Measurement and Modelling of Head-Related Transfer Function for Spatial Audio Synthesis

Wen Zhang

B.E. (Xidian University, China)
M.E. (Hons 1) (The Australian National University, Australia)

August 2010

A thesis submitted for the degree of Doctor of Philosophy
of The Australian National University

Applied Signal Processing Group
School of Engineering
College of Engineering and Computer Science
The Australian National University

Declaration

The contents of this thesis are the results of original research carried out by myself, under the supervision of A/Prof. Thushara D. Abhayapala, and Prof. Rodney A. Kennedy. These have not been submitted for a higher degree to any other university or institution.

Much of the work in this thesis has been published or has been submitted for publication in referee journal papers and conference proceedings. In some cases the conference papers contain material overlapping with the journal publications. The following is a list of these publications.

Journal Publications

1. W. Zhang, R. A. Kennedy, and T. D. Abhayapala, "Efficient continuous HRTF model using data independent basis functions: Experimentally guided approach", *IEEE Trans. Audio, Speech and Language Processing*, vol. 17, no. 4, pp. 819-829, May 2009.
2. W. Zhang, T. D. Abhayapala, R. A. Kennedy, and R. Duraiswami, "Insights into head-related transfer function: Spatial dimensionality and continuous representation", *The Journal of the Acoustic Society of America*, vol. 127, no. 4, pp. 2347-2357, Apr. 2010.
3. 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 and Language Processing*, (submitted Sep. 2010).

Conference Publications

1. W. Zhang, T. D. Abhayapala, and R. A. Kennedy, "Horizontal plane HRTF reproduction using continuous Fourier-Bessel functions", in *Proc. 31st Audio Engineering Society (AES) international conference on "New directions in high resolution audio"*, London, UK, June 2007, number 4, pp. 9 pages.
2. W. Zhang, R. A. Kennedy, and T. D. Abhayapala, "Signal estimation from incomplete data on the sphere", in *Proc. IEEE 9th Australian Communication Theory Workshop (AusCTW07)*, Christchurch, New Zealand, Feb. 2008, pp. 39-44.
3. W. Zhang, R. A. Kennedy, and T. D. Abhayapala, "Iterative extrapolation algorithm for data reconstruction over sphere", in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2008*, Las Vegas, USA, Apr. 2008, pp. 3733-3736.
4. R. A. Kennedy, W. Zhang, and T. D. Abhayapala, "Spherical harmonic analysis and model-limited extrapolation on the sphere: Integral equation formulation", in *Proc. 2nd International Conference on Signal Processing and Communication Systems*, Gold Coast, Australia, Dec. 2008, pp. 6 pages.
5. W. Zhang, T. D. Abhayapala, R. A. Kennedy, and R. Duraiswami, "Modal expansion of HRTFs: Continuous representation in frequency-range-angle", in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2009*, Taipei, Taiwan, Apr. 2009, pp. 285-288.
6. M. Zhang, W. Zhang, R. A. Kennedy, and T. D. Abhayapala, "HRTF measurement on KEMAR manikin", in *Proc. ACOUSTICS 2009 (Australian Acoustical Society)*, Adelaide, Australia, Nov. 2009, pp. 8 pages.

Wen Zhang
College of Engineering and Computer Science,
The Australian National University,
Canberra,
ACT 0200,
Australia.

DEDICATED
TO
MY FAMILY MEMBERS
WITH ALL MY LOVE

Acknowledgements

During My Ph.D. education at the Australian National University, I was fortunate to have A/Prof. Thushara Abhayapala and Prof. Rodney Kennedy as my supervisors.

Thushara introduced me into the research areas in audio and acoustics and provided valuable suggestions and guidance throughout. We had many useful discussions which deepened my understanding on acoustic signal processing and provided insights into my research. I take this opportunity to sincerely acknowledge his help, guidance and encouragement.

