

Spectral Analysis of Pathological & Normal Speech Signal

Prof. Syed Mohammad Ali, Dr. Pradeep Tulshiram Karule

Abstract— Due to nature of jobs, unhealthy social habits people are subjected to risk of voice problem [9]. If there is some neurological disorder then also there is problem of voice disorder. Therefore, voice signal can be a useful tool to diagnose them. The awkwardness of analog equipments has simulated development of digital computer techniques for processing and analysis of pathological speech signal in patients care system.

In this paper, normal & pathological speech signals are taken & a system is designed to differentiate normal from abnormal signals. These signals are first preprocessed.

Preprocessing techniques involves passing signal through high pass filter, moving average filter (ma), framing & windowing .The windowed signal is given for spectral analysis. In spectral analysis, various methods like logarithmic spectrum, cepstrum, auto correlation of speech signal, spectrogram are applied to differentiate normal & abnormal speech signals. It has been seen that with the above method one can clearly differentiate these signals.

Index Terms— Pathological speech signals; preprocessing; spectral analysis.

1 INTRODUCTION

Speech disorder detection has received great momentum in the last decade. Digital signal processing has become an important tool for voice disorder detection [3]. Pathological voice signal & Normal signals are taken. The pathological speech signals are taken from Govt. Medical College & Hospital, Nagpur & Dr. Naresh Agrawal Hospital, Nagpur. The signals are recorded keeping mic two inch away from the mouth using Voice recorder of Window XP. The sampling frequency is chosen to 11025 samples/sec, 8 bit stereo 21 kb/sec. The patients are told to pronounce vowel 'a', vowel consonant 'ah' & word 'Hello'. Physicians often use invasive techniques like endoscopy to diagnose symptoms of vocal fold disorders however, it is possible to diagnose disease using certain feature of speech signal [3].

The speech signal is noisy & noise needs to be removed. So signal is pre-processed by passing it through preprocessing system. Speech signal is sinusoidal signal having different frequency, different amplitude & different phase. It is given by the expression given below [6].

$$\sum_{i=1}^N A_i(t) \sin[2\pi F_i(t)t + \theta_i(t)] \quad (1)$$

Where, $A_i(t)$, $F_i(t)$ & $\theta_i(t)$ are the sets of amplitudes, frequencies & phases respectively, of the sinusoids & speech production requires close cooperation of numerous organs which from the phonetic point of view may be divided into organ.

1. Lungs, Bronchi, Tracheas (producing expiration air steam necessary for phonation)
2. Larynx (amplifying the initial tone)
3. Root of the tongue, throat, nasal cavity, oral cavity (forming tone quality & speech sound) [7].

The use of non invasive techniques to evaluate the larynx and vocal tract helps the speech specialists to perform accurate diagnose [10]. Speech signal in non intrusive in nature & it has potential for providing quantitative data with reasonable analy-

sis time. So study of speech signal of pathological voice has become an important topic for research as it reduces work load in diagnoses of pathological voices [8].

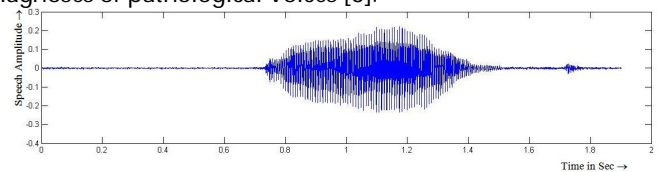


Fig. 1 (a). A speech signal

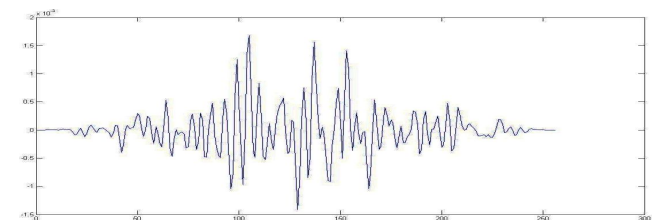


Fig. 1 (b). A processed speech signal

Figure 1 shown above is the speech signal of a neurological disorder patient. The algorithm shown figure 2 below shows the flow of control. Here in this paper we have taken speech samples of neurological disorder and normal persons the speech samples are passed through moving average filter and high pass Filter. The filtered output is framed and then each frame is passed through window. The output signal which is framed and windowed is used for spectral analysis. In spectral analysis logarithmic spectrum of framed window signal is found. Then logarithmic spectrum is used to get cepstrum. Framed signal is also used for finding autocorrelation of speech signal.

