Numerical Modeling of Head-Related Transfer Functions Using the Boundary Source Representation

Mingsian R. Bai[1]
e-mail: msbai@mail.nctu.edu.tw

Teng-Chieh Tsao

Department of Mechanical Engineering,
National Chiao-Tung University,
1001 Ta-Hsueh Road,
Hsin-Chu 300 Taiwan, Republic of China

A technique based on the virtual source representation is presented for modeling head-related transfer functions (HRTFs). This method is motivated by the theory of simple layer potential and the principle of wave superposition. Using the virtual source representation, the HRTFs for a human head with pinnae are calculated with a minimal amount of computation. In the process, a special regularization scheme is required to calculate the equivalent strengths of virtual sources. To justify the proposed method, tests were carried out to compare the virtual source method with the boundary element method (BEM) and a direct HRTF measurement. The HRTFs obtained using the virtual source method agrees reasonably well in terms of frequency response, directional response, and impulse response with the other methods. From the numerical perspectives, the virtual source method obviates the singularity problem as commonly encountered in the BEM, and is less computationally demanding than the BEM in terms of computational time and memory storage. Subjective experiments are also conducted using the calculated and the measured HRTFs. The results reveal that the spatial characteristics of sound localization are satisfactorily reproduced as a human listener would naturally perceive by using the virtual source HRTFs. [DOI: 10.1115/1.2203337]

1 Introduction

Head-related transfer function (HRTF) is a measurement of the transformation for a specific source direction relative to the head, and describes the filtering process associated with the diffraction of sound by the torso, head, and pinna. The HRTF is a chief ingredient of spatial sound reproduction. Three-dimensional sound fields can be created by convolving the source signal with an appropriate pair of head-related impulse responses (HRIRs) defined as the inverse Fourier transform of HRTFs.

There are basically two ways for obtaining HRTFs: Direct measurement and numerical modeling. Several measured HRTF databases are available on the internet [1,2]. Despite its effectiveness, the measured HRTFs have several disadvantages. First, the measurement requires a special environment, such as an anechoic chamber, and the fact that the measuring process is time consuming. Second, errors due to instrumentation, data acquisition, positioning of transducers, postprocessing, etc., may affect the localization performance using such HRTFs. Apart from the measured HRTFs, people are seeking to model HRTFs using numerical means. The shortcomings associated with the above-mentioned measured HRTFs are totally eliminated when using numerical approaches. Batterau modeled the external ear as a three-channel two-delay and sum acoustic coupler [3,4]. In a series of papers, Duda suggested various structural models of HRTF [5–7]. Diffractions due to head, pinna, and torso are modeled by simple function blocks. Among the previous research of numerical HRTF modeling, the work by Kahana et al. [8,9] must be mentioned. They conducted a thorough investigation on many aspects of HRTFs by using the boundary element method (BEM). Although BEM is somewhat a brute-force approach, fairly good agreement between the calculated and the measured HRTFs was obtained in their study. The HRTFs obtained in this work were not actually implemented to evaluate the performance in source localization. Another BEM-based model was also reported in the paper by Walsh et al. [10]. The effects of external head, ear canal, and eardrum are accounted for in their model. In viewing these numerical modeling approaches with varying complexity, there is apparently a trade-off that one has to reconcile between the computation efficiency and the modeling accuracy.

A new modeling technique of HRTFs based on virtual source representation is proposed in this paper to reach the best compromise of the computation efficiency and the modeling accuracy. The virtual source representation is motivated by the layer potential theory, where the acoustic field radiated by an arbitrarily shaped radiator can be described by using the principle of wave superposition [11–14]. Using this method, HRTFs at different directions are represented by a finite number of discrete virtual sources with different source strengths. Complex strengths of the virtual sources can be calculated by solving a matrix inversion problem with appropriate regularization [15]. The HRTFs for a given head with pinna are calculated with minimal amount of computation.

To justify the proposed method, tests were carried out to compare the virtual source method with the indirect BEM (IBEM) [8] and a direct HRTF measurement. LMS SYSNOISE [16,17] is employed for the IBEM simulation. The comparison of the calculated HRTFs and the measured HRTFs is focused on the frequency response, directional response, and impulse response. In addition, a subjective localization experiment was also conducted. The results of the comparison will be discussed in terms of localization performance and computational requirements.

2 Numerical Modeling of HRTFs

Calculation of HRTFs is a scattering problem in which case a scattered field at the ear position is being sought when the source is located in the far field. Owing to acoustic reciprocity, this scattering problem can be transformed to a radiation problem by interchanging the roles of the source and receiver. That is, the source is positioned at the ear opening of the listener, while the

[1]Author to whom correspondence should be addressed.

Contributed by the Technical Committee on Vibration and Sound of ASME for publication in the JOURNAL OF VIBRATION AND ACOUSTICS. Manuscript received March 4, 2005; final manuscript received April 4, 2006. Assoc. Editor: Roger Ohayon.

receiver is positioned at the far field. It follows that the problem to be solved can be stated as: Given the embedded source and the boundary condition of the head such as surface velocity, one seeks to determine the unknown field representation. This problem *per se* is an inverse problem. In this paper, two numerical methods are employed to model HRTFs. The first method is IBEM that is detailed in literature such as Ref. [8] and is thus omitted here. In this section, we shall focus primarily on the numerical technique based on the virtual source representation.

The virtual source method is motivated by the layer potential theory and the principle of wave superposition [11,14]. The single layer potential representation is selected because of several reasons. First, the connection between the continuous integral and the discrete monopole source representation is easily established using such an approach. Second, the source strengths of monopoles serve only as intermediate variables that bear no physical meaning to the actual pressure or velocity on the boundary. Third, it is easier to handle the single layer representation than the double layer representation which may present mathematical complications, such as singularity of gradient evaluation. In the single layer potential representation, the pressure inside or outside of the surface can be represented by the single layer of monopoles

$$p(\mathbf{x}) = \int_s \sigma(\mathbf{x}_0) G(\mathbf{x},\mathbf{x}_0) dS(\mathbf{x}_0) \tag{1}$$

where $G(\mathbf{x},\mathbf{x}_0)$ is the free space Green's function corresponding to the source and the field points, $\mathbf{x}$ and $\mathbf{x}_0$, respectively,

$$G(\mathbf{x},\mathbf{x}_0) = \frac{e^{-jkr}}{4\pi r} \tag{2}$$

with $r=|\mathbf{x}-\mathbf{x}_0|$, $k=\omega/c$ is the wave number, c is the speed of the sound, and σ is an unknown source strength of the monopole distribution on the boundary. The pressures are continuous across the boundary in this representation, while the velocities across the boundary are discontinuous

$$\frac{\partial p}{\partial n}(\mathbf{x}) = \alpha\sigma(\mathbf{x}) + \int_s \sigma(\mathbf{x}_0) \frac{\partial G}{\partial n}(\mathbf{x},\mathbf{x}_0) dS(\mathbf{x}_0) \tag{3}$$

where

$$\alpha = \begin{cases} -1/2 & \mathbf{x} \in V_o \text{ (exterior)} \\ 1/2 & \mathbf{x} \in V_i \text{ (interior)} \\ 0 & \text{otherwise} \end{cases}$$

Elaborate numerical schemes, such as the finite element method or the BEM can be used for discretization of the above continuous integral equations. However, additional insights can be gained by simply discretizing the integral in Eq. (1) into M elements with area S_m:

$$\begin{aligned} p(\mathbf{x}) &= \int_S (\mathbf{x}_0) G(\mathbf{x},\mathbf{x}_0) dS(\mathbf{x}_0) \approx \sum_{m=1}^{M} \int_{s_m} \sigma(\mathbf{x}_0) G(\mathbf{x},\mathbf{x}_0) dS(\mathbf{x}_0) \\ &\approx \sum_{m=1}^{M} \sigma(\mathbf{s}_m) G(\mathbf{x},\mathbf{s}_m) S_m = \sum_{m=1}^{M} (\sigma(\mathbf{s}_m) S_m) G(\mathbf{x},\mathbf{s}_m) \\ &= \sum_{m=1}^{M} Q_m G(\mathbf{x},\mathbf{s}_m) \end{aligned} \tag{4}$$

