ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering
Vol. 2, Issue 7, July 2013

Enhancement of Speech in Noisy Conditions

Anuprita P Pawar[1], Asst.Prof.Kirtimalini.B.Choudhari[2]

PG Student, Dept. of Electronics and Telecommunication, AISSMS C.O.E., Pune University, India[1]

Assistant Professor, Dept. of Electronics, AISSMS C.O.E., Pune University, India[2]

Abstract: The term "Speech Enhancement" refereed as to improve quality or intelligibility of speech signal. Speech signal is often degraded by additive background noise like babble noise, train noise, restaurant noise etc. In such noisy environment listening task is very difficult at the end user. Many times speech enhancement is used for pre processing of speech for computer speech recognition system. This paper presents speech enhancement methods like Spectral Subtraction, Modified Spectral Subtraction and Least Mean Square to reduce additive background noise. Basically these methods are single channel speech enhancement methods. The performance of SS algorithm and LMS algorithm is evaluated by object speech measure like, Signal to Noise Ratio, Mean Square Error, Root Mean Square Error and Normalized Root Mean Square. From result we conclude that the performance of SS algorithm and Modified SS algorithm is better than LMS algorithm. So SS algorithm is widely used in personal communication due its simplicity.

Keywords: Speech enhancement, spectral subtraction, least mean square algorithm, objective quality measure

I. INTRODUCTION

In speech communication, the speech signal is always accompanied by some noise. In most cases background noise of the environment where the source of speech lies, is the main component of noise that adds to the speech signal. Though the obvious effect of this noise addition is to make the listening task difficult for a direct listener, there are many more far reaching negative effects when we process the degraded speech for some other applications. A related problem is processing degraded speech in preparation for coding by a bandwidth compression system. Hence speech enhancement not only involves processing speech signals for human listening but also for further processing prior to listening. Main objective of speech enhancement is to improve the perceptual aspects of speech such as overall quality, intelligibility, or degree of listener fatigue. Research on speech enhancement techniques started more than 40 years ago at AT&T Bell Laboratories by Schroeder as mentioned in [2].Schroeder proposed an analog implementation of the spectral magnitude subtraction method. Then, the method was modified by Schroeder's colleagues in a published work. However, more than 15 later, the spectral subtraction method as proposed by Boll is a popular speech enhancement technique through noise reduction due to its simple underlying concept and its effectiveness in enhancing speech degraded by additive noise. The technique is based on the direct estimation of the short-term spectral magnitude. Noise reduction or speech enhancement algorithms in general, attempt to improve the performance of communication systems when their input or output signals are corrupted by noise. The main objective of speech enhancement is to improve one or more perceptual aspects of speech, such as the speech quality or intelligibility. In this paper, a speech enhancement algorithm using spectral subtraction and LMS algorithm are proposed for hearing aids. NOIZEUS database are used for testing

The structure of this paper is as follows, in section II speech enhancement algorithms are described ,implemented in III section, objective quality measure compared in section IV, section V gives conclusion.

II.SPEECH ENHANCEMENT ALGORITHMS

Speech enhancement process can be achieved by using various speech enhancement algorithms..In this process first speech signal is segmented for 20-30ms i.e. short term Fourier transform (STFT) is taken and windowed. Hamming window is used for windowing. Then Discrete Fourier transforms (DFT) or Fast Fourier Transform (FFT) of segmented and windowed. Generally in speech enhancement Fast Fourier Transform (FFT) is used. Noisy speech signal is taken. FFT of noisy signal is then given to noise estimation block and speech enhancement block. Noise estimation block estimate the noise during the speech pauses and find the noise spectrum. In most speech enhancement algorithms, it is make assumed that an estimate of the noise spectrum is available. The noise estimate can have a major impact on the quality and intelligibility of the enhanced signal. If the noise estimate is too low, unwanted residual noise will be audible, if the noise estimate is too high, speech will be distorted. The simplest approach is to estimate and update the noise spectrum during the silent segments of the signal. Speech enhancement block enhance noisy speech spectrum to generate clean speech signal.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering
Vol. 2, Issue 7, July 2013

