ADVANCED PARTICLE FILTERING FOR MULTIPLE OBJECT TRACKING IN DYNAMIC FLUORESCENCE MICROSCOPY IMAGES

Ihor Smal, Wiro Niessen and Erik Meijering

Biomedical Imaging Group Rotterdam
Erasmus MC – University Medical Center Rotterdam
Email: i.smal@erasmusmc.nl

ABSTRACT

Quantitative analysis of dynamical processes in living cells by means of fluorescence microscopy imaging requires tracking of hundreds of bright spots in noisy image sequences. Deterministic approaches that perform object detection prior to tracking usually produce many incorrect tracks. We propose an improved, completely automatic tracker, built in a Bayesian probabilistic framework. It fully exploits spatiotemporal information and prior knowledge, yielding more robust tracking also in case of photobleaching and object interaction. Results from a preliminary quantitative evaluation based on highly realistic synthetic image sequences as well as real fluorescence microscopy image data in comparison with manual tracking indicate superior performance.

Index Terms— Particle filtering, sequential Monte Carlo, multiple object tracking, microtubules, fluorescence microscopy.

1. INTRODUCTION

In the past decade, fluorescence microscopy has proven to be a groundbreaking and effective imaging tool for studying intracellular dynamics. Nowadays, high-throughput experiments generate vast amounts of dynamic image data, which cannot be analyzed manually with sufficient speed, accuracy and reproducibility. Hence, the development of robust and highly reproducible automated tracking methods, which eliminate the bias and variability in human judgment, is of great importance [1, 2]. Recently we have shown [3] that for this purpose, probabilistic tracking approaches, also known as sequential Monte Carlo (SMC) methods or particle filters (PF), are accurate and more robust alternatives to conventional deterministic tracking techniques, which separate the object detection and object linking stages and usually break down at signal-to-noise ratios (SNR)$<$5 [4, 5].

In this paper we propose several important improvements, both fundamental and practical, over the standard PF approach [6, 3]. Specifically, we present a new dynamic model which can incorporate object interaction and photobleaching. In addition, we improve the robustness and reproducibility of the algorithm by using data-dependent importance sampling and a new, completely automatic initialization procedure. The performance of the algorithm is demonstrated on real fluorescence microscopy data sets showing moving vesicles and growing microtubules.

2. TRACKING FRAMEWORK

2.1. Bayesian Tracking and Particle Filtering

The Bayesian tracking approach deals with the problem of inferring knowledge about the object state $\mathbf{x}_t$, which changes over time, using a sequence of noisy measurements $\mathbf{z}_{1:t}$ up to time t, by recursively estimating a time evolving posterior distribution (or filtering distribution) $p(\mathbf{x}_t|\mathbf{z}_{1:t})$. The exact solution to this problem can be constructed by specifying the Markovian probabilistic model of the state evolution $D(\mathbf{x}_t|\mathbf{x}_{t-1})$ and the likelihood $L(\mathbf{z}_t|\mathbf{x}_t)$ that relates the noisy measurements to any given state. The required probability density function (pdf) $p(\mathbf{x}_t|\mathbf{z}_{1:t})$ may be obtained recursively (assuming the initial pdf $p(\mathbf{x}_0|\mathbf{z}_0) \equiv p(\mathbf{x}_0)$ is available) using the Chapman-Kolmogorov equation and Bayes' rule [7]

$$p(\mathbf{x}_t|\mathbf{z}_{1:t}) \propto L(\mathbf{z}_t|\mathbf{x}_t) \int D(\mathbf{x}_t|\mathbf{x}_{t-1}) p(\mathbf{x}_{t-1}|\mathbf{z}_{1:t-1}) d\mathbf{x}_{t-1}. \tag{1}$$

This recursion can be processed sequentially so that it is not necessary to first load the complete data set, nor to reprocess existing data if a new measurement becomes available. The filtering distribution embodies all available statistical information and an optimal (with respect to any criterion) estimate of the state (such as expectation, maximum a posteriori (MAP), or minimum mean square error (MMSE) estimates) may be obtained from the pdf.

