Global Journal of Researches in Engineering: F
Electrical and Electronics Engineering
Volume 14 Issue 6 Version 1.0 Year 2014
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 2249-4596 & Print ISSN: 0975-5861

Speech Enhancement using Boll's Spectral Subtraction Method based on Gaussian Window

By S. China Venkateswarlu, A. Subba Rami Reddy & K. Satya Prasad
Jyothishmathi Institute of Technology & Science, India

Abstract- This paper investigates the effect of Gaussian window frequency response Side lobe Attenuation on the improvement of Speech quality in terms of six objective quality measures. In Speech Enhancement process, signal corrupted by noise is segmented into frames and each segment is Windowed using Gaussian window with variation in the side lobe attenuation parameter "**α**". The Windowed Speech segments are applied to the Boll's Spectral Subtraction Speech Enhancement algorithm and the Enhanced Speech signal is reconstructed in its time domain. The focus is to investigate the effect of Gaussian window frequency response side lobe level on the Boll's Spectral Subtraction Speech enhancement. For various side lobe attenuations of the Gaussian window frequency response, speech quality objective measures have been computed. From this study, it is observed that the Side lobe Attenuation parameter "**α**" plays an important role on the Speech enhancement process in terms of six objective quality measures. The results are compared with the measures of Hamming window and an optimum side lobe attenuation parameter value for the Gaussian window is proposed for better speech quality.

Keywords: boll's spectral subtraction, dft, gaussian window, objective measures, speech enhancement, side lobe attenuation.

GJRE-F Classification : FOR Code: 170204

SPEECHENHANCEMENTUSINGBOLLSSPECTRALSUBTRACTIONMETHODBASEDONGAUSSIANWINDOW

Speech Enhancement using Boll's Spectral Subtraction Method based on Gaussian Window

S. China Venkateswarlu [α], A. Subba Rami Reddy [σ] & K. Satya Prasad [ρ]

Abstract- This paper investigates the effect of Gaussian window frequency response Side lobe Attenuation on the improvement of Speech quality in terms of six objective quality measures. In Speech Enhancement process, signal corrupted by noise is segmented into frames and each segment is Windowed using Gaussian window with variation in the side lobe attenuation parameter "**α**". The Windowed Speech segments are applied to the Boll's Spectral Subtraction Speech Enhancement algorithm and the Enhanced Speech signal is reconstructed in its time domain. The focus is to investigate the effect of Gaussian window frequency response side lobe level on the Boll's Spectral Subtraction Speech enhancement. For various side lobe attenuations of the Gaussian window frequency response, speech quality objective measures have been computed. From this study, it is observed that the Side lobe Attenuation parameter "**α**" plays an important role on the Speech enhancement process in terms of six objective quality measures. The results are compared with the measures of Hamming window and an optimum side lobe attenuation parameter value for the Gaussian window is proposed for better speech quality.

Keywords: boll's spectral subtraction, dft, gaussian window, objective measures, speech enhancement, side lobe attenuation.

I. Introduction

In Speech Processing, Speech enhancement is one of the most important fields and finds many applications such as mobile phones, teleconferencing systems, speech recognition and hearing aids. The processed speech signals are supposed to be more comfort for listening and also should give better performance in tasks like automatic speech and speaker recognition [1]. Several algorithms are proposed in the literature for speech enhancement such as spectral subtraction methods, MMSE methods, Weiner algorithm etc. [2]. This paper attempts the Boll's Spectral Subtraction method of Speech Enhancement [3]. In this Method, the noisy speech signal is partitioned into frames. Each frame is multiplied by a window function prior to the spectral analysis and applied to the speech enhancement algorithm. This work investigates the effect of windowing the speech signal in the process of Speech Enhancement in terms of six Objective Speech Quality measures using Boll's Spectral Subtraction Method for Speech Enhancement process.

The purpose of windowing is to reduce the effect of discontinuity introduced by the framing process. Commonly used windows include Hamming and Hanning [4]. Although these windows have a reduced side lobe levels they have also reduced frequency resolution. Hence several factors enter into the choice of Window selection to frame the Speech for Enhancement. In this paper an attempt has been made to explore the possibility of improving the quality of speech signal by employing Gaussian window with different "**α** "values. To study the performance of any algorithm, combinations of subjective and objective measures have to be carried on. Currently, the accurate method for evaluating speech quality is through subjective listening tests. But it is costly and time consuming. Hence, six Objective measures are chosen to evaluate the performance of the Gaussian window in the enhancement system. P. Loizou has presented a correlation analysis of Objective Quality measures for evaluating speech enhancement algorithms [5]. In this paper six measures namely SNR, Segmental SNR (Seg-SNR), Log Likelihood Ratio(LLR), Weighted spectral slope distance(WSS), Frequency weighted segmental SNR (fwseg-SNR) and Cepstral Distance (Cep) are selected for performance evaluation test, considering the fact that Fwseg-SNR, LLR, Cep and WSS have high correlation with overall speech quality. The correlation coefficients for these measures with speech quality are 0.84, 0.85, 0.79 and 0.64 respectively [5]. These objective measures also have good correlation with subjective scores. Although the correlation coefficient of SegSNR is 0.36, it is chosen as a time domain measure where as the above measures are of frequency domain. This paper explains the effect of the shape parameter of the Gaussian window on the noisy speech for Enhancement in terms of the six Objective measures using Speech Enhancement algorithm. The rest of the paper is organized as follows: Section-2 briefly explains the various windows for noisy Speech Enhancement. In Section-3, the Six Objective measures used in this study are presented. In Section-4 Boll's Spectral Subtraction method for noisy speech enhancement is explained. Implementation of the scheme is presented in Section-5, Section-6 explains the results and discussions, Section- 7 presents the Simulation Results and finally Section-8 describes the conclusions.

