A Laplace-based particle filter for track-before-detect

Christian Musso
ONERA – The French Aerospace Lab
F-91761 Palaiseau, France
Email: christian.musso@onera.fr

Paul Bui Quang
CEA, DAM, DIF
F-91297 Arpajon, France
Email: paul.bui-quang@cea.fr

Achille Murangira
IFPEN
F-92852 Rueil-Malmaison, France
Email: achille.murangira@ifpen.fr

Abstract—Track-before-detect (TBD) is an efficient approach for tracking a target with low signal-to-noise-ratio. Unlike classical detection methods, TBD does not require thresholding. TBD integrates the signal of the target over space and time. In this paper, a particle filter is proposed, where the proposal density is constructed thanks to local Laplace approximations at the modes of the posterior density. This particle filter is tested on a simulated TBD problem, where the observations are a sequence of infrared images.

I. INTRODUCTION

Track-before-detect is an efficient method to track and detect a target with low signal-to-noise ratio observations. The tracking can be addressed as a Bayesian filtering problem. In TBD, the uncertainty on the target state remains high until a sufficient amount of observations has been integrated. In Bayesian terms, it means that the posterior distribution, i.e., the distribution of the target state, has many modes (local maxima). Particle filtering algorithms are adapted to this type of models. There exists many particle filter versions, in which the particles are sampled and weighted in different ways. See [1], [2] or [3] for reviews.

Various particle filters have been proposed for TBD : in radar applications ([4], [5]), in imaging sensor ([6], [7]). TBD can also include the phase information [8]. In this paper, we propose a new particle filter and apply it to TBD. We focus on the tracking of a target in long-term horizon with low signal-to-noise-ratio. The detection stage will be included in a future paper. The proposed algorithm uses an importance sampling distribution, which approximates the posterior density as a mixture of densities. Each density of the mixture is a local Laplace approximation of the posterior density around a mode. The mean of the density is the mode and its covariance matrix is the inverse of the posterior information matrix evaluated at the mode.

The computation of the modes requires to solve an optimization problem whose dimension is the state space dimension. Hence, when the state space dimension is large, this optimization can be costly. We propose a method to perform the optimization on a subspace of lower dimension, using the fact that only a few components of the state vector depend nonlinearly on the observation.

In Section II, we describe the formulation of TBD as a Bayesian filtering problem. In Section III, we present the importance sampling density, or proposal density, which we use in the proposed particle filter. We also present the technique to reduce the dimensionality of the mode computation optimization problem. The particle filter implementation is described in Section IV. Simulation results are provided in Section V.

II. PROBLEM FORMULATION

It is assumed that the target follows a constant velocity model, with a unknown intensity \( I_k \).

\[
X_k = [x_k, y_k, I_k, \dot{x}_k, \dot{y}_k]^T = [x^p_k, I_k, \dot{x}_k, \dot{y}_k]^T
\]

where \( x^p_k = [x_k, y_k]^T \) is the position of the target at time \( (k) \) in the image plane. The state-transition matrix is defined as

\[
F = \begin{bmatrix}
1 & 0 & 0 & \Delta T & 0 \\
0 & 1 & 0 & 0 & \Delta T \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}
\]

The target evolves according to

\[
X_k = F X_{k-1} + \eta_k
\]

where \( \eta_k \) is a white Gaussian noise process. At each time \( (k) \) the sensor provides an image of \( n \) cells. The measurement \( Y_k (\tau) \) is the intensity of each cell \( \tau = [\tau_x, \tau_y]^T \). It is modelled according to a radially decaying and isotropic function centered around the target’s position

\[
Y_k (\tau) = \frac{I_k}{\sqrt{\pi} \sigma_{psf}} \exp \left( - \frac{1}{2 \sigma^2_{psf}} |\tau - x^p_k|^2 \right) + \epsilon_k
\]

where \( \phi_{psf}(\tau, x^p_k) \) is the intensity contribution in the cell \( \tau \) of the target localized in \( x^p_k \). This Point Spread Function (psf) describes the response of the imaging system to the target (Figure 1). The psf is normalized such that \( \int_{\mathbb{R}^2} \phi_{psf}^2 (\tau, x) d\tau = 1 \).