This equation is a valid approximation of the integral in Eq. (1) when $M \to \infty$, $S_m \to 0$. The sound field produced by an actual source can be considered equivalent to a collection of the point sources, each with appropriate strength determined by additional conditions. All this boils down to the idea that the acoustic field of a complex radiator can be constructed as a superposition of fields generated by an array of simple sources (monopole) distributed on the virtual surface [11], which is more tractable numerically than the original setting. According to Eq. (4), the total sound pressure produced by these virtual sources is given by

$$p(\mathbf{x}) = \sum_{m=1}^{M} Q_m G(\mathbf{x},\mathbf{s}_m) \tag{5}$$

where G is the free space Green's function as defined before, $\mathbf{x}$ and $\mathbf{s}_m$ are the field and the source points, respectively, the subscript m signifies the mth virtual source, and M is the total number of virtual sources. These virtual sources are chosen to distribute on a fictitious surface that is removed from the radiator of interest. The unknown strength Q_m can be determined by imposing additional conditions, such as velocity, or, pressure gradient conditions. The pressure gradient at the normal direction on the source surface can be evaluated at the position $\mathbf{x}$ according to

$$\frac{\partial p}{\partial n}(\mathbf{x}) = \sum_{m=1}^{M} Q_m \frac{\partial G}{\partial n}(\mathbf{x},\mathbf{s}_m) \tag{6}$$

In Eq. (6), the term $\alpha\sigma(\mathbf{x})$ is not present because the fictitious source surface can be chosen to be inside the real radiator such that the field points never coincide with the source points, and so there is no velocity discontinuity. This simple method provides advantages in that the numerical complications inherent to BEM-based approaches, such as the singularity problem, can be circumvented. Then, choosing a finite number of field points, $\mathbf{x}_n$, $n=1,2,\ldots,N$, Eqs. (5) and (6) can be rewritten into the following equations

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

and

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

In matrix form,

$$\mathbf{p} = \mathbf{G}\mathbf{q} \tag{7}$$

and

$$\mathbf{p}_n = \mathbf{G}_n\mathbf{q} \tag{8}$$

Given the velocity boundary condition $\mathbf{p}_n$, the unknown source strength matrix $\mathbf{q}$ can now be determined by solving the matrix inverse problem of Eq. (8). The radiation field at any point can then be calculated using a direct substitution into Eq. (7).

2.1 Technical Issues of Virtual Source Representation

2.1.1 Regularization of Virtual Source Representation. Solving the foregoing problem in Eq. (8) for the source strengths of monopoles is a typical inverse problem. The quality of solutions hinges entirely on the conditioning of the matrix $\mathbf{G}_n$. This matrix could be very ill conditioned when the distance between the head and the virtual source surface is large, when the number and geo-

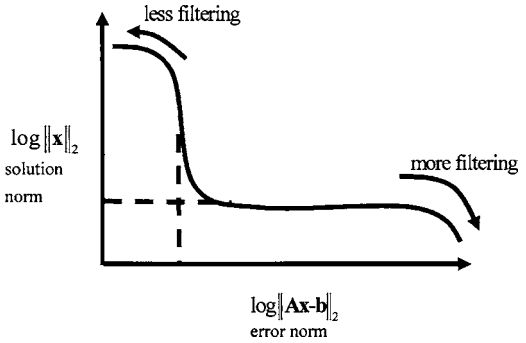


Fig. 1 A generic L-curve. The solution norm is plotted versus the error norm with varying regularization parameter in the log-log scale.

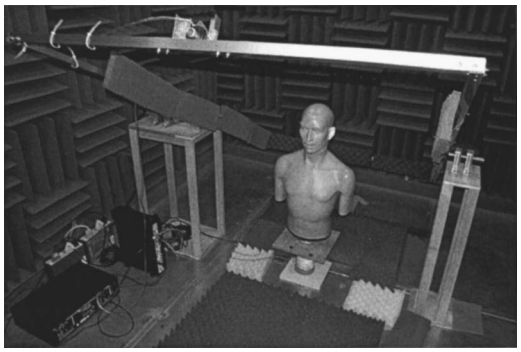


Fig. 2 The photo of experimental arrangement for the HRTF measurement

metrical arrangement of the boundary points and the virtual source points are significantly different, or when the number of the boundary points and the virtual source points are large. Under these circumstances, a straightforward inversion problem is considered impractical and special regularization is required. In the paper, the Tikhonov regularization is exploited to solve the inverse problem [15]. For a matrix inverse problem $\mathbf{Ax}=\mathbf{b}$, the solution can be found by solving the following optimization problem

$$\min_x\{\|\mathbf{Ax}-\mathbf{b}\|_2^2+\beta^2\|\mathbf{x}\|_2^2\} \tag{9}$$

The first term $\|\mathbf{Ax}-\mathbf{b}\|_2$ in the bracket denotes the residual error in 2-norm, whereas the second term $\|\mathbf{x}\|_2$ in the bracket denotes the size of solution in 2-norm. The parameter β is used to control the degree of regularization such that a reasonable compromise between the error norm and the solution norm can be reached. It can be shown that the solution obtained using the Tikhonov regularization procedure can be written in terms of the singular value decomposion (SVD) of $\mathbf{A}$ as

$$\mathbf{x}=\sum_{i=1}^{n}\frac{\sigma_i^2}{\sigma_i^2+\beta^2}\frac{\mathbf{u}_i^H\mathbf{b}}{\sigma_i}\mathbf{v}_i \tag{10}$$

where $\mathbf{u}_i$ and $\mathbf{v}_i$ are the left and right singular vectors, and σ_i is the singular value. As is common practice, the regularization parameter β is chosen to lie between the highest and smallest singular value. In the paper, the exact value of β is selected according to the L-curve criterion [15]. The L-curve is drawn by plotting the solution norm versus the error norm in log-log scale for all regularization parameters, as shown in Fig. 1. The horizontal leg of the curve is characterized by solutions that have been overly smoothed, whereas the vertical leg is characterized by solutions dominated by the fitting errors. The regularization parameter corresponding to the corner position represents a balance between the two and is thus taken to be the value for regularization.

3 Numerical and Experimental Investigations

In the section, numerical simulations and experimental investigations are presented to examine the virtual source method. For comparison purposes, the HRTFs obtained using the BEM and direct measurement are also adopted. Indices including frequency responses, impulse responses, and directivity patterns derived from the HRTFs are compared. The measurement of HRTFs is carried out using the Knowles Electronics Mannequin for Acoustic Research (KEMAR) mannequin fitted with pinnas (DB65 and DB60) and a 0.5 in. microphone inside the head. The experimental arrangement is shown in Fig. 2. The loudspeaker is mounted on a boom placed 1.2 m from the KEMAR. A program was developed for data acquisition via a B&K 3109 pulse multianalyzer and a two-axis motor control via a Pentium-4 computer. The measured results include 36 equally spaced directions on the horizontal plane and 7 directions on a plane at 60 azimuth degrees.

HRTFs are numerically modeled by using the virtual source representation and the IBEM, respectively. The SYSNOISE was employed for IBEM computation which refers to a double layer potential. The mesh of the KEMAR with pinnas required in the modeling is constructed with the aid of a scanning laser. As already mentioned, the scattering problem of HRTF computation has been converted to a radiation problem according to the acoustic reciprocity. In this setting, the source is specified as a small patch located in the ear, vibrating with constant acceleration. Thus, the source behaves as a loudspeaker operating at the mass controlled region [18]. Apart from this vibrating area, the boundary condition is set to be rigid elsewhere. The mesh used in SYSNOISE is shown in Figs. 3(*a*) and 3(*b*). In total, 10,776 mesh nodes and 21,529 elements are used in the computation.

