Modeling of Pinna Related Transfer Functions (PRTF) Using the Finite Element Method (FEM)

Manan Joshi[*1], Navarun Gupta[1], and Lawrence V. Hmurcik[1]
[1]University of Bridgeport, Bridgeport, CT
*Corresponding author: Engineering Technology Building, University of Bridgeport, 221 University Avenue, Bridgeport, CT 06604; mjoshi@my.bridgeport.edu

Abstract: Pinna Related Transfer Functions (PRTFs) are signal processing models that represent the modifications undergone by the acoustic signal as it interacts with the listener's pinna (outer ear). PRTFs are used to generate binaural sounds, especially those coming from an elevation. The complex shape of the pinna causes reflections, diffractions and resonances that give rise to a unique PRTF for each individual. One can measure individual PRTFs using specialized recording systems; however, these systems are prohibitively expensive and restrict the portability of the 3-D sound system. This paper provides an alternative to real PRTF measurement, where the 3-D images of artificial human ears are used to generate the PRTFs using COMSOL – which is a simpler, efficient, accurate and cost effective way. The synthesized PRTFs are compared with the real PRTFs measured at Florida International University's DSP lab. The comparison shows a percentage match and a cross correlation higher than 80%.

Keywords: prtf, hrtf, prtf modeling, binaural sound, 3-D sound.

1. Introduction

1.1 Head Related Transfer Functions (HRTFs)

Head Related Transfer Functions are signal processing models that represent the modifications undergone by the acoustic signal as it travels from a sound source to each of the listener's eardrums. These modifications are due to the interaction of the acoustic waves with the listener's torso, shoulders, head and pinnae, or outer ears. As such, HRTFs are somewhat different for each listener. For a listener to perceive synthesized 3-D sound cues correctly, the synthesized cues must be similar to the listener's own HRTFs.

One can measure individual HRTFs using specialized recording systems; however, these systems are prohibitively expensive and restrict the portability of the 3-D sound system. HRTF-based systems also face several computational challenges. To overcome these problems, sometimes generic HRTFs are used [1]. Due to the loss of individual characteristics, binaural sounds generated using generic HRTFs suffer from higher errors and lower accuracy in localization. Another approach is to customize HRTFs based on anthropomorphic measurements. The sound entering the pinna undergoes several reflective, diffractive and resonant phenomena, which determine the HRTF. Using signal processing tools and statistical analysis, empirical equations can be derived that describe how the HRTFs can be determined by the shape and size of the pinna, head and torso. Pinna plays an important role in localization of sound in the frontal and median plane. Elevation effects can be created by modeling HRTFs based on pinna measurements. This paper studies a new method of modeling HRTFs where a 3-D image of an ear is used. This method is simpler, and cost effective.

1.2 Pinna & Pinna Related Transfer Functions (PRTFs)

The outer ear (pinna) is a complex shaped organ responsible for shaping much of the HRTF in the higher frequencies. The folds of the pinnae cause minute time delays within the range of 0-300 μs [2] that cause the spectral content at the eardrum to differ significantly from that of the sound source. The complex shape of the pinna causes reflections, diffractions and resonances that give rise to a unique PRTF for each individual. The concha is the most important part of the outer ear when it comes to reflections and resonances.

1.3 Research Problem and Scope

One can measure individual PRTFs using these ways:

I. Using specialized recording systems; however, these systems are prohibitively expensive and restrict the portability of the 3-D sound system. PRTF-based systems also face several computational challenges.
II. Another approach is to customize PRTFs based on anthropomorphic measurements. The sound entering the pinna undergoes several reflective, diffractive and resonant phenomena, which determine the PRTF. Using signal processing tools and statistical analysis, empirical equations can be derived that describe how the PRTFs can be determined by the shape and size of the pinna.

To overcome these problems, sometimes generic PRTFs are used. However, due to the loss of individual characteristics, binaural sounds generated using generic HRTFs suffer from higher errors and lower accuracy in localization.
III. Only the 3-D images are used in Finite Element Analysis (FEA) software (COMSOL) to find the PRTF – simpler, efficient, accurate and cost effective way (proposed by this paper).

