A Tool for Volumetric Visualization and Sonification of Head-Related Transfer Functions (HRTF's)

Corey I. Cheng
coreyc@eecs.umich.edu
(734)-763-9074
http://www.eecs.umich.edu/~coreyc

Gregory H. Wakefield
ghw@eecs.umich.edu
(734)-763-1254
http://www.eecs.umich.edu/~ghw

University of Michigan
Department of Electrical Engineering and Computer Science
1101 Beal Ave.
Ann Arbor, MI 48109

ABSTRACT

This letter documents the design and implementation of a tool which facilitates a new method of visualizing and sonifying Head-Related Transfer Function (HRTF) data sets. HRTF's are empirically measured digital filters used in binaural spatialized audio systems which simulate the direction-dependant acoustic filtering properties of the human external ear. Because HRTF data sets are often reported as functions of three variables (azimuth angle, elevation angle, frequency), we suggest that three-dimensional, volumetric visualization tools can be effectively used to gain quick intuition about the structure of these large, complex data sets. Specifically, our visualization tool allows simultaneous visual and aural exploration of HRTF data by processing sounds with HRTF's selected directly from volumetric displays of HRTF data.

Keywords

Head Related Transfer Function, HRTF, spatial audio, visualization, sonification

1 INTRODUCTION

1.1 Head-related transfer functions (HRTF's)

The ability of humans to use sonic cues to estimate the spatial location of a target is of great practical and research importance. Recently, advances in computational power and acoustic measurement techniques have made it possible to empirically measure, analyze, and synthesize the spectral cues which influence spatial hearing [7]. These spectral cues are called Head-Related Transfer Functions (HRTF's), and summarize the direction-dependant acoustic filtering properties of the head, torso, and external ear (pinna). HRTF's are often used to simulate spatial audio in "virtual reality" audio systems by filtering monaural sounds with HRTF's corresponding to particular spatial locations.

Formally, a single Head-Related Transfer Function is defined to be a specific subject's left or right ear far-field frequency response, as measured from a specific point in the free field to a specific point in the ear canal. Typically, both left and right ear HRTF's are empirically measured from humans or mannequins for several spatial locations at a fixed radius from the head. Spatial location is designated by an ordered pair of angles (azimuth θ, elevation ϕ), where (-90°,0°) and (+90°,0°) correspond to spatial locations directly opposite the subject's left and right ears; (0°,-45°) and (0°,+45°) correspond to spatial locations slightly in front of/below and in front of/above the subject, etc. Thus, for notational convenience, we define a specific subject's entire set of measured left and right ear HRTF's to be samples of the three-dimensional functions $HRTF_L(\theta,\phi,f)$ and $HRTF_R(\theta,\phi,f)$. Consequently, we define $HRTF_{L,R}(\theta,\phi,f)$ to be the magnitude frequency response in dB of the left and right ears as a function of azimuth θ and elevation ϕ for a given frequency f. The three dimensional functions $HRTF_{L,R}(\theta,\phi,f)$ can be loosely interpreted as the relative amount of energy received by the left and right ears when presented with a pure tone of frequency f originating from spatial location (θ,ϕ). Usually, HRTF's are measured one location at a time for all frequencies of interest using FFT-based system identification techniques, so that $HRTF_{L,R}(\theta,\phi,f)$ are often irregularly sampled in spatial location but uniformly sampled in frequency.

