

Multitarget Tracking with Split and Merged Measurements

Zia Khan, Tucker Balch, Frank Dellaert
Georgia Institute of Technology College of Computing
{zkhan,tucker,dellaert}@cc.gatech.edu

Abstract

In many multitarget tracking applications in computer vision, a detection algorithm provides locations of potential targets. Subsequently, the measurements are associated with previously estimated target trajectories in a data association step. The output of the detector is often imperfect and the detection data may include multiple, split measurements from a single target or a single merged measurement from several targets. To address this problem, we introduce a multiple hypothesis tracker for interacting targets that generate split and merged measurements. The tracker is based on an efficient Markov chain Monte Carlo (MCMC) based auxiliary variable particle filter. The particle filter is Rao-Blackwellized such that the continuous target state parameters are estimated analytically, and an MCMC sampler generates samples from the large discrete space of data associations. In addition, we include experimental results in a scenario where we track several interacting targets that generate these split and merged measurements.

1. Introduction

In many multitarget tracking applications in computer vision, detection precedes tracking. At each time step, a detector provides locations of potential targets and, subsequently, in a data association step the measurements are associated with known target trajectories. However, the detection step can be imperfect and several common types of errors are often prevalent in the detection data: (1) A target may return more than one, split measurement per target. (2) Interacting targets may return a single merged measurement. (3) A target may fail to return a measurement at the current time step. (4) A detector may return a false detection. A number of heuristic solutions have been proposed to address this problem, motivated by specific tracking applications (e.g. surveillance [9]).

Probabilistic algorithms for tracking and data association, which emerged from the radar tracking literature (see [15] for a review), fail to adequately address this

problem as they make the following assumptions: (1) a target can generate at most one measurement at every time step (2) a measurement could have originated from at most one target. While reasonable for radar, these assumptions are often violated in multitarget tracking applications in computer vision.

Merged, or unresolved, targets have been addressed only to a limited extent in the radar tracking literature. In [6], Chang and Bar-Shalom describe a version of the Joint Probabilistic Data Association Filter (JPDAF) which can address situations where two targets are unresolved, or merged. However, the JPDAF represents the belief over the state of the targets as a Gaussian, and may not accurately capture the a multi-modal distribution over the target states. Consequently, a multiple hypothesis tracker was introduced in [11] to handle these unresolved targets, but the approach was also limited to two closely spaced targets.

Recently, Genovesio and Olivo-Marin considered the general case where the set of detector measurements can be augmented by a set of “virtual measurements” that account for feasible split and merged detections in the context of visual tracking [8]. While this is a compelling idea, introducing virtual measurements does not provide a model for split and merged measurements providing flexibility in modelling their formation.

In this paper, we introduce a multiple hypothesis tracker for interacting targets that generate split and merged measurements. The tracker is based on an efficient Markov chain Monte Carlo (MCMC) based auxiliary variable particle filter [14]. The particle filter is Rao-Blackwellized such that the continuous target state parameters are integrated out analytically [12, 5]. The advantage of this is that fewer samples are needed since part of the posterior over the state is analytically calculated, rather than being approximated using a more expensive and noisy sample representation.

The particle filter relies on a Markov chain Monte Carlo (MCMC) sampler that generates samples from the large discrete space of data associations. MCMC samplers have used to address data association problems in multitarget tracking [2, 13] and structure from motion

[7]. Both occlusion and spurious measurements, were addresses in previous work. This work represents an advance in that it also addresses the case of merged and split measurements.

The algorithm we describe updates a set of hybrid particles at each time step as measurements arrive. Each hybrid particle contains a multivariate Gaussian over the joint configuration of the targets. To update the hybrid particles, we sample from the posterior over data associations, those most consistent with the data, by running an MCMC sampler over the space of possible data associations. Using the sampled data associations, we can compute a hybrid particle approximation of the posterior over the current target positions.

In the remainder of this paper, we first review a model for tracking with split and merged measurements in Section 2, and express the tracking problem recursively as a Bayes filter. Next, we detail an efficient method for approximate inference in this problem that leverages an auxiliary variable particle filter in Section 3. Last, in Section 4 we detail experimental results in a scenario where we track several interacting targets that generate split and merged measurements.

