

马尔科夫随机场与实时多目标 跟踪算法

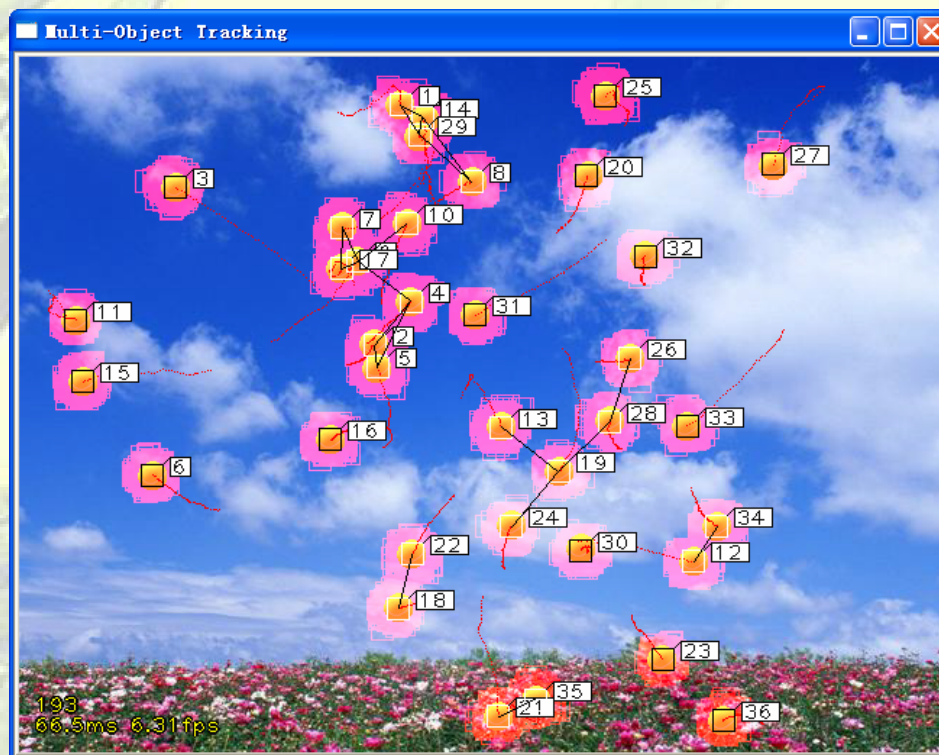
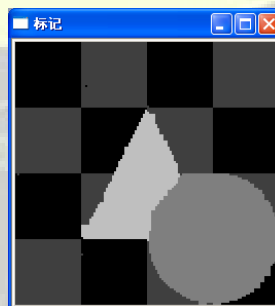
沈乐君

sljcool@sina.com

<http://www.shenlejun.cn>

MRF的应用实例

- 利用马尔科夫随机场进行噪声图像的分割
- 马尔科夫随机场结合粒子滤波器进行实时多目标跟踪



目录

- 马尔科夫随机场（MRF）的介绍
- 基于MRF的噪声图像分割
- 粒子滤波器的介绍
- 综合MRF与粒子滤波器进行多目标跟踪
- 自助重要性采样（BISP）的实时多目标跟踪

MRF简介

- 马尔可夫随机场（Markov Random Field）是计算机视觉、图像处理等研究的重要工具。
- 它在图像处理中的应用广泛。它不仅可处理底层视觉（图像分割、边缘检测），还可用于中高层视觉处理（多目标跟踪、场景分析和理解）。

MRF的应用

- 图像除噪和图像增强



MRF的应用

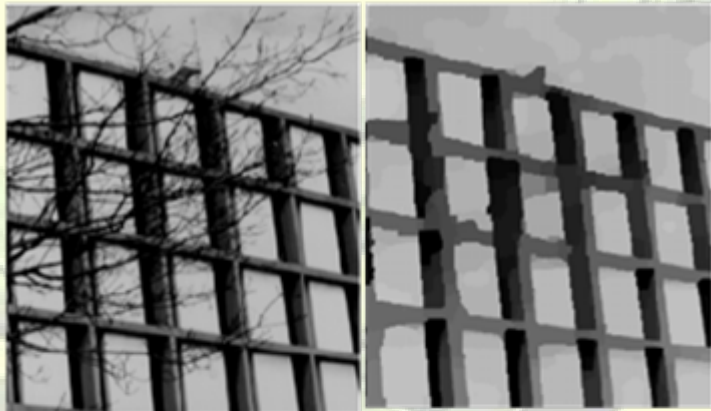
- 图像恢复



Roth and Black, CVPR 2005

MRF的应用

- 去除杂波 (Clutter Removal)



Zhu and Mumford PAMI, 1997.

