

New Insights into the Role of the Head Radius in Model-Based Binaural Speaker Localization

Mehdi Zohourian, Rainer Martin
 Institute of Communication Acoustics
 Ruhr-Universität Bochum, Germany
 Email: {mehdi.zohourian, rainer.martin}@rub.de

Nilesh Madhu
 Mobile Audio Solutions,
 NXP Semiconductors, Leuven, Belgium
 Email: nilesh.madhu@nxp.com

Abstract—In this work we evaluate the effects of the head radius on binaural localization algorithms. We employ a spherical head model and the null-steering beamforming localization method. The model characterizes the binaural cues in the form of HRTFs. One of the main parameters in this model is the head radius. We propose to optimize jointly for both the source location and the head radius. In contrast to the free-field configuration where it is difficult to estimate the source location and microphone distance simultaneously, the binaural algorithm yields a unique solution to the head radius. Moreover, for real recordings we show that the commonly-assumed size of the head achieves a fairly reliable performance. For applications with non-typical size of the head, e.g., hearing-impaired children the adaptation of the head radius using the proposed algorithm would improve the accuracy of the binaural localization algorithm.

Index Terms—Binaural speaker localization, beamforming, hearing aid, DOA

I. INTRODUCTION

Binaural speaker localization has received significant attention in various application, e.g., hearing aids (HAs). With the advent of the wireless link between the left and the right HA it is possible to estimate direction-of-arrival (DOA) even for concurrent speakers.

Some localization algorithms consider the relative phase difference of the left and the right microphone, known as *interaural phase difference* (IPD) [1], [2] whereas others have combined *interaural level difference* (ILD) with the IPD to improve the DOA estimation [3]–[7]. These cues are mostly characterized in the form of *head-related transfer functions* (HRTFs). Some features of the human body such as the size of the head, pinnae, and torso affect the HRTFs and cause uncertainties on the DOA estimation. In order to reduce the sensitivities to these factors some researchers have developed HRTF models [8]–[10] which basically approximate the head by a rigid sphere. One of the main parameters in the model is the head radius that characterizes the inter-microphone distance. The topic of microphone position calibration has been broadly discussed for microphones arrays. In [11] a multi-dimensional scaling approach has been used which assigns

This work has received funding from the People Programme (Marie Curie Actions) of the European Unions Seventh Framework Programme FP7/2007-2013/ under REA grant agreement no PITN-GA-2012-31752, and from the European Fund for Regional Development, grant no EFRE-0800372 (RaVis-3D).

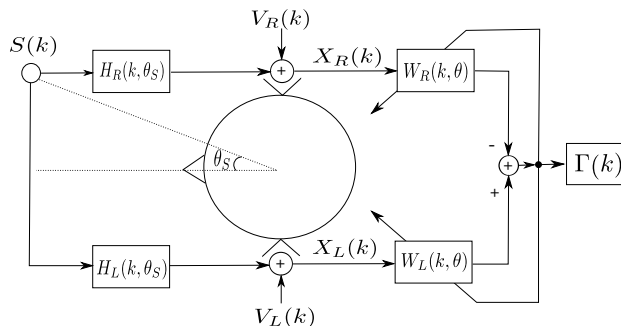


Fig. 1. Binaural signal model and localization algorithm.

microphones to basis points determined by speakers. The utilization of a diffuse noise field is another solution proposed in [12] in which the authors fit the measured noise coherence with the theoretical model and estimate the inter-microphone distance. The authors in [13] have investigated this problem for HAs in the absence of the head. They use two microphones per HA device and calibrate using the mean-square error across all possible source locations around the head.

In this paper we examine the sensitivity of a binaural localization system w.r.t. the head radius. We use a spherical head model [9] in a binaural localization algorithm based on the null-steering beamforming approach [14]. We prove that the joint optimization of DOA and the head radius is analytically possible when the model perfectly matches the real HRTFs. Furthermore, for real recordings we show that the common assumption for the radius of the head of $a = 8.75\text{cm}$ [9] is a valid assumption for localization at all frequencies and across all source positions.

