31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Classification of Respiratory Signals by Linear Analysis

Sergul Aydore, Ipek Sen, Yasemin P. Kahya, M. Kivanc Mihcak

Abstract— **The aim of this study is the classification of wheeze and non-wheeze epochs within respiratory sound signals acquired from patients with asthma and COPD. Since a wheeze signal, having a sinusoidal waveform, has a different behavior in time and frequency domains from that of a non-wheeze signal, the features selected for classification are kurtosis, Renyi entropy, f_{50}/f_{90} ratio and mean-crossing irregularity. Upon calculation of these features for each wheeze and non-wheeze portion, the whole data scattered as two classes in four dimensional feature space is projected using Fisher Discriminant Analysis (FDA) onto the single dimensional space that separates the two classes best. Observing that the two classes are visually well separated in this new space, Neyman-Pearson hypothesis testing is applied. Finally, the correct classification rate is $\%95.1$ for the training set, and leave-one-out approach pursuing the above methodology yields a success rate of $\%93.5$ for the test set.**

I. INTRODUCTION

Two main methods used in the diagnosis of respiratory diseases are computerized imaging techniques and auscultation, the latter being cheap, simple and patient-friendly. In many respiratory disorders (e.g. pneumonia, emphysema, bronchiectasis, COPD and asthma), adventitious sounds special to the disorder are observed in the sound data, and this generally provides invaluable information that can eliminate the need for imaging techniques. However, the stethoscope does not perform well within the frequency band where the human ear is most sensitive and the sound data can not be recorded with this method. Thus, this method highly depends on the subjective evaluation of the physician and it is impossible to make an objective quantification.

In order to make auscultation a more valuable diagnostic tool, computerized methods have been applied more so in recent studies. The first studies on this area are summarized in [1] and [2]. The computerized auscultation and analysis methods of respiratory sound signals can be summarized as; capturing the respiratory sounds as analog signals via a transducer (or an array of transducers), digitization after preprocessing, and interpretation of the digitized data using various analysis techniques in the computer environment.

In this study, a method has been developed for the detection of a wheeze signal (having a continuous and sinusoidal waveform) which is a clinical indicator of obstructive respiratory diseases (e.g. COPD, asthma) and is mostly observed in acute asthma episodes. The method is intended to be a solution to the detection problem of a wheeze, i.e. whether it exists or not given a segment of recorded lung sound signal. According to the performance of classification carried out with the training data set, any segment chosen within a recorded lung sound signal can be classified as wheeze or non-wheeze within a certain (foretold) error margin. Although this study considers the classification of wheeze and non-wheeze signal portions, the main aim is the detection of a wheeze. The existence of a wheeze and its characteristics (e.g. main frequency component, or ratio of its time duration to the total breath cycle) can be accepted an indicator of the degree of the bronchial obstruction [3]. This relationship should be especially emphasized since it reveals the importance of the detection (therefore classification) problem.

In this study, the features which are expected to be distinctive have been extracted from the windows which had been labeled by an expert after a visual inspection as wheeze and non-wheeze. Then, supervised classification has been applied in this feature space, and both training and test errors have been calculated. In similar studies carried previously by other researchers, artificial neural networks were generally adopted; and as the features used, either the original signal itself and its Fourier transform coefficients [4], or power spectrum components [5-7], or, descriptive statistical features derived from wavelet transform coefficients [8] have been employed. Moreover, there are some studies which apply peak detection on time-frequency representations like spectrogram or continuous wavelet transform to identify the existence of a wheeze [9-12]. The features proposed in this study are kurtosis, Renyi entropy and mean-crossing irregularity calculated in the time domain, and f_{50}/f_{90} ratio calculated in the frequency domain. After feature extraction, the wheeze and non-wheeze portions can be considered each as an element of one of the two vector sets in the four-dimensional feature space. Fisher discriminant analysis and Neyman Pearson hypothesis testing are applied for classification and detection, and after the error rates have been calculated for the training set, the test error has been calculated for the same data set using leave-one-out method.

