ISSN: 2278 – 909X
International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)
Volume 5, Issue 4, April 2016

A Brief Review on Speech Enhancement Algorithms

Ambalika, Er. Sonia Saini

***Abstract*— Speech is a fundamental and common medium to communicate. Since, background or additive noise is present in the channel that degrades the signal quality and its performance. Speech enhancement techniques have been widely used for minimizing the undesirable background noises. This review paper deals with the speech enhancing techniques i.e, spectral subtraction, weiner algorithm, MMSE, and decision directed approach. These techniques are used to improve the quality of speech and enhanced the signal.**

***Index Terms*— Speech enhancement, Spectral subtraction, Weiner Filtering, Decision directed approach.**

I. Introduction

Speech plays an important role in communication, and it is one of the essential functions of human beings. Speech is the most efficient and reliable form of exchanging information or thoughts among human[1]. Listening to speech or audio signals becomes more difficult because of the background noise level dominates[12]. Noise can be defined as an unwanted signal that get added to speech signal and signal become noisy. There are several types of noise, it may be stationary i.e., remains unchanged over time, or non-stationary. One of the most common sources of noise is the background noise, which is always present in different degrees and in any location. Therefore, speech signals get distorted by the ambient noise or additive noise. These distorted or degraded speech signals are known as noisy speech signals.

Enhancement may be defined as the improvement in the value or Quality of something. Speech enhancement is defined as the improvement in quality/intelligibility of a degraded speech signal and is achieved using signal processing tools. Speech enhancement algorithms are the noise suppression techniques[4]. There are three main objectives of speech enhancement:- (a) to improve the perceptual aspects such as quality and/or intelligibility of the processed speech, (b) to improve the robustness of speech coders which is affected by the presence of noise, and (c) to increase the accuracy of the speech recognition systems.

II. BASIC BLOCK DIAGRAM

The basic block diagram for speech enhancement is shown below in Fig. 1

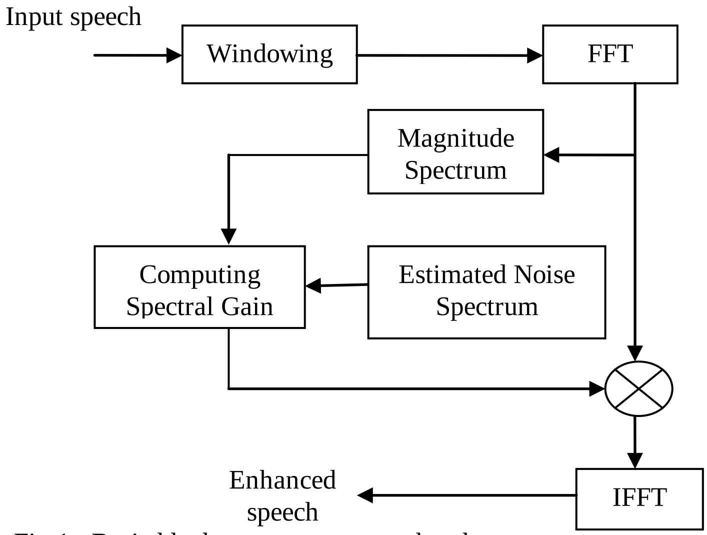

Fig.1 : Basic block diagram for speech enhancement

III. Speech Enhancement Algorithms

There are different types of speech enhancement techniques which are as follows.

1. Spectral Subtraction Method: The Spectral subtraction method is most widely used method because of the simplicity of implementation and lower computational load. This approach has some assumption: (i) noise is additive, (ii) signal is absent. It is based on simple principle. Assuming additive noise, one can obtain an estimate of the clean signal spectrum by subtracting an estimate of the noise spectrum from the noisy speech spectrum. Note that the noise spectrum is estimated, and updated, only when signal is absent or when only noise is present[2].

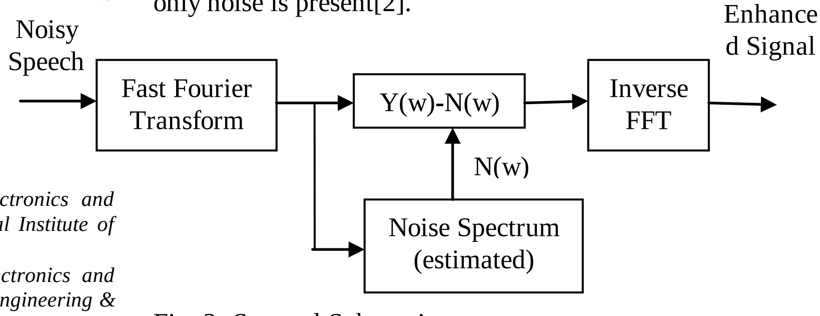

Fig. 2- Spectral Subtraction

Manuscript received April, 2016.