II. BINAURAL SIGNAL MODEL

In our scenario we consider binaural signals from a single source received by the microphones of two HAs. Using the convolution operator $*$ the received signal at the left (L) and right (R) microphone is written as

$$x_{L/R}(n) = h_{L/R}(n) * s(n) + \nu_{L/R}(n) \quad (1)$$

where $s(n)$ denotes a point source signal, $h_{L/R}(n)$ indicates a binaural room impulse response (BRIR) from the source to the left and right microphones, $\nu_{L/R}(n)$ is the noise at each microphone, and n is the sampling index. To analyze signals

in the STFT domain, we take a K -point discrete Fourier transform (DFT) on overlapped and windowed signal frames. Using matrix notation we thus obtain

$$\mathbf{X}(k, b) = \mathbf{H}(k, \theta_S)S(k, b) + \mathbf{V}(k, b) \quad (2)$$

where $\mathbf{H}(k, \theta_S) = [H_L(k, \theta_S), H_R(k, \theta_S)]^T$ is a vector of HRTFs from the source at position θ_S to both the left and right ears. In this equation $\mathbf{X}(k, b) = [X_L(k, b), X_R(k, b)]^T$ and $\mathbf{V}(k, b) = [V_L(k, b), V_R(k, b)]^T$ are vectors of the received signal and the ambient noise, where (k, b) indicate frequency and frame indices. For the sake of simplicity, the time index b is eliminated in the equations below.

The binaural cues used in our localization system are characterized in the form of HRTFs and are derived from a spherical head model as described below.

A. Spherical head model

In [9] Brown and Duda propose a spherical head model for the approximation of the ITD and ILD cues. The head model is formed by cascading a first-order recursive head-shadow filter and a propagation delay. The HRTF for the right ear and for a given source angle θ where $\theta = 0$ is at front and $\theta = -\pi/2$ is to the right ear is expressed as

$$\hat{H}_R(\omega, \theta) = \frac{1 + j\frac{\omega}{2\omega_0}\gamma_R(\theta)}{1 + j\frac{\omega}{2\omega_0}} e^{-j\omega\hat{\tau}_R(\theta)}. \quad (3)$$

In this equation we have $\omega_0 = c/a$, where c is the speed of sound, a is the radius of the head. $\gamma_R(\theta)$ and $\hat{\tau}_R(\theta)$ are two angle-dependent parameters that are defined as

$$\gamma_R(\theta) = 1.05 + 0.95 \cos(1.2(\theta - \pi/2)) \quad (4)$$

$$\hat{\tau}_R(\theta) = \begin{cases} -\frac{a}{c} \sin(\theta) & \text{if } -\pi/2 \leq \theta < 0 \\ -\frac{a}{c} \theta & \text{if } 0 \leq \theta < \pi/2. \end{cases} \quad (5)$$

Then, for the left ear we have

$$\gamma_L(\theta) = 1.05 + 0.95 \cos(1.2(\theta + \pi/2)) \quad (6)$$

$$\hat{\tau}_L(\theta) = \begin{cases} \frac{a}{c}(\theta) & \text{if } -\pi/2 \leq \theta < 0 \\ \frac{a}{c} \sin(\theta) & \text{if } 0 \leq \theta < \pi/2. \end{cases} \quad (7)$$

III. BINAURAL SPEAKER LOCALIZATION

Our approach is based on null-steering beamformer (NBF) where the null-steering vector first equalizes the binaural signals and then scans the room and searches for candidates minimizing its output power. This approach is within the class of least-mean-square localization algorithms and its variants for instance *adaptive eigenvalue decomposition (AED)* method [2], [15]. When we apply this method to the binaural localization context [14], the NBF cost function is given by

$$\Gamma_{NBF}(k, \theta) = E\{|X_L(k)W_L(k, \theta) + X_R(k)W_R(k, \theta)|^2\}. \quad (8)$$

