Citation for the published Journal paper:

Title:

Calibration errors of uniform linear sensor arrays for DOA estimation: an analysis with SRP-PHAT

Author:

Mikael Swartling, Nedelko Grbic

Journal:

Signal Processing

Year:
2011
Vol.
91
Issue:
4
Pagination:
1071-1075
URL/DOI to the paper:
10.1016/j.sigpro.2010.09.018

Calibration Errors of Uniform Linear Sensor Arrays for DOA Estimation: an Analysis with SRP-PHAT

M. Swartling[a,*], N. Grbić[a]

[a] *Department of Electrical Engineering, Blekinge Institute of Technology, SE-371 79 Karlskrona, Sweden*

Abstract

This article presents an analysis of the sensitivity of geometrical sensor errors in acoustic source localization using the well-established SRP-PHAT method. The array in this analysis is a uniform linear array and the intended source is human speech in the far field. Two major results are presented: inner-sensor geometrical errors in the linear array produce smaller localization errors than corresponding geometrical errors do in the two end-point sensors, and the localization error rises sharply for a total geometrical error exceeding the equivalence of the acoustic propagation distance of 2/3 of the sample time instance (approximately 3 cm at 8 kHz). The article also provides a mathematical and graphical explanation of the results.

Keywords: direction of arrival estimation, calibration

1. Introduction

A geometrical calibration is often needed for sensor arrays, on top of the phase calibration of individual sensors, when used in source localization. The geometrical calibration is necessary in order to correctly map the modeled source position to a real physical location of the source. An important question here is to what degree is it necessary to calibrate the sensor array geometry for a sufficiently accurate mapping. This article presents an analysis of the time-delay

*Corresponding author, tel. +46 455 385581, fax +46 455 385057

Email addresses: msw@bth.se (M. Swartling), ngr@bth.se (N. Grbić)

errors caused by the geometrical errors when using the steered response power with phase transform (SRP-PHAT) method in a uniform linear sensor array used in far-field speech source localization.

A similar analysis has previously been undertaken in relation to the MUSIC method [1], which presents a sensitivity analysis of localization and sensor calibration errors. In other contexts, the problem with uncertain sensor arrays has been studied, for example the effect on beamformers [2, 3]. Various approaches to handle uncertain sensor arrays have also been proposed, for example self-calibrating and partly calibrated sensor arrays [4, 5].

This work is organized as follows: Section 2 presents the assumed signal model. Section 3 presents a model that approximates the SRP-PHAT method for small sensor displacements, and provides the theory and simulations for how calibration errors affect the time difference of arrival (TDOA) and direction of arrival (DOA) estimation. Section 4 explains why and under conditions, the SRP-PHAT method fails for large sensor displacements as indicated by the presented simulations.

2. Signal Model and TDOA Estimation

The propagation path is assumed to be anechoic, where the sensor signals are only subject to a time delay and an attenuation due to the propagation distance. For a small sensor array, the attenuation is assumed to be uniform over the array, and is modeled with a relative gain of unity:

$$x_m(t) = s(t - \tau_m) + \nu_m(t). \tag{1}$$

The source signal $s(t)$ is subject to a propagation delay τ_m from the source to the sensor m. The sensor noise $\nu_m(t)$ is modeled as independent white Gaussian noise. Since the source signal is not available, it is not possible to estimate the propagation delay. Instead, the TDOA for the sensors p and q, $\tau_{p,q}$, is estimated.

