Comparison of Spherical Harmonics based 3D-HRTF Functional Models

Rodney A. Kennedy
Research School of Engineering,
The Australian National University,
Canberra, ACT 0200, Australia
Email: rodney.kennedy@anu.edu.au

Wen Zhang
Research School of Engineering,
The Australian National University,
Canberra, ACT 0200, Australia
Email: wen.zhang@anu.edu.au

Thushara D. Abhayapala
Research School of Engineering,
The Australian National University,
Canberra, ACT 0200, Australia
Email: thushara.abhayapala@anu.edu.au

***Abstract*—The modeling performance of three models for the 3D Head Related Transfer Function (HRTF) are compared. One of these models appeared recently in the literature whilst the other two models are novel. All models belong to the class of functional models whereby the 3D-HRTF is expressed as an expansion in terms of basis functions, which are functions of azimuth, elevation, radial distance and frequency. The expansion coefficients capture the 3D-HRTF individualization. The models differ in the choice of basis functions and the degree of orthogonality that is possibly given the constraint that for each frequency the HRTF needs to satisfy the Helmholtz wave equation. One model introduced in this paper is designed to provide a functional representation that is orthonormal on a sphere at some nominal radius and approximately so around that nominal radius. This model is shown to be superior to the other two in being able to reconstruct most efficiently the 3D-HRTF derived from a spherical head 3D-HRTF model. For all cases we show that there is a unified technique to estimate expansion coefficients from measurements taken on a sphere of arbitrary radius.**

***Index Terms*—HRTF, 3D-HRTF, 2D-HRTF, Head Related Transfer Functions, Functional Modeling, Series Expansion, Dimensionality, Helmholtz Equation.**

I. Introduction

A. Background

Physical models of pinnae, head and torso naturally lead to models of the Head Related Transfer Function (HRTF) which are functional in form, meaning they are functions of continuous spatial and frequency variables, including: the delay line model [1], the concha model [2], the spherical head model [3], the boundary element model [4], [5], and the snowman model [6], [7] as well as filter-band models: FIR model [8] and IIR model [9]), and the combined snowman filter model [10].

More mathematically motivated, but still physically consistent models, include [11], where the HRTFs are expressed in terms of surface spherical harmonics, which are a complete set of continuous functions which are orthonormal on the sphere [12]. Using spherical harmonics analysis, a preconditioned fast multipole accelerated boundary element method for HRTF computation was investigated in [13], where the model was derived by solving the wave equation based on Helmholtz reciprocity principle. Finally, a spherical harmonic based continuous functional HRTF model was developed in [14], which separated the frequency components from the spatial components and can realize both near-field and far-field HRTFs.

Rodney A. Kennedy was supported under the Australian Research Council's Discovery Projects funding scheme (Project No. DP1094350).

Thushara D. Abhayapala and Wen Zhang were supported under the Australian Research Council's Discovery Projects funding scheme (Project No. DP110103369.

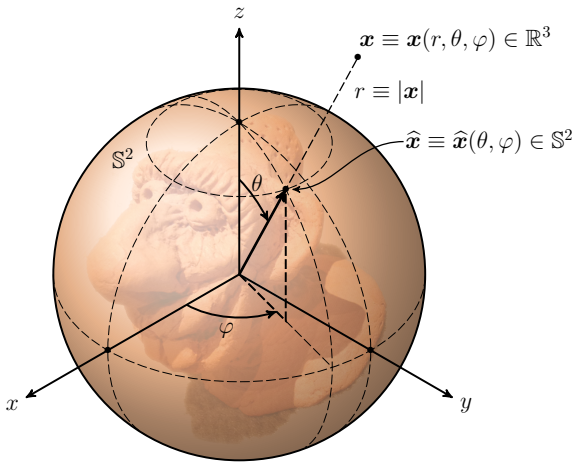

Fig. 1. Coordinate system showing a point in 3D space $\boldsymbol{x}$, a unit vector $\widehat{\boldsymbol{x}}$, which lies on the 2-sphere, $\mathbb{S}^2$, and the spherical polar coordinate system with $\theta \in [0,\pi]$ the co-latitude, $\varphi \in [0,2\pi)$ the longitude and $r \equiv |\boldsymbol{x}|$ the radius.

The situation where a model is given as a function of the three spatial variables, typically azimuth, elevation and radial distance, is referred to as a 3D-HRTF. A 3D-HRTF is always a function of frequency so there are generally four variables in play. When an HRTF is not given as a function of radial distance it is called a 2D-HRTF given it has two spatial variables plus it has a frequency variable.

B. Physical Model

In notation, see Fig. 1, we write $\boldsymbol{x} \in \mathbb{R}^3 \equiv \boldsymbol{x}(r,\theta,\varphi)$,

$$\widehat{\boldsymbol{x}} \equiv \widehat{\boldsymbol{x}}(\theta,\varphi) = (\cos\varphi\sin\theta, \sin\varphi\sin\theta, \cos\theta) = \frac{\boldsymbol{x}}{|\boldsymbol{x}|} \in \mathbb{S}^2,$$

where $r \equiv |\boldsymbol{x}|$ is the magnitude or 3D Euclidean distance, and the 2-sphere is given by $\mathbb{S}^2 \triangleq \{\boldsymbol{x} \in \mathbb{R}^3 \colon |\boldsymbol{x}| = 1\}$.

A spherical harmonics based 3D-HRTF functional model is an expansion in terms of signal independent basis functions and where the coefficients hold the signal information, that is, the HRTF individualization. In the expansion the angular portion uses the spherical harmonics, but the radial and frequency portions are constrained to have the 3D-HRTF fulfill the Helmholtz or reduced wave equation [14]–[16]

$$\Delta^2 u + k^2 u = 0$$

where $u(\boldsymbol{x})$ is a function of space $\boldsymbol{x} \in \mathbb{R}^3$ and

$$k \triangleq 2\pi/\lambda = 2\pi f/c$$

is the wavenumber, which is proportional to the frequency f and where c is the speed of sound assumed independent of frequency.

The Helmholz equation based broadband 3D-HRTF model presented in [14] is given by

$$H(r, \widehat{\boldsymbol{x}}; k) = \sum_{n,m} \eta_n^{m[s]}(r; k) Y_n^m(\widehat{\boldsymbol{x}}), \quad r > s, \tag{1}$$

where s is an effective head radius, $\widehat{\boldsymbol{x}} \equiv \{\theta, \varphi\} \in \mathbb{S}^2$ is the angular part (co-latitude and longitude) and $r \equiv |\boldsymbol{x}|$ is the radial part in a spherical polar coordinate system, $Y_n^m(\widehat{\boldsymbol{x}})$ is the spherical harmonic of degree $n \in \{0, 1, \ldots\}$ and order $m \in \{-\ell, \ldots, \ell\}$,

$$\sum_{n,m} \triangleq \sum_{n=0}^{\infty} \sum_{m=-n}^{n},$$

and

$$\eta_n^{m[s]}(r; k) \triangleq ik\, j_n(ks) \int_{\mathbb{S}^2} \rho_s(\widehat{\boldsymbol{y}}; k) \overline{Y_n^m(\widehat{\boldsymbol{y}})}\, ds(\widehat{\boldsymbol{y}})\, h_n^{(1)}(kr), \tag{2}$$

where $\rho_s(\widehat{\boldsymbol{y}}; k)$ is a virtual source distribution at the effective head radius $r = s$, $ds(\widehat{\boldsymbol{y}}) \triangleq \sin\theta\, d\theta\, d\varphi$, $j_n(\cdot)$ is the spherical Bessel function of order n, and $h_n^{(1)}(\cdot)$ is the spherical Hankel function of the first kind and of order n. It is assumed that $k \in [0, k_u]$ where k_u is the upper wavenumber assumed finite. Finally, the spherical harmonics are given by [16]

$$Y_n^m(\theta, \varphi) \triangleq \sqrt{\frac{2n+1}{4\pi} \frac{(n-|m|)!}{(n+|m|)!}} P_n^m(\cos\theta) e^{im\varphi}$$

where the associated Legendre functions are

$$P_n^m(z) \triangleq \frac{1}{2^n n!} (1 - z^2)^{m/2} \frac{d^{n+m}}{dz^{n+m}} (z^2 - 1)^n, \quad m \in \{0, 1, \ldots, n\}.$$

These definitions, for the spherical harmonics and associated Legendre functions, are not the most common in the literature but chosen to be consistent with the work in [14] from which this paper derives. They differ only in a trivial way from the most common definitions [12, pp.198-201].

C. Observations

- The wave nature is captured by noting that there are effective terms of the form

$$\left\{h_n^{(1)}(kr) Y_n^m(\widehat{\boldsymbol{x}})\right\} \tag{3}$$

  which individually are the so-called *radiating solutions* to the Helmholtz equation [16, Theorem 2.10]. We note that the frequency and radius are coupled because of the radial portion is a function of product kr. The use of the spherical Hankel function of the first kind $h_n^{(1)}(\cdot)$ can be replaced by the spherical Hankel function of the second kind $h_n^{(2)}(\cdot)$ depending on the definition of an outward traveling wave; here we stick to the convention used in [14].
- The 3D-HRTF $H(r, \widehat{\boldsymbol{x}}; k)$ is evidently completely determined by $\rho_s(\widehat{\boldsymbol{y}}; k)$ whose domain dimension is smaller than that of the 3D-HRTF. This is explained by the observation that $H(r, \widehat{\boldsymbol{x}}; k)$ cannot be arbitrary and must satisfy the Helmholtz equation for each frequency k.
- By similar arguments knowing $H(r_1, \widehat{\boldsymbol{x}}; k)$, where $r_1 > s$ is fixed, is sufficient to compute $H(r, \widehat{\boldsymbol{x}}; k)$ for all $r > s$. This is the theoretical basis of the sufficiency to perform HRTF measurements at a single fixed distance to determine, in principle, the full 3D-HRTF.
- If one then has $H(r_1, \widehat{\boldsymbol{x}}; k)$, or equivalently $\eta_n^{m[s]}(r_1; k)$, then from the expressions (1) and (2) we can deduce

$$\begin{aligned} H(r_2, \widehat{\boldsymbol{x}}; k) &= \sum_{n,m} \eta_n^{m[s]}(r_2; k) Y_n^m(\widehat{\boldsymbol{x}}) \\ &= \sum_{n,m} \frac{h_n^{(1)}(kr_2)}{h_n^{(1)}(kr_1)} \eta_n^{m[s]}(r_1; k) Y_n^m(\widehat{\boldsymbol{x}}) \\ &= \sum_{n,m} \frac{h_n^{(1)}(kr_2)}{h_n^{(1)}(kr_1)} \int_{\mathbb{S}^2} H(r_1, \widehat{\boldsymbol{y}}; k) \overline{Y_n^m(\widehat{\boldsymbol{y}})}\, ds(\widehat{\boldsymbol{y}}) Y_n^m(\widehat{\boldsymbol{x}}). \end{aligned}$$

  So if measurements are taken at fixed radius $r_1 > s$ to give us $H(r_1, \widehat{\boldsymbol{x}}; k)$ then we can compute $H(r_2, \widehat{\boldsymbol{x}}; k)$ for any other radius $r_2 > s$ [17].

II. Orthogonality and Basis Expansions

A. Rationale and Defining Criteria

Expansion of signals in terms of orthogonal functions is an effective representation because the orthogonal functions are independent of the signal (they form the basis functions) and the coefficients depend only on the signal [12]. Using basis functions, even when they are not orthogonal, yields coefficients that form a sequence, or if they are finite they form a vector, and they present a more accessible way to process and store the signal. For example, comparing signals in such models reduces to comparing sequences or vectors, rather than functions.

B. Orthogonality and Orthonormality

To be rigorous before we can talk about orthogonality we need to specify the inner product. There are three natural inner product depending of the function domains under study. We

also give expressions assuming the functions are complex-valued and these are easily modified in the real-valued case, that is, when we chose orthogonal functions which are real-valued.

For functions of frequency k on the interval $[0, k_u]$ the inner product is unweighted such that

$$\langle f, g\rangle \triangleq \int_0^{k_u} f(k)\overline{g(k)}\, dk. \tag{4}$$

For angular functions on the sphere $\mathbb{S}^2$ the inner product is given by

$$\langle f, g\rangle \triangleq \int_{\mathbb{S}^2} f(\widehat{\boldsymbol{x}})\overline{g(\widehat{\boldsymbol{x}})}\, ds(\widehat{\boldsymbol{x}}) \tag{5}$$

where $ds(\widehat{\boldsymbol{x}})$ is the area element on the sphere. These two inner products combine naturally for functions on $\mathbb{S}^2 \times [0, k_u]$ to yield

$$\langle f, g\rangle \triangleq \int_0^{k_u}\int_{\mathbb{S}^2} f(\widehat{\boldsymbol{x}}; k)\overline{g(\widehat{\boldsymbol{x}}; k)}\, ds(\widehat{\boldsymbol{x}})\, dk \tag{6}$$

Induced norms are defined through $\|f\| \triangleq \langle f, f\rangle^{1/2}$.

As usual f and g are *orthogonal* if $\langle f, g\rangle = 0$ and if in addition they satisfy $\|f\| = 1$ and $\|g\| = 1$ they are called *orthonormal*. Generally, it is clear from the context which of the three inner products, (4), (5) or (6), is in use so we do not reflect this in the notation.

III. Model Variants

The physical model in (1) and (2) does not satisfy the properties of an orthogonal expansion outlined above because the signal part is not in the form of a sequence of coefficients but takes the functional form $\rho_s(\widehat{\boldsymbol{y}}; k)$. What we need to do is find different models whose representations and coefficient parameterizations are equivalent to (1), where the signal is only present in the coefficients and the expansion involves functions that are orthonormal in some domain.

A. Model 1 – Source Distribution Expansion

Rather than attempting an orthonormal expansion for (1), which is complicated by the Helmholtz constraint that the 3D-HRTF cannot be arbitrary, we can instead attempt to represent the 2D source distribution, $\rho_s(\widehat{\boldsymbol{x}}; k)$ which lies on a sphere of radius s. In comparison to the 3D-HRTF, $\rho_s(\widehat{\boldsymbol{x}}; k)$ is unconstrained and finding an orthonormal function expansion is straightforward.

Firstly, directly in (2), we can define

$$\alpha_n^{m[s]}(k) \triangleq \int_{\mathbb{S}^2} \rho_s(\widehat{\boldsymbol{y}}; k)\overline{Y_n^m(\widehat{\boldsymbol{y}})}\, ds(\widehat{\boldsymbol{y}}), \tag{7}$$

which for fixed k are the spherical harmonic coefficients of $\rho_s(\cdot\,; k)$. This implies (2) can be written

$$\eta_n^{m[s]}(r; k) = ik\, j_n(ks)\, h_n^{(1)}(kr)\, \alpha_n^{m[s]}(k). \tag{8}$$

However, although we have implicitly dealt with the angular portion by relying on the orthonormality of the spherical harmonics, $\alpha_n^{m[s]}(k)$ remains a function of k for each n, m.

Secondly, we can introduce complete orthonormal functions,[1]

$$\big\{\varphi_q(k)\big\}_{q=1}^{\infty},$$

defined on the interval $k \in [0, k_u]$ to expand

$$\alpha_n^{m[s]}(k) = \sum_q \gamma_{n;q}^m \varphi_q(k), \tag{9}$$

where

$$\sum_q \triangleq \sum_{q=1}^{\infty},$$

with Fourier coefficients, using (4),

$$\gamma_{n;q}^m \triangleq \langle \alpha_n^m, \varphi_q\rangle.$$

The calculation of these coefficients using $\alpha_n^{m[s]}(k)$ is practically infeasible (as we do not have access to the virtual source distribution) but can be done via the measurements of the 3D-HRTF at a fixed radius (as well show later).

The source distribution can then be synthesized by the inverse of (7)

$$\begin{aligned}\rho_s(\widehat{\boldsymbol{x}}; k) &= \sum_{n,m} \alpha_n^{m[s]}(k) Y_n^m(\widehat{\boldsymbol{x}}) \\ &= \sum_{n,m;q} \gamma_{n;q}^m\, \varphi_q(k) Y_n^m(\widehat{\boldsymbol{x}}),\end{aligned}$$

where

$$\sum_{n,m;q} \triangleq \sum_q \sum_{n=0}^{\infty} \sum_{m=-n}^{n},$$

the $\{Y_n^m(\widehat{\boldsymbol{x}})\varphi_q(k)\}$ are orthonormal on $\mathbb{S}^2 \times [0, k_u]$, and thereby the full 3D-HRTF, (1) and (2), can be synthesized through

$$\begin{aligned}H(r, \widehat{\boldsymbol{x}}; k) = \sum_{n,m;q} \gamma_{n;q}^m \times \\ ik\, j_n(ks) h_n^{(1)}(kr) \varphi_q(k) Y_n^m(\widehat{\boldsymbol{x}}), \quad r > s,\end{aligned} \tag{10}$$

where

$$\big\{ik\, j_n(ks) h_n^{(1)}(kr) \varphi_q(k) Y_n^m(\widehat{\boldsymbol{x}})\big\}$$

for fixed r, is not an orthogonal on $\mathbb{S}^2 \times [0, k_u]$.

To summarize, the coefficients, $\gamma_{n;q}^m$, parametrize the source distribution referenced to a sphere of radius s (being the effective head radius) and are sufficient to compute the 3D-HRTF. But when interpreted in terms of the 3D-HRTF the effective basis is not orthonormal and so this representation is not ideal.

[1] Different sets of orthonormal functions can be used for different values of m and n, but here we use a single common set. The choice of different sets of orthonormal functions for different degrees n (but not different order m) was a feature of the model considered in [14] and used in our later investigations.

B. Model 2 – Radiating Solution Expansion

The second model is essentially that presented in [14] but varied slightly. The rationale for the model is that it is an expansion explicitly in terms of the radiating solutions (3) of the Helmholtz equation

$$H(r,\widehat{\boldsymbol{x}};k)=\sum_{n,m}\beta_n^m(k)\,h_n^{(1)}(kr)Y_n^m(\widehat{\boldsymbol{x}}),\quad r>s, \tag{11}$$

where, to be consistent with (1) we have

$$\beta_n^m(k)\triangleq ik\,\alpha_n^{m[s]}(k)j_n(ks),$$

and where $\alpha_n^{m[s]}(k)$ was defined in (7). This time we expand $\beta_n^m(k)$ by introducing complete orthonormal functions[2], $\{\varphi_q(k)\}$, defined on the interval $k\in[0,k_u]$, as follows

$$\beta_n^m(k)=\sum_q\zeta_{n;q}^m\varphi_q(k), \tag{12}$$

with Fourier coefficients

$$\zeta_{n;q}^m\triangleq\langle\beta_n^m,\varphi_q\rangle. \tag{13}$$

In this model the full 3D-HRTF, equivalent to (1) and (2), can be synthesized through

$$H(r,\widehat{\boldsymbol{x}};k)=\sum_{n,m;q}\zeta_{n;q}^m\times h_n^{(1)}(kr)\varphi_q(k)Y_n^m(\widehat{\boldsymbol{x}}),\quad r>s, \tag{14}$$

where

$$\left\{h_n^{(1)}(kr)\varphi_q(k)Y_n^m(\widehat{\boldsymbol{x}})\right\}$$

(for fixed r) are not orthogonal on $\mathbb{S}^2\times[0,k_u]$. That is, orthogonality in k is damaged by the presence of $h_n^{(1)}(kr)$.

C. Model 3 – Nominal Radius Expansion

1) 2D-HRTF Expansion: For the previous 3D-HRTF models they are expansions in terms of functions which are not orthogonal (for fixed r). The presence of the $h_n^{(1)}(kr)$ term makes separation of the frequency part and radial part difficult and its retention ultimately destroys frequency orthogonality. When the functional part lacks orthogonality then it is fundamentally inefficient. The third model does not completely alleviate the difficulties induced by the $h_n^{(1)}(kr)$ term but has the advantage of being locally orthonormal in a sense to be developed next.

Consider the 3D-HRTF for a fixed radius, $H(r_0,\widehat{\boldsymbol{x}};k)$, that is, $r_0>s$. As it is a 2D-HRTF, there is no restriction in principle that needs to be placed on it, unlike the Helmholtz constrained 3D-HRTF, and as a general function on $\mathbb{S}^2\times[0,k_u]$ it is expressible in the form

$$H(r_0,\widehat{\boldsymbol{x}};k)=\sum_{n,m;q}\rho_{n;q}^{m[r_0]}\varphi_q(k)Y_n^m(\widehat{\boldsymbol{x}}) \tag{15}$$

[2]There is no need to introduce new notation for these functions as the same ones can be used as for model 1.

where $\{Y_n^m(\widehat{\boldsymbol{x}})\varphi_q(k)\}$ are orthonormal using (6), and

$$\rho_{n;q}^{m[r_0]}\triangleq\left\langle H(r_0,\cdot\,;\cdot),Y_n^m\varphi_q\right\rangle=\int_0^{k_u}\!\!\int_{\mathbb{S}^2}H(r_0,\widehat{\boldsymbol{y}};k)\overline{Y_n^m(\widehat{\boldsymbol{y}})}\,ds(\widehat{\boldsymbol{y}})\,\overline{\varphi_q(k)}\,dk \tag{16}$$

are the Fourier coefficients.

2) 3D-HRTF Expansion: We can introduce a radial dependence to extend the above 2D-HRTF model to make it equivalent to the 3D-HRTF model (14) (and thereby the (1) and (10) models). That is, we find the expression for the 3D-HRTF expression in terms of the 2D-HRTF coefficients (16) referenced to radius r_0.

For this 2D-HRTF we can infer from Model 2, (14) where we use the coefficients from (13), and using inner product (5), that

$$\left\langle H(r_0,\cdot\,;k),Y_n^m\right\rangle=\sum_q\zeta_{n;q}^m h_n^{(1)}(kr_0)\varphi_q(k)$$

whereas, in comparison, (15), we have

$$\left\langle H(r_0,\cdot\,;k),Y_n^m\right\rangle=\sum_q\rho_{n;q}^{m[r_0]}\varphi_q(k).$$

Their equality implies

$$\sum_q\zeta_{n;q}^m\varphi_q(k)=\frac{1}{h_n^{(1)}(kr_0)}\sum_q\rho_{n;q}^{m[r_0]}\varphi_q(k),$$

and, guided by synthesis equation (14), we multiply by $h_n^{(1)}(kr)$ to obtain

$$h_n^{(1)}(kr)\sum_q\zeta_{n;q}^m\varphi_q(k)=\frac{h_n^{(1)}(kr)}{h_n^{(1)}(kr_0)}\sum_q\rho_{n;q}^{m[r_0]}\varphi_q(k),$$

which relates the Model 2 and Model 3 coefficients. So the full 3D-HRTF model (14), but now expressed in terms of the radius r_0 Model 3 2D-HRTF coefficients (16), can be written

$$H(r,\widehat{\boldsymbol{x}};k)=\sum_{n,m;q}\rho_{n;q}^{m[r_0]}\times\frac{h_n^{(1)}(kr)}{h_n^{(1)}(kr_0)}\varphi_q(k)Y_n^m(\widehat{\boldsymbol{x}}),\quad r>s, \tag{17}$$

which reduces to (15) when $r=r_0$. The coefficients $\{\rho_{n;q}^{m[r_0]}\}$ indeed hold the signal information of the full 3D-HRTF.

When $r=r_0$ the functional part reduces to $Y_n^m(\widehat{\boldsymbol{x}})\varphi_q(k)$ which is orthonormal on $\mathbb{S}^2\times[0,k_u]$. For other values of r the functional part has a ratio of spherical Bessel functions distorting the frequency k portion and destroying the orthogonality (and orthonormality). However, for kr close to kr_0 there is near orthogonality given that the Taylor series is

$$h_n^{(1)}(z)=h_n^{(1)}(z_0)+(z-z_0)\frac{\partial}{\partial z}h_n^{(1)}(z)+\cdots$$

and, therefore, with $z=kr$ and $z_0=kr_0$,

$$\frac{h_n^{(1)}(kr)}{h_n^{(1)}(kr_0)}=1+O(kr-kr_0),\ \text{as}\ kr\to kr_0.$$

The constant associated with the correction term grows with n, following similar analysis given in [18].

TABLE I
MODEL PARAMETERS IN THE GENERAL FORM (18)

Model	Expansion	Coefficients $\iota_{n;q}^{m}$	Weighting $g_n(k,r)$
1	Source Distribution (10)	$\gamma_{n;q}^{m}$	$ik\, j_n(ks)h_n^{(1)}(kr)$
2	Radiating Solution (14)	$\zeta_{n;q}^{m}$	$h_n^{(1)}(kr)$
3	Nominal Radius (17)	$\rho_{n;q}^{m[r_0]}$	$h_n^{(1)}(kr)/h_n^{(1)}(kr_0)$

IV. Computing the Coefficients for the Models from Measurements on a Sphere

A. Recovering the Coefficients through Orthonormality

For a given HRTF we show how to determine the coefficients in the three models (10), (14) and (17). Spherical harmonics based HRTF functional models are particularly well-catered for measurements taken on a sphere at a fixed radius. The three models are sufficiently similar that they can all be written in the form

$$H(r,\widehat{\boldsymbol{x}};k)=\sum_{n,m;q}\iota_{n;q}^{m}\, g_n(k,r)\varphi_q(k)Y_n^m(\widehat{\boldsymbol{x}}), \qquad (18)$$

where the coefficients and functions corresponding to the three models are given in Table I.

Following the methodology in [14] we assume we have the HRTF available at some radius $r=r_1$ and then the task to estimate $\{\iota_{n;q}^{m}\}$ from such $H(r_1,\widehat{\boldsymbol{x}};k)$ is possible via

$$\iota_{n;q}^{m}=\int_0^{k_u}\frac{1}{g_n(k,r_1)}\times \int_{\mathbb{S}^2}H(r_1,\widehat{\boldsymbol{x}};k)\overline{Y_n^m(\widehat{\boldsymbol{x}})}\,ds(\widehat{\boldsymbol{x}})\,\overline{\varphi_q(k)}\,dk. \qquad (19)$$

Such analysis reveals that we can exploit the orthonormality of both the spherical harmonics and the $\{\varphi_q(k)\}$ by freezing the radius in this case to $r=r_1$.

B. Nominal Radius Expansion Example

In the case of the nominal radius expansion, or Model 3, there is no need to identify the measurement, which is at a radius $r=r_1$, with the nominal radius $r=r_0$ about which we have local orthonormality. Indeed, from Table I, we have

$$\rho_{n;q}^{m[r_0]}=\int_0^{k_u}\frac{h_n^{(1)}(kr_0)}{h_n^{(1)}(kr_1)}\times \int_{\mathbb{S}^2}H(r_1,\widehat{\boldsymbol{x}};k)\overline{Y_n^m(\widehat{\boldsymbol{x}})}\,ds(\widehat{\boldsymbol{x}})\,\overline{\varphi_q(k)}\,dk. \qquad (20)$$

This has the interpretation that we acquire the measurement as radius r_1 and determine the coefficients in the nominal radius expansion where there is a localized orthonormality about radius $r_0\neq r_1$. Of course, (20) reduces to (16) when $r_1=r_0$ which presents the simplest case.

C. Recovering the Coefficients from Incomplete Measurements

The integral transforms in (19) need to be typically computed from discrete measurements on the sphere and in frequency or time [19]. There are a multiplicity of methods and approaches and the focus of this paper is not to develop new techniques. However, there is one aspect of HRTF measurements which needs some care and that is the measurements are rarely over the whole sphere and similarly may not be made at all frequencies. We call these incomplete measurements and we give general pointers how to treat this problem by reducing the distortion from missing measurements.

The estimation of coefficients given incomplete measurements can be approached from a generalization of the Papoulis algorithm [20]–[22] or conjugate gradient method [23]. Alternatively the estimation can be done by the least squares method given in [17], [24] and exploited in [14].

V. Spherical Head Model Based Evaluation

We use the spherical head HRTF as an example to evaluate the efficiency of the three proposed models. The closed form HRTF for a spherical head is given by [3], [14]

$$H_{\rm s}(r,\widehat{\boldsymbol{x}};k)=-\frac{4\pi}{ks^2}re^{-ikr}\sum_{n,m}\frac{h_n^{(1)}(kr)}{h_n'^{(1)}(ks)}\overline{Y_n^m(\widehat{\boldsymbol{e}})}Y_n^m(\widehat{\boldsymbol{x}}),$$

where $\widehat{\boldsymbol{e}}\in\mathbb{S}^2$ is the vector to the measurement point on the surface of the sphere, or the assumed ear position, and s is the spherical head radius.

Substituting the above spherical head model analytical expression into the three proposed models leads to the corresponding coefficients written as

$$\sum_q \iota_{n;q}^{m}\,\varphi_q(k)=-\frac{4\pi\, re^{-ikr}\overline{Y_n^m(\widehat{\boldsymbol{e}})}}{ks^2\, h_n'^{(1)}(ks)}\times\begin{cases}\dfrac{1}{ikj_n(ks)} & \text{Model 1}\\ 1 & \text{Model 2}\\ h_n^{(1)}(kr_0) & \text{Model 3}\end{cases}.$$

Following [14], for the spatial part we use the spherical harmonic degree n truncation [25], [26]

$$N(s;k)=\lceil eks/2\rceil,$$

where $e=2.71818\ldots$, k is the wavenumber and s is the effective head radius. That is, $n=0,1,\ldots,N(s;k)$. In [14], s is set as a function of k in the sense that below a threshold frequency is set at value 0.2m because the torso comes into play and above the value 0.09m is used for the head only effect. In the spherical head case the s is fixed at 0.09m. Despite its complications, because the truncation is common to the three models, it provides a fair basis for comparison.

We apply the spherical Fourier Bessel series as the complete orthonormal functions on $[0,k_u]$ [14], that is,

$$\varphi_q^{[n]}(k)=j_n(Z_q^{(n)}k/k_u),$$

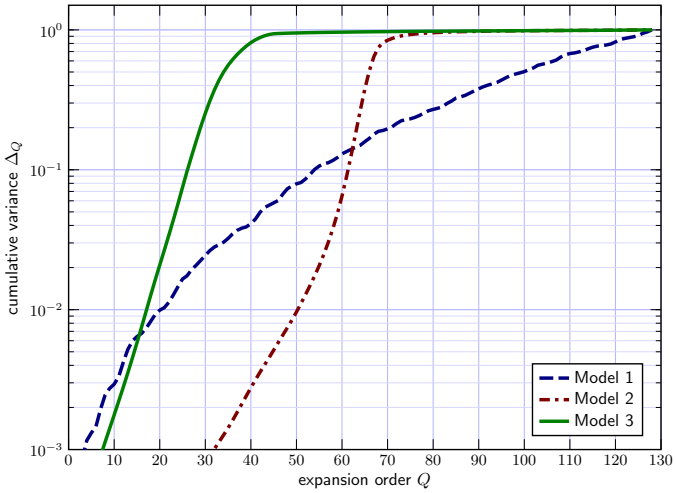

Fig. 2. Comparison of the cumulative variance of the expansion coefficients for the three models. This shows Model 3 can use less terms in the spherical Fourier Bessel series to model the reference spherical head model data.

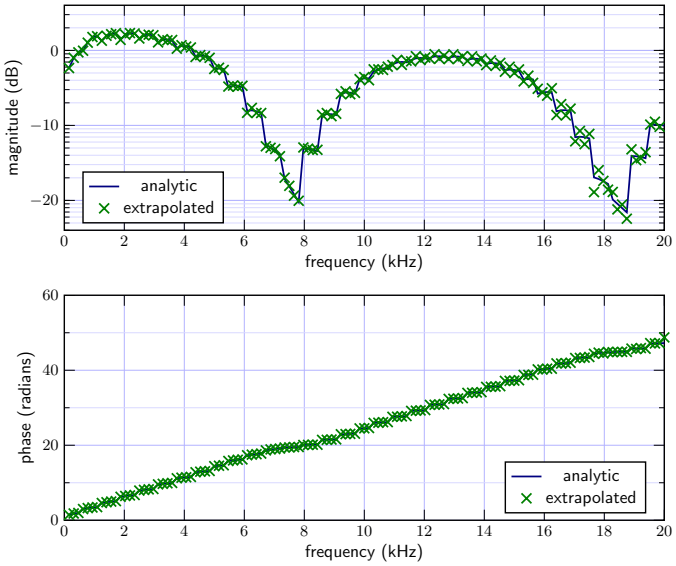

Fig. 3. Example of the spherical head HRTF reconstructed using the proposed Model 3. The spherical head model generated synthetic data at $r_0 = 1.2$ m and used to determine the Model 3 parameters. Then Model 3 was used to extrapolate to $r = 1.5$ m for $(\theta = 90°, \varphi = 80°)$ and shown marked with $\times$. This was compared with the $r = 1.5$ m analytical spherical head model reference data shown as a solid line.

where we have chosen a different series to apply for each n, and define the cumulative variance of the expansion coefficients as the evaluation metric

$$\Delta_Q = \frac{\mathbb{E}\{\sum_{q=1}^{Q} |\iota_{n;q}^m|^2\}}{\mathbb{E}\{\sum_{q=1}^{Q_{\max}} |\iota_{n;q}^m|^2\}}.$$

$\mathbb{E}\{\cdot\}$ represents the expectation operator performed on different spherical harmonic expansion orders (n, m) and $Q_{\max}$ denotes the maximum expansion order (the number of frequency samples).

Fig. 2 shows the cumulative variance of the decomposed coefficients of the three models. In this example, the HRTFs are obtained at $r_0 = 1.2$ m and extrapolated to $r = 1.5$ m. Model 3 demonstrates the fastest convergence in terms of representing the HRTF spectra within a bandwidth of 20 kHz, which corroborates an assertion that it is the most efficient representation among the three models proposed in this paper. We need 45 spherical Fourier Bessel series in the Model 3 (including more than 93% cumulative variance) to represent the 20 kHz HRTF spectra ($Q_{\max} = 128$ frequency samples). This means that the model can achieve around 65% data reduction. Fig. 3 plots the spherical head model and reconstructed HRTFs at co-latitude $\theta = 90°$ and longitude $\varphi = 80°$ using the Model 3. It is clear that the Model 3 extrapolated response closely matches the analytic spherical head data.

VI. Discussion and Conclusions

Spherical harmonics based 3D-HRTF functional models have been developed in this work. The angular portion of these basis functions is chosen for pragmatic reasons to use the spherical harmonics. The frequency and radial portions of these basis functions are complicated because they are constrained by the requirement to satisfy the Helmholtz wave equation. Ultimately this means that the 3D-HRTF which is defined on the four dimensional domain — two angular dimensions, one radial dimension and one frequency dimension — is not easily amenable to an expansion where the basis functions are fully orthonormal across all four variables. This means there is a trade-off between what portions of the functional expansion are made orthonormal (in fewer than the four variables) and the impact on the efficiency of the representation caused by the non-orthogonality.

Three expansions were shown which are all exact in the sense that with an infinite number of terms they equal the physically based 3D-HRTF model which satisfies the Helmholtz equation. However, each of the three models correspond to different choices of basis functions. Previously one of these models has appeared in the literature, Model 2, and corresponds to using the radiating solutions to the Helmholtz equation as the basis functions [14], [16]. Two new models were introduced in this work. The first, Model 1, considers an orthonormal representation of the virtual source distribution on a sphere but the orthonormality is not carried over to the 3D-HRTF. However, it does have a unambiguous physical interpretation. The second novel model, Model 3, seeks to find a representation that is orthonormal on a sphere at some nominal radius of interest (and approximately so around that nominal radius). For all three cases we showed that there is a unified technique to estimate the expansion coefficients from measurements taken on a sphere of arbitrary radius.

To compare the models we employed the analytic spherical head model and obtained closed-form expressions for the coefficients in each of the expansions. Our numerical investigations corroborated our assertions that the most efficient model — approximating the HRTF with the fewest number of

truncated terms — should correspond to the one with the best orthogonality which was Model 3. Further work is required to fully compared the models presented and determine their relative effectiveness through analysis.

Finally, in terms of open problems, we note that the radial normalization developed in [14] can be incorporated into each model and this may help in finding a superior model. Also from an abstract perspective the Helmholtz equation can be considered as a linear subspace constraint and there should exist an intrinsic orthonormal functional 3D-HRTF expansion in some three-dimensional domain. However, the exact form of this intrinsic expansion currently eludes us or its existence needs to be falsified.

REFERENCES

[1] D. W. Batteau, "The role of the pinna in human localization," *Proc. R. Soc. Lond. B*, vol. 168, no. 1011, pp. 158–180, Aug. 1967.

[2] E. A. Lopez-Poveda and R. Meddis, "A physical model of sound diffraction and reflections in the human concha," *J. Acoust. Soc. Am.*, vol. 100, no. 5, pp. 3248–3259, Nov. 1996.

[3] R. O. Duda and W. L. Martens, "Range dependence of the response of a spherical head model," *J. Acoust. Soc. Am.*, vol. 105, no. 5, pp. 3048–3058, Nov. 1998.

[4] B. F. G. Katz, "Boundary element method calculation of individual head-related transfer function. I. rigid model calculation," *J. Acoust. Soc. Am.*, vol. 110, no. 5, pp. 2440–2448, Nov. 2001.

[5] B. F. G. Katz, "Boundary element method calculation of individual head-related transfer function. II. impedance effects and comparisons to real measurements," *J. Acoust. Soc. Am.*, vol. 110, no. 5, pp. 2449–2455, Nov. 2001.

[6] N. A. Gumerov, R. Duraiswami, and Z. Tang, "Numerical study of the influence of the torso on the HRTF," in *Proc. 2002 IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 2, Orlando, FL, May 2002, pp. 1965–1968.

[7] V. R. Algazi, R. O. Duda, R. Duraiswami, N. A. Gumerov, and Z. Tang, "Approximating the head-related transfer function using simple geometric models of the head and torso," *J. Acoust. Soc. Am.*, vol. 112, no. 5, pp. 2053–2064, Nov. 2002.

[8] F. L. Wightman and D. J. Kistler, "Headphone simulation of free-field listening. I: Stimulus synthesis," *J. Acoust. Soc. Am.*, vol. 85, no. 2, pp. 858–867, Feb. 1989.

[9] A. Kulkarni and H. S. Colburn, "Infinite-impulse-response models of the head-related transfer function," *J. Acoust. Soc. Am.*, vol. 115, no. 4, pp. 1717–1728, Apr. 2004.

[10] V. R. Algazi, R. D. Duda, and D. M. Thompson, "The use of head-and-torso models for improved spatial sound synthesis," in *Proc. 113th Audio Engineering Society Conv.*, Los Angeles, CA, Oct. 2002, pp. 1–18.

[11] M. J. Evans, J. A. S. Angus, and A. I. Tew, "Analyzing head-related transfer function measurements using surface spherical harmonics," *J. Acoust. Soc. Am.*, vol. 104, no. 4, p. 2400, Oct. 1998.

[12] R. A. Kennedy and P. Sadeghi, *Hilbert Space Methods in Signal Processing*. Cambridge, UK: Cambridge University Press, Mar. 2013.

[13] N. A. Gumerov, A. E. O'Donovan, R. Duraiswami, and D. N. Zotkin, "Computation of the head-related transfer function via the fast multi-pole accelerated boundary element method and its spherical harmonic representation," *J. Acoust. Soc. Am.*, vol. 127, no. 1, pp. 370–386, Jan. 2010.

[14] 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. Am.*, vol. 127, no. 4, pp. 2347–2357, Apr. 2010.

[15] 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. Acoustics, Speech, and Signal Processing, ICASSP'2009*, Taipei, Taiwan, Apr. 2009, pp. 285–288.

[16] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, 3rd ed. New York, NY: Springer, 2013.

[17] M. Pollow, K.-V. Nguyen, O. Warusfel, T. Carpentier, M. Müller-Trapet, M. Vorländer, and M. Noisternig, "Calculation of head-related transfer functions for arbitrary field points using spherical harmonics decomposition," *ACTA Acust. United Ac.*, vol. 98, no. 1, pp. 72–82, 2012.

[18] R. A. Kennedy, D. B. Ward, and T. D. Abhayapala, "Nearfield beamforming using radial reciprocity," *IEEE Trans. Signal Process.*, vol. 47, no. 1, pp. 33–40, Jan. 1999.

[19] W. Zhang, M. Zhang, R. A. Kennedy, and T. D. Abhayapala, "On high resolution head-related transfer function measurements: An efficient sampling scheme," *IEEE Trans. Audio Speech Language Process.*, vol. 20, no. 2, pp. 575–584, Feb. 2012.

[20] A. Papoulis, "A new algorithm in spectral analysis and band-limited extrapolation," *IEEE Trans. Circuits Syst.*, vol. CAS-22, no. 9, pp. 735–742, Sep. 1975.

[21] W. Zhang, R. A. Kennedy, and T. D. Abhayapala, "Iterative extrapolation algorithm for data reconstruction over sphere," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing, ICASSP'2008*, Las Vegas, Nevada, Mar. 2008, pp. 3733–3736.

[22] R. A. Kennedy, W. Zhang, and T. D. Abhayapala, "Spherical harmonic analysis and model-limited extrapolation on the sphere: Integral equation formulation," in *Proc. International Conference on Signal Processing and Communication Systems, ICSPCS'2008*, Gold Coast, Australia, Dec. 2008, p. 6.

[23] Z. Khalid, R. A. Kennedy, S. Durrani, and P. Sadeghi, "Conjugate gradient algorithm for extrapolation of sampled bandlimited signals on the 2-sphere," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing, ICASSP'2012*, Kyoto, Japan, Mar. 2012, pp. 3825–3828.

[24] D. N. Zotkin, R. Duraiswami, and N. A. Gumerov, "Regularized HRTF fitting using spherical harmonics," in *Proc. of 2009 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY, October 2009, pp. 257–260.

[25] H. M. Jones, R. A. Kennedy, and T. D. Abhayapala, "On dimensionality of multipath fields: Spatial extent and richness," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing, ICASSP'2002*, vol. 3, Orlando, FL, May 2002, pp. 2837–2840.

[26] 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.