IEICE TRANS. INF. & SYST., VOL.E107–D, NO.4 APRIL 2024

PAPER

# Construction of Ergodic GMM-HMMs for Classification between Healthy Individuals and Patients Suffering from Pulmonary Disease

**Masaru YAMASHITA**[†a)], ***Member***

**SUMMARY** Owing to the several cases wherein abnormal sounds, called adventitious sounds, are included in the lung sounds of a patient suffering from pulmonary disease, the objective of this study was to automatically detect abnormal sounds from auscultatory sounds. To this end, we expressed the acoustic features of the normal lung sounds of healthy people and abnormal lung sounds of patients using Gaussian mixture model (GMM)-hidden Markov models (HMMs), and distinguished between normal and abnormal lung sounds. In our previous study, we constructed left-to-right GMM-HMMs with a limited number of states. Because we expressed abnormal sounds that occur intermittently and repeatedly using limited states, the GMM-HMMs could not express the acoustic features of abnormal sounds. Furthermore, because the analysis frame length and intervals were long, the GMM-HMMs could not express the acoustic features of short time segments, such as heart sounds. Therefore, the classification rate of normal and abnormal respiration was low (86.60%). In this study, we propose the construction of ergodic GMM-HMMs with a repetitive structure for intermittent sounds. Furthermore, we considered a suitable frame length and frame interval to analyze acoustic features. Using the ergodic GMM-HMM, which can express the acoustic features of abnormal sounds and heart sounds that occur repeatedly in detail, the classification rate increased (89.34%). The results obtained in this study demonstrated the effectiveness of the proposed method.

***key words:*** *hidden Markov model, lung sound, patient detection, abnormal respiration*

## 1. Introduction

Auscultation of the lungs is used for detecting patients with pulmonary diseases. Despite other noninvasive and inexpensive methods, auscultation using a stethoscope can obtain valuable information regarding the health status of an individual. In several cases, abnormal sounds (called adventitious sounds [1]) are included in the lung sounds of patients with pulmonary disease, and auscultation is currently an effective method for diagnosing pulmonary disease. However, this method requires expert knowledge and expertise. Therefore, identifying the difference between healthy people and patients is difficult for non-medical personnel, and this may be the reason auscultation is not used in common households. Furthermore, it is difficult for the elderly or individuals living in depopulated areas to visit hospitals. Thus, the distinction between healthy individuals and patients performed at home can facilitate early detection of pulmonary diseases. Several studies have focused on automatically detecting adventitious sounds from lung sounds [2]–[4]. These studies either detected a specific adventitious sound using a wavelet transform or distinguished the frame of an adventitious sound using a short-time spectrum. However, the time of occurrence and duration of adventitious sounds vary. Therefore, it is desirable to discriminate sounds using the features of the entire respiration process and its inflection. Furthermore, the features of adventitious and respiratory sounds depend on the individual and the progression of the disease. Therefore, we believe that these features should be expressed statistically. Recently, convolutional neural networks [5]–[7] and recurrent neural networks [8], [9] have been used to analyze lung sounds. Furthermore, in the field of speech recognition, end-to-end models becoming the prominent approach [10]–[12]. However, these methods require a large volume of training data to achieve good performance. To overcome these issues, in our previous studies, the features should be expressed statistically. The time series of the acoustic features of lung sounds are expressed by constructing Gaussian mixture model (GMM)-hidden Markov models (HMMs) to discriminate between normal and abnormal respiratory sounds [13]–[18]. However, we did not consider the suitable state transition of the GMM-HMMs, analysis frame length, and frame intervals.

Adventitious sounds are divided into two classes: continuous and discontinuous. Figure 1 shows the respirations including discontinuous adventitious sounds, which are called fine crackles. Figure 2 shows the respirations including continuous adventitious sounds, which are called wheezes. A distinctive feature of discontinuous adventitious sounds is that short sounds occur repeatedly. Although the acoustic features of adventitious sounds differ by type, we considered the discontinuous adventitious sound period to be a steady state and expressed it using a left-to-right GMM-HMM. Therefore, the classification rates of normal and abnormal respiration were low. Furthermore, in auscultation, noise hinders the detection of adventitious objects with high accuracy. Auscultatory sounds often include noise from the body and the rustle of the stethoscope. A typical noise from the body is the sound of the heart. Figure 3 shows examples of respiratory sounds, including adventitious sounds, heart sounds (S1 and S2), and other noise. The frequency of the appearance of heart sounds auscultated near the heart is high. The database used in our study included several heart sounds; consequently, several normal respiratory sounds were identified as abnormal. To distinguish adventitious sounds from heart sounds, we constructed a heart sound model using

Manuscript received March 2, 2023.
Manuscript revised August 4, 2023.
Manuscript publicized December 12, 2023.

