ISTANBUL UNIVERSITY –
JOURNAL OF ELECTRICAL & ELECTRONICS ENGINEERING

YEAR : 2009 (809-817)
VOLUME : 9
NUMBER : 1

COMPARATIVE STUDY OF A PRIORI SIGNAL-TO-NOISE RATIO (SNR) ESTIMATION APPROACHES FOR SPEECH ENHANCEMENT

Md. Jahangir ALAM[1] **Md. Faqrul Alam CHOWDHURY[2]**
Md. Fasiul ALAM[3]

[1,3]Department of Electrical & Electronic Engineering, Khulna University of Engineering and Technology, Khulna-920300, Bangladesh
[2]Department of Electrical & Electronic Engineering, Bangladesh University of Engineering and Technology, Dhaka-1000, Bangladesh

E-mail: jahangir_kuet@yahoo.com faqrul.buet@yahoo.com fosi_bd@yahoo.com

ABSTRACT

The problem of noise reduction has attracted a considerable amount of research attention over the past several decades. Among the numerous techniques that were developed, 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. So an a priori SNR estimator is a very important component of the Speech enhancement algorithm, especially if the algorithm should be capable of handling non-stationary noise. This paper presents a comparative study of different a priori SNR estimation methods and a performance evaluation of a Wiener denoising technique using those methods by the extensive objective quality measures under various noisy environments.

Key words: *A priori SNR, Wiener filter, decision-directed, speech enhancement*

1. INTRODUCTION

The problem of enhancing speech degraded by noise remains largely open, even though many significant noise reduction algorithms have been introduced over the past decades. This problem is more severe when no additional information on the nature of noise degradation is available, in which case the enhancement technique must exploit only the specific properties of the speech and noise signals.

Since the demand for speech communication systems in mobile environments is increasing, effective speech enhancement is considered as an indispensable speech processing tool. Relevant speech enhancement techniques can be expressed as a spectral noise suppression gain based on the signal-to-noise ratio (SNR) [1]-[4]. Among numerous techniques that were developed, 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.

Received Date: 16.09.2007
Accepted Date: 05.01.2009

In the single channel case the a posteriori SNR (defined after the Wiener filter) is greater than or equal to the a priori SNR (defined before the Wiener filter), indicating that the Wiener filter is always able to achieve noise reduction [14]. However, the amount of noise reduction is in general proportional to the amount of speech degradation. That is, it is difficult to completely reduce the background noise without introducing speech distortion. This may seem discouraging as we always expect an algorithm to have maximal noise reduction without much speech distortion. In [15] it is shown that speech distortion can be better managed in three different ways. If we have some a priori knowledge (such as the linear prediction coefficients) of the clean speech signal, this a priori knowledge can be exploited to achieve noise reduction while maintaining a low level of speech distortion. When no a priori knowledge is available, we can still achieve a better control on noise reduction and speech distortion by properly manipulating the Wiener filter.
An important parameter of numerous speech enhancement algorithms is the *a priori* signal-to-noise ratio (SNR). It is the key point behind the reduction in musical noise by the minimum-mean-square-error (MMSE) estimators [7]. The most practical and computationally efficient approach to determine this parameter is to use the decision-directed (DD) estimator of Ephraim and Malah [2] which results in significant elimination of musical noise. The musical noise comes from the residual noise composed of sinusoidal components randomly distributed over successive frames and sounds disturbing the listener.

Apart from being extremely annoying to the listeners, the musical noise also hampers the performance of the speech coding algorithms to a great extent. In [5] the performance of the DD estimation is analyzed and it is demonstrated that the musical noise is strongly reduced by the a priori SNR corresponding to a highly smoothed version of the a posteriori SNR in noise frames, while the a priori SNR follows the a posteriori SNR with a delay of one frame in speech frames. Therefore, in a conventional DD scheme the estimated noise suppression gain using the delayed a priori SNR having a fixed weighting factor matches the previous frame rather than the current frame and thus it degrades the quality of the enhanced speech signal especially in abrupt transient parts. Different alternative approaches have been proposed to solve this problem while maintaining the benefits of the DD approach. This paper presents a comparative study of those approaches and evaluates the performance of Wiener denoising technique employing those a priori SNR estimators by the extensive objective speech quality measures.

