ISSN: 2277-9655
Impact Factor: 4.116

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

SPEECH ENHANCEMENT BASED ON SELF ADAPTIVE LAGRANGE MULTIPLIER WITH WEIGHTED PERCEPTUAL WIENER DE-NOISING TECHNIQUE

S. Arjuna Rao*, K. Murali Krishna
* Dept. of ECE Gudlavalleru Engg. College (GEC) Gudlavalleru, krishna-521356, AP, India
Dept. of ECE Gudlavalleru Engg. College (GEC) Gudlavalleru, krishna-521356, AP, India

DOI: 10.5281/zenodo.59536

ABSTRACT

The most voice based communication systems facing many problems such as lack of perceptual clarity, musical noise or residual noise, speech distortion and noise distortion. The main objective of speech enhancement is to improve the speech quality and intelligibility. By using wiener filter with Lagrange multiplier makes tradeoff between the speech distortion and residual noise, when the value of Lagrange multiplier is greater than or equal to zero otherwise causes speech distortion and residual noise. The perceptual wiener filter also contains some residual noise and there is a nonlinear relationship between the Lagrange multiplier and threshold value causes noise distortion. In this a Psycho acoustically motivated method is used for choosing better Lagrange multiplier value and to avoid nonlinear relationship. The objective evaluation showed that the proposed method performance is better than different existing methods.

KEYWORDS: intelligibility, wiener filter, Lagrange multipliers, threshold value, psychoacoustically.

INTRODUCTION

Speech is the most important parameter for human communication. Most of the speech based application systems faces the problem of degradation of speech quality and intelligibility [13] due to additive noise. Speech enhancement is a challenge to the many researchers to avoid additive noise (speech distortion and noise distortion).

The spectral subtraction [2] method subtracts the estimated power spectrum or magnitude spectrum of the noise from the power spectrum or magnitude spectrum of noisy speech signal. The main problem of this methodical is musical noise that is from the quicken coming and going of waves over consecutive frames.

Wiener filter reduces the estimation error but the drawback is the fixed frequency response at all frequencies leads musical noise [12].

The wiener filter with Lagrange multiplier [1,12] makes tradeoff between speech distortion and residual noise only when the value of Lagrange multiplier is greater than or equal to zero. If it is large would produces more speech distortion and less residual noise or if it is small would produces less speech distortion and more residual noise.
The perceptual speech enhancement [6,7] performs better than non perceptual enhancement method. By the use of Lagrange multiplier and perceptual wiener filter for minimizing the speech distortion while constraining the noise distortion fall below a constant threshold value leads a non linear relationship between Lagrange multiplier and threshold value causes noise distortion [12].

The proposed method uses the Lagrange multiplier with weighted perceptual wiener de-noising technique to choose the better Lagrange multiplier maintains linear relationship between Lagrange multiplier and threshold value then it results better perceptual quality, the speech distortion and noise distortion are reduced without degrading the clarity of enhanced speech signal.

BASIC WIENER FILTER IS USED FOR NOISE DIMINISHING IN SPEECH ENHANCEMENT

Let the input speech signal be the noisy speech signal [12] can be expressed as

$$y(n) = c(n) + a(n) \quad (1)$$

Where *c(n)* is the original clean speech signal and *a(n)* is the additive contingent noise signal, interrelated with the original signal. By applying DFT to the observed signal gives

$$Y(i,k) = C(i,k) + A(i,k) \quad (2)$$

Where *i*=1,2,…….I is the frame index, *k*=1,2,…….K is the frequency bin index, I is the total number of frames and K is the frame length. The short time spectral components of the *y(n),c(n)* and *a(n)*represented as *Y(i,k),C(i,k)* and *A(i,k)*respectively.

An estimate of clean speech spectrum $\hat{C}(i,k)$ is obtained by multiplying the filter gain function with noisy speech spectrum is given as

$$\hat{C}(i,k) = H(i,k)Y(i,k) \quad (3)$$

Where *H(i,k)* is the noise suppression gain function of conventional wiener filter and can be expressed as

$$H(i,k) = \frac{\xi(i,k)}{1+\xi(i,k)} \quad (4)$$

Where *(i,k)* is a priori SNR calculation defined as

$$\xi(i,k) = \frac{\Gamma_c(i,k)}{\Gamma_a(i,k)} \quad (5)$$

$$\Gamma_a(i,k) = E\{|A(i,k)|^2\} \quad (6)$$

$$\Gamma_c(i.k) = E\{|C(i,k)|^2\} \quad (7)$$

Equation (6) and (7) represents the predicted panorama of noise power and panorama of clean speech power respectively. A posteriori SNR can be estimated as

$$\gamma(i,k) = \frac{|Y(i,k)|^2}{\Gamma_a(i,k)} \quad (8)$$

In DD approach an estimate of the current a priori SNR is estimated by using the speech spectrum estimated in the previous frame and the a priori SNR accompanies the a posteriori SNR with a delay of one frame. This delay causes undesired gain distortion and thus generates the audible distortion during abrupt transient periods. To avoid this we can use modified a priori SNR, in this α will be changed dynamically and is expressed as follows.

If Δ(k) > Thrld

$\propto_M (i.k) = \propto$ then

$$\hat{\xi}_M(i,k) = \propto \frac{|H(i-1,k)Y(i-1,k)|^2}{\Gamma_a(i.k)} + (1-\propto)P(\gamma(i,k)-1) \quad (9)$$

else

$$\hat{\xi}_M(i,k) = \propto_M (i,k) \frac{|H(i-1,k)Y(i-1,k)|^2}{\Gamma_a(i.k)} + (1-\propto_M (i,k)P(\gamma(i,k)-1) \quad (10)$$

Where $0 < \alpha_M(i, k) < 1$ is the modified factor depends on the previous a posteriori SNR and is having a chance by the following affinity

$$\propto_M (i,k) = \frac{1}{1+\left[\frac{\Delta\gamma(k)}{\max(\gamma(i,k),\gamma(i-1,k))+1}\right]^2} \quad (11)$$

Where $\Delta\gamma(k) = (\gamma(i,k) - \gamma(i-1,k))$, the threshold Thrld= $E\{\gamma(i,k)\}$, *k*=1,2,…K is the spectral bin index and *i*= 1,2,……I is the frame index, K is the length of frame and I is number of frames. The noise suppression gain function is

$$H(i,k) = \frac{\hat{\xi}_M(i,k)}{1+\hat{\xi}_M(i,k)} \quad (12)$$

GAIN OF MODIFIED PERCEPTUAL WIENER FILTER

The gain function of the modified perceptual wiener filter $H_M(i,k)$ is calculated by using cost function, J which is expressed as

$$J = [\left|\hat{C}(i,k) - C(i,k)\right|^2] \quad (13)$$

ISSN: 2277-9655
Impact Factor: 4.116

Substituting equation (2) & (3) in (13) gives

$$J = E\{|(H_M(i,k)-1)C(i,k) + H_M(i,k)A(i,k)|^2\}$$

$$J = e_d + e_r \quad (14)$$

Where $e_d = (H_M(i,k)-1)^2 E[|C(i,k)|^2]$ and $e_r = H_M^2(i,k)E[|A(i,k)|^{2]}$ represents the distortion energy and residual noise energy. If the residual noise is less than the auditory masking threshold then only we can make it inaudible otherwise it is audible. To make this inaudible the constraint is given as

$$e_r \leq T_M(i,k) \quad (15)$$

By using above constraint and substituting $\Gamma_a(i,k) = E\{|A(i,k)|^2\}$ and $\Gamma_c(i,k) = E\{|C(i,k)|^2\}$ in the (13) cost function will results as

$$J = (H_M(i,k)-1)^2\Gamma_c(i,k) + H_M^2(i,k)\{\max[(\Gamma_a(i,k) - T_M(i,k)),0]\} \quad (16)$$

