ACOUSTIC NODE CALIBRATION USING HELICOPTER SOUNDS AND MONTÉ CARLO MARKOV CHAIN METHODS

Volkan Cevher and James H. McClellan

Georgia Institute of Technology
Atlanta, GA 30332-0250
{gte460q, james.mcclella}@ece.gatech.edu

ABSTRACT

A Monté-Carlo method is used to calibrate a randomly placed sensor node using helicopter sounds. The calibration is based on using the GPS information from the helicopter and the estimated DOA's at the node. The related Cramér-Rao lower bound is derived and the effects of the GPS errors on the position estimates are derived. Issues related to the processing of the field data, e.g., time synchronization and data nonstationarity are discussed. The effects of the GPS errors are shown to be negligible under certain conditions. Finally, the results of the calibration on field data are given.

1. INTRODUCTION

Acoustic arrays with directionally sensitive or omnidirectional microphones can be used to localize and track targets using the direction-of-arrival (DOA) estimates from their sounds [1, 2]. If an acoustic node is defined as a collection of omnidirectional microphones whose relative positions are known with respect to each other, the node calibration problem refers to the determination of the unknown array center and its orientation. Note that this problem differs from the calibration of the individual microphone positions previously considered in the literature [3, 4]. In [3], accurate localization of the individual microphone positions is done by considering the effects of the calibration on the array manifold matrix, which is used in determining the target direction-of-arrival estimates (DOA's). On the other hand, as the DOA's are estimated, the node positions and orientations should be known to determine the target position (Fig. 1), which is the objective of this paper.

There has been some work in calibrating acoustic nodes using known calibration targets [5, 6]. The scenario considered in this paper is similar to one considered in [5]. A helicopter deploys the node and then transmits its GPS positions to the node. As the node receives these GPS estimates, it uses its DOA estimates of the helicopter to determine its position. It is assumed that the acoustic nodes themselves do not have GPS on them due to battery or jamming reasons (justifications can be found in [6].) The GPS position estimates are modelled as noisy and the effects of the GPS noise on the estimation performance of the node positions are also considered.

Node calibration is performed using a Monte-Carlo Markov Chain (MCMC) method. In this framework, a candidate node position is proposed and the Metropolis-Hastings scheme is used to
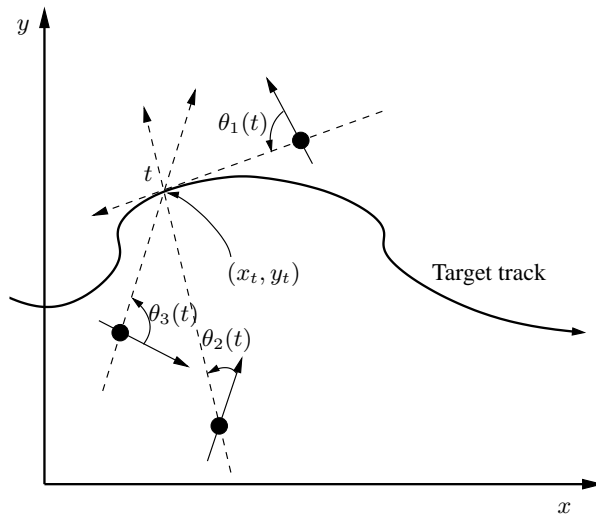
Fig. 1. Black dots are the acoustic nodes (arrays) and the solid arrows through the nodes represent their reference orientations for the local DOA estimates. The DOAs are measured counterclockwise. If the node positions are known, then it is possible to determine the target position (x_t, y_t).

sample a target distribution proposed in this paper. For this algorithm to obtain high resolution estimates, the GPS estimates and the acoustic measurements must be time-synchronized. Since the target is, in most cases, in the far-field of the sensor, the target sounds will not arrive instantaneously to the acoustic node. However, since the GPS estimates are transmitted electronically, acoustic data will be delayed because the speed of sound is relatively slow. This acoustic propagation delay, if not taken into account, will increase the estimation errors. A time-warping scheme is demonstrated that synchronizes the GPS estimates with the DOA estimates.