It is convenient for future calculations (see section III-B) to vectorize the involved matrices

\[
\begin{align*}
Y_k &= [Y_k (\tau_1), Y_k (\tau_2), \ldots, Y_k (\tau_n)]^T \\
\Psi_k (x^p_k) &= [\phi_{psf}(\tau_1, x^p_k), \phi_{psf}(\tau_2, x^p_k), \ldots, \phi_{psf}(\tau_n, x^p_k)]^T
\end{align*}
\]
where \( \tau^i = [\tau_x, \tau_y]^T \) represents the Cartesian coordinates of the cell \( \tau^i \). Now, the measurement \( Y_k \) is a vector of length \( n \), the number of cells. \( \Psi \) is the known target signature. If the target is located at \( x_k^p \) in the image, equation (3) becomes

\[
Y_k = I_k \Psi(x_k^p) + \epsilon_k
\]

(4)

The noise \( \epsilon_k \) is Gaussian distributed with zero mean and spatial covariance matrix \( R \) of dimension \( n \times n \). Let us now calculate the likelihood

\[
P(Y_k|X_k) \propto \exp \left( -\frac{1}{2} [Y_k - I_k \Psi(x_k^p)]^T R^{-1} [Y_k - I_k \Psi(x_k^p)] \right)
\]

(5)

The computation of the likelihood is performed in the vicinity of \( x_k^p \) where \( \Psi(\cdot) \) does not vanish as suggested in [6].

III. PROPOSAL DENSITY

In this section, we consider a static setting where the unknown d-dimensional state \( x \) is distributed according to a prior \( q \) and is observed through a measurement

\[
y = h(x) + \epsilon
\]

(6)

\( h \) being a non-linear function and \( \epsilon \) a zero mean Gaussian noise. The prior \( q \) is not expressed in closed form, only a sample from \( q \) is available (it is the case in particle filtering). We suppose in this section that the posterior has a predominant mode.

A. Laplace-based proposal density

We recall the Laplace method [9] which is useful to design an efficient proposal function. The likelihood \( p(y|x) \) is denoted by \( g(x) \), the posterior is written as

\[
P(X = x | Y = y) \triangleq p(x|y) \propto g(x) q(x)
\]

(7)

We wish to compute an estimate of the posterior. The importance sampling (IS) estimator \( \hat{p} \) of \( p(x|y) \) is obtained by drawing \( N \) samples \( X^i \) from a proposal distribution \( \hat{q} \) so that

\[
p(x|y) \approx \sum_{i=1}^{N} u^i \delta_{x=X^i} \triangleq \hat{p}(x)
\]

(8)

where \( u^i = \frac{\hat{q}(X^i)}{\sum_{i} \hat{q}(X^i)} \) with the following importance weights

\[
\hat{w}^i = \frac{g(X^i)q(X^i)}{\hat{q}(X^i)}
\]

(9)

A good criterion for evaluating the accuracy of the approximation (8) is the asymptotic variance of the unnormalized weights \( \hat{w}^i \) [10]

\[
V(\hat{q}) = \frac{1}{N} \left( \frac{\int_{R^d} g(x)^2 q(x)^2 dx}{\left( \int_{R^d} g(x)q(x) dx \right)^2} - 1 \right)
\]

(10)

It is well-known that the optimal importance function is the posterior \( \hat{q}_{opt}(x) = p(x|y) \) for which \( V(\hat{q}_{opt}) = 0 \). Thus, \( \hat{q} \) must be chosen close to the posterior. The choice of the proposal distribution is crucial for controlling the Monte Carlo error which can be large if the prior and the likelihood have little overlap. This is the case for the TBD application when the target only contributes to a few cells (possibly a single cell), i.e. when \( \sigma_{xy} \) (3) is small.

