

國立交通大學

機械工程研究所

碩士論文

頭部轉移函數之數值計算與量測

Computation and measurement of

Head Related Transfer Function

研究生：曹登傑

指導教授：白明憲

中華民國九十三年六月

頭部轉移函數之數值計算與量測

Computation and measurement of

Head Related Transfer Function

研究生: 曹登傑

Student: Tengchieh Tsao

指導教授: 白明憲

Advisor: Mingsian Bai

國立交通大學

機械工程研究所



Submitted to Department of Mechanical Engineering

Collage of Engineering

National Chiao Tung University

in Partial Fullfillment of Requirements

for the Degree of Master of Science

in

Mechanical Engineering

June 2004

Hsinchu, Taiwan, Republic of China

中華民國九十三年六月

頭部轉移函數之數值計算與量測

研究生: 曹登傑

指導教授: 白明憲

國立交通大學機械工程學系

摘要

本文在頭部轉移函數(HRTF)數值計算方面提出一以階層勢能定理為基礎的簡單虛擬聲源重現法。在此方法下，由任意形狀輻射聲源所造成的聲場可以聲波疊加之原理描述。在使用聲源互易性和適當的矩陣規則化之下，每一個方位上的頭部轉移函數可藉由設定外在頭部網格之邊界條件進行數值計算而得到。任意方位上的頭部轉移函數可被描述成是擁有不同體積速度之自由空間格林函數的總和。另外在本文中利用間接邊界元素法和量測所得之頭部轉移函數也用來和經由虛擬聲源重現法所計算出的頭部轉移函數進行比較。在頻率響應分析方面，足夠的網格數目和適當的矩陣規則化可以使得計算出之頭部轉移函數近乎所量測之頭部轉移函數的包絡線。合理的指向性更是驗證了頭部轉移函數數值計算的可行性。在運算量分析上，虛擬聲源重現法在合理的網點數目下展現了在運算時間與記憶體使用上較為少量的需求。在結果的分析上，數值誤差可以藉由增加網點數目來加以改善。數值計算和量測所得的頭部轉移函數也分別實現在數次主觀聆聽實驗中。實驗結果顯示，合成頭部轉移函數的重要定位特徵可以在這些聆聽實驗中成功的產生作用。

Computation and Measurement of Head Related Transfer Function

ABSTRACT

A simple virtual source representation motivated by layer potential theory is presented for head related transfer function (HRTF) computations. The acoustic field of arbitrarily shaped radiators is described by using the principle of wave superposition in virtual source representation. With the acoustic reciprocity and appropriate matrix regularization, the HRTFs at each direction are calculated by setting the additional boundary condition of external mesh. The HRTF at an arbitrary direction finally is presented as the sum of free space Green's functions with different volume velocities. Another numerical method – indirect boundary element method and a HRTF measurement are also implemented to compare with the calculated HRTFs by using virtual source representation. In the frequency response analysis, the calculated HRTFs closely approximate the envelopes of the measured HRTFs with enough mesh nodes and suitable matrix regularization. Reasonable directional responses in both numerical methods also prove the feasibility of HRTF numerical computation. From the comparison of computation loadings, the virtual source representation shows the fewer requirements in computation time and memory use under the reasonable number of mesh nodes. Performance analysis reveals the numerical errors can be improved by increasing the number of nodes. Numerous subjective experiments are also conducted by using calculated HRTFs and measured HRTFs. The results reveal the important characteristics of sound localization are successfully reproduced by the synthesized HRTFs.

誌謝

短暫的兩年研究所生涯很快地過去了，首先感謝指導教授白明憲博士的諄諄指導與教誨，使我順利完成學業與論文，在此致上最誠摯的謝意。而老師指導學生時豐富的專業知識、嚴謹的治學態度，亦是身為學生的我學習與景仰的典範。

在論文寫作上，感謝本系李安謙教授、林家瑞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，在此本人至上無限的感激。

回顧這兩年的日子，承蒙同實驗室的博士班曾平順學長、歐昆應學長、李志中學長、林家鴻學長、蘇富誠學長、陳榮亮學長與顏坤龍學長在研究與學業上的適時指點，並有幸與董志偉、廖哲緯、林振邦、及白淦元同學互相切磋討論，獲益甚多。此外學弟何柏璋、曾文亮、周中權及林建良在生活上的朝夕相處與砥礪磨練，都是我得以完成研究的一大助因，在此由衷地感謝他們。

能有此刻，我也要感謝所有在精神上給我鼓舞支持的人，感謝各位的幫忙與鼓勵。最後僅以此篇論文獻給我摯愛的雙親曹火生先生、吳寶琴女士、妹妹曹郁翎。另外感謝我的摯友洪佳惠、林立原、張孟晉，謝謝他們不斷的鼓勵我。今天我能順利取得碩士學位，要感謝的人很多，上述名單恐有疏漏，在此也致上我最深的謝意。

TABLE OF CONTENTS

摘要.....	I
ABSTRACT	II
誌謝.....	III
TABLE OF CONTENTS	IV
LIST OF TABLES	V
LIST OF FIGURES	VI
I. INTRODUCTION	1
II. HEAD RELATED TRANSFER FUNCTION	4
III. MEASUREMENT OF HEAD RELATED TRANSFER FUNCTION ..	6
A. Experimental arrangement.....	6
B. Measured HRTF analysis and post processing.....	7
IV. COMPUTATION OF HEAD RELATED TRANSFER FUNCTION ..	9
A. Indirect boundary element method.....	9
B. Virtual source representation.....	10
C. Technical issues of virtual source representation.....	15
1. Optimization of virtual source representation.....	15
2. Regularization of virtual source representation.....	17
D. Pre-processing of computational mesh model.....	19
1. Acquiring the mesh model.....	19
2. Mesh decimation.....	20
3. Mesh topology construction.....	20
4. Mesh data transformation.....	21
E. Verification of analytical solution.....	21
V. EXPERIMENTAL INVESTIGATIONS AND NUMERICAL ANALYSIS ..	23
A. Frequency response analysis.....	26
B. Directional response.....	27
C. Impulse response.....	28
D. Analysis of computation loading.....	29
E. Subjective experiment.....	30
VI. CONCLUSIONS	32
REFERENCES	34
TABLES	37
FIGURES	39

LIST OF TABLES

Table I. Number of measurement and azimuth increments at each elevation.....37
Table II. Comparison of computation loadings between IBEM and virtual source
representation.....38



LIST OF FIGURES

Figure 1.	Coordinate systems for HRTF representation:(a) Interaural-Polar coordinate system (b) Vertical-Polar coordinate system.....	39
Figure 2.	An HRTF example of KEMAR at azimuth=45° and elevation=0° (a) The impulse response (HRIR). (b) The HRTF frequency response.....	40
Figure 3.	Schematic diagram of measurement mechanism.....	41
Figure 4.	Photo of HRTF measurement mechanism.....	42
Figure 5.	Frequency response of measurement loudspeaker and power amplifier.	43
Figure 6.	The program interface of HRTF measurement written by using Visual Basic.....	44
Figure 7.	The interface of Pulse multi analyzer software - Pulse Labshop.....	45
Figure 8.	Measured HRTFs post-processing flow chart.....	46
Figure 9.	Measured HRTFs on the horizontal plane (elevation = 0°) (solid line: unequalized MIT HRTF, dash line : unequalized HRTF) (a) azimuth = 0° (b) azimuth = 30° (c) azimuth = 60° (d) azimuth = 90° (e) azimuth = 120° (f) azimuth = 150°	47
Figure 9.	Measured HRTFs on the horizontal plane (elevation = 0°) (solid line: unequalized MIT HRTF, dash line : unequalized HRTF) (g)azimuth = 180° (h) azimuth = 210° (i) azimuth = 240° (j)azimuth = 270° (k) azimuth = 300° (l) azimuth = 330°	48
Figure 10.	Measured HRTFs on the horizontal plane (elevation = 0°) (solid line: equalized MIT HRTF, dash line : equalized HRTF) (a) azimuth = 0° (b) azimuth = 30° (c) azimuth = 60° (d) azimuth = 90° (e) azimuth = 120° (f) azimuth = 150°	49
Figure 10.	Measured HRTFs on the horizontal plane (elevation = 0°) (solid line: equalized MIT HRTF, dash line : equalized HRTF) (g)azimuth = 180° (h) azimuth = 210° (i) azimuth = 240° (j)azimuth = 270° (k) azimuth = 300° (l) azimuth = 330°	50
Figure 11.	Indirect boundary element method definition (close body).....	51
Figure 12.	Generic L-curve form; a plot in the log-log scale of the discrete smoothing norm versus the residual norm as a function of regularization parameter.....	52
Figure 13.	Spacing of nodes at maximum frequency. Seven points denote the six elements in one wavelength.....	53
Figure 14.	Circular piston of radius α set in the rigid sphere. The piston vibrates with a velocity \dot{w} . The z axis passes through the center of the piston.....	54

Figure 15.	The settings of the mesh elements in the SYSNOISE. The light color areas denote the rigid baffle and others denote the vibrating circular piston.....	55
Figure 16.	The settings of mesh nodes in virtual source representation. The plus sign in (a) and (b) denotes the vibrating circular piston. Inner black points denote the virtual sources.....	56
Figure 17.	Frequency responses of circular piston in a spherical baffle with optimum and small τ in virtual source representation and analytical solution.....	57
Figure 18.	L-curve of circular piston in a spherical baffle at (a)100 Hz. (b)Distribution of value (γ) of cost function. The regularization parameter β is taken as the first singular value multiplied by the values marked on the curve.....	58
Figure 19.	Frequency responses of circular piston in a spherical baffle calculated by using analytical solution, virtual source representation and IBEM.....	59
Figure 20.	Directional response of circular piston in a spherical baffle at (a) 300 Hz (b) 500 Hz.....	60
Figure 20.	Directional response of circular piston in a spherical baffle at (c) 700 Hz (d) 1000 Hz.....	61
Figure 21.	The mesh settings of KEMAR dummy head using IBEM in the SYSNOISE (ITRI case). (a) Head mesh given by ITRI.....	62
Figure 21.	The mesh setting of KEMAR dummy head using IBEM in the SYSNOISE (ITRI case). (b) Pinna (White color areas denote the vibrating panels).....	63
Figure 22.	The mesh settings of KEMAR dummy head using IBEM in the SYSNOISE (Kahana case). (a) Head mesh given by Yuvi Kahana.....	64
Figure 22.	The mesh settings of KEMAR dummy head using IBEM in the SYSNOISE (Kahana case). (b) Pinna (White areas are the vibrating panels).....	65
Figure 23.	The distributions of KEMAR mesh nodes in virtual source representation. (Only 2% of the original scanned head mesh nodes and 17% of the original scanned pinna mesh nodes). (a)The right hand side of KEMAR (case 1) (b) The front of KEMAR (case 1) (The cluster of light color points denote the virtual sources).....	66
Figure 23.	The distributions of KEMAR mesh nodes in virtual source representation. (Only 10% of the original scanned head mesh nodes and 50% of the original scanned pinna mesh nodes). (c) The right hand side of KEMAR (case 2) (d) The front of KEMAR (case 2) (The cluster of light color points denote the virtual sources).....	67

Figure 24.	L-curve of HRTF computations using virtual source representation at (a) 100 Hz. (b) Distribution of value (γ) of cost function. The regularization parameter β is taken as the first singular value multiplied by the values marked on the curve. (Virtual source case 1).....	68
Figure 25.	Calculated HRTFs on the horizontal plane by using IBEM (ITRI case). (a) 0 azimuth degrees (b) 90 azimuth degrees.....	69
Figure 25.	Calculated HRTFs on the horizontal plane by using IBEM (ITRI case). (c) 180 azimuth degrees (d) 270 azimuth degrees.....	70
Figure 26.	Calculated HRTFs on the horizontal plane by using IBEM (Kahana case). (a) 0 azimuth degrees (b) 90 azimuth degrees.....	71
Figure 26.	Calculated HRTFs on the horizontal plane by using IBEM (Kahana case). (c) 180 azimuth degrees (d) 270 azimuth degrees.....	72
Figure 27.	Calculated HRTFs by using virtual source representation on the horizontal plane (virtual source case 1).(a) 0 azimuth degrees (b) 90 azimuth degrees	73
Figure 27.	Calculated HRTFs by using virtual source representation on the horizontal plane (virtual source case 1).(c) 180 azimuth degrees (d) 270 azimuth degrees.....	74
Figure 28.	Calculated HRTFs by using virtual source representation on the horizontal plane (virtual source case 2). (a) 0 azimuth degrees (b) 90 azimuth degrees	75
Figure 28.	Calculated HRTFs by using virtual source representation on the horizontal plane (virtual source case 2). (c) 180 azimuth degrees (d) 270 azimuth degrees.....	76
Figure 29.	Directional response of measured HRTF on the horizontal plane at (a) 500 Hz (b) 2100 Hz.....	77
Figure 29.	Directional response of measured HRTF on the horizontal plane at (c) 5100 Hz (d) 9100 Hz.....	78
Figure 30.	Directional response calculated by using IBEM (Kahana case) at (a) 500 Hz (b) 2100 Hz (dash line: measured HRTF, solid line: calculated HRTF) (elevation = 0°).....	79
Figure 30.	Directional response calculated by using IBEM (Kahana case) at (c) 5100 Hz (d) 9100 Hz (dash line: measured HRTF, solid line: calculated HRTF) (elevation = 0°).....	80
Figure 31.	Directional response calculated by using IBEM (ITRI case) at (a) 500 Hz (b) 2100 Hz (dash line: measured HRTF, solid line: calculated HRTF) (elevation = 0°).....	81

Figure 31.	Directional response calculated by using IBEM (ITRI case) at (c) 5100 Hz (d) 9100 Hz (dash line: measured HRTF, solid line: calculated HRTF) (elevation = 0°).....	82
Figure 32.	Directional response on the horizontal plane calculated by using virtual source representation (virtual source case 1) at (a) 500 Hz (b) 2100 Hz (dash line: measured HRTF, solid line: calculated HRTF).....	83
Figure 32.	Directional response on the horizontal plane calculated by using virtual source representation (virtual source case 1) at (c) 5100 Hz (d) 9100 Hz (dash line: measured HRTF, solid line: calculated HRTF).....	84
Figure 33.	Directional response on the horizontal plane calculated by using virtual source representation (virtual source case 2) at (a) 500 Hz (b) 2100 Hz (dash line: measured HRTF, solid line: calculated HRTF).....	85
Figure 33.	Directional response on the horizontal plane calculated by using virtual source representation (virtual source case 2) at (c) 5100 Hz (d) 9100 Hz (dash line: measured HRTF, solid line: calculated HRTF).....	86
Figure 34.	Head related impulse response on the horizontal plane from the measured HRTF database. (solid line: 90 azimuth degrees, dash line: 270 azimuth degrees.).....	87
Figure 35.	Head related impulse response from the calculated HRTFs by using IBEM on the horizontal plane (ITRI case).....	88
Figure 36.	Head related impulse response from the calculated HRTFs by using virtual source representation on the horizontal plane (case 1).....	89
Figure 37.	Head related impulse response from the calculated HRTFs by using virtual source representation on the horizontal plane (case 2).....	90
Figure 38.	Sound localization results of an azimuth test using white noise (a) Measured HRTFs (b) HRTFs from MIT HRTF database.....	91
Figure 38.	Sound localization results of an azimuth test using white noise (c) Calculated HRTFs by using virtual source representation (case 1) (d) Calculated HRTFs by using IBEM (ITRI case).....	92
Figure 38.	Sound localization results of an azimuth test using white noise (e) Calculated HRTFs by using virtual source representation (Case 2)....	93
Figure 39.	The total average error and average error at each direction of sound localization (a) Measured HRTFs (b) HRTFs from MIT HRTF database.	94
Figure 39.	The total average error and average error at each direction of sound localization (c) Calculated HRTF of virtual source case 1 (d) Calculated HRTF by using IBEM (ITRI case).....	95
Figure 39.	The total average error and average error at each direction of sound localization (e) Calculated HRTF of virtual source case 2.....	96

I. INTRODUCTION

Multimedia information content is changing the phase of our daily life. Three-dimensional (3D) sound technology is becoming increasingly important for scientific, commercial, and entertainment systems. More and more signal processing technique is changing with each passing day to satisfy people's requirements in 3D multimedia. One of the important 3D sound techniques is accurate spatial sound reproduction. Accurate spatial sound reproduction enhances the three dimensional multimedia, especially important for application in video and audio synchronization. In the traditional stereo systems and recently the Dolby Digital 5.1 systems¹, using multi-drives is a commonplace to reconstruct the sound images between speakers. Most spatial sound reproduction systems are still based on concept of binaural synthesis today. Despite of simplicity, those binaural synthesis approaches are just a crude approximation of the intended sound field as produced by physical sources. In order to manipulate the signals arriving the eardrum of the listener in a binaural synthesis, the directional characteristics of physically signal filtered by external ear, head and torso should be fully exploited. There is a new generation of 3D sound technologies makes positioning virtual sound images in a more accurate way. A truly immersive listening experience and accurate spatial sound reproduction can be created by this new generation 3D sound technology.

Previous work in the area of localization and spatial sound rendering dates back to the year 1907, Lord Rayleigh brought up the "duplex theory"² of sound localization. Duplex theory asserts the apparent lateral positions of low and high frequency stimuli are determined respectively by interaural time differences (ITD) and interaural level differences (ILD). The ITD and ILD cues can partially explain the human amazing ability in sound localization. Interaural differences do not account for ability of source localization in many locations which ITD and ILD are nearly the same.

People can still differentiate between these sound sources because of other localization cues arising from sound scattering. Actually the auditory system of human behaves like a complex spatial–temporal filter. The directional cues are embodied in the transformation of the sound source to the listeners. This acoustic filter is referred to as the head related transfer function (HRTF). A head related transfer function (HRTF) describes the transformation of a specific source direction relative to the head and filtering process associated with the diffraction of sound by the pinna, head and torso. The inverse Fourier transform of a HRTF is the head-related impulse response (HRIR). A 3D sound field can be created by convolving the source signal with an appropriate pair of HRIRs to position the sound image arbitrarily around the listener.

Several HRTF databases and measuring techniques have been developed around the world.³⁻⁵ Today the direct convolution using measured HRTFs is used commonly in 3D sound reproduction. Nevertheless, this approach is effective but complex and limited. There are several disadvantages in direct using measured HRTFs. First, a great amount of HRTFs needs to be stored in implementation of accurate 3D sound reproduction. This result in getting a smoothed sound impression in moving sources synthesis is particularly difficult. Second, computations of direct using measured HRTFs to synthesize multiple sources are heavy and complex. More and more efficient synthesizing techniques have been developed functionally because of these limitations. Efficient approaches are really desirable and important in synthesizing for real-time applications. Several researchers have investigated in replacing measured HRTFs with computational HRTF models. Here are some computational HRTF models have been investigated in several ways. Batterau represents the first functional approach of modeling HRTFs.⁶ The external ear could be modeled as a three-channel, two-delay and sum, acoustic coupler in Batterau’s proposed paper.

There is a series of HRTF modeling technique are proposed by Duda.⁷⁻⁹ HRTFs can be modeled into spherical models and other simple models of pinna features and shoulder echo in Duda's proposed structural model. Another important effort is that of Yuvi Kahana who investigated various aspects of numerically modeled individualized HRTFs especially with the emphasis on Boundary Element Method (BEM).¹⁰ In his thesis, further acoustic features of external ear are also investigated with emphasis on the 'mode shape' of human pinna.¹¹ Although more and more synthesizing 3D sound techniques are proposed, these are either simple but limited, or complicated but effective.

A real measurement of HRTF is realized by using a simple measurement mechanism here. Total 598 HRTFs at different directions are measured and an HRTF database is also constructed. Virtual source representation is a new HRTF modeling technique proposed in this thesis. The acoustic field of arbitrarily shaped radiators is described by using the principle of wave superposition in virtual source representation. With the acoustic reciprocity, HRTFs at different directions can be calculated by appropriate matrix regularization in virtual source representation. Compared to other complex HRTF modeling technique, virtual source representation just describes the HRTFs as the sum of free space Green's functions with different source strengths. Indirect Boundary Element Method (IBEM) is also used to calculate HRTFs via a leading software package for Computational Vibro-Acoustics – LMS SYSNOISE.^{12,13} The comparison between the measured HRTFs and the calculated ones by using IBEM and virtual source representation is also discussed. The comparison of calculated HRTFs and measured ones is focused on the frequency response, directional response and impulse response in section IV. A complete subjective experiment is implemented by using the calculated HRTFs and measured HRTFs. The experimental results will be discussed in detail in section V.

II. HEAD RELATED TRANSFER FUNCTION

Head related transfer function (HRTF) is defined as a specific individual's left and right ear far-field frequency response, as measured from a specific point in the free field to a listener's ear canal. The HRTF is a directionally dependent complex valued transfer function which is the function of azimuth, elevation and frequency. Here are two different polar coordinate systems are used, the more common used coordinate system is the polar axis oriented vertically (vertical polar coordinate system), the other is the axis running horizontally through the two ears.⁹ Fig. 1 (a) and (b) demonstrate the vertical and horizontal polar coordinate systems. An example of HRTF pair is plotted in Fig. 2 (a) and (b) is the inverse Fourier Transform of the equivalent HRTF (i.e. HRIR). In Fig. 2 (a), different time delay and magnitude level in the frontal of these two HRIRs are found to be the ITD and ILD. In Fig. 2 (b), the complex structure of peaks and notches of HRTF is made by the head shadowing effect and ear resonance.¹⁴ For instance, the distinctive notches at 8-9 kHz are mainly caused by concha reflection and increasing attenuation with frequency in the contra lateral ear. Previous measurements of the effect of torso revealed that the main contribution was in the frequency region up to 3 kHz depending on the angle of incidence.¹⁴

Many researchers have investigated several ways to model HRTFs. There are general approaches listed as the following: 1) Rational function or pole/zero models, 2) Series expansion, 3) Structure models.⁶⁻⁸ First, the modeling problem is viewed as the system identification with several classical solutions. The coefficients are usually complicated functions of azimuth and elevation that they have to be tabulated and this destroys the usefulness of the model. Second, Fourier series expansions or principal component analysis models the HRTF as a weighted sum of simpler basis functions. This is useful for inspecting the data, but the run time complexity limited

the usefulness of this modeling method. Third, structure models can be taken as composition of different physical mechanisms. HRTFs are crafted into head shadowing, shoulder reflection, and pinna effect components in this model.⁸ The structure model is simple to implementation but limited performance is also obtained. There is also a method based on acoustic principles and a simpler decomposition in Chen's paper.¹⁵ The proposed fast algorithm can reduce the computational burden by an order of magnitude compared with conventional methods. Another HRIR based on interpolation and reduced order model (HIROM)¹⁶ based on singular value decomposition (SVD) separated the spatial and temporal variation of HRTF. A great of computation loading can be reduced in implementation of multi-sources localization using this functional model. Spatial acoustic mode shapes of the human pinna were also described using a set of orthogonal basis function which is based on SVD in Kahana's paper.¹¹

Many publications have dealt with the measurements and analysis of HRTFs in the past 50 years. The definitions, methods and equalization techniques vary among studies, and a technical standard for HRTF measurements does not exist. The most famous HRTF database is constructed by the MIT Media Laboratory in 1994.⁴ The data consist of the left and right ear impulse responses from a Realistic Optimus Pro 7 loudspeaker mounted 1.4 meters from Knowles Electronics Manikin for Acoustic Research (KEMAR) dummy head. Maximum length (ML) pseudo-random binary sequences were used to obtain the impulse responses at a sampling rate of 44.1 kHz. In total, 710 different positions were sampled at elevations from -40 degrees to 90 degrees. In the past measurements of HRTF, the choices of test signal to identify the system of frequency response are different from one to others. In the HRTF measurement of MIT, maximum length sequences (MLS) is chosen to be test signal because of the list of following advantages of pseudo random signal. 1) MLS

methods employ efficient cross correlation between input and output to recover the periodic impulse response (PIR) of system being measured. 2) In the face of external noise and nonlinearities, the MLS approach is shown to be as robust as time delay spectrometry. MLS methods are also capable of rejecting or selecting nonlinear (distortion) components when measuring weakly nonlinear systems.¹⁷ An MLS is a periodic pseudo random binary sequence and actually a deterministic signal which has similar spectral properties as true random white noise. There are still other HRTF databases on the internet provided by several groups around the world, such as CIPIC⁵, Suzuki Laboratory⁶ etc.

III. MEASUREMENT OF HEAD RELATED TRANSFER FUNCTIONS

A. Experimental arrangement

Preliminary ideas of measurement mechanism and real experimental arrangement are shown in the Fig. 3 and 4 respectively. The measurement equipments are listed as the following, 1) two Parker stepper motors, 2) B&K type 3109 Pulse multi-analyzer, 3) B&K type 4191 microphone, 4) power amplifier and 5) one Meiloon satellite loudspeaker. In the Fig. 3, a simple measurement mechanism was designed to change azimuth and elevation angles as accurate as possible using two Parker stepper motors. Total 360 degrees in one circle can be divided into 25000 steps under control of these two Parker stepper motors. The ceiling steps of 1 degree were settled to be 69 and total accumulative errors were about 2.32 degrees when turning 360 degrees. The frequency responses of power amplifier and loudspeaker are shown in Fig. 5. B&K Pulse multi-analyzer is used to data acquisition and real-time data processing at a sampling rate of 65536 Hz. A user interface program written by Visual Basic combined the controls of two stepper

motors and real-time signal processing by B&K Pulse multi-analyzer. Fig. 6 and 7 are the user interfaces of automation measurement. A measurement of automation can be done via user setting up every turning azimuth and elevation angles and test signal. In the traditional modal analysis of linear system, Gaussian noise is commonly used as test signal because of impossible implementation of impulse signal. Here pseudo random noise was chosen to be test signal in this measurement. Pseudo random noise also has the same properties as MLS. Pseudo random noise can also avoid spectral leakage happened in impulse response and total 1500 linear averages were set to suppress the effect of noise in the measurement. The Knowles Electronics Manikin for Acoustic Research (KEMAR) dummy head used for measurement was mounted on a turntable which was turned around by one stepper motor. Two different ears are mounted on KEMAR. The left ear (DB60) is an average ear designed as an average of the dimensions and response of 24 subjects. These are typical of American and European females as well as Japanese males and females. Right ear (DB65) is developed with large dimensions of length and breadth. This is typical of American and European male pinna sizes. The total number of measurements and azimuth increments at each elevation were listed in Table I. The loudspeaker mounted on the measurement mechanism was 1.2 meters away from the KEMAR. The measurement experiment was processed in a semi anechoic chamber.

B. Measured HRTF analysis and post processing

In the process of measurement, the coherence function between the input pseudo random signal and received output signal is an important index of degree of linearity between two signals. The coherence function measures the linear relationship between two signals in the frequency domain. The measured results of horizontal plane (0 elevation degree) are listed from 0 to 330 azimuth degrees with a 30 degrees

interval in Fig. 9. Obviously, measured HRTFs in ipsilateral (0~180 azimuth degrees) look much smoother than the ones in contra lateral (180~360 azimuth degrees). The reason of this phenomenon is because of the shadowing of head. Under this situation, the degree of linearity of coherence function and measured HRTFs are much lower than the ones of ipsilateral. The post-processing of measured HRTF is another important point in measurement. Here the term ‘post processing’ refers to equalization of frequency response of loudspeaker, power amplifier and microphone. In measurement of total 598 different positions, resolution of every HRTF was chosen to be 800 lines in a frequency span of 25.6 kHz. Each post processing step can be plotted as a simple flow chart in Fig. 8. In Fig. 9, an obvious poor low frequency response can be obtained in unequalized measured HRTF. This bad frequency response is resulted from bad performance of low frequency of measurement loudspeaker. In order to get the “true” HRTFs in the measurement, equalizations of HRTFs must be done. Every measured HRTF can be equalized according to

$$\frac{\text{Measured HRTF}}{\text{Loudspeaker response}} = \frac{|M|_{\text{HRTF}}}{|M|_{\text{loudspeaker}}} e^{i(\theta_{\text{HRTF}} - \theta_{\text{loudspeaker}})} \quad (1)$$

The equalized results are also shown and compared to MIT’s HRTFs in Fig. 10. There are obvious some differences in 5-7 kHz at each HRTF plot. A distinct higher gain of measured HRTFs can be seen in this frequency range and this is caused by the first resonance of the pinna (DB65). The second anti resonance of pinna can also be seen at about 9 kHz. Total 598 positions and 800 points length equalized HRTF database are also constructed in this measurement.

IV. COMPUTATION OF HEAD RELATED TRANSFER FUNCTION

A. Indirect boundary element method

The indirect boundary element method can be used for cases where the domain is open, or includes both closed and open boundary surfaces.¹⁰ The indirect method also uses layer potentials that are the differences between the outside and inside values of the pressure and its normal gradient. In Fig. 11, the pressure and its gradient of the positive and negative sides of the boundary surface are showed. For open bodies, it can be shown that the indirect boundary integral formulation is given by

$$p(\mathbf{r}) = \int_{s^+} \left\{ p^+(\mathbf{r}_0) \frac{\partial G(\mathbf{r}, \mathbf{r}_0)}{\partial n} - \frac{\partial p^+(\mathbf{r}_0)}{\partial n} G(\mathbf{r}, \mathbf{r}_0) \right\} dS(\mathbf{r}_0) - \int_{s^-} \left\{ p^-(\mathbf{r}_0) \frac{\partial G(\mathbf{r}, \mathbf{r}_0)}{\partial n} - \frac{\partial p^-(\mathbf{r}_0)}{\partial n} G(\mathbf{r}, \mathbf{r}_0) \right\} dS(\mathbf{r}_0) \quad (\mathbf{r} \in V) \quad (2)$$

or in general form for open or close boundary surfaces

$$p(\mathbf{r}) = \int_s [\mu(\mathbf{r}_0) \frac{\partial G(\mathbf{r}, \mathbf{r}_0)}{\partial n} - \eta(\mathbf{r}_0) G(\mathbf{r}, \mathbf{r}_0)] dS(\mathbf{r}_0) \quad (\mathbf{r} \in V) \quad (3)$$

$$\mu(\mathbf{r}_0) = p^+(\mathbf{r}_0) - p^-(\mathbf{r}_0), \quad \eta(\mathbf{r}_0) = \frac{\partial p^+(\mathbf{r}_0)}{\partial n} - \frac{\partial p^-(\mathbf{r}_0)}{\partial n}$$

where G denotes the free space Green's function. μ is generally called the jump pressure or double layer potential. It represents a distribution of dipole sources on the surface. η is generally called the jump of normal derivative of pressure or single layer potential, and it represents a distribution of monopole source on the surface. In the discretization of S , the geometric shape functions are defined and the unknown η and μ are expressed in terms of their nodal values:

$$\eta(\mathbf{r}) = \sum \eta_i N_i(\xi) \quad (4)$$

$$\mu(\mathbf{r}) = \sum \mu_i N_i(\xi) \quad (5)$$

The first order shape function of the local coordinates $(\xi) \equiv (\xi_1, \xi_2, \xi_3)$ and are $N_i(\xi) = \xi_i$ ($i=1,2,3$). Each shape function is defined such that it has a unit value at the location of node i and that it is zero at all other node locations. Once both the single and double layers are known, the boundary integral equation can be used to post process the pressure at any field point in the volume which is then given by

$$p(\mathbf{r}) = \sum_e \int_{S^e} [N_\mu(\mathbf{r}_0) \mu \frac{\partial G(\mathbf{r}, \mathbf{r}_0)}{\partial n} - N_\eta(\mathbf{r}_0) \eta G(\mathbf{r}, \mathbf{r}_0)] dS^e(\mathbf{r}_0) \quad (6)$$

where the surface S of the body is discretized with e elements such $S \cong \sum_e S^e$ and S^e is the surface of each element. In this research, the implementation of indirect boundary element method is via the software - LMS SYSNOISE.

B. Virtual source representation

There are several methods for modeling the human spatial hearing cues in previous researches. The virtual source representation is proposed in this thesis for computing the HRTF at different azimuth and elevation. This method is motivated by layer potential theory. The acoustic field of arbitrarily shaped radiators is described by using the principle of wave superposition.¹⁸ In the Helmholtz integral equation (HIE), the surface pressure and surface velocity must be known makes it difficult to be calculated. This difficulty can be avoided by using a special Green's function in place of the free space Green's function in the Helmholtz integral equation. There is another technique to overcome this difficulty which is called the layer potential formulation. The pressure inside or outside of the surface could be determined by the single layer representation.

$$p(\mathbf{x}) = \int_s \sigma(\mathbf{x}_0) G(\mathbf{x}, \mathbf{x}_0) dS(\mathbf{x}_0) \quad (7)$$

Eq. (7) is similar to Helmholtz integral equation with one term in the integrand.

