

Approximating the head-related transfer function using simple geometric models of the head and torso

V. Ralph Algazi and Richard O. Duda^{a)}

CIPIC, Center for Image Processing and Integrated Computing, University of California, Davis, California 95616

Ramani Duraiswami, Nail A. Gumerov, and Zihui Tang

Perceptual Interfaces and Reality Laboratory, Institute for Advanced Computer Studies, University of Maryland, College Park, Maryland 20742

(Received 5 April 2002; accepted for publication 1 August 2002)

The head-related transfer function (HRTF) for distant sources is a complicated function of azimuth, elevation and frequency. This paper presents simple geometric models of the head and torso that provide insight into its low-frequency behavior, especially at low elevations. The head-and-torso models are obtained by adding both spherical and ellipsoidal models of the torso to a classical spherical-head model. Two different numerical techniques—multipole reexpansion and boundary element methods—are used to compute the HRTF of the models in both the frequency domain and the time domain. These computed HRTFs quantify the characteristics of elevation-dependent torso reflections for sources above the torso-shadow cone, and reveal the qualitatively different effects of torso shadow for sources within the torso-shadow cone. These effects include a torso bright spot that is prominent for the spherical torso, and significant attenuation of frequencies above 1 kHz in a range of elevations. Both torso reflections and torso shadow provide potentially significant elevation cues. Comparisons of the model HRTF with acoustic measurements in the horizontal, median, and frontal planes confirm the basic validity of the computational methods and establish that the geometric models provide good approximations of the HRTF for the KEMAR mannequin with its pinnae removed. © 2002 Acoustical Society of America. [DOI: 10.1121/1.1508780]

PACS numbers: 43.64.Bt, 43.66.Qp, 43.66.Pn [LHC]

I. INTRODUCTION

A. Variation of the HRTF with elevation

Head-related transfer functions (HRTFs) are central to spatial hearing, and have been studied extensively (Blauert, 1997; Carlile, 1996; Wightman and Kistler, 1997). The HRTF depends not only on the position of the sound source relative to the listener, but also on the size and shape of the listener's torso, head, and pinnae. The resulting complexity makes its behavior difficult to understand.

In this paper, we investigate the HRTFs for distant sources using very simple geometric models of the head and torso to gain insight into various features observed in acoustically measured human HRTFs. The simplest informative model is the spherical-head model. Introduced by Lord Rayleigh almost a century ago (Strutt, 1907), it has been used by many researchers to explain how the head affects the incident sound field (Hartley and Fry, 1921; Kuhn, 1977, 1987; Brungart and Rabinowitz, 1999). Although this model provides only a crude approximation to a human HRTF, it yields a first-order explanation and approximation of how the interaural time difference (ITD) and the interaural level difference (ILD) vary with azimuth and range.

However, the spherical-head model does not provide any cues for elevation.¹ It is well established that the pinna provides the major source of elevation cues (Batteau, 1967; Gardner and Gardner, 1973; Wright *et al.*, 1974).² The effect

of the pinna on the HRTF has been studied both experimentally (Mehrgardt and Mellert, 1977; Shaw, 1974, 1997; Wightman and Kistler, 1989) and computationally (Lopez-Poveda and Meddis, 1996; Kahana *et al.*, 1999; Kahana and Nelson, 2000; Katz, 2001). This work shows that the influence of the pinna is negligible below about 3 kHz, but is both significant and complicated at frequencies where the wavelength is short compared to the size of the pinna.

The torso also influences the HRTF and provides elevation dependent information (Kuhn and Gurnsey, 1983; Kuhn, 1987; Genuit and Platte, 1981). Although torso cues are not as strong as pinna cues perceptually, they appear at lower frequencies where typical sound signals have most of their energy. It has been shown that a simple ellipsoidal model of the torso can be used to calculate a torso reflection, and that such reflections provide significant elevation cues away from the median plane, even for sources having no spectral energy above 3 kHz (Algazi *et al.*, 2001a).

However, reflection is a short-wavelength or high-frequency concept, and modeling the effects of the torso by a specular reflection is only a first approximation. Furthermore, as the source descends in elevation, a point of grazing incidence is reached, below which torso reflections disappear and torso shadowing emerges. Rays drawn from the ear to points of tangency around the upper torso define a cone that we call the *torso-shadow cone* (see Fig. 1). Clearly, the specular reflection model does not apply within the torso-shadow cone. Instead, diffraction and scattering produce a qualitatively different behavior, characterized by the attenu-

^{a)}Electronic mail: rod@duda.org

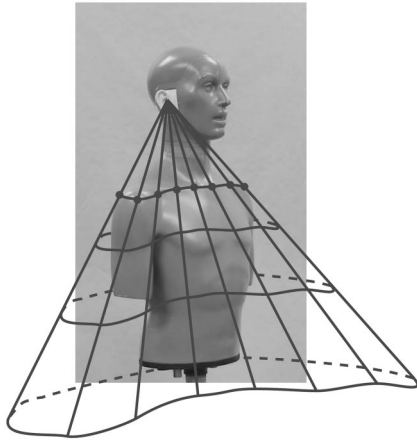


FIG. 1. The torso-shadow cone for the KEMAR mannequin. Rays from the ear are tangent to the torso at the points shown. At wavelengths where ray tracing is valid, specular reflections are produced by the torso for sound sources above the cone. The sound for sources within the cone is shadowed by the torso.

ation of high frequencies when the wavelength is comparable to or smaller than the size of the torso.

There are several reasons why the effect of the torso on the behavior of the HRTF for sources in the complete surrounding sphere has not been systematically measured or studied. First, the lengthy measurement process precludes asking human subjects to stand motionless, and seated measurements at low elevations are influenced by posture and/or the supporting chair, which introduces too many arbitrary variables. Although the use of a mannequin such as KEMAR (Burkhardt and Sachs, 1975) solves this particular problem, at very low elevations a truncated torso introduces meaningless artifacts of its own, and one would have to use a complete mannequin with intact arms and legs. Second, the torso effects appear at relatively low frequencies, where room reflections—even in anechoic chambers—make it hard to obtain accurate measurements. Third, it is experimentally difficult to place sufficiently large loudspeakers in the region directly below the subject. These obstacles have led to a lack of knowledge of HRTF behavior for sources in the torso shadow cone, a lack that may be responsible for the frequent observation that virtual sources synthesized with HRTFs rarely appear to come from really low elevations.

B. Methods for determining the HRTF for the snowman model

To gain a better understanding of the effects of the torso on the HRTF at all frequencies and elevations, a simple head-and-torso model called the snowman model was investigated. In its simplest form, the snowman model consists of a spherical head located above a spherical torso (Gumerov *et al.*, 2002). Unlike the isolated sphere, there is no elegant infinite-series solution for the scattering of sound waves by the snowman model. However, there are at least three ways to obtain the HRTF, all of which are employed in this paper:

- (1) acoustic measurements,
- (2) numerical computation using boundary-element methods, and

- (3) numerical computation using multipole reexpansions.

Each of these approaches has its characteristic advantages and disadvantages, which are summarized in turn.

Acoustic measurements entail no mathematical idealizations, are accurate over much of the audible frequency range, and can produce both HRTFs and head-related impulse responses (HRIRs) equally easily. However, room reflections make it difficult to measure the response at very low frequencies, physical constraints can make it difficult to position a loudspeaker at very low elevations, and measurement and alignment errors make it difficult to get the repeatability needed to study systematically the effects of changing the values of snowman parameters.

By contrast, the computational methods used in this study work particularly well at low frequencies, can be used for any source location, and are well suited to systematic parametric studies. However, they employ idealized assumptions, require validation, and have time/accuracy tradeoffs that limit the highest frequencies that can be used. Being frequency-domain methods, they provide the HRTF directly, but they require the computation of the HRTF at a large number of linearly spaced frequencies to invert the Fourier transform and extract the HRIR.

Boundary-element methods (or similar finite-difference and finite-element methods) can be applied to an arbitrarily shaped boundary surface (Ciskowski and Brebbia, 1991). However, the continuous surface must be approximated by a discretely sampled mesh of points in three dimensions, spaced at roughly one-tenth of the shortest wavelength of interest. It is a challenge to obtain a sufficiently accurate mesh for the human torso, head and pinnae, even without taking the possible effects of hair into account. Furthermore, to determine the response at high frequencies requires very dense sampling and correspondingly long computation times (Katz, 2001).

