ISSN: 2584-0495 Vol. 2, Issue 1, pp. 505-514

International Journal of Microsystems and IoT

ISSN :(Online) Journal homepage: https://www.ijmit.org

Performance Comparison among Different Wiener Filter Algorithms for Speech Enhancement

Saloni Dwivedi, and Ajay Khunteta

Cite as: Dwivedi, S., & Khunteta, A. (2024). Performance Comparison among Different Wiener Filter Algorithms for Speech Enhancement. International Journal of Microsystems and IoT, 2(1), 505–514. https://doi.org/10.5281/zenodo.10700982

Published online: 22 January 2024.

DOI: https://doi.org/10.5281/zenodo.10700982

International Journal of Microsystems and IoT
Vol. 2, Issue 1, pp. 505-514; DOI: https://doi.org/10.5281/zenodo.10700982

Check for updates

Performance Comparison among Different Wiener Filter Algorithms for Speech Enhancement

Saloni Dwivedi, and Ajay Khunteta

Department of Electronics Engineering, Rajasthan Technical University, Kota, Rajasthan, India

ABSTRACT
Reconstruction of original speech signal from the noisy signal is still a difficult task as nature and characteristics of noise signal vary in time and also depends on application and context for which filtering has to be performed. There are different methods individually or also combined with other filters to get better performance in speech signal processing. Wiener filter is very popular for speech signal processing, it is a linear estimator and this makes it less complex and easy to handle. There are various algorithms to solve Wiener filtering problem. In the proposed paper, performance comparison among different wiener filter algorithms is performed in context of speech signal processing. The detailed analysis is presented for comparison among Implicit Wiener filter, Two Step Noise Reduction Wiener filter and Wiener filter with Harmonic Regeneration. The performance is evaluated for both male and female speech samples. Different non-stationary noise (airport, babble, car, station, street, train, exhibition, restaurant) and stationary noise (AWGN) generated for different SNR values are considered. Performance evaluation is done through objective quality measures such as Log likelihood Ratio (LLR), Cepstrum Distance, Weighted Spectral Slope (WSS), and frequency weighted Segmental SNR in MATLAB.

KEYWORDS

Implicit Wiener Filter; Objective Measures; Two Step Noise Reduction; Harmonic Regeneration

1. INTRODUCTION

In natural environment it is nearly impossible to record speech signal in its pure form. Therefore, speech filtering and speech enhancement techniques are mandatory for faithful reproduction of desired signal in speech recognition and communication systems. Noise reduction techniques must be employed in wide variety of applications and still a very challenging issue because of varying characteristics of noise and speech signal. As power spectrum of speech changes over time it is highly non-stationary. But if it is studied in terms of short segments its' properties are nearly stationary. Noise can be of different types from different sources that we encounter in daily life such as restaurant noise, babble noise, train noise etc. Speech enhancement is the combination of algorithms and techniques which aims to improve speech quality, intelligibility and hearing fatigue to listener by removing background noise. Speech enhancement is vital step in speech processing [1]. In the speech enhancement other than applying actual filtering algorithm to speech signal various pre and post processing steps are also involved. In frame blocking process speech signal is divided into frames of N samples, which overlap with adjacent frames by M samples [2]. During blocking process each frame is multiplied with window function to minimize spectral distortion [3]. In the experimental result of this paper 50% overlap is used for framing. For speech synthesis overlap- add method is used after filtering [4].

In this paper performance comparison among different wiener algorithms based on different approaches such as estimation of a- priori SNR and estimation of noise power spectral density is performed. Rest of the paper is organized as follows. Section 2 describes related work. Section 3 describes basic overview of Wiener filter. Section 4 describes different wiener filtering algorithms. Section 5 presents objective measures used in this paper for filter performance evaluation. In section 6 results and possible discussions are presented. Section 7 describes conclusion and possible extension to the work.

2. RELATED WORK

Over the course of period various algorithms were developed through different approaches to derive wiener filter transfer function in both time and frequency domain. Yihan Wang [5] shows research progress made in speech enhancement from conventional algorithms to AI based algorithms. By replacing conventional FFT with R22 SDF FFT Mr. C. Ramesh Kumar, Dr MP Chitra [6] proposed modified wiener filter in frequency domain. Yi Hu, Philipos C. Loizou [7] evaluated the performance of various objective measures in terms of speech quality [8]. Jacob Benesty, J. Chen, Yiteng Huang [9] develops a new widely linear noise – reduction Wiener Filter based on the variance and pseudo- variance of the short time Fourier Transform coefficients of speech signal. V. Sailaja, P. Sunitha, B. Vasantha Lakshmi, V. Prasanth [10] presented an adaptive Wiener filter method to predict speech quality in presence of highly non- stationary scenarios in frequency domain. Maximum-Likelihood Estimation and BASIYAN MMSE estimation were studied and compared by Dr. China Venkateshwarlu Sonagiri, K. Satya Prakash, D. SubbaRami Reddy [12] for wiener filtering. Naveen Upadhyay, Rahul Kumar Jaiswal [13] proposed a recursive noise estimation technique for noise estimation which can be used in Wiener filter noise estimation step. Cyril Plapous, Claude Marro, Pascal Scalart [14] introduces two step noise reduction (TSNR) and Harmonic Regeneration (HRNR) [15]
.

Jaiswal, R, Romero, D. [16] develops an algorithm that recursively estimates the noise power spectral density and reconstruct the target speech signal in the frequency domain by implicit wiener filter. R. K. Jaiswal, Sreenivasa Reddy, Linga Reddy [18] proposed implicit wiener filter algorithm and analysed its performance in presence of stationary and nonstationary noise

3. WIENER FILTER

Filtering is a process where a system is used to generate an output signal with some specific requirements by reshaping the frequency components of the input. Filters can be either linear or non-linear. A linear system is completely characterized by its' transfer function [18]-[20]. By optimizing the transfer function by suitably selecting the filter parameters it can be assured that filter output is best match of desired signal. There can be two approaches to this transfer function optimization, statistical or deterministic [21]. Wiener filter is based on statistical approach where the focus is to minimize the mean squared error (MSE). Wiener filter transfer function can be expressed both in time and frequency domain [22]. In time domain noisy speech signal for speech enhancement applications can be given as equation (1):

$$y(n) = x(n) + n(n) \tag{1}$$

where x(n) = clean speech signal

n(n) = noise signal

y(n) = noisy speech signal

Wiener Filter for speech signal processing in time domain can be given as equation (2)

$$h^* = (R_{xx} + R_{nn})^{-1} r_{xx} \tag{2}$$

Where: R_{yy} = autocorrelation matrix of input noisy signal

R_{nn} = autocorrelation matrix of noise signal

Wiener filter transfer function for speech enhancement in frequency domain can be given as equation (3)

:

$$H(\omega_K) = \frac{P_{xx}(\omega_k)}{P_{xx}(\omega_k) + P_{nn}(\omega_k)} \tag{3}$$

Where, $P_{xx}(\omega_k)$ = Power spectral density of clean speech

$P_{nn}(\omega_k)$ = Power spectral density of noise signal

Constraints are here that power spectral densities $P_{xx}(\omega_k)$ and $P_{nn}(\omega_k)$,have even symmetry which makes above transfer function real, nonnegative and even

Wiener filter can also be expressed as a f-unction of a-priori SNR (ξ_k). Where a-priori SNR can be expressed as a function of power spectral density of clean speech $P_{xx}(\omega_k)$ and power spectral density of noisy speech equation (4)

$P_{nn}(\omega_k)$:

$$\xi_k \triangleq \frac{P_{xx}(\omega_k)}{P_{nn}(\omega_k)} \tag{4}$$

Equation for wiener filter transfer function can be given as equation (5)

:

$$H(\omega_k) = \frac{\xi_k}{\xi_k + 1} \tag{5}$$

In this approach Wiener filter emphasizes high SNR portions of the spectrum and attenuates low SNR portions of the spectrum or we can say that each frequency component is attenuated in proportion to the estimated SNR.

