Int J Speech Technol (2014) 17:53–64
DOI 10.1007/s10772-013-9205-5

Speech enhancement with an adaptive Wiener filter

Marwa A. Abd El-Fattah · Moawad I. Dessouky · Alaa M. Abbas · Salaheldin M. Diab · El-Sayed M. El-Rabaie · Waleed Al-Nuaimy · Saleh A. Alshebeili · Fathi E. Abd El-samie

Received: 12 February 2013 / Accepted: 1 August 2013 / Published online: 28 August 2013

Abstract This paper proposes an adaptive Wiener filtering method for speech enhancement. This method depends on the adaptation of the filter transfer function from sample to sample based on the speech signal statistics; the local mean and the local variance. It is implemented in the time domain rather than in the frequency domain to accommodate for the time-varying nature of the speech signals. The proposed method is compared to the traditional frequency-domain Wiener filtering, spectral subtraction and wavelet denoising methods using different speech quality metrics. The simulation results reveal the superiority of the proposed Wiener filtering method in the case of Additive White Gaussian Noise (AWGN) as well as colored noise.

Keywords Speech enhancement · Wiener filter · Spectral subtraction · Wavelet denoising

M.A. Abd El-Fattah · M.I. Dessouky · A.M. Abbas · S.M. Diab · S.M. El-Rabaie · F.E. Abd El-samie (✉)
Department of Electronics and Electrical Communications, Faculty of Electronic Engineering, Menoufia University, Menouf, 32952, Egypt
e-mail: fathi_sayed@yahoo.com

M.A. Abd El-Fattah
e-mail: maro_zizo2010@hotmail.com

M.I. Dessouky
e-mail: dr_moawad@yahoo.com

A.M. Abbas
e-mail: aladin_abbas@yahoo.com

S.M. Diab
e-mail: dr_salah_diab@yahoo.com

S.M. El-Rabaie
e-mail: srabie1@yahoo.com

W. Al-Nuaimy
Department of Electrical Engineering and Electronics, The University of Liverpool, Liverpool L69 3GJ, UK
e-mail: wax@liv.ac.uk

S.A. Alshebeili
Electrical Engineering Department, KACST-TIC in Radio Frequency and Photonics for the e-Society (RFTONICS), King Saud University, Riyadh, Kingdom of Saudi Arabia
e-mail: dsaleh@ksu.edu.sa

F.E. Abd El-samie
KACST-TIC in Radio Frequency and Photonics for the e-Society (RFTONICS), King Saud University, Riyadh, Kingdom of Saudi Arabia

1 Introduction

Speech signals are the most widely used signals between humans, to convey messages. Hence, the researchers gave a large attention to speech processing and proposed lots of researches in speech and hearing sciences (Ying et al. 2008; Wu et al. 2000). Speech processing systems are used in a wide variety of applications such as speech coding for communications, speech recognition for automatic information systems and speech pre-processing for aids to hearing impaired persons. These systems are designed under the assumption that corruptive background noises are absent. In a noisy environment, speech enhancement is suggested to improve the performance of these systems.

Speech enhancement is a word used to describe algorithms, which can be used to improve the quality, decrease the hearing fatigue of noisy speech, increase intelligibility, and improve the performance of the voice communication systems (Quatieri 2002). On the other hand, no speech enhancement systems can improve both speech quality and intelligibility. Basically, speech intelligibility can be viewed as an aspect of quality, since high-quality speech always gives good intelligibility, and unintelligible speech would not be classified as having high quality. In many previous researches, speech enhancement increases the quality but reduces the intelligibility (Kusumoto et al. 2005). Several methods have been proposed for this purpose like the spectral subtraction method, the signal subspace method, the Wiener filtering method, and the wavelet denoising method (Boll 1979; Berouti et al. 1979; Ephriam and Van Trees 1993; Haykin 1996; Lim and Oppenheim 1978). The improvement of the speech Signal-to-Noise Ratio (SNR) is the goal of most techniques.

Spectral subtraction is one of the traditional methods used for enhancing speech degraded by additive stationary background noise (Deller et al. 2000; Ghanbari and Karami 2004; Ghanbari et al. 2004). It can be categorized as a non-parametric method, which needs an estimate of the noise spectrum. A common problem for spectral subtraction method is the characteristic of the residual noise called musical noise. Spectral subtraction also does not attenuate noise sufficiently during the silence period.

Wiener filtering (Boll 1979) is an alternative method to spectral subtraction for enhancing the speech signal. The Wiener filter is a linear filter employed to recover the original speech signal from the noisy signal by minimizing the Mean Square Error (MSE) between the estimated signal and the original one. Wavelet de-noising (Vaseghi 2000; Manikandan 2006; Shao and Chang 2005) is another method based on wavelet decomposition of the noisy signal and thresholding in the wavelet domain to remove noise. The wavelet transform is exploited to decompose the noisy signal into sub-bands, and the reduction of noise is performed by either hard or soft thresholding. The disadvantage of this method is that it tends to distort some useful components of the original speech as well. There are also some other methods that adapt the statistical model of a recognizer to identify noise characteristics before speech enhancement (Boll 1979).

In this paper, we present an adaptive Wiener filtering method for speech enhancement. This method considers the local statistics of the speech signal, and it is carried out in the time domain. The rest of the paper is organized as follows. In Sect. 2, a review of the spectral subtraction method is presented. In Sect. 3, the traditional frequency-domain Wiener filtering method is revisited. In Sect. 4, the wavelet hard and soft thresholding is illustrated. Section 5 presents the suggested adaptive Wiener filtering method for speech enhancement. In Sect. 6, a comparison study between the proposed method and the traditional methods is presented. Finally, the concluding remarks are given in Sect. 7.

2 Spectral subtraction

Spectral subtraction is one of earliest and important methods used for speech enhancement. It is based on subtracting the noise spectrum from the noisy signal spectrum to obtain an estimate of the clean signal spectrum, and then reconstructing the signal from the estimated spectrum (Lim and Oppenheim 1978; Ghanbari and Karami 2004; Ghanbari et al. 2004; Handel 2007). Different techniques are used for calculating the spectrum of a speech signal. The signal can be passed through a filter bank to produce a series of sub-band signals, or a spectral transformer such as the Discrete Fourier Transform (DFT) can be applied to the signal. Let $x(n)$ be a noisy signal (Handel 2007):

$$x(n) = s(n) + v(n) \tag{1}$$

where $s(n)$ is the clean signal, and $v(n)$ is the noise. Assume that the clean signal and the noise are uncorrelated. The spectral subtraction method is applied as follows (Handel 2007):

$$\hat{S}(\omega) = \left(\left|X(\omega)\right| - \left|\hat{N}(\omega)\right|\right) \exp\left(j\angle X(\omega)\right) \tag{2}$$

where $\hat{S}(\omega)$ is the estimated short term spectral magnitude of the clean signal, $\hat{N}(\omega)$ is the estimated noise magnitude spectrum, and $X(\omega)$ is the noisy observation magnitude spectrum. The estimated time-domain speech signal is obtained as the inverse Fourier transform of $\hat{S}(\omega)$.

