Volume 4, Issue 8, August – 2019 International Journal of Innovative Science and Research Technology
ISSN No:-2456-2165

Double-Cosine Term Generalised Adjustable Window for Enhancing Speech Signal Denoising

Mbachu, C. B[1], Akaneme S. A.[2]
1, 2 Department of Electrical and Electronic Engineering, Chukwuemeka Odumegwu Ojukwu University, Uli, Anambra State, Nigeria.

Abstract:- Window technique is widely used in speech processing to enhance the quality of speech by using it in the design of FIR filters for removing or reducing noise components associated with speech signal propagation or transmission. Such noise components include Additive White Gaussian Noise (AWGN), Random Noise, power line noise, high and low frequency noise components. The information content of any contaminated speech signal is unreliable unless these noise components are removed. This paper presents a FIR filter designed with double-cosine term generalised adjustable window with to denoise speech signal of high frequency noise components. The speech signal is recorded as windows media audio (.wma) format and as such the optimal sampling frequency is 44100Hz while the filter order is 34. A real voice statement, "Education is the Key to the Development of any Nation" is transduced to electrical form with the in-built microphone of a laptop system, recorded in windows media audio (.wma) format and stored in a file in the system. With "audioread" instruction the stored voice is loaded into a matlab work space. A noise component of 4500Hz and above is generated with matlab and mixed with the speech to obtain a contaminated speech signal. The designed filter is able to effectively denoise the contaminated speech signal when it is applied to it.

Keywords:- *Adjustable Window, Double-Cosine Term, Speech Signal, High Frequency Noise, Power Spectral Density.*

I. INTRODUCTION

Some researchers have used double-cosine term generalized adjustable window function to denoise speech signals. A generalised adjustable window is a window that assumes the appearance of some other defined and non-defined windows as the adjustment parameter is varied. In [1] Rajput and Bhadauria used a double-cosine term generalised adjustable window function as shown in (1) with adjustment parameter of **α** = 0.07 to design low pass filter for denoising speech signals of high frequency noise components. The sampling frequency is 8000Hz, filter order, 31 and cutoff frequency, 1200Hz. The result shows that the windows is able to significantly denoise the speech signal of dot wave (.wav) format of high frequency noise components above 1200Hz. There is no indication that any researcher has used this window for denoising speech signals of windows media audio (.wma) format, which is an audio signal of double column vector and recorded at a sampling frequency of 44100Hz. Therefore in this work, a double-cosine term generalised adjustable window will be used to design an FIR filter for denoising speech signal of windows media audio (.wma) format.

II. DOUBLE-COSINE TERM GENERALISED ADJUSTABLE WINDOW

A generalised adjustable window as said above is type of window that can take different forms depending on the value of the adjustment parameter. At certain values of the parameter it takes the appearances of known windows and at other values it represents windows of no known name or appearance. A double-cosine term generalised adjustable window has two cosine terms inclusive in the function as shown in (1) [1] below in which case the adjustment parameter is **α**. When **α=0.0** the window becomes hanning window [1, 2, 3, 4] which is a fixed known window as in (2), and a Blackman window [5, 6, 7 8, 9], another fixed known window as shown in (3) when **α=0.16.** At any other value of **α,** the window takes the appearance of other windows.

$$w(n)=\frac{1-\alpha}{2}-0.5\cos\left(\frac{2\pi n}{M-1}\right)+\frac{\alpha}{2}\cos\left(\frac{4\pi n}{M-1}\right), 0 \le n \le M-1 \quad (1)$$

$$w(n) = 0.5-0.5\cos\left(\frac{2\pi n}{M-1}\right), 0 \le n \le M-1 \quad (2)$$

$$w(n)=0.42-0.5\cos\left(\frac{2\pi n}{M-1}\right)+0.08\cos\left(\frac{4\pi n}{M-1}\right), 0 \le n \le M-1 \quad (3)$$

III. Low Pass FIR Filter Design Using Double Cosine Term Generalised Adjustable Window

In this design Double-cosine term generalised adjustable window function is used to weight FIR filter that has a length of 35, that is order of 34, cutoff frequency of 3200Hz and sampling frequency of 44100Hz. Six different values of height adjustment parameter **α=(0.0**, **0.07, 0.16, 0.2, 0.3 and 0.4**) are considered and in each value the impulse, magnitude and phase responses of the filter are obtained and are depicted below. When **α=0.0** the window takes the form of a hanning window as shown in fig.1a, and when **α=0.16** it changes to Blackman window as presented in fig.1c while at **α=0.07** it becomes a different window as in fig.1b. The sampling frequency value chosen is because the speech signal in this circumstance is recorded in windows media audio (.wma) format.

