Variable Resolution Particle Filter

Vandi Verma, Sebastian Thrun and Reid Simmons
Carnegie Mellon University
5000 Forbes Ave. Pittsbugh PA 15217
{vandi, thrun, reids}@cs.cmu.edu

Abstract

Particle filters are used extensively for tracking the state of non-linear dynamic systems. This paper presents a new particle filter that maintains samples in the state space at dynamically varying resolution for computational efficiency. Resolution within statespace varies by region, depending on the belief that the true state lies within each region. Where belief is strong, resolution is fine. Where belief is low, resolution is coarse, abstracting multiple similar states together. The resolution of the statespace is dynamically updated as the belief changes. The proposed algorithm makes an explicit bias-variance tradeoff to select between maintaining samples in a biased generalization of a region of state space versus in a high variance specialization at fine resolution. Samples are maintained at a coarser resolution when the bias introduced by the generalization to a coarse resolution is outweighed by the gain in terms of reduction in variance, and at a finer resolution when it is not. Maintaining samples in abstraction prevents potential hypotheses from being eliminated prematurely for lack of a sufficient number of particles. Empirical results show that our variable resolution particle filter requires significantly lower computation for performance comparable to a classical particle filter.

1 Introduction

A number of problems in AI and robotics require estimation of the state of a system as it changes over time from a sequence of measurements of the system that provide noisy (partial) information about the state. Particle filters have been extensively used for Bayesian state estimation in nonlinear systems with noisy measurements [Isard and Blake, 1998; Fox *et al.*, 1999; Doucet *et al.*, 2001]. The Bayesian approach to dynamic state estimation computes a posterior probability distribution over the states based on the sequence of measurements. Probabilistic models of the change in state over time (state transition model) and relationship between the measurements and the state capture the noise inherent in these domains. Computing the full posterior in real time is often intractable. Particle filters approximate the posterior distribution over the states with a set of particles drawn from the posterior. The particle approximation converges to the true Bayesian posterior in the limit as the number of samples go to infinity. For real-time state estimation it is impractical to have an infinitely large number of particles. With a small number of particles, the variance of the particle based estimate can be high, particularly when there are a large number of possible state transitions.

This paper presents a new particle filter, the variable resolution particle filter (VRPF), for tracking large state spaces efficiently with low computation. The basic idea is to represent a set of states[1] rather than a single state with a single particle and thus increase the number of states (or hypotheses) that may be tracked with the same number of particles. This makes it possible to track a larger number of states for the same computation, but results in a loss of information that differentiates between the states that are lumped together and tracked in abstraction. We formalize this tradeoff in terms of bias and variance. Tracking in abstraction at a coarser resolution introduces a bias over tracking at a finer resolution, but it reduces the variance of the estimator given a limited number of samples. By dynamically varying the resolution in different parts of the state space to minimize the bias-variance error, a finite set of samples are used in a highly efficient manner. Initial experimental results support our analysis and show that the bias is more than compensated for by the reduction in variance from not eliminating potential hypotheses prematurely for lack of a sufficient number of particles.

2 Bayesian Filtering

We denote the multivariate state at time t as X_t and measurements or observations as O_t. As is commonly the case, we concentrate on the discrete time, first order Markov formulation of the dynamic state estimation problem. Hence, the state at time t is a sufficient statistic of the history of measurements, i.e. $p(X_t|X_{0:t-1}) = p(X_t|X_{t-1})$ and the observations depend only on the current state, i.e. $p(O_t|X_{0:t}) = p(O_t|X_t)$. In the Bayesian framework, the posterior distribution at time t, $p(X_{0:t}|O_{1:t})$, includes all the available information upto time t and provides the optimal solution to the state estimation problem. In this paper we are interested in estimating recursively in time a marginal of this distribution, the filtering distribution, $p(X_t|O_{1:t})$, as the measurements be-

[1] or regions of a continuous state space

come available. Based on the Markovian assumption, the recursive filter is derived as follows:

$$
\begin{aligned}
&p(X_t|O_{1:t}) \\
&= \eta_t \; p(O_t|X_t, O_{1:t-1}) \; p(X_t|O_{1:t-1}) \\
&= \eta_t \; p(O_t|X_t) \; p(X_t|O_{1:t-1}) \\
&= \eta_t \; p(O_t|X_t) \int \{ \; p(X_t|O_{1:t-1}, X_{t-1}) \\
&\quad p(X_{t-1}|O_{1:t-1}) \; \} \; dX_{t-1} \\
&= \eta_t \; p(O_t|X_t) \int p(X_t|X_{t-1}) \; p(X_{t-1}|O_{1:t-1}) \; dX_{t-1}
\end{aligned}
\tag{1}
$$