MRF的应用

- 超解析度 (Super-resolution)



Zoomed low-resolution



Full frequency original



Low-resolution

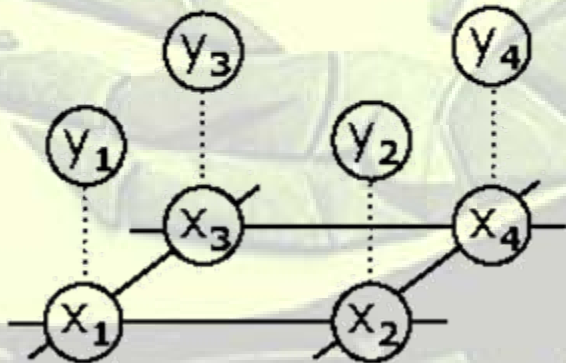


Bill Freeman ICCV99

MRF的应用

- Given two images, estimate depth at each pixel

立体视觉



Depths

(Tens of labels)



MRF的应用

- Given two images, estimate motion vector at each pixel

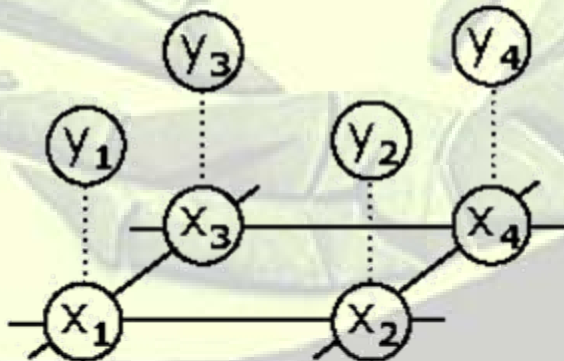
运动矢量



First image



Second image



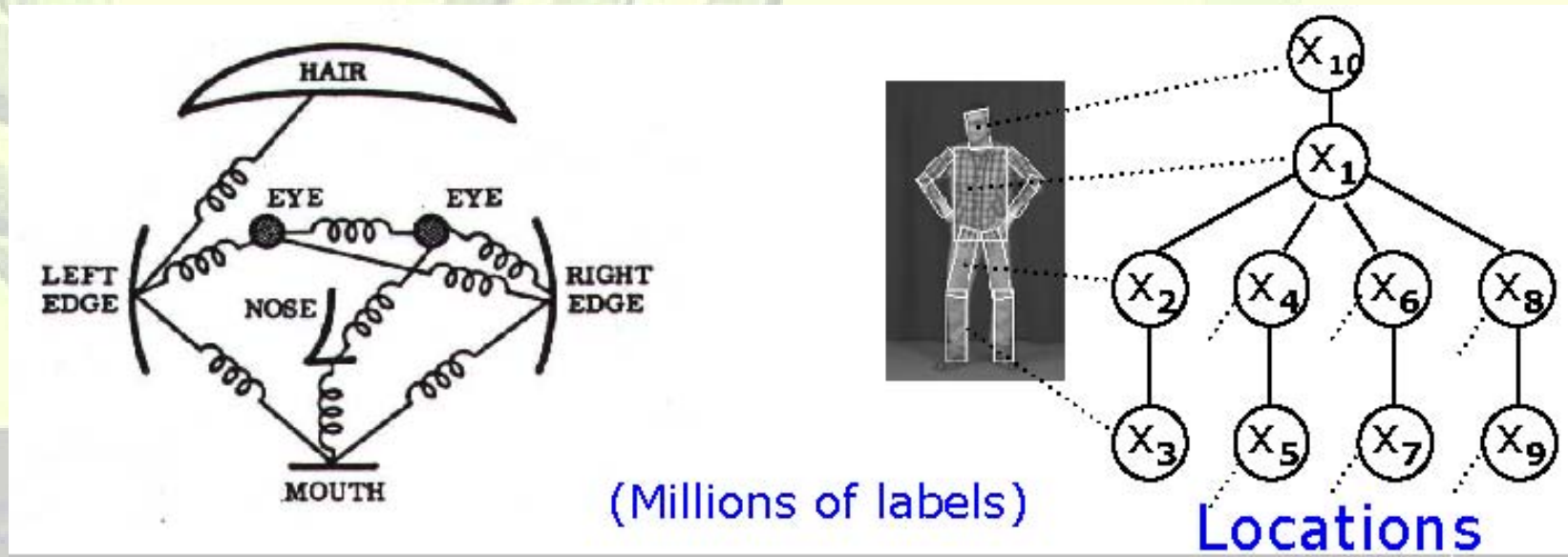
Flow vectors

(Hundreds of labels)



MRF的应用

- 目标检测与识别。对目标的部分与整体关系进行建模。



随机场

