Performance Evaluation of Noise Estimation Techniques for Blind Source Separation in Non Stationary Noise Environment

Ms. Mohini Avatade[1], Prof. Mr. S.L. Sahare[2]

[1,2]Electronics & Telecommunication Department,
Cummins College of Engineering for Women, Pune 411052, Maharashtra, India

Abstract

The Noise estimation Technique plays very important role in any speech denoising algorithm. Accuracy and intelligibility of denoised speech signal is mainly affected by noise estimation process. This paper basically deals with performance evaluation of various frequency domain noise estimation methods in the scenario where, microphone receives mixture of desired signal and non stationary noise signal. The primary solution for denoising of mixture is to adapt any one of the blind source separation process such as independent component analysis which extract statistically independent source components. Due to some artifacts in ICA techniques, small amount of residual noise will remain in extracted sources. The effectiveness and performance evaluation of noise estimation techniques is carried out over residual noise along with wavelet Thresholding. Comparison of Martin's , Minima controlled recursive averaging (MCRA), Improved MCRA (IMCRA), MCRA2, Spectral Minima Tracking method is done based on various, speech enhancement objective parameters such as Log-Likelihood Ratio (LLR), Segmental Signal to Noise Ratio (SNRseg), Weighted Spectral Slope (WSS), Perceptual Evaluation of speech Quality (PESQ), Itakura-Saito (IS) Ratio

Keywords*: Blind Source separation, Non stationary Noise, Noise estimation Technique, Objective Quality Measures.*

1. Introduction

Noise is very crucial part of while developing any speech enhancement algorithm. Noise affects intelligibility as well as quality of speech signal via channel noise, additive noise, Non stationary noise etc. the noise estimation techniques plays very important role in speech denoising process, by extracting Noise power spectrum from noisy speech signal. This paper has significance in the area where, one has to make choice between various noise estimation methods particularly in blind source separation scenario. Independent component analysis is one of the very effective solutions to blind source separation, which deals with higher order statistics of speech and extracts statistically independent components from mixture of two or more signals. We have used FASTICA algorithm [2] along with the kurtosis fourth order cumulant method, which effectively differentiate between various independent components [3]. There may be possibility of presence of small part of noise in extracted independent components. So, here is a very critical matter to adapt appropriate frequency domain noise estimation algorithm which works well in this critical condition where noise is a very small part of speech signal. Minimum statistics and Minima controlled Recursive Averaging are main principles of noise estimation [6][7]. These methods have been adapted to obtain noise spectrum, according to estimated noise value, the threshold value is varying. The adaptive wavelet domain Thresholding [1] is carried out on independent components with their respective estimated threshold values. The efficiency of noise estimation algorithms is evaluated based on objective speech evaluation parameters [9].Finally comparison [10] is made for Minimum Statistic Method, Minima Controlled Recursive Averaging (MCRA), MCRA2, Improved MCRA, Spectral Minima Tracking in sub bands (Doblinger) method based on an Objective as well as subjective parameters with respect to different types of Non stationary Noise signals. The further paper is aligned as follows: II. Independent Component Analysis III. Noise Estimation Techniques. IV Simulation Results V. Conclusion.

2. Independent Component Analysis

Independent component analysis (ICA) is a very effective mechanism for numerous applications such as blind source separation (BSS), unsupervised learning, feature extraction, and data compression. However, ICA finds a set of components that are non-Gaussian and mutually independent. Independent component analysis was originally developed to deal with problems that are closely related to the cocktail-party problem. The Basic ICA Model is given by following equation:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{v}(t) \tag{1}$$

Where, s(t)= Clean speech, A= Mixing Matrix, v(t)= Noise signal. The basic problem of ICA is then to estimate the realizations of the original speech signals using only observation of the mixture $\mathbf{x}(t)$, let us denote $\mathbf{W}$, and obtain the independent component simply by

$$s\widetilde{(t)} = W\mathbf{x}(t) \tag{2}$$

2.1 The FASTICA Algorithm:
The FASTICA proposed by Hyvanrinen is based on a fixed-point iteration scheme. Here we adopted kurtosis as the estimation rule of independence. Kurtosis has widely used as a measure of non-Gaussianity in ICA, which can be estimated simply by using the fourth moment of the sample data. Kurtosis is defined as follows:

$$\mathrm{Kurt}(S_i) = E[S_i^4] - 3(E[S_i^2])^2 \quad (3)$$

We erect adjective function:

$$\mathrm{Kurt}\,(w^T\tilde{x}i) = E[(w^T\tilde{x}i)] - 3[E\{(w^T\tilde{x}i)^2\}]^2 \quad (4)$$