First Author name- *Ambalika M.Tech student in electronics and communication department of Seth Jai Parkash Mukand Lal Institute of Engineering & Technology.*

Second Author name- *Er. Sonia Saini Lecturer in electronics and communication of Seth Jai Parkash Mukand Lal Institute of Engineering & Technology.*

2. Weiner Filtering : The Wiener filter is same as the spectral subtraction in the way that it is derived and makes an attempt to reduce the mean-square error in the frequency domain. It is generally employed in the estimation or prediction of a signal observed in noise. The Wiener filter can also be adaptively estimated used where the surrounding noise has time-varying characteristics[15]. The Filter is used to enhance the quality of speech by removing unwanted noise. The gain function of WF [1] is given by

$$H_{Wiener}(w) = \frac{Ps(w)}{Ps(w) + Pn(w)}$$

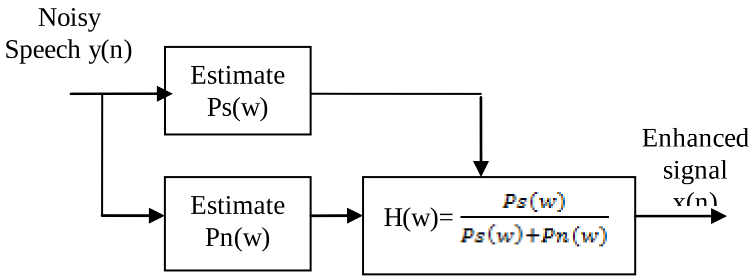

Fig. 3- Weiner Filter

3. Minimum Mean Square Error : MMSE estimation is also known as Ephraim and Malah's estimator used to overcome the problem of the background noise. MMSE method is proposed to minimize the background noise to a considerable amount and thus improved the quality of the resulting enhanced speech. The Minimum mean square error technique is implemented when the input SNR is known. It is an implementation of Weiner Filter[4]. MMSE based algorithms are mainly Minimum Mean Square Error Short-Time Spectral Amplitude(MMSE-STSA) estimator and MMSE Logarithm Spectral Amplitude(MMSE-LSA) estimator. Some power spectrum estimators are used in decision-directed approach for the calculation of *a priori* SNR [9].The only disadvantage of the MMSE processor is additional complexity in determining the linear estimator.

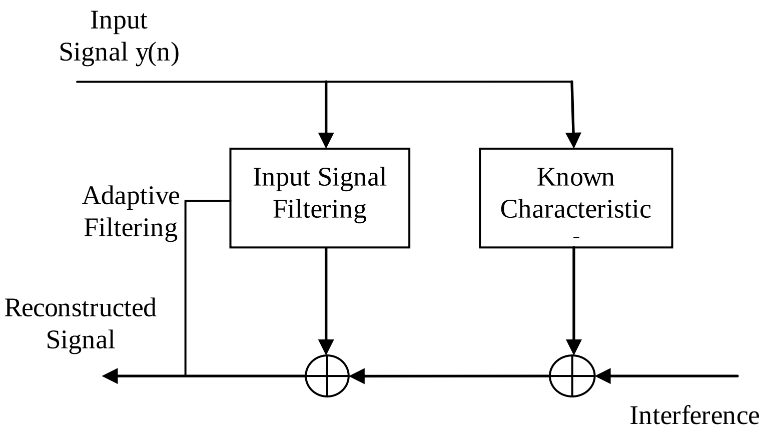

Fig. 4. MMSE Filter

4.Decision Directed Approach : A perceptual- decision directed approach is proposed to improve the performance of a gain factor for noise reduction. This method is performed to enhance the noisy and corrupted speech signal. The decision-directed method is more suitable, to reduce the effect of corrupted signal[5]. Therefore, decision-directed method is repeated again to improve the estimated *a priori* SNR. These procedures specify a two-step-decision directed approach algorithm.