For a two-sensor array, the TDOA can be estimated with the generalized cross-correlation method (GCC) [6], defined in the frequency domain as

$$\hat{\tau}_{p,q} = \arg\max_{\tau} \int_{\omega} \Psi_{p,q}(\omega) G_{p,q}(\omega) e^{j\omega\tau} d\omega \tag{2}$$

where $\Psi(\omega)$ is a general weighting function or processor and the cross-power spectrum $G_{p,q}(\omega)$ is

$$G_{p,q}(\omega) = \mathrm{E}\left[X_p(\omega) X_q^*(\omega)\right] \tag{3}$$

where $\mathrm{E}[\cdot]$ is the statistical expectation operator. The steered response power method (SRP) [7] generalizes the GCC to more than two sensors. Using the phase transform (PHAT) processor, the SRP-PHAT method is defined for a uniform linear sensor array as

$$\hat{\tau} = \arg\max_{\tau} \sum_{\{p,q\}} \int_{\omega} \frac{G_{p,q}(\omega)}{|G_{p,q}(\omega)|} e^{j\omega\tau(q-p)} d\omega \tag{4}$$

which is the sum of the GCC of the sensor pairs $\{p, q\}$ in a set $\mathbf{P}$ of all sensor pairs of the array. The SRP is effectively a beamformer, and the sum in (4) is the response power of the beamformer whose steering parameter is τ.

3. Sensor Location Calibration

This article focuses on two variants of the TDOA and DOA estimates: the true and the estimated values. The true TDOA, τ, is the value obtained if the array is perfectly calibrated according to the assumed array geometry. The estimated TDOA, $\hat{\tau}$, is the actual value obtained from the array, and is potentially affected by erroneously located sensors, introducing an offset to the estimated value compared to the true value. The true and estimated DOA, α and $\hat{\alpha}$, are defined accordingly.

The TDOA for a source at an angle α is

$$\tau = \frac{d}{c} \cdot \sin(\alpha) \tag{5}$$

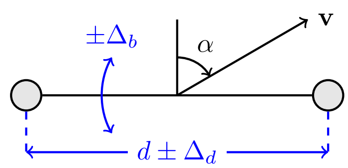
where d is the sensor separation and c is the propagation speed of sound. Solving for α, the DOA is calculated from the TDOA as

$$\alpha = \arcsin\left(\frac{c}{d} \cdot \tau\right). \tag{6}$$

The value of τ must be limited to ensure that the argument to the arcsin function is within the defined interval. Given a sensor distance d and a propagation speed c, it is known that $|\tau| \leq d/c$. This also correctly limits the argument to the interval of the arcsin function. In relation to the SRP, this is the visible region of the steered beamformer.

3.1. Two-sensor Array Displacements

The effects on the estimated TDOA and DOA in a two-sensor array can be trivially analyzed. An arbitrary displacement of the sensors is modeled by a changed sensor spacing and a rotation of the sensor array baseline that connects the two sensors. The sensor displacement is modeled as



where $\mathbf{v}$ is the DOA vector pointing towards the source, Δ_d is the change in the sensor spacing and Δ_b is the rotation of the sensor array baseline.

The estimated TDOA for a changed sensor spacing and the corresponding estimated DOA is

$$\hat{\tau} = \frac{d + \Delta_d}{c} \sin(\alpha) \tag{7}$$

$$\hat{\alpha} = \arcsin\left(\sin(\alpha) + \frac{\Delta_d}{d}\sin(\alpha)\right). \tag{8}$$

Similarly, the estimated TDOA for a rotated sensor array baseline and the corresponding estimated DOA is

$$\hat{\tau} = \frac{d}{c} \cdot \sin(\alpha + \Delta_b) \tag{9}$$

$$\hat{\alpha} = \alpha + \Delta_b. \tag{10}$$

Uncertainties in the sensor spacing and the orientation of the sensor array baseline thus directly affect the TDOA and DOA estimates. Calibration of a two-sensor array is necessary, and the accuracy of the calibration affects the accuracy of the TDOA and DOA estimates.