II. EXPERIMENTAL SETUP AND DATA

The data used in this study is selected from the data base which were recorded by the 14-channel respiratory sound data acquisition system [13] developed in the Bogazici University Lung Sounds Laboratory. The system is composed of 14 electret microphones (Sony ECM-44 BPT) attached at the posterior chest wall, an analog amplifier-filter unit (with a gain of 100 and pass band of 80-4000 Hz) which processes the signal from the microphone, a data acquisition card which digitalizes the processed signal and transmits it to the personal computer (National Instruments DAQCard - 6024E), a notebook with an interface identified in the LabView media, and a Fleisch type pneumatachograph (Validyne CD379) that synchronizes signals with the flow cycle.

978-1-4244-3296-7/09/$25.00 ©2009 IEEE

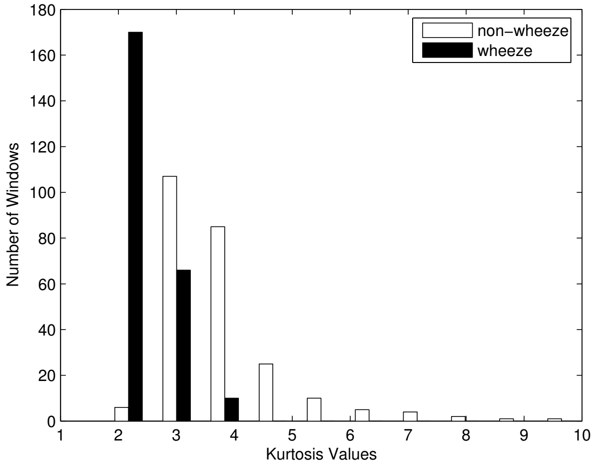

Fig. 1. Histograms of Kurtosis Values

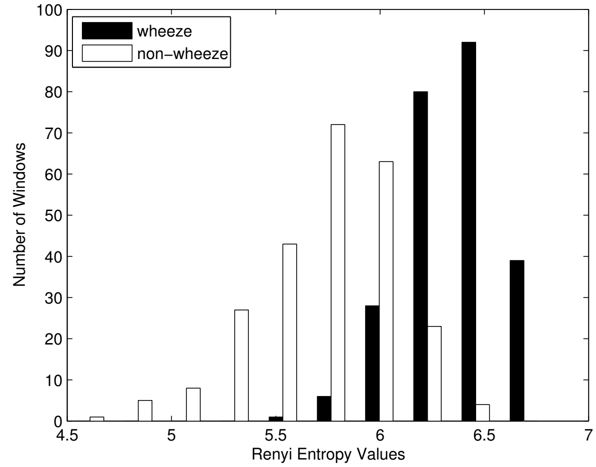

Fig. 2. Histograms of Renyi Entropy

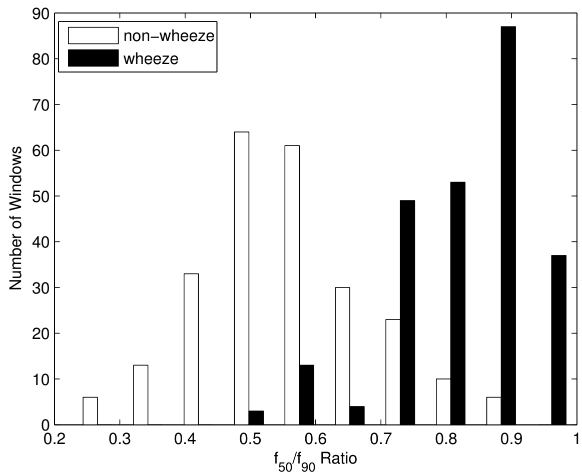

Fig. 3. Histograms of f_{50}/f_{90} ratio

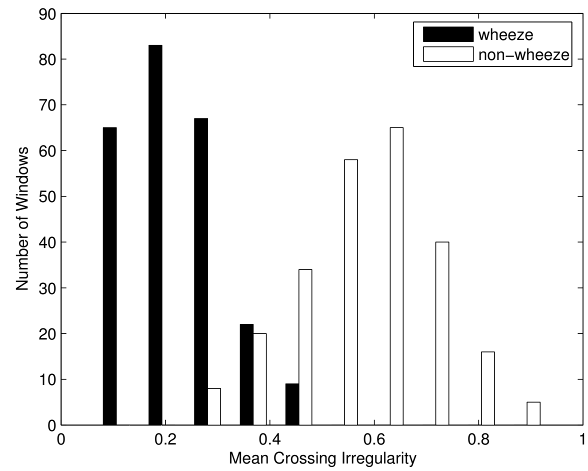

Fig. 4. Histograms of number of mean crossing

The sampling frequency of lung signals is 9600 Hz and the duration of one data acquisition session is 15 sec. More than one data acquisition sessions of 15 sec are carried out with each subject. The data in this study were taken from asthma and COPD patients who were under treatment in the Istanbul Yedikule Teaching Hospital for Chest Diseases and Thoracic Surgery. Data from four male and three female subjects in the age of 50 ± 17 were used. By visually inspecting the time-expanded sound signals together with auditory confirmation, an expert labeled the wheeze portions, and then labeled non-wheeze portions that are at comparable lengths with wheezy ones, within the same signal was made using the appropriate 15 sec sessions and appropriate channels of the 7 subjects. A total of 246 wheeze and non-wheeze portions were thereby labeled and used in this study.

III. Methodology