IV. LITERATURE SURVEY

Navneet Upadhyay et. al [1] The spectral subtraction is one of the first algorithms proposed for the enhancement of speech signal. In this method, in the absence of speech signal noise spectrum is estimated, and is subtracted from the noisy speech spectrum to estimate the clean speech. This can also be obtained by multiplying the noisy speech spectrum with a gain function and then combining it with the phase of noisy speech. The drawback or disadvantage of this method is that the presence of processing distortions, known as remnant noise. A number of variations of this method have been developed over the past years to overcome the drawback. These methods make a family of spectral subtractive-type algorithms. The main goal of this paper is to provide a comparison and simulation study of the different forms of subtraction-type algorithms that is basic spectral subtraction, spectral over-subtraction, multi-band spectral subtraction, Wiener filtering, iterative spectral subtraction, and spectral subtraction based on perceptual properties.

Miss. Anuja Chougule et. al [2] The meaning of speech enhancement is speech improvement. The speech enhancement is performed using different techniques and algorithms. Over the past years, there has been attention on the problem of enhancement of speech that affected by additive background noise. For several applications, background suppression is needed. The first algorithm suggested is spectral subtractive algorithm for additive background noise and many modifications are done with time. For SS method noise estimation is important, so that there are various noise estimation algorithms. All these noise estimation algorithms are necessary for removing background noise.

Milind U. Nemade et. al [3] Speech has been embedded into many applications such as speech recognition, development of hearing aid, VoIP, mobile and some different types of personal communication. Speech enhancement techniques have been widely used for eliminating or minimizing background noises. This paper deals with the single channel speech enhancement techniques based on Spectral Subtraction, Wavelet Transform and Adaptive Wiener Filtering. In this paper, quantitative performance of these speech enhancement techniques is compared. The parameters used for comparison are Mean Square Error (MSE), Normalised Mean Square Error(NMSE), Signal to Noise Ratio, Peak Signal to Noise Ratio and Average Absolute Distortion. The results proved the speech enhancing capability for personal communication technique where noise and echo-interference degraded the original speech signal. From the results, it can be seen that the performance of single channel speech enhancement based WT is better than AWF and SS techniques.

Jimish Dodia et. al [4] A fundamental way of communication i.e. speech has been embedded in various applications. In many unavoidable situations, it is rendered helpless trying to deduce the intelligibility of the speech where speech enhancing technique i.e. eliminating the unwanted background noise, comes into picture. In this paper, an attempt has been created towards studying Speech Enhancement techniques :- Spectral Subtraction, Minimum Mean Square Error (MMSE), Kalman and Wiener filter. Based on the observations and analysis of various performance parameters, it is concluded that which one of the methods is most suitable for speech enhancement. Graphic User Interface on MATLAB is used for implementation of the code for various filters.

Ching-Ta Lu et.al [5] The masking properties of the human ear have been applied to accommodate a speech enhancement system. The accuracy of estimated speech spectra plays a vital role in computing the noise masking threshold. Since traditional methods using the power-spectral-subtraction method can provide an appropriate performance, the estimated speech spectra is further improved for computing the noise masking threshold(NMT). The aim of this article is finding a better spectral estimate of speech by the two-step-decision-directed(TSDD) method. This estimate is occupied to compute the noise masking threshold of a perceptual gain factor. The conclusion of this paper show that the amounts of residual noise can be comfortably suppressed by embedding the TSDD algorithm in the perceptual gain factor.

Philipos C. Loizou et. al [6] Speech enhancement algorithms improves the speech quality but not speech intelligibility, and the reasons for that are unclear and uncertain. This paper presented a theoretical framework that is used to analyze potential factors that can affect the intelligibility of processed speech. This framework focuses on the fine-grain analysis of the distortions that introduced by speech enhancement algorithms. It's hypothesized that if these distortions are suitably controlled, then large gain in intelligibility can be achieved. To test this hypothesis, intelligibility tests are organized with human listeners. The aim of these tests is to assess the perceptual effect of the various distortions.

Ekaterina Verteletskaya et al. [7] This paper proposed a method for enhancing of speech corrupted by additive noise. This method is based on the spectral subtraction technique. The proposed algorithm used the weighting function that attenuates frequency spectrum components lying outside identified formants regions. The algorithm used to effects a considerable reduction of the musical noise without significantly distorting the speech.

