APPLICATION OF TONAL INDEX TO PULMONARY WHEEZES DETECTION IN ASTHMA MONITORING

Marcin Wisniewski, Tomasz P. Zielinski

Department of Telecommunications, AGH University of Science and Technology
Al. Mickiewicza 30, 30-059 Kraków, Poland
phone: + (4812) 634 55 82, fax: + (4812) 634 23 72, email: wisniewski@kt.agh.edu.pl, tzielin@agh.edu.pl
web: www.kt.agh.edu.pl

ABSTRACT

In the paper a new application of tonal index, well-known signal feature from MPEG-audio psychoacoustic model, to pulmonary wheezes detection at patients with asthma diseases is proposed. Efficiency of this approach is compared with application of some other acoustic signal features which have already been tested in asthma monitoring (e.g. spectral peaks entropy, frequency ratio and kurtosis) and superiority of the tonal index is shown. Experiments have been conducted on artificially generated wheezes-like signal that have being embedded in real pulmonary noise (hybrid signal) as well as on recorded real wheezes. The SVM classifier was used in recognition process.

1. INTRODUCTION

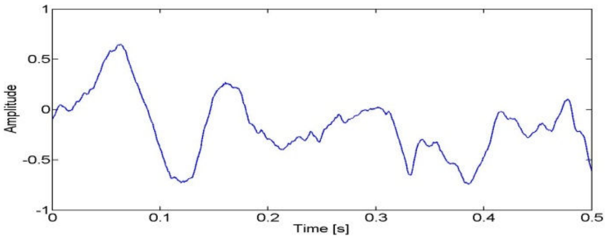
Modern medical systems can successfully detect and diagnose many disorders only on the basis of simple and none-invasive examination of different human signals. What is very important, patients can make tests without any supervision by medical doctors. Acoustic signal analysis is very popular in medicine. One of the most popular area of its usage is investigation of lung sounds obtained during stethoscope auscultation.

Asthma is one of the common chronic diseases which can be well controlled by lung sound analysis [1]. However, the auscultation analysis is only one of tasks in asthma attacks prevention. The asthma remote patient monitoring e-health systems should also offer possibility of pollution monitoring and patient self spirometry tests. On the basis of information send to external medical servers, the patient receives a general or drug recommendations from medical doctors. The history of the patient's disease can be analysed as well so the remote supervision can be more precise. In certain areas the e-health systems can inform the patient about his present health state or dangerous situations without personal medical doctor's care (intervention). The short review of asthma e-health systems can be found in [2, 3] and they will not be discussed in this paper. Such system are now commercially available [4].

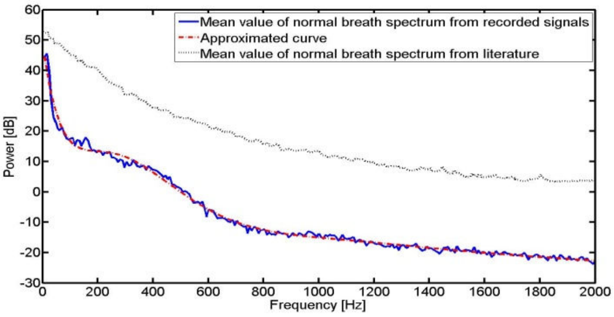
The main reason of lung auscultation in asthma is a wheeze detection. In patient with exacerbating asthma, tonal sounds, apart from breathing noise, are present in his lung sounds. Computer algorithms detect and evaluate percentage of wheezing in normal breathing, evaluate its location in the breathing cycle, or just indicate appearance of the disorder. Typically, wheeze recognition methods are based on signal features defined in three different domains: time, frequency or joint time-frequency. The time domain features are not very popular in wheezes detection but they have been used in some approaches e.g. kurtosis [5]. The most frequently are exploited Fourier spectrum frequency-based features, among them f_{50}/f_{90} ratio [5], spectral peaks entropy [6], sample entropy [7] and Renyi entropy [5]. The first three of them will be shortly described in the next section. Time-varying extensions of the frequency features based on the short-time Fourier and wavelet transform coefficients are often used too.(for example wavelet bispectrum [8]). The anomaly detection using image processing and pattern recognition performed on matrices of time-frequency spectra can be also found in literature [9]. For recognition of the breathing state linear [5] and nonlinear (e.g. based on optimally weighted Wigner–Ville distributions) [10] classifiers have been used.