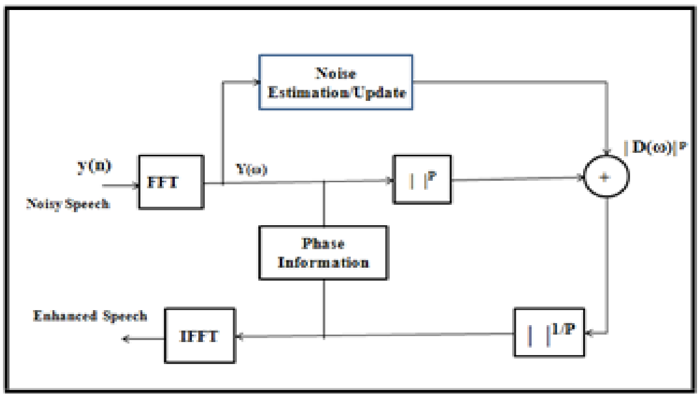

Fig.1-Block diagram of spectral Subtraction

A. Basic Spectral Subtraction Method

The spectral subtraction method is historically one of the first algorithms proposed for noise reduction [4]. 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.

Let y (n) be the noise corrupted input speech signal which is composed of the clean speech signal x(n) and the additive noise signal d(n). In mathematical equation form we can write y (n) in time domain and Fourier domain as given in equation1and 2 respectively

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

$$Y[\omega] = X[\omega] + D[\omega] \qquad (2)$$

Y[w] can be expressed in terms of Magnitude and phase as,

$$Y[\omega] = Y\,|(\omega)|\,e^{j\O y}$$

Where |Y(w)| is the magnitude spectrum and Ø is the phase spectra of the corrupted noisy speech signal. Noise spectrum in terms of magnitude and phase spectra as,

$$D[\omega] = |D[\omega]|\,e^{j\O y}$$

The magnitude of noise spectrum | D (w) | is unknown but can be replaced by its average value computed during non speech activity i.e. during speech pauses. In speech enhancement we are keeping phase spectra constant. The noise phase is replaced by the noisy speech phase Øy that does not affect speech ineligibility. We can estimate the clean speech signal simply by subtracting noise spectrum from noisy speech spectrum¸ in equation form

$$X(\omega) = [|Y(\omega)| - |D(\omega)|]\,e^{j\O y} \qquad (3)$$

Where X(ω) is estimated clean speech signal. Many spectral subtractive algorithms are there depending on the parameters to be subtracted such as magnitude (Amplitude) spectral subtraction¸ power spectral subtraction, autocorrelation subtraction. The estimation of clean speech magnitude signal spectrum is,

$$X[\omega] = |Y[\omega]| - |D[\omega]|$$

Similarly for Power' spectrum is,

$$X[\omega]^2 = |Y[\omega]|^2 - |D[\omega]|^2 \qquad (4)$$

The enhanced speech signal is finally obtained by computing the Inverse Fast Fourier Transform of the estimated clean speech $|X[\omega]|$ for magnitude. Spectrum subtractions and $|X[\omega]|^2$ for power spectrum subtraction using the phase of the noisy speech signal. The more general version of the spectral subtraction algorithms is,

$$X[\omega]^p = |Y[\omega]|^p - |D[\omega]|^p \qquad (5)$$

Where p is the power exponent when p=1 yielding the magnitude spectral subtraction algorithm and when p=2 power spectral subtraction algorithm. The general power spectral subtraction is shown in figure 2.The spectral subtraction algorithm is computationally simple as it only involves a forward and inverse Fourier transform.