Rod trained me up in unit sphere and Hilbert space signal processing. I am deeply grateful to him for his patience, for spending many hours with me, explaining bits and pieces of fundamental signal processing knowledge. I learnt many things from him including good research and good writing skills. I thank him for his encouragement, guidance, insight and enthusiasm.

I am thankful to A/Prof. Ramani Duraiswami of the University of Maryland, College Park. Discussion with him helped me enormously in broadening my knowledge on the HRTF and the spatial audio applications.

I thank everyone in the Applied Signal Processing group for providing a congenial working environment. In particular, Ying Chen, Lin Luo, Karan Zhang, Aastha Gupta, Jennifer Wu, Sean Zhou, and Sandun Kodituwakku deserve special thanks for their friendly support. It is also my pleasure to acknowledge the help from Lesley Goldburg, Elspeth Davies, and Rob Gresham on various administrative and technical matters.

I would like to thank the Australian National University for provision of the Ph.D. scholarship and National ICT Australia for the supplementary scholarship.

Finally, I thank my parents for their love and care without which I could not have completed my research work. Thanks to my husband York for his love and precious support during the most difficult times.

Abstract

There has been a growing interest in spatial sound generation arising from the development of new communications and media technologies. Binaural spatial sound systems are capable of encoding and rendering sound sources accurately in three dimensional space using only two recording/playback channels. This is based on the concept of the *Head-Related Transfer Function* (*HRTF*), which is a set of acoustic filters from the sound source to a listener's eardrums and contains all the listening cues used by the hearing mechanism for decoding spatial information encoded in binaural signals. The HRTF is usually obtained from acoustic measurements on different persons. In the case of discrete data and sets of measurements corresponding to different human subjects, it is desirable to have a continuous functional representation of the HRTF for efficiently rendering moving sounds in the virtual spatial audio systems; further this representation should be well-suited for customization to an individual listener.

In this thesis, modal analysis is applied to examine the HRTF data structure, that is to employ the wave equation solutions to expand the HRTF with separable basis functions. This leads to a general representation of the HRTF into separated spatial and spectral components, where the spatial basis functions modes account for the HRTF spatial variations and the remaining HRTF spectral components provide a new means to examine the human body scattering behavior. The general model is further developed into the HRTF continuous functional representations. We use the normalized spatial modes to link near-field and far-field HRTFs directly, which provides a way to obtain the HRTFs at different ranges from measurements conducted at only a single range. The spatially invariant HRTF spectral components are represented continuously using an orthogonal series. Both spatial and spectral basis functions are well known functions, thus the developed analytical model can be used to easily examine the HRTF data feature—individualization.

An important finding of this thesis is that the HRTF decomposition with the spatial basis functions can be well approximated by a finite number, which is defined as the HRTF spatial dimensionality. The dimensionality determines the least number of the HRTF measurements in space. We perform high resolution HRTF measurements on a KEMAR mannequin in a semi-anechoic acoustic chamber. Both signal processing aspects to extract HRTFs from the raw measurements and a practical high resolution spatial sampling scheme have been given in this thesis.

List of Acronyms

A/D	Analog to Digital
D/A	Digital to Analog
DFT	Discrete Fourier Transform
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
HRIR	Head-Related Impulse Response
HRTF	Head-Related Transfer Function
IIR	Infinite Impulse Response
ILD	Interaural Level Difference
IR	Impulse Response
ITD	Interaural Time Difference
KEMAR	Knowles Electronics Mannequin for Acoustic Research
KLE	Karhunen-Loève Expansion
LTI	Linear Time Invariant
MLS	Maximum Length Sequence
MNLS	Minimum Norm Least-Squares
MSE	Mean Square Error
PCA	Principle Component Analysis
PIE	Periodic Impulse Excitation
PRBS	Pseudo Random Binary Signal
RMS	Root Mean Square
SFRS	Spatial Frequency Response Surfaces
SNR	Signal to Noise Ratio
SPL	Sound Pressure Level

