

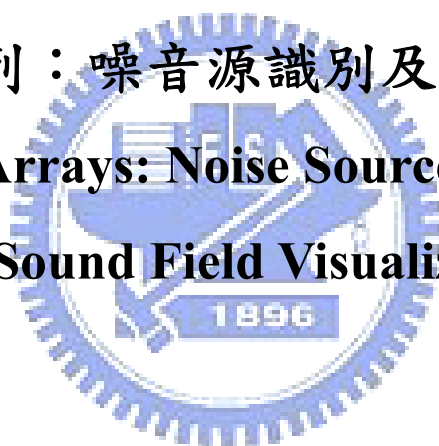
國立交通大學

機械工程學系

博士論文

麥克風陣列：噪音源識別及聲場可視化

**Microphone Arrays: Noise Source Identification
and Sound Field Visualization**



研究生：林家鴻

指導教授：鄭泗東

共同指導教授：白明憲

中華民國一百年一月

麥克風陣列：噪音源識別及聲場可視化

**Microphone Arrays: Noise Source Identification
and Sound Field Visualization**

研究生：林家鴻

Student : Jia-Hong Lin

指導教授：鄭泗東

Advisor : Stone Cheng

共同指導教授：白明憲

Co-Advisor : Mingsian R. Bai



A Thesis

Submitted to Department of Mechanical Engineering
College of Engineering
National Chiao Tung University
in Partial Fulfillment of the Requirements
for the Degree of
Doctor of Science
in
Department of Mechanical Engineering
January 2011
Hsinchu, Taiwan, Republic of China

中華民國一百年一月

麥克風陣列：噪音源識別及聲場可視化

學生：林家鴻

指導教授：鄭泗東

共同指導教授：白明憲

國立交通大學 機械工程學系 博士班

摘 要

本論文提出遠場及近場麥克風陣列來識別噪音源的位置及聲場可視化。在遠場麥克風陣列中，稀疏且隨機配置的麥克風陣列已知可以用來傳遞遠場的影像而不會產生鬼葉瓣的問題。在最佳化麥克風的配置中，全域最佳化技術包括蒙地卡羅法、模擬退火法和內部方格蒙地卡羅法被用來有效率地尋找最佳的麥克風配置。模擬結果顯示出要避免鬼葉瓣的出現，隨機配置麥克風是必要的。而結合模擬退火法和蒙地卡羅法的方法可以有效率地找到一個令人滿意的配置，這個配置能得到傑出的波束圖和相對較均勻的麥克風分布。在到達方向的估測中，平面波的聲源被視為球面波。遠場聲學影像的方法包括延遲和相加法、時間反轉法、單進多出等效聲源反逆濾波法、最小變異無失真響應法和多重信號分類法被用來估測聲源位置。結果顯示多重信號分類法在定位噪音源位置上可得到最佳的結果。在近場麥克風陣列中，提出近場等效音源影像(Nearfield Equivalence Source Imaging, NESI)來識別噪音源的位置及強度。NESI 是在時間面上設計,它除了可以應用在穩態噪音源上亦可應用在非穩定噪音源上。利用最小均方最佳化來設計出多通道反逆濾器。調整化被利用來仰制不足解模型相稱的不好條件值。設計的參數如:麥克風陣列的孔徑、麥克風的間距、焦點的間距及量測距離都會明顯地

影響到聲場影像解析度的結果。重建距離的選定是依據傳遞矩陣的條件值大小來決定。並且利用窗口矩陣的設計來成功地解決在邊界上失焦的問題。此外，本論文使用撤退焦點表面法來避免在重建聲場表面上的奇異問題。並使用黃金比例法來求得最佳的撤退距離（虛擬聲源面及重建面之間的距離）。在結果中發現，在平面活塞聲源中最佳的撤退距離為 0.4 到 0.5 倍的麥克風間距及在球面活塞聲源中最佳的撤退距離為 0.8 到 1.7 倍的麥克風間距。近場等效聲源成像法利用多通道反算濾波器可重建數個聲學變數，包含聲壓、粒子速度、主動聲強及聲功率。而當可供使用的麥克風數量不足時，可運用虛擬麥克風技巧裡的內插及外插方法分別增加解析度及減低邊緣效應。使用高效率之狀態空間最小化獲得技術來實現多通道反逆濾器。經由最佳化計算，最佳的近場麥克風配置為等間距之麥克風置配。根據模擬及實際聲源(個人電腦、空氣壓縮機、速克達及非接觸式模態測式)測試結果,本論文所提出的 NESI 技術可以有效地去識別出噪音源的位置及強度。



Microphone Arrays: Noise Source Identification and Sound Field Visualization

Student : Jia-Hong Lin

Advisor : Stone Cheng

Co-Advisor : Mingsian R. Bai

**Department of Mechanical Engineering
National Chiao Tung University**

ABSTRACT

Farfield and nearfield microphone arrays are proposed for noise source identification (NSI) and sound field visualization (SFV). In farfield array, arrays with sparse and random microphone deployment are known to be capable of delivering high quality far-field images without grating lobes. In the optimal deployment of microphone arrays, global optimization techniques including the simulated annealing (SA) algorithm and the intra-block Monte Carlo (IBMC) algorithm are exploited to find the optimal microphone positions efficiently. In direction of arrival (DOA) estimation, the planar wave sources are assumed to be spherical wave sources in this thesis. Farfield acoustic imaging algorithms including the delay and sum (DAS) algorithm, the time reversal (TR) algorithm, the single input multiple output equivalent source inverse filtering (SIMO-ESIF) algorithm, the Minimum variance distortionless response (MVDR) algorithm and the multiple signal classification (MUSIC) algorithm are employed to estimate DOA. Results show that the MUSIC algorithm can attain the highest resolution of localizing sound sources

positions. In nearfield array, a nearfield equivalence source imaging (NESI) technique is proposed to identify locations and strengths of noise sources. The NESI is based on the time-domain formulation that applies not only to stationary but also a transient noise. Multichannel inverse filters are designed using the least-square optimization. Regularization is employed to mitigate the ill-posedness inherent in the model-matching problem. Window design can also be incorporated into the inverse filters to overcome defocusing problems when the distance of reconstruction (DOR) is large or when the number of microphones is less than that of the focal points. As a basic form of the equivalent source method (ESM) applied to nearfield acoustical holography (NAH) problems, discrete monopoles are utilized to represent the sound field of interest. When setting up the virtual source distribution, it is vital to maintain a “retreat distance” (RD) between the virtual sources and the actual source surface such that reconstruction would not suffer from singularity problems. However, one cannot increase the distance without bound because of the ill-posedness inherent in the reconstruction process with large distance. How to reach the best compromise between the reconstruction errors induced by the point source singularity and the reconstruction ill-posedness is an interesting problem in its own right. This thesis revisits this issue, with the aid of an optimization algorithm based on the golden section search (GSS) and parabolic interpolation. The results revealed that the RD appropriate for the ESM ranged from 0.4 to 0.5 times the spacing for the planar piston, while from 0.8 to 1.7 times average spacing for the spherical piston. Acoustical variable including sound pressure, particle velocity, active intensity and sound power are calculated by using multichannel regularized inverse filters. In practical applications in which only patch array with scarce sensors are available, a virtual microphone approach is employed to ameliorate edge effects using extrapolation and to improve imaging resolution using interpolation. The multichannel inverse filters

are implemented in light of a highly efficient state-space minimal realization technique. A special kind of beam pattern and cost function definition is used for the multiple-input-multiple-output (MIMO) imaging problem. A striking result was also obtained that random deployment presents no particular benefit in nearfield imaging and the optimal configuration is the uniform array. As indicated by the simulation and experiment results, the proposed technique proved effective in identifying sources of many kinds, including broadband, narrowband, stationary, and transient sources.



誌謝

時光飛逝，漫長的七年博士班研究生涯終於走到尾聲了。首先感謝指導教授白明憲博士及鄭泗東博士的諄諄指導與教誨，使我順利完成學業與論文，在此對白老師及鄭老師的感激之情溢於言表。而老師指導學生時豐富的專業知識，嚴謹的治學態度以及待人處事方面，亦是身為學生的我學習與景仰的典範。更感謝白老師帶領我參與兩次國際研討會，增廣我的國際視野，這絕對是我求學生涯最難忘的回憶。

在論文寫作上，感謝中正大學鄭志鈞教授、清華大學方維倫教授、本校胡竹生教授和陳宗麟教授在百忙中撥冗閱讀並提出寶貴的意見，使得本文的內容更趨完善與充實，在此本人致上無限的感激。

回顧這七年的日子，承蒙同實驗室的博士班曾平順學長、歐昆應學長、陳榮亮學長、嚴坤龍學長、李志中同學、蔡耀坤學弟和李雨蓉學妹在研究與學業上的互相切磋討論，每在烏雲蔽空時，得以撥雲見日，獲益甚多。特別感謝與我同研究團隊的學弟，陳勁誠、劉冠良和曾智文。感謝你們在研究與計畫上盡心盡力的幫忙，是我得以順利完成研究的重大助力。還有實驗室學弟，王俊凱、徐偉智、吳俊慶、許書豪、衛帝安、羅一揚、徐和生、趙婉芝、楊浩廷、阮星瑋以及 92 級到 99 級學弟妹在生活上的朝夕相處與砥礪磨練，使我愉快度過這漫長的研究生涯。除此之外，我也要感謝三陽工業及所有在精神上給我鼓舞支持的人，謝謝各位的幫忙與鼓勵，在此也一併致上我最深的謝意。

最後僅以此篇論文，獻給我的家人，感謝你們的勉勵與無私的照顧，使我可以安心於學業與工作之上。今天我能順利取得博士學位，要感謝的人很多，上述名單恐有疏漏，在此也一致上我最深的謝意。

TABLE OF CONTENTS

	page
摘要	I
ABSTRACT	III
誌謝	VI
TABLE OF CONTENTS	VII
LIST OF TABLES	XI
LIST OF FIGURES	XIII
CHAPTER 1. INTRODUCTION	1
1.1 Background and motivation: problem statement	1
1.2 Review of prior arts: approaches for noise identification problems	4
1.3 Organization of the thesis	7
CHAPTER 2. THEORETICAL PRELIMINARIES OF ACOUSTICS	10
2.1 Fundamentals of acoustics	10
2.2 Sound field representation using basis function expansion	18
2.3 Sound field representation using Helmholtz integral equation	23
2.4 Inverse problems and ill-posedness	42
CHAPTER 3. THEORETICAL PRELIMINARIES OF SIGNAL PROCESSING	
.....	44
3.1 Linear algebra basics	44
3.2 Digital signal processing basics	47
3.3 Array signal processing basics	69
3.4 Optimization algorithms	75
CHAPTER 4. FARFIELD ARRAY SIGNAL PROCESSING ALGORITHMS ..	86
4.1 Low-resolution algorithms	86

4.1.1 Delay and sum beamformer	86
4.1.2 Time reversal beamformer	91
4.1.3 SIMO-ESIF algorithm	94
4.1.4 Optimal array: cost functions, Rayleigh's quotient.....	97
4.1.5 Choice of farfield array parameters	121
4.2 High-resolution algorithms	123
4.2.1 Minimum variance distortionless response (MVDR).....	123
4.2.2 Multiple signal classification (MUSIC).....	125
4.2.3 Choice of parameters: Akaike information criterion (AIC).....	127
4.3 Comparison of the farfield algorithms	130
CHAPTER 5. NEARFIELD ARRAY SIGNAL PROCESSING ALGORITHMS	
.....	146
5.1 Fourier NAH.....	146
5.2 BEM-based NAH (IBEM): direct and indirect formulations	151
5.2.1 Direct IBEM Formulation.....	151
5.2.2 Indirect IBEM Formulation.....	161
5.3 Equivalent source method (ESM).....	163
5.3.1 Direct ESM	165
5.3.2 Indirect ESM.....	168
5.3.3 Nearfield Equivalent Source Imaging (NESI).....	173
5.3.4 Kalman filter-based algorithm	177
5.3.5 Choice of nearfield array parameters	184
5.4 Comparison of the nearfield algorithms	188
CHAPTER 6. PRACTICAL IMPLEMENTATIONS.....	191
6.1 Inverse filter design.....	191
6.1.1 Model matching: ill-posedness and regularization	191

6.1.2 Window design.....	193
6.1.3 Parameter choice methods (PCM)	196
6.2 Multi-channel fast filtering	200
6.2.1 The time-domain processing.....	203
6.2.2 The frequency-domain processing.....	203
6.2.3 Comparison of filtering approaches.....	207
6.3 Post-processing	210
6.3.1 Acoustic variables: p, u, I, W	210
6.3.2 Miscellaneous processing items.....	212
6.4 Choice of distance of reconstruction and lattice spacing.....	219
6.5 Virtual microphone technique: field interpolation and extrapolation	220
6.6 Choice of retreat distance (RD).....	224
6.6.1 Integral approximation error vs. reconstruction ill-posedness.....	224
6.6.2 Determination of RD: Golden Section Search (GSS).....	225
6.7 Optimization of sensor deployment: uniform vs. random array	242
6.7.1 Optimal nearfield array: cost functions.....	242
6.7.2 Optimizing farfield sensor deployment	247
6.7.3 Optimizing nearfield sensor deployment.....	263
6.8 System integration and experimental arrangement.....	276
CHAPTER 7. APPLICATION EXAMPLES	279
7.1 Scooter: transient sources	281
7.2 Compressor	290
7.3 IT equipment.....	295
7.4 Wooden box	299
7.5 Non-contact modal analysis.....	303
7.6 Other applications of the ESM.....	311

CHAPTER 8. CONCLUSIONS.....313
REFERENCES.....316
ABBREVIATIONS325
PUBLICATIONS328



LIST OF TABLES

	page
Table 1	Summary of Fourier series and transform relations.....59
Table 2	FIR filter coefficients to implement Lagrange interpolation for $N = 1$ and $N = 2$60
Table 3	Distinctions of farfield and nearfield arrays.....73
Table 4	The first-order differential microphone designs.106
Table 5	Comparisons of farfield acoustic imaging algorithms.132
Table 6	Fourier based NAH for planar, cylindrical and spherical geometries.....149
Table 7	Summary of the BEM-based acoustic holography algorithms.....158
Table 8	Comparison of nearfield imaging methods.....190
Table 9	Comparison of computational complexity in terms of OPS of three multichannels filtering methods for three array configurations. The block size of FFT $N_i = 512$. The numbers of microphones and focal points are assumed to be equal, i.e., $m = j$. The DC method is used for benchmarking (100% in parenthesis).....208
Table 10	The search performance of different optimization methods for farfield array deployment with the inter-element spacing $d = 0.6$ m. The letter “w” indicates that weight optimization is performed.252
Table 11	The comparison of converged cost function Q of the URA and the optimized farfield random arrays at three different frequencies.....253
Table 12	The search performance of different optimization methods for nearfield array deployment with the inter-element spacing $d = 0.6$ m at three different frequencies.266
Table 13	The performance obtained using different numbers of focal points in a 5×6 URA with the inter-element spacing $d = 0.6$ m at three different



LIST OF FIGURES

	page
Fig. 1 Microphone array system for NSI and SFV.....	9
Fig. 2 Control volume. (a) Eulerian coordinates, (b) Lagrangian coordinates.....	16
Fig. 3 A sine wave. (a) $x = 0$, (b) $t = 0$, (c) $t = t_1$	17
Fig. 4 Three basic coordinate systems. (a) Cartesian coordinates, (b) cylindrical coordinates, (c) spherical coordinates.....	22
Fig. 5 Boundary problem of interior sound field.....	35
Fig. 6 Huygen's principle.....	36
Fig. 7 Schematic diagram to show the idea of the method of image.....	37
Fig. 8 Boundary problem of exterior sound field.....	38
Fig. 9 Virtual source representation. (a) Monopole layer, (b) dipole layer, (c) equivalent source distribution.....	39
Fig. 10 Dipole and quadrupole. (a) Coordinate figure of dipole, (b) when \mathbf{d} is aimed at y-axis (dipole), (c) sound pressure distribution of dipole, (d) general structure of quadrupole.....	40
Fig. 11 Illustration for multipole expansion.....	41
Fig. 12 Signal flow graph of direct-form I IIR structure.....	61
Fig. 13 Signal flow graph of direct-form II IIR structure.....	62
Fig. 14 Direct-form realization of an FIR system.....	63
Fig. 15 Block convolution using the OAD. (a) input $x(n)$, (b) impulse response $h(n)$, (c) expected output $y(n)$, (d) output $y_1(n)$ for block convolution of $x_1(n)$ and $h(n)$, (e) output $y_2(n)$ for block convolution of $x_2(n)$ and $h(n)$, (f) output $y_3(n)$ for block convolution of $x_3(n)$ and $h(n)$, (g) shifted block outputs, overlap is $N_h - 1 = 2$, and (h) the sum of overlapped block outputs equivalent to	

	the direct convolution result.....	64
Fig. 16	Block convolution using the OAS. (a) input signal $x(n)$ divided into overlapping sections, overlap is $N_h - 1 = 2$, (b) impulse response $h(n)$, (c) output $y(n)$ using direct convolution, (d) output $y_1(n)$ for block circular convolution of $x_1(n)$ and $h(n)$, (e) output $y_2(n)$ for block circular convolution of $x_2(n)$ and $h(n)$, (f) output $y_3(n)$ for block circular convolution of $x_3(n)$ and $h(n)$, (g) output $y_4(n)$ for block circular convolution of $x_4(n)$ and $h(n)$, and (h) sequential concatenation of block outputs after discarding the first two samples of each block, which is equivalent to the direct convolution result. " " represents concatenation.....	66
Fig. 17	LMS block diagram.....	68
Fig. 18	The illustration of the ULA. A point source is located at the farfield.	74
Fig. 19	The localized regions (dashed lines) on the microphone surface with the inter-element spacing $d = 0.6\text{m}$. The symbol " \square " indicates the microphone position.....	80
Fig. 20	Flowchart of the IBMC optimization algorithm.....	81
Fig. 21	The grid on a localized region.....	82
Fig. 22	Flowchart of the SA optimization algorithm.....	83
Fig. 23	A unimodal function to optimize.....	84
Fig. 24	The flowchart of the GSS algorithm.....	85
Fig. 25	A URA. A point sound source is located at the farfield.	89
Fig. 26	The DAS beamformer.....	90
Fig. 27	The block diagram of the TR algorithm.	93
Fig. 28	The block diagram of the SIMO-ESIF algorithm.....	96
Fig. 29	A plane wave incident from the direction κ to a farfield array.	107

Fig. 30	Diagram of first-order microphone composed of two microphones.	108
Fig. 31	FBR of first-order microphone versus the first-order differential parameter α_1	109
Fig. 32	Directivity index of first-order microphone versus the first-order differential parameter α_1	110
Fig. 33	Various first-order directional responses (a) dipole, (b) cardioids, (c) hypercardioid, (d) supercardioid.....	112
Fig. 34	The directivity pattern of first-order DMAs. (a) $\alpha_1=0.25$, (b) $\alpha_1=0.5$	113
Fig. 35	The effects of the regularization constant ε varying from 0.001 to 3 of a broadside array. (a) DI, (b) the 2-norm of the optimal array weights.....	114
Fig. 36	The contour plots of beam patterns of a broadside array. The x -axis represents the angle in degrees and the y -axis represents the frequency in kHz. (a) Superdirective array, (b) DAS array.....	115
Fig. 37	Comparison of the frequency responses of the superdirective array filters. (a) Microphones 1 and 4, (b) microphones 2 and 3, (c) phase responses.....	118
Fig. 38	The directional response of the superdirective array steered to -20 degrees. (a) 1 kHz, (b) 2 kHz, (c) 4 kHz.....	120
Fig. 39	Farfield array parameters.....	122
Fig. 40	Decide source numbers by AIC algorithm. The truncated order is corresponded to the lowest point. (a) Error, weight and AIC lines, (b) AIC lines with different weights.....	129
Fig. 41	The array configurations for DOA estimation simulations and experiments. (a) URA and (b) optimized random array.....	133
Fig. 42	The noise maps of two simulated point sources obtained using five	

acoustic imaging algorithms with 30-channel URA and random array. The squares are the preset sound source positions. The simulated whitenoise sources located at the positions $(-0.5\text{m}, 0.5\text{m})$ and $(0.5\text{m}, -0.5\text{m})$. The observed frequency is 1 kHz ($d = \lambda/4$). The power spectrums obtained using (a) DAS with URA configuration, (b) DAS with random array configuration, (c) TR with URA configuration, (d) TR with random array configuration, (e) SIMO-ESIF with URA configuration, (f) SIMO-ESIF with random array configuration, (g) MVDR with URA configuration, (h) MVDR with random array configuration, (i) MUSIC with URA configuration and (j) MUSIC with random array configuration..... 138

Fig. 43 The noise maps of two simulated point sources obtained using five acoustic imaging algorithms with 30-channel URA and random array. The squares are the preset sound source positions. The simulated white noise sources located at the positions $(-0.5\text{m}, 0.5\text{m})$ and $(0.5\text{m}, -0.5\text{m})$. The observed frequency is 7 kHz ($d = 2\lambda$). The power spectrums obtained using (a) DAS with URA configuration, (b) DAS with random array configuration, (c) TR with URA configuration, (d) TR with random array configuration, (e) SIMO-ESIF with URA configuration, (f) SIMO-ESIF with random array configuration, (g) MVDR with URA configuration, (h) MVDR with random array configuration, (i) MUSIC with URA configuration and (j) MUSIC with random array configuration..... 144

Fig. 44 Schematic diagram of a microphone array in the near field of the radiated sound from an acoustic source..... 150

Fig. 45 A set of microphones positioned close to the integration surface S 159

Fig. 46	Boundary elements used in isoparametric transformation. (a) Quadratic quadrilateral element, (b) quadratic triangular element.	160
Fig. 47	The representation of single-layer and double-layer potential.	162
Fig. 48	The definitions of important surfaces used in the ESM-based NAH. The symbol \mathbf{x}_m is the m th microphone position on the hologram surface S^h . The symbol \mathbf{z}_i is the i th source point on the actual source surface S^s . The symbol \mathbf{y}_n is the n th virtual source point on the virtual source surface S^v	164
Fig. 49	Schematic diagram for planar sources in configurations 1 and 2.	172
Fig. 50	The block diagram of inverse filtering process.	175
Fig. 51	The inverse filtering process viewed as a model matching problem.	176
Fig. 52	Block diagram of the Kalman filter.	181
Fig. 53	Scenarios of array element and actual source point distribution for simulating the point sources. (a) $S_{s,1}$ and (b) $S_{s,2}$	182
Fig. 54	The reconstructed and the desired amplitude profiles of point sources scenario. (a) $S_{s,1}$ and (b) $S_{s,2}$	183
Fig. 55	Variation of condition number of propagation matrix for a 1D array. (a) Condition number vs. kd , (b) singular value distribution of propagation matrix (frequency = 200 Hz, the spacing between microphones and between focal points $d = d_f = 0.0858$ m, $kd = 0.3142$). The parameter k is the wave number, d is the spacing of array, G is the propagation matrix, and L is the DOR. There are 30 microphones and 30 focal points. The number of null points is one at each end.	186
Fig. 56	Structure of 2D URA.	195

Fig. 57	Illustration of the GCV. The minimum is circled in the figure.....	198
Fig. 58	Illustration of the L-curve. The Circle in the figure is corner of the curve.....	199
Fig. 59	Singular values of the Hankel matrix $\mathbf{H}(0)$. The singular values above the order v are negligibly small.	202
Fig. 60	Illustration of the Overlap and add method. (a) The pressure data $p(n)$, (b) Decomposition of $p(n)$ into non-overlapping sections of length L , (c) Result of convolving each section with the inverse filter.....	206
Fig. 61	The array settings for NESI using a 4×4 URA.....	209
Fig. 62	Structure of pass-by test.	214
Fig. 63	Side-view of pass-by test.....	215
Fig. 64	Arrangement of the moving source in the pass-by test.	216
Fig. 65	Time-frequency diagram.....	217
Fig. 66	Farfield polar radiation pattern.....	218
Fig. 67	The idea of the NESI with virtual microphone technique. The symbol“ \odot ” indicates an interpolated microphone position. The symbol“ \oplus ” indicates an extrapolated microphone position. <u>A</u> The pressure data picked up by the microphones, <u>B</u> Reconstructed source strength at the focal points, <u>C</u> The pressure data interpolated at the virtual microphones, <u>D</u> Reconstructed source strength at the virtual focal points.....	222
Fig. 68	Scenarios of array element and actual source point distribution for simulating the planar baffled piston. (a) $S_{s,1}$, (b) $S_{s,2}$, (c) $S_{s,3}$, and (d) $S_{s,4}$	232
Fig. 69	Search for the optimal RD in a plane by virtual source configurations 2. The symbol“ \diamond ” indicates the coarse search of $S_{s,1}$. The	

symbol“◆” indicates the optimal point of $S_{s,1}$. The symbol “▽” indicates the coarse search of $S_{s,2}$. The symbol“▼” indicates the optimal point of $S_{s,2}$. The symbol “△” indicates the coarse search of $S_{s,3}$. The symbol“▲” indicates the optimal point of $S_{s,3}$. The symbol “⊖” indicates the coarse search of $S_{s,4}$. The symbol“●” indicates the optimal point of $S_{s,4}$233

Fig. 70 The reconstructed and the target velocity profiles of planar source scenario $S_{s,3}$ at $f= 850$ Hz.234

Fig. 71 Relative error (%) of velocity reconstructed for planar sources using virtual source configuration 2.235

Fig. 72 Configuration of a vibrating rigid piston spanned with an angle θ_0 , set at the top of a rigid sphere of radius a236

Fig. 73 Mesh configuration of the spherical baffled piston. The symbol“ ” indicates the microphones. The red mesh indicates the actual source surface. The yellow mesh indicates the virtual source surface. These three meshes are concentric.....237

Fig. 74 Search for the optimal RD of the spherical baffled piston by virtual source configuration 1. The symbol“◇” indicates the coarse search of the case $\theta_0 = 30^\circ$. The symbol“◆” indicates the optimal point. The symbol “▽” indicates the coarse search of the case $\theta_0 = 60^\circ$. The symbol“▼” indicates the optimal point. The symbol “⊖” indicates the coarse search of the case $\theta_0 = 180^\circ$. The symbol“●” indicates the optimal point.....238

Fig. 75 Polar plot of the desired and the reconstructed velocity of the spherical baffled piston with various subtending angles. The virtual source

	configuration 1 is used. (a) $\theta_0 = 30^\circ$, (b) $\theta_0 = 60^\circ$, (c) $\theta_0 = 180^\circ$ at $f = 1250$ Hz.....	240
Fig. 76	Relative error (%) of surface velocity reconstructed for the spherical sources using virtual source configuration 1. Three spanning angle $\theta_0 = 30^\circ, 60^\circ$ and 180° are examined.	241
Fig. 77	Flowchart demonstrating the procedure of nearfield beam pattern calculation.....	245
Fig. 78	The block diagram of ESM.	246
Fig. 79	The URA with inter-element spacing 0.6m (3λ at the frequency 1.7 kHz) for farfield imaging. (a) Array deployment, (b) beam pattern.....	254
Fig. 80	The farfield array optimized using the MC and the SA algorithms without the IB constraint at the frequency 1.7 kHz. For the MC simulation, maximum cost function $Q = 2.6532$ is attained at the 27596 th iteration. For the SA simulation, maximum cost function $Q = 2.7561$ is attained at the 1283 rd iteration. The circle indicates the main-lobe. (a) Learning curve of the MC search, (b) optimal array deployment obtained using the MC search, (c) beam pattern obtained using the MC search, (d) learning curve of the SA search, (e) optimal array deployment obtained using the SA search, (f) beam pattern obtained using the SA search, (g) beam pattern with weights optimized..	258
Fig. 81	The farfield array optimized using the MC algorithm and the combined SA-IBMC algorithm with the IB constraint at the frequency 1.7 kHz. Maximum cost function $Q = 2.6602$ is attained at the 1429th iteration. The circle indicates the main-lobe. (a) Learning curve, (b) optimal array deployment obtained using the SA search, (c) beam pattern	

obtained using the SA search, (d) optimal array deployment obtained using the SA-IBMC search, (e) beam pattern obtained using the SA-IBMC search, (f) beam pattern with weights optimized.261

Fig. 82 The cost function history of nearfield array optimization with inter-element spacing 0.6m (3λ at the frequency 1.7 kHz). Maximum cost function value is $Q = 27.3$ for the initial URA deployment. (a) All Q values, including those rejected during the random search by the MC algorithm, (b) all Q values, including those rejected during the random search by the IBMC algorithm, (c) only the legitimate values accepted by the SA algorithm, (d) only the legitimate values accepted by the SA-IBMC algorithm.270

Fig. 83 The beam pattern of the optimized nearfield array (URA). The inter-element spacing is 0.6m (3λ at the frequency 1.7 kHz). (a) At the frequency 500 Hz, $Q = 28.2$, (b) at the frequency 1 kHz, $Q = 26.9$, (c) at the frequency 1.7 kHz, $Q = 27.3$; (d) the beam pattern of a non-optimized nearfield array at the frequency 1.7 kHz by the SA and SA-IBMC algorithms, $Q = 3.8502$. The circles indicate the main-lobe. The symbol“ • ” indicates the focal points.272

Fig. 84 The beam pattern of a non-optimized nearfield array. The microphone deployment is identical to the optimized farfield array obtained using the SA-IBMC algorithm. (a) At the frequency 500 Hz, $Q = 3.5$, (b) at the frequency 1 kHz, $Q = 3.4$, (c) at the frequency 1.7 kHz, $Q = 3.9$. The circles indicate the main-lobe. The symbol“ • ” indicates the focal points.275

Fig. 85 The experimental arrangement for a wooden box with a loudspeaker fitted inside ,the URA, and a 30-channel random array optimized for

	farfield imaging are also shown in the picture.....	277
Fig. 86	Sensor calibration.....	278
Fig. 87	Close-up view of the 5×6 rectangular array.....	280
Fig. 88	The ERA technique. (a) The singular value plot of $\mathbf{H}(0)$, (b) the frequency responses of the inverse filter C_{13} of the original filter and the filter regenerated using ERA with $\nu = 80$, (c) the impulse responses of the inverse filter C_{13} of the original filter and the filter regenerated using ERA with $\nu = 80$	284
Fig. 89	The scooter experiment. The NESI was applied to reconstruct the sound field on the right side of the scooter running at the idle speed. (a) The unprocessed rms sound pressure image received at the microphones, (b) the rms source strength image, (c) the reconstructed rms sound pressure image, (d) the reconstructed rms particle velocity image, (e) the reconstructed rms sound intensity image.....	287
Fig. 90	The results of run-up experiment obtained using NESI with the 4×4 URA. The scooter engine was accelerated from 1500 rpm to 7500 rpm within ten seconds. (a) The unprocessed sound pressure image received at the microphones, (b) the reconstructed active intensity image, (c) the reconstructed active intensity image using the virtual microphone technique. The symbol“□” indicates the microphones. The symbol“•” indicates the focal points.....	289
Fig. 91	The results of compressor experiment obtained using the 5×6 URA. The major noise is at the air intake position situated at (0.2m, 0.3m). The observed frequencies in the algorithms are chosen to be 1.2 kHz. (a) The reconstructed sound pressure image by Fourier NAH, (b) the reconstructed sound pressure image by NESI, (c) the source image	

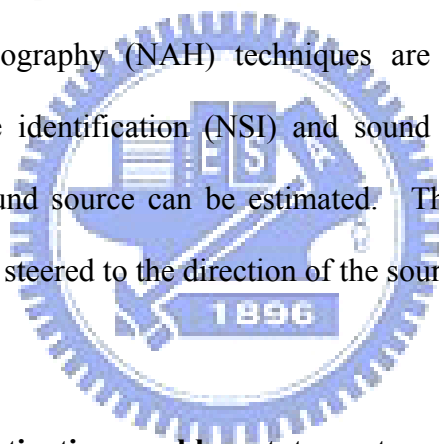
	obtained by using DAS, (d) the source image obtained by using TR, (e) the source image obtained by using MVDR, (f) the source image obtained by using MUSIC.	293
Fig. 92	The noise map within the band 1200 ~ 1300 Hz obtained using the 5×6 URA for the desktop computer. (a) The unprocessed sound pressure image received at the microphones, (b) the active intensity image reconstructed using NESI. The symbol“•” indicates the focal points.....	296
Fig. 93	The active intensity map within the band 1200 ~ 1300 Hz obtained using the random array for the desktop computer. The reconstructed active intensity image. The symbol“•” indicates the focal points.	298
Fig. 94	The experimental arrangement for a wooden box model with a loudspeaker fitted inside and holes cut on the frontal surface.	300
Fig. 95	The results of a wooden box with a loudspeaker fitted inside. The noise map is within the band 200 Hz ~ 1.6k Hz. (a) The unprocessed sound pressure image received at the microphones by 5×6 URA, (b) the particle velocity image reconstructed using NESI by the 5×6 URA.	301
Fig. 96	Experimental arrangement for an aluminum plate.	305
Fig. 97	Mode shape at the resonant frequency 594 Hz obtained using a salt sprinkle test.	306
Fig. 98	The surface velocity of the plate. (a) The velocity distribution of the plate measured by the scanning laser. The surface velocity reconstructed using various RDs (b) 1/20 <i>d</i> , (c) 1/2 <i>d</i> , (d) 1 <i>d</i> and (e) 2 <i>d</i> . The microphones are indicated in the figure using black dots.....	309
Fig. 99	The velocity data reconstructed by the ESM using <i>d</i> /2 RD and the surface velocity data measured by the scanning laser are compared for	

all lattice points.....310



CHAPTER 1. INTRODUCTION

A microphone array refers to a collection of microphones operating concurrently to achieve certain acoustic signal process purpose. Microphone array systems have received much research interest as a means of acoustic pickup utilized in various applications. Array signal processing [1]-[6] has been widely used in the areas such as non-destructive evaluation [7]-[8], underwater imaging [9]-[10], and machine diagnosis [11]-[12] and so forth. The major advantage is the enhancement of the signal-to-noise ratio (SNR). In addition, the directivity of the microphone array can be improved to be effective in eliminating background noise by beamforming techniques. Array signal processing of direction of arrival (DOA) estimation and nearfield acoustical holography (NAH) techniques are utilized to do the DOA estimation, noise source identification (NSI) and sound field visualization (SFV). The locations of the sound source can be estimated. Then, the main beam of the microphone array can be steered to the direction of the source of interest.



1.1 Background and motivation: problem statement

NSI is a vital step prior to a successful noise control program. Noise sources largely fall into two categories: vibration-induced noise and flow-induced noise. Examples of the first category include noise from rotating machinery, impact noise, noise due to structural resonance, braking squeal, etc., while examples of the second category include fan noise, pump noise, jet noise, etc. NSI techniques have been extensively studied by acoustical engineers. Among the NSI methods, SFV techniques are particularly useful in estimating the source position and the source strength as shown in Fig. 1. In addition to NSI, SFV techniques also find applications in non-destructive evaluation, underwater imaging, and machine

diagnosis and so forth. Some of simply noise problem can be solved by using the ears. However, it is difficult to characterize the problem without employing a NSI tool in many cases. The NSI tools have many kinds such as intensity probe, accelerometer, laser vibrometer, Beamforming, NAH, etc. Intensity probe measures at one point for stationary noise. The distance of measurement is very close to the source surface. In addition, intensity probe is not congruence when surface vibration is of interest. Several accelerometers are mounted on the vibrational surface to measure non-stationary vibrations. Displacement, velocity and acceleration can be measured by using accelerometers. The shortcoming of this tool is that can not used on rotational surface. Additional tool is the laser vibrometer that is a non-contact approach. Velocity is obtained at a single point on the source surface. In many cases, scanning laser is employed to measure velocity at several points within a short time interval. The shortcoming of laser vibrometer is that it must line of sight to the source surface. Hence, laser vibrometer can not measure synchronous noise source signal when noise source is non-stationary. In addition, that distance of measurement must be large to scan a large area. Hence, this tool is congruence in wide spaces. Two following tools, beamforming and NAH are based on an array of microphones. They measure pressure of the sound field to achieve NSI. Beamforming exploits array signal processing to estimate DOA. In early development, beamforming was primarily based on the farfield assumption that the source is far away and the waves become spherical or planar at the array position. It is an efficient tool for NSI by one single measurement with the microphone array technique. The limitation of beamforming is that it does not find acoustic variable near the source surface. Finally, NAH serves as powerful tools for the purposes of NSI. It is based on measurements with the microphone array in the nearfield of the sound pressure to create a map of the sound field on the source surface with a high

spatial resolution. As early development, Fourier Transform-based NAH techniques were suggested for reconstructing regularly shaped sources with planar, spherical, cylindrical geometries, etc. It is different from preceding techniques NAH can provide sound pressure, velocity, intensity and sound power on the source. In addition, it is applicable in narrow spaces and non-stationary source. The limitation of NAH is that it is costly because it must employ many channels of microphones to measure sound field at the same time.

In array implementation, transducer deployment has been one of the key issues. It is well known that, for uniform linear arrays (ULA) and uniform rectangular arrays (URA), array deployment must comply with the $\lambda/2$ -rule to avoid the spatial aliasing and the grating lobe problems. Consequently, a large number of microphones are required to cover the source area, which can render the array configuration impractical for sources at high frequencies. This prompts the development of non-uniform arrays that are capable of achieving high resolution and aliasing-free imaging with sparse sensors.

Given the fact that random deployment can be useful to farfield arrays, a question arises naturally. Can this idea of random deployment be carried over to nearfield arrays? This interesting issue is relatively unexplored in the literature of the past. The $\lambda/2$ -rule --a widely accepted criterion in deploying NAH arrays-- can lead to undesirable high channel count. It is then tempting to “randomize” the sensor positions like in the case of farfield arrays and achieve comparable performance with sparse deployment. To explore this conjecture, this thesis is employ optimization techniques for microphone deployment. Optimization of microphone deployment was carried out for both farfield and the nearfield arrays.

1.2 Review of prior arts: approaches for noise identification problems

NSI techniques [7], [8], [11] fall into two categories: farfield [13], [14] and nearfield [15]. The beamforming in the conventional sense is generally based on the assumption that the incoming waves are plane waves from the farfield. The NAH in the conventional sense is generally based on the assumption that the incoming waves are spherical waves from the nearfield. Array technology has been used in many diverse areas including radar [16], sonar [17], radio astronomy [18], telecommunications [19], and so forth. Its application encompasses purposes including signal enhancement, spatial filtering, DOA estimation, etc. Beamforming algorithms such as the delay-and-sum algorithm [20], the Minimum Variance Distortionless Response (MVDR) algorithm [21], and the multiple signal classification (MUSIC) algorithm [22] have been suggested in the past. Recently, array technology has found application in NSI with the use of microphones. Microphone arrays serve as a powerful tool for acoustic field visualization that enables effective estimation of the positions and strengths of noise sources [10], [11], [20] [23]. In comparison with the farfield arrays that are particularly useful for long-distance and large scale sources such as trains and aircrafts, nearfield arrays represent a more recent and independent development of source imaging technology. With the advent of the NAH [15], [24], many nearfield acoustic imaging techniques including the inverse boundary element method (IBEM) [25]-[32], the Helmholtz equation least squares (HELs)[33]-[36], the equivalent source method (ESM) [37]-[42], the nearfield equivalence source imaging (NESI) [43]-[45], etc., are introduced. These nearfield techniques are well suited for imaging small-scale sources such as cars and computers by virtue of high resolution focusing schemes. IBEM makes use of many numerical techniques akin to finite element analysis (FEA). Although the dimensionality is reduced by one compared to the FEA for a

three-dimensional acoustical radiation problem, its implementation is still rather computationally expensive. The HELS method that was suggested in late 90s is formulated using field representation by spherical harmonic expansion, whereas the IBEM is derived from a field representation on the basis of the Green's function theory. The IBEM can be further classified into two categories: the direct formulation and the indirect formulation. The direct formulation is derived from the Helmholtz integral equation (HIE) [25]-[29], whereas the indirect formulation stems from simple or double layer potential [30]-[32]. The ESM, also known as wave superposition method [46]-[48], were suggested for sound field calculation with far less complexity and thus higher computational efficiency. The ESM represents the sound field of interest with distribution of discrete simple sources and hence no numerical integration is required. As opposed to the actual source, these simple sources are only in virtual sense for representation purpose and are solutions of the acoustic wave equation. Although the ESM was generally used as a benchmark for BEM, it has been shown with careful choice of parameters that the ESM is capable of achieving comparable accuracy compared to other methods [39], [49]. Like IBEM, the use of ESM is not restricted to source with regular geometries. NESI was proposed for NSI and sound field reconstruction. The NESI *per se* can be considered as a time-domain ESM. Multichannel inverse filters are designed offline using truncated singular value decomposition (TSVD) or Tikhonov regularization. Since all the required processing is carried out in the time-domain, NESI eliminates many problems of Fourier based NAH. It applies not only to stationary but also a transient noise. Multichannel inverse filters are designed in advance. An Eigensystem Realization Algorithm (ERA) is employed to accelerate the multichannel filtering of the NESI [43], [45] and [50].

Fourier based NAH, a celebrated technique for NSI was pioneered by Maynard

and Williams in early 1980s [15], [24], [51] and [52]. The technique enables reconstructing the 3D sound field from the 2D hologram data scanned above the source surface. In comparison with another commonly used NSI technique, the sound intensity method [53], the Fourier based NAH provides a more global view of noise distribution and the relative strength. The Fourier based NAH has now been commercialized by Brüel and Kjael with a new name, spatial transformation of sound field (STSF) [54]-[56]. A comprehensive treatment of Fourier based NAH can be found in the monograph by Williams [57]. While the Fourier based NAH represented an elegant solution in the SFV techniques, there are a number of limitations in the original version of the Fourier-based NAH. Most of these limitations stem from the fact that the Fourier NAH transformation relies on the two-dimensional fast Fourier transform (2D FFT) between the physical space and the wave number space. This implies that stationary frequency-domain pressure phasors must be available and the scanned grid points must be equally spaced on a planar rectangular area [58]. Numerical artifacts such as aliasing problem arising in fast Fourier transform (FFT) may adversely affect the accuracy of imaging. This situation is further aggravated in practical application, where the number of sensor and data acquisition channel is usually quite limited due to cost consideration. To deal with these limitations, the Fourier based NAH was later extended to problems with irregular geometries [25], [59] and non-stationary noise [56]. The inverse reconstruction techniques have been extended to deal with irregular shaped sources by applying singular value decomposition (SVD) to the IBEM [60]. In spite of all that, one to three times of FFT has to be carried out with the microphone spacing kept less than half of a wavelength to avoid the spatial aliasing problem. This requires large number of microphones and enormous processing power to cover a reasonable source area. In addition, latency becomes more of a problem when one has to measure a

transient noise such as pass-by noise in real time. Another NAH method, statistically optimal NAH (SONAH) [61], was also suggested, based on plane wave expansion. The idea in SONAH is to expand the sound field into elementary waves and the estimate sound pressure on source surface in wave number domain. The estimated weight function can then be reused to reconstruct the sound field at another position on the source. Comprehensive coverage of NAH can be found in the monograph by Williams [57] and a recent tutorial paper by Wu [62].

1.3 Organization of the thesis

The organization of the thesis is summarized as follows. Chapter 2 gives theoretical preliminaries of acoustics. A fundamental of acoustics includes wave equation, Helmholtz equation, time harmonic conventional and acoustical boundary value problems. Secondly, sound field representation using basis function expansion is described by separable coordinates in 3D space: Cartesian, cylindrical, and spherical. Next, sound field representation using the HIE is described. It includes the Green's functions, eigenfunction expansion of the Green functions. Monopole representation, dipole representation and multipole expansion of basic source model are described. Finally, inverse problems and ill-posedness is described and discussed. Chapter 3 describes theoretical preliminaries of signal processing, which includes SVD, array signal processing basics and optimization algorithms. Array model, beam pattern, grating lobes, spatial aliasing and performance measures are introduced in array signal processing basics. Optimization algorithms include Golden ratio search, Monte Carlo (MC) and Simulated annealing (SA). Chapter 4 describes farfield array signal processing algorithms, which includes low-resolution algorithms and high-resolution algorithms. Delay and sum (DAS) beamformer and time reversal (TR) beamformer are described in the low-resolution algorithms.

Furthermore, cost functions of optimal arrays are defined. Choice of array parameters for farfield can be summarized in this section. In high-resolution, MVDR and MUSIC is described. Choice of parameters such as Akaike information criterion (AIC) is introduced in this chapter. Finally, optimal array and comparisons of FF algorithms are summarized. Chapter 5 describes nearfield array signal processing algorithms such as Fourier based NAH, IBEM and ESM. In Fourier based NAH, three coordinates that is planar, cylindrical and spherical geometries are summarized. In addition, direct and indirect formulation of IBEM is presented. Finally nearfield signal processing algorithms is ESM. In addition, direct and indirect formulation of ESM, NESI and Kalman filter-based algorithm are presented. Two virtual source configurations are explained in this section. Furthermore, comparisons of nearfield algorithms are summarized. Chapter 6 is practical implementations such as inverse filter design, multi-channel fast filtering, post-processing, choice of distance of reconstruction (DOR) and lattice spacing, virtual microphone technique, choice of RD, optimization of sensor deployment, system integration and experimental arrangement are all described in this section. In inverse filter design, model matching problem such as ill-posedness and regularization, window design and parameter choice methods (PCM) are described. The multi-channel fast filtering is includes time-domain and frequency-domain processing. As regards post-processing, acoustic variables is described. Part of choice of RD, golden section search (GSS) is employed to determine optimal RD. Optimization algorithms are used to calculate optimizing farfield and nearfield sensor deployment. Finally, system integration and experimental arrangement is described in this chapter. Chapter 7 is application examples such as scooter, information technology (IT) equipment, compressor and non-contact modal analysis. Chapter 8 restates the main conclusions of the thesis.

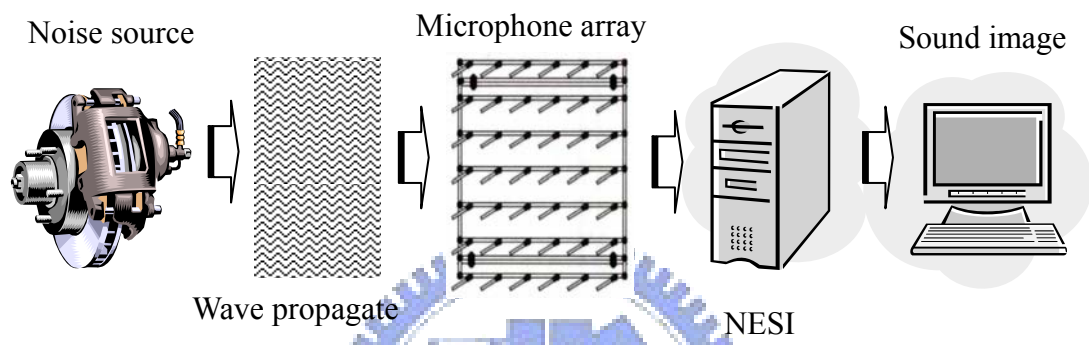


Fig. 1 Microphone array system for NSI and SFV.



CHAPTER 2. THEORETICAL PRELIMINARIES OF ACOUSTICS

2.1 Fundamentals of acoustics

Sound wave is a mechanical wave that propagates through a compressible medium. Sound wave is transmitted in the form of a longitudinal wave described by wave equation. Wave equation satisfies consideration of continuity, equilibrium of strength and thermodynamics principle. The Eulerian coordinates and Lagrangian coordinates is introduced. Eulerian coordinates fix control volume in space and Lagrangian coordinates fix control volume as the medium flows, as shown in Fig. 2. However, Lagrangian coordinates is convenient to use in conservation laws. Material derivative in Lagrangian coordinates is

$$\frac{DM}{Dt} = \frac{\partial M}{\partial t} + (\mathbf{u} \cdot \nabla)M, \quad (2.1)$$

where M is mass, t is time, \mathbf{u} is velocity of fluid and ∇ is gradient. In addition, integration and differentiation cannot only be switched by adding a correction term as

$$\frac{D}{Dt} \int_V \alpha dV = \int_V \left[\frac{\partial \alpha}{\partial t} + \nabla \cdot (\alpha \mathbf{u}) \right] dS, \quad (2.2)$$

where V is control volume, α is physical quantity and $\nabla \cdot$ is divergence. This equation is calling Reynold's transport theorem.

In deriving wave equation, medium is assumed in homogeneous, isotropic, inviscid and adiabatic. In equation of state, instant pressure of ideal gas is

$$p = p(\rho) = K \rho^\gamma, \quad (2.3)$$

where ρ is instant density, K is constant and γ is specific-heat ratio. The density is non-linear function. Linearization of Eq. (2.3) is

$$p = p_0 + \left(\frac{\partial p}{\partial \rho} \right)_0 (\rho - \rho_0) + \dots$$

Define sound speed is

$$c \triangleq \sqrt{\left(\frac{\partial p}{\partial \rho}\right)_0}.$$

Therefore, the linearized equation of state is obtained

$$p' = p - p_0 \approx c^2 \rho'. \quad (2.4)$$

That is similar Hook's law to describe relation between sound pressure p' (elasticity) and density ρ' (distortion).

In equation of continuity, the total mass in a control volume of flowing with the medium does not change (base on conservation laws)

$$\frac{DM}{Dt} = 0 \quad (2.5)$$

or

$$\frac{D}{Dt} \int_V \rho dV = 0.$$

By Reynold's transport theorem, above equation can be rewritten

$$\int_V \left[\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) \right] dV = 0.$$

Because V is arbitrary volume, so

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \quad (2.6)$$

that is equation of continuity.

Equation of momentum is base on linear momentum principle as

$$\frac{d\mathbf{L}}{dt} = \mathbf{f}_{\text{ext}}$$

$$\mathbf{L} = \int_V \rho \mathbf{u} dV,$$

where \mathbf{L} is linear momentum and \mathbf{f}_{ext} is force. Problem to sound field take

Largrangian coordinates:

$$\frac{D}{Dt} \int_V \rho \mathbf{u} dV = \int_S \mathbf{f}_s dS, \quad (2.7)$$

where \mathbf{f}_s is force on S . When medium is inviscid, the force is become

$$\mathbf{f}_s = -p\mathbf{n}, \quad (2.8)$$

where \mathbf{n} is outward normal. The two equations above can be rewritten into

$$\frac{D}{Dt} \int_V \rho \mathbf{u} dV = - \int_S p \mathbf{n} dS. \quad (2.9)$$

Using divergence theorem to obtain

$$\frac{D}{Dt} \int_V \rho \mathbf{u} dV = - \int_V \nabla p dV.$$

It can describe by Cartesian coordinates

$$\frac{D}{Dt} \int_V \rho u_j dV = - \int_V \frac{\partial p}{\partial x_j} dV, j = 1, 2, 3.$$

By Reynold's transport theorem, above equation can be rewritten

$$\int_V \left[\frac{\partial}{\partial t} (\rho u_j) + \nabla \cdot (\rho u_j \mathbf{u}) \right] dV = - \int_V \frac{\partial p}{\partial x_j} dV, j = 1, 2, 3.$$

Because V is arbitrary volume, so

$$\frac{\partial}{\partial t} (\rho u_j) + \nabla \cdot (\rho u_j \mathbf{u}) = - \frac{\partial p}{\partial x_j}, j = 1, 2, 3.$$

The equation above can be expanded as

$$\rho \frac{\partial u_j}{\partial t} + u_j \frac{\partial \rho}{\partial t} + u_j \nabla \cdot (\rho \mathbf{u}) + \rho \mathbf{u} \cdot \nabla u_j = - \frac{\partial p}{\partial x_j}, j = 1, 2, 3. \quad (2.10)$$

Terms 2 and 3 on the left-hand side can be canceled by equation of continuity (in light of Eq. (2.6), so

$$\rho \frac{\partial u_j}{\partial t} + \rho \mathbf{u} \cdot \nabla u_j = - \frac{\partial p}{\partial x_j}, j = 1, 2, 3.$$

Now, we combine these three quantities ($j = 1, 2, 3$)

$$\rho \left[\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} \right] = - \nabla p.$$

Term 2 ($(\mathbf{u} \cdot \nabla) \mathbf{u}$) is called convection term that can be neglected when flow speed is small. Therefore, Euler's equation or momentum equation is obtained

$$\rho \frac{D\mathbf{u}}{Dt} = - \nabla p. \quad (2.11)$$

Finally, wave equation is combined by equation of state, continuity and momentum. Before combination, linearization is necessary. Assume small perturbation and mean flow ($\mathbf{u}_0 = 0$):

$$p = p_0 + p', \quad |p'| \ll |p_0| \quad (2.12)$$

$$\rho = \rho_0 + \rho', \quad |\rho'| \ll |\rho_0| \quad (2.13)$$

$$\mathbf{u} = \mathbf{u}_0 + \mathbf{u}', \quad |\mathbf{u}'| \ll |\mathbf{u}_0|. \quad (2.14)$$

By Eqs. (2.12) to (2.14), the continuity equation of Eq. (2.6) can be rewritten as

$$\frac{\partial}{\partial t}(\rho_0 + \rho') + \nabla \cdot [(\rho_0 + \rho')\mathbf{u}'] = 0. \quad (2.15)$$

The linearized continuity equation is obtained by removed height order term

$$\frac{\partial \rho'}{\partial t} + \rho_0 \nabla \cdot \mathbf{u}' = 0. \quad (2.16)$$

Similarly, equation of momentum in Eq. (2.11) by linearization can be obtained

$$\rho_0 \frac{\partial \mathbf{u}'}{\partial t} = -\nabla p'. \quad (2.17)$$

For simplicity, the prime “'” is removed and equation of state, continuity and momentum with linearization can be shown as

$$p = c^2 \rho \quad (2.18)$$

$$\frac{\partial \rho}{\partial t} + \rho_0 \nabla \cdot \mathbf{u} = 0 \quad (2.19)$$

$$\rho_0 \frac{\partial \mathbf{u}}{\partial t} = -\nabla p. \quad (2.20)$$

Three equations above have five unknown elements (p , ρ , u_1 , u_2 and u_3).

Algebra is employed to combine Eq. (2.18) to (2.20) as follows:

$\frac{\partial}{\partial t}$ Eq. (2.19):

$$\frac{\partial^2 \rho}{\partial t^2} + \frac{\partial}{\partial t}(\rho_0 \nabla \cdot \mathbf{u}) = 0 \quad (2.21)$$

$\nabla \cdot$ Eq. (2.20):

$$\nabla \cdot (\rho_0 \frac{\partial \mathbf{u}}{\partial t}) = \frac{\partial}{\partial t} (\rho_0 \nabla \cdot \mathbf{u}) = -\nabla \cdot \nabla p = -\nabla^2 p \quad (2.22)$$

Wave equation with linearization is substituting Eqs. (2. 18) and (2. 22) into Eq. (2. 21)

$$\frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} = \nabla^2 p. \quad (2.23)$$

In acoustical analysis, simple harmonic motion is usually employed to the analyze frequency-domain. Base sine wave is shown in Fig. 3 that can indicated as

$$p(x,t) = A \sin(\omega t - kx), \quad (2.24)$$

where x and t are variables of space and time; A is amplitude, ω is angular frequency and k is wave number. In $x = 0$, the sine wave is

$$p(0,t) = A \sin \omega t = A \sin 2\pi ft = A \sin \frac{2\pi}{T} t, \quad (2.25)$$

where f is frequency and T is period. In addition, the sine wave in $t = 0$ is

$$p(x,0) = -A \sin kx = -A \sin \frac{2\pi}{\lambda} x. \quad (2.26)$$

From two equations above, we can draw the analogy

$$\omega = \frac{2\pi}{T} \Leftrightarrow k = \frac{2\pi}{\lambda}, \quad (2.27)$$

where λ is wave length. In addition,

$$k = \frac{2\pi}{\lambda} = \frac{2\pi f}{\lambda f} = \frac{\omega}{c}. \quad (2.28)$$

The $p(x,t)$ can be written with a complex phasor

$$A \sin(\omega t - kx) = \text{Im}\{Ae^{j(\omega t - kx)}\} = \text{Im}\{Ae^{-jkx} e^{j\omega t}\}, \quad (2.29)$$

where $j = \sqrt{-1}$ and $\text{Im}\{\}$ is the imaginary part. If remove the time-harmonic function $e^{j\omega t}$, Ae^{-jkx} is the phasor in $p(x,t)$ as $\tilde{p}(x) = Ae^{-jkx}$. Meaning of the phasor is complex Fourier coefficients. From wave equation with linearization can be rewritten in the frequency-domain

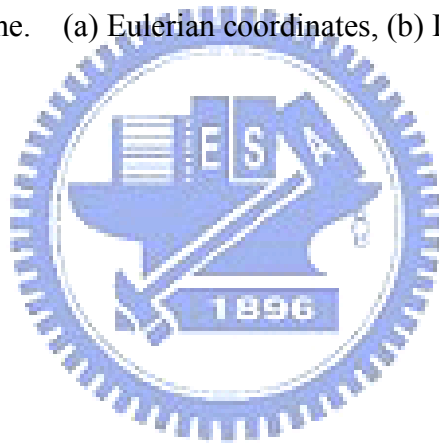
$$\begin{aligned}
\nabla^2 p &= \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} \\
\Rightarrow \nabla^2 \tilde{p} &= \frac{1}{c^2} (j\omega)^2 \tilde{p} \\
\Rightarrow \nabla^2 \tilde{p} + k^2 \tilde{p} &= 0,
\end{aligned}
\tag{2.30}$$

where $k = \omega/c$ is wave number. This reduced wave equation is also referred to as the Helmholtz equation. That is attention about the phasor, some reference employ $e^{-j\omega t}$ to express harmonic wave function. Hence, the $A \sin(\omega t - kx)$ of Eq. (2.24) can be rewritten as the $A \sin(kx - \omega t)$ when all analysis is unanimity. In thesis, the e^{-jkx} is employed to indicate of phasor.





Fig. 2 Control volume. (a) Eulerian coordinates, (b) Lagrangian coordinates.



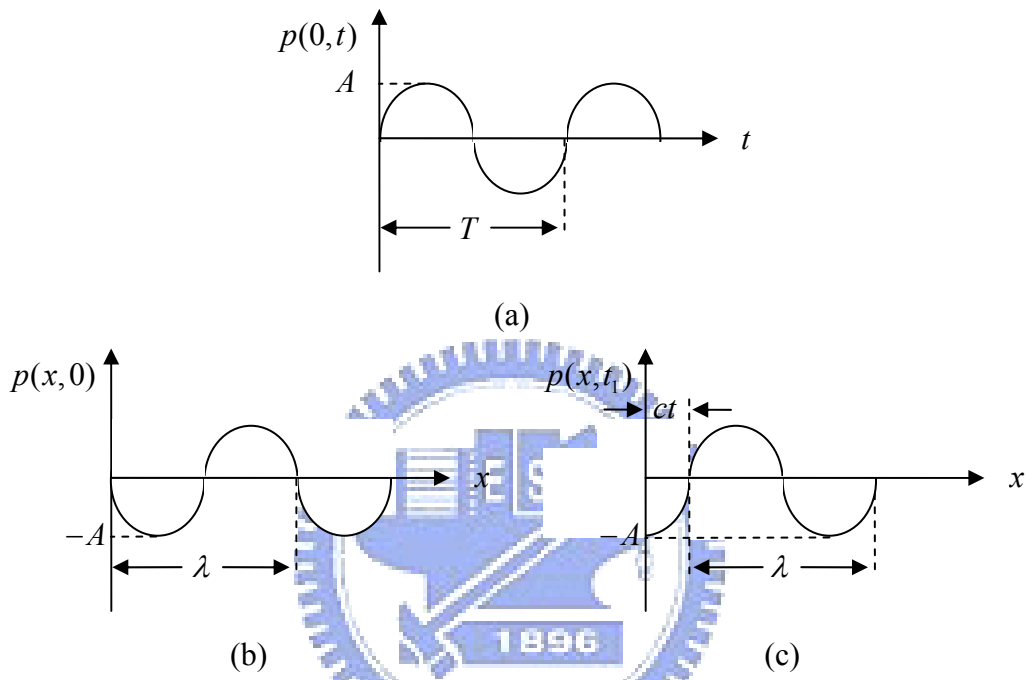


Fig. 3 A sine wave. (a) $x = 0$, (b) $t = 0$, (c) $t = t_1$.

2.2 Sound field representation using basis function expansion

Two mathematical principles underlying the solution of Laplace's equation with basis function expansions include orthonormality and completeness. The mathematical basis for completeness relations is given by Sturm-Liouville theory of second-order differential equations. Orthonormality and completeness in 1-D is instructed as following. We assume that boundary conditions are imposed at $x = a$ and $x = b$. We assume that there is a scalar product operation that takes two functions and maps them to a single number. In Sturm-Liouville theory, the inner product of two functions $f(x)$ and $g(x)$ is

$$(f, g) = \int_a^b f(x)g(x)w(x)dx, \quad (2.31)$$

where $w(x)$ that depends on the differential operator in the eigenvalue problem is a given function. For simplicity, the eigenvalues σ are discrete and $\sigma \in \{\sigma_i\}$ with $i = 1, 2, \dots$. Any function $U(x)$ satisfying the same boundary conditions as v_i may be written

$$U(x) = \sum_{i=1}^{\infty} U_i v_i(x), \quad (2.32)$$

where the coefficients U_i are unique constants and the functions $v_i(x)$ are similar basis vectors. Above equation is the expansion of a vector in basis vectors. Due to the dimensionality of a vector space is equal to the number of linearly independent basis vectors, this analogy is possible. The first condition is orthonormality of the basis functions. Sturm-Liouville theory shows that eigenfunctions with different eigenvalues are orthogonal with respect to the inner product as $(v_j, v_i) = 0$ if $j \neq i$. On the other hand, if $j = i$, the Sturm-Liouville scalar product is positive-definite. Then, the eigenfunctions may be normalized

$$(v_j, v_i) = \delta_{ji} \text{ (Orthonormality)}. \quad (2.33)$$

Scalar production of Eq. (2. 32) with v_i is taken and orthonormality is used to obtain

$$U_i = (v_i, U). \quad (2. 34)$$

The second condition expresses the completeness of the eigenfunctions. In Sturm-Liouville theory, completeness is established by showing that $(f, f) \geq 0$ for any function $f(x)$. Hence,

$$(\delta U, \delta U) = 0, \quad (2. 35)$$

where $\delta U \equiv U(x) - \sum_{i=1}^{\infty} (v_i, u)v_i(x)$. If the expansion coefficients U_i are given by Eq. (2. 34), the series given by Eq. (2. 32) converges to $U(x)$. Combining these two equations, following condition is required

$$\begin{aligned} U(x) &= \sum_{i=1}^{\infty} \left[\int_a^b U(x')v_i(x')w(x')dx' \right] v_i(x) \\ &= \int_a^b \left[\sum_{i=1}^{\infty} v_i(x)v_i(x')w(x') \right] U(x') dx'. \end{aligned} \quad (2. 36)$$

It turns out that for Sturm-Liouville problems this holds for any U in the function space, which requires

$$\sum_{i=1}^{\infty} v_i(x)v_i(x')w(x') = \delta(x-x') \quad (\text{Completeness}). \quad (2. 37)$$

In addition, orthonormlity and completeness on the sphere is introduced. The dependence on the two variables separates in spherical coordinates and the eigenfunctions are

$$U(\theta, \phi) = Y_n^m(\theta, \phi) \equiv \sqrt{\frac{2n+(n-m)!}{4\pi(n+m)!}} P_n^m(\cos\theta) e^{jm\phi}, \quad (2. 38)$$

where $Y_n^m(\theta, \phi)$ is spherical harmonics and $P_n^m(\cos\theta)$ Legendre functions. The boundary condition that U is finite on the sphere requires that n is an integer. Finally, the boundary conditions also require $-n \leq m \leq n$. Therefore, there are $2n+1$ values of the order m for each degree n . The spherical harmonic functions satisfies the following orthonormality condition

$$\int Y_n^m(\theta, \phi) Y_{n'}^{m'}(\theta, \phi) d\Omega = \delta_{nn'} \delta_{mm'}, \quad (2.39)$$

where * is complex conjugate, $d\Omega \equiv \sin\theta d\theta d\phi$, $0 \leq \theta \leq \pi$ and $0 \leq \phi \leq 2\pi$. The completeness relation for spherical harmonics is

$$\sum_{n=0}^{\infty} \sum_{m=-n}^n Y_n^m(\theta, \phi) Y_n^{m*}(\theta', \phi') = \frac{1}{\sin\theta} \delta(\theta - \theta') \delta(\phi - \phi'). \quad (2.40)$$

In the three-dimensional sound field, some coordinates is separable coordinates. The eigenfunction is existence. Eigenfunction satisfied the following Eigenvalue problem

$$\begin{aligned} (\nabla^2 + k^2)\psi_n(\mathbf{x}) &= 0 \\ \text{st. homogeneous BCs} \end{aligned} \quad (2.41)$$

Hence, eigenfunction expansion for sound pressure can be expressed as

$$p(\mathbf{x}) = \sum_{n=1}^{\infty} a_n \psi_n(\mathbf{x}). \quad (2.42)$$

Following is eigefunction for three coordinate systems:

(1) Cartesian coordinates:

In Fig. 4 (a), the eigenfunction of Cartesian coordinates is

$$\psi_n(\mathbf{x}) = e^{j(\pm k_x x \pm k_y y \pm k_z z)}, \quad (2.43)$$

where $k_x^2 + k_y^2 + k_z^2 = k^2$

(2) Cylindrical coordinates:

In Fig. 4 (b), the eigenfunction of cylindrical coordinates is

$$\psi_n(\mathbf{x}) = H_n^{(1), (2)}(k_r r) e^{\pm jn\theta e \pm jk_z z}, \quad (2.44)$$

where $k_r^2 + k_z^2 = k^2$, $H_n^{(1), (2)}$ is Hankel function of the 1st/2nd kind of order n

with a parameter k_r .

(3) Spherical coordinates:

In Fig. 4 (c), the eigenfunction of spherical coordinates is

$$\psi_n(x) = H_n^{(1), (2)}(kr)Y_n^m(\theta, \phi), \quad (2.45)$$

where $H_n^{(1), (2)}(kr)$ is spherical Hankel function of the 1st/2nd kind of order n

with a parameter k . The $Y_n^m(\theta, \phi)$ is called the spherical harmonic defined as

$$Y_n^m(\theta, \phi) = \sqrt{\frac{(2n+1)(n-m)!}{4\pi(n+m)!}} P_n^m(\cos\theta)e^{jm\phi}, \quad (2.46)$$

where $P_n^m(\cos\theta)$ is associated Legendre function of order m and degree n .



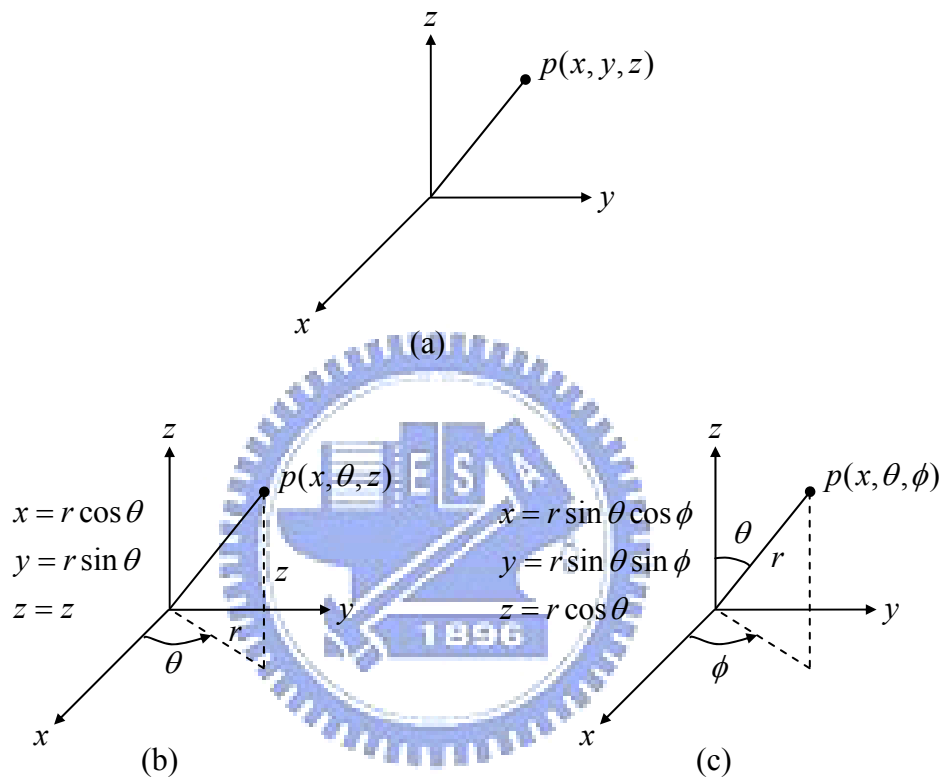


Fig. 4 Three basic coordinate systems. (a) Cartesian coordinates, (b) cylindrical coordinates, (c) spherical coordinates.

2.3 Sound field representation using Helmholtz integral equation

In this section, a mathematical tool for solving boundary value problems of sound field is introduced – the Green's function. The Green's function and eigenfunction expansion can be employed to solve inhomogeneous boundary value problem. The Green's function is that response is produced by a point source into sound. Assume mass m_s is poured into space V in Δt time at point source position \mathbf{x}_0 . Mass conservation law is used:

$$\frac{D}{Dt} \int_V \rho dV = \dot{m}_s$$

$$\int_V \left[\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) - \frac{\partial \dot{m}_s}{\partial t} \delta(\mathbf{x} - \mathbf{x}_0) \right] dV(\mathbf{x}) = 0$$

where $\delta(\mathbf{x} - \mathbf{x}_0)$ is Dirac's delta function at \mathbf{x}_0 . This Dirac's delta function satisfies the following conditions:

$$\int_V \delta(\mathbf{x} - \mathbf{x}_0) dV(\mathbf{x}) = 1, \mathbf{x}_0 \in V \quad (2.47)$$

and the sifting property

$$\int_V F(\mathbf{x}) \delta(\mathbf{x} - \mathbf{x}_0) dV(\mathbf{x}) = F(\mathbf{x}_0), \mathbf{x}_0 \in V. \quad (2.48)$$

Therefore, equation of continuity with point source is

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = \frac{\partial \dot{m}_s}{\partial t} \delta(\mathbf{x} - \mathbf{x}_0). \quad (2.49)$$

Then, inhomogeneous wave equation is

$$\nabla^2 p - \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} = -\ddot{m}_s \delta(\mathbf{x} - \mathbf{x}_0) \quad (2.50)$$

where $\ddot{m}_s = \partial^2 m_s / \partial t^2$ is mass acceleration. In harmonic sound field,

inhomogeneous Helmholtz equation can be written as

$$\nabla^2 \tilde{p} + k^2 \tilde{p} = \omega^2 \tilde{m}_s \delta(\mathbf{x} - \mathbf{x}_0). \quad (2.51)$$

Let $\omega^2 \tilde{m}_s = -4\pi$ and $\tilde{p} = g(\mathbf{x}, \mathbf{x}_0)$, equation above can be written as

$$\nabla^2 g(\mathbf{x}, \mathbf{x}_0) + k^2 g(\mathbf{x}, \mathbf{x}_0) = -4\pi\delta(\mathbf{x} - \mathbf{x}_0), \quad (2.52)$$

where function $g(\mathbf{x}, \mathbf{x}_0)$ is called the free-space Green's function. Following is the three free-space Green's functions:

(1) The one-dimensional sound field:

The free-space Green's function is

$$g(\mathbf{x}, \mathbf{x}_0) = \frac{j}{2k} e^{-jkr}, \quad (2.53)$$

where $r = |x - x_0|$.

(2) The two-dimensional sound field:

The free-space Green's function is

$$g(\mathbf{x}, \mathbf{x}_0) = \frac{j}{4} H_0^{(1)}(kr), \quad (2.54)$$

where $H_0^{(1)}$ is Hankel function of the 1st kind of order zero.

(3) The three-dimensional sound field:

The free-space Green's function is

$$g(\mathbf{x}, \mathbf{x}_0) = \frac{e^{-jkr}}{r}, \quad (2.55)$$

where $r = |x - x_0|$. This is sound field by spherical wave.

Then, how to use the Green's function to determine inhomogeneous boundary problem? We can use infinity point source set and point sources are distributed on boundary. If sound field is linear, the Green's second identity is employed as

$$\int_V (\phi_1 \hat{h} \phi_2 - \phi_2 \hat{h} \phi_1) dV = \int_S \left(\phi_1 \frac{\partial \phi_2}{\partial n} - \phi_2 \frac{\partial \phi_1}{\partial n} \right) dS, \quad (2.56)$$

where \hat{h} is a self-adjoint operator, $\partial\phi/\partial n \triangleq \nabla\phi \cdot \mathbf{n}$ is direction derivatives, \mathbf{n} is outward normal, and ϕ_1 and ϕ_2 are solution of second order differentiation. As shown in Fig. 5, our objective is to determine boundary problem of interior sound field

$$\begin{aligned}
(\nabla^2 + k^2)p(\mathbf{x}) &= f(\mathbf{x}) \\
\text{st. } p(\mathbf{x}_0) \text{ or } \frac{\partial p}{\partial n}(\mathbf{x}_0), \mathbf{x}_0 \in S,
\end{aligned} \tag{2.57}$$

where the symbol of phasor “~” is removed. If the boundary term $p(\mathbf{x}_0)$ is known, this is problem is called Dirichlet problem. This is called Neumann problem, if $\partial p / \partial n(\mathbf{x}_0)$ is know. Now, let $\phi_1 = p(\mathbf{x})$, $\phi_2 = g(\mathbf{x}, \mathbf{x}_0)$ and $\hbar \triangleq \nabla^2 + k^2$ from Eq. (2.56) as follows:

$$\begin{aligned}
& \int_V [p(\mathbf{x})(\nabla^2 + k^2)g(\mathbf{x}, \mathbf{x}_0) - g(\mathbf{x}, \mathbf{x}_0)(\nabla^2 + k^2)p(\mathbf{x})]dV(\mathbf{x}) \\
&= \int_S \left[p(\mathbf{x}) \frac{\partial}{\partial n} g(\mathbf{x}, \mathbf{x}_0) - g(\mathbf{x}, \mathbf{x}_0) \frac{\partial}{\partial n} p(\mathbf{x}) \right] dS(\mathbf{x}) \\
& \int_V [p(\mathbf{x}) - (4\pi\delta(\mathbf{x}, \mathbf{x}_0)) - g(\mathbf{x}, \mathbf{x}_0)f(\mathbf{x})]dV(\mathbf{x}) \\
&= \int_S \left[p(\mathbf{x}) \frac{\partial}{\partial n} g(\mathbf{x}, \mathbf{x}_0) - g(\mathbf{x}, \mathbf{x}_0) \frac{\partial}{\partial n} p(\mathbf{x}) \right] dS(\mathbf{x}) \\
&= -4\pi p(\mathbf{x}_0) - \int_V g(\mathbf{x}, \mathbf{x}_0)f(\mathbf{x})dV(\mathbf{x}).
\end{aligned}$$

Symmetry of the Green's function (exchange field point \mathbf{x} and source point \mathbf{x}_0) is employed to obtain the Kirchhoff-Helmholtz integral in frequency-domain

$$\begin{aligned}
4\pi p(\mathbf{x}) &= - \int_V f(\mathbf{x}_0)g(\mathbf{x}, \mathbf{x}_0)dV(\mathbf{x}_0) \\
& - \int_S \left[p(\mathbf{x}_0) \frac{\partial}{\partial n} g(\mathbf{x}, \mathbf{x}_0) - g(\mathbf{x}, \mathbf{x}_0) \frac{\partial}{\partial n} p(\mathbf{x}_0) \right] dS(\mathbf{x}_0).
\end{aligned} \tag{2.58}$$

From this equation, volume integral is influence of any source in space (source term) and surface integral is influence of boundary terms. The $g(\mathbf{x}, \mathbf{x}_0)$ is point source in free-field. The $\frac{\partial}{\partial n} g(\mathbf{x}, \mathbf{x}_0)$ is called dipole that is effect of distributed forces at boundary.

The time-domain free-field Green's function is obtained by inverse Fourier transform as follows:

$$g(\mathbf{x}, t | \mathbf{x}_0, t_0) = \frac{\delta(\tau - t_0)}{r} \tag{2.59}$$

where $r \triangleq |\mathbf{x} - \mathbf{x}_0|$, $\tau \triangleq t - \frac{r}{c}$ is retarded time and δ is Dirac's delta function. In addition, equation (2. 55) into Eq. (2. 58) to obtain time-domain Kirchhoff-Helmholtz integral by inverse Fourier transform as follows:

$$4\pi p(\mathbf{x}, t) = -\int_V \frac{f(\mathbf{x}_0, \tau)}{r} dV(\mathbf{x}_0) - \int_S \left[\mathbf{e}_R \cdot \mathbf{n} \left(\frac{1}{r^2} + \frac{1}{cr} \frac{\partial}{\partial t} \right) p(\mathbf{x}_0, \tau) - \frac{1}{r} \frac{\partial p}{\partial n}(\mathbf{x}_0, \tau) \right] dS(\mathbf{x}_0), \quad (2. 60)$$

where $\mathbf{e}_R \triangleq (\mathbf{x} - \mathbf{x}_0)/R$. First term of equation above, distributed source of sound field is more point source at \mathbf{x}_0 propagate to \mathbf{x} by sound speed c . New wave front is surface of envelope by each point source set after r/c time. This condition is called Huygen's principle, as shown in Fig. 6. For simple boundary geometries such as a plane or a sphere, homogeneous part $H(\mathbf{x}, \mathbf{x}_0)$ is added in the free-field Green's function

$$G(\mathbf{x}, \mathbf{x}_0) = g(\mathbf{x}, \mathbf{x}_0) + H(\mathbf{x}, \mathbf{x}_0) \quad (2. 61)$$

where $H(\mathbf{x}, \mathbf{x}_0)$ satisfies homogeneous Helmholtz equation

$$(\nabla^2 + k^2)H(\mathbf{x}, \mathbf{x}_0) = 0.$$

The $g(\mathbf{x}, \mathbf{x}_0)$ is singular when $\mathbf{x} = \mathbf{x}_0$. However, the $H(\mathbf{x}, \mathbf{x}_0)$ is non-singular when $\mathbf{x} = \mathbf{x}_0$. In this $G(\mathbf{x}, \mathbf{x}_0)$, that satisfies Kirchhoff-Helmholtz integral

$$4\pi p(\mathbf{x}) = -\int_V f(\mathbf{x}_0)G(\mathbf{x}, \mathbf{x}_0)dV(\mathbf{x}_0) - \int_S \left[p(\mathbf{x}_0) \frac{\partial}{\partial n} G(\mathbf{x}, \mathbf{x}_0) - G(\mathbf{x}, \mathbf{x}_0) \frac{\partial}{\partial n} p(\mathbf{x}_0) \right] dS(\mathbf{x}_0). \quad (2. 62)$$

In the Dirichlet problem, only $p(\mathbf{x})$ on the boundary is known. We can choose

$$H(\mathbf{x}, \mathbf{x}_0)|_{\mathbf{x}_0 \in S} = -g(\mathbf{x}, \mathbf{x}_0)|_{\mathbf{x}_0 \in S}. \quad (2. 63)$$

Hence, $G(\mathbf{x}, \mathbf{x}_0)$ satisfies

$$(\nabla^2 + k^2)G(\mathbf{x}, \mathbf{x}_0) = -4\pi\delta(\mathbf{x} - \mathbf{x}_0)$$

$$\text{st. } G(\mathbf{x}, \mathbf{x}_0)|_{\mathbf{x}_0 \in S} = 0 \quad (2.64)$$

Therefore, the Kirchhoff-Helmholtz integral can be rewritten as

$$4\pi p(\mathbf{x}) = -\int_V f(\mathbf{x}_0)G(\mathbf{x}, \mathbf{x}_0)dV(\mathbf{x}_0) - \int_S p(\mathbf{x}_0)\frac{\partial}{\partial n}G(\mathbf{x}, \mathbf{x}_0)dS(\mathbf{x}_0). \quad (2.65)$$

In addition, only $\frac{\partial p}{\partial n}(\mathbf{x}_0)$ on the boundary is known in Neumann problem. We can

choose

$$\frac{\partial}{\partial n}H(\mathbf{x}, \mathbf{x}_0)|_{\mathbf{x}_0 \in S} = -\frac{\partial}{\partial n}g(\mathbf{x}, \mathbf{x}_0)|_{\mathbf{x}_0 \in S} \quad (2.66)$$

Hence, $G(\mathbf{x}, \mathbf{x}_0)$ satisfies

$$(\nabla^2 + k^2)G(\mathbf{x}, \mathbf{x}_0) = -4\pi\delta(\mathbf{x} - \mathbf{x}_0) \quad (2.67)$$

$$\text{st. } \frac{\partial}{\partial n}G(\mathbf{x}, \mathbf{x}_0)|_{\mathbf{x}_0 \in S} = 0$$

Therefore, Kirchhoff-Helmholtz integral can be rewritten as

$$4\pi p(\mathbf{x}) = -\int_V f(\mathbf{x}_0)G(\mathbf{x}, \mathbf{x}_0)dV(\mathbf{x}_0) + \int_S G(\mathbf{x}, \mathbf{x}_0)\frac{\partial}{\partial n}p(\mathbf{x}_0)dS(\mathbf{x}_0). \quad (2.68)$$

That applies to close space interior sound field. Another problem is how to determine $H(\mathbf{x}, \mathbf{x}_0)$. We can use method of image to determine $H(\mathbf{x}, \mathbf{x}_0)$. In Fig. 7, a point source is arranged at \mathbf{x}_0 on an infinite reflection surface (rigidity). In free-field, the sound field $g(\mathbf{x}, \mathbf{x}_0)$ is produced by this point source. The sound field $H(\mathbf{x}, \mathbf{x}_0)$ is produced by image source, as shown in Fig. 7. The free-field

Green's function can be shown to be

$$G(\mathbf{x}, \mathbf{x}_0) = g(\mathbf{x}, \mathbf{x}_0) + H(\mathbf{x}, \mathbf{x}_0) = \frac{e^{-jkr}}{r} + \frac{e^{-jkr'}}{r'},$$

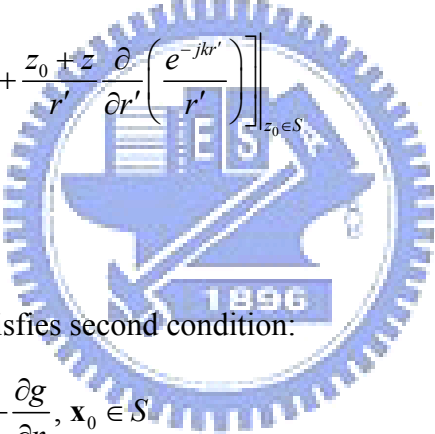
where

$$\begin{aligned}\mathbf{r} &= (x-x_0)\bar{i} + (y-y_0)\bar{j} + (z-z_0)\bar{k}, r = |\mathbf{r}| \\ \mathbf{r}' &= (x-x_0)\bar{i} + (y-y_0)\bar{j} + (z+z_0)\bar{k}, r' = |\mathbf{r}'|.\end{aligned}\quad (2.69)$$

Hence, $H(\mathbf{x}, \mathbf{x}_0)$ satisfies first condition: $(\nabla^2 + k^2)H = 0$. Then, r and r' are given in Eq. (2.69) and into

$$\begin{aligned}\frac{\partial}{\partial n} G(\mathbf{x}, \mathbf{x}_0) \Big|_{\mathbf{x}_0 \in S} &= \frac{\partial}{\partial z_0} \left(\frac{e^{-jkr}}{r} + \frac{e^{-jkr'}}{r'} \right) \Big|_{\mathbf{x}_0 \in S} \\ &= \left[\frac{\partial r}{\partial z_0} \frac{\partial}{\partial r} \left(\frac{e^{-jkr}}{r} \right) + \frac{\partial r'}{\partial z_0} \frac{\partial}{\partial r'} \left(\frac{e^{-jkr'}}{r'} \right) \right] \Big|_{\mathbf{x}_0 \in S}\end{aligned}\quad (2.70)$$

Equation (2.70) can be simplified as

$$\begin{aligned}\frac{\partial}{\partial n} G(\mathbf{x}, \mathbf{x}_0) \Big|_{\mathbf{x}_0 \in S} &= \left[\frac{z_0 - z}{r} \frac{\partial}{\partial r} \left(\frac{e^{-jkr}}{r} \right) + \frac{z_0 + z}{r'} \frac{\partial}{\partial r'} \left(\frac{e^{-jkr'}}{r'} \right) \right] \Big|_{\mathbf{x}_0 \in S} \\ &= \frac{-z}{r} \frac{e^{-jkr}}{r} + \frac{z}{r} \frac{e^{-jkr}}{r} \\ &= 0\end{aligned}$$


Therefore, $H(\mathbf{x}, \mathbf{x}_0)$ satisfies second condition:

$$\frac{\partial H}{\partial n} = \frac{\partial}{\partial n} (G - g) = -\frac{\partial g}{\partial n}, \mathbf{x}_0 \in S$$

Thus, the free-field Green's function can be expressed as

$$G(\mathbf{x}, \mathbf{x}_0) \Big|_{\mathbf{x}_0} = \frac{2e^{-jkr}}{r}\quad (2.71)$$

For a baffled rigid piston, this free-field Green's function is substituted into Kirchhoff-Helmholtz integral to yield

$$\begin{aligned}p(\mathbf{x}) &= j \frac{\rho_0 \omega}{4\pi} \int_S G(\mathbf{x}, \mathbf{x}_0) u(\mathbf{x}_0) dS(\mathbf{x}_0) \\ &= j \frac{\rho_0 c k u_0}{2\pi} \int_S \frac{e^{-jkr}}{r} dS,\end{aligned}\quad (2.72)$$

where $\omega = ck$ is temporal frequency. This equation called Rayleigh's integration.

In addition, boundary problem of open space exterior sound field, as shown in Fig. 8,

can be written as

$$4\pi p(\mathbf{x}) = -\int_V f(\mathbf{x}_0)G(\mathbf{x}, \mathbf{x}_0)dV(\mathbf{x}_0) - \int_S + \int_{S_\infty} \left[p(\mathbf{x}_0) \frac{\partial}{\partial n} G(\mathbf{x}, \mathbf{x}_0) - G(\mathbf{x}, \mathbf{x}_0) \frac{\partial}{\partial n} p(\mathbf{x}_0) \right] dS(\mathbf{x}_0).$$

In natural physics, it is zero to nearby influence in the boundary condition at infinite distance. Hence, mathematical condition is added

$$\int_{S_\infty} \left[(\mathbf{x}_0) \frac{\partial}{\partial n} G(\mathbf{x}, \mathbf{x}_0) - G(\mathbf{x}, \mathbf{x}_0) \frac{\partial}{\partial n} p(\mathbf{x}_0) \right] dS(\mathbf{x}_0) = 0.$$

Rewrite this equation by sphere integral

$$\int_{r \rightarrow \infty} \left(p \frac{\partial G}{\partial r} - G \frac{\partial p}{\partial r} \right) r^2 \sin \phi \, d\phi \, d\theta$$

or

$$2\pi \lim_{r \rightarrow \infty} r^2 \left(p \frac{\partial G}{\partial r} - G \frac{\partial p}{\partial r} \right) = 0.$$

Let $G = e^{-jkr/r}$ into equation above can obtain

$$\lim_{r \rightarrow \infty} r \left[p \left(jk + \frac{1}{r} \right) + \frac{\partial p}{\partial r} \right] = 0$$

or

$$\lim_{r \rightarrow \infty} r \left(jkp + \frac{\partial p}{\partial r} \right) = 0. \quad (2.73)$$

This equation is termed the Sommerfeld radiation condition. Therefore, problem of open space at infinite distance must satisfy not only $\lim_{r \rightarrow \infty} p = 0$ but also Eq. (2.73).

In addition, equation of momentum is employed to rewrite this radiation condition

$$\lim_{r \rightarrow \infty} r(\rho_0 c u_r - p) = 0,$$

where u_r is partial velocity on normal direction. Spherical wave at infinite distance in part become plane wave ($p = \rho_0 c u_r$). Finally, Kirchhoff-Helmholtz integration can be rewritten as

$$\alpha p(\mathbf{x}) = -\int_V f(\mathbf{x}_0)G(\mathbf{x}, \mathbf{x}_0)dV(\mathbf{x}_0) + \int_{S_\infty} \left[G(\mathbf{x}, \mathbf{x}_0) \frac{\partial p}{\partial n'} - p(\mathbf{x}_0) \frac{\partial p}{\partial n'} G(\mathbf{x}, \mathbf{x}_0) \right] dS(\mathbf{x}_0), \quad (2.74)$$

where \mathbf{n}' is outward normal at S and

$$\alpha = \begin{cases} 4\pi & \mathbf{x} \text{ is on } S \text{ outside} \\ 2\pi & \mathbf{x} \in S \\ 0 & \mathbf{x} \text{ is on } S \text{ inside} \end{cases}.$$

Usually, equation of boundary integral ($\alpha = 2\pi$) is employed to solve for unknown boundary condition and then calculate sound pressure at any field points ($\alpha = 4\pi$). This method is called the boundary element method (BEM). First, place the field point on the boundary ($\alpha = 2\pi$). The boundary integral equation can be Discretized in the matrix form

$$\mathbf{D}_s \mathbf{p}_s = \mathbf{M}_s \mathbf{v}_s, \quad (2.75)$$

where vector \mathbf{p}_s and \mathbf{v}_s is sound pressure and speed on boundary, and matrix \mathbf{M}_s and \mathbf{D}_s is monopole and dipole. Next, field point is moved to exterior field ($\alpha = 4\pi$). Discretization is used for equation of integral in the matrix form

$$\mathbf{p}_f = \mathbf{D}_f \mathbf{p}_s + \mathbf{M}_f \mathbf{v}_s, \quad (2.76)$$

where vector \mathbf{p}_f is sound pressure of exterior field, and matrix \mathbf{M}_f and \mathbf{D}_f is monopole and dipole of exterior field. Assume speed of boundary is known and the Eq. (2.75) can be written

$$\mathbf{p}_s = \mathbf{D}_s^{-1} \mathbf{M}_s \mathbf{v}_s = \mathbf{Z} \mathbf{v}_s, \quad (2.77)$$

where \mathbf{Z} is matrix of impedance (if \mathbf{D}_s^{-1} is existence). Matrix \mathbf{Z} is only involved in frequency and geometric configuration. Finally, sound pressure of exterior field at any field points is obtained

$$\mathbf{p}_f = (\mathbf{M}_f + \mathbf{D}_f \mathbf{Z}) \mathbf{v}_s = \mathbf{H} \mathbf{v}_s, \quad (2.78)$$

where \mathbf{H} is transfer matrix between speed of boundary and sound pressure of exterior field. The transfer matrix \mathbf{H} is dependent of frequency, geometric configuration and field point.

Numerical method of discrete Kirchhoff-Helmholtz integral is called direct boundary element method (DBEM). There are $\partial p / \partial n$ and p in integral of DBEM. Other method is use layer potential that only one term in integral. According to theory, sound field can be shown by simple layer potential, as shown in Fig. 9 (a)

$$p(\mathbf{x}) = \int_S \sigma(\mathbf{x}_0) G(\mathbf{x}, \mathbf{x}_0) dS(\mathbf{x}_0), \quad (2.79)$$

where G is the free-field Green's function and

$$\sigma = \frac{\partial p^+}{\partial n} - \frac{\partial p^-}{\partial n} \quad (2.80)$$

is the unknown strength of the monopole. In this method, sound pressure p is continuous at the boundary, whereas $\frac{\partial p}{\partial n}$ is discontinuous at the boundary.

$$\frac{\partial p}{\partial n}(\mathbf{x}) = \alpha \sigma(\mathbf{x}) + \int_S \sigma(\mathbf{x}_0) \frac{\partial G}{\partial n}(\mathbf{x}, \mathbf{x}_0) dS(\mathbf{x}_0)$$

$$\alpha = \begin{cases} \frac{1}{2} & \mathbf{x} \in V_0 \\ -\frac{1}{2} & \mathbf{x} \in V_i \\ 0 & \text{otherwise} \end{cases} \quad (2.81)$$

Or, sound field can be represented by double layer potential, as shown in Fig. 9 (b)

$$p(\mathbf{x}) = \int_S \mu(\mathbf{x}_0) \frac{\partial G}{\partial n}(\mathbf{x}, \mathbf{x}_0) dS(\mathbf{x}_0), \quad (2.82)$$

where $\mu = p^- - p^+$ is unknown strength of dipole. In this method, $\frac{\partial p}{\partial n}$ is continuous at inside and outside boundary, whereas sound pressure p is

discontinuous at the boundary.

$$p(\mathbf{x}) = \alpha\sigma(\mathbf{x}) + \int_S \mu(\mathbf{x}_0) \frac{\partial G}{\partial n}(\mathbf{x}, \mathbf{x}_0) dS(\mathbf{x}_0)$$

$$\alpha = \begin{cases} \frac{1}{2} & \mathbf{x} \in V_0 \\ -\frac{1}{2} & \mathbf{x} \in V_i \\ 0 & \text{otherwise} \end{cases} \quad (2.83)$$

That discrete layer potential of BEM is called indirect BEM. Finally, equivalent source distribution is introduced, as shown in Fig. 9 (c)

$$p(\mathbf{x}) = \int_S \sigma(\mathbf{x}_0) G(\mathbf{x}, \mathbf{x}_0) dS(\mathbf{x}_0)$$

$$\approx \sum_{n=1}^N \int_{S_n} \sigma(\mathbf{x}_0) G(\mathbf{x}, \mathbf{x}_0) dS(\mathbf{x}_0) \approx \sum_{n=1}^N \sigma(\mathbf{x}_n) G(\mathbf{x}, \mathbf{x}_n) S_n \quad (2.84)$$

$$= \sum_{n=1}^N [\sigma(\mathbf{x}_n) S_n] G(\mathbf{x}, \mathbf{x}_n) = \sum_{n=1}^N Q_n G(\mathbf{x}, \mathbf{x}_n)$$

where Q is source strength. This equation approximates continuous integral as $n \rightarrow \infty$. Virtual source representation is simpler than BEM. Effect of virtual source representation has not bad effect in some application.

