Hybrid Filtering : A Case Study

Rohini R. Mergu
Department of Electronics and Telecommunication Engineering
WIT Solapur

Dr. Shantanu K. Dixit
Professor
Organization Name, City, State, Country

Abstract- **The paper presents speech enhancement scheme for suppression of different background noises. The objective of speech enhancement is to improve the quality of the processed speech. This paper also investigate the use of two different transforms for speech enhancement. Speech enhancement using wiener filtering approach is presented using transforms DFT and DCT. The type of transform shows different amount of quality improvement using same type of filtering i.e, Hybrid Wiener Filtering**

Keywords – Speech enhancement, Hybrid Filter, DFTF, DCTF, DWTF, wiener filter, transform

I. INTRODUCTION

Speech enhancement aims to improve speech quality by using various algorithms. The objective of enhancement is improvement in intelligibility and/or overall perceptual quality of degraded speech signal using audio signal processing techniques. Enhancing of speech degraded by noise, or noise reduction, is the most important field of speech enhancement, and used for many applications such as mobile phones, VoIP, teleconferencing systems, speech recognition, and hearing aids. The algorithms of speech enhancement for noise reduction can be categorized into three fundamental classes: filtering techniques, spectral restoration, and model-based methods. This paper discusses the filtering based speech enhancement scheme.

Traditionally, many speech algorithms operate in the frequency domain or some other transform domains. The Speech enhancement system aims to improve the quality of speech for various different applications. Performance gain in speech enhancement using basic transform domain has reached saturation point. Many researchers have tried to exploit the correlation among adjacent time frames. Boll in [1] proposes use of time averaging or taking minimum of the coefficients of the present and it's neighboring frames. A weighted average of several frames is also adapted in [2]. Neighboring frames are also exploited in [3] which uses one dimensional interpolation and in [4] utilizes one dimensional interpolation and two dimensional smoothing scheme. Hence the correlation existing between frames can be utilized to achieve further gain in performance. One to exploit this is use two dimensional transform. Hence multiple time frames are arranged into a block and windowed before transformation. Two dimensional transform enables the use of two dimensional filtering techniques. The arrangement of speech data in 2D form and its appropriate windowing such that 2D transform can be applied. Various filtering techniques applicable in 2D Fourier transform domain. The better scheme is 2D wiener filter. The best scheme is combination of 1D and 2D wiener filter.

Use of the two-dimensional (2-D) Fourier transform for speech enhancement presented [5]. Also, include magnitude spectral subtraction, 2-D Wiener filtering as well as a hybrid filter which effectively combines the one-dimensional (1-D) Wiener filter with the 2-D Wiener filter. This is apparent in many recent works which view speech as a 2D time–frequency signal, especially in the form of a spectrogram. The transform domain also plays vital role in speech enhancement. The effect of transform domain on quality of speech recovered using wiener filtering is shown in [6,7].Enhancement in speech quality can be determined using different objective measures such as signal to noise ratio (SNR), segmental signal to noise ratio (segSNR), frequency weigthed segmental SNR (fwsegSNR) and Perceptual Evaluation of Speech Quality (PESQ) as suggested in [8].

II. METHODOLOGY

A. Hybrid Filtering -

Hybrid filtering means combination of one dimensional (1D) filtering and two dimensional (2D) filtering gives better speech quality than individual methods. Here wiener filtering is used in 1D domain and in 2D domain. In this paper an attempt is made to see the effect of change of transform on the quality of enhanced speech. The detailed block diagram of the proposed work is shown in fig. 1. Three types of transforms effect is studied with 4 different types of noises with different SNRs on the single speech sentence uttered by male speaker. The transforms investigated are Discrete Fourier Transform (DFT), Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT).

The noisy speech is formed into frames of 256 and applied 1D wiener filtering and overlapped and added. The same noisy speech is applied with framing, blocking for 2D speech block and then applied 2D wiener filtering. After overlap and add of this 2D filtered speech it is compared with 1D filtered speech and minimum magnitude of two is considered in the enhanced speech. With this better speech enhancement is achieved compared to individual filtering techniques. This is called Hybrid filtering which is combination of 1D and 2D filtering. This Hybrid filtering is also applied using three transforms DFT filtered (DFTF), DCT filtered (DCTF), DWT filtered (DWTF).