Functions and Operators

$\lceil \cdot \rceil$	Integer ceiling function
$\lfloor \cdot \rfloor$	Integer floor function
$*$	Convolution
$(\cdot)^*$	Adjoint operator
$\overline{(\cdot)}$	Complex Conjugate
$(\cdot)^T$	Transpose operator
$(\cdot)^H$	Hermitian transpose operator
$e^{(\cdot)}$	Exponential function
$E_m(\cdot)$	Normalized Exponential function
$\delta(\cdot)$	Dirac delta function
$Y_n^m(\cdot)$	Spherical harmonics of degree n and order m
$P_\ell(\cdot)$	Legendre polynomial of the ℓ-th degree
$P_n^{\lvert m\rvert}$	Associated Legendre function of degree n and order $\lvert m\rvert$
$\mathscr{P}_n^{\lvert m\rvert}(\cdot)$	Normalized Legendre function of degree n and order $\lvert m\rvert$
$U_\ell(\cdot)$	Chebyshev polynomial of the ℓ-th degree
$T_n(\cdot)$	n-th order Chebyshev function of the first kind
$J_n(\cdot)$	n-th order Bessel function of the first kind
$j_n(\cdot)$	n-th order spherical Bessel function of the first kind
$h_n^{(1)}(\cdot)$	n-th order spherical Hankel function of the first kind
$\mathscr{F}$	Fourier transform
$\mathscr{F}^{-1}$	Inverse Fourier transform
$\big(\mathscr{L}f\big)(\widehat{\boldsymbol{x}})$	General homogeneous Fredholm integral operator on function f
$\mathscr{B}_N$	Mode limiting operator
$\mathscr{D}_\varGamma$	Spatial truncation operator

Mathematical Symbols

i	$\sqrt{-1}$
H	Head-related transfer function
$\widetilde{H}$	Reconstructed Head-related transfer function
H_l, H_r	Head-related transfer functions of left and right ears
h_l, h_r	Head-related impulse responses of left and right ears
x	Variable
$x(t)$	Sound signal
r, θ, ϕ	Sound source position in spherical coordinates, i.e., distance, elevation and azimuth
$\Delta\phi, \Delta\theta$	Azimuthal and Elevation sampling interval
w	Angular frequency (radians per second, or rad/s)
f	Frequency (Hz)
k	Wavenumber
n, m, ℓ, j, u	Enumeration variables for summations and expansions
N, M, L	Number of terms or order of expansion
a_n, b_n	HRTF filter model coefficients
w_j	HRTF statistical model weights
d_j	HRTF statistical model principle components
β_n^m	HRTF spherical harmonic expansion coefficients
$A_m(f)$	Horizontal plane HRTF spectral components
$C_{n;\ell}^m$	HRTF continuous model coefficients
a	Head radius (approximately 0.09 m)
c	Speed of sound in space (approximately 340 m/s)
$\varphi_\ell(\cdot)$	Complete set of orthonormal functions indexed by ℓ
ε	Mean square error
$Z_\ell^{(n)}$	ℓ-th positive roots of Bessel function $J_n(x) = 0$
$\mathbb{S}^2$	Unit sphere in 3D, also called the 2-sphere
$\mathit{\Gamma}$	Subregion of the unit sphere, $\mathit{\Gamma} \subset \mathbb{S}^2$
Ω	Domain of interest

$L^2(\cdot)$	Hilbert space with an inner product
$\mathbf{x}$	Position vector
$\widehat{\boldsymbol{x}}$	Unit vector to represent a direction in space
$s(\widehat{\boldsymbol{x}})$	Surface element on the unit sphere
$f(\widehat{\boldsymbol{x}})$	Function defined on the surface of a sphere
F_n^m	Function spherical harmonic coefficients of degree n and order m
E	Signal energy
ϵ	Arbitrary small value
$\mathbf{B}$	Matrix
$\mathbf{f}, \mathbf{g}$	Vectors of signals
$\check{\mathbf{f}}$	Least-squares solution
$\mathbf{f}^{+}$	Minimum norm least-squares solution
$\mathbf{f}^{\ddagger}$	Regularized minimum norm least-squares solution
η	Noise or distortion
μ	Tikhonov regularization parameter
$\rho(\widehat{\boldsymbol{x}}, k)$	Equivalent source field as a function of angular position and wavenumber
$a_m(\theta, k)$	Azimuth harmonics
κ	Relative power ratio

