33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Feature Extraction using Time-Frequency/Scale Analysis and Ensemble of Feature Sets for Crackle Detection

Gorkem Serbes, C. Okan Sakar, Yasemin P. Kahya, *Member, IEEE,* and Nizamettin Aydin, *Member, IEEE*

***Abstract*—Pulmonary crackles are used as indicators for the diagnosis of different pulmonary disorders. Crackles are very common adventitious sounds which have transient characteristic. From the characteristics of crackles such as timing and number of occurrences, the type and the severity of the pulmonary diseases can be obtained. In this study, a novel method is proposed for crackle detection. In this method, various feature sets are extracted using time-frequency and time-scale analysis. The extracted feature sets are fed into support vector machines both individually and as an ensemble of networks. Besides, as a preprocessing stage in order to improve the success of the model, frequency bands containing no-information are removed using dual tree complex wavelet transform, which is a shift invariant transform with limited redundancy and an improved version of discrete wavelet transform. The comparative results of individual feature sets and ensemble of sets with pre-processed and non pre-processed data are proposed.**

I. Introduction

CHEST auscultation of pulmonary sounds via a stethoscope is a widely used, inexpensive and noninvasive method for the evaluation of the respiratory disorders. However it is considered to be an inadequate diagnostic method due to its inherent subjectivity and limited frequency response of the stethoscope (the stethoscope attenuates frequencies above 120 Hz). In recent years, the analysis of pulmonary sound signals with computers has become an established research area with the improvements in digital recording systems and advanced digital signal processing techniques [1-3].

Although the exact mechanism is still unknown, the pulmonary sounds are assumed to be produced due to air turbulence in the airways of the lungs. Pulmonary sounds can be studied in two classes, vesicular sounds and adventitious sounds. Vesicular sounds are the respiratory sounds heard over the chest wall and are synchronous with air flow in the airways. Adventitious sounds, on the other hand, are additional sounds which usually occur with respiratory disorders [4].

Crackles are discontinuous, adventitious non-musical respiratory sounds which are attributed to sudden bursts of air within bronchioles. Their duration is less than 20 ms and their frequency range is between 100 to 2000 Hz. Crackles occur in pathological conditions and are superimposed on vesicular sounds. Crackles are explosive and transient in character, and occur frequently in respiratory diseases. The characteristics of pulmonary crackles such as timing, epochs of occurrence, and pitch can be used in the diagnosis for various types of pulmonary diseases such as pneumonia, bronchiectasis, fibrosing alveolitis and asbestosis [5-8].

For a computerized analysis of pulmonary diseases, proper detection of crackles is very important. In this study, a novel method is proposed for pulmonary crackle detection. For the analysis, a pulmonary dataset consisting of 3000 256-point crackle signals and 3000 256-point non-crackle (healthy) signals, are used. By applying time-frequency (TF) and time-scale (TS) analysis to these 6000 signals, four different feature subsets are obtained and the original signals are used as the fifth feature subset. The extracted feature subsets are fed into support vector machines (SVM) classifier as inputs both individually and as an ensemble of networks. Besides, in order to improve the generalization and crackle detection capability of the model, frequency components of processed signals containing no-information (below 100 Hz and above 2000Hz) are removed using dual tree complex wavelet transform (DTCWT), which is an improved version of discrete wavelet transform (DWT) with better shift invariance property, as a pre-processing stage. The comparative results of individual and ensemble feature sets with pre-processed and non pre-processed data using SVMs are presented.

II. Materials and Methods

A. Data Acquisition System

In the data acquisition system fourteen air-coupled electret microphones (Sony-ECM 44) are placed on the posterior chest, and airflow is recorded using Fleisch-type flowmeter (Validyne CD379) to synchronize on the inspiration-expiration

Manuscript received June 20, 2011. This work was supported by Bogazici University Research Fund under Project No. 09A203D.

G. Serbes is with the Mechatronics Engineering Department, Bahcesehir Univesity, Istanbul, Turkey (e-mail: gorkem.serbes@bahcesehir.edu.tr). His work is supported by the Ph.D. scholarship (2211) from Turkish Scientific Technical Research Council (TÜBİTAK).

