Horizontal Plane HRTF Interpolation using Linear Phase Constraint for Rendering Spatial Audio

Sandeep Reddy C, and Rajesh M Hegde
Department of Electrical Engineering
Indian Institute of Technology, Kanpur
Email:{csreddy, rhegde}@iitk.ac.in

***Abstract*—In this paper, a novel method of horizontal plane Head Related Transfer Function (HRTF) interpolation is proposed. As Interaural Time Difference (ITD) manifests itself as Interaural phase difference in spectral domain, linear phase constraints are imposed to compute the interpolated HRTFs. Hence the phase of the interpolated HRTF is constrained as a weighted linear combination of the phases of adjacent HRTFs. A weighted $l2$ norm error minimization is performed using these linear phase constraints. The performance of the proposed HRTF interpolation method is evaluated by computing Root Mean Squared Error (RMSE) between the interpolated HRTF and ground truth HRTF obtained from the CIPIC and SYMARE databases. Subjective evaluation is performed on the binaural audio that is rendered using interpolated HRTFs. It is noted from both RMSE and subjective evaluations that the proposed method performs reasonably better than the conventional HRTF interpolation methods.**

I. INTRODUCTION

Binaural synthesis is a technique to synthesize spatial audio, by applying spatial filters called as Head Related Impulse Response (HRIR) to a monophonic sound [1]. Spectral representation of HRIR is called Head Related Transfer Function (HRTF). HRTFs are usually measured by placing an Impulse source at a 3D position, and a miniature microphone in the ear canal [2], [3]. HRTFs accumulate all scattering effects of the pinna, the head and torso and, as a consequence, are able to create the perception of spatially located sound [4], [5]. In rendering binaural audio, it is necessary to have HRTFs at higher angular resolution so that HRTF transition is smooth and thereby avoid any audible artifacts. But physical measurement of HRTF is a tedious task. Therefore, interpolation is required to compute HRTF for arbitrary directions from a set of sparse HRTF measurements.

A number of approaches have been earlier proposed to study this problem. They can be classified as Direct interpolation method and Indirect interpolation method [6]. In direct interpolation method, HRTFs for an arbitrary directions are computed from the HRTFs of adjacent directions. Bilinear interpolation [1], Inter positional transfer function based interpolation [7] are some of the direct approaches. Whereas, in indirect method, HRTFs for different angular directions are combinedly represented using appropriate basis functions. Fourier Bessel Series (FBS) based interpolation [8], [9], principle component weight based interpolation [6], [10], spline interpolation [11] are some of the indirect methods.

Both time and frequency domain interpolation techniques have been investigated in earlier works. Time domain interpolation techniques often result in spurious peaks due to averaging of HRIRs with different onset delay [1]. Spectral domain interpolation is usually performed as linear interpolation of HRTF magnitude spectrum. Here the phase spectrum is generally ignored, and onset delay information is alone considered [12], [13]. Other HRTF interpolation methods have also utilized complex basis functions to represent HRTFs. In these methods HRTFs corresponding to both the ipsilateral and contralateral directions are used to compute the weights of the basis functions. Subsequently, the interpolated HRTFs are obtained from this weights [8], [6].

This paper addresses the issue of horizontal plane HRTF interpolation. Interaural Time Difference (ITD) and Interaural Level Difference (ILD) are dominant cues for sound source localization in the horizontal plane [14]. ITD depends on the absolute time delay at both left and right ear canals, and with the azimuthal angle. This absolute time delays which are captured in HRIR for different azimuthal angles manifest as different slopes in the phase response of HRTF [15]. In earlier techniques HRTF phase component is completely approximated by onset delay. But it has been shown in [16] that, onset delay can only capture the excess phase component of HRTF, but not the phase contributed by the minimum phase component. Therefore, in this work, the complete phase response is utilized when computing interpolated HRTFs over the horizontal plane. This method assumes that the phase spectrum of HRTF varies linearly with azimuthal angles. Hence the interpolated HRTFs are computed by imposing a linear phase constraint. The interpolated HRTF thus obtained are objectively evaluated using Root Mean Square Error (RMSE) analysis. Experiments on rendering binaural audio are also performed using the interpolated HRTFs.

The rest of the paper is organized as follows. Section II describes HRTF interpolation using linear phase constraints. Section III describes the performance evaluation of the proposed method with the conventional techniques. Section IV concludes the paper.

