**Far East Journal of Electronics and Communications**

http://www.pphmj.com
http://dx.doi.org/10.17654/EC017051065
Volume 17, Number 5, 2017, Pages 1065-1078 **ISSN: 0973-7006**

# COMPARISON OF DISCRETE WAVELET TRANSFORM AND WAVELET PACKET DECOMPOSITION FOR THE LUNG SOUND CLASSIFICATION

**Achmad Rizal**[1,2], **Risanuri Hidayat**[1] **and Hanung Adi Nugroho**[1]

[1]Department of Electrical Engineering and Information Technology
Universitas Gadjah Mada
Yogyakarta, Indonesia

[2]School of Electrical Engineering
Telkom University
Bandung, Indonesia

## Abstract

Auscultation is still the main procedure by the physician in determining the health condition of a person's lungs. Auscultation heavily depends on the physician's skill and experience. Electronics auscultation using computer assistance is used to identify abnormalities in lung sounds for reducing the subjectivity. One of the signal processing methods that is often used to determine the lung sounds is wavelet decomposition method. This study aimed to compare several methods of lung sound classification using wavelet analysis. Some methods combined wavelet decomposition techniques and features extraction to obtain a method that produces the highest accuracy with the fewest number of features. The results showed that the DWT order 7 with DB2 mother wavelet and 46 features produce the highest accuracy of 97.98%. This method was tested on five classes of lung sound data.

Received: February 17, 2017; Accepted: May 15, 2017

Keywords and phrases: wavelet transform, lung sound, feature extraction, multilayer perceptron.

## 1. Introduction

Lung sound is one of the information used by physicians to detect abnormalities in the heart. Lung sound is heard using a stethoscope based on the expertise of the doctor [1]. Although the currently available diagnostic techniques of various lung diseases are using X-ray, CT-scan, and MRI, lung auscultation remains the first choice because of the simplicity of the device and the procedure [2]. Various techniques are used by people to analyze lung sound automatically to reduce the subjectivity [1].

Various digital signal processing techniques have been developed to recognize the lung sound using a computer. Time domain signal processing techniques often used by people such as autoregressive-modeling [3], empirical mode decomposition (EMD) [4], and fractals [5]. Some researchers used frequency analysis to identify lung sounds such like quantile vector frequency [6] or power spectral density (PSD) [7, 8]. Lung sound analysis in the time-frequency domain was also the focus of many researchers due to the non-stationary nature of lung sound. The method used such as short-time Fourier transform (STFT) [9], Wigner-Ville distribution [10], and the Hilbert-Huang transform [11].

Kandaswamy et al. [12] used wavelet decomposition order of 7 and back propagation neural network (BP-NN) to classify lung sounds into six classes. Meanwhile, the same method with some additional features and multilayer perceptron (MLP) as a classifier was proposed by Hashemi et al. [13]. This method was used to distinguish between monophonic and polyphonic wheezing sound. The method in [12] was also used in [14] with support vector machine (SVM) as the classifier. A slightly different method was used by Rizal et al. [15] to recognize the lung sounds using adaptive resonance theory 2 (ART-2 NN) as a classifier. The study in [12, 13, 15] used almost the same wavelet decomposition techniques but using different data and classification techniques, so its performance could not be compared.

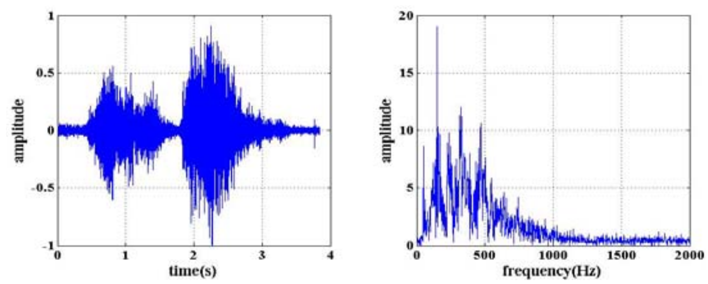
In this study, a comparison of feature extraction technique in the three studies was presented using the same lung sound data and classifier. We used the best parameter for wavelet decomposition method in each method. Due to the different sampling frequency, the feature extraction methods would also be performed at each sampling frequency. The results will be used as a recommendation for selecting the best wavelet decomposition method for lung sound feature extraction.