The multipole reexpansion method used in this paper is similar to the T-matrix method (Waterman and Truell, 1961), and it extends the classical infinite series solution for a single sphere (Morse and Ingard, 1968) to scattering by multiple spheres. The technique used employs new expressions for reexpansion of multipole solutions (Gumerov and Duraiswami, 2001a). Coupled with a procedure for enforcing boundary conditions on the sphere surfaces, it can be used to solve multiple scattering problems in domains containing multiple spheres (Gumerov and Duraiswami, 2001b). No meshes are required. Although reexpansion requires the use of numerical methods to solve the linear equations that define the boundary conditions, space and frequency can be sampled with arbitrarily fine resolution. In the particular case where the spheres are coaxial, multipole reexpansion can be several orders of magnitude faster than boundary-element methods. However, in the current version, convergence problems limit the highest frequencies that can be investigated.

In this paper, each method is used for a different purpose. Acoustic measurements are used to obtain the HRTF for the KEMAR mannequin and to validate the numerical methods. Multipole reexpansion is used for systematic studies of the snowman with a spherical torso. The boundary-



FIG. 2. The physical snowman model, which is composed of a 4.15-cm-radius bocce ball resting on top of a 10.9-cm-radius bowling ball. The probe tube microphone is inside the bocce ball, with the probe tip flush with the surface. The bowling ball is supported by a 0.5-cm-radius cylindrical rod.

element method is used for the snowman with an ellipsoidal torso. In the process, we (a) mutually validate the computational methods used, (b) identify the features of HRTFs for simple geometric models of the head and torso, (c) evaluate the adequacy of the snowman as an approximation to the human head and torso, and (d) use the snowman model to reveal the first-order effects of the torso and to identify possible localization cues.

II. METHODS

A. Measurement procedure

Acoustically measured HRTFs were obtained for two objects: the KEMAR mannequin shown in Fig. 1 and the physical snowman model shown in Fig. 2. Because head and torso effects are obscured by the presence of the pinnae, KEMAR's pinnae were removed and the exposed cavities were filled with putty and tape. Two Etymotic Research ER-7C microphones were placed inside the head, with the probe tips emerging at the entrance of the ear canals flush with the surface of the head. The Golay-code technique was used to measure the HRIRs (Zhou *et al.*, 1992). The test sounds were played through 3.2-cm-radius Bose Acoustimass™ Cube speakers mounted on a 1-m-radius hoop that was rotated about a horizontal axis through the midpoint of the interaural axis. The sampling rate for the measurements was 44.1 kHz. To remove room reflections, the resulting impulse responses were windowed using a modified Hamming window that eliminated everything occurring 2.5 ms after the initial pulse. The windowed responses were free-field equalized to compensate for the loudspeaker and microphone transfer functions. Because the small loudspeakers used were inefficient radiators at low frequencies, the low-frequency signal-to-noise ratio was poor, and it was not possible to completely restore the response below 500 Hz. As a result, measured HRTF values below 500 Hz should either be treated with suspicion or ignored.

To extend the useful range by another octave, the radius of the head of the physical snowman was made to be about half that of a human head, and the results were subsequently scaled in frequency accordingly. Specifically, the physical

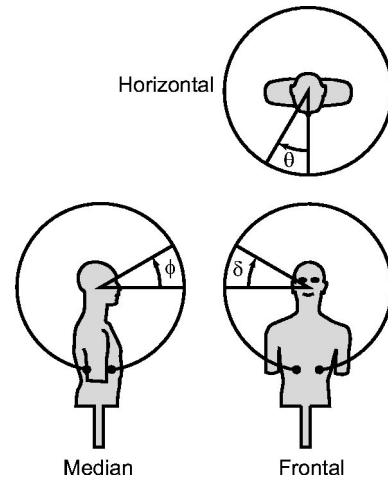


FIG. 3. The coordinate planes. In the horizontal plane, the azimuth θ ranges from 0 to 360 degrees. Support structures limit the range of the experimentally measurable elevation angles ϕ and δ . The entire 360-degree range is covered in the computational solutions.

snowman model consisted of a 4.15-cm-radius bocce ball resting on top of a 10.9-cm-radius bowling ball, with a small collar added to keep the head from rolling off. The model was supported by a 0.5-cm-radius metal rod. The bocce ball was drilled to accommodate one ER-7C microphone, with the probe tip emerging at a horizontal diameter.

For both KEMAR and the physical snowman model, measurements were made in the horizontal, median and frontal planes (see Fig. 3). The hoop was rotated in uniform steps of 360 degrees/128 \approx 2.8 degrees. For horizontal-plane measurements, the azimuth θ covered a full 360 degrees. For median-plane and frontal-plane measurements, the hoop support structure limited the elevation angles ϕ and δ to the interval from -81.6 to $+261.6$ degrees, so that no measurements could be made in a ± 8.4 -degree cone directly below the subject.

B. Computational procedures

Both of the computational methods used in this paper solve the Helmholtz equation (the Fourier transform of the wave equation) at a specified frequency in an infinite domain containing one or more scattering bodies. The Helmholtz equation is given by

$$\nabla^2 p + k^2 p = 0, \quad (1)$$

where p is the Fourier transform of the acoustic pressure, $k = \omega/c$ is the wave number, ω is the circular frequency, and c is the speed of sound. The incident pressure field p_{inc} is typically prescribed as the field from an isotropic point source, and the goal is to compute the scattered field, $p_{\text{scat}} = p - p_{\text{inc}}$, subject to boundary conditions at the surfaces of the scatterers and at infinity. We assume that the surfaces are "sound hard" ($\partial p / \partial n = 0$), and that the scattered sound field is outgoing at infinity.

When there is a single spherical scatterer, the scattered field at a point specified by the spherical coordinates (r, θ, ϕ) can be written in the form

$$p_{\text{scat}} = \sum_{l=0}^{\infty} \sum_{m=-l}^l a_{lm} h_l(kr) Y_{lm}(\theta, \phi), \quad (2)$$

where $h_l(\cdot)$ is the l th-order spherical Hankel function, $Y_{lm}(\cdot, \cdot)$ are the spherical harmonics, and the coefficients a_{lm} are determined by requiring that $p = p_{\text{inc}} + p_{\text{scat}}$ satisfies the boundary conditions on the sphere. Such a solution was used by Duda and Martens (1998) to represent the HRTF of a spherical head. If there are N spheres in the domain, one can exploit the linearity of the Helmholtz equation and write the solution as

$$p_{\text{scat}} = p_1 + p_2 + \cdots + p_N, \quad (3)$$

where

$$p_j = \sum_{l=0}^{\infty} \sum_{m=-l}^l a_{lm}^j h_l(kr_j) Y_{lm}(\theta_j, \phi_j). \quad (4)$$

Each of the functions p_j is centered at the corresponding sphere, and is expressed in a local spherical coordinate system. These series are truncated at some finite number of terms, and the coefficients a_{lm}^j are found by requiring that the boundary conditions at the surface of each sphere be satisfied. The procedure for doing this using multipole translation and reexpansion is presented in Gumerov and Duraiswami (2001b).

When the scattering surfaces are not spherical and the multipole reexpansion technique cannot be used, the boundary element method is used instead. This method works by using Green's identity to write Eq. (1) as an integral equation for the acoustic pressure. This equation specifies the pressure at a *field* point X on the surface of the acoustic domain as

$$C^X p^X = \int_{\Gamma_Y} \left[G^{XY} \frac{\partial p^Y}{\partial n^Y} - \frac{\partial G^{XY}}{\partial n^Y} p^Y \right] d\Gamma_Y, \quad (5)$$

where Γ is the surface of the acoustic domain, n is the unit outward normal vector to the acoustic domain at a surface (*source*) point Y , G is the free-space Green's function, and C is the jump term that results due to the treatment of the singular integral involving the derivative of the Green's function. The Green's function for the three-dimensional free-space problem, expressed in terms of the wave number k and the distance r between the source and field points, is

$$G^{XY} = \frac{\exp\{-ikr\}}{4\pi r}, \quad (6)$$

where $i = \sqrt{-1}$ is the complex constant. The surface of the scatterers is discretized using plane triangular elements. The equation is written at each boundary element, and a linear system of equations is obtained, which can be symbolically represented as

$$[F]\{P\} = [G] \left\{ \frac{\partial P}{\partial n} \right\}, \quad (7)$$

where $[F]$ and $[G]$ are matrices whose coefficients are obtained by evaluating integrals involving $\partial G/\partial n$ and G kernels, respectively; $\{P\}$ is the vector of acoustic pressures at the surface nodes, and $\{\partial P/\partial n\}$ is the vector of normal derivatives of the pressure. Imposition of the boundary condi-

tions leads to a system of linear equations that can be solved for the pressure.

