MICROPHONE ARRAY SYSTEM FOR SPEECH ENHANCEMENT IN LAPTOPS

A major project report submitted in partial fulfilment of the requirements for the award of the degree of

MASTER OF SCIENCE
IN
ELECTRICAL ENGINEERING WITH EMPHASIS ON SIGNAL PROCESSING

By

NAVEEN KUMAR THUPALLI
nath10@student.bth.se
19880406-3652

Under the esteemed guidance of

NEDELKO GRBIC

Department of Electrical Engineering
Blekinge Institute of Technology,
SE-371 79, Karlskrona, Sweden.

2011-2012

Contact Information:

Author:

Mr. Naveen Kumar Thupalli
E-mail: naveenkumarthupalli@gmail.com

Supervisor:

Dr. Nedelko Grbic
School of Engineering (ING)
E-mail: nedelko.grbic@bth.se
Phone: +46455385727

School of Engineering
Blekinge Institute of Technology,
Karlskrona-37179,
Sweden.

Internet: www.bth.se/ing
Phone : +46455385000
Fax : +46455385057

Abstract:

Recognition of speech at the receiver end generally gets degraded in distant talking atmospheres of laptops, teleconfereing, video conferences and in hands free telephony, where the quality of speech gets contaminated and severely disturbed because of the additive noises. To make useful and effective, the exact speech signals has to be extracted from the noise signals and the user has to be given the clean speech. In such conditions the convenience of microphone array has been preferred as a means of civilizing the quality of arrested signals. A consequential growth in laptop technology and microphone array processing have made possible to improve intelligibility of speech while communication. So this contention target on reducing the additive noises from the original speech, beside design and use of different algorithms. In this thesis a multi-channel microphone array with its speech enhancement of signals to Wiener Beamformar and Generalized side lobe canceller (GSC) are used for Laptops in a noisy environment.

Systems prescribed above were implemented, processed and evaluated on a computer using Mat lab considering SNR, SNRI as the main objective of quality measures. Systems were tested with two speech signals, among which one is Main speech signal and other is considered as Noise along with another random noise, sampling them at 16 KHz .Three Different source originations were taken into consideration with different input SNR’s of 0dB, 5dB, 10dB, 15dB, 20dB, 25dB.

Simulation Results showed that Noise is been attenuated to a great extent. But Variations in SNR and SNRI has been observed, because of the different point origination of signals in the respective feilds.Variation in SNR and SNRI is been observed when the distance between the main speech originating point and microphone is too long compared to the noise signals. This states that origination of signals plays a huge role in maintaining the speech quality at the receiver end.

Acknowledgements

I wish to thank, first and foremost, My professor and supervisor **Dr.Nedelko Grbic** who has attitude and the substance of a genius .He encouraged and guided me continually in each and every step of this thesis work and an showed an excitement in regard to teaching. Without his guidance and constant help this dissertation would not have been possible. I consider it an honour to work under him.

I owe my deepest gratitude to Sridhar Bitra who spent time for me and gave a lot of support during the thesis work.

I also want to thank my coordinate classmates Jeevan reddy Yarraguddi, Pardhasaradhi reddy for their kind support during the thesis work. I express my gratitude to all the authorities, faculty members of BTH, karlskrona

I offer my dutiful respect and love to my grandparents, parents, friends for their marvellous support and encouragement throughout my study period. I owe to them throughout my life time.

-Naveen Kumar Thupalli

TABLE OF CONTENTS

List of Figures:

List of Tables:

List of Abbreviations:

ANC	Active Noise Cancellation
DSB	Delay and Sum Beam forming
FD	Fractional Delay
GSC	Generalized Side lobe Canceller
PSD	Power Spectral Density
SNR	Signal to Noise Ratio
SNRI	Signal to Noise Ratio Improvement
WBF	Wiener Beam Former
LMS	Least Mean Square
NLMS	Normalized Least Mean Square
TDE	Time delay estimates
ULA	Uniform linear array
FIR	Finite Impulse Response
LCMV	Linear constraint Minimum Variance
BSS	Blind Source separation

Introduction and motivation

Chapter 1

1.1 Introduction and Motivation:

Most of the speech recognition systems behave fairly well in the noise free circumstances using a close talking microphone worn near the mouth of the speaker. Excessive growth of such systems in the modern day made them expandable to variety of uses. As the performance of these kinds of system's increases it eventually demands for these types of systems in several applications. However many of the target applications for this technology do not take place in noise free environments. To further compound this problem it is often inconvenient for the speaker to wear a close talking microphone. As the distance between speaker and micro phone increases degradation of speech takes place because of the back ground noise and interferences. This is especially problematic in situations where the locations of the microphones or the users dictated by physical constraints of the operating environment as in meetings or automobiles.

A common example that is used for demonstrating the extraordinary abilities of the system is by well-known cocktail party effect [25]. The cocktail party effect describes the ability to focus one's listening attention on a single talker among a mixture of conversations and background noises, ignoring other conversations. Concretely, and this is the reason for the name of the mentioned effect, this phenomenon is easily appreciated in a crowded party. In this context, it is able to maintain a fluent conversation with another person, even when other interfering conversations are happening around us or when close to loudspeakers playing music. In other words, it is able to focus our attention just to the desired ”target” and to consider the rest of the interfering sources as background noise that can be ignored. In this case, the hearing reaches a high noise suppression of the background noise and enhances our focus of attention. In such a situation, a microphone recording placed exactly in our position will show a big difference.

This problem can be greatly alleviated or decreased by the use of multiple microphones to grab the speech signal. Microphone arrays record the speech signal simultaneously over a number of spatially separated channels. Many array-signal-processing techniques have been developed to combine the signals in the array to achieve a substantial improvement in the signal-to-noise ratio (SNR) of the output signal.

One field of growing interest to reduce problems introduced by distant microphone recordings consists in taking advantage of the multi-microphone availability. More concretely, microphone array processing has been broadly investigated as a pre-processing stage in order to enhance the recorded signal that might be used for any speech application.

Unfortunately, most of the speech enhancement techniques based on multi-microphone processing relies on one fundamental cue that is mostly unknown: the source position. The need for reliable target position estimation in the beam forming applications is one of the reasons for the increasing interest in the acoustic source localization and tracking topic. Furthermore, accurate knowledge of the position of the events or the speakers present in a room is also useful. A simple wiener Beam forming concept is implemented in this work.

As the beam forming solution is inadequate to achieve desired results, a generalized side lobe canceller (GSC) is developed which incorporates a beam former. The side lobe canceller is evaluated using both LMS adaptation [26].

The motivation behind this thesis is in detecting and executing an exclusive system that is compound of various enhancement algorithms to make an advantage of strengths in each algorithm to work combine, to obtain the goal in attenuating noise adequately with conservation of speech quality and its accuracy. Motivation for supporting this assertion could vary extremely and it is not the object for this argument to listen them or to argue in favour or against of some of the reasons. In fact the intention is to discuss only one answer.

1.2 Outline:

This thesis report is fresh constructed in seven definite chapters accommodating two or more subdivisions with in a chapter. It is submitted as part of Double Degree in Master of Technology (M-Tech) and Master of Science (MSc) in Electrical Engineering with Emphasis on Signal Processing.

Chapter 1

The chapter-1 deals with the introduction for the overall thesis work, motivation and constituting of six sections which are described briefly about the respective chapters.

Chapter 2

The chapter-2 deals with the Microphone array set up and essential approaches for analyzing the thesis work. It is portioned into four sections, section-2.1 answers the basic introduction to Microphone arrays and it is followed the section-2.2 involves what exactly is microphone array and why it is needed and the solution for it And it is followed by 2.3 which contains the working procedure for the microphone array and followed by 2.4 which contains the properties of the arrays. Section 2.5 contains the microphone array geometry the contemplation in designing the microphone arrays such as spacing between microphones, source filed i.e. near or far filed. The spacing is prescribed by spatial sampling theorem to avoid spatial aliasing. In section-2.6 and 2.7 fractional delays filter is discussed and designed to generate a signal having non-integer delay and fractional delay. The sinc-windowing filter is designed.

Chapter 3

In chapter-3, Section-3.1 totally deals with beam forming basic and introduction to it and its types, Section-3.2 is described about speech extraction by array processing and also discuss about how beam forming is constructed, section 3.3 finally explains about Wiener Beam Former mathematically with equations..

Chapter 4

Chapter -4 deals with the speech enhancement methods to the GSC in brief and generalized side lobe canceller and LMS are discussed with equations mathematically.

Chapter 5

The chapter-5 deals with implementation issues of microphone array, sinc windowing FD filter, and the three speech enhancement techniques i.e. Wiener Beam Former, Generalized side lobe canceller, Signal to Noise Ratio.

Chapter 6

In chapter-6, the implemented systems are evaluated to attenuate noise with different objective measures such SNR SNRI, and analyzing of simulation results were done.

‘Chapter 7

The chapter-7 explored with conclusions of different enhancement systems in attenuating the fan noise and also the future work is suggested.

Chapter -8 is all about References.

CHAPTER 2
MICROPHONE ARRAY SETUP

2.1 Introduction:

It is not an easy task of providing a dense audio capture experience. It requires a comprehensive way that takes into examination the entire life of audio signal. A fault at any one point in the avenue of the signal results in degradation of the signal at the receiver end. It is caused by the interference from components inside the laptop itself or may be have back ground noises. By considering all these, may find that signal that is entering into the microphone may not be good to start out with. When both physical interferences and background noises mix each other the chances of a high quality signal at the receiver end can be bleak. The solution for this is by using system equipped by a microphone array, and then the results may be dramatic.

2.2 So what exactly is a microphone array?

Simply stated, array of microphones is just like a normal microphone but instead of having one microphone will have them in multiple to record the input signals. Microphones in the array work combine in a balanced way to record the sound simultaneously. The big advantage of using one or more microphones is that it helps in determining the position of the sound source in the room by allowing the software that is processing the microphone signals. This is achieved by analyzing the arrival times of the sound to each of the microphones in array. For example if the sound arrives into the microphone on the right before it enters the microphone on left, then it comes to know that sound source is to right of the system. During the capturing of sound, the microphone array software searches for the sound source and aims at making a beam in that direction. If the concerned sound source moves the capture beam will follow it eventually. It’s like having two high directional microphones one being scanning the workspace measuring the sound level and other being pointing out to the direction with highest sound level i.e. is to the source of the sound. In addition to this the huge directivity of the microphone array reduces the surrounding noises and reverberation which results in the much clearer representation of the speaker’s voice. A general layout of the microphone array is shown in the Fig.1

2.3 How does it work’s?

In any direct or indirect form Microphone array processing consists of two main procedures one is sound source localization and beam former. First once helps in finding where the sound source is and should work certainly under reverberation and noisy conditions and tells the beam former where to focus the microphone array “beam”. There are many ways in finding the direction of the sound source’s and among them one is Time delay estimates (TDE) based methods which uses the facts that the sound source’s reaches the microphones at different times. Delays are easily calculated using the cross correlation function between the signals from the different microphones. Another method is to steer the beam and to measure the direction based on the maximum output signal. This method gives the similar results as that of the time delay estimate

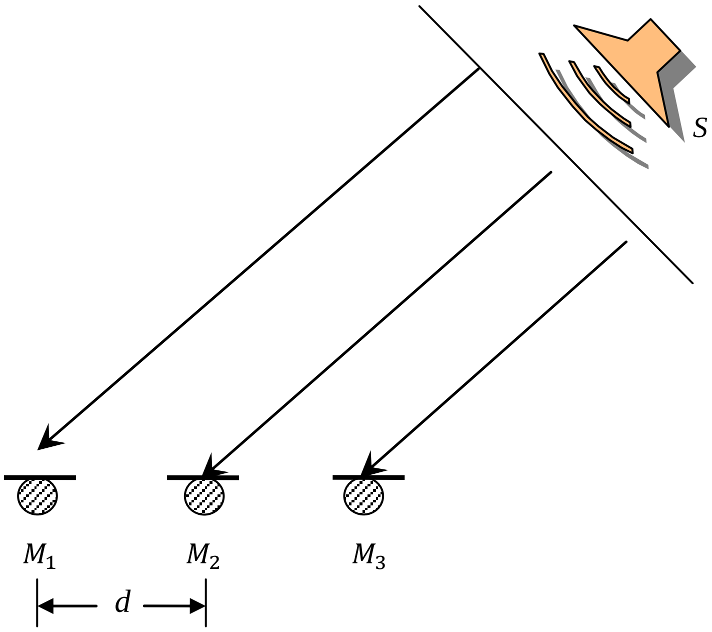


Figure 1 Microphone array layout

2.4 Array properties:

Microphones used in laptops generally are 20 to 30 centimeters away from sound source. Apart from the speech signals, the noise sources are strong enough when used in hands-free microphones as in laptops instead of using headset. As the distance between the microphone and sound source increases the quality speech gets degraded. Which means the desired signal gets weaker than the noise source signal. If this is this case it is very harder for the system to suppress the noise to operate for the enhancing desired speech signal. Depending on the how the array will be used, it may be important that the microphones used be able to receive sound from all directions. A uniform linear array is created (ULA), in order to determine the source of specific frequency sound and to listen to such sounds in certain directions while blocking the on the other directions. As the ULA is Omni directional, there is a surface ambiguity on which it is unable to determine information about signals. When an example is considered it always suffers from “front-Back ambiguity,’ meaning that signals incident from ‘mirror locations’ at equal angles on the front and back sides of the array are undistinguishable [1].