Here G is the free space Green's function and σ is an unknown strength of the monopole distribution on the boundary. The pressures are continuous across the boundary in this representation, and Eq. (7) is a valid representation anywhere in space. On the other hand, the velocities across the boundary are discontinuous.

$$\frac{\partial p}{\partial n}(\mathbf{x}) = \alpha \sigma(\mathbf{x}) + \int_S \sigma(\mathbf{x}_0) \frac{\partial G}{\partial n}(\mathbf{x}, \mathbf{x}_0) dS(\mathbf{x}_0), \quad \alpha = \begin{cases} \frac{1}{2} & \mathbf{x} \in V_o \\ -\frac{1}{2} & \mathbf{x} \in V_i \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

$G(\mathbf{x}, \mathbf{x}_0)$ is the free space Green's function,

$$G(\mathbf{x}, \mathbf{x}_0) = \frac{A}{r} e^{-jkr}, \quad (9)$$

where the $(\mathbf{x}, \mathbf{x}_0)$ are the field point and source point, respectively, $\mathbf{r} = \mathbf{x} - \mathbf{x}_0$, $r = |\mathbf{r}|$, A is the source amplitude and $k = \omega/c$ is the wave number, c is the speed of the sound. Elaborate numerical schemes such as FEM or BEM can be used for discretization of the above inherently continuous formulation. However, additional insights can be gained by simply approximating the integral in Eq. (7) as follows

$$\begin{aligned} p(\mathbf{x}) &= \int_S \sigma(\mathbf{x}_0) G(\mathbf{x}, \mathbf{x}_0) dS(\mathbf{x}_0) \approx \sum_{m=1}^M \int_{s_m} \sigma(\mathbf{x}_0) G(\mathbf{x}, \mathbf{x}_0) dS(\mathbf{x}_0) \\ &\approx \sum_{m=1}^M \sigma(\mathbf{s}_m) G(\mathbf{x}, \mathbf{s}_m) S_m = \sum_{m=1}^M (\sigma(\mathbf{s}_m) S_m) G(\mathbf{x}, \mathbf{s}_m) = \sum_{m=1}^M Q_m G(\mathbf{x}, \mathbf{s}_m) \end{aligned} \quad (10)$$

This equation is a valid approximation when $M \rightarrow \infty, S_m \rightarrow 0$. The sound field produced by actual source can be equivalent to a collection of the point sources, each with appropriate strength determined by additional conditions. This idea is originally from that the acoustic field of a complex radiator can be approximated as a superposition of fields generated by a collection of simple sources (monopole, dipole, or combination of both) distributed on the virtual surface, which is more tractable numerically than the original setting. In this thesis, a synthesis technique using

virtual source representation in HRTF is based on the forgoing idea the exterior sound field is transformed to an equivalent combination of finite virtual sources distributed inside the boundary. According to Eq. (10), the total sound pressure produced by these virtual sources is given by

$$p(\mathbf{x}) = \sum_{m=1}^M Q_m G(\mathbf{x}, \mathbf{s}_m) \quad (11)$$

where G 's are the free space Green's function as defined before, the subscript m signify the virtual sources. The \mathbf{x} are position vectors (i.e. field point position vectors), and M is the chosen number of virtual sources. Virtual sources are distributed in another virtual boundary which is away from the real one. Imposing additional conditions can make the unknown strength Q_m to be determined. Here, direct evaluation of the pressure gradient along the normal direction at the real boundary mesh node \mathbf{x} leads to

$$\frac{\partial p}{\partial n}(\mathbf{x}) = \sum_{m=1}^M Q_m \frac{\partial G}{\partial n}(\mathbf{x}, \mathbf{s}_m) \quad (12)$$

In Eq. (12) the term $\alpha\sigma(\mathbf{x})$ does not appear because of the field points never coincide with the source points and there is no velocity discontinuity in Eq. (8). The location vectors of the source points can be made always smaller than the ones of the boundary. Under this condition, there is no singularity in Eq. (12). Consequently, the singularity problems common to the boundary element formulation are avoided altogether. Then, choosing finite number of field points, \mathbf{x}_n , $n=1, 2 \dots N$, Eq. (11) and (12) can be rewritten into

$$\begin{bmatrix} p(\mathbf{x}_1) \\ p(\mathbf{x}_2) \\ \vdots \\ p(\mathbf{x}_N) \end{bmatrix} = \begin{bmatrix} G(\mathbf{x}_1, \mathbf{s}_1) & G(\mathbf{x}_1, \mathbf{s}_2) & \dots & G(\mathbf{x}_1, \mathbf{s}_M) \\ G(\mathbf{x}_2, \mathbf{s}_1) & G(\mathbf{x}_2, \mathbf{s}_2) & \dots & G(\mathbf{x}_2, \mathbf{s}_M) \\ \vdots & \vdots & \ddots & \vdots \\ G(\mathbf{x}_N, \mathbf{s}_1) & G(\mathbf{x}_N, \mathbf{s}_2) & \dots & G(\mathbf{x}_N, \mathbf{s}_M) \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \\ \vdots \\ Q_M \end{bmatrix}$$

and

$$\begin{bmatrix} \frac{\partial p(\mathbf{x}_1)}{\partial n} \\ \frac{\partial p(\mathbf{x}_2)}{\partial n} \\ \vdots \\ \frac{\partial p(\mathbf{x}_N)}{\partial n} \end{bmatrix} = \begin{bmatrix} \frac{\partial G(\mathbf{x}_1, \mathbf{s}_1)}{\partial n} & \frac{\partial G(\mathbf{x}_1, \mathbf{s}_2)}{\partial n} & \dots & \frac{\partial G(\mathbf{x}_1, \mathbf{s}_M)}{\partial n} \\ \frac{\partial G(\mathbf{x}_2, \mathbf{s}_1)}{\partial n} & \frac{\partial G(\mathbf{x}_2, \mathbf{s}_2)}{\partial n} & \dots & \frac{\partial G(\mathbf{x}_2, \mathbf{s}_M)}{\partial n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial G(\mathbf{x}_N, \mathbf{s}_1)}{\partial n} & \frac{\partial G(\mathbf{x}_N, \mathbf{s}_2)}{\partial n} & \dots & \frac{\partial G(\mathbf{x}_N, \mathbf{s}_M)}{\partial n} \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \\ \vdots \\ Q_M \end{bmatrix}$$

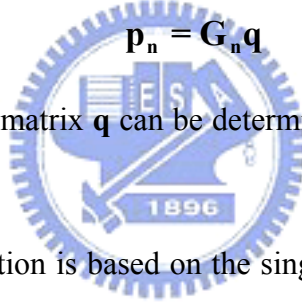
In matrix form

$$\mathbf{p} = \mathbf{G} \mathbf{q} \quad (13)$$

and

$$\mathbf{p}_n = \mathbf{G}_n \mathbf{q} \quad (14)$$

The unknown source strength matrix \mathbf{q} can be determined by solving a matrix inverse problem.



Virtual source representation is based on the single layer representation. There is another similar form based on hybrid combination of single and double layer potentials to improve the robustness of single layer representation. A hybrid combination of monopole and dipole sources is used in this method.^{19,20} This modification allows the inherent nonuniqueness of the method at certain frequencies to be overcome. It is possible to write Eq. (11) and (12) in terms of a single layer potential along with an associated expression for the pressure distribution in terms of a double layer potential. To eliminate the problem of uniqueness at critical wave numbers, the formulation for the superposition problem becomes

$$P(\mathbf{x}) = \int_S \sigma(\mathbf{x}_0) \left(G(\mathbf{x}, \mathbf{x}_0) + i \frac{1}{k} \frac{\partial G(\mathbf{x}, \mathbf{x}_0)}{\partial n_o} \right) dS(\mathbf{x}_0) \quad (15)$$

$$\frac{\partial p}{\partial n}(\mathbf{x}) = \int_S \sigma(\mathbf{x}_0) \left(\frac{\partial G(\mathbf{x}, \mathbf{x}_0)}{\partial n} + i \frac{1}{k} \frac{\partial^2 G(\mathbf{x}, \mathbf{x}_0)}{\partial n \partial n_o} \right) dS(\mathbf{x}_0) \quad (16)$$

With the same things as Eq. (10), Eq. (15) and (16) can also be approximated as

$$P(\mathbf{x}) = \sum_{m=1}^M Q_m \left(G(\mathbf{x}, \mathbf{s}_m) + i \frac{1}{k} \frac{\partial G(\mathbf{x}, \mathbf{s}_m)}{\partial n_o} \right) \quad (17)$$

$$\frac{\partial p}{\partial n}(\mathbf{x}) = \sum_{m=1}^M Q_m \left(\frac{\partial G(\mathbf{x}, \mathbf{s}_m)}{\partial n} + i \frac{1}{k} \frac{\partial^2 G(\mathbf{x}, \mathbf{s}_m)}{\partial n \partial n_o} \right) \quad (18)$$

With finite number of field points, $\mathbf{x}_n, n=1, 2 \dots N$, Eq. (17) and (18) can also be written as

$$\begin{bmatrix} p(\mathbf{x}_1) \\ p(\mathbf{x}_2) \\ \vdots \\ p(\mathbf{x}_N) \end{bmatrix} = \begin{bmatrix} G(\mathbf{x}_1, \mathbf{s}_1) + i \frac{1}{k} \frac{\partial G(\mathbf{x}_1, \mathbf{s}_1)}{\partial n_o} & G(\mathbf{x}_1, \mathbf{s}_2) + i \frac{1}{k} \frac{\partial G(\mathbf{x}_1, \mathbf{s}_2)}{\partial n_o} & \dots & G(\mathbf{x}_1, \mathbf{s}_M) + i \frac{1}{k} \frac{\partial G(\mathbf{x}_1, \mathbf{s}_M)}{\partial n_o} \\ G(\mathbf{x}_2, \mathbf{s}_1) + i \frac{1}{k} \frac{\partial G(\mathbf{x}_2, \mathbf{s}_1)}{\partial n_o} & G(\mathbf{x}_2, \mathbf{s}_2) + i \frac{1}{k} \frac{\partial G(\mathbf{x}_2, \mathbf{s}_2)}{\partial n_o} & \dots & G(\mathbf{x}_2, \mathbf{s}_M) + i \frac{1}{k} \frac{\partial G(\mathbf{x}_2, \mathbf{s}_M)}{\partial n_o} \\ \vdots & \vdots & \ddots & \vdots \\ G(\mathbf{x}_N, \mathbf{s}_1) + i \frac{1}{k} \frac{\partial G(\mathbf{x}_N, \mathbf{s}_1)}{\partial n_o} & G(\mathbf{x}_N, \mathbf{s}_2) + i \frac{1}{k} \frac{\partial G(\mathbf{x}_N, \mathbf{s}_2)}{\partial n_o} & \dots & G(\mathbf{x}_N, \mathbf{s}_M) + i \frac{1}{k} \frac{\partial G(\mathbf{x}_N, \mathbf{s}_M)}{\partial n_o} \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \\ \vdots \\ Q_M \end{bmatrix}$$

and

$$\begin{bmatrix} \frac{\partial p(\mathbf{x}_1)}{\partial n} \\ \frac{\partial p(\mathbf{x}_2)}{\partial n} \\ \vdots \\ \frac{\partial p(\mathbf{x}_N)}{\partial n} \end{bmatrix} = \begin{bmatrix} \frac{\partial G(\mathbf{x}_1, \mathbf{s}_1)}{\partial n} + i \frac{1}{k} \frac{\partial^2 G(\mathbf{x}_1, \mathbf{s}_1)}{\partial n \partial n_o} & \frac{\partial G(\mathbf{x}_1, \mathbf{s}_2)}{\partial n} + i \frac{1}{k} \frac{\partial^2 G(\mathbf{x}_1, \mathbf{s}_2)}{\partial n \partial n_o} & \dots & \frac{\partial G(\mathbf{x}_1, \mathbf{s}_M)}{\partial n} + i \frac{1}{k} \frac{\partial^2 G(\mathbf{x}_1, \mathbf{s}_M)}{\partial n \partial n_o} \\ \frac{\partial G(\mathbf{x}_2, \mathbf{s}_1)}{\partial n} + i \frac{1}{k} \frac{\partial^2 G(\mathbf{x}_2, \mathbf{s}_1)}{\partial n \partial n_o} & \frac{\partial G(\mathbf{x}_2, \mathbf{s}_2)}{\partial n} + i \frac{1}{k} \frac{\partial^2 G(\mathbf{x}_2, \mathbf{s}_2)}{\partial n \partial n_o} & \dots & \frac{\partial G(\mathbf{x}_2, \mathbf{s}_M)}{\partial n} + i \frac{1}{k} \frac{\partial^2 G(\mathbf{x}_2, \mathbf{s}_M)}{\partial n \partial n_o} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial G(\mathbf{x}_N, \mathbf{s}_1)}{\partial n} + i \frac{1}{k} \frac{\partial^2 G(\mathbf{x}_N, \mathbf{s}_1)}{\partial n \partial n_o} & \frac{\partial G(\mathbf{x}_N, \mathbf{s}_2)}{\partial n} + i \frac{1}{k} \frac{\partial^2 G(\mathbf{x}_N, \mathbf{s}_2)}{\partial n \partial n_o} & \dots & \frac{\partial G(\mathbf{x}_N, \mathbf{s}_M)}{\partial n} + i \frac{1}{k} \frac{\partial^2 G(\mathbf{x}_N, \mathbf{s}_M)}{\partial n \partial n_o} \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \\ \vdots \\ Q_M \end{bmatrix}$$

In matrix form,

$$\mathbf{p} = \mathbf{D} \mathbf{q} \quad (19)$$

and

$$\mathbf{p}_n = \mathbf{D}_n \mathbf{q} \quad (20)$$

Every entry in matrix \mathbf{D} and \mathbf{D}_n can be written as

$$D_{ij} = \left\{ G(\mathbf{x}_i, \mathbf{s}_j) + i \frac{1}{k} G'(\mathbf{x}_i, \mathbf{s}_j) \left[\frac{\mathbf{r} \cdot \mathbf{gn}_j}{r} \right] \right\}$$

$$(D_n)_{ij} = G'(\mathbf{x}_i, \mathbf{s}_j) \frac{\mathbf{r} \cdot \mathbf{g}_i}{r} + i \frac{1}{k} \left[\frac{\mathbf{r} \cdot \mathbf{g}_i}{r} \frac{\mathbf{r} \cdot \mathbf{g}_j}{r} (G''(\mathbf{x}_i, \mathbf{s}_j)) - \frac{G'(\mathbf{x}_i, \mathbf{s}_j)}{r} \right] - \frac{G'(\mathbf{x}_i, \mathbf{s}_j)}{r} \mathbf{n}_i \cdot \mathbf{g}_j$$

$$\mathbf{r} = \mathbf{r}_i - \mathbf{r}_j, \quad r = |\mathbf{r}|$$

The specification of boundary conditions plays an important role in solving unknown strength of virtual source. The boundary conditions (i.e. surface normal velocity) can be set into rigid and locally reacting areas respectively. According to the momentum equation, the surface normal velocity can be written as the matrix as follows

$$\begin{bmatrix} \frac{\partial p(\mathbf{x}_1)}{\partial n} \\ \frac{\partial p(\mathbf{x}_2)}{\partial n} \\ \vdots \\ \frac{\partial p(\mathbf{x}_N)}{\partial n} \end{bmatrix} = \begin{bmatrix} -j\rho_0\omega u \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (21)$$

In this matrix, the surface normal velocity of locally reacting area is set to be $-j\rho_0\omega u$ and that of rigid area is set to be zero. u is the particle velocity and ρ_0 is the density of air at room temperature. With the decision of boundary conditions, the volume velocity (i.e. source strength) of every virtual source can be calculated via a matrix inversion problem. Once the source strength is known, the acoustic field of desired field points also can be approximated as a superposition of fields generated by a collection of these virtual sources. Substitute the calculated volume velocity to Eqs. (11) and (17), pressure of desired field points can be calculated finally.

C. Technical issues of virtual source representation

1. Optimization of virtual source representation

The relative position of virtual source surface to that of the boundary influences

the computation accuracy very much.¹⁸ For a fixed N , the locations of simple sources inside the radiator can be adjusted to improve the computation accuracy. The parameter (τ) which is the relative position of virtual sources to mesh boundary is varied from nearly zero to nearly one time the characteristic length of the boundary. For instances, a distribution of spherical boundary nodes also have the same distribution of virtual sources inside the boundary except the sphere radius. According to Eq. (10), the computation accuracy can be improved as τ becomes smaller. There is an optimum value of τ in locating the virtual sources inside the boundary. In order to choose optimum position of virtual sources inside the mesh boundary, a measure of solution accuracy of volume velocity solved by matrix inversion is needed. When the boundary conditions are decided, the measure of accuracy of the virtual source representation can be the extent of the reconstruction of boundary normal velocity by virtual source distribution.²¹ The reconstructed surface normal velocity is given by



$$\hat{\mathbf{p}}_n = \mathbf{G}_n \mathbf{G}_n^{-1} \mathbf{p}_n \quad (22)$$

for single layer representation and

$$\hat{\mathbf{p}}_n = \mathbf{D}_n \mathbf{D}_n^{-1} \mathbf{p}_n \quad (23)$$

for hybrid combination of single and double layer potentials. There will however be some difference in the prescribed boundary condition and calculated surface normal velocity. A measure of this difference is the surface velocity reconstruction error norm,

$$\mathbf{e} = \|\mathbf{p}_n - \hat{\mathbf{p}}_n\| \quad (24)$$

This error norm will be closely linked to the error in the boundary surface pressure and can be used to select the optimum relative position of mesh boundary to virtual sources. The value of τ under the minimum sum of error norms at each frequency

is chosen to set the optimum position of virtual sources.

2. Regularization of virtual source representation

In Eqs. (14) and (20), the condition number of matrix \mathbf{G}_n is always large under small τ and this means that matrix \mathbf{G}_n is very ill conditioned. The condition number of \mathbf{G}_n can be expressed as the ratio of the largest and the smallest singular value. This quantity measures the perturbations in the matrix itself as well as on the right hand side. Singular or ill-conditioned matrices contain null or small singular values. This results in infinite or huge terms in the inverse matrix. Whereas the infinite terms clearly lead to an infinite solution, the huge terms lead to a solution that will be very sensitive to the right hand side variations. Therefore, small errors on the right hand side will result in huge errors on the solution. To avoid this problem, the small or null singular values need to be dropped. Consequently, a straightforward solution of the inverse problem is not satisfactory, and special regularization technique should be employed in order to solve Eq. (14) and (20). Here one regularization technique based on singular value decomposition (SVD) is used. In MATLAB technical language, a command function '**pinv**'²² can be used to solve volume velocity vector of virtual sources with tolerance regularization. In the regularization via '**pinv**', the SVD of matrix is taken and then any singular values of \mathbf{G}_n less than tolerance are treated as zero.²² There is also another method called Tikhonov regularization to do matrix regularization here.²³ Tikhonov regularization is one way to solve the ill-conditioned problem in a nearby sense with a constraint on the size of the sought solution. The solution size is linked to the solution smoothness, because ignoring components with a high index will lead to a smooth solution with reasonable size as measured by an appropriate norm. In mathematical form, this can be written as a minimization problem, where the functional to be minimized involves a residual norm

and a discrete smoothing norm $\|\mathbf{Lx}\|_2$ acting on the solution vector:

$$\min_{\mathbf{x}} \{ \|\mathbf{Ax}-\mathbf{b}\|_2^2 + \beta^2 \|\mathbf{Lx}\|_2^2 \} \quad (25)$$

This form of regularization is known as Tikhonov regularization. The regularization parameter β is the only input when a suitable smoothing norm has been defined. A large β favors a small solution size at the cost of a large residual norm. Hence, β controls the degree to which the sought regularized solution should fit to the measured data in \mathbf{b} or have a small solution size. There is a trade-off between the residual norm and the solution size. For the choice of $\mathbf{L}=\mathbf{I}_n$, the Tikhonov regularized solution can be written conveniently in terms of the SVD of \mathbf{A} as

$$\mathbf{x}_\beta = \sum_{i=1}^n \frac{\sigma_i^2}{\sigma_i^2 + \beta^2} \frac{\mathbf{u}_i^H \mathbf{b}}{\sigma_i} \mathbf{v}_i \quad (26)$$

where \mathbf{u}_i and \mathbf{v}_i are called the left and right singular vectors and σ_i is the singular value. In practice, the regularization parameter is chosen to lie between the highest and smallest singular value and the filter gradually dampens singular values smaller than β . The approach to estimate the regularization parameter in Tikhonov regularization is called the L-curve criterion. The idea behind the L-curve is to plot the discrete smoothing norm of the regularized solution versus the residual norm in log-log scale for all regularization parameter. The generic form of the L-curve is plotted for all valid regularization parameters in Fig. 12. The smoothing norm here is $\|\mathbf{Lx}_\beta\|_2$ and the residual norm is $\|\mathbf{Ax}_\beta-\mathbf{b}\|_2$, i.e. the two terms occurring in the functional to be minimized in Tikhonov regularization. The horizontal part of the curve is characterized by solutions that have been smoothed too much (over-regularized), whereas the vertical part is characterized by solutions dominated by the effects of errors (under regularized). In between two parts, the solutions representing a balance between fitting the solution to data and keeping the solution's

smoothing norm small can be found. The optimum value of regularization parameter defined in L-curve criterion at the corner of L-curve.

D. Pre-processing of Computational mesh model

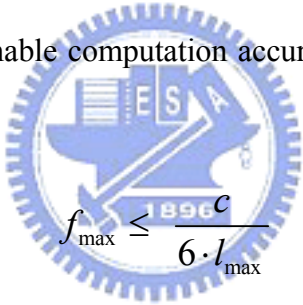
In this thesis, a “**real**” head mesh of KEMAR is chosen to be the mesh model (i.e. mesh boundary) in virtual source representation and IBEM. The pre-processing of computational mesh model includes several steps – 1) Acquiring the mesh model, 2) Mesh decimation, 3) Mesh topology construction, 4) Mesh data transformation. The final numerical result is influenced by mesh manipulation deeply. Mesh manipulation can make calculation more efficient and increase the computation accuracy.

1. Acquiring the mesh model

A real human head is an ideal mesh model to carry out HRTF numerical computation. In the above mentioned HRTF measurements, the HRTFs of KEMAR dummy head are measured and many databases are already constructed. In order to get more convenient verifications for numerical results, the KEMAR dummy head with left (DB60) and right (DB65) pinnae is chosen to be the computational mesh model. In the HRTF numerical computation, the computation is focused on the head with two pinnae and the torso is ignored. A detailed acquisition of mesh data is proceeding via three 3D laser scanners provided by Industrial Technology Research Institute (ITRI). Three laser scanners at different positions capture spatial information of KEMAR dummy head at the same time and transform them into 3D node coordinate data via an image processing software-*TriD*. Total mesh nodes close to one hundred and sixty thousands are captured in KEMAR dummy head.

2. Mesh decimation

Mesh decimation plays an important role in HRTF numerical computations. Computation loadings will be proportional to the number of mesh nodes in virtual source representation implemented by MATLAB and IBEM implemented by LMS SYSNOISE. In order to look after both computation accuracy and computation loadings, mesh decimation is a necessary and important step in mesh pre-processing. A reasonable number of mesh nodes can be estimated by the number of elements per wavelength in IBEM. In Fig. 13, seven points can be spaced-out from start to finish of the wave in order to identify the locations of peaks. The default value in SYSNOISE is six elements per wavelength and there will be a trade-off between modeling accuracy and model size if the number of elements is less than six. The maximum frequency of reasonable computation accuracy in IBEM can be calculated by


$$f_{\max} \leq \frac{1896c}{6 \cdot l_{\max}} \quad (27)$$

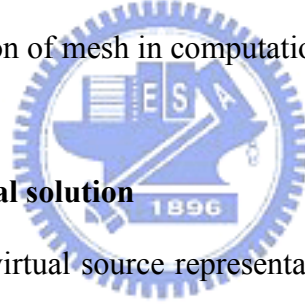
c is the speed of sound at room temperature. f_{\max} is the estimated maximum frequency and l_{\max} is the maximum element length.

3. Mesh topology construction

Before the transformation of the scanned mesh data, the topology of each mesh node must be constructed. Different types of mesh grid can be chosen when the processing the mesh data in *TriD*. In HRTF computations, triangular mesh grids are chosen to be the mesh elements in IBEM. The topology of each mesh element is especially important in calculating the surface normal velocity in virtual source representation. Here, topology information of each element can be used to revise each virtual source scattering direction to the correct normal direction of surface mesh node.

4. Mesh data transformation

Before importing mesh data to MALAB or SYSNOISE, exact data type is essential to start the computation. Especially in SYSNOISE, there is a complete two-way interface with formatted data or result is called the SYSNOISE free format. Only the node coordinates and element topologies contained in the file are considered under this format. The original data type of scanned mesh is universal Wavefront OBJ (object) which contains much more mesh information than the requirements in free format of SYSNOISE. A data transformation program is written by MALAB to transform important node coordinates and element topologies into SYSNOISE free format. In importing mesh data to MATLAB, the polygonal data in ASCII form of Wavefront OBJ gives a great help. The file can be read directly and just discard the unnecessary spatial information of mesh in computation.



E. Verification of Analytical solution

In order to validate the virtual source representation, verification to the solution of circular piston in a spherical baffle is implemented by IBEM and virtual source representation. In this case, the piston is placed at the pole of a sphere which radius is a and the circular piston of radius α set vibrates with a velocity W_0 . The z-axis passes through the center of the piston. A simple plot is shown in Fig. 14. The velocity on the sphere (independent of ϕ) is

$$w(a, \theta) = \begin{cases} W_0, & 0 \leq \theta \leq \alpha \\ 0, & \alpha \leq \theta \leq \pi \end{cases} \quad (28)$$

The radiation pressure can be described as²⁴

$$p(r, \theta) = \frac{i\rho_0 c W_0}{2} \sum_{n=0}^{\infty} [P_{n-1}(\cos \alpha) - P_{n+1}(\cos \alpha)] \frac{h_n(kr)}{h_n'(kr)} P_n(\cos \theta), \quad (29)$$

Where P_{n-1} and P_{n+1} are the Legendre polynomials, h_n is spherical Hankel function

and c is the sound velocity at room temperature. Here all the objective conditions (i.e. vibrating velocity, α and a) are set the same in analytical solution, IBEM and virtual source representation. In virtual source representation, the sphere is discretized into total 182 mesh nodes. There are 13 mesh nodes distributed uniformly at every 5 degrees from elevation 90 to 80 and -90 to -80 at each pole, and other 156 mesh nodes are distributed uniformly every 30 azimuth degrees at every 10 elevation degrees. The virtual source distributions are the same as the mesh sphere except the radius of virtual source sphere. If the area of each segment S_m in integrand of Eq. (10) becomes smaller, the accuracy will improve as expected. In order to get high accuracy in computation, the radius of virtual sphere is set to be the 15% of the mesh sphere. Under this condition, the distance between two neighbor sources are very close and the reasonable maximum frequency can be achieved to 1.09 kHz according to the four elements per wavelength criterion. In SYSNOISE, there are total 357 mesh nodes and 650 triangular elements in this computation. The maximum reasonable frequency is 1064 Hz estimated by the six elements per wavelength criterion. Two different mesh settings of IBEM and virtual source representation are shown in Fig. 15 and Fig. 16. The sphere is discretized into 182 circle points in Fig. 16. In virtual source representation, criterions of matrix regularization and optimization mentioned above are also applied to this analytical solution verification. First, according to the minimum velocity error criterion, the calculated optimum τ of mesh sphere and virtual source sphere is 0.89. The radius of virtual sphere changes to 0.89 times of the mesh sphere. In this case, how to get a “real” optimum τ becomes a complexity problem because of variable regularization parameter at each frequency. In order to simplify the calculation of optimum τ , a stationary regularization parameter based on L-curve criterion is applied to every frequency before using the minimum velocity error criterion to find optimum ratio.

The calculated result with optimum ratio is shown in Fig. 17 and it demonstrates that computation result is not as good as expected. One important reason to explain this phenomenon is that optimum ratio is not small enough to satisfy the Eq. (10). Second, a virtual sphere radius which is 0.15 times of external one is also set. The average regularization parameter of this case is set to be 10^{-6} of first singular value of \mathbf{G}_n by the L-curve criterion respectively. Fig. 18 shows the L-curve at 100 Hz. In fig. 18 (a), the ambiguous L-curve can not be used to find optimum regularization parameter. In order to find optimum regularization parameter, a cost function which can be set as

$$\gamma = \text{Smoothing norm} + \text{Residue norm} \quad (30)$$

Figure 18 (b) shows the distribution of each value of cost function. An obvious generic form of L-curve can be seen in this figure. The values (γ) on the right of the corner are dominated by the residue norm and the left ones are dominated by the smoothing norm. The value at the corner is taken to be the optimum regularization parameter β . The numerical results calculated by IBEM and virtual source representation are compared to analytical solution here. Under the same objective conditions, the radiation pressures of the field point up the center of mesh sphere 1.2 meters (i.e. ipsilateral side) calculated by different numerical methods are shown in Fig. 19. In this figure, the results calculated by numerical methods are all close to the analytical solution. Directional responses at different frequencies are shown in Fig. 20 (a)-(d). The directivity is increasingly obvious in the ipsilateral side of the sphere with increasing the frequency. In the verification to analytical solution, virtual source representation and IBEM both work well under reasonable frequency range and the computation accuracies are also good enough.

V. EXPERIMENTAL INVESTIGATIONS AND NUMERICAL ANALYSIS

In this section, HRTF numerical computations are implemented by the virtual source representation and IBEM. Besides the numerical computations, HRTF experimental results become the important analysis reference in frequency response, impulse response and directivity pattern. In the previous section, IBEM and virtual source representation worked well in analytical solution verification. Here the specification boundary conditions in HRTF computation of virtual source representation are set to be constant acceleration which is different from constant velocity setting in analytical verification case. This setting of boundary conditions can be explained by analogy with a simple mass controlled loudspeaker. Under this condition, constant acceleration is set when the constant force is applied. The positions of the field points and scattering source are also exchanged according to the acoustic reciprocity in HRTF computation. With the acoustic reciprocity, the received signal can be remain the same even though the locations of field point and sound source are interchanged in an unchanging environment. Subject to the acoustic reciprocity, the location of sound source changes close to the ear hole and set it as a constant acceleration vibrating panel both in IBEM and virtual source representation. The settings of KEMAR mesh in SYSNOISE are shown in Fig. 21. The light color areas are denoted as the constant acceleration vibrating panels and total 10776 mesh nodes and 21529 elements are used in computation. There is also another KEMAR mesh data given by Kahana for HRTF computation. In Kahana's KEMAR mesh, the resolution of mesh grids is much more detailed than the one given by ITRI. Only half of the KEMAR head is given in Kahana's KEMAR mesh and there are 12137 mesh nodes and 23799 mesh elements. In Kahana's HRTF computation, symmetry of the head is assumed and the variation between the ears

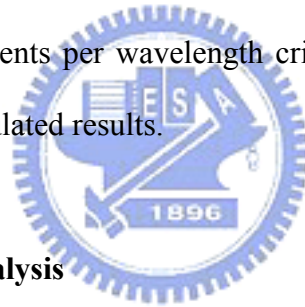
does not affect the response at only one of them. Fig. 22 shows the Kahana's KEMAR head mesh settings.

There are some differences in the setting of mesh in IBEM and virtual source representation. The nodes of head mesh except the pinna imported to MATLAB are first decimated to 2% of the ITRI head mesh and the nodes of pinna mesh are decimated to one sixth of the original ones. The total mesh nodes in HRTF of virtual source representation are 581 points for whole head and this is classified to virtual source case 1. There is another mesh setting called virtual source case 2 which the nodes of head mesh are just only the one tenth of the original ones and nodes distributed on pinna are half of the original ones. The total mesh nodes in virtual source case 2 are 2074 points for whole head mesh. The number of mesh nodes is limited to no more than 2500 because of the limitation in SVD of MATLAB. Fig. 21 (a)-(d) show the mesh settings of virtual source case 1 and case 2 in the MATLAB respectively. The location of constant acceleration source is also set around the ear hole. In order to compare to the measured HRTFs, the distances between field points and sound source are chosen to be 1.2 meters in both numerical methods. In virtual source representation, the sound sources are changed to constant acceleration ones so that the boundary conditions of the mesh surface changes to

$$\begin{bmatrix} \frac{\partial p(\mathbf{x}_1)}{\partial n} \\ \frac{\partial p(\mathbf{x}_2)}{\partial n} \\ \vdots \\ \frac{\partial p(\mathbf{x}_M)}{\partial n} \\ \frac{\partial p(\mathbf{x}_N)}{\partial n} \end{bmatrix} = \begin{bmatrix} -j\rho_0\omega \frac{a}{j\omega} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} -\rho_0 a \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (31)$$

The boundary conditions of the vibrating panels are set as $-\rho_0 a$. The distributions of virtual sources in Fig. 23 are the retracted version of the imported mesh inside the

head mesh. The value of retracted ratio is set as 0.1 of the external mesh to satisfy Eq. (10). Tikhonv regularization is also applied to HRTF computations. Depending on the experiences in previous section, the retracted ratio of the virtual sources to external mesh must be small enough to get good accuracy. An average regularization parameter is also decided by the simplified L-curve criterion mentioned before. There are also two L-curves under the present mesh settings at different frequencies in Fig. 24. The average regularization parameter is applied to each frequency in computation and there are three different parts of calculated results which will be discussed in the following. Here the maximum reasonable frequency of IBEM using Kahana's KEMAR mesh can be achieved to 10 kHz according to the six elements per wavelength criterion. The other cases can also be achieved to 10 kHz with a coarser two elements per wavelength criterion and this means the error will be occurred in these calculated results.



A. Frequency response analysis

The measured HRTFs of every 30 azimuth degrees on the horizontal plane are shown in Fig. 10 and there are two important and common characteristics in each sub plot. The results calculated by SYSNOISE of two different resolution meshes (i.e. ITRI and Kahana cases) are shown in Fig. 25 (a)-(d) and Fig. 26 (a)-(d) and the one by MATLAB is shown in Fig. 27 and Fig. 28. In measured HRTFs, every plot shows that the KEMAR with DB65 ear has an obvious first peak around 5 kHz and a deep notch around 9 kHz. These are obvious and important characteristics of HRTFs with DB65 under 10kHz. In numerical calculated results, four HRTFs from 0 to 270 degree with 90 azimuth degrees interval at the horizontal plane are taken to be compared to measured ones. First the calculated results of IBEM using Kahana's symmetric KEMAR mesh are shown with the measured ones in Fig. 26. The

numerical results only show a good match to measured ones at the ipsilateral side because of the symmetry settings. In Fig. 26 (b), a few unknown resonances are also shown in the Fig. below 4 kHz. This phenomenon is because of the non-unique solutions caused by the excitation frequency which corresponds to a resonance frequency of the associated interior problem in exterior problem when using IBEM. In the results using Kahana's KEMAR mesh, calculated HRTF at 270 azimuth degrees is the same as 90 azimuth degrees because of the symmetry calculation. In Fig. 25, the calculated results can not match the envelope of the measured ones well. The errors have already been done by the non-uniform mesh elements. The errors are resulted from the coarser mesh elements. There are 20% mesh elements which the estimated maximum reasonable frequency only can be achieved to 2 kHz and other 80 % mesh elements the estimated maximum ones can be achieved to 7 kHz under the six elements per wavelength criterion. In all plots of the Fig. 25, the calculated results can still catch the characteristics of the measured ones and head shadowing effect can be seen at the 270 azimuth degrees. The results calculated by virtual source representation are also shown in Fig. 27 and Fig. 28. In the (b) of both Fig.27 and Fig. 28, well-fitted envelopes can be seen especially in virtual source case 2. Other calculated results can not match the measured results well because of the errors made by fewer nodes of head mesh. These calculated results also prove that virtual source representation works in HRTF computations.

B. Directional response

Directional response is a good index to obtain the head shadowing effect. The directional responses of measured HRTFs at four different frequencies are shown in Fig. 29 (a)-(d). The patterns show clearly that the sound source is omni-directional at low frequency. Under this situation, interaural time difference becomes an

important cue to the localization ability of people. The wavelength is much longer than the size of head at low frequency and there is hardly any difference in sound pressure at the two ears. Only the interaural time difference can make the human auditory systems detect the waveforms unambiguously at low frequency. From the directional responses of measured HRTFs, the direction of the main lobe can also be explained as the normal direction of the external ear. The main lobe is focused on the fewer azimuth degrees than the expected angle (i.e. 90 azimuth degrees). This phenomenon is resulted from the effect of external ear. The calculated directional responses are also shown in the Fig. 30, Fig. 31, Fig. 32 and Fig.33 for IBEM Kahana case, IBEM ITRI case, virtual source case 1 and case 2. Only half of the directivity pattern is shown in IBEM Kahana case because of just using the half of KEMAR dummy head. In all calculated directional responses, main lobes focused on the ipsilateral sides clearly at high frequency. The differences between the measured HRTF and calculated ones are resulted from the computational error. Great parts of errors are resulted from the decimated nodes of KEMAR mesh in computation. However, a clear general direction can be found that every numerical result focused on the reasonable direction and the head shadowing effect can also be clearly explored. In the directional responses calculated by using virtual source representation, the unexpected leakage at the contra-lateral side are imputed to the decimated head mesh points and numerical errors.

