J.Alam et al. / IU-JEEE Vol. 13(1), (2013), 1589-1595

WIENER DENOISING BASED ON PERCEPTUAL FREQUENCY WEIGHTING AND NOISE SPECTRUM SHAPING

Md. Jahangir ALAM[1], Md. Faqrul Alam CHOWDHURY [2], Md. Fasiul ALAM [1]

[1]Department of EEE, KUET, Khulna, Bangladesh
[2]Department of EEE, BUET, Dhaka, Bangladesh
jahangir_kuet@yahoo.com, faqrul_buet@yahoo.com, fosi_bd@yahoo.com

Abstract: *Among the numerous noise reduction techniques that were developed over the past several decades, the Wiener filter can be considered as one of the most fundamental noise reduction approaches, which has been delineated in different forms and adopted in various applications. An important parameter of numerous speech enhancement algorithms is the a priori signal-to-noise ratio (SNR). The Wiener filter emphasizes portions of the noisy signal spectrum where SNR is high and attenuates portions of the spectrum where the SNR is low. An adaptive time varying filter can be used for whitening the noisy speech signal corrupted by narrow-band noise whereas by enhancing the signal using Perceptual frequency weighting filter (PFWF), formant regions of the noisy speech spectrum can be made less affected for a given SNR. Incorporation of PFWF and/or NSSF (Noise spectrum shaping filter) into the Weiner denoising technique improves the performance of the speech enhancement system.*

Keywords: *A priori SNR, Wiener filter, speech enhancement, perceptual weighting, noise spectreum shaping***.**

1. Introduction

The removal of additive noise from speech has been an active area of research for several decades. Numerous methods have been proposed by signal processing community. Among the most successful signal enhancement techniques have been spectral subtraction [1, 2] and Wiener filtering [3, 4]. Although these techniques improve speech quality, they generally results in random narrowband fluctuations in the residual noise called musical noise caused by randomly spaced spectral peaks that come and go in each frame and at random frequencies, which is annoying and disturbing to perception of the enhanced signal. The quality and the intelligibility of the enhanced speech signal could be improved by reducing or in better cases eliminating this kind of musical residual noise.

Many variations have been developed to cope with the musical residual noise phenomena including spectral subtraction techniques based on masking properties of the human auditory system. A number of methods have been developed to improve intelligibility by modeling several aspects of the enhancement function present in the auditory system [5]–[8]. These attractive methods use a noise masking threshold (NMT) as a crucial parameter to empirically adjust either thresholds or gain factors. This auditory system is based on the fact that the human ear cannot perceive additive noise when the noise level falls below the NMT.

In this paper, instead of empirically adjusting the parameters by the NMT for various types of noise, a perceptual frequency-weighting algorithm is derived based on the spectral envelope information of noisy input signal. The formant regions of the noisy speech spectrum will be less affected by WF for a given SNR value, if it is previously enhanced by the PFWF [17-18]. This maintains more properties of the original clean speech at formant peaks while leaves more noise at the same regions. The noise elements are considered to be masked by the speech power in the formant regions and conversely unmasked in the valleys between the formants. Therefore, the gain factor, which decides the amount of estimated noise subtracted from the noisy input signal, is controlled to be lower in formants and higher in valleys.

For some narrow-band noise or noise with evident and stable spectral peaks, an adaptive time varying filter can be used additionally to suppress the frequencies where noise energy is high. This can help to improve the speech enhancement performance of the WF by whitening the noisy speech signal especially for some signals corrupted by narrow-band noise, e.g., highway noise. The time varying noise spectrum shaping (NSSF) filter is proposed for this purposes [16]. Experimental results show that incorporating PFWF and/or NSSF filter into the Wiener denoising technique improve the performance of the speech enhancement system.

The remaining part of this paper is organized as follows: section 2 provides a description of the baseline speech enhancement system, in section 3, descriptions of the a priori SNR estimation, noise estimation and proposed

Received on: 22. 12.2010
Accepted on: 13. 01. 2013

method are given and discussion on the experimental results and conclusion are drawn in section 4 and section 5 respectively.

2. Wiener Denoising Technique

Let the distorted signal be expressed as

$$y(n) = x(n) + d(n), \quad (1)$$

where $x(n)$ is the clean signal and $d(n)$ is the additive random noise signal, uncorrelated with the original signal. If at the mth frame and kth frequency bin, $Y(m,k)$, $X(m,k)$ and $D(m,k)$ represent the spectral component of $y(n)$, $x(n)$ and $d(n)$, respectively, then the distorted signal in the transformed domain is

$$Y(m,k) = X(m,k) + D(m,k). \quad (2)$$

An estimate $\hat{X}(m,k)$ of $X(m,k)$ is given by

$$\hat{X}(m,k) = H(m,k)Y(m,k), \quad (3)$$

where $H(m,k)$ is the noise suppression gain (denoising filter), which is a function of *a priori* SNR and/or *a posteriori* SNR, given by

$$H(m,k) = \left(\frac{\xi(m,k)}{\mu + \xi(m,k)} \right)^{\beta}, \quad (4)$$

where μ is a constant, β is the order of the filter and $\xi(m,k)$ is the *a priori* SNR. If $\mu = 1$ and $\beta = ½$ then (4) corresponds to power spectrum filtering. In our case (i.e., for a Wiener filter) $\mu = \beta = 1$.

