On High-Resolution Head-Related Transfer Function Measurements: An Efficient Sampling Scheme

Wen Zhang, *Member, IEEE*, Mengqiu Zhang, *Student Member, IEEE*, Rodney A. Kennedy, *Fellow, IEEE*, and Thushara D. Abhayapala, *Senior Member, IEEE*

***Abstract*—This paper deals with two important questions associated with HRTF measurement: 1) "what is the required angular resolution?," and 2) "what is the most suitable sampling scheme?." The paper shows that a well-defined finite number of spherical harmonics can capture the head-related transfer function (HRTF) spatial variations in sufficient detail, which is defined as the HRTF spatial dimensionality. For the 20-kHz audible frequency range, the value of the dimensionality means a high-directional resolution HRTF measurement is required. Considering such a high-resolution measurement, a number of sampling criteria have been identified from both mechanical setup and data processing aspects. Different sampling candidates are then compared to demonstrate that the best method which satisfies all requirements is the class termed as IGLOO. A fast spherical harmonic transform algorithm based on the IGLOO scheme is developed to accelerate the high-resolution data analysis. The proposed method is validated through simulation and experimental data acquired from a KEMAR mannequin.**

***Index Terms*—Angular resolution, head-related transfer function (HRTF) measurement, sampling scheme, spatial dimensionality.**

I. Introduction

The binaural synthesis of 3-D spatial sound is based on a stored database of head-related transfer functions (HRTFs), which describe the ratio of the sound from a point in space developed at each of the listener's ears to that presented at the center of the listener's head when the listener is absent [1]. This acoustic function contains the listening cues used by human hearing mechanism for decoding spatial information encoded in the binaural signals. Therefore, the perception of a sound source in three dimensions can be realistically synthesized by filtering the source signal with the HRTFs of the desired direction, room reverberations and effects of head movement [2]. The resulting binaural signals are presented to the listener using two playback channels, typically a pair of headphones [3].

Manuscript received November 10, 2010; revised March 11, 2011; accepted June 27, 2011. Date of publication July 18, 2011; date of current version December 14, 2011. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Lauri Savioja.

W. Zhang is with CSIRO, Process Science and Engineering, Sydney, NSW 2232 Australia. (e-mail: w.zhang@csiro.au).

M. Zhang, R. A. Kennedy, and T. D. Abhayapala are with the Research School of Engineering, College of Engineering and Computer Science, The Australian National University, Canberra, ACT 0200 Australia. (e-mail: karan.zhang@cecs.anu.edu.au; rodney.kennedy@anu.edu.au; thushara.abhayapala@anu.edu.au).

Digital Object Identifier 10.1109/TASL.2011.2162404

Many research groups have tried to empirically measure the HRTF on human subjects [3]–[8] or a KEMAR mannequin [9]–[11]. In such measurements, the HRTFs are usually obtained on a sphere of constant radius for a predefined set of elevation and azimuth angles. It is clear that the success of binaural synthesis and the quality of spatial sound generated depend strongly on the physical aspects of measurement process. Since HRTFs cannot be measured in all directions, it is important to determine the minimum spatial resolution required to achieve a spatial audio rendition of sufficient fidelity.

Ajdler *et al.* [12], [13] developed an angular theorem for reconstructing sound field along a circle by using plenacoustic functions and further applied the theorem to the HRTF sampling and interpolation in the horizontal plane. Their results showed that an angular spacing of 5° or less is necessary to reconstruct the data up to a bandwidth of 22 kHz. Based on the signal processing theory, Zhong and Xie [14] studied maximal azimuthal resolution, or equivalently the minimum number of azimuthal measurements (MNAM), by applying azimuthal Fourier analysis separately on HRTFs at each elevation plane. MNAM was determined by evaluating the contribution of the decomposed azimuthal harmonics, which is found to be a function of frequency and elevation. For a given elevation plane, it increases with increasing frequency, i.e., from 5 measurements at the lowest frequency to more than 60 measurements at 20 kHz in the horizontal plane; while for a fixed frequency it decreases as the elevation deviates from the horizontal plane.

Other than azimuthal resolution analysis, Minnaar *et al.* [15] investigated the directional resolution of HRTF measurement in the horizontal, frontal, and median plane, which is based on the previous work of high-resolution HRTF measurements [11]. The resolution of 8° over the sphere around the head was deemed to be enough in HRTF measurement without introducing audible interpolation errors in binaural synthesis, but this resolution threshold was largely based on listening experiments.

The present work answers the questions of what is the required angular resolution and what is the most suitable sampling scheme/configuration for HRTF measurement, taking into account the two dimensional angular direction and the wide audio frequency range. Considering that the HRTF is a function of the source direction and the HRTF measurement is in fact an example of *data measured on the sphere*, the questions are naturally answered using the complex Fourier transform on the sphere, which are based on the *spherical harmonics*. The spherical harmonics are orthonormal basis functions on the unit sphere and the HRTF (independent of range) can be expanded through this function [16]–[18]. As the spherical harmonics provide a natural continuous representation in the angular domain, the degrees of freedom in the HRTF spherical harmonic expansion lead to a straightforward solution to the required angular resolution in HRTF measurement, which was elaborated in our previous work [19]. The required angular resolution combined with other practical criteria are used to find the most suitable HRTF sampling scheme.

Three main contributions of this paper are summarized as follows.

- In Section II, we review the spherical harmonic analysis of the HRTF and demonstrate that a finite number of spherical harmonics are sufficient to accurately represent HRTFs corresponding to all directions. We call this number a *spatial dimensionality* [19] of the HRTF, which is also the least required number of spatial samples for HRTF measurement.
- We next consider how the HRTF data should be sampled on a sphere. Section III identifies a list of requirements for the determination of the HRTF measurement grid and the structure of discrete HRTF data. We next compare different sphere sampling methods and identify the one, called IGLOO from the astrophysics literature [20], which best satisfies all of these requirements. Section IV explains how to measure the HRTF over sphere according to the IGLOO scheme.
- In Section V, according to IGLOO, we proposed a fast spherical harmonic transform algorithm to evaluate the HRTF spherical harmonics coefficients from the sampled high resolution data. From these spherical harmonic coefficients, one can synthesize the HRTF in any direction.

II. Spherical Harmonic Analysis and Spatial Dimensionality of HRTF

The spherical harmonics [21], as a function of elevation θ and azimuth ϕ, are characterized by two indices, degree n and order m:

$$Y_n^m(\theta,\phi) \triangleq \sqrt{\frac{2n+1}{4\pi}\frac{(n-|m|)!}{(n+|m|)!}}\,P_n^{|m|}(\cos\theta)\,e^{im\phi}$$
$$n=0,1,2,\ldots,\quad m=-n,\ldots,n \qquad (1)$$

where $P_n^{|m|}(\cdot)$ are the associated Legendre functions. The spherical harmonics form a complete orthonormal basis[1] with respect to the natural inner product on the sphere $\mathbb{S}^2$

$$\int_{\mathbb{S}^2} Y_n^m(\theta,\phi)\overline{Y_{n'}^{m'}(\theta,\phi)}\,ds = \delta_{nn'}\delta_{mm'} \qquad (2)$$

where $ds = \sin(\theta)\,d\theta\,d\phi$ and $\overline{(\cdot)}$ stands for the complex conjugate.

