THEORETICAL FRAMEWORK FOR THE DESIGN OF MICROPHONE ARRAYS FOR ROBOT AUDITION

V. Tourbabin and B. Rafaely

Ben-Gurion University of the Negev
Department of Electrical and Computer Engineering
Beer-Sheva 84105, Israel
{tourbabv,br}@ee.bgu.ac.il

ABSTRACT

An important part of a human-like robot is robot audition. Previous work presented systems capable of sound localization and source segregation based on microphone arrays of various configurations. However, no theoretical framework for assessing the quality of these array configurations has been presented. In the current paper such a measure is proposed based on the generalized HRTFs that account for microphone positions other than the ears. The measure is analyzed theoretically with respect to beamforming robustness and DOA estimation accuracy. The measure is then used to find the optimal location of a single microphone and a pair of microphones based on the generalized HRTF database obtained by means of BEM simulation. The results are not surprising, showing that the best position of a single microphone is the ear canal. For a pair of microphones, the results generally show that the sensors should be maximally spatially separated.

***Index Terms*—** Microphone array, robot audition, generalized HRTF, beamforming, DOA estimation.

1. INTRODUCTION

One of the fast evolving fields in recent engineering research is concerned with humanoid robots. These are studied and developed for various applications including service, welfare and entertainment (for a review see, for example [1]). A growing interest in this field is the robotic audition system. This system should be capable of performing or outperforming the everyday tasks of the human auditory system including sound localization [2], source segregation [3] and scene analysis [4].

There are publications describing human-like robotic systems capable of 3D sound localization [5–7] and speech interaction [8], while the sound field is acquired by means of two microphones located at the ears. There are also publications describing systems for sound localization [9] and blind source separation [10] using more than two microphones. The performance of an auditory system that acquires the sound field by means of a microphone array, depends on its configuration which is: (i) the number of microphones and (ii) their positions. From the microphone array processing literature it is generally known that performance improves with increasing the number of microphones [11]. However, to the knowledge of the authors, there is no research reported that is concerned with optimal microphone positioning in the context of robotic audition. Therefore, this question is addressed in the current communication.

In our previous publication [12], a measure of sound localization performance based on the head related transfer functions (HRTFs) was proposed. Here, the definition of the measure is expanded for the generalized HRTFs that account for microphone positions other than the ears. The proposed measure is investigated theoretically for its applicability to beamforming and sound localization systems. Then, a numerical simulation involving the calculation of the generalized HRTFs by means of the boundary element method (BEM) is presented. Finally, using the results of this simulation and the proposed measure of array performance, initial results of optimal microphone positioning for one and two microphones are presented.

2. A MEASURE FOR MICROPHONE DISTRIBUTION QUALITY

Consider a humanoid robot head submerged in a sound field produced by D spatially separated sources. The complex amplitude pressure measured by L microphones positioned on the surface of the head can be modeled by:

$$\mathbf{p} = \mathbf{H}\mathbf{s}, \tag{1}$$

where $\mathbf{p} = [\mathbf{p}_1^T \ \mathbf{p}_2^T \ \cdots \ \mathbf{p}_L^T]^T$ with $\mathbf{p}_n = [p_n(\omega_1) \ \ p_n(\omega_2) \ \cdots \ p_n(\omega_K)]^T$ contains complex pressure amplitudes measured by L microphones at selected frequencies $\{\omega_m\}_{m=1}^K$. The vector $\mathbf{s} = [s(\Omega_1) \ s(\Omega_2) \ \cdots \ s(\Omega_D)]^T$ contains the complex amplitudes of the signals arriving from selected directions $\{\Omega_j\}_{j=1}^D$ as measured at the center of the head when it is removed. Here $\Omega_j = (\theta_j, \phi_j)$ stands for azimuth θ_j and elevation ϕ_j and $(\cdot)^T$ stands for the matrix transpose operator. Matrix $\mathbf{H} \in \mathbb{C}^{LK \times D}$ is the generalized HRTF matrix with element h_{ij} that describes the transfer function between the source indexed by j and the microphone and frequency indexed by i. Measurements taken by different microphones and at different frequencies are all placed along the same dimension in $\mathbf{H}$, to signify the fact that information in $\mathbf{H}$ can be extended either along space or frequency. In addition, this model assumes that the arriving signals are flat, i. e. have the same magnitude and phase in the frequency range of interest. This is in order to analyze the performance of the system in a way that does not depend on the source signal properties.

