Adaptive Speech Enhancement Technique using Time Variable LMS Algorithm

JyoshnaGirika, Md. Zia Ur Rahman

Abstract*: Elimination of noise from the speech signal is the majortask in applications like communications, hearing aids, speech therapy etc. This facilitates to provide good resolution speech signal to the user. The speech signals are mainly affected due to the various natural as well as manmade noises. As the nature of these noises random in its nature fixed coefficient filtering techniques are not suitable for clutter elimination task. Hence, in this work an adaptive algorithm has driven noise canceller for speech enhancement applications which has an innate ability to change its weight coefficients depending on the statistical nature of the unwanted component in the original speech signal. In our experiments in order achieve better convergence as well as filtering capability we propose Time Variable Least Mean Square (TVLMS) algorithm rather than constant step parameter. The computational complexity of the speech enhancement process is also a key aspect due to the excessive length of the speech signals in practical scenario. Hence, to lower the computational complexity of the speech enhancement process we propose Sign Regressor TVLSM (SRTVLMS), which is a hybrid realization of familiar sign regressor algorithm and the proposed TVLMS. Using these two techniques noise cancellation models are developed and tested on real speech signals with unwanted noise contaminations. The experimental outputsconfirm that the SRTVLMS based signal enhancement unit performs better than its counterpart with respect to convergence rate, computational complexity and signal to noise ratio increment.

Index Terms*: Adaptive Noise Cancellation, Convergence Rate, Computational Complexity, Speech Signal, Signal Enhancement, Variable Step Size.

I. INTRODUCTION

Speech enhancement is a key phenomenon in the present-day technological evaluation scenario. This has many applications in the fields of mobile communications, battle filed communications, hearing aids, speech re-habitation, air traffic control, speech recognition, speech biometrics, teleconference systems, etc. Several researchers have contributed various techniques in the contest of speech enhancement. In [1] Ali Ameri Mahabadi et al. reviews a modern compositional Adaptive Noise Canceller (ANC), fits for movable artifact surroundings. Those outcomes prove that mixing of adaptive and conservative techniques with good characteristics with low calculation difficulty analyzed to various adaptive ways.However, in [2] Jae JinJeong et al. proposed a changeable step size for standardized sub-band adaptive filters is drawn by reducing the mean-square separation among the minimal weight vector and guessed weight vector at every point can efficiently find the movable surroundings.Extended in [3] Ali O. Abid Noor et al. proposed a minimized sub-band artifact remover for voice wave is improved which utilizes two-fold over slicing purifier bank and this patterned filter bank is reduced for less setup interference. Here full-band variant is taken as basic for examination. While at [4] Raymond H. Kwong et al. studies a change in step size as mean-square error changes, permitting the adaptive filter to record the variations in structure, also makes a little steady state error. Later, evaluates the results with LMS technique which forms an improved algorithm. Furthermorein [5] Zayed Ramadan et al. presented a method for adaptive noise canceller is replicated by utilizing various artifact strength intensities for immovable as well as movable artifact surroundings provides efficient velocities of meetings. Even though in [6] SungEun Jo et al. justified a reliable LMS-type method for figures least square make the step-size stabilization. While in [7] Mrityunjoy Chakraborty et al. introduced a numerous source compound LMS method utilized to guide a number of adaptive filters to oppose an arbitrary set of periodic artifact causes under meeting order, the strength of each fault series is free from voice frequencies, thus reduces the fault power. For standardization in [8] Shin'ichi Koike et al. studies declare the need of the differential equations in finding the cleaned characteristics with better results. To eliminate the noise from signal in [9] Yonggang Zhang et al. introduced a steady-state presentation for LMS based adaptation. This presentation provides nearness in the act of the FT technique with its widened realistic purposes. Data Normalization isin [10] Joonwan Kim et al. explained the speed in meeting of waves by the wavelet based LMS (WLMS) method to artifact voice difficulty and drawn effects are taken to assess the efficiency of the changeable step size normalized least mean squares (NLMS), wavelet based LMS techniques with set step size. According to unchangeable step size in [11] Leonardo Rey Vega et al. scrutinized a strong alterable step-size NLMS technique can minimize the square of a posteriori fault and also proved the relation among the suggested technique and the other obtained by utilizing a strong guide advance and found better results.In extension they introduced in [12] Scott C. Douglas et al. projected an NLMS technique is widespread, leading to a group of ledge-like technique supported on the L_p-reduced purifier coefficient variance. A total origin of algorithm group is set, and results are

Revised Manuscript Received on June 05, 2019

JyoshnaGirika, Department of Electronics and Communication Engineering, KoneruLakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India.

