PARTICLE FILTERS FOR INFINITE (OR LARGE) DIMENSIONAL STATE SPACES- PART 1

Namrata Vaswani, Anthony Yezzi**, Yogesh Rathi**, Allen Tannenbaum***

*Dept. of ECE, Iowa State University, Ames, IA 50011, namrata@iastate.edu
**School of ECE, Georgia Tech, Atlanta, GA 30332

ABSTRACT

We propose particle filtering algorithms for tracking on infinite (or large) dimensional state spaces. We consider the general case where state space may not be a vector space, we assume it to be a separable metric space (Polish space). In implementation, any such space is approximated by a finite but large dimensional vector, whose dimension may vary at every time. Monte carlo sampling from a large dimensional system noise distribution is computationally expensive. Also, the number of particles required for accurate particle filtering increases with the number of independent dimensions of the system noise, making particle filtering even more expensive. But as long as the number of independent system noise dimensions is small, even if the total state space dimension is very large, a particle filtering algorithm can be implemented. In most large dim applications, it is fair to assume that "most of the state change" occurs in a small dimensional basis, which may be fixed or slowly time varying (approximated as piecewise constant). We use this assumption to propose efficient PF algorithms. These are analyzed and extended in [1].

1. INTRODUCTION

We propose practically implementable particle filtering [2, 3] algorithms for tracking on infinite (or large) dimensional state spaces. Tracking is the problem of causally estimating a hidden state sequence, $\{X_t\}$ (that is Markovian with state transition pdf $p(X_t|X_{t-1})$), from a sequence of observations, $\{Y_t\}$, that satisfy the Hidden Markov Model (HMM) assumption ($X_t \to Y_t$ is a Markov chain for each t, with observation likelihood denoted $p(Y_t|X_t)$). Some examples of applications where large dim state spaces occur are (i) tracking the boundary (contour) of deforming objects in image sequences (e.g. space sequences such as consecutive slices in medical imaging or time sequences), whose dimension, in the worst case, may be as large as that of the image (space filling curve); (ii) tracking the Spectro-Temporal Receptive Fields (STRFs) [4] which are time-frequency plots used to characterize the time varying input-output transfer function of the auditory neuron in [4], a typical STRF dimension is $15 \times 13 = 195$; or (iii) tracking optical flow [5] as a function of time. Optical flow, for an image at time t, $I_t(x, y)$ gives the motion of every point x, y during one frame time. Its dimension is twice the image dimension. But in all these cases, at any time t, the number of dimensions in which most of the change (contour deformation or STRF intensity change or optical flow magnitude change) occurs is much smaller.

The problem of tracking on large dim state spaces has been studied by many researchers in the context of contour tracking, [6, 7, 8, 9]. Many of these tracking algorithms, e.g. the ones given in [7, 10], can be understood as approximate "posterior mode trackers" for a state space model following the HMM assumptions.They approximate the posterior, $p(X_t|Y_{1:t})$, as a Dirac delta function at its largest mode, i.e. $p(X_t|Y_{1:t}) \approx \delta(X_t - m_t)$ where $m_t = \arg\max_{X_t} p(X_t|Y_{1:t}) = \arg\max_{X_t} p(Y_t, X_t|Y_{1:t-1})$ (the largest mode m_t of $p(X_t|Y_{1:t})$ is also the global maximizer of $p(Y_t, X_t|Y_{1:t-1})$). $p(Y_t, X_t|Y_{1:t-1})$ is evaluated using Bayes recursion and the delta function approximation of $p(X_{t-1}|Y_{1:t-1})$ as: $p(Y_t, X_t|Y_{1:t-1}) = \int_{x_{t-1}} p(Y_t|X_t)p(X_t|x_{t-1})p(x_{t-1}|Y_{1:t-1})dx_{t-1} \approx p(Y_t|X_t)p(X_t|X_{t-1} = m_{t-1})$. Thus m_t is evaluated as