Our goal is to develop a measure of array quality as a function of microphone positioning that will allow optimization of microphone distribution. One way to do this, is to evaluate the amount of information contained in the measurement $\mathbf{p}$ for the reconstruction of the arriving signal amplitudes $\mathbf{s}$ as a function of the generalized HRTF matrix $\mathbf{H}$. In [12] we proposed to measure this information by defining *effective rank* of the HRTF matrix. Here, the definition is generalized by introducing the parameter α:

$$\mathcal{R}_\alpha(\mathbf{H}) = \exp\left(-\sum_{i=1}^{q} \bar{\sigma}_i \cdot \log \bar{\sigma}_i\right),$$
$$\bar{\sigma}_i = \sigma_i^\alpha / \sum_{j=1}^{q} \sigma_j^\alpha, \ \ \alpha \in [0, \infty). \tag{2}$$

In (2), $\{\sigma_j\}$ denote the singular values of $\mathbf{H}$, $\log(\cdot)$ is the natural logarithm and $q = \mathrm{rank}\{\mathbf{H}\}$.

As discussed in [12], the effective rank is bounded by:

$$1 \leq \mathcal{R}_\alpha(\mathbf{H}) \leq q. \tag{3}$$

Effective rank measures the uniformity of the singular values of $\mathbf{H}$. It is maximized when all singular values are equal, i.e. $\sigma_j = \sigma_1, \ j = 1, 2, ..., q$. In this case the effective rank and the rank of $\mathbf{H}$ are equal. Effective rank is minimized when the first singular value (it is a common practice to arrange the singular values in a descending order [13]) is much larger than the other singular values, i.e. $\sigma_1 >> \sigma_j, \ j = 2, 3, ..., q$. In this case the effective rank of the matrix is near unity.

The parameter α scales the distribution of singular values, i.e. for $\alpha > 1$ the differences between the singular values are emphasized, while for $0 \leq \alpha < 1$ the differences become less prominent. The discussion in [12] was based on efficient rank with $\alpha = 2$. Here, we use $\alpha = 1$, to keep the original distribution of singular values unchanged.

3. SIGNIFICANCE OF THE PROPOSED MEASURE

The complex pressure amplitudes $\mathbf{p}$ measured on the head surface can be utilized for various tasks including spatial filtering and localization. In this section we discuss the significance of uniformity of singular values of the HRTF matrix and its relation to the performance of beamforming and direction of arrival (DOA) estimation algorithms.

3.1. Beamforming

One widely studied data-independent beamforming approach is the maximum-directivity beamformer [11, 14]. Beamformer robustness can be analyzed by means of the beamformer sensitivity measure [11]. Sensitivity of the maximum-directivity beamformer which also maintains a distortionless-response constraint at the look direction, is given by [11]:

$$\mathcal{T}_{maxDI} = \frac{\mathbf{b}^H \mathbf{C}^{-2} \mathbf{b}}{(\mathbf{b}^H \mathbf{C}^{-1} \mathbf{b})^2}, \tag{4}$$

where $\mathbf{b} = \mathbf{v}(\Omega_l)$ is the array steering vector in look direction Ω_l, $(\cdot)^H$ denotes the conjugate transpose operator and $\mathbf{C}$ is the following matrix:

$$\mathbf{C} = \frac{1}{4\pi} \int_{\Omega \in S^2} \mathbf{v}(\Omega)\mathbf{v}^H(\Omega) d\Omega. \tag{5}$$

The integral $\int_{\Omega \in S^2} d\Omega = \int_0^{2\pi} \int_0^{\pi} \sin\theta d\theta d\phi$ covers the entire surface of the unit sphere, denoted by S^2. This implies that $\mathbf{C}$ represents the average of $\mathbf{v}(\Omega)\mathbf{v}^H(\Omega)$ over the unit sphere surface. Assuming that the average can be approximated by a finite sum, we get:

$$\mathbf{C} \approx \frac{1}{D} \sum_{j=1}^{D} \mathbf{v}(\Omega_j)\mathbf{v}^H(\Omega_j) = \frac{1}{D}\mathbf{H}\mathbf{H}^H, \tag{6}$$