Contents

Declaration **i**

Acknowledgements **v**

Abstract **vii**

List of Acronyms **ix**

Functions and Operators **x**

Mathematical Symbols **xii**

List of Figures **xix**

List of Tables **xxvii**

1 Introduction **1**
- 1.1 Motivation and Background 1
- 1.2 Aim and Scope 4
- 1.3 Structure of the Thesis 4
- 1.4 Thesis Contributions 6

2 Acoustics of Spatial Sound **9**
- 2.1 The Basics of Human Hearing 9
  - 2.1.1 Primary Localization Cues: Interaural Differences 10
  - 2.1.2 Spectral Cues 12
  - 2.1.3 Distance Cues 13
- 2.2 Head-Related Transfer Function (HRTF) 14
  - 2.2.1 What is the HRTF? 14
  - 2.2.2 Synthesis of Spatial Audio using HRTF 17
  - 2.2.3 Measurement of HRTF 18
  - 2.2.4 HRTF Data Representation 20

2.2.5 HRTF Modelling 26
2.3 Summary 29

3 Continuous HRTF Model on Horizontal Plane using Data Independent Basis Functions 31
3.1 Introduction 31
3.2 Functional Modelling of Horizontal Plane HRTF 34
3.2.1 HRTF Spatial Azimuth Modelling 34
3.2.2 HRTF Spectral Modelling 35
3.2.3 HRIR Representation 41
3.3 Model Validation 42
3.3.1 Methodology 42
3.3.2 Continuous Model Performance 45
3.3.3 Model Performance Analysis 49
3.4 Comparison with Statistical Modelling Methods 51
3.5 Summary and Contributions 54
3.A Appendix: Complete Orthogonal Functions: Fourier Bessel Series . 57

4 Spherical Harmonic Analysis and Model-Limited Extrapolation on the Sphere 59
4.1 Introduction 59
4.2 Problem Formulation 61
4.2.1 Fredholm Integral Equation 62
4.2.2 Spherical Harmonics 62
4.2.3 Two Projections on the Sphere 63
4.3 Papoulis Algorithm on the Unit Sphere 65
4.3.1 Algorithm Description 65
4.3.2 Illustration of the Algorithm 66
4.3.3 Algorithm Convergence 71
4.4 Discrete Extrapolation Problem 74
4.4.1 Definitions and Notation 75
4.4.2 MNLS Extrapolation 76
4.4.3 Iterative Extrapolation Algorithm 78
4.5 Extrapolation and Noise 80
4.5.1 Noise Sensitivity of the Mode-limited Extrapolation 81
4.5.2 Stable Results for the Extrapolation of Noisy Signals 81
4.6 Summary and Contributions 83
4.A Appendix: Least-squares and Minimum Norm Least-squares 85

4.A.1 Least-squares Solution of Over-determined Equations 85
4.A.2 Minimum Norm Least-squares of Under-determined Equations 86
4.A.3 General Norm Minimization with Equality Constraints . . . 86
4.B Appendix: Tikhonov Regularization 87

5 Insights into Head-Related Transfer Function: Spatial Dimensionality and Continuous Representation 89
5.1 Introduction 90
5.2 Modal Analysis of HRTF 91
5.2.1 Theoretical Development 91
5.2.2 Dimensionality of HRTF as a Mode-limited Function 93
5.3 HRTF Continuous Representation 96
5.3.1 Normalized Modes for HRTF Spatial Representation 96
5.3.2 Fourier Spherical Bessel Series for HRTF Spectral Representation 98
5.3.3 Proposed Continuous HRTF Model 101
5.4 Implementation Analysis 101
5.4.1 Typical HRTF Measurement Setup 101
5.4.2 Practical Modal Decomposition Method 102
5.5 Simulation Results 106
5.5.1 HRTF Database 106
5.5.2 Results for Dimensionality and Analysis 107
5.5.3 Continuous Model Performance 111
5.6 Summary and Contributions 116

