www.ijecs.in
International Journal Of Engineering And Computer Science ISSN:2319-7242
Volume 4 Issue 4 April 2015, Page No. 11143-11147

Speech Enhancement Using Beamforming

Dr. G. Ramesh Babu[1], D. Lavanya[2], B. Yamuna[2], H. Divya[2], B. Shiva Kumar[2], B. Ashok Kumar[2]

[1]Prof. and HOD, Dept of ECE, SSCE, Srikakulam.
[2]Under Graduate, Department of ECE, SSCE, Srikakulam.

Abstract: ***The problem of enhancing speech degraded by uncorrelated additive noise, when the noise speech alone available, has recently received much attention. Beamforming is one possible method of speech enhancement, because, the beamformer minimizes the output signal power but maintains signals from the desired direction. Beamforming techniques basically approach the problem from a spatial point of view. A microphone array is used to form a spatial filter which can extract a signal from a specific direction and reduce the contamination of signals from other directions. In this paper we survey some Beamforming techniques used for minimize the noise power in the output signal.***

Keywords: Beamforming, Spatial Filter, Additive Noise.

1. INTRODUCTION

Signal quality might significantly deteriorate in the presence of interference, especially when the signal is also subject to reverberation. Multisensor - based enhancement algorithms typically incorporate both spatial and spectral information. Hence, they have the potential to improve on single sensor solutions that utilize only spectral information. In particular, when the desired signal is speech, single microphone solutions are known to be limited in their performance. Beamforming methods have therefore attracted a great deal of interest in the past three decades. Applications of beamforming to the speech enhancement problem have also emerged recently.

Beamforming is the process of trying to concentrate the array to sounds coming from only one particular direction. Spatially, this would look like a large dumbbell shaped lobe aimed in the direction of interest. Making a Beamformer is crucial to meet one of the goals of our paper, which is to listen to sounds in one direction and ignore sounds in other directions. The best way to not listen in 'noisy' directions, is to just steer all your energy towards listening in one direction.

2. SPEECH ENHANCEMENT

Speech enhancement aims to improve speech quality by using various algorithms. The central methods for enhancing speech are the removal of background noise, echo suppression and the process of artificially bringing certain frequencies into the speech signal. When the background noise is suppressed, it is crucial not to harm or garble the speech signal.

Multi-channel enhancement algorithms [1], [2] and [3] exploit the spatial diversity. This diversity can be taken advantage of e.g., by steering a null towards the noise source and a beam towards the signal source. In this paper, a brief overview of one common multi-channel noise reduction technique known as beamforming technique is provided.

3. BEAMFORMING

A Beamformer is a signal processor used together with a microphone array to provide the capability of spatial filtering. The microphone array produces spatial samples of the propagating wave, which are then manipulated by the signal processor to produce the Beamformer output signal. Beamforming is accomplished by filtering the microphone signals and combining the outputs to extract (by constructive combining) the desired signal and reject (by destructive combining) interfering signals according to their spatial location.

Beamforming can separate sources with overlapping frequency content that originate at different spatial locations. Beamforming is a means of performing spatial filtering [4]. In the frequency domain, beamforming can be viewed as a linear combination of the sensor outputs:

$$Z(k) = \sum_{i=1}^{M} b_i(k) y_i(k) \text{ ----------------(1)}$$

$b_i(k)$ is the beamformer weight corresponding to the i^{th} sensor, and M is the total number of sensors. In vector notation, we have

$$Z(k) = b^T(k)Y(k) \text{-------------------------(2)}$$

where $b(k) = [b_1(k) \ldots b_M(k)]^T$

Beamforming can be classified into two categories fixed, where the weights are fixed across time, and adaptive, where the weights vary in response to changes in the acoustic environment.

FIXED BEAMFORMING

In fixed beamforming, the weights bi(k) are fixed over time, and are determined by minimizing the power of the signal at the output of the beamformer subject to a constraint that ensures that the desired signal is undistorted. The weights multiplication and output clearly shown in the Figure 1.The main purpose of weights multiplication is to enhance the signal strength. The basic objective of a beam former is to adjust the complex weights at the output of each array element so as to produce a pattern that optimizes the reception of a target signal along the direction of interest, in some statistical sense.