Yang Lu et. al [8] The traditional power spectral subtraction algorithm is computationally simple to implement however it suffers from musical noise distortion, also the subtractive rules are based on incorrect assumptions concerning the cross terms being zero. In this paper a new geometric approach to spectral subtraction is proposed that addresses these drawbacks of the spectral subtraction algorithm. A method for calculating the cross terms involving the phase differences between the noisy signals and noise is presented. Analysis of the gain function of the proposed algorithm indicated that it contains similar characteristics as the traditional MMSE algorithm. Objective analysis of the proposed algorithm included that it performed considerably better than the traditional spectral subtractive algorithm.

R.Martin et. al [9] described Gaussian statistical model gives a good approximation for the noise DFT coefficients. For speech signals, however, whose typical DFT frame size used in mobile communication are short (10ms -40ms) that assumption is not well fulfilled. It is valid only if the DFT frame size is much longer than the span of correlation of the signal under consideration.

Anthony Stark et. al [10] presented the use of the minimum mean square error (MMSE) spectral energy estimator for use in robust automatic speech recognition (ASR). In the past, it has been common to use the MMSE log-spectral amplitude estimator for this work. Though, this estimator was originally derived under subjective human listening criteria. Hence, its involved suppression rule may not be optimal for use in ASR. On the other hand, it can be presented that the MMSE spectral energy estimator is intimately related to the MMSE Mel-frequency cepstral coefficient (MFCC) estimator. Although, the spectral energy estimator has influenced to suffer from the problem of excessive residual noise. In this, examined the consideration of this residual noise and show that the introduction based speech presence uncertainty (SPU) can quite increase its performance as a front-end ASR enhancement regime.

Mingyang Wu et.al [11] Under noise-free conditions, the quality of reverberant speech is depend on two distinct perceptual components: coloration and long-term reverberation. This correspond to two physical variables: signal-to-reverberant energy ratio and reverberation time, respectively. Inspired by this observation, proposed a two-stage reverberant speech enhancement algorithm that used one microphone. In the first stage, an inverse filter is estimated to minimize coloration effects or increase SRR. The second stage employed the spectral subtraction to reduce the influence of long-term reverberation. The proposed algorithm significantly enhance the quality of reverberant speech.

Premananda B S et.al [12] In mobile phones, perceived quality of speech signal deteriorates significantly in the existence of background noise since near-end or surrounding noise also arrives at the near-end listener's ears. The quality and characteristics of the received signal varies extensively depending upon signal strength and unavoidable background noise in the user environment. It is essential to improve the quality of received speech signal in noisy conditions by establishing the speech enhancement algorithms. This paper focused on the impact of the various background noises on signal degradation and mechanisms to reduce the noise impact for improved speech signal perception. Gain adjustment process in simple time domain and frequency domain approach employing psychoacoustic has been adapted to upgrade the quality and/or intelligibility of the speech signal in the noisy environments by significantly enhancing the speech signals when the noise dominates.

S.China Venkateswarlu et.al [13] Speech enhancement aims to improve speech quality by using various algorithms. Wiener filter are quite simple and workable, but after the estimation of the background noise, one neglects the certainty that the signal is actually speech. Moreover, the phase component of the signal is left untouched. On the other hand, this is possibly not such a bad problem; after all, human ear is not very conscious to phase changes. The third restriction in spectral subtraction techniques is the processing of the speech signal in frames, so the proceeding from one frame to another should be handled with care to avoid discontinuities. A number of techniques or algorithms have been earlier developed in the frequency domain like as an optimal short-time spectral amplitude estimator proposed by Ephraim and Malah containing the estimation of the a priori SNR. This approach reduced the disturbing noise and gives enhanced speech with colorless residual noise. In this paper, proposed a technique based on a Wiener filtering in the uncertainty of signal presence in the noisy observation.

Chabane Boubakir et.al [14] In this study, DFT-based speech enhancement via Minimum Mean-Square Error (MMSE) amplitude estimators was considered. Several variants of the basic approach (MMSE-STSA) have been proposed over the years to address certain drawbacks, the quality of the remnant noise and its trade-off with speech distortion. In this study, presented a comparative study between MMLSA and the estimators based on the Gamma model, followed by an implementation in MATLAB of these techniques and an objective evaluation using a corpus of speech.

Milind U. Nemade et.al [15] The speech, being a fundamental way of communication for the humans, has been embedded in different necessary applications like speech recognition, voice-distance-talk and different forms of personal communications. There are several applications of speech still to be far from reality due to lack of efficient and reliable noise removal mechanism and improving the intelligibility of speech signals. The wide categories of speech enhancement techniques are speech filtering techniques, beam forming techniques, and active noise cancellation methods. In this paper, an effort has been stepped towards surveying the methodologies for speech improvement, and also investigate, how these techniques influence the performance of various application systems such as speech recognition and speech communication.