[†]The author is with the Graduate School of Engineering, Nagasaki University, Nagasaki-shi, 852–8521 Japan.
a) E-mail: masaru@nagasaki-u.ac.jp

DOI: 10.1587/transinf.2023EDP7040

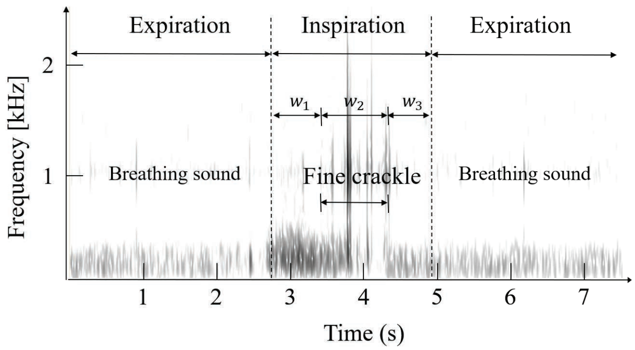



**Fig. 1** Respirations including discontinuous adventitious sounds [19], [20].

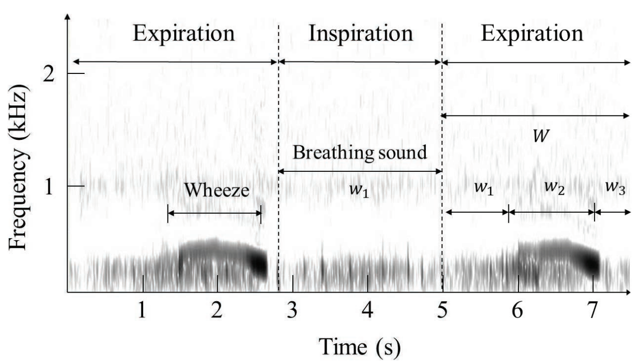



**Fig. 2** Respirations including continuous adventitious sounds [19], [20].

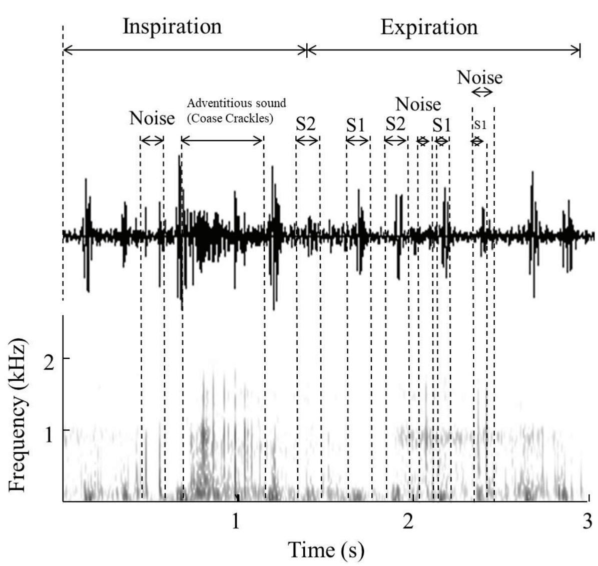



**Fig. 3** Respiratory sounds, including adventitious sounds, heart sounds, and other noise [16].

heart sounds for model learning [16], [17]. As a result, normal respiratory sounds were correctly identified. However, the accuracy decreased in the case of abnormal respiratory sounds. We assumed that these models were unsuitable.

Therefore, we focused on analyzing the topology of the acoustic models and the frame lengths of adventitious and heart sounds. In our previous study [18], we constructed an ergodic GMM-HMM for discontinuous adventitious sounds. In this study, we apply the construction method to respirations, including continuous adventitious and heart sounds, and set a suitable analysis frame length and appropriate frame intervals. We then examine the combination of ergodic GMM-HMMs for each sound. As a result, we confirmed that the construction of the ergodic GMM-HMM for respiration, including short sounds, is suitable for the detection of abnormal respiration in patients.

## 2. Lung Sound Database

### 2.1 Dataset

The lung sounds were recorded by a medical doctor using an electronic stethoscope. They were recorded in WAVE format, sampled at 5 kHz, and quantized at 16 bits. The doctor judged the recording points for each subject. Therefore, the number of recording points differed between subjects. The respiratory count was 5 breaths, and the average of recorded time was 15.3 s. The medical doctor provided diagnoses and classified the subjects as healthy and patients. As the result, the data included 134 healthy subjects and 109 patients.

