

Computation of the head-related transfer function via the fast multipole accelerated boundary element method and its spherical harmonic representation

Nail A. Gumerov,^{a)} Adam E. O'Donovan, Ramani Duraiswami, and Dmitry N. Zotkin
Perceptual Interfaces and Reality Laboratory, Institute for Advanced Computer Studies, University of Maryland, College Park, Maryland 20742

(Received 18 June 2009; revised 6 October 2009; accepted 7 October 2009)

The head-related transfer function (HRTF) is computed using the fast multipole accelerated boundary element method. For efficiency, the HRTF is computed using the reciprocity principle by placing a source at the ear and computing its field. Analysis is presented to modify the boundary value problem accordingly. To compute the HRTF corresponding to different ranges via a single computation, a compact and accurate representation of the HRTF, termed the *spherical spectrum*, is developed. Computations are reduced to a two stage process, the computation of the spherical spectrum and a subsequent evaluation of the HRTF. This representation allows easy interpolation and range extrapolation of HRTFs. HRTF computations are performed for the range of audible frequencies up to 20 kHz for several models including a sphere, human head models [the Neumann KU-100 (“Fritz”) and the Knowles KEMAR (“Kemar”) manikins], and head-and-torso model (the Kemar manikin). Comparisons between the different cases are provided. Comparisons with the computational data of other authors and available experimental data are conducted and show satisfactory agreement for the frequencies for which reliable experimental data are available. Results show that, given a good mesh, it is feasible to compute the HRTF over the full audible range on a regular personal computer. © 2010 Acoustical Society of America. [DOI: 10.1121/1.3257598]

PACS number(s): 43.66.Pn, 43.20.Fn [NX]

Pages: 370–386

I. INTRODUCTION

Scattering of sound off the bodies of human and animal listeners, especially their heads and external ears, provides them with cues that enable localization of the sound source.^{1,2} The scattering process can be modeled as a linear filter applied to the sound emanating from the source. The “head related transfer function (HRTF)” captures the scattering behavior of the ear, head, and body of the listener. If the head is centered at a point P and the sound source located at elevation θ , azimuth φ , and distance r in a head-centered spherical coordinate system, then the HRTF $H(r, \theta, \varphi, f)$ is the ratio of the Fourier transform of the signal at the ear $F_e(f)$ to the Fourier transform of the signal that would have been received at the point P in free-field $F_P(f)$, where f is the signal frequency. Roughly speaking, the size of the head is important above 1 kHz, the general characteristics of the torso are important below 3 kHz, and the detailed structure of the head and pinnae becomes significant above 3 kHz, with the details of the pinnae itself becoming important at frequencies over 7 kHz.

Often the HRTF is measured. Both direct and reciprocal HRTF measurements have been reported. In the former method a source (a speaker) is moved to various locations and the response measured at a microphone placed at the entrance to the blocked ear canal (see, e.g., Refs. 3 and 4). In the latter, an outwardly directed source (a microspeaker) is

placed at the entrance to the blocked ear canal, and the HRTF determined using Helmholtz’ reciprocity principle by simultaneously measuring the received signal at microphones placed at locations of interest.⁵

Since the HRTF results from a scattering process, many computational acousticians have attempted to compute it by solving the wave equation subject to boundary conditions on the surface of the head.^{6–10} Another area of interest has been the numerical simulation of the HRTF of various animals, e.g., the dolphin¹¹ and the bat.¹² The numerical HRTF is usually obtained by solving the Fourier transform of the wave equation (the Helmholtz equation) at wavenumbers corresponding to frequencies of interest, though direct simulation of the wave equation in the time domain has also been attempted.¹³ When the solution is obtained using the direct experimental configuration, the simulations must be repeated for each source position.

Numerical simulation is attractive since it offers the possibility of extracting HRTFs without subjecting the user to measurement (beyond those needed to create the discretization of their body). Further, numerical simulation, if it were easy and accurate, offers the promise of allowing one to relate features in the HRTF with the anatomical structure of the head, pinna, and body. By manipulating the mesh and observing the resulting computed HRTF it might be possible to explore the sensitivity of the HRTF to particular features. However, the match between experiment and simulation over the entire auditory range has not yet been presented, and such sensitivity analysis still remains to be done. Moreover,

^{a)}Author to whom correspondence should be addressed. Electronic mail: gumerov@umiacs.umd.edu

TABLE I. A list of recent work on numerical HRTF computation describing the numerical method used, whether a direct or reciprocal formulation was employed, the maximum frequency employed, the size of the mesh, the time taken, and whether the computation results are compared with experiments. Data in italics refer to projections rather than actual computations. Reference 9 uses a different numerical method. None of the papers referred to include the torso in the computations.

Ref.	Method	Dir/Rec	f_{\max} (kHz)	Mesh	Time	Expt.
7	BEM	Recip.	5.4	22 000	28 h/freq	Y
6	BEM	Recip.	10.3	30 570	9.15 h/freq	Y
13	FDTD-PML	Direct	8	[130] ³	0.66 h/direction	N
9	BEM	***	20	28 000	480 h	N
14	BEM	Direct	15	21 693	52.5 h	Y
This paper	FMBEM	Recip.	20	445 276	70 h	Y

computations are relatively time consuming (taking days of computer time), and simulations cannot yet be used for parametric study.

As can be seen from Table I, progress on simulations has been gradual, and the ability to completely synthesize HRTFs from meshes and simulations is not yet available. This paper presents a first such computation. Progress has been hindered by several factors. First, simulations are relatively time consuming. Since the wavelengths that must be resolved are relatively small, and approximately six to ten elements are needed per wavelength, large meshes result. When the boundary element method (BEM) is solved directly, the solution time scales as the cube of the discretization size leading to long simulation times, making numerical simulation unattractive in comparison with experiments. Further, the amount of memory needed to achieve a simulation at higher frequencies means that these simulations can only be done on supercomputers, and this has meant that most studies do not perform the simulations for high frequencies. Second, the comparison between experimentally measured HRTFs and numerically computed ones has not been demonstrated satisfactorily at all frequencies of interest. The reasons for this could lie in inaccuracies in the simulation procedure or in the quality of the mesh discretization.

For the computations to be useful, the HRTF information must be extracted and presented in a manner similar to the experiments. Further, the results of the computations and experiments must be compared to see that they agree, and any discrepancies explained. Table I summarizes the results of several papers. Due to the inability of previous authors to simulate the entire frequency range, such comparisons were attempted only for restricted ranges, and the results showed less than satisfactory agreement between experimental and the computed data.¹⁴

In this paper we use the preconditioned fast multipole accelerated boundary element method,¹⁵ a new formulation of the HRTF problem, and good meshes of the head and torso to compute the HRTF. We then compare the computations to experiments. In Sec. II, the theoretical background for HRTF computations is discussed. In particular, we derive the equations necessary to use the reciprocal set-up for computations. In Sec. III, we introduce the spherical representa-

tion of the HRTF corresponding to insonification from plane-waves from various directions. It is shown how one may then derive the HRTF corresponding to a source at a finite range from this representation. In Sec. IV we briefly describe the numerical methods used in this paper, namely, the preconditioned fast multiple method (FMM) BEM, and its use to compute the spherical harmonic spectrum. Section V presents numerical results for computations with an analytical solution (a sphere), the Neumann KU-100 “Fritz” head, and the Knowles Electronics Manikin for Acoustic Research (KEMAR) head and head-and-torso models. Section VI concludes the paper.

II. PROBLEM FORMULATION

A. Direct computations

To compute the HRTF for a given frequency $f = \omega/(2\pi)$ the three dimensional (3D) Helmholtz equation must be solved to determine the scattered field for different incident fields. Let the origin of the reference frame be placed at some point inside the head, which is referred to as the “head center” and let \mathbf{r} denote the radius-vector of any spatial point in this frame. Along with the Cartesian coordinates (x, y, z) the spherical coordinates (r, θ, φ) defined as

$$\mathbf{r} = (x, y, z) = r(\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta),$$

$$0 \leq \varphi < 2\pi, \quad 0 \leq \theta \leq \pi, \quad r = |\mathbf{r}| \quad (1)$$

are also used. Let S denote the surface of the scatterer and V the infinite external domain.

When the incident field is generated by a source, the sphere S' or radius R (range) centered at the origin, which includes surface S and the source location, must also be considered. This sphere is characterized by the vector \mathbf{r}' , or alternately by the radius R and a unit vector \mathbf{s} ,

$$\mathbf{r}' = R\mathbf{s}, \quad \mathbf{s} = (\sin \theta' \cos \varphi', \sin \theta' \sin \varphi', \cos \theta'), \quad |\mathbf{s}| = 1. \quad (2)$$

Thus, the incident complex pressure field generated by a unit intensity monopole source placed on S' is

$$\Phi^{\text{in}}(\mathbf{r}; \mathbf{s}) = \Phi_R^{\text{in}}(\mathbf{r}; \mathbf{r}') = G(|\mathbf{r} - \mathbf{r}'|) = G(|\mathbf{r} - R\mathbf{s}|), \quad (3)$$

$$G(r) = \frac{e^{ikr}}{4\pi r}, \quad k = \frac{\omega}{c}. \quad (4)$$

Here G is the free-field Green’s function for the Helmholtz equation, k is the wavenumber, and c is the speed of sound in air.

The other case considered is when the field is a plane-wave incident from direction \mathbf{s} :

$$\Phi^{\text{in}}(\mathbf{r}; \mathbf{s}) = e^{-iks \cdot \mathbf{r}}, \quad |\mathbf{s}| = 1. \quad (5)$$

This can be considered as the limit of the field generated by the monopole at large R since

$$\begin{aligned} \Phi^{\text{in}}(\mathbf{r}; \mathbf{s}) &= e^{-iks \cdot \mathbf{r}} = \lim_{R \rightarrow \infty} [4\pi R e^{-ikR} G(\mathbf{r} - R\mathbf{s})] \\ &= \lim_{R \rightarrow \infty} [4\pi R e^{-ikR} \Phi_R^{\text{in}}(\mathbf{r}; R\mathbf{s})]. \end{aligned} \quad (6)$$

The total acoustic field can be represented as

$$\Phi = \Phi^{\text{in}} + \Phi^{\text{scat}}, \quad (7)$$

where the scattered field Φ^{scat} is a regular function in V which can be found by solution of the Helmholtz equation

$$\nabla^2 \Phi^{\text{scat}} + k^2 \Phi^{\text{scat}} = 0, \quad \mathbf{r} \in V, \quad (8)$$

with the Sommerfeld radiation condition

$$\lim_{r \rightarrow \infty} \left[r \left(\frac{\partial \Phi^{\text{scat}}}{\partial r} - ik \Phi^{\text{scat}} \right) \right] = 0, \quad (9)$$

and boundary conditions on S , which for simplicity and as is customary in HRTF calculations, is assumed to be a sound-hard surface (the impedance boundary conditions could also be imposed):

$$\left. \frac{\partial \Phi^{\text{scat}}}{\partial n} \right|_S = - \left. \frac{\partial \Phi^{\text{in}}}{\partial n} \right|_S, \quad \left(\left. \frac{\partial \Phi}{\partial n} \right|_S = 0 \right). \quad (10)$$

This problem is closed and its solution is unique. However, this formulation is wasteful of computations since to compute the HRTF we need to compute Φ only at one point on the surface S , corresponding to the receiver location, whereas this procedure returns the solution everywhere on S . A more efficient formulation is presented later.

Let the location of the receiver be denoted as \mathbf{r}_* . We must determine the function

$$H(\mathbf{s}) = \begin{cases} \Phi(\mathbf{r}_*; R\mathbf{s}) \\ \Phi(\mathbf{r}_*; \mathbf{s}) \end{cases}, \quad \mathbf{r}_* \in S, \quad (11)$$

where the upper and lower formulas correspond to the monopole and plane-wave incident fields.

When the complex valued function $H(\mathbf{s})$ above is normalized by the value of the incident field at the head center, which by the choice of the reference frame is the origin, we obtain the conventional HRTF. Note that the unitary plane-wave satisfies $e^{-ik\mathbf{s}\cdot\mathbf{r}}=1$ at $\mathbf{r}=\mathbf{0}$, so $H(\mathbf{s})$ is already normalized and is the complex HRTF.

For the incident field generated by a source we have

$$\begin{aligned} \text{HRTF}(\mathbf{s}; R) &= \frac{H(\mathbf{s}; R)}{G(|R\mathbf{s}|)} = \frac{H(\mathbf{s}; R)}{G(R)} = 4\pi R e^{-ikR} H(\mathbf{s}; R) \\ &= 4\pi R e^{-ikR} \Phi(\mathbf{r}_*; R\mathbf{s}). \end{aligned} \quad (12)$$

Due to the linearity of the scattering problem and Eq. (6),

$$\text{HRTF}(\mathbf{s}; \infty) = \text{HRTF}^{(\text{pw})}(\mathbf{s}; \infty) = H^{(\text{pw})}(\mathbf{s}). \quad (13)$$

The superscript (pw) is used to show that the quantity is related to plane-wave incidence.

