MIXED-STATE PARTICLE FILTERS FOR MULTIASPECT TARGET TRACKING IN IMAGE SEQUENCES

Marcelo G. S. Bruno

Instituto Tecnológico de Aeronáutica
São José dos Campos SP 12228-900, Brazil
ph:(55-12) 3947-5878; e-mail:bruno@ele.ita.br

ABSTRACT

We introduce in this paper new mixed-state particle filter algorithms for direct target tracking in image sequences in a scenario where the true target template is unknown and changes randomly from frame to frame. We present two versions of the mixed-state particle filter tracker using respectively the sampling/importance resampling (SIR) technique and the alternative auxiliary particle filter (APF) method. Monte Carlo simulation results with heavily cluttered image sequences generated from real infrared airborne radar (IRAR) data show that the proposed algorithms have good performance and compare favorably to an alternative grid-based HMM filter by yielding similar steady-state root mean-square error (RMSE) at a much lower computational cost.

1. INTRODUCTION

Conventional target trackers [1] based on the suboptimal association of image correlation filters and linearized Kalman-Bucy tracking filters (KBfs) are known [2] to perform poorly in scenarios of heavily obscured targets and unknown, randomly-changing target aspect. To overcome the limitations of the correlation filter/KBf association, we propose in this paper an alternative Bayesian approach to multiaspect target tracking using particle filters [3]. The proposed algorithms enable direct tracking from the image sequence fully incorporating the statistical models for target motion, target aspect, and background clutter correlation.

We assume a real-valued kinematic state vector that collects the position and velocity of the target centroid in the two dimensions of the plane. The target aspect state is on the other hand discrete-valued and defined on a finite library of possible target models. The optimal minimum mean-square error (MMSE) estimate of the kinematic state vector at each frame is approximated then using a mixed-state particle filter that automatically takes into account the target aspect changes from frame to frame. We introduce two versions of the proposed mixed-state particle filter tracker using respectively the sampling/importance resampling (SIR) [4] technique and the alternative auxiliary particle filter (APF) [5] method. We incorporate the target signature and the background clutter correlation models into the design of the tracker using the likelihood function model introduced in [2].

The paper is divided into 5 sections. Section 1 is this introduction. In section 2, we present the target motion, target aspect, background clutter and likelihood function models that underly the derivation of the proposed algorithms. In section 3 , we present the SIR and APF mixed-state trackers. In section 4, we examine the performance of the proposed algorithms and compare them to the point-mass HMM tracker previously introduced in [6]. Finally, we present our conclusions in section 6.

This work was partially supported by FAPESP, Brazil.

2. PROBLEM FORMULATION

Let $\mathbf{x}_n$ be a four-dimensional, real-valued kinematic state vector that collects at instant n the position and velocity of the target centroid in the two dimensions of the plane, denoted respectively dimension $i = 1$ and $i = 2$. Let s_n be a discrete-valued target aspect state defined on the finite set $\mathcal{I} = \{1, 2, \ldots, K\}$ where each state $s_n = l$ represents a pointer to one of K possible target template models corresponding to rotated, scaled and/or sheared versions of the target's mother template. Assuming for simplicity that the sequences $\{\mathbf{x}_n\}$ and $\{s_n\}$ are statistically independent for $n \geq 0$, we model the sequence $\{\mathbf{x}_n\}$ as a real-valued, first-order Markov random sequence specified by the transition probability density function (pdf) $p(\mathbf{x}_n \mid \mathbf{x}_{n-1})$, and by the pdf of the initial kinematic state $p(\mathbf{x}_0)$. The sequence of aspect states $\{s_n\}$ is in turn modeled as a discrete-valued, first-order Markov chain defined on $\mathcal{I}$ and specified by the transition probability mass function $P(s_n = k \mid s_{n-1} = l)$, $(k, l) \in \mathcal{I} \times \mathcal{I}$, and by the initial probability mass function $P(s_0 = l), l \in \mathcal{I}$.

2.1. Observation Model

The nth frame in the digital image sequence generated from the raw sensor measurements is modeled as the $L \times M$ matrix

$$\mathbf{Y}_n = \mathbf{H}(\mathbf{x}_n^*, s_n) + \mathbf{V}_n \quad (1)$$

where matrix $\mathbf{V}_n$ represents the background clutter, and matrix $\mathbf{H}(\mathbf{x}_n^*, s_n)$ is the clutter-free target image model, which is a function of the 2D pixel location of the target centroid, $\mathbf{x}_n^*$ and the target aspect state, s_n. The two-dimensional random vector $\mathbf{x}_n^*$ takes values on the finite image grid $\mathcal{L} = \{(r, j) \mid 1 \leq r \leq L,\ 1 \leq j \leq M\}$ and is obtained from the four-dimensional continuous-valued state vector $\mathbf{x}_n$ by quantizing the real-valued position coordinates in dimensions $i = 1$ and $i = 2$ to the image grid.