Author α: Professor-ECE, Jyothishmathi Inst. of Tech. & Science & Research Scholar, JNTUK-Kakinada, A. P, India.
e-mail: cvenkateswarlus@gmail.com
Author σ: Principal, SKIT, Srikalahasti, Chittoor Dt, A.P, India.
e-mail: skitprincipal@gmail.com
Author ρ: Professor-Dept. of ECE JNTUK-Kakinada, A. P, India.
e-mail: Prasad_kodati@yahoo.co.in

II. Data Weighting Windows

a) Windowing

Windows are time-domain weighting functions that are used to reduce Gibbs' oscillations resulting from the truncation of a Fourier series [6-7]. Their roots date back over one-hundred years to Fejer's averaging technique for a truncated Fourier series and they are employed in a variety of traditional signal processing applications including power spectral estimation, beam forming, and digitalfilter design. The effect of a time window can be described in the frequency domain as a convolution of the frequency response of the window with the frequency response of the signal. The convolution smears frequency features, with the amount of smearing depending on the width of the main lobe of the window frequency response. For signal frequencies, observed through the rectangular window, which do not correspond exactly to one of the sampling frequencies, the pattern is shifted such that non-zero values are projected onto all sampling frequencies. This phenomenon of spreading signal power from the nominal frequency across the entire width of the observed spectrum is known as spectral leakage. In addition, spectral leakage from distant frequency components will occur if the side lobe level of the window response is large [8-18]. Ideally, the window spectrum should have a narrow main-lobe and small side-lobes. However, there is an inherent trade-off between the width of the main-lobe and the side-lobe attenuation. A wide main-lobe will average adjacent frequency components and large side lobes will introduce contamination (or spectral leakage) from other frequency regions. For rectangular window, the main lobe is narrower than that of the Hamming window, while its side-lobes are higher. Some of the commonly used windows in speech processing are symmetric (e.g., Hamming and Hanning windows) or asymmetric (such as the hybrid Hamming-Cosine window). The goal of asymmetric windows is to reduce the algorithmic delay in speech codes.

b) Gaussian Window

In contrast to the other fixed windows, the Gaussian Window [8] has two parameters: the length of the sequence N and a shape parameter "α". In short-time spectral amplitude (STSA), the length of the window is fixed and is equal to the speech signal frame length and hence the side lobe attenuation parameter "α" can be varied. As the parameter increases the side lobe level of the frequency response decreases. In this paper, the speech enhancement in terms of objective measures as a function of side lobe attenuation parameter "α" has been investigated.

The Gaussian Window can be obtained using (1). Figures 2.1 and 2.2 indicate the time and frequency description of a Gaussian Window with α=2.5 and α=3.5 respectively.

$$w(n) = e^{-\frac{1}{2}\left(\alpha\frac{n}{N/2}\right)^2} \qquad 0 \le |n| \le N/2 \qquad (1.1)$$

Where, "α" is inversely proportional to the standard deviation of a Gaussian random variable. The exact correspondence with the standard deviation, "σ", of a Gaussian probability density function is

$$\sigma = \frac{N}{2\alpha} \qquad (2.2)$$

The width of the window is inversely related to the value of "α", a larger values of "α" produces a narrower window.

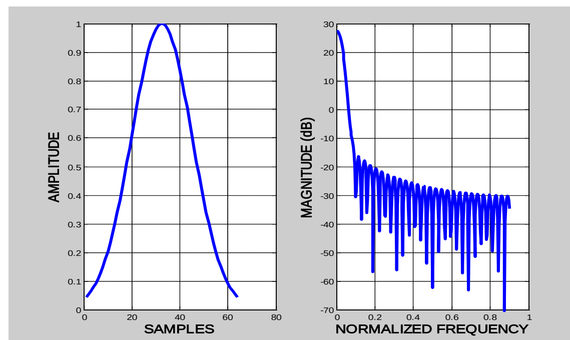

Figure 2.1 : The time and frequency description of a Gaussian Window with "α=2.5"

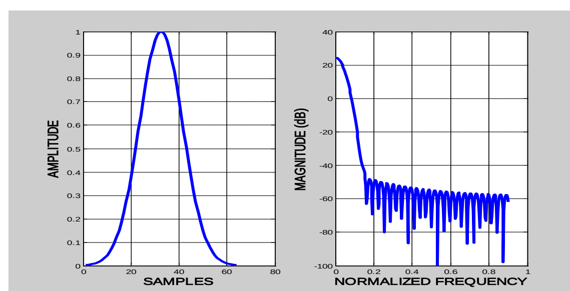

Figure 2.2 : The time and frequency description of a Gaussian Window with "α=3.5"

III. Objective Measures

a) Signal-to–Noise Ratio

The Signal-to-Noise Ratio (SNR) is the ratio of signal energy to noise energy and is expressed in decibels dB given by [2,4-5] as

$$SNR_{dB} = 10\log_{10}\left[\frac{\sum_n s^2(n)}{\sum_n [s(n)-\hat{s}(n)]^2}\right] \qquad (3.1)$$

Where s(n) is the undistorted or clean signal and $\hat{s}$ (n) is the degraded or processed/enhanced speech signal , N is the frame length.

b) The Seg-SNR

The Seg-SNR is the frame-based SNR and is estimated as It is defined [2,4-5] as

$$SegSNR_{dB} = \frac{1}{M}\sum_{m=0}^{M-1} 10\log_{10}\left[\frac{\sum_{n=0}^{N-1}|s(n+mN)|^2}{\sum_{n=0}^{N-1}|s(n+mN)-\hat{s}(n+mN)|^2}\right] \qquad (3.2)$$

Where s(n) is the undistorted or clean signal and $\hat{s}$(n) is the degraded or processed / enhanced speech signal , N is the frame length. M represents the number of frames. The Seg-SNR poses a problem if there are intervals of silence in the speech utterance. In segments in which the original speech is nearly zero, any amount of noise can give rise to a large negative SNR for that segment, which could appreciably bias the overall measure of Seg-SNR. This problem is resolved by including the SNR of the frames only if the frame energy is above a specified threshold. Generally, the frames with segmental SNR between –10 dB to 35 dB are only considered in the average.

c) Weighted Spectral Slope Distance

WSS distance measure computes the weighted difference between the spectral slopes in each frequency band. The spectral slope is obtained as the difference between adjacent spectral magnitudes in decibels. The WSS measure is defined and evaluated [19] as

$$WSS = \frac{1}{M}\sum_{m=0}^{M-1}\left[\frac{\sum_{j=1}^{k} W(j,m)[s_c(j,m)-s_p(j,m)]^2}{\sum_{j=1}^{k} W(j,m)}\right] \quad (3.3)$$

Where W (j, m) are the weights computed. S_c (j, m) and S_p(j, m) are the spectral slopes for j^{th} frequency Band at m^{th} frame of clean and processed speech signals respectively.

d) Log Likelihood Ratio

The LLR measure is defined [20] as

$$\text{LLR}=\log_{10}\left[\frac{a_p R_s a_p^T}{a_s R_s a_s^T}\right] \quad (3.4)$$

Where a_p and a_s are the LP coefficient vectors for the clean and degraded or enhanced speech segments, respectively. Rs denote the autocorrelation matrix of the clean speech segment.

e) Cepstrum Distance

The Cepstrum distance [6] provides an estimate of the log spectral distance between two spectra. It is defined as

$$WSS = \frac{1}{M}\sum_{m=0}^{M-1}\left[\frac{\sum_{j=1}^{k} W(j,m)[s_c(j,m)-s_p(j,m)]^2}{\sum_{j=1}^{k} W(j,m)}\right] \quad (3.5)$$