Expanding this equation and using matrix notation we have

$$\Gamma_{NBF}(k, \theta) = \mathbf{W}^H(k, \theta)\Phi_{\mathbf{X}\mathbf{X}}(k)\mathbf{W}(k, \theta) \quad (9)$$

where $\Phi_{\mathbf{X}\mathbf{X}}(k) = E\{\mathbf{X}(k)\mathbf{X}^H(k)\}$ is the spatial covariance matrix and $\mathbf{W}(k, \theta) = [W_L(k, \theta), W_R(k, \theta)]^T$ is the null-steering beamformer. In this approach the NBF filters are designed such that each channel is equalized independently. An optimal solution to the minimization problem in (9) is the eigenvector corresponding to the zero eigenvalue of $\Phi_{\mathbf{X}\mathbf{X}}$ in the noiseless case. Therefore, the beamformer is given by

$$\mathbf{W}_{NBF}(k, \theta) = \left[\frac{1}{H_L(k, \theta)}, \frac{-1}{H_R(k, \theta)} \right]^H \quad (10)$$

Since the head model is employed to approximate HRTFs, we replace $H_{L/R}(k, \theta)$ with $\hat{H}_{L/R}(k, \theta)$ in the above equation. We also use an instantaneous estimate of the covariance matrix in (9) and exploit the phase of the received signals. Therefore, we write

$$\Gamma_{NBF}(k, \theta) = \left| \frac{X_L(k)}{\hat{H}_L(k, \theta)} \right|^2 + \left| \frac{X_R(k)}{\hat{H}_R(k, \theta)} \right|^2 - 2\Re \left(\frac{|X_L(k)|}{|\hat{H}_L(k, \theta)|} \frac{|X_R(k)|}{|\hat{H}_R(k, \theta)|} e^{j\Delta\phi(k, \theta)} \right), \quad (11)$$

where $\Delta\phi(k, \theta) = (\phi_{X_R}(k) - \phi_{X_L}(k) - \Omega_k(\hat{\tau}_R(\theta) - \hat{\tau}_L(\theta)))$ and $\Omega_k = 2\pi k f_s / M$, where M is the number of frequency bins and f_s is the sampling frequency. Here, \Re denotes the real part operator.

We now introduce $A(k, \theta) = \frac{|X_L(k)||\hat{H}_R(k, \theta)|}{|X_R(k)||\hat{H}_L(k, \theta)|}$ and express the cost function as

$$\Gamma_{NBF}(k, \theta) = \frac{|X_L(k)||X_R(k)|}{|\hat{H}_L(k, \theta)||\hat{H}_R(k, \theta)|} \times \left(A(k, \theta) + \frac{1}{A(k, \theta)} - 2 \cos(\Delta\phi(k, \theta)) \right) \quad (12)$$

Since $A(k, \theta) \geq 0$, the function $f(A) = A(k, \theta) + \frac{1}{A(k, \theta)}$ which represents the effects of ILD deviations is always positive. It thus attains its minimum value of $f(A) = 2$ for $A(k, \theta) = 1$. Therefore, $\min\{f(A) - 2 \cos(\Delta\phi(k, \theta))\} = 0$ when both the amplitudes and the phases match the HRTF prototype. Since the minimum value of $f(A) - 2 \cos(\Delta\phi(k, \theta))$ is zeros, we may ignore $\frac{1}{|\hat{H}_L(k, \theta)||\hat{H}_R(k, \theta)|}$. We also ignore the microphone signal $|X_L(k)||X_R(k)|$ which are independent of θ and simplify the cost function to

$$\tilde{\Gamma}_{NBF}(k, \theta) = A(k, \theta) + \frac{1}{A(k, \theta)} - 2 \cos(\Delta\phi(k, \theta)). \quad (13)$$

IV. SENSITIVITY OF THE LOCALIZATION ALGORITHM VERSUS THE HEAD RADIUS

From the localization cost function it is easy to see that the radius of the head has an effect on the accuracy of the DOA estimation algorithm. In this section we analyse the robustness of the localization algorithm in terms of varying the assumed head radius. To motivate the approach we use real HRTFs measured with in-the-ear (ITE) hearing aids [16] in order to render the binaural signals. Figure 2 and 3 illustrate the effect of the variation in the head radius on the the root mean-square error (RMSE) of the narrowband and broadband