Assuming time-synchronization, we then show a Cramér-Rao lower bound on the unbiased estimators which use the DOA estimates for calibration and we also consider the effect of the GPS errors on the estimation. Another important issue is the array data non-stationarity problem caused by the rapid target movement which was considered in [1]. Rapid target movement causes the DOA estimates to be biased, which can lead to increased calibration errors. We demonstrate how to compensate for the target motion and get unbiased DOA estimates.

Organization of the paper is as follows. Section 2 formulates the problem and presents an MCMC solution. Section 3 derives the Cramér-Rao lower bound for the position estimation problem under some assumptions. Section 4 demonstrates that the GPS errors are negligible in most cases. Calibration results on field data are then given in Sec. 5.

2. THE CALIBRATION PROBLEM

We define ξ as the vector of unknown node position (x, y) and orientation φ in the 2D plane:

$$\xi = [x, y, \varphi]^T \tag{1}$$

and also define $\chi_t = [x_T(t), y_T(t)]^T$ as the known (noisy) calibration target track supplied by the GPS. The target bearing angle θ_t (measured c.c.w. with respect to the x-axis) and its range R_t at time t are given by

$$\theta_t(\xi, \chi_t) \triangleq -\varphi + \tan^{-1}\left(\frac{y - y_T(t)}{x - x_T(t)}\right) \tag{2}$$
$$R_t \triangleq ||\xi_{x,y} - \chi_t||$$

We will use the following classical narrow-band model for the acoustic microphone outputs:

$$\mathbf{y}_t = \mathbf{a}(\theta_t)s_t + \mathbf{n}_t \tag{3}$$

2.1. An MCMC Solution

The M-H scheme [7] is a general sampling algorithm and provides a basis from which other well-known sampling algorithms such as the Acceptance-Rejection (AR) and the Gibbs sampling can be derived as special cases [8, 9]. The algorithm assumes that it is possible to assign probabilities from the distribution to a given realization of the state vector, whose distribution is of interest. These probabilities need not be exact; they can also be given up to a proportionality constant. Note that calculating probabilities given a realization is different (and much easier) than generating the realizations directly from the, possibly intractable and multivariate, distribution itself.

The objective of the Metropolis-Hastings (MH) scheme [8] is to distribute the particles (discrete state samples ξ_i) according to a target distribution π. Hence, at each iteration k, the algorithm recursively redistributes its states around so that, asymptotically, the resulting Markov chain is distributed according to the target distribution. The new chain candidates γ are generated by the proposal function $q(\xi, \gamma)$, which is usually the spherically symmetric random walk:

$$q(\xi, \gamma) = q(|\xi - \gamma|) \propto \exp\left\{-\frac{(\xi - \gamma)^2}{2\sigma_q^2}\right\} \tag{4}$$

Once the new candidates are generated, the algorithm accepts the moves or keeps the current state according to the acceptance ratio $\alpha(\xi, \gamma)$ derived from the stochastic reversibility condition [8]:

$$\alpha(\xi, \gamma) = \min\left\{\frac{\pi(\gamma)q(\gamma, \xi)}{\pi(\xi)q(\xi, \gamma)}, 1\right\} \tag{5}$$

Note that the candidate generating (or proposal) function has a significant impact on the efficiency of the algorithm. It should be constructed so that the generated candidates display most of the structural dependence between the different dimensions. σ_q in (4) is defined as the jump size and is the other important variable affecting the algorithm speed. If it is too low, the algorithm takes a longer time to converge since the chain moves very slowly along the target distribution. On the other hand, if it is too high, the algorithm mostly rejects the new candidates and stays frozen. The current MCMC literature concentrates on these two important components of the algorithm for its efficiency [10].

For the calibration problem, the following target function can be used :

$$\pi(\xi) \propto \exp\left\{-\frac{1}{2\sigma_\theta^2}\sum_{t=1}^{L}\frac{1}{R_t}\left(\psi_t(\xi) + \frac{1}{L}\sum_{t=1}^{L}(\psi_t(\xi) - \theta_t) - \theta_t\right)^2\right\} \tag{6}$$

where $\psi_t(\xi)$ is the DOA calculated using the proposed vector ξ from the DOA definition in (2) and θ_t is the DOA estimate from the acoustic data using a beamformer. Theoretically, if a standard beamformer such as MVDR or MUSIC is used, the DOA estimates will be biased [1]. The following maximum-likelihood cost function based on the data model (3) can be used the estimate the DOA's with the assumption of constant velocity motion on the target:

$$J_{ML}(\theta) = \sum_{t=1}^{M}\text{tr}\left\{\left[\mathbf{I} - \frac{1}{P}\mathbf{a}(\theta_t)\mathbf{a}^H(\theta_t)\right]\hat{R}_y(t)\right\} \tag{7}$$

where

$$\theta_t = \tan^{-1}\left\{\frac{F_s\sin\theta + q(t-1)\cos\phi}{F_s\cos\theta + q(t-1)\sin\phi}\right\} \tag{8}$$

where M is number of batch samples, $\hat{R}_y(t)$ is one sample autocorrelation estimate, $q = v/R$ (v is helicopter speed estimated using the GPS), and ϕ is the approximate target orientation for the estimation batch using the proposed node position. The following MVDR based beamformer can also be used:

$$J_{MVDR}(\theta) = \sum_{i=1}^{L}\frac{1}{\mathbf{a}^H(\theta_i)\hat{R}_i^{-1}\mathbf{a}(\theta_i)} \tag{9}$$

where the estimation batch M is partitioned into L data points (e.g., $L = 256$ where $M = 2048$) and the autocorrelation estimates are taken every M/L samples and θ_i are the corresponding DOA estimates using (8).

2.2. Time-Synchronization Issue

Incorrect time-synchronization can severely degrade the performance of a high resolution calibration method. The problem is caused by the faster arrival of the target GPS estimates than the arrival of the sounds by the target at the respective GPS positions. In the M-H scheme, this issue can easily be incorporated into the solution using the proposed position ξ, the GPS estimates χ_t, and the speed of sound c.

Denote t_{gps} and t_n as the GPS and the node time frames, respectively. Then, the following time-warping needs to be applied to the GPS points:

$$t_{gps\rightarrow n} = t_{gps} - ||\xi^{(k)} - \chi_{t_{gps}}||/c \tag{10}$$

This implies that the GPS points given on a regular time frame will correspond to a time-warped irregular time frame on the node ref-

erence and needs to be interpolated on the node time frame where the DOA's are calculated.

3. THE CRAMÉR-RAO LOWER BOUND

Under the *i.i.d* Gaussian assumption on the array noise, the probability density function (pdf) of the observed data is given by

$$p(\mathbf{Y}_W|\xi, \chi_0, \dots, \chi_t) = \prod_{t=0}^{W-1} \frac{1}{\pi^P \sigma^{2P}} \exp\left[-\frac{1}{\sigma^2}|\mathbf{y}(t) - \mathbf{a}_t s(t)|^2\right] \quad (11)$$

where σ^2 is the array noise $\mathbf{n_a}(t)$ variance, W is the total number of observations with the sampling frequency F_s, P is the number of microphones in the node, $\mathbf{a}_t \triangleq \mathbf{a}(\theta_t(\xi, \chi_t))$, and $\mathbf{Y}_M$ is the aggregate data vector formed by stacking all the observed data as follows

$$\mathbf{Y}_W = \begin{bmatrix} \mathbf{y}(t) \\ \mathbf{y}(t+\tau) \\ \vdots \\ \mathbf{y}(t+(W-1)\tau) \end{bmatrix}, \quad \text{where } \tau = \frac{1}{F_s}. \quad (12)$$

Equation (11) models the data likelihood and leads to the ML solution. First, the negative log-likelihood function is obtained

$$L^- \doteq WP\log(\pi\sigma^2) + \frac{1}{\sigma^2}\sum_{t=0}^{W-1}||\mathbf{y}(t) - \mathbf{a}_t s(t)||^2 \quad (13)$$

where $\doteq$ denotes equality up to a constant. The ML estimates can then be obtained by maximizing the log-likelihood function, which is equivalent to minimizing L^-. Fixing ξ and $s(t)$ and minimizing L^- with respect to σ^2, the ML noise variance can be estimated. Then, the signal estimate is given by substituting the noise estimate into L^-:

$$\sigma^2_{ML} = \frac{1}{WP}\sum_{t=0}^{W-1}||\mathbf{y}(t) - \mathbf{a}_t s(t)||^2 \quad (14)$$

$$s_{ML}(t) = \frac{1}{P}\mathbf{a}_t^H \mathbf{y}(t)$$

The Cramér-Rao lower bound (CRLB) is an information theoretical inequality, which provides a lower bound for the variances of the unbiased estimators. If an estimator achieves the CRLB, then it is possible to prove that it is also a solution of the likelihood equation. However, it is not always true that the ML solution will achieve the CRLB (at least, for finite sample sizes) or that it will be unbiased [11]. The CRLB is still a useful metric to compare the performance of the algorithm, and is derived for the calibration problem in this section.

