MVA2011 IAPR Conference on Machine Vision Applications, June 13-15, 2011, Nara, JAPAN

An Adaptive Particle Filter Method for Tracking Multiple Interacting Targets

Isabella Szottka and Matthias Butenuth
Remote Sensing Technology, Technische Universität München
80333 München, Germany
{isabella.szottka, matthias.butenuth}@bv.tum.de

Abstract

In this paper, we present a new adaptive particle filter method for tracking multiple interacting targets. We introduce a sampling strategy to consecutively sample particles and adapt their spreading according to current measurements. For multiple targets, we develop a new concept of guiding the particles by current measurements of adjacent targets. Our method does not increase the computational complexity. Experimental results for tracking of multiple vehicles demonstrate that the search space is explored more efficiently. Further, we show that our method improves the robustness of the tracker when the target motion abruptly changes or the target is totally occluded.

1 Introduction

Visual tracking of multiple objects is of broad interest, because it provides the input for various surveillance and scene analysis applications. In this work, we present a new approach to efficiently sample the particles in a particle filter framework by integrating the behavior of interacting targets.

Particle filtering for visual object tracking is a well-established method, which can handle non-linear/non-Gaussian dynamics and observation models of the target [4]. The choice of the importance function, which guides the placement of the samples, is crucial for the performance of the particle filter. Conventional particle filters use the motion model of the target as the importance function. This choice requires the motion model to represent the dynamics of the target properly. In practice, these assumptions may not be valid due to a low frame rate of the video sequence, which lets the motion of the targets appear abruptly. Other sampling schemes that aim to lead the samples to promising regions of the sample space are found in literature. For example, [10] proposes an adaptive sampling scheme, which requires a specification of a density grid to partition the 2D sample space. In the multiple target case, such methods are prone to distract the samples by false positive alarms from similar objects. The combination of swarm intelligence with particle filtering is introduced in [9]. This approach suffers from high computational cost, because the observation model needs to be evaluated multiple times for each particle until convergence.

Classical particle filter approaches to track multiple targets can be separated into two categories: Independent particle filters that run one instance of the particle filter for each target [1, 8] and joint particle filters that merge the states of all targets into one state vector [6]. The latter is not practicable, because the state space grows exponentially in the number of targets to be tracked, leading to an infeasible computational complexity [5]. Therefore, we use one particle filter for each target.

It can be observed that adjacent targets may affect the motion of each other. Recent work showed that incorporating additional knowledge about the target context improves the tracking performance [3].

Our method consists of two parts: First, we introduce an adaptive importance sampling algorithm. We propose a new strategy to handle abrupt motion by consecutively increasing the noise of the motion model based on current measurements. Second, we include the dynamics of adjacent targets into the particle filter. For this, current measurements of neighboring targets are integrated into the sampling process. Our proposed method comes at almost no additional computational cost, because no extra measurements are required.

2 Particle Filter Basics

The problem of tracking can be defined as estimating the state of the target for every time step based on all available measurements up to this time step. Particle filtering for visual tracking has been introduced by Isard and Blake with the Condensation algorithm [4]. The key idea of particle filtering is to represent the posterior $p(\mathbf{x}_t|\mathbf{y}_{1:t})$ of the state vector $\mathbf{x}_t$ at time step t given the sequence of measurements $\mathbf{y}_{1:t} = (\mathbf{y}_1, \ldots, \mathbf{y}_t)$ by a set of particles. At each time step, new samples $\{\mathbf{x}_t^i, i = 1, \ldots, N\}$ are randomly drawn from the importance function. The associated weights $\{w_t^i, i = 1, \ldots, N\}$ are computed by the principle of sequential importance sampling [2] and normalized such that $\sum_{i=1}^N w_t^i = 1$. The computation of the weights involves the evaluation of the likelihood function $p(\mathbf{y}_t|\mathbf{x}_t^i)$, which is given by the observation model. Conveniently, the importance function is chosen to be the prior $p(\mathbf{x}_t|\mathbf{x}_{t-1})$, which describes the motion model of the target. The approximation of the posterior at time step t is given as

$$p(\mathbf{x}_t|\mathbf{y}_{1:t}) \approx \sum_{i=1}^{N} \delta(\mathbf{x}_t - \mathbf{x}_t^i) w_t^i. \qquad (1)$$

After some iterations of the particle filter algorithm, all but one particle will have negligible weights. To overcome this degeneracy problem, a resampling procedure is applied, which aims to replace the set of weighted particles by a set of uniformly weighted particles. The resampled set is generated by simulating new samples according to the distribution of the old sample weights.

