Computers in Biology and Medicine 104 (2019) 175–182

Contents lists available at ScienceDirect

Computers in Biology and Medicine

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

Wheeze type classification using non-dyadic wavelet transform based optimal energy ratio technique

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

Keywords:
Respiratory sounds
Pulmonary sounds
Discrimination
Wheezing
Monophonic
Polyphonic

ABSTRACT

Background and objective: Wheezes in pulmonary sounds are anomalies which are often associated with obstructive type of lung diseases. The previous works on wheeze-type classification focused mainly on using fixed time-frequency/scale resolution based on Fourier and wavelet transforms. The main contribution of the proposed method, in which the time-scale resolution can be tuned according to the signal of interest, is to discriminate monophonic and polyphonic wheezes with higher accuracy than previously suggested time and time-frequency/scale based methods.
Methods: An optimal Rational Dilation Wavelet Transform (RADWT) based peak energy ratio (PER) parameter selection method is proposed to discriminate wheeze types. Previously suggested Quartile Frequency Ratios, Mean Crossing Irregularity, Multiple Signal Classification, Mel-frequency Cepstrum and Dyadic Discrete Wavelet Transform approaches are also applied and the superiority of the proposed method is demonstrated in leave-one-out (LOO) and leave-one-subject-out (LOSO) cross validation schemes with support vector machine (SVM), *k* nearest neighbor (*k*-NN) and extreme learning machine (ELM) classifiers.
Results: The results show that the proposed RADWT based method outperforms the state-of-the-art time, frequency, time-frequency and time-scale domain approaches for all classifiers in both LOO and LOSO cross validation settings. The highest accuracy values are obtained as 86% and 82.9% in LOO and LOSO respectively when the proposed PER features are fed into SVM.
Conclusions: It is concluded that time and frequency domain characteristics of wheezes are not steady and hence, tunable time-scale representations are more successful in discriminating polyphonic and monophonic wheezes when compared with conventional fixed resolution representations.

1. Introduction

Auscultation with a traditional stethoscope is a widely used tool with low diagnostic value due to its limited frequency response which attenuates frequencies above 120 Hz [1], its subjectivity which depends on the physician's experience and expertise, and its inability to record sounds for further analysis. The need for a patient specific clinical decision support system has become vital during recent years in the diagnosis of lung disorders with the additional need to decrease healthcare expenses [2,3].

Lung sounds may be categorized into two basic groups: vesicular and adventitious sounds. Adventitious sounds which are usually indicators of various lung diseases are either discontinuous, i.e. crackles, or continuous, i.e. wheezes. Unlike crackles, wheezes are musical and continuous in nature and have narrow representations in frequency domain.

A lung sound segment is accepted as a wheeze, according to American Thoracic Society (ATS) and Computerized Respiratory Sound Analysis (CORSA), if its main frequency is higher than 400 and 100 Hz and its duration is longer than 250 and 100 ms, respectively [4–7]. On the other hand, in the studies of [8,9], reported minimum duration is 80 ms.

Wheezes are closely associated with diseases such as asthma and chronic obstructive pulmonary disease (COPD) [7,10]. The severity of the disease may be related to the duration, number and main frequency of wheezes within a respiration cycle [7,11].

In literature, wheezes along with other adventitious lung sounds classification problems are extensively examined and a summary can be found in Refs. [8,12]. Wheeze/non-wheeze discrimination problem is analysed using various feature extraction techniques such as cepstral

* Corresponding author. 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).

https://doi.org/10.1016/j.compbiomed.2018.11.004
Received 29 March 2018; Received in revised form 3 November 2018; Accepted 5 November 2018
0010-4825/

analysis in Gaussian Mixture Model [13], de-trend and peak detection in the time-frequency domain [14], Mel Frequency Cepstral Coefficients and Gaussian Mixture Model [15], averaging power spectrum components at certain frequencies [16], image processing based spectrogram analysis [17], autocorrelation function based delay-coordinate embeddings [18], spectro-temporal analysis based on auto-regressive averaging [19], wheeze signature in the spectrogram space and music information retrieval [20], Short Time Fourier Transform [21], Ensemble Empirical Mode Decomposition [22], compressive sensing recovered Short Time Fourier Spectrum and Hidden Markov Model [23]. Wheezes can be classified as two types due to their time-frequency behaviours, namely polyphonic and monophonic types. It is observed that polyphonic wheezes are usually caused by the pathology of small airways and monophonic wheezes are caused by the pathology of larger airways [24]. Monophonic (MP) wheezes, comprised of either single pitch frequency or multiple pitch frequencies starting and ending at different times, are believed to stem from single bronchial narrowing and may related to asthma [25–27]. Polyphonic (PP) wheezes, composed of harmonically unrelated multiple pitch frequencies starting and/or ending simultaneously, on the other hand, are supposed to originate from multiple central bronchial compression and are commonly related to COPD [25,26,28]. An MP and a PP wheeze sample represented in time-frequency (TF) domain may be depicted in Fig. 1. Despite advances made in the analysis of lung sounds, discrimination of multiple MP and PP wheezes is still an open problem [29] since both are sinusoidal in nature.

In Ref. [25], it is reported that there are statistically significant differences between MP and PP wheezes of the same pathology (asthma or COPD) using wavelet based features, paving the way for classification studies. In Ref. [25], continuous wavelet transform is applied to monophonic and polyphonic wheezes with the aim of extracting nonlinear characteristics of wheezes and associating these characteristics with the underlying pathology. Studies on MP-PP wheeze classification problem are very few in literature. In Ref. [30], only nine monophonic-polyphonic wheezes are detected using spectrogram based peak continuity, resulting in 89% accuracy. In Ref. [31], 92% F_1 score (which is the harmonic mean of precision and recall rates) is reached using dominance spectrogram based on instantaneous frequency on normal,

