PARTICLE FILTERS FOR INFINITE (OR LARGE) DIMENSIONAL STATE SPACES- PART 2

Namrata Vaswani

Dept. of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011
namrata@iastate.edu

ABSTRACT

We study particle filtering algorithms for tracking on infinite (in practice, large) dimensional state spaces. Particle filtering (Monte carlo sampling) from a large dimensional system noise distribution is computationally expensive. But, in most large dim tracking applications, it is fair to assume that "most of the state change" occurs in a small dimensional basis and the basis itself may be slowly time varying (approximated as piecewise constant). We have proposed a PF algorithm with basis change detection and re-estimation steps that uses this idea. The implicit assumptions in defining this algorithm are very strong. We study here the implications of weaker assumptions and how to handle them. We propose to use a simple modification of the asymptotically stable Adaptive Particle Filter to handle errors in estimating the basis dimension.

1. INTRODUCTION

This paper is part-2 of a two part paper [1]. We propose particle filtering (PF) [2, 3] algorithms for tracking on infinite (in practice, large) dimensional state spaces. Tracking is defined as the problem of causally estimating a hidden state sequence, $\{X_t\}$ (that is Markovian), from a sequence of observations, $\{Y_t\}$ that satisfy the HMM assumption ($X_t \rightarrow Y_t$ is a Markov chain at each t). The state space may not be a vector space, we assume it to be a separable metric space (Polish space). Particle filtering (PF) on large dimensional state spaces is expensive. But in most large dim tracking applications, it is fair to assume that at any time t, "most of the state change" occurs in a small dimensional "effective basis", and the basis dimension is either constant or slowly time varying (approximated as piecewise constant). In [1, 4], we have proposed an efficient algorithm for the constant K-dim effective basis case. It assumes that conditioned on $X_{t,s} \in \mathbb{R}^K$ and on X_{t-1}, the posterior of X_t is unimodal. In [4], the algorithm was used to track deforming object contours from image sequences (observation). We used a 6-dim space of affine deformations as the effective basis for contour deformation. But in certain other applications, such as in medical imaging, this assumption may not hold, and there may be two (or more) contours of interest at roughly the same "affine location" (have same affine parameters). In other domains, there may not even be a natural constant basis approximation that can be used.

To handle such applications, we consider a generalization of the above assumption that allows the basis dimension for $X_{t,s}$ to be slowly time varying. The state space model is repeated in Section 2. In Section 3, we propose a modified PF algorithm that includes a basis change detection and re-estimation step. Its application to contour tracking is shown in Figure 1. We analyze the implicit assumptions in defining this algorithm. In Section 4, we discuss how they can be relaxed and propose an easy-to-implement modification of the Adaptive Particle Filter [5] to handle errors in estimating the basis dimension, K, which is a piecewise constant parameter. Application to contour tracking and conclusions are given in Section 5.

The problem of tracking on large dim state spaces occurs in many domains. It has been studied by many authors in the context of tracking outer contours of deforming objects from image sequences [6, 7, 4, 8]. Another large dim tracking problem is estimating the spectro-temporal receptive fields which are time-frequency plots (short time Fourier transform at a set of time instants) that characterize the time varying input-output transfer function of auditory neurons. An example STRF size is $15 \times 13 = 195$, even though domain knowledge tells us that only a small part of it undergoes significant changes for a given time period. A third possible application is tracking optical flow. Optical flow [9], denoted $u(x,y), v(x,y)$, for an image at time t, $I_t(x,y)$ gives the motion of every point x, y in one frame interval, i.e. it is defined by $I_{t+1}(x+u, y+v) = I_t(x,y)$. Thus, $C_t(x,y) = u(x,y), v(x,y)$ has a dimension which is twice the size of the image, even though motion is highly correlated.

