SPEAKER LOCATIONS FROM INTER-SPEAKER RANGE MEASUREMENTS: CLOSED-FORM ESTIMATOR AND PERFORMANCE RELATIVE TO THE CRAMÈR-RAO LOWER BOUND

Scott Wilson, Jeff Walters, Jonathan Abel

Stanford University
Center for Computer Research in Music and Acoustics
{rswilson, jhw, abel}@ccrma.stanford.edu

ABSTRACT

The problem of determining the relative positions of speakers in an array given noisy measurements of inter-speaker ranges is considered. A closed-form position estimator which minimizes a weighted equation error norm is presented. The information inequality is used to bound the position estimate mean square error and to gauge the accuracy of the closed-form estimator, which is shown to be nearly efficient.

1. INTRODUCTION

Many current techniques for multichannel rendering such as Ambisonics [1], VBAP [2], and wavefield synthesis [3] are benefited by or are dependent upon accurate speaker location information. There are a wide variety of methods available to propagate signals between array elements and estimate inter-element distances. In this paper we study the problem of determining the positions of multiple speakers using noisy measurements of inter-speaker ranges.

Estimating the position of an object from range measurements to a set of fixed positions—the problem of intersecting spheres—arises in navigation systems [4]. One estimation approach is to form an equation error relating the unknown positions and range measurements [4]. The equation error is selected such that its norm is easily minimized and it produces hypothesized ranges close to the measured values when its norm is small.

The case of uncertainty in the reference element positions has been studied for the related problem of range difference localization, with the result that the position estimate variance is increased as if the measurements were more noisy [5, 6, 7].

The problem at hand is perhaps more closely related to that of multidimensional scaling analysis (MDS) [8]. Measurements of distance between pairs of points in an unknown space are available and MDS seeks to make a map of the relative positions of the unknown points in the lowest dimension space which captures the topology implied by the input distances. Iterative methods are generally used.

Section 2 presents a closed-form speaker position estimator as the minimizer of a weighted equation error norm. In form, it is essentially a weighted MDS using an L2 norm.

In section 3 we study the accuracy with which the unknown speaker positions may be estimated, using the information inequality to gauge estimate mean square error. The information inequality, often referred to as the Cramèr-Rao Lower Bound (CRLB) states that the mean square error of any unbiased estimator is at least as large as the Fisher Information inverse [9, page 115, ff]. We argue that for any given speaker, its position bound variance is similar to that if all other speaker positions were known and the measurement noise were increased. Finally, simulation results show that the closed-form speaker position estimator has accuracy comparable to the bound.

2. POSITION ESTIMATION

Let $\mathbf{X}$ be the $N \times P$ matrix of speaker locations $\mathbf{x}_n$ where N is the number elements and P is the dimension of the space

$$\mathbf{X} \triangleq \begin{bmatrix} \mathbf{x}_0^\top \\ \vdots \\ \mathbf{x}_{N-1}^\top \end{bmatrix}. \tag{1}$$

The inter-element ranges r_{ij} are given by

$$r_{ij} = \| \mathbf{x}_i - \mathbf{x}_j \| . \tag{2}$$

We assume that the measured inter-element ranges ρ_{ij} are corrupted with additive, independent Gaussian noise with known variance

$$\rho_{ij} = r_{ij} + \epsilon_{ij} \sim \mathcal{N}(r_{ij}, \sigma_{ij}^2), \tag{3}$$

as would be the case using any number of accurate inter-element signal arrival time estimates, and in the presence of small estimation errors. (For instance, in the limit of large Fisher Information, the maximum likelihood estimator is known to be unbiased and normally distributed.) Under this assumption, the estimation problem becomes one of finding the parameters that determine the mean of a Gaussian-distributed measurement.

The maximum likelihood estimate is known to be efficient (unbiased with minimum variance) in the limit of small estimate errors. In the case of estimating the parameters determining the mean of a Gaussian distributed random variable, the maximum likelihood estimate minimizes the weighted sum of square hypothesized measurement errors—the differences between measured and hypothesized ranges. Unfortunately the measurements are non-linear in the parameters of interest, and the maximum likelihood estimate is difficult to compute directly.

Below, an equation error is developed such that it is linearly related to the measurement error when errors are small. The equation error has a weighted norm that can be minimized over all speaker position sets directly via a singular value decomposition of a matrix formed from the inter-speaker range measurements.

