A Study of the Relationship between Head Related Transfer Functions and Elevations

M. Iyer, M. Joshi, N. Gupta and L. Hmurcik,

International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV'10)

Hamid R. Arabnia and Leonidas Deligiannidis, Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV'10), ISBN #: 1-60132-154-6, http://www.worldacademyofscience.org/worldcomp10/ws

M. Iyer, M. Joshi, N. Gupta and L. Hmurcik, "A Study of the Relationship between Head Related Transfer Functions and Elevations," Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV'10), Las Vegas, NV, 2010.

A Study of the Relationship between Head Related Transfer Functions and Elevations

M. Iyer, M. Joshi*, N. Gupta and L. Hmurcik
University of Bridgeport, 221 University Avenue, Bridgeport, CT 06604
{madhumii, mjoshi, navarung, hmurcik}@bridgeport.edu
*Contact Author

Abstract

Head Related Transfer Functions **(**HRTFs) are signal processing models that represent the transformations undergone by acoustic signals, as they travel from their source to the listener's eardrums. The study of HRTFs is a rapidly growing area with potential uses in virtual environments, auditory displays, the entertainment industry, human-computer interface for the visually impaired, aircraft warning systems, etc. The positioning of the sound source plays a major role in the resonant frequency of the HRTFs. In this paper, we examine the effect of changing the elevations of these sources; we examine the effect on the first peak and the first notch of HRTFs. We use the HRTF database at FIU DSP lab. This database hosts the HRTFs from 15 subjects and their 3-D images of conchas. For each subject, the database contains the Head Related Impulse Responses (HRIRs) for the sound sources placed at six elevations (54°, 36°, 18°, 0°, -18° and -36°) and twelve azimuths (180°, 150°, 120°, 90°, 60°, 30°, 0°, -30°, -60°, -90°, -120° and -150°). A relationship between the first peak or notch and the elevation can help us model HRTFs mathematically. This can reduce the size of a HRTF database and can increase the speed of HRTF related computations.

Key Words: HRTF, binaural, elevation, HRTF peaks, HRTF notches

I. INTRODUCTION

A. *HRTF (Head Related Transfer Functions)*

Head-Related Impulse Responses (HRIRs) are used in signal processing to model the synthesis of spatialized audio which is used in a wide variety of applications, from computer games to aids for the visually impaired [1]. HRTFs represent the transformation undergone by the sound signals, as they travel from their source to the listener's eardrums. This transformation is due to the interaction of sound waves with the torso, shoulder, head and outer ear of a listener [2]. Therefore, the two components of these HRTF pairs (left and right) are typically different from each other, and pairs corresponding to sound sources at different locations around the listener are different. Furthermore, since the physical elements that determine the transformation of the sounds reaching the listener's eardrums (i.e., the listener's head, torso and pinnae) are different for different listeners, the HRTF sets should be different also [3].

B. *HRTF Measurements and FIU Database*

The Ausim3D's HeadZap HRTF Measurement System was used for measurement in FIU lab [4]. This system measures a 256-point impulse response for both the left and the right ear using a sampling frequency of 96 KHz. Golay codes are used to generate a broad-spectrum stimulus signal delivered through a Bose Acoustimass speaker. The response is measured using miniature blocked meatus microphones placed at the entrance to the ear canal on each side of the head. Under control of the system, the excitation sound is issued and both responses (left and right ear) are captured (Figure 1).