This process is known as Bayesian filtering, optimal filtering or stochastic filtering and may be characterized by three distributions: (1) a transition model $p(X_t|X_{t-1})$, (2) an observation model $p(O_t|X_t)$, and, (3) an initial prior distribution, $p(X_0)$. In a number of applications, the state space is too large to compute the posterior in a reasonable time. Particle filters are a popular method for computing tractable approximations to this posterior.

2.1 Classical Particle Filter

A Particle filter (PF) [Metropolis and Ulam, 1949; Gordon *et al.*, 1993; Kanazawa *et al.*, 1995] is a Monte Carlo approximation of the posterior in a Bayes filter. PFs are nonparametric and can represent arbitrary distributions. PFs approximate the posterior with a set of N fully instantiated state samples or particles, $P_t = \{x_t^{[1]} \dots x_t^{[N]}\}$ as follows:

$$
p(X_t|O_{1:t}) \approx \frac{1}{N} \sum_{i=1}^{N} \delta(X_t - x_t^{[i]}) \tag{2}
$$

where $\delta(.)$ denotes the Dirac delta function. It can be shown that as $N \to \infty$ the approximation in (2) approaches the true posterior density [Doucet and Crisan, 2002]. In general it is difficult to draw samples from, $p(X_t|O_{1:t})$; instead, samples are drawn from a more tractable distribution, $q(.)$, called the proposal or importance distribution. Each particle is assigned a weight, $w^{[i]}$ to account for the fact that the samples were drawn from a different distribution [Rubin, 1988; Rubinstein, 1981]. There are a large number of possible choices for the proposal distribution, the only condition being that its support must include that of the posterior. The common practice is to sample from the transition probability, $p(X_t|X_{t-1})$, in which case the importance weight is equal to the likelihood, $p(O_t|X_t)$. The experiments in this paper were performed using this proposal distribution. The PF algorithm is:

set $P_t = P_t^{\mathrm{aux}} = \emptyset$
for $i = 1$ to N do
    pick the i-th sample $x_{t-1}^{[i]} \in P_{t-1}$
    draw $x_t^{[i]} \sim p(X_t|x_{t-1}^{[i]})$
    set $w_t^{[i]} = p(O_t|x_t^{[i]})$
    add $\langle x_t^{[i]}, w_t^{[i]} \rangle$ to P_t^{aux}
end
for $i = 1$ to N do
    draw $x_t^{[i]}$ from P_t^{aux}
    with probability proportional to $w_t^{[i]}$
    add $_t^{[i]}$ to P_t
end

3 Variable Resolution Particle Filter

A well known problem with particle filters is that a large number of particles are often needed to obtain a reasonable approximation of the posterior distribution. For real-time state estimation maintaining such a large number of particles is typically not practical. However, the variance of the particle based estimate can be high with a limited number of samples, particularly when the process is not very stochastic and parts of the state space transition to other parts with very low, or zero, probability. Consider the problem of diagnosing locomotion faults on a robot. The probability of a stalled motor is low and wheels on the same side generate similar observations. Motors on any of the wheels may stall at any time. A particle filter that produces an estimate with a high variance is likely to result in identifying some arbitrary wheel fault on the same side, rather than identifying the correct fault.

The variable resolution particle filter introduces the notion of an abstract particle, in which particles may represent individual states or sets of states. With this method a single abstract particle simultaneously tracks multiple states. A limited number of samples are therefore sufficient for representing large state spaces. A bias-variance tradeoff is made to dynamically refine and abstract states to change the resolution, thereby abstracting a set of states and generalizing the samples or specializing the samples in the state into the individual states that it represents. As a result reasonable posterior estimates can be obtained with a relatively small number of samples. In the example above, with the VRPF the wheel faults on the same side of the rover would be aggregated together into an abstract fault. Given a fault, the abstract state representing the side on which the fault occurs would have high likelihood. The samples in this state would be assigned a high importance weight. This would result in multiple copies of these samples on resampling proportional to weight. Once there are sufficient particles to populate all the refined states represented by the abstract state, the resolution of the state would be changed to the states representing the individual wheel faults. At this stage, the correct hypothesis is likely to be included in this particle based approximation at the level of the individual states and hence the correct fault is likely to be detected.

