**Spatial-Temporal Subband Beamforming for Near Field
Adaptive Array Processing**

by

Yahong Rosa Zheng, B.Eng., M.Eng.

A thesis submitted to
the Faculty of Graduate Studies and Research
in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

Carleton University
Ottawa, Ontario, Canada, K1S 5B6

©Copyright
2002, Yahong R. Zheng

National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-79447-4

Canada

The undersigned recommend to
the Faculty of Graduate Studies and Research
acceptance of the thesis

Spatial-Temporal Subband Beamforming for Near Field Adaptive Array Processing

submitted by **Yahong Rosa Zheng, B.Eng., M.Eng.**
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

Chair, Department of Systems and Computer Engineering

Thesis Supervisor

Thesis Supervisor

External Examiner

Carleton University

September, 2002

Abstract

This thesis investigates broadband adaptive beamforming for signal targets located in the near field of an array. The primary application of this research is hands-free sound pickup and speech enhancement for wideband telephony. The technical challenges are three-fold. Broadband beamformers are difficult to design due to large frequency dependent beampattern variations and reduced performances for low frequencies. Near field curvature prohibits the simplified far field assumption and many established far field beamforming techniques are not applicable to near field beamforming. Conventional adaptive beamformers experience desired signal cancellation in reverberant environments where coherent interference is dominant.

As a compromise solution to the three problems encountered in *near field broadband adaptive* beamforming, a Spatial-Temporal Subband (STS) adaptive beamforming structure has been proposed in this thesis. It incorporates a spatial subband array with temporal subband multirate filters and employs a near field adaptive beamformer in each subband. It enables parallel processing of the subband systems, improves the computational efficiency and enhances the performances of the near field broadband beamformers. Three specific STS adaptive beamformers are developed, namely (1) the Nested Array Quadrature Mirror Filter (NAQMF) beamformer which uses a nested array with critically sampled QMF banks and near field Generalized Sidelobe Canceler (GSC) adaptive beamformers, (2) the Nested Array Multirate Generalized Sidelobe Canceler (NAM-GSC) which uses a nested array with non-critically sampled multirate filter banks and near field GSC adaptive beamformers, and (3) the Nested

Array Switched Beam Adaptive Noise Canceler (NASB-ANC) which incorporates a nested array with non-critically sampled multirate filter banks and near field Delay-Filter-and-Sum beamformers followed by adaptive noise cancelers. The three STS systems are shown, via computer simulation and experimental evaluation, to reduce the frequency dependent beampattern variations to the extent which occurs within an octave frequency band. They can achieve higher noise reduction using less adaptive weights than the fullband beamformers. They can improve the convergence of adaptation and reduce the computational complexity. The use of near field beamforming also improves the de-reverberation performance of the STS systems.

Several new algorithms are also proposed in the thesis. A simplified implementation is developed for GSC adaptive beamformers to reduce the computational load by 80%. A robust near field GSC design method is developed to improve the robustness of the near field adaptive beamformer against the location errors. A near field Spatial Affine Projection (SAP) algorithm is proposed for adaptive beamformers to suppress coherent interferences and combat desired signal cancellation.

Acknowledgments

The research work reported in this thesis was carried out with the Department of Systems and Computer Engineering at Carleton University, Ottawa, Canada, from September 1997 to June 2002. I am most grateful to my supervisors, Professor Rafik A. Goubran and Professor Mohamed El-Tanany. I truly appreciate their valuable encouragement and guidance during the course of the research, their generous support on research funding, their effort on providing numerous opportunities of academic interactions with industry partners, and their understanding of the special challenges that I have encountered.

I gratefully acknowledge the financial support from **Communications and Information Technology Ontario (CITO)**, Canada, and the **Ontario Graduate Scholarship in Science and Technology** (1998–1999) from the Ontario Ministry of Education and Training, and the **Nortel Networks Scholarship** (2000–2002) from Nortel Networks Inc., Ottawa, Canada. I would also like to acknowledge the support of **Research Assistantship** awarded by the Faculty of Graduate Studies and Research, Carleton University and the Department.

I would also like to extend my appreciation to Mrs. Christine Lariviere and Dr. Osamu Hoshuyama for their careful proofreading of the manuscript, to Mr. Marco Nasr for helping with experimental recordings, to Mr. Lijing Ding for helping with DSLA test, and to Dr. James (Jim) G. Ryan for his helpful suggestions and discussions at the initial stage of this research.

I am indebted to my father Zhengfu Zheng and my mother Kaiyun Su, who have

set my roots and given me wings, and who have always believed in me and encouraged me. I am equally indebted to my parents and my mother-in-law for their valuable support to my family and their priceless loving care for my children.

I am particularly thankful to my daughter Fangjian and my son David for bearing with me through the “long school years without vacations”.

My deepest gratitude goes to my husband Dr. Chengshan Xiao, whose enthusiasm on research gained my admiration and inspired the idea of doing my Ph.D. program. Throughout the years, he not only showed great patience and understanding but also encouraged me to work for excellence.

Contents

Abstract	iii
Acknowledgments	v
Contents	vii
List of Tables	xi
List of Figures	xiii
List of Abbreviations	xviii
1 Introduction	1
1.1 Array Beamforming	1
1.2 Problem Statement and Research Objectives	2
1.3 Outline of the Thesis	6
1.4 Summary of Contributions	8
2 Introduction to Near Field Array Processing	10
2.1 Signals in Space and Time	10
2.1.1 Plane Waves and Spherical Waves	11
2.1.2 Signals Received at Sensor Array	13
2.2 Array Beamforming Basics	15
2.2.1 Beamforming and Spatial Filtering	15

2.2.2	Fixed Beamforming via Weight Selection	21
2.2.3	Adaptive Beamforming via Weight Selection	24
2.3	Near Field Beamforming	30
2.3.1	Near Field versus Far Field Beamforming	30
2.3.2	Distance Criterion for Near/Far Field Assumption	33
2.3.3	Near Field Fixed Beamforming Techniques	35
2.3.4	Near Field Adaptive Beamforming Techniques	39
3	Overview of Broadband Adaptive Beamforming	42
3.1	Technical Challenges in Broadband Adaptive Beamforming	43
3.1.1	Frequency Dependent Beampattern Variation	43
3.1.2	Desired Signal Cancellation Phenomena	44
3.2	Current Approaches to Broadbanding	48
3.2.1	Regular Array Weight Selection Approach	48
3.2.2	Unequally Spaced Array Design Approach	49
3.2.3	Nested Array Approach	51
3.3	Current Approaches to De-reverberation	54
3.3.1	Decorrelation Preprocessor	54
3.3.2	Robust Beamforming	58
4	Near Field Spatial-Temporal Subband Beamforming Systems	60
4.1	Near Field STS Adaptive Beamforming	61
4.1.1	General Structure of the STS Beamforming Systems	61
4.1.2	Advantages of the STS Beamforming Systems	68
4.1.3	Design and Implementation of the Near Field GSC Adaptive Beamformer	69
4.2	The NAQMF Adaptive Beamformer	75
4.2.1	Design of the NAQMF Beamformer	75
4.2.2	Performances of the NAQMF Beamformer	76

4.2.3	Improvements on the NAQMF Beamformer	81
4.3	The NAM-GSC Adaptive Beamformer	86
4.3.1	Nested Array Multirate Beamformers with Non-critical Sampling	86
4.3.2	Performances of the NAM-GSC Adaptive Beamformer	88
4.3.3	Robustness of the NAM-GSC Against Location Errors	93
4.4	The Nested Array Switched Beam Adaptive Noise Canceler	98
4.4.1	General Structure of the NASB-ANC Scheme	98
4.4.2	Performances of the NASB-ANC Scheme	99
5	De-reverberation Performances of the STS Beamformers	106
5.1	Reverberation Modeling	107
5.2	De-reverberation Performances	108
5.2.1	Beampatterns	109
5.2.2	The Signal Power and SINR	114
5.2.3	SNR and NR versus the Frequency	119
5.2.4	Energy Decay Curves	122
5.3	Remarks on De-reverberation Performances	122
6	Spatial Affine Projection (SAP) Algorithm	126
6.1	The SAP Algorithm for Coherent Interference Suppression	127
6.2	Performances of the SAP Algorithm in Far Field Beamforming	132
6.3	Spatial Averaging Algorithms in Near Field Beamforming	139
7	Experimental Evaluation of the STS Beamformers	144
7.1	Description of the Experiment	144
7.1.1	Measurement Apparatus	144
7.1.2	Measurement Procedures and Environments	148
7.2	Data Analysis and Results	151
7.2.1	Noise Reduction Performances	151

7.2.2	De-reverberation Performances	154
7.2.3	The PAMS Test	156
8	Conclusion	160
	Bibliography	166
A	The Image Model	178
B	Affine Projection Algorithms	183
C	List of Publications	188

List of Tables

3.1	Sensor locations of a 17-element Frequency Invariant (FI) linear array	51
4.1	Output power and SINR of the NAM-GSC beamformer and the full-band GSC for noise rejection	92
4.2	Number of constraints (L) and degree of freedom ($N - L$) in the robust GSC beamformer	97
4.3	Output power and SINR of the NASB-ANC and the fullband SB-ANC for noise rejection	102
5.1	Power and SINR of the NAM-GSC beamformers and the NASB-ANC for de-reverberation	114
5.2	Power and SINR of the Four Subarrays for De-reverberation	117
6.1	Summary of the SAP algorithm	132
6.2	Comparison of computational complexity of the SAP and SPSS algorithm	133
7.1	The experimental apparatus	146
7.2	SINR of the NASB-ANC and its subbands for noise rejection using experimental data	155
7.3	The MOS standard	158
7.4	Listening Effort (LE) and Listening Quality (LQ) scores obtained by the PAMS test	158
8.1	Performances of the NAM-GSC and the NASB-ANC via simulation .	164

8.2	Performances of the NAM-GSC and the NASB-ANC via experimental evaluation	164
A.1	The number of low order image sources in a rectangular room	180
B.1	Summary of the AP algorithm	184
B.2	Summary of the FAP algorithm	185
B.3	The simplified FAP algorithm	185
B.4	Alternate formulation of the AP algorithm	187

List of Figures

1.1	Wavefront curvature observed by an array	3
1.2	Reverberation in a rectangular room	5
2.1	Coordinate systems	11
2.2	A plane wave impinging on an array	14
2.3	A spherical wave impinging on an array	14
2.4	Structure of a common broadband beamformer	16
2.5	Near field beamforming at a focus point $\mathbf{x}_f$	17
2.6	Delay-and-Sum beamformer	21
2.7	A frequency domain beamformer	22
2.8	Delay-Filter-and-Sum beamformer	23
2.9	Multiple sidelobe canceler	25
2.10	Generalized sidelobe canceler	28
2.11	Observation paths for near field array response	31
2.12	Near field array response evaluated at different paths	32
2.13	Far field array response evaluated at different paths	32
2.14	Array optimization by stochastic region contraction (SRC)	40
3.1	Frequency dependent beampattern variation for an 11-element ULA	44
3.2	Performances of the conventional adaptive beamformer with correlated interference	46

3.3	Power spectra of the conventional adaptive beamformer with correlated interference	47
3.4	A nested array with inter-sensor spacing ratio = 3	52
3.5	A harmonically nested array with inter-sensor spacing ratio = 2	52
3.6	Subgrouping in the spatial smoothing (SS) algorithm	55
3.7	Array beampattern with and without the SS algorithm	56
3.8	Signal power spectra with and without the SS algorithm.	57
3.9	Block diagram of the CSST adaptive beamformer	58
4.1	Structure of Spatial-Temporal Subband (STS) beamformers	62
4.2	Configuration of an 11-element harmonically nested array	64
4.3	Frequency bands covered by the nested subarrays	64
4.4	Tree-structured QMF filters for critical sampling	66
4.5	Tree-structured analysis and synthesis filters for non-critical sampling	67
4.6	Adaptive beamformer implemented by a Generalized Sidelobe Canceler	72
4.7	Simplified implementation of GSC with pre-steering	74
4.8	Frequency responses of a 3-stage tree-structured QMF bank.	77
4.9	Beampattern variations of the NAQMF beamformer compared to the fullband beamformer with the same array geometry.	79
4.10	Converged nulling beampatterns of the NAQMF beamformer. The desired signal is S_1 and the interfering signals are S_2 and S_3	80
4.11	Excess MSE of the NAQMF adaptive beamformer.	82
4.12	Array geometry of the NAQMF beamformer with 5 subbands.	84
4.13	Beampatterns of the 5-subband NAQMF adaptive beamformer.	85
4.14	Excess MSE of the NAQMF beamformer with 5 subbands.	85
4.15	Frequency responses of the 3-stage tree structure FIR filters	88
4.16	Beampattern variations of the NAM-GSC beamformer compared to the fullband beamformer with the same array geometry.	90

4.17	Noise rejection performances of the NAM-GSC beamformer without location errors, where S_1 is the desired signal, S_2 and S_3 are the interference.	91
4.18	Excess MSE of the NAM-GSC adaptive beamformer using the NLMS algorithm with $\mu = 0.001$	93
4.19	Excess MSE of the NAM-GSC adaptive beamformer using the NLMS algorithm with $\mu = 0.01$	94
4.20	Sensitivity of the NAM-GSC beamformer to signal location errors. . .	95
4.21	Spatial region to be constrained by the robust GSC beamformer . . .	96
4.22	Responses of the robust NAM-GSC adaptive beamformer when the desired signal has small location errors.	98
4.23	Structure of the Switched Beam Adaptive Noise Canceler (SB-ANC) .	100
4.24	Fixed DFS beams of the NASB-ANC with the 11-element nested array	101
4.25	Fixed DFS beams of the 11-element nested array fullband SB-ANC. .	103
4.26	Excess MSE of the NASB-ANC scheme using the NLMS algorithm with $\mu = 0.01$	104
5.1	A nested array in a reverberant room. The figure is not to scale. . . .	108
5.2	De-reverberation beampatterns of the NAM-GSC beamformer $\mathbf{W}_{gsc}$ adapted at the presence of the desired signal.	111
5.3	De-reverberation beampatterns of the NASB-ANC $\mathbf{W}_{anc}$ with its ANCs switched off by a VAD.	112
5.4	De-reverberation beampatterns of the best achievable beamformer $\mathbf{W}_{bst}$ adapted at the absence of the desired signal.	113
5.5	PSD of the low subband beamformer outputs in a reverberant room. .	118
5.6	$SNR(f)$ of the adaptive beamformers in a reverberant room.	120
5.7	Reverberant noise reduction $NR(f)$ of the adaptive beamformers in a reverberant room.	123

5.8	Energy Decay Curves of the adaptive beamformers in a reverberant room.	124
6.1	An adaptive GSC beamformer with a subtractive pre-processor	128
6.2	An adaptive GSC beamformer using Spatial Smoothing (SS) algorithm	130
6.3	An adaptive GSC beamformer using Spatial Affine Projection (SAP) algorithm	131
6.4	Beampatterns of the SAP and SPSS algorithms with far field narrow band coherent interference	134
6.5	Convergence of the SAP and SPSS algorithms with far field narrow band coherent signals.	135
6.6	Responses of the SAP algorithm with far field broadband coherent signals, where $S1$ is the desired signal, $S2, S3$ and $S4$ are the coherent interference.	136
6.7	Responses of the SSPS algorithm with far field broadband coherent signals, where $S1$ is the desired signal, $S2, S3$ and $S4$ are the coherent interferences.	137
6.8	Convergence of the SAP and SPSS-NLMS algorithms with far field broadband coherent signals.	138
6.9	Subgrouping of a near field linear array. The figure is not to scale. . .	140
6.10	Responses of the near field SAP and SS-NLMS algorithm with near field broadband coherent signals.	143
7.1	The multi-channel microphone array recording system	145
7.2	Signal locations in the anechoic chamber	149
7.3	Measurement environment of the conference room	150
7.4	PSD of the three audio input signals. $S1$ was the desired signal. $S2$ and $S3$ were the interference.	153
7.5	Waveforms of the speech signals for de-reverberation	157

A.1	Image model	179
A.2	Impulse response of a reverberant room.	181
A.3	Frequency response of a reverberant room.	181
A.4	Energy decay curve of the room impulse response	182
B.1	General structure of an adaptive filter	183
B.2	Convergence of the FAP algorithm	186
B.3	Decorrelation property of the AP algorithm	187

List of Abbreviations

ANC	Adaptive Noise Canceler
AoA	Angle of Arrival
APA	Affine Projection Algorithms
CD	Compact Disk
CSST	Coherent Signal-Subspace Transformation
DoA	Direction of Arrival
DFS	Delay-Filter-and-Sum (beamformer)
DFT	Discrete Fourier Transform
DSP	Digital Signal Processing
DSL	Digital Speech Level Analyzer
EDC	Energy Decay Curve
FAP	Fast Affine Projection
FI	Frequency Invariant (beamformer)
FIR	Finite Impulse Response
FTF	Fast Transversal Filter
GSC	Generalized Sidelobe Canceler
LCMV	Linearly Constrained Minimum Variance (beamforming)
LMS	Least-Mean-Square
MOS	Mean Opinion Score
MSC	Multiple Sidelobe Canceler
MSE	Mean Squared Error

NAM	Nested Array Multirate
NAQMF	Nested Array Quadrature Mirror Filter (beamformer)
NASB	Nested Array Switched Beam
NLMS	Normalized Least-Mean-Square
NR	Noise Reduction
PAMS	Perceptual Analysis/Measurement System
PSD	Power Spectrum Density
QMF	Quadrature Mirror Filter
RLS	Recursive Least Square
SAP	Spatial Affine Projection
SINR	Signal to Interference and Noise Ratio
SNR	Signal to Noise Ratio
SRC	Stochastic Region Contraction
SS	Spatial Smoothing
SPSS	Subtractive Pre-processor Spatial Smoothing
STS	Spatial-Temporal Subband
SVD	Singular Value Decomposition
TBWP	Time BandWidth Product
TDL	Tapped Delay Line
ULA	Uniform Linear Array
VAD	Voice Activity Detector

Chapter 1

Introduction

1.1 Array Beamforming

Beamforming is a signal processing method using an array of sensors to provide an effective means of spatial filtering. Analogous to a temporal filter which processes data collected over a temporal aperture, a spatial filter processes data received over a spatial aperture, and “filters” signals and interference originating from separate spatial locations.

Since the invention of the acoustic array by Sergeant Jean Perrin in World War I [41, p.2], array beamforming has found a wide range of applications. These include RADAR and air traffic control, SONAR and underwater signal processing, wireless communications and satellite communications, ultrasonic and optical imaging, seismic signal processing in geophysical exploration, microphone arrays in teleconferencing, computer telephony, hearing aids and other biomedical applications.

The primary application of the research in this thesis is microphone array beamforming for hands-free speech and audio pick-up. The convenience and safety provided by hands-free communications is desirable in many fields, such as teleconferencing, teleworking, computer telephony, wireless communication in automobiles, and voice-only data entry, etc. However, hands-free recording systems may suffer from degrada-

tion of sound quality, due to reverberation introduced in acoustic environments and interference generated by loudspeakers and other disturbing sources. These interfering sources usually originate from separate spatial locations, thus can be suppressed by various spatial filtering techniques. Compared to the directional microphone approach and the conventional adaptive noise cancellation method, microphone array beamforming is considered more attractive for spatial as well as temporal filtering [79]. The rapid development of high performance Digital Signal Processors (DSPs) and their low cost also increase the interest in the microphone array beamforming approach for hands-free communications.

1.2 Problem Statement and Research Objectives

The application of microphone arrays generally requires *near field*, *broadband*, *adaptive* beamforming techniques. First of all, *broadband* beamforming is required because of the nature of microphone array applications. The speech and audio signals are broadband, with the band ratio (the ratio of the upper and lower frequency edges of the passband) being 10:1 or larger. Arrays for such broadband signals are more difficult to design than narrowband arrays, because both spatial variables and temporal frequencies have to be taken into account and they are coupled with each other. Furthermore, arrays with a limited number of sensors usually do not provide sufficient spatial sampling. This causes performance degradation due to frequency dependent beampattern variations. Frequency dependent beampattern variations exhibit widened mainlobe beamwidth and reduced effective aperture for lower frequencies. This problem can cause frequency distortion of the desired signal and impair the array's capability of suppressing broadband interference.

Meanwhile, many microphone applications require near field beamforming. In such applications as teleconferencing, hands-free telephony and voice-only data entry, signal sources are located well within the near field of the array. The wavefront

Figure 1.1: Wavefront curvature observed by an array

curvature can be significant within the array's aperture [79], as illustrated in Figure 1.1. Despite this, the majority of beamforming literature assumes that all sources are located far away from the array, and all waves impinging on the array are planar. This far field assumption greatly simplifies beamformer design and research. But using the far field assumption in the near field of an array can result in severe degradation in array performance [45].

Near field beamforming imposes greater technical challenges than its far field counterpart. Because of spherical wave propagation, the received signals at the array undergo complicated changes in magnitude and phase which are non-linear functions of the source/sensor locations. This nonlinear relationship increases the difficulties in near field beamforming. Many established far field beamforming techniques can not be extended to near field array. Several recent reports [45, 79] have shown that exploiting the near field curvature can significantly enhance the performance of near

field arrays. Researches in this emerging area are challenging and promising.

Adaptive beamforming is also desirable for microphone array applications. By adaptive beamforming, we mean that the beamformer weights are changing adaptively according to the statistics of the signals, interference and noises. In contrast, a *fixed* beamformer chooses its weights independently of the received data. A fixed beamformer with super-directive gain may be sufficient for speech pickup in applications where the location of the desired signal is known a priori. However, for a given number of array elements, an adaptive system usually achieves better performance than a comparable fixed beamformer, due to the fact that the background noises and interference change from time to time. Environment setups also vary greatly from case to case. A fixed beamformer designed for one room may not function properly in another room. Adaptive beamformers can provide the flexibility of implementation and the capacity of noise suppression in these situations.

However, adaptive beamformers experience great technical difficulties in real acoustic environments. Particularly, the desired signal cancellation [100][79, p.6] is significant because of the strong reverberation in small enclosures. In typical offices, reverberant interference contributes 10 dB to 15 dB more power than electret microphone background noises [36, chapter 43]. Reverberant signals are the reflected sound waves of the direct path signal, as shown in Figure 1.2. They are strongly correlated with the desired direct path signal and can cause cancellation of the desired signal in adaptive beamforming.

The conventional solution to the correlated interference problem is to add a white noise of comparable power in the data covariance matrix. Although this method guarantees the proper functioning of the adaptive system, it has no de-reverberation gain because the output Signal-to-Interference-and-Noise Ratio (SINR) is limited to the input power ratio of the direct path signal and the reflected signals. Recently, strong research interest in this area is focused on new approaches which either decorrelate the coherent interference or (partially) eliminate the effect of the coherent

Figure 1.2: Reverberation in a rectangular room

interference.

Based on the discussion above, this thesis addresses the combination of the three problems associated with near field, broadband, adaptive array processing. The research objective is to develop some new beamforming scheme which improves the sound quality of hands-free pickup. The developed system is to provide the best trade-off among

- suppressing uncorrelated interference and environmental noises;
- combating reverberation and desired signal cancellation;
- satisfying system requirements for wideband telephony;
- being robust against location error and array imperfections;
- being practical for office implementation and terminal installation.

Our approach to the research goal is to subband the broadband adaptive beamformer in both space domain and time domain. A new structure of spatial-temporal

subband (STS) beamforming system is proposed, which incorporates a harmonically nested array with multirate filter banks. The harmonically nested array spatially samples the source signals into several subbands and each subband array (subarray) is processed adaptively using near field beamforming techniques. Temporal multirate sampling is also employed by each adaptive subarray to enhance the performance and improve the computational efficiency [111]. The STS systems also enable parallel processing of every subband beamformer. Under the main framework of the STS system, three specific nested array multirate (NAM) systems are developed:

1. the Nested Array Quadrature Mirror Filter (NAQMF) beamformer using adaptive beamformers and QMF banks [114];
2. the Nested Array Multirate Generalized Sidelobe Canceler (NAM-GSC) beamformer using adaptive GSC beamformers and non-critical sampling multirate subband filters [113];
3. the Nested Array Switched Beam Adaptive Noise Canceler (NASB-ANC) using fixed beamformers and adaptive noise cancelers (ANC) with non-critical sampling multirate subband filters [115].

The three novel NAM adaptive beamformers are investigated in terms of their noise rejection performance, robustness against location errors, convergence of adaptation and de-reverberation performance.

1.3 Outline of the Thesis

This thesis is organized into 8 chapters. Chapter 1 introduces the idea of near field broadband adaptive beamforming for microphone array applications. It describes the technical problems addressed in this thesis. A summary of thesis contributions is also included.

Chapter 2 provides some background material on near field array processing, including the signal representation in the spherical wave propagation model, the basic concepts and terminologies of beamforming, and the near field beamforming techniques using the established weight selection methods.

Chapter 3 illustrates the two typical problems in broadband adaptive beamforming and the technical challenges associated with them, namely, the frequency dependent beampattern variation in broadband arrays and the desired signal cancellation phenomena with adaptive beamformers. It also presents a review of the current approaches to these problems.

In Chapter 4, a new structure of spatial-temporal subband (STS) beamforming system is proposed for a near field broadband adaptive array. The general structure of the STS system and its design procedures are discussed in detail. Three specific STS systems are designed using nested array multirate beamformers. They are the NAQMF beamformer, the NAM-GSC beamformer, and the NASB-ANC scheme. Their improved performances are demonstrated in terms of noise rejection, convergence of adaptation and robustness against location errors.

The de-reverberation performance of the NAM-GSC beamformer and the NASB-ANC are evaluated in Chapter 5.

Chapter 6 describes the Spatial Affine Projection algorithm proposed for coherent interference suppression. Its design for far field and near field adaptive beamforming is illustrated. The performance of the SAP algorithm is evaluated for far field and near field coherent interference suppression.

Chapter 7 presents the experimental evaluation of the performance of the NAM-GSC beamformer and the NASB-ANC in real room applications. The experimental results agree with the simulation ones and verify the effectiveness of the designs.

Chapter 8 draws the conclusions. A comparison of the NAM-GSC beamformer and the NASB-ANC is also included.

In addition, Appendix A discusses the image model for computer simulation of

room reverberation. Appendix B provides an introduction to the Affine Projection (AP) algorithm, which can be used in the implementation of adaptive beamformers and coherent interference suppression algorithms. A list of publications resulting from this thesis can be found in Appendix C.

1.4 Summary of Contributions

This thesis contributes to the body of knowledge on near field broadband adaptive beamforming. The fundamental contribution of this research is the development of the general structure of the spatial-temporal subband adaptive beamforming systems. The specific contributions include:

1. Incorporation of spatial subband nested arrays with temporal multirate subband filters and development of the three STS systems:
 - the NAQMF beamformer using near field adaptive GSC beamformers and critically sampled QMF banks [114];
 - the NAM-GSC beamformer using near field adaptive GSC beamformers and non-critical sampling multirate subband filters [113];
 - the NASB-ANC scheme using near field fixed beamformers and adaptive noise cancelers (ANC) with non-critical sampling multirate subband filters [115].
2. Evaluation of the performances of the three STS systems in terms of
 - noise rejection,
 - robustness against location error,
 - convergence of adaptation, and
 - de-reverberation performance.

3. Verification of the developed systems through experimental tests in an anechoic chamber and a real reverberant room;
4. Development of specific algorithms to improve the performance and implementation of the near field adaptive GSC beamformer:
 - proposal of a near field robust GSC beamforming design [113];
 - proposal of a simplified implementation for GSC beamformer to improve its computational efficiency [112];
 - innovation of the new Spatial Affine Projection (SAP) algorithm for coherent interference suppression in adaptive beamforming [109];
 - reformulation of the SAP algorithm to near field beamforming for near field coherent interference suppression [116];

Chapter 2

Introduction to Near Field Array Processing

2.1 Signals in Space and Time

In array signal processing, information is carried to the sensors by propagating waves. These signals are thus functions of position as well as time and have properties governed by the laws of physics, in particular the wave equation.

In most situations, a 3-dimensional Cartesian coordinate system representing space, with time being the fourth dimension, is used to describe a space-time signal $s(\mathbf{x}, t)$, where $\mathbf{x}$ denotes the triple of spatial variables (x, y, z) as shown in Figure 2.1. Let the unit vectors in the three spatial directions be $\boldsymbol{\iota}_x$, $\boldsymbol{\iota}_y$ and $\boldsymbol{\iota}_z$, then

$$\boldsymbol{\iota}_x \cdot \boldsymbol{\iota}_x = \boldsymbol{\iota}_y \cdot \boldsymbol{\iota}_y = \boldsymbol{\iota}_z \cdot \boldsymbol{\iota}_z = 1$$

$$\boldsymbol{\iota}_x \cdot \boldsymbol{\iota}_y = \boldsymbol{\iota}_y \cdot \boldsymbol{\iota}_z = \boldsymbol{\iota}_z \cdot \boldsymbol{\iota}_x = 0$$

$$\boldsymbol{\iota}_x \times \boldsymbol{\iota}_y = \boldsymbol{\iota}_z$$

In other situations, spherical coordinates may be used more appropriately to represent space. Here a point $\mathbf{x}$ is represented by its distance r from the origin, its azimuth θ within an equatorial plane containing the origin, and its elevation ϕ down

from the vertical axis (Figure 2.1). The spherical coordinates of a point are related to the Cartesian coordinates by trigonometric formulas:

$$x = r \sin \phi \cos \theta$$

$$y = r \sin \phi \sin \theta$$

$$z = r \cos \phi$$

Figure 2.1: Coordinate systems

2.1.1 Plane Waves and Spherical Waves

The physics of a propagating wave $s(\mathbf{x}, t)$ is described by wave equation. It can be expressed in Cartesian coordinates as [41, p.11]

$$\frac{\partial^2 s}{\partial x^2} + \frac{\partial^2 s}{\partial y^2} + \frac{\partial^2 s}{\partial z^2} = \frac{1}{c^2} \frac{\partial^2 s}{\partial t^2} \quad (2.1)$$

or in spherical coordinates as [14, p.8]

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial s}{\partial r} \right) + \frac{1}{r^2 \sin \phi} \frac{\partial}{\partial \phi} \left(\sin \phi \frac{\partial s}{\partial \phi} \right) + \frac{1}{r^2 \sin^2 \phi} \frac{\partial^2 s}{\partial \theta^2} = \frac{1}{c^2} \frac{\partial^2 s}{\partial t^2} \quad (2.2)$$

where c is the propagation speed. For sound waves in air, $c = 343\text{m/s}$.

The solutions to the wave equations (2.1) and (2.2) provide the commonly used mathematic models of wave propagation: the plane wave model and the spherical wave model. Although other coordinate systems and propagation models exist, it is sufficient to consider these two models for the purpose of beamforming in free-space.

A plane wave is one in which the value of $s(\mathbf{x}, t_0)$, at any instant of time t_0 , is constant over all points on a plane drawn perpendicular to the direction of propagation. Let $s(t)$ be an arbitrary plane wave propagating along direction $\mathbf{v}$ with speed c . The observed signal at a point $\mathbf{x}$ satisfies the wave equation and can be expressed as

$$s(\mathbf{x}, t) = s(t - \mathbf{v} \cdot \mathbf{x}/c) \quad (2.3)$$

And a monochromatic plane wave solution can be written as

$$s(\mathbf{x}, t) = A \exp\{j(\omega t - \kappa \mathbf{v} \cdot \mathbf{x})\} \quad (2.4)$$

where ω is the angular frequency and $\kappa = \frac{\omega}{c}$ is the wavenumber.

In contrast, the spherical wave model is more complicated, as shown in the wave equation (2.2). A general solution to the wave equation in spherical coordinates involves the half integer order spherical Hankel functions and associated Legendre function. In most situations, however, we are only interested in solutions which exhibit spherical symmetry. In these cases, $s(\mathbf{x}, t)$ does not depend on θ or ϕ . So the wave equation (2.2) can be simplified as

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial s}{\partial r} \right) = \frac{1}{c^2} \frac{\partial^2 s}{\partial t^2} \quad (2.5)$$

and the monochromatic solution is

$$s(r, t) = \frac{A}{r} \exp\{j(\omega t - \kappa r)\} \quad (2.6)$$

Now suppose that a spherical wave $s(t)$ with an arbitrary shape is propagating outwards from a point $\mathbf{x}_0$. The observed signal at another point $\mathbf{x}_1$ satisfies (2.2) and can be expressed as

$$s(\mathbf{x}_1, t) = \frac{s(t - |\mathbf{x}_1 - \mathbf{x}_0|/c)}{|\mathbf{x}_1 - \mathbf{x}_0|} \quad (2.7)$$

2.1.2 Signals Received at Sensor Array

Array processing algorithms vary according to whether the signal sources are located in the *far field* or in the *near field*. If the source is far away from the array and the direction of propagation is approximately equal at each sensor, then the propagating field within the array aperture consists of plane waves. If the source is located in the near field of an array, then the wave front of the propagating wave is perceptibly curved with respect to the dimensions of the array, and the propagation direction depends on the sensor location. In this case, the spherical wave model must be used for array processing.