$$m_t = \arg\max_{X_t} p(Y_t|X_t)p(X_t|X_{t-1} = m_{t-1})$$

This maximizer is typically found by starting with $X_t = m_{t-1}$ as initial guess and running "some" iterations of gradient descent to minimize $-\log p(Y_t|X_t)$. The implicit assumption in doing this is that $p^* \triangleq p(X_t|X_{t-1}, Y_t)$, has a single local maximizer (is unimodal). If $p(Y_t|X_t)$ is multimodal (e.g. contours of multiple moving objects in the image, or presence of spurious edges or large intensity variations resulting in temporary false modes), p^* will be unimodal only if the spread of $p(X_t|X_{t-1})$ is small enough. For large dim state spaces, this may not hold. When p^* is not unimodal, it is not clear how to find the global maximizer. Also, one would like to track all the "significant" modes, not just the largest mode.

But, in applications involving tracking on large dim state spaces, it is fair to assume that conditioned on a small part of the current state, denoted $X_{t,s}$, the above, i.e. $p(X_t|X_{t-1}, X_{t,s}, Y_t)$, is unimodal. This, as we discuss in Section 3, follows from the assumption that in most large dim tracking applications, at a given time, "most of the state change" occurs in a small dim basis. For many applications, this small dim basis can be assumed to be fixed and known. The contour tracking algorithm of [8] can be understood as one application of this idea. There we chose $X_{t,s}$ to be the 6-dim space of affine deformations - which approximates a global deformation model for the contour - and we used a particle filter to track $X_{t,s}$. Conditioned on $X_{t,s}$ and Y_t (image at t), the non-affine deformation can be assumed to be unimodal - this assumption is valid in many problems of object tracking, since typically multiple contours are separated by translation or scale. So, we used a posterior mode tracker for tracking the non-affine deformation. This idea as we explain later, can be understood as a modification of the algorithm of [13]. But in certain other applications, such as in medical imaging, there may be two (or more) contours of interest at roughly the same "affine location". In other applications such as the STRF tracking problem, there may not be a single constant basis (like affine basis for contour tracking) where most of the state change occurs. To handle such applications, we consider a generalization of the above assumption that allows the dimension of $X_{t,s}$ to be slow time varying.

We present particle filtering algorithms for tracking on large dim state spaces based on the above ideas. Our algorithm is an efficient importance sampling technique that can also be applied to reduce complexity of smaller dim tracking problems as long as they satisfy

Assumptions 2 and 3. The general form of the state space model is explained in Section 2. The algorithm for fixed and known basis is given in Section 3. The algorithm for time varying basis is introduced in Section 4, more details are given in [1]. Design issues, application to contour tracking and conclusions are discussed in Section 5.

2. STATE SPACE MODEL

State Space: We use the subscript t to denote the discrete time instants. Consider a state space model with state $X_t = [C_t, v_t]$ where v_t denotes the time "derivative" of C_t. C_t can be a large finite dim vector, or an infinite dim vector, $C_t = C_t(p),\ p \in [a, b]$. Or $C_t(p)$ can itself be a finite dim vector (e.g. $C_t(p) = [C_t^x(p), C_t^y(p)]^T$ with $p \in [a, b]$). Also, C_t may not even lie in a vector space (e.g. $[C_t^x(p), C_t^y(p)]^T$ may denote one parametrization of a contour [11]). Also, p can itself belong to a compact subset of $\mathbb{R}^2$ or $\mathbb{R}^3$ etc (e.g. optical flow, $C_t(x, y) = [u(x, y), v(x, y)]^T,\ \ x \in [0, a],\ y \in [0, b]$). Thus to incorporate all these cases, assume that $C_t = C_t(p)$, with p belonging to a compact subset of $\mathbb{R}^n$. Assume that C_t belongs to a Polish space $\mathcal{S}$ (a complete separable metric space) with distance metric d. 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.*

State Dynamics (System Model): 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 $B_r(p)$. The basis directions can be a function of the previous state C_{t-1}. Their dimension can also vary with t (piecewise constant with t). But to simplify notation, we do not use the subscript t with B_s.