If resampling is applied at every time step and the prior is used as the importance function, the weights are determined by evaluating the likelihood function at that particular sample position:

$$w_t^i \propto p(\mathbf{y}_t|\mathbf{x}_t^i). \tag{2}$$

The choice of the importance function plays an important role for the performance of particle filter trackers. The approximation of the posterior will give poor results, if the importance function places the particles in the tails of the likelihood function. The "optimal" importance function is given by $p(\mathbf{x}_t|\mathbf{x}_{t-1}^i, \mathbf{y}_t)$ [2]. However, in most cases this function can not be sampled from and other strategies are required to integrate the current measurements into the sampling process.

3 Adaptive Importance Sampling

In this section, we introduce a new method to adapt the motion model during the sampling process. Our approach is based on the motion model of the target as the importance function.

3.1 Adaptive Noise of the Motion Model

The motion model of the target can either be described by the prior $p(\mathbf{x}_t|\mathbf{x}_{t-1})$ or as a function of the state $\mathbf{x}_{t-1}$ and the process noise $\mathbf{u}_{t-1}$:

$$\mathbf{x}_t = f(\mathbf{x}_{t-1}, \mathbf{u}_{t-1}) \tag{3}$$

The process noise $\mathbf{u}_{t-1}$ determines the amount of diffusion by which the particles are distributed across the search space. Clearly, a high process noise enables the particles to capture abrupt motions of the target at the risk of being confused by false positive alarms. On the other hand, a low process noise concentrates the particles at the position predicted by the motion model of the target. For the sake of simplicity, in the following we assume that the noise follows a Gaussian distribution, $\mathbf{u}_{t-1} \sim \mathcal{N}(\mu, \sigma^2)$. We describe our sampling method for the single target case. The extension to multiple targets is straightforward.

In order to integrate current measurements into the sampling process, we consecutively sample the particles and measure each particle immediately to obtain its weight w_t^i from (2). Without loss of generality, it is assumed that w_t^i ranges between 0 and 1. Starting with zero noise, the noise increases monotonically until it reaches its maximum value $\sigma_{\max}$, which is defined by physical constraints on the motion of the target. The amount by which the noise increases depends on the weight w_t^i of the previously sampled particle, the previous noise standard deviation σ^i and the number of steps i:

$$\sigma^{i+1} = \sigma^i + \frac{1}{N-i}(1 - w_t^i)(\sigma_{\max} - \sigma^i) \tag{4}$$

The term $1/(N-i)$ ensures, that σ increases slowly in the beginning of the particle sampling and approximates $\sigma_{\max}$ as i approaches the total number of particles N. As long as the particles obtain a high weight, the noise increases moderately, because in this case we assume that the target has been captured. This approach has the advantage that if the target lies near to the predicted position, the tracker is less likely to be confused with false positive alarms. Still, the noise is guaranteed to reach its maximum to account for the case that the predicted position contains a false positive alarm. If the particle weights are low, the noise increases strongly to allow the particles to spread out widely across the feasible search space and capturing the target. In this way, the noise adapts to the quality of the sampled particles, which leads to a more efficiently explored search space.

The particles are proceeded in descending order according to their weight (before the resampling process) in the previous time step. If particles with a high weight represent the true target state and the motion model without noise reflects the true motion of the target, then their propagation according to the motion model (3) will produce again highly weighted particles.

3.2 Context-Aware Particle Sampling

The motion of one target normally influences the motion of adjacent targets. More precisely, the dynamics of the respective target are constrained by the position and movement of adjacent targets to avoid their collision. To incorporate this context knowledge into the particle filter, we consider the measurements of neighboring tracked targets during the sampling process of the current target. Thus, we sample the particles in rounds, where each round comprises the sampling of one particle for each target. Again, the particles are weighted immediately after sampling. We assume that the spatial configuration of neighboring targets is constant between two consecutive video frames.

