FAST MULTIPOLE ACCELERATED BOUNDARY ELEMENTS FOR NUMERICAL COMPUTATION OF THE HEAD RELATED TRANSFER FUNCTION

Nail A. Gumerov

Fantalgo, LLC,
Elkridge, MD
info@fantalgo.com

Ramani Duraiswami, Dmitry N. Zotkin,

Perceptual Interfaces and Reality Laboratory
University of Maryland, College Park
{dz,ramani}@umiacs.umd.edu

ABSTRACT

The numerical computation of head related transfer functions has been attempted by a number of researchers. However, the cost of the computations has meant that usually only low frequencies can be computed and further the computations take inordinately long times. Because of this, comparisons of the computations with measurements are also difficult. We present a fast multipole based iterative preconditioned Krylov solution of a boundary element formulation of the problem and use a new formulation that enables the reciprocity technique to be accurately employed. This allows the calculation to proceed for higher frequencies and larger discretizations. Preliminary results of the computations and of comparisons with measured HRTFs are presented.

***Index Terms*—** Head related transfer function, Boundary element method, Fast Multipole Method, Reciprocity

1. INTRODUCTION AND PREVIOUS WORK

The Head Related Transfer Function (HRTF) is the Fourier transform of the impulse response of a human being in anechoic (or infinite) space to a source of sound placed at a location (r, θ, φ) in a head centered coordinate system measured at the entrance to the ear canal. Knowledge of the HRTF allows reintroduction of the cues that are caused by the scattering of sound off the body into headphone based reproduction of music and auditory reality. As the HRTF arises from a scattering process, it depends on the geometry (and composition) of the scatterer. Humans display a remarkable diversity in their sizes and shapes. In particular, their external ears (pinnae) exhibit considerable inter-personal variability. As a consequence, HRTFs are individual, and it is necessary to obtain users' HRTFs for recreation of auditory reality for them.

Experimental Work: Since the HRTF must be determined for each individual, there have been many attempts to measure the HRTF. A first series of approaches are based on acoustics and attempt to measure the HRTF. A conventional technique moves a source sequentially in space and measures the received sound for each source position at microphones placed at the entrance to the ear canals. A more recent approach, based on the principle of acoustical reciprocity, places a source at the entrance to the ear canal and measures the sound at a network of microphones placed in space. Since these measurements are made in parallel at many locations, this method is much faster. These approaches (with additional references) are summarized in [1, 14].

Numerical Approaches: A second approach attempts to obtain HRTFs via numerical computation. It is the subject of this paper.

RD & DZ gratefully acknowledge partial support of NSF Award III-0205271. NG acknowledges partial support of TEDCO, Maryland.

Since the HRTF results from a scattering process, it can be computed if a mesh of the human being is available. Any calculation must resolve the smallest wavelengths of interest, and to satisfy the Nyquist criterion, the discretization must involve at least two points per wavelength. In practice 6 to 10 points per wavelength are required. Since the wavenumber k is inversely proportional to wavelength, if a numerical method is a surface based method (such as the boundary element method), then problem size scales as $O\left(k^2 D^2\right)$, where D is the size of the domain. If the problem is solved via a volume based approach, then the number of discretization points scales as $O\left(k^3 D^3\right)$. In the former case dense matrices appear and typical direct solution procedures require $O(k^4 D^4)$ memory and $O(k^6 D^6)$ computation, while iterative solution can be achieved for $O\left(k^4 D^4\right)$ memory and $O\left(k^4 D^4\right)$ per iteration cost.