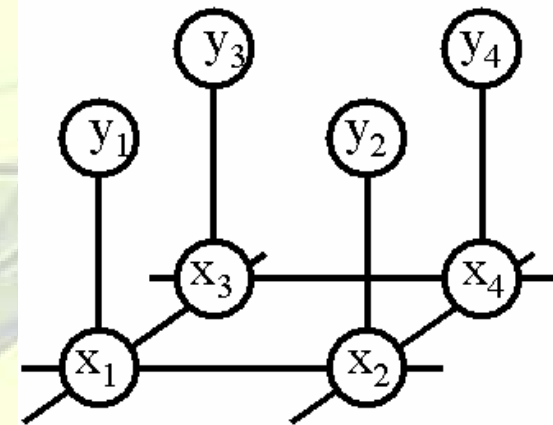
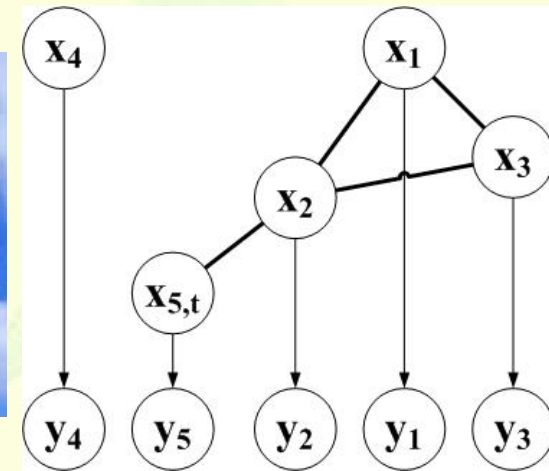
- 随机场（Random Field）是一种被广泛应用的随机数学模型。之所以它的应用广泛，是因为它的定义非常简洁和通用。
- 下面介绍其核心定义。

随机场

- Collection of n **sites S**
- Hidden **random variable x_i** at each site i
- **Label set L**
 - Each site takes on label $l \in L$
- **Neighborhood system N**
 - N_i neighbors of site i
 - Explicit dependencies between neighbors

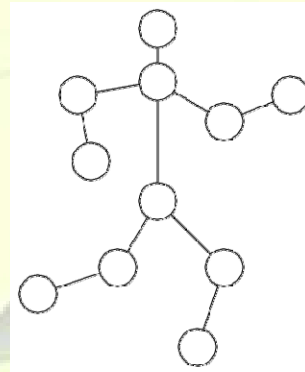
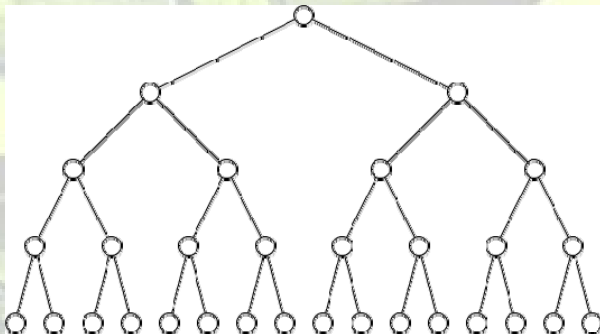
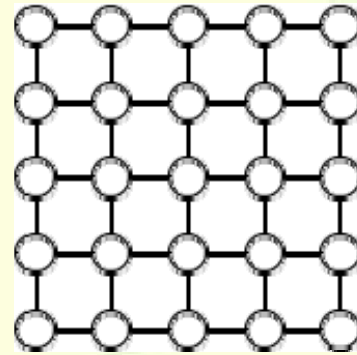
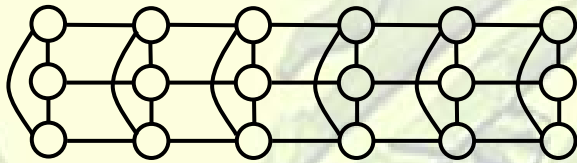
随机场