From Eq. (2), the estimation of the clean speech signal is based on obtaining an accurate spectral estimate of the noise. In fact, the noise spectrum is not available; an averaged estimate of the noise is different from the actual noise contents in the instantaneous speech spectrum. By subtracting a smoothed estimate of the noise spectrum, some sinusoidal energy appears at various frequencies in the estimated speech. This energy represents musical tones, which affect the quality of the enhanced speech, and are impossible to remove completely (Boll 1990).

The degradation of the intelligibility of the enhanced speech is the main limitation of the spectral subtraction technique, especially at low SNR levels (Boll 1990). Efficiency could be improved by improving the performance of the speech silence detection algorithm.

3 Wiener filter

Wiener filtering is another method to suppress noise in speech signals. It is based on minimizing the MSE between the estimated signal magnitude spectrum $\hat{S}(\omega)$ and the original signal magnitude spectrum $S(\omega)$. The formulation of the optimal wiener filter is as follows (Lim and Oppenheim 1979):

$$H(\omega) = \frac{S_s(\omega)}{S_s(\omega) + S_n(\omega)} \tag{3}$$

where $S_s(\omega)$ and $S_n(\omega)$ represent the estimated power spectra of the noise-free signal and the background noise, which are assumed uncorrelated and stationary. After calculating the transfer function of the Wiener filter, the speech signal is recovered through (Lim and Oppenheim 1979):

$$\hat{S}(\omega) = X(\omega)H(\omega) \tag{4}$$

In a modified Wiener filter, an adjustable parameter α has been included in the following generalized form Lim and Oppenheim (1979):

$$H(\omega) = \left(\frac{S_s(\omega)}{S_s(\omega) + \beta S_n(\omega)}\right)^{\alpha} \tag{5}$$

where β is the noise suppression factor.

4 Wavelet denoising

As wavelet analysis has its basis emulating the front-end auditory periphery, efforts have been made to take advantage of this signal-processing tool for speech enhancement. The mostly used approach is based on the non-linear thresholding of the wavelet coefficients, which bridges the multi-resolution analysis and non-linear filtering. The Thresholding process is a denoising process.

The wavelet transform decomposes the noisy speech signal into two component coefficients; approximation or low-pass coefficients and details or high-pass coefficients. Each of the approximation or the detail components has half the length of the original speech signal. Most of the speech signal energy is concentrated in the approximation component (Sheikhzadeh 2001). So, the effect of noise on the approximation component is small and on the detail component is large. If a thresholding process is performed on the detail component, it reduces the noise significantly, leaving the signal energy unaffected.

4.1 Thresholding principles

Assume that the wavelet transform of the noisy signal $x(n)$ in Eq. (1) is given by X. Thresholding is performed on the detail components of X. There are generally two ways of thresholding; hard thresholding and soft thresholding. Hard thresholding is defined as follows (Johnstone and Silverman 1997):

$$Thr_{Hard}(X,T) = \begin{cases} X & |X| > T \\ 0 & |X| < T \end{cases} \tag{6}$$

where T is the selected threshold value.

And soft thresholding is defined as follows (Johnstone and Silverman 1997):

$$Thr_{soft}(X,T) = \begin{cases} \operatorname{sgn}(X)(|X| - T) & |X| > T \\ 0 & |X| < T \end{cases} \tag{7}$$

where

$$\operatorname{sgn}(x) = \begin{cases} -1 & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ 1 & \text{if } x > 0 \end{cases} \tag{8}$$

Both of these two methods suffer from distortion of the speech, because they set coefficients that may carry some useful information to zero, resulting in observable sharp time frequency discontinuities in the speech spectrogram.

4.2 How to choose the threshold

The choice of the threshold value can be made in several ways. Donoho derived the following formula based on an AWGN assumption (Johnstone and Silverman 1997):

$$T = \sigma\sqrt{2\log(N)} \tag{9}$$

where T is the threshold value, N is the length of the noisy signal, and $\sigma = MAD/0.6745$, with MAD denoting the absolute median estimated on the first scale of the wavelet coefficients.

Johnstone and Silverman proposed a level dependent thresholding method to deal with correlated noise, where for each frequency interval, the threshold is proportional to the standard deviation of the noise in that interval (Bahoura and Rouat 2001; Shao and Chang 2005):

$$\lambda_a = \sigma_a\sqrt{2\log(N_a)} \tag{10}$$

with $\sigma_a = MAD_a/0.6745$, N_a is the number of samples in scale a, and MAD_a is the absolute median estimated at scale a.

Although the wavelet denoising method does not require a speech or noise model, and can be applied to a broader class of signals, merely a general thresholding on the wavelet coefficients does not guarantee a good performance.

5 Adaptive Wiener filtering

This proposed adaptive Wiener filter benefits from the varying local statistics of the speech signal. A block diagram of the adaptive Wiener filtering method is illustrated in Fig. 1. In the filtering process, the estimated local mean m_x and local variance σ_x^2 of the signal $x(n)$ are exploited.

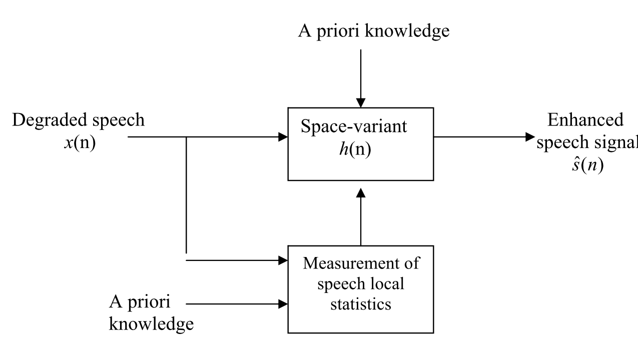


Fig. 1 Speech enhancement with adaptive Wiener filter

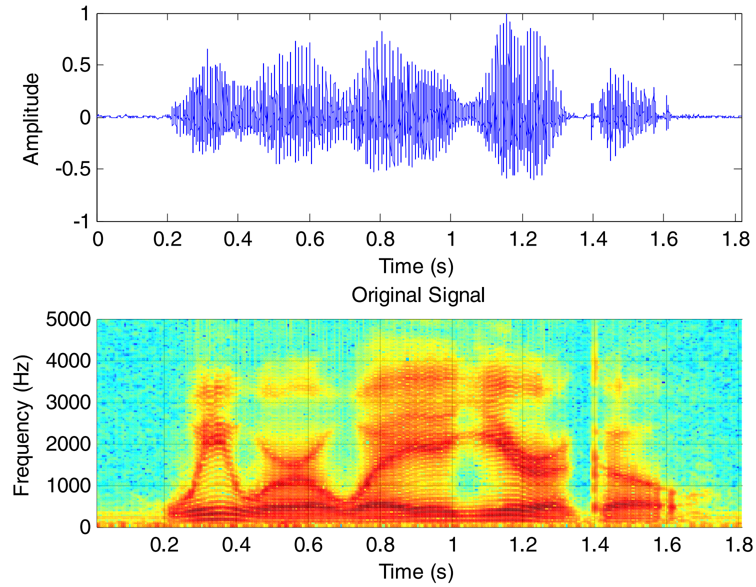


Fig. 2 Time-domain waveform and spectrogram of the clean male signal

It is assumed that the additive noise $v(n)$ is of zero mean and has a white nature with variance σ_v^2. Thus, the power spectrum $P_v(\omega)$ can be approximated by:

$$P_v(\omega) = \sigma_v^2 \tag{11}$$

Consider a small segment of the speech signal in which the signal $x(n)$ is assumed to be stationary, The signal $x(n)$ can be modeled by:

$$x(n) = m_x + \sigma_x w(n) \tag{12}$$

where m_x and σ_x are the local mean and standard deviation of $x(n)$. $w(n)$ is a unit variance noise with a zero mean. So, the mean of the original signal m_s is equal to m_x.

Within this small segment of speech, the Wiener filter transfer function can be approximated by:

$$H(\omega) = \frac{P_s(\omega)}{P_s(\omega) + P_v(\omega)} = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_v^2} \tag{13}$$

From Eq. (13), the impulse response of the wiener filter can be obtained by:

$$h(n) = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_v^2}\delta(n) \tag{14}$$

From Eq. (14), the enhanced speech signal $\hat{s}(n)$ within this local segment can be expressed as:

$$\begin{aligned} \hat{s}(n) &= m_s + \big(x(n) - m_s\big)\frac{\sigma_s^2}{\sigma_s^2 + \sigma_v^2}\delta(n) \\ &= m_s + \frac{\sigma_s^2}{\sigma_s^2 + \sigma_v^2}\big(x(n) - m_s\big) \end{aligned} \tag{15}$$

If it is assumed that m_s and σ_s are updated at each sample, we can say that:

$$\hat{s}(n) = m_s(n) + \frac{\sigma_s^2(n)}{\sigma_s^2(n) + \sigma_v^2}\big(x(n) - m_s(n)\big) \tag{16}$$

We can estimate $m_s(n)$ in Eq. (16) from $x(n)$ as follows:

$$\hat{m}_s(n) = \frac{1}{(2M+1)}\sum_{k=n-M}^{n+M} x(k) \tag{17}$$

where $(2M+1)$ is the number of samples in the short segment used in the estimation.

To measure the local signal statistics in the system of Fig. 2, we have $\sigma_x^2 = \sigma_s^2 + \sigma_v^2$, and hence,

$$\hat{\sigma}_s^2(n) = \begin{cases} \hat{\sigma}_x^2(n) - \hat{\sigma}_v^2, & \text{if } \hat{\sigma}_x^2(n) > \hat{\sigma}_v^2 \\ 0, & \text{otherwise} \end{cases} \tag{18}$$

where

$$\hat{\sigma}_x^2(n) = \frac{1}{(2M+1)}\sum_{k=n-M}^{n+M}\big(x(k) - \hat{m}_x(n)\big)^2 \tag{19}$$

By this proposed method, we guarantee that the filter transfer function is adapted from sample to sample based on the speech signal local statistics.

6 Simulation results

For the evaluation purpose, we have used a speech signal for the sentence "We were away year ago" for a male and for a female. We used speech quality metrics such as the SNR, segmental SNR (SNRseg), Log-Likelihood Ratio (LLR) and Spectral Distortion (SD). We first began with the male signal and added AWGN to it with SNRs of −5 and 5 dB. The results of all enhancement methods explained above on the male speech signal for SNR of 5 dB are shown in Figs. 2 to 8.

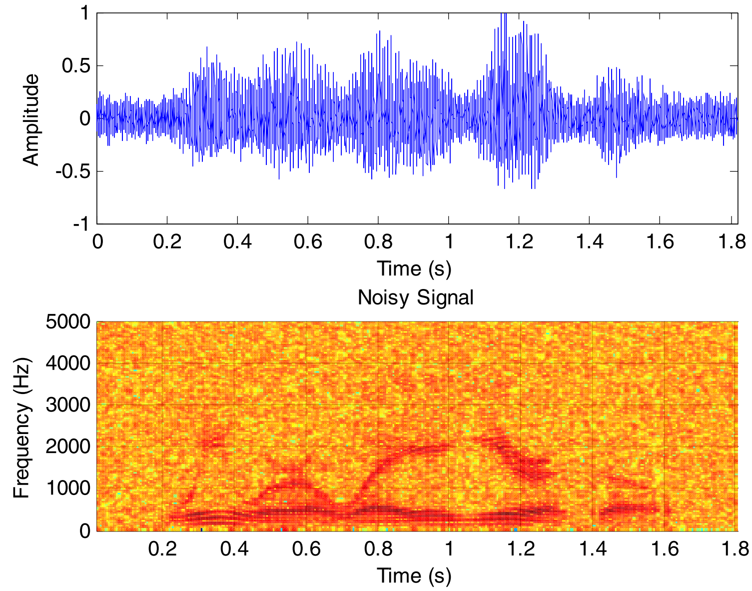


Fig. 3 Time-domain waveform and spectrogram of the noisy male signal with AWGN at SNR = 5 dB

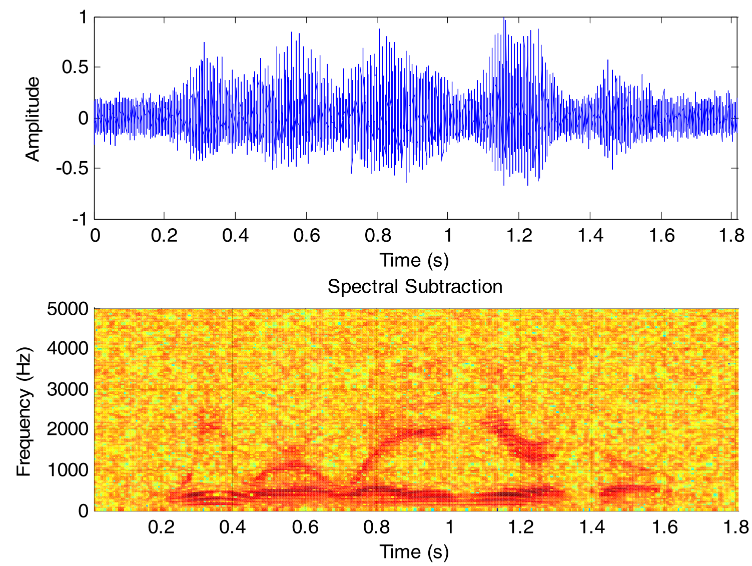


Fig. 4 Time-domain waveform and spectrogram of the enhanced signal using the spectral subtraction method. SNR = 5.0439 dB, SNR_{seg} = 5.0164 dB, LLR = 0.2336, SD = 8.4721 dB