## 2. Materials and Methods

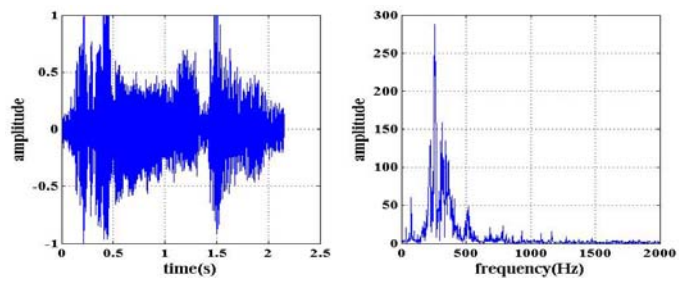
### 2.1. Lung sound data

Lung sound is generated from the air flow into and out of the respiratory tract. Changes in lung sound indicate a change in the respiratory tract which is sometimes an indicator of lung disease. Lung sound was mostly laid in the frequency of 200-250 Hz because usually above 250 Hz its energy reduces significantly [16]. Usually, the frequency range used for analyzing lung sounds is 75-1200 Hz.

In this study, we used 99 lung sound data recording consist of 22 normal data, 18 wheezing data, 21 crackles, 18 friction rub, and 20 stridor. Most of the lung sound data were used in a previous study [17, 18]. Wheezing is continuous adventitious lung sound with a duration > 250 ms. The dominant frequency of wheezing lies in the frequency of 400 Hz [16]. Crackle is an explosive and discontinuous adventitious lung sound. Crackle had very short duration and divided into fine and coarse crackle [19]. Stridor is a high-pitched wheezing heard in the upper respiratory tract [20]. Meanwhile, friction rub is adventitious lung sound that is nonmusical, explosive, and often heard in the basal region. Friction rub is often associated with inflammation of the pleura [20]. The length of each data was one respiratory cycle with a sampling frequency of 8000 Hz. For methods that require data sampling frequency of 11025 Hz, we did a resample process.




(a)




(b)

**Figure 1.** (a) Normal bronchial sound, (b) wheeze sound.

### 2.2. Preprocessing

Preprocessing on lung sounds intended to eliminate amplitude variations due to different recording environments, removing the DC component due to DC-offset and eliminate the noise. The first preprocessing in this study was mean removal to remove the DC component. At the signal $s(i)$ along $N$, the process is given by:

$$s(i) = s(i) - \frac{1}{N}\sum_{i=1}^{N} s(i), \quad i = 1, 2, \ldots N. \tag{1}$$

The next process was the amplitude normalization:

$$s(i) = \frac{s(i)}{\max|\, s(i)\,|}. \tag{2}$$

By normalization process, the difference in the signal due to the recording process and zero shifting due to noise could be minimized. The subsequent preprocessing was noise removal. One of the main noises that often occurs was the heart sound. The frequency range of heart sound was 20-150 Hz occupied a significant overlap with the low-frequency component of the lung sound [21]. The simplest technique to eliminate heart sounds was using band pass filter (BPF) with a passband of 100-2000 Hz [22, 23]. In this study, we only used normalization in preprocessing stage because the data was clean of noise.

### 2.3. Discrete wavelet transform

In general, discrete wavelet transform (DWT) can be expressed as:

$$X(a, b) = \frac{1}{\sqrt{b}} \int_{-\infty}^{\infty} x(t) \Psi\left(\frac{t-a}{b}\right) dt, \tag{3}$$

where $a$ is the time shift while $b$ is the scale (often referred to as width modulation), while $\Psi(t)$ is called the *mother wavelet* [24]. Practical implementation of DWT can be seen in Figure 2. The signal is filtered through the LPF and HPF which splits the signal into half of the original frequency range. The output filter is downsampled so that the number of samples of signals become half. In DWT, only the output of LPF called the *approximation* to be decomposed further. Selection order of DWT is chosen by the needs.

