VOICE ANALYSIS BASED DISEASES IDENTIFICATION USING SPEECH PROCESSING

Ms.Deepali S. Shrivastava(P.G Student), Prof.Saurabh Shah, Prof.Ketan Bhavsar

Electronics and Communication Engineering Department
BITS edu campus , Varnama Vadodara, India
Affiliated to Gujarat Technological University, Gujarat , India

Abstract-The proper analysis of the human voice has created a vital field of research for its various applications in medical as well as engineering sectors. The analysis of a voice signal basically deals with the extraction of some parameters from voice signal for processing of voice in desirable application by using suitable and various types of algorithms. Out of the many algorithms MFCC is a type of algorithm that gives the best and accurate results. This paper is used to identify one or more diseases in human beings like Parkinson's, asthma, etc. By comparing the voice signals of an infected person and a normal person at a lower bit rate.

Keywords- Parkinson's diseases, MFCC, spectral characteristics.

1. INTRODUCTION

In today's life style vocal problems like Parkinson's diseases, problem in utterances of some specific words are increasing day by day in adults as well as children. Parkinson's disease is a chronic and progressive movement disorder meaning that the symptoms continue and worsen over time. At present, there is no permanent cure of this disease but if it is determined at an early stage it's effect can be reduced to a very low level. Parkinson's involves the malfunction and death of vital nerve cells in the brain called neurons. Thus ,this system focuses on identifying the vocal diseases like Parkinson's , Alzheimer diseases , asthma etc. which can be treated at the early stage by taking the voice samples of the infected person and then comparing them with the normal person's voice.The detection and identification can be done by the help of various algorithms like MFCC, LPC, PLP NEURON CLASSIFEIER etc. The following paper gives the detailed note on how the spectral characteristics like pitch , jitter , shimmer , formants etc. Parameters of voice samples can be extracted from frequency as well as time domain by using one of the appropriate algorithms. The whole system works on MATLAB software. This system tries to compute the result with the highest efficiency, at low bit rate , with highest accuracy with minimum use of filters.

Section 2 consists of the method of extracting features from voice samples by comparing various algorithms and finding out the most appropriate one. Section 3 just gives the proposed work of system Section 4 gives the MATLAB simulation of voice and audio sample and section 5 gives conclusion and section 6 gives the refrences.

2. TYPES OF ALGORITHMS

A) LPC (linear predictive coding)

Linear predictive coding is a tool used mostly in signal processing and speech processing for representing the speech envelope of a digital signal of speech in compressed form.LPC starts with the assumption that the speech signal is produced by a buzzer at the end of the tube. The glottis (the space between the vocal folds) produces the buzz , which is characterized by its intensity and frequency. The vocal tracts forms tube , which is characterized by its resonances , which gives rise to formants.LPC determines the speech signal by estimating the formants, removing their effects from speech signal and estimating the intensity and frequency of the remaining buzz. Thus, LPC synthesizes the speech signal by reversing the procees , it uses the formants to create a filter and run the source through the filter , resulting in speech. But accuracy is not so much considerable as compared to other algorithms.

2)PLP (Perception linear prediction)

PLP discards irrelevant information of the speech and thus improves speech recognition rate.PLP is identical to LPC except that its spectral characteristics have been transformed to match characteristics of human auditory system.PLP has three main aspects: the critical band resolution curves , the equal loudness curve , and the intensity-loudness curve , and the intensity- loudness power law relation , which are known as cubic root. The PLP method is more suitable for human hearing ,in comparison to the classic linear prediction coding. The main difference between PLP and LPC method is that the LP model considers the all – pole transfer function of the vocal tract. Thus this method is also not suitable for extracting features as it needs a type of coversion.The power spectrum of the windowed signal is calculated as:

$$P(w)=Re(S(w))^{2} + Im(S(w))^{2}$$

A frequency warping into the bark scale is applied. The first step is a conversion from frequency to bark , which is a better representation of the human hearing resolution in frequency. The bark frequency corresponding to an audio frequency is,

$$\Omega(\varphi)=6 \ln[\varphi/1200\pi+[(\varphi/1200\pi)^{2}+1]^{0.5}]$$

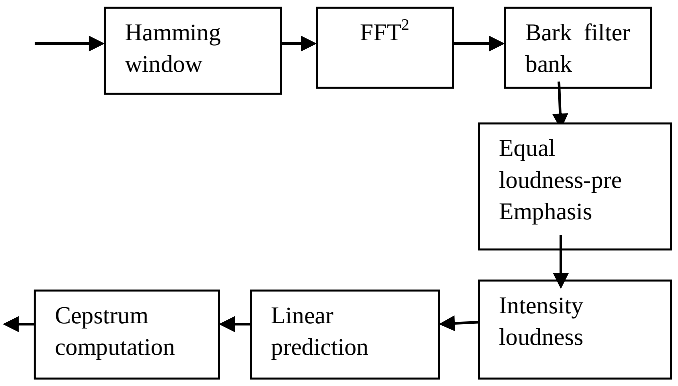


Fig.1"PLP BLOCK DIAGRAM"

3) MFCC (MEL- FREQUENCY CEPSTRAL COEFFICIENT)

This is one of the most accurate methods for extracting features from voice signal. It is more accurate and simple to understand as compared to other algorithms. MFCC is based on the human peripheral auditory system. The human perception of the frequency contents for speech signal does not follow a linear scale , infect it follows a Mel scale , which is there in MFCC algorithm. The following detail will show how MFCC coefficients can be obtained.

STEP 1: FRAME BLOCKING: First of all this block will divide the incoming signal into frames of N samples , in which each frame is of 15-20 ms.

