Classifying Respiratory Sounds using Electronic Stethoscope

Yongpeng Liu[*†], Yusong Lin[†], Xingjin Zhang[†], Zongmin Wang[†]
[*]Information Engineering School of Zhengzhou University,
[†]Cooperative Innovation Center of Internet Healthcare,
Zhengzhou, China.
{ypliu, yslin, xjzhang, zmwang}@ha.edu.cn

Yang Gao[‡], Guanling Chen[‡], Haoyi Xiong[§]
[‡]Department of Computer Science, University of Massachusetts Lowell, Lowell, MA.
[§]Department of Computer Science, Missouri University of Science and Technology, Rolla, MO.
{yang_gao, guanling_chen}@cs.uml.edu, xiongha@mst.edu

***Abstract*—In this paper, we develop a computer-based solution for automatic analysis of respiratory sounds captured using the stethoscope, which has many potential applications including telemedicine and self-screening. Three types of respiratory sounds (e.g. wheezes, crackles, and normal sounds) were captured from 60 patients by a custom-built prototype device. We extracted 46 features from time, frequency and Cepstral domain from window frames and the optimal features are selected. Then a two-stage pipeline on Gaussian Mixture Model to classify these three respiratory sounds is proposed and the optimal initial parameters of GMM for each sound type are empirically calculated. By comparing with 24 FMCC features, the evaluation results show that all features proposed in this paper improved accuracy by 7.4% for the crackles and 3% for wheeze classification. On average the method for classifying wheezes, crackles and normal sounds achieved the accuracy of 98.4%, which means the models could be used in the real-world situation for the diagnosis of pulmonary diseases.**

Keywords—respiratory sounds classification; GMM; feature extraction; wheezes; crackles

I. INTRODUCTION

Respiratory sounds, also called lung sounds or breath sounds, are one of the most important physiological signals of human body. The sounds are produced internally when breathing and could be auscultated cross the anterior and posterior chest walls with a stethoscope. Traditionally physicians mainly use the stethoscope for diagnose pulmonary diseases. This process subjectively depends on the physician's ability, experience, training levels and auditory perception that differ from one physician to another. To overcome this limitation, computerized respiratory sound analysis [1] works as a reliable tool for many years to diagnose pulmonary disorders and abnormalities.

According to the definition of the ATS (American Thoracic Society), respiratory sounds are divided into two categories – the normal breathing sounds and the adventitious sounds [2] based on their acoustic properties. The adventitious sounds (abnormal sounds) are subdivided into continuous and discontinuous sounds depending on their duration and sound properties. The continuous sounds usually include stationary events that present patterns similar with music and last longer than 250 ms [3], such as the wheezes. The discontinuous sounds are constituted with some transiently explosive sounds (e.g. fine crackles and coarse crackles).

Wheezes and crackles constitute to two key classes of adventitious sounds. The mechanism of producing wheezes relates to narrowing the airway caliber where they could be detected in many diseases, such as mucosal edema, bronchospasm, asthma, Chronic Obstructive Pulmonary Disease (COPD) [4]. Crackles are generated by pressure equalization or elastic stress changing when abnormally the closed airway reopens. Crackles could be detected from the patients with cardiorespiratory disorders, pulmonary fibrosis, pneumonitis, alveolitis [3]. Therefore computerized respiratory sound analysis of wheezes, crackles and normal sounds could help the physicians discover and diagnose patients who suffer from pulmonary diseases.

The goal of our study is to present a method that could be used in telemedicine or self-screening, such as for the elderly to diagnose and analyze respiratory disorders at home using the sounds from the stethoscope. For this purpose, we firstly set up a simple respiratory sound recording device (Fig.1). We collect normal sounds, wheezes and crackles recorded with that device from many older patients. Then we propose to extract sound features that could be used for lung sounds classification and rank the top 6 features using the feature selection methods. Next we design two-stage GMMs (Gaussian Mixture Model) for classifying these three types of respiratory sounds (normal sounds, wheezes, crackles). For the first stage we train two models for recognizing normal sounds and adventitious sounds. Here the adventitious sounds are made up by wheezes and crackles. For the second stage the wheezes and crackles are separated by their own GMM-based classifiers.

The contribution of this paper include: 1) 46 representing features are extracted from 3 different types of respiratory sounds. By comparing with the Mel-Frequency Cepstrum Coefficients (MFCC) features, we find the Chroma features perform best in representing normal sounds, wheezes and crackles; 2) We propose a two-stage pipeline on GMM models to classify 3 types of respiratory sounds. The optimal mixtures of the GMM models related to each type of respiratory sounds are analyzed using the BIC criterion; 3) By contrasting with the 24 MFCC features that have been tested as the optimal feature set [5], the evaluation results of our two-stage GMM-based