Let us consider a sensor array having M elements, located on a plane at $\{\mathbf{x}_m = (r_m, \theta_m); m = 1, 2, \dots, M\}$. It is conventional to choose the origin of the coordinate system to be the phase center of the array, that is

$$\sum_{m=1}^M \mathbf{x}_m = 0 \quad (2.8)$$

First, assume a plane wave impinging on the array from direction $\mathbf{v}$ with propagation speed c , as shown in Figure 2.2. Note that $\mathbf{v}$ has unit norm and angle θ_s . Let the signal observed at the origin be $s(t)$. Then the received signal at the m th sensor is

$$\begin{aligned} u_m(t) &= s(t - \mathbf{x}_m \cdot \mathbf{v}/c) \\ &= s(t - r_m \cos(\theta_m - \theta_s)/c) \end{aligned} \quad (2.9)$$

If there are D plane waves $\{s_i(t), i = 1, 2, \dots, D\}$ impinging on the array from directions $\Theta_s = [\theta_{s1}, \theta_{s2}, \dots, \theta_{sD}]$, then the signal received at the m th sensor is

$$\begin{aligned} u_m(t) &= \sum_{i=1}^D s_i(t - \tau_{i,m}) \\ \tau_{i,m} &= \frac{r_m}{c} \cos(\theta_m - \theta_{si}) \end{aligned} \quad (2.10)$$

where $\tau_{i,m}$ is the propagation delay of the i th source at the m th sensor.

Figure 2.2: A plane wave impinging on an array

Figure 2.3: A spherical wave impinging on an array

Secondly, if the signal sources are in the near field of the array, then the spherical wave model is used, as illustrated in Figure 2.3. Assume the signal sources $s_i(t)$ are located at $\mathbf{x}_{si} = (r_{si}, \theta_{si})$. The received signal at m -th sensor is then

$$\begin{aligned} u_m(t) &= \sum_{i=1}^D \frac{s_i(t - |\mathbf{x}_m - \mathbf{x}_{si}|/c)}{|\mathbf{x}_m - \mathbf{x}_{si}|} \\ |\mathbf{x}_m - \mathbf{x}_{si}| &= \sqrt{r_{si}^2 + r_m^2 - 2r_{si}r_m \cos(\theta_m - \theta_{si})} \end{aligned} \quad (2.11)$$

If the signal sources are monochromatic, then

$$u_m(t) = \sum_{i=1}^D \frac{A}{|\mathbf{x}_m - \mathbf{x}_{si}|} \exp\{j(\omega t - \kappa|\mathbf{x}_m - \mathbf{x}_{si}|)\} \quad (2.12)$$

2.2 Array Beamforming Basics

In this section, we will first review the basic ideas of beamforming and spatial filtering, then discuss some well established beamformer design techniques.

2.2.1 Beamforming and Spatial Filtering

The primary goal of array beamforming is to pass the desired signal within the band of interest with specified gain and phase, while suppressing the interfering signals originating from different spatial locations and/or occupying different frequency bands.

Figure 2.4 depicts a common broadband beamformer where each element attaches a tapped delay line of length K . A narrowband beamformer may be considered as a special case where K is set to 1. To determine whether an array or a signal source is narrowband or broadband, the observation *Time BandWidth Product* (TBWP) is used as the fundamental parameter [91]. An array is considered narrow band if the observation TBWP is much less than one for all possible source directions. Otherwise, it is a broadband processing. The observation TBWP is denoted ρ and defined as the product of the signal bandwidth and the temporal aperture of the source propagating across the array.

$$\rho = B \cdot T_a \quad (2.13)$$

where $B = (\omega_b - \omega_a)/(2\pi)$ is the bandwidth of the signal source, and T_a is the time interval over which the signal is propagating across the array.

Figure 2.4: Structure of a common broadband beamformer

Let the M -dimensional snapshot vector of the signal received at the sensor array be

$$\mathbf{u}(k) = [u_1(k), u_2(k), \dots, u_M(k)]^T \quad (2.14)$$

and the $N(= MK)$ -dimensional vector of the concatenated snapshot samples be

$$\mathbf{U} = [\mathbf{u}^T(k), \mathbf{u}^T(k-1), \dots, \mathbf{u}^T(k-K+1)]^T \quad (2.15)$$

where superscript $(\cdot)^T$ represents transpose. Then the beamformer output $v(k)$ is a linear combination of the sensor outputs and can be expressed in matrix form as

$$v(k) = \mathbf{W}^H \mathbf{U}(k) \quad (2.16)$$

where $(\cdot)^H$ represents complex conjugate transpose, and $\mathbf{W}$ is the concatenated weight vector defined as

$$\begin{aligned}\mathbf{W}^H &= [\mathbf{w}_1^H, \mathbf{w}_2^H, \dots, \mathbf{w}_K^H] \\ \mathbf{w}_l^H &= [W_{1l}^*, W_{2l}^*, \dots, W_{Ml}^*] \quad l = 1, 2, \dots, K.\end{aligned}\tag{2.17}$$

where $(\cdot)^*$ denotes complex conjugate.

The performance of a beamformer is evaluated by its beamformer response. Similar to the impulse response of a Finite Impulse Response (FIR) filter, the beamformer response of an array is defined as the amplitude and phase presented to a monochromatic complex plane wave as a function of frequency and location. Location is three dimensional in general for near field beamforming. Let the input signal $s(t)$ be a monochromatic spherical wave $e^{j\omega t}$ with an angular frequency ω . It originates from a point $\mathbf{x}_s = (r_s, \theta_s, \phi_s)$, as depicted in Figure 2.5. A near field beamformer is to focus at a point $\mathbf{x}_f = (r_f, \theta_f, \phi_f)$, by compensating for the curved wavefront propagation delay. This is accomplished by choosing the delays

$$\Delta_m = (r_f - |\mathbf{x}_m - \mathbf{x}_f|)/c\tag{2.18}$$

Figure 2.5: Near field beamforming at a focus point $\mathbf{x}_f$

The beamformer output is then

$$\begin{aligned}
 v(t) &= \sum_{m=1}^M w_m u_m(t - \Delta_m) \\
 &= \sum_{m=1}^M \frac{w_m}{r_{ms}} s\left(t - \frac{r_{ms} + r_f - r_{mf}}{c}\right) \\
 &= \sum_{m=1}^M \frac{w_m}{r_{ms}} \exp\{j\omega t - j\kappa(r_{ms} + r_f - r_{mf})\} \\
 &= e^{j\omega t} \sum_{m=1}^M \frac{W_m^*(\omega)}{r_{ms}} \exp\{-j\kappa r_{ms}\}
 \end{aligned}$$

where κ is the wavenumber, $W_m(\omega)$ is the beamformer weights, r_{ms} is the distance between the sensor $\mathbf{x}_m$ and the source $\mathbf{x}_s$, and r_{mf} is the distance between the sensor $\mathbf{x}_m$ and the focal point $\mathbf{x}_f$.

$$\begin{aligned}
 \kappa &= \omega/c \\
 W_m(\omega) &= w_m \exp\{j\kappa(r_f - r_{mf})\} \\
 r_{ms} &= |\mathbf{x}_m - \mathbf{x}_s| \\
 &= \sqrt{r_s^2 + r_m^2 - 2r_s r_m \cos(\theta_m - \theta_s)} \\
 r_{mf} &= |\mathbf{x}_m - \mathbf{x}_f| \\
 &= \sqrt{r_f^2 + r_m^2 - 2r_f r_m \cos(\theta_m - \theta_f)}
 \end{aligned}$$

The beamformer response becomes

$$\begin{aligned}
 b(\mathbf{x}_s, \omega) &= \sum_{m=1}^M W_m^*(\omega) \frac{r_s}{r_{ms}} \exp\{-j\kappa(r_{ms} - r_s)\} \\
 &= \mathbf{W}^H \mathbf{a}(\mathbf{x}_s, \omega)
 \end{aligned} \tag{2.19}$$

The near field steering vector $\mathbf{a}(\mathbf{x}_s, \omega)$ for the source located at $\mathbf{x}_s$ is defined as

$$\mathbf{a}^H(\mathbf{x}_s, \omega) = \frac{r_s}{e^{-j\kappa r_s}} \left[\frac{e^{-j\kappa r_{1s}}}{r_{1s}}, \frac{e^{-j\kappa r_{2s}}}{r_{2s}}, \dots, \frac{e^{-j\kappa r_{Ms}}}{r_{Ms}} \right] \tag{2.20}$$

If a tapped delay line is attached to each sensor, as depicted in Fig. 2.4, the steering vector becomes a concatenated $N \times 1$ vector

$$\mathbf{a}^H(\mathbf{x}_s, \omega) = \frac{r_s}{e^{-j\kappa r_s}} \left[\frac{e^{-j\kappa r_{1s}}}{r_{1s}}, \dots, \frac{e^{-j\kappa r_{Ms}}}{r_{Ms}}, \frac{e^{-j\kappa(r_{1s} + cT)}}{r_{1s}}, \dots, \frac{e^{-j\kappa(r_{Ms} + (K-1)cT)}}{r_{Ms}} \right] \tag{2.21}$$

In far field cases, however, three-dimensional (3-D) location is reduced to one (or two) dimensional direction of arrival (DoA). Let $s(t) = e^{j\omega t}$ be the monochromatic complex plane wave with direction of arrival θ . The far field beamformer output due to $s(t)$ can be simplified as

$$\begin{aligned} v(k) &= e^{j\omega k} \sum_{m=1}^M \sum_{l=0}^{K-1} W_{m,l}^* e^{-j\omega(\tau_{s,m}+l)} \\ &= e^{j\omega k} \mathbf{W}^H \mathbf{a}(\theta, \omega) \end{aligned}$$

The far field beamformer response is then a function of θ and ω

$$b(\theta, \omega) = \mathbf{W}^H \mathbf{a}(\theta, \omega) \quad (2.22)$$

where $\mathbf{a}(\theta, \omega)$ is the far field steering vector

$$\mathbf{a}(\theta, \omega) = [e^{j\omega\tau_{s,1}}, \dots, e^{j\omega\tau_{s,M}}, e^{j\omega(\tau_{s,1}-1)}, \dots, e^{j\omega(\tau_{s,M}-K+1)}]^H \quad (2.23)$$

$$\tau_{s,m} = \frac{r_m}{c} \cos(\theta_m - \theta), \quad m = 1, 2, \dots, M. \quad (2.24)$$

Common to both near field and far field beamforming, the vector notation introduced in (2.19) and (2.22) suggests a vector space interpretation of beamforming. The weight vector $\mathbf{W}$ and the steering vector $\mathbf{a}(\mathbf{x}_s, \omega)$ are vectors in an N -dimensional vector space. The angles between $\mathbf{W}$ and $\mathbf{a}(\mathbf{x}_s, \omega)$ determine the array response $b(\mathbf{x}_s, \omega)$. If the angle between $\mathbf{W}$ and $\mathbf{a}(\mathbf{x}_s, \omega)$ is 90° for some $(\mathbf{x}_s, \omega)$, then the beamformer response is zero. If the angle is close to 0° , then the response magnitude will be relatively large.

The *beam pattern* is defined as the magnitude squared of $b(\mathbf{x}_s, \omega)$. The weight coefficients in $\mathbf{W}$ affect both temporal and spatial responses of the beamformer. As a multiple input single output system, a beamformer is a spatio-temporal filter which is a result of mutual interaction between spatial and temporal sampling.

The general effects of spatial sampling are similar to temporal sampling. Spatial aliasing corresponds to an ambiguity in source locations. This occurs when $\mathbf{a}(\mathbf{x}_{s1}, \omega_1) = \mathbf{a}(\mathbf{x}_{s2}, \omega_2)$, that is, a source at one location and frequency cannot be distinguished from a source at a different location and frequency. For example, spatial

aliasing occurs in a Uniform Linear Array (ULA) when inter-element spacing is larger than a half wavelength of the highest frequency of interest, in which case, grating lobes (periodic repetitions of the main beam) occur in the array beampattern.

A primary focus of beamforming research is on designing response via weight selection. Beamformers can be classified as either data independent (fixed) or statistically optimum (adaptive), depending on how the weights are chosen. The weights in a fixed beamformer do not depend on the array input data. They are chosen to present a specified response for all signal and interference scenarios. Fixed beamformers allow relatively simpler design and implementation, with the ability of interference suppression to some extent. The weights in an adaptive beamformer are chosen based on the statistics of the array data to *optimize* the array response. An adaptive beamformer places nulls in the directions of interfering signals in an attempt to minimize the interference and noise power at the beamformer output. These two types of beamformers will be discussed in some detail in Section 2.2.2 and Section 2.2.3. The general principles described in these two sections are applicable to both near field and far field beamforming, unless specified otherwise.

Besides weight selection, the beampattern equations and the steering vectors indicate that beamformer response is also a function of array geometry. Sensor locations provide additional degrees of freedom in designing a desired response. When sensor locations are selected properly, the steering vector can be well dispersed in the N dimensional vector space over the range of $(\mathbf{x}_s, \omega)$ of interest, and the ability to discriminate between sources at different $(\mathbf{x}_s, \omega)$ will be increased, especially for broadband signals. Utilization of these degrees of freedom is very complicated due to the multi-dimensional nature of spatial sampling and the nonlinear relationship between $b(\mathbf{x}_s, \omega)$ and sensor locations. We will discuss this further in Chapter 3.

2.2.2 Fixed Beamforming via Weight Selection

The weights in a fixed beamformer are designed so the beamformer response approximates a desired response independent of the array input data. This design objective is the same as that for classical FIR filter design. The analogies between beamforming and FIR filter design have been exploited to develop a series of array design methods.

Delay-and-Sum Beamforming

A classical beamforming method for narrowband signals is delay-and-sum. Assume a desired signal with frequency ω_0 is impinging on the array from a known location $\mathbf{x}_0$. The beamformer weight vector $\mathbf{W}$ has to be equal to the steering vector $\mathbf{a}(\mathbf{x}_0, \omega_0)$. In other words, the received signal at each sensor is phase shifted prior to summation, as shown in Figure 2.6. The main beam may be steered electronically to different spatial locations with the pre-steering processors Δ_m , but the beamformer weights w_m usually remain unchanged; so does the beampattern. If the array is linear equispaced, then the beamformer is equivalent to a 1-D FIR filter and the same techniques for choosing tap weights w_m are applicable to either problem.

Figure 2.6: Delay-and-Sum beamformer

Frequency Domain Beamforming

If the beamformer is broadband, two approaches are generally used for beamformer design: frequency domain beamforming and “delay-filter-and-sum” beamforming.

A frequency domain beamformer is implemented by a narrowband decomposition structure, as illustrated in Figure 2.7. A discrete Fourier transform (DFT) is performed for the signals received at each sensor to obtain the frequency domain data. The data at each frequency bin are processed by their own narrowband beamformer $\mathbf{W}_p$, for $p = 1, 2, \dots, P$. With proper selection of $\mathbf{W}_p$ and careful data partitioning, the frequency domain beamformer outputs $v(f_p)$ can be made equivalent to the DFT of the broadband beamformer output in Figure 2.4. This equivalence is analogous to implementing FIR filters by circular convolution with the DFT.

Figure 2.7: A frequency domain beamformer

Delay-Filter-and-Sum Beamforming

A broadband beamformer can also be implemented by delay-filter-and-sum beamforming, as depicted in Figure 2.8. The delays are chosen to steer the beam to

the focal point or the look direction. Then the FIR filter coefficients are designed to approximate a desired temporal response. Spatial and temporal responses of a broadband beamformer interact with each other, so they cannot be synthesized completely independently. Techniques for 2-D FIR filter design are often used for broadband beamformer design.

Figure 2.8: Delay-Filter-and-Sum beamformer

Some established FIR filter design techniques utilizing L_p norm approximation may be exploited. The commonly used techniques are L_∞ (min-max) and L_2 (least squares) optimization, including:

1. Windowing of an ideal filter's impulse response
(minimizes L_2 norm over continuous ω);
2. Frequency sampling and linear weighted least squares
(minimizes L_2 norm over discrete ω);
3. Min-max design with Remez exchange algorithm
(minimizes L_∞ norm over discrete ω);
4. Min-max complex and magnitude response design
(minimizes L_∞ norm over discrete ω).

To illustrate beamformer design via L_2 norm approximation, consider choosing weight vector $\mathbf{W}$ so the actual beamformer response $b(\mathbf{x}, \omega)$ approximates an arbitrary

desired response $b_d(\mathbf{x}, \omega)$. The desired response is then sampled at the P points $\{(\mathbf{x}_p, \omega_p), 1 \leq p \leq P\}$. Choosing P much larger than N (N is the dimension of $\mathbf{W}$), we obtain the over-determined least squares minimization problem

$$\min_{\mathbf{W}} |\mathbf{A}^H \mathbf{W} - \mathbf{b}_d|^2 \quad (2.25)$$

where

$$\begin{aligned} \mathbf{A} &= [\mathbf{a}(\mathbf{x}_1, \omega_1) \ \mathbf{a}(\mathbf{x}_2, \omega_2) \ \cdots \ \mathbf{a}(\mathbf{x}_P, \omega_P)] \\ \mathbf{b}_d &= [b_d(\mathbf{x}_1, \omega_1) \ b_d(\mathbf{x}_2, \omega_2) \ \cdots \ b_d(\mathbf{x}_P, \omega_P)]^H \end{aligned}$$

The solution to (2.25) is classical and can be expressed as [91]

$$\mathbf{W} = \mathbf{A}^\dagger \mathbf{b}_d \quad (2.26)$$

where $\mathbf{A}^\dagger = (\mathbf{A} \mathbf{A}^H)^{-1} \mathbf{A}$ is the pseudo inverse of $\mathbf{A}$.

2.2.3 Adaptive Beamforming via Weight Selection

In adaptive beamforming, the weights are chosen based on the statistics of the data received at the array to optimize the beamformer response so the output contains minimal contributions due to noise and interference. The general assumptions here are

- the data received at the sensors are zero mean, wide sense stationary;
- the signal, interference and noise sources are statistically non-coherent.

Although we often deal with non-stationary data, the wide sense stationary assumption is used in designing optimal beamformers and in evaluating steady state performance.

There are several different approaches for the optimization: Multiple Sidelobe Canceler (MSC), Maximization of Signal-to-Noise Ratio (Max SNR), Linearly Constrained Minimum Variance (LCMV) and Quadratically Constrained Adaptive Beam-

former, etc. We will briefly discuss the different adaptive beamforming schemes with emphasis on the LCMV beamformer.

At this point, it is worth noting that fixed beamformer design techniques are often used in adaptive beamforming. For example, the main channel and auxiliary channels in MSC are often implemented by several fixed beamformers. The constraint design in the LCMV beamforming is essentially a fixed beamformer design, too.

Multiple Sidelobe Canceler

A multiple sidelobe canceler (MSC) consists of a “main channel” and one or more “auxiliary channels”, as shown in Figure 2.9.

Figure 2.9: Multiple sidelobe canceler

The main channel has highly directional response pointing at the desired signal. It can be either a single high gain directional sensor or a fixed beamformer. Interfering signals are presented in the main channel through the sidelobes. The auxiliary channels receive only the interfering signals. The adaptive weights are applied to the auxiliary channels to minimize the total output power and cancel the main channel

interference components. The MSC problem is formulated as

$$\min_{\mathbf{W}_a} E\{|u_d - \mathbf{W}_a^H \mathbf{u}_a|^2\} \quad (2.27)$$

and the optimum solution is

$$\mathbf{W}_{aopt} = \mathbf{R}_a^{-1} \mathbf{p}_{ad} \quad (2.28)$$

where $\mathbf{R}_a = E\{\mathbf{u}_a \mathbf{u}_a^H\}$, $\mathbf{p}_{ad} = E\{\mathbf{u}_a u_d^H\}$.

Minimization of output power can cause cancellation of the desired signal, if the auxiliary channels contain the desired signal components. So MSC is very effective in applications where the desired signal is very weak relative to interference, or when the desired signal is absent during certain time periods. The weights can be adapted in the absence of the desired signal and frozen when it is present.

A good example of the MSC method is *beam-space* adaptive beamforming [27, 82] used in smart antennas. A set of 6 to 12 fixed narrow beams are pre-designed to point at different directions over the spatial aperture. A selector will pick up a beam which contains the strongest component of the desired signal as the main channel, and several other beams as auxiliary channels. Then the MSC method is employed to adaptively filter the signal. To ensure the performance of the MSC, identical beampatterns are required for all fixed beams at all in-band frequencies. So the fixed beamformers are designed using the FAN filter method [82], as we mentioned in Section 2.2.2.

Maximization of Signal-to-Noise Ratio

Maximization of signal-to-noise ratio is formulated as

$$\max_{\mathbf{W}} \frac{\mathbf{W}^H \mathbf{R}_s \mathbf{W}}{\mathbf{W}^H \mathbf{R}_n \mathbf{W}} \quad (2.29)$$

where $\mathbf{R}_s = E\{\mathbf{s}\mathbf{s}^H\}$ and $\mathbf{R}_n = E\{\mathbf{n}\mathbf{n}^H\}$ are covariance matrices of desired signal $\mathbf{s}$ and noise (plus interference) $\mathbf{n}$, respectively. Obviously, prior knowledge of both the desired signal and noise are required or need to be estimated. When $\mathbf{R}_n$ is

nonsingular, the optimum weight vector is obtained for the operating frequency ω as

$$\mathbf{W}_{opt}(\omega) = \mathbf{R}_n^{-1} \mathbf{S}(\omega) \quad (2.30)$$

where $\mathbf{S}(\omega)$ is the spectrum of the desired signal.

Linearly Constrained Minimum Variance (LCMV)

The basic idea behind linearly constrained minimum variance (LCMV) beamforming is to constrain the beamformer response so signals from the direction of interest are passed with specified gain and phase. The weights are chosen to minimize output power or variance subject to response constraints. That is

$$\min_{\mathbf{W}} \mathbf{W}^H \mathbf{R}_u \mathbf{W} \quad \text{subject to} \quad \mathbf{C}^H \mathbf{W} = \mathbf{f} \quad (2.31)$$

where $\mathbf{R}_u = E\{\mathbf{U}(k)\mathbf{U}^H(k)\}$ is $N \times N$ covariance matrix of the received data, $\mathbf{C}$ is the constraint matrix, and $\mathbf{f}$ is the response vector. $\mathbf{C}^H \mathbf{W} = \mathbf{f}$ are a set of linear equations controlling the beamformer response. Each column of $\mathbf{C}$ imposes a linear constraint on the weight vector $\mathbf{W}$ and uses one degree of freedom. With L constraints, $\mathbf{C}$ is $N \times L$ and $\mathbf{f}$ is L -dimensional, and there are $N - L$ degrees of freedom available for adaptation.

The optimum solution to the LCMV beamformer weight vector is

$$\mathbf{W}_{opt} = \mathbf{R}_u^{-1} \mathbf{C} [\mathbf{C}^H \mathbf{R}_u^{-1} \mathbf{C}]^{-1} \mathbf{f} \quad (2.32)$$

Constraint design plays an important role in LCMV beamformer and provides flexible control over beamformer response. Without any constraints, an adaptive array will try to minimize the output power and give the trivial solution of all weights being zero. Several different approaches can be employed for linear constraint design, namely point [43], derivative [20] and eigenvector [8] constraints.

Point constraints specify the beamformer response at points of spatial direction and temporal frequency with fixed gain and phase. It is the most commonly used

constraint design method. Obviously the number of constrained points is limited to N . If N constraints are used, then there are no degrees of freedom left for adaptation and a fixed beamformer is obtained.

Derivative constraints force the derivatives of the beamformer response at some points of direction or frequency to be zero. They are usually employed in conjunction with other constraints to influence the beamformer response over a region of direction or frequency and improve the robustness of the beamformer.

Eigenvector constraints approximate the desired response over regions of direction and frequency in a least squares sense. The beamformer response at a large number of points may be specified, but only a small number of constraints are chosen to minimize the mean-squared error between the desired and actual beamformer response. So eigenvector constraints are very efficient, especially for broadband beamformers.

When an LCMV beamformer is implemented by an adaptive scheme, a Generalized Sidelobe Canceler (GSC) is often used. A GSC consists of a fixed beamformer $\mathbf{W}_q$, a signal blocking matrix $\mathbf{C}_a$ and an unconstrained adaptive weight vector $\mathbf{W}_a$, as illustrated in Figure 2.10. The similarity between GSC and MSC is obvious by comparing Figure 2.10 with Figure 2.9.

Figure 2.10: Generalized sidelobe canceler

The signal blocking matrix $\mathbf{C}_a$ can be obtained from the constraint matrix $\mathbf{C}$, using any of the orthogonalization procedures such as Gram-Schmidt, QR decomposition or singular value decomposition (SVD). The fixed beamformer $\mathbf{W}_q$ is an $N \times 1$ vector, given by

$$\mathbf{W}_q = \mathbf{C}(\mathbf{C}^H \mathbf{C})^{-1} \mathbf{f} \quad (2.33)$$

The unconstrained adaptive weight vector $\mathbf{W}_a$ is updated iteratively using one of the adaptation algorithms, such as the Normalized Least Mean Squares (NLMS) [35, chapter 9], the Recursive Least Squares (RLS) [35, chapter 11] or the family of Affine Projection Algorithms (APA) (see Appendix B). The optimum solution to $\mathbf{W}_a$ is

$$\mathbf{W}_{aopt} = [\mathbf{C}_a^H \mathbf{R}_u \mathbf{C}_a]^{-1} \mathbf{C}_a^H \mathbf{R}_u \mathbf{W}_q \quad (2.34)$$

Quadratically Constrained Adaptive Beamformer

Instead of constraining the weight vector by a set of linear equations in LCMV beamforming, quadratically constrained adaptive beamforming uses constraints in quadratic form of $\mathbf{W}$. Quadratic constraints are often used in conjunction with linear constraints to improve a beamformer's robustness against steering error, or to control the mainlobe response, or to enhance interference suppression capability.

For example, Er and Cantoni[21] proposed a quadratically constrained far field beamformer to control the mainlobe response over a small region $\Delta\theta$ about the look direction θ_0 . The beamformer is formulated as

$$\min_{\mathbf{W}} \mathbf{W}^H \mathbf{R}_u \mathbf{W} \quad (2.35)$$

subject to

$$\mathbf{W}^H (\mathbf{a}_0 \mathbf{a}_0^H + \frac{\Delta\theta^2}{12} \mathbf{a}_1 \mathbf{a}_1^H) \mathbf{W} - (\mathbf{a}_0^H \mathbf{W} + \mathbf{W}^H \mathbf{a}_0) + 1 < \epsilon \quad (2.36)$$

where ϵ is a small value, $\mathbf{a}_0$ and $\mathbf{a}_1$ are the Taylor series of the steering vector $\mathbf{a}(\theta, \omega)$ satisfying

$$\mathbf{a}(\theta, \omega) = \mathbf{a}_0 + (\theta - \theta_0) \mathbf{a}_1 \quad (2.37)$$

Alternative formulations of quadratic constraints were also reported in [92] and [75].

2.3 Near Field Beamforming

In this section, we will discuss the basic difference between near field beamforming and far field beamforming, distance criterion for near/far field assumptions and current approaches to near field beamforming research.

2.3.1 Near Field versus Far Field Beamforming

The majority of array processing literature deals with the case in which signal sources are in the far field of the array. This assumption significantly simplifies the beamformer design problem. In many practical situations, however, signal sources are located well within the near field of the array. This scenario arises in many applications of microphone arrays, such as computer telephony, voice only data entry, mobile telephony and teleconferencing, etc. Using the far field assumption for beamformer design results in severe degradation in the array performance, and near field beamforming has to be employed.

To illustrate the difference between near field and far field beamforming, an example of a 7-element linear array is considered. The array is equi-spaced at the half-wavelength of the operating frequency and is steered at broadside ($\theta = 90^\circ$) of the array.

A near field delay-and-sum beamformer is designed to focus on the point B in Figure 2.11. So we have $\mathbf{x}_f = (r_f, \theta_f) = (0.75(R+d), 90^\circ)$, where d is the inter-element spacing and R is the dimension of the array. For uniform linear arrays, $R = (M-1)d$. The beampatterns are evaluated along the circular paths in Figure 2.11, with radii being $r_1 = r_f$, $r_2 = 2r_f$ and $r_3 = 15r_f$, respectively.

Figure 2.12 shows the beampatterns obtained by the near field beamformer. The

Figure 2.12: Near field array response evaluated at different paths

Figure 2.13: Far field array response evaluated at different paths

or region.

2.3.2 Distance Criterion for Near/Far Field Assumption

As we showed in Figure 2.12 and Figure 2.13, far field beamforming is not a proper method when the signal source is close to the array. Using far field beamforming in the near field of the array will result in severe degradation in performance. Using near field beamforming in the far field of the array will unnecessarily increase the design complexity. An important issue is then the distance criterion for which the far field or near field assumption is valid. This issue has been addressed by several researchers [33, 34, 78, 107], and it is understood that defining the borderline between near field and far field depends on what “negligible error” is.

For spatial filtering purposes, it is found that the error in beampattern due to the far field assumption is closely related to the basic parameter $\frac{R^2}{\lambda}$. To elaborate on this, consider a monochromatic wave source $s(t) = e^{j\omega t}$ emitting from a point $\mathbf{x}_s$. The received signal is given by

$$u_m(t) = \frac{\exp(j\omega t - j\kappa|\mathbf{x}_m - \mathbf{x}_s|)}{|\mathbf{x}_m - \mathbf{x}_s|} \quad (2.38)$$

where

$$|\mathbf{x}_m - \mathbf{x}_s| = \sqrt{r_s^2 + r_m^2 - 2r_s \mathbf{v} \cdot \mathbf{x}_m} \quad (2.39)$$

Let

$$b = \left(\frac{r_m}{r_s}\right)^2 - 2\frac{\mathbf{v} \cdot \mathbf{x}_m}{r_s}. \quad (2.40)$$

Using a *binomial expansion*, it can be shown that

$$\begin{aligned} |\mathbf{x}_m - \mathbf{x}_s| &= r_s \sqrt{1 + b} \\ &= r_s \left(1 + \frac{b}{2} - \frac{b^2}{8} + \dots\right), \quad |b| < 1. \\ &= r_s + \mathbf{v} \cdot \mathbf{x}_m + \frac{r_m^2 - (\mathbf{v} \cdot \mathbf{x}_m)^2}{2r_s} + \frac{r_m^2 (\mathbf{v} \cdot \mathbf{x}_m)}{2r_s^2} - \frac{1}{8} \left(\frac{r_m}{r_s}\right)^4 + \dots \end{aligned}$$

It is satisfactory to approximate the amplitude term in (2.38) by the first term of the binomial expansion and as a result,

$$\frac{1}{|\mathbf{x}_m - \mathbf{x}_s|} \approx \frac{1}{r_s} \quad (2.41)$$

However, it requires the first 3 terms of the binomial expansion to approximate the phase term $\exp(-j\kappa|\mathbf{x}_m - \mathbf{x}_s|)$, since small changes in the range $|\mathbf{x}_m - \mathbf{x}_s|$ can lead to large changes in phase. This leads to the near field expansion of the received signal

$$u_m(t) = \frac{\exp(-j\kappa r_s)}{r_s} \cdot \exp(-j\kappa \mathbf{v} \cdot \mathbf{x}_m) \cdot \exp\left(-j\kappa \frac{r_m^2 - (\mathbf{v} \cdot \mathbf{x}_m)^2}{2r_s}\right) \quad (2.42)$$

The far field assumption uses only the first two terms of (2.42)– the first term is the signal observed at the coordinate origin, the second term is the far field phase adjustment at the sensor. Thus the third term is the quadratic phase error for the far field assumption.

The quadratic phase error takes its maximum value when $\mathbf{v} \cdot \mathbf{x}_m$ is zero, or equivalently, when the angle between $\mathbf{v}$ and $\mathbf{x}_m$ is 90° . Replacing $\mathbf{x}_m$ by the dimension of the array R , we can obtain the quadratic phase error across the array. It has been shown [33, 34] that the distance $\frac{2R^2}{\lambda}$ gives the beampattern error of 0.1 dB, corresponding to the quadratic phase error of $\pi/8$. It is also shown that a distance of $\frac{6R^2}{\lambda}$ or greater is required when sidelobes are as low as -40 dB.

Ryan [78] derived the distance formula as a function of array size R , operating wavelength λ and impinging angle θ . When the quadratic phase error is $\pi/2$, which corresponds to 1 dB beampattern error, the borderline distance is given by

$$r = \frac{(R \sin \theta)^2}{2\lambda} + \frac{R}{2} |\cos \theta| - \frac{\lambda}{8} \quad (2.43)$$

As an estimate, this formula gives the distance criterion of $\frac{R^2}{2\lambda}$ for an impinging angle of 90° with 1 dB beampattern error.

Based on the discussion above, we will use the distance $\frac{2R^2}{\lambda}$ as the borderline between near field and far field beamforming for all angles of impinging.

2.3.3 Near Field Fixed Beamforming Techniques

Although the fixed beamforming principles described in Section 2.2.2. are generally applicable to both near field and far field beamforming, near field fixed beampattern design has proved to be more complicated than its far field counterpart. Some special near field fixed beamforming methods have been reported in the near field beamforming literature. These methods include near field compensation [47], radial beampattern transformation or reciprocity [45, 46], and multi-dimensional Chebyshev optimization [61], which will be reviewed in this section.

Near Field Compensation

One common design method for fixed near field beamforming is near field compensation proposed by Khalil *et al.* [47]. For a specified beampattern, this method uses a delay compensation factor on each sensor to account for the near field spherical wave fronts and converts the near field beampattern into a far field beampattern. Then, far field beampattern design techniques can be used to derive appropriate sensor weights.