Since the observation signal has been pre-whitening, thus equation (8) can be simplified as:

$$\mathrm{Kurt}\,(w^T\tilde{x}i) = E[(w^T\tilde{x}i)^4] - 3\|w\|^4 \quad (5)$$

Seeking the gradient of equation (9), we get the following:

$$\Delta w\alpha\, E\,[\tilde{x}i\,(w_i(k)^T\tilde{x}i)^3] - 3\|wi(k)\|^2\, w_i(k) \quad (6)$$

Using the fixed point algorithm, the iteration of fixed point algorithm can be expressed:

$$w_i(k) = E[\tilde{x}i\,(w_i(k\text{-}1)^T\,\tilde{x}i)^3] - 3w_i\,(k\text{-}1) \quad (7)$$

Thus we obtain the FASTICA algorithm as follows:
(1)Center the data to make its mean zero.
(2)Whiten the data to get $\tilde{x}i$ (t)
(3)Make i=1;
(4)Choose an initial orthogonal matrix foe W and make k=1;
(5)Make $w_i(k) = E\,[\tilde{x}i\,(w_i(k\text{-}1)^T\,\tilde{x}i)^3] - 3w_i\,(k\text{-}1)$
(6)Make $w_i(k) = \dfrac{w_i(k)}{\|w_i(k)\|}$
(7)If not converged, make k=k+1 and go back to step (5)
(8)Make i=i+1
(9)When i<number of original signals, go back to step (4)
Until $|w_i(k)^T\, w_i(k\text{-}1)|$ is equal or close to 1, the iteration finish.

3. Noise Estimation Techniques

3.1 Optimal smoothing and minimum Statistics Method (Martin's Method):

Martin proposed a novel noise estimation algorithm based on an optimal signal power spectral density smoothing method and on minimum statistics. The smoothing algorithm utilizes a first order recursive system with a time as well as frequency dependent smoothing parameter. The smoothing parameter is optimized for tracking non stationary signals by minimizing a conditional mean square error criterion and a bias compensation algorithm for minimum power spectral density estimates. Results with various noise types show that the time varying smoothing significantly improves the minimum statistics approach [6]. An input Noisy speech signal is transformed in the frequency domain by first applying a hamming window function w(n) to M samples of y(n) and then computing the M-point FFT of the windowed signal.

$$Y(\lambda,k) = y(\lambda M + m)w(m)e^{-j2\pi mk/M} \quad (8)$$

Where λ is the frame index and k indicates frequency bin index varient from k = {0, 1, 2 ... M-1}. Y(λ, k) is the short term Fourier Transform (STFT) of y(n). Periodogram of the noisy speech is approximately equal to the sum of periodogram of clean and noise signal given as

$$|Y(\lambda,k)|^2 \approx |X(\lambda,k)|^2 + |D(\lambda,k)|^2 \quad (9)$$

Where $|Y(\lambda, k)|^2$ is the periodogram of noisy speed signal, $|X(\lambda, k)|^2$ is the periodogram of clean speed signal and $|D(\lambda, k)|^2$ is the periodogram of Noise signal.
Following are the steps to obtain an estimate of the power spectrum of the noise by tracking the minimum of smoothed power spectrum P (λ, k).