Katz [8] used a setting which replicated the conventional HRTF measurement setup and required a solution of the BEM equations for each source location and for each frequency. Kahana et al. [7] proposed the use of the reciprocal setup with a singular source placed at the entrance to the ear canal and with the computed field obtained in a single calculation at all points of interest. The calculation had to be repeated for each frequency of interest. Once the HRTF is obtained at a set of frequencies, an inverse Fourier transform could be used to obtain the time-domain head-related impulse response (HRIR), which is needed in applications. Both authors replicated the blocked-meatus setup that is necessary for headphone-based rendering applications. Algazi et al. [2] applied a BEM method to compare measurements of KEMAR HRTFs at low frequencies with a "snowman" model and validated an analytical model for the HRTF at low frequencies. Walsh et al. [12] used a more sophisticated boundary element approach (Galerkin formulation, Burton Miller approach to avoiding stability problems, parallel computing) to solve the equations using direct methods (PLAPACK). In addition, he studied ear canal resonance effects. All these authors could achieve calculations only up to about 6 kHz and required several hours of processing per frequency. Otani and Ise [9] shifted costs to the pre-processing phase in order to speed up calculation of the HRTFs for new source positions. They spent 420 hours to compute HRTFs for frequencies up to 12.5 kHz.

For the volume case, the resulting matrices are sparse. Despite the increased number of points, an iterative solution that is competitive with or faster than the surface techniques can be achieved, with typically $O(k^4 D^4)$ solution time and $O\left(k^3 D^3\right)$ memory. Xiao and Liu [13] implemented a finite-difference time domain approach to compute HRIRs directly. This approach allows a single computation to obtain all frequency components. Their calculations were valid to about 10 kHz and took about 5 hours of computing time. On the other hand, reciprocity was not used, and computations had to be repeated for each direction. Indeed, while singularities can be easily incorporated in the BEM, they are harder to include in a mesh based approach, and the use of reciprocity would require a sophisticated code. Furthermore, in neither the surface-based papers nor the volume-based ones (except the low frequency computations reported in [2]) was the torso included, which would further increase the computation time necessary.

Present Contribution: In many applications, it would be beneficial to have a capability to compute the HRTF given a mesh, even though HRTFs can be measured quite quickly [14]. Since HRTFs are well modeled for frequencies below 3kHz via an analytical anthropometric model, the computations are likely important only for frequencies higher than this. However high frequencies require finer meshes, and the storage and time required for computation increases. Thus faster algorithms with a capability to handle torsos are needed. Ideally the algorithm would execute relatively quickly and would allow one to use computation as a means of exploration of the factors influencing spatial hearing. While we do not achieve such result completely, we have made considerable progress towards it.

We employ an algorithm based on the fast multipole method (FMM) using coaxial translation operators described in [5, 6] to consider the reciprocal HRTF setup. With fast multipole acceleration based on coaxial translation, the cost of the surface methods can be reduced to $O(k^2D^2 \log kD)$ memory and per iteration cost at low frequencies ($kD \precsim 80$) and to $O(k^3D^3)$ for larger kD up to ($kD \simeq 400$) [6]. While the FMM can be used to accelerate the matrix-vector product (the dominant cost in an iteration), the number of iterations needed depend upon the condition of the matrix and the iteration method chosen. Preconditioned Krylov methods, such as the flexible generalized minimum residual method (FGMRES), ensure a convergent iteration, which, with an appropriate choice of a preconditioner, can help achieve quick convergence. With the resulting code, we are able to compute HRTFs for relatively high frequencies (up to 14 kHz with the torso and up to 22 kHz without). We also present a preliminary comparison of the computed HRTFs with measured ones for the KEMAR manikin. Our method employs the reciprocity approach, and we provide a better numerical formulation for it, avoiding some numerical errors caused by improper placement of the source in the domain as opposed to the boundary by previous authors.

2. NUMERICAL FORMULATION

We first set up the boundary value problem for the general BEM, and then for the reciprocal case. In the latter a source singularity is placed in the ear canal and the HRTF is measured at locations of interest by measuring the computed potential at those. This idea was first used in HRTF calculations by Kahana et al. [7]. Ideally, to reproduce the blocked meatus configuration, we need to place the source right on the mesh surface at the ear location. In [7] the source was placed at a location in the domain close to this surface. In comparisons of the results for a sphere (to be reported in an extended version), whose analytical HRTF is known, we found that this approximation performs poorly. In fact for perfect reciprocity we will require the source to be placed where the microphone would be in the direct measurement setup. Thus the best place would be on the boundary of the blocked meatus.

