Comparison of Two Particle Filter Methods
For Image Plane Tracking of Multiple Objects

Ernest J. Ohlmeyer¹
Aero Science Applications LLC, King George, VA 22485

and

P. K. Menon²
Optimal Synthesis Inc., Los Altos, CA 94022

Abstract

This paper examines the capabilities of the Particle Filter (PF) for the tracking of multiple objects moving on a focal plane array image plane. The study considers the tracking of four objects for which the dynamics are near constant velocity with random perturbations in acceleration. The filtering structure uses four separate PFs rather than a concatenated filter with four sets of target states. There is a data association problem in which each measurement must be assigned to the proper PF. A new association logic based on weighted nearest neighbor strategy is employed to solve this problem. Two new methods are considered for calculating the PF importance weights. The first is based on a weighted distance between the PF position estimate and the perceived location of maximum measurement intensity. The second method employs a likelihood ratio to compute the weights. The first method performed well at a signal-to-noise ratio (SNR) of 15 dB or higher but was unsatisfactory below this. The second method was found to perform well at very low SNR approaching 0 dB. The paper discusses the multi-object tracking scenario, describes the association logic and PF implementations, discusses the analysis details, and shows the performance obtained from the PF.

¹ Principal Scientist, Associate Fellow AIAA.
² Chief Scientist and President, Fellow AIAA.
I. Introduction

The Particle Filter (PF) is a newly-emerged nonlinear estimation technique based on Bayesian Sequential Monte Carlo Methods which has been found to be very effective in difficult tracking problems where traditional filters such as the Extended Kalman Filter (EKF) break down. Although related ideas have been in the literature since the 1950’s, particle filtering is generally considered to have emerged as a new paradigm in nonlinear estimation in 1993 with the publication of [1]. Since then, PF research has seen explosive growth. The algorithms for implementing a PF have been described extensively in the literature and will not be repeated here. See [1-10] for more detailed and complete descriptions.

The basic premise of the PF is to represent the evolving pdf by a fairly large set of random samples, or particles. The particles are stochastic realizations of the state vector. The particle set is propagated in time and updated with new measurements in the familiar predictor-corrector cycle. However, a re-sampling step is usually required during the update [1]. While powerful, the challenge of implementing an estimator such as the PF is to contain the computational complexity while maintaining accuracy so as to permit on-line, real-time processing. The basic version of the PF is simple to implement and often very effective. However, for some difficult tracking problems, a very large set of particles may be required for satisfactory performance. This can occur when most of the prior samples are rejected and just a small subset of the population is re-sampled many times, leading to a phenomenon known as degeneracy. Techniques for dealing with this problem are discussed in the references cited previously.

The authors of the present paper have been actively involved in the development of advanced nonlinear estimation techniques for target tracking applications over the past decade [11-16]. The current research is focused on using the PF for tracking of multiple objects (such as a primary target in the presence of one or more secondary objects, sometimes in a background of clutter). These can occur frequently in military scenarios involving target tracking with infrared imaging or radar sensor systems. Further treatment of this advanced PF multi-object tracking topic is given in [17-29].

II. Image Tracking Scenario

We consider the tracking of multiple, dynamic targets as they migrate across the image plane. A set of four targets was chosen for analysis and their motions are illustrated in Fig. 1. The targets are initially located in the corners of the array, converge toward the center with some crossing each other, and then continue on to the array edges. Also shown are the target velocity and heading angle histories. The array dimensions are 256 x 256 pixels and the maximum scenario time is 10 s. The targets experience random normal and longitudinal accelerations which introduce perturbations to the baseline constant velocity, straight line motion. The
locations of the four targets at selected time points are shown superimposed on the array’s background noise for a 15 dB SNR in Fig. 2.