From distinctive properties of wheeze and non-wheeze signals in time and frequency domains, four features are defined for classification of wheeze and non-wheeze signals. These features are calculated for each window which has been labeled either as wheeze and non-wheeze. Thus, the two classes can be considered as two vector sets with 246 elements in four-dimensional space. The line which separates the two classes best when the four-dimensional vectors are projected onto it is found by Fisher linear discriminant method. Thus, the four-dimensional space is reduced to a one-dimensional space. The training and test performances are tested using the distribution of these projected values in one dimension.

The detailed explanation about features, Fisher discriminant method and Neyman-Pearson Hypothesis test are given in the following subsections.

A. Features

Kurtosis Kurtosis gives the degree of peakedness of a probability distribution and is defined as $k = \frac{\mathrm{E}(X-\mu)^4}{\sigma^4}$ for a random variable X. It is known that the kurtosis value of the normal distribution is 3, whereas it is less than 3 for sub-gaussian distributions (e.g. uniform distribution). Thus,

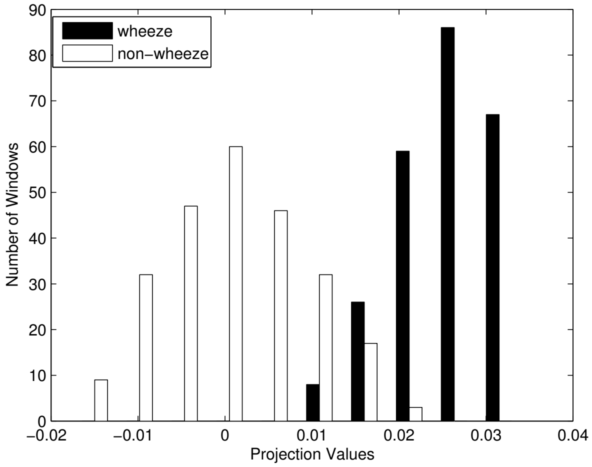

Fig. 5. Histograms of projection values

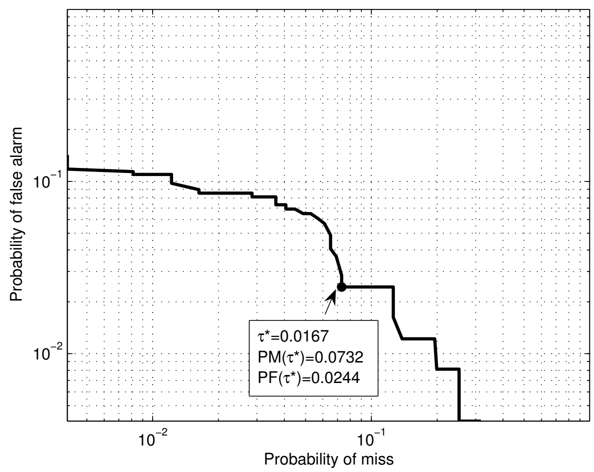

Fig. 6. ROC Curve for projection values

in practice estimated kurtosis can be used in order to measure the similarity of a given empirical distribution to the normal distribution.In this study, the kurtosis of the distributions of the values of wheeze and non-wheeze windows in time domain are calculated. The histograms of the kurtosis values for both classes are given in Fig. 1. As it can be seen from the figure, the kurtosis values are mostly distributed around 3 for non-wheeze signals and they are less than 3 for wheeze signals. This result confirms the fact that a different behavior is expected for the kurtosis of non-wheeze signals which are expected to have the normal distribution as compared to wheeze signals which are expected to have a uniform distribution.

