The Particle Filter

The particle filter is a sequential Monte Carlo algorithm, i.e. a sampling method for approximating a distribution that makes use of its temporal structure. A "particle representation" of distributions is used. In particular, we will be concerned with the distribution $P(x_t|z_{0:t})$ where x_t is the unobserved state at time t, and $z_{0:t}$ is the sequence of observations from time 0 to time t.

In the previous lecture on Kalman filters, this distribution $P(x_t|z_{0:t})$ was a multivariate Gaussian due to assumptions regarding the transition model $P(x_{t+1}|x_t)$ and sensor model $P(y_t|x_t)$. The particle filter is more general, and makes few assumptions on these models (we only need to be able to evaluate these distributions, and draw samples from $P(x_{t+1}|x_t)$ or $P(x_{t+1}|x_t, y_{t+1})$).

Without restrictive linear Gaussian assumptions regarding the transition and sensor models, $P(x_t|z_{0:t})$ cannot be written in a simple form. Instead, we will represent it using a collection of N weighted samples or particles, $\{x_t^{(i)}, \pi_t^{(i)}\}_{i=1}^N$, where $\pi_t^{(i)}$ is the weight of particle $x_t^{(i)}$, a particle representation of this density:

$$P(x_t|z_{0:t}) \approx \sum_i \pi_{t-1}^{(i)} \delta(x_t - x_{t-1}^{(i)}) \tag{1}$$

Consider the integral that needed to be performed at each filtering step from the previous lecture:

$$P(x_t|z_{0:t}) \quad = \quad \alpha P(z_t|x_t) \int P(x_{t-1}|z_{0:t-1}) P(x_t|x_{t-1}) dx_{t-1} \tag{2}$$

As before, we are using this recursive definition to compute the filtered distribution $P(x_t|z_{0:t})$ given the distribution $P(x_{t-1}|z_{0:t-1})$.

With a particle representation for $P(x_{t-1}|z_{0:t-1})$, Equation 2 can be approximated as:

$$P(x_t|z_{0:t}) \quad \approx \quad \alpha P(z_t|x_t) \sum_i \pi_{t-1}^{(i)} P(x_t|x_{t-1}^{(i)}) \tag{3}$$

How do we create the "right" set of particles for representing the distribution $P(x_t|z_{0:t})$? One answer is to use importance sampling. The particle filter can be viewed as operating as an importance sampler on this distribution. The technique of importance sampling is a method for generating fair samples of a distribution $P(x)$. Suppose $P(x)$ is a density from which it is difficult to draw samples, but it is easy to evaluate $P(x_i)$ for some particular x_i. Then, an approximation to $P(x)$ can be given by:

$$P(x) \approx \sum_{i=1}^N \pi^{(i)} \delta(x - x^{(i)}) \tag{4}$$

where

$$\pi^{(i)} = \frac{P(x)}{q(x^{(i)})} \tag{5}$$

Note that any distribution $q(\cdot)$, known as a *proposal distribution*, can be used here. In particular, a uniform sampling of state space x. However, with such a uniform sampling strategy, most samples will be wasted, having small $\pi^{(i)}$ values. Instead, we use a more direct proposal distribution, our approximation to $P(x_t|z_{0:t-1})$ (the integral in Equation 2). With this proposal distribution, the weights $\pi^{(i)}$ end up being relatively simple due to cancellation.

Concretely, the particle filter consists of the following steps (from [3], equivalent to Algorithm 4, SIR, in [1]):

1. Draw N samples $x_t^{(j)}$ from the proposal distribution $q(x_t)$:

$$x_t^{(j)} \sim q(x_t) = \sum_i \pi_{t-1}^{(i)} P(x_t|x_{t-1}^{(i)}) \tag{6}$$

   by selecting a random number r uniformly from $[0,1]$, choosing the corresponding particle i, and then sampling from $P(x_t|x_{t-1}^{(i)})$. This transition model is typically a linear Gaussian model, but any model from which samples can easily be drawn will suffice.

2. Set the weight $\pi_t^{(j)}$ as the likelihood:

$$\pi_t^{(j)} = P(y_t|x_t^{(j)}) \tag{7}$$

   The samples $\{x_t^{(j)}\}$ above are fair samples from $P(x_t|z_{0:t-1})$. Reweighting them in this fashion accounts for evidence z_t.

3. Normalize the weights $\{\pi_t^{(j)}\}$:

$$\pi_t^{(j)} = \frac{\pi_t^{(j)}}{\sum_k \pi_t^{(k)}} \tag{8}$$

Another point is that there is an optimal proposal distribution, which is not the one used here. The optimal proposal distribution, minimizing variance in weights $\pi^{(i)}$, turns out to be $p(x_t|x_{t-1}, z_t)$.

The most important property of the particle filter is its ability to handle complex, multi-modal (non-Gaussian) posterior distributions. However, it has difficulties when x_t is high-dimensional. Essentially, the number of particles N required to adequately approximate the distribution grows exponentially with the dimensionality of the state space. This poses difficulties in applications such as articulated human body tracking or SLAM.

Acknowledgements

Portions of these notes are adapted from [1, 4, 3, 5, 2]. Isard and Blake [3] introduced particle filtering to the computer vision community as the "CONDENSATION" algorithm, with contour tracking as an application. This is the reference to start with for computer vision types.

Arulampalam et al. [1] provide a tutorial on particle filtering, and describe its many variants.

References

[1] 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*, 50(2):174–188, February 2002.

[2] A. Doucet, S. J. Godsill, and C. Andrieu. On sequential monte carlo sampling methods for bayesian filtering. *Statistics and Computing*, 10:197–208, 2000.

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

[4] M. Jordan and C. Bishop. An introduction to graphical models.

[5] Z. Khan, T. Balch, and F. Dellaert. A rao-blackwellized particle filter for eigentracking. In *Proc. IEEE Comput. Soc. Conf. Comput. Vision and Pattern Recogn.*, 2004.