B. Modified boundary value problem for the reciprocal case

Even though direct solution can provide the HRTF, it is obvious that HRTF computations will be excessively long, as a separate solution is needed for each direction. Further, the only part of the solution retained is the acoustic pressure at one receiver point, with the rest discarded! To avoid this, the reciprocity principle¹⁶ can be used (see Refs. 6 and 7). This

principle states that the positions of the source and receiver in a given domain can be exchanged, while not affecting the value of the potential at the receiver location. Hence, in our case the source is placed on the domain boundary. However, the singularity of the source can cause numerical and conceptual problems unless it is explicitly treated, as here. In this section we will derive a technique that treats these problems by separating the leading order singular behavior of the scattered field by using an analytical expression for it (the image source), and then obtaining a better-conditioned boundary value problem for the remaining part of the field.

To derive appropriate forms of the boundary value problem for this situation, we place the source at the point

$$\mathbf{r}_{*\epsilon} = \mathbf{r}_* + \epsilon \mathbf{n}, \quad (14)$$

where \mathbf{n} is the normal to the boundary directed inside V and ϵ is assumed small enough so that due to the continuity of the pressure, the values obtained for $\Phi(\mathbf{r}_*; \mathbf{s})$ and $\Phi(\mathbf{r}_{*\epsilon}; \mathbf{s})$ are close. We will then consider the limiting case $\epsilon \rightarrow 0$ which will provide the result for the source placed exactly on the boundary *independent* of ϵ .

1. Incident field generated by a source

Let $\phi(\mathbf{r}; \mathbf{r}_{*\epsilon})$ be a solution of the following problem:

$$\phi = \phi^{\text{in}} + \phi^{\text{scat}}, \quad \phi^{\text{in}} = G(|\mathbf{r} - \mathbf{r}_{*\epsilon}|),$$

$$\nabla^2 \phi^{\text{scat}} + k^2 \phi^{\text{scat}} = 0, \quad \mathbf{r} \in V,$$

$$\lim_{r \rightarrow \infty} \left[r \left(\frac{\partial \phi^{\text{scat}}}{\partial r} - ik \phi^{\text{scat}} \right) \right] = 0,$$

$$\left. \frac{\partial \phi^{\text{scat}}}{\partial n} \right|_S = - \left. \frac{\partial \phi^{\text{in}}}{\partial n} \right|_S, \quad \left(\left. \frac{\partial \phi}{\partial n} \right|_S = 0 \right). \quad (15)$$

Consider the case when the incident field in the direct problem formulation is generated by a source located at \mathbf{r}' , the solution to which we denote as $\Phi(\mathbf{r}; \mathbf{r}')$. Due to the reciprocity principle, we have

$$\phi(\mathbf{r}'; \mathbf{r}_{*\epsilon}) = \Phi(\mathbf{r}_{*\epsilon}; \mathbf{r}'). \quad (16)$$

Therefore, $H(\mathbf{s})$ (11) can be computed as

$$H(\mathbf{s}) = \lim_{\epsilon \rightarrow 0} \Phi(\mathbf{r}_{*\epsilon}; \mathbf{r}') = \lim_{\epsilon \rightarrow 0} \phi(\mathbf{r}'; \mathbf{r}_{*\epsilon}) = \lim_{\epsilon \rightarrow 0} \phi(R\mathbf{s}; \mathbf{r}_{*\epsilon}). \quad (17)$$

Decomposing the total field into incident and scattered parts (15) this can be rewritten as

$$H(\mathbf{s}) = G(|\mathbf{r}_* + R\mathbf{s}|) + \lim_{\epsilon \rightarrow 0} \phi^{\text{scat}}(R\mathbf{s}; \mathbf{r}_{*\epsilon}). \quad (18)$$

Since $\mathbf{r}' \notin S$, and this point is far from the boundary, the function $\phi^{\text{scat}}(\mathbf{r}'; \mathbf{r}_{*\epsilon})$ is regular and well behaved. The limit in Eq. (18) exists and is finite. However, in numerical solution of the problem the boundary value $\phi^{\text{scat}}(\mathbf{r}_*; \mathbf{r}_{*\epsilon}) \sim \epsilon^{-1}$, which causes substantial numerical errors for small ϵ . (In

fact, previous authors have attempted to deal with the locally high gradients by modifying the mesh to account for this, see, e.g., Ref. 6.)

A straightforward use of Green's identity for determining $\phi^{\text{scat}}(\mathbf{r}'; \mathbf{r}_* \epsilon)$ is also problematic. To treat the singular behavior of the scattered field, we will use the fact that for a sound hard flat boundary the scattered field for a source located at $\mathbf{r}_* \epsilon$ can be found by placing the source of the same intensity (*image*) at $\mathbf{r}_*^- \epsilon = \mathbf{r}_* - \epsilon \mathbf{n}$. So when we have a smooth curved boundary we will represent ϕ^{scat} as a sum of the leading image term and an additional regular radiating function ψ , which satisfies the Helmholtz equation in V ,

$$\phi^{\text{scat}} = G(|\mathbf{r} - \mathbf{r}_*^- \epsilon|) + \psi, \quad (19)$$

and the following boundary condition:

$$\begin{aligned} \left. \frac{\partial \psi}{\partial n} \right|_S &= \left[-\frac{\partial}{\partial n} G(|\mathbf{r} - \mathbf{r}_*^- \epsilon|) - \frac{\partial \phi^{\text{in}}}{\partial n} \right] \Big|_S \\ &= -\left. \frac{\partial}{\partial n} [G(|\mathbf{r} - \mathbf{r}_*^- \epsilon|) + G(|\mathbf{r} - \mathbf{r}_* \epsilon|)] \right|_S. \end{aligned} \quad (20)$$

This formula allows us to take the limit $\epsilon \rightarrow 0$, which shows that

$$\left. \frac{\partial \psi}{\partial n} \right|_S = \begin{cases} 0, & \mathbf{r} = \mathbf{r}_* \\ -2 \frac{\partial}{\partial n} G(|\mathbf{r} - \mathbf{r}_*|), & \mathbf{r} \neq \mathbf{r}_*. \end{cases} \quad (21)$$

The required function $H(\mathbf{s})$ can be found from Eqs. (18) and (19) as

$$H(\mathbf{s}) = 2G(|\mathbf{r}_* - R\mathbf{s}|) + \psi(R\mathbf{s}; \mathbf{r}_*). \quad (22)$$

2. Use of Green's identity

Since ψ is a regular function in V , Green's identity can be applied to determine ψ at any field point including \mathbf{r}' via its boundary values, which can be computed using the BEM:

$$-\psi(\mathbf{r}') = L[q] - M[\psi], \quad q = \left. \frac{\partial \psi}{\partial n} \right|_S, \quad (23)$$

where L and M denote the boundary operators

$$\begin{aligned} L[q] &= \int_S q(\mathbf{r}) G(|\mathbf{r} - \mathbf{r}'|) dS(\mathbf{r}), \\ M[\psi] &= \int_S \psi(\mathbf{r}) \frac{\partial G(|\mathbf{r} - \mathbf{r}'|)}{\partial n(\mathbf{r})} dS(\mathbf{r}), \end{aligned} \quad (24)$$

where the normal orientation is consistent with Eq. (14).

Use of Green's identity allows us to link the function $H(\mathbf{s})$ (22) with the boundary values of ψ :

$$\begin{aligned} H(\mathbf{s}) &= 2G(|\mathbf{r}_* - R\mathbf{s}|) + \int_S \left[\psi(\mathbf{r}; \mathbf{r}_*) \frac{\partial G(|\mathbf{r} - R\mathbf{s}|)}{\partial n(\mathbf{r})} \right. \\ &\quad \left. - q(\mathbf{r}; \mathbf{r}_*) G(|\mathbf{r} - R\mathbf{s}|) \right] dS(\mathbf{r}). \end{aligned} \quad (25)$$

3. Plane-wave incident field

For plane-wave incidence (5) $H(\mathbf{s})$ can be found via the solution of the source on the boundary as well. Consider the solution of the problem where instead of a unit intensity source we have a source of intensity $4\pi R e^{-ikR}$. Due to linearity $H(\mathbf{s})$ can be computed according to Eq. (25) as

$$\begin{aligned} H(\mathbf{s}) &= 4\pi R e^{-ikR} \left[2G(|\mathbf{r}_* - R\mathbf{s}|) \right. \\ &\quad \left. + \int_S \left[\psi(\mathbf{r}; \mathbf{r}_*) \frac{\partial G(|\mathbf{r} - R\mathbf{s}|)}{\partial n(\mathbf{r})} \right. \right. \\ &\quad \left. \left. - q(\mathbf{r}; \mathbf{r}_*) G(|\mathbf{r} - R\mathbf{s}|) \right] dS(\mathbf{r}) \right]. \end{aligned} \quad (26)$$

Consider now the limit $R \rightarrow \infty$. Equation (6) shows that in this case the incident field for the direct problem will be $e^{iks \cdot \mathbf{r}}$. The same formula applied to Eq. (26) results in

$$\begin{aligned} H^{(\text{pw})}(\mathbf{s}) &= 2e^{-iks \cdot \mathbf{r}_*} - \int_S e^{-iks \cdot \mathbf{r}} [ik\mathbf{n}(\mathbf{r}) \cdot \mathbf{s} \psi(\mathbf{r}; \mathbf{r}_*) \\ &\quad + q(\mathbf{r}; \mathbf{r}_*)] dS(\mathbf{r}), \end{aligned} \quad (27)$$

since

$$\begin{aligned} \lim_{R \rightarrow \infty} \left[4\pi R e^{-ikR} \frac{\partial G(|\mathbf{r} - R\mathbf{s}|)}{\partial n} \right] \\ &= \mathbf{n} \cdot \nabla_{\mathbf{r}} \lim_{R \rightarrow \infty} [4\pi R e^{-ikR} G(|\mathbf{r} - R\mathbf{s}|)] \\ &= \mathbf{n} \cdot \nabla_{\mathbf{r}} e^{iks \cdot \mathbf{r}} = -ik\mathbf{n} \cdot \mathbf{s} e^{iks \cdot \mathbf{r}}. \end{aligned} \quad (28)$$

III. SPHERICAL REPRESENTATIONS

The HRTF corresponding to plane-wave incidence is a function of direction. In this section we will represent this function, $H(\mathbf{s})$, as a regular function defined on the unit sphere. Since the spherical harmonics form a basis on this surface, it is natural to represent $H(\mathbf{s})$ as

$$H(\mathbf{s}) = \sum_{n=0}^{\infty} \sum_{m=-n}^n H_n^m Y_n^m(\mathbf{s}), \quad (29)$$

where the orthonormal spherical harmonics for $\mathbf{s} = (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta)$ are defined as

$$\begin{aligned} Y_n^m(\mathbf{s}) &= (-1)^m \sqrt{\frac{2n+1}{4\pi} \frac{(n-|m|)!}{(n+|m|)!}} P_n^{|m|}(\cos \theta) e^{im\varphi}, \\ n &= 0, 1, 2, \dots, \quad m = -n, \dots, n, \end{aligned} \quad (30)$$

and $P_n^{|m|}(\mu)$ are the associated Legendre functions consistent with that in Ref. 17.

In this basis the coefficients of expansion (29) can be found as

$$H_n^m = \int_{S_u} H(\mathbf{s}) Y_n^{-m}(\mathbf{s}) dS(\mathbf{s}), \quad (31)$$

(where S_u is the unit sphere surface), using orthonormality

$$\int_{S_u} Y_{n'}^{m'}(\mathbf{s}) Y_n^{-m}(\mathbf{s}) dS(\mathbf{s}) = \delta_{nm'} \delta_{nn'}, \quad (32)$$

where δ is the Kronecker symbol.

