Biomedical Signal Processing and Control 33 (2017) 132–140

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

Cough detection by ensembling multiple frequency subband features

Mingyu You [a,*], Zeqin Liu [a], Chong Chen [a], Jiaming Liu [a], Xiang-Huai Xu [b], Zhong-Min Qiu [b]

[a] Department of Control Science & Engineering Tongji University, Shanghai, China
[b] Department of Respiratory Medicine, Tongji Hospital Tongji University School of Medicine, Shanghai, China

ARTICLE INFO

Article history:
Received 15 April 2016
Received in revised form 13 October 2016
Accepted 3 November 2016

Keywords:
Cough detection
Multiple frequency subbands
Feature extraction
Gammatone filterbank
Ensemble method

ABSTRACT

Cough is a common symptom in respiratory diseases. Objectively evaluating the quantity and intensity of cough by pattern recognition technologies can provide valuable clinical information for cough diagnosis and monitoring. Cough detection is the basis of cough diagnosis and analysis. It aims at detecting cough events and their exact boundaries from an audio stream. From signal characteristics, it is found that energy distribution scatters in the cough spectrum, which is obviously different from speech signals. However, almost all feature extraction methods for cough detection in previous works are derived from the speech recognition domain. In this article, subband features are obtained by using gammatone filterbank and an audio feature extraction method. Support Vector Machine (SVM), K-Nearest Neighbors (KNN) and Random Forest (RF) are trained with the corresponding subband features and ensemble method combines the outputs to make the final decision. Experiments are conducted on both synthetic data and real data. The real data is collected from 18 patients with respiratory diseases in clinical environments and annotated by human experts. Experiment results demonstrate that ensembling multiple frequency subbands helps to impove performance in cough detection. Compared with other methods, our method can improve the accuracy by 3.2%.

1. Introduction

Cough, a common symptom in respiratory diseases, has a serious impact on a patient's daily life. In clinical diagnosis and treatment, physicians need the precise description and progression of patient's symptoms. However, current methods in clinic depend on the patients' subjective judgments. Due to the incomparable personal criteria, subjective results can be inaccurate and unreliable. Some medical institutions have proposed to record daily sound continuously by a microphone attached to a patient's collar. Trained physicians will then be employed to identify and evaluate the cough event [1]. The manual identification of coughs is time consuming and laborious. An automatic method for detecting coughs is necessary.

Researches on automatic cough detection began in 1950s [1]. Cough detection is also called cough recognition or cough monitoring in healthcare field. The purpose of cough detection is to identify the cough sections from audio recordings. Researchers have found that traditional audio signal processing methods without pattern recognition are not a good way to deal with it [1]. Today, there are two popular frameworks for accurate cough detection. The first framework judges audio signals, frame by frame, and combines continuous cough frames as a cough event [2]. The second consists of two steps: event detection and cough classification [3]. Event detection aims at identifying the cough event candidates. The individual candidates will then be identified in the cough classification step. Event detection highly influences the overall performance, while it is difficult to be conducted. In the article, the cough detection system follows the first framework, and the feature extraction step before frame judgement is mainly focused.

Work [3] presents a system for automatically analyzing 24-h, continuous, ambulatory recordings containing coughs. The experimental data is collected from 18 patients and 8 volunteers. A Hidden Markov Model (HMM) is built using Mel Frequency Cepstral Coefficients (MFCC) as features for cough detection. They achieved a better result (median sensitivity value of 85.7%, median false positive rate of 0.8 events/h) compared with [2]. Nevertheless, this system requires a great deal of "filler" models for modeling noise signals, which brings a high computational cost. In another experiment [4], an automatic system monitoring health condition in real time reports that cough signal has a broad spectral distribution between 350 Hz and 4 KHz. The system combined Artificial Neural Network (ANN) and HMM. It extracted Energy Cepstral Coefficients (ECC) features instead of MFCC. The proposed HMM-ANN

* Corresponding author.
E-mail address: myyou@tongji.edu.cn (M. You).

http://dx.doi.org/10.1016/j.bspc.2016.11.005
1746-8094/© 2016 Published by Elsevier Ltd.

model have a good performance when the input SNR is between −10 dB to 0 dB. However, the result is just based on synthesized data by mixing cough signal with pink noise or a movie sound. Common noises such as speech, laugh and sneeze in realistic environments are ignored. The system is far from the application in clinical environments.

The goal of research [5] is to compare the performance of several sensors for cough detection: ECG, thermistor, chest belt, accelerometer, contact, and audio microphones. It extracts multiple types of features from different sensor signals in the same environment. Eventually audio microphones achieves the best performance with 94.5% both averaged sensitivity and specificity. However, there are two weaknesses in the research. First, the experimental data are recorded from 32 healthy subjects. It is doubtful whether the system could work for sick patients in clinic. Second, it is improper to apply audio feature extraction methods to other types of signals. Besides, another work [6] exploits microphone and vibration sensor to assess the accuracy of cough classification. Because of the efficiency and convenience, we use audio micophone to collect data.