The optimal Bayesian solution, defined by the recurrence relation (1) is analytically tractable in a very restrictive set of cases [8]. For most practical models of interest, SMC method [6], [8] is used as an efficient numerical approximation technique, which represents the required posterior density function $p(\mathbf{x}_t|\mathbf{z}_{1:t})$ with a set of N_s random samples, or particles, and associated weights $\{\mathbf{x}_t^{(i)}, w_t^{(i)}\}_{i=1}^{N_s}$. Thus, the

filtering distribution can be approximated as

$$p(\mathbf{x}_t|\mathbf{z}_{1:t}) \approx \sum_{i=1}^{N_s} w_t^{(i)} \delta(\mathbf{x}_t - \mathbf{x}_t^{(i)}), \quad (2)$$

where $\delta(\cdot)$ is the Dirac delta function and the weights are normalized: $\sum_{i=1}^{N_s} w_t^{(i)} = 1$. The samples and weights are then propagated through time to give an approximation of the filtering distribution at subsequent time steps. The weights are recursively updated using sequential importance sampling,

$$w_t^{(i)} \propto \frac{L(\mathbf{z}_t|\mathbf{x}_t^{(i)}) D(\mathbf{x}_t^{(i)}|\mathbf{x}_{t-1}^{(i)})}{q(\mathbf{x}_t^{(i)}|\mathbf{x}_{t-1}^{(i)}, \mathbf{z}_t)} w_{t-1}^{(i)}, \quad (3)$$

where the importance function $q(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{z}_t)$ (subject to some weak constraints) describes which areas of the state-space contain most information about the posterior $p(\mathbf{x}_t|\mathbf{z}_{1:t})$, and $\mathbf{x}_t^{(i)} \sim q(\mathbf{x}_t|\mathbf{x}_{t-1}^{(i)}, \mathbf{z}_t), i = \{1, ..., N_s\}$ are the particles drawn from that importance function. A more detailed formulation of $q(\cdot|\cdot)$ is given below. For very large numbers of samples, this MC characterization becomes an equivalent representation to the usual functional description of the posterior pdf. Having this kind of representation, the MMSE estimate can be approximated as $\hat{\mathbf{x}}_t^{\text{MMSE}} \approx \sum_{i=1}^{N_s} \mathbf{x}_t^{(i)} w_t^{(i)}$. During operation of the particle filter, resampling of the particles according to the importance weights is necessary in order to avoid the degeneracy problem [7].

2.2. Multiple Object Tracking

It is straightforward to generalize the Bayesian formulation to the problem of multi-object tracking. However, due to the increase in dimensionality this formulation leads to an exponential explosion of computational demands. The primary goal in a multi-object tracking application is to determine the multi-modal posterior distribution over the current *joint* configuration of the objects. Multiple modes are caused either by ambiguity about the object state due to insufficient measurements, or by the measurements coming from multiple objects being tracked. To capture the multi-modal nature, which is inherent to our application, the filtering distribution was modeled as an M-component mixture model [9],

$$p(\mathbf{x}_t|\mathbf{z}_{1:t}) = \sum_{m=1}^{M} \pi_{m,t} p_m(\mathbf{x}_t|\mathbf{z}_{1:t}), \quad (4)$$

with $\sum_{m=1}^{M} \pi_{m,t} = 1$ and a non-parametric model (2) is assumed for the individual mixture components $p_m(\mathbf{x}_t|\mathbf{z}_{1:t})$. In this case, the particle representation of the filtering distribution, $\{\mathbf{x}_t^{(i)}, w_t^{(i)}\}_{i=1}^N$ with $N = MN_s$ particles, is augmented with the set of component indicators $\{c_t^{(i)}\}_{i=1}^N$, with $c_t^{(i)} = m$ if particle i belongs to mixture component m. The non–parametric mixture (4) can be updated in the same fashion as the standard Bayesian sequential estimation (1) [9].

3. MODELING AND EXPERIMENTAL RESULTS

3.1. Microtubule Growth Study and Photobleaching

The proposed technique was applied to 2D fluorescence microscopy image sequences of moving vesicles and microtubules (MT) tagged with green fluorescent protein. MTs are protofilaments (diameter ∼25nm) of α and β tubulin, which are studied in terms of structure, localization and dynamic behavior under different experimental conditions. The analysis of the data is tedious and complicated by photobleaching, a dynamic process in which fluorochrome molecules undergo photo-induced chemical destruction upon exposure to excitation light and thus lose their ability to fluoresce. The mechanisms of photobleaching in biological objects are not yet well understood. Commonly, a single-exponential process $I(t) = Ae^{-\alpha t} + B$ is used as a basis for the photobleaching modeling in microscopy [10], where A, B and α are experimentally determined constants.