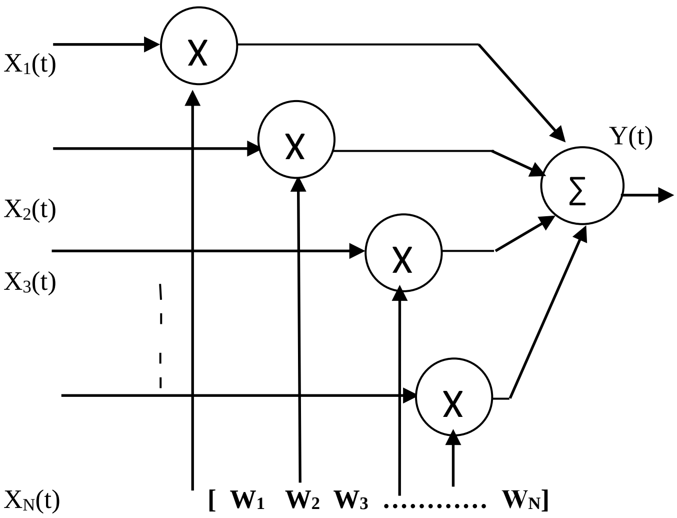

Figure 1: A Fixed Weight Beam Forming

From the figure the fixed beamforming output can be written in the mathematical form is given as,

$$Y(t)=W^TX(t) \text{ ----------------------------------(3)}$$

where,

X(t) is input signal matrix

W^T is transpose of weights

Y(t) is output signal matrix

The fixed beamforming can be obtained by using the Delay and Sum algorithm.

3.1.1 DELAY AND SUM BEMFORMER

In delay-and-sum beamforming, delays are inserted after each microphone to compensate for the arrival time differences of the speech signal to each microphone (Figure 3-1). The time aligned signals at the outputs of the delays are then summed together. This has the effect of reinforcing the desired speech signal while the unwanted off-axis noise signals are combined in a more unpredictable fashion.

The major disadvantage of delay-and-sum beamforming systems is the large number of sensors required to improve the SNR. Each doubling of the number of sensors will provide at most an additional 3 dB increase in SNR, and this is if the incoming jamming signals are completely uncorrelated between the sensors and with the desired signal. Another disadvantage is that no nulls are placed directly in jamming signal locations.

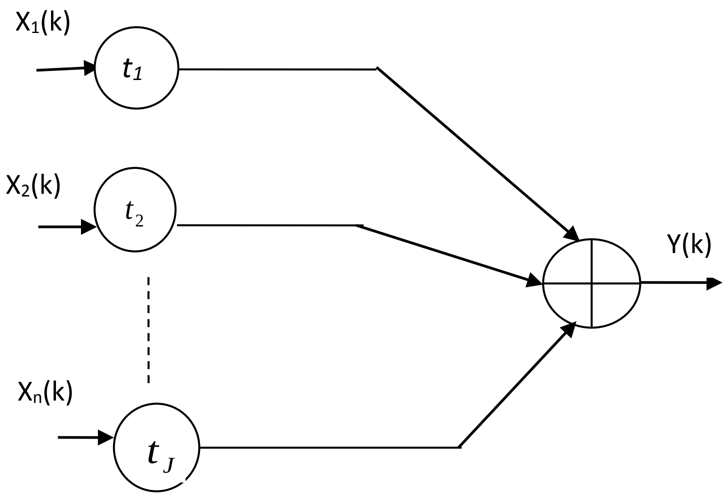

Figure 2: Delay and Sum Beamformer

The delay-and-sum Beamformer seeks only to enhance the signal in the direction to which the array is currently steered.

3.2 ADAPTIVE BEAMFORMING

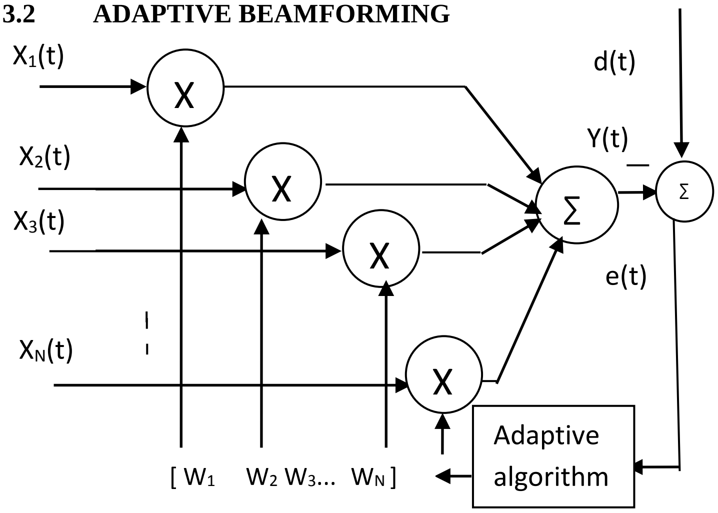

Figure 3: Adaptive Beam Forming Block Diagram

In adaptive beamforming[6][7], the beamformer weights adapt to changes in the acoustic environment over time. The optimal weights are obtained by minimizing the variance of the output signal. Here we are taking the reference signal d(t), the reference signal is used to calculate the error signal. The error signal is difference between the reference signal and output signal.

From the figure above the output response of the uniform linear array is given as,

$$e(t)=d(t)-W^TX(t) \text{ ------------------------------(4)}$$

where,

e(t) is error signal matrix
d(t) is reference signal matrix.