Visualization Strategy	Subset of HRTF's Visualized	Ref
1. Study changes in HRTF's as azimuth/elevation varies. Graph all HRTF's sharing a given elevation/azimuth on a one dimensional plot of frequency vs. magnitude in dB.	$\{HRTF_{L,R}(\theta_i,\phi,f) \mid \phi=C, C\in[-90°,+90°], i=1\ldots N_{azimuths}\}$ $\{HRTF_{L,R}(\theta,\phi_i,f) \mid \theta=C, C\in[-180°,+180°], i=1\ldots N_{elevations}\}$	[4]
2. Study changes in HRTF's as azimuth/elevation varies. Graph all HRTF's sharing a given elevation/azimuth on a two-dimensional plot of azimuth/elevation vs. frequency vs. magnitude in dB, using interpolation to create a continuous two-dimensional surface.	$\{HRTF_{L,R}(\theta,\phi,f) \mid \phi=C, C\in[-90°,+90°]\}$ $\{HRTF_{L,R}(\theta,\phi,f) \mid \theta=C, C\in[-180°,+180°]\}$	[1]
3. Study the amount of energy received by the ear at a fixed frequency as a function of spatial location. Graph all HRTF data sharing the same frequency on a two dimensional plot of azimuth vs. elevation vs. magnitude in dB, using interpolation to create a continuous two-dimensional surface.	$\{HRTF_{L,R}(\theta,\phi,f) \mid f=C, C\in[0,f_{max}]\}$	[2], [3], [6]
4. Study similarity of HRTF's by plotting a "centroid average" for clusters of HRTF's. Plot clusters as spatial regions on a sphere while plotting the "centroid average" HRTF for each cluster below the sphere using ribbon plots. Stereographic visualization available.	varies	[5]

Table 1. Some one and two dimensional HRTF visualization strategies.

Visualization Goals	**Sonification Goals**
1. Volumetric display of HRTF data: isosurfaces. plot in 3D: $\{HRTF_{L,R}(\theta,\phi,f)=C, C\in\Re$ and user selectable$\})$ 2. Two-Dimensional display of HRTF data: probe planes. plot in 2D: $\{HRTF_{L,R}(\theta,\phi,f) \mid \alpha\theta+\beta\phi+\gamma f = C, \alpha,\beta,\gamma,C\in\Re$ and user selectable$\}$ 3. One-dimensional display of HRTF data: magnitude response. plot in 1D: $\{HRTF_{L,R}(\theta,\phi,f) \mid \theta=C, \phi=D, C\in[-180°,+180°], D\in[-90°,+90°]$; C, D user selectable$\}$	1. Allow user to select HRTF's from volumetric, two-dimensional, or one-dimensional displays. 2. Present left, right HRTF's and interaural time delay (ITD) all on the same screen to facilitate cross-comparisons. 3. Allow filtering of arbitrary monaural audio signals with the selected HRTF's. a. Provide quick but medium quality playback of short, processed sounds through built-in sound cards. b. Provide high-quality playback and real-time processing of sounds using a high-performance, outboard D/A and DSP filtering device such as the Tucker-Davis PD-1 Power SDAC.

Table 2. Summary of design goals for a volumetric visualization and sonification tool for HRTF's.

1.2 Previous attempts at HRTF visualization

Visualization of HRTF data is important in discovering how macroscopic features of HRTF's, such as peaks and notches in HRTF magnitude frequency responses, correspond to perceptually meaningful directional cues, such as azimuth and elevation. Visualization of HRTF data is also important in performing cross-subject comparison of HRTF data, as more structural insight can be gained by studying the individual differences among several HRTF data sets. However, HRTF data sets are hard to visualize and compare in their entirety for several reasons. HRTF data sets are very large, as HRTF's are often measured at several hundred irregularly spaced spatial locations for a single subject using a high sampling rate. Thus, several authors have tried to uncover structure in HRTF data by visually comparing subsets of HRTF's which share the same azimuth, elevation, or frequency. Consequently, although $HRTF_{L,R}(\theta,\phi,f)$ are functions of three variables, most visualization techniques have produced one or two dimensional graphs of HRTF data. Several recent attempts are listed in Table 1, and have proven useful in gaining insight into the structure of HRTF data. For example, [1] shows how the location of a spectral notch near 7 kHz varies as a function of elevation. [4] shows how diffraction effects due to the head can be seen as secondary echoes in time-domain versions of HRTF's. [2], [3], and [6] suggest how peaks in plots of spatial location vs. $HRTF_{L,R}(\theta,\phi,f)$ for fixed f could correspond to perceptually preferred directions in space. There are many other ways to display HRTF data, including cluster analysis [5] and other time-domain methods [4].

