AN ACOUSTIC MULTIPLE TARGET TRACKER

Volkan Cevher and James H. McClellan

School of Electrical & Computer Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0250

ABSTRACT

We propose a particle filter acoustic tracker to track *multiple* maneuvering targets using a state space formulation with a locally linear motion model. The observations are a *batch* of direction-of-arrival (DOA) estimates at various frequencies. The data likelihood incorporates the possibility of missing data as well as spurious DOA observations. By imposing smoothness constraints on the target motion, the particle filter is able to avoid data association problems. To make the filter computationally efficient, a proposal strategy based on approximating the full posterior with Newton's method is employed. Computer simulations show the algorithm's performance.

1. INTRODUCTION

Tracking the bearing angles of multiple maneuvering targets using acoustic arrays is usually formulated in terms of state-space models where target DOAs are related to the acoustic microphone outputs through an observation equation. The state update (which may include variables other than the DOA) is based on a locally linear motion model [1–3]. The performance of tracking algorithms relies heavily on how accurately these models represent the observed natural phenomena.

One common model is the far-field narrow-band observation model where the array response is related to a source with a constant narrow-band frequency at DOA, θ, through a steering vector, assuming an isotropic and non-dispersive medium [4]. This well-understood model leads to many practical DOA estimation techniques, e.g., MUSIC, MVDR, etc. However, the estimation performance of algorithms using this model deteriorates when the target signals are wide-band, or when rapid target motion spreads the array spatial spectrum.

The presence of multiple targets also increases the tracking complexity because data association must be implemented to sort the received data for each target. The association problem could be handled in several ways: (i) probabilistic data association methods estimate the states by summing over all the association hypotheses weighted by the probabilities obtained by the likelihood [5], (ii) smoothness assumptions on the target (motion) states allows a natural ordering of the data [6], (iii) computationally costly ML/EM methods use the likelihood functions to search for a global maximum, or (iv) nearest neighbor methods provide easy heuristics to perform measurement updates. Most of these methods use the mean and covariance approximation on the sufficient statistics for the state, which may be estimated with a Kalman filter. However, for nonlinear state-spaces with general noise assumptions, Monté-Carlo methods should be used to adequately capture the dynamic, possibly multi-modal, statistics.

A particle filter is a natural solution for state-space problems where the observations arrive in sequence. The state probability density function is represented by discrete state samples (particles) distributed according to the underlying distribution either directly or by proper weighting [7]. Convergence results guarantee that the particles can approximate any statistics of the distribution with arbitrarily fine accuracy by increasing the number of particles. In the particle filtering framework, the data association problem is attacked implicitly by the interaction of all the particles through the state-space model. However, to increase the efficiency of the algorithm, various simplifications have been proposed, such as the partitioning approach [2], or other Bayesian approaches [8].

In this paper, a particle filter tracker is presented for multiple target DOA tracking using a constant velocity motion model. Instead of directly using the signals as explained by the narrow-band model, the algorithm uses sufficient statistics for the state vector: a batch of DOA estimates $\mathbf{y}_{t,f} = \{y_{t+m\tau,f}(j_m)\}_{m=0}^{M-1}$ computed by a beamformer appropriate for the local frequency

Prepared through collaborative participation in the Advanced Sensors Consortium sponsored by the U. S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-02-0008.

characteristics of the target signals. The frequency dependence of the batch DOAs is typical for the acoustic sources. The number of DOA estimates j_m at time $t + m\tau$ is variable because changing frequency content might give rise to spurious DOAs and missed DOAs. This entire batch of DOAs will be used to perform the state update at regular time intervals of $T = M\tau$.

The DOA observations are assumed normally distributed around the true DOA tracks with constant missing data probability and clutter rate. To increase the efficiency of the algorithm, the filter uses an approximation of the full posterior to generate particles and a robust Newton search method is used to approximate the data-likelihood. Data association is handled automatically in the particle filter by imposing smoothness constraints on the target motion. Lastly, all DOA observations within the batch are assumed locally stationary, i.e., angle spread due to target motion during the interval τ is assumed small.

Section 2 gives a brief overview of the motion model and provides the needed data-likelihood expressions. The particle filter details are covered in Sect. 3 where a constrained Newton method is introduced to approximate the filter proposal function. Finally, simulation results are given in Sect. 4.