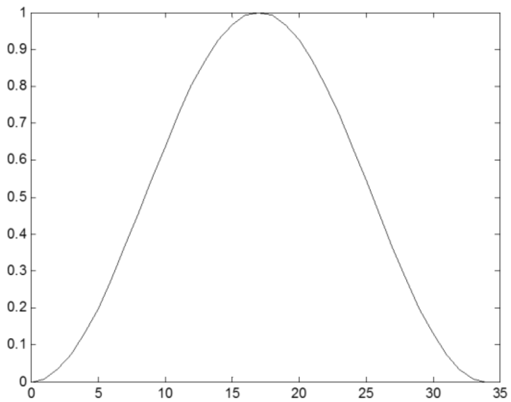

Fig. 1a: Double-cosine Term Generalised Adjustable Window When **α=0.0**

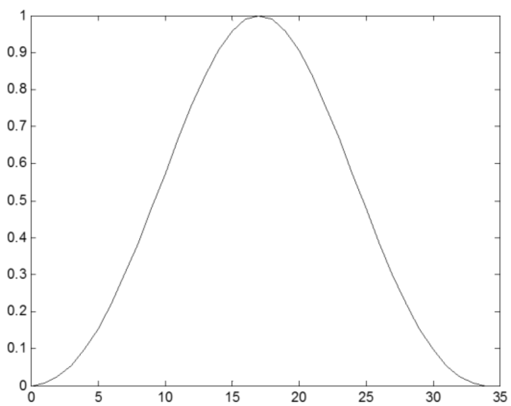

Fig. 1b: Double-cosine Term Generalised Adjustable Window When **α=0.07**

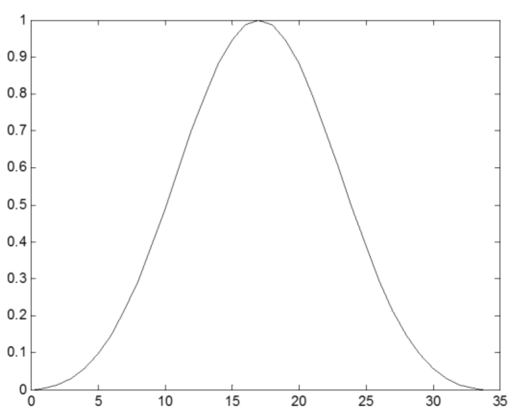

Fig. 1c: Double-cosine Term Generalised Adjustable Window When **α=0.16**

A. Responses When α =0.0

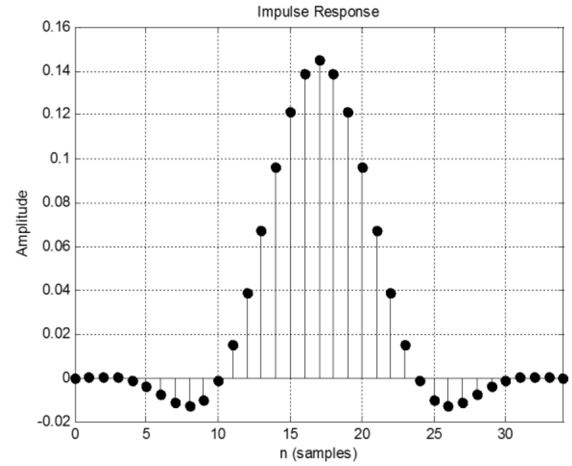

Fig 2a: Impulse Response When **α**=0.0

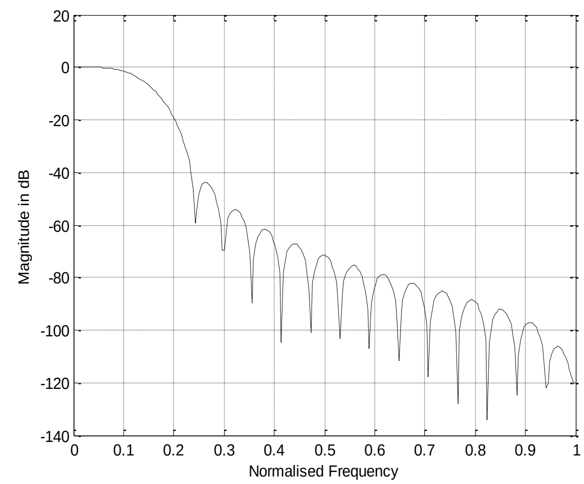

Fig 2b: Magnitude Response When **α** =0.0

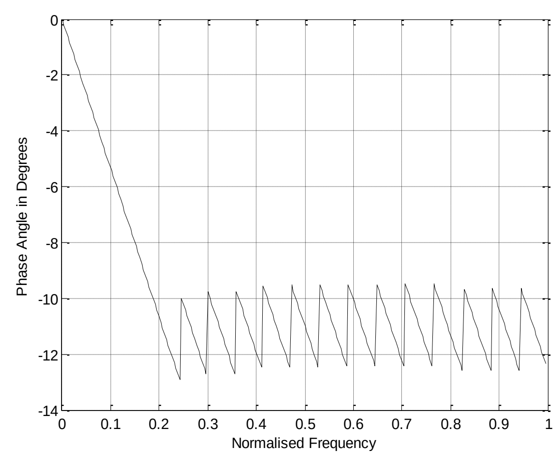

Fig 2c: Phase Response When **α** =0.0

ISSN No:-2456-2165

B. *Responses When α=0.07*

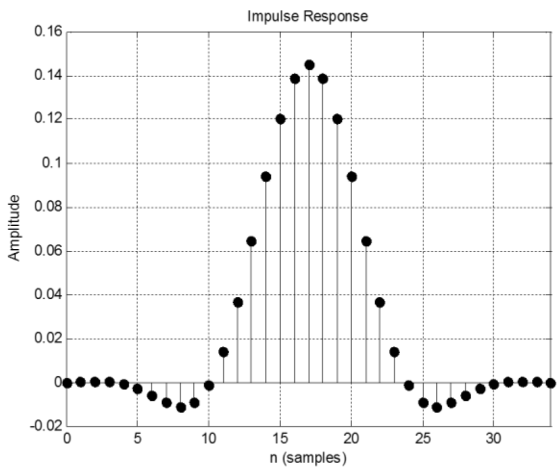

Fig 2d: Impulse Response When **α=0.07**

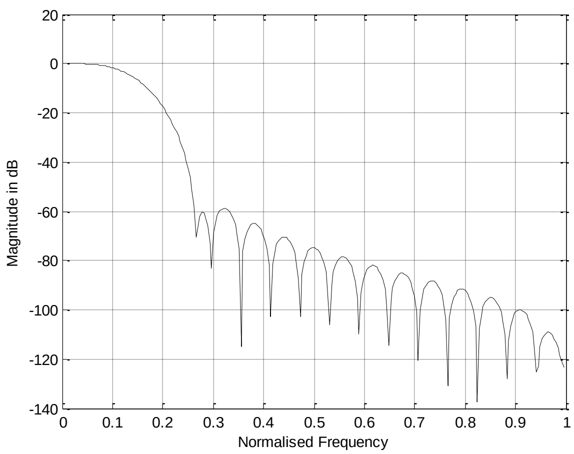

Fig 2e: Magnitude Response When **α=0.07**

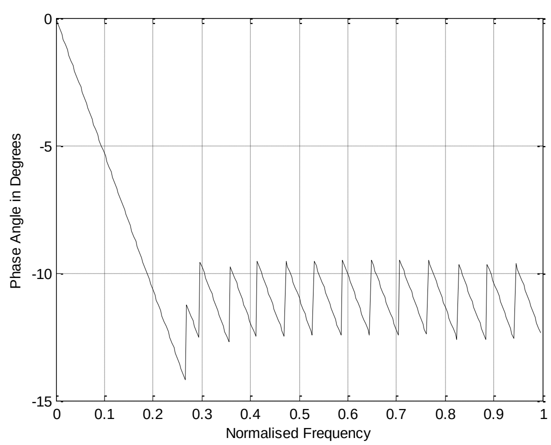