The first parameter of the noise suppression rule is the *a posteriori* SNR given by

$$\gamma(m,k) = \frac{|Y(m,k)|^2}{\Gamma_d(m,k)}, \quad (5)$$

where $\Gamma_d(m,k) = E\{|D(m,k)|^2\}$ is the noise power spectrum estimated during speech pauses. The *a priori* SNR, which is the second parameter of the noise suppression rule, is expressed as

$$\xi(m,k) = \frac{\Gamma_x(m,k)}{\Gamma_d(m,k)}, \quad (6)$$

where $\Gamma_x(m,k) = E\left\{ |X(m,k)|^2 \right\}$.

The estimation of $\xi(m,k)$ is given by the well known decision-directed approach [7] and is expressed as

$$\xi = \xi_{DD}(m,k) = \max \left(\begin{array}{l} \alpha \dfrac{|H_{DD}(m-1,k)Y(m-1,k)|^2}{\Gamma_d(m,k)} \ldots \\ + (1-\alpha) P'[\vartheta(m,k)], \xi_{\min} \end{array} \right), \quad (7)$$

where $P'[x] = x$ if $x \geq 0$ and $P'[x] = 0$ otherwise. In this paper we have chosen $\alpha = 0.98$ and $\xi_{\min} = 0.0032$ (i.e., -25 dB) by the simulations and informal listening tests.

The instantaneous SNR can be defined as [11]

$$\vartheta(m,k) = \frac{|Y(m,k)|^2}{\Gamma_d(m,k)} - 1. \quad (8)$$

The temporal-domain denoised speech is obtained with the following relation

$$x(n) = IFFT\left(|X(m,k)| . e^{j \arg(Y(m,k))} \right). \quad (9)$$

We have used noisy signal phase to obtain temporal-domain denoised speech, because the human ear is fundamentally phase deaf and perceives speech primarily based on the magnitude spectrum.

3. Proposed Method

2.1. Speech Enhancement based on PFWF

Different parts of the speech spectrum have different levels of perceptual importance on the basis of our knowledge of human perception. The difference between the clean speech spectrum and the noise speech spectrum is larger in the spectral valleys than in spectral peaks i.e., formant regions. The SNR is much higher around spectral peaks than that is near spectral valleys. Noise auditory impressions are generally provided by the parts of the spectrum with a low SNR, such as the spectral valleys. On the other hand, spectral peaks carry the most information. Therefore attenuation of spectral valleys is thought to be very effective due to reduction of speech distortion to a human listener. This encourages us to think about treating spectral peaks and spectral valleys differently with the help of a Perceptual Weighting Filter (PFWF), an IIR filter which shapes the overall spectrum of noisy speech to exploit the masking properties of the human ear [17].

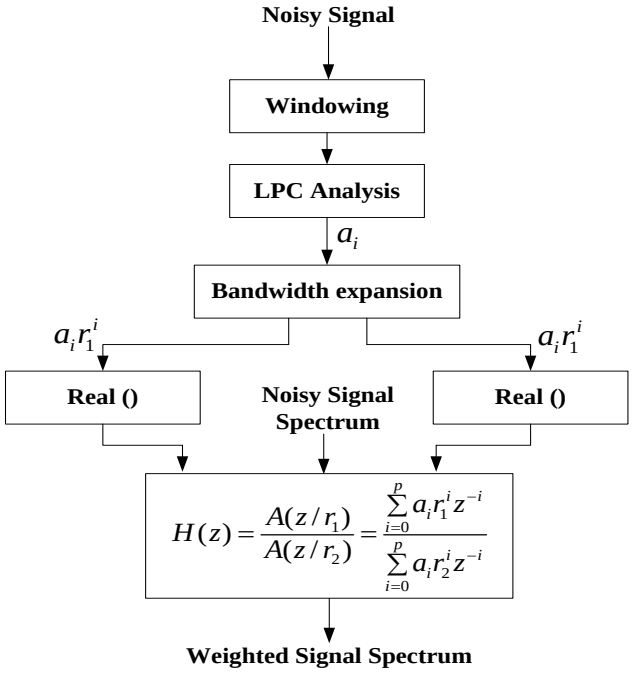


Figure 1. Block diagram of the Perceptual weighting filter (PFWF)