- Collection of n sites S
- 位置 S ，刻画随机变量 x_i 的空间关系。
 - 网格 eg.4个像素
 - 遮挡 eg.5个目标



随机场

- 其它各种类型的S



随机场

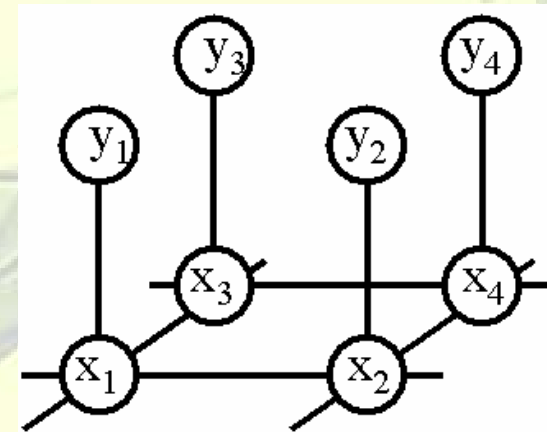
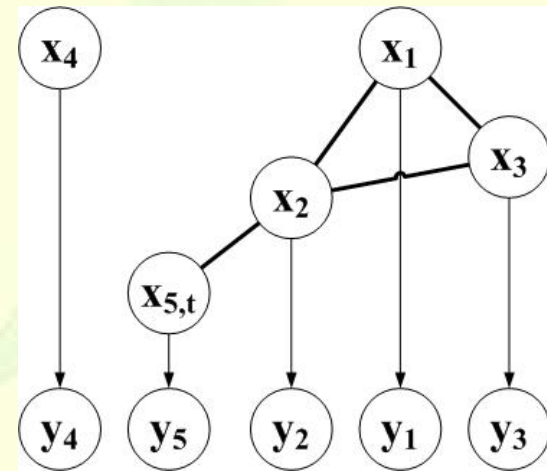
- Hidden **random variable** x_i at each site i
- 随机变量 x_i 是我们在图像处理过程中感兴趣的量。
 x_i 是隐变量，也就是随机场中的未知量。 x_i 常被噪声污染，所以是随机变量。 x_i 共有多个，所以叫做随机场。
 - 像素的真实类别标签，如：1,2,3,4
 - 目标的真实位置，如：横坐标=10.2, 纵坐标=4.5
- 为了求解 X ，常引入可观测变量 y ，它是已知量，代表CCD获取的真实图像信息。
- **Label set L**
- 标签 L 是变量 x 的取值空间，它是离散或连续的。

随机场

- Neighborhood system N

- 隐变量 x_i 常常与其它变量相关。用 N 刻画变量与其它变量的相关性。

- 依赖关系对应于无向图中的边。

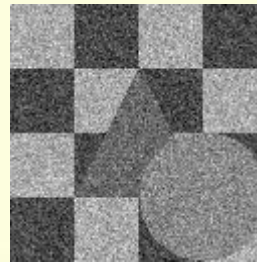
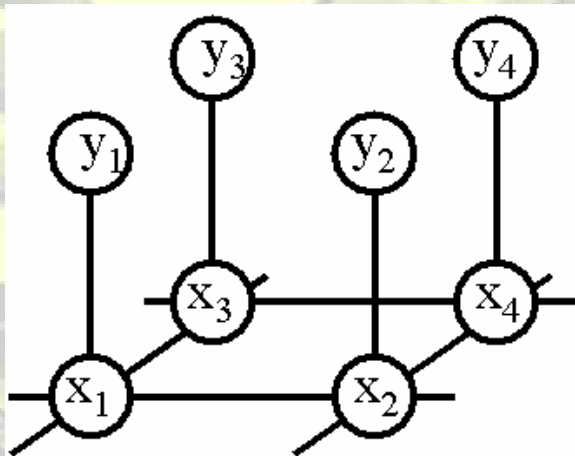