In this paper a new application of tonal index, well-known from MPEG-audio psychoacoustic model [11], to pulmonary wheezes detection at patients with asthma diseases is proposed. Efficiency of this approach is compared with application of some other acoustic signal features which have already been tested in asthma monitoring (e.g. spectral peaks entropy, frequency ratio and kurtosis) and superiority of the tonal index is shown. In addition, the spectral flatness, the feature from speech analysis, is tested. Experiments have been conducted on artificially generated wheezes-like signal that have being embedded in real pulmonary noise as well as on recorded real wheezes. Since we aimed at features not classifier testing, the time-consuming but efficient (precise) SVM classifier was used in the recognition process.

Structure of the paper is as follows. In section 3 the short description of normal and wheezy breath signal is given, and all features being compared with the tonal index are presented. In section 4 the tonal index definition and calculation algorithm are shortly described and a piece of information about introduced modifications is given. Section 5 presents the methodology and results of particular tests, the process of data modelling is presented in it as well. In the end, the final conclusions are written in section 6.



a)



b)

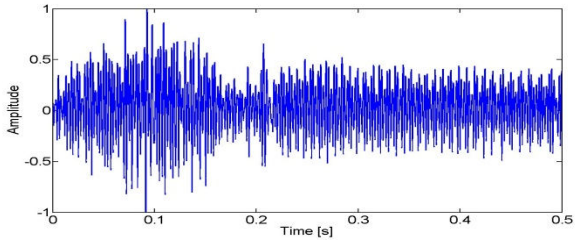
Figure 1 – a) Signature of the recorded normal breath sound, b) typical Fourier spectrum curve of a normal breath taken from literature [1] and calculated (a mean-value one) by us using recorded signals

2. ACKNOWLEDGEMENTS

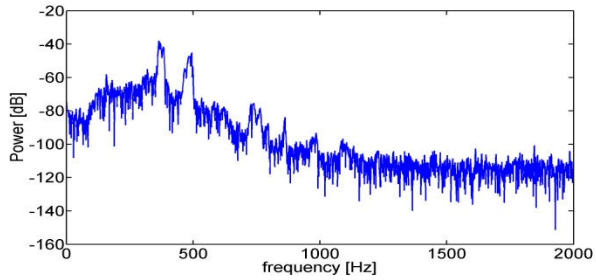
This work has been (partially) supported by the Polish Ministry of Science and Higher Education under the European Regional Development Fund, Grant No. POIG.01.01.02-00-045/09-00 *Future Internet Engineering*.