2 A TOOL FOR VOLUMETRIC VISUALIZATION OF HRTF'S

2.1 Motivation and Design Goals

We have developed a tool which uses three-dimensional, volumetric visualization tools, such as isosurfaces and probe planes, to display and sonify HRTF data. Because $HRTF_{L,R}(\theta,\phi,f)$ are functions of three variables, we suggest that three-dimensional, volumetric visualization tools can be used effectively to gain quick intuition about the structure of these large, complex data sets. Although previous one and two dimensional attempts at HRTF visualization have provided some structural intuition about HRTF data, common three-dimensional visualization tools associated with medical and geophysical imaging applications could offer new methods for uncovering structure in HRTF data. In addition, volumetric visualization is a unifying visualization strategy, as many of the one and two dimensional visualization strategies in Table 1 can be exactly recreated from volumetric plots. Furthermore, many current HRTF visualization methods do not provide the aural feedback needed to associate systematically varying structures found in visual representations of selected HRTF's with spatialized sounds produced with those selected HRTF's. Therefore, our visualization tool was designed to allow the investigator to listen to sounds processed with HRTF's selected directly from volumetric displays of HRTF data. In order to facilitate audio-visual cross-comparison of HRTF data, the tool was also designed so that all HRTF information for a single subject (left and right HRTF's and the interaural time delay (ITD) – another important spatial cue) could be presented on a single screen. The specific design goals and capabilities of our volumetric visualizion tool are summarized in Table 2; the user interface to the tool can be found in Figure 2; and several screenshots from the tool's display can be found in Figures 3, 4, and 5.

2.2 Applications

By definition, an isosurface of a three-dimensional function $f(x,y,z)$ is the locus of all points for which $f(x,y,z)=C$, a user-defined constant. Examining isosurfaces of $HRTF_{L,R}(\theta,\phi,f)$ could be useful for several reasons. Isosurfaces for which $HRTF_{L,R}(\theta,\phi,f)\approx\max(HRTF_{L,R}(\theta,\phi,f))$ and $HRTF_{L,R}(\theta,\phi,f)\approx\min(HRTF_{L,R}(\theta,\phi,f))$ could provide insight into how spectral peaks and notches vary as a function of θ,ϕ, and f. Simple, well-separated structures in these isosurfaces could suggest how HRTF's can be efficiently parameterized in θ,ϕ, and f simultaneously. For example, the volumetric display in Figure 3 shows the isosurface $\{HRTF_L(\theta,\phi,f)=8\text{dB}$) ($8\text{dB}\approx\max(HRTF_L(\theta,\phi,f)))\}$, which contains several well-separated structures for $f\in[6$ kHz, 10 kHz$]$ and $\theta\in[-180°,+180°]$. The notch at 7 kHz in several HRTF's thought to be associated with elevation perception is consistent with this separation, as there is a clear "dividing line" in the isosurface from 5-9 kHz.

Probe planes are also useful tools in the visualization of volumetric data, as they allow users to look at arbitrary "slices" of volumetric data. We define a probe plane to be the set $\{HRTF_{L,R}(\theta,\phi,f)\} \mid \alpha\theta+\beta\phi+\gamma f = C$, where $\alpha,\beta,\gamma,C\in\Re$ are user selectable}. Visualization methods 2 and 3 in Table 1 are specific cases of "axis-aligned" probe planes for which $HRTF_{L,R}(\theta,\phi,f)\}$ has been plotted in two dimensions for constant θ,ϕ, or f. Probe planes may be useful in exploring how HRTF's change with small, linear perturbations of θ,ϕ, or f. By studying the nature of these perturbations for several subjects, probe planes could highlight how low-order parameterized models for HRTF's can be mathematically customized to a specific individual. Individual differences in probe planes and volumetric plots and may be easier to recognize, and may provide a larger context within which to understand individual differences in HRTF magnitude responses. For example, Figure 4 contains a cross-comparison between two subjects' left-ear HRTF's. The volumetric displays for subject Cheng clearly contain a structure in the 8 dB isosurface at 15 kHz which subject Mellody's 8 dB isosurface lacks, and this effect can be seen in the two-dimensional and one-dimensional profiles of the HRTF data at this frequency.