2. Tracking Model

In this section, describe a model for tracking, stating assumptions needed to obtain an efficient inference algorithm. First, we specify the joint distribution over the actual measurements $Z_{1:T}$, data associations $\mathbf{J}_{1:t}$, and states $X_{0:T}$ of the targets between time steps 0 to T . We assume that the number of targets N is fixed, but the number of measurements M can change at each time step t . If we assume that the target motion is Markov, the state at the current time step depends only on the previous time step, we can factor the joint $P(Z_{1:T}, X_{0:T}, \mathbf{J}_{1:t})$ distribution as follows

$$P(X_0) \prod_{t=1}^T P(X_t | X_{t-1}) P(Z_t, \mathbf{J}_t | X_t)$$

To further simplify the model, we make some additional assumptions. We assume the initial joint state is Gaussian

$$P(X_0) = \mathcal{N}(X_0; m_0, V_0)$$

where the covariance matrix V_0 is a diagonal and targets move according to a Gaussian random-walk model

$$P(X_t | X_{t-1}) = \mathcal{N}(X_t; X_{t-1}, \Gamma)$$

Since measurements arrive a random order, the actual states of the targets do not provide us with any information on which data associations are likely. As a conse-

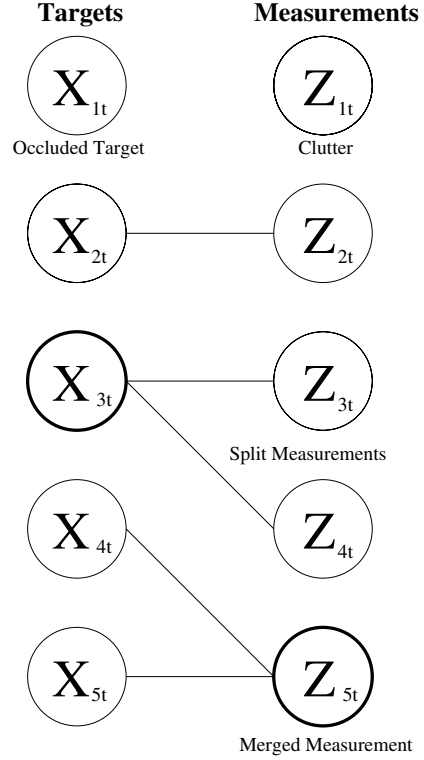


Figure 1: We sample over bipartite graphs between targets and measurements. We assume that merge and split events are mutually exclusive. Targets cannot merge with another target and split at the same time.

quence, we assume that the data associations do not depend on the target states

$$P(Z_t, \mathbf{J}_t | X_t) = P(Z_t | \mathbf{J}_t, X_t) P(\mathbf{J}_t)$$

The data associations themselves are represented by a bipartite graph $\mathbf{J}_t = (X_{(1:N)t}, Z_{(1:M)t}, \mathbf{E}_t)$ which consist of a set of target nodes $X_{(1:N)t}$ and measurement nodes $Z_{(1:M)t}$ each connected by a set of edges \mathbf{E}_t as shown in Figure 1. We restrict the possible data associations to the space of the bipartite graphs in which any tree formed by taking a split target or merged measurement node has a height of one. Our choice reflects the assumption that merge and split events are mutually exclusive. Targets cannot merge with another target and split at the same time.

The prior over the data associations can be used to favor simple data associations where few targets are merged or split

$$P(\mathbf{J}_t) = \exp(-\#(\text{merged})\alpha + -\#(\text{split})\beta)$$

Next, we can use the data associations \mathbf{J}_t to divide the measurements into clutter and observations (Z_c, Z_o) =

$f(\mathbf{J}_t, Z_t)$ respectively

$$P(Z_t|X_t, \mathbf{J}_t) = P(\mathbf{Z}_c)P(\mathbf{Z}_o|\mathbf{J}_t, X_t)$$

We assume the density over clutter is uniform and does not depend on the state of the targets or data associations $P(\mathbf{Z}_c) = \mathcal{U}(\cdot)$.