3. LUNG WHEEZES

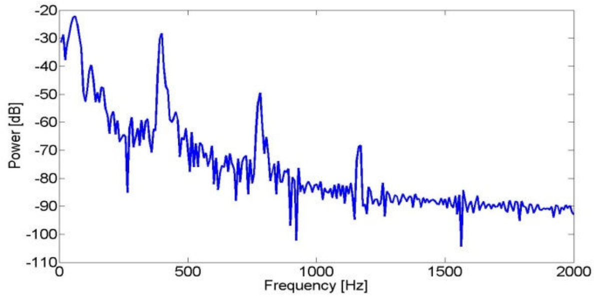
The normal breath sounds taken from the chest wall is a noise that peaks in frequency approximately around 100 Hz and drops sharply between 100 and 200 Hz but it can also be detected near 1 kHz by sensitive microphones as well. The signal is mixed with muscles and cardiovascular sounds. The sounds taken from trachea have wider range of frequency and the spectrum shape is a little bit different. In this paper tests were performed only on acoustical chest auscultation signals and the trachea sounds are not discussed. In order to find experimentally the theoretical spectrum of normal breath, the mean spectrum of 80 realisations of real breathing was calculated and approximated by polynomial fitting (Fig. 1b – red dashed line). The wheeze is a single or multi tone (frequency) sound with duration higher than 80 ms and frequency between 100 Hz and 1 kHz [1]. The wheezing can be monophonic or polyphonic but it is usually the polyphonic one. In Fig. 2a a wheezing patient recording is shown and its tonal nature is visible. The shape of the wheeze spectrum (Figs. 2b and 2c) is different from the normal breath . Noticeable hump of frequencies between 50Hz and 1kHz is usually observed in it and lower frequencies are often suppressed (Fig. 2b). However, the worst case for wheezes identification is when this low-frequency suppression does not occur (Fig. 2c)



a)



b)



c)

Figure 2 – a) a wheeze signal, b) c) examples of Fourier spectra of two wheezes taken from [12].

and when the recording is consisted of the mixture of normal noise-like lung sounds, having very strong low-frequency noise part, and a multi-frequency tonal wheeze. Such situation was assumed in our experiments.

4. WHEEZE DETECTION FEATURES

To detect a multi-tone component of the breathing sound till now different signal features have already been used. The most important ones and tested in this paper are briefly described below.

Kurtosis [5] **–** *(K)* is a feature which describes data in time domain. It measures a level of peakedness of a probability distribution and it is defined as follows:

$$k = \frac{E(x-\mu)^4}{\sigma^4} \tag{1}$$

where x denotes an input signal while μ and σ^2 - its mean and variance, respectively. The value of kurtosis of noisy signal with normal or sub-gaussian distribution reach approximately 3 while the signal with wheezes has values above 3. We found that more beneficial is calculation of the kurtosis not of a time signal but its signal spectrum expressed in decibels and results for these two method will be presented later.

Spectral peaks entropy [6] **–** *(SPE)* is defined also in frequency domain. In this case, first, a ratio between each peak value (c_1, c_2,..., c_N) of the signal frequency spectrum and the total sum of these peaks is calculated:

$$p_n = \frac{c_n}{\sum_{n=1}^{N} c_n}, \tag{2}$$

and then the entropy is to be found:

$$En = -\sum_{n=1}^{N} p_n \log_{10}(p_n) \tag{3}$$

f_{50}/f_{90} ratio [5] **–** *(FR)* is the last proposed frequency feature of our interest. It is given as a ratio of two frequencies f_{50} and f_{90} for which the area under the power spectral density (from $f = 0$ to $f = f_{50 \text{ or } 90}$) to the total signal power (the whole area under the PSD curve) reaches values of 50% and 90%, respectively. The signal with wheezes has higher values of this ratio than normal lung sounds.

Spectral flatness [13] **–** *(SF)* is a signal feature defined in frequency domain giving information whether the signal is more tone-like or more noise-likeone. It is defined as a ratio of geometrical and arithmetical mean values of the signal spectrum $X(k)$:

$$Fl = \frac{\sqrt[N]{\prod_{k=0}^{N-1} X(k)}}{\frac{\sum_{k=0}^{N-1} X(k)}{N}} \tag{4}$$

The feature has not been used in wheezes detection till now but in this article is additionally tested.

5. TONAL INDEX

In this paper a new application of the tonal index *(TI)* to lung wheezes detection is proposed and efficiency of such method is evaluated in respect to the signal features listed above. The tonal index is well known from MPEG-audio standard [11] in which it is used in the psychoacoustic model no 2 to switch a working mode of the second filter bank in the MP3 encoder. After the first decomposition of the acoustic signal into 32 frequency subbands, the second division is done in which every one of 32 frequency band-pass signals is additionally divided into 18 sub-channels sampled twice for tonal sounds (for them better frequency resolution is required) or into 6 sub-channels sampled 6-times each for noisy sounds (in this case better time resolution is advantageous. The signal splitting is based on the perceptual entropy of the FFT spectrum of the signal: small entropy (<1800) characterizes the tonal not noisy sounds. In turn, the tonal index is used in MPEG-audio standard for perceptual entropy calculation.