Fig 2f: Phase Response When **α**=0.07

C. *Responses When α=0.16*

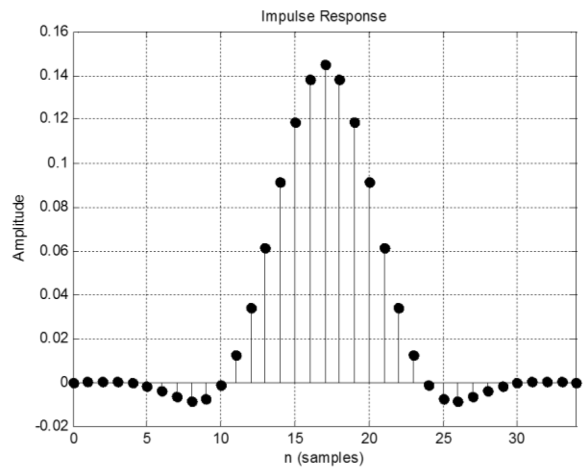

Fig 2g: Impulse Response When **α**=0.16

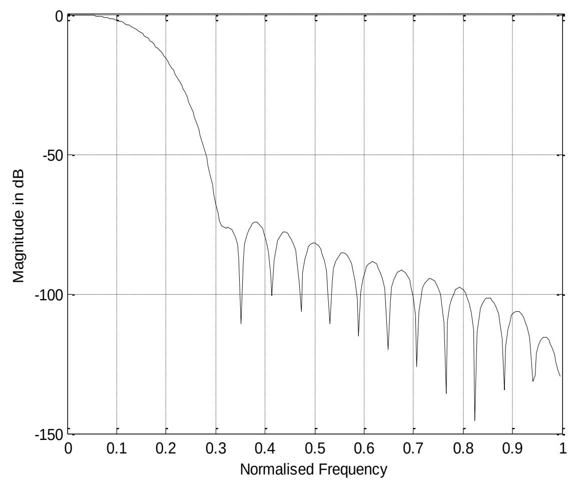

Fig 2h: Magnitude Response When **α**=0.16

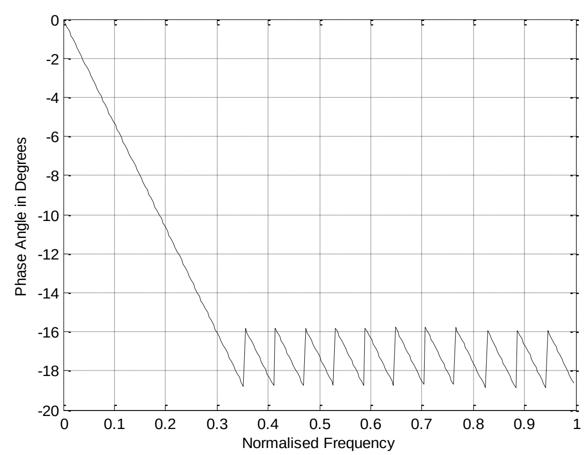

Fig 2i: Phase Response When **α**=0.16

ISSN No:-2456-2165

D. Responses When α=0.2

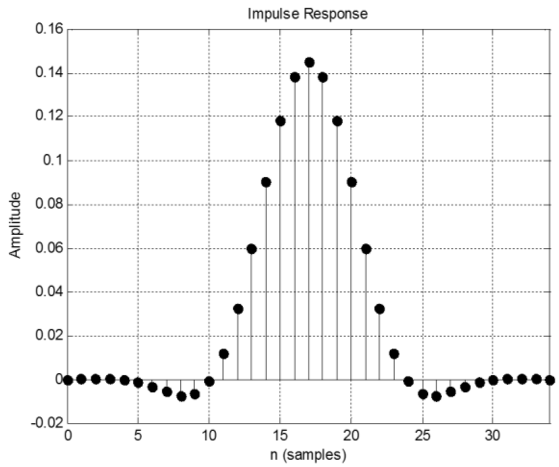

Fig 2j: Impulse Response When **α**=0.2

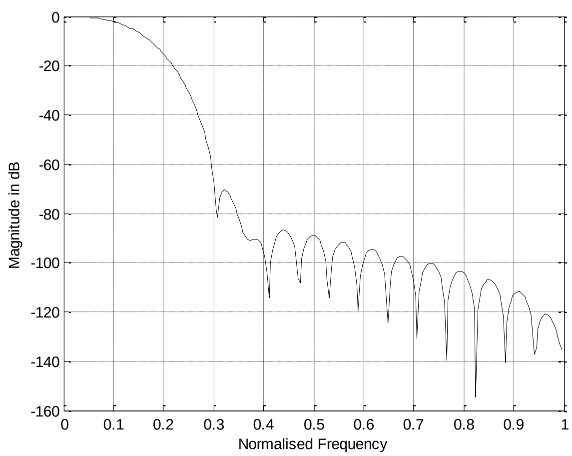

Fig 2k: Magnitude Response When **α**=0.2

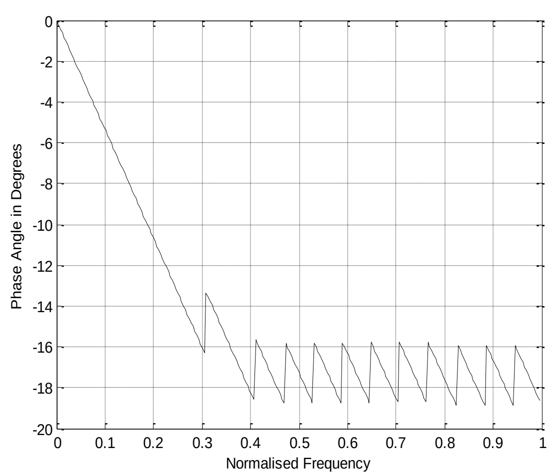

Fig 2l: Phase Response When **α**=0.2

E. Responses When α=0.3

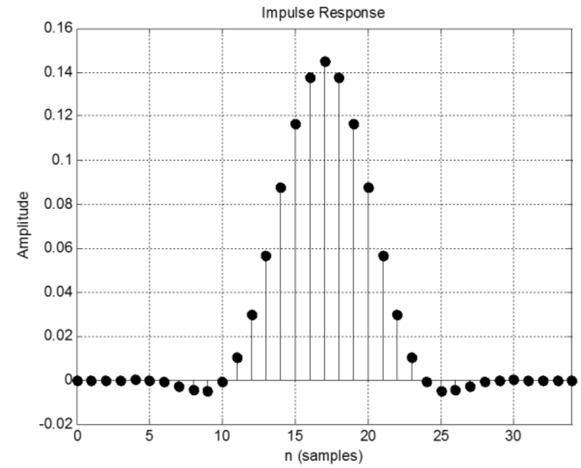

Fig 2m: Impulse Response When **α**=0.3

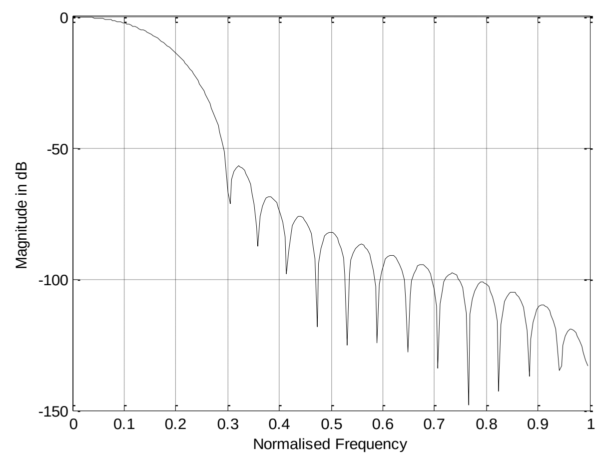

Fig 2n: Magnitude Response When **α**=0.3

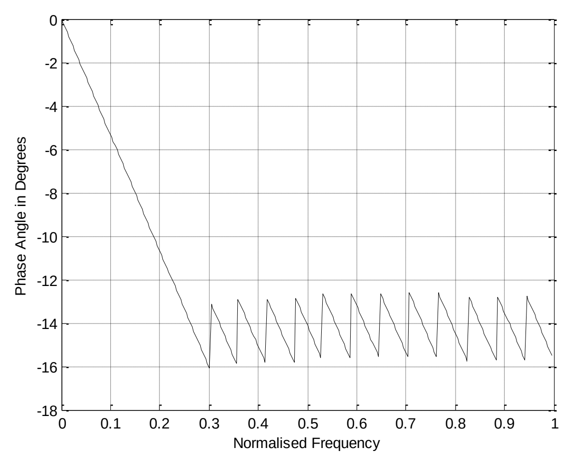

Fig 2p: Phase Response When 0.3

