

**VISUALIZATION, MEASUREMENT, AND INTERPOLATION OF  
HEAD-RELATED TRANSFER FUNCTIONS (HRTF'S)  
WITH APPLICATIONS IN ELECTRO-ACOUSTIC MUSIC**

**by**

**Corey I. Cheng**

A dissertation submitted in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
(Electrical Engineering: Systems)  
in The University of Michigan  
2001

Doctoral Committee:

Associate Professor Gregory H. Wakefield, *Chair*  
Associate Professor Jeffrey A. Fessler  
Associate Professor Mary Simoni  
Professor William J. Williams

白日依山盡黃河入海  
流欲窮千里目更上一  
層樓

板橋道人  
是同鄉

英偉要見玩味

父字

丁丑



*An old pro told me that originality does not consist of saying what has never been said before; it consists of saying what you have to say that you know to be the truth.*

Harvey Penick  
Head Golf Professional Emeritus  
Austin Country Club

© Corey I. Cheng 2001  
All Rights Reserved

For my grandfather, who was at the University of Michigan in 1940

and

For my father, who was at the University of Michigan in 1975

## Acknowledgments

I am very grateful to the many people who have encouraged me to begin, continue, finish, and cherish my time as a graduate student here at the University of Michigan. I offer my deepest thanks:

To my advisor, Gregory H. Wakefield, for all of his support and encouragement in completing this thesis. Although sometimes at odds with me, Greg always gave me a wide latitude to pursue, stumble on, and learn from my own thoughts at my own pace, no matter how strange or tangential the ideas became. I greatly appreciate the five years of protected freedom he provided for me.

To the members of my doctoral committee, Jeff Fessler, Bill Williams, and Mary Simoni, for taking their valuable time to read, comment on, and support my work.

To the Office of Naval Research (ONR), the Naval Submarine Medical Research Laboratory (NSMRL), Groton, Connecticut, and the Graduate Assistance in Areas of National Need (GAANN) program for continued funding and support of this binaural hearing research.

To Dr. John Middlebrooks of the Kresge Hearing Research Institute at the University of Michigan for providing some of the HRTF data used in this thesis.

To Mike Blommer and Paul Runkle for providing important details about HRTF measurements that I could not have obtained alone.

To Abhi Kulkarni and Bill Rabinowitz at Bose Corporation of Framingham, Massachusetts for helping with initial design considerations for the HRTF measurement apparatus built for this thesis.

To Jim Johnston at IBM for encouraging me to present my material at the 107<sup>th</sup> Audio Engineering Society Convention, New York.

To Patricia MacDonald at the *Journal of the Audio Engineering Society*, for editing portions of this thesis with critical eyes and great care.

To Colby Leider and Doug Keislar at the *Computer Music Journal*, for editing portions of this thesis with critical eyes and great care. Colby, it is a pleasure to have known you as a fellow student, a professional colleague, and a friend.

To my academic uncle, Tom Buell, at the Naval Submarine Medical Research Laboratory (NSMRL), Groton, Connecticut, for the time, resources, hospitality, patience, enthusiasm, emotional support, and wisdom he extended to me in bringing me out to Connecticut to construct the HRTF measurement apparatus. Thanks to all of the researchers and staff at NSMRL for hosting me. Special thanks to Tom for the Lissajous figures, ginger snaps, and cognac, and Tom – try not to scare any more prospective secretaries with that thing hanging near your belt buckle, okay?

To Heather Kelley, my co-conspirator at NSMRL, for dealing with the daily realities of acoustic foam, fishing line, and Tom.

To Erica Omundsen, #1(?) daughter, for hosting me during my stay at NSMRL.

To Maureen Mellody, my Bra-lady / applied physicist / aerobics instructor labmate, for her careful renditions of the letter “P,” and for waking me up in the mornings at the obscene hour of 11:00 a.m. What?! Miss America was in the bathroom?!?!

To Eric Durant, for fruitful discussions on bowling balls.

To Andrew Sterian for being wǒ mēn shí yàn shì de dà gè (our lab's big brother).

To Bryan Pardo, Chris Rozell, Tim Hoffman, Ryan Smee, Erik Skow, Marco Statt, and other members of the extended MusEn/STAIRS family for their discussions on audio, engineering, music, juggling, mutually winding screws, and the evils of NT system installation.

