Classification Between Abnormal and Normal Respiration Through Observation Rate of Heart Sounds Within Lung Sounds

Kimitake Ohkawa, *Masaru Yamashita* and *Shoichi Matsunaga*
Department of Computer and Information Sciences
Nagasaki University
Nagasaki, Japan

***Abstract*—This paper proposes an effective classification method to differentiate between normal and abnormal lung sounds, which takes into account the detection level of heart sounds. Abnormal lung sounds frequently contain adventitious sounds; however, misclassification between heart sounds and adventitious sounds makes it difficult to achieve a high level of accuracy. Furthermore, the classification performance of conventional methods, which use the detection function of heart sounds, becomes worse for those lung sounds which contain a low level of heart sounds. To address this problem, our proposed method changes the classification method according to the detection rate of heart sounds, whereby if the rate was high, the heart-sound models in the HMM-based classification method were used. In addition to spectral information, temporal information of heart sounds and adventitious sounds were also used to obtain the rate more precisely. When using lung sounds from three auscultation points, the proposed method achieved a higher classification performance of 89.9% (between normal and abnormal respiration) compared to 88.7% for the conventional method, which used the detection function of heart sounds. Our approach to the classification of healthy and unhealthy subjects also achieved a higher classification rate of 86.6%, compared to 83.1% when using the conventional method having the detection function of heart sounds.**

Keywords—lung sound; HMM; classification; heart sound; adventitious sound

I. Introduction

Lung sounds of individuals with respiratory disorders frequently contain abnormal respiratory sounds known as adventitious sounds [1]. Auscultation of lung sounds is both an effective and economic method of identifying respiratory illnesses, by observing the presence of adventitious sounds; however, it is difficult for individuals without expertise in auscultation to provide an accurate diagnosis. For this reason, a more automated determination of respiratory diseases, using respiratory sounds, is useful.

Several studies on the detection of specific adventitious sounds, such as “crackles”, by using spectral features through conventional pattern matching, or recent machine learning techniques, were conducted [2-8]. Our studies aimed to develop effective and robust methods for identifying respiratory problems, by differentiating abnormal respiratory sounds from normal sounds. Previously, we also developed a classification procedure for distinguishing between a healthy and unhealthy subject, based on a maximum likelihood approach, using hidden Markov models (HMMs) [9, 10]. However, automatic and accurate detection of adventitious sounds is difficult owing to pollution created by the heart. The main heart sound observed is the first one (S1), and the spectral features of S1 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 and adventitious sounds, a heart-sound model was designed by using the spectral features of S1. Furthermore, to capture heart sounds precisely, we used the distribution of time intervals within heart sounds and the differences in the durations between the adventitious and heart sounds [12]. These trials improved both the recognition of adventitious sounds from heart sounds, and the classification performance of respiration data, which contains many heart sounds. However, with respect to the classification for respiration data that did not include significant amounts of heart sounds, the classification experiments indicated a decrease in the performance. This was a serious problem, because an observation rate of heart sounds can vary greatly, depending on both individuals and auscultation points.

To address this problem, we propose a classification technique which chooses one of two methods, according to the detection rate of heart sounds. In our method, first, an acoustic segmentation of a target lung sound is performed using HMMs, including heart-sound models, then the observation rate of the heart sounds is estimated using both spectral and temporal features, related to both heart and adventitious sounds. If the detection rate is low, HMM-based classification is performed without using heart-sound models. Otherwise, the classification result using the heart-sound models (where this result is obtained in the segmentation process), becomes the final result. Furthermore, the temporal information of both sounds is used in acoustic likelihood calculation for the classification process. The effectiveness of the proposed method is confirmed in two ways: the classification

ISBN 978-90-827970-1-5

experiment between normal and abnormal respiration, and that between healthy subjects and those with pulmonary emphysema.

II. ACOUSTIC LABELS OF LUNG SOUND DATA

A. Lung Sounds for Training and Evaluation

We recorded lung sounds from unhealthy subjects with pulmonary emphysema, and also from healthy subjects, by using an electronic stethoscope that incorporates a piezoelectric microphone. We recorded at three auscultation points (L1, L2, and L3), on the front left side of the subjects. These auscultation points are shown in Fig. 1, where L1 is the second intercostal space. For the auscultation point, one lung-sound sample was recorded for each subject, and each sample consisted of successive respiratory phase segments (inspiratory and expiratory periods):

$$W_1 W_2 \cdots W_i \cdots W_I \qquad (1)$$

where W_i is the i-th respiratory phase $(1 \leq i \leq I)$ in which the period was manually detected. The average number of respiratory segments per sample was approximately 10. The number of recorded samples for each point is shown in Table I, which also shows the number of samples containing heart sounds and its ratio. The table also shows that the samples recorded at L2 contained the most heart sounds among the three points.