The PFWF used in this study is a 6th order IIR filter defined by the transfer function [18]

$$H(z)=\frac{A(z/r_1)}{A(z/r_2)}=\frac{\sum_{i=0}^{p}a_i r_1^i z^{-i}}{\sum_{i=0}^{p}a_i r_2^i z^{-i}}, \quad (10)$$

where $A(z)$ is the LPC synthesis filter, $\{a_i,\ i=1,2,.....,p,\ \ a_0=1\}$ are the LPC coefficients and r_1 and r_2 are weighting factors between 0 and 1. In this study we have chosen $r_1=0.85$ and $r_2=0.9$ by experiment. The reason for using the 6th order filter is that this filter is sufficient for calculating the spectral envelope, which consists of the 1st - 3rd formants [18].
Figure 1 shows the block diagram of the perceptual weighting filter. The LPC coefficients $\{a_i\}$ are first generated via LPC analysis of on the input noisy speech, and then scaled by multiplying the powers of the weighting factors r_1 and r_2 , so as to expand the bandwidth of the peaks and valleys respectively. The scaled LPC coefficients are considered as the forward and backward coefficients of the desired PFWF.

2.2. Speech Enhancement based on NSSF Filter

The Wiener filter (WF) algorithm introduced in section 1 has been shown to be asymptotically near optimal for the signals corrupted by additive white noise [4]. For some narrow band noise or noise with evident and stable spectral peaks, an adaptive time varying filter can be used additionally to suppress the frequencies where the noise energy is high. This can help to improve the speech enhancement performance of the WF by "whitening" the noisy speech signals, especially for some signals corrupted by narrow-band noise, e.g., highway noise. The time varying Noise Spectrum Shaping Filter (NSSF) is proposed for this purpose [17-18]. The design of NSSF has taken into account that the spectral characteristics of the noise change remarkably more slowly than those of the clean speech, as well as the bandwidth broadening effect of the noisy speech. After filtered by the NSSF, amplitudes of the noisy speech at the frequencies where the noise has spectral peaks are de-emphasized slightly.
The NSSF used in this study is a 6th order IIR filter defined by the transfer function [18]

$$H(z)=\frac{A(z/r_1)}{A(z/r_2)}=\frac{\sum_{i=0}^{p}\overline{a}_i r_1^i z^{-i}}{\sum_{i=0}^{p}\overline{a}_i r_2^i z^{-i}}, \quad (11)$$

where $A(z)$ is the LPC synthesis filter, $\{\overline{a}_i,\ i=1,2,.....,p,\ \ \overline{a}_0=1\}$ are the LPC coefficients and r_1 and r_2 are weighting factors between 0 and 1.

In this study we have chosen $r_1=0.95$ and $r_2=1$ by experiment.

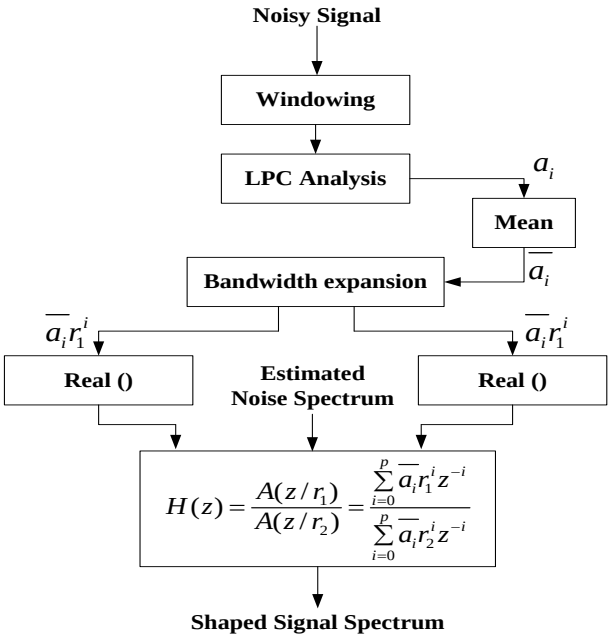


Figure 2. Block diagram of the Noise spectrum shaping filter (NSSF)