Let $\hat{\mathbf{x}}_{t-1,k}$ and $\hat{\mathbf{x}}_{t-1,l}$ denote the estimated positions of the current target k and its neighbor l in the previous time step. In each round i, a highly weighted particle $\mathbf{x}_{t,l}^i$ may become a guide for the particle to be sampled of a neighboring target k in the next round. This guide is determined as

$$\mathbf{g}_{t,l}^{i+1} = \mathbf{x}_{t,l}^i + (\hat{\mathbf{x}}_{t-1,k} - \hat{\mathbf{x}}_{t-1,l}). \tag{5}$$

Propagating the current particle position by the motion model, the resulting position is defined as

$$\mathbf{b}_{t,k}^i = f(\mathbf{x}_{t-1,k}^i, \mathbf{u}_{t-1,k}). \tag{6}$$

For the final sampled particle position $\mathbf{x}_{t,k}^i$, the guides from all neighboring targets and the motion model of the respective target are considered by a weighted combination of both (Figure 1):

$$\mathbf{x}_{t,k}^i = \mathbf{b}_{t,k}^i + \left(\sum_l s_l(\mathbf{g}_{t-1,l}^i - \mathbf{b}_{t,k}^i)\right)/M, \tag{7}$$

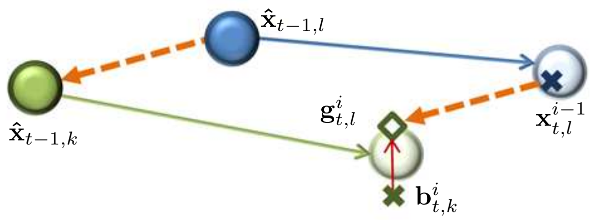


Figure 1. Sampling of the particle for the current target k using the guide from neighbor l.

Algorithm 1 Adaptive Importance Sampling

$\sigma^0 \leftarrow 0$
for $i = 1$ to N **do** ▷ rounds
  for $k = 1$ to K **do** ▷ targets
    $M \leftarrow 0$ ▷ number of guides
    for all neighbors l of k **do**
      if $w_{t,l}^{i-1}$ is large enough **then**
        • $\mathbf{g}_{t,l}^i \leftarrow \mathbf{x}_{t,l}^{i-1} + (\hat{\mathbf{x}}_{t-1,k} - \hat{\mathbf{x}}_{t-1,l})$
        • Compute the weight s_l
        • $M \leftarrow M + 1$
      end if
    end for
    • $\mathbf{b}_{t,k}^i \leftarrow f(\mathbf{x}_{t-1,k}^i, \mathbf{u}_{t-1,k})$,
      $\mathbf{u}_{t-1,k} \sim \mathcal{N}(\mu, {\sigma_k^i}^2)$
    • $\mathbf{x}_{t,k}^i \leftarrow \mathbf{b}_{t,k}^i + \left(\sum_l s_l(\mathbf{g}_{t-1,l}^i - \mathbf{b}_{t,k}^i)\right)/M$
    • Calculate the weight $w_{t,k}^i$ according to (2)
    • $\sigma_k^{i+1} \leftarrow \sigma_k^i + \frac{1}{N-i}(1 - w_{t,k}^i)(\sigma_{\max} - \sigma_k^i)$
  end for
end for

where M denotes the number of guides. The weights s_l determine how far the particle sampled by the conventional motion model (6) should be shifted toward the guided position. Typically, these weights depend on the weight $w_{t,l}^{i-1}$ of the particle that became a guide in the previous round. We model this property as a sigmoid function (Figure 3).

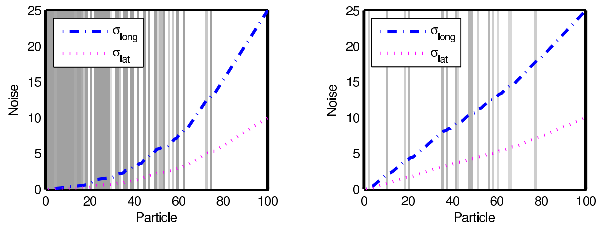
The definition of neighboring targets typically includes the distance to the current target and the relative velocity in the last time step, if collision of the targets should be avoided.

An algorithm that contains both the adaptive noise strategy and the particle sampling method guided by neighboring targets is shown in Algorithm 1. Note that no additional evaluations of the likelihood function are necessary.

4 Experimental Results

We conducted experiments on real data to show the effectiveness of the adaptive importance sampling algorithm. Our dataset contains image sequences showing moving vehicles in a complex inner city environment. Because of the low framerate (2 fps), abrupt motion is likely to occur. Our basic particle filter tracking framework is adopted from [7]. We use a constant velocity model for the motion model, where the noise represents the longitudinal and lateral acceleration of the vehicle. The observation model contains both a shape based matching and a color histogram matching. The parameter estimation is done by a weighted mean shift clustering algorithm. For all experiments, the number of particles for each vehicle is set to 100.