Therefore, the tonal index is a spectral feature that can be treated as alternative for the measures of spectral flatness and spectral peaks entropy presented above. It is based on calculation of FFT module $\overline{r_\omega}$ and phase $\overline{p_\omega}$ prediction:

$$\overline{r_\omega} = 2r_\omega(t-1) - r_\omega(t-2) \tag{5}$$

$$\overline{p_\omega} = 2p_\omega(t-1) - p_\omega(t-2) \tag{6}$$

in which 2 previous FFT blocks (t–1) and (t–2) are taken into account. The module and phase prediction are necessary to calculate the degree of unpredictability c_ω (7) energy e_b (8) and weighted unpredictability c_b (9) where ωlow and $\omega high$ denotes the border of frequency region of interest.

$$c_\omega = \frac{\sqrt{\left[r_\omega \cdot \cos(p_\omega) - \overline{r_\omega} \cdot \cos(\overline{p_\omega})\right]^2 + \left[r_\omega \cdot \sin(p_\omega) - \overline{r_\omega} \cdot \sin(\overline{p_\omega})\right]^2}}{r_\omega + \left|\overline{r_\omega}\right|} \tag{7}$$

$$e_b = \sum_{\omega=\omega low}^{\omega high} r_\omega^2, \quad c_b = \sum_{\omega=\omega low}^{\omega high} c_\omega r_\omega^2 \tag{8)(9}$$

The main difference between calculating the tonal index in the MPEG psychoacoustic model and in our algorithm for wheezes detection relies on missing a spreading function (simulating a human ear) in our approach. In wheezes detection, there is no need to such simulation, so the spreading function is skipped. The final index Ti is calculated as follows:

$$Ti = \log_{10}\left(\frac{c_b}{e_b}\right) \tag{10}$$

6. EXPERIMENTS

6.1. Wheezes modelling and recognition

Due to the fact that we are having a very small data base of recorded real wheezing lung signals taken from Internet, an initial phase of our research aiming at testing wheeze recognition features was based on modelled wheezes. According to the mentioned before fact that the worst case of a wheezy signal to detect is a sum of the tonal multi-frequency wheeze and the noisy-like breath, in the first step artificial wheezes were generated as multi-frequency signals with random three frequencies lying in the range [100, 1200] Hz (tones with higher frequencies had lower but random amplitudes and random phases). Then, tonal signals were mixed in a proper scale (Signal-to-Noise Ratio) with normal breathing signals: 1) real recorded (hybrid signal) or 2) generated artificially (artificial signal). After signal (data) modeling application of each feature was tested individually for different SNR levels. All spectral features, except TI, were calculated on smoothed (averaged in time direction) spectra scaled in decibels. Such smoothing allowed to highlight the tonal peaks and suppressed the noisy character of the spectrum. The TI feature did not use such operation because its merit relies on tracking the spectrum unpredictability (dynamic change). For recognition the SVM (Support Vector Machine) classifier was used with 3rd order polynomial kernel function. Wheeze detection efficiency was calculated using the following formula:

$$AC = \frac{TN + TP}{TN + TP + FN + FP} \cdot 100\% \tag{11}$$

where TP – true positive, TN – true negative, FP – false positive and FN – false negative.

6.1.1 Wheeze recognition in hybrid signal

First, usefulness of different signal features was verified on artificially generated wheeze-like three-component/frequency signals which were added to real breath samples.