The distance between the microphones play a major role in the array setup, Our array consists of three Omni directional microphones which are placed equidistant from each other i.e. d= 0.04 cm. As known from sampling theorem that aliasing occurs in the frequency domain if the signal is not sampled at high enough rate. So, in order to avoid aliasing need to have

$$d < \frac{\lambda_{min}}{2} \tag{1}$$

Where λ_{min} is the minimum wavelength corresponding to the maximum frequency f_{max} and

$$\lambda_{min} = \frac{c}{f} \tag{2}$$

$$f_{max} = \frac{f_s}{2} \tag{3}$$

This is due to the fact that the velocity of sound v=$\boldsymbol{\lambda}$ is fixed and thus, when the frequency is maximum, the wave length is minimum .The highest frequency that the array is capable of processing: f_{max}= 1600 Hz. All of these properties generalize to determining the design of any ULA or the design of any array, though other designs may have greater capabilities and thus would require that you consider additional signal properties and how they affect the array [2].

2.5 Microphone array geometry

In this section discussion is done on how the geometry has been applied to the current work. In order for this calculated distance between the sources and microphones by using mathematical geometry formulae. Considered a uniform linear array (ULA) consisting of 3 microphones which are a placed in a three dimensional co-ordinate system with their respective coordinates as m1= $(x_1, y_1 z_1)$, m2=$(x_2, y_2 z_2)$ and $m3 = (x_3, y_3 z_3)$.also assumed the positions of the sources in the same 3 dimensional space. Distance between these microphones is 0.04 cm and it is arranged in such a way that to avoid aliasing in spatial frequency domain. The distance between the sources and the microphones can be calculated are shown below. Concerned Fig 2 is shown below.

$$Bm_1 = \sqrt{(x - x_1\ )^2\ + (y - y_1\ )^2\ \ + (z - z_1\ )^2} \tag{4}$$

$$Bm_2 = \sqrt{(x - x_2\ )^2\ + (y - y_2\ )^2\ \ + (z - z_2\ )^2} \tag{5}$$

$$Bm_3 = \sqrt{(x - x_3\ )^2\ + (y - y_3\ )^2\ \ + (z - z_3\ )^2} \tag{6}$$

$$Am_1 = \sqrt{(x_1^* - x_1\ )^2\ + (y_1^* - y_1\ )^2\ \ + (z_1^* - z_1\ )^2} \tag{7}$$

$$Am_2 = \sqrt{(x_1^* - x_2\ )^2\ + (y_1^* - y_2\ )^2\ \ + (z_1^* - z_2\ )^2} \tag{8}$$

$$Am_3 = \sqrt{(x_1^* - x_3\ )^2\ + (y_1^* - y_3\ )^2\ \ + (z_1^* - z_3\ )^2} \tag{9}$$

$$Cm_1 = \sqrt{(x_2^* - x_1\ )^2\ + (y_2^* - y_1\ )^2\ \ + (z_2^* - z_1\ )^2} \tag{10}$$

$$Cm_2 = \sqrt{(x_2^* - x_2\ )^2\ + (y_2^* - y_2\ )^2\ \ + (z_2^* - z_2\ )^2} \tag{11}$$

$$Cm_3 = \sqrt{(x_2^* - x_3\ )^2\ + (y_2^* - y_3\ )^2\ \ + (z_2^* - z_3\ )^2} \qquad (12)$$

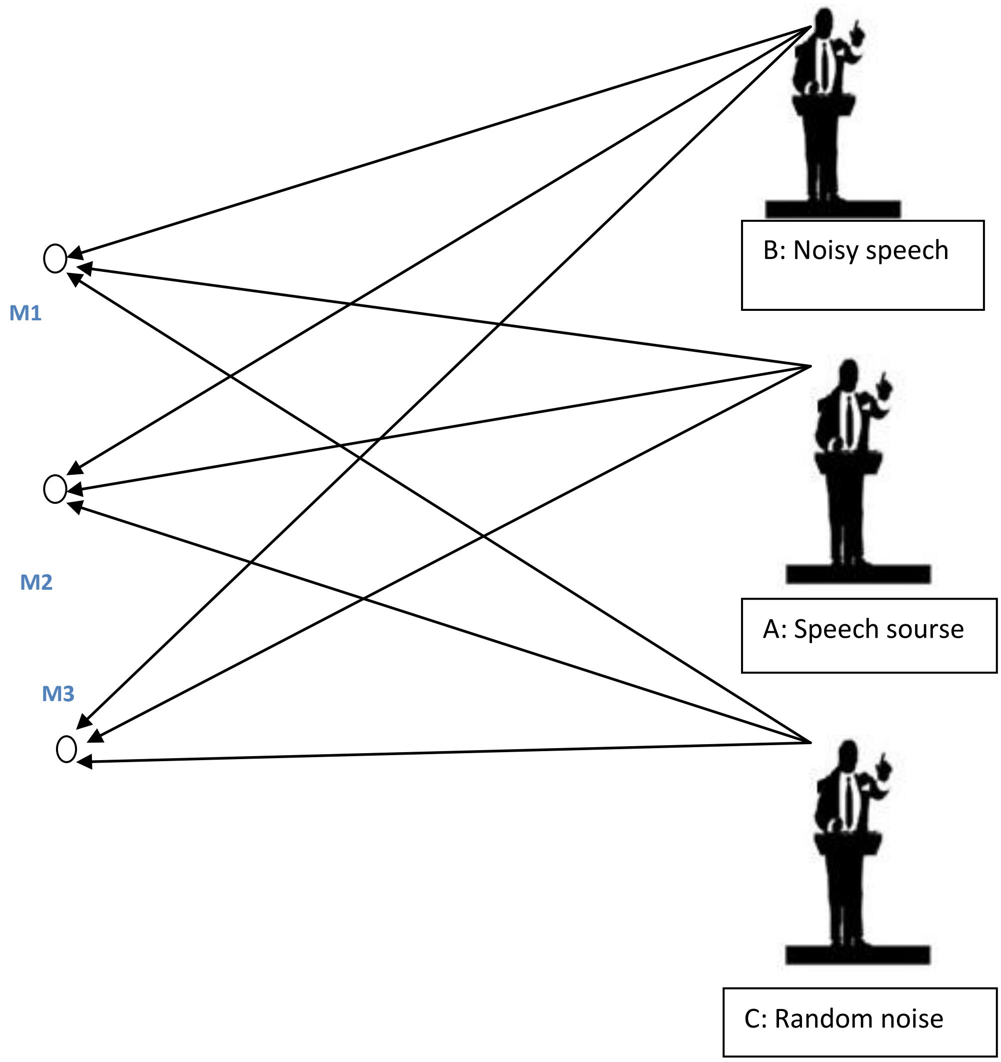


Figure 2 microphone array geometry

2.6 Fractional delay filters

Fractional delay filters (FD) approach much deeps into the Digital signal processing applications i.e. in the field of speech coding and synthesis, communication, music technology [11][10].The standard applications of FD filter are time delay estimation, audio and music technology, speech coding and synthesis, time delay estimation etc[12].It is not only the sampling frequency but also the sampling instants that plays a huge role in these applications.Fd filters provides their contribution building blocks that can be used for tuning the sampling instants i.e. executes the required band limited interpolation [4], which means a signal sample at any approximate point in time even though the point is situated in between two points.

The FD filter is commonly applied in matching of data bits or symbols when dispatched through systems like digital modems. The main function at the receiving end is to find the appropriate dispatched data symbols as accurate as possible. Matching of the sampling frequency and sampling instants are mandatory for attenuating defective decision in digital communication, as it plays a vital role in concluding the decisions of the receiving bits or symbols by considering the samples from the incoming received continuous-time pulse sequence [11].

Designing of the FD filters is to delay the input signals samples by a fractional amount of sampling period. As the delay is in fractional amount, the intersample, performance of the genuine signal becomes too important. The expectation in designing the FD filter is that incoming consecutive time signals are fully band limited up to the nyquist frequency and it is constructed in discrete time domain.

2.7 Classic FD Delay using FIR filter

When a signal x(n) is delayed, it's delayed version of the signal x^*(n) is represented as

$$x^*(n) = x(k - \beta) \tag{13}$$

The delay is sample is calculated as

$$\beta_{samples} = \frac{f_s * d}{c} \tag{14}$$

Where c velocity of sound i.e. is equal to 343.3

Where k is integer or sample index and β is the amount of delay familiarized in the signal which is a s intefer part. Fractional part is represented as

$$d = \beta - floor(\beta) \tag{15}$$

The function floor helps in finding the greatest integer part which is less than or equal to β.It is possible to reconstruct the original signal from sampled data by multiplying each sample by a scaled sinc function Based on the Nyquist-shannon theorem. The exclusive condition is that the original waveform is band limited to have a maximum frequency component of less than that of half the sampling rate. So for a signal which is sampled at 16000 samples per second the maximum component frequency must be less than 8kHz [4].

The most popular way of handy implementing fractional delay is by utilizing sinc filter that shifted by the fractional amount [5]. This can be carried out using a standard FIR structure. In order to calculate the signal instants values at any point in time can be detected by using sinc interpolator according to Shannon's sampling theorem. So By convolving the delayed signal $x^{*}(n)$ with $sinc(k-\beta)$ to given signal. Sample β At any arbitary time and k is a sample index.

$$y(\beta) = \sum_{n=-\propto}^{\propto} x^{*}(n)\ sin(k-\beta) \tag{16}$$

The delayed sic function is assigned to as a ideal fractional delay interpolator [11] [13]

$$h_{\beta}(k) = \frac{sinc(\pi(k-\beta))}{\pi(k-\beta)} \tag{17}$$

Where k is the time index ranges in between 0 and N-1 and $h_{\beta}(k)$ is the impulse response sequence corrsponding to H($e^{j2\pi v}$).The sinc functions is infinite along the x axis, however a FIR implementation requires a finite number of taps. Generally in Fd filter the delay slightly moves the impulse response in time domain, hence the moved and sample sinc function is the impulse response of ideal fractional delay filter. Ideal fractional delay is shown in Fig 3.

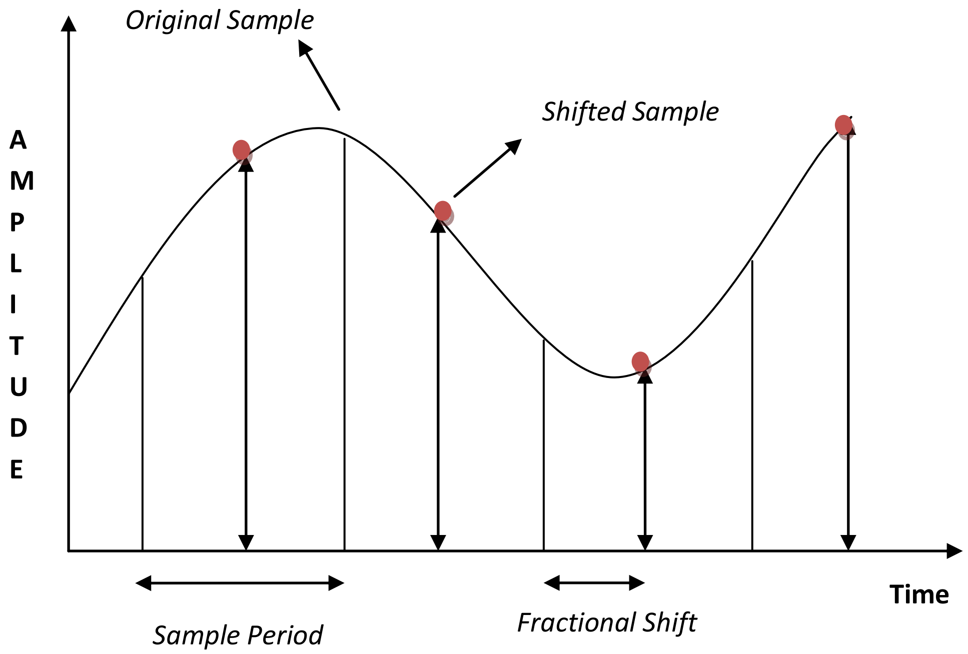


Figure 3 Ideal fractional delay

Chapter 3
Beamforming Basics

3.1 Introduction to Beamforming

Beam forming is a spatial filtering technique that divides sound sources based on the position in space [6]. The basic idea of beam forming is to concentrate on the array of sounds arriving from only one particular direction to an array of microphones and this would look like a large dumbel shaped lobe aimed in the direction of interest. Making a beam former is one of the important tasks for our work, which is to listen to the sounds particularly from one direction and making rest of them to be ignored. The Figure 5 below shows the general visualization of the beam former in lab veiw.The best way to listen to sound only one direction is to steer all your energy towards it.Beamforming technique was first originated in radio astronomy around 1950’s as path of combining antenna information from collection of antenna dishes, it started to explore as generalized signal processing in numerous applications involving spatially distributed sensors by 1970’s. Examples of this expansion include sonar, to allow submarines greater ability to detect enemy ships using hydrophones, or in geology, enhancing the ability of ground sensors to detect and locate tectonic plate shifts [7].It was around during that time microphone array have become an active area of research which made to keep the virtual microphone at some position instead of physical sensor movement. Applications of beam forming include Laptop’s, Hands free telephony, conference mikes etc. This is an important concept, because it is not just used for array signal processing, it is also used in many sonar systems as well. RADAR is actually the complete opposite process, so will not deal with that. Figure 4 shows its general view.