Work [7] develops an automatic system for cough monitoring of pulmonary tuberculosis patients, which achieves sensitivity of 75.5% and specificity of 99.6%. Sound event detection and Support Vector Machine (SVM) are combined in the detection process. The sensitivity is unsatisfactory because many cough events are missed in the event detection step. Recently, cough detection work [8] using Deep Neural Networks (DNN) compares the performances of different network parameters and obtains a better result than the traditional Gaussian Mixture Model (GMM) in a clinical environment. However, from a practical point of view, most cough detection systems are still at the research stage, a more efficient and applicable cough detection system needs to be developed.

The cough is a kind of transient and nonstationary sound, like environmental noise. Studies have found that common audio features are suitable for speech but not environment sounds, due to different characteristics [9]. In the field of environmental sound recognition, Subband Temporal Envelope (STE) is extracted as features followed by improved SVM model with probabilistic distance [9]. Study [10] extracts features by Gabor Transformation and classifies sounds by kernel fisher discriminant analysis. In [11], features extracted from coughs are divided into three categories: spectral contents, measures of noise and prosody-related features. Feature selection is used to obtain the best combination of these features. However, all of these features are also directly derived from the fields of speech or musical application, which is different from clinical cough detection. Another interesting work [12] finds that extremely high frequencies (15 kHz–90 kHz) of cough signal carried useful information. But it needs the device's higher specifications. Hull Automatic Cough Counter (HACC) [13] exploits cepstral coefficients and linear predictive coding (LPC) as features, both of which are widely used in speech recognition. Another interesting work proposes a novel wavelet feature, and combines it with other common features to analyze cough for the childhood pneumonia diagnosis [14]. As for features in time domain, the aim of study[15] is to discriminate the nature between productive and non-productive coughs by using analysis in time domain. So far, there have been few studies aimed at proposing specific features for cough detection, concerning the spectral distribution of cough signals.

In this paper, subband features are obtained by using gammatone filterbank and an audio feature extraction method. We also consider the ensemble approach to the different feature extraction. We use classical classification model SVM on our experiment. The ensemble method is exploited to deal with multiple subbands and obtain a better result. The following of the paper is organized as follows: An overview of the system followed by subband feature extraction and ensemble method modeling are all presented in Section 2. Section 3 reports the experimental results and the related analysis. And the paper is concluded in Section 4.

Table 1
Basic information of patients.

ID	Gender	Age	Disease	ID	Gender	Age	Disease
1	F	21	CAP	10	F	56	CAP
2	M	47	CAP	11	M	56	COPD
3	F	48	CAP	12	F	59	CAP
4	F	49	CAP	13	M	61	COPD
5	F	51	CAP	14	M	69	CAP
6	F	52	CAP	15	M	72	CAP
7	F	53	COPD	16	F	76	BA
8	F	54	CAP	17	F	79	BA
9	F	55	CAP	18	M	66	COPD

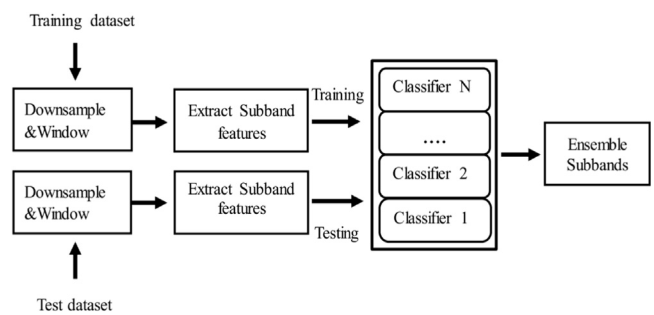

Fig. 1. Framework of cough detection.

2. Material and methods

2.1. Data collection

Open source datasets for cough detection are not yet available. The cough data used in the article is collected from patients with chronic respiratory diseases. The data acquisition device consists of a portable digital recorder (SONY ICD-LX30) and a microphone (ECM-CS10) attached to patients' collars. The microphone frequency response is from 100 to 16000 Hz. The sampling rate of the device is set to 44100 Hz. Recording lasts continuous 24 h. Patients are encouraged to go about their daily routines in hospital environment, in order to obtain natural coughing sounds. SVM is used as the classification model. The details of patients are shown in Table 1. ID represents different patients. About Gender, F means female and M for male. The majority of patients in the column are old age. All recordings are divided into a period of 10 min to save for future processing. Considering the fact that data amount affects computation time and accuracy of results, we only choose a subset of data. Experimental data are random selected from 18 patients, 80 min each from the recordings, about 1440 min. All patients had various chronic respiratory diseases, including community acquired pneumonia (CAP), bronchial asthma (BA) and chronic obstructive pulmonary disease (COPD). Besides real data, synthetic data are also employed in the experiment. The detailed information of synthetic data will be introduced below.

2.2. Framework

As shown in Fig. 1, our framework for cough detection consists of downsampling, windowing, subband feature extraction, classification model construction and subband ensemble. The original audio sampling rate is 44100 Hz, which is downsampled to 16000 Hz for less computation cost. The training dataset is exploited for building classification model and the test dataset is for performance