iJETRM
International Journal of Engineering Technology Research & Management

HUMAN VOICE PROCESSING TO EXTRACT MFC COEFFICIENTS FOR EARLY DETECTION OF ASTHMA

Sumit Kumar
Mardi Srikar
Mr. T.P. Mithun
RV College of Engineering, Bengaluru, India
kumarsumit.7w@gmail.com
srikarmardi@gmail.com
mithuntp@rvce.edu.in

ABSTRACT
Asthma is one of the top five chronic diseases globally, especially among children. It is a lung disease, which affects airflow back and forth from lungs. The airways i.e. bronchial tubes in the human lungs swells up, due to which breathing becomes hard. Infants and children below 10 years are more vulnerable to it. Due to lack of cooperation from children, the usual methods of asthma analysis fail to provide good results. A clinical diagnosis of asthma can be made with certainty by age of five, so early monitoring and diagnosis of respiratory symptoms is essential. Thus, a computerized respiratory sound analysis is an important diagnostic aid. The method utilizes a speech processing technique known by the short form MFCCs. It stands for Mel-frequency cepstral coefficients. They are like 'features' extracted from audio signals, similar to edges that constitute as a feature extracted from images. Repeated episodes of wheezing is a symptom universally accepted as the asthma diagnosis in children. MFCC technique is applied for extracting a set of features in a test person and by using MFCC as features; wheezing can be detected with accuracy higher than 95%. In this method, an algorithm is devised to calculate a Mel Frequency Cepstrum Coefficients vector of length 13 for a short voice message of a test person to compute his/her probability of being asthmatic as a result.

KEYWORDS:
MFCC, Asthma, Mel, Filterbank, Cepstrum, Discrete Cosine Transform (DCT), Framing, Fourier Transform

INTRODUCTION

Asthma is a public health problem not just for high-income countries; it occurs in all countries regardless of the level of development. The World Health organization (WHO) estimates a number of 235 million people currently suffering from asthma [6]. Asthma is considered the most common chronic disease among children in nearly all industrialized countries. Since a clinical diagnosis of asthma can usually be made with certainty by the age of five, early diagnosis, monitoring and treatment of respiratory symptoms are essential [2]. The presence of wheezing in infants is used as an important parameter in assessing the predisposition to asthma [3]. History of repeated episodes of wheezing is a symptom universally accepted as the starting point for asthma diagnosis in children [4]. A wheeze can be described as an unintentional and continuous sound [5]. Acoustically, it is characterized by periodic waveforms with a dominant frequency usually between 100 Hz to 400Hz and with duration of ≥100 ms.

Recently developed computer based respiratory sound analysis methods serve as a powerful tool to diagnose the whole spectrum of disorders and abnormalities in the lungs, including asthma [1]. However, special attention should be made when this kind of analysis is applied to signals recorded on children. Due to their lack of cooperation, such signals have a number of artefacts and the usual methods of analysis often do not provide good results [7]. The authors Sirko Molau, Michael Pitz, et. al. has presented a method to derive Mel-frequency cepstral coefficients directly from the power spectrum of a speech signal. The presented approach simplifies the speech recognizer's front end by merging subsequent signal analysis steps into a single one [8]. The ever increasing cost spend on asthma cure or better said prevention, motivated us to develop a method to detect it in early stages and take appropriate steps to prevent it in further. The aim is to design a technique for detection and starting of the treatment as soon as possible.