Volumetric visualization techniques may also be useful in comparing measured HRTF's with theoretical HRTF's computed from physical models. For example, Figure 5 shows a comparison between measured left-ear HRTF's and theoretical left-ear HRTF's computed from a rigid spherical model of the head. The theoretical HRTF's were computed from pseudo-code in [4] using the following parameters: head radius = .1m, source distance = 1.5m, speed of sound = 330 m/s, right ear location = (100°,10°), left-ear location = (100°,-10°). In this example, there are vast differences between the measured and theoretical 5 dB isosurfaces, as the measured isosurface is much more complex than the theoretical isosurface. However, the theoretical and measured magnitude response profiles match fairly well for the selected elevation, and suggest that simple models of isosurfaces may be good approximations to measured HRTF's in some cases. Therefore, low-order parameterizations for HRTF's might be found from a combination of simple structures strategically placed in the volume domain.

2.3 Implementation Details

Figure 2 shows a block diagram of the entire visualization tool. We chose to use two software packages to build our visualization tool: 1) Advanced Visual Systems' AVS/Express visualization environment and 2) MATLAB version 5.3 for the PC, an interpreted, immediate mode command line engineering data processing environment. AVS/Express provided a rich set of graphics routines with which to build a complex, interactive visualization front-end to a MATLAB-based computation and sound processing engine. Inter-process communication was accomplished via AVS/Express' and

MATLAB's C++ API's to ActiveX, a Microsoft protocol for inter-process communication. MATLAB v5.3 provided ready-made utilities for volumetric visualization which allowed fast prototyping and mathematical manipulation of volumetric HRTF data. We also chose to use MATLAB because its sound support subsystems allowed for different audio rendering methods. We implemented the two methods below, giving the user the choice to use either method during visualization:

1) Use the built-in filtering capabilities of MATLAB v5.3 to filter small soundfiles off-line with selected HRTF's, then play back the processed sounds through generic soundcards.

2) Use a dedicated, high-quality D/A and DSP hardware device, such as the Tucker-Davis Technologies PD-1 Power SDAC, to filter arbitrary monaural audio streams and/or arbitrary soundfiles with selected HRTF's in real time.

3. CONCLUSIONS AND FUTURE DIRECTIONS

This letter introduced a tool for the visualization of HRTF data in three dimensions (azimuth angle, elevation angle, and frequency). Volumetric visualization provides a new method of exploring HRTF data for structure, and facilitates cross-subject comparison of these large, complex data sets. We plan to use this tool for subsequent investigations involving 1) the parameterization of HRTF's in three dimensions, 2) comparison of theoretical HRTF's and measured HRTF's.

4. ACKNOWLEDGMENTS

We thank Dr. John C. Middlebrooks at the Kresge Hearing Research Institute of the University of Michigan and Dr. Thomas Buell at the Naval Submarine Medical Research Laboratory (NSMRL) for providing the HRTF data used in this research. We would also like to thank the ICAD2000 reviewers, whose comments were extremely helpful in completing this paper. This research was supported by a grant from the Office of Naval Research in collaboration with NSMRL, Groton, CT.

5. REFERENCES

[1] Carlile, S. and Pralong, D. "The location-dependent nature of perceptually salient features of the human head-related transfer functions." *Journal of the Acoustical Society of America*, **95**(6): June 1994.

[2] Cheng, Corey I. and Wakefield, Gregory H. "Spatial Frequency Response Surfaces: An Alternative Visualization Tool for Head-Related Transfer Functions (HRTF's)." *Proceedings of the 1999 International Conference on Acoustics, Speech, and Signal Processing (ICASSP99)*, Phoenix, Arizona: 1999.

[3] Cheng, Corey I. and Wakefield, Gregory H. "Introduction to Head-Related Transfer Functions (HRTF's): Representations of HRTF's in Time, Frequency, and Space (invited tutorial)." *Proceedings of the Audio Engineering Society (AES) 107th Convention (Preprint 5026)*, New York: 1999.

[4] Duda, Richard O. and Martens, William M. "Range dependence of the response of a spherical head model." *Journal of the Acoustical Society of America*, 104(5): November 1998.

[5] Martens, William. "Cluster Analysis of the Head-Related Transfer Function." http://wwwsv1.u-aizu.ac.jp/~wlm/cluster.html.

