

Speech Enhancement in Presence of Noise using Spectral Subtraction and Wiener Filter

¹Gupteswar Sahu, ²D. Arun Kumar, ³M. Bala Krishna and ⁴Jami Venkata Suman
Assistant Professor, Department of ECE,
GMR Institute of Technology, Rajam, India.

Abstract— The process of suppressing the back ground noise in speech signals can be improved by subtraction of an estimate of the average noise spectrum from the noisy signal spectrum. The main objective of this paper is to implement and evaluate speech enhancement techniques based on spectral subtraction methods in presence of noise. The two techniques discussed in this paper are spectral subtraction filter and wiener filter. The simulation results reveal the superiority of the proposed methods for different back ground noises.

Keywords- Spectral Subtraction, Additive white Gaussian Noise, Weiner Filter, Power spectrum.

I. INTRODUCTION

The main objective of the speech enhancement is to minimize the effects of noise on speech by improving the perceptual quality of noisy speech. In a real world environment most of the times the speech signal is accompanied by back ground noise. The presence of the additive back ground noise like car noise, babble noise, train noise affects the quality of the speech signal. A number of speech enhancement methods are proposed to enhance the quality of the degrade speech. A majority of the speech enhancement techniques can be grouped into temporal processing and spectral processing methods. In temporal processing techniques, the processing is done in the time domain, and in spectral processing techniques, the enhancement achieved by processing the degraded speech signal in frequency domain. The proper selection of the enhancement method depends on the type of degradation.

Generally, the speech enhancement problem consists of a family of problems characterized by the type of noise source, the way noise is interfered with the clean signal, the number of microphone outputs and number of voice channels available for enhancement [1]. Time domain techniques are developed based on Infinite Impulse Response (IIR) filters, Finite Impulse Response (FIR) filters, Linear Predictive Coefficients (LPC), Hidden Markov Models (HMM) and Kalman filters. In transformation domain technique, transformation is

performed over degraded speech before filtering. Followed by an inverse transformation to reinstate the original speech. The main advantage of performing noise reduction lies in the relative simplicity of distinguishing and removing noise from the speech signal [3]. Most of the research work carried out on speech enhancement uses the additive noise model to model background noise. In the time domain one of the simplest methods used to reduce noise is the comb filters, which exploits the periodicity of the voiced signal. The basic approach of this method is to use a FIR filter whose coefficients are separated by the pitch period. Linear predictive coding (LPC) is a widely used technique, which is an iterative scheme to estimate the LP parameters of the speech signal. This is based on autoregressive model of the speech signal. The LPC proven to be an effective and computationally efficient choice for speech enhancement. The kalman filtering is one of the important time domain method used for speech enhancement. The main objective for using kalman filters comes from the following advantages. (a). It can adapt to non stationary signals (e.g. Speech signal); (b). It operates on finite data set; (c) it makes use of speech and noise production model. Hidden Markov Model is one of the dominant techniques used for acoustic models. A general HMM model consists of several interconnected states, which are designed to capture time varying signal characteristics and considered to be the generalization of Gaussian Mixture Model (GMM) The model jumps from one state to another according to the signal and the state transition relationship between the states. The HMM based enhancement techniques yields reasonably well enhanced speech. Spectral subtraction can be categorized as non parametric speech enhancement method. This is used for enhancing speech degraded by additive stationary background noise. Wiener filtering is one of the alternative methods to spectral subtraction to enhance the degraded speech signal.

In this paper we focused on spectral subtraction noise removal approach in speech processing along with Wiener filtering approach.. The spectral subtraction is a popular method that is used to enhance the speech in presence of back ground noise. Its main advantage

comes from its minimal complexity and low computational load. In this method the speech spectrum enhanced by subtracting an average noise spectrum from the noisy speech spectrum. Here it is assumed that, noise is uncorrelated and additive to the noise.