Renyi Entropy: Renyi entropy is the generalized Shannon entropy which gives the degree of uncertainty in the system and is defined as $H_\alpha(X) = \frac{1}{1-\alpha} log(\sum_{i=1}^{n} p_i^\alpha)$ for a given random variable X. It can be considered as a measure of the hidden information or uncertainty for signals and uniformity for distributions. In this study, the distributions of the time-domain values of wheeze and non-wheeze signals are considered as probability mass functions and the Renyi entropies of these functions are calculated. The histograms of the Renyi entropy values when $\alpha = 2$ for both classes are given in Fig. 2. The figure is consistent with the expectation that the wheeze signals have a distributions closer to the uniform distributions compared to non-wheeze signals.

f_{50}/f_{90} ratio: f_{50} and f_{90} denote the frequencies where the ratio of the area under the power spectral density function to the total area is 50 % and 90 %, respectively. Ideally, i.e. for an accurate estimation of power spectral density, since the power spectral density is concentrated on a single frequency for wheeze signals, f_{50} and f_{90} values are expected to be close to each other. According to this expectation, the f_{50}/f_{90} ratio is supposed to be larger for wheeze signals than for non-wheeze signals. The histograms of the f_{50}/f_{90} ratios for both classes are given in Fig. 3. The figure confirms the expectation that the f_{50}/f_{90} ratio is larger for wheeze signals.

Here, Welch method is used in order to estimate the power spectral densities. The window length and the number of Fourier Transform points are 256, overlapping ratio is % 50 and window type is Hamming.

Mean Crossing Irregularity: It is expected that there should be a difference in the mean crossing behaviors of waveforms of the wheeze and non-wheeze signals which have regular and irregular oscillations, respectively. If the interval between successive mean crossing indexes of the signals is defined as a random variable, the deviation from the mean value is expected to be larger for non-wheeze signals than for wheeze signals. Mathematically, if X denotes the random variable for the interval between two successive mean crossing indices, the mean crossing irregularity is defined as $\frac{\sqrt{\mathrm{Var}(X)}}{\mathrm{E}(X)}$, i.e. the ratio of standard deviation of this variable to its' mean value. The histograms of the mean crossing irregularities for both classes are given in Fig. 4. As is expected, the irregularity of mean crossing values are larger for non-wheeze signals and there is a significant difference between the two classes in terms of mean crossing irregularity.

B. Fisher Discriminant Analysis and Neyman Pearson Hypothesis Testing

After the extraction of four features for each window from both classes, Fisher discriminant analysis is applied in order to separate the two classes in one-dimension.

