Biomedical acoustics: Paper ICA2016-324

Detection of abnormal lung sounds considering spectral and temporal features of heart sounds

Megumi Taguchi\textsuperscript{(a)}, Masaru Yamashita\textsuperscript{(b)}, Shoichi Matsunaga\textsuperscript{(c)}

\textsuperscript{(a)} Nagasaki University, Japan, b312023@cis.nagasaki-u.ac.jp
\textsuperscript{(b)} Nagasaki University, Japan, masaru@cis.nagasaki-u.ac.jp
\textsuperscript{(c)} Nagasaki University, Japan, mat@cis.nagasaki-u.ac.jp

Abstract

In this paper, we propose a robust classification method for lung sounds contaminated with heart sounds in order to distinguish between healthy subjects and abnormal patients with pulmonary emphysema. We previously developed a classification procedure based on a maximum-likelihood approach by using hidden Markov models (HMMs). However, contaminated heart sounds caused difficulties in achieving a highly accurate classification, because it was difficult to generate HMMs that distinguished between adventitious sounds and heart sounds with high accuracy, by using power and spectral acoustic features only. To address this problem, we propose a classification technique that is based on the use of spectral features and temporal features related to heart sounds: distributions of durations and time intervals of heart (S1) sounds. A validity score for detected adventitious sounds and heart sounds in the classification process is designed by considering the distribution of time intervals of heart sounds and differences in the durations between the adventitious sounds and the heart sounds. In the proposed method, the total likelihood of each respiratory sound is obtained by summing the spectral likelihood derived from the HMMs and the validity score. In the classification of healthy subjects and patients using 94 lung sound samples from 94 subjects, the proposed method achieved a higher classification rate (90\%) than the baseline method (84\%) using only the spectral features, thus demonstrating the superiority of the proposed method.

Keywords: lung sounds, adventitious sounds, heart sounds, pulmonary emphysema
Detection of abnormal lung sounds considering spectral and temporal features of heart sounds

1 Introduction

Individuals with respiratory disorders usually exhibit abnormal respiratory sounds known as adventitious sounds [1]. Auscultation of lung sounds is an effective method for identifying respiratory illnesses. However, it is difficult for individuals without expertise concerning this method to provide an accurate diagnosis since in-depth experience and knowledge of a doctor are required for the same. Hence, automated determination of respiratory diseases by using respiratory sounds is useful for the early detection of lung disorders.

Several studies on the acoustic analysis of breathing sounds by using spectral features were conducted to assist doctors in detecting and diagnosing pulmonary emphysema [2-5]. Our studies aimed to develop effective and robust methods for identifying respiratory illnesses by differentiating abnormal respiratory sounds from normal lung sounds. Our previous studies developed a classification procedure for distinguishing between a patient and a healthy subject based on a maximum likelihood approach by using hidden Markov models (HMMs) [6-10]. This procedure used spectral features as acoustic parameters in HMMs, and demonstrated the usefulness of a stochastic approach in the detection of abnormal respiratory sounds in patients.

However, it is difficult to automatically and accurately detect adventitious sounds through a stethoscope owing to noise pollution during auscultation. Most respiration sounds are accompanied by noises from a stethoscope or internal organs. Examples of typical noises from the internal organs include heart sounds. The main heart sound detected during a cardiac cycle is the first heart sound, S1. This is produced by the closure of valves. In addition, the spectral features of several noises are very similar to those of certain types of adventitious sounds. In order to address this issue, extant research proposed a classification method using a stochastic heart-sound model [11]. In this study, in addition to the HMMs of breathing sounds and adventitious sounds, a heart-sound model was designed by using the spectral feature of S1. Although this method increased the classification performance, the increase in the range was insufficient.

This study proposed a classification technique based on the use of spectral features and temporal features related to heart sounds in order to address the fore-mentioned problem. An investigation on the respiratory sounds revealed significant difference between the duration of heart sounds and that of adventitious sounds. Furthermore, it is widely known that the occurrence of heart sounds is periodic. The technique proposed in this study considered these two characteristics of temporal information in conjunction with each other to decrease the frequency of misrecognition of heart sounds as adventitious sounds. The total likelihood of each respiratory sound was obtained by adding the spectral likelihood derived from HMMs and the validity score calculated by using the periodicity of heart sounds and the distribution functions of duration for heart sounds and adventitious sounds. The validity of the proposed method was
confirmed through a classification experiment including healthy subjects and patients with respiratory disorders.