6 HRTF High Resolution Measurements on KEMAR Mannequin 119
6.1 Introduction 119
6.2 Experiment Procedure 120
6.2.1 HRTF Measurement Facility 120
6.2.2 HRTF Data Collection 122
6.3 Signal Processing Aspects of HRTF Measurement 123
6.3.1 Choice of the System Identification Method 123
6.3.2 Design of Test Signal 126
6.3.3 Post-signal Processing to Extract HRTFs 129
6.4 High Resolution HRTF Sampling over Sphere 134
6.4.1 Requirements for Design of HRTF Measurement 136
6.4.2 Meeting the Requirements: The IGLOO Grid 139
6.4.3 HRTF Sampling Arrangement based on IGLOO 143

6.4.4 Fast Spherical Harmonic Transform Algorithm 143
6.5 Simulation and Experimental Results 145
6.5.1 Analytical Solutions 145
6.5.2 Experimental Data 148
6.6 Summary and Contributions 149

7 Conclusions and Future Research Directions 151
7.1 Conclusions 151
7.2 Further Research Directions 152

Bibliography 155

List of Figures

1.1 Virtual spatial sound rendering in a tele-collaboration environment. 2
1.2 Major thesis contribution areas. 7

2.1 Low frequency interaural time difference for a spherical head of radius $a = 0.09$ m in the horizontal plane. Same ITD can be seen for front and back directions, which means we can not only use ITD to determine the direction of a sound source. 10
2.2 The “cone of confusion”. All points on the cone have the same distance from the cone’s apex and share the same ITD and ILD. They can only be distinguished using the spectral differences. Illustration taken from [1]. 11
2.3 Perception of elevation relies on the spectral coloration of a sound produced by torso, head and external ear, pinna. Illustration taken from [2]. 12
2.4 HRTF source positions defined by spherical coordinates. 14
2.5 The (a) HRIR and (b) HRTF magnitude for a human subject (left ear of subject 3 in the CIPIC database) in the horizontal plane. The abscissa in each plot is azimuth, the ordinate is time or frequency and the grayscale is amplitude or magnitude in dB. 15
2.6 The (a) HRIR and (b) HRTF magnitude in the median plane for a human subject (left ear of subject 3 in the CIPIC database). The abscissa in each plot is elevation, the ordinate is time or frequency and the grayscale is amplitude or magnitude in dB. 16
2.7 Headphone rendering of spatial sound systems using a pair of HRTFs. 17
2.8 Apparatus for HRTF Measurement. 18
2.9 Frequency domain representation of the HRTF as a function of azimuth in the horizontal plane (elevation = 90°). Diffraction effects can be seen on contralateral side at azimuths around 270° and 90° for left ear (a) and right ear (b). 21

2.10 Frequency domain representation of the HRTF as a function of elevation in then median plane (azimuth = 0°). Elevation effects can be seen, such as a notch around 7 kHz that migrates upward as elevation goes towards the north pole and a shallow peak around 13 kHz which flattens out at the upper sphere. (a) Left-ear measured HRTFs (b) Right-ear measured HRTFs. 22

2.11 Time domain representation of the HRTF as a function of azimuth in the horizontal plane (elevation = 90°). Diffraction effects can be seen on the contralateral side with lower amplitude initial peaks and longer delays at azimuths around 270° and 90° for left ear (a) and right ear (b). 23

2.12 Time domain representation of the HRTF as a function of elevation in the median plane (azimuth = 0°). Note there is a slight difference in the arrival times and amplitudes for the upper sphere and the lower sphere HRIRs. 24