A. Plane-wave incident field

Substituting Eq. (27) in Eq. (31)

$$H_n^{(pw)m} = \int_{S_u} \left[2e^{-i\mathbf{k}\cdot\mathbf{r}_*} - \int_S e^{i\mathbf{k}\cdot\mathbf{r}} [i\mathbf{k}\mathbf{n}(\mathbf{r}) \cdot \mathbf{s}\psi(\mathbf{r};\mathbf{r}_*) + q(\mathbf{r};\mathbf{r}_*)] dS(\mathbf{r}) \right] Y_n^{-m}(\mathbf{s}) dS(\mathbf{s}). \quad (33)$$

The Gegenbauer formula for plane-wave expansion and the addition theorem for spherical harmonics (e.g., see Ref. 18) provide

$$e^{-i\mathbf{k}\cdot\mathbf{r}} = 4\pi \sum_{n'=0}^{\infty} \sum_{m'=-n'}^{n'} i^{-n'} Y_{n'}^{m'}(\mathbf{s}) R_{n'}^{-m'}(\mathbf{r}),$$

$$R_{n'}^{m'}(\mathbf{r}) = j_{n'}(kr) Y_{n'}^{m'}\left(\frac{\mathbf{r}}{r}\right),$$

$$-i\mathbf{k}\mathbf{n} \cdot \mathbf{s} e^{-i\mathbf{k}\cdot\mathbf{r}} = \mathbf{n} \cdot \nabla_{\mathbf{r}} e^{-i\mathbf{k}\cdot\mathbf{r}}$$

$$= 4\pi \sum_{n'=0}^{\infty} \sum_{m'=-n'}^{n'} i^{-n'} Y_{n'}^{m'}(\mathbf{s}) [\mathbf{n} \cdot \nabla R_{n'}^{-m'}(\mathbf{r})], \quad (34)$$

where $j_n(kr)$ are the spherical Bessel functions of the first kind and $R_n^m(\mathbf{r})$ are the regular spherical wave functions. Substituting these in Eq. (33) and using orthonormality (32),

$$H_n^{(pw)m} = 4\pi i^{-n} \left\{ 2R_n^{-m}(\mathbf{r}_*) + \int_S [\psi(\mathbf{r};\mathbf{r}_*) \mathbf{n}(\mathbf{r}) \cdot \nabla R_n^{-m}(\mathbf{r}) - q(\mathbf{r};\mathbf{r}_*) R_n^{-m}(\mathbf{r})] dS(\mathbf{r}) \right\}. \quad (35)$$

B. Incident field generated by a source

A regular incident field can be represented as a sum of plane-waves, and so the coefficients of the expansion of $H(\mathbf{s})$ over the spherical harmonics basis, H_n^m , can be expressed via the corresponding coefficients for plane-wave incidence $H_n^{(pw)m}$. For example, the incident field generated by a source (3) inside the sphere S' can be represented outside it as (e.g., Ref. 18)

$$\Phi^{\text{in}}(\mathbf{r};\mathbf{s}) = G(|\mathbf{r} - R\mathbf{s}|) = ik \sum_{n=0}^{\infty} \sum_{m=-n}^n S_n^m(R\mathbf{s}) R_n^{-m}(\mathbf{r}),$$

$$S_n^m(\mathbf{r}) = h_n(kr) Y_n^m\left(\frac{\mathbf{r}}{r}\right), \quad (36)$$

where $h_n(kr)$ are the spherical Hankel functions of the first kind and $S_n^m(\mathbf{r})$ are the regular spherical wave functions.

Compared with Eq. (34), we see that for the incident field of a monopole source

$$H_n^m = \frac{ik}{4\pi} i^n h_n(kR) H_n^{(pw)m}. \quad (37)$$

This expression provides an easy way of relating HRTFs measured at finite range with those corresponding to plane-wave incidence and with those corresponding to other ranges as well.

IV. COMPUTATIONAL METHODS

The FMM accelerated BEM formulation for the Neumann problem for the Helmholtz equation in 3D (Ref. 15) is at the heart of the computational technique. Computation is done in two parts: solution and evaluation. The first step is solution of the problem for the scattered field ψ subject to the boundary conditions (21). The second step is determination of the expansion coefficients $H_n^{(pw)m}$ of the function $H^{(pw)}(\mathbf{s})$, which is done using Eq. (35). Note that the coefficients $H_n^{(pw)m}$ (whose number depends on the frequency) provide the theoretically most compact representation of $H^{(pw)}(\mathbf{s})$ within a given bandwidth. Further, the use of the spherical representation provides a fast and accurate interpolation of $H^{(pw)}(\mathbf{s})$ within a given bandwidth for any other direction.

The evaluation part can then be performed relatively quickly for an arbitrary range R . The case $R \rightarrow \infty$ is, of course, just $H_n^{(pw)m}$ itself. Computation of the coefficients $H_n^{(pw)m}$ for fields generated by monopole sources is fast, as it can be done using Eq. (37). After that the function $H(\mathbf{s})$ can be determined for any given direction using Eq. (29), and the HRTF or any other related functions can be computed.

A. BEM accelerated by the FMM

The boundary element method used in the present study is described in detail in Ref. 15, where a review of related work can also be found. This software uses the Burton–Miller¹⁹ formulation. The boundary integral equation solved is

$$-\frac{1}{2}[\psi(\mathbf{r}') + \lambda q(\mathbf{r}')] = (L + \lambda L')[q] - (M + \lambda M')[\phi], \quad \mathbf{r}' \in S, \quad (38)$$

where \mathbf{r}' is a point on the boundary, λ is a complex valued regularization parameter, the operators L and M are provided by Eq. (24), while the other operators are

$$L'[q] = \int_S q(\mathbf{r}) \frac{\partial G(|\mathbf{r} - \mathbf{r}'|)}{\partial n(\mathbf{r})} dS(\mathbf{r}),$$

$$M'[\psi] = \frac{\partial}{\partial n(\mathbf{r})} \int_S \psi(\mathbf{r}) \frac{\partial G(|\mathbf{r} - \mathbf{r}'|)}{\partial n(\mathbf{r})} dS(\mathbf{r}). \quad (39)$$

The boundary is discretized by a mesh with N_{vert} vertices and N_{el} flat triangular elements. The discretized equation is collocated either at the vertices or at the centers of elements using a constant panel approximation. The flexible preconditioned GMRES (Ref. 26) was used to solve the resulting linear system.

The FMM used in the BEM solver has the following scaling for CPU time: for problems with large kD , where D is the maximum size of the computational domain and k the wavenumber, this algorithm has theoretical time complexity $O((kD)^3)$ for $kD \gg 10^3$ and $O((kD)^{2+\alpha})$ with a relatively small α for $kD \leq 10^3$. These scalings were validated experimentally on benchmark problems for $0.0001 < kD < 500$ in Ref. 15. At high frequencies the algorithm has memory complexity $O((kD)^2 \log(kD))$. At low kD the time complexity depends on kD weakly and is determined rather by the size of the system N [complexity $O(N \log N)$] not by kD . It also requires $O(N \log N)$ memory. This is in contrast to the scaling of conventional BEM solvers which have a time complexity of $O((kD)^6)$ and a memory complexity of $O((kD)^4)$. Because of this, the HRTF calculation presented here can be done for the entire audible frequency range on a mesh that includes the torso on a conventional desktop workstation.

B. Spherical harmonic analysis

1. Truncation error and error bound

The number of spherical harmonics needed to represent the function $H(\mathbf{s})$ for any given frequency or wavenumber k can be deduced from the Gegenbauer plane-wave decomposition (34), where the infinite sum over n is truncated to a maximum value $n=p-1$. This provides a p^2 term approximation via an expansion over the spherical functions R_n^m . Using the addition theorem for spherical harmonics, we can rewrite Eq. (34) as

$$\begin{aligned} e^{-i\mathbf{k}\mathbf{s}\cdot\mathbf{r}} &= 4\pi \sum_{n=0}^{\infty} i^{-n} j_n(kr) \sum_{m=-n}^n Y_n^m(\mathbf{s}) Y_n^{-m}\left(\frac{\mathbf{r}}{r}\right) \\ &= \sum_{n=0}^{\infty} (2n+1) i^{-n} j_n(kr) P_n\left(\mathbf{s} \cdot \frac{\mathbf{r}}{r}\right) \\ &= \sum_{n=0}^{p-1} (2n+1) i^{-n} j_n(kr) P_n\left(\mathbf{s} \cdot \frac{\mathbf{r}}{r}\right) + \epsilon_p(kr), \end{aligned} \quad (40)$$

where $\epsilon_p(kr)$ is the truncation error and P_n are the Legendre polynomials, $|P_n| \leq 1$. This error for $n \geq p > kr$, for which $j_n(kr)$ are strictly positive functions, can be bounded as

$$\begin{aligned} |\epsilon_p(kr)| &= \left| \sum_{n=p}^{\infty} (2n+1) i^{-n} j_n(kr) P_n\left(\mathbf{s} \cdot \frac{\mathbf{r}}{r}\right) \right| \\ &\leq \sum_{n=p}^{\infty} (2n+1) j_n(kr) < \sum_{n=p}^{\infty} (2n+1) j_n(ka), \end{aligned} \quad (41)$$

where a is the maximum possible value of r . Since for computation of H_n^m parameter r is the distance from the scatterer center to a point located on the scatterer surface [see Eq. (33)], a is the minimal possible value of range R for which spherical decomposition can be used, so we set $a=R_{\min}$.

The error bounds shown in the literature for the sum (41) are related to decompositions of plane-wave over the wave functions (e.g., Ref. 20). A more strict bound valid for any ka for this sum is found in Ref. 21 and is

$$\begin{aligned} |\epsilon_p| &\leq \sum_{n=p}^{\infty} \frac{(ka)^n}{(2n-1)!!} < \left(\frac{ka}{2}\right) \sum_{n=p-1}^{\infty} \frac{(ka)^n}{2^n n!} \\ &= \frac{e^{ka/2}}{(p-1)!} \left(\frac{ka}{2}\right)^p, \quad p \geq 4. \end{aligned} \quad (42)$$

This substantially overestimates the p required to obtain a given accuracy $|\epsilon_p|$ for large ka for which the asymptotic bound²¹

$$\begin{aligned} |\epsilon_p| &\leq (2p+1) j_p(ka) \leq \frac{p^{2/3}}{(ka)^{1/2}} \exp\left(-\frac{1}{3}(2\eta_p^{(a)})^{3/2}\right), \\ \eta_p^{(a)} &= \frac{p-ka+1/2}{(ka)^{1/3}}, \quad p > ka \gg 1 \end{aligned} \quad (43)$$

is more appropriate. This formula results in the following estimate of p for prescribed error $\epsilon=|\epsilon_p|$:²¹

$$p \geq ka + \frac{1}{2} \left(3 \ln \frac{1}{\epsilon} + \frac{1}{2} \ln(ka) \right)^{2/3} (ka)^{1/3}, \quad ka \gg 1. \quad (44)$$

We also note a more simplistic approximation of $p(ka)$ for any ka used in the literature (e.g., Ref. 22)

$$p = ka + 4(ka)^{1/3} + 3. \quad (45)$$

The error of this approximation grows with ka [as it misses a slow growth $\sim \ln(ka)$ of the coefficient near $(ka)^{1/3}$, which actually can be estimated using Eq. (43) (which is a more accurate estimate than Eq. (44))] with p from Eq. (45).

We also note that for computation of the HRTF for finite range we have more restrictive conditions on p which increase it due to the requirement of the accuracy of source field expansion (36). A detailed error bound analysis for this case is provided in Ref. 18, where also the following approximation for p was obtained and checked for a broad range of frequencies:

$$\begin{aligned} p \approx & \left\{ \left[\frac{1}{\ln \delta} \ln \frac{\delta^{3/2}}{\epsilon(\delta-1)^{3/2}} + 1 \right]^4 + \left[ka + \frac{1}{2} \left(3 \ln \frac{1}{\epsilon} \right. \right. \right. \\ & \left. \left. \left. + \frac{1}{2} \ln(ka) \right)^{2/3} (ka)^{1/3} \right]^4 \right\}^{1/4}, \quad \delta = \frac{R}{a}. \end{aligned} \quad (46)$$

Note that this formula is modified for the HRTF, which is a normalized function (12) (in Ref. 18 the equation is derived based on the absolute error). It shows that for $\delta=1$, i.e., when the source is located on the boundary of the expansion domain, $p=\infty$, while p is finite for $\delta > 1$. For ranges $R > 2a$ the order of magnitude of the truncation error will be the same as that for the expansion of the plane-wave and more simplistic estimates can be used.

2. Computation of expansion coefficients

As soon as the required bandwidth for a given error bound, frequency, and domain size for the spherical harmonic representation is known, the p^2 coefficients $H_n^{(pw)m}$ can be computed using the direct formula (35). Note that computation of functions $R_n^m(\mathbf{r})$ and their directional derivatives $\mathbf{n}(\mathbf{r}) \cdot \nabla R_n^m(\mathbf{r})$ is done via a standard recursive procedure used

in the FMM for the Helmholtz equation, which is inexpensive with a computational complexity of $O(Np^2)$. Also the integrals should be computed in a way consistent with the procedure of computation of the BEM integrals (e.g., see treatment of singular integrals¹⁵). While computation of integrals (35) is usually much faster than solution of the linear system, for large problems these computations can be accelerated via the FMM using the fact that function $H(s)$ is a sum of dipoles and monopoles [see Eq. (25)] of given intensities, and coefficients of expansion of this function over the spherical basis can be retrieved immediately from the expansions appeared in the FMM (see Ref. 15).