3.2. Dynamic Model and Object Interactions

The dynamic behavior of the vesicles and visible ends of microtubules is modeled using a nearly constant velocity model [11] with the state vector $\mathbf{x}_t = (x_t, \dot{x}_t, y_t, \dot{y}_t, I_t)^T$. The photobleaching is modeled as a first-order Gauss-Markov process, $I_t = (1 - \alpha)I_{t-1} + u_t$, where I_t is the object intensity at time t, u_t is zero-mean Gaussian noise, and $\alpha \leq 1$ is an experimentally obtained rate of photobleaching, which can also be estimated from image data by model fitting. In this case, $D(\mathbf{x}_t|\mathbf{x}_{t-1})$ is a linear Gaussian model [7], which can easily be evaluated pointwise in (3), and is given by

$$D(\mathbf{x}_t|\mathbf{x}_{t-1}) \propto \exp\left(-\tfrac{1}{2}(\mathbf{x}_t - \mathbf{F}\mathbf{x}_{t-1})^T \mathbf{Q}^{-1} (\mathbf{x}_t - \mathbf{F}\mathbf{x}_{t-1})\right),$$

where $\mathbf{F}$ and $\mathbf{Q}$ are the process transition and covariance matrices respectively [11]. This model captures small accelerations in the object motion and fluctuations in the object intensity present in our image data [3].

To avoid track coalescence and to obtain a more realistic motion model in case of multiple object tracking, we propose to explicitly model the interaction between objects using a Markov random field (MRF) [12]. Here we use a pairwise MRF, expressed by means of a Gibbs distribution

$$\psi_t(\mathbf{x}_t^{(i)}, \mathbf{x}_t^{(j)}) \propto \exp(-d_t^{i,j}), \quad i, j \in \{1, ..., N\}, \quad c_t^{(i)} \neq c_t^{(j)}, \quad (5)$$

where $d_t^{i,j}$ is a penalty function that penalizes the states of two objects $c_t^{(i)}$ and $c_t^{(j)}$ that are closely spaced at time t. That is, $d_t^{i,j}$ is maximal when two objects coincide and gradually falls off as they move apart. This simple pairwise representation is easy to implement yet can be made quite sophisticated. Using this form, we still retain the predictive motion model of each individual target. For that, we sample N_s times the pairs $(\mathbf{x}_{m,t-1}^{(i)}, \mathbf{x}_{m,t}^{(i)})$, $m = \{1, ..., M\}$ from $p_m(\mathbf{x}_{t-1}|\mathbf{z}_{1:t-1})$ and

$q(\mathbf{x}_t|\mathbf{x}^{(i)}_{m,t-1}, \mathbf{z}_t)$ respectively. Taking into account (5), the weights (3) in this case are given by

$$w^{(i)}_{m,t} \propto \frac{L(\mathbf{z}_t|\mathbf{x}^{(i)}_{m,t}) D(\mathbf{x}^{(i)}_{m,t}|\mathbf{x}^{(i)}_{m,t-1})}{q(\mathbf{x}^{(i)}_{m,t}|\mathbf{x}^{(i)}_{m,t-1}, \mathbf{z}_t)} \prod_{k=1,k\neq m}^{M} \psi_t(\mathbf{x}^{(i)}_{m,t}, \mathbf{x}^{(i)}_{k,t}).$$

In our application we have found that an interaction potential based on object positions only is sufficient to eliminate most tracking failures.