F. Responses When α=0.4

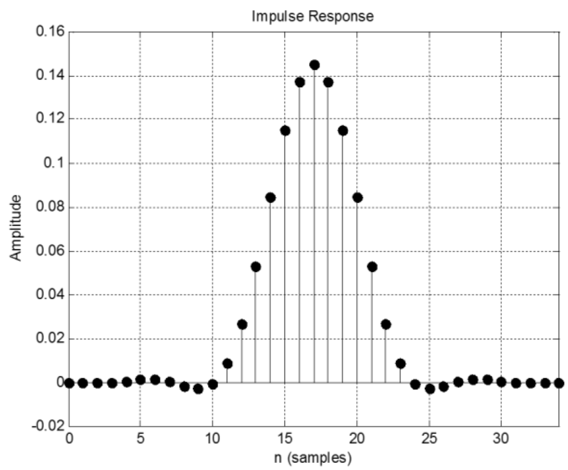

Fig 2q: Impulse Response When **α**=0.4

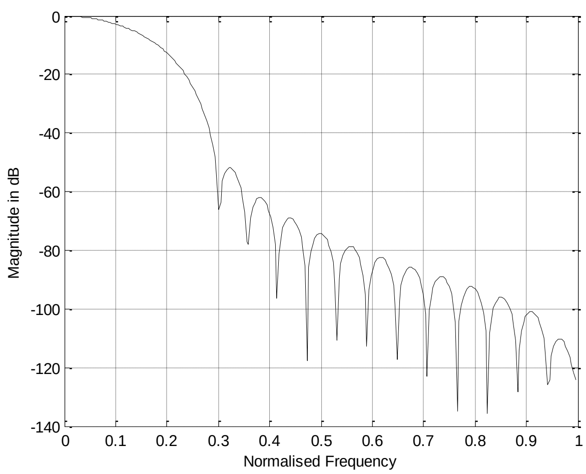

Fig 2r: Magnitude Response When **α**=0.4

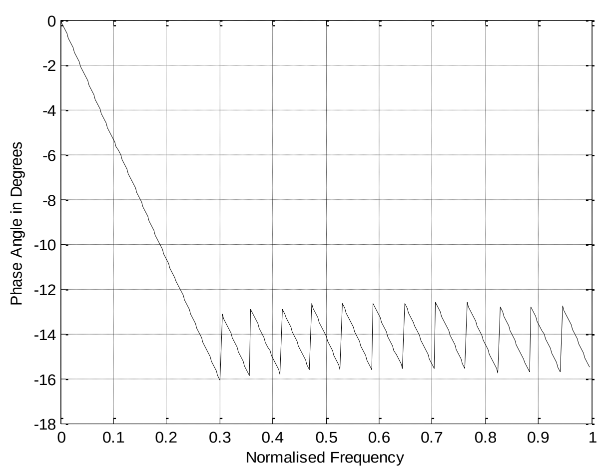

Fig 2s: Phase Response When **α**=0.4

From the responses above, all the impulse responses exhibit stability because they collapse to zero as they approach the last impulse response sample values. The magnitude responses also exhibit stability in that there are no sustained oscillations, but the main lobes of the magnitude responses for values of **α=0.16, 0.2** and **0.3** show higher attenuation values than others which imply that one of them will be the optimum value. The precise optimum value will be obtained from power spectral densities of the filtered signals. The phase responses exhibit clear linearity within the required frequency bandwidth.