To Vicki Yee, the Queen of Despair, the Übervicki, the butt-rock musicologist extraordinaire, for being the twin in the mirror, my ballroom dance partner, my confidant and car-seat conversationalist. Remember, Vic – use your powers for good, not for evil!

To Angie Snyder and David Watson for feeding me ALL THE TIME!

To the University of Michigan Ballroom Dance Team, for being an oasis of fun.

To my roommate Paul Pelzl for engaging discussions on technology and EECS210 late at night, and for introducing me to the trials and tribulations of TiVo. Bad TiVo! Pain in the \*\*\*! >WHACK<! Oh, and try not to kick my head off when I do the cha-cha in the living room, Paul?

To Kristin Martin, for sharing the first two years of my journey at Michigan, and for her good and generous heart. Honey-baked ham?

To all of my teachers, past and present, in and out of school, for caring for me and training me.

And most importantly, to my mother and father, for their undying and unconditional love, support, and encouragement. I love you, Mom and Dad!

CIC  
Ann Arbor, August 2001

## Table of Contents

Dedication	ii
Acknowledgments	iii
List of Tables	viii
List of Figures	ix
List of Sound Examples	xiii

\* \* \*

<b>Chapter 1.</b> Introduction and Background.	1
1.1 Introduction	1
1.2 Scope of thesis	5
1.3 Background	8
1.3.1 Duplex theory	8
1.3.2 Head-Related Transfer Functions (HRTF's)	12
1.3.3 Analytical and numerical models for HRTF's	16
1.3.4 Synthesis of spatial audio using HRTF's	18
1.3.5 Problems with HRTF-based synthesis of spatial audio	19
<b>Chapter 2.</b> 2-D and 3-D Visualization of HRTF's	21
2.1 Overview	21
2.2 Comparison of HRTF visualization methods	22
2.2.1 Frequency domain representation of HRTF data	26
2.2.2 Time domain representation of HRTF data	28
2.2.3 Spatial domain representation of HRTF data	34
2.2.4 Volume domain representation of HRTF data	51
2.3 Summary	58
<b>Chapter 3.</b> Error Analysis of HRTF's Measured with Complementary (Golay) Codes	61
3.1 Overview	61
3.2 System identification methods for HRTF measurement	62
3.2.1 Maximal-length shift register (MLS) method	65
3.2.2 Golay code method	67
3.3 Post-processing of HRTF signals: ITD, CTF's, and minimum phase DTF's	70
3.4 HRTF Measurement apparatus at the Naval Submarine Medical Research Laboratory (NSMRL), Groton, CT	72
3.5 Error analysis of system identification using Golay codes	78
3.5.1 Time-invariant nuisance parameters: additive noise	79
3.5.2 Time-invariant nuisance parameters: DC offsets	85

3.5.2.1	Upper bound for squared error	85
3.5.2.2	“DC-blocking” Golay codes	91
3.5.3	Time-variant nuisance parameters: variable group delay	98
3.5.4	Time-variant nuisance parameters: variable gain	107
3.6	Comparison to previous experiments and general discussion	112
3.6.1	Comparison of Golay code estimated HRTF’s	113
3.6.2	Comparison of MLS-estimated HRTF’s	118
3.7	Summary	125
<b>Chapter 4. A Beamforming Model for Cognitive HRTF’s</b>		127
4.1	Overview	127
4.2	Cognitive HRTF’s and adaptive beamformers	128
4.3	Beamformer model and adaptive algorithm description	131
4.3.1	Beamformer model description	131
4.3.2	Adaptive algorithm description	132
4.3.3	Determination of initial weight vector $\underline{w}_0$	134
4.3.4	Convergence properties of the algorithm as a function of the gain $\mu$	134
4.4	Experiments and results	135
4.4.1	Comparison of beamformer weight initialization methods	136
4.4.2	Modeling a cognitive frequency shift	137
4.4.3	Modeling the effect of spatial disorientation on auditory localization	138
4.4.4	Modeling the “cocktail party” effect	139
4.5	Summary	140
<b>Chapter 5. Moving Sound Source Synthesis for Binaural Electro-acoustic Music Using Interpolated Head-Related Transfer Functions (HRTF’s)</b>		150
5.1	Overview	150
5.2	An SFRS-based HRTF interpolation algorithm	152
5.2.1	Algorithm description	152
5.2.2	Perceptual evaluation	154
5.3	Composing in the spatial domain with moving sound sources	160
5.3.1	Implementation of a MATLAB-based spatialization GUI	161
5.3.2	Strengths and weaknesses of HRTF-based moving sound synthesis	164
5.3.2.1	Left-right motion	166
5.3.2.2	Front-back motion	169
5.3.2.3	Up-down motion	172
5.3.2.4	Dependence of spatial trajectory effectiveness on sound source content	173
5.3.3	Working with binaural sounds: compositional technique for moving sound sources	175
5.4	Summary	180
<b>Chapter 6. Conclusions and Future Directions</b>		182
<b>Bibliography</b>		186