In the virtual source method, only 2074 nodes were used for constructing the mesh of the head and pinnas. Figure 4 shows the mesh used in the virtual source method. The virtual source method was coded by using the MATLAB [19]. The distance between the field points and sound source are chosen to be 1.2 m, which is the same as the foregoing HRTF measurement. The constant acceleration source is located at the same position as that in IBEM. Therefore, in the virtual source method, the boundary condition of the mesh surface including the vibrating source and the rigid area is formulated as the following matrix equation

$$\begin{bmatrix}\frac{\partial p(\mathbf{x}_1)}{\partial n}\\ \frac{\partial p(\mathbf{x}_2)}{\partial n}\\ \vdots\\ \frac{\partial p(\mathbf{x}_N)}{\partial n}\end{bmatrix}=\begin{bmatrix}-j\rho_0\omega\frac{a}{j\omega}\\ 0\\ \vdots\\ 0\end{bmatrix}=\begin{bmatrix}-\rho_0 a\\ 0\\ \vdots\\ 0\end{bmatrix} \tag{11}$$

As indicated in Fig. 4, the distributions of virtual sources are the "shrunk" version of the head mesh. The reduction ratio is set to be 0.1 of the external mesh with reference to the previous work [11,12]. Tikhonv regularization is applied to the HRTF computations. In Fig. 5(*a*), an L-curves under the present mesh settings are calculated for the frequency 100 Hz. This curve deviates from generic L-shape considerably, and thus no corner point can be seen to locate a meaningful regularization parameter β. To overcome this problem, we introduced a cost function defined as

$$\gamma=\text{solution norm}+\text{error norm} \tag{12}$$

Figure 5(*b*) shows the plot of the cost function γ versus a progressing regularization parameter. An adequate regularization parameter is found at the corner point that can be clearly identified in this modified L-curve. The maximum frequency that the IBEM applies is approximately 10 kHz according to the six elements per

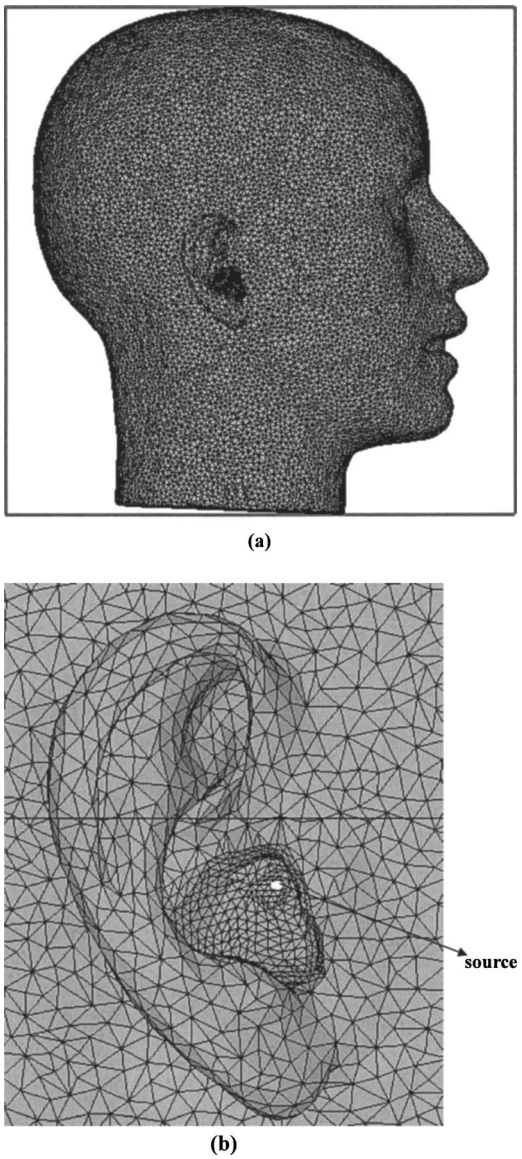


Fig. 3 The mesh of the KEMAR dummy head for the IBEM: (*a*) head mesh; (*b*) pinna mesh. The vibrating source in the ear is indicated.

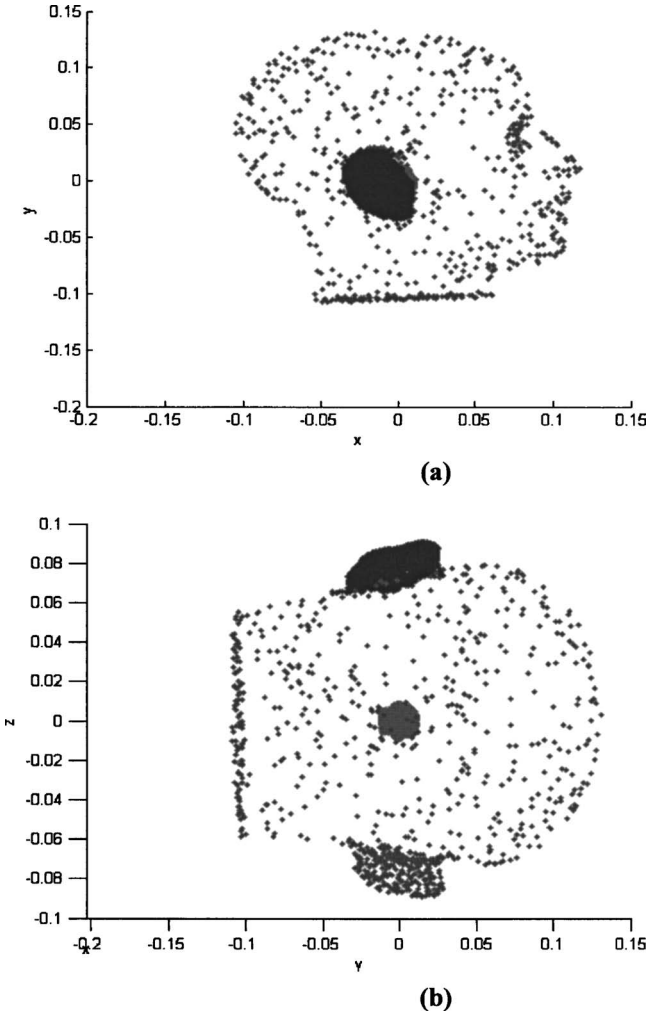


Fig. 4 The distribution of the surface nodes and virtual sources of the KEMAR: (*a*) side view; (*b*) front view

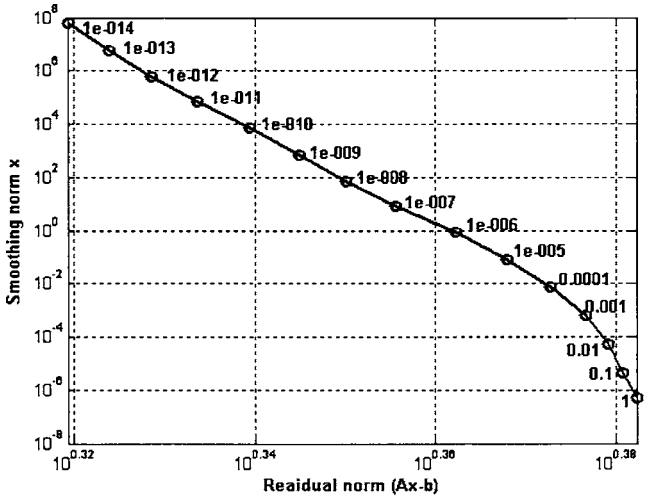
wavelength criterion recommended by the SYSNOISE.