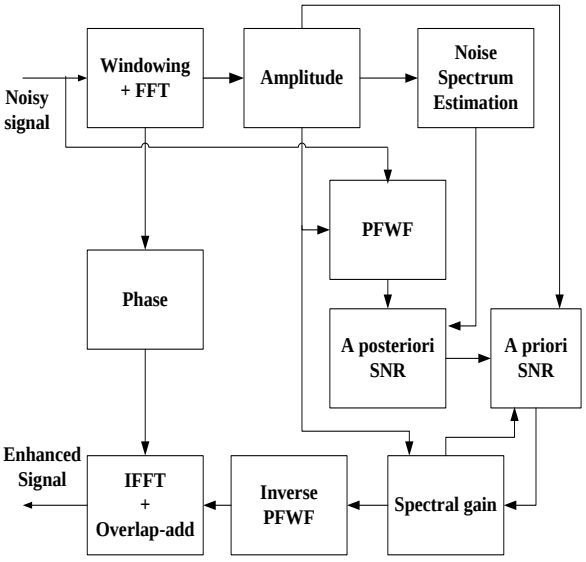


Figure 3. Block diagram showing various stages of the Wiener denoising technique based on the PFWF.

Compared with (10) for the case of the PFWF, we notice that the new coefficients $\{\overline{a}_i\}$ in (11) are average values from the adjacent analysis frames of the reference noise, rather than the instantaneous predictor coefficients from a single analysis frame of noisy speech.

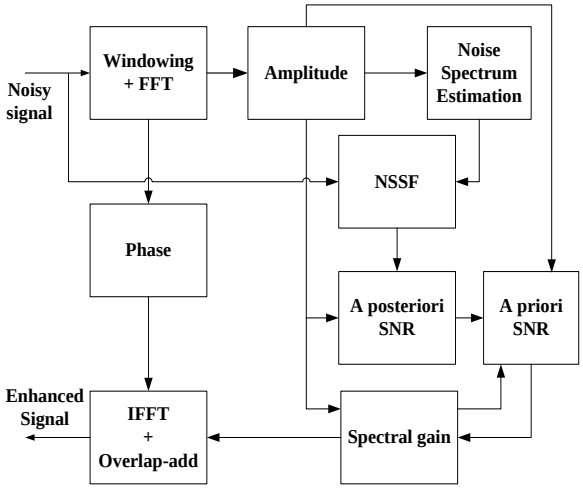


Figure 4. Block diagram showing various stages of the Wiener denoising technique that incorporates the NSSF.

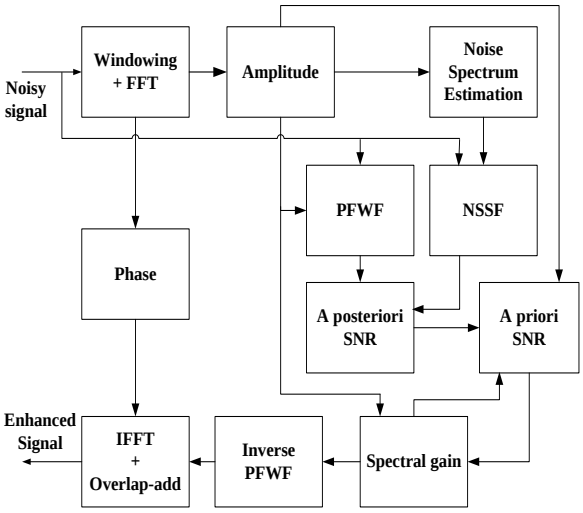


Figure 4. Block diagram showing various stages of the Wiener denoising technique that incorporates both the PFWF and NSSF.

4. Experimental Results and Discussion

In order to evaluate the performance of the proposed PFWF and NSSF-based Wiener denoising technique, we conducted extensive objective quality tests under various noisy environments. The objective quality measures used for the performance evaluation are: Segmental signal-to-noise-ratio (SSNR), Weighted spectral slope (WSS), Log likelihood ratio (LLR) and Log spectral distance (LSD). Wiener filter (WF) is chosen to compare the performance of the proposed methods. The analysis frame lenggth was chosen to be 32 msec long with an overlap of 40%; a sampling frequency of 8 kHz and a hamming window were applied.