2.13 Spatial representation of the HRTF as a function of azimuth and elevation at 2 kHz for left ear (a) and right ear (b). 24

2.14 Spatial representation of the HRTF as a function of azimuth and elevation at 8 kHz for left ear (a) and right ear (b). 25

3.1 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 that given 72 samples (source positions) equally spaced along the circle with $\Delta\phi = 5°$ in both data sets, the highest Fourier series order we can unambiguously derive is $M = 35$ with $f_{\max} = 21$ kHz. 38

3.2 Examples to demonstrate the structural similarities between the magnitude spectrum of $A_m(f)$ and the Bessel functions of the first kind. (a) the magnitude spectrum of $A_m(f)$ from MIT data at $m = 0, 10, 20$, and 30 and (b) the magnitude of the Bessel functions of the first kind $J_n(x)$ at the corresponding orders $n = 0, 10, 20$, and 30 with against arguments x from 0 to 30. 39

3.3 Magnitude of model coefficients $C_{m;\ell}$ solved from (a) MIT HRTFs of far-field $r = 1.4$ m and (b) analytically simulated HRTFs from the spherical head model of near-field $r = 0.7$ m. The model coefficients energy are kept in low order components. 44

3.4 Example of MIT measured and model reconstructed HRTFs using the frequency domain model (a) left ear measured HRTF at $\phi = 120°$ and (b) right ear measured HRTF at $\phi = 110°$. Measurements: star dot-dash line, Reconstruction: solid line. 45

3.5 The 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. 46

3.6 The 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 performances. 47

3.7 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). 47

3.8 Example of model interpolated HRTFs with original MIT measurements overlaid. (a) left ear HRTF at $\phi = 105°$ and (b) right ear HRTFs at $\phi = 165°$. Measurements: star dot-dash line, Interpolation: solid line. 49

3.9 Example of model reconstructed and interpolated HRIR from MIT data with original measurements overlaid. (a) reconstructed MIT left ear HRIR at $\phi = 180°$ and (b) interpolated MIT right ear HRIR at $\phi = 275°$. Measurements: star dot-dash line, Synthesis results: solid line. 50

3.10 (a) Example of PCA model reconstructed MIT left ear HRTF at $\phi = 120°$ and (b) example of KLE model reconstructed MIT left ear HRTF at $\phi = 120°$. Measurements: star dot-dash line, Reconstruction: solid line. 52

3.11 Example of KLE model using spline interpolation interpolated MIT (a) left ear HRTF at $\phi = 105°$ and (b) right ear HRTF at $\phi = 165°$. Measurements: star dot-dash line, Interpolation: solid line. 53

3.12 Rat distortion curves of (a) KLE model and (b) continuous model for MIT HRTF interpolation. 53

4.1 An example of data reconstruction on the sphere using Papoulis iterative algorithm. The observations in (b) are given for $0 \leq \theta \leq 100°$, $0 \leq \phi \leq 200°$ with $N = 9$. Extrapolation results in (c) are the estimates after 30 iterations with MSE of 0.3804. 67

4.2 Illustration of the algorithm as projections involving two subspaces, the mode-limited subspace and the space selection subspace, corresponding to the mode-limited signal and given limited measurements. 68

4.3 An example of the dominant two eigenvectors for the subregion $\mathit{\Gamma} \equiv (0 \leq \theta \leq 120°, 0 \leq \phi \leq 200°)$ and the mode limiting operator up to degree $N = 4$. 70

4.4 An example of λ_0 increasing with the subregion $\mathit{\Gamma}$, where the colatitude observation increases from 0° to 180° and the full range of azimuthal observation is assumed. The fraction area of the subregion is defined as the solid angles in fractions of the sphere, i.e., $\int_\Gamma d\Omega/4\pi$. 71

4.5 Mean square error between the estimated and the original signal for different data observation regions. FA stands for fractional area and the spherical harmonic coefficients are resolved to degree $N = 4$. . . 74