C. Impulse response

A comparison of HRIRs here is a way to obtain the ILD and ITD. HRIRs at the ipsilateral (90 degrees) and contra-lateral (270 degrees) sides are good indexes in validating the computation results. Fig. 34 shows the measured HRIRs at 90 azimuth degrees and 270 azimuth degrees. The time differences between them are

about 1 ms and can be converted into 0.343 meters with 343 meters per second sound speed at room temperature. A clear level difference can also be obtained between these two HRIRs and this is resulted from the head shadowing effect. There are three other comparisons of HRIRs by different numerical methods in Fig. 35, Fig. 36 and Fig. 37. Reasonable ITD and ILD can be seen in these Fig. 35 and the time differences can be converted into a distance which should be close to the diameter of the KEMAR head. The calculated results by virtual source representation are different from others. ITD seems not to be included in the calculated results. Nevertheless, there is one important reason to explain this phenomenon. The distributions of virtual sources affect the ITD deeply in virtual source distribution. The virtual sources are the retracted version of external mesh. In order to satisfy Eq. (10), the retracted ratio between virtual sources and external mesh was set to be 0.1. All virtual sources are close to each other and the distances between field points and all virtual sources are almost the same when the field points are far away relative to the virtual sources. Under this situation, the field points at 90 azimuth degrees and 270 azimuth degrees get the close distance to virtual sources and the equivalent distances make ITD ambiguously in the calculated HRIRs. However, a clear ILD can still be obtained in the Fig. 36 and Fig. 37. The comparison of HRIRs is not shown in IBEM Kahana case because that symmetry conditions are set in computation and the HRIR at 270 azimuth degrees is the same as the one at 90 azimuth degrees. Other numerical results except the symmetric IBEM Kahana case can be found that important ITD and ILD cues are included.

D. Analysis of computation loading

In the IBEM of using Kahana's KEMAR mesh, two times finer and uniform mesh was set in HRTF computations. Huge memory loading and long computation

time were paid to get nearly the same HRTFs with measured ones. Table II lists the relation of the computation loading of two numerical methods in this thesis. The computation time, the number of mesh nodes and memory loading are all listed in this table. In the most advanced PC, which is Pentium 4 3 GHz equipped with 1 GB DDR ram, about fifty-one and thirty-five hours computation time must be taken in order to finish the 52 frequency points in SYSNOISE. Not only the long computation time but also the heavy memory loadings which take 897 and 879 MB are both occurred in IBEM Kahana case and ITRI case respectively. Here, the computation time in the decimated distributions of mesh nodes using virtual source representation is much less than the IBEM. The maximum memory loading is always restricted and not exceed 100 MB in virtual source case 1. Obviously the number of nodes in the IBEM is much more than the one in virtual source representation and computation time also take more than sixty times of case 1 of virtual source representation. Nevertheless, the calculated results between these two numerical methods get the close envelopes in HRTF computation especially in the calculated results of ITRI case and two different cases of virtual source representation and this means a much more computational saving can be taken by using virtual source representation.

E. Subjective experiment

A subjective experiment was carried out using the two different numerical results, the measured HRTF and the MIT HRTF databases. The calculated results of IBEM using Kahana's KEMAR mesh can not be used to subjective experiment because of lacking the HRTFs of contra-lateral sides. This is resulted from only half of the KEMAR head which is used in computation. The test stimuli were generated by using a random noise and the test locations were focused on the horizontal plane.

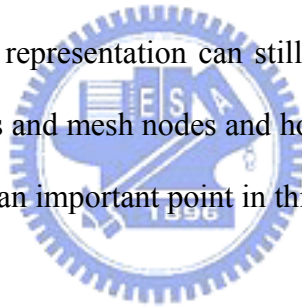
Only azimuth was investigated in the experiments²⁵. Twelve directions with a 30° interval on the horizontal plane were examined in azimuth test. Each signal in the tests was played for 5 seconds. The signals were played through the ATH-PRO6 monitor headphones connected directly to the soundcard of a personal computer. Here the perceived angle versus the presented angle in the localization tests is shown in Fig. 38 (a)-(e). The area of each circle is proportional to the number of the listeners who localized the same perceived angle. The 45° dash line indicates that the presented angle coincides with the perceived angle. Two other dot lines show that the presented sound images are incorrectly localized at symmetric direction, which is generally referred as the “front-back reversal”. Figure 38 (a)-(e) show the results of azimuth sound localization using two different numerical results and two different HRTF databases. Good agreement can be found between the perceived and presented directions when two measured HRTF are used. The discrepancies of the presented and perceived directions are often occurred on the front-back directions. The other three plots show the experimental results using calculated HRTFs, which are the case 1 and case 2 of virtual source representation and ITRI case of IBEM. The distributions of circles show a not bad agreement of the perceived and presented directions. The performance by using IBEM is almost the same as the measured HRTF databases and great parts of discrepancies are resulted from the effect of front-back reversal. Compared to the experiment using measured HRTF databases, the performance virtual source case 1 is not as good as IBEM. Some discrepancies in the experimental results are the ambiguity of the left and right sides occurred to some people. Although this only occurred to a few people, these serious errors are resulted from the ambiguous ITD made by small retracted ratio of virtual sources. In Fig. 38 (e), the experimental result of virtual source case 2 is as good as the measured ones and gets a much better than virtual source case 1. The results of sound

localization are improved much more by increasing the mesh nodes. Figure 39 (a)-(e) also show the perceived average errors in each presented directions and average total errors. As expected, the minimum total average error occurs in the measured HRTF and the total average error of ITRI case of IBEM is also close to the measured ones. The total average error of the measured HRTF database and IBEM are all between 20 and 40 degrees. In the subjective experiment, this means that the average discrepancy error of directions is only about one interval. The total average error in the virtual source representation is 59 degrees which is almost equivalent to three interval errors in the experiment. With increasing the mesh nodes in virtual source case 2, the total average error is improved much more and the performance is almost the same as measured HRTF. This result proves that the performance can be improved much more in virtual source representation with increasing the mesh nodes.

VI. CONCLUSIONS

Two different numerical methods are compared in this work. One is the simple numerical method, virtual source representation. This method is motivated by the layer potential theory. The acoustic field of arbitrarily shaped radiators is described by using the principle of wave superposition in this method. The other is indirect boundary element method which is implemented by using a Vibro-Acoustic software-LMS SYSNOISE. The measurement of HRTFs is also implemented by using one simple measurement mechanism here. The goal of this investigation is to calculate the HRTFs of each azimuth and elevation efficiently. Both of these numerical methods can work and the calculated results which can catch the obvious characteristics of HRTFs. In the comparison between the numerical results and the measured HRTF databases, important characteristics of KEMAR dummy head with DB65 pinna can be captured by the two numerical methods in the frequency response

analysis. In directional response, a reasonable and clear direction was focused by the main lobe in all numerical results. Obvious ILD and ITD can also be found in the calculated HRIRs. In the process of HRTF computation, large efforts were paid in processing the ill-conditioned matrix in virtual source representation. By using the Tikhonov regularization in virtual source representation, the suitable regularization parameters can be selected. Appropriate regularization to the ill-conditioned matrix can make the numerical results close to the measured ones. Although both the numerical results are close to the measured ones in IBEM and virtual source representation, the computation loadings are very different between these methods. The computation loadings in virtual source representation are much lighter than the IBEM in memory use and computation time. However, the errors made by using virtual source representation can still be improved by increasing the number of both virtual sources and mesh nodes and how to process the ill-conditioned matrix effectively will be still an important point in this method.



REFERENCES

- ¹ Dolby Laboratories, Inc., “5.1-Channel production guidelines.”
(www.dolby.com)
- ² J. W. Strutt (Lord Rayleigh), “On our perception of sound direction,” *Philosophical Mag.* **13**, 214-232(1907).
- ³ B. Gardner and K. Martin, “HRTF measurements of KEMAR dummy-head microphone,” MIT Media Lab (1994).
(<http://sound.media.mit.edu/KEMAR.html>)
- ⁴ V. R. Algazi, R. O. Duda, D. M. Thompson and C. Avendano, “The CIPIC HRTF Database,” *Proc. 2001 IEEE Workshop on Applications of Signal Processing to Audio and Electroacoustics*, Mohonk Mountain House, New Paltz, NY, pp. 99-102, Oct. 21-24 (2001).
- ⁵ Suzuki Laboratory(Acoustic Information Systems), Tohoku University (2001).
(<http://www.itakura.nuee.nagoya-u.ac.jp/>)
- ⁶ D. W. Batteau, “The role of the pinna in human localization,” *Proc. Royal Society of London* **168**(B), 158-180(1967).
- ⁷ C. Phillip Brown and Richard O. Duda, “A Structural Model for Binaural Sound Synthesis”, *IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING*, **6**, 476-488, 1998
- ⁸ V. R. Algazi, R. O. Duda, R. Duraiswami, N. A. Gumerov, and Z. Tang, “Approximation the head-related transfer function using simple geometric models of the head and torso”, *J. Acoust. Soc. Am.* **112**, 2053-2064 (2002).
- ⁹ R. O. Duda, “3-D audio for HCI,” Department of Electrical Engineering, San Jose State University (2000).
(http://www.engr.sjsu.edu/~knapp/HCIROD3D/3D_home.htm)
- ¹⁰ Y. Kahana, “Numerical modeling of the head related transfer function”, a thesis

- submitted for the degree of doctor philosophy, University of Southampton, Faculty of engineering and applied science, Institute of sound and vibration research, 2000.
- ¹¹P. A. Nelson, Y. Kahana, “Spherical harmonics, singular-value decomposition and the head-related transfer function”, J. Sound and Vib., **239**, 607-637 (2001).
- ¹²LMS SYSNOISE release notes and getting started manual.
- ¹³LMS SYSNOISE user manual.
- ¹⁴E. A. G. Shaw, “Transformation of sound pressure level from the free-field to the eardrum in the horizontal plane”, J. Acoust. Soc. of Am. **56**, 1848-1861 (1974).
- ¹⁵J. Chen, B. D. Van Veen, and K. E. Hecox, ”A Spatial feature extraction and regularization model for the head-related transfer function,” J. Acoust. Soc. Am. **97**, 439-452 (1995).
- ¹⁶黃嘉文, “Headphone Reproduction of 3D Immersive Spatial Sound Field”, 國立交通大學機械工程研究所碩士論文.
- ¹⁷Douglars D. Rife and John Vanderkooy, “Transfer-Function Measurement with Maximum-Length Sequences”, J. Audio Eng. Soc., **37**, 419-443 (1989).
- ¹⁸G. H. Koopmann, L. Song, and J. B. Fahnlne, “A method for computing acoustic fields based on the principle of wave superposition”, J. Acoust. Soc. Am., **86**, 2433-2438, (1989).
- ¹⁹R. Jeans and I. C. Mathews, “The wave superposition method as a robust technique for computing acoustic fields”, J. Acoust. Soc. Am., **92**, 1156-1166, (1992).
- ²⁰白明憲, *聲學理論與應用：主動式噪音控制*, 修訂版 (全華書局, 2001).
- ²¹L. Song, G. H. Koopmann, and J. B. Fahnlne, “Numerical errors associated with the method of superposition for computing acoustic fields”, J. Acoust. Soc. Am. **89**, June 1991.
- ²²MATLAB 6.5 Help contents, MathWorks. Inc.
- ²³A. Schuhmacher and J. Hald, K. B. Rasmussen, P. C. Hansen, “Sound source

reconstruction using inverse boundary element calculations”, *J. Acoust. Soc. Am.* **113**, 114-127 (2003).

²⁴E. G. Williams, *Fourier Acoustics*, (Academic Press, 1999).

²⁵A. V. Oppenheim, R. W. Schaffer, J. R. Buck, *Discrete-time signal processing*, second edition, (Prentice-Hall, 1999).



TABLES

Table I. Number of measurement and azimuth increments at each elevation

Elevation	Number of Measurements	Azimuth Increment
-20	72	5.00
-10	72	5.00
0	72	5.00
10	72	5.00
20	72	5.00
30	60	6.00
40	60	6.00
50	45	8.00
60	36	10.00
70	24	15.00
80	12	30.00
90	1	X

Table II. Comparison of computation loadings between IBEM and virtual source representation

Virtual source representation case 1 :

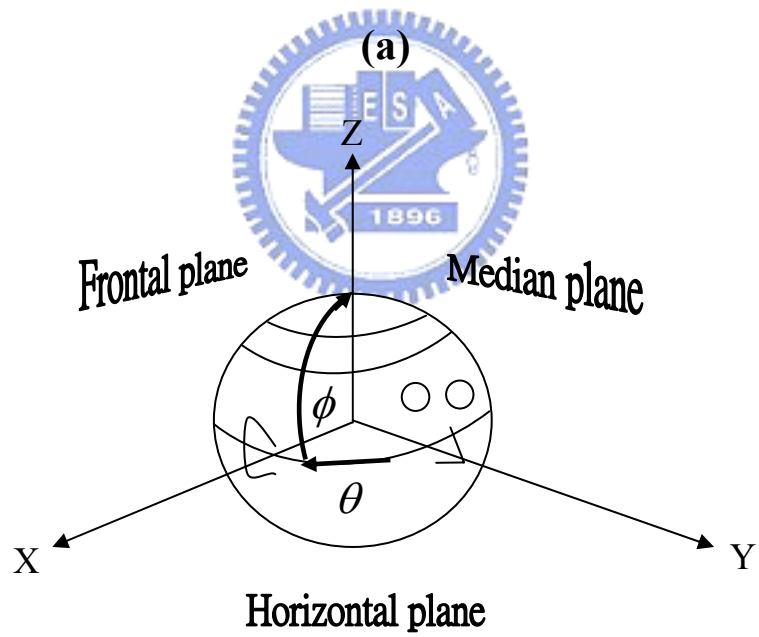
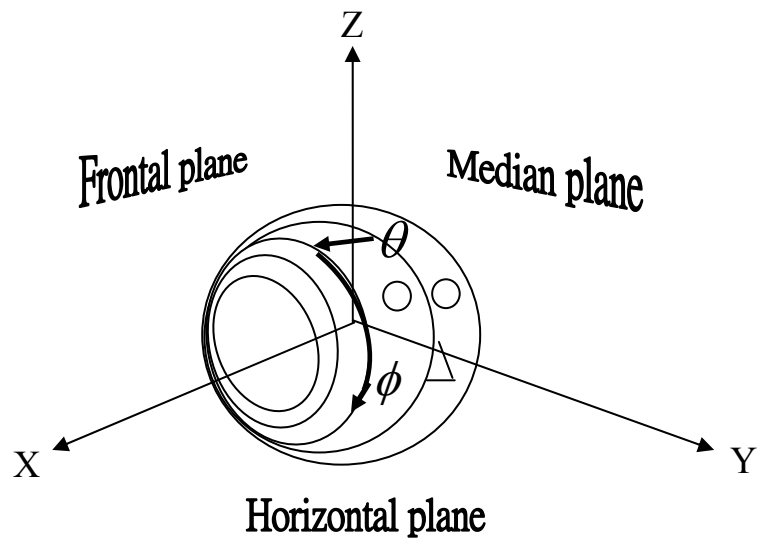
The nodes of head mesh are decimated into 2% of the original ones and the nodes of pinna mesh are also decimated into 17% of the original ones).

Virtual source representation case 2 :

The nodes of head mesh are decimated into 10% of the original ones and the nodes of pinna mesh are also decimated into 50% of the original ones).

Case	Number Of Elements	Number Of Nodes	Computation Time (seconds/per frequency point)	Memory Requirement (MB)
IBEM Kahana case	23799	12137	3560.77	897.65
IBEM ITRI Case	21529	10776	2475.13	879.51
Virtual Source Representation Case 1		581	40.52	83
Virtual Source Representation Case 2		2074	1003.23	311

FIGURES



(b)

Fig. 1 Coordinate systems for HRTF representation
(a) Interaural-Polar coordinate system (b) Vertical-Polar coordinate system

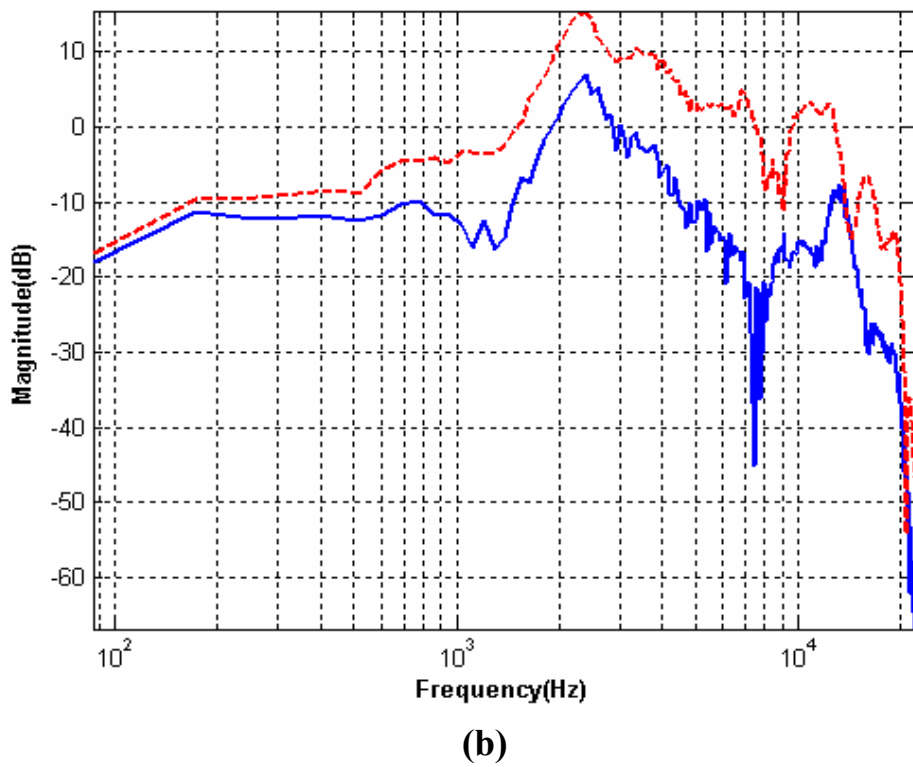
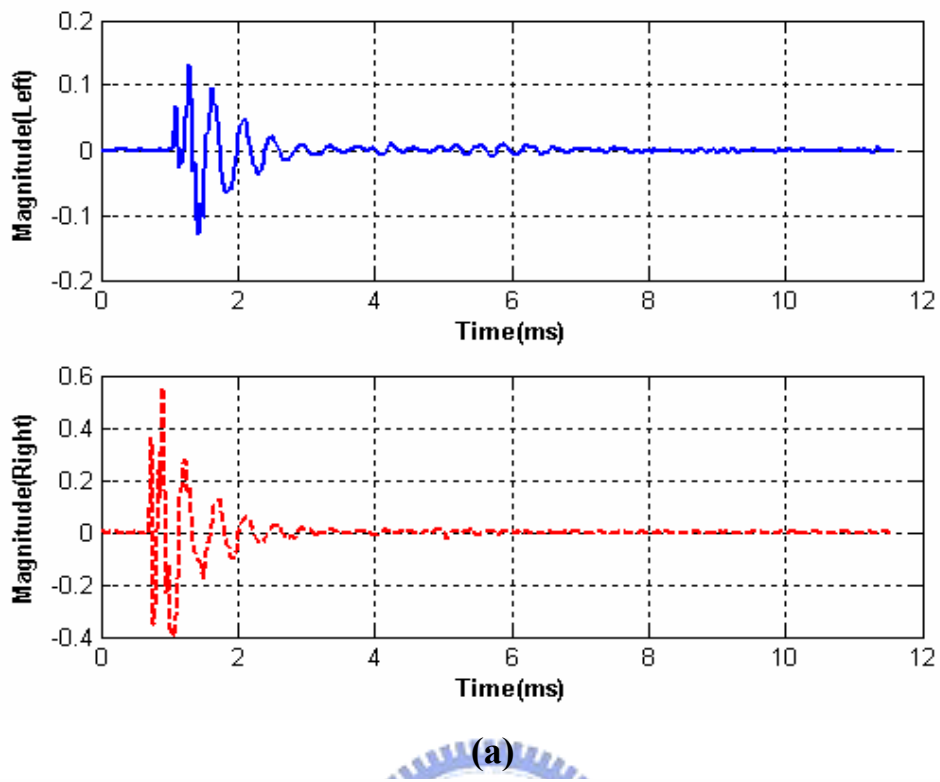


Fig. 2 An HRTF example of KEMAR at azimuth= 45° and elevation= 0°
 (a) The impulse response (HRIR). (b) The HRTF frequency response