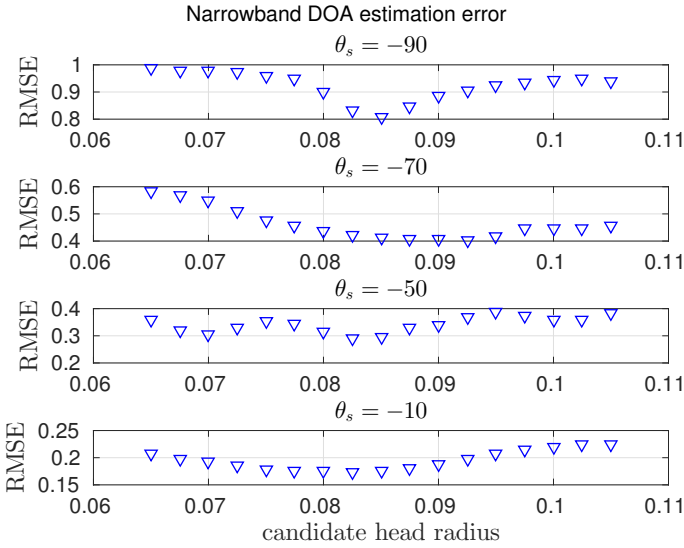


Fig. 2. Narrowband DOA estimation error versus head radius for several DOAs θ_S . Acoustic data was rendered using real HRTFs [16].

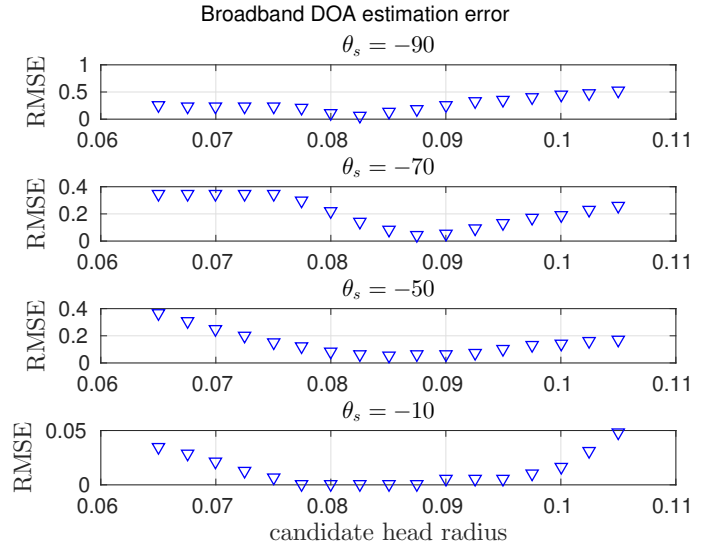


Fig. 3. Broadband DOA estimation error versus head radius for several DOAs θ_S . Acoustic data was rendered using real HRTFs [16].

DOA estimation, respectively. The broadband DOA estimation is determined by averaging the cost function (13) across all frequency bins and then taking the global minimum and discarding local minima. It is observed that the error increases with a small variation in the head radius in both cases. To analyse this issue theoretically, we insert the binaural cues given by the spherical head model in (13). Furthermore, we add the head radius a as a parameter to the cost function. For simplicity we drop the frequency index k . Hence, we rewrite the ILD and IPD contribution in the cost function as

$$A(\theta, a) = \frac{|X_L|}{|X_R|} \sqrt{\frac{1 + \mu^2 a^2 \gamma_R^2}{1 + \mu^2 a^2 \gamma_L^2}} \quad (14)$$

$$\Delta\phi(\theta, a) = \phi_{X_L} - \phi_{X_R} - \frac{\omega a}{c}(\theta + \sin\theta)$$

where $\mu = 2\omega/c$. Note that for $\theta = 0$ both $A(\theta, a)$ and $\Delta\phi(\theta, a)$ do not depend on the head radius a .

In order to optimize jointly the cost function (13) w.r.t. the angle and the head radius we set the gradient of $\tilde{\Gamma}$ to zero,

$$\begin{pmatrix} \frac{\partial \tilde{\Gamma}}{\partial \theta} \\ \frac{\partial \tilde{\Gamma}}{\partial a} \end{pmatrix} = \begin{pmatrix} \left(\frac{A^2-1}{A^2}\right) \frac{\partial A}{\partial \theta} + 2 \sin(\Delta\phi) \frac{\partial \Delta\phi}{\partial \theta} \\ \left(\frac{A^2-1}{A^2}\right) \frac{\partial A}{\partial a} + 2 \sin(\Delta\phi) \frac{\partial \Delta\phi}{\partial a} \end{pmatrix} = \mathbf{0} \quad (15)$$