### 2.2 Hand Labeling

We manually performed segmentation based on recorded sounds, waveforms, spectrograms, and power. First, lung sounds were divided into inspiration and expiration sound segments (respiratory sound segments). Next, the respiratory sound segments were divided into adventitious and other breathing sound segments. The adventitious sound segments were classified into continuous adventitious and discontinuous sounds. Additionally, we marked the heart sound segments on the lung sounds recorded from auscultation points near the heart. Because the first (S1) and second sounds (S2) can be clearly observed, we marked them as heart sounds. If the occurrence interval of adventitious sounds and heart sounds was shorter than 100 ms, it was considered as one segment.

### 2.3 Definition of Normal and Abnormal Respiration

The acoustic features of some noises were similar to those of the adventitious sounds. Some respiratory sounds from healthy individuals included adventitious sounds. Therefore, it is difficult for a nonmedical person to diagnose this condition. Conversely, some of the respiratory sounds from the patient did not include adventitious sounds. However, they cannot be referred to as normal respiratory sounds. Respiratory sounds were grouped into four categories and defined as normal and abnormal respiration as follows:

- Abnormal respiration by patients (AP): respirations that include adventitious sounds.
- Abnormal respiration by healthy individuals (AH): respirations that include noises resembling adventitious

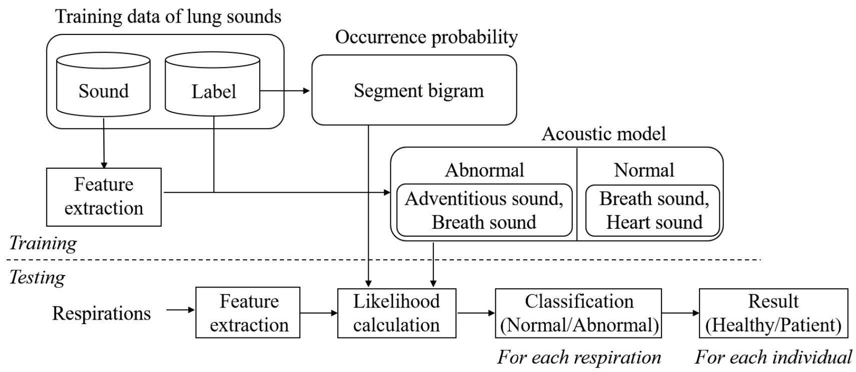



**Fig. 4** Architecture of the classification system for normal and abnormal respiration.

**Table 1** Number of respiratory sounds of each category.

| Category | Number |
|---|---|
| AP | 2135 |
| AH | 899 |
| NP | 3035 |
| NH | 4271 |

sounds.

- Normal respiration by patients (NP): respirations that do not include adventitious sounds or noises resembling adventitious sounds from patients.
- Normal respiration by healthy individuals (NH): respirations that do not include adventitious sounds or noises resembling adventitious sounds from patients.

Table1 shows the number of respirations of each category. In this study, we performed two types of experiments: detection of abnormal respiration and detection of patients. In the detection experiment of abnormal respiration and the training data for normal and abnormal respiration, we used only NH as normal respiration and AP as abnormal respiration. That is, AH and NP were not used for the experiment or training data. However, in the detection experiment of patients, all respirations were used. We used a series of respirations for each test individual. That is, we used AP, AH, NP, and NH for the detection experiment of the patient; however, we used only NH and AP for the training data. The classification procedure for the two experiments is described in Sect. 3.

## 3. Fundamental Classification Procedure

### 3.1 Detection of Abnormal Respiration

Generally, in the field of speech recognition, acoustic models of phonemes (the smallest unit of speech) and the occurrence probability of words are used to construct stochastic models. We applied this technique to the lung sounds. Figure 4 shows the architecture of the classification system for normal and abnormal respiration [16]. It comprises training and testing processes. In the training process, the GMM-HMMs are trained as the acoustic and segment sequence models, which define the occurrence probability of the divided segments. In the test process, the input respiration is classified as normal or abnormal based on the maximum likelihood approach. If we assume that sample respiration $W$ consists of $N$ segments, it can be expressed as $W = w_1\ w_2 \cdots\ w_i \cdots\ w_N$, where $w_i$ is the $i$-th segment of $W$.