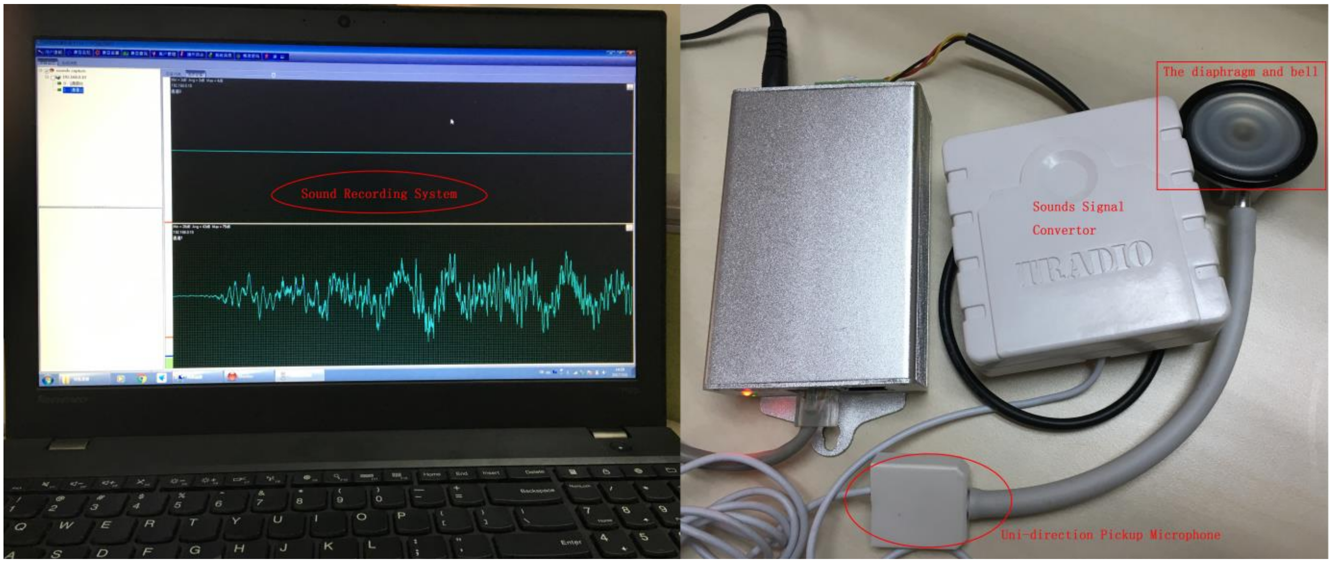

Fig. 1. Respiratory sound capturing system. It consists of a stethoscope, Uni-direction Pickup Microphone, an Analog-to-Digital Converter and the storage software on the laptop.

models show that the accuracy of classifying crackles and wheezes improved by 7.4% from 89% to 96.4% with 46 features, and normal or abnormal sounds classification accuracy reached 99%. On average, the method for classifying normal sounds, wheezes and crackles achieved the accuracy of 98.4%.

The rest of the paper is organized as follows. We describe related work in Section II. Feature extraction and GMM models for respiratory sounds classification are described in Section III. We discuss the performance evaluation in Section IV and we conclude in Section V.

II. Related Work

Just as the typical pattern recognition system, respiratory sound analysis is mainly divided into two processes: the feature extraction and the classification. Feature extraction involves the process which transforms the original high dimensional sound signals into lower dimensional vectors. And classification process involves model training and testing.

There has been many works on feature analysis of respiratory sounds. L. Pesu et al [6] used wavelet packet-based method for feature extraction of wheezes, fine crackles and coarse crackles. There the best set of wavelet packet coefficients about the abnormal respiratory sounds is selected by using best basis search algorithm. Feng Jin et al [7] propose a method based on instantaneous frequency and construct a temporal-spectral dominance spectrogram which could classify wheezes at high noise level. Semra Icer [8] applies Welch method on crackle and rhonchus lung sounds to extract power spectral density (PSD) and uses the ratio of max/min to be the feature classification. [9][10] analyzes fine and coarse crackles by Hilbert-Huang spectrum which decomposes the sound into intrinsic mode function and the instantaneous frequency is obtained as the features. Mohammed Bahoura [5] compares features extracted from such as Fast Fourier transform, Linear Predictive Coding, MFCC and Wavelet transform and finds that MFCC performs the best to classify normal sounds and wheezes.

k-Nearest Neighbor (k-NN) algorithms is used on diagnosis respiratory disease in [11]. They use the microphone with a preamplifier to get sounds from some pathological and healthy subjects. With the best model order 6, the performance of the classifier achieved 93.7%. Artificial Neural Network (ANN) model is applied into lung sounds signals in [12] with the wavelet features extracted, which they classify normal, wheeze, crackle, squawk, stridor, and rhonchus, totally six categories. The evaluation results are that the optimal ANN architecture is 19-40-6 and the wavelet of order 8 gives optimal classification efficiency. Similar work is done using the Neural Network and Genetic Algorithm [13] which is based on the feature extraction of 256-point Fourier Power Spectrum Density (PSD). By taking genetic algorithm, 129 features are chosen to be applied to a neural network.