4. WIENER FILTER APPROACHES

Filtering is a process where a system is used to generate an output signal with some specific requirements by reshaping the frequency components of the input [23]. Wiener filtering theory is based on stochastic framework which says that minimization of MSE can provide optimum filtering coefficients. The noisy speech signal for speech enhancement applications can be given as equation (6)

$$y(n) = x(n) + n(n) \tag{6}$$

where, x(n) is clean speech signal, n(n) is noise signal and d(n) is desired speech signal.

Wiener filter can also be expressed as a function of a-priori SNR. Where a-priori SNR can be expressed as a ratio of noise and clean signal power densities.

4.1 Two Step Noise Reduction:

In two step noise reduction algorithm a- priori SNR is computed in two steps which leads to improved noise reduction. Also, speech onsets and offsets are preserved and reverberation effect is reduced up to certain extent. In TSNR approach first gain is computed using Decision directed approach. In second step, a-priori SNR at next frame is estimated using this gain function using following mathematical relation equation (7)

:

$$SNR^{TSNR}_{prio,est}(p,k) = SNR^{DD}_{prio,est}(p+1,k)$$

$$= \beta' \frac{|G_{DD}(p,k)X(p,k)|^2}{\gamma_{n,est}} + (1 - \beta')P\,[SNR_{post,est}(p+1,k) - 1] \tag{7}$$

Above relation requires knowledge of future frame which can increase computational complexity and processing delay. Hence, we choose $\beta' = 1$, which is valid only for first frame. This choice of value for β'avoids delay caused due to term of future reference and also reduces musical noise caused due to decision directed approach equation (8):

$$G_{TSNR}(p,k) = \frac{SNR^{TSNR}_{prio,est}(p,k)}{1 + SNR^{TSNR}_{prio,est}(p,k)} \tag{8}$$

4.2 Harmonic Regeneration:

In harmonic regeneration step filtered signal with previous filtering step is further processed so that a fully harmonic signal can be created by regenerating all the speech harmonics which were suppressed by previous steps assuming them to be noise [24]. Then this regenerated signal will be used to refine a- priori SNR. A new spectral gain function $G_{HRNR}(p,k)$ using this refined a- priori SNR $SNR^{HRNR}_{prio,est}(p,k)$ is defined. The spectral gain is equation (9)

:

$$G_{HRNR}(p,k) = \frac{SNR^{HRNR}_{prio,est}(p,k)}{1 + SNR^{HRNR}_{prio,est}(p,k)} \tag{9}$$

4.3 Implicit Wiener:

Implicit Wiener filter is a variation of modified or parametric wiener filter. The mathematical relation for Implicit wiener filter can be given as equation (10)
:

$$H_{WF}(\omega) = \left[\frac{P_{xx}[\omega]}{P_{xx}[\omega]+\gamma\, P_{nm}[\omega]}\right]^{\beta} \quad (10)$$

Where, β is noise suppression factor and γ which is noise adjustable parameter is adjusted using segmental SNR and it control what amount of noise should be perceived. The equation for calculating γ can be given as: $\gamma = 4 - 0.15\, SNR_{seg}$.
The segmental SNR can be defined as

The segmental signal to noise ratio can be calculated in both time and frequency domain.

$$\text{SNRseg} = \frac{10}{M}\sum_{m=0}^{M-1} log_{10} \frac{\sum_{n=N_m}^{N_m+N-1} x^2(n)}{\sum_{n=N_m}^{N_m+N-1}(x(n)-y(n))^2} \quad (12)$$

where ,
x(n) = clean signal
y(n) = enhanced signal
N = frame length
M = number of frames
This equation for calculating segmental SNR is has a drawback that signal energy during silent frames is very small which leads to negative value of segmental SNR.
The segmental SNR values can also be calculated in frequency domain by following equation 13:

$$\text{fwSNRseg} = \frac{10}{M}\sum_{m=0}^{M-1}\frac{\sum_{j=1}^{K} W_j\, log_{10}\left[X^2(j,m)\big/(X(j,m)-Y(j,m))^2\right]}{\sum_{j=1}^{K} W_j}$$

(13)
where,
fwSNRseg = Segmental Signal to Noise Ratio in frequency domain
W_j = weight placed on jth frequency band
K = number of bands
M = Total number of frames
X (j, m) = filter bank amplitude of the clean signal in jth frequency band at the mth frame
Y (j, m) = filter bank amplitude of the enhanced signal in jth frequency band at the mth frame
Segmental SNR in frequency domain is more flexible than its time domain calculation as there is a scope for placing different weights for different frequency bands.

4.4 Log Likelihood Ratio

It is an objective measure based on dissimilarity between all pole models of clean and enhanced speech which can be defined as equation (14)

$$d_{LLR}(a_x, \bar{a_{\hat{x}}}) = \log\left(\bar{a}_{\hat{x}}^T R_x \bar{a_{\hat{x}}} / a_x^T R_x a_x\right) \quad (14)$$

Where
$a_x^T = [1, -\alpha_x(1), -\alpha_x(2), \ldots\ldots, -\alpha_x(p))$
$\bar{a}_{\hat{x}}^T = [1, -\alpha_{\hat{x}}(1), -\alpha_{\hat{x}}(2), \ldots\ldots, -\alpha_{\hat{x}}(p))$

$$SNR_{seg} = \frac{10}{M}\sum_{m=0}^{M-1}\log\left(1\ \frac{\sum_{n=N_m}^{N-1+N_m} x^2[n]}{\sum_{n=N_m}^{N-1+N_m}(x[n]-x_{enh}[n])^2}\right) \quad (11)$$

5 OBJECTIVE MEASURES

To calculate objective measures framing should be done and then distortion measures should be computed between original and processed speech signal. By averaging the distortion measures for each frame, a single global measure can be computed.