Both the multipole reexpansion method and the boundary-element procedure yield the magnitude and phase of the HRTF at a particular frequency f . Computational time/accuracy tradeoffs limit the maximum allowable value of ka , where a is the head radius and $k = 2\pi f/c$ is the wave number, c being the speed of sound. For the snowman models that were investigated, the maximum useable value of ka was approximately 10. For the standard 8.75-cm head radius, this corresponds to a maximum frequency of about 6 kHz. To be conservative, all HRTF calculations were limited to exactly 5 kHz.

To obtain the HRIRs, the HRTFs for 500 frequencies uniformly spaced from 0 Hz to 5 kHz were calculated, and the ifft function in MATLABTM was used to calculate the inverse discrete Fourier transform. Because this procedure implicitly assumes that values of the HRTF above 5 kHz are all zero, direct use of the inverse transform leads to significant and distracting Gibbs phenomenon ripples in the impulse response. For graphical display, these ripples were removed by applying a standard Hamming window to the magnitude spectrum, leaving the phase unchanged (Oppenheim and Schaffer, 1969). Like low-pass filtering, windowing smooths the impulse response, reducing the height and increasing the width of pulses. For the window used, a unit impulse that would have a duration of 0.2 ms when sharply band limited to 5 kHz has its peak height approximately halved and its duration approximately doubled. Although this results in some loss of information, it greatly increases the clarity of the graphs.

III. THE FRONTAL-PLANE HEAD-AND-TORSO HRTF

To investigate the behavior of the HRTF, we start with the familiar spherical-head model and subsequently consider a sequence of progressively more complex cases. Because the results for the frontal plane exhibit a greater variety of behavior, we focus on it first.

A. The spherical head model

Figure 4 shows two image representations of the magnitude of the right-ear, frontal-plane HRTF for a sphere having the standard 8.75-cm head radius. These particular images were created using the algorithm presented in Duda and Martens (1998), but the same results were also produced by both the multipole reexpansion code and the boundary element code. Brightness corresponds to dB magnitude as shown by the grayscale bar at the top of the image. In Fig. 4(a), each vertical line corresponds to the frequency response at a particular elevation angle, δ . At low frequencies the response is 0 dB for any elevation angle. The largest response (which is approximately 6 dB at high frequencies) occurs at $\delta = 0$ degrees, where the source points directly at the right ear. As expected, the response is generally large on the ipsilateral side (-90 degrees $< \delta < 90$ degrees) and small on the contralateral side (90 degrees $< \delta < 270$ degrees). However, on the contralateral side the sphere exhibits a "bright spot," which appears in Fig. 4(a) as a bright vertical streak centered at $\delta = 180$ degrees. The dark bands on each side of this

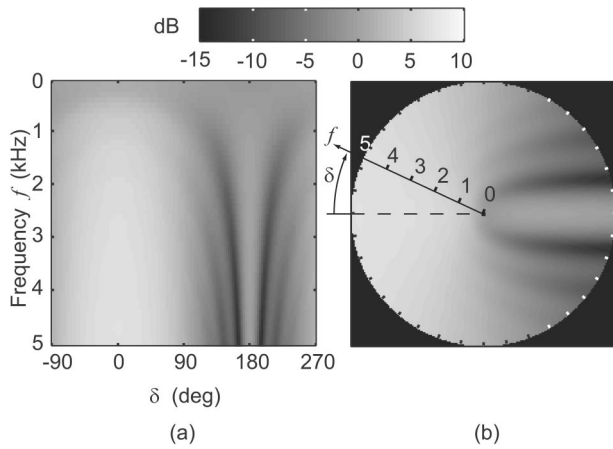


FIG. 4. Image representations of the magnitude of the HRTF for an ideal rigid sphere. The magnitude in dB is represented by brightness. The data is for the right ear of an 8.75-cm-radius sphere. The images are for the frontal plane (see Fig. 3). In (a), the left half of the image corresponds to the ipsilateral side, and the right half to the contralateral side. Note the bright spot that appears as a vertical streak at $\delta=180$ degrees. (b) shows the same data in polar coordinates. The elevation angle δ corresponds directly to the frontal view in Fig. 3. Thus, once again the left half corresponds to the ipsilateral side and the right half to the contralateral side, but the bright spot appears as a broad horizontal streak.

streak are interference patterns whose regularity is due to the perfect symmetry of the sphere; more irregularly shaped surfaces have the same general behavior, but the interference patterns become smeared.

Figure 4(b) is a useful alternative remapping of the information in Fig. 4(a). In this polar plot, frequency ranges from 0 to 5 kHz along any radial line. The center of the image corresponds to $f=0$, where the response is exactly 0 dB. The frequency response for incident sound waves arriving at an angle δ is found along the radial line at the angle δ as shown. This puts the HRTF display into direct correspondence with the coordinate system shown in Fig. 3.

The HRIR, which includes both the magnitude and the phase response of the HRTF, is particularly useful for revealing multipath components of the response. Figure 5(a) shows a family of HRIR curves corresponding to Fig. 4. Note that the ipsilateral response (-90 degrees $< \delta < 90$ degrees) is not only stronger than the contralateral response (90 degrees $< \delta < 270$ degrees), but it also occurs sooner. The approximately 0.7 ms difference in the arrival times at $\delta=0$ degrees and $\delta=180$ degrees is the maximum ITD. The

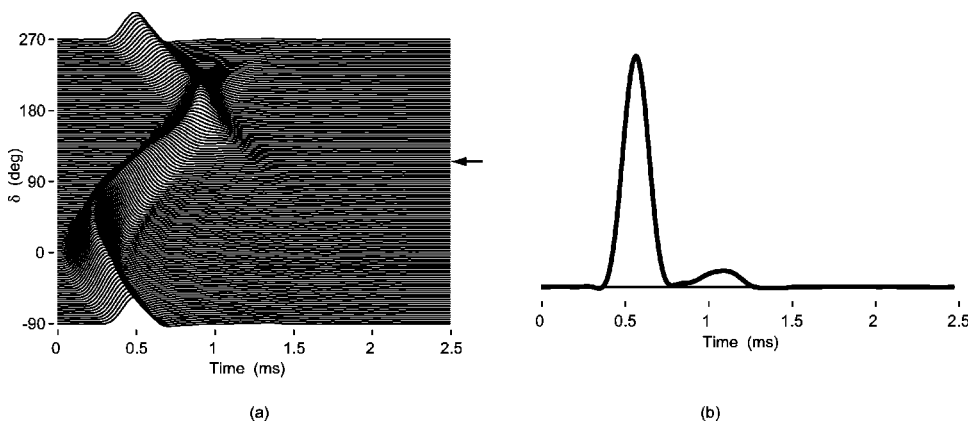


FIG. 5. (a) The HRIR for the sphere. The waveforms have a 5-kHz bandwidth, and the spectrum was smoothed with a Hamming window before inversion. The bright spot appears as a local maximum at $\delta=180$ degrees. (b) An expanded plot of the impulse response at $\delta=150$ degrees. This illustrates that near the bright spot the impulse response is bimodal. The weaker peak can be attributed to waves that travel around the contralateral side of the sphere. This second pulse is the source of the interference patterns seen in Fig. 4.

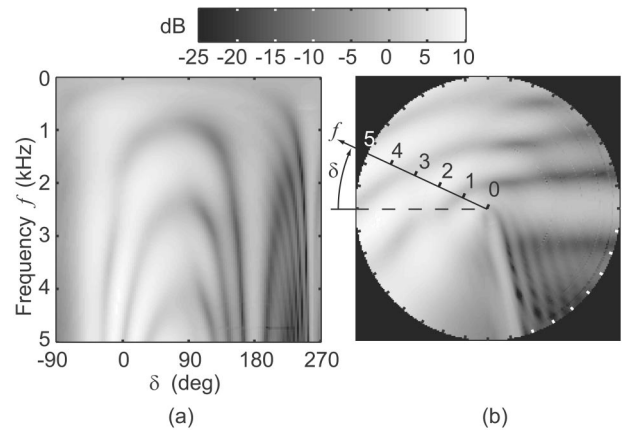


FIG. 6. The computed HRTF for the physical snowman model shown in Fig. 2, scaled for a head radius of 8.75 cm. The arch-shaped notches that are symmetric about $\delta=90$ degrees are due to specular reflections from the upper torso. The deeper notches around 210 to 250 degrees are caused by torso shadow. A torso bright spot can be seen around $\delta=255$ degrees.

bright spot appears as a local maximum in the response at $\delta=180$ degrees. In the vicinity of the bright spot, one can see a second pulse that follows the first pulse [see Fig. 5(b)]. It is this second pulse that is the source of the interference patterns seen in the frequency domain. A rough interpretation is that one pulse is composed of wave components traveling around the ipsilateral side of the sphere, and the other is composed of components traveling around the contralateral side; their in-phase confluence at $\delta=180$ degrees is the source of the bright spot.³

B. The physical snowman model

We now examine the effects produced by the introduction of the torso. Using the same 8.75-cm head radius, a 23-cm spherical torso is added directly below and tangent to the head. This results in a ratio of head size to torso size that is the same as that for the physical snowman shown in Fig. 2.