B.Modified Spectral Subtraction

Modifications made to the original spectral subtraction method are subtracting an over estimate of the noise power spectrum and preventing the resultant spectrum from going below a preset minimum level (spectral floor).This modifications lead to minimizing the perception of the narrow spectral peaks by decreasing the spectral excursions and thus lower the musical noise effect. Berouti [4] has taken a different approach that does not require access to future information. This Method consists of subtracting an overestimate of the noise power spectrum and presenting the resultant spectral components from going below a preset minimum spectral floor

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering
Vol. 2, Issue 7, July 2013

value. This algorithm is given in equation (6), where $|X_j(\omega)|$ denotes the enhanced spectrum estimated in frame j and $|D(\omega)|$ is the spectrum of the noise obtained during non speech activity.

$$|X_j(\omega)|^2 = |Y_j(\omega)|^2 - |D(\omega)|^2 \quad \text{if } |Y_j(\omega)|^2 - (\alpha+\beta)|D(\omega)|^2$$
$$= \beta\,|D(\omega)| \quad \text{else} \qquad (6)$$

With $\alpha > 1$ and $0 < \beta \leq 1$ .Where α is over subtraction factor and β is the spectral floor parameter. Parameter β controls the amount of residual noise and the amount of perceived Musical noise. If β is too small, the musical noise will became audible but the residual noise will be reduced .If α is too large, then the residual noise will be audible but the musical issues related to spectral subtraction reduces. Parameter α affects the amount of speech spectral distortion. If α is too large then resulting signal will be severely distorted and intelligibility may suffer. If α is too small noise remains in enhanced speech signal. When $\alpha > 1$, the subtraction can remove all of the broadband noise by eliminating most of wide peaks. The parameter α varies from frame to frame according to Burouti [4] as given below,

$$= \alpha o - 3/20\ \text{SNR} - 5\ \text{dB}$$

$$< \text{SNR} \qquad 20\text{dB}$$

C. Least Mean Square Algorithm

The Least Mean Square (LMS) algorithm was first developed by Widow and Hoff in 1959 through their studies of pattern recognition. From there it has become one of the most widely used algorithms in adaptive filtering. The LMS algorithm is a type of adaptive filter known as stochastic gradient-based algorithms as it utilizes the gradient vector of the filter tap weights to converge on the optimal wiener solution [2-4]. It is well known and widely used due to its computational simplicity. It is this simplicity that has made it the benchmark against which all other adaptive filtering algorithms are judged. With each iteration of the LMS algorithm, the filter tap weights of the adaptive filter are updated according to the following formula.

$$W(n+1) = w(n) + 2\mu e(n)\, x(n) \qquad (7)$$

Here x (n) is the input vector of time delayed input values, x (n) = [x(n) x(n-1) x(n-2) .. x(n-N+1)]T. The vector w (n) = [w0 (n) w1 (n) w2 (n) .. wN-1(n)]T represents the coefficients of the adaptive FIR filter tap weight vector at time n. The parameter μ is known as the step size parameter and is a small positive constant. This step size parameter controls the influence of the updating factor. Selection of a suitable value for μ is imperative to the performance of the LMS algorithm, if the value is too small the time the adaptive filter takes to converge on the optimal solution will be too long; if μ is too large the adaptive filter becomes unstable and its output diverges.

D. Implementation of the LMS Algorithm:

Each iteration of the LMS algorithm requires three distinct steps in this order:

1. The output of the FIR filter, y(n) is calculated using equation 8.

$$y(n) = \sum_{i=0}^{N-1} w(n)x(n-i)$$
$$= W^T(n)x(n) \qquad (8)$$

2. The value of the error estimation is calculated using equation 9.

$$e(n) = d(n) - y(n) \qquad (9)$$

3. The tap weights of the FIR vector are updated in preparation for the next iteration, by equation 10

$$W(n+1) = w(n) + 2\mu e(n)\, x(n) \qquad (10)$$

III.SPEECH OBJECTIVE QUALITY MEASURES

The objective comparison of three single channel speech enhancements is carried by evaluating performance of parameters such as, Mean Square Error (MSE), Normalized Mean Square Error (NRMSE), Signal to Noise Ratio (SNR), and Root Mean Square Error. It is based on mathematical comparison of the original and processed speech signal.