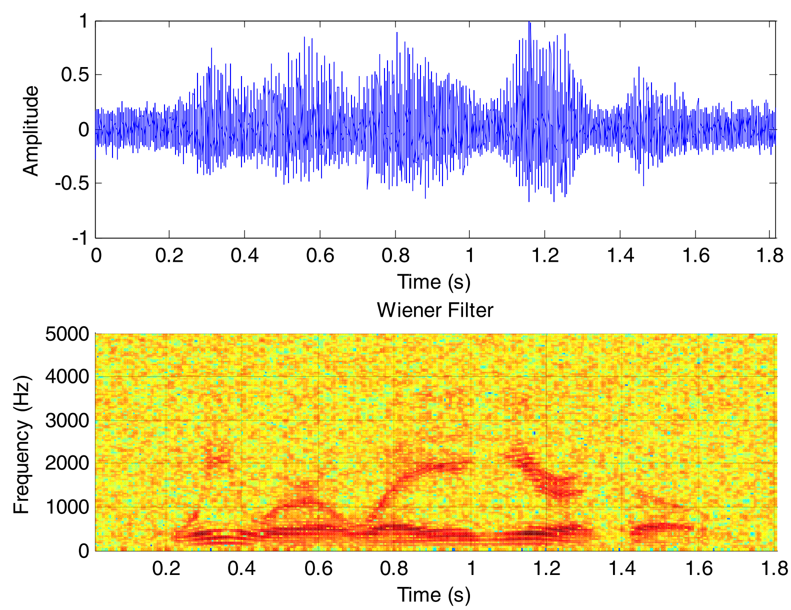


Fig. 5 Time-domain waveform and spectrogram of the enhanced signal using the Wiener filtering method. SNR = 4.9880 dB, SNR_{seg} = 4.9604 dB, LLR = 0.2383, SD = 8.5090 dB

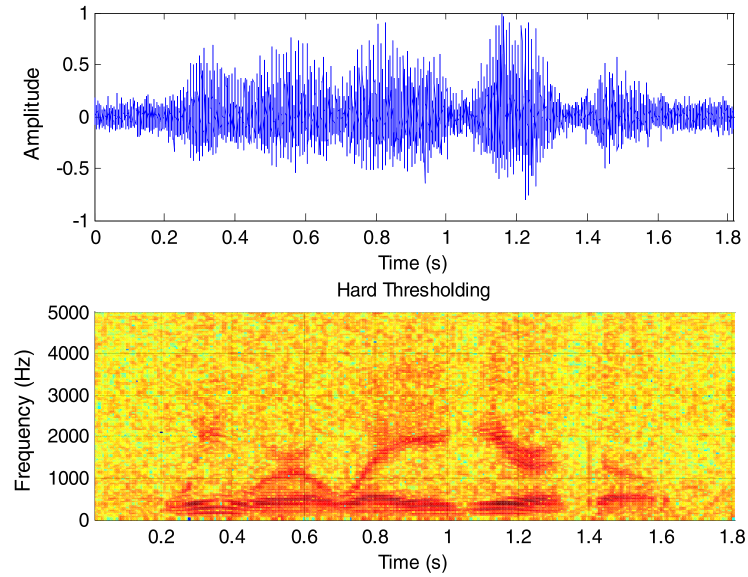


Fig. 6 Time-domain waveform and spectrogram of the enhanced male signal using the hard thresholding technique. SNR = 6.5002 dB, $\mathrm{SNR_{seg}}$ = 6.4605 dB, LLR = 0.1945, SD = 7.6423 dB

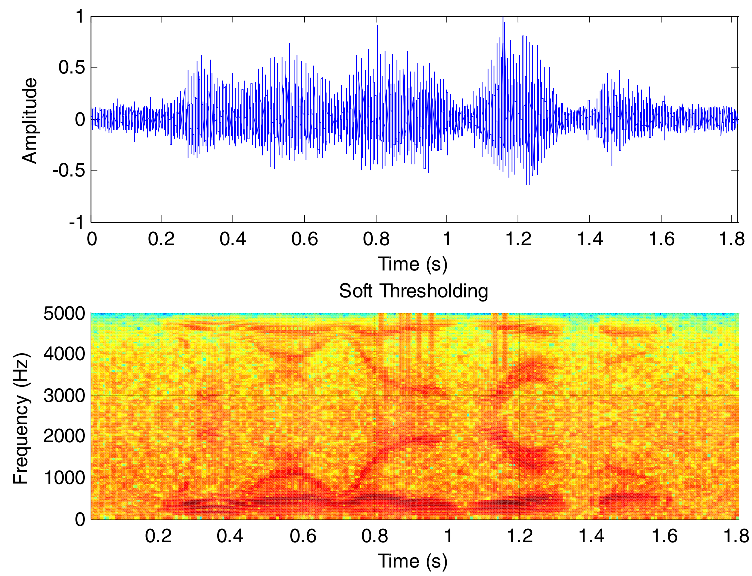


Fig. 7 Time-domain waveform and spectrogram of the enhanced male signal using the soft thresholding technique. SNR = 6.4884 dB, $\mathrm{SNR_{seg}}$ = 6.4506 dB, LLR = 0.1942, SD = 7.6463 dB

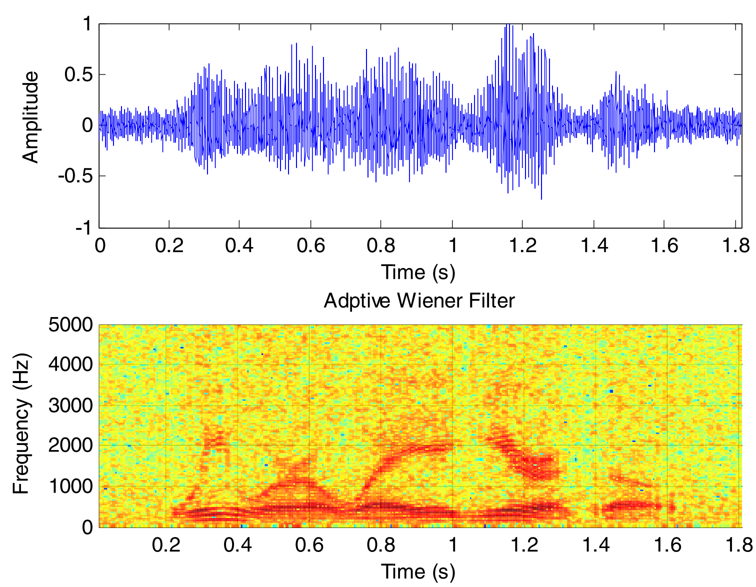


Fig. 8 Time-domain waveform and spectrogram of the enhanced male signal using the adaptive Wiener Filtering method. SNR = 6.8726 dB, $\mathrm{SNR_{seg}}$ = 6.8423 dB, LLR = 0.1609, SD = 7.3006 dB