We solve the first equation for $\left(\frac{A^2-1}{A^2}\right)$ and insert the result in the second equation which is then written as:

$$2 \sin(\Delta\phi) \left(-\frac{\frac{\partial \Delta\phi}{\partial \theta} \frac{\partial A}{\partial a}}{\frac{\partial A}{\partial a}} + \frac{\partial \Delta\phi}{\partial a} \right) = 0. \quad (16)$$

We also solve the first equation for $2 \sin(\Delta\phi)$ and insert the result in the second equation which results in

$$\frac{A^2-1}{A^2} \left(\frac{\partial A}{\partial a} - \frac{\frac{\partial A}{\partial \theta} \frac{\partial \Delta\phi}{\partial \theta}}{\frac{\partial \Delta\phi}{\partial \theta}} \frac{\partial \Delta\phi}{\partial a} \right) = 0 \quad (17)$$

Solving (16, 17) together yields the following solutions

$$\begin{cases} \left(\frac{A^2-1}{A^2} = 0, \quad 2 \sin(\Delta\phi) = 0 \right) \\ \text{or} \\ \frac{\partial \Delta\phi}{\partial \theta} \frac{\partial A}{\partial a} - \frac{\partial \Delta\phi}{\partial a} \frac{\partial A}{\partial \theta} = 0, \end{cases} \quad (18)$$

which are simplified to

$$\begin{cases} (A^2 = 1, \quad \Delta\phi = 0) \\ \text{or} \\ \frac{1 + \cos\theta}{\theta + \sin\theta} = \frac{(\gamma_R \frac{\partial \gamma_R}{\partial \theta} - \gamma_L \frac{\partial \gamma_L}{\partial \theta}) + \mu^2 a^2 \gamma_L \gamma_R (\gamma_L \frac{\partial \gamma_R}{\partial \theta} - \gamma_R \frac{\partial \gamma_L}{\partial \theta})}{\gamma_R^2 - \gamma_L^2}. \end{cases} \quad (19)$$

We solve these equations under the assumption that the model perfectly matches the real HRTF except for the head radius. In other words, the head model is employed to render the binaural signal from the source at θ_S with an actual head radius a_S in a noise-free environment, i.e., $\mathbf{X}(k, b) = \hat{\mathbf{H}}(k, \theta_S, a_S)S(k, b)$. Therefore, the first condition in (19) leads to

$$\begin{cases} \frac{1 + \mu^2 a_S^2 \gamma_L^2}{1 + \mu^2 a_S^2 \gamma_R^2} = 1 \\ \frac{\omega a_S}{c}(\theta_S + \sin\theta_S) - \frac{\omega a}{c}(\theta_S + \sin\theta) = 0, \end{cases} \quad (20)$$

which may be written as

$$\begin{cases} a = \sqrt{\frac{a_S^2 (\gamma_L^2 - \gamma_R^2)}{\mu^2 a_S^2 (\gamma_L^2 \gamma_R^2 - \gamma_R^2 \gamma_L^2) + \gamma_L^2 - \gamma_R^2}} \\ a = a_S \frac{\theta_S + \sin\theta_S}{\theta + \sin\theta}. \end{cases} \quad (21)$$

The unique solution for this equation is $[\theta = \theta_S, a = a_S]$.

Moreover, the second condition in (19) may be written as

$$a^2 = \frac{g(\theta) (\gamma_L^2 - \gamma_R^2) - 5.8 \cos(2.4\theta) - 2.2 \cos(1.2\theta)}{\mu^2 \gamma_L \gamma_R (2.2 \cos(1.2\theta) + 0.54 \sin(2.4\theta) - 0.31)}, \quad (22)$$

where $g(\theta) = \frac{1 + \cos\theta}{\theta + \sin\theta}$. However, according to our numerical analysis (22) yields $a > 13\text{cm}$ which are unrealistic values for