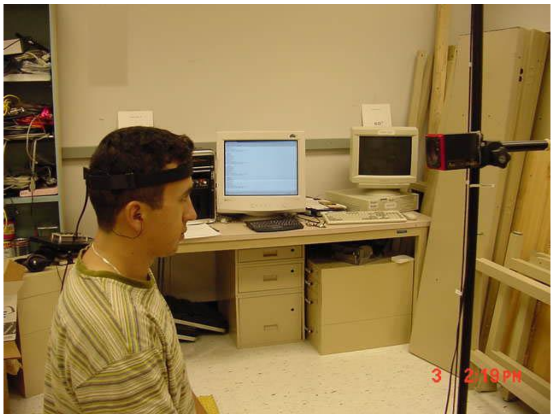


Figure 1. Measurement of HRTFs using Ausim3D's HeadZap System

Since the Golay code sequences that are played are meant to represent a broad-band excitation equivalent to an impulse, the sequences captured in each ear are the impulse responses corresponding to the HRTFs. The system provides these measured HRIRs as a pair of 256-point minimum-phase vectors, and an additional delay value that represents the Interaural Time Difference (ITD), i.e., the additional delay observed before the onset of the response collected from the ear that is farthest from the speaker position. In addition to the longer onset delay of the response from the "far" or "contralateral ear" (with respect to the sound source), this response will typically be smaller in amplitude than the response collected in the "near" or "ipsilateral ear". The difference in amplitude between HRIRs in a pair is referred to as the Interaural Intensity Difference (IID).

Our protocol records HRIR pairs from source locations at the 72 possible combinations of $\phi = \{-36^o, -18^o, 0^o, 18^o, 36^o, 54^o\}$ and $\theta = \{0^o, 30^o, 60^o, 90^o, 120^o, 150^o, 180^o, -150^o, -120^o, -90^o, -60^o, -30^o\}$. The left (L) and right (R) HRIRs collected for a source location at azimuth θ and elevation ϕ are symbolized by $h_{L,\theta,\phi}$ and $h_{R,\theta,\phi}$, respectively. The corresponding HRTFs are $H_{L,\theta,\phi}$ and $H_{R,\theta,\phi}$. The creation of a spatialized binaural sound (left and right channels) involves convolving the single-channel digital sound to be spatialized, s(n), with the HRIR pair corresponding to the azimuth and elevation of the intended virtual source location [5]:

$$y_{L,\theta,\phi}(n) = \sum_{k=-\infty}^{\infty} h_{L,\theta,\phi}(k) \cdot s(n-k), \quad \text{and} \quad y_{R,\theta,\phi}(n) = \sum_{k=-\infty}^{\infty} h_{R,\theta,\phi}(k) \cdot s(n-k) \qquad (1)$$

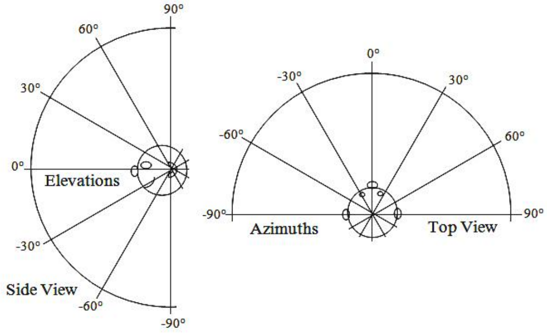


Figure 2. Diagram of spherical coordinate system

FIU HRTF database consists of head-related transfer functions based on ear/head measurements measured at the FIU DSP Lab. It contains HRTF data from 15 subjects at 12 different azimuths and 6 different elevations. This database also includes 3-D images of the subject's pinnae (outer ears). Anthropometric measurements of the various parts of the pinna are included [6].

C. Importance of the First Peak

The first peak of an HRTF corresponds to constructive interference of the incident and reflected sound waves in the pinna. Similarly, the first notch corresponds to the destructive interference between the incident and reflected sound waves. The relationship between the first peak/notch and the elevation can help us model HRTFs mathematically. This can reduce the size of a HRTF database and can increase the speed of HRTF related computations.