IV. RESULTS

Results are obtained by converting a real voice statement "Education is the Key to the Development of any Nation" to electrical speech signal using the system in-built microphone, recorded in windows media audio (.wma) format and stored in one of the files of the system. The signal is transferred to a matlab workspace using "audioread" instruction. In the workspace a noise component of 4500Hz and above is added to the speech to constitute contaminated speech signal. The corrupt-free speech signal is shown in fig.3 while fig.4 depicts the noise component, and the contaminated speech signal depicted in fig.5. The contaminated speech signal is denoised with each of the designed low pass filters and the outputs recorded. Figures 6, 7, 8, 9, 10 and 11 show the speech signal after denoising. Comparing the clean speech signal of fig.3, the contaminated speech signal of fig.5 and the filtered speech signals of fig6 to fig.11 it can be seen that the filter denoised the speech signal for each value of **α**. The optimum value of **α** in this circumstance can be determined from power spectral analysis of the filtered signals from fig.12 to fig.19 below.

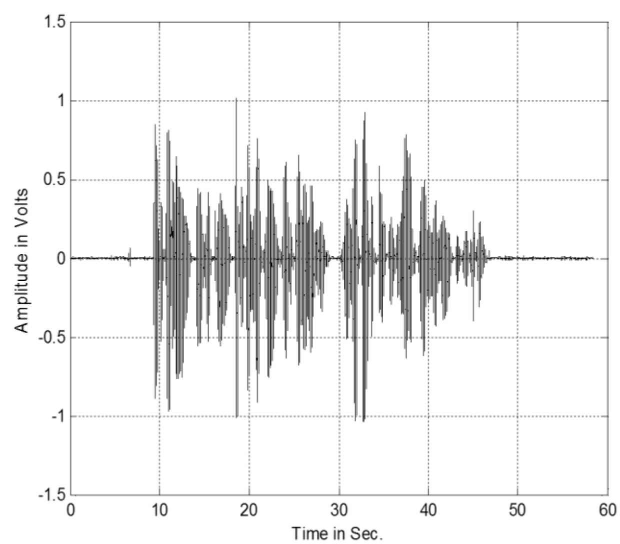

Fig. 3: Noise Free Voice Signal

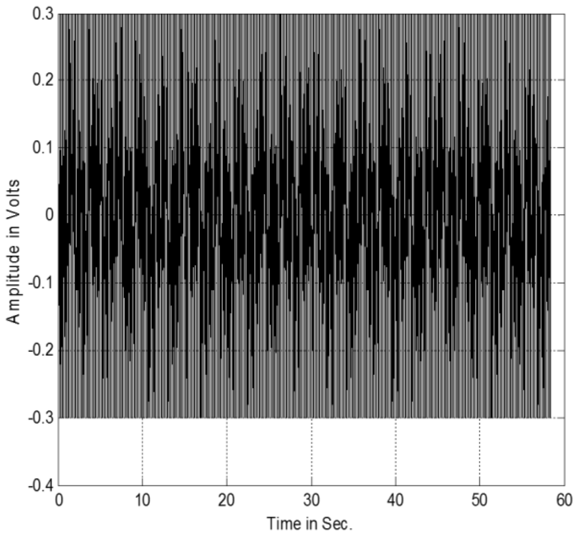

Fig. 4: Noise Signal

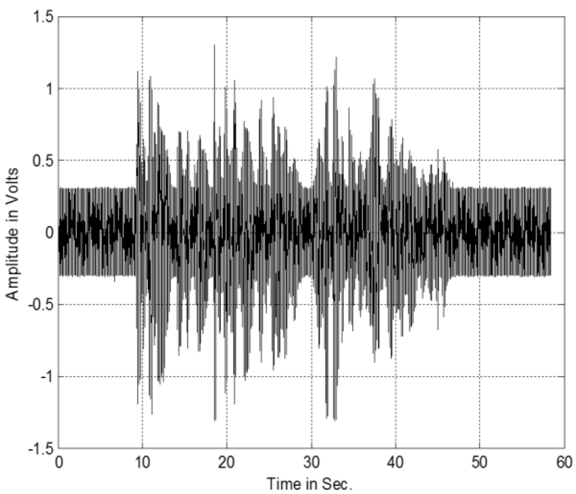

