International Journal of Electronics Communication and Computer Engineering
Volume 5, Issue (4) July, Technovision-2014, ISSN 2249–071X

Technovision-2014: 1st International Conference at SITS, Narhe, Pune on April 5-6, 2014

Review of Speech Enhancement Techniques using Statistical Approach

Vyankatesh Chapke
Department of E & TC
Indira College of Engineering and Management
Parandwadi, Pune
Email: chapkevyankatesh@gmail.com

Prof. Harjeet Kaur
Department of E & TC
Indira College of Engineering and Management
Parandwadi, Pune.
Email: mail2hkaur@gmail.com

***Abstract* – Enhancing of speech affected by the noise is the most important factor of speech enhancement. The main objective of speech enhancement is to improve perceptual aspects of speech such as overall quality, intelligibility and degree of listener fatigue [1]. Speech enhancement also results in clarity speech, pleasantness of the speech signal, compatibility with some other method in speech processing. In this paper comparative performance analysis of different speech enhancement method has been explained. Basic spectral subtraction method, minimum (MMSE), subspace methods, speech enhancement using an adaptive wiener filtering approach are the different methods that has been presented in this paper. In section 1 we have explained the introduction of speech enhancement and in section 2 and section 3 we have explained different methods and give references.**

***Keywords* – Speech Enhancement, Statistical Approach, Basic Spectral Subtraction Method, Minimum (MMSE), Subspace Methods.**

I. INTRODUCTION

In speech communication speech signal is always degraded by the noise. In order to achieve the Speech enhancement several techniques has been proposed using different algorithms with the help of mathematical approach as well as stimulation. Spectral subtraction method is one of the methods that we have presented in this paper. it is used to remove the effect of the background noise. There are different methods of spectral subtraction like spectral subtraction with over subtraction, non linear spectral subtraction (NSS), Nonlinear weighted noise subtraction (NWNS), multiband spectral subtraction [1]. The performance of wiener filtering approach is depend on the quality of the output speech signal, its intelligibility and its clarity.[1-5].Signal subspace approach is the another technique for enhancing the speech signal degrade by uncorrelated additive noise or colored noise[5,7]. The MMSE estimator is also one of the algorithms proposed for removal of additive background noise. It is a single channel speech enhancement technique for the enhancement of speech degraded by additive background noise.[6].

II. SPEECH ENHANCEMENT

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 [11]. 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[11].

III. SPEECH ENHANCEMENT METHODS

3.1 - Single channel speech enhancement
3.2 - Multiple channel speech enhancement
3.1.1- Basic Spectral Subtraction Method

The spectral subtraction method is historically one of the first algorithms proposed for noise reduction [7].It is very simple method and easy to implement, it based on the principle that we can obtain an estimate of the clean signal spectrum by subtracting an estimate of the noise spectrum from the noisy Speech spectrum. The noise spectrum can be estimated, and updated, during the periods when the signal is absent or when only noise is present i.e. during speech pauses'. Basic assumption is noise is additive, its spectrum does not change with time means noise is stationary or it's slowly time varying signal, whose spectrum does not change significantly between the updating periods. Block diagram as shown in figure 1.

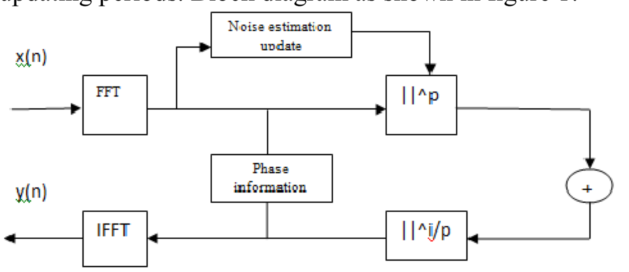

Figure 1 basic spectral subtraction method
X(n)=noisy speech
Y(n)=enhanced speech

3.1.2- speech enhancement using an adaptive wiener filtering approach.

International Journal of Electronics Communication and Computer Engineering
Volume 5, Issue (4) July, Technovision-2014, ISSN 2249–071X

The Wiener filter is a popular technique that has been used in many signal enhancement methods. The basic principle of the Wiener filter is to obtain an estimate of the clean signal from that corrupted by additive noise [8]. This estimate is obtained by minimizing the Mean Square Error (MSE) between the desired signal $s(n)$ and the estimated signal $\hat{s}(n)$. The frequency domain solution to this optimization Problem gives the following filter transfer function [8]

$$H(\omega) = \frac{Ps(\omega)}{Ps(\omega) + PV(\omega)} \tag{1}$$

Where $Ps(\omega)$ and $Pv(\omega)$ are the power spectral densities of the clean and the noise signals, respectively. This formula can be derived considering the signal s and the noise v as uncorrelated and stationary signals. The SNR is defined by: [8]

$$SNR = \frac{Ps(\omega)}{Pv(\omega)} \tag{2}$$

3.1.3- Signal subspace method

The subspace enhancement algorithms make an explicit compromise between signal distortion and noise and several authors have suggested basing this compromise on perceptual models to permit higher noise in spectral regions where it will be masked, thereby allowing a reduction in distortion [9].