2 Lung sound data

2.1 Recording of lung sound data

Lung sounds in both patients with pulmonary emphysema and in healthy subjects were recorded by using an electronic stethoscope that incorporated a piezoelectric microphone. The lung sounds were collected by indicating a sign for the beginning of each respiration phase to the subjects on a computer monitor. Two auscultation points were considered, namely the second (L2) and fourth (L4) intercostal spaces on the front left sides of the subjects. The intercostal space L2 was closer to the subjects’ hearts than L4. A lung sound sample for each auscultation point was recorded for each subject. The number of samples is listed in Table 1. The segments were tagged according to the respiratory phase (inspiratory or expiratory), diagnostic state (normal or abnormal), and subject’s health (healthy individuals or individuals with pulmonary emphysema). A subject’s health was determined by a doctor on the basis of auscultation and the presence of other medical conditions.

A spectrogram of a typical lung sound recorded from a patient is shown in Figure 1. This lung sound contained adventitious sounds (fine crackles), several heart sounds, and noises. Thus, detecting the exact periods of the sounds was difficult because the spectral features of the sounds were frequently very ambiguous and unclear. Approximately 80% of all respiration phases in the study database included some noises. Heart sounds appeared in approximately 90% of the lung sound samples recorded at L2. Hence, it was necessary to develop a robust classification method against heart-sound contamination.

Table 1: Number of respiration samples for each auscultation point

<table>
<thead>
<tr>
<th>Auscultation point</th>
<th>Healthy subjects</th>
<th>Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2</td>
<td>47</td>
<td>47</td>
</tr>
<tr>
<td>L4</td>
<td>39</td>
<td>39</td>
</tr>
</tbody>
</table>

Figure 1: Example of lung sound from a patient with pulmonary emphysema
2.2 Labels for acoustic segments

Each lung sound sample $S$ consisted of several successive respiratory phases $W$ expressed as follows:

$$S = W_1 W_2 \cdots W_i \cdots W_M,$$

where $W_i$ was the $i$-th respiratory phase for which the beginning and ending times were manually detected. The average number of respiratory segments per sample was approximately 10.0. The phase boundaries were used in a training process and a test process.

The segment types were defined according to their acoustic and temporal features, and a symbol was assigned to each segment. It was assumed that an inspiratory/expiratory sound period $W$ comprises $N$ segments, and the $i$-th segment was denoted as $w_i (1 \leq i \leq N)$. This can be expressed as follows:

$$W = w_1 w_2 \cdots w_N$$

A normal respiratory period $W$ in a respiration sample from a healthy subject consisted of normal breath segments and heart-sound segments. An abnormal respiratory period $W$ in a sample from a patient was composed of several segments that included adventitious sound segments, heart sound segments and normal breath segments. The main heart sounds detected in recorded lung sounds included the first heart sound (S1) and the second heart sound (S2). In this study, only S1 was considered to represent heart sounds because the sound intelligibility of S1 exceeded that of S2. A continuous or discontinuous sound segment was used to represent each adventitious sound. Typical examples of discontinuous sound segments included coarse crackles, fine crackles, and pleural friction rubs. Rhonchus or wheezing sounds are examples of continuous segments [1]. If an adventitious sound and a heart sound appeared simultaneously as shown in Figure 1, the acoustic period was treated as an adventitious sound segment.

2.3 Investigation for adventitious sounds and heart sounds

The duration of the heart sounds and adventitious sounds was investigated by using labels of the lung sound data. Table 2 shows the mean values and the standard deviations for adventitious sounds and heart sounds. A statistical difference in the duration of adventitious and heart sounds was obtained by comparing the figures in Table 2. The duration of most heart sounds was less than 150ms. Conversely, there was a considerable variation in the duration of each adventitious sound because of the diversity of the adventitious sounds. We considered this difference as useful in reducing the misrecognition of heart sounds as adventitious sounds.