The training process was as follows. First, we extracted acoustic features and trained each segment. In the case of normal respiration, if we assume that it does not include heart sounds, it consists of one segment ($N = 1$). If it includes heart sounds, it consists of at least two segments ($N \geq 2$). Conversely, abnormal respiration, including adventitious sounds, consists of at least two segments ($N \geq 2$) even if it does not include heart sounds. For example, the case of expiration shown in Fig. 2 consists of one wheeze segment and two breathing segments ($N = 3$). The inspiration shown in Fig. 2, which does not include adventitious sounds, consists of one breathing sound segment ($N = 1$). The training of the segment sequence model can be explained as follows: We calculated the occurrence probability of segments $P(W)$ by using a segment bigram. $P(W)$ can be expressed as

$$P(W) = P(w_1) \times \Pi_{i=2}^{N} P(w_i|w_{i-1}) \tag{1}$$

Let $P(w_i|w_{i-1})$ be defined as

$$P(w_i|w_{i-1}) = C(w_{i-1}, w_i)/C(w_{i-1}), \tag{2}$$

where $C(w_{i-1})$ is the number of $w_{i-1}$, and $C(w_i,w_{i-1})$ is the number of the segment $w_i$ after $w_{i-1}$ in the training database.

The test process can be explained as follows: The maximum likelihood among the calculated likelihoods was determined, and the corresponding segment sequence $\hat{W}$ was selected to recognize the sample respiration sound. If the sequence included at least one adventitious sound, the sample respiration was identified as an abnormal sound. If not, the sample respiration was identified as a normal sound. $\hat{W}$ can be expressed as

$$\hat{W} = \arg\max_{W}(\log P(X|W) + \alpha \log P(W)) \tag{3}$$

where $X$ is the sample respiration and $\log P(X|W)$ is the acoustic likelihood. The weight factor $\alpha$ was experimentally obtained.

### 3.2 Detection of Patient Suffering from Pulmonary Disease

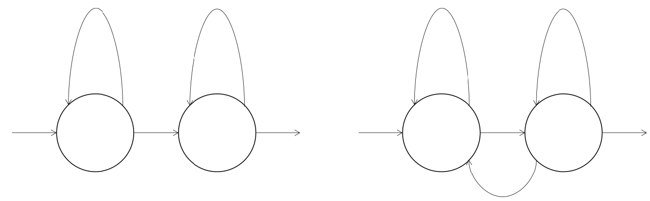
This section describes the detection of the patients suffering from pulmonary disease. Noise from the outside of the body occurs irregularly. In contrast, adventitious sounds occur periodically. Therefore, in the case of healthy individuals, most of the likelihood values for normal respiration are higher than those for abnormal respiration, even if some respirations are classified as abnormal. To detect patients, we calculated the likelihood $L(W_{no})$ for the segment sequence $\hat{W}_{no}$ that does not include adventitious sounds, and the maximum likelihood $L(W_{ab})$ for the segment sequence $\hat{W}_{ab}$ that includes adventitious sound segments for each respiration. If the total of $L(W_{ab})$ was greater than or equal to the total of $L(W_{no})$, then the individual was classified as a patient. That is, $\Sigma_j L(W_{j,ab}) \geq \Sigma_j L(W_{j,no})$, where $L(W_{j,ab})$ is the likelihood of the segment sequence that includes adventitious sound segments for the $j$-th respiration of the individual and $L(W_{j,no})$ is the likelihood of the segment sequence that does not include adventitious sound segments for the $j$-th respiration of the individual.

### 3.3 Classification Procedure Using the Heart-Sound Model

To distinguish adventitious sounds from heart sounds, we constructed a heart sound model in addition to the breathing sound and adventitious models [16], [17]. We trained the acoustic models during the training process. In the case of normal respiration sounds, we trained the normal sound model using breathing and heart sound segments. In the case of abnormal sounds, we trained the model as described in the fundamental classification procedure. In the test process, the maximum likelihood among the calculated likelihoods was determined, and the corresponding segment sequence $\hat{W}$ was selected to recognize the sample respiration sound, similar to the fundamental classification procedure. The difference from the fundamental classification procedure was that even if the sequence included heart sounds, the sample respiration was identified as normal respiration.