General formulation: Let the complex potential ϕ represent the Fourier transform of the pressure. It satisfies the Helmholtz equation on an infinite domain V in 3D and is subject to boundary conditions on S, the surface of the individual whose HRTF is being computed:

$$\nabla^2\phi + k^2\phi = 0, \quad \mathbf{x} \in V \subset \mathbb{R}^3, \quad k \in \mathbb{R}, \tag{1}$$

$$\mathbf{n}(\mathbf{x}) \cdot \nabla\phi(\mathbf{x}) = 0, \quad \mathbf{x} \in S,. \tag{2}$$

Here, as is customary we have assumed the sound hard boundary condition on the surface of the mesh, though a more general impedance boundary conditions is available in the software. On the boundary at infinity we assume the Sommerfeld condition

$$\lim_{|\mathbf{x}|\to\infty}\left(|\mathbf{x}|\left(\frac{\partial\phi}{\partial|\mathbf{x}|} - ik\phi\right)\right) = 0. \tag{3}$$

Arbitrary solutions to these equations can be expressed as sums of the single and double layer potentials [3]

$$\phi(\mathbf{y}) = \mathbf{K}(\sigma(\mathbf{x})) + \mathbf{L}(p(\mathbf{x})), \tag{4}$$

$$\mathbf{K}(\sigma(\mathbf{x})) = \int_S \sigma(\mathbf{x})\, G(\mathbf{x}-\mathbf{y})\, dS(\mathbf{x}),$$

$$\mathbf{L}(p(\mathbf{x})) = -\int_S p(\mathbf{x}) \frac{\partial G(\mathbf{x}-\mathbf{y})}{\partial n(\mathbf{x})} dS(\mathbf{x}),$$

where $\mathbf{x} \in S$, $\mathbf{y} \in V$. The values σ and p are surface densities, while G is the free space Green's function:

$$G(\mathbf{x}-\mathbf{y}) = \frac{e^{ik|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|}. \tag{5}$$

Green's identity holds, which is equation (4) with $\sigma(\mathbf{x}) = -q(\mathbf{x})$ and $p(\mathbf{x}) = -\phi(\mathbf{x})$, and which can be used to obtain ϕ in the domain if the potential and its normal derivative are known on the boundary.

Setting $\mathbf{x} \in S$, $\mathbf{y} \in S$ leads to the integral equation

$$\pm\frac{1}{2}\phi(\mathbf{y}) = \mathbf{K}(q(\mathbf{x})) + \mathbf{L}(\phi(\mathbf{x})), \tag{6}$$

which can be used together with Eq. (2) for determination of the boundary values. One approach to avoid spurious eigenvalues for the external BEM is to stay within the layer potential formulation [3], which, using the jump conditions, leads to:

$$\phi^{\pm}(\mathbf{y}) = \pm\frac{1}{2}p(\mathbf{y}) + \mathbf{K}(\sigma(\mathbf{x})) + \mathbf{L}(p(\mathbf{x})),, \tag{7}$$

$$q^{\pm}(\mathbf{y}) = \pm\frac{1}{2}\sigma(\mathbf{y}) + \mathbf{K}'(\sigma(\mathbf{x})) + \mathbf{L}'(p(\mathbf{x})), \quad \mathbf{x},\mathbf{y} \in S,$$

$$\mathbf{K}'(\sigma(\mathbf{x})) = \frac{\partial\mathbf{K}(\sigma(\mathbf{x}))}{\partial n(\mathbf{y})};\ \mathbf{L}'(\sigma(\mathbf{x})) = \frac{\partial\mathbf{L}(\sigma(\mathbf{x}))}{\partial n(\mathbf{y})}. \tag{8}$$

This can be used for solution of the Helmholtz equation, with $\sigma(\mathbf{x}) = i\eta p(\mathbf{x})$,where η is some complex parameter. Particularly for the external (scattering) problems, which solution is unique, this avoids spurious internal resonances.

