# An Improved Bi-Level Thresholding Based Uncertainty Evaluation for Speech Enhancement in Non-Stationary Noises

**S. Nageswara Rao[1], K. Jaya Sankar[2], C.D. Naidu[3]**

*[1]Assistant Professor, Department of ECE, Sri Venkateswra Engineering College, Suryapet, Telangana, India*
*[2]Prof. & Head, Department of ECE, Vasavi College of Engineering, Hyderabad, India*
*[3]Professor, Department of ECE, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, India*

## Abstract

This paper proposes a new speech enhancement framework to improve the quality of speeches recorded under adverse acoustic environments based on the speech presence uncertainty. Since the uncertainty evaluation gives a more and clear discrimination about the speech and noise, this paper proposes a new uncertainty evaluation mechanism as a preprocessing mechanism to the noise suppression methods. This mechanism relates with energies of a noisy speech signal and classifies the speech segments and noise segments more perfectly. In addition to the quality enhancement, this approach also reduces the unnecessary computational burden over the speech processing system. Extensive simulations are carried out over the speech signals with different types of non-stationary noises like babble noise, exhibition noise, restaurant noise and train station noises and the performance is measured with the performance metrics namely the Output SNR, AvgSegSNR, PESQ and COMP. The comparative analysis of proposed approach over the conventional approaches shows an outstanding performance in all environments.

*Keywords: Speech enhancement, Noise estimation, EMD, Thresholding, Babble noise, Output SNR, and PESQ.*

## 1. Introduction

In recent years, due to the rapid growth in the speech oriented applications including Automatic Speech Recognition (ASR), modern mobile communications and human-computer interaction systems through voice communications, speech enhancement has become more essential component [1], [4], [5]. All these speech based applications becomes ineffective in the presence of a speech signal with unnecessary disturbances. For example, in the hands free telephony speech communication systems, the microphones are typically placed at certain distant from the speaker's mouth. In such cases, various noise sources makes the speech signal corrupted, by which the performance of a speech oriented devices are not able to process the signal effectively. Generally the speech enhancement techniques consists of noise or disturbance estimation algorithms by which the speech quality and intelligibility increases significantly. However the complete elimination of noises form a noise contaminated speech signal results in the loss of speech related information also. Hence there is need to design an efficient speech enhancement approach with less information loss followed by greater noise removal.

Several approaches have been developed in earlier to perform speech enhancement and basically they are grouped into two classes, they are subspace methods [2] or time domain methods and frequency domain methods [3]. In the case of subspace methods, the speech enhancement is applied directly over the speech signals whereas in the transform domain, the speech enhancement is applied over the transformed values of speech signal. Both the methods have their own advantages and disadvantages.

For instance the subspace methods are simple but the information loss is observed to be high due to the direct manipulations over the speech signals. On the other hand, the transform domain approaches are more efficient but they consume high computational resources. Further some more methods are developed based on the evaluations of signal and noise uncertainties i.e., measuring the probabilities of the existence or non-existence of speech or noise in the speech signals frames. In such types cases, the speech signals are subjected to uncertainty evaluation and based on the uncertainties, a simple noise filtering technique is proposed to remove the noise more effectively form every frame.

This paper proposes a novel speech enhancement approach based on the signal uncertainty in Empirical Mode Decomposition (EMD) domain. The enhancement in the probability of detection provides a clear discrimination between the speech and noise by which the quality of speech can be increased by only passing the noise dominant samples for further noise suppression. To achieve this objective, a perfect thresholding strategy is required and this paper proposed the improved bi-level threshold estimator logic through the energy characteristics of speech signals. Compared with a single threshold, a dual threshold can achieve more effective probability of detection followed by a reduced probability of false alarm. The proposed logic also tries to reduce the missing probability by which the information loss reduces effectively. Extensive simulation are carried out over the proposed approach through different speech samples contaminated with different noises and at every test, the performance is measured and compared with conventional approaches.

Rest of the paper is organized as follows: Section II describes the literature survey details. Section III illustrates the details of proposed approach. Details of simulation results are described in section IV and the conclusions are provided in section.

## 2. Literature Survey

Several recent studies have proposed various noise estimation algorithms for unknown noise signals. Among the most established estimators are those based on minimum statistics [6], [7]. For instance, in [6] the power of the noisy signal is estimated and observed over a time-span of about 1-3 seconds. The spectral noise power is then inferred from the minimum of the estimated power of the noisy signal, assuming that speech is absent at least for a short duration within the observed time-span. However, if the noise power rises within the observed time-span, the noise power will be underestimated. While in [7] mechanisms are proposed that allow for a tracking of rising noise powers within the observed time-span, rising noise powers as caused e.g. by passing cars, are usually tracked with a rather large delay. The local underestimation of the noise power is likely to result in annoying artifacts, so-called *musical noise*, when the noise power estimate is applied in a speech enhancement framework. The further most popular techniques are established based on the Minimum Mean Square Error (MMSE) [8], [9]. In the MMSE based estimator [9], first a limited maximum likelihood (ML) estimate of the *a priori* signal-to noise ratio (SNR) is used to estimate the periodogram of the noise signal [10]. However, this simple estimate results in a bias, which is then compensated based on a second estimate of the *a priori* SNR. Though, the second estimate provides the compensated results, the first bias results in extra complexity.

Some more approaches are developed in earlier by assuming a statistics signal distribution and processing the entire signal in the short time Fourier transform (STFT) [11]. However in the algorithms, speech is assumes to be present in the every time-frequency bin, which reduces the performance. Hence some approaches are focused to analyze the speech signal in the short-time spectral amplitude (STSA) through the signal presence probability or uncertainty [12], [13] which provides more efficient performance in the noise estimation because the gain function is multiplied with the probability of speech presence. In this case some approaches tries to detect the signal and some tried to estimate the noise through Log-Spectral Amplitudes (LSA) [14]. In [13] combined detection and estimation approach is proposed for improving performance. This method used a non-continuous cost function based on the spectral magnitude square error, which was not enough subjectively. A new log spectral amplitude estimator based on joint detection and estimation theory is proposed in [15] to achieve efficient subjective evaluation. By defining the cost function on the log-spectral amplitude error, we determine a gain function in the form of a generalized binary mask, which enables improved speech intelligibility [16]. Recently the binary mask techniques have been proposed as a signal processing tool to study and analyze the time-frequency analysis and grouping process of the auditory system [17]-[20]. However, perhaps motivated by the significant intelligibility improvements achievable in this ideal setting, where local target-to-noise energy ratios are known with certainty, the binary mask framework has more recently been adapted to the practical problem of retrieving a target speech signal from a noisy mixture in the non-ideal situation where the local target-to-noise energy ratios are unknown, but must be estimated from the noisy mixture signal.