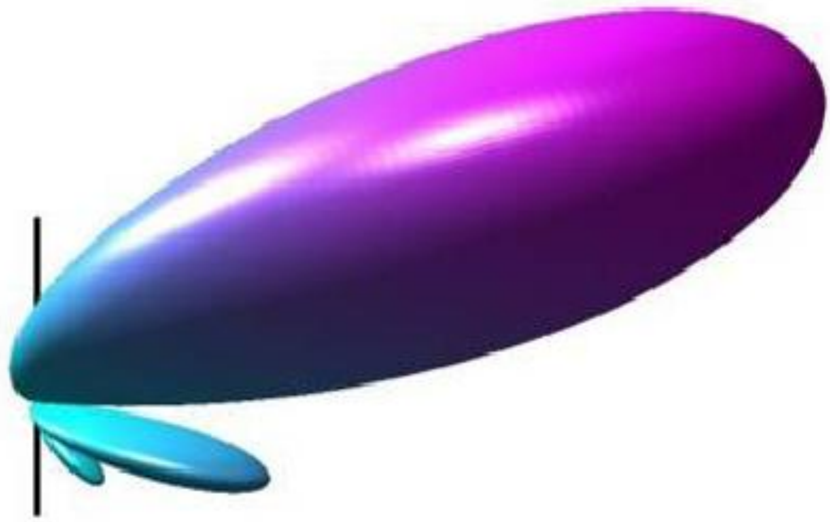


Figure 4 Beam former visualization

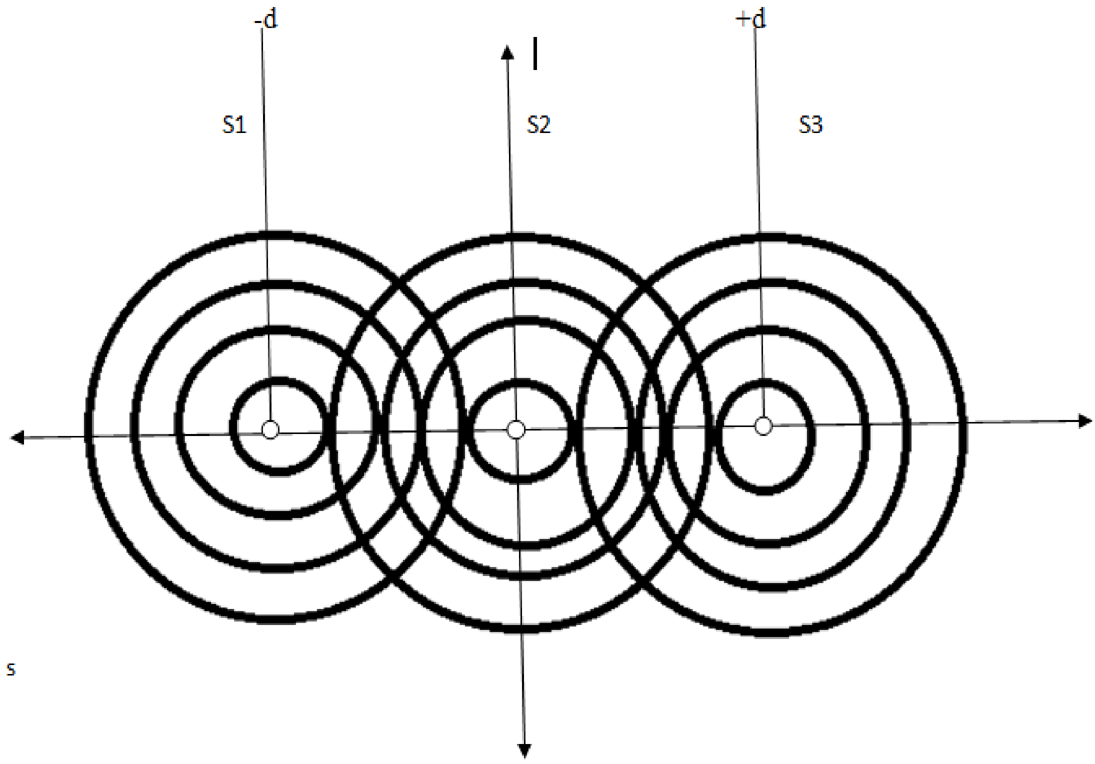


Figure 5: general view of beam former

3.2 Speech Extraction by Array Processing:

As far as speech extraction is concerned beam forming is divided into two, one is narrow band beam forming and the other is broad band beamforming.In narrow band beam forming the sound source signal is a point source or else it can seen as number of point sources collection closely in space, a pair of microphones can be arranged in such a way that the combination of the microphone signals will have high response at the point [8]. A small change in the placement of the microphones will leads to different responses for the same location and this is due to change in the distances to the microphones. By debuting the variable signal delays it is possible to control the peak response of the array and this why this procedure commonly referred as the narrow band beamforming.However as it deals with the speech signals in this work the point source signal is a wide band signal. In this case the beam former need to have the ability to deal with wide range of frequencies variously. A casual way to achieve this feature is by using a linear filter at each microphone signal. The main reason for constructing this filter is to make the signals from the certain locations pass the system and to make the rest of the signals from the other locations be cancelled or attenuated. The filter used here is a Wiener filter so it is called as Wiener beamforming. There are mainly two groups of beam forming as far as the researches are concerned one is fixed beam forming(data independent) and the other is adaptive beam forming(data dependent beam forming). Adaptive beam formers update their parameters to better suit the input signal, adapting to changing conditions. Fixed beam formers are very easier to implement than the adaptive ones, but are more limited in their ability to eliminate high directive noises. Moreover there are several beam former techniques to implement, but among them simplest is Wiener beamformer.However applied the adaptive beam former after Wiener filter is been applied i.e. when GSC is introduced. Figure 6 shows the structure of the Wiener filter beam former. The output of the beam former is y (n). The procedure really goes on with the Wiener concept is explained in the following section.

$$y(n) = \sum_{m=1}^{N} W_{opt}^{H}\, x(n) \tag{18}$$

3.3 Wiener Solution:

Considered the model of the signal in such a way that one of the speakers is situated at one position and the other two noise sources arrives from the various other directions from different points. The output at the sensor consist of speech component s(n) and the other noise components v1(n) and v2(n).Have constructed the filters after the sensors in such a way that output of the beam former resembles the signal component and the rest of the noise signals are attenuated or cancelled.

The optimal filter weight vector based on the Wiener solution is given by

$$w_{opt}[n] = [\hat{R}(n)]^{-1} \hat{r}_{dx}(n) \tag{19}$$

Here the array weight vector $w_{opt}^{(k)}[n]$ is arranged as

$$w_{opt}[n] = [w_1\ , w_2, w_3\ ,. . . . . w_I] \tag{20}$$

Where $\hat{R}^{(k)}(n)$ is a combined correlation matrix estimate?

$$\hat{R}(n) = \hat{R}_{SS}(n) + \hat{R}_{NN}(n) \tag{21}$$

$$\hat{R}_{ss} = \begin{pmatrix} R_{s_1 s_1} & \cdots & R_{s_1 s_I} \\ \vdots & \ddots & \vdots \\ R_{s_I s_1} & \cdots & R_{s_I s_I} \end{pmatrix} \tag{22}$$

where

$$R_{sisj} = \begin{pmatrix} r_{s_i s_j}[0] & \cdots & r_{s_i s_j}[L-1] \\ \vdots & \ddots & \vdots \\ r_{s_i s_j}^{*}[L-1] & \cdots & r_{s_i s_j}[0] \end{pmatrix} \tag{23}$$

$$R_{nn} = \begin{pmatrix} R_{n_1 n_1} & \cdots & R_{n_1 n_I} \\ \vdots & \ddots & \vdots \\ R_{n_I n_1} & \cdots & R_{n_I n_I} \end{pmatrix} \tag{24}$$

Where $$R_{n_in_j} = \begin{pmatrix} r_{n_in_j}[0] & \cdots & r_{n_in_j}[L-1] \\ \vdots & \ddots & \vdots \\ r^*_{n_in_j}[L-1] & \cdots & r_{n_in_j}[0] \end{pmatrix} \quad (25)$$

$$\hat{R}_{SS}(n) = \sum X_S(n) X_S(n)^H \quad (26)$$

$$\widehat{R}_{NN}(n) = \sum X_N(n) X_N(n)^H \quad (27)$$

$$\hat{r}_{dx}(n) = \sum X_S(n) X_S(n)^* \quad (28)$$

Where

$$X_S(n) = [X_{S1}(n)\ X_{S2}(n) \ldots\ldots\ldots\ldots\ X_{SI}(n)]^T \quad (29)$$

$$X_N(n) = [\,X_{N1}(n)\ X_{N2}(n) \ldots\ldots\ldots\ldots X_{NI}(n)]^T \quad (30)$$

The signal $X_S(n)$, $X_N(n)$is the received data at the i: th microphone when only the interested source signal of is active and only the Noise is active respectively.

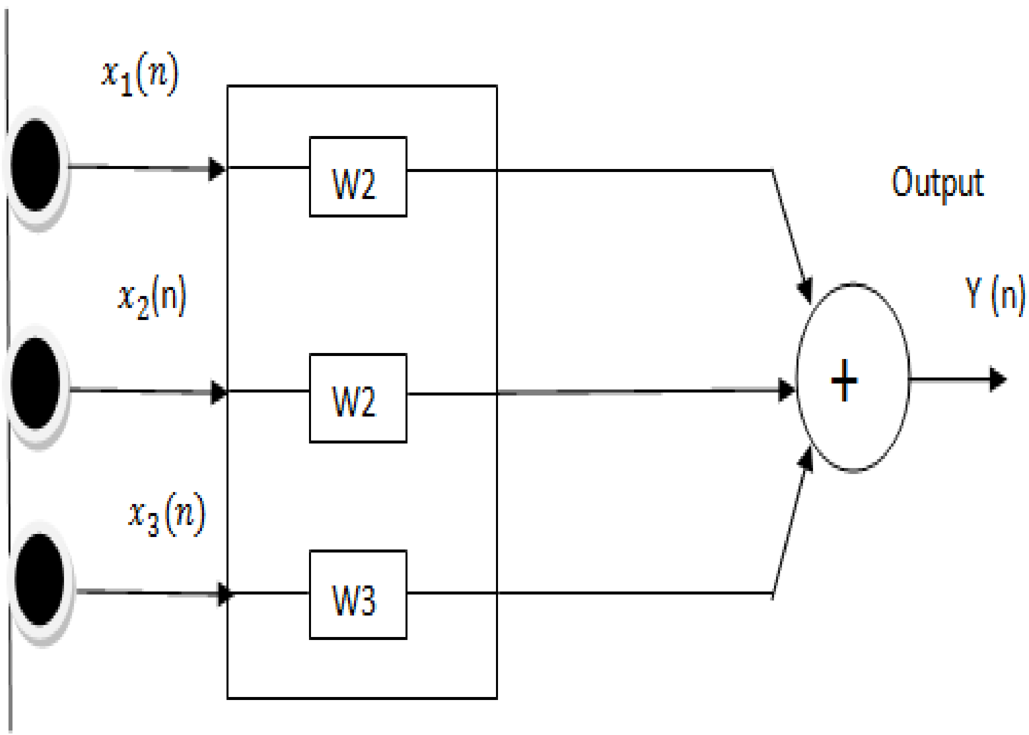


Figure 6: Wiener beamformer implementation

The output of the Wiener beam former is given by

Y (n) =

$$\sum_{i=1}^{I} \sum_{j=0}^{L-1} w_i(j) x_i\,(n-j) \qquad (31)$$

And it can be written as

$$Y\ (n) = w_{opt}^{(k)H} x^{(k)}[n] \qquad (32)$$

Chapter 4

Speech Enhancement Methods

The term speech enhancement refers to methods aiming at recovering speech signal from a noisy observation. There are number of ways to classify these types of methods, each method has various profession that are based on the certain assumptions and constraints that particularly depends on various conditions and scenarios. Therefore, it is highly impossible for a specific algorithm to perform optimally across all noise types. Speech enhancement techniques can be broadly divided into two types one is single channel speech enhancement and the other is multichannel speech enhancement techniques. As here in this thesis deal with two or more speech sources with multichannel speech enhancement techniques. This method usually performs well when the speech sources are non stationary and in low SNR conditions than the single channel speech enhancement technique. This technique usually uses the beam forming algorithm or spatiotemporal filtering, which are given below.

- Generalized side lobe cancellation(GSC)
- Blind Source separation(BSS)
- Linear constraint Minimum Variance(LCMV)
- Adaptive Noise Cancellation(ANC)
- Delay and Sum Beam Forming(DSB)

The Adaptive Noise Cancellation (ANC) is a well known speech enhancement technique that uses a primary channel containing corrupted signal and a reference channel containing noise correlated with primary channel noise to cancel highly correlated noise [19]. In order to get the exact desired signal the Mentioned input is filtered by an adaptive algorithm and then remove from the input signal. This algorithm has some leakage problem; if the primary signal is leaked into the reference signal then some original speech is cancelled and thus the speech quality decreases. A difficult and well-known problem for adaptive noise cancellation arises when there are plant resonances blocking the noise cancellation path [20].

The Blind Source Separation (BSS) is performed in the conditions where the signal and noise are independent; basically it is used to separate the mixed signals where the signals come from different directions.[21].

The Delay and Sum beam forming (DSB) is quite a simple algorithm and its ability depends on the number of microphones used in a system and helps in separating multiple sound source signals [22].

The Linear Constrained Minimum Variance (LCMV) beam forming is able to perfectly cancel a number of dominant interferers while other interferers remain same [23] .The case here deal with in this thesis is Generalized Side Lobe Cancellation (GSC).

4 .1 Adaptive Filters:

Adaptive filter is a digital filter which aspires to transfer the information carrying signal into improved form, by adjusting the characteristics according to the given input signals. As far as the machine learning is concerned adaptive filter is the easiest and among the algorithms..These filters are generally preferred over their particular characteristics counterpart, which are elementally unable to adjust to changing signal conditions. The self ruling, simple structure, good stability features and adaptability of adaptive filters made them widespread in different signal processing applications namely in adaptive control, Radar, system identification etc [18].