For the variable resolution particle filter we need: (1) A variable resolution state space model that defines the relationship between states at different resolutions, (2) an algorithm for state estimation given a fixed resolution of the state space, (3) a basis for evaluating resolutions of the state space model, and (4) and algorithm for dynamically altering the resolution of the state space.

3.1 Variable resolution state space model

We could use a directed acyclic graph (DAG) to represent the variable resolution state space model, which would consider every possible combination of the (abstract)states to aggregate or split. But this would make our state space exponentially large. We must therefore constrain the possible combinations of states that we consider. There are a number of ways to do this. For the experiments in this paper we use a multi-layered hierarchy where each physical (non-abstract) state only exists along a single branch. Sets of states with similar state transition and observation models are aggregated together at each level in the hierarchy. In addition to the physical state set $\{X_k\}$, the variable resolution model, M consists of a set of abstract states $\{S_j\}$ that represent sets of states and or other abstract states.

$$S_j = \begin{cases} \{X_k\} \\ \cup_i S_i \end{cases} \tag{3}$$

Figure 1(a) shows an arbitrary Markov model and figure 1(b) shows an arbitrary variable resolution model for 1(a). Figure 1(c) shows the model in 1(b) at a different resolution.

We denote the measurement at time t as z_t and the sequence of measurement $z_1 \dots z_t$ as z^t. From the dynamics, $p(x_t|x_{t-1})$, and measurement probabilities $p(z_t|x_t)$, we compute the stationary distribution (Markov chain invariant distribution) of the physical states $\pi(x_k)$ [Behrends, 2000].

3.2 Belief state estimation at a fixed resolution

This section describes the algorithm for estimating a distribution over the state space, given a fixed resolution for each state, where different states may be at different fixed resolutions. For each particle in a physical state, a sample is drawn from the predictive model for that state $p(x_t|x_{t-1})$. It is then assigned a weight proportional to the likelihood of the measurement given the prediction, $p(z_t|x_t)$. For each particle in an abstract state, S_j, one of the physical states, x_t, that it represents in abstraction is selected proportional to the probability of the physical state under the stationary distribution, $\pi(x_t)$. The predictive and measurement models for this physical state are then used to obtain a weighted posterior sample. The particles are then resampled proportional to their weight. Based on the number of resulting particles in each physical state a Bayes estimate with a Dirichlet(1) prior is obtained as follows:

$$\hat{p}(x_t|z^t) = \frac{n(x_t) + \pi(x_t)}{\mid N_t \mid +1}\ , \quad \sum_{x_t} \pi(x_t) = 1 \tag{4}$$

where, $n(x_t)$ represents the number of samples in the physical state x_t and $\mid N_t \mid$ represents the total number of particles in the particle filter. The distribution over an abstract state S_j at time t is estimated as:

$$\hat{p}(S_j|z^t) = \sum_{x_t \in S_j} \hat{p}(x_t|z^t) \tag{5}$$

3.3 Bias-variance tradeoff

The loss l, from a particle based approximation $\hat{p}(x_t|z^t)$, of the true distribution $p(x_t|z^t)$ is:

$$\begin{aligned} l &= \mathbb{E}[p(x_t|z^t) - \hat{p}(x_t|z^t)]^2 \\ &= \{p(x_t|z^t) - \mathbb{E}[\hat{p}(x_t|z^t)]\}^2 + \\ &\quad \{\hat{p}(x_t|z^t)^2 - \mathbb{E}[\hat{p}(x_t|z^t)]^2\} \\ &= b(\hat{p}(x_t|z^t))^2 + v(\hat{p}(x_t|z^t)) \end{aligned} \tag{6}$$

where, $b(\hat{p}(x_t|z^t))$ is the bias and $v(\hat{p}(x_t|z^t))$ is the variance.

