International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 4 Issue 7, July 2015

Intelligibility Evaluation of Noise Reduction Algorithms

Rahul Deshpande, Vanita Raj Tank, ArtiKhaparde

***Abstract*-The hearing process is based on time localization and frequency localization. In auditory signal processing this kind of time-frequency type transformation is done. The present paper deals with evaluation of intelligibility for noise reduction algorithms in speech signal processing and a comparative analysis is done with existing algorithms like Spectral subtraction, Wiener filter using an objective measure of Normalized Covariance Measure.**

***Keywords:*Normalized Covariance Metric, intelligibility, time-frequency localization.**

I. INTRODUCTION

To recognize the speech data correctly from the noisy input, noise reduction or model adaptation to the sudden noise is required. However, it is difficult to remove such noises because it is not known where the noise is overlapped on speech and the type of noise introduced in the signal. Innoise reduction background noise plays main role for the intelligibility of speech, if presence of background noise ismorethe intelligibility is reduced. The statistical parameters of noise are evaluated and updated from the slot where only noise is present.

Audio and speech signal processing techniques applies Fourier transform (FT).The resolution (time-frequency) of FT is fixed and has linearfrequency distribution.This introduces the problems such as the pitch harmonics, computational noise and sensitivity to background noise. For noise reduction the traditional algorithms such as Spectral Subtraction and Wiener filter algorithm have been used.

In this paper Section II describes the spectral subtraction algorithm. Section III describes the Wiener Filter algorithm. Section IV describes results. Section Vconcludes paper.

II. SPECTRAL SUBTRACTION ALGORITHM

The time-domain signals are reconstruction using some process; in which the magnitude spectrumof signal andphase of the noisy signalis taken at that instant, and then using Inverse Discrete Fourier Transform converted back to the time domain signal.The spectral subtraction is more complex in computation.However, considering the random varying nature of noise, spectral subtraction can give magnitude or power spectrum estimate in negative. These are non-negative variables,so any negative value is mappedto non-negative values [3, 8].

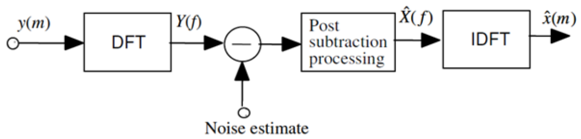

Fig 1: Illustration of Spectral subtraction

The noisy signal model in the time domain is given by

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

Where y(m) is the noisy signal, x(m) is original signal, n(m) is the additive noise and *m* is the discrete time index. Thesame model in the frequency domain is expressed as

$$S(f) = I(f) + N(f) \qquad (2)$$

Where S(f), I(f) and N(f) are the Fourier transforms of the noisy signal y(m), the signal x(m) and the noise n(m) respectively. *f is* the frequency variable.

The original signal x(m) is divided into N segments. Then windowing is applied to each segment, using a Hamming or a Hanning window and then Discrete Fourier Transform (DFT) is used to transform the signal to *N* spectral samples. The discontinuities at the endpoints of each segment are eliminated because of windowing. The windowed signal is given by (3).

$$\begin{aligned} y_w(m) &= w(m)y(m) = w(m)[x(m)+n(m)] \\ &= x_w(m)+n_w(m) \qquad (3) \end{aligned}$$

The windowing operation can be expressed in the frequency domain as below,

$$Y_w(f) = X_w(f) + N_w(f) \qquad (4)$$

$$X(f) = Y_w(f) - N_e(f) (5)$$

Where $N_e(f)$ is estimated noise. Equation (5) shows post subtraction process. In which estimated noise issubtracted form windowed signal. Last step is retrieving original signal back via Inverse Fourier Transform [3].

Manuscript received July, 2015.
Rahul Deshpande, *Electronics and Telecommunication, Maharashtra Institute of Technology, Pune, India.*
Vanita Tank*, Electronics and Telecommunication, Maharashtra Institute of Technology, Pune, India).*
ArtiKhaparde*, Electronics and Telecommunication, Maharashtra Institute of Technology, Pune, India.*

III. WIENER FILTER ALGORITHM

Another approach for noise reduction of noisy speech signal is Wiener filtering. The Wiener Filtering uses MMSE (Mean square error algorithm) which is based on the idea of minimizing the Mean square error between the input signal X (w) and estimated signal magnitude spectrum $\hat{X}$ (w) [3].

$$H(w) = \frac{X_x(w)}{X_x(w)+X_n(w)} (6)$$

Where $X_x(w)$ represent the estimated power spectrum of noise-free signal and $X_n(w)$ is background noise.Input to the wiener filter X (t) is corrupted by the additive noise N (t). The output $\hat{X}$ (t) is calculated using of Wiener filter function H (t).