4.2 The Generalized Side lobe Canceller (Griffiths-Jim Beam former)

The generalized side lobe canceller is a simplification of the Frost Algorithm Presented by Griffiths and Jim and some ten years after Frost's original paper was published [9]. This section discusses the layout of the GSC and its implementations. Generalized side lobe canceller is the most frequent and achievable approach used in microphone applications. It is used to decrease the noises or interferences form non target location in array beam forming and can be used as a adaptive noise canceller in array processing [14] . A general lay out of the GSC is shown in fig 444.

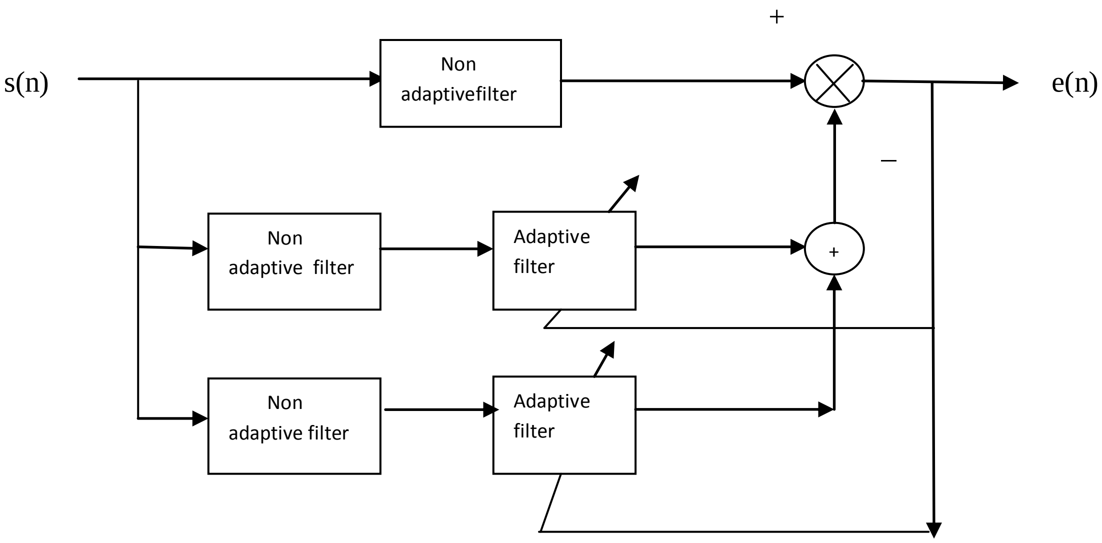


Figure 7 Block diagram of generalized side lobe canceller

This structure consists of a three non adaptive filters(Wiener filters) and two adaptive filters(LMS).The two non adaptive filters which are at the down side of the above figure are connected to the adaptive filters ,which means the adaptive part is mixture of both adaptive and non adaptive filters. As the GSC is adaptive technique the weights keep on changing based on the input signal given. Adaptive techniques present a higher capacity at reducing noise interference but are much more sensitive to steering errors due to the approximation of the channel delays. There are three fixed beam formers in the structure. The top branch of the structure produces beam formed signal which is a fixed one, even the rest two of the branches also produces the beam formed signals respectively. The outputs from the second and the third Wiener beam formers are given to the two LMS algorithms. The adaptive part of the GSC is to reduce the noise, which generally is used to match the interference in adaptive branch to as close a possible to interference in non adaptive branch.

4.3 Least mean square algorithm (LMS):

One of the most popular algorithms in adaptive signal processing is Least Mean Square (LMS) algorithm. Least Mean square algorithm (LMS) was proposed by Widrow and Hoff in 1960 with a small calculation, simple structure and robustness [18]. It has been extensively analyzed in the literature, and a large number of results on its steady state mi It is one of the most important adaptive algorithm in adaptive signal processing [15],[16] .It has been extremely evaluated in the literature, and a large number of results on its steady state maladjustment and it's tracking performance has been obtained[17].

$$w_{n+1} = w_n + \mu E[e(n)x^*(n)] \tag{33}$$

$$e(n) = d(n) - X^T(n)W(n) \tag{34}$$

The step size range of the Basic LMS algorithm is $0 < \mu < \frac{1}{\lambda_{max}}$ .Where λ_{max} is the largest value of input correlation matrix. The value of the µ should be very small.

A constructive limitation with this algorithm is that the expectation E[e(n) $x^*(n)$] is generally not known.so, it must be compensate with an estimate such as the sample mean

$$E^{\wedge}[e(n)x^*(n)] = \frac{1}{L}\sum_{l=0}^{L-1} e(n-l)\, x^*(n-l) \tag{35}$$

Associating this estimate into the steepest descent algorithm, the up\date for w_n will turn into

$$w_{n+1} = w_n + \frac{\mu}{L}\sum_{l=0}^{L-1} e(n-l)\, x^*(n-l) \tag{36}$$

The above equation is possible if is use L=1 which is alone point sample mean

$$E^{\wedge}[e(n)x^*(n)] = e(n)x^*(n) \tag{37}$$

In this particular condition the weight vector update equation consider a appropriate simple form

$$w_{n+1} = w_n + \mu e(n)x^*(n) \tag{38}$$

And the above equation is known as LMS algorithm. The easier about this algorithm comes from the truth that the update for the K th coefficient.

$$w_{n+1}(k) = w_n(k) + \mu e(n)x^*(n-k) \tag{39}$$

The LMS is based on steepest descent algorithm. The most excited factor is how the weights derive in time beginning with an approximate initial weight vector and thus updates the weight vector. The coefficients began to alter about the optimum value as the weight vector begins to converge in the mean. One of the main problems in the design and implementation of LMS algorithm is the exact selection of step size μ.The majority of the works done examining LMS algorithm with constant step size. The selection of the step size emulates a conclusion between the maladjustment and the speed of adaptation [15].Here μ should be smaller.

CHAPTER 5
Implementation Matlab

5.1 Developing Array Model (Microphone Array Setup)

The basic assumption of the array processing is illustrated as follows. Considered a (plain signals) room structure where no reflection of signals or reverberation (plain signals) is taken into consideration. The length, breadth and height of the room are taken as 5 meters each respectively. A signal model in Fig.8 is considered in such a way that positions of the microphone are fixed at a certain positions in the room whose co-ordinate axis in 3-D space are taken as follows.

$$m1 = \ (2.46,1.24,2.50)$$

$$m2 = \ (2.50,1.24,2.50)$$

$$m2 = \ (2.54,1.24,2.50)$$

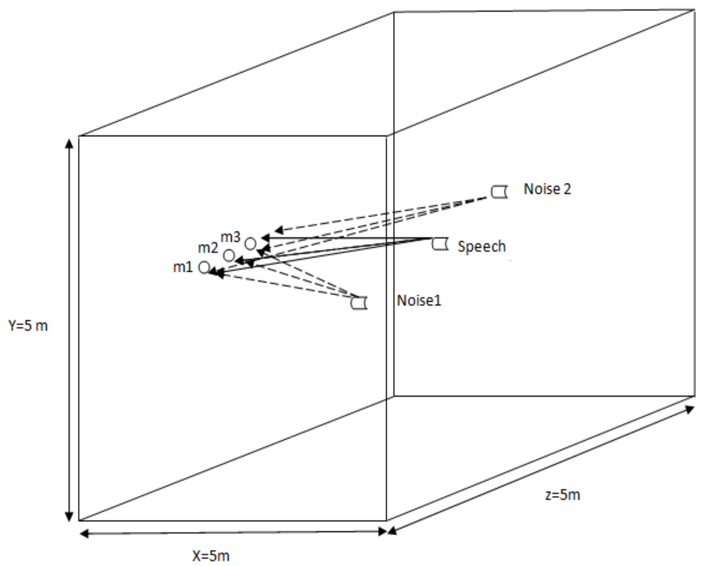
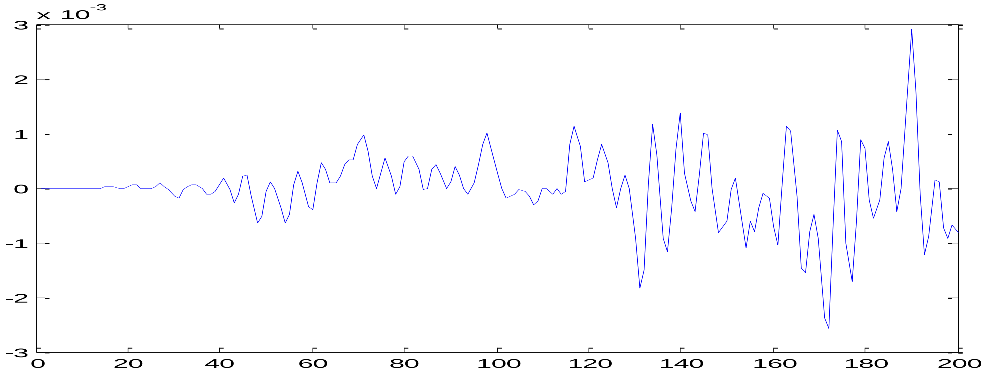


Figure 8: coordinate system for sources

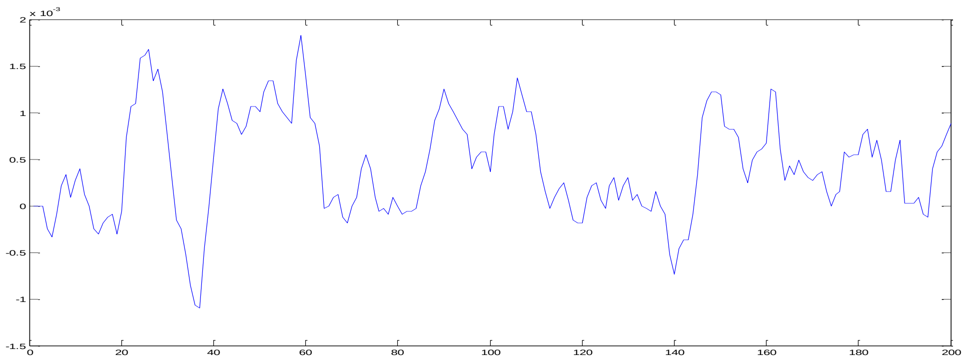
Which means the distance between each microphone is $d = 0.04\, m(4\, cm)$.This distance is fixed because to avoid aliasing effect [11].Speech sources and Noise sources are situated in a certain positions in room corresponding to the sensor arrays. Where the noise environments consists of interference sources which will have ambient noise field [Nedelko]. Microphone array receives the speech component from the speaker$s(n)$, and the sum of the noise sources, together with the ambient noise field, $v_1(n) + v_2(n)$ .It can be represented in the equation form as

$$p(n) = s(n) + \sum_{i=1}^{2} v_i\,(n) \qquad (40)$$

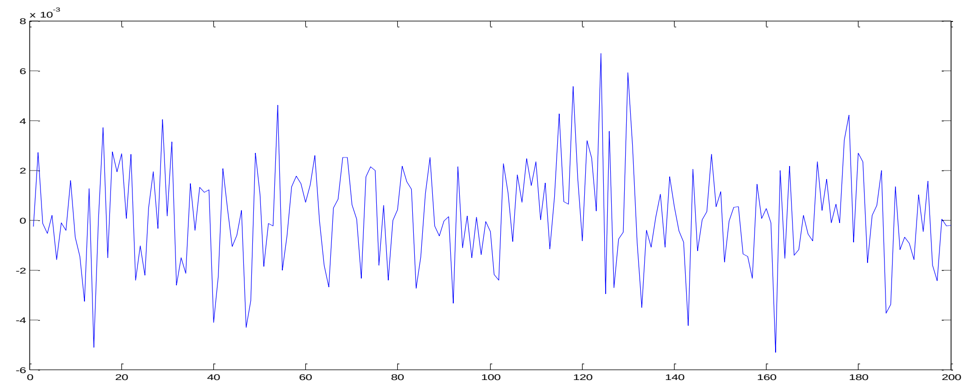
The figure below shows the speech and the noise signals at their respective initiation points.



(a)



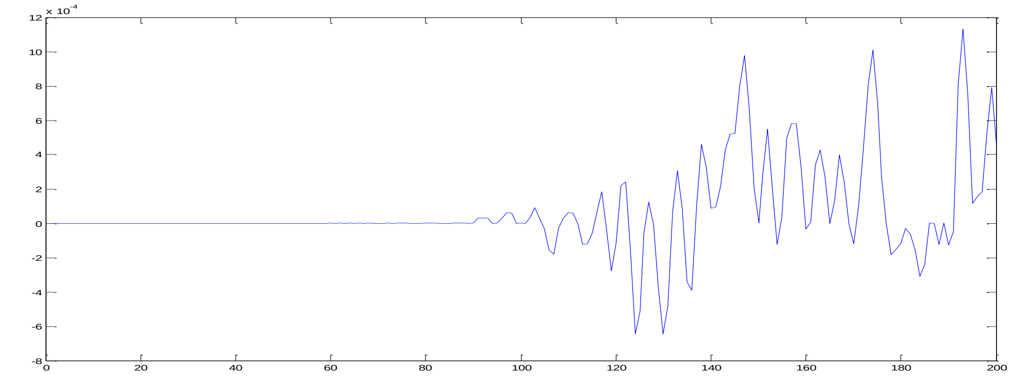
(b)



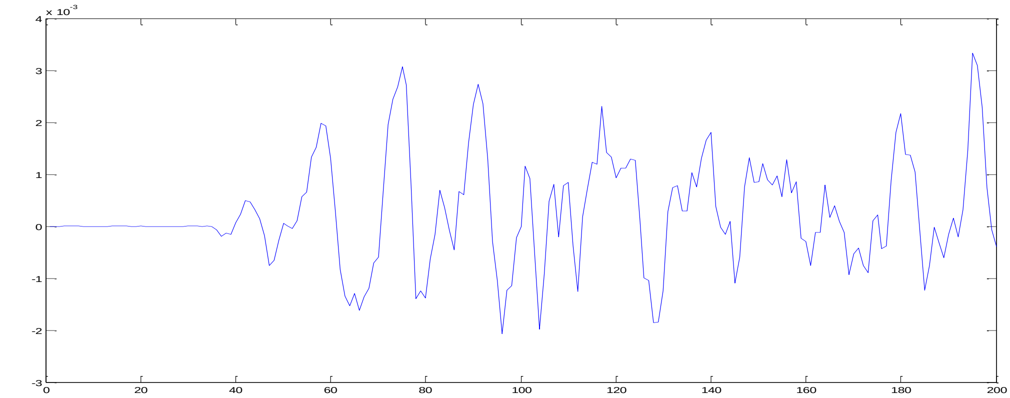
(c)

Figure 9: a) speech signal 1 at its origin b) speech signal 2 (N1) at its origin c) random noise (N2) at Its origin

When the sources and noise are at particular positions in the room i.e. **Speech(B)=(2.50,1.34,2) ,Noise1(A)=(1.25,2,1.34), Noise(C)=(1.28,2.3,1.34)**. The concerned speech signals and the interference signals reaches the corresponding arrays with some time delay as they are originating from the different source positions. The delay is a mixed part which contains both the integer part and the fractional part and it is calculated by using the sinc windowing by considering filter length as 64. The integer part of the delay is very simple to achieve with basic buffer. However, the fractional part is more complicated. The Fig shows the speech signals and noise signals after they are delayed.



