HIGH ACUITY SOUND-SOURCE LOCALIZATION BY MEANS OF A TRIANGULAR SPHERICAL ARRAY

Amir A. Handzel

42 Hillcrest Circle
Watertown MA 02472, USA
handzel@isr.umd.edu

ABSTRACT

Sound source localization systems typically measure differences in time-of-arrival between pairs of microphones in free field arrays. Using a different concept, we previously designed and built a localization system that mimics nature's solution of harnessing wave diffraction about the head while relying only on two sensors positioned antipodally. One of its important advantages is the generation of intensity difference information, in addition to time/phase differences. This sensor configuration is limited, however, by its intrinsic axial symmetry. Here I depart from the constraint of two sensors while retaining the advantages of a diffracting "head", by introducing a symmetric array of three microphones placed equidistantly on a diffracting sphere. Detailed computations reveal that this design is capable of superior broadband localization of 1° resolution in the plane of the microphones. Experimental work is under way to verify these calculations.

1. INTRODUCTION

The localization of sound sources is useful for various purposes, and it can also aid the separation of signals from multiple sources and in their identification. Designed systems typically comprise free field sensor arrays for extraction of directional information. Most measure differences in time of arrival between combinations of pairs of microphones. Applications include the localization and tracking of speakers in conference rooms and improved hearing aids having directional sensitivity; see [1] for a comprehensive overview. Several groups installed free field microphone rigs on mobile robots also using time of arrival differences between microphone pairs [2–4]. A recent robotic localizing device was based on a pair of free field microphones which were rendered directional by means of reflectors placed around the microphones [5]. This augmented the usual time difference information with intensity difference.

In nature, directional acoustic sensing evolved to rely on diffraction about the head with only two sensors — the ears. The impinging sound waves are modified by the head in a frequency and direction dependent way. The inner ear decomposes the sound pressure signal into frequency bands. The brain then uses interaural differences in phase (IPD) and intensity level (ILD) in the various frequency bands to infer the location of a source [6, 7].

Inspired by human sound localization, we previously designed and built a localization system that mimics nature's solution of harnessing wave diffraction about the head while relying only on two sensors positioned antipodally [8–10]. One of its important advantages is the generation of intensity difference information, in addition to time/phase differences. This sensor configuration is limited, however, by its intrinsic axial symmetry: it allows localization only up to a circle around the interaural axis. In order to break this symmetry, we introduced dynamic localization [8] which enabled our robot mounted system to successfully distinguish between the front and back directions and to localize sources with accuracy of 2° [10].

We also conducted preliminary computational studies aimed at breaking the axial symmetry with a different approach, by positioning the two microphones in configurations that deviate from the antipodal. We found that such asymmetric positioning of two microphones at angular separation of about $100^\circ - 120^\circ$ provides unique horizontal localization [8]. It seemed, however, that localization precision would be non-uniform in the various directions.

Here I depart from the constraint of two sensors while retaining the advantages of a diffracting "head", by introducing a symmetric array of three microphones placed equidistantly on a diffracting sphere. In the present paper I consider planar, or approximately planar localization problems (as often relevant for mobile robots), so the source is in or close to the plane determined by the microphone and the center of the sphere. The arc distance between each pair of microphones is therefore 120°, which we previously found to be optimal.

Section 2 describes the acoustic calculations and the localization algorithm, followed by results and a discussion.

2. THE LOCALIZATION ALGORITHM

Consider a point source at $\mathbf{r_0} = (r_0, \theta, \phi)$, in spherical polar coordinates, which emits sound pressure flux of $F(\omega)e^{-i\omega t}$ in the angular frequency component ω. The resultant pressure on the surface of a diffracting sphere of radius r is obtained for each frequency component ω by solving the Helmholtz equation with Neumann boundary conditions [11, 12]:

$$p_\omega = \frac{i\rho c F(\omega) e^{-i\omega t}}{4\pi r^2} \sum_n (2n+1) P_n(\cos\theta) \frac{h_n(kr_0)}{h'_n(kr)} \quad (1)$$

where ρ is the mass density of air, P_n are the Legendre functions, h_n are the spherical Hänkel functions, and $k = \omega/c$ is the wave number (c being the speed of sound); primes denote derivatives. The nature of the solution is such, that it does not depend on the azimuthal angle, ϕ, i.e. the angle of elevation relative to the plane determined by the source, the center of the sphere and the microphone [8]. In addition, except for the near-field, i.e. sources very close to the sphere, the pressure on its surface is insensitive to the source distance r_0.