We now derive the relation for the Fisher information matrix (FIM). Assume that the noise variance σ^2 is known. The likelihood function (11) for the parameter vector ξ simplifies to the following relation:

$$L(\xi) \doteq -\frac{1}{\sigma^2}\sum_{t=0}^{W-1}||\mathbf{y}(t) - \mathbf{a}_t s(t)||^2 \quad (15)$$

where $\mathbf{a}_t$ is as defined earlier. The $(i,j)^{th}$ element of the FIM is given by derivatives of the (15) by the ith and j^{th} parameter of the vector ξ

$$F_{i,j} = E\left\{\frac{\partial^2 L_\chi(\xi)}{\partial\xi_i\partial\xi_j}\right\} = -\frac{2}{\sigma^2}\sum_t \text{Re}\left\{\left(\frac{\partial\mathbf{a}_t}{\partial\xi_i}\right)^H \frac{\partial\mathbf{a}_t}{\partial\xi_j}\right\} \quad (16)$$

where $\xi_i = [\xi]_i$. The derivation of a similar problem is given by [3]. The required derivatives $\frac{\partial\mathbf{a}_t}{\partial\xi_i}$ are straightforward and are omitted. The Cramér-Rao lower bound is the inverse of this expression [11].

4. EFFECTS OF THE GPS ERRORS ON THE ESTIMATION PERFORMANCE

The GPS measurements are usually not supplied at the same rate as the acoustic data is sampled. It is assumed that the GPS measurements are given every T seconds (i.e., $T = 1s$), which corresponds to a much slower sampling rate than the acoustic array output sampling rate of F_s (i.e., $F_s = 1024Hz$). For simplicity, we consider the 2D problem where the GPS outputs χ_t have only two components in the x and y directions. We model the GPS noise as zero mean *i.i.d.* Gaussian: $\mathbf{n}_\chi \sim \mathcal{N}(\mathbf{0}, \sigma_\chi^2\mathbf{I})$. The target is assumed to have constant velocity between GPS measurements. This is a valid approximation usually used in practice [1, 2, 12].

Using the Taylor series, the first order effect of the GPS noise on the auxiliary variable θ_t defined in (2) can be modelled as follows:

$$\theta_t(\xi, \chi_t + \mathbf{n}_\chi(t)) \approx$$

$$\theta_t(\xi, \chi_t) + \frac{\partial\theta_t}{\partial n_\chi^x(t)}\bigg|_{n_\chi^x(t)=0} n_\chi^x(t) + \frac{\partial\theta_t}{\partial n_\chi^y(t)}\bigg|_{n_\chi^y(t)=0} n_\chi^y(t) \quad (17)$$

By taking the necessary derivatives in (17) and noting that the noise $\mathbf{n}_\chi(t)$ is independent in the x and y directions, it is easy to show that the DOA θ_t is approximately Gaussian:

$$\theta_t(\xi, \chi_t + \mathbf{n}_\chi(t)) \sim \mathcal{N}\left(\theta_t(\xi, \chi_t), \frac{1}{R_t^2}\sigma_\chi^2\right) \quad (18)$$

This is a very intuitive result: since the GPS errors are much smaller than the target range (i.e., $R_t \gg \sigma_\chi$ because the acoustic node is in the far-field), the position errors will translate into an approximate angle error of $\tan^{-1}(\sigma_\chi/R_t) \approx \sigma_\chi/R_t$.