(a)



(b)

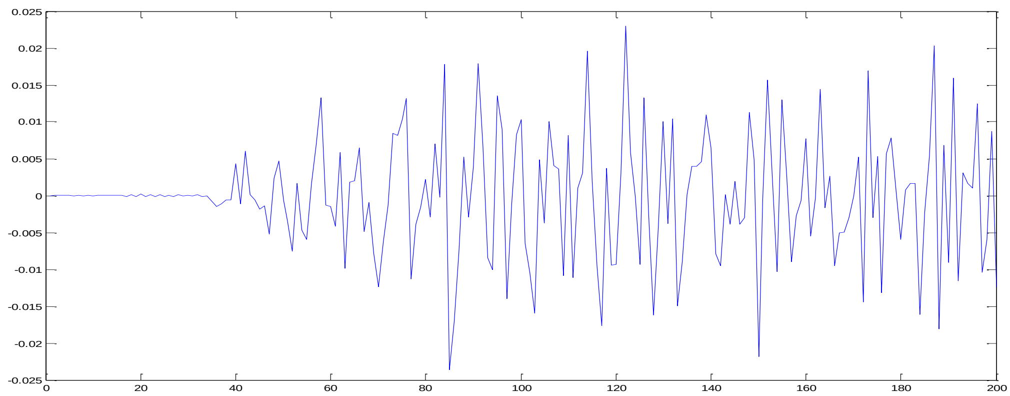


Figure 10: a) delayed speech signal 1 at mic b) delayed speech signal 2 (N1) at mic C) delayed random noise (N2) at mic

5.2 Wiener Beam former Implementation:

Three Wiener beamformer systems are considered. All the signals from different positions will reach the microphones, so the output of microphone array consists mixture of all the signals. These outputs from the microphone array will encounter with Wiener filter exercise tendering the equation mentioned 19 to 32.The outputs from the three different Wiener beamformers is $y_1(n)$, $y_2(n)$,$y_3(n)$ respectively.

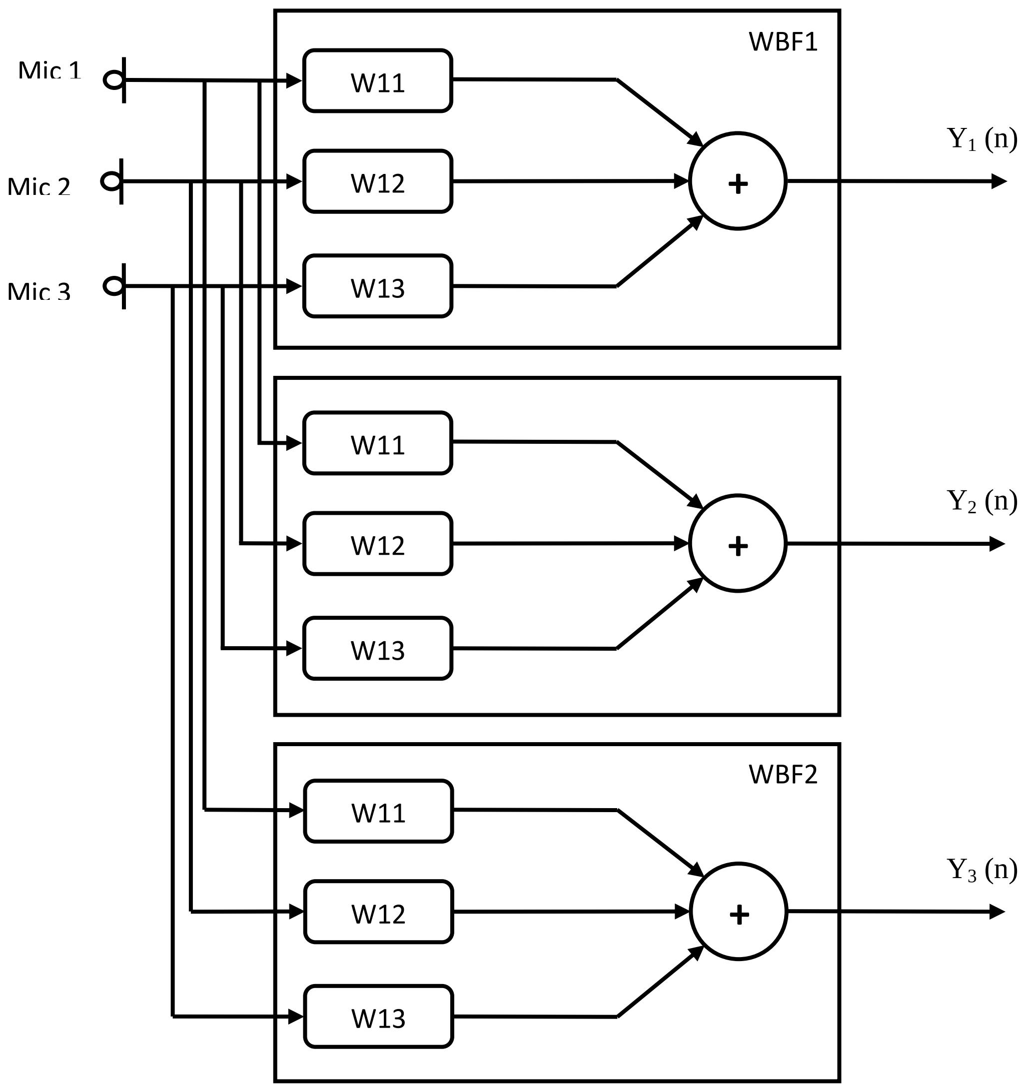


Figure 11: Implementation of Wiener Beam formers

For the first Wiener Beam former 1 ,one of the speech from the input are taken as main speech and the rest of the two as interferences noises which means that the output $y_1(n)$ resembles the concerned signal component, while the rest of the speech signals are considered to be s noise which gets attenuated or cancelled. The same procedure is done in the next Wiener beam former 2(WBF2) and Wiener beam former 3 (WBF3) respectively, but by considering the rest of speech signals as main signal components in the both the cases and their respective outputs are $y_2(n)$and $y_3(n)$ respectively.The Outputs from the $y_2(n)$and $y_3(n)$ are given to the Lms and it will be dicussed in the following section.

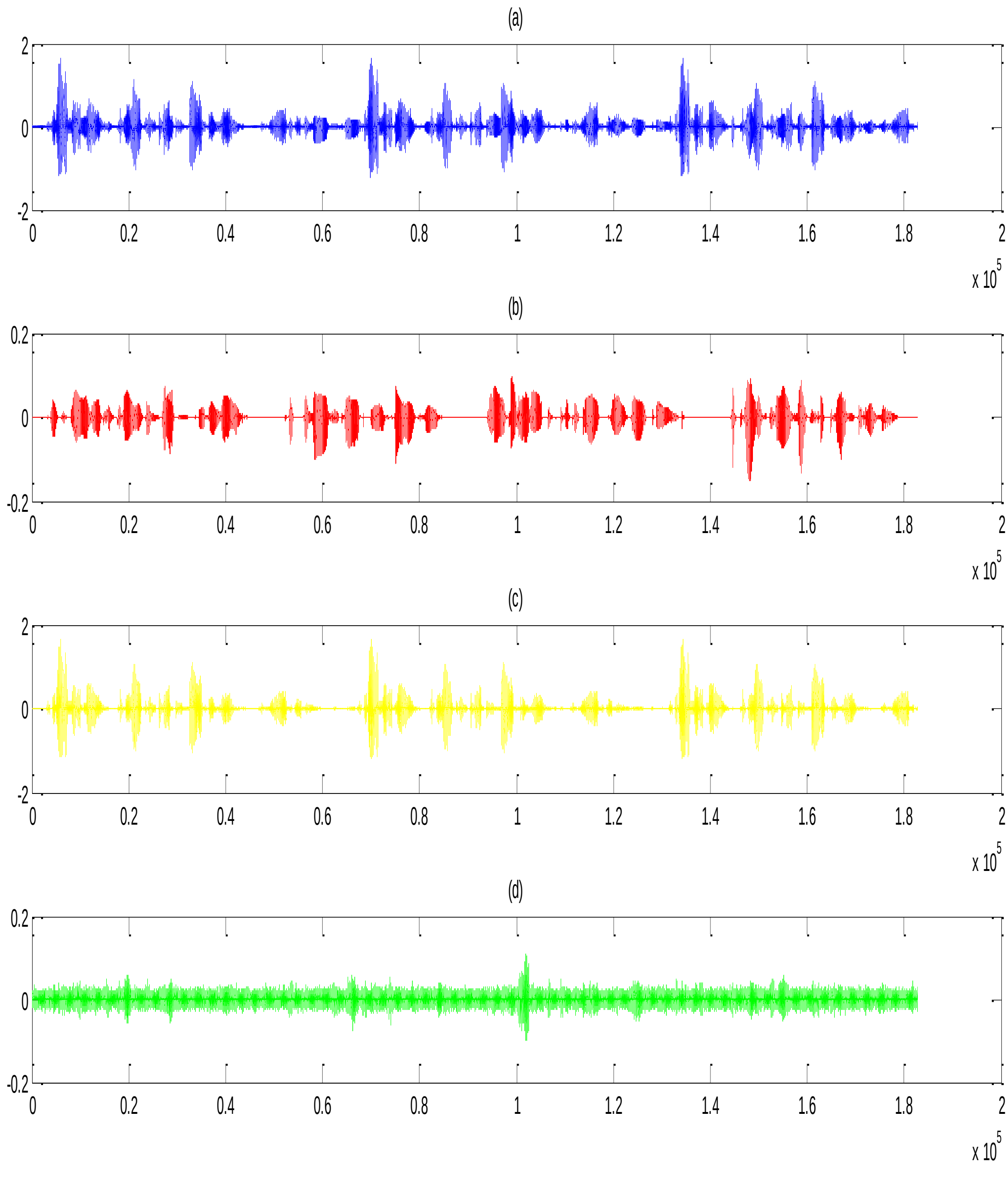


Figure 12: a) at mic b) at WBF 1c) At WBF 2 d)at WBF 3

5.3 Implementing GSC for the system:

This structure shown in the Fig. 12 express WBF based GSC.The Preferred system consists of a three non adaptive filters (Wiener filters) and two adaptive filters (LMS) along with a microphone array. The numbers of microphones used are 3. The input to the micro phones is a combination of main speech signal, noisy speech signal and a random noise. They reach the microphones with delay as they originate from different positions, which are discussed in the previous sections. The mixture of signals from the microphones is given to the Wiener Beam formers (non adaptive filters). The top branch of the structure produces beam formed signal which is a fixed one and its output is taken as $y(n)$, even the rest two of the branches also produces the beamformed signals respectively.The outputs from the second and the third Wiener beam formers are given to the Adaptive part which consists of two LMS algorithms. The description and concerned equations of the LMS are discussed in the previous chapters. As the GSC is adaptive technique the weights keep on changing based on the input signal given. Adaptive techniques present a higher capacity at reducing noise interference but are much more sensitive to steering errors due to the approximation of the channel delays. Figure ....shows the output of the system comparison with the input speech.