Stacking squared instances of (1), and taking $\mathbf{x}_0 = 0$, we have

$$2\mathbf{X}\mathbf{X}^\top = \boldsymbol{\zeta}\mathbf{1}^\top + \mathbf{1}\boldsymbol{\zeta}^\top - \mathbf{R} + \boldsymbol{\epsilon}_{EE}, \tag{4}$$

where $\boldsymbol{\zeta}$ is a column of the square range measurements from $\mathbf{x}_0$

$$\boldsymbol{\zeta} = \left[\rho_{0,1}^2 \cdots \rho_{0,N-1}^2\right]^\top, \tag{5}$$

$\mathbf{R}$ is a matrix of square range measurements excluding those from $\mathbf{x}_0$

$$\mathbf{R} = \begin{bmatrix} \rho_{1,1}^2 & \cdots & \rho_{1,N-1}^2 \\ \vdots & \ddots & \vdots \\ \rho_{N-1,1}^2 & \cdots & \rho_{N-1,N-1}^2 \end{bmatrix}, \tag{6}$$

and $\boldsymbol{\epsilon}_{EE}$ is an equation error. In the presence of separate range measurements ρ_{ij} and ρ_{ji} it is suggested that a symmetric $\mathbf{R}$ be formed using $[\rho_{ij}+\rho_{ji}]/2$ as the corresponding entries of $\mathbf{R}$.

We estimate the positions of the speaker elements $\hat{\mathbf{X}}$ to within an orthogonal transformation $\mathbf{Q}$ by constructing a matrix using the P largest singular values of the singular value decomposition of $\mathbf{X}\mathbf{X}^\top$. (It should be pointed out that with only measurements of inter-element range, there is no information to fix the orientation of the array.) This matrix is known to minimize the sum of square equation error elements over all position sets $\mathbf{X}$ in P dimensions [10, appendix F]. Defining by

$$\mathbf{U} \cdot \mathbf{D} \cdot \mathbf{V}^\top = \mathbf{X}\mathbf{X}^\top \tag{7}$$

the singular value decomposition of $\mathbf{X}\mathbf{X}^\top$, our estimate of position is given by

$$\hat{\mathbf{X}} = \mathbf{U} \cdot \mathbf{D}^{\frac{1}{2}} \cdot \mathbf{Q}. \tag{8}$$

3. INFORMATION INEQUALITY

The information inequality states that the variance of any unbiased estimator will be greater than or equal to the inverse of the Fisher Information, also known as the Cramèr-Rao Lower Bound (CRLB)

$$var\{\hat{\mathbf{X}}\} \geqq \mathbf{J}_{\hat{\mathbf{X}}}^{-1}. \tag{9}$$

The bound is useful in gauging the performance of the equation error minimizer, and in developing insight into the information contained in the range measurements.

The Fisher Information is given by

$$\mathbf{J}_{\hat{\mathbf{X}}} = \frac{\partial \boldsymbol{\mu}^\top}{\partial \hat{\mathbf{X}}} \cdot \boldsymbol{\Sigma}_R^{-1} \cdot \frac{\partial \boldsymbol{\mu}}{\partial \hat{\mathbf{X}}^\top}, \tag{10}$$

with $\boldsymbol{\mu}$ defined as the collection of all inter-element range measurements

$$\boldsymbol{\mu}^\top = [\cdots r_{ij} \cdots]. \tag{11}$$

Defining the unit vector pointing from element j to element i as $\boldsymbol{\beta}_{ij}$

$$\boldsymbol{\beta}_{ij} \triangleq \frac{\mathbf{x}_i - \mathbf{x}_j}{\|\mathbf{x}_i - \mathbf{x}_j\|}, \tag{12}$$

the sensitivity of the mean to changes in the ith element position is

$$\frac{\partial \boldsymbol{\mu}^\top}{\partial \mathbf{x}_i} = \left[0 \cdots 0 \, \boldsymbol{\beta}_{ij} \, 0 \cdots 0\right]. \tag{13}$$

When all but one, say $\mathbf{x}_0$, of the speaker positions are known, the Fisher Information is the outer product of direction vectors pointing from the unknown speaker to the other $N-2$ speakers, weighted by the respective inverse measurement variances:

$$\mathbf{J}_0 = \mathbf{B}^\top \boldsymbol{\Sigma}^{-1} \mathbf{B} = \sum_{i=1}^{N-1} \frac{1}{\sigma_i^2} \, \boldsymbol{\beta}_{0i} \boldsymbol{\beta}_{0i}^\top. \tag{14}$$

This interpretation is consistent with the notion that accurate estimates are possible along directions having a large number of elements or directions having small range estimate variances.