To derive the effect of the GPS noise on the steering vector, a first order approximation can again be used:

$$\mathbf{a}(\xi, \chi_t + \mathbf{n}_\chi(t)) \approx \mathbf{a}(\xi, \chi_t) + \frac{\partial\mathbf{a}}{\partial\theta}n_\theta \quad (19)$$

As an example, we explicitly derive the effect for a narrow-band calibration target, which has constant center frequency f_0. For the mathematical brevity of the expressions, assume that i^{th} microphone position is given in the polar coordinates: (ρ_j, ϕ_j). With these assumptions, the array steering vector can be written as

$$[\mathbf{a}(\xi, \chi_t)]_j = \exp\left[j\frac{2\pi f_0\rho_j}{c}\sin(\theta_t(\xi, \chi_t) + \phi_j)\right] \quad (20)$$

where c is the speed of sound and $[\mathbf{a}]_j$ indicates the jth element of the vector $\mathbf{a}$. By taking the derivative of (20), it can be shown $\frac{\partial\mathbf{a}}{\partial\theta} = \frac{j2\pi f_0}{c}\mathbf{a}(\xi, \chi_t)\lambda(\theta)$ with $\lambda(\theta) = \text{diag}\{[\rho_1\cos(\theta +$

$\phi_1), \ldots, \rho_P \cos(\theta + \phi_P)]\}$. Also define $\mathbf{\Lambda}(\theta) = \lambda(\theta)\lambda(\theta)$. Then, the array outputs for a target signal with constant envelope magnitude 1 can be shown to obey the following Gaussian distribution

$$\mathbf{y}(t) \sim \mathcal{N}\Big(\mathbf{a}(\xi, \chi_t), \mathbf{\Sigma}(\xi, \chi_t)\Big) \qquad (21)$$

where the autocorrelation matrix $\mathbf{\Sigma}$ is a function of the array noise as well as the GPS noise:

$$\mathbf{\Sigma}(\xi, \chi_t) = \sigma^2 \mathbf{I} + \Big(\frac{2\pi f_0 \sigma_\chi}{cR_t}\Big)^2 \mathbf{a}(\theta_t)\mathbf{\Lambda}(\theta_t)\mathbf{a}^H(\theta_t) \qquad (22)$$

Note that the last term in (22) is the perturbation due to the GPS errors. If $\sigma \gg \frac{2\pi f_0 \sigma_\chi}{cR_t}$ then it can be argued that GPS has very small effect on the estimation performance because the data likelihood (11) is not effected. This is a very reasonable assumption in most cases since the target narrow-band frequencies are usually much less than $100Hz$ and the GPS error standard deviation σ_χ is on the order of a few meters.

5. FIELD DATA RESULTS

We apply the Monté-Carlo calibration scheme on field data from a small acoustic array. A helicopter flies sorties around the acoustic node for the calibration purposes (Fig. 2). For the experiment, the array was hand-emplaced so that the true true location and orientation would be known. The acoustic node has six omnidirectional microphones placed uniformly on a circle with a radius of 1.219m, which also corresponds to the inter-microphone distance. The spatial aliasing frequency is around 135Hz corresponding to the half-wavelength (equal to the radius of the array). For the MVDR beamforming results, we tracked the ten highest peaks using the short time Fourier transform (Fig. 3) with their respective heights, and averaged the estimated DOA's accordingly. For the time-synchronization, third-order b-splines were used to interpolate the irregular time-grid for the algorithm's proposed positions for each particle.

The GPS track of the helicopter used for the calibration is shown in Fig. 2. Figure 4 demonstrates the results of the Metropolis-Hastings scheme with the Mode-Hungry modification [13]. A listing of the pseudo-code for the M-H scheme is given. To establish a baseline, we found the CRLB using (16) to approximate the lower bounds for the variances for the parameter vector as $[\sigma_x, \sigma_y, \sigma_\theta] = [1.3931m, 1.3597m, 0.13296°]$.

Note that the approximate bound assumes the single sinusoid array model for the observations. This assumption is not satisfied for the helicopter signal. It is also possible to show that the same lower bound can also be derived using a Gaussian assumption on the DOA estimates [14]. In Fig. 4, the DOAs estimated with the field data between $t = 50s$ and $t = 120s$ violates the Gaussian assumption since they have the incorrect mean. Without post-processing of the data, it is difficult to detect this incorrect mean since the orientation estimates can also bias the DOA distribution. The algorithm still managed to do well because the target distribution (6) weights the calculated DOAs according to their estimated range. When the target is far away from the node, the estimates are expected to get worse. Hence, the algorithm puts less importance on the DOA estimates corresponding to large target ranges. Lastly, Fig. 4 also demonstrates that the calibration results are significantly worse without the time-synchronization step.