2. STATE-SPACE FORMULATION

The filter state consists of the DOA $\theta_k(t)$, heading direction $\phi_k(t)$, and (logarithm of) velocity over range $Q_k(t) = \log v_k/r_k(t)$ for each target k where the total number of targets K is assumed known. The particle filter state vector $\mathbf{x}_k(t)$ consists of the concatenation of the individual motion vectors

$$x_k(t) \triangleq [\ \theta_k(t),\ Q_k(t),\ \phi_k(t)\ ]^T$$

as $\mathbf{x}_t = \left[\ x_1^T(t),\ \cdots\ x_K^T(t)\ \right]^T$. The parameters $\theta_k(t)$ and $\phi_k(t)$ are measured clockwise w.r.t. the x-axis. The state update equation, derived from the geometry imposed by the assumed constant velocity motion, is nonlinear [1, 3]:

$$x_k(t+T) = h_T\left(x_k(t)\right) + u_k(t) \tag{1}$$

where $u_k(t) \sim \mathcal{N}\left(0, \Sigma_u\right)$ with $\Sigma_u = \operatorname{diag}\{\sigma^2_{\theta,k}, \sigma^2_{q,k}, \sigma^2_{\phi,k}\}$ and $h_T(x_k(t)) =$

$$\begin{bmatrix} \tan^{-1}\left\{\dfrac{\sin\theta_k(t) + T\exp Q_k(t)\sin\phi_k(t)}{\cos\theta_k(t) + T\exp Q_k(t)\cos\phi_k(t)}\right\} \\ Q_k(t) - \frac{1}{2}\log\left\{\hat{Q}_k(t)\right\} \\ \phi_k(t) \end{bmatrix} \tag{2}$$

where $\hat{Q}_k(t) = 1 + 2Te^{Q_k(t)}\cos(\theta_k(t) - \phi_k(t)) + T^2e^{2Q_k(t)}$.

The observations $\mathbf{y}_{t,f} = \{y_{t+m\tau,f}(j_m)\}_{m=0}^{M-1}$ consist of all the DOA estimates out of a beamformer at each batch index m, where the acoustic data of length T is segmented into M segments of duration τ. The DOAs, indexed by j_m, need not be ordered or associated with the previous index $m-1$ and the number of peaks to retain can even be time-dependent.

The choice of the parameter τ for beamforming is determined by various physical constraints: (i) target frequency range, (ii) target speed, and (iii) target maneuvers. For reasonable beamforming, at least two cycles of the narrow-band target signals should be observed. This stipulates that $\tau > 2/F_{\min}$, where $F_{\min}$ is the minimum beamforming frequency for the target. Moreover, to keep the worst case beamforming bias[1] bounded for each DOA by an angle threshold denoted by θ_{B}, we approximately have $\tau < 2\theta_{\mathrm{B}}e^{-Q}$. It is necessary to have at least three DOA estimates to determine the state vector, but we use $M > 5$ to decrease the probability that the state is not observable due to missing DOAs, and to improve the robustness of the tracker. Putting these together, we get

$$\frac{2}{F_{\min}} < \tau < \min\left\{\frac{2\theta_{\mathrm{B}}}{\exp\{Q\}}, \frac{T}{M}\right\}. \tag{3}$$

Lastly, the target motion should satisfy the constant velocity assumption during the batch period T. For slow-moving ground vehicles, $T = 1$s is a reasonable choice.

Figure 1 shows the observation model. It is assumed that the batch of DOAs, $\mathbf{y}_{t,f}$, form a normally distributed cloud around the true target DOA tracks with a constant missing data probability[2] κ; and may have spurious peaks Poisson distributed with rate λ. The variance of the DOAs, σ^2, is assumed constant and may be estimated using the DOAs from the previous estimation period in conjunction with the ML estimation techniques for the specific beamformer used. For example, for an ML estimator using the narrow-band model, there is a formula that calculates the additive array noise variance [1]. The array noise then can be related to an expected DOA noise through the formulas in [10].

This specific observation model is very similar to the one used in active contour image tracking problems [9]. It can be shown that the data likelihood given the

[1] The bias is calculated by taking the angular average of the target track, so it also depends on the heading direction. The worst case bias happens when the target heading and DOA sum up to π.