1) Spectral Analysis of noisy speech signal with window FFT analysis results in a set of frequency domain signals which can be written as:

$$Y(\lambda,k) = \sum_{\mu=0}^{L-1} y(\lambda R + \mu)h(\mu)e^{-j2\pi k\mu/L} \quad (10)$$

Where λ is the sub sampled time index, and *K* is the frequency bin index, K Є {0,1,2……L-1}

2) Compute smoothing parameter: The multiplication of the correction factor with the optimal smoothing parameter then yields the final smoothing parameter $\hat{\alpha}$(λ,k)

$$\hat{\alpha}(\lambda,k) = \frac{\alpha \max \alpha c(\lambda)}{1 + (P(\lambda-1,k)/\hat{\sigma}^2\,(\lambda-1,k) - 1)^2} \quad (11)$$

3) Compute smoothed power: The recursive smoothed periodogram is considered to highlight some of the obstacles which are encountered in such an approach

$$P(\lambda,k) = \alpha(\lambda,k)P(\lambda-1,k) + (1-\alpha(\lambda-k))|Y(\lambda,k)|^2 \quad (12)$$

Where α is the smoothing constant. The above recursive equation in recognized as Low pass filter, which provides a smoothed version of periodogram $|Y(\lambda, k)|^2$.

4) To compute bias correlation: approximate the inverse mean of the minimum by

$$\beta \min(\lambda,k) \approx 1 + (D-1)\frac{2}{\hat{Q}eq(\lambda,k)}\Gamma(1 + \frac{2}{\hat{Q}eq(\lambda,k)})^{H(D)} \quad (13)$$

Where, the inverse normalized variance $\hat{Q}eq(\lambda,k)$ is also called "equivalent degrees of freedom" since (moving average) smoothing of $\hat{Q}eq(\lambda,k)$ independent squared Gaussian variates would yield an estimate with the same variance. Where $\hat{Q}eq(\lambda,k)$ is a scaled version of $Qeq(\lambda,k)$

5) An optimized results were obtained by choosing the smoothing parameter β(λ,k)= α^2(λ,k) and by limiting β(λ,k) to values less or equal to 0.8 [6] Finally, $1\backslash Qeq(\lambda,k)$ is estimated by,

$$\frac{1}{Qeq\,(\lambda,k)} \approx \frac{\widehat{var}\,\{P(\lambda,k)\}}{2\hat{\sigma}^4(\lambda-1,k)} \tag{14}$$

3.2 Minima Controlled Recursive Averaging (MCRA)

Cohen Proposed MCRA Noise estimation algorithm. The commonly used Principle for noise spectral estimation is Recursive averaging Process. It is very effective method rather than employing Voice Activity Detector based techniques which restricts the updates of noise estimator to the particular periods of speech absence. MCRA method derives smoothing parameter in time as well as frequency according to speech presence probability which is again controlled by minima values of smoothed periodogram of noisy speech signal. According to method explained in [7], the conditional speech presence probability $\hat{p}$ (λ, k) is computed by comparing the ratio of the noisy speech power spectrum to its local minimum against a threshold value. This Algorithm is called as Minima Controlled Recursive Averaging Algorithm (MCRA) due to reason that, probability estimate value $\hat{p}$ (λ, k) and the time smoothing factor α (λ, k), is controlled by the estimate of spectral minimum. [11]This Algorithm is modified by researchers and some of them are MCRA-2 Algorithm explained in [8], improved MCRA Algorithm explained in [9].
The basic MCRA noise estimation algorithm steps are as follows:

1) Noise Spectrum Estimation

$$Y(k,l) = \sum_{n=0}^{N-1} y(n+lM)h(n)e^{-j(2\pi/N)nk} \tag{15}$$

Where K is the frequency bin index, l is the time frame index, h is an analysis window of size, M and N is the frame update step in time.

2) Calculate Signal Presence Probability
In frequency Domain Representation, we use a window function whose length is $2w+1$

$$Sf(k,l) = \sum_{i=-w}^{w} b(i)|Y(k-i,l)|^2 \tag{16}$$

In time Domain, the smoothing is performed by a first order recursive averaging, given by

$$S(k,l) = \alpha S(k,l-1) + (1-\alpha)Sf(k,l) \tag{17}$$

Where, $\alpha(0 < \alpha < 1)$ is a smoothing parameter.

3) Compute the ratio between the local energy of the noisy speech and its derived minimum. A Bayes minimum-cost decision rule is given by:

$$\frac{p\,(S_r|H_1)}{p\,(S_r|H_0)} \underset{H_0'}{\overset{H_1'}{\gtrless}} \frac{c_{10}P\,(H_0)}{c_{01}P\,(H_1)} \tag{18}$$

Where $P(H_0)$ and P (H_1) are the *a priori* probabilities for speech absence and presence, respectively.