The near field compensation method depends on the array geometry and takes its simplest form when the sensors are linear equi-spaced. In this case, the compensation factors g_m for a fixed focal point (r_f, θ_f) are selected as

$$g_m = \frac{r_{mf}}{r_f} \exp\{j\kappa(r_f - r_{mf} + r_m \cos(\theta_m - \theta_f))\} \quad (2.44)$$

Including the compensation factors g_m in the near field beampattern (2.19) results in a resemblance to the far field beampattern

$$b_{far}(\mathbf{x}_f, \mathbf{x}_s, \omega) = \sum_{m=1}^M W_m^*(\omega) g_m \frac{r_s}{r_{ms}} \cdot \exp\{j\kappa(r_{ms} - r_s)\} \quad (2.45)$$

$$= \sum_{m=1}^M W_{m_{far}}^*(\omega) \cdot \exp\{j\kappa r_m \cos(\theta_m - \theta_f)\} \quad (2.46)$$

The far field weights $W_{m_{far}}(\omega)$ are related with the near field weights $W_m(\omega)$ by

$$W_{m_{far}}(\omega) = W_m(\omega) \frac{r_s}{r_{ms}} \frac{r_{mf}}{r_f} \cdot \exp\{-j\kappa(r_{ms} - r_s + r_f - r_{mf})\} \quad (2.47)$$

$W_{m_{far}}(\omega)$ are obtained by synthesizing the far field beampattern, using far field techniques.

Near field compensation only achieves the desired near field beampattern over a limited range of angles at the mainlobe. It lacks control over sidelobes because it only compensates the delay associated with the focal point.

Radial Beampattern Transformation or Reciprocity

The radial beampattern transformation/reciprocity method exploits the general solutions to the wave equation (2.2) in spherical coordinates. The spherical harmonic solution to the wave equation is given in beampattern form (synthesis equation) by [45, 46]

$$b_r(\theta, \phi) = r^{-1/2} \sum_{n=0}^{\infty} \sum_{m=-n}^n \alpha_n^m \cdot H_{n+1/2}^{(1)}(\kappa r) \cdot P_n^{|m|}(\cos \phi) \cdot e^{jm\theta} \quad (2.48)$$

where m and n are integers, $\kappa = 2\pi f/c$ is the wavenumber, $P_n^m(\cdot)$ is the associated Legendre function, and $H_{n+1/2}^{(1)}(\cdot)$ is the half odd integer order spherical Hankel function of the first kind, which is defined by

$$H_{n+1/2}^{(1)}(\cdot) = J_{n+1/2}(\cdot) + jY_{n+1/2}(\cdot) \quad (2.49)$$

where $J_{n+1/2}(\cdot)$ is a half integer order Bessel function of the first kind, and $Y_{n+1/2}(\cdot)$ is a half integer order Neumann function. The Fourier-like complex constants α_n^m can be expressed (analysis equation) explicitly as

$$\alpha_n^m = \frac{\zeta_n^m}{r^{-1/2} H_{n+1/2}^{(1)}(\kappa r)} \int_{n=0}^{2\pi} \int_0^{\pi} b_r(\theta, \phi) \cdot P_n^{|m|}(\cos \phi) \cdot \sin(\phi) \cdot e^{-jm\theta} d\phi d\theta \quad (2.50)$$

and

$$\zeta_n^m \equiv \sqrt{\frac{2n+1}{4\pi} \frac{(n-|m|)!}{(n+|m|)!}} \quad (2.51)$$

Using (2.50) followed by (2.48), one can transform the beampattern prescribed at r_1 (near field) to a beampattern at $r_2 = \infty$ (far field), then design the beamformer using far field techniques. This method is suitable for arbitrary near field beampatterns

with arbitrary array geometry, provided that the beampattern is achievable by the array geometry. The desired near field beampattern is achieved exactly over all angles, not just the primary look direction.

But this radial transformation involves multidimensional integration necessary from (2.50), and is very computationally difficult – even for the simplest case of linear array. Further development with this approach[46] has found the reciprocity relationship between the beampatterns transformed at two distances r_1 and r_2 . This leads to a novel design scheme reducing the computational burden.

The proposition of the reciprocity relationship is stated as follows:

Proposition: *If $b_{r_1}(\theta, \phi) = b$ and $b_{r_2}(\theta, \phi) = b^*$, then*

$$\tilde{b}_{r_1|r_2}^*(\theta, \phi) = \tilde{b}_{r_2|r_1}(\theta, \phi) \left(1 + O\left(\frac{1}{\kappa^2 r_2^2} - \frac{1}{\kappa^2 r_1^2}\right) \right) \quad (2.52)$$

as $\min(r_1, r_2) \rightarrow \infty$.

where $b_{r_1}(\theta, \phi)$ denotes the specified beampattern at r_1 , and $\tilde{b}_{r_2|r_1}(\theta, \phi)$ denotes the beampattern transformed from r_1 to r_2 . Similarly, $b_{r_2}(\theta, \phi)$ represents the specified beampattern at r_2 and $\tilde{b}_{r_1|r_2}(\theta, \phi)$ the re-synthesis from r_2 to r_1 .

Let $r_1 = r$ and $r_2 = \infty$, the far field beampattern corresponding to a desired near field beampattern satisfies the asymptotic equivalence

$$\tilde{b}_\infty(\theta, \phi) \simeq \tilde{b}_{r_1}^*(\theta, \phi) \quad \text{as } r_1 \rightarrow \infty \quad (2.53)$$

Then the approximation design procedure for near field beampattern is summarized as follows.

- Step 0. Specify the desired near field beampattern $b_{r_1}(\theta, \phi) = b$ at distance r ;
- Step 1. Synthesize the far field beampattern b^* at $r_2 = \infty$, i.e., $\tilde{b}_\infty(\theta, \phi) = b^*$;
- Step 2. Evaluate the near field beampattern $\tilde{b}_r(\theta, \phi) = a$ at r , using the sensor weights obtained in Step 1.;
- Step 3. Synthesize a far field beampattern a^* at $r_2 = \infty$. The resultant weights will produce the desired beampattern b at distance r .

The near field beampattern is determined from far field data sandwiched between two far field designs. Although reduced a lot from the radial transformation method, the computational complexity of the radial reciprocity method is still quite high. The design procedures are also very complicated.

Multi-dimensional Chebyshev Optimization

Nordebo *et al.* [61, 64] treated the near field beampattern design as a multi-dimensional digital FIR filter design problem. As we noted in Section 2.2.2, the min-max design of 1-D and 2-D linear phase FIR filters has been successfully applied to far field broadband beampattern design for linear equi-spaced arrays, where linear programming techniques and exchange algorithms are used for design optimization. In the near field case, the min-max design of a broadband beamformer has to be formulated as a quadratic programming of a weighted Chebyshev approximation.

The weighted Chebyshev optimization method tries to approximate the desired beampattern $b_d(\mathbf{x}, \omega)$ by the actual beampattern $b(\mathbf{x}, \omega)$, defined in spatial point $\mathbf{x}$ and frequency ω . The actual beampattern is given by $b(\mathbf{x}, \omega) = \mathbf{W}^H \mathbf{a}(\mathbf{x}, \omega)$, where $\mathbf{W}$ is the weight vector and $\mathbf{a}(\mathbf{x}, \omega)$ is the near field steering vector defined in (2.21).

Define a dense grid of P points in a space-frequency region. Evaluate the function $b_d(\mathbf{x}, \omega)$ and $\mathbf{a}(\mathbf{x}, \omega)$ at these points and denote them b_{di} and $\mathbf{a}_i$, $i = 1, 2, \dots, P$. The min-max near field design problem is to find the weight vector $\mathbf{W}$ that solves the Chebyshev optimization problem (COP):

$$\min_{\mathbf{W}} \max_i g_i |\mathbf{W}^H \mathbf{a}_i - b_{di}| \quad (2.54)$$

where g_i 's are positive weighting factors.

The quadratic programming method is then used to solve the COP numerically. The solution is, however, generally non-unique since the Haar condition may not hold. To avoid the extensive investigation of the uniqueness, some simple and applicable constraints are added to obtain a unique weighted Chebyshev solution. Minimum

Euclidean weight norm is a good choice for the constraint, since it implies minimum white noise amplification, less sensitivity to coefficient quantization errors, and less sensitivity to model imperfections in array processing, such as errors in array geometry and estimates of source location.

The advantage of this design approach is that the beampattern specified over a space-frequency region may be well controlled by weighting factors and the design of a general beampattern is usually achievable. The disadvantage, on the other hand, is the numerical complexity. “The execution time for fairly small size problems was ... not insignificant”, as described in [61].

2.3.4 Near Field Adaptive Beamforming Techniques

Research in near field adaptive beamforming is scarce to find in the array processing literature, since adaptive beamformers are sensitive to the hypotheses made on signal characteristics and errors in source localizations, and the complexity of near field processing also penalizes the implementation in real time, which is generally critical to adaptive schemes. The reported adaptive beamforming methods for near field application include array optimization using stochastic region contraction (SRC) proposed by Berger and Silverman [5], unconstrained near field gain optimization by Goulding [32, 65], and constrained near field gain optimization by Ryan and Goubran [79]. All of them are statistical optimization methods with no iterative adaptation algorithm involved.

Array Optimization using Stochastic Region Contraction (SRC)

The array optimization using stochastic region contraction (SRC) proposed by Berger and Silverman [5] tried to optimize a linear array by changing the sensor weights as well as sensor spacings. The problem was formulated as a min-max optimization of a cost function called the power spectral dispersion function (PSDF). The PSDF is derived using the spherical propagation model for the scenario in Figure 2.14., where

the desired speech signal is fixed at point $\mathbf{x}_s = (0, y)$, and white noises are presented on a line parallel to the array axis and passing the point $\mathbf{x}_s$. The noise sources are restricted in the region starting 0.3 meter away from the point $\mathbf{x}_s$ and ending 2.0 meters away from that point, on both sides. The min-max problem is formulated as

$$\min_{\mathbf{W}, \mathbf{x}} \max_{0.3 \leq |x_n| \leq 2.0} \Psi(\mathbf{W}, \mathbf{x}; \mathbf{x}_s, \mathbf{x}_n) \quad (2.55)$$

where $\mathbf{W}$ is sensor weights, $\mathbf{x}$ is sensor spacings, $\Psi(\mathbf{W}, \mathbf{x}; \mathbf{x}_s, \mathbf{x}_n)$ is the PSDF defined by [84]

$$\Psi(\mathbf{W}, \mathbf{x}; \mathbf{x}_s, \mathbf{x}_n) = \frac{1}{\omega_2 - \omega_1} \int_{\omega_1}^{\omega_2} |b(\mathbf{x}_s, \mathbf{x}_n, \omega)|^2 d\omega \quad (2.56)$$

and $b(\mathbf{x}_s, \mathbf{x}_n, \omega)$ is the near field beamformer response evaluated at noise sources. The PSDF is in fact the averaged noise power over the band of interest at the output of the array beamformer.

Figure 2.14: Array optimization by stochastic region contraction (SRC)

The optimization procedure has $2(M - 1)$ variables involved: $M - 1$ variables representing the sensor spacings, and $M - 1$ for sensor weights. In this case, the min-max cost function (the PSDF) exhibits multiple local minima (hundreds or thousands). So it is *multi-modal*. Finding the global optimum solution becomes a difficult numerical problem. The dynamic programming method used for the plane wave model [84] was

found to be very difficult or impossible for the spherical wave case. The SRC method is then developed to reduce the computational complexity. It is a kind of “random search” method which exploits the contour structure of a subclass of the cost function and avoids the search in the higher level regions at intermediate stages. So the available search effort is directed to smaller volumes which are more relevant to the global optimum. The SRC method is more efficient than the commonly used “simulated annealing” method, by a speedup factor of 30 to 50. It is also very well suited for parallel processing. However, its computational complexity makes the design very difficult even for large scale, high speed computers.

Constrained Near Field Optimization

The near field array gain optimization methods reported in [32, 65, 79] are, in fact, a maximization of SNR approach applied in near field beamforming. This is similar to the far field case described in Section 2.2.3. The unconstrained near field optimization [32, 65], however, is found to be impractical to implement, since the array gain at the end fires of the array is extremely large, resulting in unacceptable white noise amplification. Quadratic constraint [79] is then chosen for the optimization process by adding a small diagonal component to the noise covariance matrix. The optimum weight vector is then

$$\mathbf{W}_{opt}(\omega) = (\mathbf{R}_n + \gamma \mathbf{I})^{-1} \mathbf{S}(\omega) \quad (2.57)$$

where $\mathbf{I}$ is identical matrix. γ is the constraint parameter.

This method has been successfully applied to linear equi-spaced microphone arrays for near field sound pickup. By varying the constraint parameter γ with frequency, this method achieves 2 to 6 dB of improvement [80] in array gain for the low frequency end (300Hz to 2000Hz), using a 16-element uniform linear array. Unfortunately, there are no simple rules or theory on the selection of γ . An iterative procedure of trial and error has to be used.

Chapter 3

Overview of Broadband Adaptive Beamforming

The basic concepts and general methods of array beamforming have been addressed in Chapter 2, including far field and near field beamforming, narrowband and broadband beamforming, and fixed and adaptive beamforming. The emphasis has been placed on near field beamforming techniques. In this Chapter, we will direct our attention to *broadband adaptive beamforming*.

The technical challenges in broadband adaptive beamforming include frequency dependent beampattern variations associated with broadband beamforming, and the desired signal cancellation phenomena encountered with adaptive beamforming. These issues will be discussed in Section 3.1. Current approaches to broadbanding will be reviewed in Section 3.2, and remedies to desired signal cancellation phenomena are outlined in Section 3.3.

3.1 Technical Challenges in Broadband Adaptive Beamforming

Broadband adaptive beamforming imposes many technical challenges. We will discuss the frequency dependent beampattern variation with broadband beamforming and the desired signal cancellation phenomena due to reverberation and coherent interference in adaptive beamforming.

3.1.1 Frequency Dependent Beampattern Variation

With broadband signals, the problem of *broadbanding* a sensor array arises due to the frequency dependent array properties. Arrays with limited number of sensors are not able to densely sample the appropriate spatial aperture, resulting in large variations of frequency dependent beampatterns. More specifically, the variation in mainlobe width may cause spectral distortion in received signals. Frequency dependent null locations may impair the ability to cancel broadband interference.

To illustrate the frequency dependent beampattern variation, consider an 11-element uniform linear array designed for speech frequency band $B = [0.3, 3.4]$ kHz. To avoid spatial aliasing, the inter-sensor spacing is at most a half wavelength of the highest frequency, i.e. $d = \frac{c}{2f_b} = 5$ cm. An LCMV adaptive beamformer is designed with $K = 30$ taps attached to each element. To achieve the beampattern control at look direction $\theta = 90^\circ$ and over the entire frequency band, 30 constraints are designed using the eigenvector method [8]. The quiescent response of the beamformer is evaluated at five frequency points: 0.3 kHz, 0.8 kHz, 1.3 kHz, 2.3 kHz, and 3.3 kHz, as shown in Figure 3.1. Obviously, the beamwidth widens as the frequency decreases. The mainlobe beamwidth at 3.3 kHz and 300 Hz is approximately 15° and 170° , respectively. The frequency dependent variation is more than 150° .

The effective aperture measured by the number of $\lambda/2$ also varies widely, where λ is the wavelength of the operating frequency. The aperture at the high frequency

Figure 3.1: Frequency dependent beampattern variation for an 11-element ULA

edge is equal to the number of elements, while at the lowest frequency point, it is less than one. In other words, the 11 elements are equivalent to 2 elements with about $\lambda/3$ spacing for low frequencies. The reduced gain/aperture at low frequency results in very low efficiency in uniform linear arrays. Conventional delay-filter-and-sum beamformers also give a similar performance. Changing the length of the tapped delay line or the number of constraints will not improve the situation.

3.1.2 Desired Signal Cancellation Phenomena

The desired signal cancellation phenomena occur in adaptive array processing when the interference is coherent or highly correlated with the desired signal. The problem was discovered by Widrow, *et al.* [100]. Conventionally, all adaptive beamforming schemes have a key assumption that the interfering signals are non-coherent. However, if the desired and interfering signals are coherent or highly correlated, then the coherence can cause cancellation of the desired signal components and destroy the performance of conventional adaptive beamformers.

Signal cancellation can occur even when the adaptive beamformers are working perfectly. Taking a two-element MSC beamformer [100] as an example, the desired signal $s(t)$ received at the main channel is a bandpass signal with normalized passband $[0.2, 0.3]$, impinging at the broadside of the array. The interference (J_1) received at the auxiliary channel is a sinusoid with a normalized frequency 0.25, impinging at 45° . The behavior of the converged beamformer is plotted in Figure 3.2. The beam pattern in Figure 3.2(a) shows that the beamformer works effectively by placing a -40 dB null in the interference direction and forming the main beam at the look direction. The frequency response at 45° in Figure 3.2(b) shows the big notch at interference frequency 0.25, and the frequency response at 90° has all pass response. All these plots indicate that the beamformer works perfectly.

However, the signal at the beamformer output is problematic, as shown in Figure 3.3. The power spectrum of the output signal has a notch at the interference frequency. The signal components around frequency 0.25 are canceled by the adaptive beamformer.

The signal cancellation phenomena have also been found in other adaptive beamforming schemes. It can be understood that an adaptive beamformer is designed to minimize its output power, so without knowing what the desired signal is, it manipulates the correlated interference to cancel part of the desired signal to achieve its goal.

Coherent interference can arise when multipath propagation is present. In microphone array applications, reflected sound waves (reverberation) are in fact coherent interference of the direct sound wave. Reverberation not only causes degradation of speech quality, but also causes desired signal cancellation in adaptive beamformers.

(a) Beampattern at normalized frequency 0.25

(b) Frequency response for $\theta = 45^\circ$ and 90°

Figure 3.2: Performances of the conventional adaptive beamformer with correlated interference

Figure 3.3: Power spectra of the conventional adaptive beamformer with correlated interference

3.2 Current Approaches to Broadbanding

To reduce the frequency dependent variations of a broadband beamformer, a number of so-called *constant beamwidth beamforming* methods have been reported for broadband beamforming in array processing literature. These methods may be classified into 3 categories:

1. regular array weight selection approach;
2. unequally spaced array design approach;
3. nested array approach.

These three approaches will be discussed in the following subsections.

3.2.1 Regular Array Weight Selection Approach

The regular array weight selection approach uses an array with a fixed and regular geometry, such as Uniform Linear Arrays (ULA), circular arrays or planar arrays. The desired beampattern and the reduced frequency dependent variations are achieved only by means of weight selection. This approach generally requires a large number of sensors to achieve satisfactory performance over a wide frequency range. The number of sensors increases linearly with the bandwidth of interest.

In far field cases, many 2-D filter design methods may be used directly for linear array broadband beamformer design with proper frequency mapping. Frequency mapping means treating the tapped-delay line of a broadband beamformer as one frequency domain, and the spatial sampling of the linear array as another frequency domain. For example, the FAN filter method has been used for uniform linear arrays to achieve identical beampatterns over an octave passband (frequency band ratio of 2:1) [60]. The idea of the FAN filter is to use a 1-D FIR prototype to design a 2-D linear phase FIR filter by mapping 1-D frequency to some lines in 2-D frequency domain. With the number of sensors on the order of 30, the beamformer designed by the FAN filter method can obtain the desired beampatterns with very little frequency

dependent variations in mainlobe and sidelobes [82]. Other constant beamwidth beamformers use either the 2-D frequency sampling filter design method [11] or the Chebyshev shading method [31]. They both achieve a near-constant mainlobe over an octave passband with an 11 element linear array.

In near field cases, it is more difficult to apply the 2-D filter design method to a broadband beamformer. As we have discussed in Section 2.3.3, a broadband beamformer has to be formulated as a quadratic programming of a weighted Chebyshev approximation problem [64], which means enormous computational complexity. Other weight selection methods for near field broadband beamforming include the near field compensation method [47], and the constrained optimization method [80]. They have been discussed in Section 2.3.3 and Section 2.3.4, respectively.

3.2.2 Unequally Spaced Array Design Approach

A disadvantage with the regular array weight selection approach is that an equispaced array properly sampled at the highest frequency is grossly oversampled at the lowest frequency, because of the decade range of frequencies involved in most broadband applications. Although a constant beamwidth beamformer is achievable by weight selection, the number of elements implied by the oversampling is excessive and unnecessary. A more appropriate approach is then to consider a nonuniform array. This is called a “thinned” array in contrast to a “filled” array in antenna literature.

One such approach is unequally spaced array design utilizing the optimization of sensor locations (as well as tap weights) [5, 17, 25, 85, 88, 97]. This approach is proven to be very difficult due to the multi-dimensional nature of spatial sampling and the nonlinear relationship between the steering vector $b(\mathbf{x}_s, \omega)$ and sensor locations [91]. The problem was first targeted by numerical methods of multidimensional optimization. More specifically, dynamic programming [85, 84] has been used for far field arrays and Stochastic Region Contraction [5] for near field arrays, as we

mentioned in Section 2.3.4. These methods are very computationally intensive and have to rely on large-scale digital computers. They are also very limited in that little guidance can be provided for new designs other than those tried. Nevertheless, these trial-and-error type of techniques has produced quite satisfactory results.

Several theoretical researches in unequally spaced array design have been reported for far field beamformers. The method proposed by Unz [88] first expresses the beampattern in a series expansion, then truncates the expansion and inverts a matrix to obtain the sensor spacings. Another method is space taping [102], in which the density of sensors is made proportional to the amplitude of the aperture illumination of a continuous sensor array. Sensor spacings are chosen deterministically (rather than statistically) for arrays with a small number of elements [85]. The asymptotic theory was also developed [39] to express the relationships between beampattern properties and array design. The functional requirements on sensor spacings and weightings are derived from these relationships and then lead to the broadband array design. This method results in arrays having very little or no frequency dependence in their beampattern [17].

Recently, a more general theory and design method was proposed in [97]. This frequency invariant (FI) design approach uses a continuously distributed sensor to derive a frequency invariant beampattern property, and then approximate this continuous sensor with a finite set of unequally spaced discrete sensors. It was shown that the frequency response of the continuous sensor can be factored into two parts: (1) a primary filter response which is related to a slice of the desired aperture distribution; (2) a secondary filter which is independent of the sensor location and depends only on the dimension of the array. This provides the guidance on choosing nonuniform spacings which simultaneously avoid spatial aliasing and minimize the number of sensors. For a linear array designed with uniform aperture size M over frequency band $[f_a, f_b]$, the minimum number of sensors required is given by

$$N = M + 1 + \lceil \log \left(\frac{f_b}{f_a} \right) / \log \left(\frac{M}{M-1} \right) \rceil \quad (3.1)$$

Table 3.1: Sensor locations of a 17-element Frequency Invariant (FI) linear array

i	0	1	2	3	4	5	6	7	
$\frac{x_i}{\lambda_b}$	0	0.5	1	1.5	2	2.5	3.1	3.9	
i	8	9	10	11	12	13	14	15	16
$\frac{x_i}{\lambda_b}$	4.9	6.1	7.6	9.5	11.9	14.9	18.6	23.3	25

where $\lceil \cdot \rceil$ is the ceiling function, and the optimal sensor spacings are

$$x_i = \begin{cases} (\lambda_b/2)i, & \text{for } 0 \leq i \leq M \\ M \left(\frac{\lambda_b}{2} \right) \left(\frac{M}{M-1} \right)^{i-M}, & \text{for } M < i < N-1 \\ M(\lambda_a/2), & \text{for } i = N-1. \end{cases} \quad (3.2)$$

where λ_a and λ_b are the wavelength of the frequencies f_a and f_b . As an example, a speech band linear array was designed having 17 elements with the sensor locations given in Table 3.1.

The FI design method is suitable for one-, two- and three-dimensional sensor arrays, and it can cope with arbitrarily wide bandwidth and arbitrary desired beampatterns. Unfortunately, this method is only valid for far field beamforming. To extend it to near field array design, the radial beampattern transformation or reciprocity method (see Section 2.3.3) has to be used, resulting in very complicated implementation and very high computational complexity.

3.2.3 Nested Array Approach

Another approach to broadband beamforming is to use a set of nested arrays. This approach has become favorable, especially in microphone array signal processing [11, 57, 59, 72].

The nested array approach was first proposed by Morris and Hands [59] in the early 1960's. Three uniform subarrays are used, one for midband and one for each

band edge, as depicted in Figure 3.4. The ratio of inter-element spacings between the subarrays is 3. These three subarrays are then superimposed, after suitable filtering, to form a compound array which covers the whole frequency band.

Figure 3.4: A nested array with inter-sensor spacing ratio = 3

Figure 3.5: A harmonically nested array with inter-sensor spacing ratio = 2

Similarly, when subarrays have the inter-element spacing ratio of 2, the compound array is called a harmonically nested array. One such example is shown in Figure 3.5.

The harmonically nested array is designed for frequency band $[0.5, 4.0]$ kHz, consisting of 4 subarrays. This structure has been reported in [11], [47] and [57]. A large planar microphone array utilizing the harmonical nesting has also been implemented in the Murray Hill auditorium at AT & T Bell Labs [23]. It used 380 elements to cover the 3-octave frequency band.

Generally, to design a harmonically nested array, choose the first subarray to be an M -element Uniform Linear Array (ULA) for the highest frequency range $[f_b/2, f_b]$. To avoid grating lobes, the inter-sensor spacing d is at most half the wavelength of the high frequency edge, that is $d = c/(2f_b)$, where c is the speed of propagation. The second subarray is then designed for frequency range $[f_b/4, f_b/2]$ with inter-sensor spacing being $2d$. The first subarray is nested within the second subarray with $(M + 1)/2$ superimposed elements, assuming M is odd. The third and additional subarrays are designed similarly until the lowest frequency f_a is covered or the sensor spacing limit is reached. The number of total elements is a logarithmic function of the band ratio

$$N = M + (M - 1) \log_2 \frac{f_b/f_a - 1}{2} \quad (3.3)$$

In contrast, a single ULA requires $M(f_b/f_a)$ elements to achieve the same aperture for all frequencies.

Beampatterns of nested arrays are identical only at the high frequency edges of each subarray, but vary at intermediate frequencies. The effect of nesting is to reduce the extent of the beampattern variation to that which occurs within a subband. Frequency-dependent sensor weights are then used to interpolate to the frequencies in between. The reduced interpolation bandwidth implies reduced difficulties and improved performance.

3.3 Current Approaches to De-reverberation

Current approaches to de-reverberation fall into 3 categories: 1) blind equalization; 2) fixed beamforming with near field focusing; 3) adaptive beamforming with coherent interference suppression. The first approach is outside the scope of this research while the second approach has been discussed in Section 2.3.3. The third approach includes the method of decorrelation preprocessors and the method of robust beamforming. They will be reviewed in Section 3.3.1 and Section 3.3.2, respectively.

3.3.1 Decorrelation Preprocessor

Decorrelation preprocessors for coherent interference suppression generally rely on either spatial averaging [83] or spectral averaging (for broadband signals) [98, 103] to destroy the correlation.

Spatial Smoothing

First proposed for bearing estimation, then developed for spatial filtering, spatial smoothing (SS) is the most successful spatial averaging method for coherent interference suppression. The basic idea is to form p subgroups from an M element linear array, as depicted in Figure 3.6. So each subgroup has q elements and $q = M - p + 1$. At each time instant k , the data of these subgroups are fed into an adaptive beamformer in sequence. In other words, the $(N = qK)$ -dimensional weight vector of the adaptive beamformer is updated p times for each time instant k . Note K is the length of the transversal filters attached to the q channels of the beamformer.

It is proven [83] that the covariance matrix of the spatially smoothed data is the average of the covariance matrices of the subgroups. It decorrelates the covariance matrix of the input vector for coherent interference and signals, provided that the number of coherent signals D is less than p and q , or equivalently

$$M \geq 2D \tag{3.4}$$

Figure 3.6: Subgrouping in the spatial smoothing (SS) algorithm

Therefore, the decorrelation property of spatial smoothing is obtained at the expense of reduced aperture.

To simplify the analysis of an adaptive beamformer for coherent interference suppression, sinusoidal signals with fixed phase differences are used as desired and coherent interfering signals [83]. Beampatterns and frequency responses due to these signals will not form nulls properly if the signal cancellation occurs. As an example, Figure 3.7 shows the beampatterns of adaptive beamformers with and without an SS preprocessor. The desired signal is $s_1(t) = \sin(0.4\pi t)$. There are four interfering signals: $J1$ and $J3$ are two coherent ones having the same frequency as the desired signal; $J2$ and $J4$ are non-coherent interference. The amplitude of all interference is 10. The array without SS preprocessor has $M = 6$ elements. The array with SS preprocessor has a total of 10 elements divided into 5 subgroups. Each subgroup has 6 elements. The SS beamformer has nulls at all interference directions, but the conventional beamformer only forms nulls at directions of $J2$ and $J4$. Figure 3.8 (a) and (b) show the power spectral density (PSD) of the desired signal and the interfering signals. Figure 3.8 (c) shows the PSD of the conventional array output after conver-

gence. Note the big change in scale. Signal cancellation occurs with the conventional beamformer. Figure 3.8 (d) is the PSD of the SS beamformer output. It is clear that the desired signal is preserved by the SS preprocessor.

Figure 3.7: Array beampattern with and without the SS algorithm

Recent developments in the SS approach include the generalized eigenspace-based beamformers [106] and the eigenspace-based method using multiple shift-invariant subarrays [105], etc.

Spectral Averaging

The spectral averaging method proposed by Yang and Keveh [103] uses a coherent signal-subspace transformation (CSST) preprocessor $\mathbf{T}(\theta, f_j)$ for broadband coherent interference suppression. Let the broadband signal received by the array be transformed by discrete Fourier transform (DFT) to produce J narrowband frequency bins within the design bandwidth $B = [f_a, f_b]$. The CSST preprocessor is chosen to transform the frequency dependent array response into a frequency invariant response,

Figure 3.8: Signal power spectra with and without the SS algorithm.

that is

$$\mathbf{T}(\Theta, f_j) \mathbf{A}(\Theta, f_j) = \mathbf{A}(\Theta, f_0) \quad (3.5)$$

where $\mathbf{A}(\Theta, f_j)$ and $\mathbf{A}(\Theta, f_0)$ are the array steering matrix at frequency point f_j and the central frequency $f_0 = (f_a + f_b)/2$, respectively.

$$\mathbf{A}(\Theta, f_j) = [\mathbf{a}(\theta_1, f_j), \mathbf{a}(\theta_2, f_j), \dots, \mathbf{a}(\theta_D, f_j)] \quad (3.6)$$

where $\mathbf{a}(\theta_i, f_j)$ is the steering vector of the i th source at frequency f_j .

The CSST preprocessor $\mathbf{T}(\Theta, f_j)$ is obtained by

$$\mathbf{T}(\Theta, f_j) = \mathbf{A}(\Theta, f_0) \mathbf{A}^{-1}(\Theta, f_j) \quad (3.7)$$

The block diagram of the CSST beamformer is depicted in Figure 3.9. It has been proven that, after the CSST preprocessor, the data covariance matrix $\mathbf{R}_v$ is the spectral averaging of the covariance matrices $\mathbf{R}_u(f_i)$ of the array data

$$\mathbf{R}_v = E\{\mathbf{V}(k) \mathbf{V}^H(k)\} = \frac{1}{J} \sum_{j=0}^{J-1} \mathbf{R}_u(f_j) \quad (3.8)$$

Figure 3.9: Block diagram of the CSST adaptive beamformer

The spectral averaging reduces the correlation between the coherent signal and interference to a negligible level [103].

The advantage of CSST method is no loss of array aperture. But from the definition of $\mathbf{T}(\Theta, f_j)$, it is obvious that the CSST preprocessor requires the knowledge of all impinging angles Θ [96]. A modified scheme which does not require the estimates of DoA was proposed [98], based on the frequency invariant broadband beamforming method [97] described in Section 3.2.2. But this scheme is only valid for far field beamforming.

3.3.2 Robust Beamforming

The robust beamforming approach to coherent interference suppression tries to limit the level of desired signal cancellation to a tolerable amount through constraint design. It can also combat the effect of other model imperfections, such as steering error, location errors of the array and variations in propagation medium, etc.

A number of robust beamforming methods have been reported for coherent interference suppression. Qian and Van Veen [74, 75] have proposed a quadratically constrained partially adaptive beamformer for correlated interference rejection. In this method, some quadratic constraints are constructed based on the estimates of the

interference parameters, and then added to a linearly constrained minimum variance (LCMV) adaptive beamformer to prevent signal cancellation. The proper selection of the quadratic constraints ensures that signal cancellation is reduced to a specified small level, while satisfactory interference rejection is maintained. This approach does not require a uniform array structure, and is applicable to both narrowband and broadband signals. However, the design of quadratic constraints requires an estimate of the interference covariance matrix, which is a drawback of the method. Besides, this method has only been studied for far field beamforming. Its effectiveness to near field beamforming remains open.

Other far field robust beamforming methods include the time-domain adaptive beamformer with constrained power minimization [37], the constrained adaptive blocking matrix GSC method proposed by Hoshuyama, et al. [38], and the robust beamforming via target tracking method [26], etc.

Few near field de-reverberation techniques are reported in the literature. Ryan [81] has proposed a near field array optimization scheme to increase the array's capability of distance discrimination. This scheme uses quadratic constraints to reduce the array gain at the far field locations right behind the near field focal point. In microphone applications, this scheme can suppress the image sources behind the focal point by an additional 6 dB while maintaining the array gain loss within 2 dB.