2. CLASSICAL METHOD OF SPEECH ENHANCEMENT

Let the distorted signal be expressed as

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

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

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

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

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

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

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

where μ is a constant, β is the order of the filter and $\xi(m,k)$ is the a priori SNR. If μ =1 and β =1/2 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

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

where $\Gamma_d(m,k) = E\{|D(m,k)|^2\}$ is the noise power spectrum estimated during speech pauses using the classical recursive relation

$$\Gamma_d(m,k) = \lambda_D \Gamma_d(m-1,k) + (1-\lambda_D)|Y(m,k)|^2\,, \qquad (6)$$

where $0 \le \lambda_D \le 1$ is the smoothing factor. In this paper we have chosen $\lambda_D = 0.9$ for all cases. $E\{.\}$ is the expectation operator.

The a priori SNR, which is the second parameter of the noise suppression rule, is expressed as

$$\xi(m,k)=\frac{\Gamma_x(m,k)}{\Gamma_d(m,k)}, \quad (7)$$

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

The instantaneous SNR [6] can be defined as

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

The temporal domain denoised speech is obtained by the following relation

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

The phase of the input noisy signal is used for reconstruction of the estimated speech spectrum based on the fact for human perception the short-time spectral amplitude (STSA) is more important than the phase for intelligibility and quality.

3. ESTIMATION OF A PRIORI SNR

3.1. Decision-Directed (Dd) Approach

A widely used method to determine the a priori SNR from distorted speech is the decision-directed (DD) approach. In [4] the DD approach was defined as a linear combination of (7) and (8). With a weighting parameter α that is constrained to be $0<\alpha<1$, the linear combination results in

$$\xi(m,k)=E\{\alpha\frac{|X(m,k)|^2}{\Gamma_d(m,k)}+(1-\alpha)\vartheta(m,k)\}, \quad (10)$$

However, as this expression is hard to implement in practice, approximations were made. This led to [4]

$$\hat{\xi}_{DD}(m,k)=\max\{\alpha\frac{|H_{DD}(m-1,k)Y(m-1,k)|^2}{\Gamma_d(m,k)}+(1-\alpha)P'[\vartheta(m,k)],\xi_{\min}\} \quad (11)$$

where $P'[x]=x$ if $x\geq 0$ and $P'[x]=0$ otherwise. In this paper we have chosen $\alpha=0.98$ and $\xi_{\min}=.0032$ (i.e. -25dB) by the simulations. The multiplicative gain function for this approach is

$$H_{DD}(m,k)=\frac{\hat{\xi}_{DD}(m,k)}{1+\hat{\xi}_{DD}(m,k)}, \quad (12)$$

Then the enhanced speech spectrum is obtained using (3).

An important characteristic of the DD approach is the dependency on previously enhanced frames which results in biased estimates of the a priori SNR during speech transitions. This method results in significant elimination of musical noise.

3.2. Two-Step Snr (Tsnr) Approach

In the well known DD approach the speech spectrum estimated at the previous frame is used to compute the current a priori SNR. Therefore, the gain function matches the $\xi_M(m,k)$ previous frame rather than current frame, which degrades the noise reduction performance. In order to improve the performance of the noise reduction process in [6] a refinement of the a priori SNR estimated in the DD approach (11) was proposed. In this approach the multiplicative gain function is computed using the DD approach as described in section 3.1. That is the first step of the TSNR approach is the same as the DD approach. In second step of the TSNR approach, the multiplicative gain obtained using (11) and (12) is used to refine the a priori SNR estimation using the following relation

$$\hat{\xi}_{TS}(m,k)=\frac{|H_{DD}(m,k)Y(m,k)|^2}{\Gamma_d(m,k)}, \quad (13)$$

The multiplicative gain for TS approach is given by

$$H_{TS}(m,k)=\frac{\hat{\xi}_{TS}(m,k)}{1+\hat{\xi}_{TS}(m,k)}, \quad (14)$$

The a priori SNR estimated in the first step provides interesting properties but suffers from a delay of one frame which is removed by the second step of TSNR approach [6]. This technique can provide fast response to an abrupt increase in the speech signal without introducing musical noise.

