FEATURE EXTRACTION FROM ASTHMA PATIENT'S VOICE USING MEL-FREQUENCY CEPSTRAL COEFFICIENTS

Rachna[1], Dinesh Singh[2], Vikas[3]

[1] *M.Tech Student, ECE Department, DCTM Palwal, Haryana, India*
[2] *Professor, ECE Department, DCTM Palwal, Haryana, India*
[3] *M.Tech Student, School of VLSI Design & Embedded System, NIT Kurukshetra, Haryana, India*

Abstract

Asthma affects the airways to and fro from human lungs, due to blockage of air breathing problem also occurs. People suffering from this type of problems are said to be asthmatic. In this paper a feature extraction technique is applied for extracting similar features in asthmatic patients and normal human being. Mel – frequency Cepstral Coefficient technique is used for feature extraction process. As there is no proper cure for this disease and treatment is too costlier, here is a start up of diagnosing this disease in initiation stages using this technique.

Key Words: *Mel – frequency Cepstral Coefficient (MFCC), voice, asthma, disease, feature extraction.*

---***---

1. INTRODUCTION

Asthma is a lung disease that affects airflow to and fro from lungs. A whistling sound comes when asthmatic patient breathes. Major symptoms of asthma are chest stiffness, breathe shortness, cough production during night and morning etc. It affects all ages of persons, around 20 million persons in India affected from asthma including children's [1] and around 25 million persons affected in U.S including 7 million of children's [2]. When the airways starts working muscles around airways becomes tight and less air can flow into the lungs, also due to swallowing of muscles around airways. When airways blocked then asthma attack may occur as shown clearly in fig. 1. Since, 1970's bronchial asthma has increased continuously that affects approximately 4 to 7 % of the persons in the world [3]. Asthma has no cure, just it can be controlled [2]. Major risk factors are bedding dust, carpet, furniture dust, also family history or allergy. According to the study of South Atlantic Island of Tristan da Cunha, asthma problems has been found in children's those parents are asthmatic [2]. It can be controlled during asthma stages by doing long term meditation daily, regular check up by doctor in case of serious patients, taking some drugs through inhalers when asthma attack came etc. These treatment method may be very costlier to a normal or poor persons. This paper doing a start up for diagnosing asthma in initial stages, so that it can be controlled. Mel-frequency Cepstral Coefficients (MFCC) are the very popular parameter of voice that can view auditory system of human more closely as other parameters do. In this paper MFCC are calculated using PRAAT and MATLAB for feature extraction process. Further these extracted coefficients will be analyzed for finding similarities between patients and normal persons.

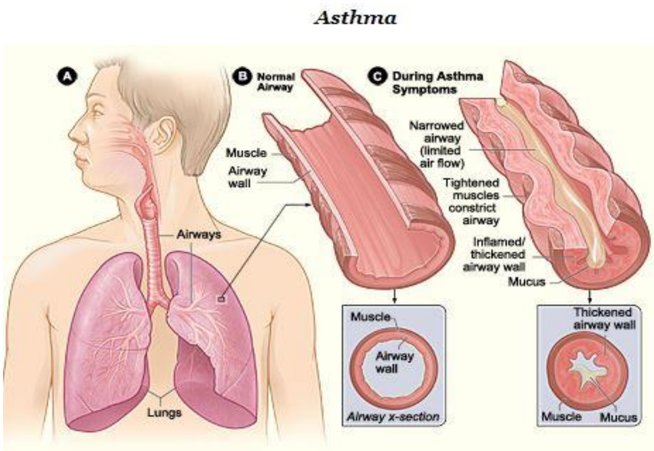

Fig – 1: 'A' showing structure of lungs and airways in human body, 'B' showing normal airways structure and 'C' showing airways when asthma occurs.

2. Mel-frequency Cepstral Coefficient (MFCC)

Mel- frequency Cepstral Coefficeints are the most commonly used acoustic parameter of voice for speech recognization. Procedure/steps to find out MFCC is shown in fig. 2.