For this purpose, we wish to design a proposal, for instance a Gaussian, which has means nearly equal to those of the posterior. The posterior expectation \( \mathbb{E}[X|Y] \) and the posterior covariance matrix \( \mathbb{V}[X|Y] \) are well approximated by the Laplace formula [9] if the posterior has a predominant mode

\[
\mathbb{E}[X|Y] \approx \hat{x} - F_0(\hat{J}, \hat{J}^T)
\]

\[
\mathbb{V}[X|Y] \approx \hat{J}^{-1} - G_0(\hat{J}, \hat{J}^T, \hat{J}^T)
\]

(11)

where \( \hat{x} \) is the maximum a posteriori (MAP)

\[
\hat{x} = \arg \max_{x \in \mathbb{R}^d} \{ g(x)q(x) \}
\]

(12)

and where \( \hat{J} = J(\hat{x}) \) is the (positive definite) posterior information matrix with

\[
J(x) = -(\log q)'(x) - (\log q)''(x)
\]

(13)

\( J' \) and \( J'' \) are the first and second derivatives with respect to \( x \). Notice that \( J \) depends on the measure \( y \) through the MAP \( \hat{x} \).

The approximations (11) are very accurate [9] even if the prior and the likelihood have little overlap. The functions \( F_0 \) and \( G_0 \), which are related to the asymmetry of the posterior, are described in [9]. For the sake of simplicity, these functions will be neglected in the practical algorithm while it is possible to compute the accurate approximations (11). But, for the TBD application the following approximations are sufficient

\[
\begin{align*}
\mathbb{E}[X|Y] &\approx \hat{x} \\
\mathbb{V}[X|Y] &\approx \hat{J}^{-1}
\end{align*}
\]

(14)
The proposal density \( \tilde{q} \) is chosen so that it has \( \hat{x} \) as mean and \( J^{-1} \) as covariance matrix

\[
\tilde{q} \sim \tilde{q}(\hat{x}, J^{-1})
\]  

(15)

Of course, \( \tilde{q} \) must be chosen among densities that are easy to sample from. For example Gaussian densities and Student densities. The latter give theoretically more robust results since they have heavy tails which implies that \( V(\tilde{q}) (10) \) is bounded.

B. Implementation of Laplace-based IS for TBD

There are three difficulties encountered when calculating the Laplace-based proposal:

(i) The computation of the MAP (12). It involves a maximization in a d-dimensional space: \( d = 5 \) for TBD (1)

(ii) The computation of the matrix \( J(x) \) (26). It requires the evaluation of the derivatives of the prior \( q(x) \)

(iii) The computation of the importance weights (9). It involves the evaluation \( q(X'') \)

(i) In most filtering problems, only a small part of the state vector \( X_k \) takes part in the likelihood \( P(Y_k|X_k = x) = g(x) \). This part, specifically its nonlinear contribution, can be handled separately in the maximization (12) under an assumption on the prior. We adopt here a methodology similar to the one developed in [11] and [12]. For the TBD state vector (1) let us define

\[
x_1 = x_k^0 = [x_0, y_0]^T \quad \text{and} \quad x_2 = [I_k, \dot{x}_k, \dot{y}_k]^T
\]  

(16)

Dropping momentarily the time index, we have

\[
p(x|y) \propto p(y|x_1, x_2) q(x_2|x_1) q(x_1)
\]

Thus, the calculation of the MAP can be split as follows

\[
\max_{(x_1, x_2)} p(x_1, x_2|y) = \max_{x_1} q(x_1) \max_{x_2} \underbrace{p(y|x_1, x_2) q(x_2|x_1)}_{F(x_1, x_2)} \quad \text{(17)}
\]

It is assumed that the prior \( q(x_1, x_2) \) is Gaussian with mean \( [E_1, E_2]^T \) and covariance matrix

\[
P = \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix}
\]  

(18)

Under this assumption, \( F(x_1, x_2) \) can be maximized explicitly with respect to \( x_2 \). Indeed, the conditional probability density function (pdf) \( q(x_2|x_1) \) given \( x_1 \) is Gaussian with mean \( E_{2|1}(x_1) = E(x_2|x_1) \) and covariance matrix \( P_{2|1} \)

\[
E_{2|1}(x_1) = E_2 + P_{21} P_{11}^{-1} (x_1 - E_1)
\]

\[
P_{2|1} = P_{22} - P_{21} P_{11}^{-1} P_{12}
\]  

(19)