## 4. Construction of Ergodic GMM-HMMs

In our previous studies [13]–[18], we constructed a left-to-right GMM-HMM with limited states for each segment, as shown in Fig. 5 (a), and assumed that the models were not suitable. Therefore, we focused on analyzing the topology of the acoustic models. For example, the duration of the stationary sound period was significantly different between heart and adventitious sounds. Table 2 shows the mean and standard deviation (SD) of the adventitious and heart sound durations. The duration of heart sounds was shorter than that of adventitious sounds, and it was too small to analyze using a long frame length and long frame intervals. Therefore, we focused on the adventitious and heart sounds for the model. For example, to construct an GMM-HMM that is appropriate for discontinuous adventitious sounds, we assumed that the discontinuous sound period consists of the repetition of an abnormal sound period and a silent period. We then constructed an ergodic GMM-HMM that transitions to the former state, as shown in Fig. 5 (b). Furthermore, to construct the ergodic GMM-HMM, we selected the suitable analysis frame length and frame intervals because an abnormal sound period and a silent period are too short to analyze through the typical values used in the frequency analysis of speech.



(a) Left-to-right GMM-HMM (b) Ergodic GMM-HMM

**Fig. 5** Topology of GMM-HMMs [21].

**Table 2** Mean and standard deviation (SD) of duration for adventitious and heart sounds [16], [20].

| Source sound | Mean (s) | SD (s) |
|---|---|---|
| Adventitious sound | 0.53 | 0.31 |
| Heart sound | 0.12 | 0.03 |

## 5. Classification Experiments

### 5.1 Experimental Conditions

Every 10 ms, six mel-frequency cepstral coefficients (MFCCs) and power values were extracted as acoustic features using a 25 ms Hamming window. Figure 6 shows the auscultation points. In this study, a heart sound model was used to auscultate lung sounds from three points near the heart ($P_4$-$P_6$). Conversely, a heart sound model was not used to auscultate lung sounds from six points away from the heart ($P_1$-$P_3$, $P_7$-$P_9$). Table 3 lists the number of abnormal respiratory sounds that included adventitious sounds and the number of patients that included at least one adventitious sound; as many normal respirations or healthy individuals were randomly selected for each detection experiment of abnormal respiration and patients. The number of each sound segment is listed in Table 4. These were used as the training data for the GMM-HMM. The respiration data recorded at the same point as the testing data were used for training. Thus, we constructed a point-dependent model. We performed leave-one-out cross-validation to construct an individual-independent model. In other words, we did not use respiration data from the same or another recording point of the same individual.

### 5.2 Classification Experiments between Normal and Abnormal Respirations

We compared the left-to-right and ergodic GMM-HMMs

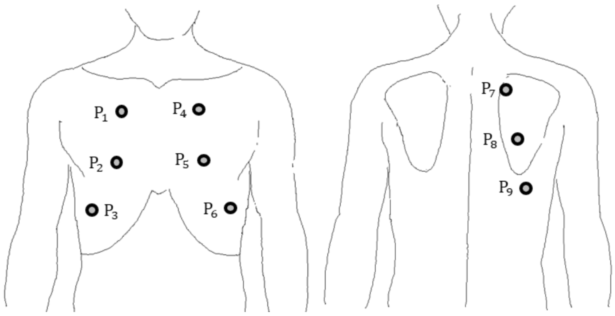



**Fig. 6** Auscultation points [21].

**Table 3** Number of abnormal respiratory sounds and patients [21].

| Points | No. Abnormal respirations | No. Patients |
|---|---|---|
| $P_1$ | 219 | 44 |
| $P_2$ | 161 | 89 |
| $P_3$ | 254 | 53 |
| $P_4$ | 217 | 47 |
| $P_5$ | 312 | 62 |
| $P_6$ | 206 | 52 |
| $P_7$ | 182 | 46 |
| $P_8$ | 329 | 62 |
| $P_9$ | 260 | 62 |
| Total | 2135 | 517 |

**Table 4** Number of each sound segment [20].

| Points | No. of Sound segments |
|---|---|
| Heart sound | 4949 |
| Discontinuous adventitious sound | 1753 |
| Continuous adventitious sound | 397 |
| Breathing sound | 4285 |
| Normal respiration | 2135 |

**Table 5** Combinations of frame length and frame interval [21].

| Conditions | Frame length | Frame interval |
|---|---|---|
| A | 5 | 2 |
| B | 10 | 4 |
| C | 15 | 6 |
| D | 20 | 8 |
| E | 25 | 10 |
| F | 30 | 12 |

and selected the frame length and frame intervals for analysis. In our previous studies [13]–[18], we set the analysis frame length to 25 ms and frame interval to 10 ms. In this study, we selected several combinations of frame length and frame interval, as presented in Table 5. Figure 7 shows the classification rate between normal and abnormal respirations using a left-to-right GMM-HMM and ergodic GMM-HMM for heart sounds. When the analysis frame length and interval were too small, the classification rate decreased because the frequency resolution was low. First, we constructed an ergodic GMM-HMM for heart sounds. The models of the other sounds were constructed using a left-to-right model. Figure 8 shows the classification rate between normal and abnormal respiration using the ergodic GMM-HMM for heart sounds. By comparing Figs. 7 and 8, we can observe that the accuracy using the ergodic GMM-HMM for heart sounds is