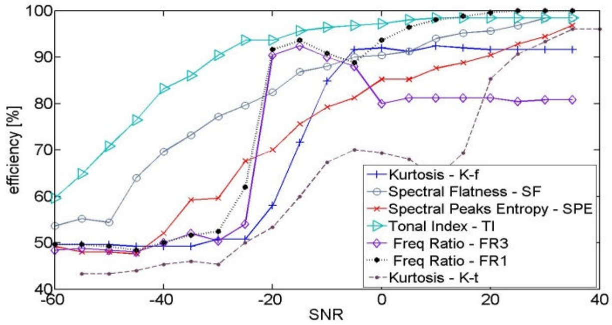


Figure 3 – Wheeze recognition efficiency for single feature and semi artificial data. K-t *and* K-f *denotes kurtosis calculated in time and frequency, respectively while FR-1 and FR-3 stands for simulated single or three component wheeze.*

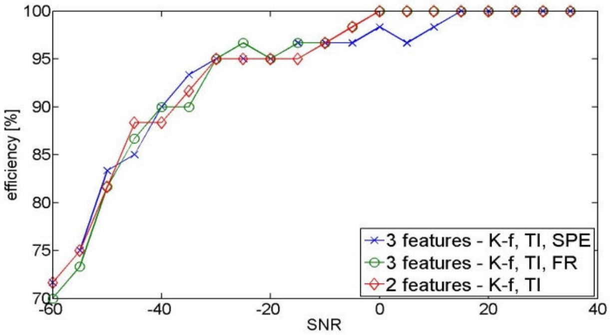


Figure 4 – Wheeze recognition efficiency for selected 2 or 3 features and semi artificial data

Signals of real breath were recorded from chest auscultation using 8 kHz/16-bit recorder and Panasonic WM-61 microphones. The microphones were connected to the stethoscope heads. Efficiency of each feature was evaluated using 50 samples of normal lung sounds (25 for training and 25 for testing) and 50 samples of mixed normal breath and artificial wheeze signal (again, 25 + 25). 10 different selections were done and an average efficiency value was calculated. Each sound sample was 1024-point long. Obtained results are presented in Fig.3: recognition efficiency AC (11) in percentage against different SNR. The learning and recognition process was made on 50 different signals each. One can see a superiority of the TI method in the whole SNR range. The second is the SF method. The FR feature works well only for single component wheezes. The K-f method outperforms the K-t one (which also works well only for single component wheezes). After feature test, the best sets of features were chosen to multi feature recognition tests. The verification process was made on the same data as before. Obtained results are shown in Fig.4.

6.1.2 Wheeze recognition in artificial signal

In the second test, on the basis of the real lung recordings the mean value of normal breath spectrum was calculated (see Fig. 1b) and its curve was next used to model a normal noisy breath sound: generated random white noise was filtered using it. This way artificial normal lung sounds were

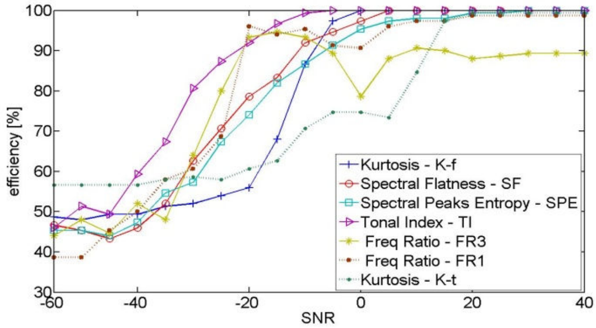


Figure 5 – Recognition for each feature for artificial data. K-t *and* K-f *denotes kurtosis calculated in time and frequency, respectively while FR-1 and FR-3 stands for simulated single or three component wheeze*