II. Spectral Subtraction Method

It is based on the basic principal of restoring the signal by subtracting an estimate of the noise spectrum from the noisy speech spectrum. The noise spectrum can be estimated from the 'pauses' and 'quite' periods in the speech signal. When there is no speech being said and only noise is present. Basic assumption is noise is additive. Spectral subtraction of the signal takes place in the frequency domain. The data from the signal are segmented and windowed by using Hamming window followed by FFT to transform the signal from time domain to frequency domain.

Let $p(k)$ and $n(k)$ be represented by a windowed speech signal and noise signal respectively. The sum of the two is then denoted by $x(k)$,

$$x(k) = p(k) + n(k). \quad (1)$$

Where $x(k)$, $p(k)$ and $n(k)$ are the noisy signal, the original signal and the noise respectively.

By applying Fourier transform on both sides of equation (1).

$$X(e^{j\omega}) = P(e^{j\omega}) + N(e^{j\omega}) \quad (2)$$

Where $X(e^{j\omega})$, $P(e^{j\omega})$ and $N(e^{j\omega})$ are the fourier transforms of the noisy signal, the original signal and the additive noise.

Further the incoming signal is $x(k)$ is divided into segments of length N . and each segment is windowed by using a hamming window and then transformation is performed via FFT. The window signal is given as follows

$$\begin{aligned} x_w(m) &= w(m) x(m) \\ &= w(m) [p(m) + n(m)] \\ &= p_w(m) + n_w(m) \end{aligned} \quad (3)$$

The frequency domain representation for the windowed operation can be represented as

$$X_w(f) = P_w(f) + N_w(f) \quad (4)$$

In the frequency domain, with their respective Fourier Transforms, The power spectrum of the noisy signal is given by

$$|X_w(w)|^2 = |P_w(w)|^2 + |N_w(w)|^2 + P_w(w).N_w(w)^* + P_w(w)^*.N_w(w) \quad (5)$$

If we assume the noise component $n(k)$ is un-correlated to $p(n)$. Then the terms $P_w(w).N_w(w)^*$ and $P_w(w)^*.N_w(w)$ are reduced to zero. From the above based assumptions the estimation of clean speech can be estimated as follows

$$|P_w(w)|^2 = |X_w(w)|^2 - |N_w(w)|^2 \quad (6)$$

A more general form can be derived by generalizing the exponent from 2 to b .

$$|P_w(w)|^b = |X_w(w)|^b - \alpha |N_w(w)|^b \quad (7)$$

Where b represents the power exponent, for magnitude spectral subtraction, the exponent $b=1$, and for power spectral subtraction, $b=2$. The parameter α controls the amount of noise subtracted from the noisy signal. For full noise subtraction $\alpha=1$, and for over subtraction $\alpha>1$.

III. Algorithm

The overview of the algorithm presented in Fig. 1. The input signals are digitized with a sampling rate of 8 KHz. The analysis window is a 128 point Hamming window and the overlap between two successive windows is set to be 50%. In order to avoid wrap around errors each frame is zero padded to 256 point. The frequency transformation of the signal is done by using FFT.

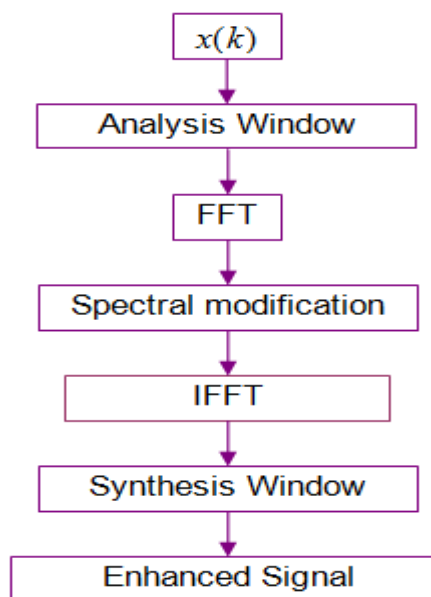


Fig. 1: Block Diagram of Spectral Subtraction Method

IV. Spectral Subtraction Via Wiener Filtration

Weiner filtering was first proposed by Norbert Wiener in 1949. This type of filter generally is used to estimate or to predict the signal in presence of noise. For the implementation of the Wiener filter requires the power spectra of the signal and the noise process. The frequency domain representation of the noisy signal, the original signal and the additive white noise is represented by equation (5). The formulation for the Wiener filter is as follows.