A.Signal to Noise Ratio (SNR)

It is most widely used and popular method to measure the quality of speech. It is ratio of signal to noise power in decibels.

$$\text{SNR}_{dB} = 10\log_{10}((\sigma x)^2/(\sigma_d)^2)$$

Where $(\sigma x)^2$ is the mean square of speech signal and $(\sigma_d)^2$ is the mean square difference between the original and reconstructed speech.

B. Mean Square Error

The Mean Squared Error (MSE) is another method classically used to measure a degree of likeliness between signals. It is defined as,

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering
Vol. 2, Issue 7, July 2013

$$MSE = \frac{1}{N} (r (n) - x (n)^2)$$

Where N is length of reconstructed signal, r(n) reconstructed speech signal and x(n) inputl signal.

C. Root Mean Square Error

$$RMSE = \sqrt{\frac{r(n)- \ x(n)2}{n}}$$

D. Normalized Root Mean Square Error (NRMSE)

$$NRMSE= \sqrt{\frac{[X(n)-r(n)]2}{[x(n)-\mu x(n)]2}}$$

Where N is length of input speech signal, x(n) is input speech signal and r(n) is reconstructed speech signal

IV. RESULT OF SS ALGORITHM AND LMS ALGORITHM

The Spectral Subtraction algorithm was simulated using Matlab. Figure 1 shows the result of Rectangular and Hamming window with 512 point FFT. Generally Hamming window is mostly used for speech enhancement. Figure 2 shows the spectrogram of noisy speech signal corrupted by Babble noise at 0dB SNR and enhanced speech signal by Spectral Subtraction. The performance of spectral subtraction algorithm and LMS algorithm is evaluated by according to NOZEUS database. The LMS algorithm was simulated using Matlab. Figure 3 shows the input speech signal which is collected from the computer system through microphone. Figure 4 shows the desired signal derived from the input signal. Figure 5 shows the output of filter which is highly stable as compare to spectral subtraction algorithm.Figure 6 shows the mean square error signal calculated from the filter output signal. The adaptive filter is a 1025th order FIR filter. The step size was set to 0.02. The MSE shows that as the algorithm progresses the average value of the cost function decreases.

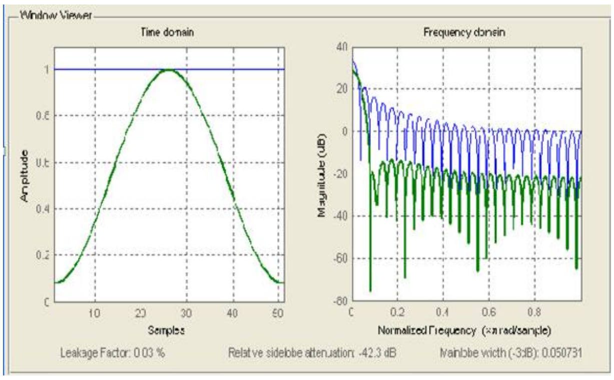

Fig.1 Result of Hamming and Rectangular Window

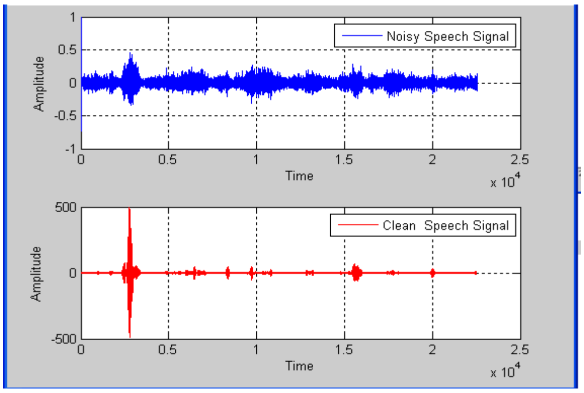

Fig.2 Result of Spectral subtraction algorithm

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering
Vol. 2, Issue 7, July 2013