Md Zia Ur Rahman, Department of Electronics and Communication Engineering, KoneruLakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India.

carried out to explain the union characteristics of the methods.

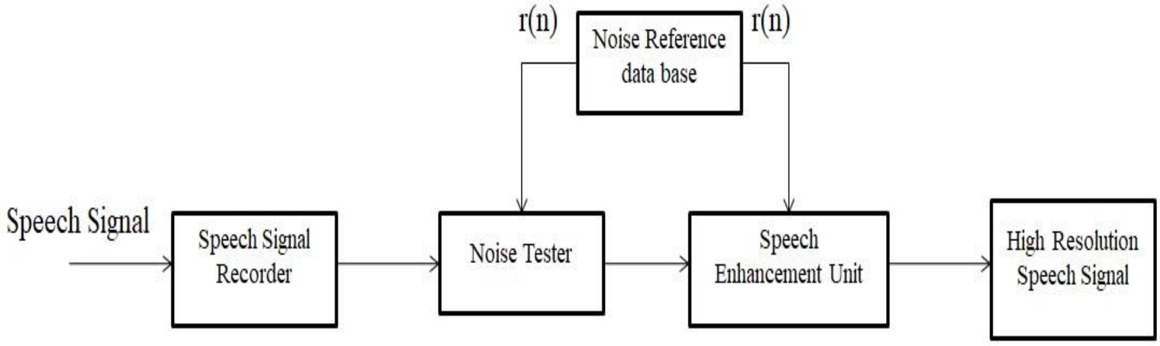

Fig.1: Proposed block diagram of Speech Signal Enhancement Unit used in our experiments.

In[13] Dr. D. Deepa et al. scrutinized clattery, wanted voice waves were double altered with discrete cosine and Hadamard changes and given to adaptive filter by utilizing NLMS method as a result, speed assembly of clattery voice to wanted voice with nice act when evaluated to conventional LMS method. A new innovative idea is futured in [14] Jwu-sheng Hu et al. described a modern advance with minimum variance distortion less response (MVDR) beam former which minimizes the least functioning and gives best strength beside a random but norm-bounded wanted wave routing vector variance. Also, second-order extended (SOE) $H\infty$ cleans the noise. Adifferentmethodhas been used to identify the pure speech signal from the artifact corrupted signal is explained briefly in [15] Yue Xian Zou et al. discussed an efficient spatial-frequency area voice development technique wiener post-filtering (WPF) methods in which WPF is the potent approximator reduces the spatial distortion. Not only LMS we have a new adaptive technique is also shown in [16] L. Zao et al. presented the empirical mode decomposition (EMD) to unwanted voice wave and gets a rest of intrinsic mode functions (IMF). Here EMD and Hurst-based (EMDH) analysis is estimated with voice improvement researches taking surroundings echo artifacts in various parameters of mobility which gets better results.To extract the speech, signal a modernized algorithm is used in [17] Brady N. M. Laska et al. proposed that by utilizing element clean outline lets the needed technique to form the voice spectral voltages as an autoregressive method with laplace spread excitation. Two variables, one is interacting various schemes for changes and the other permits for angle variations and hence developing slicing competence.In [18] Upal Mahbub et al. justified a two-level method to pact with acoustic echo cancellation (AEC), firstly the holdup edition of echo and cleaned wave is made as reference in the next level a gradient-based adaptive purifier method and minimum wiener-hopf result is formed and gave better results when compared with TIMIT database. They introduced the new implementation in [19] Amin Zehtabian et al. explained a modern advancewith voice wave development for Singular Value Decomposition (SVD) based along with Genetic Algorithm (GA) which proceed outcomes in a required decrement of the distorted wave among drawing the worth of the unique voice.In [20] M. Raghu Ram et al. described an easy and capable advance with the help of adaptive step-size least mean squares (AS-LMS) adaptive filter to lessen the motion artifacts (MAs) at distorted photoplethysmography (PPG) waves competently. Nowadays, due to the usage of high rate data in mobile communications and video calls, the speech enhancement algorithms should have less computational complexity. If the computational complexity of the speech enhancement algorithm is large, the input data samples overlap at the input port of the speech analysis block and causes inter symbol interference. This causes ambiguity in the speech signal at the receiver end. Hence, computational complexity of the speech enhancement algorithm is also a crucial aspect. So for to the best of authors knowledge this aspect is not addressed in literature in the contest of noise removal from speech signal. So, in our work we focused to minimize the computational complexity of the noise cancellation technique by combining the proposed TVLMS with Sign Regressor algorithm.Variousexperiments are performed to enhance speech signals using TVLMS, SRTVLMS and the performance is compared with reference to the familiar LMS based speech enhancement technique.