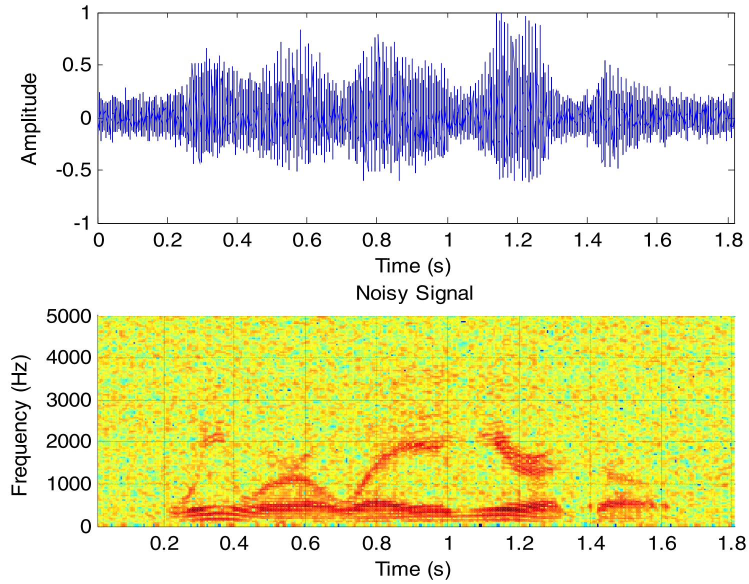


Fig. 9 Time-domain waveform and spectrogram of the noisy male signal with colored noise at SNR = 5 dB

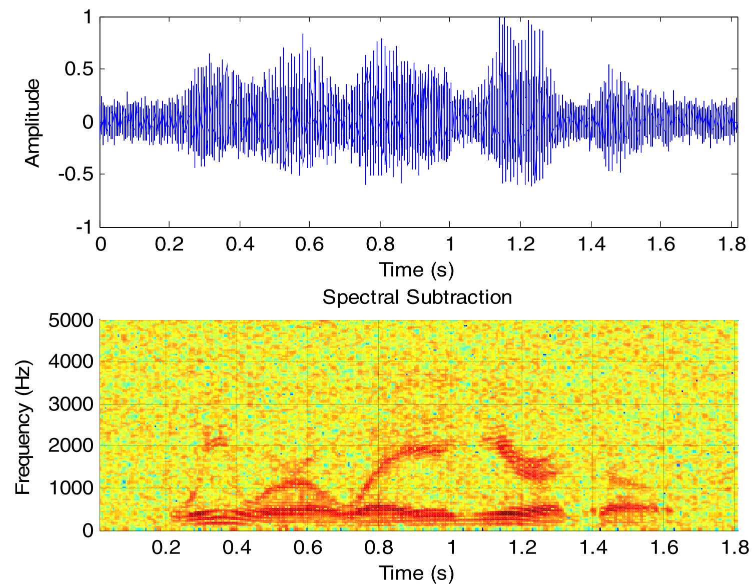


Fig. 10 Time-domain waveform and spectrogram of the enhanced male signal using the spectral subtraction method. SNR = 5.0571 dB, $\mathrm{SNR_{seg}}$ = 5.0235 dB, LLR = 0.1340, SD = 8.5551 dB

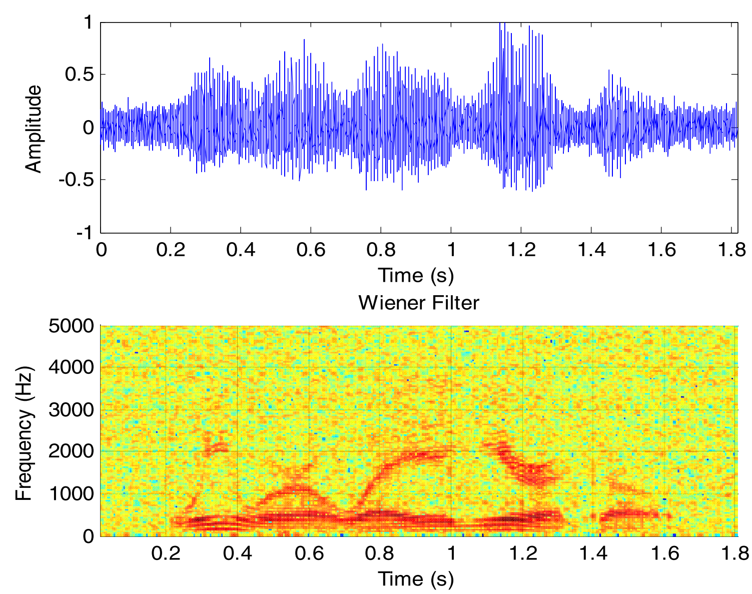


Fig. 11 Time-domain waveform and spectrogram of the enhanced male signal using the Wiener filtering method. SNR = 5.0005 dB, $\mathrm{SNR_{seg}}$ = 4.9668 dB, LLR = 0.1333, SD = 8.5926 dB

Fig. 12 Time-domain waveform and spectrogram of the enhanced male signal using the hard thresholding method. SNR = 5.5011 dB, SNR_{seg} = 5.4711 dB, LLR = 0.3625, SD = 8.1469 dB

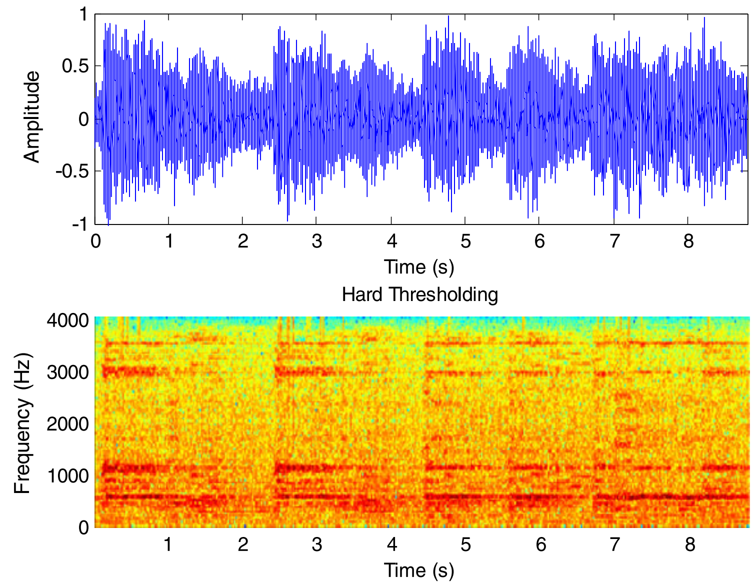


Fig. 13 Time-domain waveform and spectrogram of the enhanced male signal using the soft thresholding method. SNR = 6.5776 dB, SNR_{seg} = 6.5449 dB, LLR = 0.2606, SD = 7.5809 dB