The Noisy speech is the speech signal corrupted by different types of background noises as fan noise, car noise, aeroplane noise, train noise. The additive noise model is described by the following equation,

$$y(t) = x(t) + n(t) \qquad (1)$$

Where,
y(t) = observed noisy speech,
x(t) = clean speech
n(t) = additive background noise.

The noisy speech is divided into overlapping frames of length of 256 samples in each frame and 75% overlapping is used. The n^{th} frame, can be represented by a column vector described by the following equation

$$fL = [y(64L)y(64L+1)y(64L+2)\dots y(64L+255)]T \qquad (2)$$

This signal is windowed using Hamming window. Then the transform can be applied onto the speech block.
A speech block can be obtained by arranging a number of frames together to form a matrix .Suitable numbers of frames are found experimentally to be 8,16and32.In this paper, the number of frames used is 16 throughout. Similarly each block overlaps its neighboring block by50%. Then the speech block can be represented as

$$bn = [f_{8n} \;\; f_{8n+1} \;\; f_{8n+2} \;\; \dots\dots f_{8n+15}] \qquad (3)$$

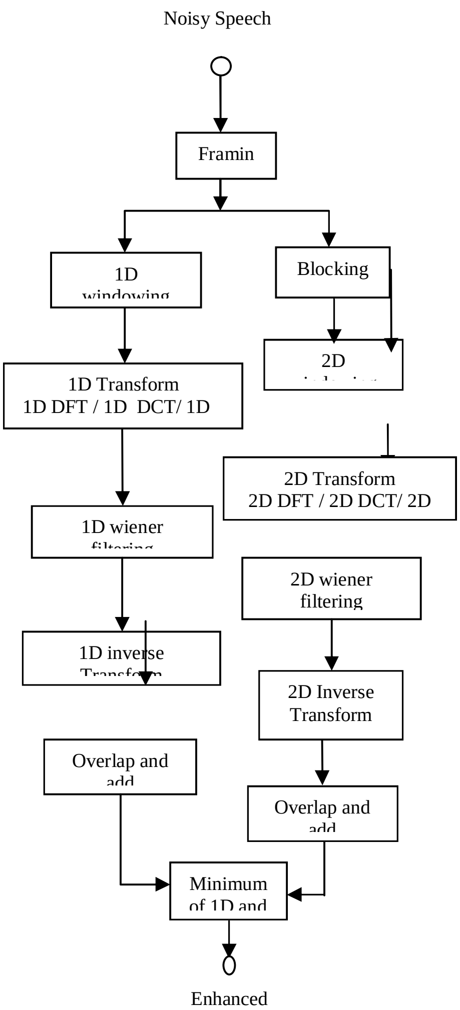

Fig. 1 : Block diagram of Hybrid speech Enhancement System

B. Wiener Filtering –

The wiener filter produces the highest noise attenuation.

$$\hat{x}(t) = g(t) * y(t) \qquad (4)$$

Where,

$\hat{x}(t)$ = Estimated signal after Filtering

$g(t)$ = impulse response of Wiener Filter

And the difference between clean speech & estimated signal indicates amount of error. This is given as,

$$e(t) = x(t) - \hat{x}(t) \qquad (5)$$

Where,
$e(t)$ = error signal

The prime focus in this paper is to reduce error in turn background noise in speech signal. This is done by wiener filtering i.e, 1D wiener and 2D wiener filtering and combination of both i.e, hybrid wiener filtering. Three transforms are used Discrete Fourier Transform filtered (DFTF) and Discrete Cosine Transform filtered wiener (DCTF) and Discrete Wavelet Transform Filtered (DWTF). The results for speech sp01 corrupted by train noise are shown in figs. 2 to 5. The speech sample sp01 is obtained from noiseus database.

C. *Transform*

It is easier to remove noise from the noisy speech in frequency domain. Hence, covert time domain speech signal to frequency domain using transform. Generally used and the most popular transform used is Discrete Fourier Transform (DFT). But this paper experiments the use of Discrete Cosine Transform (DCT) , Discrete Wavelet Transform (DWT) shows the effect of change of transform on the speech quality and intelligibility.