II. ADAPTIVE ALGORITHMS FOR SPEECH SIGNAL ENHANCEMENT

In order to eliminate the noise components, form the speech signals we propose a Speech Enhancement Unit (SEU) based on adaptive algorithm in our work. The main theme of this SEU is the proposed adaptive algorithm for the operation of speech enhancement minimizes the noise contamination in the actual speech signal. Figure 1 shows a schematic block diagram of SEU for speech signal enhancement. The input to the SEU is the recorded speech signal with noise contamination. The noise tester measures type of contamination using power spectral estimation and choose a correlated reference signal. If the correlated reference signal is not available then a random signal is given reference signal to the SEU. The adaptive algorithm used in the filtering process has the innate ability to change the filter coefficients to train the reference signal such that it becomes close to the noise contamination buried in the speech signal statistically. This process finally facilitates high resolution speech signal without noise contaminations.

Based on the steepest descent algorithm [21], ,$\nabla x(n)$.

$$\mathbf{v}(n+1) = \mathbf{v}(n) - Q\nabla x(n) \quad (2)$$

$$x(n) = E[m^2(n)]$$

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

As the negative gradient vector position's in the path of steepest descent for the N dimensional quadratic cost function, each recursion shifts the value of the filter coefficients nearer to their optimum value, which match up to the least attainable value of the cost function, x(n). The LMS algorithm is a random process [22] accomplishment of the steepest descent algorithm. Here the probability of the error signal is not acknowledged so the instant value is utilized as an approximation. The steepest descent algorithm then becomes

$$\mathbf{v}(n+1) = \mathbf{v}(n) - Q\nabla x(n) \quad (3)$$

Wherever $x(n) = m^2(n)$

The weight update recursion for the LMS algorithm is.

$$\mathbf{v}(n+1) = \mathbf{v}(n) + Q m(n)\mathbf{f}(n) \quad (4)$$

Where, $\mathbf{v}(n)$ is weight vector, Q is step size, m(n) is error component, $\mathbf{f}(n)$ is data vector.The mathematical expression for the Sign Regressor LMS (SRLMS)algorithm is.

$$\mathbf{v}(n+1) = \mathbf{v}(n) + Q m(n) sign(\mathbf{f}(n)) \quad (5)$$

In critical situations when the signal power is varying with respect to time randomly the step size parameter will not bound to a constant value. In such scenarios, time variable LMS is a better candidature. In this algorithm the step size parameter varies with respect to the instantaneous time value. The analysis of this algorithm is presented in [20]. TVLMS method is described by the following weight update mechanism,

$$v(n+1) = v(n) + Q(n) \times \mathbf{m}(n) \times \mathbf{f}(n) (6)$$

The step size parameter instantaneously varies according to the following function,

$$\mathbf{Q}(n) = \alpha(n) \times Q(0)$$

where $\alpha(n) = C\left(\frac{1}{(1+b^n \times d)}\right)$ is the decaying factor and C, b and d are positive constants that will determine the value of decaying factor. At each step the decaying factor is multiplied with initial step size. This method can achieve a faster convergence rate compared to LMS algorithm with constant step size and also can remove the artifacts effectively.The sign regressor version of this algorithm is written as,

$$v(n+1) = v(n) + Q(n) \times \mathbf{m}(n) \times sign(\mathbf{f}(n)) (7)$$

This SR is computationally less complex and the number of multiplications required for carrying the noise cancellation process.