3.1 The Objective Comparison. The frequency response magnitude of HRTF at three different directions on horizontal plane calculated by using the IBEM and the virtual source method are compared to the measured HRTF in Figs. 6 and 7, respectively. The IBEM result agrees reasonably well with the measured data. In Fig. 6(*b*), a few fluctuations due to nonuniqueness of solutions can be seen in the IBEM result below 4 kHz. In Fig. 7, the result obtained using the virtual source method follows the general trend of the measured data with a slightly larger error than the IBEM result.

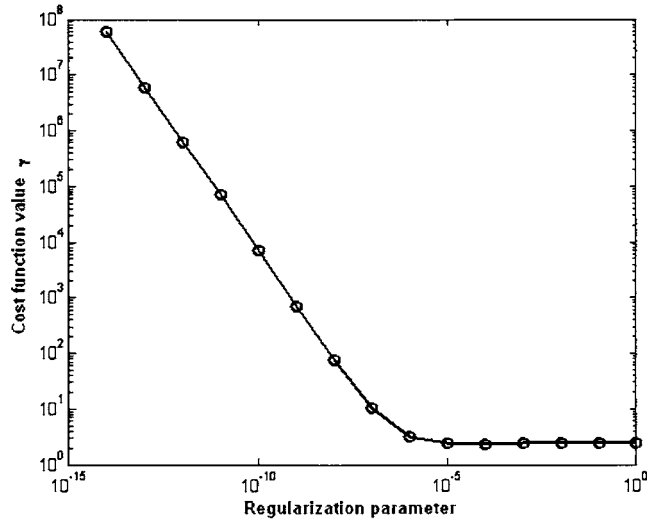
Directional responses of the calculated HRTFs and the measured HRTFs at four different frequencies are compared in Fig. 8 for IBEM, and Fig. 9 for the virtual source method, respectively. The vertical polar coordinate system shown in Fig. 8 is employed. The azimuth and the elevation angles are both 0 deg in the front of head. The directional responses are nearly omnidirectional at low frequencies and increasingly directional at high frequencies. The head shadowing effect is also more pronounced as frequency increases. Main lobes are clearly visible on the ipsilateral sides at high frequencies. Larger numerical errors are found in the directional responses obtained using the virtual source method, where the unexpected leakage on the contralateral side occurs due to the very coarse mesh used in the method.

Next, the HRIRs are calculated from the HRTFs of IBEM using the inverse fast Fourier transform. Figure 10 shows the directional responses of HRTFs on the horizontal plane calculated using the virtual source method. HRIRs at the ipsilateral (90 deg) and contralateral (270 deg) sides are shown in Figs. 11 and 12. The Interaural Time Difference (ITD) estimated from the result is 1 ms, which is close to the prediction for the given head diameter of the KEMAR (30 cm). The amplitude on the contralateral side is also smaller than that of the ipsilateral side, as expected. The ITD of calculated HRIRs by using virtual source representation are different from the ones calculated by using IBEM. The ITD calculated by using the virtual source method seems to be smaller than the IBEM result. This phenomenon could be attributed to the fact that the overly reduced virtual source surface (10% of the actual head) may decrease the effective diameter of the head. Nevertheless, the amplitude on the contralateral side is also smaller than that of the ipsilateral side, as expected.

Table 1 compares the computation loadings of two numerical methods. Computation time, the number of mesh nodes and memory usage are compared in Table 1. On the personal computer (PC), a Pentium-4 3 GHz equipped with 1 GB DDR RAM, about 35 h computation time are required to finish the 52 frequency



(a)

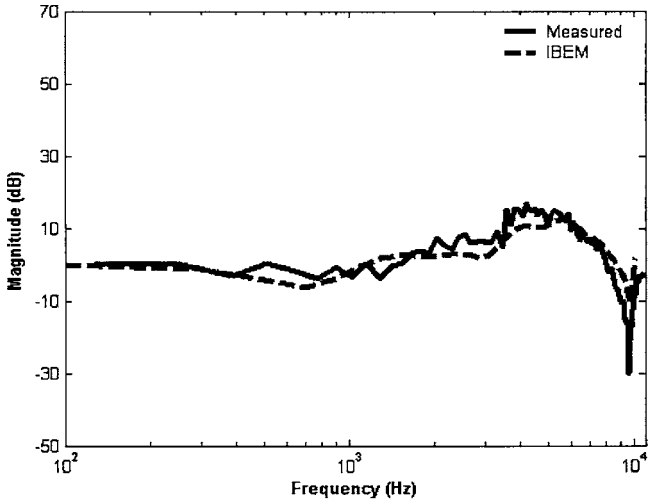


(b)

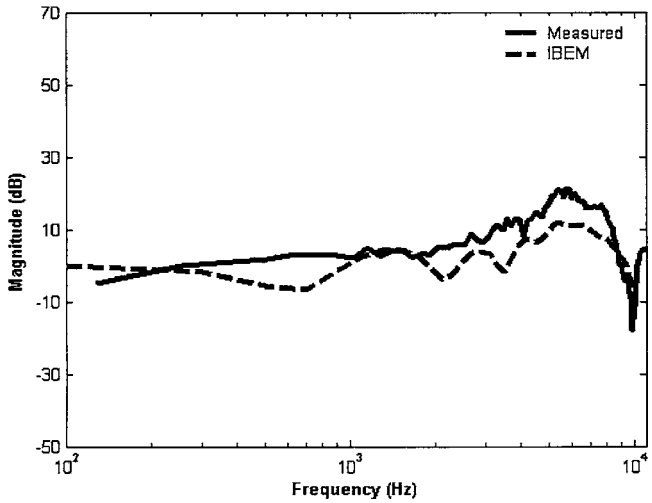
Fig. 5 L-curves of HRTF at 100 Hz, obtained using the virtual source method. (*a*) The L-curve calculated using the original definition (the regularization parameter β multiplied by the first singular value is indicated on the curves); (*b*) the modified L-curve (the regularization parameter β is multiplied by the first singular value).

points in SYSNOISE. Not only the long computation time but also the heavy memory loading are taken in IBEM. The maximum memory loading does not exceed 400 MB in virtual source representation. The number of nodes in the IBEM is about five times greater than in virtual source representation. However, the computation time of IBEM is just about two times to the virtual source representation because of the time-consuming SVD of the matrix $\mathbf{G}_n$. The response envelopes calculated by using these two numerical methods are in close agreement with the measured data.

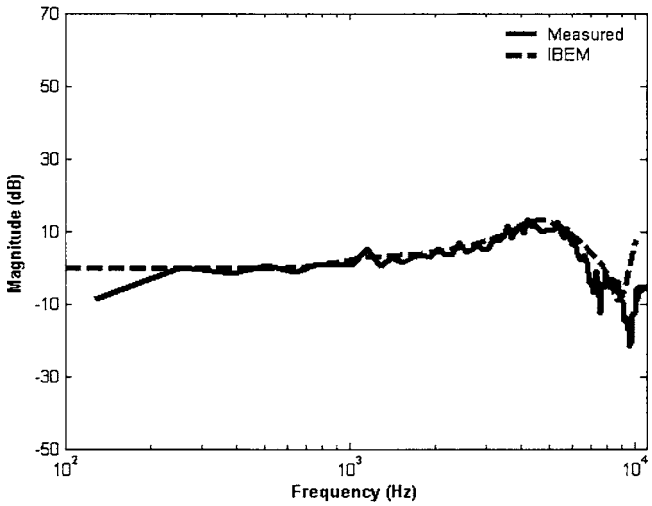
3.2 The Subjective Comparison. To further compare the HRTF models, a subjective experiment was carried out for the IBEM, the virtual source technique, and the measured HRTFs. The HRTFs are implemented as 512-tapped FIR filters [20] with the sampling rate 48 kHz. The test stimulus was a random noise. Each signal was played 5 s in duration. The signals were filtered by the intended HRTF before rendered via the ATH-PRO6 monitor headphones connected to the soundcard of the PC. Both azimuth and elevation localizations were examined. Twelve direc-



