Biomedical Signal Processing and Control 38 (2017) 322–336

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

Overcomplete discrete wavelet transform based respiratory sound discrimination with feature and decision level fusion

Sezer Ulukaya [a,b,*], Gorkem Serbes [c], Yasemin P. Kahya [a]

[a] Department of Electrical and Electronics Engineering, Boğaziçi University, 34342 Istanbul, Turkey
[b] Department of Electrical and Electronics Engineering, Trakya University, 22030 Edirne, Turkey
[c] Department of Biomedical Engineering, Yildiz Technical University, 34220 Istanbul, Turkey

ARTICLE INFO

Article history:
Received 30 December 2016
Received in revised form 17 May 2017
Accepted 30 June 2017
Available online 23 July 2017

Keywords:
Pulmonary sound
Rational dilation wavelet transform
Q-factor
Ensemble learning
Adventitious lung sounds
Statistical feature extraction

ABSTRACT

Background and objective: Crackle, wheeze and normal lung sound discrimination is vital in diagnosing pulmonary diseases. Previous works suffer from limited frequency resolution and lack of the ability to deal with oscillatory signals (wheezes). The main objective of this study is to propose a novel wavelet based lung sound classification system that is capable of adaptively representing crackle, wheeze and normal lung sound signal time-frequency properties.
Methods: A method which is based on rational dilation wavelet transform is proposed to classify lung sounds into three main categories, namely, normal, wheeze and crackle. Six different feature extraction methods were used with five different classifiers all of which were compared with the proposed method on 600 lung sound episodes in a cross validation scheme. Six statistical subset features were extracted from raw features and fed into classifiers. After comparative evaluation of the proposed method, an ensemble learning scheme was built to increase the performance of the proposed method.
Results: It has been shown that performance of the proposed method was superior to previous methods in terms of accuracy. Moreover, its computational time was far less than its nearest competitor (S transform). It has also been shown that the proposed method was able to cope with oscillatory type signals as well as transient sounds performing 95.17% average accuracy for energy subset and 97.38% ensemble average accuracy showing a promising time-frequency tool for biological signals.
Conclusions: The proposed method has shown better performance even using only one subset of extracted features. It provides better time-frequency resolution for all types of signals of interest and is less redundant than continuous wavelet transform and significantly faster than its nearest competitor.

1. Introduction

Stethoscope is a traditional tool in diagnosing respiratory dysfunctions and disorders. However, it is regarded to have low diagnostic value due to its limited frequency response which attenuates frequencies greater than 120 Hz and due to the subjectivity [1] involved in the evaluation of the auscultated sounds. Moreover, the traditional stethoscope offers no option to record pulmonary data for further analysis. Interdisciplinary efforts in medicine and engineering as summarized in [2], which aim to make auscultation a more valuable diagnosis tool, use advanced machine learning and signal processing algorithms to be utilized in treatment follow-ups and remote diagnosis.

Lung sounds (LS) are believed to be produced by the turbulent flow in the lung airways, and are essentially classified as adventitious (abnormal) sounds and vesicular (normal) sounds [1,3]. The normal breath sounds heard over the chest wall, which are synchronous with air flow in the airways, are defined as vesicular sounds. The frequency spectrum of vesicular sounds in healthy people has 200–600 Hz dominant frequency range. Adventitious lung sounds (ALS), which are specific markers of various respiratory diseases, are superimposed on vesicular sounds. Discontinuous adventitious lung sounds (DALS) and continuous adventitious lung sounds (CALS) are the two main categories of ALS. In literature, crackles and wheezes can be exemplified as the most known components of the CALS and DALS, respectively [4,5].

Crackles are non-musical instantaneous bursts, explosive in nature and divided as coarse (lower pitch) or fine (higher pitch).

* Corresponding author at:Electrical and Electronics Engineering Department, Boğaziçi University, 34342, Bebek, Istanbul, Turkey.
E-mail addresses: sezer.ulukaya@boun.edu.tr, sezer.ulukaya@gmail.com (S. Ulukaya), gserbes@yildiz.edu.tr (G. Serbes), kahya@boun.edu.tr (Y.P. Kahya).

http://dx.doi.org/10.1016/j.bspc.2017.06.018
1746-8094/

List of abbreviations

ALS	adventitious lung sounds
CALS	continuous adventitious lung sounds
CWT	continuous wavelet transform
DALS	discontinuous adventitious lung sounds
DT	decision tree
DWT	discrete wavelet transform
ELM	extreme learning machines
FT	Fourier transform
LOOCV	leave-one-out cross validation
LS	lung sounds
MFCC	Mel frequency cepstral coefficients
NB	Naive Bayes
PLP	perceptual linear prediction
PSD	power spectral density
RADWT	rational dilation wavelet transform
S Transform	Stockwell transform
SVM	support vector machine

Crackles are believed to be generated by abnormally closed airway openings [6]. The frequency spectrum of crackles varies between 200 and 2000 Hz range while their duration is usually less than 20 ms. In most of the pulmonary diseases, the severity of the disease has a strong correlation with the number of crackles per breath. Moreover, different lung diseases can be diagnosed by using the occurrence times, durations and types of crackles within a breath cycle [6]. For example, in bronchopneumonia and bronchiectasis coarse crackles exist typically whereas the common symptoms of pneumonia and interstitial fibrosis are fine crackles [7]. Crackles in chronic obstructive pulmonary disease are coarse and occur in early to mid inspiration whereas crackles in fibrosing alveolitis are fine and occur in late inspiration [6]. Due to their transient waveform and scattered energy content over frequencies (200–2000 Hz), crackles have barely noticeable impact on the total power spectrum [7]. The frequency and time domain characteristics of vesicular and crackle sounds overlap in both domains.