Where C_s(n) and C_p(n) represent the cepstrum of clean and the degraded or enhanced speech respectively.

$$C_s(k,m)=\text{Re}[\text{IDFT}\{\log|DFT(s(k,m)|\}] \quad (3.6)$$

The cepstrum coefficients can also be obtained recursively from the LPC coefficients using the following expression [2, 5]

$$c(m) = a_m + \sum_{k=1}^{m-1}\frac{k}{m}c(k)a_{m-k} \quad \text{for } 1\le m\le p \quad (3.7)$$

f) Frequency Weighted Segmental SNR

It is computed [5, 21] using the following equation

$$fwSNRseg = \frac{10}{M}\sum_{m=0}^{M-1}\left\{\frac{\sum_{j=1}^{k} W(j,m)\log_{10}\frac{s(j,m)^2}{[s(j,m)-\hat{s}(j,m)]^2}}{\sum_{j=1}^{k} W(j,m)}\right\} \quad (3.8)$$

where W (j, m) is the noise-dependent weight applied on the jth frequency band, K is the number of bands, M is the total number of frames in the signal, s(j, m) is the weighted (by a Gaussian-shaped window) clean signal spectrum in the jth frequency band at the mth frame, and $\hat{s}$(j, m) in the weighted enhanced signal spectrum in the same band.

IV. Boll's Spectral Subtraction Method

The noise corrupted speech is processed by the Spectral Subtraction method to get processed or enhanced speech. Spectral Subtraction [3, 22, 24-25] is a popular frequency domain method to reduce the effect of additive uncorrelated noise in a signal. The noise spectrum is estimated, and updated, from the periods when the signal is absent and only the noise is present. For restoration of the time-domain signal, an estimate of the instantaneous magnitude spectrum is combined with the phase of the noisy signal, and then transformed via an Inverse Discrete Fourier Transform (IDFT) to the time domain.

If y(n) is the discrete noise corrupted input signal which is composed of the clean speech signal s(n) and v(n) the uncorrelated additive noise signal, then the noisy signal can be represented as:

$$y(n) = s(n) + v(n) \quad (4.1)$$

Since the speech is not a stationary signal, the processing is carried on short-time basis (frame-by-frame).

$$y(n, k) = s(n, k) + v(n, k) \quad (4.2)$$

Where n is the time index and k is the frame index, y(n, k) is the k th frame. In the frequency domain, with their respective Fourier transforms, the power spectrum of the noisy signal can be represented as:

$$P_{yy}(w,k)=P_{ss}(w,k)+P_{vv}(w,k) \quad (4.3)$$

$$|Y(w,k)|^2=|S(w,k)|^2 + |V(w,k)|^2 \quad (4.4)$$

Here Y(w, k) is the DFT of y(n, k) given by

$$Y(w,k) = \sum_{n=0}^{N-1} y(n,k)e^{-j\frac{2\pi wn}{N}} = |Y(w,k)|\, e^{j\varphi(w)} \quad (4.5)$$

Where φ(w) is the phase of the corrupted noisy signal, and N is the number of samples in the framed speech signal.

Thus from Eq.4.4 the estimation of clean speech signal can be given as

$$|\hat{S}(w,k)|^2=|Y(w,k)|^2 - |\hat{V}(w,k)|^2 \quad (4.6)$$

Once the estimate of the clean speech is obtained in the spectral domain, the enhanced speech signal is obtained according to:

$$\hat{s}(n,k) = \mathrm{IDFT}\{|\hat{S}\,(w,k)|e^{j\varphi(w)}\} \qquad (4.7)$$

Here, the phase information from the corrupted signal is used to reconstruct the time domain signal by taking the IDFT.

One may generalize the technique of spectral subtraction by replacing the magnitude squared of the DFT by some power of the magnitude. The exponent, '2', in Equation.4.6 can be replaced by 'a' as given below:

$$|\hat{S}\,(w,k)|^{a} = |Y\,(w,k)|^{a} - |\hat{V}\,(w,k)|^{a} \qquad (4.8)$$

To estimate the noise, a method of exponential averaging proposed in [23] is used to estimate the noise. The frame-by-frame update scheme using the exponential averaging method is given below:

$$|\hat{V}\,(w,k)|^{a} = \begin{cases} \mu|\hat{V}\,(w,k-1)|^{a} + (1-\mu)|\hat{Y}\,(w,k)|^{a} & \text{for } \textit{Noise only} \\ |\hat{V}\,(w,k)|^{a} & \text{Speech and Noise} \end{cases} \qquad (4.9)$$

In this paper, all the measures were evaluated with a = 1 by restricting the study for magnitude spectrum only.

Where 0≤ μ ≤ 1 is exponential averaging constant. In this work "μ" is selected as 0.9. The block diagram for the overall spectral subtraction algorithm is shown in Figure5.1 below.