III. RESULTS AND DISCUSSIONS

The noise reduction for different speech sentences from noiseus database in different noise conditions with different SNRs for three transforms are investigated. The experiment is carried out for airport, exhibition, restaurant, station, street, train noise. But here the results are presented only for sp04 corrupted by airport noise. Also Table I shows results obtained by airport and station noise.

From the Figs. 2 to 7 shows clean speech and enhanced speech along with corresponding spectrograms of 1D, 2D and hybrid filtered speech for DFTF, DCTF and DWTF. These figures shows that hybrid filtering gives good quality of speech than 1D filtering and 2D filtering alone. Figures 8 to 12 of SNR, segSNR, fwsegSNR and PESQ and MOS shows improvement in quality of speech by DFTF than DCTF and DWTF. The plots indicate that speech quality improvement is more using DFTF than DCTF and DWTF. Also from the results shown in Table 1 we can observe that noise reduction is good in case of DFTF than DCTF and DWTF the values for SNR, segSNR, fwsegSNR and PESQ and MOS are shown in Table I for two noise conditions.

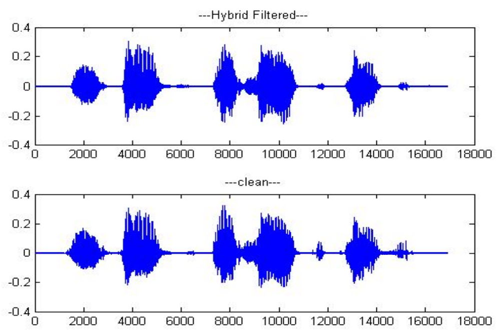

Fig.2 : DFTF Speech sp04, SNR10

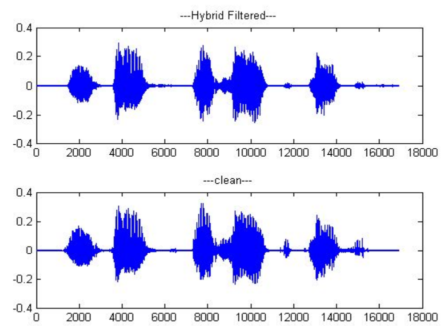

Fig 3 : DCTF Speech sp04, SNR10

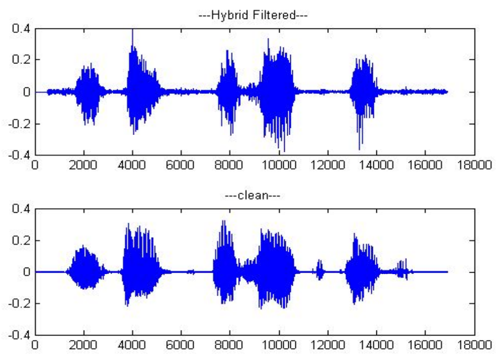

Fig.4 : DWTF Speech sp04, SNR10

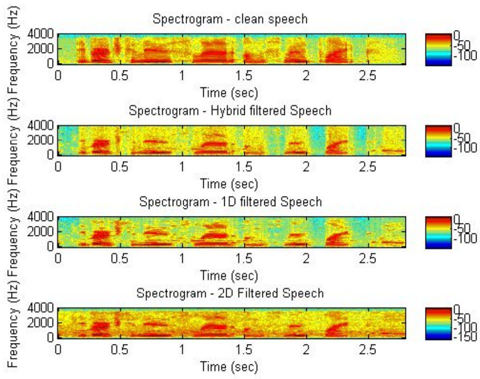

Fig.5 : DFTF spectrogram Speech sp04, SNR10

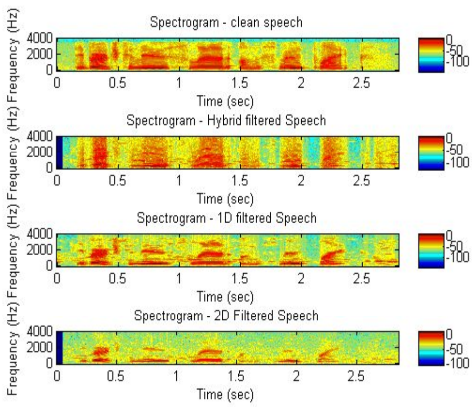

Fig.6 : DCTF Spectrogram Speech sp04, SNR10

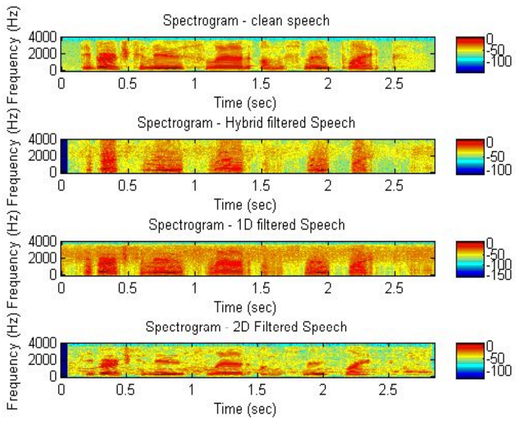

Fig. 7 : DWTF Spectrogram Speech sp04, SNR10

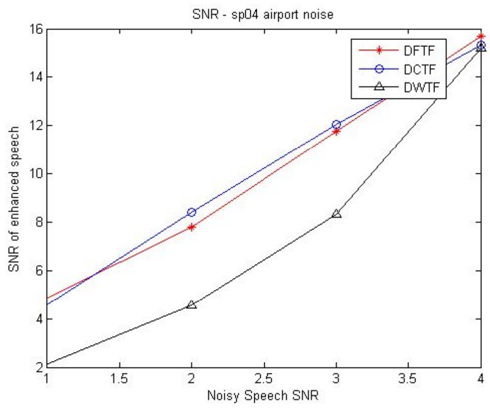

Fig.8 : SNR