Recalling that \( Y_k \) given \( X_k \) is also Gaussian (4), we have (up to an additive constant)

\[
G(x_1, x_2) \triangleq -2 \log [F(x_1, x_2)] = [Y_k - I_k \Psi(x_1)]^T R^{-1} [Y_k - I_k \Psi(x_1)]
\]

\[
+ [x_2 - E_{2|1}(x_1)]^T P_{2|1}^{-1} [x_2 - E_{2|1}(x_1)] \quad \text{(20)}
\]

Define the nx3 matrix \( M \equiv M_q(x_1) = \Psi(x_1) \times [1, 0, 0] \) so that \( M \equiv x = I_k \Psi(x_1) \) and denote \( E_{2|1} = E_{2|1}(x_1) \), equation (20) gives

\[
G(x_1, x_2) = [Y_k - M \gamma \gamma]^T R^{-1} [Y_k - M \gamma \gamma] + [x_2 - E_{2|1}]^T P_{2|1}^{-1} [x_2 - E_{2|1}] \quad \text{(21)}
\]

We can easily minimize \( G(x_1, x_2) \) with respect to \( x_2 \) for \( x_1 \) fixed, since this positive function is quadratic with respect to \( x_2 \). It can be shown that \( G(x_1, x_2) \) can be written as the sum of 3 positive terms

\[
G(x_1, x_2) = [Y_k - M \gamma \gamma]^T R^{-1} [Y_k - M \gamma \gamma] + [\gamma \gamma - E_{2|1}]^T P_{2|1}^{-1} [\gamma \gamma - E_{2|1}] + [x_2 - \gamma \gamma]^T \Sigma^{-1} [x_2 - \gamma \gamma] \quad \text{(22)}
\]

where

\[
\Sigma^{-1} = M_q^T R^{-1} M_q + P_{2|1}^{-1}
\]

(23)

We can see in (22) that the minimum of \( G(x_1, x_2) \) with respect to \( x_2 \) is reached in

\[
\hat{x}_2(x_1) = \arg \min_{x_2} G(x_1, x_2) = \gamma \gamma = \gamma \gamma(x_1, Y_k)
\]

(24)

Replacing \( x_2 \) by \( \gamma \gamma \) in (22) and recalling that \( q(x_1) \) is a marginal of the Gaussian \( q(x_1, x_2) \), we can maximize the whole criterion (17) in 2 steps.

Computing the MAP

- Minimize the conditional criterion with respect to \( x_1 \)
  \[
  \hat{x}_1 = \arg \min_{x_1 \in \mathbb{R}^2} G(x_1)
  \]

- Evaluate \( \hat{x}_2 \)
  \[
  \hat{x}_2 = \gamma \gamma(\hat{x}_1)
  \]

with \( x_1 \in \mathbb{R}^2 \) and where \( G(x_1) \) is defined by

\[
G(x_1) \triangleq [x_1 - E_1]^T P_{11}^{-1} [x_1 - E_1] + [Y_k - M \gamma \gamma(x_1, Y_k)]^T R^{-1} [Y_k - M \gamma \gamma(x_1, Y_k)]
\]

\[
+ [\gamma \gamma(x_1, Y_k) - E_{2|1}(x_1)]^T P_{2|1}^{-1} [\gamma \gamma(x_1, Y_k) - E_{2|1}(x_1)]
\]

(25)
The MAP is estimated by \( \hat{x} = [\hat{x}_1, \hat{x}_2]^T \). Finally, the initial d-dimensional maximization (12) to get the MAP boils down to a 2-dimensional minimization. The moments \((E_1, E_2, P_{11}, P_{12}, P_{22})\) of the prior (18) are estimated empirically using samples drawn from \( q \) (predicted particles). It should be noted that the search space for obtaining \( \hat{x}_1 \) is restricted to the \((1 - \epsilon)\) confidence ellipse given by \( P_{11} \), the covariance matrix of the position components of the predicted particles. Notice that a moderately precise estimation of the MAP is generally sufficient for the pdf approximation (8). The Monte Carlo variance (10) may increase a little but the bias remains zero.