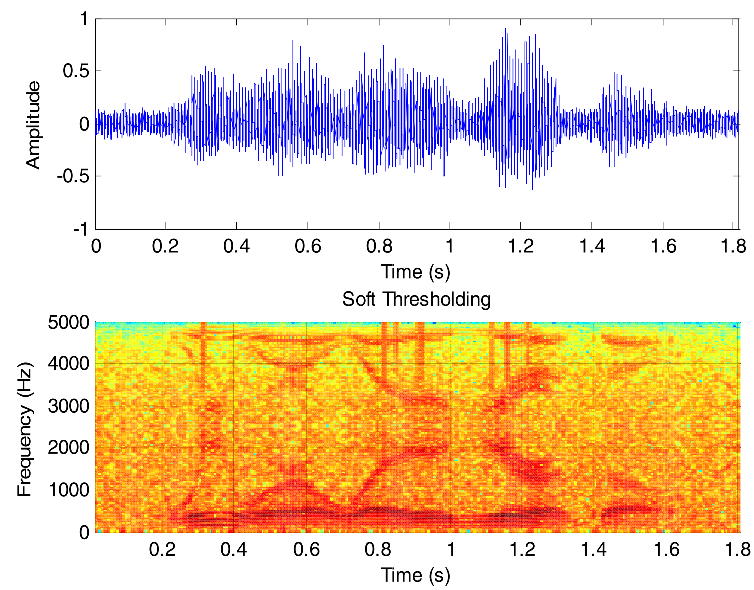
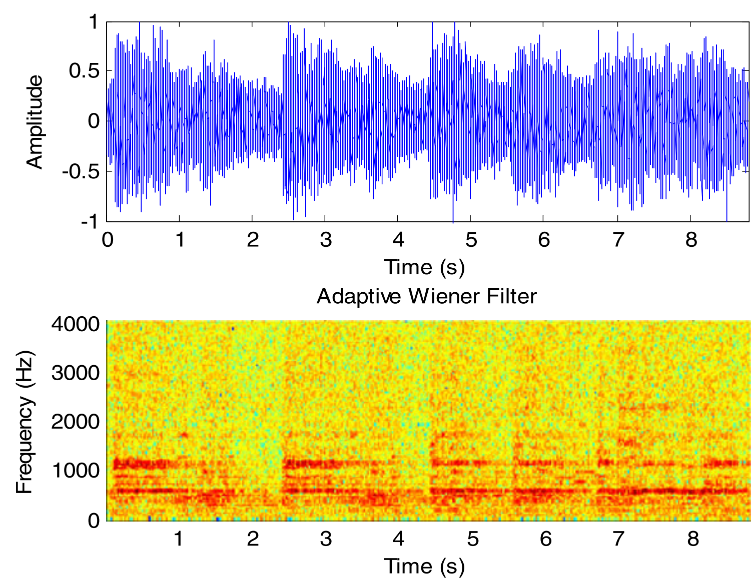


Fig. 14 Time-domain waveform and spectrogram of the enhanced male signal using the adaptive Wiener Filtering method. SNR = 6.1153 dB, SNR_{seg} = 6.0893 dB, LLR = 0.3831, SD = 7.7822 dB



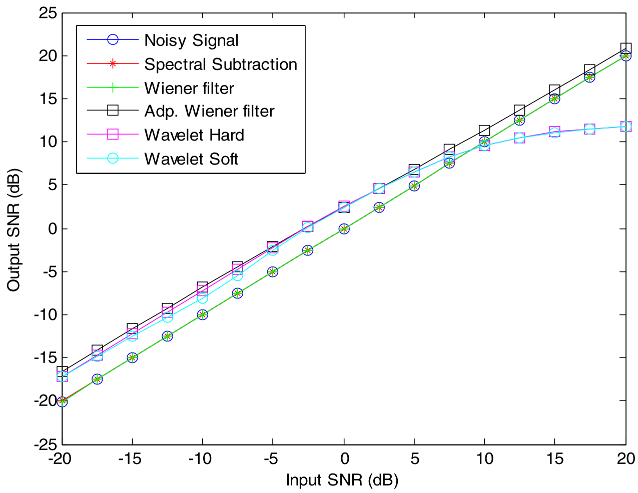


Fig. 15 Output SNR vs. input SNR for all methods on the male signal in the AWGN case

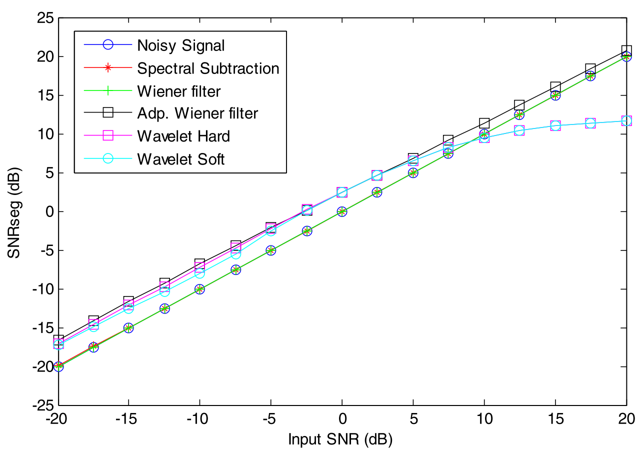


Fig. 16 SNR_{seg} vs. input SNR for all methods on the male signal in the AWGN case

