Neurocomputing 141 (2014) 139–147

Contents lists available at ScienceDirect

Neurocomputing

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

Content-based classification of breath sound with enhanced features

Baiying Lei [a,b,*], Shah Atiqur Rahman [b], Insu Song [b]

[a] National-Regional Key Technology Engineering Laboratory for Medical Ultrasound, Guangdong Key Laboratory for Biomedical Measurements and Ultrasound Imaging, Department of Biomedical Engineering, School of Medicine, Shenzhen University, Shenzhen 518060, China

[b] James Cook University Australia, School of Business/IT, Singapore

ARTICLE INFO

Article history:
Received 26 July 2012
Received in revised form
18 January 2014
Accepted 1 April 2014
Communicated by: R. Capobianco Guido
Available online 29 April 2014

Keywords:
Breath sound
Content-based classification
Support vector machine
Artificial neural network

ABSTRACT

Since breath sound (BS) contains important indicators of respiratory health and disease, analysis and detection of BS has become an important topic, with diagnostic and assessment of treatment capabilities. In this paper, the identification and classification of respiratory disorders based on the enhanced perceptual and cepstral feature set (PerCepD) is proposed. The hybrid PerCepD feature can capture the time-frequency characteristics of BS very well. Thus, it is very effective for the exploration and classification of normal and pathological BS related data. The classification models based on support vector machine (SVM) and artificial neural network (ANN) have been adopted to achieve automatic detection from BS data. The high detection accuracy results validate the performance of the proposed feature sets and classification model. The experimental results also demonstrate that the high accuracy of the pathological BS data can provide reliable diagnostic suggestions for breath disorders, such as flu, pneumonia and bronchitis.

1. Introduction

Breath sound (BS) has been widely used in diagnosing respiratory diseases, such as flu, pneumonia and bronchitis because BS contains important indicators of respiratory health and diseases. The World Health Organization (WHO) has defined pneumonia solely on the basis of clinical findings obtained by visual inspection and timing of the respiratory rate [1]. In one study, diagnosing a total of 222 children with pneumonia, fast breath was found to be the most useful sign for pneumonia in all age groups [2]. To overcome the problems of subjective auscultation-based diagnosis of respiratory problems, various automated signal processing methods have been proposed [3–19]. The significance of medical applications of this study has been highlighted in various computerized and automatic sound/voice detection systems [3–19]. The findings of these studies suggest that computerized sound analysis is a useful and effective way to provide information of a wide variety of health conditions including respiratory diseases. For example, in [5], Azarbarzin et al. proposed an automatic snore extraction algorithm with the unsupervised fuzzy C-means clustering method from the respiratory sound, which demonstrate high accuracy both for tracheal sound recordings and ambient microphone. Similar system was proposed by Yadollahi et al. [3] where multiple spectral and prosodic features are combined to analyze the signal and fisher linear discriminant analysis with Bayesian threshold were used to identify snores. A method for discriminating between normal and pathological BS is proposed by Wang et al. [6] where they achieved high accuracy by using mel-frequency cepstral coefficient (MFCC) with a hybrid classification model of Gaussian mixture model and support vector machine (GMM-SVM). Cough signals were detected in continuous audio recording by means of intensity and frequency information of the sound in [7]. Automatic voice disorder was detected by means of cepstral features [12] or by means of combination of spectral and prosodic features [11]. Tapliduo et al. [13] proposed a method which automatically detects wheeze episodes by employing time-frequency analysis on the BS. Banora [10] proposed a method for wheeze detection which employed different types of features such as Fourier transform, wavelet transform, MFCC and various classifiers, for example, vector quantization, GMM, Artificial Neural Network (ANN). In [14], the adventitious respiratory sound was identified and extracted to facilitate physician analysis of pulmonary dysfunction based on the temporal-spectral dominance feature. In [15], Shin et al. presented a detection system for the diagnosis of pathological conditions based on cough sounds. This automatic system is built on the hybrid ANN and Hidden Markov Model (HMM). The developed prototype had high performance when the signal to noise ratio was below 5 dB and could be used

* Corresponding author at: National-Regional Key Technology Engineering Laboratory for Medical Ultrasound, Guangdong Key Laboratory for Biomedical Measurements and Ultrasound Imaging, Department of Biomedical Engineering, School of Medicine, Shenzhen University, Shenzhen 518060, China.
Tel.: +65 6576 6833; fax: +65 6455 2833.
E-mail addresses: leiby666@gmail.com (B. Lei), atiqur.rahman@jcu.edu.au (S.A. Rahman), insu.song@jcu.edu.au (I. Song).

http://dx.doi.org/10.1016/j.neucom.2014.04.002
0925-2312/

for real-time processing. In [16], the respiratory sound was detected at the external ear, which also validated the effectiveness of detecting and monitoring the breath and respiration over an extended period of time.

However, almost all computerized respiratory sound analysis methods are based on the data collected by using special sensors attached to the body. This is not only invasive, but also requires trained personnel to properly acquire the data. Currently, to our best knowledge, there are no computerized diagnosis methods available for pneumonia using breathing sounds collected by hand-held microphones (i.e., a non-contact method). Therefore, there is a strong need for such a method: Moreover, the existing literature strongly suggests that BS collected via non-contact methods would enable early diagnosis of respiratory diseases [20].