The frontal plane HRTF computed by the multipole re-expansion method is shown in Fig. 6. Comparison of Figs. 4(a) and 6(a) reveals two major differences. First, three arch-shaped notches centered near $\delta=80$ degrees appear on the ipsilateral side. The lowest-frequency notch occurs around 1 kHz at $\delta=80$ degrees, where the response dips to -5 dB. As was shown in Algazi *et al.* (2001a), these elevation-dependent notches are comb-filter interference patterns

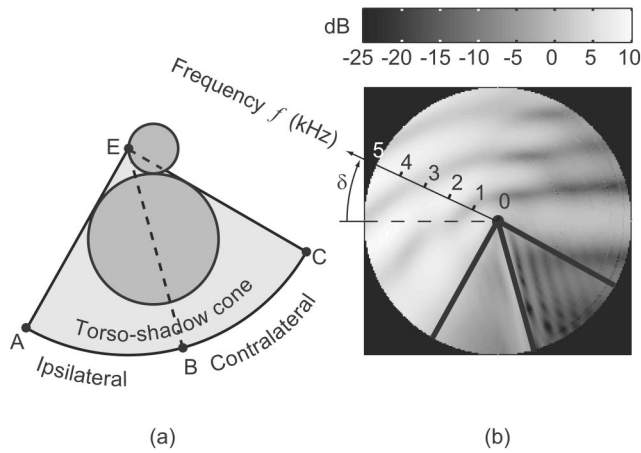


FIG. 7. Boundaries of the torso-shadow cone for the physical snowman model. The bisecting ray BE through the center of the torso to the right ear shown in (a) defines the direction of the torso bright spot. The tangent rays AE and CE are shown superimposed on the computed HRTF in (b). There is a close correspondence between the geometrical boundaries and the region of reduced response.

caused by torso reflection. They extend throughout the audible frequency range, and provide cues for elevation away from the median plane.

A second major difference is the appearance of deeper and more closely spaced notches on the contralateral side below the horizontal plane, where δ ranges from roughly 195 to 250 degrees. These low-elevation notches, which form a pattern of parallel lines in Fig. 6(b), fall in the torso-shadow cone; combined with head shadow, they cause the response for frequencies above 1 kHz to be much lower on the contralateral side than on the ipsilateral side.

Somewhat surprisingly, the lowest response does not occur when the source is directly below ($\delta = -90$ or $+270$ degrees). Instead, another bright spot appears at very low elevations. This “torso bright spot” is particularly clear in Fig. 6(b), where it forms a bright radial ridge near $\delta = 255$ degrees. Thus, the snowman model exhibits two bright spots, one due to the head around $\delta = 180$ degrees, and one due to the torso around $\delta = 255$ degrees.

There is a simple explanation for the torso bright spot. If a sound source below the torso were directed along a line from the center of the torso to the location of the right ear, and if the head did not disturb the sound field, a bright spot would be formed on the contralateral side of the torso and would strongly “illuminate” the right ear. For the dimensions of the physical snowman, the elevation angle for this line is 254.6 degrees, which is consistent with this interpretation.

The torso-shadow cone for the physical snowman is shown drawn to scale in Fig. 7(a). The ipsilateral limit is defined by the ray AE tangent to the torso on the ipsilateral side, and the contralateral limit is defined by the ray CE tangent to the torso on the contralateral side. BE, the ray through the center of the torso to the right ear, bisects the torso-shadow region. Sources in the ipsilateral zone are shadowed by just the torso, while sources in the contralateral zone are shadowed by both the torso and the head. In Fig. 7(b) these three rays are shown superimposed on the polar

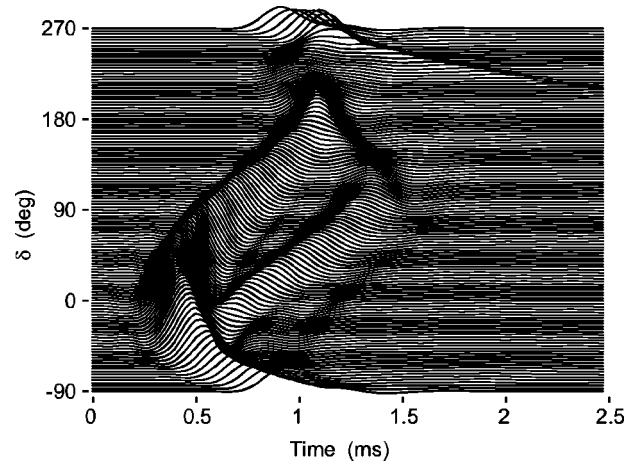


FIG. 8. The computed HRIR for the physical snowman. The torso reflection is prominent when the elevation δ is between -30 and 150 degrees. The torso bright spot can be seen near $\delta = 180$ degrees, and the torso bright spot near $\delta = 255$ degrees. The response for δ between 200 and 250 degrees is flattened and broadened by torso shadow. The long thin “tail” in this region is responsible for the strong interference notches seen in Fig. 6(b). A symmetric tail for δ between -90 and -40 degrees is barely visible.

HRTF plot of Fig. 6(b). Although the correspondence is not perfect, AE is closely aligned with the edge of reduced ipsilateral response, BE is closely aligned with the torso bright spot, and CE is roughly aligned with the edge of the even more reduced contralateral response. Thus, the geometric torso-shadow cone is in good agreement with the zones of reduced response. A similar geometric argument helps to explain why the torso reflection notches attain their lowest frequencies near $\delta = 80$ degrees. If the head is removed, the diagram in Fig. 7(a) is symmetric about the ray from B to E. Outside of the torso shadow cone, a pulse of sound directed to the ear at E is followed by a subsequent torso reflection. For pulses directed along the tangent ray from A to E or along the tangent ray from C to E, the delay between the initial pulse and the reflection is zero. By symmetry, the maximum time delay occurs for an overhead ray directed from E to B, and this leads to the lowest frequency for the interference notch. For the dimensions of the physical snowman, this overhead ray would be found at $\delta = 74.6$ degrees. Although the head disturbs the symmetry, this argument explains why the angle for the lowest notch frequency is biased below 90 degrees.

Additional insight can be obtained by comparing the HRIRs for the sphere and the snowman. The 5-kHz bandwidth HRIR, computed by inverting the multipole reexpansion, is shown in Fig. 8. Comparing these results to those for the isolated sphere [Fig. 5(a)], three prominent differences can be seen. First, the torso reflection is clearly present in the snowman HRIR in the general range of elevations from about -30 to $+150$ degrees. The maximum time delay between the main pulse and the torso reflection occurs when the source is overhead, and is approximately $\Delta T = 0.7$ ms. In the frequency domain, this corresponds to the first notch that occurs at $f_0 = 1/(2\Delta T) = 700$ Hz. This value is in good agreement with the location of the lowest frequency arch at $\delta = 80$ degrees in Fig. 6(a).