Clutter-Free Target Signature Model For each pixel centroid position $\mathbf{x}_n^* = (r_n, j_n) \in \mathcal{L}$, the nonlinear function $\mathbf{H}$ in (1) returns a spatial distribution of (real-valued) pixel intensities $\{a_{k,l}(s_n)\}$, $-r_i \leq k \leq r_s$, $-l_i \leq l \leq l_s$, centered at (r_n, j_n) and dependent on the aspect state s_n.

Clutter Model The random clutter returns at frame n, $V_n(r, j)$, $1 \leq r \leq L, 1 \leq j \leq M$, are described by the 2D first-order, noncausal Gaùss-Markov random field (GMrf) model [7]

$$\begin{aligned} V_n(r, j) &= \beta_v^c \left[V_n(r-1, j) + V_n(r+1, j)\right] \\ &+ \ \beta_h^c \left[V_n(r, j-1) + V_n(r, j+1)\right] + \varepsilon_n(r, j) \quad (2) \end{aligned}$$

where $E\left[V_n(r, j)\, \varepsilon_n(k, l)\right] = \sigma_c^2\, \delta_{r-k,\, j-l}$. The unknown parameters β_h, β_v and σ_c^2 are estimated from the data using a simplified approximate maximum likelihood (AML) estimator, see [7] for details. The assumption of zero-mean clutter implies a pre-processing of the data that subtracts the mean of the background. We also assume that, after the subtraction of the mean, the clutter frames $\mathbf{V}_n$ and $\mathbf{V}_m$ are independent for $m \neq n$.

Likelihood Function Let $\mathbf{y}_n$ be a 1D long vector representation of the image frame $\mathbf{Y}_n$. Assuming a GMrf background as in (2) and assuming that the signature parameters $\{a_{k,l}\}$ are deterministic and known, we use the results in [2] to write the likelihood function of the hidden states $\mathbf{x}_n$ and s_n at instant n as

$$p(\mathbf{y}_n \mid \mathbf{x}_n, s_n) \infty \exp\left[\frac{2\lambda(\mathbf{x}_n, s_n) - \rho(s_n)}{2\sigma_c^2}\right]. \quad (3)$$

In (3), the symbol ∞ denotes "proportional to", $\rho(s_n)$ is a target energy term that depends on the aspect state, but that does not vary with $\mathbf{x}_n$ away from the image borders, see [2] for further details, and $\lambda(\mathbf{x}_n, s_n)$ is a data-dependent term such that

$$\lambda(\mathbf{x}_n, s_n) = \sum_{k=-r_i}^{r_s} \sum_{l=-l_i}^{l_s} a_{k,l}(s_n) d(x_n^*(1) + k,\ x_n^*(2) + l) \quad (4)$$

where $x_n^*(i)$, $i = 1, 2$, is, as explained before, the quantized centroid coordinate in dimension i, and $d(r, j)$ is the output of the linear differential operator

$$\begin{aligned} d(r,j) &= Y_n(r,j) - \beta_h^c \left[Y_n(r, j-1) + Y_n(r, j+1)\right] \\ &- \ \beta_v^c \left[Y_n(r-1, j) + Y_n(r+1, j)\right] \quad (5) \end{aligned}$$

with Dirichlet (identically zero) boundary conditions. Equation (4) is valid for $r_i + 1 \leq x_n^*(1) \leq L - r_s$ and $l_i + 1 \leq x_n^*(2) \leq M - l_s$. For centroid positions close to the image borders, the summation limits in (4) must be changed accordingly.

3. MIXED-STATE PARTICLE FILTER TRACKERS

We use a recursive Monte Carlo simulation strategy known as sequential importance sampling (SIS) [3] to represent the mixed posterior $p(\mathbf{x}_n, s_n \mid \mathbf{Y}_1^n)$ by a properly weighted set of particles $\left\{\mathbf{x}_n^{(j)}, s_n^{(j)}\right\}$, $1 \leq j \leq N_p$, drawn from an alternative mixed importance function. The minimum mean-square error (MMSE) estimate of the kinematic state $\mathbf{x}_n$ at instant n is approximated then by the weighted average of the corresponding samples $\left\{\mathbf{x}_n^{(j)}\right\}$.