$$W(f) = \left(\frac{P_{xx}(f)}{P_{xx}(f) + P_{yy}(f)} \right) \quad (8)$$

Where $P_{xx}(f)$ and $P_{yy}(f)$ represents the estimated power spectra of the noise free signal and the background noise. Which are assumed to be uncorrelated and stationary. After calculating the transfer function $W(f)$. Dividing equation (8) by $P_{yy}(f)$.

$$W(f) = \frac{SNR(f)}{SNR(f) + 1} \quad (9)$$

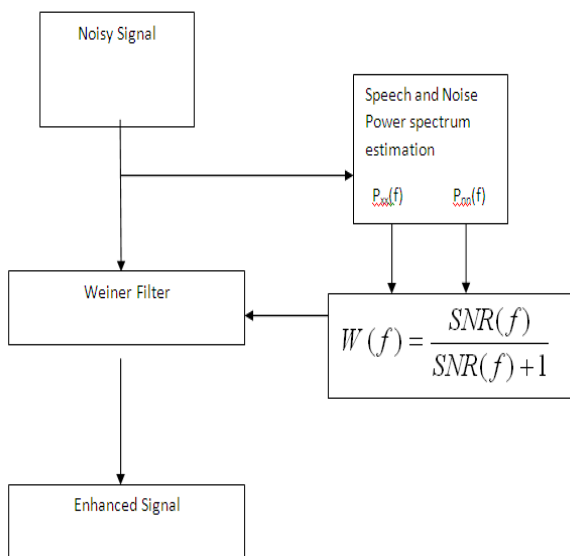


Fig. 2 shows the block diagram of Wiener filter

V. Results and Discussions

For the evaluation of performance of spectral subtraction and wiener filter, a time spectrum of several utterances in presence of noise is analyzed and to show the improvement of the noisy speech signal, to conduct experiments we have used speech signals constituted by sentences pronounced in English language by male and female speakers. The speech signals are sampled with 8kHz and are corrupted by three types of additive noise

(car noise, train noise and babble noise) with different signal to noise ratio (10, 5, 0dB). The amount of noise reduction is generally measured with the SNR improvement.

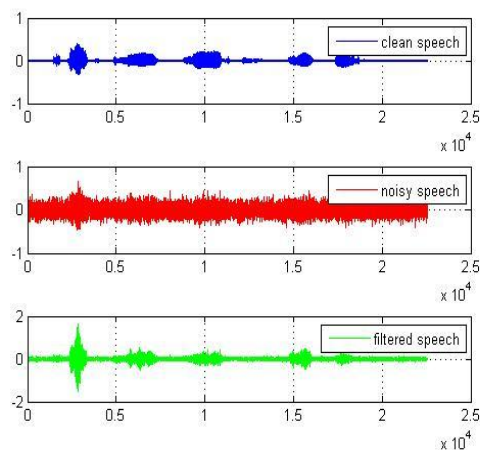


Fig. 3. Clean Speech signal (blue), Clean speech signal added with 5 db AWGN (red), and filtered signal (green).