Pseudo Code for the Metroplis-Hastings Calibration

At time k, for each particle i $(i = 1, \ldots, N)$, $\xi_i^{(k)}$:

i. Generate a candidate $\gamma_\mathbf{i}$ using $q(\xi_i, \gamma_i)$, which represents a spherical random walk.

ii. Estimate the time-reference frame for data synchronization using the proposed position, the target GPS track, and the speed of sound c.

iii. Calculate the DOAs, $\theta(t)$, using a beamformer such as (7).

iv. Calculate the acceptance ratio, where the target distribution $\pi(.)$ is as given in (6)

$$\alpha(\xi_i, \gamma_i) = \min\left(\frac{\pi(\gamma_i)}{\pi(\xi_i)}, 1\right)$$

v. Sample $u \sim \mathcal{U}(0, 1)$

vi. If $u \le \alpha(\xi_i, \gamma_i)$, set $\xi_i^{(k+1)} = \gamma_i$, else, $\xi_i^{(k+1)} = \xi_i^{(k)}$. ∎
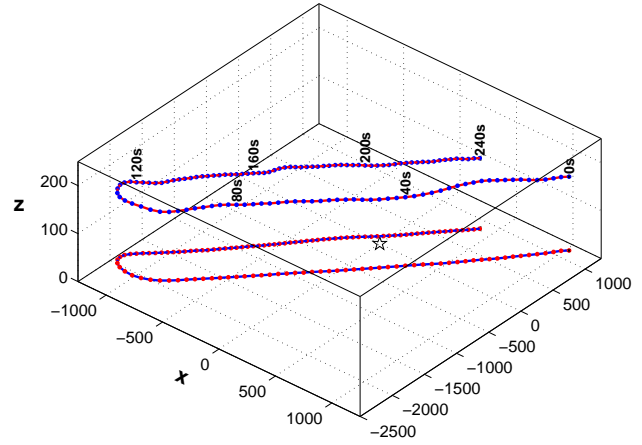

Fig. 2. The acoustic node (star) is situated at the origin. The calibration helicopter completes two sorties around the node for the calibration, corresponding to a four-minute run. The actual helicopter track as well as its projection on the x-y plane are shown.

Fig. 3. The helicopter spectrum displays strong harmonic lines. The ten highest peaks in the time-frequency plane are picked using the magnitude of the Fourier transform once per second. These frequencies as well as their time-frequency amplitudes are used in determining the DOA estimates.

Fig. 4. Top figure shows the GPS track coming from the helicopter (dashed line), the time-warped GPS track using (10), and the MVDR beamformer estimates of the field data. The bottom left plot is the resulting MHMH algorithm distribution with estimate $\xi = [\ 6.43,\ -6.52\ ]$m using time synchronization, whereas, at the bottom right, the distribution with estimate $\xi = [\ -31.66,\ 43.77\ ]$m is the MHMH result without time synchronization. The estimated orientation for both cases is $1°$. The true node location is shown with the diamond.

6. CONCLUSIONS

An MCMC calibration method is demonstrated to solve the problem of node calibration using acoustic sources. The solution takes into account the practical issues of time-synchronization, non-Gaussianity, and array data non-stationarity typically encountered when the calibration source is a rapidly moving target such as a helicopter. Lastly, the GPS errors are shown to be negligible relative to the microphone sensor noise in terms of overall calibration performance.

7. REFERENCES

[1] Y. Zhou, P.C. Yip, and H. Leung, "Tracking the direction-of-arrival of multiple moving targets by passive arrays: Algorithm," *IEEE Transactions on Signal Processing*, vol. 47, no. 10, pp. 2655–2666, October 1999.

[2] V. Cevher and J. H. McClellan, "General direction-of-arrival tracking with acoustic nodes," To appear in *IEEE Transactions on Signal Processing*.