978-0-9928-6265-7/16/$31.00

II. HRTF INTERPOLATION USING LINEAR PHASE CONSTRAINTS

In this section HRTF interpolation using linear phase constraints is discussed. The problem formulation is initially proposed. Subsequently an unconstrained solution is discussed. Finally a linear phase constrained solution is described to solve the HRTF interpolation problem. Analytical results comparing magnitude and phase plots are presented.

A. Problem Formulation

Conventional methods for linear HRTF interpolation utilize a weighted linear combination of HRTF magnitudes of the adjacent angles [12]. However, when HRTF interpolation is performed in horizontal plane, phase spectrum of HRTF assumes significance since onset delay manifests as the slope of the HRTF phase spectra. In this work, the HRTF interpolation problem is formulated as an optimization problem with linear phase constraints. The minimization is performed over HRTFs of adjacent azimuthal angles with linear phase constraints as,

$$\begin{aligned} \underset{H(f_n,\phi_i)\in\mathbb{C}^{\mathbb{N}}}{minimize} \quad & \alpha||H(f_n,\phi_i)-H(f_n,\phi_{i-1})||_2 \\ & +(1-\alpha)||H(f_n,\phi_i)-H(f_n,\phi_{i+1})||_2 \\ subject\ to: \quad & tan^{-1}\left(\frac{H_I(f_n,\phi_i)}{H_R(f_n,\phi_i)}\right)=\alpha tan^{-1}\left(\frac{H_I(f_n,\phi_{i-1})}{H_R(f_n,\phi_{i-1})}\right) \\ & +(1-\alpha)tan^{-1}\left(\frac{H_I(f_n,\phi_{i+1})}{H_R(f_n,\phi_{i+1})}\right) \quad \forall f_n \end{aligned} \tag{1}$$

where $H = H_R + jH_I$ is the complex N-point HRTF and $|| \ ||_2$ represents $l2$ norm. $H(f_n,\phi_{i-1})$ and $H(f_n,\phi_{i+1})$ are known HRTFs of two different azimuthal angles of ϕ_{i-1} and ϕ_{i+1}, respectively. $H(f_n,\phi_i)$ is the unknown HRTF of angle ϕ_i, where $\phi_{i-1} < \phi_i < \phi_{i+1}$. α in Equation 1 is the weight associated with each error term and it can be found by representing ϕ_i as a linear combination of ϕ_{i-1} and ϕ_{i+1} as given in Equation 2.

$$\alpha = \frac{\phi_i - \phi_{i+1}}{\phi_{i-1} - \phi_{i+1}}, \qquad 0 \le \alpha \le 1 \tag{2}$$

B. Unconstrained Solution

A closed form solution to the minimization problem in Equation 1 is not possible, since the constraint is non-convex. An unconstrained solution (UCI) to the minimization problem is discussed herein. It is to be noted that, in Equation 1, minimization is performed with a complex variable. Therefore for an unconstrained problem, the solution is not a weighted linear combination of complex adjacent HRTFs, but as given in Equation 3.

$$H(f_n,\phi_i) = \begin{cases} \frac{H(f_n,\phi_{i-1})+H(f_n,\phi_{i+1})}{2} & if\ \alpha = \frac{1}{2} \\ H(f_n,\phi_{i-1}) & if\ \alpha > \frac{1}{2} \\ H(f_n,\phi_{i+1}) & otherwise \end{cases} \quad \forall f_n \tag{3}$$

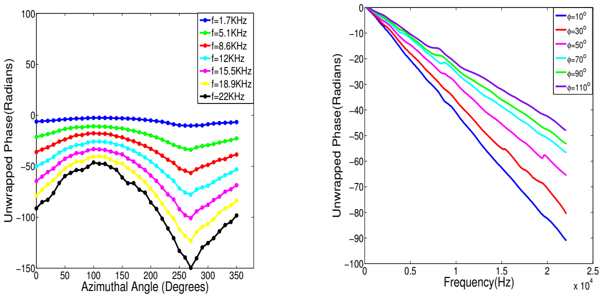

Fig. 1: Unwrapped phase spectrum of HRTF obtained from SYMARE Database. (a)Variation of unwrapped phase with azimuthal angle for different frequencies (b)Variation of unwrapped phase with frequencies for different azimuthal angles.