V. NUMERICAL TESTS

We first conducted some tests on simple geometrical models for which the HRTF can be obtained either analytically or semi-analytically without using the BEM, reciprocity principle, and decompositions/synthesis over spherical basis functions and compare it with the present method which uses all these. These tests were used to validate the method and figure out the settings appropriate for computations with more complex geometries. After that we performed several tests on a complex head geometry in the presence and absence of the torso and compared this with the results from simpler geometries. Also, we tested the BEM solutions against analytical results for arbitrary meshes, such as by computing the field generated by a source placed inside the body. For the meshes used the relative errors in the solution for the worst cases were of order 1% which are consistent with usual BEM errors. These tests provided us with the confidence that the method can be used for modeling more complex situations, and we then compared the results of the computations with the HRTF measured in experiments.

A. Tests for head mesh

1. Mesh generation

Computation of scattering off complex shapes, such as human head, requires quality surface meshes. A quality triangular mesh satisfies the requirements that the mesh is fine enough to satisfy the requirement of at least five elements per wavelength, that all triangles be as close to equilateral as possible, that the number of triangles associated with each vertex of the mesh be as close to 6 as possible, and finally that the rate of change in triangle areas across different regions of the mesh be gradual. The use of the version of the FMM accelerated BEM described in Ref. 15 requires high quality meshes, which ideally should have a uniform distribution of the mesh vertices over the surface.

On the other hand, the data coming from laser scans of heads and meshed using standard triangulation techniques, such as those based on the Delaunay triangulation, are acceptable for graphics, but have a poor quality for the BEM. Several such meshes were available to the researchers. To make them appropriate for the BEM we developed a remeshing technique which produced meshes of the desired quality. Note that in the literature several techniques for generation of quality meshes are available [e.g., in the Computational

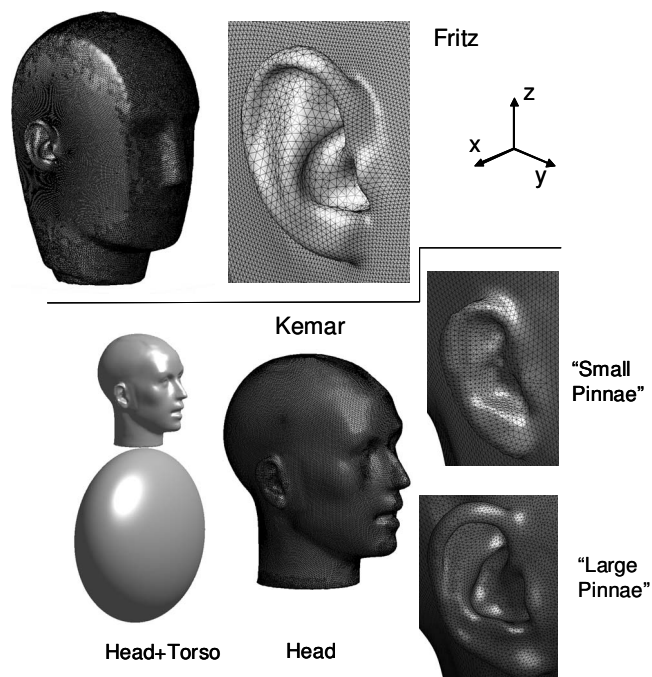


FIG. 1. The meshes used for computation of the HRTF using the FMM accelerated BEM for the Fritz and Kemar models. The Fritz head mesh contains 152 666 triangular panels, while Kemar head+torso meshes contain 426 846 and 445 276 panels for “small pinnae” and “large pinnae” cases, respectively.

Geometry Algorithms Library (CGAL) project]. A detailed description of our technique is beyond the scope of the present paper, and we only mention that the technique is based on the 3D radial basis function (RBF) interpolation combined with a relaxation technique, which allows the mapping of meshes between topologically equivalent objects. Details of the technique can be found in our technical report.²³ We note that for high frequency computations up to 20 kHz the head mesh contained of order 10^5 elements, and the maximum value of the computational parameter kD was 181. The RBF technique for such a large data set was enabled by the use of the FMM for the biharmonic equation and $O(N \log N)$ preconditioning technique, as described in Ref. 24.

The resulting mesh was checked for quality using several geometrical parameters (minimal and maximum size of the triangles, aspect ratios, etc.) and for several elements an additional subdivision procedure was applied. In this way it was guaranteed that the surface is discretized with not less than six elements per wavelength for the highest frequency computed. Figure 1 shows the computational mesh for the Neumann KU 100 Dummy Head model (referred to as Fritz). This mesh contains 152 666 triangles. We also did computations with two Kemar head meshes which contain 99 166 triangles (“small pinnae” mesh) and 117 596 triangles (“large pinnae” mesh), respectively (see Fig. 1). Note that the small pinnae Kemar mesh was generated from a laser scan, while the large pinnae mesh was obtained in the following way. A mesh of good quality was generated from the CT-scan of the pinnae provided by Duda and Algazi.²⁵ The pinna areas of the original Kemar head were cut off, and the pinnae meshes were positioned at an angle corresponding to measurements

TABLE II. Performance for the Fritz and Kemar (head with small pinnae) models.

Frequency (kHz)	0.172	0.517	1.034	2.067	4.996	9.991	14.987	20.155
	BEM solution (s)							
Fritz	131	125	159	334	558	919	1254	1483
Kemar	62	64	86	160	341	417	520	628
	Computation of $H_n^{(pw)m}$ (s)							
Fritz	1.2	1.8	2.9	5.6	13	35	64	103
Kemar	0.9	1.3	1.8	3.3	8	21	36	58
	Memory (Gbyte)							
Fritz	2.07	2.10	2.14	2.20	2.30	2.37	2.44	2.60
Kemar	3.12	3.16	3.20	3.25	3.30	3.32	3.34	3.43

in the UC Davis Center for Image Processing and Integrated Computing (CIPIC) database.³ After that a mesh “stitching” and relaxation technique was used to merge the meshes. Below, if it is not mentioned otherwise specifically, the term “Kemar head mesh” means the original model with the small pinnae.

2. Performance

In the present study we used the BEM with the panel-center collocation method, since it is more appropriate to the image source method. The algorithm was executed on a four core PC (Intel Core 2 Extreme QX6700 2.66 GHz, 8 Gbyte RAM) using OpenMP parallelization. Table II shows the wall clock time (includes input/output time for each procedure) and memory required for several cases. The total execution time for the run, which includes BEM solution, and computation of the spherical spectrum over 117 frequencies equispaced from 172 to 20 155 Hz took about 32 h for Fritz and 14 h for the Kemar head.

3. Spherical spectra

We analyzed the spherical spectra for the computed cases and found that the truncation occurs in the region of spectral (exponential) convergence, and the values of the truncated modes are several orders of magnitude smaller than the values of the modes stored for the HRTF representation. Computed spectra were compared with analytical results for a sphere and showed a good agreement if the size of the effective sphere is selected as the minimal radius of the head (the shortest distance from the head surface to the center). Particularly, the limiting value of the mean mode $H_0^{(pw)0}(0) = (4\pi)^{1/2}$ was achieved with an accuracy of four digits; good quantitative match was found for $ka < 1$ (frequencies below 440 Hz) and for frequency of oscillations of the mean mode $H_0^{(pw)0}$ as a function of ka .

4. HRTF

Figure 2 shows a comparison between the plane-wave HRTF for the Fritz model and for a sphere of radius a_{\min} . The bright side of the imaging sphere corresponds to azimuth angles $-90^\circ \leq \varphi \leq 90^\circ$ [see Eq. (1) and the reference frame in Fig. 1], while the dark side corresponds to azimuths $90^\circ \leq \varphi \leq 270^\circ$. It is seen that at low frequencies, such as at 1 kHz the HRTF for the sphere and for the Fritz head is

qualitatively similar. It is also seen that at these frequencies “the bright spot” on the dark side of the imaging sphere is well defined for both the head shapes. At higher frequencies, such as 10 kHz the HRTFs for the Fritz model and sphere are very different. On the ipsilateral side (the side of the ear) the maximum amplitude of the HRTF for the sphere is at the ear location. At the same time for the Fritz model a notch is observed at this location, where the HRTF amplitude can be 20 dB or so less than its maximum value. The pictures for the contralateral side are also different. If for the sphere we can see the bright spot surrounded by a circular interference pattern, the Fritz model demonstrates rather irregular filamentary structure of the notches, where the amplitude of the HRTF can be tens of decibels lower than the average (in fact, the bright spot and circular interference pattern are present also in the Fritz model, but it is substantially masked by the notch structure).

The frequency dependence of the plane-wave HRTF along with its dependence on the spherical angles is presented in Fig. 3. Again we can see the difference between the HRTF for the sphere and for the Fritz and Kemar head models. The dynamic range of the HRTF for Fritz and Kemar is much larger than that for the sphere. The amplitude at the notches can be tens of decibels lower than the average values. It is also seen that while the HRTF for the sphere is clearly symmetric for elevation angles $\theta = 90^\circ \pm \alpha$, the HRTF for the head models is different for the planes defined by elevation angles $\theta = 45^\circ$ and $\theta = 135^\circ$ which cut the head above and below the elevation angle at the plane of ear location, $\theta = 90^\circ$. The figure also shows a difference between the HRTFs for Fritz and Kemar head models and the effect of the pinnae shape since the difference between the two Kemar head meshes is only due to the different pinnae.

5. Comparison with the computations of other authors

Fortunately, there are computational results for the Fritz head model computed by other authors. Two groups computed the HRTF for frequency range from 100 Hz to 15 kHz using conventional BEM on substantially non-uniform meshes containing 21 693 and 12 054 panels, respectively, and compared their results.¹⁴ The results for the “gain” in decibels for two incident angles were scanned from the cited paper and renormalized to obtain true HRTFs. As the HRTF is plotted using a logarithmic scale (decibels), any renormal-

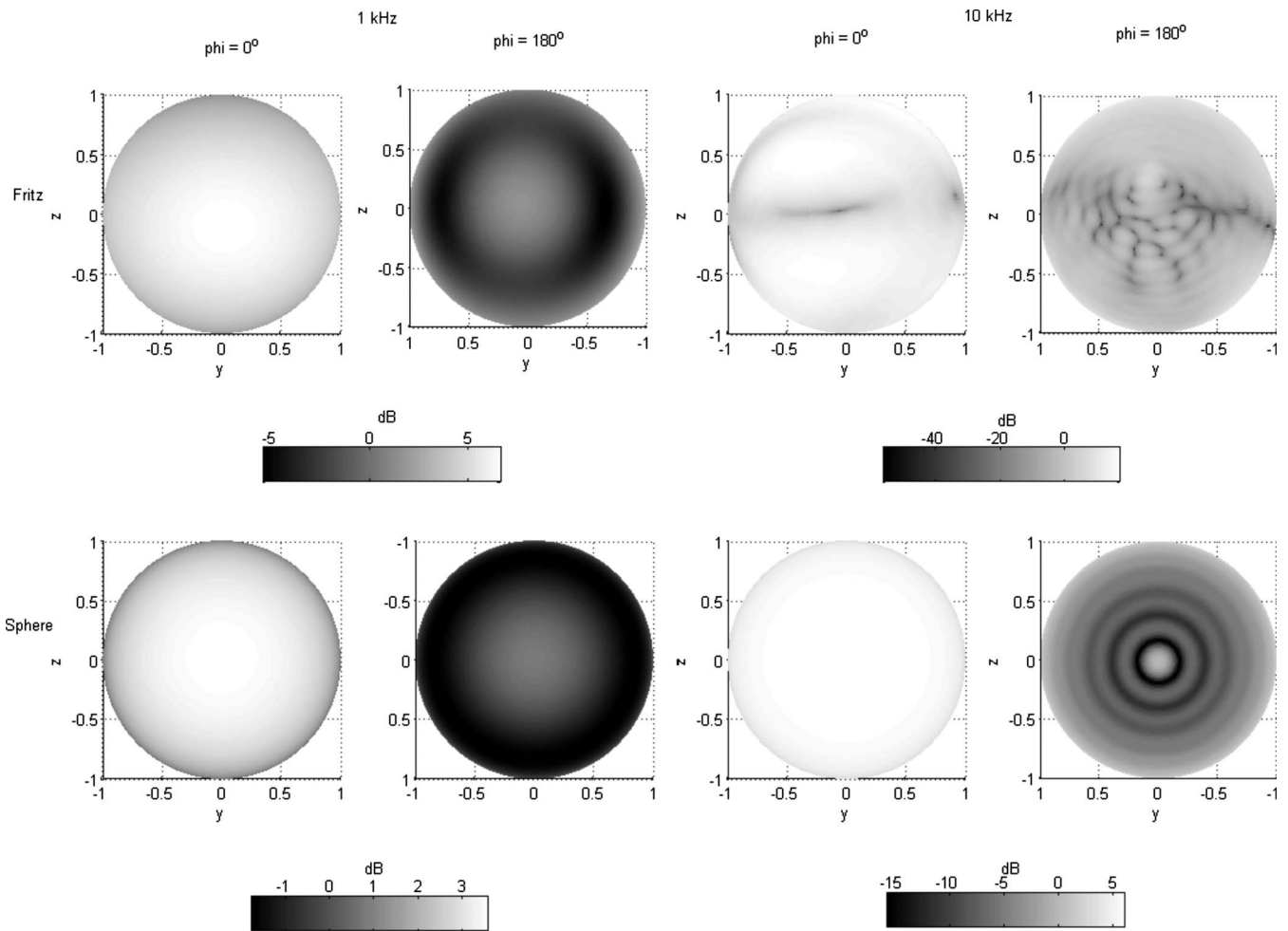


FIG. 2. The plane-wave HRTF ($20 \log_{10}|H^{(pw)}|$) computed for the Fritz model (the top row) and analytical for a sphere of radius $a = a_{\min} = 0.0705$ m (the bottom row) at frequencies 1.034 and 9.991 kHz. The HRTF is plotted as a function of spherical angles θ and φ [see Eq. (1)] on a unit sphere. The center of the right ear is located at the $\theta = 90^\circ$ and $\varphi = 0^\circ$. The views corresponding to $\varphi = 0^\circ$ are from the side of the right ear, while $\varphi = 180^\circ$ are from the side of the left ear. Axes (x, y, z) are oriented as shown in Fig. 1.

