to classify segments that contained heart sounds or no heart sounds was assessed (Table 5). Each classification rate is shown as the average value for all samples recorded at two auscultation points. The purpose of our method was to decrease the misrecognition of lung sounds containing heart sounds as abnormal sounds; this purpose was achieved to some extent because the classification rate of lung sounds that included heart sounds increased from 86.0% to 87.6%. Furthermore, the performance was not decreased when respiratory phases that did not contain heart sounds were classified (93.5%). This indicates that, although our proposed method was designed for the classification of lung sounds contaminated with heart sounds, this method would be also applicable to lung sounds with no heart sounds. While our method improved the classification rate to 87.6% for lung sounds containing heart sounds, this rate is far lower than the rate for sounds that do not contain heart sounds (93.5%). This indicated that the performance of this method could be improved considerably.

4.4. Classification of healthy subjects and patients

4.4.1. Classification using a single sample

To verify our proposed method, the healthy subjects and the patients were classified for each auscultation point. If the total likelihood of abnormal respiration was greater than that of normal respiration for the test sample, the subject was regarded as a patient. The evaluation data were 74 lung sound samples from 37 patients and 37 healthy subjects for each auscultation point. Each sample from the patients included at least one adventitious sound.

The classification performance in this experiment is shown in Table 6. For each auscultation point, the classification performance of the proposed method (91%) was superior to that of the baseline method (88%). The results shown in Tables 4 and 6 indicate that the acoustic modeling of heart sounds yielded better results for the classification between healthy and patient subjects as well as between normal and abnormal respiration.

4.4.2. Classification using samples from two auscultation points

Finally, a classification experiment was performed to distinguish between healthy subjects and patients by using two lung sound samples recorded from different auscultation points. In this experiment, we prepared 56 samples from 56 patients and 56 samples from 56 healthy subjects for each auscultation point. These samples included those that were used in the previous experiments. The samples recorded from both L2 and L4 contained adventitious sounds for 19 out of 56 patients, whereas only the sample from L2 or from L4 contained these sounds for the other 37 patients. In this experiment, we compared the normal and abnormal average likelihoods by using two samples for each subject where the total likelihood was averaged by the number of respiratory periods. The subject was regarded as a patient if the average likelihood for abnormal respiration was greater than that for normal respiration. The usefulness of this classification criterion utilizing multiple lung sounds was shown in our previous work [13] in which lung sounds that did not contain heart sounds were used. In this paper, we adopted this classification criterion to confirm the validity of the proposed classification method for the use of two lung sounds contaminated with heart sounds.

The upper part of Table 7 shows the average classification performance when using one sample from L2 or L4, and the lower part of this table shows the performance when using both samples. These data indicated that the proposed method achieved a higher classification rate than the baseline method when using either a single lung sound or two lung sounds from different auscultation points. Further, the classification rate of the proposed method showed better performance when using two lung-sound samples compared to the method using a single sample, even when the lung sounds were highly contaminated with heart sounds.

Table 6. Classification performance for each auscultation point [%] (No. of detected subjects/No. of all subjects)

Aus. points	Methods	Healthy	Patients	Average
L2	Baseline	89 (33/37)	86 (32/37)	88 (65/74)
	Proposed	95 (35/37)	86 (32/37)	91 (67/74)
L4	Baseline	89 (33/37)	86 (32/37)	88 (65/74)
	Proposed	97 (36/37)	84 (31/37)	91 (67/74)

Table 7. Average classification performance between patients and healthy subjects [%] (No. of detected subjects/No. of all subjects)

No. of aus. points	Methods	Healthy	Patients	Average
1 (L2 or L4)	Baseline	85.7	73.2	79.0
	Proposed	88.4	73.2	80.8
2 (L2 and L4)	Baseline	88 (49/56)	75 (42/56)	81.3
	Proposed	93 (52/56)	73 (41/56)	83.0

5. CONCLUSIONS

This paper proposed a new method for discriminating between normal respiration from healthy subjects and abnormal respiration containing adventitious sounds from patients for the automatic auscultation of lung sounds contaminated with heart sounds. The key characteristic of the proposed classification method was the explicit use of the heart-sound model. In the proposed method, stochastic likelihoods were calculated based on the maximum likelihood approach by using acoustic-segment HMMs for acoustic spectral features and segment bigrams for the occurrence feature of acoustic segments, where heart-sound segments were taken into account. The heart-sound HMM was generated by using the first heart sound, S1, during a cardiac cycle.

According to the classification experiments for distinguishing between normal and abnormal respiration, the proposed classification method, which accounted for the spectral feature of heart sounds, achieved a better performance than both the conventional method [7], which did not account for heart sounds, and the duration-constrained method [9], which accounted for the difference duration distribution of heart sounds and adventitious sounds. Furthermore, the proposed method achieved a better classification performance when distinguishing between healthy subjects and patients.

With respect to the classification of normal and abnormal respiration, however, our experiments indicated that the performance of this method could be improved considerably. This is a subject for 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 ICA*, 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.

[10] J. Gnitecki and Z. Mousasavi, “Separating heart sounds from lung sounds – accurate diagnosis of respiratory deseases depends on understanding noises,” *IEEE Eng. Med. Bio. Mag.,* Vol.26, 1, pp.20-29, 2007

[11] C. Ahlstrom, O. Liljefeldt, P. Hult and P. Ask, “Heart sound cancellation from lung sound recordings using recurrent time statistics and nonlinear prediction,” *IEEE Signal Processing Letters,* vol.12, no.12, pp.812-815, 2005.

[12] J. Feng and F. Sattar, “A multiscale mean shift localization approach for robust extraction of heart sounds in respiratory signals,” *Proc. of IEEE ICASSP*, pp. 1291-1295, 2013.

[13] 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.