On the other hand, the thresholding based noise suppression is also attained a much effective results in the speech enhancement research [21], [22]. Different preprocessing approaches including the wavelet transform [24], Empirical Mode decomposition [24], Discrete Cosine transform [23] etc., are applied over the speech signal to study the time-frequency characteristics and then deriving a novel threshold to suppress the external noise in the noise contaminated speech signal. A new probabilistic speech enhancement filter is presented in [25] considering the three state possibilities of discrete cosine transform (DCT) coefficients of noisy speech: speech absence, speech and noise are constructive, and destructive. In this case, the design of an optimal threshold to make define the signal is constructive (speech presence) or destructive (speech absence) is very important, because the wrong assumption results in the speech related information loss. A novel threshold design strategy is proposed in [26] through the Generalized signal detector [27] based on the speech signal uncertainty. However the threshold derived in this approach is a binary threshold by which the signal presence and absence are detected with high probability of false alarm. Further the proposed speech enhancement mechanism in [22] assumes that the signal follows a chi-square distribution in which the samples are assumed to be independent, but in real there exists a strong correlation between speech samples.

## 3. Proposed Approach

This section illustrates the details of proposed speech enhancement technique. Initially the basic speech enhancement process based on the Generalized Signal Detector (GSD) is described and next, the determination of optimal threshold value is described clearly. Here the proposed speech enhancement process applies EMD over the noisy speech signal and then obtained IMFs are processed for noise estimation. Considering the speech uncertainty as a main contribution this approach estimates the noise samples in every IMF, perfectly.

### A. Generalized signal detector [27]

Let y(n) be a noisy speech signal that is the sum of clean speech signal s(n) and a noise signal d(n); y(n)=s(n)+d(n). Apply empirical mode decomposition (EMD) over the noisy speech and divide the every IMF into K segments with every segment having equal number of samples.

$$Y(m,k) = S(m,k) + D(m,k) \tag{1}$$

Where $m = 1,2,\dots,M$ is the frequency bin and the $k = 1,2,\dots,K$ is the segment index. $S(m,k)$ and $D(m,k)$ denotes the speech signal and noise signal of $m_{th}$ frequency bin in the $k_{th}$ segment, respectively.

Consider two hypotheses $H_0(m,k)$ and $H_1(m,k)$ which indicates the absence of speech and presence of speech respectively, as

$$\begin{aligned} H_0(m,k): Y(m,k) = D(m,k) \\ H_1(m,k): Y(m,k) = S(m,k) + D(m,k) \end{aligned} \tag{2}$$

In the case of $H_0(m,k)$, the speech signal is absent and only the noise signal is present. Whereas in the case $H_1(m,k)$ denotes the presence of both speech signal and noise signal.

Let $\sigma_d(m,k) = E[|D(m,k)|^2]$ be the noise variance in the $k_{th}$ segment, the speech signal can be detected by the Generalized signal detector (GSD) [27] as

$$W' = \sum_{i=1,k=constant}^{N} \left(\frac{Y(i,k)}{\sigma_d(k)}\right)^p \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \delta' \tag{3}$$

Where N is the total number of samples in $k_{th}$ segment, $p > 0$ is an arbitrary constant and $\delta'$ is the threshold to be determined. Based on the obtained GSD, the probability of detection and probability of false alarm are determined as

$$P_d = \Pr(W' > \delta'|H_1) \tag{4}$$
$$P_F = Pr(W' > \delta'|H_0) \tag{5}$$

Here the $P_d$ is defined as the probability of correct detection, i.e., the result shown by eq.(3) and the assumption are same. Otherwise it results under $P_F$.

In the earlier GSD based speech enhancement approach, only one threshold is determined to differentiate the signal dominant frequency bins form noise dominant frequency bins. Though the GSD achieves an efficient performance, the high correlation between the speech signal and noise results in the increased false

alarm rate. i.e., the noise dominant frequency bins can be classified as signal dominant frequency bins and vice versa. In such case the probability of detection becomes poor. Hence this approach develops a new signal detector through the improved Bi-level Estimator logic.

## B. Improved Bi-level Thresholding

In order to obtain an improvement in the estimator logic, bi-level estimator logic is presented in this paper. In [27], a two level threshold approach is presented. In two level threshold approach, the energy detector operates by considering discrete samples of the IMFs and process them to form a test statistic, and later these test statistics are compared to a pre-calculated threshold, where $\delta^`$ is defined as the predetermined limit for the signal-threshold energy detectors and $\delta_0^`$ and $\delta_1^`$ are for the bi-level threshold ones. When $W' \geq \delta^`$ it does mean that the speech is presence by which the hypothesis is $H_1$, there may be a chance of very closer to the threshold such that it denotes the presence of speech but closer to the threshold. In other case, a similar assumption can be a valid but in the opposite direction, i.e., the value denotes the absence of speech but closer to the threshold. Otherwise, it represents the absence of speech which is the hypothesis $H_0$. Hence the proposed improved Bi-level estimator logic introduces two thresholds $\delta_1^`$ and $\delta_0^`$, and the signal detection based on their energy values is determined through the mechanism shown in figure.2
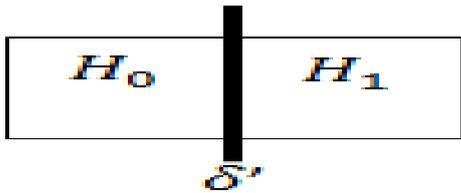


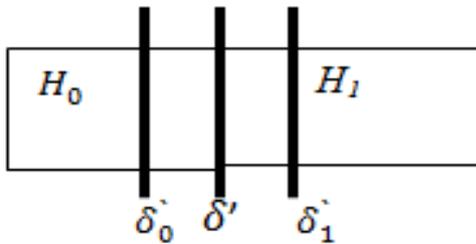**Fig. 1:** conventional threshold approach for energy estimator [27]



**Fig. 2:** Proposed Bi-level energy estimator logic

According to the figure.2, in the proposed Improved Bi-level Estimator, under the condition when $W'>\delta_1^`$, it indicates the hypothesis $H_1$, and when $W' \leq \delta_0^`$, it indicates the hypothesis $H_0$. When $\delta_0^` < W^` \leq \delta_1^`$ indicates the indeterminacy region. This indeterminacy region results in the miss probability, $P_m$ and can be defined as

$$P_m = \Pr\{\delta_0^` < W' < \delta_1^`\} \tag{6}$$

In the above expression, the two threshold are derived based on the single threshold value $\delta^`$ through an arbitrary scalar, denoted as $\beta$, where $\beta \geq 1$, consequently defining $\delta_0^` = \delta^`/\beta$ and $\delta_1^` = \delta^` * \beta$ [28]. Here the improved bi-level estimator logic provides a clear distinction between the $H_1$ and $H_0$.