5.1 Segmental Signal to Noise Ratio

R_x = (p+ 1) * (p +1) autocorrelation matrix
Log likelihood ratio gives difference in formants peak locations of clean and filtered speech signal. . LLR must range from 0 to 2 in case of faithful reproduction of speech signal.[1]

5.1.1 Cepstrum Distance

It is a distance measure based on cepstrum coefficients and LPC coefficients. The cepstrum coefficients can be calculated from LPC coefficients. Cepstrum distance gives an estimate of the log spectral distance between two spectra [25.] If p is the order of LPC analysis, the mathematical equation to derive cepstrum coefficients can be given as following equation (15)

$$c(m) = a_m + \sum_{k=1}^{m-1}\frac{k}{m}\, c(k)\, a_{m-k} \qquad 1\le m\le p \qquad (15)$$

The expression for calculating cepstrum distance based on cepstrum coefficients derived in previous equation can be written as equation (16):

$$d_{cep}(c_x, \bar{c_{\hat{x}}}) = \frac{10}{log_e 10}\sqrt{2\sum_{k=1}^{p}[c_x(k) - c_{\hat{x}}(k)]^2} \qquad (16)$$

Where,
c_x(k) = cepstrum coefficients for clean signal
$c_{\hat{x}}$(k) = cepstrum coefficients for enhanced signal

5.2 Weighted Spectral Slope Distance Measure

WSS is an objective speech quality measure which is based on human auditory speech perception models and can produce results which are more, close to subjective quality measures. WSS was given by Klatt and is based on weighted differences between the spectral slopes in each band. It is based on differences in spectral peak locations and penalize it heavily while ignoring all other factors.
The weighted spectral slope can be computed with following mathematical expression equation (17):
:

$$d_{WSM}(C_x, \bar{C}_x) = \sum_{k=1}^{36} W(k)\,(S_x(k) - \bar{S}_{\hat{x}}(k))^2 \qquad (17)$$

Where, W(k) = weights for band.
By averaging the WSS value obtained across all frames a mean WSS value can be computed.

6 RESULTS AND DISCUSSIONS

Comparison among different wiener filtering algorithms is performed by implementing them in MATLAB. Noisy speech samples for stationary and non-stationary noises for both male and female speakers are considered for calculating the values of objective speech quality measures for the purpose of comparison. Noisy speech samples are taken from NOIZEUS [12] noisy speech corpus. They used noise from Aurora dataset [13]. Stationary noise is generated by MATLAB function "awgn" at 0dB, 5dB and 10 dB SNR levels. The male speech sample is "A good book informs of what we ought to know" and the female speech sample is "Let us all join as we sing the last chorus". The speech samples are narrow-band, with 8KHz frequency and duration of 2-3 seconds. Speech samples are saved in .wav (16bit PCM, mono) format. The experimental results for all 8 noises (airport, babble, car, exhibition, train, restaurant, street, station) for these two speech samples is based on objective measures which are, Cepstrum distance (CEP), Log Likelihood Ratio (LLR), Segmental SNR calculated in frequency domain (Seg SNR), Weighted Spectral Slope (WSS). By plotting spectrogram and time domain graphs for clean, noisy and filtered speech signal as shown in Fig. 1, Fig. 2 and Fig. 3 results are evaluated.

From Table 1 to Table 3 experimental results are shown for female speaker for non-stationary noises at 0 dB, 5 dB and 10 dB SNR. It can be said by results of Table 1 and plots for objective measures for female speaker 0 dB SNR noise values that HRNR Approach performs better and comparable in almost all cases in terms of cepstrum distance. For all other parameters Implicit Wiener filter is superior to all other algorithms.Table 2 reflects that HRNR approach performs better in terms of cepstrum distance for car, restaurant noise and in terms of LLR for station noise. For station and street noise in case of female speaker HRNR approach performs better or comparable for 0 dB and 5dB. Implicit wiener filter approach performs better for 10 dB SNR for female speaker speech samples. Also, spectrogram and time domain plots for different noise scenarios in each case were evaluated which confirms increased quality and intelligibility.

From Table 4 to Table 6 results for different noise samples for male speakers at 0dB, 5dBand 10 dB SNR values are shown. We can say for male speaker at 0dB SNR level HRNR and TSNR approach are not giving faithful values of LLR but they are performing better or comparable in terms of cepstrum distance. For male speaker voice at 5 dB and 10 dB cepstrum distance have better values for both street and exhibition noise. Tabel 7 and 8 shows results for stationary noise for both male and female speakers at 0dB, 5dB and 10dB noise level. The scenario of better values for cepstrum distance in case of non-stationary noise by HRNR and TSNR algorithm is not visible in case of stationary noise. But WSS values for stationary noise for male speaker at 0dB noise level and female speakers for 10 dB noise level are better or comparable. For street noise male speaker speech samples, HRNR algorithm gives better performance in terms of Cepstrum distance. Similarly for airport noise TSNR algorithm gives better LLR values for 5 dB and 10 dB noise levels. From above tables and graphs. showing comparative results in case of stationary noise we can also conclude that implicit wiener algorithm performs well in non- stationary noise condition than stationary noise conditions. Superiority of implicit wiener filter is not that much when stationary noise conditions are of concern and values for objective measures don't show much variation. We can also derive from experimental results that for many noise conditions harmonic regeneration step can only improve segmental snr values and WSS values. But cepstrum distance and LLR doesn't seem to improve much in many cases.