The modified perceptual wiener filter gain function can be obtained by differentiating the J with respect to

$$H_M(i,k) = \frac{\Gamma_c(i,k)}{\Gamma_c(i,k)+max\Gamma_a(i,k)-T_M(i,k),0)} \quad (17)$$

By multiplying and dividing the above equation with $\Gamma_a(i,k), H_M(i,k)$ will gives

$$H_M(i,k) = \frac{\hat{\xi}_M(i,k)}{\hat{\xi}_M(i,k)+\frac{\max(\Gamma_a(i,k)-T_M(i,k),0)}{\Gamma_a(i,k)}+\mu(i,k)} \quad (18)$$

$$\hat{\xi}_N(i,k) = \hat{\xi}_M(i,k) + \frac{\max(\Gamma_a(i,k)-T_M(i,k),0)}{\Gamma_a(i,k)} \quad (19)$$

Substituting the eq.(19) into eq.(18) we get

$$H_M(i,k) = \frac{\hat{\xi}_M(i.k)}{\hat{\xi}_N(i,k)+\mu(i,k)} \quad (20)$$

Where $T_M(i,k)$ is the noise masking threshold, it is estimated based on noisy speech spectrum.

THE LAGRANGE MULTIPLIER

To minimize the speech distortion energy in the frequency domain while maintaining the energy of residual noise below the preset threshold the Lagrange multiplier is in [12] used.

The Lagrange multiplier [12,1] creates the tradeoff between the speech distortion and residual noise. If the value of μ is large would produce more speech distortion and less residual noise. If the value of μ is small would produce less speech distortion and more residual noise.

The value of μ have to made based on the estimated a priori SNR $\hat{\xi}_M(i,k)$ is derived as

$$\mu(i,k) = 1 + U_0(1 - \frac{1}{1+e^{-\xi_{db}(i,k)}}) \quad (21)$$

Where $\xi_{db}(i,k) = 10\ log_{10}\hat{\xi}_M(i,k)$ and U_0 is constant chosen experimentally.

WEIGHTED PERCEPTUAL WIENER FILTER

The perceptual wiener filter causes some residual noise, due to fact that only noise greater than the noise masking threshold is percolated and below the noise masking threshold is remain, there is no guarantee for this whether it is

audible or not. The Lagrange multiplier value should be better for reducing speech distortion and residual noise, also for avoiding the non linear relationship between the μ and threshold value.
To overcome these drawbacks we proposed to weight the perceptual wiener filter with Lagrange multiplier using a psychoacoustic motivated weighting filter [14] and is given as

$$W(i,k) = \begin{cases} H(i,k), if\ ATH(i,k) < \Gamma_a \leq T_M(i,k) \\ 1,\ otherwise \end{cases} \quad (22)$$

Where *ATH (i,k)* is the absolute threshold of hearing. The gain function for the proposed weighting factor is given as

$$H_{M1}(i,k) = H_M(i,k)W(i,k) \quad (23)$$

PERFORMANCE EVALUATION AND COMPARISON

To compare and judge the value of the performance of the proposed speech enhancement scheme with different existing schemes, the simulation results are carried out with NOIZEOUS. It is a noisy speech corpus for evaluation of speech enhancement algorithms and database [11]. The data base is made up of 30 IEEE sentences by three manly and three womanly speakers corrupted by eight different real world noises with different SNRs. The quality of speech signals were degraded by several types of noises at global SNR levels of 0dB, 5dB, 10dB and 15dB. In this evaluation only seven noises are considered those are Babble, airport, car, exhibition, station, train and street.

The objective quality measures for the proposed speech enhancement method are the segmental SNR, the PESQ and the WSS measures. These parameters are more accurate to indicate speech distortion than overall SNR. The higher value of segmental SNR indicates debilitated speech distortion. The superior PESQ score reveals the better perceptual quality. The lower value of WSS indicates the weaker speech distortion. The performance of the proposed method is compared with the spectral subtraction, wiener filter and wiener filter with Lagrange multiplier.

The simulation results are compared in the Table1, Table2 and Table3. The observation of the simulation results in the table shows the proposed method have the better and accurate readings compared to existing methods.

Table.1 The Output Average Segmental SNR values of Enhanced Signals

Noise Type	Input SNR(dB)	Spectral Subtraction(dB)	Wiener Filter(dB)	Wiener With Lagrange(dB)	Lagrange with WPWF (dB)
Babble	0	-3.7366	-1.8405	-0.9066	-0.2191
	5	-2.2737	-0.3774	0.0451	0.0930
	10	-0.7052	0.4416	0.7345	1.2208
	15	-0.7395	2.0856	2.1762	2.1962
Airport	0	-3.7925	-1.5080	-0.3972	-0.0386
	5	-2.1992	-0.0356	0.4456	0.4995
	10	-0.3340	1.5529	1.4877	1.5529
	15	1.7845	2.1822	2.1717	2.1864
Car	0	-3.6392	-0.6348	-0.0475	0.1025
	5	-2.1566	0.3487	0.3589	0.3619
	10	-0.8882	0.5247	1.0078	1.0266
	15	0.9473	2.7556	2.7535	2.7556
Exhibition	0	-3.7311	-0.8684	-0.0706	-0.0550
	5	-2.2019	0.0223	0.1606	0.1662
	10	-1.1983	0.8914	0.8372	0.9131
	15	0.8604	2.2319	1.9237	1.9569
Station	0	-3.8588	-0.8558	-0.6168	-0.2135
	5	-2.2019	0.3735	0.4093	0.4113

ISSN: 2277-9655
Impact Factor: 4.116

	10	-1.1983	1.3478	1.3461	1.4257
	15	0.8604	1.1201	1.1799	1.2588
Train	0	-3.8072	-1.4851	-0.7514	-0.4405
	5	-2.2472	0.0446	0.3496	0.6896
	10	-0.3731	1.5905	1.5905	1.7507
	15	0.7646	2.2021	2.2305	2.3521
Street	0	-3.6131	-0.1100	0.1753	0.1838
	5	-1.6857	-1.8200	0.2395	0.4751
	10	-0.4977	2.3536	2.3541	2.4219
	15	1.2401	1.8658	1.9311	2.0837

Table.2 The Output PESQ values of Enhanced Signals

Noise Type	Input SNR(dB)	Wiener Filter(dB)	Wiener With Lagrange(dB)	Lagrange with WPWF (dB)
Babble	0	1.2206	1.2467	1.2846
	5	1.7275	1.6316	1.8263
	10	2.0344	1.8970	2.0544
	15	2.1269	2.0014	2.2673
Airport	0	1.4723	1.5026	1.5173
	5	1.4925	1.6160	1.7964
	10	2.0259	2.0132	2.0559
	15	2.2498	2.2239	2.2562
Car	0	1.1658	1.3978	1.4124
	5	1.6946	1.7250	1.7824
	10	1.9212	1.8935	1.9932
	15	2.2653	2.2623	2.2656
Exhibition	0	0.9098	1.2181	1.2185
	5	1.4547	1.4562	1.4769
	10	1.9846	1.8863	1.9956
	15	2.1307	2.1486	2.2132
Station	0	0.9176	1.0134	1.0193
	5	1.6663	1.6671	1.6772
	10	2.0880	2.0663	2.0981
	15	1.9949	1.9957	2.1020
Train	0	1.4509	1.4923	1.5332
	5	1.6808	1.6685	1.7116
	10	2.0087	2.0088	2.0179
	15	2.0040	2.0042	2.0164
Street	0	1.6364	1.6524	1.7052
	5	1.6797	1.7019	1.7057
	10	2.1197	2.1448	2.1577
	15	2.3809	2.3167	2.3994