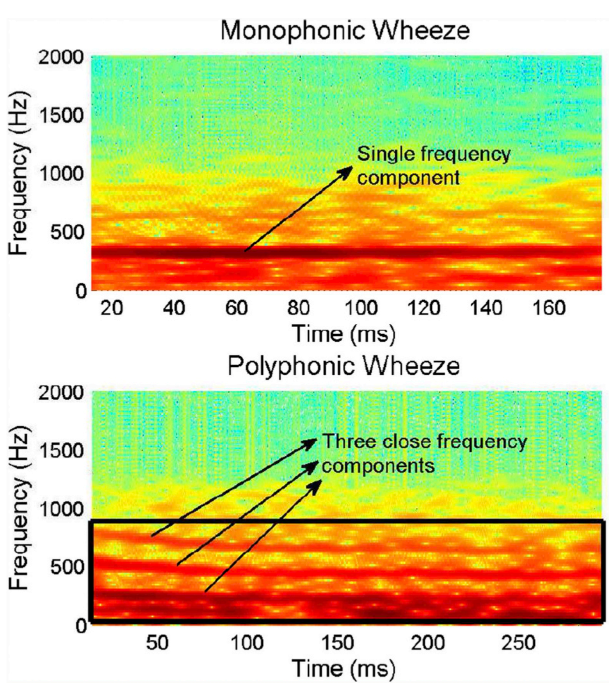

Fig. 1. Time-frequency domain representation of MP (top) and PP (lower) wheezes (Best viewed in color).

monophonic, polyphonic and stridor classes whereas 72% F_1 score is obtained using the classical spectrogram on various numbers of wheezes, ranging between 70 and 155. A recent work [32] using time domain based higher order statistics (with a genetic algorithm based feature selection step) reaches 91% classification accuracy using 102 wheezing sounds. In Ref. [33], 89% accuracy is obtained on 140 wheezes using dyadic Discrete Wavelet Transform (DWT) based features with an additional feature selection step.

Our previous study [34] on this classification problem has led us to explore robust and discriminative features for PP wheezes. In this work, unlike previous studies which use fixed TF resolution based on Fourier transform, we propose a tunable Rational Dilation Wavelet Transform (RADWT) based technique to discriminate MP and PP wheezes by using localized energy peaks which are calculated from wavelet coefficients. The effectiveness of the proposed feature extraction method has been demonstrated using three approaches; i) The accuracies of the classifiers using the proposed tunable features and the fixed RADWT parameters are compared. ii) The feature set extracted using the proposed method is fed to support vector machine (SVM), extreme learning machine (ELM) and k-nearest neighbor (k-NN) classifiers with a view to measure its robustness. iii) Previously suggested feature extraction methods for MP/PP wheeze classification such as dyadic discrete wavelet transform (DWT) [33], Quartile Frequency Ratios (QFR) [34], Mean Crossing Irregularity (MCI) [34], feature level fusion of QFR-MCI and Multiple Signal Classification (MUSIC) [35] are applied to the same dataset for a comparison with the results of the proposed method both in LOO and LOSO cross validation settings. Properties of the database are described in Section 2.1, while Section 2.2 gives the details of the wavelet based method and Section 2.3 introduces the proposed method. Section 3 and Section 4 consist of experimental results and discussion, respectively whereas Section 5 is the conclusion.

2. Methods

2.1. Data acquisition and database

The 14-channel data acquisition system [36] designed at Boğaziçi University Lung Acoustics Laboratory (BU-LAL) was utilized to record wheeze sounds. Sampling rate was 9600 Hz with each data recording session lasting 15 s. Each subject had a nose clip and a flow-meter was employed to measure airflow. An informed consent was taken from all subjects before data acquisition. The data acquisition procedure had the consent of the second Ethical Committee on Clinical Research of Istanbul (in accordance with the Declaration of Helsinki). Wheeze sound was collected from asthma and COPD patients who were under treatment at the Istanbul Yedikule Teaching Hospital for Chest Diseases and Thoracic Surgery. Database was comprised of seven male and four female subjects at the age of 52 ± 19. The subjects breathed spontaneously and were not forced to breathe at a predetermined flow rate in order to be coherent with real clinical settings. Resulting average flow rate was 0.72 ± 0.34 L/s. The mean and standard deviation of the pulmonary function test parameter scores were 66 ± 32 for FVC, 52 ± 42 for FEV_1 and 61 ± 21 for FEV_1/FVC. Wheeze sounds were labeled by visual verification of time expanded waveforms and auditory inspection by an expert. The database consisted of 147 MP and 153 PP wheezes, where the duration of each segment was at least 80 ms, to be consistent with literature. The samples were measured from different locations on the chest wall of the subjects to enrich the database in terms of spatial variability.

2.2. Rational Dilation Wavelet Transform (RADWT)

According to definition, wheezes that occur with a single peak or with the harmonics of a single basal peak are called MP wheezes, while those with variable peaks that differ in harmonics are called PP wheezes [29]. Due to the similarities between MP and PP wheezes,