3.1.4- Minimum Mean Square Estimators (MMSE)

In an influential paper [10] proposed an optimal MMSE estimation of the short time spectral amplitude (STSA); its structure is the same as that of spectral subtraction but in contrast to the Wiener filtering motivation of spectral subtraction, it optimizes the estimate of the real rather than complex spectral amplitude In this section, we explain a historically important speech enhancement method, that is, the MMSE-STSA method [12] proposed by Ephraim and Malah in 1984. Ephraim and Malah have proposed not only an efficient spectral gain, but also an efficient estimation technique to get the a priori SNR. The MMSE-STSA method is derived by minimizing a conditional mean square value of the short time spectral amplitude. The cost function to be
minimized is given by

$$\begin{aligned} J_{MMSE} &= E\left[\left|X - \hat{X}\right|^2 \mid Y\right] \\ &= \int_{-\infty}^{\infty} |X|^2 p(X \mid Y)\,dx + |\hat{X}|^2 - \hat{X}\int_{-\infty}^{\infty} X^* p(X \mid Y)\,dx \\ &\quad - \hat{X}^* \int_{-\infty}^{\infty} X p(X \mid Y)\,dx, \end{aligned} \tag{3}$$

Where $p(X \mid Y)$ denotes the conditional PDF of X . The estimated speech spectrum which minimizes (mmse)is given as

$$\hat{X}_{MMSE} = \int_{-\infty}^{\infty} X p(X \mid Y)\,dx = E[X \mid Y] \tag{4}$$

3.2 -Multiple channel techniques—

3.2.1 -Adaptive filter

In a multichannel system (p>1) it is possible to remove noise and interference signals by applying sophisticated adaptive filtering techniques that use spatial or redundant information. However there are a number of noise and distortion sources that cannot be minimized by increasing the number of microphones. Examples of this are the surveillance, recording, and playback equipment. There are several classes of adaptive filtering (Honig & Messerschmitt, 1984)[14] that can be useful for speech enhancement. In the estimator application .the internal parameters of the adaptive filter are used as estimate. In the predictor application , the filter is used to filter an input signal, x(n),in order to minimize the output signal, e(n)=x(n)-y(n), within the constraints of the filter structure. A predictor structure is a linear weighting of some finite number of past input samples used to estimate or predict the current input sample. In the joint-process estimator application there are two inputs x(n) and d(n). The objective is usually to minimize the size of the output signal, e(n)=d(n)-y(n) , in which case the objective of the adaptive filter itself is to generate an estimate of d(n), based on a filtered version of x(n),y(n).[14]

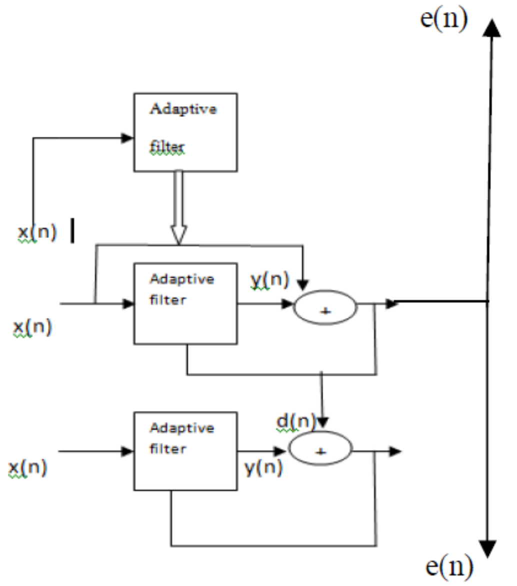

Fig.2. Classes of adaptive filters

3.2.2- Multichannel wiener filtering

Wiener filtering is most widely used technique for the enhancement of speech. In multichannel [14], broadband noisy speech signal is divided into multiple sub bands through filterbank called analysis. Each subband then filtered from Wiener filter for enhancement. All the outputs from Wiener filter recombined through inverse filterbank are called synthesis.[14] In order to formulate Wiener filter, it is required to have a desired signal d[k][n]. We need to have autocorrelation matrix and cross correlation vector, which can be calculated as,

International Journal of Electronics Communication and Computer Engineering
Volume 5, Issue (4) July, Technovision-2014, ISSN 2249–071X