Hidden Markov Models(HMM) for acoustic spectral features is used in [14] [15] for Classification between normal and abnormal respiratory sounds and the results show the accuracy of normal and abnormal respiratory sounds achieves at 84.2% where the abnormal sounds are made up by crackles, rhonchus and wheezes. But the problem of their work is that the wheeze or rhonchus only consisted about 2% among the abnormal data. So it is hard to use their model for real world testing.[16] [17] use Gaussian Mixture Models to evaluate the lung sounds by capturing them from the electronic stethoscopes and demonstrate the ways of electronic stethoscopes could be a promising potential for disease diagnosis although the data quantity is insufficient and the accuracy result is only 52.5% for identifying Asthma disease.

The most similar work with us is [5]. The conclusion from that paper is that GMM/MFCC combination is well-adapted for wheezes and normal sounds classification when comparing with the methods of vector quantization or multilayer

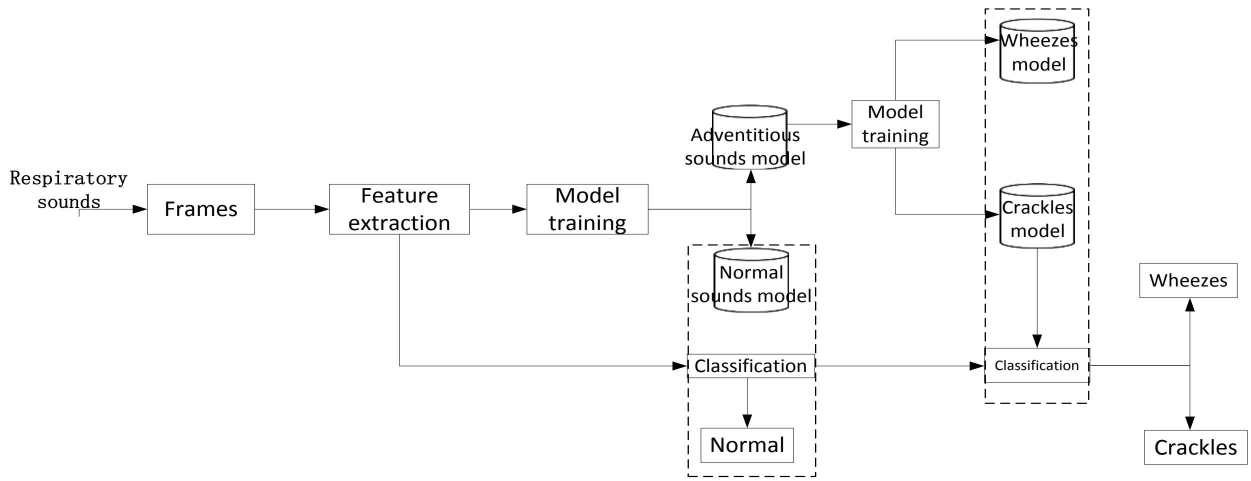

Fig. 2. Two-stage pipeline on GMMs to classify 3 types of respiratory sounds. For the stage 1st, Normal sounds or adventitious sounds are identified while for the 2nd stage wheezes or crackles are classified.

perceptron. The type of covariance matrix used is diagonal type and mixture components are 8 which could cause over fitting problem when it is applied on GMM models of both normal sounds and wheezes. The features extracted are MFCC with 24 coefficients.

III. Feature Extraction and GMM Models for Respiratory Sounds

A. Setup Devices

We setup a respiratory-sounds-device for the purpose of capturing lung sounds in home environment. Fig. 1 shows the components of the device which is consisted of the diaphragm and bell, the tubing, and the uni-direction pickup microphone [18] with sound recording system on a laptop. The diaphragm is responsible for capturing vibrations from inner body when being touched onto the chest. The captured sound wave is created in the bell with the hollow interior and is magnified by traveling along the tubing. At the end of the tube, a uni-direction pickup microphone is connected and could convert the analog waves into digital signal. Finally the sounds are recorded in the computer by the recording system. The uni-direction pickup microphone with the signal to noise ratio (SNR) about 100dB guarantees the sounds captured to be fidelity and almost without any noise reduction. This function is critical and helpful because some of adventitious sounds are week or strange and may be reduced if it is recorded by the commercial electronic stethoscope such as 3M Littmann [19]. And the sensitivity of the pickup microphone reaches -50dB. The high sensitivity could help improve the sound recognition rate during the sound preprocessing stage.

B. Classification Pipeline

When we get the respiratory sounds with the device in the section A, the audio files are first reprocessed since the raw data includes much invalid information that need to be removed. With the help of respiratory physicians who could accurately identify wheezes and crackles from the first affiliated hospital of Zhengzhou university, every type of respiratory sound is segmented manually from the begin of an inspiration where wheezes or cracks come up to the end where the sounds disappear either during the same inspiration or during the next expiration by using the software Adobe Audition [20]. After the preprocessing, the final data is assembled, which is made up by the clear and effective sounds where the non-wheezes or the non-crackles are eliminated. Most importantly a two stage classifying scheme for normal sounds, wheezes and crackles is designed just as shown in Fig.2. First the finally effective data is segmented into window frames. We extract a feature vector on each frame. For the first stage of the pipeline a classifier for normal and adventitious sounds through extracted feature vectors is modeled. The adventitious here are mainly consisted of wheezes and crackles. At the end of the first stage the normal respiratory sounds are actually identified. The pipeline turns into the second stage with the classification result of adventitious sounds, where a different model for classifying the wheezes and crackles is built. Eventually three types of respiratory sounds are recognized by the model of the first and the second stage

C. Feature Extraction.

The features are extracted on each window frame. Here we slit the sound sample segments into frames of the length of 85.3 microseconds (ms) with 42.6 ms overlap for the consecutive frames. We extract 46 features from time, frequency and cepstral domain. The features are described as below and their feature indices are shown in the table I. Each type of features is normalized with mean and variance equal to 0 and 1.