From Equation 3, it can be noted that when $\alpha \neq \frac{1}{2}$, the solution $H(f_n,\phi_i)$ is quantized to HRTF of the nearest angle. In such cases $H(f_n,\phi_i)$ is computed by finding the average of HRTFs of adjacent azimuthal angles in an iterative manner. It is ensured that the average of the angles converge to ϕ_i at the end of iterative procedure. Hence the interpolated HRTF for an arbitrary α for the unconstrained method is given by,

$$H(f_n,\phi_i) = \bar{\alpha}H(f_n,\phi_{i-1}) + (1-\bar{\alpha})H(f_n,\phi_{i+1}) \quad \forall f_n \tag{4}$$

Selection of $\bar{\alpha}$ in Equation 4 is significant in the proposed unconstrained method. The value of $\bar{\alpha}$ is selected as the nearest element to α and belonging to the set Ω.

$$\Omega = \{0, 1/8, 1/4, 3/8, 1/2, 5/8, 3/4, 7/8, 1\}$$

It is to be noted that the solution in Equation 4 is different when compared to weighted average of complex HRTF. Because weights α in Equation 1 and $\bar{\alpha}$ in Equation 4 are different.

C. Solution using Linear Phase Constraints

In section II-B, we have presented a solution to the unconstrained error minimization to compute interpolated HRTF. It must be noted in Equation 4 that, the weighted average of complex HRTFs is not equal to weighted average of phase responses. Owing to this disadvantage of the unconstrained solution, a solution to the problem in Equation 1 is proposed herein. The Interpolated HRTF phase at angle ϕ_i is constrained as a weighted linear combination of adjacent angle HRTF phases. This weighted linear combination is based on the assumption that HRTF phase varies continuously without any impulsive fluctuations with different azimuthal directions. In order to illustrate this linear variation of HRTF phase with different azimuthal angles, the variation of unwrapped HRTF phase with azimuthal angles for different frequencies is plotted in Figure 1a. The variation of unwrapped HRTF phase with frequency for different azimuthal angles is also illustrated in Figure 1b. It can be noted from Figure 1 that there is a continuous variation of HRTF phase with different azimuthal angles. On the other hand, a change in the slopes of HRTF phases is also noted when the azimuthal angle is varied. Hence the variation of slope of the interpolated HRTF phase is represented as a weighted linear combination of its adjacent angle phases. Having developed an insight on the linear

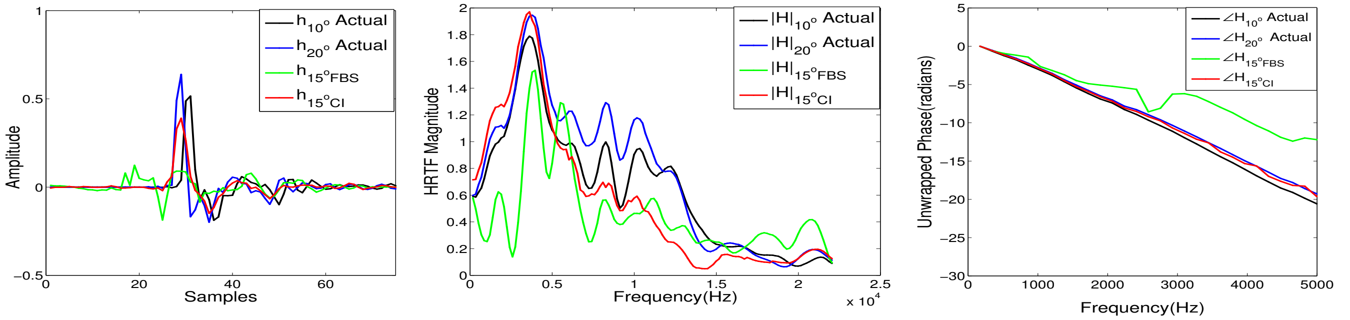

Fig. 2: HRIR, Magnitude and Unwrapped phase spectrum of interpolated HRTF obtained for CI, and FBS using SYMARE Database. The interpolated HRTF was obtained for a angle of 15°whose adjacent angles are 10°and 20°.

variation of HRTF phases, Equation 1 can now be solved as given in Equation 5.