In this case, the determination of signal uncertainty is more critical and also important because if there is a large number of samples are in the same situation, the probability of detection will be suffered. In such case, the correlation can give relation with this problem. Finding the correlation between the current segment and its neighboring segments will provides the required information

about the uncertainty detection. Depends on the earlier segments state ($H_1$ or $H_0$), the current segment can be determined easily even though it is under the constraint shown in eq.(6).

For each segment $k$, for $k=1,...,K$, perform the signal uncertainty evaluation individually, i.e., the energy detection with a results of $W'_k$. If $\delta_0^` < W'_k < \delta_1^`$, then the $kth$ segment is determined as a speech present or speech absent according to the fusion value obtained by fusing the energy detections of all the preceding segments, $1$ to $k-1$. Depends on the preceding segments Hypothesis Strategies, the current $kth$ segment is determined. Here we use $R_k$ to denote the information that the fusion strategy acquired from the kth segment and it is defined by

$$R_k = \begin{cases} W'_k & \delta_0^` < W'_k < \delta_1^` \\ L_k & Otherwise \end{cases} \tag{7}$$

Where $L_k$ is the local decision for kth segment and is defined as

$$L_k = \begin{cases} 0 & 0 < W' \leq \delta_0^`, \\ 1 & W^` > \delta_1^`, \end{cases} \tag{8}$$

Based on the obtained local decision, the fusion strategy defines an upper bound according to the *N-k* energy detection values, which is given by

$$D = \begin{cases} 0 & 0 \leq \sum_{i=1}^{N-k} W'_i < \delta^` \\ 1 & \sum_{i=1}^{N-k} W'_i > \delta^` \end{cases} \tag{9}$$

Where $\delta^`$ is the energy detection threshold value of the fusion strategy according to the appropriate false alarm rate, defined in equation (5). Further to achieve the optimal decision static for the kth segment, the SNR of every segment is considered and the distribution is formulated according to the energy detection theory [29] as

$$\sum_{i=1}^{N-k} W'_i \sim \begin{cases} \chi^2_{2(N-k)u} & H_0 \\ \chi^2_{2(N-k)u}(2\gamma_0) & H_1 \end{cases} \tag{10}$$

Where $\gamma_0 = \sum_{i=1}^{N-k} \gamma_i$ is the sum of SNR for all the *N-k* segments. Finally the fusion strategy makes a final decision according to the decision fusion [30], as follows;

$$F = \begin{cases} 1 & D + \sum_{i=1}^{K} L_i > 1 \\ 0 & Otherwise \end{cases} \tag{11}$$

Based on the improved bi-level estimator logic discussed above, the all $K$ segments of every IMF are properly divided into noise dominant bins and signal dominant bins. Further noise dominant bins are processed for thresholding according to the process mentioned in [23]. Finally the noise suppressed features are combined and reconstructed into a noise free signal.
**Note:** A complete description through an algorithm is represented in Appendix at the end of the paper.

## 4. Simulation Results

To illustrate the effectiveness of the proposed speech enhancement algorithm, extensive computer simulations were carried out with 10 male and 10 female utterances, randomly selected from the TIMIT database. The performance evaluation of proposed framework is carried out over different speech signals contaminated with different types of noise like Babble Noise, Exhibition Noise, Restaurant Noise and Train Station noise. Different speech signals are also considered having different Signal to Noise Ratio (SNR) values like 0 dB, 5 dB, 10 dB and 15 dB.

To check the performance evaluation of proposed mechanism, different performance metrics are measured at different instants. The performance metrics considered for performance evaluation are overall output SNR and average segmental SNR. Further to check the quality of the enhanced speech, a subjective evaluation is carried out through Perceptual evaluation of speech quality (PESQ) test. Further one more metric, composite speech quality measure (COMP) [31] was used for objective performance comparison. The COMP quality index is a linear combination of different objective quality measures such as Log-Likelihood Ratio (LLR) and Weighted Spectral Slop (WSS). The correlation coefficients for the linear combination are determined from listening tests. The index is given by,

$$COMP = 1.594 + 0.805PESQ - 0.512LLR - 0.007WSS \quad (12)$$

This is known to be highly correlated with human listening.

To show the performance improvement of proposed speech enhancement algorithm, it was compared with conventional approaches proposed by Tafiq.et.al, [22] S.N. Rao et.al, [23] and the Saggurti.et.al, [26]. The obtained simulation results for different noise types are represented form figure.3 to figure.6. Further the performance evaluation is measured through the performance metrics for all the above test cases and the obtained results are shown in the following figure.7.
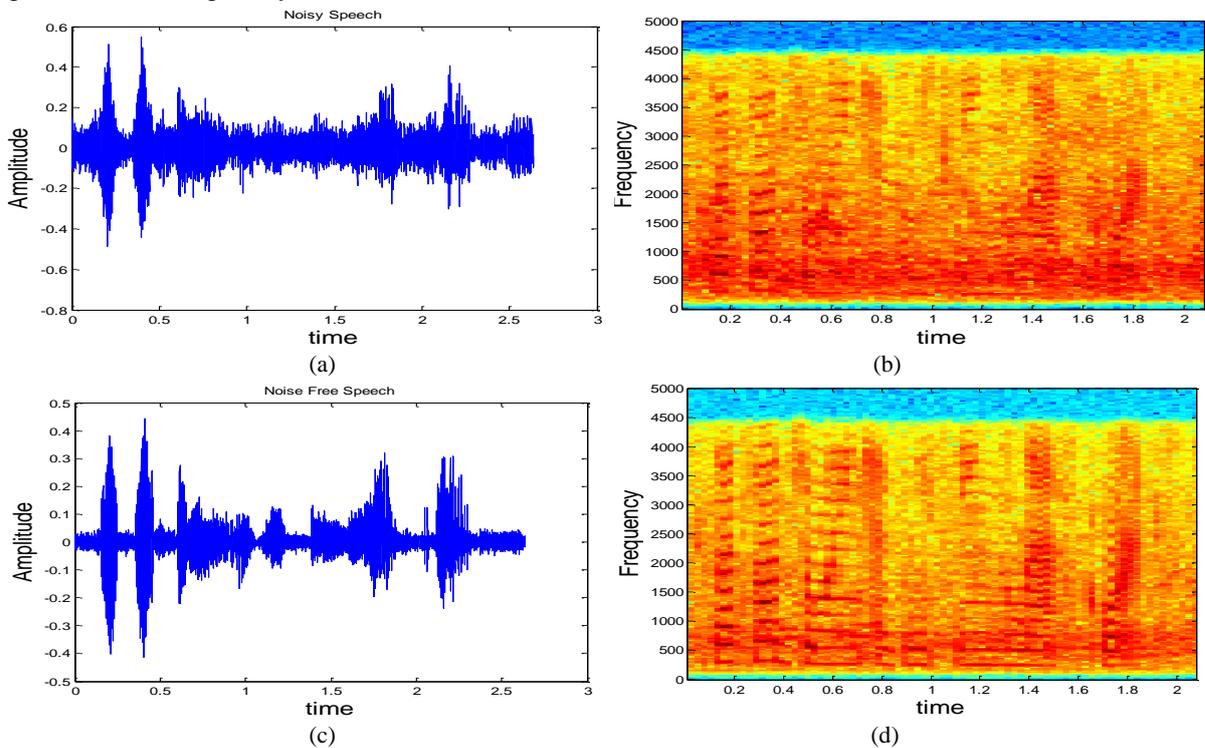