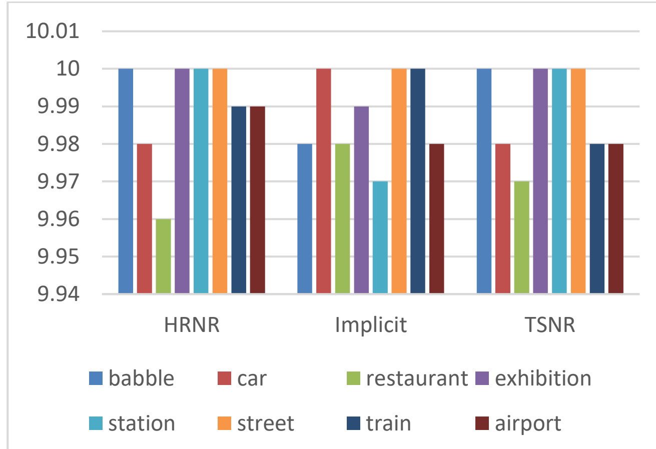

Fig. 1 CEP for female voice 0 dB for different algorithms

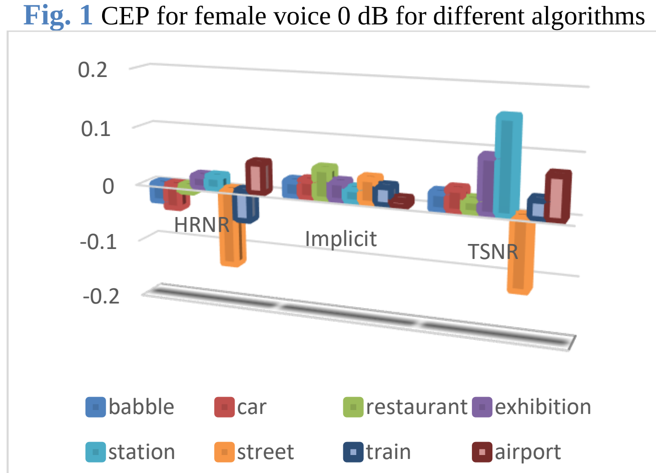

Fig. 2 LLR for female voice at 0 dB for different algorithms

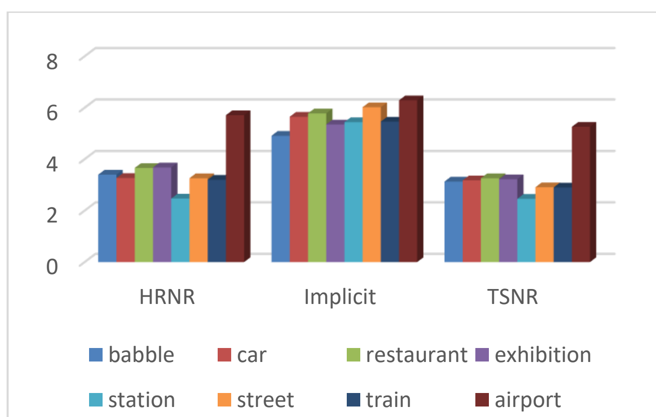

Fig. 3 Segmental SNR for female voice at 0 dB for different algorithms

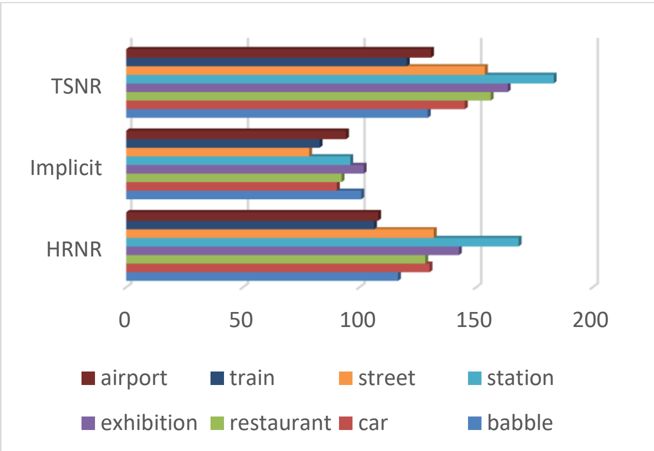

Fig. 4 WSS for female voice at 0 dB for different algorithms

6.1 Comparison among different Wiener Filter approaches for different noise levels

Comparison among different Wiener Filter approaches for different noise levels are given in Fig. 5, Fig. 6, Fig.7, Fig. 8, Fig. 9, and Fig. 10.

Fig. 5 CEP for different noise levels for female voice for street noise

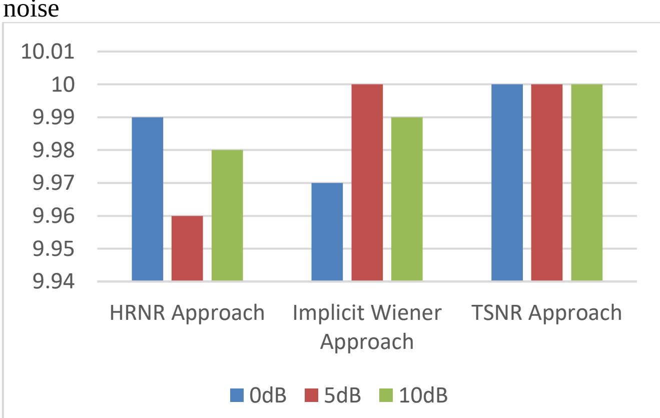

Fig. 6 CEP for different noise levels for male voice for street noise

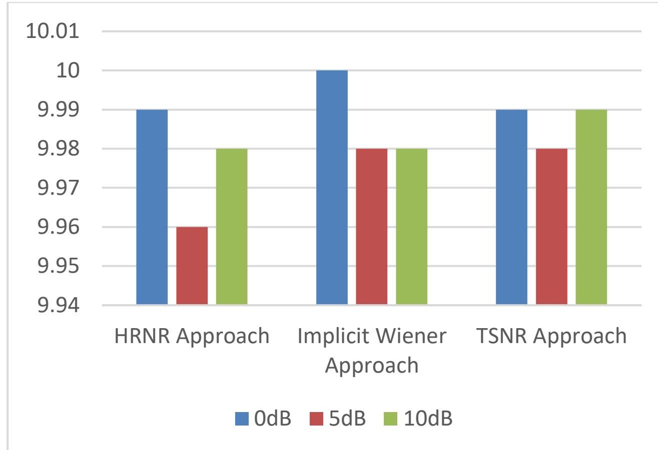

Fig. 7 CEP for different noise levels for female voice for restaurant noise

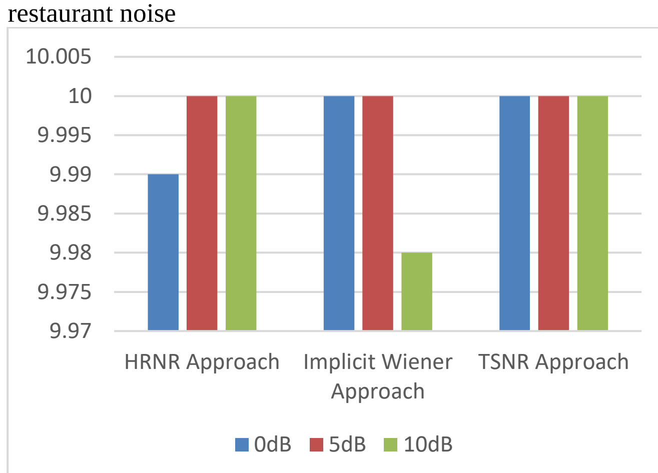

Fig. 8 CEP for different noise levels for male voice for restaurant noise

Fig. 9 CEP for different noise levels for female voice for car noise

Above graphs show comparison for different noise levels for street, restaurant and car noise for cepstrum distance measure.

Fig. 10 CEP for different noise levels for male voice for car

For street noise scenario for female voice a-priori based approaches perform better for 0dB and 10dB noise levels and for male voice they perform better at all noise levels. For restaurant noise scenario a -priori based approaches perform better at all noise levels for female voice and for male voice only 0dB noise level values are better. For car noise scenario a-priori based approaches perform better at 5dB and 10 dB noise levels for both male and female voice. HRNR, TSNR, Implicit Wiener, Approach results for babble noise of female speaker at 5 dB noise are shown in Fig. 11, Fig.12, and Fig.13 respectively. Table 7 and Table 8 shows Experimental results for Female, and male speaker for AWGN noisy speech respectively.