Second, the response in the interval from 200 to 250

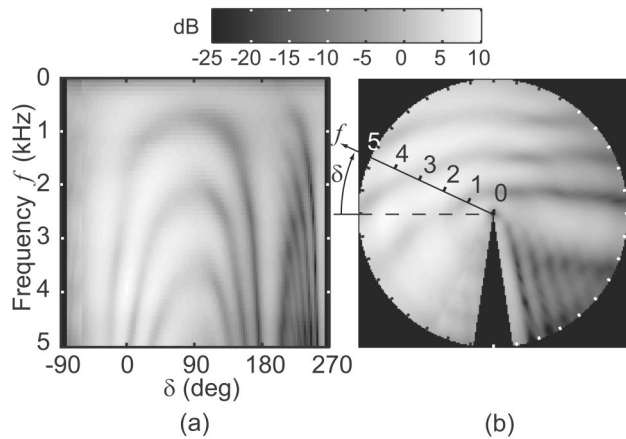


FIG. 9. The measured HRTF for the physical snowman model shown in Fig. 2, scaled for a head radius of 8.75 cm. These results should be compared with the computational results in Fig. 6.

degrees is significantly broader and flatter for the snowman than for the isolated sphere. This corresponds to the presence of torso shadowing on the contralateral side. Third, in this same interval there is a long, thin “tail” whose arrival time changes rapidly with elevation. Like the torso reflection, this second pulse produces notch filter interference patterns, and is the source of the deep notches that appear in the contralateral torso shadow zone in Fig. 6. This presence for this tail can be explained by imagining that waves traveling over the torso from the contralateral side of the torso-shadow cone can be divided into two groups. The group that arrives first travels over the contralateral side [along the line CE in Fig. 7(a)], where it is further shadowed by the head, while the group that arrives later takes a longer path over the ipsilateral side [along the line AE in Fig. 7(a)], and is not subject to head shadow. Although this is a crude explanation of a complicated diffraction and scattering phenomenon, it also explains why the same “tail” is much less prominent when the source is on the ipsilateral side, where the first arriving group of waves is not subject to head shadow but the second arriving group (the “tail”) is attenuated by head shadow.

IV. VALIDATION OF THE MULTIPOLE REEXPANSION TECHNIQUE

To confirm the validity of the multipole reexpansion method and of the characteristics of the snowman HRTF just presented, we now compare those results to the results of

acoustic measurements of the physical snowman model shown in Fig. 2. Figure 9 shows the measured HRTFs, scaled in frequency to match the standard 8.75-cm head radius. As was mentioned in Sec. II A, experimental constraints limited the elevation angles to the interval from -81.6 to $+261.6$ degrees; regions for which no data are available are shown in solid black. Ignoring those regions, a comparison of Figs. 6 and 9 shows a generally very good correspondence between the computational and the acoustic results.

In particular, the torso reflections and the bright spot for the head are in excellent agreement. There is also good general agreement regarding the behavior in the contralateral torso shadow region, including the torso bright spot. The main difference is that the interference notches are a bit deeper and more closely spaced in the computational data than in the measured results. Because such interference patterns are quite sensitive to small changes in path lengths, this difference could be caused by any of a number of imperfections in the experimental conditions, such as the presence of a supporting rod, the presence of the collar, or small angular misalignments.

Comparison of the HRIRs reveals some differences that are not as evident in the frequency domain. As was mentioned in Sec. II A, it was not possible to compensate for the low response of the loudspeakers below 500 Hz without exposing low-frequency noise. The measured data is effectively a high-pass-filtered version of the true responses, with a 3-dB corner frequency around 500 Hz. This is revealed by a slight darkening at the top of Fig. 9(a) that is not visually prominent in the frequency domain, but that shows up clearly in the time domain as a negative overshoot following the main pulse. The high-pass filtering hides the low-frequency noise that is present in the measured data.

Thus, in the comparison of the computed and the measured impulse responses shown in Fig. 10, the computed response was filtered by a single-pole high-pass filter with a 500-Hz corner frequency to introduce a comparable distortion to the curve [see Fig. 10(a)]. In addition, a Hamming window was used to band-limit the measured impulse responses to 5 kHz to account for the lack of high-frequency energy in the computed results [see Fig. 10(b)]. With these corrections understood, there is again very good agreement between the computed and the measured data. This gives confidence that the multipole reexpansion technique is providing correct HRTF values in both magnitude and phase.

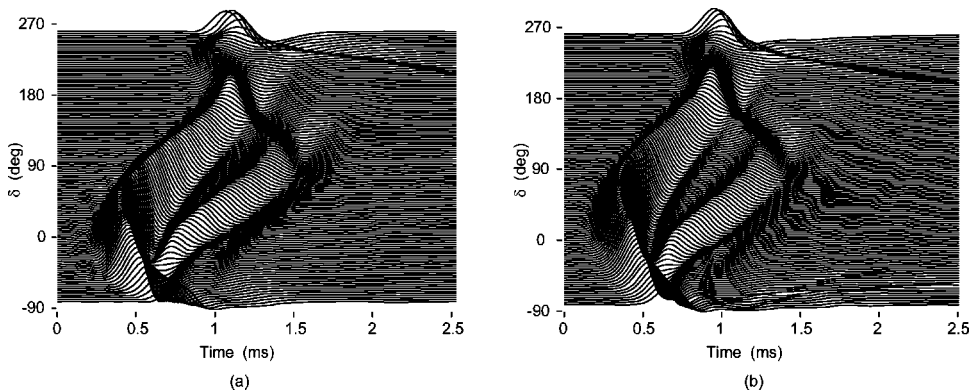


FIG. 10. Comparison of the (a) computed and (b) measured impulse responses for the physical snowman. The computed HRIR was high-pass filtered to simulate the effects of loss of low-frequency information in the measured data. This accounts for the negative overshoot following the main pulse. Similarly, the measured HRIR was low-pass filtered (bandlimited to 5 kHz) to simulate the absence of high-frequency information in the computed results. The range of elevations is limited to the range for which measurements are available.

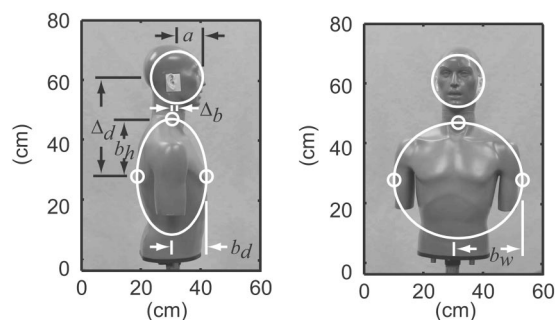


FIG. 11. Fitting a spherical head and ellipsoidal torso to the KEMAR mannequin.

V. EVALUATING THE SNOWMAN MODEL

To investigate how well a human HRTF can be approximated by the HRTF of the snowman model, we decided to compare the HRTF for the KEMAR mannequin shown in Fig. 1 to two snowman models, one with a spherical torso and one with an ellipsoidal torso. This process involves two steps: (a) geometrically fitting the snowman models to KEMAR, and (b) comparing the resulting HRTFs. Each process is explained in turn.

A. Fitting the snowman model to KEMAR

The optimum snowman model for KEMAR is the one whose HRTF best matches KEMAR's HRTF. However, there is no obvious or well-established measure of error in comparing two HRTFs. As a consequence, it was decided simply to match anthropometric characteristics. The head and torso components were fit separately, placing more emphasis on fitting the upper torso than the lower torso. Although better results could be obtained with a more sophisticated fitting procedure, the method used had the virtue of being easy to understand and apply.

The spherical head is defined by its center and radius. The center was located at the center of a bounding box for KEMAR's head (see Fig. 11). When the regression equation in Algazi *et al.* (2001b) was used to estimate the optimum head radius, a value of 8.70 cm was obtained, which is remarkably close to the standard 8.75-cm value. To fit the torso, the five control points shown as open circles in Fig. 11 were used.⁴ This was almost sufficient to define an ellipsoidal torso model whose principal axes were parallel to the coordinate axes, but left one degree of freedom free. This was resolved by arbitrarily placing the center of the ellipsoid at the intersection of the lines joining the left/right and front/back control points.

The resulting spherical head and ellipsoidal torso are shown in Fig. 11. The dimensions for this "best-fitting" model were as follows:

- Head radius, a 8.7 cm
- Torso half height, b_h 19.3 cm
- Torso half width, b_w 21.5 cm
- Torso half depth, b_d 11.6 cm
- Torso geometric mean radius, b 16.9 cm

Torso displacement from head, back, Δ_b 1.7 cm

Torso displacement from head, down, Δ_d 30.9 cm.

The resulting snowman model provides just a first approximation to KEMAR. The spherical head is wider, shorter, and shallower than the real head. The ellipsoidal torso matches the real torso fairly well on top, but is badly mismatched below the control points. This was an intentional compromise, and was made for three reasons. First, an ellipsoid that would provide a better match to the lower torso would necessarily sacrifice the more important upper region. Second, in the torso shadow region where the match is poor, KEMAR is no longer a realistic model for a human. Third, at very low elevations, the wavelengths that are unattenuated by torso shadow are sufficiently long that there is reason to hope that a crude fit will be adequate and that the specific shape is relatively unimportant. Finally, the head is not tangent to the torso, leaving a gap where the neck should be. Although a more realistic model would include a neck component, it was omitted in the interest of simplicity.