$$\begin{aligned} \underset{H(f_n,\phi_i)\in\mathbb{C}^{\mathbb{N}}}{minimize} \quad & \alpha||H(f_n,\phi_i)-H(f_n,\phi_{i-1})||_2 \\ & +(1-\alpha)||H(f_n,\phi_i)-H(f_n,\phi_{i+1})||_2 \\ subject\ to: \quad & \frac{H_I(f_n,\phi_i)}{H_R(f_n,\phi_i)} = \alpha\frac{H_I(f_n,\phi_{i-1})}{H_R(f_n,\phi_{i-1})} + (1-\alpha)\frac{H_I(f_n,\phi_{i+1})}{H_R(f_n,\phi_{i+1})} \\ & \forall f_n : \frac{-1}{2} <= \frac{H_I(f_n,\phi)}{H_R(f_n,\phi)} <= \frac{1}{2} \end{aligned} \tag{5}$$

It must be noted that the constraints in Equation 5 are a relaxed form of the original constraints used in Equation 1. This relaxation is applied since the original constraints in Equation 1 involve a $\tan^{-1}(x)$ function which is concave for positive x and convex for negative x and therefore can be relaxed as

$$\tan^{-1}(x) \approx x, \quad -1/2 <= x <= 1/2 \tag{6}$$

Equation 5 is solved using numerical methods [17], [18]. The solution is called constraint interpolation (CI) solution.

HRTF interpolation is performed using UCI and CI using Equations 4 and 5, respectively. FBS interpolation is performed using modal parameters M=30 and K=55. FBS is a indirect method, therefore, all the angles of a plane are considered. Interpolated HRTF magnitude and unwrapped phase response computed using CI, UCI, and FBS for CIPIC Database [2] is shown in Figure 3. HRTFs for azimuthal angles -5° and 5° are considered to be known and HRTFs at 0° is computed using above mentioned techniques. It is observed that CI, performs better in both magnitude and phase response. Similar experiments are performed on SYMARE database [19] with known HRTFs at angles 10° and 20°. Interpolation is performed with $\alpha = 0.5$ which resulted in HRTF at 15°. Interpolated HRIR, and its corresponding HRTF magnitude and phase response are shown in Figure 2. It can be seen that phase response of the proposed technique is closer to the actual response, unlike FBS where there is a inconsistency of its variation with respect to actual response. This inconsistency can be seen at lower frequencies in Figure 3(b) which is highlighted as mini plot.

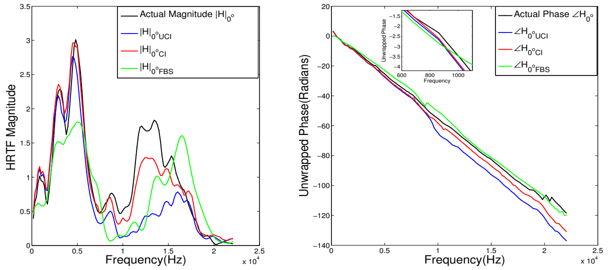

Fig. 3: Magnitude and Unwrapped phase spectrum of interpolated HRTF obtained for UCI,CI, and FBS using CIPIC Database. The interpolated HRTF was obtained for a angle of 0°whose adjacent angles are -5°and 5°

III. PERFORMANCE EVALUATION

In this section performance of the proposed technique is presented. Root mean square error analysis and subjective performance of the proposed technique is compared with the conventional techniques.

A. *Database*

Two sets of databases namely, CIPIC HRTF database [2], and SYMARE database [19] are used in the experiments. CIPIC database include HRIRs for 1250 different directions following interaural polar coordinate system. In this work, HRIRs of the horizontal plane with elevation 0° are considered. Horizontal plane contain 25 HRIRs spaced unequally as, $\phi = [-80^\circ, -65^\circ, -55^\circ, -45^\circ : 5^\circ : 45^\circ, 55^\circ, 65^\circ, 80^\circ]$. Symare database include both acoustic and simulated HRIRs. Acoustic data include HRIRs for 393 directions. Horizontal plane contains HRIRs for 36 directions with an equiangular separation of 10° ranging from 0° to 350°. HRTF interpolation performance is compared using these two databases.