At a fixed radius, the HRTF expansion in spherical harmonics [16], [17] is written as

$$H(\theta,\phi,k) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n}\beta_n^m(k)Y_n^m(\theta,\phi) \qquad (3)$$

[1]A notation of complete basis means that any finite energy function on the sphere can be represented over such basis.

at wavenumber $k=2\pi f/c$, where f is the frequency and c is the speed of sound propagation. The complex coefficients $\beta_n^m(k)$ in (3) are given by the *spherical harmonic transform*

$$\beta_n^m(k) = \int_{\mathbb{S}^2} H(\theta,\phi,k)\overline{Y_n^m(\theta,\phi)}\,ds. \qquad (4)$$

However, in practice, we usually numerically approximate the above integral

$$\beta_n^m(k) = \sum_{i=1}^{I} H(\theta_i,\phi_i,k)\overline{Y_n^m(\theta_i,\phi_i)}\sin\theta_i + \varepsilon \qquad (5)$$

where i is the index of a set of sample points, I is the total points over the sphere, and ε is called the quadrature error. The key concern in designing a sampling scheme in spatial audio based on spherical harmonics is to minimize this quadrature error [22], [23].

It has been demonstrated that the HRTF decomposition in (3) can be well approximated by choosing a sufficiently large truncation degree depending on the wavenumber, $N(k)$, *viz.*,

$$\widetilde{H}(\theta,\phi,k) \cong \sum_{n=0}^{N(k)}\sum_{m=-n}^{n}\beta_n^m(k)Y_n^m(\theta,\phi) \qquad (6)$$

based on the fact that high-order components $\beta_n^m(k)$ represent the higher frequencies in the response and make very small contributions to the HRTF [17], [18]. Therefore, the reconstruction error includes the quadrature error and truncation error.

According to [17], [19], the required number $(N(k)+1)^2$ of spherical harmonics to represent the HRTF spatial variations should be determined by the wave number k and the scattering object size s (in our case, the human head). Hence, an approximate truncation number is given by an average inter-element spacing of half a wavelength over the surface area of the sphere, which produces a required number of $N(k)=\lceil ks\rceil$ [17], [19]. A more accurate value is obtained from the fact that the HRTF spherical harmonic coefficients $\beta_n^m(k)$ can be represented by the spherical Bessel functions [19], which have a high-pass character, the bounds on which determines the number of terms in HRTF spherical harmonic expansion [24], [25]. This yields a slightly larger value [19]

$$N(k) = \left\lceil \frac{eks}{2} \right\rceil \qquad (7)$$

where $e=\exp(1)\approx 2.7183$.

In summary, the analysis above demonstrates that the HRTF is essentially mode-limited (or has a finite spherical harmonic decomposition) [26]. We only need a set of $(N(k)+1)^2$ coefficients $\{\beta_n^m(k)\}$ to represent HRTFs corresponding to all directions, where $N(k)$ is determined by (7). The minimum number of coefficients required in the HRTF spherical harmonic decomposition is essentially the least number of HRTF measurement required on the sphere $\mathbb{S}^2$, which, for frequency f, is given by

$$M \triangleq \left\lceil \left(\frac{e\pi s f}{c}+1\right)^2 \right\rceil. \qquad (8)$$

Therefore, to synthesize the HRTF over the entire audible frequency range ([0.02, 20] kHz), at least 2209 HRTF measurements over the sphere are required based on the assumption that the typical head radius is $s = 9$ cm.

III. Requirements for Design of HRTF Measurement

As stated in Section II, the spatial components of HRTF can be expanded using spherical harmonics. Further, we can factorize the spherical harmonics into the azimuthal part and the elevation part by

$$Y_n^m(\theta,\phi) = \mathcal{P}_n^m(\cos\theta)e^{im\phi} \tag{9}$$

where

$$\mathcal{P}_n^m(\cos\theta) \triangleq \sqrt{\frac{2n+1}{4\pi}\frac{(n-|m|)!}{(n+|m|)!}}\,P_n^{|m|}(\cos\theta). \tag{10}$$

Then the HRTF at any arbitrary position (θ,ϕ) and wave number k can be synthesized by

$$\widetilde{H}(\theta,\phi,k) = \sum_{n=0}^{N(k)}\sum_{m=-n}^{n}\beta_n^m(k)\mathcal{P}_n^{|m|}(\cos\theta)\,e^{im\phi} \tag{11}$$

as long as a set of $\beta_n^m(k)$ is known.

To evaluate $\beta_n^m(k)$ with k corresponding to 20 kHz, where $N(k) = 46$, at least 2209 HRTF measurements over the sphere are required. Generally, the measurements are performed on the sphere at positions (θ_q,ϕ_v), $q = 1,\ldots,Q$, $v = 1,\ldots,V$. Thus, given the least number of positions, what is the best choice for Q and V? In addition, a significant problem with the HRTF measurement is that measurements usually cannot be done at lower elevations due to mechanical limitations of measurement apparatus and ground reflections. Clearly, how the high-resolution HRTF data is to be measured over sphere is an important question; the right choice of measurement grid will both significantly improve data acquisition and accelerate data analysis. We next identify three important considerations.

First, in Section II, we have shown that the HRTF measurements performed on a sphere is required to evaluate a number of spherical harmonic coefficients. The spherical harmonic transform becomes prohibitively computationally demanding if the sampling on the sphere and the corresponding structure of the discrete data set are not designed carefully.

Second, the HRTF is a mode-limited function; however, the measurement process will inevitably induce noise on the data set due to the apparatus. This noise is assumed to be random, white and with a spatial bandwidth significantly exceeding that of the HRTF data.

The last but not least, the consideration of the human subjects taking part in the measurement. Taking measurements at all angles typically requires rotating either a sound source or listener or both. It is desirable to keep the rotations to a minimum number of steps.

Taking these considerations into account, we propose the following list of requirements for the design of the HRTF measurement and the structure of the discrete HRTF data.

1) *Least number of measurements*: For efficiency the number of points on the proposed sampling grid must at least equal the required number of measurements (at least 2209 sample positions for the 20-kHz audible frequency range).
2) *Equal area division*: The proposed sampling grid should have nearly equal area division of the sphere, which makes all the measurements contribute nearly equally when used in the discrete spherical harmonic transform. Moreover, equal area sampling avoids over-sampling near the pole. In addition, with the property of nearly equal area division, the measured HRTF data need only to be sampled within a region rather than at a specific point, which provides more tolerance for the sampling strategy.
3) *Hierarchical structure of data*: The database measured on the proposed sampling grid should be structured such that a low spatial resolution data set suitable for low frequencies is imbedded in the high spatial resolution data set.
4) *Iso-longitude measurement setup*: The proposed sampling strategy should measure all elevations at each azimuth in order that the listener and apparatus experience the least rotations.

Many alternative sampling distributions on the sphere have been suggested for the discretization and analysis of functions [20], [23], [27]–[29]. The IGLOO [20]/HEALPix [29] schemes were originally designed in the astrophysics literature for sampling the sky maps. The specific criteria are to achieve uniform sphere sampling and meanwhile minimize the distortions by taking sample region area and shape into account, which provides a new idea for HRTF sampling based on almost equal area.

In [23], Li and Duraiswami designed a flexible and optimal spherical microphone array for beamforming using the Fliege grid. The grid requires only $(N+1)^2$ points for Nth-order harmonics, the exact minimum number of points required in HRTF measurement. However, HRTF measurements are usually unavailable at low elevations, which means at least $(N+1)^2$ points are required on high elevations over the sphere rather than on the whole sphere. More importantly, it does not meet the requirements 2, 3, and 4 for the HRTF measurement.[2] We next compare the four methods (Equiangular grid, Gauss-Legendre sampling, HEALPix, IGLOO) to assess which one best meets the above requirements.