[6] Middlebrooks, John C. "Narrow-band sound localization related to external ear acoustics." *Journal of the Acoustical Society of America*, **92**(5): November 1992.

[7] Wightman, Frederic L. and Kistler, Doris J. "Headphone simulation of free-field listening. I: Stimulus synthesis." *Journal of the Acoustical Society of America*, **85**(2): February 1989.

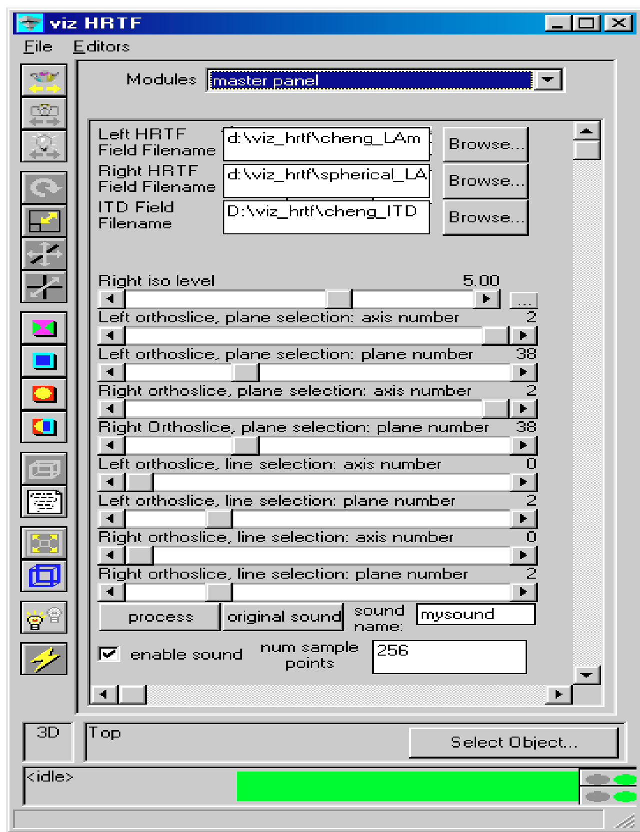

Figure 1. User Interface to Visualization / Sonification Tool. This interface is used to control volumetric visualization parameters, such as isosurface and probe plane selection. The user can control other graphic rendering parameters, such as scaling, camera, lighting, smoothing, occlusion, and opacity options.