Note that the algorithms given in this work, are also applicable to problems of time varying but finite dimensional state tracking. PF for time varying state dimensions has been studied by many authors. [10, 11] study algorithms for filtering clean speech from an observed noisy speech signal which is modeled using an order-K AR model whose coefficients as well as order are slowly time varying. In [10] a slow time varying partial correlation model is proposed which models β size blocks of any time series, its partial correlation coefficient vector (PARCOR) and the order (length of PARCOR) as the state vector. It assumes the PARCOR vector is slow time varying and hence estimates it sparsely (every few time instants) and interpolates these samples to get the PARCOR vector at each t. This idea is similar to our proposed interpolation of the velocity vector (see Section 2). Many techniques based on MCMC [12, 11] or Reversible Jump MCMC[13] or treating model order as a discrete Markov chain [10] have been proposed for model order selection in PF algorithms. Our algorithm can also be used for tracking the shape of a time-varying number, K, of landmarks (objects)[14].

2. STATE SPACE MODEL

We use the subscript t to denote the discrete time instants. p denotes probability density functions (pdfs). We briefly describe the state space model detailed in [1]. Consider a state space model with state $X_t = [C_t, v_t]$ where v_t denotes the time "derivative" of C_t. Assume that $C_t = C_t(p)$ where p belongs to a compact subset of $\mathbb{R}^n$. Assume that C_t belongs to a Polish space $\mathcal{S}$ (a complete separable metric space). v_t now denotes the time "derivative" of C_t (defined in the corresponding tangent space at C_t, denoted $\mathcal{TS}_{C_t}$). Thus v_t belongs to a vector space. In implementing any algorithm for infinite dim state spaces, the number of points at which C_t is defined is always large but finite (and can change at every t). For example, if the parameter $p \in [0,1]$, C_t is defined at M_t points $p = 0, 1/M_t, ...1$ at time t. *Hence in the rest of this paper, we assume that $\mathcal{S}$ is a large but finite dim space with dimension M_t at time t.* We split v_t as $v_t = [v_{t,s}, v_{t,r}]$, where $v_{t,s} \in \mathbb{R}^K$ denotes the coefficients along the K basis directions representing the K-dim subspace ($\mathcal{S}_s$), in which "most of the state change" is assumed to occur, and $v_{t,r}$ denotes the state change in the rest of the state space ($\mathcal{S}_r$) which is assumed "small". The basis directions for $\mathcal{S}_s$ are denoted by $B_s(p) = [b_1(p), ..b_K(p)]$ and the basis for $\mathcal{S}_r$ are denoted by $B_r(p)$. We assume the following general form of the discrete time state dynamics (with time discretization interval denoted as τ):

$$C_t(p) = \hat{C}_t(p) + g(\hat{C}_t, B_r(p)v_{t,r}) \quad (1)$$

$$\hat{C}_t(p) = C_{t-1}(p) + \tau g(C_{t-1}, B_s(p)v_{t,s}), \; B_s \triangleq B(C_{t-1}) \quad (2)$$

$$v_{t,s} = f_t(\tau, v_{t-1,s}) + \nu_{t,s}, \;\; \nu_{t,s} \sim p_{v,t,s}(.) \quad (3)$$

$$v_{t,r} = \nu_{t,r}, \; \nu_{t,r} \sim p_{v,t,r}(.) \quad (4)$$

g defines the mapping from $\mathcal{TS}_{C_{t-1}}$ (tangent space at C_{t-1}) to $\mathcal{S}$. The dimension of $\mathcal{S}_s$, K, can be fixed or slowly time varying (modeled as piecewise constant). We have assumed a first order Markov model on $v_{t,s}$ while $v_{t,r} = \nu_{v,t,r} \sim p_{v,t,r}(.)$ is independent over t, and so can be excluded from the state space. *Thus, in this paper we assume the state to be $X_t = [C_t, v_{t,s}]$.* The pdf $p_{v,t,r}(.)$ is unimodal.Assume an observation model where the observations, Y_t depend only on C_t, i.e. the observation likelihood, $p(Y_t|X_t) = p(Y_t|C_t)$ and where $C_t \rightarrow Y_t$ is a Markov chain for each t. The observation likelihood, $p(Y_t|C_t)$, obtained from above model can, in general, be multimodal.