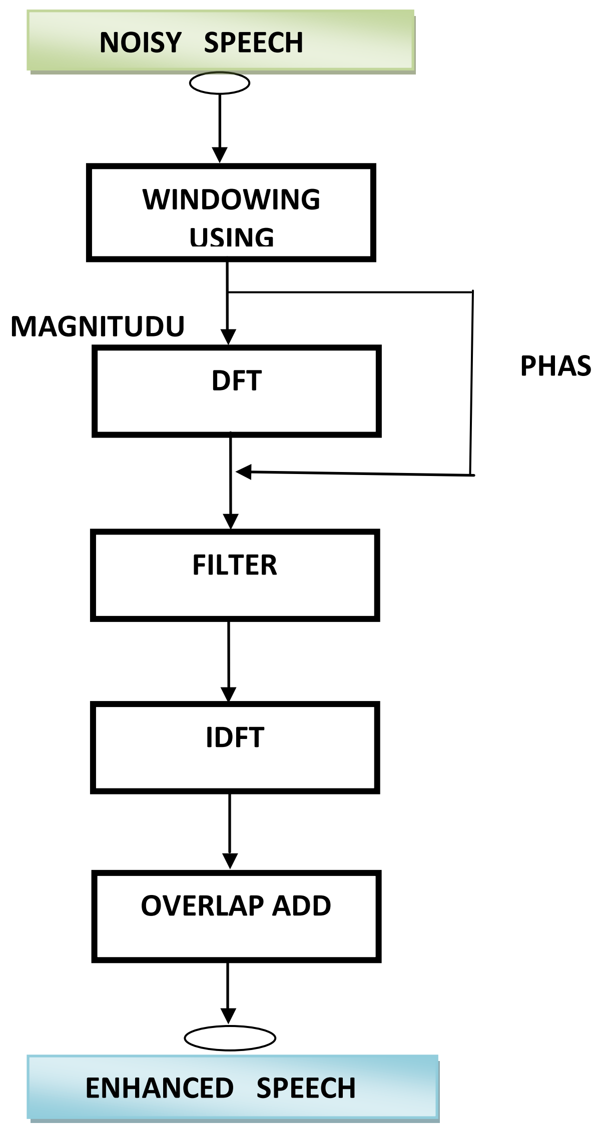

Figure 4.1 : Block diagram for Spectral Subtraction Algorithm

V. Implementations

Phonetically balanced clean speech signals and noise corrupted signals at different SNR levels have been taken from a speech corpus called "NOIZEUS". Noise corrupted speech signal is segmented into frames containing 256 samples of 32ms length (at 8 KHz Sampling rate). 256- point Discrete Fourier Transform (DFT) of each segment is obtained after applying the Gaussian window with variable shape parameter **"α"**.

Spectral Subtraction Algorithm is applied to the spectral components of each segment using Eq.4.1-Eq.4.9. The signal is reconstructed in its time domain with the help of IDFT and overlap add method (with 50% overlap between frames). The signal thus obtained is the enhanced signal. The performance of the enhanced signal is analyzed by using six objective measures for speech enhancement. The measures are WSS, LLR, fwseg-SNR, Cep, Seg-SNR, and SNR defined in Eq.3.1-Eq.3.8. All the measures are computed by segmenting the sentences using 32-ms duration Hamming windows with 75% overlap between adjacent frames. A tenth order LPC analysis was used in the computation of LPC-based objective measure LLR. The performance of Gaussian windowed signal is studied under two real world noise conditions namely "Car noise" and "Airport noise" at 0dB, 5dB, 10dB and 15dB SNR levels and presented in Table.1 (a)-1(h) and Table.2 (a)-2(h)

VI. Results and Discussions

The Objective measure scores are shown in Fig. 2(a) - 2(b) as a function of the shape parameter **"α"** of the Gaussian window used as analysis window. The following observations can be made based on these results. For the narrow main lobe width of the analysis Window's frequency response, the Objective measure scores are good and as expected. The measures WSS, LLR and CEPSTRAL DISTANCE increase with the increase of **"α"**. This result indicates that, the main lobe width of the analysis Window increases and contributes the unwanted noise for the noise corrupt signal which results degradation in the Speech Enhancement process.

It is important to note that Fig. 2(a)-2(h) shows steep variation in Objective measures is observed when the shape parameter **"α"** is in between 0.3 and 0.8 for Car Noise and between 1.5 and 3.5 for Airport noise, the shape of the analysis window's frequency response plays significant role on the speech quality measures in terms of the above objective measures. It is also observed that, further increase beyond 1.0 for Car noise and beyond 3.5 for Airport noise in the shape parameter **"α"** there is no significant improvement in the objective intelligibility scores as a function of analysis window shape variable. This can be attributed by the fact that the main lobe width and side lobe levels of the window's frequency response are compliment to each other. Narrower level main lobe width will tend to increase the side lobe levels and vice versa. Hence the contribution of undesired spectral components due to the main lobe width will be compensated due to the reduced side lobe levels. Hence the overall improvement in the Objective measures is not observed as in the case of variation of the shape parameter values beyond the above specified values. Based on objective intelligibility scores, it can be seen that the optimum window shape parameter for speech analysis is between 0.3 and 0.8 for Car noise and between 1.5 and 3.5 for Airport noise. For speech applications based solely on the short-time magnitude spectrum, the Gaussian window with the above shape parameter is expected to be the right choice. This study proposes that the optimum shape parameter for the Gaussian window is $\alpha=0.5$ for Car Noise and 2.8 for Airport Noise.

Comparative analysis of Gaussian window with the Hamming window using six objective measures of speech quality is made for the case of speech signal contaminated with Car noise and Airport noise at various SNRs of 0dB, 5dB, 10dB and 15dB along with clean speech signals. The results are presented in Table.1 (a)-1(h) and Table.2 (a)-2(h). Considering the fact that, higher SNR, Seg-SNR and fwseg-SNR values give better quality where as WSS, Cep and LLR measures, lower values indicate a better quality. From the Table.1 (a)-1(h) and Table.2 (a)-2(h), shown, it can be noticed that majority of the objective measures are improved for Gaussian Window with shape parameter **"α"** when compared with Hamming windowed measure. From WSS measure it is observed that, a significant improvement is achieved with the proposed window scheme. Observing the results presented in the Tables, the proposed window method shown is able to remove the residual noise in better manner compared to the Hamming window method and observing all the results, one can have a judicious choice for **"α=0.5"** Car noise and **"α=2.8"** for Airport noise as optimum values and can be used for Speech Enhancement process for better results using Spectral Subtraction method.

Further from the Fig (2a) for the case of Car noise it is observed that under low noise conditions (0dB), comparing with Hamming window the Gaussian window with above optimum shape parameter gives better Results where this speech activity is clearly visible. But the same is absent when using Hamming window. This speech activity visible using Gaussian window is encircled with Red mark.

Where as in the case of airport noise, the enhanced speech signal using Gaussian window looks identical with the enhanced speech signal using Hamming window. But from the objective measures with the above shape parameter values majority of the objective measures are in favour of Gaussian window.

Global Journal of Researches in Engineering (F) Volume XIV Issue VI Version I Year 2014 13

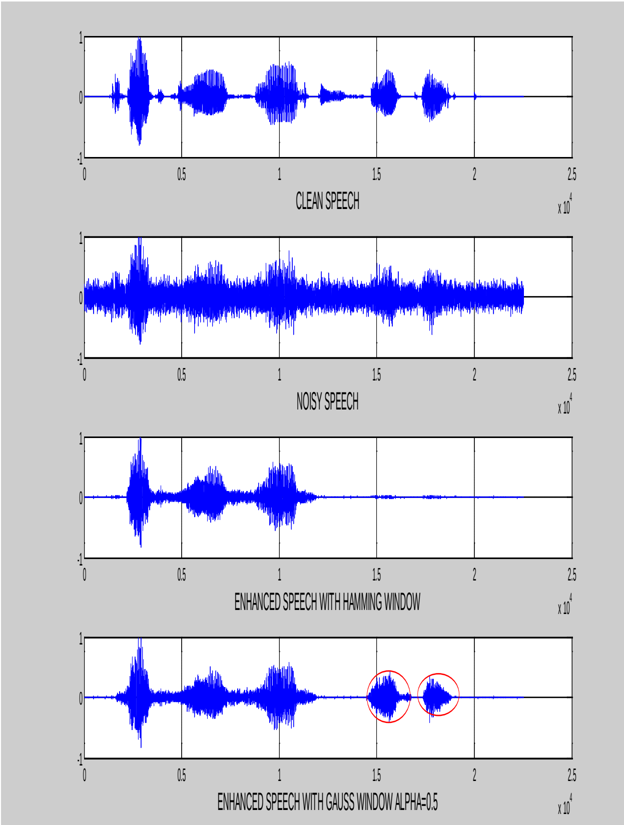

Fig. 2 (a) : Speech Signal Enhancement comparison (α=0.5, CAR-NOISE of 0dB SNR)