3.3. Modified Tsnr (M-Tsnr) Approach

The delay problem has been removed by the second step of the TSNR approach. The a priori SNR estimated using the TSNR approach varies with the gain function, $H(m,k)$ and thus the noise reduction performance is affected. In order to overcome this dependency a modified two step

SNR (M-TSNR) approach has been proposed in [9]. In this approach the MMSE estimation for $X^2(m,k)$ can be obtained from $Y(m,k)$ as follows:

$$\hat{X}^2(m,k) = E\left\{X^2(m,k)\middle|Y(m,k)\right\}$$
$$= \frac{\int_{-\infty}^{\infty} X^2 P\{Y|X\} P\{X\} dX}{\int_{-\infty}^{\infty} P\{Y|X\} P\{X\} dX}, \quad (15)$$

where $P\{.\}$ denotes the probability density function (PDF). For simplicity of notation the frame index, m and frequency index k are dropped. Assuming Gaussian distributions $P\{Y|X\}$ and $P\{X\}$ are expressed as:

$$P\{Y|X\} = \frac{1}{\sqrt{2\pi\Gamma_d}} e^{\left\{-\frac{(Y-X)^2}{2\Gamma_d}\right\}}, \quad (16)$$

$$P\{X\} = \frac{1}{\sqrt{2\pi\Gamma_x}} e^{\left\{-\frac{X^2}{2\Gamma_x}\right\}}, \quad (17)$$

where $\Gamma_x = E\{X^2\}$. Now from (15)

$$\hat{X}^2 = \frac{\int_{-\infty}^{\infty} X^2 e^{\left\{-\frac{(Y-X)^2}{2\Gamma_d} - \frac{X^2}{2\Gamma_x}\right\}} dX}{\int_{-\infty}^{\infty} e^{\left\{-\frac{(Y-X)^2}{2\Gamma_d} - \frac{X^2}{2\Gamma_x}\right\}} dX}$$

$$= \frac{\int_{-\infty}^{\infty} X^2 e^{\left\{-\left(\sqrt{\frac{\Gamma_x+\Gamma_d}{2\Gamma_x\Gamma_d}}X - \sqrt{\frac{\Gamma_x^2}{2\Gamma_x\Gamma_d+\Gamma_d^2}}Y\right)^2\right\}} dX}{\int_{-\infty}^{\infty} e^{\left\{-\left(\sqrt{\frac{\Gamma_x+\Gamma_d}{2\Gamma_x\Gamma_d}}X - \sqrt{\frac{\Gamma_x^2}{2\Gamma_x\Gamma_d+\Gamma_d^2}}Y\right)^2\right\}} dX}, \quad (18)$$

Taking $Z = \sqrt{\frac{\Gamma_x+\Gamma_d}{2\Gamma_x\Gamma_d}}X - \sqrt{\frac{\Gamma_x^2}{2\Gamma_x\Gamma_d+\Gamma_d^2}}Y$ we have from (18)

$$\hat{X}^2 = \frac{\int_{-\infty}^{\infty} \left(\sqrt{\frac{2\Gamma_x\Gamma_d}{\Gamma_x+\Gamma_d}}Z + \frac{\Gamma_x}{\Gamma_x+\Gamma_d}Y\right)^2 e^{\{-Z^2\}} dZ}{\int_{-\infty}^{\infty} e^{\{-Z^2\}} dZ}$$

$$= \frac{\frac{2\Gamma_x\Gamma_d}{\Gamma_x+\Gamma_d}\int_{-\infty}^{\infty} Z^2 e^{\{-Z^2\}} dZ}{\int_{-\infty}^{\infty} e^{\{-Z^2\}} dZ} + \frac{\Gamma_x^2 Y^2}{(\Gamma_x+\Gamma_d)^2}$$

$$= \frac{\Gamma_x\Gamma_d}{\Gamma_x+\Gamma_d} + \frac{\Gamma_x^2 Y^2}{(\Gamma_x+\Gamma_d)^2}, \quad (19)$$