For an even simpler approximation, the head was centered over the torso, and the ellipsoid was approximated by a sphere whose radius (16.9 cm) was the geometric mean of the semi-axes of the ellipsoid, so that the volume was unchanged. The resulting spherical torso is a coarser approximation, exhibiting a symmetry about the vertical polar axis that is not realistic. However, it provides a model that is easy to understand and evaluate.

B. Comparing the KEMAR and the snowman HRTFs

We now compare the pinnaless KEMAR HRTF to that of the spherical-torso and ellipsoidal-torso models. In all cases, the KEMAR HRTFs were measured acoustically, the spherical-torso results were computed using the multipole reexpansion method, and the ellipsoidal-torso results were computed using the boundary-element method. We compare frontal-plane results in some detail, and compare horizontal-plane and median-plane results more briefly. We start by comparing the KEMAR HRTF to the HRTF for the spherical-torso approximation, and then examine the improvement that the ellipsoidal-torso model can provide.

1. Frontal plane

The frontal-plane KEMAR HRTF is shown in Fig. 12(a), and the corresponding spherical-torso HRTF is shown in Fig. 12(b). The gross characteristics are very similar. In particular, in both cases one can see the presence of a strong response on the ipsilateral side, torso reflection notches at upper elevations, head shadowing on the upper contralateral side, and even stronger torso shadowing on the lower contralateral side. Thus, the model provides an approximation to KEMAR that captures important, first-order effects. However, there are numerous differences as well, caused either by the simplicity of the model or the choices made in fitting the model to KEMAR.

On the upper contralateral side, the head bright spot is almost as well defined for KEMAR as it is for the model. However, it occurs around $\delta=155$ degrees for KEMAR versus 180 degrees for the model, which is a 25-degree shift.

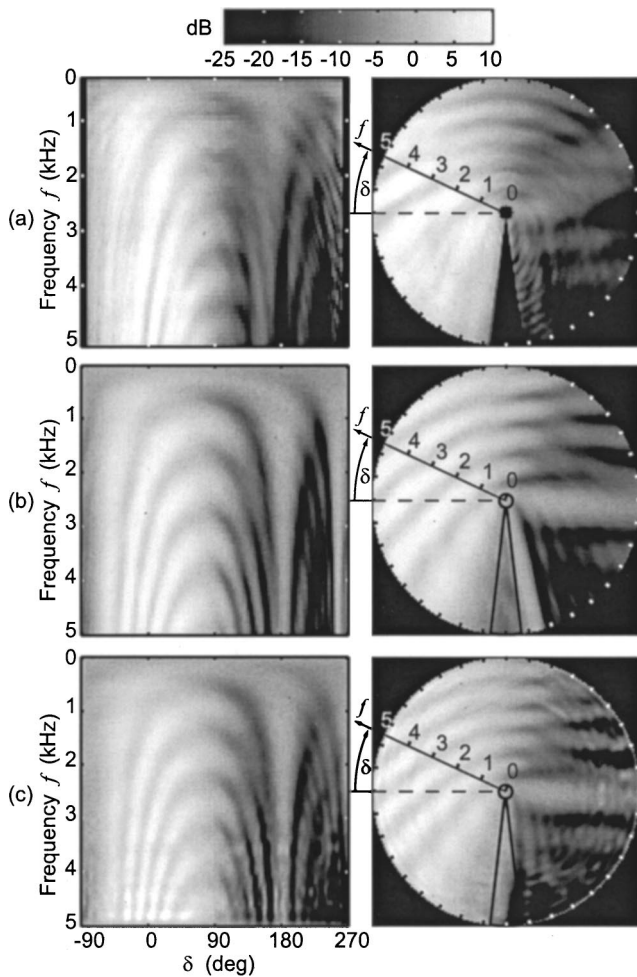


FIG. 12. Measured and computed frontal-plane HRTFs. (a) KEMAR with the pinnae removed. (b) The spherical-torso snowman approximation. (c) The ellipsoidal-torso snowman approximation.

The reason for this shift can be traced to the way that the spherical head was fit to KEMAR's head. The actual top of KEMAR's head is about 3 cm above the top of the sphere (see Fig. 11). If the spherical head were to be shifted up 3 cm without moving the z coordinate of the ears in space, the ears would no longer be across a diameter, but would be down by $\sin^{-1}(3/8.7) \approx 20$ degrees, which would introduce the corresponding shift in the position of the bright spot and would account for most of the discrepancy. However, shifting the head up results in a poorer fit to the data at low elevations. If this difference is important, a better solution would be to use an ellipsoidal head model (Duda *et al.*, 1999).

On the upper ipsilateral side, the torso-reflection arches for KEMAR and the model are very similar, with the highest-frequency fifth arch being even better defined for KEMAR than for the model. However, there is a systematic shift in the elevation at which the notches reach their lowest frequency, being around $\delta = 95$ degrees for KEMAR and around 75 degrees for the model. This shows up clearly as a 20-degree rotation of the torso reflection notches in the polar plots. Using the dimensions for the model, the heuristic argument given in Sec. III B predicts that the elevation for the lowest frequency should appear at $\delta = 90$ degrees $-\tan^{-1}(8.7/30.9) \approx 74$ degrees, which is in close agreement

with what is observed. The exact reason why this point is different for KEMAR is not completely understood, but it is probably a combination of the effect of the neck and a mismatch between the surface orientation of the ellipsoid and the top of the torso in the vicinity of the neck.

Although a small zone directly below KEMAR could not be measured, the ipsilateral torso shadow that is visible matches the model well. The contralateral torso shadow is more complex for KEMAR than it is for the model, and—not surprisingly—there is no sign of a torso bright spot. The new interference patterns that appear for $\delta > 240$ degrees are undoubtedly due to scattering from the truncated lower torso, and are not expected to be encountered in human HRTFs. However, the general presence of strong torso shadow on the contralateral side for both KEMAR and the snowman model presents a strong contrast with the behavior of the spherical head model alone (cf. Fig. 4).

The boundary-element method was used to compute the frontal-plane HRTF for the ellipsoidal-torso model, and the results are shown in Fig. 12(c). As might be expected, the results are closer to those for the spherical-torso model than to KEMAR. There are two primary effects of stretching the spherical torso into an ellipsoid: (a) introducing angular anisotropy, and (b) breaking up the perfect symmetry that leads to the torso bright spot. Because there is little difference between the height and width of the ellipsoid, the first effect is not very prominent in a frontal-plane response. The second effect can be seen, however, in that the interference patterns in the contralateral torso shadow region are not as sharply defined for the ellipsoidal torso as for the spherical torso. Indeed, in this region, the polar plot for the ellipsoidal torso seems to be intermediate between that for the spherical torso and that for KEMAR.

2. Horizontal plane

The torso has relatively little effect on the HRTF in the horizontal plane. The HRTF for the isolated spherical head looks exactly the same in the horizontal plane as it does in the frontal plane (see Fig. 4). The horizontal-plane results for KEMAR, the spherical-torso snowman, and the ellipsoidal-torso snowman are shown in Fig. 13. The effects of torso reflection are more irregular in the KEMAR data than in the models. The difference between torso width and torso depth for the ellipsoidal-torso snowman introduces an anisotropy that can be seen as a flattening of the torso reflection arches in Fig. 13(c), but the effect is small. Perhaps the greatest difference between KEMAR and the models is the relative weakness of the head bright spot in KEMAR. As was explained in the discussion of frontal-plane results, this can be attributed to the fact that KEMAR's ears are displaced below the center of the head.