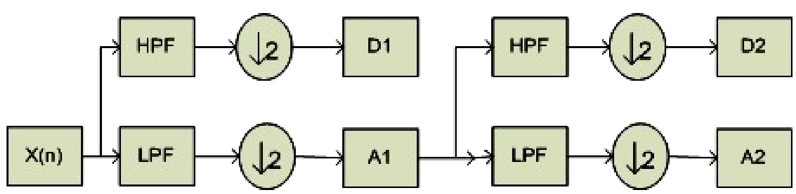



**Figure 2.** DWT order of 2.

DWT order of 7 was proposed by Kandaswamy et al. [12] for lung sound decomposition. The same scheme was proposed by Hashemi et al. [13]. Because DWT was highly dependent on the sampling frequency, the frequency range of subband in both studies would be different as shown in Table 1.

**Table 1.** Frequency range of each subband for DWT order 7 using different sampling frequency

| Subband | Kandaswamy $Fs = 11025$ Hz | Hashemi $Fs = 8000$ Hz |
|---|---|---|
| A7 | 0-43.07 | 0-31.25 |
| D7 | 43.07-86.13 | 31.25-62.5 |
| D6 | 86.13-172.26 | 62.5-125 |
| D5 | 172.26-344.53 | 125-250 |
| D4 | 344.53-689.06 | 250-500 |
| D3 | 689.06-1378.13 | 500-1000 |
| D2 | 1378.13-2756.25 | 1000-2000 |
| D1 | 2756.25-5512.50 | 2000-4000 |

Some parameters of lung sound were calculated in each subband. In Kandaswamy's method [12], the parameters were:

(1) the average of the absolute value of each subband coefficient ($\mu i$),

(2) the average power of each subband ($Pi$),

(3) standard deviation of each subband coefficient ($\sigma i$),

(4) the ratio of the average absolute value of the adjacent subband ($\mu i / \mu i + 1$),

while the Hashemi's method used the features (1)-(4) coupled with [13]:

(5) skewness of subband coefficient ($ski$),

(6) kurtosis of subband coefficient ($kui$).

Because at $Fs = 11025$ Hz, the coefficient values in D1-D2 and A7 were close to zero, the subband was not used by Kandaswamy. Total features employed by Kandaswamy method were 19 [12]. Meanwhile, Hashemi method used the whole subband so that the total of features was 46 [13].

### 2.4. Wavelet packet decomposition

Wavelet packet decomposition (WPD) is an extension of the DWT. At

WPD, the detail component was decomposed further as to the approximation component. WPD order *n* will generate 2*n* subband. WPD order of 5 was used for lung sound classification by Rizal et al. [15]. Subbands for lung sound feature extraction were selected according to the need. The energy of each subband was used as the lung sound's features. The subband used in feature calculation is shown in Figure 2. The frequency range of each subband for $Fs = 8000\,\text{Hz}$ and $Fs = 11025\,\text{Hz}$ is shown in Table 2. The results obtained 15 features.