2 ALGORITHM FOR SPECTRAL ANALYSIS

The figure 2 is algorithm of spectral analysis

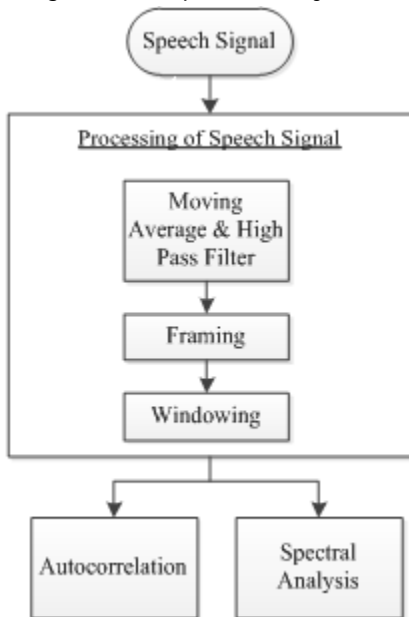


Fig. 2. Flowchart of spectral analysis of speech signal.

3 MOVING AVERAGE FILTER DESIGN

The noisy voice signal is passed through filter. Like moving average filter, preprocessing filter. The moving average filter takes the average of sample for filtering the noise signal. The expression for output of such filter is given below [6].

$$Y(n) = \frac{X(n) + X(n-1) + X(n-2)}{3} \quad (2)$$

Where, $X(n)$ is the input speech sample.

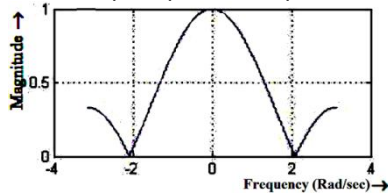


Fig. 3(a). Magnitude v/s Frequency plot of moving average filter

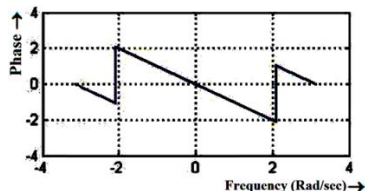


Fig. 3(b). Phase v/s Frequency plot of moving average filter

The figures 3(a) & 3(b) are the magnitude and the phase plot of moving average filter. The cut-off frequency of low pass moving average filter is 2 rad/sec. whenever magnitude changes its sign a phase jump of $+\pi$ radians on right hand side and $-\pi$ radians on left hand side will occur.

3 PREEMPHASIS FILTER FILTER DESIGN

The output of moving average filter is given to pre-emphasis filter, which is high pass filter. This filter is used to flatten the speech signal spectrum & to make the speech signal less sensitive to finite processing effects later in speech signal processing [4]. The pre-emphasis filter amplifies the area of spectrum. Thus improving the efficiency of spectral analysis [2].

The time domain presentation of filter will be

$$Y(n) = X(n) - \lambda X(n-1) \quad (2)$$

Where $y(n)$ is the output, $x(n)$ is input speech sample & λ is the filter coefficient with $\lambda = 0.9375$ optimum result of filtering is received [5]. The output if this filter is framed & passed through window this is done as speech signals are analyzed for short period of time (5 msec to 100msec). The signal is fairly stationary & windowing is done to avoid problem due to truncation of signal & window helps in smoothing of signal [1].

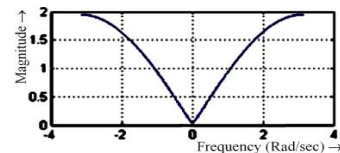


Fig. 4(a). Magnitude v/s Frequency plot of pre-emphasis filter

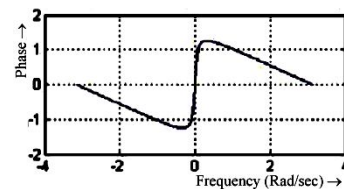


Fig. 4(b). Phase v/s Frequency plot of moving average filter

The figures 4(a) & 4(b) are the magnitude and the phase spectrum of pre-emphasis filter. At the origin there in magnitude plot, the magnitude changes its sign, so phase jump of $+\pi$ radians is clearly seen in the phase plot.