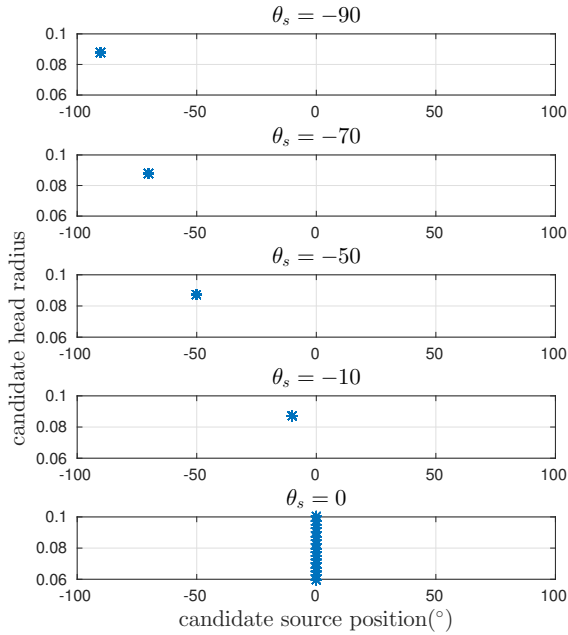


Fig. 4. The joint estimation of DOA and head radius when the head model perfectly matches the real HRTF. The head radius is set to $a = 8.75$ cm.

the head radius. Therefore, the solution attained from the first condition, i.e., (θ_S, a_S) is a unique solution in the assumed search interval for the radius of the head.

In practice however, the head model does not match the real HRTF perfectly and the mismatch of the head radius is not only the source of DOA estimation errors. Therefore, when the source signal is rendered using real HRTFs, the joint optimization of the DOA and head radius does not lead to a closed form solution as in (21) and may show a large error due to this mismatch. Therefore, we investigate the performance of the localization algorithm by adjusting the head radius to the observed data. Our proposal is as follows: 1- An initial head radius is selected and based on this the narrowband and the broadband DOA are estimated. 2- Based on oracle information, reliable frequency bands are selected for which the estimated RMSE of the DOA estimation is less than 5° . 3- The DOA cost function determined by (13) is computed for each possible head radius candidate and summarized across the reliable bins. The head radius candidates are in the interval $[6.5, 10.5]$ cm in steps of 2.5 mm. 4- The resulting cost function is minimized across these head radius candidates. 5- The DOA is re-estimated by substituting the estimated head radius in the DOA cost function (13). The algorithm can be also executed in frequency subbands.

V. EXPERIMENTAL VALIDATION

In the first experiment we examine the joint estimation of DOA and head radius. We render the binaural signals using a white Gaussian noise multiplied by HRTFs for the left and right ear taken from the spherical head model [9]. For this

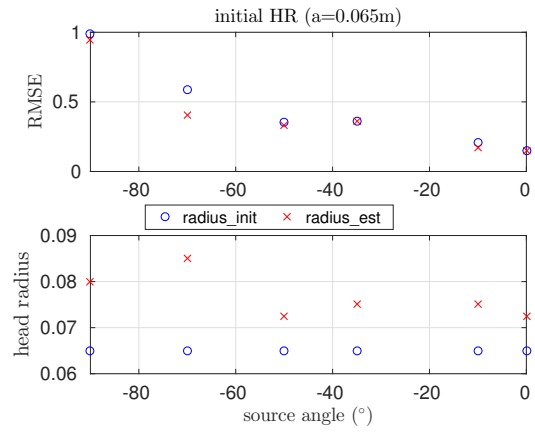


Fig. 5. The effect of adaptation of the head radius on narrowband DOA estimation when using binaural signals measured with ITE hearing aids.