The posterior belief state estimate from tracking states at the resolution of physical states introduces no bias. But the variance of this estimate can be high, specially with small sample sizes. An approximation of the sample variance at the resolution of the physical states may be computed as follows:

$$v(x_t) = \hat{p}(x_t|z^t)\ \frac{\hat{p}(x_t|z^t)\ [1 - \hat{p}(x_t|z^t)]}{n(x_t) + \pi(x_t)} \tag{7}$$

The loss of an abstract state S_j, is computed as the weighted sum of the loss of the physical states $x_t \in S_j$, as follows[2]:

$$l(S_j) = \sum_{x_t \in S_j} \hat{p}(x_t|z^t)\ l(x_t) \tag{8}$$

The generalization to abstract states biases the distribution over the physical states to the stationary distribution. In other words, the abstract state has no information about the relative posterior likelihood, given the data, of the states that it represents in abstraction. Instead it uses the stationary distribution to project its posterior into the physical layer. The projection of the posterior distribution $\hat{p}(S_j|z^t)$, of abstract state S_j, to the resolution of the physical layer $\tilde{p}(x_t|z^t)$, is computed as follows:

$$\tilde{p}(x_t|z^t) = \frac{\pi(x_t)}{\pi(S_j)}\ \hat{p}(S_j|z^t) \tag{9}$$

where, $\pi(S_j) = \sum_{x \in S_j} \pi(x)$.

As a consequence of the algorithm for computing the posterior distribution over abstract states described in section 3.2, an unbiased posterior over physical states x_t is available at no extra computation, as shown in equation (4). The bias $b(S_j)$, introduced by representing the set of physical states $x_t \in S_j$, in abstraction as S_j is approximated as follows:

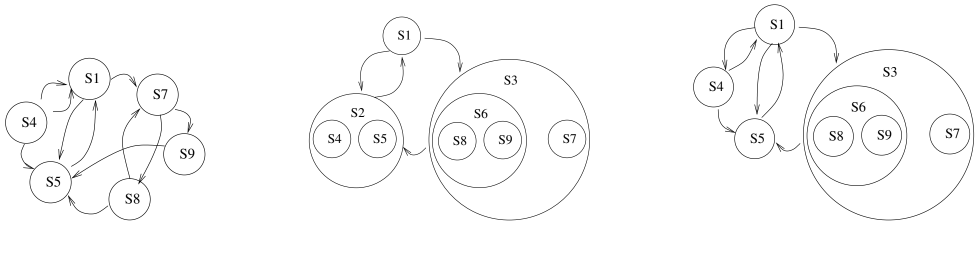
$$b(S_j) = \sum_{x_t \in S_j} \hat{p}(x_t|z^t)\ \ [\hat{p}(x_t|z^t) - \tilde{p}(x_t|z^t)]^2 \tag{10}$$

It is the weighed sum of the squared difference between the unbiased posterior $\hat{p}(x_t|z^t)$, computed at the resolution of the physical states and the biased posterior $\tilde{p}(x_t|z^t)$, computed at the resolution of abstract state S_j.

An approximation of the variance of abstract state S_j is computed as a weighted sum of the projection to the physical states as follows:

$$v(S_j) = \sum_{x_t \in S_j} \hat{p}(x_t|z^t) \left[\frac{\pi(x_t)}{\pi(S_j)}\right]^2 \frac{\hat{p}(S_j|z^t)[1 - \hat{p}(S_j|z^t)]}{n(S_j) + \pi(S_j)} \tag{11}$$

[2]The relative importance/cost of the physical states may also be included in the weight



(a) Example

(b) Variable resolution model

(c) S2 at a finer resolution

Figure 1: *(a)Arbitrary Markov model (b) Arbitrary variable resolution model corresponding to the Markov model in (a). Circles that enclose other circles represent abstract states. S2 and S3 and S6 are abstract states. States S2, S4 and S5 form one abstraction hierarchy, and states S3, S6, S7, S8 and S9 form another abstraction hierarchy. The states at the highest level of abstraction are S1, S2 and S3. (c) The model in (b) with states $S4$ and $S5$ at a finer resolution.*

The loss from tracking a set of states $x_t \in S_j$, at the resolution of the physical states is thus:

$$l_p = 0 + \sum_{x_t \in S_j} v(x_t) \tag{12}$$

The loss from tracking the same set of states in abstraction as S_j is:

$$l_a = b(S_j) + v(S_j) \tag{13}$$

There is a gain in terms of reduction in variance from generalizing and tracking in abstraction, but it results in an increase in bias. Here, a tradeoff between bias and variance refers to the process of accepting a certain increase in one term for a larger reduction in the other and hence in the total error.