B. *Statistical Analysis*

To study the performance of this techniques for various angles, interpolated HRTF using the proposed method is computed for angles $-40^\circ : 10^\circ : 40^\circ$ using adjacent HRTFs separated by 5°. For example, HRTF corresponding to $\phi = -40^\circ$ is computed using HRTFs of angles -45° and -35°. Root Mean Square Error (RMSE) performance analysis

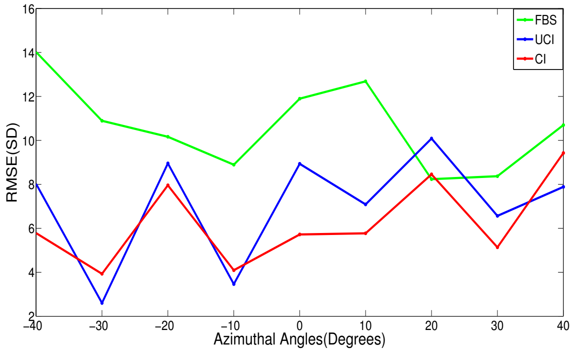

Fig. 4: Variation of RMSE with azimuthal angle

is performed on the obtained and actual HRTF. This error is calculated on log magnitude spectrum [6], [20] as given in Equation 7.

$$SD = \sqrt{\frac{1}{N}\sum_{n=1}^{N}[20log_{10}|H(f_n,\phi)| - 20log_{10}|\hat{H}(f_n,\phi)|]^2} \tag{7}$$

$|H(f_n,\phi)|$ and $|\hat{H}(f_n,\phi)|$ represent the magnitude of the interpolated and actual HRTF respectively at angle ϕ. SD is the RMSE at angle ϕ. This error is plotted for various azimuthal angles for each of the above mentioned techniques as shown in Figure 4. As the RMSE analysis requires ground truth, experiments are performed only on CIPIC database. It can be observed from Figure 4 that RMSE of the proposed technique is lesser when compared to FBS across all the azimuthal angles.

C. Experiments on Rendering Binaural Audio

Binaural audio is synthesized by filtering a monophonic sound using HRTFs of both left and right ears. In order to render binaural audio continuously, interpolation is performed to obtain both left and right ear HRTFs at higher angular resolution. Successive segments of monophonic sound are filtered using HRTFs of successive azimuthal angles. This combined filtered response played through a headphone and evaluated subjectively. The procedure for rendering binaural audio is listed in Algorithm 1. In this work, binaural audio is rendered for two different sound files namely, Chorus and Airplane. An interpolated HRTF is obtained for an azimuthal angle equal to 5°. Adjacent azimuthal angles equal to 0°and 10°are considered in obtaining the interpolated HRTFs. All the three HRTFs corresponding to 0°, 5°and 10°are used in rendering binaural audio as discussed in Algorithm 1. Fifteen subjects are chosen to evaluate the synthesized binaural audio. Each subject was made to listen the rendered audio using actual measured HRTFs of all the three angles. Subsequently, measured HRTFs of 0°, 10°, and interpolated HRTF of 5°are rendered and played through the headphones. Subjects were asked to rate them on a scale of 1 (Bad) to 5 (Excellent) based on five spatial attributes. This attribute set encompassed Naturalness (N), Presence (Ps), Preference (Pf), Source Envelopment (SE) and Perception of Motion (PoM) [21], [22], [23]. The description of attributes is listed below [21], [24].

Algorithm 1 Rendering Binaural Audio using interpolated HRTFs

```
1: Input: Input angles $\phi_{i-1}$,$\phi_i$,$\phi_{i+1}$ and a monophonic sound, x.
2: Pick left and right ear HRTFs of angles $\phi_{i-1}$ and $\phi_{i+1}$ from HRTF Database.
3: Calculate $\alpha$. $\alpha = \frac{\phi_i - \phi_{i+1}}{\phi_{i-1} - \phi_{i+1}}$.
4: If interpolation method is unconstrained, follow step 5-8 Else, jump to step 9.
5: Quantize $\alpha$ to the nearest element of set $\Omega$, say it as $\bar{\alpha}$
6:     $\Omega = \{0, 1/8, 1/4, 3/8, 1/2, 5/8, 3/4, 7/8, 1\}$.
7: $H(f_n,\phi_i) = \bar{\alpha}H(f_n,\phi_{i-1}) + (1-\bar{\alpha})H(f_n,\phi_{i+1})$.
8: Compute left and right HRTFs $(Hl, Hr)$ using Step 7. Jump to step 10.
9: Compute left and right HRTFs $(Hl, Hr)$ using CVX Toolbox.
10: Divide monophonic sound in to three successive segments.
11: Compute the FFT of each segment of monophonic sound, say as $X_{i-1}$,$X_i$,$X_{i+1}$.
12: Filter each segment with HRTFs of successive angles.
13: $X_{lj} = X_j Hl(f_n,\phi_j)$, where $j = i-1, i, i+1$.
14: $X_{rj} = X_j Hr(f_n,\phi_j)$, where $j = i-1, i, i+1$.
15: Concatenate all segments of filtered left channel signal successively to obtain $X_l$ and right channel to obtain $X_r$.
16: Play the left and right channel output through headphones.
17: Output: Binaural Audio.
```