V. DISCUSSION

In this paper, a survey on speech enhancement algorithms are shown. There are different types of techniques used to improve the signal quality. The spectral subtraction technique is most widely used method because this algorithm is computationally simple to implement. But there are some drawbacks in spectral subtraction technique such as : this method is not applicable for non-stationary signal, it depends on the VAD accuracy, musical noise due to imperfect noise estimation. So, to overcome these problems other techniques are come into use. The MATLAB tool is used to compare the performance of these different algorithms.

REFERENCES

[1] Navneet Upadhyay, Abhijit Karmakar, "Speech Enhancement using Spectral Subtraction-type Algorithms: A Comparison and Simulation Study", Eleventh International Multi-Conference on Information Processing-2015

[2] Miss. Anuja Chougule, Dr. V.V. Patil, "Survey of Noise Estimation Algorithms for Speech Enhancement Using Spectral Subtraction", International Journal on Recent and Innovation Trends in Computing and Communication, vol. 2, Issue 12, December 2014.

[3] Milind U. Nemade, Prof. Satish K. Shah, "Performance Comparision of Single Channel Speech Enhancement Techniques for Personal Communication", International Journal of Innovative Research in Computer and Communication Engineering Vol. 1, Issue 1, March 2013.

[4] Jimish Dodia, "GUI based Performance Analysis of Speech Enhancement Techniques", International Journal of Scientific and Research Publications, vol. 3, Issue 9, September 2013.

[5] Ching-Ta Lu, "Enhancement of Single Channel Speech using Perceptual-Decision-Directed Approach", Science Direct, Speech Communication, Vol. 53, pp. 495-507, December 2011.

[6] Philipos C. Loizou, Gibak Kim, "Reasons Why Current Speech-Enhancement Algorithms do not Improve Speech Intelligibility and Suggested Solutions", IEEE Transactions On Audio, Speech, And Language Processing, Vol. 19, No. 1, January 2011.

[7] Ekaterina Verteletskaya, Boris Simak, "Noise Reduction Based on Modified Spectral Subtraction Method", IAENG International Journal of Computer Science, 38:1, IJCS_38_1_10, 10 February 2011.

[8] Yang Lu, Philipos C. Loizou, "A geometric approach to spectral subtraction", Science Direct, Speech Communication, Vol. 50, pp. 453-466, January 2008.

[9] R. Martin, "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimation," in Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing, pp. 504-512, 2002.

[10] Anthony Stark, Kuldip Paliwal, "Use of Speech Presence Uncertainty with MMSE Spectral Energy Estimation for Robust Automatic Speech Recognition", Science Direct, Speech Communication 53 (2011) p.p. 51–61

[11] Mingyang Wu and DeLiang Wang, "A Two-Stage Algorithm for One-Microphone Reverberant Speech Enhancement", IEEE Transactions On Audio, Speech, And Language Processing, Vol. 14, No. 3, May 2006

[12] Premananda B S and Dr. Uma B V , "Speech Enhancement Algorithm to Reduce the Effect of Background Noise in Mobile Phones", International Journal of Wireless & Mobile Networks (IJWMN) Vol. 5, No. 1, February 2013

[13] S.China Venkateswarlu, Dr. K.Satya Prasad, "Improve Speech Enhancement Using Weiner Filtering", Global Journal of Computer Science and Technology Volume 11 Issue 7 Version 1.0 May 2011

[14] Chabane Boubakir and Daoud Berkani, "Speech Enhancement Using Minimum Mean-Square Error Amplitude Estimators Under Normal and Generalized Gamma Distribution", Journal of Computer Science 6 (7): 700-705, 2010

[15] Milind U. Nemade and Satish K. Shah, "Speech Enhancement Techniques: Quality vs. Intelligibility", International Journal of Future Computer and Communication, Vol. 3, No. 3, June 2014

First Author *Ambalika M.Tech student in electronics and communication department of Seth Jai Parkash Mukand Lal Institute of Engineering & Technology.*

Second Author *Er. Sonia Saini Lecturer in electronics and communication of Seth Jai Parkash Mukand Lal Institute of Engineering & Technology.*