<table>
<thead>
<tr>
<th>Sound</th>
<th>Mean</th>
<th>S. D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adventitious sound</td>
<td>0.53</td>
<td>0.31</td>
</tr>
<tr>
<td>Heart sound (S1)</td>
<td>0.12</td>
<td>0.03</td>
</tr>
</tbody>
</table>
This was followed by an investigation of the periodicity of the heart sounds. All heart sounds were not audible in the recorded auscultation sounds. In the samples recorded at L2 and L4, the ratios of the number of detected S1 to the number of assumed occurrences of S1 were 37% and 29%, respectively. The time interval between two neighbouring S1 sounds is shown in Table 3. The high periodicity was expected to be effective in the detection of heart sounds. Thus, in this study it was assumed that the accurate detection of the heart sounds led to the accurate detection of the adventitious sounds.

Table 3: Mean value and standard deviation of the time interval between two S1 sounds [s]

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>S. D.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.61</td>
<td>0.11</td>
</tr>
</tbody>
</table>

3 Classification methods

3.1 Likelihood calculation for each respiratory phase

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

$$P(W) = P(w_1w_2 \cdots w_i \cdots w_n)$$

(3)

A segmental bigram was used to calculate $P(W)$ as given below:

$$P(W) \approx \sum_{i=2}^{N} P(w_i \mid w_{i-1})$$

(4)

The total likelihood composed of the acoustic likelihood derived from segment HMMs and the segmental sequence likelihood derived from (4) was calculated for each respiratory period $W$. The derived diagnostic state (normal/abnormal) for a respiratory input that provided the segment (sequence) $\hat{W}$ with the highest likelihood $\log P(\hat{W} \mid X)$ is expressed as follows:

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

(5)

where $X$ is the respiratory input and $P(X \mid W)$ is the acoustic likelihood derived from HMMs. The weight factor $\alpha$ controlled the contribution of the bigram. In this study, $\alpha$ was experimentally acquired. The likelihood $\log P(\hat{W}_j^{ab} \mid X_j)$ of an abnormal respiration candidate $\hat{W}_j^{ab}$ and the likelihood $\log P(\hat{W}_j^{no} \mid X_j)$ of a normal respiration candidate $\hat{W}_j^{no}$ for the $j$-th respiratory $X_j$ ($1 \leq j \leq M$) were calculated to classify patients and healthy subjects.

3.2 Criteria for detection of patient subjects

A subject was considered to be a patient when the total likelihood of the abnormal respiration candidate for each respiratory period exceeded that of candidates with normal respiration [7]:
$$\sum_{j=1}^{M} \log P(\hat{W}_j | X_j) \geq \sum_{j=1}^{M} \log P(\hat{W}_j | X_j)$$  \hspace{1cm} (6)$$

3.3 Validity score as adventitious sounds

In the proposed method, the characteristics of temporal information including adventitious sounds and heart sounds were considered. The method comprises two steps: first, reliable heart sounds were selected in the recognized segment sequence by using the periodicity information of S1, and second, misclassified heart sounds (mistakenly classified as adventitious sounds) were detected using the duration information of the heart sounds and adventitious sounds. Thereafter, total likelihood was adjusted by using the results of the detection. The flow of classification including the proposed method is shown in Figure 2.

3.3.1 Heart sound detection

First, all recognized segment sequences with the highest likelihood were concatenated:

$$\hat{W}_1 \hat{W}_2 \ldots \hat{W}_i \ldots \hat{W}_M$$  \hspace{1cm} (7)$$

For each detected heart sound in the sequence, the correctness of occurrence timing was evaluated by using other detected heart sounds. In the evaluation, the mean and standard deviation of the interval of S1 (shown in Table 3) was used to confirm the correctness. For each detected heart sound, the number of other detected heart sounds that occurred periodically was counted. Given that the mean and standard deviation of the duration of S1 are $\mu_D$ and $\sigma_D$ and that the mean and standard deviation of the interval of S1 are $\mu_I$ and $\sigma_I$, and if the target heart sound is detected correctly, then it is assumed that the onset time $t$ of each of the heart sounds is given by

$$n(\mu_D - 3\sigma_D) + (n-1)(\mu_D - 3\sigma_D) \leq t \leq n(\mu_D + 3\sigma_D) + (n-1)(\mu_D + 3\sigma_D)$$  \hspace{1cm} (8)$$

where $n$ represents an integer. The heart sound sequence with the maximum number of heart sounds within an adequate period was defined as the detection result.
3.3.2 Validity score

Further, in the study, if misrecognition between heart sounds and adventitious sounds was detected, then the total likelihood of the recognized candidate was modified based on the duration of each sound. A normal distribution was adopted to describe the occurrence probability of the duration for heart sounds or adventitious sounds.