2. Research Plan

The research includes these steps:

(1) Measurement of Real PRTF of an artificial pinna; (2) Find PRTF of a similar ear model in COMSOL. The model is also developed using COMSOL; (3) Filter PRTF from COMSOL using Moving Average Filter if required; (4) Comparing the results from steps (1) and (3) for similarity and errors (percentage match & cross correlation)

Figure 1. Artificial ears made of paper.

3. Measurement of Real PRTF

The effective empirical measurement of PRTF pairs is carried out in the following sequence: A speaker is placed at known relative positions with respect to the ear for which the PRTFs are being determined, and a known, broad-band audio signal is used as excitation. At FIU’s DSP laboratory, the Ausim3D’s HeadZap HRTF Measurement System [3] is used. 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. Since the Golay codes sequences played are meant to represent a broad-band excitation equivalent to an impulse, the sequences captured in the artificial ear are the impulse responses corresponding to the PRTFs. Therefore these responses are called Pinna-Related Impulse Responses (PRIRs). The system provides these measured PRIRs as a pair of 256-point minimum-phase vectors, and an additional delay value that represents the Interaural Time Difference (ITD).

AUSIM3D HeadZap system is designed for use in reflective, noisy settings typical of offices and laboratories. Instead of using an array of speakers, HeadZap uses a single speaker mounted on a vertically sliding arm. The loudspeaker’s position is fixed for every ring of measurements around the subject, changing only every time measurements for a new elevation are needed. The subject is seated on a rotating stool and adjusts his/her position according to the desired HRTF location.

To increase the signal-to-noise ratio (SNR), the ear is outfitted with blocked meatus microphones which gather considerably more signal than do probe microphones, and permit louder test sequences to be used. To overcome the problem of reflection from the walls and other objects in the measurement room, HeadZap only extracts the direct path contribution from the measurements. The measurements are carried out in a room padded with foam to minimize reflections. Extraction of raw PRTF data is performed by windowing the measurements such that no room reflections are incorporated in the compilation of the final data.

Figure 2. PRTF measurements of the artificial ear using the HeadZap System.

The measurement sequence must be completed for each of the source locations for which a PRIR pair is desired. The different sound source positions are characterized by two descriptive angles of the spherical coordinate system: azimuth and elevation. The horizontal plane is divided into twelve sections, resulting in HRTFs separated by 30º in azimuth. The vertical plane is divided into six intervals at elevations of: +54º, +38º, +18º, 0º and -18º. Thus 60 impulse responses are measured using HeadZap.

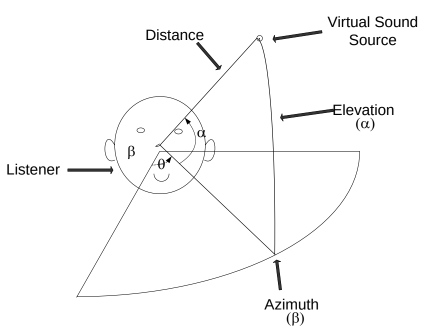

Figure 3. Spherical coordinate system used.

4. Generation of Artificial PRTF using COMSOL Acoustic Module

In this model, a point source generates a pressure wave. The sound level is measured at another point (inside the ear canal of the artificial ear) and at an audible range of frequencies (from 0 Hz to 10 kHz). The measured sound level can be studied to see the changes that the sound waves go through after interacting with the ear/pinna walls. In this paper, instead of using a scanned image of the artificial ear, a model with similar dimensions was drawn in COMSOL. In the second phase, the 3D scanned images from FIU DSP lab will be used.