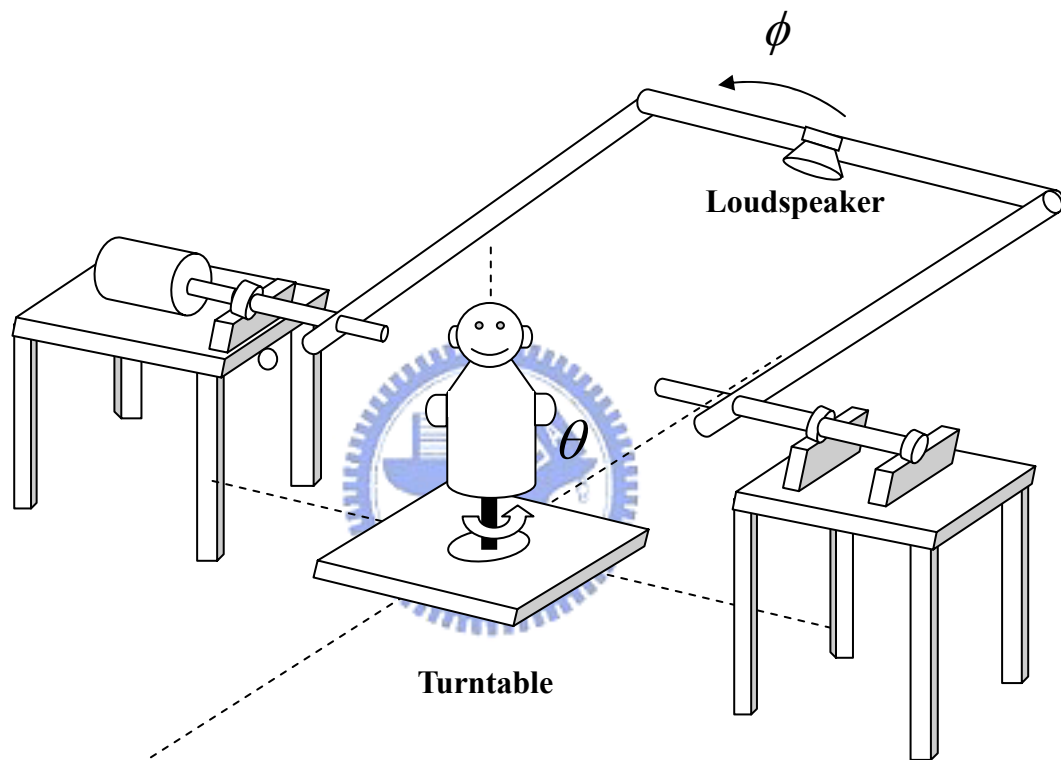


Fig. 3 Schematic diagram of measurement mechanism

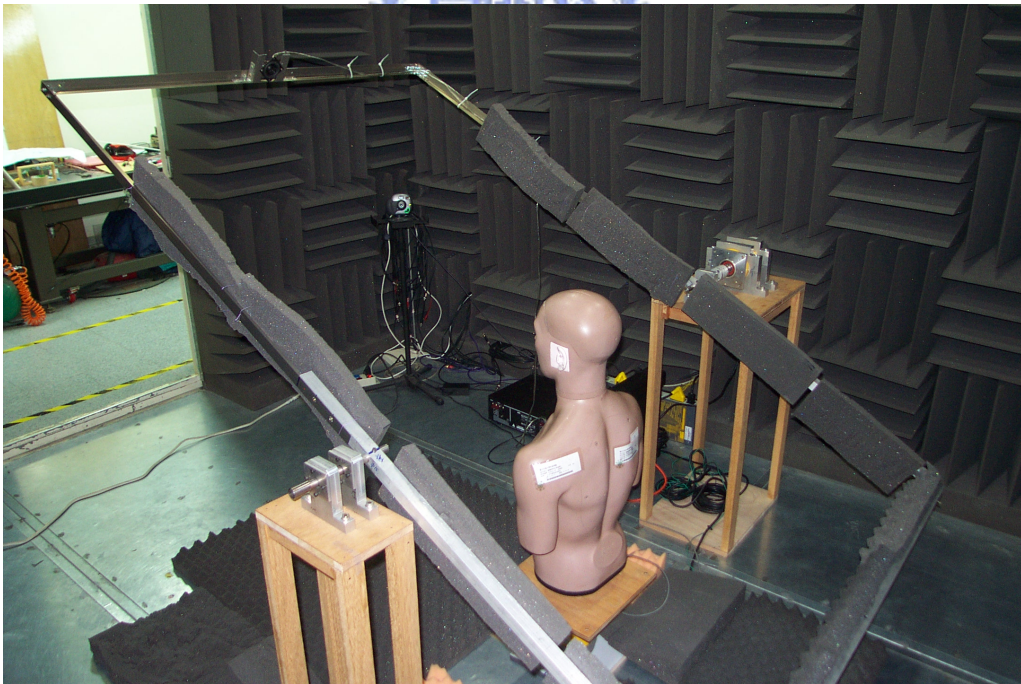


Fig. 4 Photo of HRTF measurement mechanism

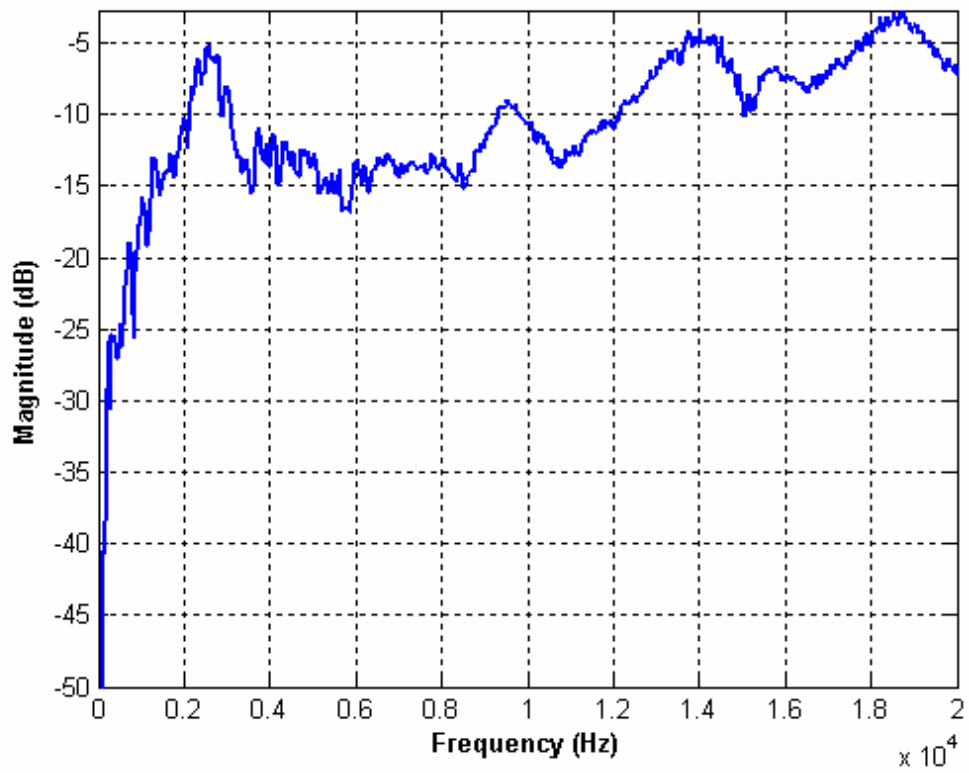


Fig. 5 Frequency response of measurement loudspeaker and power amplifier

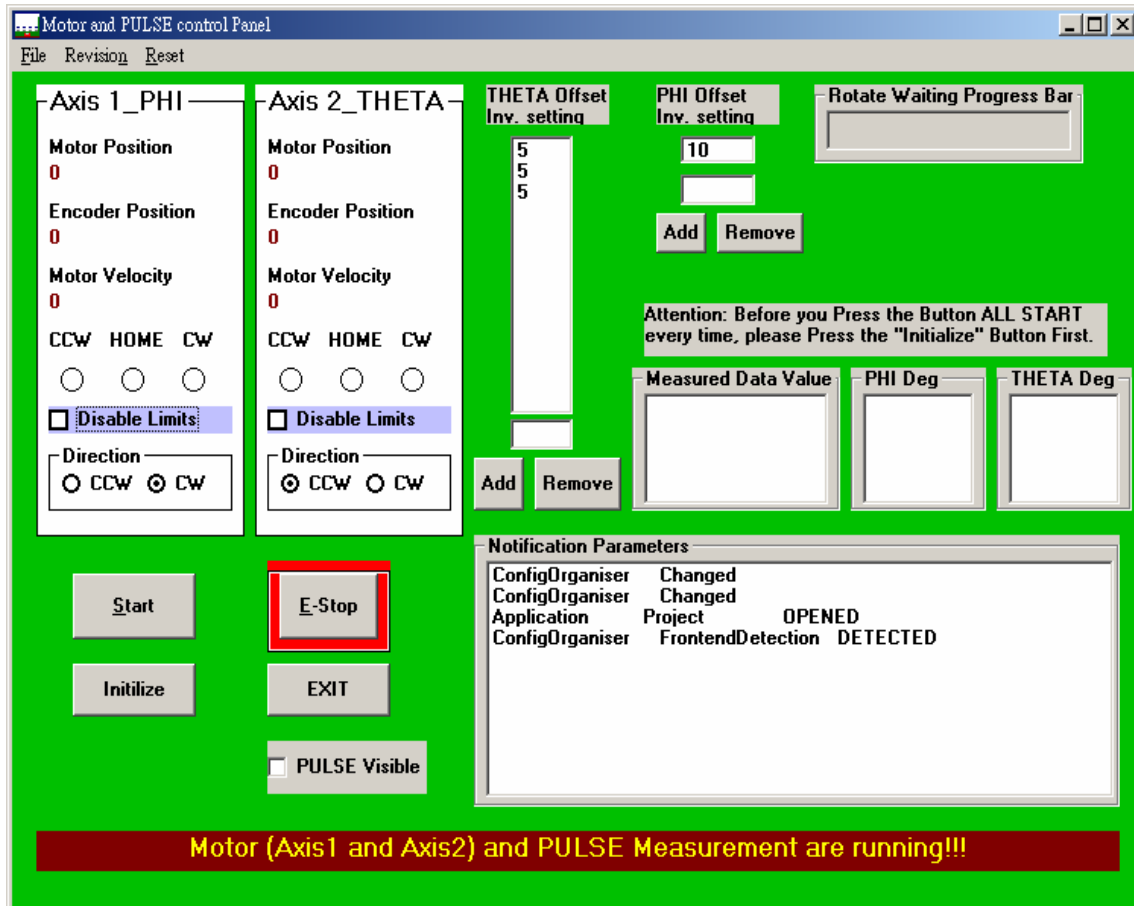


Fig. 6 The program interface of HRTF measurement written by using Visual Basic

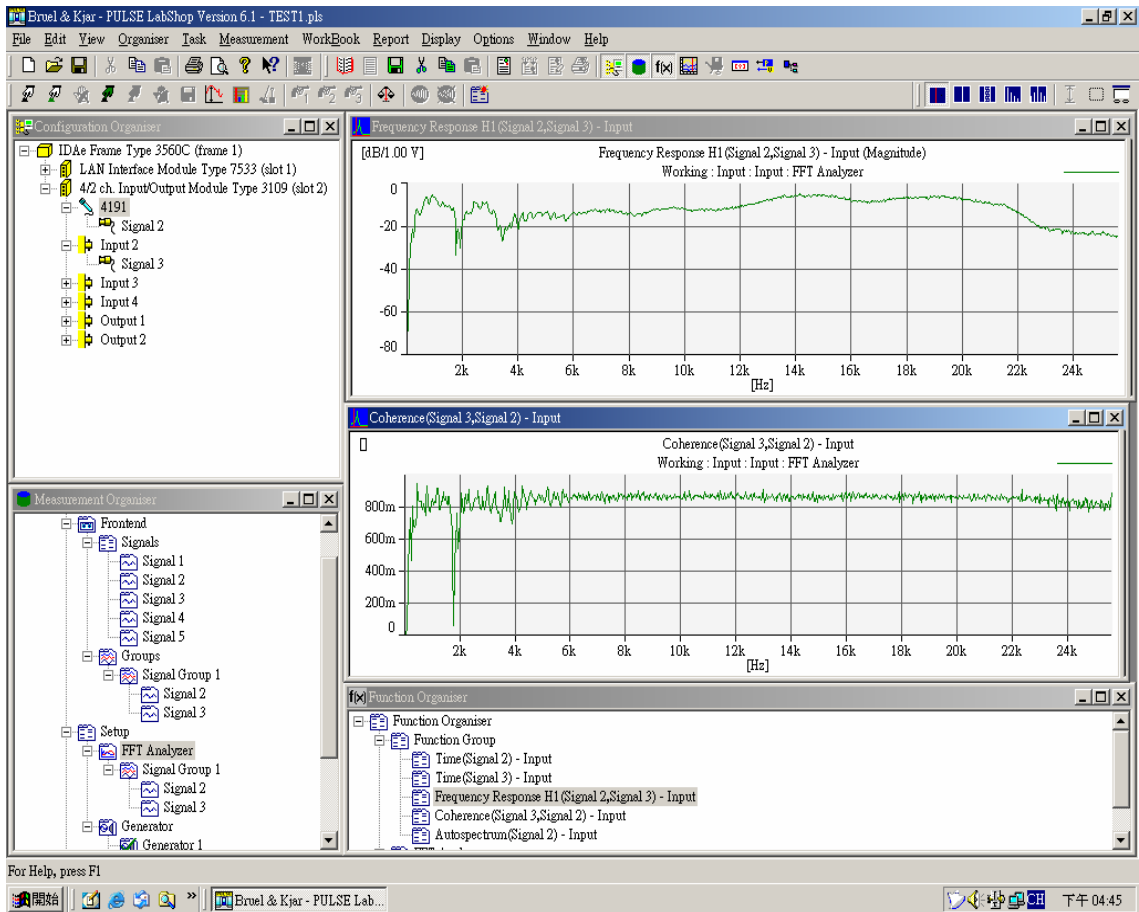


Fig. 7 The interface of Pulse multi analyzer software – Pulse Labshop

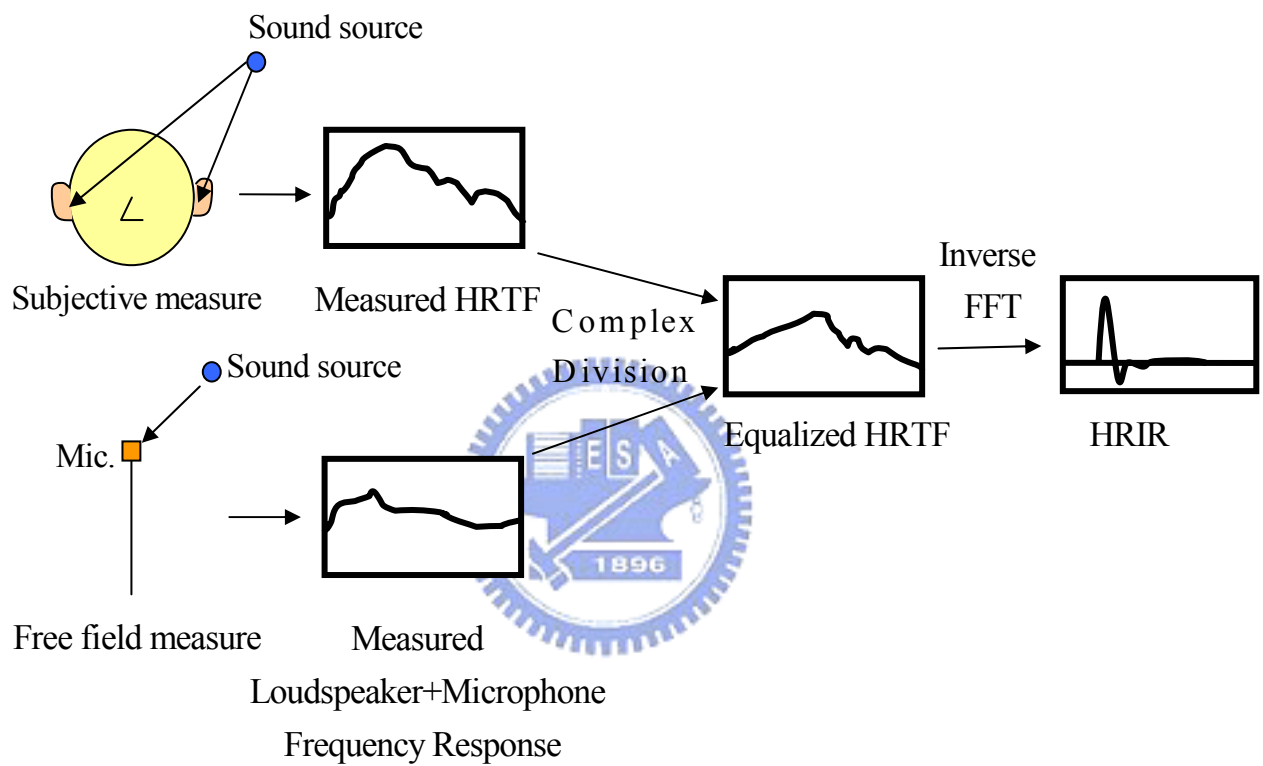


Fig. 8 Measured HRTFs post-processing flow chart

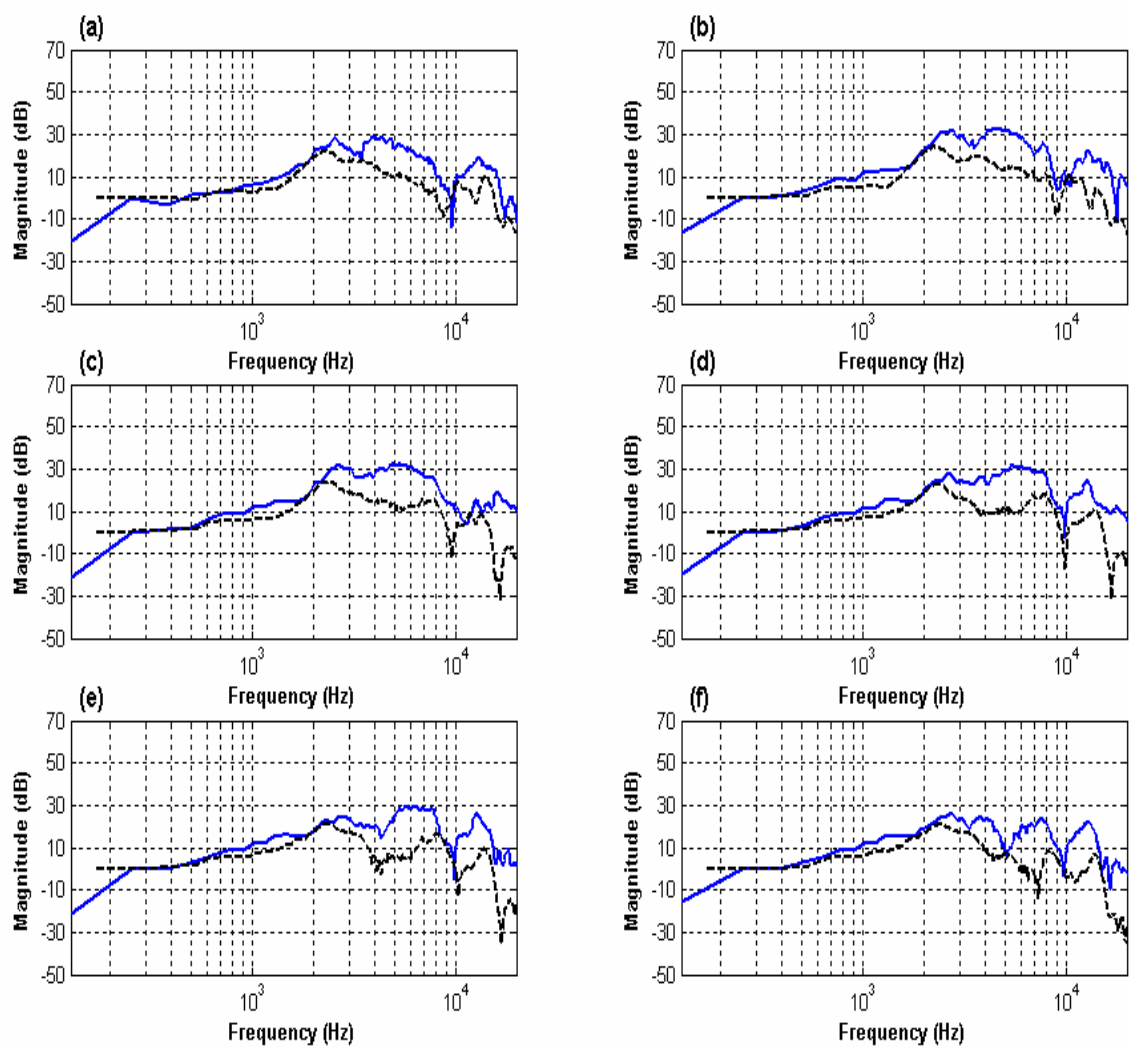


Fig. 9 Measured HRTFs on the horizontal plane (elevation = 0°)
 (dash line: unequalized MIT HRTF, solid line : unequalized HRTF)
 (a) azimuth = 0° (b) azimuth = 30° (c) azimuth = 60°
 (d) azimuth = 90° (e) azimuth = 120° (f) azimuth = 150°

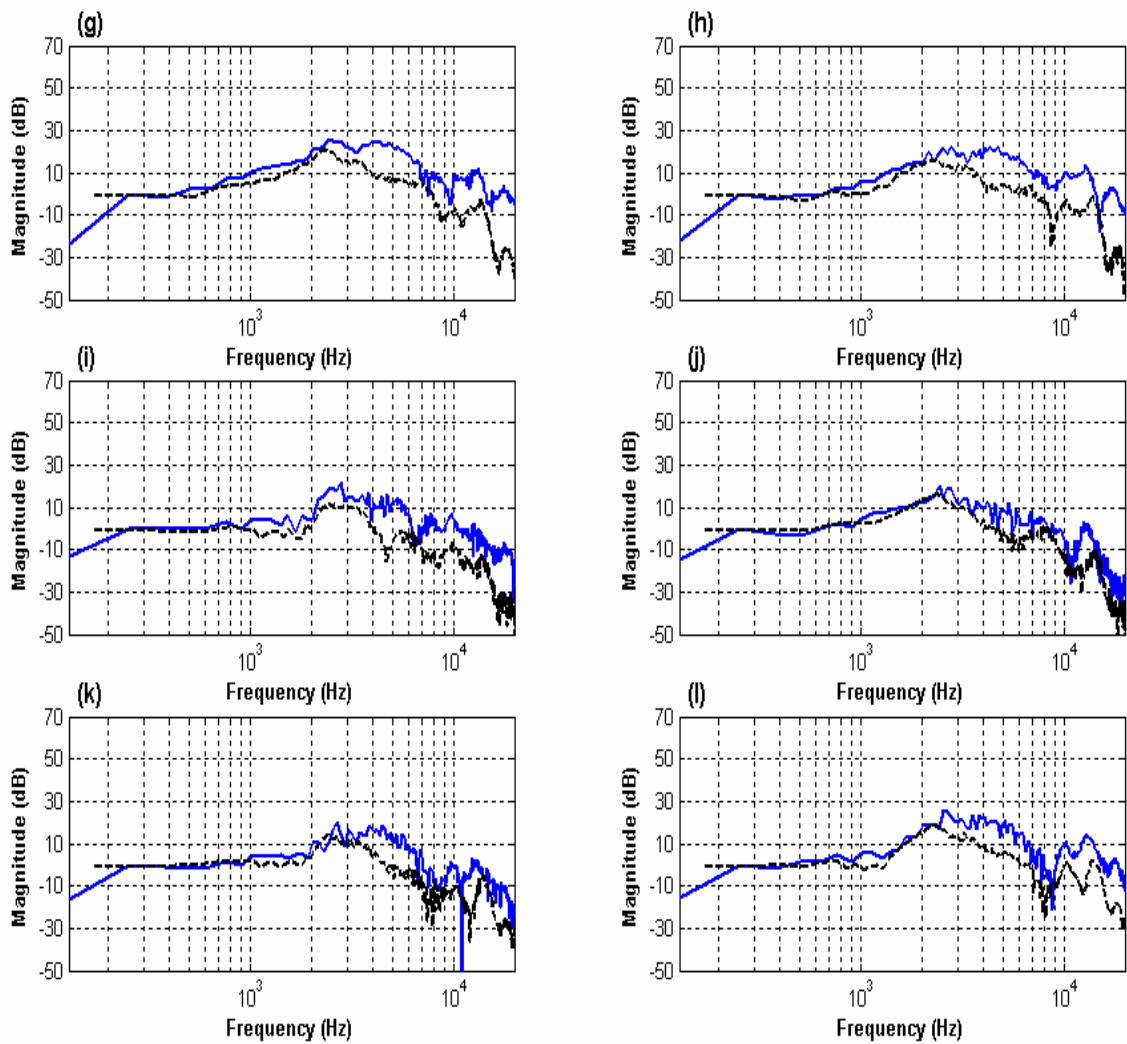


Fig. 9 Measured HRTFs on the horizontal plane (elevation = 0°)
 (dash line: unequalized MIT HRTF, solid line : unequalized HRTF)
 (g) azimuth = 180° (h) azimuth = 210° (i) azimuth = 240°
 (j) azimuth = 270° (k) azimuth = 300° (l) azimuth = 330°

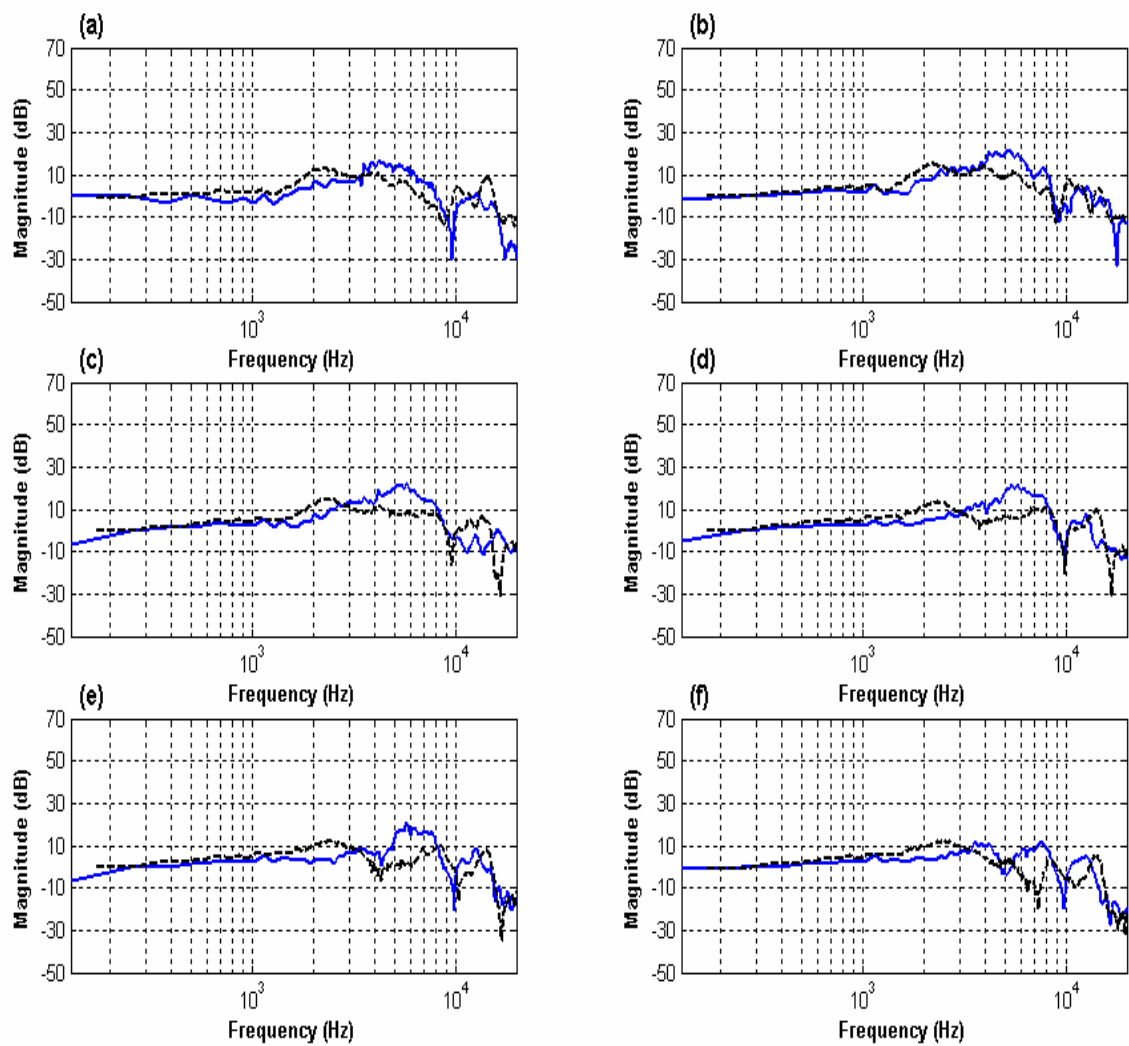


Fig. 10 Measured HRTFs on the horizontal plane (elevation = 0°)
 (dash line: equalized MIT HRTF, solid line : equalized HRTF)
 (a) azimuth = 0° (b) azimuth = 30° (c) azimuth = 60°
 (d) azimuth = 90° (e) azimuth = 120° (f) azimuth = 150°

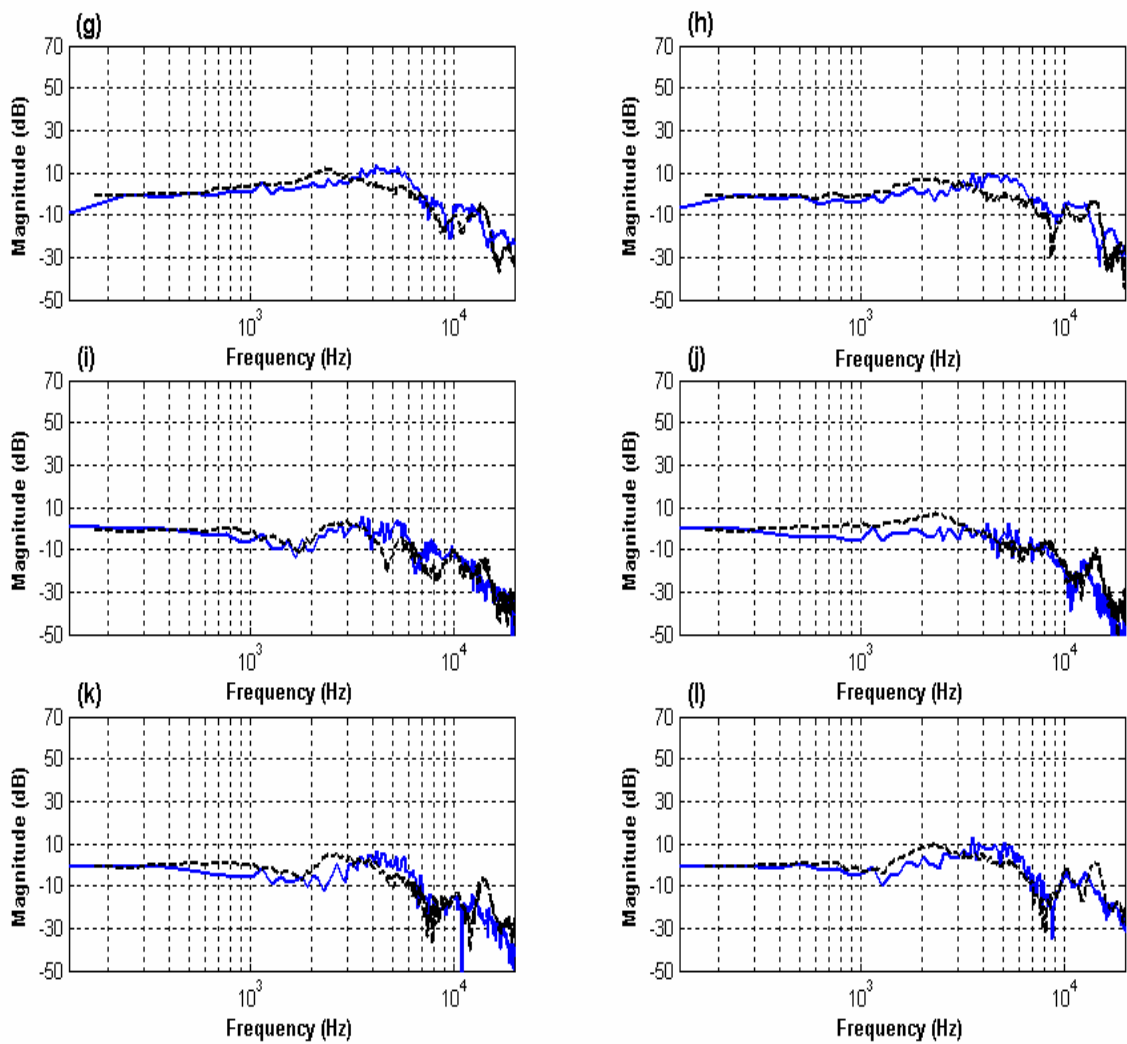


Fig. 10 Measured HRTFs on the horizontal plane (elevation = 0°)
 (dash line: equalized MIT HRTF, solid line : equalized HRTF)
 (g)azimuth = 180° (h) azimuth = 210° (i) azimuth = 240°
 (j)azimuth = 270° (k) azimuth = 300° (l) azimuth = 330°

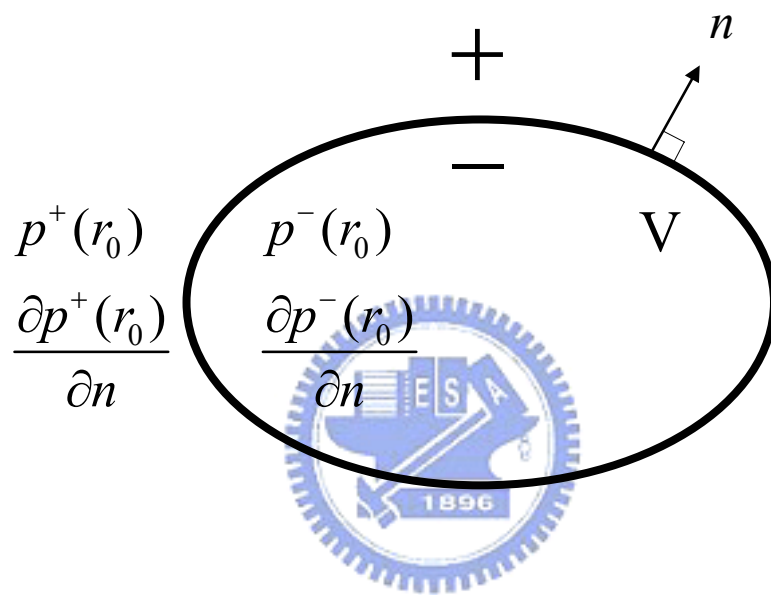


Fig. 11 Indirect boundary element method definition (close body)

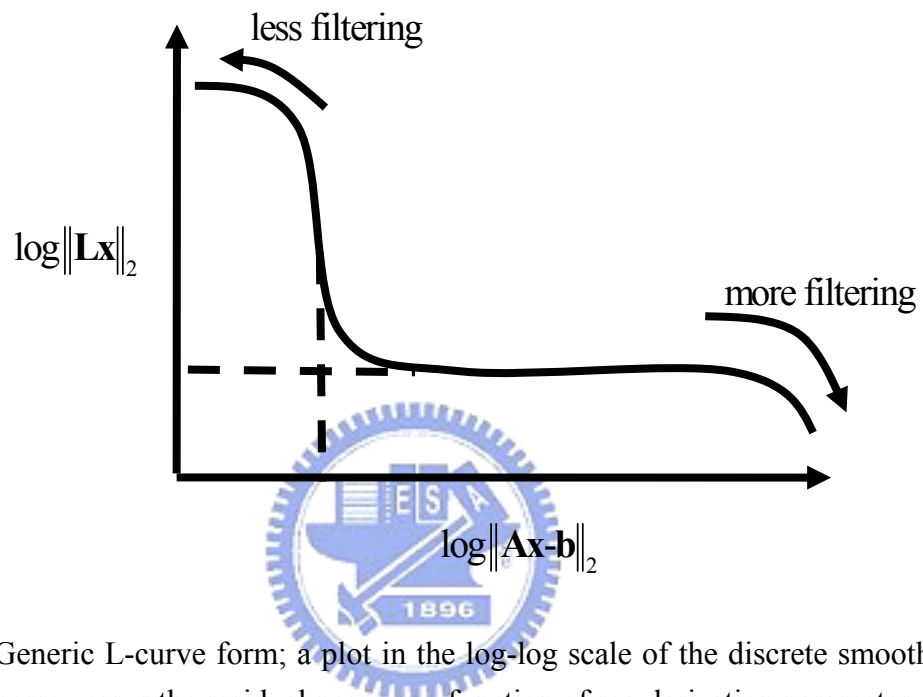


Fig. 12 Generic L-curve form; a plot in the log-log scale of the discrete smoothing norm versus the residual norm as a function of regularization parameter

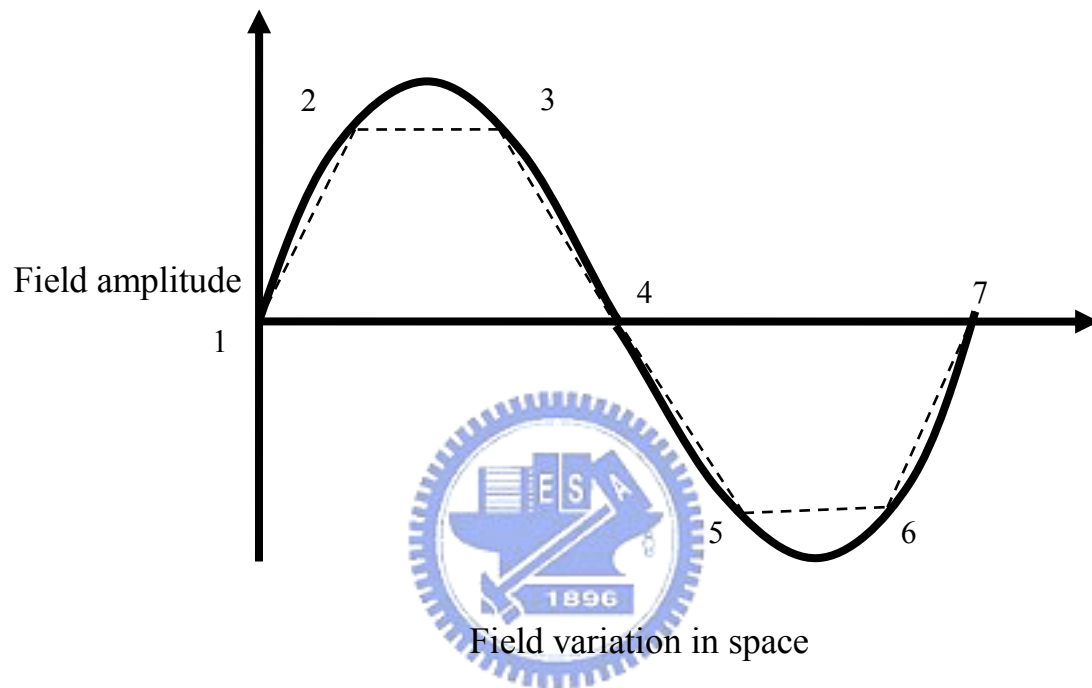


Fig. 13 Spacing of nodes at maximum frequency. Seven points denote the six elements in one wavelength.

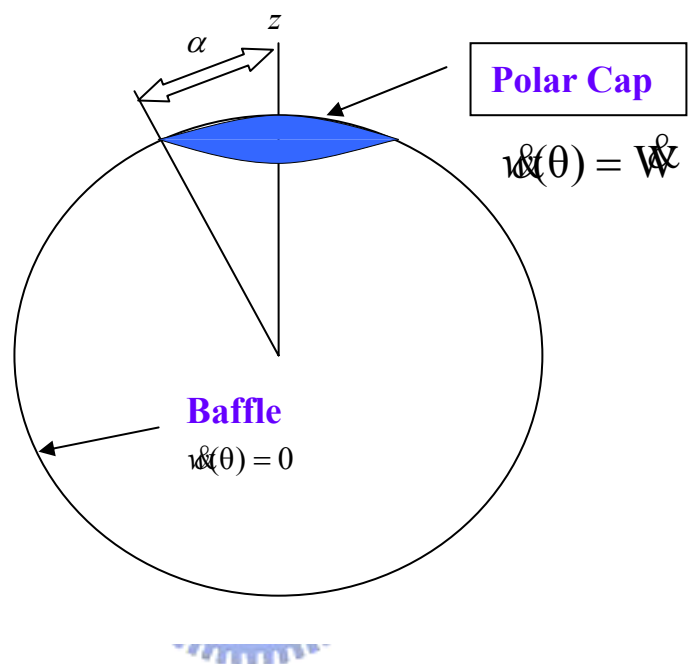


Fig. 14 Circular piston of radius α set in the rigid sphere. The piston vibrates with a velocity \hat{w} . The z axis passes through the center of the piston.

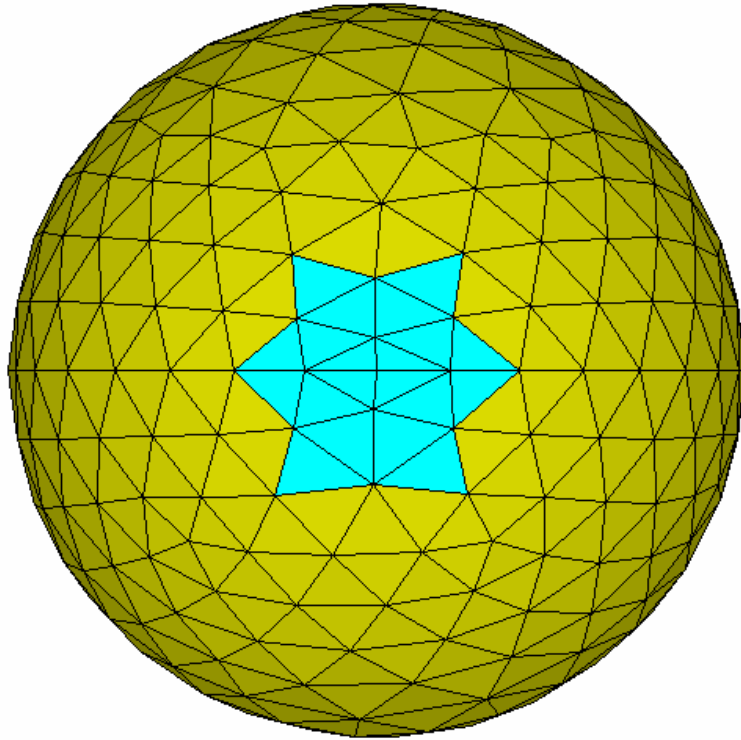
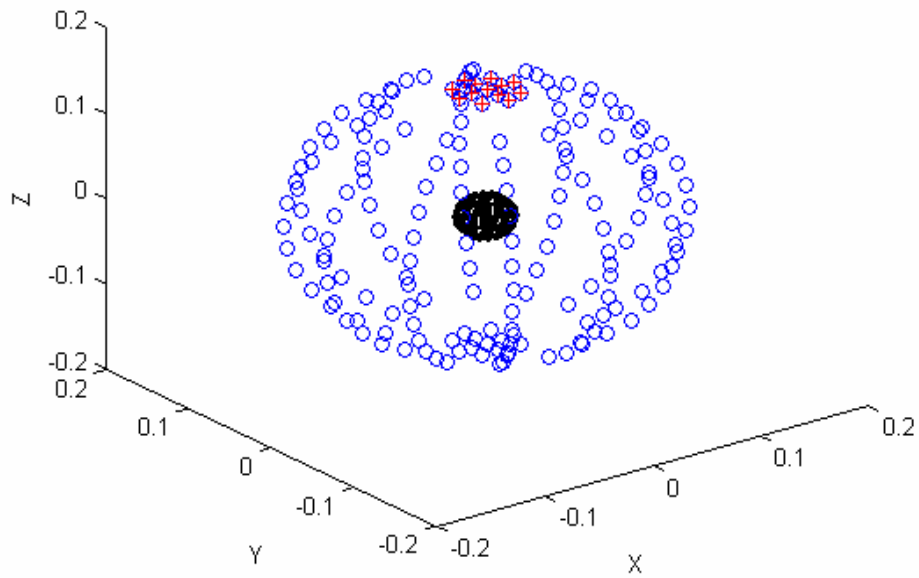
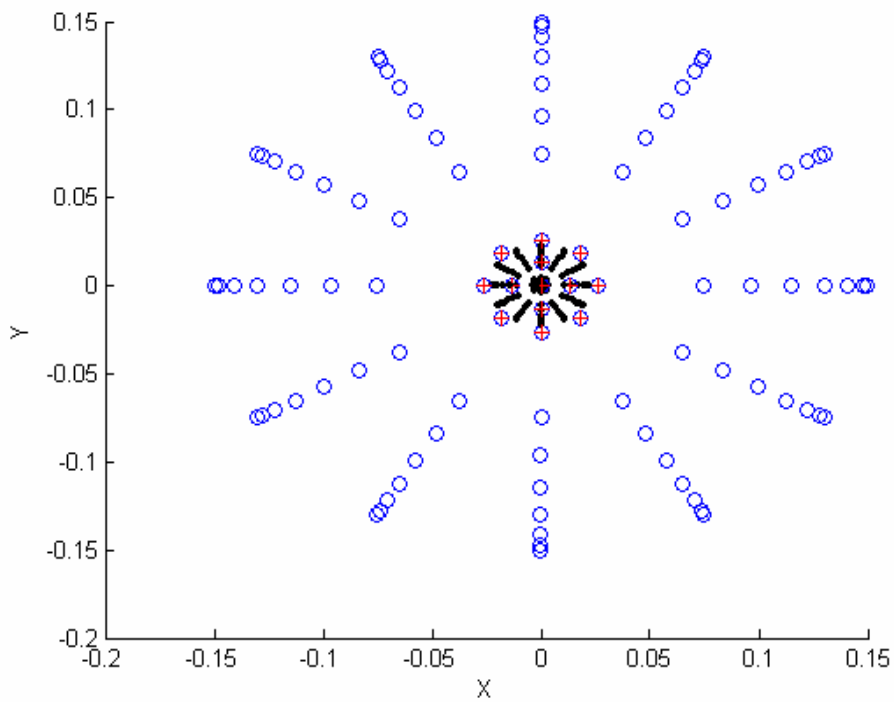


Fig. 15 The settings of the mesh elements in the SYSNOISE. The light color areas denote the rigid baffle and others denote the vibrating circular piston.



(a)



(b)

Fig. 16 The settings of mesh nodes in virtual source representation. The plus sign in (a) and (b) denotes the vibrating circular piston. Inner black points denote the virtual sources.