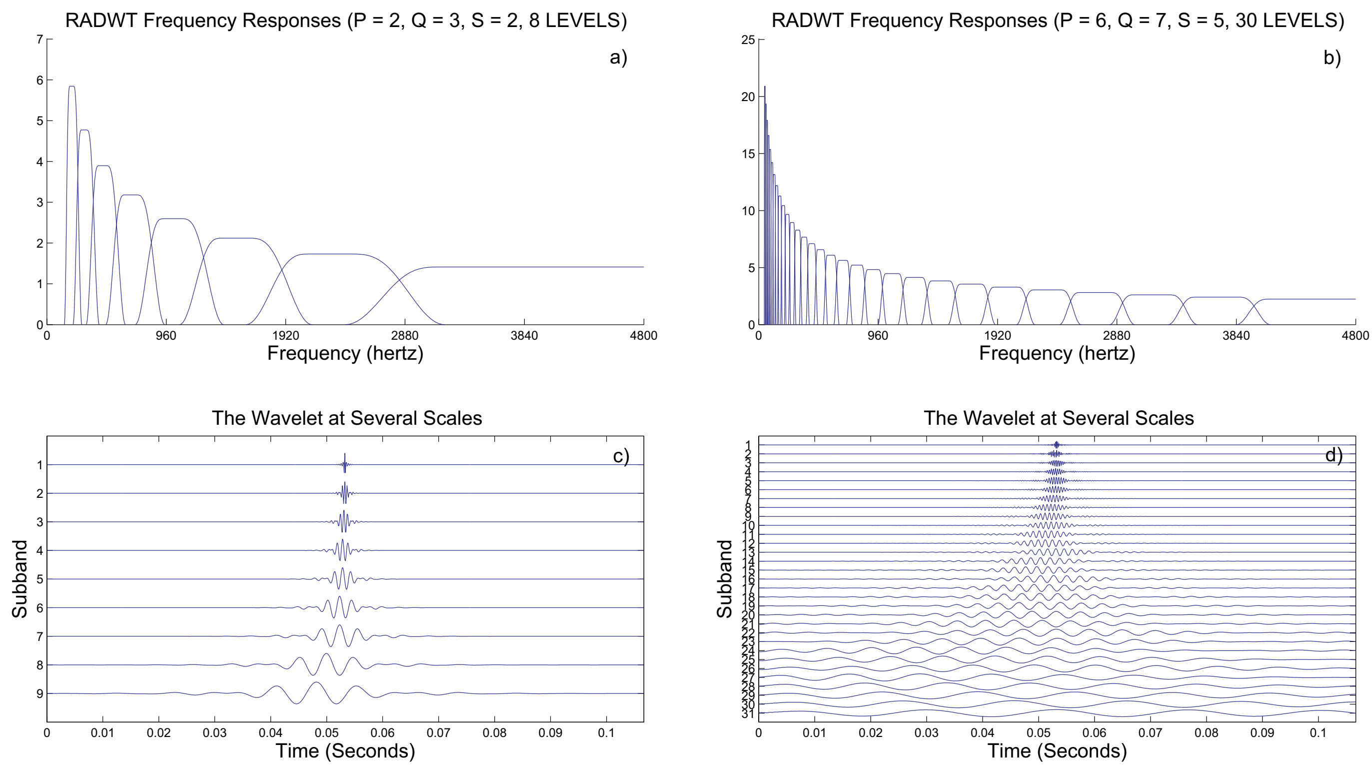

Fig. 2. Wavelet and frequency responses of two different (low-Q, a-c and high-Q factor, b-d) p, q, s and J parameters.

discriminating multiple MP wheezes from PP wheezes is still an open and important problem. When the severity of the pathology is very strong, a fundamental (basal peak with high energy) signal can occur with accompanying harmonics (peaks with lower energy) in MP wheezes [25,29]. This MP pattern may be confused with PP wheezes in which various peaks with relatively close energies show up. In order to discriminate the MP and PP wheezes in time-scale domain, a wavelet transform, in which the frequency selectivity of the sub-bands can be adjusted, is needed. Therefore in this study, the Rational Dilation Wavelet Transform (RADWT) [37], which has finer and adjustable frequency resolution with acceptable redundancy, is proposed as a suitable feature extractor for processing lung sounds.

The RADWT [37] is a frequency-domain (FFT based) design transform which does not employ rational transfer functions and offers greater design flexibility. Moreover, the RADWT is a rational (based on non-dyadic dilations), fully discrete, approximately shift-invariant and easily invertible transform. The non-dyadic (rational) behaviour of the RADWT yields a range of Q-factors and redundancy factors. In the RADWT, the Q-factor of wavelets, which controls the frequency resolution of the transform, is built upon three positive integers p, q and s satisfying $1 \le p < q$ and $p/q + 1/s \ge 1$, where p and q are co-prime, and their ratio (q/p) is the dilation factor of the wavelet transform.

In RADWT, the relation between the scaling $(\phi(t))$ and wavelet $(\psi(t))$ functions, where $h_0(n)$ and $g_0(n)$ are low and high pass filters respectively, can be given as,

$$\phi(t) = (q/p)^{1/2} \sum_{n\in\mathbb{Z}} h_0(n)\phi\left(\frac{q}{p}t - n\right) \tag{1}$$

$$\psi(t) = (q/p)^{1/2} \sum_{n\in\mathbb{Z}} g_0(n)\phi\left(\frac{q}{p}t - n\right) \tag{2}$$

Mathematically, the frequency responses of $h_0(n)$ $(H_0(\omega))$ and $g_0(n)$ $(G_0(\omega))$ are given as [37],

$$H_0(\omega) = \begin{cases} \sqrt{pq} & \omega \in \left[0, \left(1 - \frac{1}{s}\right)\frac{\pi}{q}\right], \\ \sqrt{pq}\,\theta\left(\frac{\omega - a}{b}\right) & \omega \in \left[\left(1 - \frac{1}{s}\right)\frac{\pi}{q}, \frac{\pi}{q}\right], \\ 0 & \omega \in \left[\frac{\pi}{q}, \pi\right], \end{cases} \tag{3}$$

$$G_0(\omega) = \begin{cases} 0 & \omega \in \left[0, \left(1 - \frac{1}{s}\right)\pi\right], \\ \sqrt{s}\,\theta_c\left(\frac{\omega - pa}{pb}\right) & \omega \in \left[\left(1 - \frac{1}{s}\right)\frac{\pi}{q}, \frac{p}{q}\pi\right], \\ \sqrt{s} & \omega \in \left[\frac{p}{q}\pi, \pi\right], \end{cases} \tag{4}$$

where

$$a = \left(1 - \frac{1}{s}\right)\frac{\pi}{p},\ b = \frac{1}{q} - \left(1 - \frac{1}{s}\right)\frac{1}{p} \tag{5}$$

the transition function $\theta(\omega)$ is,

$$\theta(\omega) = \frac{1}{2}(1 + cos(\omega))\sqrt{2 - cos(\omega)} \quad \text{for} \quad \omega \in [0, \pi] \tag{6}$$

and $\theta_c(\omega)$ is

$$\theta_c(\omega):=\sqrt{1 - \theta^2(\omega)} \tag{7}$$

The transition function, $\theta(\omega)$, which is used to construct the transition bands of $G_0(\omega)$ and $H_0(\omega)$, originates from Daubechies orthonormal wavelet filters with two vanishing moments. As it can be seen from above equations, the bandwidth, center frequency and transition bands of high-pass and low-pass filters are determined by using p, q and s values. As the q/p ratio approaches unity, higher number of decomposition levels are needed. Therefore, the number of subbands (J) must also be considered as an important parameter in the analysis. As seen in

Table 1
Various p, q, s and J values used in analysis.

Set #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
p	2	3	4	5	5	5	6	6	7	7	7	8	8	8	8	8	9	10	10	10	11	11
q	3	4	5	6	6	6	7	7	8	8	8	9	9	9	9	9	10	11	11	11	12	12
s	2	2	3	4	5	5	5	6	5	6	7	3	4	5	6	7	5	6	7	8	7	8
J	8	10	15	20	25	30	30	35	35	35	40	35	35	35	35	40	35	40	45	50	45	50

Fig. 2 when the q/p ratio approaches one (right side, b-d) the frequency selectivity of the subbands gets better for middle and high frequencies when compared with the low Q-factor analysis (left side, a-c) in which the frequency selectivity of the subbands is limited.