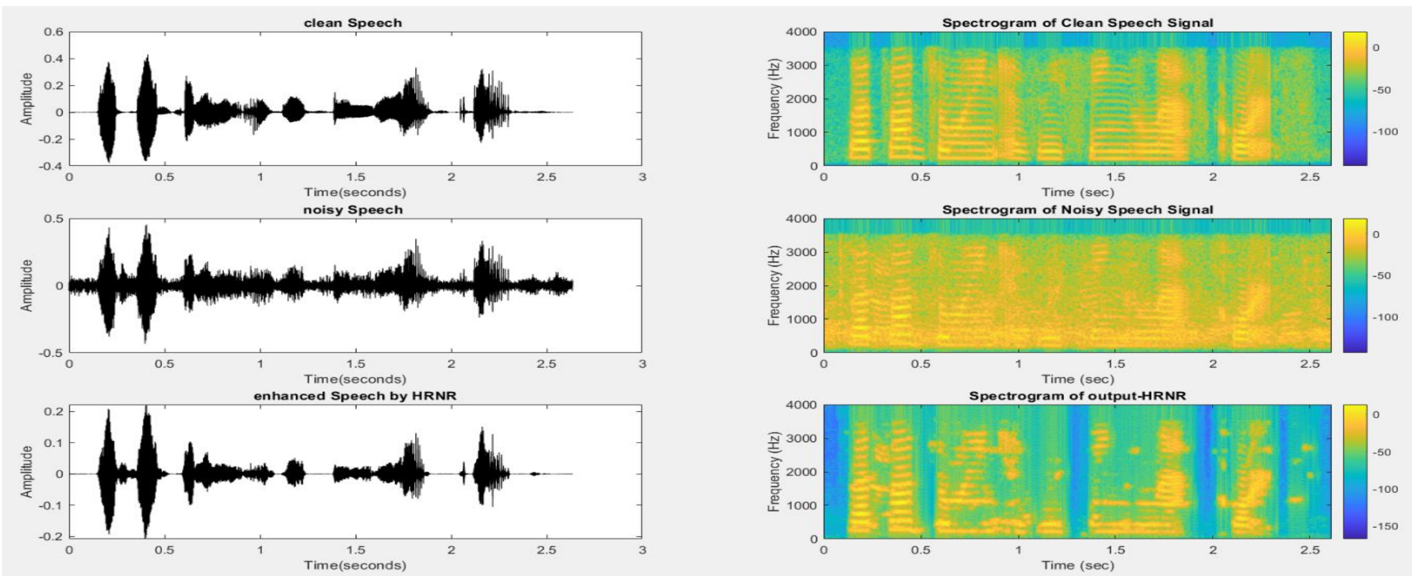

Fig. 11 HRNR Approach results for babble noise of female speaker at 5dB noise

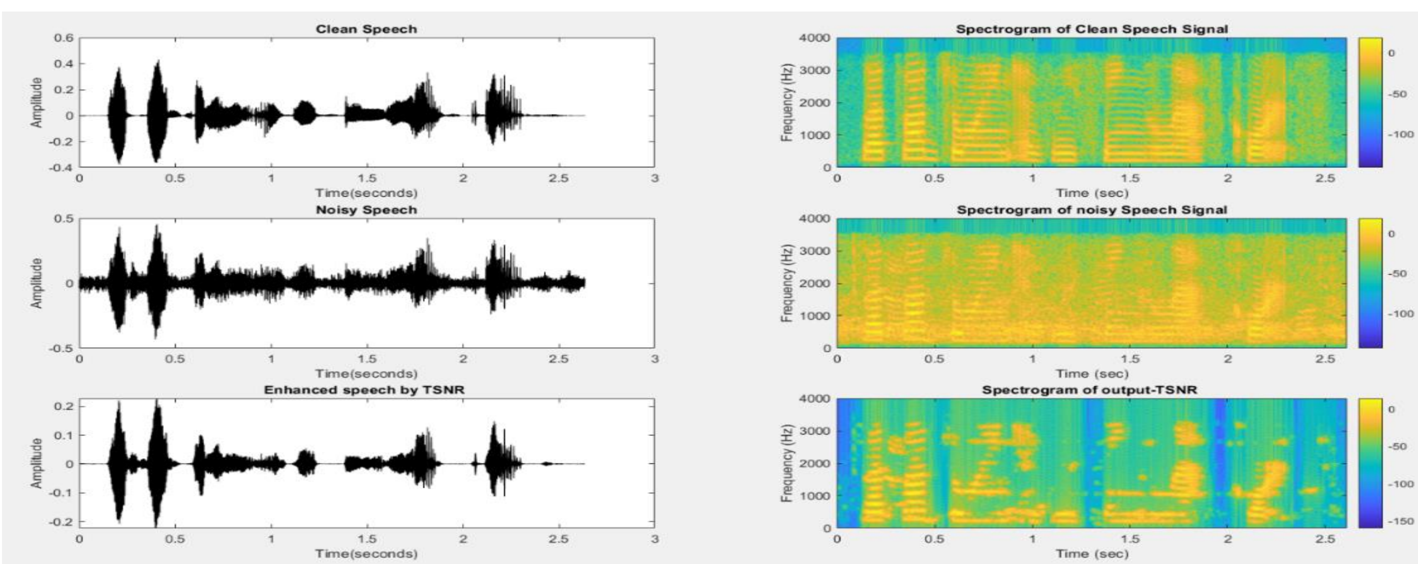

Fig. 12. TSNR Approach results for babble noise of female speaker at 5 dB noise

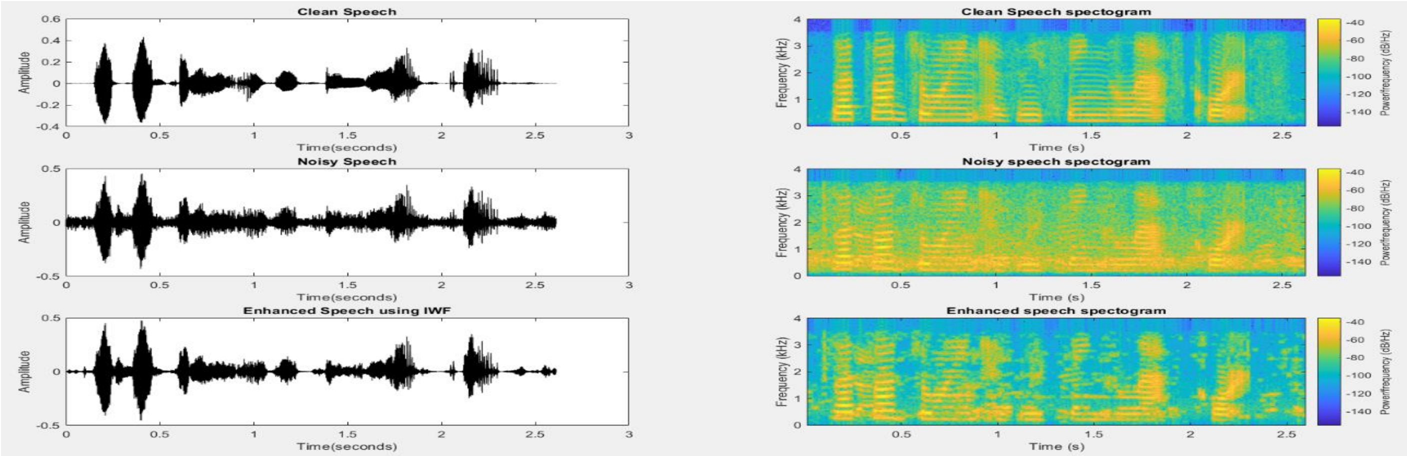

Fig.13 Implicit Wiener Approach results for babble noise of female speaker at 5 dB noise