马尔科夫随机场

- Random field with **Markov property**
- $P(x_i | x_{S \setminus i}) = P(x_i | x_{N_i})$
 - Where $S \setminus i$ denotes set S excluding element i
- **Conditional probabilities** depend only on neighborhood
 - Probability of x_i taking on some value same given all other nodes as given just neighbors
- 赋予随机场马尔科夫性，最重要的目的就是降低模型的计算代价。cf.蝴蝶效应、颜色革命与利比亚

马尔科夫随机场

- 在点对马尔科夫随机场（Pairwise Markov Random Fields）中引入观测变量 y_i ，对隐变量 x_i 进行贝叶斯推理。下面先介绍贝叶斯公式：
- $P(X|Y) = P(Y|X) P(X) / P(Y)$



Y: 图像



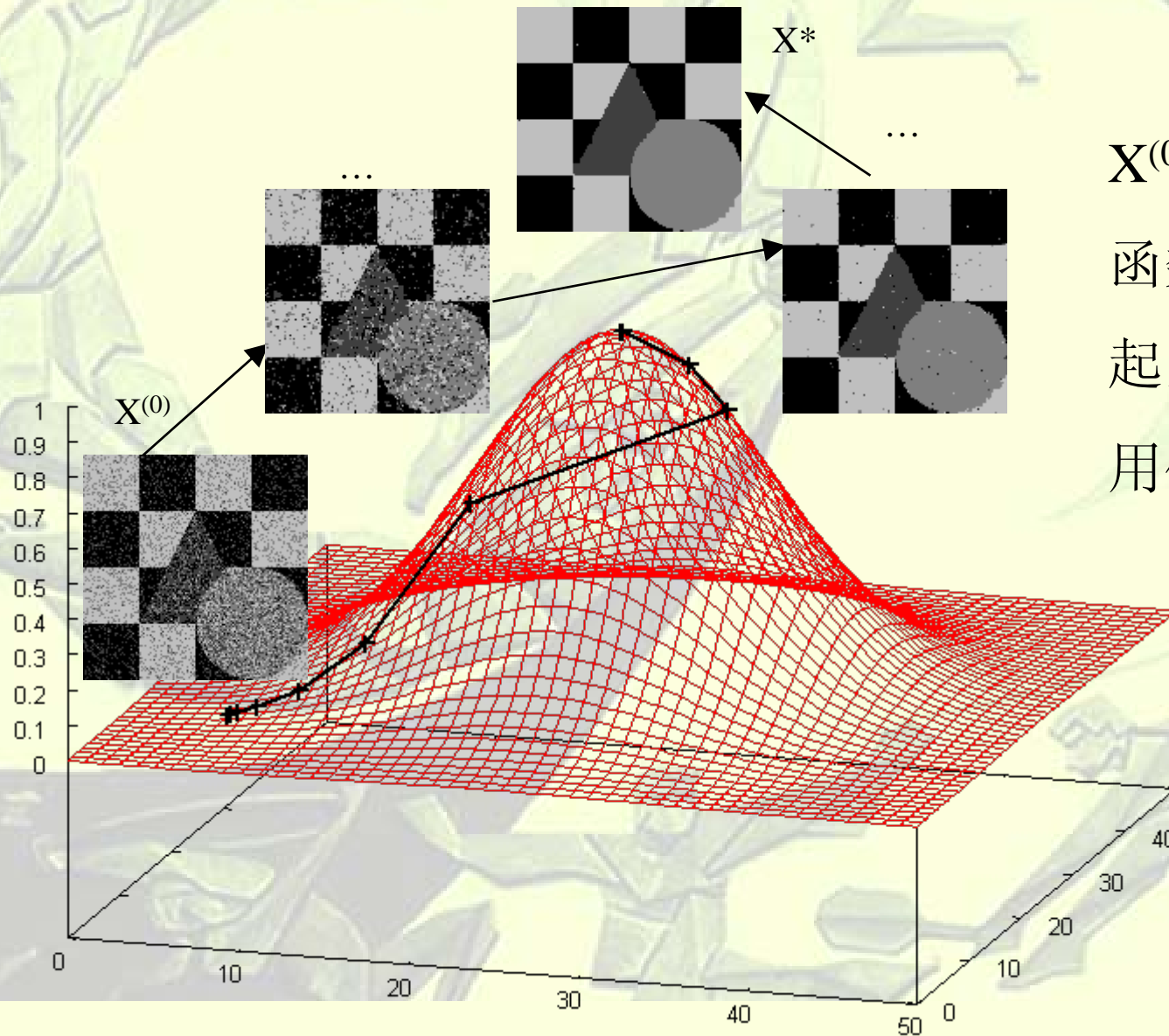
X: 分类标签

(下页解释X,Y)

马尔科夫随机场

- 其中， X 是全部隐变量组成的向量， Y 是全部观测变量组成的向量。 Y 是已知图像， X 是未知的图像分割结果。如果图像大小是 128×128 ，那么， X 是16384维向量。假设每个像素有4种状态，那么 X 共有 $4^{128 \times 128} \approx 10^{9864}$ 种可能。
- 如何在庞大的可能空间中寻找最佳的图像分割结果？显然，用贪婪法求解是不现实的。这是一个典型的组合优化问题。于是，图像分割问题转换为最优化问题。大量的图像处理问题，最终都能被转化成为最优化问题。（下页是最优化示意图）