- Naturalness: How true to life the audio listening was.
- Presence: Presence in audio source environment.
- Preference: Degree of pleasantness or harshness.
- Source Envelopment: Sound being all around a person.
- Perception of motion: The precision and correctness with which the trajectory of the source is perceived.

Mean Opinion Scores (MOS) of fifteen subjects for CIPIC and SYMARE Database are listed in Table 1 & 2 respectively. It is observed that although all the techniques crossed the average performance, CI has an edge for the attributes, Perception of Motion and Naturalness over other techniques. It can also be seen that the preference attribute, has a below average performance for FBS method. This can be due to the indirect methodology used in FBS method. Since indirect method utilizes information from HRTFs of all the angles, basis coefficients cannot capture the variation of all the HRTFs. It is also noted that imposing linear phase constraints results in an improvement in both RMSE and subjective evaluation. It is also important to study the computational complexity of CI, as it does not involve a closed form solution. In this work, it is found that CI can perform near real time when computed on a PC with Intel i5 processor and 4Gb RAM.

	N	**Ps**	**Pf**	**SE**	**PoM**
CI-Chorus	4.1	3.6	3.5	3.5	3.7
UCI-Chorus	3.5	3.4	3.1	2.9	3.2
FBS-Chorus	3.2	2.7	2.7	2.3	2.5
CI-Airplane	3.8	3.6	3.2	3.9	3.8
UCI-Airplane	3.2	3.5	3.2	3.9	3.4
FBS-Airplane	3.4	2.8	2.8	2.7	3.2

TABLE I: MOS Scores on CIPIC Database

	N	**Ps**	**Pf**	**SE**	**PoM**
CI-Chorus	4.1	3.3	3.8	3.4	3.7
UCI-Chorus	3.4	3.1	3	2.5	2.7
FBS-Chorus	3.9	3.3	3.3	3.4	3.5
CI-Airplane	3.8	3.5	3.1	3.5	3.4
UCI-Airplane	3.8	2.9	3.4	3.1	3
FBS-Airplane	3.6	3.5	2.9	3.4	3.1

TABLE II: MOS Scores on SYMARE Database

IV. CONCLUSION

In this paper a new method for HRTF interpolation in the horizontal plane is proposed. Since the onset delay manifests itself as the slope of HRTF phase spectrum, a method that uses linear phase constraints is used in the HRTF computation. The proposed technique performs better when compared to the conventional methods for rendering binaural audio. A constrained optimization formulation for HRTF interpolation has been proposed first time in this work. It may be noted that, change in the slopes of HRTF phases is observed when azimuthal angle is varied. A solution to the constrained optimization for HRTF interpolation is proposed by imposing linear phase constraints. This method would be further studied in future for median plane HRTF interpolation.

V. ACKNOWLEDGEMENT

This work was funded in part by TCS Reasearch Scholarship under project number TCS/CS/2011191E and in part by DST project EE/SERB/20130277.

REFERENCES

[1] D.R. Begault, *3-D Sound for Virtual Reality and Multimedia*, AP Professional, 1994.

[2] V.R. Algazi, R.O. Duda, D.M. Thompson, and C. Avendano, “The cipic hrtf database,” in *IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics*, 2001, pp. 99–102.

[3] Dmitry N. Zotkin, Ramani Duraiswami, Elena Grassi, and Nail A. Gumerov, “Fast head-related transfer function measurement via reciprocity,” *The Journal of the Acoustical Society of America*, vol. 120, no. 4, 2006.

[4] V.R. Algazi, R.O. Duda, R.P. Morrison, and D.M. Thompson, “Structural composition and decomposition of hrtfs,” in *IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics*, 2001, pp. 103–106.

[5] A. Sohni, C. Ahuja, and R.M. Hegde, “Extraction of pinna spectral notches in the median plane of a virtual spherical microphone array,” in *4th Joint Workshop on Hands-free Speech Communication and Microphone Arrays (HSCMA)*, May 2014, pp. 142–146.