We assume the following general form of the discrete time state dynamics (with time discretization interval denoted as τ):

$$\begin{aligned} C_t(p) &= \hat{C}_t(p) + g(\hat{C}_t, B_r(p)v_{t,r}) && (1)\\ \hat{C}_t(p) &= C_{t-1}(p) + \tau g(C_{t-1}, B_s(p)v_{t,s}),\ B_s \triangleq B(C_{t-1}) && (2)\\ v_{t,s} &= f_t(\tau, v_{t-1,s}) + \nu_{t,s},\ \ \nu_{t,s} \sim p_{v,t,s}(.) && (3)\\ v_{t,r} &= \nu_{t,r},\ \nu_{t,r} \sim p_{v,t,r}(.) && (4) \end{aligned}$$

g defines the mapping from $\mathcal{TS}_{C_{t-1}}$ (tangent space at C_{t-1}) to $\mathcal{S}$. For e.g., if C is a planar contour [11], $g(C, v(p)) = v(p)\,\mathbf{N}(C(p))$ where $\mathbf{N}(C(p))$ denotes the normal to C at point $C(p)$. For this application, B_s can be a K=6-dim basis for affine deformation as in [8] or it can be a K-dim B-spline basis for interpolating contour velocity at K control points as in [12].

The dimension of $\mathcal{S}_s$, K, can be fixed or slowly time varying (modeled as piecewise constant). The system noise sequences $\{\nu_{v,t,s},\ \nu_{v,t,r}\}$ are independent of each other and over time. 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.

Observation Model: 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 $C_t \to Y_t$ is a Markov chain for each t. The observation likelihood, $p(Y_t|C_t)$, obtained from above model can, in general, be multimodal (e.g. multiple target tracking problems or problems where multiple false target modes may get generated due to sensor error or background clutter).

3. CONSTANT FINITE-DIM BASIS

We assume here that $\mathcal{S}_s$ is constant for all t. For many applications, such as the contour tracking problems shown in [8] (where $\mathcal{S}_s$ is taken to be the 6-dim basis of affine deformation), this assumption suffices. We have considered extensions of the algorithms proposed here to time-varying basis in Section 4 and [1].

The optimal importance sampling (IS) distribution (one that minimizes the variance of weights $w_t^{(i)}$ conditioned on particles $x_{1:t-1}^{(i)}$ and past observations $Y_{1:t-1}$) for particle filtering has been shown to be $p^* \triangleq p(X_t|X_{t-1}, Y_t)$ in [13] and other works. But this cannot be evaluated analytically for most state space models. In [13], the authors suggest approximating p^* by a Gaussian about its mode, when it is unimodal. When $p(Y_t|X_t)$ is multimodal, p^* will be unimodal only if the spread of $p(X_t|X_{t-1})$ is small. For a large dim state space, the spread of $p(X_t|X_{t-1})$ may not be small enough in all dimensions to ensure unimodality of p^* and hence the algorithm of [13] cannot be used. *But if the change in the "rest of the state space", $\mathcal{S}_r$, is "small enough" (quantified in Assumption 3), then, $p^{**} \triangleq p(X_t|X_{t-1}, v_{t,s}, Y_t) = p(C_t|\hat{C}_t, Y_t)$ can be shown to be unimodal.* Under this assumption, we propose the following modification to the algorithm of [13]: For each particle i, (i) sample $v_{t,s}^{(i)}$ from its state transition pdf, $p(v_{t,s}|v_{t-1,s}^{(i)})$ (defined by (3)), in order to sample possible multiple modes of X_t, and (ii) sample $C_t^{(i)}$ from a Gaussian approximation to p^{**} about its mode (denoted $m_t^{(i)}$). The Gaussian approximation of p^{**} is denoted $\mathcal{N}(C_t; m_t^{(i)}, \Sigma^{(i)})$, where the mode, $m_t^{(i)} = m_t(C_{t-1}^{(i)}, v_{t,s}^{(i)}, Y_t)$, is obtained as