2.1 Pre- Emphasis

In this block sound signal is passed through a high pass filter. Let sound signal is x(n), then

$$x_2(n) = x(n) – a*x(n-1) \qquad (1)$$

where $x_2(n)$ is the output of filter and a is normalisation factor that varies between 0.9 to 1.0.

Pre-emphasis compensate the suppressed part of signal during sound production and it also improves or amplify the sound signal for better results [4].

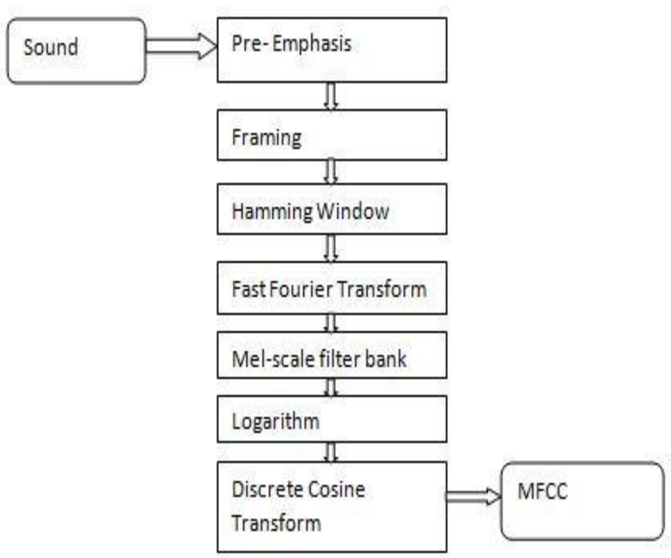

Fig -2: Procedure for extracting Mel-frequency Cepstral Coefficients [5].

2.2 Framing

In this block pre-emphasized signal is segmented into frames of 30-40 ms with optional overlap of 1/3-1/2 of the frame size. Framing is very important part for good results because variation of amplitude is more in larger signals as compared to smaller signals. If the sample rate is 15 kHz and frame size is 300 sample points, then frame duration is 300/15000 = 0.02 seconds i.e. 20 ms [4].

2.3 Hamming Window

In this block all frames will be multiplied with a hamming window in order to keep the continuity of the first and last points in the frame.
Here sound signal is denoted by x(n), where n = 0,1,2.....N-1, then after multiplying signal with hamming window output of this block is

$$w(n,\alpha) = x(n) * w(n) \qquad (2)$$

$$w(n,\alpha) = (1-\alpha) - \alpha \cos((2*pi*n)/(N-1)), \quad \text{where } 0<=n<=N-1 \qquad (3.)$$

Values of 'α' will be different for different windows, normal it is used as 0.46. Basic plot of hamming window shown in fig. 3 [4].

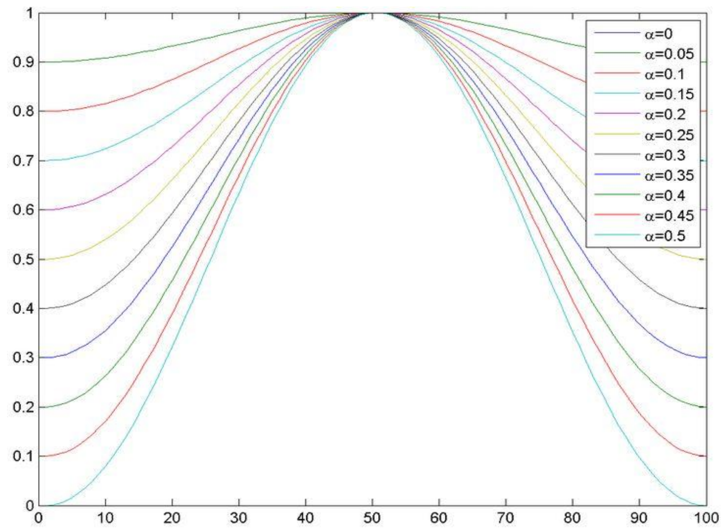

Fig -3: Basic plot of hamming window using MATLAB.

2.4 Fast Fourier Transform

Fast fourier transform converts time domain signal in frequency domain signal. While doing this transform we will assume that signal is periodic within frame. If signal is not periodic within the frame, then also we compute this transform but discontinuity comes at start and end points of the frame. For dealing with this situation we have two options.