- **Equiangular grid** simply uses equal divisions in latitude and longitude (which is also known as Equidistance Cylindrical Projection in sky pixelizations [31]). For a mode-limited function, such as the HRTF of the highest modes up to degree $N(k)$, the equiangular method needs $4(N(k)+1)^2$ samples to approximate the spherical harmonic transform [27].
  The equiangular grid satisfies requirements 3 and 4 but fails with requirements 1 and 2. The biggest drawback of the equiangular grid is that the region near the pole is overly densely sampled.
- **Gauss–Legendre sampling** is based on optimal methods of evaluating definite integrals in the spherical harmonic transform [16], [28]. For mode-limited functions up to degree $N(k)$, in the θ direction, $N(k)+1$ measurement points are taken as the roots of the Legendre polynomial of

[2]The sample positions obtained from an optimization shows no obvious regular structure [30].

degree $N(k)$, i.e., $P_{N(k)}(\cos\theta_q) = 0$, $q = 0, \ldots, N(k)$. The corresponding weights w_q^{GL} are determined using the Gaussian–Legendre method. As for the ϕ variable, $2N(k)+1$ points are equally spaced in $[0, 2\pi)$, that is, $\phi_v = 2\pi v/(2N(k)+1)$, $v = 0, 1, \ldots, 2N(k)$. So in total, the Gauss–Legendre method requires $2N(k)^2+3N(k)+1$ samples.

The Gauss–Legendre sampling grid provides an accurate way to approximate the spherical harmonic transform. However, it does not meet requirements 2 and 3. The shape of sampling region is not regular and the sampling grids for different resolution data are totally different, which causes trouble for data acquisition and complicate data analysis. In common with the equiangular grid, the areas near the poles are also heavily oversampled.

- **HEALPix** stands for hierarchical equal area isolatitude pixelization, which is a pixelization scheme based on a rhombic dodecahedron [29]. This scheme samples the surface of a sphere into equal area subregions, which means all samples have same weights in approximating the spherical harmonic transform. HEALPix is also hierarchical: the base resolution has twelve subregions in three elevations (two around the pole and one around the equator) and the high resolution data are constructed by subdividing each base region into four. Therefore, a HEALPix grid always has $N_{\mathrm{pix}} = 12 \times 4^p$, $p = 0, 1, \ldots$, subregions of the same area. Thus, HEALPix grid meets requirements 2, 3. However, to perform fast spherical harmonic transform to evaluate the HRTF spherical harmonic coefficients up to the desired degree, an excess number of samples, 12 288, relative to the least number, 2209, are required[3]. The other problem of HEALPix for HRTF measurement is the azimuthal positions change with elevation, that is the sampling points are not iso-longitude.
- **IGLOO** is a general sampling grid developed from the equiangular method [20]. Analysis shown in [20] demonstrated that the 12 base regions configured as 3: 6: 3 is the best compromise between distortion and hierarchy (simplicity inherent in having few base regions). Based on this configuration, only 3072 samples are required to synthesis the HRTF over the entire audible frequency range.[4] The sampling edges are defined along constant lines of elevation θ and azimuth ϕ, thus allowing for simplified determination of the spherical harmonics. Different from the equiangular grid, the division of sphere in IGLOO is closer to the square and nearly equal area; thus, the number of samples at each elevation must decrease near approaching poles, meaning that the IGLOO sampling is less distorted. In addition, IGLOO also has built-in hierarchy. Since IGLOO scheme has almost equal area divisions, it does not satisfy the iso-longitude criterion for the whole sphere (but does in each sub-region).

Table I summarizes the characteristics of the above four sampling methods.

TABLE I
COMPARISON OF DIFFERENT METHODS FOR HRTF SAMPLING OVER SPHERE

	Equiangular	Gauss-Legendre	HEALPix	IGLOO
Number of samples (for 20 kHz bandwidth)	8836	4371	12288	3072
Equal area division	No	No	Yes	Almost
Hierarchical	Yes	No	Yes	Yes
Iso-longitude	Yes	Yes	No	No

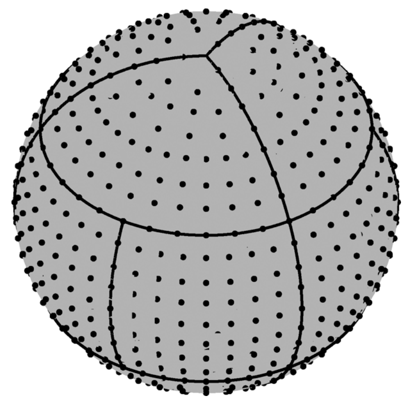


Fig. 1. Picture of the 3:6:3 equal area division, which divides the sphere into 12 base regions, three at either cap and six $60^\circ \times 60^\circ$ equatorial regions. Here, each base region is sampled with 64 points.

IV. Meeting the Requirements: The IGLOO Grid

All the requirements introduced in Section III are best satisfied by IGLOO according to Table I.

A. IGLOO Sampling Scheme

In this scheme, one first divides the sphere into a number of base regions subject to the minimum distortion criteria.[5] There is a tradeoff between the number of base regions and the sample distortion: the measurements can be made less distorted by increasing the number of base regions at the expense of increased scheme complexity.

The best compromise has been demonstrated to be 12 base regions configured as 3:6:3 divisions on the sphere [20], as shown in Fig. 1. The sphere is firstly divided into three layers, two cap layers near each pole and one equatorial layer. The cap layers are further divided into three subregions and the equatorial layer is divided into six subregions. One can create a finer division of the sphere with the required resolution by dividing each subregion into four, configured as 12:24:12.

Mathematically, the resolution of the grid can be expressed by a parameter $M_d = 0, 1, 2, 3, \ldots$, which is the subdivision level to higher resolution pixels. Then, each base region has

[3]The formula for determining the number of samples in HEALPix scheme is $N_{\mathrm{pix}} = 12 \times 4^p$. Based on the discrete spherical harmonic transform algorithm, at least $2N(k)+1$ azimuthal samples are required over equatorial region to calculate the azimuth harmonics [(15)] to the desired degree ($N(k) = 46$ for the HRTF), and in HEALPix scheme, the numbers of azimuthal samplings over equatorial region is 4×2^p. Hence, $p = 5$. Then the required number of samples are 12 288.

[4]In IGLOO scheme, the total number of samples is also determined by $N_{\mathrm{pix}} = 12 \times 4^p$, but, the numbers of azimuthal samplings over equatorial region is 6×2^p. Hence, $p = 4$ and the required number of samples are 3072.

[5]In [20], the distortion of a sample region is defined to be the length of longest diameter divided by the ideal value $D = 4/\sqrt{N_{\mathrm{tot}}}$, where N_{tot} is the total number of samplings. The principle is that the division of a sphere suppresses modes with wavelengths shorter than the sample region dimension. To minimize this effect, we should make the largest sample region diameter as small as possible. The ideal case would be to have circular sample regions where the largest diameter is D.

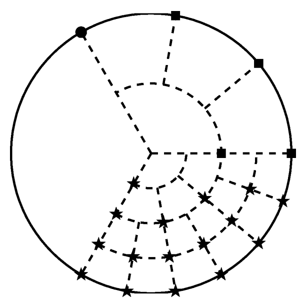


Fig. 2. Picture of the polar cap region in the igloo scheme, showing the subdivision with $M_d = 2$. The dot indicates the sampling position of the base region, four squares indicate the sampling positions with $M_d = 1$, and 16 stars indicate sampling positions with $M_d = 2$.

4^{M_d} pixels and the whole sphere has 12×4^{M_d} pixels in total. Fig. 1 shows an example of 3:6:3 equal area IGLOO scheme with $M_d = 3$, which means each base region is sampled with 64 points.

From the resolution parameter M_d, one can define the sample positions. The locations of the sample points are defined by (θ, ϕ), where $\theta \in [0°, 180°]$ from the north pole and $\phi \in [0°, 360°)$ counterclockwise in accordance with having a right hand coordinate system and 0° being the direct front direction. The divisions between the base polar caps and equatorial belt are $\theta = 60°$, 120°.

