Computers in Biology and Medicine 75 (2016) 118–129

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/cbm

Lung sound classification using cepstral-based statistical features

Nandini Sengupta [a,*], Md Sahidullah [b], Goutam Saha [a]

[a] Department of Electronics and Electrical Communication Engineering, Indian Institute of Technology, Kharagpur, Kharagpur 721302, India
[b] Speech and Image Processing Unit, School of Computing, University of Eastern Finland, Joensuu 80101, Finland

ARTICLE INFO

Article history:
Received 12 March 2016
Received in revised form
18 May 2016
Accepted 20 May 2016

Keywords:
Artificial neural network (ANN)
Auscultation
Discrete wavelet transform (DWT)
Mel-frequency cepstral coefficients (MFCCs)
Spectral features
Statistical features

ABSTRACT

Lung sounds convey useful information related to pulmonary pathology. In this paper, short-term spectral characteristics of lung sounds are studied to characterize the lung sounds for the identification of associated diseases. Motivated by the success of cepstral features in speech signal classification, we evaluate five different cepstral features to recognize three types of lung sounds: normal, wheeze and crackle. Subsequently for fast and efficient classification, we propose a new feature set computed from the statistical properties of cepstral coefficients. Experiments are conducted on a dataset of 30 subjects using the artificial neural network (ANN) as a classifier. Results show that the statistical features extracted from mel-frequency cepstral coefficients (MFCCs) of lung sounds outperform commonly used wavelet-based features as well as standard cepstral coefficients including MFCCs. Further, we experimentally optimize different control parameters of the proposed feature extraction algorithm. Finally, we evaluate the features for noisy lung sound recognition. We have found that our newly investigated features are more robust than existing features and show better recognition accuracy even in low signal-to-noise ratios (SNRs).

1. Introduction

Lung sound characteristics and their diagnoses form an indispensable part of pulmonary pathology [1]. Medical practitioners use diverse techniques to identify the lung sound characteristics but most of the available methods are not always convenient. The simplest and most popular technique to determine the pathological conditions from chest auscultation is to use a stethoscope. However, it is unreliable owing to the factors like: (a) the inexperience of the physician leading to his/her inability to recognize the lung sound abnormalities and dysfunctions and (b) low sensitivity of the human ear to the lower frequency band of a lung sound [2]. Advanced techniques like chest X-rays, spirogram, arterial blood gas analysis, which physicians heavily rely on, are also not without limitations. Arterial blood gas analysis is invasive and expensive; X-ray radiation is harmful for body and spirogram is a subjective process. Furthermore, the non-stationary nature of lung sounds, which gets severe for the case of abnormal subjects [2], makes the detection task an even more difficult for the physicians [3]. Ergo, more economical, patient-friendly, non-invasive, objective and automated or computerized techniques are desirable. Recent studies in computerized procedures have not only improved over simple auscultation processes, but they have also furnished new insights into the analysis of lung sounds for diagnostic purposes.

An automatic lung sound analysis system mainly progresses through the steps of pre-processing, feature extraction and classification. In pre-processing step, collected signal is prepared for the subsequent processing by reducing the heart sound effect [4], filtering or sample rate conversion and amplitude normalization. Feature extraction involves deriving a compact representation of a large set of data without losing distinguishable information. On the other hand, classification assigns different signals to their corresponding groups. A systematic and extensive study on various features and classifiers used for computer-based lung sound recognition can be found in a recent paper [5]. Wavelet-based features proved to be the most popular among these, since they focus on non-stationarity of lung sounds [2]. Compared to this, however, cepstral features along with Gaussian mixture model (GMM) classifier have shown better classification results [6]. Though Gaussian mixture model (GMM) and mel-frequency cepstral coefficients (MFCCs) yielded better recognition accuracy in the case of crackle, it was when MFCCs were used along with artificial neural network (ANN), that comparatively better results for wheeze and normal lung sounds were registered [7]. Over the years, a number of studies have been published where wavelet and cepstral features are used for this task [8–15]. Auto-regressive (AR) coefficients have also been used in certain cases [16–21]. But being highly sensitive to numerical precision, they were transformed

* Corresponding author.
E-mail addresses: nandinisengupta@iitkgp.ac.in (N. Sengupta), sahid@cs.uef.fi (M. Sahidullah), gsaha@ece.iitkgp.ernet.in (G. Saha).

http://dx.doi.org/10.1016/j.compbiomed.2016.05.013
0010-4825/

into cepstral coefficients [22]. Another frequently used set of features for lung status recognition are those derived from the power spectrum of the signal [23–29]. Yet, being overly sensitive to noise, those power-spectrum based features show poor results.

Motivated by the great success of cepstral features in speech pattern classification, in our work, we explore the cepstral features further for lung sound recognition. In speech processing, they are compact representation of spectral envelope that separates source and filter [30]. The source signal is due to the vibrations of the vocal cords whereas the vocal tract acts as a filter [22]. In the context of lung sounds, the chest wall or thoracic wall acts as a filter through which the lung sounds sift through [31,32]. Besides, cepstral coefficients are suitable for pattern recognition task for they are uncorrelated with each other.