Sampling/Importance Resampling (SIR) Tracker Let $\mathbf{X}_n = \left[\mathbf{x}_n^T\ s_n\right]^T$ be the extended state vector at frame n. A simple SIS strategy [8] is to use the Markovian transition kernel $p(\mathbf{X}_n \mid \mathbf{X}_{n-1})$ to generate the particle set at instant n and then update the particle weights [3, 8] using the likelihood function $p(\mathbf{y}_n \mid \mathbf{X}_n)$.

A practical problem associated with the use of sequential SIS filters is however that, as the number of iterations increase, the distribution of particle weights may get increasingly skewed, resulting in particle degeneracy. The sampling/importance resampling (SIR) [4] technique attempts to mitigate this problem by introducing an additional particle selection step that consists of resampling from the original particle set with replacement according to the importance weights. After the selection step, the particle weights are all reset to $1/N_p$. Using the GMRf-based likelihood function model from section 2.1, we summarize in Table 1 the SIR filter for direct, multiaspect target tracking from image sequences in the special case when we add the simplifying assumption of statistical independence between target motion and target aspect such that we can generate the particle sets $\left\{\mathbf{x}_n^{(j)}\right\}$ and $\left\{s_n^{(j)}\right\}$ by sampling independently from $p(\mathbf{x}_n \mid \mathbf{x}_{n-1}^{(j)})$ and $P(s_n \mid s_{n-1}^{(j)})$.

Auxiliary Particle Filter (APF) Tracker The SIR tracker described before has the disadvantage of sampling blindly from $p(\mathbf{X}_n \mid \mathbf{X}_{n-1})$ ignoring the current observation $\mathbf{y}_n$. An alternative strategy to reduce the sensitivity of the tracker to outliers is to use auxiliary particle filtering (APF) [5] In intuitive terms, APF can be viewed as resampling (or pre-selecting) the particles at instant $n-1$ based on the likelihood of certain point estimates $\widehat{\mathbf{X}}_n^{(j)}$ that represent $p(\mathbf{X}_n \mid \mathbf{X}_{n-1}^{(j)})$. Formally, that is accomplished by introducing an auxiliary integer variable k defined on the set $\{1, 2, \ldots, N_p\}$.

1. Initialization For $j=1,\ldots,N_p$
• Draw $\mathbf{x}_0^{(j)} \sim p(\mathbf{x}_0)$
and $s_0^{(j)} \sim P(s_0)$.
• Make $w_0^{(j)} = 1/N_p$ and set $n=1$.
2. Importance Sampling For $j=1,\ldots,N_p$
• Draw $\tilde{\mathbf{x}}_n^{(j)} \sim p(\mathbf{x}_n \mid \mathbf{x}_{n-1}^{(j)})$.
• Draw $\tilde{s}_n^{(j)} \sim P(s_n \mid s_{n-1}^{(j)})$.
• Compute the importance weights
$\tilde{w}_n^{(j)} \propto w_{n-1}^{(j)}\, p(\mathbf{y}_n \mid \tilde{\mathbf{x}}_n^{(j)}, \tilde{s}_n^{(j)}), \quad \sum_{j=1}^{N_p} \tilde{w}_n^{(j)} = 1$
using equations (3), (4), and (5).
3. Selection
• Generate a new set of samples
$\left\{\mathbf{x}_n^{(j)}, s_n^{(j)}\right\} \quad 1 \le j \le N_p$
such that $P(\left[\mathbf{x}_n^{(j)}, s_n^{(j)}\right] = \left[\tilde{\mathbf{x}}_n^{(k)}, \tilde{s}_n^{(k)}\right]) = \tilde{w}_n^{(k)}$.
• Make $w_n^{(j)} = 1/N_p, \ 1 \le j \le N_p$.
• Set $n = n+1$ and go back to step 2.

Table 1. Algorithm I: Mixed-State SIR Filter for Multiaspect Target Tracking in Images.