![Fig. 1. Assumed Target Motions in Image Plane](image1)

Fig. 1. Assumed Target Motions in Image Plane

![Fig. 2. Target Locations in Measurement Plane at Selected Times (SNR = 15 dB)](image2)

Fig. 2. Target Locations in Measurement Plane at Selected Times (SNR = 15 dB)

The true target dynamics are shown in Fig. 3 and consist of a perturbed constant acceleration model. States are x and y position, total velocity (V) and heading angle (γ), and longitudinal (aL) and normal (aN) accelerations with respect to the velocity vector. The jerk, or rate of change of acceleration, is modeled as random Gaussian noise with standard deviations of \( \sigma_L = 0.5 \) pixels/s².
and $\sigma_N = 1.0$ pixels/s$^2$ at a time step of $\Delta t = 0.1$ s. The tracking is performed by four separate particle filters where the model used in each filter is given in Fig. 4. In this case, each filter has four states consisting of $x$, $y$ positions and velocities in Cartesian coordinates. This nearly constant velocity model is driven by random $x$, $y$ accelerations ($w$) with a standard deviation of $\sigma_w = 0.5$ pixels/s$^2$ at $\Delta t = 0.1$ s.

\[
\begin{bmatrix}
\dot{x} \\
\dot{y} \\
\dot{V} \\
\dot{\gamma} \\
\dot{a}_L \\
\dot{a}_N
\end{bmatrix} = \begin{bmatrix}
V \cos(\gamma) \\
V \sin(\gamma) \\
a_L \\
a_N/V \\
0 \\
0
\end{bmatrix} + \begin{bmatrix}
0 & 0 & 0 & 0 & w_L \\
0 & 0 & 0 & 0 & w_N
\end{bmatrix}
\]

\[
\dot{X} = f(X) + BW(t)
\]

![Fig. 3. Target Truth Model](image)

\[
\begin{bmatrix} x \\ y \\ v_x \\ v_y \end{bmatrix} = \begin{bmatrix} r^T \\ v^T \end{bmatrix}
\]

\[
\begin{bmatrix} r_{k+1} \\ v_{k+1} \end{bmatrix} = \begin{bmatrix} I_2 & T \cdot I_2 \\ 0_2 & I_2 \end{bmatrix} \begin{bmatrix} r_k \\ v_k \end{bmatrix} + \begin{bmatrix} \frac{1}{2} T^2 \cdot I_2 \\ T \cdot I_2 \end{bmatrix} w_k
\]

\[
X_{k+1} = \Phi X_k + \Gamma w_k
\]

![Fig. 4. Filter Model](image)

The measurement set $Z$ consists of a 256 x 256 set of measured intensities in the image plane. The distortion of the measured intensities is modeled by a Gaussian point spread function with blur parameter $\Sigma = 2$ pixels and nominal amplitude $A = 1$. In the analysis, the signal-to-noise...
ratio (SNR, dB) was parametrically varied and the corresponding image noise standard deviation \( \sigma_Z \) was calculated using the relation: \( \sigma_Z = A/10^{\text{SNR}/20} \).

**III. Data Association Problem**

In the multi-target tracking scenario, the possibility of incorrect association between the various filters and measurements must be considered. The current analysis employed a data association logic based on the nearest neighbor method. For each of the four filters, the distances from each of the \( N \) prior particle positions to each of the four measurement peak intensity locations (\( x^*, y^* \)) are computed. Using the last prior weight set from each filter, the weighted mean of the distances from the \( i \)-th filter to the \( j \)-th peak is calculated over the \( N \) particles. Then, each filter associates with itself the measurement having the smallest mean weighted distance. *This technique for data association was used in both multi-target tracking methods described in Sections IV and V below.* Details of the association algorithm are as follows:

Let the prior particle sets and their associated normalized weights for the \( n \)-th particle and the \( i \)-th filter be denoted by:

\[
\{ \hat{x}_i^n, \hat{y}_i^n, w_i^n \}_{i=1:4}^{n=1:N}
\]

The coordinates of the \( j \)-th peak measured intensity on the image plane are written as:

\[
\{ x_j^*, y_j^* \}_{j=1:4}
\]

The distance from the \( n \)-th particle of the \( i \)-th filter to the \( j \)-th peak is given by:

\[
r_{i,j}^n = \sqrt{(\hat{x}_i^n - x_j^*)^2 + (\hat{y}_i^n - y_j^*)^2}
\]