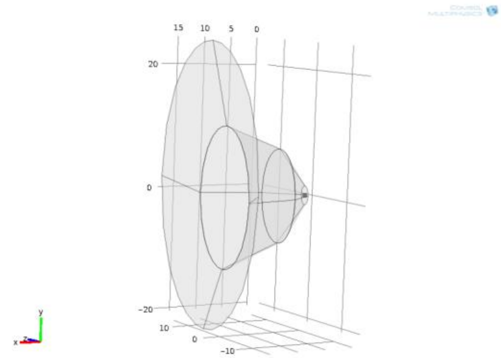

Figure 4. Artificial Ear model in Comsol – 30 mm W x 20 mm D x 50 mm H.

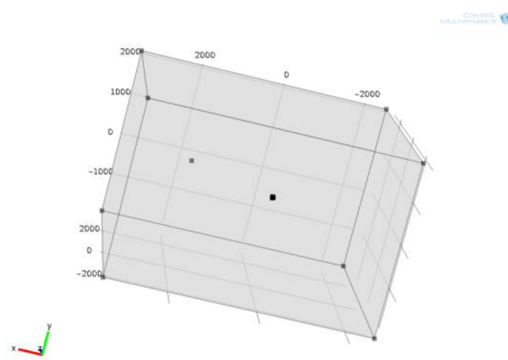

Figure 5. Room (6 m W x 6 m D x 4 m H) – filled with air; Two points seen – (1) Artificial Ear (2) monopole sound source.

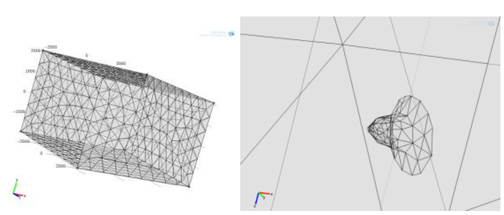

Figure 6. Mesh creation to solve using Finite Element Method

For the harmonic sound waves of acoustic pressure, $p(x,t) = p(x)e^{iwt}$ that is seen in the model above, the following frequency domain Helmholtz equation applies for p(x) [4]:

$$\nabla \cdot \left(-\frac{1}{\rho_0}\nabla p\right) - \frac{\omega^2 p}{\rho_0 c_s^2} = Q \qquad \text{........ (1)}$$

where ρ_0 is the density (kg/m^3), $\omega = 2\pi f$ denotes the angular frequency (rad/s), cs refers to the speed of sound (m/s), and Q (1/s2) is a monopole source. A point source flow of strength S = 10−5 m^3/s located at the point R_0 = (x, y, z) drives the system, so that:

$$Q = \omega S\,\delta^{(3)}(\mathbf{R} - \mathbf{R}_0) \qquad \text{......... (2)}$$

where δ(3)(R) is the 3D Dirac delta function. Assuming, that the ear walls are perfectly reflecting, and the room walls are radiating so that no waves are reflected back into the room from the walls, the sound level is measured at the point R1 = (0, 0, 0) at a range of frequencies from 0 Hz Hz to 10 kHz. The sound pressure at point R1 is studied by varying the location of R0. The angle between R1 and R0 are varied along the horizontal plane (azimuth) and the vertical plane (elevation) as shown below. The sound pressure at R1 represents the simulated Pinna Related Transfer Function (PRTF).

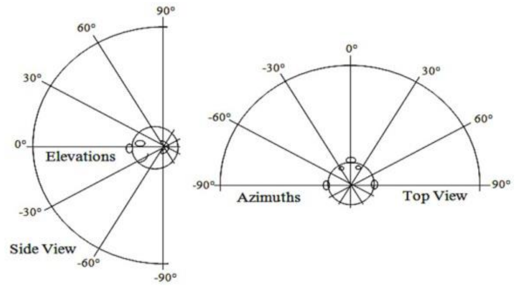

Figure 7. Azimuth and Elevations

The COMSOL generated PRTF is smoothened using Moving Average Filter if required. The moving average filter operates by averaging a number of points from the input signal to produce each point in the output signal.

5. Results

5.1. Measured Real PRTFs of the artificial ear

The real PRIRs are measured for following sound source locations (combination of azimuths & elevations) - Azimuths: 180º, 150º, 120º, 90º, 60º, 30º, 0º, -30º, -60º, -90º, -120º, -150º; Elevations: 54º, 36º, 18º, 0º, -18º.