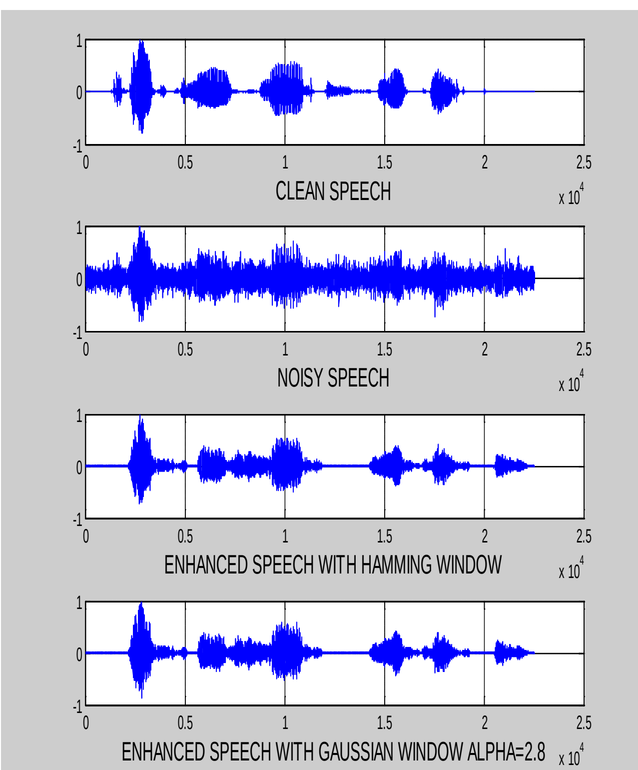

Fig. 2 (b) : Speech Signal Enhancement comparison (α=2.8, AIRPORT-NOISE of 0dB SNR)

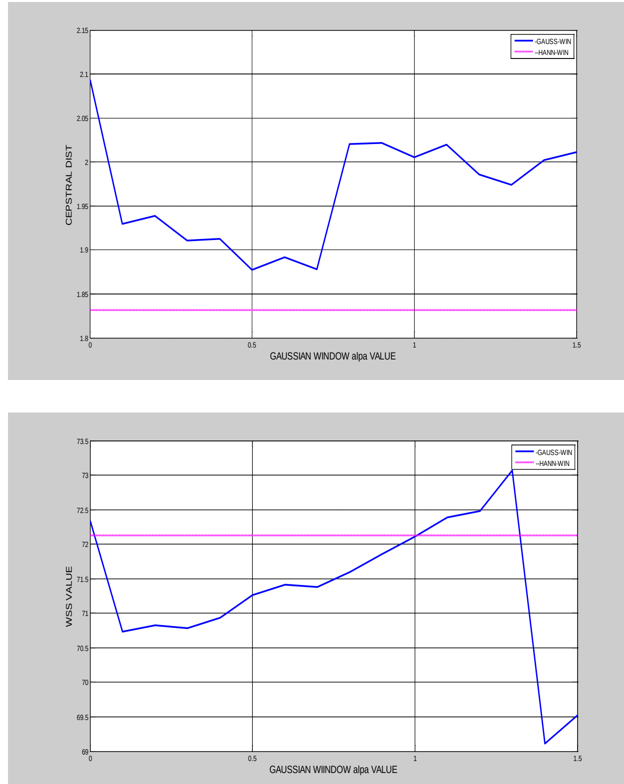

Fig. 2 (c) : Speech quality Objective Measures CEP and WSS Variation With "α", (CAR-NOISE of 0dB SNR)

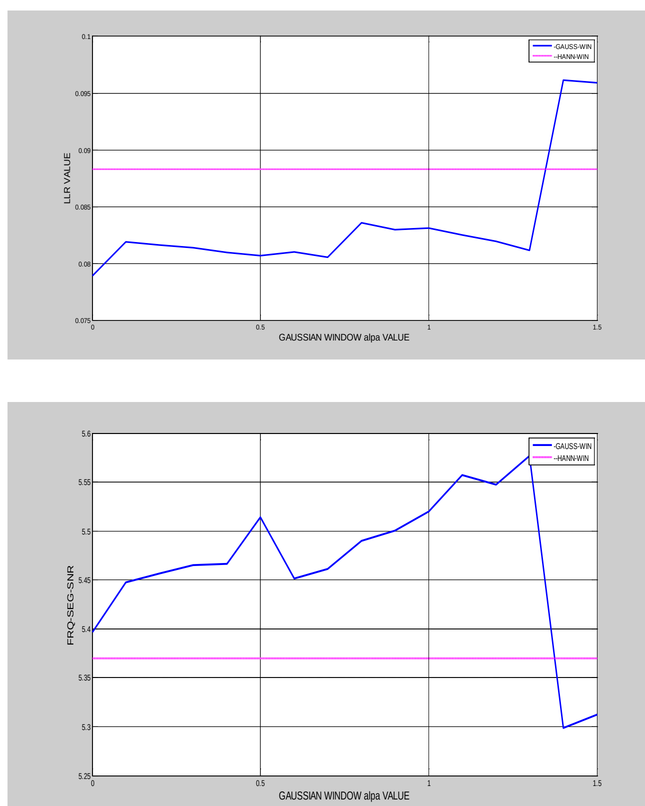

Fig. 2 (d) : Speech quality Objective Measures LLR and FSG-SNR Variation With "α", (CAR-NOISE of 0dB SNR)