4) The following is an estimator function for $\hat{p}(k,l)$

$$\hat{p}(k,l) = \alpha\hat{p}(k,l-1) + (1-\alpha)I(k,l) \tag{19}$$

Where, $I(k,l)$ denotes an indication function.

4. Simulation Results

The Simulation is carried out on NOIZEUS database (sp01.wav: The birch canoe slid on the smooth planks) clean signal Mixed with Various Non stationary noise signals from SpEAR database. The independent component analysis separates original source signal, which is again denoised using different Noise estimation techniques. Table 1, Table2, Table3 Corresponds to Objective Quality measure parameters [10] of denoised speech signals evaluated under influence of Car noise, Factory Noise and Pink Noise respectively for various Noise estimation Techniques listed below:

	Log Likelihood Ratio (LLR)	Segmental Signal to Noise Ratio (SNRseg)	Weighted Spectral Slope (WSS)	Perceptual Evaluation of speech quality (PESQ)	Itakura-Saito Ratio (IS Ratio)
Initial Parameters	1.0858	-6.2871	89.0565	1.7886	36.5663
Martin	0.1231	-6.0759	10.7062	4.2657	4.1139
MCRA	0.1257	-6.0701	10.8195	4.2612	4.5030
MCRA2	0.1437	-6.0274	12.0165	4.2205	8.5063
IMCRA	0.1243	-6.0734	10.7603	4.2639	4.2996
Doblinger	0.2179	-5.8280	22.2925	4.0364	17.6289
Hirsch	0.1232	-6.0757	10.7097	4.2657	4.1280

Table I. Comparison of Noise estimation methods in case of Volvo noise

	Log Likelihood Ratio (LLR)	Segmental Signal to Noise Ratio (SNRseg)	Weighted Spectral Slope (WSS)	Perceptual Evaluation of speech quality (PESQ)	Itakura-Saito Ratio (IS Ratio)
Initial Parameters	0.9809	-6.3403	92.8873	1.5035	4.7853
Martin's	0.1190	-6.0745	7.5846	4.2418	3.6027
MCRA	0.1222	-6.0678	7.7164	4.2350	3.9751
MCRA2	0.1400	-6.0264	8.6911	4.1934	7.3994
IMCRA	0.1207	-6.0710	7.6642	4.2382	3.8191
Doblinger	0.2158	-5.8224	18.6970	4.0068	16.9864
Hirsch	0.1196	-6.0732	7.6148	4.2406	3.6710

Table II. Comparison of Noise estimation methods in case of Factory noise

IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 3, No 2, May 2013
ISSN (Print): 1694-0814 | ISSN (Online): 1694-0784
www.IJCSI.org

	Log Likelihood Ratio (LLR)	Segmental Signal to Noise Ratio (SNRseg)	Weighted Spectral Slope (WSS)	Perceptual Evaluation of speech quality (PESQ)	Itakura-Saito Ratio (IS Ratio)
Initial Parameters	1.1943	-6.3542	83.8480	1.4889	3.2883
Martin's	0.1174	-6.1511	8.8775	4.1852	0.4966
MCRA	0.1245	-6.1296	9.0484	4.1708	0.7208
MCRA2	0.1365	-6.0952	9.4547	4.1405	1.2911
IMCRA	0.1198	-6.1438	8.9346	4.1806	0.5613
Doblinger	0.2176	-5.8793	18.7042	3.9123	10.1160
Hirsch	0.1179	-6.1495	8.8888	4.1843	0.5091

Table III. Comparison of Noise estimation methods in case of Pink noise

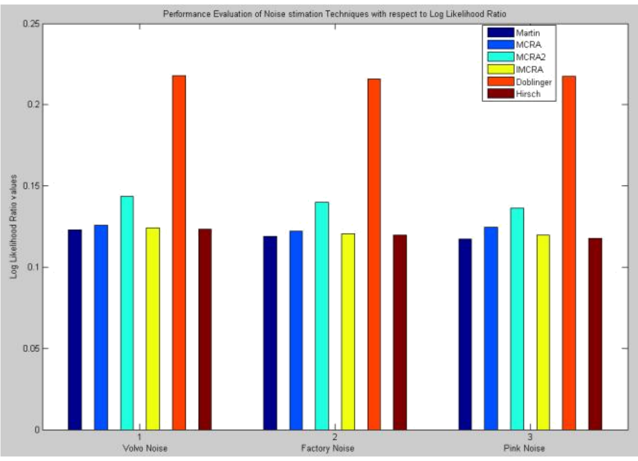

Figure 1. Graph of LLR values for different noise estimation methods in the effect of Volvo, Factory and Pink Noise