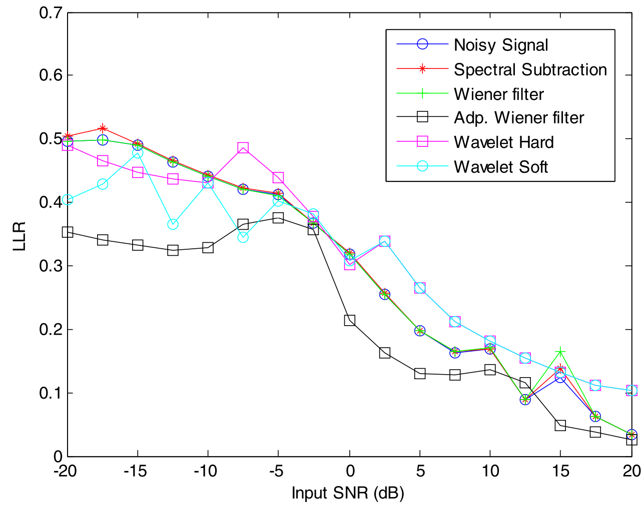


Fig. 17 LLR vs. input SNR for all methods on the male signal in the AWGN case

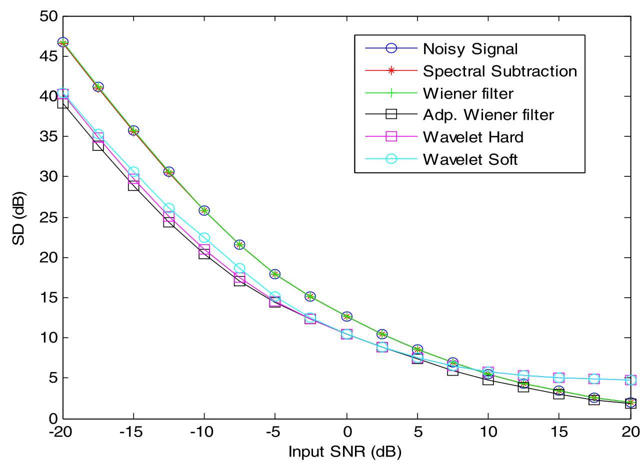


Fig. 18 SD vs. input SNR for all methods on the male signal in the AWGN case

For the colored noise case, we simulated colored noise by lowpass filtering of the AWGN prior to adding it to the signal. We also tested all speech enhancement methods on the male and the female signals in the presence of colored noise. The results of the tests for the male signal at SNR equal 5 dB are shown in Figs. 9 to 14. Figure 15 shows the output SNR versus the input SNR for all methods on the male signal. Figure 16 shows the SNR_{seg} versus the input SNR for all methods on the male signal. Figure 17 shows the variation of the LLR versus the input SNR for all methods on the male signal. Figure 18 shows the variation of the SD versus the input SNR for all methods on the male signal. The case of the colored noise has also been studied in the comparison and its results are given in Figs 19 to 22. A similar study has been repeated on the female signal, and the results are tabulated in Tables 1 to 4.These results are all in favor of the proposed adaptive Wiener filtering method.

7 Conclusion

An adaptive Wiener filtering method for speech enhancement has been presented in this paper. This method depends on the adaptation of the filter impulse response from sample to sample based on the speech signal statistics. The results show that the proposed adaptive Wiener filtering method has the best performance as compared to all other speech enhancement methods mentioned in this paper at both low and high SNR values. The proposed filter succeeds in both the AWGN and the colored noise cases. This is attributed to the adaptive nature of the filter impulse response. This proposed adaptive Wiener filter has another advantage of being dependent only on the noisy signal as a single input.

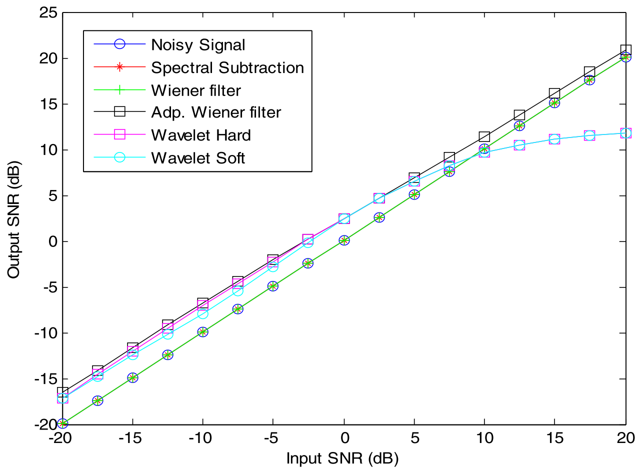


Fig. 19 Output SNR vs. input SNR for all methods on the male signal in the colored noise case

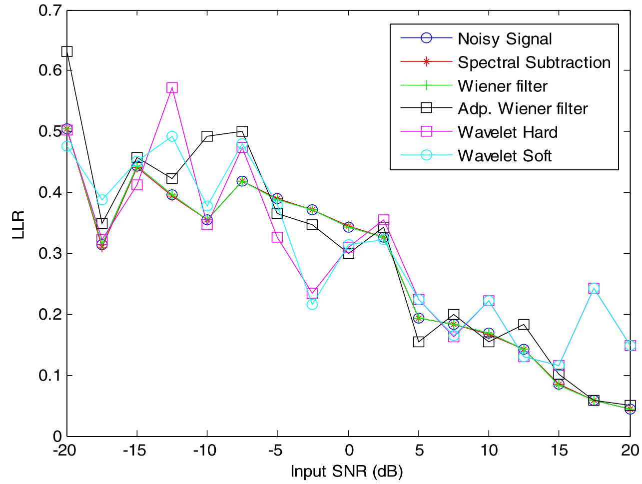


Fig. 21 LLR vs. input SNR for all methods on the male signal in the colored noise case

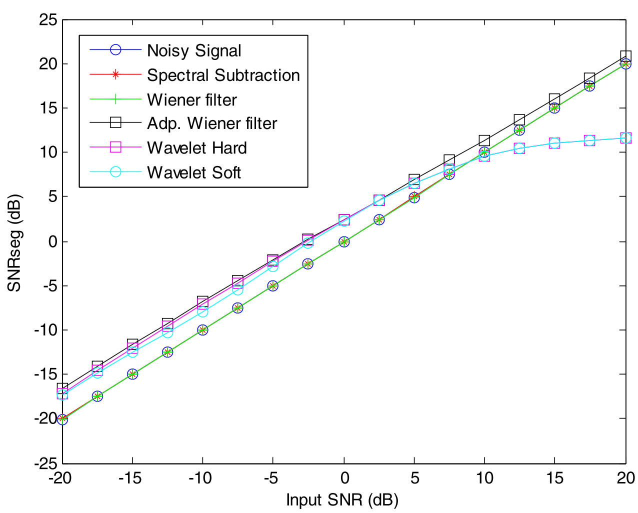


Fig. 20 SNR_{seg} vs. input SNR for all methods on the male signal in the colored noise case

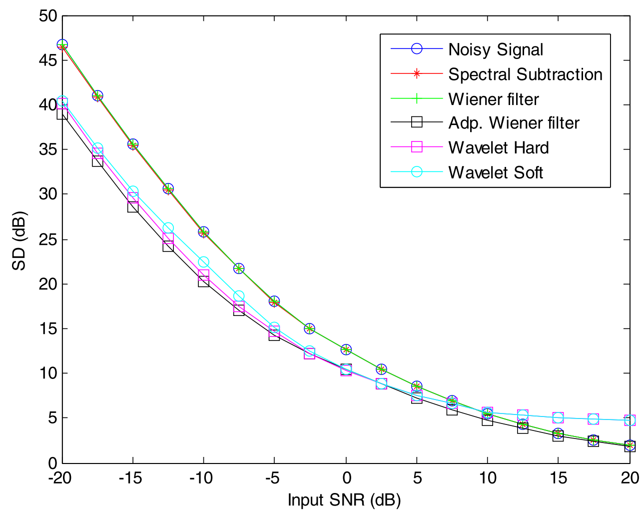


Fig. 22 SD vs. input SNR for all methods on the male signal in the colored noise case