Incorporating the source in the reciprocal set-up: As discussed earlier, we would like to move the source from a location near the surface to the location directly on it. We can represent $\phi(\mathbf{x};\mathbf{x}_*)$ as

$$\phi(\mathbf{x};\mathbf{x}_*) = G(\mathbf{x}-\mathbf{x}_*) + \phi^{scat}(\mathbf{x};\mathbf{x}_*);\ \phi^{scat} = G(\mathbf{x}-\mathbf{x}_*) + \phi'$$

We must find out how the scattered field is when the singularity is at the surface. A source near a hard boundary generates image source of the same intensity. As the source moves to the boundary, the part of the image source moves to the boundary as well. Accordingly

ϕ^{scat} contains a monopole source that is radiating outward, and a regular, radiating part ϕ'. The latter satisfies the boundary conditions

$$\frac{\partial \phi'}{\partial n(\mathbf{x})} = -2\frac{\partial G(\mathbf{x}-\mathbf{x}_*)}{\partial n(\mathbf{x})}, \quad \mathbf{x} \in S, \quad \mathbf{x} \neq \mathbf{x}_*.$$

At the location $\mathbf{x}_*$ on a flat surface it satisfies

$$\left.\frac{\partial \phi'(\mathbf{x};\mathbf{x}_*)}{\partial n(\mathbf{x})}\right|_{\mathbf{x}=\mathbf{x}_*} = 0,$$

to provide perfect reflection. With these boundary conditions, the BEM software described in [5] can be used unaltered.

BEM speed up with the FMM: Our goal is to use an iterative method that requires a matrix-vector product in every step. This product can be obtained quickly to a specified arbitrary precision using the FMM. Our approach to do this is described in [5]. In our BEM code we used surface discretization with flat triangular elements with a linear variation of the unknown function (potential or its normal derivative) over the panel. The basic iterative method we used is the Generalized Minimal Residual Method (GMRES) and its modifications (flexible GMRES, fGMRES) [10]. In this method there is an external iteration and several internal preconditioning iterations for each step of the external iteration. The choice of the preconditioner has an important bearing on the achievable accuracy. It should solve a similar system but at a much lower accuracy very quickly. We employ the FMM itself at a lower precision setting to create a preconditioner. Further details will be provided in a larger journal version of this paper.

Fig. 1. A mesh of KEMAR used in our runs. The mesh has 54945 nodes and 109882 elements. The ellipsoidal torso dimensions are those in [2]. For the computations above 12 kHz the torso was meshed using a half ellipsoid, with the overall mesh having the same total number of nodes and elements.

KEMAR mesh: We solve the problem of computing the HRTF of the KEMAR manikin,[1] which is widely used for testing in acoustics.

[1] http://www.gras.dk/00012/00330/

The mesh of the right half of the head of KEMAR was obtained courtesy of Dr. Yuvi Kahana, who originally scanned and processed the mesh. We processed the mesh a bit further for our purposes, and then reflected the mesh to obtain a whole head mesh. This mesh is able to resolve wavelengths of interest up to 9 kHz. Nevertheless, we used it for higher frequency computations. It is also somewhat noisy. For proper computations, a finer and smoother mesh is being currently developed.