3.2. Multi-sensor Uniform Linear Array Approximation

The GCC integral in (2) is the inverse Fourier transform of $\Psi(\omega)G(\omega)$, where τ is the point at which the inverse Fourier transform is evaluated. Given the assumed signal model, the model of the cross-power spectrum becomes

$$G_{p,q}(\omega) = \mathrm{E}\left[|S(\omega)|^2\right] e^{-j\omega\tau_{p,q}}. \tag{11}$$

Combined with the PHAT processor to normalize the amplitude, the generalized cross-power spectrum becomes

$$\Psi_{p,q}(\omega)G_{p,q}(\omega) = e^{-j\omega\tau_{p,q}} \tag{12}$$

which is a unit amplitude linear phase signal whose inverse Fourier transform is

$$\mathcal{F}^{-1}\left\{e^{-j\omega\tau_{p,q}}\right\} = \mathrm{sinc}(\tau - \tau_{p,q}). \tag{13}$$

The sinc function is here approximated by a first order Taylor series expansion around the origin:

$$\begin{aligned} \mathrm{sinc}(\tau) &= \frac{\sin(\pi\tau)}{\pi\tau} && (14) \\ &= \sum_{n=0}^{\infty}(-1)^n \frac{(\pi\tau)^{2n}}{(2n+1)!} && (15) \\ &\approx 1 - \frac{\pi^2}{6}\tau^2. && (16) \end{aligned}$$

The SRP is the sum of the GCC for all sensor pairs, so the maximum of the SRP is located at the maximum of the sum of the GCC for all sensor pairs. The SRP-PHAT is thus approximated by the maximization of the sum of second degree polynomial functions, which has a trivial analytical solution.

Assume that the set of sensor pairs in an array of M sensors contains only pairs of adjacent sensors, such that

$$\mathbf{P} = \left\{\{m, m+1\} | m = 1 \ldots M-1\right\}. \tag{17}$$

The SRP-PHAT method is then approximated by

$$\hat{\tau} = \arg\max_{\tau} \sum_{m=1}^{M-1} \operatorname{sinc}(\tau - \hat{\tau}_{m,m+1}) \tag{18}$$

$$\approx \arg\max_{\tau} \sum_{m=1}^{M-1} 1 - \frac{\pi^2}{6}(\tau - \hat{\tau}_{m,m+1})^2 \tag{19}$$

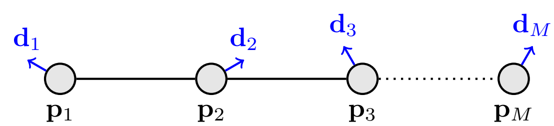
where $\hat{\tau}_{m,m+1}$ is the TDOA for the sensor pair $\{m, m+1\}$. The maximum of the first order Taylor series expansion approximation is solved analytically, and is

$$\hat{\tau} \approx \frac{1}{M-1} \cdot \sum_{m=1}^{M-1} \hat{\tau}_{m,m+1}. \tag{20}$$

Thus, the TDOA for the SRP-PHAT is the average of the TDOA for all individual sensor pairs.

3.3. Multi-sensor Uniform Linear Array Displacements

The sensor displacement of a uniform linear sensor array with three or more sensors is modeled as



where $\mathbf{p}_m$ is the position of the sensor m, and $\mathbf{d}_m$ is a random sensor displacement vector. The DOA vector $\mathbf{v}$ is a vector pointing from the sensor array towards the source and is defined as

$$\mathbf{v} = \begin{bmatrix} \sin(\alpha) & \cos(\alpha) \end{bmatrix}^{\mathrm{T}}. \tag{21}$$

Due to the far-field assumption, $\mathbf{p}_m$ can be projected onto the DOA vector (or, in general, arbitrarily displaced perpendicular to the DOA vector) without affecting the TDOA.

The TDOA for a sensor pair $\{m, m+1\}$ can be calculated from the difference in sensor position as projected onto the DOA vector:

$$\tau_{m,m+1} = \frac{(\mathbf{p}_{m+1} - \mathbf{p}_m)^{\mathrm{T}} \cdot \mathbf{v}}{c}. \tag{22}$$

