SPEECH ENHANCEMENT USING GENERALISED SIDE LOBE CANCELLER

MATTAPARTHI SAI VENKATA AKSHAY
SATYANARAYANA NAMUDURI
AMIRISETTI SRAVYA

ABSTRACT

The main aim of this project is to provide an efficient Speech Enhancement technique using a Generalized Side-lobe Canceller(GSC)[1][2]. Speech is a broadband signal. It is affected by different kinds of noise. Some of the noises are thermal noise, pink noise, white noise, etc. Speech can be enhanced in many ways. One of them is to use a GSC. A GSC consists of a Fixed Beam Former(FBF) whose main lobe is directed towards the speaker, but there exist some side lobes in the beam pattern of the FBF that collects the noise. And then the noise is isolated and cancelled with the help of a blocking filter and an adaptive interference canceller(AIC).In this project we implement a broadside broadband beam former, using a delay and sum beam-former, the Jim and Griffith blocking matrix[3].

CONTENTS

List of Figures

Chapter 1

INRODUCTION

Speech is a broadband signal. It is mainly affected by noise which is correlated with it. Speech enhancement is a signal processing technique in which the quality of a desired signal is enhanced. Speech Enhancement aims to improve speech quality by using various algorithms. The speech enhancement can be done in many ways. Two of the techniques we use are

- Single Channel Speech Enhancement
- Multi Channel Speech Enhancement

Single channel:

Single channel speech enhancement aims at enhancing the quality of a speech signal based on the signal itself.

Multi channel:

Where as in multi channel speech enhancement, for example in beam-forming we use an array of microphones, for acquisition and enhancement of the speech signals.

Beam-forming or spatial filtering is a signal processing technique in which a set of sensors as an array is used for directional transmission or reception [4]. Beam-forming can be divided into many types based on different parameters taken into consideration.

Types of beam formers:

Based on range of frequencies:

- Narrowband
- Broadband

Based on the type of spatial filters used:

- Fixed Beam-former
- Partially Adaptive Beam-former
- Fully Adaptive Beam-former

A Fixed Beam-former(FBF) is a microphone array calibrated for spatially filtering in a particular direction, and does not change the direction of main lobe based on the movement of the source, where as an adaptive beam-former updates the weights of the spatial filters to maintain the direction of the main lobe in the direction of the source for having an optimal signal to noise ratio(SNR). The weights of the spatial filters are updated based on adaptive filtering algorithms such as Least mean squares filter(LMS), Normalized least mean squares filter(NLMS), etc.

1.1 Problem Formulation

Consider a uniform linear microphone array in a noisy environment. A Speech source is considered to be present on the broadside of the array and located along the axis of the array. A noise source is considered to be present in the end- re direction of the array and also stationary. The speech signal is to be enhanced by reducing the background noise.

Chapter 2

Problem Solution

Speech Enhancement can be achieved in a number of ways. One of them is through the use of a Generalized Side-lobe Canceller(GSC). A GSC is an advanced spatial filtering algorithm in which the direction of the source is known, which is not same as that of the noise. A GSC consists of a Fixed beam-former, in parallel with a blocking filter and an adaptive interference canceller. A voice activity detector(VAD) is incorporated into the system in order to detect the voice activity of the system.

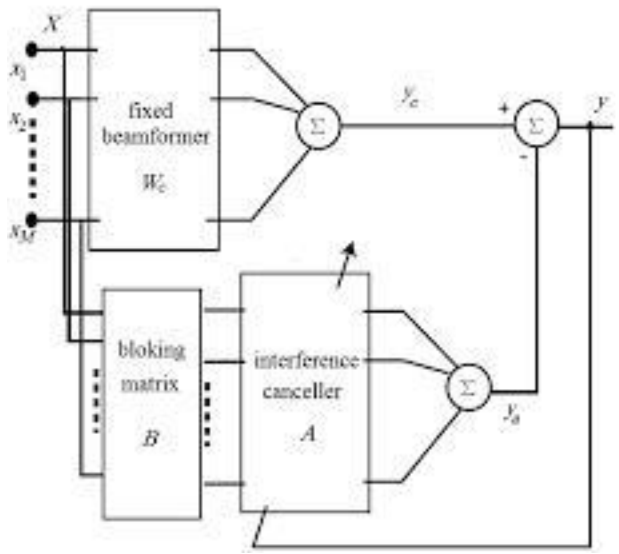

Figure 2.1: Block diagram of a Generalized Side-lobe Canceller

The Fixed beam-former is considered to be delay and sum(DAS) beam-former. The blocking filter is considered to be the Jim and Griffith blocking matrix. The adaptive interference canceller is an NLMS adaptive filter. Source of the speech is considered to be in far field with respect to the microphone array.