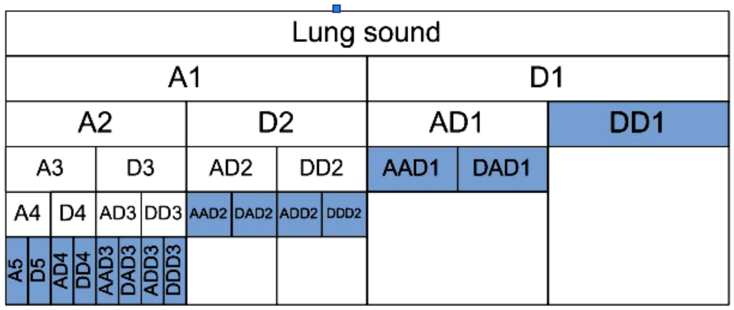



**Figure 3.** Subband for feature extraction using WPD.

### 2.5. Classifier

Many researchers chose artificial neural network (ANN) to be a classifier. Kandaswamy et al. used BP-ANN, Hashemi et al. used MLP, while Rizal et al. used adaptive resonance theory-2 (ART-2NN) as a classifier [12, 13, 15]. BP-ANN, MLP and ART-2 NN differed regarding their training process. BP-ANN and MLP included in supervised learning while ART-2 NN included in unsupervised learning. In this study, we used MLP as a classifier because in previous work very high lung sound classification was obtained [13]. In this experiment, we used MLP with *N*-40-5 configuration, where *N* was the number of input feature as used in [13].

**Table 2.** Frequency range of selected subband for WPD using different sampling frequency

| Subband | ***Fs*** **= 11025 Hz** | ***Fs*** **= 8000 Hz** |
|---|---|---|
| A5 | 0-172.27 Hz | 0-125 Hz |
| D5 | 172.27-344.53 Hz | 125-250 Hz |
| AD4 | 344.53-516.8 Hz | 250-375 Hz |
| DD4 | 516.8-689.96 Hz | 375-500 Hz |
| AAD3 | 689.96-861.33 Hz | 500-625 Hz |
| DAD3 | 861.33-1033.6 Hz | 626-750 Hz |
| ADD3 | 1033.6-1205.86 Hz | 750-875 Hz |
| DDD3 | 1205.86-1378.125 Hz | 875-1000 Hz |
| AAD2 | 1378.125-1607.81 Hz | 1000-1250 Hz |
| DAD2 | 1607.81-1837.5 Hz | 1250-1500 Hz |
| ADD2 | 1837.5-2067.19 Hz | 1500-1750 Hz |
| DDD2 | 2067.19-2756.25 Hz | 1750-2000 Hz |
| AAD1 | 2756.25-3445.31 Hz | 2000-2500 Hz |
| DAD1 | 3445.31-4134.38 Hz | 2500-3000 Hz |
| DD1 | 4134.38-5512.5 Hz | 3000-4000 Hz |

### 2.6. Cross validation

Cross validation (CV) is used to see if the model of ANN is working properly. A standard technique is used to divide the data into the training and testing data. This method produces a very high variation accuracy value depending on the distribution of the data [26]. $N$-fold CV is used to solve the problem. In the $N$-fold CV, the data is divided into $N$ data sets, with one data set becomes testing data while $N$-1 is used as training data. This process is repeated $N$ times with the final accuracy is the average of $N$ measurement accuracy. We used 3-fold CV as in [12]. The parameters used for the performance assessment is the accuracy as described below:

$$Accuracy(\%) = \frac{Number\ of\ correcctly\ classified\ data}{Total\ data}. \tag{4}$$

## 3. Result and Discussion

### 3.1. Combination of wavelet decomposition and feature extraction method

In this research, we performed a combination of wavelet decomposition method and feature extraction techniques used in studies mentioned previously [12, 13, 15]. The important parameters used in three studies are shown in Table 3. The parameters were included wavelet decomposition method, the feature parameter, *Fs*, and mother wavelet.

**Table 3.** Parameter comparison of each method

| Feature extraction method | Data | *Fs* (Hz) | Decomposition | Mother wavelet | Features |
|---|---|---|---|---|---|
| Rizal et al. [15] | 324, 20 classes | 8000 | WPD | Db2 | 15 on WPD |
| Kandaswamy et al. [12] | 120, 6 classes | 11025 | DWT | Db8 | 19 on DWT |
| Hashemi et al. [13] | 140, 2 classes | 8000 | DWT | Bior1.5 | 46 on DWT |

### 3.2. Lung sound feature extraction using DWT

In this section, DWT order of 7 was performed on lung sounds to produce eight subbands. Thus Rizal's method generated eight features, fewer than when used WPD. The accuracy of each feature extraction method is shown in Table 4.