A displaced position $\tilde{\mathbf{p}}_m$ of the sensor is

$$\tilde{\mathbf{p}}_m = \mathbf{p}_m + \mathbf{d}_m. \tag{23}$$

The estimated TDOA for the array with displaced sensors is, according to the approximation in (20), the average of the TDOA for all sensor pairs:

$$\hat{\tau} = \frac{1}{M-1} \sum_{m=1}^{M-1} \frac{(\tilde{\mathbf{p}}_{m+1} - \tilde{\mathbf{p}}_m)^{\mathrm{T}} \cdot \mathbf{v}}{c} \tag{24}$$

$$= \frac{1}{M-1} \sum_{m=1}^{M-1} \frac{(\mathbf{p}_{m+1} + \mathbf{d}_{m+1} - \mathbf{p}_m - \mathbf{d}_m)^{\mathrm{T}} \cdot \mathbf{v}}{c}. \tag{25}$$

By separating the terms with the true sensor positions and the displacements, the estimated TDOA for the true sensor positions reduces to the true TDOA:

$$\hat{\tau} = \tau + \frac{1}{M-1} \left(\sum_{m=1}^{M-1} \frac{\mathbf{d}_{m+1}^{\mathrm{T}} \cdot \mathbf{v}}{c} - \sum_{m=1}^{M-1} \frac{\mathbf{d}_m^{\mathrm{T}} \cdot \mathbf{v}}{c} \right). \tag{26}$$

If it is assumed that only inner sensors are displaced, so that $\mathbf{d}_1 = \mathbf{0}$ and $\mathbf{d}_M = \mathbf{0}$, then the summation ranges can be adjusted so that the terms cancel each other:

$$\hat{\tau} = \tau + \frac{1}{M-1} \left(\sum_{m=1}^{M-2} \frac{\mathbf{d}_{m+1}^{\mathrm{T}} \cdot \mathbf{v}}{c} - \sum_{m=2}^{M-1} \frac{\mathbf{d}_m^{\mathrm{T}} \cdot \mathbf{v}}{c} \right) \tag{27}$$

$$= \tau. \tag{28}$$

The average of the estimated TDOA of the displaced sensor pairs is reduced to the true TDOA, independently of the displacement vectors $\mathbf{d}_m$. Besides the errors of the approximation in (20), variations in inner sensor locations do not affect the estimated TDOA or DOA.

By displacing only the inner sensors in the sensor array, there are sensor pairs that counteract the TDOA offsets. Each inner sensor exists in two sensor pairs in $\mathbf{P}$, and each sensor pair counteracts the TDOA offset of the other. Because only one sensor pair has end-point sensors, there are no sensor pairs that counteract their TDOA offset when the end-point sensors are displaced. The effect of displacing an end-point sensor is analyzed by assuming that the end-point sensors are not displaced. Instead, a uniform change in the sensor spacing and a rotation of the sensor array baseline is assumed, as described in Section 3.1.

3.4. Simulations

The presented theory is evaluated using real speech signals in a simulated anechoic environment according to the assumed signal model given in Section 2. The inner sensor of a uniform linear sensor array with $d = 0.04$ m sensor spacing, consisting of three sensors, is displaced randomly around its true location. The simulated sensor signals are sampled at 8 kHz, and decomposed into the time-frequency domain for the cross-power spectrum estimation using a two-times oversampled uniform DFT filterbank [8] with 256 subbands.

Figure 1 shows the average DOA estimation error for a varying sensor location displacement distance and direction. The displacement direction is defined relative to the true source DOA vector:

$$\Delta_\alpha = \arccos\left(\frac{\mathbf{d}^{\mathrm{T}} \cdot \mathbf{v}}{|\mathbf{d}|}\right). \tag{29}$$

A displacement direction of 0° corresponds to a sensor displacement along the DOA vector, while a 90° displacement direction corresponds to a sensor displacement perpendicular to the DOA vector. The simulations also show that the DOA estimation error is symmetric with respect to positive and negative displacement angles and distances.

The DOA estimation error is small for most of the test cases, especially for displacements near perpendicular to the DOA vector. However, the error increases significantly for some cases when the sensor is significantly displaced along the DOA vector. These large estimation errors are geometrical errors introduced by the destructive addition of shifted sinc functions, as shown in Section 4.

