Markov chain Monte Carlo methods for visual tracking

Ray Luo
rluo@cory.eecs.berkeley.edu
Department of Electrical Engineering and Computer Sciences
University of California, Berkeley
Berkeley, CA 94720

Abstract

Tracking articulated figures in high dimensional state spaces require tractable methods for inferring posterior distributions of joint locations, angles, and occlusion parameters. Markov chain Monte Carlo (MCMC) methods are efficient sampling procedures for approximating probability distributions. We apply MCMC to the domain of people tracking and investigate a general framework for sample-approximation tracking based on the Particle Filter, MCMC, and simulated annealing. A tutorial discussion of MCMC is provided.

1 Introduction

Tracking complex moving emsembles in noisy environments with highly quantized video sequences produces difficult computational problems that cannot be fixed simply by making the object or motion models more complex. Complicated object models involve a greater number of state variables that parameterize the object orientation or position. The high dimensionality of the state space makes probabilistic inference for the posterior distribution of the object location or angle intractable. Detailed motion models fail miserably when asked to track a similar but different motion sequence. For example, built-in walking models for tracking walking humans fail when the subject turns a corner [10].

Conventional approaches to visual tracking involves a sampling and resampling procedure commonly known as Condensation or Particle Filtering [4]. Unlike linear prediction algorithms like the Kalman Filter [5], Particle Filters do not assume Gaussian transition and observation noise. Its problem lies in the exponentially increasing number of particles needed for high dimensional noisy problems.

Markov chain Monte Carlo (MCMC) is a set of techniques for efficient sampling of probability distributions. We apply MCMC to the domain of people tracking to construct better, more efficient approximations to the posterior distributions of states. We begin with a review of tracking as a probabilistic prediction and inference problem. After discussing the Particle Filter, we move on to MCMC methods, culminating in the hybrid Monte Carlo method that was implemented. Finally, we describe each piece of the tracker as a Bayesian filtering step and discuss extensions and future research directions.

2 Probabilistic foundation

Let $\mathbf{S}_t \in \mathbb{K}$ be a vector of variables we want to track, e.g. torso position, joint angle, and body orientation. $\mathbb{K}$ is the appropriate domain for each variable, e.g. $[0, \pi]$ for arm elbow angle and $\mathbb{R}^3$ for location of body centroid. In general, each ith variable $\mathbf{S}_t^i$ can be a vector in a vector space.

Our observed data at time t is a vector $\mathbf{D}_t$. At any given time t we are given the observations $\mathbf{D}_{0:t} = (\mathbf{D}_0, \mathbf{D}_1, \ldots, \mathbf{D}_t)$. We're interested in the posterior distribution $p(\mathbf{S}_t|\mathbf{D}_{0:t})$ which assigns a probability to each configuration $\mathbf{S}_t = S_t$.

2.1 Bayesian filtering

We factor the posterior distribution by Bayes Rule [7],

$$\begin{aligned} p(\mathbf{S}_t|\mathbf{D}_{0:t}) &= \frac{p(\mathbf{D}_t|\mathbf{S}_t, \mathbf{D}_{0:t-1})\, p(\mathbf{S}_t|\mathbf{D}_{0:t-1})}{p(\mathbf{D}_t|\mathbf{D}_{0:t-1})} \\ &= \alpha\, p(\mathbf{D}_t|\mathbf{S}_t, \mathbf{D}_{0:t-1})\, p(\mathbf{S}_t|\mathbf{D}_{0:t-1}), \end{aligned}$$

where $p(\mathbf{D}_t|\mathbf{D}_{0:t-1})$ is the constant α^{-1} set after calculations to a value that normalizes the distribution.

Next we make some assumptions about the observations to simplify our problem. Given our current state, past observations should be independent of present observations. For example, suppose your little brother puts an indeterminant number of coins into the piggy bank every week. Your state estimate is the observed number of coins from the previous week. Then your observations this week is independence of past observations as long as your brother is a capricious money saver. Note that in this case, the state is simply a record of observations. In general, the state is a function of the observations, but the independence relationship of the observations may still hold. This gives us a factorization,

$$p(\mathbf{S}_t|\mathbf{D}_{0:t}) = \alpha\, p(\mathbf{D}_t|\mathbf{S}_t)\, p(\mathbf{S}_t|\mathbf{D}_{0:t-1}), \tag{1}$$

where $p(\mathbf{D}_t|\mathbf{S}_t)$ is the likelihood function and $p(\mathbf{S}_t|\mathbf{D}_{0:t-1})$ is the prediction distribution which represents current state estimates given the past observations only. Equation (1) specifies the correction made to $\mathbf{S}_t$ given the data. It represents the control process of our dynamical system due, for example, to sensory feedback and temporal corrections.