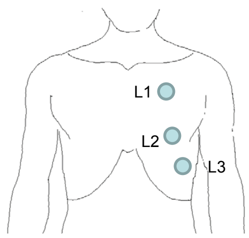

Fig. 1. Auscultation pointns on the front left side of subjects.

TABLE I. NUMBER OF RECORDED SAMPLES

Points	Unhealthy subjects	Healthy subjects
L1	89 (33, 37%)[a]	89 (37, 42%)
L2	47 (33, 70%)	47 (36, 77%)
L3	38 (18, 47%)	38 (22, 58%)

a. ( ):Number of samples containing heart sounds and its ratio

B. Manual Labeling

Manual labeling for the acoustic segments was performed, based on the acoustic and segmental features, such as adventitious sounds and heart sounds. A respiratory phase segment (period) comprised several successive acoustic segments, and one abnormal respiratory phase comprised breath, adventitious-sound, and heart-sound segments. There was also one normal respiratory period, which comprised both breath and heart-sound segments. A continuous or discontinuous sound segment was used to represent each adventitious sound. Typical examples of discontinuous sound segments are coarse crackles, fine crackles, and pleural friction rubs. Rhonchus or wheezing sounds are both examples of continuous segments [1]. With regard to discontinuous adventitious sounds, a sequence of instantaneous “snap” sounds, over a short duration, was labeled as a single discontinuous sound segment.

Each first heart sound (S1) was also labeled as one segment. When an adventitious sound and a heart sound appeared at same time, we labeled this acoustic period as an “adventitious-sound segment”. Fig. 2. shows a spectrogram and label sequence of a typical lung sound taken from an unhealthy subject in our database.

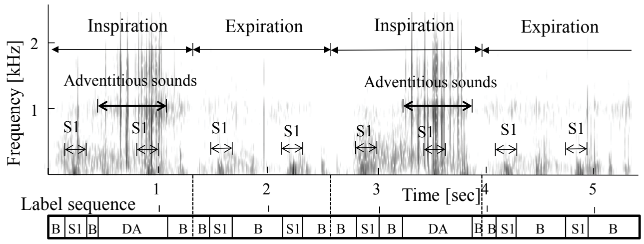

Fig. 2. A spectrogram and a label sequence of a typical lung sound from an unhealthy subject. *(B: breath segment, S1:first heart-sound segment, DA: discontinuous adventitious-sound segment)*

C. Adventitious Sounds and Heart Sounds

In our proposed classification method, we took into account the duration of both adventitious-sound and heart-sound segments. Table II shows both the mean value and standard deviation for each type of sound segment. It also shows that the duration of heart-sound segments was much shorter than those of adventitious sounds. Compared to the heart sounds, there was a considerable variation in the duration of each adventitious sound, because of the diversity between the continuous and discontinuous adventitious sounds. The time interval between two consecutive S1 sounds, from the end of one S1 to the begging of the next S1, was also used in the classification. This time interval is also shown in Table II. These values were calculated using lung-sound samples in which all occurrences of heart sounds were observed.

In our strategy, the choice of which classification method to use, from the two methods available, was based on the estimated detection rate of heart sounds. The previously mentioned temporal information was used to estimate the detection rate precisely.

TABLE II. DURATION OF HERT SOUNDS AND ADVENTITIOS SOUNDS AND TIME INTERVAL OF HEART SOUNDS [S]

	Duration		Interval
Sound	*Adventitious*	*Heart sound (S1)*	*Heart sound (S1)*
Mean	0.53	0.12	0.61
S.D.	0.31	0.03	0.11

III. Classification Considering the Observation Rate of Heart Sounds

A. Classification Flow

Initially, the input lung-sound sample was segmented using acoustic HMMs (breath-sound, continuous and discontinuous adventitious-sound, and heart-sound models), from which the most likely acoustic-segment sequence was derived. This sequence usually contains misrecognized segments, a typical example being the confusion between heart sounds and adventitious sounds, caused because of the similar spectral features of these sounds. In the proposed method, two characteristics of temporal information were used: significant differences between the duration of heart and adventitious sounds, and the periodicity of heart sounds. The first characteristic was also used to calculate the acoustic likelihood for abnormal respiration.