2.3. Proposed optimal peak-energy-ratio parameter selection method

In order to discriminate wheeze types using energy localizations of the harmonics, the peak energy points of these harmonics must be determined. However, the location in the frequency axis, the amplitude and the bandwidth of these peaks differ for each sample due to the physiological properties of the lung and the mechanism of the pathology. This results in a need for an optimal algorithm that can locate peaks in TF domain for processing MP and PP wheezes. In the proposed algorithm, the RADWT is applied to MP and PP wheezes by using a set of various p, q, s and J values, which are given in Table 1, with an aim to achieve an optimum representation. These 22 different parameter sets are used as suggested in Ref. [37] and each different parameter set induces a unique time-frequency representation. Using the same predefined set of parameters offers a more objective comparison basis of our proposed method. In this sense, optimum representation is defined as two distinct and non-consecutive peaks, where one belongs to either basal peak in MP wheezes or first peak in PP wheezes and the other belongs to either a weak harmonic in MP wheezes or a second peak in PP wheezes. In this context, energy of a peak indicates the localized sub-band energy of wavelet coefficients obtained from the decomposition of MP or PP wheezes. The normalized energy of each sub-band is calculated using the following steps:

(i) finding the absolute value of wavelet coefficients,
(ii) taking the square of these terms,
(iii) summing the squared terms giving the relevant sub-band energy and
(iv) dividing the calculated energy of relevant sub-band by the total signal energy for normalization.

The selected first and second peaks should not be consecutive, since consecutive peaks can occur due to energy leakage between successive bandpass filters of the wavelet transform as shown in Fig. 2 (upper side, a-b).

The first selected peak, therefore, corresponds to the highest energy carrying peak and the second is the second highest energy carrying non-consecutive peak.

Then a metric termed as the peak-energy-ratio (PER) is defined as,

$$\mathrm{PER} = \frac{\text{Energy of the first selected peak}}{\text{Energy of the second selected peak}} \tag{8}$$

In Fig. 3, energy distributions of wheezes across sub-bands and peaks are represented. For the RADWT analysis, 22 various p, q, s and J value combinations, which are used in the decomposition stage of each wheeze sample, are given in Table 1. For each set, two distinct and non-consecutive peaks are found and the PER is calculated. As a result, for each signal, 22 different PER values are obtained. The minimum PER value is selected as the indicator of the best representation because it means that the two peaks are correctly located while preserving maximum amount of their energies. In order to quantify the performance of the proposed method, the chosen minimum PER metrics are employed as features for discriminating MP and PP wheezes. The flowchart of the proposed algorithm can be found in Fig. 4.

As a fair quantitative validation of the proposed PER feature, three classifiers, k-NN, SVM and ELM are applied to both PER features and previously used features with the aim of MP/PP wheeze discrimination. k-NN is employed as a fast classifier which predicts the class of a new test data by using the class labels of its k nearest neighbors in the training feature space. k value is chosen from a set changing from 1 to 10 empirically, and Euclidean/city-block distances are used for measuring the distance between the test and train samples. SVM [38], in which an optimal separating hyperplane is employed to generate a maximum margin between MP and PP wheeze samples, is used as the second classifier. The optimum cost (C) and the kernel width parameters (γ) are set by using grid search for linear and radial basis function (RBF) kernels [39].

Lastly, as a non-iterative method, ELM [40] with single hidden layer feed forward neural networks is employed. In ELM, the input weights are chosen randomly whereas the output weights are determined analytically. Due to its non-iterative character, ELM achieves a high generalization performance with an improved learning speed. For the ELM case, the number of hidden neurons is set empirically from a range of 1–100 and sigmoid function is used for activation.

2.4. Previously used MP/PP wheeze discrimination approaches

MP/PP wheeze classification, which is a critical step in the diagnosis of asthma [25–27] and COPD diseases [25,26,28], is a novel and open problem [25,29] in pulmonary sound analysis literature [30–35]. The feature extraction methods that are employed in MP/PP wheeze classification are mostly based on time domain [32,34] and frequency domain [30,31,33–35] representations of lung signals. In our study, in order to validate the performance of the proposed RADWT based PER feature, previously suggested wavelet based methods are also applied to our data-set with the aim of feature extraction. As the first comparison method, Quartile Frequency Ratios (QFR [34]), which are the percentile frequencies denoting the exact frequencies where the cumulative power reaches 25, 50, 75 and 90% of total power in power spectral density (PSD) of each sample, are calculated. In this approach, the aim is to measure the power concentration of peaks in PSD which is more localized for a single oscillation in MP wheezes and more distributed for multiple oscillations in PP wheezes. In the calculation of PSD, Welch method [41] using 50% overlapping for 256 points Hamming windows is employed and Fast Fourier Transform size is 256. As the second comparison, mean crossing irregularity (MCI [34]) measure is calculated to quantify the periodicity of MP and PP wheezes. In MCI, the successive mean crossing intervals are defined as a random variable X, and the standard deviation of this X, which is normalized by the mean of X, is used as a feature. Additionally, as the third comparison approach, the feature level fusion of QFR and MCI features is employed in order to benefit from both time and frequency domain information. The dyadic DWT is a well known signal decomposition method in which the frequency resolution is changed with the powers of two (dyadic). In Ref. [33], six statistical features which were extracted from wavelet coefficients are fed to an artificial neural network with the aim of wheeze type discrimination, and therefore in our study the dyadic DWT

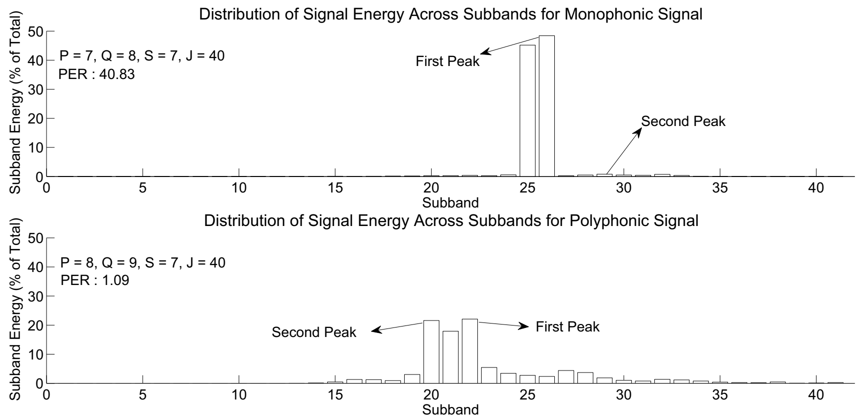

Fig. 3. Energy distribution of MP and PP wheezes.