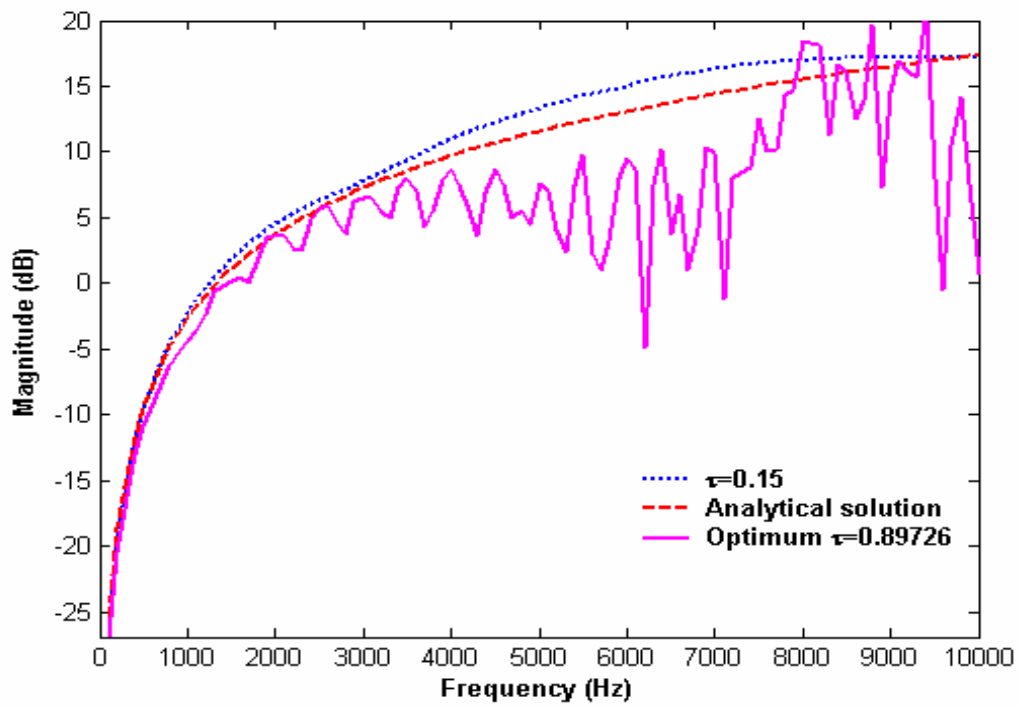
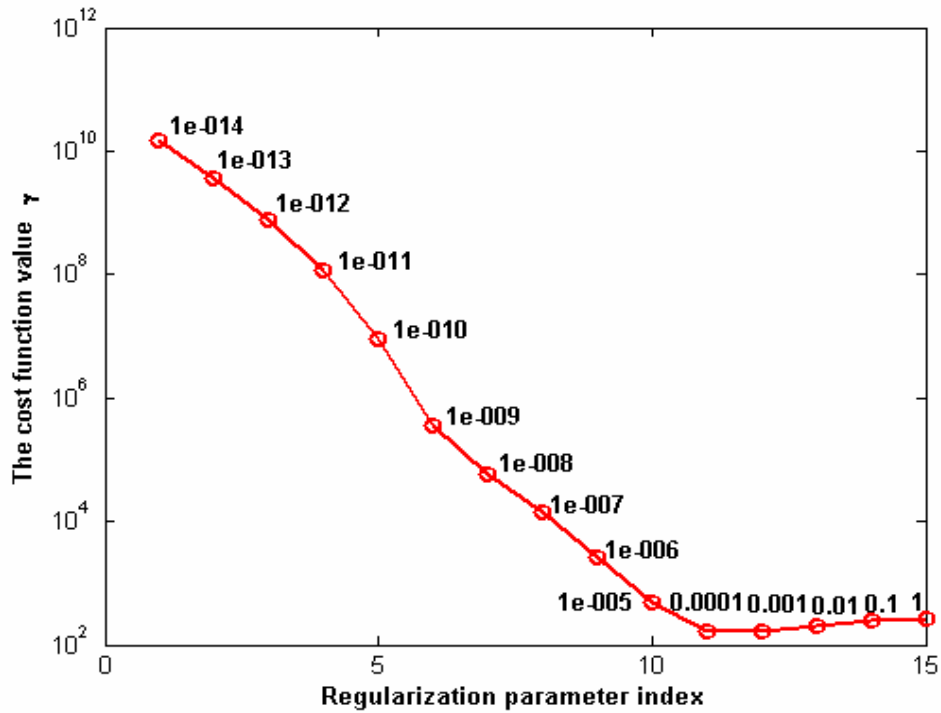
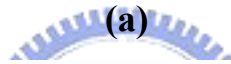
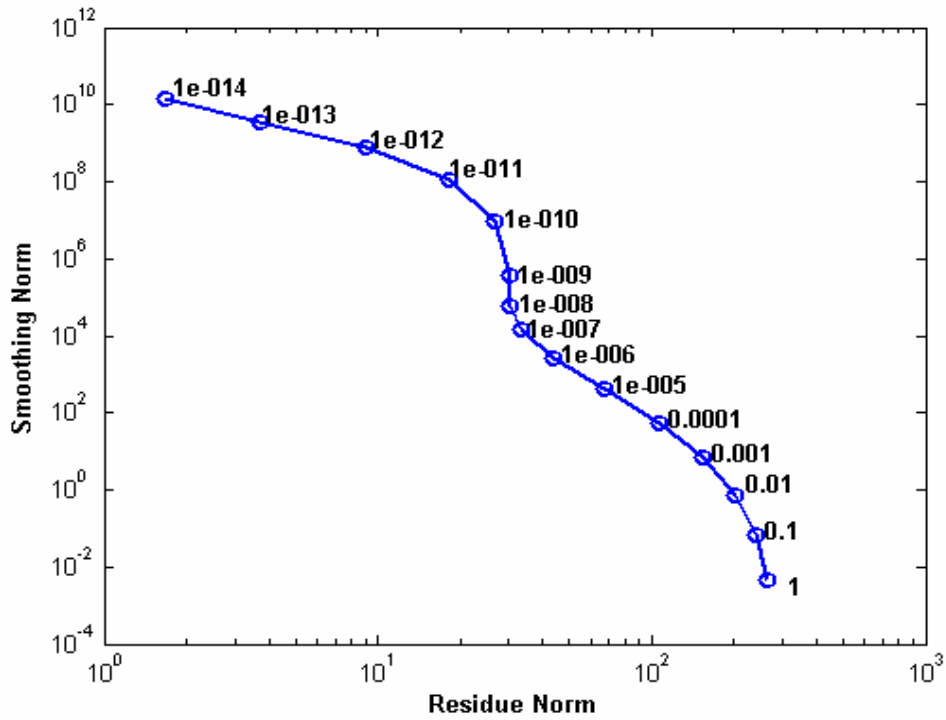


Fig. 17 Frequency responses of circular piston in a spherical baffle with optimum and small τ in virtual source representation and analytical solution.



(b)

Fig. 18 L-curve of circular piston in a spherical baffle at
 (a) 100 Hz. (b) Distribution of the value (γ) of cost function.
 The regularization parameter β is taken as the first singular value multiplied by the values marked on the curve.

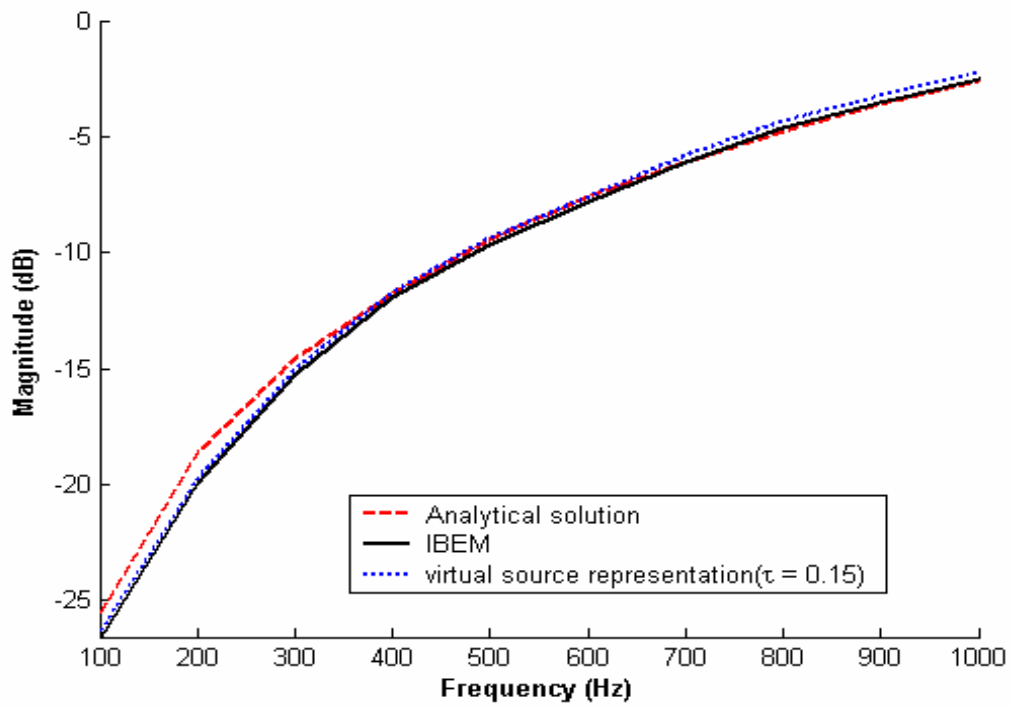


Fig. 19 Frequency responses of circular piston in a spherical baffle calculated by using analytical solution, virtual source representation and IBEM.

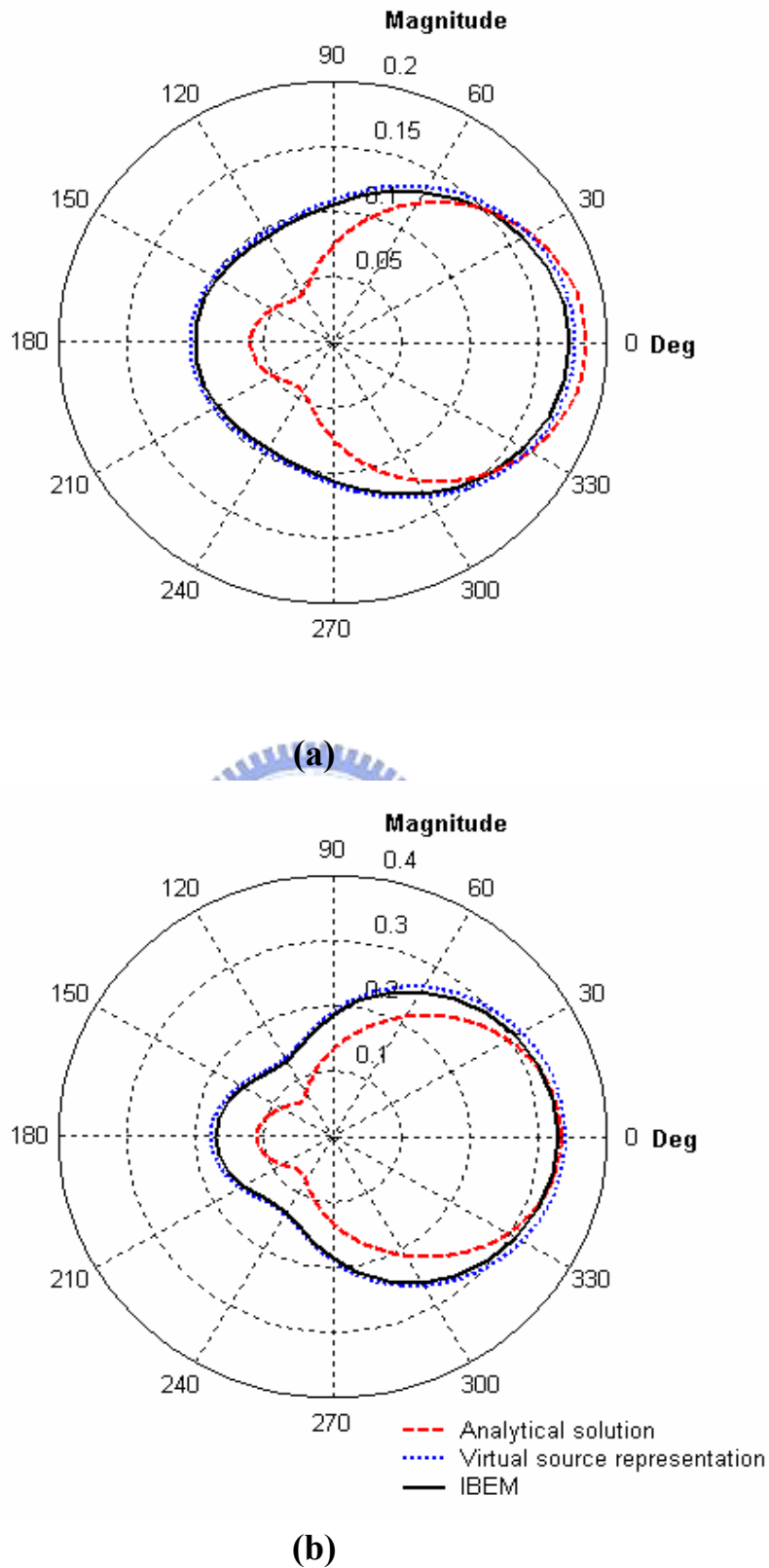
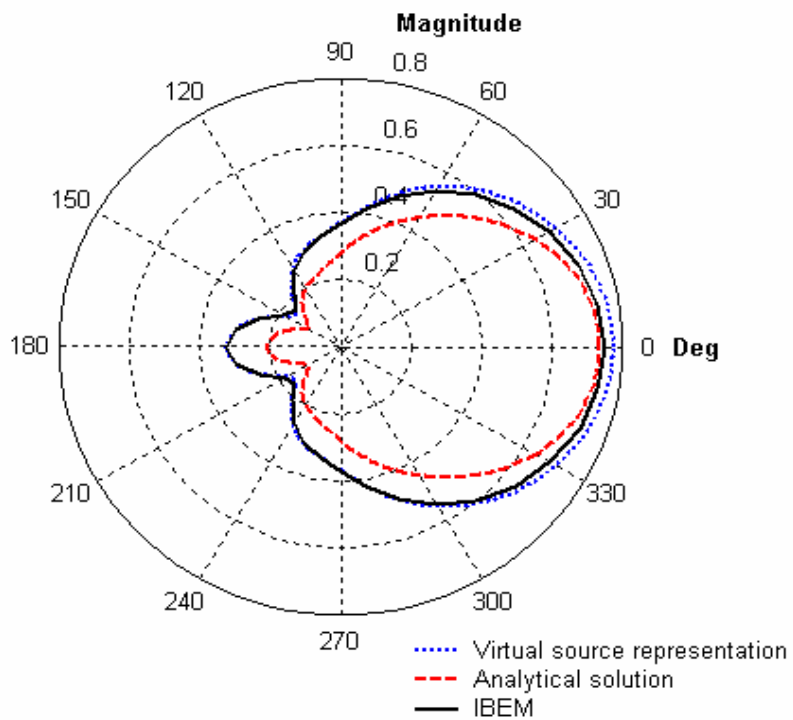
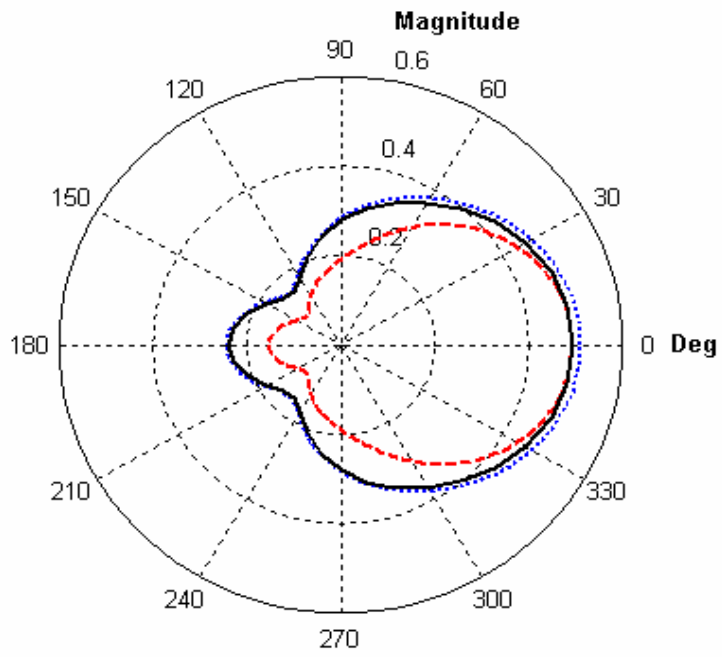
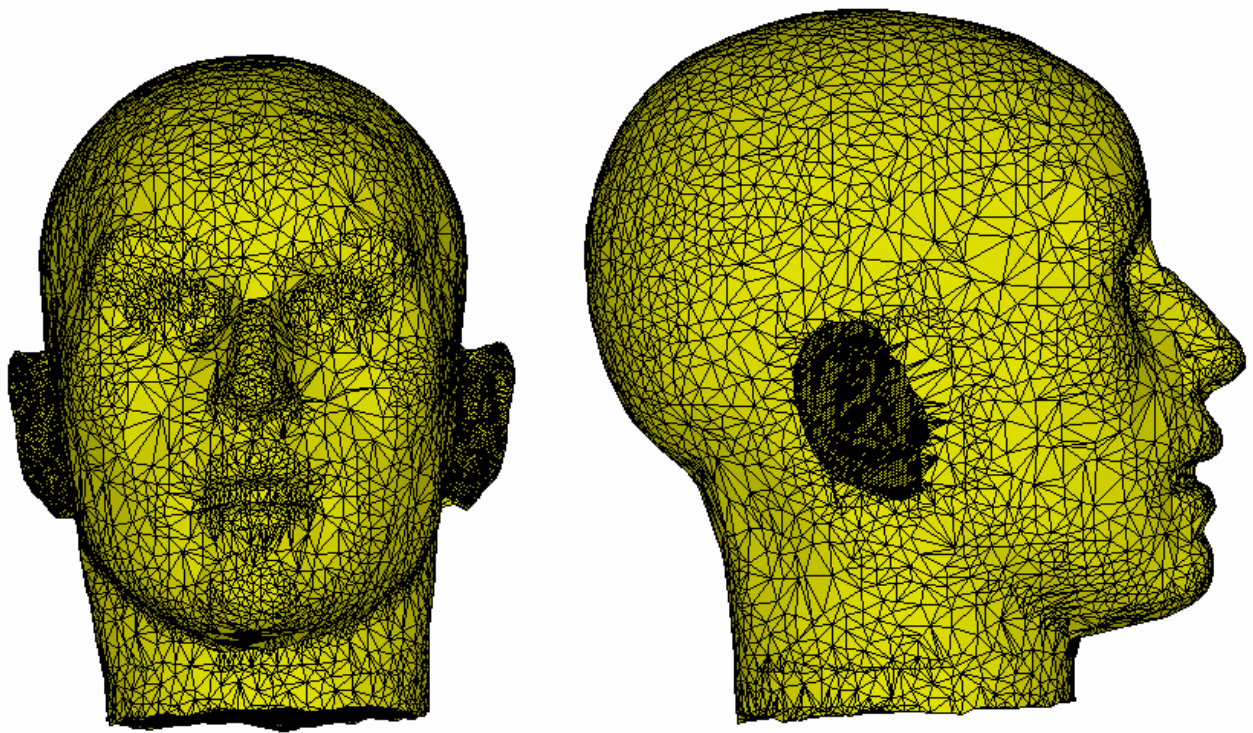


Fig. 20 Directional response of circular piston in a spherical baffle at (a) 300 Hz (b) 500 Hz



(d)

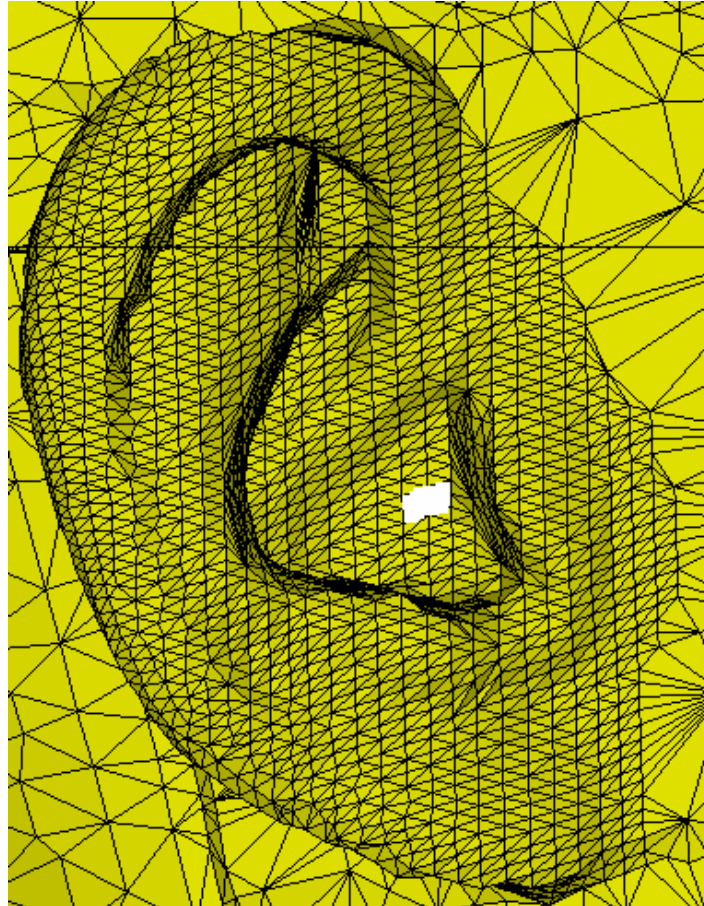
Fig. 20 Directional response of circular piston in a spherical baffle at (c) 700 Hz (d) 1000 Hz



(a)

Fig. 21 The mesh settings of KEMAR dummy head using IBEM in the SYSNOISE (ITRI case).

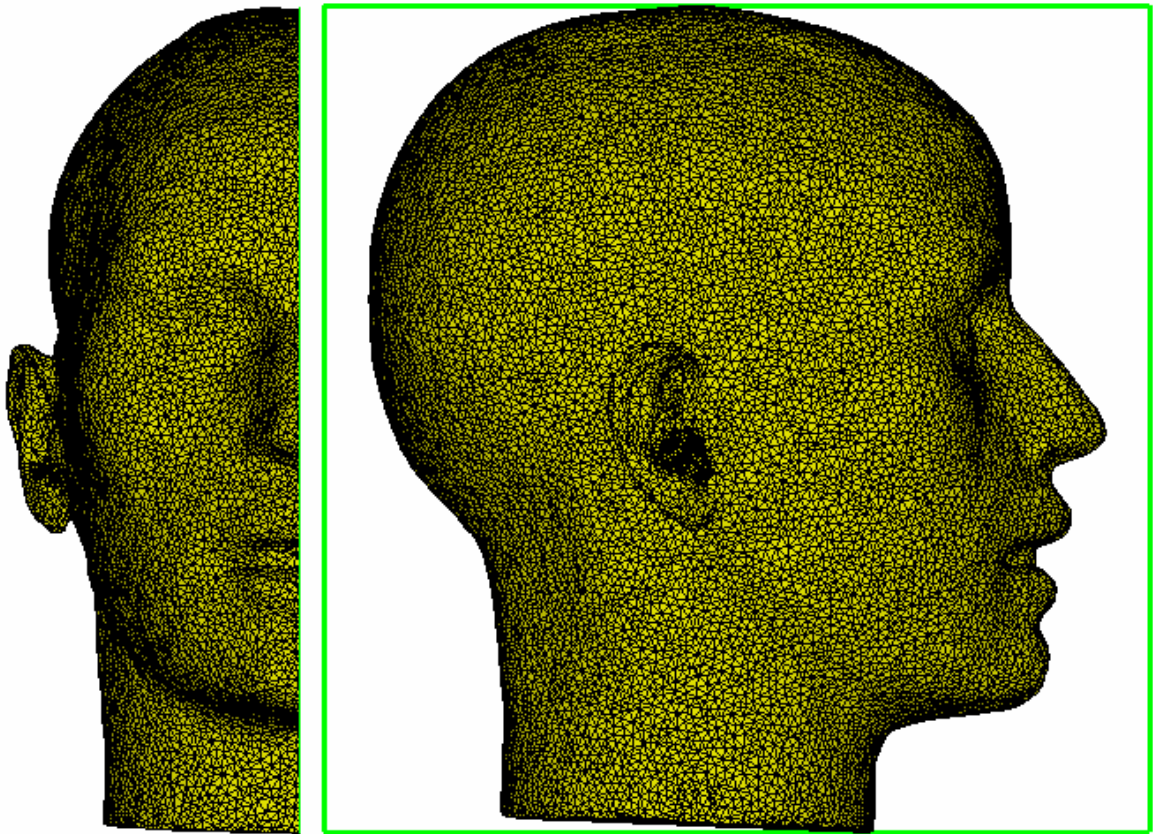
(a) Head mesh given by ITRI



(b)

Fig. 21 The mesh setting of KEMAR dummy head using IBEM in the SYSNOISE (ITRI case).

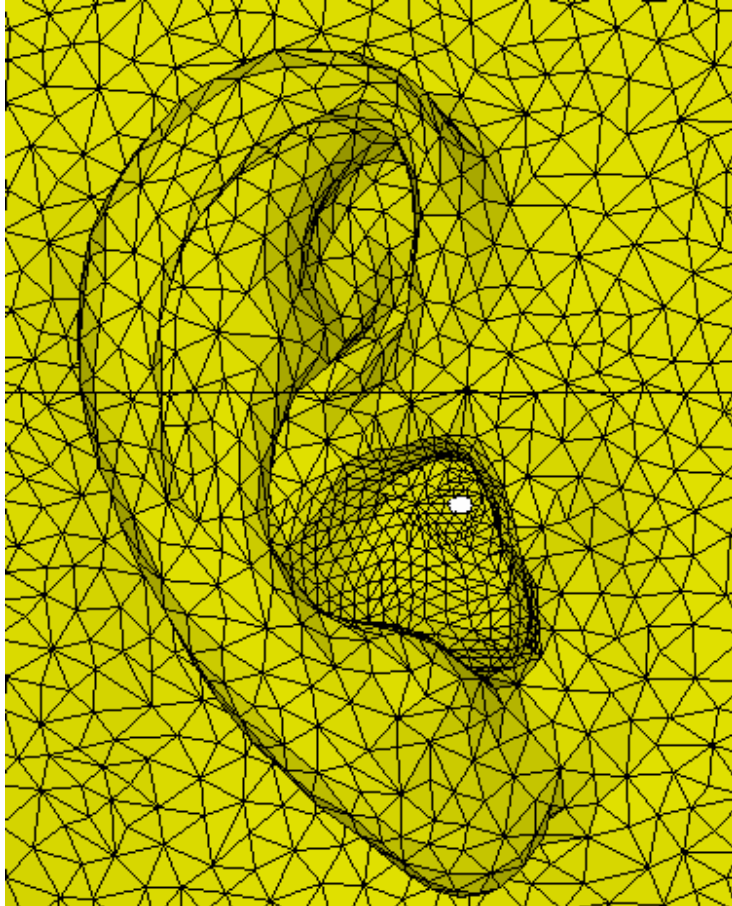
(b) Pinna (White color areas denote the vibrating panels).



(a)

Fig. 22 The mesh settings of KEMAR dummy head using IBEM in the SYSNOISE (Kahana case).

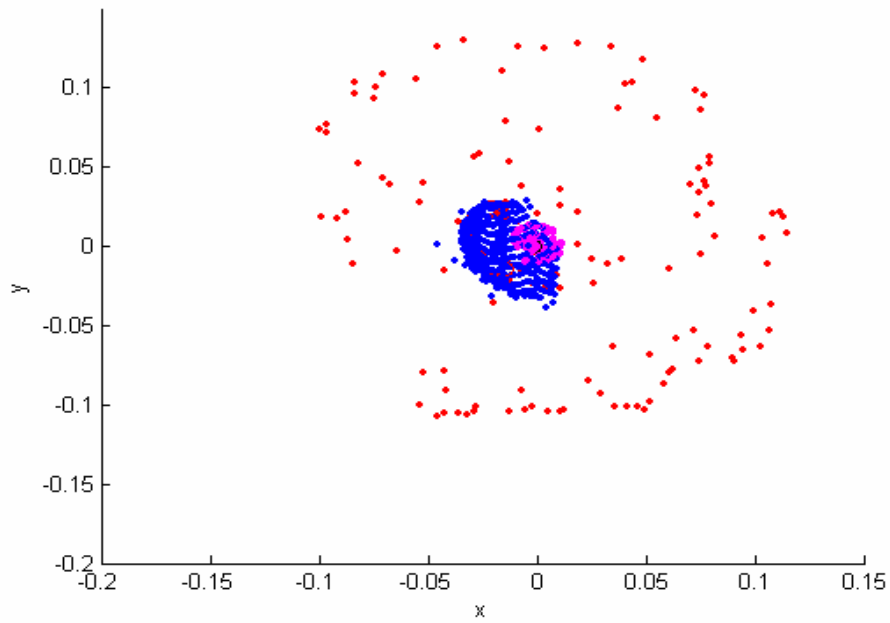
(a) Head mesh given by Yuvi Kahana



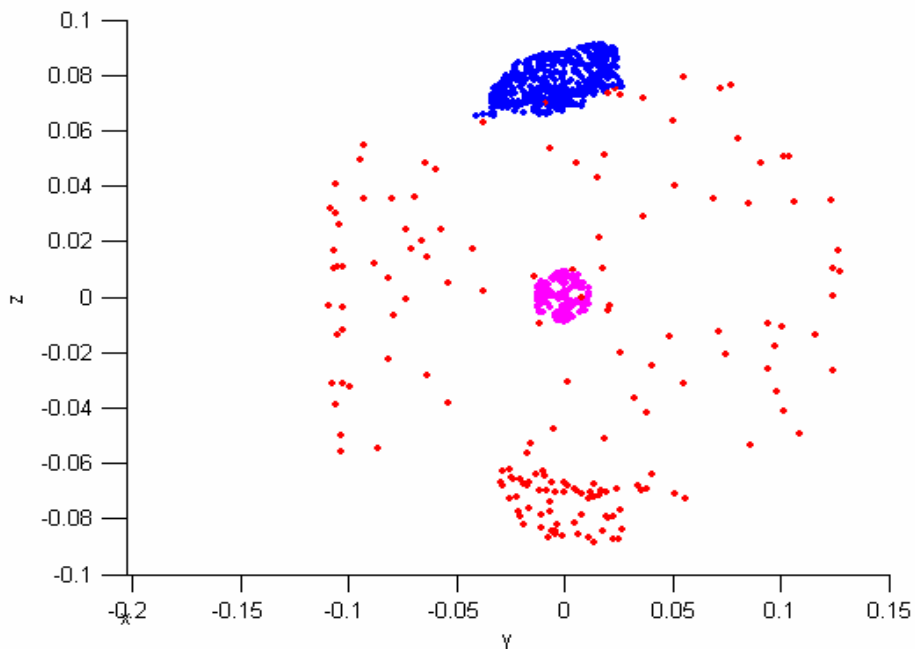
(b)

Fig. 22 The mesh settings of KEMAR dummy head using IBEM in the SYSNOISE (Kahana case).

(b) Pinna (White areas are the vibrating panels).



(a)



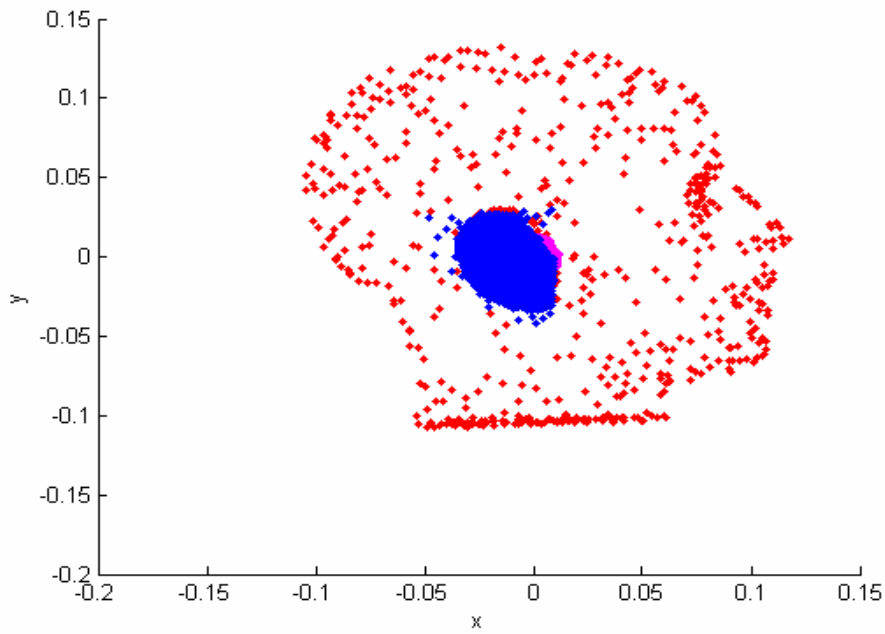
(b)

Fig. 23 The distributions of KEMAR mesh nodes in virtual source representation. (Only 2% of the original scanned head mesh nodes and 17% of the original scanned pinna mesh nodes).

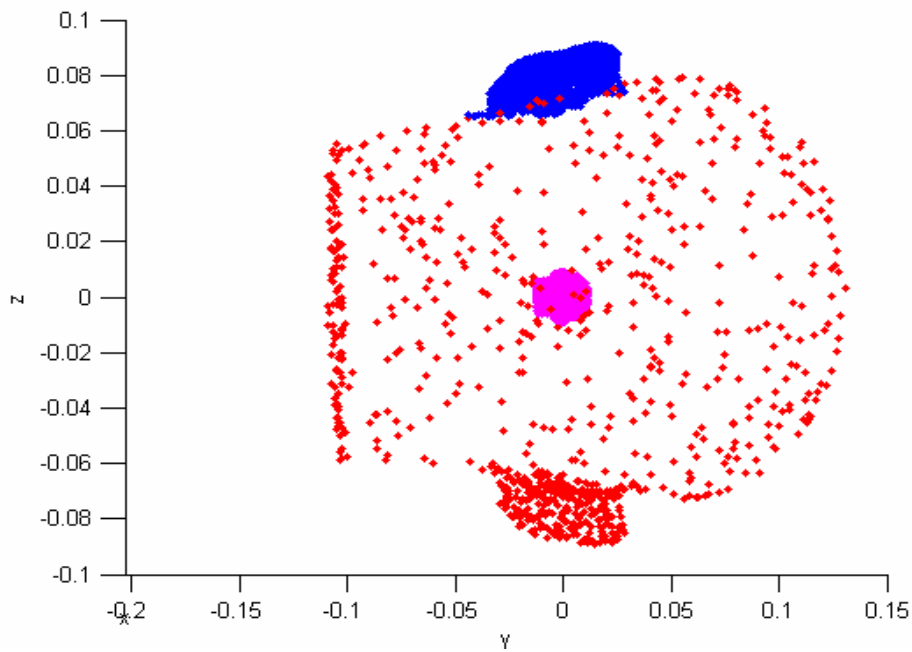
(a)The right hand side of KEMAR (case 1)

(b)The front of KEMAR (case 1)

(The cluster of light color points denote the virtual sources)



(c)



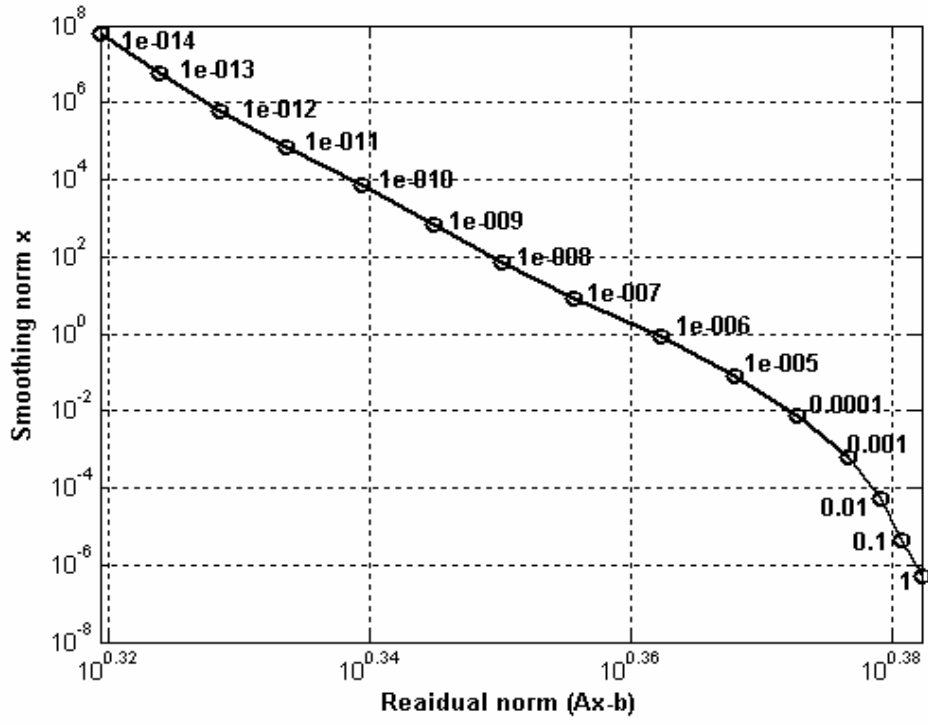
(d)

Fig. 23 The distributions of KEMAR mesh nodes in virtual source representation. (Only 10% of the original scanned head mesh nodes and 50% of the original scanned pinna mesh nodes).