$$\begin{aligned} m_t^{(i)} &= \arg\min_{C_t} L(C_t) \triangleq -\log p^{**} + const \\ &= -\log p(Y_t|C_t) - \log p(C_t|\hat{C}_t^{(i)}) + const. \quad (5) \end{aligned}$$

The $M_t \times M_t$ covariance matrix, $\Sigma^{(i)}$, can be chosen as suggested in [13] to be $\Sigma^{(i)} = L''(m_t^{(i)})$. In summary, we propose to use as importance sampling density:

$$q(X_t|X_{t-1}^{(i)}, Y_t) = p(v_{t,s}|v_{t-1,s}^{(i)})\mathcal{N}(C_t; m_t^{(i)}, \Sigma^{(i)}) \qquad (6)$$

Note that when $\mathcal{S}$ is not a vector space, we use $\mathcal{N}(C_t; m_t, \Sigma)$ to denote the following sampling scheme: Sample $v_t \sim \mathcal{N}(0, \Sigma)$ and compute $C_t = m_t + \tau g(m_t, B_r v_t)$. The stepwise algorithm is summarized in Algorithm 1. We now discuss sufficient conditions under which p^{**} will be unimodal.

Sufficient Conditions for Unimodality of p^{**}

Assumption 1 *To simplify the derivation below, assume $p_{v,t,r}(v_{t,r}) = \mathcal{N}(v_{t,r}; 0, \Sigma_r)$ with $\Sigma_r = \Delta\, I$, i.e. the change in $\mathcal{S}_r$ is spatially i.i.d. Gaussian distributed with variance Δ. The derivation can be easily generalized to any unimodal pdf.*

Define $D_\Sigma(C_2, C_1)$ as[1]

$$D_\Sigma(C_2, C_1) \triangleq v_r^T \Sigma^{-1} v_r,\ \ v_r \triangleq B_r^T g^{-1}(C_1, C_2 - C_1) \qquad (7)$$

[1] If Σ is singular, Σ^{-1} denotes the pseudo-inverse.

Algorithm 1 Particle Filter with Efficient Importance Sampling

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 $m_t^{(i)}$ defined in (5).
   (c) If $\Sigma \approx 0$ is valid, set $C_t^{(i)} = m_t^{(i)}$, else sample $C_t(i) \sim \mathcal{N}(C_t; m_t^{(i)}, \Sigma)$.
   (d) Set $X_t^{(i)} = [C_t^{(i)}, v_{t,s}^{(i)}]$
4. **Weighting :** For $i = 1$ to N,
   (a) If $\Sigma \approx 0$ is valid, $\tilde{w}_t^{(i)} = \tilde{w}_{t-1}^{(i)} p(Y_t|C_t^{(i)}) p(C_t^{(i)}|C_{t-1}^{(i)}, v_{t,s}^{(i)})$, else $\tilde{w}_t^{(i)} = \tilde{w}_{t-1}^{(i)} \frac{p(Y_t|C_t^{(i)}) p(C_t^{(i)}|C_{t-1}^{(i)}, v_{t,s}^{(i)})}{\mathcal{N}(C_t^{(i)}; m_t^{(i)}, \Sigma)}$.
   (b) Set $w_t^{(i)} = \frac{\tilde{w}_t^{(i)}}{\sum_{j=1}^{N} \tilde{w}_t^{(j)}}$