3. Median plane

In the median plane, the HRTF for the isolated spherical head is uninteresting, because there is no variation with the elevation angle, ϕ . The median-plane results for KEMAR, the spherical-torso snowman, and the ellipsoidal-torso snowman are shown in Fig. 14. Above the horizontal plane, all

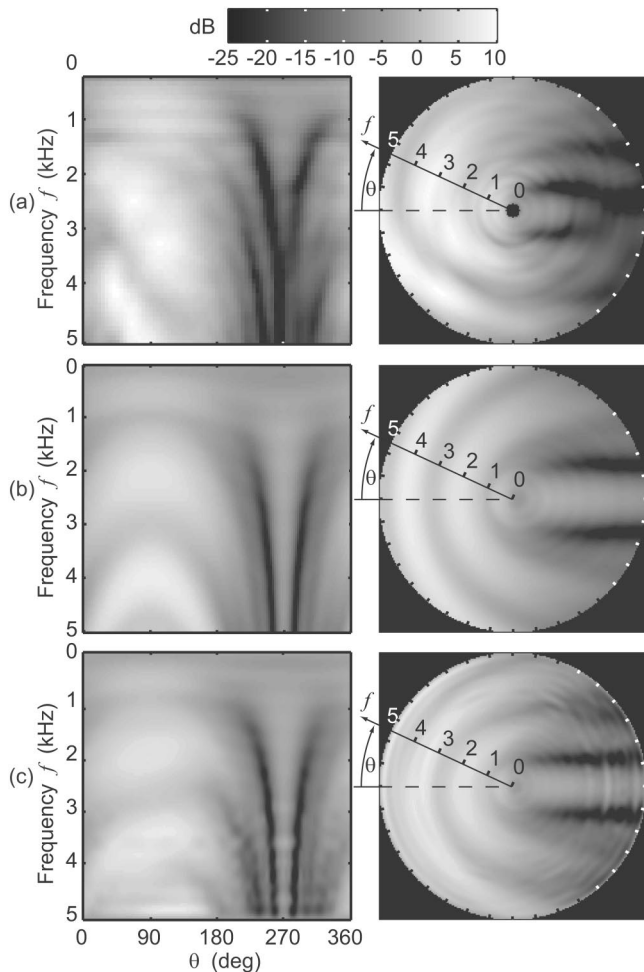


FIG. 13. Measured and computed horizontal-plane HRTFs. (a) KEMAR with the pinnae removed. (b) The spherical-torso snowman approximation. (c) The ellipsoidal-torso snowman approximation.

three results are very similar, with the slightly flatter torso reflection arches for the ellipsoidal model providing a slightly better fit to the KEMAR data. At lower elevations, the differences between the spherical torso and the ellipsoidal torso models are greater than might be expected. In particular, the ellipsoidal-torso model exhibits two moderately deep interference notches, with a bright spot at the point directly below, while the spherical-torso model exhibits only a general shadowing. The low-elevation shadowing is relatively weak compared to what is observed in the frontal plane. This is explained by the fact that head and torso shadow both occur in the frontal plane, while only torso shadowing occurs in the median plane. The ellipsoidal-torso results, although less easily explained, seem to be closer to what is observed in the KEMAR data. However, once again KEMAR's sharply truncated torso introduces some complex interference patterns at low elevations that are not really relevant to human HRTFs.

VI. DISCUSSION AND CONCLUSIONS

The results of this investigation show that simple models of the head and torso can explain the major features found in the pinnaless KEMAR HRTF for distant sources. The addi-

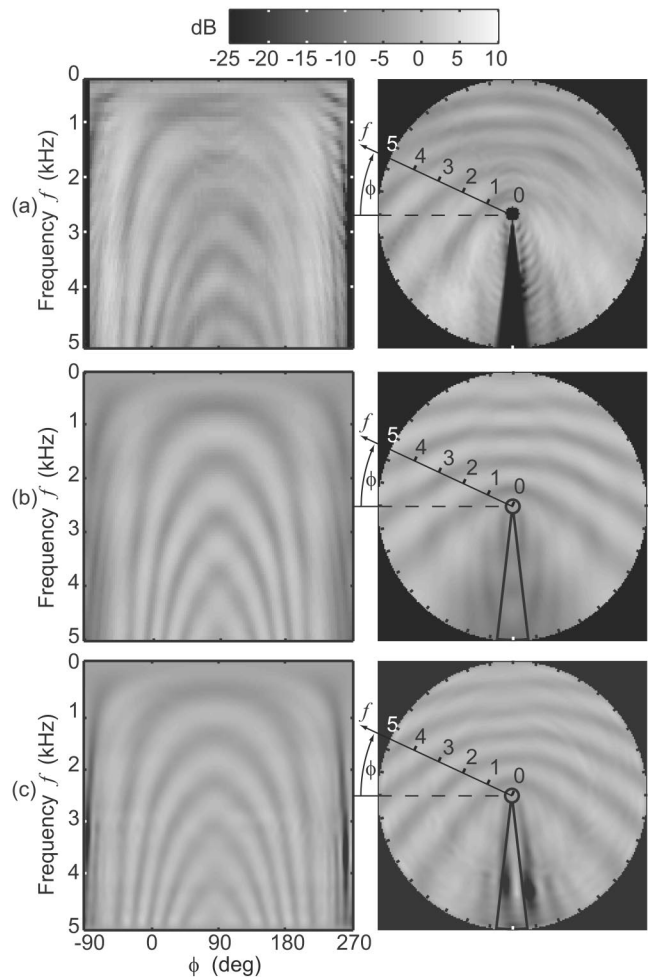


FIG. 14. Measured and computed median-plane HRTFs. (a) KEMAR with the pinnae removed. (b) The spherical-torso snowman approximation. (c) The ellipsoidal-torso snowman approximation.

tion of either a spherical or an ellipsoidal torso to the standard spherical-head model changes the HRTF significantly, bringing the behavior of the model substantially closer to that of a pinnaless KEMAR. In particular, the torso introduces reflections when the source is at high elevation, and shadow when the source is at low elevations. These phenomena are not seen with the isolated spherical head model. More elaborate geometric models should produce even better approximations, but the simplicity of the spherical-torso snowman model facilitates systematic studies.

Our previous work showed that torso reflections provide relatively weak but genuine elevation cues, particularly when the source is away from the median plane (Algazi *et al.*, 2001a). The elevation cues provided by torso shadow have not been subjected to systematic psychoacoustic tests, but informal listening experiments using HRIRs generated by the models indicate that torso shadow does increase the sense that a virtual auditory source is at a low elevation. Because these elevation cues occur at low frequencies, they are particularly important for sources such as footsteps or thunder that have little high-frequency energy.

Simple head and torso models have uses in addition to providing insight into HRTF behavior. As part of a structural model of the HRTF, they can be customized to particular

individuals by matching their parameters to anthropometry (Algazi *et al.*, 2001d). In addition, they can be used to compensate acoustically measured HRTFs, filling in important low-frequency information that is difficult to measure experimentally.

There are several open issues that need further investigation. Perhaps most important, it is clear that the detailed behavior of human HRTFs at low elevations is sensitive to posture. However, relatively little is known as to how sensitive people are to the acoustic changes that accompany postural changes. Another open question concerns the effect of the neck on the HRTF and its perceptual importance. Finally, the sensitivity of the results to displacements of the head relative to the torso and displacements of the ears relative to the center of the head need to be better understood, particularly because these displacements have been found to be quite substantial for human subjects (Algazi *et al.*, 2001c). Although the geometric modeling approach cannot answer the perceptual questions, it offers an attractive way to investigate the effects of different components of the body on HRTFs.

ACKNOWLEDGMENTS

The authors would like to thank Dennis Thompson for his help with much of the experimental work. Support was provided by the National Science Foundation under Grant Nos. IIS-00-97256 and ITR-00-86075. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the view of the National Science Foundation.

¹To be more precise, the ITD and ILD for the spherical-head model are constant on a cone of confusion. In the interaural-polar coordinate system, a cone of confusion is a cone of constant interaural-polar azimuth θ_{ip} (Algazi *et al.*, 2001a). Thus, the ITD and ILD uniquely determine the interaural-polar azimuth, but provide no information about the interaural-polar elevation ϕ_{ip} . A simple coordinate conversion shows that the interaural-polar azimuth is related to the standard vertical-polar azimuth θ_{vp} and vertical-polar elevation ϕ_{vp} by $\theta_{ip} = \sin^{-1}[\sin \theta_{vp} \cos \phi_{vp}]$. Thus, strictly speaking, the ITD and ILD do not determine either θ_{vp} or ϕ_{ip} , but instead determine the product $\sin \theta_{vp} \cos \phi_{vp}$. However, the main point is that the spherical-head model does not provide sufficient cues to determine both the azimuth and the elevation.

²Butler (1975) gives an excellent survey of the extensive research on pinna cues. In addition to the static acoustic cues that are captured by the HRTF, there are also dynamic elevation cues generated by head motion (Perrett and Noble, 1997). This paper is concerned with static cues only.