III. EXPERIMENTAL RESULTS AND DISCUSSION

In order to test the ability of the speech enhancement units using TVLMS and SRTVLMS algorithms in practical we developed two SEUs and experimentation is carried. In all the SEUs the tap length is chosen as ten. In this experiment initially the concept of noise cancellation is proved by applying additive Gaussian noise and then several speech signals with real noise are applied. To prove the ability of the proposed adaptive algorithms speech signals are chosen for filtering. For that purpose, five sample speech signals are taken from the data base. Both synthetic and real noises are taken to prove the performance analysis of the proposed adaptive algorithms and the non-stationary tracking performance of the algorithms. The speech signal-I is anc.wav which is practically recorded signal with 53569 samples. The speech signal -II is male signal obtained from database and it has 95232 samples. The speech signal -III is a male voice recorded one with 100864 samples, the speech signal -IV has 103936 samples and the speech signal -V are female speech signals from data base records with 114176 samples respectively. These speech signals are contaminated with various types of noises namely cockpit noise, elevator noise, high voltage murmuring, gun firing noise and random noise. The experimental results for gun firing noise in battle filed scenario are shown for all the five considered samples in the figures 2-6. The performance of these techniques is calculated in terms of signal to noise ratio improvement (SNRI) and are tabulated in Table I. The performance of these techniques is calculated in terms of signal to noise ratio improvement (SNRI) and are tabulated in Table I. The comparison of calculated SNRI is shown in Figure 7.

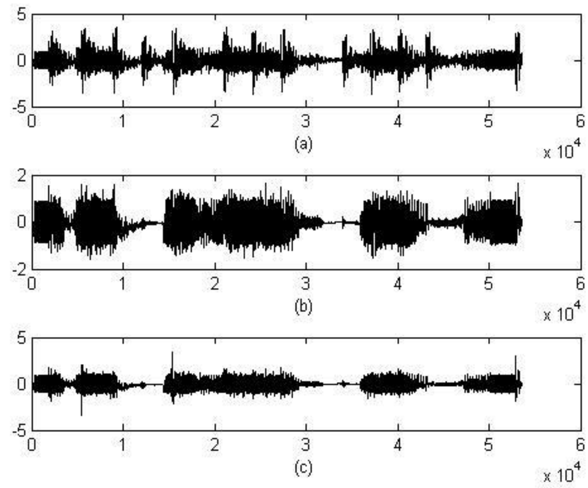

Fig.2: Enhancement results for Speech Signal I: (a). Speech Signal with gun firing noise in battle field, (b). enhanced signal using TVLMS algorithm-based Speech Enhancement Unit, (c). enhanced signal using SRTVLMS algorithm-based Speech Enhancement Unit.

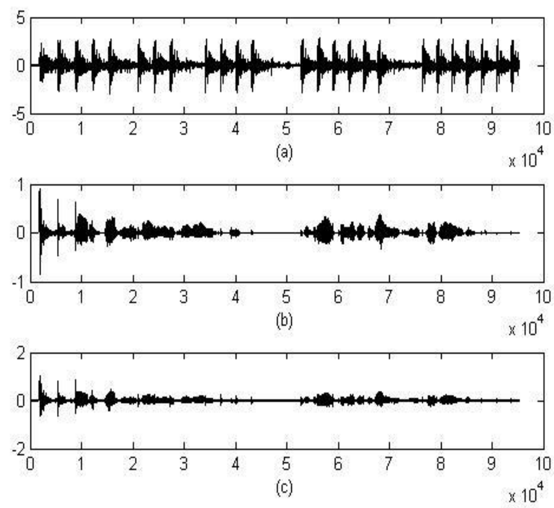

Fig.3: Enhancement results for Speech Signal II: (a). Speech Signal with gun firing noise in battle field, (b). enhanced signal using TVLMS algorithm-based Speech Enhancement Unit, (c). enhanced signal using SRTVLMS algorithm-based Speech Enhancement Unit.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Table I: Signal to Noise RatioImprovementcomputations for the proposed Speech Enhancement Units based on LMS, TVLMS and SRTVLMS algorithms (all values are in dBs)

