Binaural Source Localization using a HRTF Data Model with Enhanced Frequency Diversity

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

***Abstract*—A novel method for binaural source localization using a HRTF data model with enhanced frequency diversity is proposed in this work. The method proposed herein is developed in the frequency domain and enhances the directional information in the sound signal for efficient broad band sound source localization even under conditions of low SNR. The directional information is enhanced by cancelling the spectrum of the direct sound source. Subsequently the sound source is localized using a sub space based method. Binaural source localization experiments are performed using HRIR from the CIPIC database and the HRTFs recorded from a customized mannequin developed using 3D printing technology. Experimental results on sound source localization indicate a reasonable improvement in terms of RMSE when compared to state of the art methods.**

I. Introduction

Sound source localization has been investigated extensively over the last decade using an array of sensors. Some of the widely used techniques are correlation [1], [2], beamforming [3], [4] and subspace [5], [6] based methods. Localization has also been performed using different array configurations, like linear [7], planar, and spherical arrays [8]. However, all these techniques differ significantly from the binaural hearing mechanism of the human ear [9], [10]. Conventional methods of source localization utilize time delay of arrival between sensors. On the other hand, the human ear utilizes both time delay and direction dependent spectral cues present in the received sound. The head, torso and asymmetrical shape of outer ear called as pinna are responsible for providing these spectral cues [11], [12]. Additionally, phenomena like torso reflection, head diffraction and pinna reflections are also involved in the complex process [13], [14], [15] of binaural hearing in humans. This complex binaural hearing mechanism is generally characterized by a directional filter and is modelled by the Head Related Impulse Response (HRIR) [16], [17]. Spectral representation of HRIR is called as Head Related Transfer Function (HRTF). HRTF of the two human ears together capture both inter aural time difference and inter aural level difference [18] and can therefore be used in binaural localization of sound sources.

A novel binaural source localization is proposed in this work by utilising HRTF data model with enhanced frequency diversity. A similar method has been proposed in [19]. A broadband direction of arrival (DOA) estimation using spectral cues provided by complex-shaped rigid bodies is also proposed in [20]. In contrast to these methods, the proposed localization method is developed in the frequency domain. Additionally, the proposed method enhances the frequency diversity and directional information in the signal for efficiently localizing broad band sources under low SNR. A frequency domain formulation of the HRTF data model is first provided. The directional information is not completely captured due to limited frequency diversity provided by this model. Hence a HRTF data model which enhances the frequency diversity and directional information in a broad band source signal is proposed next. The directional information is enhanced by cancelling the spectrum of the direct sound source with no directional information. The spectrum of direct sound source with no directional information is obtained from an additional sensor. Source localization is then performed using the proposed HRTF data model and using a sub space based method like MUSIC [5]. Binaural source localization experiments are performed using HRIR from the CIPIC database [21] and the HRTFs recorded from a customized mannequin developed at IIT Kanpur. The results are reasonably better than state of the art localization methods.

The rest of the paper is organized as follows. Section II describes the development of the proposed HRTF model with enhanced frequency diversity. Section III presents the experimental procedures and results of performance evaluation. Section IV concludes the work.

II. A framework for binaural source localization using HRTFs

In this section, a framework for sound source localization utilizing a HRTF model with enhanced frequency diversity is discussed. A frequency domain formulation of the HRTF data model with limited frequency diversity is first provided. Subsequently a HRTF data model which enhances directional information in a broadband source signal is discussed. Source localization is then performed using a subspace based method like MUSIC [5].

A. The HRTF data model for source localization

Sound arriving from a particular direction as received at the entrance of ear canal is given by,

$$y^i(t) = h^i(\Omega_q, t) * s(t), \quad i \in \{l, r\}, \quad \Omega_q = (\theta_q, \phi_q) \quad (1)$$

where $y^l(t)$ and $y^r(t)$ represents the sound received at left and right ear respectively. $h^l(\Omega_q, t)$ and $h^r(\Omega_q, t)$

978-0-9928-6265-7/16/$31.00 ©2016 IEEE

represents left and right HRIRs for an elevation angle θ_q and azimuth ϕ_q. $s(t)$ is the sound emitted from the source and * is a convolution operator. However for processing a broadband signal $s(\tau)$, it needs to be decomposed into equally spaced, non overlapping subband signals as [20], [22].