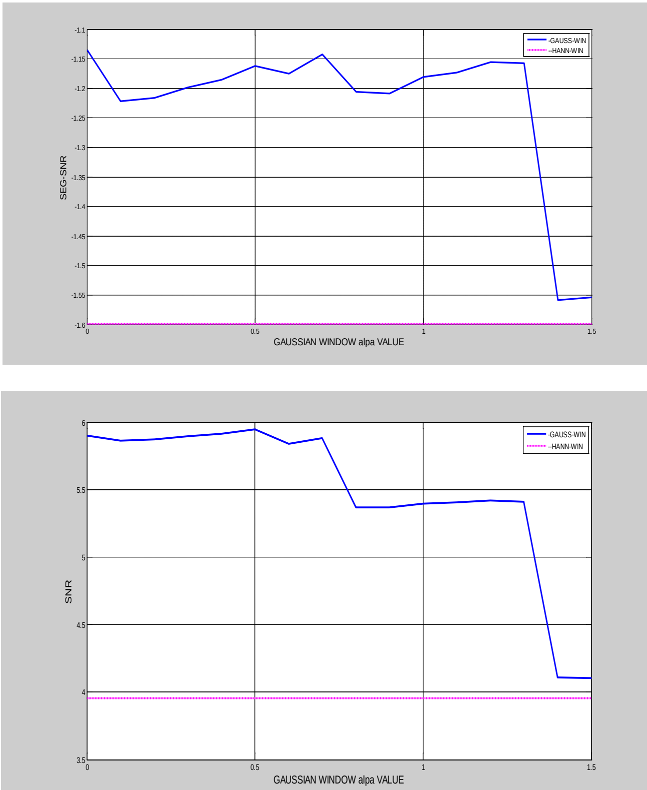

Fig. 2 (e) : Speech quality Objective Measures SEG-SNR and SNR Variation With "α", (CAR-NOISE of 0dB SNR)

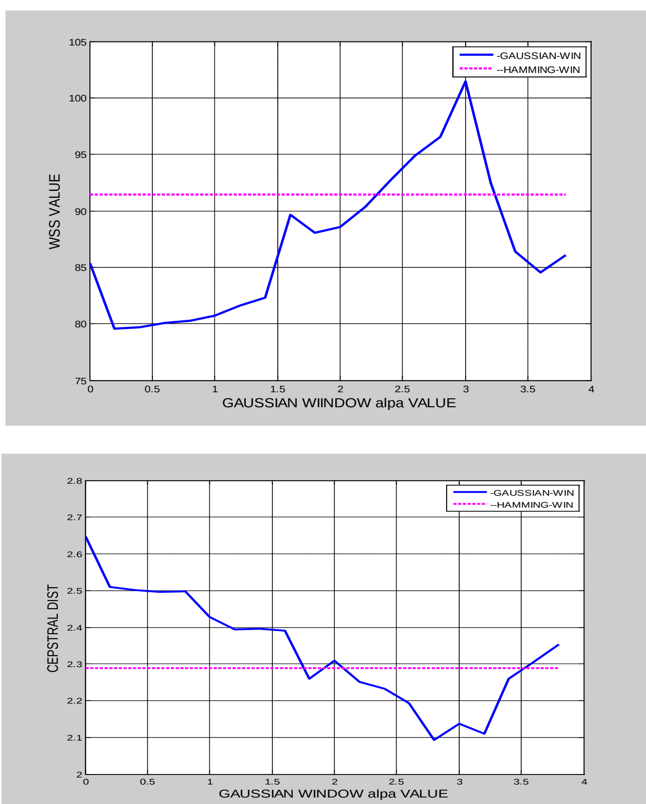

Fig. 2 (f) : Speech quality Objective Measures CEP and WSS Variation With "α", (AIRPORT-NOISE of 0dB SNR)

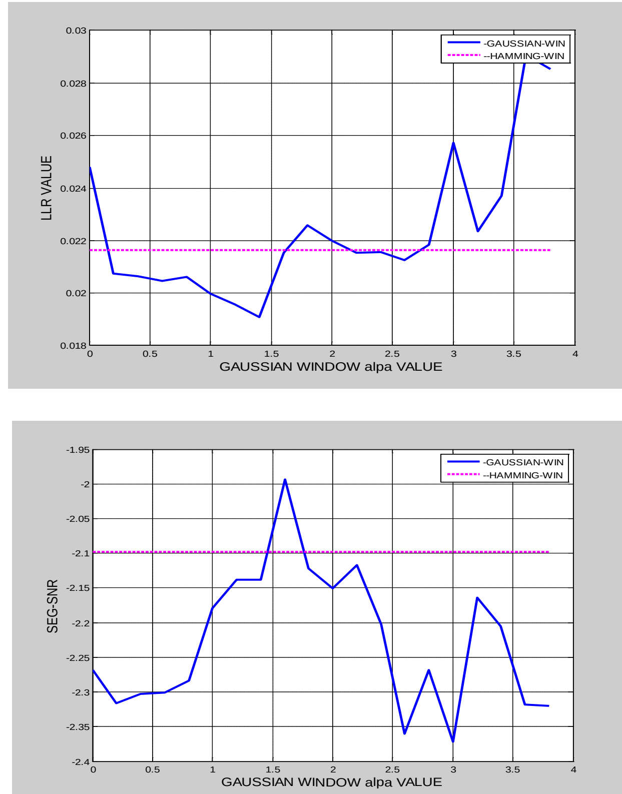

Fig. 2 (g) : Speech quality Objective Measures LLR and SEG-SNR Variation With "α", (AIRPORT-NOISE of 0dB SNR)

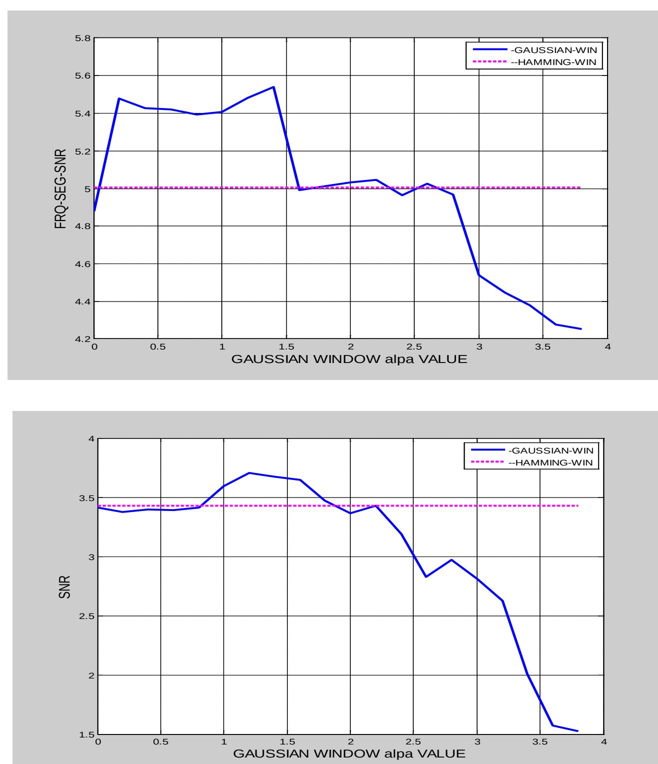

Fig. 2 (h): Speech quality Objective Measures FRS-SNR and SNR Variation With "α", (AIRPORT-NOISE of 0dB SNR)