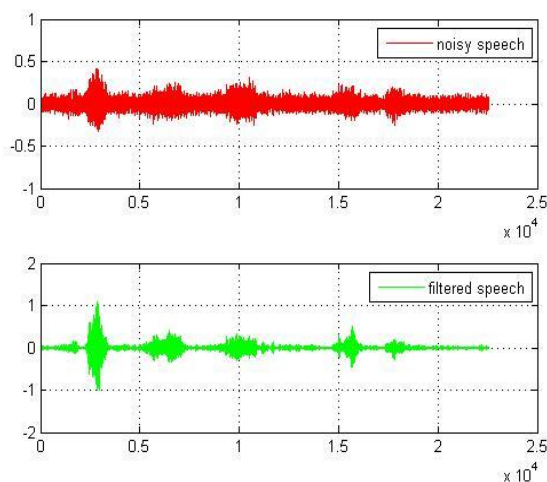


Fig. 4. Speech signal with car noise of 0 db (red) and filtered signal (Green)

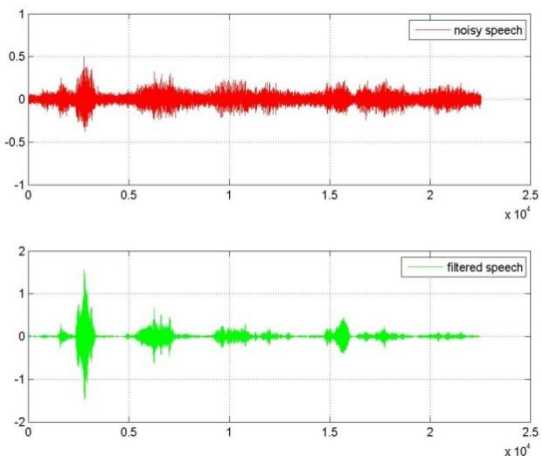


Fig. 5. Speech signal with train noise of 0 db (red) and filtered signal (Green).

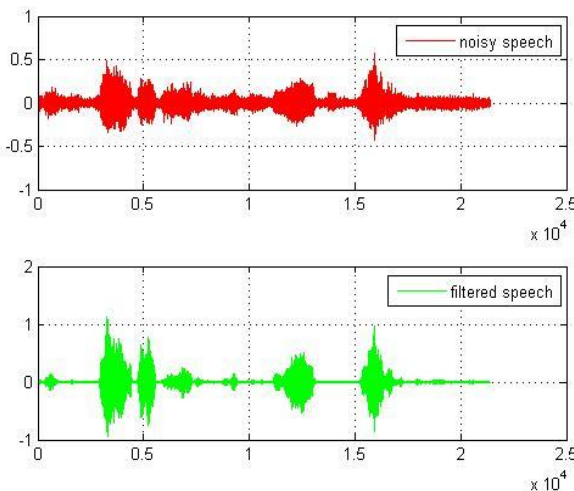


Fig. 8. Speech signal with car noise of 5 db (red) and filtered signal (Green).

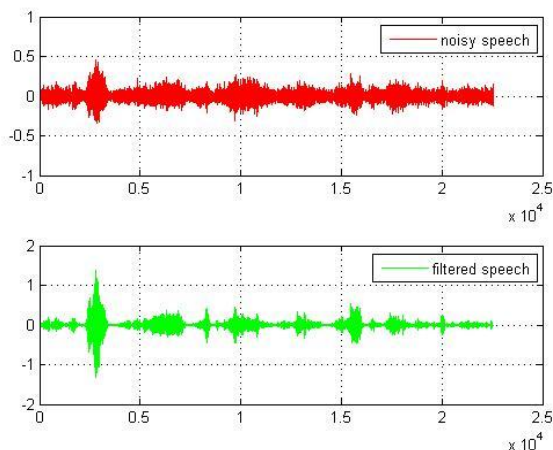


Fig. 6. Speech signals with babble noise of 0 db (red) and filtered signal (Green).