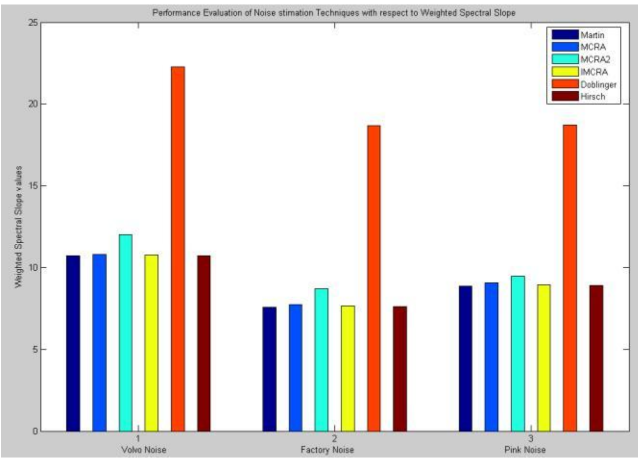

Figure 2. Graph of WSS values for different noise estimation methods in the effect of Volvo, Factory and Pink Noise

Respective Algorithms are simulated on MATLAB version-7.0.Simulation results are used for evaluation of quality of enhanced speech signal. The intelligibility and speech quality measures reflect the true performance of any speech enhancement algorithm as well as Noise estimation Technique. Quality assessment is done using subjective and

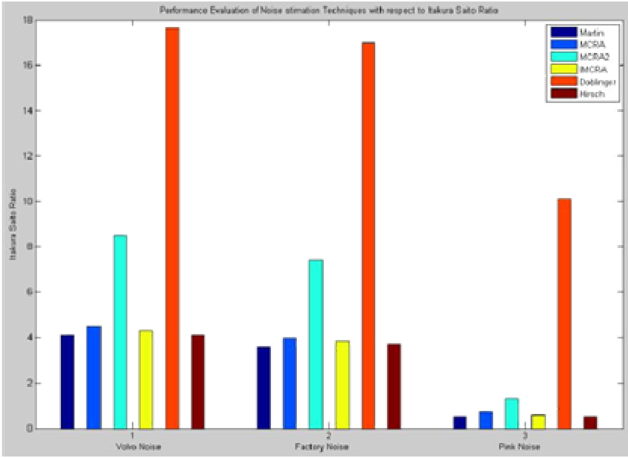

Figure 3. Graph of IS Ratio values for different noise estimation methods in the effect of Volvo, Factory and Pink Noise

Objective quality measures, which involve a mathematical Model, used for comparison of the original and processed Speech signals. Objective measures signify Quality of enhanced speech by measuring the numerical distances between original and processed speech signal [10]. Log-Likelihood Ratio (LLR), Itakura- Saito ratio, Weighted Spectral Slope , Segmental Signal to Noise ratio are an objective quality measures corresponding to each frame of speech signal. Perceptual Evaluation of speech quality (PESQ) is subjective quality measure parameter; this is estimated with the help of various listening tests and Mean opinion score (MOS) of corresponding tests. PESQ measures suitable mainly for predicting signal distortion, noise distortion and overall speech quality. LLR provides distance between two frames by means of Log function of auto correlation ratio of corresponding clean and processed speech. The IS ratio measures distance between two frames based on various spectral levels in signal. Weighted Spectral Slope is obtained as difference between current and adjacent spectral magnitudes. Small values of LLR, IS and WSS are required for better quality enhanced signal. The Statistics shown in bar graphs indicates performance level of every Noise estimation method verified for Volvo, Factory and Pink Noise. Martin's Minimum Statistics process works better and provides satisfied level of parameters as compared with other competitive methods.