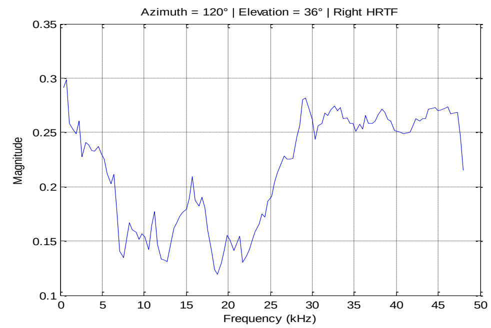

Figure 8. PRTF for artificial ear – sound source location: azimuth: 120º, elevation: 36º

5.2. COMSOL Synthesized PRTFs of an artificial ear with similar structure & dimensions

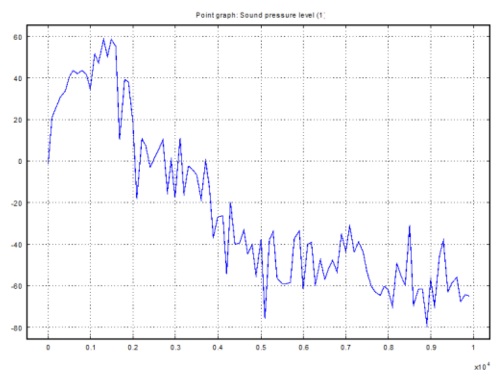

Figure 9. Sound pressure level at R1 (0º azimuth / 30º elevation) – or PRTFs – Range: 0 Hz to 10 kHz at a step size of 100 Hz. The data from COMSOL is exported as a text file

5.3. Comparison

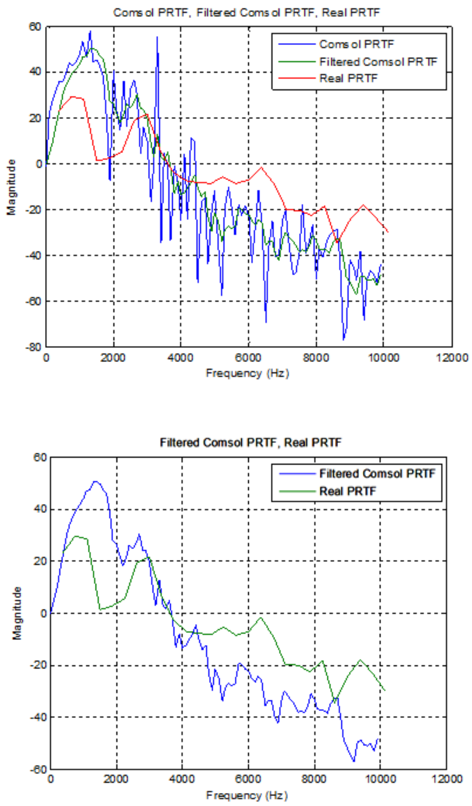

Figure 10. Plots showing the Comsol PRTF, Filtered Comsol PRTF and Real PRTF for sound source location 180º azimuth and 0º elevation. The filtered COMSOL PRTF (Green) and Real PRTF (Red) are similar in the range of 0 to 10 kHz (frequency range of interest). The Moving Average Filter window size is 3

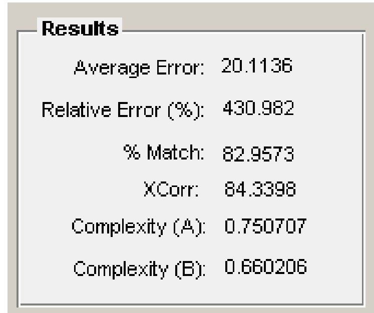

Figure 11. Parameters compared

The cross correlation between the real and artificial simulated HRTF is 84% and the percentage match between the two signals are 82.96.

The accuracy of FEM depends on the largest element size hmax used in the model. The maximal frequency fmax for which the predictions are still reliable is given by:

$$f_{max} = c / (6 * h_{max}) \;\ldots\ldots\ldots\ldots\; (3)$$

Decreasing the element size will reduce the error.