3.3. Observation Model and Likelihood

The complete measurement recorded by a CCD camera at time t is an $N \times M$ image, $\mathbf{z}_t = \{z_t(i,j) : i = 1, ..., N, j = 1, ..., M\}$, with $z_t(i,j)$ the measured intensity at pixel (i,j), which corresponds to a rectangular region of dimensions $\Delta_x \times \Delta_y$ nm^2. Due to limited spatial resolution ($\sim$200nm) in light microscopy, subcellular structures (typically of size <20nm) are imaged as blurred spots. The blurring is characterized by the point-spread function (PSF) of the microscope. We use a Gaussian approximation of the PSF [4]. To model the elongation in the intensity profile of MTs we use the velocity components from the state vector $\mathbf{x}_t$ as parameters in the PSF. Thus, for an object of intensity I_t at position (x_t, y_t), the intensity contribution to pixel (i,j) is approximated as:

$$h_t(i,j;\mathbf{x}_t) = I_t \exp\left(-\tfrac{1}{2}\mathbf{m}^T\mathbf{R}^T\mathbf{\Sigma}^{-1}\mathbf{R}\mathbf{m}\right) + b_t \qquad (6)$$

where $\mathbf{R} = \mathbf{R}(\phi)$ is a rotation matrix, $\tan\phi = \dot{y}_t/\dot{x}_t$, $\mathbf{m} = (i\Delta_x - x_t, j\Delta_y - y_t)^T$, $\mathbf{\Sigma} = \mathrm{diag}[\sigma^2_{\text{max}}, \sigma^2_{\text{min}}]$ and b_t is an estimated background level, which is defined as the average image intensity at time t (the contribution of object intensity values is negligible). The parameters σ_{max} and σ_{min} represent the amount of blurring and, at the same time, model the elongation of the object along the direction of motion.

The proposed likelihood is defined as

$$L_S(\mathbf{z}_t|\mathbf{x}_t) \propto \frac{1}{\sigma_\eta(\mathbf{x}_t)} \exp\left(-\frac{\left(S^z_t(\mathbf{x}_t) - S^o_t(\mathbf{x}_t)\right)^2}{2\sigma^2_\eta(\mathbf{x}_t)}\right), \quad (7)$$

with $S^z_t(\mathbf{x}_t) = \sum_{(i,j)\in C(\mathbf{x}_t)} z_t(i,j)$ and $S^o_t(\mathbf{x}_t) = \sum_{(i,j)\in C(\mathbf{x}_t)} h_t(i,j)$,

where $C(\mathbf{x}_t) = \{(i,j) \in \mathbb{Z}^2 : h_t(i,j;\mathbf{x}_t) - b_t > 0.1 I_t\}$, and $\sigma^2_\eta = S^o_t$ is taken to approximate the Poisson distribution. The recursive Bayesian solution is applicable as long as the statistics of the measurement noise are known for each pixel.

3.4. Data-Dependent Sampling

Standard PF [6] uses $q(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{z}_t) = D(\mathbf{x}_t|\mathbf{x}_{t-1})$ and usually performs poorly because too few samples are generated in regions where the desired posterior $p(\mathbf{x}_t|\mathbf{z}_{1:t})$ is large. In order to construct the proposal distribution which alleviates this problem and takes into account the most recent measurements $\mathbf{z}_t$, we propose to transform the image sequence into probability distributions, knowing that the PSF blurs the point light sources and adds uncertainty to their positions. For each object we use the following convolution transformation

$$\tilde{p}_m(x,y|\tilde{z}_{m,t}) = \frac{(G_\sigma * \tilde{z}_{m,t})^r}{\int_{C_{m,t}} (G_\sigma * \tilde{z}_{m,t})^r dxdy}, \quad r > 1, \quad (8)$$

where G_σ is the Gaussian kernel with scale σ and $\tilde{z}_{m,t}(x,y)$ is the first-order interpolation of $\mathbf{z}_t$ in the circular region $C_{m,t}$ (the radius is defined by the covariance matrix of $D(\mathbf{x}_t|\mathbf{x}_{t-1})$, e.g. 3-standard-deviation level) centered at the object position predicted from the previous time step. The new data dependent proposal distribution for object m is defined as

$$\tilde{q}_m(\mathbf{x}_t|\tilde{z}_{m,t}) = \tilde{p}_m(x_t, y_t|\tilde{z}_{m,t}) \mathcal{N}(I_t|\tilde{z}_{m,t}(x_t,y_t) - b_t, \sigma^2_I) \times \mathcal{N}(\dot{x}_t|x_t - \hat{x}^{\text{MMSE}}_{m,t-1}, \sigma^2_{\dot{x}_t}) \mathcal{N}(\dot{y}_t|y_t - \hat{y}^{\text{MMSE}}_{m,t-1}, \sigma^2_{\dot{y}_t}),$$

where $\mathcal{N}(\cdot|\mu, \sigma^2)$ indicates a real normal distribution. For better performance we use a mixture of both proposals,

$$q_m(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{z}_t) = \gamma D(\mathbf{x}_t|\mathbf{x}_{t-1}) + (1-\gamma)\tilde{q}_m(\mathbf{x}_t|\tilde{z}_{m,t}),$$

where $0 < \gamma < 1$ balances the two proposals. We have found that this proposal distribution is uniformly superior to the regular one and scales much better to smaller sample sizes.