ization results in a shift of graphs along the y -axis. Since at low frequencies we have $|H^{(pw)}| \rightarrow 1$, or 0 dB, we shifted the cited results to this level at the lowest frequency (100 Hz).

Figure 4 shows comparison between the present computations and that of Ref. 14. It is seen that the computations reproduce the patterns in the HRTF. Quantitatively a good agreement is observed for the range of frequencies below 5 kHz, while for larger frequencies some mismatch takes place. There can be several reasons for this mismatch (differences in meshes and methods and accuracy, slight misalignment of the model [e.g., our measurement point was located actually at angle $\theta \approx 89^\circ$ instead of ideal $\theta = 90^\circ$]). The comparison shows that the HRTF in all cases is computed well enough, and the difference in computations is smaller than usual difference between experiments and computations.

6. Comparison with experiments

Two sets of experimental data for the Fritz model were available to the authors via their own measurements, using both a direct and a reciprocal set-up. These measurements were conducted as a part of the “Club Fritz” activity.²⁷ Comparisons with both direct and reciprocal experiments pro-

vided approximately the same quality of agreement, so only data on the direct measurements which were conducted with a denser sampling are reported below for the sake of space.

In these experiments, the impulse response function was measured at both ears using external microphones and then its spectrum was analyzed to determine the HRTF for 823 directions. The signal generation/processing technique used windowing, so the experimental data for frequencies above 14 kHz are not reliable and frequencies above 16 kHz are completely suppressed. Also, as is the case with most HRTF measurements (see, e.g., Ref. 28 for a discussion) the experimental spectra were not reliable for low frequencies below 700 Hz. So for comparisons with computations only experimental HRTF for the range of frequencies from 700 Hz to 14 kHz was used.

We compared experiments and computations for all measurement locations, and several typical comparisons are shown in Fig. 5, where all non-normalized experimental data were shifted by 12 dB to provide consistent magnitudes with computations. A match in the patterns produced by the experiments and computations is clearly seen, while a quantitative discrepancy is also observed. Such a discrepancy may arise due to several factors, including modeling issues and

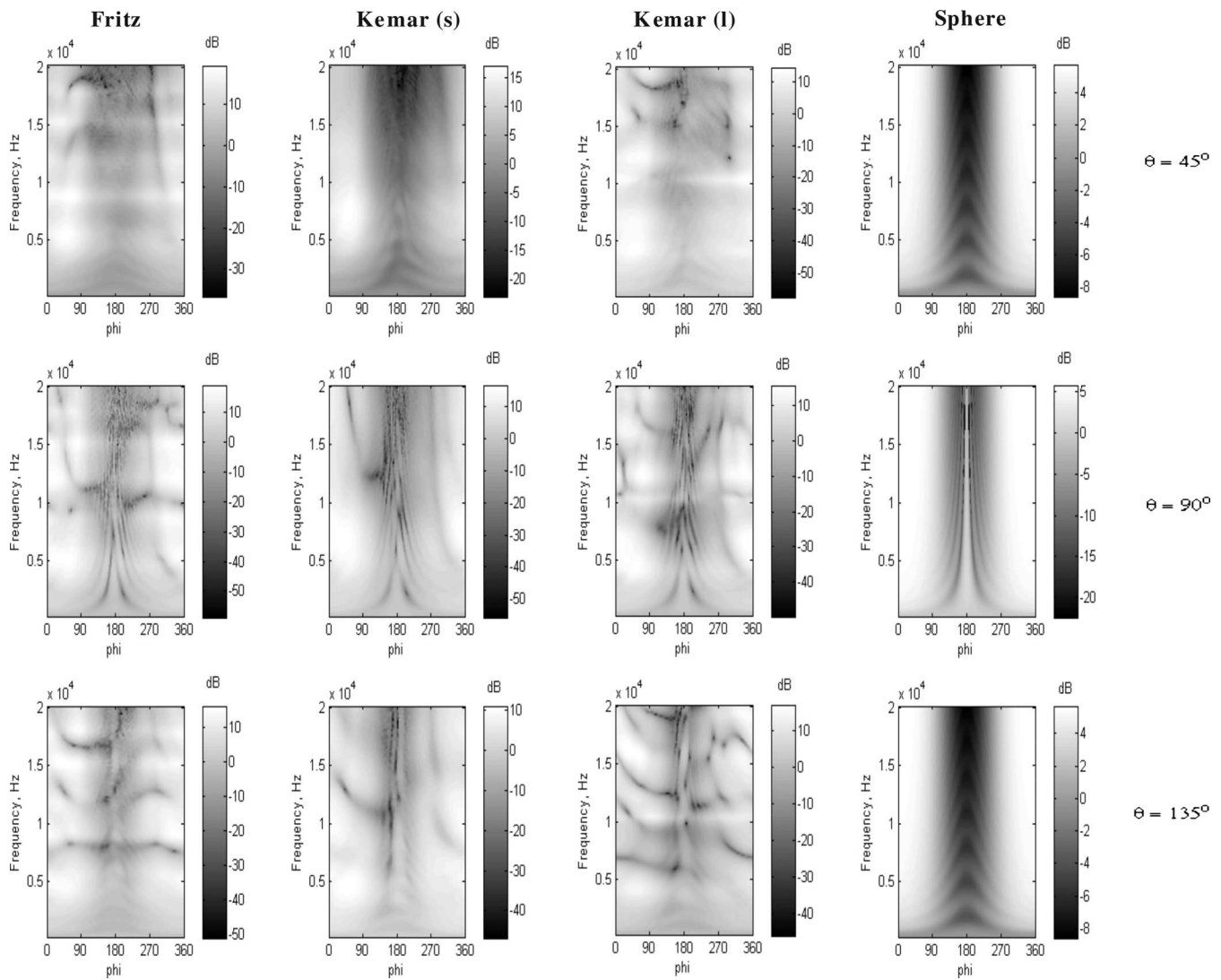


FIG. 3. The plane-wave HRTF ($20 \log_{10}|H^{(pw)}|$) for the right ear ($\theta=90^\circ$, $\varphi=0^\circ$) computed for Fritz (the left column) and Kemar heads (the left center column for the “small pinnae” model and the right center column for the “large pinnae” model) and analytical for a sphere of radius $a=a_{\min}=0.0705$ m (the right column) at fixed spherical angles θ for the range of audible frequencies and spherical angles φ [see Eq. (1) and Fig. 1 for definitions].

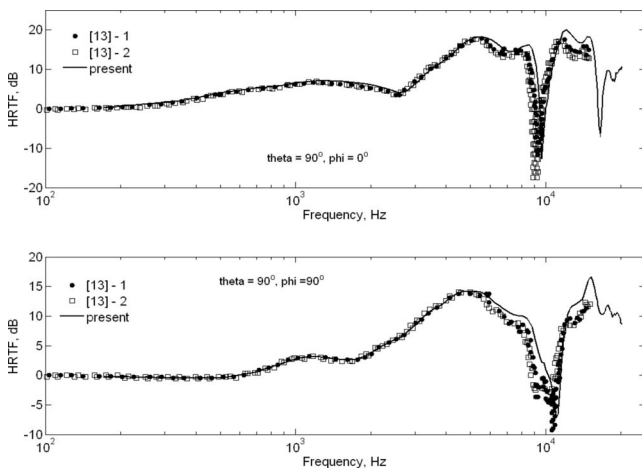


FIG. 4. Comparison of the plane-wave HRTFs ($20 \log_{10}|H^{(pw)}|$) for two incident angles computed for the Fritz model by two groups and reported in Ref. 14 and using the present method. Original normalized data of Greff and Katz are renormalized to match the HRTF asymptotics at zero frequency ($|H^{(pw)}| \rightarrow 1$).

experimental errors. The latter, for example, include imprecision in recording of the tilt in the model position, which may result in a few degrees of inconsistency in the directions for which the HRTF was measured and computed. To evaluate such a type of the error, we plotted in Fig. 5 by the dashed lines measurements for the left ear for which directions were modified to get the “right ear data from the left ear measurements.” This was easy to do since the 823 directions were symmetric with respect to the azimuth $\varphi=90^\circ$, and the same set could be used for evaluation of the response of both ears (pairs $\varphi=90^\circ \pm \alpha$ with the same θ were swapped). In the case of perfect symmetry and ideal measurements the actual data for the right ear and derived from the data for the left ear should coincide for this symmetric model. However, Fig. 5 shows discrepancy of these data, which can be used to judge the experimental errors and model misalignment. It is seen then that in some cases inconsistency of the experimental data is of the same order of magnitude as the inconsistency between the experiment and computations, while in some other cases the experimental error is smaller than between the experiments and computations (e.g., 1–2 kHz

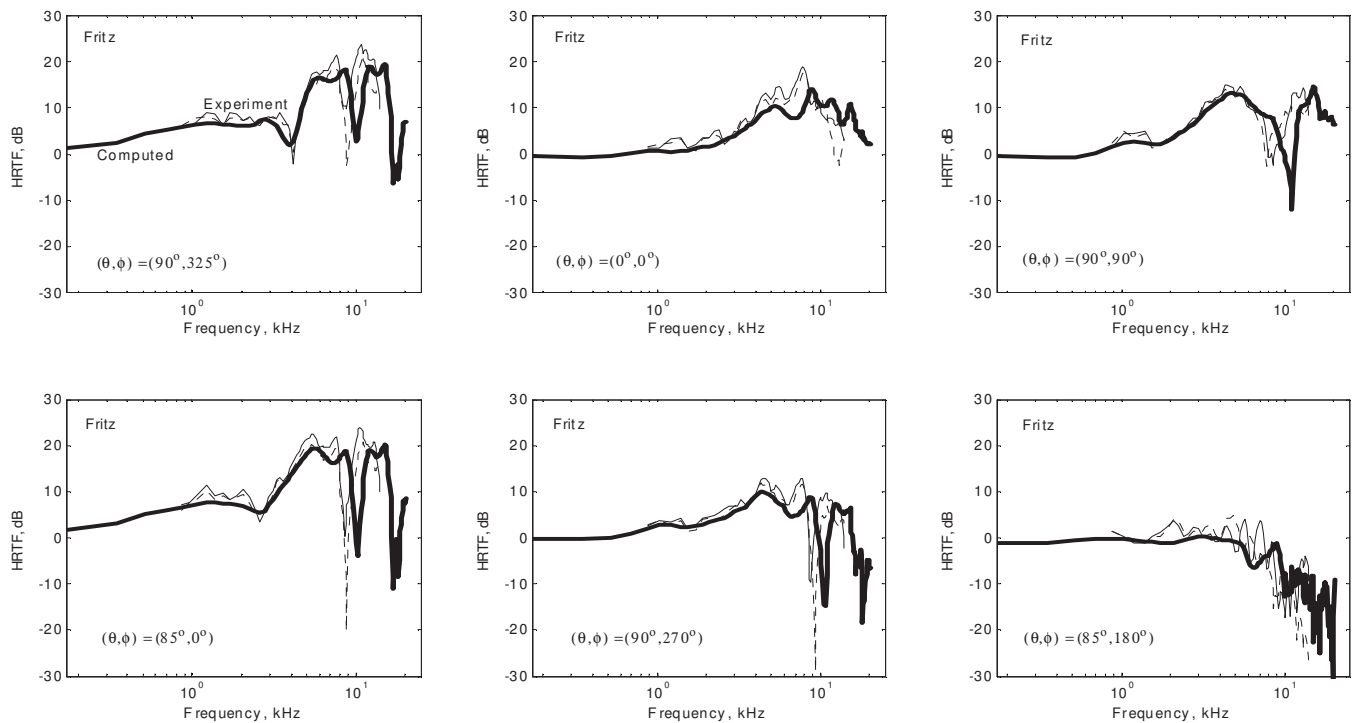


FIG. 5. Comparison of the computed and experimental HRTFs for the Fritz model for several angles. Computed data for the right ear are shown by the thick lines. The thin solid lines show experimental data for the right ear, while the dotted lines show data for the right ear derived from the experimental data for the left ear assuming symmetry of the model.

shift in the HRTF notches), which may be explained, e.g., by the differences in the computational and actual shapes. Figure 6 shows the data measured and computed at a constant $\theta=90^\circ$ elevation (coronal plane through the ears) for the frequency range of validity of the experimental data. Experimental data for $\theta=90^\circ$ measured at 58 azimuth angles ϕ were used for computations. This figure shows that an overall match exists between the computations and experiments, while quantitative discrepancy (in terms of notch locations) is clear for frequencies above 7–8 kHz.