Table 1. Average match between Real and Simulated PRTFs for the artificial ear

Azimuth	180º	180º	180º	180º	180º
Elevation	54º	36º	18º	0º	-18º
% Match	70.8	81.4	**82.9**	70.6	72

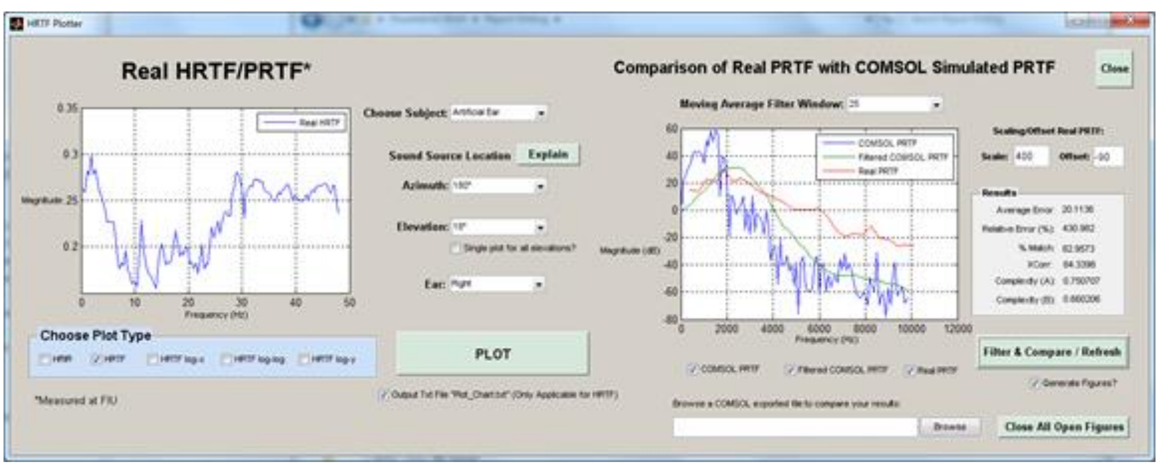

Figure 12. GUI developed using MATLAB to perform comparison and error analysis

6. Conclusions

Pinna Related Transfer Functions (PRTFs) are signal processing models that represent the modifications undergone by the acoustic signal as it travels from a sound source to each of the listener's eardrums. These modifications are due to the interaction of the acoustic waves with the pinnae, or outer ears. PRTFs are somewhat different for each listener. For a listener to perceive synthesized 3-D sound cues correctly, the synthesized cues must be similar to the listener's own PRTFs. To achieve the best localization, the PRTFs have to be measured individually. At the present time, the PRTF recording process is cumbersome, requiring the use of an anechoic chamber and other specialized equipment, which is both expensive and time consuming. An alternative to using individual PRTFs is the use of a standard set of "generic" PRTFs. This gives rather poor elevation results for some percentage of the population [5], but it is all that is practical for inexpensive systems. This paper provides an easier way of measuring PRTFs, where the 3-D images of pinna are used to obtain the PRTFs. % Match and Correlation with the real PRTF up to 80%.

8. References

[1] Gardner, Bill and Martin, Keith. "HRTF measurements of a KEMAR dummy-head microphone", MIT Media Lab, Perceptual Computing, 1994.

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

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

[4] COMSOL Multiphysics Users Guide, Introduction To Acoustics Module, from http://www.comsol.com

[5] Wenzel, E. M., Arruda, M., Kistler, D. J., and Wightman, F. L., "Localization Using Nonindividualized Head-Related Transfer Functions", J. Acoust. Soc. Am., vol. 94, pp. 111-123, 1993

[6] Navarun Gupta, Armando Barreto, Manan Joshi and Juan Carlos Agudelo, "HRTF Database at FIU DSP Lab". In the proceedings of 2010 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2010), March 14-19, 2010, Dallas, Texas.