$$s(\tau) = \frac{1}{\sqrt{T}} \sum_{k=-\infty}^{\infty} S(k,t) e^{jk\omega_0 \tau} \quad t - T \leq \tau \leq t \tag{2}$$

$$where, \quad S(k,t) = \frac{1}{\sqrt{T}} \int_{t-T}^{t} s(\tau) e^{-jk\omega_0 \tau} d\tau \tag{3}$$

$\omega_0 = 2\pi/T$ is the frequency spacing between subbands. HRIRs $h^l(\Omega_q, \tau)$ and $h^r(\Omega_q, \tau)$ are time limited functions. Hence the Fourier series representation of HRIRs is given by,

$$h^i(\Omega_q, \tau) = \frac{1}{\sqrt{T}} \sum_{k'=-\infty}^{\infty} H_q^i(k') e^{jk'\omega_0 \tau} \quad 0 \leq \tau \leq T, i \in \{l, r\} \tag{4}$$

$$where, \quad H_q^i(k') = \frac{1}{\sqrt{T}} \int_0^T h^i(\Omega_q, \tau) e^{-jk'\omega_0 \tau} d\tau \tag{5}$$

From Equation 1, 2, and 4, $y^i(t)$ can be expressed as,

$$y^i(t) = \sum_{k=-\infty}^{\infty} H_q^i(k) S(k,t) e^{jk\omega_0 t} \quad i \in \{l, r\} \tag{6}$$

However, the subband representation of $y^i(\tau)$ is given by,

$$y^i(\tau) = \frac{1}{\sqrt{T}} \sum_{k=-\infty}^{\infty} \tilde{Y}^i(k,t) e^{jk\omega_0 \tau} \quad t - T \leq \tau \leq t \tag{7}$$

At $\tau = t$, Equation 7 can be rewritten as,

$$y^i(t) = \sum_{k=-\infty}^{\infty} Y^i(k,t) e^{jk\omega_0 t} \tag{8}$$

$$where, \quad Y^i(k,t) = \frac{1}{\sqrt{T}} \tilde{Y}^i(k,t) \quad i \in \{l, r\} \tag{9}$$

From Equation 6 and 8, $Y^i(k,t)$ can be expressed as,

$$Y^i(k,t) = H_q^i(k) S(k,t), \quad i \in \{l, r\} \quad \forall k \tag{10}$$

Hence subband coefficients of the observed signal is equal to the product of HRTF coefficients and subband coefficients of source signal. In reality sound received at the ear canal also contain the noise of the background environment. Therefore, Equation 1 can now be expressed as,

$$y^i(t) = h^i(\Omega_q, t) * s(t) + n^i(t) \quad i \in \{l, r\} \tag{11}$$

where $n^l(t)$ and $n^r(t)$ represents noise observed at left and right ear respectively. Hence the subband representation of observed signal can be given by,

$$Y^l(k,t) = H_q^l(k) S(k,t) + N^l(k,t) \quad k = 1, ..K \tag{12}$$

$$Y^r(k,t) = H_q^r(k) S(k,t) + N^r(k,t) \quad k = 1, ..K \tag{13}$$

It may be noted that, Equation 12 and Equation 13 represent an HRTF model for source localization and is comparable to the standard data model in source localization. This model is carried forward to develop an HRTF model with enhanced frequency diversity in the ensuing section.

B. HRTF data model with enhanced frequency diversity

Binaural source localization is usually performed by exploiting the frequency diversity of the HRTFs embedded in the observed signal spectrum. However from Equation 12 and 13 it is seen that, the directional information in the spectral subbands of the observed signal is hidden due to the source and noise spectra. In order to enhance the directional information in the observed subbands, spectrum of the direct sound source is cancelled using an additional sensor. This additional sensor captures the free field response of the source signal with delay t_o and noise components as given in Equation 14.

