

# A survey on various multichannel speech enhancement algorithms

Soumasunderaswari.D<sup>1</sup> Prashanthini.K<sup>2</sup>

Dean / Professor, Department of ECE, SNS college of Technology, Coimbatore, India<sup>1</sup>

PG scholar, Department of ECE, SNS college of Technology, Coimbatore, India<sup>2</sup>

**Abstract:** Speech enhancement has become inevitable in recent communication systems. While transmitting a signal in a multichannel the quality of the signal may be degraded due to many factors like white Gaussian noise, acoustic additive noise, and acoustic reverberation. The speech signal should be enhanced and the noise should be filtered. In this paper we discuss various techniques that are used for enhancement of speech signal. We discuss main four techniques that are optimal STSA, novel dual microphone technique, suboptimal generalized side lobe canceller and spectro-temporal filtering.

**Keywords:** white Gaussian noise, optimal STSA, novel dual microphone technique, suboptimal generalized side lobe canceller and spectro-temporal filtering.

## INTRODUCTION

The objective speech enhancement is to improve the quality or the intelligibility of the degraded speech signal by using various algorithms. In long distance transmission or in a noisy environment the speech quality will be degraded. In the field of communication speech enhancement techniques are used to reduce the noise, so that it will be helpful in many applications such as in hearing aids, speech recognition, teleconferencing systems and in mobile phones.

The algorithm used for noise reduction in speech signal can be classified into different classes:

- Spectral restoration
- Filtering techniques
- Model-based methods

Widely used method is spectral subtraction. In this method we estimate the noise spectrum, which can be calculated from the region of no speech. Generally noise is added to a greater extent when there is no speech. Considering the noise spectrum to be uniform, it is subtracted from the noisy signal thus giving a clean signal. Speech enhancement is closely related to speech restoration. When speech is degraded, its restoration to the original speech signal often leads to speech enhancement. But there is a difference between speech restoration and speech enhancement. When a speech signal is degraded,

In optimal short time spectral amplitude estimator (STSA), it assumes that the phase of speech and the noise is equal. The optimal spectral amplitude estimation is projected on to a 1-D subspace of complex plane, this simplifies the problem. Based on the criteria of ML, MMSE or MAP for speech and noise spectral magnitudes, separate families of novel estimators are derived assuming the generalized gamma distribution (GGD). By using GGD particular solution can be obtained by substituting statistical shape parameters of the speech and noise signal. Some special cases of GGD method results in Wiener filter (WF) and magnitude spectral subtractor. Quantitative analysis of selected subset of the STSA

restoration involves bringing the processed signal as close as possible to the original signal. Whereas the idea of speech enhancement is to give processed signal better than the original signal. Originally un-degraded signal cannot be further restored, but can be enhanced.

Speech enhancement is desirable in wide variety of contexts. For example, environment in which interfering background noise has been introduced in offices, street and motor vehicles, for example/are common, and the interfering noise generally degraded the intelligibility and the quality of the speech signal.

Other application of speech enhancement include correcting form reverberation, correcting for distortion of the speech of deep/sea divers breathing a helium/oxygen mixture, correcting for the distortion of speech due to physical difficulties of speaker. Another very important application of speech enhancement is the connection between speech enhancement and speech bandwidth compression systems, because of increasing role of digital communication channels coupled with the need for decoding of speech and increased emphasis on integrated voice/data networks, speech bandwidth compression system are expected to play an important role in speech communication system.

### A. Optimal short time spectral amplitude estimator:

estimator shows better improvement compared to traditional log spectral MMSE estimator. This method can be applied to signal, sound and images.

### B. Novel dual –microphone technique

The novel dual microphone technique is used to enhance the speech signal in reverberating environment. In this algorithm coherence between the input signal is estimated considering the direct, reverberant signal and noise components received at the sensors. This algorithm is capable of dealing with coherence and diffuse noise. Once the coherence is estimated, the signal to noise ratio (SNR) can be predicted by solving the quadratic equation

obtained from the real and the imaginary part of the function. Evaluating in room with reverberation time  $T_{60} = 465\text{ms}$ , show improvement in SNR and quality of processed signal compared to coherence based noise reduction algorithm. The advantage of this algorithm is, it is simple in computation and can be implemented in real time. When combined with effective noise reduction in realistic situation, this can be used in hearing aids and communication devices.