马尔科夫随机场



$X^{(0)} \rightarrow \dots \rightarrow X^*$

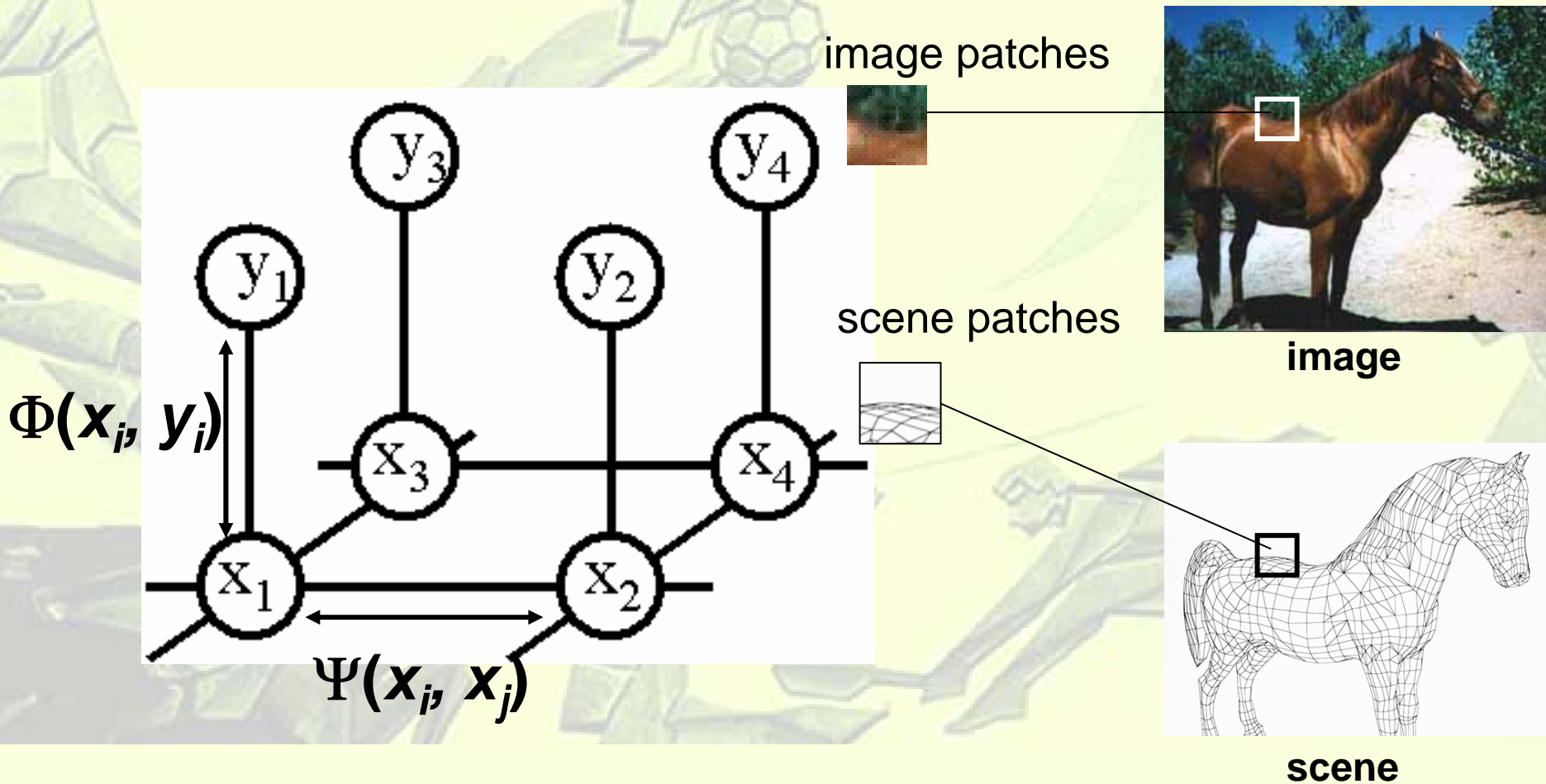
函数 $f(X)$ 是什么？

起点 $X^{(0)}$ 如何得到？

用什么最优化方法？

函数形态

- 首先，考虑 $f(X)$ ，也就是联合概率 $P(X, Y)$



函数形态

- 点对马尔科夫随机场的联合概率公式

$$P(\mathbf{x}, \mathbf{y}) = \frac{1}{Z} \prod_{i,j} \Psi(\mathbf{x}_i, \mathbf{x}_j) \prod_i \Phi(\mathbf{x}_i, \mathbf{y}_i)$$

scene image

Scene-scene compatibility function

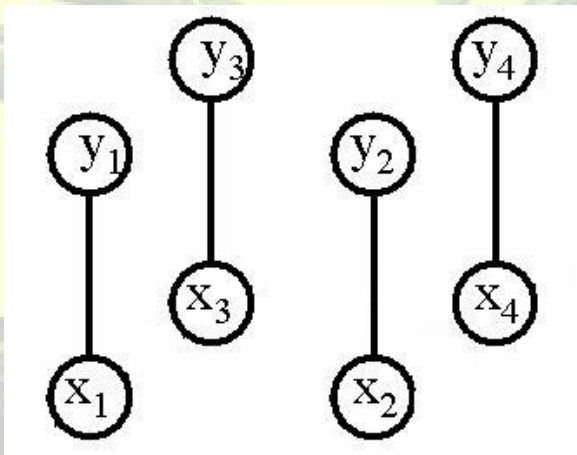
neighboring scene nodes

Image-scene compatibility function