To model the observations, we map $\mathbf{H} = g(\mathbf{J}_t)$ the data associations to a *sparse* measurement matrix in a Gaussian observation model

$$P(\mathbf{Z}_o|\mathbf{J}_t, X_t) = \mathcal{N}(\mathbf{Z}_o; \mathbf{H}X_t, \Sigma)$$

The columns of the matrix H correspond to individual target states and the rows observed measurements. For every edge connected to a measurement node, an identity matrix is placed in the corresponding row and column. For merged measurements, where there are several edges connecting the measurement to multiple targets. The identity matrix is multiplied by the inverse of the number of edges. This choice reflects the assumption that merged measurements occur at the centroid of merged targets. An example measurement matrix is shown in Figure 2.

$$H = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 \end{bmatrix}$$

$$X_t = \begin{bmatrix} x_{1t} \\ x'_{1t} \\ x_{2t} \\ x'_{2t} \\ x_{3t} \\ x'_{3t} \\ x_{4t} \\ x'_{4t} \\ x_{4t} \\ x'_{4t} \end{bmatrix} \quad \mathbf{Z}_o = \begin{bmatrix} z_{2t} \\ z'_{2t} \\ z_{3t} \\ z'_{3t} \\ z_{4t} \\ z'_{4t} \\ z_{5t} \\ z'_{5t} \end{bmatrix} \quad \mathbf{Z}_c = \begin{bmatrix} z_{1t} \\ z'_{1t} \end{bmatrix}$$

Figure 2: H shows the sparse measurement matrix for the bipartite graph in Figure 1. X_t shows the joint target state. Both the individual measurements and individual targets are 2-dimensional. The observed measurements \mathbf{Z}_o and the clutter measurements \mathbf{Z}_c are shown as well.

In tracking, we observe the measurements $Z_{1:T}$ but we do not observe the actual states of the targets $X_{0:T}$. Consequently, the objective of all tracking algorithms is to infer exactly or approximately the probability distribution over the current position X_t of the targets given all of the measurements observed so far.

It is convenient to write inference in this model recursively via the Bayes filter. The posterior distribution $P(X_t|Z^t)$ over the joint state X_t of *all* present targets is given all observations $Z^t = \{Z_1, \dots, Z_t\}$ up to and including time t , is updated according to

$$P(X_t|Z^t, M^t) = k \sum_{\mathbf{J}_t} P(Z_t, \mathbf{J}_t|X_t) \times \int_{X_{t-1}} P(X_t|X_{t-1}) \times P(X_{t-1}|Z^{t-1})$$

Note that we marginalize over the data associations \mathbf{J}_t . The space of data associations is quite large. We can approximate this marginalization using Markov chain Monte Carlo (MCMC) methods.

3. Inference

In this section, we describe a MCMC-based variant of the auxiliary variable particle filter to approximately infer the position of the targets [14]. The filter samples over the data associations and sample indices jointly, discarding the sample indices. In doing so, the target ratio calculation in the filter becomes computationally efficient.

Next, we show how we can obtain a Rao-Blackwellized version of the particle filter for the described tracking model [5]. In the Rao-Blackwellized filter, the continuous target state parameters are integrated out analytically [12, 5]. This reduces the variance of the Monte Carlo approximation of the posterior. Consequently, fewer samples are needed since part of the posterior over the state is analytically calculated, rather than being approximated using a more expensive and noisy sample representation.

3.1. MCMC-based Auxiliary PF

In a typical joint particle filter, one inductively assumes that the posterior distribution over the joint state of the targets at the previous time step is approximated by a set of N samples

$$P(X_{t-1}|Z^{t-1}) \approx \{X_{t-1}^{(r)}\}_{r=1}^N$$

Given this approximate representation, we obtain the following Monte Carlo approximation of the Bayes filter:

$$P(X_t|Z^t) \approx k P(Z_t|X_t) \frac{1}{N} \sum_{i=1}^N P(X_t|X_{t-1}^{(r)}) \quad (1)$$

In a MCMC-based particle filter this becomes the target distribution from which we sample at each time step.

Note the summation over the target states specifies a prior over the current state

$$P(X_t|Z^{t-1}) \approx \frac{1}{N} \sum_r P(X_t|X_{t-1}^{(r)})$$

One typically needs to evaluate this summation in the target ratio at every time step which becomes costly as the number of samples increase [3, 10].

Auxiliary variable particle filters address this problem by sampling over the the joint distribution over the sample index r , the auxiliary variable, and the state X_t , discarding the sample indices - as this is equivalent to sampling from the mixture prior on X_t as in (1). By eliminating the summation from the target density, the auxiliary variable particle filter offers considerable computational savings.