## List of Tables

<u>Table</u>	<u>Description</u>	<u>Page</u>
2.1	Some one and two dimensional HRTF visualization strategies.	23
2.2	Summary of cross-subject similarity in SFRS's from 1-13 kHz.	41
2.3	Summary of design goals for a volumetric visualization and sonification tool for HRTF's.	54
5.1	Perceptual Evaluation of SFRS-based HRTF interpolation.	156

## List of Figures

<u>Figure</u>	<u>Description</u>	<u>Page</u>
1.1	Application of auditory display technology.	2
1.2	Directional hearing in the financial world?	3
1.3	Analogy of HRTF's to graphic equalizers.	4
1.4	Using Interaural Time Differences (ITD's) to estimate the azimuth of a sound source.	10
1.5	Using Interaural Intensity Differences (IID's) to estimate the azimuth of a sound source.	10
1.6	Ambiguity of ITD's in determining lateral position for higher frequencies.	11
1.7	The Cone of Confusion.	11
1.8	Spatial coordinate system and terminology used in much HRTF literature.	14
1.9	Block diagram of a simple HRTF-based spatial sound synthesis system.	18
2.1	Frequency domain comparison of theoretical HRTF's as a function of azimuth in the horizontal plane (elevation = 0°).	29
2.2	Frequency domain comparison of theoretical HRTF's as a function of elevation in the median plane (azimuth = 0°).	30
2.3	Frequency domain comparison of measured HRTF's as a function of azimuth in the horizontal plane (elevation = 0°).	31
2.4	Frequency domain comparison of measured HRTF's as a function of elevation in the median plane (azimuth = 0°).	32
2.5	Time domain comparison of measured and theoretical left ear HRTF's as a function of azimuth in the horizontal plane (elevation = 0°).	35
2.6	Time domain comparison of measured and theoretical right ear HRTF's as a function of azimuth in the horizontal plane (elevation = 0°).	36
2.7	Time domain comparison of measured and theoretical left ear HRTF's as a function of elevation in the median plane (azimuth = 0°).	37

2.8	Time domain comparison of measured and theoretical right ear HRTF's as a function of elevation in the median plane (azimuth = 0°).	38
2.9	Construction of SFRS's from existing HRTF data.	39
2.10	Interpretation of SFRS's.	40
2.11	Subject CC's SFRS's for several frequencies.	42- 43
2.12	Cross-subject comparison of SFRS's for three subjects at two different frequencies.	44
2.13	Diffraction effects in theoretical and measured HRTF's at 1.9 kHz as expressed in the spatial domain.	47
2.14	Diffraction effects in theoretical and measured HRTF's at 2.4 kHz as expressed in the spatial domain.	48
2.15	Elevation effects in theoretical and measured HRTF's at 6.8 kHz as expressed in the spatial domain.	49
2.16	Elevation effects in theoretical and measured HRTF's at 8.7 kHz as expressed in the spatial domain.	50
2.17	Volumetric visualization of HRTF's.	51
2.18	+8dB isosurface for $HRTF_L(\theta, \phi, f)$ empirically measured from subject CC.	52
2.19	-3dB isosurface for $HRTF_L(\theta, \phi, f)$ computed from a rigid, spherical model of the head.	53
2.20	A volumetric visualization and sonification tool for HRTF's.	54
2.21	Volumetric visualization of left and right HRTF's for a single subject.	55
2.22	Cross-comparison visualization of left-ear HRTF's taken from two different subjects.	56
2.23	Cross-comparison of measured and theoretical left-ear HRTF's.	57
3.1	Measurement of HRTF's.	63
3.2	Comparison of different system identification methods for measuring HRTF's.	64
3.3	Block diagram of Golay code system identification method with no nuisance parameters.	69
3.4	HRTF Measurement apparatus at the Naval Submarine Medical Research Laboratory (NSMRL), Groton, CT.	76
3.5	HRTF Measurement Apparatus at the Naval Submarine Medical Research Laboratory (NSMRL), Groton, CT.	77
3.6	Block diagram of Golay code system identification method with time-varying HRTF, additive noise, and DC offset contaminants.	79