The measured sound pressure (1) is a complex response to the excitation in frequency ω:

$$p_\omega = A e^{i\alpha - i\omega t} \quad (2)$$

where A is the amplitude, and α is the part of the phase containing spatial information. For microphones j and k in the array, we define the Interaural Level Difference and Interaural Phase Difference:

$$\mathrm{ILD}_{jk} = \log A_k - \log A_j \qquad \mathrm{IPD}_{jk} = \alpha_k - \alpha_j, \quad (3)$$

which are both smooth functions of frequency. We consider the ILD-IPD plane as a basic feature space for localization. Every source direction and emission frequency induce an "active" point in the ILD-IPD plane of a microphone pair. Since ILD and IPD depend smoothly on frequency, a broadband sound source generates a whole curve $\sigma(\omega)$ in this plane. This curve is the source's specific *signature* which depends on its location [8].

The picture, so far, can be summarized as follows: a source at position $\mathbf{r_0}$ in the plane emits sound which is mapped through the scattering process, S, to a pair of sound pressure measurements, i.e. a pair of smooth complex functions of some frequency interval Ω. Extracting the interaural — i.e. relative — phase and intensity reduces them to a pair of Real functions (operation I):

$$\begin{array}{ccccc} \mathbb{R}^2 & \xrightarrow{S} & C_{\mathbb{C}}(\Omega) \times C_{\mathbb{C}}(\Omega) & \xrightarrow{I} & C_{\mathbb{R}}(\Omega) \times C_{\mathbb{R}}(\Omega) \\ \mathbf{r_0} & \overset{S}{\longmapsto} & (p_k, p_j) & \overset{I}{\longmapsto} & (\mathrm{ILD}_{jk}, \mathrm{IPD}_{jk}) \end{array} \quad (4)$$

The task is to prescribe a localization operator that by using all such pairwise measurements would, in effect, invert (4) to recover the source direction. We do so by defining the squared L^2 norm distance between the measured interaural functions (ILD_{jk}, IPD_{jk}) and the theoretical functions $(\mathrm{IPD}_{jk}(\theta), \mathrm{ILD}_{jk}(\theta))$. These theoretical signatures, which were previously calculated using the analytical solution of the full 3D acoustic scattering problem (1), are stored in a table. The IPD component of the metric for micrphone pair jk is:

$$\begin{aligned} \mathbf{D}^{\mathrm{IPD}}_{jk}(\theta) &\equiv \|\mathrm{IPD}_{jk}(\theta) - \mathrm{IPD}_{jk}\|_2^2 \\ &= \textstyle\sum_\omega \left(\mathrm{IPD}_{jk}(\theta, \omega) - \mathrm{IPD}_{jk}(\omega)\right)^2, \end{aligned} \quad (5)$$

and similarly for ILD. Marking the three microphones numerically from 1 to 3, ILD and IPD are thus calculated for all pairs of microphones with positive permutation, namely 2-1, 3-2, and 1-3, thereby accounting for all three possible pairs. ILD and IPD terms are then summed separately:

$$\mathbf{D}^{\mathrm{ILD}} = \sum_{\pi(jk)>0} \mathbf{D}^{\mathrm{ILD}}_{jk}, \qquad \mathbf{D}^{\mathrm{IPD}} = \sum_{\pi(jk)>0} \mathbf{D}^{\mathrm{IPD}}_{jk}, \quad (6)$$

to produce the overall ILD and IPD metric components. In order to avoid unbalanced contribution from the two, each is then normalized with respect to its maximal value (over θ). The two components are then combined to produce the total distance function:

$$\mathbf{D}^{\mathrm{Tot}}(\theta) = \mathbf{D}^{\mathrm{IPD}} + \mathbf{D}^{\mathrm{ILD}}. \quad (7)$$

Finally, the angle which obtains the minimum of the metric is assigned to the source direction:

$$\theta_0 = \mathrm{argmin}_\theta(\mathbf{D}^{\mathrm{Tot}}(\theta)). \quad (8)$$