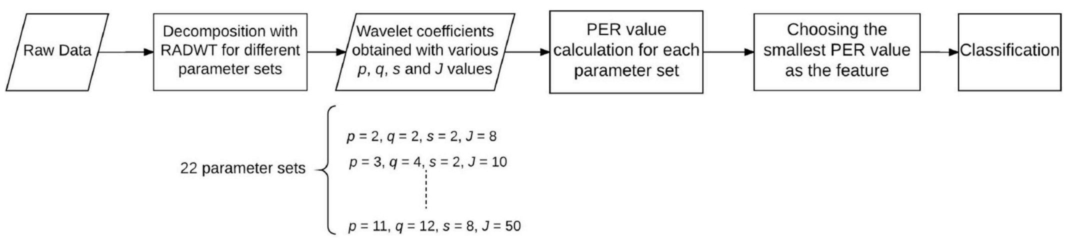

Fig. 4. The flowchart of the proposed algorithm on MP-PP wheeze classification.

based statistical features are extracted from our dataset and used as the fourth comparison method. According to best parameters reported in Ref. [33], bi-orthogonal 1.5 wavelet type with seven details and one approximation is employed and energy, mean and standard deviation statistical features are computed from wavelet coefficients and fed into classifiers after normalization. As a baseline method [42], statistical features based cepstral method is also experimented on our dataset. In Ref. [42], the statistical features (mean and standard deviation) are extracted from Mel Frequency Cepstral Coefficients (MFCCs) with the order of 20 which are computed using 20, 50 and 70 ms length frames with 50% overlap and 1024-point Fast Fourier Transform (FFT). After the normalization of extracted 20 features between [− 1,+ 1], classification is performed in a leave-one-out cross validation scheme. Up to this point, the time and frequency/scale based features are extracted in the form to be used in a supervised learning paradigm. As a different approach, Multiple Signal Classification (MUSIC [35]) algorithm, which does not need a pre-training step and can be used in an unsupervised way, is tested as the last comparison method. The MUSIC algorithm can be employed to detect multiple frequencies by using the power density spectrum of the processed signal even if the signal-to-noise ratio is very low (in high-level noise) [43]. In MP/PP wheeze classification problem, we use the MUSIC algorithm to detect the mono (in MP wheeze) or poly (in PP wheeze) frequencies that form the signal of interest. Later, the detected frequencies are counted and the signal is labeled as MP if the number of the detected peak frequency number is one or PP if the number of the detected peak frequencies is two. In MUSIC algorithm, the maximum number of possible sinusoidal frequencies must be given as a pre-defined parameter.

3. Results

The performance of the proposed method is initially evaluated by examining the calculated PER values. Two non-consecutive distinct energy peaks for both MP and PP wheezes are depicted in Fig. 3 with the corresponding p, q, s, J and PER values. It is seen that for MP wheezes high PER values are obtained, whereas relatively small PER values are obtained for PP wheezes. In the PP case, each peak corresponds to the activity of small airways whose energy capacity is close while in the MP case the highest peak corresponds to a large airway activity and latter peak mostly corresponds to a normal lung activity or a leakage which has very small amount of energy. Therefore, in the PP case, the ratio of distinct small airway energies results in small PER values whilst in the MP case the ratio of a relatively large airway energy and a small normal breath sound or leakage energy results in high PER values. In Fig. 5, whisker plot of PER values with respect to wheeze types is depicted where the PER values correspond to the minimum values of the tested data samples after all the calculated PER values using the parameter sets of Table 1 are scanned. The median values of PER metrics for monophonic and polyphonic wheezes are clearly far from each other with values of 22.01 and 2.41, respectively, and this difference shows the discriminative power of PER metric in the MP and PP classification problem.

In order to validate the need of an optimal method, the behaviour of p, q, s and J combinations (parameter sets) having minimum PER values in the proposed method is also investigated and it is seen that a fixed parameter set did not perform distinctively better than the other parameter sets. For each individual parameter set in Table 1, a significant number of wheeze samples having the minimum PER values was obtained and the distribution of wheeze sample number over various parameter sets was nearly homogeneous as shown in Fig. 6.

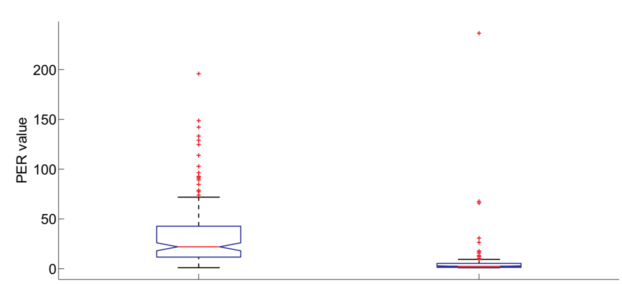

Fig. 5. Comparison of PER values with respect to wheeze types when the optimum parameters are employed.

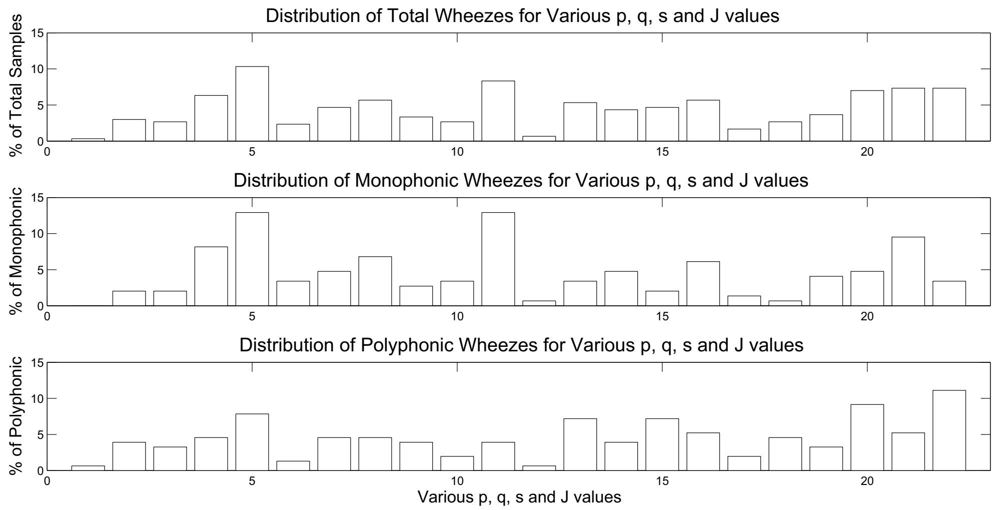

Fig. 6. Histogram of total, monophonic and polyphonic wheezes with minimum PER values obtained from the p, q, s and J parameter combinations given in Table 1.