The weighted mean distance from the \( i \)-th filter to the \( j \)-th measurement peak is computed as:

\[
R_j^i = \sum_{n=1}^{N} w_i^n r_{i,j}^n
\]

Then, the \( i \)-th filter uses the measurement \( j^* \), where \( j^* \) is the index corresponding to:

\[
R_{i}^{j^*} = \min_{j} \{ R_j^i \}
\]
IV. PF Tracking Method #1

Use of the PF for multi-target tracking was implemented using two methods. In Method #1, the importance weights were generated as shown in Fig. 5. The location \((x^*, y^*)\) of the maximum measured intensity for each measurement set is first determined. Then, the distances \((r_n)\) between the estimated position coordinates of the n-th prior particle \((x_n, y_n)\) and the peak intensity locations \((x^*, y^*)\) are calculated. The weights are computed using the formula at the bottom of Fig. 5, which is an exponential function of \((r_n/r_1)^2\) where the design parameter \(r_1\) was chosen as 1.0 pixel.

\[
W_n = \exp\left\{-\frac{1}{2}\left[(x^*-x_n)^2 + (y^*-y_n)^2\right]/r_1^2\right\} = \exp\left\{-\frac{1}{2}r_n^2/r_1^2\right\}
\]

Fig. 5. Generation of Importance Weights from Image Data (Method #1)

Tracking results using Method #1 are shown for SNR = 15 dB. For this case the PF used 10,000 particles. Fig. 6 displays the true target positions and the estimated particle clouds at selected time points. The position estimation errors and 2-\(\sigma\) bounds for each target are shown in Fig. 7. The results indicate accurate and consistent tracking with Method #1 at this SNR. The performance of the data association logic is illustrated in Fig. 8. The weighted distances are shown from each filter (X1-X4) to each measurement peak (Z1-Z4). It is seen that for this SNR each filter is associated with the correct measurement, as identified by the smallest distance.
Fig. 6. Method #1 Tracking Results, SNR = 15 dB

Fig. 7. Method #1 Position Estimation Errors and 2-σ Bounds, SNR = 15 dB
Further analysis of Method #1 indicated that while excellent tracking performance was obtained at fairly high SNR (≥ 15 dB), the filter performance rapidly deteriorated below this point with eventual collapse of the filter solution. This seems somewhat intuitive since Method #1 relies on being able to distinguish the peak intensity location on the array in the presence of noise. As SNR decreases, the ability to discern this peak becomes increasingly difficult, as evidenced by the problem of visually picking the target out of low SNR images. As a result of this limitation, an alternate Method #2 was examined.

V. PF Tracking Method #2

This method borrows ideas from the track-before-detect literature [3, 4 (chapter 11)] but extends them by considering multiple targets and by employing a data association logic based on nearest neighbor using mean weighted distance (see Section III). Method #2 uses a technique in which the weights are computed from a likelihood ratio, which will be defined below.

A number of assumptions have been employed in the present analysis and are listed below:

- The targets have been previously detected and always exist thereafter. (i.e., the detection problem is not explicitly addressed but could be considered in an expanded analysis)
- The number of targets is constant and known (in this case, four).
- There are four PFs, each assigned to one target.
- The filters are initialized with reasonably small errors: ± 3 pixels and ± 0.5 pixels/s maximum errors, uniformly distributed. (The sensitivity to initialization errors will be examined later.)
The measurement model for the n-th particle in cell (i, j) is a Gaussian point spread function
with amplitude A and blur parameter \( \Sigma \):

\[
z^{(i,j)} = h^{(i,j)}(\hat{x}^n, \hat{y}^n) = A \exp \left\{ -\frac{(i - \hat{x}^n)^2 + (j - \hat{y}^n)^2}{2\Sigma^2} \right\}
\]

In the analysis, the assumed filter values of A and \( \Sigma \) were chosen to be different from the true
values.