$$\hat{X}(t) = H(t) * (X(t) + N(t))(7)$$

Where X (t) is the original signal.N (t) is the noise, $\hat{X}(t)$ is the estimated and H (t) is the impulse response of filter. The error function is defined as

$$E(t) = X(t+\alpha) - \hat{X}(t)(8)$$

α is the delay of the wiener filter. Using equation (8) the error is calculated. The drawback of the Wiener filter is the fixed frequency response at all frequencies and the requirement to estimate the power spectral density of the clean signal and noise prior to filtering. Consider the simplestway when the signal and noise are assumed statistically independent. The priori SNR in maximum likelihood (ML) approach iscalculated by subtracting unity from Noisy SNR or a posterior SNR. The decision-directed approach is calculated using the SNR of this MLby taking a weighted average and SNR of the previous frame's determined from the enhanced speech. Both, maximum likelihood ratio algorithm and decision–direct algorithm, assume that the noise power spectrum is known [3].

IV. RESULTS

NOIZEUS is a noisy speech database.The noisy database contains 30 IEEE sentences and mixedby different real-world noises,like station noise, airport noise, restaurant noise at different SNRs [10]. In this paper the spectrogram result are shown for Restaurant noise of 0dB with sampling frequency of 8000Hz[10].

A. Normalized Covariance Measure

Previous studies by researchers have emphasized on the preservation of temporal envelope cues for speech perception. It becomes essential for hearing impaired listeners as they have reduced ability to process the fine temporal structure and temporal cues. Hence to assess the performance of noise reduction techniques this factor need to be determined utilizing an objective measure. One such measure namely normalized covariance metric (NCM) has been proposed which is based on temporal cues.

The NCM measure uses covariance between the clean signal and processed speech signals. It is difficult to perform subjective testing to assess the performance of noise reduction techniques being expensive and time consuming. While objective measures based testing replaces the listeners with a computational algorithm which leads to fast, repeatable and cost effective testing procedure [7].

The NCM objective measure showed high correlation with subjective testing in the investigations. Hence it can be used to estimate the speech intelligibility [7].

NCM uses the correlation between input X_i (t) and output Y_i (t), and then the SNR is computed.

From the SNR transmission indices(TI) in each band is computed and using TI the SNR values are mappedbetween 0 and 1 linearly [2, 8].

$$R_i = \frac{\sum_i (X_i(t)-\mu_i).(Y_i(t)-v_i)}{\sum_i (X_i(t)-\mu_i)^2 . \sum_i (Y_i(t)-v_i)^2} (9)$$

$$SNR_i = 10\log(R_i^2/(1-R_i^2))(10)$$

$$TI_i = \frac{SNR_i+15}{30} (11)$$

$$NCM = \frac{1}{\sum_{i=1}^{k} w_i} \sum_{i=1}^{k} w_i . TI_i (12)$$

Where w_i is the band-importance weights applied to each of K bands. The comparison between the three noises is done by calculating NCM for Spectral subtractionand Wiener filter methods.

a) Restaurant Noise

TABLE 1: NCM comparison for Restaurant Noise

File Name	Noisy	Spectral	Wiener
sp07	0.516	0.5552	0.536
sp08	0.5446	0.508	0.4774
sp11	0.4808	0.4848	0.4715
sp12	0.518	0.4947	0.4255
sp13	0.4624	0.4879	0.4426
sp15	0.3108	0.3124	0.3096
sp22	0.5229	0.5127	0.504
sp23	0.4252	0.4323	0.3942
sp25	0.3386	0.492	0.3161
sp27	0.5436	0.5573	0.5155

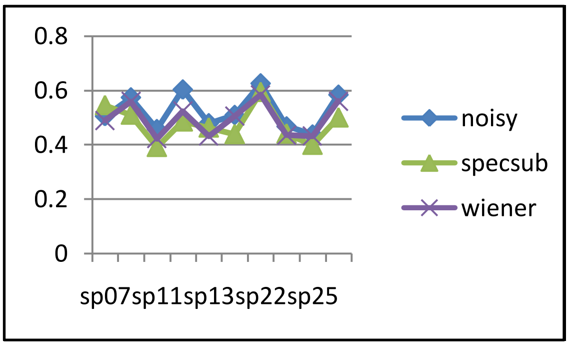

Fig 2: plot of NCM comparison for Restaurant Noise

b) Airport Noise

TABLE 2: NCM comparison for Airport Noise

File Name	Noisy	Spectral	Wiener
sp07	0.5924	0.5825	0.5703
sp08	0.5203	0.5189	0.5079
sp11	0.5468	0.4767	0.4948
sp12	0.5144	0.4582	0.405
sp13	0.4813	0.4637	0.4586
sp15	0.4625	0.4408	0.4292
sp22	0.5144	0.5614	0.5194
sp23	0.4611	0.4221	0.4427
sp25	0.5197	0.4841	0.522
sp27	0.5657	0.5696	0.5441