$$y^a(t) = s(t - t_o) + n^a(t) \tag{14}$$

where $y^a(t)$ represent observed signal at the additional sensor and $n^a(t)$ is the noise observed at additional sensor. The directional information in the observed subband spectrum is enhanced by utilizing the subband spectrum obtained from the additional sensor as,

$$\hat{Y}^l(k,t) = \frac{Y^l(k,t)}{Y^a(k,t)} \quad \hat{Y}^r(k,t) = \frac{Y^r(k,t)}{Y^a(k,t)} \quad \forall k \tag{15}$$

$$where, \quad Y^a(k,t) = S(k,t) e^{-jk\omega_o t_o} + N^a(k,t) \tag{16}$$

After substituting Equation 12 and 13 in Equation 15 we have,

$$\hat{Y}^l(k,t) = H_q^l(k) \hat{S}(k,t) + \hat{N}^l(k,t) \quad k = 1, ..K \tag{17}$$

$$\hat{Y}^r(k,t) = H_q^r(k) \hat{S}(k,t) + \hat{N}^r(k,t) \quad k = 1, ..K \tag{18}$$

$$where, \quad \hat{S}(k,t) = \frac{S(k,t)}{S(k,t) e^{-jk\omega_o t_o} + N^a(k,t)} \quad and, \tag{19}$$

$$\hat{N}^i(k,t) = \frac{N^i(k,t)}{S(k,t) e^{-jk\omega_o t_o} + N^a(k,t)} \quad i \in \{l, r\} \tag{20}$$

It can be observed from Equation 19 that, if the delay term is cancelled and under low noise assumption at additional sensor, the observed subband coefficients consist of the directional component $H_q^i(k)$ and noise $\hat{N}^i(k,t)$. As the noise level at additional sensor increases, directional component in the observed subband coefficients is decreased and thus decreases the localization performance. Delay term $e^{-jk\omega_o t_o}$ is cancelled in Equation 23, which is described in Section II-C.

The new observed subband coefficients of left and right ear as given in Equation 17 and 18, are concatenated and expressed as a matrix form. This results in a HRTF data model with enhanced frequency diversity. This HRTF model can now be written as,

$$\mathbf{Y}_{2K \times 1} = \mathbf{D_q} \mathbf{S}_{K \times 1} + \mathbf{N}_{2K \times 1} \tag{21}$$

$$where, \quad \mathbf{S} = [\hat{S}(1,t) \quad .... \quad \hat{S}(K,t)]^T$$

$$\mathbf{Y} = [\hat{Y}^l(1,t) \quad \hat{Y}^r(1,t) \quad .... \quad \hat{Y}^l(K,t) \quad \hat{Y}^r(K,t)]^T$$

$$\mathbf{N} = [\hat{N}^l(1,t) \quad \hat{N}^r(1,t) \quad .... \quad \hat{N}^l(K,t) \quad \hat{N}^r(K,t)]^T$$

$$and, \mathbf{D_q} = \begin{bmatrix} H_q^l(1) & 0 & ... & 0 \\ H_q^r(1) & 0 & ... & 0 \\ 0 & H_q^l(2) & ... & 0 \\ 0 & H_q^r(2) & ... & 0 \\ 0 & 0 & ... & . \\ . & . & ... & . \\ . & . & ... & . \\ . & . & ... & H_q^l(K) \\ 0 & 0 & ... & H_q^r(K) \end{bmatrix} \quad (22)$$

The final HRTF data model as given in Equation 21 is similar to a standard data model used in source localization. Hence the MUSIC method can be used for source localization in this context. Binaural source localization using MUSIC is discussed in the ensuing section.

C. Binaural source localization using subspace based methods

For uncorrelated source and noise subband coefficients, the covariance of the observed subband coefficients can be decomposed in to two orthogonal subspaces as given below [5] .

$$\mathbf{R} \triangleq E\{\mathbf{YY^H}\} = \mathbf{D_{\hat{S}}\Lambda_{\hat{S}}D_{\hat{S}}^H} + \mathbf{D_{\hat{N}}\Lambda_{\hat{N}}D_{\hat{N}}^H} \quad (23)$$

where $E\{\cdot\}$ is the expectation operator. It may be noted from Equation 23 that, in $\mathbf{YY^H}$, the exponential term $e^{-jk\omega_o t_o}$ present in the source component $\hat{S}(k,t)$ is cancelled due to multiplication with its conjugate under the condition of zero noise at additional sensor. In Equation 23, it can be noted that the noise subspace of the covariance matrix $\mathbf{R}$ is orthogonal to the underlying HRTF at the direction of arrival. Hence the source position is identified as that direction Ω_q for which $\hat{P}(\Omega_q)$ will have a peak value [19]. $\hat{P}(\Omega_q)$ is calculated as,