Next, the classification method was chosen based on the value of the detection rate. If the detection rate was lower than the predefined threshold, HMM-based classification was used without the heart-sound models. Conversely, if the detection rate was higher, the likelihood for a normal respiration candidate and that for an abnormal candidate were both obtained in the segmentation process and were used to obtain the recognition result.

B. Detection of Heart-Sound Sequence

The most likely acoustic segment sequence was obtained by using the Viterbi algorithm, based on HMMs. This process is the same as the process of likelihood calculations using HMMs, including heart-sound models, as described in Section III.C. The i-th segment s_i in the sequence W had an acoustic label and temporal information (start time $B(s_i)$ and end time $E(s_i)$). Next, heart-sound segments were extracted from the sequence, and from this extraction, heart-sound segments satisfying the following constraints, concerning duration, were extracted;

$$\mu_H - 3\sigma_H \leq E(s_i) - B(s_i) \leq \mu_H + 3\sigma_H \quad (2)$$

where μ_H and σ_H are the mean value and standard deviation of the duration of the heart sound, respectively (shown in Table II). This constraint indicates that the segment mistakenly classified as a heart sound was excluded by taking into account the duration of the heart-sound segment.

Some of the extracted heart-sound segments were contaminated with short noises, that were mistakenly recognized. Heart sounds occur periodically, but short noises appear randomly. Then, for each extracted heart-sound segment, the number of other heart sounds that occurred periodically was counted using the mean and standard deviation of the time interval between neighboring heart-sound segments (Table II). This second constraint indicates that short noise segments were excluded, by taking into account the significant periodicity of heart sounds.

Finally, the maximum number M of detected heart-sound segments in the respiratory sample was obtained. In our study the detection rate R of heart sounds is defined as follows.

$$R = {}^{M}/_{N} \quad (3)$$

where N is the estimated number of heart-sound segments when all segments of heart sounds were accurately observed. This number was obtained using the time length of the sample and the mean value of the interval of heart sounds.

C. Likelihood Calculation for Normal/Abnormal Respiration Candidate

Our classification method comprised a training process and a test process. In the training process, two sets of both acoustic HMMs and an acoustic segment bigram [10], were generated using the prepared labels of acoustic segments. One set considered the contamination of heart sounds, and the other ignored this contamination.

In the test process, the likelihood of a normal/abnormal respiratory phase was calculated, and also the occurrence probability $P(W)$ of the segment sequence W was calculated using the segment bigram. The total likelihood was composed of the acoustic likelihood, calculated from the HMMs, and the segment sequence likelihood. The segment sequence $\widehat{W}$ with the highest likelihood $\log P(\widehat{W}|X)$ for an unknown respiratory input X is expressed using Bayes' theorem, as follows:

$$\begin{aligned}\widehat{W} &= \mathrm{argmax}_W[\log P(W|X)] \\ &\approx \mathrm{argmax}_W[\log P(X|W) + \alpha \log P(W)] \quad (4)\end{aligned}$$

where $\log P(X|W)$ is the acoustic likelihood. The weight factor α controlled the contribution of the segment bigram, and this factor was obtained experimentally.

D. Likelihood Adjustment

When a heart-sound model was used for classification, the likelihood of the recognized candidate was adjusted based on the duration of each recognized segment, taking into account misrecognition between heart sound and adventitious sound [12]. A normal distribution model was adopted to describe the distribution for heart-sound segment duration and adventitious-sound segment duration. The crossing point of the two probability density functions (pdf), for the two types of sound segments, was set as the threshold *Td*. If the duration of the measured adventitious-sound segment (that occurred at the timing of a heart sound) was shorter than the threshold, it was assumed that the probability of misrecognizing the heart sound as adventitious sound was high, in which case the likelihood for the abnormal candidate was adjusted (decreased). The validity score *V*, of the duration, was defined by using the pdf $f(x)$ (where x indicates duration) of the heart-sound duration, and the pdf $g(x)$ of the adventitious sound was calculated as follows:

ISBN 978-90-827970-1-5 © EURASIP 2018

$$V = \sum_{j=1}^{J} \begin{cases} \log f(x_j) - \log g(x_j) & \text{if } x_j < Td \\ 0 & \text{if } x_j \geq Td \end{cases} \quad (5)$$

where x_j is the onset time of the j-th adventitious-sound segment $(1 \leq j \leq J)$ detected in the segmentation process. Equation (5) indicated that if the probability of misrecognition of a heart-sound as an adventitious-sound segment, was high, from the viewpoint of segment duration, then the validity score V became large.