(ii) The computation (see section III-B) of \( J(x) \) (26) can be handled in the following way. The first term \(-\langle \log q \rangle'(x)\) is easy to determine as \( g(x) \), the likelihood, is expressed in closed form (5). The second term \(-\langle \log q \rangle''(x)\) needs to be approximated. Again, under the Gaussian assumption of the prior \( q(x_1, x_2) \), this term is equal to \( P^{-1} \) where \( P \) is the covariance matrix of \( q \) (18). The posterior information matrix is thus approximated by

\[
J(x) \approx -\langle \log g \rangle''(x) + P^{-1}
\]  

(26)

(iii) The evaluation \( q(x^i) \) can be done by using kernel estimation [13]. This method has two drawbacks for our application. The first one is that the computational cost is high as the kernel evaluation is performed for each sample \( x^i \) (9) drawn from \( \tilde{q} \). The second one is that the kernel estimation is not accurate for a moderate size sample when the dimension of the state vector is high \((d \geq 5)\) [13]. Therefore, the evaluation of \( q(.) \) is done simply assuming that \( q \) is a Gaussian. The mean and covariance matrix of \( q \) are estimated empirically using the predicted sample. Theoretically this introduces a bias in the estimation of the importance weights (9). In practice (see section V) this bias is not observed.

IV. PARTICLE FILTER FOR TBD

In TBD when the SNR is low, the posterior is generally multimodal particularly in the first iterations when the initial uncertainty of the state is large. Roughly speaking, “false” modes can arise due to the noise which is often stronger than the target’s signal. In this section we propose a particle filtering algorithm which addresses the problem of multimodality in the TBD context using local Laplace-based proposals.

A. Initialisation

At the initial time \((k = 0)\) we place a particle (position component) on each cell of the image. The remaining components \(x_{2,k} \) are drawn independently according to a Gaussian law with mean \( E_{2,k} \) and covariance matrix \( P_{22,k} \) (18). The prior \( q_k \) is assumed Gaussian with the moments estimated by the particles. This leads to an i.i.d initial sample of particles \( \{X^1_k, X^2_k, ..., X^N_k\} \) where

\[
X^i_k = [x^i_{1,k}, x^i_{2,k}, I^i_k, d^i_k, y^i_k]^T
\]  

(27)

The measure \( Y_k \) is available. We aim to design a proposal close to the (possibly multimodal posterior using Laplace-based importance sampling (see section III-B). For this, we compute a number of maxima of the posterior

\[
p(x_{1,k}, x_{2,k} | Y_k) = g_k(x_{1,k}, x_{2,k}) q_k(x_{1,k}, x_{2,k})
\]  

(28)

where \( g_k(x_{1,k}, x_{2,k}) = p(Y_k | x_{1,k}, x_{2,k}) \) is the likelihood (5).

Consider the \((1 - \epsilon)\) confidence ellipse \( \Sigma_{1,k} \) for the \( x_{1,k} \) components. At the initialisation phase, this ellipse contains the whole image. For each cell \( \tau^j = [\tau^j_1, \tau^j_2] \) \( j = 1, ..., \tilde{m}_k \) belonging to \( \Sigma_{1,k} \), we compute the conditional criterion (25)

\[
G_j = G(\tau^j) \quad \text{for} \quad j = 1, ..., \tilde{m}_k
\]  

(29)

Now, we select the \( m_k \) best cells \( \tau^j \) \( j = 1, ..., m_k \) which have the \( m_k \) minimum values of \( G_j \). Remember that for each local minimum \( \tau^j \) we can compute a local maxima of the posterior for the whole state vector (24)

\[
\tilde{x}_k^j = [\tau^j, \gamma_\psi(\tau^j)]^T \quad \text{for} \quad j = 1, ..., m_k
\]  

(30)

The idea of placing particles on the cells with high signal level has been introduced in [7]. But dealing with the criterion (25) is more informative as it represents the level of the matching between the signature and the measurement given \( x_{2,k} \).

We could sample around each local minimum \( \tilde{x}_k^j \) using a mixture of proposals centered on these minima. But, the number of the best cells \( m_k \) is chosen deliberately high (for example : \( \tilde{m}_k/3 < m_k \leq \tilde{m}_k \)) especially when the SNR is low. Moreover, observing that these minima form clusters, it is easier and sufficient to cluster them first.