Example: Consider object contour tracking from image sequences [15]: $C_t(p),\; p \in [0,1]$ denotes a parametrization of a 2D closed contour, $B(C_{t-1})(p) = [b_1(p), ..b_K(p)]$ denotes the K B-spline basis functions defined on the contour C_{t-1}, $v_{t,s}$ denotes the velocity of the K control points of the B-spline. $g(C, Bv) = Bv\,\mathbf{N}(C)$ where $\mathbf{N}(C)$ denotes the normal to C. We use a linear Gauss-Markov model for $v_{t,s}$, i.e. $f_t(v) = Av$, $p_{v,t,s}$ is a zero mean Gaussian. Also, Y_t is the image at t and $p(Y_t|C_t) \propto e^{-E_{cv}(Y_t,C_t)}$ where E_{cv} denotes the Chan-Vese energy[4]. We show results for contour tracking using Algorithm 1 in Figure 1.

3. PARTICLE FILTER WITH TIME VARYING BASIS

The basic idea of PF with time-varying basis was introduced in [1].

Fact 1 *Since $\mathcal{S}$ is a Polish space, $\forall\, \epsilon > 0$, any $C_t \in \mathcal{S}$ can be approximated by a $\hat{C}_t = C_{t-1} + g(C_{t-1}, B_K v_s)$, s.t. $d(C_t, \hat{C}_t) < \epsilon$ by choosing $K = K(\epsilon, C_t, C_{t-1})$ large enough and choosing $v_s = v_s(K, C_t, C_{t-1}) \in \mathbb{R}^K$.*

Let us replace distance by average distance i.e. we look for one K and one v_s (depending on C_{t-1}) that works for all C_t on average. Also, we consider a piecewise constant effective basis dimension, i.e. the same K works for all $C_{t-1} \in \mathcal{S}$ and for all $t \in [T_1, T_2]$, i.e.

Assumption 1 *Given a Δ^* and a time interval $[T_1, T_2]$, $\exists\; K = K(\Delta^*, [T_1, T_2])$ s.t. for every $C_{t-1} \in \mathcal{S}$ and $\forall t \in [T_1, T_2]$, $\exists\; v_{t,s} = v_{t,s}(K, C_{t-1})$ so that $\Delta_t = \mathbb{E}[d(C_t, \hat{C}_t)|C_{t-1}, v_{t,s}] \leq \Delta^*$.*

In addition, we also need the assumptions discussed in [1] that ensure that $p(C_t|C_{t-1}, v_{t,s}, Y_t)$, with $v_{t,s}$ belonging to the current K dim basis, is "effectively" unimodal (it has only one mode with significantly nonzero pdf value). Based on these assumptions, we modified Algorithm 1 in [1] to include a basis change detection step at every t and a basis dimension estimation step whenever a change is detected. The new algorithm is summarized in Algorithm 1. Basis dimension change detection and new dimension estimation has be application dependent, e.g. [15]. It has been discussed in [1].

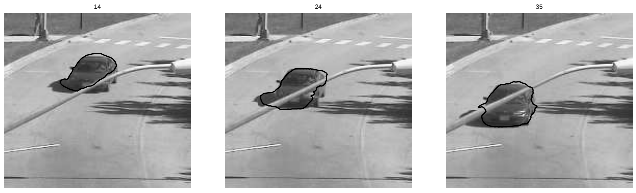
Algorithm 1 Particle Filter with Time Varying Basis

1. At $t = 0$, for $i = 1$ to N, set $C_0^{(i)} = C_0$, sample $v_{0,s}^{(i)} \sim \mathcal{N}(v_{0,s}; 0, \Sigma_0)$. Set $X_0^{(i)} = [C_0^{(i)}, v_{0,s}^{(i)}]$
2. At any t, assume that $p(X_{t-1}|Y_{1:t-1}) \approx \sum_{i=1}^{N}(1/N)\delta(X_{t-1} - X_{t-1}^{(i)})$ is available.
3. **Importance Sampling:** For $i = 1$ to N,
   (a) Sample $\nu_{t,s}^{(i)} \sim p_{v,t,s}(.)$. Compute $v_{t,s}^{(i)}$ using (3) and $\hat{C}_t^{(i)}$ using (2).
   (b) Compute $m_t^{(i)} = \arg\max_{C_t \in \mathcal{S}} p(Y_t|C_t)p(C_t|C_{t-1}, v_{t,s})$ as explained in Section 3 of [1].
   (c) Set $C_t^{(i)} = m_t^{(i)}$, since $\Sigma \approx 0$ for large dim spaces [1]. Set $X_t^{(i)} = [C_t^{(i)}, v_{t,s}^{(i)}]$
4. **Weighting :** For $i = 1$ to N,
   (a) Set $\tilde{w}_t^{(i)} = \tilde{w}_{t-1}^{(i)} \frac{p(Y_t|C_t^{(i)})p(C_t^{(i)}|C_{t-1}^{(i)}, v_{t,s}^{(i)})}{\mathcal{N}(m_t^{(i)}; m_t^{(i)}, \Sigma)}$. Note denominator is a constant (can be removed).
   (b) Set $w_t^{(i)} = \frac{\tilde{w}_t^{(i)}}{\sum_{j=1}^{N} \tilde{w}_t^{(j)}}$

   Now $p(X_t|Y_{1:t}) \approx \sum_{i=1}^{N}(w_t^{(i)})\delta(X_t - X_t^{(i)})$, where δ is the Dirac delta function
5. **Detect Basis Change:** Detect if basis change required. If yes, then go to step 6, else go to step 7.
6. **Change Basis:** Compute the new basis dimension K_{new}. For $i = 1$ to N,
   (a) Compute the new basis $B_{new,i} = B_{K_{new}}(C_t^{(i)})$ and old basis $B_{t,i} = B_K(C_t^{(i)})$.
   (b) Project $v_{t,s}^{(i)}$ into new basis as: $v_{t,s}^{(i)} \longleftarrow (B_{new,i}^T B_{new,i})^{-1} B_{new,i}^T B_{t,i} v_{t,s}^{(i)}$.
   (c) Set $K_t \longleftarrow K_{new}$, $B_{t,i} \longleftarrow B_{new,i}$.
7. **Resampling:** For all $i = 1$ to N:
   (a) Sample the index $I(i) \sim \{i, w_t^{(i)}\}_{i=1}^{N}$
   (b) Set $X_t^{(i)} \longleftarrow X_t^{(I(i))}$, $\tilde{w}_t^{(i)} \longleftarrow 1$, $w_t^{(i)} \longleftarrow 1/N$.

   Now $p(X_t|Y_{1:t}) \approx \sum_{i=1}^{N}(1/N)\delta(X_t - X_t^{(i)})$.
8. Set $t \longleftarrow t+1$, go to step 3

Implicit Assumptions: Algorithm 1 implicitly assumes the following: First, it assumes that there is no error in estimating the new basis dimension. This is a very strong assumption. Another very strong assumption is that there is zero delay in detecting the basis change and also that there is no error in projecting $v_{t,s}^{(i)}$ in the new basis. Also, if at a basis change time, T_1, the basis dimension is reduced, it means that a certain part of the state space remains constant during $[T_1, T_2]$. For example if the dimension reduces from K



(a) Left part (b) Full (starting) (c) Full

Fig. 1. Algorithm 1 for tracking through an occluding street light

to $K-1$, it means that posterior value of one component of $v_{t,s}$ (denoted $(v_{t,s})_K$) is approximately zero with zero variance or it means that $(v_{t,s})_K$ is deterministically known conditioned on some other component of $v_{t,s}$ and $Y_{1:t}$. Both of these imply that the region of C_t which is affected by $(v_{t,s})_K$ (denote this region by $C_t(K)$) remains constant during $[T_1, T_2]$, except for small deterministic deformation introduced by mode finding. Summarizing, Algorithm 1 can be shown to converge, if in addition to the mixing assumptions for the the state space model [16], the following hold:

Assumption 2 *Assume that there is zero error in estimating the new basis dimension.*

Assumption 3 *There is zero delay in detecting the need for basis change and there is zero error in projecting $v_{t,s}^{(i)}$ in the new basis.*

Assumption 4 *At $t = T_1$, if the basis dimension reduces from K to $K-1$, then the posterior of $(v_{t,s})_K$ conditioned on the rest of the components of $v_{t,s}$, i.e. $p((v_{t,s})_K | Y_{1:t}, (v_{t,s})_{1:K-1})$ has either converged to a Dirac delta function at zero, $\delta((v_{t,s})_K)$ or to $\delta((v_{t,s})_K - f_{det}((v_{t,s})_{1:K-1}, Y_{1:t}))$ where f_{det} is some deterministic function of its arguments. Also, either of these imply that the posterior of the region of C_t affected by $(v_{t,s})_K$ (denoted $C_t(K)$) has converged to a delta function too, i.e. it remains constant during $[T_1, T_2]$.*

4. HANDLING WEAKER ASSUMPTIONS

Assume that assumption 2 does not hold and K_{new} is a random variable with a prior distribution (that depends on K_{old}) at $t = T_1$. If one simply treated K as part of the state space, then the PF resampling step will introduce the usual problems of resampling for static parameters (loss of a good particle due to resampling, new particle cannot be generated because no randomness) [5]. To avoid these problems, we use a modification of the adaptive PF [5] for K.

Adaptive PF: In [5], the author defines an M particle Adaptive Particle Filter (APF) (the standard particle filter without the resampling step) for the unknown static parameter. For each particle of the unknown static parameter, they run a regular PF for the rest of the state space. Since there is no resampling between different static parameter particle sets, the weight, $\tilde{W}_\tau^m$, of a static parameter particle depends on several ($q(M)$, where $q(M)$ is an increasing function of M) past observations, i.e. $\tilde{W}_\tau^m = \prod_{t=\tau-q(M)+1}^{\tau} \sum_{j=1}^{N/M} \tilde{w}_t^{m,j}$. The two main assumptions required are: (i) the static parameter belongs to a compact set (its prior distribution has compact support) and (ii) conditioned on the value of the parameter, the regular PF for the rest of the state space is uniformly convergent. Under these main assumptions, it has been proved [5] that for t large enough, the posterior for the static parameter converges to a delta function at its true value (or at a set of true values). Consequently, the estimated posterior of the rest of the state space also converges to the true posterior.

Adaptive PF for K: Consider a weaker version of Assumption 2:

Assumption 5 *Assume*

1. *The error in estimating the change in the new basis dimension from the previous dimension is bounded, i.e. $\exists\, A < \infty$ s.t. $|K_{new} - K_{old}| < A$.*
2. *The Adaptive PF for K converges in finite time, denoted by T_{conv}, to one or (in case of multiple targets) a finite number of possible true values.*
3. *The time interval between two basis change times, i.e. the duration $T_2 - T_1 + 1$, is larger than T_{conv}.*
4. *For any value of K, the PF for the rest of the state space is uniformly convergent.*

We treat K as a piecewise static parameter and run a simple modification of the Adaptive PF for it (summarized in Algorithm 2). Given: At a basis change time, $t = T_1$, only $M = 2$ modes of K have survived (corresponding to two observable targets), each with N_m, $m = 1, 2$ particles ($N_1 + N_2 = N$). (i) For each mode, estimate the P new values of K (for e.g. $K = K_{old}-1, K_{old}, K_{old}+1, P = 3$). (ii) "Split each PF into P parts", i.e. resample $\{j, w_t^{m,j}\}_{j=1}^{N_m}$ N_m/P times (instead of the usual N_m times) and allocate new particles accordingly. Thus, starting at $t = T_1$, we run $M' = M * P$ PFs, each with $N'_m = N_{\lceil m/P \rceil}/P$ particles. (iii) For each PF, run the importance sampling and weighting steps, evaluate $p_{y,t}^m$ and the posterior weight of the m^{th} PF, $\tilde{W}_t^{(m)}$, and then resample within each PF.