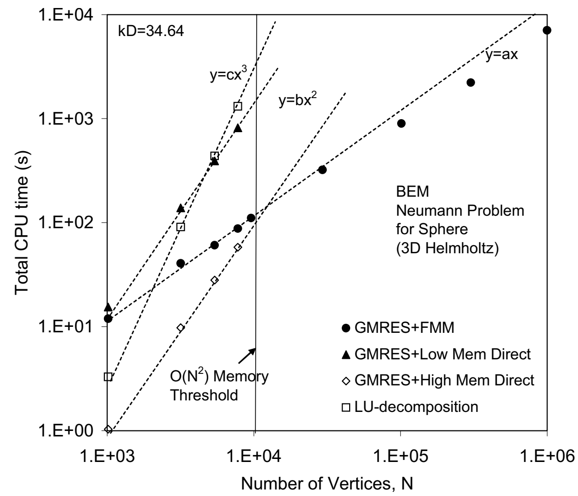

Fig. 2. The performance of the various components of our FMM BEM software on a test problem at various sizes. For comparison the performance of a typical BEM direct software package, and a BEM iterative package are also shown.

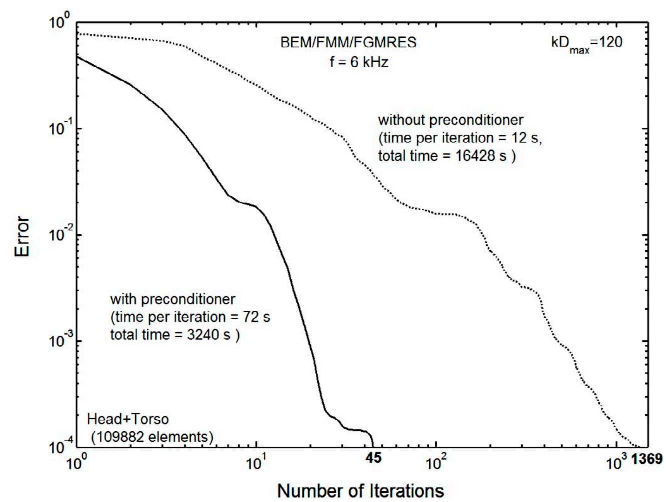

Fig. 3. Performance of the preconditioner for $kD = 120$.

While previous authors have not included a torso, we included a torso in our runs. A scanned KEMAR torso mesh was not available. However, it was shown in [2] using extensive measurements and simulations that an ellipsoidal body fit for KEMAR does a very good job. We accordingly meshed an ellipse of that size and used that for the torso in our runs. Fig. 1 shows the mesh used. As the frequency increases the influence of the torso decreases, and at high frequencies it only participates by providing a shoulder reflection for some directions. On the other hand the FMM algorithm used is not as efficient when the size of the domain is large for high frequency computations and the speed would improve with a smaller sized domain. Accordingly, for frequencies beyond 11 kHz we use just the top half of the ellipsoidal mesh, with bottom capped off. The total number of vertices and elements for this mesh is the same as for the full ellipsoid. For frequencies above 14kHz the head alone was used.

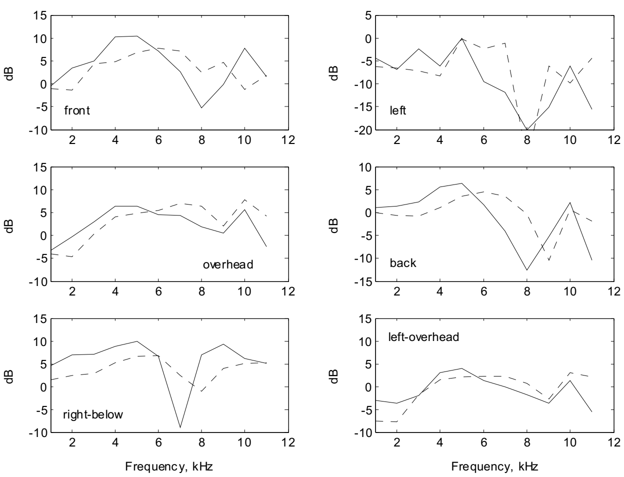

Fig. 4. Comparison of computed and measured HRTFs in the CIPIC database [1] for selected positions.