where D is the number of selected directions and $\mathbf{H} \in \mathbb{C}^{L \times D}$ is the HRTF matrix at a single frequency, with columns representing the steering vectors of the microphone array positioned on the head surface. In addition, it is assumed that $D \geq L$ and $\mathbf{H}\mathbf{H}^H$ is non-singular. Thus, sensitivity of the-maximum directivity beamformer depends on the HRTF matrix, which in turn depends on microphone positioning. By substituting (6) into (4) and using the SVD of $\mathbf{H}$, i.e. $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$, we obtain:

$$\mathcal{T}_{maxDI} = \frac{\mathbf{b}^H \mathbf{U}\mathbf{\Sigma}^{-4}\mathbf{U}^H \mathbf{b}}{(\mathbf{b}^H \mathbf{U}\mathbf{\Sigma}^{-2}\mathbf{U}^H \mathbf{b})^2} = \frac{\sum_{i=1}^{q} \frac{1}{\sigma_i^4} |\mathbf{u}_i^H \mathbf{b}|^2}{\left(\sum_{i=1}^{q} \frac{1}{\sigma_i^2} |\mathbf{u}_i^H \mathbf{b}|^2\right)^2} \tag{7}$$

where $\mathbf{\Sigma} = \mathrm{diag}\{\sigma_1, \sigma_2, ..., \sigma_q\}$, $\mathbf{U} = [\mathbf{u}_1 \ \mathbf{u}_2 \ \cdots \ \mathbf{u}_L]$ and $\mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \cdots \ \mathbf{v}_D]$ contain the singular values $\{\sigma_i\}$, the left singular vectors $\{\mathbf{u}_i\}$ and the right singular vectors $\{\mathbf{v}_i\}$ of matrix $\mathbf{H}$, respectively. Now, recall that maximizing the effective rank of $\mathbf{H}$ will tend to produce a more uniform distribution of its singular values. Observe, that in the limiting case where all singular values are equal, i.e. $\sigma_i = \sigma_1, \ i = 1, 2, ..., q$, the expression in (7) reduces to

$$\mathcal{T}_{maxDI} = \frac{1}{\sum_{i=1}^{q} |\mathbf{u}_i^H \mathbf{b}|^2} = \frac{1}{\mathbf{b}^H \mathbf{U}\mathbf{U}^H \mathbf{b}} = \frac{1}{\mathbf{b}^H \mathbf{b}}. \tag{8}$$

The result in (8) is the lower bound on sensitivity for the maximum-directivity, distortionless-response beamformer [11], over all variations of the steering vector characterizing the array. This result implies that maximizing the effective rank of the HRTF matrix will generally tend to reduce the sensitivity and therefore improve the robustness of maximum-directivity beamformer.

3.2. DOA estimation

One of the widely studied DOA estimation approaches is based on the multiple signal classification (MUSIC) algorithm [15]. This approach is narrowband and is based on the following measurement model:

$$\mathbf{p} = \mathbf{H}\mathbf{s} + \mathbf{n}, \tag{9}$$

which is identical to (1) with the following exceptions: additive noise $\mathbf{n} \in \mathbb{C}^{L \times 1}$ is included which has zero mean and covariance $\mathrm{E}[\mathbf{n}\mathbf{n}^H] = \sigma\mathbf{I}$, $\mathbf{H}$ is the transfer matrix at a single frequency and it is assumed that $L > D$. A comprehensive statistical analysis of the performance of this method as a function of arriving signal properties and the number of sensors is provided in [16]. Here, we focus on the relation between the performance and the singular values of $\mathbf{H}$. In addition to the assumptions made in [16] we assume that $\mathrm{E}[\mathbf{s}\mathbf{s}^H] = \alpha\mathbf{I}$. In this case the covariance matrix of the measurements $\mathbf{p}$ is given by:

$$\mathrm{E}[\mathbf{p}\mathbf{p}^H] = \alpha\mathbf{H}\mathbf{H}^H + \sigma\mathbf{I} = \alpha\mathbf{U}\mathbf{\Sigma}^2\mathbf{U}^H + \sigma\mathbf{I} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H, \tag{10}$$

where $\mathbf{U}$ and $\mathbf{\Sigma}$ contain the left singular vectors and the singular values of $\mathbf{H}$, as above. Matrix $\mathbf{\Lambda} = \mathrm{diag}\{\lambda_1, \lambda_2, ..., \lambda_L\}$, with $\{\lambda_i\}$ denoting the eigenvalues of $\mathrm{E}[\mathbf{p}\mathbf{p}^H]$, from (10) given by:

$$\lambda_i = \begin{cases} \alpha\sigma_i^2 + \sigma, & i = 1, 2, ..., D \\ \sigma, & i = D+1, ..., L \end{cases}, \tag{11}$$

where $\{\sigma_i\}$ denote the singular values of $\mathbf{H}$ as above. In the special case where the arrival direction is described by a single parameter (e.g. elevation), the variance of estimation error is given by [16]:

$$\mathrm{E}(\hat{\Omega}_j - \Omega_j)^2 = c\sigma \sum_{i=1}^{D} \frac{\lambda_i}{(\sigma - \lambda_i)^2} |\mathbf{h}_j^H \mathbf{u}_i|^2, \tag{12}$$

where $\hat{\Omega}_j$ is the MUSIC estimator of Ω_j, $\mathbf{h}_j$ is the j^{th} column of $\mathbf{H}$, and $\{\mathbf{u}_i\}_{i=1}^{D}$ are the first D left singular vectors of $\mathbf{H}$. The constant c depends on the number of snapshots used for the estimation and the derivative of $\mathbf{h}_j$ with respect to Ω_j (see [16] for the details). We assume that c is approximately independent of j. By substituting (11) into (12) and summing estimation error variances for all arriving directions, we obtain:

$$\begin{aligned} \epsilon &= \sum_{j=1}^{D} \mathrm{E}(\hat{\Omega}_j - \Omega_j)^2 = \frac{c\sigma}{\alpha} \sum_{i=1}^{D} \frac{\sigma_i^2 + \frac{\sigma}{\alpha}}{\sigma_i^4} \sum_{j=1}^{D} |\mathbf{h}_j^H \mathbf{u}_i|^2 \\ &= \frac{c\sigma}{\alpha} \sum_{i=1}^{D} \frac{\sigma_i^2 + \frac{\sigma}{\alpha}}{\sigma_i^4} \|\mathbf{H}^H \mathbf{u}_i\|^2 = \frac{cD}{\mathrm{SNR}} + \frac{c}{\mathrm{SNR}^2} \sum_{i=1}^{D} \frac{1}{\sigma_i^2}, \end{aligned} \tag{13}$$

where $\mathrm{SNR} = \alpha/\sigma$. Only the second factor in (13) depends on the singular values of $\mathbf{H}$. Thus, for high SNR values, the singular values will have no effect on the total estimation error ϵ. However, for low SNR values, the second factor in (13) will be dominant. By considering the constraint of constant matrix gain, i.e. $\sum_{i=1}^{D} \sigma_i^2 = const$, it can be shown using the Lagrange multipliers method that the second factor, i.e. $\sum_{i=1}^{D} 1/\sigma_i^2$ is minimized when $\sigma_i = \sqrt{const/D}$, $i = 1, 2, ..., D$. This implies that total estimation error ϵ is minimized for uniform singular values. Thus, maximizing the effective rank of $\mathbf{H}$ will tend to minimize the MUSIC DOA estimation variance under the above assumptions, especially for low SNR values.

4. GENERALIZED HRTF DATABASE

As a first approach for obtaining the generalized HRTFs through a simulation, the boundary element method (BEM) was utilized. MATLAB code was written implementing the method as described in [17]. Head geometry model used in this simulation was kindly provided by Brian F. G. Katz who used this model for the study reported in [18]. The model was resampled in order to reduce the number of faces to 4000. The calculations were performed up to 5 kHz, although this model is valid up to 3 kHz according to the rule of 6 elements per wavelength [18]. The pressure was obtained for 240 nearly uniformly [19] distributed sources at a distance of 1 m from the head center. It was assumed that the head is acoustically rigid based on results reported in [20]. The obtained database represents the generalized HRTFs for 4000 microphone positions over the head at the frequencies $\{0, 100, 200, ..., 5000\}$ Hz to 240 nearly uniformly [19] distributed directions.

Comparison of the results of current simulation to the results of numerical simulations reported in [18] is presented in Fig. 1. It can be seen that there is a disagreement between the results above 3.5 kHz. This disagreement is most probably due to the fact that the requirement for 6 elements per wavelength is not fulfilled above 3 kHz. Although there is a deviation for higher frequencies, in general the results fit very well. Small differences in the results for the lower frequencies can be explained by the differences in the mesh and the algorithms used to obtain the solutions.