Next, monopole, dipole and multipole expansion of basic source model are described. Dipole is combined by two equal amplitude but 180° out-of-phase point sources, as shown in Fig. 10 (a). Let

$$g(\mathbf{x}, \mathbf{x}_s) = \frac{e^{-jkR}}{R},$$

where $R = |\mathbf{x} - \mathbf{x}_s|$. Sound pressure of dipole at \mathbf{x} is

$$p(\mathbf{x}) = Ag(\mathbf{x}, \mathbf{x}_s + \mathbf{d}/2) - Ag(\mathbf{x}, \mathbf{x}_s - \mathbf{d}/2)$$

$$= A|\mathbf{d}| \frac{\Delta g(\mathbf{x}, \mathbf{x}_s)}{|\mathbf{d}|}.$$

When $|\mathbf{d}| \rightarrow 0$, the sound pressure can be written as

$$p(\mathbf{x}) \approx \mathbf{D} \cdot \nabla_s g(\mathbf{x}, \mathbf{x}_s) = -\mathbf{D} \cdot \nabla g(\mathbf{x}, \mathbf{x}_s) = -\mathbf{D} \cdot \left(\frac{\mathbf{x} - \mathbf{x}_s}{R} \right) \frac{d}{dR} \left(\frac{e^{-jkR}}{R} \right), \quad (2.85)$$

where ∇_s and ∇ is gradient of \mathbf{x}_s and \mathbf{x} respectively and $\mathbf{D} = A\mathbf{d}$ is called dipole moment. According to equation above, the \mathbf{d} is aimed at y-axis, as shown in Fig. 10 (b). Sound pressure is given as

$$p(\mathbf{x}) = Ad \cos \theta \left(jk + \frac{1}{R} \right) \frac{e^{-jkR}}{R} \\ \approx jkAd \cos \theta \frac{e^{-jkR}}{R} \quad (\text{when } R \gg d).$$

Sound field has maximum value on y-axis, as shown in Fig. 10 (c). The $\frac{\partial g}{\partial n}$

represents a dipole in Kirchhoff-Helmholtz integral as

$$\frac{\partial g}{\partial n} = \cos \theta \left(jk + \frac{1}{R} \right) g.$$

Next, a quadrupole is composed of two dipoles. Sound pressure as shown in Fig. 10 (d) is given as

$$p(\mathbf{x}) = (\mathbf{D} \cdot \nabla)(\mathbf{d} \cdot \nabla) \frac{e^{-jkR}}{R}. \quad (2.86)$$

Or, in Cartesian tensor

$$p(\mathbf{x}) = Q_{\mu\nu} \frac{\partial^2}{\partial x_\mu \partial x_\nu} \frac{e^{-jkR}}{R}, \quad \mu \text{ and } \nu = 1, 2, 3, \quad (2.87)$$

where $Q_{\mu\nu} \triangleq D_\mu d_\nu$ is called quadrupole moment.

In complex sound source, sound pressure at \mathbf{x} in farfield as shown in Fig. 11 is given as

$$p(\mathbf{x}) = \sum_{n=1}^N A_n \frac{e^{-jkR_n}}{R_n}$$

where $R_n \triangleq |\mathbf{x} - \mathbf{x}_n|$ and \mathbf{x}_n is the position vector of the n th point source.

Multipole expansion is obtained from Eq. (2.87) by Taylor's expansion

$$f(\mathbf{x} - \mathbf{x}_n) = f(\mathbf{x}) - (\mathbf{x}_n \cdot \nabla) f(\mathbf{x}) + \frac{1}{2!} (\mathbf{x}_n \cdot \nabla)^2 f(\mathbf{x}) - \dots$$

Therefore, sound pressure can be expanded as

$$p(\mathbf{x}) = \sum_{n=1}^N A_n \left[1 - (\mathbf{x}_n \cdot \nabla) + \frac{1}{2!} (\mathbf{x}_n \cdot \nabla)^2 - \dots \right] \frac{e^{-jkr}}{r}$$

$$= \left(S - \mathbf{D} \cdot \nabla + \sum_{\mu, \nu} Q_{\mu\nu} \frac{\partial^2}{\partial x_\mu \partial x_\nu} - \dots \right) \frac{e^{-jkr}}{r}, \quad (2.88)$$

where $S \triangleq \sum_{n=1}^N S_n$, $\mathbf{D} \triangleq \sum_{n=1}^N \mathbf{x}_n S_n$ and $Q_{\mu\nu} \triangleq \frac{1}{2!} \sum_{n=1}^N x_{n\mu} x_{n\nu} S_n$.



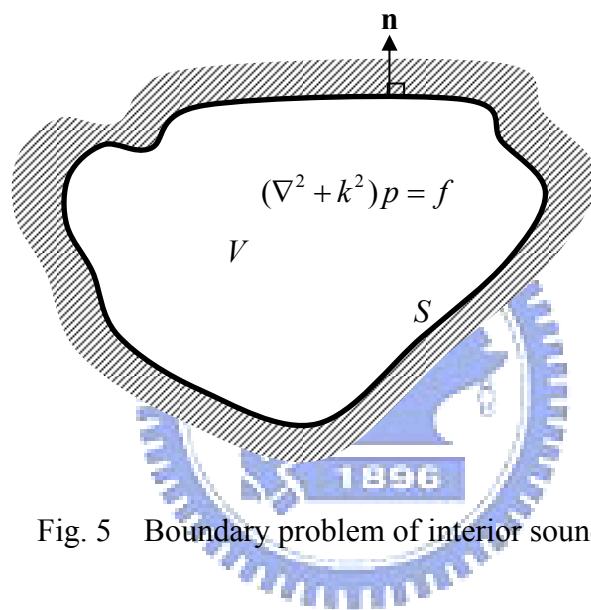


Fig. 5 Boundary problem of interior sound field.

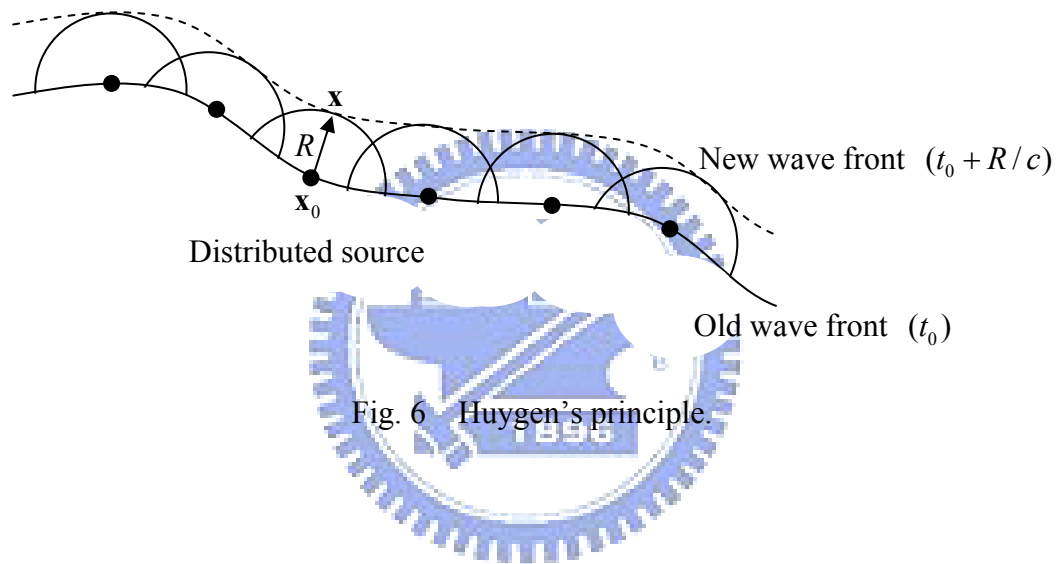


Fig. 6 Huygen's principle.

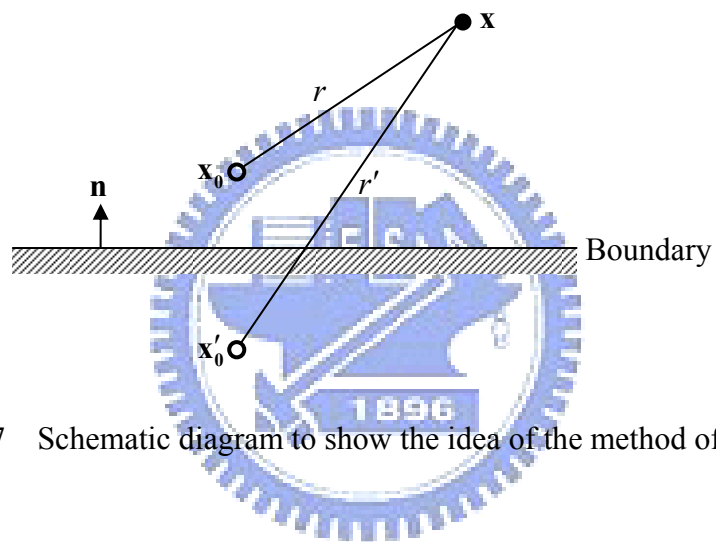


Fig. 7 Schematic diagram to show the idea of the method of image.

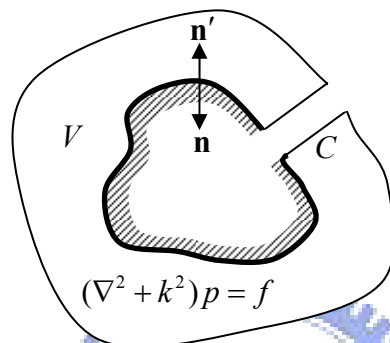
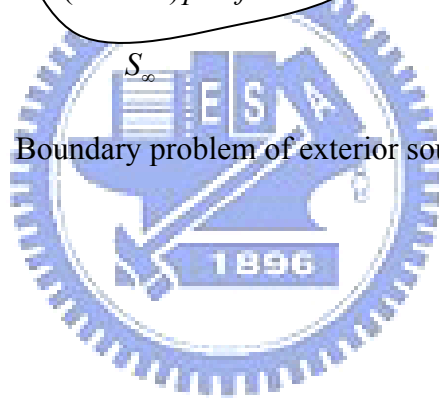


Fig. 8 Boundary problem of exterior sound field.



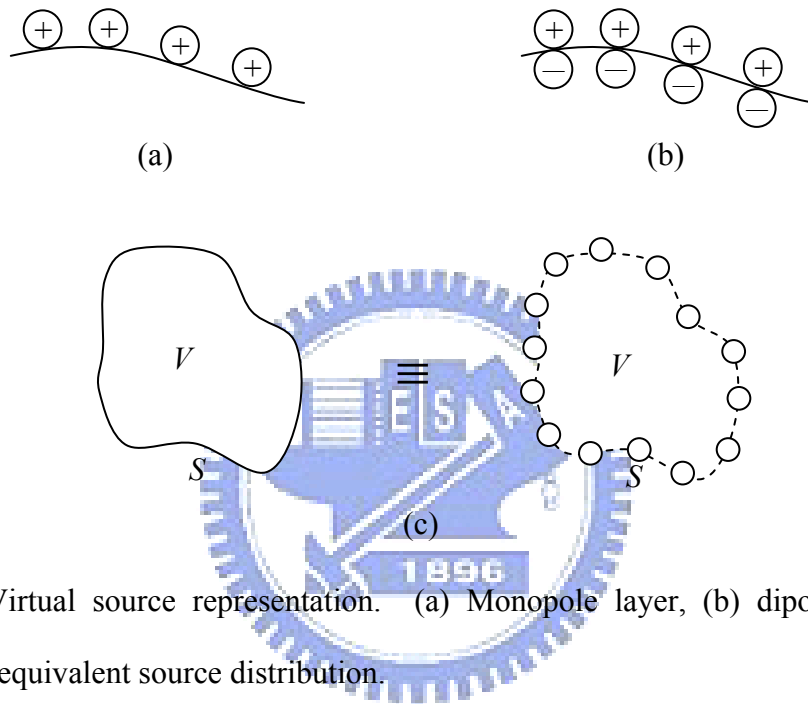


Fig. 9 Virtual source representation. (a) Monopole layer, (b) dipole layer, (c) equivalent source distribution.

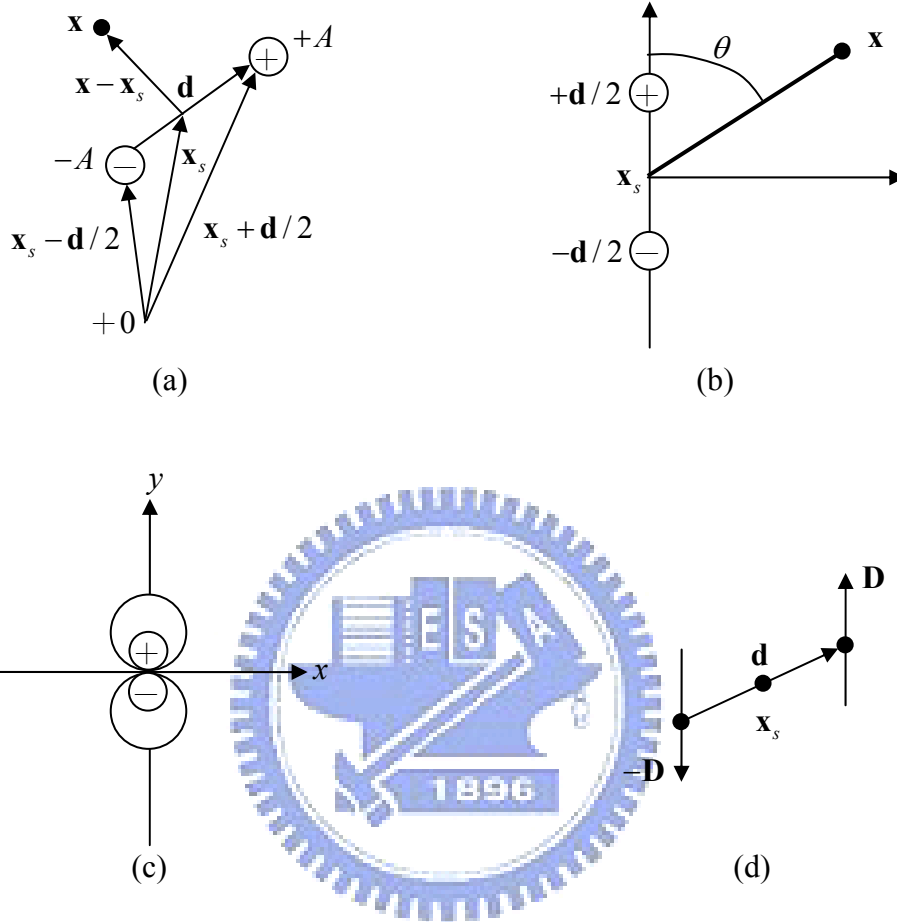


Fig. 10 Dipole and quadrupole. (a) Coordinate figure of dipole, (b) when \mathbf{d} is aimed at y -axis (dipole), (c) sound pressure distribution of dipole, (d) general structure of quadrupole.

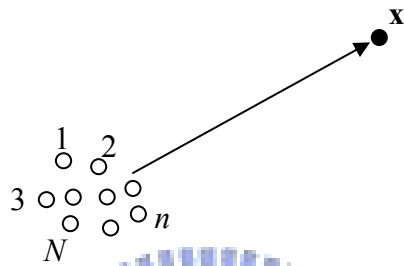


Fig. 11 Illustration for multipole expansion.



2.4 Inverse problems and ill-posedness

In general, inverse problems are derived from their associated forward problem. The forward problem of NAH is a sound field model that describes the acoustic radiation from a source. The forward problem of NAH can be expressed as

$$\mathbf{p} = \mathbf{G}\mathbf{q}, \quad (2.89)$$

where matrix \mathbf{G} contains the model parameter such as propagation matrix, vector \mathbf{q} is a input such as source strength or particle velocity on the source, and vector \mathbf{p} is the resulting sound pressure at field points. Vector \mathbf{q} is the unknown and \mathbf{p} is the measured quantity. Hence the inverse problem is to be solved. The modeled sound field by \mathbf{G} includes both propagating wave components and wave components that decay in the direction away from the source such as evanescent waves. In fact, there are no real evanescent waves in the nearfield of a finite size radiator. The wave mode decomposition describes part of the radiation as evanescent waves, but this is more a question of representation (cut off modes in a duct, on the other hand, are truly evanescent). In other words, it is merely reflection of the convolution/smoothing process as waves propagate to the farfield. The exact quantification is more tied to the decomposition (such as Fourier transform, SVD, etc.) one uses. The evanescent waves consist of spatial frequency components higher than the propagating waves. The decay rate of the evanescent waves generally increases with the spatial frequencies. Contribution of evanescent wave component is decay when moving away from the source. In this condition, contribution from the evanescent components can be very low on the measurement surface. Noise components in the measurement data that contain high spatial frequencies will be treated as evanescent waves. Unwanted amplifications is occurred in the reconstruction when inverting the matrix \mathbf{G} . In other words, small perturbations in \mathbf{p} may have a huge effect on the solution vector \mathbf{q} , that is said ill-posedness. Regularization is needed when

solving the inverse problem. Regularization is employed to limit the influence of the measurement noise.



CHAPTER 3. THEORETICAL PRELIMINARIES OF SIGNAL PROCESSING

3.1 Linear algebra basics

In general, inverse problems are based on an associated forward problem that describes the unknown effect on the basis of the known cause. In light of NAH, the acoustic radiation problem can always be formulated into the following matrix form:

$$\mathbf{G}\mathbf{q} = \mathbf{p}, \quad (3.1)$$

where \mathbf{p} and \mathbf{q} are the hologram data and source data, respectively, which are related by the propagation matrix \mathbf{G} . It can be shown for some perturbations $\delta\mathbf{p}$ (measurement noise, numerical error, etc.) that the perturbation $\delta\mathbf{q}$ satisfies

$$\frac{\|\delta\mathbf{q}\|}{\|\mathbf{q}\|} \leq \text{cond}(\mathbf{G}) \frac{\|\delta\mathbf{p}\|}{\|\mathbf{p}\|}, \quad (3.2)$$

where $\text{cond}(\mathbf{G}) = \sigma_{\max} / \sigma_{\min}$ is the condition number of the matrix \mathbf{G} provided 2-norm is assumed.

The SVD is useful in order to understand the mechanisms behind the inverse problem. The SVD of the matrix \mathbf{G} is given as

$$\mathbf{G} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H, \quad (3.3)$$

where \mathbf{U} is the left singular vectors, $\mathbf{\Sigma}$ is a diagonal matrix composed of singular values of the matrix \mathbf{G} and \mathbf{V} is the right singular vectors. As a theory of the orthonormality of the matrices, \mathbf{U} and \mathbf{V} are unitary matrices as $\mathbf{U}^H\mathbf{U} = \mathbf{U}\mathbf{U}^H = \mathbf{I}_U$ and $\mathbf{V}^H\mathbf{V} = \mathbf{V}\mathbf{V}^H = \mathbf{I}_V$, where \mathbf{I}_U and \mathbf{I}_V are identity matrices.

Vector \mathbf{q} in Eq. (3.1) can be found as

$$\mathbf{q} = \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^H\mathbf{p} = \sum_{i=1}^n \frac{\mathbf{u}_i^H\mathbf{p}}{\sigma_i} \mathbf{v}_i, \quad (3.4)$$

where n is the smallest dimension of \mathbf{G} , \mathbf{u}_i is the i th left singular vector, \mathbf{v}_i is

the i th right singular vector and σ_i is the i th singular value. However, the measure data \mathbf{p} is perturbed by noise ($\mathbf{p} = \tilde{\mathbf{p}} + \mathbf{e}$), where $\tilde{\mathbf{p}}$ is the true pressure and \mathbf{e} is the noise. Hence, equation (3. 4) can be rewritten as

$$\mathbf{q} = \sum_{i=1}^n \frac{\mathbf{u}_i^H \tilde{\mathbf{p}}}{\sigma_i} \mathbf{v}_i + \sum_{i=1}^n \frac{\mathbf{u}_i^H \mathbf{e}}{\sigma_i} \mathbf{v}_i . \quad (3. 5)$$

Because the singular vectors with high index correspond to evanescent wave components and the magnitude of $\mathbf{u}_i^H \tilde{\mathbf{p}}$ reaches the noise level, the magnitude of the $\mathbf{u}_i^H \tilde{\mathbf{p}}$ will decrease when the index i increases. The second term of Eq. (3. 5) will dominate the solution for the high indices of i (corresponding to small singular values). Equation (3. 5) will cause huge errors in the reconstruction result when taking the reciprocal of σ_i the noise term. Regularization is required in order to prevent these unwanted amplifications of the noise. In solving the inverse problem, the regularization is employed to avoid the amplification of the noise. Hence the solution vector can be regularized in Eq. (3. 4). Regularized solution is given as

$$\mathbf{q}_{\text{reg}} = \sum_{i=1}^n f_i \frac{\mathbf{u}_i^H \mathbf{p}}{\sigma_i} \mathbf{v}_i , \quad (3. 6)$$

where the filter factors f_i constitute a low-pass filter. The filter factors is allowing components with spatial frequencies (in low indices) in the solution and damping the high frequencies (in high indices). The filter coefficients is to be $f_i = 1$ for $i \leq I$ and $f_i = 0$ for $i > I$, where I is the discrete regularization parameter. This regularization parameter I determines how much the solution is regularized in Eq.(3. 6). This method of regularization is called TSVD that can be written as

$$\mathbf{q}_I = \sum_{i=1}^I \frac{\mathbf{u}_i^H \mathbf{p}}{\sigma_i} \mathbf{v}_i . \quad (3. 7)$$

The solution \mathbf{q}_I will be over-smoothed when I is chosen too low due to the removal of the high frequency components. If I is chosen too high, this solution \mathbf{q}_I has distortion because of the amplification of the noise. Hence, how to choose

optimal regularization parameter is a problem. This problem will be described in chapter 6.

In mathematical optimization, the method of Lagrange multipliers (named after Joseph Louis Lagrange 1736~1813) provides a strategy for finding the maxima and minima of a function subject to constraints. Assume we want to optimize (find the minimum and maximum) of a function, $f(x, y)$, subject to the constraint $g(x, y) = K$. Solve the system of equation as follows

$$\nabla f(x, y) = \lambda_1 \nabla g(x, y) \quad (3.8)$$

$$g(x, y) = K, \quad (3.9)$$

where $\lambda_1 \neq 0$ is Lagrange multiplier that is constant and K is constant of constraint.

The system of equations actually has three equations; the system can be rewritten in a simpler form. The first equation put in the definition of the gradient vector to see what we get

$$(f_x, f_y) = \lambda_1 (g_x, g_y) = (\lambda_1 g_x, \lambda_1 g_y). \quad (3.10)$$

In order for these two vectors to be equal the individual components must be equal.

Hence, two equations are

$$f_x = \lambda_1 g_x \quad (3.11)$$

and

$$f_y = \lambda_1 g_y. \quad (3.12)$$

These two equations along with the constraint, $g(x, y) = K$, give three equations with three unknowns x , y and λ_1 . For example, find the maximum and minimum of $f(x, y) = 5x - 3y$ subject to the constraint $x^2 + y^2 = 136$. Hence, we need to solve

$$\begin{aligned} 5 &= 2\lambda_1 x \\ -3 &= 2\lambda_1 y \\ x^2 + y^2 &= 136. \end{aligned}$$

The variables x and y can be expressed in terms of the Lagrange multipliers

$$\begin{aligned} x &= \frac{5}{2\lambda_1} \\ y &= \frac{-3}{2\lambda_1}. \end{aligned}$$

Plugging x and y into the constraint gives

$$\frac{25}{4\lambda_1^2} + \frac{9}{4\lambda_1^2} = \frac{17}{2\lambda_1^2} = 136.$$

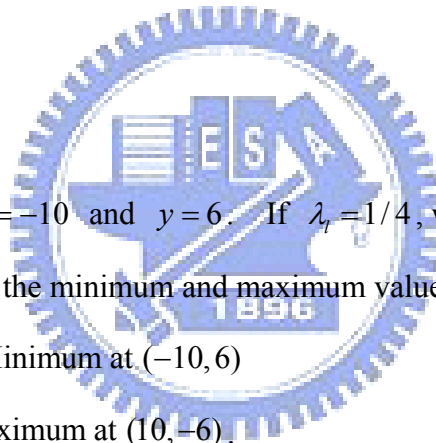
The Lagrange multiplier λ_1 can be solved

$$\begin{aligned} \lambda_1^2 &= \frac{1}{16} \\ \Rightarrow \lambda_1 &= \pm \frac{1}{4}. \end{aligned}$$

If $\lambda_1 = -1/4$, we get $x = -10$ and $y = 6$. If $\lambda_1 = 1/4$, we get $x = 10$ and $y = -6$. Thus, here are the minimum and maximum values of the function as

$$f(-10, 6) = -68, \text{ Minimum at } (-10, 6)$$

$$f(10, -6) = 68, \text{ Maximum at } (10, -6).$$



3.2 Digital signal processing basics

Under-sampling a continuous-time signal can introduce a distortion into the signal, which is generally referred to as “aliasing” (a relatively high-frequency component in the analog input signal appears at a lower frequency in the reconstructed output signal). Nyquist showed that sampling rate must exceed two-times of the cutoff frequency of the bandlimited signal to fully recover the continuous signal. Let $x_c(t)$ be continuous and bandlimited signal with

$$X_c(j\Omega) = 0 \quad \text{for } |\Omega| \geq \Omega_N, \quad (3.13)$$

where Ω_N is the cutoff frequency and also referred to as the Nyquist frequency.

$x_c(t)$ is uniquely determined by its samples $x[N] = x_c(nT), n = 0, \pm 1, \pm 2, \dots$, if

$$\Omega_s = \frac{2\pi}{T} \geq 2\Omega_N, \quad (3.14)$$

where Ω_s is the sampling rate and T is the period of signal. Also oversampling may be desirable because interpolation during signal alignment can degrade resolution. Oversampling is also required if you intend to use deconvolution to increase resolution; the final deconvolved signal must also meet the Nyquist Sampling Criterion.

Many of the phenomena studied in engineering and science are periodic in nature eg. the sound signal in an alternating current circuit. These periodic functions can be analysed into their constituent components (fundamentals and harmonics) by a process called Fourier analysis. We are aiming to find an approximation using trigonometric functions for various square, saw tooth, etc waveforms that occur in electronics. We do this by adding more and more trigonometric functions together. The sum of these special trigonometric functions is called the Fourier series. The Fourier series was shown as a way of representing the spectrum of a periodic signal as a series of discrete lines in the frequency-domain. In this module, the concept of the Fourier transform will be introduced and applied to nonperiodic functions. The Fourier transform defines a relationship between a signal in the time-domain and its representation in the frequency-domain. Summary of Fourier series and transform relations is shown in Table 1.

A transfer function (TF) is a mathematical representation, in terms of spatial or temporal frequency, of the relation between the input and output of a linear time-invariant system. With optical imaging devices, for example, it is the Fourier

transform of the point spread function. Let $x(t)$ is the input and $y(t)$ is the output. Take the bilateral Laplace transform of $x(t)$ and $y(t)$

$$X(s) = \int_{-\infty}^{\infty} x(t)e^{-st} dt \quad (3.15)$$

$$Y(s) = \int_{-\infty}^{\infty} y(t)e^{-st} dt. \quad (3.16)$$

The output is related to the input by the TF $H(S)$ as

$$Y(s) = H(s)X(s) \quad (3.17)$$

and the TF is

$$H(s) = \frac{Y(s)}{X(s)}. \quad (3.18)$$

If a complex harmonic signal with a sinusoidal component with amplitude $|X|$, angular frequency ω and phase $\arg(X)$

$$x(t) = Xe^{j\omega t} = |X|e^{j(\omega t + \arg(X))}, \quad (3.19)$$

where $X = |X|e^{j\arg(X)}$ is input to a linear time-invariant system. Frequency response function (FRF) is the measure of any system output spectrum in response to an input signal. The frequency response $H(j\omega)$ describes this change for every frequency ω in terms of gain:

$$|H(j\omega)| = \frac{|Y|}{|X|} \quad (3.20)$$

and phase shift:

$$\phi(\omega) = \arg(Y) - \arg(X) = \arg(H(j\omega)). \quad (3.21)$$

The TF can also be shown using the Fourier transform which is only a special case of the bilateral Laplace transform for the case where $s = j\omega$.

The impulse response, or impulse response function (IRF), of a dynamic system is its output when presented with a brief input signal, called an impulse. The Laplace transform of the IRF is known as the TF. It is usually easier to analyze

systems using TFs as opposed to IRFs. The Laplace transform of a system output may be determined by the multiplication of the TF with the input function in the complex plane, also known as the frequency domain. An inverse Laplace transform of this result will yield the output function in the time domain. In array applications, impulse responses enable the acoustic characteristics of a location.

Infinite impulse response (IIR) filter is a type of a signal processing filter whose impulse response is of infinite length of time. This is in contrast to finite impulse response (FIR) filters, which have fixed-duration impulse responses. Direct-form IIR filters are often described and implemented in terms of the difference equation that defines how the output signal is related to the input signal:

$$y(n) = \sum_{i=1}^{\Lambda} b_i x(n-i) + \sum_{i'=0}^{\Lambda'} a_{i'} y(n-i'). \quad (3.22)$$

where $x(n)$ is the input signal, $y(n)$ is the output signal, Λ is the feedforward filter order, b_i are the feedforward filter coefficients, Λ' is the feedback filter order and $a_{i'}$ are the feedback filter coefficients. The direct-form I structure is shown in Fig. 12 and the direct-form II structure is shown in Fig. 13. The corresponding rational system function is

$$H(z) = \frac{\sum_{i=1}^{\Lambda} b_i z^{-i}}{1 - \sum_{i'=0}^{\Lambda'} a_{i'} z^{-i'}}. \quad (3.23)$$

A FIR filter is a type of a signal processing filter whose impulse response is of finite duration, because it settles to zero in finite time. Direct-form FIR filters are described by the following difference equation, which defines the output $y(n)$ in terms of its input $x(n)$:

$$y(n) = \sum_i^{\Lambda} b_i x(n-i), \quad (3.24)$$

where $x(n)$ is the input signal, $y(n)$ is the output signal, b_i are the filter coefficients and Λ is the filter order. This can be recognized as the discrete convolution of $x(n)$ with the impulse response

$$h(n) = \begin{cases} b_i & i = 0, 1, \dots, \Lambda, \\ 0 & \text{otherwise.} \end{cases} \quad (3.25)$$

The direct-form structure is shown in Fig. 14.

Convolution is the mathematical process that relates the output $y(t)$ of a linear, time-invariant system to its input $x(t)$, and impulse response $h(t)$. The output is

$$y(t) = x(t) * h(t) = h(t) * x(t), \quad (3.26)$$

where “*” represents the commutative convolution operation. For two finite discrete sequences of length N_x and N_h , the linear or aperiodic convolution sum takes on a slightly different form

$$y(n) = \sum_{k_n}^n h(k_n)x(n-k_n), \quad (3.27)$$

where n is sample index, $h(k_n)$ and $x(n-k_n)$ are zero outside their appropriately defined intervals. For $N_x > N_h$, each summation need only be calculated for the $0 \leq k_n \leq N_h - 1$ terms. The output $y(n)$ will have length $N_x + N_h - 1$. In frequency domain convolution, multiplication in the frequency domain translates to *circular* convolution in the time domain in the discrete case. The output is

$$y(n) = \text{IFFT}(Y(n_f)) = \text{IFFT}(X(n_f)H(n_f)), \quad (3.28)$$

where n_f is the discrete frequency variable and IFFT is inverse FFT. In the Block convolution using the overlap-add method (OAD) as shown in Fig. 15, the input blocks need not be precisely N_h samples long. But it is generally a good idea to keep $N_{x\Lambda}$ on the order of N_h to avoid unnecessarily long block convolutions. The overlap between block outputs must remain N_{h-1} , regardless. Mathematically, $x(n)$ and $y(n)$ can be represented as

$$x(n) = \sum_{i=0}^{\infty} x_{\Lambda}(n), \text{ where } x_{\Lambda}(n) = x(n) \text{ for } N_{\text{block}} \leq n \leq (\Lambda + 1)N_{\text{block}} \quad (3.29)$$

$$x(n) = 0 \text{ otherwise}$$

$$y(n) = h(n) * x(n)$$

$$y(n) = \sum_{n_k}^n h(n_f) \sum_{i=1}^{\infty} x_{\Lambda}(n - n_f) = \sum_{i=1}^{\infty} h(n) * x_{\Lambda}(n) = \sum_{i=0}^{\infty} y_{\Lambda}(n), \quad (3.30)$$

where $\Lambda = 1, 2, \dots, \hat{\Lambda}$ is the frame index and $\hat{\Lambda}$ is the number of frames. Unlike the OAD, the overlap-save method (OAS) requires that the *input* blocks overlap as shown in Fig. 16. Then the input blocks are *circularly* convolved with the impulse response. Because of the overlap redundancy at the input, the circular artifacts in the output (the first $N_h - 1$ samples) can simply be discarded. Mathematically, symbolic representations of $x(n)$ and $y(n)$ are rather cumbersome, but can be expressed as

$$x(n) = \sum_{\Lambda=0}^{\infty} x_{\Lambda}(n) - x_{\Lambda-1}(n_f) \quad (3.31)$$

where $x_{\Lambda}(n) = x(n)$ for $N_{\text{block}} - (\Lambda + 1)(N_h - 1) \leq n \leq (\Lambda + 1)N_{\text{block}} - (\Lambda + 1)(N_h - 1) - 1$ and $x_{\Lambda}(n) = 0$ otherwise. Also, $x_{\Lambda-1}(n_f) = x_{\Lambda-1}(n)$ for

$$N_{\text{block}} - (\Lambda + 1)(N_h - 1) \leq n_f \leq \Lambda N_{\text{block}} - \Lambda(N_h - 1) - 1.$$

$$y_{\Lambda}(n) = h(n) (*) x_{\Lambda}(n), \quad (3.32)$$

where (*) denotes circular convolution. The output is

$$y(n) = y_0(n_f) | y_1(n_f) | y_2(n_f) | \dots, \quad (3.33)$$

where " | " denotes concatenation and m indexes the last $N_{\text{block}} - (N_h - 1)$ samples of each block.

In digital signal processing, time delay is common problem. The delay usually is not an integer in digital signal processing. There are many ways to deal with these fractional delay problems. The simplest approach is Lagrange interpolation method. Firstly we divide τ_m by sampling period T to acquire the fractional delay Ψ_m .

The delay is separated into two parts

$$\frac{\tau_m}{T} = \Psi_m = D_m + e_m, \quad (3.34)$$

where D_m and e_m are the integer and fractional component of Ψ_m , respectively.

FIR filter coefficients to implement the Lagrange interpolation can be calculated by

$$w_{mk} = \prod_{\substack{l=0 \\ l \neq k}}^N \frac{e_m - l}{k - l}, \quad k = 0, 1, 2, \dots, N. \quad (3.35)$$

The coefficients for the Lagrange filters of order $N = 1, 2$ are given in the Table 2.

The case $N = 1$ corresponds to the linear interpolation using two samples.

In signal processing, the Wiener filter is a filter. Its purpose is to reduce the amount of noise present in a signal by comparison with an estimation of the desired noiseless signal. A Wiener filter is not an adaptive filter because the theory behind this filter assumes that the inputs are stationary.

The input to the Wiener filter is assumed to be a signal $x(t)$ corrupted by additive noise $v(t)$. The output $y(t)$ is calculated by means of a filter $h(t)$ as

$$y(t) = h(t) * [x(t) + v(t)], \quad (3.36)$$

where $x(t)$ is the original signal, $v(t)$ is the noise, $y(t)$ is the estimated signal and $h(t)$ is the Wiener filter's impulse response. The error is defined as

$$e(t) = x(t + \tau_w) - y(t), \quad (3.37)$$

where τ_w is the delay of the Wiener filter. Writing $y(t)$ as a convolution integral:

$$y(t) = \int_{-\infty}^{\infty} h(\tau) [x(t - \tau) + v(t - \tau)] d\tau, \quad (3.38)$$

where τ is time delay. Taking the expected value of the squared error results in

$$\begin{aligned} E\{e^2\} &= R_x(0) - 2 \int_{-\infty}^{\infty} h(\tau) R_{\hat{x}x}(\tau + \tau_w) d\tau \\ &\quad + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(\tau) h(\theta) R_x(\tau - \theta) d\tau d\theta, \end{aligned} \quad (3.39)$$

where $\hat{x}(t) = x(t) + v(t)$ is the observed signal, R_x is the autocorrelation function of $x(t)$, $R_{\hat{x}}$ is the autocorrelation function of $\hat{x}(t)$ and $R_{\hat{x}x}$ is the cross-correlation

function of $\hat{x}(t)$ and $x(t)$. The goal is to minimize $E\{e^2\}$ by finding the optimal $h(t)$ that is the Wiener filter IRF.

The Wiener filter problem has solutions for three possible cases. Firstly, non-causal filter is acceptable. Next, causal filter is desired. Finally, finite amount of past data is used. The first case is simple to solve but is not suited for real-time applications. The non-causal solution is

$$H(s) = \frac{X_{\hat{x},x}(s)}{X_x(s)} e^{\tau_w s} . \quad (3.40)$$

Provided that $h(t)$ is optimal, then the minimum mean-square error equation reduces to

$$E\{e^2\} = R_x(0) - \int_{-\infty}^{\infty} h(\tau) R_{\hat{x},x}(\tau + \tau_w) d\tau , \quad (3.41)$$

and the solution $h(t)$ is the inverse two-sided Laplace transform of $H(s)$. The causal solution is

$$H(s) = \frac{G(s)}{X_x^+(s)} , \quad (3.42)$$

where $G(s)$ consists of the causal part of $e^{\tau_w s} X_{\hat{x},x}(s)/X_x(s)$ and $X_x^+(s)$ is the causal component of $X_x(s)$.

The causal FIR Wiener filter, instead of using some given data matrix \mathbf{X} and output vector \mathbf{Y} , finds optimal tap weights by using the statistics of the input and output signals. It populates the input matrix \mathbf{X} with estimates of the auto-correlation of the input signal \mathbf{T} and populates the output vector \mathbf{Y} with estimates of the cross-correlation between the output and input signals \mathbf{V} .

In order to derive the coefficients of the Wiener filter, we consider a signal $w(n)$ being fed to a Wiener filter of order Λ and with coefficients $a_i, i = 0, \dots, \Lambda$. The output of the filter is denoted $x(n)$ which is given by the expression

$$x(n) = \sum_{i=0}^{\Lambda} a_i w(n-i). \quad (3.43)$$

The residual error is denoted $e(n)$ and is defined as $e(n) = x(n) - v(n)$. The Wiener filter is designed so as to minimize the mean square error (MMSE) which can be stated concisely as follows:

$$a_i = \arg \min E\{e^2(n)\}, \quad (3.44)$$

where $e\{\cdot\}$ denote the expectation operator. In the general case, the coefficients a_i may be complex and may be derived for the case where $w(n)$ and $v(n)$ are complex as well. For simplicity, we will only consider the case where all these quantities are real. The mean square error may be rewritten as:

$$\begin{aligned} E\{e^2(n)\} &= E\{[x(n) - v(n)]^2\} \\ &= E\{x^2(n)\} - 2E\{x(n)v(n)\} + E\{v^2(n)\} \\ &= E\left\{\left[\sum_{i=0}^{\Lambda} a_i w(n-i)\right]^2\right\} - 2E\left\{\sum_{i=0}^{\Lambda} a_i w(n-i)v(n)\right\} + E\{v^2(n)\}. \end{aligned} \quad (3.45)$$

If we suppose that $w(n)$ and $v(n)$ are each stationary and jointly stationary, we can introduce the following sequences $R_w(n')$ and $R_{wv}(n')$ known respectively as the autocorrelation of $w(n)$ and the cross-correlation between $w(n)$ and $v(n)$ defined as follows:

$$\begin{aligned} R_w(n') &= E\{w(n)w(n+n')\} \\ R_{wv}(n') &= E\{w(n)v(n+n')\}. \end{aligned} \quad (3.46)$$

The derivative of the MSE may therefore be rewritten as

$$\frac{\partial}{\partial a_i} E\{e^2(n)\} = 2 \sum_{j=0}^{\Lambda} R_w(j-i)a_j - 2R_{wv}(i) \quad i = 0, \dots, \Lambda. \quad (3.47)$$

Notice that $R_{wv}(-i) = R_{vw}(i)$.

Letting the derivative be equal to zero, we obtain

$$\sum_{j=0}^{\Lambda} R_w(j-i)a_j = R_{wv}(i) \quad i = 0, \dots, \Lambda, \quad (3.48)$$

which can be rewritten in matrix form

$$\mathbf{T}\mathbf{a} = \mathbf{v} \quad (3.49)$$

or

$$\begin{bmatrix} R_w(0) & R_w(1) & \cdots & R_w(\Lambda) \\ R_w(1) & R_w(0) & \cdots & R_w(\Lambda-1) \\ \vdots & \vdots & \ddots & \vdots \\ R_w(\Lambda) & R_w(\Lambda-1) & \cdots & R_w(0) \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_\Lambda \end{bmatrix} = \begin{bmatrix} R_{vw}(0) \\ R_{vw}(1) \\ \vdots \\ R_{vw}(\Lambda) \end{bmatrix}. \quad (3.50)$$

These equations are known as the Wiener-Hopf equations. The matrix \mathbf{T} appearing in the equation is a symmetric Toeplitz matrix. These matrices are known to be positive definite and therefore non-singular yielding a unique solution to the determination of the Wiener filter coefficient vector $\mathbf{a} = \mathbf{T}^{-1}\mathbf{v}$.

Least mean squares (LMS) algorithms are a class of adaptive filter used to mimic a desired filter by finding the filter coefficients that relate to producing the LMS of the error signal. Most linear adaptive filtering problem is shown in Fig. 17. That is, an unknown system $H(z)$ is to be identified and the adaptive filter attempts to adapt the filter $\hat{H}(z)$ to make it as close as possible to $H(z)$, while using only observable signals $x(n)$, $d(n)$ and $e(n)$; but $y(n)$, $v(n)$ and $\hat{y}(n)$ are not directly observable. Its solution is closely related to the Wiener filter. The idea behind LMS filters is to use steepest descent to find filter weights $H(n)$ which minimize a cost function. We start by defining the cost function as

$$\xi(n) = E\{|e(n)|^2\}, \quad (3.51)$$

where $E\{\cdot\}$ denotes the expected value.

The objective of the adaptive filter is to minimize the instantaneous mean square error and according to LMS algorithm updating the coefficient vector in the negative direction with step size μ :

$$h'(n+1) = h'(n) - \frac{\mu}{2} \nabla \hat{\xi}, \quad (3.52)$$

where $\mu/2$ is the step size, ∇ is the gradient operator, $\nabla \hat{\xi}$ is an instantaneous estimate of the mean square error (MSE) gradient at time n , and hence $\nabla \hat{\xi}$ can be expressed as

$$\nabla \hat{\xi}(n) = -2E\{\mathbf{x}(n)e^*(n)\}, \quad (3.53)$$

where $\mathbf{x}(n) = [x(n), x(n-1), \dots, x(n-p+1)]^T$ and p is filter order. The update algorithm follows as

$$h'(n+1) = h'(n) + \mu E\{\mathbf{x}(n)e^*(n)\}. \quad (3.54)$$

That means we have found a sequential update algorithm which minimizes the cost function. For most systems the expectation function $E\{\mathbf{x}(n)e^*(n)\}$ must be approximated. This can be done with the following unbiased estimator

$$\hat{E}\{\mathbf{x}(n)e^*(n)\} = \frac{1}{N} \sum_{i=1}^{N-1} \mathbf{x}(n-i)e^*(n-i), \quad (3.55)$$

where N indicates the number of samples we use for that estimate. The simplest case is $N=1$

$$\hat{E}\{\mathbf{x}(n)e^*(n)\} = \mathbf{x}(n)e^*(n). \quad (3.56)$$

For that simple case the update algorithm follows as

$$h'(n+1) = \hat{h}(n) + \mu \mathbf{x}(n)e^*(n). \quad (3.57)$$

Indeed this constitutes the update algorithm for the LMS filter.

The purposed of Kalman filter is to use measurements that are observed over time that contain noise and other inaccuracies, and produce values that tend to be closer to the true values of the measurements and their associated calculated values.

There are two equations in Kalman filtering. The first equation is called the process equation:

$$\mathbf{x}(n+1) = \mathbf{F}(n+1, n)\mathbf{x}(n) + \mathbf{v}_1(n), \quad (3.58)$$

where $\mathbf{F}(n+1, n)$ is a known $M \times M$ state transition matrix relating the state of the

system at times $n+1$ and n . The $M \times 1$ vector $\mathbf{v}_1(n)$ represents process noise. The vector $\mathbf{v}_1(n)$ is a zero-mean, white-noise process whose correlation matrix is defined by

$$E[\mathbf{v}_1(n)\mathbf{v}_1^H(n')] = \begin{cases} \mathbf{Q}_1, & n = n' \\ \mathbf{0}, & n \neq n' \end{cases} \quad (3.59)$$

The second equation is called the measurement equation:

$$y(n) = \mathbf{C}(n)\mathbf{x}(n) + \mathbf{v}_2(n) \quad (3.60)$$

where $\mathbf{C}(n)$ is a known $N \times M$ measurement matrix. The $N \times 1$ vector $\mathbf{v}_2(n)$ is the measurement noise modeled as a zero-mean, white-noise process whose correlation matrix is defined by

$$E[\mathbf{v}_2(n)\mathbf{v}_2^H(n')] = \begin{cases} \mathbf{Q}_2, & n = n' \\ \mathbf{0}, & n \neq n' \end{cases} \quad (3.61)$$

The noise vectors $\mathbf{v}_1(n)$ and $\mathbf{v}_2(n)$ are statistically independent so that we have

$$E[\mathbf{v}_1(n)\mathbf{v}_2^H(n')] = \mathbf{0}, \quad (3.62)$$

for all n and n' .

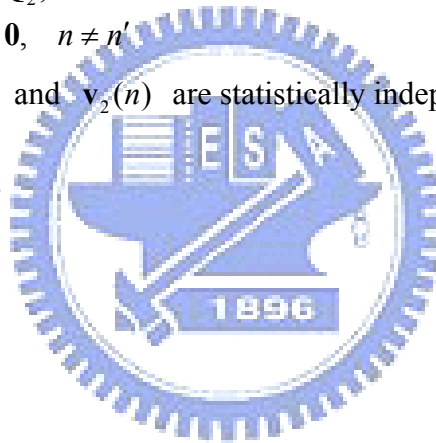


Table 1 Summary of Fourier series and transform relations.

	Continuous time		Discrete time	
	Time-domain	Frequency-domain	Time-domain	Frequency-domain
Fourier series	$x(t) = \sum_{k=-\infty}^{\infty} a_k e^{jk\omega_0 t}$	$a_k = \frac{1}{T_0} \int_{T_0} x(t) e^{-jk\omega_0 t} dt$	$x[n] = \sum_{k=\langle N \rangle} a_k e^{jk\Omega_0 n}$	$a_k = \frac{1}{N} \sum_{n=\langle N \rangle} x[n] e^{-jk\Omega_0 n}$
	Continuous time	Discrete frequency	Discrete time	Discrete frequency
	T_0 periodic	Aperiodic	N periodic	N periodic
	$\omega_0 = 2\pi / T_0$	$\omega_0 = 2\pi / T_0$	$\Omega_0 = 2\pi / N_0$	$\Omega_0 = 2\pi / N_0$
Fourier transform	$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(j\omega) e^{j\omega t} d\omega$	$X(j\omega) = \int_{-\infty}^{\infty} x(t) e^{-j\omega t} dt$	$x[n] = \frac{1}{2\pi} \int_{2\pi} X(e^{j\Omega}) e^{j\Omega n} d\Omega$	$X(e^{j\Omega}) = \sum_{n=-\infty}^{\infty} x[n] e^{-j\Omega n}$
	Continuous time	Continuous frequency	Discrete time	Continuous frequency
	Aperiodic	Aperiodic	Aperiodic	2π periodic

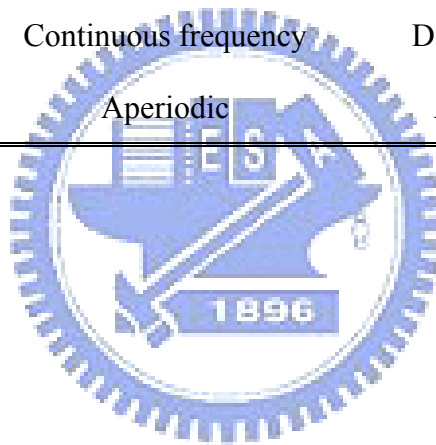


Table 2 FIR filter coefficients to implement Lagrange interpolation for $N = 1$ and $N = 2$.

	w_{m0}	w_{m1}	w_{m2}
$N = 1$	$1 - e_m$	e_m	
$N = 2$	$(e_m - 1)(e_m - 2)/2$	$-e_m(e_m - 2)$	$e_m(e_m - 1)/2$



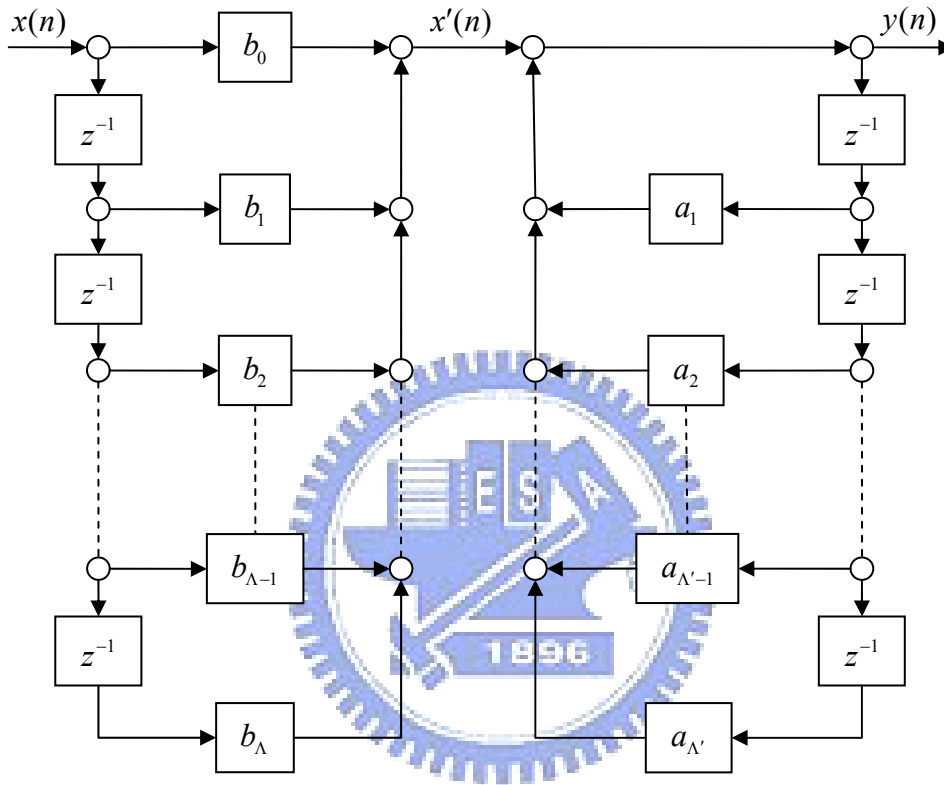


Fig. 12 Signal flow graph of direct-form I IIR structure.

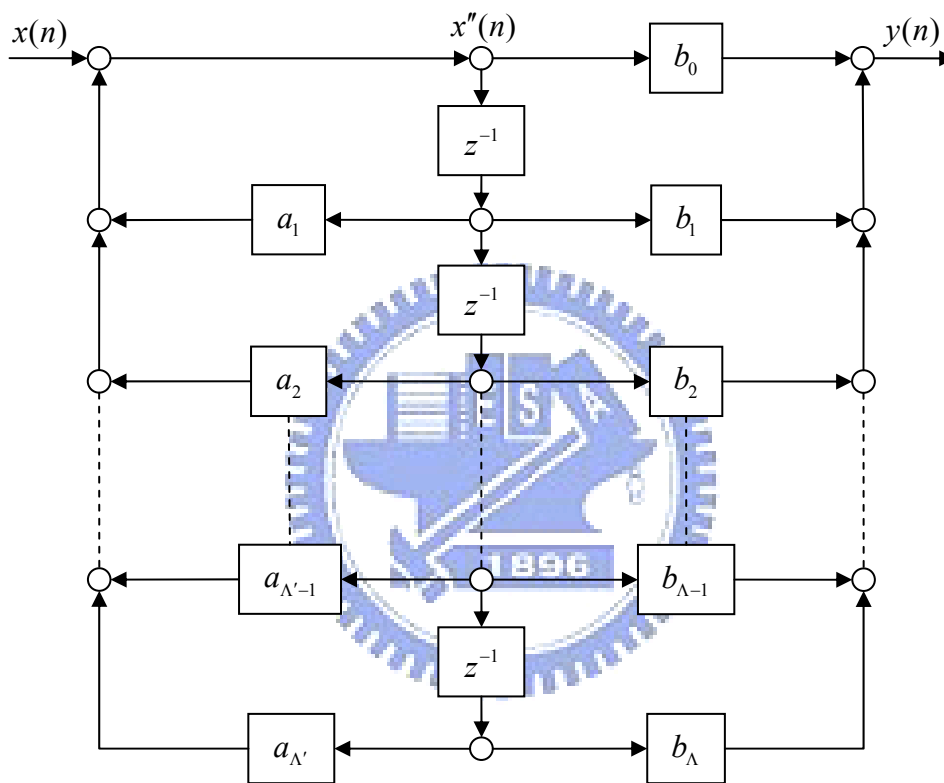


Fig. 13 Signal flow graph of direct-form II IIR structure.

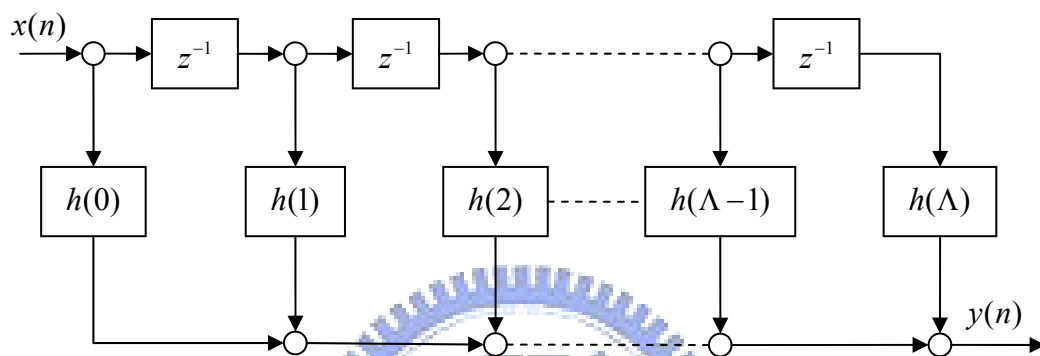
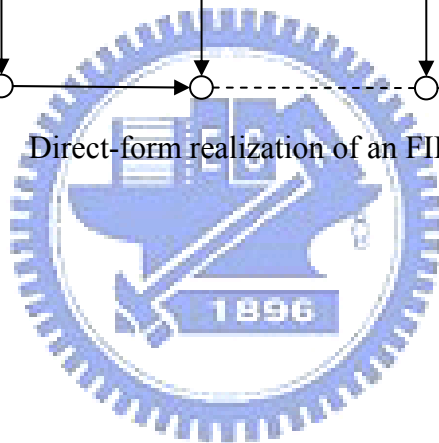


Fig. 14 Direct-form realization of an FIR system.



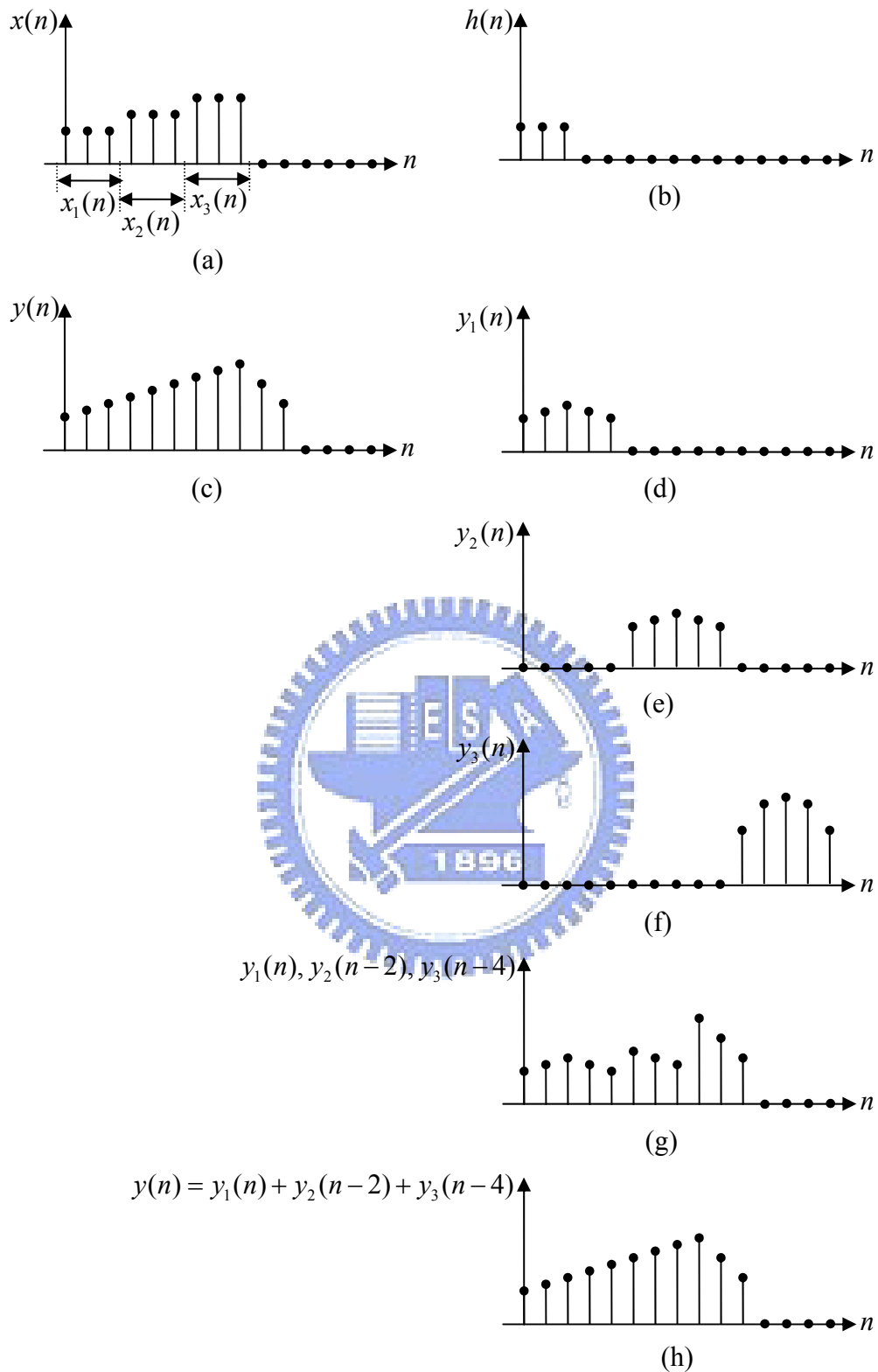


Fig. 15 Block convolution using the OAD. (a) input $x(n)$, (b) impulse response $h(n)$, (c) expected output $y(n)$, (d) output $y_1(n)$ for block convolution of $x_1(n)$ and $h(n)$, (e) output $y_2(n)$ for block convolution

of $x_2(n)$ and $h(n)$, (f) output $y_3(n)$ for block convolution of $x_3(n)$ and $h(n)$, (g) shifted block outputs, overlap is $N_h - 1 = 2$, and (h) the sum of overlapped block outputs equivalent to the direct convolution result.



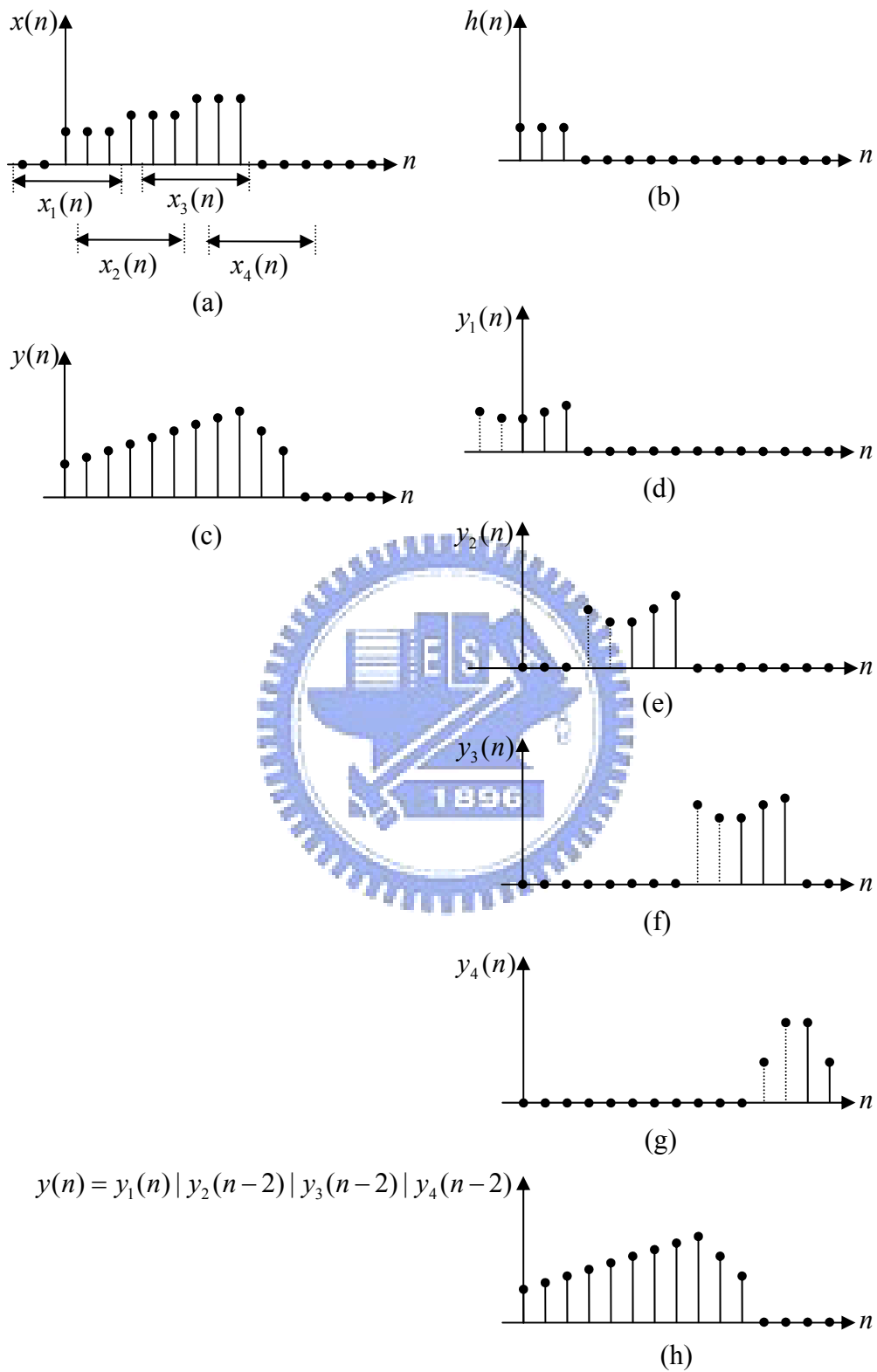


Fig. 16 Block convolution using the OAS. (a) input signal $x(n)$ divided into overlapping sections, overlap is $N_h - 1 = 2$, (b) impulse response $h(n)$, (c)

output $y(n)$ using direct convolution, (d) output $y_1(n)$ for block circular convolution of $x_1(n)$ and $h(n)$, (e) output $y_2(n)$ for block circular convolution of $x_2(n)$ and $h(n)$, (f) output $y_3(n)$ for block circular convolution of $x_3(n)$ and $h(n)$, (g) output $y_4(n)$ for block circular convolution of $x_4(n)$ and $h(n)$, and (h) sequential concatenation of block outputs after discarding the first two samples of each block, which is equivalent to the direct convolution result. "|" represents concatenation.



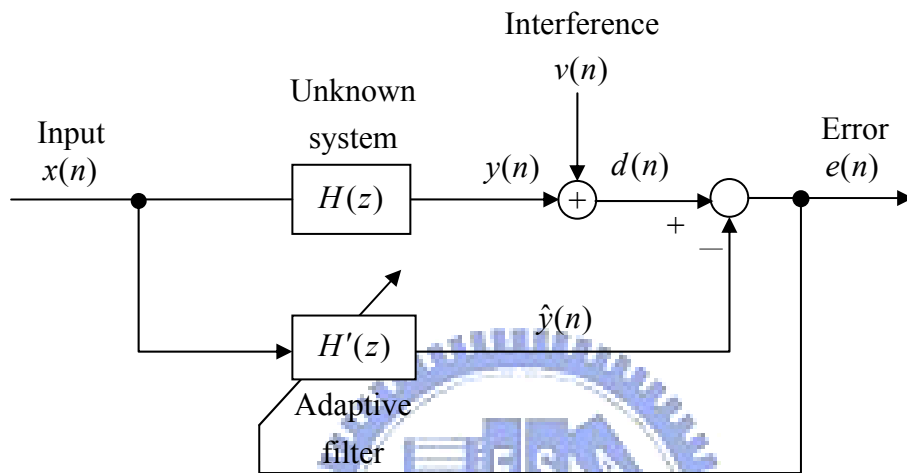


Fig. 17 LMS block diagram.

3.3 Array signal processing basics

Array signal processing techniques are utilized for DOA estimation, beamforming and NSI. Definitions of farfield and nearfield array are shown in Table 3. There is a one important thing we have to know. To avoid spatial aliasing, the spacing between each microphone must be less than half wavelength of the signals. In the following array signal processing, we assume that signals received at the reference point are of farfield and narrowband characteristics. Farfield assumes that the signals are located far enough away from the array that the wavefronts impinging on the array can be modeled as plane waves. On the other hand, the effects of distance on field intensity can be neglected. Narrowband assumes that the incident signal that the Beamformer is trying to capture has a narrow bandwidth centered at a particular frequency. The ULA is shown in Fig. 18 and the array model can be constructed. The spacing between adjacent microphones is d . A typical choice $d \leq \lambda/2$, which corresponds to Nyquist spatial sampling rate. Assume that the signal $r(t)$ at a reference point is a narrowband with center frequency ω_c :

$$r(t) = s(t)e^{j\omega_c t}, \quad (3.63)$$

where $s(t)$ is the phasor of $r(t)$. The signal received at the m th array element located at \vec{x}_m is denoted as $x_m(t)$, and let \vec{r} be the unit vector pointing to the sound source direction. If the speed of sound is c , the signal $x_m(t)$ can be written as

$$x_m(t) = r\left(t + \frac{\vec{x}_m \cdot \vec{r}}{c}\right) + n_m(t) = s\left(t + \frac{\vec{x}_m \cdot \vec{r}}{c}\right) e^{j\omega_c \frac{\vec{x}_m \cdot \vec{r}}{c}} e^{j\omega_c t} + n_m(t), \quad (3.64)$$

where $n_m(t)$ is the noise signal of the m th component in the array. In general, $s\left(t + \frac{\vec{x}_m \cdot \vec{r}}{c}\right) \approx s(t)$ for far field approximation. For M sensor signals $x_1(t), \dots, x_M(t)$, the data vector can be expressed as

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_M(t) \end{bmatrix} = \begin{bmatrix} e^{j\omega_c \frac{\vec{x}_1 \cdot \vec{r}}{c}} \\ \vdots \\ e^{j\omega_c \frac{\vec{x}_M \cdot \vec{r}}{c}} \end{bmatrix} s(t)e^{j\omega_c t} + \begin{bmatrix} n_1(t) \\ \vdots \\ n_M(t) \end{bmatrix} = \mathbf{a}(\vec{r})r(t) + \mathbf{n}(t), \quad (3.65)$$

where $\mathbf{a}(\vec{r})$ is called the array manifold vector. The unit vector \vec{r} for a sound source at the θ direction is given by

$$\vec{r} = (\sin \theta, \cos \theta). \quad (3.66)$$

The position vector of the m th element can be expressed as

$$\vec{x}_m = ((i-1)d, 0). \quad (3.67)$$

Hence, the inner product of the position vector and the unit vector yields

$$\vec{x}_m \cdot \vec{r} = (i-1)d \sin \theta \quad (3.68)$$

The array manifold vector $\mathbf{a}(\vec{r})$ can be rewritten from Eq. (3.68)

$$\mathbf{a}(\omega_c, \theta) = \begin{bmatrix} 1 & e^{j\omega_c \frac{d \sin \theta}{c}} & \dots & e^{j\omega_c \frac{(M-1)d \sin \theta}{c}} \end{bmatrix}^T \quad (3.69)$$

Extension from the narrowband to broadband formulation is straightforward. The center frequency ω_c is replaced by ω , where ω means a broadband frequency variable. The beamformer output $y(t)$ is the weighted sum of the delayed input signals $x_m(t), m = 1 \dots M$,

$$y(t) = \sum_{m=1}^M \sum_{k=0}^N w_{mk} x_m(t - kT). \quad (3.70)$$

This equation can be rewritten in the frequency-domain for a particular direction θ as

$$y(\omega, \theta) = \mathbf{h}(e^{j\omega}) \mathbf{x}(\omega, \theta) = \mathbf{h}(e^{j\omega}) (\mathbf{a}(\omega, \theta) r(\omega) + \mathbf{n}(\omega)), \quad (3.71)$$

where the manifold vector is given by

$$\mathbf{a}(\omega, \theta) = \begin{bmatrix} 1 & e^{j\omega \frac{d \sin \theta}{c}} & \dots & e^{j\omega \frac{(M-1)d \sin \theta}{c}} \end{bmatrix}^T \quad (3.72)$$

and $\mathbf{h}(e^{j\omega}) = [h_1(e^{j\omega}) \ h_2(e^{j\omega}) \ \dots \ h_M(e^{j\omega})]$ consists of Discrete Time Fourier

Transforms (DTFT) of each tapped-delay line channel. The frequency response is given by

$$h_m(e^{j\omega}) = \sum_{k=0}^N w_{mk} e^{-j\omega k T}, \quad m = 1 \cdots M, \quad (3.73)$$

where N and T are the filter order and the sampling period, respectively. The dimensions of $\mathbf{h}(e^{j\omega})$ and $\mathbf{a}(\omega, \theta)$ are $1 \times M$ and $M \times 1$, respectively.

In our problem, the delay-sum algorithm is tried to find $\mathbf{h}(e^{j\omega})$ that maximizes signal-to-noise ratio gain (SNRG), i.e.

$$\max_{\mathbf{h}(\omega)} \frac{\mathbf{h}(e^{j\omega}) \mathbf{a}(\theta) \mathbf{a}^H(\theta) \mathbf{h}^H(e^{j\omega})}{\mathbf{h}(e^{j\omega}) \mathbf{h}^H(e^{j\omega})} \quad (3.74)$$

which is equivalent to

$$\max_{\mathbf{h}(\omega)} \mathbf{h}(e^{j\omega}) \mathbf{a}(\theta) \mathbf{a}^H(\theta) \mathbf{h}^H(e^{j\omega}), \quad \text{subject to } \mathbf{h}(e^{j\omega}) \mathbf{h}^H(e^{j\omega}) = 1. \quad (3.75)$$

This problem can be solved by the Lagrange multiplier method, the solution is obtained as follows

$$\mathbf{h}(e^{j\omega}) = \mathbf{a}^H(\omega, \theta). \quad (3.76)$$

Equation (3.75) explains that each channel filter equals to the conjugate of each component in the manifold vector.

$$h_m(e^{j\omega}) = e^{-j\omega \frac{(m-1)d \sin \theta}{c}} = e^{-j\omega \tau_m}, \quad m = 1 \cdots M, \quad (3.77)$$

where $\tau_m = \frac{(m-1)d \sin \theta}{c}$ is the delay of each channel according to the difference

between m th sensor and the reference point. The delay is implemented using Lagrange interpolation. Once we obtain the filters, the output signals $y(\omega, \theta)$ can be calculated from Eq. (3.71). The square of $y(\omega, \theta)$ is called the spatial power spectrum, which is given by

$$S(\theta) = |y(\omega, \theta)|^2 = \left| \mathbf{h}(e^{j\omega}) \mathbf{x}(\omega, \theta) \mathbf{x}^H(\omega, \theta) \mathbf{h}^H(e^{j\omega}) \right|. \quad (3.78)$$

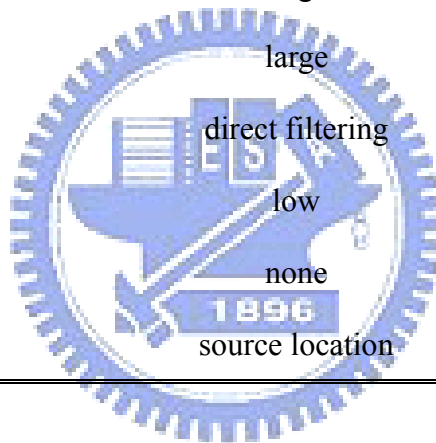
The maximum magnitude of the spatial power spectrum is the direction of the sound source. A quantity of interest in array signal processing is data correlation matrix:

$$\begin{aligned}
R_{xx} &= E\{x(n)x^H(n)\} \\
&= E\{[\mathbf{A}\mathbf{s}(n) + \mathbf{n}(n)][\mathbf{A}\mathbf{s}(n) + \mathbf{n}(n)]^H\} \\
&= \mathbf{A}E\{\mathbf{s}(n)\mathbf{s}^H(n)\}\mathbf{A}^H + E\{\mathbf{n}(n)[\mathbf{A}\mathbf{s}^H(n)]^H\} \\
&\quad + E\{[\mathbf{A}\mathbf{s}(n)]\mathbf{n}^H(n)\} + E\{\mathbf{n}(n)\mathbf{n}^H(n)\}.
\end{aligned}
\tag{3.79}$$



Table 3 Distinctions of farfield and nearfield arrays.

	Farfield array	Nearfield array
Source	simple source (plane or spherical)	distributed
Optimizing sensor deployment	random	uniform
Input-output domain	MISO	MIMO
Focal point	single	multiple
Main lobe	single	multiple
Distance	large	small
Source scale	large	small
Algorithm	direct filtering	inverse filtering
Resolution	low	high
Contemplation	none	evanescent wave
Sound image	source location	velocity distribution



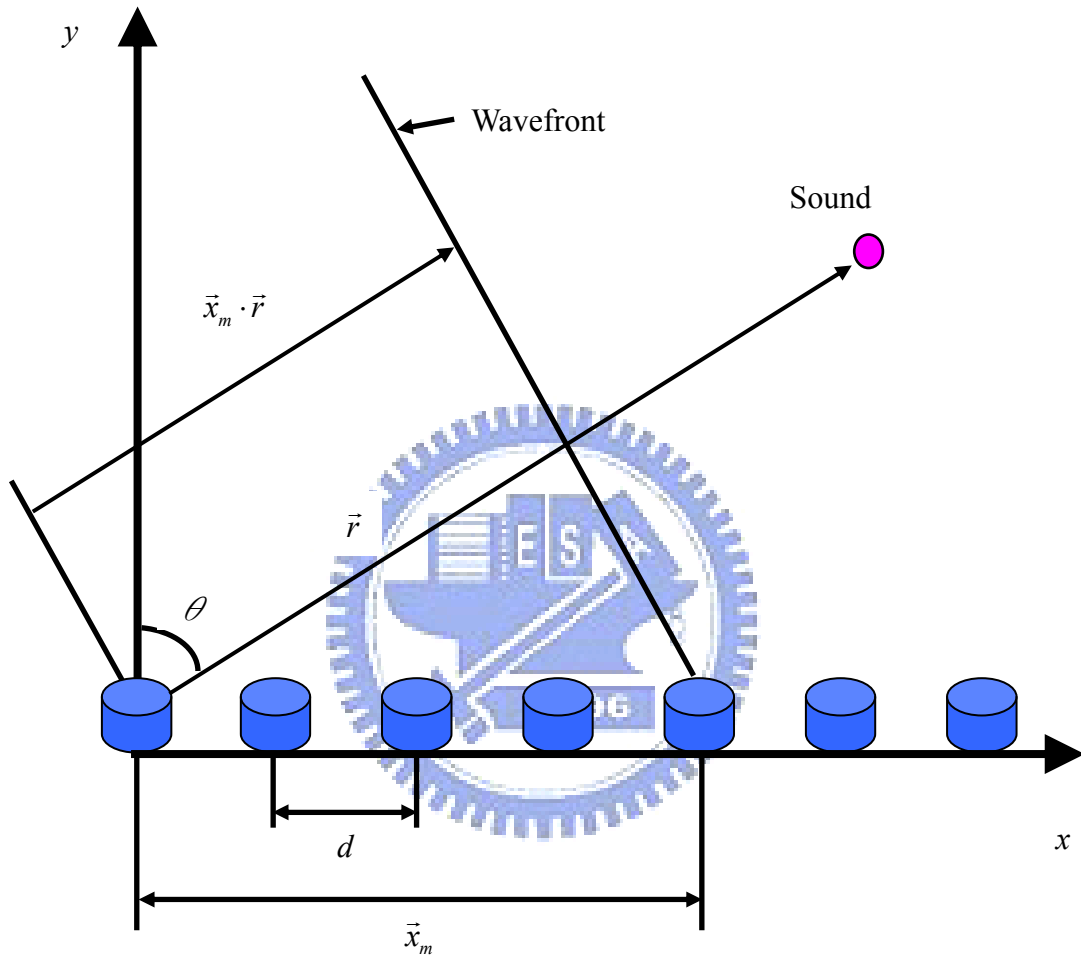


Fig. 18 The illustration of the ULA. A point source is located at the farfield.

3.4 Optimization algorithms

In this section, global optimization methods for microphone deployment and RD are presented. Arrays with sparse and random sensor deployment are known to be capable of delivering high quality farfield images without grating lobes. This raises the question of whether or not this idea can be applied to nearfield imaging as well. To answer this question that has not yet been widely investigated in previous research, numerical simulations of chapter 6.7 are undertaken in this thesis to optimize the microphone deployment for both farfield and nearfield arrays with the latter being the main focus. Next, we shall examine the issue of RD from an optimization perspective. The optimization algorithm we adopted is the GSS with Parabolic Interpolation [63].

The basic MC algorithm is based on straightforward random search. For M microphones to be allocated to $(m+1) \times (n+1)$ rectangular grid points, the number of possible combinations is $C_M^{(m+1) \times (n+1)}$, which is known to be an *NP-complete* problem [64]. Due to the blind search nature, the MC algorithm can be very inefficient and result in non-uniform distribution of microphones that concentrate at certain areas. To address these problems, a modified method intra-block Monte Carlo (IBMC) is proposed. By “Intra-block” (IB), we mean the localized region designated to each microphone on the surface,, as shown in Fig. 19. The MC search is only conducted within each block with random positions generated inside this designated region. The M microphone elements will be designated to M localized regions. Hence, each region necessarily contains one and only one microphone. The flowchart of IBMC is shown in Fig. 20. Initially, $m \times n$ divisions of a rectangular grid are set up on the microphone surface. Next, M localized search regions are designated to microphones,, as shown in Fig. 19. Each localized region in Fig. 21 has the

dimensions $d_{mx} = d_{my} = d$, whereas the inter-element spacing of the grid points is chosen to be $d_x = 4d/m$ and $d_y = 5d/n$, respectively. The localized regions are centered at the microphone positions of the URA that is selected to be the initial configuration in the optimization. The associated data including the microphone positions \mathbf{x}_i , the beam pattern \mathbf{b}_i , and the cost function Q_i are calculated. Next, each of the M microphone positions \mathbf{x} is randomly assigned to one of the search points on the localized region. The new beam pattern \mathbf{b} and the cost function Q are calculated for the assigned microphone positions \mathbf{x} . The optimal solutions \mathbf{x}_{opt} , \mathbf{b}_{opt} and Q_{opt} are then replaced by the new solutions \mathbf{x} , \mathbf{b} and Q if $Q > Q_{opt}$; otherwise the solutions are discarded. The simulation is continued until the number of iterations I exceed the preset value I_{IBMC} .

The IBMC algorithm is more efficient than the MC algorithm in that the search area for each microphone is far smaller. In addition, the IBMC algorithm generally results in microphone positions that are more uniformly distributed than those of the MC algorithm.

The MC algorithm can be very time-consuming and result in deployment that is far from optimal. Instead of blind search like the MC method, another efficient SA algorithm is used in this study. SA is a generic probabilistic meta-algorithm for the global optimization problem, namely locating a good approximation to the global optimum of a given function in a large search space [65]-[68]. SA is well suited for solving problems with many local optima. Each point in the search space is analogous to the thermal state of the annealing process in metallurgy. At high temperatures, atoms with high internal energy are free to move to the other positions. As temperature drops, the internal energy is decreased to a lower state to gradually

form a crystalline structure. The objective function Q to be maximized is likened to the internal energy in that state. One important feature of the SA approach is that it allows the search to move to a new state that is “worse” than the present one in the initial high-temperature stage. It is this mechanism that prevents the search from being trapped in a local maximum. The probability of accepting bad solutions decreases as temperature is decreased according to the Boltzmann distribution and the algorithm finally converges to the optimum solution.

The flowchart of SA is illustrated in Fig. 22. For the problem of maximizing the array cost function, the array is initially set to be the URA with microphone positions \mathbf{x}_i . The corresponding beam pattern \mathbf{b}_i and cost function Q_i are calculated. The microphone surface is partitioned into $m \times n$ divisions in a rectangular grid. The localized regions and the associated grid points are defined in the same way as the IBMC. Accordingly, each microphone can be assigned to any position within the localized region in the simulation. The initial temperature T_i , the final temperature T_f , and the annealing factor a are selected accordingly. A typical value of a is in the range of 0.8 and 0.99. Initially, set $\mathbf{x}_{\text{opt}} = \mathbf{x}_i$, $\mathbf{b}_{\text{opt}} = \mathbf{b}_i$ and $Q_{\text{opt}} = Q_i$. Next, M microphone positions \mathbf{x} are tentatively assigned. Each microphone is randomly assigned to one of the grid points with respect to the localized region. The beam pattern \mathbf{b} and the cost function Q are evaluated for a new microphone positions \mathbf{x} . Calculate the difference between the present and the optimal cost function,

$$\Delta Q = Q - Q_{\text{opt}}. \quad (3.80)$$

If $T > T_f$ and $\Delta Q > 0$, replace the optimized solutions \mathbf{x}_{opt} , \mathbf{b}_{opt} and Q_{opt} with the new solutions \mathbf{x} , \mathbf{b} and Q . Otherwise, if $\Delta Q \leq 0$, evaluate the following

probability function:

$$P(\Delta Q, T) = e^{\Delta Q/T}. \quad (3. 81)$$

The above probability will be compared with a random number $0 \leq \gamma \leq 1$ generated subject to the uniform distribution. A tentative solution is accepted when the probability function P is greater than the random number γ ; otherwise, the solution is rejected. Namely,

$$\begin{cases} P(\Delta Q, T) > \gamma, \text{ accepted} \\ P(\Delta Q, T) < \gamma, \text{ rejected} \end{cases} \quad (3. 82)$$

Note that the larger the cost function difference ΔQ or the higher the temperature T , the higher is the probability to accept a worse solution.

As the search proceeds, the temperature is decreased according to an exponential annealing schedule that begins at some initial temperature T_0 and decreases the temperature in steps

$$T_{k+1} = a \times T_k, \quad (3. 83)$$

where $0 < a < 1$ is the annealing coefficient. The annealing process will be terminated if the temperature is lower than a preset final temperature T_f . As the annealing process proceeds and T decreases, the probability of accepting a bad move becomes increasingly small until it finally settles to a stable solution.

The GSS algorithm is a bracketing method for finding the extremum of a one-dimensional function. One advantage of the method is that it enables efficient search for the extremum within a finite number of steps, without the need to evaluate numerical gradients. In a typical step there are seven points, $x_1, x_2, x_3, x_4, x_5, x_6$ and x_7 , not all distinct. One possible configuration is shown in Fig. 23. A local minimum is known to lie between x_1 and x_2 . Note that x_7 is the point with the least value of $Q(x)$ evaluated at x_1, x_2 and x_3 ; x_5 is the point with the next lowest of $Q(x)$; x_4 is the previous value of x_5 ; x_3 is the point at which the function was evaluated last; x_6

is the midpoint between x_1 and x_2 . The tolerance is chosen to be a combination of a relative and an absolute tolerance:

$$tol = \varepsilon^{1/2} |x_7| + \frac{tol_x}{3}. \quad (3.84)$$

The parameter ε is the floating-point relative accuracy, and the parameter tol_x is the tolerance of design variable.

The flowchart is shown in Fig. 24. The scheme to detect optimal minimum point has to be robust in the potential interval $[x_1, x_2]$, the attention must be paid to the optimal minimum evaluation. The following parameters p and q are calculated

$$p = \pm \left\{ (x_7 - x_4)^2 [Q(x_7) - Q(x_5)] - (x_7 - x_5)^2 [Q(x_7) - Q(x_4)] \right\}, \quad (3.85)$$

$$q = \mp 2 \left\{ (x_7 - x_4) [Q(x_7) - Q(x_5)] - (x_7 - x_5) [Q(x_7) - Q(x_4)] \right\}. \quad (3.86)$$

Let e be the value of p/q . If $|e| \leq tol$, $q = 0$ or $x_7 + p/q \notin (x_1, x_2)$, then a “GSS” step is performed. The next value of x_3 is

$$x_3 = \begin{cases} \alpha x_7 + (1 - \alpha)x_1 & \text{if } x_7 \geq x_6 \\ \alpha x_7 + (1 - \alpha)x_2 & \text{if } x_7 < x_6 \end{cases}. \quad (3.87)$$

Where $\alpha = (\sqrt{5} - 1)/2 \approx 0.618$ being the golden ratio. On the contrary, the parabola interpolation is performed when $|e| > tol$. Note that $x_7 + p/q$ is the extremum of the parabola interpolated through $[x_4, Q(x_4)]$, $[x_5, Q(x_5)]$, and $[x_7, Q(x_7)]$. The points x_1, x_2, x_3, x_4, x_5 , and x_7 are updated as necessary at next stage, so that it conforms to the definition of the seven points. The procedure above is repeated until the desired stop criterion ($|x_7 - x_6| \leq 2 \cdot tol - (x_2 - x_1)/2$) is reached.

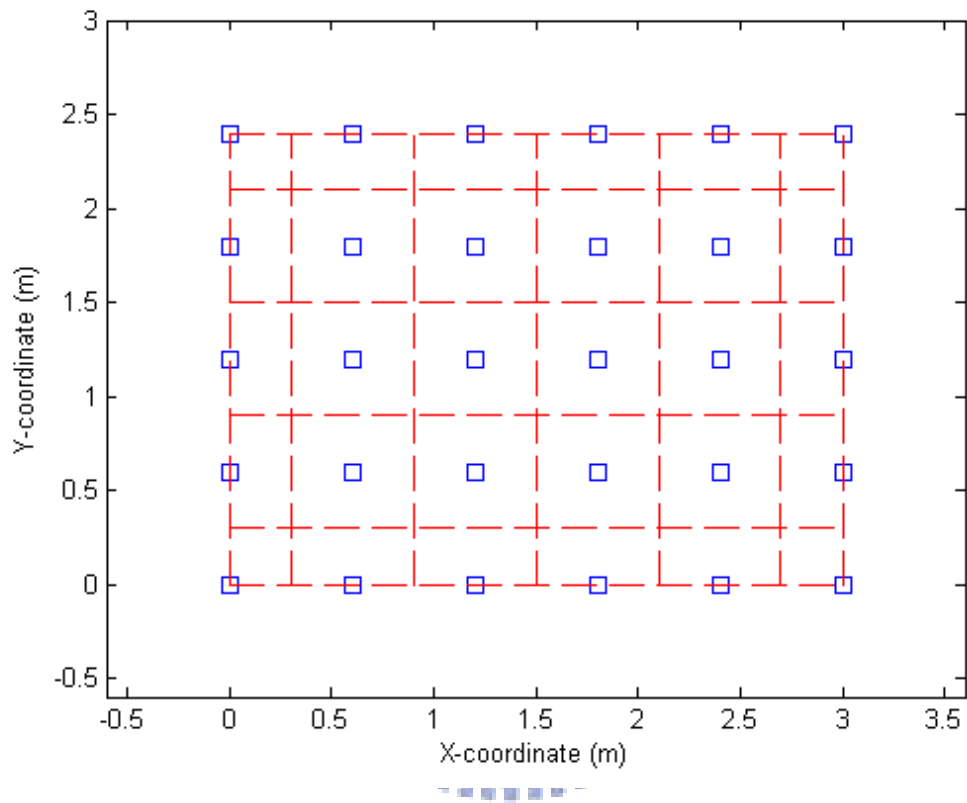


Fig. 19 The localized regions (dashed lines) on the microphone surface with the inter-element spacing $d = 0.6\text{m}$. The symbol “□” indicates the microphone position.

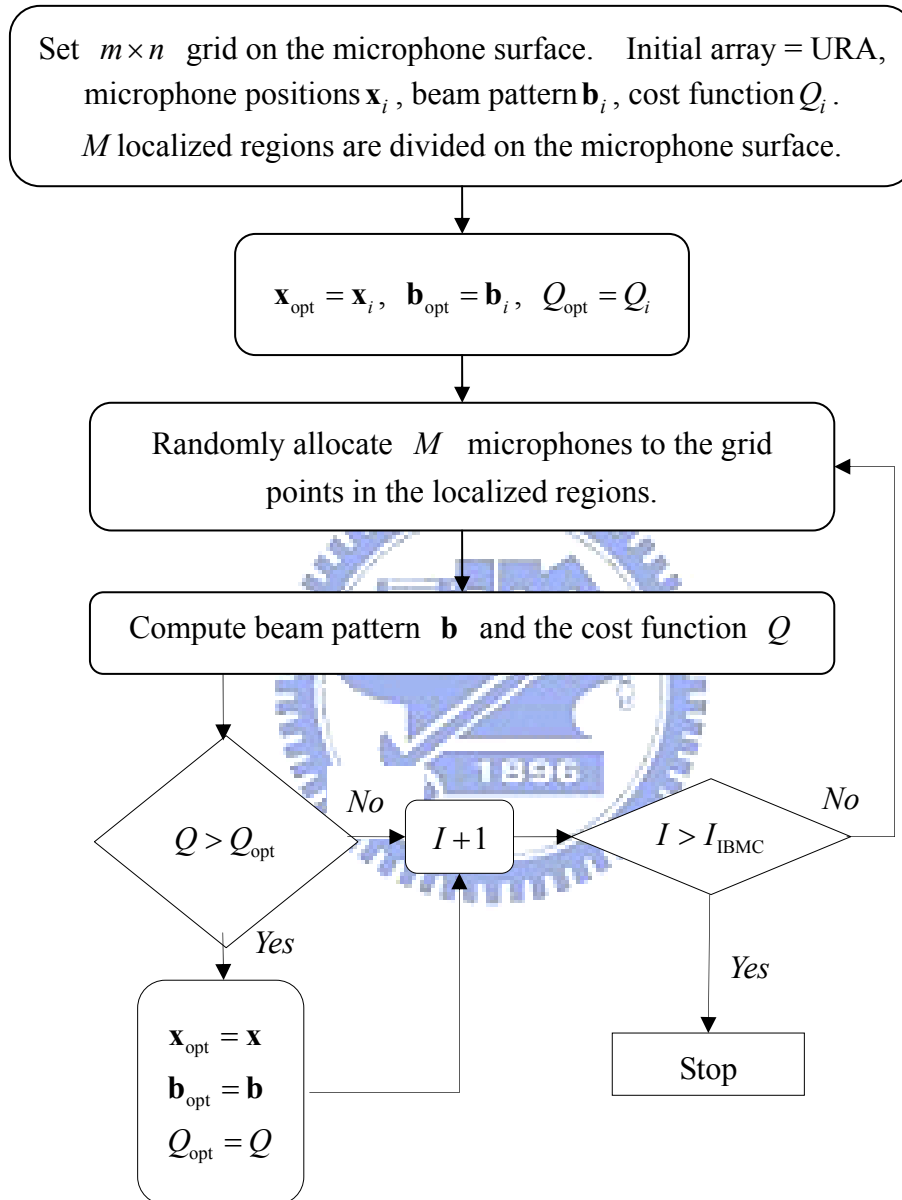


Fig. 20 Flowchart of the IBMC optimization algorithm.

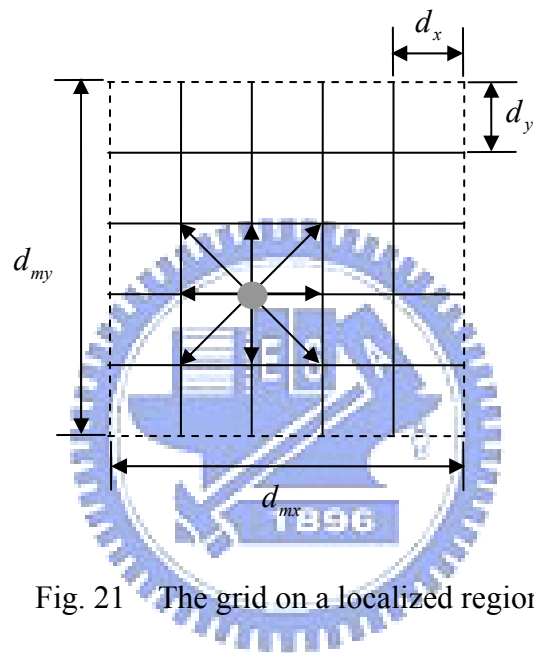


Fig. 21 The grid on a localized region.