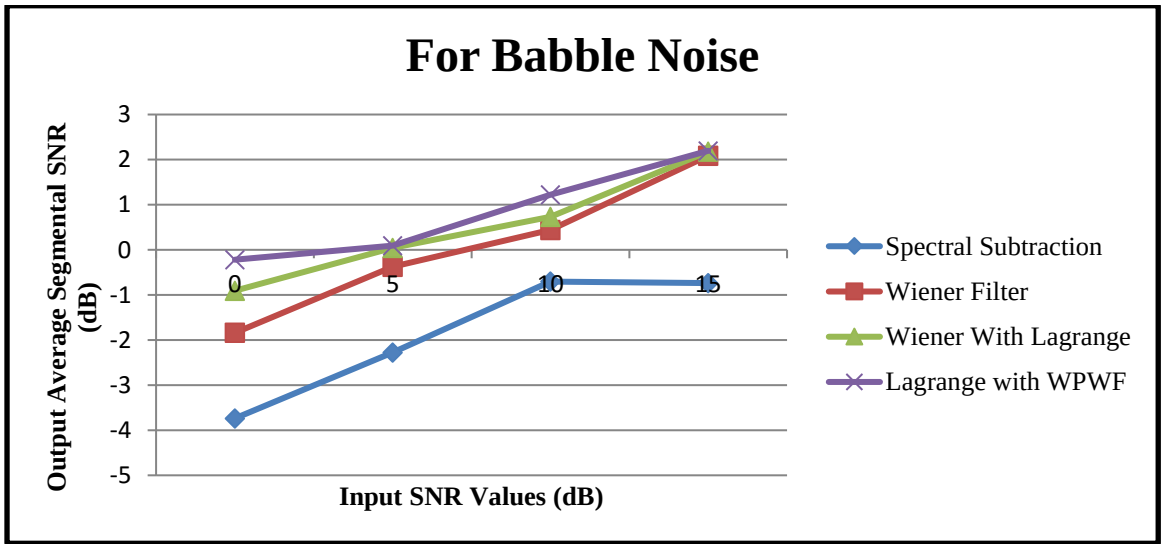
Table.3 The Output WSS values of Enhanced Speech Signals

Noise Type	Input SNR(dB)	Wiener Filter(dB)	Wiener With Lagrange(dB)	Lagrange with WPWF (dB)
Babble	0	119.809	113.3895	109.2569
	5	112.2422	105.6359	101.5174
	10	93.4644	91.5866	87.8359
	15	83.5530	76.9853	73.5549

ISSN: 2277-9655
Impact Factor: 4.116

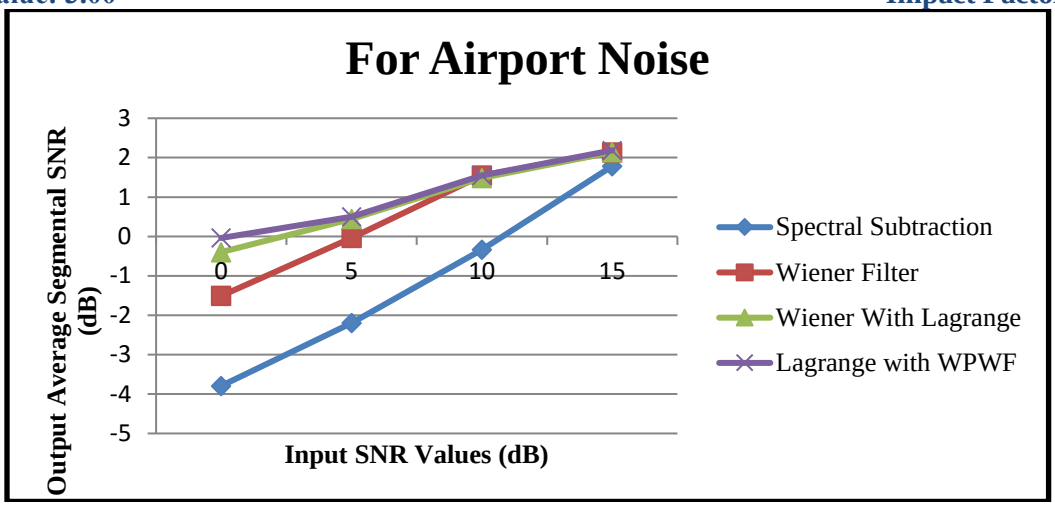
Airport	0	119.7087	112.8996	105.9685
	5	115.5162	107.1717	100.2342
	10	83.8046	77.1962	70.2193
	15	75.7421	73.6003	69.5217
Car	0	117.6957	109.1681	104.7279
	5	97.8260	92.7606	87.6798
	10	89.7359	84.6176	79.1301
	15	79.5697	70.7176	67.6631
Exhibition	0	119.1785	109.2467	104.6194
	5	117.5452	107.0087	98.2400
	10	99.9124	91.2562	83.3784
	15	81.6266	73.4110	70.9322
Station	0	122.0165	112.4626	109.0839
	5	103.5956	96.4847	91.9303
	10	86.1690	79.0295	74.9699
	15	96.9370	92.8625	83.7680
Train	0	109.3028	106.3934	101.0337
	5	98.9948	95.8433	89.3867
	10	87.0933	85.5740	80.5306
	15	73.7194	70.8510	67.2351
Street	0	104.9760	99.8952	94.1306
	5	102.8184	97.6329	91.8218
	10	81.6440	79.5334	75.2053
	15	75.7298	74.8156	71.9770

The graphical representation for the measured values in the Tables 1, 2, 3, are as follows in figure 1, 2, 3, respectively. Each figure consists of different input SNR values are compared with respective parameter. Figure 1 represents the Average Segmental SNR, figure 2 represents the PESQ values, figure 3 represents the WSS values of the enhanced signals.