$$\hat{P}(\Omega_q) = \left\{ \sum_{k=1}^{K} \frac{|\mathbf{d}_q(k)^H \mathbf{P_N} \mathbf{d}_q(k)|}{|\mathbf{d}_q(k)^H \mathbf{d}_q(k)|} \right\}^{-1} \quad (24)$$

where $\mathbf{P_N} = \mathbf{D}_{\hat{N}}\mathbf{D}_{\hat{N}}^H$ and $\mathbf{d_q}(k)$ is the k^{th} column of matrix $\mathbf{D}_q$.

III. Experiments on Source Localization

In this section, performance of the proposed localization method is investigated. Localization experiments are performed using HRIRs from CIPIC database first. Subsequently localization experiments are conducted on a customized mannequin developed at IIT Kanpur (MIPS Mannequin). Results obtained are compared with existing binaural source localization techniques.

A. Experiments on CIPIC Database

Performance of the proposed technique is evaluated using HRIR from the CIPIC database. CIPIC contains HRIRs for 45 subjects and 1250 directions in the 3D space. CIPIC HRIRs were recorded at an elevation resolution of 5.625^o following interaural polar coordinate system. Using CIPIC database proposed technique is evaluated. Spatialized sound is synthesized by convolving a broadband sound signal (chirp signal) with HRIRs of any particular elevation and azimuth angle from CIPIC database. This spatialized signal added with an uncorrelated noise is considered as an observation signal. Since the length of HRIRs in CIPIC database is 200 samples, subband extraction of observed and additional sensor signal is performed on frames of 200 samples. The continuous representations given in Equation 3 and 5 are computed by using their discrete versions over a fixed period. Covariance matrix is computed for observed subband coefficients at different time snapshots. Eigen value decomposition of covariance matrix is performed to obtain the eigen vectors corresponding to noise subspace. Source direction is then computed by identifying a peak in $\hat{P}(\Omega_q)$ computed using Equation 24.

BW↓ SNR	5dB	10dB	20dB
100-1KHz	0.027	0.010	0
1-2KHz	0.209	0.041	0
2-3KHz	0.173	0.050	0
3-4KHz	0.296	0.164	0
4-5KHz	0.295	0.102	0
5-6KHz	0.297	0.297	0.051
6-7KHz	0.297	0.147	0
7-8KHz	0.297	0.297	0.134
8-9KHz	0.297	0.278	0.002

TABLE I
RMSE FOR VARIOUS SOURCE BANDWIDTHS AT DIFFERENT SNR LEVELS

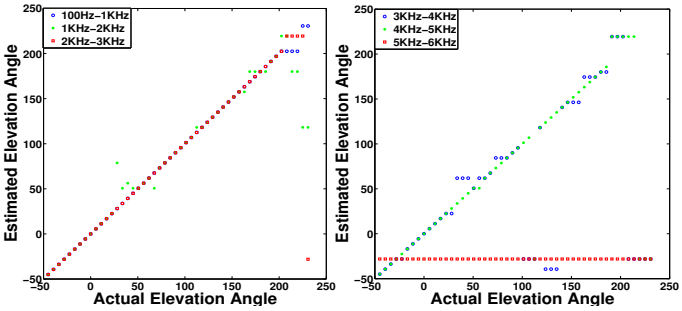

Fig. 1. Scatter plots of estimated and actual elevation angles for various source bandwidths

Performance of the proposed method is evaluated using total root mean square error (TE) as given below.

$$TE = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\sqrt{\sum_{l=1}^{L}\frac{1}{L}[(\theta_{il}-\hat{\theta}_{il})^2 + (\phi_{jl}-\hat{\phi}_{jl})^2]},$$
$$where, \quad \hat{\theta_{il}}, \theta_{il} \in (-\pi, \pi) \quad \& \quad \hat{\phi_{il}}, \phi_{il} \in (-\pi/2, \pi/2) \quad (25)$$

where, M and N are total number of azimuthal and elevation angles respectively and L is the number of iterations performed in estimating elevation $\hat{\theta_{il}}$ and azimuth $\hat{\phi_{il}}$. RMSE is computed using chirp signal as a source of various bandwidths and at different SNR levels as enumerated in Table I. It has to be noted that observed signal bandwidth depends on source signal bandwidth. Hence RMSE results of Table I indicate the performance due to the localization cues present in the