[6] Lin Wang, Fuliang Yin, and Zhe Chen, “Head-related transfer function interpolation through multivariate polynomial fitting of principal component weights,” *Acoustical Science and Technology*, vol. 30, no. 6, pp. 395–403, 2009.

[7] Fabio P. Freeland, Luiz W. P. Biscainho, and Paulo S. R. Diniz, “Efficient hrtf interpolation in 3d moving sound,” in *Audio Engineering Society Conference: 22nd International Conference: Virtual, Synthetic, and Entertainment Audio*, Jun 2002.

[8] Wen Zhang, Thushara D Abhayapala, and Rodney A Kennedy, “Horizontal plane hrtf reproduction using continuous fourier-bessel functions,” in *Proc. 31st Audio Eng. Soc. Int. Conf.: New Directions in High Resolution Audio*. Citeseer.

[9] Wen Zhang, Thushara D. Abhayapala, Rodney A. Kennedy, and Ramani Duraiswami, “Insights into head-related transfer function: Spatial dimensionality and continuous representation,” *The Journal of the Acoustical Society of America*, vol. 127, pp. 2347 – 2357, 2010.

[10] Bo-Sun Xie, “Recovery of individual head-related transfer functions from a small set of measurements,” *The Journal of the Acoustical Society of America*, vol. 132, no. 1, 2012.

[11] T. Nishino, S. Kajita, K. Takeda, and F. Itakura, “Interpolating head related transfer functions in the median plane,” in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 1999, pp. 167–170.

[12] D.N. Zotkin, R. Duraiswami, and L.S. Davis, “Rendering localized spatial audio in a virtual auditory space,” *IEEE Transactions on Multimedia*, vol. 6, no. 4, pp. 553–564, Aug 2004.

[13] A. Kulkarni, S.K. Isabelle, and H.S. Colburn, “On the minimum-phase approximation of head-related transfer functions,” in *IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics*, Oct 1995, pp. 84–87.

[14] Lord Rayleigh, “On our perception of the direction of a source of sound,” *Proceedings of the Musical Association*, vol. 2, pp. pp. 75–84, 1875.

[15] Sungmok Hwang and Youngjin Park, “Time delay estimation from hrtfs and hrirs,” in *International Conference on Motion and Vibration Control MOVIC*. IEEE, 2006.

[16] Juhan Nam, Miriam A. Kolar, and Jonathan S. Abel, “On the minimum-phase nature of head-related transfer functions,” in *Audio Engineering Society Convention 125*, Oct 2008.

[17] Inc. CVX Research, “CVX: Matlab software for disciplined convex programming, version 2.0,” Aug. 2012.

[18] M. Grant and S. Boyd, “Graph implementations for nonsmooth convex programs,” in *Recent Advances in Learning and Control*, V. Blondel, S. Boyd, and H. Kimura, Eds., Lecture Notes in Control and Information Sciences, pp. 95–110. Springer-Verlag Limited, 2008.

[19] C.T. Jin, P. Guillon, N. Epain, R. Zolfaghari, A. van Schaik, A.I. Tew, C. Hetherington, and J. Thorpe, “Creating the sydney york morphological and acoustic recordings of ears database,” *IEEE Transactions on Multimedia*, vol. 16, no. 1, pp. 37–46, Jan 2014.

[20] Ramani Duraiswami and Vikas C Raykar, “The manifolds of spatial hearing,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing*. IEEE, 2005, vol. 3, pp. iii–285.

[21] Jan Berg and Francis Rumsey, “Systematic evaluation of perceived spatial quality,” in *Audio Engineering Society Conference: 24th International Conference: Multichannel Audio, The New Reality*, Jun 2003.

[22] C. Guastavino, V. Larcher, G. Catusseau, and P. Boussard, “Spatial audio quality evaluation: comparing transaural, ambisonics and stereo,” *13th International Conference on Auditory Display*, 2003.

[23] Catherine Guastavino and Brian F. G. Katz, “Perceptual evaluation of multi-dimensional spatial audio reproduction,” *The Journal of the Acoustical Society of America*, vol. 116, no. 2, 2004.

[24] A. Silzle, S. George, E.A.P. Habets, and T. Bachmann, “Investigation on the quality of 3d sound reproduction,” *Proc. of the International Conference on Spatial Audio, Detmold, Germany,*, 2011.