**Fig. 3:** Obtained results for clean speech with Babble Noise,
(a) Noisy Speech, (b) Spectrogram of (a), (c) Noise Free Speech, (d) Spectrogram of (c)
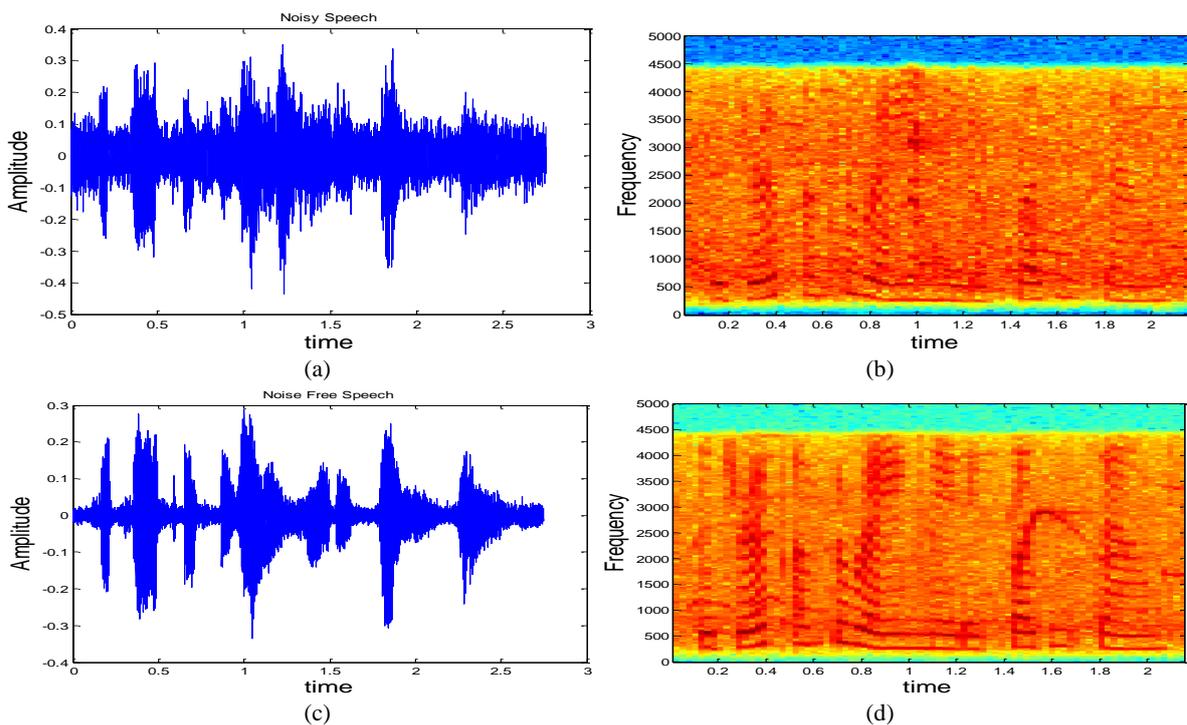


**Fig. 4:** Obtained results for clean speech with Exhibition Noise,
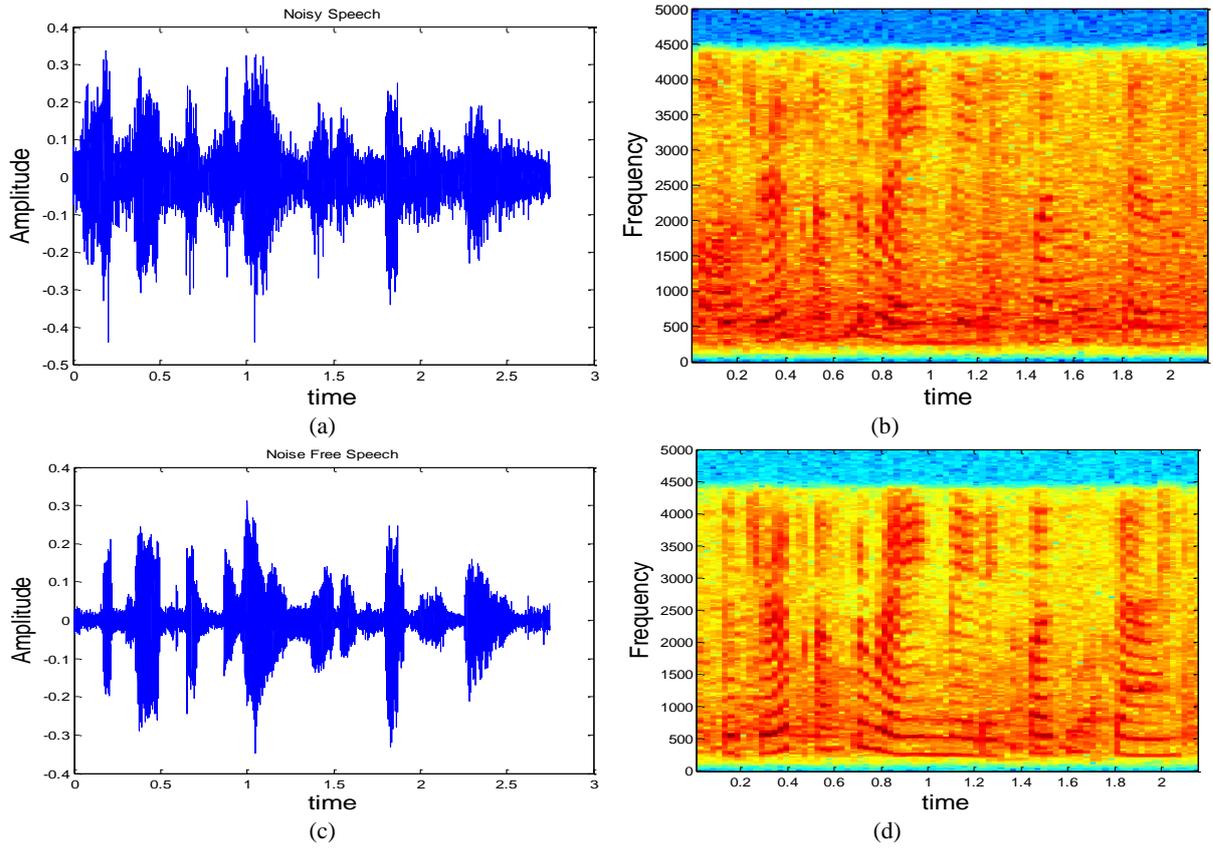(a) Noisy Speech, (b) Spectrogram of (a), (c) Noise Free Speech, (d) Spectrogram of (c)