The likelihoods in cell (i, j) for signal plus noise and for noise alone are given by:

\[
p_{S+N}^{(i,j)} = p(z^{(i,j)} | \hat{x}^n, \hat{y}^n) = N(z^{(i,j)}; h^{(i,j)}, \sigma_Z^2)
\]

\[
p_N^{(i,j)} = p(z^{(i,j)}) = N(z^{(i,j)}; 0, \sigma_Z^2)
\]

where

\[
N(z^{(i,j)}; h^{(i,j)}, \sigma_Z^2) \propto \exp \left\{ -\frac{1}{2} \left( z^{(i,j)} - h^{(i,j)} \right)^2 / 2\sigma_Z^2 \right\}
\]

Division of these gives the likelihood ratio:

\[
\ell^{(i,j)} = \frac{p_{S+N}^{(i,j)}}{p_N^{(i,j)}} = \exp \left\{ -\frac{h^{(i,j)}(h^{(i,j)} - 2z^{(i,j)})}{2\sigma_Z^2} \right\}
\]

Method #2 uses the likelihood ratio to calculate the unnormalized weights. Assuming the
measurement noise is independent from pixel to pixel, these are given by:

\[
\tilde{W}^n = \prod_{i \in C} \prod_{j \in C} \ell^{(i,j)}
\]

where C is a region of \( \pm D \) pixels about the (x, y) position estimate of the n-th prior particle. In
the current analysis \( D = 10 \) pixels was used. A summary of the parameters used in the evaluation
of Method #2 is provided in Table 1.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Assumed Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of particles, N</td>
<td>10,000</td>
</tr>
<tr>
<td>Time step, T (for both propagate &amp; update)</td>
<td>0.1 s</td>
</tr>
<tr>
<td>Target 1-σ normal accel, σ_{AN}</td>
<td>1.0 pixel/s^2</td>
</tr>
<tr>
<td>Target 1-σ longitudinal accel, σ_{AL}</td>
<td>0.5 pixel/s^2</td>
</tr>
<tr>
<td>Filter process noise 1-σ x, y accel, σ_w</td>
<td>0.5 pixel/s^2</td>
</tr>
<tr>
<td>Initial conditions ±(x, y)max (uniform)</td>
<td>3.0 pixels</td>
</tr>
<tr>
<td>Initial conditions ±(vx, vy)max (uniform)</td>
<td>0.5 pixels/s</td>
</tr>
<tr>
<td>Signal-to-noise ratio, SNR</td>
<td>2, 6, 15 dB</td>
</tr>
<tr>
<td>True amplitude, A_{SIGNAL}</td>
<td>1.0</td>
</tr>
<tr>
<td>Estimated amplitude</td>
<td>0.8 * A_{SIGNAL}</td>
</tr>
<tr>
<td>True 1-σ image noise, σ_z</td>
<td>A_{SIGNAL} / (10^{SNR/20})</td>
</tr>
<tr>
<td>Estimated 1-σ image noise</td>
<td>1.5 * σ_z</td>
</tr>
<tr>
<td>True blur parameter, Σ</td>
<td>2.0 pixels</td>
</tr>
<tr>
<td>Estimated blur parameter</td>
<td>1.5 * Σ</td>
</tr>
<tr>
<td>Focal plane array dimensions</td>
<td>256 x 256 pixels</td>
</tr>
<tr>
<td>Box (± D) about estimated (x, y) used in weight calc</td>
<td>10 pixels</td>
</tr>
</tbody>
</table>

Table 1. Assumed Parameters for Method #2

Results for Method #2 are now provided as a function of SNR. The tracking performance of Method #2 at SNR = 15 dB was found similar to Method #1 but was somewhat more accurate. Fig. 9 shows the true and estimated target tracks and Fig. 10 shows the position estimation errors and their 2-σ bounds. The tracking errors are seen to be very small, a small fraction of a pixel. The data association algorithm is assessed in Fig. 11, which displays flags for each filter that indicate whether the filter-target association was correct (value = 1) or incorrect (value = 0). Next to the filter number is shown the probability of correct associations (PCA) computed over the tracking interval. For this SNR, the PCA is very high, ranging from 98-100 percent.
Fig. 9. Method #2 Tracking Results, SNR = 15 dB