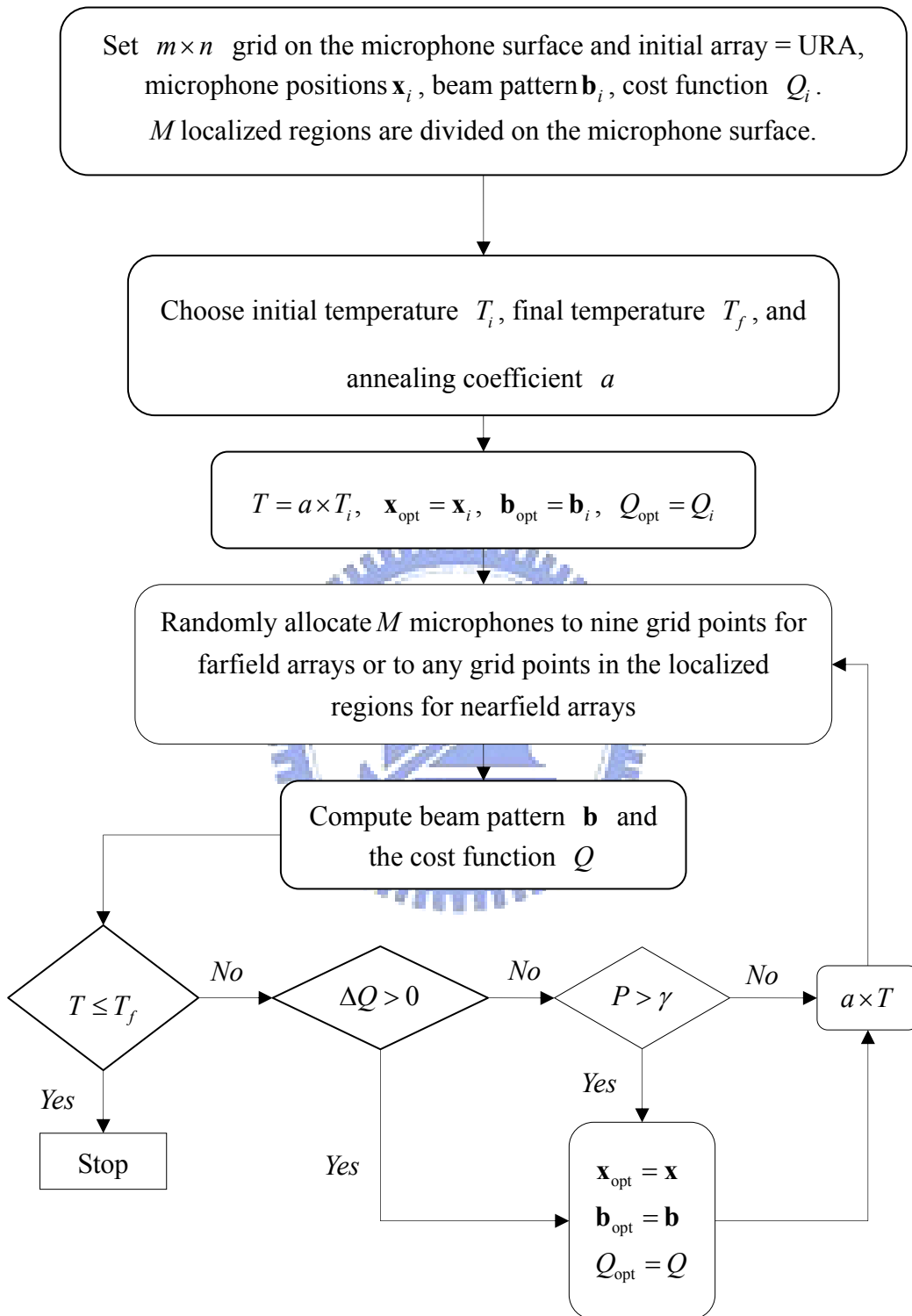


Fig. 22 Flowchart of the SA optimization algorithm.

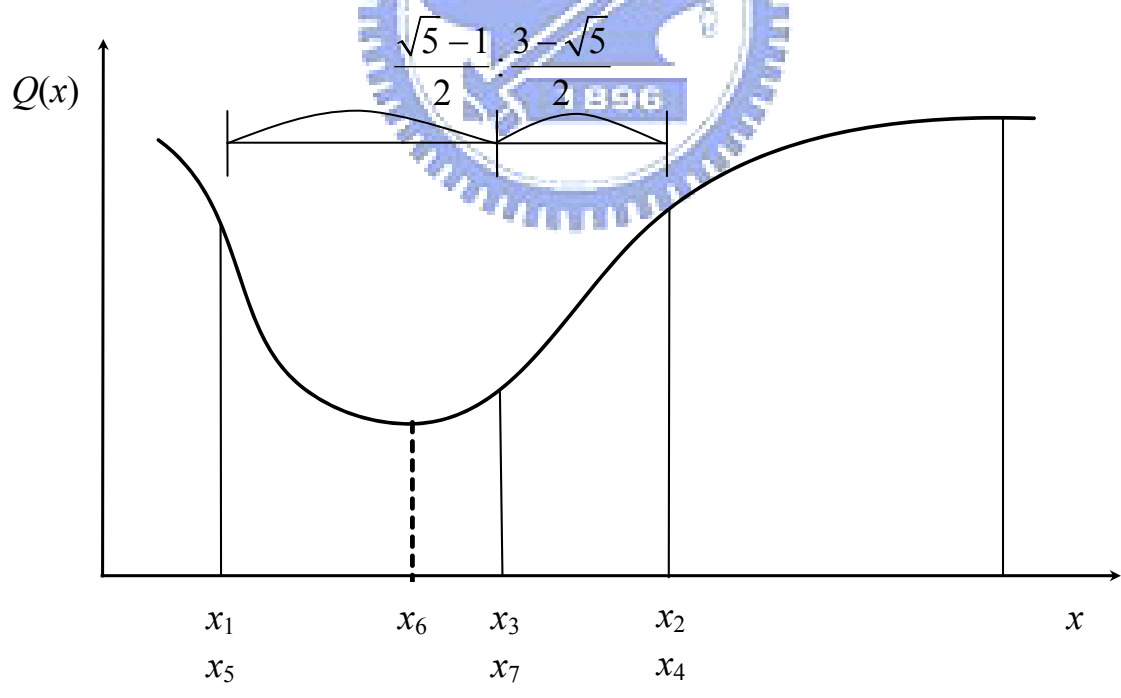
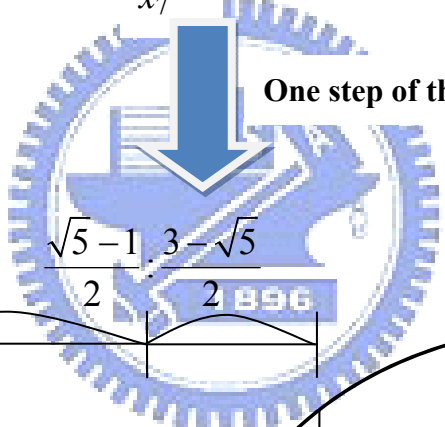
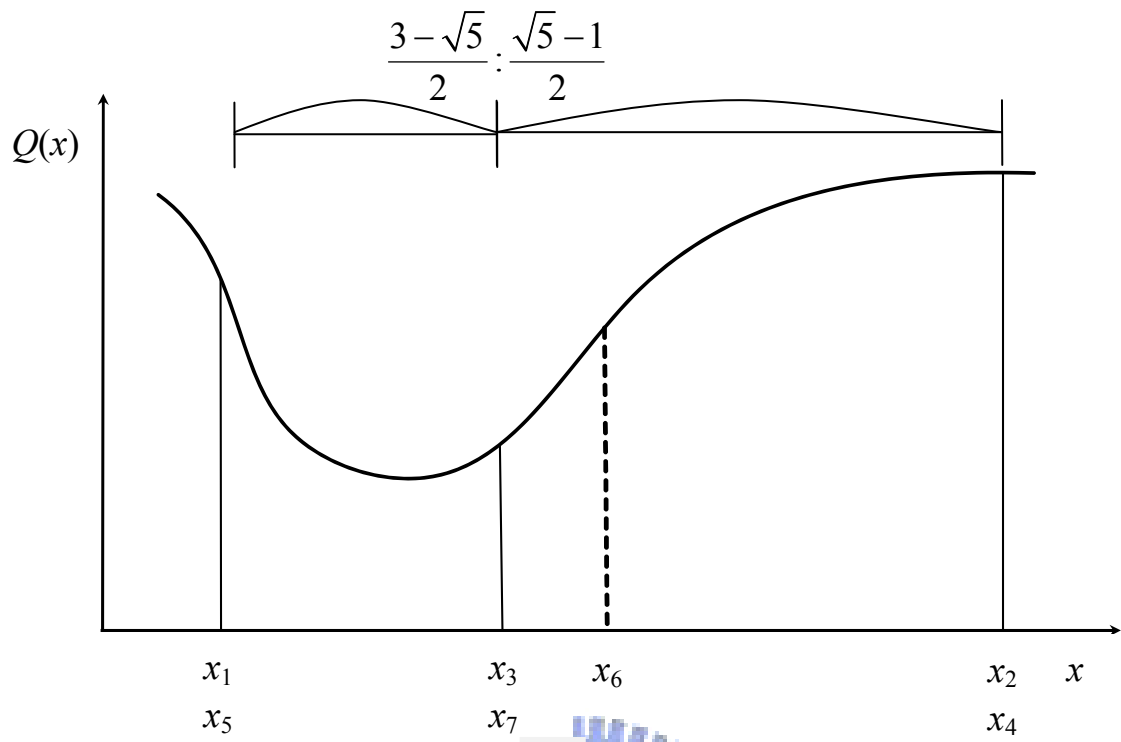


Fig. 23 A unimodal function to optimize.

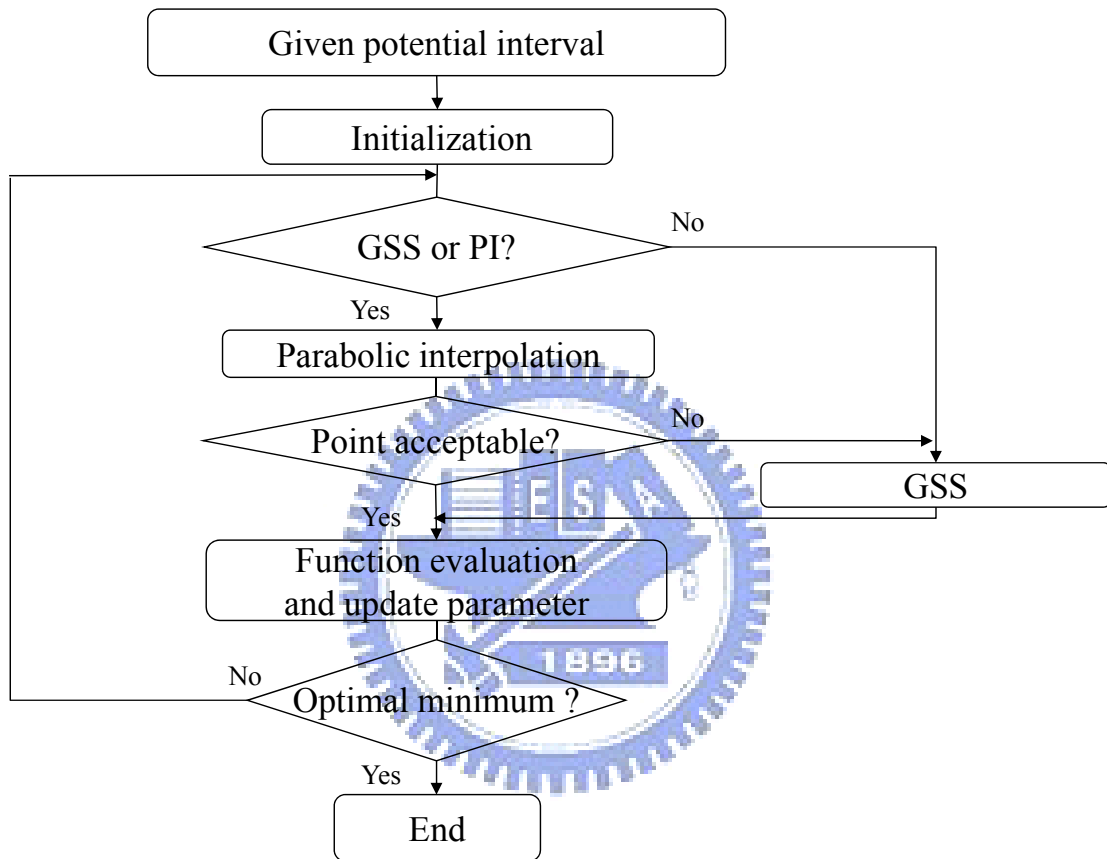


Fig. 24 The flowchart of the GSS algorithm.

CHAPTER 4. FARFIELD ARRAY SIGNAL PROCESSING ALGORITHMS

This section describes farfield array signal processing algorithms, which includes low-resolution algorithms and high-resolution algorithms. In addition, comparisons of farfield algorithms are summarized in finally.

4.1 Low-resolution algorithms

4.1.1 Delay and sum beamformer

Before discussing DOA algorithms, an array model should be established. A URA with inter-element spacing d can be constructed. Assume $r(t)$ is a broadband frequency ω at a reference point:

$$r(t) = s(t)e^{j\omega t}, \quad (4.1)$$

where $s(t)$ is the phasor of $r(t)$. Let \vec{r} be the unit vector pointing to the sound source direction. The signal received at the m th microphone located at \vec{x}_m is denoted as $x_m(t)$:

$$x_m(t) = r\left(t + \frac{\vec{x}_m \cdot \vec{r}}{c}\right) + n_m(t) = s\left(t + \frac{\vec{x}_m \cdot \vec{r}}{c}\right) e^{j\omega \frac{\vec{x}_m \cdot \vec{r}}{c}} e^{j\omega t} + n_m(t), \quad (4.2)$$

where $n_m(t)$ is the noise signal of the m th microphone and c is the speed of sound.

In general, $s\left(t + \frac{\vec{x}_m \cdot \vec{r}}{c}\right) \approx s(t)$ for farfield approximation. For M microphone signals $x_1(t), \dots, x_M(t)$, the data vector can be formed as

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_M(t) \end{bmatrix} = \begin{bmatrix} e^{j\omega \frac{\vec{x}_1 \cdot \vec{r}}{c}} \\ \vdots \\ e^{j\omega \frac{\vec{x}_M \cdot \vec{r}}{c}} \end{bmatrix} s(t)e^{j\omega t} + \begin{bmatrix} v_1(t) \\ \vdots \\ v_M(t) \end{bmatrix} = \mathbf{a}(\vec{r})r(t) + \mathbf{n}(t), \quad (4.3)$$

where $\mathbf{a}(\vec{r})$ is termed the array manifold vector.

A URA comprising I and J microphones with inter-element spacing d_x and d_y in the x and y axis respectively as shown in Fig. 25. Let the left and the upper corner element is the reference point. The vector of pointing to each microphone from the

reference point in the polar coordinate is given as

$$\vec{x}_{ij}(t) = ((i-1)d_x, (j-1)d_y, 0), \quad (4.4)$$

where $i = 1, 2, \dots, I$ and $j = 1, 2, \dots, J$. The unit vector \vec{r} pointing to a sound source at the look directions θ and Φ is given by

$$\vec{r} = (\sin \theta \sin \phi, \sin \theta \cos \phi, \cos \theta). \quad (4.5)$$

The delay of each microphone will be given by

$$\tau_{ij} = \frac{\vec{x}_{ij} \cdot \vec{r}}{c} = \frac{(i-1)d_x \sin \theta \sin \phi + (j-1)d_y \sin \theta \cos \phi}{c}. \quad (4.6)$$

Then the array manifold vector can be written as

$$a(\omega, \theta, \phi) = \begin{bmatrix} 1 \\ e^{j\omega \frac{d_x \sin \theta \sin \phi}{c}} \\ \vdots \\ e^{j\omega \frac{(I-1)d_x \sin \theta \sin \phi}{c}} \\ e^{j\omega \frac{d_y \sin \theta \cos \phi}{c}} \\ \vdots \\ e^{j\omega \frac{d_x \sin \theta \sin \phi + d_y \sin \theta \cos \phi}{c}} \\ \vdots \\ e^{j\omega \frac{(I-1)d_x \sin \theta \sin \phi + (J-1)d_y \sin \theta \cos \phi}{c}} \end{bmatrix}. \quad (4.7)$$

With reference to the Fig. 26, the output of DAS beamformer can be expressed as

$$y(t, \theta_0) = \sum_{m=1}^M x_m(t - \tau_m), \quad (4.8)$$

where $x_m(t)$ is the signal received by m th microphone, as shown in Eq.(4.3).

In Eq. (4.8), τ_m are the steering delays appropriate for focusing the array to the look direction, θ_0 , and compensation for the direct path propagation delay associated with the desired signal at each microphone. The delay of each channel in Eq. (4.8) can be calculated by

$$\tau_m = \frac{\Delta x_m}{c},$$

where Δx_m is the distance between the reference source positions θ_s and m th microphone. This time delay is implemented using Lagrange interpolation.



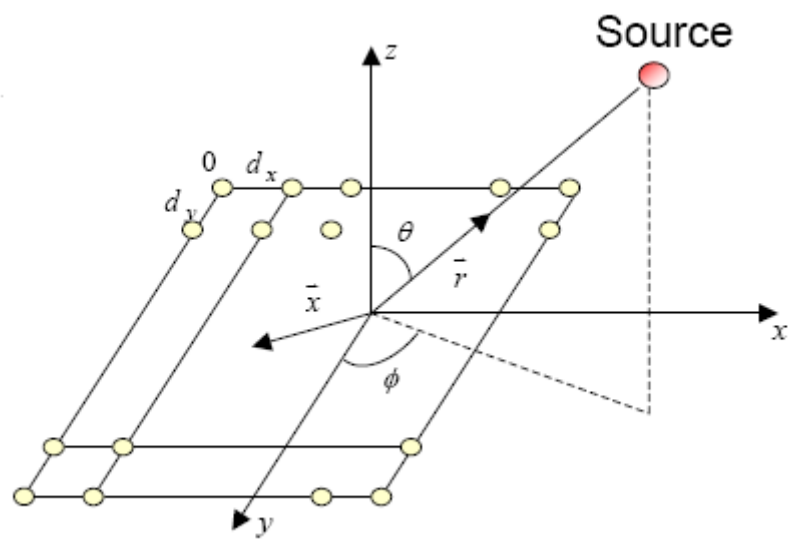


Fig. 25 A URA. A point sound source is located at the farfield.

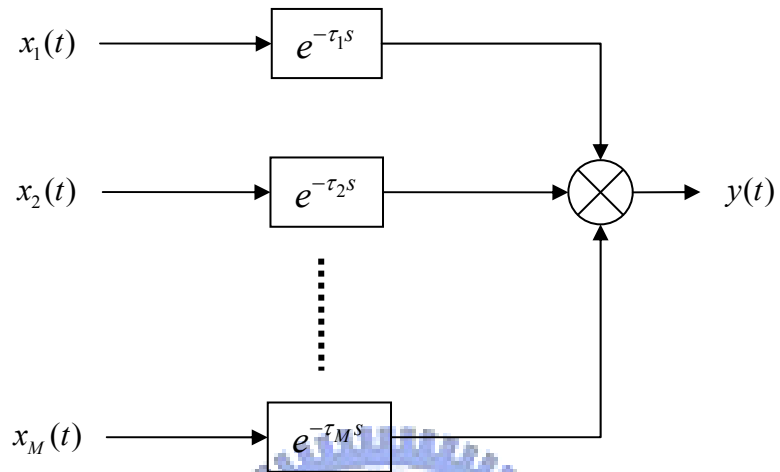
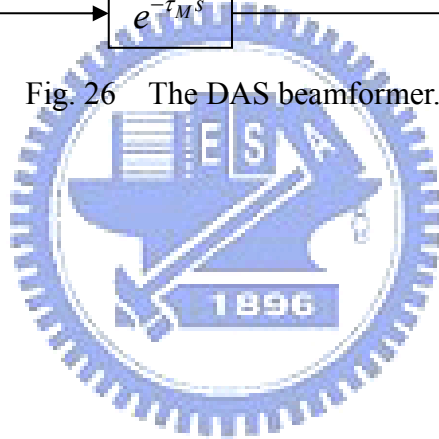


Fig. 26 The DAS beamformer.



4.1.2 Time reversal beamformer

The TR algorithm is based on an idea of time-reversing the received signal. The block diagram of TR algorithm is shown in Fig. 27. First, use a microphone array to receive and save sound data. The signal $x(t)$ is reversed to $x(-t)$ by TR block. Then the reversed signal $x(-t)$ is played by a loudspeaker array.

The phase-conjugate field at the field location \vec{r} in frequency domain is written as

$$p(\vec{r}) = \sum_{i=1}^I g^*(\vec{r}_i | \vec{r}_{ps}) g(\vec{r} | \vec{r}_i) = \mathbf{g}^+(\mathbf{r}_a | \vec{r}_{ps}) \mathbf{g}(\vec{r} | \mathbf{r}_a), \quad (4.9)$$

where $g(\vec{r}_i | \vec{r}_{ps})$ represents the received acoustic pressure at the i th array element location \vec{r}_i propagated from the probe source position \vec{r}_{ps} . Likewise, $g(\vec{r} | \vec{r}_i)$ represents the field propagated from the i th array element location \vec{r}_i to the arbitrary receiver location \vec{r} , where I is the number of array elements and superscripts $()^*$ and $()^+$ denote complex conjugate and Hermitian transpose, respectively. In a vector notation, \mathbf{g} and \mathbf{r}_a are $(I \times 1)$ column vectors. Note that the position vectors are written in *italic* letters with arrows and the column vectors and matrices are written in boldface letters.

In the following, we are aiming at finding an optimal filter with the impulse response \mathbf{h} that processes the measured signal

$$\mathbf{x} = \mathbf{s} + \mathbf{v} \quad (4.10)$$

to maximize the SNR, where \mathbf{s} and \mathbf{v} denote the signal and noise, respectively.

Hence,

$$\mathbf{y} = \mathbf{h}^H \mathbf{x} = \mathbf{h}^H \mathbf{s} + \mathbf{h}^H \mathbf{v}. \quad (4.11)$$

Note that the vectors are all ‘time-reversed,’ e.g., $\mathbf{s} = [s(n) \ s(n-1) \ \dots \ s(n-N)]^T$.

$$SNR = \frac{|\mathbf{h}^H \mathbf{s}|^2}{E\{|\mathbf{h}^H \mathbf{v}|^2\}} = \frac{\mathbf{h}^H \mathbf{s} \mathbf{s}^H \mathbf{h}}{\mathbf{h}^H \mathbf{R}_v \mathbf{h}}, \quad (4.12)$$

where \mathbf{R}_v is the covariance matrix of noise (diagonal and positive definite). The SNR expression is identified as the Rayleigh's quotient. Maximal SNR can be obtained by finding the λ_{\max} of the following eigenvalue problem

$$\mathbf{s} \mathbf{s}^H \mathbf{h} = \lambda \mathbf{R}_v \mathbf{h}. \quad (4.13)$$

Since $\mathbf{s} \mathbf{s}^H$ is of unit rank and positive definite, it has only one positive eigenvalue. The associated eigenvector must lie in the range space of $\mathbf{s} \mathbf{s}^H$, or the span of \mathbf{s} , which satisfies

$$\mathbf{R}_v \mathbf{h} = \alpha \mathbf{s} \quad \text{or} \quad \mathbf{h}_{opt} = \alpha \mathbf{R}_v^{-1} \mathbf{s}. \quad (4.14)$$

Thus, the optimal filter is closely related to the time-reversed version (phase conjugation) of the measured signal. With the optimal filter, the following maximal SNR is attained from Eqs. (4.12) and (4.14)

$$SNR_{\max} = \lambda_{\max} = \mathbf{s}^H \mathbf{R}_v^{-1} \mathbf{s}. \quad (4.15)$$

If the filter coefficient is normalized with the constraint

$$\mathbf{h}^H \mathbf{R}_v \mathbf{h} = 1 \quad (4.16)$$

Combining Eqs. (4.14) and (4.16) gives

$$\alpha = \frac{1}{\sqrt{\lambda_{\max}}} = \frac{1}{\sqrt{\mathbf{s}^H \mathbf{R}_v^{-1} \mathbf{s}}} \Rightarrow \mathbf{h}_{opt} = \frac{\mathbf{R}_v^{-1} \mathbf{s}}{\sqrt{\mathbf{s}^H \mathbf{R}_v^{-1} \mathbf{s}}}. \quad (4.17)$$

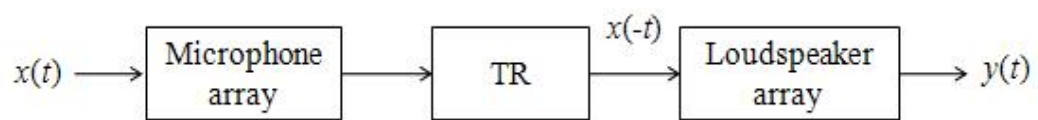
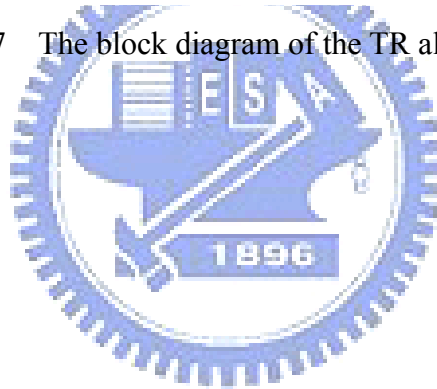


Fig. 27 The block diagram of the TR algorithm.



4.1.3 SIMO-ESIF algorithm

The SIMO-ESIF algorithm is introduced in this section. In Fig. 28, M microphones are employed to pick up the sound emitting from a source positioned in the farfield. In the frequency-domain, the sound pressure received at the microphones and the source signal can be related by a $M \times 1$ transfer matrix \mathbf{H}

$$\mathbf{p} = \mathbf{H}q(\omega), \quad (4.18)$$

where $q(\omega)$ is the Fourier transform of a scalar source strength,

$\mathbf{p} = [p_1(\omega) \cdots p_M(\omega)]^T$ is the pressure vector with “ T ” denoting matrix transpose, and

$\mathbf{H} = [h_1(\omega) \cdots h_M(\omega)]^T$ is the $M \times 1$ propagation matrix. The aim here is to estimate the source signal $q(\omega)$ based on the pressure measurement \mathbf{p} by using a set of inverse filters

$$\mathbf{C} = [c_1(\omega) \cdots c_M(\omega)]^T \quad (4.19)$$

such that $\mathbf{C}^T \mathbf{H} \approx \mathbf{I}$ and therefore

$$\hat{q} = \mathbf{C}^T \mathbf{p} = \mathbf{C}^T \mathbf{H}q \approx q. \quad (4.20)$$

On the other hand, this problem can also be written in the context of the following least-squares optimization problem

$$\min_q \|\mathbf{p} - \mathbf{H}q\|_2^2, \quad (4.21)$$

where $\|\cdot\|_2$ denotes vector 2-norm. This is an overdetermined problem whose least-squares solution is given by

$$\hat{q} = (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H \mathbf{p} = \frac{\mathbf{H}^H \mathbf{p}}{\|\mathbf{H}\|_2^2}, \quad (4.22)$$

where the superscript “ H ” denotes Hermitian transpose. Comparison of Eqs. (4.20)

and (4.22) yields the following optimal inverse filter

$$\mathbf{C}^T = \frac{\mathbf{H}^H}{\|\mathbf{H}\|_2^2}, \quad (4.23)$$

If the scalar $\|\mathbf{H}\|_2^2$ is omitted, the inverse filters above reduce to the “phase-conjugated” filters, or the “time-reversed” filters in the free-field context. Specifically, for a point source in the free field, it is straightforward to show that

$$\|\mathbf{H}\|_2^2 = \sum_{m=1}^M \frac{1}{r_m^2}, \quad (4.24)$$

where r_m is the distance between source and the m th microphone. Since $\|\mathbf{H}\|_2^2$ is a frequency-independent constant, the inverse filters and the time-reversed filters differ only by a constant. In a reverberant environment, these filters are different in general. Being able to incorporate the reverberant characteristics in the measured acoustical plant model represents an advantage of the proposed approach over conventional methods such as the DAS beamformer.

In real-time implementation, the inverse filters are converted to the time-domain FIR filters with the aid of inverse FFT and circular shift. Thus, the source signal can be recovered by filtering the pressure signals with the inverse filters $\mathbf{c}(k)$:

$$\hat{q}(k) = \mathbf{c}^T(k) * \mathbf{p}(k), \quad (4.25)$$

where k is discrete-time index, $\mathbf{c}(k)$ is the impulse response of the inverse filter, and “*” denotes convolution.

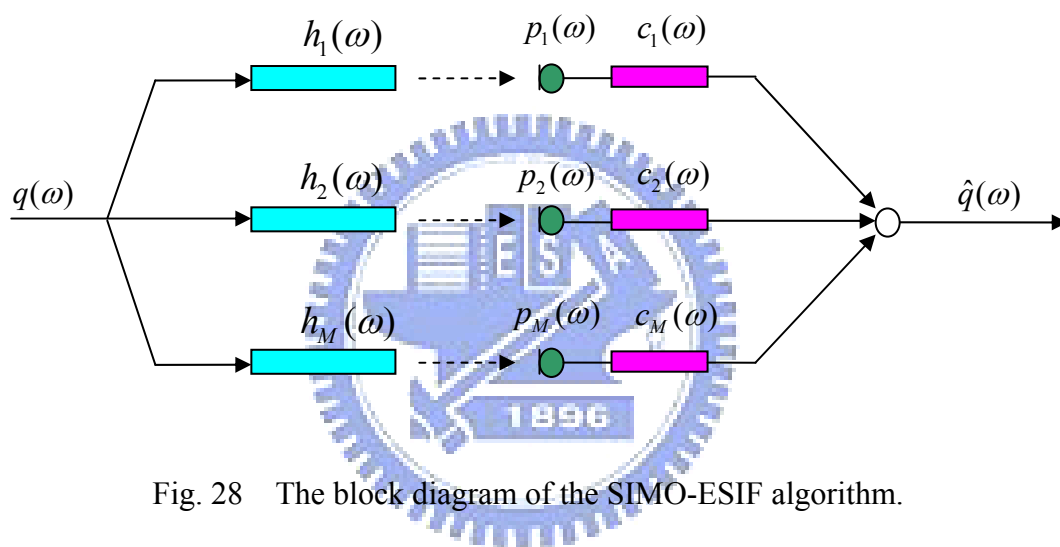


Fig. 28 The block diagram of the SIMO-ESIF algorithm.

4.1.4 Optimal array: cost functions, Rayleigh's quotient

In this section, the beam patterns and the cost functions are defined to facilitate the following array optimization formulation. In addition, super-directive microphone arrays are introduced in last section.

For a farfield array, the beam pattern can be defined in the wave number domain

$$b = \frac{1}{M} \sum_{m=1}^M e^{j\mathbf{k} \cdot \mathbf{r}_m}, \quad (4.26)$$

where $k = \omega/c$ is the wave number, ω is angular frequency, c is the speed of sound, and \mathbf{r}_m is the position vector of the m th microphone and $\mathbf{k} = -k\boldsymbol{\kappa}$ is the wave number vector of a plane wave incident from the direction represented by the unit vector $\boldsymbol{\kappa}$, as shown in Fig. 29.

In optimizing farfield performance, the aim is to minimize the maximum side-lobe level (MSL) of the beam pattern. First, a circle with radius r_m is drawn on the k_x - k_y plane to define the scope of the main-lobe, which is a judicious choice based on the beam pattern observations. The exterior of this circle is considered the side-lobe region. The cost function for farfield arrays is defined as

$$Q = \frac{\tilde{m}}{\tilde{s}}, \quad (4.27)$$

where \tilde{m} and \tilde{s} denote the maxima of the main-lobe and the side-lobes, respectively. Because $\tilde{m} = 1$, maximizing the cost function Q amounts to minimizing the MSL.

Super-directive microphone arrays are introduced in this section. It begins with first-order differential microphone array (DMA), a simple kind of super-directive microphone array. Second, a method of optimization of array beam pattern is introduced. There are three objective functions to be maximized, including directive index (DI) and front-to-back ratio (FBR). Due to their directional and close-talking properties, they have proven essential for the reduction of feedback in public address

systems. In telephone applications, such as speakerphone teleconferencing, directional microphones are very useful but at present are seldom utilized. Since small differential arrays can offer significant improvement in typical teleconferencing configurations, it is expected that they will become more prevalent in years to come.

The first-order DMAs have been discussed for more than 50 years [69]-[71]. Owing to the small size of first-order DMAs, they can be used in hands-free telecommunications where the distance between microphones and speakers are quite short. Another benefit is that the directivity of first-order DMAs is independent of frequency. The block diagram of the first-order DMA is shown in Fig. 30. For a plane wave with amplitude A and wave number k incident on a two-element array, the magnitude of output can be written as

$$|P_d| = A\omega \left(\tau + \frac{d \cos \theta}{c} \right) = A\omega \left(\tau + \frac{d}{c} \right) \left[\frac{\tau}{\tau + \frac{d}{c}} + \frac{\frac{d}{c}}{\tau + \frac{d}{c}} \cos \theta \right], \quad (4.28)$$

where τ is the incorporated delay, c is the speed of sound, and θ is the polar angle. It is found from Eq. (4.28) that the response of first-order DMAs is in direct proportion with frequency. This implies that the frequency response of the first-order DMAs need to be equalized to compensate for the low-frequency loss and high-frequency noost. Let

$$\alpha_1 = a_0 = \frac{\tau}{\tau + \frac{d}{c}} \quad (4.29)$$

and

$$1 - \alpha_1 = a_1 = \frac{\frac{d}{c}}{\tau + \frac{d}{c}}. \quad (4.30)$$

It follows that

$$a_0 + a_1 = 1. \quad (4.31)$$

Thus, the normalized directional response is

$$p_{Nd}(\theta) = a_0 + a_1 \cos \theta. \quad (4.32)$$

Accordingly, the directional response can be designed by adjusting the parameter a_0 . In order to get a better directivity, there should be objective measures for analyzing the array performance. One possible measure is FBR, the microphone gain for signals propagating to the front of the microphone relative to the rear, and it is defined as

$$FBR(\omega) = \frac{\frac{1}{2\pi} \int_0^{2\pi} \int_0^{\pi/2} |H(\omega, \theta, \phi)|^2 \sin \theta d\theta d\phi}{\frac{1}{2\pi} \int_0^{2\pi} \int_{\pi/2}^{\pi} |H(\omega, \theta, \phi)|^2 \sin \theta d\theta d\phi}, \quad (4.33)$$

where the angles θ and ϕ are the spherical coordinate angles and $H(\omega, \theta, \phi)$ is the frequency response of the array. The relativity between the parameter α_1 and FBR is shown in Fig. 31. The maximum FBR occurs when α_1 is equal to 0.366, and in this situation, the array can reject the noise from rear well.

The other measure is DI, the ratio of intensity of the acoustic beam in the measured axis to that of the entire distributed omnidirectional sound energy. It is defined as

$$DI(\omega, \theta_0, \phi_0) = 10 \log_{10} \left(\frac{|H(\omega, \theta_0, \phi_0)|^2}{\frac{1}{4\pi} \int_0^{2\pi} \int_0^{\pi} |H(\omega, \theta, \phi)|^2 \sin \theta d\theta d\phi} \right), \quad (4.34)$$

where θ_0 and ϕ_0 are the angles at which DI is being measured. The DI of first-order DMAs varies with the parameter α_1 , and the maximum DI reaches at $\alpha_1 = 0.25$, which is shown in Fig. 32.

The polar plot of the absolute value of the responses is shown in Fig. 33. The first-order DMAs that correspond to the maximum DI is given the name hypercardioid, and the maximum FBR value corresponds to the supercardioid design.

When $\alpha_1 = 0$, the 1st order differential system is a dipole. At $\alpha_1 = 1$, the microphone is an omnidirectional microphone with 0 dB DI. A special case of $\alpha_1 = 0.5$ is the cardioids pattern. Although the cardioids microphone is not optimal in directional gain or FBR, it is the most commonly manufactured differential microphone. Table 4 summarizes the results for first-order microphone.

The directivity pattern of the first-order DMA in different frequency is shown in Fig. 34. When α_1 is fixed, the shape of the directivity pattern is almost the same no matter what frequency it is, but the gain increases with frequency.

Assuming the signals are narrowband and incident from the farfield such that the wavefronts impinging on the array can be modeled as plane waves. To avoid spatial aliasing, the spacing must be less than half of the wavelength of the signals. Two important applications of array signal processing are DOA estimation and beamforming.

The first algorithm concerns the DOA estimation of the source (the driver). This algorithm has to be able to on one hand accurately localize the sound source and on the other hand withstand the boundary reflections in the car. The simplest way for the DOA estimation is based on the steered response power (SRP) obtained using the delay-sum method. In the method, the received signals are delayed to compensate for the difference in arrival time at each microphone such that all signals are time-aligned in phase with respect to a given direction and summed together to form a single output.

$$y(t, \theta) = \sum_{n=1}^N x_n(t + \Delta_n), \quad (4.35)$$

where N is the number of array elements, $x_n(t)$ is the time signal received by the n -th microphone, Δ_n is the steering delay appropriate for steering the beam of the array to the angle θ . Ideally, with no additive noise and channel effects, the output

of the delay-sum beamformer is optimal and represents a scaled and potentially delayed version of the desired signal. In practice, channel effects and the additive noise can be nontrivial. The delay-sum concept can be extended to the more general filter-and-sum approach. Define the spatial power spectrum

$$Y(\omega, \theta) = \sum_{n=1}^N G_n(\omega) X_n(\omega) e^{j\omega\Delta_n}, \quad (4.36)$$

where $X_n(\omega)$ and $G_n(\omega)$ are the Fourier transforms of the signal received at the n -th microphone and its associated array filter, respectively. The output power of the array can be obtained by integrating $Y(\omega, \theta)$ over the frequency band of interest.

$$S(\theta) = \int_{-\infty}^{+\infty} |Y(\omega, \theta)|^2 d\omega, \quad (4.37)$$

The maximum magnitude of the output power corresponds presumably the DOA.

The aforementioned algorithms are based on the free field assumption. In practice, the performance of DOA could be degraded in reverberant environments such as the car interior considered in the paper. It is then desirable to develop a DOA method that is less susceptible to coherent multipath reflections from the boundaries. To this end, an idea based whitening filters in the steered response power-phase transform (SRP-PHAT) beamformer is employed in the DOA estimation. The signal received at each microphone is whitened by its own power spectral density before the DOA processing. Similar to the delay-sum algorithm, the array output power $S(\theta)$ is maximized over a region of the potential source location. Why this simple operation will improve the DOA estimation in a reverberant environment will be explained as follows.

Define the cross correlation of signals $x_i(t)$ and $x_j(t)$ received at the i -th and j -th microphone:

$$c_{ij}(\tau) = \int_{-\infty}^{\infty} x_i(t)x_j(t+\tau)dt. \quad (4.38)$$

The peak of the cross-correlation function indicates the time difference of arrival of the signals picked up by the two microphones. If filters are used to process the received signals, with the Fourier transforms of these filters denoted by $G_i(\omega)$ and $G_j(\omega)$, respectively, the generalized cross-correlation (GCC) can be expressed in terms of the Fourier transforms of the microphone signals

$$R_{ij}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} (G_i(\omega)X_i(\omega))(G_j(\omega)X_j(\omega))^* e^{j\omega\tau} d\omega \quad (4.39)$$

Suppose we now choose zero-phase filters with FRFs:

$$G_n(\omega) = \frac{1}{|X_n(\omega)|}. \quad (4.40)$$

This in effect “whitens” the spectra and ends up with only phase differences. In other words, the GCC reduces to a sharp peak (δ) at the time difference (τ_{ij}) between the two microphone signals:

$$R_{ij}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{X_i(\omega)X_j^*(\omega)}{|X_i(\omega)X_j^*(\omega)|} e^{j\omega\tau} d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-j\omega\tau_{ij}} e^{j\omega\tau} d\omega = \delta(\tau - \tau_{ij}) \quad (4.41)$$

As a result, by enriching signals spectrally by the whitening filter, the time difference carrying the DOA information is emphasized with improved spatial resolution, even in the face of background reverberation.

It should be noted that, however, the robust DOA algorithm must be implemented in the frequency-domain by block processing because of the zero-phase whitening filter which is also signal-dependent.

In this section, the design of the superdirective beamformer is introduced. Assume that the signal $r(t)$ received at the reference microphone is a narrowband signal with the center frequency ω_c

$$r(t) = s(t)e^{j\omega_c t}, \quad (4.42)$$

where $s(t)$ is the phasor of $r(t)$. The signal received at the m -th microphone

located at \vec{x}_m is denoted as $x_m(t)$, and let \vec{r} be the unit vector pointing to the source direction. If the speed of sound is c , the signal $x_m(t)$ can be written as

$$x_m(t) \approx s(t)e^{j\omega_c \frac{\vec{x}_m \cdot \vec{r}}{c}} e^{j\omega_c t} + n_m(t), \quad (4.43)$$

where $n_m(t)$ is the noise signal of the m -th component in the array. For M sensor signals $x_1(t), \dots, x_M(t)$, the vector of received signals can be written as

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_M(t) \end{bmatrix} = \begin{bmatrix} e^{j\omega_c \frac{\vec{x}_1 \cdot \vec{r}}{c}} \\ \vdots \\ e^{j\omega_c \frac{\vec{x}_M \cdot \vec{r}}{c}} \end{bmatrix} s(t)e^{j\omega_c t} + \begin{bmatrix} n_1(t) \\ \vdots \\ n_M(t) \end{bmatrix} = \mathbf{d}(\vec{r})\mathbf{s}(t) + \mathbf{n}(t). \quad (4.44)$$

Or, in the frequency-domain,

$$\mathbf{x}(e^{j\omega}) = \mathbf{s}(e^{j\omega})\mathbf{d}_s + \mathbf{n}(e^{j\omega}), \quad (4.45)$$

where \mathbf{d}_s is the vector of the look direction and \mathbf{n} is the noise vector. Let \mathbf{w} be the vector of the frequency-domain array coefficients. The array output y can be written as

$$y = \mathbf{w}^H \mathbf{x}, \quad (4.46)$$

where the operator H denotes the Hermitian operator.

The array-gain is defined as the ratio of the SNR of one sensor element over that of the array.

$$G = \frac{SNR_{Array}}{SNR_{Sensor}}. \quad (4.47)$$

The SNR of one sensor is given by the ratio of the power spectral densities (PSD) of the signal ϕ_{ss} and the noise ϕ_{nn} . Assuming a spherically isotropic and uncorrelated noise field, the array gain can be shown to be

$$G = \frac{|\mathbf{w}^H \mathbf{d}_s|^2}{\mathbf{w}^H \mathbf{\Gamma}_{nn} \mathbf{w}}, \quad (4.48)$$

where $\mathbf{\Gamma}_{nn}$ is the coherence matrix of a diffuse noise field whose entry ab is

$\Gamma_{n_a n_b}(e^{j\omega})|_{Diffuse} = \frac{\sin[k(a-b)d]}{k(a-b)d}$, k is the wave number, and d is the spacing between adjacent microphones. Expressing the above array gain in dB scale leads to the definition of the directivity index (DI) which indicates the ability of a beamformer to suppress noise from all but the principal direction.

$$DI = 10 \log_{10} \left(\frac{|\mathbf{w}^H \mathbf{d}_s|^2}{\mathbf{w}^H \mathbf{\Gamma}_{nm} \mathbf{w}} \right) \quad (4.49)$$

The superdirective design is achieved by maximizing the DI of array, which is equivalent to the following constrained optimization problem:

$$\begin{aligned} \min_{\mathbf{w}} \quad & \mathbf{w}^H \mathbf{\Gamma}_{nm} \mathbf{w} \\ \text{subject to} \quad & \mathbf{w}^H \mathbf{d}_s = 1 \end{aligned} \quad (4.50)$$

The solution to the above problem is the well-know MVDR beamformer:

$$\mathbf{w} = \frac{\mathbf{\Gamma}_{nm}^{-1} \mathbf{d}_s}{\mathbf{d}_s^H \mathbf{\Gamma}_{nm}^{-1} \mathbf{d}_s}. \quad (4.51)$$

In case of nearly singular coherence matrices, a more robust version of this formula with the introduction of a regularization constant ε is given as

$$\mathbf{w}|_{\text{Regularized}} = \frac{(\mathbf{\Gamma}_{nm} + \varepsilon \mathbf{I})^{-1} \mathbf{d}_s}{\mathbf{d}_s^H (\mathbf{\Gamma}_{nm} + \varepsilon \mathbf{I})^{-1} \mathbf{d}_s} \quad (4.52)$$

The constant ε can vary from zero to infinity, which corresponds to the ideally superdirective beamformer and the delay-sum beamformer, respectively. The Fig. 35 (a) and Fig. 35 (b) show the effects of ε varying from 0.001 to 3 on the DI and the 2-norm of the optimal weights in a 4-element broadside array, respectively. By ‘‘broadside,’’ we mean that the main beam of the array is perpendicular to the array axis. It can be observed that the regularized superdirective array does have higher directivity than the delay-sum array up to the frequency 3 kHz, where the constant ε can be used as a means to reconcile the DI and the array weights. As ε decreases, the DI and the array weights increase at low frequencies. Exceedingly large array weights will be difficult to implement in the filter design. Experience suggests that

$\varepsilon = 0.01$ is an adequate choice.

The Fig. 36 shows the contour plots to compare the beam patterns of the delay-sum array and the superdirective array with $\varepsilon = 0.01$. The beam width of the superdirective array is distinctively narrower than the delay-sum array. The Fig. 37 compares the frequency responses obtained from simulation and measurement of the superdirective array filters. Since the filter coefficients are symmetrical, the frequency responses are plotted in pairs. In addition, a phase flipping phenomenon can be seen in the phase responses of the microphones 1 and 2, which is similar to the behavior of differential microphones.

Once the DOA of the source (the driver) is found, the orientation of the main beam of array can be electronically steered to the desired look direction. This is accomplished by introducing appropriate phase delays to the array channels. In the frequency-domain, the beamsteering filter can be written as

$$w_n = e^{-j \frac{n\omega f_s d \sin\theta_M}{c}}, \quad (4.53)$$

where n is array index, ω is the digital frequency, d is the spacing between adjacent elements, θ_M is the steering angle, and f_s is the sampling frequency. However, $f_s d \sin\theta_M / c$ generally is tantamount to fractional samples, which requires interpolation of some sort. In the present work, the second-order Lagrange interpolation is used to implement the fractional delays for beamsteering.

The Fig. 38 shows the directional response of the superdirective array steered to -20 degrees for 1, 2, 4 kHz. We see a main lobe appear at -20 degree, as desired. The mirror image also appears at 200 degrees due to the axial symmetry of the array.

Table 4 The first-order differential microphone designs.

Microphone type	DI (dB)	FBR (dB)	3dB Beamwidth	Nulls (degrees)
Dipole	4.77	0.00	90.00°	90.00
Cardioid	4.77	8.45	131.06°	180.00
Hypercardioid	6.02	8.45	104.90°	109.47
Supercardioid	5.72	11.44	114.90°	125.26



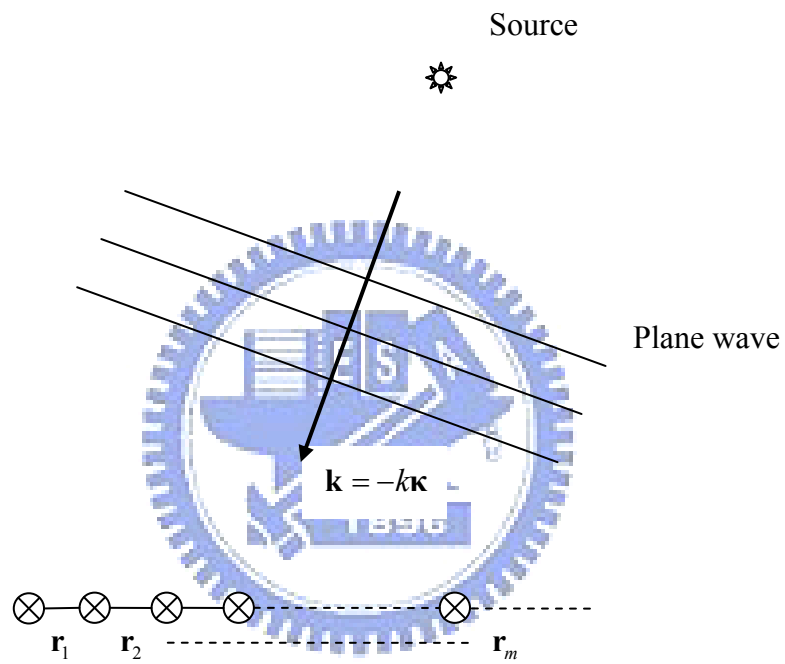


Fig. 29 A plane wave incident from the direction $\boldsymbol{\kappa}$ to a farfield array.

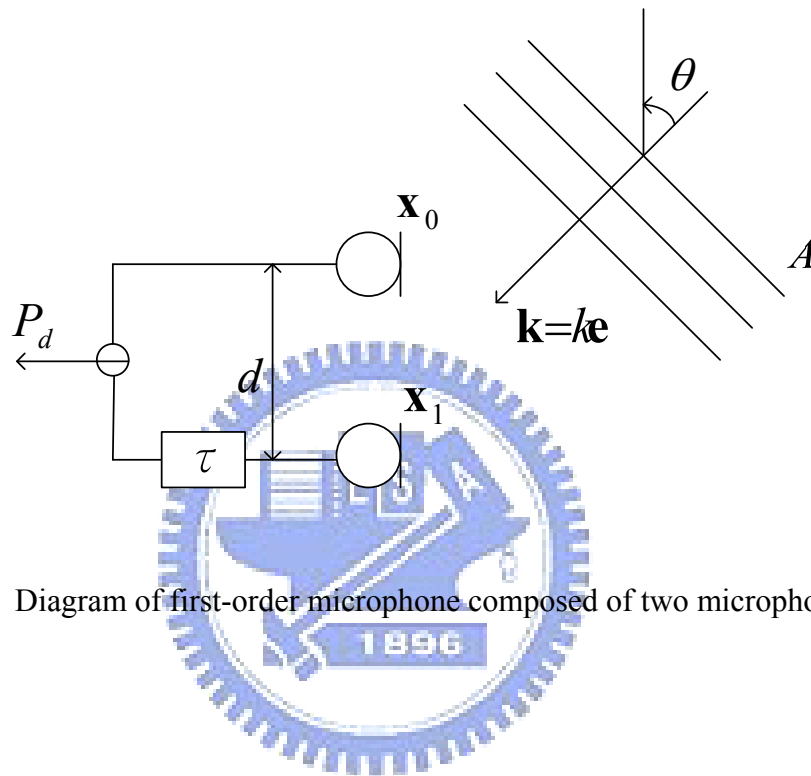


Fig. 30 Diagram of first-order microphone composed of two microphones.

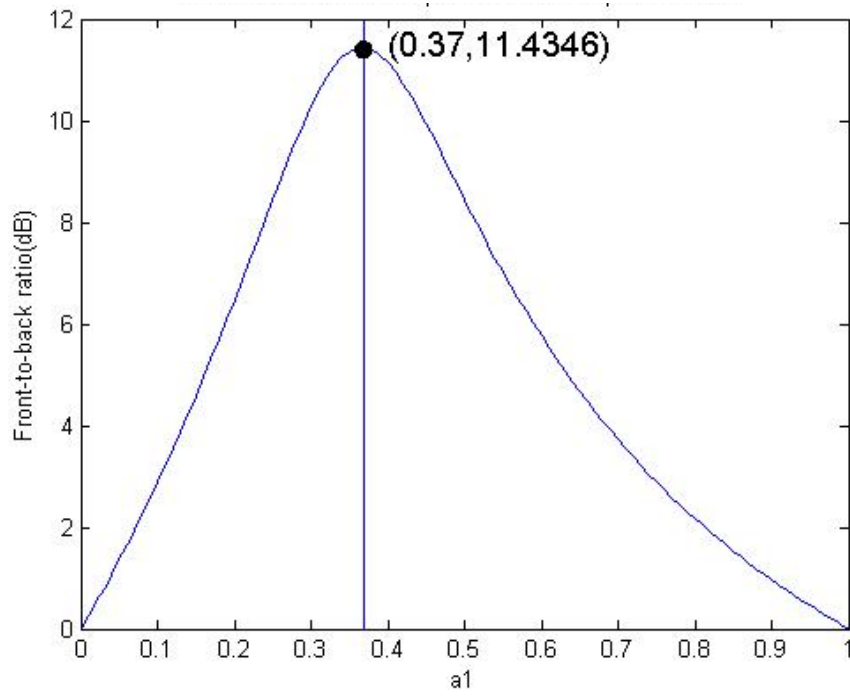
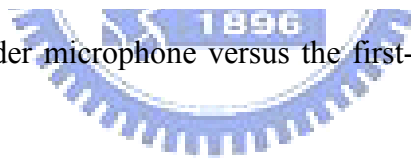


Fig. 31 FBR of first-order microphone versus the first-order differential parameter

α_1 .



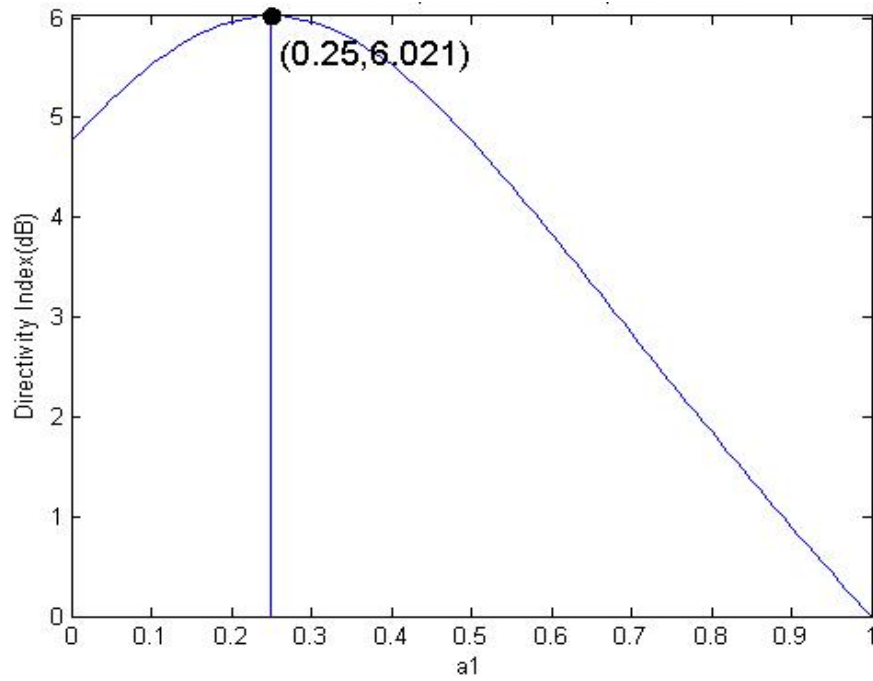
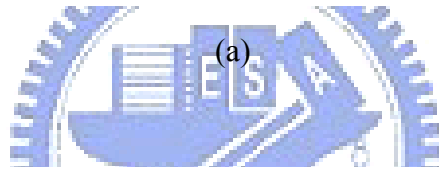
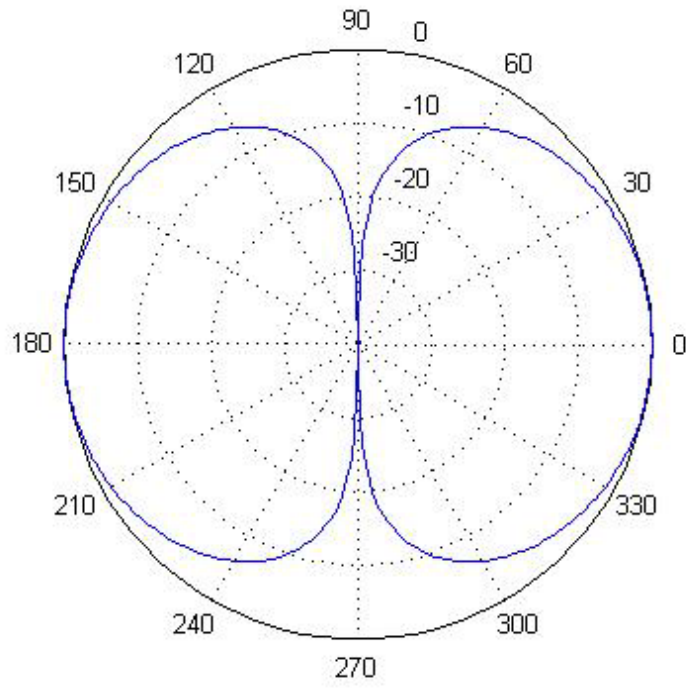
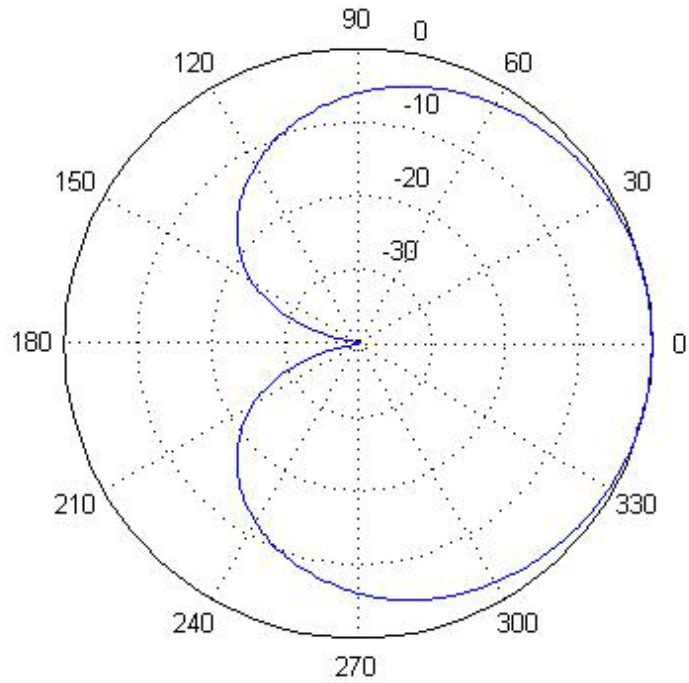


Fig. 32 Directivity index of first-order microphone versus the first-order differential parameter α_1 .

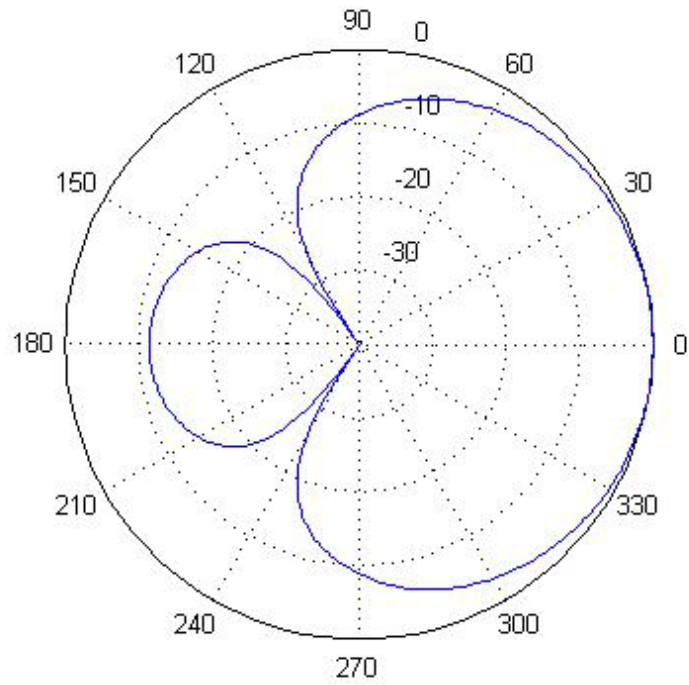
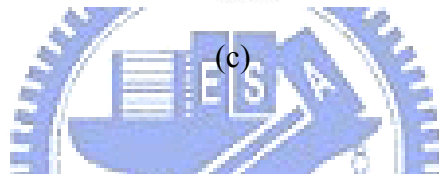
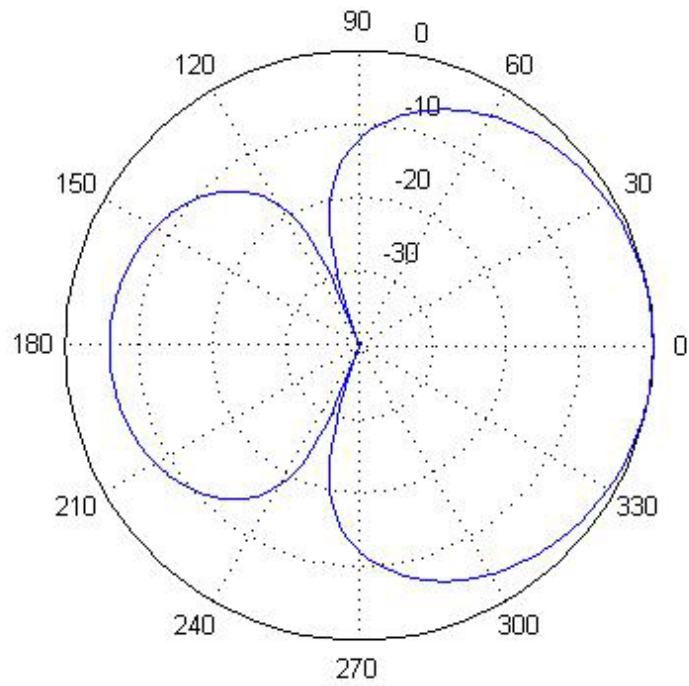




(a)



(b)



(d)

Fig. 33 Various first-order directional responses (a) dipole, (b) cardioids, (c) hypercardioid, (d) supercardioid.

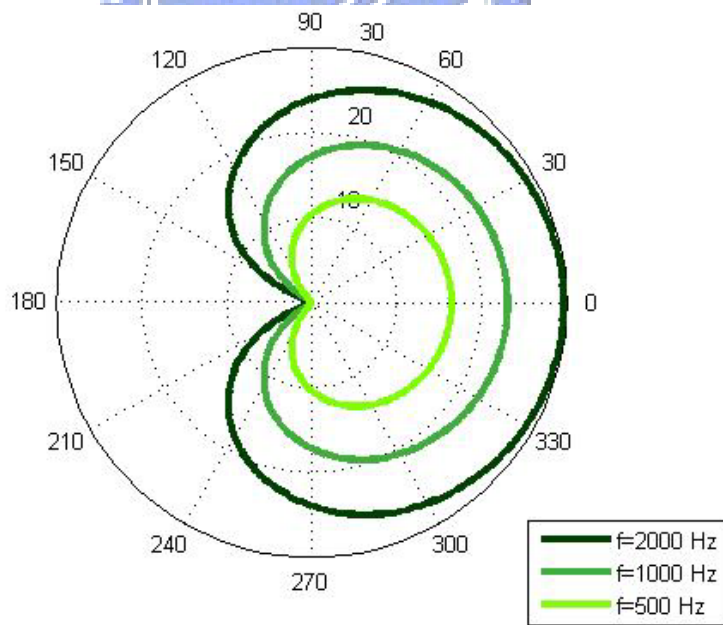
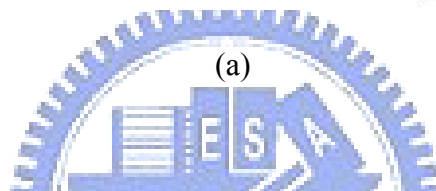
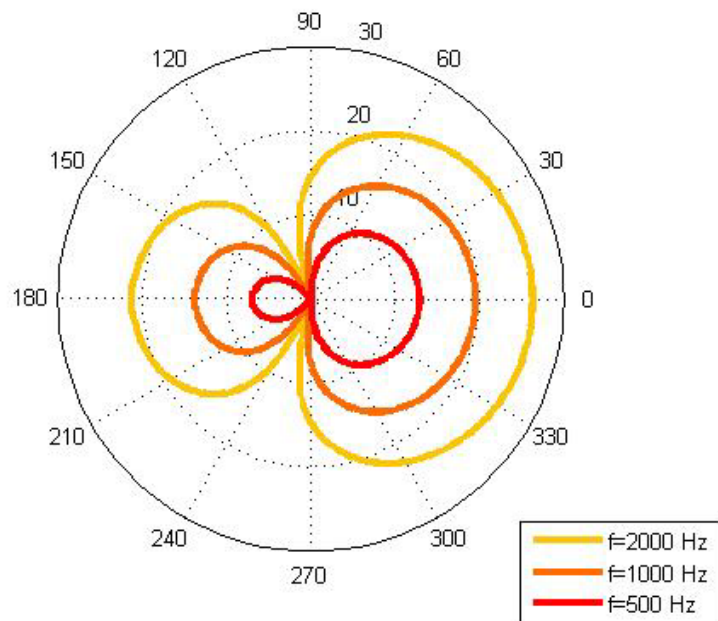
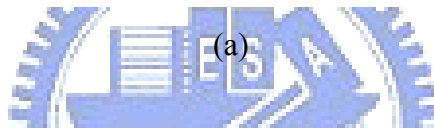
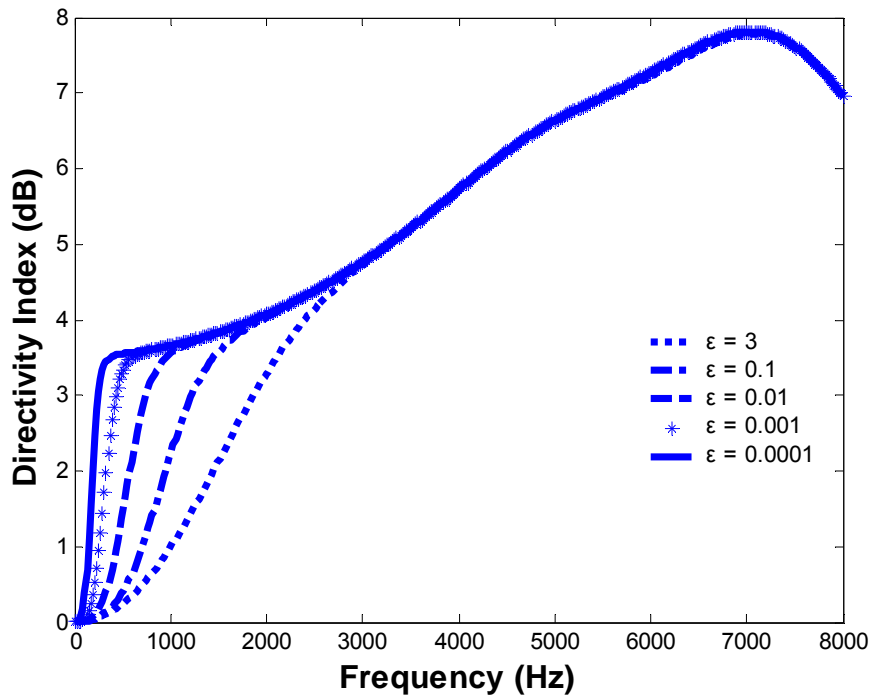
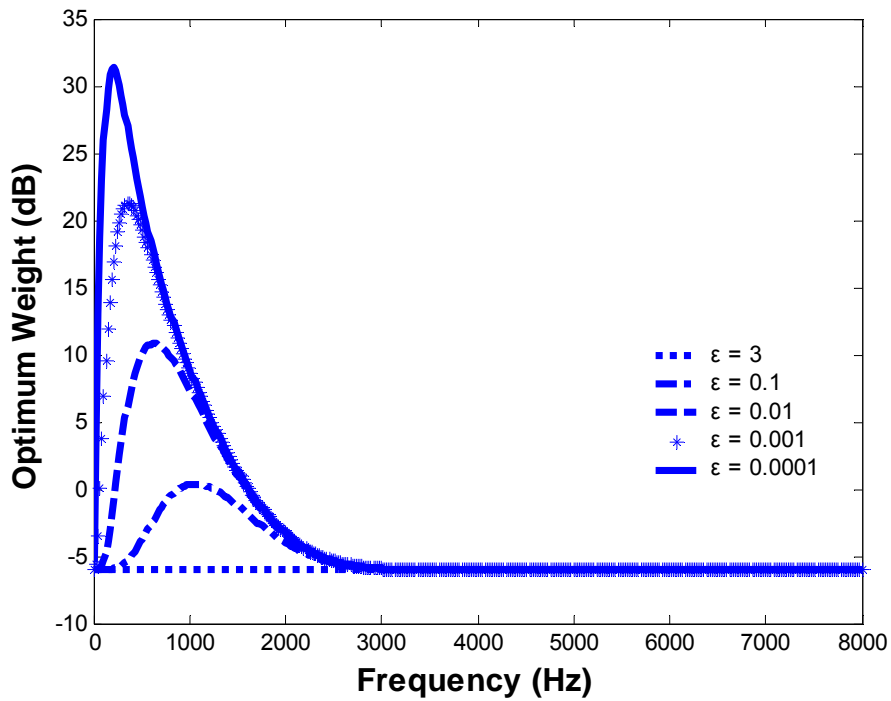


Fig. 34 The directivity pattern of first-order DMAs. (a) $\alpha_1=0.25$, (b) $\alpha_1=0.5$.



(a)



(b)

Fig. 35 The effects of the regularization constant ϵ varying from 0.001 to 3 of a broadside array. (a) DI, (b) the 2-norm of the optimal array weights.

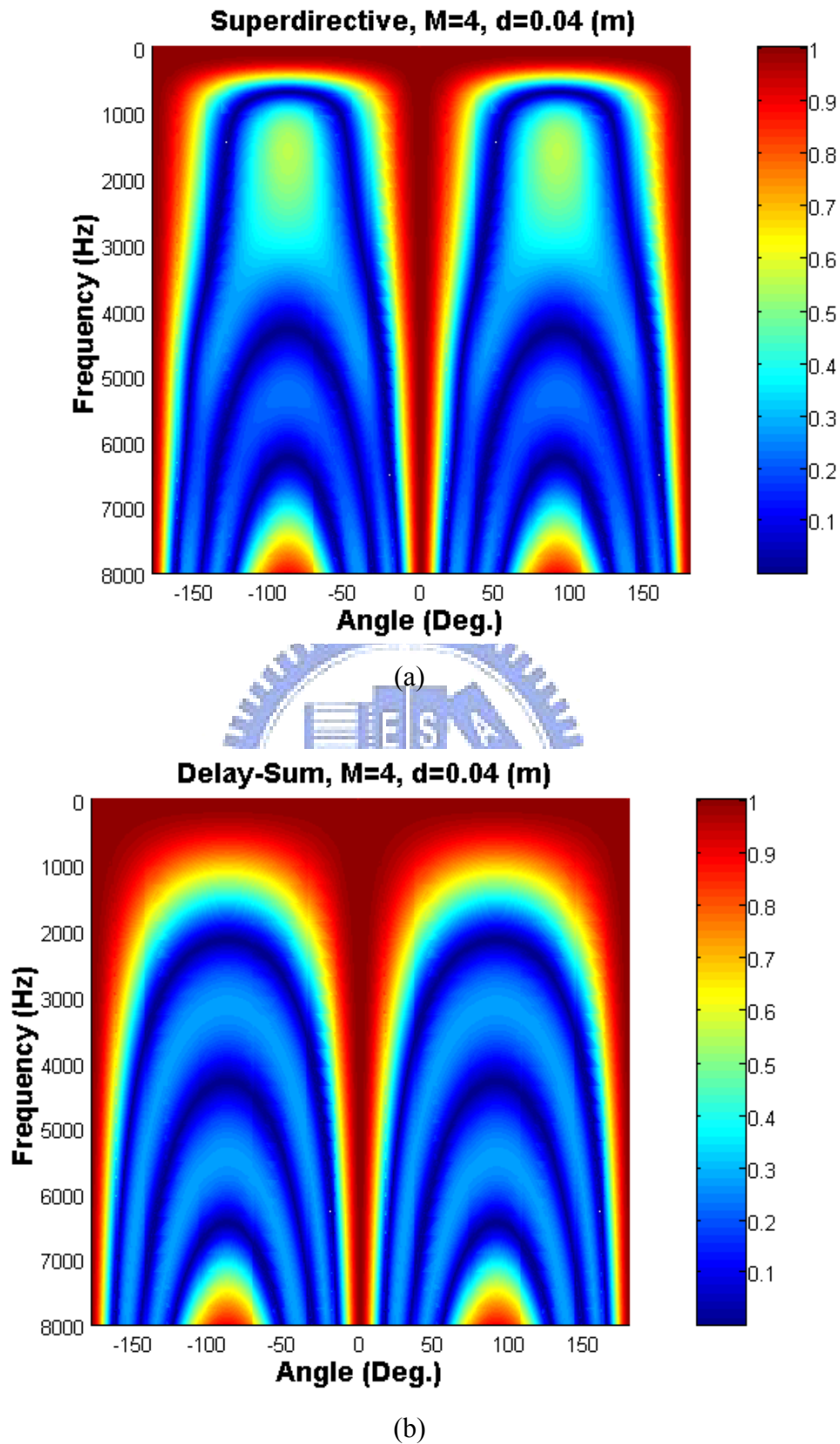
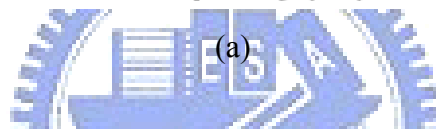
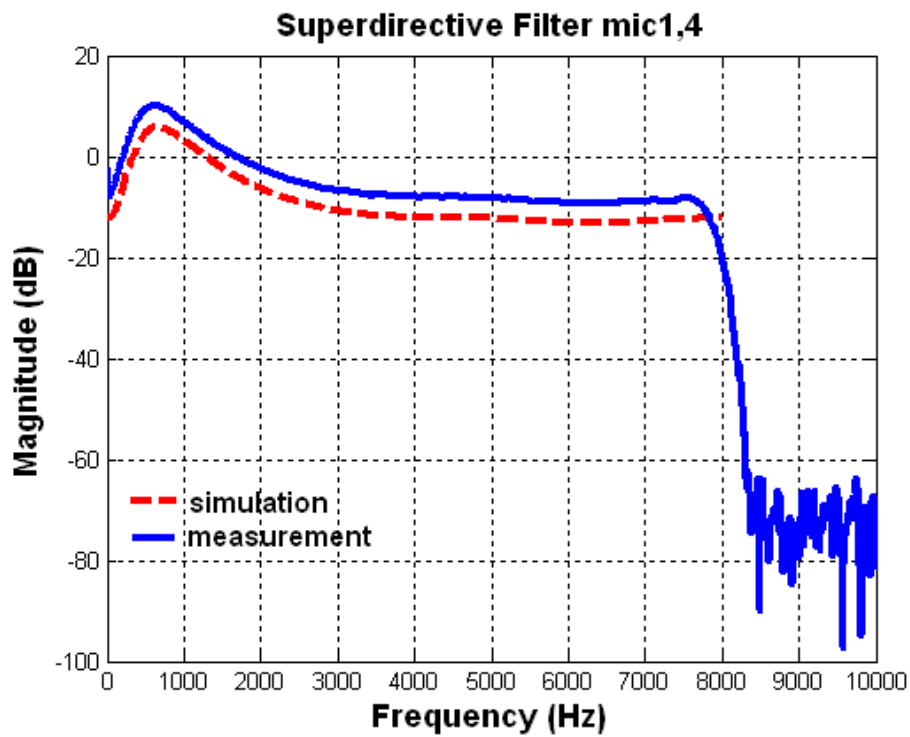


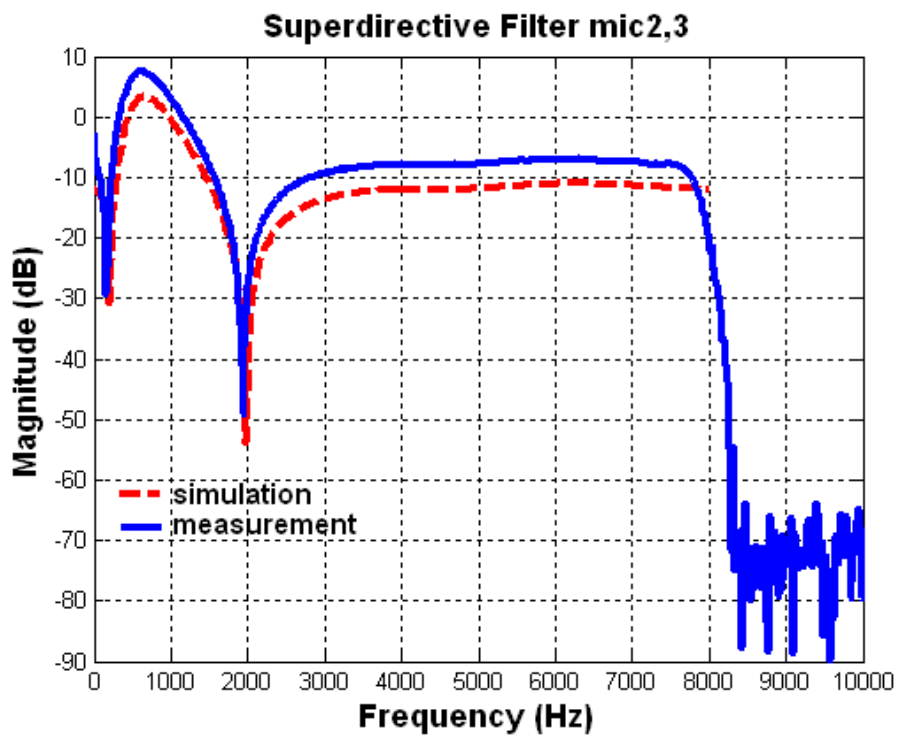
Fig. 36 The contour plots of beam patterns of a broadside array. The x -axis represents the angle in degrees and the y -axis represents the frequency in

kHz. (a) Superdirective array, (b) DAS array.

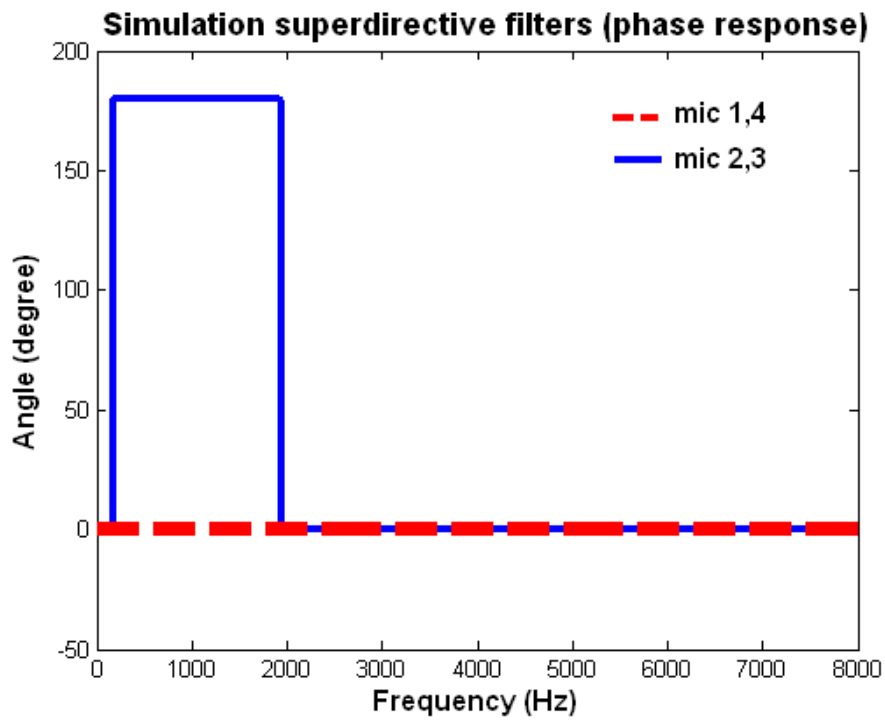




(a)



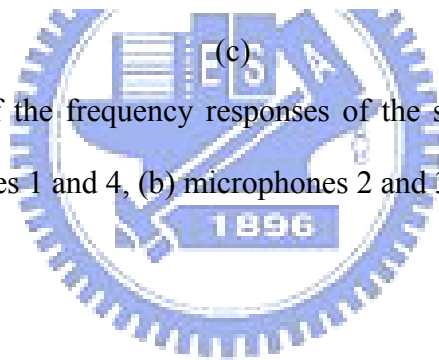
(b)

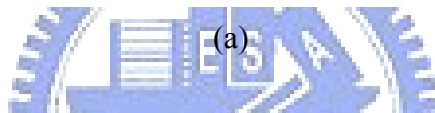
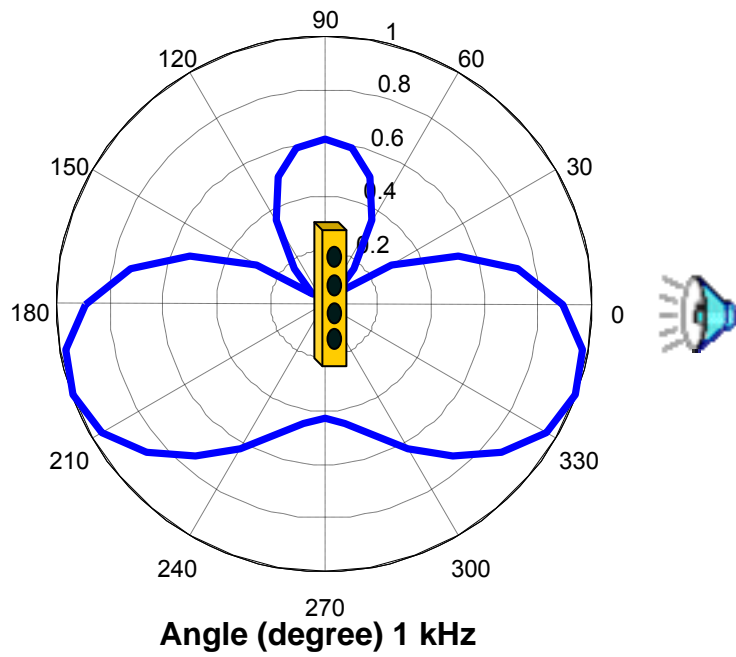


(c)

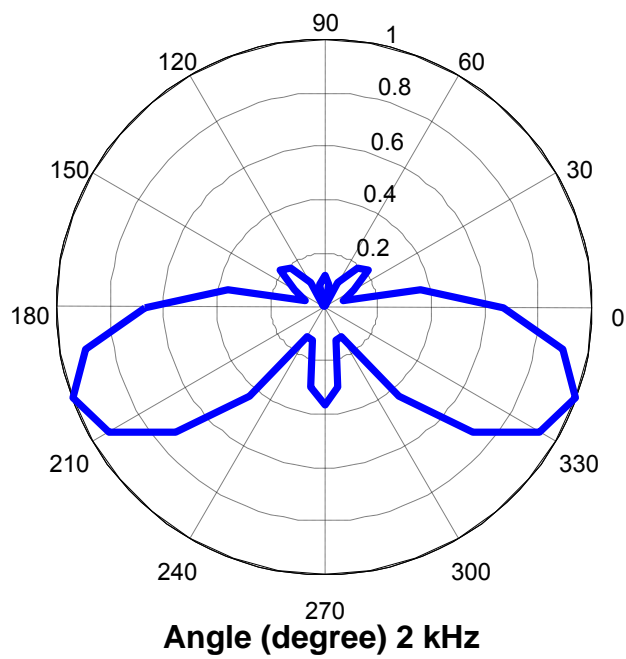
Fig. 37 Comparison of the frequency responses of the superdirective array filters.

(a) Microphones 1 and 4, (b) microphones 2 and 3, (c) phase responses.





(a)



(b)

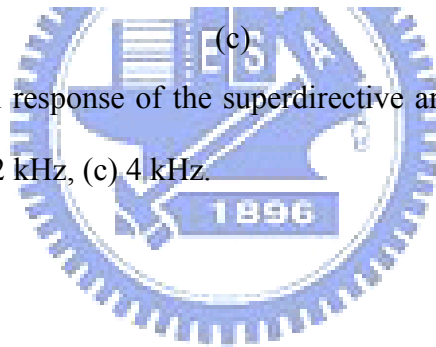
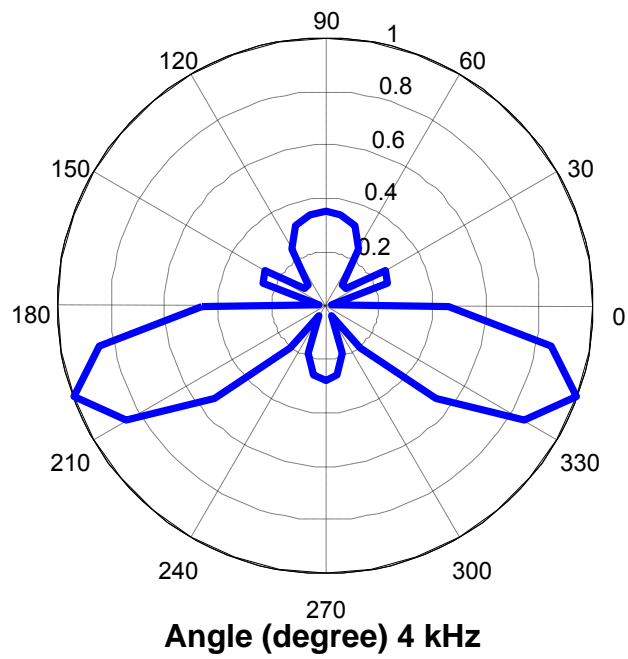
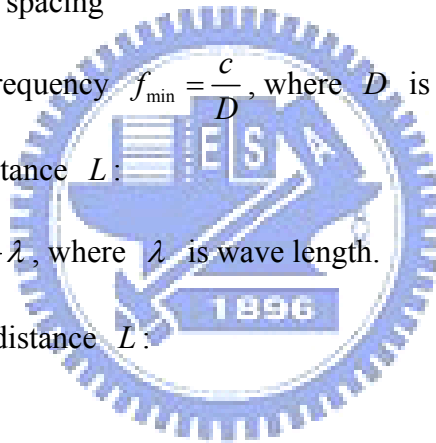


Fig. 38 The directional response of the superdirective array steered to -20 degrees.
 (a) 1 kHz, (b) 2 kHz, (c) 4 kHz.

4.1.5 Choice of farfield array parameters

In this section, the farfield array parameters (as shown in Fig. 39) can be summarized as follows:

1. Array:
 - a. Choose the microphone spacing d according to the maximum frequency (f_{\max}). A conservative rule is $d = \lambda/2$, where λ is wave length.
2. Frequency range:
 - a. Maximum frequency $f_{\max} = \frac{c}{2d}$, where c is speed of sound and d is microphone spacing
 - b. Minimum frequency $f_{\min} = \frac{c}{D}$, where D is array aperture.
3. Resolution of distance L :
 - a. $R_L = 1.22 \frac{L}{D} \lambda$, where λ is wave length.
4. Area covered at distance L :
 - a. $z = 1.15L$.



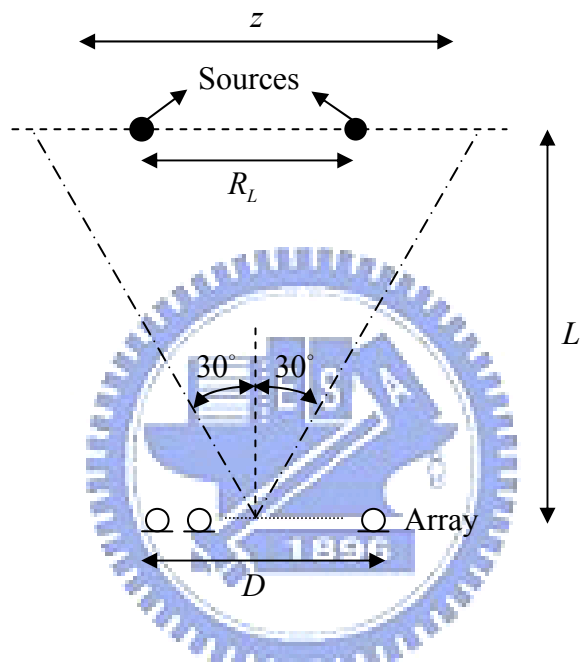


Fig. 39 Farfield array parameters.

4.2 High-resolution algorithms

4.2.1 Minimum variance distortionless response (MVDR)

Another approach has been proposed using the data covariance matrix. This method has been shown to provide higher resolution in DOA estimations than the DAS algorithm. In order to facilitate digital processing, we simultaneously sample all array inputs to form digital data $x_m(t) = x_m(kT)$, $k=1,2,\dots$. For D sources, we may invoke the principle of superposition to write

$$\begin{aligned} \mathbf{x}(k) &= \sum_{i=1}^D \mathbf{a}(\omega, \theta_i) r_i(k) + \mathbf{n}(k) = [\mathbf{a}(\omega, \theta_1) \cdots \mathbf{a}(\omega, \theta_D)] \begin{bmatrix} r_1(k) \\ \vdots \\ r_D(k) \end{bmatrix} + \mathbf{n}(k) \\ &= \mathbf{A}\mathbf{r}(k) + \mathbf{n}(k) \end{aligned} \quad (4.54)$$

where θ_i is the direction of the i th source, $\mathbf{r}(k)$ is the source signal vector and \mathbf{A} is DOA matrix. A beamformer output is a linear combiner that produces an output signal by weighting and summing all components.

$$y(k) = \sum_{m=1}^M w_m^* x_m(k) = \mathbf{w}^H \mathbf{x}(k) \quad (4.55)$$

where \mathbf{w} is the weight vector given by $\mathbf{w} = [w_1 \cdots w_M]^T$. Because the MVDR method exploits the correlation between array input signals, it is necessary to calculate the array signal correlation matrix.

$$\mathbf{R}_{xx} = E \{ \mathbf{x}(k) \mathbf{x}^H(k) \} \quad (4.56)$$

Suppose that the noise is uncorrelated with signals $E \{ \mathbf{r}(k) \mathbf{n}^H(k) \} = 0$ and the noise is spatially white $E \{ \mathbf{n}(k) \mathbf{n}^H(k) \} = \sigma_n^2 \mathbf{I}$. By the preceding assumption, the Eq. (4.56) can be rewritten as

$$\begin{aligned}
\mathbf{R}_{xx} &= \mathbf{A}E\{\mathbf{r}(k)\mathbf{r}^H(k)\} + E\{\mathbf{n}(k)\mathbf{n}^H(k)\} \\
&= \mathbf{A}\mathbf{R}_{rr}\mathbf{A}^H + \mathbf{R}_{nn} \\
&= \mathbf{A}\mathbf{R}_{rr}\mathbf{A}^H + \sigma_n^2\mathbf{I}
\end{aligned} \tag{4.57}$$

where \mathbf{R}_{rr} and \mathbf{R}_{nn} are the source and noise correlation matrices, respectively. In practice, the data correlation matrix \mathbf{R}_{xx} is usually approximated by the data covariance covariance

$$\mathbf{R}_{xx}(p) = \alpha \mathbf{x}_p^H \mathbf{x}_p + (1-\alpha)\mathbf{R}_{xx}(p-1), \quad p=1,2,\dots,P, \quad \mathbf{R}_{xx}(0) = \mathbf{0} \tag{4.58}$$

At this recursive equation, α is a constant which satisfied $0 \leq \alpha \leq 1$. The received signal is divided to p frames and rearranged to the data vector $\mathbf{x} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_p]$.

In the following, the aim is to find the MVDR weight vector \mathbf{w}_{MV} . An optimization problem is given for solving the unknown vector \mathbf{w}_{MV} . The MVDR beamformer attempts to minimize the output power

$$E\{|y(k)|^2\} = E\{|\mathbf{w}_{MV}^H \mathbf{x}(k)|^2\} = \mathbf{w}_{MV}^H \mathbf{R}_{xx} \mathbf{w}_{MV} \tag{4.59}$$

Another constraint is to maintain unity gain in the look direction $\mathbf{w}_{MV}^H \mathbf{a}(\theta_0) = 1$. The MVDR beamforming suppresses the undesired interference from $\theta \neq \theta_0$ and the noise. The problem can be expressed as follows

$$\begin{aligned}
&\min_{\mathbf{w}_{MV}} \mathbf{w}_{MV}^H \mathbf{R}_{xx} \mathbf{w}_{MV} \\
&\text{subject to } \mathbf{w}_{MV}^H \mathbf{a}(\omega, \theta_0) = 1
\end{aligned} \tag{4.60}$$

This problem can be solved by Lagrange multiplier method

$$\begin{cases} \nabla_{\mathbf{w}_{MV}} \mathbf{w}_{MV}^H \mathbf{R}_{xx} \mathbf{w}_{MV} - \lambda \nabla_{\mathbf{w}_{MV}} [\mathbf{w}_{MV}^H \mathbf{a}(\omega, \theta_0) - 1] = 0 \\ \mathbf{w}_{MV}^H \mathbf{a}(\omega, \theta_0) = 1 \end{cases} \tag{4.61}$$

If \mathbf{R}_{xx} is nonsingular, \mathbf{R}_{xx} can be inverted to solve the unknown vector by

$$\mathbf{w}_{MV} = \lambda \mathbf{R}_{xx}^{-1} \mathbf{a}(\omega, \theta_0), \tag{4.62}$$

where $\lambda = \mathbf{w}_{MV}^H \mathbf{R}_{xx} \mathbf{w}_{MV} = \frac{1}{\mathbf{a}^H(\omega, \theta_0) \mathbf{R}_{xx}^{-1} \mathbf{a}(\omega, \theta_0)}$ is the beamformer output power.

Next, λ is substituted in Eq. (4. 62) to obtain the MVDR weight

$$\mathbf{w}_{MV} = \frac{\mathbf{R}_{xx}^{-1} \mathbf{a}(\omega, \theta_0)}{\mathbf{a}^H(\omega, \theta_0) \mathbf{R}_{xx}^{-1} \mathbf{a}(\omega, \theta_0)} \quad (4. 63)$$

In the preceding results, it is convenient to obtain the spatial power spectrum $S_{MV}(\theta)$ by continuing altering θ

$$S_{MV}(\theta) = \mathbf{w}_{MV}^H \mathbf{R}_{xx} \mathbf{w}_{MV} = \frac{1}{\mathbf{a}^H(\omega, \theta) \mathbf{R}_{xx}^{-1} \mathbf{a}(\omega, \theta)} \quad (4. 64)$$

The spatial power spectrum $S_{MV}(\theta)$ exhibits J peaks approximately at $\theta_1 \cdots \theta_D$.