C. O. Sakar is with the Department of Computer Engineering, Bahcesehir Univesity, Istanbul, Turkey (e-mail: okan.sakar@bahcesehir.edu.tr). His work is supported by the Ph.D. scholarship (2211) from TÜBİTAK.

Y. Kahya is with the Department of Electrical Engineering, Bogazici University, Istanbul, Turkey (e-mail: kahya@boun.edu.tr).

N. Aydin is with the Department of Computer Engineering, Yildiz Technical Univesity, Istanbul, Turkey (e-mail: naydin@yildiz.edu.tr).

phases. A low-noise preamplifier, 8^{th} order Butterworth low-pass filters with 4 kHz cut-off frequency and 6^{th} order Bessel high-pass filters with 80 Hz cut-off frequency are used in order to minimize frictional noise and heart sound interference and for an anti-aliasing filter. The amplified signals are digitized by a 12-bit ADC Card (NIDAQ500) at a 9.15 kHz sampling rate and stored [9]. The details of the system are described in [10]. A sample of a lung signal containing crackles and a time expanded part of that signal containing two crackles are shown in Figure 1.

B. Feature Extraction

The spectral characteristics of lung sounds show different behaviors according to the state and pathology of the lung. The pathological sounds appear in higher frequency bands, i.e. as crackles which are explosive and transient in time. The frequency characteristic of crackles is used for the feature extraction part of the proposed method using TF and TS analysis.

The signals collected from 26 subjects (13 healthy, 13 pathological) are divided into 6000 samples, 3000 non-crackle, i.e. healthy, and 3000 crackle, each consisting of 256-point. In the preparation of crackle samples, 3000 crackles which were identified previously by physicians are randomly placed into 256 point windows. The window size is chosen as 256 points because the duration of crackles is less than 20 ms and due to sampling frequency of the data acquisition system, which is 9150 Hz, 256 points sized window is equal to 26.9 ms. In the preparation of healthy signals, 3000 256-point windows were randomly created from healthy subjects. In Figure 2, an example consisting of two 256-point crackle windows and two healthy windows can be seen.

In order to use frequency characteristics of crackles, TF and TS analysis are applied to both crackle windows and healthy windows. For the TF and TS analyses, 64 points complex Fourier transform with Gaussian window and 64 scales complex wavelet transform with Morlet wavelet are used, respectively.

The output of TF analysis gives the information about behavior of analyzed signals depending on both time and frequency. Then as a first feature set, the outputs of TF analysis are integrated over frequency, and so the behavior of signal upon time is obtained. As a second feature set, the outputs of TF analysis are integrated over time, and the behavior of signal upon frequency is obtained. Same procedure was also carried out for TS analysis. As the third and fourth feature sets, the outputs of TS analysis are integrated over scale and time, and the behaviors of signal upon time and scale, respectively, are obtained. The original signals are used as the fifth feature set.

Vesicular sounds have the frequency components between 0-200 Hz and crackles have the frequency components between 150 - 2000 Hz. In order to improve the performance of the proposed method, a preprocessing step which removes the frequency components having no crackle information, is applied to the dataset before extracting the feature sets. For this task, a five level DTCWT is applied on both crackle and healthy signals. The DTCWT is developed to overcome the lack of shift invariance property of ordinary discrete wavelet transform (DWT). Moreover it has limited redundancy (2^m:1 for m dimensional signals, which is a very good ratio as compared with undecimated DWT). In the analysis of non-stationary crackles, which are transient signals, DTCWT removes undesirable signal components more successfully than DWT because of its shift invariance property. With DTCWT the frequency bands below 150 Hz and above 2400 Hz are replaced with null vectors and then the processed signal is reconstructed. The details of DTCWT can be seen in [11-13].