Fig. 5: Contaminated Speech signal

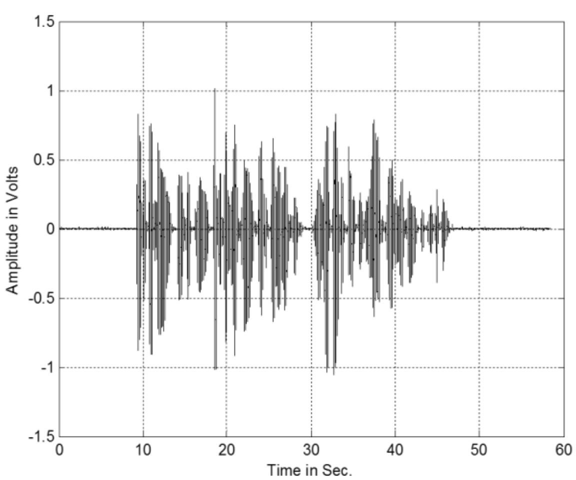

Fig.6: Voice Signal Filtered With Low Pass Filter When α=0.0

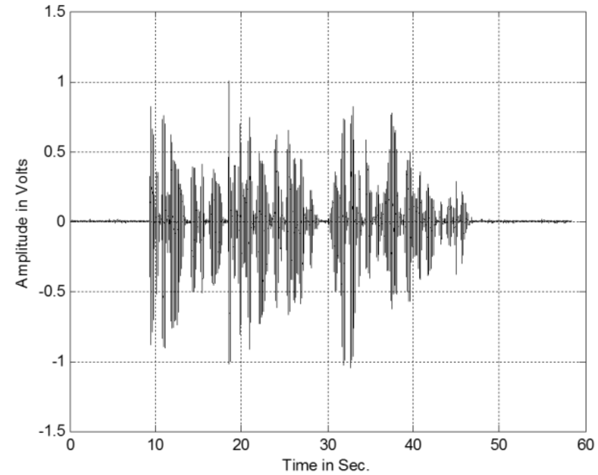

Fig.7: Speech Signal Filtered With Low Pass Filter When α=0.07

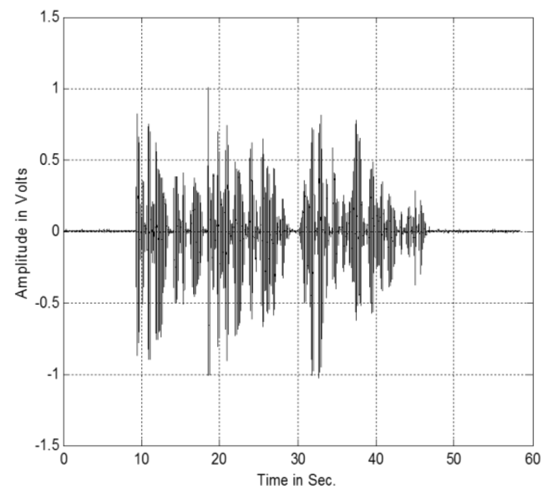

Fig.8: Speech Signal Filtered With Low Pass Filter When α=0.16

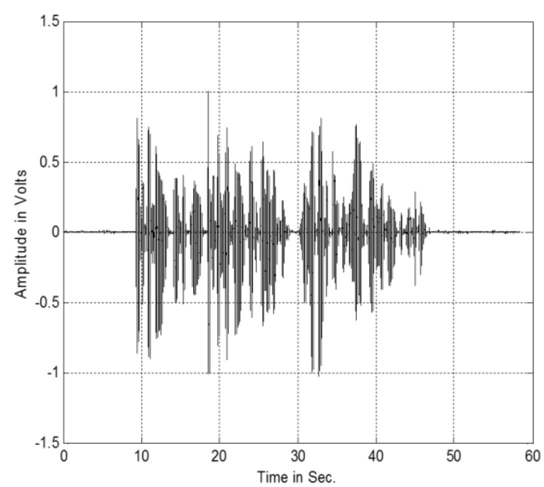

Fig.9: Speech Signal Filtered With Low Pass Filter When α=0.2

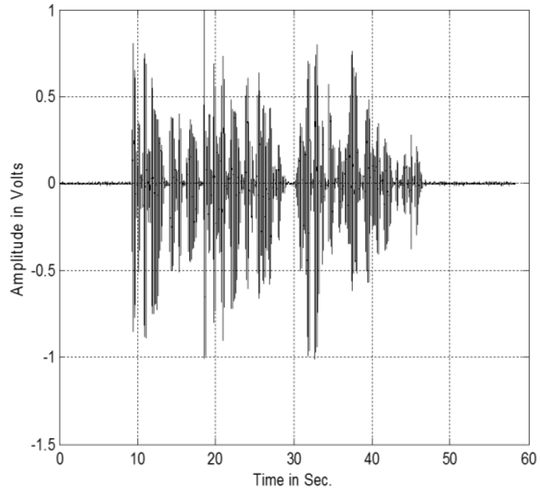

Fig.10: Speech Signal Filtered With Low Pass Filter When α=0.3

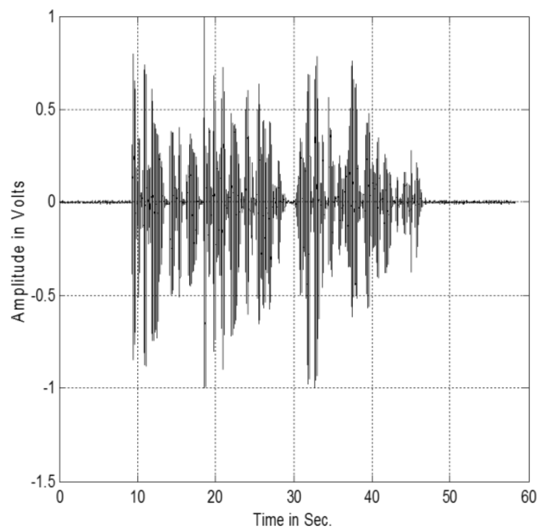

Fig.11: Speech Signal Filtered With Low Pass Filter When α=0.4

➢ *Signal Power Level*

The performance analysis of the filters can be carried out by considering the power levels of the filtered signals [10, 11]. Fig. 12 is the spectral density of a noise-free speech signal while the spectral density of the contaminated speech signal is depicted in fig. 13. Fig.14 to fig.19 is the power spectral densities of the filtered voice signal as α changes value. From fig. 13 the noise component in the contaminated speech signal is most at normalised frequency of 0.875. Therefore it is meaningful to carry out analysis at this frequency position. Table 1 below shows the summary of the power levels in dB of the filtered signals as α varies.