[2] This model assumes that only one DOA at each f belongs to the target or the target is missed. If the probability of the true measurement being in the observed data is equal for each j_m, a constant data miss probability κ may be assumed [9].

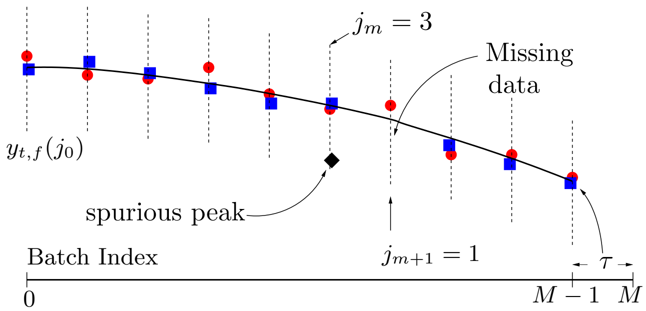

Fig. 1. The circles and squares are the DOA estimates at different frequencies calculated using the acoustic data received during an interval of duration τ. Given the observations $\mathbf{y}_{t,f}$, the objective of the tracker is to determine the state $x_k(t)$ which completely parameterizes the solid curve.

state under the assumptions described above can be written as $p(\mathbf{y}_t|\mathbf{x}_t) \propto$

$$\prod_f \prod_k \prod_m \left\{ 1 + \frac{1}{\sqrt{2\pi}\kappa\lambda} \sum_{j_m} e^{-(h^\theta_{m\tau}(\mathbf{x}_t) - y_{t+m\tau,f}(j_m))^2/2\sigma^2} \right\} \tag{4}$$

where $\mathbf{y}_t = \{\mathbf{y}_{t,f}\}_{f=1}^F$ denotes the cumulative DOA data calculated in time interval $[t, t+T)$ and h^θ refers to the DOA component generated by the state update. Equation (4) also assumes that DOAs for different frequencies f are independent. If there happens to be a known internal correlation structure, a joint density can be formulated to replace (4).

3. PARTICLE FILTER DETAILS

The efficiency of the particle filtering algorithm depends on the proposal functions that determine the random support of the particles to be *properly* weighted for estimation. In this paper, a proposal function, denoted as $g(\mathbf{x}_t|\mathbf{y}_t, \mathbf{x}_{t-1})$, is derived to approximate the target posterior density directly:

$$g(\mathbf{x}_t|\mathbf{y}_t, \mathbf{x}_{t-1}) \approx p(\mathbf{x}_t|\mathbf{y}_t, \mathbf{x}_{t-1}) \propto p(\mathbf{y}_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{x}_{t-1}) \tag{5}$$

where $p(\mathbf{x}_t|\mathbf{x}_{t-1}) \sim \mathcal{N}(f_T(\mathbf{x}_{t-1}), \Sigma_u)$ and $p(\mathbf{y}_t|\mathbf{x}_t)$ is given by (4). Moreover, the proportionality in (5) is independent of the current state $\mathbf{x}_t$. This approximation, in effect, moves the particle stream towards high probability regions of the posterior so that more particles survive the final resampling step, producing better future states as the system evolves [7].

The posterior density should be approximated such that the resulting proposal function is as close to the posterior as possible and, at the same time, is easy to sample from. Hence, various Gaussian approximations to the full posteriors are commonly used [11]. In our case, we first approximate the data-likelihood by a Gaussian so that the proposal function will also be Gaussian. Hence, a mean μ_y and a covariance Σ_y are first determined from the observed data $\mathbf{y}_t$. Then, by using the Gaussian parameters of the state update, an analytical relation for the proposal function is given.

The mode of $p(\mathbf{y}_t|\mathbf{x}_t)$, denoted as $\mathbf{x}_M$, is a good candidate for the parameter μ_y. Then, the Hessian H of (4) at the mode can also be used as the covariance estimate: $\Sigma_y = H^{-1}$. To calculate these parameters, a Newton search algorithm can be used on the negative log-likelihood of the data [1]. However, the resulting analytical relations have numerical sensitivity issues, so we propose to use the following alternative cost function to determine the mode $\mathbf{x}_M$:

$$J = -\sum_f \sum_k \sum_m \sum_{j_m} e^{-(h^\theta_{m\tau}(\mathbf{x}_t) - y_{t+m\tau,f}(j_m))^2/2\sigma^2} + \tfrac{1}{2}(\mathbf{x}_t - \mathbf{x}_0)^T \Sigma^{-1} (\mathbf{x}_t - \mathbf{x}_0) \tag{6}$$

This cost function consists of two terms: the first has the same minima as the negative log-likelihood function of the data distribution; and the second is a regularization term forcing the solution $\mathbf{x}_M$ to lie close to some vector $\mathbf{x}_0$ w.r.t. the Σ-weighted distance measure.