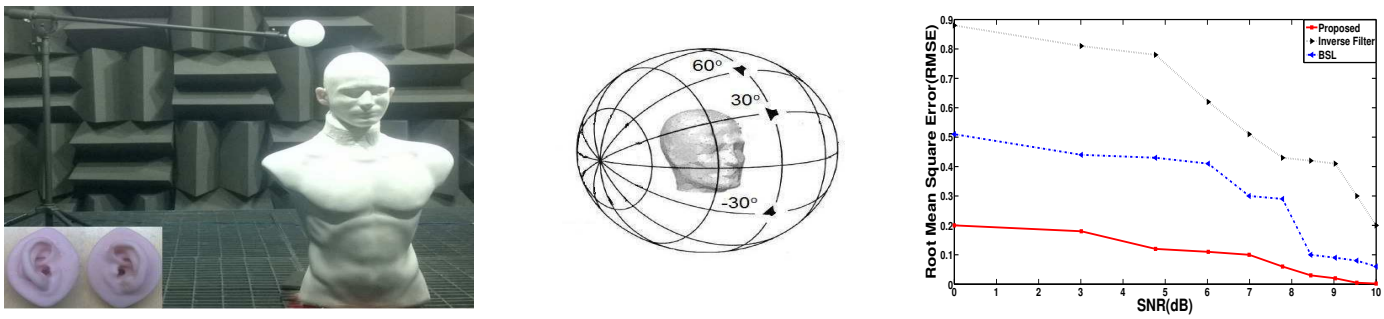

Fig. 2. (a) Figure Illustrating (a) MIPS mannequin in anechoic chamber (b) Sound source placed at different elevation angles in the median plane (c) Total RMSE for the HRIRs measured using MIPS mannequin

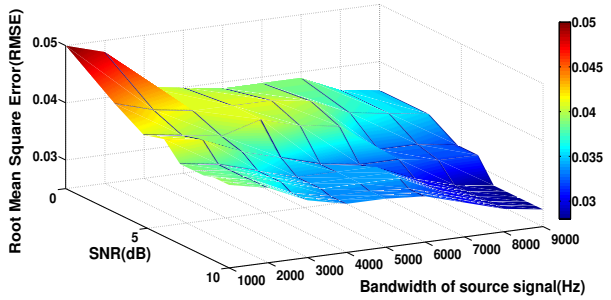

Fig. 3. Total RMSE for various source bandwidth and SNR utilizing CIPIC database

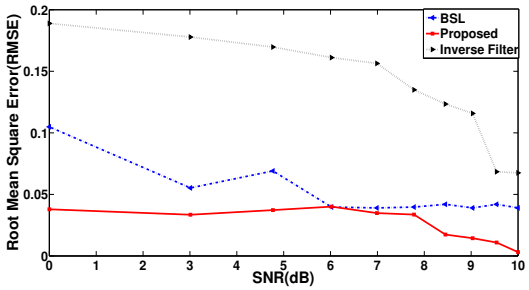

Fig. 4. Total RMSE for various SNR for a source bandwidth of 100Hz-9000Hz utilizing CIPIC Database.

observed signal for that corresponding bandwidth. It can be observed from Table I that, spectral cues present in the lower frequencies (100 Hz to 2 KHz) provide best localization performance. Performance due to individual directions for various bandwidths is illustrated in Figure 1. It can be observed that estimated elevation angle matches with actual elevation angle for most of the directions, especially in the lower frequencies. It must be noted that horizontal plane localization between -90^o and 90^o is performed using time delay techniques [2]. Since the conventional techniques cannot perform elevation identification using just three sensors, results are explicitly described for elevation angles ranging from -45^o and 225^o.