When there are two unknown speaker positions ($\mathbf{x}_0$ and $\mathbf{x}_1$), we write the Fisher Information relative to element $\mathbf{x}_0$, $\tilde{\mathbf{J}}_0$, as the sum of the information in contained in the ranges measured from $\mathbf{x}_0$ to the known speakers, and the information in the range measured between the unknown elements. Denoting by $\mathbf{J}_0$ and $\mathbf{J}_1$ the information contained in noisy range measurements to the remaining known speaker locations relative to elements $\mathbf{x}_0$ and $\mathbf{x}_1$, the desired Fisher Information $\tilde{\mathbf{J}}_0$ may be written as:

$$\tilde{\mathbf{J}}_0 = \mathbf{J}_0 + \frac{\boldsymbol{\beta}_{01}\boldsymbol{\beta}_{01}^\top}{\sigma_{01}^2 + \boldsymbol{\beta}_{01}^\top \mathbf{J}_1^{-1} \boldsymbol{\beta}_{01}}. \tag{15}$$

Note that the additional information provided by the range measurement between the two unknown positions takes on a form similar to that of a summand in (14) with an increased effective variance. This can be seen by comparing the term

$$\frac{1}{\sigma_i^2}$$

appearing in (14), with

$$\frac{1}{\sigma_{01}^2 + \boldsymbol{\beta}_{01}^\top \mathbf{J}_1^{-1} \boldsymbol{\beta}_{01}}$$

appearing in (15)—the difference being an increase in the estimate variance due to the unknown locations of both endpoints providing the range measurements.

The CRLB for speaker element $\mathbf{x}_0$ is

$$var\{\hat{\mathbf{x}}_0\} \geqq \mathbf{J}_0^{-1} - \frac{\mathbf{J}_0^{-1}\boldsymbol{\beta}_{01}\boldsymbol{\beta}_{01}^\top \mathbf{J}_0^{-1}}{\sigma_{01}^2 + \boldsymbol{\beta}_{01}^\top(\mathbf{J}_0^{-1} + \mathbf{J}_1^{-1})\boldsymbol{\beta}_{01}}. \quad (16)$$

When the range measurement to the unknown position is available the variance is decreased. The decrease is large when $\boldsymbol{\beta}_{01}$ aligns with a major axis of the the error ellipse implied by $\mathbf{J}_0$, i.e., when the new range measurement provides position information along a direction that was otherwise poorly measured.

4. SIMULATION

The speaker array geometries shown in Figures 1 and 2 were produced and simulated inter-element range measurements were created by adding Gaussian noise to the actual ranges. Using these simulated measurements, the closed form position estimator was applied to estimate positions, which were then compared to the actual speaker positions.

In addition, the theoretical Cramèr-Rao Lower Bounds were calculated and 90% confidence ellipses were plotted. The bounds were calculated using three scenarios for each speaker location: one assuming that all other speaker locations were known, one assuming that all other locations were unknown, and one assuming that all other locations were unknown, but the relative orientation of the speaker array was known. As expected, as seen in Figure 3, the bound is tightest for the case of all other speaker positions known. One can also observe that the information added by including knowledge of the angular orientation shrinks the confidence ellipse only slightly.

Simulation results not presented here show that as the number of elements increases, the bound for the case of all element positions being unknown approaches that of only a single position being unknown. This is consistent with the notion that with more elements come more range measurements and therefore more position information. Additionally, as the elliptical bounds become more flattened, the spread between the all unknown and one unknown cases is widened. As expected, arrays with elements in a line result in variances that are high in a direction perpendicular to the line of the array.

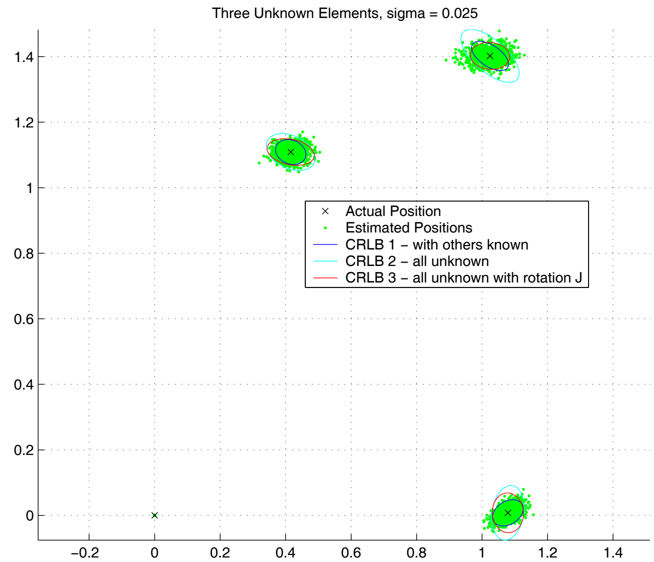

Fig. 1. Simulation result for three unknown locations.

Error ellipses based on both including and excluding array rotation information are very similar, except when the array has a small number of elements.

In 1000-trial Monte Carlo simulations the closed form estimator (8) was seen to be approximately unbiased with variance equal to 1.5 times the Cramèr-Rao Lower Bound.