E. Criteria for the Detection of Abnormal Respiratory Phase

For each respiratory period X in the test sample, the likelihood $\log P(\widehat{W}^{no}|X)$ for a normal respiratory candidate $\widehat{W}^{no}$, and the likelihood $\log P(\widehat{W}^{ab}|X)$ for an abnormal respiratory candidate $\widehat{W}^{ab}$, a sequence of acoustic segments (including adventitious-sound segments) were calculated. For the classification that did not use the heart-sound HMM, and when the likelihood for the abnormal candidate was larger than that of the normal candidate, we applied the expression:

$$\log P(\widehat{W}^{ab}|X) + \beta \geq \log P(\widehat{W}^{no}|X) \quad (6)$$

where the test respiratory X was considered abnormal respiration. For the classification that did use the heart-sound HMM, we applied the expression:

$$\log P(\widehat{W}^{ab}|X) - \gamma V + \beta' \geq \log P(\widehat{W}^{no}|X), \quad (7)$$

where γ is a weight factor for the validity score of duration, and β, β' are offset values.

F. Criteria for the Detection of an Unhealthy Subject

We defined two criteria for distinguishing between unhealthy and healthy subjects, as follows. A subject was considered to be unhealthy when the total likelihood of an abnormal respiration candidate for each respiratory period exceeded that of a candidate with normal respiration [7]. For the classification when not using the heart-sound HMM, the expression used was:

$$\sum_{i=1}^{I} \log P(\widehat{W}_i^{ab}|X_i) + \beta'' \geq \sum_{i=1}^{I} \log P(\widehat{W}_i^{no}|X_i) \quad (8)$$

where I is the number of the respiratory phase segments in the input sample of lung sound. For the classification using the heart sound HMM, the expression was defined as follows:

$$\sum_{i=1}^{I} (\log P(\widehat{W}_i^{ab}|X_i) - \gamma V) + \beta''' \geq \sum_{i=1}^{I} \log P(\widehat{W}_i^{no}|X_i) \quad (9)$$

where β'' and β''' are offset values. These values were obtained experimentally, to yield the highest performance for each method.

IV. EALUATION EXPERIMENTS

To confirm the effectiveness of the proposed method, three classification methods were examined: Method I) classification not using a heart-sound HMM, Method II) classification using heart-sound HMM where a likelihood adjustment was performed, and Method III) proposed classification method that selects either Method I or Method II, depending on the detection rate of heart sound.

A. 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 modelling the HMMs [11, 12]. In the experiments, two sets of HMMs were generated, with one set including the heart-sound model and the other set excluding the heart-sound model. 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 were used to train the model for the breath segment of normal respiration, whereas the models for adventitious sound were generated by using the sounds obtained from unhealthy subjects. The models for heart sounds were developed using both normal respiration and abnormal respiration. The HMMs with three states (left-to-right topology) and two Gaussian probability density functions were used. Two segment bigram models were also trained by using the segment labels of the training samples, both with and without heart-sound segments.

In our experiments, we assumed that the number of respiratory periods for each lung sound sample and the respiratory boundaries were both known. We also performed a leave-one-out cross validation for each test sample. In addition, our experiments were subject-independent, because the samples recorded from a subject who was used as the test sample were excluded from the training process.

We set the threshold *Th* of the detection rate of heart sounds 0.65 in Method III. If $R < Th$ for each lung sound sample (Equation (3)), Method I was performed. Otherwise $(R \geq Th)$, Method II was carried out.

B. Classification of Normal and Abnormal Respirations

Three classification methods, which distinguish between abnormal and normal respirations, were examined. The number of respiratory phase segments used for evaluation is shown in Table III, which were randomly selected from the recorded samples (Table I). The classification results for each auscultation point and each method are shown in Table IV. The average classification rate, weighted with the data amount, is represented by the "Average."

The average classification rate of the proposed method (89.9%) was the highest among the three methods, which validated the proposed selection method. Concerning the sound samples recorded from L1, this observation rate of heart sounds was lowest from the three auscultation points. For this data set, the proposed method achieved the highest performance, which is equivalent to that of Method I not using heart-sound models. On the other hand, the observation rate for the samples recorded from L3 was highest. For these samples, the highest performance was obtained by using the proposed method.