In our study, we consider three types of lung sounds: *normal*, *wheeze* and *crackle*. Towards this end, five popular cepstral features are studied falling under two broad categories: (i) features based on all-pole modeling and (ii) features based on filterbank method. Under the first head, *linear prediction cepstral coefficient* (LPCC) and its perceptual form, *perceptual linear prediction coefficient* (PLPCC) [33] are considered. In the case of filterbank-based approach, *linear frequency cepstral coefficient* (LFCC), its perceptual form, *mel-frequency cepstral coefficient* (MFCC) [34] and a complementary feature of MFCC, i.e., *inverted mel-frequency cepstral coefficient* (IMFCC) [35] are studied. Next, we have proposed a new feature extraction method that uses statistical properties of the cepstral coefficients. The proposed feature outperforms the wavelet-based features in terms of classification accuracy. In addition to that, the proposed features require significantly less time for classification task as compared to the baseline cepstral features. Finally, we optimize different parameters for our proposed features and evaluate their performance in presence of different noises.

The rest of the paper is organized as follows. Section 2 briefly describes the database used in our study. In Section 3, short-term spectral characteristics of lung sound are studied. In Section 4, cepstral features are described for the completeness of this paper. In Section 5, we study the statistical properties of cepstral coefficients and propose new features from this analysis. Experimental setups are described in Section 6. Results are discussed in Section 7. Finally, we conclude the paper in Section 8 by summarizing the work and by providing some future directions.

2. Database description

Systematic collection of lung sound samples with reliable ground truth is an important part of this research. Our database consists of recorded lung sounds obtained from three different resources: RALE database,[1] Audio and Bio-signal Processing Lab (IIT Kharagpur),[2] Institute of Pulmocare and Research (Salt Lake, Kolkata).[3] The sampling frequency of the original recording is 8000 Hz in all the three cases. The ground-truths were verified by experienced pulmonologists. The database includes three types of lung sounds: normal, crackle (both fine and coarse crackle) and wheeze. As the heart sound is mixed with the recorded lung sound, the lung sound separation is an essential step before the samples are modeled and characterized. In this work, the effect of heart sound has been reduced from the collected data by adopting the method proposed in [36]. In this approach, the signal is first decomposed into intrinsic mode functions (IMFs) using empirical mode decomposition (EMD) technique. After that, an energy-based heart sound peak detection method is used in each IMFs within the LS signals followed by a boundary estimation algorithm to localize the spread of the heart sounds. Finally, localized heart sounds are reduced from the signal using filtering technique. The lung sound cycles were extracted using a Hilbert envelope based algorithm [37] and we use 72 cycles (24 from each class) for our experiments collected from 30 different subjects.

3. Spectral characteristics of lung sounds

Lung sounds can be broadly categorized into *normal*, *abnormal*, and *adventitious* [38,39]. *Normal* lung sound refers to the respiratory sound of a healthy subject, hardly audible without a stethoscope. On the contrary, *abnormal* sound indicates the absence or decrease of normal sound [38] while *adventitious* sound appertains to the superposition of additional sound with normal sound. Adventitious sounds can again be continuous or discontinuous. Among the several adventitious lung sound instances we come across in the medical field viz. wheeze, crackle, cough, rhonchus, squawk, stridor and so on, two most common adventitious sounds wheeze (continuous) and crackle (discontinuous) are considered for detailed study in this paper. Often, either wheeze or crackle is found to be a characteristic feature of most of the other adventitious pulmonary sounds mentioned earlier. Acoustically, wheezes characterize cough sounds, even the asthmatic ones. Rhonchus is a low-pitched wheeze; squawk is a short wheeze preceded by a crackle; and stridor discernibly is low-frequency wheeze [39].

Wheeze: Wheeze is a continuous and adventitious waveform of duration 250 ms or more, perceptibly displaying a musical character [40]. Wheezing, usually the outcome of localized or diffused airway narrowing or obstruction in the passageway from the larynx to the small bronchi, can be induced by multiple causes like mucosal edema, external compression, partial obstruction by a tumor or foreign body, and so on [41]. The dominant frequency range typifying a wheeze normally exceeds 100 Hz [2,41]. Wheezing is a symptom of diseases predominantly associated with the obstruction of airways. Whereas reversible obstruction characterizes asthmatic tendencies [42], irreversible obstruction denotes chronic obstructive pulmonary diseases (COPDs) [43], for example, emphysema (caused by smoking), chronic bronchitis and cystic fibrosis.

Crackle: Discontinuous in nature, crackles normally last for 1–10 ms [40,31]. In general, crackles are caused by explosive opening of small airways [44]. They are short, explosive, non-musical sounds heard on inspiration and sometimes during expiration. It has a distinctly wide frequency range up to 2000 Hz [38,44]. Two types of crackles are commonly found: *fine* and *coarse*. Fine crackles are high pitched and last for less than 5 ms as distinguished from coarse crackles, which are low pitched and have a duration of about 10 ms [45]. Crackles detected in patients can be symptomatic of alveolitis, interstitial lung diseases [46], congestive heart failure [47], carotid arterial stiffness [48], etc.

Various types of lung sounds have their different spectral characteristics [49]. From the spectrogram plot in Fig. 1, it can be observed that, crackles are distinguishable from normal and wheeze for their wide frequency range that usually extends up to 2000 Hz. This is also evident from Fig. 1 that unlike crackle or wheeze, a normal respiratory sound has a narrower frequency range. Then we have done statistical analysis with power-spectrum coefficients to study the class-separability in different frequency components of short-term power spectrum. We use *F*-ratio based method which gives separability measure between the data of two or more classes [50,51]. For a particular frequency

[1] http://www.rale.ca/repository.htm
[2] http://www.ecdept.iitkgp.ernet.in/index.php/home/labs/bio-sig-proc
[3] http://www.pulmocareindia.org/