$$Wn+1=Wn-\mu(n)e(n)x^*(n) \tag{2.1}$$

where $\mu(n)$ is

$$\mu(n)=x(n)x^*(n) \tag{2.2}$$

$$
\mathbf{B} = \begin{bmatrix} 1 & -1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & 0 & 0 & \\ 0 & 0 & 0 & 0 & \cdots & 1 & -1 \end{bmatrix}
$$

Figure 2.2: The Jim and Griffith Blocking Matrix

The fixed beam-former spatially selects the speech signal coming from the broadside of the microphone array and the blocking filter isolates the noise signal coming from the end-fire direction of the microphone array. This Isolated noise is then used in the cancellation of the side-lobes present in the gain pattern of the microphone array with the help of an adaptive filter. Weights of the adaptive filter are updated using fixed beam-former output and the voice activity detector output.

Chapter 3

Implementation

The Implementation of this project was accomplished in two stages.

- Matlab Implementation
- Real Time Implementation on a DSP

3.1 Matlab Implementation

While implementing the GSC on MATLAB some theoretical assumptions were made. They are as follows:

- Number of Microphones was assumed to be two.
- Speech source was considered to be in the broadside of the microphone array located along the axis.
- Noise was assumed to come from the end- re directions.
- A block size of 32 samples was assumed with a sampling frequency of 8KHz.

Based on these assumptions a delay and sum beam-former was designed. The blocking filter was assumed to be the Jim and Griffith blocking matrix[4]. This was designed using a set of FIR filters. An adaptive interference canceller was designed using Normalized Least Mean Squares filtering algorithm. A Voice activity detection algorithm was incorporated to decide whether there was voice activity in a block of samples or not, based on which the weights of the NLMS filter[6] were updated.

Input to the system was a noisy speech signal generated by adding two pre-recorded clean speech and noise signals. First the voice activity detector decides whether there exists speech in a block of data or not, followed by a fixed beam-former which equalizes the delay between the two channels of the input and sums them up to give the fixed beam-former output. This acts as a reference signal to the Adaptive Interference Canceller(AIC).

Blocking matrix is fed with the input signals coming into the microphone array, and the output is the signals that are spatially filtered which are not from the desired speech direction. This acts as the input to the AIC. The weights of the NLMS filter in the AIC are updated when the output of the VAD is '0', i.e. when there exists no speech in the input data. This helps us to train the AIC to filter out noise and not the speech. Output of the AIC is then subtracted from the fixed beam-former output and is given out as the final output of the GSC algorithm. based on this output the weights of the NLMS filter are updated for the next block of data. This whole process is carried out for total signal block by block.

3.2 Implementation on the DSP

The real-time implementation of this project was achieved in the Visual DSP++ environment. The programming of the DSP processor was written in C language. The block size for the system was assumed to be 32 samples and the sampling frequency was set to 8 Khz.