Table 4 shows that Hashemi's method produced the highest accuracy of 97.98% in *Fs* = 8000 Hz and Db2 or Bior1.5. Meanwhile, Kandaswamy's method produced the highest accuracy of 94.95% at Db8 and Rizal's method yielded 95.96% accuracy in Db2. All the highest accuracy of each method were obtained at *Fs* = 8000 Hz. This result showed that *Fs* = 8000 Hz generated appropriated subband for DWT so that it generated higher accuracy compared with *Fs* = 11025 Hz.

**Table 4.** Accuracy of each feature extraction method using DWT order of 7, three-fold validation, 40 nodes hidden layer of MLP

| Feature extraction method | Number of features | *Fs* = 8000 Hz | | | *Fs* = 11025 Hz | | |
|---|---|---|---|---|---|---|---|
| | | Db2 | Db28 | Bior1.5 | Db2 | Db8 | Bior1.5 |
| Rizal et al. [15] | 8 | 95.96% | 91.92% | 89.90% | 91.92% | 94.95% | 89.90% |
| Kandaswamy et al. [12] | 19 | 90.91% | 94.95% | 90.91% | 90.91% | 90.91% | 86.87% |
| Hashemi et al. [13] | 46 | 97.98% | 96.97% | 97.98% | 96.97% | 96.97% | 96.97% |

**Table 5.** Accuracy of each feature extraction method using WPD order of 5, three-fold validation, 40 node hidden Layer of MLP

| Feature extraction method | Number of features | *Fs* = 8000 Hz | | | *Fs* = 11025 Hz | | |
|---|---|---|---|---|---|---|---|
| | | Db2 | Db28 | Bior1.5 | Db2 | Db8 | Bior1.5 |
| Rizal et al. [15] | 15 | 91.92% | 96.97% | 92.93% | 96.97% | 93.94% | 95.96% |
| Kandaswamy et al. [12] | 59 | 97.98% | 97.98% | 95.96% | 97.98% | 97.98% | 97.98% |
| Hashemi et al. [13] | 89 | 95.96% | 96.97% | 95.96% | 94.95% | 94.95% | 95.96% |

### 3.3. Lung sound feature extraction using WPD

The WPD scenario was designed to produce 15 subbands. This method would change the number of features produced by Kandaswamy's method and Hashemi's method. The number of subbands was more than reported in the preliminary study. The accuracy of each method of feature extraction on the WPD is shown in Table 5.

Table 5 demonstrates that the Kandaswamy's method produced the highest accuracy of 97.98%. The Kandaswamy's method was superior to the Hashemi's method due to fewer features. Meanwhile, Rizal's method produced the highest accuracy of 96.97% for the mother wavelet DB2 and *Fs* = 11025 Hz using 15 features.

### 3.4. Discussion

The selection criteria for the better methods were the accuracy and the number of features. If the accuracy had the same value, the method that had the least amount of features was chosen. In Table 4, even though the Hashemi's method was originally only intended to distinguish the monophonic and polyphonic wheeze sound but was able to distinguish five classes of data. In this case, although the Kandaswamy's method had fewer features but generated lower accuracy. This result showed that the skewness and kurtosis in each subband, A7, D2, and D1 subbands had a contribution to the classification accuracy.

In Table 4, with a sampling frequency of 11025 Hz, Hashemi's method produces the highest accuracy of 96.97%. Eliminate skewness and kurtosis in each subband were shown to reduce the accuracy significantly. The energy of subband as features was only capable of producing up to 94.95% accuracy.

Table 5 showed that Kandaswamy's method was better than the Hashemi's method. Kandaswamy's method produced a classification accuracy of 97.98% with fewer features than the Hashemi's method. These results were not influenced by the mother wavelet and the sampling frequency selection.