3.5. Track Initialization and Management

The proposed PF can be initialized manually in the first image frame by specifying regions of interest (ROI). For completely automatic initialization in the first frame and also for detection of newly appearing objects for tracking in the subsequent frames, the image space is divided into rectangular cells and N_s particles are sampled according to importance function (8), where $\bar{z}_{m,t}$ is the whole image. The number of sampled particles in each cell represents the degree of belief in object birth. In cells (not containing any of the M existing objects) with counts larger than some threshold N_{td}, new tracks are initiated with initial mixture weights π_{bd}. The threshold N_{td} can be estimated experimentally and depends on N_s, cell area and the number of bright spots in the image data.

Whenever objects pass close to one another, the object with the best likelihood score typically "hijacks" particles of nearby mixture components. This problem is partly solved using the MRF model for object interactions. To better resolve the ambiguity in such situations, the Hough transformation is used for each spatiotemporal ROI of 3-5 frames $C_{m,t-\tau:t+\tau}$ of track m to correctly model the velocity changes. If object m passes close to the object which was not tracked yet, the distribution $p_m(\mathbf{x}_t|\mathbf{z}_{1:t})$ becomes too diffuse in a few time steps and the reclustering procedure $(\{c'^{(i)}_t\}, M') = F(\{\mathbf{x}^{(i)}_t\}, \{c^{(i)}_t\}, M)$ [9] is performed to initiate new tracks. The merging of the objects does not occur in our application and is therefore forbidden. If the mixture weight $\pi_{m,t}$ is below some predefined threshold level π_{td}, the component m is removed from the mixture and the track m is terminated at time t.

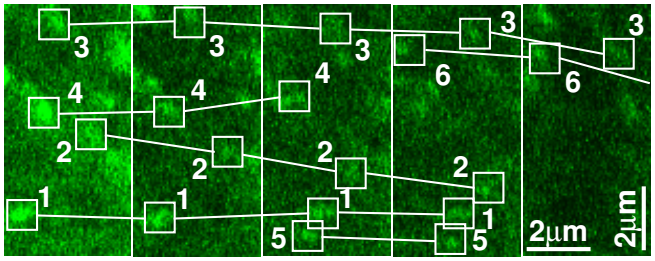

Fig. 1. Visualization of the results (six tracks) of tracking microtubules (bright spots) in presence of photobleaching using our technique (single frames, taken every 2 sec, from 2D time-lapse studies). This example illustrates the capability of our algorithm to capture newly appearing objects (tracks 5 and 6) and to detect object disappearance (e.g. track 4). It also shows the robustness of the algorithm in case of closely passing objects (tracks 1 and 5).

3.6. Evaluation on Real Data

The tracker was tested on three 2D+T image data sets of size $512 \times 512 \times 20$ showing moving vesicles and growing microtubules in the presence of photobleaching. The data sets were preselected from larger volumes by manually choosing the ROIs. The results of tracking MTs are presented in Figs. 1 and 2. In the experiments the tracker correctly followed on average 15-20 automatically detected spots simultaneously. The evaluation of the algorithm was done by visual inspection and by quantitative comparison with the results of manual tracking of the same spots. Lacking ground truth for the real data, we calculated the localization deviations between manual and our automatic tracking. These ranged from 50nm to 180nm depending on SNR, which is in good agreement with our previous results on synthetic image data [3]. The velocity estimates derived from the trajectories compare well to the estimates obtained by expert biologists by means of manual tracking in the same image data.