**Fig. 5:** Obtained results for clean speech with Restaurant Noise,
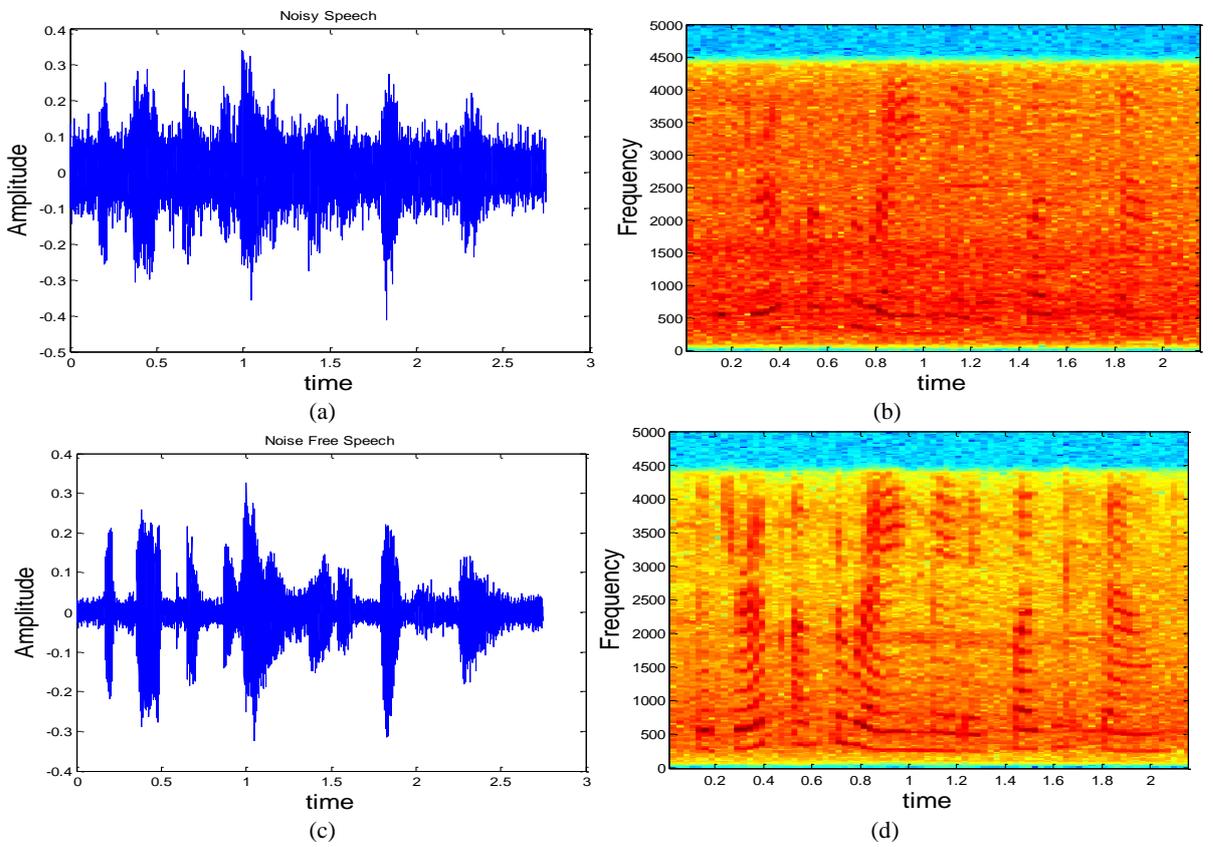(a) Noisy Speech, (b) Spectrogram of (a), (c) Noise Free Speech, (d) Spectrogram of (c)