5. CONCLUSION

A closed-form estimator of speaker element locations using inter-speaker range measurements was presented. Additionally, the Fisher Information for the problem was calculated and used to lend geometrical insight to the estimation problem. The estimator was applied to simulated noisy range measurements and was shown to exhibit a variance about 1.5 times the theoretical minimum.

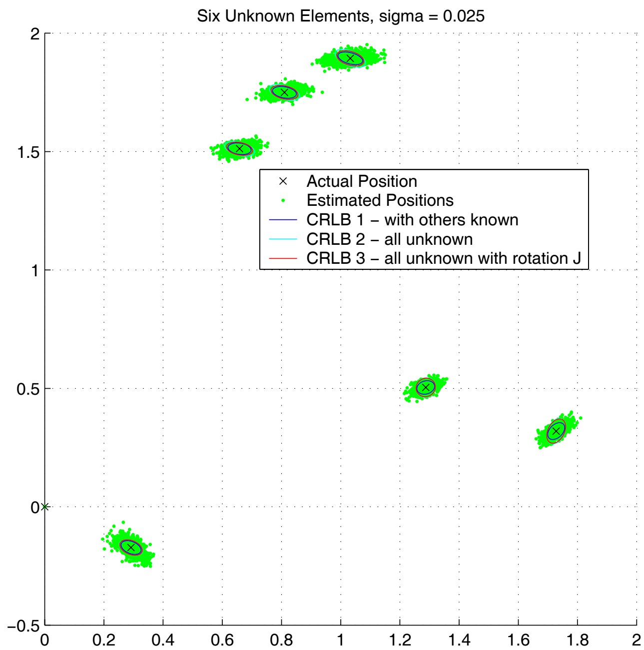

Fig. 2. Simulation result for six unknown locations.

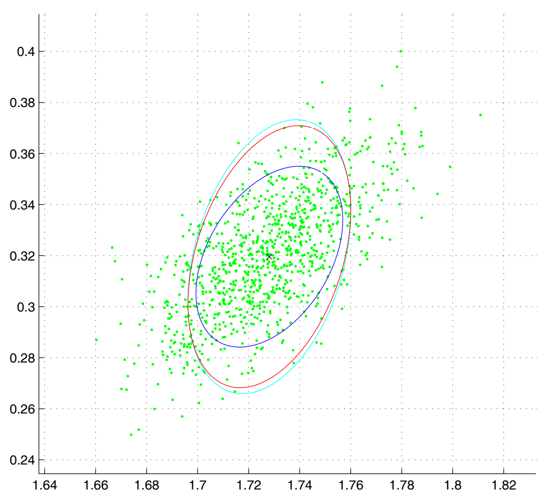

Fig. 3. Magnification of lower right element in Figure 2

6. REFERENCES

[1] M.A. Gerzon, "Ambisonics. part one. general system description," *Studio Sound*, vol. 17, no. 8, pp. 20–22, August 1975.

[2] V. Pukki, "Virtual source positioning using vector base amplitude panning," *Journal of the Audio Engineering Society*, vol. 45, no. 6, pp. 456–466, 1997.

[3] D. B. Ward and T. D. Abhayapala, "Reproduction of a plane-wave sound field using an array of loudspeakers," *IEEE Trans Speech and Audio Processing*, vol. 9, no. 6, pp. 697–708, September 2001.

[4] J. S. Abel and J. W. Chaffee, "Direct GPS solutions," in *Proc Institute of Navigation 49th Annual Meeting*, Cambridge, Mass., June 1993.

[5] P. M. Schultheiss and J. P. Ianniello, "Optimum range and bearing estimation with randomly perturbed arrays," *Journal Acoustical Soc of America*, vol. 68, no. 1, pp. 167–173, July 1980.

[6] Y. Rockah and P. M. Schultheiss, "Array shape calibration using sources in unknown locations—part 1: Far-field sources," *IEEE Trans Acoust Sig Proc*, vol. ASSP-35, no. 3, March 1987.

[7] Y. Rockah and P. M. Schultheiss, "Array shape calibration using sources in unknown locations—part 2: Near-field sources and estimator implementation," *IEEE Trans Acoust Sig Proc*, vol. ASSP-35, no. 6, June 1987.

[8] T. F. Cox and M. A. Cox, *Multidimensional Scaling, 2nd Edition*, Chapman-Hall, London, 2001.

[9] E. L. Lehmann, *Theory of point estimation*, John Wiley and Sons, New York, NY, 1983.

[10] J. O. Smith III, *Techniques for Digital Filter Design and System Identification with Application to the Violin*, Ph.D. thesis, Elec. Engineering Dept., Stanford University (CCRMA), June 1983.