Impact of Technology in Filter Design for Noise Removal from Pathological Noisy Speech Signal & its Preprocessing

Syed Mohammad Ali
Department of Electronics & Telecommunication Engineering, Anjuman College of Engineering & Technology, Nagpur, India
Email: ali_acet2003@yahoo.co.in

Pradeep Tulshiram Karule
Department of Electronics Engineering, Yeshwantrao chavan College of Engineering, Wanadongri, Nagpur, India
Email: ptkarule@rediffmail.com

***Abstract*—In the recent year the trend towards automated analysis of pathological noise signal has gain momentum. The awkwardness of analog equipment has simulated development of digital computer techniques for processing and analysis of pathological speech signal in patient care system. The above filter design techniques & prepossessing of speech signal can be used in any speech processing application.**

This paper discusses pathological speech signal of patients and their preprocessing. In prepossessing, Speech signal is passed through Moving Average (M.A) filter[1], High pass (H.P) filter for removal of noise. The output of filter is framed & these frames are passed through window. Typically, hamming window is used. This preprocessed output can be used for pathological voice recognition, speech identification, speaker identification & many more application.

***Index Terms*—Speech signal, moving average filter, highpass filter, framing, windowing**

I. Introduction

A computerized technique for pathological voice recognition has received a greater attention from researchers in the last decade. Speech processing has proved to be excellent tool for speech disorder detection [1]. Pathological voice signal of patient from Dr. Naresh Agrawal Hospital, ENT Surgeon & Government Medical College & Hospital, Nagpur has been taken. The signals are recorded keeping mic two inch away from mouth using voice recorder of window XP. The sampling frequency is chosen to 11025 samples /sec. The patient has pronounced, 'a' which is vowel, then 'ah' which is vowel with consonant and a word 'Hello' for two sec, these signals are noisy & noise needs to be removed. So filters are designed. The signal is passed through filter & then framing & windowing is done. The output of window is called prepossessed output, which can be used for further application like diagnosis of disease, speaker recognition, & speech recognition. 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 [1]. Speech signal is sinusoidal signal having different frequency, different amplitude & different phase. It is given by the expression given below [2].

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

where, *Ai (t), Fi (t)* & *θi (t)* are the sets of amplitudes, frequencies & phases respectively, of the sinusoids as shown in Fig. 1.

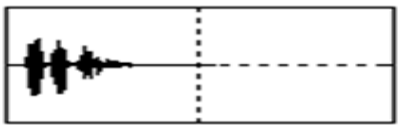


Figure 1. A speech signal

Voice & speech production requires close cooperation of numerous organs which from the phonetic point of view may be divided into organ.

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

Speech signal in non-intrusive in nature & it has potential for providing quantitative data with reasonable analysis 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 [4].

[1] Averaging of three samples gives best result.

Manuscript received December 31, 2012; revised January 25, 2013; accepted February 15, 2013.

II. BLOCK DIAGRAM

The figure below shows block diagram of preprocessing of speech signal. The above three blocks are must for better preprocessing which will give noise free, framed & windowed signal. Moving average filter removes the low frequency noise signal and high frequency noise is removed by pre-emphasis filter.

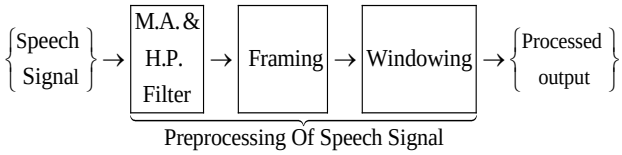


Figure 2. Block diagram of preprocessing of speech signal.

III. FILTER DESIGN

The noisy speech signal is passed through filters like moving average filter, high pass filter, which is also called as pre-emphasis filter. The moving average filter takes average of samples for filtering the noise from signal. The expression for output of such filter is given below [5].

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

where, *X(n)* is the input speech sample.

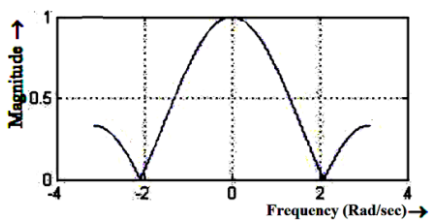


Figure 3. Magnitude v/s Frequency plot of moving average filter

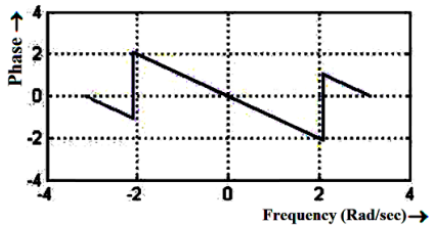


Figure 4. Phase v/s Frequency plot of moving average filter

Analyses with different averages of samples are taken. It is found that with average of three samples, proper filtering is done. The system function *H(z)* & the impulse response *h(n)* of filter are given below:

$$Y(z) = \frac{1}{3}[X(z) + z^{-1}X(z) + Z^{-2}X(z)] \tag{3}$$

$$H(z) = \frac{Y(z)}{X(z)} = \frac{1}{3}[1 + z^{-1} + Z^{-2}] \tag{4}$$

$$\therefore\ h(n) = [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}] \tag{5}$$

This FIR filters shown by magnitude & phase plots in Fig. 3, Fig. 4, Fig. 5 & Fig. 6 are more stable.