Table 2
Classification accuracies (in %) of fixed p, q, s and J parameters for 22 combinations as defined in Table 1 using different classifiers. Bold values indicate maximum accuracy of each row.

Classifier	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
SVM Linear	56	50.7	57	54.3	50.7	56.3	54.3	50.3	52.7	51.3	74.7	60.7	63	61.7	54	54.3	56	**76.7**	76.3	76	57	56
SVM RBF	56	54.3	62	72.3	76.3	77.3	75	79	77.3	76.3	79.3	75.3	79.3	76.3	79.3	77.7	75.7	80.3	**81.7**	80	76.3	75
k-NN	54.7	52.3	61	69.3	72.7	71	69.3	75	73.7	71.3	75.7	69.3	75.7	72.3	77	68.3	71.3	78.7	**79**	76	71	70.3
ELM	55.3	53	61.3	71	74.3	76	68.7	79.3	78	71	79.3	72.7	79	77.7	78.7	72.7	74.3	79.7	**81.3**	80.3	72.7	74.3

Additionally, in order to evaluate the performance of the proposed method, the classification accuracies of fixed p, q, s and J parameter sets (the same order as in Table 1) on the whole dataset are also given in Table 2. In each column, the given accuracies are obtained when a fixed (same) parameter set is employed during PER value calculation process for the whole dataset in contrast to proposed optimal method. It is shown that the highest classification accuracy (81.7%) is obtained with 19^{th} parameter set in Table 1 using SVM radial basis function (RBF) kernel in LOO setting. This is due to the fact that in Table 1, on the left-side, the relatively lower Q-factor combinations and on the right side relatively higher Q-factor combinations are represented therefore frequency resolution increases from left to right.

Therefore, it is concluded that an optimal system is needed for optimum localization of different peaks due to subject specific TF properties of wheezes. Additionally, the accuracy of the SVM classifier obtained with the proposed method is 82.6% and 86% respectively, when the linear and RBF kernels are employed in a LOO scheme. This shows that PER metric can be used as an indicator for discriminating MP and PP wheezes and better TF representation can be achieved with the proposed method when compared to fixed parameter method. Since RBF kernel performs better than linear kernel, in Table 3, only SVM with RBF kernel results are presented. A statistical test on the classifier results of the proposed method is performed with paired samples t-test, with p-values (the largest 0.0141) providing $p < 0.05$ condition.

In order to make a fair comparison among related methods and the proposed method, leave-one-subject-out (LOSO) cross validation results are given in Table 3 to provide person independent classification results. The proposed method achieved the best accuracy results (82.9% with SVM) with all classifiers in LOSO setting which shows the generalization ability of the proposed method.

In order to show the effect of number of features on classifier learning model, computation time results (in milliseconds) for SVM classifier in a LOO setting which compare the computational complexity of the all tested methods, are also shown in Table 3. The proposed method has lower computation time (26 ms.) for classifier learning model when compared with related works since the proposed method employs merely one feature.

Table 3
Performance comparison of related works on MP-PP classification problem with the proposed method using various classifiers, cross validation techniques and computation time (in ms.) of the classifier learning model.

Related Works	Features	LOO CV			LOSO CV			Comp.Time
		SVM	k-NN	ELM	SVM	k-NN	ELM	
Hashemi et al. [33]	Dyadic DWT	77	74.7	76.3	74.2	70.6	71.8	105
Ulukaya et al. [34]	Time domain	76.3	74.3	76.7	72.9	69.3	71.2	26
Ulukaya et al. [34]	Frequency domain	62.3	59	61.7	59.9	55.2	58.9	27
Ulukaya et al. [34]	TF domain	78	75	77.3	76.1	72.5	75.3	32
Sengupta et al. [42]	Statistical cepstral	81	78	79.7	78.3	74.6	76.2	202
Proposed	Non-dyadic wavelet	**86**	84.3	85.3	**82.9**	80.6	81.9	26

Bold values indicate highest accuracies of cross validation methods over all classifiers.

Table 4
Confusion matrix and individual accuracies (in %) of the proposed method on MP-PP wheeze classification. True labels are in the first column of the table and estimated labels are in the first row of the table.

Labels	MP	PP
MP	**81.6**	18.4
PP	9.1	**90.9**

Bold values indicate highest true classification rates.

4. Discussion

There are relatively few groups [30–33] working on MP/PP wheeze classification, although it is considered a critical problem [25,28,29] in lung sound analysis. Therefore, a publicly available dataset, which can be utilized to compare the performance of the proposed method in relation to previous works, does not exist. Our proposed method is based on wavelet transform therefore we have compared our results with our previous works and other wavelet and Fourier based studies used in the wheeze type classification. So the methods in Refs. [33–35,42] are tested on our dataset for comparison. As represented in Table 3, the proposed method using SVM with RBF kernel performs significantly better than related works in literature. As seen in Table 3, proposed method achieved higher accuracies than related works even using k-NN classifier which shows the discriminative feature extraction ability of the proposed method both in LOO and LOSO setting.