3.2. Rao-Blackwellization

Note that we can write the posterior over the target positions and data associations $P(X_t, \mathbf{J}_t|Z^t)$ as

$$kP(X_t|\mathbf{J}_t Z^t)P(\mathbf{J}_t|Z^t)$$

The target motion follows a Gaussian. Consequently, the first term can be computed analytically using a Kalman filter, provided we know the data association \mathbf{J}_t . If we had samples from the posterior on data associations $P(\mathbf{J}_t|Z^t)$, a hybrid Monte Carlo approximation of the posterior over the target positions can be obtained. This reasoning provides the basis for a Rao-Blackwellized sampling scheme.

To obtain a Rao-Blackwellized particle filter, we use a hybrid approximation of the posterior over the current state X_t and current data associations \mathbf{J}_t , instead of using a sample set over the states

$$P(X_t, \mathbf{J}_t|Z^t) \approx \frac{1}{N} \sum_{r=1}^N P(X_t|\mathbf{J}_t^{(r)}, Z^t) \delta(\mathbf{J}_t^{(r)})$$

To obtain the hybrid approximation for the tracking model, we must be able to sample from the posterior on the data associations $P(\mathbf{J}_t|Z^t)$. Because the target motion is linear Gaussian, the posterior can be written in terms of a tractable integral over the current state

$$P(\mathbf{J}_t|Z^t) \propto P(\mathbf{J}_t) \times \int_{X_t} P(Z_t|\mathbf{J}_t, X_t) P(X_t|Z^{t-1})$$

where the predictive prior on the current state is a mixture of Gaussians

$$P(X_t|Z^{t-1}) \approx \frac{1}{N} \sum_{r=1}^N \mathcal{N}(X_t; m_{t-1}^{(r)}, P_{t-1}^{(r)} + \Gamma)$$

Using the mixture naively in an MCMC-based particle filter would require us to compute the integral for each mixture component. To address this problem, we use an auxiliary variable particle filter approach detailed in Section 3.1 to obtain an efficient Rao-Blackwellized sampler. We sample the indices and data associations jointly

$$P(\mathbf{J}_t, r|Z^t) \propto P(\mathbf{J}_t)P(\mathbf{Z}_c) \times \int_{X_t} \mathcal{N}(\mathbf{Z}_o; \mathbf{H}X_t, \Sigma) \times \mathcal{N}(X_t; m_{t-1}^{(r)}, P_{t-1}^{(r)} + \Gamma)$$

To compute the integral we use the fact that the integral of any function $q(X_t)$ proportional to a Gaussian, in our case the function is the product of the likelihood and the prior

$$q(X_t) = \mathcal{N}(Z_t; \mathbf{H}X_t, \Sigma) \mathcal{N}(X_t; m_{t-1}^{(r)}, P_{t-1}^{(r)} + \Gamma)$$

is equal to the a maximum of that function X_t^* times a proportionality constant

$$\int_{X_t} q(X_t) = \sqrt{|2\pi P_t^{(r)}|} q(X_t^*)$$

We obtain X_t^* by solving the normal equations after mapping the data associations to an observation matrix $\mathbf{H} = g(\mathbf{J}_t)$ where $Q = P_{t-1}^{(r)} + \Gamma$

$$(\mathbf{H}^\top \Sigma^{-1} \mathbf{H} + Q^{-1}) X_t^* = \mathbf{H}^\top \Sigma^{-1} \mathbf{Z}_o + Q^{-1} m_{t-1}^{(r)} \quad (2)$$

by computing the covariance $P_t^{(r)} = (\mathbf{H}^\top \Sigma^{-1} \mathbf{H} + Q^{-1})^{-1}$. This procedure describes a method for solving recursive least squares problems, and is equivalent to one forward iteration of the Kalman filter.

Hence, the Rao-Blackwellized MCMC sampling algorithm, roughly, consists of the following: (1) propose to use a certain sample index r , (2) propose a new data association \mathbf{J}_t , (3) compute the covariance $P_t^{(r)}$, (4) finding the optimal state X_t^* , and (5) accept or reject the new data association by computing the following target density

$$\pi(\mathbf{J}_t, r) = P(\mathbf{J}_t)P(\mathbf{Z}_c) \sqrt{|2\pi P_t^{(r)}|} q(X_t^*)$$

Note the updated mean and variance for the s th sample is the optimal state $m_t^{(s)} = X_t^*$ and computed covariance $P_t^{(s)} = P_t^{(r)}$. The sampling algorithm can be spelled out precisely by defining a Markov chain on the space of bipartite graphs.

