AN ONLINE MOTION-BASED PARTICLE FILTER FOR HEAD TRACKING APPLICATIONS

Nidhal Bouaynaya, Wei Qu and Dan Schonfeld

University of Illinois at Chicago
Department of Electrical and Computer Engineering
nbouay1@uic.edu, wqu@ece.uic.edu, ds@ece.uic.edu

ABSTRACT

Particle filtering framework has revolutionized probabilistic tracking of objects in a video sequence. In this framework the proposal density can be any density as long as its support includes that of the posterior. However, in practice, the number of samples is finite and consequently the choice of the proposal is crucial to the effectiveness of the tracking. The CONDENSATION filter uses the transition prior as the proposal density. We propose in this paper a motion-based proposal. We use Adaptive Block Matching (ABM) as the motion estimation technique. The benefits of this model are two fold. It increases the sampling efficiency and handles abrupt motion changes. Analytically, we derive a Kullback-Leibler (KL)-based performance measure and show that the motion proposal is superior to the proposal of the CONDENSATION filter. Our experiments are applied to head tracking. Finally, we report promising tracking results in complex environments.

1. INTRODUCTION

Reliable object tracking in complex environments is a challenging task. Its applications include video surveillance [1], autonomous vehicle navigation [2], and virtual reality [3] among many others. Recently temporal Bayesian filtering [4], [5] has become very popular for object tracking. In this probabilistic framework, the goal is to estimate the system's current state given its past and current observations. However, except for the linear gaussian case (Kalman filter [5]) the problem does not admit an analytical solution. Moreover, real world object tracking does not satisfy kalman filter requirements: the system dynamics can be highly nonlinear and the observation density is multimodel due to clutter. Particle filters can handle non-linear and non-Gaussian systems. The idea is to approximate the posterior density by its sample set. Since it is hard to sample directly from the posterior, Particle filter employs the Importance Sampling technique [6], [7], [8]. In Importance Sampling, a proposal density, also called importance function, is used to generate samples. Each sample is then assigned a proper weight to make up the difference between the posterior density and the proposal density. It can be shown that (i) the compensated sample set is a fair approximation of the posterior and (ii) if number of samples is sufficiently large, the sample approximation of the posterior density can be made arbitrarily accurate [9], [10]. However, in practice the resources are finite. To make the situation worse, if a good dynamic model is not available or if the state dimension of the tracked object is high, the number of required samples becomes even larger and Particle filter can be computationally prohibitive. Choosing the right proposal density is one of the most important issues in particle filters' design.

In this paper we propose to use a motion-based proposal density. The benefits of this model are two fold. First the motion estimation allows efficient allocation of the samples. Second the tracker is adaptive to all kinds of motion. Particularly it handles sudden and unexpected motion; e.g., a motion that is not captured by the state transition model. We choose Adaptive Block Matching technique because of its simplicity in implementation [11]. However, our model can be generalized with any motion estimation algorithm. The rest of the paper is organized as follows. In Section 2, we construct the motion proposal using the motion vector estimated by the ABM. In Section 3, we set up an optimization problem based on the KL performance measure and prove that the motion proposal is superior to the proposal used in the CONDENSATION filter, i.e., the transition prior. In Section 4, we apply our algorithm to head tracking using challenging real-world video sequences. Concluding remarks and future work are given in Section 5.

2. A MOTION-BASED PARTICLE FILTER

We assume a Markovian discrete-time state space model. Let X_k represent the target characteristics at discrete time k (position, velocity, shape, etc). The state space model is described by a state transition and measurement equations. The goal is to estimate the posterior density $p(\mathbf{X})$ of the target given its past and current observations. In what follows, we use a subscript to denote the time index and a superscript to denote the sample index; e.g., $X_k^{(n)}$ is the n^{th} sample at time k. Let $\{X^{(n)}, n \in \mathbb{N}\}$ be a sample set from a proposal density $q(\mathbf{X})$. The importance weights are then given by

$$\pi^{(n)} = \frac{p(X^{(n)})}{q(X^{(n)})}\,\tilde{\pi}^{(n)}, \tag{1}$$

where $\tilde{\pi}^{(n)}$ are the un-compensated weights associated with the sampling of $q(\mathbf{X})$. The state estimate is then given by the sample mean:

$$\hat{X}_k = \sum_{i=1}^{N} \pi_k^{(i)} X_k^{(i)}. \tag{2}$$

Our experiments are applied to head tracking. The human head is well modelled by an ellipse [12]. This domain knowledge helps avoid erroneous shape evolvement therefore greatly improving the tracking results. Specifically, we use a four dimension parametric ellipse to represent the head contour:

$$X_k = [x_c(k), y_c(k), b(k), \phi(k)]^T, \tag{3}$$

where $(x_c(k), y_c(k))$ is the center of the ellipse at time k, $b(k)$ and $\phi(k)$ are the minor axis and the orientation of the ellipse at time k, respectively. The ratio of the major and minor axis of the ellipse is held constant equal to its value computed in the first frame. This is a very reasonable assumption that allows us to reduce the dimensionality of the state vector by 1. The ABM module output is a mask of the object at the current time. The motion vector that interests us is given by

$$\Delta X_k = [x_c(k) - x_c(k-1), y_c(k) - y_c(k-1), 0, 0]. \tag{4}$$

If we model computational errors by a zero-mean white Gaussian noise, we have the following sampling scheme

$$X_k^{(n)} = X_{k-1}^{(n)} + \Delta X_k + v_k^{(n)}, \tag{5}$$

where

$$v_k \sim N(0, \Sigma_G), \tag{6}$$

and

$$\Sigma_G = \begin{pmatrix} \sigma_x^2 & 0 & 0 & 0 \\ 0 & \sigma_y^2 & 0 & 0 \\ 0 & 0 & \sigma_s^2 & 0 \\ 0 & 0 & 0 & \sigma_r^2 \end{pmatrix}, \tag{7}$$

where σ_x^2, σ_y^2, σ_s^2 and σ_r^2 are the variances of the motion diffusion in the (x-y) direction, scaling and rotation respectively. The motion proposal q_m is then given by

$$q_m(X_k|X_{k-1}, Z_k) \equiv \sum_{i=1}^{N} N_{X_k}(X_{k-1}^{(i)} + \Delta X_k, \Sigma_G), \tag{8}$$

where we made use of the following notation that we will be also using for the rest of this paper

$$N_X(\mu, \Sigma) \equiv \frac{1}{2\pi|\Sigma|} exp(-\frac{1}{2}(X-\mu)^T \Sigma^{-1}(X-\mu)). \tag{9}$$

In addition to the proposal evaluation, we also need to calculate the particle likelihood $p(Z_k|X_k)$ and transition probability $p(X_k|X_{k-1})$. We use both face color and edge detection as visual clues for particle weighting. For the gradient module, we use the model developed by Isard in [4]. For the color module, we use the histogram model developed in [13]. We adopt a simple random walk model for the prior dynamics.

3. OPTIMAL IMPLEMENTABLE IMPORTANCE FUNCTION

As pointed out in [8] and [10], the optimal proposal density q_{opt} is the one that minimizes the variance of the importance weights. In practice, however, sampling from the optimal density is very difficult if not impossible. Though we cannot practically sample from q_{opt}, we can assess the performance of any proposal density q by computing some kind of similarity measure between q_{opt} and q. The Kullback-Leibler (KL) measure between two continuous functions q and q_{opt} is defined as the multiple integral:

$$I(q_{opt}, q) = \int q_{opt}(x) \log(\frac{q_{opt}(x)}{q(x)})dx, \tag{10}$$

where log denotes the natural logarithm.
$I(q_{opt}, q)$ is the "information" lost when a given proposal density q is used to approximate the optimal proposal density q_{opt}. Ideally, the general optimization problem would be to minimize $I(q_{opt}, q)$ over the set $\mathbf{Q}$ of practically implementable density functions:

$$q^* = \min_{q \in \mathbf{Q}} I(q_{opt}, q). \tag{11}$$

Notice that $\mathbf{Q}$ does not contain the optimal density q_{opt}. However this general problem is intractable. For simplification we consider the subset $\mathbf{Q}_2$ of $\mathbf{Q}$ containing only two densities. $\mathbf{Q}_2 = \{q, q_c\}$, where q is a given implementable density and q_c is the proposal density in the CONDENSATION filter; i.e.,

$$q_c(X_k|X_{k-1}^{(i)}) = p(X_k|X_{k-1}^{(i)}).$$

Let us choose a set $\mathcal{S}$ that contains the region of support of the optimal density q_{opt}. Let Ω be the set of densities q satisfying the inequality $\int_{\mathcal{S}} \frac{q_{opt} q_c}{q} dX_k \leq 1$. Then we have the following result:

Proposition. *If $q \in \Omega$ then $I(q_{opt}, q) \leq I(q_{opt}, q_c)$.*

Proof.