### C. Suboptimal generalized sidelobe canceller

In suboptimal method of generalized sidelobe canceller, sensor array is located in an enclosure. This array is used to reduce the interference that affects the signal. Constrained minimum power adaptive beamforming method is widely used for signal enhancement. Even though good interference suppression can be achieved, in complicated acoustic environment the arbitrary transfer function (TF) is a major problem. The generalized sidelobe canceller that adapts to arbitrary Transfer function. suboptimal algorithm estimate transfer function ratio instead of calculating the transfer function. These TF ratio are calculated based on the nonstationary characteristics of input signal. The use of TF ratio improves efficiency and robustness of algorithm. This algorithm can be implemented in both frequency and time domain. Although implementation in time domain is complex comparatively

### D. spectro-temporal filtering algorithm:

In spectro temporal filtering algorithm we consider the spectral and temporal correlation between adjacent frames in a speech signal. In multiplicative transfer function approach we need to know the transfer function of the source and the receiver in advance so that the noise in the speech signal can be reduced. Practically this is not possible as the transfer function of the receiver cannot be known in advance. This makes it difficult to eliminate the noise completely in the signal. This is the disadvantage of the multiplicative TF approach.

To overcome this disadvantage we use spectro-temporal filtering method. In this method we calculate the speech and noise power spectral densities (PSD), extended parameterized non-causal wiener filtering is applied to the speech and noise PSDs. From the resulting value speech and noise are estimated along with speech presence probabilities. This formulation can be applied to the input microphone array so that the noise will be reduced and the speech will be enhanced.

The advantage of this algorithm over multiplicative TF approach is that, the PSD of speech and noise is calculated for the whole spectrum. Whereas in multiplicative TF approach the PSD of speech and noise should be calculated for each signal separately. the overhead in the channel can be reduced by computing the PSD of speech and noise in first frame and calculating the PSD of speech and noise of adjacent frames from the value of the first frame.

## CONCLUSION

In this paper we have discussed various algorithms for speech enhancement. Each algorithm have their own advantages and disadvantages. The purpose of speech enhancement can be achieved by these algorithm based on the type of application. The purpose of enhancing the signal is to give a clear signal removed from noise so that it can be used in various applications such as in hearing aid, digital communication field, teleconferencing field. This paper gives an insight to different types of speech enhancement algorithms.

## REFERENCE

- [1] Yu Gwang Jin, Jong Won Shin, and Nam Soo Kim, "Spectro-Temporal Filtering for Multichannel Speech Enhancement in Short-Time Fourier Transform Domain" *IEEE signal processing Letters*, vol. 21, no. 3, march 2014.
- [2] L. J. Griffiths and C.W. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Trans. Antennas Propagat.*, vol. AP-30, no. 1, pp. 27–34, Jan. 1982.
- [3] S. Gannot, D. Burstein, and E. Weinstein, "Signal enhancement using beamforming and nonstationarity with applications to speech," *IEEE Trans. Signal Process.*, vol. 49, no. 8, pp. 1614–1626, Aug. 2001.
- [4] S. Gannot and I. Cohen, "Adaptive beamforming and postfiltering," in *Springer Handbook of Speech Processing*, J. Benesty, Y. Huang, and M. M. Sondhi, Eds. New York, NY, USA: Springer-Verlag, 2007, ch. 47, pp. 945–978.
- [5] M. Souden, J. Benesty, and S. Affes, "On optimal frequency-domain multichannel linear filtering for noise reduction," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 2, pp. 260–276, Feb. 2010.
- [6] M. Souden, J. Chen, J. Benesty, and S. Affes, "An integrated solution for online multichannel noise tracking and reduction," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 7, pp. 2159–2169, Sep. 2011.
- [7] Y. Avargel and I. Cohen, "System identification in the short-time Fourier transform domain with crossband filtering," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 15, no. 4, pp. 1305–1319, May 2007.
- [8] Y. Avargel and I. Cohen, "Adaptive system identification in the short-time Fourier transform domain using cross-multiplicative transfer function approximation," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 16, no. 1, pp. 162–173, Jan. 2008.
- [9] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, no. 6, pp. 1109–1121, Dec. 1984.
- [10] I. Cohen, "Noise spectrum estimation in adverse environments: Improved minima controlled recursive averaging," *IEEE Trans. Speech Audio Process.*, vol. 11, no. 5, pp. 466–475, Sep. 2003.
- [11] M. Souden, J. Chen, J. Benesty, and S. Affes, "Gaussian model-based multichannel speech presence probability," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 5, pp. 1072–1077, Jul. 2010.
- [12] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Amer.*, vol. 65, pp. 943–950, Apr. 1979.
- [13] E. A. Lehmann, "Image-source method for room acoustics," [Online]. Available: [http://www.eric-lehmann.com/ism\\_code.html](http://www.eric-lehmann.com/ism_code.html).