iJETRM
International Journal of Engineering Technology Research & Management

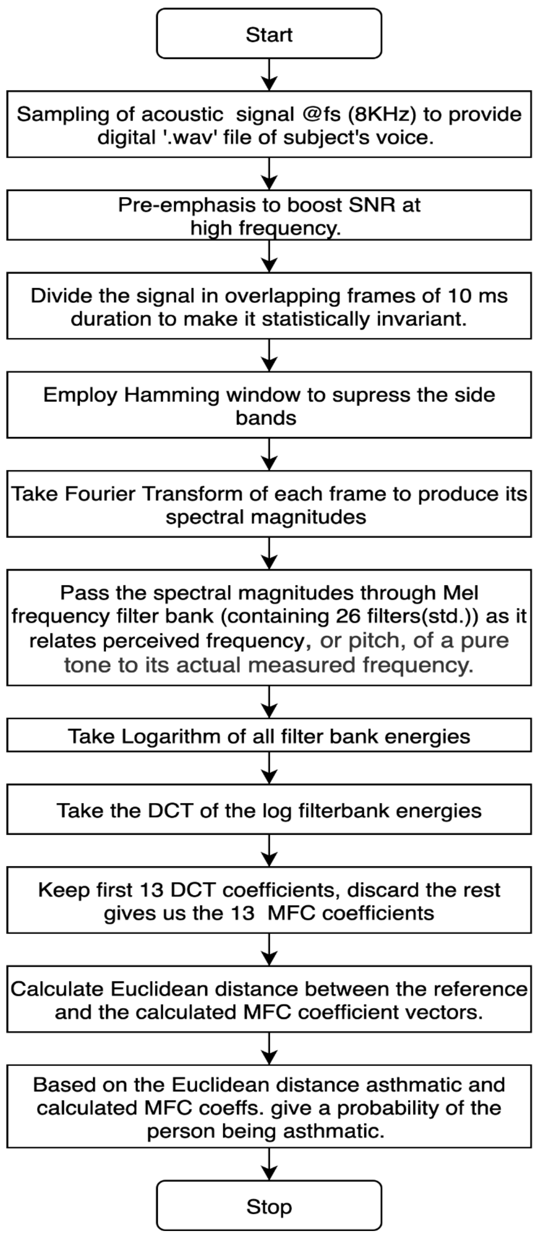
OBJECTIVES

The objective of this work is to study upon and understand the concepts of voice signal processing, apprehend the concepts of MFC Coefficients, implement a code for the proposed algorithm in reference paper and study its application in the field of telemedicine for construction of a low cost automated system for monitoring asthma, based on a mobile device and appropriate application.

.

METHODOLOGY

This flow chart specify the step-by-step process employed in the algorithm coded in MATLAB to obtain the final results. For most of the steps, reason of its presence is also specified.



iJETRM
International Journal of Engineering Technology Research & Management

IMPLEMENTATION MINUTAE

one cannot discern the difference between two closely spaced frequencies and this effect becomes more pronounced as the frequency increases. For this reason, we take clumps of periodogram bins and sum them up to get an idea of how much energy exists in various frequency regions. This is performed by our Mel filterbank: the first filter is very narrow and gives an indication of how much energy exists near 0 Hertz. As the frequencies get higher, our filters get wider as we become less concerned about variations. We are only interested in roughly how much energy occurs at each spot. The Mel scale tells us exactly how to space our filterbanks and how wide to make them. Mel filterbank is a set of 26 triangular filters that we apply to the periodogram power spectral estimate. To get the filterbank we first have to choose a lower and upper frequency. Good values are 300Hz for the lower and 8000Hz for the upper frequency. Given below is a relation between linear and mel scale of frequency.

$$\text{mel}(f) = 2595 * \log\left(1 + \frac{f}{700}\right) \quad \ldots (1)$$

Using equation 1, convert the upper and lower frequencies to mels. In our case, 300Hz is 401.97 mels and 8000Hz is 2840.02 Mels. For 26 triangular filters in the filterbanks, we need 28 points. This means we need 26 additional points spaced linearly between 401.97 and 2840.02. Now using equation 1, convert these back to Hertz to get the peak frequencies for triangular filters. We do this to reduce the dimensionality of our input vector (amplitude spectrum), as well as capture its envelope. Triangular filters are spaced over the Mel scale most of the information is located in lower frequencies.

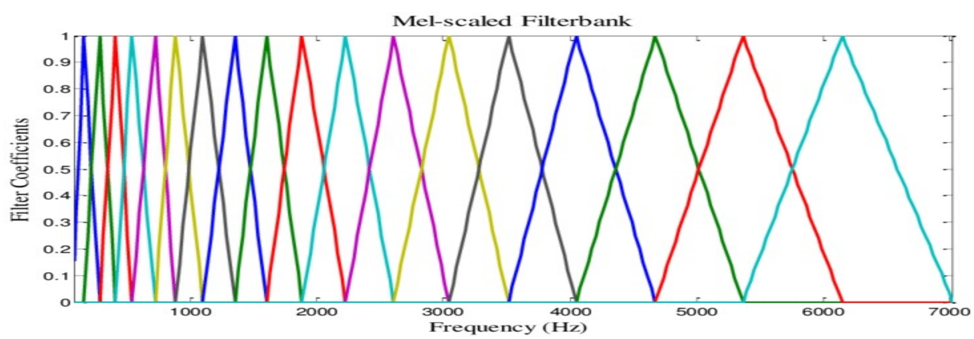


Figure 1: A pictorial representation of mel-scaled filterbank used in the methodology [5].