(a)



(b)



(c)

Fig. 6 HRTFs on the horizontal plane calculated using the IBEM: (*a*) Azimuth 0 deg; (*b*) azimuth 90 deg; (*c*) azimuth 270 deg

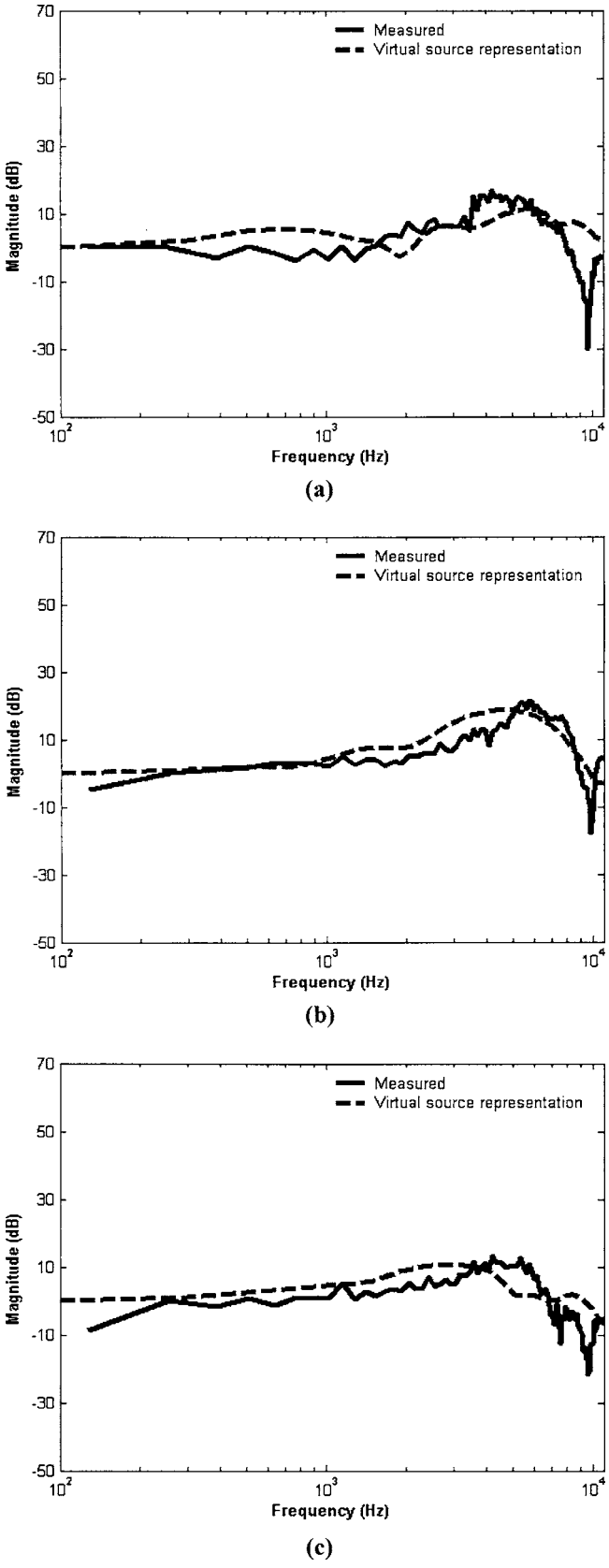


Fig. 7 HRTFs calculated using the virtual source method on the horizontal plane: (*a*) azimuth 0 deg; (*b*) azimuth 90 deg; (*c*) azimuth 270 deg

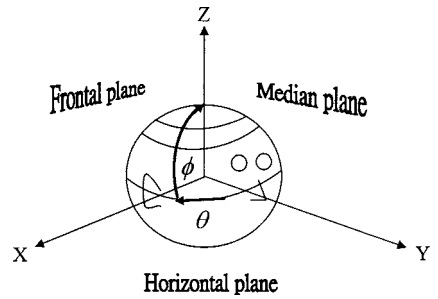


Fig. 8 The vertical polar coordinate system for the HRTF modeling

tions with a 30 deg interval on the horizontal plane were selected in the azimuth test. Similarly, six directions with a 20 deg interval from elevation −20 deg to 80 deg were selected on the plane at azimuth 60 deg in the elevation test. There are ten subjects participating in the subjective experiment.

The results of the perceived angle versus the presented angle in the subjective experiments are shown in Figs. 13 and 14. Figure 13(*a*) compares the results of azimuth localization using the measured HRTFs, the virtual source method and the IBEM. Good agreement was found in the result of Fig. 13(*a*) between the perceived and presented directions when the measured HRTF was used, albeit front-back reversals occurred to several subjects. On the other hand, despite the seemingly poor match with the measured HRTFs, the calculated HRTFs yielded comparable subjective localization performance as the measured HRTF. Some discrepancies in distinguishing the left and right using the virtual source method are due to inaccurate ITD estimation as previously mentioned. Table 2 summarizes the average and the standard deviation of localization error for the azimuth test, as calculated according to the following formulas:

$$\text{Average localization error} = \frac{1}{IJ}\sum_{j=1}^{J}\sum_{i=1}^{I}(\theta'_{ij} - \theta_{ij}) \qquad (13)$$

$$\text{Standard deviation of localization error} = \sqrt{\frac{1}{IJ}\sum_{j=1}^{J}\sum_{i=1}^{I}(\theta'_{ij} - \theta_{ij})^2} \qquad (14)$$

where θ_{ij} and θ'_{ij} are the presented and the perceived angles, respectively, corresponding to the ith angle (I angles in total) and the jth subject (J subjects in total). Although the calculated HRTFs demonstrated similar localization performance to the measured HRTFs in terms of average error, the smaller standard deviation indicated that the measured HRTFs resulted in more consistent performance than the calculated HRTFs.

Figure 14 compares the results of elevation localization using the measured HRTFs, the virtual source method, and the IBEM. Table 3 summarizes the average and the standard deviation of localization error for the elevation test. The average performance

Table 1 Comparison of computation loading between the IBEM and the virtual source method

Method	Number of Elements	Number of Nodes	Computational Time (seconds per frequency)	Memory Requirement (MB)
IBEM	21529	10776	2475.13	879
Virtual Source method		2074	1003.23	311