S.No	Noise type	Sample	LMS	TVLMS	SRTVLMS
1.	Cockpit Noise	Speech Signal I	8.5795	22.3672	20.4578
		Speech Signal II	8.1673	22.4823	20.3245
		Speech Signal III	8.8705	22.3494	20.8743
		Speech Signal IV	8.3691	22.3914	20.3765
		Speech Signal V	8.7753	22.4392	20.4376
2.	Elevator Noise	Speech Signal I	6.1468	20.5924	19.3287
		Speech Signal II	6.3617	20.4720	19.2176
		Speech Signal III	6.5784	20.4934	19.3287
		Speech Signal IV	6.8665	20.8043	19.8743
		Speech Signal V	6.0346	20.4938	19.6532
3.	High Voltage Murmuring	Speech Signal I	5.6385	19.4870	18.9843
		Speech Signal II	5.1893	19.1369	18.6521
		Speech Signal III	5.3866	19.4593	18.7632
		Speech Signal IV	5.9582	19.4626	18.2198
		Speech Signal V	5.7418	19.4354	18.3265
4.	Gun Firing Noise	Speech Signal I	7.9127	21.6463	19.2149
		Speech Signal II	7.0836	21.3290	19.2387
		Speech Signal III	7.3353	21.2579	19.3469
		Speech Signal IV	7.7538	21.3254	19.2167
		Speech Signal V	7.5253	21.8756	19.2398
5.	Random Noise	Speech Signal I	9.0314	23.8845	21.2180
		Speech Signal II	9.9735	23.2379	21.2387
		Speech Signal III	9.2946	23.9843	21.3274
		Speech Signal IV	9.5904	23.3409	21.0943
		Speech Signal V	9.7733	23.4598	21.5689

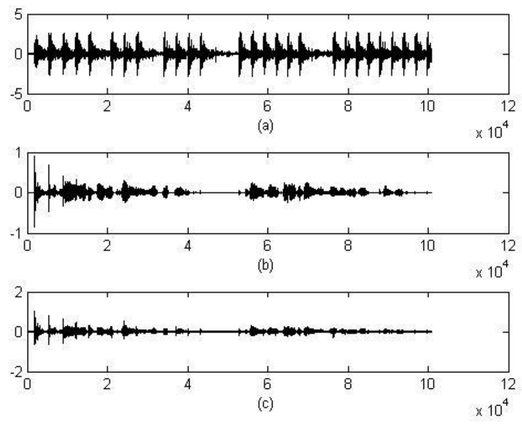

Fig.4: Enhancement results for Speech Signal III: (a). Speech Signal with gun firing noise in battle field, (b). enhanced signal using TVLMS algorithm-based Speech Enhancement Unit, (c). enhanced signal using SRTVLMS algorithm-based Speech Enhancement Unit.

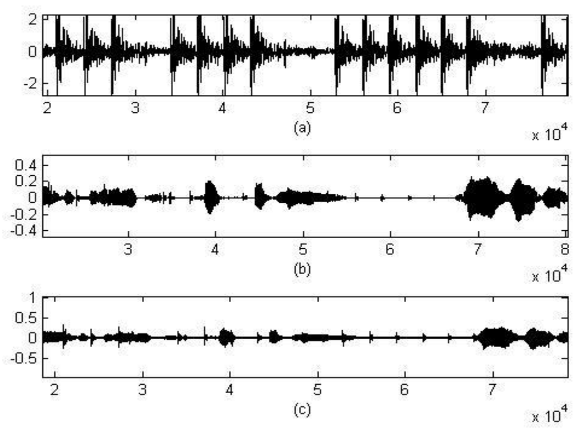

Fig.5: Enhancement results for Speech Signal IV: (a). Speech Signal with gun firing noise in battle field, (b). enhanced signal using TVLMS algorithm-based Speech Enhancement Unit, (c). enhanced signal using SRTVLMS algorithm-based Speech Enhancement Unit.