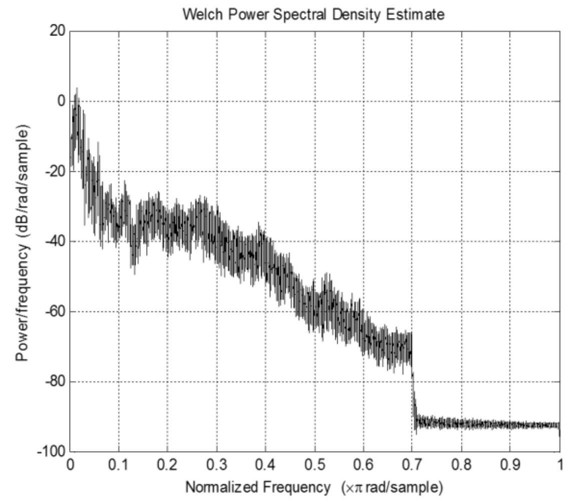

Fig.12: Power Spectral Density of Noise Free Speech Signal

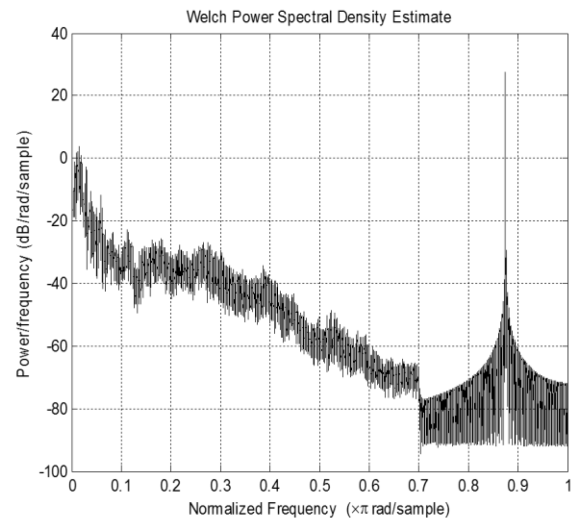

Fig.13: Power Spectral Density of Contaminated Signal

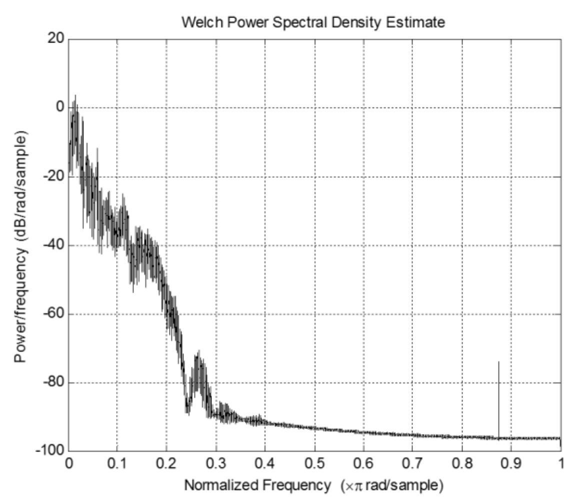

Fig.14: Power Spectral Density of Speech Signal Filtered With Low Pass Filter When α=0.0

ISSN No:-2456-2165

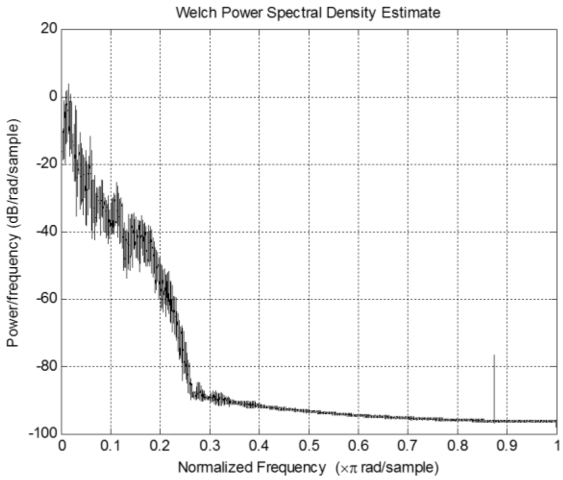

Fig.15: Power Spectral Density of Speech Signal Filtered With Low Pass Filter When **α**=0.07

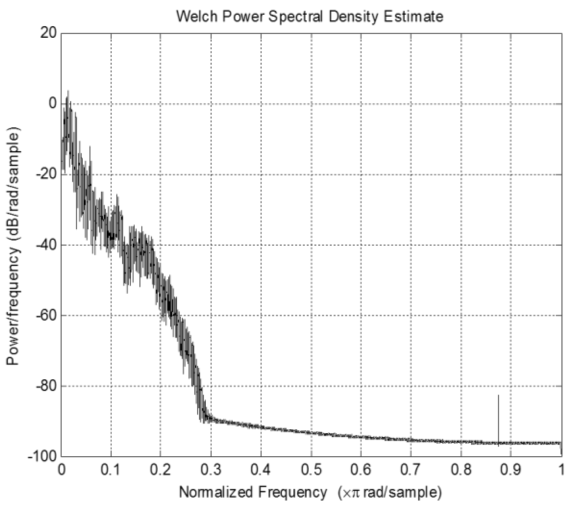

Fig.16: Power Spectral Density of Speech Signal Filtered With Low Pass Filter When **α**=0.16

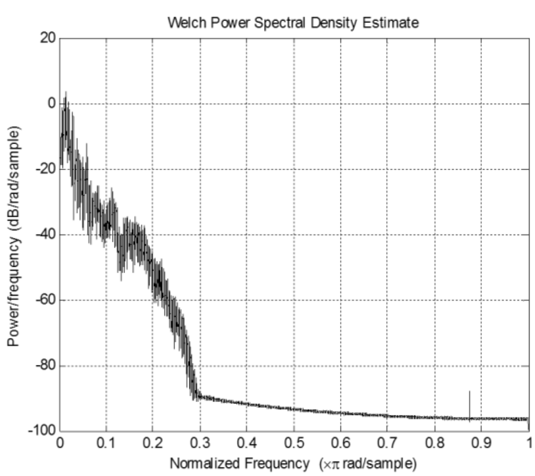

Fig.17: Power Spectral Density of Speech Signal Filtered With Low Pass Filter When **α**=0.2