First, we demonstrate the efficacy of the adaptive noise strategy of the motion model. For a single target, Figure 2 shows how the noise adaptively increases depending on the quality of the sampled particles. In the first case (Figure 2(a)), high weights of early sampled particles indicate the capture of the target. Therefore, the noise is slowly increasing to lower the risk of the particles capturing a false positive alarm. In the second case (Figure 2(b)), the early sampled particles feature low weights. Apparently, the target is not located around the position predicted by the motion model and the noise increases strongly to cover a wider area of the search space. Compared to the first case, the noise is twice as high after the first half of the particles has been sampled. This strategy enables the sampling process to finally hit the target and successfully continue tracking.



(a) (b)

Figure 2. Increasing noise of the motion model. Dark regions represent high values of the particle weight.

Second, we prove that our context-aware particle sampling method can increase the robustness of the tracker. In this experiment, the image sequence contains bridges where the vehicles are totally occluded for up to two frames. The adaptive noise of the motion model as in section 3.1 is used. The sigmoid function for computing the weights s_l is defined as

$$s_l = 1/\left(1.1 + \exp(-25 w_{t,l}^{i-1} + 15)\right), \qquad (8)$$

where the parameters are determined experimentally (Figure 3).

Without using the guiding from adjacent vehicles, the tracker loses one target (Figure 4, left, cf. red trajectory). When the positions of adjacent vehicles are taken into consideration during the sampling process, the tracker successfully maintains all trajectories (Figure 4, right). Using the context-aware sampling strategy, the tracker may be distracted by false positive alarms, though (Figure 4, right, cf. green trajectory). However, when the vehicle is visible again, the tracker correctly assigns the trajectory.

Finally, we evaluate the performance of our tracker with an image sequence showing a number of vehicles under poor illumination conditions. In this example, both the adaptive noise strategy of the motion model and the context-aware sampling are used. Figure 5 shows that all vehicles are tracked successfully during this sequence.

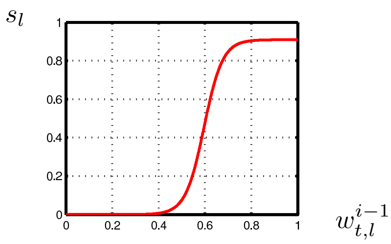


Figure 3. Weight s_l as a sigmoid function of the particle weight.

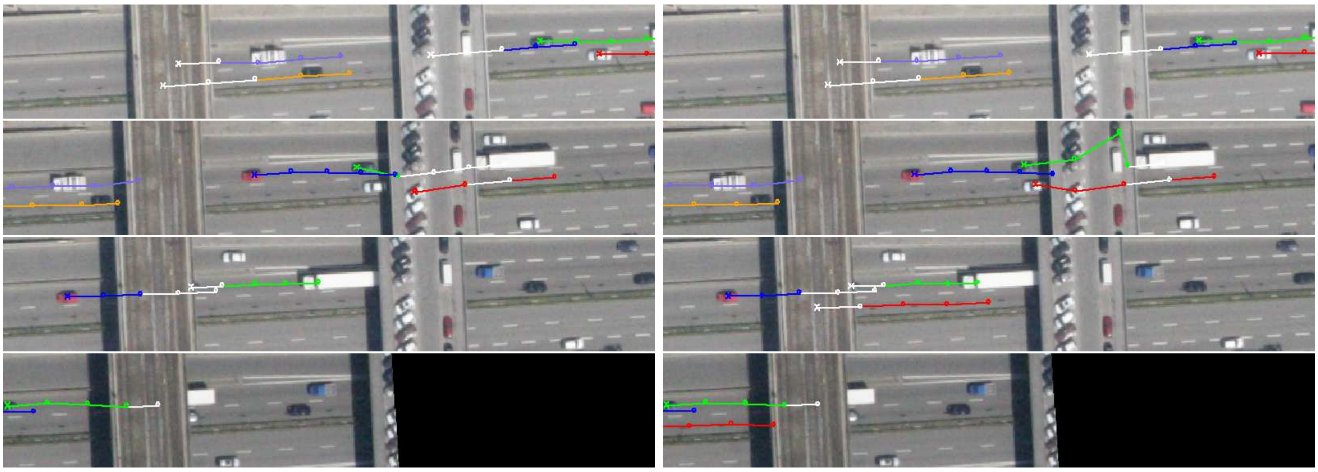


Figure 4. Test sequence with 2 fps showing frames 4, 9, 14, 19 in full resolution. Left column: without context-aware sampling. Right column: With context-aware sampling. The current vehicle position together with the recent trajectory is shown. Extrapolated trajectories are displayed in white.