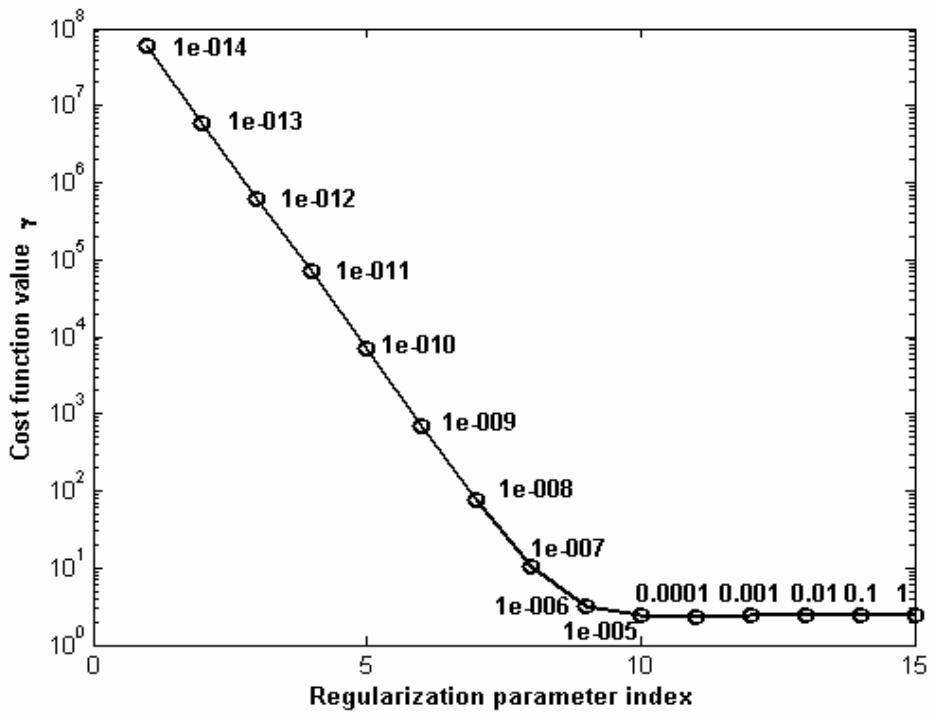
(c)The right hand side of KEMAR (case 2)

(d)The front of KEMAR (case 2)

(The cluster of light color points denote the virtual sources)

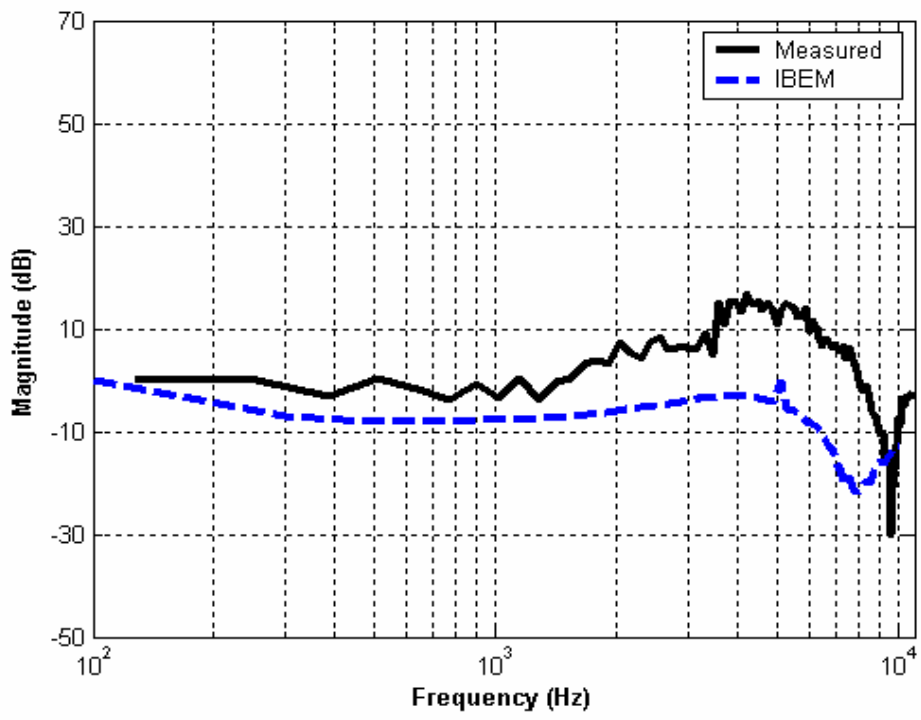


(a)

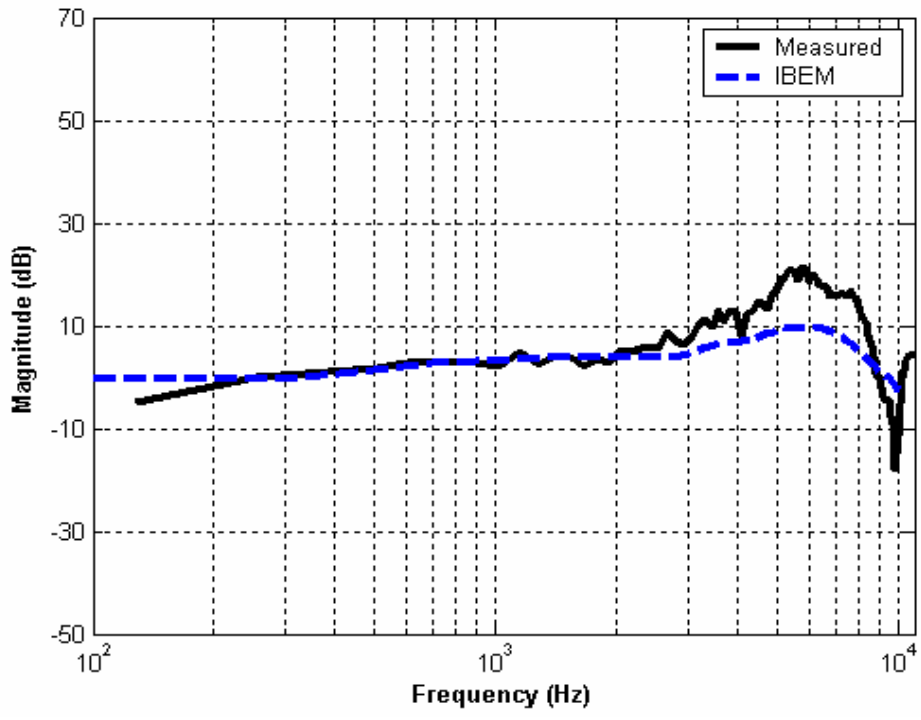


(b)

Fig. 24 L-curve of HRTF computations using virtual source representation at (a) 100 Hz. (b) Distribution of value (γ) of cost function. The regularization parameter β is taken as the first singular value multiplied by the values marked on the curve. (Virtual source case 2)

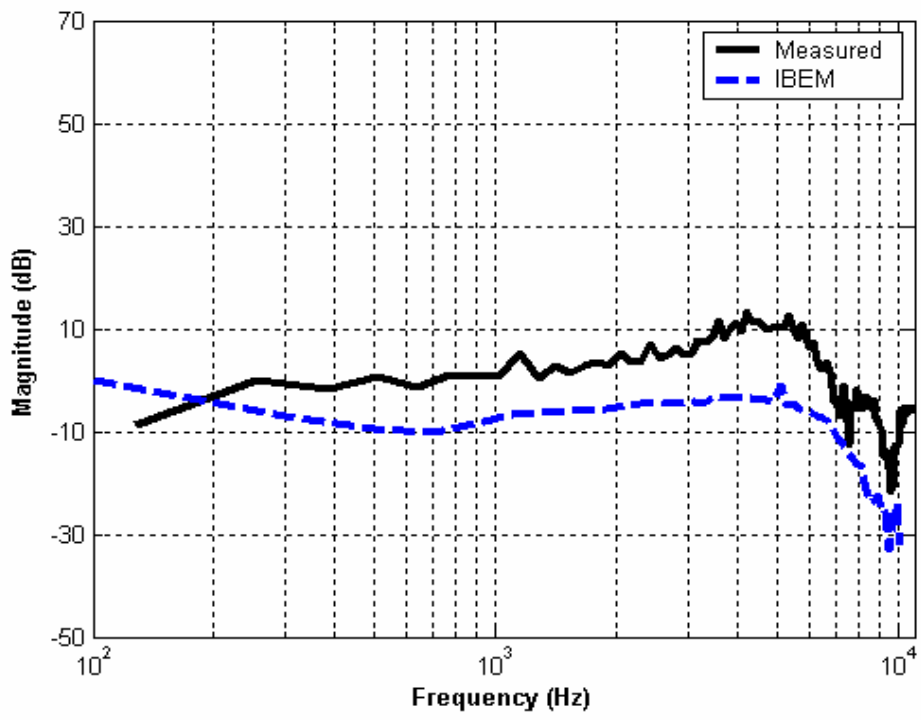


(a)

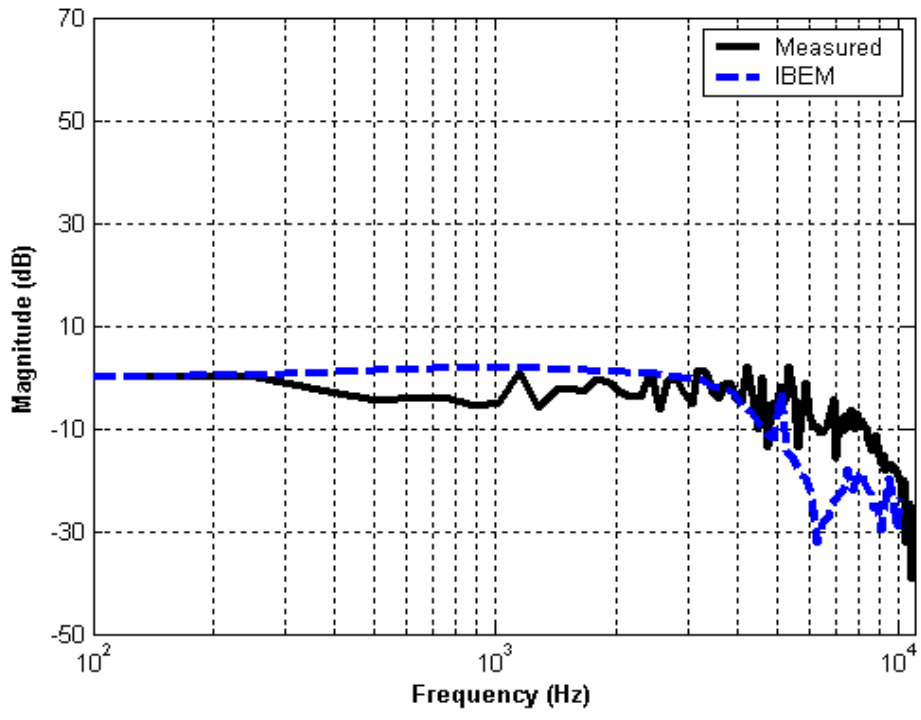


(b)

Fig. 25 Calculated HRTFs on the horizontal plane by using IBEM (ITRI case).
 (a) 0 azimuth degrees (b) 90 azimuth degrees

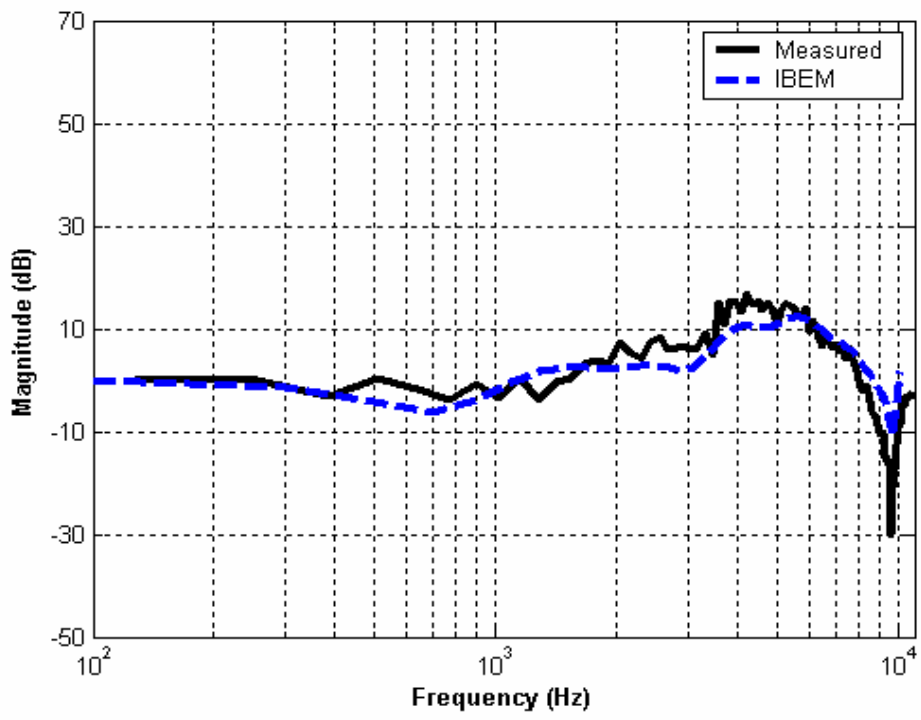


(c)

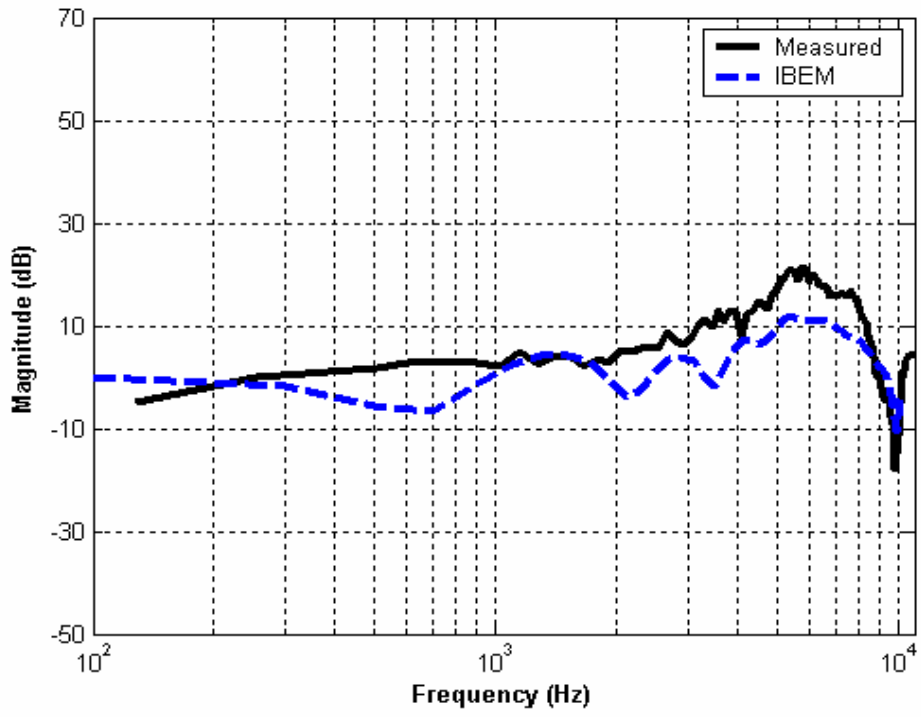


(d)

Fig. 25 Calculated HRTFs on the horizontal plane by using IBEM (ITRI case).
 (c) 180 azimuth degrees (d) 270 azimuth degrees

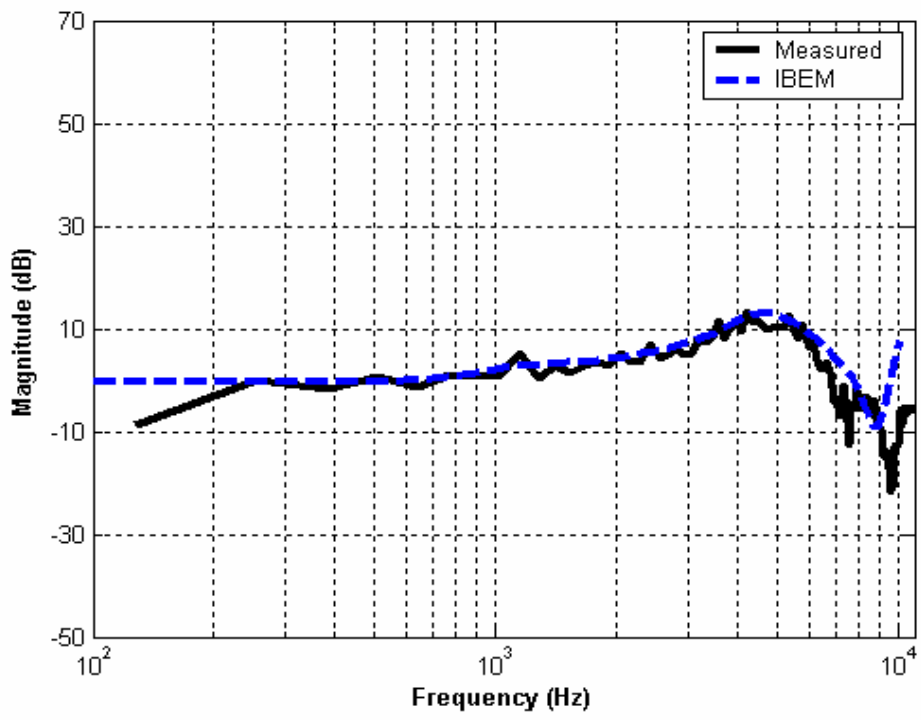


(a)

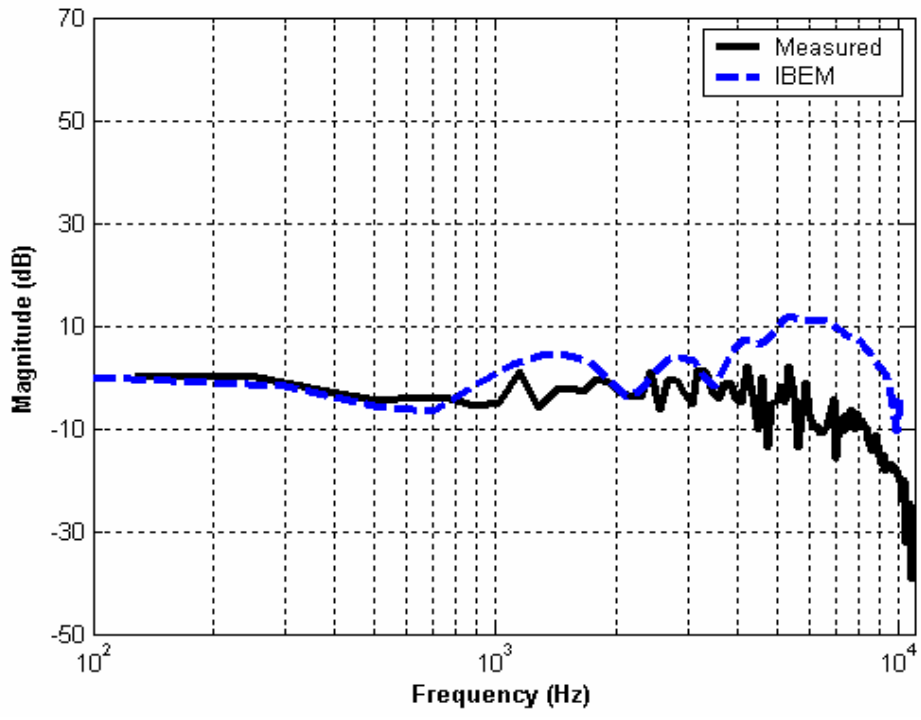


(b)

Fig. 26 Calculated HRTFs on the horizontal plane by using IBEM (Kahana case).
 (a) 0 azimuth degrees (b) 90 azimuth degrees

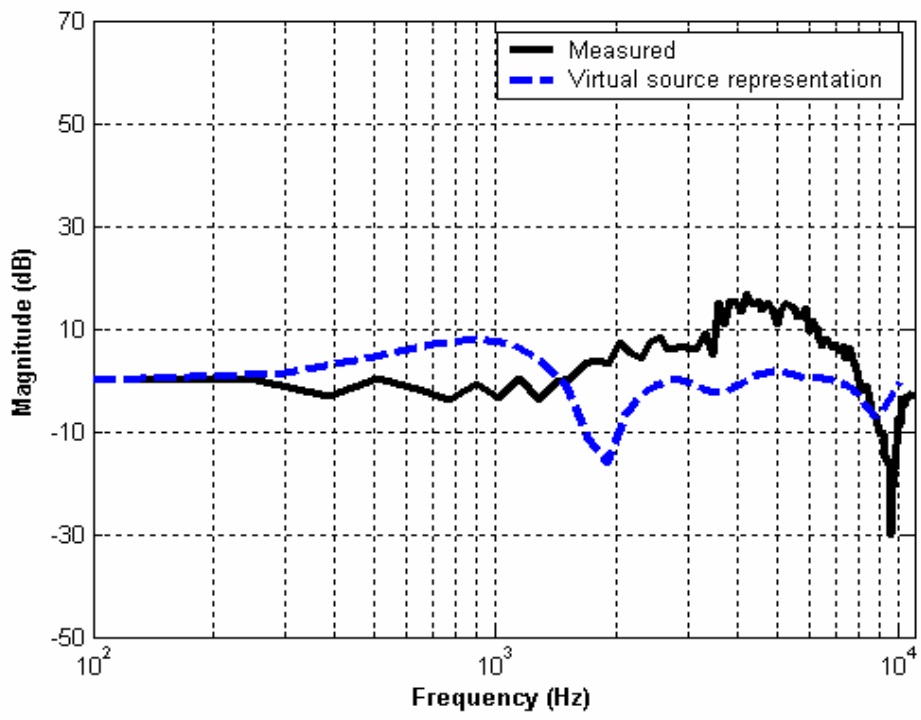


(c)

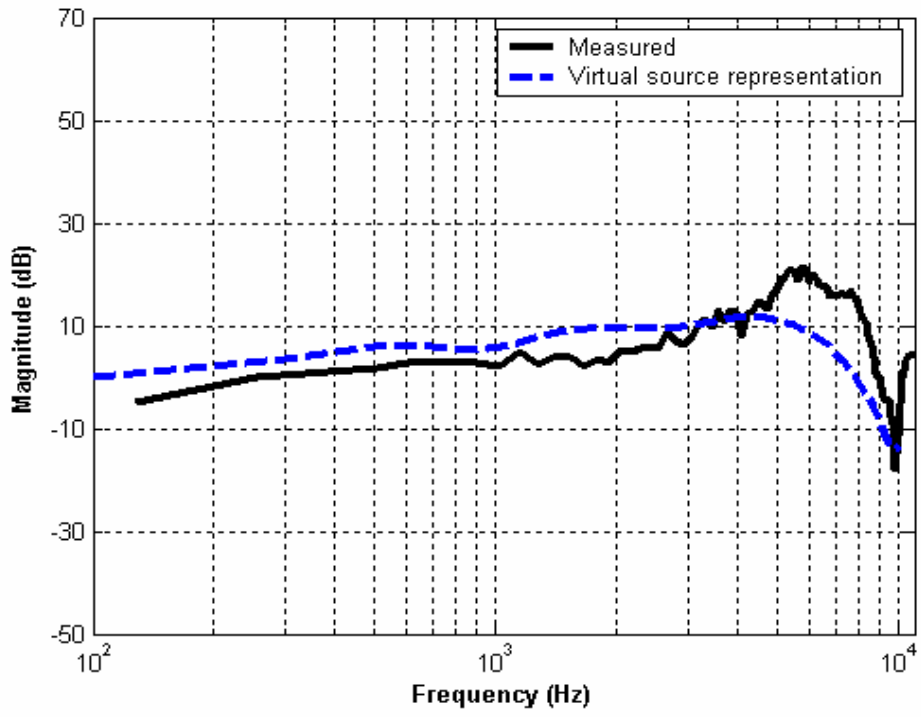


(d)

Fig. 26 Calculated HRTFs on the horizontal plane by using IBEM (Kahana case).
 (c) 180 azimuth degrees (d) 270 azimuth degrees



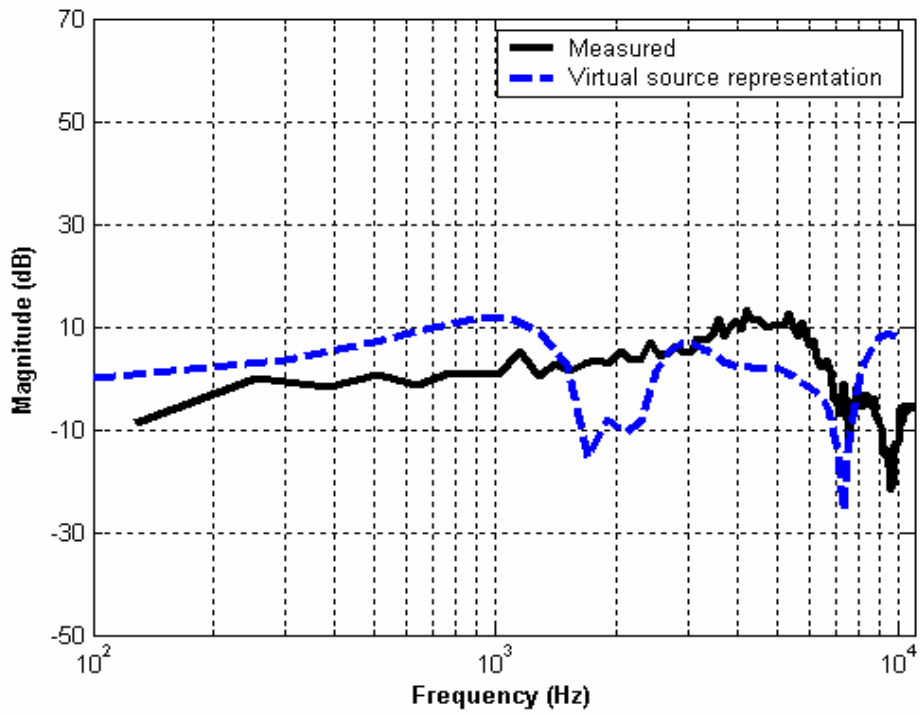
(a)



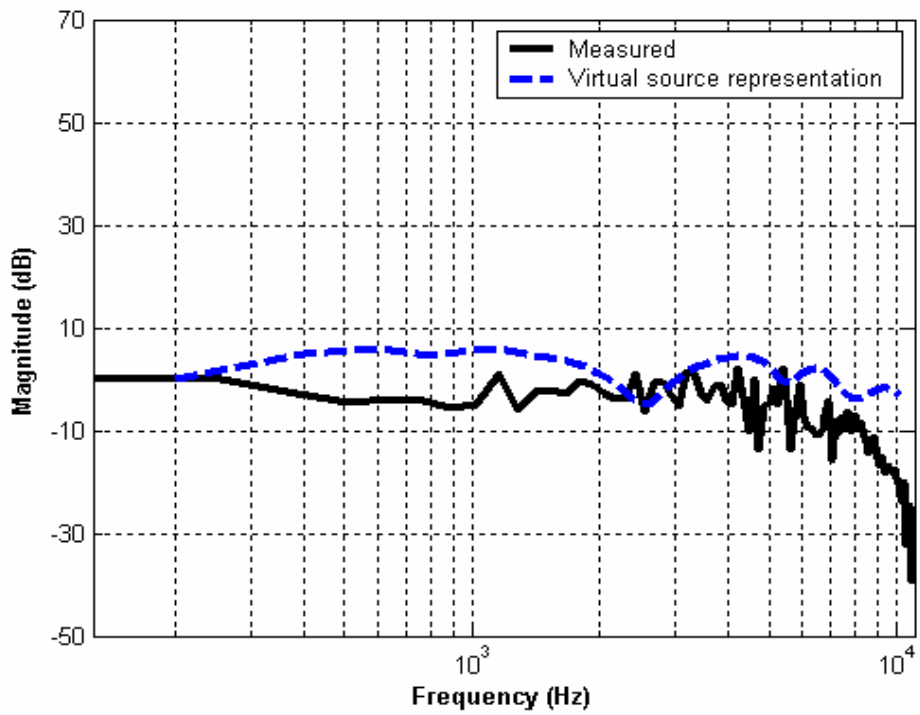
(b)

Fig. 27 Calculated HRTFs on the horizontal plane by using virtual source representation (virtual source case 1).

(a) 0 azimuth degrees (b) 90 azimuth degrees



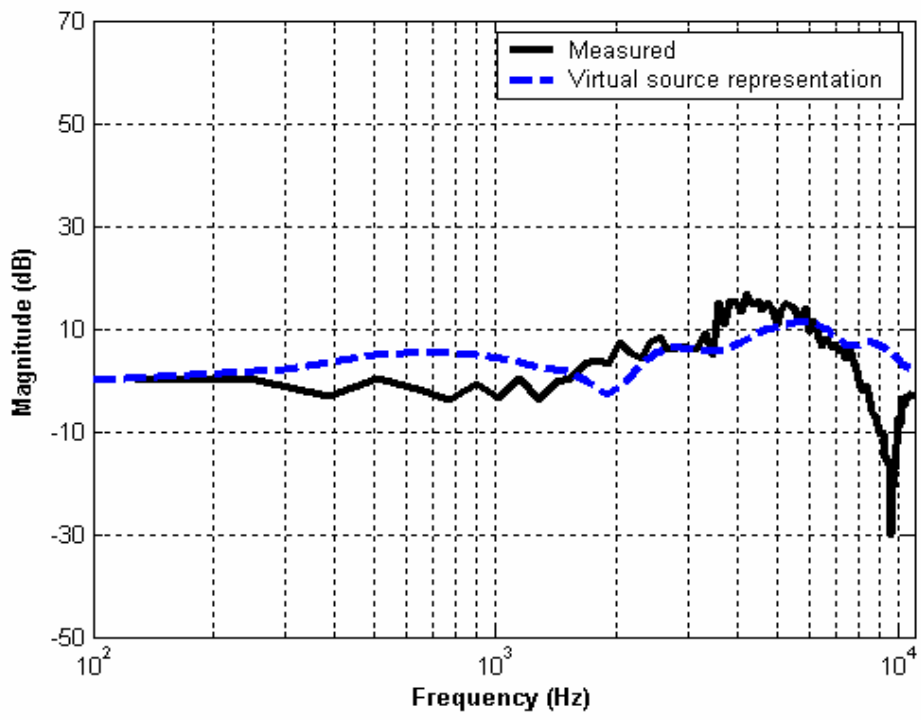
(c)



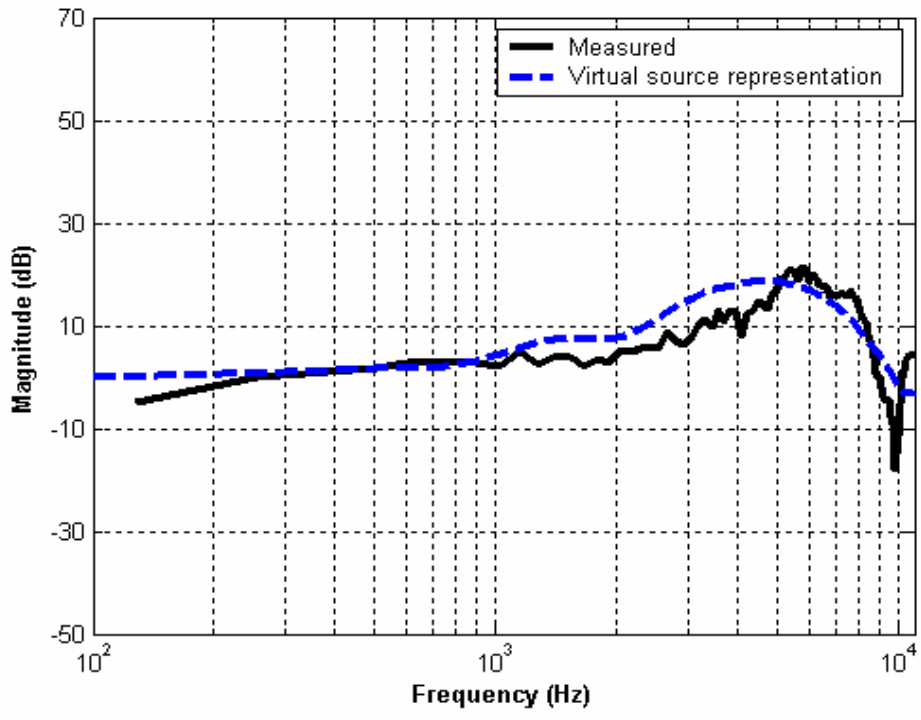
(d)

Fig. 27 Calculated HRTFs on the horizontal plane by using virtual source representation (virtual source case 1).