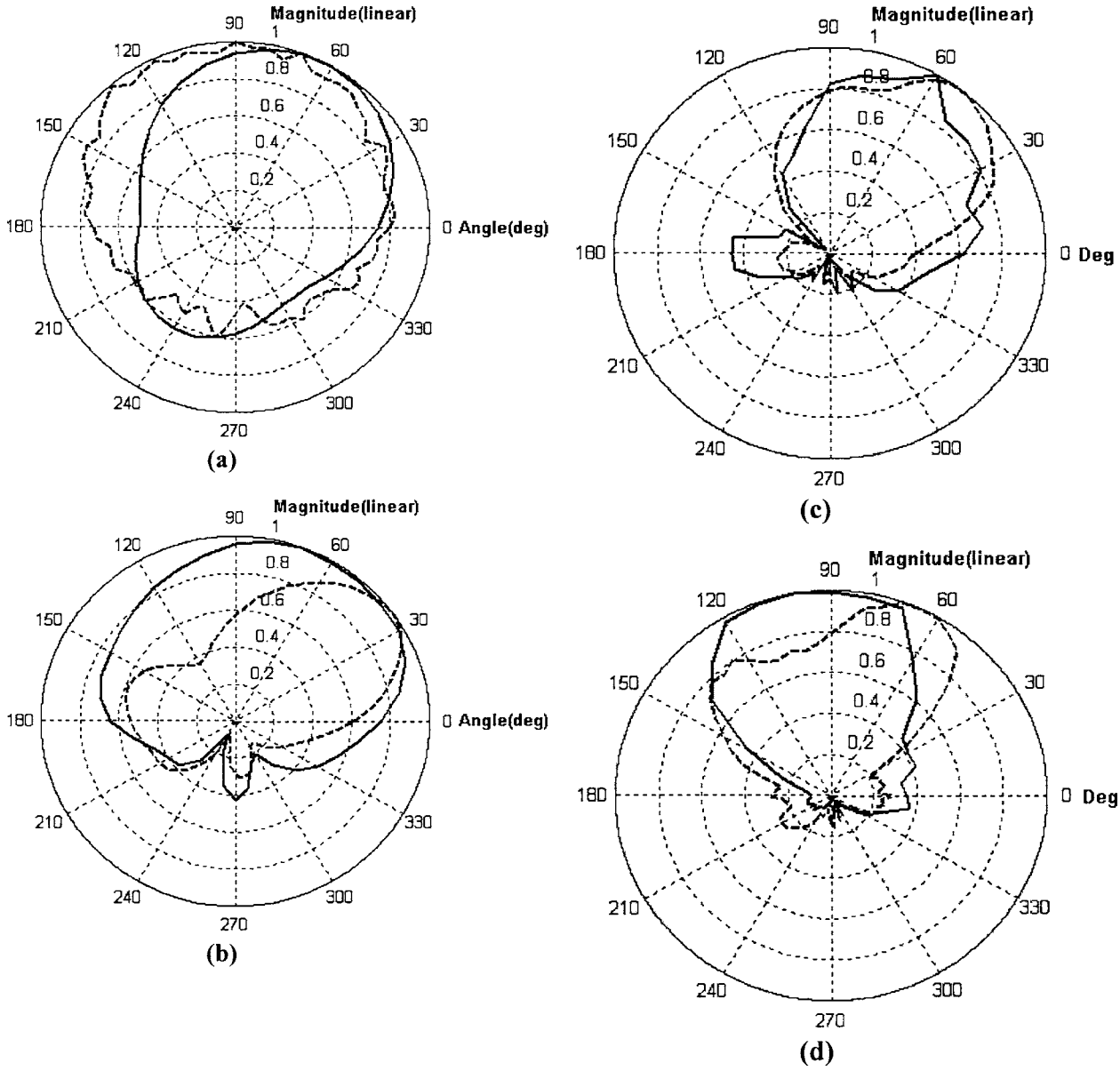


Fig. 9 The directional responses of HRTFs on the horizontal plane calculated using the IBEM. The magnitude is in linear scale and normalized with respect to the maximum. (*a*) HRTF at 500 Hz; (*b*) HRTF at 2100 Hz; (*c*) HRTF at 5100 Hz; (*d*) HRTF at 9100 Hz. (Dashed line: measured HRTF, solid line: calculated HRTF.)

of elevation localization was not as good as azimuth localization, regardless of which HRTF is used. This indicated that human localization in elevation is not as sensitive as in azimuth. In general, the localization performances of the measured HRTFs and the calculated HRTFs were comparable. The average errors of elevation localization are all around 20 deg; that is only one interval in the subjective test.

4 Conclusions

Two numerical approaches of modeling HRTFs are compared in this paper. In particular, the technique based on the virtual source representation is investigated. Using the virtual source representation, the HRTFs for a human head with pinnae are calculated with a minimal amount of computation. The feasibility of the virtual source method has been justified by extensive objective as well as subjective experiments. The numerical approaches are also compared to the direct measurement. Using numerical modeling, cost and time-consuming field measurement can be waived and a scalable prototype can be rapidly obtained with sufficiently fine mesh. This is not to mention the potential errors due to noise, transducer dynamics, position inaccuracy, and the like, during field measurement.

From the objective comparison, both numerical methods are

Table 2 Comparison of the average and standard deviation of localization errors in the azimuth test

Case	Measured HRTF database	IBEM	Virtual Source method
Average Error	23°	37°	31°
Standard Deviation	14°	34°	22°

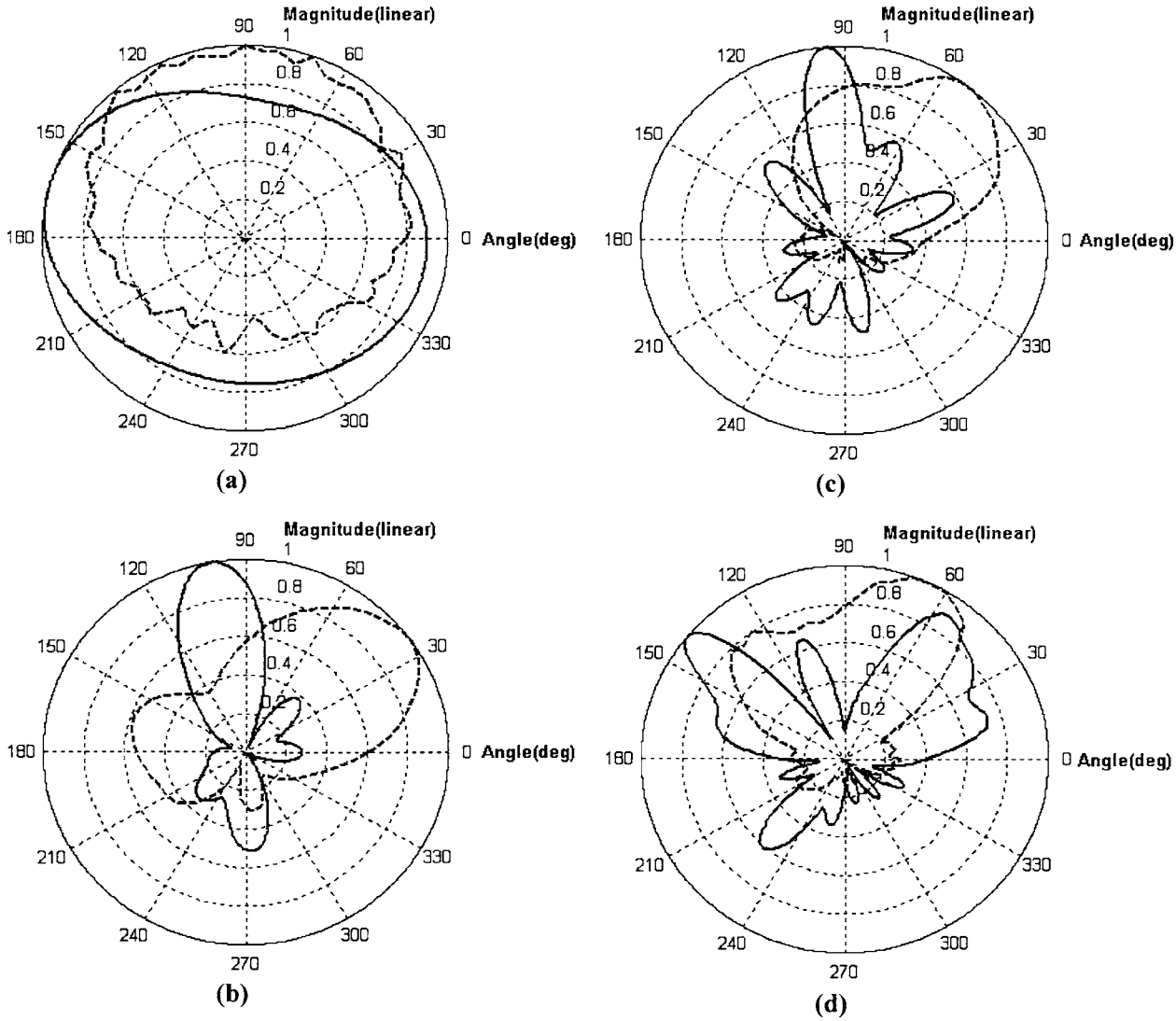


Fig. 10 The directional responses of HRTFs on the horizontal plane calculated using the virtual source method. The magnitude is in linear scale and normalized with respect to the maximum. (*a*) HRTF at 500 Hz; (*b*) HRTF at 2100 Hz; (*c*) HRTF at 5100 Hz; (*d*) HRTF at 9100 Hz. (Dashed line: measured HRTF, solid line: calculated HRTF.)