IV. HIGH PASS FILTER

Pre-emphasis filter, which is a high pass filter is used to flatten speech signal spectrum & to make the speech signal less sensitive to finite precision effects later in speech signal processing [6].

Here we have used single coefficient FIR filter. The system function *H(z)* of filter is

$$H(z) = 1 - \lambda z^{-1} \tag{6}$$

$$\frac{Y(z)}{X(z)} = 1 - \lambda z^{-1} \tag{7}$$

$$Y(z) = X(z) - \lambda z^{-1} X(z) \tag{8}$$

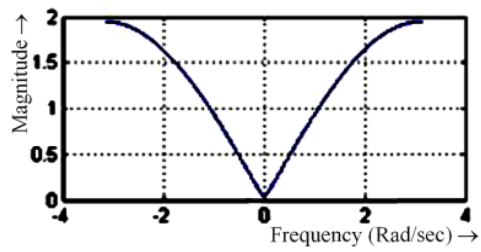


Figure 5. Magnitude v/s Frequency plot of Pre-emphasis Filter

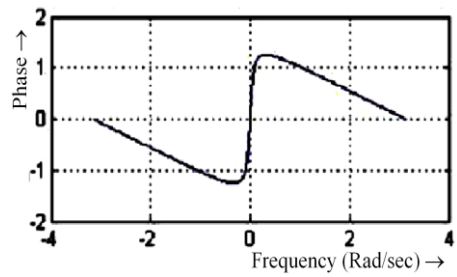


Figure 6. Phase v/s Frequency plot of Pre-emphasis Filter

The time domain representation of filter will be

$$Y(n) = X(n) - \lambda X(n-1) \tag{9}$$

Where, λ is the filter coefficient and the value of pre-emphasis coefficient is $\lambda\ \varepsilon(0.9 - 1.0)$ with $\lambda = 0.9375$, best optimum result of filtering is received [7]. The pre-emphasis filter serves to offset this natural slope before spectral analysis, thereby improving the efficiency of the analysis secondly the hearing is more sensitive above 1 KHz region of spectrum. The pre-emphasis filter amplifies the area of spectrum. Thus improving the efficiency of spectral analysis [8].

V. FRAMING & WINDOWING

Human speech signal is slowly varying over time (quassi stationary) that is when the signal is examined over a short period of time (5 msec to 100 msec). The signal is fairly stationary.

Therefore speech signals are often analyzed in short time segment which are referred as short term spectral analysis. This practically means the signal is blocked in frames of typically 20-30 msec. Adjacent frames typically overlap each other with 30% to 50 %. This is done in order not to lose any information due to the windowing. Also Human speech signal is slowly time varying & can be treated as stationary process when considered under a short time frame. Therefore, the speech signal is usually separated into small duration blocks called frames. The neighbourhood blocks are overlapped by 1/2 to 2/3 length of the frame & frame shift is frame length minus the frame overlap. The commonly used frame length & frame shifts are 20 to 30 msec & 10 msec respectively. After the signal has been framed each frame is multiplied by window function W(n) with length N. The frame has N= 256 samples & adjacent frame separated by M=128 samples. Where *N* is the length of the frame. Typically hamming window W(n) is used & given by

$$W(n) = 0.5 - 0.46\cos(2\pi n / N - 1) \quad 0 \le n \le N-1$$

Where, N is total number of samples. The windowing is done to avoid problems due to truncation of signal as window helps in smoothing of signal [9].