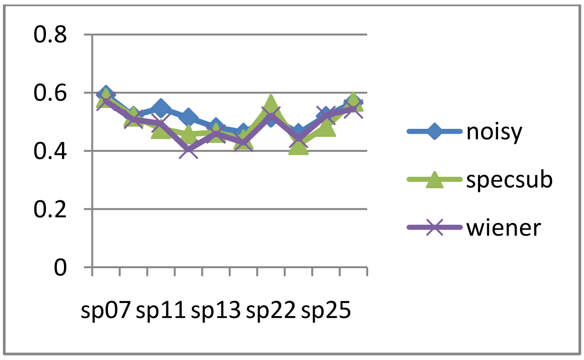

Fig 3: plot of NCM comparison for Airport Noise

c) Station Noise

TABLE 3: NCM comparison for Station Noise

File Name	Noisy	Spectral	Wiener
sp07	0.5052	0.5461	0.4899
sp08	0.5739	0.511	0.5595
sp11	0.4575	0.3922	0.4229
sp12	0.6029	0.4859	0.5243
sp13	0.4791	0.4636	0.4324
sp15	0.5094	0.4394	0.5049
sp22	0.6254	0.5946	0.5843
sp23	0.4674	0.4415	0.4359
sp25	0.4376	0.4001	0.4313
sp27	0.5838	0.5007	0.5605

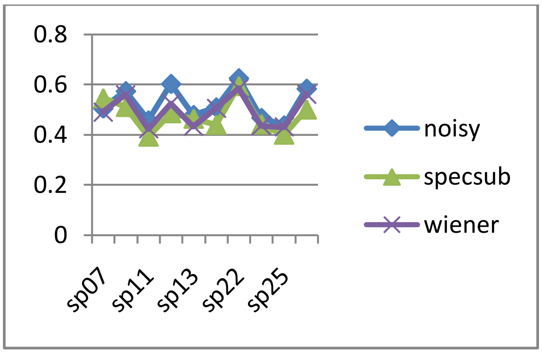

Fig 4: plot of NCM comparison for Station Noise

As discussed above the value of NCM must fall between the range 0<NCM<1. Compared with wiener filter value of NCM for spectral subtraction is large, so it can be concluded that in all 30 cases NCM of spectral subtraction is greater than wiener filter [2].

The main advantage of this objective measure i.e. NCM is that, it is a covariance between the output speech and clean speech, so it is a direct distortion measure. Higher the value of NCM greater is the intelligibility of output.

B. PSNR Comparison

Peak signal-to-noise ratio, often abbreviated -PSNR, is the ratio between the maximum possible power of a signal and the power of corrupting noise.

$$\mathrm{PSNR} = 10\log((Max^2/MSE)) \quad (13)$$

Where MSE is Mean Square Error and Max is the maximum value of amplitude of sample.

Table 4, 5, 6 shows that, the PSNR values are greater for Wiener Filter than spectral subtraction, which shows better noise reduction in Wiener Filter than Spectral Subtraction

TABLE 4: PSNR comparison for Restaurant Noise

File Name	Spectral	Wiener
sp07	32.6457	36.9653
sp08	38.3473	42.1164
sp11	35.5596	38.9167
sp12	27.7929	32.5811
sp13	31.0849	34.5189
sp15	26.3601	30.219
sp22	27.2364	30.8578
sp23	32.5407	35.3456
sp25	33.8595	37.5601
sp27	33.3313	37.2879

TABLE 5: PSNR comparison for Airport Noise

File Name	Spectral	Wiener
sp07	35.7597	39.7655
sp08	34.4625	37.9364
sp11	36.6146	40.1365
sp12	29.0188	33.3532
sp13	28.9247	32.2594
sp15	30.7246	35.2024
sp22	33.4285	36.5943
sp23	31.4071	34.977
sp25	31.6339	34.5971
sp27	35.8537	39.7059

TABLE 6: PSNR comparison for Station Noise

File Name	Spectral	Wiener
sp07	33.7966	37.406
sp08	34.3781	37.9494
sp11	37.0622	40.2712
sp12	29.2625	32.9027
sp13	29.1842	32.8773
sp15	30.7844	34.4297
sp22	28.4846	31.7428
sp23	31.9736	34.8081
sp25	33.5302	36.7082
sp27	34.672	38.2486

C. Spectrogram Comparison

The spectrogram is a visual representation of the spectrum of the frequencies of sound. Figure 5 is spectrogram of clean speech. Figure 6 is spectrogram of noisy input signal which contains 0dB restaurant noise.

Figure 7 is spectrogram of spectral subtraction in which some of noise gets reduced. Figure 8 is spectrogram of output of wiener filter in which the harmonic's are overlapped which causes loss of intellibility.