Wheezes are oscillatory waveforms, which last more than 80–250 ms, and represent narrow-horizontal lines in the time-frequency domain (>100 Hz). Time-frequency representation of two samples of crackles and wheezes is illustrated in Fig. 1. Various pulmonary diseases such as asthma and chronic obstructive pulmonary disease, can be diagnosed by using the presence of wheezes. As given in [8,9], the degree of airway obstruction may be related to wheeze properties such as duration and pitch frequency.

The motivation of the proposed LS, CALS and DALS discrimination system is to overcome the subjectivity and low-performance drawbacks of traditional non-automatic systems while increasing the performance of automated systems by enhancing the frequency selectivity of wavelet filters resulting in a better separation of overlapped components of these three lung sounds in time-frequency domain.

In previous studies, crackle/non-crackle and wheeze/non-wheeze episode discrimination has been extensively examined and a summary can be found in [10,4]. In addition to binary discrimination approaches, there were also a few three class (crackle, wheeze and normal classes) discrimination studies such as [11–16]. In [11], discrete wavelet transform (DWT) coefficients and artificial neural networks (ANN) based classification system was used to solve a six-class (squawk, stridor and rhonchus in addition to crackle, normal and wheeze classes) problem on 265 episodes using mean and standard deviation based statistical features. 100% and 94.02% accuracies were achieved for the training and validation sets, respectively. However, when the episode number was increased from 422 to 5786, the general accuracy decreased to 59.15% on the validation set. In [12], the power spectral density (PSD) features, obtained from three lung sound types, were fed into genetic algorithm for feature selection. The selected features were forwarded to multilayer perceptron neural network and 91.7% average accuracy was achieved when 96 subjects were used. In [13], the data was modelled with maximum-likelihood approach and, Hidden Markov Model (HMM) based classification was employed resulting in an average classification rate of 83% on 1544 episodes. In [14], using multilayer perceptron neural network on 20 test epochs, confidence levels of 90%, 87% and 89% were obtained for normal, wheeze and crackle classes, respectively. In [15], a combination of Mel frequency cepstral coefficients (MFCC) and Gaussian mixture model (GMM) were used for classification resulting in 98.75% and 52.5% accuracies for the reference (50 epochs) and cross-validation (24 epochs) set, respectively. Using 225 samples and S-transform (which is a phase corrected version of continuous wavelet transform (CWT) with a scalable Gaussian window) based statistical features (mean and standard deviation), the study in [16] reached overall classification accuracy of 94.99% with k nearest neighbor (k-NN), 96.85% with support vector machine (SVM) and 98.52% with extreme learning machines (ELM) classifiers. A summary of the related works in literature is shown in Table 1.

Feature extraction in a classifier is one of the most critical steps since the features' discrimination ability is more important than their number. The works [12,14] used Fourier Transform (FT) based power spectral density features for classification. Moreover, the study in [12] used an extra feature selection step based on genetic algorithm. These methods, however, lacked localizing waveform features in time domain. In [13,15] MFCC features were extracted where these features were designed to represent human auditory perception and used frequently in speech recognition. However, as depicted in Table 1 the lowest accuracies were obtained using MFCC features, the reason being that lung sound characteristics need not be person specific as human voice [17]. The works of [11,16] used wavelet transform based features where wavelet based features had better time-frequency resolution than the others. However, the wavelets employed in [11] had low Q-factors and this resulted in poor frequency selectivity, which was a dramatic drawback in the modelling of wheeze signals. In [11] a shrinkage denoising technique was additionally used with the performance decreasing severely with increasing number of episodes. In [16], best results were reported in literature; however, database was relatively small and all the episodes had the same, fixed length while crackle and wheeze sounds had different durations in real cases.

In literature, wavelet transform based methods have been successfully employed for feature-extraction and/or de-noising in the pulmonary sounds [7,11,18]. The meaning of the Q-factor in wavelet terminology is the ratio of bandpass filter's center frequency to its bandwidth. In the previous DWT based systems, mostly, constant low Q-factor wavelets having poor frequency-resolution were employed and satisfactory results were obtained in the analysis of piecewise smooth signals. On the other hand, a DWT having better and controllable frequency resolution must be utilized in the analysis of oscillatory signals like wheezes to achieve optimum time-frequency representation of signal of interest. Hence in this study, unlike previous pulmonary signal processing systems, rational dilation wavelet transform (RADWT) [19], whose analysis and synthesis filters' Q-factors can be tuned with respect to the signal of interest, was proposed for feature extraction. Shannon entropy, standard deviation, energy, maximum/minimum, mean and skewness/kurtosis values of each decomposed sub-band were calculated as statistical feature-subsets. Later, these statistical feature-subsets were given into decision tree (DT), Naive Bayes (NB), k-NN, SVM, and ELM models with the final aim of classifying wheeze, crackle and normal lung sounds. Experimental results