B. Head-and-torso model computations

While computations for the head alone provide important information on the HRTF, the effect of torso on the HRTF is substantial. Modeling of this effect using a simple geometric approximation was performed in Refs. 28 and 29, while the shape influence of the head was considered in Ref. 30. In the present study we used a hybrid model consisting of detailed head shape of Kemar and torso modeled as an ellipsoid disjoint from the head. The parameters of the ellipsoid fitting the actual torso can be found in Ref. 28.

1. Meshes

The finest small pinnae mesh used for computations had 426 846 panels (445 276 for the finest “large pinnae mesh”). Both head-and-torso meshes satisfy the quality requirements for the BEM/FMM. Figure 1 illustrates the model and meshes used.

These meshes are suitable for computations of the HRTF in the entire audible range and provide not less than six elements per the shortest wavelength computed (frequency of

20.155 kHz). The dimensionless size of the computational domain for this frequency was $kD=410$, which is in the range of the BEM/FMM capabilities for personal computers (PCs). To reduce the computational cost we used two more meshes (“mid frequency” and “low frequency” meshes), which had 181 086 and 119 646 panels, respectively (for the small pinnae). These meshes had the same head meshing as the high frequency mesh, while we used coarser discretizations for the torso. The meshes were applied for frequency ranges 7.2–14.2 and below 7.2 kHz, respectively. In these ranges the head mesh certainly satisfied six elements/per wavelength criteria, while the torso discretizations were a bit rougher, but anyway not less than 5.5 elements/per wavelength.

2. Performance

Adding the torso substantially increases the computation time, as the size of the computational domain increases. The total time for the BEM solution was slightly below 70 h and required about 8 Gbyte RAM at the peak for the highest frequency.

3. Effect of torso

As a result of larger model size, plus the fact that the spectral representations can be applied to the region outside a large sphere concentric with the head, which includes the full torso, the number of harmonics for the head+torso model was substantially larger than that for the head alone. We found that for all frequencies several low frequency spatial modes of models with and without the torso match almost perfectly. The difference arises at high frequency spatial

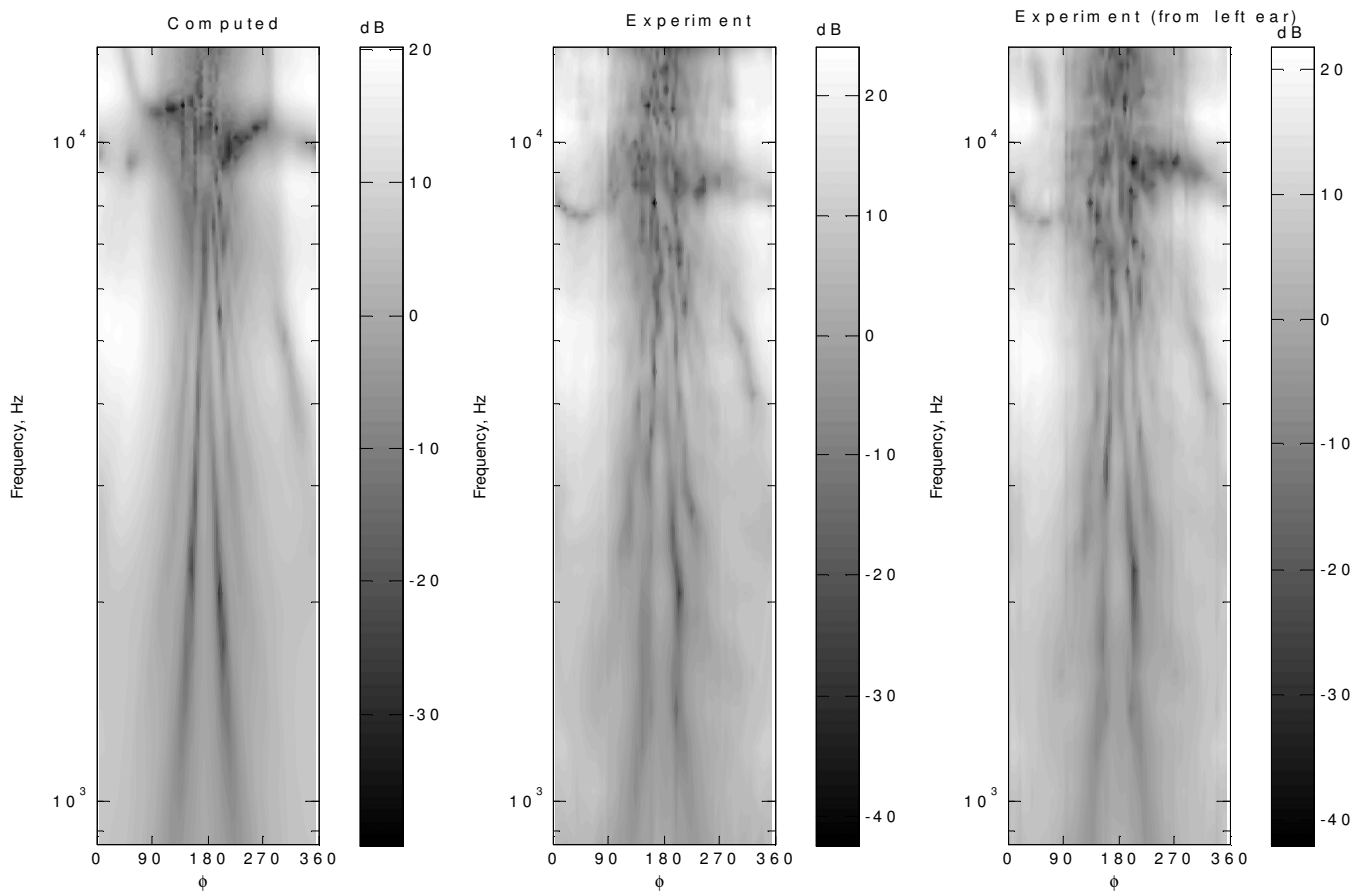


FIG. 6. Computed and measured HRTFs (in decibels) for the Fritz model (right ear) at angle $\theta=90^\circ$ for frequency range 800 Hz–14 kHz. The middle plot is obtained from the right ear measurements, while the right plot is obtained from the left ear measurements by applying symmetry.

modes starting from some value dependent on the frequency of the acoustic field. The higher the acoustic frequency the higher is this starting value. So, e.g., at 500 Hz spectral differences are observable for modes with $n \geq 3$, while for 5 kHz the difference is visible for $n \geq 10$ and for 15 kHz for $n \geq 25$. The high frequency spatial modes in the presence of torso have a considerable magnitude in a wider range.

Figure 7 shows the plane-wave HRTF. It is seen that the presence of torso results in interference patterns due to reflections from the torso, which spatial frequency is higher at higher acoustic frequencies. According to the spectral analysis at high acoustic frequencies the torso does not affect low frequency spatial modes, while high frequency spatial component shows up as a wave reflected from the torso, which is superposed with the HRTF for the head alone. At low frequencies when the wavelength is comparable with the head size the torso affects low frequency spatial components, which create effects of totally different spatial dependence of the HRTF.

Figure 8 represents the HRTF along the centerlines when the azimuth varies at constant elevation ($\theta=90^\circ$) and the elevation varies at constant azimuth ($\varphi=0^\circ$, $\delta=90^\circ-\theta$ and $\varphi=180^\circ$, $\delta=90^\circ+\theta$; angle δ is the same as in Ref. 28). The effect of torso is seen on both pictures, while for the latter case especially clear characteristic pattern for angles $180^\circ < \delta < 270^\circ$ is visible. This pattern is similar to that

measured for the snowman model and the Kemar model without ears reported in Ref. 28 at frequencies below 5 kHz.

4. Comparisons with experimental data

Several sets of experimental data from different researchers were compared with the computed data for Kemar. The basic conclusion we came to is that computations for the small pinnae model do not agree well with the experiments with the small pinnae at frequencies above 6 kHz or so, while computations for the large pinnae agree good enough with the large pinnae experiments up to 20 kHz. The major reason for this we see in the ways how the ear meshes were produced. In the former case that was done via the laser scans, which loses information on the depth and so the mesh has low fidelity to the real shape. In the latter case the CT-scans were used for mesh generation, which provide much better information on the features of the real shape. The importance for the use of quality scans for meshes and effect on the computation was also discussed in Ref. 9.

The comparisons presented below use experimental data from the database of HRTFs that was presented in Ref. 3, and which is referred to as the ‘‘CIPIC’’ database in the literature. All the data there are standardized in terms of measurement technique and data are presented in the hoop coordinates φ_{hoop} (azimuth) and θ_{hoop} (elevation). These angles are different from the spherical polar angles we used since in

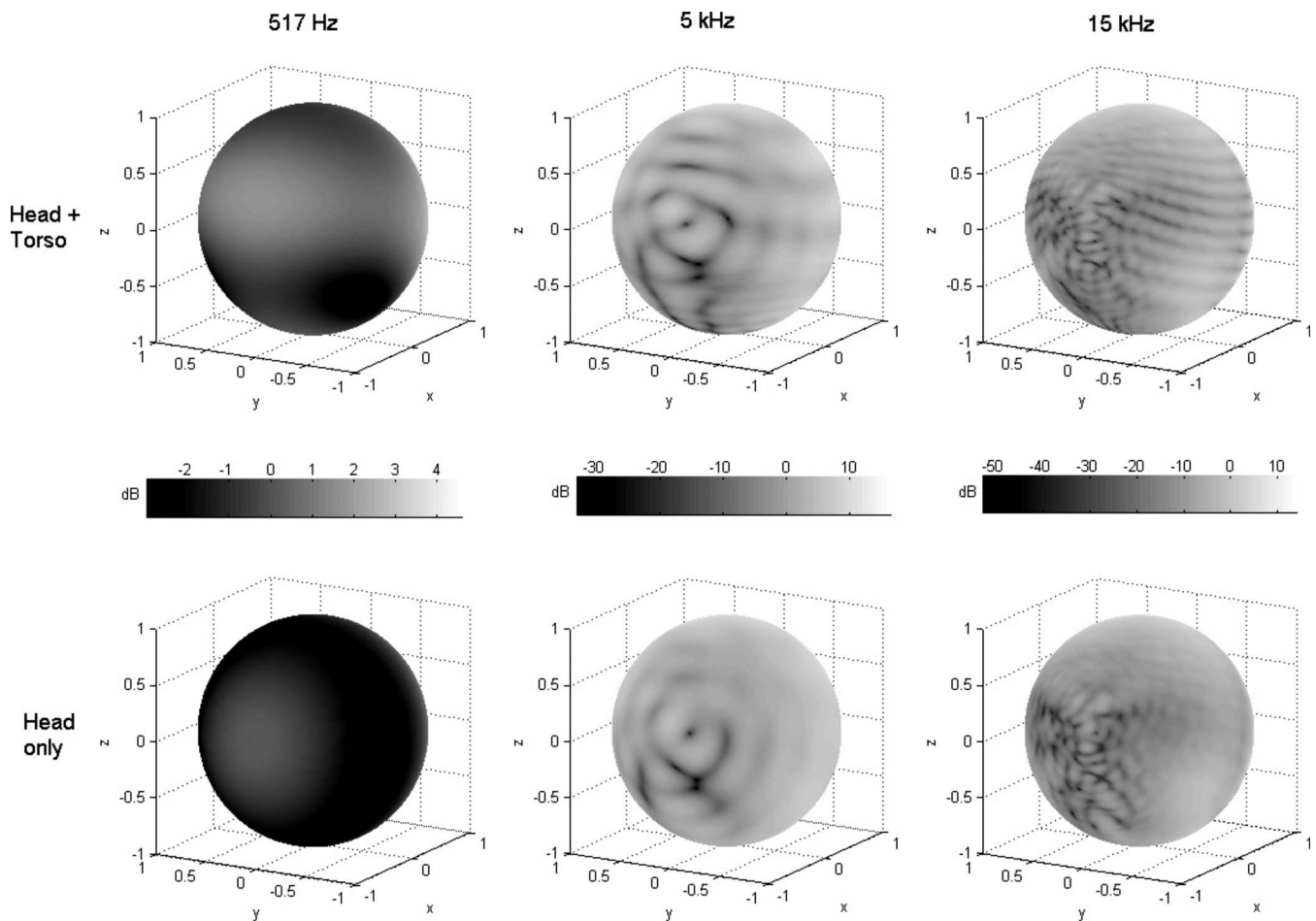


FIG. 7. The plane-wave right ear HRTF ($20 \log_{10}|H^{(pw)}|$) computed for the “small pinnae” Kemar with torso (the top row) and without torso (the bottom row) at different frequencies. The HRTF is plotted as a function of spherical angles θ and φ [see Eq. (1)] on a unit sphere. The center of the right ear is located at the $\theta=90^\circ$ and $\varphi=0^\circ$ ($x=1, y=0, z=0$). The views are from the side of the left ear. Axes (x, y, z) are oriented as shown in Fig. 1.