Note that the cost function without the regularization term only depends on angle differences. The gradients in that case may lead to physically infeasible (motion) changes in the parameters $Q(t)$ and $\phi(t)$ to account for fractional angle errors while determining $\mathbf{x}_M$. The regularization prevents this issue by constraining the solution space to lie in the Σ-neighborhood of $\mathbf{x}_0$, and, at the same time, imposing smoothness on the target motion. The parameter Σ also bounds the covariance of the data-likelihood approximation.

Nominally, the mode should be within the particle cloud coming from the previous iteration after being propagated through the state update. Hence, an easy way to determine the mode would be to choose the particle best explaining the current data set (denoted as $\mathbf{x}_0$ in (6)). Unfortunately, when the targets maneuver, $\mathbf{x}_0$ may fall outside actual data observations. This necessitates a correction accomplished by a Newton search algorithm to determine the actual mode $\mathbf{x}_M$.

If we define $G = \partial J/\partial \mathbf{x}$ and $\hat{H} = \partial^2 J/\partial \mathbf{x} \partial \mathbf{x}^T$, the Newton recursion is given by the familiar iteration:

$$\mathbf{x}_M^{l+1} = \mathbf{x}_M^l - \mu_l \hat{H}_l^{-1} G_l$$

with its starting value at $\mathbf{x}_M^l = \mathbf{x}_0$. The step size μ_l should be decreased adaptively making sure that the cost function is always decreasing. Although time-consuming, it is straightforward to derive analytical expressions for G and $\hat{H}$ (similar calculations can be found in [1].)

Even with the available analytical relations, the calculation of the Hessian still poses problems. If the Hessian of (6) is directly calculated from the exact formulas, the resulting expression for $\hat{H}$ is not guaranteed to be positive definite and modifications are necessary to make the Newton correction $\hat{H}^{-1}G$ effective at each iteration[3]. Hence, while calculating the final expression of the Hessian, the terms including second-order derivatives are neglected in the analytical formula. In this case, the Hessian is a function of the outer product of the gradient, and it is possible to prove that it is positive definite.

After the Gaussian approximation to the data likelihood described above (note that $\Sigma_y = H^{-1} \approx \hat{H}^{-1}$), the final expression for the proposal function to be used in the particle filter is given by

$$g(\mathbf{x}_t|\mathbf{y}_t, \mathbf{x}_{t-1}) \sim \mathcal{N}\left(\mu_g, \Sigma_g\right), \quad \text{where} \tag{7}$$

$$\begin{aligned} \Sigma_g &= \left(\Sigma_y^{-1} + \Sigma_u^{-1}\right)^{-1}, \quad \text{and} \\ \mu_g &= \Sigma_g \left(\Sigma_y^{-1}\mathbf{x}_M + \Sigma_u^{-1} h_T(\mathbf{x}_{t-1})\right). \end{aligned} \tag{8}$$

Then, the particle filter incremental weights are given by

$$u^{(i)} = \frac{p(\mathbf{y}_t|\mathbf{x}_t^{(i)})p(\mathbf{x}_t^{(i)}|\mathbf{x}_{t-1}^{(i)})}{g(\mathbf{x}_t^{(i)}|\mathbf{y}_t, \mathbf{x}_{t-1}^{(i)})} \tag{9}$$

The implementation of the filter used in the simulations also has a resampling (with replacement) stage after estimation.

4. SIMULATIONS

Three simulations show the performance of the algorithm using data generated according to the assumed models. To initialize the trackers, we used a particle cloud with the correct mean, and a covariance of $\Sigma_0 = \text{diag}\{(2°)^2,\ 0.1^2,\ (4°)^2\}$. For the Newton algorithm that approximates the data-likelihood for the particle proposal, the initial step size was 0.1, but it is decreased adaptively until 1000 iterations are reached. In all cases, we used $\Sigma = \sqrt{2}\Sigma_u$.