3.3. Markov Chain

In this section, we define a procedure for running a reversible Markov chain on the space of bipartite graphs with the properties described in Figure 1. The stationary distribution of the Markov is the posterior over data associations $P(\mathbf{J}_t|Z^t)$. The chain operates exclusively on the set of edges. We use \mathbf{E}^* to denote the set of all possible edges between target $X_{(1:N)_t}$ and measurement $Z_{(1:M)_t}$ nodes. The procedure used to run the the Markov chain is described below:

1. Pick a sample r uniformly at random (u. a. r.).
2. Pick an edge $e = \{i, j\} \in \mathbf{E}^*$ u. a. r.
3. *Edge Deletion*: If the edge e is in the current edge set \mathbf{E}_t , delete the edge $\mathbf{E}'_t \leftarrow \mathbf{E}_t \setminus \{e\}$ to obtain a new data association \mathbf{J}'_t .
4. *Edge Addition*: If the edge e is not in the current edge set \mathbf{E}_t , add the edge to the edge set $\mathbf{E}'_t \leftarrow \mathbf{E}_t \cup \{e\}$ to obtain a new data association \mathbf{J}'_t . Verify that adding the edge does not introduce a situation where a target is splitting and merging at the same time. This can be done efficiently by computing the the tree height relative to the measurement node Z_{jt} and the target node X_{it} . If height of either tree exceeds one or the height of both trees is one, then go back to step 1.
5. *Acceptance Ratio*: Accept \mathbf{J}'_t with probability a , where

$$a = \min \left\{ \frac{\pi(\mathbf{J}'_t, r)}{\pi(\mathbf{J}_t, r)}, 1 \right\}$$

3.4. Algorithm Summary

In this section, we summarize the steps of the tracking algorithm:

1. Start with a set of R hybrid samples

$$\{\mathcal{N}(X_{t-1}; m_{t-1}^{(r)}, P_{t-1}^{(r)})\}_{r=1}^R$$

which approximate the posterior over the target state at the previous time step $P(X_{t-1}|Z^{t-1})$.

2. Starting with a random data association where targets do not merge and split at the same time. Run the Markov chain described in Section 3.3 until convergence. Take M sampled data associations $\{\mathbf{J}_t^{(s)}\}_{s=1}^S$ at widely separated iterations of the Markov chain to limit correlation between the samples.

3. Map the data association to an observation matrix $\mathbf{H} = g(\mathbf{J}_t^{(s)})$, and separate the measurements into observations and clutter $(\mathbf{Z}_c, \mathbf{Z}_o) = f(\mathbf{J}_t, Z_t)$. Solve equation (2) for each data association $\mathbf{J}_t^{(s)}$ to obtain a new hybrid sample set

$$\{\mathcal{N}(X_t; m_t^{(s)}, P_t^{(s)})\}_{s=1}^S$$

which approximates the posterior distribution over the current state $P(X_t|Z^t)$.

4. Experimental Results

Because the application of visual tracking technologies to monitoring the movement of animals has important implications in the study of behavior, we applied the algorithm to tracking interacting ants in a behavioral experiment [1]. The application presents a substantial challenge as targets deform and frequently interact. Consequently, merged and split measurements are common during tracking.

To obtain detections we use the efficient and simple color segmentation algorithm detailed in [4]. During each run, the measurement noise and movement noise were set to $\Sigma = 8I_{M_t}$ and $\Gamma = 35I_d$ respectively. The posterior was represented by a total of $R = S = 50$ hybrid samples. The sampler was run 500 iterations where every 50th sample was taken until the total 50 a samples were obtained. The estimated position was obtained by computing the mean of the means of the hybrid samples.

The results of two different tracking runs are shown in Figure 3. The tracking runs shown in Figure 3(a) and Figure 3(b) spanned 500 and 1000 frames respectively. In run (a), the tracker correctly estimated the target positions. In the tracking run (b) the tracker temporarily lost track during a sustained interaction between all three targets. Dealing with data associations in such complex tracking situations is an area of future work.