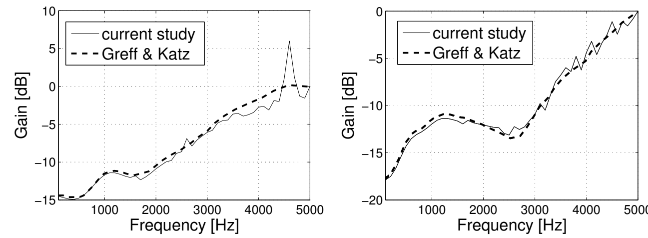

Fig. 1. Comparison of HRTF gain between the current study and the study reported in [18] (a) - front source, left ear, (b) - left source, left ear.

5. OPTIMAL POSITIONING OF ONE AND TWO MICROPHONES

The optimal microphone positioning problem can be formulated as a selection problem. Consider the set of possible microphone positions $\mathcal{M}$. We are to choose the optimal subset $\mathcal{L} = \{l_1, l_2, ..., l_L\} \subset \mathcal{M}$ of distinct positions. The objective function for the selection is the effective rank (see (2)) of the HRTF matrix $\mathbf{H} = [\mathbf{H}_{l_1}^T \ \mathbf{H}_{l_2}^T \ \cdots \ \mathbf{H}_{l_L}^T]^T$, which is the column concatenation of individual sensor transfer matrices as is suggested by (1).

The effective rank of the HRTF matrices of all possible subsets for $L = 1$ and $L = 2$ was evaluated. The elements close to the neck (white elements in the bottom of Fig. 2c) were excluded from calculations. The results are summarized in Fig. 2. Results of single sensor positioning are shown in figures 2a-2c. Recall that the HRTF matrices were obtained for 240 distinct nearly distributed arrival directions. Fig. 2a displays the results where the HRTF matrices with all 240 arrival directions are considered, while Fig. 2b and Fig. 2c represent the results for selected arrival directions: Fig. 2b - only the directions in the horizontal plane, Fig. 2c - only the directions in the median plane. The maximum value of 1 (black) and the minimum value of 0 (white) represent 100% and 80% of the maximum effective rank that was obtained, respectively. The elements for which the HRTF matrices had effective rank less than 80% of the maximum value are colored in white.

It can be seen that the best position of a single microphone (Fig. 2a) is the very entrance to the ear canal. This result is not surprising, as one would expect that the shape of the outer ear and its position is preferable from evolutionary point of view. Fig. 2b shows that when only the sources located in the horizontal plane are important, the preferable microphone positions are distributed on a horizontal strip surrounding the head. The response of these microphones is expected to vary considerably as a function of azimuth while being nearly independent of elevation. Therefore, microphones positioned in the horizontal strip are preferable for discrimination between sources positioned in the horizontal plane. For the same reason, when only sources located in the median plane are important (Fig. 2c), the picture is reciprocal, i.e. the preferable positions are distributed in a vertical strip surrounding the head.

Results of positioning of a pair of microphones are summarized in figures 2d-2f. In order to draw these figures the effective rank of all possible pairs (total $5,822,578$) was calculated. Then, best 1% $(58,226$ pairs) of all pairs was used in order to calculate the presented colormap. Each of the elements was assigned a value in the range of $[0, 1]$ based on the number of its repetitions in the best pairs. Values of 1 and 0 correspond to the maximum an the minimum