Technovision-2014: 1st International Conference at SITS, Narhe, Pune on April 5-6, 2014

$$R_x^{[k]} = E\{x^{[k]}[n]x^{[k]H}[n]\}$$
$$r_{dx}^{[k]} = E\{d^{[k]*}[n]x^{[k]}[n]\} \quad (5)$$

Practically, these are calculated according to the array model and sources. Wiener Hopf can be calculated as [15],

$$R_x^{[k]}w^k = r_{dx}^{[k]}$$
$$w^{[k]} = R_x^{[k]-1}r_{dx}^{[k]} \quad (6)$$

Error signal can be calculated from the difference between desired signal and the beam forming output signal as,

$$e[n] = d^{[k]}[n] - w^{[k]H}x^{[k]}[n] \quad (7)$$

IV. Comparison Table for Single Channel and Multi Channel Techniques

Single Channel Speech Enhancement Techniques	**Multiple Channel Speech Enhancement Techniques**
Basic Spectral subtraction method— It is based on the principle that we can obtain an estimate of the clean signal spectrum by subtracting an estimate of the noise spectrum from the noisy Speech spectrum [7].	**Adaptive filter** In a multichannel system it is possible to remove noise and interference signals by applying sophisticated adaptive filtering techniques [14]. Examples of this are the surveillance, recording, and playback equipment [14].
Signal subspace method- The subspace enhancement algorithms make an explicit compromise between signal distortion and noise [9].	
MMSE-(minimum mean square estimation)- Its structure is the same as that of spectral subtraction but in contrast to the Wiener filtering motivation of spectral subtraction, it optimizes the estimate of the real rather than complex spectral amplitude.	**Wiener filter** Wiener filtering is most widely used technique for the enhancement of speech. In multichannel [14], broadband noisy speech signal is divided into multiple sub bands through filterbank called analysis. Each subband then filtered from Wiener filter for enhancement [14] All the outputs from Wiener filter recombined through inverse filterbank are called synthesis[14]
Adaptive wiener filtering The basic principle of the Wiener filter is to obtain an estimate of the clean signal from that corrupted by additive noise. This estimate is obtained by minimizing the Mean Square Error (MSE) between the desired signal $s(n)$ and the estimated signal $\hat{s}(n)$[8].	

V. Conclusion

Various speech enhancement approaches has been approached in this paper. Studied the basic filter subtraction method. Also studied the filter transfer function for an adaptive wiener filtering approach also the signal to noise ratio for the same. Also seen the signal subspace method and minimum mean square estimators. thus In this paper comparative performance analysis of different speech enhancement method has been explained.

References

[1] Anuprita P. Pawar Electronics and Telecommunication Department AISSMS, College of Engineering Kenedy Road Pune, Maharashtra, India
[2] Boll, S. F., "Suppression of acoustic noise in speech using spectral Subtraction," *IEEE Trans. Acoust., Speech, SignalProcessing*, Vol. ASSP-27, 113–120, 1979.
[3] Berouti, M., R. Schwartz, and J.Makhoul, "Enhancement of speech corrupted by acoustic noise," *Proc. IEEE Int. Conf. Acoust., Speech Signal Processing*, 208–211, 1979.
[4] Ephriam, Y. and H. L. van Trees, "A signal subspace approach for speech enhancement," *Proc.International Conference on Acoustic, Speech and Signal Processing*, Vol. 2, 355–358, Detroit, MI, USA., May 1993.
[5] Haykin, S., *Adaptive Filter Theory*, Prentice- Hall, ISBN 0-13-322760-X, 1996.
[6] Lim, J. S. and A. V. Oppenheim, "All-pole modelling of degraded speech," *IEEE Trans. Acoust., Speech, Signal Processing*, Vol. ASSP-26, June 1978.
[7] Savita Hooda and Smriti Aggarwal Maharishi Markandeshwar University, Mullana (Ambala), INDIA
[8] Berouti,M. Schwartz,R. and Makhoul,J.,"Enhancement of Speech Corrupted by Acoustic Noise", Proc ICASSP 1979, pp208-211,.
[9] Lim, J. S. and A. V. Oppenheim, "Enhancement and band width compression of noisy speech," *Proc. of the IEEE*, Vol. 67, No. 12, 1586–1604, Dec. 1979. [Huetal., 2003, Jabloun and Champagne, 2003, Kim et al., 2003].
[10] Ephraim and Malah [1984]
[11] J. Benesty, S. Makino, J. Chen (ed). *Speech En hancement*. pp.1-8. Springer, 2005. ISBN 978-3-540- 24039-6.
[12] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator," IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 32, no. 6, pp. 1109–1121, 1984
[13] Honig & Messerschmitt, 1984).
[14] B. Sallberg, "Applied methods for blind speech enhancement," Ph.D. dissertation, Sweden, 2008
[15] S. Haykin, Adaptive Filter Theory. John Wiley and Sons, 2002.