Simulation parameters are summarized in Table 1. Figure 2 demonstrates a single target tracking scenario. The observed DOAs are Gaussian distributed around the true DOA track with variance $(2°)^2$. The filter does a good job of catching the target as it maneuvers because the model is given a large heading process noise ($\sigma_{u,\phi} = 10°$). Figure 3 shows a much more difficult scenario for the tracker where two independent layers

[3]The same issue applies to the negative log-likelihood of (4).

Table 1. Simulation Parameters

Fig.	N	T	M	F	σ	$\sigma_{u,\theta}$	$\sigma_{u,Q}$	$\sigma_{u,\phi}$
2, 4	100	1	10, 20	1	2°	1°	0.05	10°
3	100	1	10	2	3°	1°	0.05	10°

of DOA estimates are given each with the correct mean and a variance of $(3°)^2$. There is a small bias in the filter DOA estimates between $t = 4$s and $t = 6$s where the targets are crossing. This bias seems to be a result of the target maneuvers, which start at $t = 4$s. The filter maintains the track coherence in this difficult case using the independent frequency observations (without frequency information, the filter can confuse the targets). Although the estimates deteriorate in the region where targets are crossing as well as maneuvering, the filter locks back on the targets after $t = 6$s.

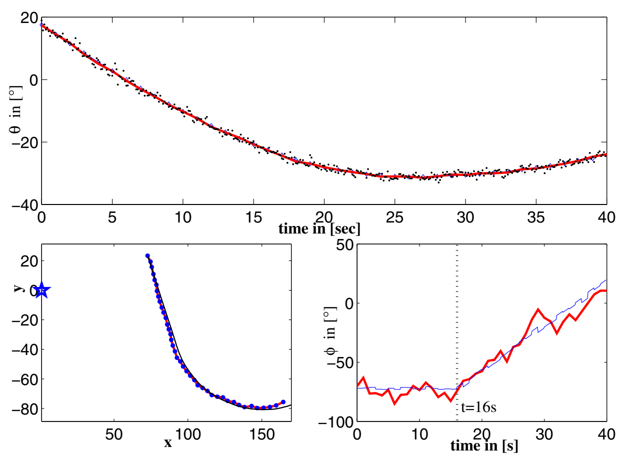

Fig. 2. *Top*: Black dots are the DOA observations generated by adding Gaussian noise to the true target DOA track. The filter estimate is hard to distinguish from the true DOA track. *Lower Left*: Ground truth vs. filter estimate. The sensor array is shown with the star. The filter track estimate is constructed by using the filter motion outputs with the correct initial target position. *Lower Right*: The filter heading estimates vs. the true target heading.

For the last example (Fig. 4), acoustic data sampled at $F_s = 1000$Hz is generated for two narrow-band targets, one at $f_1 = 40$Hz and the other at $f_2 = 80$Hz, using the narrow-band observation model in [4]. The filter state is also augmented to include a frequency variable: $x_k(t) \triangleq [\ \theta_k(t),\ Q_k(t),\ \phi_k(t),\ f_k(t)\ ]^T$ [3]. Then, Gaussian noise is added to the array data where the noise standard deviation is equal to one-tenth of the sinusoid amplitudes. The microphone array used for the simulation has 15 microphones situated uniformly on

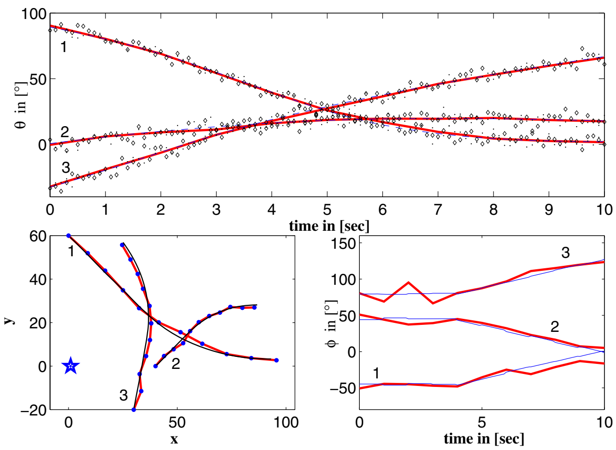

Fig. 3. DOAs, represented by diamonds and dots, are generated by independent noise and are input to the filter unsorted.