(c) 180 azimuth degrees (d) 270 azimuth degrees



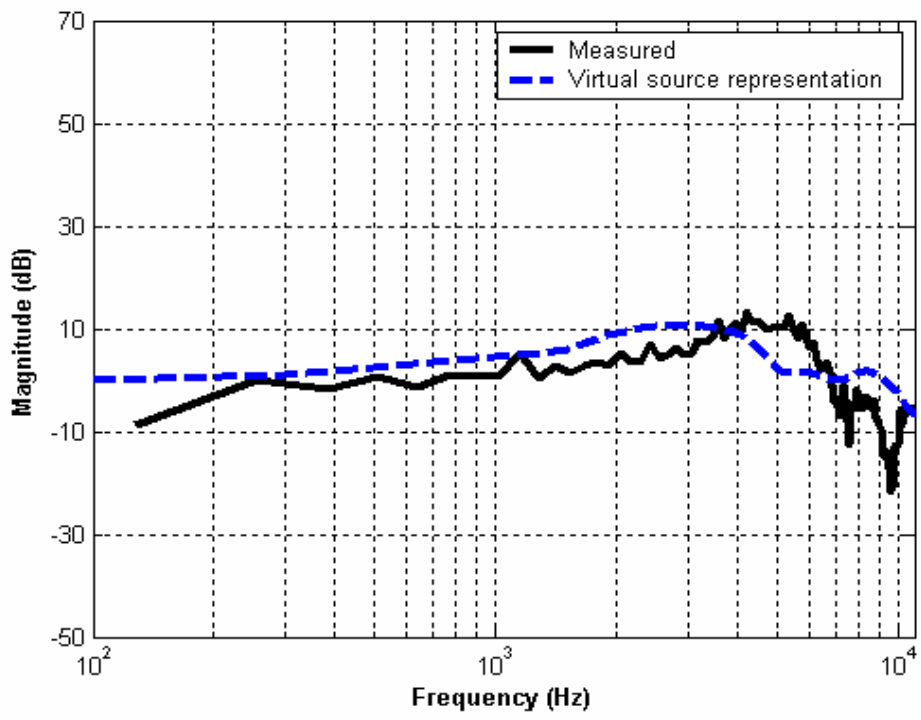
(a)



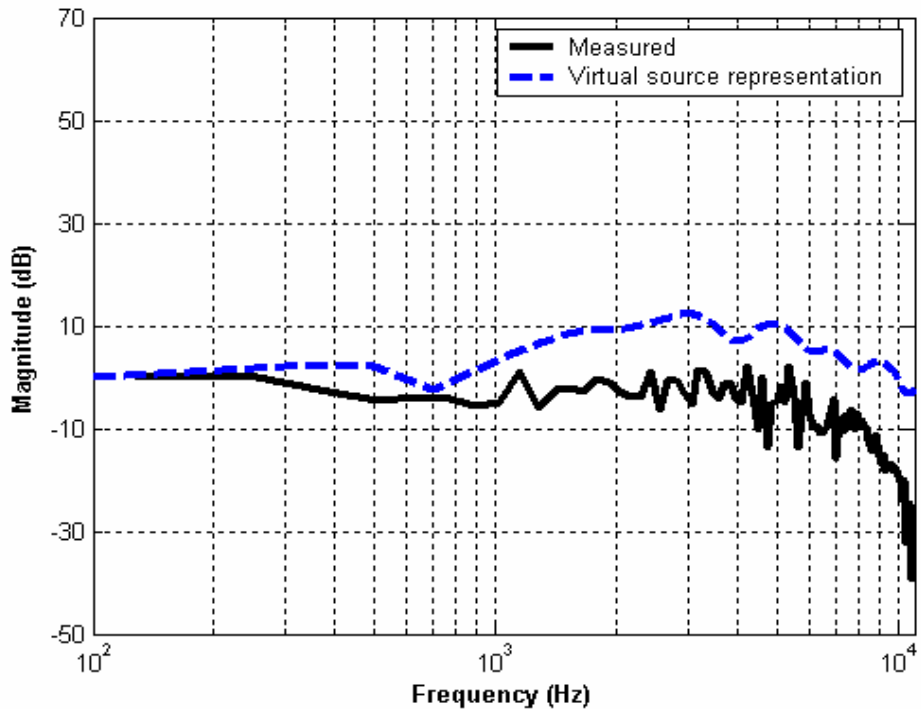
(b)

Fig. 28 Calculated HRTFs on the horizontal plane by using virtual source representation (virtual source case 2).

(a) 0 azimuth degrees (b) 90 azimuth degrees



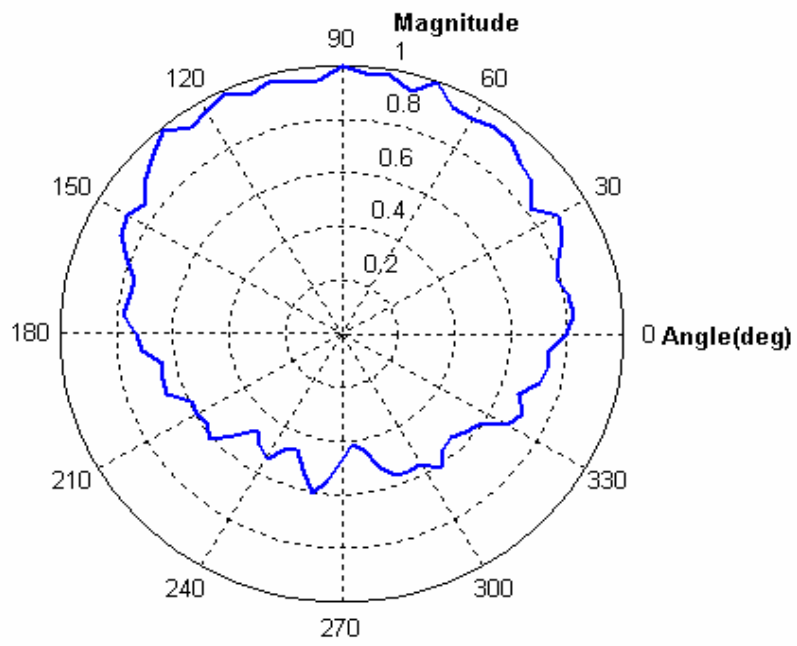
(c)



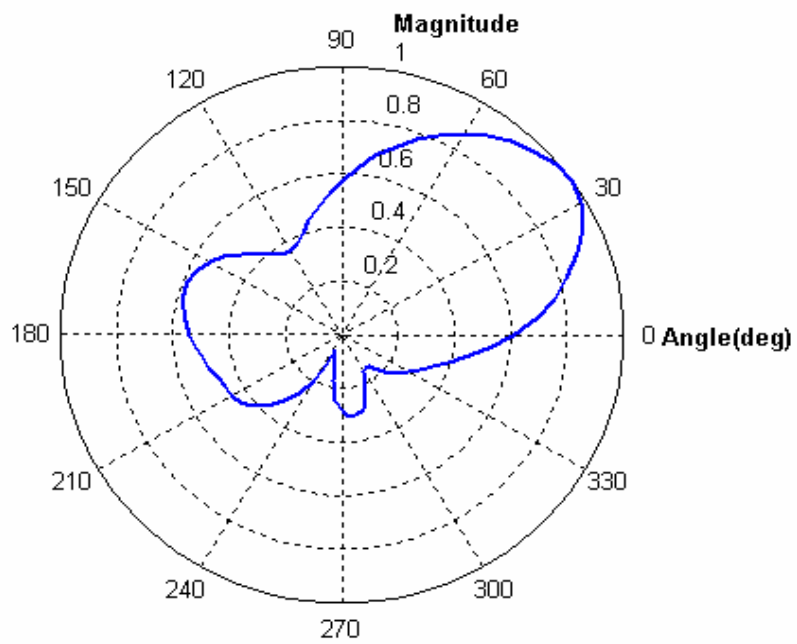
(d)

Fig. 28 Calculated HRTFs on the horizontal plane by using virtual source representation (virtual source case 2).

(c) 180 azimuth degrees (d) 270 azimuth degrees

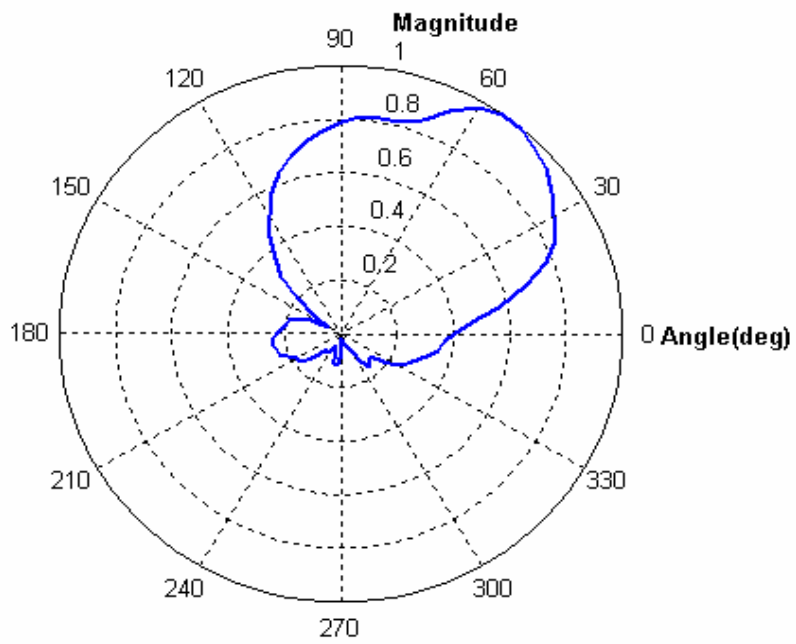


(a)

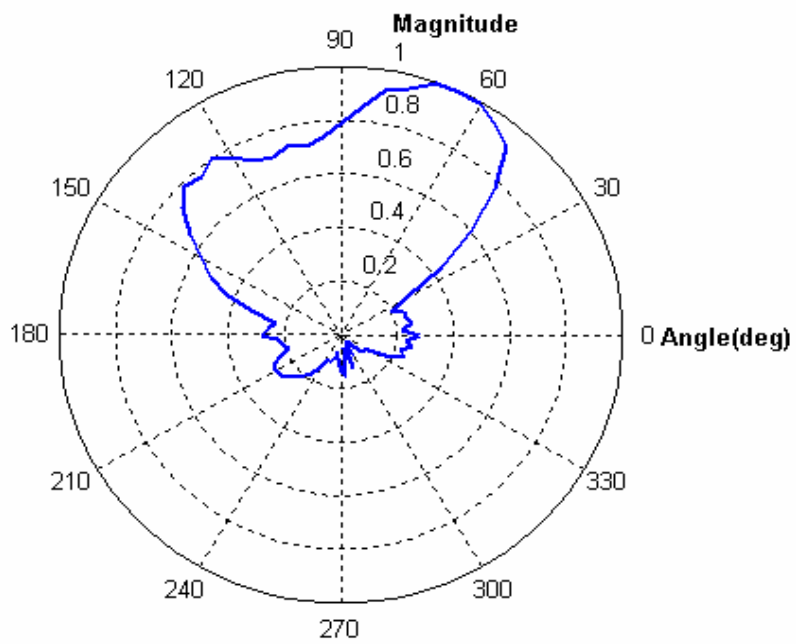


(b)

Fig. 29 Directional response of measured HRTF on the horizontal plane at
 (a) 500 Hz (b) 2100 Hz

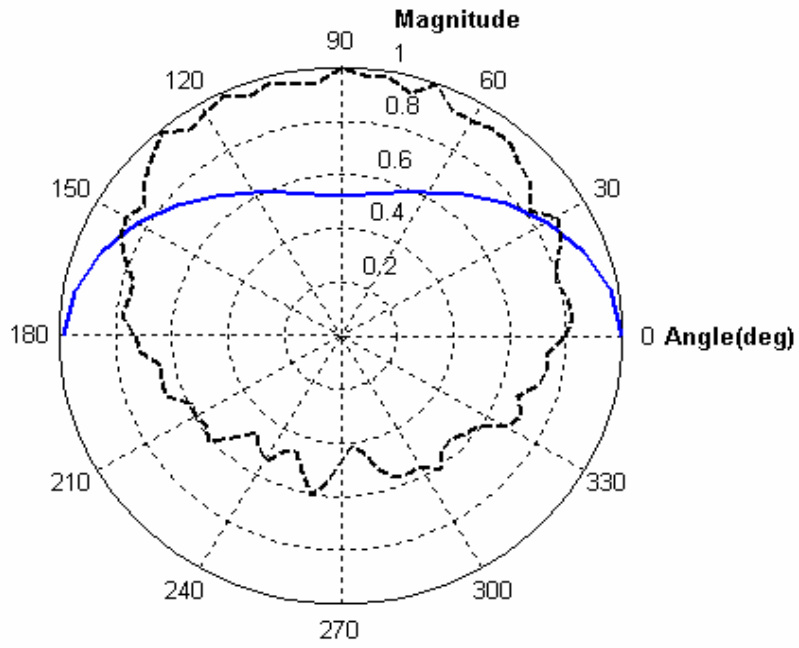


(c)

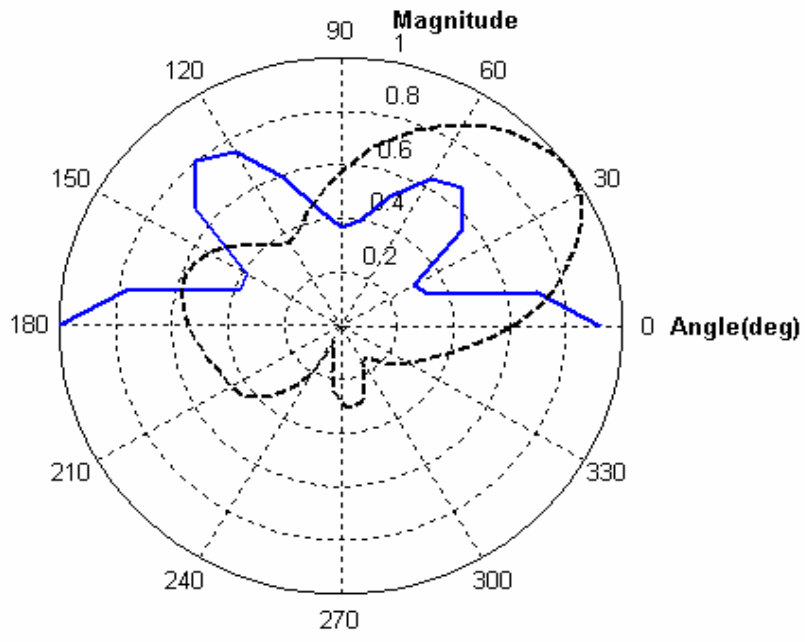


(d)

Fig. 29 Directional response of measured HRTF on the horizontal plane at
 (c) 5100 Hz (d) 9100 Hz

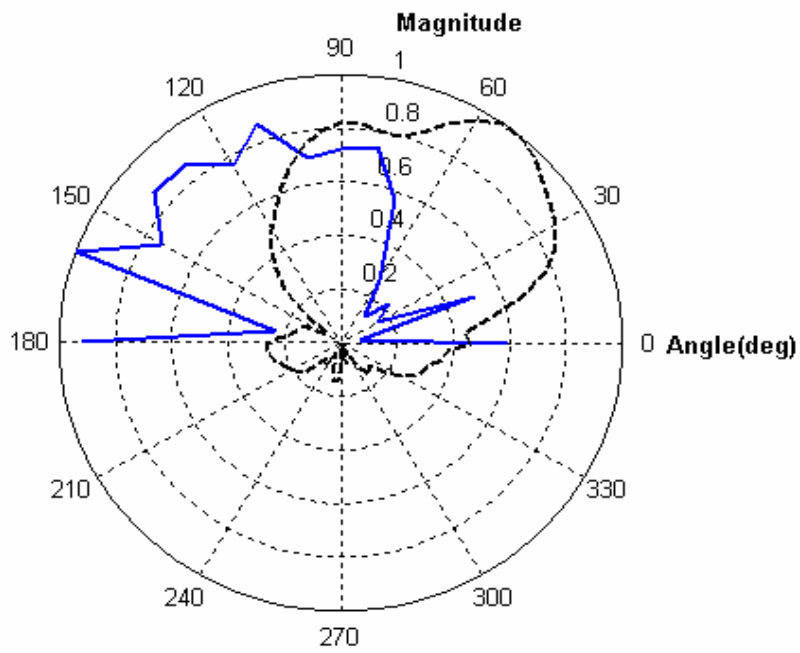


(a)

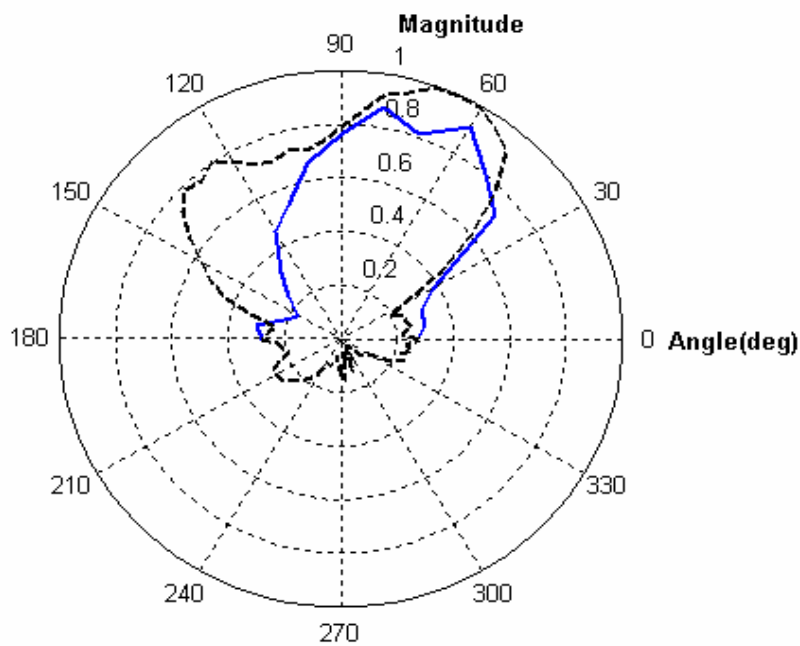


(b)

Fig. 30 Directional response calculated by using IBEM (Kahana case) at
 (a) 500 Hz (b) 2100 Hz
 dash line: measured HRTF solid line: calculated HRTF
 (elevation = 0°)

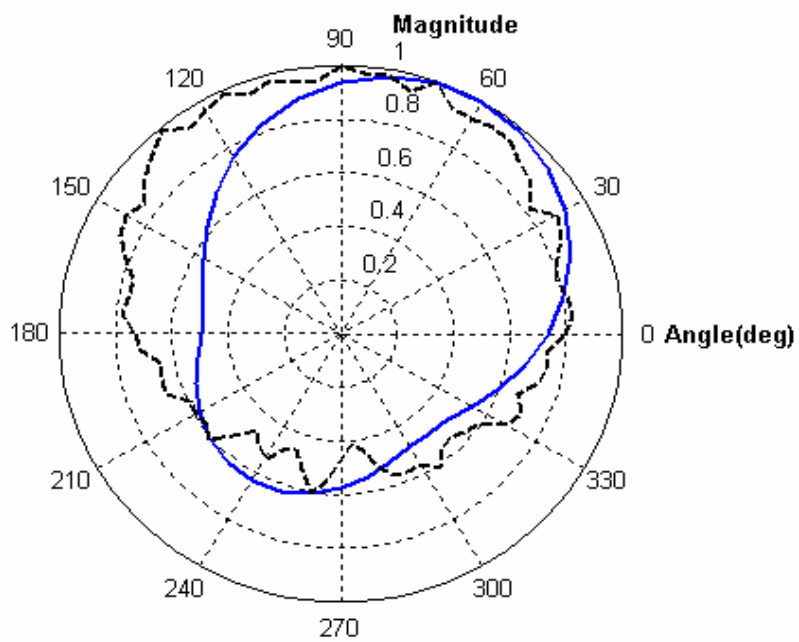


(c)

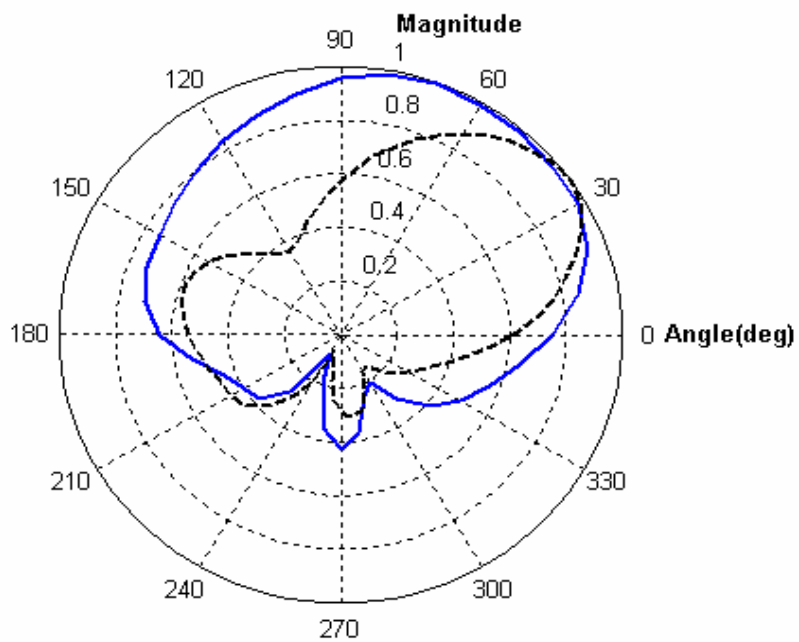


(d)

Fig. 30 Directional response calculated by using IBEM (Kahana case) at
 (c) 5100 Hz (d) 9100 Hz
 dash line: measured HRTF solid line: calculated HRTF
 (elevation = 0°)

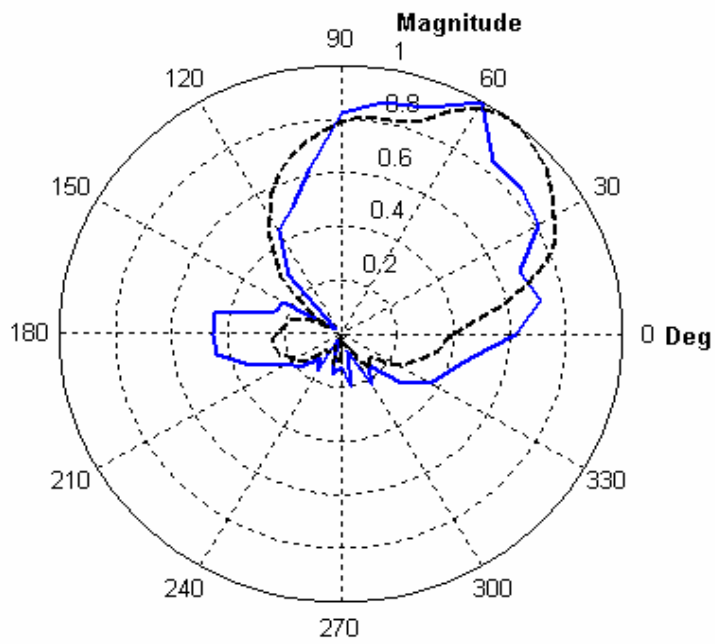


(a)

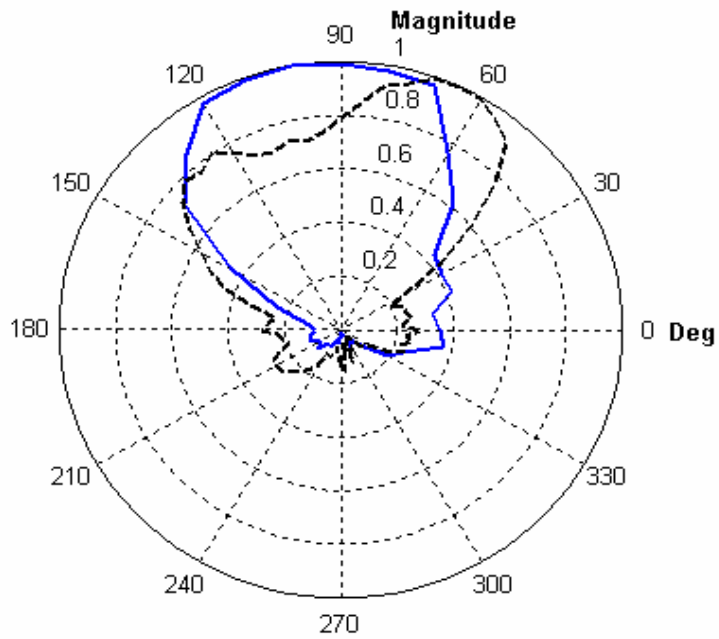


(b)

Fig. 31 Directional response calculated by using IBEM (ITRI case) at
 (a) 500 Hz (b) 2100 Hz
 dash line: measured HRTF, solid line: calculated HRTF
 (elevation = 0°)

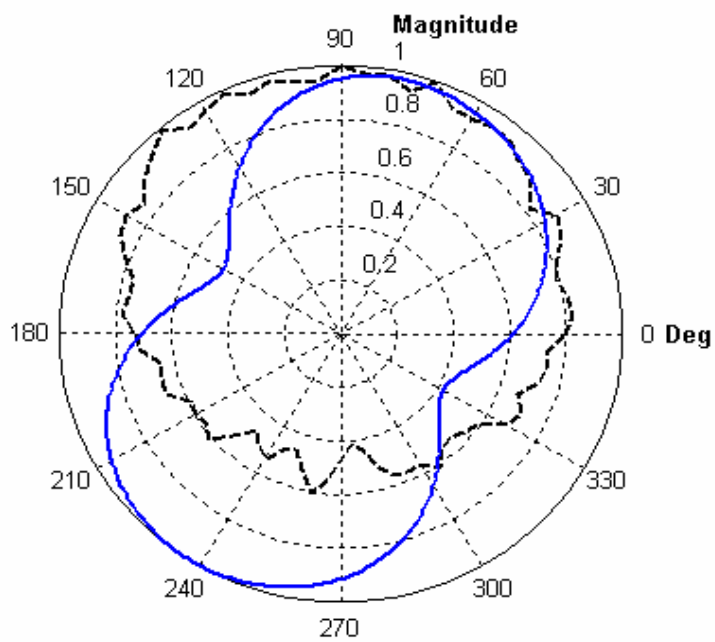


(c)

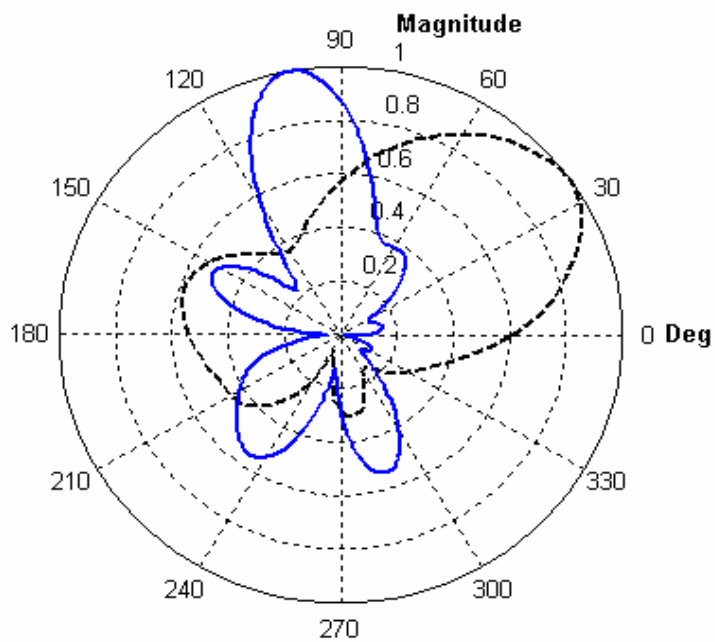


(d)

Fig. 31 Directional response calculated by using IBEM (ITRI case) at
 (c) 5100 Hz (d) 9100 Hz
 dash line: measured HRTF, solid line: calculated HRTF
 (elevation = 0°)

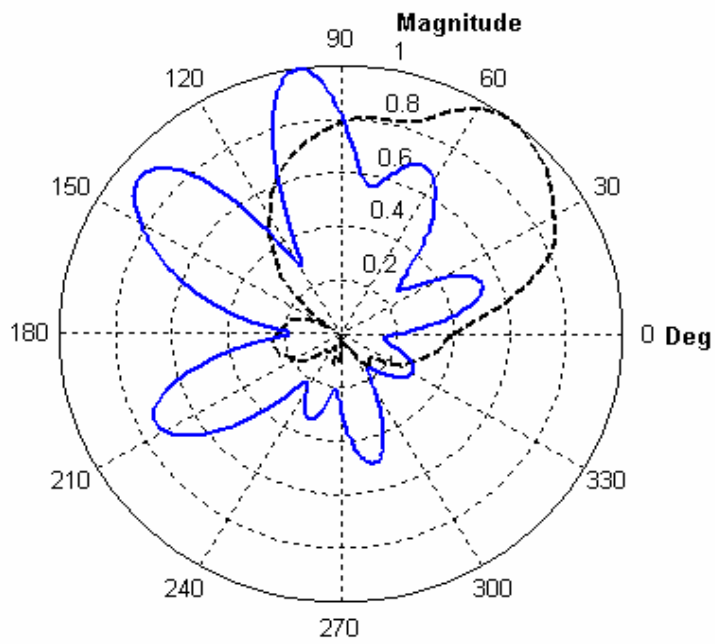


(a)

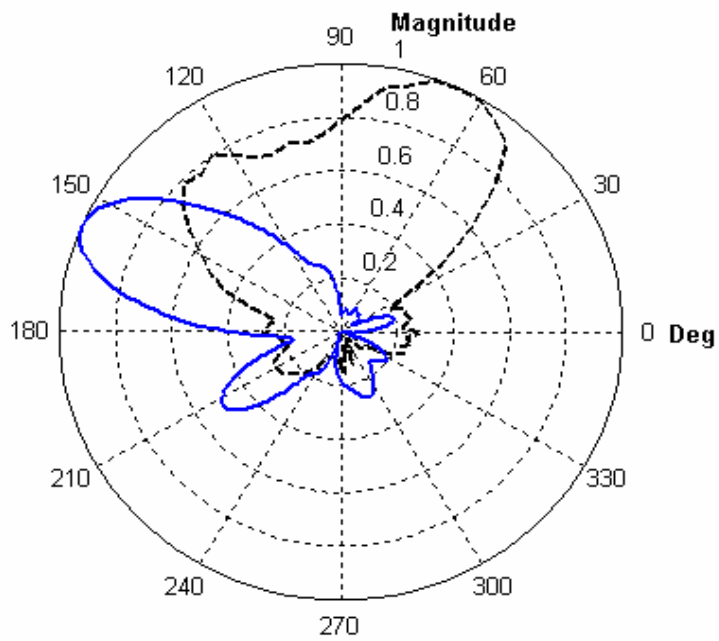


(b)

Fig. 32 Directional response on the horizontal plane calculated by using virtual source representation (virtual source case 1) at
 (a) 500 Hz (b) 2100 Hz
 (dash line: measured HRTF, solid line: calculated HRTF)

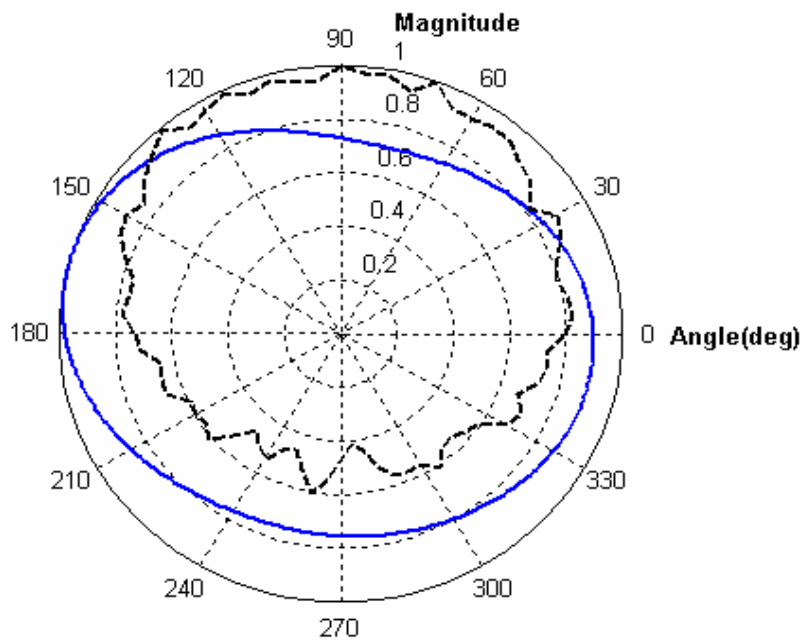


(c)

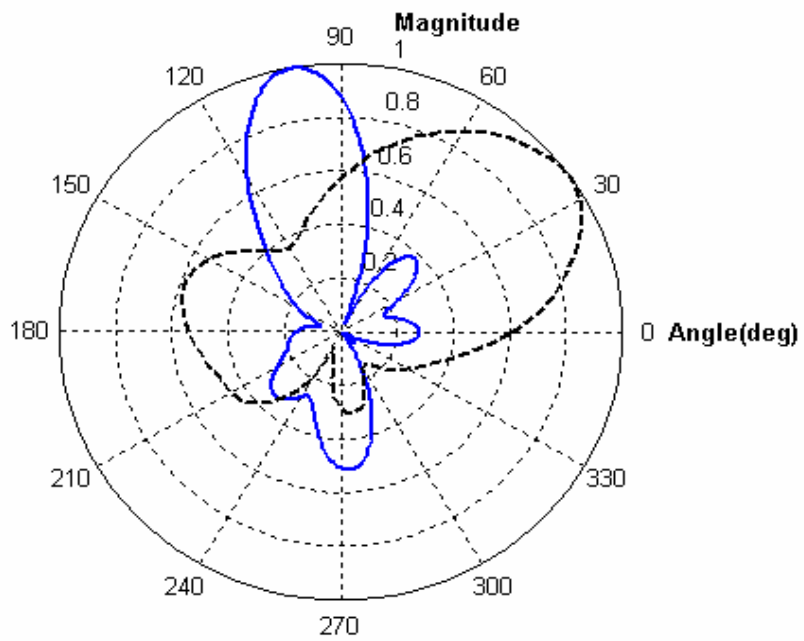


(d)

Fig. 32 Directional response on the horizontal plane calculated by using virtual source representation (virtual source case 1) at
(c) 5100 Hz (d) 9100 Hz
(dash line: measured HRTF, solid line: calculated HRTF)