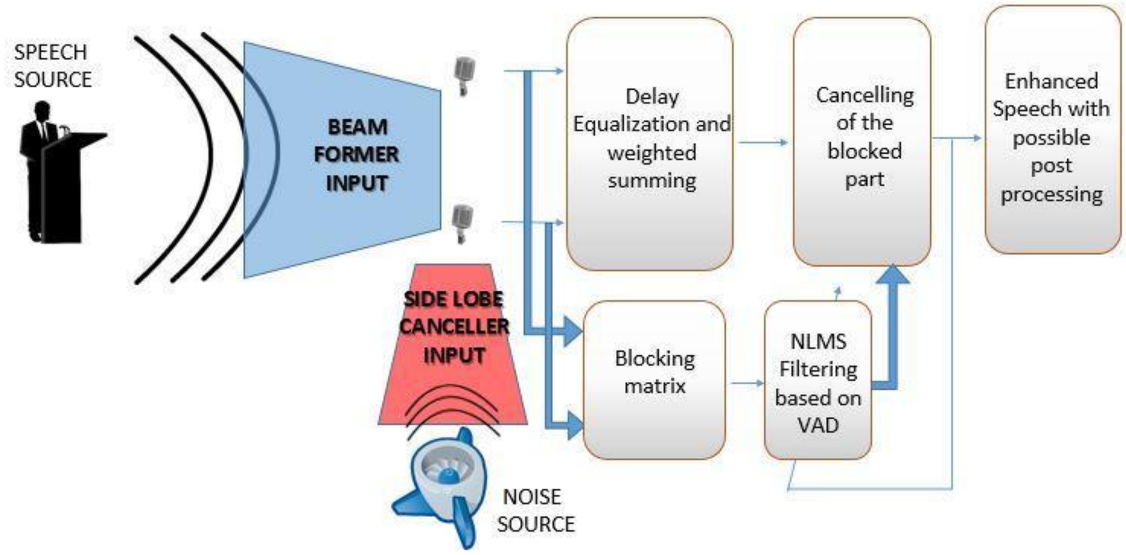

Figure 3.1: Block diagram of the Generalized Side-lobe Canceller

The input was given through an AUX cable to the DSP processor. The blocks of input were acquired using an audio interrupt to the DSP Processor. The whole operation of GSC was written in a function called process. The Process function was initiated whenever there was an audio interrupt. A pair of push button were used for enabling and disabling the GSC algorithm whenever necessary.

The final NLMS weights from the MATLAB implementation were used as an FIR filter instead of the adaptive NLMS filter and so no feedback was given and VAD was not used while implementing the algorithm in real-time. This was done because there was a stability problem in the NLMS filter in real-time implementation.

Chapter 4

Results

Matlab Results:

Beamformer output using different arrays of microphones:

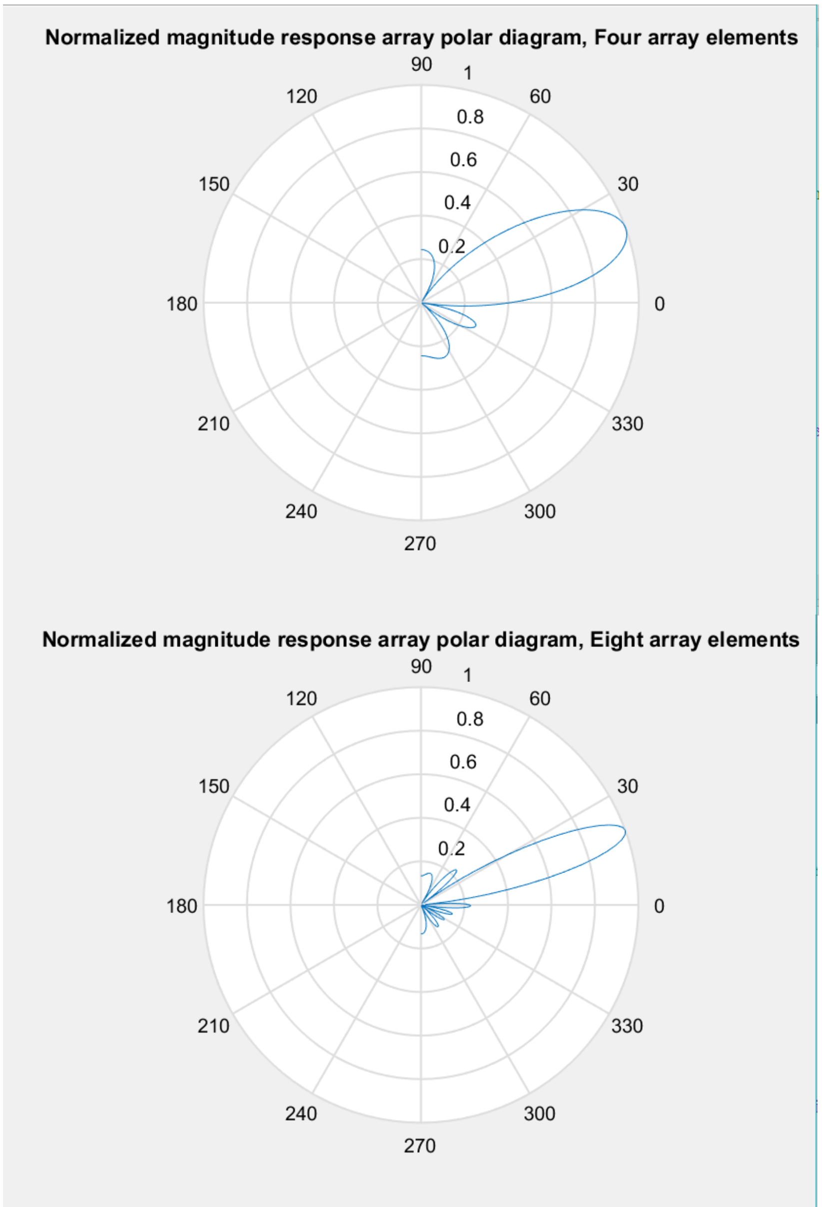

Fig 4.1.1: Beamforming output for 4 and 8 microphones

Normalized output for different microphone arrays:

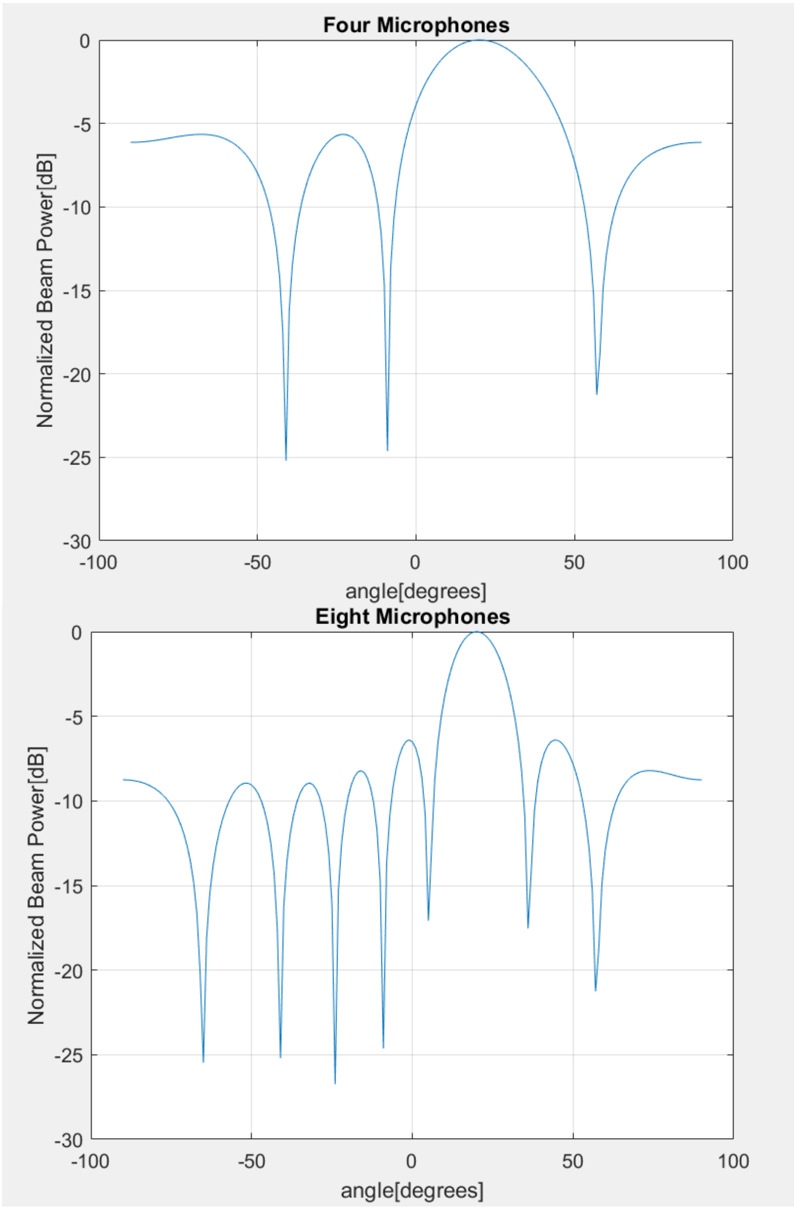

Fig 4.1.2: Normalized output for 4 and 8 microphones

Voice Activity Detector Output:

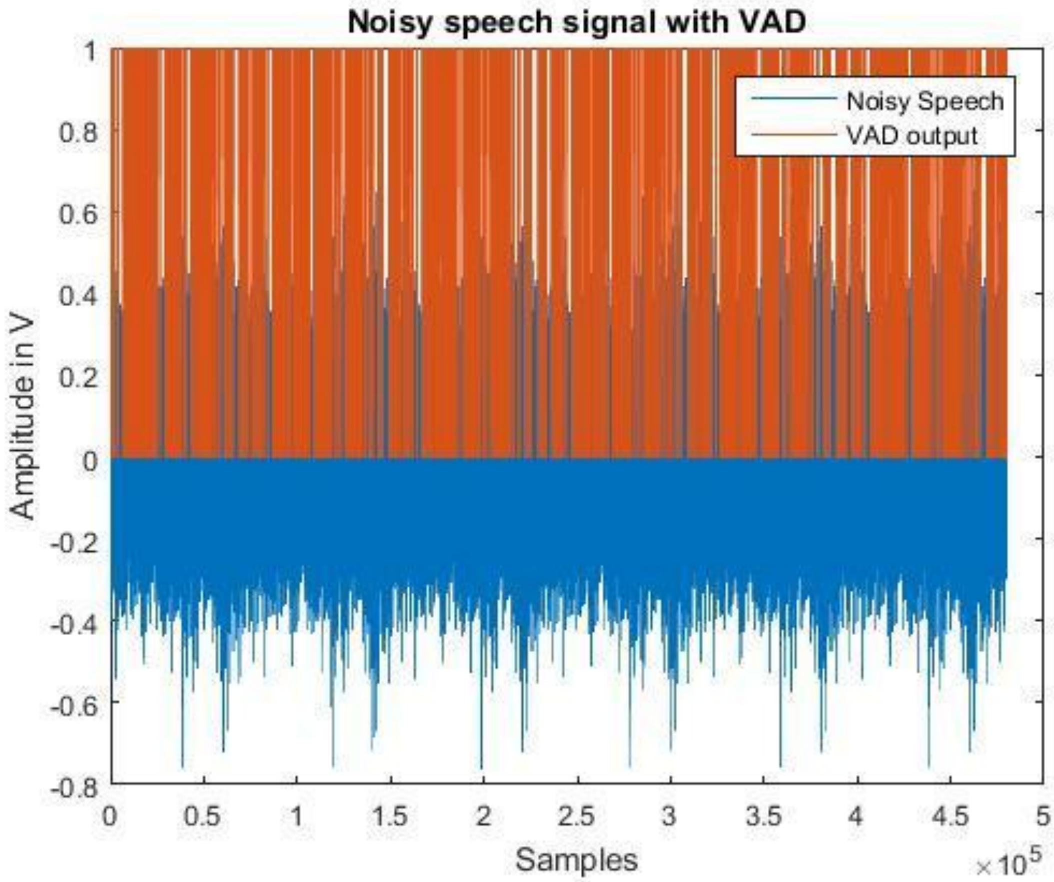

Figure 4.1.3: Voice Activity Detector Output

The enhanced speech output:

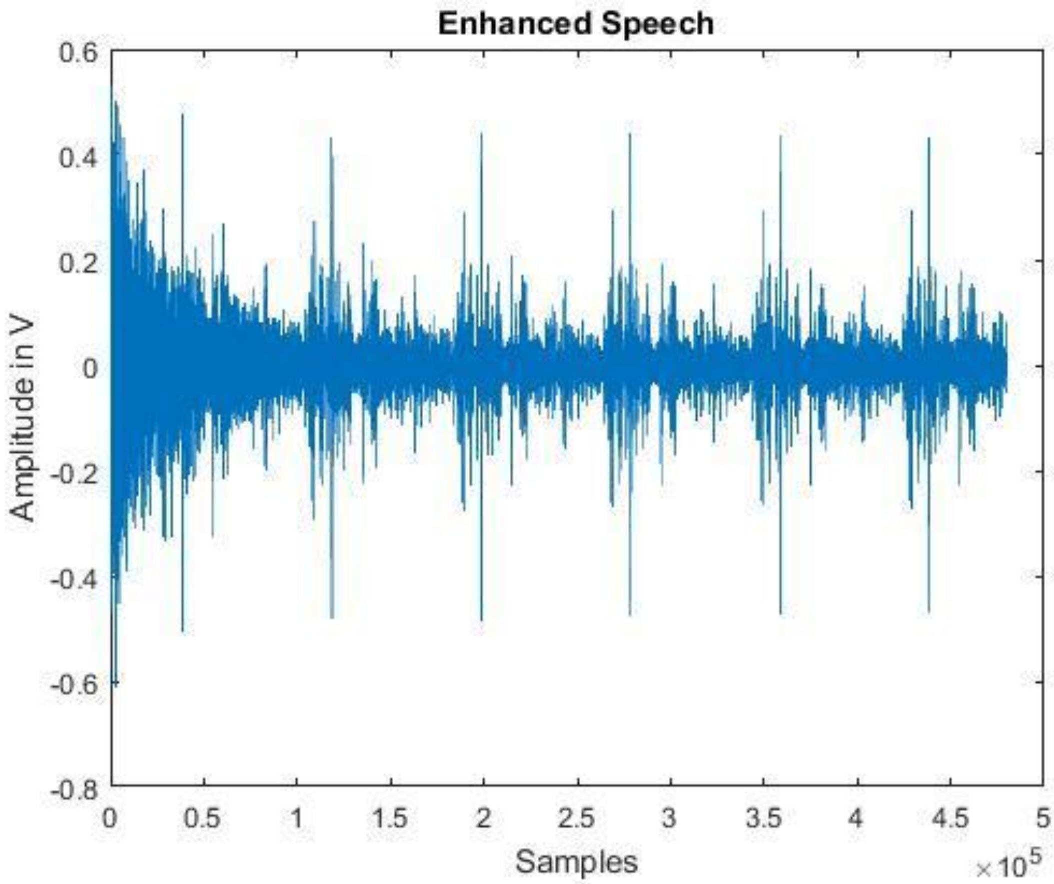

Figure 4.1.4: Enhanced Speech output of the system

The Input Signal to Noise ratio of the system was measured to be 12:1832dB and the Output Signal to Noise ratio was measured to be 1:9412dB. From this we could observe that the improvement in the SNR was 14:1244dB.

4.2 DSP Implementation Results

Simulating the MATLAB Equivalent implementation on a DSP simulator using the Visual DSP++ environment, by giving a block of data input from Matlab, it was ob-served that the outputs of both the Matlab and Visual DSP++ simulations are in good agreement.

```
E[0] = 0.301700
E[1] = 0.307877
E[2] = 0.308285
E[3] = 0.305238
E[4] = 0.298028
E[5] = 0.283623
E[6] = 0.259508
E[7] = 0.225134
E[8] = 0.185460
E[9] = 0.138767
E[10] = 0.087920
E[11] = 0.036648
E[12] = -0.014722
E[13] = -0.064246
E[14] = -0.117096
E[15] = -0.169508
E[16] = -0.218535
E[17] = -0.261705
E[18] = -0.296628
E[19] = -0.319253
E[20] = -0.326302
E[21] = -0.319592
E[22] = -0.303620
E[23] = -0.281931
E[24] = -0.256456
E[25] = -0.231427
E[26] = -0.207867
E[27] = -0.189141
E[28] = -0.176526
E[29] = -0.168910
E[30] = -0.168353
E[31] = -0.176704
```

Figure 4.2.1: Enhanced Speech output of the system using Visual DSP++ simulator

```
Command Window
>> E(1,:)

ans =

  Columns 1 through 16

    0.3017    0.3108    0.3129    0.3109    0.3043    0.2904    0.2664    0.2318    0.1916    0.1440    0.0922    0.0396   -0.0131   -0.0640   -0.1183   -0.1722

  Columns 17 through 32

   -0.2228   -0.2673   -0.3036   -0.3273   -0.3351   -0.3287   -0.3128   -0.2909   -0.2650   -0.2395   -0.2153   -0.1959   -0.1826   -0.1746   -0.1736   -0.1819
```

Figure 4.2.2: Enhanced Speech output of the system using MATLAB

4.3 Problems in Implementation:

While hearing to the audio output of the system at different stages, it was observed that the blocking filter was not as efficient as required and the desired speech signal was still present in the blocking matrix output, which depicts a signal leakage in the system, leading to a distortion in the enhanced speech output.

Chapter 5

Conclusions & Scope for Further work

Hence a Generalized Side-lobe Canceller has been simulated in MATLAB and Visual DSP++ and implemented in real-time on a DSP Processor. The Improvement in the Signal to Noise ratio between the input and output was observed to be around 14dB. It was observed that there exists a signal leakage in the Broadband Generalized side lobe Canceller that leads to a distortion in the Enhanced output.

Performance of the system can further be improved by using more number of microphones in the array and different array configurations. Blocking output can be further improved by the usage of adaptive blocking filters.

Bibliography

[1] M. Ahmed, "Adaptive sub band gsc beam forming using linear microphone-array for noise reduction/speech enhancement.."

[2] P. Khayeri, H. Abutalebi, and V. Abootalebi, "A nested superdirective generalized sidelobe canceller for speech enhancement," in Information, Communications and Signal Processing (ICICS) 2011 8th International Conference on, pp. 1-5, Dec 2011.

[3] L. J. Griffiths and C. Jim, "An alternative approach to linearly constrained adaptive beamforming," Antennas and Propagation, IEEE Transactions on, vol. 30, pp. 27-34, Jan 1982.

[4] "Beamforming," Nov. 2015. Page Version ID: 688881500.