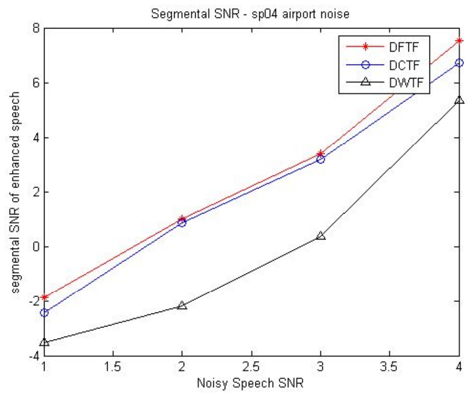

Fig.9 : Segmental SNR

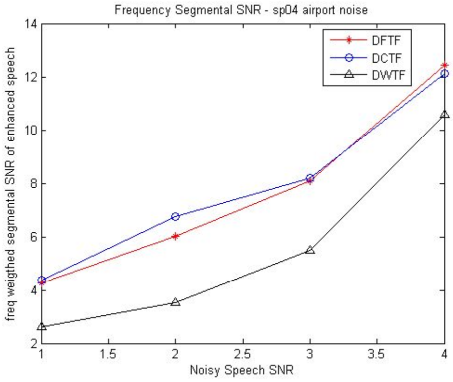

Fig.10 : Frequency Weighted Segmental SNR

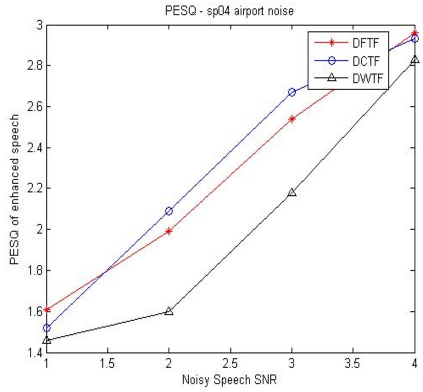

Fig.11 : PESQ

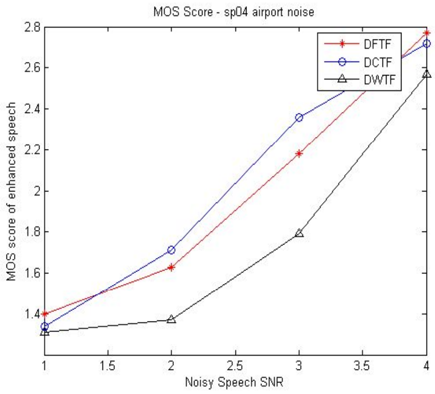

Fig.12 : MOS Score

Table -1 : Objective measures for sp04 speech enhancement using Hybrid Filtering

Type of Noise	Noisy speech SNR	Transform	SNR	segSNR	fwsegSNR	PESQ	MOS
Sp04							
airport	**0 dB**	**DFTF**	4.87	-1.89	4.26	1.61	1.40
		DCTF	4.59	-2.44	4.37	1.52	1.34
		DWTF	2.16	-3.54	2.60	1.46	1.31
	5 dB	**DFTF**	7.79	1.00	6.02	1.99	1.63
		DCTF	8.42	0.87	6.75	2.09	1.71
		DWTF	4.55	-2.21	3.52	1.60	1.37
	10 dB	**DFTF**	11.76	3.38	8.10	2.54	2.18
		DCTF	12.07	3.16	8.21	2.67	2.36
		DWTF	8.31	0.36	5.48	2.18	1.79
	15 dB	**DFTF**	15.71	7.52	12.44	2.96	2.77
		DCTF	15.32	6.72	12.13	2.93	2.72
		DWTF	15.20	5.34	10.58	2.83	2.57
Station	**0 dB**	**DFTF**	5.82	-0.64	5.01	1.92	1.57
		DCTF	5.57	-1.35	5.47	1.89	1.55
		DWTF	2.15	-3.60	2.92	1.56	1.35
	5 dB	**DFTF**	6.57	0.43	5.97	2.03	1.66
		DCTF	5.99	-0.13	5.67	1.92	1.57
		DWTF	2.42	-2.63	2.23	1.52	1.33
	10 dB	**DFTF**	11.35	3.53	7.74	2.49	2.13
		DCTF	11.23	3.03	7.88	2.63	2.29
		DWTF	8.09	0.48	5.41	2.21	1.82
	15 dB	**DFTF**	15.46	6.01	10.98	2.49	2.13
		DCTF	14.68	5.36	11.09	2.46	2.09
		DWTF	14.95	4.76	9.86	2.56	2.21

IV.CONCLUSION

It can be concluded from fig. 2 to 7 that hybrid filtering gives best results than 1D filtering and 2D filtering. The speech quality obtained is best of hybrid filtered speech. Along with filtering transform domain plays important role in improvement in speech quality. By observation of the values of SNR, segSNR, fwsegSNR and PESQ and MOS shown in Table 1 and from the figs. 8 to 12 we can