Table 1. Experimental results for Female speaker noisy speech with SNR 0 dB noise

	HRNR Approach				**Implicit Wiener Approach**				**TSNR Approach**			
Type of Noise	***CEP***	***LLR***	***Seg SNR***	***WSS***	***CEP***	***LLR***	***Seg SNR***	***WSS***	***CEP***	***LLR***	***Seg SNR***	***WSS***
babble	10	-0.03	3.4	116.53	9.98	.025	4.91	100.78	10	0.027	3.13	129.3
Car	10	-0.04	3.27	130.06	10	.028	5.65	90.31	10	0.035	3.18	145.2
Restaurant	9.99	-0.01	3.66	128.23	10	.05	5.78	92.32	9.99	0.02	3.26	156.3
exhibition	9.97	0.02	3.68	142.67	10	.03	5.35	101.83	10	0.09	3.22	163.6
Station	10	0.021	2.47	168.24	10	.021	5.44	96.60	10	0.154	2.46	183.3
Street	9.97	-0.13	3.26	131.87	10	.04	6.02	78.47	9.98	-0.12	2.91	153.8
train	9.99	-0.05	3.20	106.24	9.96	.029	5.46	82.91	9.99	0.029	2.90	120.3
airport	9.97	0.05	5.71	108.01	10	.01	6.29	94.24	10	0.07	5.26	130.7

Table 2. Experimental results for Female speaker noisy speech with SNR 5 dB noise

	HRNR Approach				**Implicit Wiener Approach**				**TSNR Approach**			
Type of Noise	***CEP***	***LLR***	***Seg SNR***	***WSS***	***CEP***	***LLR***	***Seg SNR***	***WSS***	***CEP***	***LLR***	***Seg SNR***	***WSS***
babble	10	-0.04	5.32	91.91	9.98	0.04	6.82	75.04	10	0.02	4.44	118.36
Car	9.98	-0.06	5.29	73.03	10	0.02	6.74	65.67	9.98	-0.01	4.42	93.29
Restaurant	9.96	0.07	7.81	78.86	9.98	0.04	8.17	68.72	9.97	0.10	7.34	95.35
exhibition	10	0.002	5.61	84.20	9.99	.037	7.20	77.75	10	0.05	4.89	109.82
Station	10	0.008	3.92	117.9	9.97	.017	6.73	69.25	10	0.11	3.45	123.16
Street	10	-0.02	6.18	102.6	10	.018	7.31	69.87	10	-0.01	5.81	140.45
train	9.99	0.019	6.60	82.09	10	.014	8.24	68.50	9.98	0.052	6.04	93.17
airport	9.99	-0.03	4.91	113	9.98	0.02	7.00	70.05	9.98	0.03	4.21	134.92

Table 3. Experimental results for Female speaker noisy speech with SNR 10 dB noise

	HRNR Approach				**Implicit Wiener Approach**				**TSNR Approach**			
Type of Noise	***CEP***	***LLR***	***Seg SNR***	***WSS***	***CEP***	***LLR***	***Seg SNR***	***WSS***	***CEP***	***LLR***	***Seg SNR***	***WSS***
babble	9.98	-0.08	8.64	69.03	9.98	0.008	10.5	53.91	10	0.025	7.61	85.44
Car	10	0.002	8.46	82.25	9.95	-0.01	10.65	55.77	9.93	0.039	7.68	105.8
Restaurant	9.98	0.01	7.86	68.98	9.98	0.03	11.24	59.08	9.99	0.03	6.58	92.01
exhibition	10	-0.04	7.55	63.96	9.99	0.009	10.62	49.51	10	-0.02	6.48	87.66
Station	9.99	0.02	7.69	67.28	9.96	0.004	10.43	50.06	9.99	0.47	7.07	85.10
Street	10	0.07	10.2	53.54	10	0.03	10.46	40.00	9.99	0.09	9.94	62.18
train	10	-0.03	9.64	57.24	9.93	0.012	10.52	46.21	10	-0.01	8.84	74.81
airport	10	0.04	7.49	72.88	9.97	0.005	10.86	48.40	9.98	0.062	7.18	89.93

Table 4. Experimental results for Male speaker noisy speech with SNR 0 dB noise

	HRNR Approach				Implicit Wiener Approach				TSNR Approach			
Type of Noise	*CEP*	*LLR*	*Seg SNR*	*WSS*	*CEP*	*LLR*	*Seg SNR*	*WSS*	*CEP*	*LLR*	*Seg SNR*	*WSS*
babble	10	-0.02	3.79	129.86	9.99	0.02	5.30	89.14	10	0.00	3.15	142.76
Car	10	-0.08	4.14	107.67	9.98	0.02	5.28	77.21	10	-0.01	3.44	131.52
Restaurant	9.99	-0.02	3.60	124.52	10	0.08	4.19	89.47	10	-0.00	3.35	153.31
exhibition	10	-0.09	3.12	103.66	10	0.03	4.95	81.15	10	-0.09	2.32	137.19
Station	9.99	-0.10	3.18	107.61	10	0.02	5.20	73.60	9.99	-0.01	2.62	126.27
Street	9.99	-0.07	3.21	115.09	9.97	.007	5.74	73.43	10	-0.02	2.96	142.26
train	10	-0.11	3.27	111.15	9.98	0.04	4.61	76.10	9.99	-0.05	2.28	138.89
airport	10	-0.01	4.47	115.21	9.98	0.02	6.02	82.85	9.99	-0.00	3.83	128.71

Table 5. Experimental results for Male speaker noisy speech with SNR 5 dB noise

	HRNR Approach				Implicit Wiener Approach				TSNR Approach			
Type of Noise	*CEP*	*LLR*	*Seg SNR*	*WSS*	*CEP*	*LLR*	*Seg SNR*	*WSS*	*CEP*	*LLR*	*Seg SNR*	*WSS*
babble	10	-0.01	4.80	113.40	10	0.02	6.59	78.22	10	-0.04	4.08	126.85
Car	10	-0.05	5.25	83.07	10	0.02	6.70	66.67	9.99	-0.03	4.43	97.63
Restaurant	10	-0.02	5.03	107.24	10	0.05	6.07	79.73	10	-0.01	4.19	124.34
exhibition	9.98	-0.04	4.63	120.24	10	0.03	6.72	78.24	10	0.02	4.02	134.31
Station	10	-0.07	5.28	88.85	10	0.00	7.04	67.85	10	-0.03	4.30	108.27
Street	9.96	0.05	4.72	117.76	10	0.02	6.32	75.43	10	0.08	4.21	131.30
train	10	-0.03	5.72	80.49	9.98	0.03	6.28	66.28	10	-0.01	4.74	91.90
airport	9.99	-0.00	5.40	112.97	9.99	0.00	7.83	82	9.99	0.005	5.24	131.24