Suppose we have n d-dimensional $\{\mathbf{x}_1, \mathbf{x}_2 ... \mathbf{x}_n\}$ vectors and n_1 (n_2) of them are in the subset $\mathscr{D}_1$ $(\mathscr{D}_2)$ and labeled as w_1 (w_2), respectively. If the dot product of these d-dimensional vectors with $\mathbf{w} \in \mathbb{R}^d$ is taken, we have $y_i = \mathbf{w}^T \mathbf{x}_i$ $(i \in 1, 2, ..., n)$ scalars contained in $\mathscr{Y}_1$ or $\mathscr{Y}_2$ subspaces. Here, $n = 492$, $n_1 = n_2 = 246$, $d = 4$ and w_1 and w_2 denote wheeze and non-wheeze classes, respectively. Geometrically, if $||\mathbf{w}|| = 1$, each y_i is the projection of the corresponding $\mathbf{x}_i$ vector to the line in the direction of the vector $\mathbf{w}$. Fisher discriminant analysis finds the best $\mathbf{w}$ which separates two classes best when they are projected onto it. One of the separation measures between the two classes of projection values is the distance between the mean values of the classes. If the class index is denoted as k, for $k=\{1,2\}$, the mean vector in d-dimensional space is defined as $\mathbf{m}_k = \frac{1}{n_k}\sum_{\mathbf{x}\in\mathscr{D}_k}\mathbf{x}$ and the mean of the projected values is defined as $\tilde{m}_k = \frac{1}{n_k}\sum_{y\in\mathscr{Y}_k} y = \mathbf{w}^T\mathbf{m}_k$. Thus, the distance which is desired to be maximized turns out to be $|\tilde{m}_1-\tilde{m}_2| = |\mathbf{w}^T(\mathbf{m}_1-\mathbf{m}_2)|$. However, in order to get a good separation between the projected classes, the distance between the mean values needs to be normalized according to the standard deviation of each class. If the scattering of the projected values for each class is defined as $s_k = \sum_{y\in\mathscr{Y}_k}(y-\tilde{m}_k)^2$, then the cost function desired to be maximized becomes $J(\mathbf{w}) = \frac{|\tilde{m}_1-\tilde{m}_2|^2}{\tilde{s}_1^2+\tilde{s}_2^2}$. In order to write $J(.)$ as an explicit function of $\mathbf{w}$, $\mathbf{S}_k = \sum_{\mathbf{x}\in\mathscr{D}_k}(\mathbf{x}-\mathbf{m}_k)(\mathbf{x}-\mathbf{m}_k)^T$, $\mathbf{S}_W = \mathbf{S}_1+\mathbf{S}_2$ and $\mathbf{S}_B = (\mathbf{m}_1-\mathbf{m}_2)(\mathbf{m}_1-\mathbf{m}_2)^T$ scattering matrices are defined. After trivial constructions, the cost function reduces to $J(\mathbf{w}) = \frac{\mathbf{w}^T\mathbf{S}_B\mathbf{w}}{\mathbf{w}^T\mathbf{S}_W\mathbf{w}}$ [14]. The histograms of the projected values of $\mathbf{x}$ vectors onto $\mathbf{w}$ vector are given in Fig. 5. For any selected value of the threshold (τ), Neyman Pearson hypothesis test can be written as; $(i\in 1,2,...,n)$

$$\begin{array}{llll} H_0: & y_i\in\mathscr{D}_2 & \text{if} & y_i<\tau \\ H_1: & y_i\in\mathscr{D}_1 & \text{if} & y_i>\tau \end{array}$$

. Here, H_0 hypothesis means that the given window contains a wheeze and H_1 hypothesis means that it is a non-wheeze window. For all possible values of τ, the ROC curve which shows the probability of false alarm $\mathrm{P}_F(\tau) = \mathrm{P}(y_i>\tau|H_0)$ vs the probability of miss $\mathrm{P}_M(\tau) = \mathrm{P}(y_i<\tau|H_1)$ in logarithmic axes is given in Fig. 6. The best value for the threshold (τ^*) is selected as the value where the total probability of error $\mathrm{P}_T(\tau) = \mathrm{P}(y_i>\tau|H_0)+\mathrm{P}(y_i<\tau|H_1)$ is minimized.

After evaluating success over the training data set, the classification performance for a test data set should also be evaluated for completeness. Since the data set is not large, leave-one-out approach is adopted for this study. Each time, one member out of 492 is left out and classified according to the new optimum threshold τ_i^* which is calculated over the remaining data set for $i=\{1,2,\cdots,492\}$. Next, the ratio of misclassified samples to the total number of data gives the test error rate.

IV. Results and Discussion

For the training data set, the number of misses, i.e. the number of wheeze windows which are detected as non-wheeze, is 18 among 246 windows and the number of false alarms, i.e. the number of non-wheeze windows which are detected as wheeze is 6 among 246 windows for τ^*. Hence, the percentage error rate for the training set is found as $24/492 = 4.88\%$. For the test data set, the number of misses is counted as 20 and the number of false alarm as 12. Thus, the percentage error rate for the test set is calculated as $32/492 = 6.5\%$. To state the results in terms of percentage success rates, the training success is 95.1% and the test success is 93.5%.