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

Then, since $\Sigma_r = \Delta\, I$,

$$p(C_t|\hat{C}_t) = p_{v,t,r}(B_r^T g^{-1}(\hat{C}_t, C_t - \hat{C}_t)) \propto \exp[-\frac{D_I(C_t, \hat{C}_t)}{2\Delta}] \quad (8)$$

Thus

$$L(C_t) = E(C_t) + \frac{D_I(C_t, \hat{C}_t)}{2\Delta}, \;\; E(C_t) \triangleq -\log p(Y_t|C_t) \quad (9)$$

Assumption 2 *Assume*

1. *$E(C_t)$ and $D_I(C_t, \hat{C}_t)$ are continuously differentiable functions of C_t.*
2. *D_Σ is a strictly convex function of C_t.*
3. *E is Lipschitz continuous everywhere.*
4. *$\hat{C}_t$ lies in a region where E is locally convex, i.e. the Hessian $E''(\hat{C}_t) > 0$ (positive definite).*

Definition 1 (Region R) *Denote by C_{min}^* the minimizer of E whose distance $D_I(C_{min}^*, \hat{C}_t)$ from $\hat{C}_t$ is the least among all minimizers of E. Define the* **region** *R as the largest continuous region around C_{min}^* that contains $\hat{C}_t$ and where E is locally convex, i.e. the Hessian $E''(C_t) \geq 0$ (positive definite).*

Fact 1 *With Assumption 2, the region R always exists and $L(C_t)$ is strictly convex inside R. Also, the minimizer of L inside R, m, satisfies $E(C_{min}^*) \leq E(m) \leq E(\hat{C}_t)$.*

Now if we can assume that, there is no extremum point of L outside R (denoted R^c), then m will be the only minimizer of L (i.e. p^{**} will be unimodal). A sufficient condition for this is that for every $C \in R^c$, there exists some p for which $(\nabla_C L)(p) = (\nabla_C E)(p) + (\nabla_C D_I)(p)/2\Delta \neq 0$. The only places where $(\nabla_C L)(p)$ can equal 0 for all p, will be in regions where $(\nabla_C E)(p)$ and $(\nabla_C D_I)(p)$ differ in sign for all p, i.e. $(\nabla_C E)(p)(\nabla_C D_I)(p) < 0,\ \forall p$. But if Δ is such that it is strictly smaller than $\max_p \frac{|(\nabla_C D_I)(p)|}{|(\nabla_C E)(p)|}$ for all C in these regions, then $(\nabla_C L)(p)$ will never be zero for all p in these regions.

Assumption 3 *Let $A \triangleq \{C \in R^c : (\nabla_C E)(p)(\nabla_C D_I)(p) < 0,\ \forall p\}$. Assume*

$$\Delta < \min_{C \in A} \max_p \frac{|(\nabla_C D_I)(p)|}{|(\nabla_C E)(p)|} \triangleq \Delta^* \quad (10)$$

Fact 2 *By assumption 2, Δ^* is strictly positive. If assumptions 1, 2 and 3 are true, then L has a single minimizer, denoted by m, which lies inside R. Equivalently $p^{**} = p(C_t|\hat{C}_t, Y_t) \propto p(Y_t, C_t|\hat{C}_t) = e^{-L}$ is unimodal.*

Practical choice of Σ in (6) for Large Dim State Spaces: Since conditioning reduces average variance, $\mathbb{E}[Covar[p(C_t|\hat{C}_t, Y_t)]] \leq Covar[p(C_t|\hat{C}_t)] = \Delta\, I \leq \Delta^* I$. But $\Sigma \approx Covar[p(C_t|\hat{C}_t, Y_t)]$. Thus in situations where Δ^* is small, the average (taken over Y_t) eigenvalues of Σ will be still smaller. Also, it is observed that the value of Δ^* decreases as the dimension M_t increases (for fixed K). This happens because the minimization in (10) will be performed over a larger dim space. Thus for large M_t, the approximation $\Sigma \approx 0$ is valid. *When M_t is large, importance sampling from $\mathcal{N}(C_t; m_t, \Sigma)$ is approximately equivalent to deterministically setting the particle $C_t^{(i)} = m(C_{t-1}^{(i)}, v_{t,s}^{(i)}, Y_t)$.*