We sample then from the importance function

$$q(k, \mathbf{X}_n \mid \mathbf{y}_n) \propto w_{n-1}^{(k)}\, p(\mathbf{y}_n \mid \widehat{\mathbf{X}}_n^{(k)}) p(\mathbf{X}_n \mid \mathbf{X}_{n-1}^{(k)}) \ . \quad (6)$$

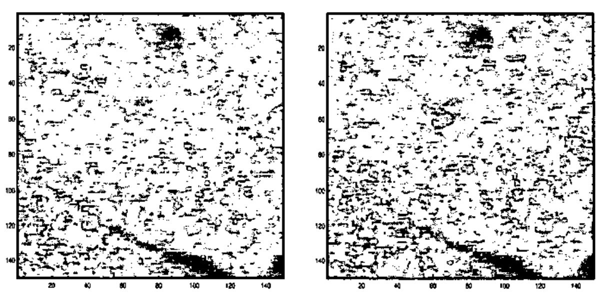
where $\widehat{\mathbf{X}}_n^{(j)}$ is e.g. the mean of or a draw from $p(\mathbf{X}_n \mid \mathbf{X}_{n-1}^{(j)})$. Table 2 summarizes the APF filter for multiaspect target tracking using the models developed in section 2 and assuming as before statistical independence between target motion and target aspect. The initialization step is identical to the same step in Table 1 and is omitted here for conciseness.

4. SIMULATION RESULTS

We examine in the sequel the performance of algorithms introduced in section 3 using a simulated image sequence generated from real infrared airborne radar (IRAR) data obtained from the Center for Imaging Sciences at Johns Hopkins University. The simulated background sequence consists of synthetic GMrf sample frames whose spatially variant local means and correlation parameters are estimated from a real IRAR base image. We add then to the background sequence a simulated moving target whose centroid position and velocity change from frame to frame according to a linear white noise acceleration motion model [1]. The unknown initial target position is uniformly distributed between pixels 20 and 60 in the vertical dimension and between pixels 20 and 40 in the horizontal dimension. The initial target velocity in both dimensions is a sample from a Gaussian random variable with mean $10m/s$ and standard

For n=1,2,...
Importance Sampling: for $j=1,\ldots N_p$
• Draw $\widehat{\mathbf{x}}_n^{(j)} \sim p(\mathbf{x}_n \mid \mathbf{x}_{n-1}^{(j)})$.
and $\widehat{s}_n^{(j)} \sim P(s_n \mid s_{n-1}^{(j)})$.
• Compute the first-stage weights
$\Lambda_n^{(j)} \propto w_{n-1}^{(j)} p(\mathbf{y}_n \mid \widehat{\mathbf{x}}_n^{(j)}, \widehat{s}_n^{(j)}) \quad \sum_{j=1}^{N_p} \Lambda_n^{(j)} = 1$
using equations (3), (4), and (5).
• Draw $k^{(j)} \sim \{1, 2, \ldots, N_p\}$
with $\left\{P(k^{(j)} = i) = \Lambda_n^{(i)}\right\}$, $i = 1,\ldots,N_p$.
• Draw $\mathbf{x}_n^{(j)} \sim p(\mathbf{x}_n \mid \mathbf{x}_{n-1}^{(k^{(j)})})$,
and $s_n^{(j)} \sim P(s_n \mid s_{n-1}^{(k^{(j)})})$.
• Compute the second-stage weights
$w_n^{(j)} \propto \frac{p(\mathbf{y}_n \mid \mathbf{x}_n^{(j)}, s_n^{(j)})}{p(\mathbf{y}_n \mid \widehat{\mathbf{x}}_n^{(k^{(j)})}, \widehat{s}_n^{(k^{(j)})})} \qquad \sum_{j=1}^{N_p} w_n^{(j)} = 1$
using equations (3), (4), and (5).

Table 2. Algorithm II: Mixed-State Auxiliary Particle Filter for Multiaspect Target Tracking in Images.

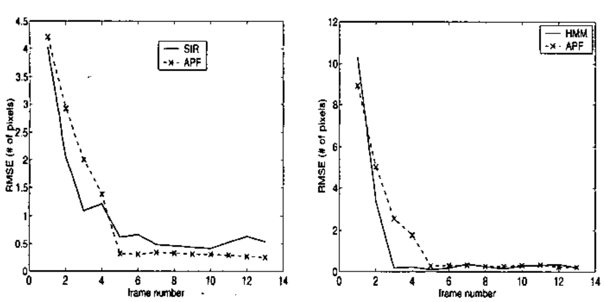


(a) (b)

Fig. 1. Cluttered target sequence, PTCR=3.6 dB: (a) first frame, (b) tenth frame with random target translation, rotation, scaling, and shearing.