3.4 Dynamically varying resolution

The variable resolution particle filter uses a bias-variance tradeoff to make a decision to vary the resolution of the state space. A decision to abstract to the coarser resolution of abstract state S_j, is made if the state space is currently at the resolution of states S_i, and the combination of bias and variance in abstract state S_j, is less than the combination of bias an variance of all its children S_i, as shown below:

$$b(S_j) + v(S_j) \leq \sum_{S_i \in \{children(S_j)\}} [b(S_i) + v(S_i)] \tag{14}$$

On the other hand if the state space is currently at the resolution of abstract state S_j, and the reverse of equation (14) is true, then a decision to refine to the finer resolution of states S_i is made. The resolution of a state is left unaltered if its bias-variance combination is less than its parent and its children. To avoid hysteresis, all abstraction decisions are considered before any refinement decisions.

Each time a new measurement is obtained the distribution of particles over the state space is updated. Since this alters the bias and variance tradeoff, the states explicitly represented at the current resolution of the state space are each evaluated for gain from abstraction or refinement. Any change in the current resolution of the state space is recursively evaluated for further change in the same direction.

4 Experimental results

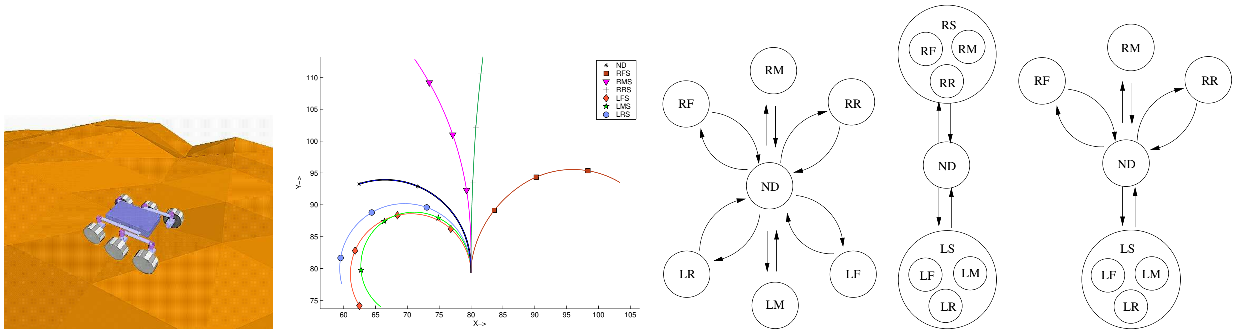
The problem domain for our experiments involves diagnosing locomotion faults in a physics based simulation of a six wheel rover. Figure 2(a) shows a snapshot of the rover in the Darwin2K [Leger, 2000] simulator.

The experiment is formulated in terms of estimating discrete fault and operational modes of the robot from continuous control inputs and noisy sensor readings. The discrete state, x_t, represents the particular fault or operational mode. The continuous variables, z_t, provide noisy measurements of the change in rover position and orientation. The particle set P_t therefore consists of N particles, where each particle $x_t^{[i]}$ is a hypothesis about the current state of the system. In other words, there are a number of discrete fault and operational states that a particle may transition to based on the transition model. Each discrete fault state has a different observation and predictive model for the continuous dynamics. The probability of a state is determined by the density of samples in that state.

The Markov model representing the discrete state transitions consists of 7 states. As shown in figure 2(c) the normal driving (ND) state may transition back to the normal driving state or to any one of six fault states: right front (RF), right middle (RM), right rear (RR), left front (LF), left middle (LM) and left rear (LR) wheel stuck. Each of these faults cause a change in the rover dynamics, but the faults on each side (right and left), have similar dynamics.

Given that the three wheels on each side of the rover have similar dynamics, we constructed a hierarchy that clusters the fault states on each side together. Figure 2(d) shows this hierarchical model, where the abstract states right side fault (RS), and left side fault (LS) represent sets of states $\{$RF, RM, RR$\}$ and $\{$LF, LM, LR$\}$ respectively. The highest level of abstraction therefore consists of nodes $\{$ND, RS, LS$\}$. Figure 2(e) shows how the state space in figure 2(d) would be refined if the bias in the abstract state RS given the number of particles outweighs the reduction in variance over the specialized states RF, RM and RR at a finer resolution.