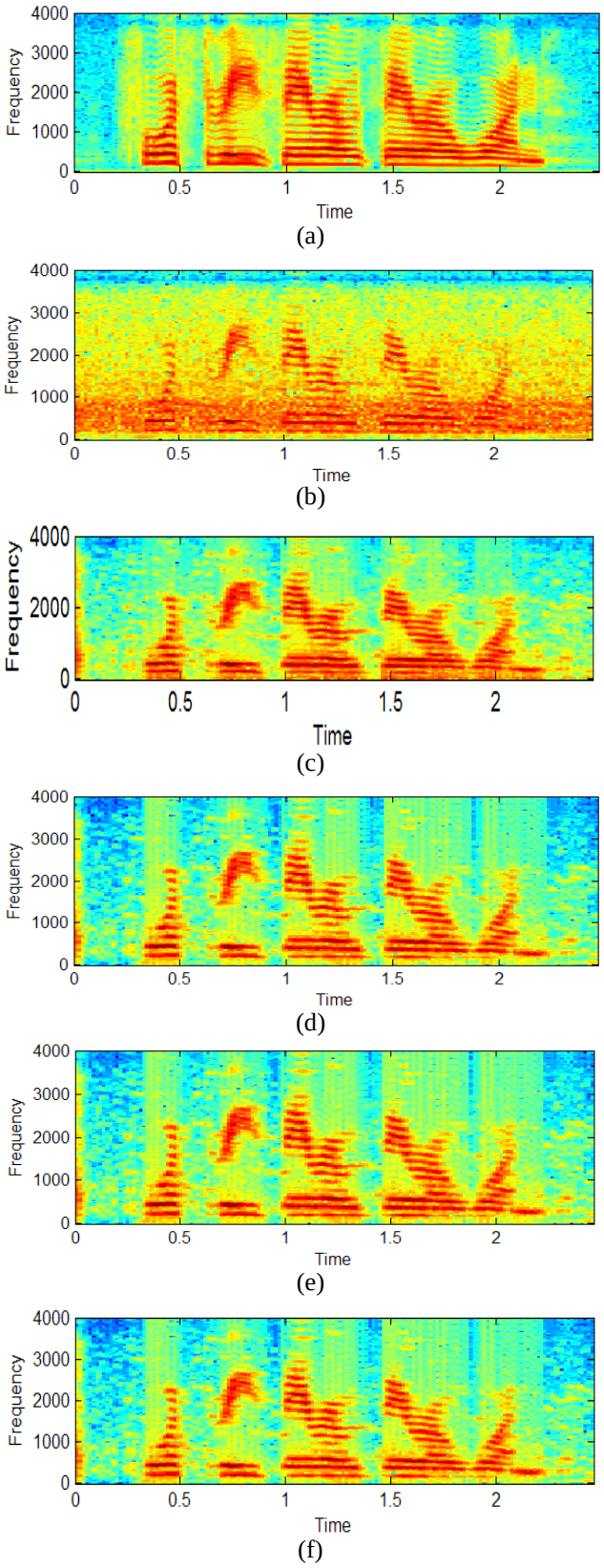
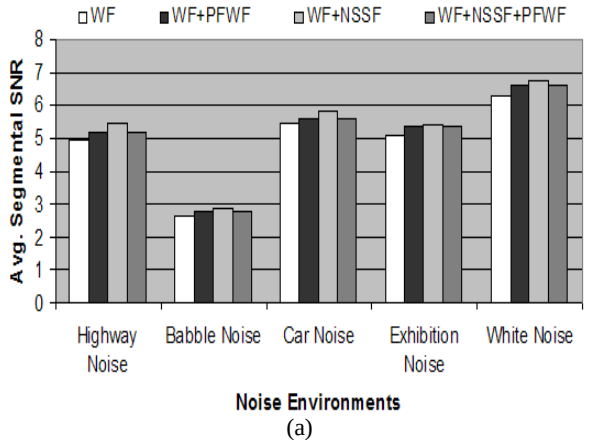
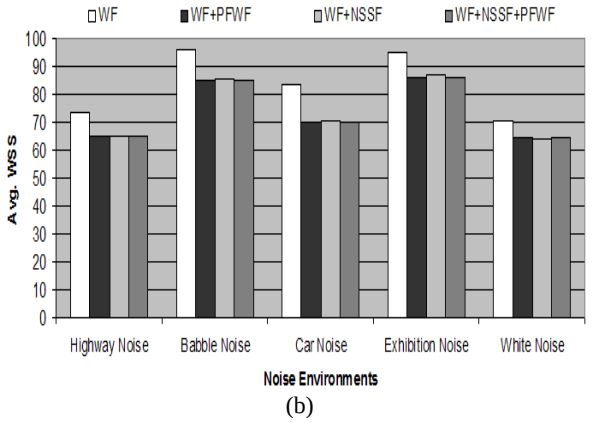


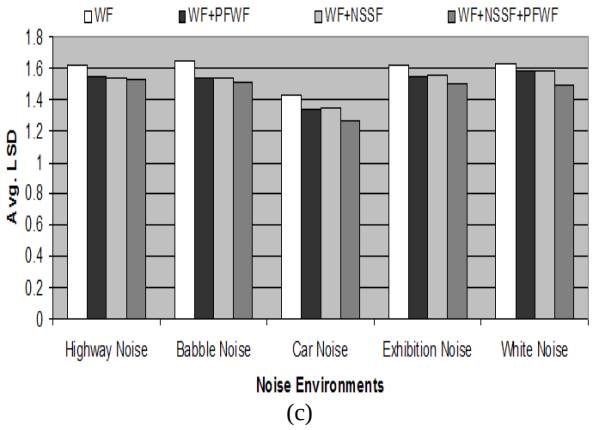
Figure 6: Speech spectrograms of (a) clean speech signal, (b) noisy speech signal corrupted with car noise, SNR = 5dB, and enhanced speech signals using (c) the Wiener denoising method, (d) PFWF-based Wiener denoising method, (e) NSSF-based Wiener denoising method, and (f) Wiener denoising method based on the PFWF and NSSF.