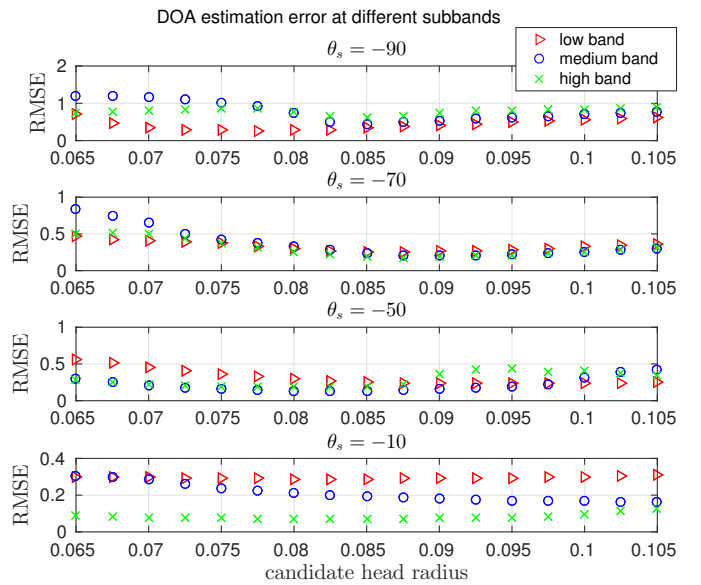


Fig. 6. The effect of variation of head radius on the performance of narrowband DOA estimation in three frequency subbands when using binaural signals measured with ITE hearing aids.

model we consider the radius $a_S = 8.75$ cm, the sampling frequency 16 kHz and 512 FFT bins. Figure 4 shows the result of the grid search for different source positions and across all frequency bins. It reveals that the joint estimate of DOA and head radius occurs uniquely at (θ_S, a_S) except for $\theta = 0^\circ$ where the head radius is not observable. When the head radius is the only source of mismatch between the HRTF and the head model, the estimation of the head radius results in an accurate DOA estimation.

We also investigate the benefit of calibrating the head radius in the DOA estimation for the case where the binaural signals are rendered using real HRTFs. In this experiment we use a speech signal concatenated from male and female utterances with the length of 25 s. The speech signal is convolved with in-the-ear (ITE) hearing aid BRIRs for certain angles taken from the database [16]. Results are presented in Fig. 5. In

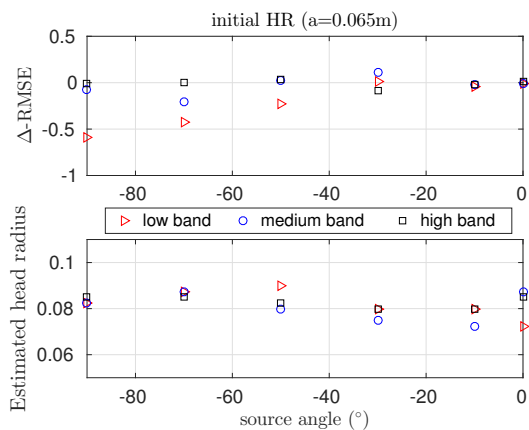


Fig. 7. The effect of tuning the head radius on DOA estimation in three frequency subbands when using binaural signals measured with ITE hearing aids.

In this figure we consider an initial head radius $a = 6.5\text{cm}$ and then optimize the head radius within an interval of $[6.5, 10.5]$ cm in steps of 2.5mm using the proposed algorithm. Results are shown in terms of RMSE for different source positions. According to this figure the optimum head radius that gives best DOA estimates is not necessarily 8.75cm and it depends on where the source is located. For lateral angles the adaptation of the radius has more impact on the performance of the DOA estimation than the frontal angles.

In Fig. 6 we analyze the effect of the variation of the head radius in DOA estimates for different frequency bands. We consider three subbands including low, medium and high frequencies. These frequency subbands are defined as $0 - 2.6\text{kHz}$, $2.6 - 5.2\text{kHz}$, and $5.2 - 8\text{kHz}$. As it is expected the low band is more affected by the variation of the radius as compared to the medium and high bands. The reason is that IPD cues which play major roles in low frequencies are more affected by the variation of the head radius than the ILD cues. In general, we find, however, that the minima in the RMSE measure is relatively shallow.

We also investigate the effect of the adaptation of the head radius in different bands for a given size of the head in the BRIR database [16]. Results are shown in Fig. 7. In this figure, we measure the difference of RMSE of the DOA estimation before and after calibrating the head radius. Again we set the initial head radius to 6.5cm . As can be seen the RMSE improves ($\Delta\text{-RMSE}$ becomes negative) when the radius is optimized. The RMSE reduction is more pronounced for lateral angles and in low frequencies.