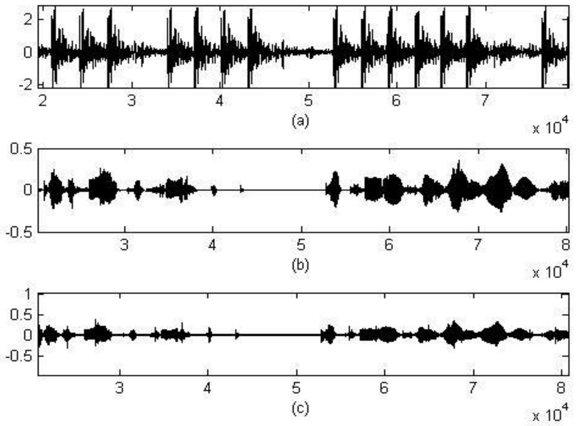

Fig.6: Enhancement results for Speech Signal V: (a). Speech Signal with gun firing noise in battle field, (b). enhanced signal using TVLMS algorithm-based Speech Enhancement Unit, (c). enhanced signal using SRTVLMS algorithm-based Speech Enhancement Unit.

IV. CONCLUSION

This research comprises of adaptive noise cancellation of voice waves by overcoming the composition of different categories of noises. With selected step-size the conventional LMS technique changes the gradient noise component. In order to cope up the drawbacks associated with constant step size parameter, in this work a variable step size parameter with respect to time variable is developed in the contest of speech signal enhancement. To overcome the problem of inter symbol interference at the input port of the noise canceller, a strategy is applied to minimize the computational burden of the noise canceller. This is possible by developing the hybrid realization of time variable step size LMS and sign regressor algorithm. Therefore, in our realizations, we developed TVLMS and SRTVLMS algorithms for the development of proposed speech signal enhancement. The experimental results carried during the noise elimination process are shown in Figures 2-6, the calculated signal to noise ratio improvement in these experiments are showed in Table I. From these results it is clear that SRTVLMS based SEU achieved little bit inferior SNRI than TVLMS. But, due to the signum function applied to the data vector we are able to minimize number of multiplications equal to the tap length of SEU. Therefore, based on this it is conclude that SRTVLMS based SEU performs better than TVLMS as well as conventional LMS. Hence, SRTVLMS based SEU can be used in all practical applications in real time environment.

