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

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
$\mathscr{D}_\Gamma$	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
$\varGamma$	Subregion of the unit sphere, $\varGamma \subset \mathbb{S}^2$
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

List of Figures

List of Tables