Zero Crossing Rate (ZCR). Zero-crossing rate of a frame is the rate where the signal changes from the positive to negative and vice versa. The ZCR of voiced and non-voiced sounds reflects observable fundamental changes. So ZCR is one key feature and often used in both speech recognition and audio classification.

Energy. The energy reflects the change of the signal amplitude along time. It is a basic and important audio feature.

TABLE I. EXTRACTED FEATURE INDEX AND NAME

Feature No.	Feature Name	Feature No.	Feature Name
1	ZRC	7	Spectral Flux
2	Energy	8	Spectral Rolloff
3	Entropy of Energy.	9-20	Chroma vector
4	Spectral Centroid	21	Fundamental Frequency
5	Spectral Spread	22	Harmonic Ratio
6	Spectral Entropy	23-46	24 MFCC Features
7	Spectral Flux		

We calculate the mean energy of all the point within a frame as the energy feature of that frame.

Entropy of Energy. The entropy of energy could be interpreted as the abrupt changes in the energy level of an audio [21]. If the energy of the signal changes rapidly, the entropy turns out to be very small. With this quality it is used to detect crackles and wheezes.

Spectral Centroid and Spectral Spread. Spectral centroid and spread measure the shape and position of the spectrum within an audio frame. Spectral centroid is the center of mass of the spectrum while the other measures the distribution that the spectrum is around the centroid.

Spectral Entropy. Spectral entropy is similar to energy entropy except that the entropy is computed in the frequency domain. The spectrum of the frame is divided into sub sections. And the energy of each sub section is normalized by the total spectral energy which is formulated to where i means the subsection number and is the energy of every subsection. The spectral entropy is calculated by .This feature is robust to noise because multi-band entropy feature tries to capture the location of the formants which are less affected by noise [22].

Feature Harmonic. Feature harmonic refers to Fundamental Frequency and Harmonic Ratio. The lowest frequency produced by any particular instrument is known as the fundamental frequency. Harmonics are periodic at the fundamental frequency. And the harmonic ratio is the orders based on the fundamental frequency.

Spectral Flux. Spectral flux measures the changes between two successive frames and is calculated by L2-norm of spectral amplitude difference between two adjacent frames [23]. It indicates how quickly the power spectrum changes.

Spectral Rolloff. Spectral Rolloff represents the concentration of the magnitude of the spectrum and measures the skewness of the spectral distribution. In our work we use a 90% threshold which means the frequency below 90% of the spectrum is concentrated. The work [24] shows that Spectral Rolloff is useful to discriminate speech and other audio signals.

Chroma vector. Chroma vector is obtained from frequency spectrum. And the spectral energy is summed up into 12 bins which represent 12 different pitches within a musical octave. Chroma vector actually shows the ratio of spectral energy [25]. It also removes the distinction between different octaves. Since Chroma vector capture both musical information and harmonic information, many musical audio analysis uses it as feature representation.

Mel-frequency Cepstral Coefficients (MFCCs). MFCCs are the discrete cosine transform coefficients of the mel-scaled log-power spectrum. MFCCs have been used to the study of the generation and transmission process of the respiratory sound [26]. MFCCs have also been shown as the best features for classifying wheezes and normal sounds in [27]. In this paper we calculated the first 24 coefficients from each frame as a feature vector.

D. Gaussian Mixture Model (GMM)

In our approach we use the GMM to model the distribution of the features extracted from lung sound classes. Each class is modeled according to a finite number of Gaussian mixture densities and its Probability Density Function (PDF) is

$$p(\vec{x} \mid \lambda) = \sum_{i=1}^{k} w_i \bullet N(\vec{x}; \vec{\mu}_i, \Sigma_i)$$

Where $\vec{x}$ is a D-dimensional feature vector, λ is the model of respiratory sound class. And $i(1 \ldots k)$ is the number of the Gaussian components. w_i is the weight of each Gaussian component such that $\sum_{i=1}^{k} w_i = 1$. $N(\vec{x}; \vec{\mu}_i, \Sigma_i)$ is the PDF of normal distribution from each component as:

$$N(\vec{x}; \vec{\mu}_i, \Sigma_i) = \frac{1}{(2\pi)^{|x|/2}\sqrt{|\Sigma_i|}} \exp\left(-\frac{1}{2}(\vec{x} - \vec{\mu}_i)^T \Sigma_i^{-1} (\vec{x} - \vec{\mu}_i)\right)$$

Where $\vec{\mu}_i, \Sigma_i$ are the mean and the covariance matrix.