Based on the criteria mentioned earlier, we recommended that the best method was DWT order of 7 with the mother wavelet Db2 and $Fs = 8000$ Hz as in Table 4. The parameters used in the method proposed by Hashemi with a different mother wavelet. Although the accuracy was 97.98%, similar to that produced by the Kandaswamy's method in Table 5, the Hashemi's method used 46 features, less than Kandaswamy's method on WPD.

The same signal decomposition using DWT and WPD method were used in [27]. Modified gray level difference matrix (GLDM) was used as features [28]. In the study reported up to 100% accuracy for 81 lung sound data in five classes. Testing with the same lung sound data with this research resulted in 98.99% accuracy using 40 features DWT while the accuracy of 98.99% for WPD using the 15 features. These results indicate that using appropriate feature extraction method, DWT and WPD methods proposed in [12, 13, 15] could be improved in their performance for classification of lung sounds. Research using another classifier and bigger lung sound database can be done in future studies.

## 4. Conclusion

In this study, we presented a comparison of the three wavelet decomposition and feature extraction methods for lung sound classification. The results showed that combination of DWT order of 7 and feature extraction method proposed by Hashemi et al. provided the highest accuracy and fewest features. Testing using larger lung sound database is planned to be done in the next study. Various studies on lung sound analysis are still in progress. Researchers are trying for better methods to increase the accuracy and reduce computation time for lung sound analysis for automatic pulmonary lung disease diagnosis.

## References

[1] A. Gurung, C. G. Scrafford, J. M. Tielsch, O. S. Levine and W. Checkley, Computerized lung sound analysis as diagnostic aid for the detection of abnormal lung sounds: a systematic review and meta-analysis, Respir. Med. 105(9) (2011), 1396-1403.

[2] H. Melbye, Auscultation of the lungs: still a useful examination, Tidsskr. Nor. Laegeforen. 121(4) (2001), 451-454.

[3] Y. P. Kahya, E. C. Guler and B. Sankur, Statistical analysis of lung sound data, Proceedings of the First Joint BME/EMBS Conference, Atlanta, 1999, p. 1015.

[4] S. Charleston-Villalobos, R. González-Camarena, G. Chi-Lem and T. Aljama-Corrales, Crackle sounds analysis by empirical mode decomposition, IEEE Eng. Med. Biol. Mag. 26 (2007), 40-47.

[5] J. Gnitecki, Z. Moussavi and H. Pasterkamp, Classification of lung sounds during bronchial provocation using waveform fractal dimensions, Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society EMBS 2004, San Francisco, 2004, pp. 3844-3847.

[6] P. Mayorga, C. Druzgalski, O. H. Gonzalez and H. S. Lopez, Modified

classification of normal lung sounds applying quantile vectors, Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society EMBS, San Diego, 2012, pp. 4262-4265.

[7] R. Jané, S. Cortés, J. A. Fiz and J. Morera, Analysis of wheezes in asthmatic patients during spontaneous respiration, Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Francisco, 2004, pp. 3836-3839.

[8] M. Oud, E. H. Dooijes and J. S. van Der Zee, Airways obstruction assessment based on detailed analysis of respiratory sound spectra, IEEE Trans. Biomed. Eng. 47(11) (2000), 1450-1455.

[9] A. Rizal and V. Suryani, Lung sound recognition using spectrogram and adaptive resonance theory 2 neural network (ART2), Proceedings Biomedical Engineering Day 2008, Surabaya, 2008.

[10] M. Maciuk, W. Kuniszyk-Jóźkowiak, A. Doboszyńska and M. Maciuk, Analysis of lung auscultatory phenomena using the Wigner-Ville distribution, Ann. UMCS, Inform. 12(1) (2012), 7-16.

[11] X. Chen, J. Shao, Y. Long, C. Que, J. Zhang and J. Fang, Identification of Velcro rales based on Hilbert-Huang transform, Phys. A Stat. Mech. Appl. 401(8) (2014), 34-44.