**Fig. 6:** Obtained results for clean speech with Train Station Noise,
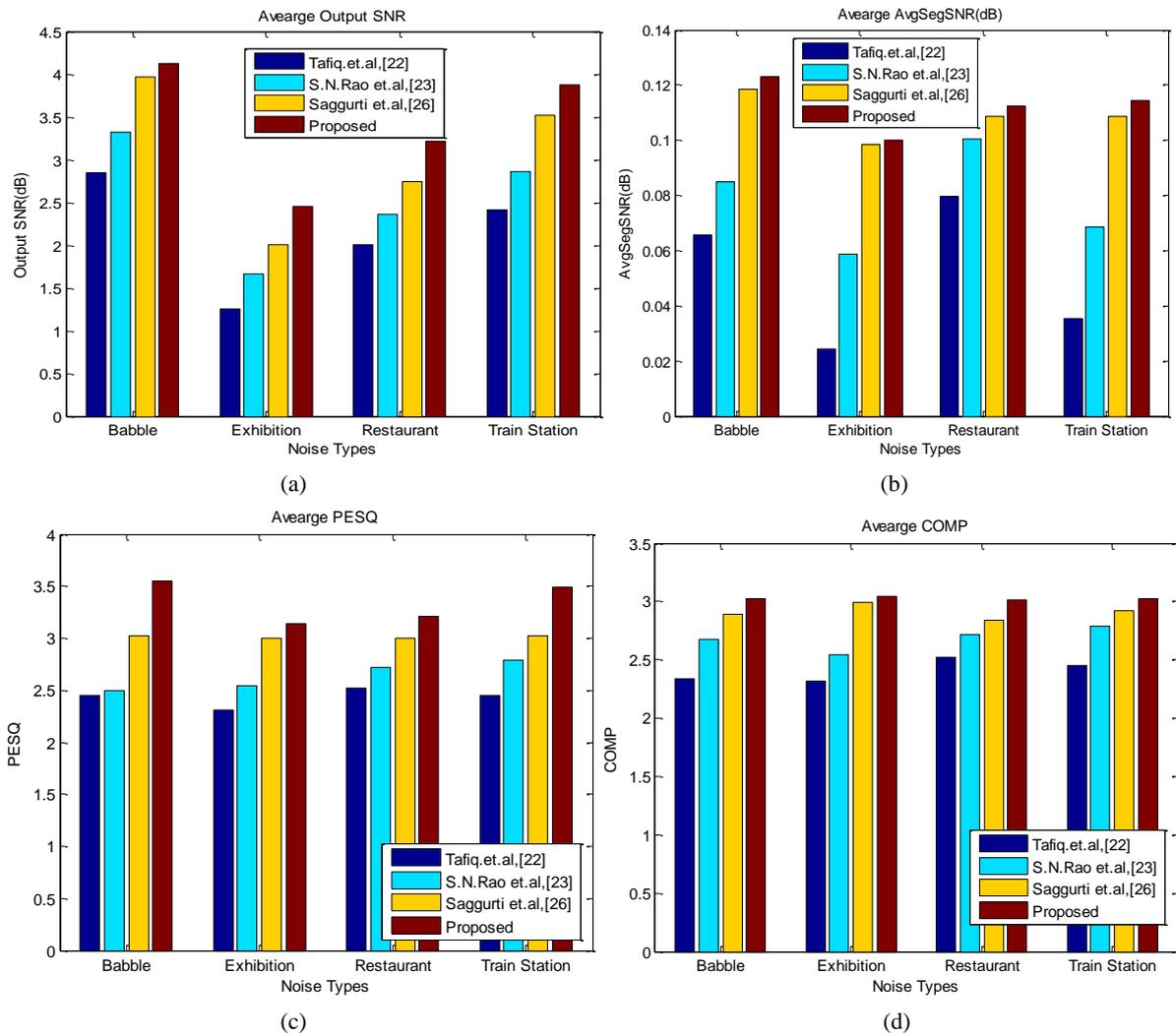(a) Noisy Speech, (b) Spectrogram of (a), (c) Noise Free Speech, (d) Spectrogram of (c)

**Fig. 7:** Performance metrics evaluation and comparison, Average value of
(a) Output SNR, (b) AvgSegSNR, (c) PESQ and (d) COMP for different noise types

Figure.3 to figure.6 shows the results obtained after the simulation of proposed approach. Totally the simulation has considered 10 female and 10 male speech samples and only few of them are represented above. Figure.3 represents the results related to the babble noise, figure.4 describes the exhibition noise related results, figure.5 is of restaurant noise related results and the figure.6 belongs to the Train station noise related results. In all the figures, the spectrograms of noisy speech along with the spectrogram of noise free speech are represented. As it can be observed that the noise free speech is more efficient compared to the noisy free speech.

Further the figure.7 shows the comparative analysis between the proposed and conventional approaches through the obtained performance metrics such as Output SNR, AvgSegSNR, PESQ and COMP. All these metrics are measured first for the proposed approach over its accomplishment on the clean speech signals contaminated with different types of noises such as Babble, Exhibition, Restaurant and Train Station noises at different SNR values. Further they are subjected to averaging and compared with conventional approaches. The comparative analysis shown in the figure.7 describes the efficiency of proposed approach. On an average the proposed approach achieved an improvement of 1.2856dB, 0.8681 dB and 0.35593 dB in the overall output SNR form the conventional approaches Tafiq.et.al.,[22] S.N.Rao et.al., [23] and the Saggurti.et.al., [26], respectively. Similarly on an average the proposed approach achieved an improvement of 0.0612dB, 0.0345 dB and 0.0041 dB in the Average AvgSegSNR form the conventional approaches Tafiq.et.al.,[22] S.N.Rao et.al., [23] and the Saggurti.et.al., [26], respectively. Further the performance improvement with respect to the PESQ is observed to

be 0.9158, 0.7109, and 0.3389 form the conventional approaches Tafiq.et.al.,[22] S.N.Rao et.al., [23] and the Saggurti.et.al., [26], respectively. Finally the COMP metrics is observed as 0.6212, 0.3450 and 0.1151 respectively.

## 5. Conclusion

A new speech enhancement approach is developed in this paper to enhance the quality of speech recorded under adverse acoustic environments. This approach is completely oriented to the uncertainty evaluation of speech presence and absences. Unlike the conventional approaches those applies noise estimation directly over the noisy contaminated speech signals, this approach first discovers the noise dominant bins and speech dominant bins more accurately and then applies the noise suppression method to remove the noise. The proposed improved bi-level estimator logic is presented in this paper to perform an accurate segmentation through the energy based threshold. This approach mainly aims to increases the probability of detection such that the noise dominant bins are only subjected to noise filtering, results in the quality enhancement. Experimental evaluations are done over different types of noises and the performance is measured with some performance metrics including Output SNR, PESQ and observed an effective improvement over the conventional approaches.

# References

[1] S. F. Boll, "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.

[2] Y. Hu, Subspace and multitaper methods for speech enhancement [Ph.D. dissertation], University of Texas at Dallas, Richardson, Tex, USA, 2003.

[3] Firas Jabloun and Benoît Champagne, "Incorporating the Human Hearing Properties in the Signal Subspace Approach for Speech Enhancement", IEEE Transactions on Speech and Audio Processing, Vol. 11, No. 6, November 2003.

[4] D. O'Shaughnessy, Speech Communications Human and Machine, 2nd ed. New York: IEEE Press, 2000

[5] Philipos C. Loizou, "Speech Enhancement: theory and practice", 2nd Ed. CRC press, Taylor and Francis Group, 2013.