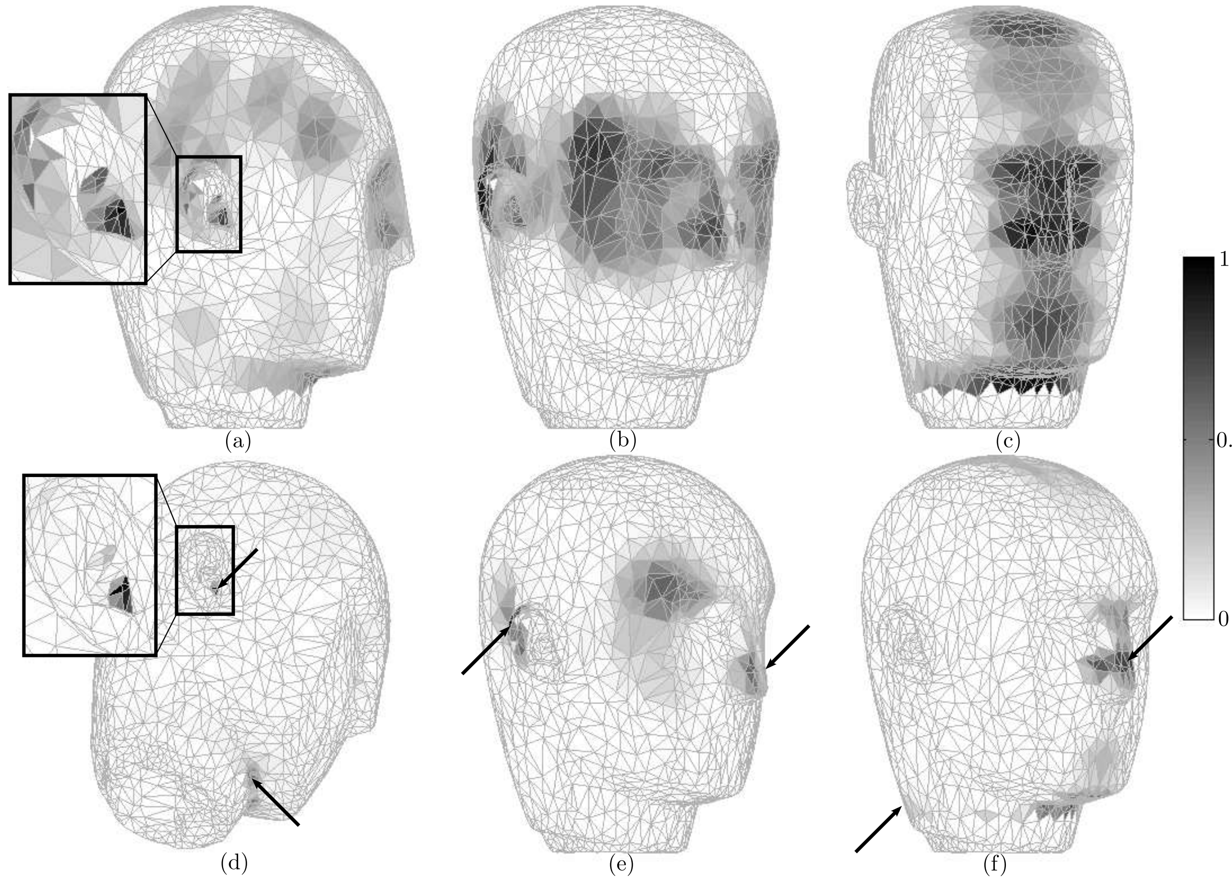

Fig. 2. Best positioning maps for one and two microphones. (a)-(c) - single microphone, color represents the effective rank of the HRTF matrix. (d)-(f) - pair of microphones, color represents the likelihood of being in an optimal pair. (a),(d) - all arrival directions, (b),(e) - horizontal plane, (c),(f) - median plane. In (d)-(f) an example of optimal pair is indicated by arrows.

number of repetitions, respectively. Similarly to the results presented for single microphone, figures 2e and 2f represent the results with the HRTF matrices at selected arrival directions.

When all arrival directions are considered (Fig. 2d), it can be seen that the most likely pairs will include sensors at the entrance to ear canals and under the chin. Note that the first few optimal pairs are not the left and the right ears. These pairs (indicated by arrows) consist of one element in the (either left or right) ear and an element positioned beneath the chin (or less likely, nose and forehead). This is probably due to the valuable information added because of the vertical displacement between these positions and their non symmetric location. The results of optimal positioning of a pair of microphones for selected source positions, either horizontal (Fig. 2e) or vertical (Fig. 2f), are similar to the results of single microphone positioning. The most likely microphone pairs are distributed along either vertical or horizontal strips surrounding the head, while the two microphones are maximally dislocated. These pairs include the (either left or right) outer ear and the opposite side of the nose, for the sources in the horizontal plane, and the upper part of the nape and the top of the nose, for the sources located in the median plane.

6. RELATION TO PRIOR WORK

In [12] an objective measure of sound localization was proposed based on the HRTFs and validated against the human sound localization theory. Here, this measure is adopted in order to quantify the quality of a microphone array configuration and a theoretical and numerical analysis related to this measure is presented.

7. CONCLUSION

A measure of microphone array configuration quality was proposed. It was analyzed theoretically showing that positioning of the microphones according to this measure can generally improve the robustness of beamforming algorithms and reduce the MUSIC DOA estimation variance. The results of one and two optimal microphone positioning on the surface of a simulated dummy head are presented and discussed validating the feasibility of the proposed measure. Future work will include improvement of the generalized HRTF database and development of the algorithms allowing fast solution of the sensor positioning problem.