Unlike the audio/voice signal, BS is often regarded as a band-limited or broadband noise [21]. Thus, it is necessary to comprehend its unique time and frequency features before classifying or detecting the BS signal. Feature extraction may be the most significant part of the BS feature classification stage. The effectiveness of BS detection obviously depends on its ability to classify sound data properties or contents. A reliable, accurate, fast and content-based method for BS data classification is essential for providing treatment and diagnosis of different respiratory diseases. A recent trend is the effectiveness of the content-based classification [22–25] techniques for the audio data. As an important part of the classification and separation of signal, the content-based classification strategy [24] shows a high accuracy compared to the traditional method. Therefore, the popular content-based classification algorithm is employed for the feature extraction of our BS data. It is known that some features cannot be captured well by just one feature set. For example, the perceptual feature has only spectral characteristics, while the cepstral coefficient cannot be reconstructed as it only captures the shape of the frequency spectrum of BS. Obviously [24], the enhanced perceptual and mel-cepstral (PerCepD) feature, which explores more features such as the delta and delta–delta coefficients to capture time dynamical information, is investigated in our study. Therefore, BS data characteristics can be better represented by only one feature set.

Compared to the pattern recognition field, the classification or separation technique is a relatively new recognition technique. Many techniques have been proposed, such as k nearest neighbor (KNN), ANN [12,26–29], Gaussian mixture model (GMM) [6,30], hidden Markov model (HMM) [7,15] and support vector machine (SVM) [6,11,23,29]. Some works in the literature use a hybrid model to achieve optimum classification and detection performance [6,31]. An overview and performance evaluation of different classification methods for the respiratory sound can be found in [10]. To be more effective and accurate in the detection of the BS data, popular kernel based algorithms, such as ANN and SVM, are both adopted in the classification model, because both methods can achieve high accuracy and have more variability for solving practical problems [32–35]. Furthermore, this algorithm has been extended to other sound related classification and detection areas for the medical applications, such as cough sound, heart sound, respiratory sound and crying sound.

In this paper, the hybrid feature set and content-based classification model is proposed for automatic detection of the pathological BS data collected by hand-held microphones. The main contributions of this paper is to provide (1) an analysis of BS collected at the mouth, (2) acoustic feature extraction methods for BS so that the diagnosis of major respiratory disease can be achieved with diagnostic feature selection and (3) classification methods from BS. To the best of our knowledge, this is the first report on a method for diagnosing major respiratory diseases using BS data collected by microphones held at the mouth without any direct contact with patients. Moreover, the set of features, named PerCepD, has not been experimented before, especially in combination with the proposed classifiers, for the diagnosis of breath sounds. The organization of this paper is as follows. Section 2 introduces the methodology of feature extraction in detail. The pattern classification algorithms, such as SVM and ANN, are presented in Section 3. Section 4 demonstrates the proposed method with the experimental results. Finally, conclusions and future directions are given in Section 5.

2. Feature extraction

For the classification procedures, three steps are involved. The first step is to find the optimal parameters to discriminate the BS data. The second step is to design the reference model to evaluate the similarity of normal and pathological BS and implement it practically. The final step is to classify and recognize the BS data statistically or directly with the developed model.

Fig. 1 plots a schematic diagram of a BS detection algorithm. The detection algorithm is composed of data acquisition, signal pre-processing, feature extraction, model training and pattern classification, and final classification results.

2.1. Data acquisition

During the data acquisition stage, the BS signals are directly collected from the participants with digital voice recorders. The captured BS signals are usually different from center to center depending on data collection methods and locations, the available devices, the established practice, the training level and the particular applications. The collected BS data are from a variety of different sources with different formats. The recorded BS is resampled to 8000 Hz with 16 bit resolution to facilitate the analysis techniques.

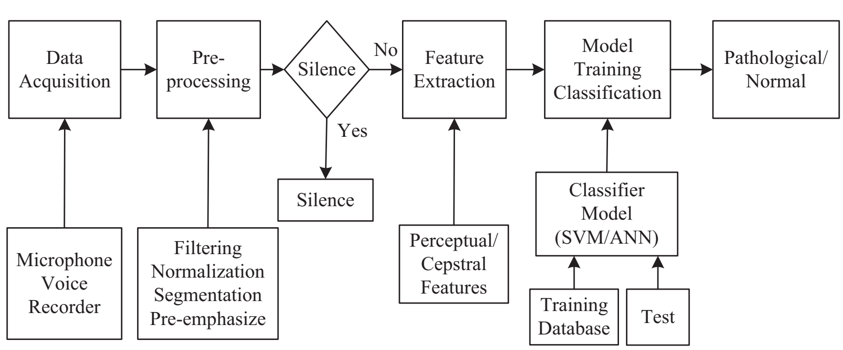

Fig. 1. Schematic diagram of the BS detection algorithm.