[6] R. Martin, "Noise power spectral density estimation based on optimal smoothing and minimum statistics," IEEE Trans. Speech Audio Process., vol. 9, no. 5, pp. 504–512, July 2001.

[7] I. Cohen, "Noise spectrum estimation in adverse environments: Improved minima controlled recursive averaging," IEEE Trans. Speech Audio Process., vol. 11, no. 5, pp. 466–475, Sept. 2003.

[8] R. Yu, "A low-complexity noise estimation algorithm based on smoothing of noise power estimation and estimation bias correction," in IEEE ICASSP, 2009, pp. 4421–4424.

[9] R. C. Hendricks, R. Heusdens, and J. Jensen, "MMSE based noise PSD tracking with low complexity," IEEE ICASSP, pp.4266–4269, Mar. 2010.

[10] J. Taghia, N. Mohammadiha, J. Sang, V. Bouse, and R. Martin, "An evaluation of noise power spectral density estimation algorithms in adverse acoustic environments," IEEE ICASSP, May 2011.

[11] Mahdi Parchami, Wei-Ping Zhu, Benoit CHampange, "Recent Developments in Speech Enhancement in the Short-Time Fourier Transform Domain", IEEE Circuits and Systems Magazine, Volume: 16, Issue: 3, third quarter 2016.

[12] J. S. Erkelens, R. C. Hendriks, R. Heusdens, and J. Jensen, "Minimum mean-square error estimation of discrete Fourier coefficients with generalized gamma priors," IEEE Trans. Audio, Speech, Lang. Process., vol. 15, no.6, pp. 1741–1752, 2007.

[13] J. Jensen and R. C. Hendriks, "Spectral magnitude minimum mean-square error estimation using binary and continuous gain functions," IEEE Trans. Audio, Speech, Lang. Process., vol. 20, no. 1, pp. 92–102, 2012.

[14] I. Cohen, "Optimal speech enhancement under signal presence uncertainty using log-spectral amplitude estimator," IEEE Signal Process. Lett., vol. 9, no. 4, pp. 113–116, 2002.

[15] Van-Khanh MAI, Dominique PASTOR, Abdeldjalil AÏSSA-EL-BEY, Raphaël LE BIDAN, "Combined Detection and Estimation Based on Mean-Square Error Log-Spectral Amplitude for Speech Enhancement", GRETSI 2017.

[16] Y. Hu and P. C. Loizou, "Techniques for estimating the ideal binary mask," in Proc. 11th Int. Workshop Acoust. Echo Noise Control, 2008, pp. 154–157.

[17] U. Kjems et al., "Role of mask pattern in intelligibility of ideal binary-masked noisy speech," J. Acoust. Soc. Amer., vol. 126, no. 3, pp.1415–1426, Sep. 2009.

[18] Y. Hu and P. C. Loizou, "Techniques for estimating the ideal binary mask," in Proc. 11th Int.Workshop Acoust., Echo, Noise Control, 2008.

[19] G. Kim, Y. Lu, and P. C. Loizou, "An algorithm that improves speech intelligibility in noise for normal-hearing listener," J. Acoust. Soc. Amer., vol. 126, no. 3, pp. 1486–1494, Sep. 2009.

[20] Leo Lightburn, Enzo De sena, Alastair Moore, "Improving the perceptual quality of ideal binary masked speech", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2017.

[21] Abhay Upadhyay, "Speech enhancement based on mEMD-VMD method", IEEE Electronics Letters, Vol.53, No.7, 2017.

[22] Taufiq Hasan, "Suppression of Residual Noise From Speech Signals Using Empirical Mode Decomposition", IEEE signal processing letters, Vol. 16, No. 1, 2009.

[23] S. Nageswara Rao, K.J. Shankar and C.D. Naidu, "An adaptive speech enhancement approach based on DCT and empirical mode decomposition", In Proc. of International Conference on Communication and Signal Processing (ICCSP), 2016.

[24] Yannis Kopsinis, Member, IEEE, and Stephen McLaughlin, "Development of EMD-Based Denoising Methods

[25] Inspired by Wavelet Thresholding", IEEE Transactions on Signal Processing, Vol. 57, No. 4, April 2009.

[26] Taufiq Hasan and Md. Kamrul Hasan, "A Probabilistic Speech Enhancement Filter Utilizing the Constructive and Destructive Interference of Noise", 15th European Signal Processing Conference (EUSIPCO 2007), Poznan, Poland, September 3-7, 2007.

[27] Saggurti Nageswararao, K Jaya sankar, C.D Naidu, "Suppression of Non-Stationary Noises Through the Generalized Signal Detector", International Journal of Intelligent Engineering and Systems, Vol.11, No.1, 2018.

[28] S. S. Kalamkar and A. Banerjee, "On the Performance of Generalized Energy Detector under Noise Uncertainty in Cognitive Radio", In: Proc. of National Conference on Communications (NCC), 2013.

[29] Jun Zhu, Yun Bai, "Analytical Optimization for Collaborative Double Threshold Energy Detection in Cognitive Radio Network", "Journal of Information & Computational Science 9: 13 (2012) 3875–3882.

[30] H.Urkowitz, "Energy detection of unknown deterministic signals", Proceedings of IEEE, vol.55, pp. 523-531, April 1967.

[31] Sun.C, Zhang.W, Letaief.K.B, "Cluster-based cooperative spectrum sensing in cognitive radio systems", in Proc. IEEE ICC'07, pp.2511-2515, 2007.

[32] Y. Hu and P. C. Loizou, "Evaluation of objective measures for speech enhancement," in Proc. INTERSPEECH, Sep. 2006.

[33] Avinash Yadlapati, Dr. Hari Kishore Kakarla, "An Advanced AXI Protocol Verification using Verilog HDL", Wulfenia Journal, ISSN: 1561-882X, Volume 22, Number 4, pp. 307-314, April 2015.

[34] P Ramakrishna, K. Hari Kishore, "Design of Low Power 10GS/s 6-Bit DAC using CMOS Technology "International Journal of Engineering and Technology(UAE), ISSN No: 2227-524X, Vol No: 7, Issue No: 1.5, Page No: 226-229, January 2018.

[35] A Murali, K. Hari Kishore, "Efficient and High Speed Key Independent AES Based Authenticated Encryption Architecture using FPGAs "International Journal of Engineering and Technology(UAE), ISSN No: 2227-524X, Vol No: 7, Issue No: 1.5, Page No: 230-233, January 2018.