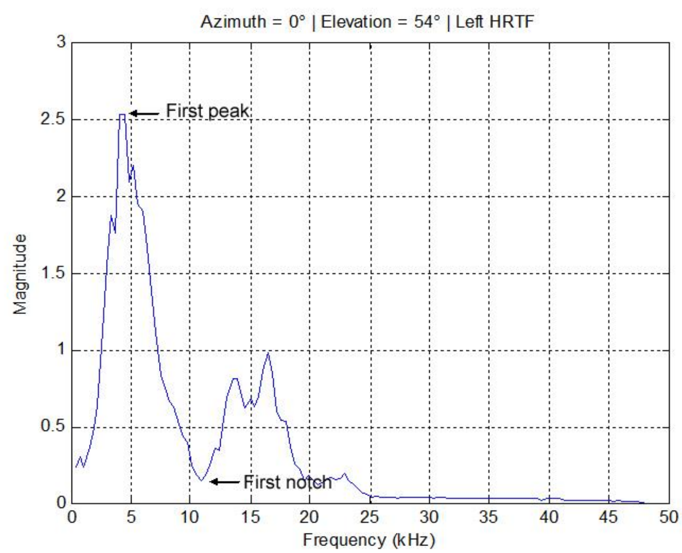


Figure 3. HRTF measured at Azimuth = 0°and Elevation = 54°

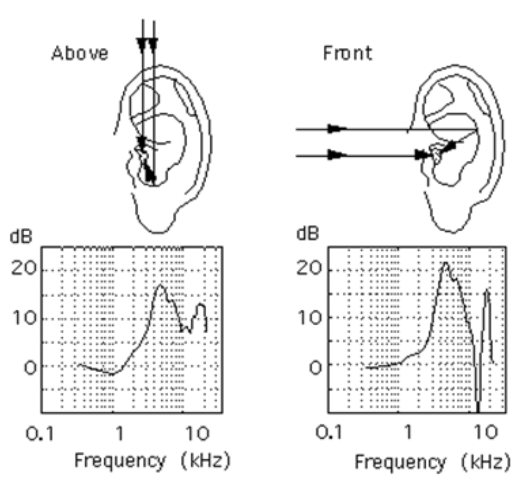


Figure 4. Constructive and destructive interference

II. PROCEDURE

MATLAB (Version 7.4 (R2007a)) was used to plot figures from the FIU database. Data from eight subjects (both left and right ears) were analyzed. HRTFs show distinct peaks and notches and they move as the elevation changes (see Figure 5).

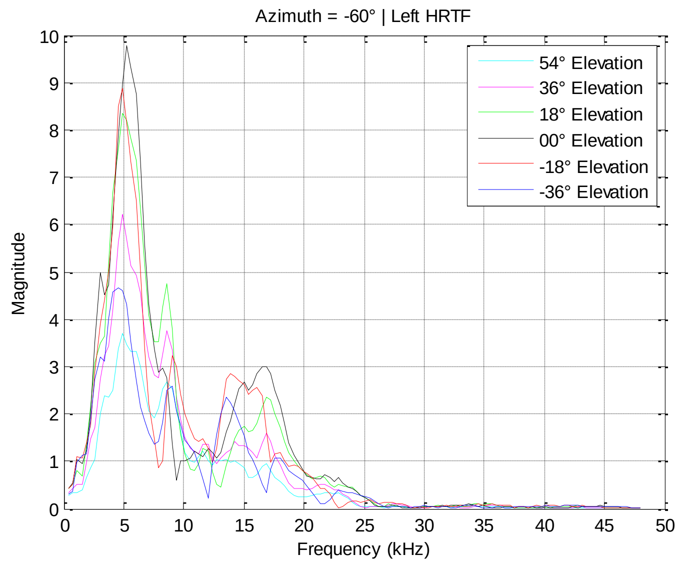
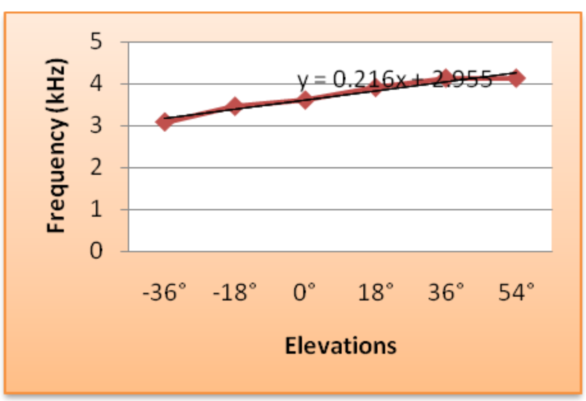