4.6 An example of data reconstruction on the sphere using the MNLS extrapolation. A mode-limited signal (a) is artificially generated by assigning the spherical harmonic coefficients up to $N = 13$. The limited observations in (b) are given for $0 \leq \theta \leq 90°$, $0 \leq \phi \leq 180°$. Extrapolation in (c) is the MNLS solution with MSE of 0.034. . . . 79

4.7 An example of noisy data extrapolation on the sphere using the regularization technique. A mode-limited signal (a) is artificially generated by assigning the spherical harmonic coefficients being one up to $N = 13$. The noisy observations in (b) are given for $0 \leq \theta \leq 90°$, $0 \leq \phi \leq 180°$. Extrapolation results in (c) and (d) are the regularized extrapolation solutions. 83

5.1 Geometry of HRTF measurement based on the reciprocity principle. 91

5.2 Dependence of the spherical Bessel function $j_n(ks)$ vs degree n at different ks shown on the vertically shifted curves. 94

5.3 Calculated the required truncation order N for HRTF representation as a function of frequency. 95

5.4 Modal decomposition of the HRTF with radial invariant frequency components. 96

5.5 Examples to demonstrate the structural similarities between the HRTF spectral components $\beta_n^m(k)$ and the spherical Bessel functions of the first kind. Top plots and middle plots are the real and imaginary parts of $\beta_n^m(k)$ with (a) $n = 0, m = 0$ and (b) $n = 12, m = 0$; and the bottom plots are the spherical Bessel functions $j_n(\cdot)$ at the corresponding degree $n = 0$ and $n = 12$ against arguments from 0 to 30. 99

5.6 Magnitude of the HRTF spectral components over spatial modes and wavenumber for the spherical head case. The spatial modes of degree n and order m correspond to number of $n^2 + n + m + 1$ on x-axis. 99

5.7 Synthetic HRTF reconstruction error performance for the audible frequency range of $[0.2, 20]$ kHz. 108

5.8 (a) MIT KEMAR mannequin measurements of frequency range $[0.2, 12]$ kHz. (b) HRTFs of CIPIC subject 3 of frequency range $[0.2, 8]$ kHz. 108

5.9 The relative power distribution of the FSB series components for HRTF spectral representation. (a) Analytically simulated HRTFs from the spherical head model; (b) MIT KEMAR mannequin HRTFs. For all spatial modes, the relative contribution of lower order FSB series is significant. 110

5.10 Examples of analytical simulated and measured HRTFs reconstructions using the proposed continuous model. (a) Analytical simulated HRTFs at elevation 90° and azimuth of 80° and (b) Left ear MIT KEMAR data at elevation 60° and azimuth of 0°. Original: dotted line $(\cdot)$, Reconstruction: solid line $(-)$. 110

5.11 MIT left ear HRTF polar response at 2 kHz: (a) original data and (b) synthesized response over the sphere. 112

5.12 MIT left ear HRTF polar response at 8 kHz: (a) original data and (b) synthesized response over the sphere. 112

5.13 MIT left ear HRTF polar response at 4.15 kHz, which is not measured but interpolated by applying the modal decomposed coefficients to the continuous spectral modelling basis functions. 113

5.14 Analytically simulated HRTFs at $r = 1.0$ m (top plots) and the extrapolated HRTFs at $r = 0.5$ and $r = 1.5$ m (b) compared to the reference (a). The horizontal axis is frequency and the vertical axis is azimuth from $[0°, 360°]$. 114

5.15 MIT HRTF magnitude spectrum reconstruction performance for direction of azimuth 120° and elevation 120° (a) left ear and (b) right ear. Original: dot-dash line, Reconstruction: solid line. 115

5.16 CIPIC HRTF magnitude spectrum reconstruction performance at azimuth 115° and elevation 45° (a) left ear and (b) right ear. Original: dot-dash line, Reconstruction: solid line. 115