8. ACKNOWLEDGMENTS

The authors would like to thank Brian F. G. Katz from LIMSI-CNRS, France, for kindly providing the model of the head geometry that was used in this study, and his valuable advices that helped to develop and validate the generalized HRTF database.

9. REFERENCES

[1] M. Akhtaruzzaman and A. Shafie, “Evolution of humanoid robot and contribution of various countries in advancing the research and development of the platform,” in *ICCAS 2010*, Oct. 2010, pp. 1021–1028.

[2] J. C. Middlebrooks and D. M. Green, “Sound localization by human listeners,” *Annu. Rev. Psychol.*, vol. 42, pp. 135–159, Feb. 1991.

[3] R. Y. Litovsky, “Spatial release from masking,” *Acoustics today*, Apr. 2012.

[4] R. L. Robart and L. D. Rosenblum, “Hearing space: Identifying rooms by reflected sound,” in *Studies in perception and action VIII*, 2005, pp. 153–156.

[5] P. Strumillo, Ed., *Advances in Sound Localization*. InTech, Apr. 2011.

[6] F. Keyrouz, W. Maier, and K. Diepold, “A novel humanoid binaural 3D sound localization and separation algorithm,” in *2006 6th IEEE-RAS International Conference on Humanoid Robots*, Dec. 2006, pp. 296–301.

[7] J. Hornstein, M. Lopes, J. Santos-Victor, and F. Lacerda, “Sound localization for humanoid robots - building audio-motor maps based on the HRTF,” in *IROS 2006*, Oct. 2006, pp. 1170–1176.

[8] M. Heckmann, H. Brandl, J. Schmuedderich, X. Domont, B. Bolder, I. Mikhailova, H. Janssen, M. Gienger, A. Bendig, T. Rodemann, M. Dunn, F. Joublin, and C. Goerick, “Teaching a humanoid robot: Headset-free speech interaction for audio-visual association learning,” in *RO-MAN 2009*, Oct. 2009, pp. 422–427.

[9] N. Schmitz, C. Spranger, and K. Berns, “3D audio perception system for humanoid robots,” in *ACHI 2009*, Feb. 2009, pp. 181–186.

[10] M. Maazaoui, K. Abed-Meraim, and Y. Grenier, “Adaptive blind source separation with HRTFs beamforming preprocessing,” in *SAM 2012*, June 2012, pp. 269–272.

[11] H. L. Van Trees, *Optimum Array Processing (Detection, Estimation and Modulation Theory, Part IV)*. New York: Wiley Interscience, 2002.

[12] V. Tourbabin and B. Rafaely, “Objective measure for sound localization based on head-related transfer functions.” *to appear at the 2012 IEEE 27-th Convention of Electrical and Electronics Engineers in Israel*, 2012.

[13] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed., ser. Johns Hopkins Studies in Mathematical Sciences. Johns Hopkins University Press, Oct. 1996.

[14] V. Tourbabin, M. Agmon, B. Rafaely, and J. Tabrikian, “Optimal real-weighted beamforming with application to linear and spherical arrays,” *IEEE Trans. Audio, Speech, Language Process.*, vol. 20, no. 9, pp. 2575–2585, Nov. 2012.

[15] R. Schmidt, “Multiple emitter location and signal parameter estimation,” *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Mar. 1986.

[16] P. Stoica and A. Nehorai, “MUSIC, maximum likelihood, and Cramer-Rao bound,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 5, pp. 720–741, May 1989.

[17] S. Chandler-Wilde and L. Steve, *Boundary Element Methods for Acoustics*, Department of Mathematics, University of Reading, Whiteknights PO Box 220, 2007.

[18] R. Greff and B. Katz, “Round robin comparison of HRTF simulation results: Preliminary results,” in *Proc. 123rd Convention of the Audio Engineering Society*, Oct. 2007, pp. 5–8.

[19] R. H. Hardin and N. J. A. Sloane, “Mclarens improved snub cube and other new spherical designs in three dimensions,” *Discrete and Computational Geometry*, vol. 15, pp. 429–441, 1996.

[20] B. F. G. Katz, “Acoustic absorption measurement of human hair and skin within the audible frequency range,” *J. Acoust. Soc. Am.*, vol. 108, no. 5, 2000.