Figure 5. HRTFs for left ear measured at Azimuth -60°and six elevations

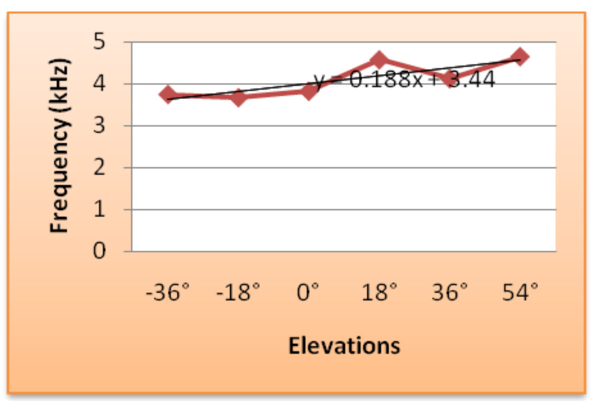
The first peak/notch of the HRTFs was measured manually and tabulated. They were plotted using MS-Excel. To study the trend of the movement of the peak/notch with frequency, best fit lines were plotted and their slopes (kHz/degree) were calculated.

III. RESULTS

There is a pronounced correlation between elevation and the frequency of the first peak/notch in HRTFs.

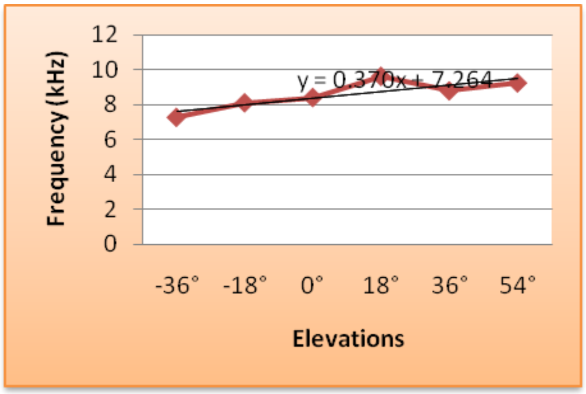


6a. Left Ear | Azimuth 90° (Slope=0.216 kHz/degree)

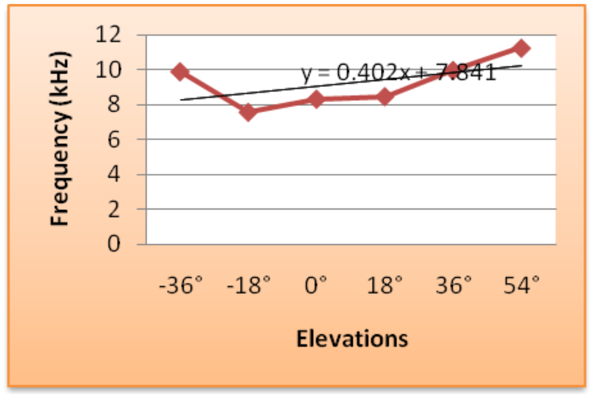


6b. Right Ear | Azimuth -90° (Slope=0.188 kHz/degree)

Figure 6a and 6b. HRTFs first peak frequency vs. elevations



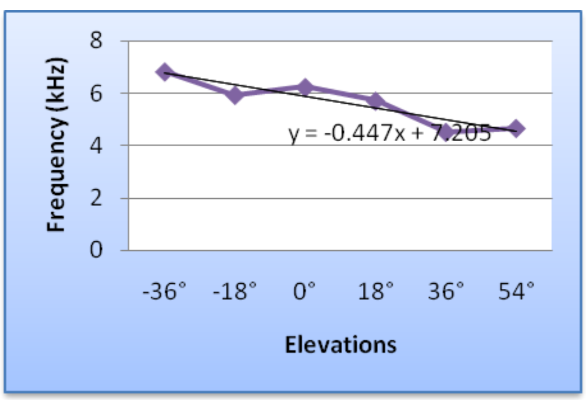
7a. Left Ear | Azimuth 150° (Slope=0.370 kHz/degree)