Dyadic DWT [33], which has limited frequency resolution at higher frequencies, performed poorly when dealing with oscillatory types of signals (high Q-factor) obtaining an accuracy of 77% for LOO and 74.2% for LOSO cross validation as shown in Table 3. Frequency domain approach QFR, time domain approach MCI and their feature level fused time-frequency domain approach perform with 62.3%, 76.7% and 78% accuracies for LOO and 59.9%, 72.9% and 76.1% accuracies for LOSO, respectively. Additionally, as a different approach, multiple signal classification (MUSIC) algorithm [35] which does not need to employ a classifier (unsupervised method) is also tested on our dataset and 59.7% accuracy is obtained. Statistical features [42] extracted from MFCCs showed poor performance (81% accuracy for LOO, 78.3% accuracy for LOSO) as shown in Table 3. This is probably due the fact that MFCC method based on Fourier Transform which has fixed time-frequency resolution, may not be able to capture the discriminative properties of the oscillatory types of sounds. From these experiments, we observed the following inferences; i) the assumption in the use of QFR is that in MP wheezes the total power is concentrated around a single frequency, but in PP wheezes the total power is expected to spread over several frequencies. Therefore, the percentile frequencies are expected to be closer to each other for MP wheeze episodes as compared to PP wheezes. However in reality, as depicted in Fig. 3, the frequency of the two periodic waveforms may be very close, resulting in a localized energy for PP wheezes and this closeness reduces the discriminative power of this feature. ii) the assumption in the use of MCI is that although both MP and PP wheezes are periodic in time domain, MCI is expected to be lower for an MP wheeze episode since it is composed of a single main frequency component. However, it is observed that the frequency of the main component in MP wheeze is not fixed and can get higher or equal frequency values when compared with PP case, which may cause confusion in classifier learning models. iii) the assumption in the use of MUSIC algorithm is that it can detect the single (for MP wheezes) and multiple (for PP wheezes) periodic components in pseudo-power spectrum of signals. However, in some PP cases, the second or third periodic components have relatively lower power. Therefore, in these cases, only one peak in the pseudo-spectrum is found like in the MP cases. Generally, in the light of the foregoing arguments, it can be reasoned that the proposed method achieves better time-frequency representation and generalization ability than other Fourier, time-domain and dyadic Discrete Wavelet Transform (DWT) based methods due to its Q-parameter tunability property [33–35,42].

Moreover, to show the improvement achieved with the application of the proposed approach, the classification results of our method are compared with the classification results obtained when fixed p, q, s and J parameters (in this context fixed means, the p, q, s and J parameters are not changed during the feature extraction and classification procedure in contrast to the proposed method where a search is made for the optimum set of parameters based on the minimum PER value) are employed in the wavelet analysis. It is observed that approximately 4% improvement is obtained by using the proposed method (86% accuracy) when it is compared with the best fixed (81.7% accuracy in Table 2) parameter set ($p = 10$, $q = 11$, $s = 7$ and $J = 45$).

Individual classification accuracies obtained with the proposed approach for MP and PP wheezes are 81.6% and 90.9% (the harmonic mean of these two values gives the overall accuracy as 86%) respectively as given in Table 4 in more detail as a confusion matrix. This demonstrates the effectiveness of the proposed method for representing PP wheezes when compared with other wavelet based methods. In the classification method context, it is shown that the RBF kernel gives better results than linear kernel, and it is suggested to use non-linear kernel when the number of features (only one PER value in the proposed approach) is small [39].

When the advantages of the proposed method over previous studies are examined, the following points may be noted. In the proposed method with only one feature (PER value), high accuracy values in MP and PP wheeze classification problem are obtained, while in Ref. [32] 166 features and in Ref. [33] 15 features were used. Additionally, as opposed to the proposed method, to increase the classification performance, in Refs. [32,33], an extra feature selection step was employed. In the proposed method, the number of wheeze samples (300 segments) used is significantly higher than in the other works given in literature (in Ref. [30] 9, in Ref. [31] 155, in Ref. [32] 102 and in Ref. [33] 140 segments). As a final remark, the proposed method has lower computational cost in the classification step when compared with the other previously suggested methods since their classifiers need a higher number of features to train their models.

5. Conclusion

Monophonic (MP) and/or polyphonic (PP) wheezes in a respiration cycle occur in pathologies such as asthma and chronic obstructive pulmonary disease. When the time-frequency (TF) characteristics of MP and PP wheezes are investigated, it is observed that the MP wheezes are composed of signal components which have single pitch frequency or multiple pitch frequencies starting and ending at different times. On the other hand, PP wheezes are composed of signal components which have harmonically unrelated multiple pitch frequencies and have the same starting and/or ending points. In this study, different TF behaviours of MP and PP wheezes have led to a non-dyadic wavelet based feature extraction method which is proposed to solve the MP-PP wheeze classification problem. The performance of the proposed method is compared with traditional time, frequency, TF and dyadic DWT based feature extraction methods quantitatively by using various classification algorithms, and the results verify the superiority of the proposed method over other wavelet based feature extraction methods in wheeze type classification literature in terms of accuracy, feature number and hence computational time. Moreover, it is inferred that using merely one feature decreases the computational complexity of the classification model, making the proposed algorithm a strong candidate for real time medical decision support systems.

Conflicts of interest

No conflict of interest is declared.

Acknowledgments

This work was supported by Boğaziçi University Research Fund, Turkey under grant number 16A02D2. S. Ulukaya was supported by the Ph.D. scholarship (2211) from Turkish Scientific and Technological Research Council (TUBITAK), Turkey.