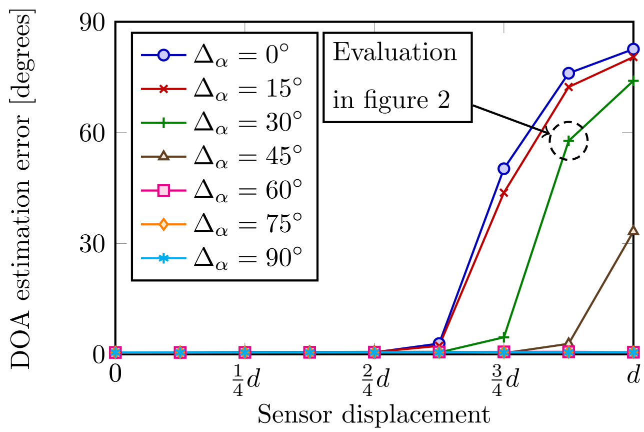


Figure 1: The inner sensor of an array is displaced. The graph shows the errors in the DOA estimation as the difference between the estimated DOA and the true DOA for a varying sensor displacement direction and distance.

4. The Geometrical Errors

4.1. Modeling the Error

Figure 2 shows the cause of the estimation error in Figure 1 for a case corresponding to the dashed circle. The solid lines are the sinc functions for the two sensor pairs and their peaks correspond to the pair's TDOA. The peaks are separated because the inner sensor is displaced, causing the different estimated TDOAs for the sensor pairs. The dashed line is the sum of the two delay functions and corresponds to the beamformer's output power. There are two peaks in the output power—although one of the peaks is outside the visible region of the beamformer, indicated by the vertical dashed lines—because the sinc functions are separated to such an extent that they no longer are able to constructively produce a single global maximum. Instead, the beamformer response becomes a local minimum at the correct source location.

The maximum allowed separation, which occurs before the global maximum turns into a local minimum, can be derived with the second derivative of the

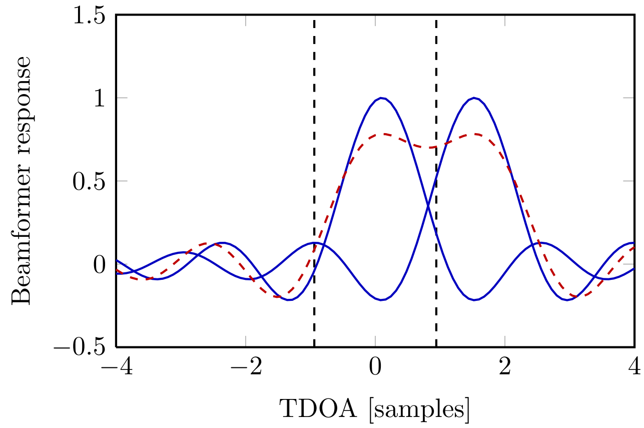


Figure 2: The geometrical effect causing the estimation error in figure 1 for a case corresponding to the dashed circle. The sinc functions no longer add constructively, causing erroneous global peaks in the beamformer response.

sum of the two sinc functions. If the two sinc functions are separated by τ, then

$$f(t, \tau) = \text{sinc}(t + \tau) + \text{sinc}(t - \tau). \tag{30}$$

By letting $t = 0$, and recognizing the even symmetry of the sinc function, the maximum separation τ_{max} is

$$\frac{\partial^2 f(t = 0, \tau)}{\partial \tau^2} = 0 \rightarrow \tau_{\text{max}} = \tau. \tag{31}$$

A numerical search reveals that the maximum separation is $\tau_{\text{max}} \approx 0.6626$ samples. As long as a sensor displacement results in a TDOA displacement less that τ_{max} samples, the sum of the two sinc functions maintains the global maximum given by the SRP-PHAT method.