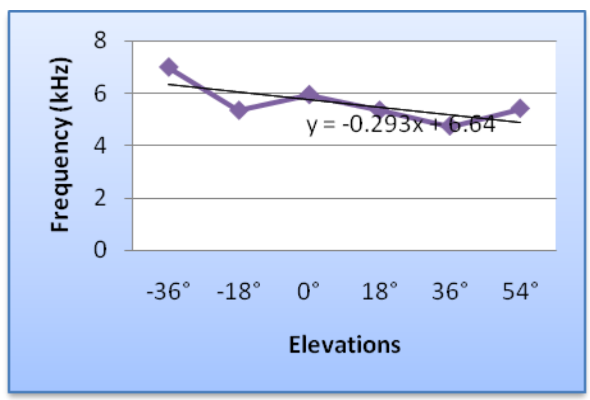
7b. Right Ear | Azimuth -150° (Slope=0.402 kHz/degree)

Figure 7a and 7b. HRTFs first notch frequency vs. elevations

As seen in Figures 6a and 6b, the first peak moves towards higher frequencies as we increase the elevation. The rate of increase is 0.216 kHz/degree and 0.188 kHz/degree for azimuth 90° (left ear HRTFs) and -90° (right ear HRTFs) respectively. In Figures 7a and 7b, we see a similar trend of the first notch moving toward higher frequencies with an increase in elevation. The rate of increase is 0.370 kHz/degree and 0.402 kHz/degree for azimuth 150° (left ear HRTFs) and -150° (right ear HRTFs) respectively. The reason for this move is not clear, but is likely related to the shape of the concha and how the incident and the reflected sound waves interact [7]. This trend of notches moving towards higher frequency with increase in elevation is more consistent and uniform than the movement of the peak.



8a. Left Ear | Azimuth -90° (Slope=-0.447 kHz/degree)



8b. Right Ear | Azimuth 90° (Slope=-0.293 kHz/degree)

Figure 8a and 8b. HRTFs first peak frequency vs. elevations

Figures 8a and 8b show anomalous results for the peak movement with respect to elevation. Here we see a decrease in frequency (-0.447 kHz/degree and -0.293 kHz/degree for azimuth -90° (left ear HRTFs) and 90° (right ear HRTFs) respectively) with an increase in elevation. As previously stated the movement of notches with frequency is more consistent than the peaks. The Figures 9 and 10 below show an alternative view of the movement of peak and notch with elevation. The relationship between peaks/notch movement and elevations is not linear. However, this relationship can be fairly accurately described using a best fit line.

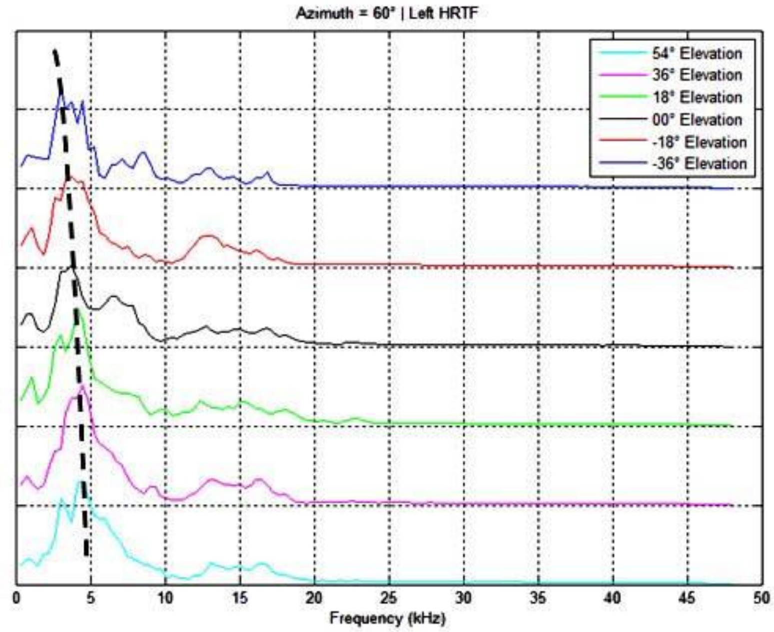


Figure 9. HRTF first peaks moved towards higher frequency with increase in elevation