Table 6. Experimental results for Male speaker noisy speech with SNR 10 dB noise

	HRNR Approach				Implicit Wiener Approach				TSNR Approach			
Type of Noise	*CEP*	*LLR*	*Seg SNR*	*WSS*	*CEP*	*LLR*	*Seg SNR*	*WSS*	*CEP*	*LLR*	*Seg SNR*	*WSS*
babble	10	0.010	5.32	99.52	9.98	0.02	8.29	67.99	10	0.016	4.66	103.9
Car	9.99	-0.04	7.39	69.65	10	0.01	9.28	57.74	10	-0.02	6.15	85.91
Restaurant	10	-0.02	7.91	99.25	9.98	0.02	9.07	73.17	10	-0.05	6.73	114.6
exhibition	10	-0.00	7.84	77.24	9.99	0.01	10.4	58.64	9.98	-0.04	6.81	87.77
Station	10	-0.04	6.65	91.82	9.99	0.01	8.89	57.76	10	-0.02	5.57	105.7
Street	9.98	-0.03	6.33	94.01	9.99	0.02	8.90	65.65	10	-0.00	5.01	104
train	10	-0.03	6.94	70.60	10	0.01	9.33	55.39	9.98	-0.02	5.92	76.94
airport	9.99	-0.00	5.90	112.97	10	0.00	10.6	62.13	9.97	0.02	8.08	97.23

Table 7. Experimental results for Female speaker for AWGN noisy speech

	HRNR Approach				Implicit Wiener Approach				TSNR Approach			
SNR Levels	*CEP*	*LLR*	*Seg SNR*	*WSS*	*CEP*	*LLR*	*Seg SNR*	*WSS*	*CEP*	*LLR*	*Seg SNR*	*WSS*
0dB	10	-0.08	3.17	123	10	-0.05	5.91	107	10	0.00	2.83	127
5dB	10	-0.06	5.74	92.1	10	-0.02	8.34	81.2	9.99	-0.01	5.12	100
10dB	10	-0.05	7.99	70.3	9.96	-0.03	10.3	71.8	9.99	-0.01	6.84	76.4

Table 8. Experimental results for Male speaker for AWGN noisy speech

	HRNR Approach				Implicit Wiener Approach				TSNR Approach			

SNR Levels	CEP	LLR	Seg SNR	WSS	CEP	LLR	Seg SNR	WSS	CEP	LLR	Seg SNR	WSS
0dB	10	-0.07	4.41	104	10	-0.01	4.65	107	10	-0.04	3.47	113
5dB	10	-0.07	4.16	112	10	0.00	4.27	106	10	-0.02	3.66	118
10dB	10	-0.03	7.58	81.1	9.99	0.02	9.12	75.8	10	-0.02	5.32	77.7

7 CONCLUSION

Wiener filter is the most promising candidate for speech filtering application as a linear estimator because of less complexity in calculations and easier implementation. During the course of years methods have been developed to implement wiener filtering problem. In previous work [17] Implicit Wiener filter algorithm is compared with spectral subtraction and its' variants. Moreover, comparison among different Wiener filter algorithms was not studied. In this paper we studied and compared algorithms for Wiener filter implementation for speech enhancement purpose. Majorly experimental results and discussion of this report is focused on approaches based on a- priori SNR estimation and an approach based on power spectral density estimation. Wiener filter gain can be obtained from these methods, application of which on noisy speech signal provides us with enhanced speech signal as output. Objective quality measures (Cepstrum distance, Log Likelihood ratio, Weighted Spectral slope and segmental SNR), subjective listening, time domain plots and spectrogram plots are used for performance evaluation purpose. Two step Noise Reduction (TSNR) is an improvement over conventional decision directed approach used for a- priori SNR estimation. In HRNR approach speech harmonics which were suppressed during filtering process are regenerated to improve speech quality. Implicit Wiener filter is a variation of frequency domain parametric wiener filter which uses first order recursive equation to estimate noise PSD. Experimental results shows that implicit wiener filter outperform other two algorithms but cepstral distance measure values for HRNR and TSNR algorithms are better or comparable as compared to Implicit Wiener filter. Various speech and emotion recognition algorithms and approaches are based on cepstral distance. Wiener filtering approaches based on a – priori SNR estimation can be used in this scenario.

Research is a continuous process. So, it is important to think about the scopes of further extension of the current work. In this report comparison among Wiener filter algorithms has been performed using experimental results. As an extension to this work performance comparison of Wiener filtering algorithms in combination with other filtering algorithms and in different transfer domain can be performed. Also, evaluation of these algorithms for suitability for different applications such as speech recognition, emotion recognition, speech coding etc can be done.