4.2.2 Multiple signal classification (MUSIC)

The MUSIC approach of DOA estimation has been proposed by exploiting the eigenvalue decomposition (EVD) of the covariance matrix. Firstly, the array covariance matrix \mathbf{R}_{xx} in Eq. (4. 58) is represented by EVD

$$\mathbf{R}_{xx} = \mathbf{A} \mathbf{R}_{rr} \mathbf{A}^H + \sigma_n^2 \mathbf{I} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^{-1} \quad (4. 65)$$

where \mathbf{U} is a unitary matrix and comprise M linearly independent eigenvectors $\mathbf{u}_1 \cdots \mathbf{u}_M$. The eigenvector associate with M eigenvalues $\alpha_1 \cdots \alpha_M$. The array correlation matrix can be represented as

$$\begin{aligned} \mathbf{R}_{xx} &= \mathbf{U} \mathbf{\Lambda} \mathbf{U}^{-1} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^H \\ &= [\mathbf{u}_1 \ \mathbf{u}_2 \ \cdots \ \mathbf{u}_M] \begin{bmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_M \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^H \\ \mathbf{u}_2^H \\ \vdots \\ \mathbf{u}_M^H \end{bmatrix} = \sum_{m=1}^M \alpha_m \mathbf{u}_m \mathbf{u}_m^H \end{aligned} \quad (4. 66)$$

The diagonal terms of $\mathbf{\Lambda}$ have been arranged with $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_M$. The noise term $\sigma_n^2 \mathbf{I}$ can be written as

$$\sigma_n^2 \mathbf{I} = \sigma_n^2 \mathbf{U} \mathbf{U}^{-1} = \sigma_n^2 \mathbf{U} \mathbf{U}^H = \sigma_n^2 \sum_{m=1}^M \mathbf{u}_m \mathbf{u}_m^H \quad (4. 67)$$

Because \mathbf{A} is derived from D sources, we assume that \mathbf{A} and \mathbf{R}_{rr} are of full rank D . Subsequently the signal-only correlation matrix \mathbf{C}_{xx} is generated by subtracting the noise component from \mathbf{R}_{xx}

$$\mathbf{C}_{xx} = \mathbf{A}\mathbf{R}_{rr}\mathbf{A}^H = \sum_{m=1}^M (\alpha_m - \sigma_n^2) \mathbf{u}_m \mathbf{u}_m^H \quad (4.68)$$

If \mathbf{R}_{rr} is rank D and small than the array size M , the smallest $M - D$ eigenvalues $\alpha_{D+1} \cdots \alpha_M$ are equivalent to the noise power. Therefore the range of \mathbf{C}_{xx} are spanned by \mathbf{u}_1 to \mathbf{u}_D . If the array has no coherent source between any of two received signals, \mathbf{R}_{rr} only has nonzero values on the diagonal terms which represent the power of the D sources. Note that the range of \mathbf{C}_{xx} is identical to the range of \mathbf{A} which is spanned by the manifold vectors $\mathbf{a}(\omega, \theta_1) \cdots \mathbf{a}(\omega, \theta_D)$. The relation between

\mathbf{C}_{xx} and \mathbf{A} is

$$R\{\mathbf{A}\} = \text{span}\{\mathbf{a}(\omega, \theta_1), \cdots, \mathbf{a}(\omega, \theta_D)\} = \text{span}\{\mathbf{u}_1, \cdots, \mathbf{u}_D\} \quad (4.69)$$

and

$$R\{\mathbf{A}\}^\perp = \text{span}\{\mathbf{u}_{D+1}, \cdots, \mathbf{u}_M\}, \quad (4.70)$$

where $\text{span}\{\mathbf{u}_1, \cdots, \mathbf{u}_D\}$ and $\text{span}\{\mathbf{u}_{D+1}, \cdots, \mathbf{u}_M\}$ are called the signal subspace and noise subspace, respectively. Because the subspace is orthogonal to the noise subspace such that

$$\mathbf{u}_m^H \mathbf{a}(\omega, \theta_{d_s}) = 0, \quad d_s = 1, 2, \cdots, D; \quad m = D+1, D+2, \cdots, M \quad (4.71)$$

The MUSIC technique is to exploit Eq. (4.71) to improve the DOA estimations.

The eigenvectors $\mathbf{u}_{D+1}, \cdots, \mathbf{u}_M$ is used to construct the projection matrix as follows

$$\sum_{m=D+1}^M \mathbf{u}_m \mathbf{u}_m^H = \mathbf{P}_N \quad (4.72)$$

From Eq. (4.70), the direction of the source θ_i ($i = 1, \cdots, D$) can be found by solving

$$\mathbf{P}_N \mathbf{a}(\omega, \theta) = \sum_{m=J+1}^M \mathbf{u}_m \mathbf{u}_m^H \mathbf{a}(\omega, \theta) = \mathbf{0}, \quad \theta = \theta_{d_s} \quad (4.73)$$

The projection matrix has the properties of $\mathbf{P}_N^2 = \mathbf{P}_N$ and $\mathbf{P}_N^H = \mathbf{P}_N$. The problem of Eq. (4.73) can be extended to solve

$$\|\mathbf{P}_N \mathbf{a}(\omega, \theta)\|_2^2 = \mathbf{a}(\omega, \theta)^H \mathbf{P}_N^H \mathbf{P}_N \mathbf{a}(\omega, \theta) = 0, \quad \theta = \theta_i \quad (4.74)$$

Equivalently, the inverse of Eq. (4.74) has the infinitely value when $\theta = \theta_i, i = 1, \dots, D$. The inverse of Eq. (4.74) is referred to as the MUSIC spectrum.

$$S_{MU}(\theta) = \frac{1}{\mathbf{a}^H(\omega, \theta) \mathbf{P}_N \mathbf{a}(\omega, \theta)} \quad (4.75)$$

The peaks of the MUSIC spectrum indicate the directions of sources. Note that the MUSIC spectrum does not exhibit infinitely high peaks due to noises in practice.

4.2.3 Choice of parameters: Akaike information criterion (AIC)

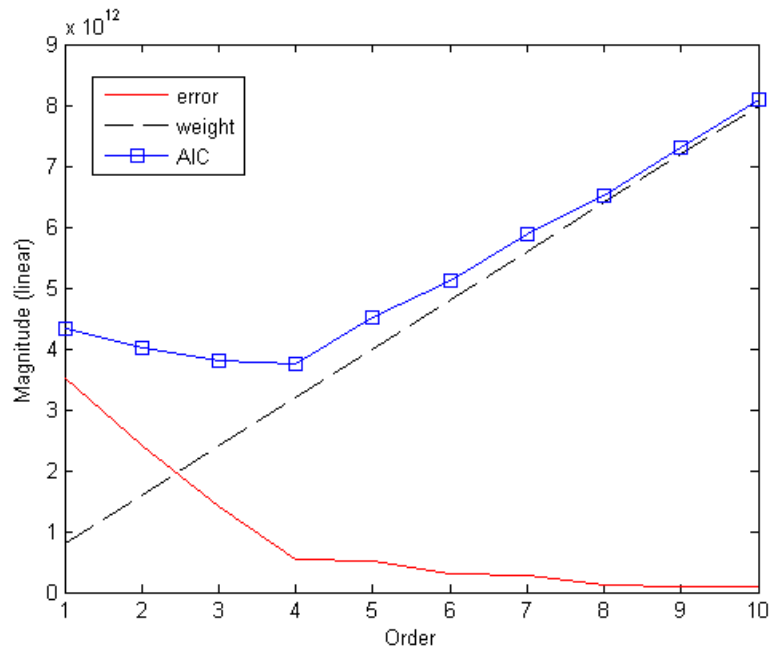
How to determine D of Eq. (4.71) is an important issue. It would rather be overestimated than underestimated. The AIC [72] can be employed to choose D . The notion of the AIC is to calculate matching error and weight the truncated order, the equation can be defined as

$$AIC(m) = \|\mathbf{R}_{xx} - \mathbf{R}_{xx}^*(m)\|_F + w_A m, \quad \mathbf{R}_{xx}^*(m) = \sum_{i=1}^m \alpha_i \mathbf{u}_i \mathbf{u}_i^H \quad (4.76)$$

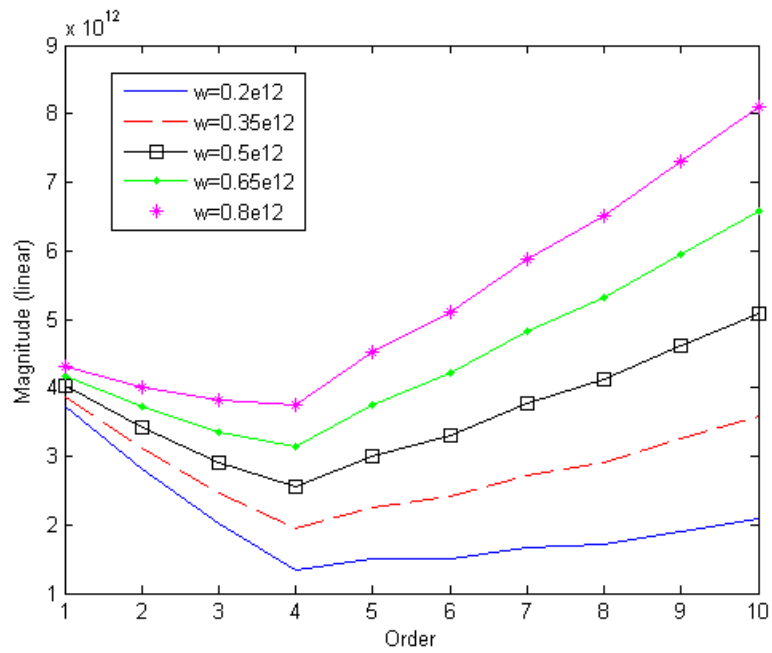
The EVD of data covariance matrix is used to calculate the matching error, which is $\|\mathbf{R}_{xx} - \mathbf{R}_{xx}^*(m)\|_F$. The weight part is to weight order by w_A in order to make the order with lowest AIC value would be the same order in the error line which has an apparent turning point. For a preset two point sources simulation, the error, weight and AIC lines are shown in Fig. 40 (a). From the figure, the turning point of error line is at the 4th order. Thus the weight w_A should be chosen to make the AIC line

will have a lowest point at that order. In our simulation, the value is chosen to be 0.5×10^{12} . The AIC line with different weight is shown in Fig. 40 (b).





(a)



(b)

Fig. 40 Decide source numbers by AIC algorithm. The truncated order is corresponded to the lowest point. (a) Error, weight and AIC lines, (b) AIC lines with different weights.

4.3 Comparison of the farfield algorithms

In order to validate and compare the DOA estimation methods, numerical simulations are conducted for a 30-channel URA and a random array optimized by the SA-IBMC optimization method. The aperture of array is $0.4\text{m}\times 0.5\text{m}$ ($d = 0.1\text{m}$) for URA and $0.5\text{m}\times 0.6\text{m}$ for random array, as shown in Fig. 41 (a) and Fig. 41 (b). In the simulation, we assume two white-noise sources located at the positions $(-0.5\text{m}, 0.5\text{m})$ and $(0.5\text{m}, -0.5\text{m})$. The sources are 1m from array surface. Assume the sound velocity c_s is 343 m/s. Consider the $\lambda/2$ rule, the maximum measurement frequency with inter-element spacing 0.1m is $f_{\max} = c/2d = 1.7\text{ kHz}$. Therefore, we choose point sources with the frequencies 1 kHz ($d = \lambda/4$) and 7 kHz ($d = 2\lambda$) to be the observed frequencies in simulations. The magnitude of beam pattern or spectrum of each approach is normalized to a range from 0 to 1. This makes the results of five methods can easily be compared in main-lobe width and side-lobes levels.

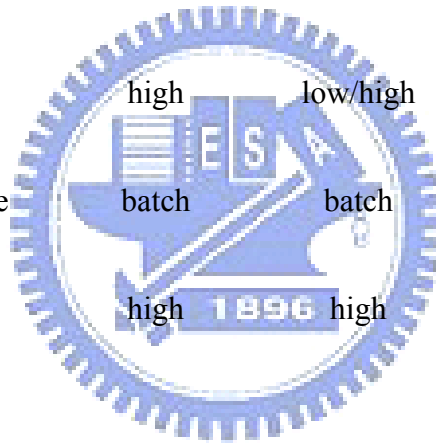
The noise maps of two simulated point sources obtained using different acoustic imaging algorithms with a URA or an optimized random array in the frequency 1 kHz, as shown in Fig. 42 (a)-Fig. 42 (j). At this frequency, the spacing is less than a half of wave length. Therefore, grating lobes are not occurring in the simulated results of URA and random configurations. The noise maps obtained using the DAS and TR algorithms are shown in Fig. 42 (a)-Fig. 42 (d). Both of these figures are with poor resolution. In addition, the source positions are not focus at the preset source positions. The noise maps of another lower resolution algorithm SIMO-ESIF is shown in Fig. 42 (e) and Fig. 42 (f). Compared with results of DAS and TR, SIMO-ESIF also has large main lobes but can correctly point the source positions. The Fig. 42 (g)-Fig. 42 (j) show the noise maps obtained using MVDR and MUSIC algorithms with two array configurations. As predicted, the results validated that the

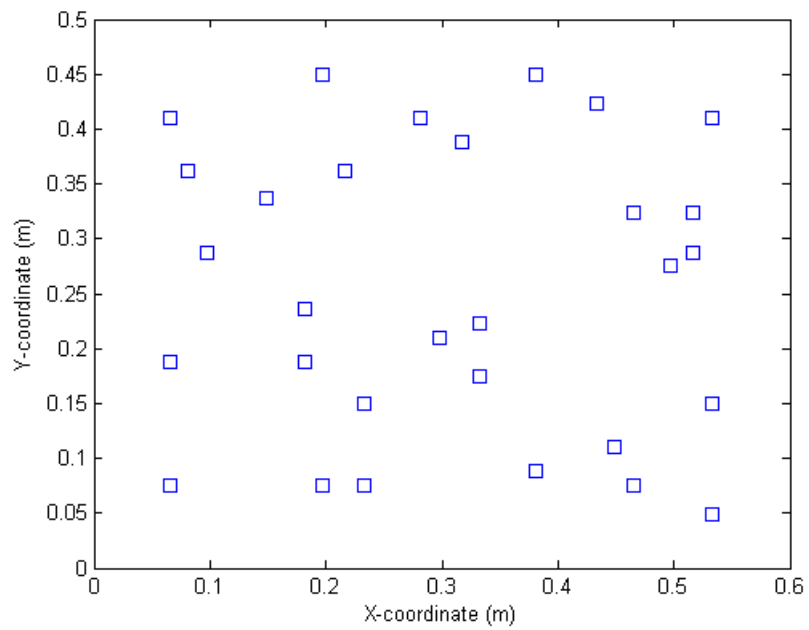
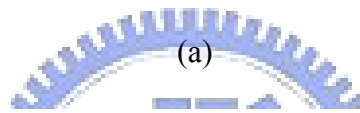
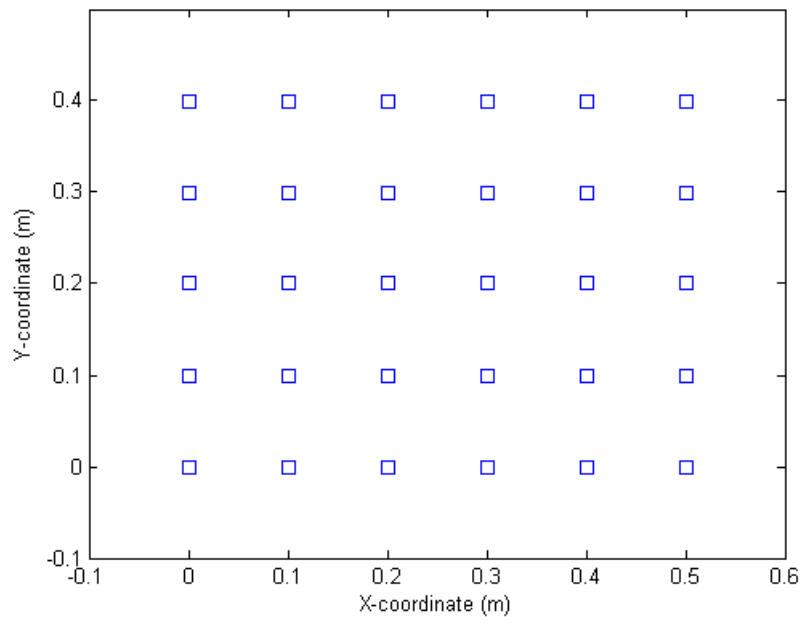
MVDR and MUSIC are the methods which can achieve higher resolutions, especially MUSIC. They can correctly localize the preset source points with narrow main lobes. The side-lobes of MVDR are higher than MUSIC.

Apart from simulations at frequency 1 kHz, we also run some simulations in a higher frequency to make the spacing exceeds a half of a wave length. Clearly, the frequency is chosen to be 7 kHz. At this frequency, the spacing is approximate two times of a length. The noise maps of two simulated point sources of DOA estimation using different approaches with a URA and an optimized random array in the frequency 7 kHz is shown in Fig. 43 (a)-Fig. 43 (j). The simulated results are largely identical but minor differences with the results in the frequency 1 kHz except the grating lobes appeared at those power spectrums with URA configuration. In SIMO-ESIF and MUSIC cases, the noise maps have no clearly visible grating lobes with URA configuration. Nevertheless, the noise maps obtained using SIMO-ESIF still have large main lobes and much higher side lobes than MUSIC. Summary, the MUSIC is the algorithm which can obtain highest resolution in the frequency from low to high. The MVDR is worse than MUSIC but still can get relatively higher resolution than other algorithms. The proposed SIMO-ESIF is the only in low resolution algorithms which can use URA to localize high frequency noise sources. Farfield acoustic imaging algorithms are compared in Table 5.

Table 5 Comparisons of farfield acoustic imaging algorithms.

	DAS	TR	SIMO-ESIF	MVDR	MUSIC
Algorithm	delay-sum	time-reversed	inverse filtering	MVDR	MUSIC
Resolution	low	low	low	High	Very high
Complexity	low	low	low	High	High
Area covered	large	large	small/large	large	large
Processing Domain	time	time	frequency	frequency	frequency
Frequency range	high	high	low/high	low/high	low/high
Sample/Batch	sample	batch	batch	batch	batch
Robustness to reverberation	poor	high	high	low	low
Acoustic variables	no	no	yes	no	no

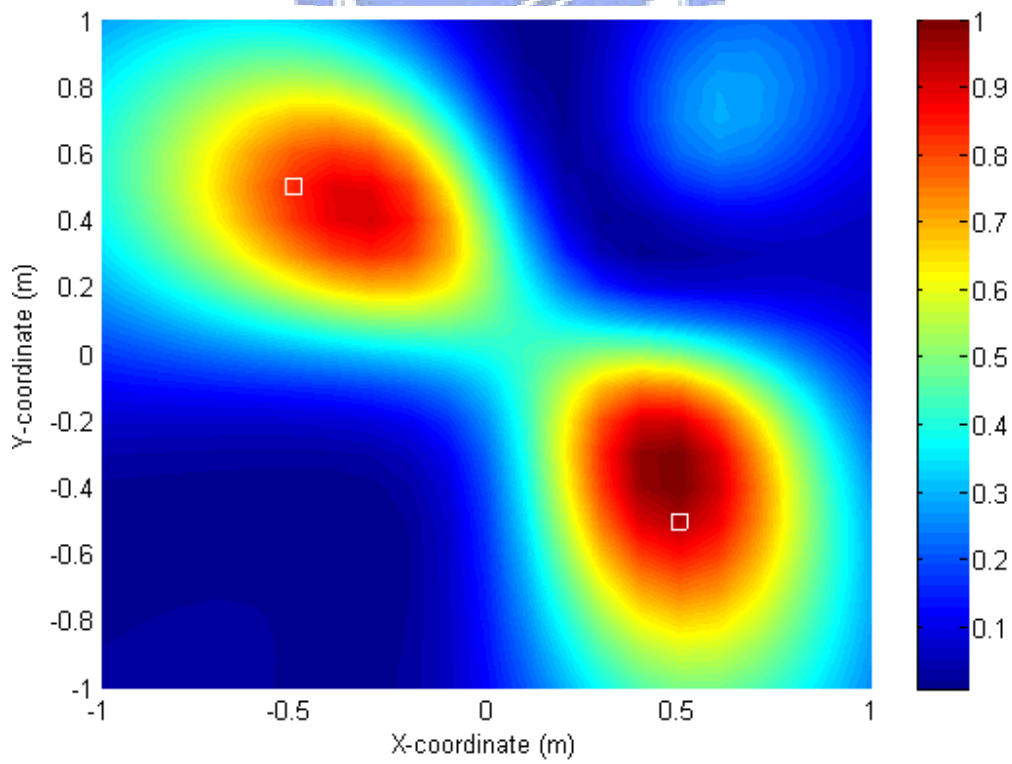
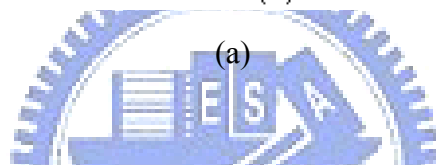
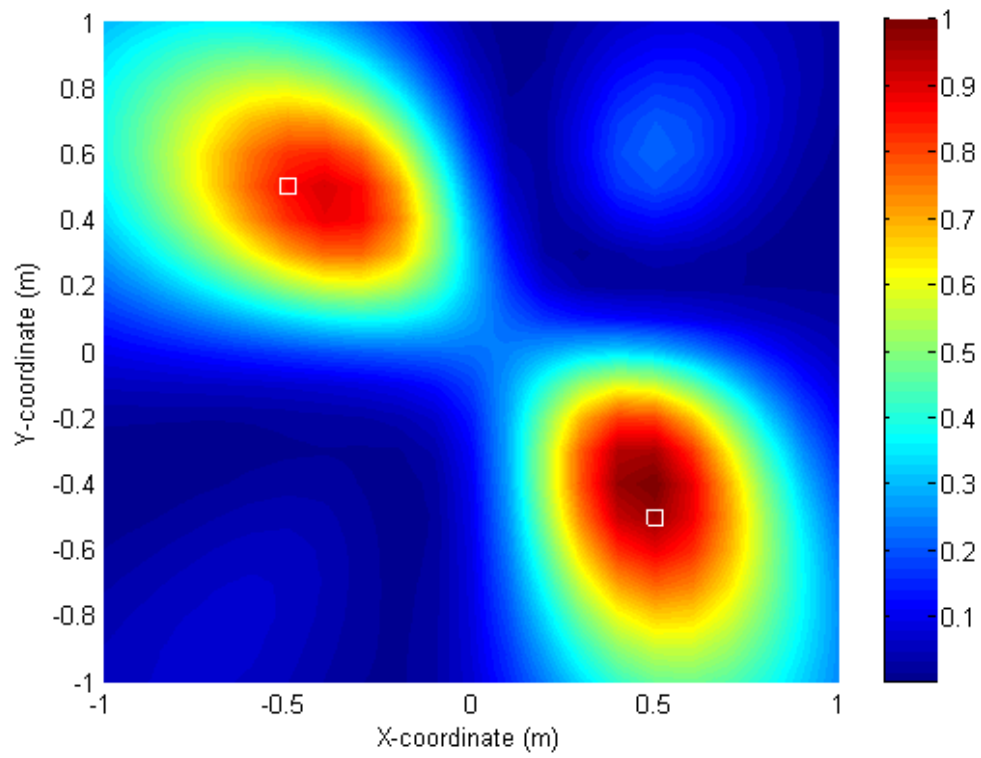




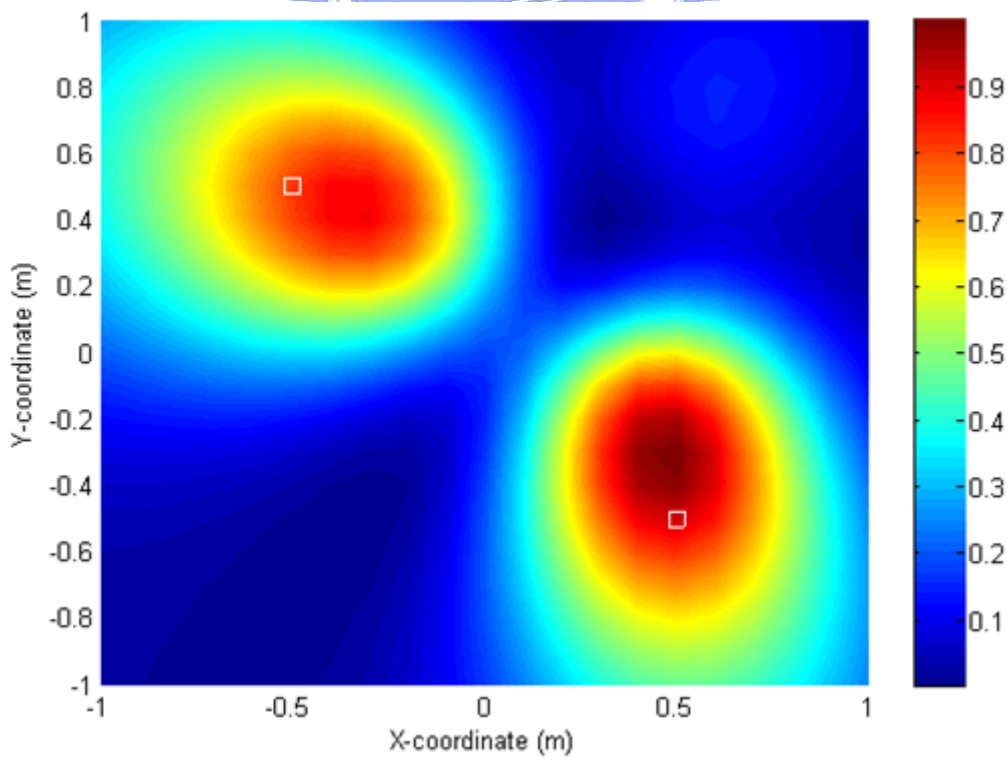
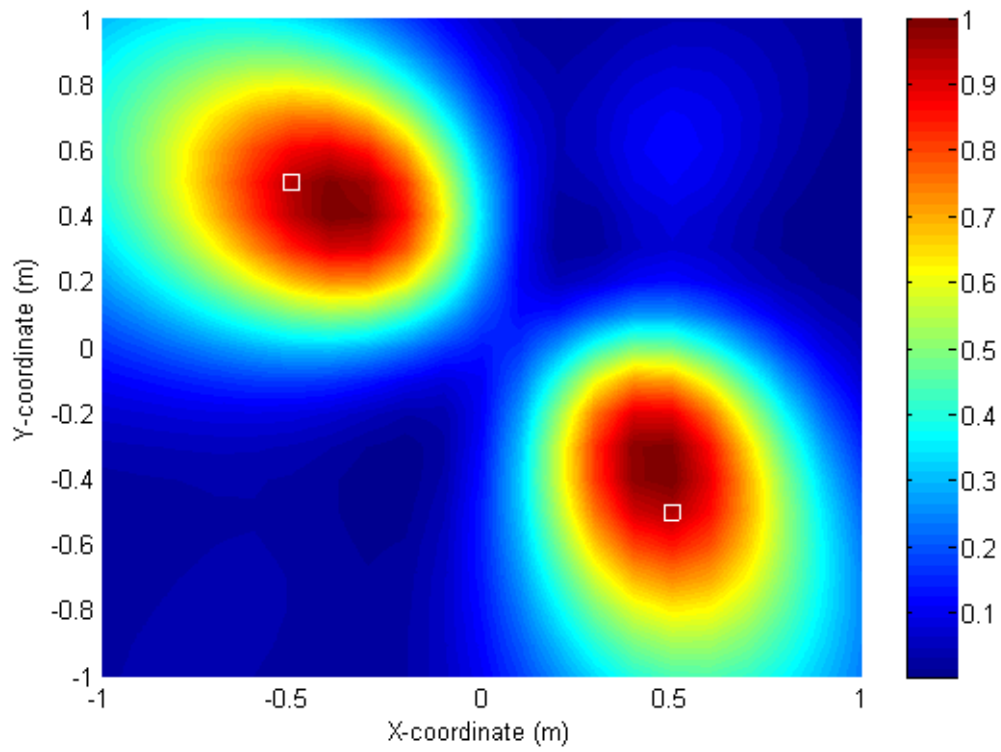
(b)

Fig. 41 The array configurations for DOA estimation simulations and experiments.

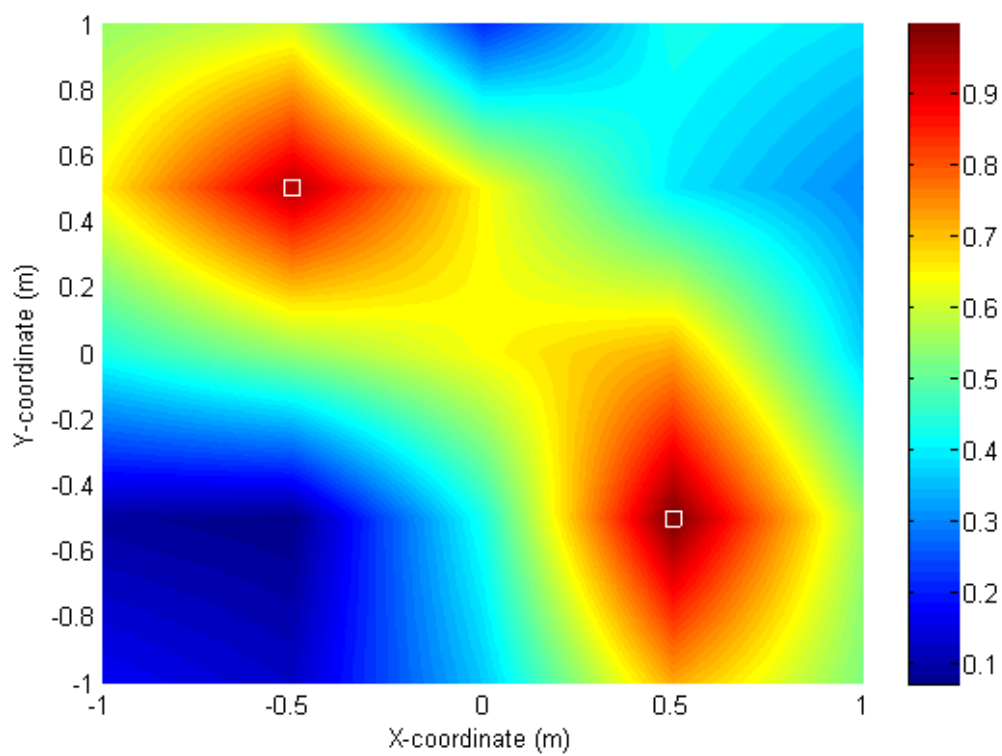
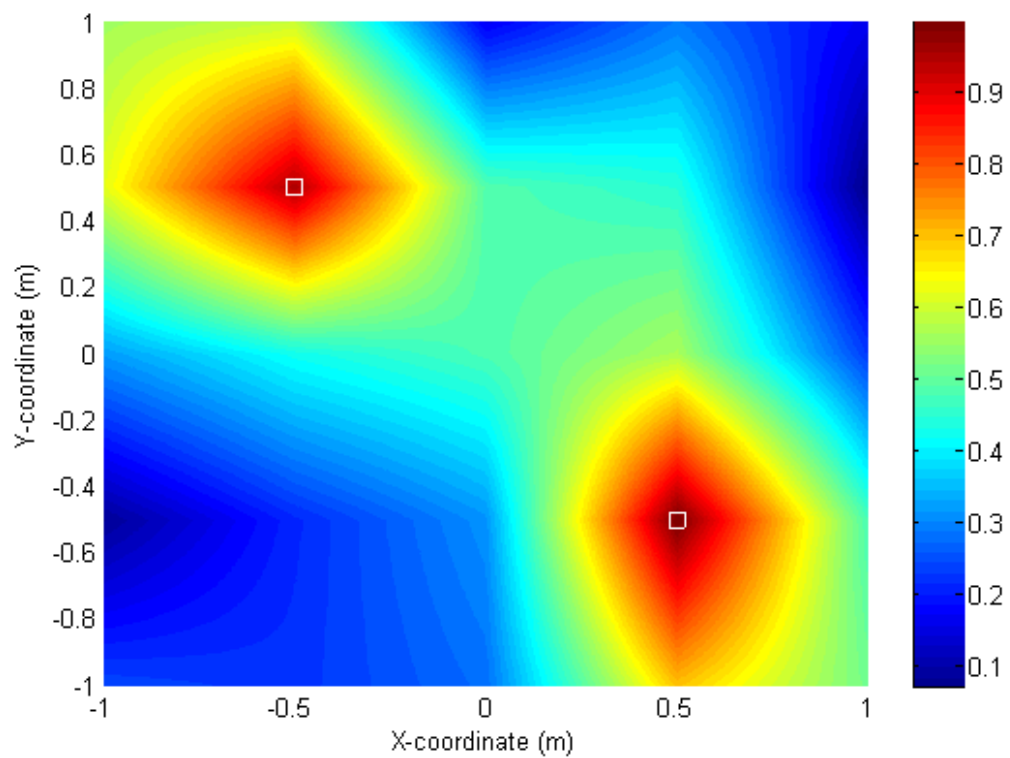
(a) URA and (b) optimized random array.



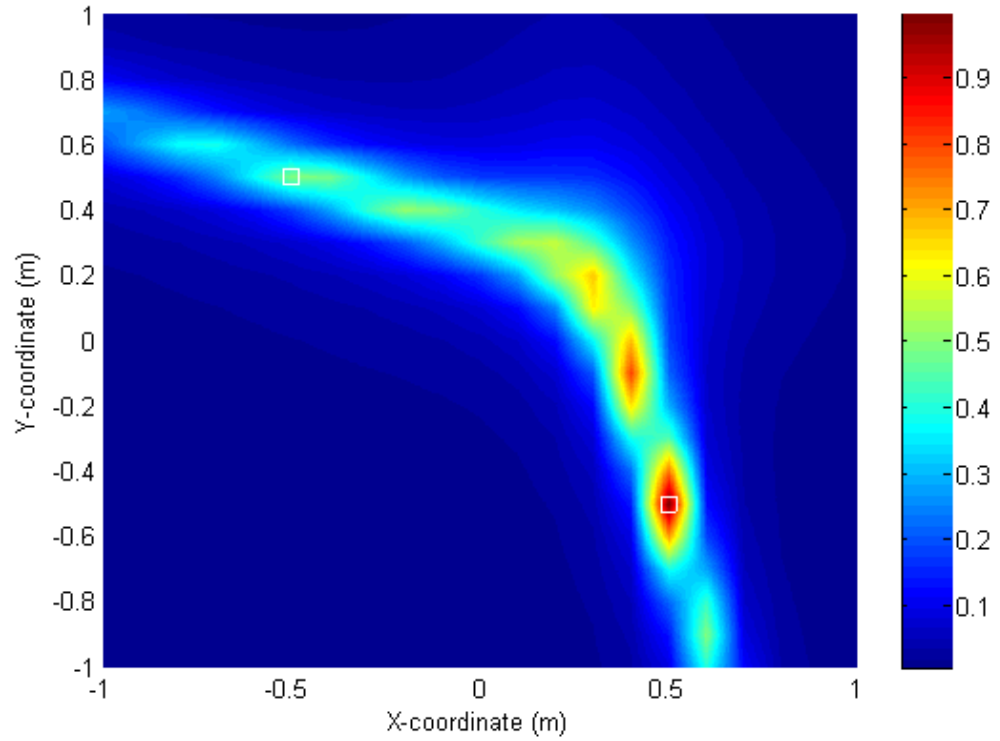
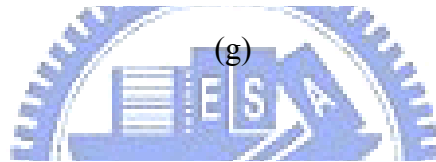
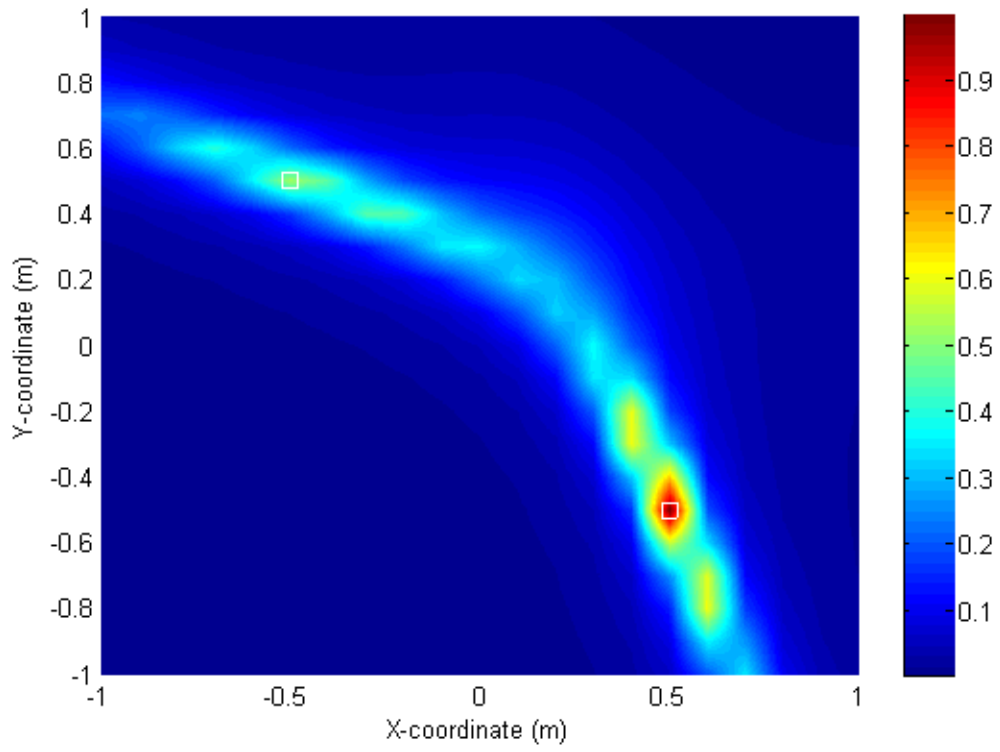
(b)



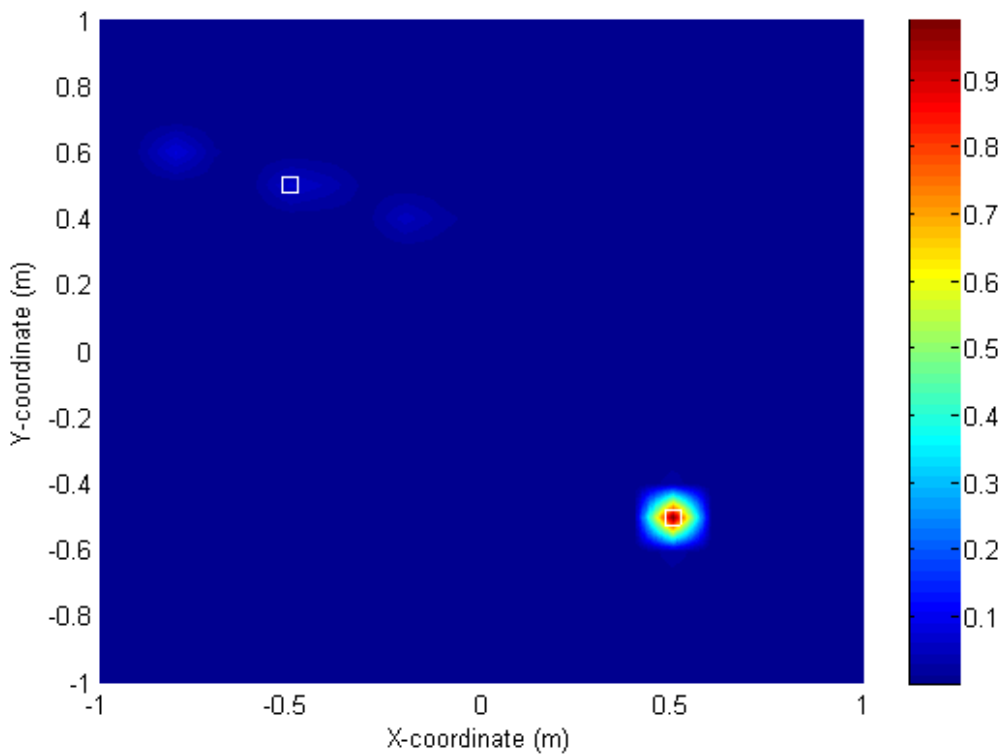
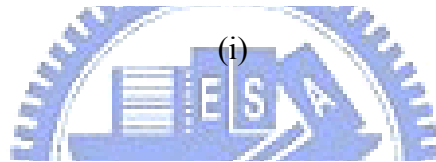
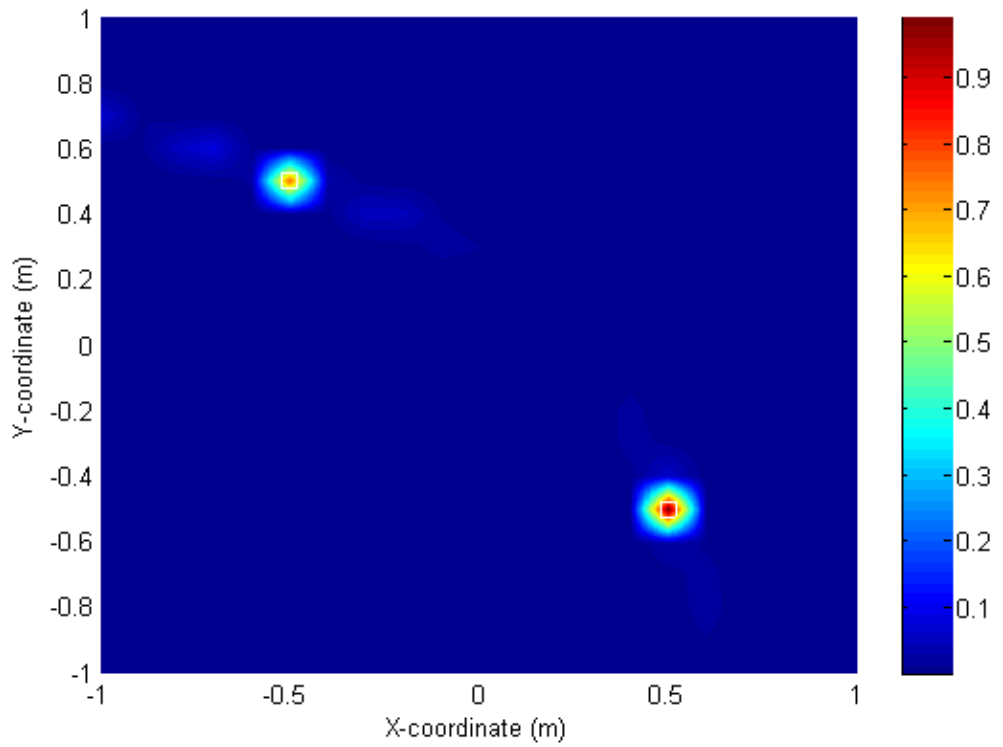
(d)



(f)



(h)

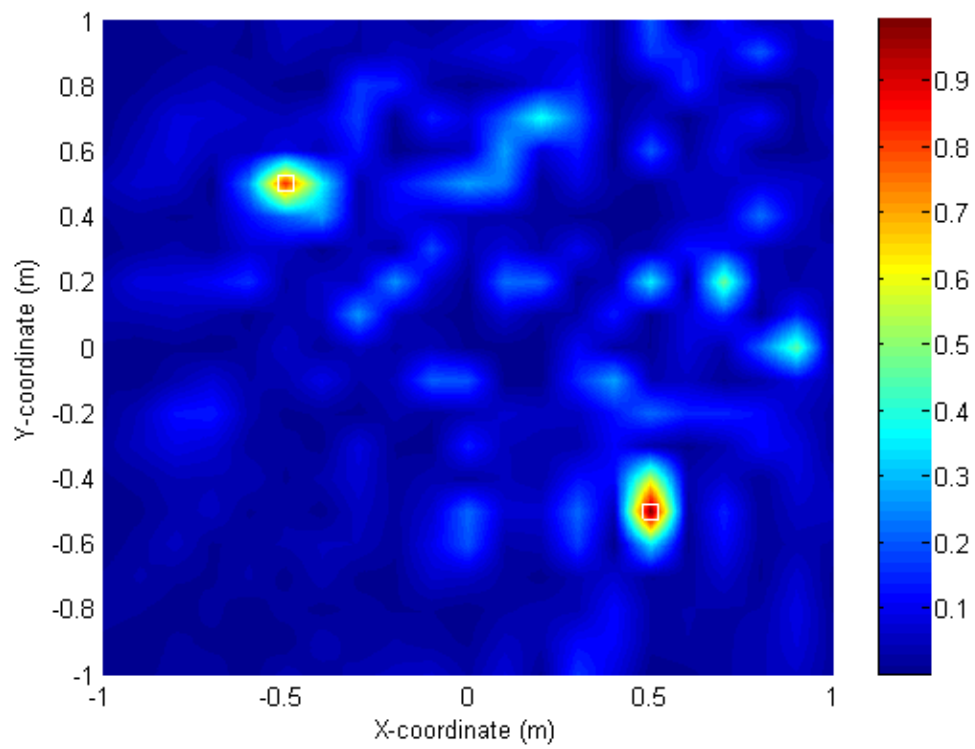
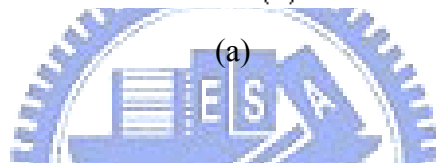
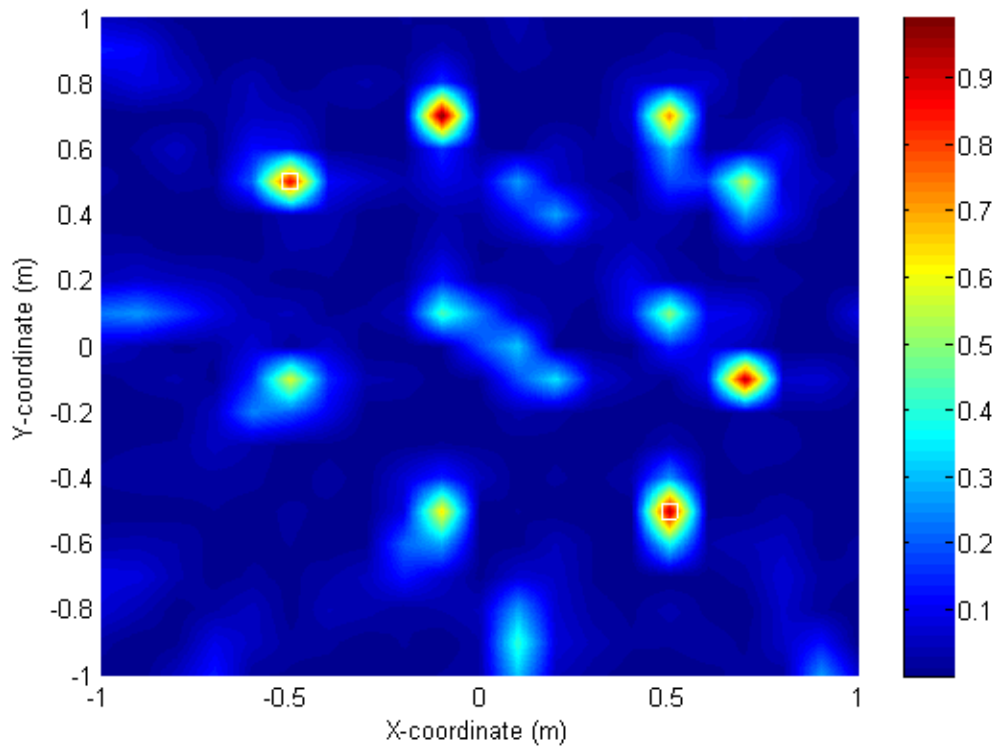


(j)

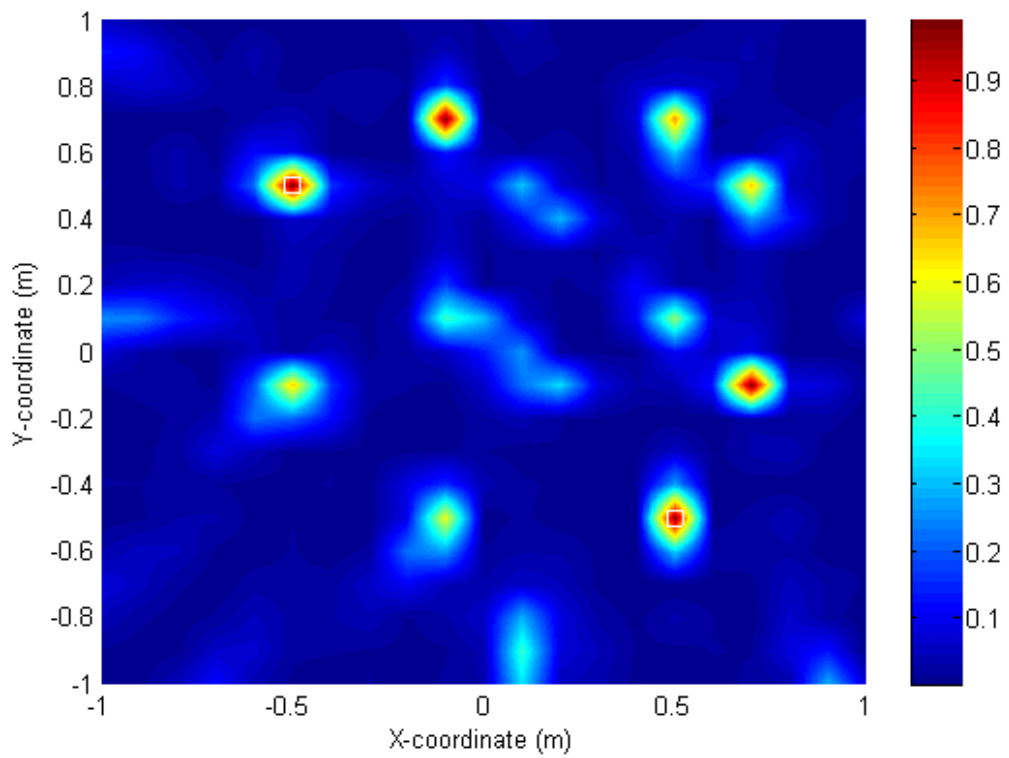
Fig. 42 The noise maps of two simulated point sources obtained using five acoustic

imaging algorithms with 30-channel URA and random array. The squares are the preset sound source positions. The simulated whitenoise sources located at the positions (-0.5m, 0.5m) and (0.5m, -0.5m). The observed frequency is 1 kHz ($d = \lambda/4$). The power spectrums obtained using (a) DAS with URA configuration, (b) DAS with random array configuration, (c) TR with URA configuration, (d) TR with random array configuration, (e) SIMO-ESIF with URA configuration, (f) SIMO-ESIF with random array configuration, (g) MVDR with URA configuration, (h) MVDR with random array configuration, (i) MUSIC with URA configuration and (j) MUSIC with random array configuration.

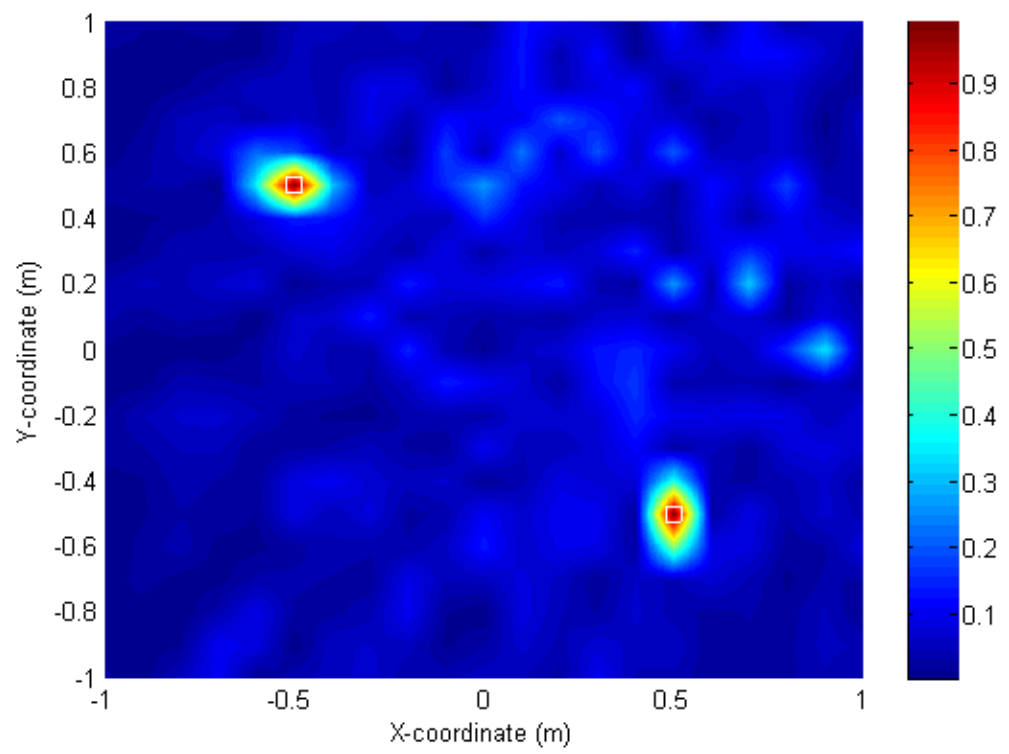




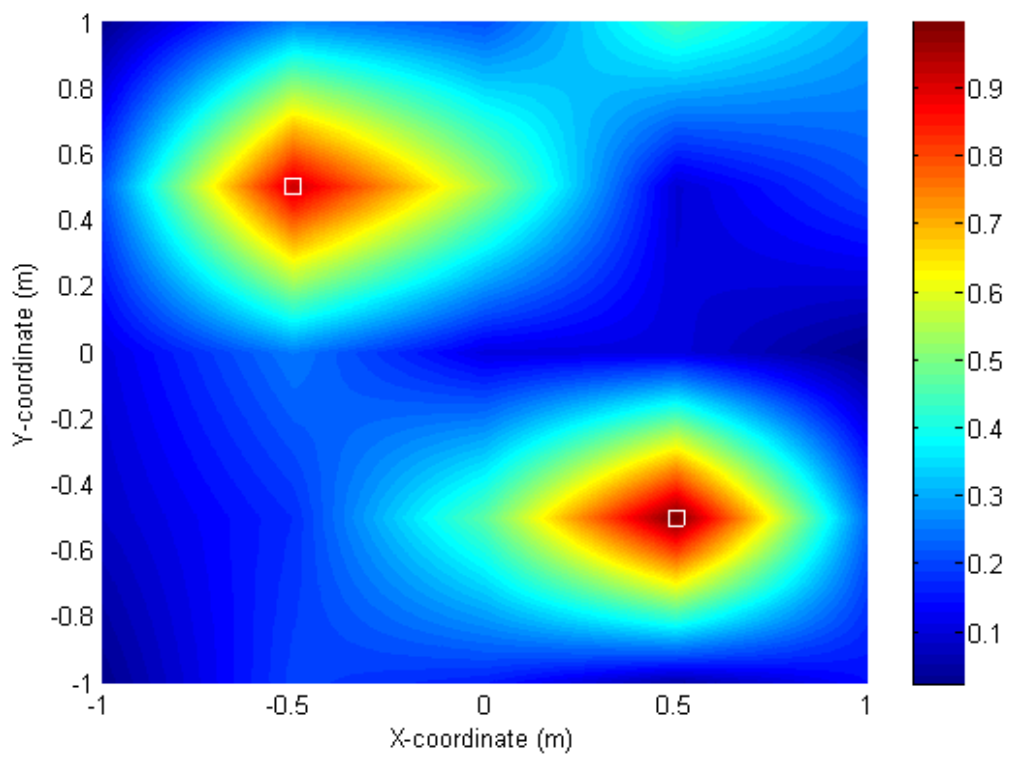
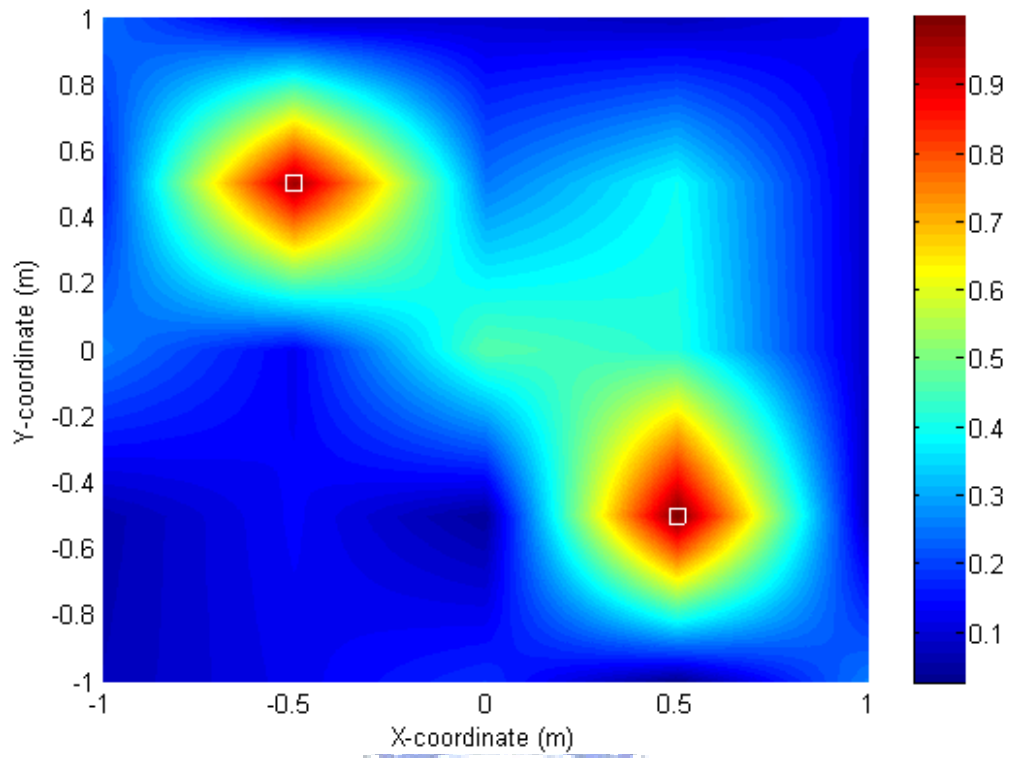
(b)



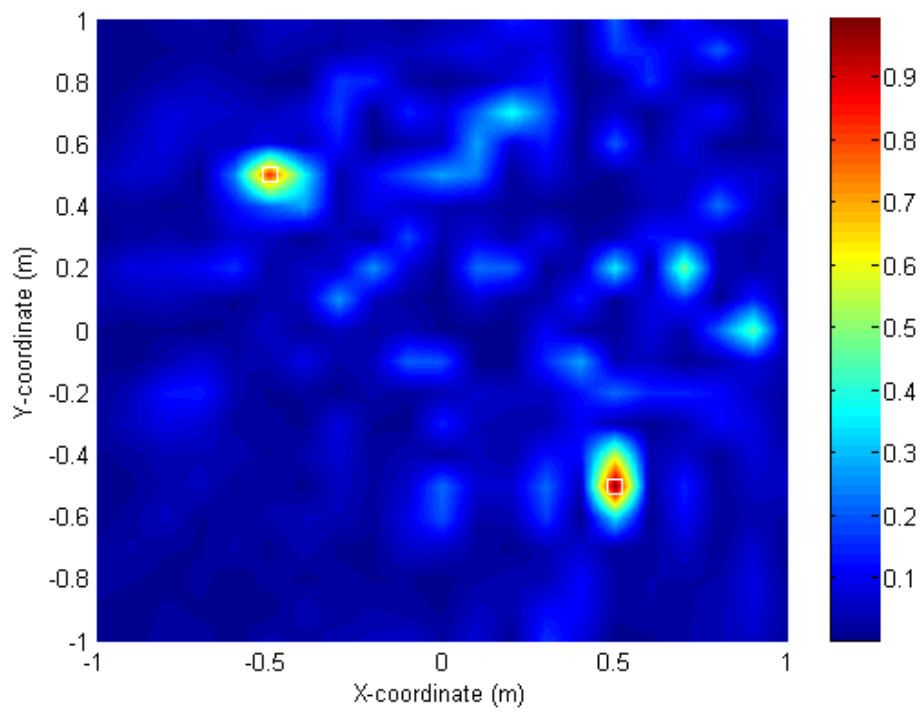
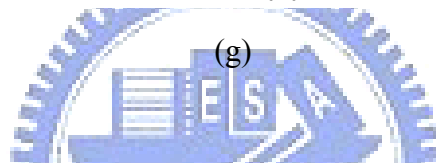
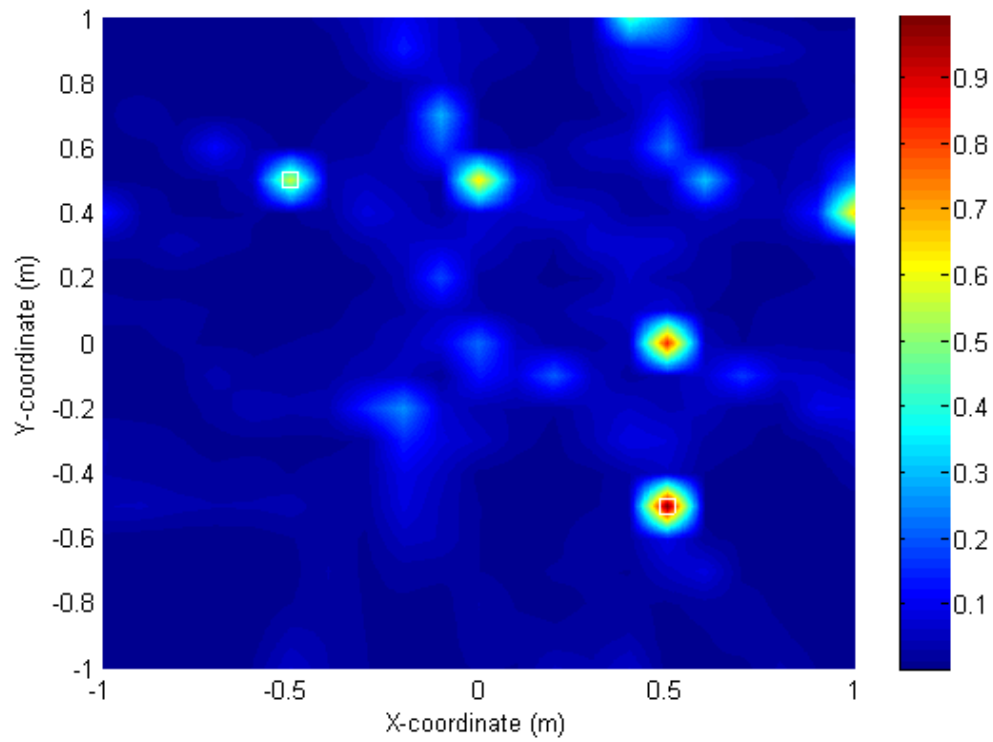
(c)



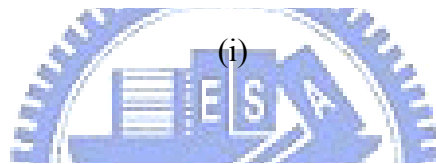
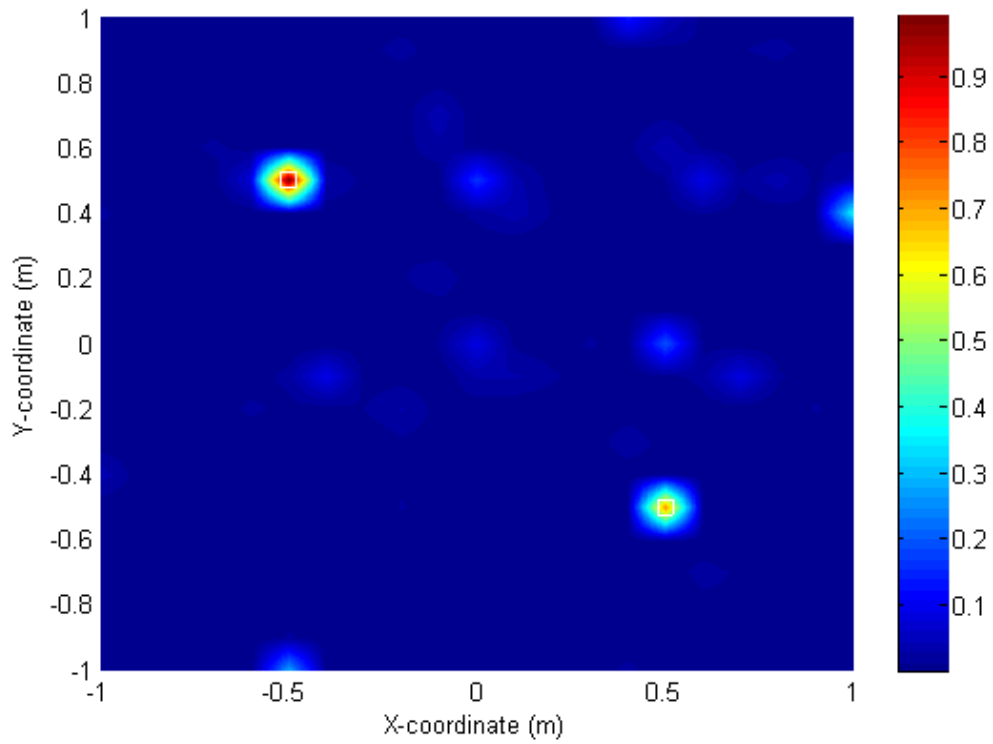
(d)



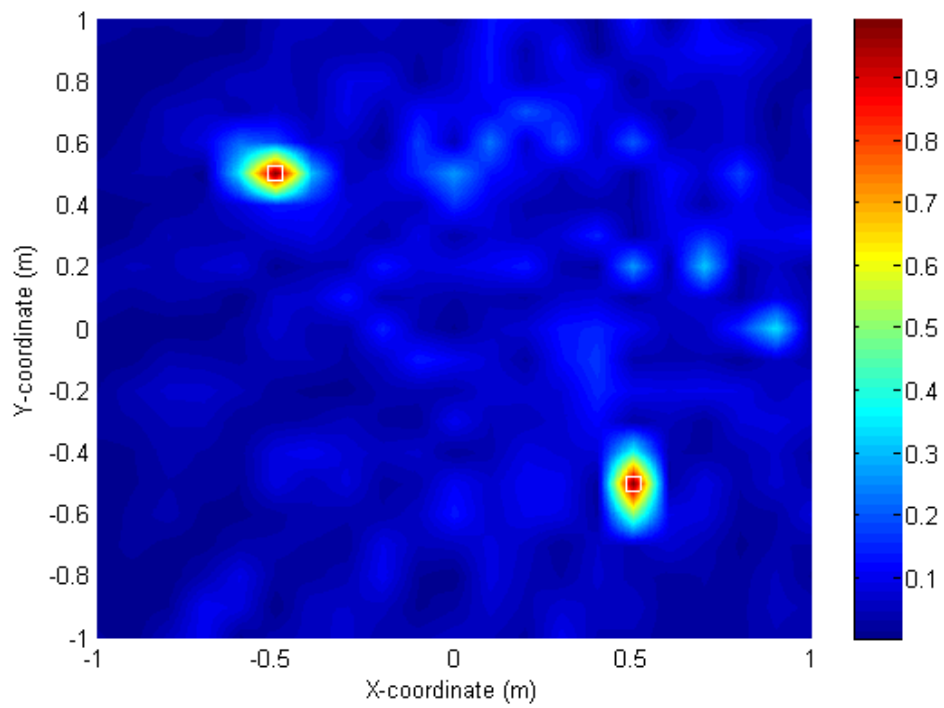
(f)



(h)



(i)



(j)

Fig. 43 The noise maps of two simulated point sources obtained using five acoustic

imaging algorithms with 30-channel URA and random array. The squares are the preset sound source positions. The simulated white noise sources located at the positions $(-0.5\text{m}, 0.5\text{m})$ and $(0.5\text{m}, -0.5\text{m})$. The observed frequency is 7 kHz ($d = 2\lambda$). The power spectrums obtained using (a) DAS with URA configuration, (b) DAS with random array configuration, (c) TR with URA configuration, (d) TR with random array configuration, (e) SIMO-ESIF with URA configuration, (f) SIMO-ESIF with random array configuration, (g) MVDR with URA configuration, (h) MVDR with random array configuration, (i) MUSIC with URA configuration and (j) MUSIC with random array configuration.



CHAPTER 5. NEARFIELD ARRAY SIGNAL PROCESSING ALGORITHMS

This section describes nearfield array signal processing algorithms such as the Fourier based NAH, IBEM (direct and indirect formulations) and ESM (direct and indirect formulations), NESI and Kalman filter-based algorithm. Comparison of the nearfield algorithms is shown in finally.

5.1 Fourier NAH

In this section, the traditional Fourier based NAH is reviewed. The Fourier NAH serves to reconstruct a three-dimensional sound field from the two-dimensional hologram data scanned above the source surface, as shown in Fig. 44.

The 2D spatial Fourier transformation is employed to transform the space domain to the wave number domain, and vice versa.

$$\tilde{p}(k_x, k_y, z) \triangleq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y, z) e^{j(k_x x + k_y y)} dx dy \quad (5.1)$$

$$p(x, y, z) \triangleq \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{p}(k_x, k_y, z) e^{-j(k_x x + k_y y)} dk_x dk_y, \quad (5.2)$$

where x , y and z are the Cartesian coordinates, d_x and d_y are the spacing of microphones in the x and y directions, and k_x and k_y are the wave number components in the x and y directions. In the k -domain, the sound pressure data of the reconstruction plane and the hologram plane can be related by

$$\tilde{p}(k_x, k_y, z) = \tilde{p}(k_x, k_y, z_H) e^{-jk_z(z-z_H)}, \quad (5.3)$$

where k_z is the wave number in the z direction. Let $H(k_x, k_y) = e^{-jk_z(z-z_H)}$. The sound field is expressed as propagating and evanescent wave components in the k -domain, i.e., a plane wave expansion. The Wiener inverse filter is employed to mitigate the ill-posedness during inverse reconstruction on boundary using Fourier NAH:

$$W(k_x, k_y) = \frac{1}{H(k_x, k_y)} \frac{1}{1 + (\alpha / |H(k_x, k_y)|)^2}, \quad (5.4)$$

where α is a regularization parameter. With the Wiener inverse filter, the Eq. (5.3) can be rewritten as

$$\tilde{p}(k_x, k_y, z) = \tilde{p}(k_x, k_y, z_H) W(k_x, k_y). \quad (5.5)$$

After multiplying with inverse propagator and k -domain filter, sound pressure can be reconstructed in the spatial domain by the inverse Fourier transform in Eq. (5.2).

The particle velocity can be calculated by using

$$\tilde{\mathbf{u}}(k_x, k_y, z) = \frac{\mathbf{k}}{\rho\omega} \tilde{p}(k_x, k_y, z), \quad (5.6)$$

where $\omega = 2\pi f$ is the angular frequency and $\mathbf{k} = (k_x, k_y, k_z)$ is the wave vector in Cartesian coordinates. The active intensity can be calculated with

$$\mathbf{I} = \frac{1}{2} \text{Re}\{\tilde{p}\tilde{\mathbf{u}}\}. \quad (5.7)$$

Although the Cartesian coordinates are most often employed in connection with the Fourier based NAH, the method can just as well be applied in spherical or cylindrical, as shown in Table 6. In the table, $H_n^{(1)}$ is the Hankel function of the first kind.

The Fourier based NAH approach is more efficient, but it has some drawbacks in practical application. First drawback is the microphone array must be larger than the source to avoid windowing effect from the spatial Fourier transform because the sound pressure must be sufficiently low at the boundary of the array. Other drawback is the grid spacing must be less than half a wavelength at the highest frequency of interest and the scanned grid points must be equally spaced on a planar rectangular area. Hence, it requires a large number of microphones to cover a reasonable source area. In addition, the measurement distance should be less than

$(R_x \frac{S}{N}) / (20\pi \log e)$, where R_x is spatial resolution in the x -axis direction and S/N is signal to noise ratio.



Table 6 Fourier based NAH for planar, cylindrical and spherical geometries [62].

Plane	$p(x, y, z; \omega) = F_x^{-1} F_y^{-1} \{F_x F_y [p(x, y, z_H; \omega)] e^{jk_z(z_H - z)}\},$ $u_n(x, y, z; \omega) = F_x^{-1} F_y^{-1} \{F_x F_y [p(x, y, z_H; \omega)] \frac{k_z}{\rho_0 c k} e^{jk_z(z_H - z)}\},$ <p>where</p> $F_x F_y [\varphi] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \varphi e^{-jk_x x} e^{-jk_y y} dx dy,$ $F_x^{-1} F_y^{-1} [\Phi] = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Phi e^{jk_x x} e^{jk_y y} dk_x dk_y, \text{ and } k_z = \sqrt{k^2 - k_x^2 - k_y^2}.$
Cylinder	$p(r, \phi, z; \omega) = F_\phi^{-1} F_z^{-1} \{F_\phi F_z [p(r_H, \phi, z; \omega)] \frac{H_n^{(1)}(k_r r)}{H_n^{(1)}(k_r r_H)}\},$ $u_n(r, \phi, z; \omega) = F_\phi^{-1} F_z^{-1} \{F_\phi F_z [p(r_H, \phi, z; \omega)] \frac{k_r}{j \rho_0 c k} \frac{H_n^{(1)}(k_r r)}{H_n^{(1)}(k_r r_H)}\},$ <p>where</p> $F_\phi F_z [\varphi] = \frac{1}{2\pi} \int_0^{2\pi} \int_{-\infty}^{\infty} \varphi e^{-jn\phi} e^{-jk_z z} d\phi dz,$ $F_\phi^{-1} F_z^{-1} [\Phi] = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \int_{-\infty}^{\infty} \Phi e^{jn\phi} e^{jk_z z} dk_z, \text{ and } k_r = \sqrt{k^2 - k_z^2}.$
Sphere	$p(r, \theta, \phi; \omega) = F_\theta^{-1} F_\phi^{-1} \{F_\theta F_\phi [p(r_H, \theta, \phi; \omega)] \frac{H_n^{(1)}(kr)}{H_n^{(1)}(kr_H)}\},$ $u_n(r, \theta, \phi; \omega) = F_\theta^{-1} F_\phi^{-1} \{F_\theta F_\phi [p(r_H, \theta, \phi; \omega)] \frac{1}{j \rho_0 c} \frac{H_n^{(1)}(kr)}{H_n^{(1)}(kr_H)}\},$ <p>where</p> $F_\theta F_\phi [\varphi] = \int_0^{2\pi} \int_0^\pi \varphi Y_n^m(\theta, \phi)^* \sin \theta d\theta d\phi,$ $F_\theta^{-1} F_\phi^{-1} [\Phi] = \sum_{n=0}^{\infty} \sum_{m=-n}^n \Phi Y_n^m(\theta, \phi).$

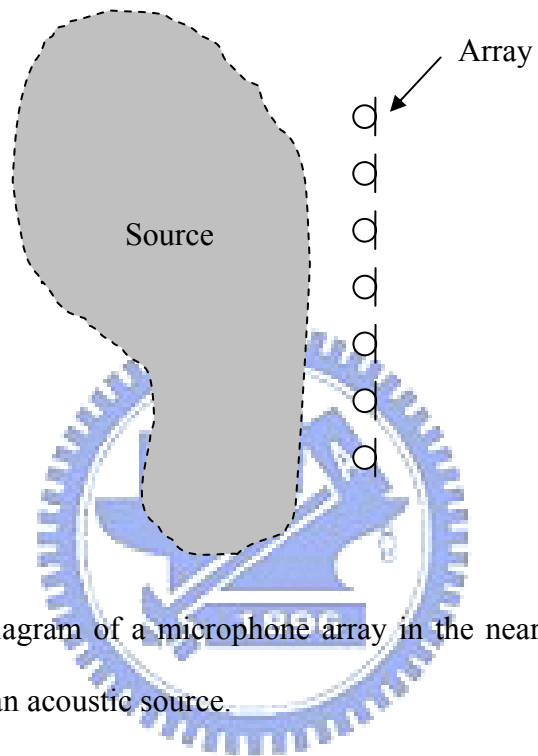


Fig. 44 Schematic diagram of a microphone array in the near field of the radiated sound from an acoustic source.

5.2 BEM-based NAH (IBEM): direct and indirect formulations

The BEM is a numerical method based on integral equations. The surface of an arbitrary object is discretized small elements and an interpolation scheme which uses an integral representation to produce the Green's function. The TF shows a relation between the source point and the field point in matrix form. In order to reduce implement formulation process to generate the measurement errors, it has to inverse the TF by regularization. This methodology of NAH is termed the IBEM. The IBEM is divided into two methods: The direct formulation can be derived from the HIE using the Green's function. In the indirect formulations, we survey two indirect formulations which include the single-layer potential and the double-layer potential. The disadvantage of the direct formulation is that it is not suited for thin structures.

5.2.1 Direct IBEM Formulation

The direct formulation can be derived from the HIE using the Green's function. Therefore, from the Green's theory, it can be recast into the Helmholtz integral for an exterior boundary value problem:

$$\alpha p(\mathbf{x}) = \int_S \left[p(\mathbf{x}_0) \frac{\partial G(\mathbf{x}, \mathbf{x}_0)}{\partial n} - G(\mathbf{x}, \mathbf{x}_0) \frac{\partial p}{\partial n}(\mathbf{x}_0) \right] dS(\mathbf{x}_0), \quad (5.8)$$

where \mathbf{x} and \mathbf{x}_0 are the field and source points on $S(\mathbf{x})$ and $S(\mathbf{x}_0)$ respectively; $S(\mathbf{x}_0)$ denotes the source surface, $S(\mathbf{x})$ denotes the field points surface, $\partial/\partial n$ the derivative operator in a normal direction to the surfaces, $G(\mathbf{x}, \mathbf{x}_0) = e^{-jkr(\mathbf{x}, \mathbf{x}_0)} / 4\pi r(\mathbf{x}, \mathbf{x}_0)$ is the free-space Green's function corresponding to the Helmholtz equation and

$$\alpha = \begin{cases} 1, & \text{if } \mathbf{x} \text{ is outside } S \\ \frac{1}{2}, & \text{if } \mathbf{x} \text{ is on } S \\ 0, & \text{if } \mathbf{x} \text{ is inside } S \end{cases}. \quad (5.9)$$

Equation (5.8) is used to express the sound field at a set of microphone positions

close to the source (see Fig. 45). The numerical is required due to integration cannot be solved analytically. This thesis chooses the simple discretization scheme in which the surface S is broken up into N small elements of area ΔS_l , $l=1, \dots, N$ as follows:

$$\int_S [\dots] dS \approx \sum_{n=1}^N \int_{S_n} [\dots] dS_n, \quad (5.10)$$

where S_n denotes the surface of n^{th} element, and N denotes number of element. The boundary elements used for approximating the surface integral are schematically shown in Fig. 46 (a) and Fig. 46 (b). Triangular elements and quadrilateral elements are introduced in this thesis for construction of meshes. The global and local coordinate system and are related by isoparametric transformation. Quadratic shape functions are used for interpolation the global coordinates as

$$X_i(\xi) \approx \sum_{l=1}^L N_l(\xi) x_{il}, \quad i=1, 2, 3; L=6 \text{ or } 8, \quad (5.11)$$

$$P_m(\xi) \approx \sum_{l=1}^L N_l(\xi) P_{ml}, \quad m=1, 2, \dots, M; L=6 \text{ or } 8, \quad (5.12)$$

$$\frac{\partial P_m}{\partial n_q}(\xi) \approx \sum_{l=1}^L N_l(\xi) \frac{\partial P_{ml}}{\partial n_q}, \quad m=1, 2, \dots, M; L=6 \text{ or } 8, \quad (5.13)$$

where x_{il} is the i th coordinate component of the l th node. $N_l(\xi)$ are the quadratic functions and $\xi \equiv (\xi_1, \xi_2)$ are the local coordinate. P_{ml} and $\partial P_{ml} / \partial n_q$ are sound pressure and pressure gradient of the l th node on the m th element, and M is the number of the element. The integration over each element can be computed using numerical integration. Each element will be found the pressure and the pressure gradient by node values, respectively.

Assume that the sound pressure data $p(\mathbf{x}^h)$ are measured on N locations of the hologram S_h . The integral equation relating the source surface and the hologram is discretized and assembled into the matrix form as in Eq. (5.8) with $\alpha=1$:

$$\mathbf{P}^h = \mathbf{D}^{hs} \mathbf{P}^s - \mathbf{S}^{hs} \mathbf{P}_n^s, \quad (5.14)$$

where \mathbf{P}^h , \mathbf{P}^s and \mathbf{P}_n^s are $N \times 1$ column vectors corresponding to the given sound pressure $p(\mathbf{x}^h)$ and the unknown sound pressure $p(\mathbf{x}^s)$ and its gradient $\partial p(\mathbf{x}^s)/\partial n$; the superscript h and s denote, respectively, the hologram S_h and the source surface S_s , the subscript n denotes the directional derivative, \mathbf{D}^{hs} and \mathbf{S}^{hs} are both $N \times N$ square matrices corresponding to the integrals in Eq. (5.8) that relate the N measuring points \mathbf{x}^h and the N surface nodes \mathbf{x}^s , the superscript hs denotes a spatial transformation between the hologram and the source surface.

Boundary integral equation (5.8) with $\alpha = 1/2$ takes the following matrix form if the field point is taken to the source surface S_s :

$$\alpha \mathbf{P}^s = \mathbf{D}^{ss} \mathbf{P}^s - \mathbf{S}^{ss} \mathbf{P}_n^s \quad (5.15)$$

or

$$\bar{\mathbf{D}}^{ss} \mathbf{P}^s = \mathbf{S}^{ss} \mathbf{P}_n^s, \quad (5.16)$$

where \mathbf{D}^{ss} and \mathbf{S}^{ss} are both $N \times N$ square matrices corresponding to the integrals in Eq. (5.14) that relate the N field points and the N nodes \mathbf{x}^s on source surface S_s , the superscript ss denotes both the field and source points which are located on the source surface, $\bar{\mathbf{D}}^{ss} \equiv (\mathbf{D}^{ss} - \alpha \mathbf{I})$, with \mathbf{I} being an identity matrix.

Equations (5.14) and (5.16), respectively, constitute the main equation and the constraint equation with which one performs backward reconstruction of the sound field for irregularly shaped sources. This technique is termed pressure-based conformal holography with a hologram and a source surface coupling (PCHHS), since its formulation is based on the sound pressure measured on the hologram.

There are two different ways of solving the unknowns \mathbf{P}^s and \mathbf{P}_n^s in Eqs. (5.

14) and (5. 16). First, one may combine these two sets of equations into a single linear system with $2N$ equations and $2N$ unknowns. Second, one may substitute the constraint Eq. (5. 16) into the main Eq. (5. 14) to yield a single linear system with N equations and N unknowns. The second approach is adopted here because it requires less memory space and produces better-conditioned matrices than the first approach. That is,

$$[\mathbf{D}^{hs} (\bar{\mathbf{D}}^{ss})^{-1} \mathbf{S}^{ss} - \mathbf{S}^{hs}] \mathbf{P}_n^s = \mathbf{P}^h, \quad (5. 17)$$

with \mathbf{P}^s eliminated. \mathbf{P}^s can then be recovered from Eq. (5. 16) after \mathbf{P}_n^s is solved in Eq. (5. 17). One may alternative eliminate alternate \mathbf{P}_n^s from Eqs. (5. 14) and (5. 16) to obtain

$$[\mathbf{D}^{hs} - \mathbf{S}^{hs} (\mathbf{S}^{ss})^{-1} \bar{\mathbf{D}}^{ss}] \mathbf{P}^s = \mathbf{P}^h. \quad (5. 18)$$

\mathbf{P}_n^s can then be recovered from Eq. (5. 16) after \mathbf{P}^s is solved in Eq. (5. 18).

In some cases, it is more desirable to measure the sound pressure gradient, or equivalently, the particle velocity, by using an intensity probe. This situation arises when, for example, one seeks to avoid the influences from disturbing sources from the background in carrying out an acoustic measurement. This application motivates the development of the following velocity-based conformal holography with a hologram and a source surface coupling (VCHHS).

Suppose that the sound pressure gradient $\partial p(\mathbf{x}^h) / \partial n$, or the particle velocity $u_n(\mathbf{x}^h)$, is measured at N locations on the hologram. The integral equation relating the source surface and the hologram then takes the matrix form with $\alpha = 1$:

$$\mathbf{p}_n^h = \mathbf{D}_n^{hs} \mathbf{p}^s - \mathbf{K}^{hs} \mathbf{p}_n^s, \quad (5. 19)$$

where \mathbf{p}_n^h is an $N \times 1$ column vector corresponding to the known quantity

$\partial p(\mathbf{x}^h)/\partial n$; \mathbf{D}_n^{hs} and \mathbf{K}^{hs} are both $N \times N$ square matrices corresponding to the integrals that relate the N measuring points \mathbf{x}^h and the N surface nodes \mathbf{x}^s . As with the pressure-based conformal holography, the Eq. (5. 16) can be used again as a constraint equation for eliminating the unknown \mathbf{P}_s from the main Eq. (5. 19) to arrive at a linear system with N equations and N unknowns:

$$\left[\mathbf{D}_n^{hs} (\bar{\mathbf{D}}^{ss})^{-1} \mathbf{S}^{ss} - \mathbf{K}^{hs} \right] \mathbf{p}_n^s = \mathbf{p}_n^h. \quad (5. 20)$$

Here, \mathbf{P}^s can then be recovered from Eq. (5. 16) after \mathbf{P}_n^s is solved in Eq. (5. 20).

One may alternatively eliminate \mathbf{P}_n^s from Eq. (5. 19) to obtain

$$\left[\mathbf{D}_n^{hs} - \mathbf{K}^{hs} (\mathbf{S}^{ss})^{-1} \bar{\mathbf{D}}^{ss} \right] \mathbf{p}^s = \mathbf{p}^h. \quad (5. 21)$$

Then \mathbf{P}_n^s can then be recovered from Eq. (5. 16) after \mathbf{P}^s is solved in Eq. (5. 21).

In addition to the previously mentioned algorithms that utilize the source surface integral as a constraint, an arbitrarily chosen surface located in the interior of the source may be used for setting up a constraint equation (see Fig. 45). This interior surface is only a fictitious one which is not required to be physically accessible in field measurement. One advantage of choosing an interior surface instead of a source surface as a constraint is: the integration kernels can never become singular when using the interior surface since the distances between the source points and field points are always greater than zero. Slow convergence in carrying out Gaussian quadrature integration for singular elements is then avoided to some degree. The choice of interior points is arbitrary, except for those points coinciding with the nodal points of the eigenmodes of the corresponding interior problems. Interior points with N pressure data measured on the hologram are chosen here to be located at equal distances along the inward normal directions to the source surface points. N interior points are thus located on a surface that almost conforms to the source surface.

The following pressure-based matrix equation is obtained here in accordance with the HIE with $\alpha = 0$:

$$\mathbf{D}^{is} \mathbf{P}^s = \mathbf{S}^{is} \mathbf{P}_n^s, \quad (5. 22)$$

where \mathbf{D}^{is} and \mathbf{S}^{is} are both $N \times N$ square matrices corresponding to the integrals that relate the N field points and the N nodes \mathbf{x}^s on the source surface S_s ; the superscript is denotes the spatial transformation between the interior surface S_i and the source surface S_s .

Parallel to the development of the aforementioned two algorithms, the matrix Eq. (5. 22) can be used as a constraint equation for eliminating either the unknown surface pressure \mathbf{P}^s , or the surface pressure gradient \mathbf{P}_n^s , from the main Eq. (5. 14) to solve for the remaining unknown quantity. This approach is then termed pressure-based conformal holography with a hologram and a source interior coupling (PCHHI).

Similar reasoning can finally be applied to obtain a holography transformation algorithm based on the sound particle velocity \mathbf{P}_n^h measured on the hologram, by using the source interior equation (Eq. (5. 22)) as a constraint. One may choose to eliminate either the surface pressure \mathbf{P}^s or the surface pressure gradient \mathbf{P}_n^s to recover the other variable. This approach leads to velocity-based conformal holography with a hologram and a source interior coupling (VCHHI).

The resulting matrix equations of these four acoustic holography algorithms (with the different types of constraints presented in this section) are summa in

Table 7. Backward reconstruction and forward propagation of the sound pressure, the particle velocity, and the sound intensity at any field point of interest can be performed by applying either the pressure-based approach or the velocity-based approach.

Some noteworthy numerical aspects arise in the implementation phase of the acoustic holography algorithms. The first aspect is associated with evaluation of the singular integrals. The assembled coefficient matrices in the aforementioned integral formulation, i.e., $\bar{\mathbf{D}}^{ss}$ and \mathbf{S}^{ss} , involve evaluation of singular integrals. Element integrals can be conveniently evaluated by using the Gaussian quadrature algorithm when the field point \mathbf{x}_p and the source point \mathbf{x}_q are apart from each other. Integrand becomes singular when the field point \mathbf{x}_p coincides with the source point \mathbf{x}_q . Singularity of the integrand can be reduced by means of polar transformation in order to improve convergence in carrying out the Gaussian quadrature integration. This is a typical procedure detailed in BEM literature.

In applying the IBEM to reconstructing sound fields of a vibrating object, more than the number of discrete nodes considering a wavelength of interest is required to avoid distortions and take a considerable amount of measurements, since the surface quantities are described by spatial discretization. Such a procedure may be impractical, especially for a complex-shaped structure due to the fact that the discrete nodes and corresponding measurements may be excessive and the reconstruction process can be extremely complex and time consuming.

Table 7 Summary of the BEM-based acoustic holography algorithms.

Model	Known	Unknown	Fundamental matrix equations	Variable eliminated	Resultant system equations
PCHHS	\mathbf{p}^h		Main equation: $\mathbf{P}^h = \mathbf{D}^{hs} \mathbf{p}^s - \mathbf{S}^{hs} \mathbf{p}_n^s$	\mathbf{p}^s	$[\mathbf{D}^{hs} (\bar{\mathbf{D}}^{ss})^{-1} \mathbf{S}^{ss} - \mathbf{S}^{hs}] \mathbf{P}_n^s = \mathbf{P}^h$
			Constraint equation: $\bar{\mathbf{D}}^{ss} \mathbf{p}^s = \mathbf{S}^{ss} \mathbf{p}_n^s$	\mathbf{p}_n^s	$[\mathbf{D}^{hs} - \mathbf{S}^{hs} (\mathbf{S}^{ss})^{-1} \bar{\mathbf{D}}^{ss}] \mathbf{P}^s = \mathbf{P}^h$
VCHHS	\mathbf{p}_n^h		Main equation: $\mathbf{P}_n^h = \mathbf{D}_n^{hs} \mathbf{p}^s - \mathbf{K}^{hs} \mathbf{p}_n^s$	\mathbf{p}^s	$[\mathbf{D}_n^{hs} (\bar{\mathbf{D}}^{ss})^{-1} \mathbf{S}^{ss} - \mathbf{K}^{hs}] \mathbf{p}_n^s = \mathbf{P}_n^h$
		\mathbf{p}^s and \mathbf{p}_n^s	Constraint equation: $\bar{\mathbf{D}}^{ss} \mathbf{p}^s = \mathbf{S}^{ss} \mathbf{p}_n^s$	\mathbf{p}_n^s	$[\mathbf{D}_n^{hs} - \mathbf{K}^{hs} (\mathbf{S}^{ss})^{-1} \bar{\mathbf{D}}^{ss}] \mathbf{p}^s = \mathbf{P}_n^h$
PCHHI	\mathbf{p}^h		Main equation: $\mathbf{P}^h = \mathbf{D}^{hs} \mathbf{p}^s - \mathbf{S}^{hs} \mathbf{p}_n^s$	\mathbf{p}^s	$[\mathbf{D}^{hs} (\mathbf{D}^{is})^{-1} \mathbf{S}^{is} - \mathbf{S}^{hs}] \mathbf{p}_n^s = \mathbf{P}^h$
			Constraint equation: $\mathbf{D}^{is} \mathbf{p}^s = \mathbf{S}^{is} \mathbf{p}_n^s$	\mathbf{p}_n^s	$[\mathbf{D}^{hs} - \mathbf{S}^{hs} (\mathbf{S}^{is})^{-1} \mathbf{D}^{is}] \mathbf{p}^s = \mathbf{P}^h$
VCHHI	\mathbf{p}_n^h		Main equation: $\mathbf{P}_n^h = \mathbf{D}_n^{hs} \mathbf{p}^s - \mathbf{K}^{hs} \mathbf{p}_n^s$	\mathbf{p}^s	$[\mathbf{D}_n^{hs} (\mathbf{D}^{is})^{-1} \mathbf{S}^{is} - \mathbf{K}^{hs}] \mathbf{p}_n^s = \mathbf{P}_n^h$
			Constraint equation: $\mathbf{D}^{is} \mathbf{p}^s = \mathbf{S}^{is} \mathbf{p}_n^s$	\mathbf{p}_n^s	$[\mathbf{D}_n^{hs} - \mathbf{K}^{hs} (\mathbf{S}^{is})^{-1} \mathbf{D}^{is}] \mathbf{p}^s = \mathbf{P}_n^h$

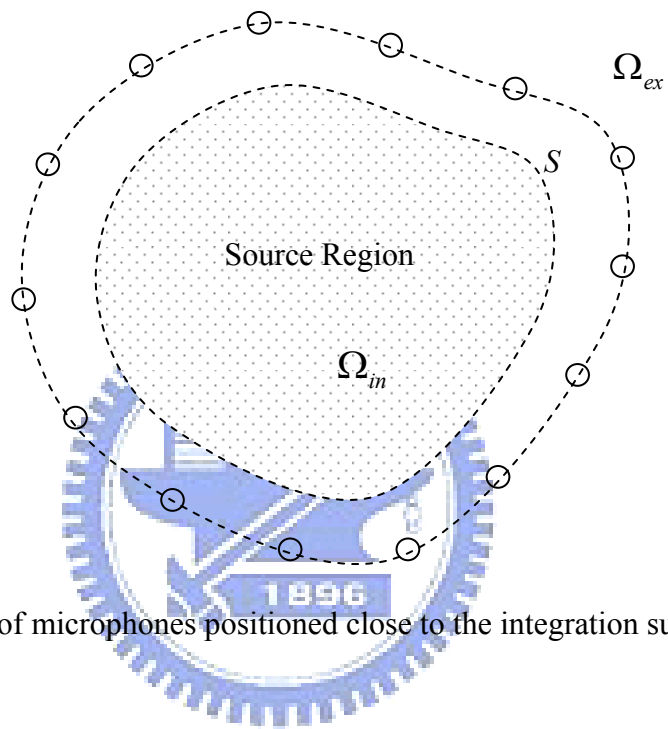


Fig. 45 A set of microphones positioned close to the integration surface S .