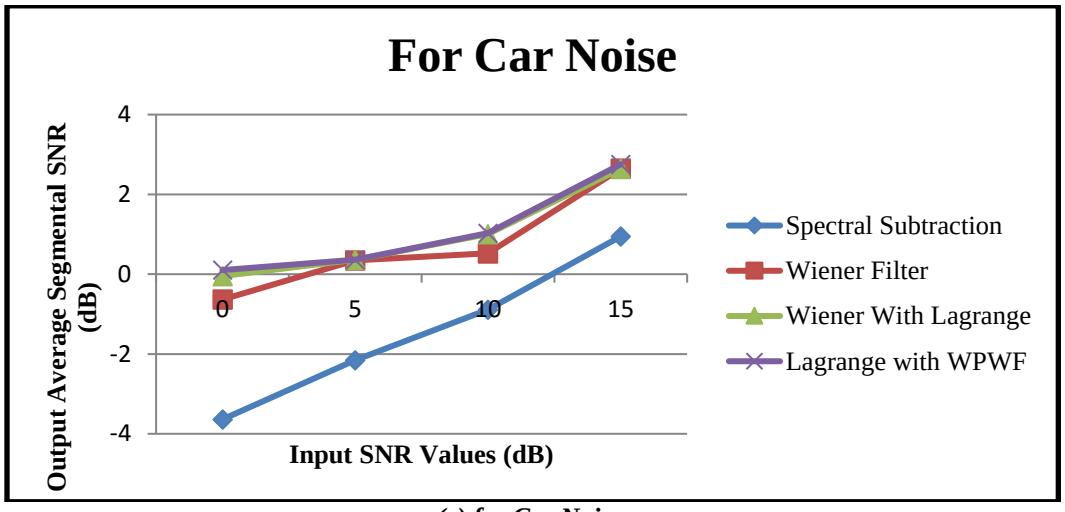


(a) for Babble Noise

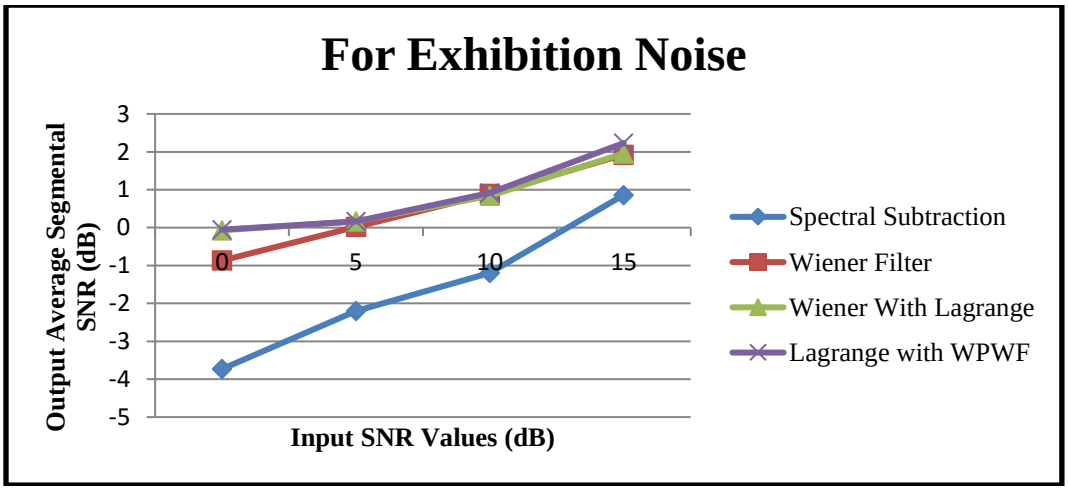
ISSN: 2277-9655
Impact Factor: 4.116



(b) for Airport Noise.

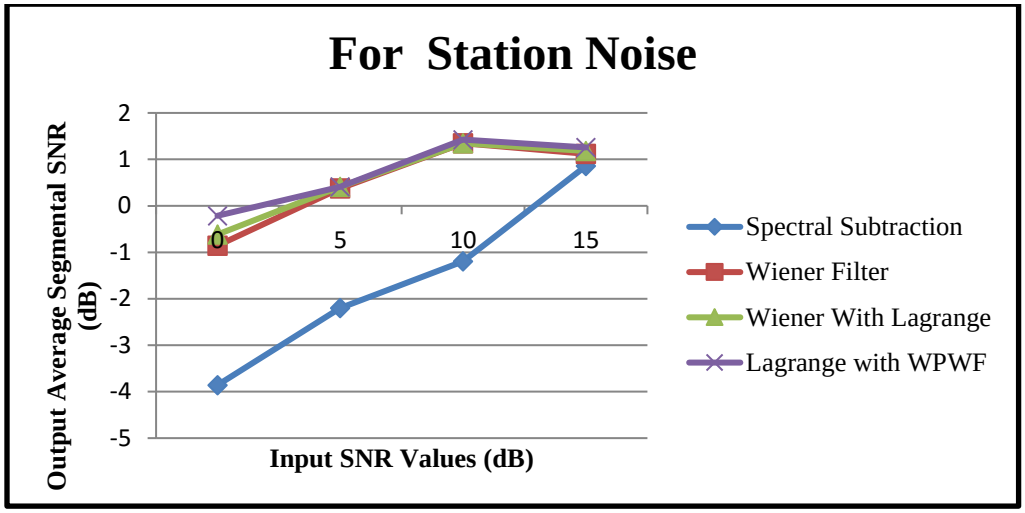


(c) for Car Noise.

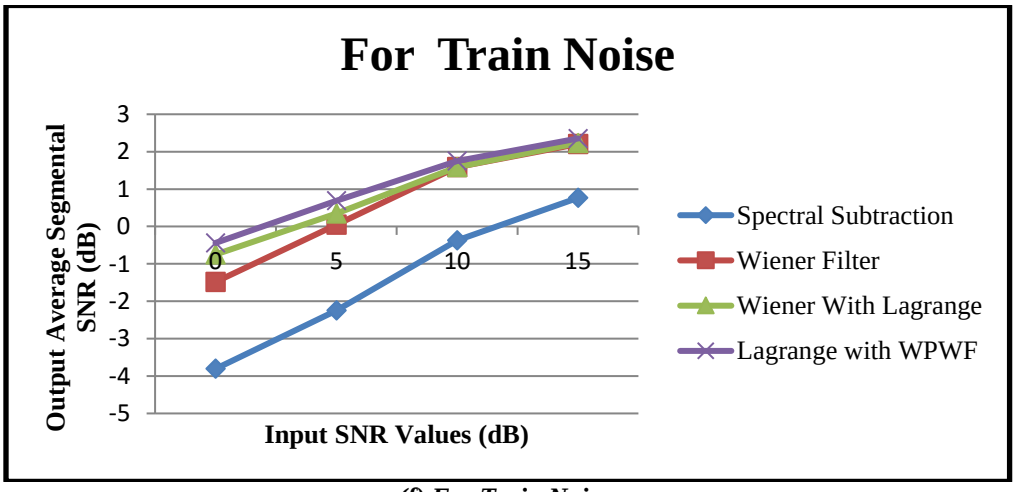


(d) For Exhibition Noise

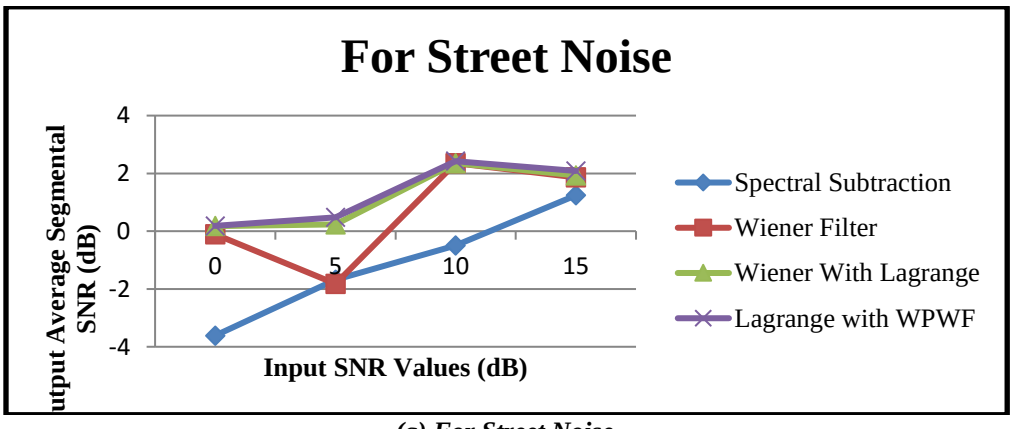
ISSN: 2277-9655
Impact Factor: 4.116



(e) For Station Noise.