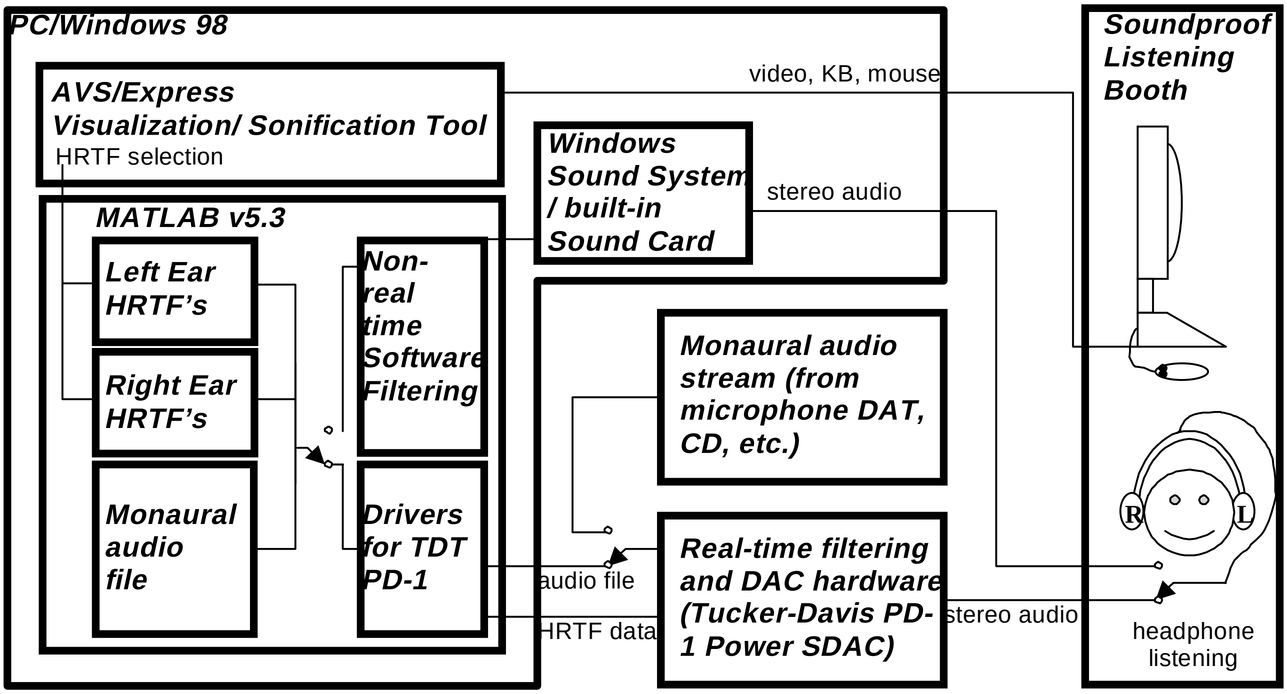

Figure 2. Block diagram of a volumetric visualization and sonification tool for HRTF's.