The output from the adaptive filters is y_b

$$y_b = \sum_{k=1}^{2} w_k^T v_k \tag{41}$$

The output of the GSC is given as

$$Y(N) = y_1 - y(n) \tag{42}$$

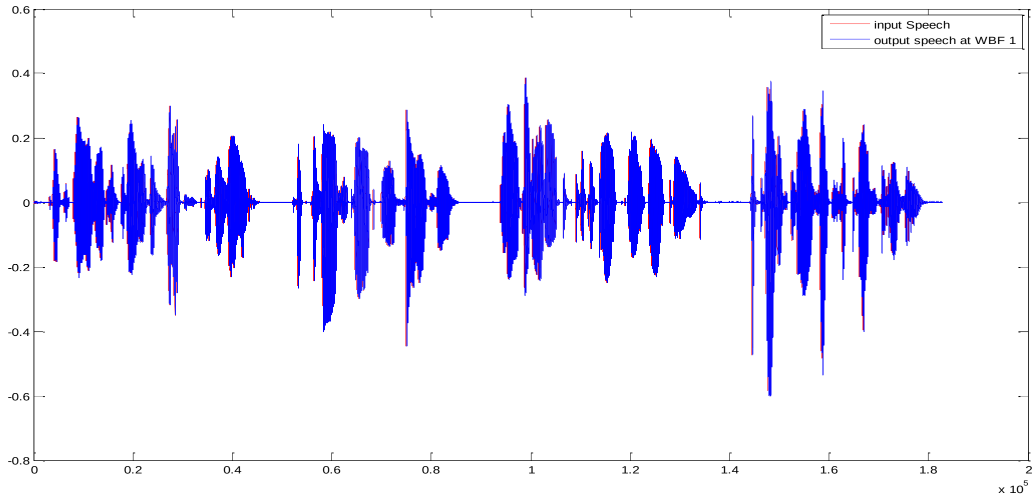


Figure 13: speech at the system output

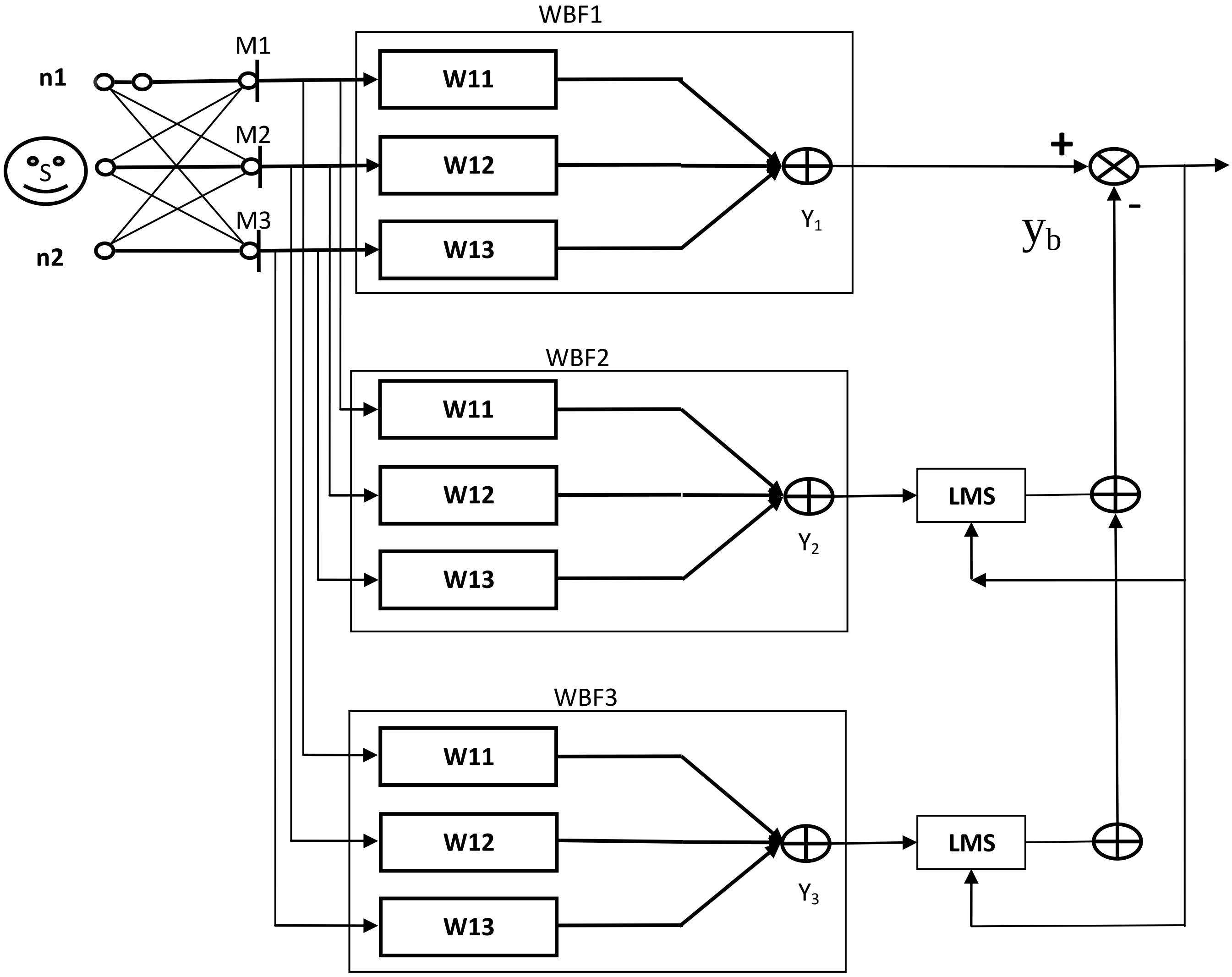


Figure 14: detailed structure of GSC implementation

5.4 Signal to Noise Ratio

Objective test used to measure the performance of the above mentioned system at WBF 1 and at system output is Signal to Noise Ration (SNR).Signal to Noise Ration (SNR) is calculated for them. Signal to noise ratio is defines as the variance of the output power to the various of the noise power and

$$\text{SNR}=\frac{\text{Signal output power}}{\text{Noise Output power}} \tag{43}$$

Output at the Mic is mixed of main speech signals, Noisy speech signal and random Noise, where the noises are been multiplies with $\propto$ to attenuate the noise.

$$mic(2) = S(n) + \propto v_1\ (\mathrm{n}) + \propto v_2\ (\mathrm{n})$$

$$\propto = \sqrt{\frac{1}{10^{\frac{SNR_{In} - SNR}{10}}}}$$

In the equation form it is written as SNR=$10\log_{10}(\frac{var(speech)}{var(noise)})$ (44)

$$SNR_{In} = 10\log_{10}[\frac{var(speech)}{\propto^2 var(n1) + \propto^2 var(n2)}] \quad (45)$$

$$SNR_{In} = 10\log_{10}[\frac{var(speech)}{\propto^2 (var(n1+n2))}) \quad (46)$$

$$SNR_{In} = 10\log_{10}(\frac{var(speech)}{var(n1+n2)}) + 10\log_{10}(\frac{1}{\propto^2}) \quad (47)$$

Chapter 6

Results and Analysis

The evaluation of results is been done by choosing two different speech signals and a random noise (source points) at three different cases entirely in Matlab environment. These three signals are sampled at 16 KHz, which are of 182824 samples of data. Phrases of the two speech signals are **"The time is now twenty five to one in the afternoon (N2)" and "(s) it's easy to tell the Depth of a well... Kick the ball straight and follow through ...blue the sheet to the dark Blue background** and third one is a random noise (N3).Among these three signals one is considered as a main speech and the rest two as a interferences or noises. The main aim of our work is to suppress the noises. The tests are performed at various SNR inputs by changing the noise power based on the equations 43 to 47.For the evaluation there are three different consideration points from where the source and noises originates as far as the room scenario is concerned.

And the cases are as follows according to the room dimensions (non reverberant room) to the microphone array.

1) S= [2.50 1.24 2.20], N1= [2.60 0.50 2.20], N2= [2.40 0.50 2.20]

2) S= [2.50 0.50 2.20], N1= [2.50 1.34 2.20], N2= [2.70 0.50 2.20]

3) S= [2.85 0.20 2.20], N1= [2.30 0.50 2.20], N2= [2.55 0.45 2.20]

The distance between the signal sources and the microphones are calculated based on the equations 4 to 12.For the considered point 1 the respective distances between them are as follows. m1S=0.3187, m2S=0.300, m3S=0.3027 and m1N1=0.8107, m2N1=0.8047, m3N1=0.8007, and m1N2=0.8007, m2N2=0.8047, m3N2=0.8107,

For considered point 2. m1S=0.799, m2S=0.7985, m3S= 0.7995and m1N1=0.3187, m2N1=0.3162, m3N1=0.3262, and m1N2=0.838, m2N2=0.8232, m3N2=0.8114,

For considered point 3. m1S=1.5105, m2S=1.1376, m3S=1.1259 and m1N1=0.9215, m2N1=0.9421, m3N1=0.9640, and m1N2=0.144, m2N2=0.8232, m3N2=.08338.

For these three different consideration points the analysis were achieved at their respective Systems i.e. at WBF 1 and at the system output(after WBF 2,WBF 3 are combined with LMS) and their respective Results by varying input SNR's from 0,5,15,20,25 are shown in following sections and their values are tabulated in Table-1 ,Table-2 ,Table-3 respectively.SNR is calculated based on the equations from 43 to 47. For Case 1, the main speech source is nearer to the microphone array and the noises are far from it, at this point the performance of the system is effective in cancelling the noise with an improvement around 35 to 40 dB and 42 to 47 dB at WBF 1 and System output Respectively and it respective figure 15 for the case 1 is shown below.

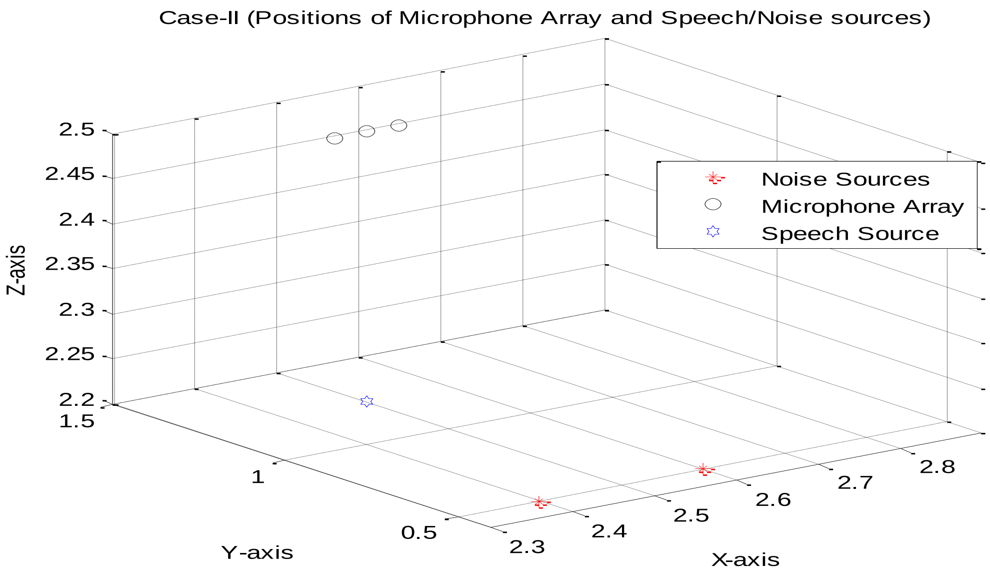


Figure 15 : case 1 setup

For Case 2, the main speech source is somehow farer to the microphone array compared to the one Noisy signal and nearby to the other noise signal, at this point the performance of the system is slightly degrades when compared to the consideration 1with an improvement around 30 dB and 37 dB at WBF 1 and System output respectively. The respective figure for this case is shown in Figure 16

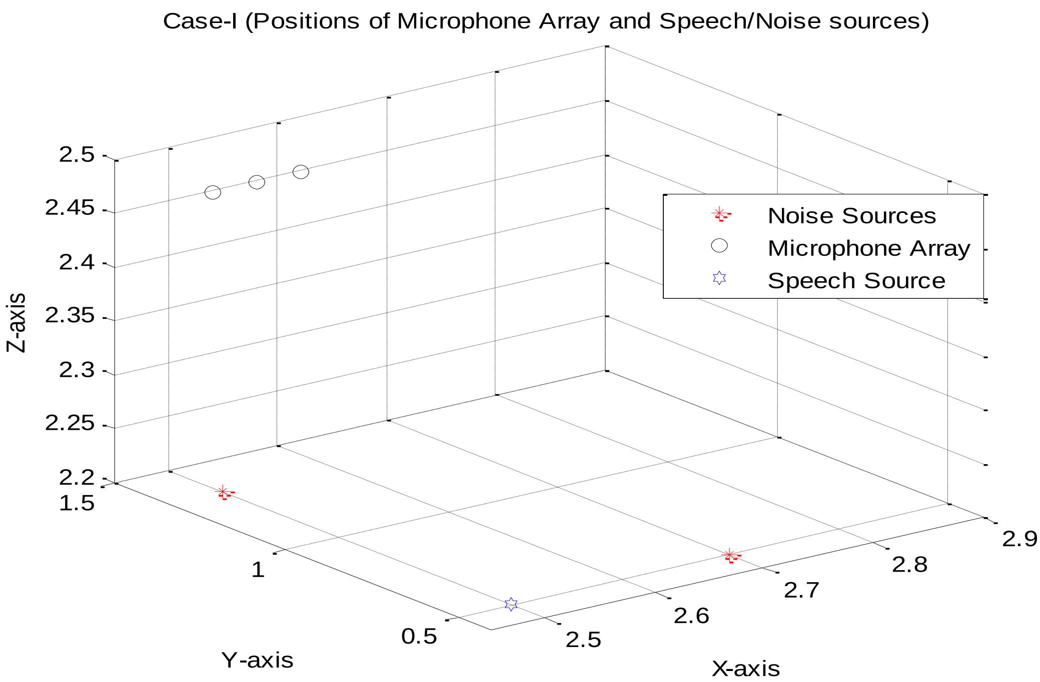


Figure 16: case 2 setup

For case 3, the main speech source is too farer to the microphone array and Noisy speech signal and random noise is far too but not as Main speech source. At this point the performance of the system is degraded when compared to the consideration 1 and consideration 2 as the main speech is too long to the microphone array. The SNR and SNRI of the system at this position are with an improvement around 15 dB and 21 dB at WBF 1 and System output respectively.