Wheezing in infants is an important parameter in assessing their predisposition to asthma. Wheezes are associated with airways that are obstructed due to asthma. A wheeze can be described as an unintentional and continuous sound. Acoustically, it is characterised by a periodic waveform with dominant frequency between 100 to 400Hz and duration greater than 100ms. on taking the log of the magnitude of Fourier spectrum, and then again taking the spectrum of this log by a cosine transformation, we observe a peak wherever there is a periodic element in the original time signal. Since we apply a transform on the frequency spectrum itself, the resulting spectrum is neither in the frequency domain nor in the time domain. Hence Bogert et al. decided to .call it the "quefrency" domain. And this spectrum of the log of the spectrum of the time signal was named cepstrum. Cepstrum was first introduced to characterize the seismic echoes resulting due to earthquakes. It serves as a tool to investigate periodic structures within frequency spectra. Cepstrum and occurrence of harmonic frequency are related. There is no standard number of MFCCs for analysing lung sounds. For maximum classification accuracy, an optimal number of 13 is used.

Speech signal is a convolution of source (vocal cord excitation) and filter (articulation cavities). Spectrum is a product of two. We want to get rid of excitation (for us this is a noise). Homomorphic filtering (taking logarithm) is used to linearly separate the components given in product form. Spectrum becomes a sum of components. Applying a Discrete cosine transform on log-abs spectrum, we get cepstrum. A positive coefficient represents a sonorant sound while negative represents a fricative sound. Low-order cepstral coefficients correspond to spectral envelope i.e. represents acoustic filters formed by resonant cavities of vocal tract.

iJETRM
International Journal of Engineering Technology Research & Management

RESULTS AND DISCUSSION

Features were extracted for the voice sample (collected through Audacity software in .wav format) using Mel-frequency Cepstral Coefficients in MATLAB. After analysing all the 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 1st and 2nd coefficients. We designed a reference vector of asthmatic person containing 13 MFCC coefficients with the data made available by WHo. When the algorithm is fed with a voice clip of preferably unvoiced sounds like letter s, f etc., it calculates the MFCC coefficients. It retain first 13 coefficients and discard the rest of them. Further, it calculates the Euclidean distance between the reference and the test vector. More is the Euclidean distance between the reference and test vector less the probability of test person being asthmatic. This probability-of-being-asthmatic function worked remarkably well for test persons given below.

P (an asthmatic person) = 0.9102

P (a healthy person) = 0.0364

This analysis to find probability of a person of being asthmatic can be done on any person of interest by just recording their voice data digitally and feeding it to the algorithm.

ACKNOWLEDGEMENT

We thank all the teaching and technical staff of Telecommunication Engineering department, RV College of Engineering, and our guide Mr. T.P. Mithun, Assistant Professor who provided insight and expertise that greatly assisted the research. We are expressing our gratitude to our families for being an inspiration. Above all, to God.

CONCLUSION

The required Mel frequency Cepstral Coefficients are obtained using MATLAB software. The obtained MFCC are analyzed to give a probability of a person being asthmatic. . 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 first and second coefficients. The proposed method can be applied to the field of telemedicine, a much need of the hour. The proposed project work is good for construction of a low cost automated system for monitoring asthma, based on a mobile device and appropriate application.

REFERENCES

[1] Siddhant C. Joshi and Dr. A.N.Cheeran , MATLAB Based Feature Extraction Using Mel Frequency Cepstrum Coefficients for Automatic Speech Recognition , International Journal of Science, Engineering and Technology Research (IJSETR), Volume 3, Issue 6, June 2014 .

[2] Rachna, Dinesh Singh and Vikas , Feature extraction from asthma patients voice using Mel Frequency Coefficients, IJRET: International Journal of Research in Engineering and Technology, eISSN: 2319-1163 | pISSN: 2321-7308 .

[3] Sirko Molau, Michael Pitz, Ralf Schluter, and Hermann Ney, Computing Mel-frequency cepstral coefficients on the power spectrum, University of Technology, 52056 Aachen, Germany.

[4] Lindasalwa Muda, Mumtaj Begam and I. Elamvazuthi, Voice recognition algorithm using MFCC & DTW techniques, Journal of Computing, Volume 2, Issue 3, March 2010, ISSN 2151-9617, pp. 138-143.

[5] Davis, S. Mermelstein , P. (1980) Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences. In IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 28 No. 4, pp. 357-366

[6] X. Huang, A. Acero , and H. Hon. Spoken Language Processing: A guide to theory, algorithm, and system development. Prentice Hall, 2001.

[7] Prevalence of Bronchial Asthma in Indian Children Ranabir Pal, Sanjay Dahal and Shrayan Pal, Indian J Community Med. 2009 oct; 34(4): 310–316, PMCID: PMC2822191,PMID: 20165624, doi : 10.4103/0970-0218.58389

[8] 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.

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

[10] https://www.nhlbi.nih.gov/health-topics/asthma