deviation $\sigma = 0.1m/s$. The random target aspect is initialized with an unknown template model chosen uniformly from the template library $\mathcal{I}$ and is subsequently changed from frame to frame according to a discrete Markov chain model with five states and a 40 % probability of transition to an adjacent state. The target pixel intensity is time-invariant and known and was set according to a desired low level of contrast between the template and the background. Figures 1 (a) and (b) show two simulated frames, respectively at instants $n=0$ and $n=9$, with peak target-to-clutter ratio (PTCR) equal to 3.6 dB. We tracked the simulated target over 13 consecutive frames using respectively the SIR filter in Table 1, the APF filter in Table 2, and the grid-based multiaspect HMM tracker introduced in [6]. Both the SIR filter and the APF use $N_p = 5{,}000$ particles. Figure 2(a) shows the root mean-square error (RMSE) of the horizontal position estimate for the SIR (solid line) and APF (dashed line) trackers in a scenario where the PTCR was lowered to -3.6 dB. Both trackers diverged in 5 out of 100 Monte Carlo runs. The error curves in Figure 2 were obtained excluding the divergent tracks from the average. We see from the plots that, excluding the rare occasions when the particle filters diverge, both algorithms show good performance quickly acquiring the target after an initial error and converging to a low steady-state RMSE. The APF algorithm seems to outperform the SIR tracker slightly in terms of steady-state error.

Figure 2(b) shows the RMSE in number of pixels this time for the vertical position estimates obtained by the HMM (solid line) and APF (dashed line) trackers. We see from the plots in Figure 2(b) that the auxiliary particle filter has a slower target acquisition time, but eventually converges to a steady-state RMSE that is indistinguishable from the error of the HMM filter. The HMM tracker is however computationally expensive requiring at each frame, see [6], the evaluation the likelihood function in all points of the image grid for all aspect states. For an $L \times L$ grid, that means a computational cost of order $O(\alpha L^2)$ where usually $\alpha << L$. By contrast, each iteration of the particle filter trackers requires the evaluation of the likelihood function only for the N_p samples in the current particle set, thus reducing the computational cost to $O(\beta N_p)$. If $N_p << L^2$, the computational savings may be considerable. Overall, our simulations suggest that the particle filter trackers compare favourably to the grid-based HMM tracker in [6] by yielding similar RMSE performance roughly 95 % of the time, but at a much lower computational cost.



(a) (b)

Fig. 2. (a) Horizontal coordinate estimate RMSE for the SIR filter (solid) and the Auxiliary Particle Filter (dashed), PTCR= -3.6 dB; (b) Vertical coordinate estimate RMSE for the HMM filter (solid) and the Auxiliary Particle Filter (dashed), PTCR=-3.6 dB.

5. CONCLUSIONS

We introduced in this paper a new Bayesian framework for multiaspect target tracking in image sequences using mixed-state particle filters that fully incorporate the statistical models for target motion, target aspect and background clutter correlation. Two versions of the mixed-state particle filter tracker were introduced, using respectively the sampling/importance resampling (SIR) technique and the alternative auxiliary particle filter (APF) method. Monte Carlo simulation results with heavily cluttered image sequences generated from real infrared airborne radar (IRAR) data show that the proposed algorithms have good performance and compare favorably to an alternative grid-based HMM filter by yielding similar steady-state root mean-square error (RMSE) at a much lower computational cost.

6. REFERENCES

[1] Y. Bar-Shalom and X. Li, *Multitarget-Multisensor Tracking: Principles and Techniques.* YBS, Storrs, CT, 1995.

[2] M. G. S. Bruno and J. M. F. Moura, "Multiframe detection/tracking in clutter: optimal performance," *IEEE Transactions on Aerospace and Electronic Systems*, vol.37, n.3, pp. 925-946, July 2001.

[3] M. S. Arulampalam, S. Maskell, N. Gordon and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. on Signal Processing*, vol.50, n.2, pp. 174-188, February 2002.

[4] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *IEE Proceedings-F*, vol.140, n.2, pp. 107-113, 1993.

[5] M. K. Pitt and N. Shephard, "Filtering via simulation: Auxiliary particle filters," *Journal of the American Statistical Association*, 94(446), pp. 590-599, 1999.

[6] M. G. S. Bruno and J. M. F. Moura, "Clutter Adaptive Tracking of Multiaspect Targets in IRAR Imagery," in *Proceedings IEEE ICASSP 2002*, Orlando FL, May 2002.

[7] J. M. F. Moura and N. Balram, "Noncausal Gauss-Markov random fields: parameter structure and estimation," *IEEE Transactions on Information Theory*, IT-39(4), pp. 1333-1355, July 1993.

[8] M. Isard and A. Blake, "A mixed-state condensation tracker with automatic model switching," *Proceedings of the 6th International Conference on Computer Vision* 1998, pp. 107-112.