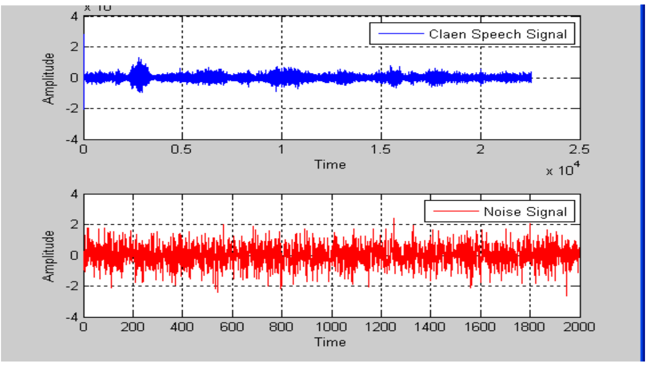

Fig.3 Clean speech and noisy signal

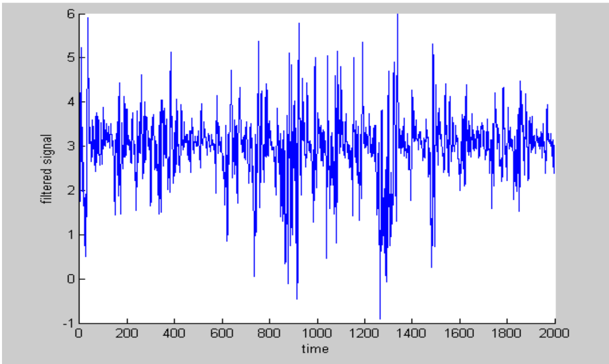

Fig.4 Desired signal

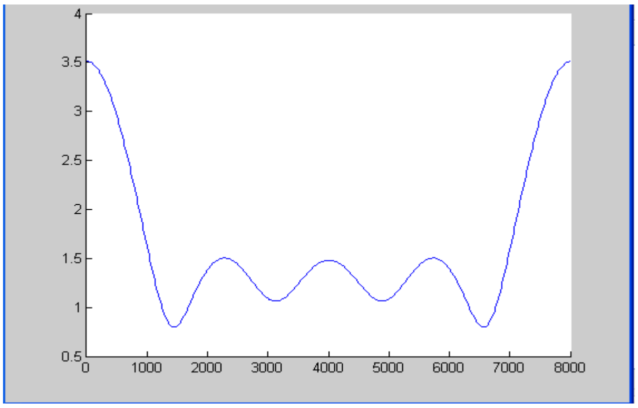

Fig.5 Output of filter

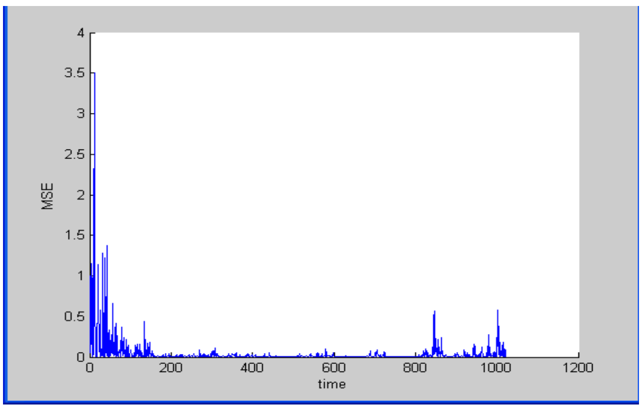

Fig.6 Mean square error

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering
Vol. 2, Issue 7, July 2013

V.PERFORMANCE OF OBJECTIVE QUALITY MEASURES

Table.1 Performance Measures of SS and LMS algorithm based on enhanced SNR

Input SNR dB	Enhanced SNR dB	
	SS Algorithm	LMS Algorithm
0 dB	85.86	29.12
5 dB	152.11	34.84
10 dB	105.67	39.45
15 dB	97.58	43.34