3.7	Example of data recorded directly from ear canals.	80
3.8	Comparison between actual HRTF and simulations of HRTF's measured in the presence of additive Gaussian noise.	83
3.9	Dependence of the variance of the estimator in (3.19) on $A$ , $L$ , and $\sigma^2$ .	84
3.10	Dependence of squared error on $C$ , $A$ , $L$ , and $(T-L)^2$ .	90
3.11	Effects of mild time-variance (variable group delay) on HRTF measurements for varying $n_0$ .	104
3.12	Effects of mild time-variance (variable group delay) on HRTF measurements for varying $L$ .	105
3.13	Effects of mild time-variance (variable group delay) on HRTF measurements for varying $A$ . Note there is no change to $\Delta dB(\omega)$ .	106
3.14	Effects of mild time-variance (variable gain) on HRTF measurements for varying $G$ .	109
3.15	Effects of mild time-variance (variable gain) on HRTF measurements for varying $L$ .	110
3.16	Effects of mild time-variance (variable gain) on HRTF measurements for varying $A$ .	111
3.17	Measured left and right HRTF's for $(\theta, \phi)=(0^\circ, 0^\circ)$ .	114
3.18	Simulation of small head movement effects observed in HRTF's empirically measured by Zahorik with Golay codes.	117
3.19	Magnitude frequency domain comparison of left-ear, $(0^\circ, 0^\circ)$ HRTF's measured with Golay codes contaminated by actual and simulated head movement.	119
3.20	Time-domain comparison of left-ear, $(0^\circ, 0^\circ)$ HRTF's measured with Golay codes contaminated by actual and simulated head movement.	120
3.21	Magnitude frequency domain comparison of left-ear, $(0^\circ, 0^\circ)$ HRTF's measured with MLS codes contaminated by actual and simulated head movement.	123
3.22	Time-domain comparison of left-ear, $(0^\circ, 0^\circ)$ HRTF's measured with MLS codes contaminated by actual and simulated head movement.	124
4.1	A simple model for the physical and cognitive HRTFs' roles in spatial encoding and decoding.	129
4.2	Using a beamformer to approximate SFRS's.	130
4.3	Beamformer geometry and coordinate system used to simulate SFRS's and HRTF's.	136
4.4	Comparison of beamformer initialization techniques.	142

4.5	Comparison of beamformer initialization techniques.	143
4.6	Adaptation of beamformer from one carrier frequency to a different carrier frequency.	144
4.7	Adaptation of beamformer from one carrier frequency to a different carrier frequency.	145
4.8	Adaptation of beamformer due to a quick spatial shift of the entire beamformer.	146
4.9	Adaptation of beamformer due to a quick spatial shift of the entire beamformer.	147
4.10	Adaptation of beamformer due to changes in the spatial penalty function.	148
4.11	Adaptation of beamformer due to changes in the spatial penalty function.	149
5.1	Triangulation and linear interpolation used in computing SFRS's.	153
5.2	Using SFRS's to compute interpolated HRTF's	154
5.3	Design of listening task for perceptual evaluation of SFRS-interpolated HRTF's.	155
5.4	Perceptual evaluation of interpolated and non-interpolated HRTF's for subject CC.	158
5.5	Perceptual evaluation of interpolated and non-interpolated HRTF's for subject MM.	159
5.6	A simple headphone-based spatialization system which relies on ITD's and HRTF's only.	161
5.7	Graphical User Interface (GUI) constructed to facilitate the design and processing of spatial trajectories.	162
5.8	Informal comparison of spatialization and externalization properties for various areas in virtual space.	165
5.9	Intended orientations of the listener and spatial locations of a tennis ball in Sound Examples 1-5, sounds from a virtual tennis game.	167
5.10	Analogy of pivot locations to pivot chords.	170
5.11	Intended and implemented down-up/up-down spatial trajectories.	172
5.12	The "Slingshot" technique.	179

## List of Sound Examples

*Note:* There are several sound examples included on the accompanying CD at the back of this volume which demonstrate some of the binaural spatialization techniques discussed in this thesis. These binaural sound examples have been specifically prepared for listening over a good pair of headphones. Nonetheless, some of these examples are more successful than others, and the author realizes that not all listeners may be able to immediately hear the intended spatial effects. Because the effects can be delicate and may vary somewhat from person to person, we suggest that listeners hear each sound example in a quiet environment several times over headphones, closing their eyes to better concentrate on the sound. Sound examples 1-9 contain demonstrations which isolate certain spatialization techniques, while sound examples 10-16 contain spatialized excerpts from *Fishbowl*, a short piece of binaural electro-acoustic music composed to exploit these spatialization techniques. Sound example 17 contains the entire piece *Fishbowl*.