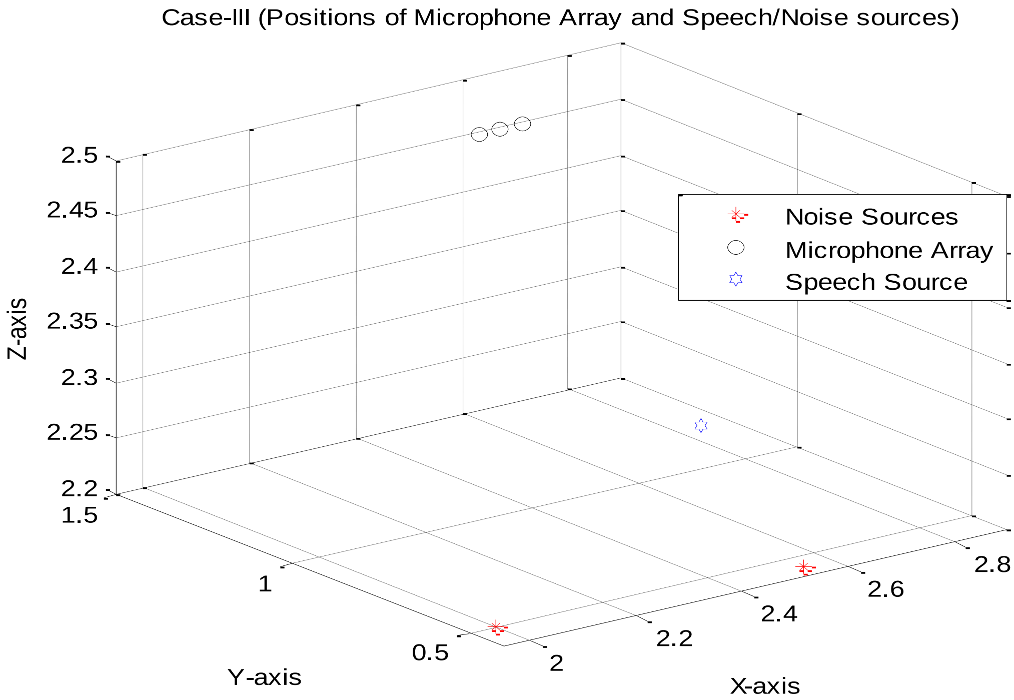


Figure 17: case 3 setup

Evaluation of SNR of the system at WBF 1 for three different cases:

The following table 1 shows the SNR of the system at WBF 1 and SNR improvement for three different cases that has been considered and it is followed by Figures 18 and 19 with their respective bar graphs for the obtained SNR and SNRI values.

At different cases of sources points	Input SNR(dB)	SNR at WBF 1(dB)(Y1)	SNRI(dB)
Case 1	0	46.9180	46.9180
	5	49.7229	44.7229
	10	53.4974	43.4974
	15	57.5437	42.5437
	20	63.4974	43.4974
	25	67.0496	42.0496
Case 2	0	31.3221	31.3221
	5	36.2719	31.2719
	10	41.2489	31.2489
	15	46.2201	31.2201
	20	51.1013	31.1052
	25	55.8644	30.6644
Case 3	0	17.2890	17.2890
	5	27.3055	16.1867
	10	31.7312	15.7106
	15	35.6123	14.6417
	20	40.0929	14.0723
	25	44.0723	14.0540

Table 1: Evaluation of SNR at three different cases at WBF 1(Y1)

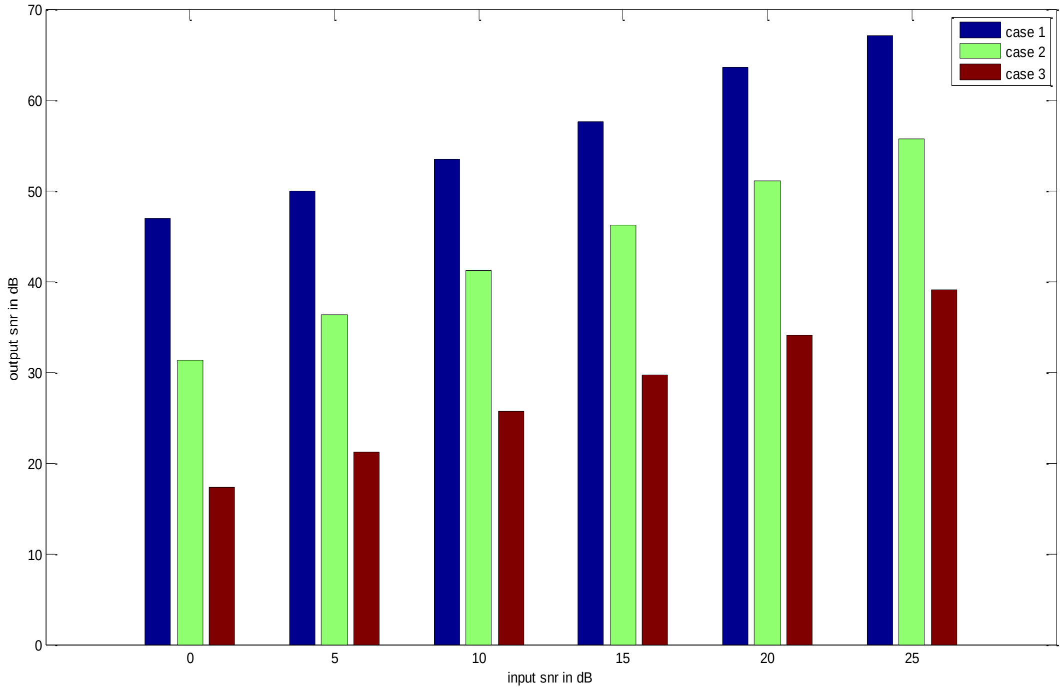


Figure 18: Plot between Input SNR and SNR at WBF 1 for three different cases

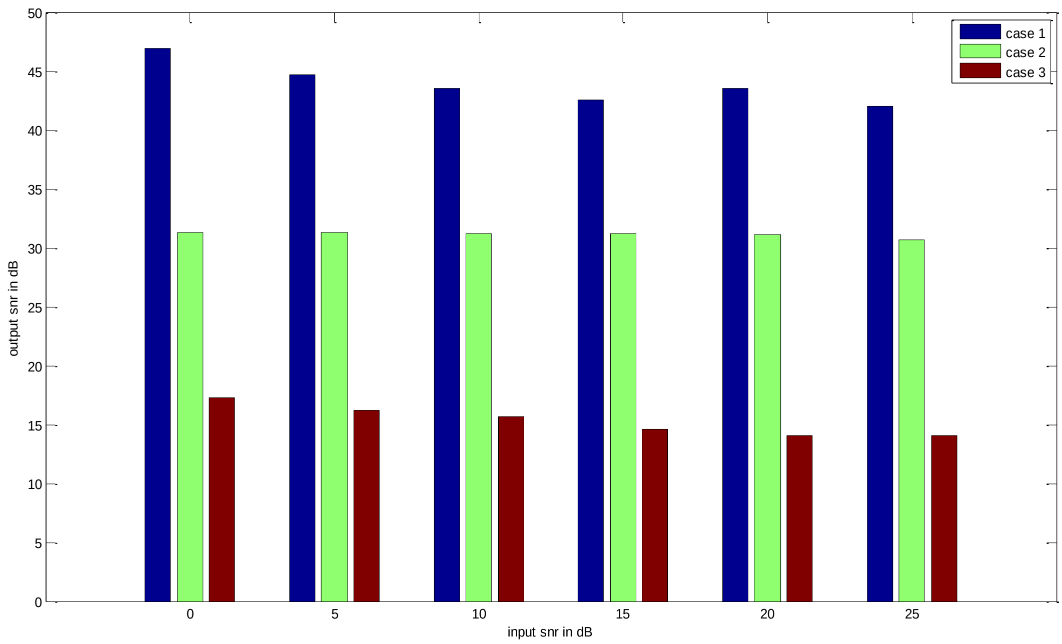


Figure 19: Plot between Input SNR and SNR Improvement for WBF 1 at three different Cases

Table for the SNR at the output (Evaluating entire system) for three different cases

The following table 2 shows the SNR for the system and SNR improvement for three different cases that has been considered and it is followed by Figures 21 and 22 with their respective bar graphs for the obtained SNR and SNRI values.

At different cases	Input SNR(dB)	Output SNR (dB) for the entire system	SNRI (dB)
Consideration1	0	52.9386	52.9386
	5	56.7435	51.7435
	10	59.8679	49.8679
	15	63.5180	48.5180
	20	67.5643	47.5643
	25	71.0702	46.0702
Consideration2	0	37.3427	37.3427
	5	42.2925	37.2925
	10	47.2659	37.2659
	15	52.2407	37.4070
	20	57.1219	37.1219
	25	61.6850	36.6850
Consideration3	0	23.2120	23.2120
	5	27.3055	22.3055
	10	31.7312	21.7312
	15	35.6123	20.6129
	20	40.0929	20.0929
	25	44.0723	19.0723

Table 2: SNR Evaluation at o/p of system at three different cases (e)

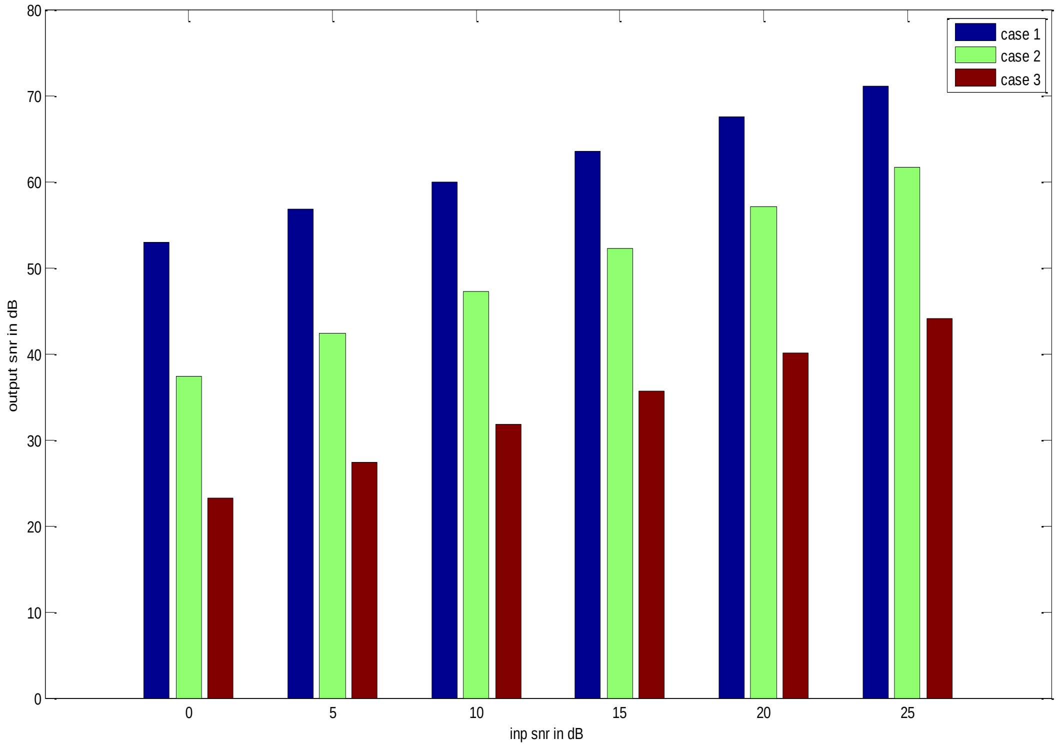


Figure 21: Plot B/w input SNR and Output SNR at three different cases for entire System

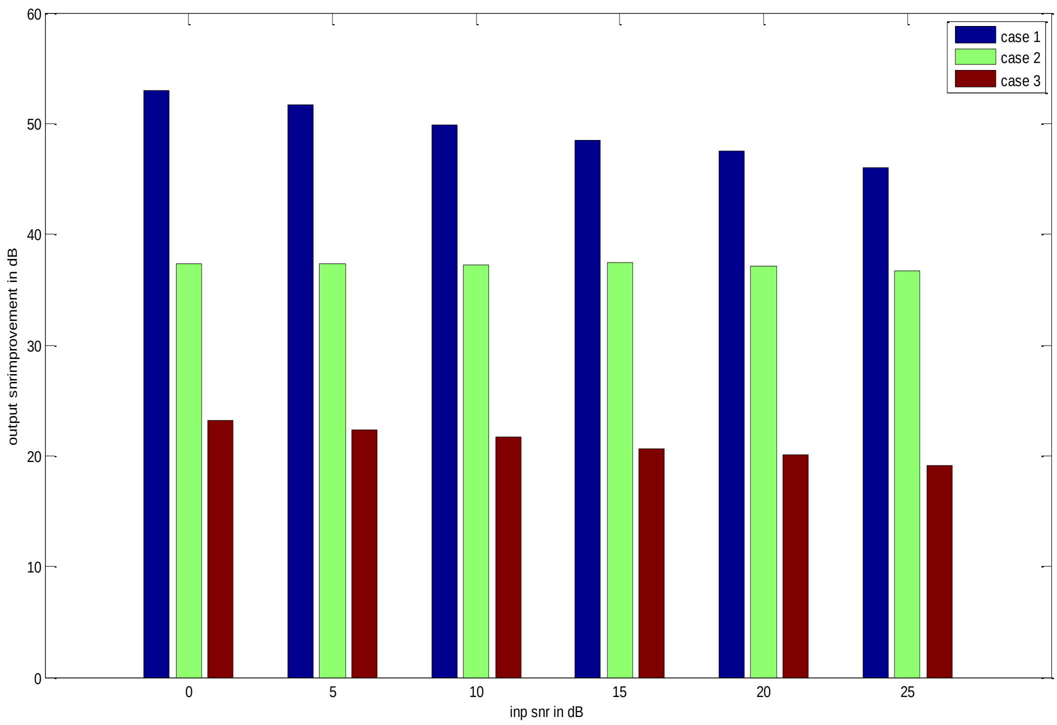


Figure 20: Plot B/w i/p SNR & SNR Improvement at three different cases for entire syste

Evaluation of the system:

Comparison of SNR at WBF 1 and system Output:

The following table 3 shows SNR and SNRI of the system at WBF 1, and SNR, SNRI for entire system and it is followed by respective bar plots(figure 24 and 25) for obtained SNR and SNRI values.