the experiments the poles of spherical coordinates were located at the ear positions. Relation between the angles (φ, θ) and ($\varphi_{\text{hoop}}, \theta_{\text{hoop}}$) follows from Eq. (1) and expression of the Cartesian coordinates via the hoop coordinates

$$\begin{aligned} \mathbf{r} &= (x, y, z) \\ &= r(\sin \varphi_{\text{hoop}}, \cos \varphi_{\text{hoop}} \cos \theta_{\text{hoop}}, \cos \varphi_{\text{hoop}} \sin \theta_{\text{hoop}}), \\ -\pi/2 < \theta_{\text{hoop}} \leq 3\pi/2, \quad -\pi/2 \leq \varphi_{\text{hoop}} \leq \pi/2, \quad r = |\mathbf{r}|, \end{aligned} \quad (47)$$

so the right ear location, for which all comparisons are made, corresponds to the pole $\varphi_{\text{hoop}} = \pi/2$.

Figure 9 shows comparison of the computation (plane-wave HRTF) and experiments for the frequency range 0.5–20 kHz for frontal direction ($\varphi_{\text{hoop}} = \theta_{\text{hoop}} = 0$). It is seen that for frequencies above 6 kHz or so there is a substantial deviation of the computations from experiments for the small pinnae case. On the other hand, computations for the large pinnae produce a HRTF pattern that matches the experiments as all maxima and minima of the frequency dependence are reproduced. Moreover, a good quantitative match is observed for frequencies up to 10 kHz and for the frequency and the depth of the strongest notch. It is also interesting that computations for head alone quantitatively are in better agree-

ment with experiments for frequencies above 10 kHz, while computations with the torso agree only qualitatively (the experimental data on this figure were normalized for 3 dB shift to match maximum value of the HRTF for the large pinnae case). This shows that to reproduce effect of torso quantitatively a better discretization of the torso than a simple fitting ellipsoid should be used. Such a mesh was unavailable to us.

Figure 10 supports these observations for the entire range of azimuths at fixed elevation and for all elevations at fixed azimuths. It is clear that computations for the large pinnae match experiments in terms of the notch structure (dark filamentary patterns), while experiments for the small pinnae show more complex notch structure than the computed one. The effect of torso is well seen on the charts plotted for the fixed azimuth. Both experiments and head-and-torso computations show the ripples due to torso reflections, which are not present in the computations with head alone. The computed and experimental ripple patterns and ripple frequencies agree well, while as we noted above the quantitative agreement for a given point cannot be very good, which happens due to some ripple phase mismatch and slight overall distortion of the ripple pattern due to approximate torso shape modeling.

Figure 11 shows comparisons of computed (head and torso) and experimental data for three fixed frequencies

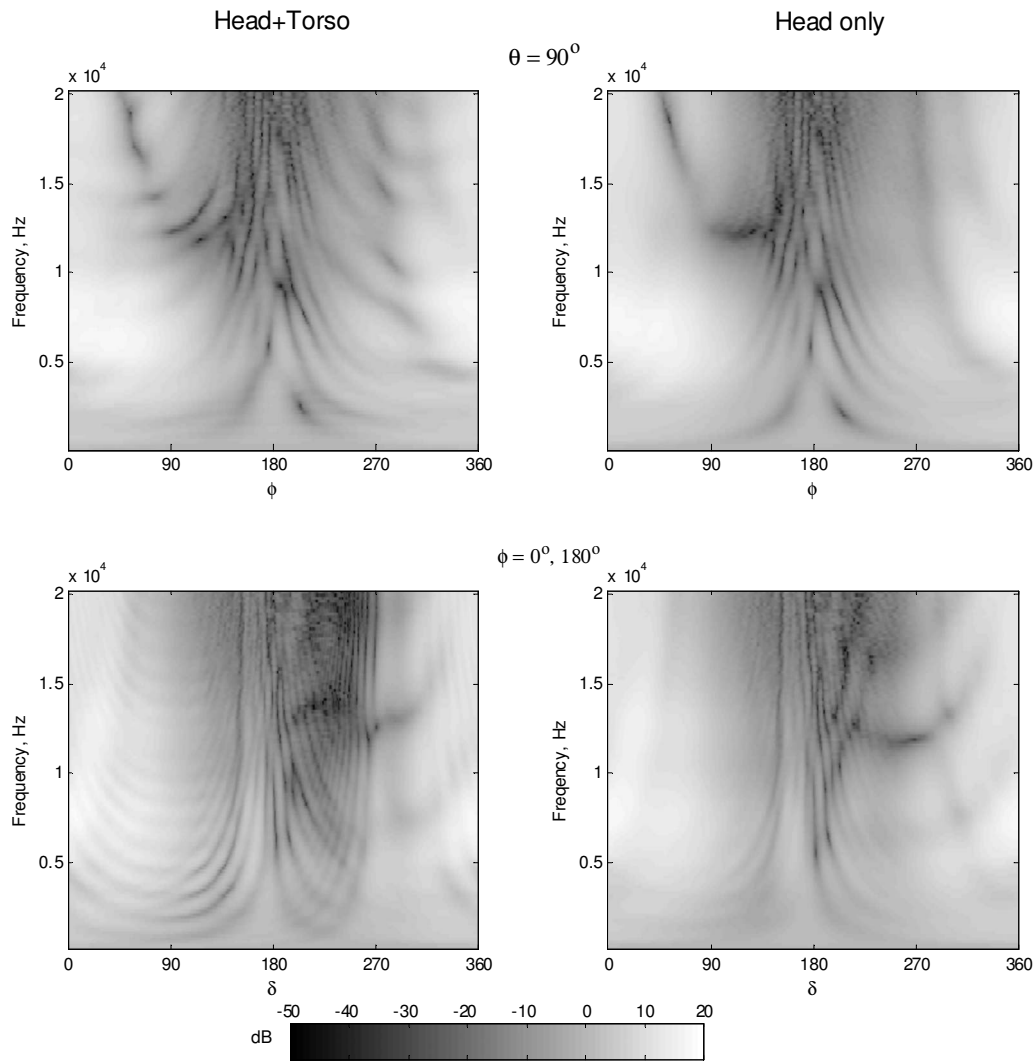


FIG. 8. The plane-wave right ear HRTF ($20 \log_{10}|H^{(pw)}|$) computed for the “small pinnae” Kemar with torso (the left column) and without torso (the right column) for fixed elevation ($\theta=90^\circ$, the top row) and for the centerline going through the ears and the top of the head (the bottom row). In the latter case angle δ is measured from the right ear; $\delta=(90^\circ-\theta) \bmod 360^\circ$ for $\varphi=0^\circ$ and $\delta=90^\circ+\theta$ for $\varphi=180^\circ$ ($0 \leq \theta \leq 180^\circ$).

(about 5, 14, and 19 kHz) and all measured azimuths and elevations, which gives an idea on spatial dependence of the HRTF for a given frequency. Again we can see consistency of the computed and experimental data for the “large pinnae

case,” while such consistency is observed for the small pinnae case only at frequency 5 kHz. At larger frequencies the latter case shows a different notch structure. We did not plot on this graph the computed data for the head alone, which are similar to computations taking into account the presence of the torso in terms of notch locations, but do not have reflection ripples. Such ripples are clearly seen in experiments, and the ellipsoid torso approximation reproduces this structure good enough.

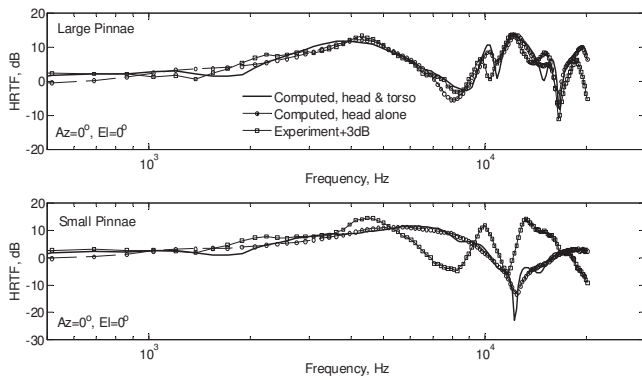


FIG. 9. Comparison of the computed HRTF for head and torso, head alone, and CIPIC experiments (the right ear) for the frontal direction (azimuth $=0^\circ$ and elevation $=0^\circ$). The top graph corresponds to the “large pinnae” model, while the bottom graph shows comparison for the “small pinnae” model. The experimental data are shifted up by 3 dB.

VI. CONCLUSIONS

The main purpose of the present study was development and testing of an efficient method for computation of the HRTF over the audible frequency range. We believe that this objective was achieved. Indeed, comparing our results with those of the direct boundary element methods, the present FMM accelerated BEM shows a superior performance and is suitable for computation of the HRTF in the full audible range including not only the head but the head-and-torso models. The method appeared to be stable and produces good results for meshes containing hundreds of thousands of

