

# PARTICLE FILTER WITH EFFICIENT IMPORTANCE SAMPLING AND MODE TRACKING (PF-EIS-MT) AND ITS APPLICATION TO LANDMARK SHAPE TRACKING

*Namrata Vaswani and Samarjit Das*

Dept. of Electrical & Computer Engineering, Iowa State University, Ames, IA 50011  
{namrata,samarjit}@iastate.edu

## ABSTRACT

We develop a practically implementable particle filtering (PF) method called “PF-EIS-MT” for tracking on large dimensional state spaces. Its application to tracking the shape change of a large number of “landmark” (feature) points from image sequences is shown. Two issues common to most large dimensional problems are (a) observation likelihood is often multimodal and the state transition prior is often broad in at least some dimensions and (b) direct application of PF requires an impractically large number of particles. PF-EIS-MT combines the advantages of two recently proposed ideas which address both of these issues. Improved performance of PF-EIS and PF-EIS-MT over existing PF algorithms is demonstrated for landmark shape tracking.

## 1. INTRODUCTION

Tracking is the problem of causally estimating a hidden state sequence from a sequence of observations that satisfy the Hidden Markov Model (HMM) assumption. A tracking algorithm recursively computes the “posterior” at time  $t$  (probability density function of the current state conditioned on all observations until the current time) using the posterior at  $t - 1$  and the current observation. For most nonlinear or non-Gaussian state space models, the posterior cannot be computed analytically. But, it can be efficiently approximated using the particle filter (PF) [1, 2] which is a sequential Monte Carlo technique. A PF outputs at each time  $t$ , a cloud of  $N$  “particles” (Monte Carlo samples), along with their corresponding weights, whose empirical measure closely approximates the true posterior for large  $N$ .

An important issue in PF design is the choice of an importance sampling density that reduces the variance of the particle weights and thus improves “effective particle size” [3]. The first PF algorithm [1] used the state transition prior as the importance density. This assumes nothing and is easiest to implement. But since it does not use knowledge of the current observation, the weights’ variance can be large, particularly when the observations are more reliable than the prior model. The “optimal” importance density [3] is the posterior conditioned on the previous state (denote it by  $p^*$ ). But in most

problems,  $p^*$  cannot be computed analytically. When it is unimodal, PF-Doucet [3] approximates it by a Gaussian about its mode (Laplace’s approximation [4]) and samples from the Gaussian. Other work that also implicitly assumes that  $p^*$  is unimodal includes [2, 5, 6]. But very often, the observation likelihood (OL) is multimodal or heavy-tailed, e.g. due to clutter, occlusions or low contrast/blur in image tracking problems or due to failed or nonlinear sensors in other problems. If the state transition prior (STP) is broad compared to the distance between OL modes (in case of multimodal OL) or compared to the distance between OL and STP mode (in case of heavy-tailed OL),  $p^*$  will be multimodal [7]. See Fig. 1 for an example. The STP is broad when either the state sequence is changing fast compared to the observation arrival rate or when the system model is unreliable. For such problems, we proposed the PF-EIS algorithm in [7] which combines the advantages of PF-Gordon [1] and PF-Doucet [3].

When in addition to multimodality, the state space dimension is large (typically more than 10 or 12), the effective particle size reduces [1, 8], thus making any regular PF impractical. If the state space model is conditionally linear-Gaussian, or if some states can be vector quantized into a few discrete centers, Rao Blackwellization (RB-PF) [9, 8] can be used. In general, neither assumption may hold. But in most large dimensional problems, the state change variance is large in only a few dimensions i.e. the “LDSS property” [7] holds. We exploited this in [7] to introduce a mode tracking (MT) approximation of importance sampling (IS) along the “residual” directions, which greatly reduced the IS dimension.

In this work, we combine the EIS and MT ideas to obtain the PF-EIS-MT algorithm and develop its application to landmark shape tracking. Landmark shape tracking is the problem of tracking the shape change of a set of 2D “landmarks” (feature points of interest) from an image sequence. Shape of a group of discrete points (landmarks) is the geometric information that remains when location, scale and rotational effects are filtered out [10]. An important example of landmark shape change is the shape change of the set of centroid locations of different body parts of the human body. Modeling shape change separately from global motion is useful here because the global motion is usually due to random camera zoom or motion (random forward/backward motion

or in-plane rotation due to the camera lying on an unstable platform or due to it being hand-held). By separating shape and motion dynamics, it is possible to learn the motion prior based on the type of camera and platform, while learning the shape dynamics for the particular action separately.

Another example is shape change of a group of persons or vehicles viewed from a distance, so that each person/vehicle forms a landmark. The camera may be a surveillance camera placed on the roof of a room or a camera placed in an unmanned air vehicle in a military application. In both of the above examples, we assume a scaled orthographic camera model which is valid either when the camera is looking down at the scene (its principal axis is perpendicular to the shape plane and intersects the center of the shape), or when the scene is far from the camera. A third application domain is medical image analysis where landmarks correspond to points of interest to the medical practitioner. For example, brain landmarks imaged during surgery (the brain shape deforms since the skull is opened) using an optical camera that may have some random motion, either due to zooming or due to forward/backward random motion w.r.t. the patient.

In [12], we introduced models for nonstationary shape change and used PF-Gordon [1] to track the shape and motion change. In the current work, we revisit and modify the model introduced in [12] and derive the PF-EIS and PF-EIS-MT algorithms for it. Tracking performance is compared with PF-Gordon and PF-Doucet. Our observation model is more realistic than that of [12], because it also models the occurrence of false landmarks due to clutter and missed landmarks due to occlusions/blur. Because of this, the OL is often multimodal or is heavy-tailed at the outlier mode (far from the STP mode). The state vector consists of the shape parameters and the global motion parameters (scale, rotation and translation). Usually, the STP of scale and rotation is broad to allow for occasional large camera motion. This combined with multimodal or heavy tailed OL often results in a multimodal  $p^*$ .

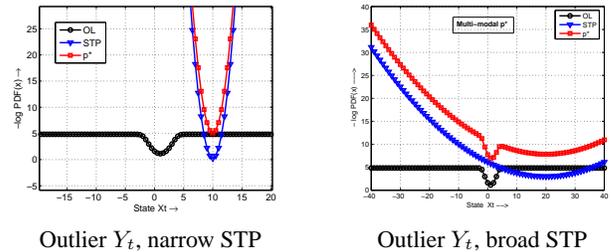
We give the problem formulation and develop the PF-EIS-MT algorithm in Sec. 2. In Sec. 3, the system and observation model for landmark shape change is given followed by deriving PF-EIS and PF-EIS-MT for that model. Tracking comparisons are given in Sec. 4.

## 2. PF WITH EFFICIENT IMPORTANCE SAMPLING (EIS) AND MODE TRACKING (MT)

### 2.1. General Problem Formulation

The goal is to sequentially estimate (track) a hidden sequence of states,  $X_t$ , from a sequence of observations,  $Y_t$ , which satisfy the Hidden Markov Model (HMM) property, i.e.

1. For each  $t$ , the dependence  $X_t \rightarrow Y_t$  is Markovian, with observation likelihood (OL) represented as  $p(Y_t|X_t)$ .
2. For each  $t$ , the dependence  $X_{t-1} \rightarrow X_t$  is Markovian, with state transition pdf (STP),  $p(X_t|X_{t-1})$ .



**Fig. 1.** Consider a scalar problem with OL,  $p(Y_t|X_t) = 0.8\mathcal{N}(Y_t; X_t, \sigma_o^2) + 0.2\mathcal{N}(Y_t; 0, 100\sigma_o^2)$  and STP  $p(X_t|X_{t-1}^i) = \mathcal{N}(X_t; 0.5X_{t-1}^i, \sigma_s^2)$ . The OL is a raised Gaussian, i.e. it is heavy tailed with mode at  $Y_t$ . Whenever  $Y_t$  is generated by the outlier component, the OL mode is far from the STP mode. If the STP is broad, this results in a bimodal  $p^*$ . We plot the  $-\log$  of the OL, STP and  $p^*$  for  $Y_t = 1$ ,  $\sigma_o^2 = 1$ , and  $0.5X_{t-1}^i = 10$ ,  $\sigma_s^2 = 0.25$  (narrow STP) and  $0.5X_{t-1}^i = 20$ ,  $\sigma_s^2 = 64$  (broad STP).

The posterior,  $\pi_t(X_t) \triangleq p(X_t|Y_{1:t})$ , needs to be recursively computed at each  $t$ , using  $\pi_{t-1}$  and the current observation  $Y_t$ . Once the posterior is available, any “optimal” state estimate, e.g. MAP or MMSE, can be computed. The tracking problem is complicated by the following two issues.

1. For a given observation,  $Y_t$ , the OL (as a function of the state,  $X_t$ ) is often multimodal or heavy-tailed at the outlier mode. If the STP is broad even in some dimensions, it will result in the posterior given previous state,

$$p^*(X_t) \triangleq p(X_t|X_{t-1}, Y_t) \quad (1)$$

being multimodal. For e.g., see Fig. 1. In such problems, PF-Doucet [3] as well as [2, 5, 6] that implicitly assume that  $p^*$  is unimodal cannot be used.

2. A second issue is applying the PF when the state space dimension is large. The effective particle size reduces with dimension, thus requiring a larger number of particles for a given accuracy as dimension increases.

### 2.2. The PF-EIS-MT Algorithm

We first explain the PF-EIS algorithm [7]. PF-EIS splits  $X_t$  into  $[X_{t,s}, X_{t,r}]$ , in such a way that  $p^*$  is unimodal conditioned on  $X_{t,s}^i$  ( $i^{th}$  particle of  $X_{t,s}$ ), i.e.

$$p^{*,i}(X_{t,r}) \triangleq p^*(X_t|X_{t,s}^i) = p(X_{t,r}|X_{t-1}^i, X_{t,s}^i, Y_t) \quad (2)$$

is unimodal. We sample  $X_{t,s}^i$  from its STP but approximate  $p^{*,i}$  by a Gaussian about its mode [4, 3] and sample  $X_{t,r}$  from it, i.e. we sample  $X_{t,r}^i$  from  $\mathcal{N}(m_t^i, \Sigma_{IS}^i)$  where

$$m_t^i = \arg \min_{X_{t,r}} [-\log p^{*,i}(X_{t,r})] = \arg \min_{X_{t,r}} L^i(X_{t,r}) \quad (3)$$

$$L^i \triangleq [-\log p(Y_t|X_{t,s}^i, X_{t,r})] + [-\log p(X_{t,r}|X_{t-1}^i, X_{t,s}^i)]$$

$$\Sigma_{IS}^i = [\nabla^2 L^i(m_t^i)]^{-1} \quad (4)$$

---

**Algorithm 1 PF-EIS-MT. Going from  $\pi_{t-1}^N$  to  $\pi_t^N(X_t) = \sum_{i=1}^N w_t^{(i)} \delta(X_t - X_t^i)$ ,  $X_t^i = [X_{t,s}^i, X_{t,r}^i]$ ,  $X_{t,r}^i = [X_{t,r,s}^i, X_{t,r,r}^i]$**

---

1. *Importance Sample  $X_{t,s}$* :  $\forall i$ , sample  $X_{t,s}^i \sim p(X_{t,s}^i | X_{t-1}^i)$ .
  2. *Efficient Importance Sample  $X_{t,r,s}$* :  $\forall i$ ,
    - (a) Compute  $m_t^i$  and  $\Sigma_{IS}^i$  using (3), (4). Let  $m_t^i = \begin{bmatrix} m_{t,s}^i \\ m_{t,r}^i \end{bmatrix}$  and  $\Sigma_{IS}^i = \begin{bmatrix} \Sigma_{IS,s} & \Sigma_{IS,s,r} \\ \Sigma_{IS,r} & \Sigma_{IS,r,s} \end{bmatrix}$ .
    - (b) Sample  $X_{t,r,s}^i \sim \mathcal{N}(m_{t,s}^i, \Sigma_{IS,s}^i)$ .
  3. *Mode Track  $X_{t,r,r}$* :  $\forall i$ , set  $X_{t,r,r}^i = m_{t,r}^i + \Sigma_{IS,r,s}^i (\Sigma_{IS,s}^i)^{-1} (X_{t,r,s}^i - m_{t,s}^i)$
  4. *Weight*:  $\forall i$ , compute  $w_t^i = \frac{\tilde{w}_t^i}{\sum_{j=1}^N \tilde{w}_t^j}$  where  $\tilde{w}_t^i = w_{t-1}^i \frac{p(Y_t | X_t^i) p(X_{t,r}^i | X_{t-1}^i, X_{t,s}^i)}{\mathcal{N}(X_{t,r}^i; m_t^i, \Sigma_{IS}^i)}$  where  $X_{t,r}^i = [X_{t,r,s}^i, X_{t,r,r}^i]$ .
  5. *Resample using any standard algorithm [2]*. Set  $t \leftarrow t + 1$  & go to step 1.
- 

As shown in [7], unimodality of  $p^{*,i}$  is ensured if the variance of the STP of  $X_{t,r}$  is small enough compared to distance between the modes of OL given  $X_{t,s}^i$  in any direction. Even if  $X_{t,s}$  is chosen so that this holds for most particles, at most times, the proposed algorithm will work [7].

Now, because of the LDSS property (at any given time, “most of the state change” occurs in a small number of dimensions, while the change in the rest of the state space is small),  $X_{t,r}$  can further be split into  $[X_{t,r,s}; X_{t,r,r}]$  so that the covariance of the STP of  $X_{t,r,r}$  is small enough to ensure that there is little error in approximating the conditional posterior of  $X_{t,r,r}$ ,  $p^{*,i}(X_{t,r,r})$ , by a Dirac delta function at its mode. We call this the Mode Tracking (MT) approximation of importance sampling (IS), or IS-MT. *When MT is combined with PF-EIS, the resulting algorithm is called PF-EIS-MT. It is summarized in Algorithm 1.*

The IS-MT approximation introduces some error in the estimate of  $X_{t,r,r}$  (error decreases with decreasing spread of  $p^{*,i}(X_{t,r,r})$ ). But it also reduces the importance sampling dimension from  $\dim(X_t)$  to  $\dim([X_{t,s}; X_{t,r,s}])$  (a significant reduction for large dimensional problems), thus improving the effective particle size. For carefully chosen dimension of  $X_{t,r,r}$ , this results in smaller total error, especially when the available number of particles,  $N$ , is small.

### 3. LANDMARK SHAPE TRACKING

A “configuration” is an ordered set of  $K$  landmark locations. In the 2D case, it is the x and y coordinates of the landmarks arranged as a  $K$  dimensional complex vector (x location + j y location) [13]. In the current work, we assume that the configuration is centered and there is zero translation over time. In this case the configuration is completely described by its shape,  $z$ , the global scale,  $e^s$ , and the global in-plane rotation,  $\theta$ . It is computed as  $ze^{s+j\theta}$  where  $j = \sqrt{-1}$ . The shape space is a  $K - 2$  dimensional manifold in  $C^K$  (complex  $K$ -

dimensional space)[13]. A tangent space to any point in shape space is thus a  $K - 2$  dimensional hyperplane in  $C^K$  [13].

#### 3.1. System Model

The system model involves a first order model on the logarithm of global scale,  $s_t$ , and on global 2D rotation angle,  $\theta_t$ , and a second order model on shape,  $z_t$  (which is equivalent to a first order model on “shape velocity” coefficients,  $c_t$ ). We describe the details below.

Motion in shape space requires moving by small amounts in the tangent space to the current shape. Denote the tangent space to  $z_t$  by  $T_{z_t}$ . Then  $z_{t+1}$  is obtained by moving  $z_t$  by an amount  $\tilde{v}_t \in T_{z_t}$  as follows:  $z_{t+1} = (1 - \tilde{v}_t^* \tilde{v}_t)^{1/2} z_t + \tilde{v}_t$  [13]. We refer to  $\tilde{v}_t$  as the “shape velocity”. Let  $(U_{t+1})_{K \times (K-2)}$  denote the matrix of basis directions that span  $T_{z_t}$ . Then  $\tilde{v}_t \in C^K$  is  $\tilde{v}_t = U_{t+1} \tilde{c}_t$  where  $\tilde{c}_t \in C^{K-2}$  contains the coefficients along  $U_{t+1}$ . Let  $c_t \in \mathbb{R}^{2K-4}$  denote the real vector obtained by arranging the real and imaginary components of  $\tilde{c}_t$  as a vector, i.e.  $\tilde{c}_t = S c_t$ , where  $S = [I_{K-2} \ j I_{K-2}]$  and  $I_q$  is the  $q \times q$  identity matrix.

To maintain correspondence between elements of  $c_t$  over time, we obtain  $U_t$  by starting with  $U_{t-1}$  and obtaining an orthogonal matrix that is perpendicular to  $z_{t-1}$  by using a Gram Schmidt procedure applied to the columns of  $U_{t-1}$  i.e.

$$U_{t,m} = [I_K - z_{t-1} z_{t-1}^* - \sum_{\tilde{m}=1}^{m-1} U_{t,\tilde{m}} U_{t,\tilde{m}}^*] U_{t-1,m},$$

$\forall m = 1, \dots, (K - 2)$  where  $U_{t,m}$  denotes the  $m^{\text{th}}$  column of  $U_t$ . Denote this transformation by  $U_t = g(U_{t-1}, z_{t-1})$ . Note that each  $U_t$  is also perpendicular to the vector of 1s,  $1_K$  (shape is translation normalized). This is enforced by making  $U_1$  perpendicular to  $1_K$ .

We assume a first order autoregressive (AR) model on  $c_t$ . We also assume an AR model on log-scale,  $s_t$ , and on rotation

angle,  $\theta_t$ . Thus the state consists of  $X_t = [s_t, \theta_t, c_t, z_t, U_t]$  and the dynamical model for  $X_t$  is

$$\begin{aligned} s_t &= \alpha_s s_{t-1} + \nu_{s,t}, \nu_{s,t} \sim n(0, \sigma_s) \\ \theta_t &= \alpha_\theta \theta_{t-1} + \nu_{\theta,t}, \nu_{\theta,t} \sim n(0, \sigma_\theta) \\ c_t &= A_c c_{t-1} + \nu_{c,t}, \nu_{c,t} \sim \mathcal{N}(0, \Sigma_c) \\ U_t &= g(U_{t-1}, z_{t-1}) \\ z_t &= f(z_{t-1}, U_t, c_t), \\ f(z_{t-1}, U_t, c_t) &\triangleq (1 - c_t^T c_t)^{1/2} z_{t-1} + U_t [I \ jI] c_t \end{aligned} \quad (5)$$

The resulting state transition prior (STP) is

$$\begin{aligned} p(X_t | X_{t-1}) &= \mathcal{N}(\alpha_s s_{t-1}, \sigma_s^2) \mathcal{N}(\alpha_\theta \theta_{t-1}, \sigma_\theta^2) \times \\ &\mathcal{N}(A_c c_{t-1}, \Sigma_c) \times \\ &\delta(U_t - g(U_{t-1}, z_{t-1})) \delta(z_t - f(z_{t-1}, U_t, c_t)) \end{aligned}$$

where  $\delta$  denotes the Dirac delta function. The STP of  $s_t, \theta_t$  is usually broad to allow for occasional large camera motion or zoom.

### 3.2. Observation Model

The landmarks' configuration is obtained from the shape, scale and rotation by the transformation  $h(s_t, \theta_t, z_t) = z_t e^{s_t + j\theta_t}$ . There are various ways to extract landmarks from image sequences - e.g. edge detection followed by extracting the  $K$  strongest edges or the edges closest to predicted landmark locations or using block optical flow estimation (KLT) [14].

The simplest observation model is of the form  $Y_t = h(s_t, \theta_t, z_t) + w_t$ ,  $w_t \sim \mathcal{CN}(0, \sigma_o^2 I)$  where  $w_t$  is a complex Gaussian noise vector [12]. This assumes that there is no background clutter: each of the  $K$  strongest edges or the  $K$  KLT-feature points are always generated by the true landmark location plus some error modeled as Gaussian noise. But this is often a simplistic model. It may happen that out of the  $K$  observed landmark locations, some landmark at some time is actually generated by clutter (e.g. if a true landmark is blurred or occluded, while a nearby clutter point has a stronger edge). We model this as follows: with a small probability  $p$ , the  $k^{th}$  landmark,  $Y_{t,k}$ , is generated by a clutter point and with probability  $(1 - p)$  it is generated by a Gaussian noise-corrupted actual landmark, i.e. for all  $k = 1, \dots, K$ ,

$$Y_{t,k} \sim (1 - p) \mathcal{CN}([h(s_t, \theta_t, z_t)]_k, \sigma_o^2) + p \mathcal{CN}(0, 100\sigma_o^2) \quad (6)$$

independent of other landmarks. We denote the resulting OL by  $p(Y_t | X_t) = p(Y_t | h(s_t, \theta_t, z_t))$ . The above is a valid model whenever the equivalent STP of  $[h]_k$  is broad (allows landmarks to be far from their predicted location). The resulting OL is heavy-tailed as a function of  $[h]_k$  with mode at  $Y_{t,k}$ . If the equivalent STP of  $[h]_k$  is broad (e.g. this will happen if the STP of scale,  $s_t$ , is broad), whenever  $Y_{t,k}$  is generated by clutter (i.e. is far from the STP mode), the resulting  $p^*$  will be multimodal. See Fig. 1 for a simple 1D example.

### 3.3. PF-EIS-MT for Landmark Shape Tracking

We first derive the PF-EIS algorithm and then develop PF-EIS-MT. Since the STP of  $s_t, \theta_t$  is usually broad (to allow for occasional large camera motion or zoom), we use  $X_{t,s} = [s_t, \theta_t]$  and  $X_{t,r} = [c_t, z_t, U_t]$ . Note that for the purpose of importance sampling only  $s_t, \theta_t, c_t$  are the "states" since  $z_t, U_t$  are deterministically computed from  $c_t$  and  $X_{t-1}$ . The particles of  $X_{t,s}$  are sampled from its STP, i.e. using the first two equations of (5). Conditioned on the sampled scale and rotation,  $X_{t,s}^i = [s_t^i, \theta_t^i]$ , it is much more likely that  $p^*$  is unimodal, i.e.  $p^{*,i}(c_t, U_t, z_t)$  defined below is unimodal

$$\begin{aligned} p^{*,i}(c_t, z_t, U_t) &= \zeta p(Y_t | h(s_t^i, \theta_t^i, z_t)) \mathcal{N}(A_c c_{t-1}^i, \Sigma_c) \times \\ &\delta(U_t - g(U_{t-1}^i, z_{t-1}^i)) \delta(z_t - f(z_{t-1}^i, U_t, c_t)) \end{aligned}$$

Since the pdfs of  $U_t, z_t$ , conditioned on  $c_t, X_{t-1}$ , are Dirac delta functions, the above simplifies to:

$$\begin{aligned} p^{*,i} &= \zeta p(Y_t | h(s_t^i, \theta_t^i, f(z_{t-1}^i, g^i, c_t))) \mathcal{N}(A_c c_{t-1}^i, \Sigma_c) \times \\ &\delta(U_t - g^i) \delta(z_t - f(z_{t-1}^i, g^i, c_t)) \\ &\triangleq p^{*,i}(c_t) \delta(U_t - g^i) \delta(z_t - f(z_{t-1}^i, g^i, c_t)) \end{aligned} \quad (7)$$

The importance sampling part of  $X_{t,r}$  is only  $c_t$ . We compute the importance density for  $c_t$  by approximating  $p^{*,i}(c_t)$  by a Gaussian at its unique mode. The mode is computed by minimizing  $L^i(c_t) = -\log p^{*,i}(c_t)$  defined below

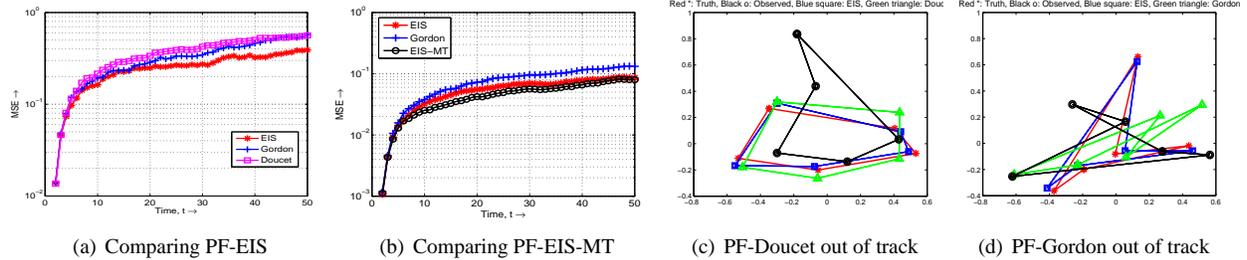
$$\begin{aligned} L^i(c_t) &= [-\log p(Y_t | h(s_t^i, \theta_t^i, f(z_{t-1}^i, g^i, c_t)))] + \\ &[-\log \mathcal{N}(c_t; A_c c_{t-1}^i, \Sigma_c)] \end{aligned} \quad (8)$$

Thus the PF-EIS algorithm for landmark shape tracking becomes

1. Imp. sample  $s_t^i \sim \mathcal{N}(\alpha_s s_{t-1}^i, \sigma_s^2)$ ,  $\theta_t^i \sim \mathcal{N}(\alpha_\theta \theta_{t-1}^i, \sigma_\theta^2)$
2. Compute  $m_t^i = \arg \min_{c_t} L^i(c_t)$ ,  $\Sigma_{IS}^i = [\nabla^2 L^i(m_t^i)]^{-1}$  where  $L^i$  is defined in (8).
3. Imp. sample  $c_t^i \sim \mathcal{N}(m_t^i, \Sigma_{IS}^i)$ .
4. Compute  $U_t^i = g(U_{t-1}^i, z_{t-1}^i)$  and  $z_t^i = f(z_{t-1}^i, U_t^i, c_t^i)$ .
5. Weight and Resample. Compute  $w_t^i = \frac{\tilde{w}_t^i}{\sum_{j=1}^N \tilde{w}_t^j}$  where  $\tilde{w}_t^i = w_{t-1}^i \frac{p(Y_t | h(s_t^i, \theta_t^i, z_t^i)) \mathcal{N}(c_t^i; A_c c_{t-1}^i, \Sigma_c)}{\mathcal{N}(c_t^i; m_t^i, \Sigma_{IS}^i)}$ .

PF-EIS-MT will have improved performance over PF-EIS when some directions of  $c_t$  have small enough variance to satisfy the IS-MT approximation (the total error due to replacing IS by MT for these directions is smaller than otherwise). Let  $c_t = \begin{bmatrix} c_{t,s} \\ c_{t,r} \end{bmatrix}$  where  $c_{t,r}$  satisfies IS-MT. Similarly split  $m_t^i$  and  $\Sigma_{IS}^i$  as shown in Algorithm 1. We set  $X_{t,r,s} = [c_{t,s}]$  and  $X_{t,r,r} = [c_{t,r}, U_t, z_t]$ . The PF-EIS-MT algorithm is the PF-EIS algorithm listed above, with step 3 replaced by

- 3'. Imp. sample  $c_{t,s}^i \sim \mathcal{N}(m_{t,s}^i, \Sigma_{IS,s}^i)$  and set  $c_{t,r}^i = m_{t,r}^i + \Sigma_{IS,r,s}^i (\Sigma_{IS,s}^i)^{-1} (c_{t,s}^i - m_{t,s}^i)$ . Set  $c_t^i = [c_{t,s}^i, c_{t,r}^i]$ .



**Fig. 2.** Landmark Shape Tracking results. Figs. 2(a), 2(b): MSE plots comparing PF-EIS, PF-EIS-MT with PF-Doucet and PF-Gordon. Figs. 2(c), 2(d): Examples where PF-Doucet, PF-Gordon lose track but PF-EIS does not.

#### 4. SIMULATION RESULTS

We simulated landmark shape change of a set of  $K = 5$  landmarks (a deforming pentagon) and tracked it using PF-EIS, PF-Gordon [1] and PF-Doucet [3] with  $N = 50$  particles. The mean squared error (MSE) plot averaged over 80 Monte Carlo runs is shown in Fig. 2(a). The initial shape,  $z_0$ , was a regular pentagon. The shape and global motion change of the configuration followed (5) with  $\Sigma_c = 0.0025I_6$ ,  $\sigma_s^2 = 0.0001$ ,  $\sigma_\theta^2 = 0.25$ ,  $A_c = 0.6I_6$ ,  $\alpha_s = 0.9$ ,  $\alpha_\theta = 0.9$ . The observations followed (6) with  $\sigma_o^2 = 0.04$  and  $p = 0.2$ .

Note that the STP of scale ( $e^{s_t}$ ) is a log-normal distribution and hence even  $\sigma_s^2 = 0.0001$  results in a fairly broad STP of  $e^{s_t}$ . The STP of  $\theta_t$  is also broad. Whenever one or more landmarks are generated by clutter, the OL of log-scale ( $s_t$ ) is either heavy-tailed with the wrong (outlier) mode or is multimodal. When many landmarks are generated by clutter, the same happens for the OL of  $\theta_t$ . This combined with a broad STP of  $s_t, \theta_t$ , results in multimodal  $p^*(s_t, \theta_t)$ . Whenever this happens, most particles of PF-Doucet end up sampling from a Gaussian about the wrong mode of  $p^*(s_t, \theta_t)$  or of  $p^*(s_t)$ , resulting in loss of track. See Fig. 2(c) for an example. But PF-EIS does not suffer from this problem since it samples from the STP of  $s_t, \theta_t$ . Also, since the STP of  $c_t$  is narrow compared to the OL,  $p^{*,i}(c_t)$  is usually unimodal and thus sampling from its Laplace approximation is valid. On the other hand, PF-Gordon often loses track because it samples all states from its STP, thus resulting in small effective particle size, especially when total particles  $N = 50$  itself is small.

In Fig. 2(b), we compare PF-EIS and PF-Gordon with PF-EIS-MT for a landmark shape sequence with  $K = 8$  landmarks. Only the first 2 dimensions of  $c_t$  had large variance, i.e.  $\Sigma_c = \text{diag}(0.0004I_2, 0.000025I_{10})$ . So we used the IS-MT approximation for the last 10 dimensions. Thus, for this 16-dimensional problem, the importance sampling dimension was only 6, which resulted in improved effective particle size when compared with PF-EIS (and greatly improved effective particle size compared to PF-Gordon). The tracking used  $N = 20$  particles and thus effective particle size is very critical here. Other model parameters were  $\sigma_s^2 = 0.0001$ ,  $\sigma_\theta^2 = 0.16$ ,  $A_c = 0.6I_{12}$ ,  $\alpha_s = 0.9$ ,  $\alpha_\theta = 0.9$ ,  $\sigma_o^2 = 0.01$ ,  $p = 0.2$ .

#### 5. REFERENCES

- [1] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to nonlinear/nongaussian bayesian state estimation," *IEE Proceedings-F (Radar and Signal Processing)*, pp. 140(2):107–113, 1993.
- [2] S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for on-line non-linear/non-gaussian bayesian tracking," *IEEE Trans. Sig. Proc.*, vol. 50, no. 2, pp. 174–188, Feb. 2002.
- [3] A. Doucet, "On sequential monte carlo sampling methods for bayesian filtering," in *Technical Report CUED/F-INFENG/TR. 310, Cambridge University Department of Engineering*, 1998.
- [4] L. Tierney and J. B. Kadane, "Accurate approximations for posterior moments and marginal densities," *J. Amer. Stat. Assoc.*, vol. 81, no 393, pp. 82–86, March 1986.
- [5] R. van der Merwe, N. de Freitas, A. Doucet, and E. Wan, "The unscented particle filter," in *Advances in Neural Information Processing Systems 13*, Nov 2001.
- [6] V. Cevher and J. H. McClellan, "Proposal strategies for joint state-space tracking with particle filters," in *IEEE Intl. Conf. Acoustics, Speech, Sig. Proc. (ICASSP)*, 2005.
- [7] N. Vaswani, "Pf-eis & pf-mt: New particle filtering algorithms for multimodal observation likelihoods and large dimensional state spaces," in *IEEE Intl. Conf. Acoustics, Speech, Sig. Proc. (ICASSP)*, 2007.
- [8] T. Schn, F. Gustafsson, and P.J. Nordlund, "Marginalized particle filters for nonlinear state-space models," *IEEE Trans. Sig. Proc.*, 2005.
- [9] R. Chen and J.S. Liu, "Mixture kalman filters," *Journal of the Royal Statistical Society*, vol. 62(3), pp. 493–508, 2000.
- [10] D.G. Kendall, D. Barden, T.K. Carne, and H. Le, *Shape and Shape Theory*, John Wiley and Sons, 1999.
- [11] N. Vaswani, A. RoyChowdhury, and R. Chellappa, "Shape Activity": A Continuous State HMM for Moving/Deforming Shapes with Application to Abnormal Activity Detection," *IEEE Trans. Image Proc.*, pp. 1603–1616, October 2005.
- [12] N. Vaswani and R. Chellappa, "Nonstationary shape activities," in *IEEE Conf. Decision and Control (CDC)*, December 2005.
- [13] I.L. Dryden and K.V. Mardia, *Statistical Shape Analysis*, John Wiley and Sons, 1998.
- [14] Jianbo Shi and Carlo Tomasi, "Good features to track," in *IEEE Conf. on Comp. Vis. and Pat. Recog. (CVPR)*, 1994, pp. 593–600.