Fig. 10. Method #2 Position Estimation Errors and 2-σ Bounds, SNR = 15 dB
To further verify algorithm consistency, Fig. 12 provides plots of the measurement residuals and 2-σ bounds. Although there is actually a residual for each cell in the array, and for each of the N particles, the residuals here were computed as the differences between the measured and modeled intensities in the cell nearest to each prior particle location. Then a weighted mean and variance calculation was performed over the N residual values. The residual means and standard deviations so derived are plotted in Fig. 12 and indicate consistent behavior.
In sharp contrast to Method #1, Method #2 was able to maintain accurate tracking at very low SNR. Results for SNR = 6 dB are shown in Figs. 13 through 15. The true and estimated target tracks are displayed in Fig. 13. The position estimation errors in Fig. 14 are larger than those at 15 dB, but, except for a few transient excursions, tracking is being performed to within 1-2 pixels most of the time and the tracking errors are generally contained within their statistical bounds. The data association indicator history in Fig. 15 shows that at SNR = 6 dB, the PCA is between 26-40 percent, with an average of 30.7.

![Fig. 13. Method #2 Tracking Results, SNR = 6 dB](image)

![Fig. 14. Method #2 Position Estimation Errors and 2-σ Bounds, SNR = 6 dB](image)
Fig. 15. Method #2 Correct Association Indicators for Each Filter, SNR = 6 dB

Method #2 was also found to give very satisfactory performance at SNR = 2 dB. These results are shown in Figs. 16 through 18. Correct data associations were made between 23-34 percent of the time with an average of 27.5.

The PCA (averaged over the tracking interval and over four filters) was computed for a range of SNR values and these results are shown in Fig. 19. For Method #2, correct associations are made about 27 percent of the time at SNR = 0 dB, 50 percent of the time at SNR = 9 dB, and 100 percent of the time for SNR $\geq$ 15 dB. In contrast, Method #1 is seen to have much poorer association performance.

Fig. 16. Method #2 Tracking Results, SNR = 2 dB
Fig. 17. Method #2 Position Estimation Errors and 2-σ Bounds, SNR = 2 dB

Fig. 18. Method #2 Correct Association Indicators for Each Filter, SNR = 2 dB
Fig. 19. Probability of Correct Association vs SNR

The sensitivity to initial condition uncertainty was evaluated using a slightly different scenario, as shown in Fig. 20. It was assumed that the particles were initialized according to a uniform distribution over a region $\pm X_{\text{max}}$ pixels about the true target x, y positions and over $\pm V_{\text{max}}$ pixels/s about the true x, y velocities. The values of $X_{\text{max}}$ and $V_{\text{max}}$ were made increasingly larger until the accuracy of the target tracking started to break down. As seen in Fig. 20 for SNR = 2 dB, acceptable tracking with Method #2 is maintained for initial uncertainties up to about $X_{\text{max}} = 7\text{-}10$ pixels and $V_{\text{max}} = 1.5\text{-}2.0$ pixels/s.
Initial Particle Distribution is Uniform $\rightarrow$ True Position $\pm X_{\text{max}}$, True Velocity $\pm V_{\text{max}}$

Fig. 20. Tracking Sensitivity to Initial Condition Uncertainty, SNR = 2 dB

VI. Summary and Conclusions

This paper investigated two methods for performing multi-target tracking with a PF. The first method computed importance weights based on an exponential function of the distance from the PF estimated position to the location of the peak measured intensity. Method #1 was found to perform well at SNR $\geq 15$ dB, but rapidly deteriorated below that. Since good performance at low SNR was desired, a second method was investigated. Method #2 utilized track-before-detect techniques and computed the importance weights based on a likelihood ratio function. This method was shown to give excellent tracking performance down to low SNR approaching 0 dB. A new data association algorithm was proposed based on a weighted nearest neighbor strategy and was evaluated in the study. The probability of correct association was determined as a function of SNR. The influence of initial condition uncertainty was also evaluated.
VII. Acknowledgement

This research was supported under MDA Contract No. HQ0006-10-C-7267, with Dr. Demetrios Serakos of the Naval Surface Warfare Center, Dahlgren Division as the technical monitor. We thank Dr. Serakos for his interest and suggestions on several aspects of our work.

References