6.1 HRTF measurement rotary hoop system 121

6.2 Operation diagram of HRTF data collection procedure. 123

6.3 Spectrogram of an exponential sweep from 300 Hz to 20 kHz with a duration of 4.2 ms. 127

6.4 Pre-emphasis filter: (a) Magnitude response (b) Phase response. . . 127

6.5 Test signal: (a) Original sweep in time domain (b) Spectrum of the original sweep (c) Pre-emphasized sweep in time domain (d) Spectrum of the pre-emphasized sweep. 128

6.6 Aligned measurements in 15 sections. 130

6.7 Low pass filter (LPF): (a) Magnitude response of LPF (b) Phase property of LPF (c) Signal before LPF (d) Signal after LPF. 132

6.8 Equalization of left ear HRTF data at direction of $\theta = 90°, \phi = 0°$ (a) System frequency response (b) Rough HRTF data before equalization (c) Final HRTF after equalization. 133

6.9 Main spectral characteristics of the HRTF. 134

6.10 Frequency-domain comparision of HRTFs measured at 13 elevations in the median plane (azimuth = 0°) (a) Left ear measured HRTFs (b) Right ear measured HRTFs. 135

6.11 Frequency-domain comparision of HRTFs measured at 36 azimuths in the horizontal plane (elevation = 90°) (a) Left ear measured HRTFs (b) Right ear measured HRTFs. 135

6.12 Picture of the polar cap region in the igloo scheme, showing three levels of subdivision to higher resolution samples. 140

6.13 Picture of the 3:6:3 equal area division, which divides the sphere into twelve base regions, three at either cap and six 60° × 60° equatorial regions. Here, each base region is sampled with 64 points. 142

6.14 An example of synthetic sampled HRTFs based on the IGLOO scheme and the reconstruction results at elevation of 79° and azimuth of 90°. Sampled data: dot-dash line, Reconstruction: solid line. 146

6.15 An example of HRTF interpolation at elevation of 62° and azimuth of 14° compared with reference of the synthetic solutions. Reference data: dot-dash line, Interpolation: solid line. 146
6.16 Synthetic HRTFs reconstruction error performance at different elevations. (a) $\theta = 30^\circ$ (b) $\theta = 60^\circ$ (c) $\theta = 90^\circ$ (d) $\theta = 120^\circ$. 147
6.17 An example of KEMAR left ear HRTF sampled based on the IGLOO scheme and the reconstruction results at elevation of 45° and azimuth of 110°. Sampled data: dot-dash line, Reconstruction: solid line. . . 148
6.18 An example of KEMAR left ear HRTF interpolation at elevation of 85° and azimuth of 0° compared with measurement reference. Sampled data: dot-dash line, Interpolation: solid line. 149
6.19 KEMAR HRTFs reconstruction error performance at different elevations. (a) Right ear $\theta = 60^\circ$ (b) Right ear $\theta = 90^\circ$ (c) Right ear $\theta = 120^\circ$ (d) Left ear $\theta = 60^\circ$ (e) Left ear $\theta = 90^\circ$ (f) Left ear $\theta = 120^\circ$.150

List of Tables

2.1 Terminology used in most HRTF literature. 15

3.1 Candidate closed form orthogonal functions $\varphi_\ell^{(m)}(f)$ defined on $f \in (0, f_{\max})$ for modelling the HRTF spectral components $A_m(f)$ with the argument normalized as $f' = f/f_{\max}$. 37

3.2 Percent mean square error of approximating $A_m(f)$ using four orthogonal functions with truncation number $L = 100$ based on MIT measurement data. 38

3.3 Horizontal plane HRTF reconstruction/interpolation using PCA, KLE, and continuous model. 54

5.1 MIT KEMAR data measurement steps (angles in degrees). 107

5.2 Summary of the number of spatial modes and the number of FSB series for the three sets of HRTF database. 111

6.1 Comparison of different methods for HRTF sampling over sphere. . 139

6.2 The IGLOO scheme based HRTF sampling for 20 kHz audible bandwidth (angles in degrees). 143