Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.12785/amis/090217

Noise Tracking Algorithm for Speech Enhancement

M. Kalamani[1,*], *S. Valarmathy*[1] *and M. Krishnamoorthi*[2]

[1] Department of ECE, Bannari Amman Institute of Technology, Sathyamangalam, Tamilnadu, India.
[2] Department of CSE, Bannari Amman Institute of Technology, Sathyamangalam, Tamilnadu, India.

Received: 19 May 2014, Revised: 18 Aug. 2014, Accepted: 20 Aug. 2014
Published online: 1 Mar. 2015

Abstract: In this paper, the improved noise tracking algorithm for speech enhancement is proposed. This method is used to detect the speech presence probability based on chi square distribution. During speech presence period, the time varying smoothing factor is adjusted. In addition, the estimated noise variance is recursively smoothed then averaged for various noises. This proposed method can track the noise signal with different input SNR (0dB and 5dB) levels. The performance of the proposed and the existing methods are evaluated by various noise conditions. From these evaluated results, it is observed that the proposed method reduces the performance measures as 6% - 58% of MSE and 3% - 97% of LogErr as compared to that of the various existing algorithms under various noise conditions with optimal smoothing factors $\alpha_p = 0.97$ and $\alpha_d = 0.7$. When this is integrated into the speech enhancement, it improves the speech signal quality and intelligibility with less speech distortion and residual noise.

Keywords: speech enhancement, noise tracking, noise variance, speech distortion, residual noise, speech quality, intelligibility

1 Introduction

Major challenging problem in the speech processing applications like mobile phones, hands-free phones, car communication, teleconference systems, hearing aids, voice coders, automatic speech recognition and forensics etc., is to eliminate the background noise. Speech enhancement algorithms are widely used for these applications in order to remove the noise from degraded speech in the noisy environment. Hence, the conventional noise reduction methods introduce more residual noise and speech distortion. So, it has been found that the noise reduction process is more effective to improve the speech quality but it affects the intelligibility of the clean speech signal.

The noise estimation method plays the major role in speech enhancement. For stationary noise conditions, the noise statistic is estimated by averaging the noisy spectrum, which is detected during the silence period. In non-stationary noise conditions, the noise spectrum should vary rapidly over time. The estimated noise spectrum is updated by the use of voice activity detector (VAD). In this, it is very difficult to decide whether the speech is present or absent. Due to sudden rise in the noise power, it may be misinterpreted as speech present period.

Martin (2001) proposed an algorithm for tracking the noise based on Minimum Statistics (MS) [22]. This method failed when the noise signal level is higher than the clean speech signal. Cohen (2002) proposed a Minima Controlled Recursive Averaging (MCRA) in which the noise is estimated by averaging the past power spectrum based on smoothing parameter [18,19,20,21]. In this case, there is no hard decision about the speech presence probability. In addition, the noise estimation is continuously updated during speech absence period. This type of noise estimator is computationally efficient and robust with respect to SNR. Cohen (2003) further improved the MCRA method based on speech presence probability estimation which is called as Improved Minima Controlled Recursive Averaging (IMCRA) method [15,16]. In this method, smoothing and minimum tracking is carried out in the two iterations. In order to reduce the speech leakage, it requires a large window sequences for minimum tracking which limits the ability to track the sudden rise in the noise level.

Rangachari (2006) et al. introduced an algorithm which estimates the noise using time-frequency smoothing factors computed based on speech presence probability [12,13,14,17]. The computed local minimum is independent to window length, which improves the tracking speed when rapid variations in the noise signal.

* Corresponding author e-mail: kalamani.mece@gmail.com

Erkelens (2008) et al. proposed the Minimum Mean Square Error (MMSE) based noise estimation method which reduces the speech leakage and allows for faster tracking as comparison with the MS based algorithms [4, 5, 6, 10, 11]. This requires a hard decision about speech presence probability and bias compensation in order to improve the maximum likelihood.

Timo Gerkmann (2011) et al. introduced noise estimator which replaces the VAD by soft Speech Presence Probability (SPP) based on the Gaussian distribution [1, 2, 3, 8, 9] which is computationally and memory wise more efficient. In this, the speech and noise spectral coefficients are Gaussian distributed in which it is symmetric with respect to the mean value. This introduces the speech leakage because of non stationary noise conditions.

In this paper, the noise estimation algorithm by speech presence probability based on chi square distribution is proposed. This distribution provides the goodness to fit of an original and the estimated noise signal. This paper is organized as follows. Section 2 introduces the signal modeling and Section 3 provides some of the existing noise tracking algorithms. The proposed noise tracking algorithm is described in Section 4. Section 5 presents the performance evaluation of the existing and proposed algorithms and Section 6 provides conclusions.