When particle filtering is performed with the variable resolution particle filter, the particles are initialized at the highest



(a) Rover in D2K (b) Change in trajectory (c) (d) (e)

Figure 2: *(a) Snapshot from the dynamic simulation of the six wheel rocker bogie rover in the simulator, (b) An example showing the normal trajectory (ND) and the change in the same trajectory with a fault at each wheel. (c) Original discrete state transition model. The discrete states are: Normal driving (ND), right and left, front, middle and rear wheel faulty (RF, RM, RR, LF, LM, LR) (d) Abstract discrete state transition model. The states, RF, RM and RR have been aggregated into the Right Side wheel faulty states and similarly LF, LM and LR into Left Side wheel faulty states (RS and LS). (e) State space model where RS has been refined. All states have self transitions that have been excluded for clarity.*

level in the abstraction hierarchy,i.e. in the abstract states ND, RS and LS. Say a RF fault occurs, this is likely to result in a high likelihood of samples in RS. These samples will multiply which may then result in the bias in RS exceeding the reduction in variance in RS over RF, RM and RR thus favoring tracking at the finer resolution. Additional observations should then assign a high likelihood to RF.

The model is based on the real-world and is not very stochastic. It does not allow transitions from most fault states to other fault states. For example, the RF fault does not transition to the RM fault. This does not exclude transitions to multiple fault states and if the model included multiple faults, it could still transition to a “RF and RM” fault, which is different from a RM fault. Hence, if there are no samples in the actual fault state, samples that end up in fault states with dynamics that are similar to the actual fault state may end up being identified as the fault state. The hierarchical approach tracks the state at an abstract level and does not commit to identifying any particular specialized fault state until there is sufficient evidence. Hence it is more likely to identify the correct fault state.

Figure 3(a) shows a comparison of the error from monitoring the state using a classical particle filter that tracks the full state space, and the VRPF that varies the resolution of the state space. The X axis shows the number of particles used, the Y axis shows the KL divergence from an approximation of the true posterior computed using a large number of samples. 1000 samples were used to compute an approximation to the true distribution. The KL divergence is computed over the entire length of the data sequence and is averaged over multiple runs over the same data set [3]. The data set included normal operation and each of the six faults. Figure 3(a) demonstrates that the performance of the VRPF is superior to that of the classical filter for small sample sizes. In addition figure 3(b) shows the Kl-divergence along the Y axis and wall clock time along the X axis. Both filters were coded in matlab and share as many functions as possible.

5 Discussion and Future Work

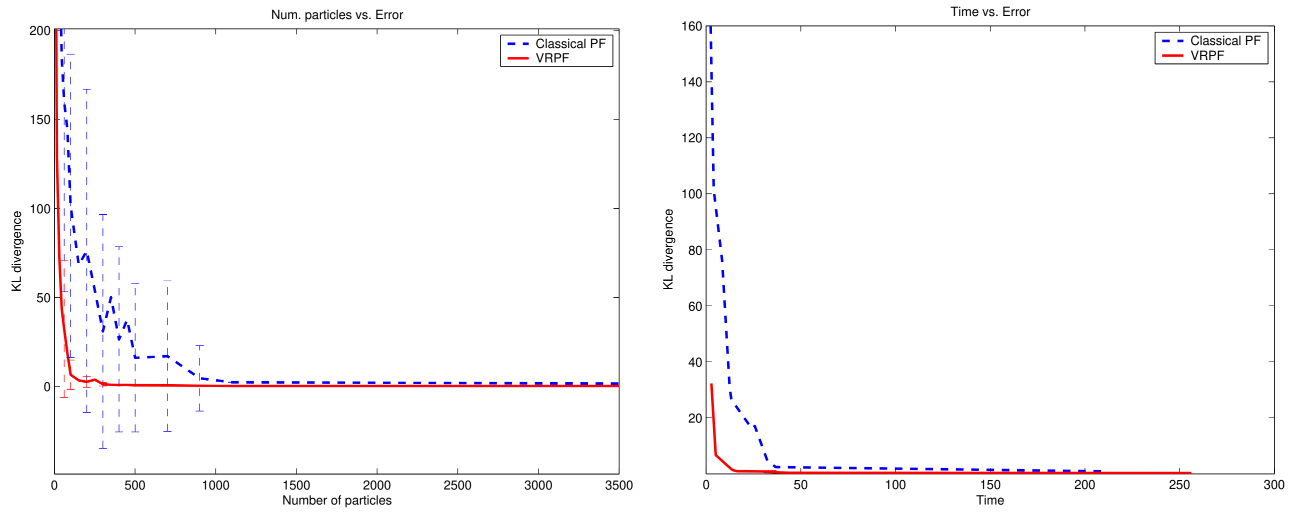
This paper presents a novel approach for state abstraction in particle filters that allows efficient tracking of large discrete, continuous or hybrid state spaces. The advantage of this approach is that it makes efficient use of the available computation by selecting the resolution of the state space based on an explicit bias-variance tradeoff. It performs no worse than a classical particle filter with an infinite number of particles [4]. But with limited particles we show that its performance is considerably superior. The VRPF does not prematurely eliminate potential hypotheses for lack of a sufficient number of particles. It maintains particles in abstraction until there are a sufficient number of particles to provide a low variance estimate at a lower resolution.