Chapter 4

Near Field Spatial-Temporal Subband Beamforming Systems

In this chapter, we propose a novel spatial-temporal subband (STS) beamforming structure for near field broadband adaptive array processing. The proposed STS structure incorporates a spatial subband array with temporal subband multirate filters and obtains the advantages of both subband systems. It enables parallel processing of the subband systems, improves the computational efficiency and enhances the performances of the near field broadband beamformers.

In the structure of the spatial-temporal subband beamforming system, a harmonically nested array is used for spatial subbanding; while the temporal subband system employs either the Quadrature Mirror Filter (QMF) banks with maximum decimation or the non-critical sampling multirate subband filters. Three specific STS systems are developed using the nested array and the multirate subband filters:

1. the Nested Array Quadrature Mirror Filter (NAQMF) beamformer using near field adaptive GSC beamformers and critically sampled QMF banks;
2. the Nested Array Multirate Generalized Sidelobe Canceler (NAM-GSC) using near field adaptive GSC beamformers and non-critically sampled multirate sub-

band filters;

3. the Nested Array Switched Beam Adaptive Noise Canceler (NASB-ANC) using fixed beamformers with adaptive noise cancelers (ANC) and non-critically sampled multirate subband filters.

The three STS adaptive beamforming systems will be discussed in detail in this chapter. Section 4.1 describes the general structure of the Spatial-Temporal Subband adaptive beamforming system. Section 4.2 details the design, implementation and the noise rejection performance of the NAQMF beamformer. The problem of the high residual adaptation error caused by the maximum down-sampling of the NAQMF beamformer is also discussed. Section 4.3 describes the details of the NAM-GSC beamformer and its difference from the NAQMF beamformer. It also proposes a novel solution for improving the robustness of the NAM-GSC adaptive beamformer against location errors. Section 4.4 demonstrates the design and performances of the NASB-ANC scheme.

4.1 Near Field STS Adaptive Beamforming

4.1.1 General Structure of the STS Beamforming Systems

A novel Spatial-Temporal Subband (STS) adaptive beamforming system is proposed for near field adaptive arrays to overcome the frequency dependent beam pattern variation encountered by broadband beamformers. The general structure of the STS system is illustrated in Figure 4.1. It incorporates a spatial subband array with temporal subband multirate filters, and employs an adaptive beamformer or an adaptive noise canceler in each subband. It consists of a harmonically nested array, several analysis filters and down-samplers, near-field adaptive beamformers, up-samplers and synthesis filters.

Signals received by the nested array are sampled at a high frequency F_s . The

Figure 4.1: Structure of Spatial-Temporal Subband (STS) beamformers

sampled data are grouped into several subarrays. Each subarray is processed by its corresponding analysis filter $H_i(z)$ ($i = 1, 2, \dots, 4$), and then decimated by D_i . After the decimation, the adaptive beamformer of each subarray operates at a lower sampling rate F_i , where $F_i = F_s/D_i$. The outputs of the beamformers are interpolated by the up-samplers I_i and combined via the synthesis filters $G_i(z)$.

The harmonically nested array is a spatial subband system. It is used to cover a broad frequency range $B = [f_1, f_2]$, as shown in Figure 4.2. The nested array is composed of several equi-spaced linear subarrays, each having M elements. Subarray1 is designed for the highest frequency range $[f_2/2, f_2]$. The inter-element spacing d is at most half the wavelength of the high frequency edge, that is $d = c/(2f_2)$, where c is the speed of propagation. Subarray2 is designed for the frequency range $[f_2/4, f_2/2]$ with inter-element spacing being $2d$. Subarray1 is nested within Subarray2 with $(M+1)/2$ superimposed elements, assuming M is odd. More subarrays are designed similarly until the lowest frequency edge f_1 is covered.

Theoretically, the total number of elements of the composed array is a logarithmic function of the band ratio, that is $M + \frac{M-1}{2}(\log_2 \frac{f_2}{f_1} - 1)$. In practice, fewer elements and fewer subarrays may be used at the cost of performance degradation over the lower frequency range. The trade off can be made between the complexity of the beamformer and the performance of the array at low frequencies. For example, in the application of microphone arrays, the bandwidth of the wideband telephony is $B = [50, 7000]$ Hz, according to the G.722 standard [58]. The band ratio is as high as 140. It requires at least 8 harmonically nested subarrays to obtain optimum performance. Practically, however, a system of 4 to 6 subarrays will provide satisfactory performance with reasonable complexity.

The frequency bands covered by the 4-subarray system are depicted in Figure 4.3. Clearly the nested array is a spatial subband sampling system.

The analysis and synthesis filters are temporal subband systems. Each subarray requires an analysis filter and a synthesis filter to avoid aliasing and imaging. With

Figure 4.2: Configuration of an 11-element harmonically nested array

Figure 4.3: Frequency bands covered by the nested subarrays

smaller bandwidth covered by each subarray, temporal multirate sampling is incorporated with spatial subbanding via down-samplers and up-samplers. The analysis filters $H_i(z)$ and the down-samplers D_i can be implemented by a multistage tree structure, as depicted in Figure 4.4(a) or Figure 4.5(a). The structure in Figure 4.4(a) is the maximum decimation QMF bank, and the one in Figure 4.5(a) depicts the non-critical sampling multirate filter. Each stage of the tree consists of a high-pass filter $HP_i(z)$, a low-pass filter $LP_i(z)$ and down-samplers. The high-pass and low-pass filters are related with the parallel filters $H_i(z)$ in Figure 4.1 as

$$\begin{aligned}
 H_1(z) &= HP_1(z) \\
 H_2(z) &= LP_1(z) * HP_2(z^2) \\
 H_3(z) &= LP_1(z) * LP_2(z^2) * HP_3(z^4) \\
 H_4(z) &= LP_1(z) * LP_2(z^2) * LP_3(z^4).
 \end{aligned} \tag{4.1}$$

The synthesis filters $G_i(z)$ are the mirror images of the analysis filters and can also be implemented by a tree structure, as shown in Figure 4.4(b) or Figure 4.5(b).

The non-critical sampling filters in Figure 4.5 are slightly different from the maximum decimation QMF bank in Figure 4.4. The difference is that the high-pass branches of the analysis filter are not followed by down-samplers and those of the synthesis filter have no up-samplers, either. So the sampling frequencies of the subarrays are higher than the QMF scheme.

The output of each path of the tree-structured filter is fed into the corresponding subarray beamformer. In practice, not all paths in the tree are to be implemented for each sensor. For those sensors used by one or two subarrays, only the paths corresponding to the subarrays are needed. For example, only path HP_1 is necessary for sensor x_1 and x_{-1} which are only used in Subarray1.

In each subband, an adaptive beamformer is designed using near field beamforming techniques. A Generalized Sidelobe Canceler is used for the NAQMF and the NAM-GSC schemes. The design and implementation of the GSC are illustrated in Section

(a) analysis QMF filters

(b) synthesis QMF filters

Figure 4.4: Tree-structured QMF filters for critical sampling

(a) analysis FIR filters

(b) synthesis FIR filters

Figure 4.5: Tree-structured analysis and synthesis filters for non-critical sampling

4.1.3. For the NASB-ANC scheme, several Delay-Filter-and-Sum beamformers and an adaptive noise canceler are employed in each subband. The details of the DFS beamformers and the ANC will be discussed in Section 4.4.

4.1.2 Advantages of the STS Beamforming Systems

The proposed spatial-temporal subband beamformers may appear to be complicated at first glance, but they actually ease the difficult task of the near field broadband beamformer design. First, the use of a nested array splits a broadband beamformer into several subarray beamformers of smaller bands, so each subarray covers only an octave frequency band. They can be designed separately and processed in parallel. Without complicated design techniques, the nested array can provide spatial subbanding and reduce the frequency dependent beampattern variations to the extent which occurs within an octave frequency band. Different design methods and parameters may be employed in each subarray to best suit the characteristics of the subband. For example, different inter-element spacings and adaptation step sizes may be selected to optimize the performance of the whole array.

Secondly, nested arrays are easy to design, to scale and to implement. Changing the number of elements in a nested array or scaling the nested array for different frequency bands is straightforward. It does not require complicated redesign of the whole array.

Thirdly, the use of temporal multirate sampling techniques provides decimation in the time domain, so less taps are needed in each subarray beamformer than the full band schemes having high sampling rates and wide frequency bands. Temporal multirate sampling reduces the cost of the adaptive beamformers and leads to a higher computational efficiency. It also improves the tracking performance over the full band adaptive beamformers.

Furthermore, temporal multirate sampling relaxes the design requirements of the subband filters. Without multirate sampling, as proposed in [57], an analysis filter is

still needed for each element in each subarray, and more stringent filter specifications are required to avoid aliasing. With multirate sampling, the analysis and synthesis filters can be implemented by multistage tree-structured QMF banks or FIR filters, and the requirements for these filters can be relaxed [90].

Finally, the proposed spatial-temporal subband beamformers can significantly improve the performances of interference rejection, de-reverberation, convergence of adaptation, and robustness against location errors. These improvements will be detailed in Section 4.2 through Section 4.4, and in Section 5.2.

4.1.3 Design and Implementation of the Near Field GSC Adaptive Beamformer

In the STS beamforming systems, a near field broadband beamformer is employed in each subarray. To design the near field broadband adaptive beamformer, the far field LCMV method outlined in Section 2.2.3 is successfully adopted to near field adaptive beamforming using the eigenvector constraint method proposed by Buckley [8]. It is generally agreed that near field beamforming is much more complicated than far field beamforming. But using the eigenvector constraint design method, we developed a simple and elegant structure [112] for near field beamformers without increasing the computational complexity. This method also enables real arithmetic implementation which guarantees real coefficients and real outputs.

The goal of the constraint design is to find the constraint matrix $\mathbf{C}$ and the response vector $\mathbf{f}$, so the desired signal source is passed with specified gain and linear phase, and the interference and noises from other directions can be suppressed adaptively by minimizing the power of the array output. That is

$$\min_{\mathbf{W}} \mathbf{W}^T \mathbf{R}_u \mathbf{W} \quad \text{subject to} \quad \mathbf{C}^T \mathbf{W} = \mathbf{f} \quad (4.2)$$

where $\mathbf{R}_u = E\{\mathbf{U}\mathbf{U}^T\}$ is the covariance matrix of the input vector.

To design the constraint matrix $\mathbf{C}$ and the response vector $\mathbf{f}$, the eigenvector

constraint method first selects a large number of frequency points $\{f_j, j = 1, 2, \dots, J\}$ ($J \gg L$) within the passband, and forms the equation

$$\mathbf{A}^T \mathbf{W} = \mathbf{d} \quad (4.3)$$

$$\begin{aligned} \mathbf{A} &= [\mathbf{c}(f_1), \dots, \mathbf{c}(f_J) \quad \mathbf{s}(f_1), \dots, \mathbf{s}(f_J)] \\ \mathbf{d} &= [d_1 \cos(2\pi f_1 \tau_0), \dots, d_J \cos(2\pi f_J \tau_0) | \\ &\quad d_1 \sin(2\pi f_1 \tau_0), \dots, d_J \sin(2\pi f_J \tau_0)]^T \end{aligned} \quad (4.4)$$

where d_j and τ_0 are the desired gain and group delay respectively. And $\mathbf{c}(f_j)$ and $\mathbf{s}(f_j)$ are, respectively, the real and imaginary part of the steering vector, which is defined by (2.21) for near field beamforming, and by (2.23) for far field beamforming. The formulation of $\mathbf{A}$ and $\mathbf{d}$ guarantees that the designed LCMV beamformer has a real-valued weight vector and can be implemented with real arithmetic.

Secondly, the eigenvector constraint method decomposes $\mathbf{A}$ via singular value decomposition (SVD)

$$\mathbf{A} = \mathbf{P} \mathbf{\Sigma} \mathbf{Q}^T \quad (4.5)$$

where $\mathbf{\Sigma}$ is the $2J \times 2J$ diagonal matrix containing all singular values. $\mathbf{P}$ and $\mathbf{Q}$ are corresponding singular vectors. A rank L approximation of $\mathbf{A}$ is obtained as

$$\mathbf{A} \approx \mathbf{A}_L = \mathbf{P}_L \mathbf{\Sigma}_L \mathbf{Q}_L^T \quad (4.6)$$

where $\mathbf{\Sigma}_L$ is the diagonal matrix containing the L largest singular values of $\mathbf{A}$. The columns of $\mathbf{P}_L$ and $\mathbf{Q}_L$ are, respectively, the L columns of $\mathbf{P}$ and $\mathbf{Q}$ corresponding to these singular values.

To choose L , Buckley [8] has shown that it is sufficient to use the largest L singular values containing 99% of the total energy to enforce a unit gain at the look direction; while the largest singular values containing 99.99% of the total energy are required to force a 40 dB null at the interference direction. In far field beamforming, the

observation Time BandWidth Product (TBWP) provides a guideline on choosing L . The observation TBWP is denoted by ρ and defined by (2.13) in Section 2.2.1. Buckley [8] has also shown that over 99.99 % of the signal energy is concentrated in the first $\lceil 2\rho \pm 1 \rceil$ eigenvalues of the covariance matrix of the source, where $\lceil x \rceil$ represents the smallest integer greater than x . As a rule of thumb, it is sufficient to choose L such that

$$\lceil 2\rho \pm 1 \rceil \leq L \leq K. \quad (4.7)$$

In near field beamforming, this guideline is not as accurate as that in the far field case.

After choosing L , the rank L matrix $\mathbf{A}_L$ in (4.6) is used to replace $\mathbf{A}$ in (4.3). Then it yields

$$\mathbf{P}_L^T \mathbf{W} = \Sigma_L^{-1} \mathbf{Q}_L^T \mathbf{d} \quad (4.8)$$

Finally, the desired eigenvector constraints are obtained as

$$\begin{aligned} \mathbf{C} &= \mathbf{P}_L \\ \mathbf{f} &= \Sigma_L^{-1} \mathbf{Q}_L^T \mathbf{d}. \end{aligned} \quad (4.9)$$

The columns of $\mathbf{P}_L$ correspond to the eigenvectors of $\mathbf{A}\mathbf{A}^T$, hence the name eigenvector constraints.

An adaptive LCMV beamformer is usually implemented by a Generalized Sidelobe Canceler (GSC), as depicted in Figure 4.6. It consists of a fixed beamformer $\mathbf{W}_q$, a signal blocking matrix $\mathbf{C}_a$ and an unconstrained adaptive weight vector $\mathbf{W}_a$. The signal blocking matrix $\mathbf{C}_a$ can be obtained from $\mathbf{C}$ by solving $\mathbf{C}^H \mathbf{C}_a = \mathbf{0}$. The fixed beamformer $\mathbf{W}_q$ is given by $\mathbf{W}_q = \mathbf{C}(\mathbf{C}^T \mathbf{C})^{-1} \mathbf{f}$.

With L constraints, the dimensions of $\mathbf{C}_a$, $\mathbf{W}_q$ and $\mathbf{W}_a$ are $N \times (N - L)$, $N \times 1$ and $(N - L) \times 1$, respectively. Using internal steering, a GSC beamformer has the computational complexity of

$$O(N^2) = N(N - L) + N + (N - L) = N^2 + 2N - L(N + 1) \quad (4.10)$$

Figure 4.6: Adaptive beamformer implemented by a Generalized Sidelobe Canceler

real multiplications and real additions for each iteration.

When pre-steering and beam shaping are employed, however, the fixed beamformer $\mathbf{W}_q$ and the signal blocking matrix $\mathbf{C}_a$ are of the special form

$$\mathbf{W}_q = \begin{bmatrix} W_{q1} \cdot \mathbf{1}_M^T & \cdots & W_{qK} \cdot \mathbf{1}_M^T \end{bmatrix}^T \quad (4.11)$$

$$\mathbf{C}_a = \left[\begin{array}{ccc|c} \mathbf{ca} & & & \\ & \ddots & & \mathbf{cb} \\ & & \mathbf{ca} & \end{array} \right] \quad (4.12)$$

where W_{q_i} are scalars, and $\mathbf{1}_M$ is an $M \times 1$ unit vector. The left block of $\mathbf{C}_a$ consists of K folds of $\mathbf{ca}$ and the right block $\mathbf{cb}$ is an $N \times (M-1)L$ sparse matrix.

$$\mathbf{ca} = \begin{bmatrix} 1 & & & \\ -1 & 1 & & \\ & -1 & \ddots & \\ & & \ddots & 1 \\ & & & -1 \end{bmatrix}_{M \times (M-1)} \quad (4.13)$$

$$\mathbf{cb} = \begin{bmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,J-L} \\ \mathbf{0}_{M-1} & \mathbf{0}_{M-1} & \cdots & \mathbf{0}_{M-1} \\ \vdots & \vdots & \vdots & \vdots \\ b_{L,1} & b_{L,2} & \cdots & b_{L,J-L} \\ \mathbf{0}_{M-1} & \mathbf{0}_{M-1} & \cdots & \mathbf{0}_{M-1} \\ 1 & & & \\ \mathbf{0}_{M-1} & & & \\ & & \ddots & \\ & & \ddots & 1 \\ & & & \mathbf{0}_{M-1} \end{bmatrix} \quad (4.14)$$

where $\mathbf{0}_{M-1}$ is an $(M-1) \times 1$ zero vector.

Based on the sparse forms of $\mathbf{W}_q$ and $\mathbf{C}_a$, we have developed a simple and elegant implementation structure [112], as in Figure 4.7. The pre-steering and beam shaping are employed at each element of the array by the complex weighting factors a_m^* . The fixed beamformer $\mathbf{W}_q$ is implemented by a filter of length K instead of length N . The unconstrained adaptive weight $\mathbf{W}_a$ is split into two parts: one is the K -tap vectors $\mathbf{wa}_1, \mathbf{wa}_2, \dots, \mathbf{wa}_{(M-1)}$ corresponding to the K folds of $\mathbf{ca}$; another is the weights $W_{a1}, W_{a2}, \dots, W_{a(K-L)}$ corresponding to $\mathbf{cb}$. The L -tap filters $\mathbf{b}_j$ correspond to the $b_{i,j}$ values in $\mathbf{cb}$.

$$\mathbf{b}_j = \begin{bmatrix} b_{1,j} & b_{2,j} & \cdots & b_{L,j} \end{bmatrix}^T \quad (4.15)$$

The simplified structure reduces the computational complexity. The implementation of $\mathbf{W}_a$ still requires $(N-L)$ real multiplications and additions. But $\mathbf{C}_a$ requires only $L(K-L)$ real multiplications and $(L+1)(K-L) + 2(M-1)$ real additions. The fixed beamformer $\mathbf{W}_q$ also reduces to a K -tap FIR filter, requiring K real multiplications and additions. The pre-steering and beam shaping require a phase shift and M multiplications and additions. In total, this simplified structure requires only $N + (K-L)(L+1) + M$ real multiplications and $N + (L+2)(K-L) + 3M - 2$ real additions, compared with $O(N^2)$ in (4.10). For a beamformer having $M = 5$

Figure 4.7: Simplified implementation of GSC with pre-steering

elements, $K = 30$ taps and $L = 8$ constraints, the internally steered beamformer requires 2.16×10^4 multiplications and additions per iteration. The simplified structure requires only 353 real multiplications and 383 real additions, more than 95% savings.

The simplified implementation requires a good estimate of the desired signal location for accurate pre-steering and beam shaping. When there is no estimate error, the simplified implementation performs exactly the same as the internally steered beamformer. In near field applications, however, the estimate error is often large which neither implementation can tolerate. Modifications on beamformer design are generally required to cope with location errors. This robustness issue is discussed further in Section 4.3.3.

4.2 The NAQMF Adaptive Beamformer

4.2.1 Design of the NAQMF Beamformer

As a specific system employing the STS structure, we propose a Nested Array Quadrature Mirror Filter (NAQMF) beamformer [114] which uses a near field adaptive GSC beamformer in each subarray and critically sampled QMF banks as the analysis and synthesis filters. To use critical sampling in the QMF multirate filter banks and the nested array, the subband allocation is related to the selection of the sampling frequency F_s . The basic QMF banks can only be applied directly [67] when the passband edges are located at integer multiples of $F_s/(2D_i)$, where D_i is the downsampling rate of the subarray. With the subband allocation in Figure 4.3, the critically sampled QMF bank requires that the edges of the subbands must be located at $F_s/2$, $F_s/4$, and $F_s/8$, etc.

In the wideband telephony application, the G.722 standard [58] requires that the sampling frequency is 16kHz and the signal passband is $B = [50, 7000]$ Hz. If the Nyquist sampling rate of 14 kHz is used for the NAQMF beamformer, then the output has to be re-sampled to 16kHz; if we choose $F_s = 16$ kHz, then the high frequency edge

of the passband of the beamformer has to be adjusted to 8.0 kHz, with the subband edges located at 4.0 kHz, 2.0 kHz, 1.0 kHz, etc. Although complicated multirate filters are available for arbitrarily located subbands [54], they are not preferable for practical implementations.

Now let the sampling rate F_s of the NAQMF beamformer be 16 kHz. The nested array is then designed for the passband [50, 8000] Hz using four subbands with 5 elements in each subarray. With the speed of sound propagation being $c = 343$ m/s, the inter-element spacing of Subarray1 is set to $d = 2.0$ cm. The total size of the 11-element nested array is then 64.0 cm. The sampling frequencies of the 4 subarrays are $F_1 = 8$ kHz, $F_2 = 4$ kHz, $F_3 = 2$ kHz and $F_4 = 2$ kHz, respectively; while the passband of the subarrays are $B_1 = [4.0, 8.0]$ kHz, $B_2 = [2.0, 4.0]$ kHz, $B_3 = [1.0, 2.0]$ kHz and $B_4 = [0.05, 1.0]$ kHz.

A 3-stage tree-structured perfect reconstruction QMF bank [89, 90] is employed, as shown in Figure 4.4. Using a 48-tap D-type filter [15, table 7.2] in each stage, the resulting QMF bank obtains a stop band attenuation of 60 dB and a normalized transition band of 0.01, as illustrated in Figure 4.8.

An adaptive beamformer is designed for each subarray using $K = 21$ taps for each element. The constraints are designed for a focal point $\mathbf{x}_f = (r_f, \theta_f, \phi_f) = (0.6m, 90^\circ, 90^\circ)$, by the eigenvector constraint method detailed in Section 4.1.3. The number of constraints used in Subarray1 to Subarray4 are 21, 22, 23 and 23, respectively. The difference in the number of constraints is due to the different array size and the sampling rate of each subarray [13, 94].

4.2.2 Performances of the NAQMF Beamformer

The performances of the NAQMF adaptive beamformer are evaluated by its quiescent beampatterns, adaptive beampatterns and frequency responses, the output SINR and the convergence rate. For fair comparison, the performances of a full band beamformer are evaluated along with the NAQMF beamformer under the same conditions.

Figure 4.8: Frequency responses of a 3-stage tree-structured QMF bank.

The fullband beamformer has the same array geometry as the NAQMF beamformer. It uses the sampling frequency $F_s = 16$ kHz for the passband $B = [0.1, 8.0]$ kHz. It is designed to focus at the same focal point $\mathbf{x}_f$ using the same constraint method. With larger bandwidth, the fullband beamformer requires more taps at each element to achieve satisfactory beampatterns and cancellation of interference [13, 94]. The number of taps in the fullband beamformer is $K = 45$. The number of constraints used in the 11-element 45-tap fullband beamformer is 51.

The quiescent beampattern is defined as the array gain due to a white noise input. The quiescent beampatterns of the NAQMF beamformer are measured on a semi-circle of radius r_f on the $x - y$ plane, as illustrated in Figure 2.11. The four plots correspond to four in-band frequencies 0.5 kHz, 1.8 kHz, 3.5 kHz, and 6.8 kHz. The frequency dependent beampattern variations are illustrated in Figure 4.9 by comparing the quiescent beampatterns of the fullband and subband beamformers. The beampatterns of the NAQMF beamformer are shown in Figure 4.9(a). The 3 dB mainlobe beamwidth of all the plots varies between 30° to 60° . The frequency

dependent beamwidth variations are within approximately 30° . But the beamwidth of the fullband beamformer widens as the frequency decreases, as shown in Figure 4.9(b). The beamwidth variations are as large as 80° .

The reason for the reduced beampattern variation of the subband beamformer is that only 5 elements of the subarray are active for the corresponding subband beamformer, while all the 11 elements of the fullband beamformer are active for the whole frequency band.

The noise rejection performances of the fullband and the NAQMF beamformers are evaluated under three signal inputs. The desired signal $S1$ is located at the focal point $(0.6m, 90^\circ, 90^\circ)$ and the interfering signals $S2$ and $S3$ are at $(1.0m, 50^\circ, 90^\circ)$ and $(1.0m, 120^\circ, 90^\circ)$, respectively. They are uncorrelated color noises band limited to $[50, 7000]$ Hz, as specified by the G.722 standard. Each signal has a power of 20 dB with respect to the background noise. The Normalized Least-Mean-Square (NLMS) algorithm is used for both adaptive beamformers with the same step size of $\mu = 0.01$. The converged beamformers are also evaluated at the four in-band frequencies along the semi-circle of radius r_f . The NAQMF beamformer is able to place consistent nulls at the interference locations for all in-band frequencies, as illustrated by the beampatterns in Figure 4.10(a). But in Figure 4.10(b), the beampatterns of the fullband beamformer show that the nulls at the interference locations are not consistent for all frequencies and much higher sidelobes are presented for most frequencies.

The input signal at the array elements has a Signal-to-Interference-and-Noise-Ratio (SINR) of -3 dB. The subarray beamformers suppress the interference and obtain the output SINR of 25.7 dB, 24.6 dB, 23.9 dB and 9.5 dB, respectively. The synthesized output of the NAQMF beamformer achieves a SINR of 22.5 dB. The fullband beamformer, however, obtains a SINR of only 13.3 dB.

The noise reduction (NR) factor is defined as the ratio of the input noise power over output noise power. The NR factor of the NAQMF beamformer is 25.5 dB,

(a) 11-element NAQMF Beamformer

(b) 11-element Full Band Beamformer

Figure 4.9: Beampattern variations of the NAQMF beamformer compared to the fullband beamformer with the same array geometry.

(a) 11-element NAQMF Beamformer

(b) 11-element Full Band Beamformer

Figure 4.10: Converged nulling beampatterns of the NAQMF beamformer. The desired signal is $S1$ and the interfering signals are $S2$ and $S3$.

while that of the fullband beamformer is 16.3 dB.

The tracking performances of the fullband and subband adaptive beamformers are evaluated by the excess Mean Squared Error (MSE). Let $\mathbf{W}_a(k)$ denote the iterative solution of the unconstrained adaptive weights using the normalized LMS algorithm, and $\mathbf{W}_{aopt}$ denote the optimum Wiener solution. The excess MSE of an adaptive beamformer is defined as [35]

$$J_{ex}(k) = E\{|e(k)|^2\} - E\{|e_{opt}(k)|^2\} \quad (4.16)$$

where $e(k)$ and $e_{opt}(k)$ are the errors determined by

$$e(k) = [\mathbf{W}_q - \mathbf{C}_a \mathbf{W}_a(k)]^H \mathbf{u}(k) \quad (4.17)$$

$$e_{opt}(k) = [\mathbf{W}_q - \mathbf{C}_a \mathbf{W}_{aopt}]^H \mathbf{u}(k) \quad (4.18)$$

Refer to Figure 4.6. The operator $E\{.\}$ denotes the expectation. The total error of the NAQMF beamformer is the summation of $|e(k)|^2$ of all subarrays.

Figure 4.11 shows the excess MSE of the fullband adaptive beamformer and the NAQMF beamformer. Both adaptive beamformers use the NLMS algorithm with the three input signals as in Figure 4.10. The fullband adaptive beamformer converges slightly faster than the NAQMF beamformer and has smaller residual error after the convergence. The NAQMF beamformer with fixed step size $\mu = 0.01$ converges fast but has approximately 6 dB higher residual error than the fullband beamformer. Selecting a different step size between 0.1 to 0.01 for each subband beamformer, the NAQMF beamformer can reduce the residual error and obtain an excess MSE curve close to the fullband beamformer. The residual error of the NAQMF beamformer is 2 dB above the fullband beamformer.

4.2.3 Improvements on the NAQMF Beamformer

The advantages of the NAQMF adaptive beamformer include the reduced beamwidth variation and the improved computational efficiency over the full band beamformer.

Figure 4.11: Excess MSE of the NAQMF adaptive beamformer.

The use of the harmonically nested array and subband beamforming reduces the frequency dependent beampattern variation. The use of multirate QMF banks reduces the computational complexity of the adaptive beamformer because less taps are needed in each subarray beamformer. The spatial-temporal subband beamforming also enables parallel processing of the system.

However, there are also disadvantages to the NAQMF beamformer. One is the restriction of the subband frequency edges relative to the sampling frequencies. For the 11-element NAQMF beamformer with 16 kHz sampling frequency, the high frequency edge has to be 8.0 kHz which is much higher than the required G.722 passband edge of 7.0 kHz. The unnecessary stretch over the high frequency band results in the reduced aperture in the low frequency band. Another disadvantage is the unsatisfactory convergence performance and the high residual error. There are several reasons for the convergence behavior of the 11-element NAQMF beamformer:

- slow convergence of the low frequency band beamformer, especially Subarray4;
- limited capability of interference rejection at the low frequency subarray;

- aliasing errors between adjacent subbands due to critical sampling in each subband.

The unsatisfactory MSE performance of the 11-element NAQMF beamformer is mainly because of the degradation of the lowest frequency band subarray. Subarray4 has to cover more than an octave frequency band. It has limited aperture to reject interference at the low frequencies. It converges much slower and with much higher residual errors than those of Subarray1 to Subarray3. To improve the low band performance, another nested subarray may be added to the existing 4 subarrays, as shown in Figure 4.12. In the five subband NAQMF system, Subarray4 only covers the band [500, 1000] Hz, and Subarray5 covers the band below 500 Hz. The inter-element spacings of the two lower subband arrays are chosen to be $\lambda/4$ instead of $\lambda/2$, where λ is the wavelength of the high frequency edge of the corresponding band. Reducing the inter-element spacing and increasing the number of elements improve the performance of the low band subarrays, because smaller than half-wavelength spacing is required to avoid near field spatial aliasing [2].

The total number of elements in the 5 subband NAQMF beamformer is 17 and the total array size is 128 cm. The beampattern variation of the 5-subarray NAQMF beamformer is reduced to 18° , as shown in Figure 4.13. This is better than the 4-subband NAQMF beamformer. The convergence curve of the 5-subband NAQMF beamformer is plotted in Figure 4.14, along with that of the 17-element fullband beamformer. Now the NAQMF beamformer has 2 dB higher residual MSE than that of the fullband beamformer, using the same step size of $\mu = 0.01$. With the step size matched to the subbands, the 5-subband NAQMF beamformer is able to achieve lower MSE than that of the fullband beamformer.

The critical sampling used in the NAQMF beamformer still has aliasing errors between subbands. Although each subband has better convergence speed due to the smaller eigen spread of the subband input signals, the synthesized beamformer does not converge significantly faster than the fullband beamformer because of the

Figure 4.12: Array geometry of the NAQMF beamformer with 5 subbands.

Figure 4.13: Beampatterns of the 5-subband NAQMF adaptive beamformer.

Figure 4.14: Excess MSE of the NAQMF beamformer with 5 subbands.

aliasing errors. Solutions to the critically sampled adaptive filter problem have been reported in the literature, including an oversampling scheme [44] and two critical sampling schemes: the scheme using adaptive cross-terms between subbands [28] and the adaptive filter with sparse sub-filters [71, 93], etc. The two critical sampling schemes use very complicated structures to remove the aliasing errors between adjacent subbands. They result in some improvement in convergence for cases where the number of subbands is greater than 8. The conclusion drawn in [28] says: “Since both convergence gains and computational efficiency can be best achieved with oversampled schemes, oversampling is still the way to go.”

In our spatial-temporal subband beamforming schemes, oversampling in time leads to the NAM-GSC beamformer and the NASB-ANC scheme, which are detailed in Section 4.3 and Section 4.4, respectively.

4.3 The NAM-GSC Adaptive Beamformer

4.3.1 Nested Array Multirate Beamformers with Non-critical Sampling

The STS adaptive beamforming system may incorporate a harmonically nested array with non-critically sampled multirate subband filters and adaptive GSC beamformers. This type of STS beamformer is simply called the Nested Array Multirate GSC (NAM-GSC) beamformer. Without critical sampling, the nested subarrays of an NAM-GSC adaptive beamformer can be designed to best suit the desired frequency band. Using the sampling frequency 16 kHz in the wideband telephony applications, we design a 4-subarray NAM-GSC beamformer to cover the passband [50, 7200] Hz, which provides a tradeoff between the low band performance and system complexity [113].

The frequency bands covered by the 4 subarrays are $B_1 = [3600, 7200]$ Hz, $B_2 =$

[1800, 3600] Hz, $B_3 = [900, 1800]$ Hz, and $B_4 = [50, 900]$ Hz, respectively. Still, Subarray4 covers more than an octave frequency band. The inter-element spacing of Subarray1 is set to $d = 2.4$ cm, which is the half wavelength of 7200 Hz. The inter-element spacings of Subarray 2 to Subarray4 are 4.8 cm, 9.6 cm and 19.2 cm, respectively. The total size of the 11-element array is 76.8 cm. Compared to the NAQMF beamformer, the high frequency range of 7200 Hz to 8000 Hz is not covered by the NAM-GSC beamformer. The size of the array is slightly larger to provide better aperture for the low frequency end.