In the proposed method, the cross point of the two probability density functions (pdfs) for the heart sounds and adventitious sounds was initially set as the threshold $T$. If the duration of the obtained adventitious sound segment that occurred at the occurrence timing of a heart sound in (8) was shorter than the threshold, it was assumed that the probability of misrecognition of the heart sound as adventitious sound was high. The validity score $V$ of duration was defined by using the ratio of the pdf $f(x)$ ($= N(0.12, 0.03^2)$) of the heart sound duration to the pdf $g(x)$ ($= N(0.53, 0.31^2)$) of the duration of the adventitious sound as follows:

$$V(\hat{W}_1, \hat{W}_2, \ldots \hat{W}_m) = \sum_{i=1}^{K} d(x_i)$$

(9)

where $x_i$ is the onset time of the $k$-th adventitious sound segment that satisfied (8), and $d$ is the validity score for each segment. The classification criterion in (6) differentiating between patients and healthy subjects was modified as follows:

$$\sum_{j=1}^{M} \log P(\hat{W}_j | X_j) + \beta V + \gamma \sum_{j=1}^{M} \log P(W_j | X_j)$$

(10)

where $\beta$ and $\gamma$ are weight factors. The values were obtained experimentally to obtain the highest performance for the methods detailed below.

Two types of validity scores were used. Method I used the following equation:

$$d(x_i) = \begin{cases} \log \frac{g(x_i)}{f(x_i)} & \text{if } x_i < T \\ 0 & \text{if } x_i \geq T \end{cases}$$

(11)

Equation (11) indicated that if and only if the probability of misrecognition of the heart sound to the adventitious sound segment was high, then the total likelihood for the candidates with abnormal respiration reduced.

In Method II, for every duration $x$ of the detected adventitious sound segment, the total likelihood was adjusted by using the following score:

$$d(x_i) = \log \frac{g(x_i)}{f(x_i)}$$

(12)

In Method II, if the probability of misrecognition of the heart sound to the adventitious sound was high ($x_i < T$), then the total likelihood reduced, and if the probability of the correct detection of the adventitious sound was high ($x_i > T$), then the total likelihood increased.
4 Evaluation experiments

4.1 Experimental conditions

We performed classification tests to evaluate the proposed methods. The total number of test samples was 172, and they were recorded from auscultation points L2 and L4 listed in Table 1. The lung sound data were sampled at 10 kHz. A vector of 5 mel-frequency cepstral coefficients (MFCC) and power was computed every 10 ms by using a 25-ms Hamming window. The respiratory sounds from healthy subjects were used to generate the acoustic models (HMMs) for normal respiration, and the sounds from patients were used to generate the models for adventitious sound segments. All heart sounds from healthy subjects and patients were used to generate the HMM for heart sound. The method used HMMs with three states and two Gaussian probability density functions. In the experiments, it was assumed that both the respiratory phase (inspiratory or expiratory) and phase boundaries were known. A leave-one-out cross-validation for each test phase was performed. The HMMs were subject-independent because the samples recorded from the same subject used as the test sample were excluded in the training process.

4.2 Classification of healthy and patient subjects

Classification experiments for distinguishing between healthy subjects and subjects with pulmonary emphysema (patients) were performed. In these experiments, the following five methods were examined: a conventional method (C1) that only used the HMMs of normal breath sounds and adventitious sounds described by previous research [8], a conventional method (C2) that used the aforementioned HMMs and HMMs of heart sounds S1 [11], a baseline method (Baseline) that used HMMs and segmental bigrams including S1, a proposed method (Method I) that considered the validity score of incorrect recognition for adventitious sounds, and a proposed method (Method II) that considered the validity score of incorrect and correct recognition for adventitious sounds. The Baseline, Method I, and Method II used the same segment bigram model. However, the segment model in C1 excluded segment S1. In C1, the deterministic connection rules of the acoustic segments were used instead of bigrams. The characteristics of each method are listed in Table 4.