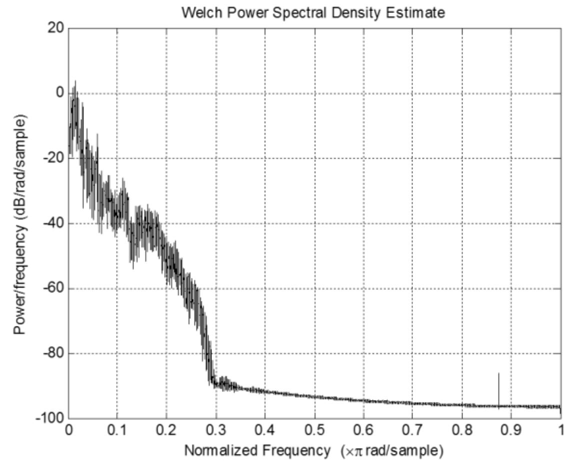

Fig.18: Power Spectral Density of Speech Signal Filtered With Low Pass Filter When **α**=0.3

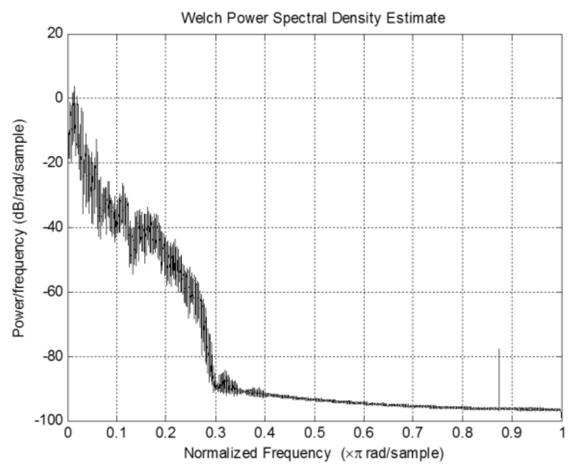

Fig.19: Power Spectral Density of Speech Signal Filtered With Low Pass Filter When **α**=0.4

Power level of unfiltered voice	+27.37dB
Power level of filtered voice when α=0.0	-77.10dB
Power level of filtered voice when α=0.07	-77.98dB
Power level of filtered voice when α=0.16	-85.12dB
Power level of filtered voice when **α** =0.2	-90.14dB
Power level of filtered voice when **α** =0.3	-88.21dB
Power level of filtered voice when **α** =0.4	-78.32dB

Table 1: Power Levels of Voice Filtered Signals at 0.875 Normalised Frequency

From table 1, power level of the corrupt speech signal is +27.42dB. Comparing it with the contaminated speech signal power level of -91.19dB which can be determined from fig. 12 shows that noise signal added a lot of noise power of 27.42-(-91.19)=118.61dB. Also from the table1 maximum attenuation of the noise is when **α=0.2,** giving a power spectral density of -90.14dB which implies that the optimum value of the adjustment parameter is 0.2 in the circumstance under consideration.

V. CONCLUSION

It can be concluded that double cosine term generalised adjustable window is a very effective window in designing FIR filters for speech signal denoising. For the speech signal, the noise type and level in this circumstance, the optimum value of the adjustment parameter is 0.2. This value may vary if a different type of signal, noise type or window length is under consideration.

REFERENCES

[1]. Saurabh Singh Rajput and Dr. S. S Bhaduaria. Implementation of FIR Filter Using Efficient Window Function and its Application in Filtering a Speech Signal. International Journal of Electrical, Electronic and Mechanical Controls, Vol.1, Issue1, 2012.
[2]. Poulami Das, Subhas Chandra, Sudip Kumar and Sankar Narayan. An Approach for Obtaining Least Noisy Signal Using Kaiser Window and and Genetic Algorithm. International Journal of Computer Applications, Vol. 150, No. 6, September, 2016, pp. 16-21.
[3]. Ritesh Pawar and Rajesh Mehra. Design and performance analysis of FIR Filter for Audio Applications. International Journal of Scientific Research in Engineering and Technology, 2014, pp. 122-126.
[4]. Ayush Gavel, Hem Lal Sahu, Gautum Sharma and Pranay Kumar Rahi. Design of Lowpass FIR Filter Using Rectangular and Hamming Window Techniques. International Journal of Innovative Science, Engineering and Technology, Vol. 3, Issue 8, August, 2016, pp. 251-256.
[5]. Gopika P. and Supriya Subash. Effect of High Pass Filters on Human Voice. Proceedings of 39th IRF International Conference, 27th March, 2016, pp. 55-59.
[6]. Akanksha Raj, Arshiyanaz Khateeb and Fakrunnisa Balaganur. Design of FIR Filter for Efficient Utilisation of Speech Signal. International Journal for Scientific Research and Development, Vol. 3, Issue 3, 2015.
[7]. S. Sangeetha and P. Kannan. Design and Analysis of Digital Filters for Speech Signals Using Multirate Signal Processing. ICTACT Journal on Microelectronics, Vol. 03, Issue 04, January, 2018, pp. 480-486.
[8]. P. Suresh Babu, Dr. D. Srinivasulu Reddy, and Dr. P. V. N. Reddy . Speech Signal Analysis Using Windowing Techniques. International Journal of Emerging Trends in Engineering Research, Vol.3, No.6, 2015, pp. 257-263.
[9]. Rohit Patel, Er. Mukesh Kumar, Prof. A. K. Jaiswal and Er. Rohini Saxena. Design Technique of Bandpass FIR Filter Using Various Window Function. IOSR Journal of Electronics and Communication Engineering, Vol. 6, Issue 6, July- August, 2013, pp. 52-57
[10]. Mbachu C. B. and Nwosu A. W. Performance Analysis of Various Infinite Impulse Response (IIR) Digital Filters in the Reduction of Powerline Interference in ECG Signals. International Journal of Scientific and Engineering Research, Vol. 5, Issue 11, November, 2014, pp. 498-506.
[11]. Marc Karam, Hasan F. Khazaal, Heshmat Aglan and Cliston Cole. Noise Removal in Speech Processing Using Spectral Subtraction. Joiurnal of Signal and Information Processing, Vol. 5, 2014, pp. 32-41.