5. Conclusions

Noise Estimation is very crucial and important process in case of Blind Source Separation Problem. Independent Component Analysis with kurtosis function produces statistically independent components with some small amount of residual noise. Comparing various Noise estimation techniques based on objective parameters we concluded that, Martin's Minimum statistic algorithm works better in this particular case of Blind source separation. We obtained significant improvement in frequency domain parameters with the minimum statistic method whereas in a case of speech enhancement algorithm in which Segmental Signal to Noise Ratio is major significant parameter, Spectral Minima tracking in sub band (Doblinger's Method) Produces better results.

Acknowledgements

Mohini Avatade Thanks to, Electronics and telecommunication department of Cummins college of Engineering.

References

[1] Hongyan Li, Huakui Wang, Baojin Xiao," Blind separation of noisy mixed speech signals based on wavelet transform and Independent Component Analysis", 0-7803-9737-1©2006 IEEE.

[2] Li Hongyan, Ren Guanglong, "Blind separation of noisy mixed speech signals based Independent Component Analysis", 2010 First International Conference on Pervasive Computing, Signal Processing and Applications, 978-0-7695-4180-8/10© 2010 IEEE.

[3] Qingyun Wang, Hui Zong, Liye Zhao, "Speech extraction method based on multiple reference signals ICA algorithm", 2011 International Conference on Transportation, Mechanical, and Electrical Engineering (TMEE) December 16-18, Changchun, China, 978-1-4577-1701-7/11/©2011 IEEE.

[4] Doblinger, G., 1995. "Computationally efficient speech Enhancement by spectral minima tracking in subbands" Proc.Euro speech 2, pp. 1513–1516

[5] Hirsch, H., Ehrlicher, C,(1995) "Noise estimation Techniques for robust speech recognition. Proc. IEEE International Conference on Acoustic Speech Signal Processing pp 153–156.

[6] Martin, R.(2001) "Noise power spectral density estimation based on optimal smoothing and minimum statistics". IEEE Trans. Speech Audio Process 9 (5), pp. 504–512

[7] Cohen, I., 2002. "Noise estimation by minima controlled recursive averaging for robust speech enhancement", IEEE Signal Proc. Letter 9 (1), pp.12–15

[8] Loizou P, Sundarajan R. (2006) "A Noise estimation Algorithm for highly non–stationary Environments", speech Communication 48 (2006) Science direct pp. 220-231

[9] Israel Cohen," Noise Spectrum Estimation in Adverse Environments: Improved Minima Controlled Recursive Averaging", IEEE transactions on speech and audio processing, vol. xx, no. y, month 2003.

[10] Yi Hu and Philipos C. Loizou, Senior Member, *IEEE, "*Evaluation of Objective Quality Measures for Speech Enhancement*",* IEEE transactions on audio, speech, and language processing, vol. 16, no. 1, january 2008, 1558-7916/© 2007 IEEE.

[11] Anuradha R. Fukane, Shashikant L. Sahare, "Noise estimation Algorithms for Speech Enhancement in highly non-stationary Environments", IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 2, March 2011, ISSN (Online): 1694-0814.

Mohini S. Avatade received Diploma in Electronics and Telecommunication Engg in 2004 from Government Polytechnic Pune, Completed Bachelors Degree in Electronics and Telecommunication in 2010, currently doing masters in Engg. (M.E.) in signal processing from Pune University ,will receive masters Degree in 2013.Associated with Pad. Dr. D. Y. Patil Institute of Technology since 2012. Associate Member of IETE. Area of interest are Speech Signal processing, Digital Signal Processing. Currently working on Different Speech Denoising and enhancement Algorithms.

Shashikant L. Sahare received bachelor's Degree in Electronics Engg. in 2001, Master's Degree (M. Tech.) in Electronics Design and Tech. in 2004 from Center for Electronics Design Tech. Aurangabad . PhD pursuing in from university of pune. Associate with Cummins College of Engineering for Women Pune, Maharashtra, India since 2004. Three papers are published in National Conferences and one paper is published in International Conference. Areas of interest are Signal processing, Electronic Design. Currently working as Assistant Professor at Cummins College of Engineering for Women Pune, Maharashtra, India.