$$\begin{aligned} & I(q_{opt}, q) - I(q_{opt}, q_c) \\ &= \int_{\mathcal{S}} q_{opt} \log(\frac{q_c}{q})\, dX_k \qquad (12) \\ &\leq \log(\int_{\mathcal{S}} \frac{q_{opt} q_c}{q})\, dX_k \qquad (13) \\ &\leq \int_{\mathcal{S}} \frac{q_{opt} q_c}{q}\, dX_k - 1 \qquad (14) \\ &\leq 0 \end{aligned}$$

Equation (13) uses Jensen's Inequality and equation (14) uses the inequality $\log(x) \leq x - 1$. □

To establish the performance properties of the motion-based proposal density, we need the following assumption: In equation (5), $v_k^{(n)} \ll \Delta X_k$. This assumption is perfectly legitime in our case since the motion proposal increases the sampling efficiency allowing for a small noise variance in the propagation process. Let Σ_P be the covariance matrix of the transition dynamics. We then have the following corollary:

Corollary. *If* $\Sigma_G = \Sigma_P$, *then* $q_m \in \Omega$ *and* $I(q_{opt}, q_m) \leq I(q_{opt}, q_c)$.

Proof.

$$\begin{aligned} & q_m(X_k = X_k^{(n)} | X_{k-1}^{(n)}, Z_k) \\ &= \frac{1}{2\pi|\Sigma_G|} \exp^{(-\frac{1}{2}(X_k^{(n)} - X_{k-1}^{(n)} - \Delta X_k)^T \Sigma_G^{-1} (X_k^{(n)} - X_{k-1}^{(n)} - \Delta X_k))} \\ &= \frac{1}{2\pi|\Sigma_G|} \exp^{(-\frac{1}{2}(X_k^{(n)} - X_{k-1}^{(n)})^T \Sigma_G^{-1} (X_k^{(n)} - X_{k-1}^{(n)}))} \\ &\quad \exp^{([(X_k^{(n)} - X_{k-1}^{(n)}) - \frac{1}{2}\Delta X_k]^T \Sigma_G^{-1} \Delta X_k)} \\ &= \frac{1}{2\pi|\Sigma_G|} \exp^{(-\frac{1}{2}(X_k^{(n)} - X_{k-1}^{(n)})^T \Sigma_G^{-1} (X_k^{(n)} - X_{k-1}^{(n)}))} \\ &\quad \exp^{([\frac{1}{2}\Delta X_k + v_k^{(n)}]^T \Sigma_G^{-1} \Delta X_k)} \end{aligned}$$

The last equation was obtained by replacing $X_k^{(n)}$ by its expression in equation (5).
Using the assumption we introduced earlier, we end up with

$$\begin{aligned} & q_m(X_k = X_k^{(n)} | X_{k-1}^{(n)}, Z_k) \\ &= \frac{1}{2\pi|\Sigma_G|} \exp^{(-\frac{1}{2}(X_k^{(n)} - X_{k-1}^{(n)})^T \Sigma_G^{-1} (X_k^{(n)} - X_{k-1}^{(n)}))} \\ &\quad \exp^{(\frac{1}{2}\Delta X_k^T \Sigma_G^{-1} \Delta X_k)} \\ &= p(X_k^{(n)} | X_{k-1}^{(n)}) \exp^{(\frac{1}{2}\Delta X_k^T \Sigma_G^{-1} \Delta X_k)} \end{aligned}$$

where the last equality uses the fact that the prior is a random walk with covariance matrix $\Sigma_P = \Sigma_G$.

Replacing q by q_m in equation (14), we have

$$\begin{aligned} & I(q_{opt}, q_m) - I(q_{opt}, q_c) \\ &\leq \int_{\mathcal{S}} \frac{q_{opt} q_c}{q_m}\, dX_k - 1 \\ &= \int_{\mathcal{S}} q_{opt} \exp^{-(\frac{1}{2}\Delta X_k^T \Sigma_G^{-1} \Delta X_k)}\, dX_k - 1 \\ &\leq \int_{\mathcal{S}} q_{opt}\, dX_k - 1 \qquad (15) \\ &= 0 \end{aligned}$$

The last equality follows from the fact that the region $\mathcal{S}$ includes the region of support of q_{opt}. □

The proposition gives a KL-based performance measure for a given density compared to the CONDENSATION filter. From an information theory viewpoint, this result means that q_m is a better approximation to q_{opt} than q_c.