Evaluating/Approximating the mode of p^{**}: If assumptions 2 and 3 hold, p^{**} is unimodal. Also, by Fact 1, its mode, $m = m(C_{t-1}, v_{t,s}, Y_t)$, satisfies $E(C_{min}^*) \leq E(m) \leq E(\hat{C}_t)$. Thus, by starting with $C_t = \hat{C}_t$ as initial guess and running gradient descent to minimize E, there will always be an iteration number k for which $C_t = m$, if the iteration scaling is small enough. *Thus, an efficient way to evaluate m is to start with $C_t = \hat{C}_t$ as initial guess and to run k iterations of gradient descent to minimize E.* In most practical applications, it is not possible to evaluate k. But as demonstrated in the experiments of [8], a heuristic choice often suffices to give an approximation to the mode.

4. TIME VARYING BASIS

As discussed earlier, the assumption of a fixed basis is restrictive in certain situations. Here we attempt to relax it with that of a piecewise constant with time basis.

Fact 3 *Since $\mathcal{S}$ is a Polish space, by definition: $\forall\ \epsilon > 0$, for any $C_t, C_{t-1} \in \mathcal{S}$, $\exists\ K = K(\epsilon, C_t, C_{t-1})$ large enough and $v_s = v_s(K, C_t, C_{t-1}) \in \mathbb{R}^K$, s.t. $d(C_t, \hat{C}_t) < \epsilon$, where $\hat{C}_t = C_{t-1} + g(C_{t-1}, B_K v_s)$. B_K is a K-dim basis for any countable dense subset of $\mathcal{TS}_{C_{t-1}}$.*

Now 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 4 *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 $\forall t \in [T_1, T_2]$, $\exists\ v_{t,s} = v_{t,s}(K, C_{t-1})$ s.t.*
$\mathbb{E}[d(C_t, C_{t-1} + g(C_{t-1}, B_K v_{t,s}))|C_{t-1}, v_{t,s}] \leq \Delta^*$.

In addition, we also need the assumptions discussed in the above section that ensure that $p(C_t|C_{t-1}, v_{t,s}, Y_t)$ is unimodal. Under these assumptions, one can modify Algorithm 1, to include a basis change detection step at every t and a basis dimension estimation step whenever a change is detected. Also, when the basis dimension changes, the velocity from the previous time step needs to be re-evaluated in the new basis. The stepwise algorithm is given and analyzed in [1].

5. DESIGN ISSUES, APPLICATIONS AND CONCLUSIONS

Basis Choices and Dimension Estimation: We would like to define a K dim basis to approximate an element of the tangent space of an element of an infinite dim Polish space. The tangent space element is an infinite dim vector. This can be also be understood as a way of sampling a continuous function at K points to define a K-dim subspace. Some possible choices are: Fourier basis (uniformly discretizes the Fourier transform of the function), B-spline basis (provides a piecewise polynomial approximation that is local in space) or the wavelet basis (local in both time and frequency). Any of these can be used to either uniformly discretize an element of $\mathcal{TS}_{C_{t-1}}$ which is the tangent space to $\mathcal{S}$ at C_{t-1} or it can be used to uniformly discretize an element of the vector space that embeds $\mathcal{S}$. In the second case, the interpolated velocity needs to be projected into $\mathcal{TS}_{C_{t-1}}$ (which is a subspace of the embedding vector space).