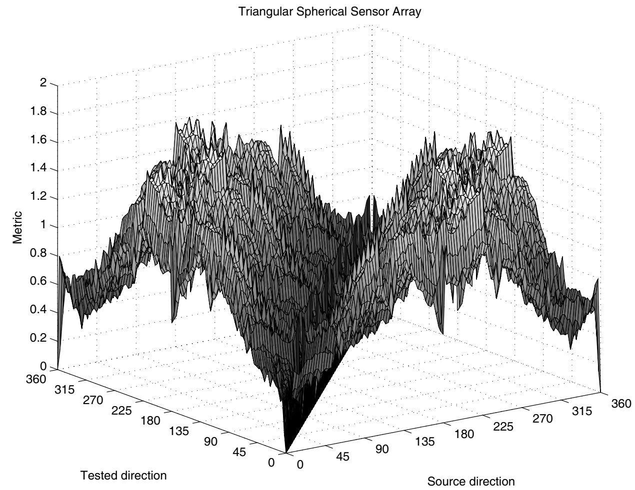

Fig. 1. Localization metric for sources at all possible directions for a triangular spherical array.

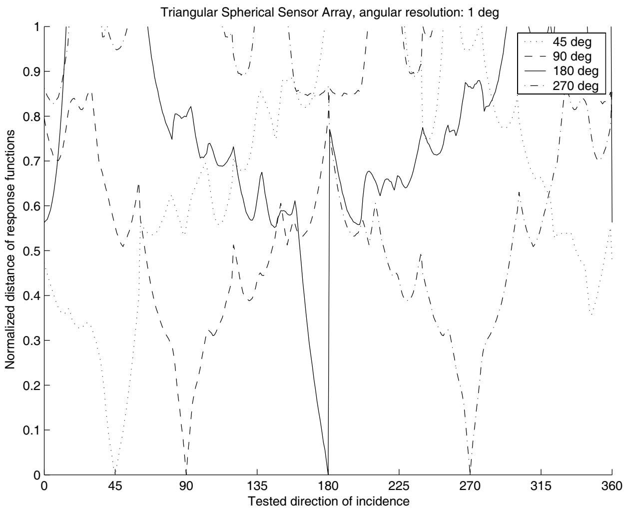

Fig. 2. Triangular spherical array: The metric provides unique localization of sources.

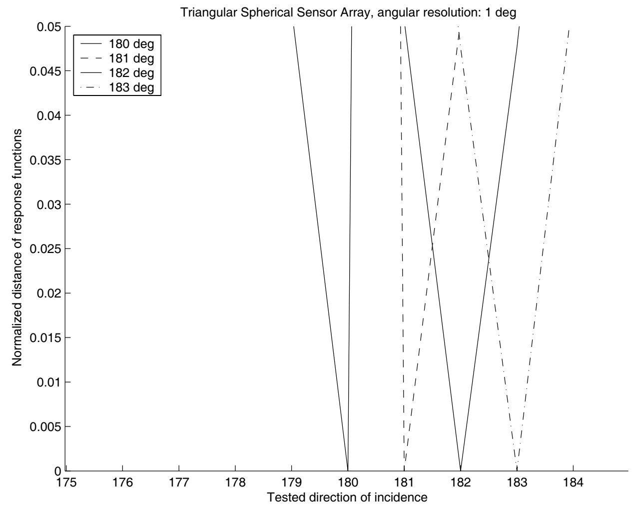

Fig. 3. Triangular spherical array: Localization metric distinguishes between sources separated by mere 1^o.

3. RESULTS

The above prescribed metric was calculated for all possible source directions at 1^o intervals, resulting in a 360×360 numerical matrix of distances between source directions and tested directions. This matrix is plotted in Figure 1. The "topography" of the metric has a broad valley along the diagonal, with a deep narrow gorge at its center exactly on the diagonal itself. Indeed, this is an ideal structure for such a function because for each possible source direction it would clearly mark the correct direction. Figure 2 shows the metric function for several source directions. For each of them, the function has a needle-like minimum at the correct value.