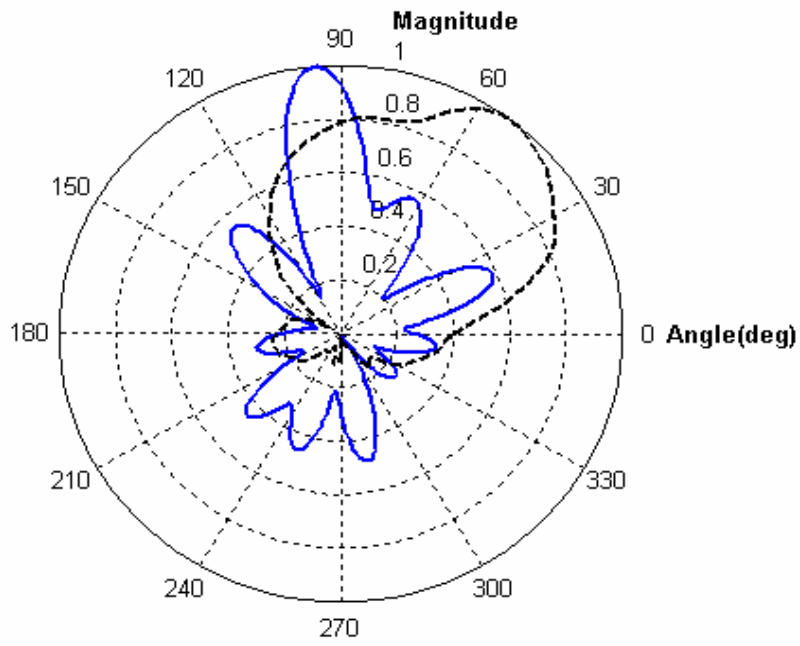


(a)

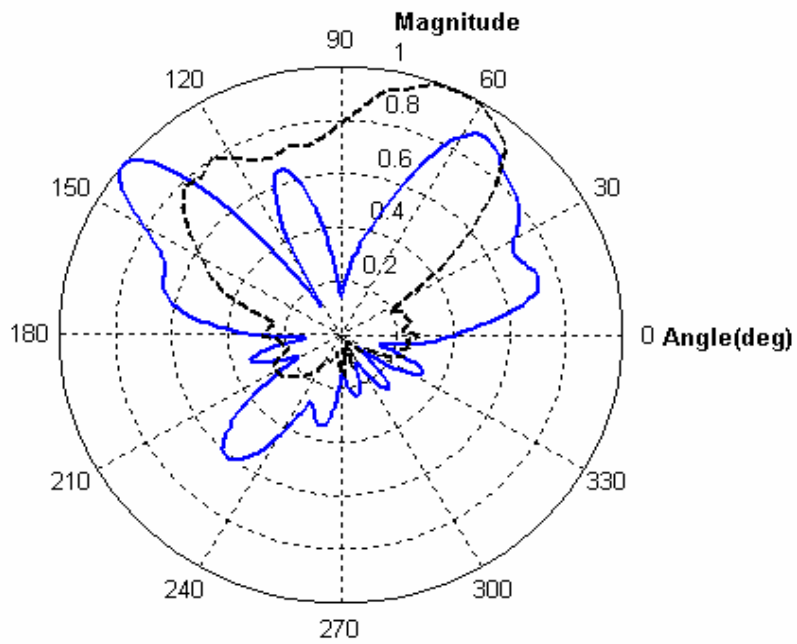


(b)

Fig. 33 Directional response on the horizontal plane calculated by using virtual source representation (virtual source case 2) at
 (a) 500 Hz (b) 2100 Hz
 (dash line: measured HRTF, solid line: calculated HRTF)



(c)



(d)

Fig. 33 Directional response on the horizontal plane calculated by using virtual source representation (virtual source case 2) at
(c) 5100 Hz (d) 9100 Hz
(dash line: measured HRTF, solid line: calculated HRTF)

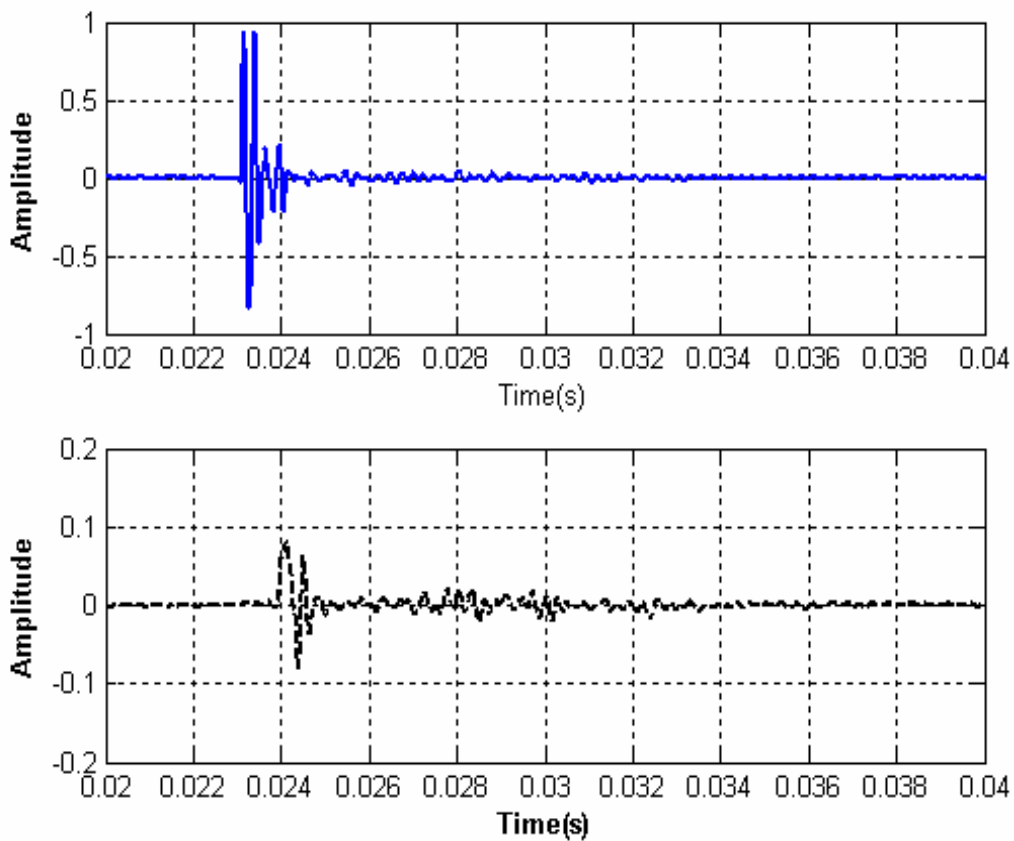


Fig. 34 Head related impulse response on the horizontal plane from the measured HRTF database.
(solid line: 90 azimuth degrees.)
(dash line: 270 azimuth degrees.)

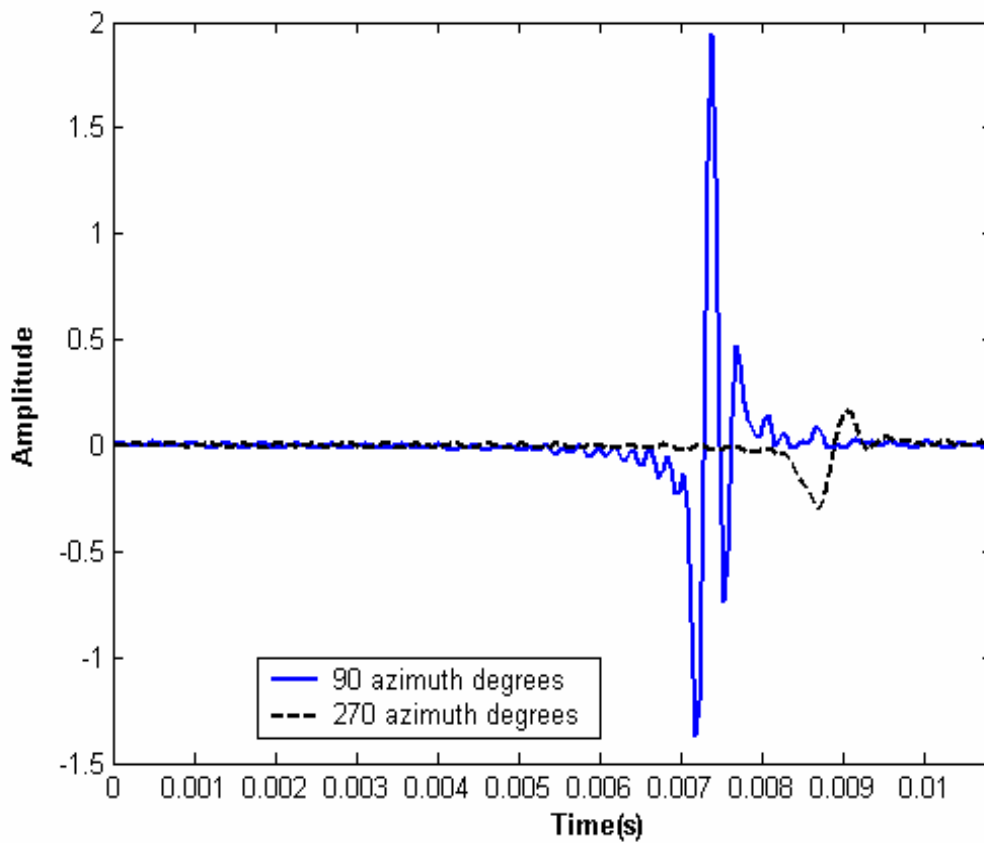


Fig. 35 Head related impulse response on the horizontal plane from the calculated HRTFs by using IBEM. (ITRI case).

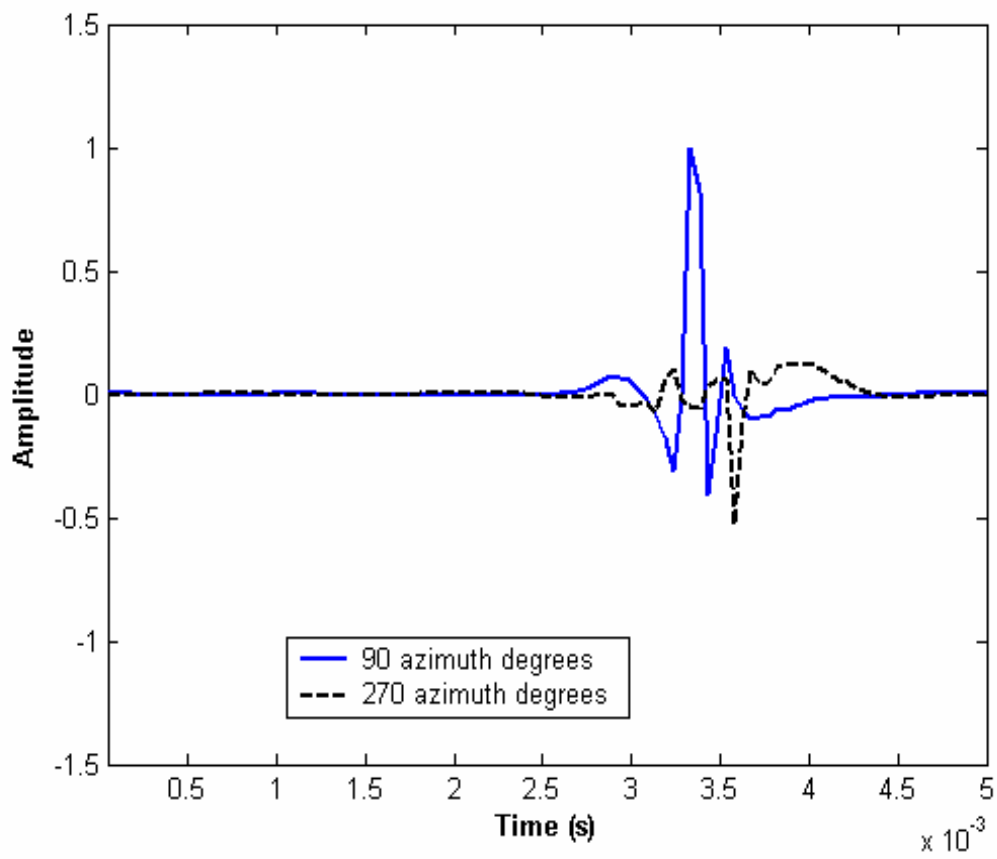


Fig. 36 Head related impulse response on the horizontal plane from the calculated HRTFs by using virtual source representation (case 1).

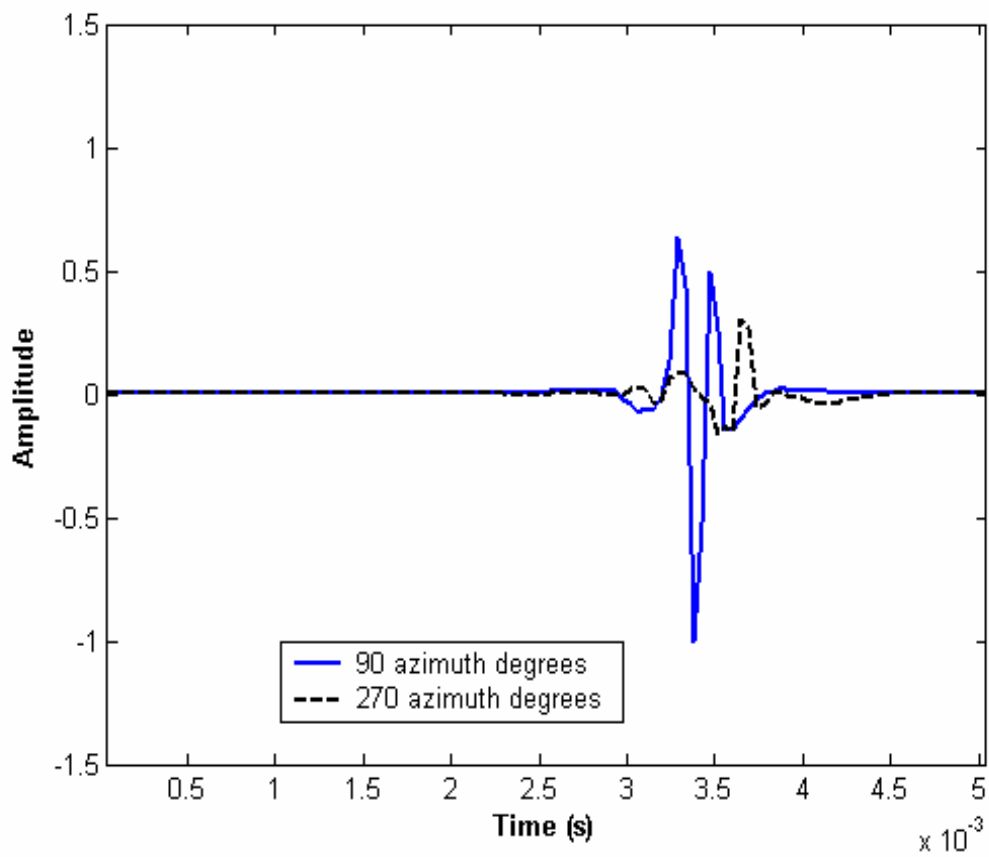
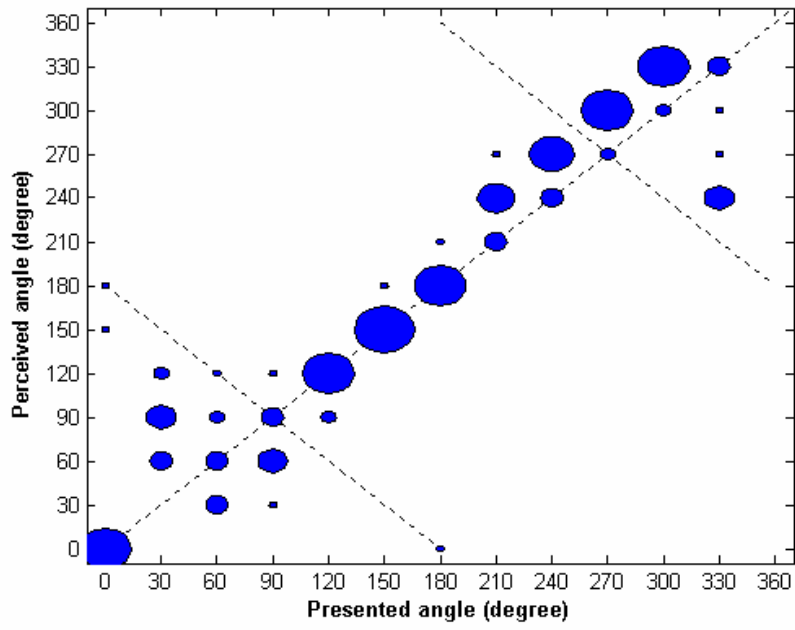
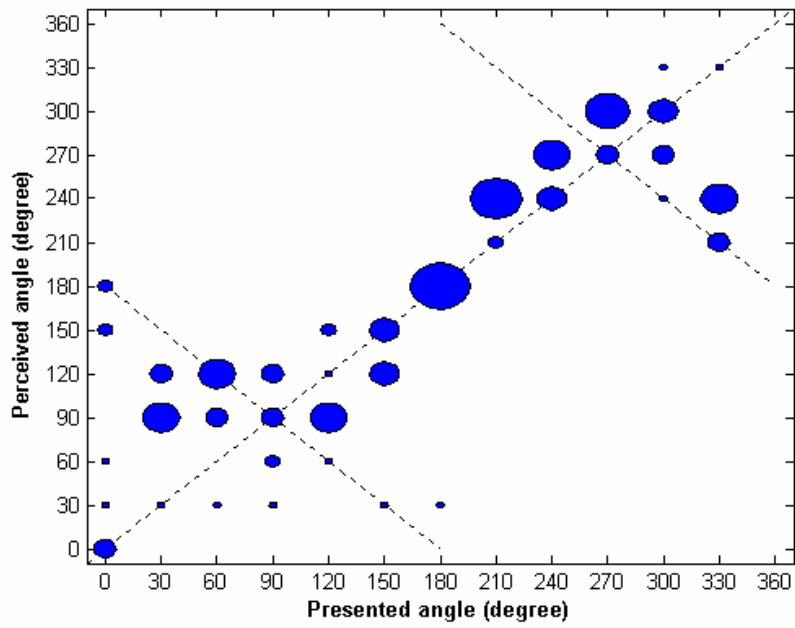


Fig. 37 Head related impulse response on the horizontal plane from the calculated HRTFs by using virtual source representation (case 2).

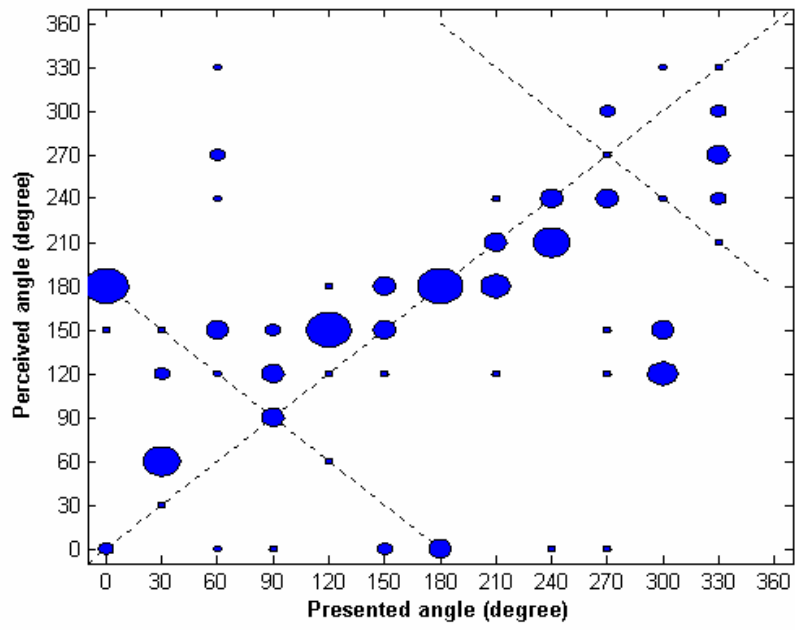


(a)

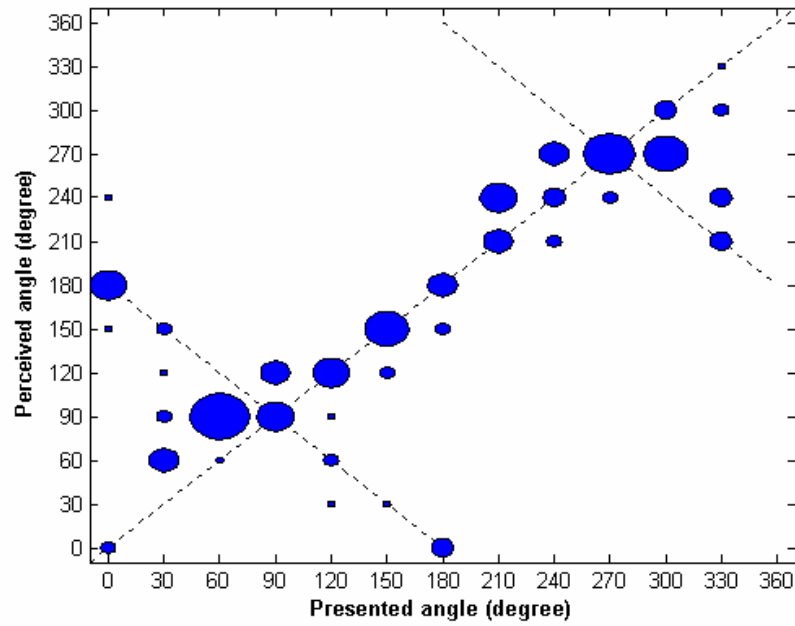


(b)

Fig. 38 Sound localization results of an azimuth test using white noise
 (a) Measured HRTFs (b) HRTFs from MIT HRTF database



(c)



(d)

Fig. 38 Sound localization results of an azimuth test using white noise
 (c) Calculated HRTFs of virtual source case 1
 (d) Calculated HRTFs by using IBEM (ITRI case)

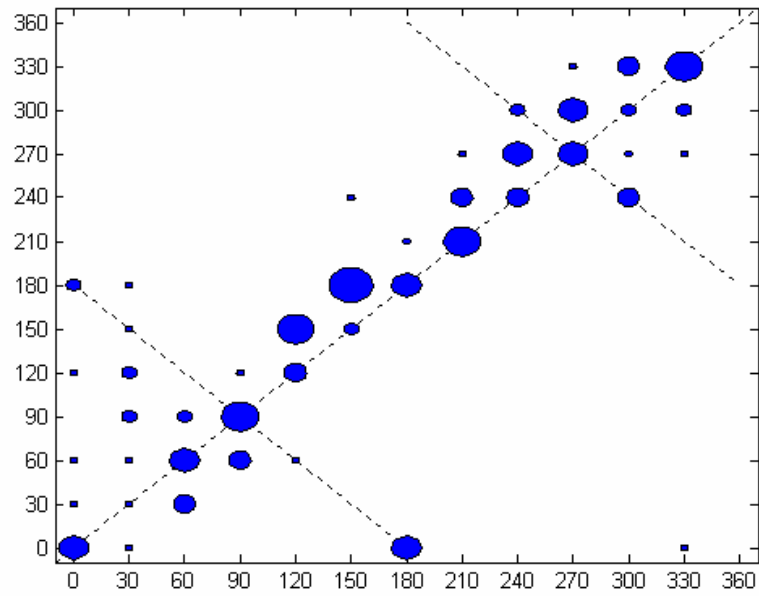
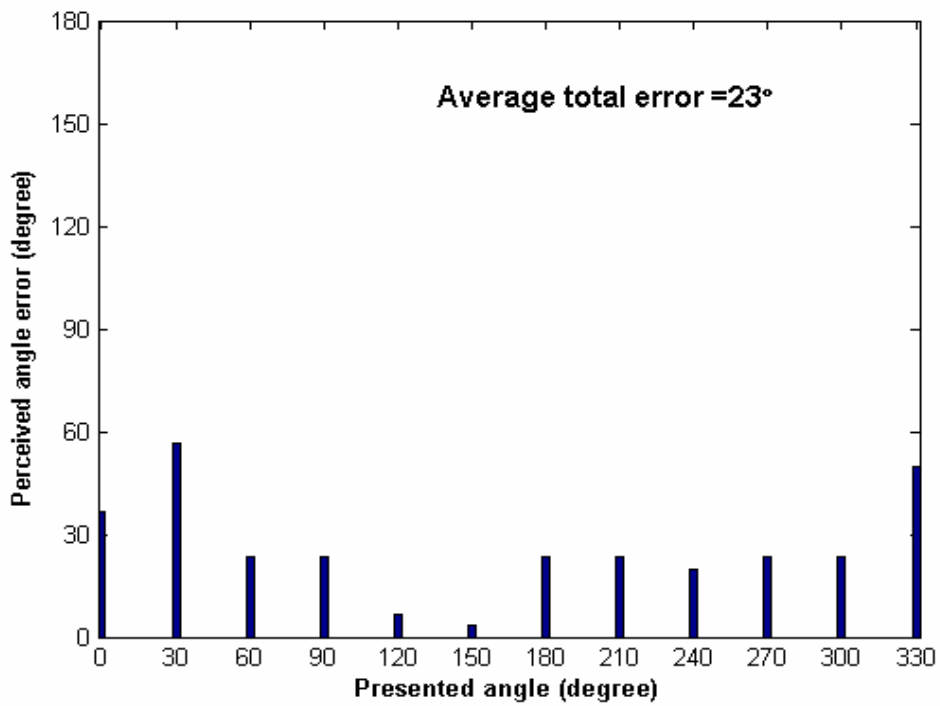
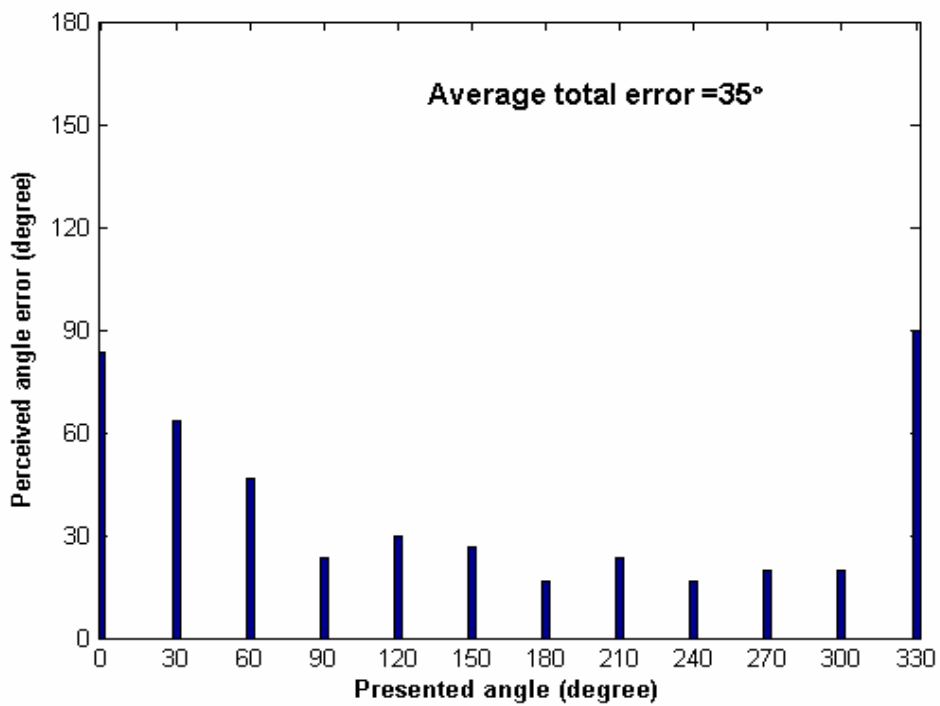


Fig. 38 Sound localization results of an azimuth test using white noise
 (e) Calculated HRTFs of virtual source case 2





(a)

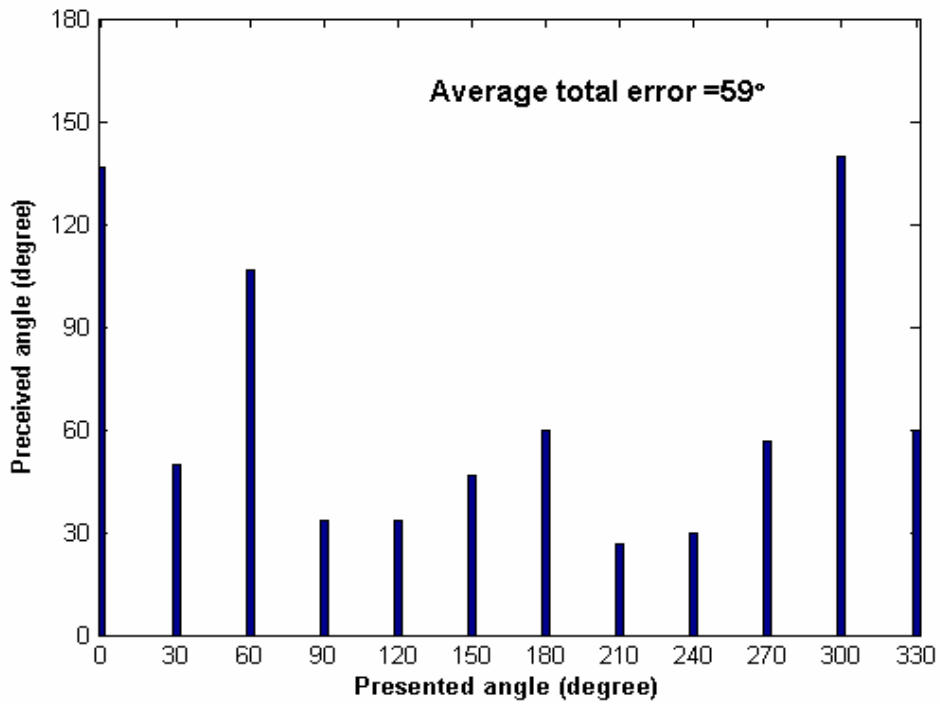


(b)

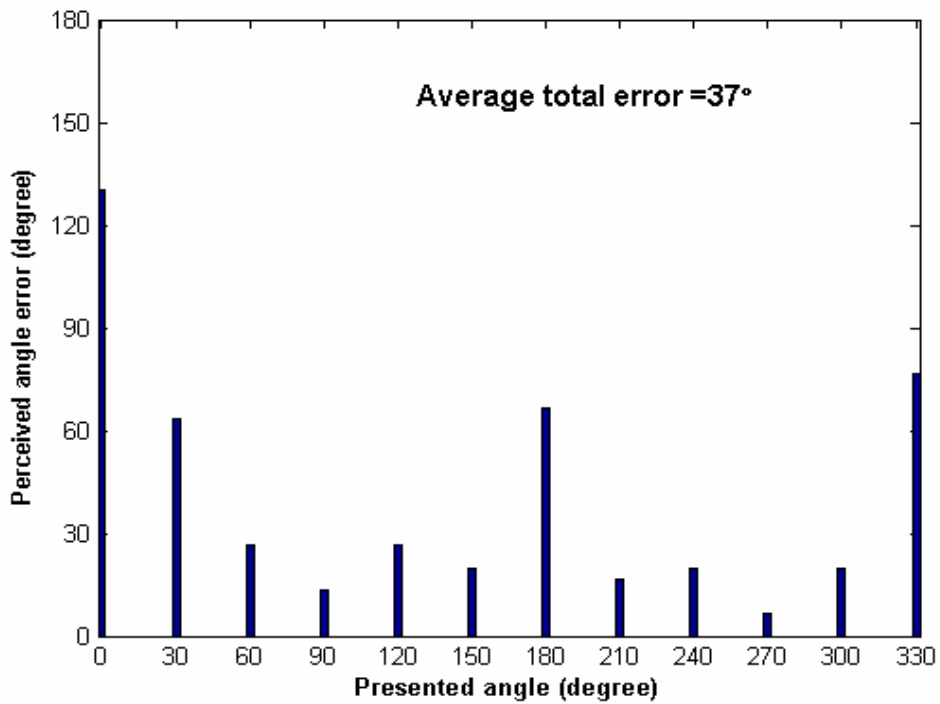
Fig. 39 The total average error and perceived angle error at each direction of sound localization

(a) Measured HRTFs

(b) HRTFs from MIT HRTF database



(c)

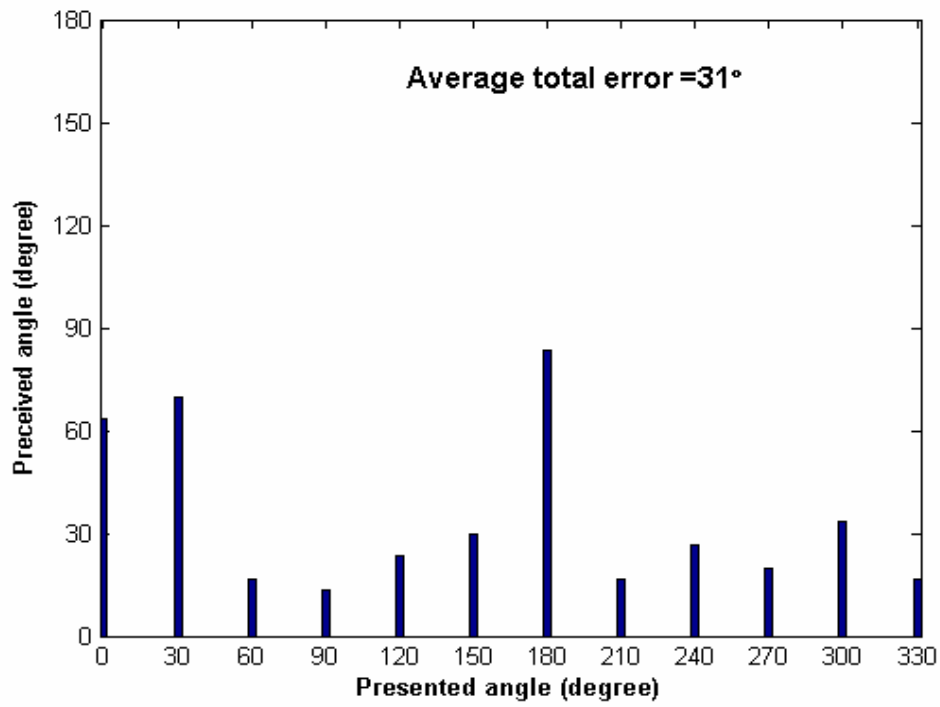


(d)

Fig. 39 The total average error and perceived angle error at each direction of sound localization

(c) Calculated HRTF of virtual source case 1

(d) Calculated HRTF by using IBEM (ITRI case)



(e)

Fig. 39 The total average error and perceived angle error at each direction of sound localization

(e) Calculated HRTF of virtual source case 2