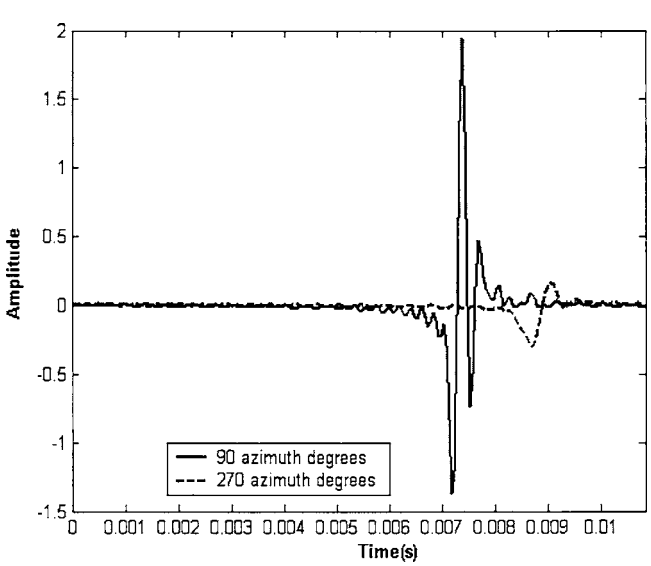


Fig. 11 Head-related impulse response on the horizontal plane obtained from the HRTF calculated using the IBEM

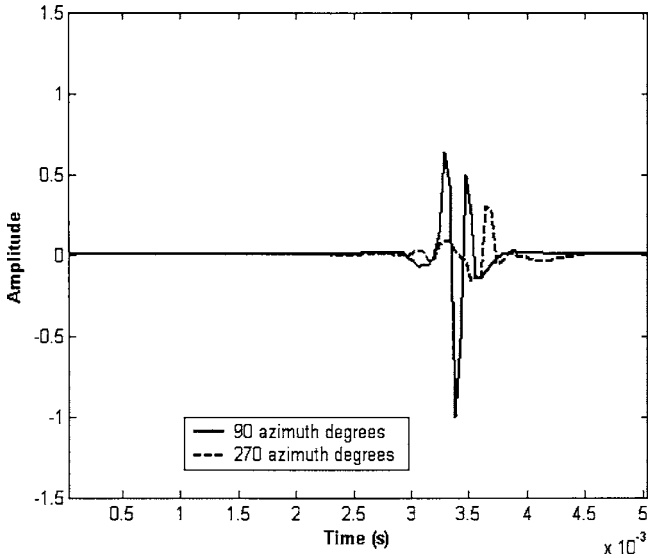


Fig. 12 Head-related impulse response on the horizontal plane obtained from the HRTF calculated using the virtual source method

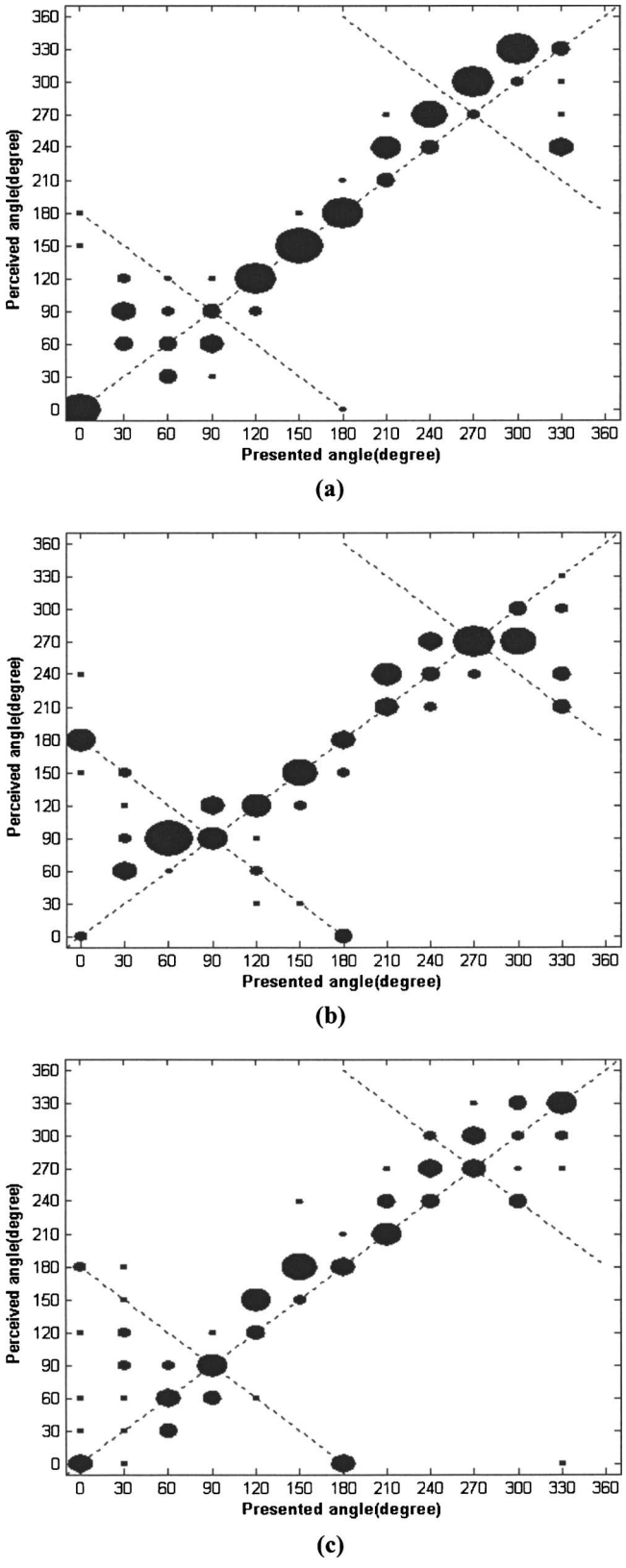


Fig. 13 Subjective experiment of azimuth localization using white noise input: (*a*) Measured HRTFs; (*b*) HRTFs calculated using the virtual source method; (*c*) HRTFs calculated using the IBEM