From the functions above, each respiratory class is modeled by a Gaussian Mixture Model parameterized by three parameters, $\lambda = \{w_i, \vec{\mu}_i, \Sigma_i\}$, $i = 1, \ldots, k$ For every class with λ, we need to train the parameters which are estimated by the Expectation Maximization algorithm [28]. When testing a unknown sound, we calculate the likelihood of every sound class by GMM PDF $p(\vec{x} \mid \lambda_c)$. And the class c is determined according to which sound class generates the maximum log likelihood through threshold $\theta = \log p(\vec{x} \mid \lambda_c) - \log p(\vec{x} \mid \lambda_{\bar{c}})$. Where λ_c means normal sounds when $\lambda_{\bar{c}}$ is the adventitious sounds or λ_c means wheezes if $\lambda_{\bar{c}}$ turns out to be crackles. The threshold $\theta > 0$ defines the class to belong to λ_c and vice versa.

IV. EVALUATION RESULTS

A. Respiratory Sound Dataset

By the device that we setup we collected respiratory sounds of 20 individuals who produced wheezes, 20 individuals who got crackles and 20 normal individuals' lung sound at the first affiliated hospital of Zhengzhou university. Some of the individuals suffered from the obstructive pulmonary emphysema, some got asthma. All the participants were

TABLE II. THE LENGTH AND FRAME SIZE ABOUT THE 60 INDIVIDUAL RESPIRATORY SOUNDS

	Wheezes			**Crackles**			**Normal sounds**	
NO	**Size of frames**	**Length (ms)**	**NO**	**Size of frames**	**Length(ms)**	**NO**	**Size of frames**	**Length(ms)**
W1	223	9.581	C1	442	18.91	N1	350	15
W2	264	11.343	C2	481	20.607	N2	376	16.11
W3	403	17.272	C3	277	11.894	N3	517	22.13
W4	714	30.518	C4	274	11.774	N4	889	38
W5	231	9.9	C5	47	2.074	N5	543	23.241
W6	92	3.988	C6	278	11.916	N6	499	21.345
W7	705	30.144	C7	64	2.778	N7	800	34.203
W8	161	6.937	C8	331	14.188	N8	679	29.05
W9	456	19.501	C9	426	18.26	N9	754	32.241
W10	351	15.037	C10	250	10.744	N10	289	12.394
W11	173	7.461	C11	58	2.537	N11	457	19.573
W12	686	29.34	C12	88	3.803	N12	881	37.64
W13	188	8.099	C13	351	15.051	N13	544	23.286
W14	47	2.053	C14	296	12.676	N14	956	40.855
W15	219	9.393	C15	182	7.826	N15	408	17.476
W16	95	4.134	C16	200	8.584	N16	1081	46.203
W17	251	10.792	C17	329	14.112	N17	391	16.76
W18	153	6.587	C18	134	5.781	N18	278	11.92
W19	120	5.2	C19	71	3.076	N19	595	25.472
W20	276	11.835	C20	47	2.088	N20	660	28.226
Total	5808	249.115	Total	4626	198.679	Total	11947	511.125

between 50-70 years old because we mainly focused on the older. We cared less about gender ratio because it influenced little to the respiratory sounds.

Reference [3] shows us the upper frequency among lung sounds is below 3000Hz, so we first preprocess the sounds as canceling noises by the 8-order Butterworth [29] band pass filter with the frequency between 50Hz to 3000Hz. And then we down-sample the data to sampling rate of 6000Hz just as [5] did. Each individual's sound recording includes wheeze/crackles sections and non-wheeze/non-crackle parts (e.g. such as silence and normal breath sounds). Since that wheezes and crackles are attained with the help of some professional respiratory doctors. They use adobe audition software segmented the wheezes and crackles one by one after listening the whole recording. And then the processed respiratory sounds are framed by the hamming window with the length of 512 points (about 85.3million seconds). The overlap length is 256 sampling points. Finally the length and frame size about the 60 individuals' sounds are described in the table II.

B. *Feature Selection*

In order to determine which features among 46 features extracted from respiratory sounds are more important for classification, we use the wrapper method for feature selection. The Random Forest with the factor of Information Gain is chosen as the classification criteria.

The conclusion of reference [27] indicates MFCC gives the best performances for classifying both normal and wheeze sounds by contrasting the features such as FFT spectrum, wavelet packet decomposition and linear prediction coefficients. So we compare the features that we proposed (feature No 1-22) with MFCCs (No 23-46) to weight their importance.

We calculate the feature importance of every feature shown in Fig.3. We choose the features with the importance above 0.04 and the table III gives the top 6 features about their names and importance. We find that most of top features are related to Chroma vector which means Chroma features perform better in classifying normal and wheeze sounds. And among the top 22 features MFCCs only have the ratio about 18.2%. 18 features in the features that we proposed (from index 1 to 22) are among the top 22. In the next section we will compare FMCC features (index 23-46) and all features (index 1-46) on GMM for their recognition accuracy.