VI. CONCLUSION

This paper explores the importance of the head radius as an essential parameter in model-based binaural speaker localization. In contrast to the free-field microphone array where the joint estimation of the DOA and microphone distance is not possible, we showed analytically and by rendering the signal using the head model that the binaural localization algorithm enables to optimize jointly these two parameters. The result is unique due to the interaction of both IPD and ILD in the localization cost function. Furthermore, we investigate

the adaptation of the head radius in DOA estimation for real scenarios. Results show that the commonly-cited radius of the human head, i.e., 8.75cm achieves a relatively good performance across all angles and for different frequencies. Nevertheless, the DOA estimates could be slightly improved by adjusting the head radius parameter. The proposed algorithm could potentially be employed in adaptive binaural DOA estimation for humanoid robots and also for users of hearing aids with head sizes deviating from the standard value. It should be noted that since the source information is not present at all frequency bands at sufficient SNRs, the adaptation of the head radius in localization needs automatic selection of reliable bands which will be a subject for future work.

REFERENCES

- [1] J.H. DiBiase, H.F. Silverman, and M.S. Brandstein, "Robust localization in reverberant rooms," in *Microphone Arrays*, M. Brandstein and D. Ward, Eds., pp. 157–180. Springer, 2001.
- [2] J. Benesty, "Adaptive eigenvalue decomposition algorithm for passive acoustic source localization," *The Journal of the Acoustical Society of America*, vol. 107, no. 1, pp. 384–391, 2000.
- [3] M. Raspaud, H. Viste, and G. Evangelista, "Binaural source localization by joint estimation of ILD and ITD," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 18, no. 1, pp. 68–77, Jan 2010.
- [4] T. May, S. van de Par, and A. Kohrausch, "A probabilistic model for robust localization based on a binaural auditory front-end," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 19, no. 1, pp. 1–13, Jan 2011.
- [5] I. Merks, G. Enzner, and T. Zhang, "Sound source localization with binaural hearing aids using adaptive blind channel identification," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, May 2013, pp. 438–442.
- [6] S. Braun, W. Zhou, and E. A. P. Habets, "Narrowband direction-of-arrival estimation for binaural hearing aids using relative transfer functions," in *Applications of Signal Processing to Audio and Acoustics (WASPAA), 2015 IEEE Workshop on*, Oct 2015, pp. 1–5.
- [7] M. Zohourian, G. Enzner, and R. Martin, "On the use of beamforming approaches for binaural speaker localization," in *Speech Communication; 12. ITG Symposium*, Oct 2016, pp. 1–5.
- [8] R. S. Woodworth and H. Schlosberg, *Experimental psychology*, Oxford and IBH Publishing, 1955.
- [9] C.P. Brown and R.O. Duda, "A structural model for binaural sound synthesis," *Speech and Audio Processing, IEEE Transactions on*, vol. 6, no. 5, pp. 476–488, Sep 1998.
- [10] W. Zhang, R. A. Kennedy, and T. D. Abhayapala, "Efficient continuous HRTF model using data independent basis functions: Experimentally guided approach," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 17, no. 4, pp. 819–829, May 2009.
- [11] S. T. Birchfield and A. Subramanya, "Microphone array position calibration by basis-point classical multidimensional scaling," *IEEE Transactions on Speech and Audio Processing*, vol. 13, no. 5, pp. 1025–1034, Sept 2005.
- [12] I. McCowan, M. Lincoln, and I. Himawan, "Microphone array shape calibration in diffuse noise fields," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 16, no. 3, pp. 666–670, March 2008.
- [13] M. Farmani, R. Heusdens, M. S. Pedersen, and J. Jensen, "TDOA-based self-calibration of dual-microphone arrays," in *2016 24th European Signal Processing Conference (EUSIPCO)*, Aug 2016, pp. 617–621.
- [14] M. Zohourian and R. Martin, "Binaural speaker localization and separation based on a joint ITD/ILD model and head movement tracking," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2016, pp. 430–434.
- [15] G. Enzner, I. Merks, and T. Zhang, "Adaptive filter algorithms and misalignment criteria for blind binaural channel identification in hearing-aids," in *2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*, Aug 2012, pp. 315–319.
- [16] H. Kayser, S. D. Ewert, J. Anemüller, T. Rohdenburg, V. Hohmann, and B. Kollmeier, "Database of multichannel in-ear and behind-the-ear head-related and binaural room impulse responses," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, pp. 6, 2009.