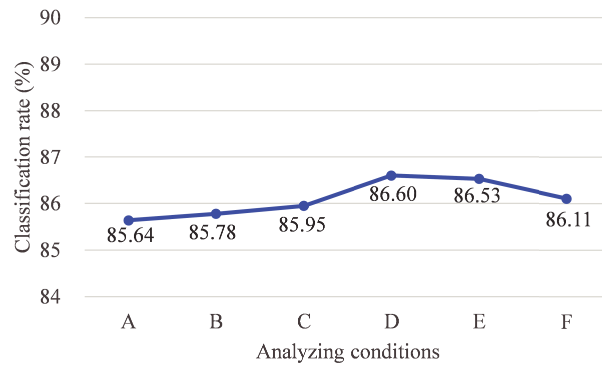



**Fig. 7** Classification rate between normal and abnormal respiration using left-to-right GMM-HMM [21].

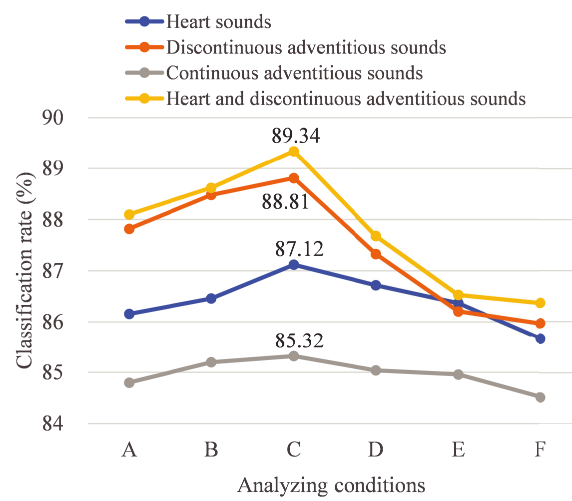



**Fig. 8** Classification rate between normal and abnormal respiration using ergodic GMM-HMM for each sounds.

slightly higher. This is because the ergodic GMM-HMM for approximately half of the respirations, including heart sounds, used the transition to the former state for the calculation of likelihood, and the model was valid. However, the ergodic GMM-HMM for the remaining half respirations did not transition to the former state to calculate the likelihood.

Subsequently, we constructed an ergodic GMM-HMM for the discontinuous sounds. The models of the other sounds were constructed using left-to-right models. In both ergodic GMM-HMMs for heart and discontinuous adventitious sounds, the accuracy was higher than that of the left-to-right GMM-HMMs for which the analysis frame length was set to 25 ms and the frame interval was set to 10 ms. This is because the analysis frame length and interval were too large to express the acoustic features of each intermittent sound. We then set the frame length and interval to a smaller value (Condition C), which resulted in an increase in the classification accuracy. When a suitable analysis frame length and appropriate frame interval were set, the ergodic GMM-HMM could express the acoustic features of intermittent sounds. Furthermore, we constructed an ergodic GMM-HMM for continuous sounds. The other sound models were