[3] B.C. Ng and C.M.S. See, "Sensor array calibration using a maximum-likelihood approach," *IEEE Transactions on Antennas and Propagation*, vol. 44, pp. 827–835, June 1996.

[4] B.C. Ng and A. Nehorai, "Active array sensor localization," in *ICASSP 1993*, 27-30 April 1993, vol. 4, pp. 21–24.

[5] V. Cevher and J. H. McClellan, "Sensor array calibration via tracking with the extended Kalman filter," in *Proc. of the Fifth Ann. Fed. Lab. Symp. on Adv. Sensors*, College Park, MD, 20-22 March 2001, pp. 51–56.

[6] R.L. Moses, D. Krishnamurthy, and R. Patterson, "An auto-calibration method for unattended ground sensors," in *ICASSP 2002*, Orlando, FL, May 2002, vol. 3, pp. 2941–2944.

[7] W.K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, vol. 57, pp. 97–109, 1970.

[8] S. Chib and E. Greenberg, "Understanding the Metropolis-Hastings algorithm," *The American Statistician*, vol. 49, no. 4, pp. 327–335, 1995.

[9] L. Tierney, "Markov chains for exploring posterior distributions," *The Annals of Statistics*, vol. 22, no. 4, pp. 1701–1728, 1994.

[10] A. Gelman, G.O. Roberts, and W.R. Gilks, "Efficient Metropolis jumping rules," *Bayesian Statistics*, vol. 5, 1996.

[11] H.V. Poor, *An Introduction to Signal Detection and Estimation*, Springer-Verlag, 1994.

[12] M. Orton and W. Fitzgerald, "A Bayesian approach to tracking multiple targets using sensor arrays and particle filters," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 216–223, February 2002.

[13] V. Cevher and J. H. McClellan, "Fast initialization of particle filters using a modified Metropolis-Hastings algorithm: Mode-Hungry approach," in *ICASSP 2004*, Montreal, CA, 17–22 May 2004.

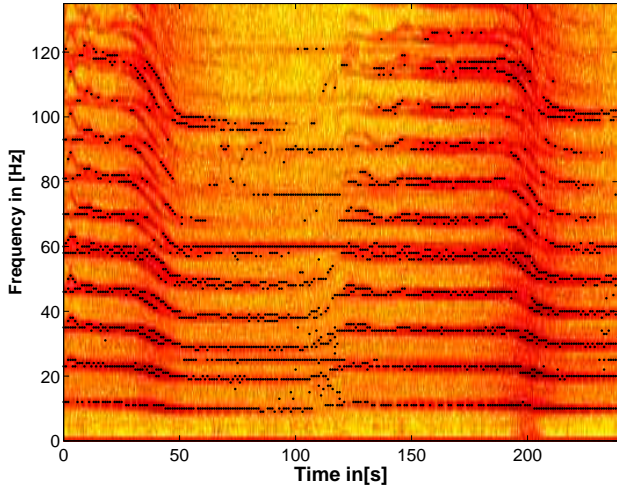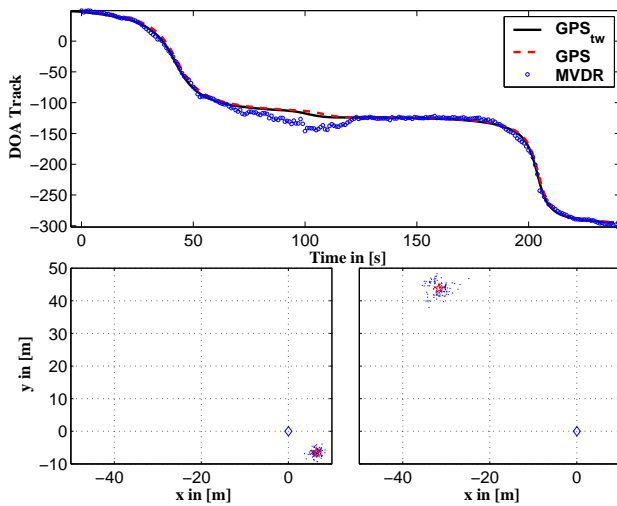[14] V. Cevher and J. H. McClellan, "Acoustic node calibration using moving sources," in preparation for *IEEE Aerospace and Electronic Systems*.