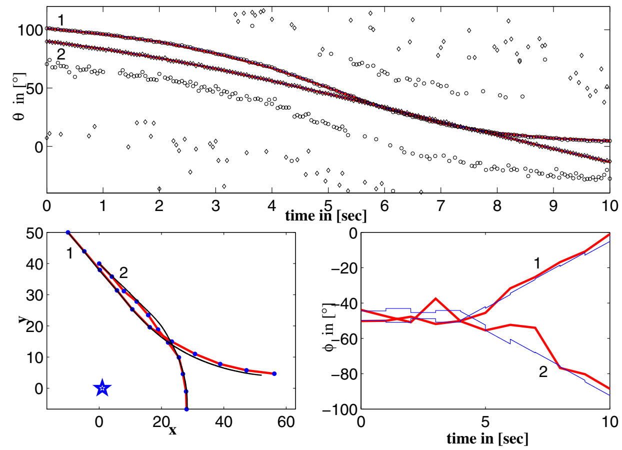

Fig. 4. The filter automatically separates DOAs with respect to the frequency planes defined by f_k in the state. This enables the filter to survive the time period between $t = 5$s and $t = 8$s where the motion vectors for both targets are very similar. Some of the spurious peaks (circles) form a track similar to the true target track due to array spatial aliasing.

a circle such that the minimum inter-microphone distance is 0.45 times the wavelength of the second signal. The acoustic data is processed by a MVDR beamformer and the three highest peaks are picked (in no particular order). Fifty time samples are used to calculate each DOA, hence $M = 20$ for the one-second batch interval.

Note that in this case, ignoring the model dependent angle bias of ($\approx 0.2°$), the calculated DOAs are distributed around the true DOA track with much less variance ($\approx (0.4°)^2$) than the assumed variance $\sigma^2 = (2°)^2$. Hence, the actual data-likelihood is narrower than what is assumed by the tracker. This has the impact of decreasing the number of effective particles that contribute to the estimation accuracy. This is to be expected since the filter is not matched to the data; however, this also demonstrates the robustness of the algorithm under the unmatched prior case.

5. CONCLUSIONS

In this paper, a robust acoustic tracker is formulated in a flexible framework that has minimal assumptions on the observations. Even though the filter requires a batch of data for its update, it can be implemented for online applications because the data-likelihood approximation can be done as the data is being accumulated. Since the filter uses the Bayesian framework, it can avoid the data association problems commonly encountered in target tracking. The multiple target tracking performance can be further improved by altering the data likelihood so that a confusion probability is assigned to each data while the targets are crossing.

6. REFERENCES

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

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

[3] V. Cevher and J. H. McClellan, "General direction-of-arrival tracking with acoustic nodes," *IEEE Trans. on Signal Processing*, vol. 53, no. 1, pp. 1–12, January 2002.

[4] D.H. Johnson and D.E. Dudgeon, *Array Signal Processing: Concepts and Techniques*, Prentice Hall, 1993.

[5] Y. Bar-Shalom and T. Fortmann, *Tracking and Data Association*, Academic-Press, 1988.

[6] R. Karlsson and F. Gustafsson, "Monte Carlo data association for multiple target tracking," in *IEE Target Tracking: Algorithms and Applications*, Netherlands, 16-17 Oct. 2001.

[7] J.S. Liu and R. Chen, "Sequential Monte Carlo methods for dynamic systems," *Journal of the*

American Statistical Association, vol. 93, pp. 1032–1044, September 1998.

[8] M. Isard and J. MacCormick, "BraMBLe: A Bayesian multiple-blob tracker," in 8^{th} *International Conference on Computer Vision*, 2001.

[9] M. Isard and A. Blake, *Active Contours*, Springer, 2000.

[10] P. Stoica and A. Nehorai, "Music, maximum likelihood, and Cramér-Rao bound," *IEEE Trans. on ASSP*, vol. 37, no. 5, pp. 720–741, May 1989.

[11] A. Doucet, N. Freitas, and N. Gordon, Eds., *Sequential Monte Carlo Methods in Practice*, Springer-Verlag, 2001.