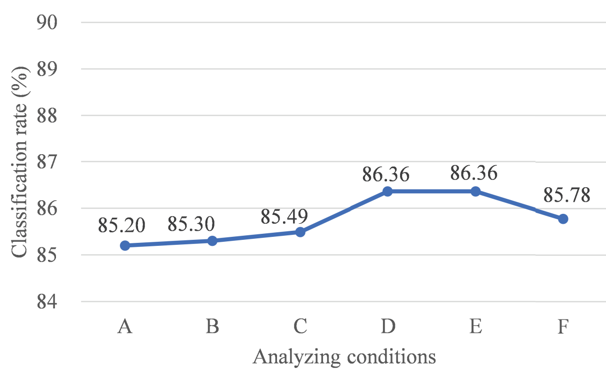



**Fig. 9** Classification rate between healthy individual and patient using left-to-right GMM-HMM.

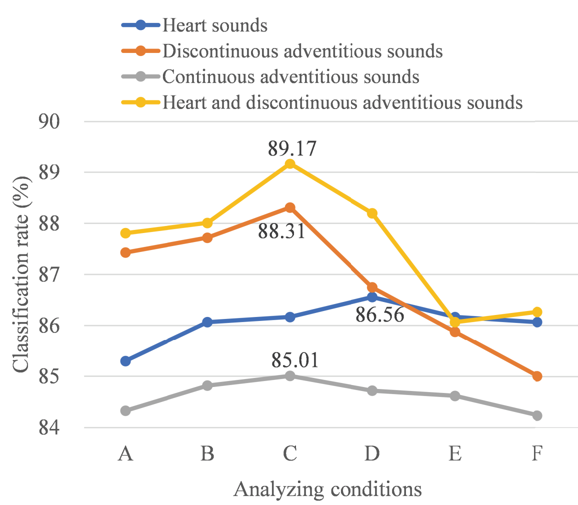



**Fig. 10** Classification rate between healthy individual and patient using ergodic GMM-HMM for each sounds.

constructed using the left-to-right model. The accuracy using the ergodic GMM-HMM for continuous adventitious sounds was lower than that of the left-to-right GMM-HMM under all the conditions. We considered that this is because the amount of training data was smaller than that of discontinuous adventitious sounds, and the acoustic features were constant throughout the continuous adventitious sound periods. We then constructed a combined ergodic GMM-HMM for heart and discontinuous sounds. The models of the other sounds were constructed using the left-to-right model. Accuracy using the ergodic GMM-HMM for both heart and discontinuous adventitious sounds was the highest for the methods mentioned above. We conducted a t-test and the result indicates the significant effectiveness ($p < 0.01$) of constructing an ergodic GMM-HMM for heart and discontinuous adventitious sounds.

### 5.3 Classification Experiments between Healthy Individuals and Patients

Finally, an experiment for the classification of healthy individuals and patients is discussed. Figures 9, 10 show the classification rates of healthy individuals and patients. When we compared the results, we observed that there was the same tendency between normal and abnormal respiration as in the previously mentioned classification experiments. The results obtained by the experiment for classifying healthy individuals and patients also indicated the significant effectiveness ($p = 0.026$) of constructing a combined ergodic GMM-HMM of heart and discontinuous adventitious sounds, wherein the frame length and interval were set to as in Condition C.

## 6. Conclusions

This study proposed the construction of an ergodic GMM-HMM for sounds that occur intermittently and repeatedly, with the aim of classifying normal and abnormal respirations with high accuracy. The ergodic GMM-HMMs for heart and discontinuous sounds were valid, and the accuracy using both models showed an improvement over using the models separately or left-to-right GMM-HMM. Furthermore, to construct an ergodic GMM-HMM with a repetitive structure, we set a suitable analysis frame length and appropriate frame intervals. The results obtained through the classification experiment confirmed that the classification rate improved when the frame length and interval were set slightly smaller than the typical values used in the frequency analysis of speech. We consider that the ergodic model is valid for sounds that occur intermittently and repeatedly. Furthermore, the improvement was also significant in the experiment for classifying healthy individuals and patients. Thus, the effectiveness of the proposed approach was demonstrated.

In future work, we will clarify the suitable topology of GMM-HMMs using a deep neural network, which has proven to be effective in the field of speech recognition.

## Acknowledgments

This study was supported by The Telecommunications Advancement Foundation.

### References

[1] N. Gavriely and D.W. Cugell, Breath Sounds Methodology, CRC Press, Boca Raton, FL, USA, 1995.
[2] Y.P. Kahya, S. Yerer, and O. Cerid, "A wavelet-based instrument for detection of crackles in pulmonary sounds," Conf. Proc. IEEE Eng. Med. Biol. Soc. EMBC 2001, Istanbul, Turkey, vol.4, pp.3175–3178, Oct. 2001.
[3] A. Marshall and S. Boussakta, "Signal analysis of medical acoustic sounds with applications to chest medicine," J. Franklin I., vol.344, no.3-4, pp.230–242, 2007.
[4] S.A. Taplidou and L.J. Hadjileontiadis, "Wheeze detection based on time-frequency analysis of breath sounds," Comput. Biol. Med., vol.37, no.8, pp.1073–1083, 2007.
[5] F. Demir, A. Sengur, and V. Bajaj, "Convolutional neural networks based efficient approach for classification of lung diseases," Health Information Science and Systems, vol.8, no.1, 2020.
[6] T. Nguyen and F. Pernkopf, "Lung sound classification using snapshot ensemble of convolutional neural networks," EMBC 2020 (Conf.