Thus, we perform a clustering of the 2-dimensional points \( \tilde{x}_k^j \) \( j = 1, ..., m_k \) with a fixed number \( n_c \) of clusters. This gives \( n_c \) centers of clusters \( \tilde{c}_k^j \) \( j = 1, ..., n_c \). For each center we compute each of the \( n_c \) components of the local mode and the local posterior information matrix using (24) and (26)

\[
\begin{align*}
\hat{\xi}_k^j &= [\hat{\xi}_k^j, \gamma_\psi(\hat{\xi}_k^j)]^T \\
\hat{J}_k^j &= J_k(\hat{\xi}_k^j)
\end{align*}
\]  

(31)

Finally, the proposal function is the following mixture

\[
\tilde{q}_k(x) = \frac{1}{n_c} \sum_{j=1}^{n_c} \phi \left( x, \hat{\xi}_k^j, [\hat{J}_k^j]^{-1} \right)
\]  

(32)

where \( \phi \left( x, \hat{\xi}_k^j, [\hat{J}_k^j]^{-1} \right) \) is a density with mean \( \hat{\xi}_k^j \) and covariance matrix \( [\hat{J}_k^j]^{-1} \). Let \( \{X^1_{k+}, ..., X^N_{k+}\} \) be a i.i.d sample drawn from (32). The importance weights are computed in accordance with (9)

\[
\tilde{w}^j_{k+} = \frac{q_k(X^k_{k+}) q_k(X^k_{k+})}{\tilde{q}_k(X^k_{k+})}
\]  

(33)
The various steps for sampling according to the proposal are summarized below.

Summary

- Compute for each cell $\tau^j$ \( \{ j = 1, \ldots, m_k \} \) in $\Sigma_{1,k}$ the conditional criterion $G_j$
- Select the $m_k$ best minima $\hat{\tau}^i$ \( \{ i = 1, \ldots, m_k \} \)
- Perform a clustering of $\{ \hat{\tau}^i \}$ which gives the centers $\hat{\tau}^i = \{ j = 1, \ldots, n_c \}$
- Compute the local modes of the posterior $\xi^j_k = [\hat{\tau}^j_k, \gamma_0(\hat{\tau}^j_k)]^T$ \( \{ j = 1, \ldots, n_c \} \)
- Compute the local posterior information matrix $J_k^j = J_k(\xi^j_k)$
- Sample from the proposal $q_k(x) = \frac{1}{n_c} \sum_{i=1}^{n_c} \phi \left( x, \xi^j_k, [J_k^j]^{-1} \right)$
- \( \rightarrow \{ X_{k+1}^j, \ldots, X_{k+1}^{n_c} \} \)
- Compute the local modes of the posterior $\xi^j_k$ \( \{ j = 1, \ldots, n_c \} \)
- Compute the local posterior information matrix $J_k^j = J_k(\xi^j_k)$
- Compute for each cell $\tau^j$ \( \{ j = 1, \ldots, m_k \} \) in $\Sigma_{1,k}$ the conditional criterion $G_j$
- Select the $m_k$ best minima $\hat{\tau}^i$ \( \{ i = 1, \ldots, m_k \} \)
- Perform a clustering of $\{ \hat{\tau}^i \}$ which gives the centers $\hat{\tau}^i = \{ j = 1, \ldots, n_c \}$
- Compute the local modes of the posterior $\xi^j_k = [\hat{\tau}^j_k, \gamma_0(\hat{\tau}^j_k)]^T$ \( \{ j = 1, \ldots, n_c \} \)
- Compute the local posterior information matrix $J_k^j = J_k(\xi^j_k)$
- Sample from the proposal $q_k(x) = \frac{1}{n_c} \sum_{i=1}^{n_c} \phi \left( x, \xi^j_k, [J_k^j]^{-1} \right)$
- \( \rightarrow \{ X_{k+1}^j, \ldots, X_{k+1}^{n_c} \} \)
- Compute for each cell $\tau^j$ \( \{ j = 1, \ldots, m_k \} \) in $\Sigma_{1,k}$ the conditional criterion $G_j$
- Select the $m_k$ best minima $\hat{\tau}^i$ \( \{ i = 1, \ldots, m_k \} \)
- Perform a clustering of $\{ \hat{\tau}^i \}$ which gives the centers $\hat{\tau}^i = \{ j = 1, \ldots, n_c \}$
- Compute the local modes of the posterior $\xi^j_k = [\hat{\tau}^j_k, \gamma_0(\hat{\tau}^j_k)]^T$ \( \{ j = 1, \ldots, n_c \} \)
- Compute the local posterior information matrix $J_k^j = J_k(\xi^j_k)$