The low frequency performance of the NAM-GSC beamformer is also limited by the size of the array. If a larger array size is allowed, one or more subarrays may be added in the manner similar to the 5 subband NAQMF beamformer shown in Figure 4.12. The added subarray can cover the frequency band below 450 Hz, and the low frequency performance is improved at the cost of increased system complexity.

The analysis and synthesis filters of the NAM-GSC beamformer are the 3-stage tree-structured multirate filters shown in Figure 4.5. Each stage of the tree has a 49-tap high-pass filter and a 49-tap low-pass FIR filter designed by the Remez method. The equivalent parallel filters have a stop band attenuation of 60 dB and the normalized transition band of 0.0625. The frequency responses of the analysis filters are shown in Figure 4.15. The cutoff frequencies of the filters are at 900 Hz, 1800 Hz, 3600 Hz and 7200 Hz, matching the designed subbands for the subarrays.

The analysis and synthesis filters of the NAM-GSC beamformer are different from the multirate QMF bank shown in Figure 4.4. The difference is that the high-pass branches of the analysis filter are not followed by down-samplers and those branches of the synthesis filter have no up-samplers, either. The sampling frequencies of the subarrays are $F_1 = F_s = 16$ kHz, $F_2 = 8$ kHz, $F_3 = 4$ kHz and $F_4 = 2$ kHz.

For each subarray, a near field adaptive beamformer is designed using the same eigenvector constraint method as in Section 4.2. The focal point of the NAM-GSC beamformer is $\mathbf{x}_f = (0.6\text{m}, 90^\circ, 90^\circ)$, which is the same as that of the NAQMF beam-

Figure 4.15: Frequency responses of the 3-stage tree structure FIR filters

former. For Subarray1 and Subarray2, the distance of 0.6 meters is at the boundary of the near-field and far-field, while for Subarray3 and Subarray4, it is well within the near-field of the array. The number of taps used in each subarray is 16. So the total number of weights in the NAM-GSC beamformer is 320. The number of constraints is 10 for Subarray1, and 11 for the other three subarrays.

Similarly, an 11-element fullband beamformer is also designed to use the same array geometry and the sampling rate $F_s = 16$ kHz. It covers the whole frequency band $B = [50, 7200]$ Hz. The number of taps attached at each element is 32. The total number of weights in the full band beamformer is 352. The number of constraints is $L = 36$.

4.3.2 Performances of the NAM-GSC Adaptive Beamformer

The performances of the NAM-GSC adaptive beamformer are also evaluated by its quiescent beampatterns, adaptive beampatterns and frequency responses, the output SINR and the convergence rate.

The reduction of frequency dependent variation obtained by the NAM-GSC is illustrated by its quiescent beampatterns in Figure 4.16. The frequencies of the plots are 0.5 kHz, 1.8 kHz, 3.5 kHz, and 6.8 kHz, the same as those in Section 4.2. The beampatterns of the NAM-GSC beamformer are shown in Figure 4.16(a). The mainlobe beamwidth at the 4 frequency points varies within 15° . The beamwidth variation at the lowest frequency is smaller than that of the NAQMF beamformer in Figure 4.9(a). This better performance is obtained due to the larger array size of the NAM-GSC beamformer. In Figure 4.16(b), the beamwidth of the fullband GSC beamformer also widens as the frequency decreases, similar to the beampatterns of the fullband beamformer in Figure 4.9(b), with the frequency dependent beamwidth variation being approximately 80° .

The adaptive beampatterns of the NAM-GSC beamformer are evaluated with three signal sources. The desired signal $S1$ is located at the focal point and two interfering signals $S2$ and $S3$ are at $(1.0m, 50^\circ, 90^\circ)$ and $(1.0m, 120^\circ, 90^\circ)$, respectively. They are uncorrelated, colored noises generated by passing independent white noises through an 81-tap bandpass FIR filter. The signals are band limited to $[50, 7000]$ Hz with SNR=20 dB. The Normalized Least-Mean-Square (NLMS) algorithm is used for adaptation. The converged beamformer responses are plotted in Figure 4.16. Figure 4.17(a) shows that the NAM-GSC beamformer has consistent deep nulls formed at all frequencies at the interference directions while maintaining unit gain at the desired signal direction. The fullband array also maintains the unit gain at the desired signal direction, as shown in Figure 4.17(b), but the nulls at low frequencies are not as deep as the NAM-GSC beamformer.

Using the converged NAM-GSC weight vector $\mathbf{W}_a$, the outputs of the subarrays $v_i(k)$ are obtained by filtering the array input signals with $\mathbf{W}_a$. The SINR of the 4 subarrays are 32.8 dB, 31.7 dB, 32.0 dB, and 21.9 dB, respectively. The output of the compound NAM-GSC beamformer is obtained by combining the outputs of the 4 subarrays via the synthesis filters. Although subband sampling introduces some

(a) 11-element 16-tap NAM-GSC beamformer

(b) 11-element 32-tap full band beamformer

Figure 4.16: Beampattern variations of the NAM-GSC beamformer compared to the fullband beamformer with the same array geometry.

(a) 11-element 16-tap NAM-GSC beamformer

(b) 11-element 32-tap full band beamformer

Figure 4.17: Noise rejection performances of the NAM-GSC beamformer without location errors, where $S1$ is the desired signal, $S2$ and $S3$ are the interference.

aliasing error, the NAM-GSC beamformer achieves a SINR of 30.3 dB. The fullband GSC beamformer achieves a SINR of 27.0 dB. The output power and SINR of the NAM-GSC beamformers are listed in Table 4.1, where P_s , P_d , and P_i , denote the power of the total output, the power of the desired signal output, and the power of the interference plus noise output, respectively.

The noise reduction factor of the NAM-GSC beamformer is 33.8 dB, while the NR of the fullband beamformer is 30.1 dB. The NAM-GSC beamformer achieves better noise rejection performance than the fullband beamformer with less adaptive weights.

Table 4.1: Output power and SINR of the NAM-GSC beamformer and the fullband GSC for noise rejection

	P_s	P_d	P_i	SINR
Array Input Signal	167.57	55.548	112.09	-3.0 dB
Fullband GSC Output	55.177	55.067	0.1096	27.0 dB
NAM-GSC Output	50.175	50.129	0.0466	30.3 dB

Next, the Mean-Squared-Error (MSE) is evaluated to compare the tracking performance of the subband and the fullband adaptive GSC beamformers. Figure 4.18 shows the excess MSE curves for the same step size $\mu = 0.001$. The three input signals are the same as the example in Figure 4.17. The excess MSE is obtained by (4.16) with ensemble average of 10 trails and time average of every 500 iterations. With non-critical sampling, the NAM-GSC beamformer converges faster than the fullband beamformer and provides a constant improvement of the excess MSE of 4.5 dB.

Figure 4.19 shows the excess MSE curves for the NAM-GSC beamformer and the fullband GSC beamformer with step size $\mu = 0.01$. They converge much faster but have higher residual errors than the ones in Figure 4.18. This is expected because the step size is much larger in Figure 4.19. Comparing this figure to Figure 4.11,

Figure 4.18: Excess MSE of the NAM-GSC adaptive beamformer using the NLMS algorithm with $\mu = 0.001$.

the NAM-GSC beamformer converges much faster than the NAQMF beamformer with the same step size. This is due to the fact that the NAM-GSC beamformer has smaller aliasing errors between subbands with non-critical sampling and faster convergence of the low frequency band.

4.3.3 Robustness of the NAM-GSC Against Location Errors

In real world applications, the robustness of an adaptive beamformer against location errors is an important issue in near field beamforming. It is much more difficult to estimate a 3-dimensional location for near field arrays than just the Angle of Arrival (AoA) in the far field scenario. The estimation error of the radial distance is often large, and exact estimation of angles is also difficult. The estimation error of the desired signal location may only result in slight reduction of the array gain for fixed beamformers or optimum beamformers without real time adaptation. But it will cause severe degradation in performance for iteratively adaptive beamformers. The

Figure 4.19: Excess MSE of the NAM-GSC adaptive beamformer using the NLMS algorithm with $\mu = 0.01$.

desired signal may be treated as interference and be cancelled completely.

The following example illustrates the effect of a location error on the adaptive GSC beamformers. The NAM-GSC beamformer and the signals are the same as those in Figure 4.17, except that the desired signal is now located at $\mathbf{x}_1 = (0.75m, 89^\circ, 90^\circ)$. The focal point of the adaptive beamformers is $\mathbf{x}_f = (0.6m, 90^\circ, 90^\circ)$. There is a location error of 0.15 meter in the estimated distance and 1° in the azimuth angle.

Figure 4.20 shows the NAM-GSC beamformer response after convergence. The adaptive beamformer treats the desired signal as interference and tries to cancel it by forming a null at its direction. The capability of cancelling the other two interfering signals is reduced. The output SINR also decreases dramatically to 10 dB.

To reduce the sensitivity of the GSC beamformer to location errors, we propose a new design for near field robust NAM-GSC beamformer [113]. The idea is to constrain a spatial region around the focal point. The array on the x -axis does not have resolution in the elevation angle ϕ , thus the constrained points are selected in

Figure 4.20: Sensitivity of the NAM-GSC beamformer to signal location errors.

a fan shaped region on the $x - y$ plane. The size of the region is specified by Δr and $\Delta\theta$, as illustrated in Figure 4.21. When the focal point is $\mathbf{x}_f = (r_f, \theta_f, \phi_f)$, a set of I points may be selected by varying r and θ uniformly within the range $[r_f - \Delta r, r_f + \Delta r]$ and $[\theta_f - \Delta\theta, \theta_f + \Delta\theta]$, respectively.

The set of I spatial points are denoted by $\mathbf{x}_i$. To place unit gain constraints on the I spatial locations as well as the J in-band frequencies, the constraint equation (4.4) is modified as

$$\mathbf{A} = \begin{bmatrix} \mathbf{c}(\mathbf{x}_1, \omega_1), \dots, \mathbf{c}(\mathbf{x}_i, \omega_j), \dots, \mathbf{c}(\mathbf{x}_I, \omega_J) \\ \mathbf{s}(\mathbf{x}_1, \omega_1), \dots, \mathbf{s}(\mathbf{x}_i, \omega_j), \dots, \mathbf{s}(\mathbf{x}_I, \omega_J) \end{bmatrix} \quad (4.19)$$

$$\mathbf{d} = \begin{bmatrix} d_{11} \cos(\omega_1 \tau_1), \dots, d_{ij} \cos(\omega_j \tau_i), \dots, d_{IJ} \cos(\omega_J \tau_I) \\ d_{11} \sin(\omega_1 \tau_1), \dots, d_{ij} \sin(\omega_j \tau_i), \dots, d_{IJ} \sin(\omega_J \tau_I) \end{bmatrix}^T \quad (4.20)$$

where $\mathbf{c}(\mathbf{x}_i, \omega_j)$ and $\mathbf{s}(\mathbf{x}_i, \omega_j)$ are, respectively, the real and imaginary part of the steering vector $\mathbf{a}(\mathbf{x}_i, \omega_j)$ defined in (2.21). And τ_i are the group delays corresponding to the spatial location $\mathbf{x}_i$. Unit gain is enforced by setting $d_{ij} = 1$, so that signals

Figure 4.21: Spatial region to be constrained by the robust GSC beamformer

falling within the constrained spatial region and frequency band are passed without attenuation. This leads to the robustness against location errors.

The formulation of $\mathbf{A}$ and $\mathbf{d}$ still guarantees real arithmetic. The remaining procedures of the eigenvector constraint design are unchanged, as in (4.5), (4.6), (4.8) and (4.9). With the increased spatial points being constrained, the number of the constraints increases, too.

For the numerical example of the NAM-GSC beamformer in Section 4.3, the focal point is $\mathbf{x}_f = (0.6m, 90^\circ, 90^\circ)$. We choose $\Delta r = 0.15$ meter and $\Delta\theta = 2^\circ$ for the constrained region. Five r values are selected uniformly within $[r_f - \Delta r, r_f + \Delta r]$. Three θ values are selected within $[\theta_f - \Delta\theta, \theta_f + \Delta\theta]$. The total number of the constrained spatial points is then $I = 15$. For each spatial point, a set of $J = 40$ frequency points is also chosen uniformly within the passband. This constraint design is performed for each subband adaptive GSC beamformer. The resulting robust NAM-GSC beamformer has less degree of freedom for adaptation. Table 4.2 shows the numbers of eigenvector constraints and the degree of freedom for the four subarrays of the robust NAM-GSC beamformer.

Table 4.2: Number of constraints (L) and degree of freedom ($N - L$) in the robust GSC beamformer

	Subarray1	Subarray2	Subarray3	Subarray4
L	23	27	29	29
$N - L$	57	53	51	51

The adaptive GSC beamformers designed above are more robust against location errors. When the desired signal is off the focal point and locates at $(0.75m, 89^\circ, 90^\circ)$, the robust NAM-GSC beamformer can pass the desired signal without cancellation and suppress the two interfering signals effectively. The output SINRs of Subarray1 to Subarray4 are 22.7 dB, 22.5 dB, 22.4 dB, and 16.8 dB, respectively. The compound NAM-GSC beamformer achieves the SINR of 21.5 dB. The beamformer responses are also satisfactory, as shown in Figure 4.22. Unit gain is maintained at the desired signal location and nulls are formed at the interference locations. The convergence behavior of the robust adaptive beamformer is similar to that in Figure 4.18.

The improved robustness is obtained at the cost of the reduced degree of freedom and the reduced SINR at the output. This can be justified by comparing the beam-patterns in Figure 4.22 and Figure 4.16(a). The end fires of the beampatterns in Figure 4.22 are much higher than the beampatterns in Figure 4.16(a), and the nulls at interference locations are not as deep. This is because there are not enough degrees of freedom available in the unconstrained adaptive weights. The SINR can be improved by increasing the number of taps in the tapped delay lines, and/or increasing the number of elements in the array.

Figure 4.22: Responses of the robust NAM-GSC adaptive beamformer when the desired signal has small location errors.

4.4 The Nested Array Switched Beam Adaptive Noise Canceler

4.4.1 General Structure of the NASB-ANC Scheme

In the spatial-temporal subband array system depicted in Figure 4.1, the adaptive beamformer in each subarray may be replaced by a Switched Beam Adaptive Noise Canceler (SB-ANC). The resulting system is the Nested Array Switched Beam Adaptive Noise Canceler (NASB-ANC). The nested array and the analysis and synthesis filters of the NASB-ANC remain the same as those of the NAM-GSC scheme.

The block diagram of a Switched Beam Adaptive Noise Canceler (SB-ANC) is illustrated in Figure 4.23. It consists of three functional blocks: the array beamformers, the switches, and the adaptive noise canceler (ANC). The signals received at the M -element array are fed into several pre-designed beamformers. Each beamformer

focuses at a separate spatial location without adapting to the signal environment. The switches select the desired beam as the primary channel and other beams as the auxiliary channels for the ANC. The ANC is a standard adaptive filter which adaptively cancels the noise components in the primary channel and tries to achieve a higher SINR at the output.

Near field delay-filter-and-sum (DFS) beamformers are employed for our NASB-ANC scheme. The control signals $C_1, C_2, \dots, C_P$ are used to steer the DFS beamformers. A Voice Activity Detector (VAD) is used to turn off the adaptation of the ANC when the desired signal is present. This is critical to the success of the NASB-ANC, because the coupling of the desired signal in the auxiliary channels would cause severe cancellation of the desired signal at the ANC output. A perfect VAD is assumed for our study.

4.4.2 Performances of the NASB-ANC Scheme

The noise rejection performance of the NASB-ANC scheme is evaluated by the same three signals used in Section 4.3. They are the three uncorrelated signals $S1$, $S2$ and $S3$ located at $\mathbf{x}_{s1} = (0.6m, 90^\circ, 90^\circ)$, $\mathbf{x}_{s2} = (1.0m, 50^\circ, 90^\circ)$, and $\mathbf{x}_{s3} = (1.0m, 120^\circ, 90^\circ)$, respectively. The three DFS beams (denoted Beam1, Beam2 and Beam3) are designed for each subarray focusing at the three signal locations respectively. Each DFS beamformer has $M = 5$ elements and $K = 16$ taps per element. The DFS beamformers can provide approximately 15 dB sidelobe attenuation, as illustrated by their beampatterns plotted in Figure 4.24. The beampatterns are also evaluated at 0.5 kHz, 1.8 kHz, 3.5 kHz, and 6.8 kHz — the same four in-band frequencies as those in Section 4.3.

Suppose $S1$ is the desired signal so Beam1 is selected as the primary channel of the ANC. Beam2 and Beam3 are the auxiliary channels. The ANCs have $Q = 32$ taps per auxiliary channel. The adaptive weights of the ANC is a (64×1) dimensional vector. The group delay in the primary channel is $D = Q/2$.

Figure 4.23: Structure of the Switched Beam Adaptive Noise Canceler (SB-ANC)

(a) Beam1 focusing at $(0.6m, 90^\circ, 90^\circ)$ (b) Beam2 focusing at $(1.0m, 50^\circ, 90^\circ)$ (c) Beam3 focusing at $(1.0m, 120^\circ, 90^\circ)$

Figure 4.24: Fixed DFS beams of the NASB-ANC with the 11-element nested array

In comparison, a fullband SB-ANC is also designed using the same array geometry and three DFS fullband beamformers. Each beamformer has 11-elements and 32 taps per element. The beampatterns of the three fullband beams are plotted in Figure 4.25. The fullband ANC has $Q = 100$ taps per auxiliary channel. The group delay of the primary channel is also $D = Q/2$.

With a perfect VAD, the fullband and subband ANC are converged to their optimum weights. The output of each optimum ANC is denoted $y_s(t)$. It is decomposed into the desired signal portion $y_d(t)$ and the interference and noise portion $y_i(t)$. The power of the outputs $y_s(t)$, $y_d(t)$ and $y_i(t)$ are denoted P_s , P_d and P_i , respectively. The SINR of the subarray outputs are 29.6 dB, 28.1 dB, 27.5 dB, and 27.4 dB, respectively. The compound NASB-ANC achieves a SINR of 29.0 dB. The fullband SB-ANC obtains 26.4 dB SINR at the output.

The power and SINR of the fullband SB-ANC and the subband NASB-ANC are listed in Table 4.3. The noise reduction factor of the fullband SB-ANC is 29.5 dB; while the NR factor of the NASB-ANC is 32.3 dB.

Table 4.3: Output power and SINR of the NASB-ANC and the fullband SB-ANC for noise rejection

	P_s	P_d	P_i	SINR
Array Input Signal	167.57	55.548	112.09	-3.0 dB
Fullband SB-ANC Output	55.326	55.281	0.1266	26.4 dB
Subband NASB-ANC Output	52.789	52.684	0.0663	29.0 dB

The convergence of the NASB-ANC is also compared with the fullband SB-ANC, using the NLMS algorithm with step size $\mu = 0.01$. The excess MSE curves are plotted in Figure 4.26. The excess MSE of the subband NASB-ANC has higher residual error than that of the fullband SB-ANC. Selecting the step sizes of the subband ANC between 0.1 to 0.01, the NASB-ANC can achieve faster convergence

(a) Beam1 focusing at $(0.6m, 90^\circ, 90^\circ)$ (b) Beam2 focusing at $(1.0m, 50^\circ, 90^\circ)$ (c) Beam3 focusing at $(1.0m, 120^\circ, 90^\circ)$

Figure 4.25: Fixed DFS beams of the 11-element nested array fullband SB-ANC.

than the fullband SB-ANC with compatible residual error of -29 dB.

Figure 4.26: Excess MSE of the NASB-ANC scheme using the NLMS algorithm with $\mu = 0.01$.

The MSE curves of the SB-ANC schemes in Figure 4.26 are different from those of the GSC beamformers in Figure 4.19. For the MSE curves of step size $\mu = 0.01$, the fullband SB-ANC scheme achieves a residual error of -29 dB, which is lower than the -24 dB obtained by the subband NASB-ANC scheme. The fullband GSC beamformer only obtains a residual error of -10 dB, which is little higher than the -12 dB obtained by the subband NAM-GSC beamformer. When the MSE is lower than -20 dB, the aliasing errors of the subbands become dominant and the MSE of the subband NASB-ANC is limited by the aliasing errors. Thus the fullband SB-ANC outperforms the subband NASB-ANC. In the fullband and subband GSC schemes, however, the MSEs are high and are mainly contributed by the background noises. So the subband NAM-GSC beamformer performs better than the fullband GSC beamformer. On the other hand, both fullband and subband NASB-ANC schemes achieve much lower residual errors than the fullband and subband GSC beamformers. This better performance is

obtained at the cost of the “beamformer plus ANC” structures and the assistance of the perfect VAD.

The robustness of the NASB-ANC is also examined and the results are fully satisfactory. Without changing the pre-designed beams, the NASB-ANC is able to preserve the desired signal and suppress the interference when the signals are located away from the focal points of the beams. When the location error of the desired signal $S1$ is as large as $0.5m < r_{s1} < 1.5m$ and $88^\circ \leq \theta_{s1} \leq 92^\circ$, the change of the output SINR of the NASB-ANC is within 1 dB. Moving $S2$ and $S3$ around also does not degrade the performance of the NASB-ANC, providing that $\theta_{s2} < 55^\circ$ and $\theta_{s3} > 115^\circ$. For example, when the desired signal is at $(0.75m, 89^\circ, 90^\circ)$, the NASB-ANC still achieves a SINR of 28.9 dB and a NR of 31.8 dB. This NR factor is much better than the NR of 24.5 dB obtained in the same scenario by the robust NAM-GSC beamformer in Section 4.3.3.

Chapter 5

De-reverberation Performances of the STS Beamformers

Room reverberation contributes to a large amount of interference in microphone array applications. De-reverberation is a great challenge to adaptive beamforming because

1. reverberant interference is highly correlated with the direct path signal. The correlation of the reverberant signals may cause desired signal cancellation in adaptive beamformers;
2. reverberant interference always follows the desired signal and is difficult to separate. The technique of adapting during the absence of the desired signal is not applicable to adaptive de-reverberation.

Consequently, the STS adaptive beamformers proposed in Chapter 4 may degrade their performances in reverberant environments. Therefore, the NAM-GSC beamformer and the NASB-ANC scheme will be evaluated for their de-reverberation performance in this chapter. Section 5.1 describes the simulated room reverberation by the image model [3]. The simulated reverberant signals are used to evaluate the de-reverberation performance. Section 5.2 develops the objective measures for de-reverberation performance, including the output SINR, the PSD versus frequency,

the SNR versus frequency, the Noise Reduction (NR) factor and the energy decay curve (EDC). The de-reverberation performances of the NAM-GSC beamformer and the NASB-ANC scheme are evaluated and compared with their fullband counterparts. Section 5.3 provides an analysis to the de-reverberation performances of the NAM-GSC beamformer and the NASB-ANC scheme.

5.1 Reverberation Modeling

The reverberation of a room is generally described by its impulse response. The impulse response of a real room reverberation is often difficult to simulate accurately, because reverberant sound fields are very complex. The reflections of the room boundaries vary with signal frequencies and surface materials [79, Chapter 2]. But for the simplicity of computer simulation, the image model proposed in [3] is the most appropriate method. The details of the image model and the simulation of room impulse response can be found in Appendix A.

The reverberant signals are generated by convolving a clean signal $s(t)$ with the impulse responses between the source location and the array elements. The simulated room has a size of $(L_x, L_y, L_z) = (5.0m, 4.0m, 3.0m)$. The reflection coefficients of the walls are 0.9, and those of the ceiling and floor are 0.7. The reverberation time of the simulated room is approximately $T_{60} = 250$ ms. The 11-element nested array is located on the axis x in the room, as shown in Figure 5.1. The angle between the x axis and the wall is $\beta = 45^\circ$. The phase center of the array is at point o , and it is located at $(1.0\text{ m}, 1.0\text{ m}, 1.0\text{ m})$ on the $x' - y'$ plane. The geometry of the array is the same as the nested array used in the NAM-GSC beamformer and the NASB-ANC in Chapter 4. The signal source is located in front of the array at $\mathbf{x}_s = (0m, 0.7m, 0m)$ on the $x - y$ plane.

The direct path signal is received at the array elements as the desired signal $u_d(t)$. The sum of the reflected image signals is received at the array elements as the in-

Figure 5.1: A nested array in a reverberant room. The figure is not to scale.

terfering signal $u_i(t)$. The sum of the desired signal and the interfering signal is the reverberant signal $u(t) = u_d(t) + u_i(t)$.

When the sampling frequency is $F_s = 16$ kHz, the room impulse responses among array elements involve fractional delays. In our simulation, the fractional delays are implemented by the method of FIR filter approximation [51] with the sampling rate remaining F_s . No up-sampling and down-sampling are needed.

5.2 De-reverberation Performances

The NAM-GSC adaptive beamformer and the NASB-ANC proposed in Chapter 4 are evaluated for their de-reverberation performance using the simulated reverberant signals. For comparison, the fullband adaptive GSC beamformer and the fullband SB-ANC are also evaluated along with the two subband schemes. All schemes use

the same 11-element harmonically nested array with the same location arrangement in the room, as shown in Figure 5.1.

Assume there is no uncorrelated interference present but the reverberant signals and background noises. The direct path signal located at $\mathbf{x}_s$ is received at the array phase center with SNR of 23dB. The adaptive GSC beamformers are adapted with the input signal $u(t)$. The converged optimum weights are denoted as $\mathbf{W}_{gsc}$. The fullband SB-ANC and the subband NASB-ANC are adapted with a perfect VAD. The converged optimum weights are denoted as $\mathbf{W}_{anc}$.

A simulation is also performed for the fullband and subband adaptive GSC beamformers with the input of only the interfering signal $u_i(t)$ plus background noises. The optimum weights (denote $\mathbf{W}_{bst}$) are obtained at the absence of the direct path signal so no desired signal cancellation can occur. These weights ($\mathbf{W}_{bst}$) are not attainable in real reverberant rooms because the direct path signal is separated from its reflections. But these weights provide some guidelines for the best achievable performance of de-reverberation with the given array parameters.

For the purpose of de-reverberation evaluation, the reverberant signal $u(t)$ is filtered separately by the three sets of optimum weights $\mathbf{W}_{gsc}$, $\mathbf{W}_{anc}$, and $\mathbf{W}_{bst}$. The outputs of the beamformers are denoted by $y_u(t)$. It is decomposed into the desired signal portion $y_d(t)$ and the interference plus noise portion $y_i(t)$. Several measurements are made on these outputs to demonstrate the effect of reverberation to the STS beamformers. These measurements include the beampatterns, the signal power and the SINR, the PSD of the output signals, the noise reduction factor (NR) and the energy decay curves (EDC).

5.2.1 Beampatterns

The beampatterns obtained by the adaptive weights $\mathbf{W}_{gsc}$, $\mathbf{W}_{anc}$, and $\mathbf{W}_{bst}$ are shown in Figure 5.2, Figure 5.3, and Figure 5.4, respectively. All beampatterns are evaluated on the semi-circle in front of the array with radius 0.7 meter. The four in-band

frequencies are 0.5 kHz, 1.8 kHz, 3.5 kHz and 6.8 kHz, respectively.

The shapes of the beampatterns vary greatly although all of them maintain a unit gain at the focal point which is at $(0.7m, 90^\circ, 90^\circ)$ with respect to the array axis. Figure 5.2 shows the beampatterns of the fullband GSC beamformer and the subband NAM-GSC beamformer adapted at the presence of the desired signal. The beampatterns of the fullband GSC beamformer exhibit nulls at the focal point, as shown in Figure 5.2(a). Especially the nulls of the high frequency plots are deeper than 10 dB. This indicates the desired signal cancellation and the reduced de-reverberation performance. For the subband NAM-GSC beamformer, high frequency plots are fine but valleys are formed around the focal point in low frequency plots, as shown in Figure 5.2(b). This indicates that the desired signal cancellation of the subband NAM-GSC occurs in the low frequency band.

Figure 5.3 shows the beampatterns of the fullband SB-ANC and the NASB-ANC with a perfect VAD. They have the same shapes as those of the DFS beampatterns in Figure 4.25(a) and Figure 4.24(a). This suggests that the ANC's are not active in reverberant environments. The near field DFS beamformers of the fullband SB-ANC and the subband NASB-ANC can provide more than 5 dB attenuation to the sidelobes. These near field beamformers can provide some de-reverberation gain, as illustrated in Table 5.1.

Figure 5.4 shows the beampatterns of the best achievable beamformers for de-reverberation. The fullband $\mathbf{W}_{bst}$ beamformer can form peaks at the focal point and attenuate reverberant signals from sidelobes, as shown in Figure 5.4(a). But the end fires of the low frequency beampatterns are slightly higher than those of the fullband $\mathbf{W}_{anc}$ in Figure 5.3(a). In Figure 5.4(b), the low frequency plots of the subband $\mathbf{W}_{bst}$ beamformer exhibit slightly higher sidelobes than the mainlobes. Its high frequency plots are fine, providing unit gain at the desired signal location and attenuation at the sidelobes. But its capacity of canceling individual interfering signals is limited due to the small number of elements.

(a) The 11-element fullband GSC beamformer

(b) The 11-element Subband Scheme

Figure 5.2: De-reverberation beampatterns of the NAM-GSC beamformer $\mathbf{W}_{gsc}$ adapted at the presence of the desired signal.

(a) The 11-element Fullband Scheme

(b) The 11-element Subband Scheme

Figure 5.3: De-reverberation beampatterns of the NASB-ANC $\mathbf{W}_{anc}$ with its ANCs switched off by a VAD.

(a) The 11-element fullband GSC beamformer

(b) The 11-element subband GSC beamformer

Figure 5.4: De-reverberation beampatterns of the best achievable beamformer $\mathbf{W}_{bst}$ adapted at the absence of the desired signal.

5.2.2 The Signal Power and SINR

The input and output power and the SINR of the adaptive beamformers are listed in Table 5.1. The power of the total output $y_u(t)$, the desired signal output $y_d(t)$ and the interference output $y_i(t)$ are denoted P_u , P_d and P_i , respectively. The background noise power is denoted P_n . The SINR is defined as $10 \log_{10} \frac{P_d}{P_i + P_n}$.

Table 5.1: Power and SINR of the NAM-GSC beamformers and the NASB-ANC for de-reverberation

	P_u	P_d	P_i	P_n	SINR
Array Input Signal	93.858	75.288	18.358	0.1000	6.1 dB
Fullband $\mathbf{W}_{gsc}$	76.082	71.972	4.8120	0.6400	11.2 dB
Fullband $\mathbf{W}_{anc}$	85.699	79.955	5.5080	0.0100	11.6 dB
Fullband $\mathbf{W}_{bst}$	80.725	79.322	1.3660	0.0550	17.5 dB
Subband $\mathbf{W}_{gsc}$	72.246	69.415	7.2700	0.4400	9.5 dB
Subband $\mathbf{W}_{anc}$	77.762	69.761	7.4240	0.0050	9.7 dB
Subband $\mathbf{W}_{bst}$	75.818	69.632	5.6410	0.1600	10.8 dB

From the SINR values in Table 5.1, all beamformers provide 4 dB to 8 dB de-reverberation gain over the input signals. The fullband schemes perform slightly better than the subband schemes.

The fullband and subband GSC beamformers $\mathbf{W}_{gsc}$ provide 5.1 dB and 3.4 dB SINR improvement, respectively. Their reverberant interference powers are on the same order as those obtained by the switched beam ANC schemes ($\mathbf{W}_{anc}$), but the background noises are enhanced to a level much higher than the input noises. The fullband GSC beamformer also has lower desired signal power than those of the other two fullband schemes.

The switched beam ANC schemes ($\mathbf{W}_{anc}$) suppress the background noises (P_n) to a very low level. But they have quite a high portion of the reverberant interference (P_i)

left in the output. The fullband SB-ANC scheme provides 5.5 dB de-reverberation gain and the subband NASB-ANC scheme provides 3.8 dB gain. Since the ANC's are turned off by the perfect VAD, the de-reverberation gains are provided solely by the near field DFS beamformers.

The $\mathbf{W}_{bst}$ beamformers provide the highest SINR at the output. The fullband scheme achieves 11.4 dB de-reverberation gain and the subband scheme obtains 4.7 dB. They suppress the reverberant interference to the lowest level, although slightly higher background noises than the ANC schemes are left in the output. The fullband best achievable $\mathbf{W}_{bst}$ beamformer has much higher de-reverberation gain than the fullband adaptive GSC beamformer and the fullband SB-ANC scheme. But the differences between the subband schemes are much smaller. This means that reverberation has greater impact on the fullband beamformers than on the subband beamformers.

In each row of Table 5.1, the total power P_u is approximately the sum of the desired signal power P_d , the interference power P_i and the background noise power P_n . This fact does not clearly suggest any desired signal cancellation. However, the desired signal cancellation phenomena do occur in certain frequency ranges for the GSC beamformers adapted at the presence of the direct path signal. In the fullband GSC beamformer, this occurs in the high frequency band which will be illustrated by the SNR and NR versus frequency plots. For the subband NAM-GSC beamformer, the cancellation occurs in the low frequency band. This can be verified by the input/output powers and SINR of the subarrays and the PSD plots of the lowest subband.