4 CLASSIFICATION OF SPEECH USING CEPSTRUM

The name 'cepstrum' was derived by using the first four letter of spectrum [14]. A reliable way of obtaining an estimate of the dominant fundamental frequency for long clean stationary speech signal is to use cepstrum. The cepstrum is a Fourier analysis of the logarithmic amplitude spectrum of the signal. If the log amplitude spectrum contains many regularly spaced harmonics, then Fourier analysis of the spectrum will show a peak corresponding to the spacing between the harmonics i.e. fundamental frequency. Here signal spectrum is treated as another signal, then looking for periodicity in the spectrum itself. The cepstrum is so called because it turns the spectrum inside out. The X axis of cepstrum has unit of quefrequency & peak in cepstrum is called rahmonics [13].

If $X(n)$ is the speech signal then logarithmic spectrum is given by

$$Y(n) = \text{FFT}[X(n)] \quad (3)$$

$$Y(n) = 20 \times \log_{10}[\text{abs } Y(n)] \quad (4)$$

The cepstrum is DFT of log spectrum

$$Y(n) = \text{FFT}[\log(\text{abs}(Y(n)))] \quad (5)$$

From the cepstrum, one can easily differentiate the difference between the normal and abnormal speech signal. In this paper, we have calculated logarithmic spectrum and cepstrum of patient suffering from neurological disease like schizophrenia, ocd and normal person. Seeing the spectrum, cepstrum one can classify the normal and abnormal speech.

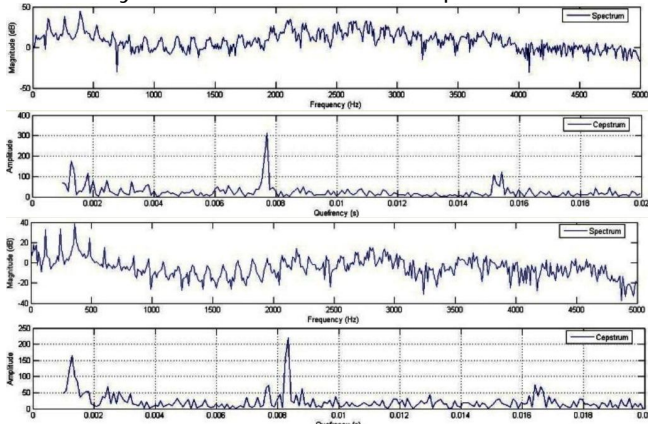


Fig. 5. The above figures show spectrum and cepstrum of normal person

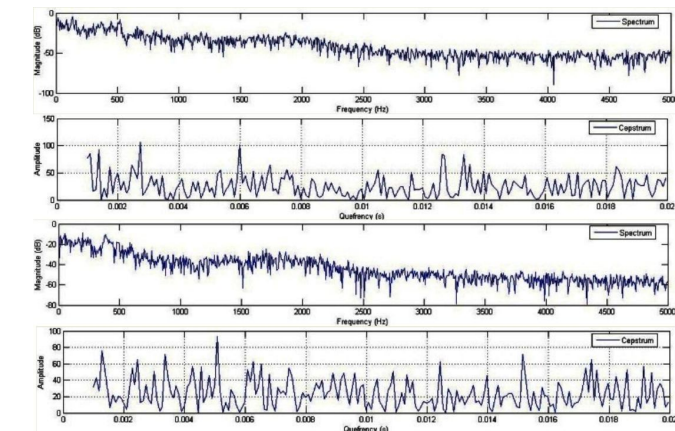


Fig. 6. The above figures show spectrum and cepstrum of neurological disorder

5 AUTO CORRELATION OF PATHOLOGICAL AND NORMAL SIGNAL

One of the other methods, which is applied for the classification of pathological speech signal from normal is Autocorrelation method. Using this method, one can easily classify the normal and abnormal speech signals. The autocorrelation of discrete time signal $X(n)$, is given by [6].