4.2. Simulations

Figure 3 shows the average TDOA estimation error for a sensor array with displaced inner sensors over 10 000 simulations. The sensors are randomly displaced with a uniform distribution up to a distance corresponding to 1 sample at a sample rate of 8 kHz. The error is estimated for different numbers of sensors within the sensor array and for a varying maximum random displacement

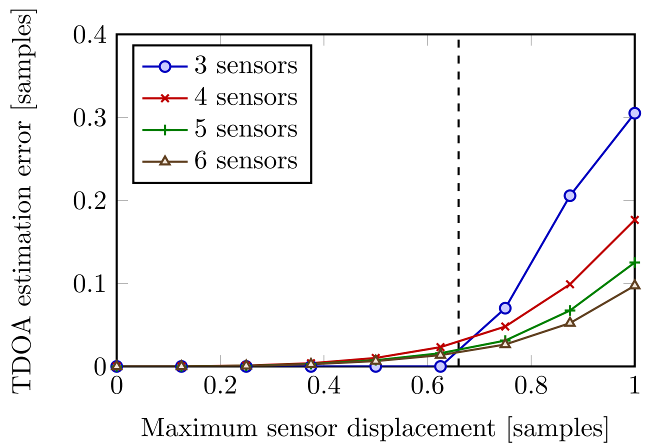


Figure 3: The TDOA estimation errors for multiple sensors. The inner sensors of the sensor array are randomly displaced, and the TDOA estimation error is the mean error of the estimated TDOA over many random displacements.

of the sensors. The simulations for the three-sensor array is consistent with the presented theory; the error is close to zero up to the vicinity of the maximum separation $\tau_{\max}$, indicated by the vertical dashed line. Beyond the limit, this algorithm effectively breaks down. Simulations for sensor arrays with more than three sensors are also consistent with the theory. The error is small for displacements where the second-degree polynomial sufficiently approximates the sinc function, and the averaging of sensor displacement errors decreases the TDOA error in larger arrays.

5. Conclusions

The article presents an analysis of the geometrical errors when performing TDOA and DOA estimations with the SRP-PHAT method, using a uniform linear sensor array and a far-field source. Both the theory and the simulations show that the SRP-PHAT method is robust in relation to small sensor placement errors. The DOA estimation fails when the TDOA estimates are significantly offset due to extreme sensor displacement. The derived maximum allowed sensor displacement $\tau_{\max}$ for a three-sensor array corresponds to a 2.82 cm location error at a sample rate of 8 kHz, or 0.47 cm at 48 kHz. The sensor location error should be related to the initially assumed 4 cm sensor spacing of the sensor array model. This becomes a significant sensor location error. To achieve sufficient DOA estimation accuracy, a rough calibration of the array may thus be sufficient.

References

[1] B. Friedlander, A sensitivity analysis of the MUSIC algorithm, IEEE Transactions on Acoustics, Speech, and Signal Processing 38 (10) (1990) 1740–1752.

[2] M. Wax, Y. Anu, Performance analysis of the minimum variance beamformer in the presence of steering vector errors, IEEE Transactions on Signal Processing 44 (4) (1996) 938–947.

[3] O. Besson, F. Vincent, Performance analysis of beamformers using generalized loading of the covariance matrix in the presence of random steering vector errors, IEEE Transactions on Signal Processing 53 (2) (2005) 452–459.

[4] M. Pesavento, A. B. Gershman, K. M. Wong, Direction finding in partly calibrated sensor arrays composed of multiple subarrays, IEEE Transactions on Signal Processing 50 (9) (2002) 2103–2115.

[5] C. M. S. See, A. B. Gershman, Direction-of-arrival estimation in partly calibrated subarray-based sensor arrays, IEEE Transactions on Signal Processing 53 (2) (2004) 329–338.

[6] C. H. Knapp, G. C. Carter, The generalized cross correlation method for estimation of time delay, IEEE Transactions on Acoustics, Speech, and Signal Processing ASSP-24 (4) (1976) 320–327.

[7] M. Brandstein, D. Ward (Eds.), Microphone Arrays: Signal Processing Techniques and Applications, Springer, 2001.

[8] P. P. Vaidyanathan, Multirate Systems and Filter Banks, Prentice Hall, 1993.