We firstly define the elevation sample positions. Since M_d is the level of subdivision of the base region, the samples are located on 2^{M_d} elevations in each layer with lines equally spaced in longitude. In total, there are 3×2^{M_d} elevations for the whole sphere. Then the sampled elevation positions are

$$\theta_q = q\,\frac{60}{2^{M_d}}, \quad q = 1, \ldots, Q \tag{12}$$

where q is the elevation sampling index and $Q = 3 \times 2^{M_d}$.

We next define the azimuth sample positions. To achieve almost equal sampling area, the number of pixels along each latitude must decrease as one approaches the poles.

As depicted in Fig. 2, the polar cap is initially divided into three equal wedges (base regions). Then, the higher resolution sampling scheme is formed by dividing each base region into four pieces. There are two cases: 1) the regions including the pole are divided into one polar wedge and three pieces surrounding it; and 2) the regions not including the pole are bisected by lines of constant θ and ϕ. The dot in Fig. 2 indicates the sampling position of the base region, four squares indicate the sampling positions with $M_d = 1$, and 16 stars indicate the sampling positions with $M_d = 2$. Sample positions over the north and south polar cap are symmetric with respect to the equator ($\theta = 90°$).

Then, the azimuthal samples in the two polar caps are structured as

$$\phi_{v^{(q)}} = v\,\frac{360}{V_q}, \quad v^{(q)} = 0, \ldots, V_q - 1 \tag{13}$$

where $v^{(q)}$ is the azimuth sampling index for each elevation and

$$V_q = \begin{cases} 3, & q = 1, 3 \times 2^{M_d} \\ 9 \times 2^{E(q)}, & 2 \le q \le 2^{M_d} \\ 9 \times 2^{E(M_d, q)}, & 2 \times 2^{M_d} + 1 \le q \le 3 \times 2^{M_d} - 1 \end{cases}$$

in which

$$E(q) = \lceil \log_2(q) \rceil - 1$$

and $$E(M_d, q) = \lceil \log_2(3 \times 2^{M_d} + 1 - q) \rceil - 1.$$

The equatorial layer initially has six base regions and the high-resolution sampling is obtained by bisecting each region into four. Given the resolution parameter M_d, the base region has 2^{M_d} pixels along each latitude. Then the azimuthal samples in equatorial layer are

$$\phi_{v^{(q)}} = v^{(q)}\,\frac{360}{V_q}, \quad v^{(q)} = 0, \ldots, V_q \tag{14}$$

where

$$V_q = 6 \times 2^{M_d}, \quad 2^{M_d} + 1 \le q \le 2 \times 2^{M_d}.$$

Combining (13) and (14), the azimutal samples over the whole sphere are

$$\phi_{v^{(q)}} = v^{(q)}\,\frac{360}{V_q}, \quad v^{(q)} = 0, \ldots, V_q - 1 \tag{15}$$

where

$$V_q = \begin{cases} 3, & q = 1, 3 \times 2^{M_d} \\ 9 \times 2^{E(q)}, & 2 \le q \le 2^{M_d} \\ 6 \times 2^{M_d}, & 2^{M_d} + 1 \le q \le 2 \times 2^{M_d} \\ 9 \times 2^{E(M_d, q)}, & 2 \times 2^{M_d} + 1 \le q \le 3 \times 2^{M_d} - 1. \end{cases}$$

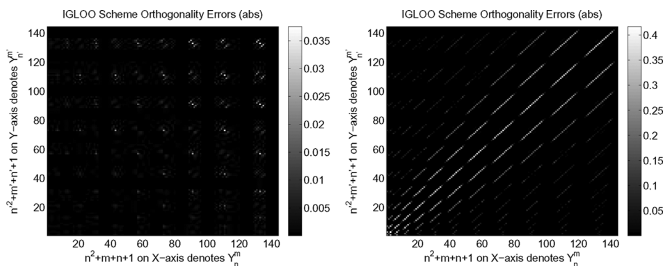
Based on this sampling arrangement, we perform a numerical integration of spherical harmonic pairs of various degrees (up to degree 11) for both the full-sphere and the grid with the bottom part cut out using the IGLOO scheme. Fig. 3 shows the orthogonality error, the difference between the theoretical value of the integral (1 when degree and order are equal and 0 in all other cases) and the numerical value, for the 192 sampling points at a pixel gray level. As expected, when the full-sphere data is available, the orthogonal errors, same as other quadrature methods, are quite small (less than 0.04); however, for the grid with bottom part cut out, larger errors appear. This means for HRTF data given its unavailability at low elevations, simply approximating spherical harmonic integral could not provide accurate spherical harmonic decomposition; instead we proposed a more efficient fast spherical harmonic transform in Section V, where regularized least-square fitting is incorporated to give more accurate estimates.

B. HRTF Sampling Arrangement

Generally, no HRTF measurements are made over the south polar cap because of mechanical interferences between measurement setup and subject's lower body. Applying the IGLOO scheme to the HRTF measurement, we actually divide the sphere into 9 base regions. From $9 \times 4^{M_d} \ge 2209$, we have

TABLE II
IGLOO SCHEME BASED HRTF SAMPLING FOR 20-kHz AUDIBLE BANDWIDTH

Layer	Elevation θ (degrees)	Azimuth Resolution (degrees)	Number of Azimuthal Samples V_q	Total number of Samples
North Cap	4	120.00	3	768
	7	40.00	9	
	11 and 15	20.00	18	
	19 to 30	10.00	36	
	34 to 60	5.00	72	
Equator	64 to 120	3.75	96	1536



(a) (b)

Fig. 3. Validity of the orthogonality relation $\int Y_n^m(s) Y_{n'}^{m'}(s) ds$ over the full-sphere (a) and the grid with the bottom part cut out (b) using the IGLOO scheme. The orthogonality error is the difference between the theoretical value of the integral (1 when degree and order are equal and 0 in all other cases) and the numerical value, for the 192 sampling points at a pixel gray level.

$M_d = 4$. Then the maximum elevation index Q is $2 \times 2^4 = 32$ and $9 \times 4^4 = 2304$ sampling positions are located at

$$(\theta_q, \phi_{v^{(q)}}) = \begin{cases} \theta_q = 3.75q, & q = 1, \ldots, 32 \\ \phi_{v^{(q)}} = v^{(q)} \frac{360}{V_q}, & v^{(q)} = 0, \ldots, V_q - 1 \end{cases} \tag{16}$$

where

$$V_q = \begin{cases} 3, & q = 1 \\ 9 \times 2^{(\lceil \log_2(q) \rceil - 1)}, & 2 \le q \le 16 \\ 96, & 17 \le q \le 32. \end{cases}$$

Table II shows the detail of the IGLOO-based HRTF sampling for 20-kHz audible bandwidth. We can see that the elevation resolution is $\Delta\theta = 3.75^\circ$. This is consistent with Blauert's study [32], which indicated that the localization blur for the changes of elevation angle in the median plane is about 4° for white noise. This means that the localization attributes of the sound moving less than 4° in elevation angle remain perceptually unchanged. Therefore, the elevation resolution of 3.75° is the reasonable choice to capture all the localization information regarding elevation angle. Considering the difficulty of setting sampling position with high accuracy to a fraction of a degree, we approximate the elevation sampling position to the nearest integer in degrees; thus, the sampling interval is changed from 3.75° to either 3° or 4° along the elevation.

According to (15), the azimuth sampling resolution in the equatorial layer is 3.75° and the resolution in the north cap layer increases from 5° to 120° as the elevation plane approaching the pole. Similarly, Zhong and Xie [14] also suggested the fine azimuth resolution near the horizontal plane and the coarse azimuth resolution in the polar area.

V. FAST SPHERICAL HARMONIC TRANSFORM

In this section, we will show one attractive feature of the IGLOO data structure, in which an exact discrete azimuthal symmetry at each elevation allows fast and accurate spherical harmonic transform. Such an ability to perform spherical harmonic transform quickly is highly desirable for the HRTF data analysis and reconstruction.

At each elevation, with the use of the orthogonality of the discrete exponential functions over circle, we get the *azimuth harmonics* [33]

$$a_m(\theta_q, k) = \sum_{n=|m|}^{N(k)} \beta_n^m(k) \mathcal{P}_n^{|m|}(cos\theta_q). \tag{17}$$

To estimate $\beta_n^m(k)$, we first calculate $\alpha_m(\theta_q, k)$.

Since there are V_q azimuth samplings at each elevation θ_q, we have

$$a_m(\theta_q, k) = \frac{1}{V_q} \sum_{v^{(q)}=1}^{V_q} H(\theta_q, \phi_{v^{(q)}}, k) e^{-im\phi_{v^{(q)}}}$$
$$m = -N(k), \ldots, N(k). \tag{18}$$

In the IGLOO scheme, given the azimuth symmetry $\phi_{v^{(q)}} = 2\pi v^{(q)}/V_q$, (18) can be implemented using the fast Fourier transform (FFT). For example, over the equatorial belt, 96 azimuthal samples at each elevation implies we need $96^2 = 9216$ multiplications and $96 \times (96 - 1) = 9120$ additions to evaluate (18). In comparison, using the FFT, only 317 multiplications and 633 additions are needed, which gives a speed improvement factor of 29.16 [34].

Then, we can solve $\beta_n^m(k)$ from $a_m(\theta_q, k)$ over the sampled Q elevations. By writing (17) for a specific order of m for all measured elevations, we can now form a system of simultaneous equations, *viz.*,

$$\mathbf{P}_m \mathbf{b}_m = \mathbf{a}_m, \quad m = -N(k), \ldots, N(k) \tag{19}$$

where

$$\mathbf{P}_m = \begin{bmatrix} \mathcal{P}_m^{|m|}(\cos\theta_1) & \cdots & \mathcal{P}_N^{|m|}(\cos\theta_1) \\ \vdots & \ddots & \vdots \\ \mathcal{P}_m^{|m|}(\cos\theta_Q) & \cdots & \mathcal{P}_N^{|m|}(\cos\theta_Q) \end{bmatrix} \tag{20}$$

$$\mathbf{b}_m = [\beta_m^m(k), \beta_{|m|+1}^m(k), \ldots, \beta_N^m(k)]^T \tag{21}$$

and

$$\mathbf{a}_m = \left[a_m(\theta_1, k), a_m(\theta_2, k), \ldots, a_m(\theta_Q, k)\right]^T. \tag{22}$$

The HRTF spectral components $\beta_n^m(k)$ can be calculated by solving the system of linear equations described by (19) for each order m. Since there will be noise in HRTF measurements, it is necessary to solve (19) in the least-squares sense by minimizing the mean squared error $\|\mathbf{P}_m\mathbf{b}_m - \mathbf{a}_m\|^2$. Further, to avoid the blowup of the unmeasured HRTFs (the HRTFs over lower elevations), we need to regularize the solution (the energy $\|\mathbf{b}_m\|^2$ may be included as the regularizer). Thus, the solution has the form

$$\mathbf{b}_m = \mathbf{P}_m^+\mathbf{a}_m \tag{23}$$

where $\mathbf{P}_m^+$ the Tikhonov regularized inverse explicitly given by

$$\mathbf{P}_m^+ = \begin{cases} [\mathbf{P}_m^T\mathbf{P}_m + \lambda\mathbf{I}_{N(k)-|m|+1}]\mathbf{P}_m^T, & Q \geq Z \\ \mathbf{P}_m^T[\mathbf{P}_m\mathbf{P}_m^T + \lambda\mathbf{I}_Q], & Q < Z \end{cases} \tag{24}$$

where λ is the regularization control parameter, $\mathbf{I}$ is the identity matrix, and $Z = N(k) - |m| + 1$. A systematic approach to specify λ for a meaningful result is given in the work [35]. In our experiments, we set a small value of $\lambda = 10^{-5}$, which was seen to achieve reasonable reconstruction and interpolation quality.

VI. Method Evaluation

The proposed IGLOO scheme-based sampling and data analysis methods are studied through simulation and experiments. Simulations are run on the 20-kHz audible frequency range and the HRTF data reconstruction and interpolation results are demonstrated. To illustrate both the accuracy of the magnitude and phase of HRTF, the error metric is defined as the mean square error (MSE) at the measurement (or interpolation) locations

$$\varepsilon(\theta_q, \phi_{v^{(q)}}) = \frac{\sum_{k=1}^{K}\left\|H(\theta_q, \phi_{v^{(q)}}, k) - \widetilde{H}(\theta_q, \phi_{v^{(q)}}, k)\right\|^2}{\sum_{p=1}^{P}\left\|H(\theta_q, \phi_{v^{(q)}}, k)\right\|^2}$$
$$q = 1, \ldots, Q, \quad v^{(q)} = 1, \ldots, V_q \tag{25}$$

where for each position HRTFs are measured at K frequency points $1, 2, \ldots, K$, $H(\theta_q, \phi_{v^{(q)}}, k)$ is the measured HRTF, and $\widetilde{H}(\theta_q, \phi_{v^{(q)}}, k)$ from (11) with estimated $\beta_n^m(k)$ is the reconstructed HRTF.

A. Analytical Solutions

The synthetic right ear HRTFs from the spherical head model [36] are generated 1.0 m away from the head center on a sphere according to the IGLOO-based HRTF sampling arrangement at a sampling rate of 44.1 kHz. Note in order to simulate the real experiments, no data are generated in the lower polar cap ($\theta \geq 120°$); and according to Table II in total there are 2304 samples on the sphere for frequencies up to 20 kHz.

The fast spherical harmonic transform introduced in Section V is applied to obtain the HRTF spherical harmonic coefficients. Then we can generate HRTFs at any direction. Fig. 4 gives an example of a sampled synthetic HRTF and its reconstruction results at position of $\theta = 79°$, $\phi = 90°$; and the plots in Fig. 5 show the interpolated HRTF compared to the reference of analytical solutions. It is clear in both cases the emulation results are reasonably accurate. In fact, the mean square error for the 2304 point data reconstruction is around 0.005 and the interpolation error is less than 0.01.

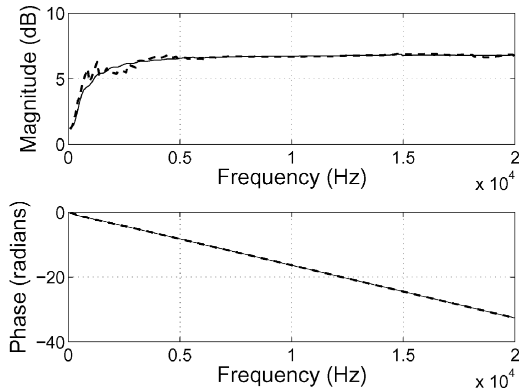


Fig. 4. Example of synthetic sampled HRTFs based on the IGLOO scheme and the reconstruction results at elevation of 79° and azimuth of 90°. Sampled data: solid line; reconstruction: dot–dash line.

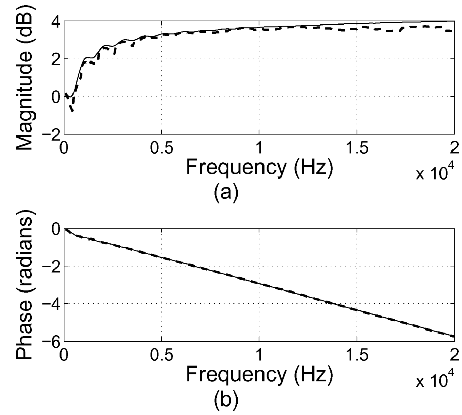


Fig. 5. Example of HRTF interpolation at elevation of 62° and azimuth of 14° compared with reference of the synthetic solutions. Reference data: solid line; interpolation: dot–dash line.

The reconstruction error performance in Fig. 6 further demonstrates that the synthesis of the right ear HRTF is usually better at the source-facing side of the head (right ear side, $\phi \in [45°, 135°]$) than the shadowed side (left ear side, $\phi \in [235°, 315°]$). This is due to the head-related diffraction effects, where the contralateral sounds have more variations, resulting in the spectral shapes that are more complicated and more difficult to model. Beside the diffraction effects, the Shaw's "bright spot" phenomenon is also observed [37]. Hence, we can see the local least reconstruction error at azimuth of 270° and elevations of 90° and 120°.

B. Experimental Data

We perform the IGLOO scheme-based HRTF measurements on a Knowles Electronics Mannequin for Acoustics Research (KEMAR) mannequin in our acoustic chamber (3.2 m × 3.2 m × 2 m) at ANU. The far-field HRTFs are measured at a distance of 1.5 m from the KEMAR at a sampling rate of 44.1 kHz. The KEMAR is placed on a turntable in the center of the measurement apparatus (a rotary hoop on which fixed loudspeakers can be placed at the specified elevations)

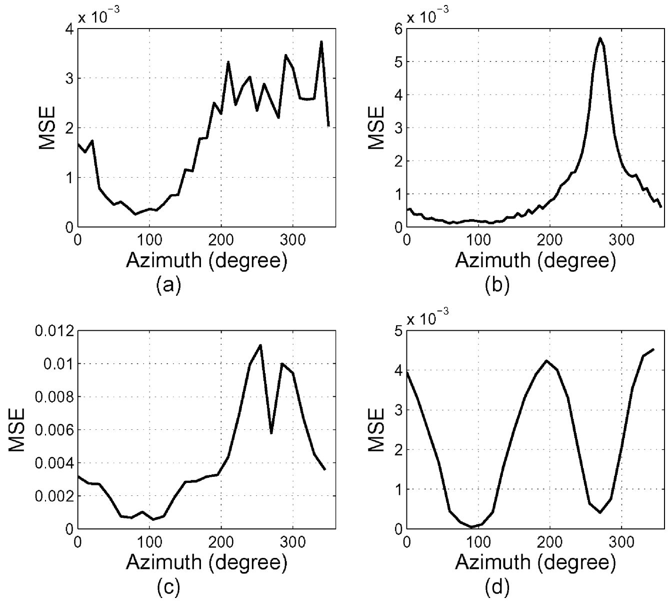


Fig. 6. Synthetic HRTFs reconstruction performance at different elevations. (a) $\theta = 30°$. (b) $\theta = 60°$. (c) $\theta = 90°$. (d) $\theta = 120°$.

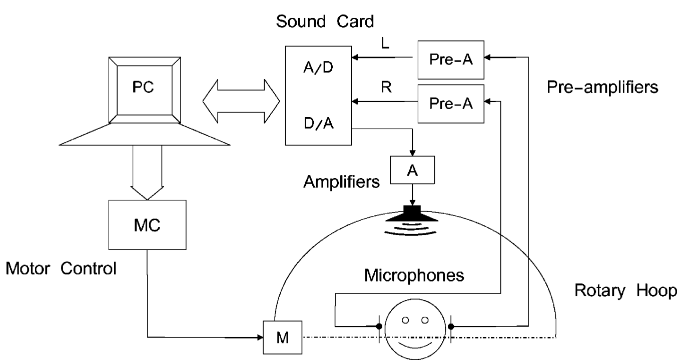


Fig. 7. Operation diagram of HRTF measurement.

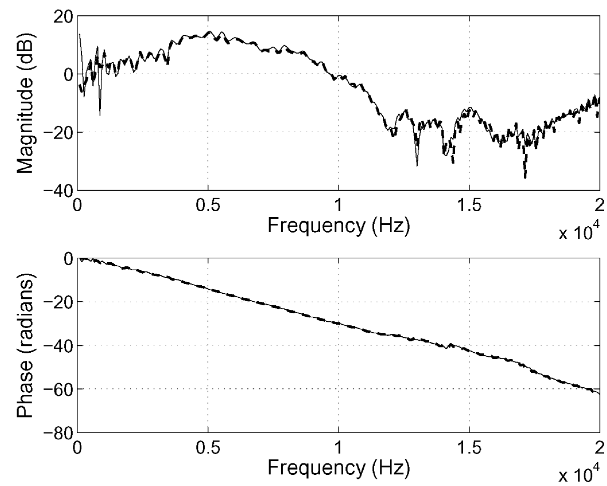


Fig. 8. Example of KEMAR left ear HRTF sampled based on the IGLOO scheme and the reconstruction results at elevation of 45° and azimuth of 160°. Sampled data: solid line; reconstruction: dot–dash line.

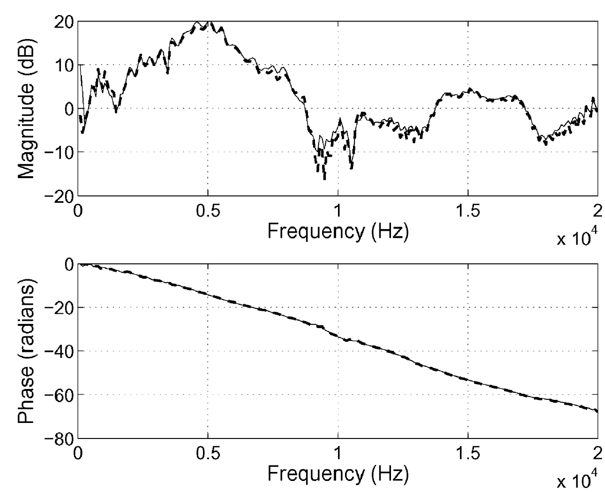


Fig. 9. Example of KEMAR left ear HRTF interpolation at elevation of 85° and azimuth of 0° compared with measurement reference. Sampled data: solid line; interpolation: dot–dash line.

and the turntable is automatically controlled to precisely set the azimuth (Fig. 7). The whole measurement procedure is carried out under a customized MATLAB program running on a desktop computer placed outside the acoustic chamber, which controls setting the measurement position, broadcasting the test signal, acquiring raw data, and performing post-signal processing [38].

We next use the experimental data to evaluate the proposed sampling method and the fast spherical harmonic analysis. The plots in Fig. 8 show a direct comparison of the measured HRTF spectrum and the reconstructed HRTF from the calculated spherical harmonic coefficients ($\theta = 45°$, $\phi = 160°$ of left ear). The reconstructed responses closely match the original measurements except some deviation at low frequencies, such as $f \leq 1$ kHz, but overall, the 2304 data reconstruction is accurate with the percent mean square error around 0.0324. Fig. 9 further demonstrates that the interpolated responses are also reasonably accurate (percent MSE of 0.0115).[6] Both left ear and right ear HRTFs reconstruction error performances shown in Fig. 10 demonstrate the similar head diffraction effects and bright spot phenomenon.

C. Comparison With Equidistance/Equiangular HRTF Samplings

The two most used samplings for the HRTF measurement are equidistance in azimuth arc (MIT HRTF database) [10] and equiangular (CIPIC HRTF database) [4]. In the former one, 710 sampling points are distributed equally in azimuth arc at all elevations, resulting in a reduction of angular resolution towards the pole of the sphere. The latter one applies the equal angular interval along both elevations and azimuths and samples the sphere with very high spatial resolution (1250 points).

[6]Here, 0.0324 is the average reconstruction error for all data while 0.0115 is the interpolation error only for the example shown in Fig. 8.

TABLE III
RECONSTRUCTION COMPARISON OF DIFFERENT HRTF SAMPLING SCHEMES

HRTF Sampling Scheme	Reconstruction Error			
	[0.2 8] kHz	[8 12] kHz	[12 16 kHz]	[16 20] kHz
Equiangular Scheme (CIPIC database)	0.1542	0.3674	0.7524	0.9905
Equidistance Scheme (MIT database)	0.0739	0.1162	0.3320	0.4292
IGLOO Scheme (ANU database)	0.0291	0.0509	0.0872	0.0973

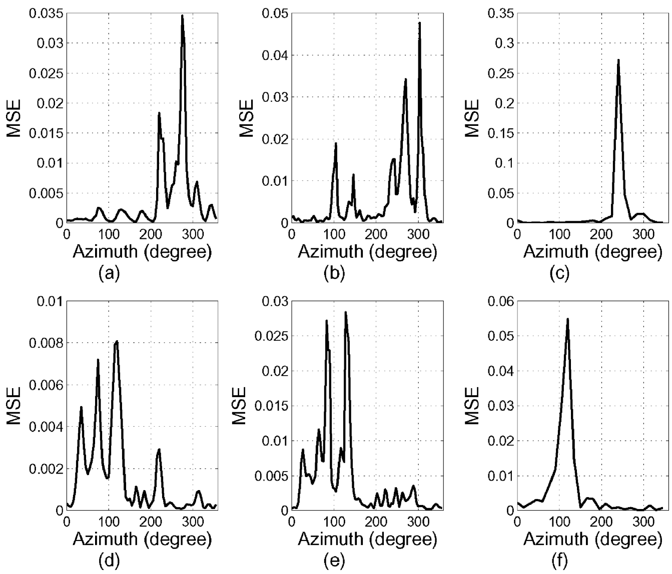


Fig. 10. KEMAR HRTFs reconstruction error performance at different elevations. (a) Right ear $\theta = 60°$. (b) Right ear $\theta = 90°$. (c) Right ear $\theta = 120°$. (d) Left ear $\theta = 60°$. (e) Left ear $\theta = 90°$. (f) Left ear $\theta = 120°$.

We further compare the proposed IGLOO sampling scheme with these two sampling arrangements. Table III shows reconstruction errors of these three HRTF datasets over different frequency ranges. The first observation is that the proposed IGLOO sampling scheme can achieve the most accurate HRTF reconstruction over the entire audio frequency range [0.2, 20] kHz because only ANU HRTF database satisfies the required sampling resolution (2209 points). CIPIC data, which is measured using human subjects, has the largest errors due to a heavy oversampling near the poles, which is impossible to control using regularization. In addition, even over low frequencies [0.2, 8] kHz (all three sets of HRTFs have enough number of samples), the proposed IGLOO scheme also shows the smallest reconstruction error. This is possibly due to the fact that the IGLOO scheme has the least distortion over the pole and uses near equal area division on the sphere.

VII. CONCLUSION

This paper has presented angular sampling and resolution criteria for HRTF measurement. A practical sampling strategy, the IGLOO scheme from the astrophysics literature [20], was shown to satisfy all the requirements for efficient, accurate, practical, and high-resolution HRTF measurement. In addition, the discrete data structure possesses an exact discrete azimuthal symmetry and thus allows fast and precise spherical harmonic transform computation. Accurate reconstruction and interpolation results up to 20 kHz are observed from both analytical solutions and experimental data using a KEMAR mannequin. The current sampling arrangement is chosen based on the representation of empirical measurements at the technical level. Psychoacoustic techniques could offer means to further reduce the number of samples.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers and the associate editor for their constructive comments and suggestions that helped to improve the presentation of this paper.

REFERENCES

[1] Y. Huang and J. Benesty, Eds., Audio Signal Processing for Next – Generation Multimedia Communication Systems. Norwell, MA, Kluwer, 2004.
[2] D. N. Zotkin, R. Duraiswami, and L. S. Davis, "Rendering localized spatial audio in a virtual auditory space," *IEEE Trans. Multimedia*, vol. 6, no. 4, pp. 553–564, Aug. 2004.
[3] F. L. Wightman and D. J. Kistler, "Headphone simulation of free-field listening I: Stimulus synthesis II: Psychophysical validation," *J. Acoust. Soc. Amer.*, vol. 85, no. 2, pp. 858–878, Feb. 1989.
[4] V. R. Algazi, R. O. Duda, D. M. Thompson, and C. Avendano, "The CIPIC HRTF database," in *Proc. IEEE Workshop Applicat. Signal Process. Audio Acoust.*, New Paltz, NY, Oct. 2001, pp. 91–102.
[5] J. Blauert, M. Brueggen, A. W. Bronkhorst, R. Drullman, G. Reynand, and L. Pellieux, "The AUDIS catalog of human HRTFs," *J. Acoust. Soc. Amer*, vol. 103, no. 5, pp. 3082–3082, May 1998.
[6] J. C. Middlebrooks and J. C. Makous, "Directional sensitivity of sound-pressure levels in the human ear canal," *J. Acoust. Soc. Amer.*, vol. 86, no. 1, pp. 89–108, Jul. 1989.
[7] H. Møller, M. F. Sørensen, D. Hammershøi, and C. B. Jensen, "Head-related transfer functions of human subjects," *J. Audio Eng. Soc.*, vol. 43, no. 5, pp. 300–321, May 1995.
[8] K. A. J. Riederer, "Head-related transfer function measurements," M.S. thesis, Helsinki Univ. of Technol., Helsinki, Finland, 1998.
[9] W. L. Martens, "Principal components analysis and resynthesis of spectral cues to perceived direction," in *Proc. Int. Comput. Music Conf.*, San Francisco, CA, Aug. 1987, pp. 274–281.
[10] W. G. Gardner and K. D. Martin, "HRTF measurements of a KEMAR," *J. Acoust. Soc. Amer.*, vol. 97, no. 6, pp. 3907–3908, Jun. 1995.
[11] B. P. Bovbjerg, F. Christensen, P. Minnaar, and X. P. Chen, "Measuring the head-related transfer functions of an artificial head with a high directional resolution," in *Proc. AES 109th Conv.*, Los Angeles, Sep. 2000, p. 17.
[12] T. Ajdler, L. Sbaiz, and M. Vetterli, "Plenacoustic function on the circle with application to HRTF interpolation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP'05)*, Philadelphia, PA, Mar. 2005, pp. 273–276.
[13] T. Ajdler, C. Faller, L. Sbaiz, and M. Vetterli, "Sound field analysis along a circle and its applications to HRTF interpolation," *J. Audio Eng. Soc.*, vol. 56, no. 3, pp. 156–175, Mar. 2008.
[14] X. L. Zhong and B. S. Xie, "Maximal azimuthal resolution needed in measurements of head-related transfer functions," *J. Acoust. Soc. Amer.*, vol. 125, no. 4, pp. 2209–2220, Apr. 2009.
[15] P. Minnaar, J. Plogsties, and F. Christensen, "Directional resolution of head-related transfer functions required in binaural synthesis," *J. Audio Eng. Soc.*, vol. 53, no. 10, pp. 919–929, Oct. 2005.
[16] M. J. Evans, J. A. S. Angus, and A. I. Tew, "Analyzing head-related transfer function measurements using surface spherical harmonics," *J. Acoust. Soc. Amer.*, vol. 104, no. 4, pp. 2400–2411, Oct. 1998.
[17] R. Duraiswami, D. N. Zotkin, and N. A. Gumerov, "Interpolation and range extrapolation of HRTFs," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. ICASSP'04*, Montreal, QC, Canada, May 2004, vol. IV, pp. 45–48.

[18] W. Zhang, T. D. Abhayapala, R. A. Kennedy, and R. Duraiswami, "Modal expansion of HRTFs: Continuous representation in frequency-range-angle," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., ICASSP'09*, Taipei, Taiwan, Apr. 2009, pp. 285–288.

[19] W. Zhang, T. D. Abhayapala, R. A. Kennedy, and R. Duraiswami, "Insights into head-related transfer function: Spatial dimensionality and continuous representation," *J. Acoust. Soc. Amer.*, vol. 127, no. 4, pp. 2347–2357, Apr. 2010.

[20] R. G. Crittenden and N. G. Turok, Exactly azimuthal pixelizations of the sky Arxiv, 1998 [Online]. Available: astro-ph/9806374v1, to be published

[21] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*. New York: Springer-Verlag, 1998.

[22] R. Duraiswami, Z. Li, D. N. Zotkin, E. Grassi, and N. A. Gumerov, "Plane-wave decomposition analysis for spherical microphone arrays," in *Proc. IEEE Workshop Applicat. Signal Process. Audio Acoust.*, New Paltz, NY, Oct. 2005, pp. 150–153.

[23] Z. Li and R. Duraiswami, "Flexible and optimal design of spherical microphone arrays for beamforming," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 15, no. 2, pp. 702–714, Feb. 2007.

[24] T. D. Abhayapala, T. S. Pollock, and R. A. Kennedy, "Characterization of 3D spatial wireless channels," in *Proc. IEEE 58th Veh. Technol. Conf., VTC 2003-Fall*, Orlando, FL, Oct. 2003, vol. 1, pp. 123–127.

[25] R. A. Kennedy, P. Sadeghi, T. D. Abhayapala, and H. M. Jones, "Intrinsic limits of dimensionality and richness in random multipath fields," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2542–2556, Jun. 2007.

[26] F. J. Simons, F. A. Dahlen, and M. A. Wieczorek, "Spatiospectral concentration on a sphere," *SIAM Rev.*, vol. 48, no. 3, pp. 504–536, 2006.

[27] J. R. Driscoll and D. M. Healy, "Computing Fourier transforms and convolutions on the 2-sphere," *Adv. Appl. Math.*, vol. 15, pp. 202–250, 1994.

[28] P. N. Swarztrauber, "On computing the points and weights for Gauss–Legendre quadrature," *SIAM J. Sci. Comput.*, vol. 24, no. 3, pp. 945–954, 2003.

[29] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann, "HEALPix: A framework for high-resolution discretization and fast analysis of data distributed on the sphere," *Astrophys. J.*, vol. 622, pp. 759–771, Apr. 2005.

[30] J. Fliege and U. Maier, "The distribution of points on the sphere and corresponding cubature formulae," *IMA J. Numer. Anal.*, vol. 19, no. 2, pp. 317–334, 1999.

[31] J. D. McEwen, M. P. Hobson, D. J. Mortlock, and A. N. Lasenby, "Fast directional continuous spherical wavelet transform algorithms," *IEEE Transactions on Signal Processing*, vol. 55, no. 2, pp. 520–529, Feb. 2007.

[32] J. Blauert, *Spatial Hearing*, revised ed. Cambridge, MA, U.K.: MIT Press, 1996.

[33] T. D. Abhayapala and A. Gupta, "Spherical harmonic analysis of wavefields using multiple circular sensor arrays," *IEEE Trans. Acoust., Speech, Lang. Process.*, vol. 18, no. 6, pp. 1655–1666, Aug. 2010.

[34] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing: Principles, Algorithms, and Applications*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[35] D. N. Zotkin, R. Duraiswami, and N. A. Gumerov, "Regularized HRTF fitting using spherical harmonics," in *Proc. IEEE Workshop Applicat. Signal Process. Audio Acoust.*, New Paltz, NY, Oct. 2009, pp. 257–260.

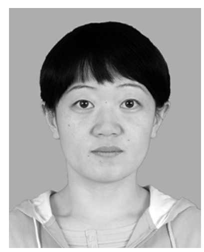
[36] R. O. Duda and W. L. Martens, "Range dependence of the response of a spherical head model," *J. Acoust. Soc. Amer.*, vol. 104, no. 5, pp. 3048–3058, Nov. 1998.

[37] E. A. G. Shaw, "The external ear," in *Handbook of Sensory Physiology: Vol V/1: Auditory System*, W. D. Keidel and W. D. Neff, Eds. New York: Springer-Verlag, 1974.

[38] M. Zhang, W. Zhang, R. A. Kennedy, and T. D. Abhayapala, "HRTF measurement on KEMAR manikin," in *Proc. ACOUSTICS'09 (Australian Acoustical Society)*, Adelaide, Australia, Nov. 2009, p. 8.

Wen Zhang (M'09) received the B.E. degree in telecommunication engineering from Xidian University, Xi'an, China, in 2003 and the M.E. degree in electrical engineering (with first class honors) and the Ph.D. degree in the area of acoustics and auditory modeling from the Australian National University, Canberra, in 2005 and 2010, respectively, where she worked on projects to measure and model the head-related transfer function for spatial sound generation.

Since February 2010, she has been an OCE Postdoctoral Fellow at CSIRO Process Science and Engineering, Sydney, Australia, and has been working on the projects to investigate acoustics of bubbles and particles in fluids. Her primary research interests include sound propagation modeling, spatial sound-field recording and reconstruction, array signal processing, and wireless channel modelling.



Mengqiu Zhang (S'10) received the M.E. degree in instrument and measurement from Xidian University, Xi'an, China, in 2006, and the B.E. degree in wireless test and measurement from China Jiliang University, Hangzhou, in 1994. She is currently working toward the Ph.D. degree in Applied Signal Processing Group, Australian National University, Canberra.

Previously, she worked for Huizhou Institute of Test and Measurement as a Senior Electronic Engineer. Her research interests are head-related transfer function (HRTF) measurement and modeling, spatial audio signal processing techniques for noise reduction in hearing aids, sound separation, and sound synthesis.

Rodney A. Kennedy (S'86–M'88–SM'01–F'05) received the B.E. degree from the University of New South Wales, Sydney, Australia, the M.E. degree from the University of Newcastle, Newcastle, Australia, and the Ph.D. degree from the Australian National University, Camberra.

He worked three years for the Commonwealth Scientific and Industrial Research Organization (CSIRO) on the Australia Telescope Project. Since 2000, he has been a Professor and is currently in the Research School of Engineering at the Australian National University. His research interests are in the fields of digital signal processing, digital and wireless communications, and acoustical signal processing.

Thushara D. Abhayapala (S'97–M'00–SM'08) received the B.E. degree (with honors) in interdisciplinary systems engineering in 1994, and the Ph.D. degree in telecommunications engineering, both from the Australian National University (ANU), Canberra, in 1999.

From 1995 to 1997, he was a Research Engineer with the Arthur C. Clarke Center for Modern Technologies, Sri Lanka. Since December 1999, he has been a faculty member at ANU. He was the Leader of the Wireless Signal Processing (WSP) Program at the National ICT Australia (NICTA) from November 2005 to June 2007. Since February 2010, he has been the Director of the Research School of Engineering, ANU. His research interests are in the areas of spatial audio and acoustic signal processing, space–time signal processing for wireless communication systems, and array signal processing. He has supervised 26 research students and coauthored over 170 peer-reviewed papers. He is currently an Associate Editor for the *EURASIP Journal on Wireless Communications and Networking*.

Dr. Abhayapala is a Member of the Audio and Acoustic Signal Processing Technical Committee (2011–2013) of the IEEE Signal Processing Society.