Table 1 Results for SNR = −5 dB for the AWGN case

Signals	Metric	Noisy signal	Spectral subtraction	Wiener filtering	Adaptive Wiener filtering	Hard threshold	Soft threshold
Male	SNR	−5.034	−5.0126	−5.0338	−2.1110	−2.3152	−2.7080
	LLR	0.5663	0.5679	0.5656	0.3935	0.2701	0.3000
	SNRseg	−5.048	−5.0263	−5.0475	−2.1412	−2.3451	2.7315
	SD	18.052	18.0208	18.0517	14.4264	14.6439	15.1937
Female	SNR	−5.070	−5.0550	−5.0699	−2.1248	−2.2984	−3.0666
	LLR	0.3021	0.2993	0.3022	0.3776	0.3861	0.3781
	SNRseg	−5.104	−5.0889	−5.1037	−2.1674	−2.3434	−3.1055
	SD	17.803	17.7814	17.8022	14.2275	14.5844	15.4801

Table 2 Results for SNR = −5 dB for the colored noise case

Signals	Metric	Noisy signal	Spectral subtraction	Wiener filtering	Adaptive Wiener filtering	Hard threshold	Soft threshold
Male	SNR	−4.9999	−4.9781	−4.9994	−2.0612	−2.2013	−2.6602
	LLR	0.3070	0.3059	0.3065	0.2518	0.2147	0.2225
	SNRseg	−5.0251	−5.0033	−5.0246	−2.0923	−2.2294	−2.6938
	SD	18.0640	18.0321	18.0634	14.5021	14.7059	15.2124
Female	SNR	−5.0002	−4.9849	−4.9998	−2.0534	−2.1545	−2.9303
	LLR	0.3484	0.3470	0.3482	0.3268	0.2979	0.3208
	SNRseg	−5.0299	−5.0146	−5.0295	−2.0865	−2.1847	−2.9638
	SD	17.7959	17.7745	17.7955	14.2892	14.2804	15.3252

Table 3 Results for SNR = 5 dB for the AWGN case

Signals	Metric	Noisy signal	Spectral subtraction	Wiener filtering	Adaptive Wiener filtering	Hard threshold	Soft threshold
Male	SNR	4.9877	5.0439	4.9880	6.8726	6.5002	6.4884
	LLR	0.2383	0.2336	0.2383	0.1609	0.1945	0.1942
	SNRseg	4.9601	5.0164	4.9604	6.8423	6.4605	6.4506
	SD	8.5092	8.4721	8.5090	7.3006	7.6423	7.6463
Female	SNR	4.9971	5.0399	4.9974	6.8373	6.9567	6.9486
	LLR	0.3232	0.3434	0.3240	0.3004	0.2061	0.2063
	SNRseg	4.9679	5.0106	4.9682	6.8021	6.9183	6.9097
	SD	8.4378	8.4092	8.4376	7.2527	7.2023	7.2019

Table 4 Results for SNR = 5 dB for the colored noise case

Signals	Metric	Noisy signal	Spectral subtraction	Wiener filtering	Adaptive Wiener filtering	Hard threshold	Soft threshold
Male	SNR	5.0001	5.0571	5.0005	6.9061	6.5832	6.5776
	LLR	0.1333	0.1340	0.1333	0.2900	0.2637	0.2606
	SNRseg	4.9665	5.0235	4.9668	6.8673	6.5489	6.5449
	SD	8.5928	8.5551	8.5926	7.3502	7.5738	7.5809
Female	SNR	4.9998	5.0425	5.0001	6.8292	7.0006	6.9737
	LLR	0.2104	0.2110	0.2105	0.2329	0.1754	0.1750
	SNRseg	4.9733	5.0161	4.9736	6.7984	6.9591	6.9291
	SD	8.4450	8.4177	8.4449	7.2977	7.1786	7.1911

References

Bahoura, M., & Rouat, J. (2001). Wavelet speech enhancement based on the teager energy operator. *IEEE Signal Proc. Letters*, *8*(1).

Berouti, M., Schwartz, R., & Makhoul, J. (1979). Enhancement of speech corrupted by acoustic noise. In *Proc. IEEE int. conf. acoust., speech signal processing* (pp. 208–211).

Boll, S. F. (1979). Suppression of acoustic noise in speech using spectral subtraction. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *ASSP-27*, 113–120.

Boll, S. (1990). Suppression of acoustic noise in speech using spectral subtraction. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *27*, 113–120.

Deller, J. R., Hansen, J. H. L., & Proakis, J. G. (2000). *Discrete-time processing of speech signals* (2nd ed.). New York: IEEE Press.

Ephriam, Y., & Van Trees, H. L. (1993). A signal subspace approach for speech enhancement. In *Proc. international conference on acoustic, speech and signal processing*, Detroit, MI, USA, May 1993 (Vol. II, pp. 355–358).

Ghanbari, Y., & Karami, M. (2004). Spectral subtraction in the wavelet domain for speech enhancement. *International Journal of Software & Information Technology (IJSIT)*, *1*, 26–30.

Ghanbari, Y., Karami, M., & Amelifard, B. (2004). Improved multi band spectral subtraction method for speech enhancement. In *Proc. 6th IASTED internat. conf. on signal image process*, USA (pp. 225–230).

Handel, P. (2007). Power spectral density error analysis of spectral subtraction type of speech enhancement method. *EURASIP Journal On Advances in Signal Processing*, 1–9.

Haykin, S. (1996). *Adaptive filter theory*. New York: Prentice-Hall. ISBN 0-13-322760-X.

Johnstone, I. M., & Silverman, B. W. (1997). Wavelet threshold estimators for data with correlated noise. *Journal of Royal Statistical Society*, *59*(2), 319–351.

Kusumoto, A., Arai, T., Kinshita, K., & Hodoshima, N. (2005). Modulation Enhancement of speech by a preprocessing algorithm for improving intelligibility in reverberant environments. *Speech Communication*, *45*(2), 101–113.

Lim, J. S., & Oppenheim, A. V. (1978). All-pole modelling of degraded speech. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *ASSP-26*.

Lim, J. S., & Oppenheim, A. V. (1979). Enhancement and bandwidth compression of noisy speech. *Proceedings of the IEEE*, *12*, 197–210.

Manikandan, S. (2006). Speech Enhancement based on Wavelet De noising. *Academic Open Internet journal ISSN 1311-4360*, *17*.

Quatieri, T. F. (2002). *Discrete-time speech signal processing: principles and practice*. New York: Prentice Hall.

Shao, Y., & Chang, C. H. (2005). A versatile speech enhancement system based on perceptual wavelet denoising. In *IEEE* (pp. 864–867).

Sheikhzadeh, H. (2001). An improved wavelet-based speech enhancement system. Eurospeech.

Vaseghi, S. V. (2000). *Advanced digital signal processing and noise reduction* (2nd ed.). Chichester: Wiley.

Wu, C. Z., Tai Guo, M., Boxiong, S., & Li, Q. (2000). Studies on a speech signal testing method based on adaptive filtering feedback technology. *Signal Processing Proceedings*, *1*, 543–546.

Ying, J., Li, G., & Wang, Z. (2008). A novel approach to speech signal synthesis. In *Audio, language and image processing*, July 2008 (pp. 680–685).