4. EXPERIMENTS

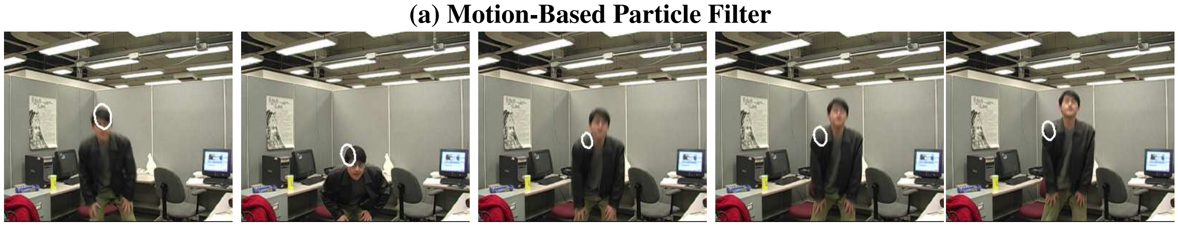
We test our algorithm in different challenging real world situations. Clutter, full rotation and fast and erratic movements are considered. In the experiment, we propagate 50 particles. The prediction model is a random walk reflecting a poor a priori knowledge of the object's dynamics. Figure 1 shows the tracking results of simultaneous jumping and full plane rotation of the head. Figure 2 compares MBPF and CONDENSATION filter in case of sudden movement. The person in the video walks slowly then suddenly bends and quickly recovers. The transition prior model does not take into account such an erratic movement and consequently CONDENSATION filter fails. On the other hand, because MBPF's proposal distribution places the limited particles more effectively, it tracks both sequences successfully.

5. CONCLUSION

Choosing the right proposal density is one of the most important issues in particle filter's design. We presented a motion based particle filter. The motion estimation increases the sample allocation efficiency and handles all kinds of motion, particulary fast and erratic motion. Our model can be generalized with any motion estimation technique. We then set up a general optimization problem to find the optimal implementable proposal density in the KL measure sense. For simplicity we constrained ourselves to a sub-optimal solution to the general optimization problem. This led us to interesting analytical results comparing the motion proposal with the CONDENSATION filter proposal in the KL measure sense.

Fig. 1. Motion-Based Particle Filter handles simultaneous jumping and full rotation of the head.

(a) Motion-Based Particle Filter



(b) CONDENSATION filter

Fig. 2. **Tracking a sudden movement**. The person in this video walks slowly then suddenly bends and quickly recovers.

6. REFERENCES

[1] I. Haritaoglu, D. Harwood, and L. Davis, "W4: Real-time surveillance of people and their activities," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 2, pp. 809–830, August 2000.

[2] E. Dickmanns and V. Graefe, "Applications of dynamic monocular machine vision," *Machine vision and Applications*, vol. 1, pp. 241–261, 1988.

[3] E. Marchand and F. Chaumette, "Virtual visual servoing: A framework for real-time augmented reality," in *Proc. Eurographics*, 2002, pp. 289–298.

[4] M. Isard and A. Blake, "Condensation conditional density propagation for visual tracking," *Int. J. Computer Vision*, vol. 29, no. 1, pp. 5–28, 1998.

[5] B. Anderson and J. Moore, *Optimal Filtering*, Englewood Cliffs, Prentice Hall, New Jersy, 1979.

[6] M.A. Tanner, *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions*, Springer-Verlag, New York, 1993.

[7] A. Doucet, S. Godsill, and C. Andrieu, "On sequential monte carlo sampling methods for bayesian filtering," *Stat. Comput.*, vol. 10, no. 3, pp. 197–208, 2000.

[8] J. Liu and R. Chen, "Sequential monte carlo methods for dynamic systems," *J. Amer. Stat. Assoc.*, vol. 93, pp. 1032–1044, 1998.

[9] N. Gordon, D. Salmond, and A. Smith, "Novel approach to nonlinear/non-gaussian bayesian state estimation," in *IEE Proc. F*, 1993, vol. 140, pp. 107–113.

[10] A. Doucet, "On sequential simulation-based methods for bayesian filtering," Tech. Rep. CUED/F-INFENG/TR.310, 1998.

[11] P. Raffy K. Hariharakrishnan, D. Schonfeld and F. Yassa, "Fast object tracking using adaptive block matching," *IEEE Transactions on Multimedia*, to appear.

[12] S. Birchfield, "Elliptical head tracking using intensity gradients and color histograms," in *Proc. IEEE Int. Conf. on Comput. Vis. and Patt. Recog.*, 1998, pp. 232–237.

[13] K. Nummiaro, E. Koller-Meier, and L.J. Van Gool, "Object tracking with an adaptive color-based particle filter," *DAGM-Symposium Pattern Recognition*, pp. 353–360, 2002.