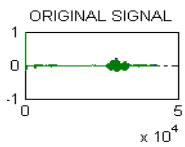
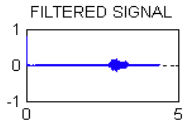
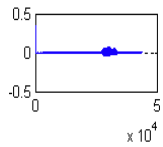
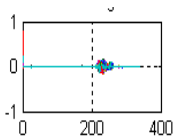
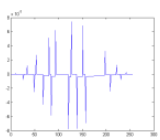
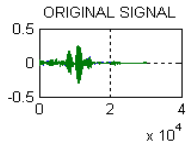
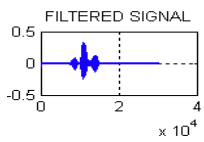
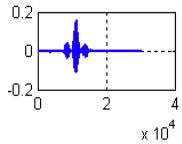
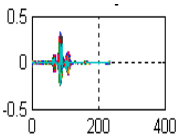
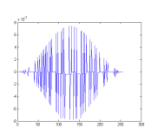
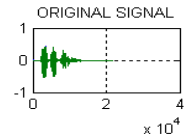
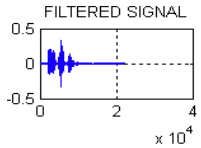
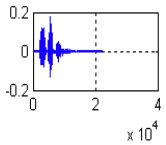
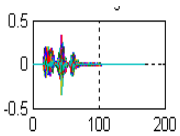
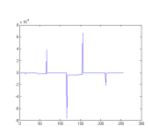
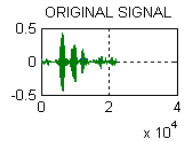
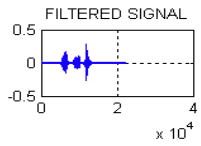
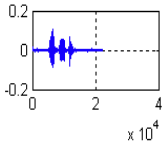
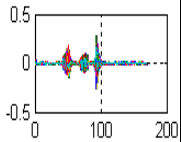
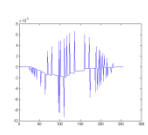
Sr. No.	Patient Samples	Disease	HP Filter Output	MA Filter Output	Framing	Windowing
1		Squamous Cell Carcinoma				
2		Adenoid				
3		Oral Fibrosis				
4		Pharyngeal Fibrosis				

Figure 7. Examples showing complete preprocessing analysis of patient's speech samples

VI. Result & Conclusion

The table on the last page (fig. 7) shows that, signals are perfectly filtered using both ma and pre-emphasis filter. The output of ma filter needs to be scaled to get amplified output. From the table it is clear that this system can be used for any speech preprocessing applications. Filtered signal is framed and passed through window.

References

[1] L. Salhi, T. Mourad, and A. Cherif (April 2010). Voice disorders identification using multilayer neural network. *The International Arab Journal of Information Technology* [Online]. 7(2), pp. 177-185. Available: www.ccis2k.org/iajit/PDF/vol.7,no.2/742.pdf

[2] J. G. Proakis, D. G. Manolakis, *Digital Signal Processing. principles, Algorithm and Applications*, Third Edition, Prentice Hall India, ch. 4, pp. 309.

[3] O. A. Izworski, R. Tadeusiewiez, 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.

[4] M. CesarE, R. HugoL, "Acoustic analysis of speech for detection of laryngeal pathologies," *IEEE proceeding of the 22nd Annual EMBS international conference Chicago IL*, 2000, pp. 2369-2372.

[5] J. G. Proakis, D. G. Manolakis, *Digital Signal Processing principles, Algorithm and Applications*, Third Edition Prentice Hall India, pp. 309.

[6] A. Lipeika, J. Lipeikiene, L. Telksnys (Month 2002). Development of isolated speech recognition system. *INFORMATICA* [Online], 13(1), pp. 3-46. Available: http://www.ijser.org/researchpaper%5CSpectral-Analysis-of-Pathological-Normal-Speech-Signal.pdf

[7] M. Sigmund, *Voice Recognition By Computer*, Tectum Verlag publication, pp. 20-22

[8] J. W. Picone, "Signal modeling techniques in speech recognition," *Proceedings of the IEEE*, Sept. 1993, vol. 81, no.9.

[9] M. Kaur, G. Vector, "Quantization based Speaker identification," *International Journal of Computer Application* (0975-8887) vol. 4, no. 2, pp. 1-4, July 2010.

Prof. Syed Mohammad Ali working as Associate Professor in Anjuman College of engineering and Technology. Graduated from Nagpur University in 1991. M. Tech in Electronics from RTM NU, Pursuing Ph.D from RTM NU. Published four International and one National Paper in the Journal. He has 21 years of teaching experience. He is expertise in digital signal processing. He worked as officiating principal in A.C.E.T. Presently working as Dean Academics and Head of Deptt. Of Electronics Engg. A.C.E.T

Dr. P. T. Karule Professor in Department of Electronics Engineering of YCCE and also looking after administrative work of Registrar. Worked as Head Department of ECE, YCCE, also Dean Industry Institute Interaction, YCCE. He has 23 years of teaching experience. He is expertized in microprocessors and embedded system. He Published 25 research papers in various international journals & international and national conferences. B. E. (Electronics and Power) from Govt. College of Engineering, Amravati. in 1986, M.Tech. (Electronics Engineering) from VNIT, Nagpur in 1992 Ph.D. in the area of Medical Image Processing & Neural Network. from Amravati University in 2010.