³This interpretation is presented by Duda and Martens (1998). The impulse responses shown in that paper were computed for a 22.05-kHz bandwidth, and their higher temporal resolution makes it easier to see the two pulses that appear in the vicinity of the bright spot.

⁴These control points were drawn from earlier work on creating an HRTF database for human subjects in which operational definitions for 17 measurements of the head and torso, including the head height and neck height, were provided (Algazi *et al.*, 2001c). This unambiguously specified the locations of four of the torso control points: the left and right arm points and the front and back torso points. The definition for the top of the torso used with the HRTF database gave a point that was too low for the current study. For this paper, the top of the torso was located by going down from the center of the bounding box for the head by the amount of one-half the head height plus one third of the neck height. Although this procedure is arbitrary, it is well defined and can be applied to subjects other than KE-MAR.

- Algazi, V. R., Avendano, C., and Duda, R. O. (2001a). "Elevation localization and head-related transfer function analysis at low frequencies," *J. Acoust. Soc. Am.* **109**, 1110–1122.
- Algazi, V. R., Avendano, C., and Duda, R. O. (2001b). "Estimation of a spherical-head model from anthropometry," *J. Audio Eng. Soc.* **49**(6), 472–478.
- Algazi, V. R., Duda, R. O., Thompson, D. M., and Avendano, C. (2001c). "The CIPIC HRTF database," in *WASPAA'01* (Proc. 2001 IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics), New Paltz, New York, pp. 99–102.
- Algazi, V. R., Duda, R. O., Morrison, R. P., and Thompson, D. M. (2001d). "Structural composition and decomposition of HRTFs," in *WASPAA'01* (Proc. 2001 IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics), New Paltz, New York, pp. 103–106.
- Batteau, D. W. (1967). "The role of the pinna in human localization," *Proc. R. Soc. London, Ser. B* **168**, 158–180.
- Blauert, J. P. (1997). *Spatial Hearing* (revised edition) (MIT, Cambridge, MA), pp. 50–93, 373–374.
- Brungart, D. S., and Rabinowitz, W. R. (1999). "Auditory localization of nearby sources. Head-related transfer functions," *J. Acoust. Soc. Am.* **106**, 1465–1479.
- Burkhardt, M. D., and Sachs, R. M. (1975). "Anthropometric manikin for acoustic research," *J. Acoust. Soc. Am.* **58**, 214–222.
- Butler, R. A. (1975). "The influence of the external and middle ear on auditory discriminations," in *Handbook of Sensory Physiology*, edited by W. D. Keidel and W. D. Neff (Springer Verlag, New York), pp. 247–260.
- Carlile, S. (ed.) (1996). *Virtual Auditory Space: Generation and Applications* (Landes, Austin, TX), pp. 38–45.
- Ciskowski, R. D., and Brebbia, C. A. (eds.) (1991). *Boundary Element Methods in Acoustics* (Computational Mechanics Publications, Southampton).
- Duda, R. O., and Martens, W. L. (1998). "Range dependence of the response of a spherical head model," *J. Acoust. Soc. Am.* **104**, 3048–3058.
- Duda, R. O., Avendano, C., and Algazi, V. R. (1999). "An adaptable ellipsoidal head model for the interaural time difference," in *ICASSP'99* (Proc. IEEE International Conference in Acoustics Speech and Signal Processing), Vol. II, pp. 965–968.
- Gardner, M. B., and Gardner, R. S. (1973). "Problem of localization in the median plane: Effect of pinna cavity occlusion," *J. Acoust. Soc. Am.* **53**, 400–408.
- Genuit, K., and Platte, H. J. (1981). "Untersuchungen zur Realisation einer richtungsgetreuen Übertragung mit elektroakustischen Mitteln (Investigations on the implementation of directionally faithful transmission by electroacoustical means)," *Fortschritte der Akustik, FASE/DAGA '81* Berlin (VDE-Verlag, Berlin), pp. 629–632.
- Gumerov, N. A., and Duraiswami, R. (2001a). "Fast, exact, and stable computation of multipole translation and rotation coefficients for the 3-D Helmholtz equation," UMIACS Technical Report TR-2001-44, University of Maryland, College Park, MD (<http://www.umiacs.umd.edu/~ramani/pubs/multipole.pdf>).
- Gumerov, N. A., and Duraiswami, R. (2001b). "Multiple scattering from N spheres using multipole reexpansion," UMIACS Technical Report TR-2001-72, University of Maryland, College Park, MD (<http://www.umiacs.umd.edu/~ramani/pubs/multisphere.pdf>).
- Gumerov, N., Duraiswami, R., and Tang, Z. (2002). "Numerical study of the influence of the torso on the HRTF," in *ICASSP'02* (Proc. IEEE Int. Conf. Acoustics Speech and Signal Processing).
- Hartley, R. V. L., and Fry, T. C. (1921). "The binaural localization of pure tones," *Phys. Rev.* **18**, 431–442.
- Kahana, Y., and Nelson, P. A. (2000). "Spatial acoustic mode shapes of the human pinna," preprint 5218, *109th Convention Audio Engineering Society*, Los Angeles, CA.
- Kahana, Y., Nelson, P. A., Petyt, M., and Choi, S. (1999). "Numerical modeling of the transfer functions of a dummy-head and of the external ear," Paper s52654, *AES 16th International Conference on Spatial Sound Reproduction*, Rovaniemi, Finland.
- Katz, B. F. G. (2001). "Boundary element method calculation of individual head-related transfer function. I. Rigid model calculation," *J. Acoust. Soc. Am.* **110**, 2440–2448.
- Kuhn, G. F. (1977). "Model for the interaural time differences in the azimuthal plane," *J. Acoust. Soc. Am.* **62**, 157–167.
- Kuhn, G. F. (1987). "Physical acoustics and measurements pertaining to directional hearing," in *Directional Hearing*, edited by W. A. Yost and G. Gourevitch (Springer Verlag, New York), pp. 3–25.

- Kuhn, G. F., and Guernsey, R. M. (1983). "Sound pressure distribution about the human head and torso," *J. Acoust. Soc. Am.* **73**, 95–105.
- Lopez-Poveda, E. A., and Meddis, R. (1996). "A physical model of sound diffraction and reflections in the human concha," *J. Acoust. Soc. Am.* **100**, 3248–3259.
- Mehrgardt, S., and Mellert, V. (1977). "Transformation characteristics of the external human ear," *J. Acoust. Soc. Am.* **61**, 1567–1576.
- Morse, P. M., and Ingard, K. U. (1968). *Theoretical Acoustics* (Princeton U. P., Princeton, NJ), pp. 418–441.
- Oppenheim, A. V., and Schaffer, R. W. (1969). *Discrete-Time Signal Processing* (Prentice Hall, Englewood Cliffs, NJ), pp. 447–450.
- Perrett, S., and Noble, W. (1997). "The effect of head rotations on vertical plane sound localization," *J. Acoust. Soc. Am.* **102**, 2325–2332.
- Shaw, E. A. G. (1974). "The external ear," in *Handbook of Sensory Physiology*, Vol. V/I: Auditory System, edited by W. D. Keidel and W. D. Neff (Springer Verlag, New York), pp. 455–490.
- Shaw, E. A. G. (1997). "Acoustical features of the human external ear," in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. H. Gilkey and T. R. Anderson (Lawrence Erlbaum, Mahwah, NJ), pp. 25–47.
- Strutt, J. W. (Lord Rayleigh) (1907). "On our perception of sound direction," *Philos. Mag.* **13**, 214–232.
- Waterman, P. C., and Truell, R. (1961). "Multiple scattering of waves," *J. Math. Phys.* **2**, 512–537.
- Wightman, F. L., and Kistler, D. J. (1989). "Headphone simulation of free-field listening. I: Stimulus synthesis," *J. Acoust. Soc. Am.* **85**, 858–867.
- Wightman, F. L., and Kistler, D. J. (1997). "Factors effecting the relative salience of sound localization cues," in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. H. Gilkey and T. R. Anderson (Lawrence Erlbaum, Mahwah, NJ), pp. 1–23.
- Wright, D., Hebrank, J. H., and Wilson, B. (1974). "Pinna reflections as cues for localization," *J. Acoust. Soc. Am.* **56**, 957–962.
- Zhou, B., Green, D. M., and Middlebrooks, J. C. (1992). "Characterization of external ear impulse responses using Golay codes," *J. Acoust. Soc. Am.* **92**, 1169–1171.