If $(t - T_1)\ mod\ T_{conv} = 0$ (APF convergence time), "eliminate zero modes" and "re-allocate N to non-zero modes". "Zero modes" will be indices m whose posterior weight $W_t^{(m)}$ is negligibly small. Do this as follows: Sample M times from $\{m, W_t^{(m)}\}_{m=1}^{M}$ and set N'_m=number of times index m gets sampled. Thus for all the "zero modes", N'_m will be zero. Now "re-allocate N to non-zero modes", i.e. allocate N'_m particles to the m^{th} PF: resample N'_m times from $\{j, w_t^{m,j}\}_{j=1}^{N_m}$ (instead of the usual N_m times) and allocate new particles accordingly. We give the stepwise algorithm in Algorithm 2.

Delay in Detecting Basis Change: Now Assumption 3 that there is no delay in detecting basis change is also unrealistic. A more practical assumption is: There is a bounded delay in detecting the new basis change and there can be error in re-estimating the new velocity. Under this assumption, one cannot show convergence of the particle filter, since there is a finite duration of system model error, whose effect can only go to zero asymptotically with time (if at all). But one can modify the stability results of [14] (which are based on the results of [16]) to show stability (asymptotic stability under strong assumptions) of the total filtering error (system model error plus particle filtering error), i.e. the total error at any t remains bounded by a function of the initial error.

Non-Uniform Sampling: Until now, we have used uniform sampling of an element of the tangent space of C_t, to generate a K dim subspace. For example, for contour tracking, at the start of any basis change time, we allocate K B-spline knot locations, uniformly on the contour and define velocity at the K control points (in between these knots) as the K dim subspace. These may become non-uniformly spaced as the contour deforms. K is reduced when a set of knots come very close to each other. In this case, Assumption 4 ($(v_{t,s})_K$ converges to a deterministic function of $(v_{t,s})_{1:K-1}$ when basis dimension reduces) holds.

Algorithm 2 Adaptive PF for K, with fixed particle budget N

$t = T_1$, M PFs, each with N_m particles, $\sum_{m=1}^{M} N_m = N$.

1. **Basis Change:** For $m = 1$ to M do
   (a) Detect need for basis change for m^{th} PF
   (b) If needed, evaluate P possible K_{new} values
   (c) "Split" m^{th} PF into P parts, i.e.: For $j = 1$ to N_m/P do, Sample $I(j) \sim \{j, w_t^{m,j}\}_{j=1}^{N_m}$ and set $Z_t^{m,j} \longleftarrow Z_t^{m,I(j)}$. Also, $\tilde{W}_t^{(m)} \longleftarrow \tilde{W}_t^{\lceil m/P \rceil}$
   (d) New value of M is $M' = M * P$ and $N'_m = \frac{N_{\lceil m/P \rceil}}{P}$.
2. **Regular PF with $W_t^{(m)}$ Calc:** For $m = 1$ to M do,
   (a) Perform importance sampling and weighting (steps 3 and 4 of Algorithm 1) for all N_m particles.
   (b) Compute $p_{y,t}^m = \sum_{j=1}^{N_m} \tilde{w}_t^{m,j}$, $\tilde{W}_t^{(m)} = \tilde{W}_{t-1}^{(m)} \frac{p_{y,t}^m}{p_{y,t-q(M)}^m}$ and $W_t^{(m)} = \frac{\tilde{W}_t^{(m)}}{\sum_{k=1}^{M} \tilde{W}_t^{(k)}}$
   (c) Resample N_m times from $\{j, w_t^{m,j}\}_{j=1}^{N_m}$ (step 7 of Algorithm 1)

   Now $p(X_t|Y_{1:t}) \approx \sum_{m=1}^{M} W_t^{(m)} \sum_{j=1}^{N_m} \frac{1}{N_m} \delta(X_t - Z_t^{m,j})$
3. $t \longleftarrow t+1$. If $(t - T_1)\ mod\ T_{conv} \neq 0$, go to step 2, else **eliminate "zero modes":**
   (a) Sample M times from $\{m, W_t^{(m)}\}_{m=1}^{M}$ and set N'_m=number of times index m gets sampled. For all "zero modes", N'_m will be zero. Thus new value of M is M'=number of non-zero modes.
   (b) **"Re-allocate N to non-zero modes":**
   For $m = 1$ to M' do,
   For $j = 1$ to N'_m do, Sample $I(j) \sim \{j, w_t^{m,j}\}_{j=1}^{N_m}$ and assign $Z_t^{m,j} \longleftarrow Z_t^{m,I(j)}$.
   (c) Go to step 1.