Although the numerical calculations were performed on a grid of $1^o \times 1^o$ granularity, this in itself does not guarantee 1^o localization acuity. Indeed, the system we built and tested previously using only one pair of antipodal microphones achieved slightly lower resolution of 2^o [10]. We therefore test here by computation whether the metric for the triangular array can distinguish between sources separated by mere 1^o. Figure 3 shows the metric function for several adjacent sources only 1^o apart. Even at this small separation the metric has sharp minima at the correct values, thus clearly distinguishing between them. The figure shows the metric around 180^o; it was tested, however, for all 360^o directions and found to be uniformly accurate throughout.

4. DISCUSSION

This paper builds on our previous work of designing sound localization systems which utilize a diffracting sphere to generate intensity information, in addition to time/phase information available in popular free field microphone arrays. Our previous work was inspired by nature's solution which relies only on two sensors. The present paper shows that by adding another sensor and arranging the array in a triangular symmetric configuration, the sharpness of localization is substantially enhanced and its acuity can increase to 1^o uniformly in the 2π circle of directions. One of the reasons such high resolution is possible with a small number of sensors is that this localization method is based on the physics of the problem and the diffraction process whereby the solution and algorithm are completely analytical.

As mentioned in the introduction, we previously found that placing two microphones on the diffracting sphere at 120^o arc distance from each other could give unique localization, but non-uniformly in directions. Compared with such single asymmetric pair, the new symmetric array of three microphones has three pair measurements of ILD and IPD. Their optimal localization directions cover different sectors of the circle. This mutual augmentation contributes to the high acuity and to its uniformity in direction. It can be contrasted with linear time-of-arrival (TOA) systems using digital signal processing. The discretization of the signal by sampling combined with the geometry of these arrays may result in non-uniform localization performance in the various directions.

Experimental work is under way to verify the computational results presented here.

5. REFERENCES

[1] Michael Brandstein and Darren Ward, Eds., *Microphone Arrays*, Digital Signal Processing. Springer-Verlag, 2002.

[2] Jie Huang, Tadawute Supaongprapa, Ikutaka Terakura, Fuming Wang, Noboru Ohnishi, and Noboru Sugie, “A model-based sound localization system and its application to robot navigation,” *Robotics and Autonomous Systems*, vol. 27, no. 4, pp. 199–209, June 1999.

[3] Juyang Weng and Kamen Y. Guentchev, “Three-dimensional sound localization from a compact non-coplanar array of microphones using tree-based learning,” *J. Acoust. Soc. Am.*, vol. 110, no. 1, pp. 310–323, July 2001.

[4] Stuart H. Young and Michael V. Scanlon, “Detection and localization with an acoustic array on a small robotic platform in urban environments,” Tech. Rep. ARL-TR-2575, Army Research Laboratory, Jan. 2003.

[5] Lorenzo Natale, Giorgio Metta, and Giulio Sandini, “Development of auditory-evoked reflexes: Visuo-acoustic cues integration in a binocular head,” *Robotics and Autonomous Systems*, vol. 39, pp. 87–106, 2002.

[6] Jens Blauert, *Spatial Hearing*, MIT Press, Cambridge, MA, 2nd edition, 1997.

[7] William M Hartmann, “How we localize sound,” *Physics Today*, pp. 24–29, Nov. 1999.

[8] Amir A. Handzel and P. S. Krishnaprasad, “Biomimetic sound-source localization,” *IEEE Sensors Journal*, vol. 2, no. 6, pp. 607–616, 2002.

[9] Amir A. Handzel, Sean B. Andersson, Martha Gebremichael, and P. S. Krishnaprasad, “A biomimetic aparatus for sound-source localization,” in *Intl. Conf. Decision and Control (CDC)*. 2003, vol. 6, pp. 5879–5884, IEEE.

[10] Sean B. Andersson, Amir A. Handzel, Vinay Shah, and P. S. Krishnaprasad, “Robot phonotaxis with dynamic sound-source localization,” in *Intl. Conf. Robotics and Automation (ICRA)*. 2004, vol. 5, pp. 4833–4838, IEEE.

[11] J.J. Bowman, Thomas B.A. Senior, and P.L.E. Uslenghi, Eds., *Electromagnetic and Acoustic Scattering by Simple Shapes*, Hemisphere Publishing Corporation, New York, 1987.

[12] Richard O. Duda and William L. Martens, “Range dependence of the response of a spherical head model,” *J. Acoust. Soc. Am.*, vol. 104, no. 5, pp. 3048–3058, Nov. 1998.