$$r_{xx}(\ell) = \sum_{n=-\infty}^{+\infty} X(n) \cdot X(n - \ell) \quad \ell = 0, \pm 1, \pm 2, \dots$$

The autocorrelation function of a signal is a transformation of signal, which is useful for displaying structure in the waveform [11]. Here it is shown how the autocorrelation function classifies the signals. For the normal signal the decay of auto-

correlation of signal with respect to time is exponential where as for abnormal decay will not be exponential result of which is shown in figure 7 & 8;

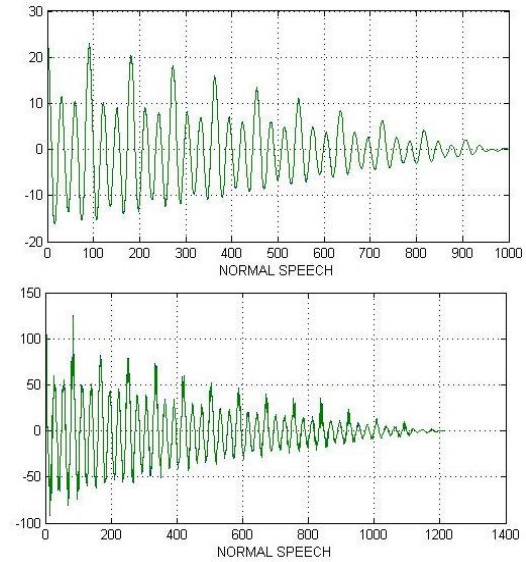


Fig. 7. Autocorrelation of normal speech

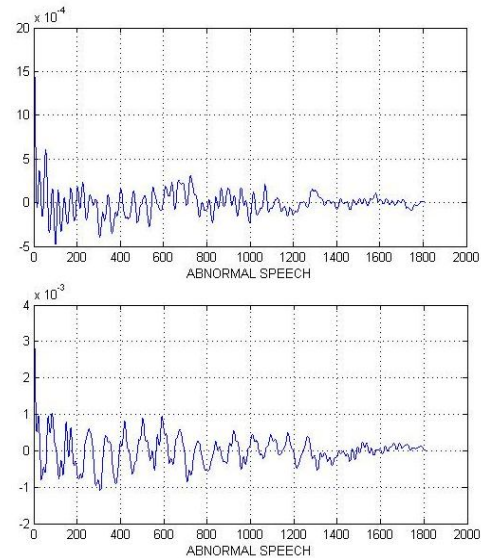


Fig. 8. Autocorrelation of neurological disorder

5 SPECTROGRAM

Among distinct signal processing techniques employed for voice analysis, the spectrogram is commonly used as it allow for visualization of variation of energy, of the signal as function of both time and frequency [15]. The study investigates the use of the global energy of the signal estimated through spectrogram as a tool for discrimination between signals obtained from healthy and pathological subjects.

A spectrogram is display of frequency content of the signal drawn so that energy content in each frequency region and time is displayed on a coloured seal The horizon axis of spectrogram is time and the picture shown how the signal develops and time . The vertical axis of the spectrogram is frequen-

cy it provides an analysis of signal into different frequency regions. You can think of each of these signal as comprising a particular kind of building blocks of signals.