conclude that DFTF gives better enhanced speech quality than DCTF and DWTF. Thus the best quality of speech can be obtained by hybrid wiener filtering using DFT transform. Also from MOS Score we can conclude that hybrid wiener filtering gives good speech quality of speech but speech intelligibility is poor.

REFERENCES

[1] Boll S.," Suppression of acoustic noise in speech using spectral subtraction", ***volume*** 27 ,Issue 2 ,***pp.*** 113 - 120 , IEEE trans. On Acoustics, Speech and Signal Processing, 1979

[2] Y.M Cheng and D. O'Shaughnessy,"Speech Enhancement Based Conceptually on Auditory Evidence", *IEEE trans. Signal Processing*, vol. 39, pp. 1943-1954, 1991

[3] I.Y. Soon and S.N. Koh, "Low Distortion Speech enhancement", *IEEE trans. Vision Image and Signal Processing*,vol 147,pp. 247-253 , 2000

[4] Jensen J. , Hansen J.H.L., "Speech enhancement using a constrained iterative sinusoidal model**",** *IEEE trans. On speech and Audio Processing,* Vol 9, Issue 7, pp. 731 – 740, 2001

[5] I.Y. Soon and S.N. Koh, "Speech enhancement using 2-D Fourier transform", *IEEE trans. On speech and Audio Processing,* v**ol** 11 , Issue: 6,pp. 717 - 724 , 2003

[6] Mrs. R. R. Mergu , Dr. S.K. Dixit , "Investigation of Transform dependency in Speech Enhancement", *Int Journal of Recent Technology and Engineering,* vol-2, Issue-3,pp.158-161, July 2013

[7] Mrs. R. R. Mergu , Dr. S.K. Dixit , "Hybrid Filtering : An approach for Speech Enhancement", *Int Journal of Advanced Technology in Engineeringand Science,* vol-2, Issue-7,pp.51-57, July 2014

[8] Yi Hu and Philipos C. Loizou,, "Evaluation of Objective Quality measures", *IEEE trans. On speech and Audio Processing,* vol 16, Issue 1, 2008