Experiments are then performed with a chirp signal by successively increasing source bandwidth of 1000 Hz at each step by retaining the lower cut of frequency at 100 Hz. RMSE obtained for various bandwidths and SNR is plotted in Figure 3. It can be observed from Figure 3 that, RMSE decreases as SNR and bandwidth increases. This implies that binaural source localization performance is better for a broadband source when compared to a narrowband source. This is so because when sound source is broad band, the observed signal will have more spectral cues as compared to narrow band source. The proposed technique is compared with Inverse filter based localization [23] and Subband BSL [19]. It can be seen from Figure 4 that, the proposed technique has lower RMSE when compared to the above mentioned techniques.

B. Experiments on MIPS-Mannequin

A dummy mannequin developed with customized 3D printed ears at MIPS Lab, IIT Kanpur is used to measure the HRIR. The experimental conditions for measuring HRIR is first described. Subsequently experiments on median plane binaural source localization are performed for various source locations with 30^o elevation resolution.

1) Experimental Conditions: Experiments are also performed on MIPS-mannequin. MIPS-mannequin is a head and torso setup of size equating to a average man. MiPS-mannequin is developed using a styrofoam based head and fibre glass based torso. The left and right ear of MIPS-mannequin are manufactured using Silicone material and 3D printing technology. HRIRs are measured for mannequin with MLS as the excitation signal, using M-Audio miniature microphones (M1255) of diameter 12mm. A 2.1 JBL voyager system is used as a sound source. Experiments are performed in an anechoic chamber of size $5m \times 5m \times 3m$. The complete setup is illustrated in Figure 2a. HRIR measurement is performed by placing sound source at different locations in the median plane with 30^o elevation resolution. A 1024 sample pseudo random MLS sequence is excited with source. HRIRs are extracted from the recorded signals using deconvolution technique as mentioned in [24], [25], [26].

2) Performance Evaluation: A broadband sound signal response (chirp signal) is recorded by placing sound source at an elevation of 30^o, 60^o, -30^o, 120^o, 150^o, and -150^o successively in the median plane. The same microphone

$\hat{\theta}$ \ θ	30^o	-30^o	150^o	-150^o	60^o	120^o
30^o	91	2	8	0	0	0
-30^o	0	86	0	14	0	0
150^o	22	0	77	1	0	0
-150^o	1	8	2	89	0	0
60^o	0	0	0	0	92	8
120^o	0	0	0	0	11	89

TABLE II
CONFUSION MATRIX FOR BROAD BAND SOURCES LOCATED AT ELEVATION ANGLES 30^o,-30^o,150^o,-150^o,60^o,120^o.

speaker setup as used for HRIR measurement is used herein. This is illustrated in Figure 2b where each speaker indicates the position of the source. Azimuthal angle is estimated using the time delay techniques [2] and elevation angle is estimated using the proposed technique. The estimated elevation angle is compared with the ground truth for 100 time snapshots and tabulated as shown in Table II. It can be observed that, estimated elevation angle matches with the actual elevation angle except in a few instances. This confusion is mainly due the unavailability of spectral cues in specific time shots. Total RMSE is plotted for various SNRs as shown in Figure 2c. It can be observed that the proposed technique as lower error compared to the Inverse filter technique [23] and Subband BSL [19]. This improvement in performance is mainly due to enhanced frequency diversity provided by the proposed HRTF data model.

IV. CONCLUSION

This work proposes a HRTF data model for binaural source localization with enhanced frequency diversity. The proposed HRTF data model enhances the frequency diversity and directional information in a broadband source signal by utilizing an additional sensor. It may be noted that the method is able to effectively localize the source in the median plane in contrast to conventional methods of source localization. However, the localization performance is found to depend on the bandwidth of the sound source. Sources with higher bandwidth provide better directional information when compared to sources with lower bandwidth. Future work will investigate issues related to the resolution of closely spaced sources and computational complexity.

V. ACKNOWLEDGEMENT

This work was funded in part by TCS Research Scholarship under project number TCS/CS/2011191E and also supported by the BITCOE IIT Kanpur.