4. CONCLUSIONS

We have presented several important improvements over the basic PF-based tracking approach presented earlier [3]. These have led to a fully automatic algorithm for robust and accurate detection and tracking of time varying numbers of objects in image sequences obtained by dynamic fluorescence microscopy imaging. While we have presented results on 2D+T image data, the algorithm can straightforwardly (simply by extending the state vector with z-components) be applied to (multi-channel) 3D+T image data. The experimental tracking results obtained with our algorithm compare well with manual tracking results of expert biologists. Our findings encourage use of the described approach to analyze complex biological image sequences not only for obtaining statistical estimates of average velocity and life span, but also for detailed analysis of complete life histories.

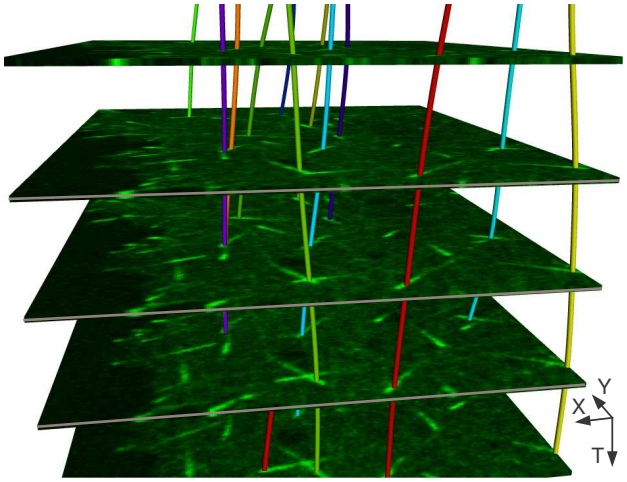

Fig. 2. Visualization of a 2D+T fluorescence microscopy image sequence showing microtubules (bright, elongated spots) and the results of tracking using our technique.

5. REFERENCES

[1] E. Meijering, I. Smal, and G. Danuser, "Tracking in molecular bioimaging," *IEEE Sig. Proc. Mag.*, vol. 23, no. 3, pp. 46–53, May 2006.

[2] A. Genovesio, T Liedl, V. Emiliani, W.J. Parak, M. Coppey-Moisan, and J.-C. Olivo-Marin, "Multiple particle tracking in 3-D+t microscopy: method and application to the tracking of endocytosed quantum dots," *IEEE Trans. Im. Proc.*, vol. 15, no. 5, pp. 1062–70, May 2006.

[3] I. Smal, W. Niessen, and E. Meijering, "Particle filtering for multiple object tracking in molecular cell biology," in *Proc. Nonlinear Statistical Signal Processing Workshop*, Sep. 2006.

[4] D. Thomann, D. R. Rines, P. K. Sorger, and G. Danuser, "Automatic fluorescent tag detection in 3D with super-resolution: application to the analysis of chromosome movement," *J. Microsc.*, vol. 208, no. 1, pp. 49–64, Oct. 2002.

[5] M. K. Cheezum, W. F. Walker, and W. H. Guilford, "Quantitative comparison of algorithms for tracking single fluorescent particles," *Biophys. J.*, vol. 81, no. 4, pp. 2378–88, Oct. 2001.

[6] M. Isard and A. Blake, "CONDENSATION – conditional density propagation for visual tracking," *Int. J. Comp. Vis.*, vol. 29, no. 1, pp. 5–28, Aug. 1998.

[7] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and Computing*, vol. 10, no. 3, pp. 197–208, July 2000.

[8] S. M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Sig. Proc.*, vol. 50, no. 2, pp. 174–88, Feb. 2002.

[9] J. Vermaak, A. Doucet, and P. Pérez, "Maintaining multi-modality through mixture tracking," in *Proc. 9th IEEE Int. Conf. Comp. Vis.*, Oct. 2003, pp. 1110–16.

[10] L. Song, E. J. Hennink, I. T. Young, and H. J. Tanke, "Photobleaching kinetics of fluorescein in quantitative fluorescence microscopy," *Biophys. J.*, vol. 68, no. 6, pp. 2588–600, June 1995.

[11] X. R. Li and V. P. Jilkov, "Survey of maneuvering target tracking: Part I: Dynamic models," *IEEE Trans. Aero. Elec. Sys.*, vol. 39, no. 4, pp. 1333–64, Oct. 2003.

[12] Z. Khan, T. Balch, and F. Dellaert, "MCMC-based particle filtering for tracking a variable number of interacting targets," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 27, no. 11, pp. 1805–918, Nov. 2005.