REFERENCES

1. Philips C. Loizou (2013). Speech Enhancement Theory and Practice, ISBN-13:978-1-4665-0422-6 (e-Book-PDF), 2nd edition CRC Press https://doi.org/10.1201/b14529
2. Komal R. Borisagar, Rohit M. Thakani, Bhavin S. Sedani, (2018). Speech Enhancement Techniques for Digital Hearing Aids ISBN 978-3-319-96821-6(eBook), Springer https://doi.org/10.1007/978-3-319-96821-6
3. Oday Kamil Hamid (2018). Frame Blocking and Windowing Speech Signal", Journal of Information, Communication, and Intelligence Systems (JICIS) Volume 4, Issue 5, ISSN:2413-6999
4. Das, N., Chakraborty, S., Chaki, J, Padhy, N. & Dey (2020). Fundamentals, Present and Future Perspective of Speech Enhancement", International Journal of Speech Technology, 24(4), 883-901 https://doi.org/10.1007/s10772-020-09674-2
5. Yihan Wang (2020). Research Progress in Speech Enhancement Technology, International Conference on Computer Vision, Image and Deep Learning https://doi.org/10.1109/cvidl51233.2020.00-97
6. Mr. C. Ramesh Kumar (2022). Dr. M.P. Chitra, Implementation of Modified Wiener Filtering in Frequency Domain in Speech Enhancement, International Journal of Advanced Computer Science and Applications, Vol. 13, No. 2 https://doi.org/10.14569/ijacsa.2022.0130251
7. Hu, Y., Loizou, P.C (2007). Evaluation of Objective Quality measures of Speech Enhancement, IEEE Transactions on Audio, Speech and Language Processing, 16(1), 229-238: https://doi.org/10.21437/interspeech.2006-84 Konrad Kowalczyk (2019). Multichannel Wiener Filter with Early Reflection Raking for Automatic Speech Recognition in Presence of Reverberation", IEEE Signal Processing Algorithms, Architectures, Arrangements and Applications: https://doi.org/10.23919/spa.2019.8936706
8. Jacob Benesty, Jingdong Chen, Yiteng Huang Study of the widely Linear Wiener Filter for Noise Reduction, International Conference on Acoustic, Speech, and Signal Processing: https://doi.org/10.1109/icassp.2010.5496033
9. V. Sailaja, P. Sunitha, B. Vasantha Lakshmi, V. Prasanth (2019). Wiener Filtering in Frequency Domain to Enhance Speech Corrupted by Colored noise, International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8, Issue-2S11
10. Scott C. Douglas, Damilo P. Mandic (2017). Single Channel Wiener Filtering of Deterministic Signals in Stochastic Noise Using the Panorma", International Conference on Acoustic, Speech, and Signal Processinghttps://doi.org/10.1109/icassp.2017.7952944
11. Dr. China Venkateshwarlu Sonagiri, K. Satya Prakash, D. SubbaRami Reddy (2011). Improve Speech Enhancement using Wiener Filtering, Global Journal of Computer Science and Technology, volume 11 Issue7Version1.0 https://computerresearch.org/index.php/computer/article/view/725
12. Naveen Upadhyay, Rahul Kumar Jaiswal (2015). Single Channel Speech Enhancement: using Wiener Filtering with Recursive Noise Estimation", 7th International conference on Intelligent Human Computer Interaction, IHCI https://doi.org/10.1016/j.procs.2016.04.061
13. Cyril Plapous, Claude Marro, Pascal Scalart (2006). Improved Signal to Noise Ratio Estimation for Speech Enhancement", IEEE Transactions on Audio, Speech and Language Processing, Vol 14, No. 6 https://doi.org/10.1109/tasl.2006.872621
14. K. Prajna, C.K. Mukopadhyay (2019). Efficient Harmonic regeneration noise reduction- based Wiener Filter for acoustic emission signal detection", IET Electronics letters, volume 55, Issue 22, p. 1163-1165 **https://doi.org/10.1049/el.2019.2196**
15. Jaiswal, R, Romero, D. (2021). Implicit Wiener Filtering for Speech Enhancement in non-stationary noise", 11th International Conference on Information Science and Technology (ICIST), IEEE (PP. 39-47) https://doi.org/10.1109/icist52614.2021.9440639
16. Hirsch, H.G., Pearce, D. (2000). The Aurora experimental framework for the performance evaluation of speech recognition systems under noisy conditions, In Proceedings of the automatic speech recognition: Challenges for the new millenium, ISCA Tutorial and Research Workshop (ITRW) https://doi.org/10.21437/icslp.2000-743
17. Rahul Kumar Jaiswal, Sreenivasa Reddy Yeduri, Linga Reddy Cenkeramaddi (2022). Single-channel speech enhancement using implicit Wiener filter for high-quality speech communication", International Journal of Speech Technology. https://doi.org/10.1007/s10772-022-09987-4

18. Siddhart Chhetri, Manjusha Sanjeev Joshi, Chaitanya Vijay Kumar Mahamuni, Repana Naga Sangeetha, Tushar Roy (2023). Speech Enhancement: A Survey of Approaches and Applications", International Conference on Edge Computing and Applications (ICECAA)https://doi.org/10.1109/icecaa58104.2023.10212180
19. Bin Wen Fan, Huan Yu Song, Ming Liu, Yong Jun Wang (2015). The Improvement and Realization of Speech Enhancement Algorithm Based on Wiener Filtering ,8th International Congress on Image and Signal Processing
20. Nayan Modhave, Yepuganti Karuna (2016). Sourabh Tonde, "Design of Multichannel Wiener Filter for Speech Enhancement in Hearing Aids and Noise Reduction Techniques, Online International Conference on Green Engineering and Technologies (IC-GET) DOI: https://doi.org/10.1109/get.2016.7916626
21. Horia- Nicolai L. Teodoreacu,(2018). Wiener filters for the reconditioning of old speech recordings under uncertainty, ISBN 978-80-0722-4, University of West Bohemia
22. P. Rajesh, K. Umamaheshwari, P. Vinodh Babu, S. Srinivasa Rao (2015). Application of Adaptive filters in digital hearing aids for Cancellation of Noise", IEEE International Conference on Computer, Communication and Signal Processing, https://doi.org/10.1109/iccsp.2015.7322540
23. V. Srinivas Rao, Umesh Ghanekar (2020). Speech Intelligibility Enhancement: A Hybrid Wiener Approach", International Journal of Speech Technology, 23(3), 517-525
24. Kristian Timm Andersen (2018). Robust Speech- Distortion Weighted Interframe Wiener Filters for Single- Channel Noise Reduction", IEEE/ACM Transactions on Audio Speech and Language Processing, Vol 26, No 1 https://doi.org/10.1109/taslp.2017.2761699
25. Wageesha Manamperi, Prasanga N. Samarasinghe, Thushara D. Abhayapala, Jihui Zhang (2022). GMM Based Multi- Stage Wiener Filtering for Low SNR Speech Enhancement", International Workshop on Acoustic Signal Enhancement (IWAENC), 2022 Conference on Telecommunications, Optics and Computer Science (TOCS) https://doi.org/10.1109/iwaenc53105.2022.9914707
26. Qiujuan Tong, Huan Han, Lu Huang, Jun Wang (2021). Musical Noise Reduction Methos Based on MMSE Combined with Wiener Filter, International Conference on Intelligent System and Knowledge Engineering,ISKE https://doi.org/10.1109/iske54062.2021.9755387
27. Ouardia Abedelli, Fatiha Merazka, (2020). Denoising Speech Signals using Decision Directed Approach, International Journal of Informatics and Applied Mathematics e-ISSN: 2667-6990, Vol 3, No.1, 70-83
28. Suzumi Kanehara, Hiroshi Saruwatari, Ryoichi Miyazaki, Kiyohiro Shikano, Kazunobu Kondo (2012). Comparative Study on Various Noise Reduction Methods with Decision- Directed a Priori SNR Estimator via Higher – Order Statics CREST, Japan.
29. Chiea, R. A., Costa, M.H., Barrualt, (2019). New Insights on the Optimality of Parameterized Wiener Filters for Speech Enhancement Applications, Speech Communications, 109, 46-54 https://doi.org/10.1016/j.specom.2019.03.005
30. Ali, Y. S. E. Parsa, V Doyle, Berkane (2020). Low – complexity disordered speech quality estimation", International Journal of Speech Technology, 23(3), 585-594 https://doi.org/10.1007/s10772-020-09688-w
31. Noizeus Speech Corpus https://ecs.utdallas.edu/loizou/speech/noizeus/

AUTHORS

Saloni Dwivedi received her BTech degree from R.N. Modi Engineering College, Rajasthan, India in 2016. She is currently pursuing MTech degree at the department of Electronics Engineering, Rajasthan Technical University, Kota, Rajasthan.

Email: salonidwivedi31gmail.com

Ajay Khunteta received his BE degree in Electronics and Communication Engineering from MBM Engineering College, Jodhpur. Hecompleted his MTech in Communication and Radar Engineering from IIT Delhi in 2002. He received his PhD from IIT Roorkhee in Electronics and Communication Engineering in 2016. His research area are signal and Image Processing, Communication systems and Wireless Communication Networks.

Email: akhunteta@rtu.ac.in