For the next correction step of filtering, at time $(k+1)$, we have to compute the underlying predictive filtering pdf $q_{k+1}$ of this sample. Each sample $X_{k+1}^j$ is connected to the cluster $C_j$ (32) propagated by the linear target’s dynamics (2). Formally, $q_{k+1}$ can be written as

$$q_{k+1}(x) \approx \sum_{j=1}^{n_c} \left( \sum_{i \in C_j} w^i_{j,k+1} \right) q_{k+1}^j(x) \quad (36)$$

where $q_{k+1}^j(x)$ is a Gaussian with the following mean and covariance matrix

$$\begin{align*}
E^j_{k+1} &= \sum_{i \in C_j} w^i_{j,k+1} X_{k+1}^i \\
P^j_{k+1} &= \sum_{i \in C_j} w^i_{j,k+1} (X_{k+1}^i - E^j_{k+1})(X_{k+1}^i - E^j_{k+1})^T
\end{align*} \quad (37)$$

The measure $Y_{k+1}$ is available which allows us to compute the likelihood $g_{k+1}(x) = P(Y_{k+1}|X_{k+1} = x)$ (5). Using (36), the posterior can be approximated as follows

$$p_{k+1}(x) \triangleq p(X_{k+1} = x|Y_{k+1}) \propto g_{k+1}(x) q_{k+1}(x) \approx \sum_{j=1}^{n_c} \left( \sum_{i \in C_j} w^i_{j,k+1} \right) g_{k+1}(x) q_{k+1}^j(x) \quad (38)$$

Let us now compute the local modes of the posterior. We suppose that the supports of the functions $q_{k+1}^j$ are disjoint, precisely, we suppose that the intersections of the supports are negligible. Hence, the local maxima of $p_{k+1}$ are approximatively those of $(\sum_{i \in C_j} w^i_{j,k+1}) g_{k+1}$ for $j = 1, \ldots, n_c$. Since $q_{k+1}^j$ is unimodal, we can apply the procedure described in section III-B to determine the local maxima of $p_{k+1}^j \propto g_{k+1} q_{k+1}^j$. We denote by $G_j(x_i)$ the criterion $G(x_i)$ (25) in replacing the prior expectation and prior covariance matrix by their local counterpart

$$\left[ E_1, E_2 \right] = [E^j_{1,k+1}, E^j_{2,k+1}]$$

$$P = P^j_{k+1}$$

where $E^j_{1,k+1}$ is the mean of the position component of $E^j_{1}$ and $E^j_{2,k+1}$ the mean of the remaining components. The various steps for sampling according to the proposal are summarized below.

Summary

- For each propagated cluster $C_j$ and for each cell $\tau^j$ in the $(1 - \epsilon)$ confidence ellipse given by $P^j_{k+1}$ compute $G_j(\tau^j) = (\sum_{i \in C_j} w^i_{j,k+1}) G_j(\tau^j)$
- Select the $m_k+1$ best cells which minimize $G_j(\tau^j)$
- Perform a clustering of $\{ \hat{\tau}^i \}$ which gives the centers $\{ j = 1, \ldots, m_k+1 \}$
- Compute the local modes of the posterior $\xi^j_{k+1} = [\hat{\tau}^j_k, \gamma_0(\hat{\tau}^j_k)]^T$ \( \{ j = 1, \ldots, m_k+1 \} \)
- Compute the local posterior information matrix $J^j_{k+1} = J_{k+1}(\xi^j_{k+1})$
Sample from the proposal
\[ \tilde{q}_{k+1}(x) = \frac{1}{N} \sum_{j=1}^{N} \tilde{\phi}(x, \hat{X}^j_{k+1}, [J^j_{k+1}]^{-1}) \]

\[ \rightarrow \{X^1_{k+1}, \ldots, X^N_{k+1}\} \]