Here we have used following relations

$$\int_{-\infty}^{\infty} t^q e^{-at^2} dt = \begin{cases} 2I_q(a) & \text{if } q = 0,2,... \\ 0 & \text{if } q = 1,3,... \end{cases} \quad (20)$$

where

$$I_q(a) = 0.5a^{-\frac{q+1}{2}}\Gamma\left(\frac{q+1}{2}\right), \quad (21)$$

$\Gamma(.)$ is the gamma function expressed by the relation,

$$\Gamma(q) = \int_0^{\infty} t^{q-1}e^{-t} dt, \quad (22)$$

Using (5) and (12) in (19) the a priori SNR for M-TSNR is given as

$$\hat{\xi}_{MTS} = \frac{\hat{X}^2}{\Gamma_d} = \frac{\xi_{DD}}{1+\xi_{DD}} + \left(\frac{\xi_{DD}}{1+\xi_{DD}}\right)^2 \frac{Y^2}{\Gamma_d}$$
$$= H_{DD}\left(1 + H_{DD}\frac{Y^2}{\Gamma_d}\right), \quad (23)$$

Using (11) the estimation of ξ_{DD} of (23) is given as

$$\hat{\xi}_{DD}(m,k) = \max\{\alpha \frac{\left|H_{MTS}(m-1,k)Y(m-1,k)\right|^2}{\Gamma_d(m,k)}$$
$$+ (1-\alpha)P'[\vartheta(m,k)], \xi_{\min}\} \quad (24)$$

where $H_{MTS}(m,k)$, is the gain for M-TSNR approach and is expressed as

$$H_{MTS}(m,k) = \frac{\hat{\xi}_{MTS}(m,k)}{1+\hat{\xi}_{MTS}(m,k)}, \quad (25)$$

In this comparative study we have adopted [11] with a slight modification.

In [9], $\hat{\xi}_{MTS}$ was expressed as

$$\hat{\xi}_{MTS} = \frac{\hat{X}^2}{\Gamma_d} = \frac{2\Gamma(1.5)\xi_{DD}}{1+\sqrt{\pi}\xi_{DD}} + \left(\frac{\xi_{DD}}{1+\xi_{DD}}\right)^2 \frac{Y^2}{\Gamma_d}, \quad (26)$$

3.4. The Noncausal Approach

The estimation of the a priori SNR heavily relies on the strong time-correlation between successive speech spectral amplitudes. Being a heuristically motivated approach, the DD approach suffers from frame delay. To overcome the drawbacks of the DD approach a noncausal a priori SNR estimator was proposed in [8] which can make a further reduction in musical noise. The approach is described below. For notational simplicity the frame index, *m*, and frequency index, *k*, are often dropped when there is no confusion. The a priori SNR in the noncausal approach is given as

$$\xi_{NC} = \frac{\Gamma_{X|m+L}}{\Gamma_d}, \tag{27}$$

where $\Gamma_{X|m+L} = E\left\{|X|^2 \middle| y_0^{m+L}\right\}$ is the conditional variance of *X* given the noisy measurements y_0^{m+L}, *L*, is the delay.

Let $\Gamma'_{X|m+L}(m,k) = E\left\{|X(m,k)|^2 \middle| y_0^{m+L} \setminus \{Y(m,k)\}\right\}$ denotes the conditional variance of *X* given y_0^{m+L} excluding the noisy measurement *Y*.

An estimate for $\Gamma_{X|m+L}$ can be obtained by computing the conditional variance of *X* given *Y* and $\hat{\Gamma}'_{X|m+L}(m,k)$ as

$$\hat{\Gamma}_{X|m+L}(m,k) = E\left\{|X|^2 \middle| \hat{\Gamma}'_{X|m+L}, Y\right\}$$

$$= \frac{\hat{\Gamma}'_{X|m+L}}{\hat{\Gamma}'_{X|m+L} + \Gamma_d}\left(\Gamma_d + \frac{\hat{\Gamma}'_{X|m+L}|Y|^2}{\hat{\Gamma}'_{X|m+L} + \Gamma_d}\right), \tag{28}$$

where $\hat{\Gamma}'_{X|m+L}(m,k)$ is obtained by employing the estimates $\hat{X}(m-1,k)$ and $\hat{\Gamma}_{X|m+L-1}(m-1,k)$ from the previous frame.