<table>
<thead>
<tr>
<th>Method</th>
<th>Heart sound S1</th>
<th>Segment bigram</th>
<th>Validity score</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1 [8]</td>
<td>-</td>
<td>V</td>
<td>-</td>
</tr>
<tr>
<td>C2 [11]</td>
<td>V</td>
<td>Deterministic rules</td>
<td>-</td>
</tr>
<tr>
<td>Baseline</td>
<td>V</td>
<td>V</td>
<td>-</td>
</tr>
<tr>
<td>Method I (proposed)</td>
<td>V</td>
<td>V</td>
<td>(11)</td>
</tr>
<tr>
<td>Method II (proposed)</td>
<td>V</td>
<td>V</td>
<td>(12)</td>
</tr>
</tbody>
</table>

Table 5 shows the classification performance of the five methods for respiratory samples recorded from each auscultation point. Method II achieved the highest performance for the auscultation point L2, which was nearer to the subjects' hearts than L4. Particularly, the
detection rate of patients increased. These results suggested that a validity score that accounts for both the misrecognition of heart sounds to adventitious sounds and the correct detection of adventitious sounds is useful. Conversely, there was no increase in the classification of respiratory data recorded from L2. It was assumed that this was because the proposed algorithm could not accurately capture the occurrence of the heart sounds. This indicated that if the frequency of heart sound contamination is infrequent, then the proposed algorithm was ineffective. This will be a topic of future research.

<table>
<thead>
<tr>
<th>Auscultation</th>
<th>Method ($\beta, \gamma$)</th>
<th>Healthy</th>
<th>Patients</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2</td>
<td>C1</td>
<td>85 (40/47)</td>
<td>85 (40/47)</td>
<td>85 (80/94)</td>
</tr>
<tr>
<td></td>
<td>C2</td>
<td>91 (43/47)</td>
<td>81 (38/47)</td>
<td>86 (81/94)</td>
</tr>
<tr>
<td></td>
<td>Baseline</td>
<td>91 (43/47)</td>
<td>77 (36/47)</td>
<td>84 (79/94)</td>
</tr>
<tr>
<td></td>
<td>Method I (50, 200)</td>
<td>91 (43/47)</td>
<td>79 (37/47)</td>
<td>85 (80/94)</td>
</tr>
<tr>
<td></td>
<td>Method II (30, 190)</td>
<td>91 (43/47)</td>
<td>89 (42/47)</td>
<td>90 (85/94)</td>
</tr>
<tr>
<td>L4</td>
<td>C1</td>
<td>82 (32/39)</td>
<td>90 (35/39)</td>
<td>88 (69/78)</td>
</tr>
<tr>
<td></td>
<td>C2</td>
<td>90 (35/39)</td>
<td>90 (35/39)</td>
<td>90 (70/78)</td>
</tr>
<tr>
<td></td>
<td>Baseline</td>
<td>95 (37/39)</td>
<td>82 (32/39)</td>
<td>88 (69/78)</td>
</tr>
<tr>
<td></td>
<td>Method I (10, 30)</td>
<td>95 (37/39)</td>
<td>82 (32/39)</td>
<td>88 (69/78)</td>
</tr>
<tr>
<td></td>
<td>Method II (30, 150)</td>
<td>90 (35/39)</td>
<td>87 (34/39)</td>
<td>88 (69/78)</td>
</tr>
</tbody>
</table>

5 Conclusions

This study proposed a robust classification method for lung sounds contaminated with heart sounds in order to distinguish between healthy subjects and patients with pulmonary emphysema. The key characteristic of the proposed classification method included the use of the validity score representing the misrecognition of heart sounds as adventitious sounds. In the proposed method, the sequence of heart sounds was estimated by using the periodicity of the heart sounds. Then, the score of misrecognition of heart sounds to adventitious sounds was calculated based on the difference in duration between heart sounds and adventitious sounds.

The classification experiments distinguishing between healthy subjects and patients revealed that the proposed classification method, which accounted for the misrecognition of heart sounds to adventitious sounds and the correct detection of adventitious sounds, indicated the best performance for the respiration data recorded at auscultation point close to subjects’ hearts. However, with respect to the classification for respiration that did not significantly include heart sounds, the experiments indicated that there was no improvement in the performance. This is a subject for future research.

References