REFERENCES

1. Ali Ameri Mahabadi, Seyed Amir Hejaziand Mohammad A. Akhaee, "A Combinational Adaptive Noise Canceller Using Filter Bank", Proceedings of the 6th International Symposium on Image and Signal Processing and Analysis, pp. 461-464, 2009.
2. Jae JinJeong, Keunhwi Koo andGyu Tae Choi, "A Variable Step Size for Normalized Subband Adaptive Filters", IEEE signal processing letters, Vol. 19, no. 12, pp. 906-909, Dec. 2012.
3. Ali O. Abid Noor, Salina Abdul Samad and Aini Hussain, "Noise Cancellation in Speech Using Optimized Subband Adaptive Filtering", Proceedings IEEE Student Conference on Research and Development (SCOReD), pp. 266-268, Nov. 16-18, 2009.
4. Raymond H. Kwongand Edward W. Johnston, "A Variable Step Size LMS Algorithm", IEEE Trans. on Signal Processing. Vol. 40, no. 7, pp. 1633-1642, July 1992.
5. Zayed Ramadan and Alexander Poularikas, "An Adaptive Noise Canceller Using Error Nonlinearities in the LMS Adaptation", IEEE South east Con, pp. 359 – 364, 2004.
6. SungEunJoand Sang Woo Kim, "Consistent Normalized Least Mean Square Filtering With Noisy Data Matrix", IEEE Trans. on signal processing, Vol. 53, no. 6, pp. 2112-2123, June 2005.
7. MrityunjoyChakrabortyand Hideaki Sakai, "Convergence Analysis of a Complex LMS AlgorithmWith Tonal Reference Signals", IEEE Trans. on speech and audio processing, Vol. 13,
8. no. 2, pp. 286-292, Mar. 2005.
9. Shin'ichi Koike, "Analysis of Adaptive Filters Using Normalized Signed Regressor LMS Algorithm", IEEE Trans. on signal processing, Vol. 47, no. 10, pp. 2710-2723, Oct.1999.
10. Yonggang Zhang, Ning Li and Jonathon A. Chambers, "Steady-State Performance Analysis of a Variable Tap-Length LMS Algorithm", IEEE Transactions on signal processing, Vol. 56, no. 2, pp. 839-845, Feb. 2008.
11. Joonwan Kim, Chae-Wook Lee and Jae-Kyun Lee, "Comparison of adaptive systems for noise reduction in speech performance", Proceedings of the 38th Southeastern Symposium on System Theory, pp. 123-127, Mar. 5-7, 2006.
12. Leonardo Rey Vega, Hernán Rey and Jacob Benesty, "A New Robust Variable Step-Size NLMS Algorithm", IEEE Trans. on signal processing, Vol. 56, no. 5, pp. 1878-1893, May 2008.
13. Scott C. Douglas, "AFamily of Normalized LMS Algorithms", IEEE Signal processing letters, Vol. 1, no. 3, pp. 49-51, Mar. 1994.
14. Dr. D. Deepa and Dr. C. Poongodi, "Speech Enhancement Using Dual Transform-Normalized LMS Algorithm for Speech Recognition Application", International Journal of Advance Research, Ideas and Innovations in Technology, Vol. 3, no. 4, pp. 729-735, 2017.
15. Jwu-Sheng Hu, Ming-Tang Lee and Chia-Hsin Yang, "Robust Adaptive Beamformer for Speech Enhancement Using the Second-Order Extended H_∞ Filter", IEEE Trans. on Audio, Speech, and Language Processing., Vol. 21, no. 1, pp. 39-50, Jan. 2013.
16. Yue Xian Zou, Peng Wang and Yong Qing Wang, "Speech enhancement with an acoustic vector sensor: an effective adaptive beamforming and post-filtering approach", EURASIP Journal on Audio, Speech, and Music Processing, 2014.
17. L. Zão, R.Coelho and P. Flandrin, "Speech Enhancement with EMD and Hurst-Based Mode Selection", IEEE/ACM Trans. on Audio, Speech, and Language Processing, Vol. 22, no. 5, pp. 899-911, May 2014.
18. Brady N. M. Laska, MiodragBolic´and Rafik A. Goubran, "Particle Filter Enhancement of Speech Spectral Amplitudes", IEEE Trans. on Audio, Speech, and Language Processing, Vol. 18, no. 8, pp. 2155-2167, Nov. 2010.
19. Upal Mahbub, Shaikh Anowarul Fattah and Wei-Ping Zhu, "Single-channel acoustic echo cancellation in noise based on gradient-based adaptive filtering", EURASIP Journal on Audio, Speech, and Music Processing, 2014.
20. Amin Zehtabian, Hamid HassanpourandShahrokhZehtabian, "A Novel Speech Enhancement Approach Based on Singular Value Decomposition and Genetic Algorithm",in: proc. SoCPaR-2010,2nd Intl. Conf. on Soft Computing and Pattern Recognition, CergyPontoise, France, pp.430-435, Dec. 7-10, 2010.
21. M. Raghu Ram, K. Venu Madhav and E. Hari Krishna, "A Novel Approach for Motion Artifact Reduction in PPG Signals Based on AS-LMS Adaptive Filter", IEEE Trans. on Instrumentation and Measurement, Vol. 61, no. 5, pp. 1445-1457, May 2012.
22. B. Widrow and S. D. Stearns, *Adaptiue Signal Processing.* Englewood Cliffs, NJ: Prentice-Hall, 1985.
23. Peng Xu, Jiangzhou Wang, Jinkuan Wang, and Feng Qi," Analysis and Design of Channel Estimation in Multicell Multiuser MIMO OFDM Systems", IEEE Transactions on Vehicular Technology, vol. 64, no. 2, Feb. 2015, PP.610-620.
24. EwedaEweda, "Analysis and Design of a Signed Regressor LMS Algorithm for Stationary and Non-stationary Adaptive Filtering with Correlated Gaussian Data", *IEEE Transactions on Circuits and Systems*, vol. 37, no. 11, pp.1367-1374, 1990.
25. Shinichi Koike, "Analysis of Adaptive Filters Using Normalized Signed Regressor LMS Algorithm", *IEEE Transactions on Signal Processing*, vol. 47, no. 10, pp.2710-2723, 1999.