The VRPF generalizes from current samples to unsampled regions of the state space. Regions of the state space that had no samples at time t may acquire samples at time $t+1$ on abstraction. This provides additional robustness to noise that may have resulted in eliminating a likely hypothesis. [Koller and Fratkina, 1998] addresses this by using the time t samples as input to a density estimation algorithm that learns a distribution over the states at time t. Samples at time $t+1$ are then generated using this generalized distribution. [Ng *et al.*, 2002] uses a factored representation to allow a similar mixing of particles. The VRPF uses a bias-variance tradeoff to generalize more selectively and efficiently.

We have presented the variable resolution particle filter for a discrete state space. We plan to extend it to a continuous

[3]The results are an average over 50 to 5 runs with repetitions decreasing as the sample size was increased.

[4]There is a minor overhead in terms of evaluating the resolution of the state space



(a) (b)

Figure 3: *Comparison of the KL divergence from the true distribution for the classical particle filter and the VRPF, against (a) number of particles used, (b) wall clock time.*

state space where regions of state space, rather than sets of states are maintained at different resolutions. A density tree may be used for efficiently learning the variable resolution state space model with abstract states composed of finitely many sub-regions.

The VRPF makes efficient use of available computation and can provide significantly improved posterior estimates given a small number of samples. However, it should not be expected to solve all problems associated with particle impoverishment. For example, it assumes that different regions of the state space are important at different times. We have generally found this to be the case, but the VRPF does not improve performance if it is not. Another situation is when the transition and observation models of all the states are highly dissimilar and there is no possibility for state aggregation.

6 Acknowledgments

We would like to thank Geoff Gordon, Tom Minka and the anonymous reviewers.

References

[Behrends, 2000] E. Behrends. *Introduction to Markov Chains with Special Emphasis on Rapid Mixing*. Vieweg Verlag, 2000.

[Doucet and Crisan, 2002] A. Doucet and D. Crisan. A survey of convergence results on particle filtering for practitioners. IEEE Trans. Signal Processing, 2002.

[Doucet *et al.*, 2001] A. Doucet, N. de Freitas, and N.J. Gordon, editors. *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, 2001.

[Fox *et al.*, 1999] D. Fox, W. Burgard, and S. Thrun. Markov localization for mobile robots in dynamic environments. In *Journal of Artificial Intelligence*, volume 11, pages 391–427, 1999.

[Gordon *et al.*, 1993] N.J. Gordon, D.J Salmond, and A.F.M Smith. Novel approach to nonliner/non gaussian bayesian state estimation. volume 140 of *IEE Proceedings-F*, pages 107–113, 1993.

[Isard and Blake, 1998] M. Isard and A. Blake. Condensation: conditional density propagation for visual tracking. *International Journal of Computer Vision*, 29(1):5–28, 1998.

[Kanazawa *et al.*, 1995] K. Kanazawa, D. Koller, and S. Russell. Stochastic simulation algorithms for dynamic probabilistic networks. pages 346–351, 1995.

[Koller and Fratkina, 1998] D. Koller and R. Fratkina. Using learning for approximation in stochastic processes. In *ICML*, 1998.

[Leger, 2000] P.C. Leger. *Darwin2K: An Evolutionary Approach to Automated Design for Robotics*. Kluwer Academic Publishers, 2000.

[Metropolis and Ulam, 1949] N. Metropolis and S. Ulam. The monte carlo method. volume 44 of *Journal of the American Statistical Association*, 1949.

[Ng *et al.*, 2002] B. Ng, L. Peshkin, and A. Pfeffer. Factored particles for scalable monitoring. In *UAI*, 2002.

[Rubin, 1988] D. B. Rubin. *Using the SIR algorithm to simulate posterior distributions*. Oxford University Press, 1988.

[Rubinstein, 1981] R. Y. Rubinstein. Simulation and Monte Carlo method. In *John Wiley and Sons*, 1981.