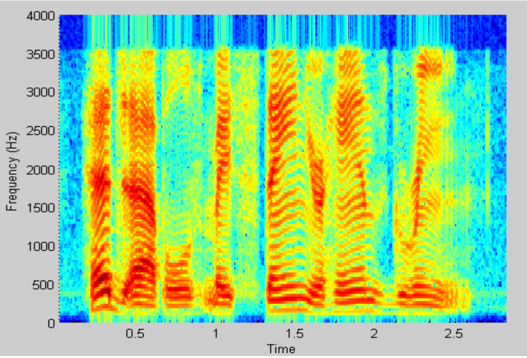

Fig 5: Spectrogram of clean Signal

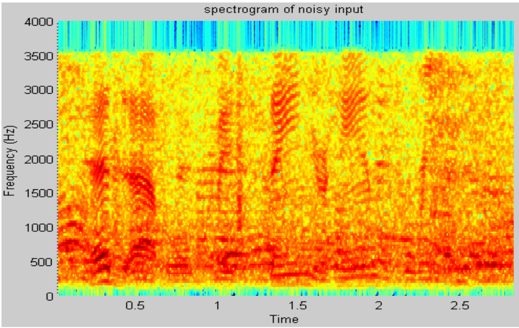

Fig 6: Spectrogram of Noisy Signal

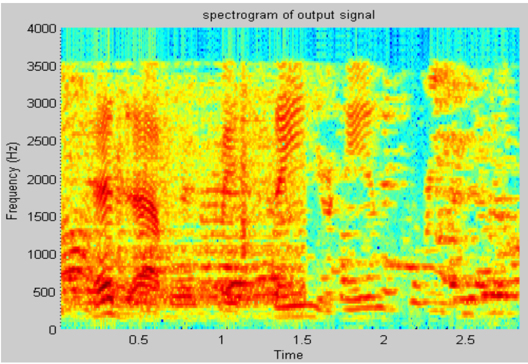

Fig 7: Spectrogram of Output of Spectral Subtraction

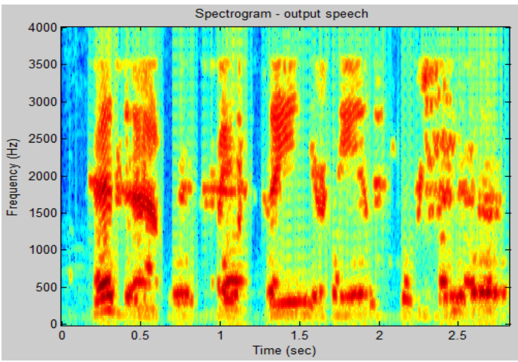

Fig 8: Spectrogram of Output of Wiener Filter

6. CONCLUSION

From the results it can be concluded that the objective measures are able to adequately predict intelligibility in the presence of noise. The value of NCM measure lies between 0 to1. NCM value closer to unity indicates increase in the intelligence of the speech. The evaluated NCM value, an objective measure, for the output of spectral subtraction shows an increased value as compare to Wiener filter method. The increased value of NCM for output of spectral subtraction indicates the better performance in terms of intelligibility of enhanced speech. The greater value of PSNR for output of wiener filter is showing the improvement in the quality of speech which adds to listening comfort. The prominent contribution of the spectral subtraction is the preservation of acoustic cues of the speech that are crucial for the intelligibility of speech.

REFERENCES

[1]. Q. Li, "An auditory-based transform for audio signal processing," IEEE Workshop (WASAAP) Oct.2009.
[2]. Joao Felipe Santos, Philipos C. Loizou, "Performance Comparison of Intrusive Objective Speech Intelligibility and Quality Metrics for Cochlear Implant Users", The University of Texas at Dallas, USA.
[3]. S.ChinaVenkateswarlu, Dr. K.Satya Prasad, Dr. A.SubbaRami Reddy, "Improve Speech Enhancement Using Weiner Filtering", Global Journal of Computer Science and Technology.Volume 11 Issue 7 Version 1.0 May 2011
[4]. Dinesh Lal Arora1, Vijay Kumar, "Enhancing speech using an Adaptive Wiener Filter based algorithm",International Journal Of Advances In Computing and Information Technology.
[5]. Philipos C. Loizou, "Objective speech intelligibility measurement for cochlear implant users in complex listening environment", Speech Communication 55 (2013) 815–824,Science Direct 2013.
[6]. Q. Li, "An Auditory-Based Feature Extraction Algorithm for Robust Speaker Identification under Mismatched Conditions," IEEE Trans. on Audio, Speech, and language processing, 2009
[7]. Q (peter) Li, "Speaker Authentication," Springer.
[8]. Philipos C. Loizou, "Speech Enhancement",second edition.
[9]. Saeed V. Vaseghi, "Advanced Digital Signal Pro-cessing and Noise Reduction", Second Edition.
[10]. http://ecs.utdallas.edu/loizou/speech/noizeus/