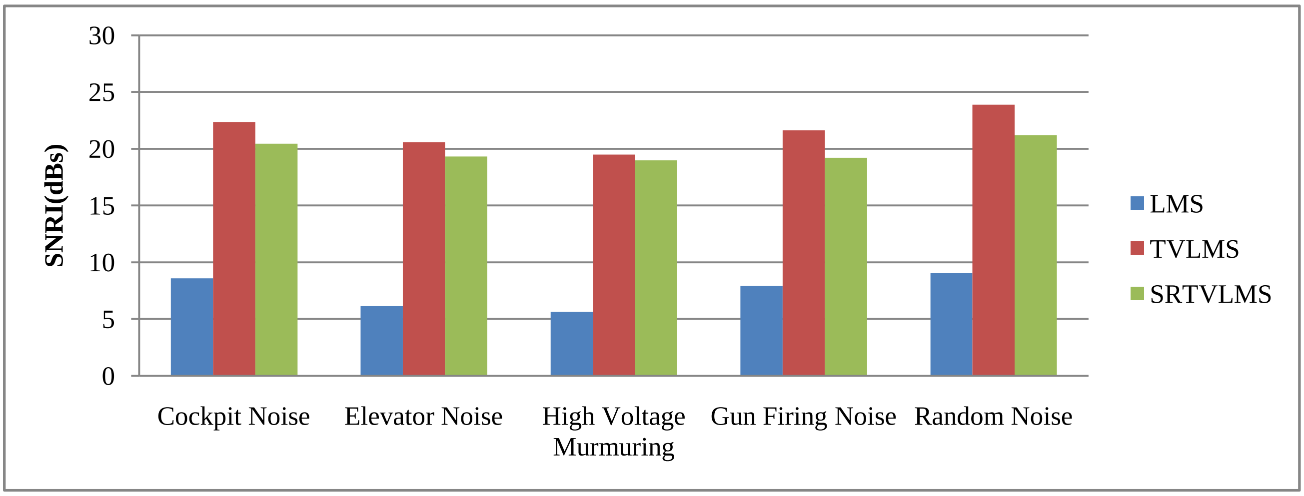

Fig.7: Comparison of performance measures in speech signal filtering due to various adaptive filters.

AUTHORS PROFILE

JYOSHNA GIRIKAis currently a Research Scholar in the Department of Electronics and Communications Engineering, Dept. of E.C.E., KoneruLakshmaiah Education Foundation, Green Fields, Guntur-522502, A.P., Inida. Her areas of interests are adaptive signal processing, speech processing and algorithms for speech analysis.

MD ZIA UR RAHMAN(M'09) (SM'16) received M.Tech. and Ph.D. degrees from Andhra University, Visakhapatnam, India. Currently, he is a Professor with the Department of Electronics and Communication Engineering,KoneruLakshmaiah Educational Foundation Guntur, India. His current research interests include adaptive signal processing, biomedical signal processing, array signal processing, MEMS,Nano photonics. He published more than 100 research papers in various journals and proceedings. He is serving in various editorial boards in the capacity of Editor in Chief, Associate Editor, reviewer for publishers like IEEE, Elsevier, Springer, IGI, American Scientific Publishers, Hindawai etc.