Global Journal of Researches in Engineering (F) Volume XIV Issue VI Version I Year 2014
15

Table 1 (a) : Variation of Objective Measures with Gaussian Window Shape Parameter "α". CAR-NOISE of 15dB SNR Value

S.No	Objective Measure	Hamming Window	Gaussian window with different Alpha values					
			α=0	α=0.25	α=0.50	α=0.75	α=1.0	α=1.25
1	LLR	0.006	0.009	0.010	0.009	0.009	0.008	0.008
2	WSS	46.614	47.055	42.986	43.119	43.214	43.729	43.879
3	Cep	1.386	1.660	1.673	1.655	1.632	1.585	1.549
4	Seg-SNR	5.257	4.942	5.311	5.332	5.353	5.400	5.435
5	Fwseg-SNR	10.480	9.754	10.365	10.365	10.41	10.391	10.436
6	SNR	11.789	12.462	13.023	13.044	13.078	13.078	13.069

Table 1 (b) : Variation of Objective Measures with Gaussian Window Shape Parameter "α". AIRPORT-NOISE of 5dB SNR Value

S.No	Objective Measure	Hamming Window	Gaussian window with different Alpha values					
			α=0	α=0.75	α=1.50	α=2.25	α=3.0	α=3.75
1	LLR	0.019	0.025	0.023	0.021	0.019	0.018	0.017
2	WSS	85.570	79.131	78.281	80.928	84.636	92.385	95.729
3	Cep	1.751	2.083	1.997	1.881	1.774	1.592	1.471
4	Seg-SNR	0.122	0.122	0.177	0.260	0.131	-0.020	-0.435
5	Fwseg-SNR	6.777	6.619	6.835	6.984	6.780	6.516	6.126
6	SNR	6.497	7.119	7.134	7.085	6.591	6.119	5.148

Table 1 (c) : Variation of Objective Measures with Gaussian Window Shape Parameter "α". AIRPORT-NOISE of 10dB SNR Value

S.No	Objective Measure	Hamming Window	Gaussian window with different Alpha values					
			α=0	α=0.75	α=1.50	α=2.25	α=3.0	α=3.75
1	LLR	0.006	0.012	0.0103	0.008	0.006	0.006	0.006
2	WSS	63.324	59.279	57.040	59.917	64.340	69.122	76.522
3	Cep	1.577	1.955	1.803	1.746	1.617	1.399	1.343
4	Seg-SNR	2.581	2.465	2.706	2.710	2.600	2.335	1.563
5	Fwseg-SNR	8.912	8.414	8.768	9.027	8.780	8.376	7.449
6	SNR	9.587	9.988	10.284	10.273	9.599	8.385	6.317

Table 1 (d) : Variation of Objective Measures with Gaussian Window Shape Parameter "α".AIRPORT-NOISE of 15dB SNR Value

S.No	Objective Measure	Hamming Window	Gaussian window with different Alpha values					
			α=0	α=0.75	α=1.50	α=2.25	α=3.0	α=3.75
1	LLR	0.005	0.009	0.007	0.006	0.005	0.005	0.006
2	WSS	52.939	50.083	46.783	48.267	52.520	60.610	71.138
3	Cep	1.373	1.733	1.665	1.565	1.425	1.265	1.214
4	Seg-SNR	4.635	4.451	4.877	4.959	4.614	3.735	2.551
5	Fwseg-SNR	10.647	9.857	10.509	10.638	10.636	9.926	8.483
6	SNR	11.635	12.166	12.819	12.777	11.726	9.497	6.806

Table 1 (e) : Variation of Objective Measures with Gaussian Window Shape Parameter "α". CAR-NOISE of 0dB SNR Value

S.No	Objective Measure	Hamming Window	Gaussian window with different Alpha values					
			α=0	α=0.25	α=0.50	α=0.75	α=1.0	α=1.25
1	LLR	0.064	0.066	0.066	0.066	0.067	0.065	0.064
2	WSS	79.107	78.541	76.820	76.606	77.004	77.804	78.035
3	Cep	1.925	2.218	2.117	2.089	2.0710	2.049	2.038
4	Seg-SNR	-1.114	-1.277	-1.247	-1.196	-1.233	-1.188	-1.155
5	Fwseg-SNR	5.027	4.828	4.936	4.966	4.951	4.999	5.012
6	SNR	4.149	4.086	4.277	4.324	4.1907	4.229	4.305

Table 1 (f) : Variation of Objective Measures with Gaussian Window Shape Parameter "α" CAR-NOISE of 5dB SNR Value

S.No	Objective Measure	Hamming Window	Gaussian window with different Alpha values					
			α=0	α=0.25	α=0.50	α=0.75	α=1.0	α=1.25
1	LLR	0.036	0.039	0.036	0.036	0.035	0.036	0.036
2	WSS	66.712	66.211	63.654	62.903	62.968	63.499	63.785
3	Cep	1.731	1.969	1.848	1.819	1.769	1.797	1.800
4	Seg-SNR	0.533	0.700	0.642	0.675	0.727	0.728	0.746
5	Fwseg-SNR	6.604	6.195	6.543	6.588	6.639	6.646	6.638
6	SNR	6.351	6.863	7.141	7.160	7.182	7.179	7.173

Table 1 (g) : Variation of Objective Measures with Gaussian Window Shape Parameter "α". CAR-NOISE of 10dB SNR Value

S.No	Objective Measure	Hamming Window	Gaussian window with different Alpha values					
			α=0	α=0.25	α=0.50	α=0.75	α=1.0	α=1.25
1	LLR	0.013	0.014	0.014	0.014	0.014	0.014	0.013
2	WSS	55.790	55.784	51.344	51.141	51.119	50.728	51.435
3	Cep	1.540	1.825	1.759	1.754	1.733	1.716	1.676
4	Seg-SNR	3.043	2.943	3.223	3.242	3.279	3.221	3.265
5	Fwseg-SNR	8.988	8.336	8.998	9.056	9.017	9.0342	9.080
6	SNR	10.145	10.531	11.045	11.095	11.116	11.019	11.038

Table 1 (h) : Variation of Objective Measures with Gaussian Window Shape Parameter "α". CAR-NOISE of 15dB SNR Value