Marginalize over the previous state to get the prediction distribution (or, a.k.a. the temporal prior),

$$\begin{aligned} p(\mathbf{S}_t|\mathbf{D}_{0:t-1}) &= \int_{\mathbf{S}_{t-1}} p(\mathbf{S}_t, \mathbf{S}_{t-1}|\mathbf{D}_{0:t-1}) \\ &= \int_{\mathbf{S}_{t-1}} p(\mathbf{S}_t|\mathbf{S}_{t-1}, \mathbf{D}_{0:t-1})\, p(\mathbf{S}_{t-1}|\mathbf{D}_{0:t-1}). \end{aligned}$$

Let us now make an assumption regarding our state description of the dynamical system. In particular, we assume that the future state of the system is independent of the past observations given the present state. That is, the present state captures everything we need to know about to calculate the state probabilities at the next time step. This is a first order Markov assumption applied to Bayesian filtering. Note that the assumption has nothing to do with the underlying physical process. Given a rich enough state description and accurate sensor readings, we can always satisfy the Markov assumption to a sufficient degree. Now the prediction is

$$p(\mathbf{S}_t|\mathbf{D}_{0:t-1}) = \int_{\mathbf{S}_{t-1}} p(\mathbf{S}_t|\mathbf{S}_{t-1})\, p(\mathbf{S}_{t-1}|\mathbf{D}_{0:t-1}), \tag{2}$$

where $p(\mathbf{S}_t|\mathbf{S}_{t-1})$ is the dynamic or transition probability and $p(\mathbf{S}_{t-1}|\mathbf{D}_{0:t-1})$ is the prior distribution, i.e. the posterior from the previous time step.

From (1) and (2), we get a recursive formulation of the posterior state distribution that depends only on previous and current observations. Starting with a flat initial prior that assigns uniform probability to every state configuration, we propagate $\mathbf{S}_{t-1}$ forward in time according to our dynamical model, examine the probability of the new data given the propagated $\mathbf{S}_t$, and calculate the new posterior based on the on the temporal prior and the likelihood. The space complexity of the general algorithm depends on the size of $\mathbf{S}_t$

2.2 Sample approximation

The Monte Carlo principle approximates a probability distribution using a weighted set of delta functions [8].

2.3 Particle Filtering

Kalman filtering [5]. Particle filtering of Isard and Blake [4].

3 Markov chain Monte Carlo

3.1 Gibbs sampling

Geman and Geman were the first to apply Gibbs sampling to an image restoration problem [3].

3.2 Metropolis algorithm

Metropolis et al. used a symmetric proposal distribution [9].

3.3 Hybrid Monte Carlo

An auxiliary variable sampler samples from an augmented distribution and obtains the desired sample approximation by marginalizing over unwanted variables. The most popular auxiliary variables algorithm is the hybrid Monte Carlo filter first introduced in the thermal physics community by Duane [2]. Duane et al. showed that the MCMC can be combined with the Metropolis test.

3.4 Simulated annealing

Kirkpatrick et al. first applied simulated annealing to problems in statistical physics [6].

4 Human tracker

Choo and Fleet built a people tracker based on hybrid Monte Carlo sampling of the posterior [1].

4.1 Likelihood

4.2 Dynamics

4.3 Temporal prior

5 Results

6 Conclusions

Acknowledgments

The author would like to thank David J. Fleet and Eunice Poon.

References

[1] K. Choo and D. J. Fleet. Tracking people using hybrid Monte Carlo. In *International Conference on Computer Vision*, volume II, pages 321–328, Vancouver, Canada, 2001.

[2] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth. Hybrid Monte Carlo. *Physics Letters B*, 195(2):216–222, 1987.

[3] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):721–741, November 1984.

[4] M. Isard and A. Blake. Contour tracking by stochastic propagation of conditional density. In *European Conference on Computer Vision*, pages 343–356, Cambridge, UK, 1996.

[5] R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the American Society of Mechanical Engineers - Journal of Basic Engineering*, 82(D):35–45, 1960.

[6] S. Kirkpatrick, Jr. C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, May 1983.

[7] A. Leon-Garcia. *Probability and Random Processes for Electrical Engineering*. Addison-Wesley, Reading, Massachusetts, 2nd edition, 1994.

[8] J. S. Liu and R. Chen. Sequential Monte Carlo methods for dynamical systems. *Journal of the American Statistical Association*, 93:1032–1044, 1998.

[9] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21(6):1087–1092, June 1953.

[10] H. Sidenbladh, M. J. Black, and D. J. Fleet. Stochastic tracking of 3d human figures using 2d image motion. In *European Conference on Computer Vision*, volume II, pages 702–718, Dublin, Ireland, 2000.