1. For increasing signal continuity at first and last points multiply all frames with hamming window.
2. Take a frame of variable size [4].

Fig 4. showing energy spectrum using FFT for original signal and windowed signal.

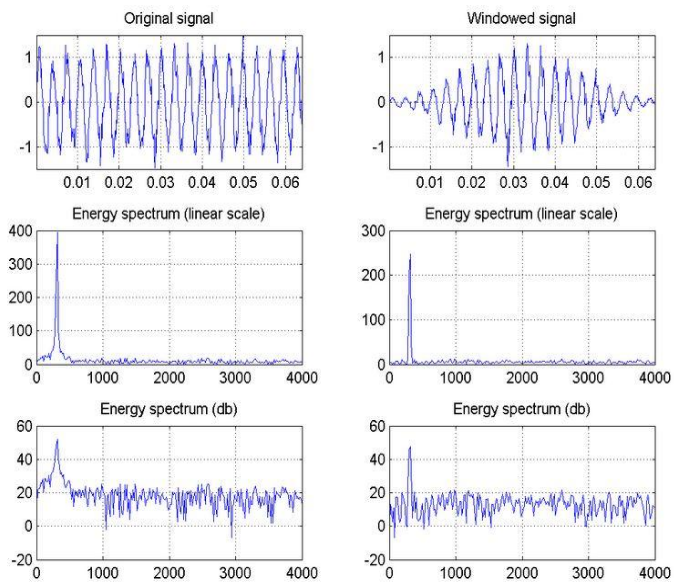

Fig -4: Showing effect of hamming window multiplication in energy spectrum.

2.5 Mel- scale filter bank

In this block output of fast fourier transform block multiplied by a set of 20 triangular bandpass filters for getting log energies of each filter. All these filters are equally spaced along mel frequency. Basic formula for converting frequency to mel-scale is as follows [4].

$$\text{Mel(freq.)} = 1125 * \ln(1 + f/700) \qquad (4)$$

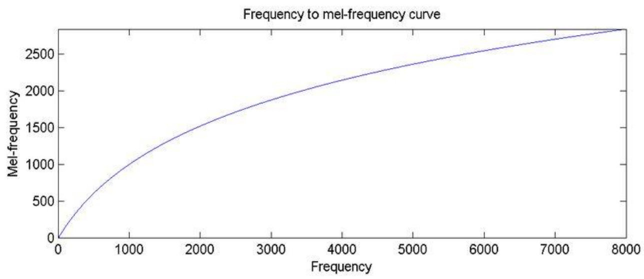

Fig -5: Plot of mel-frequency scale using MATLAB.

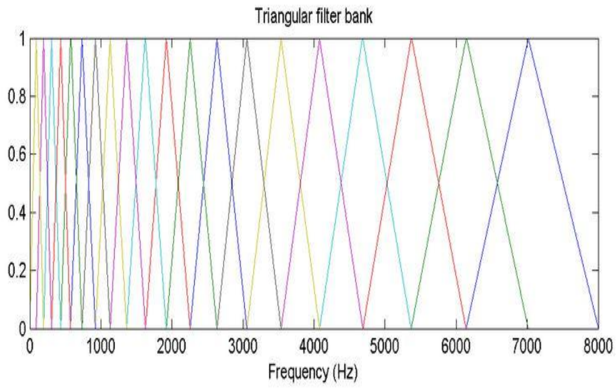

Fig -6: Plot for a set of 20 triangular bandpass filters using MATLAB.

2.6 Logarithm and Discrete Cosine Transform (DCT)

Output of Mel-scale fiter bank then passed through logarithm block. This block is used for normalization purpose. After that normalized signal is then passed through DCT block that de-correlates the log energies of filters. Finally the output of DCT block provides us the MFCC coefficient values.

2.7 Collection of asthma patient voice samples

For feature extraction process vowels (a, e, i, o, u) voice of asthmatic and normal people has been recorded using Samsung mobile recorder (Star Pro). Then all signals are converted into .wav format from .wma format for analysis purpose as MATLAB and PRAAT only runs on .wav format files.