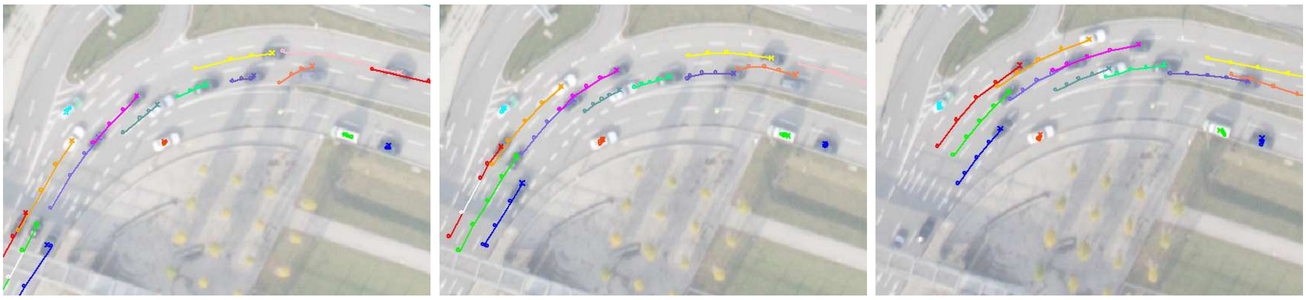


Figure 5. Test sequence with 2 fps showing frames 3, 6 and 10 in full resolution. The current vehicle position together with the recent trajectory is shown. Extrapolated trajectories are displayed in white.

5 Conclusions

In this paper, we developed a new method to adaptively sample the particles based on the motion model of the target. Experiments showed that our particle filter tracker enables the particles to explore the search space more efficiently. For tracking multiple targets, we introduced a context-aware sampling method, which considers the motion of adjacent targets during the particle sampling process of the current target. This strategy improves the robustness of the tracker, as experiments with totally occluded targets showed.

In future work, we aim to develop models for handling complex spatial relations of adjacent targets. Further, we plan to perform statistical evaluations on large complex tracking scenarios.

Acknowledgment

This work was supported by the TUM International Graduate School of Science and Engineering (IGSSE) and the German Aerospace Center (DLR).

References

[1] Y. Cai, N. de Freitas, J. J. Little: "Robust Visual Tracking for Multiple Targets," *Proc. ECCV*, pp.107-118, 2006

[2] A. Doucet: "On Sequential Simulation-Based Methods for Bayesian Filtering," Tech. Report CUED/F-INFENG/TR. 310, Cambridge Univ. Dept. Eng., 1998

[3] H. Grabner, J. Matas, L. J. Van Gool, P. C. Cattin: "Tracking the Invisible: Learning Where the Object Might Be," *Proc. CVPR*, pp.1285-1292, 2010

[4] M. Isard, A. Blake: "CONDENSATION - Conditional Density Propagation for Visual Tracking," *Int. Journal of Computer Vision*, vol.29, pp.5-28, 1998

[5] Z. Khan, T. Balch, F. Dellaert: "MCMC-Based Particle Filtering for Tracking a Variable Number of Interacting Targets," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.27, pp.1805-1918, 2005

[6] C. Rasmussen, G. D. Hager: "Probabilistic Data Association Methods for Tracking Complex Visual Objects," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.23, no.6, pp.560-576, 2001

[7] I. Szottka, M. Butenuth: "Tracking Multiple Vehicles in Airborne Image Sequences of Complex Urban Environments," *Proc. IEEE Joint Urban Remote Sensing Event*, pp.13-16, 2011

[8] C. Yang, R. Duraiswami, L. Davis: "Fast Multiple Object Tracking via a Hierarchical Particle Filter," *Proc. ICCV*, pp.212-219, 2005

[9] X. Zhang, W. Hu, and S. Maybank: "A Smarter Particle Filter," *LNCS, Computer Vision - ACCV 2009*, vol.5995, pp.236-246, 2010

[10] X. Zhou, Y. Lu: "Abrupt Motion Tracking via Adaptive Stochastic Approximation Monte Carlo Sampling," *Proc. CVPR*, pp.1847-1854, 2010