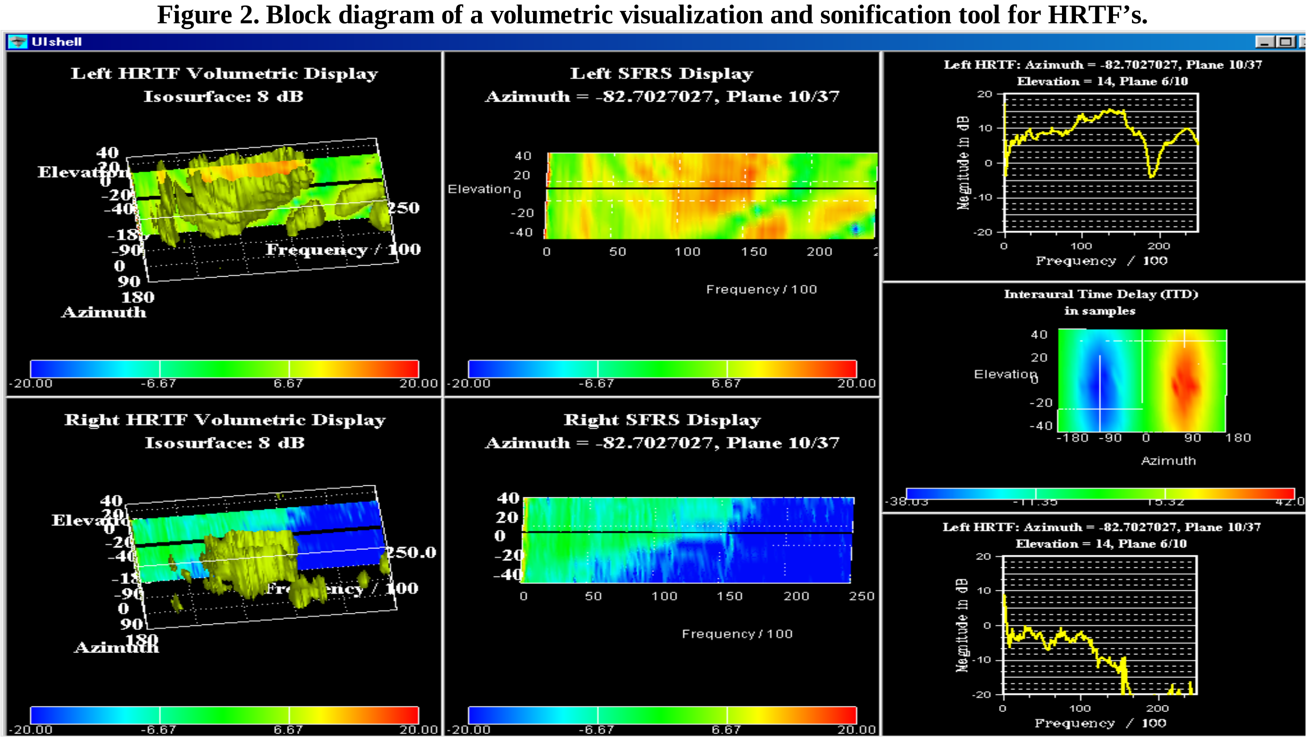

Figure 3. Visualization of Left and Right HRTF's taken from a single subject. The leftmost two panels are volumetric displays of HRTF data and contain both an isosurface and a probe plane. The center two panels are isolated displays of the same probe planes seen in the leftmost two graphs. The rightmost three panels are actual HRTF's, along with a two-dimensional graph of another important spatial cue, the interaural time delay (ITD). The black lines in the leftmost four graphs represent the trajectory through $HRTF_{L,R}(\theta,\phi,f)$ from which the HRTF's in the rightmost column have been sampled. In this figure, note how there are several well-separated structures in the left HRTF volumetric displays starting at 6-10 kHz. These structures could be important to the perception of elevation, and are consistent with the "spectral notch" at 7 kHz thought to be important to the perception of elevation.

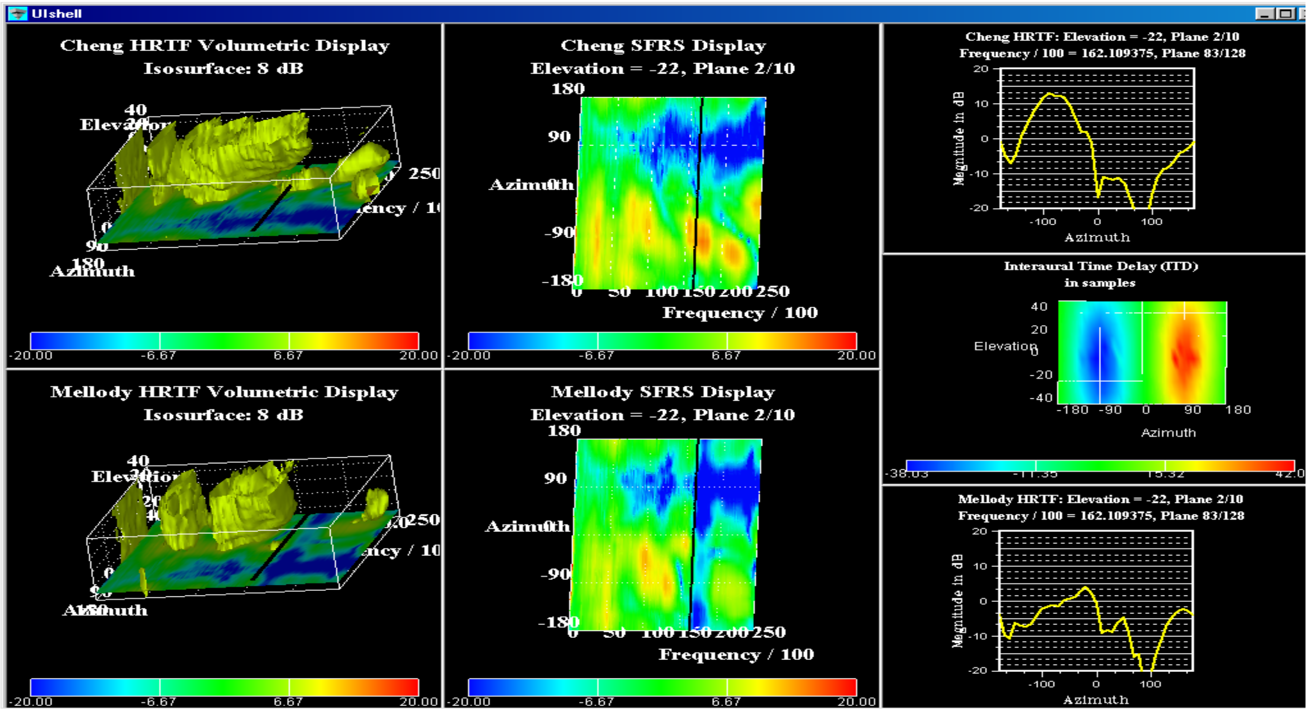

Figure 4. Cross-comparison visualization of left-ear HRTF's taken from two different subjects. In this example, one can see clear differences in the volumetric displays of the data, as there are clear structural differences between the 8 dB isosurfaces. There is a structure at 15 kHz in the upper left panel which is not present in the lower left panel. This difference can also be seen in the middle two panels, and is probed at 15 kHz in the upper and lower right panels.