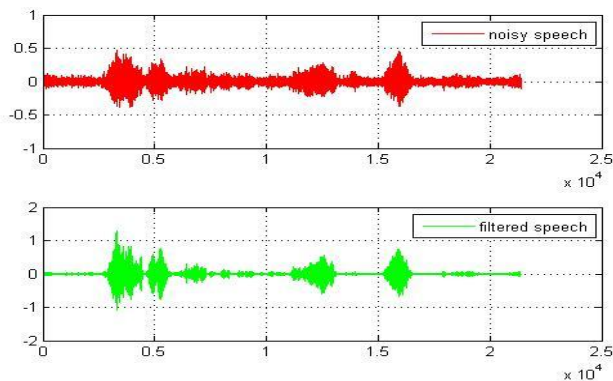


Fig. 9. Speech signal with car noise of 5 db (red) and filtered signal (Green).

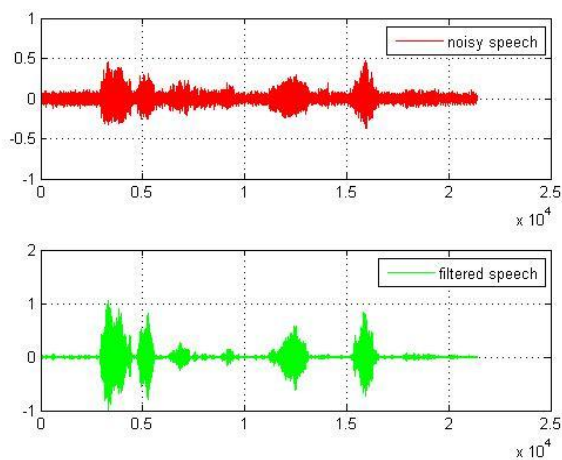


Fig. 7. Speech signal with car noise of 5 db (red) and filtered signal (Green).

Figure [3] to [6] shows the results of spectral subtraction working at SNR of 0 db for noisy speech signal, car noise, train noise and babble noise. From the figures it can be seen that the filter does indeed remove the noise. It shows the magnitude of the speech against time.

Figure [7] to [9] shows the results from the wiener filter implementation. From these figures it is clear that the noisy signal has been filtered and up to certain extent noise has been removed.

VI. TESTING

The testing method involves the listening of the filtered speech with listeners and examining the results of the filtered signal. The testing results appear to be relatively close in quality of filtered speech for both the methods.

CONCLUSION

In this paper, the speech enhancement method by using spectral subtraction and wiener filter had been implemented and analyzed. The methods improve the speech quality by increasing the signal to noise ratio. This method provides a definite improvement compared to other traditional speech enhancement methods. The results from both simulation and evaluation suggest that, this method achieves better reduction of the noise for different noisy signals such as car noise, train noise and babble noise.

REFERENCES

1. S. Boll, "Suppression of acoustic noise in speech using spectral subtraction," IEEE Transactions on Acoustics, Speech, Signal Processing vol.27, pp. 113-120, Apr. 1979.
2. Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error short-term spectral amplitude estimator," IEEE Transactions on Acoustics, Speech, Signal Processing vol. ASSP-32, No.6, pp. 1109-1121, Dec.1984.
3. Y. M. Cheng and D. O'Shaughnessy, "Speech enhancement based conceptually on auditory evidence," IEEE Trans. Signal Processing, vol. 39, Sept. 1991.
4. Y. Ephraim and H. L. Van Trees, "A signal subspace approach for speech enhancement," IEEE Trans. Speech and Audio Processing, vol. 3, July 1995.
5. Scalart, P., and Filho, J. "Speech enhancement based on a priori signal to noise estimation," in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 629-632. 1996.
6. Md. Kamrul Hasan, Sayeef Salahuddin, and M. Rezwan Khan, "A Modified A Priori SNR for Speech Enhancement Using Spectral Subtraction Rules," IEEE Signal Processing Letters, vol. 11, no. 4, April 2004.
7. B. L. Sim, Y. C. Tong, J. S. Chang, and C. T. Tan, "A parametric formulation of the generalized spectral subtraction method," IEEE Trans. Speech Audio Processing, vol. 6, pp. 328-337, July