TABLE III. TOP 6 FEATURES ABOUT THEIR NAMES AND IMPORTANCE

Feature index	Feature name	Importance
10	Chroma 2th	0.068
12	Chroma 4th	0.052
16	Chroma 7th	0.045
9	Chroma 1th	0.043
15	Chroma 6th	0.042
2	Energy	0.041

C. *GMM Training and Mixtures*

For respiratory sounds training, each GMM represents one sound class. GMM parameters are estimated iteratively by the Expectation-Maximization (EM) algorithm. It converges to the Maximum Likelihood Estimate (MLE) of the mixture parameters. The problem is that GMM estimates are sensitive to the initial parameters which could lead to the boundary parameter with inaccurate estimation. Also the mixture number of GMM could affect the overall performance of EM. We use k-means [30] clustering algorithm to initialize the parameters. For the covariance matrix and components for GMM, we use Bayesian Information Criteria (BIC) [31] to penalize the complexity of the GMM to avoid over-fitting. BIC actually introduces a penalty term for the number of parameters besides

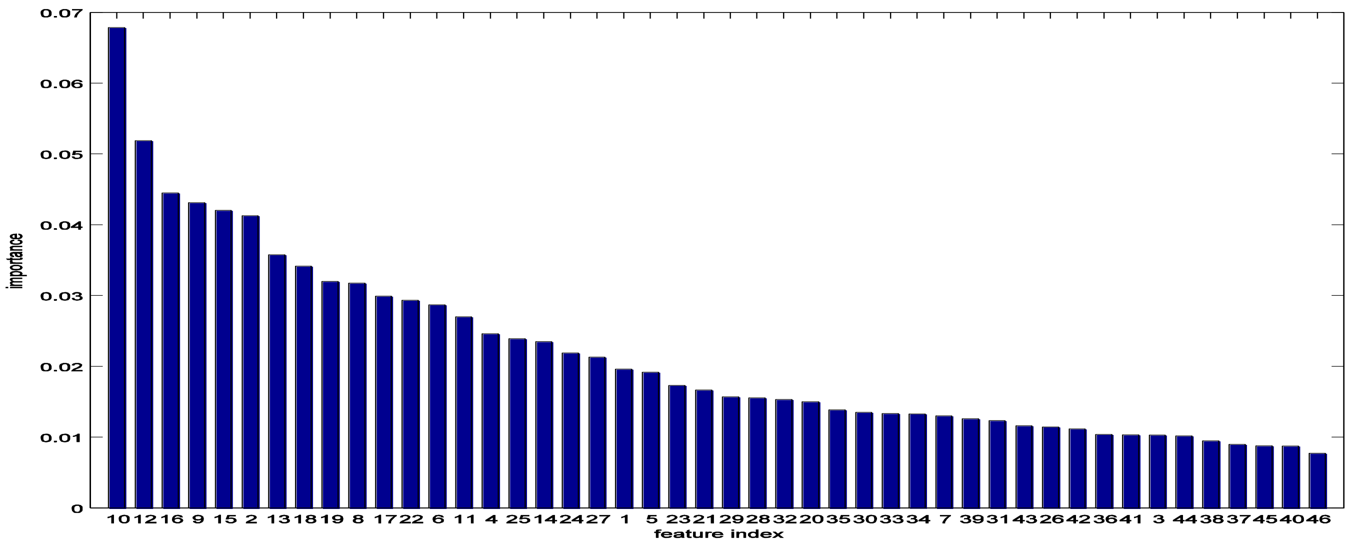

Fig. 3. Features ranked by Random Forest with the feature index

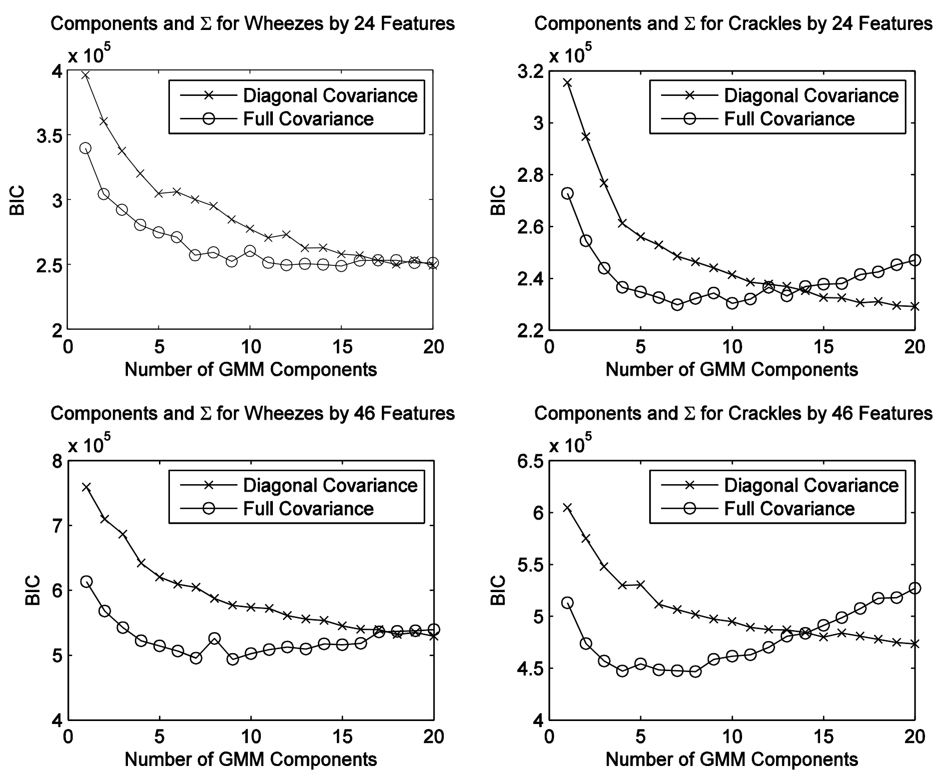

Fig. 4. The number of GMM components from 1 to 20 under BIC criteria for 24 MFCC or 46 proposed features.

the log-likelihood. For the parameters, the GMM model with the lowest BIC is preferred. The Figure 4 shows the results with two covariance matrices (the full and the diagonal) that we search for the number of GMM components from 1 to 20 under BIC criteria. And from the left-top of the figure, we could find the preferred number of GMM components is 15 by the full covariance. And by using the same method all GMMs for different respiratory sounds are researched and are shown in the table IV as the preferred parameters of GMM components.

D. *Performance Evaluation*

For the 60 participants we use 10 fold cross validation for testing the model performance. 10% of each type of respiratory sounds is randomly selected out and works as testing set. And the remaining data is used to train the GMMs. The process is repeated 10 times and the mean of all the confusion matrices is produced. For the first stage we train two GMMs for norm sounds and adventitious sounds (wheezes and crackles) and compare the confusion matrix of FMCC features with the features that is proposed in this paper. And for the second stage, we validate the confusion matrix for wheezes and crackles. The table V shows the results. When classifying about normal sounds and adventitious sounds, models of both MFCC features and all 64 features perform very well which approach the accuracy 99%. But for identifying wheezes and crackles, all 64 features based mode improves 7.4% for the crackles.