STEP 2:PRE-EMPHASIS: The output of first block is then fed to this block which is used to increase the energy of the voice signal.STEP 3:HAMMING WINDOW: This window gives best result as compared to other windows with suitable no. of filters. It reduces the discontinuities between the continuous signals so as to make the ends smoother by tapering to 0 at ends as well as at the beginnings.

STEP 4:FFT: This block converts the time domain of frames of N samples into frequency domain

STEP 5:MEL- FILTER BANK: This block consists of the combination of 40 mel filters which form the MEL-FILTER BANK. It passes only a particular set of frequencies which corresponds to the samples of each frame. Finally , the output is given to the DCT block.

STEP 6:DCT: DCT is a discrete cosine transform which performs the task of converting the frames of N Samples from frequency domain to again time domain.

STEP 7:OUTPUT MEL-COEFFICENTS: Finally in this step we get the output mel-coefficients.In order to compute mels for a given frequency *f* in Hz , we have to follow the following equation

$$Mel(f)=2595*\log10(1+f/700)$$

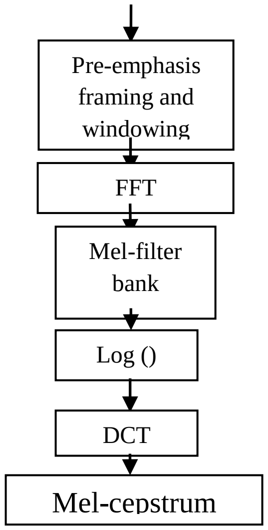


Fig.2 "MFCC BLOCK DIAGRAM"

3. PROPOSED WORK

In voice detection method , the main drawback is lack of efficiency and also the system is not able to identify more than one diseases also cannot identify at early stage. This paper focuses on the matter of increasing the accuracy as well as identification of more than one disease at early stage.

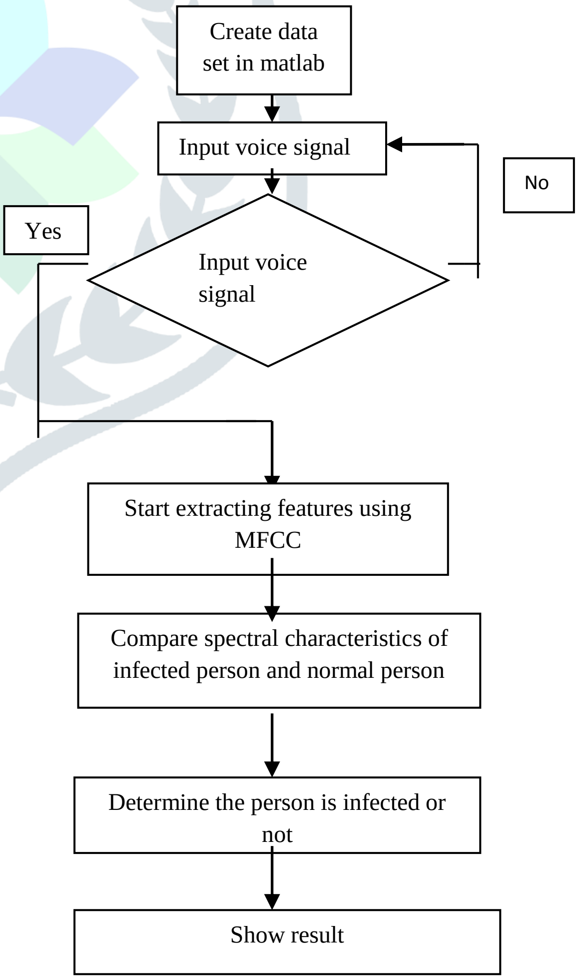


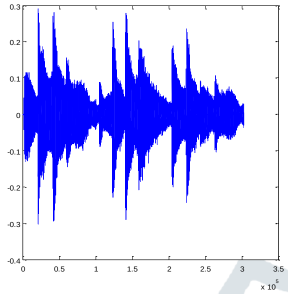
Fig.3 "PROPOSED FLOW CHART"

4. AUDIO SIMULATION

This section consists of just the normal waveform of some audio sample in MATLB just to describe how the waveform appears, when audio sample is taken in MATLAB.

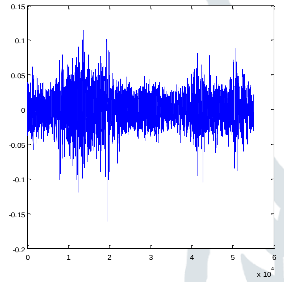
i) Mantle=wavread('C:\SOUND\Clock_mantle.wav');

Plot(Mantle)



ii) d=wavread('C:\SOUND\myvoice.wav');

plot(d)



CONCLUSION

In this paper, it is shown that how a voice signals is identified and the way of recognizing the diseases with more accuracy in future work.

REFRENCES

[1].K.K. Paliwal and B.S. Atal, 'Frequency related representation of speech,' in *Proc. EUROSPEECH*, p.p.65-68 Sep. (2003).

[2] T. Fukuda, M. Takigawa and T. Nitta, "Peripheralfeatures for HMMbase speech recognition," in *Proc. ICASSP*, **1:** 129-132(2001).

[3]. Hartelius L and Svensson P, "Speech and swallowing symptoms associated with Parkinson's disease and multiple sclerosis: a survey," Folia Phoniatr Logop 1994:46(1):9-17.

[4] Goberman A.M. and Mc Millan J, "Relative speech Timing in Parkinson's disease," commun Science Discord 2005,32:22-29

[5] Jiri Mekyska, Irena Rektorova and Zdenek , "Selection of optimal Parameters for Automatic Analysis of speech Disorders in Parkinson's disease".