$$\hat{\Gamma}'_{X|m+L}(m,k) = \max\Big\{\eta\hat{X}^2(m-1,k) + (1-\eta)$$
$$\times[\eta' \sum_{i=-w}^{w} b(i)\hat{\Gamma}_{X|m+L-1}(m-1,k- \tag{29}$$
$$+(1-\eta')\hat{\Gamma}'_{X|[m,m+L]}(m,k)], \Gamma_n$$

where $0 \le \eta \le 1$ is related to degree of nonstationarity of the random process $\{\Gamma_X(m,k)|m=0,1,...\}$, b denotes a normalized window function of length $2w+1$ (i.e., $\sum_{i=-w}^{w} b(i) = 1$) and is related to the correlation between frequency bins of Γ_X, $0 \le \eta' \le 1$ is associated with the reliability of the estimate $\hat{\Gamma}'_{X|[m,m+L]}$ in comparison with that of $\hat{\Gamma}_{X|m+L-1}$.

An estimate for $\Gamma'_{X|[m,m+L]}(m,k)$ is given by

$$\hat{\Gamma}'_{X|[m,m+L]}(m,k) = \max\{\frac{\sum_{(m',i)\in \ell} b(i)|Y(m+m',k-i)|^2}{\sum_{(m',i)\in \ell} b(i)}$$
$$-\beta\Gamma_d, 0\} \tag{30}$$

where $\ell \square \{(m',i)|0 \le m' \le L, -w \le i \le w, (m',i) \ne (0,0)\}$ designates the time frequency indices of the measurements, and $\beta' \ge 1$ is the over-subtraction factor to compensate for a sudden increase in the noise level.

The spectral gain function for this approach is given as

$$H_{NC}(m,k) = \frac{\hat{\xi}_{NC}(m,k)}{1+\hat{\xi}_{NC}(m,k)}, \tag{31}$$

where $\hat{\xi}_{NC} = \frac{\hat{\Gamma}_{X|m+L}(m,k)}{\Gamma_d(m,k)}$. In this paper we have chosen $\eta' = \eta = 0.8$, $L = 3$, $\beta' = 2$, $\Gamma_{\min} = \xi_{\min}\Gamma_d$ and $b = [.25 \quad .5 \quad .25]$ (these values were used in [8]).

Fig.1(a) represents the clean speech signal and Fig.1(b) represents the variation of a priori SNRs of different approaches with the variation of the a posteriori SNR. It is seen that the delay problem of the DD approach has been removed by the TSNR, the M-TSNR and the Noncausal (NC) a priori SNR approaches while they maintain the advantages of the DD approach.

4. EXPERIMENTAL RESULTS

In order to evaluate the performance of the a priori SNR approaches described in section 3.0, we conducted extensive objective quality tests under various noisy environments. The frame sizes were chosen to be 256 samples (32 msec) long with 40% overlap, a sampling frequency of 8 kHz and a hamming window were applied. To evaluate and compare the performance of the a priori SNR estimators, we carried out simulations with the *TEST A* database of Aurora [16]. Speech signals were degraded with five types of noise at global SNR levels of -5 dB, 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).

Table 1, Table 2 and Table 3 represent the Average Segmental SNR (AvgSegSNR), Log Spectral Distance (LSD) and Log Likelihood Ratio (LLR) respectively of enhanced signals at different approaches [10] [11] [13]. In the case of AvgSegSNR all other approaches performs better than the DD approach but the TSNR and the NC performs better than the others. In the case of the LLR measure, M-TSNR approach showed better improvements than all other approaches. For the LSD objective quality measure, the M-TSNR and the NC approaches shows better results than the others. Fig.2 represents the spectrograms of the clean speech signal and enhanced speech signals obtained using different a priori SNR estimation approaches. Speech spectrograms presented in Fig.2 use a Hamming window of 256 samples with 50% overlap and the noisy signal includes N3 (Car Noise) with SNR=5dB. It is observed from Fig.2 that musical residual noise is removed in most part of the figures 2(b), 2(c), and 2(d) specifically in 2(d) and 2(e).