Table 5.2 lists the input/output powers and SINR of each subarray. The frequency bands covered by the subarrays are $B_1 = [3.6, 7.2]$ kHz, $B_2 = [1.8, 3.6]$ kHz, $B_3 = [0.9, 1.8]$ kHz, and $B_4 = [0.3, 0.9]$ kHz, respectively. For the three high frequency bands, the SINR of the three adaptive schemes are pretty close, ranging from 9.6 dB to 11.8 dB. They are able to suppress the reverberant interference and the background

noise effectively and maintain a high desired signal power at the output. However, the low frequency subband beamformer $\mathbf{W}_{gsc}$ has an output SINR lower than the input SINR. Its interference plus noise output (P_i and P_n) are higher than those of the input signal, and its desired signal output is lower than that of the input signal. The total output power P_u is also much less than the sum of P_d , P_i and P_n . This clearly indicates that the desired signal is cancelled partially by the reverberant interfering signal.

There is no quantitative measure of the desired signal cancellation in the array processing literature yet. Here we define a desired signal cancellation rate as

$$D_c = 10 \log_{10} \left(\frac{P_d + P_i + P_n}{P_u} \right). \quad (5.1)$$

The higher the rate, the more severe the desired signal is cancelled and the worse the performance. The desired signal cancellation is negligible when $D_c < 1$ dB.

For the low frequency subband beamformer $\mathbf{W}_{gsc}$, the desired signal cancellation rate is about 3.0 dB, calculated from the data in Table 5.2. For all other schemes of the three high frequency subbands, the desired signal cancellation rates are less than 0.1 dB. Using the data in Table 5.1 for the fullband and subband beamformers, the desired signal cancellation rates are low for all schemes. The desired signal cancellation rate of the subband NAM-GSC beamformer is $D_c = 0.2$ dB, and those of the other schemes are less than 0.1 dB.

The signal cancellation phenomenon in the lowest subband GSC beamformer is also shown on the PSD plots in Figure 5.5. Note that the frequency is normalized to 1000 Hz. In Figure 5.5(a), the output PSD plots are obtained with the weights $\mathbf{W}_{gsc}$. The PSD of the reverberant interference output $y_i(t)$ is pretty high. The PSD of the total output $y_u(t)$ is lower than the PSD of the desired $y_d(t)$ over a large frequency range. This also indicates that the desired signal cancellation occurs in the low frequency subband beamformer when adapted at the presence of the direct path signal.

For comparison, the PSD plot of the low subband NASB-ANC scheme is shown

Table 5.2: Power and SINR of the Four Subarrays for De-reverberation

Subarray1 covering the subband $B_1 = [3.6, 7.2]$ kHz.

	P_u	P_d	P_i	P_n	SINR
Subarray1 Input	47.190	37.893	9.2250	0.0050	6.1 dB
$\mathbf{W}_{gsc}$ Output	41.439	37.902	3.5710	0.0480	10.0 dB
$\mathbf{W}_{anc}$ Output	42.087	37.933	4.0750	0.0010	9.6 dB
$\mathbf{W}_{bst}$ Output	41.627	37.928	3.5470	0.0560	10.5 dB

Subarray2 covering the subband $B_2 = [1.8, 3.6]$ kHz.

	P_u	P_d	P_i	P_n	SINR
Subarray2 Input	11.617	9.3620	2.2650	0.0025	6.1 dB
$\mathbf{W}_{gsc}$ Output	10.181	9.4080	0.8700	0.0160	10.3 dB
$\mathbf{W}_{anc}$ Output	10.456	9.4400	1.0180	0.0005	9.6 dB
$\mathbf{W}_{bst}$ Output	10.283	9.4340	0.8450	0.0180	10.9 dB

Subarray3 covering the subband $B_3 = [0.9, 1.8]$ kHz.

	P_u	P_d	P_i	P_n	SINR
Subarray3 Input	2.9768	2.3449	0.5436	0.0013	6.3 dB
$\mathbf{W}_{gsc}$ Output	2.5843	2.4111	0.1652	0.0100	11.4 dB
$\mathbf{W}_{anc}$ Output	2.7678	2.4302	0.2420	0.0002	10.0 dB
$\mathbf{W}_{bst}$ Output	2.6386	2.4199	0.1492	0.0090	11.8 dB

Subarray4 covering the subband $B_4 = [0.3, 0.9]$ kHz.

	P_u	P_d	P_i	P_n	SINR
Subarray4 Input	0.9732	0.7924	0.1664	0.0006	6.8 dB
$\mathbf{W}_{gsc}$ Output	0.5585	0.8511	0.2524	0.0080	5.1 dB
$\mathbf{W}_{anc}$ Output	0.9813	0.8719	0.0953	0.0002	9.6 dB
$\mathbf{W}_{bst}$ Output	0.8952	0.8602	0.0255	0.0040	14.6 dB

(a) NAM-GSC W_{gsc} adapted at the presence of the desired signal

(b) NASB-ANC W_{anc} with the ANC switched off

(c) NAM-GSC W_{bst} adapted at the absence of the desired signal

Figure 5.5: PSD of the low subband beamformer outputs in a reverberant room.

in Figure 5.5(b). The reverberant interference and noises are suppressed by more than 10 dB as indicated by the PSD of $y_i(t)$. The total output $y_u(t)$ and the desired output $y_d(t)$ have similar PSD over the passband. No desired signal cancellation is observed.

The output PSD plots in Figure 5.5(c) are obtained with the optimum $\mathbf{W}_{bst}$ beamformer adapted at the absence of the direct path signal. The PSD of the total output $y_u(t)$ now is very close to the PSD of the desired output $y_d(t)$. The PSD of the interference output $y_i(t)$ is low except the small peak at $f = 350$ Hz, where a drop of the desired signal power also occurs. But the PSD of the total output is not low at that point. This also indicates a small amount of the desired signal cancellation and leakage of the reverberant interference around that frequency point. In the rest of the passband, the optimum $\mathbf{W}_{bst}$ beamformer performs well for de-reverberation.

5.2.3 SNR and NR versus the Frequency

The SNR and NR of the adaptive beamformers are evaluated as functions of frequency. The $SNR(f)$ is defined as

$$SNR(f) = 10 \log_{10} \frac{\Phi_d(f)}{\Phi_i(f) + \Phi_n(f)} \quad (5.2)$$

where $\Phi_d(f)$ and $\Phi_i(f)$ are the PSD of the beamformer desired output $y_d(t)$ and the interference output $y_i(t)$, and $\Phi_n(f)$ is the PSD of the background noise output.

The noise reduction factor $NR(f)$ is defined as

$$NR(f) = 10 \log_{10} \frac{\Psi_{in}(f)}{\Psi_{out}(f)} \quad (5.3)$$

where $\Psi_{in}(f)$ is the input PSD of the interference plus noise, and $\Psi_{out}(f)$ is the output PSD of the interference plus noise.

Figure 5.6 plots the $SNR(f)$ curves of the fullband and subband beamformers. The $SNR(f)$ curve of the array input signal is included in both plots as a reference.

The fullband beamformers $\mathbf{W}_{anc}$, $\mathbf{W}_{gsc}$, and $\mathbf{W}_{bst}$ have different characteristics of $SNR(f)$, as shown in Figure 5.6(a). The $SNR(f)$ of the $\mathbf{W}_{anc}$ is pretty flat over the

(a) Fullband Schemes

(b) Subband Schemes

Figure 5.6: $SINR(f)$ of the adaptive beamformers in a reverberant room.

passband. It has almost a constant improvement of 4 dB over the input $SNR(f)$, except the lowest band below 900 Hz. The $SNR(f)$ of the fullband $\mathbf{W}_{bst}$ beamformer also has less improvement of SNR in this low band; while in the high frequency range [900, 7200] Hz, its improved SNR is as high as 12 dB to 18 dB. The SNR of the fullband $\mathbf{W}_{gsc}$ beamformer decreases with the increase of the frequency. It drops below the input curve at the high frequency edge. This indicates that the desired signal cancellation and/or noise enhancement over the high frequency band is quite serious for the full band GSC beamformer in the reverberant environment.

The $SNR(f)$ curves of the subband schemes are plotted in Figure 5.6(b). Unlike the fullband $\mathbf{W}_{gsc}$ beamformer which has problems over the high frequency range, the subband NAM-GSC beamformer $\mathbf{W}_{gsc}$ performs better over high frequency bands which are covered by Subarray1, Subarray2 and Subarray3. It has decreased $SNR(f)$ over the frequency band below 900 Hz, which is covered by Subarray4. The Subarray4 beamformer $\mathbf{W}_{bst}$ achieves a large improvement of SNR, while the Subarray4 $\mathbf{W}_{anc}$ scheme has a 2 dB to 3 dB SNR improvement over the input. For the high frequency subbands, the $SNR(f)$ curves of the three schemes are close to each other. They obtain a nearly constant SNR improvement of 4 to 5 dB. The $SNR(f)$ curves of the four subarrays are in close agreement with the SINR values listed in Table 5.2.

The $NR(f)$ curves in Figure 5.7 provide the measure of reverberant interference rejection of the adaptive beamformers. The $NR(f)$ curves of the three fullband schemes have similar behavior to their $SNR(f)$ curves. The $NR(f)$ of the fullband $\mathbf{W}_{anc}$ scheme is pretty flat over the passband; the $NR(f)$ of the $\mathbf{W}_{gsc}$ decreases as the frequency increases; the $NR(f)$ of the $\mathbf{W}_{bst}$ is the highest over the passband.

For the three subband adaptive schemes, on the other hand, the $NR(f)$ curves exhibit several peaks at the edges of the subbands. This suggests that better de-reverberation performance is obtained at the transition bands of the subarrays. The $NR(f)$ curves of the low subband also behave similarly to the corresponding $SNR(f)$ curves. The $NR(f)$ of the low subband beamformer $\mathbf{W}_{gsc}$ is the lowest; the $NR(f)$ of

the $\mathbf{W}_{anc}$ scheme is pretty flat over the passband; the $NR(f)$ of the $\mathbf{W}_{bst}$ is the highest over the passband. The $NR(f)$ and $SNR(f)$ curves of the NAM-GSC beamformer $\mathbf{W}_{gsc}$ show that the de-reverberation performance of the NAM-GSC beamformer is degraded mainly due to the low frequency subband. The high frequency subbands are not affected as much by the reverberation interference. The reason for this is elaborated in Section 5.3.

5.2.4 Energy Decay Curves

The output Energy Decay Curves of the adaptive beamformers are plotted together in Figure 5.8. All subband schemes decay faster than their fullband counterparts. The EDC curves of the fullband and subband ANC schemes ($\mathbf{W}_{anc}$) decay much more rapidly at the beginning of the curves. This may suggest that the switched beam ANC schemes have better suppression of the low order images which are located closer to the array. The EDC curves of the fullband and subband best achievable GSC beamformers ($\mathbf{W}_{bst}$) decay slightly slower than the switched beam ANC ($\mathbf{W}_{anc}$) schemes but faster than the GSC beamformers ($\mathbf{W}_{gsc}$). The EDCs of the fullband GSC and subband NAM-GSC $\mathbf{W}_{gsc}$ beamformers decay slowly at the beginning. They converge to the same level as the other adaptive beamformers after $t = 0.15$ second. This means that the low order images are not suppressed effectively by the adaptive GSC beamformers, due to problems such as desired signal cancellation and leakage of reverberant noises. It also suggests that the low order images play the most significant role in causing the desired signal cancellation in the adaptive NAM-GSC beamformer.

5.3 Remarks on De-reverberation Performances

Through the analysis of the de-reverberation performances, it is concluded that:

1. the subband NASB-ANC scheme obtains a flat de-reverberation gain of approx-

(a) Fullband Schemes

(b) Subband Schemes

Figure 5.7: Reverberant noise reduction $NR(f)$ of the adaptive beamformers in a reverberant room.

Figure 5.8: Energy Decay Curves of the adaptive beamformers in a reverberant room.

imately 4 dB over the passband. The perfect VAD ensures that the adaptive noise cancelers are turned off so no desired signal is cancelled at the output. The de-reverberation gain is merely the contribution of the near field DFS beamformers;

2. the subband NAM-GSC adaptive beamformer performs well in reverberant environments over the high frequency subbands covering [900, 7200] Hz. It suffers from both the desired signal cancellation and reduced reverberant interference rejection over the low frequency subband of [50, 900] Hz. It obtains approximately 4 dB SINR improvement in the high frequency bands and 1.7dB SINR loss in the low frequency band;
3. the fullband adaptive GSC beamformer outperforms the subband NAM-GSC beamformer for de-reverberation. The fullband GSC beamformer suffers from the desired signal cancellation over the high frequency range; while the desired signal cancellation occurs over the low frequency band of the subband NAM-

GSC beamformer.

The desired signal cancellation is observed in the low subband of the NAM-GSC scheme. This is because the low order image sources fall within the near field of the low frequency band subarray. The sizes of Subarray1 to Subarray4 are 9.6cm, 19.2cm, 38.4cm, and 76.8cm, respectively. Their near field distances extend to 0.4m, 0.8m, 1.6m, and 3.2m, respectively. The lower subband has larger array size. The first order images are located at 2 to 5 meters from the array center. They are far field interference for the three high frequency subarrays. But they fall within the near field of Subarray4 (low subband). They contribute the most to the desired signal cancellation and the reduced interference rejection. High order images are observed by Subarray4 as isotropic noises which are less correlated with the desired signal. For Subarray1 to Subarray3, all images are received as far field interference. The sum of the far field images is observed as the isotropic noise with low correlation to the desired signal. Thus the desired signal cancellation is negligible in the three high frequency bands. To improve the low frequency subband performance, special coherent interference suppression algorithms are required to suppress the low order image sources.

The reduced de-reverberation performance in Subarray4 is also due to its insufficient aperture. Subarray4 covers more than an octave frequency band. It has limited capacity to suppress low frequency interference. One easy method to achieve better performance at the low frequency end is by adding more elements in the low band subarray and/or splitting the low frequency band into more subbands. The cost of this method is the increased system complexity. Another de-reverberation method is to design the low band subarray using the special optimization method proposed by Ryan [79]. This task is not carried out by this thesis.

Chapter 6

Spatial Affine Projection (SAP) Algorithm

A new Spatial Affine Projection (SAP) algorithm is developed to decorrelate the coherent interference for adaptive beamforming[110]. The SAP algorithm combines the Spatial Averaging method with the Affine Projection algorithm to destroy the coherency of the interference in both space and time domains. It can effectively suppress narrowband and broadband coherent interference. It can simultaneously improve the convergence of the adaptation with a small increase in computational complexity.

The detailed structure of the SAP algorithm, as well as the existing Spatial Smoothing algorithms, is introduced in Section 6.1. Its application and performance in far field beamforming is presented in Section 6.2. Finally, Section 6.3 shows that the direct extension of the far field SAP algorithm to near field adaptive beamforming is problematic. The near field SAP algorithm is reformulated using the near field robust adaptive beamforming technique proposed in Section 4.3.3.

6.1 The SAP Algorithm for Coherent Interference Suppression

As we have discussed in Section 3.3.1, the Spatial Smoothing (SS) algorithm has some decorrelation properties which enable it for coherent interference suppression. It is noticed that a time domain adaptive algorithm, the Affine Projection (AP) algorithm, also has a decorrelation property that makes it converge much faster than the LMS algorithm. The proposed SAP algorithm applies the AP algorithm to GSC beamformers in the space domain, thus combines the decorrelation properties of the SS algorithm and the AP algorithm together, and achieves coherent interference suppression and fast convergence simultaneously.

The Affine Projection (AP) algorithm was originally proposed by Ozeki *et al.* [69] for acoustical echo and noise cancellation. The algorithm and its fast version (FAP) [24, 86] have been investigated extensively in recent years. As a generalization of the Normalized LMS (NLMS) algorithm and the Recursive Least Square (RLS) algorithm, the family of AP algorithms improves the convergence of adaptive filters with reasonably low computational complexity. The details of the AP algorithms can be found in Appendix B.

The proposed SAP beamformer uses a subtractive preprocessor, a master beamformer and slaved beamformer, as depicted in Figure 6.1. The SAP algorithm is employed in the master beamformer. After pre-steering Δ_i and subtractive preprocessor, the snapshot samples $v(1, k)$, $v(2, k)$, $\dots$, $v(M - 1, k)$ are obtained as inputs to the master beamformer. The adaptive weights obtained in the master beamformer are then copied to the slaved beamformer for adaptive filtering. It has been shown that the subtractive preprocessor preserves the phase relationship of the signals at the slaved beamformer input [100, 104], thus the copied weights are effective for coherent interference suppression in the slaved beamformer.

To avoid the desired signal cancellation in the presence of multiple coherent inter-

Figure 6.1: An adaptive GSC beamformer with a subtractive pre-processor

fering signals, the SPSS scheme proposed in [70] uses the SS algorithm in the master beamformer. Figure 6.2 shows its implementation by a Generalized Sidelobe Canceler (GSC). The input samples $v(i, k)$ are grouped into p subarrays, each having q elements. The input vector of the i -th subarray is

$$\mathbf{v}_q(i, k) = [v(i, k), v(i+1, k), \dots, v(i+q-1, k)]^T \quad (6.1)$$

where $i = 1, 2, \dots, p$ and $q = M - p$.

If a tapped-delay-line of length K is included in the adaptive GSC beamformer, then a concatenated vector is formed for each subgroup

$$\mathbf{V}_N(i, k) = [v(i, k), v(i+1, k), \dots, v(i+q-1, k)]^T \quad (6.2)$$

where $i = 1, 2, \dots, p$ and $N = q \times K$. The adaptive GSC beamformer may be designed as discussed in Section 4.1.3.

The SS or SPSS algorithm using the NLMS adaptation is summarized as follows: for $i = 1, 2, \dots, p$

$$d(i, k) = \mathbf{W}_q^H \cdot \mathbf{V}_N(i, k) \quad (6.3)$$

$$\mathbf{x}_L(i, k) = \mathbf{C}_a^H \cdot \mathbf{V}_N(i, k) \quad (6.4)$$

$$y(i, k) = \mathbf{W}_a^H(i, k) \cdot \mathbf{x}_L(i, k) \quad (6.5)$$

$$e(i, k) = d(i, k) - y(i, k) \quad (6.6)$$

$$\mathbf{W}_a(i+1, k) = \mathbf{W}_a(i, k) + \mu \frac{\mathbf{x}_L(i, k) \cdot e^H(i, k)}{\mathbf{x}_L^H(i, k) \cdot \mathbf{x}_L(i, k)} \quad (6.7)$$

$$\mathbf{W}_a(1, k+1) = \mathbf{W}_a(p+1, k) \quad \text{when } i = p \quad (6.8)$$

where μ is the step size and L is the dimension of the unconstrained adaptive weight vector $\mathbf{W}_a$.

Applying the AP algorithm to beamformers in the time domain is straightforward, simply replacing the NLMS algorithm in Figure 6.2 by the AP algorithm. It is worth noting, however, that the decorrelation property of the AP algorithm employed in the time domain is fundamentally different from the decorrelation property of the SS or SPSS algorithm. It only improves the convergence of the adaptation, not the capability of the coherent interference suppression [109].

In contrast, our newly proposed SAP algorithm performs the affine projection in the space domain, as shown in Figure 6.3. The input vectors of the subarrays $\mathbf{v}_q(i, k)$ are fed into a set of GSC's in parallel. The set of signals $d(i, k)$ and $\mathbf{x}_L(i, k)$ resulting from Equation(6.3) and Equation(6.4) are collected to form the vector $\mathbf{D}(k)$ and the matrix $\mathbf{X}(k)$, respectively. Then they are fed into the adaptive GSC beamformer $\mathbf{W}_a$ and adapted by the SAP algorithm. That means the set of p GSC beamformers are processed simultaneously.

The SAP algorithm is formulated in Table 6.1, where μ is the step size, and δ is the regulation parameter. The projection order p is equal to the number of subarrays.

The capability of coherent interference suppression is limited by the array parameters p and q . Denote the number of spatially separated coherent interferences that can be suppressed as D . It has been proved [83] that $D \leq \min(p, q)$.

The computational complexity of the SAP algorithm may be divided into two parts: The SAP algorithm may be viewed as the straightforward Block Affine Projection algorithm [87]. Therefore the fast version of the AP algorithm (FAP) may

Figure 6.2: An adaptive GSC beamformer using Spatial Smoothing (SS) algorithm

Figure 6.3: An adaptive GSC beamformer using Spatial Affine Projection (SAP) algorithm

Table 6.1: Summary of the SAP algorithm

$d(i, k) = \mathbf{W}_q^H \cdot \mathbf{V}_N(i, k)$
$\mathbf{x}_L(i, k) = \mathbf{C}_a^H \cdot \mathbf{V}_N(i, k)$
1. $\mathbf{X}(k) = [\mathbf{x}_L(1, k) \quad \mathbf{x}_L(2, k) \quad \cdots \quad \mathbf{x}_L(p, k)]$
2. $\mathbf{D}(k) = [d(1, k) \quad d(2, k) \quad \cdots \quad d(p, k)]$
3. $\mathbf{R}(k) = \mathbf{X}^H(k) \cdot \mathbf{X}(k) + \delta \mathbf{I}$
4. $\mathbf{e}(k) = \mathbf{D}^H(k) - \mathbf{X}^H(k) \cdot \mathbf{W}_a(k)$
5. $\mathbf{W}_a(k+1) = \mathbf{W}_a(k) + \mu \mathbf{X}(k) \mathbf{R}^{-1}(k) \cdot \mathbf{e}(k)$

be used for SAP to reduce the computational complexity. For each time instant k , the SPSS algorithm has to adapt p times—each subgroup adapts once. So the SPSS-NLMS algorithm requires $(2L + 1)p$ additions and multiplications. The SPSS-RLS algorithm using the Fast Transversal Filter (FTF) method [12] requires $(7L + 14)p$ additions and multiplications. But the SAP algorithm only adapts once for each k . So the SAP using FAP method requires $2L + 20p$ additions and multiplications [24]. The SAP using conventional APA method requires $(p + 1)L + O(p^3)$ multiplications [86]. The computational complexities are compared in Table 6.2 for $L = 10, p = 5$ and $L = 80, p = 6$. The computational complexity of the SAP is similar to that of the SPSS-NLMS algorithm when L and p are small. When L and p are large, however, the SAP will have lower computational complexity than the SPSS-NLMS algorithm.

6.2 Performances of the SAP Algorithm in Far Field Beamforming

To illustrate the performance of the SAP algorithm, a beamformer consisting of 12 equi-spaced elements is considered. The number of subarrays or the projection

Table 6.2: Comparison of computational complexity of the SAP and SPSS algorithm

	Formula	$L = 10, p = 5$	$L = 80, p = 6$
SAP with APA	$(p + 1)L + O(p^3)$	185	696
SAP with FAP	$2L + 20p$	120	360
SPSS-NLMS	$(2L + 1)p$	105	805
SPSS-RLS	$(7L + 14)p$	420	2670

order is $p = 5$. The desired signal is narrowband with the Signal-to-Noise Ratio (SNR) of 30dB, arrived at $\theta_1 = 90^\circ$. Four coherent interfering signals have the power of $INR = [25, 20.5, 20, 17]$ dB, relative to the background noise. They arrive at directions $\Theta = [35^\circ, 70^\circ, 110^\circ, 130^\circ]$. The desired signal is chosen to be $s_1(t) = \sin(0.4\pi t)$. The four coherent interfering signals are also sinusoidal signals of the same frequency with fixed phase differences. The beam patterns of the beamformer after convergence are plotted in Figure 6.4, where interfering signals are indicated by $J_1, J_2, \dots, J_4$. All three algorithms are able to suppress the coherent interference. But the nulls obtained by the SAP algorithm and the SPSS algorithm are much deeper than those of the SS algorithm, illustrating the effectiveness of the subtractive preprocessor. The SAP algorithm also outperforms the SPSS algorithm because it provides additional decorrelation by the means of affine projection, compared to simply subgrouping.

The excess Mean Squared Error (MSE) curves of the SAP algorithm and the SPSS algorithm are shown in Figure 6.5. The step size for the SAP and SPSS-NLMS adaptation is $\mu = 0.001$. The forgetting factor for the SPSS-RLS algorithm was $\lambda = 0.999$. The curves were averaged over 50 trials. It is clear that the SAP algorithm converges much faster than the SPSS-NLMS algorithm, and it is about 3 times slower than the SPSS-RLS algorithm. With narrowband inputs, the computer simulations show that the convergence of the SAP is comparable to that of the SPSS-

Figure 6.4: Beampatterns of the SAP and SPSS algorithms with far field narrowband coherent interference

RLS algorithm when the number of coherent signals D is close to the projection order p . If D is much smaller than p , then the convergence of the SAP algorithm becomes closer to the SPSS-NLMS algorithm.

The SAP algorithm is also evaluated under the input of broadband coherent signals. The desired signal S_1 is a colored noise band limited to $B = [0.2, 0.4]$, impinging on the array at $\theta_1 = 90^\circ$. The signal to noise ratio is $\text{SNR} = 15\text{dB}$. The three coherent interfering signals S_2 , S_3 and S_4 are delayed and scaled versions of S_1 , impinging on the array at $\Theta = [35^\circ, 70^\circ, 130^\circ]$. The Interference-to-Noise-Ratio is $\text{INR} = [8.0, 12.0, 11.2]\text{ dB}$. The array parameters are the same as that in the narrowband case, except that the number of taps at each element is $K = 15$, and the spacing between the elements is half the wavelength of the high frequency edge. Since the band ratio is 2:1, the array's aperture at the lowest frequency is reduced by a half. With the number of subgroup $p = 5$ and the number of elements $q = 7$, the array is capable of suppressing four interfering signals at the high frequency end, but

Figure 6.5: Convergence of the SAP and SPSS algorithms with far field narrowband coherent signals.

only two or three at low frequency range.

The broadband beampatterns of the SAP and SSPS algorithms are shown in Figure 6.6(a) and Figure 6.7(a). The beampatterns are evaluated at the four in-band frequencies 0.25, 0.30, 0.35 and 0.40. Both algorithms can place consistent nulls at the interference directions and maintain unit gain at the look direction. The nulls obtained by the SAP algorithm are slightly deeper than those obtained by the SSPS algorithm. The frequency responses of the two algorithms are similar, as shown in Figure 6.6(b) and 6.7(b). They have flat gain for the look direction within the designed band $[0.2, 0.4]$. The attenuation at the interference directions is more than 20 dB over the band $[0.25, 0.40]$. The coherent interference suppression is slightly degraded over the low frequency band $[0.20, 0.25]$.

The convergence behaviors of the SAP and SPSS algorithms are very close to each other under the broadband condition. When the projection order equals the number of subgroups, the narrowband SAP algorithm converges much faster than that of the

(a) Beampatterns at four in-band frequencies

(b) Frequency responses

Figure 6.6: Responses of the SAP algorithm with far field broadband coherent signals, where $S1$ is the desired signal, $S2$, $S3$ and $S4$ are the coherent interference.

(a) Beampatterns at four in-band frequencies

(b) Frequency responses

Figure 6.7: Responses of the SSPS algorithm with far field broadband coherent signals, where $S1$ is the desired signal, $S2$, $S3$ and $S4$ are the coherent interferences.

SPSS-NLMS algorithm, as shown in Figure 6.5. For the broadband cases, however, the SAP algorithm converges at the same rate as the SPSS-NLMS algorithm. This is due to the fact that the affine projection over the space domain is only able to decorrelate the spatial coherency but not the temporal correlation. To improve the convergence rate of SAP for broadband signals, the projection order may be increased to include both space and time domain vectors. For example, let the projection order over the space domain remain $p_s = 5$ and add a projection order $p_t = 2$ over the time domain. The total projection order is then $p = p_s p_t = 10$. In this case, the convergence of the SAP algorithm is faster than that of the SPSS-NLMS algorithm, as shown in Figure 6.8.

Figure 6.8: Convergence of the SAP and SPSS-NLMS algorithms with far field broadband coherent signals.

6.3 Spatial Averaging Algorithms in Near Field Beamforming

The SAP and SPSS algorithms can suppress narrowband and broadband coherent interference due to their spatial decorrelation property obtained by spatial shifting and averaging. The decorrelation property of the SAP and other spatial averaging algorithms is based on the assumption that a far field signal impinges on every element of the array with the same DoA. However, the situation is different for near field signals. A near field source travels different distances to each element and arrives at each element at different angles. This near field curvature causes problems for the near field SAP or SPSS in two aspects. First, the subtractive preprocessor in the master beamformer is not effective for near field signals. Secondly, the direct application of subgrouping and spatial smoothing to near field adaptive beamforming not only fails to destroy the coherency of the interference, but also causes malfunctioning of the adaptive GSC beamformer.

The first problem is associated with the subtractive preprocessor. In the far field case, the subtractive preprocessor in Figure 6.1 is capable of removing the desired signal and preserving the spatial relationship of the coherent interfering signals. In the near field case, however, the subtractive preprocessor can not preserve the spatial locations of the signals after removing the desired signal.

The second problem is encountered by near field spatial smoothing. The problem can be intuitively explained by Figure 6.9. Assume that the desired signal S_1 is fixed on the $x-y$ plane at the location $\mathbf{x}_{s1} = (r_{s1}, \theta_1) = (8\lambda, 90^\circ)$, where r_{s1} is the distance from the origin x_0 and θ_{s1} is the impinging angle measured with respect to the array axis. The coherent interfering signals S_2, S_3 and S_4 are located at $\mathbf{x}_{s2} = (10\lambda, 35^\circ)$, $\mathbf{x}_{s3} = (12\lambda, 70^\circ)$, and $\mathbf{x}_{s4} = (15\lambda, 130^\circ)$, respectively, where λ is the wavelength of the high frequency edge of the pass band. The array consists of 11 equally spaced elements, with spacing $d = \lambda/2$. The size of the array is 5λ . All the signals are

within the near field of the array which extends to 25λ . The array geometry and signal locations may be easily found in broadband microphone array applications. For example, the subarray for the band $[1800, 3600]\text{Hz}$ has $\lambda = 9.6\text{ cm}$. So the array size of 5λ is 48 cm and the focal point of 8λ is 76.8 cm. This is very close to the real world scenario of computer telephony applications.

Figure 6.9: Subgrouping of a near field linear array. The figure is not to scale.

Assume that the 11 elements are grouped into 5 subgroups, each having 7 elements. Subgroup3 has its phase center at the origin x_0 , as illustrated in Figure 6.9. The signal locations observed by Subgroup3 remain the same as those by the whole array:

$$\begin{aligned}
 S_1 : \quad \hat{\mathbf{x}}_{3s1} &= (8\lambda, 90^\circ) \\
 S_2 : \quad \hat{\mathbf{x}}_{3s2} &= (10\lambda, 35^\circ) \\
 S_3 : \quad \hat{\mathbf{x}}_{3s3} &= (12\lambda, 70^\circ) \\
 S_4 : \quad \hat{\mathbf{x}}_{3s4} &= (15\lambda, 130^\circ)
 \end{aligned}$$

Meanwhile, the signal locations observed by other subgroups are different because the phase center locations of the subgroups are shifted. The phase center of Subgroup1 is shifted to x_{-2} . The signal locations observed by Subgroup1 become

$$\begin{aligned} S_1 : \quad \hat{\mathbf{x}}_{1s1} &= (8.1\lambda, 89.3^\circ) \\ S_2 : \quad \hat{\mathbf{x}}_{1s2} &= (10.8\lambda, 32^\circ) \\ S_3 : \quad \hat{\mathbf{x}}_{1s3} &= (12.4\lambda, 65.6^\circ) \\ S_4 : \quad \hat{\mathbf{x}}_{1s4} &= (14.4\lambda, 126.9^\circ) \end{aligned}$$

Now applying the Spatial Smoothing (SS) algorithm to the near field array is to average the correlation matrices of the subarrays. The signal subspace of the resulting correlation matrix will be the combination of the observed signal subspaces at all five subarrays. With the LCMV adaptive beamformer designed to focus at $\mathbf{x}_{s1} = (8\lambda, 90^\circ)$, the desired signal S_1 received at Subgroup1 is treated as an interfering signal located at $\mathbf{x}_{1s1} = (8.1\lambda, 89.3^\circ)$, which is very close to the focal point. The received desired signals at Subgroup2, Subgroup4 and Subgroup5 are also treated as three different interfering signals. The signals S_2, S_3 and S_4 are processed similarly by the SS or SAP algorithm. As a result, the near field adaptive beamformer will cause severe cancellation of the desired signal.

This problem can be solved by employing the near field robust beamforming techniques proposed in Section 4.3.3. A spatial region around the point $\mathbf{x}_{s1} = (8\lambda, 90^\circ)$ is constrained for each subgroup array. Consequently, the desired signal observed by every subgroup array is passed with unit gain, and the spatial averaging will not cause the desired signal cancellation.

The near field SAP and SS algorithms using the robust beamformers are able to suppress the coherent interference. The performances of the reformulated near field SAP and SS algorithms are demonstrated by the following example. The desired signal is S_1 at $(8\lambda, 90^\circ)$ with a SNR of 15dB. The coherent interference signals are S_2 and S_3 located at $(10\lambda, 35^\circ)$ and $(15\lambda, 130^\circ)$, respectively. The interfering signals are

received at the array center with $\text{INR}=[8.0, 12.0]$ dB. The three signals are broadband coherent color noises with normalized bandwidth $B = [0.2, 0.4]$. The array has 11 elements grouped into 5 subgroups. No subtractive preprocessor but near field SAP or SS-NLMS algorithm is used. Adaptive GSC implementations of the SAP and SS are the same as that depicted in Figure 6.3 and Figure 6.2.