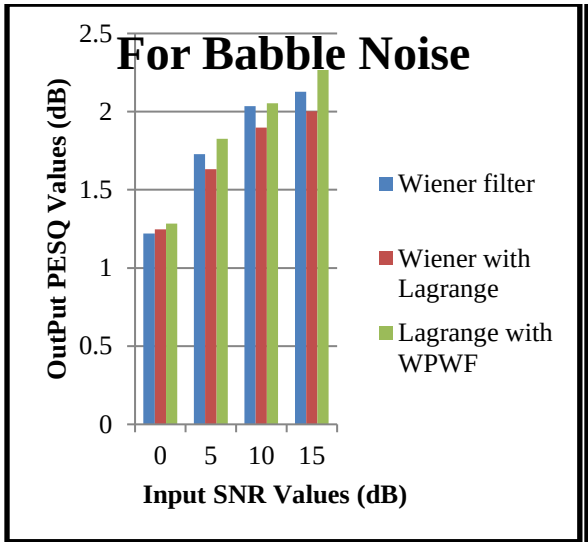
(f) For Train Noise.



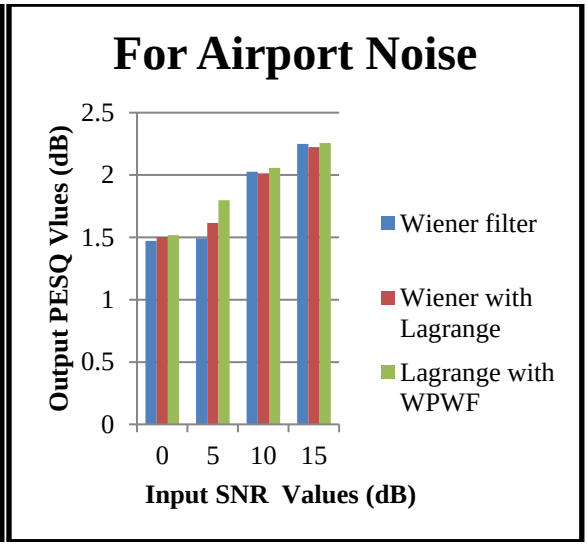
(g) For Street Noise.

Fig.1 The graphical representation of comparison of Output Average segmental with Input SNR for different Noises are in (a) to (g).