[36] G.S.Spandana, K Hari Kishore "A Contemporary Approach For Fault Diagnosis In Testable Reversible Circuits By Employing The CNT Gate Library" International Journal of Pure and Applied Mathematics, ISSN No: 1314-3395, Vol No: 115, Issue No: 7, Page No: 537-542, September 2017.

[37] K Hari Kishore, CVRN Aswin Kumar, T Vijay Srinivas, GV Govardhan, Ch Naga Pavan Kumar, R Venkatesh "Design and Analysis of High Efficient UART on Spartran-6 and Virtex-7 Devices", International Journal of Applied Engineering Research, ISSN 0973-4562, Volume 10, Number 09 , pp. 23043-23052, June 2015.

[38] K Bindu Bhargavi, K Hari Kishore "Low Power BIST on Memory Interface Logic", International Journal of Applied Engineering Research, ISSN 0973-4562, Volume 10, Number 08 , pp. 21079-21090, May 2015.

[39] Korraprolu Brahma Reddy, K Hari Kishore, "A Mixed Approach for Power Dissipation Reduction in Nanometer CMOS VLSI circuits", International Journal of Applied Engineering Research, ISSN 0973-4562 Volume 9, Number 18 , pp. 5141-5148, July 2014.

[40] Nidamanuri Sai Charan, Kakarla Hari Kishore "Reorganization of Delay Faults in Cluster Based FPGA Using BIST" Indian Journal of Science and Technology, ISSN No: 0974-6846, Vol No.9, Issue No.28, page: 1-7, July 2016.

[41] Sravya Kante, Hari Kishore Kakarla, Avinash Yadlapati,"Design and Verification of AMBA AHB-Lite protocol using Verilog HDL" International Journal of Engineering and Technology, E-ISSN No: 0975-4024, Vol No.8, Issue No.2, Page:734-741, May 2016.

[42] Bandlamoodi Sravani, K Hari Kishore, "An FPGA Implementation of Phase Locked Loop (PLL)", International Journal of Applied Engineering Research, ISSN 0973-4562, Volume 10, Number 14 , pp. 34137-34139, August 2015

[43] Avinash Yadlapati, Kakarla Hari Kishore,"Constrained Level Validation of Serial Peripheral Interface Protocol", Proceedings of the First International Conference on SCI 2016, Volume 1, Smart Computing and Informatics, Smart Innovation, Systems and

Technologies 77, ISSN No: 2190-3018, ISBN: 978-981-10-5544-7, Chapter No: 77, pp. 743-753, 25th December 2017.

[44] P Kiran Kumar, P Prasad Rao, Kakarla Hari Kishore, "Optimal Design of Reversible Parity Preserving New Full Adder / Full Subtractor", IEEE SPONSORED 3rd INTERNATIONAL CONFERENCE ON ELECTRONICS AND COMMUNICATION SYSTEMS (ICECS 2016), pp. 3465-3470, 25th and 26th February 2016.

[45] Y Avinash, K Hari Kishore ''Designing Asynchronous FIFO for Low Power DFT Implementation'' International Journal of Pure and Applied Mathematics, ISSN No: 1314-3395, Vol No: 115, Issue No: 8, Page No: 561-566, September 2017.

[46] Mahesh Mudavath and K Hari Kishore "Design of RF Front End CMOS Cascade CS Low Noise Amplifier on 65nm Technology Process" International Journal of Pure and Applied Mathematics, ISSN No: 1314-3395, Vol No: 115, Issue No: 7, Page No: 417-422, September 2017.

[47] P. Sahithi K Hari Kishore, E Raghuveera, P. Gopi Krishna "DESIGN OF VOLTAGE LEVEL SHIFTER FOR POWER-EFFICIENT APPLICATIONS USING 45nm TECHNOLOGY" International Journal of Engineering and Technology(UAE), ISSN No: 2227-524X, Vol No: 7, Issue No: 2.8, Page No: 103-108, March 2018.

[48] N Bala Dastagiri K Hari Kishore "A 14-bit 10kS/s Power Efficient 65nm SAR ADC for Cardiac Implantable Medical Devices" International Journal of Engineering and Technology(UAE), ISSN No: 2227-524X, Vol No: 7, Issue No: 2.8, Page No: 34-39, March 2018.

[49] S.V.Manikanthan and T.Padmapriya "Recent Trends In M2m Communications In 4g Networks And Evolution Towards 5g", International Journal of Pure and Applied Mathematics, ISSN NO:1314-3395, Vol-115, Issue -8, Sep 2017.

[50] S.V. Manikanthan , T. Padmapriya "An enhanced distributed evolved node-b architecture in 5G tele-communications network" International Journal of Engineering & Technology (UAE), Vol 7 Issues No (2.8) (2018) 248-254.March2018.

[51] S.V. Manikanthan, T. Padmapriya, Relay Based Architecture For Energy Perceptive For Mobile Adhoc Networks, Advances and Applications in Mathematical Sciences, Volume 17, Issue 1, November 2017, Pages 165-179

## Appendix

| Algorithm 1: Speech Enhancement Algorithm |
|---|
| **Input:** A noisy speech signal X |
| **Output:** Noise free speech signal Y |
| %%%%%%%%%%%%%%%%%%% |
| Initially apply Empirical mode decomposition over the noisy speech signal X, as |
| IMFs{N} = EMD(X)                // Apply EMD to get *N* IMFs |
| For i=1:Length(IMFs)-1        // For loop to process every IMF, except last IMF (residual) |
| IMF = IMFs{i}; |
| Ln = length(IMF)                // number of sample in the every IMF |
| Segments = Ln/SL              // SL = Segmentation Length |
| Threshold = Th [26]            // derived according to the Ref.[26] |
| For j = 1:SL:Ln |
| Seg_m = IMF(1, j:SL) |
| Seg_s = Sum(Seg_m).^2/(Variance(Seg_m)) |
| t=1;  Th0 = Th/Beta;  Th1 = Th*Beta; |
| If Seg_s > Th1 |
| Seg_sd {t} = Seg_m            // speech dominant segments |
| t=t+1; |
| Elseif Seg_s < Th0 |
| Seg_nd {t} = Seg_m            // noise dominant segments |
| t=t+1; |
| Else Th0<Seg_s<Th1 |
| Find the fusion decision according to the preceding segments and segment into one of the categories. |
| End |
| End |
| End |
| ➤        Apply Noise Suppression according to the Ref. [23] over only noise dominant segments. |
| ➤        Concatenate the noise suppressed segments with speech dominant segments and reconstruct the noise free speech signal, Y. |