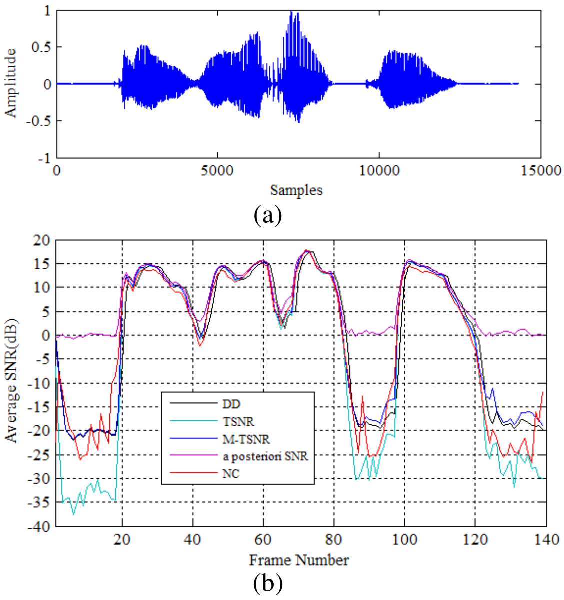

Figure 1. (a) Clean Signal, (b) Variations of a priori SNRs of different approaches with a posteriori SNR. a posteriori SNR (solid pink line), a priori SNR of DD approach (solid grey line), a priori SNR of TSNR approach (solid light green line), a priori SNR of M-TSNR approach (solid blue line), a priori SNR of NC approach (solid red line)

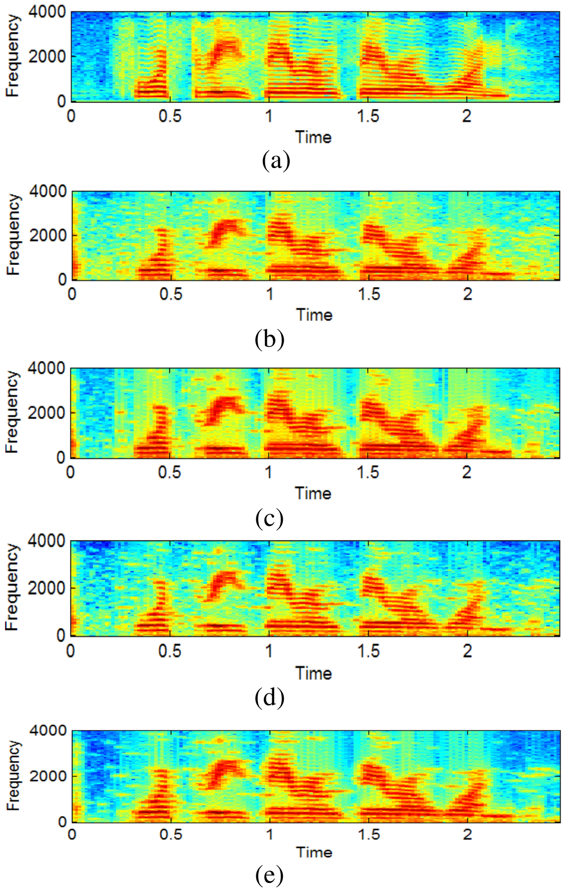

Figure 2 Speech spectrograms, N3 (Car Noise), SNR = 5 dB. (a) Clean signal, (b) enhanced signal (DD approach), (c) enhanced signal (TSNR approach), (d) enhanced signal (M-TSNR approach) and (e) enhanced signal (NC approach).

5. CONCLUSIONS

In this paper we have made a comparative study of different a priori SNR estimation approaches for speech signal enhancement. Performance evaluations of the approaches are carried out using extensive objective speech quality tests [10] [11] [13]. Almost all approaches overcome the delay problem of the conventional DD approach. In terms of SegSNR the TSNR and the NC performed the best. The approach that performed the best in terms of the LLR objective quality measure is the M-TSNR Approach but in terms of LSD measure the M-TSNR and the NC performed the best. Considering all three objective measures we can say that the NC approach performed the best.

REFERENCES

[1] Lim J.S., Oppenheim A.V., "Enhancement and bandwidth compression of noisy speech", *Proc. IEEE,* Vol: 67, pp. 1586- 1604, 1979.