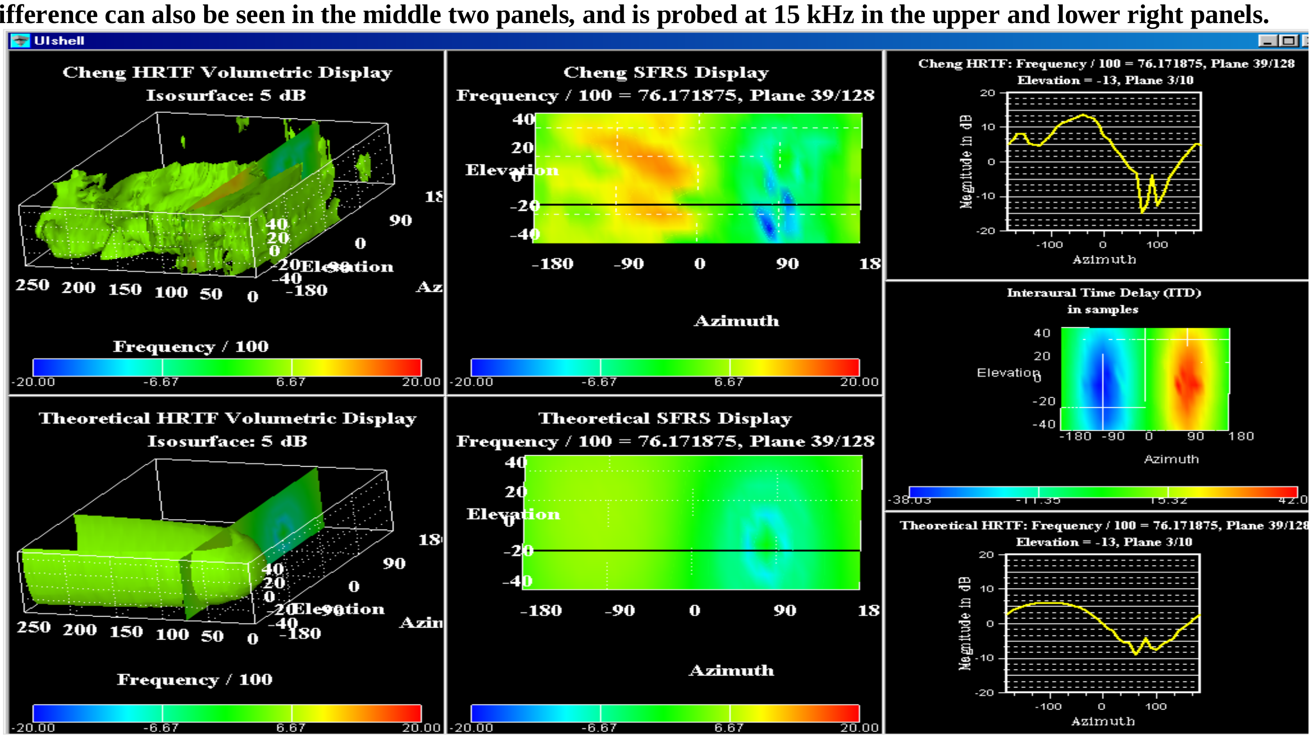

Figure 5. Cross-comparison of measured and theoretical left-ear HRTF's. The upper three panels contain measured HRTF's, while the lower three panels contain theoretical HRTF's computed from a rigid sphere model of the head. In this example, there are vast differences in the two 5 dB isosurfaces, as the measured isosurface is much more complex than the theoretical isosurface. However, the magnitude responses in the upper and lower right panels match fairly well, and suggest that simple models of isosurfaces may be good approximations to measured HRTF's in some cases.