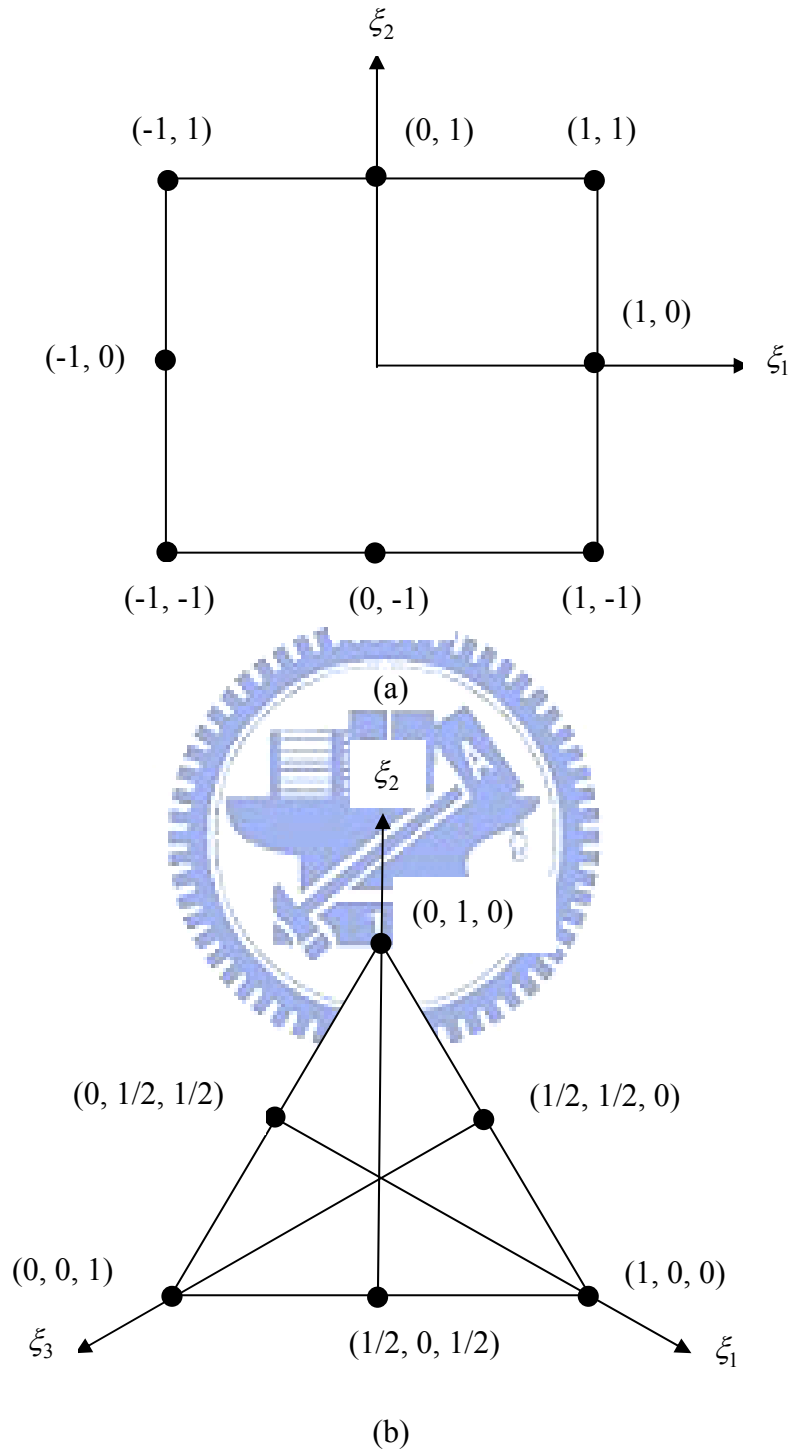


Fig. 46 Boundary elements used in isoparametric transformation. (a) Quadratic quadrilateral element, (b) quadratic triangular element.

5.2.2 Indirect IBEM Formulation

In this section, we survey two indirect formulations which include the single-layer potential and the double-layer potential. The acoustic field is expressed in terms of above layer potential theory on the source boundary S . First boundary integral method based on a distribution of single-layer on the surface is shown in Fig. 47. The pressure p can be expressed in terms of the simple-layer potential:

$$p(\mathbf{x}) = \int_S \sigma(\mathbf{x}_0) G(\mathbf{x}, \mathbf{x}_0) dS(\mathbf{x}_0), \quad (5.23)$$

where \mathbf{x} and \mathbf{x}_0 are field point and source point, respectively. $G(\mathbf{x}, \mathbf{x}_0)$ is the free-space Green's function, and the unknown source strength $\sigma(\mathbf{x}_0)$ is written as

$$\sigma = \frac{\partial p^+}{\partial n} - \frac{\partial p^-}{\partial n}. \quad (5.24)$$

We recognize the Green's function in the integrand, and we can view this formulation as a distribution of simple sources (monopoles) on the surface S . Once the source strengths $\sigma(\mathbf{x}_0)$ are known, we are able to find the pressure on the surface or in the exterior by means of surface integration.

Another is to express the field pressure by a double-layer formulation

$$p(\mathbf{x}) = \int_S \mu(\mathbf{x}_0) \frac{\partial G(\mathbf{x}, \mathbf{x}_0)}{\partial n} dS(\mathbf{x}_0), \quad (5.25)$$

where $\mu(\mathbf{x}_0)$ is the unknown source density. Such a representation of the exterior field can be shown to be distribution of double sources (dipoles) over S . The unknown source density $\mu(\mathbf{x}_0)$ is written for \mathbf{x}_0 on S as

$$\mu = p^- - p^+, \quad (5.26)$$

where p^- and p^+ are the pressures on the two sides of the surface at source point \mathbf{x}_0 .

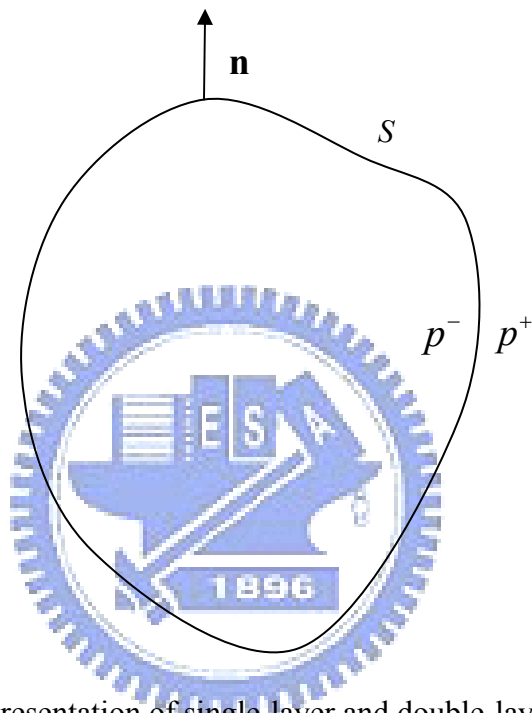


Fig. 47 The representation of single-layer and double-layer potential.

5.3 Equivalent source method (ESM)

ESM has gained much attention in the recent years. This method employed a set of virtual simple sources that is monopoles or dipoles placed inside the vibrating structure to represent the radiation from the source. ESM has been used for solving the inverse problem. The idea of ESM is modeling sound field generated by a set of the distributed simple sources placed on a vibrating structure in the interior of the structure, as shown in Fig. 48. Then use these source strengths to estimate the sound field on the source. The advantages of ESM are its ease to implement and low computation complexity, as compared to the direct IBEM approach. The disadvantage is the quandary of how to decide the distribution density and the RD between the source surface and virtual source surface (monopoles).



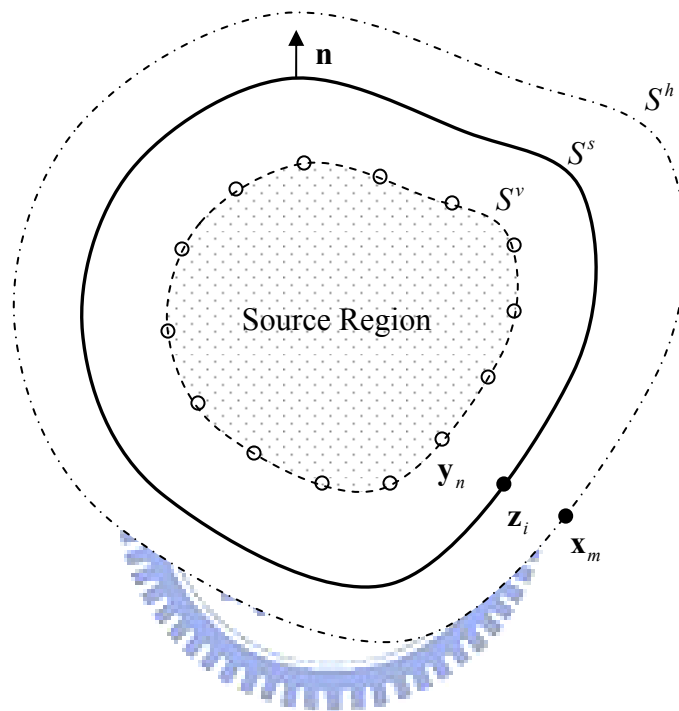


Fig. 48 The definitions of important surfaces used in the ESM-based NAH. The symbol \mathbf{x}_m is the m th microphone position on the hologram surface S^h . The symbol \mathbf{z}_i is the i th source point on the actual source surface S^s . The symbol \mathbf{y}_n is the n th virtual source point on the virtual source surface S^v .

5.3.1 Direct ESM

The direct ESM uses the multipole expansion by the discretization with simple sources placed in the interior of the structure (see Fig. 48). The ESM algorithm is shown to be equivalent to Helmholtz-integral formulation that can be represented in matrix form. In acoustic problems, the Kirchhoff–Helmholtz integral equation can be derived by applying the Green’s theorem to the volume integral equations:

$$\alpha p(\mathbf{x}) = \int_s [p(\mathbf{x}_0) \frac{\partial G(\mathbf{x}, \mathbf{x}_0)}{\partial n} - G(\mathbf{x}, \mathbf{x}_0) \frac{\partial p}{\partial n}(\mathbf{x}_0)] dS(\mathbf{x}_0), \quad (5.27)$$

where \mathbf{x} and \mathbf{x}_0 are the field and source points on $S(\mathbf{x})$ and $S(\mathbf{x}_0)$ respectively; $S(\mathbf{x}_0)$ denotes the source surface, $S(\mathbf{x})$ denotes the field points surface, $\partial/\partial n$ the derivative operator in a normal direction to the surfaces, $G(\mathbf{x}, \mathbf{x}_0) = e^{-jkr(\mathbf{x}, \mathbf{x}_0)} / 4\pi r(\mathbf{x}, \mathbf{x}_0)$ is the free-space Green’s function corresponding to the Helmholtz equation and

$$\alpha = \begin{cases} 1, & \text{if } \mathbf{x} \text{ is outside } S \\ \frac{1}{2}, & \text{if } \mathbf{x} \text{ is on } S \\ 0, & \text{if } \mathbf{x} \text{ is inside } S \end{cases}. \quad (5.28)$$

This study chooses the simple discretization scheme in which the surface S is broken up into N small elements of area ΔS_l , $l = 1, \dots, N$ as follows:

$$\alpha p_m = \sum_{l=1}^N (p_l \frac{\partial G_{ml}}{\partial n} - G_{ml} \frac{\partial p_l}{\partial n}) \Delta S_l. \quad (5.29)$$

The matrix representation of Eq. (5.29) can be obtained after utilizing the numerical integration as follows:

$$\alpha \mathbf{p}^h = \mathbf{D}^{hs} \mathbf{S} \mathbf{p}^s - \mathbf{M}^{hs} \mathbf{S} \frac{\partial \mathbf{p}^s}{\partial n} \quad \text{with} \quad \mathbf{S} = \begin{bmatrix} \Delta S_1 & & & 0 \\ & \Delta S_2 & & \\ & & \ddots & \\ 0 & & & \Delta S_N \end{bmatrix}, \quad (5.30)$$

where \mathbf{M}^{hs} and \mathbf{D}^{hs} denote the source model of monopole and dipole respectively, the form of propagation matrix as follow:

$$M^{hs}(\mathbf{z}_i, \mathbf{x}_m, \omega) = \frac{e^{-jkr}}{4\pi r}, \quad (5.31)$$

$$D^{hs}(\mathbf{z}_i, \mathbf{x}_m, \omega) = \mathbf{n}_{\mathbf{z}_i} \cdot \nabla_h \frac{e^{-jkr}}{4\pi r} = \mathbf{n}_{\mathbf{z}_i} \cdot \mathbf{e}_r \frac{1+jkr}{r} \frac{e^{-jkr}}{4\pi r}, \quad (5.32)$$

where $r = |\mathbf{x}_m - \mathbf{z}_i|$ denotes the distance between the source point \mathbf{z}_i on the source surface S^s and the microphone point \mathbf{x}_m on the surface S^h , $\mathbf{n}_{\mathbf{z}_i}$ is outward unit normal vector of \mathbf{z}_i ; ∇_h is gradient vector of \mathbf{x}_m , \mathbf{e}_r denotes the unit vector between \mathbf{x}_m and \mathbf{z}_i , k is the wave number, ω is angular frequency and ρ_0 is the density of the air. Assume $\boldsymbol{\sigma} = -S\partial\mathbf{p}^s / \partial n$ and $\boldsymbol{\mu} = S\mathbf{p}^s$ where \mathbf{p}^s is pressure on source surface S^s . Therefore, the holography equation can be rewritten as

$$\alpha\mathbf{p}^h = \mathbf{M}^{hs}\boldsymbol{\sigma} + \mathbf{D}^{hs}\boldsymbol{\mu}. \quad (5.33)$$

The \mathbf{p}^h denotes measured sound pressure on microphone surface when $\alpha = 1$ in Eq. (5.33). In addition, constraint equation (on source surface) can be obtained when $\alpha = 1/2$ as follows:

$$\frac{1}{2}\mathbf{p}^s = \mathbf{M}^{ss}\boldsymbol{\sigma} + \mathbf{D}^{ss}\boldsymbol{\mu}. \quad (5.34)$$

Assume $\Delta S_1 = \Delta S_2 = \dots = \Delta S_N = \Delta S$ and thus $\mathbf{p}^s = \boldsymbol{\mu} / \Delta S$. By doing so, the Eq. (5.34) can be rewritten as

$$\frac{1}{2\Delta S}\boldsymbol{\mu} = \mathbf{M}^{ss}\boldsymbol{\sigma} + \mathbf{D}^{ss}\boldsymbol{\mu} \Rightarrow 0 = \mathbf{M}^{ss}\boldsymbol{\sigma} + (\mathbf{D}^{ss} - \frac{1}{2\Delta S}\mathbf{I})\boldsymbol{\mu}, \quad (5.35)$$

where \mathbf{I} being an identity matrix.

In the measurement of radiated sound field by microphones, singularity of the diagonal elements are generated due to the field points are located on the source surface S^s . In order to solve the singularity of the diagonal elements \mathbf{M}^{ss} and \mathbf{D}^{ss} , it must be properly dealt with the following two integrals on a ‘singular element’. First, the \mathbf{M}^{ss} is weak singularity matrix form as

$$\int_{\Delta S} \frac{e^{-jkr}}{4\pi r} dS = \int_0^{2\pi} \int_0^R \frac{e^{-jkr}}{4\pi r} r dr d\theta = \frac{R \sin(kR/2)}{2 kR/2} e^{-jkR/2}, \quad (5.36)$$

where $r = |\mathbf{z}_i - \mathbf{z}_i|$ denotes the distance on the same source point \mathbf{z}_i on the source surface S^s , and ΔS is a small disk element with radius $R = \sqrt{\Delta S / \pi}$. Finally, the \mathbf{D}^{ss} is strong singularity matrix form as follow:

$$\lim_{r \rightarrow 0} \int_{\Delta S} \frac{\partial}{\partial n} \left(\frac{e^{-jkr}}{4\pi r} \right) dS = \lim_{r \rightarrow 0} \int_{\Delta S} \frac{\partial}{\partial r} \left(\frac{e^{-jkr}}{4\pi r} \right) \frac{\partial r}{\partial n} dS = 0, \quad (5.37)$$

where $\partial r / \partial n = \mathbf{n} \cdot \nabla r \equiv 0$ on ΔS .

The following equation combines Eqs. (5.33) and (5.34) to solve for the unknown $\boldsymbol{\sigma}$ and $\boldsymbol{\mu}$

$$\begin{bmatrix} \mathbf{M}^{hs} & \mathbf{D}^{hs} \\ \mathbf{M}^{ss} & \mathbf{D}^{ss} \end{bmatrix} \begin{bmatrix} \boldsymbol{\sigma} \\ \boldsymbol{\mu} \end{bmatrix} = \begin{bmatrix} \mathbf{p}^h \\ \mathbf{p}^s \end{bmatrix} \Rightarrow \mathbf{A} \begin{bmatrix} \boldsymbol{\sigma} \\ \boldsymbol{\mu} \end{bmatrix} = \begin{bmatrix} \mathbf{p}^h \\ \mathbf{p}^s \end{bmatrix}. \quad (5.38)$$

Apparently, equation (5.38) is the inverse problem and the matrix \mathbf{A} is usually ill-posed and even non-square. Therefore, the unknown pressure gradient $\boldsymbol{\sigma}$ and pressure $\boldsymbol{\mu}$ on the virtual surface can be calculated by

$$\begin{bmatrix} \boldsymbol{\sigma} \\ \boldsymbol{\mu} \end{bmatrix} = \mathbf{A}^+ \begin{bmatrix} \mathbf{p}^h \\ \mathbf{p}^s \end{bmatrix}. \quad (5.39)$$

where \mathbf{A}^+ is pseudo-inverse matrix from \mathbf{A} by TSVD or Tikhonov regularization.

The sound pressure \mathbf{p}^r and particle velocity \mathbf{u}^r on the reconstruction surface S^r can be calculated by

$$\begin{cases} \mathbf{p}^r = \mathbf{M}^{rv} \boldsymbol{\sigma} + \mathbf{D}^{rv} \boldsymbol{\mu} \\ \mathbf{u}^r = \mathbf{D}'^{rv} \boldsymbol{\sigma} + \mathbf{Q}^{rv} \boldsymbol{\mu} \end{cases} \quad (5.40)$$

where \mathbf{M}^{rv} denotes the monopole source model, \mathbf{D}^{rv} and \mathbf{D}'^{rv} denote the dipole source model, \mathbf{Q}^{rv} denotes the quadrupole source model. The unknown \mathbf{D}^{rv} and \mathbf{Q}^{rv} can be calculated by

$$D'^{rv}(\mathbf{y}_n, \mathbf{z}_i, \omega) = \frac{j}{\rho_0 \omega} \mathbf{n}_{\mathbf{y}_n} \cdot \nabla_r \frac{e^{-jkr}}{4\pi r} = \frac{j}{\rho_0 \omega} \frac{-1-jkr}{r} \frac{e^{-jkr}}{4\pi r}, \quad (5.41)$$

$$Q^{rv}(\mathbf{y}_n, \mathbf{z}_i, \omega) = \frac{j}{\rho_0 \omega} (\mathbf{n}_{\mathbf{y}_n} \cdot \nabla_s)(\mathbf{n}_{\mathbf{z}_i} \cdot \nabla_v) \frac{e^{-jkr}}{4\pi r}, \quad (5.42)$$

where $r = |\mathbf{z}_i - \mathbf{y}_n|$ denotes the distance between the source point \mathbf{y}_n on the virtual source surface S_v and the reconstructed point \mathbf{z}_i on the reconstruction surface S^s , $\mathbf{n}_{\mathbf{y}_n}$ and $\mathbf{n}_{\mathbf{z}_i}$ are outward unit normal vector of \mathbf{y}_n and \mathbf{z}_i respectively; and ∇_s and ∇_v are the gradient vectors of \mathbf{z}_i and \mathbf{y}_n , respectively.

5.3.2 Indirect ESM

The basic idea of the ESM is to model sound field by using distribution of virtual simple sources, as shown in Fig. 48. In the figure, \mathbf{x}_m is the m th microphone position on the hologram surface S^h , \mathbf{z}_i is the i th source point on the actual source surface S^s , and \mathbf{y}_n is the n th virtual source point on the virtual surface S^v . We shall formulate the ESM-based NAH in two source configurations illustrated in Fig. 49, which is more practical in modeling actual continuous sources. In source configuration 1, the virtual source surface (S^v) stays away from the actual source surface S^s with a non-zero RD. Source configuration 2 is applied to only planar sources for which we distribute virtual sources right on the actual source surface S^s (RD = 0 in this case).

A. Virtual source configuration 1:

The ESM can be formulated by discretizing the integral of simple layer potential.

$$p^h(\mathbf{x}) = \int_{S^v} \sigma(\mathbf{y}) G^{hv}(\mathbf{x}, \mathbf{y}) dS^v(\mathbf{y}), \quad (5.43)$$

where $\sigma(\mathbf{y})$ is an unknown source strength of the point source distribution, p^h is

the hologram pressure measured on the surface S^h , \mathbf{x} and \mathbf{y} are the position vectors of source point and field point, respectively. The propagation matrix consists of elements expressed as the free-space Green's function for the baffled piston

$$G^{hv}(\mathbf{x}, \mathbf{y}) = \frac{e^{-jkr}}{4\pi r}, \quad (5.44)$$

where j denotes $\sqrt{-1}$, $k = \omega/c$ is wave number, c is speed of sound, ω is angular frequency, and $r = |\mathbf{x} - \mathbf{y}|$. Straightforward discretization of Eq. (5.43) leads to

$$\begin{aligned} p^h(\mathbf{x}) &= \int_S \sigma(\mathbf{y}) G^{hv}(\mathbf{x}, \mathbf{y}) dS^v(\mathbf{y}) \approx \sum_{n=1}^N \int_{S_n^v} \sigma(\mathbf{y}) G^{hv}(\mathbf{x}, \mathbf{y}) dS^v(\mathbf{y}) \\ &\approx \sum_{n=1}^N [\sigma(\mathbf{y}_n) \hat{S}_n^v] G^{hv}(\mathbf{x}, \mathbf{y}_n) = \sum_{n=1}^N q_n^v G^{hv}(\mathbf{x}, \mathbf{y}_n), \end{aligned} \quad (5.45)$$

where \hat{S}_n^v is the area of the n th element, $\mathbf{y}_n \in \hat{S}_n^v$, $q_n^v = \sigma(\mathbf{y}_n) \hat{S}_n^v$ represents the strength of the n th point source, and N is the number of point source. It should be borne in mind that these virtual point sources only constitute an equivalent discrete representation in that they would produce the same field, $p^h(\mathbf{x})$. The amplitudes of these point sources represent relative source strength associated with each source location, which could be due in part to physical sources and/or the effects nearby boundary reflection. These N virtual sources are assumed to be located at the desired focal points. Equation (5.43) can be discretized into the matrix form

$$\mathbf{p}^h = j\rho_0\omega \mathbf{G}^{hv} \mathbf{q}^v, \quad (5.46)$$

where ρ_0 is the air density, \mathbf{p}^h represents the hologram pressure vector, \mathbf{q}^v represents the virtual source strength vector, and \mathbf{G}^{hv} is the propagation matrix relating the source volume velocity and the hologram pressure. The unknown virtual source strengths can be calculated by inverting Eq. (5.46) as

$$\hat{\mathbf{q}}^v = \frac{1}{j\rho_0\omega} \mathbf{G}^{hv+} \mathbf{p}^h, \quad (5.47)$$

where $\hat{\mathbf{q}}^v$ is the estimated source strength vector and \mathbf{G}^{hv+} is the pseudo-inverse

matrix of \mathbf{G}^{hv} . The TSVD or Tikhonov regularization can be used to deal with the ill-conditioned inversion process.

Once the source strength $\hat{\mathbf{q}}^v$ is obtained, acoustical variables can be reconstructed on the actual source surface. Due to singularity of virtual point sources, we need a non-zero RD between the virtual source surface (S^v) and the actual source surface S^s to assure smooth reconstruction results, which is referred to as source configuration 1. Thus, the sound pressure on S^s can be reconstructed as

$$\mathbf{p}^s = j\rho_0\omega\mathbf{G}^{sv}\hat{\mathbf{q}}^v, \quad (5.48)$$

where \mathbf{G}^{sv} denotes the propagation matrix relating the virtual source strength and the actual source surface pressure. In addition, the actual source surface normal velocity can be reconstructed as

$$u^s(\mathbf{z}_i) = \sum_{n=1}^N (\mathbf{n} \cdot \mathbf{e}_r) \frac{1 + jkr_m}{r_m} \hat{q}_n^v G^{sv}(\mathbf{z}_i, \mathbf{y}_n), \quad (5.49)$$

where $r_{na} = |\mathbf{z}_i - \mathbf{y}_n|$, $\mathbf{e}_r = (\mathbf{z}_i - \mathbf{y}_n)/r_m$ is the unit vector pointing from the n th virtual source to the i th reconstruction point, and \mathbf{n} is the outward unit vector normal to the actual source surface.

B. Virtual source configuration 2:

For planar sources, one can use source configuration 2 in which virtual sources are distributed right on the actual source surface S^s , so the RD is zero in this case. The pressure field produced by a planar source can be expressed as the Rayleigh's integral [73]

$$p^h(\mathbf{x}) = 2j\rho_0\omega \int_{S^v} G^{hv}(\mathbf{x}, \mathbf{y}) u_0(\mathbf{y}) dS^v(\mathbf{y}), \quad (5.50)$$

where u_0 is the source surface velocity of virtual sources on the surface S^v and $G^{hv}(\mathbf{x}, \mathbf{y})$ is as defined previously. Equation (5.50) can be expressed in a matrix form of Eq. (5.51) by using zeroth-order discretization with constant elements. The

measured pressures on the hologram surface and the velocities on the resulting can be represented in the matrix form

$$\mathbf{p}^h = 2j\rho_0\omega\mathbf{G}^{hv}\mathbf{q}^v \quad (5.51)$$

The velocity on the actual source surface can be calculated by

$$\mathbf{u}^v = \frac{1}{\Delta S}\mathbf{q}^v = \frac{1}{j\rho_0\omega\Delta S}\mathbf{G}^{hv+}\mathbf{p}^h, \quad (5.52)$$

where ΔS denotes the average area per “element” on the actual source surface.



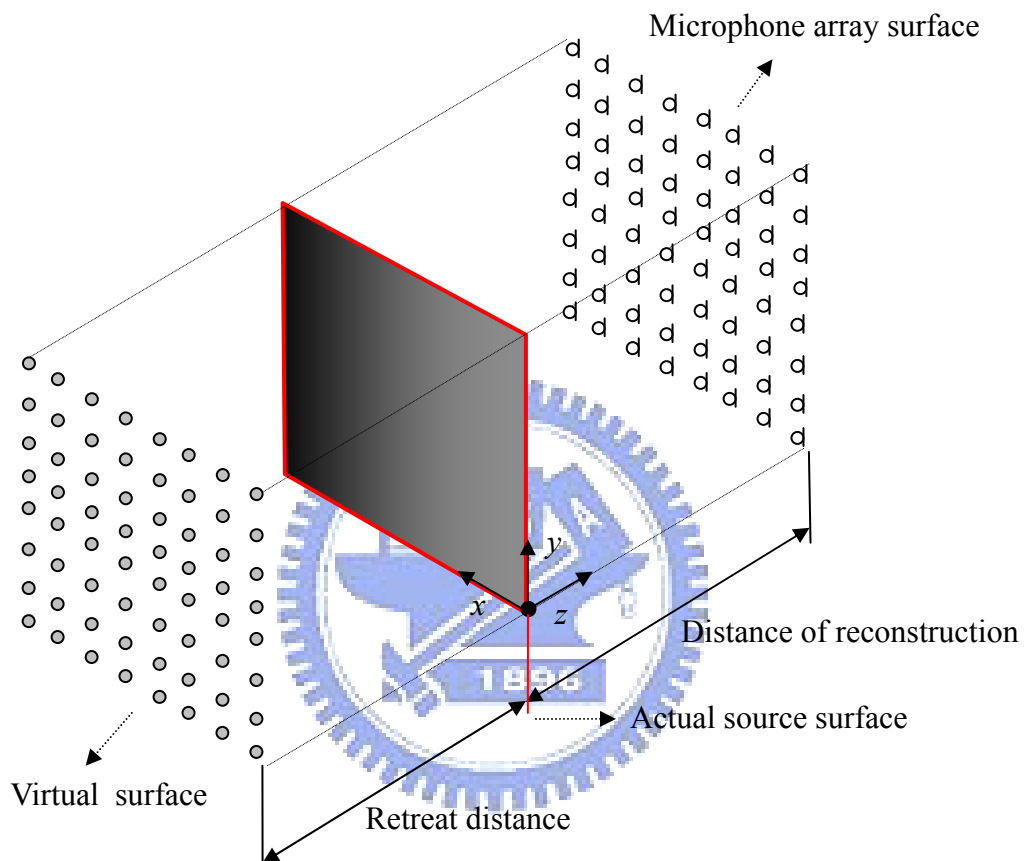


Fig. 49 Schematic diagram for planar sources in configurations 1 and 2.

5.3.3 Nearfield Equivalent Source Imaging (NESI)

In the NESI formulation, the target source is simulated by a collection of point sources distributed on a surface. For that source configuration, one can simply choose virtual sources to be the target sources. The free-space Green's function between the source point \mathbf{y} and the field point \mathbf{x} as

$$G(\mathbf{x}, \mathbf{y}) = \frac{e^{-jkr}}{r}. \quad (5.53)$$

Consider a NSI problem depicted in Fig. 50, where sound pressures radiated by a source are received at the M microphones. The measured sound pressures and the source amplitudes are related by

$$\begin{bmatrix} p^h(\mathbf{x}_1) \\ p^h(\mathbf{x}_2) \\ \vdots \\ p^h(\mathbf{x}_M) \end{bmatrix} = \begin{bmatrix} G(\mathbf{x}_1, \mathbf{y}_1) & G(\mathbf{x}_1, \mathbf{y}_2) & \cdots & G(\mathbf{x}_1, \mathbf{y}_N) \\ G(\mathbf{x}_2, \mathbf{y}_1) & G(\mathbf{x}_2, \mathbf{y}_2) & \cdots & G(\mathbf{x}_2, \mathbf{y}_N) \\ \vdots & \vdots & \ddots & \vdots \\ G(\mathbf{x}_M, \mathbf{y}_1) & G(\mathbf{x}_M, \mathbf{y}_2) & \cdots & G(\mathbf{x}_M, \mathbf{y}_N) \end{bmatrix} \begin{bmatrix} q^v(\mathbf{y}_1) \\ q^v(\mathbf{y}_2) \\ \vdots \\ q^v(\mathbf{y}_N) \end{bmatrix}. \quad (5.54)$$

Or, in the matrix form

$$\mathbf{p}^h = \mathbf{G}^{hv} \mathbf{q}^v. \quad (5.55)$$

Since the number of microphones is usually no greater than that of the focused points ($M \leq N$), the propagation matrix $\mathbf{G} \in \mathbb{C}_{M \times N}$ could be non-square and the problem could be underdetermined. The purpose here is to estimate \mathbf{q}^v based on the measurement \mathbf{p}^h . This can be regarded as a model matching problem depicted in Fig. 51, where the propagation matrix \mathbf{G} has the source amplitude $\mathbf{q}^v \in \mathbb{C}_{N \times 1}$ and the pressure measurement $\mathbf{p}^h \in \mathbb{C}_{M \times 1}$ as its input and output, $\mathbf{e} \in \mathbb{C}_{J \times 1}$ is the matching error where the J is number of focal points, $\mathbf{C} \in \mathbb{C}_{J \times M}$ is the inverse filter matrix, and $\mathbf{W} \in \mathbb{C}_{J \times N}$ is the matching model. With the most basic choice of \mathbf{W} being the identity matrix \mathbf{I} , the problem reduces to, given \mathbf{p} and \mathbf{G} , finding an inverse filter \mathbf{C} such that $\mathbf{CG} \approx \mathbf{W} = \mathbf{I}$ and hence

$$\hat{\mathbf{q}} = \mathbf{C}\mathbf{p} = \mathbf{C}\mathbf{G}\mathbf{q} \approx \mathbf{q}.$$

Inverse matrix \mathbf{C} can be obtained by TSVD of the propagation matrix as

$$\mathbf{C} = \sum_{i=1}^I \frac{\mathbf{u}_i^H}{\sigma_i} \mathbf{v}_i, \quad (5.56)$$

where I is the discrete regularization parameter., \mathbf{u}_i is the i th left singular vector, \mathbf{v}_i is the i th right singular vector and σ_i is the i th singular value.

Alternatively, inverse matrix \mathbf{C} can be obtained by Tikhonov regularization as

$$\mathbf{C} = \mathbf{W}\mathbf{G}^H (\mathbf{G}\mathbf{G}^H + \beta\mathbf{I})^{-1}, \quad (5.57)$$

where $\beta \geq 0$ is a regularization parameter that can either be a constant or frequency-dependent.

An NAH method termed the NESI was proposed for NSI and sound field visualization. The NESI *per se* can be considered as a time-domain ESM. It is noted that the inverse filters \mathbf{C} obtained are still in the frequency-domain. Inverse FFT is called for to convert and truncate these FRFs into FIR filters in the time-domain. In this step, circular shift is needed to ensure that the resulting filters are causal filters. Now that the inverse filters are available, the amplitude of the virtual sources at the virtual source points can be calculated by using multichannel convolution:

$$\hat{q}_n(n') = \sum_{m=1}^M p_m(n') * C_{nm}(n'), \quad 1 \leq n \leq N, \quad (5.58)$$

where n' denotes the discrete-time index, $C_{nm}(n')$ denotes the impulse response of the nm th inverse filter and “*” denotes the convolution operator. The thus obtained virtual source strength at the virtual source point will serve as the basis for subsequent sound field reconstruction.

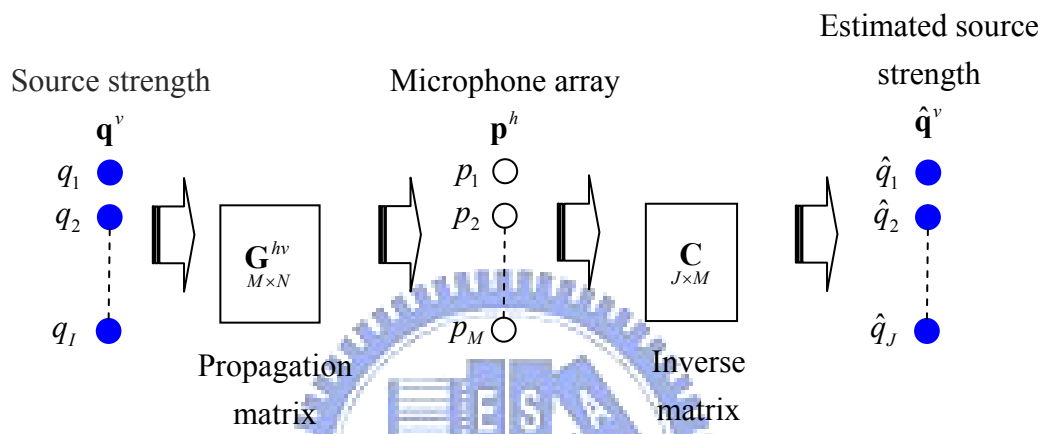


Fig. 50 The block diagram of inverse filtering process.

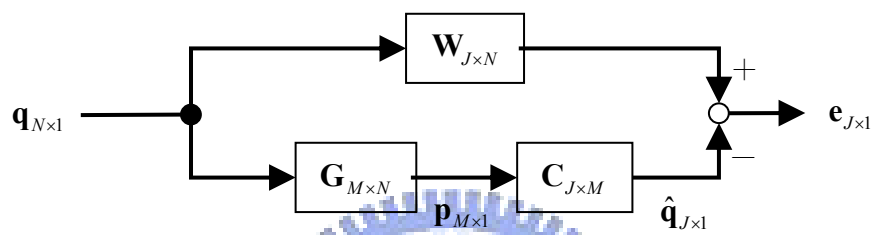


Fig. 51 The inverse filtering process viewed as a model matching problem.



5.3.4 Kalman filter-based algorithm

Kalman filter is a model-based adaptive filtering approach relying on a known propagation model. In terms of state space formulation, a generic structure of a Kalman filter involves two equations. The first equation is the *process equation*:

$$\mathbf{x}(n+1) = \mathbf{F}(n+1, n)\mathbf{x}(n) + \mathbf{v}_1(n), \quad (5.59)$$

where $\mathbf{F}(n+1, n)$ is a known $M \times M$ state transition matrix relating the state of the system at times $n+1$ and n . The $M \times 1$ vector $\mathbf{v}_1(n)$ is process noise.

The vector $\mathbf{v}_1(n)$ is a zero-mean, white-noise process whose correlation matrix is defined as

$$E[\mathbf{v}_1(n)\mathbf{v}_1^H(n')] = \begin{cases} \mathbf{Q}_1, & n = n' \\ \mathbf{0}, & n \neq n' \end{cases} \quad (5.60)$$

The second equation is the *measurement equation*:

$$y(n) = \mathbf{C}(n) \mathbf{x}(n) + \mathbf{v}_2(n), \quad (5.61)$$

where $\mathbf{C}(n)$ is a known $N \times M$ measurement matrix. The $N \times 1$ vector $\mathbf{v}_2(n)$ is a zero-mean, white-noise measurement noise whose correlation matrix is defined as

$$E[\mathbf{v}_2(n)\mathbf{v}_2^H(n')] = \begin{cases} \mathbf{Q}_2, & n = n' \\ \mathbf{0}, & n \neq n' \end{cases} \quad (5.62)$$

The noise vectors $\mathbf{v}_1(n)$ and $\mathbf{v}_2(n)$ are assumed statistically independent

$$E[\mathbf{v}_1(n)\mathbf{v}_2^H(n')] = \mathbf{0}. \quad (5.63)$$

From the perspective of the linear system theory, the *process equation* and the *measurement equation* of the acoustic testing in the time-domain can be written as follows:

$$\begin{bmatrix} p(\mathbf{x}_1, t) \\ \vdots \\ p(\mathbf{x}_M, t) \end{bmatrix} = \begin{bmatrix} \frac{\delta(t - \frac{r_{11}}{c})}{r_{11}} & \dots & \frac{\delta(t - \frac{r_{1S}}{c})}{r_{1S}} \\ \vdots & \ddots & \vdots \\ \frac{\delta(t - \frac{r_{M1}}{c})}{r_{M1}} & \dots & \frac{\delta(t - \frac{r_{MS}}{c})}{r_{MS}} \end{bmatrix} * \begin{bmatrix} a_1(t) \\ \vdots \\ a_S(t) \end{bmatrix} \Rightarrow \mathbf{p}_h = \boldsymbol{\delta}_{hv} * \mathbf{a}_v, \quad (5.64)$$

where ρ_0 is the air density, \mathbf{p}_h represents the hologram pressure vector, \mathbf{a}_v represents the virtual source amplitude vector, and $\boldsymbol{\delta}_{hv}$ is the Dirac Delta function matrix relating the source amplitude and the hologram pressure. The virtual sources are distributed right on the actual source surface S_s , so RD is zero in this case.

The ERA establishes the following minimal state space realization for linear systems:

$$\mathbf{x}_s(n+1) = \mathbf{A}_e \mathbf{x}_s(n) + \mathbf{B}_e \mathbf{a}(n), \quad (5.65)$$

$$\mathbf{p}(n) = \mathbf{C}_e \mathbf{x}_s(n) + \mathbf{D}_e \mathbf{a}(n), \quad (5.66)$$

where n is the discrete-time index, $\mathbf{x}_s(n)$ is the state vector, $\mathbf{a}(n)$ is the $S \times 1$ input vector, $\mathbf{p}(n)$ is the $M \times 1$ output vector, and \mathbf{A}_e , \mathbf{B}_e , \mathbf{C}_e and \mathbf{D}_e are constant matrices. The ERA starts with the impulse response matrices of the Dirac Delta function ($\boldsymbol{\delta}_{hv}$). Assume the source amplitude functions satisfy the smoothness condition described by the following state equation:

$$\mathbf{a}(n+1) = \mathbf{A} \mathbf{a}(n), \quad (5.67)$$

where \mathbf{A} being an identity matrix. Define compound state variables:

$$\mathbf{x}(n) = \begin{bmatrix} \mathbf{x}_s(n) \\ \mathbf{a}(n) \end{bmatrix}, \quad (5.68)$$

then it can be written as follows

$$\begin{bmatrix} \mathbf{x}_s(n+1) \\ \mathbf{a}(n+1) \end{bmatrix} = \begin{bmatrix} \mathbf{A}_e & \mathbf{B}_e \\ \mathbf{0} & \mathbf{A} \end{bmatrix} \begin{bmatrix} \mathbf{x}_s(n) \\ \mathbf{a}(n) \end{bmatrix} \quad (5.69)$$

$$\mathbf{p}(n) = \begin{bmatrix} \mathbf{C}_e & \mathbf{D}_e \end{bmatrix} \begin{bmatrix} \mathbf{x}_s(n) \\ \mathbf{a}(n) \end{bmatrix}$$

or more concisely formulation

$$\begin{aligned} \mathbf{x}(n+1) &= \mathbf{F}(n+1, n)\mathbf{x}(n) + \mathbf{v}_1(n) \\ \mathbf{p}(n) &= \mathbf{C}(n)\mathbf{x}(n) + \mathbf{v}_2(n) \end{aligned} \quad (5.70)$$

However, this Kalman filtering problem can not be solved by the steady-state Riccati equation because $\mathbf{C}(n)$ is a time-varying matrix, but rather the following recursive algorithm:

$$\mathbf{G}(n) = \mathbf{F}(n+1, n)\mathbf{K}(n, n-1)\mathbf{C}^H(n)[\mathbf{C}(n)\mathbf{K}(n, n-1)\mathbf{C}^H(n) + \mathbf{Q}_2(n)]^{-1}, \quad (5.71)$$

$$\boldsymbol{\alpha}(n) = \mathbf{p}(n) - \mathbf{C}(n)\hat{\mathbf{x}}(n | p_{n-1}), \quad (5.72)$$

$$\hat{\mathbf{x}}(n+1 | p_n) = \mathbf{F}(n+1, n)\hat{\mathbf{x}}(n | p_{n-1}) + \mathbf{G}(n)\boldsymbol{\alpha}(n), \quad (5.73)$$

$$\mathbf{K}(n) = \mathbf{K}(n, n-1) - \mathbf{F}(n, n+1)\mathbf{G}(n)\mathbf{C}(n)\mathbf{K}(n, n-1), \quad (5.74)$$

$$\mathbf{K}(n+1, n) = \mathbf{F}(n+1, n)\mathbf{K}(n)\mathbf{F}^H(n+1, n) + \mathbf{Q}_1(n), \quad (5.75)$$

To initialize the algorithm, the initial conditions are taken to be: $\hat{\mathbf{x}}(1 | p_0) = \mathbf{0}$, $\mathbf{K}(1,0) = \mathbf{I}$, with \mathbf{I} being an identity matrix. The block diagram of the recursive Kalman filter is shown in Fig. 52.

Numerical simulations are conducted to demonstrate the performance of ESM-based NAH by using the Kalman filter algorithm. In this formulation, the target source is simulated by a collection of point sources distributed on a planar surface. Source configuration 2 is applied to only planar sources for which we distribute virtual sources right on the actual source surface S_s (RD = 0 in this case). Suppose that the array comprises 5×5 microphones uniformly spaced with distance $d = 0.1\text{m}$ ($d = \lambda/2$ at $f = 1.7\text{ kHz}$), as shown in Fig. 53. The number and spacing of microphones and virtual sources are of identical settings. The point sources

embedded at the plane $z = 0\text{m}$ is arranged in four scenarios as denoted by $S_{s,i}$, $i = 1-2$ respectively, as shown in Fig. 53. In the simulations, the desired amplitude vector was chosen according to the source type. The random sources with equal strength and band-limited to 2.5 kHz are assumed to be on the actual source surface. The parameters assumed for the Kalman filter based simulation are as follows: $\mathbf{Q}_1(n)=0.1\mathbf{I}$, $\mathbf{Q}_2(n)=0.01\mathbf{I}$, with \mathbf{I} being an identity matrix and the sampling rate is $f_s=1$ kHz. The time-domain magnitudes of the surface amplitude reconstructed using inversion and Kalman filter method in $S_{s,1}$ and $S_{s,2}$ are shown in Fig. 54. In the scenario $S_{s,1}$, the relative errors of the reconstructed amplitude thus calculated are 14.5% for inversion method and 10.8% for Kalman filter method, respectively. Furthermore, the relative errors are calculated for previous methods in the scenario $S_{s,2}$. The error percentage calculated is 16.8% for inversion and 30.2% for Kalman filter. The method of Kalman filter, which can be regarded as an extended observer with the constraint of Gaussian noise, can be used to include practical imperfections such as multi-path, reflection, and sensor noise in the state model. In summary the present observer-based algorithm is able to be performed recursively over each sampling block. The outcome is still comparable to the corresponding inversion result.

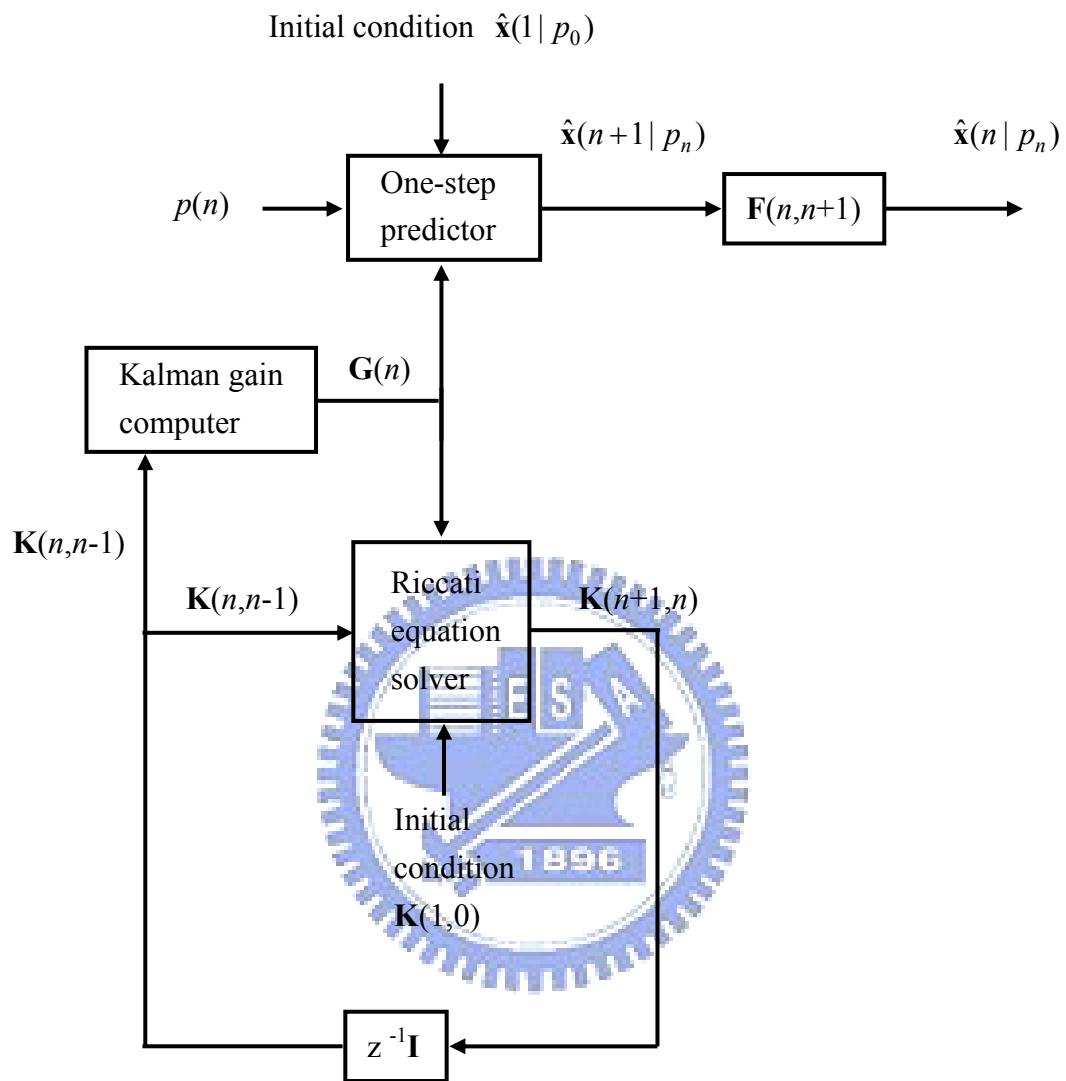
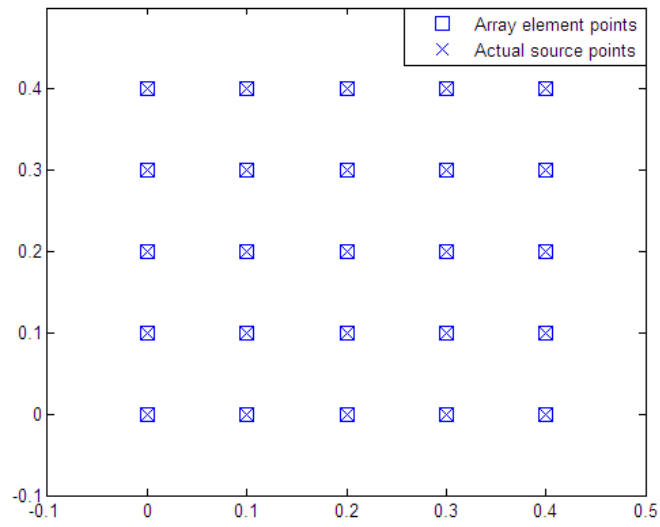
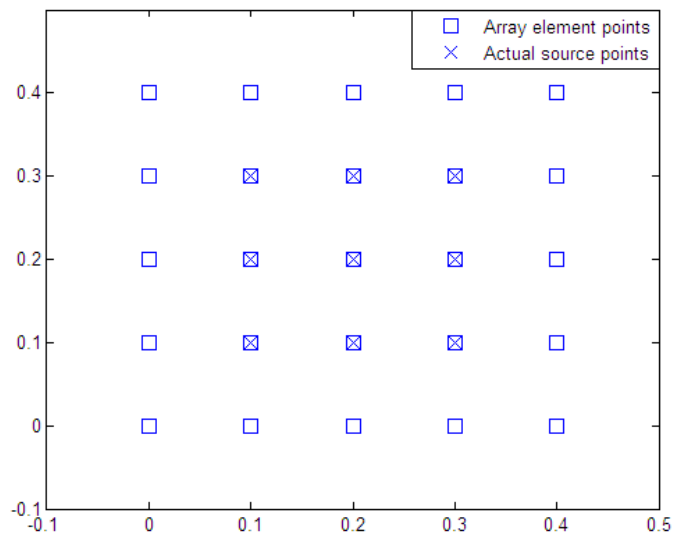


Fig. 52 Block diagram of the Kalman filter.

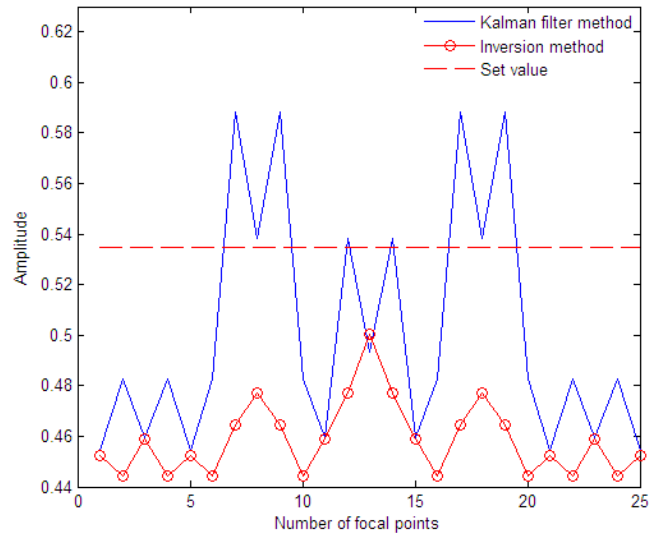


(a)

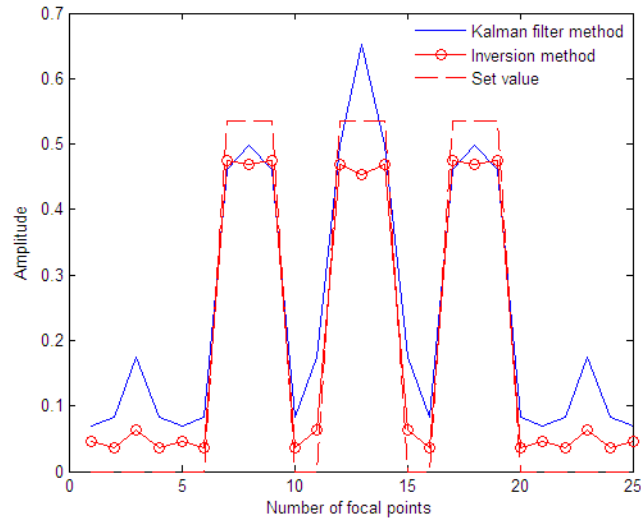


(b)

Fig. 53 Scenarios of array element and actual source point distribution for simulating the point sources. (a) $S_{s,1}$ and (b) $S_{s,2}$.



(a)



(b)

Fig. 54 The reconstructed and the desired amplitude profiles of point sources scenario. (a) $S_{s,1}$ and (b) $S_{s,2}$.

5.3.5 Choice of nearfield array parameters

In configuring the array, there are many design factors to be taken into account. An in-depth investigation conducted in Refs. [39] and [68] has arrived at the following conclusions. The reconstruction performance achievable by the array is dependent on the condition number of the propagation matrix. The inverse problem is well conditioned when the number of sources and sensors is small, when the geometrical arrangement of sensors is conformal to the source geometry in terms of location, orientation and spacing, when the sensor array is positioned symmetrically with respect to the source array, and when the DOR is small.

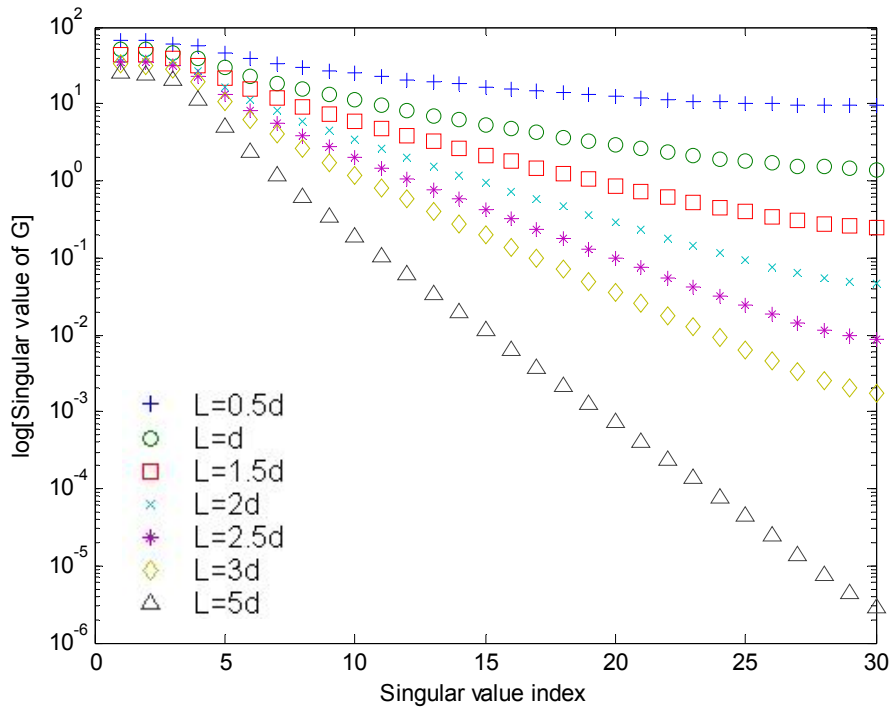
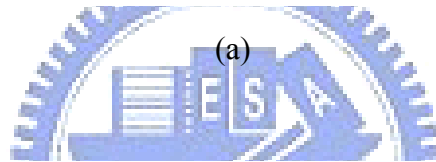
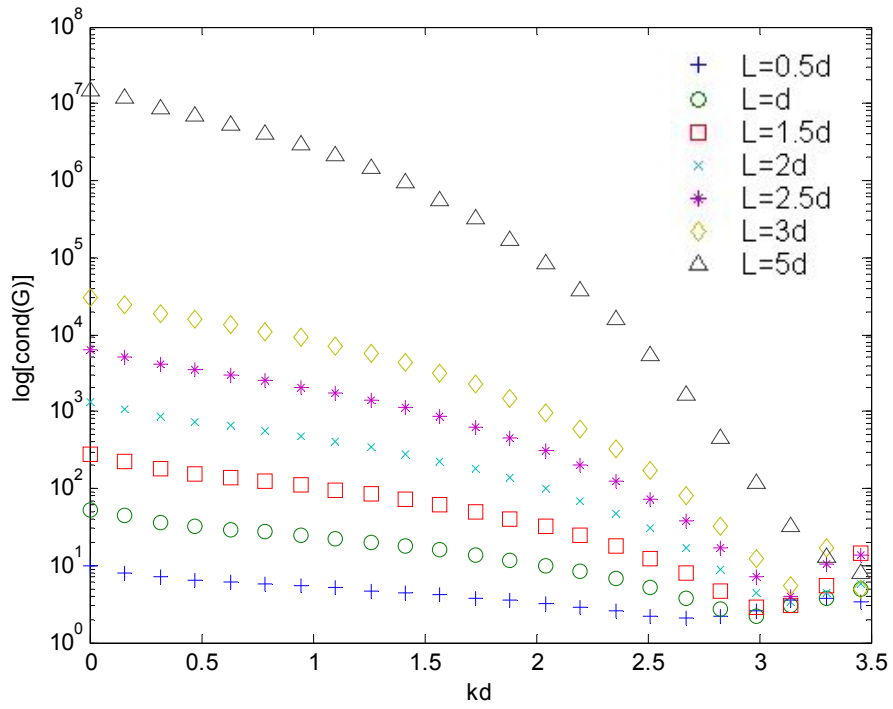
The last point regarding the DOR is crucial to the performance of the NESI. It is well known in acoustics that the evanescent waves carrying the details in the nearfield will rapidly decay with distance to the farfield. The propagation matrix then becomes ill-conditioned as one attempts to reconstruct the sound image based on an already smoothed farfield measurement. A numerical simulation conducted for a 1D array shows in more detail what the effects of some array parameters have on the matrix conditioning. In Fig. 55 (a), the condition number drops with increasing kd . The condition number of reconstruction with large L is higher than that of reconstruction with small L . Because condition number is defined as the ratio of the maximum singular value and the minimum singular value, the singular value plot in Fig. 55 (b) further exhibits the same trend of matrix conditioning versus the DOR. In this work, the condition number during sound field reconstruction is generally kept under 10^3 .

The procedures in choosing array parameters can be summarized as follows:

1. Choose the microphone spacing according to the maximum frequency (f_{\max}).
A conservative rule is $d = \lambda / 2$.
2. Choose the array aperture (D) that covers the source surface size.

3. The last two steps determine the number of microphone needed, $N_m = D/d$ for a ULA.
4. Choose the DOR according to the condition number of propagation matrix and the array parameters determined above. As a rule of thumb, we choose the condition number under 10^3 .





(b)

Fig. 55 Variation of condition number of propagation matrix for a 1D array. (a)

Condition number vs. kd , (b) singular value distribution of propagation matrix (frequency = 200 Hz, the spacing between microphones and between focal points $d = d_f = 0.0858$ m, $kd = 0.3142$). The parameter k is the wave number, d is the spacing of array, \mathbf{G} is the propagation matrix, and L is the DOR. There are 30 microphones and 30 focal points. The number of null points is one at each end.



5.4 Comparison of the nearfield algorithms

Fourier NAH enables reconstructing the 3D sound field from the 2D hologram data scanned above the source surface. Although Fourier NAH provides a clear picture of noise distribution, it suffers from a number of limitations which stem from the fact that NAH relies on the FFT of uniformly sampled pressure data on the hologram. To avoid spatial aliasing, the spacing of microphones is required to be less than one-half the wavelength. To avoid spatial wrap-around, the surrounding of the aperture would generally need to be padded with zeros. However, these two considerations often result in a large channel count.

NAH methods appropriate for arbitrarily shaped source were suggested, e.g., the NAH based on IBEM and the HELS method. The IBEM can be derived from a field representation on the basis of the Green's function theory, whereas the HELS method can be formulated using the field representation based on spherical harmonics. The IBEM can be further classified into two categories: the direct formulation and the indirect formulation. The former is derived from the HIE, whereas the latter stems from the simple layer or double layer potential. The advantages of IBEM are it allows for reconstruction of the acoustic quantities on an arbitrarily shaped structure, and it is suitable for exterior and interior regions. However, spatial discretization is employed to obtain acoustic quantities. Hence, IBEM must require a minimum number of nodes per wavelength to assure a satisfactory spatial resolution in reconstruction. For a complex structure, the number of discrete nodes required to accurately reconstruct the surface acoustic quantities can be very large. Therefore, the total number of measurement points may be large. In this condition, IBEM making the reconstruction process extremely time consuming. The major disadvantage of HELS is that can yield a good approximation for all surface geometries. HELS proved effective in imaging blunt and convex sources, but not for a highly elongated

one. The cylindrical coordinate system is ideal for a slender body, but not for a flat surface.

Another method, the ESM, also known as wave superposition method, was suggested for sound field calculation with far less complexity. The idea underlying the ESM is to represent sound field with discrete simple sources with no need to perform numerical integration. As opposed to the actual source, these solutions of simple sources deduced from the acoustic wave equation serve as the basis for sound field representation. While the ESM was often used as a benchmark for BEM, it has been shown with careful choice of parameters that the ESM is capable of achieving accuracy comparable to other methods. Like IBEM, the use of ESM is not restricted to source with regular geometries. The simplicity of the ESM lends itself very well to the implementation with digital signal processing and control paradigms. From such perspectives, the NESI previously proposed by the authors performs multichannel inverse filtering in the time-domain in light of minimal state-space realization of digital filters. Since all the required processing is carried out in the time-domain, NESI eliminates many problems of Fourier NAH. Not only sound field processing is entirely carried out in the time-domain but also nearfield details can be reconstructed. This technique is applicable to noise sources of all kinds, including narrowband, broadband, stationary, and transient types. Table 8 compares the aforementioned nearfield imaging methods.

Table 8 Comparison of nearfield imaging methods.

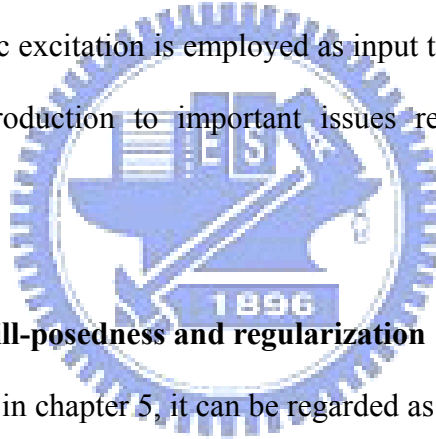
	Fourier NAH	IBEM	HELs	NESI
Algorithm	k -space FFT2D	matrix inverse	least-square	inverse filtering
Geometry	planar/regular	planar/regular	irregular	arbitrary/irregular
Mic spacing	$\lambda/2$	$\lambda/2$	$\approx \lambda/2$	$3\sim 4\lambda$
Spatial window effect	yes	no	no	no
DOR	nearfield	nearfield	nearfield	nearfield/farfield
Area covered	small	small	small	small/large
Domain	frequency	frequency	frequency	time
Real time	no	no	no	yes
Stationarity	yes	yes	yes	no
Reference	yes	yes	yes	no
Acoustic variables	yes	yes	yes	yes
Need scaling	no	no	no	no
Sensor	many	many	many	few

CHAPTER 6. PRACTICAL IMPLEMENTATIONS

This chapter outlines practical implementations such as inverse filter design, multi-channel fast filtering, post-processing, choice of DOR and lattice spacing, virtual microphone technique, choice of RD, optimization of sensor deployment, and system integration and experimental arrangement.

6.1 Inverse filter design

In general, inverse problems are based on an associated forward problem. This forward problem behind NAH is sound field model that presents the sound radiation from a source. Acoustic excitation is employed as input to predict the field pressure. This part gives an introduction to important issues regarding inverse problems explained from NAH.



6.1.1 Model matching: ill-posedness and regularization

In problem of ESM in chapter 5, it can be regarded as a model matching problem depicted in Fig. 51. Similar to the deconvolution process in NAH, the system matrix \mathbf{G} is usually ill-conditioned and even non-square. This calls for the need of optimization with proper regularization. The problem considered herein can be put into the following optimization formalism:

$$\min_{\mathbf{C}} \|\mathbf{W} - \mathbf{C}\mathbf{G}\|_F^2, \quad (6.1)$$

where $\|\cdot\|_F^2$ symbolizes the Frobenius norm [75] defined as, for an $J \times N$ matrix \mathbf{A} ,

$$\|\mathbf{A}\|_F^2 = \sum_{n=1}^N \sum_{\hat{n}=1}^{\hat{N}} |a_{\hat{n}n}|^2 = \sum_{n=1}^N \|\mathbf{a}_n\|_2^2. \quad (6.2)$$

Hence, the minimization problem of Frobenius-norm can be converted to the

minimization problem of the 2-norm by partitioning the matrices into columns. Since there is no coupling between the columns of the matrix \mathbf{C} , the minimization of the square of the Frobenius norm of the entire matrix \mathbf{G} is tantamount to minimizing each column independently.

$$\min_{\mathbf{c}_n} \|\mathbf{W} - \mathbf{C}\mathbf{G}\|_F^2 = \min_{\mathbf{c}_n} \sum_{n=1}^N \|\mathbf{w}_n^H - \mathbf{G}^H \mathbf{c}_n^H\|_2^2, \quad n=1, 2, \dots, N. \quad (6.3)$$

where \mathbf{w}_n and \mathbf{c}_n denote the n th columns of the matrices \mathbf{W} and \mathbf{C} , respectively, the superscript “ H ” denotes the Hermitian transpose. The least-squares solution to the problem above is given as

$$\mathbf{c}_n^H = (\mathbf{G}^H)^+ \mathbf{w}_n^H, \quad n=1, 2, \dots, N, \quad (6.4)$$

where the superscript “ $+$ ” denotes the pseudo-inverse. This optimal solution in least-square sense can be assembled into a more compact matrix form:

$$[\mathbf{c}_1 \quad \mathbf{c}_2 \quad \dots \quad \mathbf{c}_N]^H = (\mathbf{G}^H)^+ [\mathbf{w}_1 \quad \mathbf{w}_2 \quad \dots \quad \mathbf{w}_N]^H \quad (6.5)$$

or

$$\mathbf{C} = \mathbf{W}\mathbf{G}^+ \quad (6.6)$$

If the system matrix \mathbf{G} is of full-row rank, the pseudo-inverse is given as

$$\mathbf{G}^+ = \mathbf{G}^H (\mathbf{G}\mathbf{G}^H)^{-1} \quad (6.7)$$

Note that \mathbf{G}^+ is also the right inverse in that $\mathbf{G}\mathbf{G}^+ = \mathbf{I}$. Nevertheless, if \mathbf{G} is not of full-row rank, TSVD and Tikhonov regularization [32], [76] can be used to avoid singularity of $\mathbf{G}\mathbf{G}^H$. Inverse matrix \mathbf{C} can be obtained by TSVD of the propagation matrix as

$$\mathbf{C} = \sum_{i=1}^I \frac{\mathbf{u}_i^H}{\sigma_i} \mathbf{v}_i, \quad (6.8)$$

where I is the discrete regularization parameter., \mathbf{u}_i is the i th left singular vector, \mathbf{v}_i is the i th right singular vector and σ_i is the i th singular value.

Alternatively, inverse matrix \mathbf{C} can be obtained by Tikhonov regularization as

$$\mathbf{C} = \mathbf{W}\mathbf{G}^H (\mathbf{G}\mathbf{G}^H + \beta\mathbf{I})^{-1}, \quad (6.9)$$

where $\beta \geq 0$ is a regularization parameter that can either be a constant or frequency-dependent [77].

The right inverse is used for the inverse problem. If, instead, left inverse is used in the present formulation, the matrix product $\mathbf{G}^H\mathbf{G}$ can become extremely rank-deficient. Heavy regularization would be required to maintain numerical stability. The solution will be over-smoothed when over-regularization is chosen too low due to the removal of the high frequency components. If under-regularization is chosen, the solution has distortion because of the amplification of the noise. Hence, how to choose optimal regularization parameter is described in chapter 6.1.3.

6.1.2 Window design

One problem with the simple choice of the matching model $\mathbf{W} = \mathbf{I}$ is that defocusing problem may arise. This is a frequently encountered problem particularly in the neighborhood of the boundary of the virtual surface. To resolve the problem, a modified matching model is suggested as follows. In addition to the virtual source points, the outer ring of the virtual surface is padded with “null points” to restrain the level of reconstructed data outside the boundary, as shown in Fig. 56. This is analogous to the windowing technique in digital filter design. Thus, the matching model is modified into

$$\mathbf{W} = \begin{bmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \\ \hline 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{bmatrix}_{J \times N}, \quad (6.10)$$

where N diagonal entries “1” correspond to the virtual source points and “0” entries

at the bottom half correspond to the null points. With this choice of \mathbf{W} , it can be shown that

$$\hat{\mathbf{q}} = \mathbf{W}\mathbf{q} = [q_1 \quad q_2 \quad \dots \quad q_N \quad 0 \quad \dots \quad 0]^T. \quad (6.11)$$



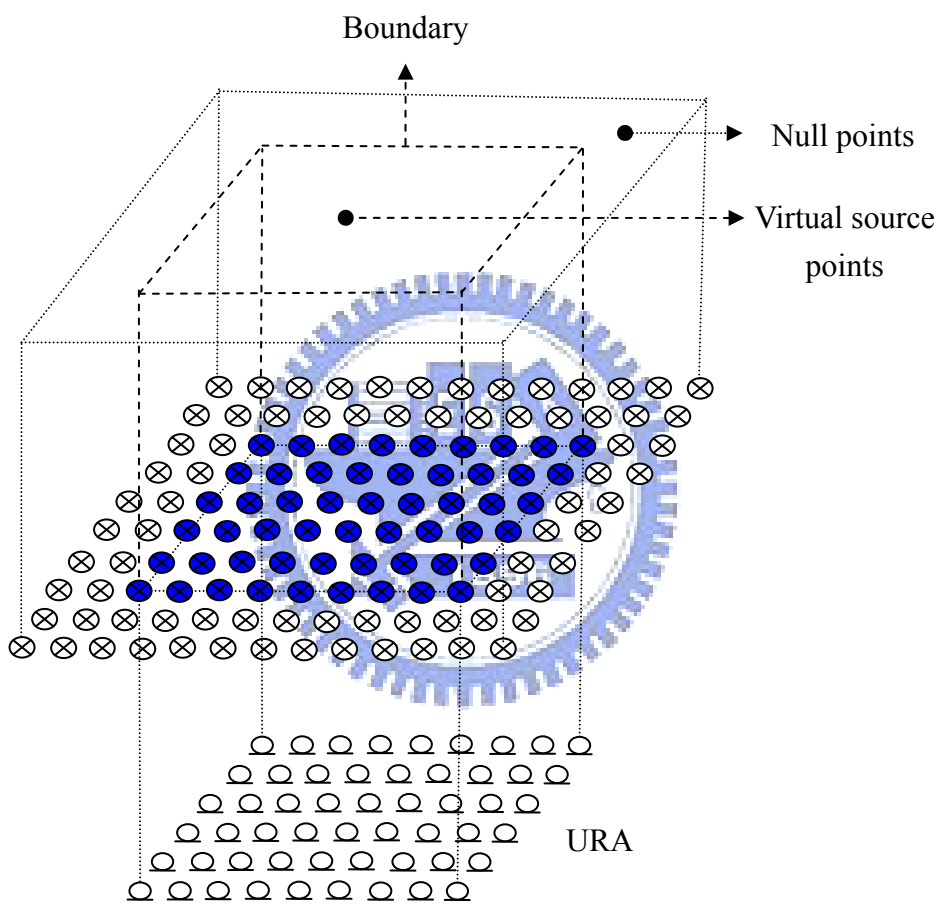


Fig. 56 Structure of 2D URA.

6.1.3 Parameter choice methods (PCM)

In this section, the generalized cross-validation (GCV) and L-curve criterion (L-C) are described to choose optimal regularization parameter. The idea of GCV is to remove an element from the vector of measurement data and then predict the missing data based on the remaining measurements [78], [79]. The optimal regularization parameter is to minimizing the GCV function that is

$$\text{GCV} = \frac{\|\mathbf{p} - \mathbf{G}\mathbf{q}_{\text{reg}}\|_2^2}{(\text{trace}(\mathbf{I} - \mathbf{G}\mathbf{G}^\#))}, \quad (6.12)$$

where \mathbf{q}_{reg} is the regularized solution vector and $\mathbf{G}^\#$ is

$$\mathbf{G}^\# = \sum_{i=1}^n f_i \mathbf{v}_i \sigma_i \mathbf{u}_i^H, \quad (6.13)$$

where n is the dimension of \mathbf{G} , \mathbf{u}_i is the i^{th} left singular vector, \mathbf{v}_i is the i^{th} right singular vector, σ_i is the i^{th} singular value and f_i is a low-pass filter. The residual vector can be calculate as

$$\begin{aligned} \mathbf{p} - \mathbf{G}\mathbf{q}_{\text{reg}} &= \mathbf{p} - \left(\sum_{i=1}^n \mathbf{u}_i \sigma_i \mathbf{v}_i^H \right) \left(\sum_{i=1}^n f_i \frac{\mathbf{u}_i^H \mathbf{p}}{\sigma_i} \mathbf{v}_i \right) \\ &= \mathbf{p} - \sum_{i=1}^n \mathbf{u}_i f_i \mathbf{u}_i^H \mathbf{p} \\ &= \sum_{i=1}^n \mathbf{u}_i \mathbf{u}_i^H \mathbf{p} - \sum_{i=1}^n \mathbf{u}_i f_i \mathbf{u}_i^H \mathbf{p} \\ &= \sum_{i=1}^n \mathbf{u}_i (1 - f_i) \mathbf{u}_i^H \mathbf{p}. \end{aligned} \quad (6.14)$$

Therefore, the residual norm in the numerator of Eq. (6.12) is

$$\|\mathbf{p} - \mathbf{G}\mathbf{q}_{\text{reg}}\|_2^2 = \sum_{i=1}^n (1 - f_i)^2 |\mathbf{u}_i^H \mathbf{p}|^2, \quad (6.15)$$

where \mathbf{u}_i from Eq. (6.14) is removed because of the orthonormality of the singular vectors. The denominator of Eq. (6.12) can be evaluated as

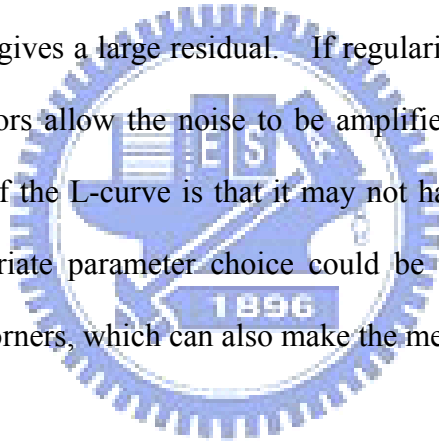
$$(\text{trace}(\mathbf{I} - \mathbf{G}\mathbf{G}^\#))^2 = \left(\text{trace}(\mathbf{I} - \sum_{i=1}^n \mathbf{u}_i f_i \mathbf{u}_i^H) \right)^2 = M - \sum_{i=1}^n f_i,$$

where M is equals the length of \mathbf{p} . The optimal regularization parameter is to minimizing the GCV function, as shown in Fig. 57.

Another PCM in this thesis is L-curve [78], [80]. The two norms are plotted against each other in a log-log scale, as shown in Fig. 58. The optimal solution is assumed to be at the corner of the L-curve. This corner is found at the point with maximum curvature. The solution norm can be found as

$$\|\mathbf{q}_{\text{reg}}\|_2^2 = \sum_{i=1}^n f_i^2 \frac{|\mathbf{u}_i^H \mathbf{p}|^2}{\sigma_i^2}$$

where \mathbf{v}_i from Eq. (3. 6) is removed because of the orthonormality of the singular vectors. The solution will be over-smoothed when regularization parameter is chosen too large, which gives a large residual. If regularization parameter is chosen too small, the filter factors allow the noise to be amplified to yield a large solution norm. The drawback of the L-curve is that it may not have a clear L-shape, which means that an inappropriate parameter choice could be made. In other instances there can be two or no corners, which can also make the method fail.



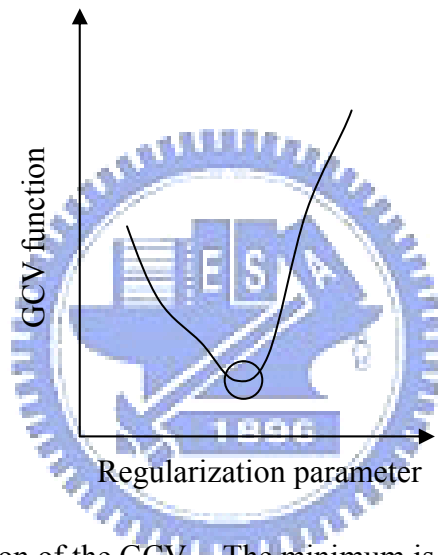


Fig. 57 Illustration of the GCV. The minimum is circled in the figure.

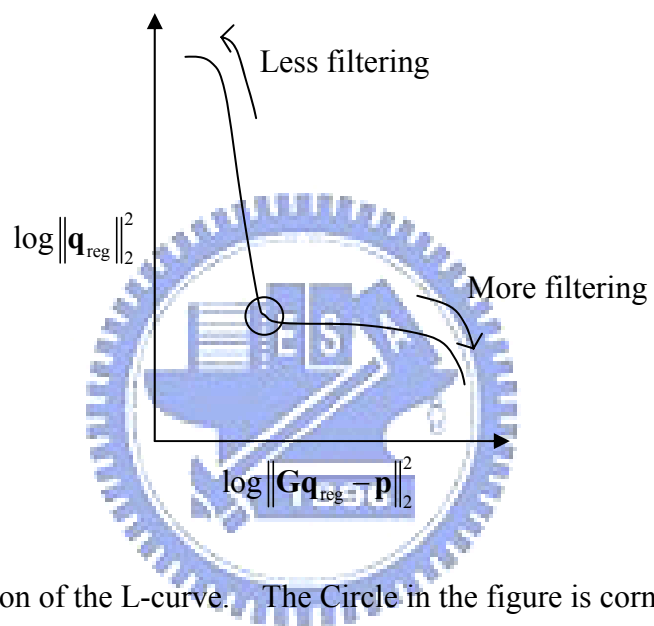


Fig. 58 Illustration of the L-curve. The Circle in the figure is corner of the curve.

6.2 Multi-channel fast filtering

The aforementioned multichannel inverse filtering is carried out entirely in the time-domain and is thus computationally intensive. To ease this problem, a technique based on the ERA is exploited to simplify the NESI processing. This method establishes the following minimal state-space realization for linear systems:

$$\mathbf{x}(n+1) = \mathbf{A}_e \mathbf{x}(n) + \mathbf{B}_e \mathbf{u}(n) \quad (6.16)$$

$$\mathbf{y}(n) = \mathbf{C}_e \mathbf{x}(n) + \mathbf{D}_e \mathbf{u}(n), \quad (6.17)$$

where n is the discrete-time index, $\mathbf{x}(n)$ is the state vector, $\mathbf{u}(n)$ is the $M \times 1$ input vector, $\mathbf{y}(n)$ is the $J \times 1$ output vector, and \mathbf{A}_e , \mathbf{B}_e , \mathbf{C}_e and \mathbf{D}_e are constant matrices.

The ERA starts with the impulse response matrices of the inverse filters:

$$\mathbf{C}(n) = \begin{bmatrix} c_{11}(n) & c_{12}(n) & \cdots & c_{1M}(n) \\ c_{21}(n) & c_{22}(n) & \cdots & c_{2M}(n) \\ \vdots & \vdots & \ddots & \vdots \\ c_{J1}(n) & c_{J2}(n) & \cdots & c_{JM}(n) \end{bmatrix}, \quad n = 0, 1, \dots, N, \quad (6.18)$$

where n is the time index and N is the length of impulse response. Assemble these impulse response matrices into a $J_S \times M_S$ Hankel matrix:

$$\mathbf{H}(n-1) = \begin{bmatrix} \mathbf{C}(n) & \mathbf{C}(n+1) & \cdots & \mathbf{C}(n+s-1) \\ \mathbf{C}(n+1) & \mathbf{C}(n+2) & \cdots & \mathbf{C}(n+s) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}(n+s-1) & \mathbf{C}(n+s) & \cdots & \mathbf{C}(n+2s-1) \end{bmatrix}, \quad (6.19)$$

where the s is an integer that determines the size of the matrix. Usually, s is taken to be $N/2$. Factor the Hankel matrix $\mathbf{H}(0)$ using SVD

$$\mathbf{H}(0) = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^H, \quad (6.20)$$

where \mathbf{U} and \mathbf{V} are unitary matrices and $\mathbf{\Sigma}$ is a diagonal matrix with singular values in its main diagonal. Great reduction is possible by observing the singular value plot. A typical example of singular value plot is shown in Fig. 59. The singular valves after $v = 100$ is very small and can be replaced by zeros. Hence, the matrices \mathbf{U} , \mathbf{V} and $\mathbf{\Sigma}$ are in effect truncated. Based on the above SVD result, the minimal

realization of \mathbf{A}_e , \mathbf{B}_e , \mathbf{C}_e and \mathbf{D}_e then follows:

$$\begin{aligned} \mathbf{A}_e &= \boldsymbol{\Sigma}^{-1/2} \mathbf{U}^H \mathbf{H}(1) \mathbf{V} \boldsymbol{\Sigma}^{-1/2}, \quad \mathbf{B}_e = \boldsymbol{\Sigma}^{1/2} \mathbf{V}^H \mathbf{E}_M, \\ \mathbf{C}_e &= \mathbf{E}_J^H \mathbf{U} \boldsymbol{\Sigma}^{1/2} \quad \text{and} \quad \mathbf{D}_e = \mathbf{C}(0), \end{aligned} \quad (6.21)$$

where $\mathbf{E}_M^H = [\mathbf{I}_M \quad \mathbf{0}_M \quad \cdots \quad \mathbf{0}_M]$ and $\mathbf{E}_J^H = [\mathbf{I}_J \quad \mathbf{0}_J \quad \cdots \quad \mathbf{0}_J]$, and the \mathbf{I} is an identity matrix and the $\mathbf{0}$ is a null matrix.

Instead of direct convolution (DC) in the time-domain, the multichannel filtering is efficiently carried out using the minimally realized state-space equation. This can give rise to considerable computational saving for the NESI processing owing to the fact that $\nu \ll Ms$ in general. For example, if $M = J = 30$ and $N = 227$, $M \times J \times 227 = 204300$ multiplications are required using DC, whereas only $\nu^2 + \nu M + J\nu + JM = 22500$ multiplications are required using the ERA with $\nu = 120$. Almost one order of reduction has been attained.



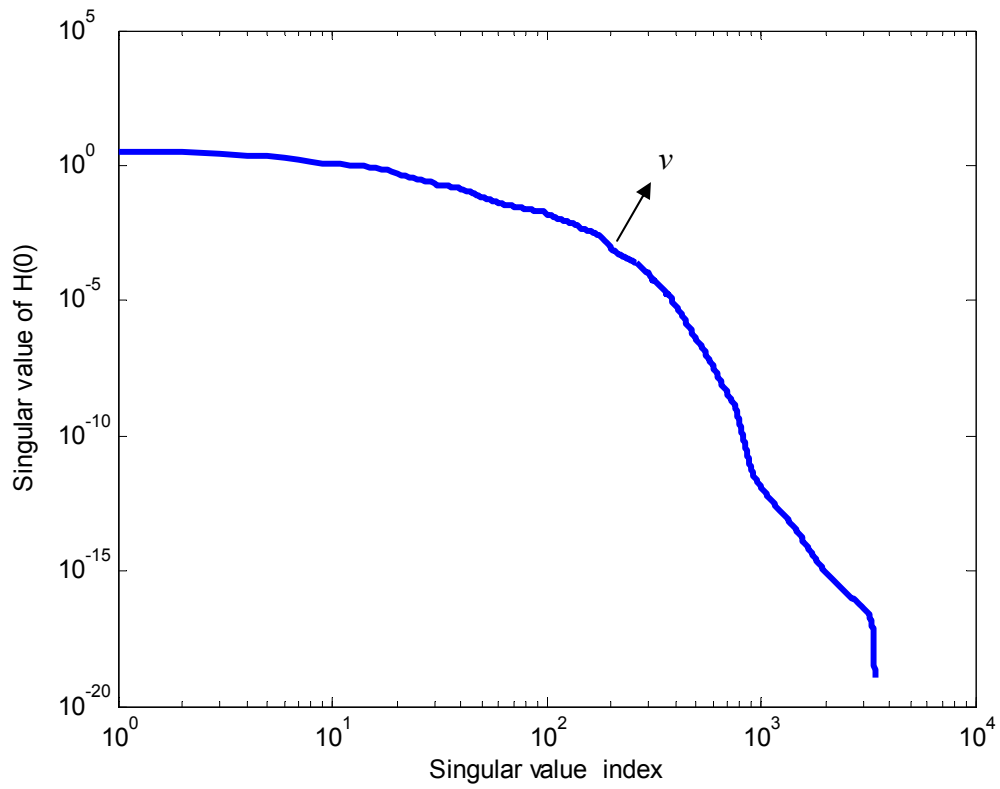


Fig. 59 Singular values of the Hankel matrix $\mathbf{H}(0)$. The singular values above the order v are negligibly small.

6.2.1 The time-domain processing

Apparently, execution of the vast number of inverse filters in NESI is very computationally demanding if all processing is carried out in the time-domain. This is becoming even more problematic in practical applications. To overcome this problem, the ERA is employed in this thesis. The ERA was described in the chapter

6.2. The Operations Per Sample (OPS) for ERA is

$$\text{OPS(ERA)} = v^2 + vM + Jv + JM . \quad (6.22)$$

In this thesis, the ERA is compared to widely used methods, the DC. Direct linear convolution denoted as DC is also used as a benchmark method. The OPS of the DC method is given by

$$\text{OPS(DC)} = M \times J \times N_f , \quad (6.23)$$

where N_f is the length of FIR filter.



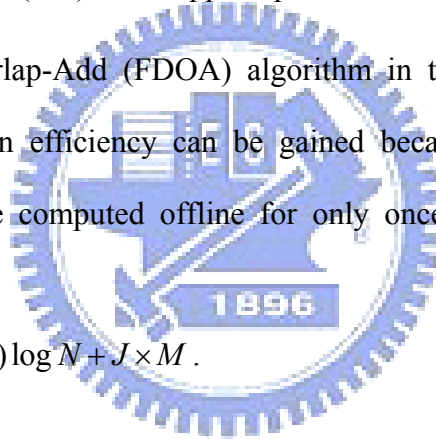
6.2.2 The frequency-domain processing

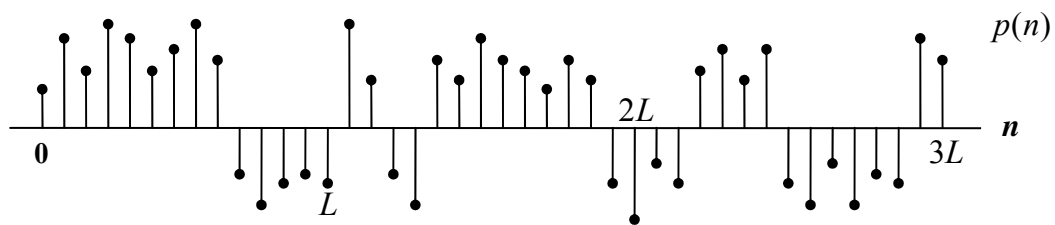
The NESI algorithm can also be efficiently implemented in the frequency-domain. First, partition the time-domain microphone pressure data $\mathbf{p}^h(n)$ into non-overlapping frames and zero-pad the frames into $\mathbf{p}_\Lambda^h(n)$, where $\Lambda = 1, 2, \dots, \hat{\Lambda}$ is the frame index and $\hat{\Lambda}$ is the number of frames, as shown in Fig. 60 (a) and Fig. 60 (b). Transform each frame to the frequency-domain by using the FFT. Next, multiply the transformed pressure data with the frequency-domain inverse matrix $\mathbf{C}(\omega)$ that can be computed offline. Finally, calculate the time-domain source amplitudes $\hat{\mathbf{q}}_\Lambda^v(n)$ for each frame by using the inverse FFT and overlap and add the consecutive frames, as shown in Fig. 60 (c).

Overlap-and-add technique can be used if continuous processing is desired.

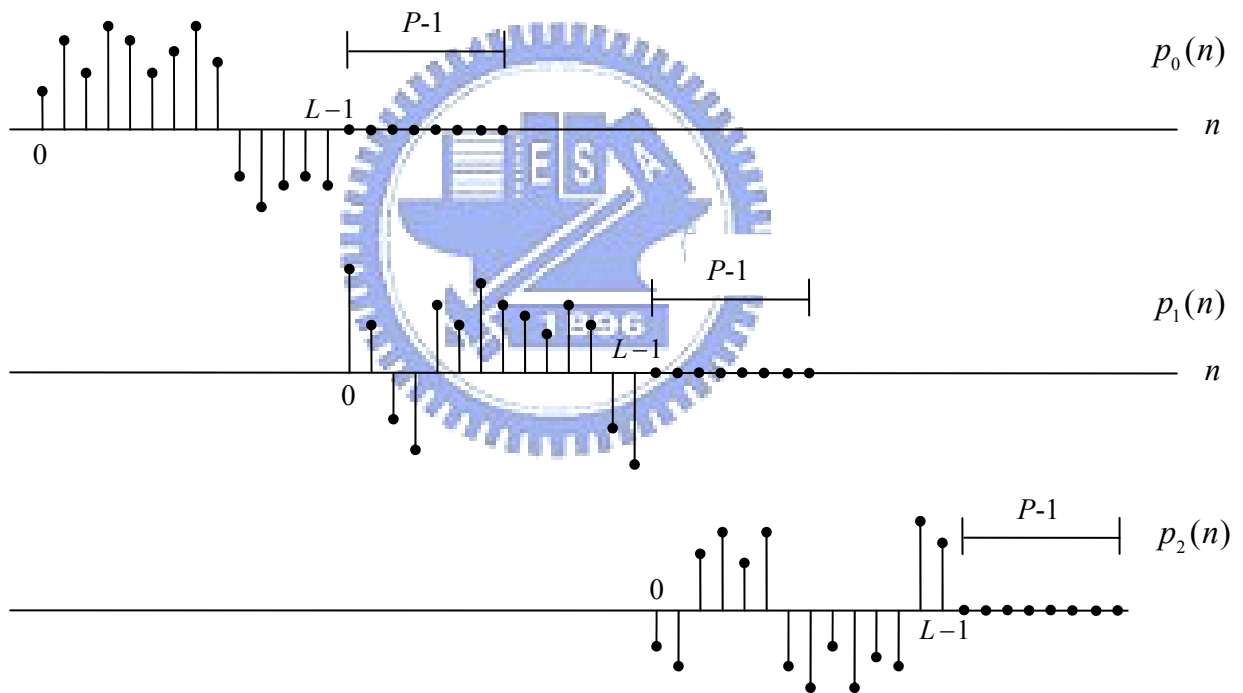
To illustrate how to choose parameters in the overlap-and-add block processing, consider the impulse response of the inverse filter matrix $\mathbf{C}(\omega)$ of length P . Assume that there are L samples in each frame $\mathbf{p}_\Lambda^h(n)$. Thus, the output of linear convolution $\hat{\mathbf{q}}_\Lambda^v(n) = \mathbf{C}(n) * \mathbf{p}_\Lambda^h(n)$ has the length $(L+P-1)$. The linear convolution can be efficiently implemented, with the aid of FFT, by calculating the product $\mathbf{C}(\omega)\mathbf{p}_\Lambda^h(\omega)$ in the frequency-domain, where $N \geq L+P-1$ point FFT must be used to avoid wraparound errors. To meet this length requirement, each frame must be padded with $(P-1)$ zeros. After inverse filtering, each frame of the source amplitude $\hat{\mathbf{q}}_\Lambda^v(n)$ is added with $(P-1)$ overlapped points. This is referred to as the Frequency-Domain-Overlap-Add (FDOA) algorithm in the following presentation. Tremendous computation efficiency can be gained because the frequency-domain inverse matrix needs be computed offline for only once. The OPS of FDOA is estimated to be

$$\text{OPS(OA)} = (M + J) \log N + J \times M. \quad (6. 24)$$





(a)



(b)

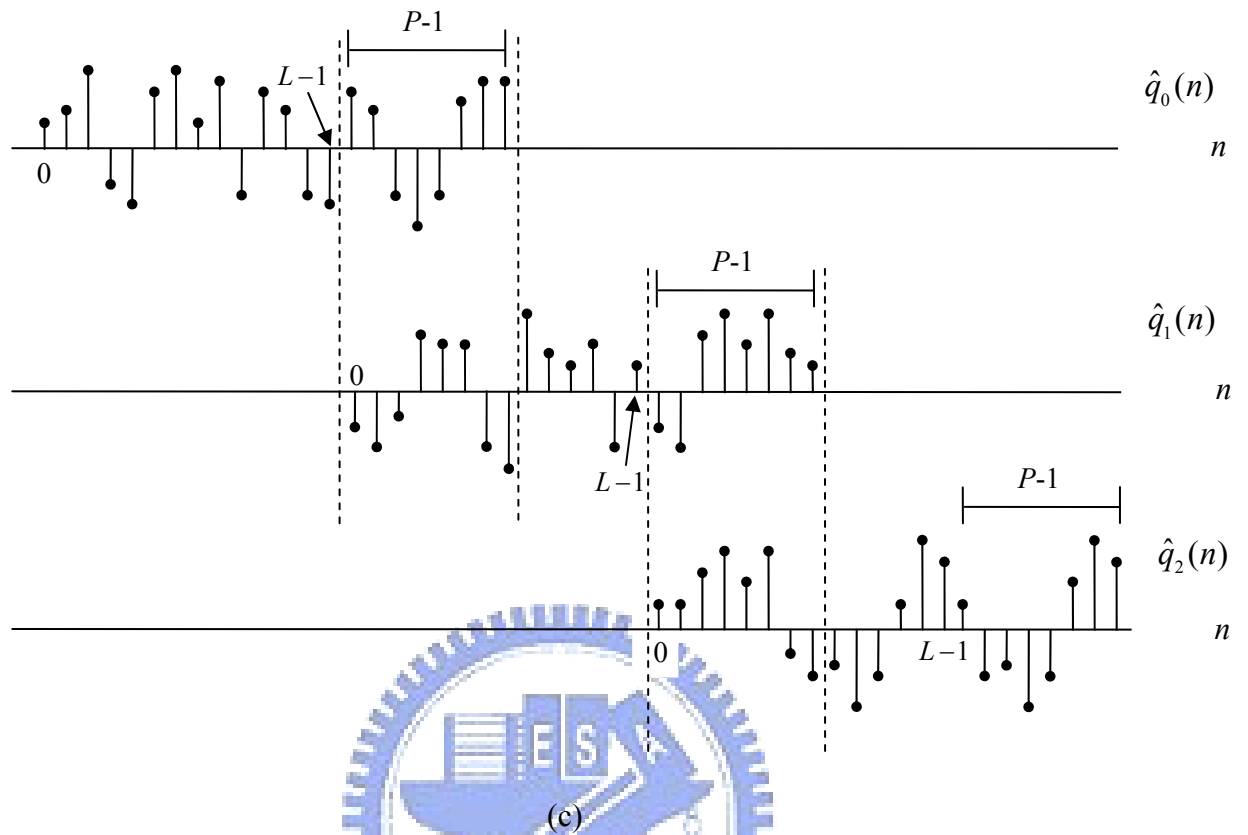


Fig. 60 Illustration of the Overlap and add method. (a) The pressure data $p(n)$, (b) Decomposition of $p(n)$ into non-overlapping sections of length L , (c) Result of convolving each section with the inverse filter.