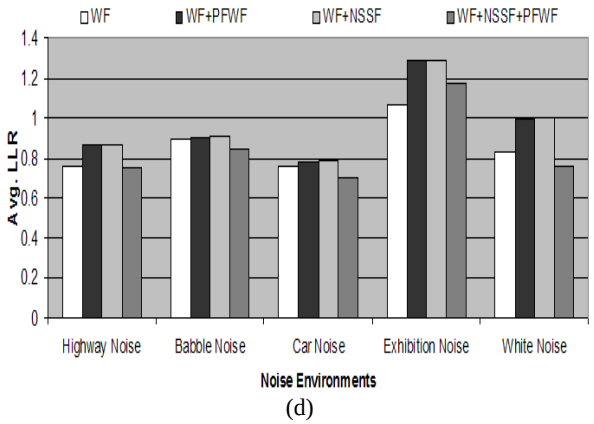
(a)



(b)



(c)



(d)

Figure 7: Histograms of the average (a) SSNR, (b) WSS, (c) LSD, and (d) LLR measures (averaged over various SNR levels, 0 - 20 dB) of the enhanced speech signals obtained using the Wiener denoising method (WF), (d) PFWF-based Wiener denoising method (WF+PFWF), NSSF-based Wiener denoising method (WF+NSSF), and Wiener denoising method based on the PFWF and NSSF (WF+PFWF+NSSF).

Table 1: Weighted spectral slope (WSS) of the enhanced signals. A lower WSS measure indicates a better speech quality.

Noise Type	Input SNR (dB)	WF	WF +PFWF	WF +NSSF	WF +PFWF +NSSF
Subway	0	103.15	**90.95**	92.13	**90.95**
	5	87.90	76.40	**75.73**	76.40
	10	71.03	**61.98**	62.99	**61.98**
	15	58.09	**51.76**	51.78	**51.76**
	20	48.82	45.41	**42.68**	45.41
Babble	0	132.76	**120.71**	121.29	**120.71**
	5	103.26	**90.09**	90.96	**90.09**
	10	78.63	69.51	**68.85**	69.51
	15	92.29	**84.18**	84.44	**84.18**
	20	73.62	**61.55**	63.89	**61.55**
Car	0	133.79	**117.48**	117.57	**117.48**
	5	92.75	**73.70**	74.72	**73.70**
	10	75.65	**59.74**	60.61	**59.74**
	15	63.10	**51.96**	52.42	**51.96**
	20	54.03	**47.72**	48.13	**47.72**
Exhibition	0	126.29	**115.32**	116.46	**115.32**
	5	102.29	**94.73**	94.79	**94.73**
	10	94.81	**86.47**	87.28	**86.47**
	15	83.05	**73.47**	74.36	**73.47**
	20	69.59	**61.73**	62.52	**61.73**
WGN	0	96.20	**85.85**	85.90	**85.85**
	5	80.60	**72.53**	72.84	**72.53**
	10	67.07	61.96	**61.70**	61.96
	15	57.50	55.00	**53.85**	55.00
	20	51.67	49.52	**47.18**	49.52

To evaluate and compare the performance of the proposed PFWF- and NSSF-based Wiener denoising techniques, we carried out simulations with the *TEST A* set of the Aurora-2 corpus [9]. Speech signals were degraded with five types of noise at global SNR levels of 0 dB, 5 dB, 10 dB, 15 dB and 20 dB. The noises were N1 (subway noise), N2 (babble noise), N3 (car noise), N4 (exhibition hall noise) and WGN (white Gaussian noise).

Figure 6 represents the spectrograms of the clean speech signal, noisy signal and enhanced speech signals obtained using the Wiener denoising technique and the proposed techniques. The speech spectrograms provide more accurate information about the residual noise and speech distortion than the corresponding time domain waveforms. We compared the spectrograms for each of the methods and confirmed a reduction of the residual noise and speech distortion. Speech spectrograms presented in Figure 6 use a Hamming window of 256 samples with 50% overlap and the noisy signals include N3 (car noise) with SNR = 5 dB. It is seen that the musical noise is almost removed for most part in figures 6 (d-f).