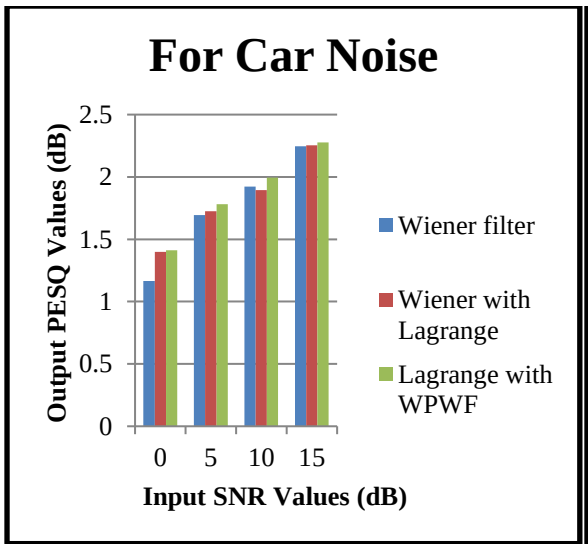
ISSN: 2277-9655
Impact Factor: 4.116



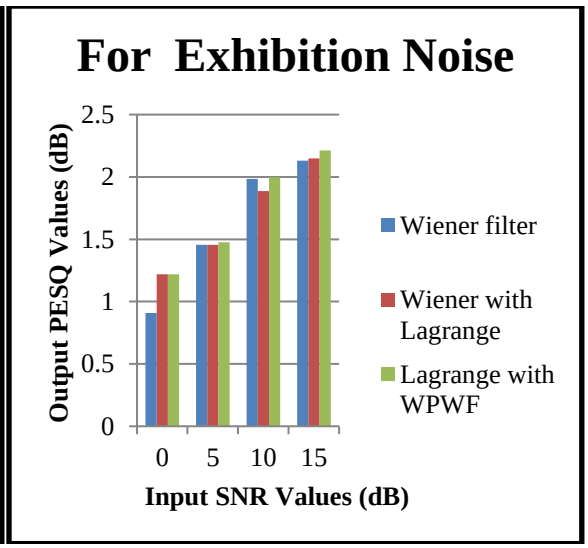
(a) Babble Noise



(b)Airport Noise

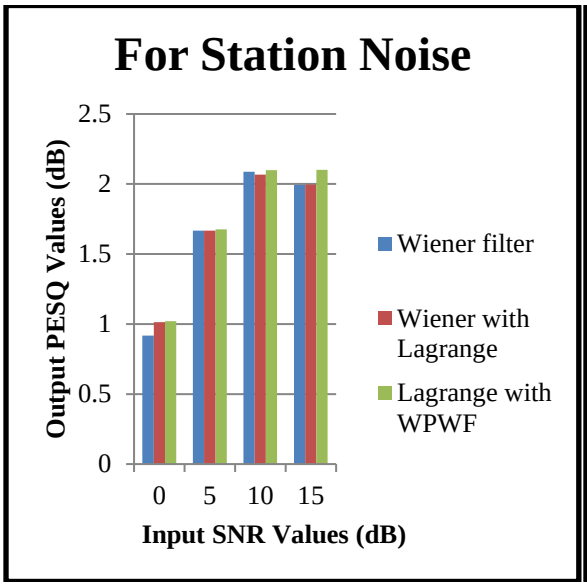


(c) Car Noise

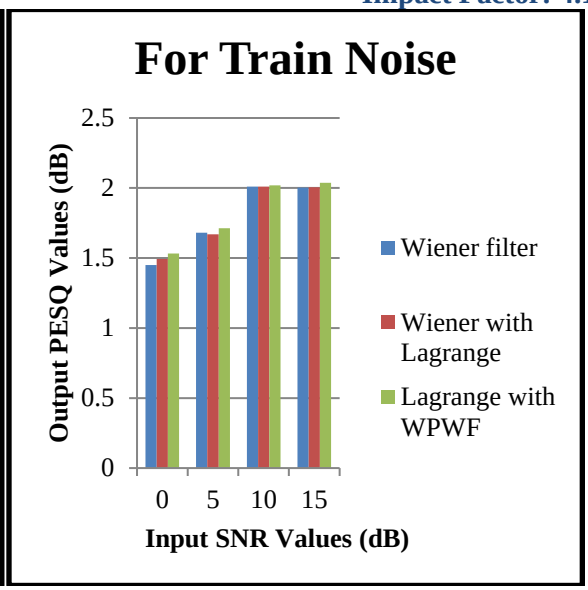


(d) Exhibition Noise

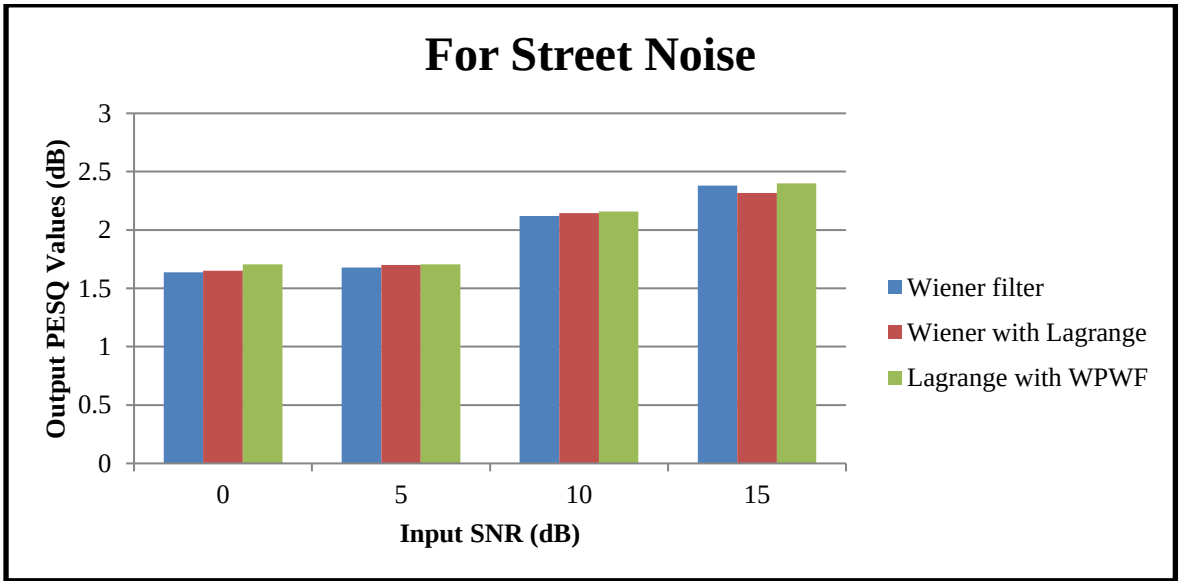
ISSN: 2277-9655
Impact Factor: 4.116



(e) Station Noise



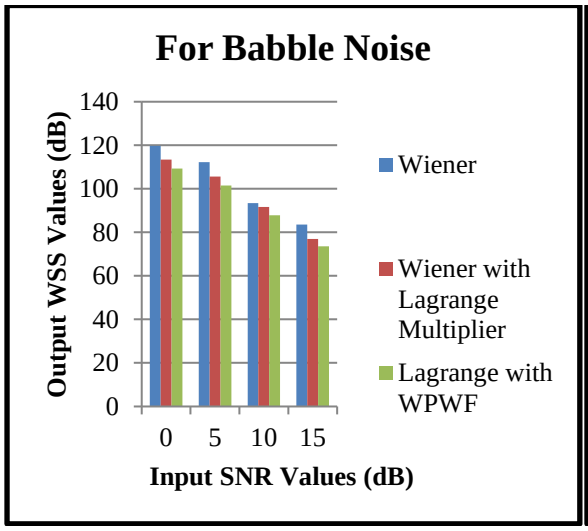
(g) Train Noise



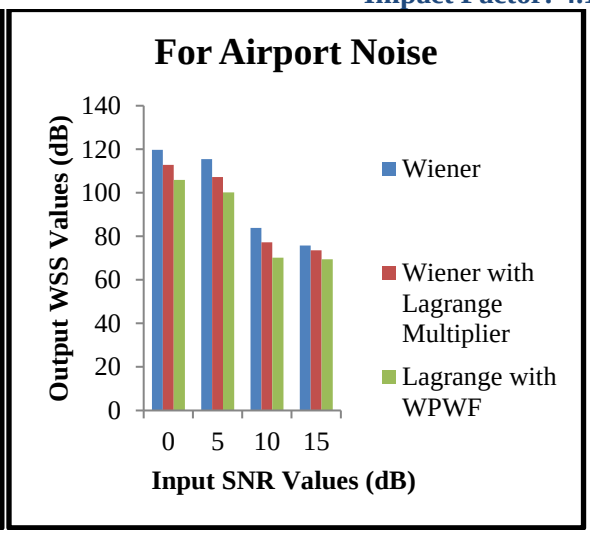
(g) Street Noise

Fig.2 The PESQ values of different methods are compared with different Input SNRs in (a) to (g).