5. Conclusions

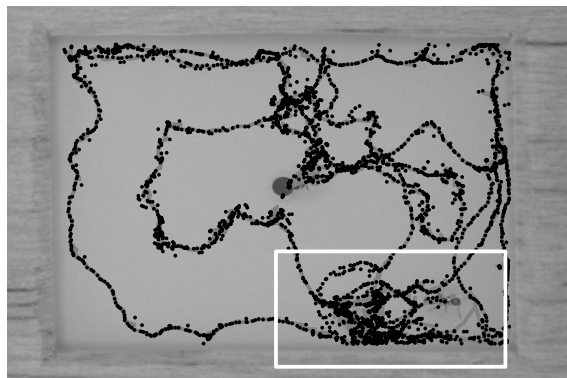
In this work, we introduced a multiple hypothesis tracker for interacting targets that generate split and merged measurements. The tracker is based on an efficient Rao-Blackwellized MCMC-based auxiliary variable particle filter. We are currently investigating models of behavior to improve both tracking and data association results during interactions.

Acknowledgments

This work was funded under NSF Award IIS-0219850, "ITR: Observing, Tracking and Modeling Social Multi-



(a)



(b)

Figure 3: (a) shows a typical tracking run where the detections are plotted in blue and the estimated trajectories of each target. (b) shows a tracking run in which a failure, shown in the box, occurred over a sustained interaction. During the interaction, targets underwent several merges and splits over time.

agent Systems”. We would like to thank Ananth Ranganathan and Eric Vigoda for their insightful comments. We would like to thank Stephan Pratt at the Princeton University Department of Ecology and Evolutionary Biology for the *Leptothorax curvispinosus* video sequence.

References

- [1] T. Balch, Z. Khan, and M. Veloso. Automatically tracking and analyzing the behavior of live insect colonies. In *Proc. Autonomous Agents 2001*, Montreal, 2001.
- [2] N. Bergman and A. Doucet. Markov chain Monte Carlo data association for target tracking. In *Intl. Conf. Acoust., Speech, and Signal Proc. (ICASSP)*, 2000.
- [3] C. Berzuini, N. G. Best, W.R. Gilks, and C. Larizza. Dynamic conditional independence models and Markov chain Monte Carlo methods. *Journal of the American Statistical Association*, 92:1403–1412, 1996.
- [4] J. Bruce, T. Balch, and M. Veloso. Fast and inexpensive color image segmentation for interactive robots. In *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2000.
- [5] G. Casella and C.P. Robert. Rao-Blackwellisation of sampling schemes. *Biometrika*, 83(1):81–94, March 1996.
- [6] Y.C. Chang and Y. Bar-Shalom. Joint probabilistic data association for multitarget tracking with possibly unresolved measurements and maneuvers. In *IEEE Trans. Automat. Contr.*, volume 29, pages 585–594, 1984.
- [7] F. Dellaert, S.M. Seitz, C.E. Thorpe, and S. Thrun. EM, MCMC, and chain flipping for structure from motion with unknown correspondence. *Machine learning*, 50(1-2):45–71, January - February 2003. Special issue on Markov chain Monte Carlo methods.
- [8] A. Genovesio and J. C. Olivo-Marin. Split and merge data association filter for dense multi-target tracking. In *Intl. Conf. on Pattern Recognition (ICPR)*, pages 677–680, 2004.
- [9] I. Haritaoglu, D. Harwood, and L. S. Davis. W4: Real-time surveillance of people and their activities. *PAMI*, 22(8):809–830, 2000.
- [10] Z. Khan, T. Balch, and F. Dellaert. An MCMC-based particle filter for tracking multiple interacting targets. In *Eur. Conf. on Computer Vision (ECCV)*, 2004.
- [11] W. Koch and G. van Keuk. Multiple hypothesis track maintenance with possibly unresolved measurements. *IEEE Trans. Aerosp. Electron. Syst.*, 33(3):589–601, 1997.
- [12] K. Murphy and S. Russell. Rao-Blackwellised particle filtering for dynamic Bayesian networks. In A. Doucet, N. de Freitas, and N. Gordon, editors, *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, New York, January 2001.
- [13] S. Oh, S. Russell, and S. Sastry. Markov chain Monte Carlo data association for general multiple target tracking problems. In *Proc. 43rd IEEE Conference on Decision and Control (CDC-04)*, 2004.
- [14] M. K. Pitt and N. Shephard. Auxiliary variable based particle filters. In Arnaud Doucet, Nando de Freitas, and Neil Gordon, editors, *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, New York, 2001.
- [15] Robert Popoli and Samuel S. Blackman. *Design and Analysis of Modern Tracking Systems*. Artech House Radar Library, August 1999.