V. Conclusion

In this study, the wheeze and non-wheeze windows in the respiratory sound signals acquired from asthma and COPD patients are classified by using Fisher discriminant method and detected by using Neyman-Pearson hypothesis testing after extraction of features. Considering the studies in [4-12], the rates can be accepted high, although the proposed methodology is simpler to implement. From these results, it can be concluded that the features defined and the proposed method are promising for using on a more extensive data set. The classification and detection performance might be increased by utilizing non-linear methods and adding new features.

References

[1] Pasterkamp H, Kraman SS, Wodicka GR., "Advances beyond the stethoscope", Am J Respir Crit Care Med, 156(3 Pt 1), pp. 974-87, September 1997.

[2] Earis J.E., Cheetham B.M.G., "Current Methods Used for Computerized Respiratory Sound Analysis", Eur. Respir. Rev., vol. 10:77, pp. 586-590, 2000.

[3] Sovijrvi A. R. A., Malmberg L. P., Charbonneau G., Vanderschoot J., Dalmasso F., Sacco C., Rossi M. and Earis J. E., "Characteristics of Breath Sounds and Adventitious Respiratory Sounds" , European Respiratory Review, vol. 10, no. 77, pp. 591-596, 2000.

[4] Forkheim K.E., Scuse D., Pasterkamp H., "A comparison of neural network models for wheeze detection" WESCANEX 95. Communications, Power, and Computing. Conference Proceedings. IEEE , vol.1, pp.214-219 vol.1, 15-16 May 1995

[5] Rietveld S., Oud M., Dooijes E. H., "Classification of Asthmatic Breath Sounds: Preliminary Results of the Classifying Capacity of Human Examiners versus Artificial Neural Networks", Computers and Biomedical Research, vol. 32, no. 5, pp. 440-448, October 1999.

[6] Dokur Z., "Respiratory sound classification by using an incremental supervised neural network", Pattern Anal Applic, 2008, ISSN 1433-755X.

[7] Gler ., Polat H., Ergn U., "Combining Neural Network and Genetic Algorithm for Prediction of Lung Sounds", Journal of Medical Systems, vol. 29, no. 3, June 2005.

[8] Kandaswamy A. , Kumar C., Ramanathan R. , Jayaraman S., Malmurugan N., "Neural classification of lung sounds using wavelet coefficients", Computers in Biology and Medicine , vol. 34 , no. 6 , pp. 523 - 537, 2004.

[9] Taplidou S.A., Hadjileontiadis L.J., Penzel T., Gross V., Panas S.M., "WED: An efficient wheezing-episode detector based on breath sounds spectrogram analysis", Engineering in Medicine and Biology Society, Proceedings of the 25th Annual International Conference of the IEEE, vol.3, pp. 2531-2534, 17-21 Sept. 2003.

[10] Taplidou S.A., Hadjileontiadis L.J., Kitsas I.K., Panoulas K.I., Penzel T., Gross V., Panas S.M., "On Applying Continuous Wavelet Transform in Wheeze Analysis"‘, Engineering in Medicine and Biology Society, 26th Annual International Conference of the IEEE, vol.2, pp. 3832-3835, 1-5 Sept. 2004.

[11] Taplidou S. A., Hadjileontiadis L. J., "Wheeze detection based on time-frequency analysis of breath sounds", Computers in Biology and Medicine , vol. 37, no. 8, pp. 1073 - 1083, 2007.

[12] Homs-Corbera A., Fiz J.A., Morera J., Jane R., "Time-frequency detection and analysis of wheezes during forced exhalation", IEEE Transactions on Biomedical Engineering, vol.51, no.1, pp.182-186, Jan. 2004.

[13] I. Sen, "A Multi-channel Device for Respiratory Sound Data Acquisition and Transient Detection", M.S. Thesis, Boazii University, 2005.

[14] Richard O. Duda, Peter E. Hart, David G. Stork, *Pattern Classification*, Wiley-Interscience Publication, Second Edition, 2001.

[15] H. V. Poor, *An Introduction to Signal Detection and Estimation*, Springer–Verlag, New York, 1988.