The converged near field SAP and SS algorithms have almost identical beam patterns and frequency responses. The beam patterns of the near field SAP algorithms are plotted in Figure 6.10(a). Consistent nulls are placed around the interference locations and a peak is formed at the desired signal location. But the sidelobes are much higher than those in the far field case, as shown in Figure 6.7(a). The frequency responses of the near field SAP and SS algorithms are illustrated in Figure 6.10(b). The two interference locations have low gain over the passband; while the unit gain at the desired signal location is preserved.

Furthermore, the near field SAP or SS algorithm is found, via computer simulations, to have reduced capacity of coherent interference suppression. An M -element array can suppress maximum of $D = (M - 1)/2$ far field coherent interfering signals. With near field coherent interference, this number reduces by approximately one third. The reason for the reduced capacity of the near field SAP and SS algorithm remains unknown.

In conclusion, the extension of the SAP algorithm and SS algorithm to near field adaptive beamforming has to involve the removal of the subtractive pre-processor and the reformulation of the algorithms by the robust near field adaptive beamforming technique. The reformulated near field SAP and SS algorithms are capable of suppressing near field coherent interference. Their capacity of coherent interference suppression is less than $D = \min(p, q)$, where p is the number of subgroups and q is the number of elements in each subgroup. In terms of coherent interference suppression, the near field SAP algorithm and SS algorithm achieve approximately 20dB attenuation to near field coherent interfering signals.

(a) Beampatterns at four in-band frequencies

(b) Frequency responses

Figure 6.10: Responses of the near field SAP and SS-NLMS algorithm with near field broadband coherent signals.

Chapter 7

Experimental Evaluation of the STS Beamformers

In this chapter, the performances of the NAM-GSC beamformer and the NASB-ANC are evaluated using the experimental data recorded in an anechoic chamber and a real conference room. Section 7.1 describes the experimental equipment, measurement procedures and environments. Section 7.2 presents the data processing techniques and the performances of the NAM-GSC and the NASB-ANC.

7.1 Description of the Experiment

7.1.1 Measurement Apparatus

A multi-channel audio recording system, as shown in Figure 7.1, was used for microphone array recordings. The system consisted of two parts: the generator and the recorder. The generator used a Compact Disk (CD) player to produce the sound sources. The recorder consisted of high quality microphones, multi-channel pre-amplifiers, multi-channel A/D converters, and a personal computer. The details of the apparatus are listed in Table 7.1.

Figure 7.1: The multi-channel microphone array recording system

Table 7.1: The experimental apparatus

	Manufacturer	Model	Description
CD Player	TASCAM	CD-150	Compact Disk Player
Loudspeaker	TANNOY	REVEAL	Omni-directional
Microphone	AUDIO-TECHNICA	AT803b or AT831	Miniature Omni-directional Condenser
Pre-Amplifier	ALLEN & HEATH	MixWizard WZ12:2DX	12-channel Pre-Amp
A/D Converter	MIDIMAN	Delta 1010	8-channel Digital Recording System
Personal Computer	Dell	Pentium 300MHz	Microsoft WIN98, 128MB RAM

The “clean” sound source was recorded on a recordable CD disk. The signal source was 16 bit PCM waveform sampled at 44.1 kHz. The length of the signal source was approximately 2 minutes. It consisted of segments of white noise, a cue frame, male speech, female speech and music clippings. Each segment of the signal was separated by 4 seconds of silence so it could be easily extracted from the recorded data. The white noise segment at the beginning of the signal source was sufficiently long to allow manual operation of the ‘play’ and ‘record’ buttons. The cue frame was designed for synchronization of non-simultaneously recorded data. It consisted of three single frequency tones of 300 Hz, 3.0 kHz and 7.0 kHz. Each single tone had a length of 0.2 second. The speech and music segments were broadband signals originally sampled at 16 kHz or higher. They were re-sampled to 44.1 kHz for the CD disk.

The microphones were omni-directional condenser microphones possessing a flat frequency response from 50 Hz to 15 kHz. They were mounted on a plywood board with the nested array geometry shown in Figure 4.2. There were a total of 11 microphones nested into four subarrays. Each subarray had 5 elements. The inter-element spacings of the four subarrays were 2.4 cm, 4.8 cm, 9.6 cm, and 19.2 cm, respectively. The total size of the nested microphone array was 76.8 cm.

Other apparatus were mounted on a metal equipment rack. The multi-channel pre-amplifier had 8 mono inputs for microphones with 70 dB gain range. Each input channel had a filter with the low cut-off frequency at 100 Hz. The MIDI digital recording system was configured as a PCI host card plus an external rack-mount unit, which housed the A/D (and D/A) converters. It had 8 data channels with bit widths and sampling rates up to 24-bit/96 kHz. A/D converters had a high dynamic range (A-weighted measured) of 109 dB, and low distortion (measured THD @0 dBFS) of less than 0.001. The PCI host card of the MIDIMAN digital recording system was installed in the personal computer. A multi-channel recording software was also installed and configured in the personal computer for 8 channel digital recording.

After the set-up of the equipment, channel calibration was performed for each

microphone channel by adjusting the gain of the pre-amplifier. The calibration ensured that no amplitude clipping occurred and the gains of the received signals were accurate within 0.25 dB.

7.1.2 Measurement Procedures and Environments

The experiments were performed in an anechoic chamber and a small conference room, respectively. The source CD was played back at several locations and the microphone recordings were done separately for each location. The data were recorded in 16 bit PCM format with a rate of 48,000 samples per second.

Multiple runs of recording were required for each location because our recording equipment had only 8 channels available and the 11-element array could not be recorded simultaneously. Each recording run used 7 channels to simultaneously record the 7 elements of two adjacent subarrays — the first recording run for the 7 elements of Subarray1 and Subarray2, the second run for Subarray2 and Subarray3, and the third run for Subarray3 and Subarray4. The second run was redundant but it turned out to be very helpful in case there were damaged data in the other two runs due to various reasons. It also helped with the cueing or synchronization of the multiple runs.

The multiple recording runs were not synchronized due to the non-synchronized manual operation of the ‘play’ and ‘record’ buttons. The synchronization was carried out for the recorded data. It was achieved by identifying and aligning the cue frames of the superimposed microphone elements. After the synchronization, the recorded data were down sampled to 16 kHz.

The microphone recordings were first carried out in an anechoic chamber in Loeb Building , Carleton University. The equipment rack was placed outside of the anechoic chamber. The microphone array and the loudspeaker were inside the anechoic chamber. The connection cables run through a small hole on a wall of the chamber and it was covered by sound absorbing forms. The sound source CD was played back

Figure 7.2: Signal locations in the anechoic chamber

at several locations in front of the array, as shown in Figure 7.2. A Cartesian coordinate system was defined such that the array center was at the origin and the elements laid along the x axis. The sound sources were on the $x-y$ plane, as illustrated in Figure 7.2. Their Cartesian coordinates were $(0 \text{ m}, 0.60 \text{ m}, 0 \text{ m})$, $(0.55 \text{ m}, 0.89 \text{ m}, 0 \text{ m})$ and $(-0.59 \text{ m}, 0.80 \text{ m}, 0 \text{ m})$, respectively. Their corresponding spherical coordinates were $(0.6 \text{ m}, 90^\circ, 90^\circ)$, $(1.05 \text{ m}, 58.4^\circ, 90^\circ)$, and $(0.99 \text{ m}, 126.4^\circ, 90^\circ)$, respectively.

Secondly, the recordings were performed in a small conference room in an engineering building at Carleton University. The size of the room was $5.0\text{m} \times 3.8\text{m} \times 3.5\text{m}$. The room was constructed with double plaster board walls, cement floor with linoleum tiles, acoustic tile drop ceiling below a corrugated steel roof, and a double wooden door. There were a square table and 6 padded chairs in the middle of the room, as shown in Figure 7.3. The equipment rack stood in a corner of the room beside the door. The microphone array was placed on a desk in another corner of the room. The phase center of the array was located at 1 meter away from the floor and the two walls. The angle between the array axis and the walls was $\beta = 45^\circ$. The sound source was located 0.6 meter away from the array center on the y axis. This arrangement was similar to the simulation in Chapter 5. The background noise level in the room was low compared to a typical office environment. Consequently, these recordings were suitable for examination of the beamformer's de-reverberation performance.

Figure 7.3: Measurement environment of the conference room

7.2 Data Analysis and Results

The two subband adaptive schemes — the NAM-GSC beamformer and the NASB-ANC, were evaluated using the recorded data. Both schemes used the 11-element nested array with the 4 subband structure.

The NAM-GSC beamformer utilized the robust beamforming design in each subarray. The robust GSC beamformers were required for two reasons: (1) the locations of the signal sources were only accurate to a few centimeters; (2) the sound source generated by the CD player was not a restricted point source. The constrained spatial region was $\Delta r = 0.1$ m and $\Delta\theta = 5^\circ$; refer to Figure 4.21. The robust GSC beamformer used 32 taps per element.

The NASB-ANC used three DFS beamformers and two auxiliary channel ANC's for the noise rejection application. Each DFS beamformer had 16 taps per element. Each auxiliary channel of the ANC had 32 taps. The parameters were the same as those used in the simulation of Section 4.4.

7.2.1 Noise Reduction Performances

The recordings made in the anechoic chamber were used for noise rejection evaluations. The desired signal S_1 was a female speech located at (0 m, 0.60 m, 0 m) in the Cartesian coordinate system. Two interfering signals S_2 and S_3 were a music signal at (0.55 m, 0.89 m, 0 m), and a mix of male speech segments and female speech segments located at (−0.59 m, 0.80 m, 0 m). The interfering female speech and the desired female speech were generated from different talkers. Each signal was of length 18 seconds and the sampling rate was 16 kHz.

The speech signals were some English sentences as follows:

- S_1 : “Welcome to the Code Composer Studio multimedia tutorial. This tutorial has been created to show developers how to utilize a few of the Code Composer Studio’s key features. It is complementary to the tutorial found both in the

on-line help and as a pdf file located on the program CD-ROM (female 1)."

- S2: "Incoming file transfer (male)"; "Incoming chat request (male)"; "This is the speaker and sound card test for the Intel configuration wizard. As you listen to this recording, adjust the volume to a comfortable level using the configuration wizard slider bar (female 2)."

The three signals were received by the array separately. They were scaled to have the same power at the array's phase center. Uncorrelated white noises were also added to each element with -20 dB power with respect to the signals. The power spectrum densities of the input signals are shown in Figure 7.4. The two speech signals ($S1$ and $S2$) had energies concentrated in the low frequency band, while the music signal ($S3$) had high energies spread in the lowest subband and the highest subband. The input SINRs of the subbanded signals were not the same in every subband. This was different from the simulations in Chapter 4 where the signal energy was uniformly distributed within the passband.

The three input signal sources were fed into the adaptive NAM-GSC beamformer simultaneously. The optimum weights were obtained for each subarray of the NAM-GSC beamformer. The subarrays achieved SINRs of 29.1 dB, 26.2 dB, 24.9 dB, 25.1 dB, respectively. The output of the NAM-GSC beamformer achieved a SINR of 24.6 dB and NR factor of 27.9 dB. This performance was better than the simulation example of the robust NAM-GSC beamformer which used only 16 taps per element. It was comparable to the performance of the simulated NAM-GSC beamformer without location errors.

The PSD of the NAM-GSC output was also shown in Figure 7.4. The output PSD was very close to the desired input $S1$. The difference between the output PSD and the desired signal PSD was the contribution of the interference power. It was low, indicating that high noise reduction factor was achieved. The experimental results verified that the design and simulation of the robust NAM-GSC beamformer was successful.

Figure 7.4: PSD of the three audio input signals. $S1$ was the desired signal. $S2$ and $S3$ were the interference.

The same three input signals were processed by the NASB-ANC scheme. The adaptation of each subband ANC was controlled by a simple power estimation VAD. The VAD estimated the power of the desired signal $S1$ in every frame of several hundred samples. A threshold was set for the VAD according to the subband signal energy. If the estimated power was above the threshold, then the VAD was on and the adaptation of the ANC was stopped. If the VAD was off, then the ANC would adapt to the signal inputs. The NLMS algorithm was used for the ANC with $\mu = 0.02$. The adaptation was converged within 10 seconds of the signal input. This corresponded to 16×10^4 samples in Subarray1 (high band), 8×10^4 samples in Subarray2, 4×10^4 samples in Subarray3, and 2×10^4 samples in Subarray4 (low band). The output power and SINR were computed using the rest of the signal segments. The results were listed in Table 7.2. The compound NASB-ANC achieved a SINR of 23.9 dB at the output, and a NR of about 26.4 dB. The results were slightly inferior to the simulated NASB-ANC scheme in Section 4.4, where 28.9 dB SINR was achieved with

location errors. The better SINR was obtained in the simulation because the input signals were colored noises with flat spectra in the passband and the adaptation of the ANC was performed at the absence of the desired signal. While in the experiment, real speech and audio signals were used and a simple VAD was implemented to turn the ANC on and off. The VAD was not perfect and the adaptation of the ANC was affected too. It resulted in 5 dB degradation of the noise rejection performance, which was within expectations.

7.2.2 De-reverberation Performances

The recordings made in the conference room were used for the de-reverberation performance evaluation. The signal source was the same female speech used in the noise rejection case. The recorded reverberant signals were processed by the NASB-ANC and the NAM-GSC beamformer. The ANC of the NASB-ANC scheme was switched off at all times by the VAD. The NAM-GSC was adaptive during the presence of the speech, and the optimum weights were obtained. The outputs of the beamformers were obtained as the total outputs.

To evaluate the SINR at the beamformers' output, the input signal had to be decomposed into the direct path and the reflected paths. This was easily performed in the simulation. However, separating the direct path signal from its reflected paths was difficult for the real room recordings. Thus the clean signal source recorded in the anechoic chamber was used as the direct path signal. All signals were of length 18 seconds. This signal was filtered by the two subband beamformers and the outputs were obtained as the desired outputs. The reverberant interference power was estimated as the difference between the total output power and the desired output power. Then the SINR was measured based on these output powers. The NASB-ANC obtained a de-reverberation gain of 3.5 dB. The NAM-GSC achieved a de-reverberation gain of 3.2 dB. The measured results were close to the simulated ones.

Figure 7.5 plotted the waveforms of the direct path signal, the reverberant signal,

Table 7.2: SINR of the NASB-ANC and its subbands for noise rejection using experimental data

Subarray1 covering the subband $B_1 = [3.6, 7.2]$ kHz.

	P_s	P_d	P_i	SINR
Beamformer Input	9.8035	3.5765	6.2270	-2.4 dB
ANC Input	4.1785	3.7961	0.3734	10.1 dB
ANC Output	3.1797	3.1744	0.0047	28.3 dB

Subarray2 covering the subband $B_2 = [1.8, 3.6]$ kHz.

	P_s	P_d	P_i	SINR
Beamformer Input	5.4712	0.7957	4.6767	-7.6 dB
ANC Input	0.7638	0.7491	0.1392	5.9 dB
ANC Output	0.7228	0.7209	0.0019	25.8 dB

Subarray3 covering the subband $B_3 = [0.9, 1.8]$ kHz.

	P_s	P_d	P_i	SINR
Beamformer Input	5.5274	1.0797	4.4476	-6.1 dB
ANC Input	1.4597	1.3051	0.1515	9.3 dB
ANC Output	1.0053	1.0028	0.0035	24.6 dB

Subarray4 covering the subband $B_4 = [0.1, 0.9]$ kHz.

	P_s	P_d	P_i	SINR
Beamformer Input	15.688	7.2600	8.4282	-0.4 dB
ANC Input	7.2087	6.4471	0.7824	9.2 dB
ANC Output	6.2442	6.2093	0.0226	24.4 dB

Compound NASB-ANC covering the band $B = [0.1, 7.2]$ kHz.

	P_s	P_d	P_i	SINR
Beamformer Input	169.37	67.260	102.09	-1.8 dB
ANC Input	69.856	62.321	7.6469	9.1 dB
ANC Output	58.600	58.302	0.2335	23.9 dB

and the output signal processed by the NAM-GSC beamformer. Only the first 1 second of the signals was plotted to show the details of the reverberation. The clean signal recorded in the anechoic chamber had very low noises in the non-speech frames, as shown in Figure 7.5(a). The reverberant signal recorded in the conference room was shown in Figure 7.5(b). The non-speech frames were covered by the reflected speech signals, except for the beginning of the signal. The waveform of the speech signal was also changed from the clean speech. Figure 7.5(c) showed the output signal processed by the NAM-GSC beamformer. The reverberation was partially suppressed and the speech waveform was restored close to the clean signal. The benefit of de-reverberation was evident.

The output signal waveform of the NASB-ANC scheme was similar to the one obtained by the NAM-GSC beamformer.

7.2.3 The PAMS Test

The Perceptual Analysis/Masurement System (PAMS) is an objective test of Listening Effort (LE) and Listening Quality (LQ) specified by ITU-T Recommendation P.800. The Mean Opinion Score (MOS) calculated by PAMS is typically within one half a MOS of that determined by a well controlled subjective test in a laboratory [117, pp.17-23]. The standard MOS gives a measure of perceptual quality, as listed in Table 7.3.

A Digital Speech Level Analyzer (DSL), made by Malden Electronics Ltd, was used to perform the PAMS test for the NAM-GSC beamformer and the NASB-ANC. The clean speech signal source was used as the reference input to the DSL. The input signal at the array's phase center and the output signals of the beamformers were the test signals fed separately to the DSL. The resulting LE and LQ scores are listed in Table 7.4.

The results of the noise rejection experiments showed that the noisy input at the array had a LQ score of only 1.0. The NAM-GSC beamformer output improved the

(a) The clean signal source recorded in an anechoic chamber

(b) The reverberant signal recorded in a conference room

(c) The output signal after de-reverberation processing

Figure 7.5: Waveforms of the speech signals for de-reverberation

Table 7.3: The MOS standard

	Listening Quality	Listening Effort
5	Excellent	Complete relaxation possible; no effort required
4	Good	Attention necessary; no appreciable effort required
3	Fair	Moderate effort required
2	Poor	Considerable effort required
1	Bad	No meaning understood with any feasible effort

Table 7.4: Listening Effort (LE) and Listening Quality (LQ) scores obtained by the PAMS test

	Noise Rejection		De-reverberation	
	LE	LQ	LE	LQ
Array Input	1.8	1.0	2.9	2.4
NAM-GSC Output	4.0	3.6	3.9	3.1
NASB-ANC Output	4.2	3.5	3.5	3.1

LQ score to 3.6, which corresponded to the NR factor of 27.9 dB. The NASB-ANC scheme improved the LQ to 3.5, which corresponded to the NR factor of 26.4 dB.

The recorded reverberant input at the array had a LQ=2.4. The outputs of the NAM-GSC and the NASB-ANC both achieved LQ of 3.1, which corresponded to the de-reverberation gains of 3.5 dB and 3.2 dB, respectively.

The LE scores were higher than the corresponding LQ scores. This was common to all experiments, according to [117]. All the results were within expectations. They verified the experiments of the NAM-GSC beamformer and the NASB-ANC.

Chapter 8

Conclusion

This thesis has investigated broadband adaptive beamforming for applications where signals are located in the near field of an array. The primary application of this research is hands-free sound pickup and speech enhancement for wideband computer telephony. The standard frequency band of interest is $[50, 7000]$ Hz and the sampling frequency is 16 kHz. The size of the array is limited by terminal installation. The desired signal target is located in 0.5 to 1 meter range from the array. The signal is usually corrupted by undesirable sound sources, environmental noises and reverberant interference. The technical challenges to the *near field broadband adaptive* beamformer include: (1) many well established far field beamforming techniques are not applicable to near field beamforming because the curvature of the near field signals are large in the array's aperture; (2) broadband beamforming is required for the wide frequency band, but frequency dependent beampattern variations impair the performance of the beamformer; (3) adaptive beamformers in reverberant environments suffer from the desired signal cancellation due to high correlation between the direct path signal and the reflected signals.

As a compromise solution to the three problems encountered in near field broadband adaptive beamforming, this thesis has proposed a spatial-temporal subband (STS) adaptive beamforming system which incorporates a spatial subband array with

a temporal multirate subband filter bank and employs near field beamforming techniques in each subband. The STS beamforming system enhances the performances of near field broadband beamformers in terms of interference rejection, convergence of adaptation, and de-reverberation. It also enables parallel processing of the adaptive subband beamformers and improves the computational efficiency.

Three specific STS adaptive beamformers have been developed using different harmonically nested arrays, multirate filter banks and near field beamformers:

1. the Nested Array Quadrature Mirror Filter (NAQMF) beamformer using near field adaptive GSC beamformers and critically sampled QMF banks;
2. the Nested Array Multirate Generalized Sidelobe Canceler (NAM-GSC) using near field adaptive GSC beamformers and non-critical sampling multirate subband filters;
3. the Nested Array Switched Beam Adaptive Noise Canceler (NASB-ANC) using fixed Delay-Filter-and-Sum (DFS) beamformers plus adaptive noise cancelers (ANC) and non-critical sampling multirate subband filters.

For the wideband telephony application, the three STS adaptive beamformers were designed using an 11-element nested array split into 4-octave subbands. The sampling frequency was 16 kHz for all three systems. The NAQMF beamformer was designed to cover the frequency band up to 8.0 kHz with an array size of 64 cm. The 4 subbands were assigned as [4.0, 8.0] kHz, [2.0, 4.0] kHz, [1.0, 2.0] Hz, and [0.05, 1.0] kHz, respectively. The subbands were critically sampled at 8 kHz, 4 kHz, 2 kHz, and 1 kHz, respectively. Both the NAM-GSC and the NASB-ANC used a 4-subband nested array of size 76.8 cm to cover the frequency band up to 7.2 kHz. The subbands were allocated as [3.6, 7.2] kHz, [1.8, 3.6] kHz, [0.9, 1.8] kHz, and [0.05, 0.9] kHz, respectively. The sampling frequencies for the subbands were 16kHz, 8 kHz, 4 kHz and 2 kHz, respectively. The oversampling rate was 2. The different subband allocation of the NAQMF beamformer is required by the practical implementation of

the critically sampled multirate system. The differences between the NAM-GSC and the NASB-ANC include:

1. In the structural aspect, each subarray of the NASB-ANC scheme uses a fixed DFS beamformer followed by an ANC, while each subarray of the NAM-GSC scheme uses a single stage adaptive GSC beamformer. So the NAM-GSC scheme has a simpler structure and better computational efficiency than the NASB-ANC scheme;
2. In the robustness aspect, the NAM-GSC beamformer requires a special constraint design to achieve limited robustness against location errors. It can only tolerate small location errors on distance and impinging angle. The NASB-ANC scheme is much more robust against location errors without loss of performance. It can easily avoid the desired signal cancellation by switching off the auxiliary channels of the ANC when needed;
3. In the application aspect, the NAM-GSC beamformer requires only the knowledge of the desired signal location or the focal point. It can adaptively suppress the interfering signals without knowing their locations and their statistics. The NASB-ANC scheme, however, requires the estimates of the desired signal location and all interfering signal locations. Two or three dimensional location estimation is a non-trivial task, requiring additional algorithms. The NASB-ANC also requires a VAD to adaptively suppress interference.

The performances of the three STS beamformers have been improved significantly compared to their fullband beamformers of the same array geometry. Computer simulations have shown that the three STS beamforming systems can reduce the frequency dependent beampattern variations to the extent which occurs within an octave frequency band. They can achieve higher noise reduction using less adaptive weights than the fullband beamformers. They can improve the convergence of adaptation and reduce the computational complexity. The use of near field beamforming

also improves the de-reverberation performance of the STS systems. The STS beamformers can reduce the frequency dependent beampattern variation to less than 30° , while the fullband GSC beamformer of the same array has 80° beampattern variation. The NAQMF beamformer achieves a noise reduction factor of 25.5 dB using 21 taps in each subband beamformer. The NAQMF beamformer has higher residual errors at convergence than its fullband counterpart, because of the higher aliasing errors inherent in the critical sampling adaptive systems.

The simulated performances of the NAM-GSC beamformer and the NASB-ANC are summarized in Table 8.1. The listed NR factors were obtained when there were one desired signal, two interfering signals and background noises. The NAM-GSC used 16 taps per element in each subband adaptive beamformer. The NASB-ANC used 16 taps per element in each subband fixed beamformer and two 32-tap auxiliary channels in each subband ANC. When there is no location error for the desired signal, the NAM-GSC beamformer can achieve a higher noise reduction factor than the NASB-ANC with a perfect VAD. But the NASB-ANC has better robustness against location errors. It obtains a higher NR factor than the NAM-GSC beamformer when the desired signal is off the focal point by 0.15 meters in distance and 1° in the azimuth angle. Using the NLMS algorithm with step size $\mu = 0.01$, the NASB-ANC converges much slower than the NAM-GSC beamformer. On the other hand, the NASB-ANC has a much lower residual error (-24 dB) than that of the NAM-GSC (-12 dB). The de-reverberation performances of the two schemes are very close, with similar de-reverberation gain and desired signal cancellation rate. The NASB-ANC achieves better performances than the NAM-GSC at the cost of the added ANC structure, the assistance of the VAD and slightly higher computational complexity.

The performances of the NAM-GSC beamformer and the NASB-ANC have been verified by experiments made in an anechoic chamber and a reverberant conference room. The results are listed in Table 8.2. The NAM-GSC used in the experiments employed robust GSC beamformers with 32 taps per element — more than those used

Table 8.1: Performances of the NAM-GSC and the NASB-ANC via simulation

	NAM-GSC	NASB-ANC
NR without Location Errors	33.3 dB	32.3 dB
NR with Location Errors	24.5 dB	31.8 dB
Convergence Speed with $\mu = 0.01$	$k = 0.2 \times 10^5$	$k = 2 \times 10^5$
Residual Error after Convergence	-12 dB	-24 dB
De-reverberation Gain	3.4 dB	3.6 dB
Desired Signal Cancellation Rate	0.2 dB	< 0.1 dB

in the simulation. The NASB-ANC used a simple power estimation VAD. All signals were real speech and audio. The experimental results are pretty close to the simulated ones. The PAMS test LQ scores suggest that the NAM-GSC and the NASB-ANC can achieve much better speech qualities in noisy and reverberant environments.

Table 8.2: Performances of the NAM-GSC and the NASB-ANC via experimental evaluation

	NAM-GSC	NASB-ANC
Noise Reduction	27.9 dB	26.4 dB
De-reverberation Gain	3.2 dB	3.5 dB
PAMS (LQ) for NR Input	1.0	1.0
PAMS (LQ) for NR Output	3.6	3.5
PAMS (LQ) for Reverberant Input	2.4	2.4
PAMS (LQ) for De-reverberation	3.1	3.1

The performances of the three STS beamformers may be further enhanced by improving the performances of the low frequency band subarray. With limited array size and system complexity, the low band subarray may be designed using some

special techniques, such as the super-directive beamformer method [19] or the near field array optimization with quarter-wavelength spacing [79].

Beside the proposed STS system, several new algorithms have also been developed in the thesis. A simplified implementation is developed for GSC adaptive beamformers to reduce the computational load by 80%. A robust near field GSC beamforming design method is also developed to improve the robustness of the near field adaptive beamformer against location errors, by constraining a small spatial region around the focal point as well as large number of frequencies in the passband. A near field Spatial Affine Projection (SAP) algorithm is proposed for adaptive beamformers to suppress the coherent interference and combat desired signal cancellation, by utilizing the near field robust beamforming technique.

Bibliography

- [1] Abhayapala, T. D.; R. A. Kennedy, and R. C. Williamson, "Noise modeling for nearfield array optimization." *IEEE Signal Process. Letters*, vol.6, no.8, pp. 210–212, Aug. 1999.
- [2] Abhayapala, T. D.; R. A. Kennedy, and R. C. Williamson, "Spatial aliasing for near-field sensor arrays." *Electronics Letters*, vol.35, no.10, pp. 764–765, 13 May 1999.
- [3] Allen, J. B.; and D. A. Berkley, "Image method for efficiently simulating small room acoustics." *J. Acoust. Soc. Amer.*, vol.65, no.4, pp. 943–950, Apr. 1979.
- [4] Benesty, J.; P. Duhamel, and Y. Grenier, "A multichannel affine projection algorithm with applications to multichannel acoustic echo cancellation." *IEEE Signal Process. letters*, vol.3, no.2, pp. 35–37, Feb. 1996.
- [5] Berger, M. F.; and H. F. Silverman, "Microphone array optimization by stochastic region contraction." *IEEE Trans. Signal Process.* , vol.39, no.11, pp. 2377–2386, Nov. 1991.
- [6] Besbes, H.; Y.B. Jamaa, and M. Jaidane, "Exact convergence analysis of affine projection algorithm: the finite alphabet inputs case." *IEEE ICASSP*, 1999, pp. 1669–1672.
- [7] Brandstein, M.; and D. Ward (ed.), *Microphone arrays: signal processing techniques and applications*, Springer, 2000.

- [8] Buckley, K. M.; "Spatial/spectral filtering with linearly-constrained minimum variance beamformers." *IEEE Trans. Acoust., Speech, Signal Processing*, vol.ASSP-35, no.3, pp. 249–266, Mar. 1987.
- [9] Bucci, O. M.; G. D'Elia, and M.D. Migliore, "Near-field far-field transformation in time domain from optimal plane-polar samples." *IEEE Trans. Antennas and Propagation*, vol.46, no.7 , pp. 1084–1088, July 1998.
- [10] Champagne, B.; S. Bedard, and A. Stephenne, "Performance of time-delay estimation in the presence of room reverberation." *IEEE Trans. Speech and Audio Processing*, vol.4, no.2 , pp. 148–156, 1996.
- [11] Chou, T.; "Frequency-independent beamformer with low response error." *Proc. IEEE ICASSP-95*, pp. 2995–2998, May 1995.
- [12] Cioffi, J. M.; and T. Kailath, "Fast, recursive-least-squares transversal filters for adaptive filtering ." *IEEE Trans. Acoust., Speech, Signal Processing*, vol.ASSP-32, no.2, pp. 304–337, April 1984.
- [13] Compton, R. T. Jr.; "The bandwidth performance of a two-element adaptive array with tapped delay-line processing," *IEEE Transactions Antennas and Propagation*, vol. 36 no. 1 . pp. 5–14, Jan. 1988.
- [14] Coulson, C. A.; and A. Jeffrey, *Waves: A Mathematical Approach to the Common Types of Wave Motion*. Longman, London, U.K. 1977.
- [15] Crochiere, R. E.; and L. R. Rabiner, *Multirate Digital Signal Processing*, Englewood Cliffs, New Jersey 07632, Prentice-Hall Inc., 1983.
- [16] DeGroat, R. D.; D. Begusic, E. M. Dowling, and D. A. Linbarger, "Spherical subspace and eigen based affine projection algorithms." *IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, May 1997, pp. 2345–2348.
- [17] Doles, J. H., III; and F. D. Benedict, "Broadband array design using the asymptotic theory of unequally spaced arrays." *IEEE Trans. Antennas Propag.*, vol.36, no.1, pp. 27–33, Jan. 1988.

- [18] Douglas, S.; "Efficient approximate implementations of the fast affine projection algorithm using orthogonal transforms." *IEEE ICASSP*, 1996, pp. 1656–1659.
- [19] Elko, G. M.; "Superdirectional microphone arrays." in *Acoustic Signal processing for Telecommunications*, chapter 10, Kluwer Academic Publishers, Boston, 2000.
- [20] Er, M. H.; and A. Cantoni, "Derivative constraints for broad-band element space antenna array processors." *IEEE Trans. Acoust., Speech, Signal Process.*, vol.ASSP-31, no.12, pp. 1378–1393, Dec. 1983.
- [21] Er, M. H.; and A. Cantoni, "An alternative formulation for an optimum beamformer with robustness capability." *IEE Proc. part F: Comm. Radar Signal Process.*, vol.132, no.6, pp. 447–460, Oct. 1985.
- [22] Flanagan, J. L.; "Analog measurements of sound radiation from the mouth." *J. Acoust. Soc. Amer.*, vol.32, no.12, pp. 1613–1620, Dec. 1960.
- [23] Flanagan, J. L.; D. A. Berkley, G. W. Elko, J. E. West, and M. M. Sondt, "Autodirective microphone systems." *Acoustica*, vol.73, pp. 58–71, 1991.
- [24] Gay, S. L.; and S. Tavadia, "The fast affine projection algorithm." *IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Detroit, MI, May 1995, pp. 3023–3026.
- [25] Gazor, S.; S. Affes, and Y. Grenier, "Robust adaptive beamforming via target tracking." *IEEE Trans. Signal Process.*, vol.44, no.6, pp. 1589–1593, Jun. 1995.
- [26] Gazor, S.; and Y. Grenier, "Criteria for positioning of sensors for a microphone array." *IEEE Trans. Speech and Audio Process.*, vol.3, no.4, pp. 294–303, July. 1996.
- [27] Ghavami, M.; and R. Kohno, "Broadband beamforming using fan filter structures with low number of antenna elements." *IEEE Int. Conf. Personal Wireless Communication*, 1999, pp. 66–70.