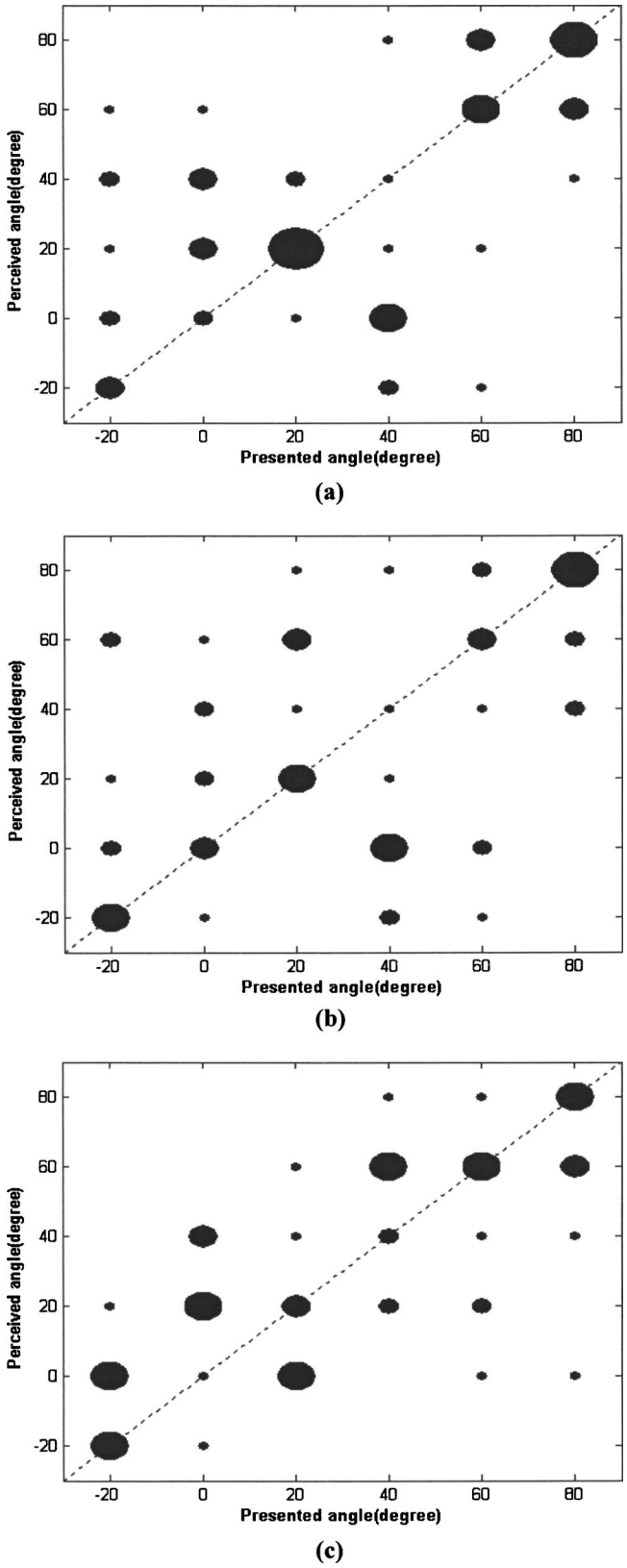


Fig. 14 Subjective experiment of elevation localization using white noise input: (*a*) Measured HRTFs; (*b*) HRTFs calculated using the virtual source method; (*c*) HRTFs calculated using the IBEM

Table 3 Comparison of the average and standard deviation of localization error of the elevation test

Case	Measured HRTF database	IBEM	Virtual Source method
Average Error	22°	25°	18°
Standard Deviation	10°	7°	3°

Table 4 Comparison of the IBEM and the virtual source method

Method / Features	IBEM	Virtual Source Method
Frequency response	Excellent	Good
Directional response	Good	Fair
Impulse response	Fair	Fair
Computation loading	High	Low
Subjective localization	Good	Excellent

capable of modeling HRTFs within a 10 kHz bandwidth. The IBEM demonstrated a slightly better numerical performance than the virtual source method in terms of frequency response, directional response, and impulse response. However, the virtual source method with appropriate regularization is capable of producing a quite impressive subjective performance of localization, as compared to the IBEM. Tikhonov regularization, in association with a modified L-curve criterion, was exploited to address the ill-conditioned inverse problem. The advantage of using the virtual source method lies in the great computational savings in memory storage and computational time. The advantages and disadvantages of the numerical methods are compared in Table 4. In conclusion, the virtual source technique offers a useful solution of "rapid-prototyping" of HRTFs. The spatial characteristics of sound localization can be reproduced as a human listener would naturally perceive by using the virtual source HRTFs.

As a limitation of the present research, the virtual source method is applied to the evaluation of HRTFs using only a coarse mesh. Indeed, the reason for the use of this coarse mesh is because there is a limit (2000×2000) to the matrix size of the SVD solver. In fact, an iterative approach that is capable of handling matrices of a larger size is being undertaken to more appropriately assess the convergence properties of the proposed technique. Further research on this aspect is currently underway.

Acknowledgment

The work was supported by the National Science Council, Taiwan, Republic of China, under Project No. NSC 92-2212-E009-030.

References

[1] Gardner, B., and Martin, K., 1995, "HRTF Measurements of a KEMAR," J. Acoust. Soc. Am., **97**, pp. 3907–3908, http://sound.media.mit.edu/KEMAR.html
[2] Algazi, V. R., Duda, R. O., Thompson, D. M., and Avendano, C., 2001, "The CIPIC HRTF Database," *Proceedings of the 2001 IEEE Workshop on Applications of Signal Processing to Audio and Electroacoustics*, Mohonk Mountain House, New Paltz, NY, pp. 99–102.
[3] Batteau, D. W., 1967, "The Role of the Pinna in Human Localization," Proc. R. Soc. London, **B168**, pp. 158–180.
[4] Batteau, D. W., 1968, *Listening With the Naked Ear: The Neuropsychology of Spatially Oriented Behavior*, Dorsey Press, Homewood, IL.
[5] Brown, C. P., and Duda, R. O., 1998, "A Structural Model for Binaural Sound Synthesis," IEEE Trans. Speech Audio Process., **6**, pp. 476–488.
[6] Algazi, V. R., Duda, R. O., Duraiswami, R., Gumerov, N. A., and Tang, Z., 2002, "Approximating the Head-Related Transfer Function Using Simple Geometric Models of the Head and Torso," J. Acoust. Soc. Am., **112**, pp. 2053–2064.
[7] Duda, R. O., 2000, "3-D Audio for HCI," Report of Spatial Audio Project, Department of Electrical Engineering, San Jose State University, http://www.engr.sjsu.edu/~knapp/HCIROD3D/3D_home.htm
[8] Kahana, Y., 2000, "Numerical Modeling of the Head-Related Transfer Function," Ph.D. thesis, University of Southampton, Institute of Sound and Vibration Research.
[9] Nelson, P. A., and Kahana, Y., 2001, "Spherical Harmonics, Singular-Value Decomposition and the Head-Related Transfer Function," J. Sound Vib., **23**(9), pp. 607–637.
[10] Walsh, T., Demkowicz, L., and Charles, R., 2004, "Boundary Element Modeling of External Human Auditory System," J. Acoust. Soc. Am., **115**, pp. 1033–1043.
[11] Koopmann, G. H., Song, L., and Fahnline, J. B., 1989, "A Method for Computing Acoustic Fields Based on the Principle of Wave Superposition," J. Acoust. Soc. Am., **86**, pp. 2433–2438.
[12] Song, L., Koopmann, G. H., and Fahnline, J. B., 1991, "Numerical Errors Associated With the Method of Superposition for Computing Acoustic Fields," J. Acoust. Soc. Am., **89**, pp. 2626–2633.
[13] Jeans, R., and Mathews, I. C., 1992, "The Wave Superposition Method as a Robust Technique for Computing Acoustic Fields," J. Acoust. Soc. Am., **92**, pp. 1156–1166.
[14] Williams, E. G., 1999, *Fourier Acoustics*, Academic Press, NY.
[15] Schuhmacher, A., Hald, J., Rasmussen, K. B., and Hansen, P. C., 2003, "Sound Source Reconstruction Using Inverse Boundary Element Calculations," J. Acoust. Soc. Am., **113**, pp. 114–127.
[16] SYSNOISE 5.6—Release Notes and Getting Started Manual, LMS International, Leuven, Belgium.
[17] SYSNOISE 5.6—User Manual, LMS International, Leuven, Belgium.
[18] D'Appolito, J., 1998, *Testing Loudspeakers*, Audio Amateur Press.
[19] MATLAB 6.5—Help Content, Mathworks, Inc.
[20] Oppenheim, A. V., Schafer, R. W., and Buck, J. R., 1999, *Discrete-Time Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ.