Efficient Continuous HRTF Model Using Data Independent Basis Functions: Experimentally Guided Approach

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

***Abstract*—This paper introduces a continuous functional model for head-related transfer functions (HRTFs) in the horizontal auditory scene. The approach uses a separable representation consisting of a Fourier–Bessel series expansion for the spectral components and a conventional Fourier series expansion for the spatial components. Being independent of the data, these two sets of basis functions remain unchanged for all subjects and measurement setups. Hence, the model can transform an individualized HRTF to a subject specific set of coefficients. A continuous functional model is also developed in the time domain. We show the efficient model performance in approximating experimental measurements by using the HRTF measurements from a KEMAR manikin and the synthetic data from the spherical head model. The statistical results are determined from a 50-subject HRTF data set. We also corroborate the predictive capability of the proposed model. The model has near optimal performance, which can be ascertained by comparison with the standard principle component analysis (PCA) and discrete Karhunen–Loeve expansion (KLE) methods at the measurement points and for a given number of parameters.**

***Index Terms*—Continuous model, Fourier–Bessel series, head-related transfer function (HRTF).**

I. Introduction

HUMANS have a remarkable ability to identify the direction of a sound source originating from any point in three-dimensional space. The perceptual cues for spatial localization include the amplitude and the time arrival of the sound at each ear and most importantly, the spectrum of the sound, which is modified by the interaction between the sound wave and a person's body (the torso, head, and external pinna shape). The head-related transfer function (HRTF) contains all the relevant spatial cues and describes how a given sound wave input (parameterized as frequency and source location) is filtered by the diffraction and reflection properties of the individual body shape before the sound reaches the listener's eardrum [1].

Manuscript received December 17, 2007; revised December 04, 2008. Current version published April 01, 2009. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Rudolf Rabenstein.

The authors are with the Department of Information Engineering, Research School of Information Sciences and Engineering, Australian National University, Canberra ACT 0200, Australia (e-mail: wen@cecs.anu.edu.au; rodney.kennedy@anu.edu.au; thushara.abhayapala@anu.edu.au).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TASL.2009.2014265

HRTFs are usually obtained from measurements on people (or dummy heads) and so there is a rich variety of HRTFs corresponding to the different human subjects being measured[1]. Such data is naturally taken from systematic measurements over a discrete set of angles and at discrete frequencies or time samples. While the data is discrete by necessity, it is understood that the underlying HRTF is fully continuous in space and frequency. It is a well studied problem of finding how to interpolate the discrete measurements to obtain synthetic or estimated data at any particular angle [2]–[4]; but these methods' validity is not related to the spatial sampling condition and the most appropriate interpolation can be still considered as an open question. The discrete Fourier transform (DFT) is commonly applied for the HRTF spectral analysis so only limited spectral resolution can be achieved. In this paper, a continuous HRTF model in both spatial and spectral domain is proposed so that the need for interpolation is removed.

A. Conventional HRTF Modeling Methods

In the case of discrete data and sets of measurements corresponding to different human subjects, there is a number of work devoted to HRTF functional representations, like the filter bank models [5], [6], and related areas like principal component analysis (PCA) [7], [8] and surface spherical harmonic expansion [9]. Results show that sound synthesis through the filter-bank models are very close to the original HRTF measurements, but the expansion weights in the model are coupled with both angle and frequency variables, which limits the usefulness of the model for HRTF reconstruction.

Statistical methods, PCA and the (discrete) Karhunen-Loeve expansion (KLE) [10], are used to find the most efficient lower dimensional representation of the statistical data where efficiency is typically measured in terms of variance. These optimal discrete methods rely on the empirical data and provide an optimal description for any given lower dimensional model. However, their dependence on the empirical data is a weakness as well as a strength. The clear strength is that the representation deals directly with the set of measurements and reality rather than an abstraction through a model which may be inaccurate. The weakness is that optimal discrete models only represent the given empirical data set and any changes to that data set change the model. That is, the principle components, which are basis vectors, will vary with any change to the data set

[1]Here, we assume that the HRTFs are measured at a fixed distance between the loudspeaker and head center and there is no head rotation movement.

(additions or omissions). Of course, if a data set is sufficiently large and rich, in the sense of capturing the true variance across the population, the basis functions may be fixed without much loss of optimality [11]. Further, because the data is discrete and taken from a measurement grid, the representation is not universal in the sense that HRTF measurements taken with differing measurement grids cannot be merged directly, whereas the underlying continuous HRTF should be consistent and is independent of how it is measured. Our model of the HTRF combines the strengths of the empirical formulation with the use of a continuous model, retaining the near optimality of PCA type formulations but with a data-independent set of basis functions.

HRTFs have also been represented as a weighted sum of surface spherical harmonics in three dimensions [9], and a series of multipoles based on the reciprocity principle have been used to express the HRTF [12]. The advantage of using this orthogonal expansion is that the basis functions, spherical harmonics, provide a natural continuous representation in the spatial domain. Therefore, it leads to a straightforward solution to the problem of HRTF interpolation in elevation and azimuth. However, as the series weights are still functions of the measured frequencies, the above two models cannot achieve any optimality in lower dimensional representations compared to statistical modeling methods.

B. Continuous Representation of the HRTF

The continuous model developed in this paper for HRTFs expands any particular HRTF in terms of a weighted sum of continuous basis functions which are separable in azimuth angle and frequency. The weights carry the information regarding the individuality. The study first uses the Fourier expansion to represent the spatial dependence embedded within the structure of horizontal domain measured HRTFs. As the series weights depend on frequency, analyzing these components can provide a new means to model the scattering behavior of the human head and pinna. As stated in [13], HRTF frequency domain data is more relevant to the elevation source localization and diffraction effects; for example, the spectral notches and peaks change with elevation. These observations require the developed HRTF model to have high spectral resolution even though most experimental measurements are limited by the sampling rate and number of samples. The novelty of this work is applying parametric spectrum analysis to model these frequency components of the HRTF. Different from traditional methods using DFT for HRTF spectral analysis, the parametric approach is to develop a model for the HRTF frequency components, and then the parameters are estimated from the empirical data. Once the model is given, the spectrum of HRTF can be determined with better spectral resolution and fewer parameters (so that the lower dimensional representation is achieved).

Our proposed approach, for the spectral representation portion of the HRTF, is to seek a preferred orthonormal function set from a range of closed form orthonormal functions as the one that best empirically represents test data (described in Section III). Each of the candidate closed form orthonormal functions is by its nature fixed and independent of any data. A member of the class of *Fourier–Bessel series* expansions is identified as the preferred representation due to the strong correlation between these basis functions and the measured spectral structure of the HRTFs (the previous work [14] found this fact and used the Fourier–Bessel series to reproduce the measured HRTFs). So the important distinction in our approach is to use the test data to make a one-off *choice* of the set of basis functions from a range of candidate sets which can then be applied to any measurements including those of different measurement grids. In contrast, PCA and related statistical approaches use a single set of data to generate empirical basis vectors which can only be used on that single set of data. Our approach is *continuous* and the basis are functions and *data independent* (the spectral representation is fixed as a *Fourier–Bessel* series, and the spatial azimuth representation is fixed as a *Fourier* series) whereas for PCA the basis is discrete, measurement grid dependent and data dependent. Thus, the proposed model can achieve continuous HRTF representation (at any frequency point and an arbitrary angle) for any HRTF data set.

The paper is organized as follows. Section II develops the HRTF model in both frequency and time domains and describes the model parameter estimation. Section III describes the acoustic validation of the model using the MIT HRTF data set [15] collected from a KEMAR manikin, and some analytically generated HRTFs from the spherical head model [16]. We also investigate the statistical results on the model performance from a 50-subject HRTF database [17]. Section IV compares the continuous model with the PCA method and shows the near optimality of performance of the proposed model. The paper concludes in Section V.

II. HRTF Continuous Function Model

We develop a continuous function representation of the horizontal plane HRTF as a function of frequency $f \in (0, f_{\max})$, where $f_{\max}$ is the maximum measurement frequency, and azimuth angle $\phi \in [0, 2\pi)$. In the following subsections we develop spatial (Section II-A) and spectral (Section II-B) continuous representations of the HRTF. Further, the time domain HRIR (head-related impulse response) model is provided in Section II-C.

A. HRTF Spatial Modeling

Using a Fourier series expansion, $H(f, \phi)$ can be expanded in the ϕ coordinate as

$$H(f,\phi) = \sum_{m=-\infty}^{\infty} A_m(f) e^{jm\phi} \tag{1}$$

where the mth order Fourier series weight is given by

$$A_m(f) = \frac{1}{2\pi} \int_0^{2\pi} H(f,\phi) e^{-jm\phi}\, d\phi. \tag{2}$$

The work of [18] shows that the horizontal HRTF has a low pass character with limited spatial bandwidth. Most of energy in $A_m(f)$ is restricted in a region which is limited by $|m| \leq 2\pi f a/c$, where c denotes the sound propagation velocity and a is the radius of the head (usually we assume $c = 340$ m/s

and $a = 0.09$ m). While outside of this region, the energy is not significant and decaying rapidly with increasing order $|m|$. Thus, the infinite series (1) can be truncated to $|m| \leq M$ to represent a band-limited HRTF function through

$$H_M(f,\phi) \triangleq \sum_{m=-M}^{M} A_m(f) e^{jm\phi}. \tag{3}$$

As a guide, based on [18], a suitable M can be determined through

$$M = \lceil 2\pi a f_{\max}/c \rceil \tag{4}$$

where $\lceil \cdot \rceil$ is the integer ceiling function, and $f_{\max}$ is the maximum frequency.

To estimate the coefficients $A_m(f)$, we replace the integral in (2) with a finite summation at discrete angles corresponding to the measurements. The standard Fourier analysis for discrete periodic signals is used here. Along the circle $0 \leq \phi < 2\pi$, given N equi-azimuthal angle increments, $\phi_0, \ldots, \phi_{N-1}$, we have

$$A_m(f) \cong \frac{1}{N} \sum_{i=0}^{N-1} H(f,\phi_i)\, e^{-jm\phi_i}, \quad |m| \leq M. \tag{5}$$

Equations (3) and (5) show that the continuous $H(f,\phi)$ can be recovered from the N discrete measurements $H(f,\phi_i)$ provided the number of samples $N \geq (2M+1)$ [19]. This is because, given N sampled discrete measurements, the Fourier series representation will contain at most N expansion components. In particular, given the maximum measurement frequency $f_{\max}$ and (4), the horizontal HRTFs should be sampled with $N \geq (2\lceil 2\pi a f_{\max}/c \rceil + 1)$ so that the function can be accurately reconstructed.

As the bandlimited HRTF $H_M(f,\phi)$ is completely determined by $2M+1$ Fourier series weights and these weights are functions of frequency only, we believe analyzing $A_m(f)$ provides a new means to investigate the HRTF spectrum and study the scattering characteristics of human head and pinna. In order to achieve the goal of a continuous representation, parametric spectral analysis is applied in Section II-B below to model $A_m(f)$ by introducing a second orthogonal series expansion in frequency f.

B. HRTF Spectral Modeling

Parametric spectral analysis seeks a functional representation of the HRTF spectrum which can give better spectral resolution from the sampling conditions—sampling rate and number of samples. Such a representation facilitates direct HRTF evaluation at any frequency point. Our approach is to find which closed form standard orthogonal functions match the experimentally measured HRTFs distribution most efficiently as basis functions and expand the HRTF spectral components with these basis functions.

We initially formulate a general orthogonal representation of the HRTF frequency components $A_m(f)$ for each m defined over the finite interval $(0, f_{\max})$, where $f_{\max}$ is the maximum frequency. Unlike the expansion for the azimuthal variable there is no obviously compelling choice for the preferred orthogonal representation so this is a crucial issue in formulating our representation. We write

$$A_m(f) = \sum_{k=1}^{\infty} C_{mk}\, \varphi_k^{(m)}(f) \tag{6}$$

where $\varphi_k^{(m)}(f)$ is a suitable complete set of orthogonal functions with weight $W(f)$ defined on the finite interval $(0, f_{\max})$ and indexed by k (note different orthogonal functions may be chosen for different values of m). Theoretically any square-integrable function can be represented by (6), [20], and arbitrarily well approximated truncating to a finite number of terms K, i.e.,

$$A_m(f) \cong \sum_{k=1}^{K} C_{mk} \varphi_k^{(m)}(f). \tag{7}$$

The order K is usually chosen as a tradeoff between accuracy and economy of representation, noting that the approximation can be made arbitrarily close by choosing K sufficiently large.

The expansion coefficients C_{mk} in (6) can be obtained from

$$C_{mk} = \frac{1}{h_k} \int_0^{f_{\max}} A_m(f) \varphi_k^{(m)}(f)\, W(f)\, df \tag{8}$$

where

$$h_k = \int_0^{f_{\max}} \left| \varphi_k^{(m)}(f) \right|^2 W(f)\, df \tag{9}$$

and $W(f)$ is the prescribed (non-negative) weighting function defined on $f \in (0, f_{\max})$. The weighting function for a given complete orthogonal basis $\{\varphi_k^{(m)}(f)\}$ is to make $\{\psi_k^{(m)}(f) = \sqrt{W(f)}\varphi_k^{(m)}(f)\}$ become orthogonal set in the specific region $(0, f_{\max})$. Table I gives an example of four standard orthogonal basis with $W(f)$ defined for HRTF spectrum modeling purpose.

1) Comparison of Different Orthogonal Basis Functions: This leaves the issue about the choice of the basis functions for each m. In principle any complete orthogonal set of functions on a finite interval can be adapted for the purpose and provide an exact representation and in that sense are equivalent. However, under *truncation* to a certain number of terms, some orthogonal sets will perform better than others. That is, under truncation, different orthogonal sets of functions are not equivalent and there will exist preferred choices.

It is our interest to find the most efficient orthogonal set for truncation (7) based on real data (Fig. 1), which is the first novel element. The second novel element is to restrict the choice of candidates to well studied classical closed form orthogonal sets defined on a finite interval (possibly with different weighting functions) for which there is a rich set of choices including: Complex Exponentials, Legendre Polynomials, Chebyshev Polynomials, and the Fourier–Bessel series of various integer orders. Table I summaries the representation of the $A_m(f)$ using four families of these orthogonal functions (refer to [20] and [21] for the detailed reference of these orthogonal functions). To choose the most effective basis function, the four functions are applied to model $A_m(f)$ calculated from the MIT

TABLE I
CANDIDATE CLOSED FORM ORTHOGONAL FUNCTIONS $\varphi_k^{(m)}(f)$ DEFINED ON $f \in (0, f_{\max})$ FOR MODELING THE HRTF SPECTRUM COMPONENTS $A_m(f)$ WITH THE ARGUMENT NORMALIZED AS $f' = f/f_{\max}$

Orthogonal Functions (general form)	Interval	Weight $W(x)$	Modified Basis functions $\varphi_k^{(m)}(f)$	h_k Equation (9)
Complex Exponentials $e^{j\pi kx}$	$x \in [-1, 1]$	1	$e^{j\pi k(2f'-1)}$	2
Legendre Polynomials $P_k(x)$	$x \in [-1, 1]$	1	$P_k(2f'-1)$	$\frac{1}{2k+1}$
Chebyshev Polynomials $U_k(x)$	$x \in [-1, 1]$	$(1-k^2)^{\frac{1}{2}}$	$U_k(2f'-1)$	$\frac{\pi}{2}$
Fourier Bessel series of order ℓ $J_\ell(\beta_k^{(\ell)}x)$	$x \in (0, 1)$	1	$J_\ell(\beta_k^{(\ell)}f')$	$\frac{2}{J_{\ell+1}(\beta_k^{\ell})^2}$

data [15] and the percent mean-square approximation error is calculated from the Q sampling frequencies, i.e.,

$$e = \frac{\sum_{q=1}^{Q} \|A_m(f_q) - \widehat{A_m}(f_q)\|^2}{\sum_{q=1}^{Q} \|A_m(f_q)\|^2} * 100\% \qquad (10)$$

where $\widehat{A_m}(f_q)$ is the approximated HRTF spectrum components. It is shown in Table II that the Fourier–Bessel series (note that different orders of Fourier–Bessel series are available and they may differ from different m; but here we use the zeroth-order Bessel function of the first kind $J_0(\cdot)$ to model the $A_m(f)$ for all m) can provide a better match to the measured HRTFs compared with the Complex Exponentials under the same expansion order $K = 100$, while the approximation made by Legendre and Chebyshev Polynomials has relatively large errors using the same truncation order.[2]

The observed efficiency of the Fourier-Bessel series can be explained by referring to Fig. 1. The energy of the HRTF frequency components does not equally spread over the Fourier series expansion for each order m, which is described as *butterfly shape* of HRTF spectrum in [22], [18] (the plots in Fig. 1 give half the butterfly shape corresponding to positive frequencies). The explanation for the special shape of the spectrum is as follows. For low frequency f, the large bandwidth requires small spatial Fourier series (m) support since the waves are spatially varying slowly. For increasing frequency, the butterfly shape is widening due to large spatial Fourier series support. It can be seen that the spatial Fourier series support increases with frequency as the smaller wavelength indicates faster changes along the azimuthal angle domain. Therefore, most of $A_m(f)$ energy is present in a butterfly shaped region; and outside the region, the energy is greatly reduced. This provides an insight why Complex Exponentials, Legendre, and Chebyshev polynomials are less efficient. In contrast, the Bessel functions of the first kind are better matched to the observed distribution of energy over frequency for various orders.

The Bessel functions $J_\ell(\cdot)$ for $\ell \geq 1$ have a high pass character. That is, the magnitude of $J_\ell(\cdot)$ for $\ell \geq 1$ is approximately zero for smaller arguments. And this zero value interval grows as ℓ increases. As shown in Fig. 1, the resemblance between the patterns of $A_m(f)$ and Bessel functions explain the reason why the Fourier–Bessel series can match the HRTF with highest accuracy for a given truncation order. Thus, in this paper, the Fourier–Bessel series are proposed to model the HRTFs spectral components $A_m(f)$ dependent on m as explained next.

2) Fourier–Bessel Series: The Fourier–Bessel series make use the orthogonality between Bessel functions of the first kind for a specific order ℓ on the interval $(0, 1)$ to expand a general function [21]. The Fourier–Bessel series expansion of the frequency components of HRTFs is given as

$$A_m(f) = \sum_{k=1}^{\infty} C_{mk} J_\ell\left(\frac{\beta_k^{(\ell)}}{f_{\max}} f\right) \qquad (11)$$

where $\beta_1^{(\ell)}, \beta_2^{(\ell)}, \ldots, \beta_k^{(\ell)}$ are the positive roots of $J_\ell(x) = 0$ and ℓ is the specific order of the Bessel function of the first kind. C_{mk} are complex coefficients whose dependence on ℓ has been suppressed. The coefficients of the Fourier–Bessel series expansion can be determined from

$$C_{mk} = \frac{2}{f_{\max}^2 \left(J_{\ell+1}\left(\beta_k^{(\ell)}\right)\right)^2} \int_0^{f_{\max}} f A_m(f) J_\ell\left(\frac{\beta_k^{(\ell)}}{f_{\max}} f\right) df. \qquad (12)$$

As the choice of the Fourier–Bessel series has an infinite number of possibilities, we need to decide which specific order ℓ to use and how this depends on m. The previous modeling comparison among different orthogonal functions (in Table II) shows that the zeroth-order Bessel function of the first-kind $J_0(\cdot)$ can achieve very accurate approximations. This is due to the dominance of $A_0(f)$ and, among all orders, only $J_0(\cdot)$ can give very efficient approximation to $A_0(f)$ in the limit as $f \to 0$. However, $J_0(\cdot)$ is not the optimum representation for higher order $A_m(f)$ when $m \neq 0$. Noticing that high order Fourier–Bessel series can match high order frequency components, we believe the representation with ℓ increasing with m is more effective. In this paper, we have not mathematically derived a specific formula relating m and ℓ, but we propose a simple linear relationship $\ell = |m|$ which avoids the significant triangular near null energy regions seen in Fig. 1. Under this heuristic, the mse of approximating $A_m(f)$ is 1.53% for $K = 100$, which is the lowest

[2]Note the approximation errors of these two functions will also be arbitrarily small when a large number of terms are included, e.g., $K > 200$.

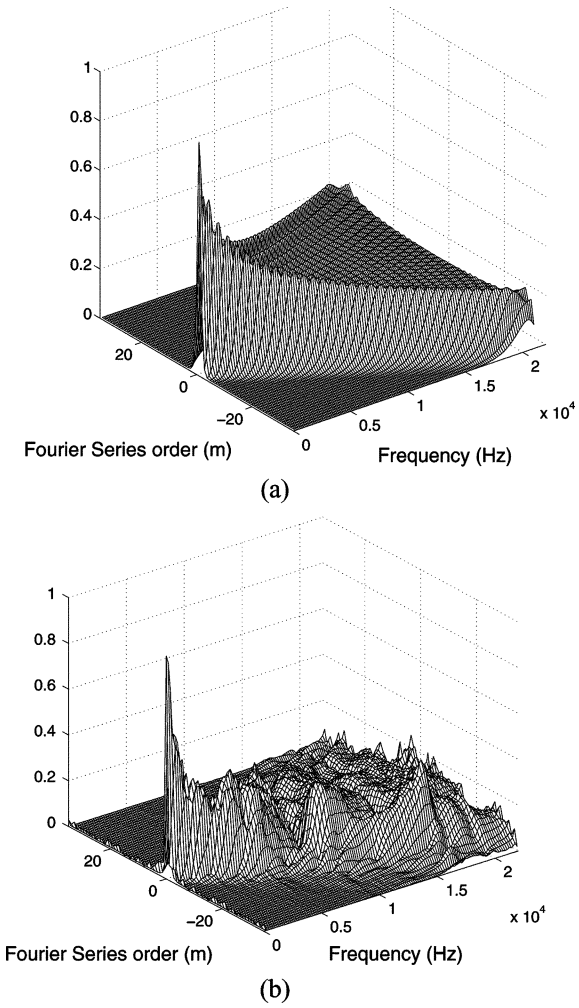

Fig. 1. Mesh plots of the magnitude spectrum of $A_m(f)$, $|m| \leq 35$ calculated from (a) analytical HRTFs from the spherical head model and (b) measured HRTFs from MIT data. Note given 72 samples (source positions) equally spaced along the circle with $\Delta\phi = 5^\circ$ in both data sets, the highest Fourier series order we can unambiguously derive is $M = 35$ with $f_{\max} = 21$ kHz.

TABLE II
PERCENT MEAN-SQUARE-ERROR OF APPROXIMATING $A_m(f)$ USING FOUR ORTHOGONAL FUNCTIONS WITH TRUNCATION NUMBER $K = 100$ BASED ON MIT MEASUREMENT DATA

Name of Orthogonal Functions	Percent MSE
Complex Exponentials	3.19%
Legendre Polynomials	79.40%
Chebyshev Polynomials	25.97%
Fourier Bessel series ($\ell = 0$)	2.27%
Fourier Bessel series ($\ell = \|m\|$)	1.53%

approximation error in Table II. The real measurements reconstruction performance in Section III establishes that this choice works very well, in the sense of high approximation accuracy and significantly reduced number of parameters (the physical basis of this simple formula will be left as on open problem).

3) Proposed Efficient Continuous HRTF Model: The above development leads to the HRTF functional model in the frequency domain written as

$$H(f,\phi) = \sum_{m=-\infty}^{\infty} \sum_{k=1}^{\infty} C_{mk} J_{|m|}\left(\frac{\beta_k^{(|m|)}}{f_{\max}} f\right) e^{jm\phi} \tag{13}$$

where

$$C_{mk} = \frac{1}{\pi f_{\max}^2 \left(J_{|m+1|}\left(\beta_k^{(|m|)}\right)\right)^2} \int_0^{f_{\max}} \int_{-\pi}^{\pi} f H(f,\phi) J_{|m|}\left(\frac{\beta_k^{(|m|)}}{f_{\max}} f\right) e^{-jm\phi} d\phi df. \tag{14}$$

Equation (14) illustrates how to calculate the model parameters from a continuous HRTF function. When using experimentally measured HRTFs, the model coefficients C_{mk} can be calculated using the left Riemann sum to approximate the integral (14). Section III shows how to choose the truncation of (13) up to provide accurate approximations to the original HRTF measurements.

The HRTF representation (13) exhibits three significant advantages.

- First, the representation has well studied closed form orthogonal sets as basis functions, which can make the HRTF approximation easily implemented and model parameters C_{mk} simply computed using (14).
- Second, using continuous basis functions, the proposed model can achieve HRTF reconstruction at any frequency point for an arbitrary azimuth angle. That is, it provides natural interpolation and extrapolation.
- Third, the basis function choice is empirically guided but the basis functions are independent of the data. As the basis is same for all subjects, the model coefficients C_{mk} carry all the information about the individuality. Thus, the model has capability to represent the individualized HRTF by assigning a subject specific set of parameters to the model through (13).

C. HRIR Representation

In this subsection, the equivalent head-related impulse response (HRIR) representation is developed. The model is based on the inverse Fourier transform of the Fourier–Bessel series, the function used to represent the frequency variable of the HRTF in (13).

The Fourier Transform of $J_\ell(t)$ is given by [23] as

$$\mathcal{F}\{J_\ell(t)\} = \frac{2(-j)^\ell T_\ell(w)}{\sqrt{(1-w^2)}}, \quad \text{for } |w| < 1 \tag{15}$$

and zero otherwise, where $T_\ell(\cdot)$ is the Chebyshev function of the first kind and is defined as

$$T_\ell(w) = \cos(\ell \cos^{-1}(w)), \quad |w| < 1. \tag{16}$$

Thereafter, the HRTF model in the time domain, the so-called HRIR model, can be developed as

$$h(t,\phi) = \sum_{m=-\infty}^{\infty} \sum_{k=1}^{\infty} C_{mk} e^{jm\phi} \sigma_m \left(\beta_k^{(|m|)}, t\right) \tag{17}$$

where based on (15)

$$\begin{aligned} &\sigma_m \left(\beta_k^{(|m|)}, t\right) \\ &\quad \triangleq \mathcal{F}^{-1} \left\{ J_{|m|}\left(\frac{\beta_k^{(|m|)}}{f_{\max}} f\right) \right\} \\ &\quad = \begin{cases} \dfrac{(-j)^{|m|} T_{|m|}\left(\frac{f_{\max}}{\beta_k^{(|m|)}} t\right)}{\pi \sqrt{\left(\frac{\beta_k^{(|m|)}}{f_{\max}}\right)^2 - t^2}}, & \text{if } 0 \le t < \frac{\beta_k^{(|m|)}}{f_{\max}} \\ 0, & \text{if } t \ge \frac{\beta_k^{(|m|)}}{f_{\max}} \end{cases} \end{aligned} \tag{18}$$

Now, both HRTF and HRIR models (13) and (17) are continuous representations and have data independent basis functions to represent the spectral (or time) and spatial dependence. Thus, for each individualized measurement set only a coefficient matrix needs to be saved and used to further reconstruct the HRTF (or HRIR) for a specific listener.

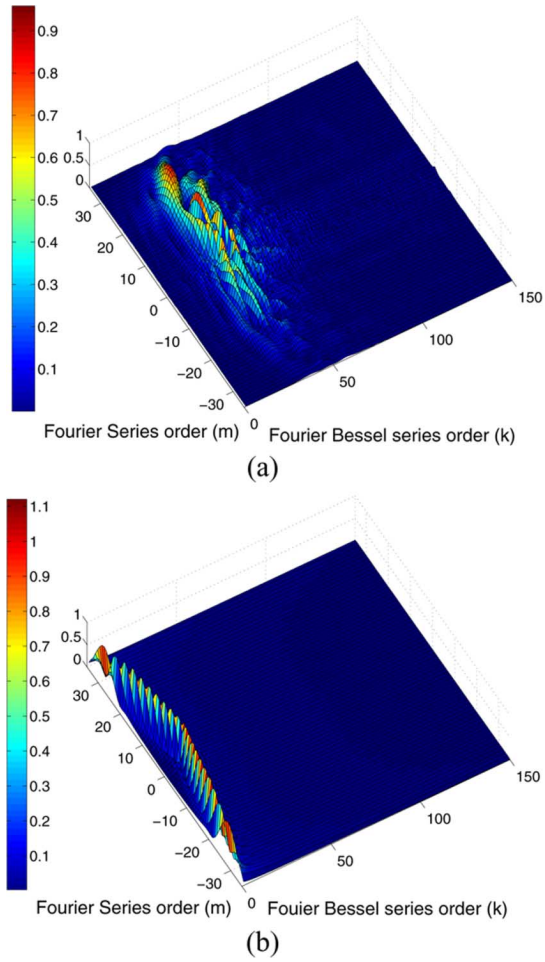

Fig. 2. Magnitude of model coefficients C_{mk} solved from (a) MIT HRTFs for far field $r = 1.4$ m and (b) analytically simulated HRTFs from the spherical head model for near field $r = 0.7$ m. The model coefficients energy are kept in low order components.

III. Model Validation

A. Methodology

The effectiveness of the proposed HRTF continuous representation is investigated by comparing the experimentally measured (or analytically simulated) HRTFs with model reconstructed results. In this section, three sets of data are employed in evaluation: 1) the MIT data acquired using a KEMAR manikin [15], 2) some analytical solutions from the spherical head model [16], and 3) a 50-subject HRTF data set [17]. A brief summary of these three sets of data is given below.

The MIT measurements were made in an anechoic chamber; and both the "small" and "large" pinna models were tested on the KEMAR. The measurements are the head related impulse response in the time domain at the 44.1 kHz sampling rate and each response is 512 samples long. The speakers were at a distance of 1.4 m away from the head center (far field measurements). In the horizontal plane, a full 360° of azimuth was sampled in equal sized increments (5° approximately). The azimuthal angle ϕ is from 0° to 360°. The direct front and back directions at ear level are defined as 0° and 180° azimuth; while the right and left sides are defined as $\phi = 90°$ and $\phi = 270°$.

The analytically simulated HRTFs is based on the spherical head model presented in [16]. We generate the HRTFs along a circle of 72 points on the horizontal plane at a distance of 0.7 m from the head center (near field values).

The HRIR measurements performed at IRCAM and AKG (two professional audio companies) anechoic chamber include 50 subjects [17]. A 8192 point logarithmic sweep (sampling frequency is also 44.1 kHz) is used as the test signal. For each subject, the HRTFs are measured at 24 equally spaced points along the azimuth angle domain (15° intervals). We use these data sets especially to check the favorable statistical performance of the model reconstruction performance in Section III-B1.

1) Choice of Truncation Number: The model coefficients C_{mk} are calculated for both MIT data and analytically simulated HRTFs as shown in Fig. 2. First, the Fourier series expansion (3) are solved up to and including the order 35 $(M = 35)$, which is the maximum solvable order upper bounded by the number of discrete measurements (72 azimuthal samples in the horizontal plane) and corresponds to the maximum measurement frequency $f_{\max} = 21$ kHz based on (4).[3] Then, the HRTF

[3]We should state that according to (4), for MIT data of $f_{\max} = 22.05$ kHz (sampling rate of 44.1 kHz), we need approximate 74 samples with $\Delta\phi = 4.86°$ for accurate reconstruction.

spectrum $A_m(f)$ are expanded with Fourier–Bessel series up to order $K = 150$.

The structure of the model coefficients guides us on how to choose the truncation number of the proposed model (13) and (17). Especially, the two sets of data correspond to two types of HRTFs, i.e., the far field HRTFs from the MIT data and the near field data from analytical simulated HRTFs. The plots in Fig. 2 show that the model coefficients energy are mainly kept in the low order expansion components so that we just need to keep these coefficients for the HRTF representation. To decide the appropriate truncation for series representation, the criteria is to choose the truncation number M and K as the lowest orders at which at least 90% of the total "energy" is contained in the approximation. Thus, the truncation number for the far field HRTF is chosen as $M = 25$ and $K = 100$ and the near field HRTF model is truncated to $M = 35$, $K = 30$ accordingly [note the large spatial variance of the near field HRTFs requires more Fourier series terms in the model as shown in Fig. 2(b)]. There is no doubt that keeping more of the series coefficients will increase approximation accuracy. Section III-C shows the shape of the synthesized responses will go arbitrarily close to that of the measurements by increasing Fourier series order M and Fourier–Bessel series order K.

2) Error Metric: Given the efficient representation with the lowest truncation numbers, the performance of the model in accurately approximating experimentally measured HRTFs has been assessed in both frequency and time domains. The error metric is defined as the percent mean-square-error (mse) in the magnitude and phase spectrum (or the impulse response) at the measured azimuthal locations

$$e_i = \frac{\sum_{n=1}^{N} \|H(f_n, \phi_i) - \widehat{H}(f_n, \phi_i)\|^2}{\sum_{n=1}^{N} \|H(f_n, \phi_i)\|^2} \times 100\% \qquad (19)$$

where for each azimuth, HRTFs are measured at N frequency points $f_1, f_2, \ldots, f_N$, $H(f_n, \phi_i)$ is the measured HRTF (or HRIR) at the nth frequency point and ith azimuth, and $\widehat{H}(f_n, \phi_i)$ is the reconstructed HRTF (or HRIR).

B. Continuous Model Performance

The performance of the proposed model is investigated in two ways: 1) in reconstruction, in Section III-B1, using the measured data (the same data used to determine the model parameters), and 2) in interpolation, using the cross-validation technique, in Section III-B2.

1) HRTF Model Reconstruction Performance: The first example to assess the HRF model performance is to reconstruct the 72 MIT measured (or analytically simulated) HRTFs spaced by 5° in the horizontal plane. The error analysis of the frequency model is conducted in each azimuthal angle to both the magnitude and unwrapped phase responses. Fig. 3 shows a direct and qualitative comparison of the measured HRTF spectrum from MIT left and right ears and model reconstructed HRTFs in the frequency domain. The errors in fitting the HRTF measurements (or analytically simulated HRTFs) are further analyzed at each of the measured positions. The distribution of errors across all positions are presented in Figs. 4(a) and 5(a) for both MIT

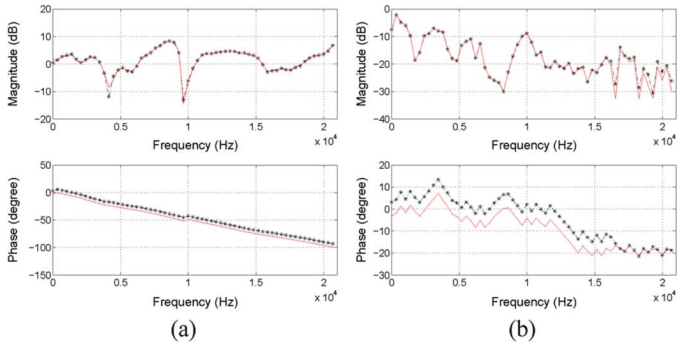

Fig. 3. Example of MIT measured and model reconstructed HRTFs using the frequency domain model (a) left ear measured HRTFs at $\phi = 240°$ and (b) right ear measured HRTFs at $\phi = 250°$. Measurements: star dot-dashed line, Reconstruction: solid line.

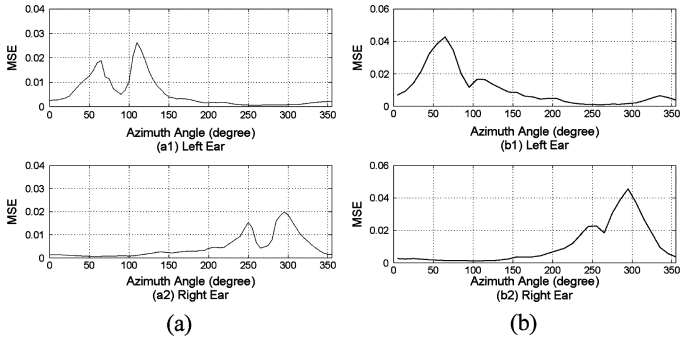

Fig. 4. Approximation error distribution as a function of the source position (azimuthal angle ϕ) for the MIT HRTFs. (a) reconstruction error distributions and (b) interpolation error distributions. The top plots correspond to the left ear and the bottom plots correspond to the right ear.

measurements and analytical simulated HRTFs. Note that in both data sets, a source located at 90° azimuth is directly across from the right ear (the left ear is the shadowed ear) and a source located at 270° is directly across from the left ear (the right ear is the shadowed ear). The calculated error in the model fitting leads to two observations. First, in general, the continuous functional model provides a good fit to the HRTF with significantly reduced number of parameters, for example, the mean-square reconstruction error is about 1.78% for the MIT measurements and 0.5% for analytically simulated HRTFs. Second, given the accurate approximation, the synthesis of HRTFs is better at the source-facing side of the head compared to the head's shadowed side.

The reconstruction of the 50-subject HRTF data set shows very good statistical results on the reconstruction performance of the proposed approach. First, given 24 samples equally spaced by 15° in the horizontal plane for each subject, according to the sampling and reconstruction analysis in Section II-A, we can accurately reconstruct the human subject HRTF data to 7.2 kHz. We check the reconstruction errors across all positions for each human subject. Fig. 6 shows the statistical reconstruction error distribution among 50 subjects with the mean value overlaid. It can be seen that the maximum reconstruction error is under 6% and the mean value distribution is very like the error distribution of the KEMAR

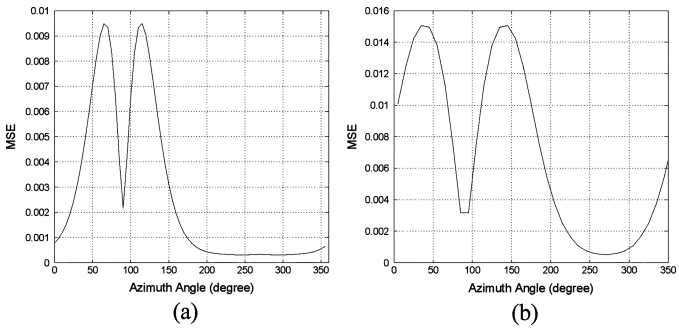

Fig. 5. Approximation error distribution as a function of the source position (azimuthal angle ϕ) for the analytically simulated HRTFs. (a) reconstruction error distributions and (b) interpolation error distributions. Note we only plot the left ear because for the spherical head model; the simulated HRTFs at left and right ears are symmetrical and have identical error performance.

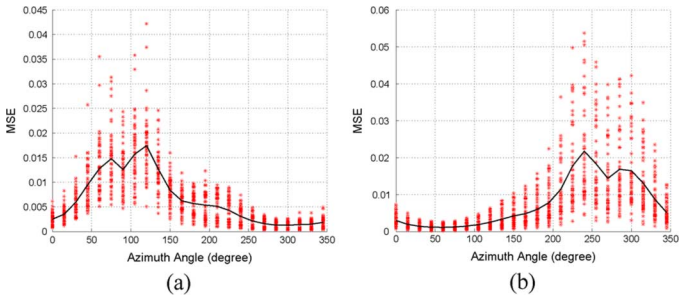

Fig. 6. Statistical performance of the 50-subject HRTF data reconstruction error distribution for both (a) left ear and (b) right ear with the mean value overlaid (solid line).

data. This corroborates our claim that we have an efficient data independent continuous functional model.

2) HRTF Model Interpolation Performance: In applications, HRTF interpolation may be required over angles not measured. In this section, the interpolation performance of the proposed continuous functional model is demonstrated. In order to use the available measurements to examine the interpolation performance, we employ cross-validation. We partition the 72 5° spaced MIT (or analytically simulated) HRTF data points in the horizontal plane into two groups with each having 36 data sets spaced by 10°. The first group is used to determine the model parameters C_{mk} which are used to synthesize/interpolate the HRTF at the angles corresponding to the second group. The difference indicates the interpolation performance. In addition, the second group locations, which are at the midpoints of those used to determine the modal parameters, represent the locations of maximum interpolation error.

For interpolation, the most important implementation issue is to record the sound field with enough samples. As shown in Section II-A, for the sound field reconstruction along the circle, the maximum measurement frequency $f_{\max}$ determines the highest magnitude of the Fourier expansion coefficients $M = \lceil 2\pi a f_{\max}/c \rceil$ as given in (4). Hence, the maximum angular sampling interval is determined by

$$\triangle\phi^{\circ}_{\max} = \frac{360}{N_{\min}} = \frac{360}{2\lceil 2\pi a f_{\max}/c \rceil + 1} \tag{20}$$

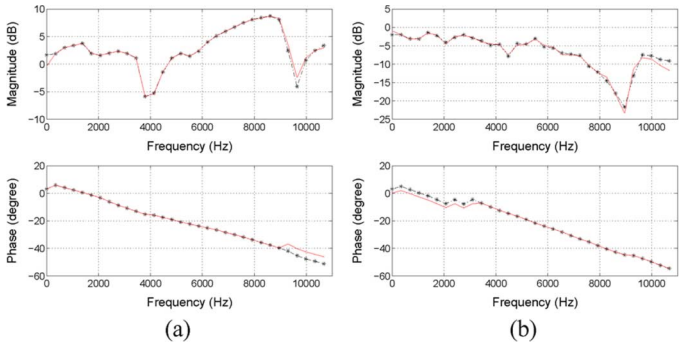

Fig. 7. Example of model interpolated HRTFs with original MIT measurements overlaid. (a) Left ear HRTF at $\phi = 255^{\circ}$ and (b) Right ear HRTFs at $\phi = 195^{\circ}$. Measurements: star dot-dashed line. Interpolation: solid line.

where a is the radius of the head. Equation (20) is based on the fact that we need $N \geq (2M + 1)$ equally spaced samples along the circle. It shows the maximum azimuthal angular spacing $\triangle\phi_{\max}$ is related to the maximum measurement frequency $f_{\max}$. For example, when the sampling frequency is 44.1 kHz (as used in MIT data set), the maximum angular sample interval is approximately 4.86°.

In our validation test of the MIT data, the second group of data sets is spaced by 10°; and we can only accurately interpolate the HRTFs up to 10.8 kHz. Fig. 7 shows the interpolated MIT left and right ear HRTFs with original measurements overlaid. The interpolation error performances in Figs. 4(b) and 5(b) demonstrate that our model is able to interpolate over large angles with high accuracy (the maximum error is about 4.3% for MIT data and 1.52% for the analytical solutions). In addition, the maximum interpolation error is also located at the contralateral ear.

3) Time-Domain HRIR Model Performance: The measured MIT HRTF database is applied to the time-domain HRIR model (17). The HRIR analysis is conducted in a similar way to the HRTF analysis, except the error is calculated for each time slot and for each azimuthal angle. Fig. 8 shows examples of the HRIR measurements with their model reconstruction and interpolation overlaid. The closeness of the reconstructed and experimental responses corroborates the effectiveness and high accuracy of the time-domain HRIR model.

C. Model Performance Analysis

Section III-B shows that the synthesis of HRTFs is usually better at the source-facing side of the head than at the head's shadowed side. Three factors contribute to this phenomenon. The first factor is that the comparative poorer signal-to-noise ratio (SNR) at the head shadowed side. The signal level at the head shadowed side is lower than that at the source facing side, resulting the SNR in the measurement is relatively poorer at the head shadowed locations. The second factor is the high complexity of the HRTF at the head's shadowed locations. Because of the diffraction around the head, the contralateral sounds have more variations. This results in the spectral shapes that are more complicated and more difficult to model. Third, the head shadowing may cause HRTF spectral components has energy outside of the butterfly shape. Hence, the truncated model can spatially not predict the large spatial variation including diffraction and

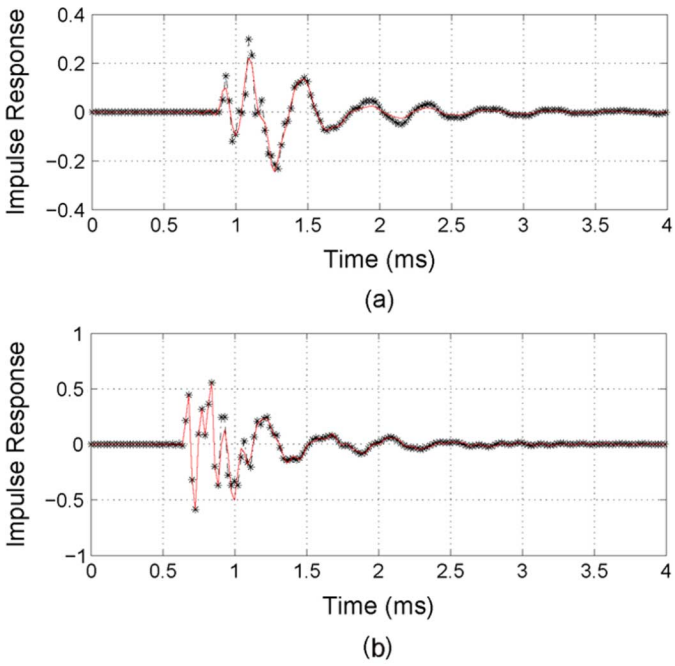

Fig. 8. Example of model reconstructed and interpolated HRIR from MIT data with original measurements overlaid. (a) reconstructed MIT left ear HRIRs at $\phi = 180°$ and (b) interpolated MIT right ear HRIRs at $\phi = 85°$. Measurements: star dot-dashed line. Synthesis results: solid line.

head shadowing. If we choose a larger M and K, we may specifically improve the reconstruction over those regions.

Beside the diffraction effects, the Shaw's "bright spot" phenomenon can also be observed [24], that is, HRTFs corresponding to contralateral locations which lie in the direct shadow of the head have relatively higher-amplitudes or a local maximum energy injected. Hence, both ears HRTF reconstructions (Figs. 4–6) have the local least approximation error with azimuths of 90° and 270°, respectively.

We summarize the performance of the proposed model in the following four aspects. First, the proposed continuous functional HRTF and HRIR models provide accurate approximations to the experimental measurements. The shape of the synthesized responses can get arbitrarily close to that of the measurements by increasing the Fourier series and Fourier–Bessel series harmonic orders. Our calculation shows that the reconstruction error of MIT data decreases to 0.52% when $M = 35, K = 150$. Second, as given in (14), each of the individualized HRTF is transformed to a set of coefficient. This coefficient set is much smaller in size compared to the original HRTF data samples. For example, original MIT databases has 36 864 sample points (72 directions and 512 frequency samples for each azimuth); now the transformed coefficient just has 5100 entries $(M = 25, K = 100)$. The data needs to be saved has been reduced by nearly 85%. Third, the computation of the model coefficients is easily implemented as the basis functions are well known orthogonal functions and their value can be calculated in advance. Hence, we can use the model to decompose any HRTF through (14) and resynthesize the data from the series representation of (13) for any specific position. Fourth, the proposed model can be regarded as noise discriminated as the

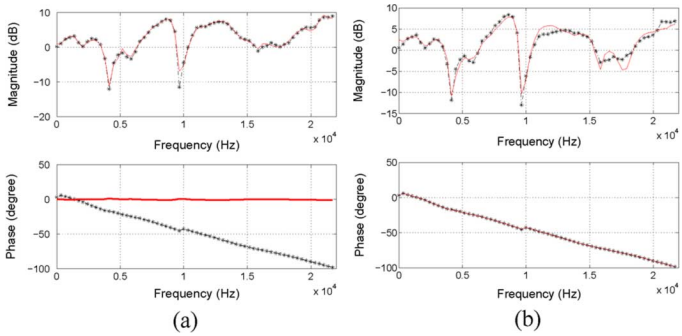

Fig. 9. (a) Example of PCA model reconstructed MIT left ear HRTFs at $\phi = 240°$ and (b) example of KLE model reconstructed MIT left ear HRTFs at $\phi = 240°$. Measurements: star dot-dashed line. Reconstruction: solid line.

TABLE III
HORIZONTAL DOMAIN HRTF RECONSTRUCTION USING PCA, KL EXPANSION, AND CONTINUOUS MODEL

Methods	Mag. spectrum appro. error	Complex resp. appro. error	Number of parameters
PCA	2.00%	Not applicable	4672
KL expansion	4.84%	2.46%	4672
Continuous model	5.29%	2.63%	4900

two orthogonal functions we choose as basis has structural similarities to the HRTF being analyzed. Thus, the unwanted components (noise or distortion) will not be represented with the same accuracy as the signal interested. For example, the noise components of high spatial bandwidth $(|m| > M)$ is removed from (5) and the noise with frequency components outside the butterfly shaped region will be significantly reduced from (12).

IV. COMPARISON WITH STATISTICAL MODELING METHODS

The statistical modeling methods, specifically PCA and the (discrete) KL expansion, are applied in an effort to reduce the redundancy (correlation) of the HRTF data set. Both methods apply eigenanalysis to a sample frequency covariance matrix constructed from the measured HRTFs to yield a set of the principle components as basis spectral functions. The PCA model, as defined in [8], constructs the sample covariance matrix from the HRTF magnitude spectrum[4] while the KLE method defined in [10] uses the complex valued HRTFs after the directionally independent frequency dependence is removed. Since the eigenvectors with largest eigenvalue of the frequency covariance matrix are chosen as principle components, the statistical modeling methods achieve optimal lower dimensional orthogonal representations. These models rely on the empirical data to provide an optimal representation. However, for these representations, the basis is discrete and data dependent which can have disadvantages.

The same HRTF reconstruction on the MIT data procedure is performed using the PCA and KLE methods. To capture 90% of the variance in the spectrum, we extract eight principal components as basis vectors. Fig. 9 shows the reconstructed HRTF magnitude and phase spectrums with original data overlaid. The

[4]The phase is reconstructed based on a minimum phase assumption of the HRTF corresponding to each magnitude function and another simple time delay filter contains the interaural time difference information.

example is for the MIT left ear HRTF reconstruction at azimuthal angle $\phi = 240°$ (Fig. 3(a) gives our continuous model reconstruction performance). The PCA model gives nearly identical approximation to the HRTF magnitude spectrum but ignores the measured phase; while the discrete KLE is to model the complex valued HRTFs.

Table III summaries the numerical HRTF reconstruction performance using the two statistical methods and our proposed continuous model. There are several observations. First, even though there are a small number of basis vectors in the statistical models, the two sets of functions (basis and weights) in these models actually require similar number of parameters as our proposed continuous model.[5] This shows that our continuous model also achieves nearly an optimal lower dimensional representation. Second, two kinds of approximation errors are investigated for the three models. As PCA is formulated to analyze the HRTF magnitude function, it assuredly gives the least magnitude spectrum approximation error. However, there is no qualitative way to show the complex valued HRTF reconstruction using the PCA method due to the minimum phase assumption (there is only a monaural phase between two ears to model the interaural phase difference). While it is true that for minimum phase systems knowledge of the magnitude response is sufficient to infer the phase response, this does not imply the PCA applied to the magnitude only of a minimum phase system has any claim to optimality when full complex data modeling is desired. The proper formulation needs to apply PCA to the full complex response like the discrete KLE method. Finally, as for a fair comparison between KLE and the proposed continuous model, the relatively bigger approximation error using the continuous model shows that the discrete statistical method is truly the optimal description for this data set.[6] However, the disadvantage of lack of strict optimality of our approach is countered by the universality (data independent and measurement grid independence of the basis) and the continuous nature of the basis functions (eliminating the need for interpolation).

V. Conclusion

A continuous functional model was developed for the horizontal plane HRTF representation. A Fourier series was used to model the spatial dependence of the HRTF while the HRTF spectral components were represented by the continuous Fourier–Bessel series of various orders. The advantage of this model is that both Fourier series and Fourier–Bessel series are well-studied complete closed form orthogonal functions and can be readily applied to accurately model the HRTF. By applying these two sets of continuous functions as basis functions, each individualized HRTF is expressed by a coefficient matrix in a form amenable to truncation to a significantly reduced number of parameters. The model can be used to synthesis/interpolate HRTFs at arbitrary azimuth and frequency points. Moreover, being independent of the data, the basis functions will remain unchanged for all subjects and measurement setups.

The development of a 3-D HRTF representation following the same methodology simply replaces the conventional Fourier series expansion with an expansion using spherical harmonics to represent the spatial dependence (adding elevation to the azimuth angular variable). A similar parametric spectral analysis can be performed. The work in [25] found similar butterfly shaped HRTF spectral component of 2-D HRTFs at other elevations. Thus, the Fourier-Bessel series may be used as an efficient functional analysis of the 3-D HRTF spectrum. However, such a study would require more complete measurement data that is currently unavailable.

We also need to state that the current approach is driven from a phenomenological perspective where we are dealing with an explanation of empirical observations at the technical level. Psychoacoustic validation has to be performed in the future. Besides the technical measure presented in this work, we also need subjective validation to consummate the error bounds and truncation orders given in the paper.

[5]From [8], the basis functions size is $p \times q$ and the weights size is $q \times N$, where p is total number of frequency samples, q is the number of principal components, and N is the total number of transfer functions. In the MIT data, $p = 512, q = 8, N = 72$; thus, the total number of parameters is $512 \times 8 + 8 \times 72 = 4672$.

[6]For a fair comparison, we reduce the parameters of the proposed model to 4900 with average reconstruction error of 2.63%. Remember previously in Section III-B1 the proposed model achieves reconstruction error of 1.78% with 5100 parameters.

References

[1] J. Blauert, *Spatial Hearing-Revised Edition: The Psychophysics of Human Sound Localization*. Cambridge, MA: MIT Press, 1996.

[2] K. Hartung, J. Braasch, and S. J. Sterbing, "Comparison of different methods for the interpolation of head-related transfer functions," in *Proc. 16th Audio Eng. Soc. Int. Conf. Spatial Sound Reproduction*, Rovaniemi, Finland, Apr. 1999, pp. 319–329.

[3] S. Carlile, C. Jin, and V. V. Raad, "Continuous virtual auditory space using HRTF interpolation: Acoustic and psychophysical errors," in *Proc. Int. Symp. Multimedia Inf. Process.*, Sydney, Australia, Dec. 2000, pp. 220–223.

[4] M. Matsumoto, S. Yamanaka, M. Tohyama, and H. Nomura, "Effect of arrival time correction on the accuracy of binaural impulse response interpolation," *J. Audio Eng. Soc.*, vol. 52, no. 1/2, pp. 56–61, Feb. 2004.

[5] 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.

[6] A. Kulkarni and H. S. Colburn, "Infinite-impulse-response models of the head-related transfer function," *J. Acoust. Soc. Amer.*, vol. 115, no. 4, pp. 1714–1728, Apr. 2004.

[7] 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.

[8] D. J. Kistler and F. L. Wightman, "A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction," *J. Acoust. Soc. Amer.*, vol. 91, no. 3, pp. 1637–1647, Mar. 1992.

[9] 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.

[10] J. S. Chen, B. D. Van Veen, and K. E. Hecox, "A spatial feature extraction and regularization model for the head-related transfer function," *J. Acoust. Soc. Amer.*, vol. 97, no. 1, pp. 439–452, Jan. 1995.

[11] C. Jin, P. Leong, A. Corderoy, and S. Carlile, "Enabling individualized virtual auditory space using morphological measurements," in *Proc. Int. Symp. Multimedia Inf. Process.*, Sydney, Australia, Dec. 2000, pp. 235–238.

[12] 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.

[13] C. I. Cheng and G. H. Wakefield, "Introduction to head-related transfer functions (HRTFs): Representations of HRTFs in time, frequency and space," *J. Audio Eng. Soc.*, vol. 49, no. 4, pp. 231–249, Apr. 2001.

[14] W. Zhang, T. D. Abhayapala, and R. A. Kennedy, "Horizontal plane HRTF reproduction using continuous Fourier-Bessel functions," in *Proc. 31st Audio Eng. Soc. Int. Conf.: New Directions in High Resolution Audio*, London, U.K., Jun. 2007.

[15] 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.
[16] 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.
[17] O. Warusfel, "Listen HRTF Database," May 2003 [Online]. Available: http://recherche.ircam.fr/equipes/salles/listen/index.html
[18] T. Ajdler, C. Faller, L. Sbaiz, and M. Vetterli, "Sound field analysis along a circle and its applications to HRTFs interpolation," *J. Audio Eng. Soc.*, vol. 56, no. 3, pp. 156–175, Mar. 2008.
[19] J. G. Proakis and D. K. Manolakis, *Digital Signal Processing: Principles, Algorithms, and Applications*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, Oct. 1995.
[20] L. E. Franks, *Signal Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1969.
[21] W. Kaplan, "Fourier-bessel series," in *Advanced Calculus*. Reading, MA: Addison Wesley , Sep. 1991, pp. 512–518.
[22] 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, vol. III, pp. 273–276.
[23] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*. New York: Dover, Jun. 1965.
[24] 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, pp. 455–490.
[25] X. L. Zhong and B. S. Xie, "Spatial characteristics of head-related transfer function," *Chinese Phys. Lett.*, vol. 22, no. 5, pp. 1166–1169, May 2005.

Wen Zhang (S'07) received the B.E. degree in telecommunication engineering from Xidian University, Xi'an, China, in 2003, and M.E. degree (with first class honors) from the Australian National University (ANU), Canberra, in 2005. She is currently pursuing the Ph.D. degree in the Research School of Information Sciences and Engineering (RSISE), ANU.

Her research interests are spatial audio signal processing techniques and their applications in head-related transfer function (HRTF) measurement and modeling, virtual spatial audio systems, and acoustic source separation from HRTF cues.

Ms. Zhang is a student member of JASA.

Rodney A. Kennedy (M'88–SM'01–F'05) received the B.E. degree in electrical engineering from the University of New South Wales, Sydney, Australia, in 1982, the M.E. degree from the University of Newcastle, Newcastle, Australia, in 1985, and the Ph.D. degree in systems engineering from the Australian National University, Acton, in 1988.

He is currently a Professor and Director of Research in the College of Engineering and Computer Science, Australian National University, Canberra. He has been an active member of the IEEE, serving on a number of technical program committees of workshops and conferences, and was an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS. His research interests include digital and wireless communications, signal processing, spatial information systems and information theory, and acoustical signal processing.

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

He is an Associate Professor 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. 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 with the Research School of Information Sciences and Engineering, ANU. His research interests are in the areas of audio and acoustic signal processing, space–time signal processing for wireless communication systems, and array signal processing. He has supervised 22 research students and coauthored approximately 130 peer-reviewed papers. He is currently an Associate Editor for the *EURASIP Journal on Wireless Communications and Networking*.