6.2.3 Comparison of filtering approaches

Numerical simulation is conducted to validate the virtual microphone technique and multichannel filtering algorithms. A URA with $M = J = 4 \times 4$ is employed in this simulation, as depicted in Fig. 61. The spacing of the microphones (d) and the focus points (d_f) were both selected to be $0.1\text{m} = \lambda / 2$ for 1.7 kHz.

In the inverse filter design, Tikhonov regularization parameter was selected according to the L-Curve method. The OPS required by three different filtering methods (DC, ERA and FDOA) is compared for three different array configurations (16, 30 and 64 channels) in Table 9. The number of FFT frequency points $N_i = 512$. The numbers of microphones and focal points are assumed to be equal, i.e., $M = J$. The most computationally expensive DC method is used for benchmarking as 100% (in parenthesis) OPS requirement. It is obvious from the comparison that the ERA filtering displayed remarkable computation efficiency, e.g., 6.65% for a 64-channel array, as compared with the DC processing. The computation efficiency is considerably improved using the FDOA approach, especially for large number of microphone channels (only 5% of the benchmark DC method for a 64-channel array).

Table 9 Comparison of computational complexity in terms of OPS of three multichannels filtering methods for three array configurations. The block size of FFT $N_i = 512$. The numbers of microphones and focal points are assumed to be equal, i.e., $m = j$. The DC method is used for benchmarking (100% in parenthesis).

Domain	Method	4×4 URA	5×6 URA	8×8 URA
Time	DC	65,536 (100%)	230,400 (100%)	1,048,576 (100%)
	ERA	1,936 (2.95%)	16,900 (7.34%)	69,696 (6.65%)
Frequency	FDOA	544 (0.83%)	1440 (0.63%)	5248 (0.5%)



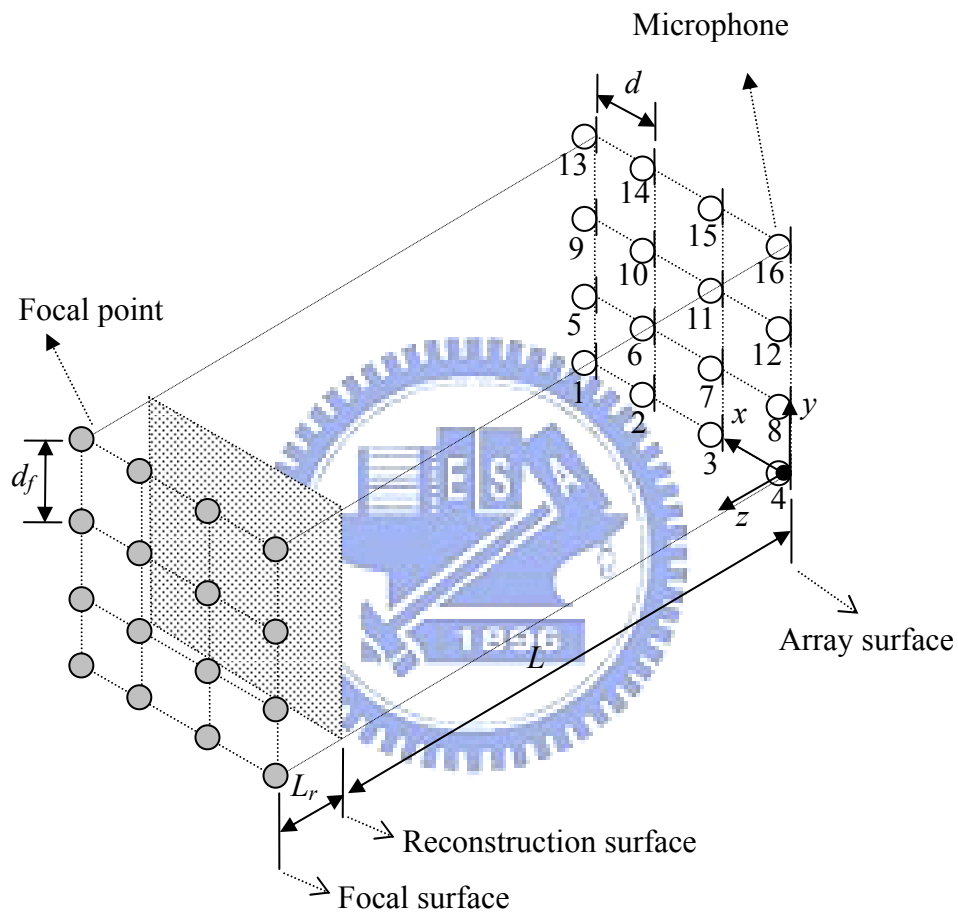


Fig. 61 The array settings for NESI using a 4x4 URA.

6.3 Post-processing

6.3.1 Acoustic variables: p , u , I , W

As the NESI algorithm, once the source amplitude $\hat{\mathbf{q}}^v$ is obtained, acoustical variables such as pressure and particle velocity can be reconstructed on the actual source surface. Due to singularity of virtual point sources, we need a non-zero RD to assure smooth reconstruction results. Therefore, the S^v is kept away from the S^s with a non-zero RD. It follows that the sound pressure on the reconstruction surface can be calculated in the time-domain using

$$p^r(\mathbf{z}, i) = \sum_{n=1}^N \frac{\hat{q}_n^v(n - \Delta_n)}{r_n}, \quad (6. 25)$$

where \mathbf{z} is the position vector of the Field point on the reconstruction surface, r_n is the distance between the n th virtual source and the Field point \mathbf{z} , and $\Delta_n = \text{int}(r_n / c)$ is the propagation time delay implemented using Lagrange interpolation. In addition to sound pressure, particle velocity, and sound intensity can be calculated by using the NESI technique. The last two acoustical variables are deemed more effective in the context of the NSI application. For simplicity, consider only one virtual point source on the focal point surface. Let \mathbf{z}_0 and \mathbf{z} be the position vectors of the virtual source and the field point on the reconstruction surface, respectively. The sound pressure at the field point produced by the point source of amplitude $\hat{q}(\omega)$ can be written as the following frequency-domain expression:

$$p^r(\mathbf{z}, \omega) = \hat{q}^v(\omega) \frac{e^{-jkr}}{r}, \quad (6. 26)$$

where $r = |\mathbf{z} - \mathbf{z}_0|$. From the Euler equation, the particle velocity at the normal direction to the reconstruction surface can be expressed as

$$\begin{aligned}
u^r(\mathbf{z}, \omega) &= \frac{-1}{j\rho\omega} \frac{\partial}{\partial n} p^r(\mathbf{z}, \omega) \\
&= \frac{-1}{j\rho_0\omega} \mathbf{n} \cdot \nabla_{\mathbf{x}} \left(\frac{\hat{q}^v(\omega) e^{-jk r}}{r} \right) \\
&= \frac{1}{j\rho_0\omega} (\mathbf{n} \cdot \mathbf{e}_r) \left(jk + \frac{1}{r} \right) p^r(\mathbf{z}, \omega),
\end{aligned} \tag{6.27}$$

where $\mathbf{e}_r = (\mathbf{x} - \mathbf{x}_0)/r$. Rewrite this equation in the Laplace domain

$$u^r(\mathbf{z}, t) = \frac{1}{\rho_0 c} (\mathbf{n} \cdot \mathbf{e}_r) \frac{s + (1/\tau)}{s} p^r(\mathbf{z}, t), \tag{6.28}$$

where $\tau = r/c$ being the time delay. Here the DC pole ($s=0$) behaves like an integrator, which could cause problems. To fix this, a highpass filter is introduced by modifying Eq. (6.28) into

$$\begin{aligned}
u^r(\mathbf{z}, t) &= \frac{1}{\rho_0 c} (\mathbf{n} \cdot \mathbf{e}_r) \frac{s + \frac{1}{\tau}}{s} \frac{s}{s + \frac{1}{\tau_0}} p^r(\mathbf{z}, t), \quad \tau_0 \gg 0 \\
&= \frac{1}{\rho_0 c} (\mathbf{n} \cdot \mathbf{e}_r) F(s) p^r(\mathbf{z}, t),
\end{aligned} \tag{6.29}$$

where

$$F(s) = \frac{s + (1/\tau)}{s + (1/\tau_0)}, \quad \tau_0 \gg 0. \tag{6.30}$$

It follows that the normal velocity u^r can be obtained by properly filtering the sound pressure with $F(s)$. To facilitate digital implementation, a discrete-time filter can be obtained by the Prewarped Bilinear Transform [17]:

$$F(z) = F(s) \Big|_{s=g \frac{z-1}{z+1}} = \frac{\tau_0}{\tau} \frac{(1 + g\tau)z + (1 - g\tau)}{(1 + g\tau_0)z + (1 - g\tau_0)}, \tag{6.31}$$

where

$$g = \frac{2\pi f_p}{\tan(\pi f_p / f_s)} \tag{6.32}$$

with f_p being the bandwidth of interest and f_s being the sampling rate. It turns out that $F(z)$ is always stable since its pole is inside the unit circle.

In summary, the normal velocity can be calculated for a point source using the following formula:

$$u^r(z) = \frac{1}{\rho_0 c} (\mathbf{n} \cdot \mathbf{e}_r) F(z) p^r(z). \quad (6.33)$$

The instantaneous normal intensity is simply

$$I_n^r(\mathbf{x}, n) = p^r(\mathbf{x}, n) u^r(\mathbf{x}, n). \quad (6.34)$$

Sound power can be calculated by integrating the normal active intensity over an area of interest S^r

$$W^r(n) = \int_{S^r} I_n^r(\mathbf{x}, n) dS^r(\mathbf{x}) \quad (6.35)$$

Similar procedure applies to a collection of point sources, where the sound field can be calculated using the principle of superposition.

6.3.2 Miscellaneous processing items

During the future trend, we hope to use the NESI for pass-by tests of vehicles. That is using source strength in nearfield by NESI to estimate farfield sound pressure of scooter or car from go-in line to go-out line at microphone R and L, as shown in Fig. 62. Side view of the pass-by test is shown in Fig. 63. The farfield sound pressure is calculated by

$$p_f(\mathbf{x}, t) = \sum_{i=1}^I \frac{\hat{q}_i(t - \Delta_i)}{r_i}, \quad (6.36)$$

where \mathbf{x} is the position vector of the field point on the microphone R or L, \hat{q}_i is the source strength of i th focal point, r_i is the distance between the i th virtual point source and the field point at \mathbf{x} , and $\Delta_i = r_i / c$ is the time delay. In Doppler effect, the frequency shift is calculated by

$$\Delta f = f \frac{v}{c}, \quad (6.37)$$

where f is frequency of source and v is speed of source. For example,

frequency of moving source is 1 kHz to estimate sound pressure of microphone, as shown in Fig. 64. Time-frequency diagram of microphone is shown in Fig. 65. The frequency shift is $\Delta f = f \cdot v / c = 1000 \times 28.5 / 343 = 83$ Hz. In addition, frequency shift has two conditions due to Doppler effect are following:

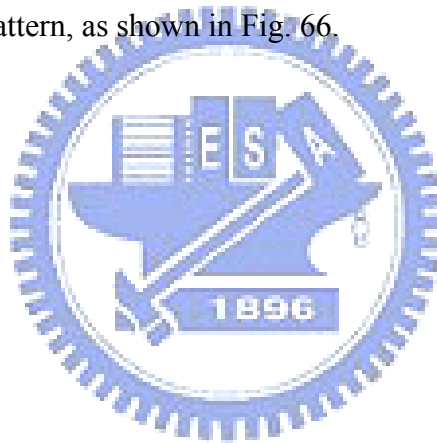
1. Fixed source, moving sensor

$$f' = f \left(\frac{c \pm v}{c} \right) \quad (6.38)$$

2. Fixed sensor, moving source (pass-by test)

$$f' = f \left(\frac{c}{c \mp v} \right) \quad (6.39)$$

Finally, sound pressure around scooter or car can be shown at appointed position by farfield polar radiation pattern, as shown in Fig. 66.



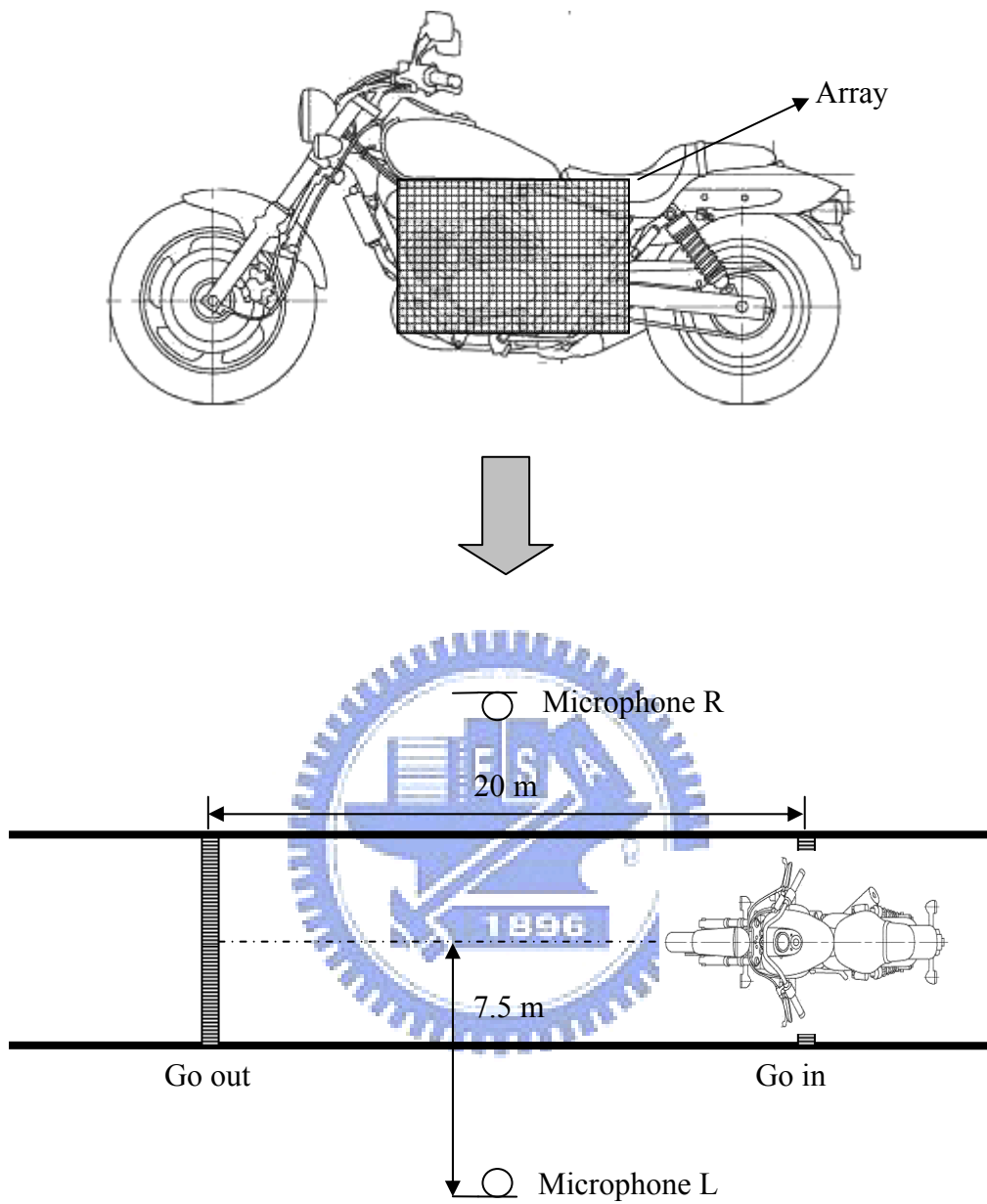


Fig. 62 Structure of pass-by test.

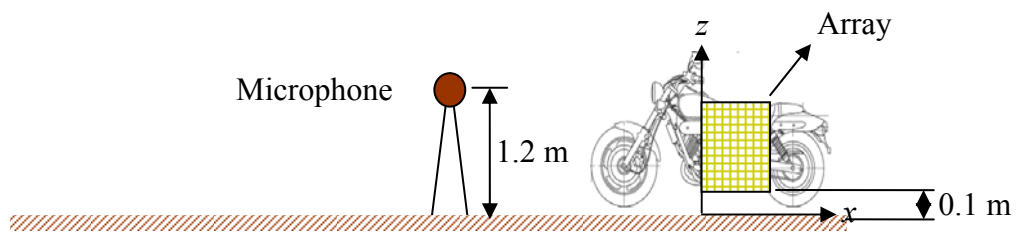
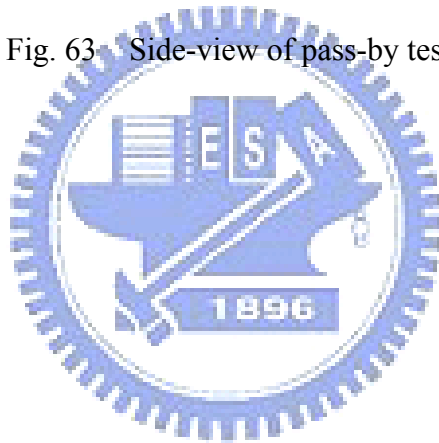


Fig. 63 Side-view of pass-by test.



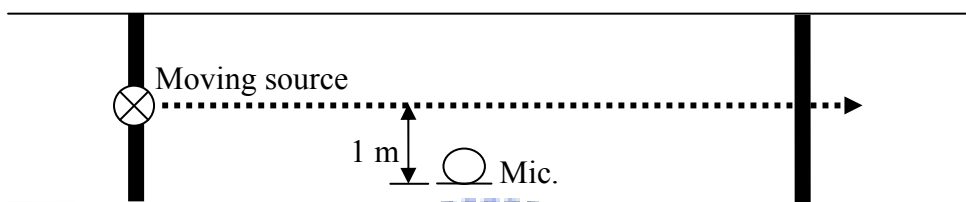
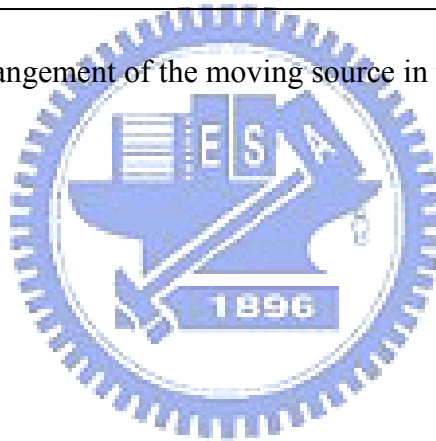


Fig. 64 Arrangement of the moving source in the pass-by test.



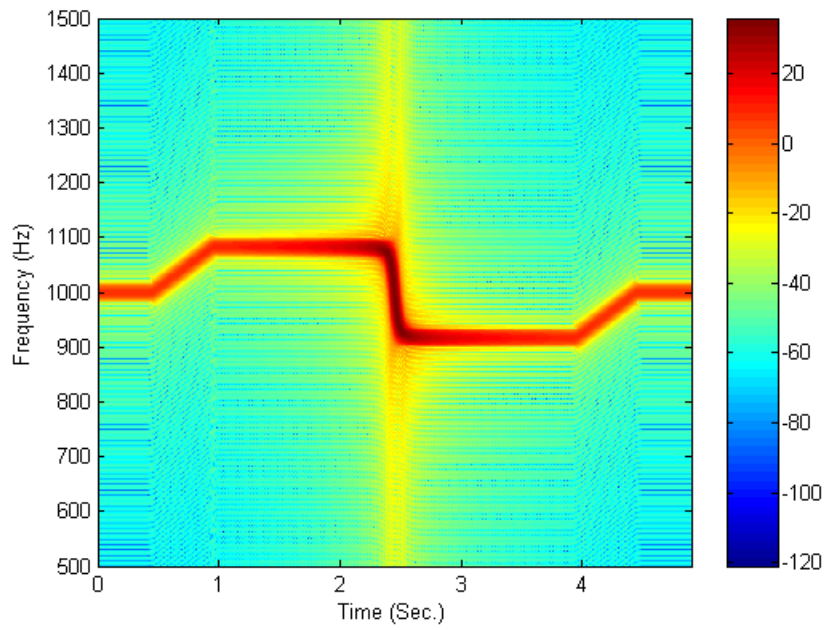


Fig. 65 Time-frequency diagram.

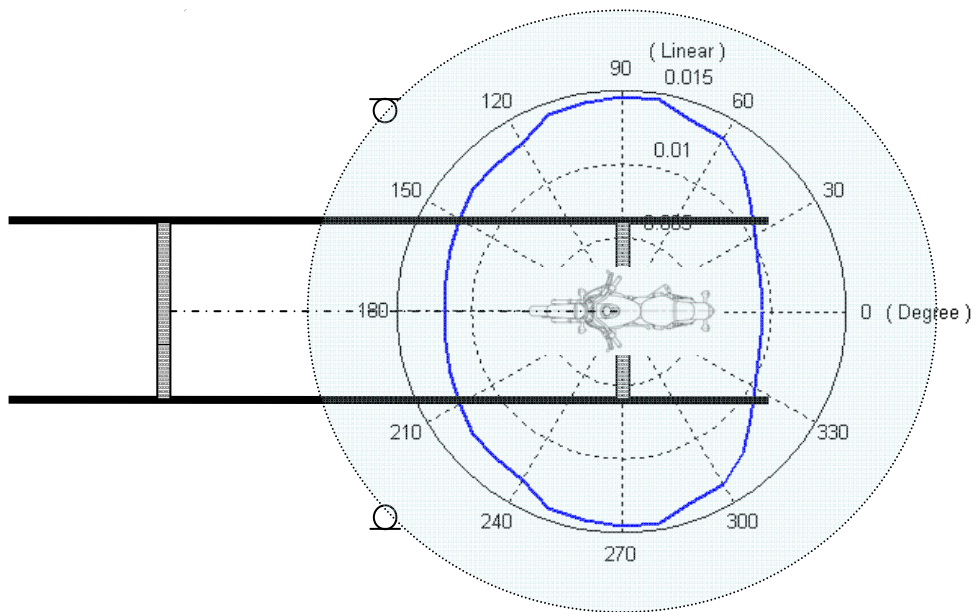


Fig. 66 Farfield polar radiation pattern.

6.4 Choice of distance of reconstruction and lattice spacing

Another criterion is needed to choose the DOR. Assume that the acoustic radiation problem can be formulated via ESM into the following matrix equation $\mathbf{G}\mathbf{q} = \mathbf{p}$, where \mathbf{p} and \mathbf{q} are the hologram data and source data, respectively, which are related by the propagation matrix \mathbf{G} . By assuming there is no uncertainty in the matrix \mathbf{G} , it can be shown that the perturbation term $\delta\mathbf{p}$ of the data vector such as measurement noise, numerical error, etc., and the perturbation term $\delta\mathbf{q}$ of the reconstructed data satisfy the following inequality [75]

$$\frac{\|\delta\mathbf{q}\|}{\|\mathbf{q}\|} \leq \text{cond}(\mathbf{G}) \frac{\|\delta\mathbf{p}\|}{\|\mathbf{p}\|}, \quad (6.40)$$

where $\text{cond}(\mathbf{G}) = \sigma_{\max} / \sigma_{\min}$ is the condition number of the matrix \mathbf{G} and $\|\cdot\|$ symbolizes vector 2-norm. Therefore, as an indicator of the ill-posedness inherent in the inverse filtering process, the condition number can also be regarded as a magnification factor of perturbations as well as loss of SNR after inverse filtering. For example, the SNR of data will be reduced by 60 dB of dynamic range after inverse filtering if $\text{cond}(\mathbf{G}) = 10^3$. It is well known that condition number of the propagation matrix increases with the DOR since the evanescent wave decays rapidly with the distance. It follows that the condition number can be used as a useful criterion for choosing the DOR. Thus, given a 60 dB tolerance of loss of SNR, a DOR that gives a condition number less than 1000 is generally deemed appropriate.

Another parameter is lattice spacing. The lattice spacing of microphones used is usually one-half the wavelength. This criterion is based on the Nyquist sampling theorem in the spatial domain. Violation of this criterion will likely cause spatial aliasing problem to the resulting image. Although this theoretical bound seems to be a somewhat aggressive rule, it has been widely accepted in the NAH community.

6.5 Virtual microphone technique: field interpolation and extrapolation

In practical implementation of the NESI technique, edge effects may occur when the physical extent of source is larger than the patch array aperture. In addition, the number of sensors may be too scarce to yield acceptable imaging resolution. To address these problems, a virtual microphone technique is employed with field interpolation (for improving resolution) and extrapolation (for reducing edge effect). This following example demonstrates this technique using 4×4 URA with microphone spacing d . This rather coarse array configuration is to be interpolated and extrapolated into 11×11 grid. The DOR is chosen to be $d/2$ so that the condition number of the propagation matrix \mathbf{G}_v was well below 1000, where \mathbf{G}_v is between the virtual microphone surface and reconstruction surface. In step \square of Fig. 67, the source amplitudes on the focal surface $\hat{\mathbf{q}}_{J \times 1}$ estimated by NESI are used to calculate sound pressure p_v for a finer grid on the microphone surface:

$$p_v(\mathbf{x}_v, n) = \sum_{j=1}^J \frac{\hat{q}_j(n - \Delta_{vj})}{r'_{vj}} \quad (6.41)$$

where \mathbf{x}_v is the position vector of the field point on the microphone surface, $r'_{vj} = |\mathbf{x}_v - \mathbf{y}_j|$, \mathbf{y}_j is the position vector of the j th point source on the focal surface, and $\Delta_{vj} = r'_{vj}/c$ is the time delay. The sound pressures regenerated using Eq. (6.41) for the interpolated and extrapolated actual/virtual sensor locations with a finer spacing can be assembled into the matrix form

$$\mathbf{p}_v(n) = \mathbf{G}_v(n) \hat{\mathbf{q}}(n) \quad (6.42)$$

where \mathbf{G}_v is the propagation matrix between the focal surface and the microphone surface, $M_v = 11 \times 11 = 121$ is the number of microphone and $J = 4 \times 4 = 16$ is the

number of point sources on the focal surface. In the frequency-domain, the sound pressure is calculated by

$$\underset{M_v \times 1}{\mathbf{p}_v(\omega)} = \underset{M_v \times J}{\mathbf{G}_v(\omega)} \underset{J \times 1}{\hat{\mathbf{q}}(\omega)} = \underset{M_v \times J}{\mathbf{G}_v(\omega)} \underset{J \times M}{\mathbf{C}(\omega)} \underset{M \times 1}{\mathbf{p}(\omega)}. \quad (6.43)$$

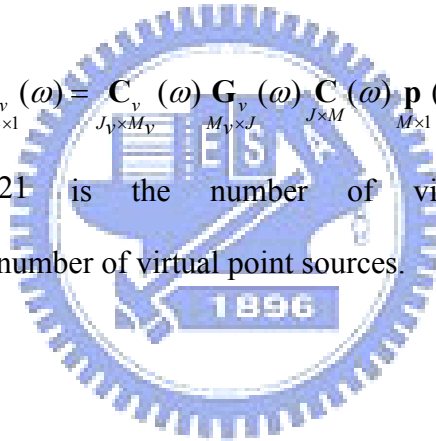
In Fig. 67, the interpolated and extrapolated microphones are indicated with the symbols “ \odot ” and “ \oplus ”, respectively. Next, choose a new point source distribution with finer spacing. The source amplitudes $\hat{\mathbf{q}}_v$ are estimated with the augmented inverse filters \mathbf{C}_v in the time-domain and the frequency-domain, as shown in step \square of Fig. 67:

$$\underset{J_v \times 1}{\hat{\mathbf{q}}_v(n)} = \underset{J_v \times M_v}{\mathbf{C}_v(n)} * \underset{M_v \times 1}{\mathbf{p}_v(n)} \quad (6.44)$$

and

$$\underset{J_v \times 1}{\hat{\mathbf{q}}_v(\omega)} = \underset{J_v \times M_v}{\mathbf{C}_v(\omega)} \underset{M_v \times 1}{\mathbf{p}_v(\omega)} = \underset{J_v \times M_v}{\mathbf{C}_v(\omega)} \underset{M_v \times J}{\mathbf{G}_v(\omega)} \underset{J \times M}{\mathbf{C}(\omega)} \underset{M \times 1}{\mathbf{p}(\omega)}, \quad (6.45)$$

where $M_v = 11 \times 11 = 121$ is the number of virtual microphones and $J_v = 11 \times 11 = 121$ is the number of virtual point sources.



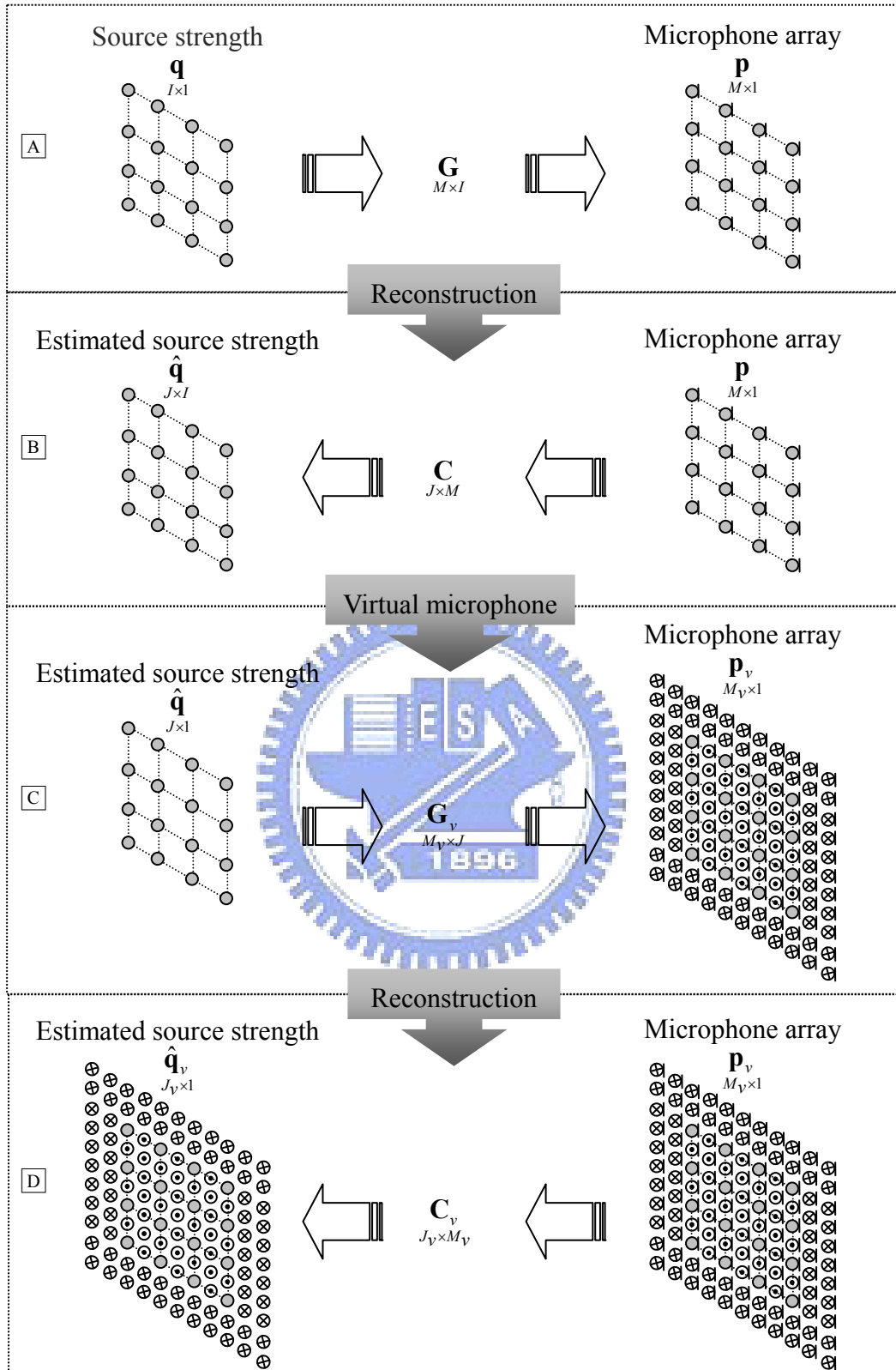


Fig. 67 The idea of the NESI with virtual microphone technique. The symbol “ \odot ” indicates an interpolated microphone position. The symbol “ \oplus ” indicates an extrapolated microphone position. **A** The pressure data picked up by

the microphones, **B** Reconstructed source strength at the focal points, **C**
The pressure data interpolated at the virtual microphones, **D**
Reconstructed source strength at the virtual focal points.



6.6 Choice of retreat distance (RD)

As a basic form of the ESM applied to NAH problems, discrete monopoles are utilized to represent the sound field of interest. When setting up the virtual source distribution, it is vital to maintain a RD between the virtual sources and the actual source surface such that reconstruction would not suffer from singularity problems. However, one cannot increase the distance without bound because of the ill-posedness inherent in the reconstruction process with large distance. In prior research, 1-2 times the lattice spacing, or the inter-element distance of microphones, is generally recommended as RD in the use of the ESM-based NAH. While this rule has shown to yield good results in many cases, the optimal choice is a complicated issue that depends on frequency, geometry of the physical source, content of evanescent waves, distribution of sensors and virtual sources, etc. How to reach the best compromise between the reconstruction errors induced by the point source singularity and the reconstruction ill-posedness is an interesting problem in its own right. This section revisits this issue, with the aid of an optimization algorithm based on the GSS.

6.6.1 Integral approximation error vs. reconstruction ill-posedness

Despite the versatility of the ESM, how to choose RD (also called retract distance or standoff distance) between the virtual sources and the reconstruction surface remains an issue that deserves attention in the formulation stage of the ESM. In the NESI formulation, the target source is simulated by a collection of point sources distributed on a surface. For that source configuration, one can simply choose virtual sources to be the target sources and the source field can be reconstructed. Assume RD is chosen very small. For continuous sources, however, it is impractical to distribute the discrete virtual sources on the source surface if the objective is to reconstruct the acoustic field such as velocity on the actual source

surface. Singularity problem will arise if the virtual sources are deployed on the actual source surface, which entails a RD between the virtual sources and the target source surface. This issue was investigated by Valdivia and Williams from a perspective of numerical integral approximation error [49]. They concluded that RD greater than one lattice is appropriate for acceptable integral approximation error. While increase of the distance has a spatial smoothing effect as a regularization mechanism, an overly large RD makes inverse reconstruction very ill-conditioned due to the loss of evanescent waves with the distance. There exists an optimal RD to best reconcile singularity and ill-posedness inherent in the inverse reconstruction process. In literature [41], [49] and [81], 1-2 times lattice spacing of RD is often recommended. For example, Sarkissian suggested one lattice spacing in his ESM simulation [41]. Also in a comparative study of the ESM and IBEM by Valdivia and Williams [49], RD greater than one lattice spacing is recommended, based on an integral approximation argument. In the same paper, it was also cautioned by the authors that the distance cannot be increased indefinitely because the inverse process will become ill-conditioned for large standoff distance.

While the 1-2 spacing criterion is a rule of thumb widely accepted in the NAH community that has shown to yield good results in many cases, the optimal choice is a complicated issue that may depend on frequency, nature of the physical source, content of evanescent waves, distribution of sensors and virtual sources, etc. How to best compromise between the reconstruction errors incurred by point source singularity and inversion ill-posedness is an interesting problem in its own right. The relative error of the velocity reconstructed on the actual source surface is used as the cost function in optimization.

6.6.2 Determination of RD: Golden Section Search (GSS)

The aim of the study is to find the optimal distance for the ESM-based NAH problem, with the aid of the GSS algorithm. We focus on only reconstruction of the surface velocity of the actual source, which proves the most effective for source identification. Define the relative velocity reconstruction error

$$E(\omega, x) = \frac{\|\mathbf{u}(\omega) - \mathbf{u}_r(\omega, x)\|_2^2}{\|\mathbf{u}(\omega)\|_2^2} \times 100\%, \quad (6.46)$$

where \mathbf{u} and \mathbf{u}_r are the desired and the reconstructed velocity vectors, respectively.

The cost function is defined as follows:

$$Q(x) = \frac{1}{I} \sum_{i=1}^I E(\omega_i, x), \quad (6.47)$$

where ω_i is the i th frequency and x is the design variable (RD in this case), is a measure of “average relative error of reconstructed velocity” summed in the frequency range 50 Hz~1.7 kHz. We aim at finding the RD that minimizes the cost function above. In the simulation, the microphone spacing d is selected to be 0.1m, complying with the $\lambda/2$ -rule for $f_{\max} = 1.7$ kHz. The frequency step is 100 Hz in the frequency range 50 Hz~1.7 kHz, hence $I = 17$. The variable tol_x is selected according to number of significant digits of the design variable. In this case, we require three significant digits for the RD, i.e., $tol_x = 10^{-3}$. The parameter $\varepsilon = 2.2 \times 10^{-15}$ is chosen according to the floating-point precision setting of the computer. The preceding parameter settings of the cost function, design variable and termination tolerance will be used in finding the optimal RD with the GSS algorithm.

Numerical simulations are conducted to demonstrate how to optimize RD by using the GSS algorithm. A baffled planer piston and a baffled spherical piston are employed as source examples. The first simulation is concerned with how to choose RD for virtual source configuration 1, while configuration 2 has no RD issue. The search proceeds in two steps. Step 1 evaluates the cost function in coarse but

uniform steps of the RD (0.1 times spacing). Inspection of the curve reveals that the minimum may exist in the interval $[1/20d, 2d]$. Step 2 searches for the optimal RD using the GSS in the potential interval found in Step 1. The second simulation utilizes the ESM-based NAH to calculate source surface velocity by using the optimal RD obtained in the first simulation. Source configuration 1 is employed in both simulations for the planar and spherical pistons, whereas source configuration 2 is applied to only the planar piston.

The first numerical simulation is concerned with the optimal RD for a planar baffled piston source. Suppose that the array comprises 8×8 microphones with spacing $d = 0.1\text{m}$ ($d = \lambda/2$ at $f = 1.7\text{ kHz}$), as shown in Fig. 49. The microphone surface is located at $z = 0.1\text{m}$. The number and spacing of microphones and virtual sources are of identical settings. A piston embedded at the plane $z = 0\text{m}$ is arranged in four scenarios as denoted by $S_{s,i}$, $i = 1-4$, shown in Fig. 68 (a)-Fig. 68 (d), respectively. The planar piston is simulated by discrete point sources distributed on the plane $z = 0\text{m}$. In the simulations, the desired velocity vector was specified to be 5 m/s . In scenario $S_{s,1}$, the actual sources are distributed in the same way as the microphone array element points. In scenario $S_{s,2}$, the actual sources are uniformly distributed at the array element points and their center midpoints. In scenario $S_{s,3}$, actual source distribution is similar to that in $S_{s,1}$, but the sources cover only part of the array aperture. Finally, in scenario $S_{s,4}$, the actual source distribution is similar to $S_{s,2}$, but the sources cover only part of the array aperture. With virtual source configuration 1, we conducted a coarse search in uniform steps for the 4 scenarios to find the potential interval in which the optimal RD may exist. Next, the GSS algorithm is employed to fine-search for the optimal RD in the potential interval. The result was shown in Fig. 69. The minimum points corresponding to the optimal

RDs are indicated in Fig. 69 with solid symbols. The optimal RDs calculated are 0.047m ($\doteq 0.5d$ or 0.25λ), 0.054m ($\doteq 0.5d$ or 0.25λ), 0.038m ($\doteq 0.4d$ or 0.20λ) and 0.049m ($\doteq 0.5d$ or 0.25λ) for scenarios $S_{s,i}$, $i = 1\sim 4$, respectively. Overall, the optimal distance ranges from 0.38 to 0.54 times spacing. It is interesting to note that the results are far less than the 1-2 times spacing recommended by the previous research. In addition, the optimal RDs for $S_{s,3}$ and $S_{s,4}$ are smaller than those of $S_{s,1}$ and $S_{s,2}$, respectively. The explanation of this phenomenon is as follows. The piston covers only part of the array aperture in scenarios $S_{s,3}$ and $S_{s,4}$. The velocity discontinuity at the edge of the piston will result in more evanescent waves, which entails smaller inversion distance (and hence smaller RD) for reconstructing surface velocity than the other two fully covered layouts. On the other hand, the RDs of $S_{s,2}$ and $S_{s,4}$ are greater than those of $S_{s,1}$ and $S_{s,3}$, respectively. A qualitative explanation is as follows. $S_{s,2}$ and $S_{s,4}$ represent more “uniform” approximation than the other two when simulating the continuous piston source. This requires larger propagation distance (and hence larger RD) to the piston surface than the other two scenarios to smooth out the reconstructed velocity field. The Fig. 70 compares the reconstructed and the desired velocity profiles for source configuration $S_{s,3}$ at $f = 850$ Hz. The relative error of the reconstructed velocity at $f = 850$ Hz in scenario $S_{s,3}$ is 21.13%. We see reasonable agreement of the profiles except near the edge of the piston where marked discrepancy occurs due to the velocity discontinuity.

While virtual source configuration 1 is applicable to arbitrary source geometries, virtual source configuration 2 is restricted only to planar sources. In configuration 2, RD is no longer an issue because the virtual sources are distributed on the actual continuous source surface when reconstructing the surface velocity. To better quantify the reconstructed results, the relative velocity reconstruction errors are calculated for four scenarios in Fig. 71. The error percentage calculated is 0% for

$S_{s,1}$, 20% for $S_{s,2}$, 26% for $S_{s,3}$, and 30% for $S_{s,4}$, in the frequency range 50 Hz~1.7 kHz. Although the reconstruction is nearly perfect in scenario $S_{s,1}$, 20-30% errors arise for scenarios $S_{s,2}$ and $S_{s,4}$ because of the limited number of microphones with respect to sources. Errors in the reconstructed velocity of scenario $S_{s,3}$ is due to the velocity discontinuity at the edge, which results in more evanescent wave components.

In the next numerical simulation, a spherical baffled piston source is adopted for further validation. For this non-planar source, only virtual source configuration 1 is applicable. Consider a vibrating piston set in a rigid sphere with axisymmetric velocity distribution (independent of azimuth angle ϕ) on its surface, as shown in Fig. 72. Assume $e^{j\omega t}$ time dependence. The piston is placed on the top of the rigid sphere of radius a , spanning an angle θ_0 with respect to the z axis. The surface velocity of the source is of the form

$$U(a, \theta) = \begin{cases} u_0 & , 0 \leq \theta \leq \theta_0 \\ 0 & , \theta_0 \leq \theta \leq \pi \end{cases} \quad (6.48)$$

The pressure field at a point (r, θ) of the piston can be calculated by [73]

$$p(r, \theta) = - \sum_{m=0}^{\infty} \frac{j\rho_0 c U_m}{h_m^*(ka)} P_m(\cos \theta) h_m^*(kr), \quad (6.49)$$

where

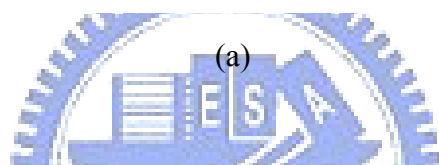
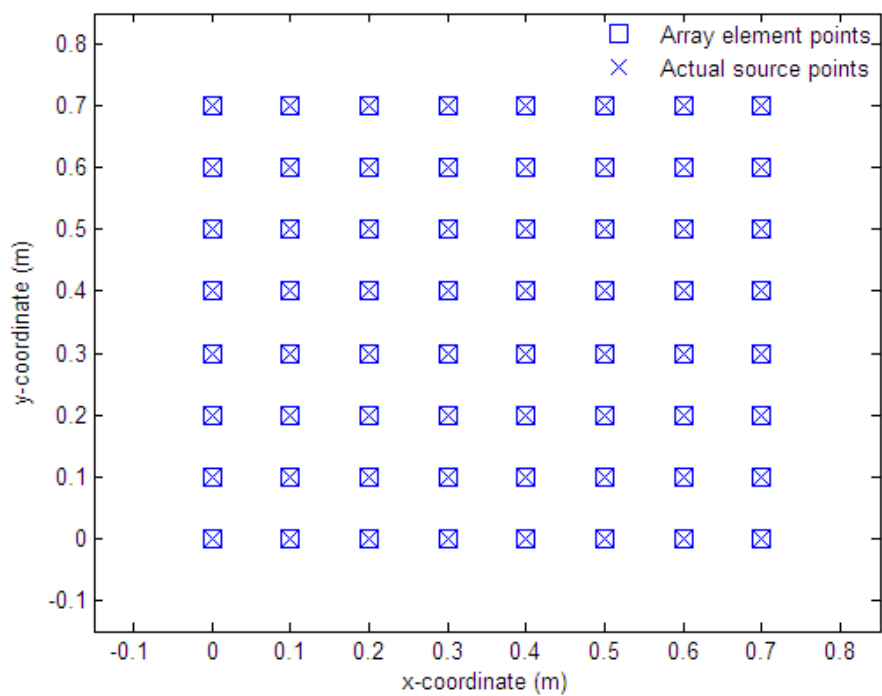
$$h_m'(\zeta) = \frac{1}{2m+1} [m h_{m-1}(\zeta) - (m+1) h_{m+1}(\zeta)], \quad (6.50)$$

$$U_m = \frac{u_0}{2} [P_{m-1}(\cos \theta_0) - P_{m+1}(\cos \theta_0)], \quad (6.51)$$

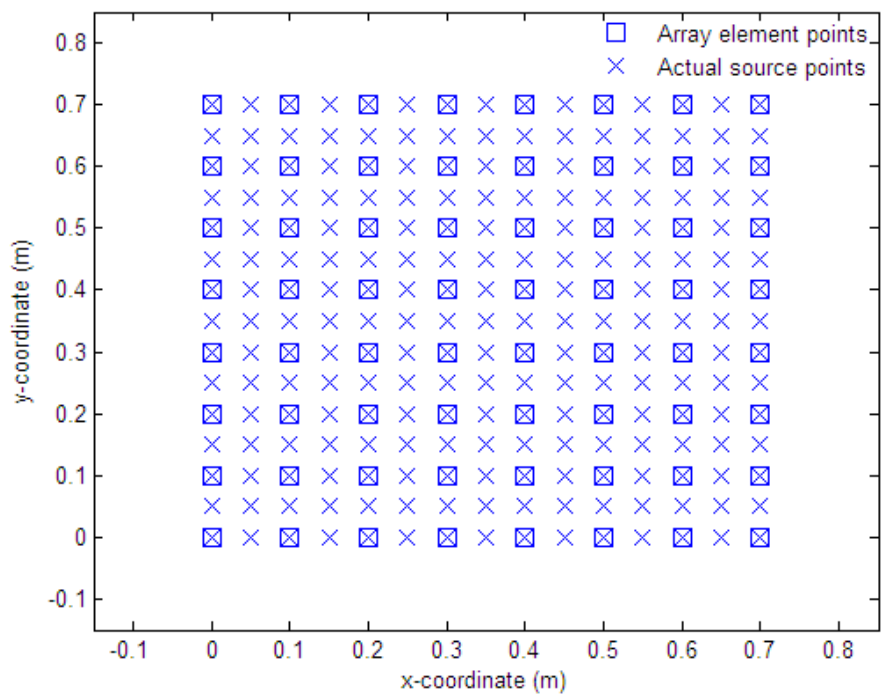
and “*” denotes complex conjugation. In Eq. (6.49), P_m is the Legendre polynomial of the first kind of order m , h_m is the spherical Hankel function of the first kind of order m , and h_m' is the derivative of the spherical Hankel function.

In the simulation, the source parameters are chosen as: $a = 0.2$ m, $r = 0.3$ m, $u_0 = 5$ m/s, $\rho_0 = 1.21$ Kg/m³ and $c = 343$ m/s. Three different spanning angles

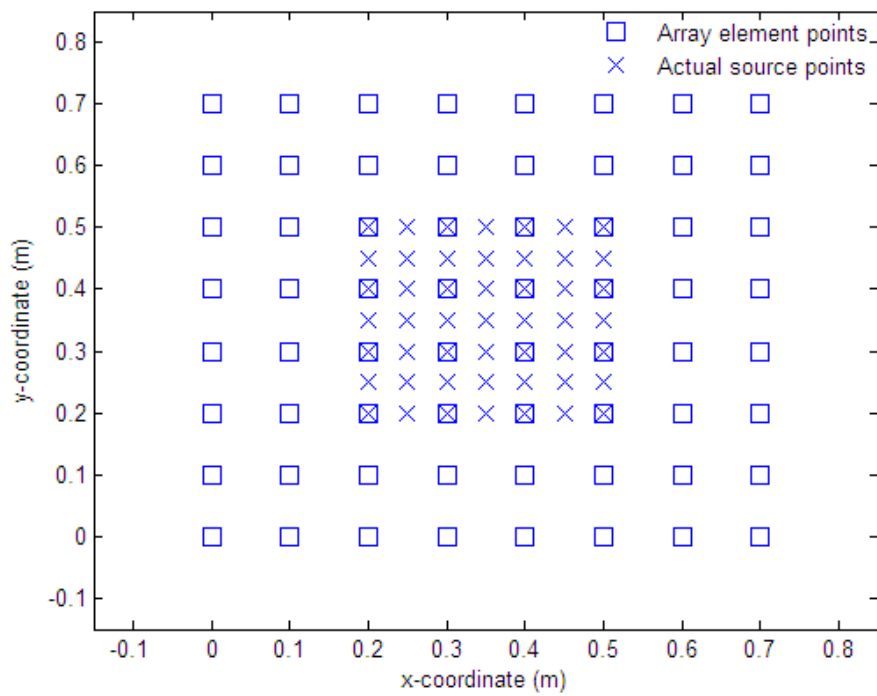
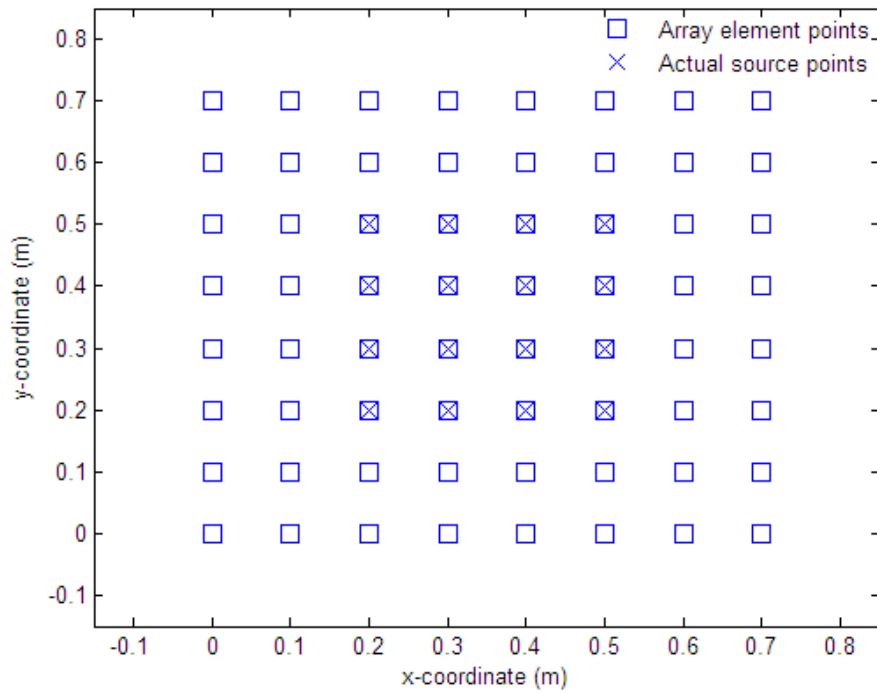
$\theta_0 = 30^\circ, 60^\circ$ and 180° are tested. Numbers of virtual sources and the sampling points for velocity reconstruction are both 58, as shown in Fig. 73. The average spacing of microphone is approximately 0.1 m (spacing $d = \lambda/2$ at $f = 1.7$ kHz). A preliminary search using coarse and uniform steps was first performed to find the potential interval in which the optimal RD may exist. The relative velocity errors for three spanning angles are shown in Fig. 74. Next, the GSS is employed to locate the optimal RD, as indicated by solid symbols in Fig. 74. The optimal RDs are 0.076m ($\doteq 0.8d$ or 0.40λ), 0.084m ($\doteq 0.8d$ or 0.40λ), and 0.165m ($\doteq 1.7d$ or 0.85λ) for $\theta_0 = 30^\circ, 60^\circ$ and 180° , respectively. Overall, the optimal distance can vary from 0.8 to 1.7 times spacing. As indicated by its not-so-smooth curve, the $\theta_0 = 180^\circ$ example seemed to be an “outlier” case that allows for RD much larger than other examples. Inspection of the relative error curves reveals that the error of the source with $\theta_0 = 180^\circ$ is significantly lower than those with $\theta_0 = 30^\circ$ and 60° . This can be explained by noting that the evanescent waves generated around the edge of the active part of the piston would result in larger reconstruction errors for small spanning angles. Matrix polar plots have been presented in Fig. 75 for the cases of $\theta_0 = 30^\circ, 60^\circ$ and 180° at $f = 1250$ Hz. Numbers of virtual sources and reconstruction points for velocity are 58 and 994, respectively. The relative error (%) of the reconstructed velocity versus frequency for the spanning angle $\theta_0 = 30^\circ, 60^\circ$ and 180° is shown in Fig. 76. The reconstructed results seemed to have captured the general trend of the surface velocity with this optimal RD.



(a)



(b)



(d)

Fig. 68 Scenarios of array element and actual source point distribution for simulating the planar baffled piston. (a) $S_{s,1}$, (b) $S_{s,2}$, (c) $S_{s,3}$, and (d) $S_{s,4}$.

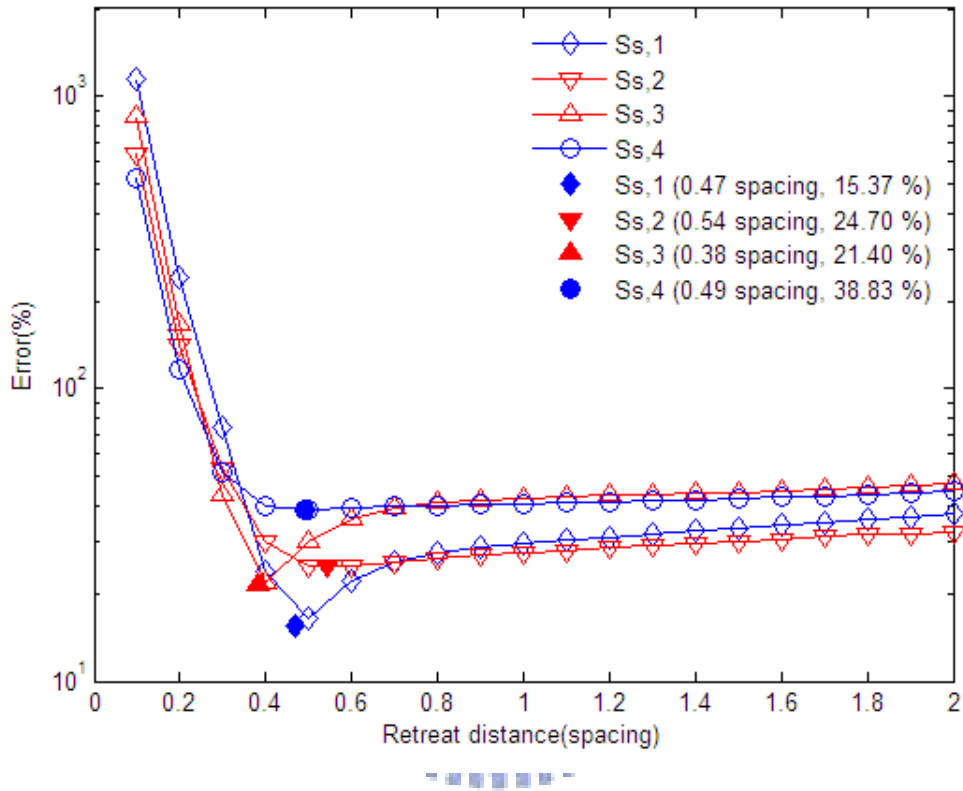


Fig. 69 Search for the optimal RD in a plane by virtual source configurations 2.

The symbol “ \diamond ” indicates the coarse search of $S_{s,1}$. The symbol “ \blacklozenge ” indicates the optimal point of $S_{s,1}$. The symbol “ ∇ ” indicates the coarse search of $S_{s,2}$. The symbol “ \blacktriangledown ” indicates the optimal point of $S_{s,2}$. The symbol “ \triangle ” indicates the coarse search of $S_{s,3}$. The symbol “ \blacktriangle ” indicates the optimal point of $S_{s,3}$. The symbol “ \bigcirc ” indicates the coarse search of $S_{s,4}$. The symbol “ \bullet ” indicates the optimal point of $S_{s,4}$.

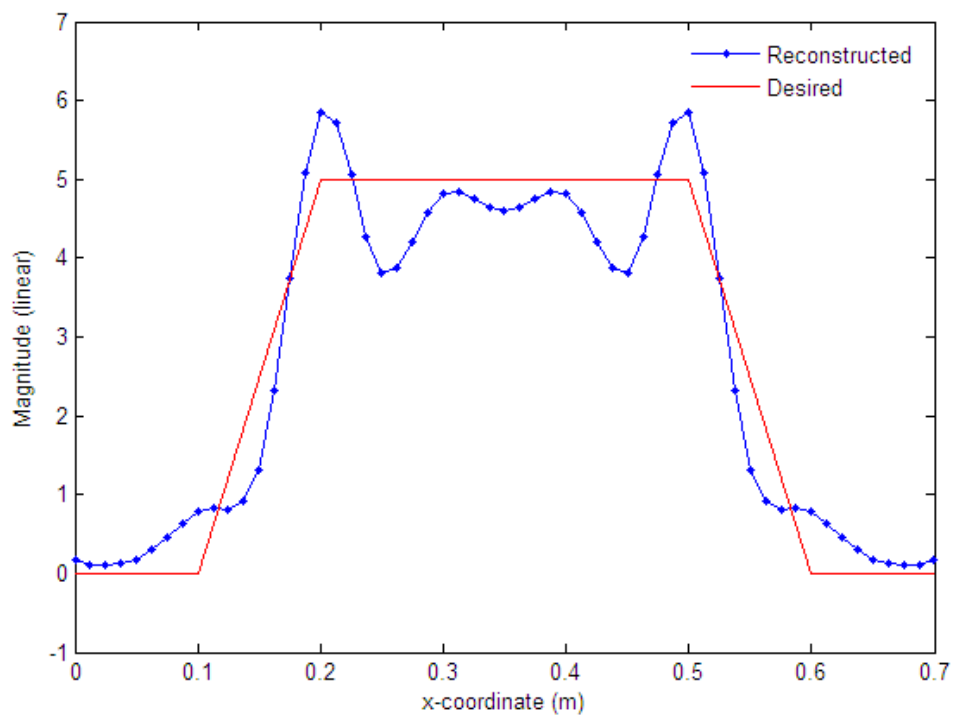


Fig. 70 The reconstructed and the target velocity profiles of planar source scenario $S_{s,3}$ at $f = 850$ Hz.

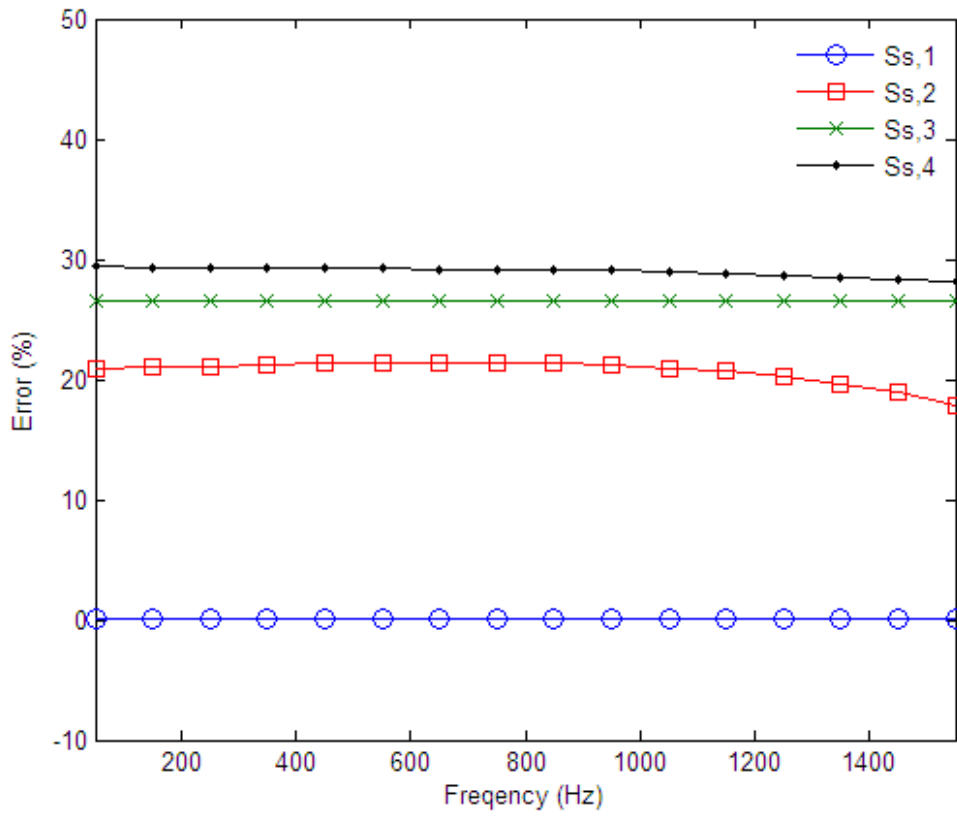


Fig. 71 Relative error (%) of velocity reconstructed for planar sources using virtual source configuration 2.

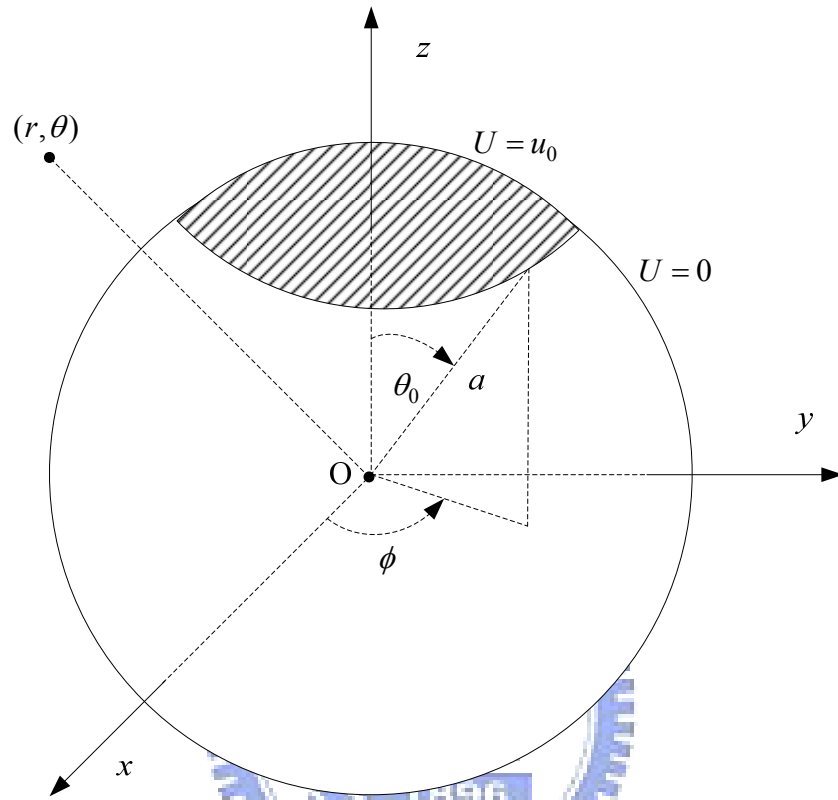


Fig. 72 Configuration of a vibrating rigid piston spanned with an angle θ_0 , set at the top of a rigid sphere of radius a .

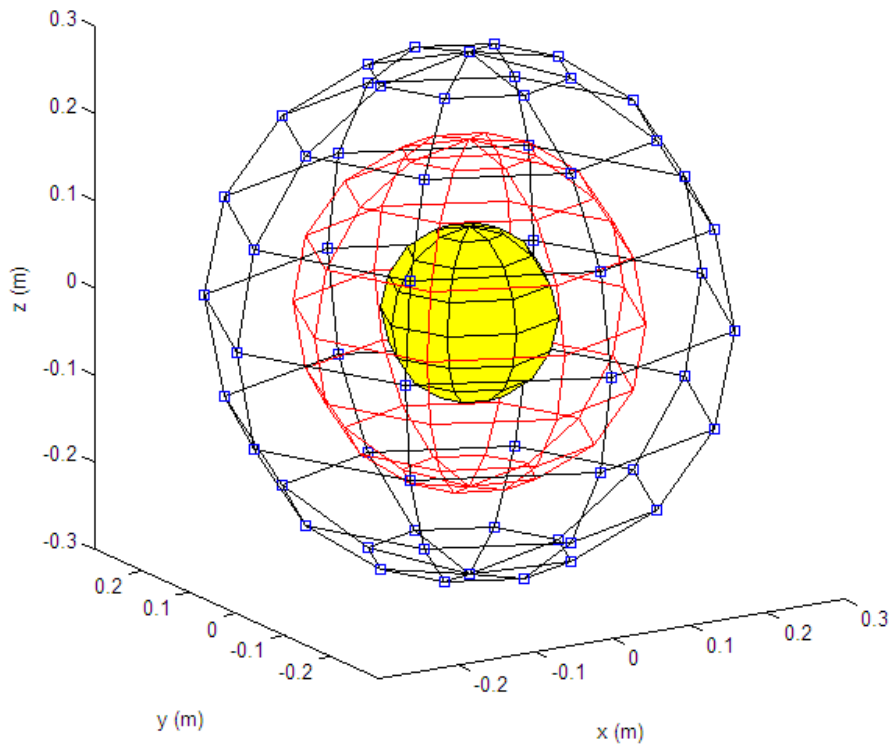


Fig. 73 Mesh configuration of the spherical baffled piston. The symbol“□” indicates the microphones. The red mesh indicates the actual source surface. The yellow mesh indicates the virtual source surface. These three meshes are concentric.

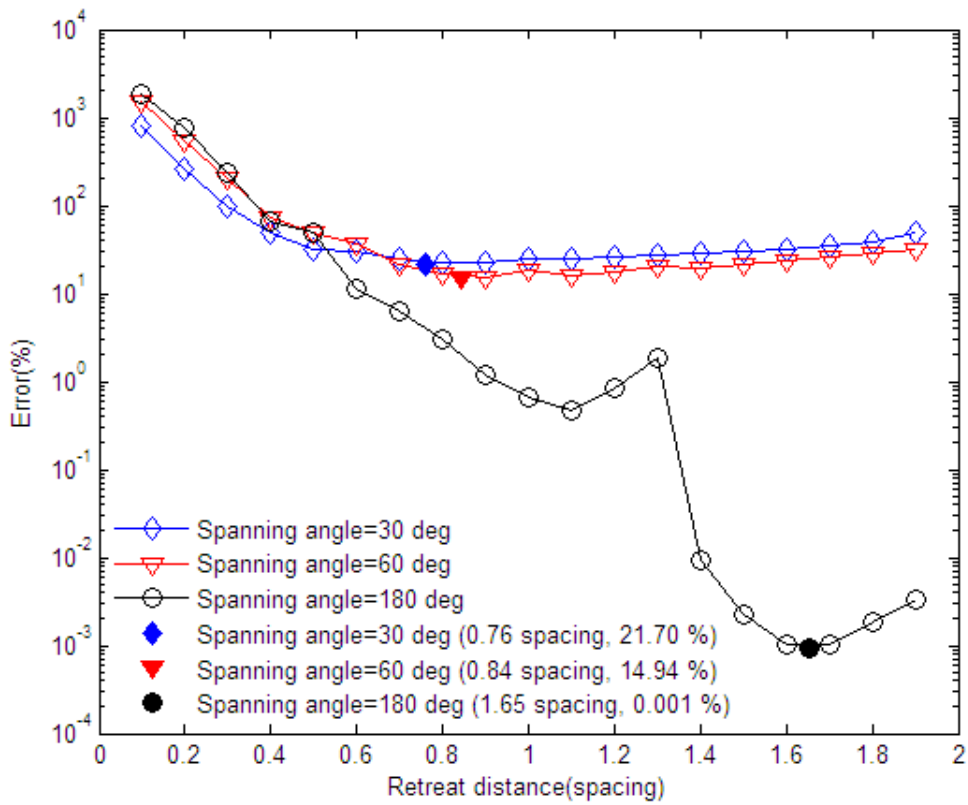
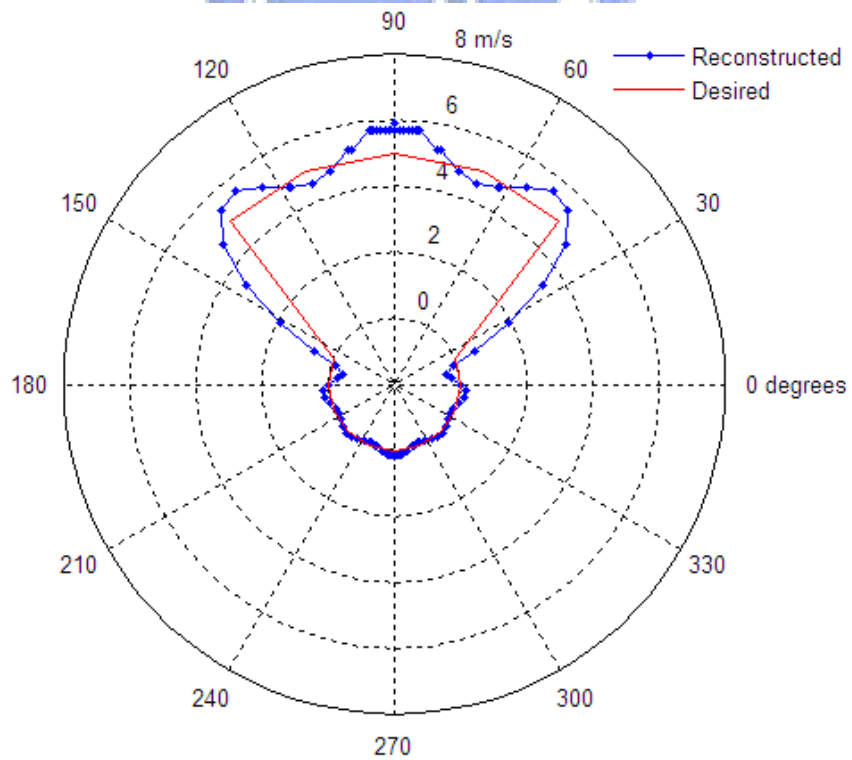
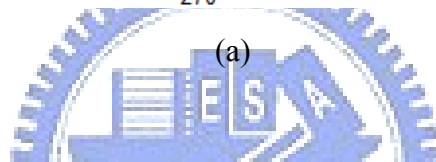
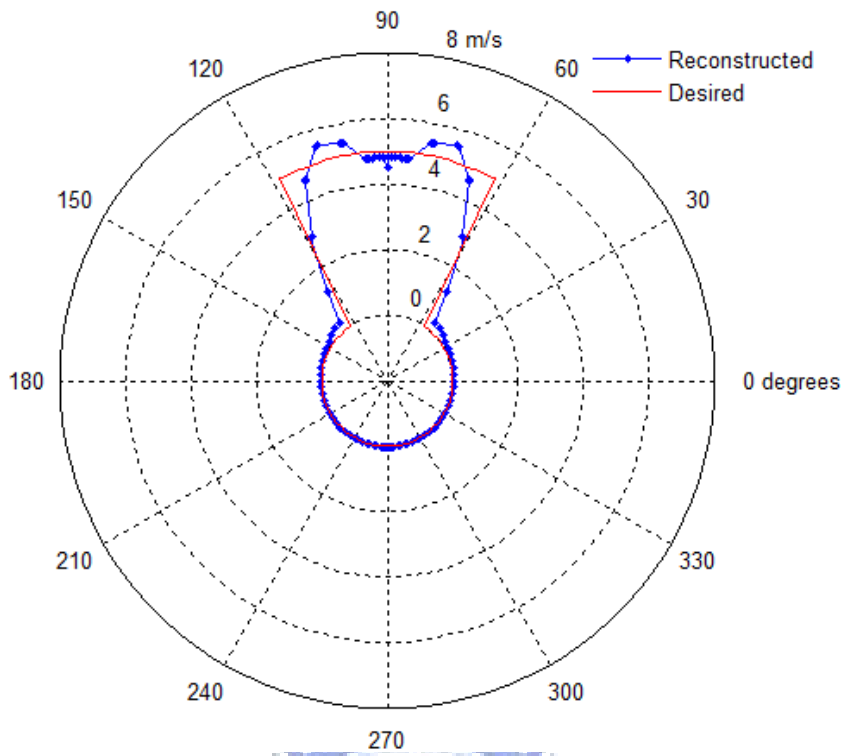
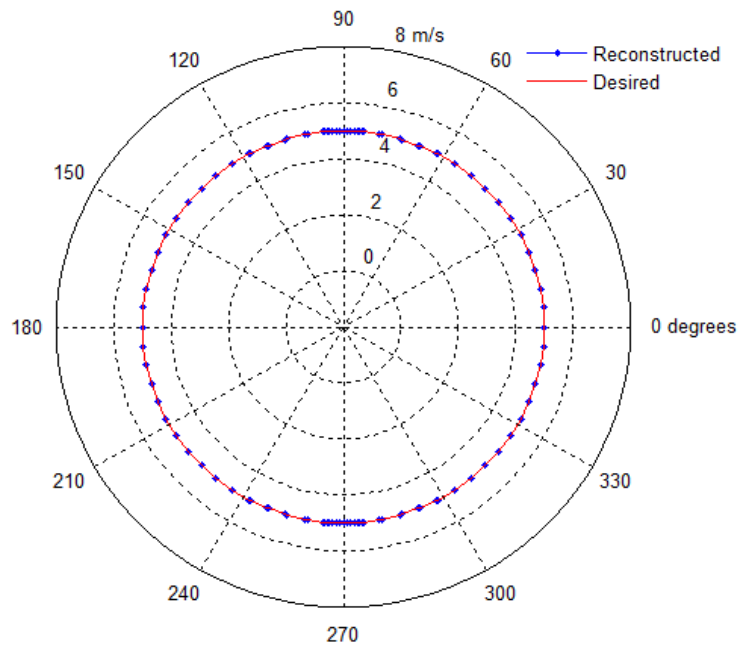


Fig. 74 Search for the optimal RD of the spherical baffled piston by virtual source configuration 1. The symbol “ \diamond ” indicates the coarse search of the case $\theta_0 = 30^\circ$. The symbol “ \blacklozenge ” indicates the optimal point. The symbol “ ∇ ” indicates the coarse search of the case $\theta_0 = 60^\circ$. The symbol “ \blacktriangledown ” indicates the optimal point. The symbol “ \ominus ” indicates the coarse search of the case $\theta_0 = 180^\circ$. The symbol “ \bullet ” indicates the optimal point.

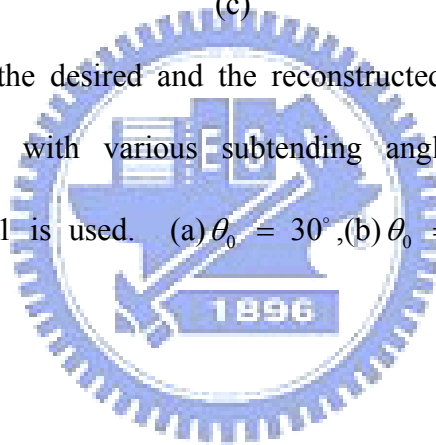


(b)



(c)

Fig. 75 Polar plot of the desired and the reconstructed velocity of the spherical baffled piston with various subtending angles. The virtual source configuration 1 is used. (a) $\theta_0 = 30^\circ$, (b) $\theta_0 = 60^\circ$, (c) $\theta_0 = 180^\circ$ at $f = 1250$ Hz.



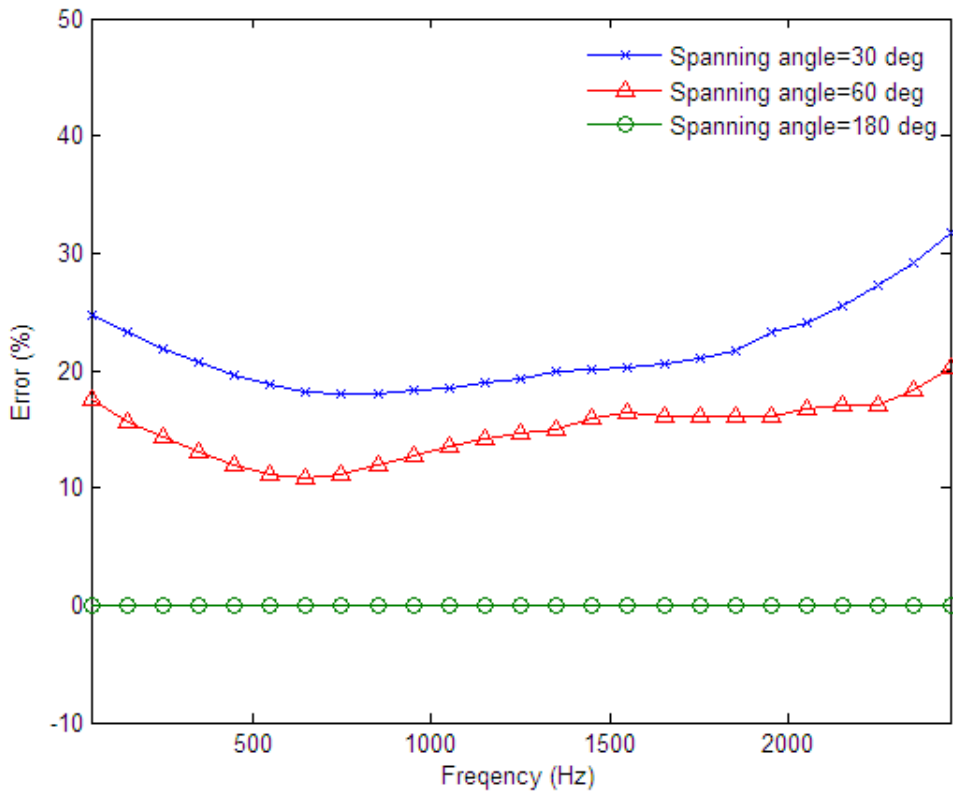


Fig. 76 Relative error (%) of surface velocity reconstructed for the spherical sources using virtual source configuration 1. Three spanning angle $\theta_0 = 30^\circ, 60^\circ$ and 180° are examined.

6.7 Optimization of sensor deployment: uniform vs. random array

In this section, the beam patterns and the cost functions are defined to facilitate the following array optimization formulation. Since our focus is the nearfield array, we only briefly review the farfield array next.

6.7.1 Optimal nearfield array: cost functions

For a nearfield array, due to the fact that the source is in the nearfield and the focal points are multiple in numbers, the definition of nearfield beam pattern is not as straightforward as in the farfield case. This thesis suggests a procedure to calculate the nearfield pressure field beam pattern by scanning on the focal surface using a test point source. The procedure is depicted in Fig. 77 and described as follows:

- (1) Design the inverse matrix \mathbf{C} for the given array configuration.
- (2) Position the test source at a grid point on the focal surface. Calculate the sound pressure vector \mathbf{p} received at the microphones.
- (3) Calculate the source strength vector at the focal points using the inverse matrix \mathbf{C} :

$$\mathbf{q} = \mathbf{C}\mathbf{p} \quad (6.52)$$

- (4) Calculate the pressure field \mathbf{p}_r^z by propagating the spherical waves emitting from the point source at the z th grid point on the focal surface to the reconstruction surface by

$$\mathbf{p}_r = \mathbf{G}_r \mathbf{q}, \quad (6.53)$$

where \mathbf{G}_r denotes the propagation matrix between the focal surface and the reconstruction surface.

- (5) Move the test source to next grid point on the focal surface and repeat steps (2)-(4). Superimpose the magnitude of the calculated pressure field for all Z

test source positions to get the following nearfield beam pattern vector

$$\mathbf{b} = \left[\sum_{z=1}^Z |p_r^z(1,1)| \quad \sum_{z=1}^Z |p_r^z(2,1)| \quad \cdots \quad \sum_{z=1}^Z |p_r^z(\Gamma,1)| \right]^T, \quad (6.54)$$

where Γ is the number of pixels on the reconstruction surface. Often, $Z > \Gamma$ is selected.

With the beam pattern obtained above, a cost function is defined for the nearfield array. Instead of the maximum main-lobe and side-lobe ratio as in the farfield array, the nearfield cost function focuses on the number and positions of the main-lobe peaks.

In the following, a nearfield cost function is devised to ensure that the resulting beam pattern has well defined main-lobe at all focal points. First, the magnitude of the beam pattern is normalized to unity. Second, define the interior of the circle centered at each focal point with the radius r_m as the scope of the main-lobe. The circle is used not only to define the main-lobe but also to check if a main-lobe is defocused. For the j th focal point, find the maximum of a main-lobe m_j that is greater than 0.707 (-3 dB). The cost function of the nearfield array is defined as

$$Q = \sum_{j=1}^J m_j. \quad (6.55)$$

The objective of the present array optimization is to find the deployment that maximizes the preceding cost function.

In addition, the cost function of the nearfield array is defined too from matrix conditioning. From the Fig. 78, let \mathbf{p} , \mathbf{v} and \mathbf{G} be the pressure vector, the source velocity vector and the propagation matrix, respectively. The measured sound pressures and the source velocities are related by

$$\mathbf{p} = \mathbf{G}\mathbf{v} + \mathbf{n}, \quad (6.56)$$

where \mathbf{n} is sensor noise. The unknown reconstructed velocity can be calculated

$$\hat{\mathbf{v}} = \mathbf{G}^+ \mathbf{p}, \quad (6.57)$$

where \mathbf{G}^+ is the pseudo-inverse matrix of \mathbf{G} . The reconstruction error of velocity is

$$\mathbf{e} = \hat{\mathbf{v}} - \mathbf{v} = \mathbf{G}^+ \mathbf{p} - \mathbf{v} = \mathbf{G}^+ (\mathbf{G}\mathbf{v} + \mathbf{n}) - \mathbf{v} = \mathbf{G}^+ \mathbf{n} - (\mathbf{I} - \mathbf{G}^+ \mathbf{G})\mathbf{v}. \quad (6.58)$$

Note that $(\mathbf{I} - \mathbf{G}^+ \mathbf{G})$ is the projection matrix onto the null space of \mathbf{G} . That is the reconstruction error due to $(\mathbf{I} - \mathbf{G}^+ \mathbf{G})\mathbf{v}$ would not contribute to the error in \mathbf{p} and $\mathbf{G}^+ \mathbf{n}$ is the term to be minimized. This can be done by using the matrix induced

2-norm of \mathbf{G}^+ :

$$\|\mathbf{G}^+\|_2^2 = \max_{\mathbf{n} \neq \mathbf{0}} \left\{ \frac{\|\mathbf{G}^+ \mathbf{n}\|_2^2}{\|\mathbf{n}\|_2^2} \right\} = \max_{\mathbf{n} \neq \mathbf{0}} \left\{ \frac{\mathbf{n}^H (\mathbf{G}^+)^H \mathbf{G}^+ \mathbf{n}}{\mathbf{n}^H \mathbf{n}} \right\} \quad (6.59)$$

That is Rayleigh's quotient. Thus, the maximum occurs at the minimum eigenvalue of $(\mathbf{G}^+)^H \mathbf{G}^+$, or equivalently, the maximum singular value of \mathbf{G}^+ . If the TSVD is used to obtain \mathbf{G}^+ , the maximum singular value of \mathbf{G}^+ will be the reciprocal of the minimum singular value of the 'truncated' \mathbf{G} .

$$\min_{\mathbf{G}} \|\mathbf{G}^+\|_2^2 = \min_{\mathbf{G}} \sigma_{\max}(\mathbf{G}^+) = \min_{\mathbf{G}} \frac{1}{\sigma_{\min}(\mathbf{G})} = \max_{\mathbf{G}} \sigma_{\min}(\mathbf{G}) \quad (6.60)$$

It follows that the optimal sensor deployment problem can be posed as a 'max-min' problem aimed at finding the deployment that maximizes the minimum singular value of the 'truncated' \mathbf{G} . Alternatively,

$$\min_{\mathbf{G}} \text{cond}(\mathbf{G}), \quad (6.61)$$

where $\text{cond}(\mathbf{G}) = \sigma_{\max}(\mathbf{G}) / \sigma_{\min}(\mathbf{G})$ is the condition number of \mathbf{G} . Therefore, another method to define the cost function is shown in Eqs. (6.60) or (6.61).

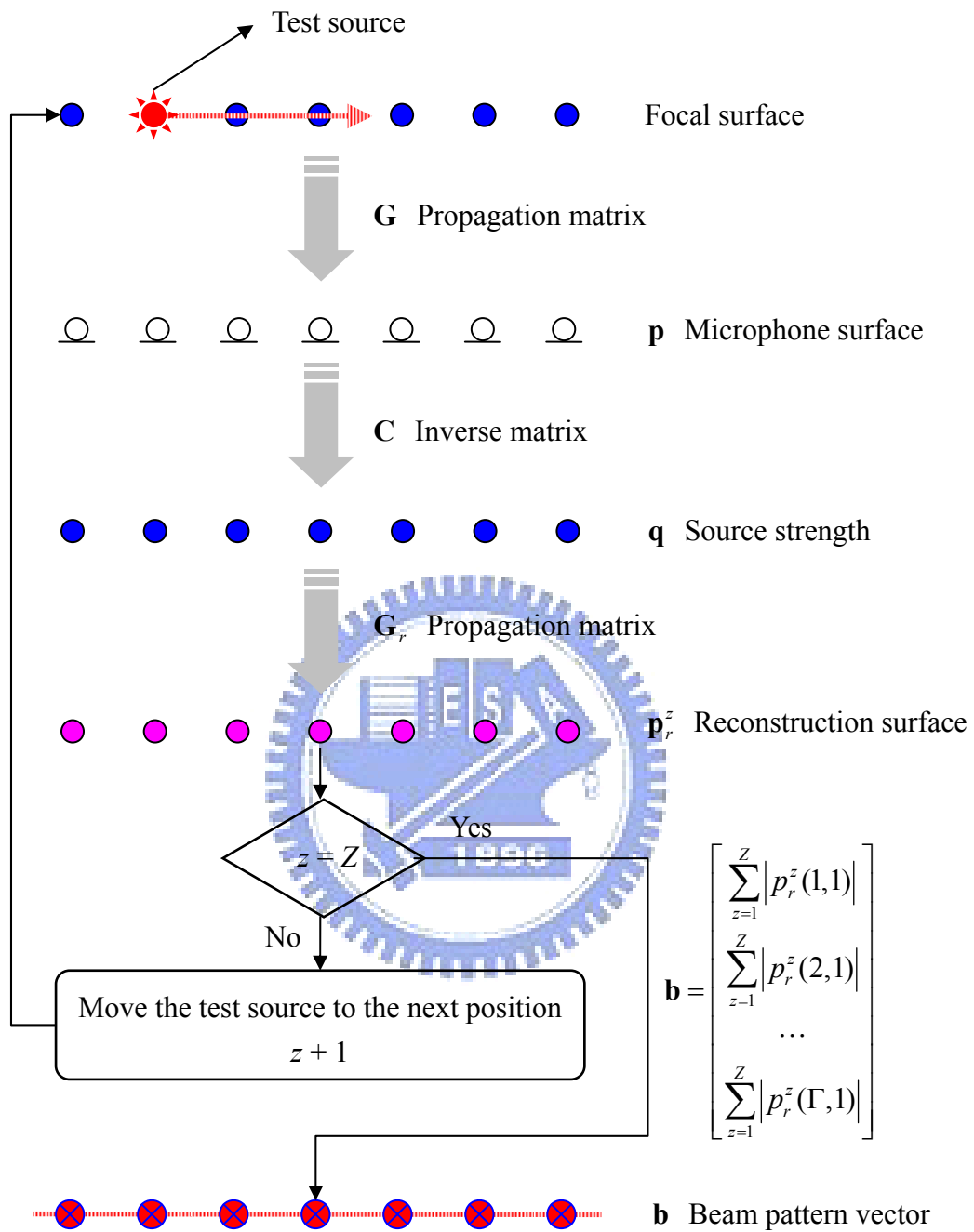


Fig. 77 Flowchart demonstrating the procedure of nearfield beam pattern calculation.

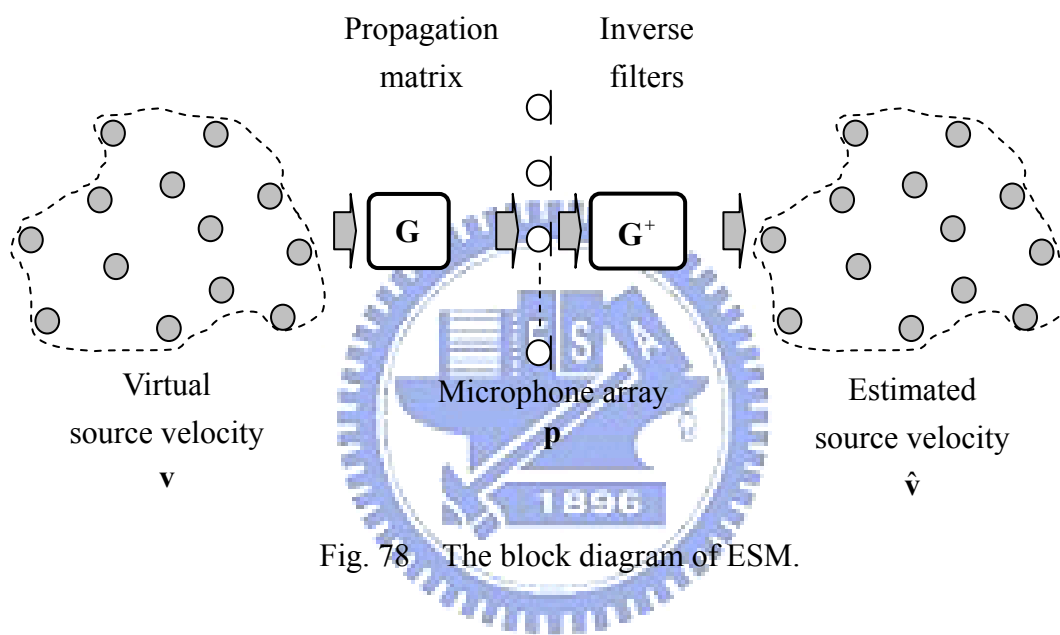


Fig. 78 The block diagram of ESM.

6.7.2 Optimizing farfield sensor deployment

In this section, array optimization is conducted in the context of farfield imaging. Simulations with and without the IB constraint are carried out. The MC and SA algorithms are exploited to optimize microphone deployment with no IB constraint. On the other hand, the SA, IBMC and a combined SA-IBMC algorithm with 5×6 URA and random arrays with 30 microphones as the initial settings are employed to optimize microphone deployment with the IB constraint. Numerical simulations are all based on ten randomly generated starting points.

In the first simulation, we try to optimize the farfield array deployment by the MC and SA. Initially, $m \times n$ divisions ($m = 24$ and $n = 30$) of a square grid are set up on the microphone surface, as shown in Fig. 79 (a). Each side of the square grid measures 0.1m. The source frequency was considered as $f = 1.7$ kHz and the speed of sound as $c = 340$ m/s, yielding the wave number $k = 2\pi f / c = 31.4 \text{ m}^{-1}$. In addition, a URA of 5×6 ($M = 30$) deployment with inter-element spacing $d = 0.6$ m is used as a benchmark in the following simulations, as shown in Fig. 79 (a). Its beam pattern calculated by Eq. (4. 26) is shown in Fig. 79 (b). As expected, the grating lobes are clearly visible because the microphone spacing violates the $\lambda/2$ -rule ($d = 3\lambda$ at $f = 1.7$ kHz). The cost function calculated by Eq. (4. 27) is only 1.0261 because of the grating lobes. This prompts the use of random deployment of microphones as follows.

In the MC simulation, the 30 microphones can freely occupy any 30 positions of the 25×31 grid points on array surface. Exhaustive search would require $16^4 \times 28^{14} \times 49^{12}$ combinations for a 30-element array, while only 10^5 iterations are carried out using this MC search. The MC search attains the optimal cost function 2.6532 at the 27596th iteration. The learning curve of the MC search is shown in Fig. 80 (a). The corresponding microphone positions and beam pattern are shown in Fig.

80 (b) and Fig. 80 (c), respectively. Apart from the extremely time-consuming MC search, the SA approach is employed next. The annealing parameters of the SA for array deployment are chosen to be $T_i = 10$ deg K, $T_f = 10^{-8}$ deg K and $a = 0.95$ [65], [66]. The learning curve of the SA search (405 iterations) is shown in the left portion (denoted as 1stSA) of Fig. 80 (d). The curve fluctuates initially and then converges to a constant value 2.5767 between the 351st and the 405th iteration. The optimal microphone deployment and beam pattern are shown in Fig. 80 (e) and Fig. 80 (f). In addition to optimizing the microphone positions, optimizing the microphone weights can further improve the value of the cost function.

On the basis of the configuration found previously by the SA, we continue to optimize the weights of microphones again using the SA algorithm. The number of iterations is increased to 1000. Starting from unity weights, the microphone weights are adjusted in each iteration with a random perturbation within the range of -0.1 to 0.1. The learning curve is shown in the right portion (denoted as 2nd SA) of Fig. 80 (d). The cost function is further increased to 2.7561 at the 1283rd iteration. The resulting beam pattern is shown in Fig. 80 (g), where a unique main-lobe is clearly visible.

In this section, the SA, IBMC and a combined SA-IBMC algorithm are exploited to optimize microphone deployment with the IB constraint. Both microphone positions and weights are to be optimized using the SA algorithm. Specifically, the combined SA-IBMC method proceeds with three stages—the 1stSA stage, the IBMC stage, and the 2ndSA stage. The parameters of the two SA stages are identical to those in above simulation. The learning curve of the 1stSA stage (405 iterations) is shown in the left portion of Fig. 81 (a). The curve fluctuates initially and then converges to a constant value 2.5328 between the 208th and the 405th iteration. The resulting microphone deployment and beam pattern are shown in Fig. 81 (b) and Fig.