Further information regarding these sound examples can be obtained by contacting the author at [coreyc@umich.edu](mailto:coreyc@umich.edu).

All Sound Examples Copyright © 2000 Corey Cheng.

<u>CD</u>			
<u>Track #</u>	<u>Title</u>	<u>Dur.</u>	<u>Note</u>
1	Sound Example 1	0'29"	Virtual tennis game (mono). Source; no spatialization. See Figure 5.9a.
2	Sound Example 2	0'29"	Virtual tennis game (stereo). Spatialized version of Sound Example 1. Left-right motion accomplished with hard left-right panning. See Figure 5.9b for details.
3	Sound Example 3	0'29"	Virtual tennis game (stereo). Spatialized version of Sound Example 1. Left-right motion accomplished with HRTF's. See Figure 5.9c for details.
4	Sound Example 4	0'29"	Virtual tennis game (stereo). Spatialized version of Sound Example 1. Front-back motion accomplished with HRTF's. See Figure 5.9d for details.

5	Sound Example 5	0'29"	Virtual tennis game (stereo). Spatialized version of Sound Example 1. Combined left-right and front-back motion accomplished with HRTF's. Example of pivot location connecting two different spatial trajectories. See Figure 5.9e for details.
6	Sound Example 6	0'33"	Footsteps walking up and down stairs (mono). Source; no spatialization.
7	Sound Example 7	0'33"	Footsteps walking up and down stairs (stereo). Spatialized version of Sound Example 6. Up-down motion accomplished with HRTF's.
8	Sound Example 8	0'06"	Male and female speech (mono). Source; no spatialization.
9	Sound Example 9	0'06"	Male and female speech (stereo). Spatialized version of Sound Example 8. Front-back motion accomplished with HRTF's.
10	Sound Example 10	0'27"	<i>Fishbowl</i> excerpt (mono). Source with reverberation; no spatialization.
11	Sound Example 11	0'27"	<i>Fishbowl</i> excerpt (stereo). Spatialized version of Sound Example 10. Externalization and left-right motion accomplished with HRTF's.
12	Sound Example 12	0'09"	<i>Fishbowl</i> excerpt (stereo). Example of left-right motion accomplished with HRTF's. Combining spatialized, reverberated sounds and non-spatialized, non-reverberated sounds to create contrast in depth.
13	Sound Example 13	0'45"	<i>Fishbowl</i> excerpt (stereo). Example of down-up motion accomplished with HRTF's. Accentuating down-up motion by matching a sound's spectral properties to a listener's spatial expectations for low and high frequency sounds.
14	Sound Example 14	0'15"	<i>Fishbowl</i> excerpt (stereo). Example of the "Chase" technique. Combining hard left-right panning and HRTF-based left-right spatialization in the same sound example to create spatial tension and release.
15	Sound Example 15	0'17"	<i>Fishbowl</i> excerpt (stereo). Example of the "slingshot" technique. Combining hard left-right panning and HRTF-based left-right spatialization in the same sound example to create spatial tension and release.

16	Sound Example 16	0'22"	<i>Fishbowl</i> excerpt (stereo). Development of the “slingshot” technique. Combining hard left-right panning and HRTF-based left-right spatialization in the same sound example to create spatial tension and release.
17	Sound Example 17	3'36"	<i>Fishbowl</i> (1999). (stereo, entire piece).  My labmates and I call our laboratory the "Fishbowl," since one of its walls is made entirely of glass. Every day, friends and strangers point at us through the glass on their way to lunch, being sure to make funny faces at the poor grad students trapped inside. On days when things aren't going well, friendly knocks or taps at the glass are just reminders that we are like a strange school of fish that can be seen but not heard. This piece is what the everyday politics of our particular fishbowl could sound like to a passerby, when the fish are scared and the bowl of water is having a bad day!
	Total CD Time	10'05"	