[2] Boll S.F., "Suppression of acoustic noise in speech using spectral subtraction," *IEEE Transactions on Acoustics, Speech, and Signal Processing,* Vol: 27, No. 2, pp. 113–120, 1979.

[3] McAulay R.J., Malpass M.L., "Speech enhancement using a soft-decision noise suppression filter," *IEEE Trans. Acoustics, Speech and Signal Processing,* Vol: 28, pp. 137–145, 1980.

[4] Y. Ephraim, Malah D., "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator," *IEEE Trans. on Acoustics, Speech and Signal Processing*, Vol: 32, pp. 1109–1121, 1984.

[5] Cappe O., "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.

[6] Plapous C., Marro C., Laurent Mauuary and Pascal Scalart, "A Two Step Noise Reduction Technique", *IEEE Trans. Acoustic, Speech and Signal Processing,* Vol: 1, pp. 289–292, 2004.

[7] Scalart P., Filho J.V., "Speech Enhancement Based on a Priori Signal to Noise Estimation," *IEEE Int. Conf. on Acoustics, Speech and Signal Proc.*, pp. 629–632, 1996.

[8] Cohen I., "Speech Enhancement using a Noncausal A Priori SNR Estimator", I*EEE signal processing letters*, Vol: 11, No. 9, pp. 725-728, September 2004.

[9] Ou S., Zhao X., Gao Y., "Speech Enhancement Employing Modified A Priori SNR Estimation", *Proc. of 8th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing,* pp. 827-831, July-August, 2007.

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

[11] Bayya A., Vis M., "Objective Measures for Speech Quality Assessment in Wireless Communications", *Proc. of IEEE international conference on Acoustics, Speech and Signal Processing*, Vol: 1, pp. 495–498, 1996.

[12] Vaseghi S.V., *Advanced Signal Processing and Digital Noise Reduction, John Wiley and Teubner*, 1996.

[13] Quackenbush S.R., Barnwell T.P., Clements M.A., *Objective measures of speech quality. Prentice-Hall*, Englewood Cliffs, New Jersey, 1988.

[14] Chen J., Banesty J., Huang Y., "On the optimal linear filtering techniques for noise reduction", *Speech communication,* Vol: 39, pp. 305-316, 2007.

[15] Chen J., Banesty J., Huang Y., Declo S., "New Insights into the Noise reduction Wiener Filter", *IEEE Trans. on Audio, Speech and Language processing,* Vol: 14, No. 4, July, 2006.

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

Table 1 Average Segmental SNR of enhanced signal

Noise Type	Input SNR(dB)	DD	TSNR	M-TSNR	NC
N1 Subway	-5	-2.8288	-2.6173	-2.7234	**-2.3320**
	0	-1.2066	-0.2194	-0.8365	**-0.2895**
	5	2.1933	2.7402	2.5992	**2.8515**
	10	5.1047	5.4858	5.4630	**5.3981**
	15	8.0351	8.6631	8.2471	**8.7369**
	20	10.6350	10.7510	10.5727	**10.7846**
N2 Babble	-5	-4.3295	**-3.9976**	-4.4160	-4.3882
	0	-2.4296	**-2.3308**	-2.3834	-2.5840
	5	0.0568	**0.3823**	0.1696	0.0584
	10	2.5243	**3.1131**	2.7563	2.9780
	15	5.2433	**6.0730**	5.3119	5.5898
	20	7.6099	**8.3742**	7.6983	8.3669
N3 Car	-5	-2.5189	**-1.8967**	-2.3989	-2.5562
	0	-0.4902	**0.9769**	-0.1870	0.4565
	5	2.5768	**3.9506**	2.9002	3.5312
	10	5.3008	**6.4610**	5.4362	5.8163
	15	8.5861	**9.3308**	8.8451	9.1285
	20	11.9238	**12.6625**	12.2007	12.3792
N4 Exhibition	-5	-1.7590	**-1.4132**	-2.0245	-1.6780
	0	0.7211	**1.2539**	0.4210	0.8123
	5	2.1753	**2.5841**	2.3611	2.5048
	10	5.2387	**5.6026**	5.2845	5.5288
	15	7.3885	7.6674	7.6935	**8.0827**
	20	10.0256	10.4616	10.5318	**10.9133**
WGN White	-5	-3.0976	-2.6039	**-1.5603**	-2.1594
	0	1.1353	**2.4886**	2.2561	2.1132
	5	3.9579	**5.2959**	4.4119	4.6715
	10	6.0814	7.5250	6.5529	**7.6625**
	15	8.4770	9.9505	9.0382	**10.2282**
	20	11.5672	12.1502	11.9415	**13.0269**