81 (c). Being able to avoid local minima by accepting “bad” solutions in the initial SA search can be a benefit and a shortcoming as well. A shortcoming of the SA algorithm is that it can miss the optimal solution in the initial stage and converges prematurely to a suboptimal one. A hybrid SA-IBMC approach is used in an attempt to address this problem.

The previous deployment obtained by the SA search is used as the input to the IBMC simulation. The microphone position can be randomly chosen from the nine grid points in the localized region. Each region necessarily contains one and only one microphone. Exhaustive search would require prohibitively 9^{30} combinations for a 30-element array, while only 100 iterations are required in the IBMC search. The learning curve of the IBMC (iteration 406-505) is shown in Fig. 81 (a). By the IBMC search, the cost function is further increased to 2.5465 at the 482nd iteration. Fig. 81 (d) and Fig. 81 (e) show the optimal microphone positions and beam pattern obtained at the 482nd iteration. Next, in the 2nd SA stage, the microphone weights are optimized based on the configuration found previously by the SA-IBMC approach. The microphone weights initially set to unity are adjusted in each iteration with a random perturbation within the range of -0.1 to 0.1. The learning curve in 506 iterations is shown in Fig. 81 (a). The cost function is further increased to 2.6602 at the 1429th iteration. The resulting beam pattern is shown in Fig. 81 (f), where a unique main-lobe is clearly visible.

Apart from the URA, the random array deployment is also used as the initial setting in the simulation. For brevity, the results of the MC, IBMC, SA and SA-IBMC simulations are summarized in Table 10. The highest value of the cost function obtained with these ten randomly generated starting points is regarded as the optimal Q in Table 10. The simulation results obtained with and without the IB constraint are compared in terms of number of iterations and the maximum cost

function values. Although the MC approach has reached the highest cost function ($Q = 2.6532$), it takes 27596 iterations to achieve this value. By comparing the results of the MC and IBMC (with the URA as the initial setting), we found that the IBMC approach can attain comparable cost function value to the MC approach with far less amount of computation ($Q = 2.5638$ at the 7662nd iteration of IBMC vs. $Q = 2.6532$ at the 27596th iteration of MC). In comparison with the results obtained using the SA algorithm with the IB constraint ($Q = 2.6602$ for the URA as the initial setting and $Q = 2.6573$ for a random array as the initial setting), the SA approach with no IB constraint has attained a slightly higher cost function ($Q = 2.7561$) with comparable computational complexity. It all boils down to the tradeoff between search time and optimality.

Incorporating the IB constraint could potentially have the following benefits. First, the IBMC algorithm is computationally more efficient than the plain MC algorithm because of smaller search areas. Second, in the hybrid SA-IBMC approach, the IB constraint could possibly improve the SA results when the SA algorithm converges prematurely to a suboptimal result. Third, the IB constraint normally results in uniform distributions of microphones. By “uniform”, we simply mean that microphones would not concentrate at only a few areas, which should not be confused with the deployment of the constant-spacing uniform arrays. In summary, it is fair to say that the IB constraint significantly reduces the computation complexity at the risk of converging to a suboptimal solution which may not be far from the global optimum. This is generally sufficient in practical applications.

Apart from the source frequency of 1.7 kHz, we also run the simulation for the frequencies of 500 Hz and 1 kHz. For brevity, we only summarize the results in Table 11. Random arrays yield unique main-lobe and higher cost function than the URA at 1 kHz. For the lower frequency of 500 Hz, no grating lobes are seen in the

beam pattern of URA, while a higher side-lobe level is found in the beam pattern of the random array. This leads to a higher value of cost function for the URA than the random array at low frequencies. In the next section, we will examine whether the same idea of random array applies to nearfield imaging as well.

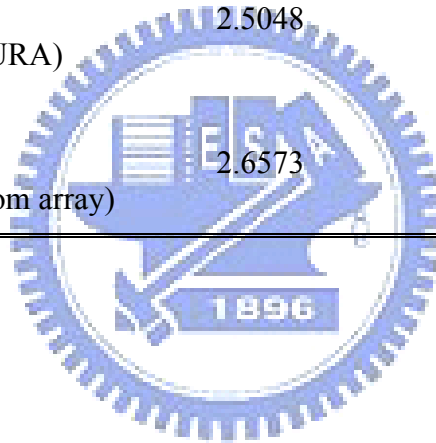


Table 10 The search performance of different optimization methods for farfield array deployment with the inter-element spacing $d = 0.6$ m. The letter “w” indicates that weight optimization is performed.

Constraint	Method	Find best Q Iterations	Best Q (Linear)
without IB (initially random array)	MC	27596	2.6532
	SA	351	2.5767
	SA + w	1283	2.7561
with IB (initially URA)	IBMC	7662	2.5638
	SA	208	2.5328
	SA + IBMC	482	2.5465
	SA + IBMC + w	1429	2.6602
with IB (initially random array)	IBMC	23285	2.5617
	SA	222	2.5224
	SA + IBMC	406	2.5224
	SA + IBMC + w	1352	2.6573

Table 11 The comparison of converged cost function Q of the URA and the optimized farfield random arrays at three different frequencies.

Array	$f = 500$ Hz	$f = 1$ kHz	$f = 1.7$ kHz
URA	4.0216	1.0192	1.0261
Random array (without IB, initially random array)	1.5961	2.5451	2.7561
Random array (with IB, initially URA)	2.5048	2.3324	2.6602
Random array (with IB, initially random array)	2.6573	2.4305	2.6573



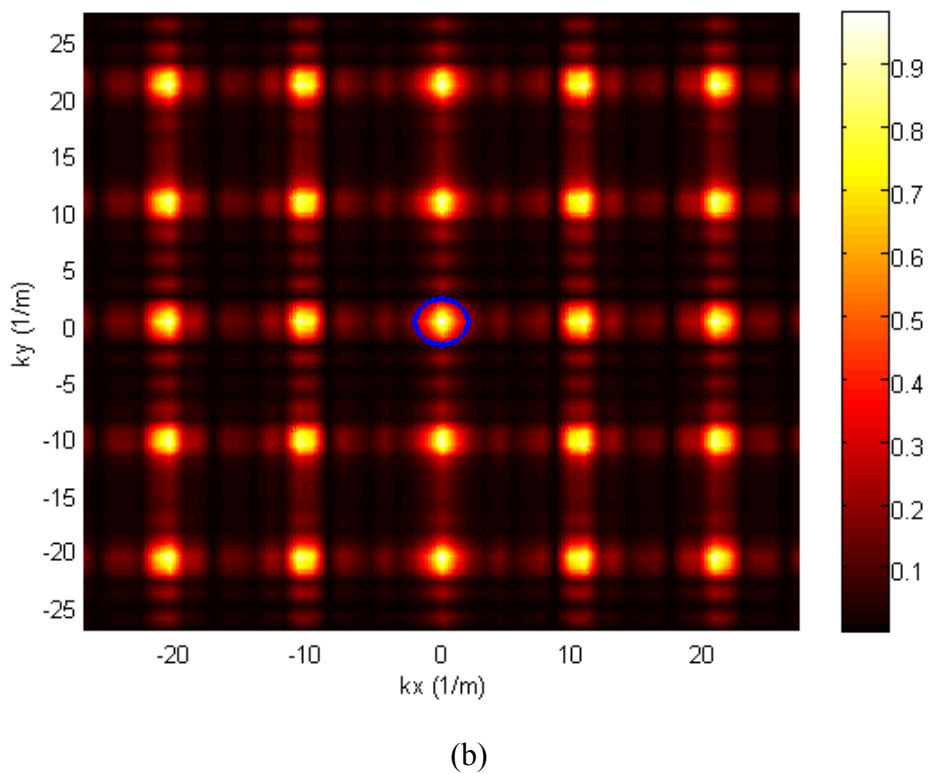
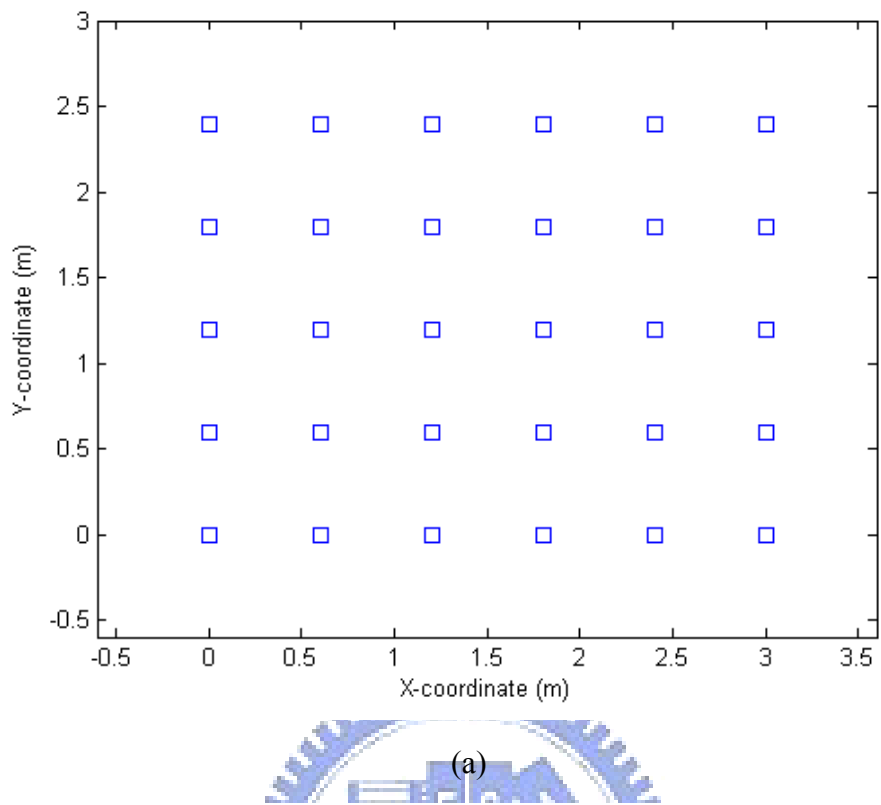
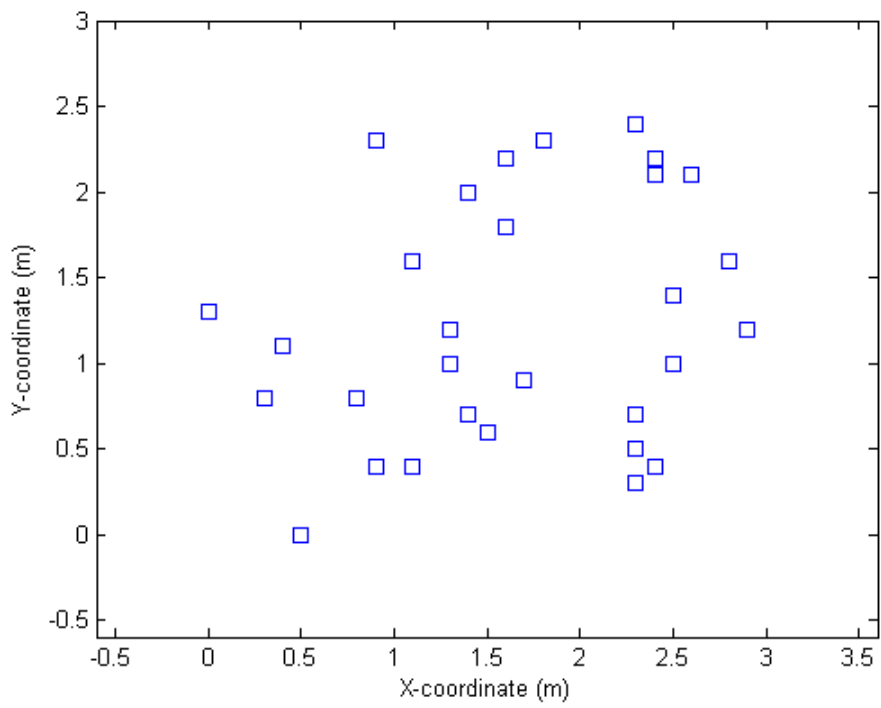
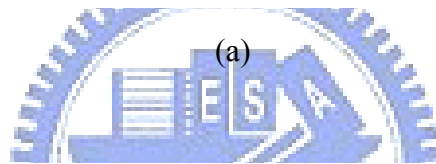
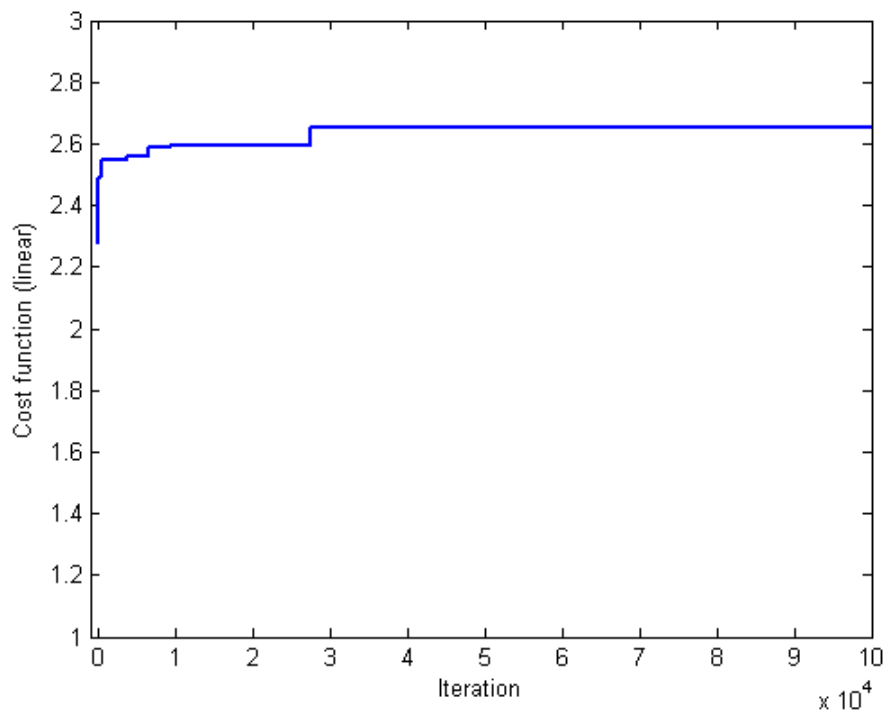
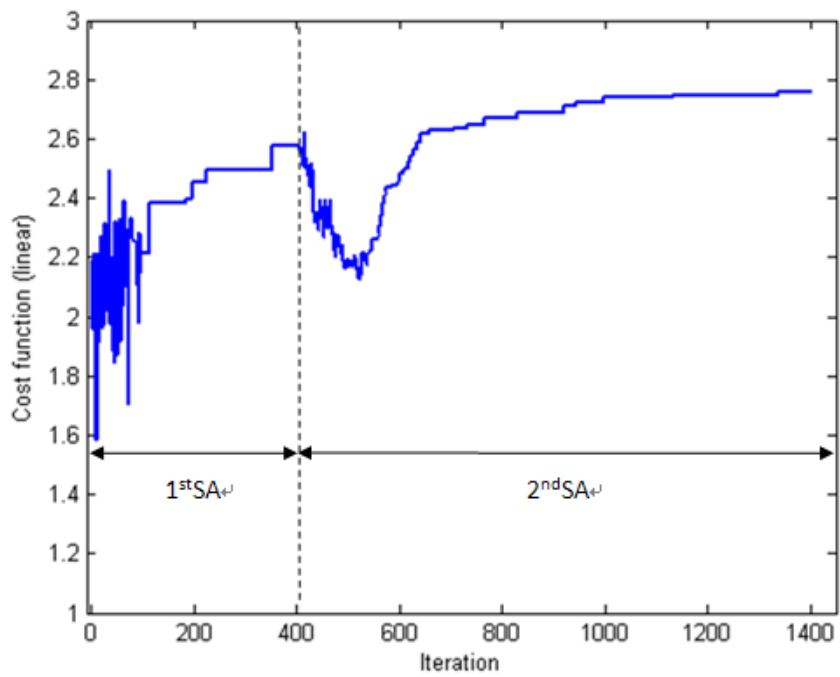
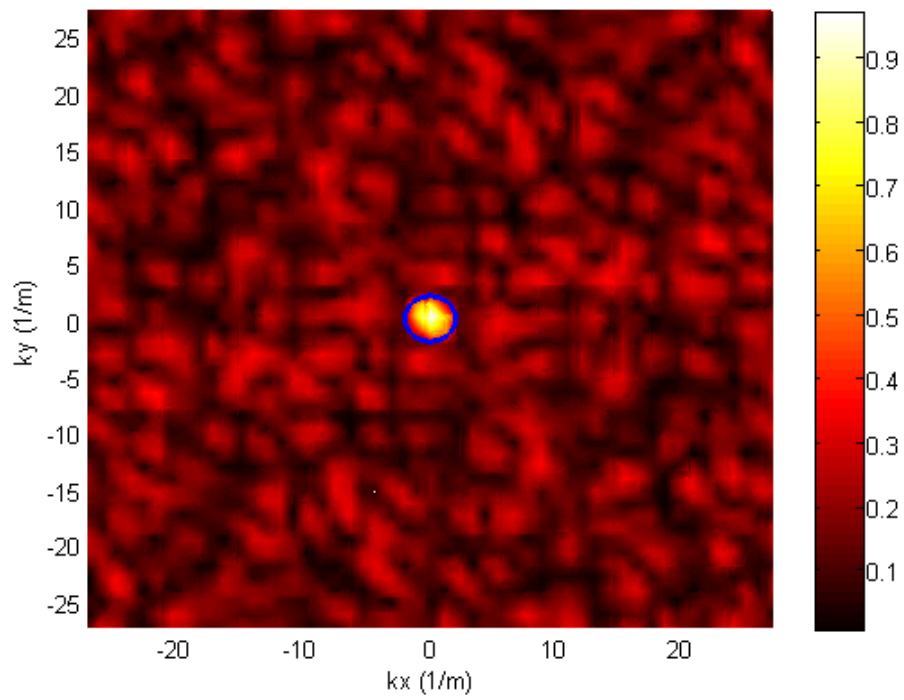


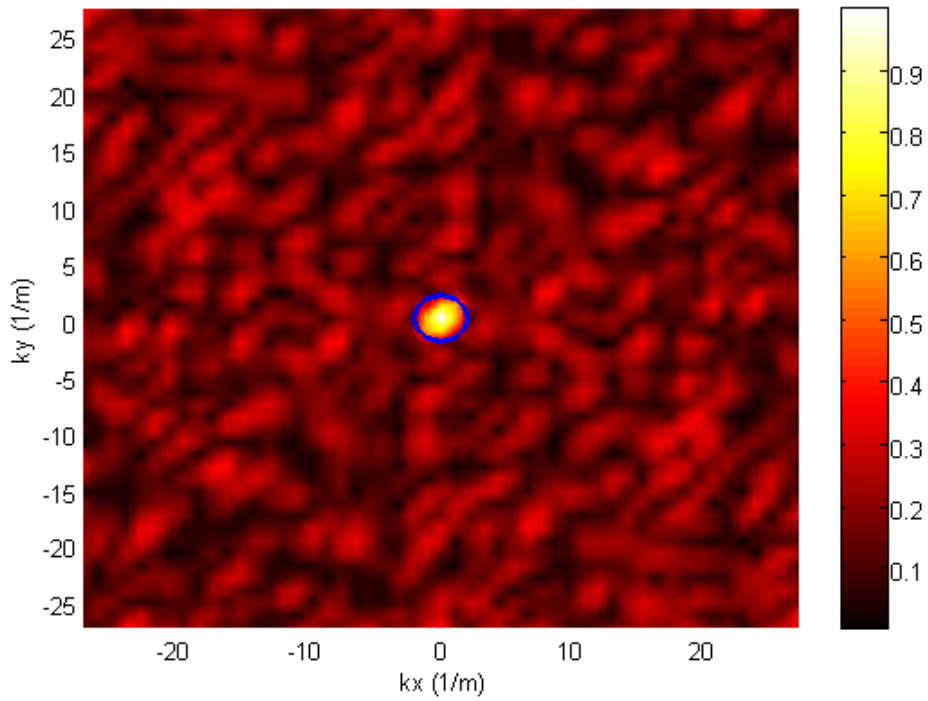
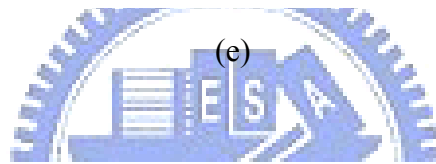
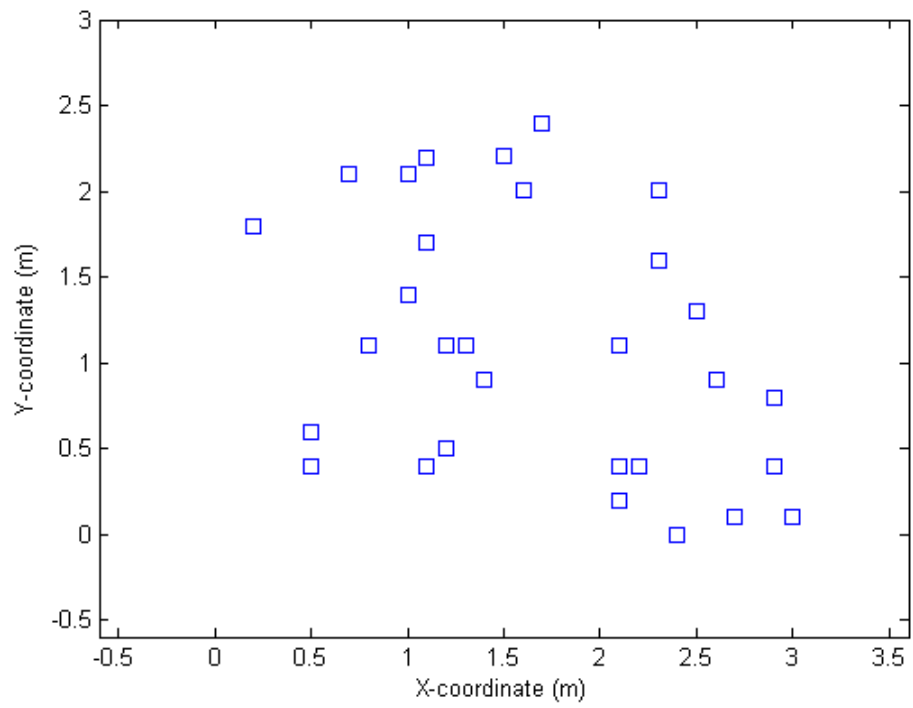
Fig. 79 The URA with inter-element spacing 0.6m (3λ at the frequency 1.7 kHz) for farfield imaging. (a) Array deployment, (b) beam pattern.



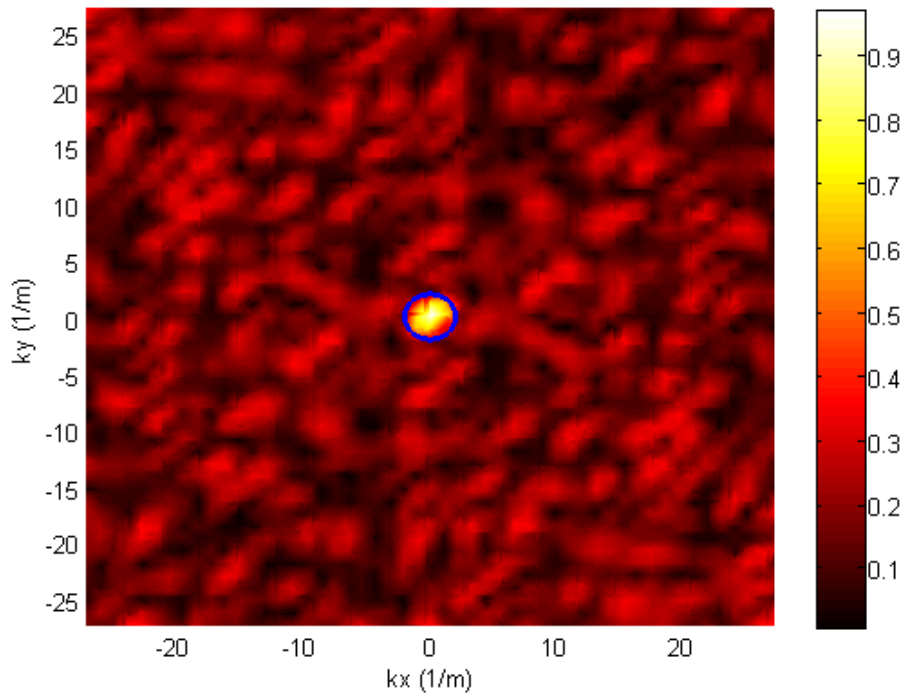
(b)



(d)

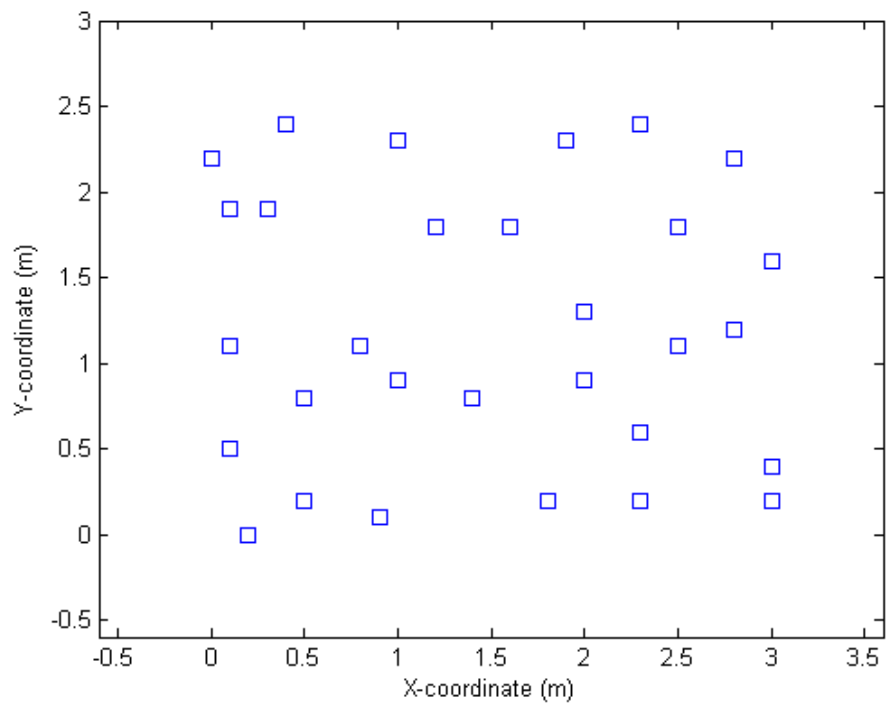
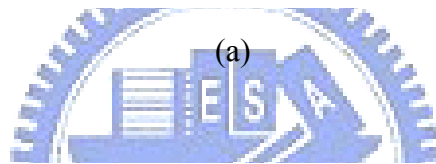
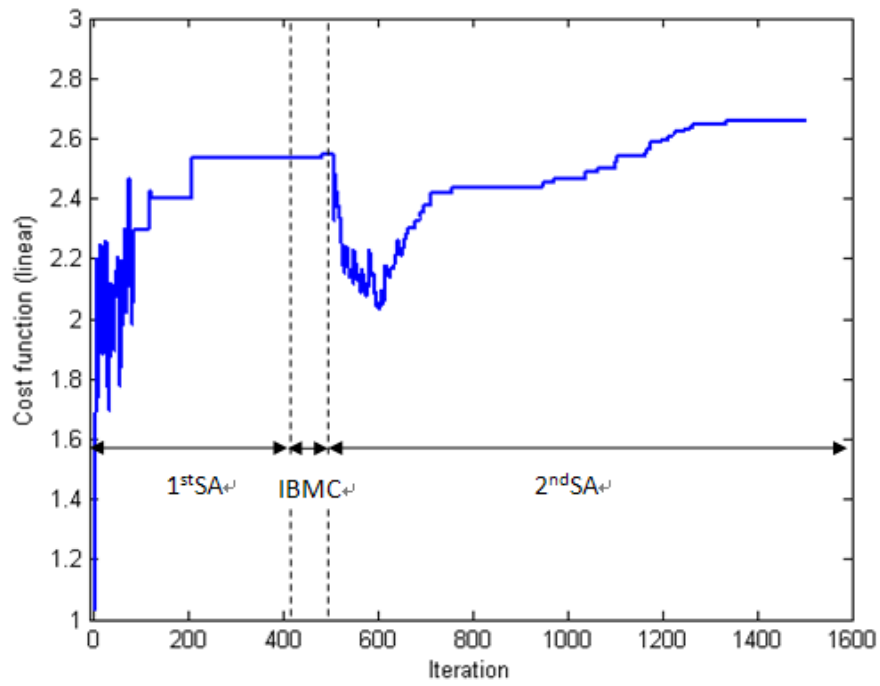


(f)

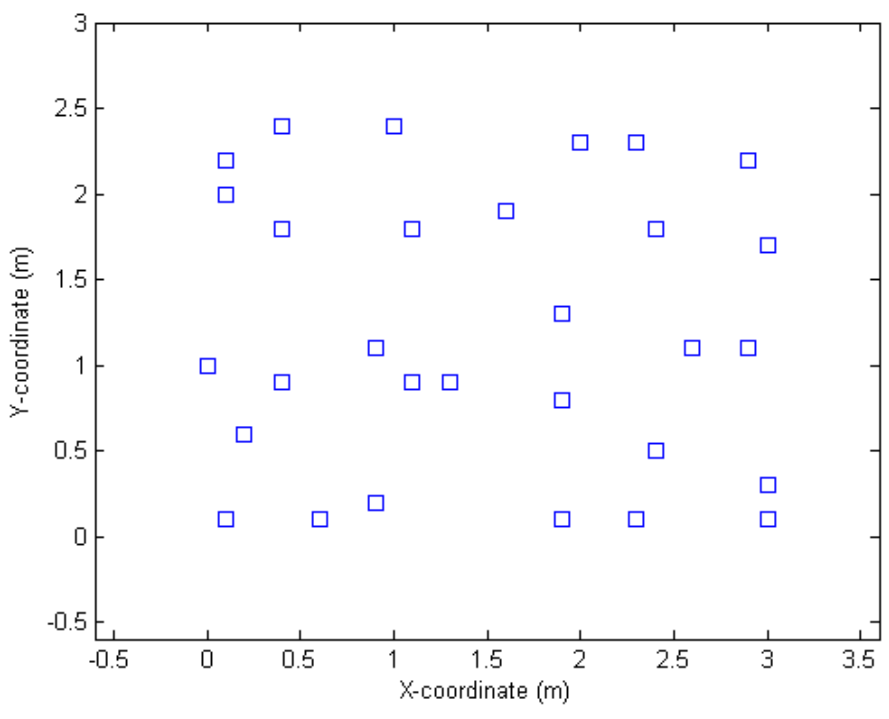
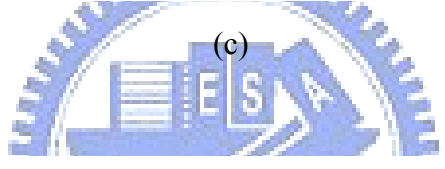
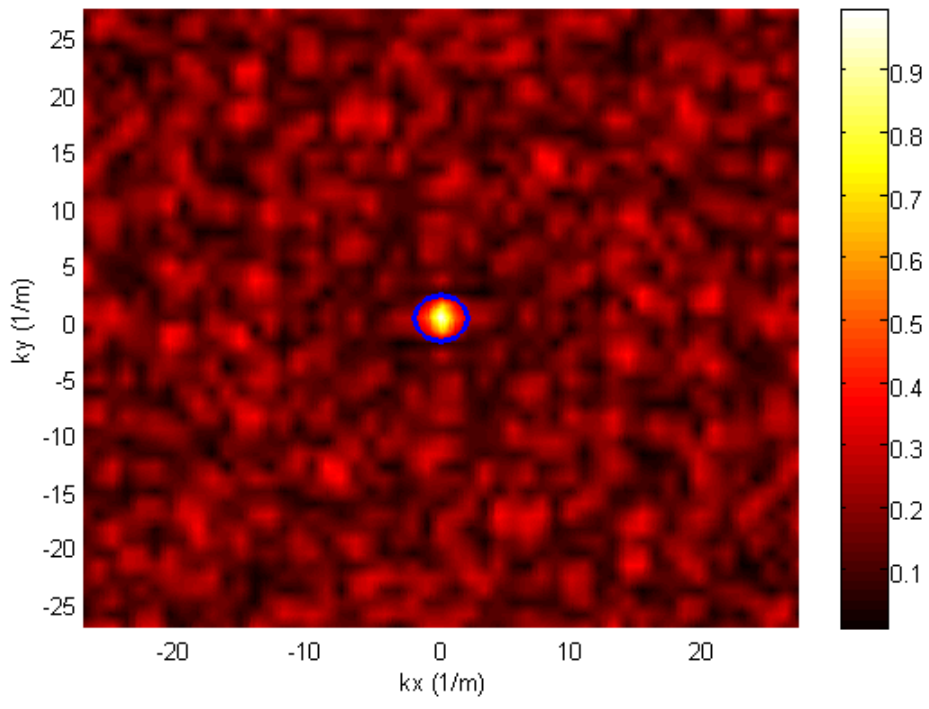


(g)

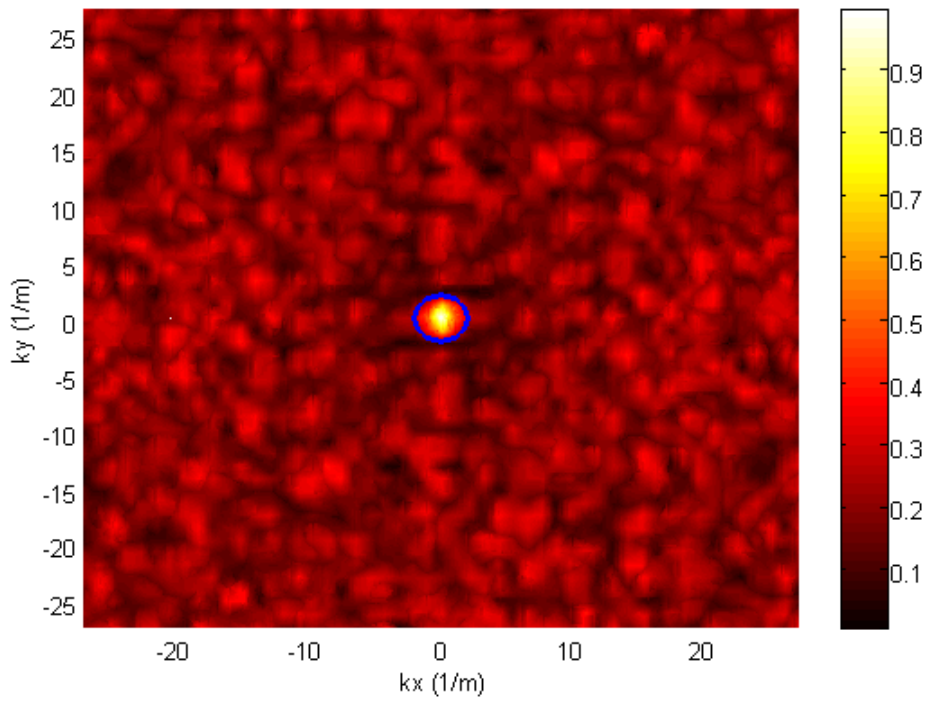
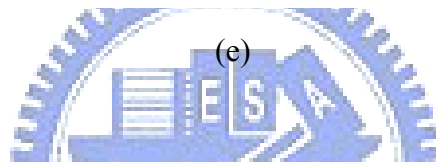
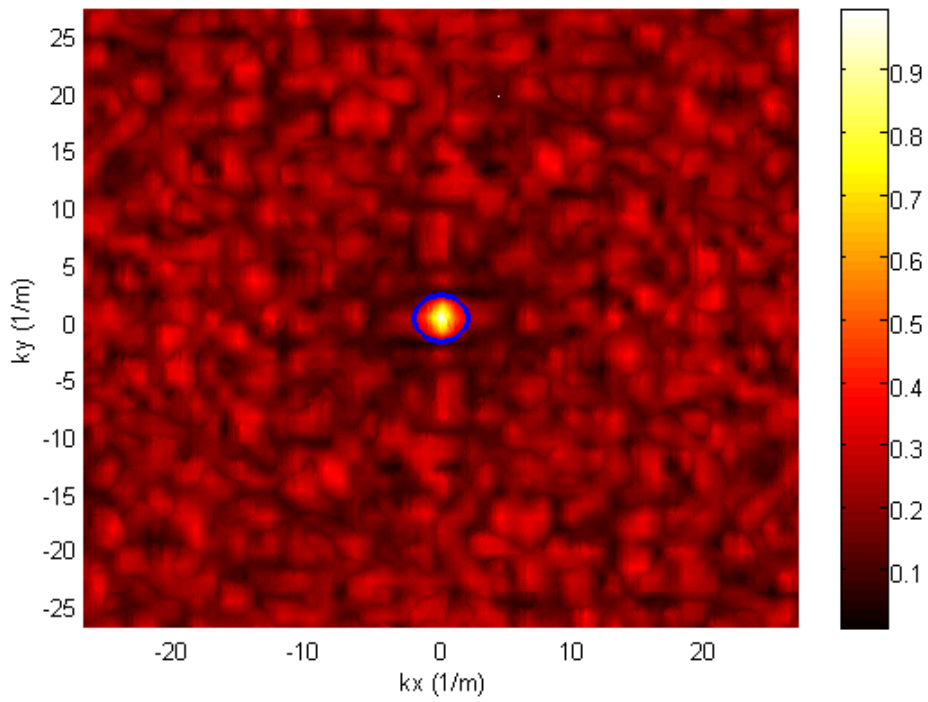
Fig. 80 The farfield array optimized using the MC and the SA algorithms without the IB constraint at the frequency 1.7 kHz. For the MC simulation, maximum cost function $Q = 2.6532$ is attained at the 27596th iteration. For the SA simulation, maximum cost function $Q = 2.7561$ is attained at the 1283rd iteration. The circle indicates the main-lobe. (a) Learning curve of the MC search, (b) optimal array deployment obtained using the MC search, (c) beam pattern obtained using the MC search, (d) learning curve of the SA search, (e) optimal array deployment obtained using the SA search, (f) beam pattern obtained using the SA search, (g) beam pattern with weights optimized.



(b)



(d)



(f)

Fig. 81 The farfield array optimized using the MC algorithm and the combined

SA-IBMC algorithm with the IB constraint at the frequency 1.7 kHz. Maximum cost function $Q = 2.6602$ is attained at the 1429th iteration. The circle indicates the main-lobe. (a) Learning curve, (b) optimal array deployment obtained using the SA search, (c) beam pattern obtained using the SA search, (d) optimal array deployment obtained using the SA-IBMC search, (e) beam pattern obtained using the SA-IBMC search, (f) beam pattern with weights optimized.



6.7.3 Optimizing nearfield sensor deployment

In this simulation, the MC, SA, IBMC algorithms and a combined SA-IBMC algorithm are exploited to optimize microphone deployment for nearfield imaging. The 5×6 URA and random array with 30 microphones are used as the initial setting for the simulation of nearfield imaging. With reference to Fig. 61, the array parameters are chosen as: microphone spacing $d = 0.6\text{m}$, focal point spacing $d_f = 0.6\text{m}$ (3λ at the source frequency of 1.7 kHz), and DOR $L = 0.1\text{m}$. The focal points are collocated underneath with the microphones on a conformal plane (the number of focal points is equal to the number of microphones). Thus, finer 25×31 grid points are set up on the microphone surface. The microphone is free to move to any grid point inside the assigned localized regions. The dimensions of the localized regions (see Fig. 21) are $d_{mx} = d_{my} = 0.6\text{m}$ and the spacing of the grid points is $d_x = d_y = 0.1\text{m}$. The main-lobes are defined as the interior of the circles centered at each focal point with the radius $r_m = 0.02\text{m}$ (the minimal distance between two adjacent grid points). The iteration limit I_{MC} and I_{IBMC} are chosen to be 5000 and 1000, respectively. The Fig. 82 (a) and Fig. 82 (b) show the cost function history of the MC search and the IBMC search with the URA as the initial setting, respectively. The maximum value of the cost function $Q = 27.3$ corresponds to the initial URA deployment cost function at the frequency of 1.7 kHz. Apart from the time-consuming MC algorithm and the IBMC algorithm, another attempt was made to find the optimal deployment using the more efficient SA algorithm. The annealing parameters are chosen to be: $T_i = 10 \text{ deg K}$, $T_f = 10^{-8} \text{ deg K}$ and $a = 0.95$. After 405 iterations, we obtained the learning curve with the URA as the initial setting, as shown in Fig. 82 (c). This learning curve converges to a stable value 3.8502 at the low-temperature stage, as a typical SA behavior. This Q value is apparently suboptimal since it is smaller than the initial $Q = 27.3$ of the URA. The maximum cost function remains the initial cost

function ($Q = 27.3$ at the frequency 1.7 kHz) corresponding to the URA deployment. Finally, the hybrid approach SA-IBMC is employed to optimize the nearfield array with the URA as the initial setting. The learning curve (405 SA iterations + 100 IBMC iterations) is shown in Fig. 82 (d). The maximal cost function remains to be that of the initial URA deployment. Table 12 summarizes the search performance of nearfield optimization methods for three different frequencies 500 Hz, 1 kHz and 1.7 kHz with the URA and the random array as the initial settings. As concluded from the table, even though the other initial setting is used for optimizing the random array, the optimal deployment with the maximal cost function remains to be the URA. These results suggest that the optimal nearfield array is the URA when the number of focal points is equal to the number of microphones. The nearfield beam patterns defined in chapter 6.7.1 calculated for the URA in Fig. 77 at the frequencies 500 Hz, 1 kHz and 1.7 kHz are shown in Fig. 83 (a)-Fig. 83 (c). Thirty main-lobes with equal height centered at the focal points are clearly visible. The beam pattern of a non-optimized nearfield array at the frequency 1.7 kHz by the SA and SA-IBMC algorithms is shown in Fig. 83 (d). Clearly, the converged pattern is a dramatic departure from the URA and the associated beam pattern is quite ugly with quite a few “defocused” points. To see the contrast, we also show an example of random deployment that corresponds to the optimized farfield deployment obtained previous using the SA-IBMC method (the nearfield cost function). The beam pattern of the deployments at the frequencies 500 Hz, 1 kHz and 1.7 kHz are shown in Fig. 84 (a)-Fig. 84 (c). It can be seen from the quite smeared beam patter that many main-lobes are defocused or even missed. These observations lead to a conclusion that contradicts the farfield experience -- random deployment degrades the multi-focusing performance and the URA is the optimal array. In addition, Table 13 summarizes the performance of different numbers of focal points using the same 5×6

URA with the inter-element spacing $d = 0.6\text{m}$ (corresponding to 3λ for the frequency 1.7 kHz) at the frequencies 500 Hz, 1 kHz and 1.7 kHz. The cost function for 5×6 focal points (the number of microphones) is 27.3 at the frequency 1.7 kHz. It increases to 28.3 when the number of focal points is increased to 9×11 . However, for 13×16 focal points the cost function decreases to 16.6, showing that it does not increase indefinitely with the number of microphones.



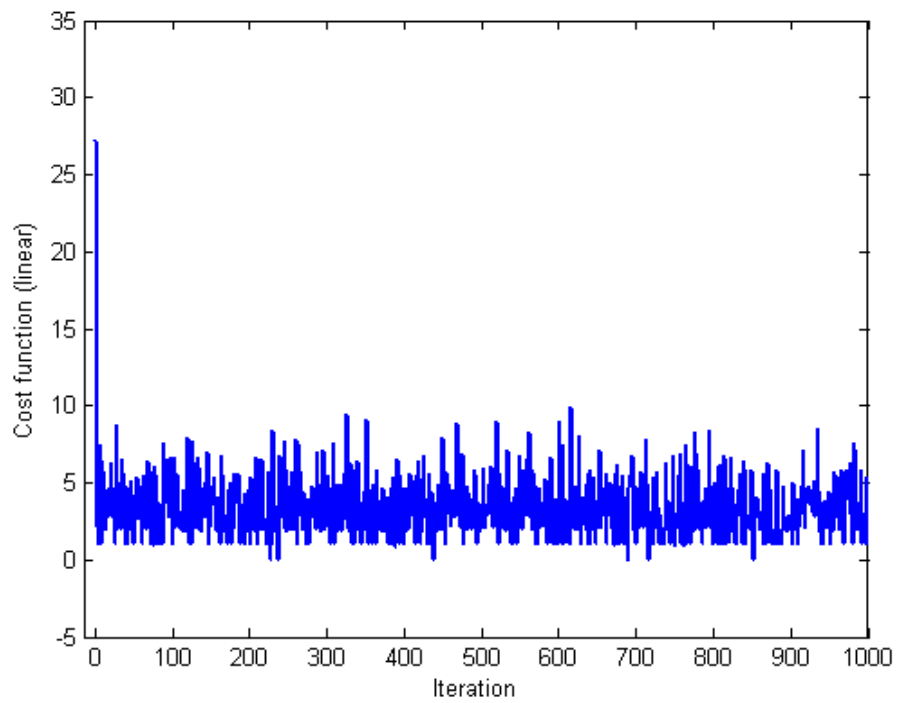
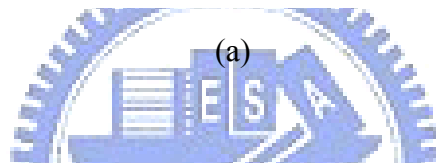
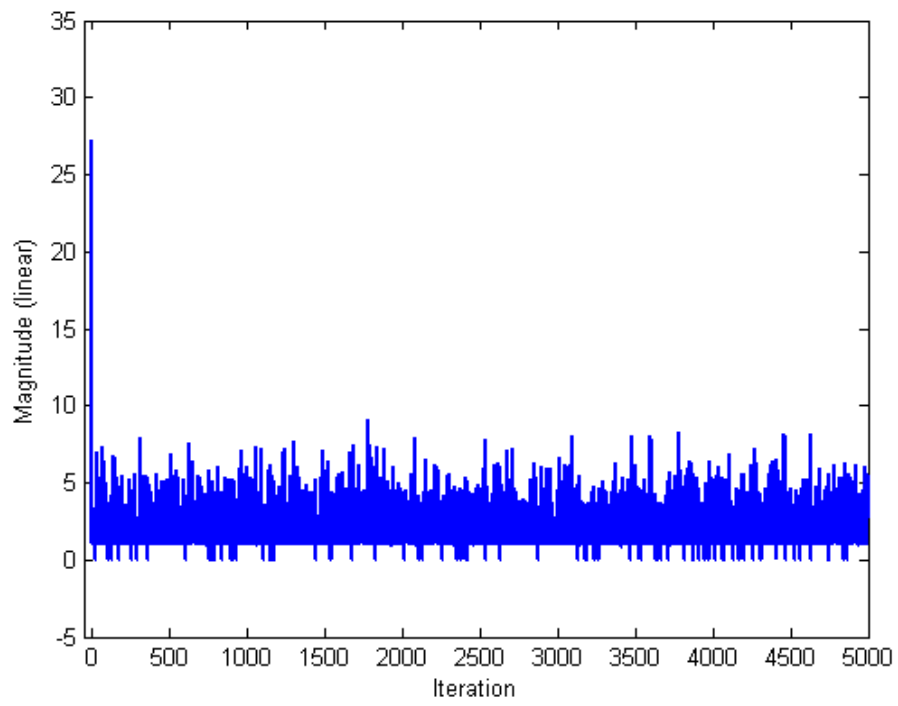
Table 12 The search performance of different optimization methods for nearfield array deployment with the inter-element spacing $d = 0.6$ m at three different frequencies.

Method	Frequency (Hz)	Best Q (Linear)	
		initially URA	initially random array
MC	500	28.2	8.5
	1000	26.9	8.3
	1700	27.3	8.2
IBMC	500	28.2	9.1
	1000	26.9	9.5
	1700	27.3	10
SA	500	28.2	4.2
	1000	26.9	5.6
	1700	27.3	4.7
SA+IBMC	500	28.2	4.2
	1000	26.9	5.6
	1700	27.3	4.7

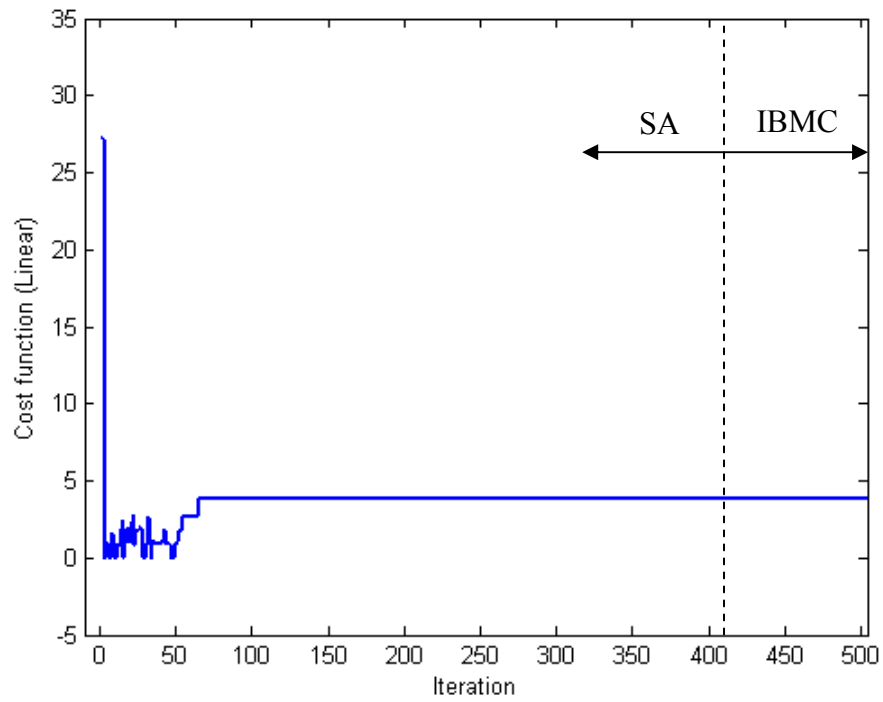
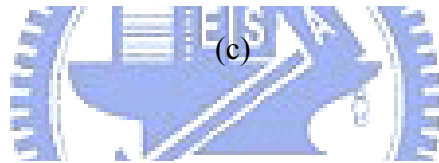
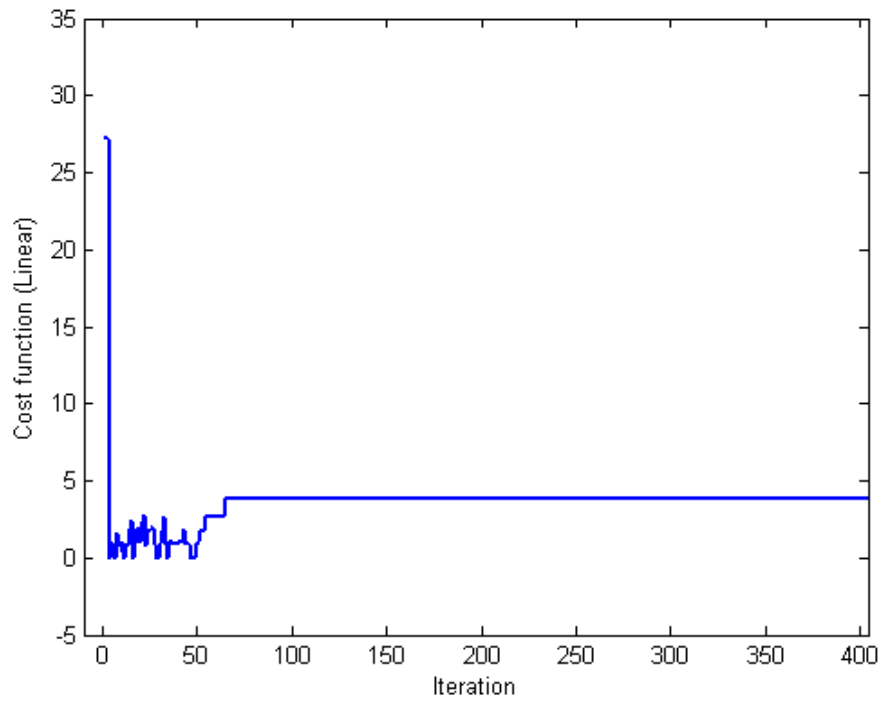
Table 13 The performance obtained using different numbers of focal points in a 5×6 URA with the inter-element spacing $d = 0.6\text{m}$ at three different frequencies.

Frequency (Hz)	Q (Linear)		
	5×6 focal points	9×11 focal points	13×16 focal points
	500	28.2	28.8
1000	26.9	27.2	21.8
1700	27.3	28.3	16.6





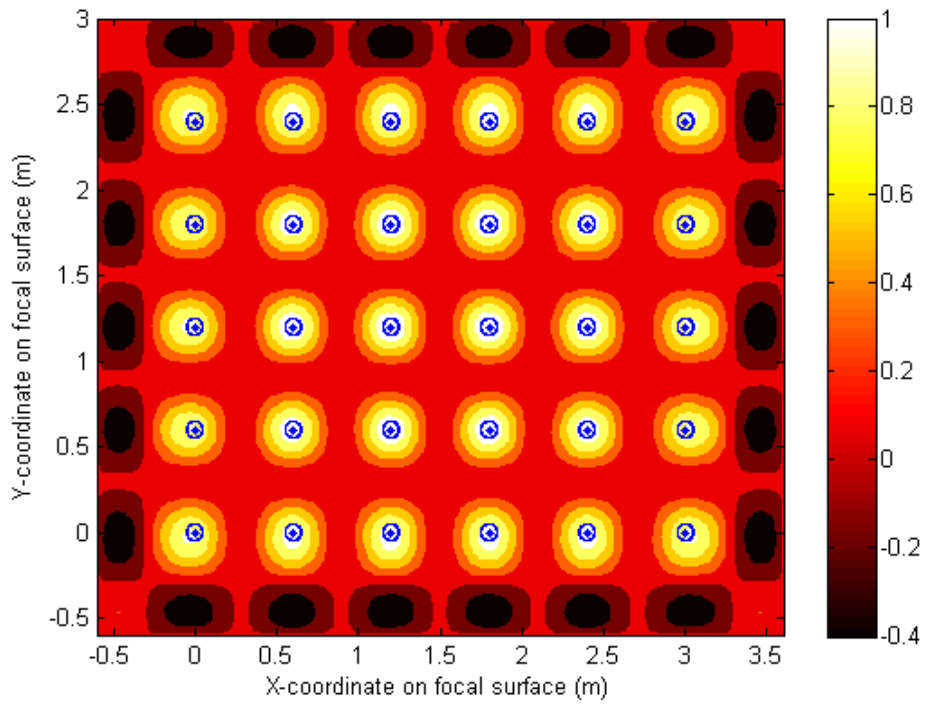
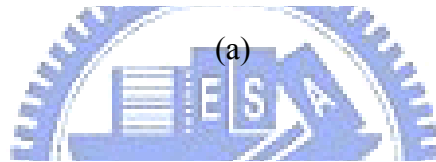
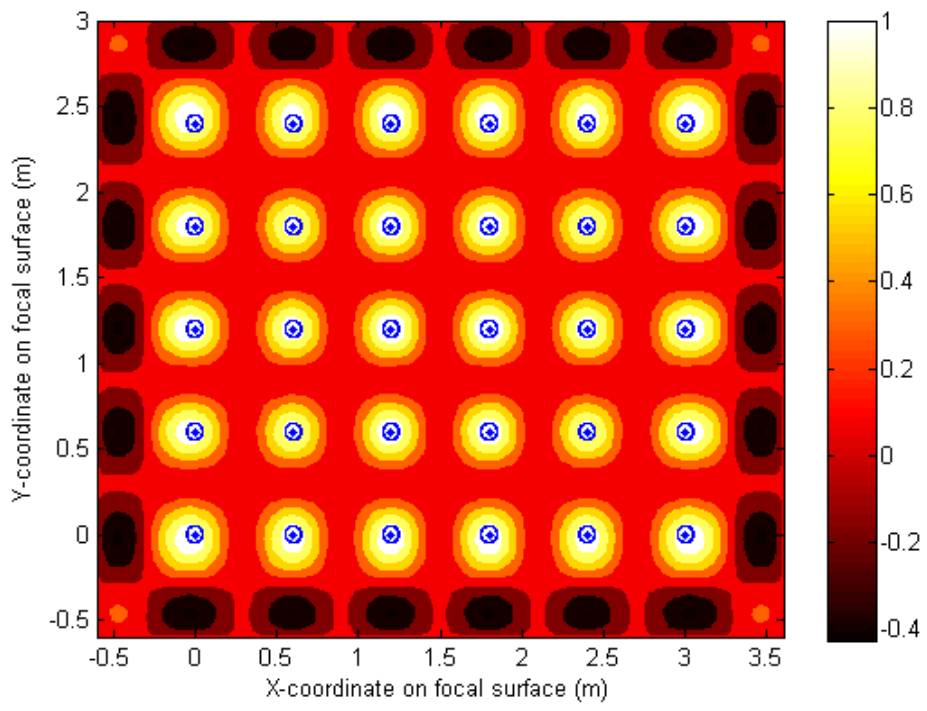
(b)



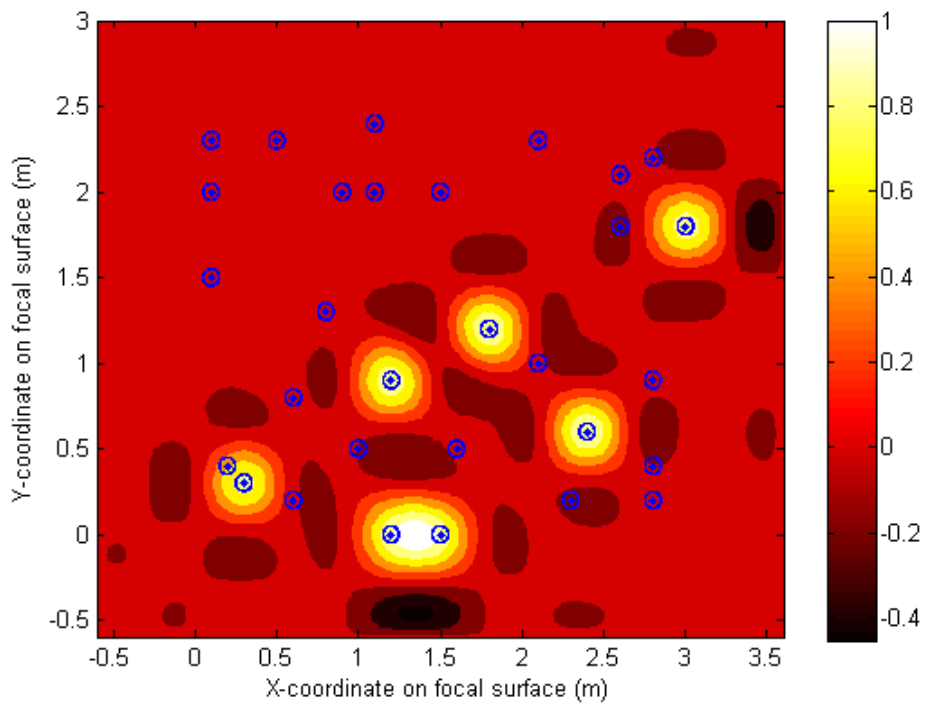
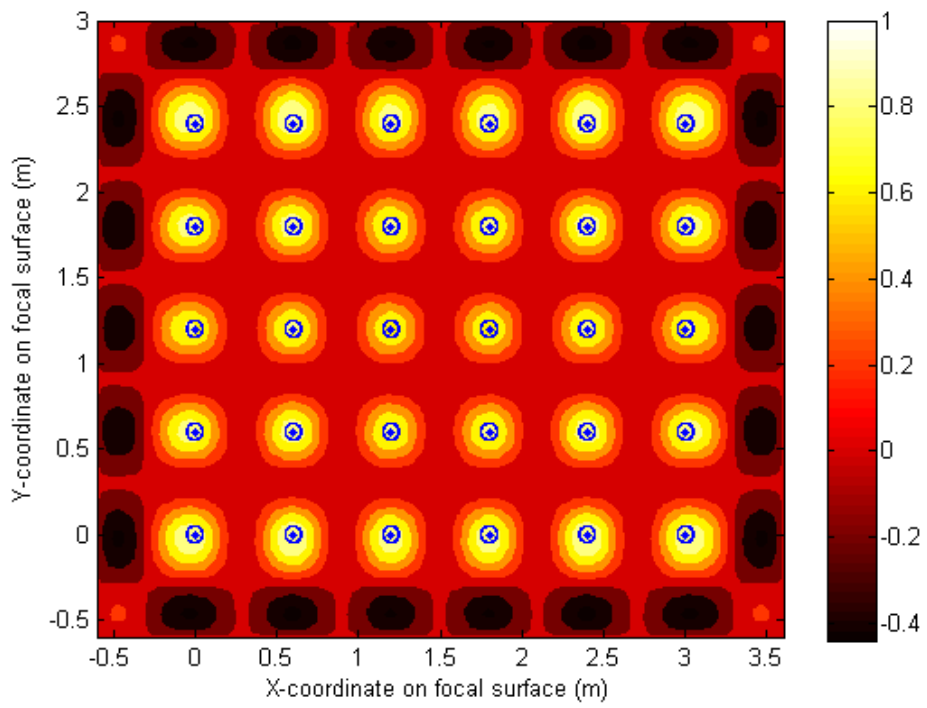
(d)

Fig. 82 The cost function history of nearfield array optimization with inter-element spacing 0.6m (3λ at the frequency 1.7 kHz). Maximum cost function value is $Q = 27.3$ for the initial URA deployment. (a) All Q values, including those rejected during the random search by the MC algorithm, (b) all Q values, including those rejected during the random search by the IBMC algorithm, (c) only the legitimate values accepted by the SA algorithm, (d) only the legitimate values accepted by the SA-IBMC algorithm.





(b)

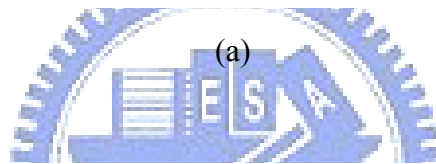
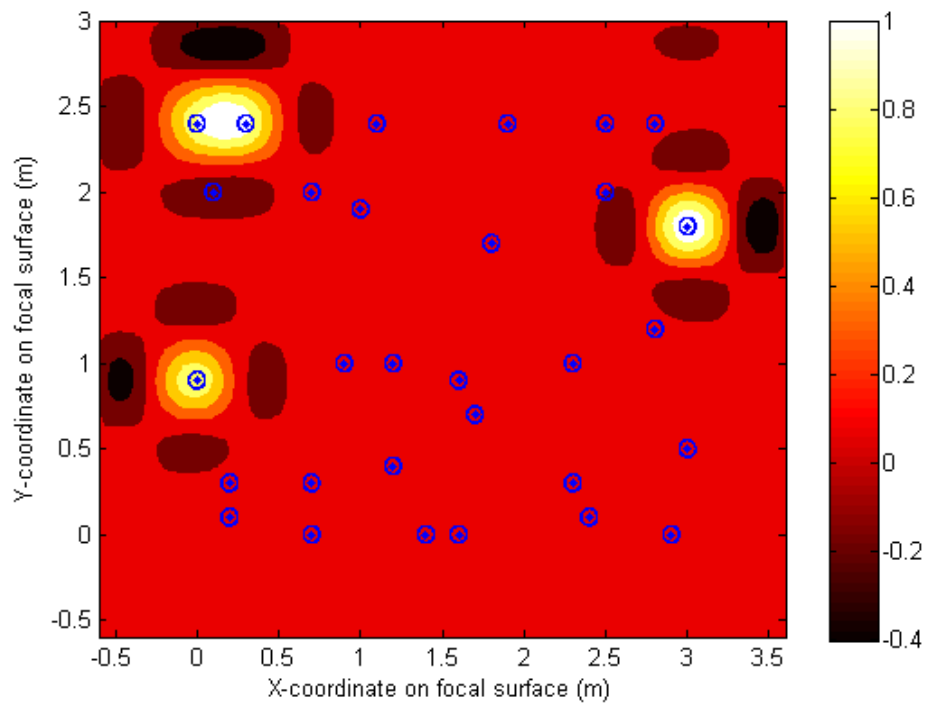


(d)

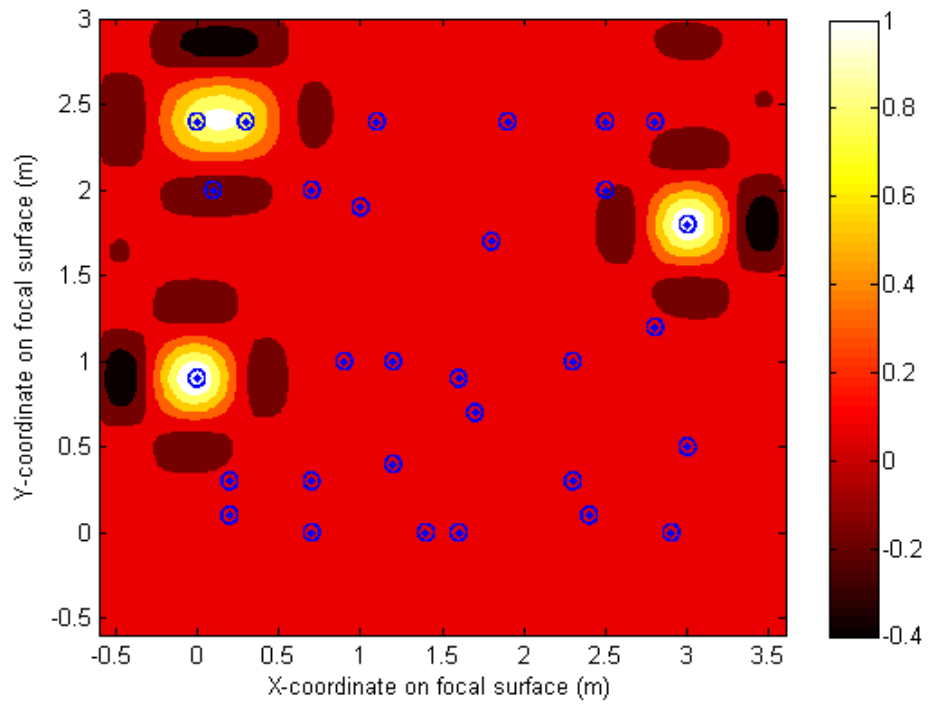
Fig. 83 The beam pattern of the optimized nearfield array (URA). The inter-element

spacing is 0.6m (3λ at the frequency 1.7 kHz). (a) At the frequency 500 Hz, $Q = 28.2$, (b) at the frequency 1 kHz, $Q = 26.9$, (c) at the frequency 1.7 kHz, $Q = 27.3$, (d) the beam pattern of a non-optimized nearfield array at the frequency 1.7 kHz by the SA and SA-IBMC algorithms, $Q = 3.8502$. The circles indicate the main-lobe. The symbol“ • ” indicates the focal points.

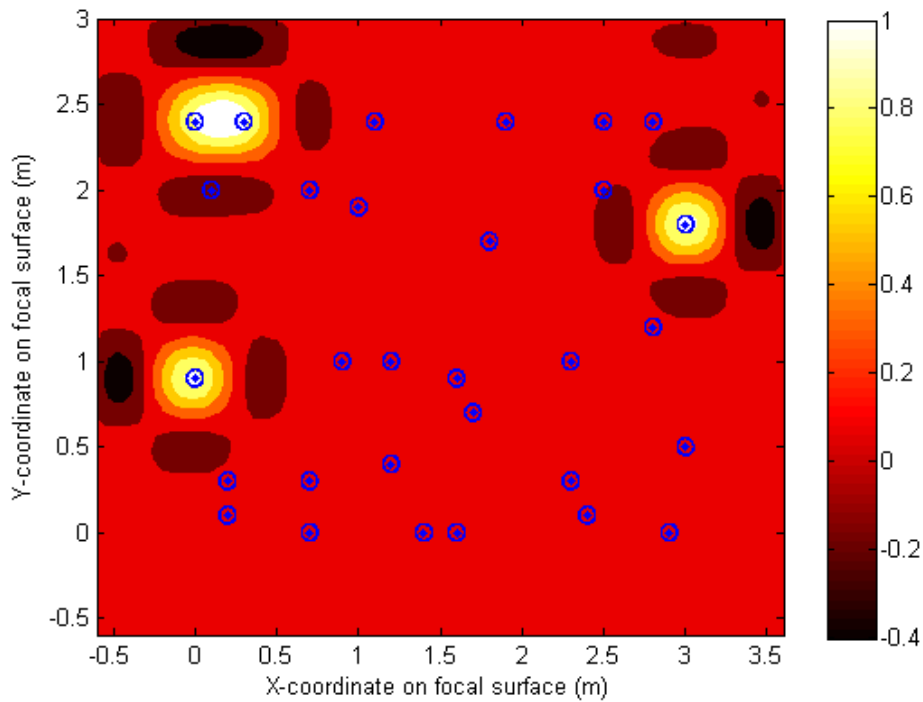




(a)



(b)



(c)

Fig. 84 The beam pattern of a non-optimized nearfield array. The microphone deployment is identical to the optimized farfield array obtained using the SA-IBMC algorithm. (a) At the frequency 500 Hz, $Q = 3.5$, (b) at the frequency 1 kHz, $Q = 3.4$, (c) at the frequency 1.7 kHz, $Q = 3.9$. The circles indicate the main-lobe. The symbol“ • ” indicates the focal points.

6.8 System integration and experimental arrangement

To validate the NAH technique, experiments were undertaken in the laboratory. The block diagram of the experimental arrangement is shown in Fig. 85. Two PXI 4496 systems [82] in conjunction with LabVIEW [82] were used for data acquisition and processing. A bandpass filter is used to prevent aliasing and errors occurring in the out-of-band frequencies. The source amplitude, source strength, sound pressure, particle velocity and sound intensity reconstructed using NESI can be displayed on the monitor.

In the experimental arrangement, sensor calibration is requisite before microphone measurement. First, set a real source (random sources with band-limited to appointed frequency that usually $f = f_s / 2$) at the center of array at far distant place, as shown in Fig. 86. The frequency response H is measured by this calibration structure. Secondly, calibration filter is calculated for the calibration processor with H_1 being the benchmark filter

$$H_m^a = \frac{H_1}{H_m}, m = 1, 2, \dots, M, \quad (6.62)$$

where M is number of microphone. Finally, it is noted that the calibration filter H^a obtained are still in the frequency-domain. Inverse FFT is called for to convert and truncate these FRFs into FIR filters in the time-domain. In this step, circular shift is needed to ensure that the resulting filters are causal filters. Now that the calibration filters are available, the response of microphones is calibrated by using calibration filter:

$$p'_m(n) = p_m(n) * H_m^a, m = 1, 2, \dots, M, \quad (6.63)$$

where n denotes the discrete-time index and p_m is pressure measurement by m th microphone.

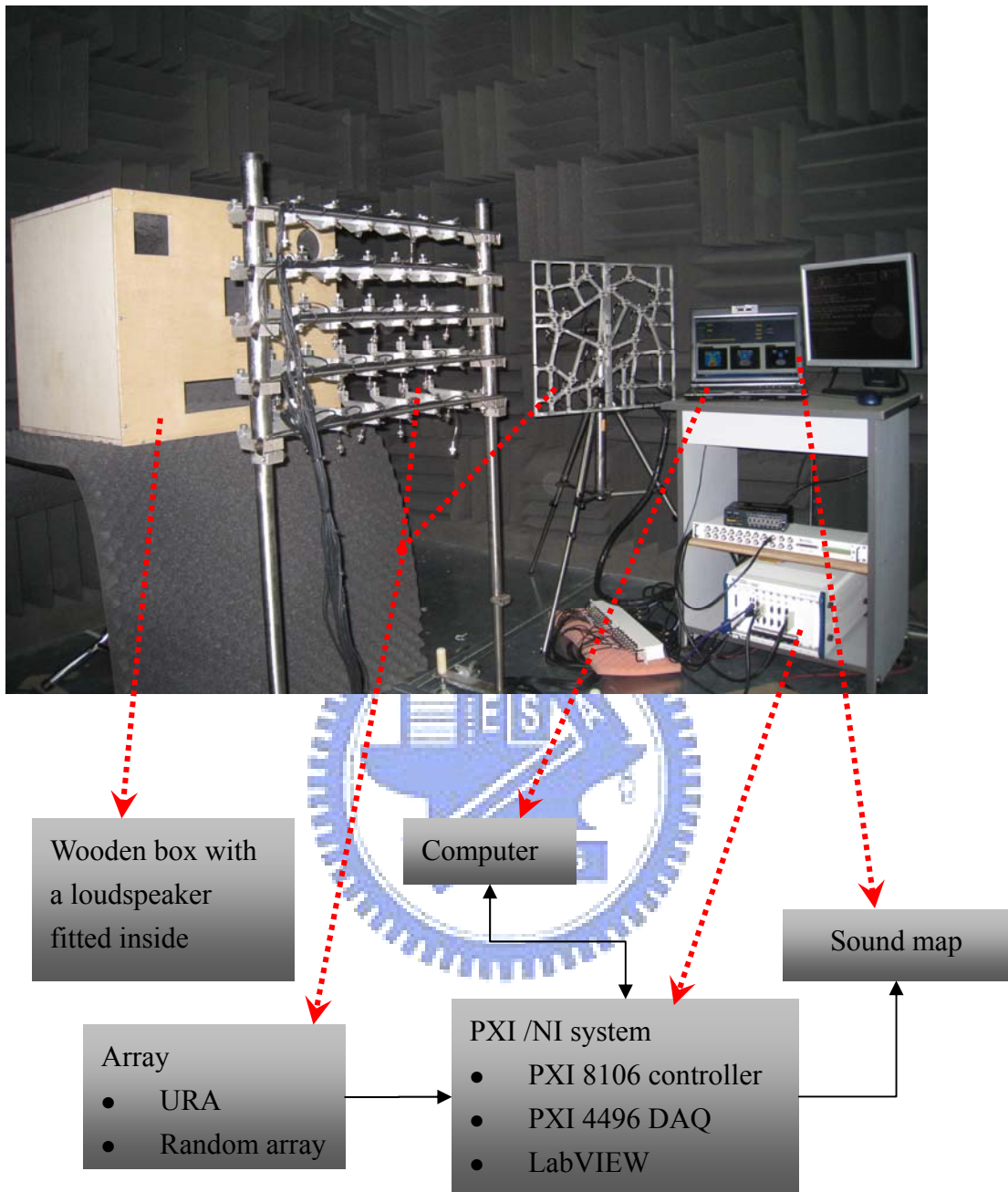


Fig. 85 The experimental arrangement for a wooden box with a loudspeaker fitted inside ,the URA, and a 30-channel random array optimized for farfield imaging are also shown in the picture.

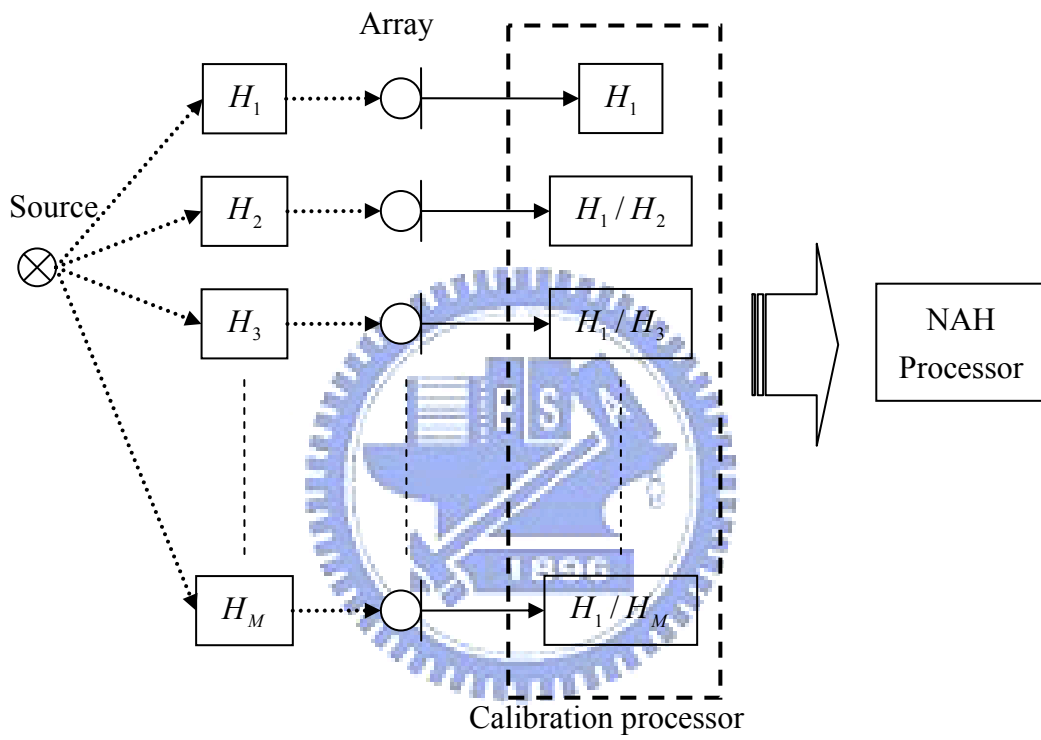


Fig. 86 Sensor calibration.

CHAPTER 7. APPLICATION EXAMPLES

Practical sources such as scooter, vehicle, IT equipment, compressor and a vibrating plate were chosen as the application examples of the NESI technique. The close-up view of the microphone array used in the experiments is shown in Fig. 87.



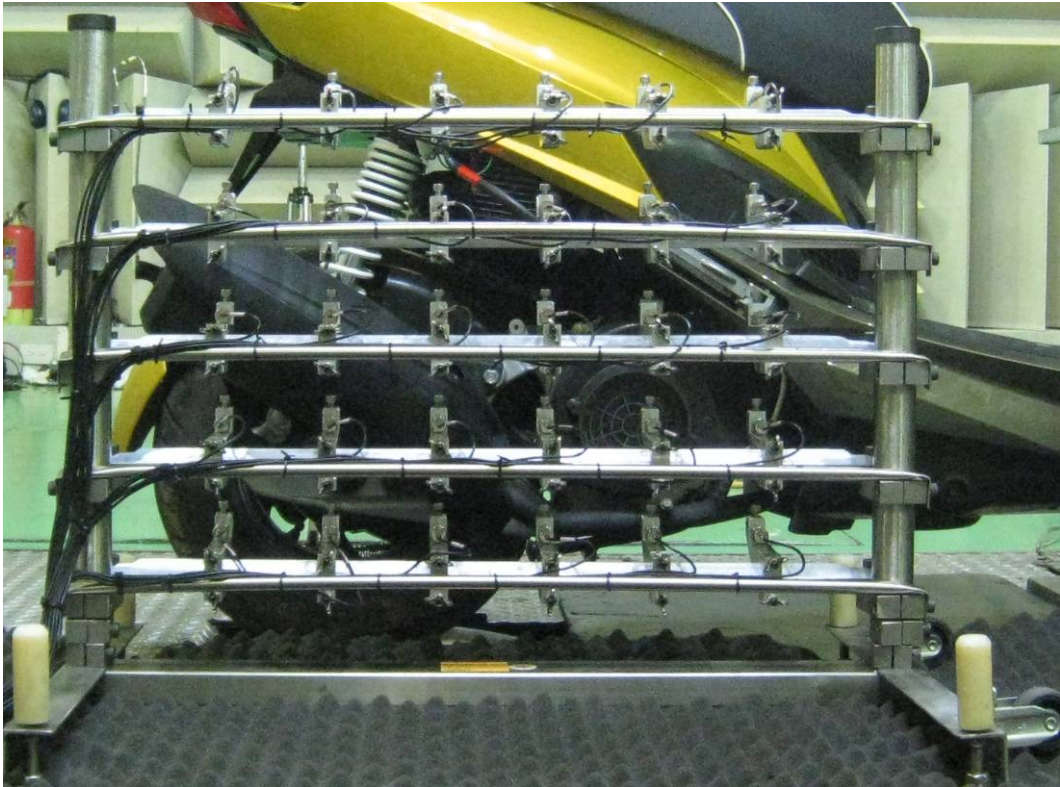


Fig. 87 Close-up view of the 5×6 rectangular array.

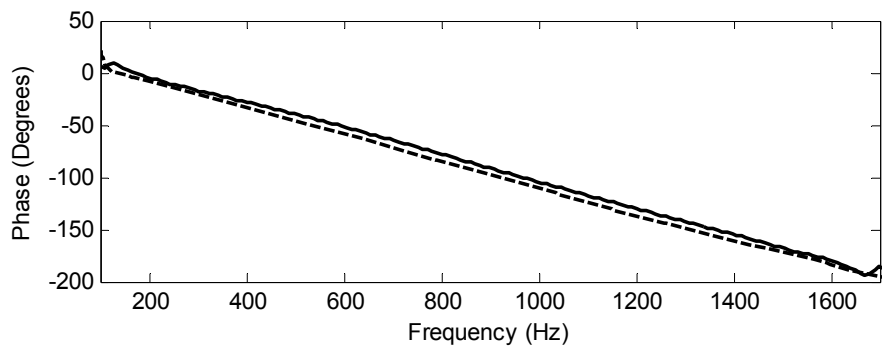
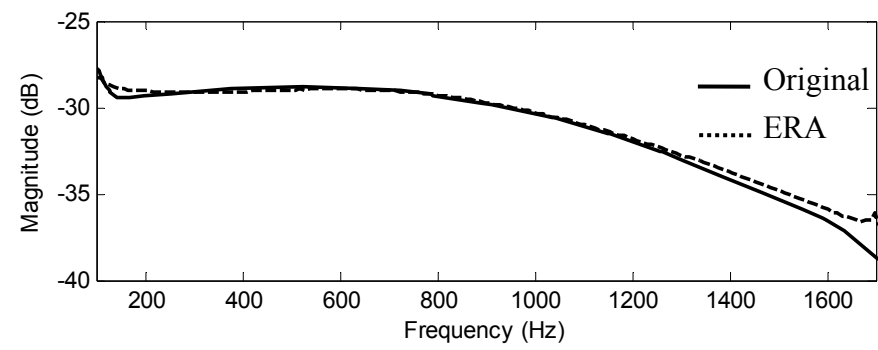
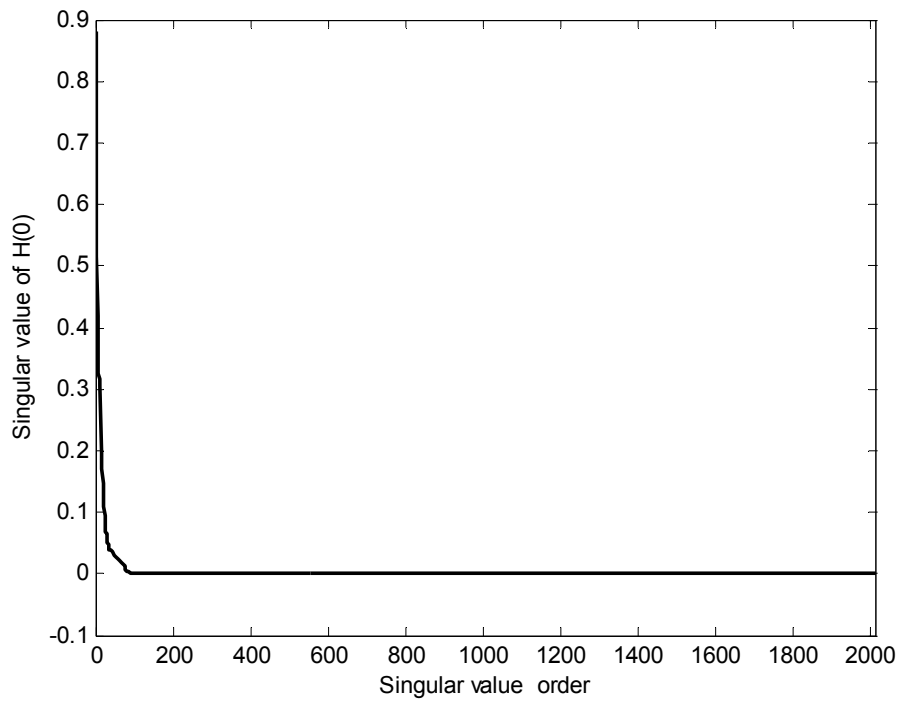
7.1 Scooter: transient sources

In the following, a 125cc SYM scooter served as a more practical source to justify performance of the NESI technique. The scooter is mounted on a dynamometer inside a semi-anechoic room. The 4×4 rectangular array was placed in front of the scooter engine case. The array parameters are selected to be $M = J = 4 \times 4$, $d = d_f = 0.1\text{m} = \lambda/2$ for 1.7 kHz and $L = d/2$. The inverse filtering was carried out by using $16 \times 16 = 256$ channels of FIR filters, each filter of order 252. In a total, $256 \times 252 = 64512$ multiplications are required for DC, which is computationally demanding. To reduce the complexity, the minimal realization obtained using the ERA is used. The singular values of $\mathbf{H}(0)$ are plotted in Fig. 88 (a), indicating the lowest 80 orders in SVD suffice to realize the multichannel inverse filter. The close agreement of the frequency responses and the impulse responses of the inverse filter C_{13} between the original filter and the filter regenerated using the ERA with $\nu = 80$ are shown in Fig. 88 (b)-Fig. 88 (c). In this case, only $\nu^2 + \nu M + J\nu + JM = 9216$ multiplications are required using the ERA state-space model. The computational efficiency has been improved by a factor of seven.

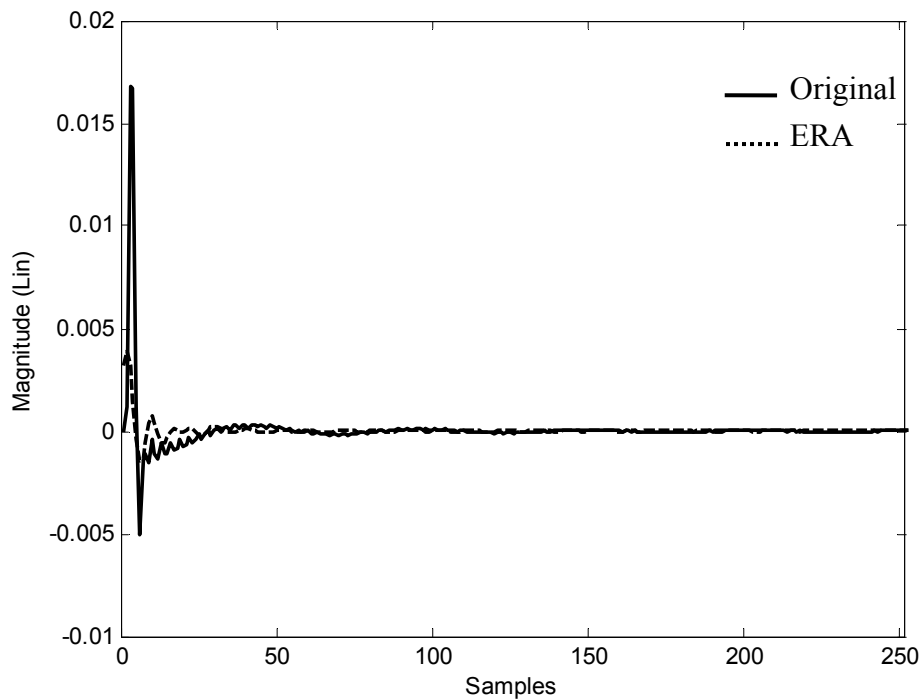
In the first experiment, the NESI was applied to reconstruct the sound field on the right side of the scooter running at the idle speed. In order to enhance computational efficiency, the aforementioned ERA technique with $\nu=28$ is used for multichannel filtering. The unprocessed rms sound pressure picked up at the microphones is shown in Fig. 89 (a). This figure would lead to the incorrect conclusion that the major noise source was located at the left boundary unless the NESI was applied. The rms source strength, sound pressure, particle velocity, and sound intensity reconstructed by using the NESI is shown in Fig. 89 (b)-Fig. 89 (e). With NESI, the dark red area in the reconstructed pressure field reveals that the cooling fan at the center was the main culprit. In addition to the cooling fan, the

reconstructed sound fields (Fig. 89 (b)-Fig. 89 (e)) indicate that there were secondary sources at (0m, 0m) and (0m, 0.2m) on the tire.

In final experiment, a 125cc scooter served as a practical source to examine the capability of NESI in dealing with non-stationary sources. The scooter is mounted on a dynamometer inside a semi-anechoic room. The NESI was used to reconstruct the sound field on the right side of the scooter in a run-up test. The engine speed increased from 1500 rpm to 7500 rpm within ten seconds. The unprocessed sound pressure received at the microphones is shown in Fig. 90 (a), while the rms velocity reconstructed using the NESI is shown in Fig. 90 (b). These results revealed that the cooling fan behind the vented engine cover was the major noise source. Next, the virtual microphone technique is employed to see if it is possible to further enhance the image quality by increasing the number of channels from $4 \times 4 = 16$ to $11 \times 11 = 121$. The inverse filters have been designed in the previous numerical investigation. The particle velocity was then reconstructed on the basis of the estimated source amplitude, as shown in Fig. 90 (c). Total sound power level is 95 dB re. 1×10^{-12} W. Clearly visible is a larger area of image with improved resolution than that of Fig. 90 (b), where again the cooling fan is the major noise source. Therefore, this experiment proved that the proposed NESI technique applies to not only broadband random but also transient noise sources.

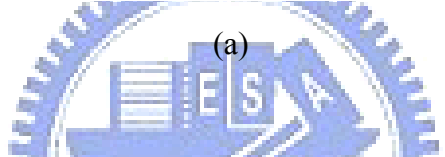
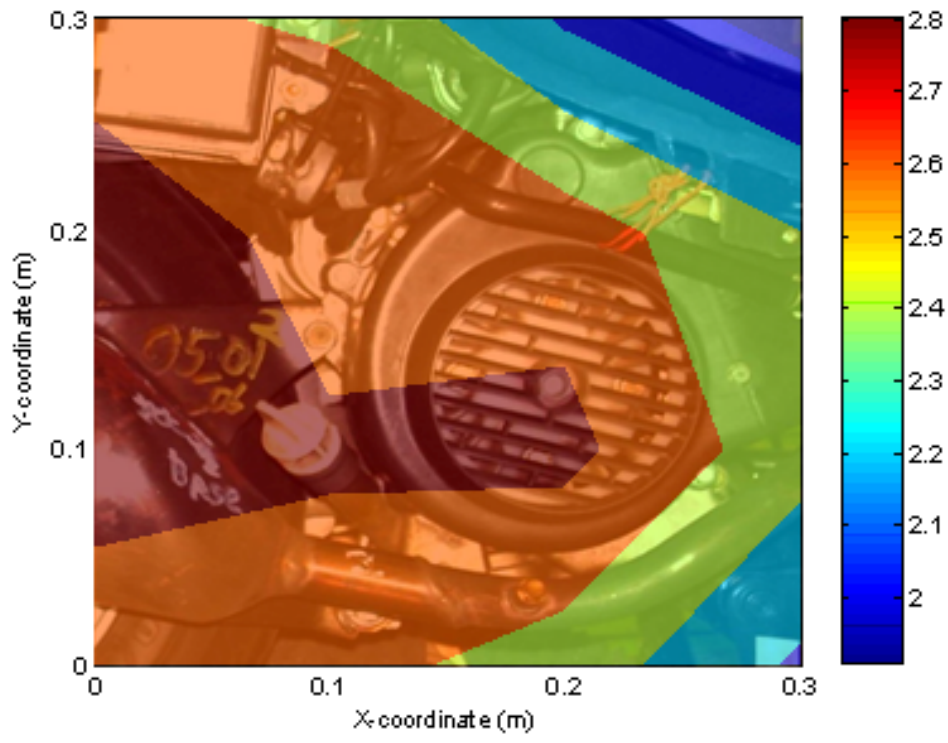


(b)

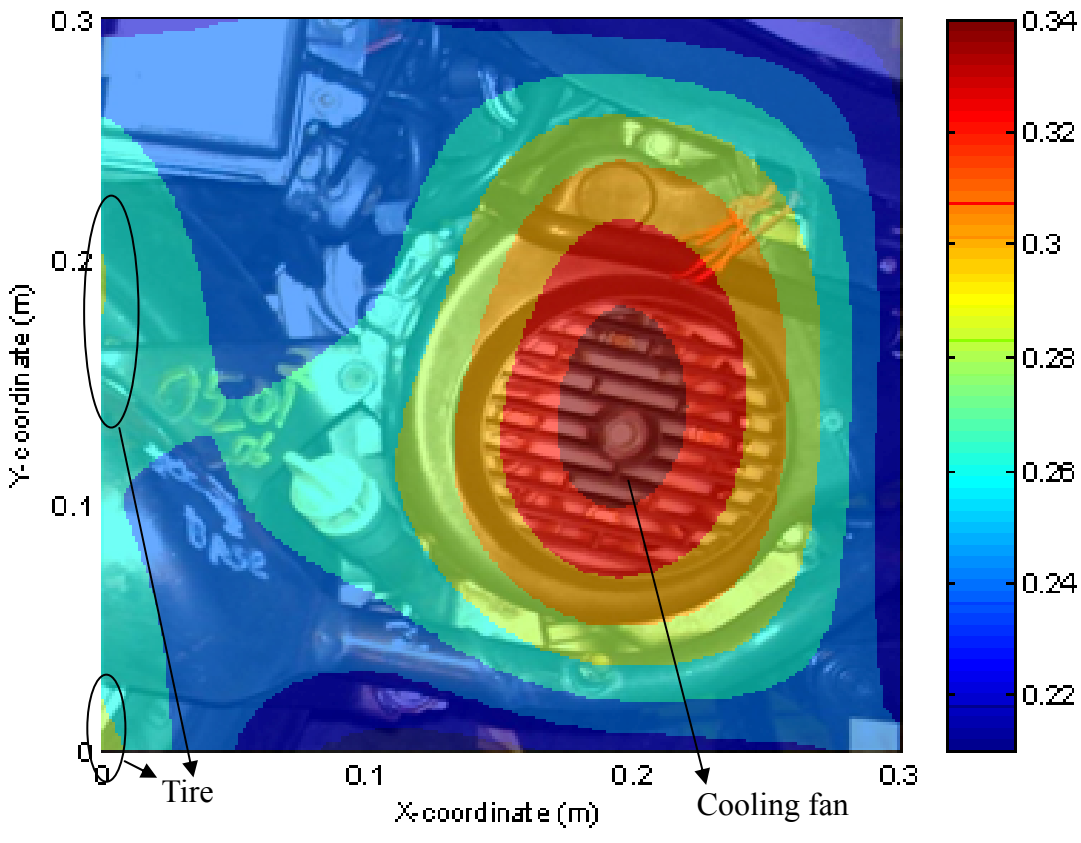


(c)

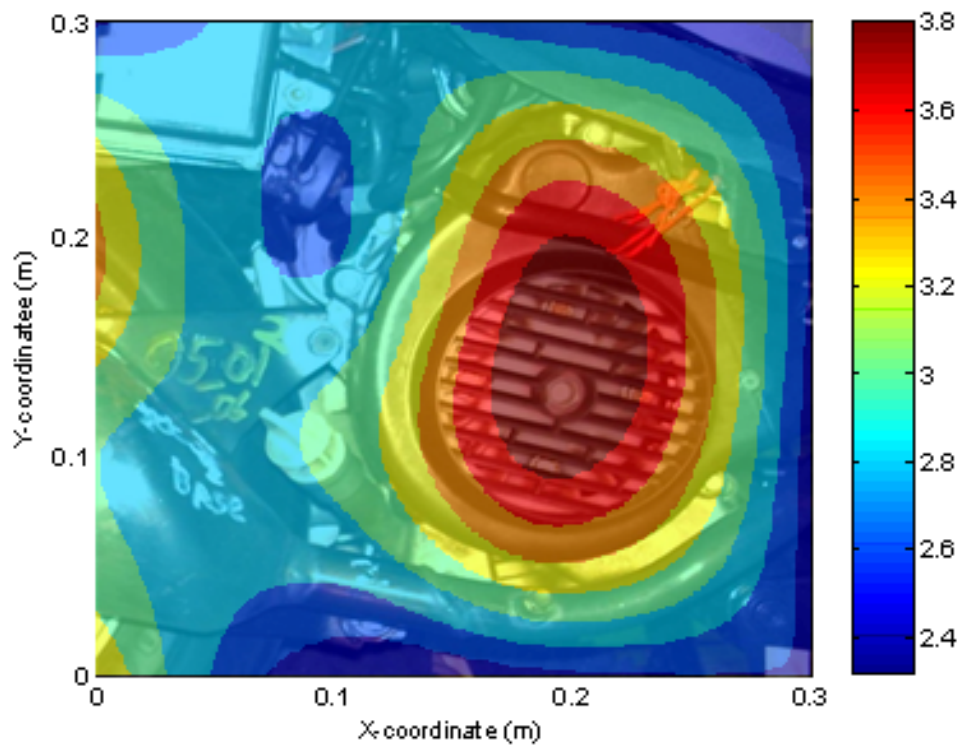
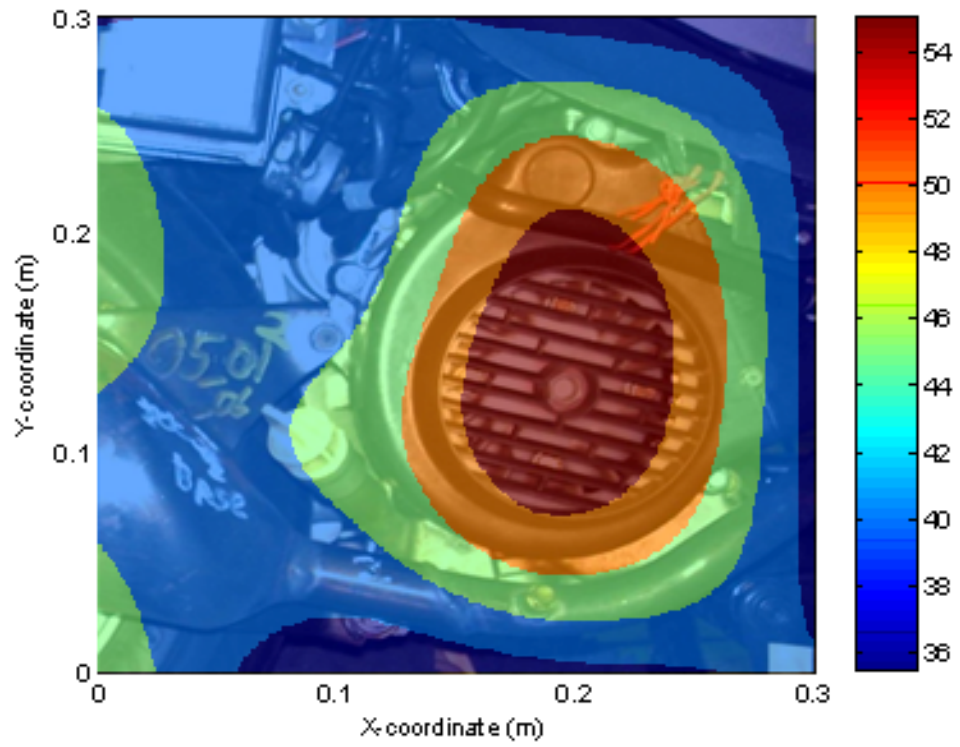
Fig. 88 The ERA technique. (a) The singular value plot of $\mathbf{H}(0)$, (b) the frequency responses of the inverse filter C_{13} of the original filter and the filter regenerated using ERA with $\nu = 80$, (c) the impulse responses of the inverse filter C_{13} of the original filter and the filter regenerated using ERA with $\nu = 80$.