3. Results and Discussion

In this analysis 20 coefficients of MFCC were computed for feature extraction purpose. After computing these coefficients using MATLAB large variation has been seen in the first coefficient of asthmatic persons. Some of the variations are shown in fig 7, 8 and table 1, 2.

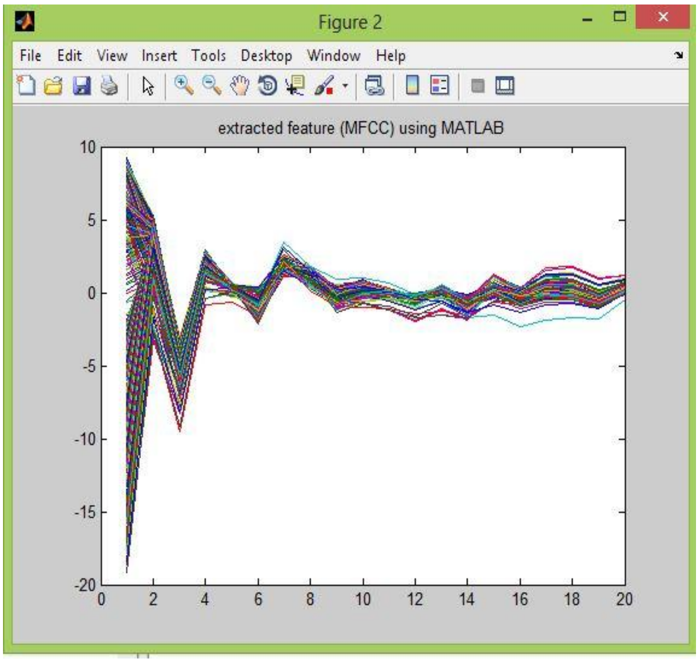

Fig -7: Extracted feature using MFCC for normal person where X axis showing 20 coefficients.

Table – 1: These are values of 20 MFCC coefficients with five frames out of 256 frames for normal person where X1-X20 are Coefficients and F1-F5 are frames.

	F1	**F2**	**F3**	**F4**	**F5**
X1	-17.003	-17.892	-18.661	-17.440	-16.521
X2	-1.127	-2.341	-3.255	-1.546	-0.222
X3	-6.772	-7.727	-8.329	-7.095	-6.293
X4	1.103	0.343	0.166	0.941	1.369
X5	0.263	-0.176	0.124	0.457	0.478
X6	0.177	-0.105	0.407	0.316	-0.035
X7	2.122	1.955	2.565	2.430	2.095
X8	1.428	1.348	1.798	1.753	1.466
X9	-0.882	-1.044	-1.002	-0.782	-0.799
X10	-0.157	-0.290	-0.656	-0.084	-0.158
X11	-0.966	-0.328	-1.194	-0.439	-0.028
X12	-0.584	-0.839	-1.973	-1.106	-0.565
X13	0.542	0.181	-1.144	-0.531	-0.151
X14	-0.459	-0.758	-1.870	-1.526	-1.269
X15	1.269	1.140	0.485	0.457	0.437
X16	0.245	0.361	0.324	0.009	-0.202
X17	0.850	1.167	1.668	1.155	0.786