ISBN 978-90-827970-1-5 © EURASIP 2018

TABLE III. NUMBER OF RESPIRATORY PHASE SEGMENTS FOR EVALUATION

Point	Normal Segments	Abnormal Segments
L1	161	161
L2	217	217
L3	206	206

TABLE IV. CLASSIFICATION PERFORMANCE BETWEEN NORMAL AND ABNORMAL RESPIRATION [%]

Point	Method	Abnormal	Normal	Average
L1	I	93.2	89.4	91.3
	II	87.0	87.6	87.3
	III	93.2	89.4	91.3
L2	I	90.3	86.2	88.2
	II	78.8	98.2	88.5
	III	88.0	89.8	88.9
L3	I	94.2	75.7	85.0
	II	87.4	92.7	90.0
	III	87.4	92.7	90.0
Average	I	92.5	83.4	87.9
	II	84.1	93.3	88.7
	III	89.2	90.6	89.9

Threshold of detection rate *Th*: 0.65

TABLE V. CLASSIFICATION PERFORMANCE BETWEEN UNHEALTY AND HEALTHY SUBJECTS [%]

Points	Method	Unhealthy	Healthy	Average
L1	I	83	85	84.3
	II	71	85	78.1
	III	83	85	84.3
L2	I	92	83	87.2
	II	79	98	88.3
	III	89	87	88.3
L3	I	95	82	88.5
	II	87	92	89.7
	III	87	92	89.7
Average	I	88.0	84.0	86.0
	II	76.0	90.3	83.1
	III	85.7	87.4	86.6

Threshold of detection rate *Th*: 0.65

C. Classification of Unhealthy and Healthy Subjects

Finally, unhealthy and healthy subjects were classified for each auscultation point, using three classification methods. The classification performance is shown in Table V. The tendency of the classification performance was almost the same as that of the classification between normal and abnormal respiration. For each auscultation point, the classification performance of the proposed method was superior or equivalent to those of Methods I and II. The results shown in Tables IV and V indicate that the selection of classification method, with or without heart-sound modeling, yielded better results for the classification between unhealthy and healthy subjects, as well as between normal and abnormal respiration.

V. CONCLUSIONS

This paper proposed an effective classification method for distinguishing between normal and abnormal respiration, regardless of the contamination level of lung sounds by the heart sounds. According to the estimated detection rate of heart sounds, one adequate classification method was selected out of two methods: HMM-based classification using the heart-sound model, and one not using the heart-sound model. To estimate the detection rate of heart sounds accurately, temporal information of both heart sounds and adventitious sounds were used. Specifically, the significant difference between the duration of heart and adventitious sounds, and the periodicity of heart sounds, were used. Evaluation experiments indicated that the proposed method yielded better results for the classification between unhealthy and healthy subjects, as well as between normal and abnormal respiration, compared to the sole use of the HMM-based classification method, whether using the heart-sound model or not.

Our proposed classification adopted a deterministic selection out of two classification methods; however, a more adaptive technique corresponding to the detection rate of heart sounds should be devised to greatly improve classification performance. This is a subject area for future investigation.

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] A. Fuchita and F. Kondo, “Detection algorithm for continuous adventitious lung sounds,” *Proc. of Lung Sounds Research Conference,* vol. 11, pp.25-31, 2007

[7] C. Daniel, et al., “Application of semi-supervised deep learning to lung sound analysis,” *Proc. of IEEE EMBC,* WeCT 1.32, 2016.

[8] M. Aykanat, O. Kilig, B. Kurt and S. Saryal. “Classificatio of lung sounds using convolutional neural networks,” *EURASIP Journal on Image and Video Processing,* 2017

[9] S. Matsunaga, et al., “Classification between normal and abnormal respiratory sounds based on maximum likelihood approach,” *Proc. of IEEE ICASSP*, pp. 517-520, 2009.

[10] H. Yamamoto, et al., “Classification between normal and abnormal respiratory sounds based on stochastic approach,” *Proc. of ICA*, 2010.

[11] M. Himeshima, et al., “Detection of abnormal lung sounds taking into account duration distribution for adventitious sounds,” *Proc. of EUSIPCO*, pp.1821-1825, 2012.

[12] M. Taguchi, M. Yamashita, and S. Matsunaga, “Detection of abnormal lung sounds considering spectral and temporal features of heart sounds,” *Proc. of ICA*, 2016.