S.No	Objective Measure	Hamming Window	Gaussian window with different Alpha values					
			α=0	α=0.25	α=0.50	α=0.75	α=1.0	α=1.25
1	LLR	0.006	0.009	0.010	0.009	0.009	0.008	0.008
2	WSS	46.614	47.055	42.986	43.119	43.214	43.729	43.879
3	Cep	1.386	1.660	1.673	1.655	1.632	1.585	1.549
4	Seg-SNR	5.257	4.942	5.311	5.332	5.353	5.400	5.435
5	Fwseg-SNR	10.480	9.754	10.365	10.365	10.41	10.391	10.436
6	SNR	11.789	12.462	13.023	13.044	13.078	13.078	13.069

VII. Conclusion and Future Enhancement

Speech signal is enhanced with the help of Spectral Subtraction method using Gaussian Window with suitable shape parameter **"α".** It is found that the speech quality in terms of the Objective measures is improved by applying the Gaussian Window with optimum shape parameter. From the figure.2 (a)-2(g), it is observed that the objective measure WSS has significant improvement for Gaussian Windowed signal when compared to Hamming Windowed signal. Majority

Global Journal of Researches in Engineering (F) Volume XIV Issue VI Version I Year 2014

the measures evaluated indicate that Gaussian Window with suitable shape parameter is superior to the Hamming Window in speech enhancement application using Spectral Subtraction method. It is also observed that the Window presented in this paper works out good for different types of noise, like babble, train, street and restaurant noises, at different SNR levels. Hence it may be concluded that Gaussian Window with suitable shape parameter is an attractive option compared with Hamming Window for the Speech Enhancement using Spectral Subtraction method for better Speech quality and intelligibility.

VIII. Acknowledgements

The work is carried out through the research facility at the Department of Computer Science & Engineering, Department of Electronics & Communication Engineering and Department of Electrical & Electronics Engineering, Jyothishmathi Institute of Technology & Science, Nustulapur, Karimnagar District, Telangana & Dept., ECE from JNTUK. The Authors also would like to thank the authorities of JNTUK, Kakinada. SKIT for encouraging this research work. Our thanks to the experts who have contributed towards development of this paper.

References Références Referencias

1. J Lim & A Oppenheim, Enhancement and bandwidth compression of noisy speech, *Proc IEEE*, vol 67, pp 1586-1604, Dec 1979.
2. P. C. Loizou, Speech Enhancement: Theory and Practice, 1st ed. CRC, 2007.
3. S Boll, Suppression of acoustic noise in speech using spectral subtraction, *IEEE Tans Acoust., Speech, Signal Processing*, vol ASSP-27, pp 113-120, April 1979.
4. A. M. Kandoz, "Digitalspeech", 2nd edition, Willey, 2002.
5. Y.Hu & P. Loizou, "Evaluation of objective measures for speech enhancement". IEEE Trans. Audio speech Lang. process, Vol .16. No.1, pp.229-238, Jan-2008.
6. S. K. Mitra, "Digital signal processing: A computer based approach", 2nd edition, MCGraw Hill, 2002.
7. V. Oppenheim and Ronald W. Schafer "Descrite Time Signal Processing" Prentice Hall International. Inc (1998).
8. F. J. Harris, "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform", *Proc. IEEE*, vol. 66, no. 1, pp. 51-83, Jan. 1978.
9. J. F. Kaiser, "Nonrecursive digital filter design using I0–sinh window function." in Proc. IEEE Int. Symp. Circuits and Systems (ISCAS '74), pp. 20-23, San Francisco, Calif, USA, April 1974.
10. T. Saram¨ki, "A class of window functions with nearly minimum side lobe energy for designing FIR filters," in Proc. IEEE Int. Symp. Circuits and Systems (ISCAS '89), vol. 1, pp. 359-362, Portland, Ore, USA, May 1989.
11. C. L. Dolph, "A current distribution for broadside arrays which optimizes the relationship between beamwidth and side-lobe level," Proc. IRE, vol. 34, pp. 335-348, June 1946.
12. T. Saram¨ki, "Finite impulse response filter design," in Hand book for Digital digital Processing, S. K. Mitra and J. F. Kaiser, Eds., Wiley, New York, NY, USA, 1993.
13. Rabiner, L. R., McClellan, J. H. and Parks, T., FIR Digital Filter Design Techniques Using Weighted Chebyshev Approximation, Proc. IEEE, Vol. 63, pp. 595 – 610, April 1975.
14. R. L. Streit, "A two-parameter family of weights for nonrecursive digitalfilters and antennas," IEEE Trans. Acoustics, Speech, and Signal Processing, vol. 32, no. 1, pp. 108-118, 1984.
15. A. G. Deczky, "Unispherical windows," in Proc. IEEE Int. Symp. Circuits and Systems (ISCAS 01), vol. 2, pp. 85-88, Sydney, NSW, Australia, May 2001.
16. T. Saram¨ki, "A class of window functions with nearly minimum side lobe energy for designing FIR filters," in Proc. IEEE Int. Symp. Circuits and Systems (ISCAS '89), vol. 1, pp. 359-362, Portland, Ore, USA, May 1989.
17. Stuart W. A. Bergen and Andreas Antoniou, "Design of Ultraspherical Window Functions with Prescribed Spectral Characteristics", EURASIP Journal on Applied Signal Processing 13, 2053-2065, (2004).
18. Kuldip Paliwal, and Kamil Wójcicki," Effect of Analysis Window Duration on Speech Intelligibility", IEEE signal processing letters, vol.15, pp785-788, 2008.
19. D. Klatt, "Prediction of perceived phonetic distance from critical band spectra," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*,1982, vol. 7, pp. 1278–1281.
20. Quackenbush, T. Barnwell, M. Clements, "objective measures of speech quality", Engle wood Cliffs, NJ: Printice Hall, 1988.
21. P. Krishnamoorthy, "An Overview of Subjective and Objective Quality Measures for Noisy Speech Enhancement Algorithms", IETE technical review, vol 28, issue 4, jul-aug 2011, pp 292-301.
22. Sim B, Tong Y, chang J., Tan C. (1998) 'A parametric formulation of the generalized spectral subtraction method' *IEEE. Trans. Speech Audio Process, 6(4), 328-337.*
23. Vaseghi, Saeed V., "Advanced Digital Signal Processing and Noise Reduction", Second Edition, John Wiley & Sons Ltd, July 2000.
24. P. Loizou & G. Kim" Reasons why current speech enhancement algorithms do not improve speech intelligibility and suggested solutions', IEEE transactions on Audio, Speech and Lang. Processing, Vol.19, No.1, Jan 2011.

25. Yi Hu and Philipos C. Loizou, "Subjective comparison and evaluation of speech enhancement algorithms" IEEE Trans. Speech Audio Proc.2007:49(7): 588–601.

This page is intentionally left blank