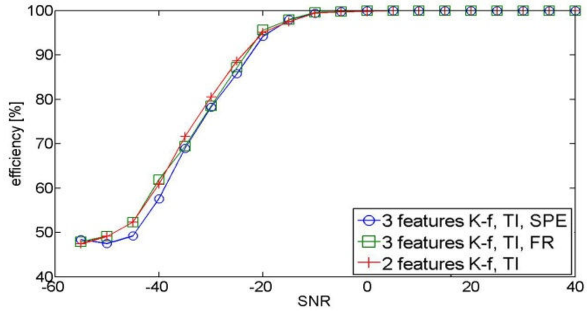


Figure 6 – Wheeze recognition for selected 2 or 3 features and artificial data

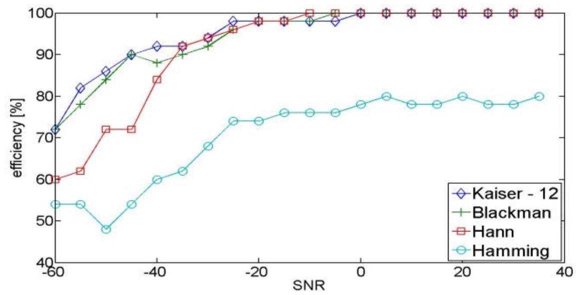


Figure 7 – Wheeze recognition efficiency using tonal index with different time windows.

1000 testing samples. Obtained results are presented in Fig. 5. Then, recognition experiment was repeated for selected different sets of features (Fig. 6). Curves from figures 5 and 6 confirm those from figures 3 and 4 and the previous remarks.

6.1.3 Window choice in tonal index calculation

To check efficiency of tonal index, different windows was applied to algorithm: Hann, Hamming, Blackman and Kaiser (β = 12). Verification was done for different SNR for real noise. The results are presented in Fig.7. For the low SNR application of the Kaiser window is the best while for high SNR (≥ -25 dB) usage of the Hann window is beneficial.

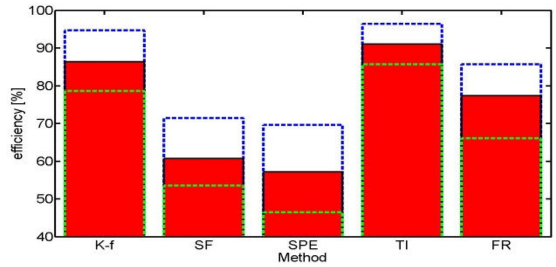


Figure 8 – Wheeze recognition efficiency for real (recorded) data for each different feature. Mean value after 10 tests; dotted line – maximal and minimal value

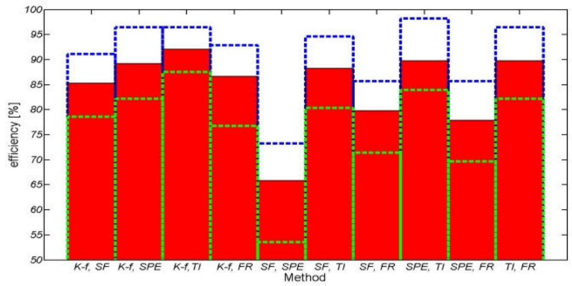


Figure 9 – Wheeze recognition efficiency for real (recorded) data for different sets of features. Mean value after 10 tests; dotted line – maximal and minimal value

6.2. Recognition of real wheezes

In order to verify the previous results on real data, some tests on recorded wheezes, obtained from the Internet [12, 14, 15], were done. To learn the algorithm , 28 normal lungs sounds and 28 wheezes have been used while for testing – 28 normal breathings and 28 wheezes. However, the database built this way was very poor since types of recordings (lungs, trachea) and recording devices (microphones, frequencies, number of bits) were different. Taking into account the small number of recordings, the test was made 10 times with random selected signals from the whole database. Therefore, the result of such test are less authoritative but they are still quite good, as can be observed in Fig. 8 for each feature and Fig. 9 for sets of features.