Table.2 Performance Measures of SS and LMS algorithm based on Mean Squared Error

Input SNR dB	Mean Square Error	
	SS Algorithm	LMS Algorithm
0 dB	0.102	3.424
5 dB	0.0100	11.70
10 dB	0.0102	1.237
15 dB	0.0101	14.23

Table.3 Performance Measures of SS and LMS algorithm based on Root Mean Square Error

Input SNR dB	Root Mean Square Error	
	SS Algorithm	LMS Algorithm
0 dB	0.1009	0.001
5 dB	0.1002	3.24
10 dB	0.1009	0.001
15 dB	0.1006	1.07

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering
Vol. 2, Issue 7, July 2013

Table.4 Performance Measures of SS and LMS algorithm based on Normalized Root Mean Square Error

Input SNR dB	Normalized Root Mean Square Error	
	SS Algorithm	LMS Algorithm
0 dB	41.32	1.070
5 dB	42.66	0.6428
10 dB	42.66	0.3119
15 dB	42.66	0.6827

From the above Table I to Table IV results, it is clearly found that SS single channel speech enhancement performed well in both clean and noisy environment.

VI. CONCLUSION

The main objective of the speech enhancement is to bring up the performance in the presence of noise and echo interference to the performance obtained with pure speech signals, which is the ideal case. Thus, our aim was to approach the performance of single channel based speech enhancement techniques to that in the case of ideal signal. Another objective of this work is to compare objective performance of LMS and SS based single channel speech enhancement techniques and the parameters used for comparison are Mean Square Error, Normalized Mean Square Error, Signal to Noise Ratio and Root Mean Square Error evaluation also proved that Spectral Subtraction enhancement technique perform better due to good speech reconstruction quality

REFERENCES

[1] Yi Hu and Philipos C. Loizou, "Subjective comparison and evaluation of speech enhancement algorithms" IEEE Trans.Speech Audio Proc.2007:49(7): 588–601.
[2] Kim W, Kang S, and ko H.(2000) "Spectral subtraction Based on phonetic dependency and masking effects" IEEE Proc.vision image signal process, 147(5),pp423-427
[3]Phillips C Loizou "Speech enhancement theory and practice"1st ed. Boca Raton, FL.: CRC, 2007. Releases Taylor & Francis
[4]Berouti,M.Schwartz,R.andMakhoul,J.,"Enhancement of Speech Corrupted by Acoustic Noise", Proc ICASSP 1979, pp208-211,.
[5] Boll,S.F.,"Suppression of Acoustic Noise in Speech using Spectral Subtraction", IEEE Trans ASSP 27(2):113-120, April 1979
[6] Lockwoord, P. and Boudy,J.," Experiments with a Nonlinear Spectral Subtractor (NSS), Hidden Markov Models and the projection, for robust speech recognition in cars", Speech Communication, 11, pp215-228, Elsevier 1992.
[7]Liuyang Gao and Yunfei "Speech enhancement algorithm based on improved spectral suntraction " IEEE 2009
[8] Anuradha R. Fukane," Different Approaches of Spectral Subtraction method for Enhancing the Speech Signal in Noisy Environments" International Journal of Scientific & Engineering Research, Volume 2, Issue 5, May-2011
[9] Paruav Goel and Anil Garg "Review of spectral subtraction technique for speech enhancement" IJECT Vol 2 Dec 2011.
[10] Kamath S. and Loizou P.(2002) "A multiband spectral subtraction methods for enhancing speech corrupted by colored noise" Proc. IEEE Intl. Conf. Acoustics, Speech, Signal Processing
[12] Sim B, Tong Y, chang J., Tan C.(1998) 'A parametric Formulation of the generalized spectral subtraction method ' IEEE. Trans.
[13] He C, and Zweig G. (1999) "Adaptive two band spectral subtraction with multi window spectral estimation" Proc. IEEE Intl. Conf. Acoustics, Speech, Signal Processing ,2,