References

[1] N. Gavriely, D.W. Cugell, Breath Sounds Methodology, CRC Press, 1995.
[2] S. Leng, R. San Tan, K.T.C. Chai, C. Wang, D. Ghista, L. Zhong, The electronic stethoscope, Biomed. Eng. Online 14 (2015) 1–37.
[3] I. Sen, M. Saraclar, Y.P. Kahya, A comparison of SVM and GMM-based classifier configurations for diagnostic classification of pulmonary sounds, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 62 (2015) 1768–1776.
[4] A.T. Society, Updated nomenclature for membership reaction, American Thoracic Society News 3 (1977) 5–6.
[5] A.R.A. Sovijarvi, F. Dalmasso, J. Vanderschoot, L.P. Malmberg, G. Righini, S.A.T. Stoneman, Definition of terms for applications of respiratory sounds, Eur. Respir. Rev. 10 (2000) 597–610.
[6] A.R.A. Sovijarvi, L.P. Malmberg, G. Charbonneau, J. Vanderschoot, F. Dalmasso, C. Sacco, M. Rossi, J.E. Earis, Characteristics of breath sounds and adventitious respiratory sounds, Eur. Respir. Rev. 10 (2000) 591–596.
[7] N. Meslier, G. Charbonneau, J.L. Racineux, Wheezes, Eur. Respir. J. 8 (1995) 1942–1948.
[8] R.X.A. Pramono, S. Bowyer, E. Rodriguez-Villegas, Automatic adventitious respiratory sound analysis: a systematic review, PloS One 12 (2017) e0177926.
[9] H. Pasterkamp, S.S. Kraman, G.R. Wodicka, Respiratory sounds: advances beyond the stethoscope, Am. J. Respir. Crit. Care Med. 156 (1997) 974–987.
[10] K.N. Priftis, OC-11 the child with recurrent wheezing, Arch. Dis. Child. 102 (2017) A4–A5.
[11] R.P. Baughman, R.G. Loudon, Lung sound analysis for continuous evaluation of airflow obstruction in asthma, Chest 88 (1985) 364–368.
[12] F.G. Nabi, K. Sundaraj, C.K. Lam, S. Sundaraj, R. Palaniappan, Wheeze Sound analysis using computer-based techniques: a systematic review, Biomed. Eng.-Biomed. TE (2017) .
[13] J.-C. Chien, H.-D. Wu, F.-C. Chong, C.-I. Li, Wheeze detection using cepstral analysis in Gaussian mixture models, Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE, IEEE, 2007, pp. 3168–3171.
[14] S.A. Taplidou, L.J. Hadjileontiadis, Wheeze detection based on time-frequency analysis of breath sounds, Comput. Biol. Med. 37 (2007) 1073–1083.
[15] M. Bahoura, Pattern recognition methods applied to respiratory sounds classification into normal and wheeze classes, Comput. Biol. Med. 39 (2009) 824–843.
[16] Z. Dokur, Respiratory sound classification by using an incremental supervised neural network, Pattern Anal. Appl. 12 (2009) 309.
[17] R. Riella, P. Nohama, J. Maia, Method for automatic detection of wheezing in lung sounds, Braz. J. Med. Biol. Res. 42 (2009) 674–684.
[18] S. Emrani, T. Gentimis, H. Krim, Persistent homology of delay embeddings and its application to wheeze detection, Signal Processing Letters, IEEE 21 (2014) 459–463.
[19] F. Jin, F. Sattar, D.Y. Goh, New approaches for spectro-temporal feature extraction with applications to respiratory sound classification, Neurocomputing 123 (2014) 362–371.
[20] L. Mendes, I. Vogiatzis, E. Perantoni, E. Kaimakamis, I. Chouvarda, N. Maglaveras, V. Tsara, C. Teixeira, P. Carvalho, J. Henriques, et al., Detection of wheezes using their signature in the spectrogram space and musical features, Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE, IEEE, 2015, pp. 5581–5584.
[21] P. Bokov, B. Mahut, P. Flaud, C. Delclaux, Wheezing recognition algorithm using recordings of respiratory sounds at the mouth in a pediatric population, Comput. Biol. Med. 70 (2016) 40–50.
[22] M. Lozano, J.A. Fiz, R. Jané, Automatic differentiation of normal and continuous adventitious respiratory sounds using ensemble empirical mode decomposition and instantaneous frequency, IEEE Journal of Biomedical and Health Informatics 20 (2016) 486–497.
[23] D. Oletic, V. Bilas, Asthmatic wheeze detection from compressively sensed respiratory sound spectra, IEEE Journal of Biomedical and Health Informatics 22 (2018) 1406–1414.
[24] V.C. Broaddus, R.C. Mason, J.D. Ernst, T.E. King, S.C. Lazarus, J.F. Murray, J.A. Nadel, A. Slutsky, M. Gotway, Murray & Nadel's textbook of respiratory medicine e-book, Elsevier Health Sciences, 2015.
[25] S. Taplidou, L.J. Hadjileontiadis, Analysis of wheezes using wavelet higher order spectral features, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 57 (2010) 1596–1610.
[26] P. Forgacs, The functional basis of pulmonary sounds, Chest 73 (1978) 399–405.
[27] M. Sarkar, I. Madabhavi, N. Niranjan, M. Dogra, Auscultation of the respiratory system, Ann. Thorac. Med. 10 (2015) 158.
[28] C. Jácome, A. Oliveira, A. Marques, Computerized respiratory sounds: a comparison between patients with stable and exacerbated COPD, The Clinical Respiratory Journal 11 (2017) 612–620.
[29] Y. Nagasaka, Lung sounds in bronchial asthma, Allergol. Int. 61 (2012) 353–363.
[30] A. Jain, J. Vepa, Lung sound analysis for wheeze episode detection, 30th International Conference of the IEEE-EMBS, 2008, pp. 2582–2585.
[31] F. Jin, S. Krishnan, F. Sattar, Adventitious sounds identification and extraction using temporal–spectral dominance-based features, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 58 (2011) 3078–3087.
[32] R. Naves, B.H.G. Barbosa, D.D. Ferreira, Classification of lung sounds using higher-order statistics: a divide-and-conquer approach, Comput. Methods Progr. Biomed. 129 (2016) 12–20.
[33] A. Hashemi, H. Arabalibiek, K. Agin, Classification of wheeze sounds using wavelets and neural networks, International Conference on Biomedical Engineering and Technology, vol. 11, IACSIT Press, 2011, pp. 127–131.
[34] S. Ulukaya, I. Sen, Y.P. Kahya, Feature extraction using time-frequency analysis for monophonic-polyphonic wheeze discrimination, 37th International Conference of the IEEE-EMBS, 2015, pp. 5412–5415.
[35] S. Ulukaya, I. Sen, Y.P. Kahya, A novel method for determination of wheeze type, Signal Processing and Communications Applications Conference (SIU), 23th IEEE, 2015, pp. 2001–2004 2015.
[36] I. Sen, Y.P. Kahya, A multi-channel device for respiratory sound data acquisition and transient detection, 27th International Conference of the IEEE-EMBS, 2005, pp. 6658–6661.
[37] I. Bayram, I.W. Selesnick, Frequency-domain design of overcomplete rational-dilation wavelet transforms, IEEE Trans. Signal Process. 57 (2009) 2957–2972.
[38] V.N. Vapnik, An overview of statistical learning theory, IEEE Trans. Neural Network. 10 (1999) 988–999.
[39] C.C. Chang, C.J. Lin, LIBSVM: a library for support vector machines, ACM Transactions on Intelligent Systems and Technology 2 (2011) 27:1–27:27.
[40] G.-B. Huang, Q.-Y. Zhu, C.-K. Siew, Extreme learning machine: theory and applications, Neurocomputing 70 (2006) 489–501.
[41] J.G. Proakis, D.K. Manolakis, Digital Signal Processing, fourth ed., Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2006.
[42] N. Sengupta, M. Sahidullah, G. Saha, Lung sound classification using cepstral-based statistical features, Comput. Biol. Med. 75 (2016) 118–129.
[43] R. Schmidt, Multiple emitter location and signal parameter estimation, IEEE Trans. Antenn. Propag. 34 (1986) 276–280.