Proc. IEEE Eng. Med. Biol. Soc.), pp.760–763, Montreal, Canada, Aug. 2020.
[7] L. Pham, I. McLoughlin, H. Phan, M. Tran, T. Nguyen, and R. Palaniappan, "Robust deep learning framework for predicting respiratory anomalies and diseases," EMBC 2020 (Conf. Proc. IEEE Eng. Med. Biol. Soc.), pp.164–167, Montreal, Canada, Aug. 2020.
[8] E. Messner, M. Fediuk, P. Swatek, S. Scheidl, F.-M. Smolle-Juttner, H. Olschewski, and F. Pernkopf, "Crackle and breathing phase detection in lung sounds with deep bidirectional gated recurrent neural networks," EMBC 2018 (Conf. Proc. IEEE Eng. Med. Biol. Soc.), pp.356–359, Honolulu, Hawaii, July 2018.
[9] D. Perna and A. Tagarelli, "Deep auscultation: Predicting respiratory anomalies and diseases via recurrent neural networks," International Symposium on Computer-Based Medical Systems (CBMS 2019), pp.50–55, Cordova, Spain, June 2019.
[10] S. Karita, N. Chen, T. Hayashi, T. Hori, H. Inaguma, Z. Jiang, M. Someki, N.E.Y. Soplin, R. Yamamoto, X. Wang, S. Watanabe, T. Yoshimura, and W. Zhang, "A comparative study on transformer vs RNN in speech applications," Proc. IEEE Automatic Speech Recognition and Understanding Work shop 2019 (ASRU 2019), pp.449–456, 2019.
[11] P. Liang, A. Bouchard-Côté, D. Klein, and B. Taskar, "An End-to-end discriminative approach to machine translation," Proc. 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics (COLING-ACL 2006), pp.761–768, Sydney, Australia, July 2006.
[12] A. Graves and N. Jaitly, "Towards end-to-end speech recognition with recurrent neural networks," Proc. 31st International Conference on Machine Learning, JMLR, vol.32, Beijing, China, pp.1764–1772, June 2014.
[13] S. Matsunaga, K. Yamauchi, M. Yamashita, and S. Miyahara, "Classification between normal and abnormal respiratory sounds based on maximum likelihood approach," International Conference on Acoustics, Speech and Signal Processing 2009 (ICASSP 2009) (IEEE SPS), Taipei, Taiwan, pp.517–520, 2009.
[14] M. Yamashita, K. Yamauchi, S. Matsunaga, and S. Miyahara, "An examination of constructive method of HMM considering the segmentation of respiratory sound and variety of adventitious sounds in classification between normal and abnormal respiratory sounds," J. IEICE, vol.J96-D, no.9, pp.2070–2077, 2013.
[15] H. Yamamoto, S. Matsunaga, M. Yamashita, K. Yamauchi, and S. Miyahara, "Classification between normal and abnormal respiratory sounds based on stochastic approach," 20th International Congress on Acoustics (ICA 2010), Sydney, Australia, pp.4144–4148, Aug. 2010.
[16] M. Yamashita, M. Himeshima, and S. Matsunaga, "Robust classification between normal and abnormal lung sounds using adventitious-sound and heart-sound models," International Conference on Acoustics, Speech and Signal Processing 2014 (ICASSP 2014) (IEEE SPS), pp.4451–4455, Florence, Italy, May 2014.
[17] M. Taguchi, M. Yamashita, and S. Matsunaga, "Detection of abnormal lung sounds considering spectral and temporal features of heart sounds," 22nd International Congress on Acoustics 2016 (ICA 2016), Buenos Aires, Argentina, pp.1858–1867, Sept. 2016.
[18] M. Yamashita, "Construction of effective HMMs for classification between normal and abnormal respiration," Proc. Asia-Pacific Signal and Information Processing Association Annual Summit and Conference 2020 (APSIPA ASC 2020), pp.914–918, Auckland, New Zealand, Dec. 2020.
[19] M. Yamashita, "Construction of Suitable DNN-HMM for Classification between Normal and Abnormal Respiration," Lecture Notes in Computer Science, vol.12861, pp.609–619, 2021.
[20] M. Yamashita, "Acoustic HMMs to detect abnormal respiration with limited training data," E. IEICE, vol.E106-D, no.3, pp.374–380, 2023.
[21] M. Yamashita, "Classification between normal and abnormal respiration using ergodic HMM for intermittent abnormal sounds," European conference on signal processing 2021 (EUSIPCO 2021), pp.1187–1190, Dublin, Ireland, Aug. 2021.

**Masaru Yamashita** received his B.S., M.S. and Ph.D. degrees in Engineering from Nagasaki University in 2002, 2004 and 2016, respectively. In 2004, he joined Nagasaki University, where he is with the Technical Staff.