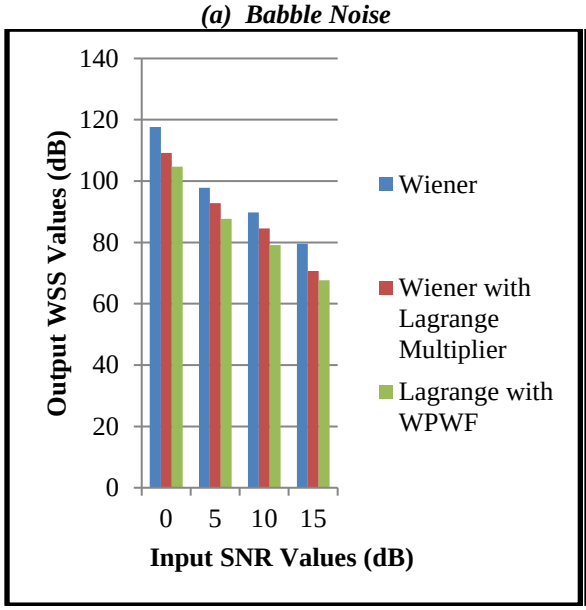
ISSN: 2277-9655
Impact Factor: 4.116



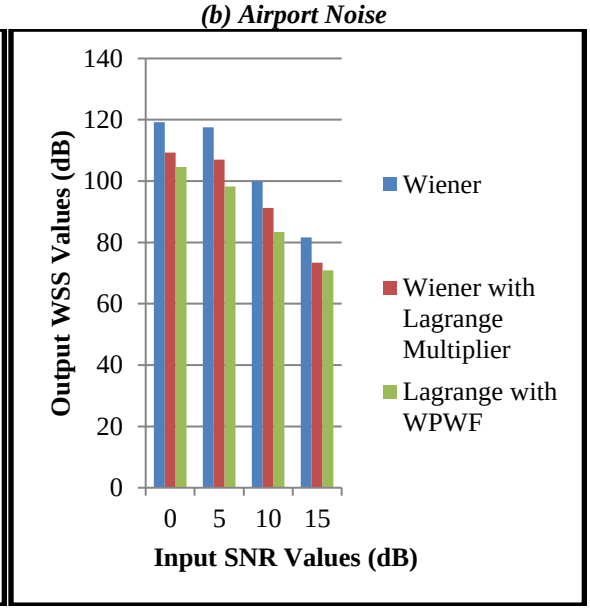
(a) Babble Noise



(b) Airport Noise

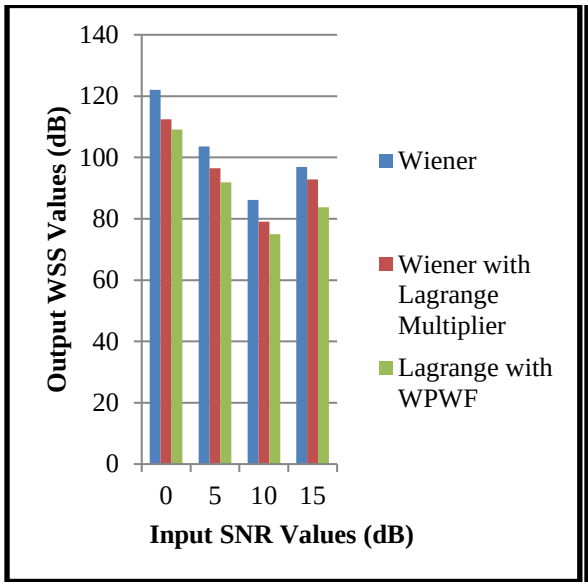


(c) Car Noise

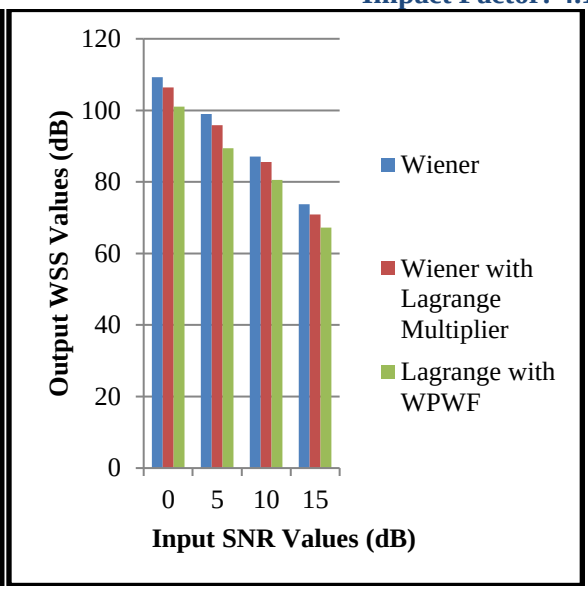


(d) Exhibition Noise

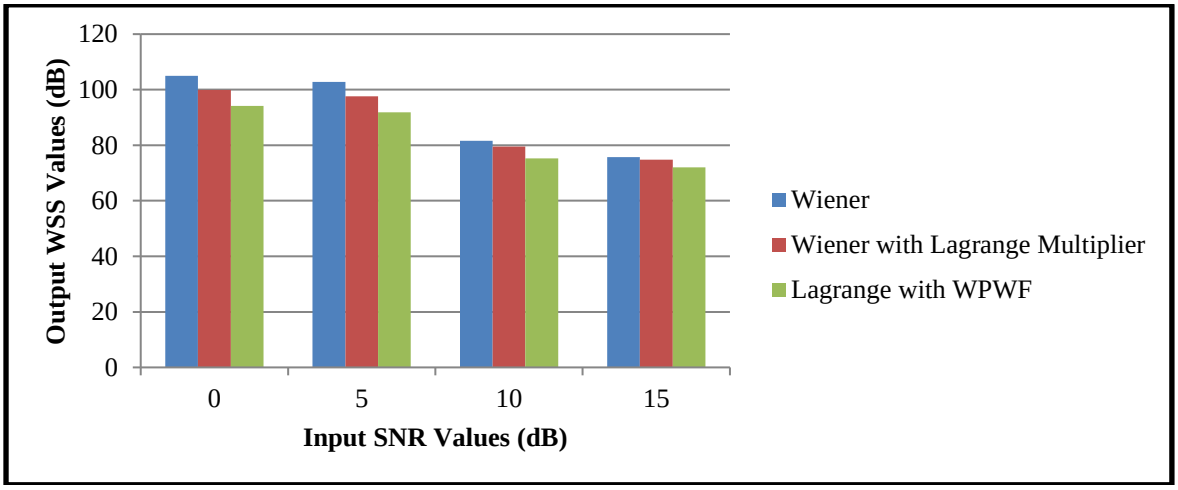
ISSN: 2277-9655
Impact Factor: 4.116



(e) Station Noise



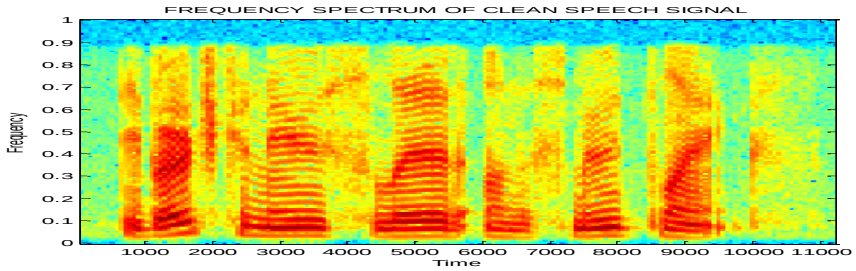
(f) Train Noise



(g) Street Noise

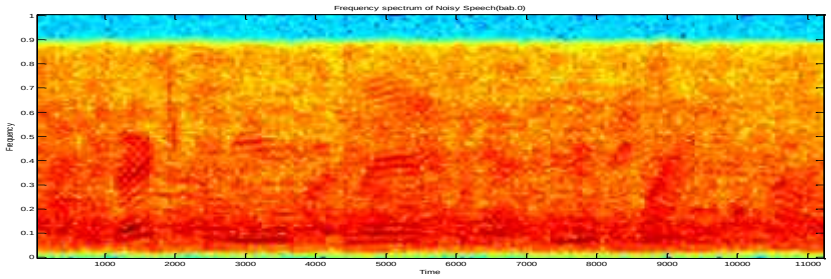
Fig.3 The graphical representation of Output WSS Vs. Input SNR are compared for different methods in (a) to (g).

The spectrograms of the clean speech signal, noisy speech signal and different enhanced speech signals tells the perceptual quality is improved, and noise is reduced compared to the existing methods are shown in figure 4.

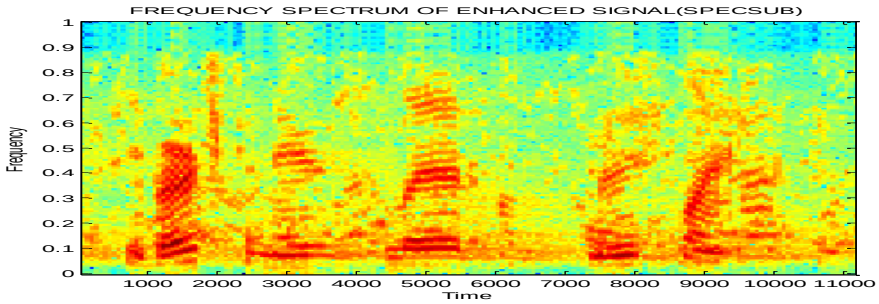


(a) Original clean speech signal.