3. RESULTS

Test Results: The FMM/BEM software we used has been tested on a large number of problems and results were reported in [5]. In particular it was tested against analytical solutions such as scattering off a sphere, and was shown to perform as expected. One crucial item in the use of the software, especially at higher frequencies, is the performance of the preconditioner. Fig. 3 presents the performance with our FMM based preconditioner, and shows that the software converges in 45 more expensive iterations, as opposed to 1369 cheaper ones.

Calculation Procedure: The calculations reported here were performed for the frequencies (1-11 kHz) with the KEMAR head plus full torso, (12-14 kHz) with head plus half torso, and for 15 & 20 kHz with the head alone. Fig. 4 shows a comparison of the KEMAR with small pinna, right ear, from the CIPIC database [1].

4. CONCLUSIONS AND DISCUSSION

We have reported on a procedure that, given an appropriately resolved mesh, allows the calculation of a full HRTF on a desktop personal computer. While measurement techniques have improved sufficiently over the past few years (see [14]) so that it is possible to measure HRTFs in a minute, the availability of a numerical model should allow one to address questions not easily addressed via measurement.

Sensitivity analysis and cause and effect relationships: Having a numerical model of the process by which cues are generated gives us an opportunity to explore what cues are available, and what the extent of their variation are. Such sensitivity testing also provides a guideline for the accuracy required in the reconstruction of the body parts, and in the numerical solution. Once the sensitive components are known, one can build possible models of how the human brain performs source localization. Similarly, it is easier to perform virtual surgery on a mesh of an individual to see how a particular feature influence their HRTF.

Use in other scattering: The spatial scattering of sound is known to be important in audio, and can be used for modeling room acoustics. The availability of numerical modeling software should help in the use of computations as a tool in design of audio systems.

5. REFERENCES

[1] V. R. Algazi, R. O. Duda, D. M. Thompson and C. Avendano, “The CIPIC HRTF Database,” Proc. 2001 IEEE WASPAA, pp. 99-102, 21-24, 2001

[2] V. 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

[3] L. H. Chen and J. Zhou, *Boundary Element Methods*. Academic Press, 1992

[4] R. O. Duda and W. L. Martens, “Range dependence of the response of a spherical head model,” J. Acoust. Soc. Am. 104:3048-3059, 1998

[5] N. A. Gumerov and R. Duraiswami. “FMM accelerated BEM for 3D Laplace & Helmholtz equations,”. *Proceedings Int. Conf. on Boundary Element Techniques*, Paris, 2006

[6] N. A. Gumerov and R. Duraiswami, *Fast Multipole Methods for the Helmholtz Equation in Three Dimensions.* Elsevier, 2005

[7] Y. Kahana, P. A. Nelson, M. Petyt, S. Choi, “Numerical modeling of the transfer functions of a dummy head and of the external ear,” Proc. AES 16th Int. Conf. Spatial Sound Reproduction, Finland, 1999

[8] B. 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

[9] M. 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, 2005

[10] Y. Saad, “A flexible inner-outer preconditioned GMRES algorithm,” SIAM J. Sci. Comput. 14:461-469, 1993

[11] T. Sakuma and Y. Yasuda. “Fast multipole boundary element method for large-scale steady-state sound field analysis. Part I: Setup and validation”, Acustica/Acta Acustica, vol. 88, pp. 513-525, 2002; “Part II: Examination of numerical items”, Acustica/Acta Acustica, vol. 89, pp. 28-38, 2003

[12] T. Walsh, L. Demkowicz, R. Charles, “Boundary element modeling of the external human auditory system,” J. Acoust. Soc. Am. 115: 1033-1042, 2004

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

[14] D. N. Zotkin, R. Duraiswami, E. Grassi, and N. A. Gumerov. “Fast head-related transfer function measurement via reciprocity,” J. Acoust. Soc. Am. 120:xxx , 2006