C. Individual Learning with Feature Sets

We used LIBSVM [14] implementation of Support Vector Machines (SVMs) [15] as the classifier. The aim of the classifier is to build a predictive model capable of distinguishing between the crackle and non-crackle, i.e. 'healthy', signals. For this purpose, we divided the dataset into three groups with equal number of samples: 2000 samples for train, 2000 for validation, and 2000 for test. The distribution of the samples to the datasets has been done such that each set contains 1000 samples from each class type.

We trained each feature subset using training set and tested on validation set in order to find the most suitable kernel type among linear, polynomial and Gaussian. The SVMs parameter values, C (cost) and g (the spread parameter), are also optimized for each of the feature set. The optimized models are finally tested on the yet unseen test sets, and the unbiased success of each feature set is proposed.

D. Ensemble of the Feature Sets

We used the ensemble of the feature sets in order to improve the overall accuracy and generalization capability of the constructed model based on the proof of Hansen and Salamon [16]: if each member of the ensemble, i.e. feature subset, can get the right answer more than half the time, and if the responses of members are independent, the likelihood of an error by a majority voting strategy will monotonically decrease with the increasing number of members. Learning from multiple sets of features, called ensemble learning, is based on employing separate classifiers on each feature subset and combining the predictions of the views using techniques such as voting and stacking [17]. We trained each of the four extracted feature sets and also the original signals on the training set, then applied the optimized model of each feature set on the test set, and combined the class posterior probability estimates using simple voting (*Algorithm 1*).

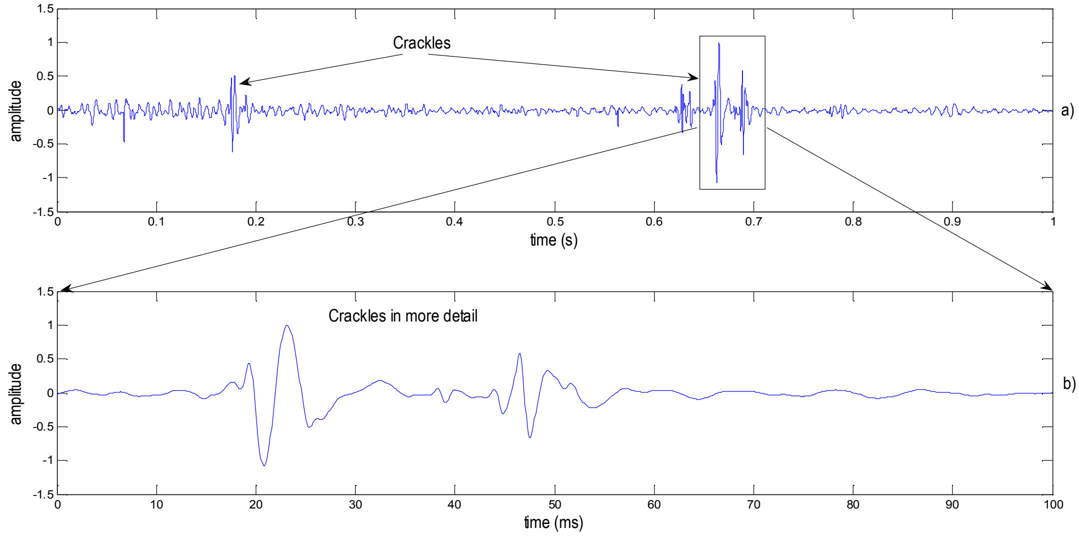

Fig. 1. A sample of a lung signal, a) containing crackles, and b) a time expanded part of it containing two crackles

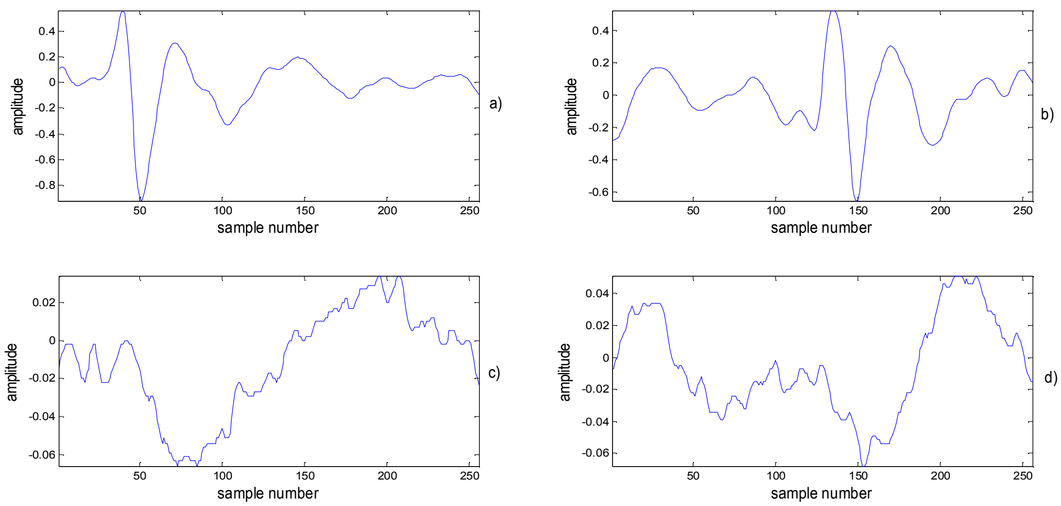

Fig. 2. Examples of 256-point windows containing crackles (a, b) and healthy signals (c, d).

III. Results

The results of optimized models on test set of each individual feature set and ensemble of sets are shown in Table 1. It is seen that the highest overall accuracy with 97.20% is obtained with our proposed method where a DTCWT is applied as a preprocessing step on each feature set, and the features sets are used as an ensemble of networks. True positive (TP) rate of the ensemble is also the second highest with 96.80 which shows that our ensemble method is successful at detecting the crackle signals. Although the highest TP rate, 97.10%, is obtained using TF analysis upon frequency, its true negative (TN) rate shows that it is not as successful as the ensemble of sets at recognizing the healthy signals. As we also see in Table 1, removing the frequency parts (i.e. noise) containing no- crackle information of processed signals using DTCWT enhances the overall accuracy of four of the five feature sets.

The results show that the highest overall accuracy with 96.35% among individual feature sets is obtained using DTCWT and TF analysis upon time feature extraction. TF analysis upon frequency without preprocessing is the next most successful feature extraction technique with 95.50% accuracy. Although the TN rate of the model that uses original signals is the highest among all networks, its TP rate and overall accuracy is the worst which shows that both TF and time TS feature extraction techniques improve the success of the model by increasing the TP rate, i.e. the crackle detection capability.

Algorithm 1 *Ensemble of feature sets*

Given F as the number of feature sets,
TRAIN and TEST as the training and test sets, respectively, consisting of all the feautres of feature sets
Individual Learning
1: **For all** $\text{train}_f \in \text{TRAIN}$, and $\text{test}_f \in \text{TEST}$ $f = 1,....,F$
2: $\text{Model}_f \leftarrow$ learn an SVM using train_f
3: $\text{Prob.Estimates}_f \leftarrow$ apply Model_f on TEST_f
4: **end for**
Combination of Predictions
5: FinalLabels $\leftarrow$ simple voting using Prob.Estimates

IV. Conclusions and Future Works

The analysis of pulmonary sound signals with computers is a recent research area due to the improvements in digital recording systems and advanced digital signal processing techniques. In this study, we propose a crackle detection method in which we use DTCWT as a pre-processing step for removing the frequency bands containing no-information, then extract various feature sets using TF and TS analysis, feed these feature sets to SVMs, and combine their class posterior probability estimates for final predictions as an ensemble of networks. We conclude that using DTCWT as a preprocessing step, extracting features instead of using the original signals, and combining the feature sets as an ensemble improve the crackle detection capability of the model.

In the future, effect of windowing on the TF analysis and effect of mother wavelet function on the TS analysis of crackles will be investigated. It is also a challenge to implement the proposed method in real time as an online crackle detection system.

Table 1. Accuracies, true positive (TP), and true negative (TN) rates of the individual feature sets and their ensembles

Preprocessing	**- (no preprocessing)**			**DTCWT**		
Feature Set	**TP (%)**	**TN (%)**	**Overall (%)**	**TP (%)**	**TN (%)**	**Overall (%)**
Time Freq. Analysis Upon Time	87.20	96.40	91.80	96.70	96.00	96.35
Time Freq. Analysis Upon Freq.	**97.10**	93.90	95.50	95.40	93.80	94.60
Time Scale Analysis Upon Time	93.20	95.70	94.45	94.50	95.60	95.05
Time Scale Analysis Upon Scale	92.00	88.60	90.30	89.30	94.90	92.10
Original Signals	65.70	96.30	81.00	69.20	**97.80**	83.50
Ensemble of Sets	94.50	**98.30**	**96.40**	**96.80**	97.60	**97.20**

References

[1] N Gavriely. *Breath Sounds Methodology*. Boca Raton, FL, CRC Press, 1995.
[2] H. Pasterkamp, S. Kraman, and G. Wodicka, "Respiratory sounds. Advances beyond the stethoscope," *Am J Respir Crit Care Med, vol. 156, Sept.* 1997, pp. 974–987.
[3] S Lehrer, Understanding Lung Sounds, 3rd ed. WB Saunders Company, 2002.
[4] M. Yeğiner, Y. P. Kahya, "Elimination of vesicular sounds from pulmonary crackle waveforms," *Computers Methods and Programs in Biomedicine*, vol.89, 2008, pp.1-13.
[5] A. Nath and L. Capel, "Inspiratory crackles: early and late," *Thorax,* vol. 29, 1974, pp. 223-227.
[6] P. Piirilä and A. R. Sovijärvi, "Crackles: recording, analysis and clinical significance," *Eur Respir J*, vol. 8(12), Dec. 1995, pp. 2139-2148.
[7] A. R. Sovijärvi, P. Piirilä, and R. Luukkonen, "Separation of pulmonary disorders with two-dimensional discriminant analysis of crackles," *Clin Physiol,* vol. 16(2), Mar. 1996, pp. 171-181.
[8] A. Vyshedskiy, F. Bezares, R. Paciej, M. Ebril, J. Shane, and R. Murphy, "Transmission of crackles in patients with interstitial pulmonary fibrosis, congestive heart failure, and pneumonia," *Chest*, vol. 128(3), Sep. 2005, pp. 1468-1474.
[9] M. Yeğiner, Y. P. Kahya, "Feature extraction for pulmonary crackle representation via wavelet networks," *Computers in Biology and Medicine,* vol.39, 2009, pp.713-721.
[10] I. Sen, Y. P. Kahya, "A Multi-Channel Device for Respiratory Sound Data Acquisition and Transient Detection", in *Proceedings of the International Conference of IEEE/EMBS*, 2005, pp. 6658-6661.
[11] N. G. Kingsbury, "The dual-tree complex wavelet transform: a new technique for shift invariance and directional filters", *IEEE Digital Signal Processing Workshop, SP 98*, Bryce Canyon, paper no 86, 1998.
[12] I. W. Selesnick, R G Baraniuk, and N. G. Kingsbury: "The dual-tree complex wavelet transform", *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 123-151, 2005.
[13] G. Serbes, N. Aydin, "Denoising Embolic Doppler Ultrasound Signals Using Dual Tree Complex Discrete Wavelet Transform," in *Proceedings of the International Conference of IEEE/EMBS, 2010, pp. 1840-1843.*
[14] C. W. Hsu, C. J. Lin, "A Comparison of Methods for Multi-Class Support Vector Machines," *IEEE Trans. Neural Networks*, vol. 13, 2002, pp. 415-425.
[15] V. N. Vapnik, *The Nature of Statistical Learning Theory*, New York: Springer-Verlag, 1995.
[16] L. Hansen, P. Salamon, "Neural network ensembles," *IEEE Trans. on Pattern Analysis and Machine Intelligence," vol. 12, 1990, pp. 993-1001.*
[17] D. Wolpert, "Stacked Generalization*," Neural Networks*, vol. 5(2), 1992, pp. 241-259.