Tables 1-4 presents the WSS, SSNR, LSD, and LLR of the enhanced signals obtained using various speech enhancement methods at various SNR levels of different noisy environments. Fig. 7 depicts the histograms of the average SSNR, average WSS, average LSD, and average

LLR (averaged over various SNR levels 0-20 dB) versus various noise environments for the speech enhancement methods considered in this work. It is evident from the reported results that the proposed methods performed better (except in the LLR measure) than the WF alone.

Table 2: Segmental SNR (SSNR) of the enhanced signals. A higher SSNR measure indicates a better speech quality.

Noise Type	Input SNR (dB)	WF	WF+ PFWF	WF+ NSSF	WF+ PFWF +NSSF
Subway	0	-1.20	-1.04	**-0.98**	-1.04
	5	2.19	**2.44**	2.42	**2.44**
	10	5.10	5.22	**5.31**	5.22
	15	8.03	8.33	**8.47**	8.33
	20	10.63	10.94	**11.88**	10.94
Babble	0	-2.42	**-2.40**	-2.44	**-2.40**
	5	0.05	**0.19**	0.17	0.19
	10	2.52	2.65	**2.88**	2.65
	15	5.24	5.57	**5.66**	5.57
	20	7.60	7.72	**8.02**	7.72
Car	0	-0.49	**-0.29**	-0.35	**-0.29**
	5	2.57	2.68	**2.73**	2.68
	10	5.30	5.47	**5.53**	5.47
	15	8.38	8.46	**8.87**	8.46
	20	11.42	11.59	**12.39**	11.59
Exhibition	0	0.72	**0.73**	0.70	**0.73**
	5	2.17	2.24	**2.27**	2.24
	10	5.23	5.42	**5.44**	5.42
	15	7.38	7.70	**7.82**	7.70
	20	10.02	10.54	**10.67**	10.54
WGN	0	1.13	**1.46**	1.45	**1.46**
	5	3.95	3.79	**3.84**	3.79
	10	6.08	**6.63**	6.57	**6.63**
	15	8.47	9.31	**9.37**	9.31
	20	11.56	11.82	**12.56**	11.82

Table 3: Log Spectral distance (LSD) measures of the enhanced signals. A lower LSD measure indicates a better speech quality.

Noise Type	Input SNR (dB)	WF	WF+ PFWF	WF +NSSF	WF +PFWF +NSSF
Subway	0	2.11	2.03	**2.02**	2.08
	5	1.68	1.58	1.58	**1.55**
	10	1.69	1.62	1.62	**1.57**
	15	1.39	1.34	1.34	**1.28**
	20	1.17	**1.14**	1.12	1.17
Babble	0	2.08	1.98	1.97	**1.97**
	5	1.74	1.66	1.66	**1.62**
	10	1.53	1.42	1.42	**1.39**
	15	1.48	1.39	1.40	**1.36**
	20	1.35	1.21	1.22	**1.18**
Car	0	1.70	1.61	1.62	**1.57**
	5	1.49	1.40	1.40	**1.35**
	10	1.39	1.29	1.30	**1.22**
	15	1.28	1.20	1.22	**1.11**
	20	1.26	1.18	1.19	**1.05**
Exhibition	0	1.82	1.75	**1.75**	1.75
	5	1.69	1.63	**1.63**	1.64
	10	1.75	1.66	1.67	**1.64**
	15	1.43	1.38	1.39	**1.28**
	20	1.38	1.33	1.33	**1.20**
WGN	0	2.05	2.00	2.00	**1.94**
	5	1.83	1.78	1.78	**1.69**
	10	1.59	1.54	1.54	**1.45**
	15	1.38	1.35	1.35	**1.25**
	20	1.25	1.23	1.22	**1.11**

Table 4: Log likelihood ratio (LLR) measures of the enhanced signals. A lower LSD measure indicates a better speech quality.

Noise Type	Input SNR (dB)	WF	WF+ PFF	WF +NSSF	WF+ PFWF +NSSF
Subway	0	**1.33**	1.46	1.48	1.39
	5	0.84	0.90	0.90	**0.76**
	10	0.72	0.86	0.86	**0.70**
	15	0.53	0.62	0.60	**0.49**
	20	**0.35**	0.46	0.46	0.41
Babble	0	1.18	1.20	1.21	**1.12**
	5	0.93	0.89	0.89	**0.81**
	10	0.77	0.80	0.82	**0.75**
	15	0.85	0.87	0.88	**0.81**
	20	0.71	0.71	0.72	**0.69**
Car	0	**1.14**	1.23	1.24	1.16
	5	0.86	0.92	0.91	**0.81**
	10	0.76	0.71	0.74	**0.63**
	15	0.53	0.52	0.52	**0.47**
	20	0.50	0.50	0.52	**0.40**
Exhibition	0	**1.32**	1.55	1.54	1.45
	5	**1.23**	1.48	1.48	1.44
	10	1.28	1.46	1.46	**1.21**
	15	**0.79**	1.03	1.03	0.93
	20	**0.68**	0.90	0.90	0.81
WGN	0	1.16	1.36	1.36	**1.12**
	5	.925	1.17	1.17	**0.90**
	10	.803	0.96	0.96	**0.73**
	15	.637	0.80	0.81	**0.57**
	20	.617	0.67	0.67	**0.48**