Compute the importance weights
\[ \tilde{w}^i_{k+1} = \frac{g_{k+1}(X^i_{k+1}) q_{k+1}(X^i_{k+1})}{\tilde{q}_{k+1}(X^i_{k+1})} \]

Compute the normalized weights
\[ w^i_{k+1} = \frac{\tilde{w}^i_{k+1}}{\sum_{i=1}^{N} \tilde{w}^i_{k+1}} \]

Finally at time \((k+1)\) we obtain a weighted sample from the posterior \(p_k\) (38)
\[ \{(w^1_{k+1}, X^1_{k+1}), \ldots, (w^N_{k+1}, X^N_{k+1})\} \]

and we can iterate this algorithm going back to (35).

The sampling procedure described here is performed only when weight degeneracy occurs. That is, when the effective sample size \(N_{eff}\) falls below a given threshold \(N_{th}\).

- Compute \(w^i = g_{k+1}(X^i_{k+1}) w^i_{k+1}\), normalize \(w^i\) such that \(\sum_{i=1}^{N} w^i = 1\)

- Compute \(N_{eff} = \frac{1}{\sum_{i=1}^{N}(w^i)^2} \)
  - If \(N_{eff} \geq N_{th}\) then \(w^i_{k+1} \propto g_{k+1}(X^i_{k+1}) w^i_{k+1}\)
  - If \(N_{eff} < N_{th}\) then perform Laplace-based resampling method

Notice that, in this particle filter, there is no more standard (multinomial) resampling, the resampling is done via the Laplace proposal method.

V. SIMULATION RESULTS

We show some simulation results of the proposed particle filter. The scenario parameters are described below. The position in the image plan is expressed in term of the number of cells and the speed is expressed in terms of number of cells per second.

Scenario parameters
- Number of cells: \(n = 10000\)
  - Sampling period: \(\Delta T = 0.8\) s
  - Number of clusters: \(n_c = 20\) before convergence and \(n_c = 4\) after convergence
  - Kernel of the proposal: \(\phi(\cdot)\) is Gaussian (32)
- Standard deviation of the point spread function: \(\sigma_{psf} = 1.3\)
  - Initial position of the target in the image plan: \(x_0 = 28, y_0 = 35\)
  - Speed of the target: \(\hat{x} = 0.75, \hat{y} = -0.25\)
  - Intensity of the target: \(I = 3\), SNR = 6dB and 3dB
  - Target’s dynamics noise = 0
  - The covariance matrix \(R\) of the measurement (4), is diagonal: no spatial correlation of the noise
- Initial uncertainty for the position: one particle in each cell. Number initial of particles: \(N=10000\)

Initial uncertainty for the speed: \(\sigma_x = \sigma_y = 2\) with a systematic bias = 1.5
Initial uncertainty for the intensity of the target: \(\sigma = 2\) with a systematic bias = 1.5.

50 Monte Carlo trials have been performed. The root mean square error (RMSE) is computed for the 5 components of the state vector: position, intensity and speed. The results are presented in the following figures. We can see that in average the filter converges after 40 iterations for SNR = 6dB (figure 2) and after 90 iterations for SNR = 3dB (figure 5). When the particle filter achieves convergence, the number of particle is reduced to \(N=500\) with \(n_c = 4\).

![Fig. 2. RMSE for the position of the target with SNR=6dB](image_url)

![Fig. 3. RMSE for the intensity of the target with SNR=6dB](image_url)

VI. CONCLUSION

TBD is a challenging problem, which can be addressed as a Bayesian filtering problem. In the context of low signal-to-noise ratio, the difficulty in TBD is that the posterior target’s state distribution has many modes. Particle filters are adapted to this type of problem. We propose in this paper a new particle filter where the proposal density is a mixture designed thanks
to local Laplace approximations at the modes of the posterior density. The mean of each density of the mixture is a mode and its covariance matrix is the inverse of the posterior information matrix at the mode. The proposed particle filter is tested on the simulated TBD problem where the observations are a sequence of infrared images. It exhibits good results with a moderate number of particles.

REFERENCES