TABLE IV. GMM COMPONENTS OF DIFFERENT FEATURES AND RESPIRATORY SOUNDS TYPES

Name	MFCC (24 features)	All 46 Features
Normal sounds	19	10
Adventitious sounds	19	14
Wheezes	15	9
Crackles	7	8

TABLE V. THE CONFUSION MATRIX OF THREE RESPIRATORY SOUNDS BY THE TWO-STAGE GMM MODEL CLASSIFICATION

		Prediction			
		MFCC (24 features)		All 46 Features	
		Normal sound	Adventitious sounds	Normal sound	Adventitious sounds
Respiratory Sounds	Normal sound	**1185**	10	**1193**	2
	Adventitious sounds	13	**1031**	1	**1043**
	Precision[%]	98.9	99.0	99.9	99.9
		Wheezes	Crackles	Wheezes	Crackles
	Wheezes	**527**	54	**564**	17
	Crackles	24	**439**	7	**456**
	Precision[%]	95.6	89.0	98.8	96.4

V. CONCLUSION

In this paper, in order to develop a method for remote telemedicine and self-screening, we built a custom device using stethoscope to capture and analyze respiratory sounds. We extracted 46 features and used them to classify respiratory sounds into normal sounds, wheezes and crackles. The 24 FMCC features that have been demonstrated to be the best features to identify normal sounds and wheezes in [5] were compared with the 46 features proposed in this paper. The performance results show that all 46 features based GMM model accuracy improved by 7.4% for the crackles and 3% for the wheeze classification. A two-stage pipeline on Gaussian Mixture Model to classify 3 respiratory sounds is proposed and the optimal initial parameters of GMM for each sound are demonstrated. The final evaluation results show that classifying the 3 types of respiratory sounds achieved the average accuracy of 98.4%. With such superior performance, our proposed method can thus be used in the real-world situation for the diagnosis of pulmonary diseases.

ACKNOWLEDGMENT

The authors would like to thank doctor Huaqi Wang and Yonghai Feng from the First and the Fifth Affiliated Hospital of Zhengzhou University, for helping carry out the sound recordings, qualitatively evaluating them and providing constructive comments that greatly contributed to the improvement of this manuscript.

REFERENCES

[1] Sovijärvi, A., J. Vanderschoot, and J. Earis. "Standardization of computerized respiratory sound analysis." Crit Care Med 156 (1997): 974-987.

[2] Palaniappan, Rajkumar, Kenneth Sundaraj, and Nizam Uddin Ahamed. "Machine learning in lung sound analysis: a systematic review." Biocybernetics and Biomedical Engineering 33.3 (2013): 129-135.

[3] Sovijarvi, A. R. A., et al. "Characteristics of breath sounds and adventitious respiratory sounds." European Respiratory Review 10.77 (2000): 591-596.