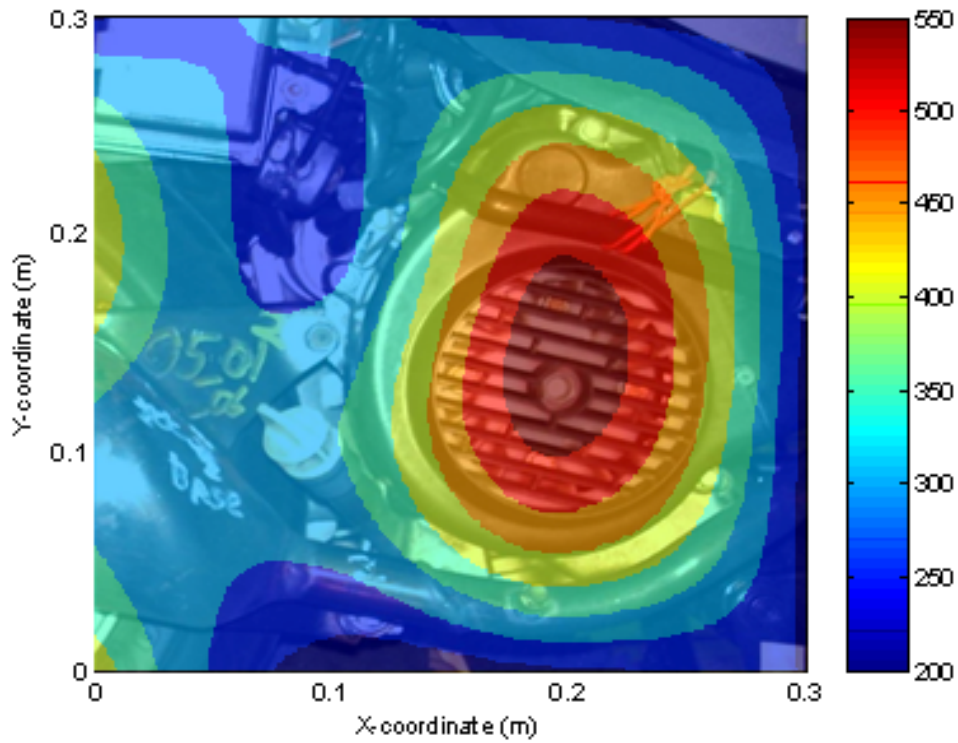
(a)



(b)

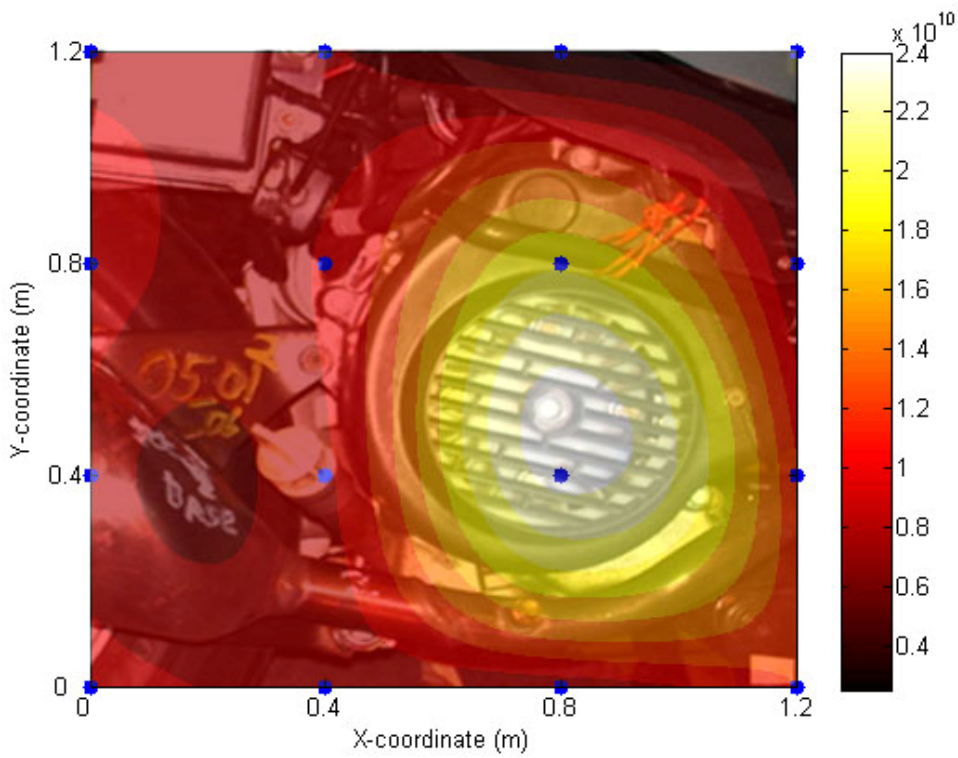
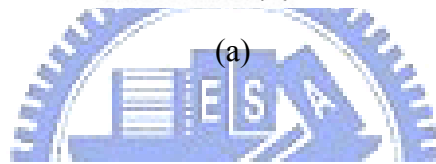
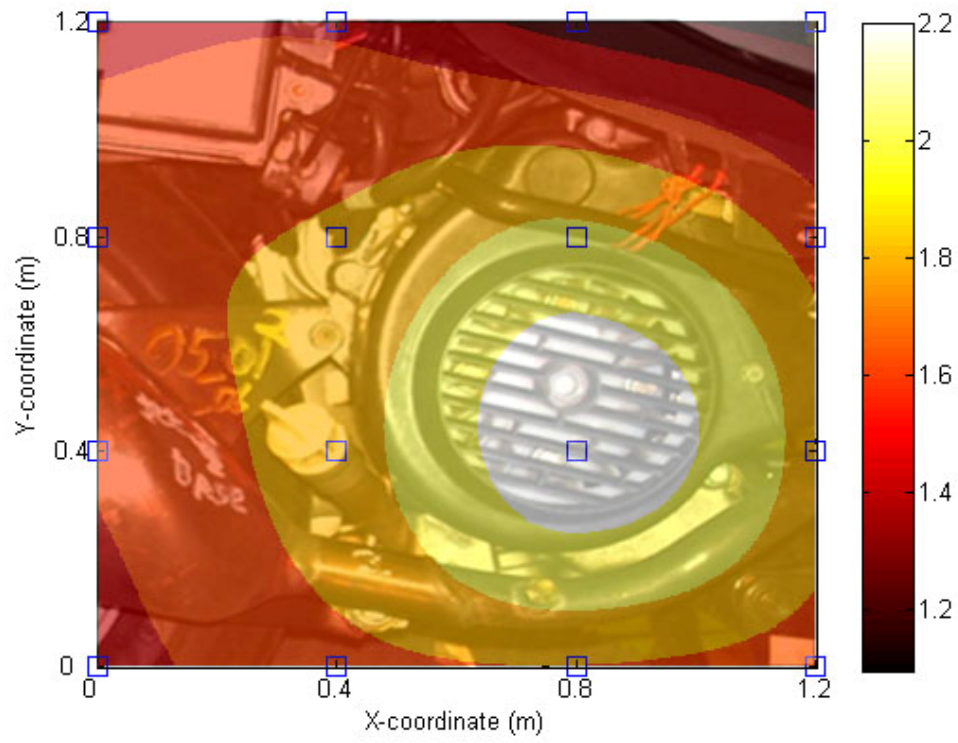


(d)

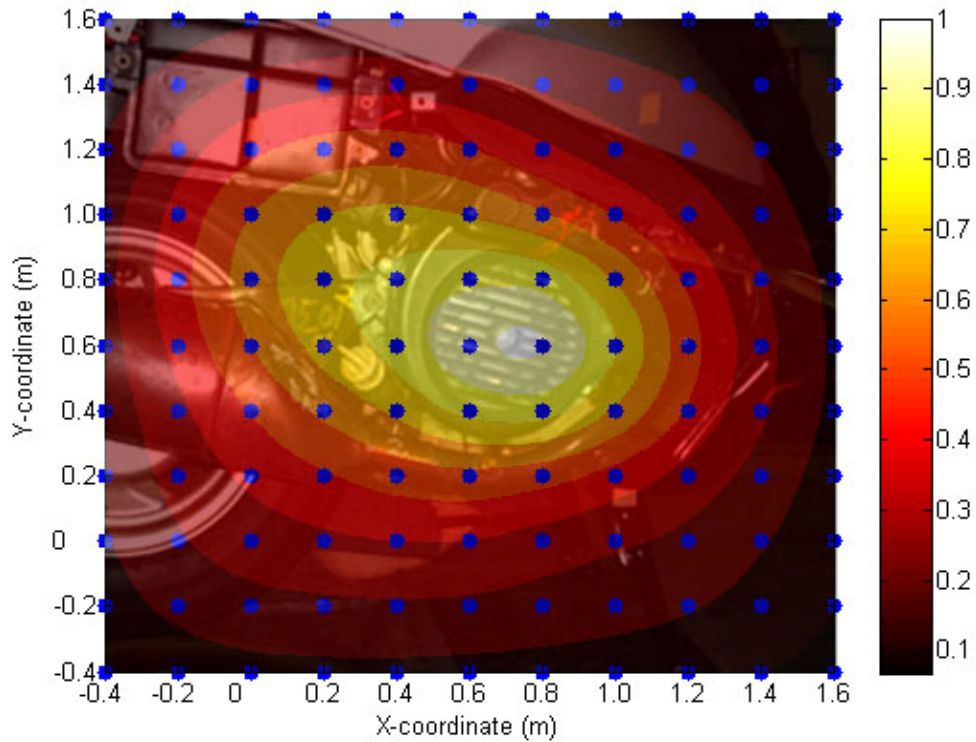


(e)

Fig. 89 The scooter experiment. The NESI was applied to reconstruct the sound field on the right side of the scooter running at the idle speed. (a) The unprocessed rms sound pressure image received at the microphones, (b) the rms source strength image, (c) the reconstructed rms sound pressure image, (d) the reconstructed rms particle velocity image, (e) the reconstructed rms sound intensity image



(b)

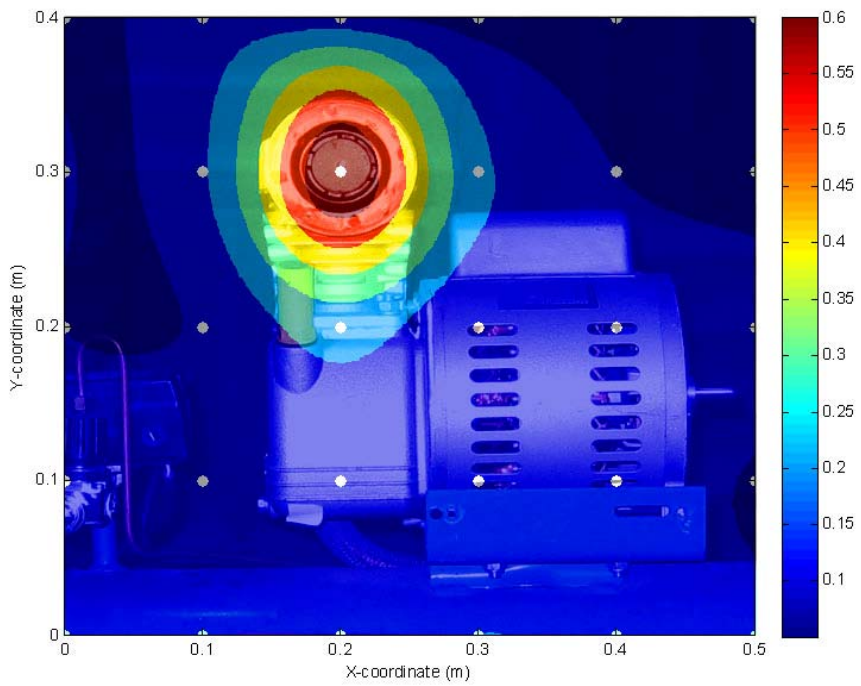
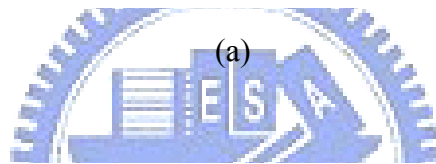
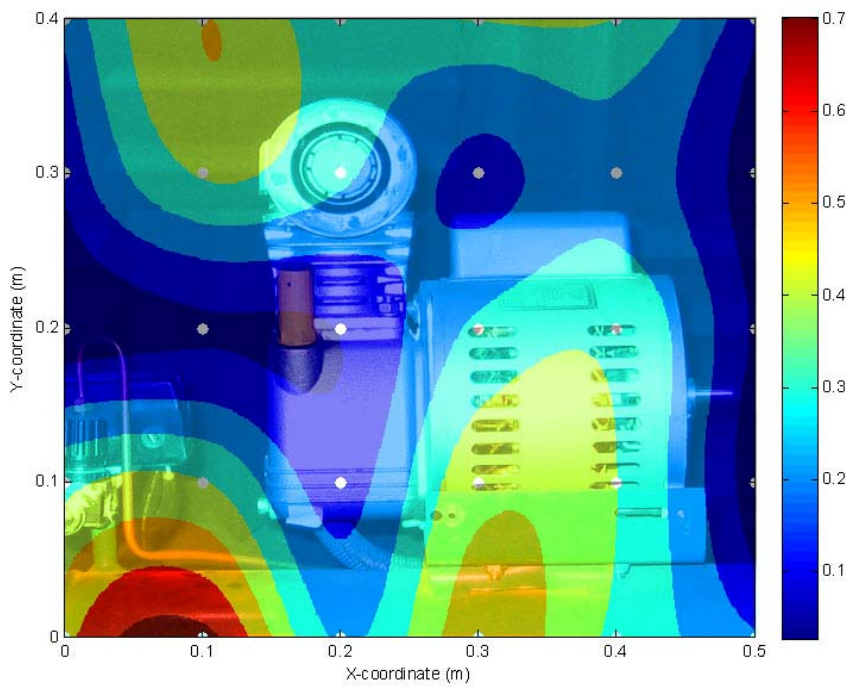


(c)

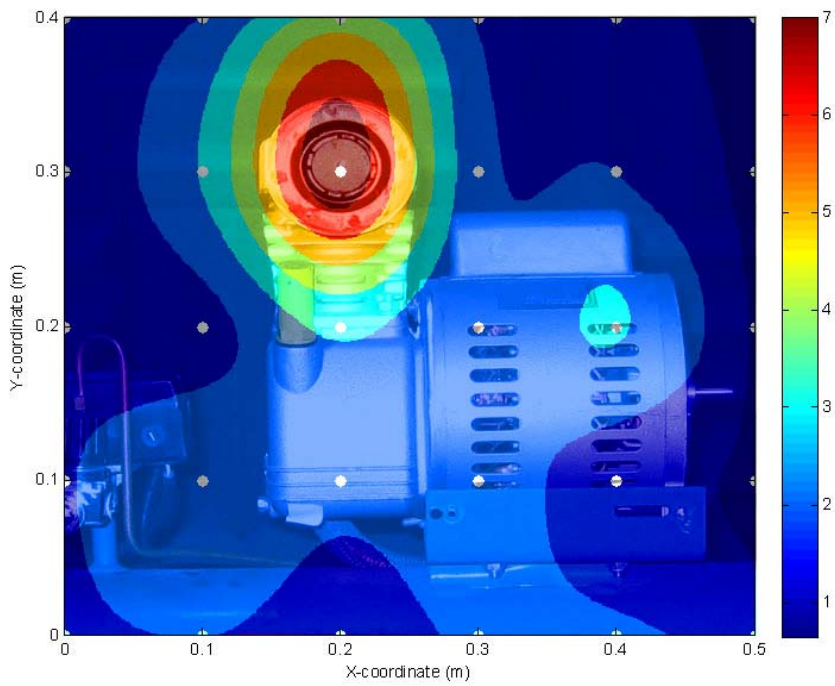
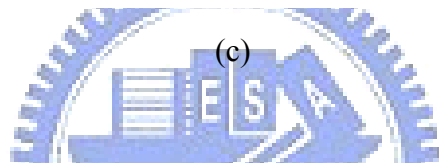
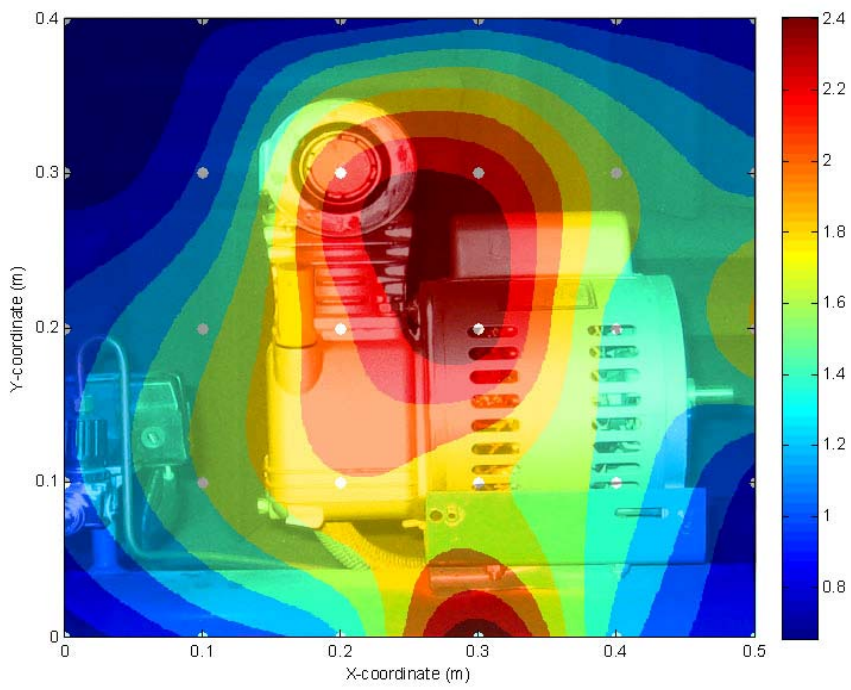
Fig. 90 The results of run-up experiment obtained using NESI with the 4×4 URA. The scooter engine was accelerated from 1500 rpm to 7500 rpm within ten seconds. (a) The unprocessed sound pressure image received at the microphones, (b) the reconstructed active intensity image, (c) the reconstructed active intensity image using the virtual microphone technique. The symbol“□” indicates the microphones. The symbol“•” indicates the focal points.

7.2 Compressor

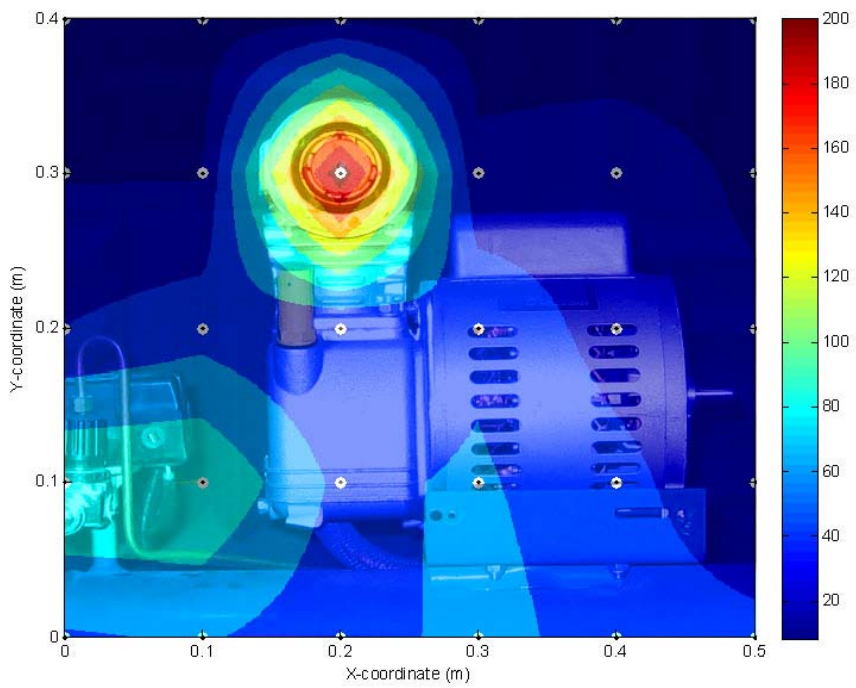
In the experiment, a compressor served as a practical source to examine the capability of the six algorithms. The compressor is mounted on a table inside a semi-anechoic room that the major noise is at the air intake position on the top of the compressor and the minor noise is low intensity vibration at the entire body. Different with loudspeaker experiment, the source of this experiment is not on the planar surface. The observed frequencies in the algorithms are chosen to be 1.2 kHz. The noise images obtained by processing of the six algorithms with URA are shown in Fig. 91 (a)-Fig. 91 (f). From Fig. 91 (a), Fourier NAH has a terrible noise source distribution, consistent with the theory that the source should be planar in Fourier NAH to identify successfully. From the reconstructed sound pressure of NESI shown in Fig. 91 (b), NESI can identify the major source at the air intake and the vibration at overall body. The result of DAS is bad by wrong location and very big main lobe but TR provide an acceptable result as shown in Fig. 91 (c)-Fig. 91 (d). In the noise images of MVDR and MUSIC, they identified the noise source at the air intake accurately and the result of MUSIC is aim at the major source.



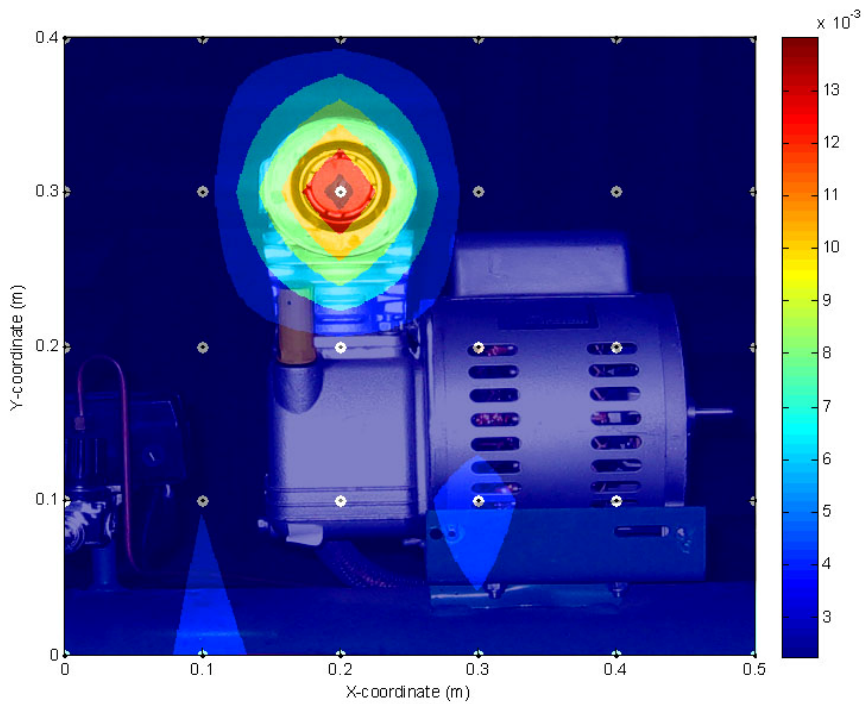
(b)



(d)



(e)



(f)

Fig. 91 The results of compressor experiment obtained using the 5×6 URA. The major noise is at the air intake position situated at (0.2m, 0.3m). The

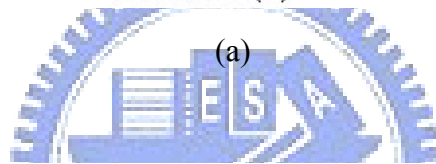
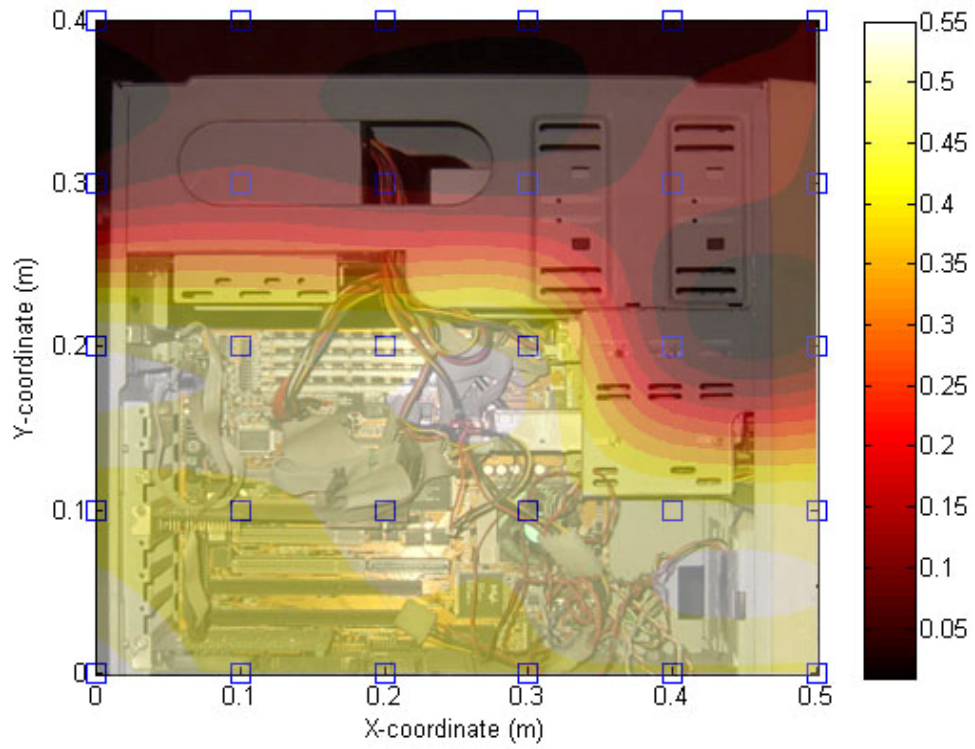
observed frequencies in the algorithms are chosen to be 1.2 kHz. (a) The reconstructed sound pressure image by Fourier NAH, (b) the reconstructed sound pressure image by NESI, (c) the source image obtained by using DAS, (d) the source image obtained by using TR, (e) the source image obtained by using MVDR, (f) the source image obtained by using MUSIC.



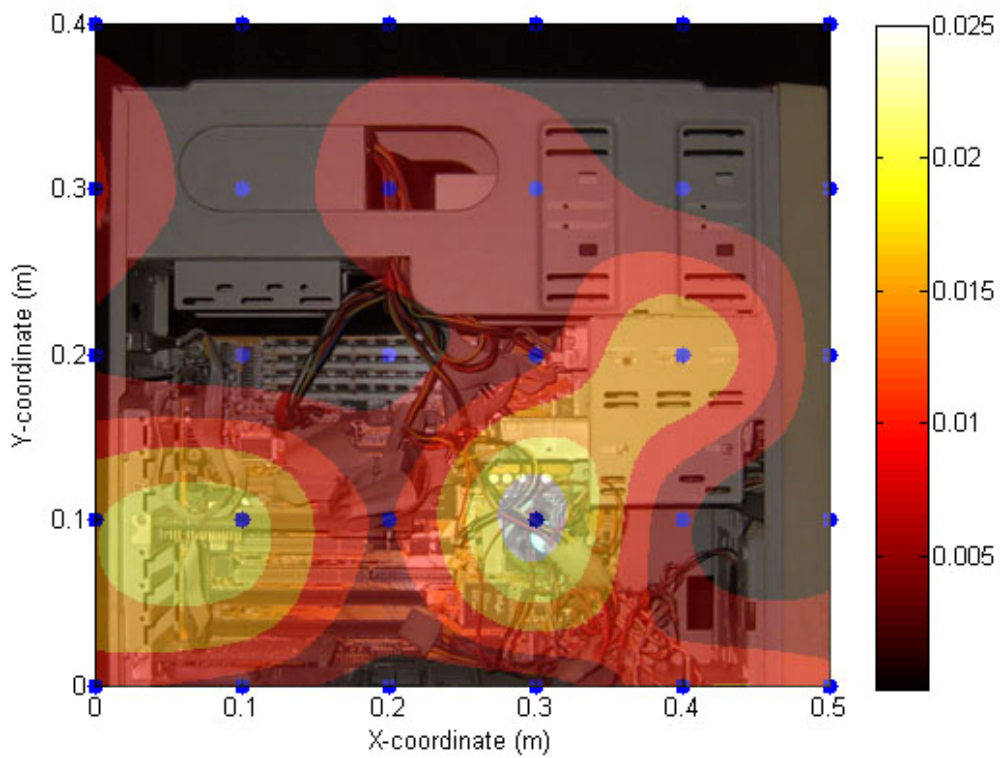
7.3 IT equipment

In this experiment, a desktop computer is used to validate the NESI technique by using a 5×6 URA. The microphone spacing d is selected to be 0.1 m ($\lambda/2$ corresponding to $f_{\max} = 1.7$ kHz).. The multichannel filtering is performed using ERA with $\nu=100$. The unprocessed sound pressure picked up at the microphones within the band 1200 ~ 1300 Hz is shown in Fig. 92 (a). This figure would lead to the incorrect conclusion that the noise source was located at the middle bottom. The active intensity was then reconstructed using the NESI, as shown in Fig. 92 (b) within the band 1200 ~ 1300 Hz. The bright areas on the intensity plot revealed that the power fan located at (0m, 0.3m), the metal plate located at (0m, 0.1m), the electric wire located at (0.3m, 0.1m), the floppy disk drive located at (0.4m, 0.2m) and the cabin mount located at (0.4m, 0m) are the major sources. Total sound power level is 90 dB re. 1×10^{-12} W. The NESI images apparently yielded more reliable information about noise sources than the unprocessed sound pressure.

To conclude this section, an experiment was undertaken to verify the conclusion obtained in the previous numerical simulation by the authors. The active intensity reconstructed using a random array optimized for farfield imaging within the band 1200 ~ 1300 Hz is shown in Fig. 93. Except that the power fan (0m, 0.3m) and the cabin mount (0.4m, 0m) in the desktop computer were correctly identified as the major noise sources in Fig. 93, the other sources previously identified by the URA in Fig. 92 were largely missed. This attests the conclusion drawn from the numerical simulation that random deployment offers little advantages for nearfield imaging and the URA is the optimal array configuration.



(a)



(b)

Fig. 92 The noise map within the band 1200 ~ 1300 Hz obtained using the 5×6

URA for the desktop computer. (a) The unprocessed sound pressure image received at the microphones, (b) the active intensity image reconstructed using NESI. The symbol“•” indicates the focal points.



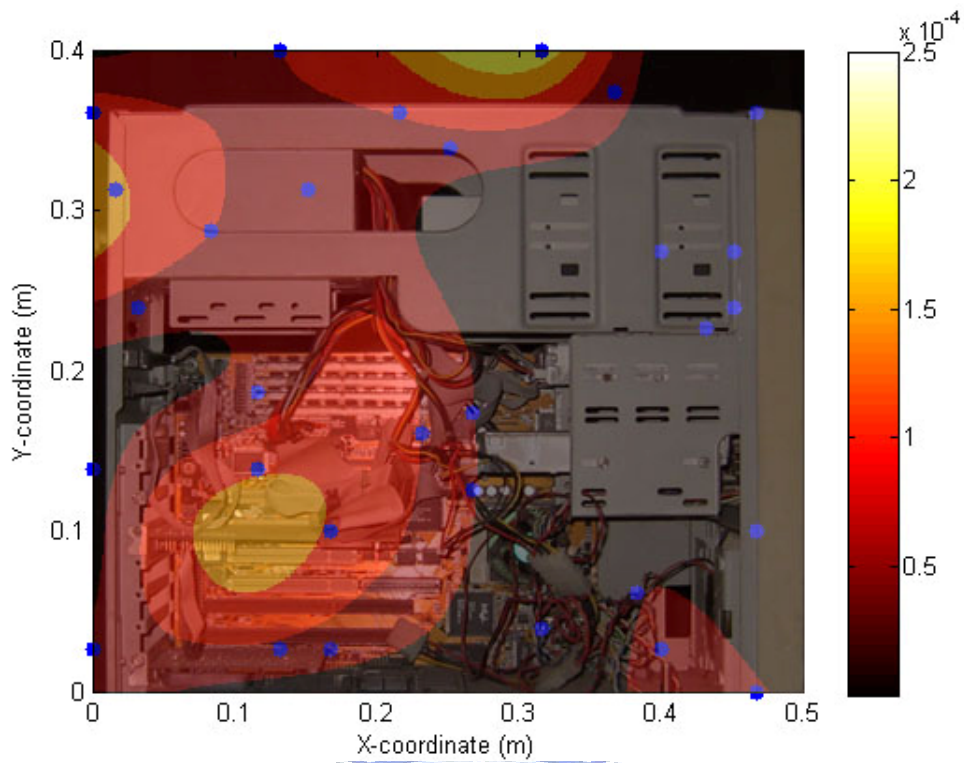
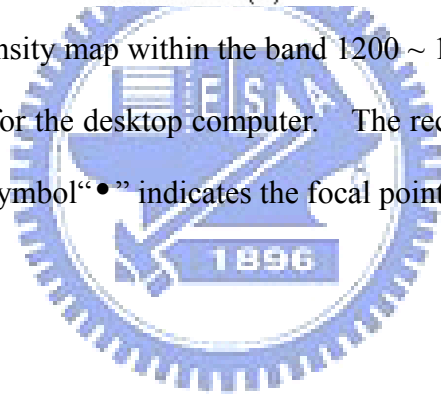


Fig. 93 The active intensity map within the band 1200 ~ 1300 Hz obtained using the random array for the desktop computer. The reconstructed active intensity image. The symbol“•” indicates the focal points.



7.4 Wooden box

In this experiment, a wooden box model with loudspeaker fitted inside is used to validate the NESI technique by using a 5×6 URA. As shown in Fig. 94, several holes with different shapes are cut in the front face of the box like a Jack_O_Lantern. A circle, two squares, and a slit are located at $(0.5\text{m}, 0.4\text{m})$, $(0\text{m}, 0.4\text{m})$, $(0.25\text{m}, 0.25\text{m})$ and $(0.25\text{m}, 0\text{m})$, respectively. The loudspeaker produces random noise band-limited to 1.7 kHz. The microphone spacing d is selected to be 0.1m ($\lambda/2$ corresponding to $f_{\max} = 1.7$ kHz).

The unprocessed sound pressure picked up at the microphones within the band 200~1600 Hz is shown in Fig. 95 (a). From the image, the noise sources were barely resolvable, particularly for the noise source at the edge - the circle, the slot and the square at upper left corner. Also, the square at the center was difficult to distinguish. Virtual microphone technique was again applied to overcome this problem by interpolate and extrapolate the pressure field on the microphone surface and increase the number of microphones and focal points from 5×6 to 13×15 . With the new setting, the particle velocity (rms) reconstructed using the NESI is shown in Fig. 95 (b). It can be clearly observed from the result that the quality of the reconstructed image was significantly improved. Problems due to edge effect and insufficient resolution were basically eliminated.

The NESI images apparently yielded more reliable information about noise sources than the unprocessed sound pressure.

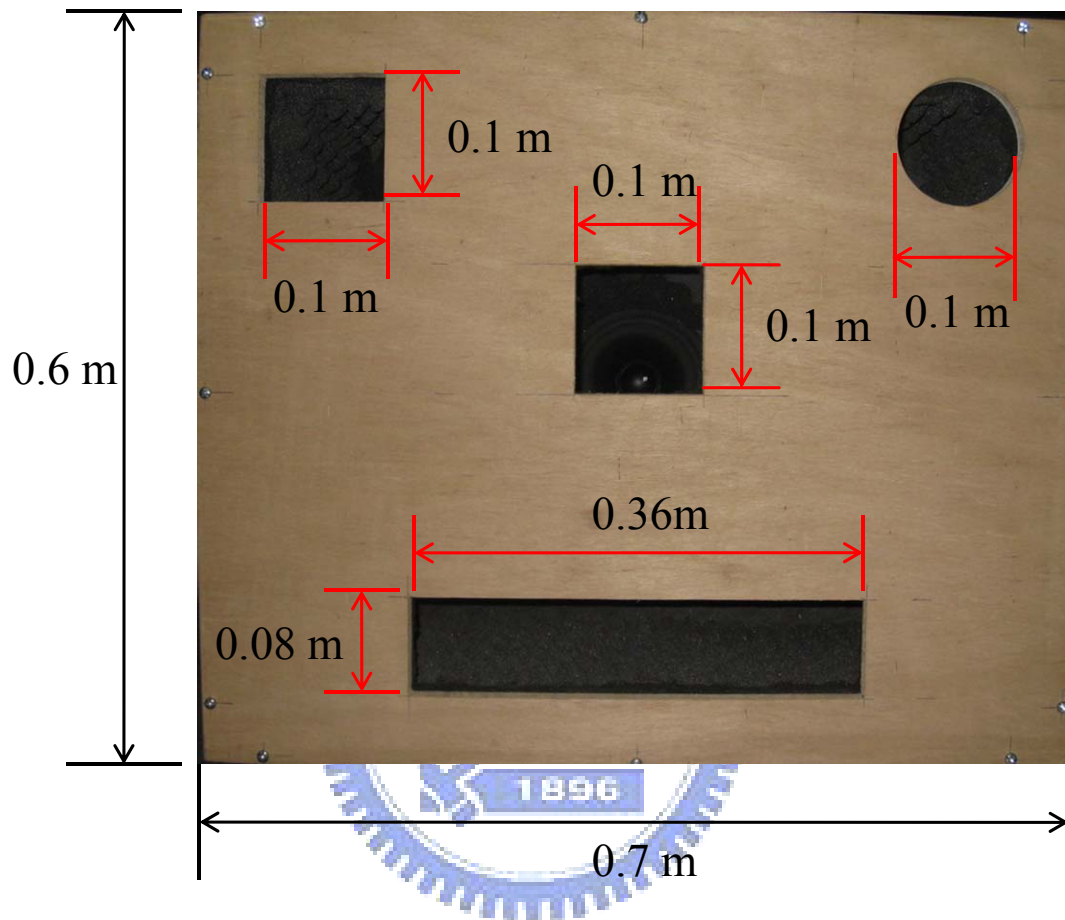
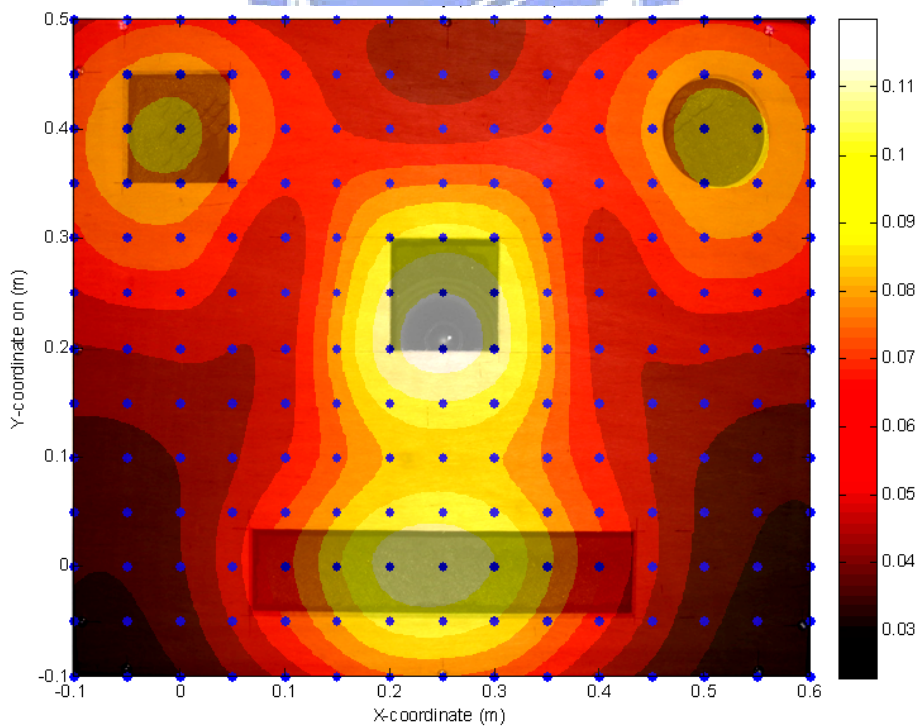
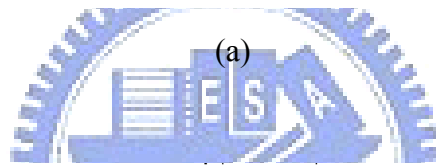
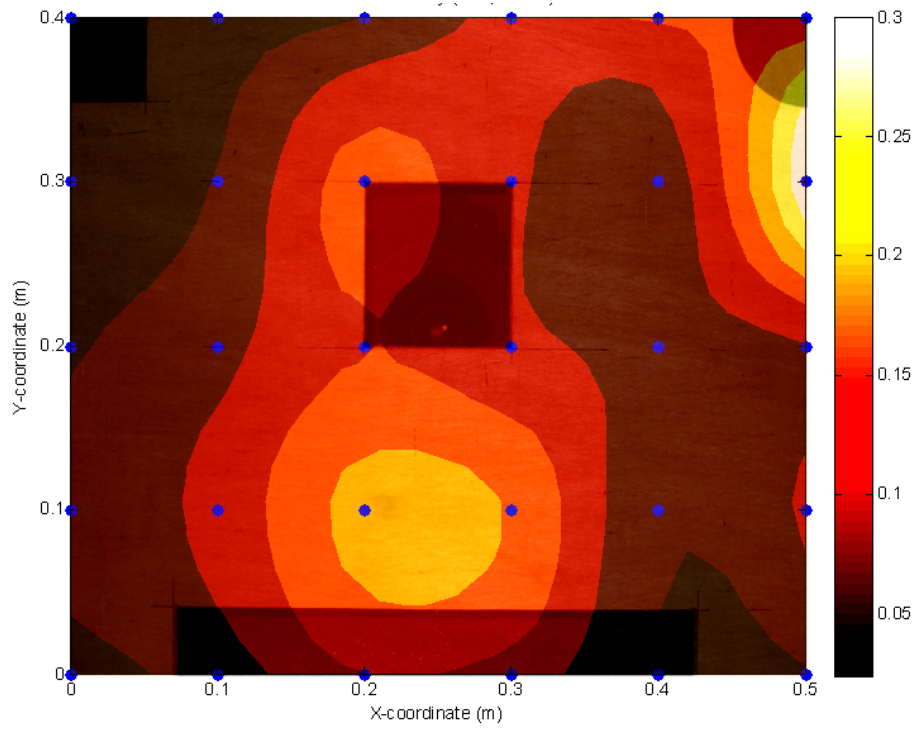


Fig. 94 The experimental arrangement for a wooden box model with a loudspeaker fitted inside and holes cut on the frontal surface.



(b)

Fig. 95 The results of a wooden box with a loudspeaker fitted inside. The noise map

is within the band 200 Hz ~ 1.6k Hz. (a) The unprocessed sound pressure image received at the microphones by 5×6 URA, (b) the particle velocity image reconstructed using NESI by the 5×6 URA.



7.5 Non-contact modal analysis

An interesting application of the NESI is non-contact modal analysis of vibrating structures. The mode shapes of the structure can be calculated by reconstructing the surface velocity without having to mount accelerometers on the surface as usually so in the traditional modal analysis. An experiment is undertaken in an anechoic room to validate the ESM NAH in reconstructing the surface velocity of a free-edged aluminum ($0.2\text{m} \times 0.2\text{m} \times 0.002\text{m}$). The experimental arrangement is shown in Fig. 96. The plate was driven at the resonant frequency 594 Hz. The apertures of the 9×11 URA and the plate are identical, with the microphone spacing $dx = 0.02\text{m}$ and $dy = 0.025\text{m}$ in x and y directions, respectively. The array is positioned 0.02m above the plate. The velocity reconstructed by the ESM NAH was compared to the surface velocity measured using a scanning laser vibrometer (PSV-400, Polytec). In addition, the Chladni pattern (mode shape) at the same frequency was obtained by a salt sprinkle test, as shown in Fig. 97. The velocity distribution of the plate measured by the scanning laser is shown in Fig. 98 (a). The frequency-domain magnitudes of the surface velocity reconstructed using various RDs, $1/20d$, $1/2d$, $1d$ and $2d$ are shown in Fig. 98 (b)-Fig. 98 (e). By inspecting the velocity patterns in Fig. 98, the best match is evidenced between the patterns in Fig. 98 (a) and Fig. 98 (c) which is obtained using $d/2$ RD, whereas the reconstructed patterns in Fig. 98 (b), Fig. 98 (d) and Fig. 98 (e) differ drastically from the laser scanned result. This result is also consistent with the Chladni pattern shown in Fig. 97. The experimental results above suggest that the choice of RD is vital to reconstruction quality in the ESM NAH. To more precisely quantify the comparison of reconstruction performance, the velocity data reconstructed by the ESM using $d/2$ RD and the surface velocity data measured by the scanning laser are compared in Fig. 99 for all lattice points. The error metric is the relative velocity reconstruction error

defined in Eq. (6. 46). Vectors \mathbf{u} and \mathbf{u}_r represent the laser-measured and the ESM-reconstructed velocity vectors, respectively. The relative errors of the reconstructed velocity thus calculated are 100% for $RD = 1/20d$, 28.11% for $RD = 1/2d$, 71.47% for $RD = 1d$, and 72.97% for $RD = 2d$. The ESM using $RD = 1/2d$ has achieved far better (28.11%) surface velocity reconstruction than the other three RD settings. This substantiates our preceding conclusion that the optimal RD for reconstructing surface velocity of planar sources using the ESM is $d/2$, which is smaller than the conventional choice of 1-2 times spacing.





Fig. 96 Experimental arrangement for an aluminum plate.

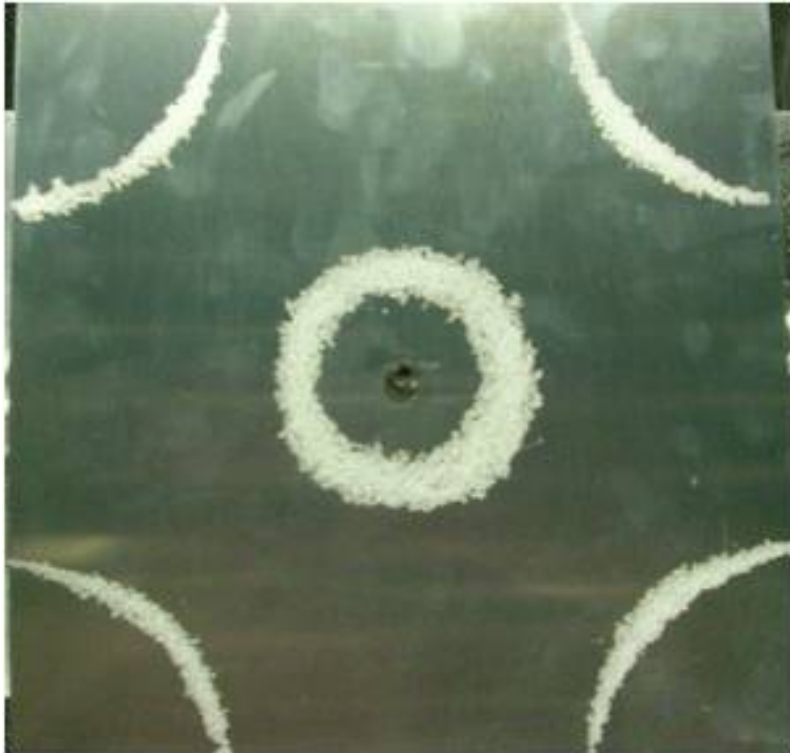
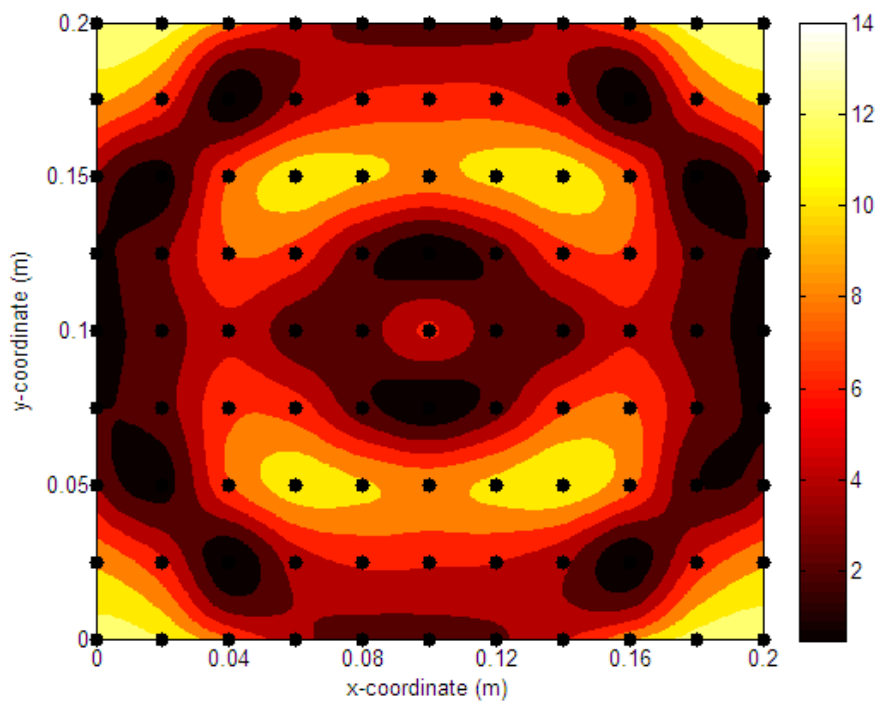
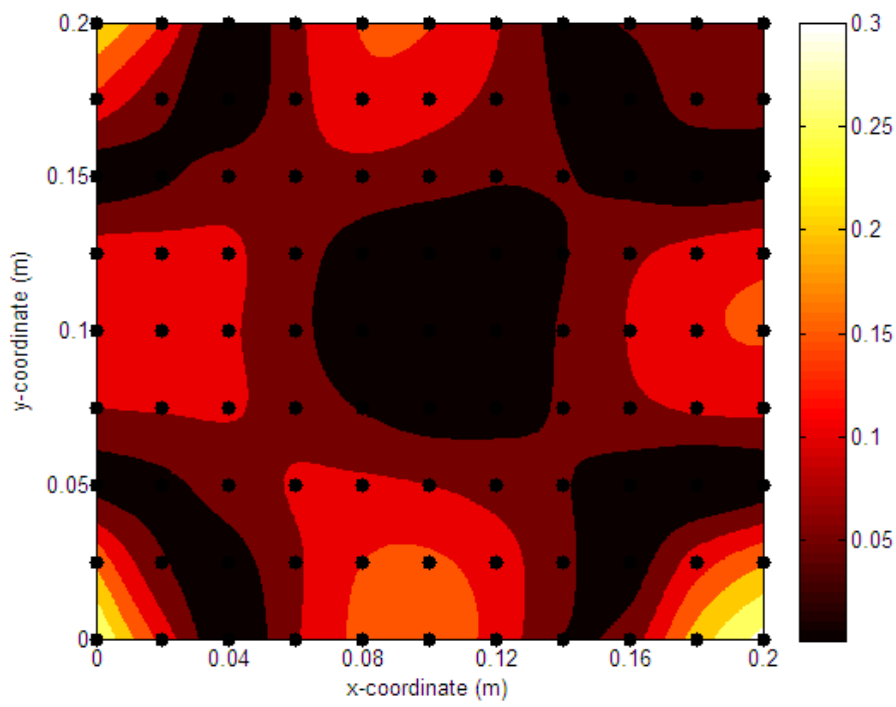
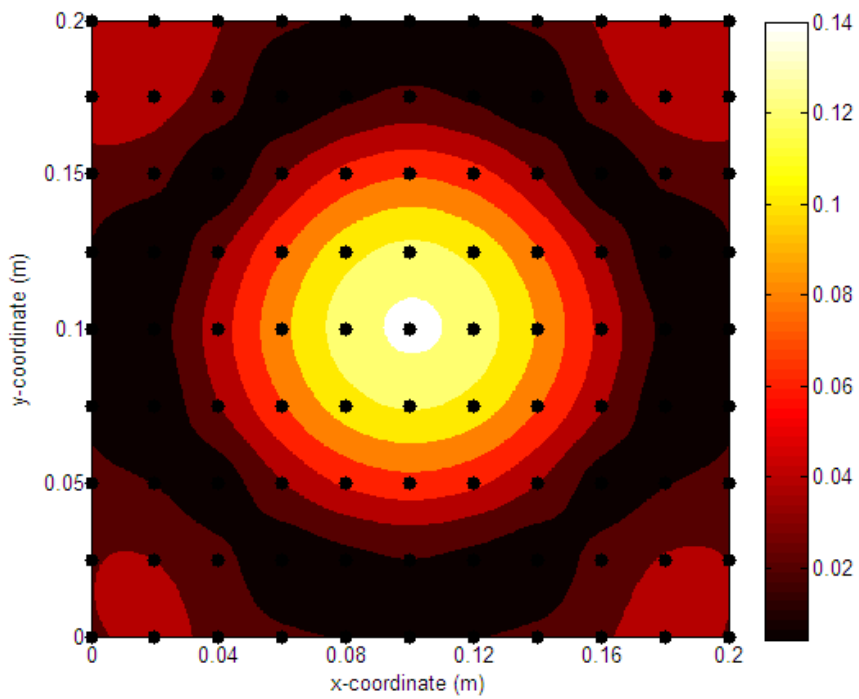
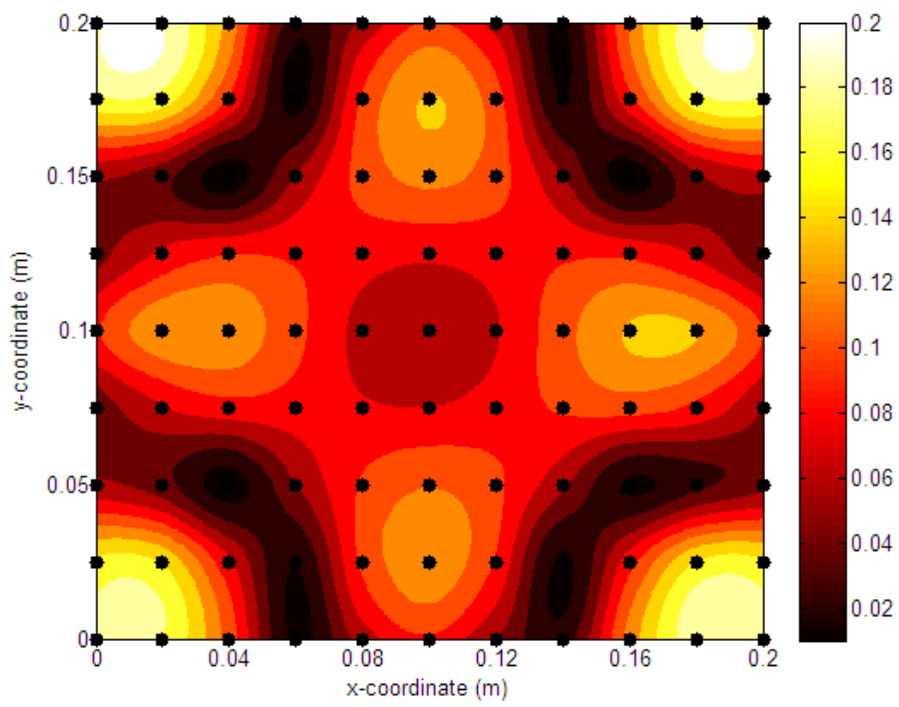


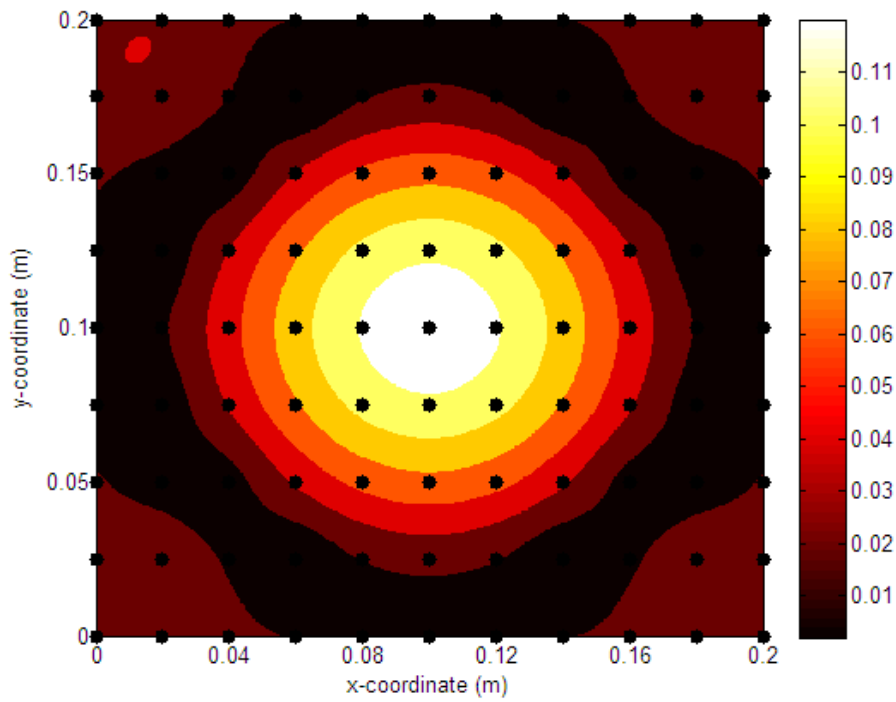
Fig. 97 Mode shape at the resonant frequency 594 Hz obtained using a salt sprinkle test.



(b)



(d)



(e)

Fig. 98 The surface velocity of the plate. (a) The velocity distribution of the plate measured by the scanning laser. The surface velocity reconstructed using various RDs (b) $1/20d$, (c) $1/2d$, (d) $1d$ and (e) $2d$. The microphones are indicated in the figure using black dots.

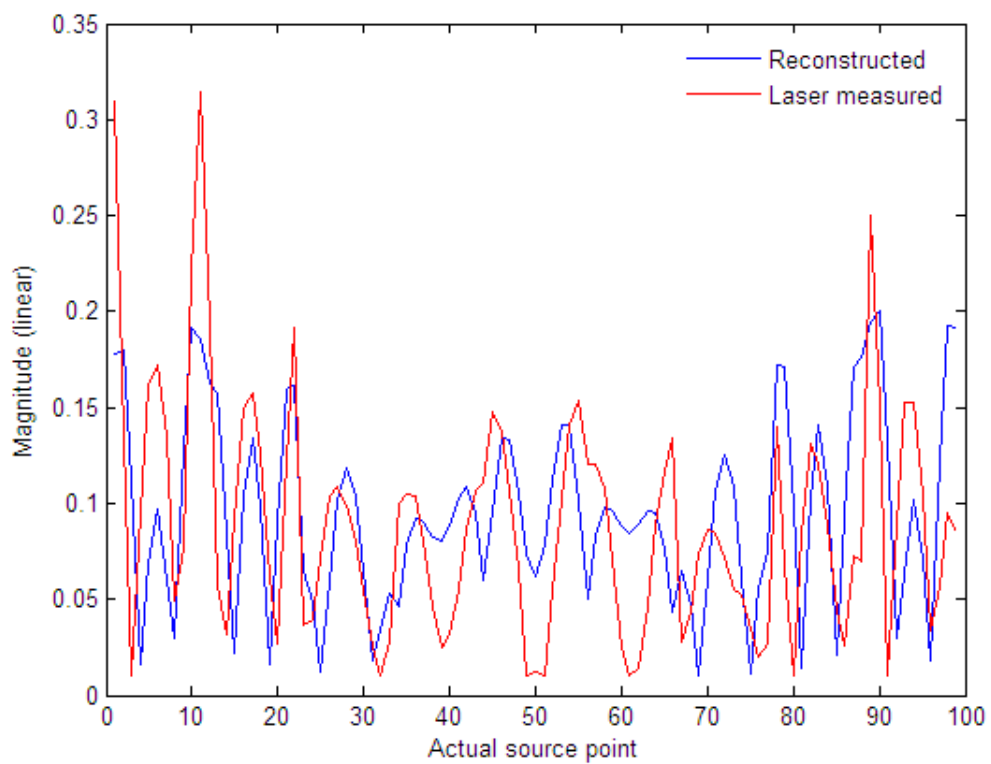


Fig. 99 The velocity data reconstructed by the ESM using $d/2$ RD and the surface velocity data measured by the scanning laser are compared for all lattice points.

7.6 Other applications of the ESM

The ESM applies not only to NSI but also other applications. In automotive hands-free system, the Single input multiple output equivalent source inverse filtering (SIMO-ESIF) algorithms are implemented to reconstruct the speech signal in a reverberant environment [83]. Specifically, the system serves two purposes: dereverberation and noise reduction. In order to further improve the noise reduction performance in spatial filtering and robustness against system uncertainties, the SIMO-ESIF algorithm is combined with an adaptive Generalized Side-lobe Canceller (GSC). As indicated by several performance measures in noise reduction and speech distortion, the proposed microphone array algorithm proved effective in reducing noise in human speech without significantly compromising the speech quality.

In application of reverberation [84], a synthesis technique is developed using discrete boundary source representation in an attempt to overcome these difficulties. In comparison with FEM and BEM that rely upon complex numerical operations, the ESM method is based on simple representation of the sound field with a distribution of discrete simple sources on the boundary. The filtering property of human hearing is also exploited in a non-uniform sampling procedure to further simplify the computation. Subjective listening experiments also demonstrate that the ESM technique is capable of conferring remarkable realism of reverberation.

In application of modeling head related transfer functions (HRTF) [85], the HRTFs for a given head with pinnae are calculated with minimal amount of computation using ESM. In the process, a special regularization scheme is required to calculate the equivalent strengths of virtual sources. The HRTFs obtained using ESM agrees reasonably well in terms of frequency response, directional response and impulse response with the other methods. The ESM obviates the singularity problem as commonly encountered in the BEM, and is less computationally demanding than

the BEM in terms of time and memory usage.

In dodecahedral loudspeaker, the ESM is employed to calculate the pressure radiation pattern of the source. The aim is design a dodecahedral loudspeaker source to approximate an omni-directional source. The ESM also find applications in Ambisonics, supersensitive array, p-v-based array and superdirective optimized array, and so forth.



CHAPTER 8. CONCLUSIONS

This thesis presents a comprehensive overview of microphone array technology encompassing all aspects from the underlying theory to implementation. Farfield and nearfield sound image have been addressed in this thesis. In the farfield sound image, there have two major researches in this thesis. First, optimized planar array deployment for source imaging is investigated in this thesis. Global optimization algorithms have been developed to facilitate the search of the optimized microphone deployment. The SA algorithm and the combined SA-IBMC algorithm prove effective in finding the optimal deployment. For far-field array with sparse deployment in which inter-element spacing is large, random deployment with optimal weights is crucial to avoid grating lobes. As predicted by the conventional wisdom, the optimized random sparse array has excellent beam pattern with a unique main-lobe. Second, several acoustic imaging algorithms including DAS, TR, MVDR, MUSIC and an inverse filter-based method SIMO-ESIF have been developed to estimate DOA. The resolutions of noise maps in low frequency are much worse than in high frequency with random array configuration. The proposed SIMO-ESIF approach can use URA to estimate DOA in high frequency without grating lobes problem. As expected, the high resolution methods such as MVDR and MUSIC can obtain much greater results than DAS, TR and SIMO-ESIF in localizing sound source positions.

In the nearfield sound image, optimized nearfield array deployment for source imaging is investigated. The outcome of the work is twofold. First, global optimization algorithms have been developed to facilitate the search of the optimized microphone deployment. Second, whether or not randomization is necessary in nearfield imaging is explored, with the aid of the above optimization techniques. The hybrid SA-IBMC algorithm proves effective in finding the optimal deployment.

To answer the question whether the idea of random deployment acquired in farfield imaging applies to nearfield imaging as well, simulation employing the MC, IBMC, SA and SA-IBMC optimization techniques were conducted. Due to its complicated nearfield and multi-focusing nature, a special kind of beam-pattern alongside the cost function is defined. It was observed with all the adopted optimization approaches that the cost function keeps fluctuating and shows no sign of convergence. The maximum cost function is always the initial cost function corresponding to the URA deployment. Random deployment seems to present no particular benefit for nearfield imaging and the optimal array is the URA. The NESI enables effective identification of noise sources based on sound pressure, particle velocity, and intensity images. Inverse filters are design using the least-squares optimization with the aid of Tikhonov regularization. Criteria for choosing array parameters are summarized. In particular, the DOR is always selected to render the condition number of propagation matrix below 10^3 . The window design is employed to alleviate boundary defocusing problem. Singularity problem is circumvented by using a retreat focal point technique. The state-space form obtained using the ERA is exploited to enhance computation efficiency for real-time implementation. The NESI proves effective in the identification of broadband random and transient noise sources. In addition, the NESI does not need as large channel count as the NAH approach. Experimental investigations have been undertaken to verify the proposed implementation technique. It is observed from the results that the practicality of NESI has been considerably enhanced by the use of the proposed techniques.

Various implementation issues of the NESI technique which is a time-domain ESM have been investigated in chapter 6 and 7. A virtual microphone technique is employed to minimize edge effects using extrapolation and to improve imaging resolution using interpolation when only patch array with scarce sensors are available.

Numerical simulations suggest that the optimal distance ranges from 0.4 to 0.5 times spacing for planar sources, whereas the optimal distance can vary from 0.8 to 1.7 times average spacing for the spherical source. Although the optimal distance is not a unique value and may well depend on many factors, the contribution of this work is to provide an effective methodology for finding the optimal distance for ESM applications. Experiments were also carried out for a vibrating aluminum plate to validate the optimization results. It appears that the optimal RD is less than that predicted by the 1-2 spacing rule except some “outlier” cases (such as 1.7 spacing in the $\theta_0 = 180^\circ$ case of the spherical piston example). From the compressor experiment, the six algorithms are compared on the point of resolution and performance. Both Fourier NAH and NESI have good performance while NESI is more robust than Fourier NAH for it can reconstruct the sound field from the source of arbitrary shape. As expected, the high resolution methods such as MVDR and MUSIC can obtain greater results than DAS and TR in localizing the source position. Although the resolution of MVDR and MUSIC are better than Fourier NAH and NESI but the performance of Fourier NAH and FDNESI are better. Most important of all, Fourier NAH and NESI can reconstruct the acoustic variables such as sound pressure, particle velocity and active intensity.

REFERENCES

- [1] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques*, Prentice-Hall, Englewood Cliffs, NJ (1993).
- [2] G. H. Koopmann and J. B. Fahnlne, *Designing Quiet Structures*, Academic Press, San Diego (1997).
- [3] M. Brandstein and D. Ward, *Microphone Arrays Signal Processing Techniques and Applications*, Springer-Verlag, NY (2001).
- [4] J. Benesty, J. Chen and Y. Huang, *Microphone Array Signal Processing*, Springer-Verlag, NY (2008).
- [5] H. L. Van Trees, *Optimum Array Processing*, Wiley & Sons, Inc., NY (2002).
- [6] B. A. Cray, V. M. Evora, and A. H. Nuttall, "Highly directional acoustic receivers," *Journal of the Acoustical Society of America*, Vol. 113, pp. 1526-1532 (2003).
- [7] Y. Fan, B. Tysoe, J. Sim, K. Mirkhani, A. N. Sinclair, F. Honarvar, H. Sildva, A. Szecket and R. Hardwick, "Nondestructive evaluation of explosively welded clad rods by resonance acoustic spectroscopy," *Ultrasonics*, Vol. 41, pp. 369-375 (2003).
- [8] M. Duquennoy, M. Ouaftouh and M. Ourak, "Ultrasonic evaluation of stresses in orthotropic materials using Rayleigh waves," *NDT & E International*, Vol. 32, pp. 189-199 (1999).
- [9] V. Murino, "Reconstruction and segmentation of underwater acoustic images combining confidence information in MRF models," *Pattern Recognition*, Vol. 34, pp. 981-997 (2001).
- [10] D. Zha and T. Qiu, "Underwater sources location in non-Gaussian impulsive noise environments," *Digital Signal Processing*, Vol. 16, pp. 149-163 (2006).
- [11] U. Benko, J. Petrovic, D. Juricic, J. Tavcar, J. Rejec and A. Stefanovska, "Fault

- diagnosis of a vacuum cleaner motor by means of sound analysis,” *Journal of Sound and Vibration*, Vol. 276, pp. 781-806 (2004).
- [12] J. D. Wu and C. Q. Chuang, “Fault diagnosis of internal combustion engines using visual dot patterns of acoustic and vibration signals,” *NDT&E International*, Vol. 38, pp. 605-614 (2005).
- [13] B. D. V. Veen, K. M. Buckley, Beamforming: a versatile approach to spatial filtering, *IEEE Signal Processing Magazine*, Vol.5, pp. 4-24 (1998).
- [14] H. Krim, M. Viberg, Two decades of array signal processing research, *IEEE Signal Processing Magazine*, Vol. 13, pp. 67-94 (1996).
- [15] E. G. Williams and J. D. Maynard, Holographic imaging without the wavelength limit, *Physical Review Letters*, Vol. 45, pp. 554-557 (1980).
- [16] M. I. Skolnik, Introduction to Radar Systems, McGraw-Hill, NY, (1980).
- [17] A. V. Oppenheim and R. W. Schaffer, Discrete-Time Signal Processing, Prentice-Hall, Englewood Cliffs, NJ, 1989.
- [18] S. Haykin, Array Signal Processing, Prentice Hall, Englewood Cliffs, NJ, (1985).
- [19] H. T. Friis and C. B. Feldman, “A multiple unit steerable antenna for short-wave reception,” *Bell System Technical Journal*, Vol. 16, pp. 337-419 (1937).
- [20] J. J. Christensen and J. Hald, “Beamforming,” Brüel & Kjær Technical Review No. 1 (2004).
- [21] J. Capon, “High-resolution frequency-wavenumber spectrum analysis,” *Proceedings of the IEEE*, Vol. 57, pp. 1408-1418 (1969).
- [22] R. O. Schmidt, “Multiple emitter location and signal parameter estimation,” *IEEE Transactions on Antennas and Propagation*, Vol. 34, pp. 276-280 (1986).
- [23] M. R. Bai and J. W. Lee, Industrial noise identification by using an acoustic beamforming system, *ASME, Journal of Sound and Vibration*, Vol. 120,

pp.426-433 (1998).

- [24] E. G. Williams, J. D. Maynard and E. Skudrzyk, "Sound source reconstructions using a microphone array," *Journal of the Acoustical Society of America*, Vol. 68, pp. 340-344 (1980).
- [25] M. R. Bai, "Application of BEM (boundary element method)-based acoustic holography to radiation analysis of sound sources with arbitrarily shaped geometries," *Journal of the Acoustical Society of America*, Vol. 92, pp. 533-549 (1992).
- [26] B. K. Kim and J. G. Ih, "On the reconstruction of vibro-acoustic field over the surface enclosing an interior space using the boundary element method," *Journal of the Acoustical Society of America*, Vol. 100, pp. 3030–3016 (1996).
- [27] S. C. Kang and J. G. Ih, "Use of non-singular boundary integral formulation for reducing errors due to near-field measurements in the boundary element method based near-field acoustic holography," *Journal of the Acoustical Society of America*, Vol. 109, pp. 1320-1328 (2001).
- [28] G. T. Kim and B. H. Lee, "3-D sound source reconstruction and field reproduction using the Helmholtz integral equation," *Journal of Sound and Vibration*, Vol. 136, pp. 245–261 (1990).
- [29] E. G. Williams, B. H. Houston, P. C. Herdic, S. T. Raveendra and B. Gardner, "Interior near-field acoustical holography in flight," *Journal of the Acoustical Society of America*, Vol. 108, pp. 1451-1463 (2000).
- [30] Z. Zhang, N. Vlahopoulos, S. T. Raveendra, T. Allen and K. Y. Zhang, "A computational acoustic field reconstruction process based on an indirect boundary element formulation," *Journal of the Acoustical Society of America*, Vol. 108, pp. 2167–2178 (2000).
- [31] Z. Zhang, N. Vlahopoulos, T. Allen and K. Y. Zhang, "A source reconstruction

- process based on an indirect variational boundary element formulation,” *Engineering Analysis with Boundary Elements*, Vol. 25, 93–114 (2001).
- [32] A. Schuhmacher, J. Hald, K. B. Rasmussen and P. C. Hansen, “Sound source reconstruction using inverse boundary element calculations,” *Journal of the Acoustical Society of America*, Vol. 113, pp. 114–126 (2003).
- [33] Z. Wang and S. F. Wu, “Helmholtz equation-least-squares method for reconstruction the acoustic pressure field,” *Journal of the Acoustical Society of America*, Vol. 102, pp. 2020–2032 (1997).
- [34] S. F. Wu, “On reconstruction of acoustic pressure fields using the Helmholtz equation least squares method,” *Journal of the Acoustical Society of America*, Vol. 107, pp. 2511-2522 (2000).
- [35] N. E. Rayess, Development of Acoustic Holography Using the Helmholtz Equation-Least Squares (HELs) Method, Ph.D. thesis, Department of Mechanical Engineering, Wayne State University, pp. 65–77 (2001).
- [36] M. Moondra and S. F. Wu, “Visualization of vehicle interior sound field using HELs based NAH,” *Noise Control Engineering Journal*, Vol. 53, pp. 146–154 (2005).
- [37] M. Ochmann, “The source simulation technique for acoustic radiation problems,” *Acustica*, Vol. 81, pp. 512–527 (1995).
- [38] M. E. Johnson, S. J. Elliott, K-H. Baek, and J. Garcia-Bonito, “An equivalent source technique for calculating the sound field inside an enclosure containing scattering objects,” *Journal of the Acoustical Society of America*, Vol. 104, pp. 1221–1231 (1998).
- [39] I. Y. Jeon and J. G. Ih, “On the holographic reconstruction of vibroacoustic fields using equivalent sources and inverse boundary element method,” *Journal of the Acoustical Society of America*, Vol. 118, pp. 3475–3484 (2005).

- [40] A. Sarkissian, "Extension of measurement surface in near-field acoustic holography," *Journal of the Acoustical Society of America*, Vol. 115, pp. 1593–1596 (2004).
- [41] A. Sarkissian, "Method of superposition applied to patch near-field acoustic holography," *Journal of the Acoustical Society of America*, Vol. 118, pp. 671–678 (2005).
- [42] M. Ochmann and R. Piscoya, "The source simulation technique with complex source points for computing acoustic radiation problems," *The thirteenth International Congress on Sound and Vibration*, Vienna, Austria (2006).
- [43] M. R. Bai and J. H. Lin, "Source identification system based on the time-domain nearfield equivalence source imaging: fundamental theory and implementation," *Journal of Sound and Vibration*, Vol. 307, pp. 202–225 (2007).
- [44] M. R. Bai, J. H. Lin and K. L. Liu, "Optimized microphone deployment for near-field acoustic holography: To be, or not to be random, that is the question," *Journal of Sound and Vibration*, Vol. 329, pp. 2809-2824 (2010).
- [45] M. R. Bai, J. H. Lin and C. W. Tseng, "Implementation issues of the nearfield equivalent source imaging microphone array," *Journal of Sound and Vibration*, Vol. 330, pp. 545-558 (2011).
- [46] G. H. Koopmann, L. Song and J. B. Fahline, "A method for computing acoustic fields based on the principle of wave superposition", *Journal of the Acoustical Society of America*, Vol. 86, pp. 2433-2438 (1989).
- [47] L. Song, G. H. Koopmann and J. B. Fahline, "Numerical errors associated with the method of superposition for computing acoustic fields," *Journal of the Acoustical Society of America*, Vol. 89, pp. 2626-2633 (1991).
- [48] J. B. Fahline and G. H. Koopmann, "A numerical solution for the general

- radiation problem based on the combined methods of superposition and singular-value decomposition,” *Journal of the Acoustical Society of America*, Vol. 90, pp. 2808–2819 (1991).
- [49] N. P. Valdivia and E. G. Williams, “Study of the comparison of the methods of equivalent sources and boundary element methods for near-field acoustic holography,” *Journal of the Acoustical Society of America*, Vol. 120, pp. 3694-3705 (2006).
- [50] J. N. Juang, *Applied System Identification*, Prentice-Hall, Englewood Cliffs, NJ (1994).
- [51] E. G. Williams, “The nearfield acoustic holography (NAH) experimental method applied to vibration and radiation in light and heavy fluids,” *Computers & Structures*, Vol. 65, pp. 323-335 (1997).
- [52] E. G. Williams and B. H. Houston, “Fast Fourier transform and singular value decomposition formulations for patch nearfield acoustical holography,” *Journal of the Acoustical Society of America*, Vol. 114, pp. 1322-1333 (2003).
- [53] J. H. Zhang and B. Han, “Analysis of engine front noise using sound intensity techniques,” *Mechanical Systems and Signal*, Vol. 19, pp. 213-221 (2005).
- [54] J. Hald, “STSF – a unique technique for scan-based nearfield acoustic holography without restriction of coherence,” *Brüel & Kjær Technical Review No.1* (1989)
- [55] J. Hald, “STSF – practical instrumentation and application,” *Brüel & Kjær Technical Review No.2* (1989).
- [56] J. Hald, “Non-stationary STSF,” *Brüel & Kjær Technical Review No.1* (2000).
- [57] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*, Academic Press, San Diego, CA (1999).
- [58] M. A. Rowell and D. J. Oldham, “Determination of the directivity of a planar

- noise source by means of near field acoustical holography, 1: theoretical background,” *Journal of Sound and Vibration*, Vol. 180, pp. 99-118 (1995).
- [59] W. A. Veronesi and J. D. Maynard, “Digital holographic reconstruction of sources with arbitrarily shaped surface,” *Journal of the Acoustical Society of America*, Vol. 85, pp. 588-598 (1988).
- [60] G. V. Borgiotti, A. Sarkissian, E. G. Williams, L. Schuetz, “Conformal generalized near-field acoustic holography for axisymmetric geometries,” *Journal of the Acoustical Society of America*, Vol. 88, pp. 199–209 (1990).
- [61] J. Hald, Combined NAH and beamforming using the same array, Brüel & Kjær Technical Review No. 1, 2005.
- [62] S. F. Wu, “Methods for reconstructing acoustic quantities based on acoustic pressure measurements,” *Journal of the Acoustical Society of America*, Vol. 124, pp. 2680-2697 (2008).
- [63] R. P. Brent, *Algorithms for Minimization without Derivatives*. Prentice-Hall, Englewood Cliffs, NJ (1973).
- [64] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W. H. Freeman, NY (1979).
- [65] S. Kirkpatrick, C. D. Gelatt, C. D. Jr, and M. P. Vecchi, “Optimization by simulated annealing,” *Science*, Vol. 220 (4598), pp.671-680 (1983).
- [66] R. T. Eduardo, J. K. Hao and T. J. Jose, “An effective two-stage simulated annealing algorithm for the minimum linear arrangement problem,” *Computer and Operation Research*, Vol. 35, pp. 3331-3346 (2008).
- [67] B. Franco, Simulated annealing overview, <http://www.geocities.com/francorbusetti/saweb.pdf> (last viewed on June 22, 2009).
- [68] V. Murino, A. Trucco, and C. S. Regazzoni, ”Synthesis of unequally spaced arrays by simulated annealing,” *IEEE Transactions Signal Processing*, Vol. 44,

pp.119–123 (1996).

- [69] S. L. Gay and J. Benesty, Acoustic signal processing for telecommunication, Kluwer Academic Publishers (2000).
- [70] H. Teutsch and G.W. Elko, “First- and Second-order adaptive differential microphone arrays,” Seventh International Workshop on Acoustic Echo and Noise Control, Darmstadt (2001).
- [71] H. Song and J. Liu, “First-Order Differential Microphone Array for Robust Speech Enhancement,” *Language and Image Processing*, pp. 1461 - 1466 (2008).
- [72] H. Akaike, “A new look at the statistical model identification,” *IEEE Transactions on Automatic Control*, Vol. 19, pp. 716-723 (1974).
- [73] M. C. Junger and D. Feit, “Sound, Structures, and their Interaction,” *MIT Press*, pp.86-112 (1986).
- [74] P. A. Nelson and S. H. Yoon, “Estimation of acoustic source strength by inverse methods: part I, conditioning of the inverse problem,” *Journal of Sound and Vibration*, Vol. 233, pp. 643-668 (2000).
- [75] B. Nobel, J.W. Daniel, Applied Linear Algebra, Prentice-Hall, Englewood Cliffs, NJ (1988).
- [76] O. Kirkeby, P. A. Nelson, and H. Hamada, “Fast deconvolution of multichannel systems using regularization,” *IEEE Transactions on Speech Audio Processing*, Vol. 6, pp. 189-194 (1998).
- [77] M. R. Bai, C. W. Tung and C. C. Lee, “Optimal design of loudspeaker arrays for robust cross-talk cancellation using the Taguchi method and the genetic algorithm,” *Journal of the Acoustical Society of America*, Vol. 117, pp. 2802-2813 (2005).
- [78] P. C. Hansen, Rank-Decient and Discrete Ill-Posed Problems: Numerical

Aspects of Linear Inversion, SIAM (1997).

- [79] G. Wahba, Spline Models of Observational Data, SIAM, Philadelphia (1990).
- [80] P. C. Hansen and D. P. O'Leary, "The use of the L-curve in the regularization of discrete ill-posed problem," *SIAM Journal on Scientific Computing*, Vol. 14, pp. 1487-1503 (1993)
- [81] J. Gomes, J. Hald, P. Juhl and F. Jacobsen, 2009, "On the Applicability of the Spherical Wave Expansion with a Single Origin for Near-Field Acoustical Holography," *Journal of the Acoustical Society of America*, Vol. 125, pp. 1529-1537 (2009).
- [82] National Instruments Corporation, PCI Extensions for Instrumentation (PXI), <http://www.ni.com/> (last viewed on 15 November 2008)
- [83] M. R. Bai, K. N. Hur and Y. T. Liu, "Speech enhancement using an equivalent source inverse filtering -based microphone array," *Journal of the Acoustical Society of America*, Vol. 127, pp. 1373-1380 (2010).
- [84] M. R. Bai and K. Y. Ou , "Synthesis of room responses using virtual source representation with application in reverberator design," *Journal of the Audio Engineering Society*, Vol. 53, pp. 297-306 (2005).
- [85] M. R. Bai and T. C. Tsao, "Numerical modeling of head related transfer functions using the boundary source representation," *American Society of Mechanical Engineers/ Journal of Vibration and Acoustics*, Vol. 128, pp. 594 – 603 (2006).

ABBREVIATIONS

Noise source identification (NSI)

Sound field visualization (SFV)

Simulated annealing (SA)

Monte Carlo (MC)

Intra-block (IB)

Intra-block Monte Carlo (IBMC)

Direction of arrival (DOA)

Delay and sum (DAS)

Time reversal (TR)

Single input multiple output equivalent source inverse filtering (SIMO-ESIF)

Minimum variance distortionless response (MVDR)

Multiple signal classification (MUSIC)

Nearfield equivalence source imaging (NESI)

Equivalent source method (ESM)

Nearfield acoustical holography (NAH)

Multiple-input-multiple-output (MIMO)

Uniform linear array (ULA)

Uniform rectangular array (URA)

Boundary element method (BEM)

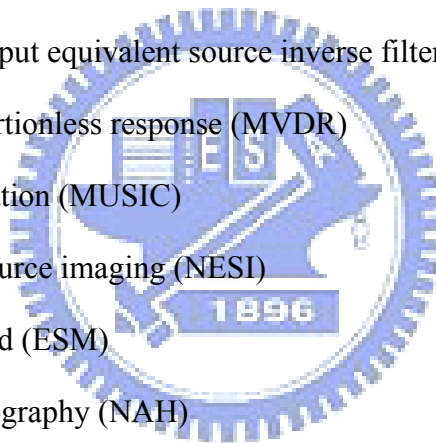
Direct boundary element method (DBEM)

Inverse boundary element method (IBEM)

Helmholtz equation least squares (HELSS)

Finite element analysis (FEA)

Helmholtz integral equation (HIE)



Truncated singular value decomposition (TSVD)

Eigensystem realization algorithm (ERA)

Spatial transformation of sound field (STSF)

Two-dimensional fast Fourier transform (2D FFT)

Fast Fourier transform (FFT)

Discrete time Fourier transforms (DTFT)

Singular value decomposition (SVD)

Statistically optimal NAH (SONAH)

Akaike information criterion (AIC)

Parameter choice methods (PCM)

Information technology (IT)

Transfer function (TF)

Frequency response function (FRF)

Impulse response function (IRF)

Infinite impulse response (IIR)

Finite impulse response (FIR)

Overlap-add method (OAD)

Overlap-save method (OAS)

Minimize the mean square error (MMSE)

Least mean squares (LMS)

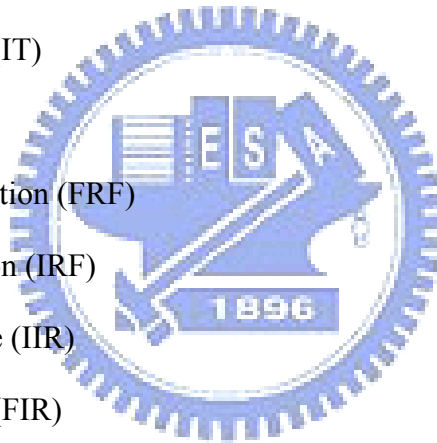
Golden section search (GSS)

Signal-to-noise ratio (SNR)

Differential microphone array (DMA)

Directive index (DI)

Front-to-back ratio (FBR)



Steered response power-phase transform (SRP-PHAT)

Generalized cross-correlation (GCC)

Eigenvalue decomposition (EVD)

Pressure-based conformal holography with a hologram and a source surface coupling (PCHHS)

Velocity-based conformal holography with a hologram and a source surface coupling (VCHHS)

Pressure-based conformal holography with a hologram and a source interior coupling (PCHHI)

Velocity-based conformal holography with a hologram and a source interior coupling (VCHHI)

Distance of reconstruction (DOR)

Generalized cross-validation (GCV)

L-curve criterion (L-C)

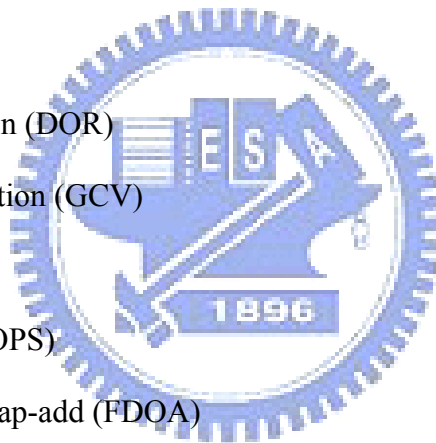
Operations per sample (OPS)

Frequency-domain-overlap-add (FDOA)

Direct convolution (DC)

Retreat distance (RD)

Head related transfer functions (HRTF)



PUBLICATIONS

Journal Papers

1. Mingsian R. Bai and Jia-Hong Lin, "Source identification system based on the time-domain nearfield equivalence source imaging: fundamental theory and implementation," *Journal of Sound and Vibration*, Vol. 307, pp. 202-225 (2007).
2. Mingsian R. Bai, Jia-Hong Lin and Kwan-Liang Liu, "Optimized microphone deployment for near-field acoustic holography: To be, or not to be random, that is the question," *Journal of Sound and Vibration*, Vol. 329, pp. 2809-2824 (2010).
3. Mingsian R. Bai, Jia-Hong Lin and Chih-Wen Tseng, "Implementation issues of the nearfield equivalent source imaging microphone array," *Journal of Sound and Vibration*, Vol. 330, pp. 545-558 (2011).
4. Mingsian R. Bai, Ching-Cheng Chen and Jia-Hong Lin, "On optimal retreat distance for the equivalent source method-based nearfield acoustical holography," *Journal of the Acoustical Society of America*, (accepted).

Conference Papers

1. Mingsian R. Bai and Jia-Hong Lin, "A new noise source identification technique: nearfield acoustical beamformer (NABF)," Proceedings of Inter-noise 2006, Honolulu (December 2006).
2. Mingsian R. Bai and Jia-Hong Lin, "Sound field reconstruction using the time-domain nearfield equivalence source imaging (NESI) technique," Proceedings of Inter-noise 2007, Istanbul (August 2007).
3. Mingsian R. Bai, Jia-Hong Lin and Kwan-Liang Liu, "Investigation on Uniform and Non-Uniform Deployment of Microphones in Near-Field and

Far-Field Source Imaging,” Proceedings of Inter-noise 2008, Shanghai (October 2008).

4. Mingsian R. Bai, Jia-Hong Lin and Ching-Cheng Chen, “On optimal retreat distance of equivalent source,” Proceedings of the 17th international congress on sound & vibration (ICSV), Cairo (July 2010).

Patents

1. 白明憲、林家鴻、甯攸威，「聲源能量分佈可視化系統及其方法」，台灣證號：I313823。
2. Mingsian R. Bai, Jia-Hong Lin and Yu-Wei Ning, ”System and Method for Visualizing Sound Source Energy Distribution,” (美國審查中)。
3. 白明憲、林家鴻、劉冠良，「聲學照相機」，TW 201039653 A1。

Contests

1. 民國 95 年以”陣列麥克風噪音源辨識系統:近場聲學波束合成器”獲頒車輛研究測試中心(ARTC)頒發第二屆車輛工程研究論文銀質獎。
2. 民國 97 年以"利用 Labview 及 PXI 實現麥克風陣列聲場可視化"榮獲美商國家儀器 2008 第七屆虛擬儀控徵文比賽學術組第一名。
3. 民國 98 年以"利用 Labview 及 PXI 實現麥克風陣列聲場可視化"獲選刊登於《機電整合》125 期中。
4. 民國 100 年以”利用時域之近場等效聲源影像法進行聲源識別和聲場重建”獲選刊登於《工程科技通訊》111 期中。