X18	0.4907	0.953	1.799	1.252	0.834
X19	-0.140	0.240	1.024	0.522	0.136
X20	0.439	0.695	1.204	0.943	0.733

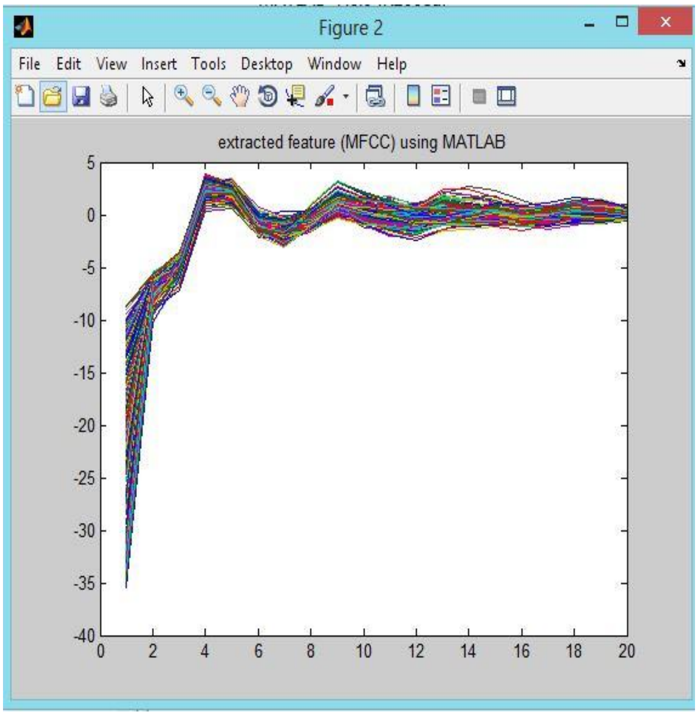

Fig -8: Extracted feature using MFCC for asthmatic person where X axis showing 20 coefficients.

Table – 2: These are values of 20 MFCC coefficients with five frames out of 256 frames for asthmatic person where X1-X20 are coefficients and F1-F5 are frames.

	F1	**F2**	**F3**	**F4**	**F5**
X1	- 33.005	- 32.904	- 32.798	- 32.186	- 32.655
X2	-6.694	-6.569	-6.569	-6.890	-7.420
X3	-4.426	-4.313	-4.302	-4.805	-5.312
X4	1.731	1.812	1.859	1.767	1.565
X5	2.209	2.295	2.407	2.464	2.462
X6	-0.962	-1.040	-1.024	-0.916	-0.762
X7	-1.722	-1.814	-1.835	-1.720	-1.528
X8	0.134	0.127	0.051	-0.029	-0.124
X9	1.782	1.750	1.587	1.3060	0.944
X10	0.810	0.811	0.730	0.500	0.143
X11	0.582	0.422	0.276	0.75	-0.209
X12	0.481	0.093	-0.214	-0.432	-0.630
X13	0.749	0.255	-0.142	-0.363	-0.479
X14	0.557	0.188	-0.148	-0.363	-0.479
X15	0.659	-0.556	0.444	0.336	0.244
X16	0.083	0.225	0.353	0.395	0.369
X17	0.067	0.274	0.468	0.580	0.614
X18	0.716	0.819	0.903	0.983	1.042
X19	0.428	0.367	0.289	0.294	0.3605
X20	0.045	-0.016	-0.090	-0.083	-0.008

4. Conclusion

Features were extracted for all voice samples in the database using Mel-frequency Cepstral Coefficients in MATLAB. After analyzing all extracted features for both asthmatic and normal persons it has been observed that there is a large variation in coefficients of asthmatic persons as compared to normal persons especially in the 1^{st} and 2^{nd} coefficients. In future feature extraction will be done for a large database for finding better results, also feature matching will be done for both asthmatic and normal persons.

5. References

[1]. http://www.who.int/mediacentre/factsheets/fs206/en/\

[2].http://www.nhlbi.nih.gov/health/health-topics/topics/asthma/

[3]. Ranabir Pal, Sanjay Dahal and Shrayan Pal, "Prevalence of Bronchial Asthma in Indian Children", Indian J Community Med. Oct 2009; 34(4): 310–316.

[4].http://mirlab.org/jang/books/audiosignalprocessing/speechFeatureMfcc.asp?title=12-2%20MFCC

[5]. Wei HAN, Cheong-Fat CHAN, Chiu-Sing CHOY and Kong-Pang PUN. "An Efficient MFCC Extraction Method in Speech Recognition", IEEE International Symposium on Circuits and Systems, 2006. ISCAS.

[6]. Kashyap Patel, R.K. Prasad, "Speech Recognition and Verification Using MFCC & VQ", International Journal of Emerging Science and Engineering (IJESE) ISSN: 2319–6378, Volume-1, Issue-7, May 2013

[7]. Vibha Tiwari, "MFCC and its applications in speaker recognition", International Journal on Emerging Technologies 1(1): 19-22(2010).