REFERENCES

[1] Jacob Benesty, Jingdong Chen, and Yiteng Huang, *Microphone array signal processing*, vol. 1, Springer Science & Business Media, 2008.
[2] C. Knapp and G. Carter, "The generalized correlation method for estimation of time delay," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 24, no. 4, pp. 320–327, 1976.
[3] Petre Stoica and Randolph L Moses, *Spectral analysis of signals*, Pearson/Prentice Hall Upper Saddle River, NJ, 2005.
[4] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proceedings of the IEEE*, vol. 57, no. 8, pp. 1408 – 1418, aug. 1969.
[5] Ralph O Schmidt, "Multiple emitter location and signal parameter estimation," in *Proceedings of RADC Spectrum Estimation Workshop, Griffiss AFB, NY*, 1979, pp. 243–258.
[6] L. Kumar, A. Tripathy, and R.M. Hegde, "Robust multi-source localization over planar arrays using music-group delay spectrum," *IEEE Transactions on Signal Processing*, vol. 62, no. 17, pp. 4627–4636, Sept 2014.
[7] I. Cohen, J. Benesty, and S. Gannot, "Microphone arrays: Fundamental concepts," *Speech Processing in Modern Communication: Challenges and Perspectives*, vol. 3, 2010.
[8] Dima Khaykin and Boaz Rafaely, "Acoustic analysis by spherical microphone array processing of room impulse responses," *The Journal of the Acoustical Society of America*, vol. 132, no. 1, pp. 261–270, 2012.
[9] John C. Middlebrooks and David M. Green, "Sound localization by human listeners," *Annual Review of Psychology*, vol. 42, no. 1, pp. 135–159, 1991, PMID: 2018391.
[10] D.R. Begault, *3-D Sound for Virtual Reality and Multimedia*, AP Professional, 1994.
[11] 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.
[12] Edgar A. G. Shaw, *Acoustical Characteristics of the Outer Ear*, John Wiley Sons, Inc., 2007.
[13] V. Ralph Algazi, Richard O. Duda, and Patrick Satarzadeh, "Physical and filter pinna models based on anthropometry," in *Audio Engineering Society Convention 122*, May 2007.
[14] Enrique A. Lopez-Poveda and Ray Meddis, "A physical model of sound diffraction and reflections in the human concha," *The Journal of the Acoustical Society of America*, vol. 100, no. 5, 1996.
[15] Vikas C Raykar, Ramani Duraiswami, and B Yegnanarayana, "Extracting the frequencies of the pinna spectral notches in measured head related impulse responses," *The Journal of the Acoustical Society of America*, vol. 118, no. 1, pp. 364–374, 2005.
[16] Jens Blauert, *Spatial Hearing - Revised Edition: The Psychophysics of Human Sound Localization*, The MIT Press, Cambridge,MA, 1996.
[17] Corey I. Cheng and Gregory H. Wakefield, "Introduction to head-related transfer functions (hrtfs): Representations of hrtfs in time, frequency, and space," *J. Audio Eng. Soc*, vol. 49, no. 4, pp. 231–249, 2001.
[18] 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.
[19] Dumidu S. Talagala, Wen Zhang, Thushara D. Abhayapala, and Abhilash Kamineni, "Binaural sound source localization using the frequency diversity of the head-related transfer function," *The Journal of the Acoustical Society of America*, vol. 135, no. 3, 2014.
[20] D.S. Talagala, Wen Zhang, and T.D. Abhayapala, "Broadband doa estimation using sensor arrays on complex-shaped rigid bodies," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 8, pp. 1573–1585, Aug 2013.
[21] 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.
[22] Alan V. Oppenheim, Alan S. Willsky, and S. Hamid Nawab, *Signals &Amp; Systems (2Nd Ed.)*, Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1996.
[23] Justin A. MacDonald, "A localization algorithm based on head-related transfer functions," *The Journal of the Acoustical Society of America*, vol. 123, no. 6, 2008.
[24] Douglas D. Rife and John Vanderkooy, "Transfer-function measurement with maximum-length sequences," *J. Audio Eng. Soc*, vol. 37, no. 6, pp. 419–444, 1989.
[25] Jeffrey Borish and James B. Angell, "An efficient algorithm for measuring the impulse response using pseudorandom noise," *J. Audio Eng. Soc*, vol. 31, no. 7/8, pp. 478–488, 1983.
[26] M. R. P. Thomas, "A novel loudspeaker equalizer," M.S. thesis, May 2006.