To ensure that the speech signal is not cancelled out or distorted, a distortion less constraint is imposed on the desired signal. This results in the linearly constrained minimum variance (LCMV) beamformer, where the adaptive beamformer weights are obtained through a constrained minimization procedure. The generalized side lobe canceller (GSC)[5] is an efficient alternative implementation of Frost's LCMV approach, that converts the constrained optimization problem into an unconstrained one. This leads to an efficient implementation for the update of the beamformer weights. The weights are adjusted based on the error signal, until the error signal will be zero. By this process we can get the clear output. The adaptive beamforming can be obtained by using Griffiths and Jim proposed one algorithm.

3.2.1 Griffiths and Jim Beamfomer algorithm

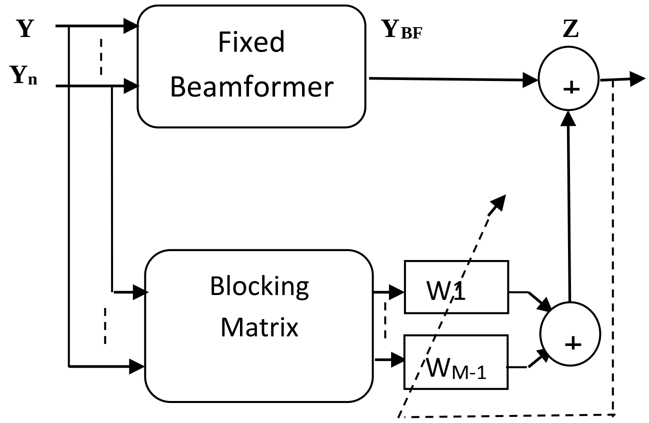

Figure 4: Frequency domain implementation of the Generalized Side lobe Canceller. The ANC is implemented by the adaptive filters $w_1 \ldots . w_{M-1}$

The GSC consists of three parts - a fixed beamformer (FBF), a blocking matrix (BM) and an adaptive noise canceller (ANC) as shown in Figure.2. The FBF includes a pre-steering module and its weights are designed to produce a speech reference YBF with a specified gain and phase response. The FBF could either be a simple delay-and-sum beamformer, or a more advanced filter-and-sum or super directive beamformer. The BM is generally orthogonal to the FBF and produces M-1 outputs, called the noise references, by steering zeros towards the desired signal direction. One way to create the noise references is to take the difference between adjacent sensor signals. The ANC (implemented by the adaptive filters w_1 . . . . w_{M-1}) in Figure 2 removes any remaining correlation between the speech reference YBF and the noise references. Thus, any residual noise in the speech reference that is correlated to the noise references is removed. In practice, the noise references are not completely free of speech. As a consequence, the ANC results in some of the speech signal being cancelled. To minimize the effect of the speech leakage on the ANC, the noise-cancelling filters are adapted only during periods of speech absence. To reduce the amount of speech leakage, some variants of the GSC employ an adaptive blocking matrix.

4. Results & Discussion

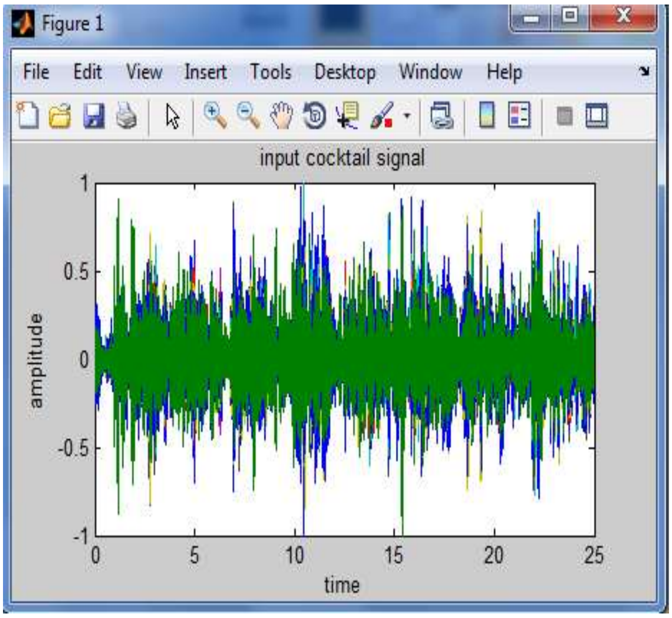

Figure 5: Array of Speech Signals - Cocktail

The array of input speech signals which is called as a Cocktail Signal is as shown in figure 5 above. And this cocktail signal is applied to the Delay-Sum Beamformer, where in this case the speech signals are multiplied with the weight coefficients in order to obtain the required original speech signal to be enhanced. Here there will be no reduction of noise. And the enhanced output speech signal is as shown in the figure 6 below.