Table 2 Log Spectral Distance of enhanced signals

Noise Type	Input SNR(dB)	DD	TSNR	M-TSNR	NC
N1 Subway	-5	2.4370	**2.2899**	2.4701	2.3519
	0	2.1199	2.1192	2.1663	**2.0786**
	5	1.6808	1.8900	**1.6790**	1.8054
	10	1.6937	1.7685	1.6980	**1.6657**
	15	1.3976	1.5282	**1.3941**	1.4157
	20	1.1780	1.4299	**1.1958**	1.2873
N2 Babble	-5	2.4353	2.4078	2.4452	**2.3588**
	0	2.0871	2.0865	2.1225	**2.0477**
	5	1.7440	1.7665	1.7897	**1.7002**
	10	1.5349	1.5451	1.5429	**1.4433**
	15	1.4892	1.5552	1.4796	**1.4242**
	20	1.3576	1.5466	1.3179	**1.2756**

N3 Car	-5	1.8677	1.9365	**1.8412**	1.9693
	0	1.7058	1.9061	**1.6715**	1.8641
	5	1.4996	1.7055	**1.4572**	1.6305
	10	1.3910	1.5753	**1.3322**	1.5081
	15	1.2880	1.5810	**1.2146**	1.4410
	20	1.2668	1.5989	**1.1793**	1.4258
N4 Exhibition	-5	2.1051	2.1091	**2.0800**	2.1702
	0	1.8213	1.8632	**1.8564**	1.8407
	5	1.6954	1.7871	**1.6890**	1.7350
	10	1.7512	1.8656	**1.7159**	1.8042
	15	1.4365	1.5981	**1.3813**	1.4503
	20	1.3860	1.5741	**1.3047**	1.4005
WGN White	-5	2.2926	**2.1068**	2.2952	2.1148
	0	2.0508	**1.9257**	2.0413	1.9782
	5	1.8324	1.8178	1.8172	**1.7653**
	10	1.5942	1.6546	**1.5496**	1.6061
	15	1.3814	1.5587	**1.3332**	1.4982
	20	1.2570	1.5615	**1.1959**	1.4174

Table 3 Log Likelihood Ratio of enhanced signals

Noise Type	Input SNR(dB)	DD	TSNR	M-TSNR	NC
N1 Subway	-5	1.279	1.665	**1.373**	1.568
	0	1.331	1.592	**1.325**	1.590
	5	0.844	0.989	**0.803**	1.002
	10	0.722	0.925	**0.700**	0.847
	15	0.535	0.644	**0.503**	0.580
	20	**0.355**	0.356	0.374	0.366
N2 Babble	-5	1.328	1.407	**1.303**	1.406
	0	1.187	1.283	**1.183**	1.304
	5	0.932	0.984	**0.914**	.973
	10	.778	.837	**.766**	.788
	15	.859	.849	**.839**	.848
	20	.713	**.632**	.725	.633
N3 Car	-5	**1.211**	1.465	1.219	1.505
	0	1.175	1.366	**1.150**	1.552
	5	.898	1.006	**.887**	1.159
	10	.764	.881	**.744**	.971
	15	.534	.603	**.457**	.608
	20	.503	.537	**.429**	.525
N4 Exhibition	-5	**1.560**	1.712	1.590	1.746
	0	**1.327**	1.545	1.360	1.584
	5	1.238	1.262	**1.182**	1.189
	10	1.289	1.386	**1.245**	1.444
	15	.797	.833	.755	**.754**
	20	.688	.783	**.648**	.713
WGN White	-5	**1.305**	1.747	1.350	1.466
	0	**1.162**	1.488	1.206	1.340
	5	.925	1.201	**.855**	1.095
	10	.803	1.021	**.774**	1.068
	15	.637	.817	**.564**	.837
	20	.617	.706	**.517**	.727