7. CONCLUSION

The presented results prove that the tonal index feature, well known from the MPEG-audio standard, can be successfully used to wheezes recognition process. The tonal index is more sensitive to tonality in noisy signals (with low SNR, weaker wheeze) than any other feature presented in this paper and it reaches full efficiency in lower SNR as well. To increase its efficiency it was necessary to choose an optimal time window. The result of different windows testing shows that the differences in efficiency can reach even 20%. The tonal index with Hamming window, used in MPEG-audio standard, was significantly worse than other methods and it did not reach 100% recognition efficiency. The results shows that the tonal index could detect the wheezes as one independent feature as well.

During comparison of different signal features the Kaiser window (β = 12) was used in tonal index computation algorithm since such choice ensure better wheeze recognition efficiency in lower SNR (–70 : – 65) than other tested windows (the efficiency for higher SNR is almost identical for different windows). It was observed that the f_{50}/f_{90} ratio works better for single-component wheezes and the K-f (kurtosis-in-frequency) method outperforms the K-t (kurtosis-in-time) one. The future research will be concentrated on verification of the presented results on extensive real data bases of lung sounds.

8. REFERENCES

[1] H. Pasterkamp, S. S. Kraman, G. R. Wodicka, *Respiratory Sounds. Advances Beyond the Stethoscope*, Am. J. Respir. Crit. Care Med., Volume 156, Number 3, pp. 974-987, September 1997

[2] Finkelstein, J. ; Cabrera, M.R. ; Hripcsak, G.; *Web-based monitoring of asthma severity: a new approach to ambulatory management*, Information Technology Applications in Biomedicine, 1998. pp: 139 - 143

[3] M. Wiśniewski, T. Zieliński, *E-health: the remote patience monitoring systems in asthma diseases*, Electrical Review 86 no 9 pp. 72–75, 2010

[4] "WHEEZE RATE - A New Paradigm in Asthma Management", Karmel Sonix Co, http://www.karmelsonix. com/

[5] Aydore, S.; Sen, I.; Kahya, Y.P.; Mihcak, M.K., *Classification of respiratory signals by linear analysis*, Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE Publication Year: 2009, pp. 2617 - 2620

[6] Jianmin Zhang; Wee Ser; Jufeng Yu; Zhang, T.T.; *A Novel Wheeze Detection Method for Wearable Monitoring Systems* Intelligent Ubiquitous Computing and Education, 2009 International Symposium on Publication Year: 2009 pp. 331 - 334

[7] Jin, Feng; Sattar, Farook; Goh, Daniel Y.T.; *Automatic wheeze detection using histograms of sample entropy,* Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE Publication Year: 2008 , pp: 1890 - 1893

[8] S.A. Taplidou, L.J. Hadjileontiadis, *Analysis of wheezes using wavelet higher order spectral features*, IEEE Transactions on Biomedical Engineering, 57 (7), pp. 1596- 1610. 2010

[9] R.J. Riella, P. Nohama, R.F. Borges, A.L. Stelle,; *Automatic wheezing recognition in recorded lung sounds,* Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE Vol: 3: 2003 , pp 2535 - 2538

[10] F. Jin, F. Sattar, *Tracking and time-frequency analysis on nonlinearity of tracheal sound*, Medical and Biological Engineering and Computing Volume 47, Number 4, pp. 457-461, February 2009

[11] J.D. Johnston, *Estimation of Perceptual Entropy Using Noise Masking Criteria*, Acoustics, Speech, and Signal Processing, ICASSP-88, 1988 , pp. 2524 - 2527 vol.5

[12] http://www.faemse.org/downloads.html

[13] A.H. Gray, J.D. Markel, *A spectral-flatness measure for studying the autocorrelation method of linear prediction of speech analysis*, IEEE Trans. Acoust. Speech Signal Process., 1974, 22, pp. 207–217

[14]http://solutions.3m.com/wps/portal/3M/en_US/Littmann/stethoscope/education/lung_sounds/

[15]http://faculty.etsu.edu/arnall/www/public_html/heartlung/breathsounds/contents.html