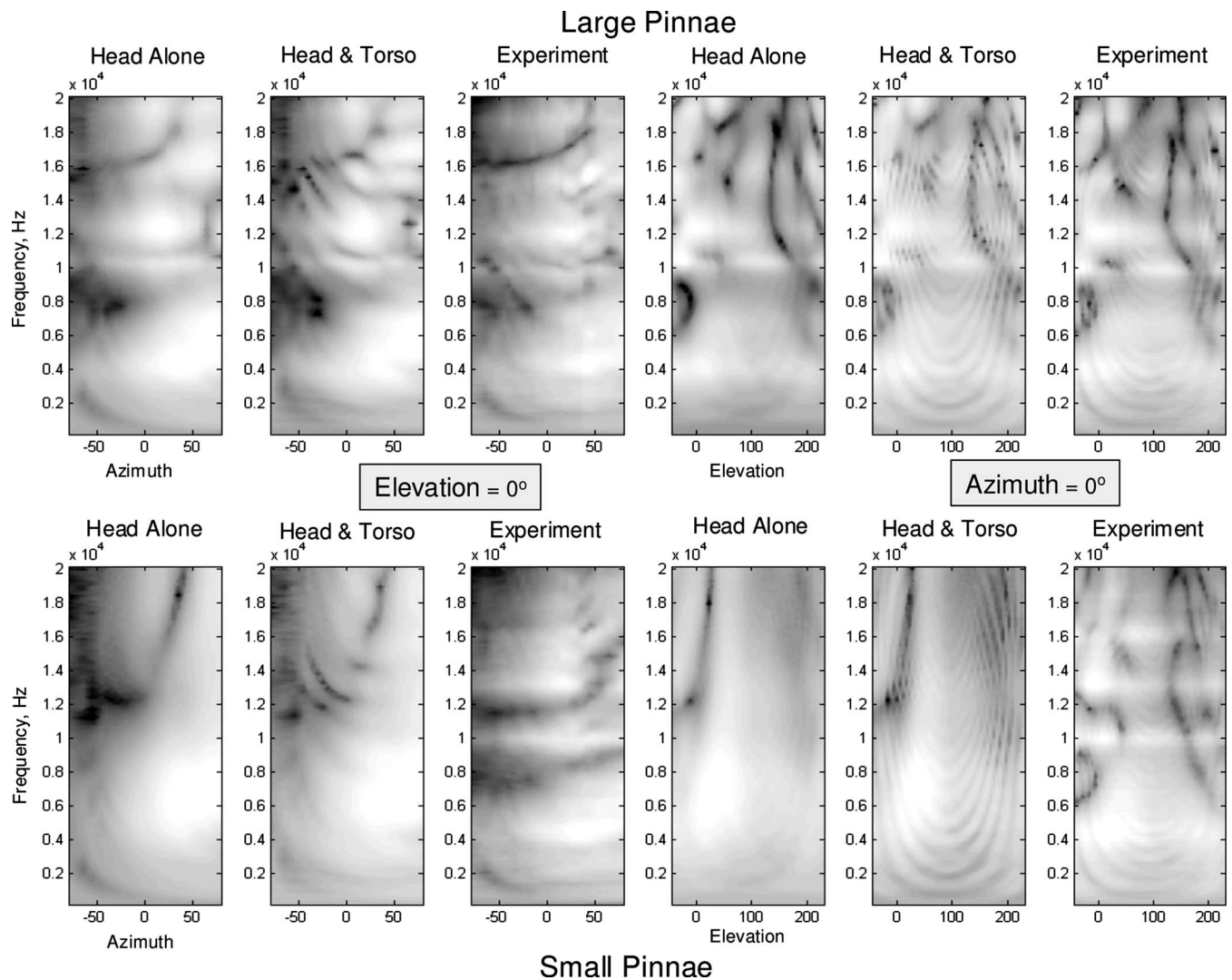


FIG. 10. Comparison of computed (head alone and head and torso) and CIPIC experimental HRTF data (in decibels) for the entire range of audible frequencies for fixed elevation (0°) and all measured azimuths ($-80^\circ \leq \varphi_{\text{hoop}} \leq 80^\circ$) (the six pictures on the left) and for fixed azimuth (0°) and all measured elevations ($-45^\circ \leq \theta_{\text{hoop}} \leq 230.625^\circ$) (the six pictures on the right). The top row shows data for the “large pinnae” case and the bottom row shows data for the “small pinnae” case. Experimental data are taken for the right ear. The dynamic range varies from about -60 dB (the darkest spots) to about $+10$ dB (the brightest areas).

boundary elements, which provides accurate computation and representation of the HRTF at frequencies as high as 20 kHz.

Representation of the HRTF via its discrete spherical spectrum, suggested in this paper (and for fitting of experimental data in Ref. 31), is rather natural. First, this is a relatively compact representation, second, it provides spectral accuracy consistent with the accuracy of computations, and third, it enables fast evaluation of the HRTF for any set of directions. In particular, this is important for rotation transforms, range dependence, and corrections of the HRTF due to body motion in the use of the HRTF in virtual reality.

Comparisons with experiments for the head alone and the head and torso showed that a good agreement can be obtained in the entire audible range. However, to achieve this goal several things should be handled with care. This relates to accurate scans of pinnae (laser scans may provide loss of important features, while the CT-scans are much more accurate though difficult to acquire), careful reproduction of ex-

perimental environment [e.g., positioning of the head model, modeling of torso (if present), etc.], and quality computational meshes.

Of course, the availability of a computational tool for the HRTF is only a first step. These should now be used to explore the dependence of HRTFs on shape, to compute HRTFs for different animals/individuals, and to study the influence of things such as clothing and hair. To aid in this effort it would be essential to combine the different research tools such as those developed in this study (the meshing procedures, the fast multipole software, and the software for analysis of the spherical spectra) into a usable tool. Then this tool should be used for computation of the HRTF for other meshes, modeling real human head/torso interactions for comparisons with data in available HRTF databases, and conducting experiments on perceptual experiences with the computed HRTFs. We were able to compute full range HRTF for tens of hours on a PC. However, for parametric studies, further accelerations are needed. These may be provided by

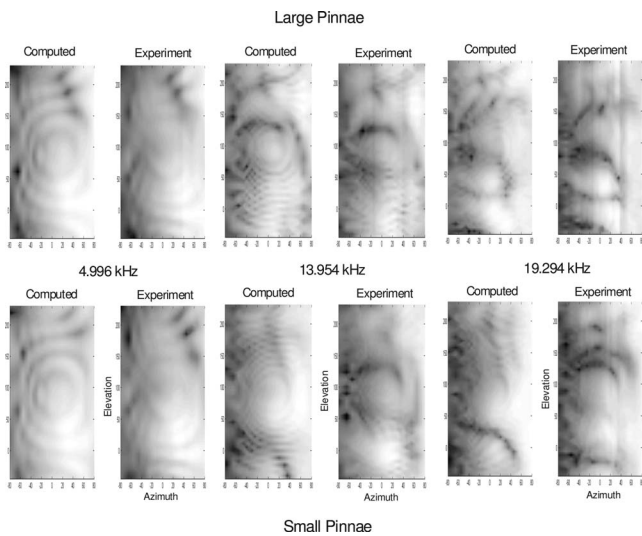


FIG. 11. Comparisons of the computed (head and torso) and CIPIC experimental HRTF data (in decibels) for three fixed frequencies (5.2, 14.1, and 19.2 kHz) and all azimuths and elevations measured in the experiments. The upper row shows the “large pinnae” case, while the bottom row shows the “small pinnae” case. The dynamic range for the highest frequency case varies from about -60 dB (the darkest spots) to about $+10$ dB (the brightest areas).

supercomputers, but we believe that the method can be further accelerated in the personal computing environments being made available, e.g., using graphic processors. Indeed, some promising results are obtained for the FMM for the Laplace equation on graphical processors,³² while the method should be developed further for the Helmholtz equation. Also we note that research is needed on the use of HRTF spherical spectra in applications.

ACKNOWLEDGMENTS

These computations are a result of a long term project that has been performed over the past 8 years at the Perceptual Interfaces and Reality Laboratory at the University of Maryland. At various stages our research has been supported by the NSF, DARPA, the state of Maryland, and Fantalgo, LLC; all are thanked for their support. We thank Professor R. O. Duda and Professor V. R. Algazi of the University of California, Davis for several discussions on the HRTF, for providing the CIPIC database, and for information on it. They are also thanked for providing the CT-scans of the Kemar large pinna. We would also like to thank Dr. Kexue Liu for creating the pinna mesh, Dr. Elena Grassi for the measurements of the Fritz HRTFs, Dr. Yuvi Kahana for the original Kemar mesh, and Dr. Brian Katz for the original Fritz mesh. We would like to acknowledge the use of software provided by Fantalgo, LLC for the fast multipole boundary element method.

¹D. W. Batteau, “The role of the pinna in human localization,” *Proc. R. Soc. London, Ser. B* **168**, 158–180 (1967).

²W. M. Hartmann, “How we localize sound,” *Phys. Today* **52**(11), 24–29 (1999).

³V. R. Algazi, R. O. Duda, D. M. Thompson, and C. Avendano, “The CIPIC HRTF database,” in *Proceedings of the 2001 IEEE WASPAA*, New Paltz, NY (2001), pp. 21–24.

⁴W. Gardner and K. Martin, “HRTF measurements of a KEMAR,” *J. Acoust. Soc. Am.* **97**, 3907–3908 (1995).

⁵D. N. Zotkin, R. Duraiswami, E. Grassi, and N. A. Gumerov, “Fast head-related transfer function measurement via reciprocity,” *J. Acoust. Soc. Am.* **120**, 2202–2215 (2006).

⁶Y. Kahana, P. A. Nelson, M. Petyt, and S. Choi, “Numerical modelling of the transfer functions of a dummy-head and of the external ear,” in *AES 16th International Conference on Spatial Sound Reproduction* (1999), pp. 1–6.

⁷B. F. G. Katz, “Boundary element method calculation of individual head-related transfer function. I. Rigid model calculation,” *J. Acoust. Soc. Am.* **110**, 2440–2448 (2001).

⁸Y. Tao, A. I. Tew, and S. J. Porter, “The differential pressure synthesis method for efficient acoustic pressure estimation,” *J. Audio Eng. Soc.* **51**, 647–656 (2003).

⁹M. Otani and S. Ise, “Fast calculation system specialized for head-related transfer function based on boundary element method,” *J. Acoust. Soc. Am.* **119**, 2589–2598 (2006).

¹⁰T. Walsh, L. Demkowicz, and R. Charles, “Boundary element modeling of the external human auditory system,” *J. Acoust. Soc. Am.* **115**, 1033–1043 (2004).

¹¹J. L. Aroyan, “Three-dimensional modeling of hearing in delphinus delphis,” *J. Acoust. Soc. Am.* **110**, 3305–3318 (2001).

¹²F. De Mey, J. Reijnen, and H. Peremans, “Simulated head related transfer function of the phyllostomid bat *Phyllostomus Discolor*,” *J. Acoust. Soc. Am.* **124**, 2123–2132 (2008).

¹³T. Xiao and Q.-H. Liu, “Finite difference computation of head-related transfer function for human hearing,” *J. Acoust. Soc. Am.* **113**, 2434–2441 (2003).

¹⁴R. Greff and B. F. G. Katz, “Round robin comparison of HRTF simulation results: Preliminary results,” *Audio Engineering Society Convention Paper*, the 123rd Convention (2007).

¹⁵N. A. Gumerov and R. Duraiswami, “A broadband fast multipole accelerated boundary element method for the 3D Helmholtz equation,” *J. Acoust. Soc. Am.* **125**, 191–205 (2009).

¹⁶P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (Princeton University Press, Princeton, NJ, 1968).

¹⁷M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (National Bureau of Standards, Washington, DC, 1964).

¹⁸N. A. Gumerov and R. Duraiswami, *Fast Multipole Methods for the Helmholtz Equation in Three Dimensions* (Elsevier, Oxford, 2004).

¹⁹A. J. Burton and G. F. Miller, “The application of the integral equation methods to the numerical solution of some exterior boundary-value problems,” *Proc. R. Soc. London, Ser. A* **323**, 201–210 (1971).

²⁰W. C. Chew, J. M. Jin, E. Michelsen, and J. Song, *Fast and Efficient Algorithms in Computational Electromagnetics* (Artech House, Boston, MA, 2001).

²¹N. A. Gumerov and R. Duraiswami, “A scalar potential formulation and translation theory for the time-harmonic Maxwell equations,” *J. Comput. Phys.* **225**, 206–236 (2007).

²²C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles* (Wiley, New York, 1983).

²³A. E. O’Donovan, N. A. Gumerov, and R. Duraiswami, “Good meshes for the fast multipole accelerated boundary element method via radial basis function interpolation and relaxation,” *University of Maryland Department of Computer Science Technical Report No. CS-TR-4937*. Also UMIACS-TR-2009-07 (2009).

²⁴N. A. Gumerov and R. Duraiswami, “Fast radial basis function interpolation via preconditioned Krylov iteration,” *SIAM J. Sci. Comput. (USA)* **29**, 1876–1899 (2007).

²⁵K. Liu, R. Duraiswami, and L. S. Davis, “A simple and optimal energy surface reconstruction algorithm from volumetric data,” *Department of Computer Science Technical Report CS-TR-4441 and UMIACS-TR-2003-11*, University of Maryland, College Park, MD, 2003.

²⁶Y. Saad, “A flexible inner-outer preconditioned GMRES algorithm,” *SIAM J. Sci. Comput. (USA)* **14**, 461–469 (1993).

²⁷B. F. G. Katz and D. R. Begault, “Round robin comparison of HRTF measurement systems: Preliminary results,” in *19th International Congress on Acoustics* (2007).

²⁸V. R. Algazi, R. O. Duda, R. Duraiswami, N. A. Gumerov, and Z. Tang, “Approximating the head-related transfer function using simple geometric models of the head and torso,” *J. Acoust. Soc. Am.* **112**, 2053–2064 (2002).

²⁹N. A. Gumerov, R. Duraiswami, and Z. Tang, “Numerical study of the

influence of the torso on the HRTF,” in Proceedings of the IEEE ICASSP, Orlando, FL, Vol. 2 (2002), pp. 1965–1968.

³⁰Y. Tao, A. I. Tew, and S. J. Porter, “A study on head-shape simplification using spherical harmonics for HRTF computation at low frequencies,” J. Audio Eng. Soc. **51**, 799–805 (2003).

³¹M. J. Evans, J. A. S. Angus, and A. I. Tew, “Analyzing head-related transfer function measurements using surface spherical harmonics,” J. Acoust. Soc. Am. **104**, 2400–2411 (1998).

³²N. A. Gumerov and R. Duraiswami, “Fast multipole methods on graphics processors,” J. Comput. Phys. **227**, 8290–8313 (2008).