2 Signal Modeling

It is considered that the noisy signal is a bandlimited and sampled speech signal which is the sum of a clean speech signal $s(i)$ and a disturbing noise $n(i)$, $y(i) = x(i) + n(i)$ where i denotes the sampling time index. Assume that noisy speech is statistically independent and zero mean.

By window technique, the noisy signal is converted into frames of L consecutive samples and then FFT is computed on windowed data. Before the next FFT computation the window is shifted by R samples. This sliding window FFT of the signal can be written as,

$$Y(\lambda,k) = \sum_{i=0}^{L-1} y(\lambda R + i)h(i)e^{\frac{-j2\pi ki}{L}} \tag{1}$$

where, λ is the sub sampled time index, k is the frequency bin index, $k \in \{0, 1, \ldots, L-1\}$ and the normalized center frequency Ω_k is given by $\Omega_k = \frac{2\pi k}{L}$.

The additive-noise signal model of the form is,

$$Y(\lambda,k) = X(\lambda,k) + N(\lambda,k) \tag{2}$$

where, $Y(\lambda,k)$, $X(\lambda,k)$ and $N(\lambda,k)$ are the short-time DFT coefficients obtained at frequency index k in each signal frame λ from the noisy speech, clean speech and noise signal respectively. The noisy amplitude is $R = |Y|$, the speech spectral amplitude is $A = |X|$ and the noise amplitude is $D = |N|$.

The noise spectral variance is $\lambda_D = E(|N(\lambda,k)|^2) = E(D^2)$ and the speech spectral variance is $\lambda_x = E(|X(\lambda,k)|^2) = E(A^2)$. The prior SNR and the posterior SNR are defined as,

$$\xi(\lambda,k) = \frac{\lambda_S(\lambda,k)}{\lambda_D(\lambda,k)}, \quad \zeta(\lambda,k) = \frac{R^2(\lambda,k)}{\lambda_D(\lambda,k)} \tag{3}$$

respectively.

3 Speech Presence Probability (SPP) based Noise Estimation Method [8]

In this method the hard decision with VAD is replaced with soft decision by means of speech presence probability. For MMSE estimator, the noise periodogram under speech presence period is given by,

$$E(|N|^2|Y) = P(H_0|Y)E(|N|^2|Y,H_0) + P(H_1|Y)E(|N|^2|Y,H_1) \tag{4}$$

where, H_0 indicates the speech absence period and H_1 indicates the speech presence period.

Both real and imaginary parts of the speech and noise spectral coefficients are Gaussian distributed. Based on Bayes theorem, assume $P(H_0) = P(H_1)$ for uniform priors. The probability of speech presence is,

$$P(H_1|Y) = (1 + (1+\xi_{opt}).exp(-\frac{|Y|^2}{\overline{\sigma}_N^2}\frac{\xi_{opt}}{1+\xi_{opt}}))^{-1} \tag{5}$$

where, $\overline{\sigma}_N^2$ is the noise variance estimate of the previous frame. Similarly, fixed optimal a priori SNR is selected $10\log_{10}(\xi_{opt}) = 15dB$ in order to minimize the total probability error when the true a priori SNR lies between $-\infty$ and 20 dB.

To derive the posterior SNR $\gamma = \frac{|Y|^2}{\overline{\sigma}_N^2}$ in terms of ξ_{opt} and $P(H_1|Y)$ is given by,

$$\gamma = \log\left(\frac{1+\xi_{opt}}{P(H_1|Y)^{-1} - 1}\right)\frac{(1+\xi_{opt})}{\xi_{opt}} \tag{6}$$

If $10\log_{10}(\xi_{opt}) = 15dB$, then posterior SNR satisfies $\gamma > 1$. From this, it can be concluded that the speech presence only when $P(H_1|Y)$ is adequately large. Under speech absence, noisy power equals to the noise power.

Spectral noise power is underestimated only when $P(H_1|Y) = 1$ and $|Y|^2$ is smaller than the true noise power. Noise power may not be updated and remains underestimated. To overcome this, recursively smoothing the speech presence probability by,

$$\overline{P}(l) = 0.9\overline{P}(l-1) + 0.1 + P(H_1|Y(l)) \tag{7}$$

If $\overline{P}(l)$ is larger than a threshold, then force the current estimate $P(H_1|Y)$ to be smaller than 1 as,

$$P(H_1|Y(l)) \leftarrow \begin{cases} \min(0.99, P(H_1|Y(l))), & \overline{P}(l) > 0.99 \\ P(H_1|Y(l)), & \text{else} \end{cases} \tag{8}$$

This step fits well and it is more memory efficient than the safety net.

The noise periodogram estimate is updated by,

$$|\overline{N}|^2 = E(|N|^2|Y) = P(H_0|Y)|Y|^2 + P(H_1|Y)\overline{\sigma}_N^2 \quad (9)$$

where, $P(H_0|Y) = 1 - P(H_1|Y)$ and $\overline{\sigma}_N^2$ is the spectral noise power estimated in the previous frame. Then the noise power estimation by temporal smoothing is given by,

$$\overline{\sigma}_N^2(l) = \alpha\overline{\sigma}_N^2(l-1) + (1-\alpha)|\overline{N}(l)|^2 \quad (10)$$

Assume smoothing factor, $\alpha = 0.8$ for this noise estimation.

4 Proposed Method Based on Chi Square Distribution

In proposed method, noise is estimated based on chi square distribution. This provides the best fit between the distribution of noisy speech and the estimated noise from previous frame. If long frames are used, this distribution converges to a Gaussian distribution. Whenever the noise only frame is found, the noise power is updated. The noise is updated based on following two hypotheses,

Null hypotheses, H0 : noise only frame
Alternate Hypotheses, H1 : noisy frame

The spectral components of the noisy speech are obtained by computing the Short Time Fourier Transform (STFT) of the Hanning windowed sequence. These spectral components of the current frame are considered as an observation sequence for the chi square statistic and the estimated noise variance of the previous frame as estimated sequence. Each frame consists of N frequency bins and these sequences are described as follows,

$$O = [o_1, o_2, \ldots, o_k, \ldots, o_N] \quad (11)$$

$$E = [e_1, e_2, \ldots, e_k, \ldots, e_N] \quad (12)$$

Then the chi square test is applied for these frequency bins and the chi square statistic is given by,

$$NS^2 = \sum_{i=1}^{N} \frac{(o_k - e_k)^2}{e_k} \quad (13)$$

The calculated statistic value is compared with the threshold value which is obtained by chi square tables with $(N-1)$ degrees of freedom. The hypotheses are tested by,

$$\left.\begin{array}{l} \text{if } (NS^2 > \text{ threshold}) \\ \quad I(\lambda,k) = 1 \\ \text{else} \\ \quad I(\lambda,k) = 0 \\ \text{end} \end{array}\right\} \begin{array}{l} \rightarrow \text{accept } H_1 \\ \\ \rightarrow \text{accept } H_0 \end{array} \quad (14)$$

For each frame, this testing is carried out. If H_1 is accepted which indicates the presence of the speech signal. During this period, the noise signal is estimated by updating the speech presence probability [17]. It is updated with a first order recursion, which depends on smoothing factor α_p,

$$\overline{p}(\lambda,k) = \alpha_p \overline{p}(\lambda-1,k) + (1-\alpha_p)I(\lambda,k) \quad (15)$$

The time-frequency smoothing parameter $\alpha_s(\lambda,k)$ depends on the speech presence probability estimate $\overline{p}(\lambda,k)$ and smoothing factor α_d,

$$\alpha_s(\lambda,k) = \alpha_d + (1-\alpha_d)\overline{p}(\lambda,k) \quad (16)$$

where, α_d lies between 0 to 1, α_s depends on both α_p and α_d which always lies between α_d to 1. When the speech presence probability estimate $\overline{p}$ is near to 1, α_s is reduced to 1 and then the noise estimate should be kept close to its previous value. This prevents the speech power to leak into the noise variance estimate. Noise variance update is faster when the speech presence probability estimate is lower. To avoid speech leakage, an accurate estimate of $\overline{p}(\lambda,k)$ is needed. In this method, the minimum value of a smoothed power spectrum of the noisy signal is controlled by the estimate of $p(\lambda,k)$ [17].

The noise variance estimate $\overline{\lambda_D}$ is obtained by recursively smoothing the noisy power with time-frequency smoothing factor and it is given by,

$$\overline{\lambda_D}(\lambda,k) = \alpha_s(\lambda,k)\overline{\lambda_D}(\lambda-1,k) + (1-\alpha_s(\lambda,k))R^2(\lambda,k) \quad (17)$$

Averaging the neighboring bins gives the strong correlation of the speech presence in neighboring frequency bins of consecutive frames [17]. The proposed noise tracking algorithm is summarized in Algorithm 1.

Algorithm 1 Proposed Noise tracking Algorithm

1: **for all** frame index λ and frequency bin index k **do**, (Assume Frame size, $N = 64$)
2: Compute Chi square statistic
$NS^2 = \sum_{i=1}^{N} \frac{(o_k-e_k)^2}{e_k}$
3: Compare the computed value with the threshold value as $(N-1)$ freedom from chi square table
if $(NS^2 >$ threshold$)$
$I(\lambda,k) = 1$ $\rightarrow$ accept H_1
else
$I(\lambda,k) = 0$ $\rightarrow$ accept H_0
end
4: From the evaluation, the optimal values for smoothing factors are found as $\alpha_p = 0.97$ and $\alpha_d = 0.7$. Update speech presence probability,
$\overline{p}(\lambda,k) = \alpha_p\overline{p}(\lambda-1,k) + (1-\alpha_p)I(\lambda,k)$
5: Recursively smoothening the time-frequency smoothing factor,
$\alpha_s(\lambda,k) = \alpha_d + (1-\alpha_d)\overline{p}(\lambda,k)$
6: Estimate the noise variance,
$\overline{\lambda_D}(\lambda,k) = \alpha_s(\lambda,k)\overline{\lambda_D}(\lambda-1,k) + (1-\alpha_s(\lambda,k))R^2(\lambda,k)$
7: **end for**

5 Performance Evaluation

In this section, the performance of the proposed noise tracking algorithm is compared with IMCRA, MMSE and MMSE with SPP methods. For the evaluation, the input noisy signal is taken from NOIZEUS database for various noise environments such as: airport, car, babble, exhibition, restaurant, street, station and train noises with 0 dB and 5 dB.

5.1 Evaluation of Mean Square Error (MSE)

The relative mean squared error between the true noise spectrum and the estimated noise spectrum is computed as follows,

$$MSE = \frac{1}{M}\sum_{\lambda=0}^{M-1}\frac{\sum_k[\overline{\lambda_D}(\lambda,k)-\sigma_D^2(\lambda,k)]^2}{\sum_k\sigma_D^2(\lambda,k)} \qquad (18)$$

where, $\overline{\lambda_D}(\lambda,k)$ is the estimated noise variance, $\sigma_D^2(\lambda,k)$ is the true noise power and M is the number of frames in the noisy speech signal.

5.2 Evaluation of Log Error

Another performance measure is the LogErr distortion measure. In this, the estimated noise signal is compared with the original noise include two terms as,

$$LogErr = LogErrOver + LogErrUnder \qquad (19)$$

where, the term LogErrOver is used to measure the contributions of an overestimation of the true noise power as,

$$LogErrOver = \frac{1}{NL}\sum_{l=0}^{L-1}\sum_{k=0}^{N-1}\left|\min\left(0, 10\log_{10}\left(\frac{\sigma_{N,k}^2(l)}{\overline{\sigma}_{N,k}^2(l)}\right)\right)\right| \qquad (20)$$

while, the term LogErrUnder is used to measure the contributions of an underestimation of the true noise power as,

$$LogErrUnder = \frac{1}{NL}\sum_{l=0}^{L-1}\sum_{k=0}^{N-1}\max\left(0, 10\log_{10}\left(\frac{\sigma_{N,k}^2(l)}{\overline{\sigma}_{N,k}^2(l)}\right)\right) \qquad (21)$$

The value of the LogErrOver term indicates the attenuation of a speech signal which produces the speech distortion. Other term, LogErrUnder indicates the noise signal that is not attenuated in the enhanced signal which results in residual noise.

The time-frequency smoothing factor in the proposed method depends on the two smoothing factors namely α_p and α_d. Fig. 1 shows the MSE and LogErr (in dB) for the various values (0.1 to 0.99) of α_p, α_d for the proposed noise tracking algorithm under car and train noises with input SNR of 5 dB. From these results, it is observed that the performance measures MSE and LogErr are decreased by varying the smoothing factor α_p from 0.1 to 0.97 then starts increasing from 0.97 onwards.

From these evaluated results, it is considered that the value 0.97 is found as the optimal value for the smoothing factor α_p. In addition, it is observed that the performance measures MSE and LogErr is decreased by varying the smoothing factor α_d from 0.1 to 0.7 then starts increasing from 0.7 onwards. It is observed from the results that the optimal value for the smoothing factor α_d is found as 0.7. These optimal values are used to compute the time-frequency smoothing factor for the proposed noise tracking algorithm. MSE values for 0 db and 5 dB levels of various noises are determined for the existing and proposed methods. These evaluation methods results are shown in Fig. 2.

From these results, it is observed that the IMCRA method produces larger MSE for all the noises. This indicates that, if there are rapid changes in the noise level then the IMCRA method fails to track the noise level which causes the speech distortion and residual noise. The performance of the MMSE method slightly improved as compared to IMCRA method. In MMSE with SPP method, noise tracking level is improved in the considerable level as compared to IMCRA and MMSE method. But, still it produces the musical noise. For various noises, the proposed method reduces the mean square error as compared to all the existing methods method and also provides better tracking. The LogErr is calculated for existing and proposed algorithms with 0 dB and 5 dB levels of the various noises and the evaluation results are shown in Fig. 3.

It can be revealed that, the IMCRA, MMSE and MMSE with Speech Presence Probability (SPP) methods produce higher LogErr measures. The proposed noise estimator is compared with the other existing approaches which produced less LogErr (in dB). From these evaluated results, it is observed that the proposed method reduces the speech distortion and residual noise. In addition, it improves the speech signal quality and intelligibility for 0 dB and 5 dB levels of stationary and non stationary noisy environments.

For various noises, Table 1 shows the comparison of the performance measures MSE and LogErr in dB for IMCRA, MMSE, MMSE with SPP algorithms and proposed noise tracking algorithm for various noises with different input SNR (0dB and 5dB) levels. From these tabulated results, it is observed that the proposed method reduces the MSE as 51% - 58%, 50% - 58% and 6% - 51% as compared to IMCRA, MMSE and MMSE with SPP methods respectively. In addition, this reduces the LogErr as 45% - 97%, 5% - 98% and 3% - 92% as compared to that of the various existing methods. This evaluated result indicates that the proposed method introduces less speech distortion and residual noise for the stationary and non stationary noise conditions. In

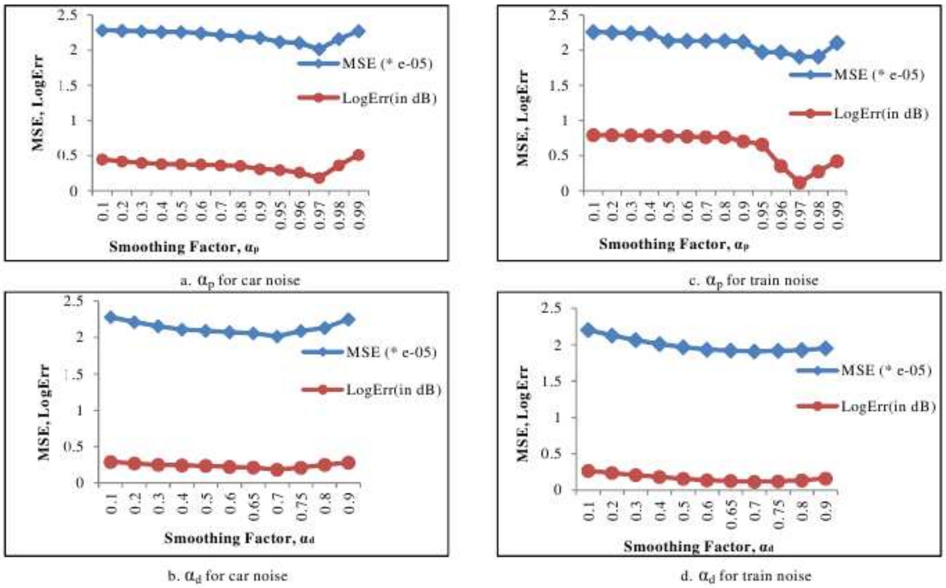

Fig. 1: (a to d) MSE and LogErr (in dB) for the various values (0.1 to 0.99) of Smoothing factors α_p, α_d for the Proposed Noise Tracking Algorithm for car and train noises with input SNR of 5dB

Table 1: Comparison of MSE and LogErr values for various noises by IMCRA, MMSE, MMSE with SPP algorithms and Proposed Noise Tracking Algorithm

Noise Type		**IMCRA Method**		**MMSE Method**		**MMSE-SPP Method**		**Proposed Method**	
Noise Name	**dB Level**	**MSE (* e-05)**	**LogErr in dB**	**MSE (* e-05)**	**LogErr in dB**	**MSE (* e-05)**	**LogErr in dB**	**MSE (* e-05)**	**LogErr in dB**
Airport	0	4.4275	0.3400	4.4284	0.2823	3.113	0.2772	2.0181	0.1876
	5	4.4273	0.7595	4.4293	1.9657	3.1647	0.1622	2.0560	0.1441
Babble	0	4.4275	1.2045	4.4278	0.8726	3.0799	0.2268	2.0412	0.1805
	5	4.4295	1.8283	4.4296	0.7401	4.2272	0.4315	2.1860	0.1387
Car	0	4.4136	4.7526	4.4191	4.0526	2.1704	0.3162	1.8577	0.1269
	5	4.4162	1.0715	4.4213	1.1805	2.0380	0.1625	1.6147	0.1444
Exhibition	0	4.3902	2.3011	4.4058	2.0508	2.8052	1.1405	1.8637	0.2745
	5	4.4295	0.7552	4.4298	3.3556	4.3919	0.9876	2.1704	0.1089
Restaurant	0	4.4297	2.2203	4.4316	3.0219	3.6033	0.2577	2.1063	0.0522
	5	4.4292	1.8904	4.4295	1.5121	3.9028	1.2137	2.0567	0.1156
Station	0	4.4221	0.5357	4.4229	0.2530	3.8497	0.4419	2.0915	0.2048
	5	4.4200	1.7940	4.4229	0.2275	2.1408	0.2242	2.0027	0.2165
Street	0	4.4108	1.8726	4.4160	2.2169	2.8047	0.5974	1.9434	0.1935
	5	4.4299	3.6510	4.4309	5.8001	3.9706	1.5994	2.1625	0.1214
Train	0	4.4268	6.3458	4.4259	4.7428	3.2871	0.4146	1.9778	0.2692
	5	4.4449	4.1614	4.4363	3.4045	3.8584	1.3189	1.9094	0.1175

addition, it provides better noise tracking for various noises with different input SNR levels.

6 Conclusions

In this paper, the improved noise tracking algorithm is presented for estimating the noise. Based on the chi square statistic, the probability of the speech presence period is identified in the noisy frame. During this period, the speech presence probability, smoothing factor and the estimate of noise variance are recursively smoothed then averaged. In this paper, the proposed and exiting algorithms are tested under airport, babble, station, exhibition, restaurant, car, street and train noises with different input SNR levels (0 dB and 5 dB) for the optimal smoothing factors $\alpha_p = 0.97$ and $\alpha_d = 0.7$. From

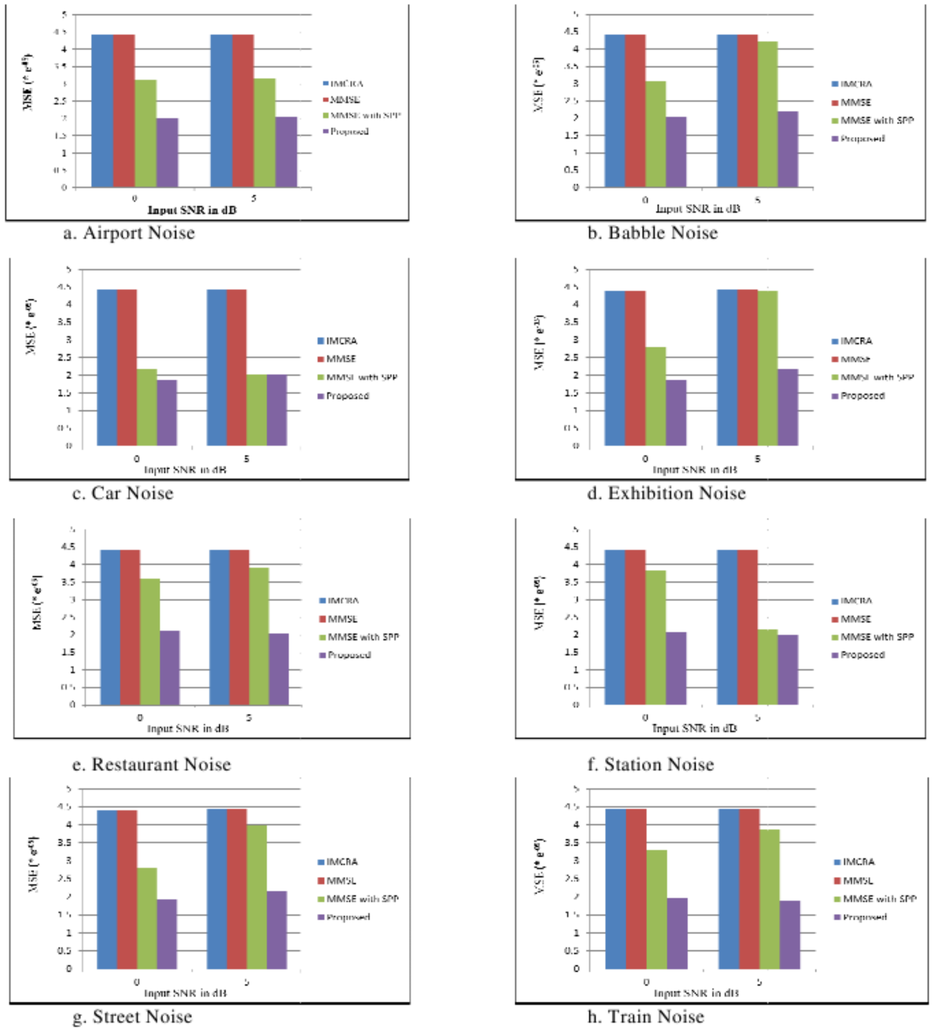

Fig. 2: (a to h) Mean Square Error values for IMCRA, MMSE, MMSE with SPP algorithms and Proposed Noise Tracking Algorithm for various noises with different input SNR levels

the evaluated results, it is seen that the proposed noise tracking algorithm reduces the performance measures as 6% - 58% of MSE and 3% - 97% of LogErr (in dB) as compared to that of the various existing algorithms. These results indicate that the proposed method produces less speech distortion and residual (musical) noise and also it improve the speech signal quality and intelligibility. In addition, it provides the best tracking for non stationary noise conditions.

Acknowledgements

The authors would like to thank the anonymous reviewers for all their valuable comments and suggestions.

References

[1] Timo Gerkmann, Martin Krawczyk, MMSE-optimal spectral amplitude estimation given the STFT-phase, IEEE Signal Processing Letters, **20**, 1–4 (2013).
[2] Timo Gerkmann, Richard C. Hendriks, Unbiased MMSE-Based Noise Power Estimation with Low Complexity and Low Tracking Delay, IEEE Transactions on Audio, Speech, and Language Processing, **20**, 1383–1393 (2012).
[3] Timo Gerkmann, Richard C. Hendriks, Improved MMSE-based noise PSD tracking using temporal cepstrum Smoothing, IEEE International Conference on Acoustics, Speech, and Signal Processing, 105–108 (2012).
[4] F. Chen, P. Loizou, Impact of SNR and gain function over- and under-estimation on speech intelligibility, Speech Communication, **54**, 272–281 (2012).
[5] Philipos C. Loizou, Gibak Kim, Reasons why Current Speech-Enhancement Algorithms do not Improve Speech

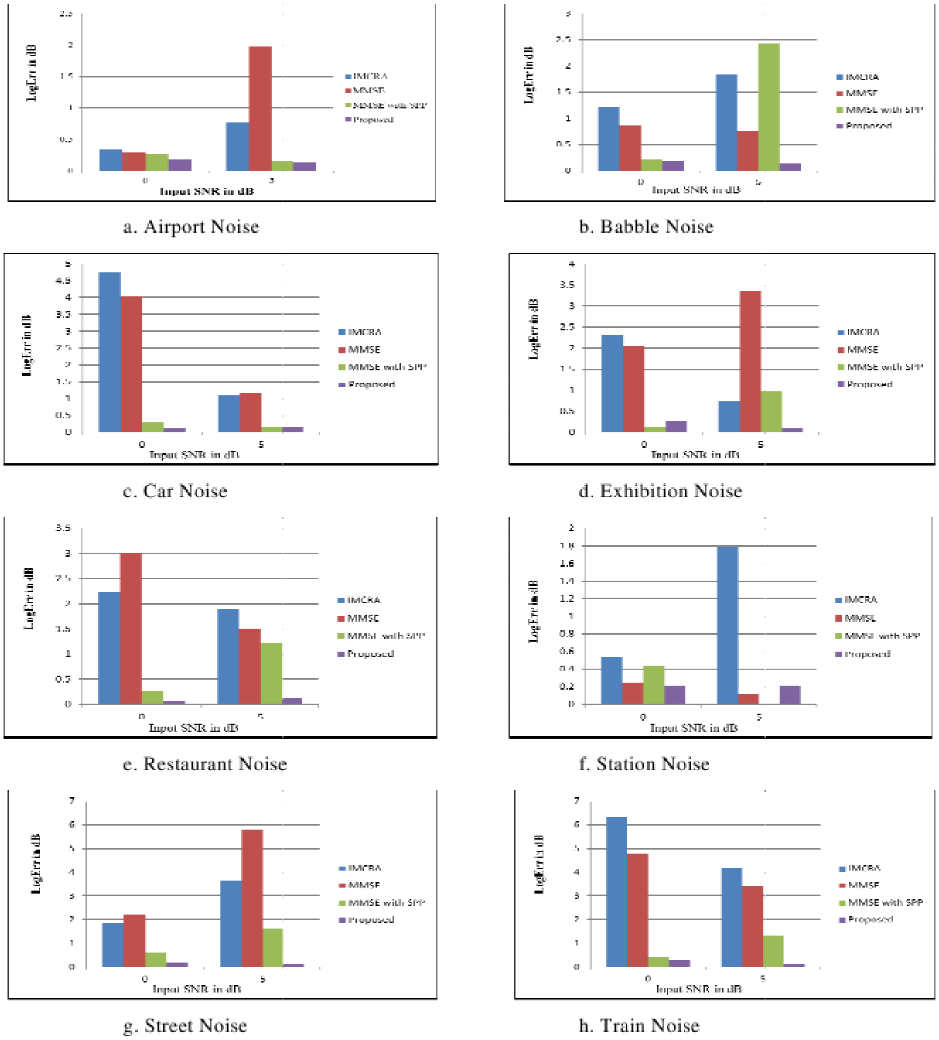

Fig. 3: (a to h) LogErr values in dB for IMCRA, MMSE, MMSE with SPP algorithms and Proposed Noise Tracking Algorithm for various noises with different input SNR levels

Intelligibility and Suggested Solutions, IEEE Transactions on Audio, Speech and Language processing, **19**, 47–56 (2011).

[6] Suhadi Suhadi, Carsten Last, Tim Fingscheidt, A Data-Driven Approach to a Priori SNR Estimation, IEEE Transactions on Audio, Speech and Language Processing, **19**, 186–195 (2011).

[7] Sandhya Hawaldar, Manasi Dixit, Speech Enhancement for Non stationary Noise Environments, Signal & Image Processing International Journal (SIPIJ), **2**, 129–136 (2011).

[8] Timo Gerkmann, Richard C. Hendriks, Noise Power Estimation based on the probability of Speech Presence, IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (2011).

[9] T. Gerkmann, C. Breithaupt, R. Martin, Improved a posteriori speech presence probability estimation based on a likelihood ratio with fixed priors, IEEE Trans. Audio, Speech, Lang. Process, **16**, 910–919 (2008).

[10] Jan S. Erkelens, Richard Heusdens, Tracking of Non-stationary noise based on Data-Driven Recursive Noise Power Estimation, IEEE Transactions on Audio, Speech, and Language Processing, **16**, 1112–1123 (2008).

[11] J.S. Erkelens, R. Heusdens, Fast noise tracking based on recursive smoothing of MMSE noise power estimates, IEEE Int. Conf. Acoust., Speech, Signal Processing, 4873–4876 (2008).

[12] S. Beheshti, N. Nikvand, X.N. Fernando, Soft Thresholding by Noise Invalidation, 24th Biennial Symposium on Communications, 235–238 (2008).

[13] Richard C. Hendriks, Jesper Jensen, Richard Heusdens, DFT domain subspace based noise tracking for speech enhancement, 8th Annual Conference of the International Speech Communication Association, 830–833 (2007).

[14] S. Rangachari, Philipos C. Loizou, A noise-estimation algorithm for highly non-stationary environments, Speech Communication, **48**, 220–231 (2006).

[15] I. Cohen, Speech Enhancement Using Super gaussian Speech Models and Non causal A Priori SNR Estimation, Speech Communication, **47**, 336–350 (2005).
[16] I. Cohen, Speech Enhancement Using a Noncausal A Priori SNR Estimator, IEEE Signal Processing Letters, **11**, 725–728 (2004).
[17] S. Rangachari, Philipos C. Loizou and Yi Hu, A noise estimation algorithm with rapid adaptation for highly non-stationary environments, IEEE International Conference on Acoustics, Speech, and Signal Processing, **1**, I-305–308 (2004).
[18] I. Cohen, Noise spectrum estimation in adverse environments: Improved minima controlled recursive averaging, IEEE Trans. Speech Audio Process., **11**, 466–475 (2003).
[19] Sridhar Krishnan, X.N. Fernando and K.H. Sun, Non-stationary noise cancellation in infrared wireless receivers, in the Proceedings of the Canadian Conference on Electrical and Computer Engineering (CCECE'2003), Montreal, Canada (2003).
[20] I. Cohen, Noise Estimation by Minima Controlled Recursive Averaging for Robust Speech Enhancement, IEEE Signal Processing Letters, **9**, 12–15 (2002).
[21] I. Cohen, B. Berdugo, Speech enhancement for non-stationary noise environments, Speech Communication, **81**, 2403–2418 (2001).
[22] R. Martin, Noise power spectral density estimation based on optimal smoothing and minimum statistics, IEEE Trans. Speech Audio Process., **9**, 504–512 (2001).

M. Kalamani received her B.E. (Electronics and Communication Engineering) from Bharathiar University, Coimbatore and M.E. (Applied Electronics) from Anna University, Chennai in April 2004 and April 2009 respectively. She is currently pursuing her research in the area of Speech signal processing under Anna University, Chennai. She is presently working as Assistant Professor (Senior Grade) in the Department of Electronics and Communication Engineering, Bannari Amman Institute of Technology, Sathyamangalam. She is having 9 years of teaching experience in various engineering colleges. She has published 4 papers in International Journals, 11 papers in International and National Conferences.

S. Valarmathy received her B.E. (Electronics and Communication Engineering) and M.E. (Applied Electronics) degrees from Bharathiar University, Coimbatore in April 1989 and January 2000 respectively. She received her Ph.D. Degree at Anna University, Chennai in the area of Biometrics in 2009. She is presently working as a Professor and Head of the Department of Electronics and Communication Engineering, Bannari Amman Institute of Technology, Sathyamangalam. She is having 21 years of teaching experience in various engineering colleges. Her research interest includes Biometrics, Image Processing, Soft Computing, Pattern Recognition and Neural Networks. She is the life member of Indian Society for Technical Education and Member in Institution of Engineers. She has published 38 papers in International and National Journals, 68 papers in International and National Conferences.

M. Krishnamoorthi received his B.E. (Electrical and Electronics Engineering) from Bharathiar University, Coimbatore and M.E. (Computer Science and Engineering) (Computer Science and Engineering) from Annamalai University, Chidambaram in April 2002 and April 2004 respectively. He is currently pursuing his research in the area of Data Mining and Optimization algorithms under Anna University, Chennai. He is presently working as Assistant Professor (Senior Grade) in the Department of Computer Science and Engineering, Bannari Amman Institute of Technology, Sathyamangalam. He is having 9 years of teaching experience. He has published 3 papers in International Journals, 7 papers in International and National Conferences.