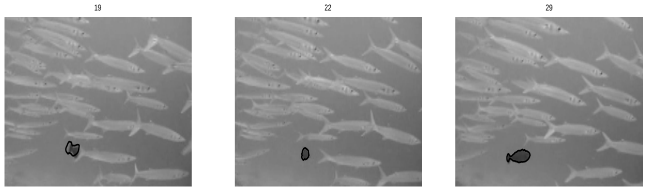
The basis dimension should be chosen to guarantee a certain spatial resolution for the change in C_t and needs to increase when the "spatial frequency content" in the "change of C_t" increases. The specific rules need to be application dependent though. But since spatial frequency in "change of C_t" cannot be measured (without tracking the complete v_t), heuristics/prior knowledge can be used.

Basis Change Detection: Detecting the need to change basis will also be very application dependent. Some ways to detect a need to change basis are: (i) when expected posterior distance of the current C_t from a reference state C_{ref}) exceeds a threshold, or (ii) when tracking error of the observation increases, or (iii) when local tracking error for certain parts of the observation vector increases or (iv) simply change basis at fixed time intervals.

Application to Contour Tracking: In [8], we used Algorithm 1 for contour tracking. We show results for tracking a moving and deforming fish in Figure 1. The fixed dim basis was a 6-dim affine basis for contour deformation. Thus $B(C_{t-1})(p)$ was

$$B_t = \mathbf{N}^T \begin{bmatrix} C^x_{t-1}(p) C^y_{t-1}(p) & 0 & 0 & 10 \\ 0 & 0 & C^x_{t-1}(p) C^y_{t-1}(p) & 01 \end{bmatrix}$$

and $g(C_{t-1}, B_t v_{t,s})(p) = B_t(p) v_{t,s}\,\mathbf{N}$ where $\mathbf{N} = \mathbf{N}(C_{t-1}(p))$ denotes the normal to C_{t-1} at $C_{t-1}(p)$. The observation, Y_t, was the image at time t, the energy functional, E was the Chan-and-Vese energy and this satisfies $E''(C) > 0$ in the neighborhood of a minimizer, as long as the spatial image gradient is non-zero. If different object contours in an image are sufficiently separated by affine parameters (e.g. translation or scale), the minimizers of E are located far apart. The practical implication of assumptions on d is: choose a continuously differentiable and strictly convex distance function d and then ensure that K is large enough so that expected value of this distance is "small".



(a) Occlusion starting (b) Recovered (c) Deformation

Fig. 1. Tracking a deforming fish through partial occlusions

Conclusions: We have proposed two practically implementable PF algorithms for tracking on infinite (or large) dim state spaces. The first assumes that there exists a known and constant finite dim basis in which most of the state change occurs. The second algorithm allows this basis to be slowly time varying (piecewise constant). We discuss the implicit assumptions in defining this algorithm and how they can be relaxed in [1]. Also, note that Algorithm 1 suggests an efficient importance sampling strategy (a generalization of [13]) that can be used whenever Assumptions 2 and 3 are satisfied (even if the state space dimension is finite and small). It can also be understood as an approximate Rao-Blackwellization [14] technique.

6. REFERENCES

[1] N.Vaswani, "Particle filters for infinite (or large) dimensional state spaces-part 2," 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] "Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex," *Nature Neuroscience*, vol. 6 n. 11, pp. 1216–1223, 2003.
[5] A.M.Tekalp, *Digital Video Processing*, Prentice Hall, 1995.
[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] 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.
[9] M. Niethammer and A. Tannenbaum, "Dynamic level sets for visual tracking," in *IEEE CDC*, 2004.
[10] S.Haker, G.Sapiro, and A.Tannenbaum, "Knowledge-based segmentation of sar data with learned priors," *IEEE Tran. on Image Processing*, vol. 9, pp. 298–302, 2000.
[11] G. Sapiro, *Geometric Partial Differential Equations and Image Processing*, Cambridge University Press, January 2001.
[12] "Blind conference submission," .
[13] 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.
[14] R. Chen and J.S. Liu, "Mixture kalman filters," *Journal of the Royal Statistical Society*, vol. 62(3), pp. 493–508, 2000.