But to make the algorithm more efficient, one could assume a non-uniform sampling, for e.g. use prior information to allocate knots only in regions where deformation is known to occur. This will introduce more static parameters (to decide where to sample the state vector) into the system model. Also, with this sampling, Assumption 4 may not hold. In principle, both these situations can be handled by using an Adaptive PF for the sampling locations and for $C_t(K)$. $C_t(K)$ is also like a piecewise static parameter with prior given by its posterior at the beginning of the interval. We will address this as part of future work.

5. CONCLUSIONS AND APPLICATIONS

This paper is part-2 of [1]. We have presented algorithms for tracking on infinite (or large) dimensional state spaces, whose effective basis dimension is assumed to be piecewise constant with time and small. The above assumption allowed us to define a particle filter with a small dimensional effective basis at any time. It required Monte Carlo sampling from only this small dim space and is practically implementable. We studied the implicit assumptions in defining this algorithm and how to relax them. We handle errors in the basis dimension, K, by treating it as a piecewise static parameter and using the Adaptive PF (APF) [5] for K during each time interval. We have modified the original APF algorithm by resampling it every T_{conv} time instants. By letting the duration between two basis change times be larger than T_{conv}, we ensure that we resample at least once between two basis change times. This prevents the number of particles or the number of APF modes from blowing up. The above algorithm was originally motivated by the problem of contour tracking (explained in Section 2). We show an example of contour tracking using Algorithm 1 in Figure 1.

6. REFERENCES

[1] N.Vaswani, A.Yezzi, Y.Rathi, and A.Tannenbaum, "Particle filters for infinite (or large) dimensional state spaces-part 1," in *IEEE ICASSP*, 2006.

[2] 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.

[3] S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for on-line non-linear/non-gaussian bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, Feb. 2002.

[4] Y. Rathi, N. Vaswani, A. Tannenbaum, and A. Yezzi, "Particle filtering for geometric active contours and application to tracking deforming objects," in *IEEE CVPR*, 2005.

[5] A. Papvasiliou, "A uniformly convergent adaptive particle filter," *Journal of Applied Probability (to appear)*.

[6] R.W. Brockett and A. Blake, "Estimating the shape of a moving contour," in *IEEE CDC*, 1994.

[7] J. Jackson, A. Yezzi, and S. Soatto, "Tracking deformable moving objects under severe occlusions," in *IEEE CDC*, 2004.

[8] M. Niethammer and A. Tannenbaum, "Dynamic level sets for visual tracking," in *IEEE CDC*, 2004.

[9] A.M.Tekalp, *Digital Video Processing*, Prentice Hall, 1995.

[10] W. Fong and S. Godsill, "Sequential monte carlo simulation of dynamical models with slowly varying parameters: Application to audio," in *IEEE ICASSP*, 2002.

[11] A. Doucet, N. deFreitas, and Eds Gordon, N., *Sequential Monte Carlo Methods in Practice*, Springer, 2001.

[12] C. Andrieu, J. de Freitas, and A. Doucet, "Sequential mcmc for bayesian model selection," in *IEEE Higher Order Statistics Workshop, Ceasarea, Israel*, 1999.

[13] A. Doucet, N. Gordon, and V. Krishnamurthy, "Particle filters for state estimation of jump markov linear systems," *IEEE Trans. Signal Processing*, pp. 613–624, 2001.

[14] N. Vaswani, "Change detection in partially observed nonlinear dynamic systems with unknown change parameters," in *American Control Conference (ACC)*, 2004.

[15] "Blind conference submission," .

[16] F. LeGland and N. Oudjane, "Stability and Uniform Approximation of Nonlinear Filters using the Hilbert Metric, and Application to Particle Filters," *Technical report, RR-4215, INRIA*, 2002.