At one fixed case	Input SNR (dB)	Output SNR (dB)	SNRI (dB)
Wiener Beamfor 1(y1)	0	31.3221	31.3221
	5	36.2719	31.2719
	10	41.2489	31.2489
	15	46.2201	31.2201
	20	51.1013	31.1052
	25	55.8644	30.6644
Entire system	0	37.3427	37.3427
	5	42.2925	37.2925
	10	47.2659	27.2659
	15	52.2407	37.4070
	20	57.1219	37.1219
	25	61.6850	36.6850

Table 3: Comparison of SNR at WBF 1, System out and its improvement

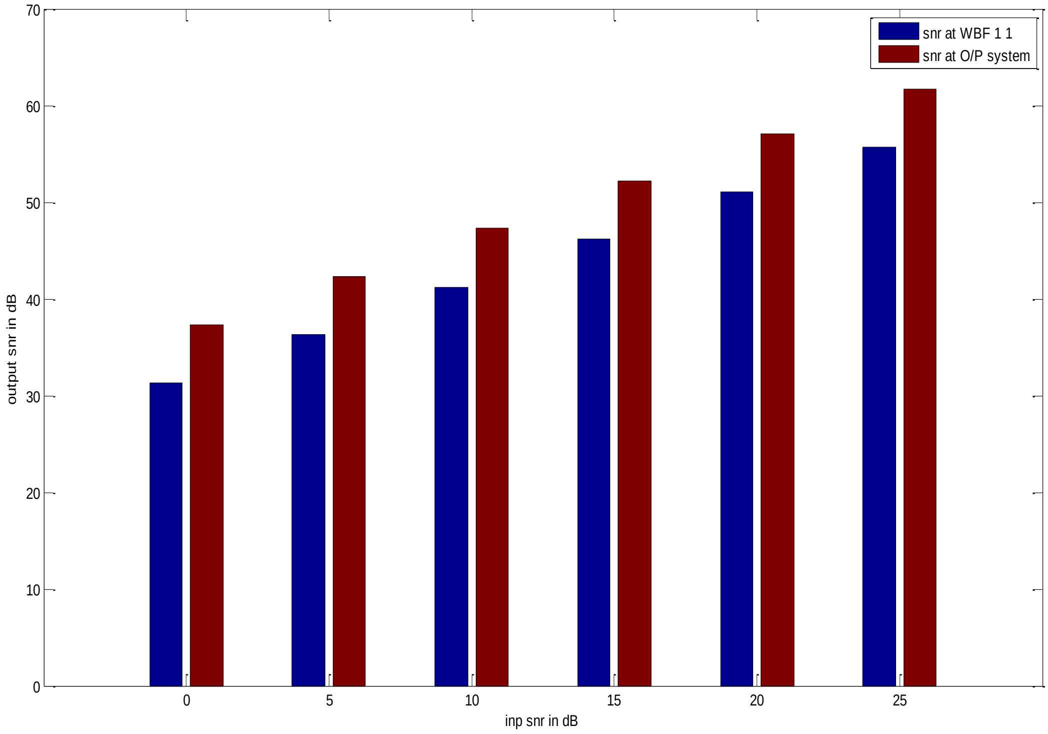


Figure 24: SNR comparison at WBF 1 (Y1) and system o/p

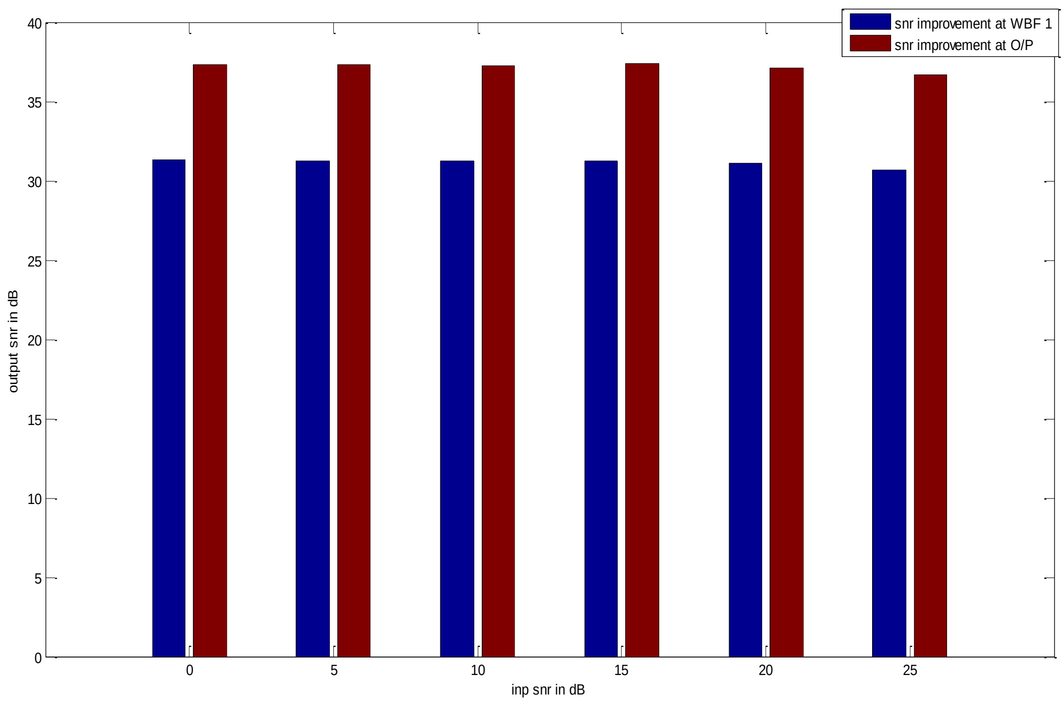


Figure 21: SNR Improvement comparison at WBF 1(Y1) and system o/p

Chapter 7

Conclusion and future work

7.1 Conclusion:

This work is focused on the design and implementation of WBF based generalized side lobe canceller for the enhancement of speech signal from the noisy atmosphere. The system has been implemented and its performance has been analyzed by considering two noisy signals (one is male voice signal and the other is random noise signal with SNR and SNRI as main objective for the evaluation of the proposed system

Simulation Results have been analyzed for different main speech source position's and Noises sources at three different cases. The Results obtained shows that origination of main speech and the noise sources play a major role in effective cancelling of the noise. When the main speech is nearer to the microphone array than that of noise fields, then both SNR and the SNR improvement recorded at WBF 1 and at System output (e) is high and better when compared with, when the noise fields are nearer to the microphone array respectively.

7.2 Future work:

Although I have covered many aspects of Wiener beam forming, and the GSC, there are areas that require further research, showing promise for improved results using a microphone array structure. However, in a real environment, multiple reflections leading to multiple target DOA's can be implemented further. Additional microphones may improve overall performance, especially in the GSC.The same structure can be implemented by taking Reverberation into consideration.

Chapter 8

Bibliography

[1] Claiborne McPheeters, James Finnigan, Jeremy Bass and Edward Rodriguez, “Array Signal Processing: An Introduction” Version 1.6: Sep 12, 2005.

[2]Amin, M.S.; Ahmed-Ur-Rahman; Saabah-Bin-Mahbub; Ahmed, K.I.; Chowdhury, Z.R.; , "Estimation of direction of arrival (DOA) using real-time Array Signal Processing , " Electrical and Computer Engineering,2008.ICECE.2008.International Conference on ,vol.,no.,pp.422-427,2022Dec.2008

[3]Cornelius, P.;Yermeche, Z.;Grbic,N.;Claesson,I.;,"A spatially constrained subband beamforming algorithm for enhancement," Sensor Array and Multichannel Signal Processing Workshop Proceedings, 2004 , vol., no., pp. 89- 93, 18-21 July 2004

[4]Valimaki, V.; Laakso, T.I.; ,"Principles of fractional delay filters," Acoustics, Speech, and Signal Processing, 2000. ICASSP '00. Proceedings. 2000 IEEE International Conference on , vol.6, no., pp.3870-3873 vol.6, 2000

[5]Cain, G.D.; Yardim, A.; Henry, P.; , "Offset windowing for FIR fractional-sample delay," Acoustics, Speech, and Signal Processing, 1995.ICASSP-95.,1995 International Conferenceon ,vol.2, no.,pp.1276-1279 vol.2,9-12 May 1995

[6]Darren B Ward, Rodney A Kennedy, and Robert C Williamson. *Microphone Arrays: Signal Processing Techniques and Applications*,chapter 1.Constant Directivity Beamforming. Springer-Verlag, 2001.

[7]Gray W Elko. *Microphone Arrays: Signal Processing Techniques and Applications*, chapter 17. Future Directions for Microphone Arrays. Springer-Verlag, 2001.

[8] Ryan, J.G.; Goubran, R.A.; , "Array optimization applied in the near field of a microphone array ," *Speech and Audio Processing, IEEE Transactions on* , vol.8, no.2, pp.173-176, Mar 2000

[9] Lloyd J Griffiths and Charles W Jim. An alternative approch to linearly constrained adaptive beamforming. IEEE Transactions on Antennas and Propagation,AP-30(1):27–34, January 1982.

[10]M. Karjalainen and U. K. Laine, "A model for real-time sound synthesis of guitar on a floating-point signal processor," Proc. IEEE Int. Conf. Acoust. Speech Signal Processing (ICASSP-91), vol. 5, pp. 3653-3656, 14-17 May 1991.

[11]Chokkarapu Anil, "Accoustic Fan Noise cancellation in a Laptop Microphone system,"BTH, Karlskrona, Sweden, 2011.

[12]M. F. Pyer and R. Ansari, "The design and apllication of optimal FIR fractionalslope phase filters," Proc. IEEE Int. Conf. Acoust. Speech Signal Processing (ICASSP-87), vol. 2, pp. 896-899, 6-9 April 1987.

[13]T. I. Laakso, V. Valimaki, M.Karjalainen and U. K. Laine, "Splitting the unit delay-tools for fractional delay filter design," IEEE Signal Processing Magazine, vol. 13, no. 1, pp. 30-60, 1996.

[14]J. P. Townsend, K. D. Donohue, "Stability Analysis for the Generalized Sidelobe Canceller," in IEEE Signal Process. Lett. Lexington, KY, USA, June 2010, pp. 603- 606.

[15]Kwong,R.H.;Johnston,E.W.;,"A variable step size LMS algorithm," Signal Processing,IEEE Transactions on ,vol.40,no.7,pp.1633-1642,Jul 1992

[16] B. Widrow,J.M.McCool,M.G.Larimore,and C.R. Johnson, "Stationary and nonstationary learning characteristics of the LMS adaptive filter," Proc.IEEE,vol. 64, pp.1151-1162, Aug. 1976.

[17]B.Widrow,J.R.Glover,Jr.,J.M.McCool,J.Kaunitz.C.S.Wil-liams,RH.Hearn.I.R.Zeidler,E. Dong, J r.,and R.C.Goodlin,"Adaptive noise cancelling: Principles and applications, " Proc.I EEE,vol.63.pp.1692-1716, Dec. 1975

[18]Ting-Ting Li;Min Shi;Qing-MingYi;,"An Improved Variable Step-Size LMS Algorithm," Wireless Communications, Networking and Mobile Computing (WiCOM), 201 7th International Conference on ,vol.,no.,pp.1-4,23-25 Sept.2011

[19]Saxena,G.;Ganesan,S.;Das,M.;,"Real time implementation of adaptive noise cancellation," Electro/Information Technology,2008.EIT 2008.IEEE International Conferenceon ,vol.,no.,pp.431-436,18-20May2008

[20] Bayard, D.S.; , "A modified augmented error algorithm for adaptive noise cancellation in the presence of plant resonances," American Control Conference, 1998. Proceedings of the 1998 , vol.1, no., pp.137-141 vol.1, 21-26 Jun 1998

[21]Xi-Ren Cao; Ruey-Wen Liu; , "General approach to blind source separation," Signal Processing, IEEE Transactions on , vol.44, no.3, pp.562-571, Mar 1996 doi: 10.1109/78.489029

[22] Buchner, H.; Kellermann, W.; , "A Fundamental Relation Between Blind and Supervised Adaptive Filtering Illustrated for Blind Source Separation and Acoustic Echo Cancellation," *Hands-Free Speech Communication and Microphone Arrays,2008. HSCMA2008* ,vol.,no.,pp.17-20,6-8May2008

[23] Sasaki,Y.;Kagami,S.;Mizoguchi,H.;Enomoto,T.;,"A predefined command recognition System using a ceiling microphone array in noisy housing environments," Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on, vol., no.,pp.2178-2184,22-26Sept.2008

[24]Hanyu Li;Yu-Dong Yao; Jin Yu;,"Outage Performance of Wireless Systems with LCMV Beamforming for Dominant Interferers Cancellation,"Communications,2007. ICC '07. IEEE International Conference on,vol.,no.,pp.190-195,24-28 June 2007

[25] Yiteng Huang; Benesty, J.; Jingdong Chen; , "Speech Acquisition and Enhancement in Reverberant,Cocktail-Party-Like Environment," Acoustics,Speech and Signal Processing, 2006.ICASSP 2006 Proceedings.2006 IEEE International Conference on ,vol.5,no.,pp.V,14-19May2006

[26] Florencio, D.A.; Malvar, H.S.; , "Multichannel filtering for optimum noise reduction in microphone arrays," *Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP '01). 2001 IEEE International Conference on*, vol.1, no., pp.197-200 vol.1, 2001