[4] R.G.Loudon. "Clinical application of wheeze analysis", Monaldi Archieves of Chest Disease,48(5) 1993.

[5] Bahoura, Mohammed. "Pattern recognition methods applied to respiratory sounds classification into normal and wheeze classes." Computers in biology and medicine 39.9 (2009): 824-843.

[6] Pesu, L., et al. "Classification of respiratory sounds based on wavelet packet decomposition and learning vector quantization." Technology and Health Care 6.1 (1998): 65-74.

[7] Jin, Feng, Sridhar Krishnan, and Farook Sattar. "Adventitious sounds identification and extraction using temporal–spectral dominance-based features." IEEE Transactions on Biomedical Engineering 58.11 (2011): 3078-3087.

[8] İçer, Semra, and Şerife Gengeç. "Classification and analysis of non-stationary characteristics of crackle and rhonchus lung adventitious sounds." Digital Signal Processing 28 (2014): 18-27.

[9] Charleston-Villalobos, Sonia, et al. "Acoustic thoracic images for transmitted glottal sounds." Engineering in medicine and biology society, 2007. EMBS 2007. 29th annual international conference of the IEEE. IEEE, 2007.

[10] Reyes, B. A., et al. "Analysis of discontinuous adventitious lung sounds by Hilbert-Huang spectrum." Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE. IEEE, 2008.

[11] Sankur, Bülent, et al. "Comparison of AR-based algorithms for respiratory sounds classification." Computers in Biology and Medicine 24.1 (1994): 67-76.

[12] Kandaswamy, A., et al. "Neural classification of lung sounds using wavelet coefficients." Computers in Biology and Medicine 34.6 (2004): 523-537.

[13] Güler, İnan, Hüseyin Polat, and Uçman Ergün. "Combining neural network and genetic algorithm for prediction of lung sounds." Journal of Medical Systems 29.3 (2005): 217-231.

[14] Yamamoto, Hitoshi, et al. "Classification between normal and abnormal respiratory sounds based on stochastic approach." Proc. 20th International Congress on Acoustics. 2010.

[15] Matsunaga, Shoichi, et al. "Classification between normal and abnormal respiratory sounds based on maximum likelihood approach." Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on. IEEE, 2009.

[16] Bahoura, Mohammed, and Charles Pelletier. "Respiratory sounds classification using Gaussian mixture models." Electrical and Computer Engineering, 2004. Canadian Conference on. Vol. 3. IEEE, 2004.

[17] Mayorga, P., et al. "Acoustics based assessment of respiratory diseases using GMM classification." Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE. IEEE, 2010.

[18] http: //www.kuaiyu.com/

[19] http: //www.littmann.com/

[20] http: //www.adobe.com/products/audition/

[21] Giannakopoulos, Theodoros, Aggelos Pikrakis, and Sergios Theodoridis. "A multi-class audio classification method with respect to violent content in movies using bayesian networks." Multimedia Signal Processing, 2007. MMSP 2007. IEEE 9th Workshop on. IEEE, 2007.

[22] Misra, Hemant, et al. "Spectral entropy based feature for robust ASR." Acoustics, Speech, and Signal Processing, 2004. Proceedings.(ICASSP'04). IEEE International Conference on. Vol. 1. IEEE, 2004.

[23] Lu, Hong, et al. "SoundSense: scalable sound sensing for people-centric applications on mobile phones." Proceedings of the 7th international conference on Mobile systems, applications, and services. ACM, 2009.

[24] Guo, Guodong, and Stan Z. Li. "Content-based audio classification and retrieval by support vector machines." IEEE transactions on Neural Networks 14.1 (2003): 209-215.

[25] Ellis, Daniel PW, and Graham E. Poliner. "Identifyingcover songs' with chroma features and dynamic programming beat tracking." Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on. Vol. 4. IEEE, 2007.

[26] Xu, Jingping, Jingzhi Cheng, and Yanjun Wu. "A cepstral method for analysis of acoustic transmission characteristics of respiratory system." IEEE transactions on biomedical engineering 45.5 (1998): 660-664.

[27] Bahoura, Mohammed, and Charles Pelletier. "New parameters for respiratory sound classification." Electrical and Computer Engineering, 2003. IEEE CCECE 2003. Canadian Conference on. Vol. 3. IEEE, 2003.

[28] Reynolds, Douglas A., and Richard C. Rose. "Robust text-independent speaker identification using Gaussian mixture speaker models." IEEE transactions on speech and audio processing 3.1 (1995): 72-83.

[29] Sen, I., and Y. P. Kahya. "A multi-channel device for respiratory sound data acquisition and transient detection." Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the. IEEE, 2006.

[30] Lee, Younjeong, Ki Yong Lee, and Joohun Lee. "The estimating optimal number of Gaussian mixtures based on incremental k-means for speaker identification." International Journal of Information Technology 12.7 (2006): 13-21.

[31] Schwarz, Gideon. "Estimating the dimension of a model." The annals of statistics 6.2 (1978): 461-464.