- [28] Gilloire, A.; and M. Vetterli, "Adaptive filtering in subbands with critical sampling: analysis, experiments, and application to acoustic echo cancellation." *IEEE Trans. Signal Process.*, vol.40, no.8, pp. 1862–1875, Aug. 1992.
- [29] Godara, L.C.; "Application of the fast Fourier transform to broadband beamforming." *J. Acoust. Soc. Amer.*, vol.98, no.1, pp. 230–240, July 1995.
- [30] Godara, L. C.; and M. R. S. Jahromi, "Limitations and capabilities of frequency domain broadband constrained beamforming schemes." *IEEE Trans. Signal Processing*, vol.47, no.9, pp. 2386–2395, Sep. 1999.
- [31] Goodwin, M. M.; and G. W. Elko, "Constant beamwidth beamforming." *IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, 1993, pp. I-169–I-172.
- [32] Goulding, M. M.; and J. S. Bird, "Speech enhancement for mobile telephony." *IEEE Trans. Vehicular Technology*, vol.39, no.4, pp. 317–326, Nov. 1990.
- [33] Hacker, P. S.; and H. E. Schrank, "Range distance requirements for measuring low and ultralow sidelobe antenna patterns." *IEEE Trans. Antennas Propagat.*, vol.AP-30, pp. 956–965, Sept. 1982.
- [34] Hansen, R. C.; "Measurement distance effects on low sidelobe patterns." *IEEE Trans. Antennas Propagat.*, vol.AP-32, pp. 591–594, June. 1984.
- [35] Haykin, S.; *Adaptive Filter Theory*, 3rd ed., Upper Saddle River, New Jersey 07458, Prentice Hall, 1996.
- [36] Harris, C. M. edited; *Handbook of acoustical measurements and noise control*, 3rd ed., New York, McGraw-Hill, 1991.
- [37] Hoffman, M. W.; and K. M. Buckley, "Robust time-domain processing of broadband microphone array data." *IEEE Trans. Speech and Audio Processing*, vol.3, no.3, pp. 193–203, May 1995.
- [38] Hoshuyama, O.; A. Sugiyama and A. Hirano, "A robust adaptive beamformer for microphone arrays with a blocking matrix using constrained adaptive filters." *IEEE Trans. Signal Process.*, vol.47, no.10, pp. 2677–2684, Oct. 1999.

- [39] Ishimaru, A.; "Theory of unequally-spaced arrays." *IRE Trans. Antennas and Propagation*, vol. AP-8, no.11, pp. 691–702, Nov. 1962.
- [40] Jeffery, Allen; *Handbook of Mathematical Formulas and Integrals*, 2nd Edition, San Diego, CA 92101, Academic Press, 1995.
- [41] Johnson, D. H.; and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques*, Englewood Cliffs, New Jersey 07632, Prentice Hall, 1993.
- [42] Jot, J.-M. "An analysis/synthesis approach to real-time artificial reverberation." *IEEE International Conference on Acoustics, Speech, and Signal Processing*, ICASSP-92., vol.2, pp. 221 -224, 1992.
- [43] Kelly, E. L., Jr.; and M. L. Levin, "Signal parameter estimation for seismometer arrays." *Mass. Inst. Tech., Lincoln Lab., Tech. Report 339*, 1964.
- [44] Kellermann, W.; "Analysis and design of multirate systems for cancellation of acoustical echoes." *IEEE ICASSP*, New York, NY, 1988, pp. 2570–2573.
- [45] Kennedy, R. A.; T. D. Abhayapala, and D. B. Ward, "Broadband nearfield beamforming using a radial beampattern transformation." *IEEE Trans. Signal Proc.*, vol.46 no.8, pp. 2147–2156, Aug. 1998.
- [46] Kennedy, R. A.; D. B. Ward, and P. T. D. Abhayapala, "Nearfield beamforming using radial reciprocity." *IEEE Trans. Signal Process.* , vol.47, no.1, pp. 33–40, 1999.
- [47] Khalil, F.; J. P. Jullien, and A. Gilloire, "Microphone array for sound pickup in teleconference systems." *J. Acoust. Soc. Amer.*, vol.42, no.9, pp. 691–700, Sep. 1994.
- [48] Khalab, J. M.; and M. K. Ibrahim, "Novel multirate adaptive beamforming technique," *Electronics Letters*, vol. 30, no. 15 , pp. 1194–1195, 21 July 1994.
- [49] Khalab, J. M.; and M. K. Ibrahim, "Efficient multirate adaptive beamforming technique," *Electronics Letters*, vol. 30, no. 25 , pp. 2102–2103, 8th Dec. 1994.
- [50] Kuttruff, H.; *Room Acoustics*, 3rd ed., Essex, England, Elsevier Applied Science, 1991.

- [51] Laakso, T. I.; V. Välimäki, M. Karjalainen, and U. K. Laine, "Splitting the unit delay." *IEEE Signal Process. Magazine*, pp. 30–60, Jan. 1996.
- [52] Li, Z.; and M. W. Hoffman, "Evaluation of microphone arrays for enhancing noisy and reverberant speech for coding." *IEEE Trans. Speech and Audio Processing*, vol.7, no.1 , pp. 91–95, 1999.
- [53] Lin, Qiguang; Ea-Ee Jan, and J. Flanagan, "Microphone arrays and speaker identification." *IEEE Trans. Speech and Audio Process.*, vol.2, no.4 , pp. 622–629, Oct. 1994.
- [54] Lin, Y-P.; and P. P. Vaidyanathan, "Periodically nonuniform sampling of band-pass signals." *IEEE Trans. Circuits and Systems II: Analog and Digital Signal Processing*, vol.45, no.3, pp. 340–351, Mar. 1998.
- [55] Liu, Q. G.; and K. C. Ho, "On the use of a modified fast affine projection algorithm in subbands for acoustic echo cancelation." *IEEE DSP Workshop* , Loen, Norway, Sep. 1996, pp. 354–357.
- [56] Mailloux, Robert J.; *Phased array antenna handbook*, Boston, Mass., Artech House. 1994.
- [57] Marro, C.; Y. Mahieux and K. U. Simmer, "Analysis of noise reduction and dereverberation techniques based on microphone arrays with postfiltering." *IEEE Trans. Speech and Audio Process.*, vol.6 no.3, pp. 240–259, May 1998.
- [58] Mermelstein, P.; "G.722, a new CCITT coding standard for digital transmission of wideband audio signals," *IEEE Communications Magazine*, vol. 26, no. 1, pp. 8–15, Jan. 1988.
- [59] Morris, J.C.; and E. Hands, "Constant-beamwidth arrays for wide frequency bands." *Acoustica*, vol.11, pp. 341–347, 1961.
- [60] Nishikawa, K.; T. Yamamoto, K. Oto, and T. Kanamori, "Wideband beam-forming using fan filter." *Proc. IEEE ISCAS*, 1992, pp. 533–536.

- [61] Nordebo, S.; I. Claesson, and S. Nordholm, "Weighted Chebyshev approximation for the design of broadband beamformers using quadratic programming." *IEEE Signal Processing Lett.*, vol.1, no.7, pp. 103–105, July 1994.
- [62] Nordebo, S.; I. Claesson, and S. Nordholm, "Adaptive beamforming: spatial filter designed blocking matrix." *IEEE J. Oceanic Engineering*, vol.19, no.4, pp. 583–590, Oct. 1994.
- [63] Nordebo, S.; and I. Claesson, "Minimum norm design of two-dimensional weighted Chebyshev FIR filters." *IEEE Trans. Circuits and Systems II: Analog and Digital Signal Processing*, vol.44, no.3, pp. 251–253, 1997.
- [64] Nordholm, S.E.; V. Rehbock, K.L. Tee, and S. Nordebo, "Chebyshev optimization for the design of broadband beamformers in the near field." *IEEE Trans. Circuits and Systems II: Analog and Digital Signal Processing*, vol.45, no.1 , pp. 141–143, Jan. 1998.
- [65] Nomura, H.; Y. Kaneda, and J. Kojima, "Optimum gains of a delay-and-sum microphone array for near sound field." *J. Acoust. Soc. Amer.*, vol.100, p.2697, Oct. 1996.
- [66] Oh, S.; D. Linebarger, B. Priest and B. Raghothaman, "A fast affine projection algorithm for an acoustic echo canceler using a fixed-point DSP processor." *IEEE ICASSP-97*, pp. 4121–4124, 1997.
- [67] Oppenheim, A. V.; A. S. Willsky, with S. H. Nawab, *Signals and Systems*. 2nd Edition, Upper Saddle River, N.J., Prentice-Hall. 1996.
- [68] Oppenheim, A. V.; R. W. Schaffer, with J. R. Buch, *Discrete-Time Signal Processing*, 2nd Edition, Upper Saddle River, N.J., Prentice-Hall. 1998.
- [69] Ozeki, K.; and T. Umeda, "An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties." *Electron. Comm. Jpn.*, vol.67-A, no.5, pp. 19–27, 1984.

- [70] Pei, S. C.; C. C. Yeh and S. C. Chiu, "Modified spatial smoothing coherent jammer suppression without signal cancellation." *IEEE Trans. Speech and Audio Processing*, vol.36, no.3, pp. 412–414, March 1988.
- [71] Petraglia, M.R.; R. G. Alves, and P. S.R. Diniz, "New structures for adaptive filtering in subbands with critical sampling." *IEEE Trans. Signal Process.* , vol.48, no.12, pp. 3316–3327, Dec. 2000.
- [72] Pirz, F.; "Design of wideband, constant beamwidth, array microphone for use in the near field." *Bell Syst. Tech. J.*, vol.58, no.8, pp. 1839–1850, Oct. 1979.
- [73] Proakis, John G.; and D. G. Manolakis, *Digital signal processing; principles, algorithms, and applications*. 3rd ed. Upper Saddle River, N.J., Prentice Hall. 1996.
- [74] Qian, F.; and B.D. Van Veen, "Partially adaptive beamforming for correlated interference rejection." *IEEE Trans. Signal Process.* , vol.43, no.2, pp. 506–515, Feb. 1995.
- [75] Qian, F.; and B.D. Van Veen, "Quadratically constrained adaptive beamforming for coherent signals and interference." *IEEE Trans. Signal Process.* , vol.43, no.8, pp. 1890–1900, Aug. 1995.
- [76] Rupp, M.; "A family of adaptive filter algorithms with decorrelating properties." *IEEE Trans. Signal Process.* , vol.46, no.3, pp. 771–115, March 1998.
- [77] Ryan, J. G.; and R. A. Goubran, "Near-field beamforming for microphone arrays." *IEEE ICASSP*, 1997, pp. 363–366.
- [78] Ryan, J. G.; "Criterion for the minimum source distance at which plane-wave beamforming can be applied." *J. Acoust. Soc. Amer.*, vol.104, no.1, pp. 595–598, July 1998.
- [79] Ryan, J. G.; "Near-field beamforming using microphone array." *Ph.D. dissertation, Dept. of Systems and Computer Engineering, Carleton Univ.*, June 1999.

- [80] Ryan, J. G.; and R. A. Goubran, "Array optimization applied in the near field microphone array." *IEEE Trans. Speech and Audio Processing*, vol.8, no.2, pp. 173–176, March 2000.
- [81] Ryan, J. G.; "Optimum near-field performance of microphone arrays subject to a far-field beampattern constraint." *J. Acoust. Soc. Amer.*, vol.108, no.5, Pt.1, pp. 2248–2255, Nov. 2000.
- [82] Sekiguchi, T.; and Y. Karasawa, "Design of FIR fan filters used for beamspace adaptive array for broadband signals." *IEEE Int. Symposium on Circuits and Systems (ISCAS)*, vol.4 , pp. 2453–2456, 1997.
- [83] Shan, T. J.; and T. Kailath, "Adaptive beamforming for coherent signals and interference." *IEEE Trans. Acoust., Speech, Signal Process.*, vol.ASSP-33, pp. 527–536, June 1985.
- [84] Silverman, H.F., "Some analysis of microphone arrays for speech data acquisition." *IEEE Trans. Acoust., Speech, Signal Processing*, vol.ASSP-35, no.12, pp. 1699–1712, Dec. 1987.
- [85] Skolnik, M.I.; G. Nemhauser, and J.W. Sherman. III "Dynamic programming applied to unequally spaced arrays." *IEEE Trans. Antennas and Propagation*, vol. AP-12, no.1, pp. 35–43, Jan. 1964.
- [86] Tanaka, M.; Y. Kaneda, S. Makino, and J. Kojima, "Fast projection algorithm and its step size control." *IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Detroit, MI, May 1995, pp. 945–948.
- [87] Tanaka, M.; S. Makino, and J. Kojima, "A block exact fast affine projection algorithm." *IEEE Trans. Signal Process.* , vol. 7, no.1, pp. 79–86, 1999.
- [88] Unz, H.; "Linear array with arbitrarily distributed elements." *IRE Trans. Antennas and Propagation*, vol. AP-8, no.3, pp. 222–223, Mar. 1960.
- [89] Vaidyanathan, P. P.; "Multirate digital filters, filter banks, polyphase networks, and applications: a tutorial." *Proc. IEEE*, vol.78, no.1, pp. 56–93, Jan. 1990.

- [90] Vaidyanathan, P. P.; *Multirate systems and filter banks*, Englewood Cliffs, N.J., Prentice Hall. 1993.
- [91] Van Veen; B. D.; and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering." *IEEE ASSP Magazine*, pp. 4–24, April 1988.
- [92] Van Veen; B. D.; "Minimum variance beamforming with soft response constraints." *IEEE Trans. Signal Process.* , vol. 39, no.9, pp. 1964–1972, Sep. 1991.
- [93] Vasconcellos, R. T. B ; M. R. Petraglia, and R. G. Alves, "A new critically sampled non-uniform subband adaptive structure." *IEEE ICASSP*, 2001, pp. 3713–3716.
- [94] F. W. Vook and R. T. Compton, Jr. "Bandwidth performance of linear adaptive arrays with tapped delay-line processing," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 28, no. 3, pp. 901–908, July. 1992.
- [95] Wallace, R. B.; and R. A. Goubran, "Improved tracking adaptive noise canceler for nonstationary environments." *IEEE Trans. Signal Process.* , vol.40, no.3, pp. 700–703, Mar. 1992.
- [96] Wang, H.; and M. Keveh, "Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wideband sources." *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-33, no.4, pp. 823–831, Aug. 1985.
- [97] Ward, D. B.; R. A. Kennedy and R. C. Williamson, "Theory and design of broadband sensor arrays with frequency invariant far-field beam patterns." *J. Acoust. Soc. Amer.*, vol.97, no.2, pp. 1023–1034, Feb. 1995.
- [98] Ward, D. B.; "Technique for broadband correlated interference rejection in microphone array." *IEEE Trans. Speech and Audio Processing*, vol.6, no.4, pp. 414–417, July 1998.

- [99] Wax, M.; T. J. Shan and T. Kailath, "Spatio-temporal spectral analysis by eigenstructure methods." *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, no.4, pp. 817–827, Aug. 1984.
- [100] Widrow, B.; K.M. Duvall, R.P. Gooch and W.C. Newman, "Signal cancellation phenomena in adaptive antennas: causes and cures", *IEEE Tran. Antennas Propagat.*, vol.AP-30, no.3, pp. 469–478, May 1982.
- [101] Widrow, B.; and S.D. Stearns, *Adaptive signal processing*. Englewood Cliffs, N.J., Prentice-Hall. 1985.
- [102] Willey, R. E.; "Space tapering of linear and planar arrays." *IRE Trans. Antennas Propagat.*, vol. AP-10, no.7, pp. 369–377, July 1962.
- [103] Yang, J.-F.; and M. Keveh, "Coherent signal-subspace transformation beamformer." *IEE Proc. part F: Comm. Radar Signal Process.*, vol.137, no.4, pp. 267–275, Aug. 1990.
- [104] Yeh, C. C.; S. C. Chiu and S. C. Pei, "On the coherent interference suppression using a spatially smoothing adaptive array." *IEEE Trans. Ant and Audio Processing*, vol.37, no.7, pp. 851–857, July 1989.
- [105] Yu, S. J.; and J. H. Lee, "Efficient Eigenspace-based array signal processing using multiple shift-invariant subarrays." *IEEE Trans. Antennas and Propagation*, vol.47 no.1, pp. 186–194, Jan. 1999.
- [106] Yu, J. L.; and C. C. Yeh, "Generalized Eigenspace-based beamformers." *IEEE Trans. Signal Process.*, vol.43 no.11, pp. 2453–2461, Nov. 1995.
- [107] Ziomek, L. J.; "Three necessary conditions for the validity of the Fresnel phase approximation for the near-field beam pattern of an aperture." *IEEE J. Oceanic Engineering*, vol.18, no.4 , pp. 73–75, 1993.
- [108] Ziomek, L. J.; *Fundamentals of acoustic field theory and space-time signal processing*, Boca Raton, Florida, CRC Press, 1994.

- [109] Zheng, Y. R.; and R. A. Goubran, "Adaptive beamforming using affine projection algorithms." *IEEE ICSP-2000*, Beijing, P.R. China, Aug. 2000. vol.3, pp. 1929–1932.
- [110] Zheng, Y. R.; R. A. Goubran, and M. El-Tanany, "Coherent interference suppression with an adaptive array using spatial affine projection algorithm." *52nd IEEE Fall VTC 2000*, Boston, MA, Sep. 2000. vol.1, pp. 105–109.
- [111] Zheng, Y. R.; R. A. Goubran, and M. El-Tanany, "A broadband adaptive beamformer using nested arrays and multirate techniques." *IEEE DSP Workshop 2000*, Hill County, TX, USA, Oct. 2000. http://spib.rice.edu/SPS/SPS_prevconf.html
- [112] Zheng, Y. R.; R. A. Goubran, and M. El-Tanany, "On constraint design and implementation for broadband adaptive array beamforming." *IEEE ICASSP*, Orlando, FL, USA, May 2002. vol.3, pp. 2917–2920.
- [113] Zheng, Y. R.; R. A. Goubran, and M. El-Tanany, "Near-field adaptive beamforming using a multirate nested array." submitted to *J. Acoustic Society America*, Feb. 2002.
- [114] Zheng, Y. R.; R. A. Goubran, and M. El-Tanany, "A broadband adaptive beamformer using nested arrays and critically sampled multirate QMF banks." submitted to *IEEE Signal Process. Letters*, June, 2002.
- [115] Zheng, Y. R.; R. A. Goubran, and M. El-Tanany, "Experimental evaluation of a near field nested microphone array with adaptive noise canceler." submitted to *IEEE Trans. Instrumentation and Measurement*, June, 2002.
- [116] Zheng, Y. R.; R. A. Goubran, and M. El-Tanany, "Broadband spatial affine projection algorithm for nearfield coherent interference suppression." *IEEE Trans. Vehicular Technology*, in preparation.
- [117] *Digital Speech Level Analyser User Guide*, Malden Electronics Ltd, Revision 3.3, 1997–2000.

Appendix A

The Image Model

The image model proposed in [3] is a commonly used method for computer simulation of a room reverberation.

Consider a rectangular room with rigid walls, ceiling and floor. All the walls have constant reflection coefficients over all frequencies. A sound source is modeled as a point source [79, 22]. The reflected sound waves can be represented by the image sources illustrated in Figure A.1. The room impulse response at a sensor location can be calculated from the image source locations, room geometry and reflection coefficients.

Let $\mathbf{x}_s = (x_s, y_s, z_s)$ denote the vector of the sound source location, $\mathbf{x}_0 = (x_0, y_0, z_0)$ denote the vector of the receiver location. The room dimensions are (L_x, L_y, L_z) . The reflection coefficients of the 6 walls are β_{x1} , β_{x2} , β_{y1} , β_{y2} , β_{z1} and β_{z2} , respectively. The room impulse response is then derived from the image model as [3]

$$h(t, \mathbf{x}_0, \mathbf{x}_s) = \sum_{\mathbf{p}=\mathbf{0}}^1 \sum_{\mathbf{r}=-\inf}^{\inf} \beta_{x1}^{|n-q|} \beta_{x2}^{|n|} \beta_{y1}^{|l-j|} \beta_{y2}^{|l|} \beta_{z1}^{|m-k|} \beta_{z2}^{|m|} \frac{\delta(t - |\mathbf{R}_p - \mathbf{R}_r|/c)}{|\mathbf{R}_p - \mathbf{R}_r|} \quad (\text{A.1})$$

where

$$\mathbf{p} = (q, j, k),$$

$$\mathbf{r} = (n, l, m),$$

$$\mathbf{R}_p = (x_0 - x_s + 2qx_s, y_0 - y_s + 2jy_s, z_0 - z_s + 2kz_s),$$

Figure A.1: Image model

$$\mathbf{R}_r = (2nL_x, 2lL_y, 2mL_z).$$

The sum $\sum$ with vector index $\mathbf{p}$ is used to indicate three sums, namely one for each of the three components of $\mathbf{p} = (q, j, k)$. The sum with index $\mathbf{r} = (n, l, m)$ is a similar sum. There are eight points in a three-dimensional lattice of points for $\mathbf{p}$ and for $\mathbf{r}$, the lattice is infinite.

Theoretically, the total number of images is infinite; but practically, only those with significant strength are included. Higher order images are attenuated more by the walls and are located farther away from the sensor, thus they contribute much less power. Table A.1 gives the number of image sources corresponding to the low order reflections.

Table A.1: The number of low order image sources in a rectangular room

Order of Images	Number of Images	Total Number of Images
1	6	6
2	18	24
3	38	62
4	66	128
5	102	230

Assume the simulated room has a size of $(L_x, L_y, L_z) = (5.0m, 4.0m, 3.0m)$. The reflection coefficients of the walls are 0.9, and those of the ceiling and floor are 0.7. The signal source is located at $\mathbf{x}_s = (1.5m, 1.5m, 1.0m)$. An omni-directional microphone receiver is at the point $\mathbf{o}' = (1.0m, 1.0m, 1.0m)$. The impulse response observed at the receiver is composed of 54,000 images for the time window of $t = [0, 0.3]$ second. Figure A.2 shows the impulse response between the sound source $\mathbf{x}_s$ and the receiver $\mathbf{o}'$ for frequency band $[50, 7000]$ Hz. The frequency characteristics of the impulse response are shown in Figure A.3.

Figure A.2: Impulse response of a reverberant room.

Figure A.3: Frequency response of a reverberant room.

The reverberation time, denoted T_{60} , is defined as the time required for the steady-state sound intensity in a room to decay by 60dB after the source is removed. It can be easily estimated from the room impulse response. The reverberation time of the simulated room is approximately $T_{60} = 250\text{ms}$.

Energy Decay Curve (EDC) is a graphic plot of the energy decay as a function of time t . The energy decay of an impulse response at a given time instant t is defined as the total remaining energy of the impulse response after time t [50, pp.116–117]. It is calculated by adding the energy of the impulse response tail from t to infinity. The longer the reverberation time, the slower the energy decays. Figure A.4 shows the energy decay curve of the simulated room. The energy contribution of the individual low order images can be seen in the EDC. The high order image sources correspond to the tail of the EDC. They can be treated as far field spherically isotropic noises [77] or near field spherically isotropic noises [1].

Figure A.4: Energy decay curve of the room impulse response

Appendix B

Affine Projection Algorithms

The Affine Projection (AP) algorithm is an adaptive algorithm, originally proposed by Ozeki *et al.* [69] for acoustical echo and noise cancellation. It is a generalization of the Normalized LMS (NLMS) algorithm and the windowed RLS algorithm. Its fast version (FAP) [24, 86] was developed, which reduced the computational complexity from $(p + 1)N + O(p^3)$ to $2N + 20p$, with N being the length of the adaptive filter. Properties and variations of the FAP algorithm have also been investigated extensively in recent years.

Consider an adaptive filter shown in Figure B.1. The input signal at time instant k is an $N \times 1$ vector $\mathbf{x}_N(k)$

$$\mathbf{x}_N(k) = [x(k), x(k-1), \dots, x(k-N+1)]^T$$

The subscript of a vector is used to indicate its dimension.

Figure B.1: General structure of an adaptive filter

Table B.1: Summary of the AP algorithm

1. $\mathbf{X}(k) = [\mathbf{x}_N(k) \quad \mathbf{x}_N(k-1) \quad \cdots \quad \mathbf{x}_N(k-p+1)]$	$N \times p$
2. $\mathbf{D}(k) = [d(k), d(k-1), \dots, d(k-p+1)]^H$	$p \times 1$
3. $\mathbf{R}(k) = \mathbf{X}^H(k) \cdot \mathbf{X}(k) + \delta \mathbf{I}$	$p \times p$
4. $\mathbf{e}(k) = \mathbf{D}(k) - \mathbf{X}^H(k) \cdot \hat{\mathbf{w}}(k)$	$p \times 1$
5. $\hat{\mathbf{w}}(k+1) = \hat{\mathbf{w}}(k) + \mu \mathbf{X}(k) \cdot \mathbf{R}^{-1}(k) \cdot \mathbf{e}(k)$	$N \times 1$

The AP algorithm is formulated in Table B.1, where $d(k)$ is the desired signal, $\hat{\mathbf{w}}(k)$ is the vector of filter coefficients, δ is a regulation parameter, μ is the step size and p is the projection order.

Choosing $p = 1$ produces the NLMS algorithm, while $p = N$ yields the windowed RLS algorithm. For $1 < p < N$, the AP algorithm provides a range of compromise solutions of medium fast convergence and low computational complexity.

The computational complexity for the AP algorithm is $(p+1)N + O(p^3)$. The Fast AP algorithm (FAP) brings the complexity down to $2N + 20p$, compared to $2N + 1$ for the NLMS and $8N$ for the fast RLS algorithm. The FAP algorithm does not explicitly calculate the weight vector $\hat{\mathbf{w}}(k)$ but $y(k)$. Matrix inversion $\mathbf{R}^{-1}(k)$ is computed by sliding window RLS algorithm (FTF) [12], but only for order $p \times p$. The FAP algorithm is formulated in Table B.2.

A proper selection of p provides the trade-off between the convergence rate and the computational complexity. Figure B.2 shows the convergence rate of the FAP algorithm with different projection order p . The algorithm is applied to an adaptive echo canceler with speech inputs. The theoretical analysis on convergence is proven to be difficult [6]. The general observation is that the projection order $p = 2$ gives the greatest gain in the convergence rate improvement. When the projection order is close to the “degree of correlation” of the input signals, the convergence is close to the RLS algorithm.

Table B.2: Summary of the FAP algorithm

0. Initialization
$\mathbf{r}_{p-1}(0) = [\mathbf{x}_N^H(0)\mathbf{x}_N(-1), \dots, \mathbf{x}_N^H(0)\mathbf{x}_N(-p+1)]^H$ $\mathbf{e}_p(0) = 0, F(0) = B(0) = \delta,$ $\mathbf{f}_{p-1}(0) = \mathbf{s}_{p-1}(0) = 0,$ $\mathbf{z}_N(0) \text{ is arbitrary}$
Start with $k = 1$
1. $\mathbf{r}_{p-1}(k) = \mathbf{r}_{p-1}(k-1) + x(k)\mathbf{x}_{p-1}(k-1) - x(k-N)\mathbf{x}_{p-1}(k-N-1)$
2. $\hat{y}(k) = \mathbf{x}_N^H(k)\mathbf{z}_N(k) + \mathbf{r}_{p-1}^H(k)\mathbf{s}_{p-1}(k-1)$
3. $e(k) = y(k) - \hat{y}(k)$
4. $\begin{bmatrix} \mathbf{e}_p(k) \\ * \end{bmatrix} = \begin{bmatrix} e(k) \\ (1-\mu)\mathbf{e}_p(k-1) \end{bmatrix}$
5. Compute $\mathbf{a}_p(k)$, $\mathbf{b}_p(k)$, $F(k)$, and $B(k)$ by the sliding window version of FTF
6. $\mathbf{g}_p(k) = (1-\mu) \begin{bmatrix} 0 \\ \mathbf{f}_{p-1}(k-1) \end{bmatrix} + \frac{\mathbf{a}_p^H(k)\mathbf{e}_p(k)}{F(k)}\mathbf{a}_p(k)$
7. $\begin{bmatrix} \mathbf{f}_{p-1}(k) \\ 0 \end{bmatrix} = \mathbf{g}_p(k) - \frac{\mathbf{b}_p^H(k)\mathbf{e}_p(k)}{B(k)}\mathbf{b}_p(k)$
8. $\begin{bmatrix} \mathbf{s}_{p-1}(k) \\ s(k) \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{s}_{p-1}(k-1) \end{bmatrix} + \mu\mathbf{g}_p(k)$
9. $\mathbf{z}_N(k+1) = \mathbf{z}_N(k) + s(k)\mathbf{x}_N(k-p+1)$

Table B.3: The simplified FAP algorithm

5. $\mathbf{R}_{pp}(k) = \mathbf{X}^H(k)\mathbf{X}(k)$
6. Compute the inverse of $\mathbf{R}_{pp}(k)$ using Direct Matrix Inversion (DMI);
7. $\mathbf{g}_p(k) = \mathbf{R}_{pp}^{-1}(k)\mathbf{e}_p(k)$ |
|--|

Figure B.2: Convergence of the FAP algorithm

The drawbacks of the FAP algorithm are the implementation difficulty and numerical instability of its embedded FTF (Fast Transversal Filter) algorithm. Hence for a small p , the simplified FAP algorithm using the Direct Matrix Inversion (DMI) [66] is very attractive. The simplified FAP algorithm replaces Line 5. to Line 7. in Table B.2 by the equations listed in Table B.3. Other variations of the FAP algorithm include the modified FAP using Discrete Cosine Transform (DCT) [18], the eigen based FAP [16] and the modified FAP using Matrix Inversion Lemma [55], etc. All of them try to replace the embedded FTF by other methods.

The block exact FAP algorithm was also developed by Tanaka *et al.* [87], using FFT techniques to achieve an exact convergence rate as the sample-by-sample FAP algorithm. The multichannel FAP algorithm was proposed by Benesty [4], which projects twice to decorrelate the cross correlation between multiple channels.

Another drawback of the AP algorithm is its noise amplification. It can be best explained by formulating the APA in an alternative form [76], as shown in Table B.4.

Table B.4: Alternate formulation of the AP algorithm

1. $\mathbf{X}(k) = [\mathbf{x}_N(k-1), \dots, \mathbf{x}_N(k-p+1)]$	$N \times (p-1)$
2. $\hat{\mathbf{a}}(k) = [\mathbf{X}^H(k)\mathbf{X}(k)]^{-1} \mathbf{X}^H(k)\mathbf{x}_N(k)$	$N \times 1$
3. $\Phi(k) = \mathbf{x}_N(k) - \mathbf{X}(k)\hat{\mathbf{a}}(k)$	$N \times 1$
4. $e(k) = d^H(k) - \mathbf{x}_N^H(k)\hat{\mathbf{w}}(k)$	scalar
5. $\hat{\mathbf{w}}(k+1) = \hat{\mathbf{w}}(k) + \mu \frac{\Phi(k)}{\Phi^H(k)\Phi(k)} e(k)$	$N \times 1$

The AP algorithm first projects the input vector $\mathbf{x}_N(k)$, to obtain the decorrelated direction vector $\Phi(k)$, then performs the NLMS adaptation in the direction of $\Phi(k)$, as depicted in Figure B.3. The decorrelated direction vector is orthogonal to the past $p-1$ input vectors, thus allowing the AP algorithm to converge fast. Meanwhile, the background noise (assumed to be white) is filtered through a filter with coefficients $\hat{\mathbf{a}}(k)$ close to the optimal solution $\mathbf{a}$. The filtered noise has a variance that is enlarged by $1 + \mathbf{a}^H \mathbf{a}$. As a rule of thumb, the projection order is chosen to be less than 10 to avoid the noise amplification effect.

Figure B.3: Decorrelation property of the AP algorithm

Appendix C

List of Publications

Conference Proceedings

1. Y. R. Zheng, R. A. Goubran, and M. El-Tanany, "On constraint design and implementation for broadband adaptive array beamforming." *IEEE ICASSP*, Orlando, FL, USA, May 2002. vol.3, pp. 2917–2920.
2. Y. R. Zheng, R. A. Goubran, and M. El-Tanany, "Coherent interference suppression with an adaptive array using spatial affine projection algorithm," *52nd IEEE Fall VTC 2000*, Boston, MA, Sep. 2000. vol.1, pp. 105–109.
3. Y. R. Zheng, R. A. Goubran, and M. El-Tanany, "A broadband adaptive beamformer using nested arrays and multirate techniques." *IEEE DSP Workshop 2000*, Hill County, TX, USA, Oct. 2000.
<http://spib.rice.edu/SPS/SPS-prevconf.html>
4. Y. R. Zheng and R. A. Goubran, "Adaptive Beamforming using Affine Projection Algorithms," *IEEE ICSP-2000*, Beijing, P.R. China, Aug. 2000. vol.3, pp. 1929–1932.

Journal Papers

1. Y. R. Zheng, R. A. Goubran, and M. El-Tanany, "Near-field adaptive beam-forming using a multirate nested array." submitted to *J. Acoustic Society America*, Feb. 2002.
2. Y. R. Zheng, R. A. Goubran, and M. El-Tanany, "A broadband adaptive beam-former using nested arrays and critically sampled multirate QMF banks." submitted to *IEEE Signal Process. Letters*, June, 2002.
3. Y. R. Zheng, R. A. Goubran, and M. El-Tanany, "Experimental evaluation of a near field nested microphone array with adaptive noise canceler." submitted to *IEEE Trans. on Instrumentation and Measurement*, June, 2002.
4. Y. R. Zheng, R. A. Goubran, and M. El-Tanany, "Broadband spatial affine projection algorithm for nearfield coherent interference suppression." *IEEE Trans. Vehicular Technology*, in preparation.