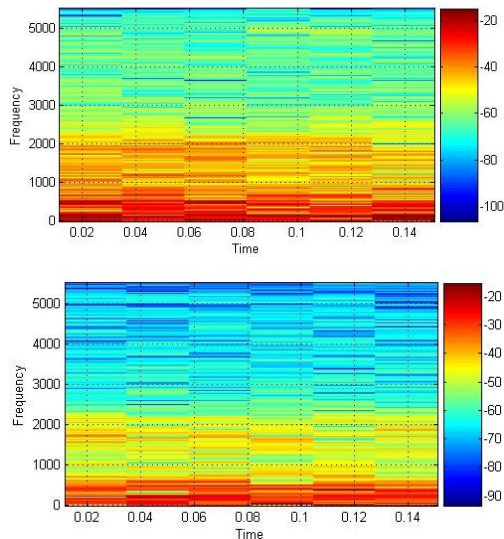


Fig. 9(a). The above figures show spectrogram of normal person

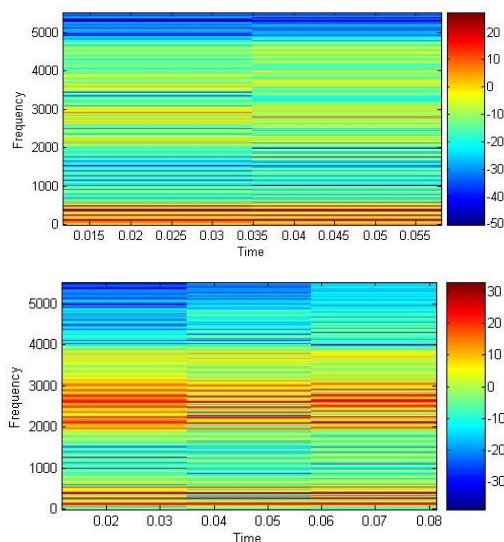


Fig. 9(b). The above figures show spectrogram of neurological disorder

RESULTS & DISCUSSION

In this paper, we have taken speech samples of neurological disorder patient and applied various techniques of classifications like logarithmic spectrum, cepstrum, autocorrelation, & spectrogram. Seeing figure from 5 to 9, one can easily differentiate the abnormal speech samples from normal.

REFERENCES

[1] Manjot Kaur Gill, Vector, "Quantization based Speaker identification", International Journal of Computer Application (0975-8887) Volume 4- No.2, pp1-4 July 2010.
[2] J.W.Picone, "Signal modeling techniques in speech recognition", Proceedings

of the IEEE, Vol.81, No.9, Sept.1993.

[3] Lotfi Salhi, Talbi Mourad, and Adnene Cherif, "Voice Disorders Identification Using Multilayer Neural Network", The International Arab Journal of Information Technology Volume 7-No.2, pp.177-185 April 2010.
[4] Antanas Lipeika, Joana Lipeikiene, Laimutis Telksnys, "Development of Isolated speech Recognition System", INFORMATICA, 2002, Vol 13, No.1, 37-46, 2002.
[5] Milan Sigmund, "Voice Recognition By Computer", Tectum Verlag publication, pp no20-22
[6] John G.Proakis, Dimitris, G.Manolakis, "Digital Signal Processing. principles, Algorithm and Applications", Prentice Hall India, pp.309, Third Edition.
[7] Orzechowski, A Izvorski, R. Tadeusiewicz K. Chmunzynska, P. Radkowski, I Gotkowska, " Processing of pathological change in speech caused by Dysarthria ", IEEE Proceedings of 2005 International symposium on intelligent signal processing & communication system, pp 49-52, Dec 13-16, 2005.
[8] Martinez Cesar E, Refiner Hugo L " Acoustic Analysis of speech for detection of Laryngeal pathologies", IEEE proceeding of the 22nd Annual EMBS international conference chicao IL, pp.2369-2372, 2000.
[9] M.Hariharan, M.P.Paulraj, Sazali, aacob, "Time Domain Features and Probabilistic Neural Network For the Detection Of Vocal Fold Pathology", Malaysian journal Of Computer Science, Vol(23) 2010 pp. 60-67.
[10] Marcelo de Oliveira Rosa, Jose Carlos Pereira and Marcos Gellet, "Adaptive Estimation of Residue Signal for Voice Pathology Diagnose" IEEE Transaction on Biomedical Engineering , Vol.47, No.1, Jan.2000.
[11] Lawrence R. Rabiner, "On the use of Autocorrelation Analysis for Pitch Detection", IEEE Transaction Acoustics, Speech and signal Processing , Vol. ASSP-25, No.1, pp24-30, February 1977.
[12] Estefancy Carrillo , Bathiya Senevirathna, " Automatic Pitch Detection Using Speech segmentation and Autocorrelation
[13] R.W. Schafer and L.R.Rabiner, "System for automatic formant analysis of voiced speech," J.Amer., vol.47, pp 634-648, Feb.1970
[14] <http://en.wikipedia.org/wiki/cepstrum>
[15] M.Fernandes, F.E.R.Mattioli, E.A.Lamounier Jr. and A. O. Andrade, "Assesment of Laryngeal Disorders Through The Global Energy of Speech," IEEE Latin American Transactions, vol.9, No.7, December 2011.