[12] A. Kandaswamy, C. S. Kumar, R. P. Ramanathan, S. Jayaraman and N. Malmurugan, Neural classification of lung sounds using wavelet coefficients, Comput. Biol. Med. 34(6) (2004), 523-537.

[13] A. Hashemi, H. Arabalibiek and K. Agin, Classification of wheeze sounds using wavelets and neural networks, International Conference on Biomedical Engineering and Technology 2011, Kuala Lumpur, 2011, pp. 127-131.

[14] S. Abbasi, R. Derakhshanfar, A. Abbasi and Y. Sarbaz, Classification of normal and abnormal lung sounds using neural network and support vector machines, 2013 21st Iranian Conference on Electrical Engineering (ICEE), Mashhad, 2013, pp. 1-4.

[15] A. Rizal, T. L. R. Mengko and A. B. Suksmono, Lung sound recognition using wavelet packet decomposition and ART2 (adaptive resonance theory 2) neural network, Proceedings Biomedical Engineering Day 2006, Bandung, 2006, pp. 2-6.

[16] A. R. A. Sovijärvi, L. P. Malmberg, G. Charbonneau, J. Vanmderschoot, F. Dalmasso, C. Sacco, M. Rossi and J. E. Earis, Characteristics of breath sounds and adventitious respiratory sounds, Eur. Respir Rev. 10(77) (2000), 591-596.

[17] A. Rizal, R. Hidayat and H. A. Nugroho, Multiscale Hjorth descriptor for lung sound classification, International Conference on Science and Technology 2015, Yogyakarta, 2015, pp. 1-7.

[18] A. Rizal, R. Hidayat and H. A. Nugroho, Multi-scale grey-level difference for lung sound classification, J. Electr. Syst. 13(3) (2016), 556-564.

[19] S. Reichert, R. Gass, C. Brandt and E. Andrès, Analysis of respiratory sounds: state of the art, Clin. Med. Circ. Respir. Pulm. Med. 2 (2008), 45-58.

[20] A. Bohadana, G. Izbicki and S. S. Kraman, Fundamentals of lung auscultation, N. Engl. J. Med. 370(8) (2014), 744-751.

[21] L. J. Hadjileontiadis and S. M. Panas, Adaptive reduction of heart sounds from lung sounds using fourth-order statistics, IEEE Trans. Biomed. Eng. 44(7) (1997), 642-648.

[22] B. S. Lin, B. S. Lin, H. D. Wu, F. C. Chong and S. J. Chen, Wheeze recognition based on 2D bilateral filtering of spectrogram, Biomed. Eng. Basis Commun. 18(3) (2006), 29-38.

[23] A. Homs-Corbera, R. Jane, J. A. Fiz and J. Morera, Algorithm for time-frequency detection and analysis of wheezes, Proceedings of the 22nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Chicago, 2000, pp. 2977-2980.

[24] S. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, IEEE Trans. Pattern Anal. Mach. Intell. 11(7) (1989), 674-693.

[25] G. A. Carpenter, S. Grossberg and D. Rosen, ART 2-A: an adaptive resonance algorithm for rapid category learning and recognition, IJCNN-91-Seattle Int. Jt. Conf. Neural Networks, Vol. II, Seattle, 1991.

[26] R. Palaniappan, Biological Signal Analysis, Ventus Publishing ApS, 2010.

[27] A. Rizal, R. Hidayat and H. A. Nugroho, Multiresolution modified gray level difference for respiratory sound classification, The 2016 Advance Research in Engineering and Information Technology International Conference (AREITIC), Bandung, 2016, pp. 1-5.

[28] J. S. Weszka, C. R. Dyer and A. Rosenfeld, A comparative study of texture measures for terrain classification, Syst. Man Cybern. IEEE Trans. SMC-6(2) (1976), 269-285.