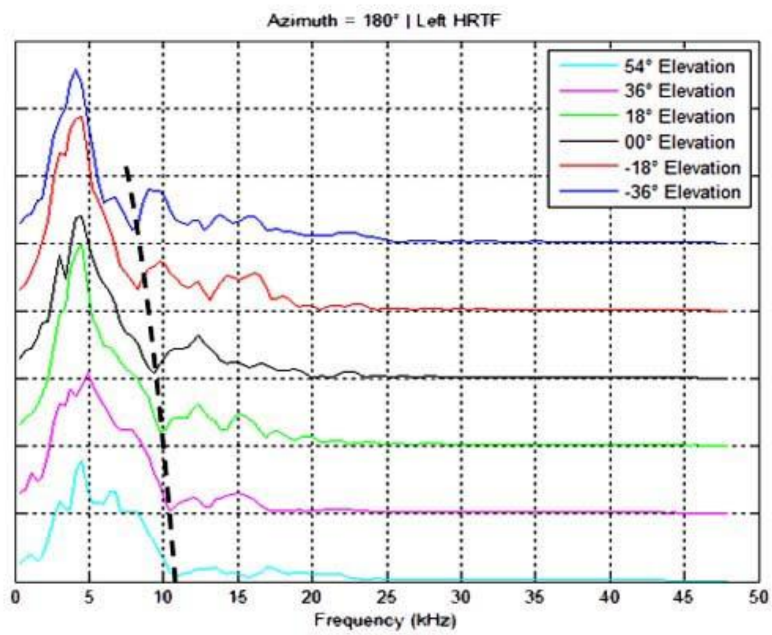


Figure 10. HRTF first notches moved toward higher frequency with increase in elevation

IV. CONCLUSION

We have summarized the movement of the HRTFs peaks and notches with elevations in this paper. We used data from five subjects from FIU HRTF database [6]. Our results show that for notches, there is a strong correlation between elevation and frequency. For the peaks, this relationship is not uniform and is inconsistent. This is due to the fact that the human concha is very complex and the interaction of the sound waves in the concha is not easy to model. This relationship can be used to build mathematical models to describe HRTFs.

References

[1] K. J. Faller II, A. Barreto, N. Gupta and N. Rishe, "Performance Comparison of Two Identification Methods for Analysis of Head Related Impulse Responses", CISSE 2005.

[2] Mills, A. W., "Auditory Localization," in Foundations of Modern Auditory Theory, Vol. II (J, V. Tobias, Ed.), pp. 303-348, Academic Press, New York, 1972.

[3] Begault, D. R., "A head-up auditory display for TCAS advisories." Human Factors, 35, 707-717, 1993.

[4] AuSIM, Inc., "HeadZap: AuSIM3D HRTF Measurement System Manual". AuSIM, Inc., 4962 El Camino Real, Suite 101, Los Altos, CA 94022, 2000.

[5] N. Gupta, A. Barreto and C. Ordonez, *"Spectral Modification of Head-Related Transfer Functions for Improved Virtual Sound Spatialization"*, Proceedings of the 2002 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2002), May 13-17, 2002, Orlando FL, pp. II-1953 to II-1956.

[6] N. Gupta, A. Barreto, M. Joshi and Juan Carlos Agudelo, "HRTF Database at FIU DSP Lab", 2010 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2010).

[7] Algazi, V. R., Duda, R. O., Avendano, C., "The Cipic HRTF database", IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 2001.