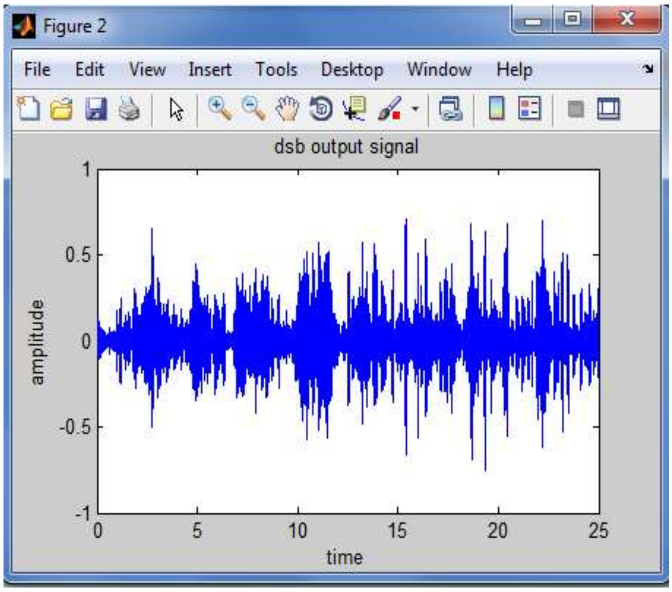

Figure 6: Delay-Sum Beamformer output

And the same input array of speech signals are applied to the Generalized Side Lobe Canceller, where in this case the Fixed Beamformer outputs are differed with the array multiplied with the adaptive weight coefficients, which is the error signal and so that the noise signal i.e., present in the side lobes are removed, and the required signal be enhanced and is as shown in the figure 7 below.

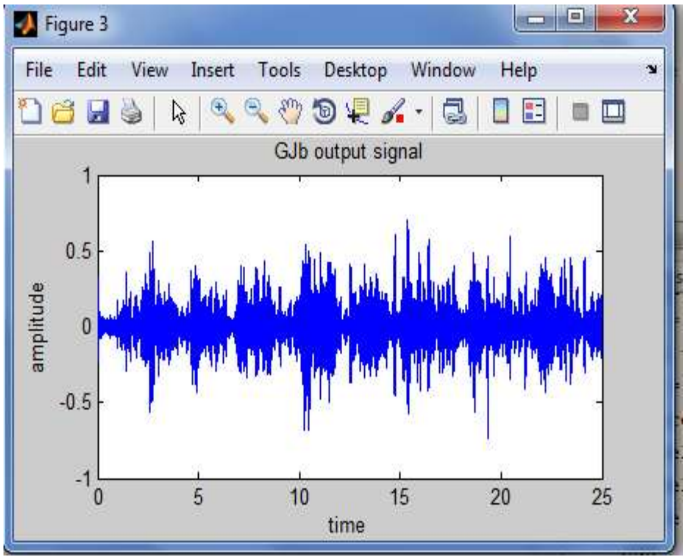

Figure 7: Griffths-Jim Beamformer output

Conclusion

The beamforming technique can be applied for enhancing an arbitrary non stationary signal corrupted by stationary noise. Although our algorithm was implemented in the frequency domain, it can also be implemented in the time domain. This applies both to the Adaptive Beamformer stage and to the system identification stage. Both versions of the algorithm yield comparable performance. However, the computational burden of the frequency domain algorithm is significantly smaller than that of the time domain version.

References

[1] M. Drews, “Speaker localization and its application to time delay estimators for multi-microphone speech enhancement systems”, *Proc. Eusipco,* 1996, pp. 483-486.

[2] M. Drews, “Construction of microphone arrays for the optimization of multi-channel speech enhancement systems”, *Frequenz* 50, 1996, 223-227 (in German).

[3] Brandstein, M.S., D.B. Ward (Eds.), *Microphone Arrays: Signal Processing Techniques and Applications*, Springer, Berlin, 2001.

[4] B.D. Van Veen, and K. M. Buckley, “Beamforming: a versatile approach to spatial filtering," *IEEE ASSP Magazine*, vol. 5, no. 2, 1988, pp. 4-24.

[5] B. R. Breed, and J. Strauss, “A short proof of the equivalence of LCMV and GSC beamforming," *IEEE Signal Processing Lett*., vol. 9, no. 6, 2002, pp. 168-169.

[6] “Robust adaptive beamforming via target tracking,” *IEEE Trans. Signal Processing*, vol. 44, pp. 1589–1593, June 1996.

[7] O. Hoshuyama and A. Sugiyama, “A robust adaptive beamformer for microphone arrays with a blocking matrix using constrained adaptive filters,” in *Proc. Int. Conf. Acoustics, Speech, Signal Process.*, Atlanta, GA, May 1996, pp. 925–928.