5. Conclusions

The aim of our study is to improve the performance of the Wiener denoising technique. In this paper we have presented PFWF and NSSF-based Wiener denoising techniques that would maximize noise reduction while minimizing speech distortion. Performance evaluations of the proposed approaches are carried out using four objective quality measures, namely, SSNR, WSS, LSD and LLR. Simulation results and plotted speech spectrograms show that the proposed algorithms give better performance for speech enhancement in various noisy environments than that of the conventional Wiener denoising method.

6. References

[1] Boll, S. F., "Suppression of acoustic noise in speech using spectral subtraction", *IEEE Trans. Acoustics, Speech, Signal Processing*, vol. 27, pp. 113–120, Apr. 1979.

[2] M. Berouti, R. Schwartz, and J. Makhoul, "Enhancement of speech corrupted by acoustic noise", in Proc. *IEEE Int. Conf. on Acoustics, Speech, Signal Processing*, vol. 1, (Washington, DC), pp. 208–211, Apr. 1979.

[3] H. L. V. Trees, Detection, Estimation, and Modulation: Part I - Detection, Estimation and Linear Modulation Theory. John Wiley and Sons, Inc., 1st ed., 1968.

[4] T.F. Quatieri and R.B. Dunn, "Speech enhancement based on auditory spectral change," *IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, vol. 1, pp. 257-260, Orlando, FL, USA, 2002.

[5] Y.M. Cheng and D. O'Shaughnessy, "Speech enhancement based conceptually on auditory evidence," *IEEE Trans. Signal Processing*, vol.39, no.9, pp.1943–1954, 1991.

[6] D. Tsoukalas, M. Paraskevas, and J. Mourjopoulos, "Speech enhancement using psychoacoustic criteria," *IEEE ICASSP*, pp.359–362, Minneapolis, MN, 1993.

[7] N. Virag, "Single channel speech enhancement based on masking properties of the human auditory system," *IEEE Trans. Speech Audio Processing*, vol.7, no.2, pp.126–137, 1999.

[8] M. R. Schroeder, B. S. Atal, and J. L. Hall, "Optimizing digital speech coders by exploiting masking properties of the human ear," *J. Acoust. Soc. Am*., vol. 66, pp. 1647–1652, Dec. 1979.

[9] E. Zwicker and H. Fastl, *Psychoacoustics: Facts and Models*. Springer-Verlag, 2nd ed., 1999.

[10] Y. Ephraim and D. Mallah, "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimation," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. ASSP-32, no. 6, pp. 1109-1121, Dec. 1984.

[11] O. Cappe, "Elimination of the musical noise phenomenon with the Ephraim and Malah noise suppressor," *IEEE Trans. Speech and Audio Processing*, vol. 2, no. 1, pp. 345–349, April 1994.

[12] Yi Hu and Philipos C. Loizou, "Evaluation of Objective Quality Measures for Speech Enhancement," IEEE Trans. on Audio, Speech and Language Processing, vol. 16, No. 1, pp. 229-238, January 2008.

[13] Quackenbush S., T. Barnwell and M. Clements, *Objective Measures of Speech Quality*, Englewood Cliffs, NJ, USA, Prentice Hall, 1988.

[14] J. H. L. Hansen and B. L. Pellom, "An effective evaluation protocol for speech enhancement algorithms", in Proc. ICSLP, vol. 7, Sydney, Australia, 1998, pp. 2819–2822.

[15] H. Hirsch and D. Pearce, "The Aurora experimental framework for the performance evaluation of speech recognition systems under noisy environments", ISCA ITRW ASR, September 2000.

[16] H. Tolba, Z. Li, and D. O'Shaughnessy, "Robust automatic speech recognition using a perceptually-based optimal spectral amplitude estimator speech enhancement algorithm in various low-SNR environments", INTERSPEECH - Eurospeech, pp. 937–940, September, 2005.

[17] Z. Li, H. Tolba, and D. O'Shaughnessy, "Robust automatic speech recognition using an optimal spectral amplitude estimator algorithm in low-SNR car environments", INTERSPEECH - ICSLP, pp. 2041–2044, October 2004.

[18] Zili Li, "Distributed Speech Recognition and Speech Reconstruction System in Noisy Environments", Ph.D. thesis, Dept. of Telecomm., INRS-EMT, Univ. of Quebec, Montreal, Canada, 2007.