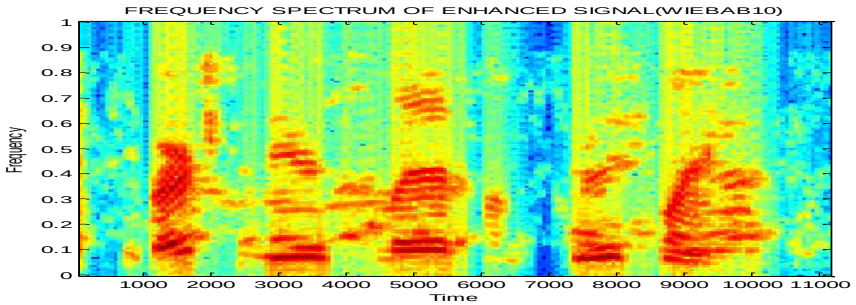
ISSN: 2277-9655
Impact Factor: 4.116



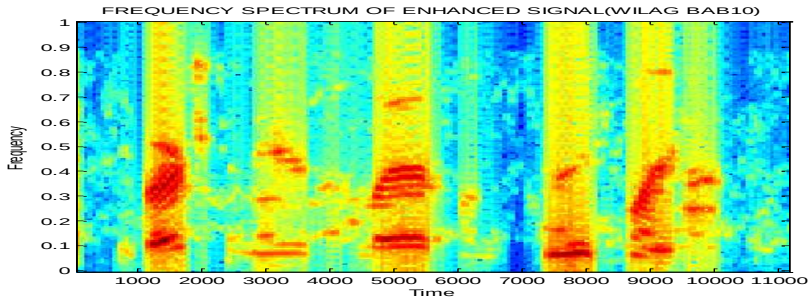
(b) Noisy signal (Babble noise SNR=10dB).



(c) Enhanced signal using Spectral Subtraction.

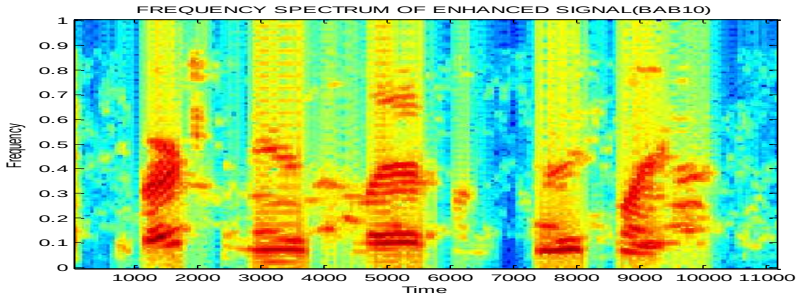


(d) Enhanced Signal using Wiener filter.



(e) Enhanced Signal using Wiener with Lagrange.

ISSN: 2277-9655
Impact Factor: 4.116



(f) Enhanced Signal using Weighted PWF with Lagrange Multiplier.

Fig 4. Speech Spectrograms, Babble Noise with input SNR=10dB (a) Original clean speech signal (b) Noisy signal(Babble noise SNR=10dB) (c) Enhanced signal using Spectral Subtraction (d) Enhanced Signal using Wiener filter (e) Enhanced Signal using Wiener with Lagrange (f) Enhanced Signal using Weighted PWF with Lagrange Multiplier.

CONCLUSION

In this speech enhancement process, by the use of Lagrange multiplier and psychoacoustic motivated weighting factor the noise below the noise masking threshold is filtered, the noise due to the non linearity between the Lagrange multiplier and threshold value is avoided; the speech distortion and residual noise are reduced, the better perceptual quality is achieved.

REFERENCES

[1] M. Priyanka, CH.V. Rama Rao, "Speech Enhancement using Self Adaptive Wiener Filter based on Hybrid a priori SNR", proceedings of ICNAE & Advanced computing(ICNEAC-2011),pp.1-6,2011.
[2] Tsai-Tsung Han, Pei-Yun Liu, "A Speech Enhancement System using Binary Mask Approach and Spectral Subtraction Method" IEEE International Symposium on computer, Consumer and Control(ISCCC),pp.1065-1068,2014.
[3] Craig A. Anderson Paul D. Teal Mark A. Poletti, "Multi channel Wiener Filter Estimation Using Source Location Knowledge for Speech Enhancement".2014 IEEE Workshop on Statistical Signal Processing (SSP),pp.57-60,2014.
[4] Laksmikanth. S, Natraj. K. R, Rekha. K. R, "Noise cancellation in Speech Signal Processing –A Review. International Journal of Advanced Research in Computer and Communication Engineering, Vol.3, Issue 1,pp.5175-5186, January 2014.
[5] Vyankatesh Chapke, Prof.Harjeet Kaur, "Review of Speech Enhancement Techniques using Statistical Approach", International Journal of Electronics Communication and Computer Engineering, Volume 5,pp.307-309, Issue(4) July, Technovision-2014,ISSN 2249-071X.
[6] S. Alaya, N. Zoghlami, and Z. Lachiri, "Speech Enhancement based on perceptual filter bank improvement" International Journal of Speech Technology,pp.1-6,2014.
[7] Astik Biswas and P. K. Sahu, Anirban Bhowmick and Mahesh Chandra, "Acoustic Feature Extraction using ERB like Wavelet Sub-band Perceptual Wiener Filtering for Noisy Speech Recognition", 2014 Annual IEEE India Conference(INDICON).
[8] V. Sunnydayal, T. Kishore Kumar, "Speech Enhancement using Sub-Band Wiener Filter with Pitch Synchronous Analysis",ICACCI,pp-20-25,IEEE-2013.
[9] R. Yu, "A Low –Complexity noise estimation algorithm based on smoothing of noise power estimation and estimation bias correction," in proc. IEEE Int. Confcoust. Speech, Signal process. Taipei, pp.4421-4424, April 2009.
[10] T. Gerkmann and R. C.Hendriks, "Unbiased MMSE-based noise power estimation with low complexity and low tracking delay," IEEE Trans. On Audio, Speech, and Language Process., Vol. 20, no. 4, pp. 1383-1393, May 2012.
[11] http:www.utdallas.edu/~loizou/speech/noizeus.
[12] P. C. Loizou, "Speech Enhancement: Theory and Practice. BocaRaton, FL: CRC press, 2007.

[13] J. Ma, Y.Hu, and P. Loizou, "Objective measures for predicting speech intelligibility in noisy conditions based on new band-importance functions," J. Acoustic. Soc. Am., Vol.125, no.5, pp.3387-3405, May 2009.
[14] Y. Hu and P. Loizou, "Incorporating a psychoacoustic model in frequency domain speech enhancement," IEEE signal processing letters, vol 11(2), pp.270-273, 2004.
[15] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean square error short-time spectral amplitude estimator," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-32, pp.1109-1121, Dec 1984.
[16] P. K. Daniel Lun, Tai-Chiu Hsung, "Improved Wavelet Based A-Priori SNR Proc. IEEE International Symposium on Circuits and Systems (ISCAS), Paris, France, pp.2382-2385, May 2010
[17] Y. Hu and P. Loizou, "Evaluation of objective quality measures for speech enhancement," IEEE Trans. On Audio, Speech, and language Process., vol. 16, no.1, pp.229-238, 2008.
[18] R. Schwartz M. Berouti and J. Makhoul, "Enhancement of speech corrupted by acoustic noise," Proc.of ICASSP, vol.1, pp.208-211, 1979.
[19] Sana Alaya, Novelene Zoghlami, Zied Lachiri, "Speech enhancement using perceptual multi-band wiener filter," 1st International Conference on ATSIP, Sousse, Tunisia, pp.468-471, IEEE 2014.
[20] J.D. Johnston, "Transform coding of audio signals using perceptual noise criteria," IEEE Jour. Selected Areas Commun, vol.6, pp.314-323, February 1988.
[21] Philipos C. Loizou, Gibak Kim, "Reasons why current speech enhancement algorithms do not improve speech intelligibility and suggested solutions," IEEE Trans. On Audio, Speech, and Language Process., vol.19, no.1, pp.47-56, January 2011.