local observations

起始解

- 其次，考虑优化的起点，也就是 X 的初始解 $X^{(0)}$ 。假设像素之间相互独立，使得像素之间相互独立地进行分类或标记。也就是暂时忽略联合概率公式中的边势函数。



$$P(X, Y) = \frac{1}{Z} \prod_{i,j} \phi(x_i, x_j) \prod_i \phi(x_i, y_i)$$

起始解

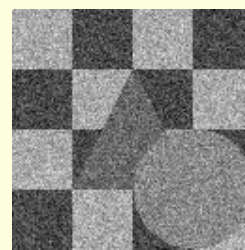
- 相互独立假设，使得我们仅考虑每个像素的分类或标记。可是，解是

$$P(X^{(0)}, Y) = \frac{1}{Z} \prod_i \phi(x_i, y_i)$$

$$\phi(x_i, y_i) = P(x_i | y_i)$$

$$P(x_i | y_i) \propto P(y_i | x_i) P(x_i)$$

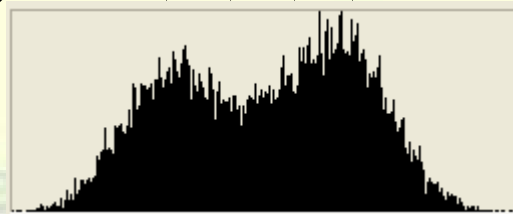
$$\frac{1}{\sqrt{2\pi}\sigma_{L_i}} \exp\left(\frac{-(y_i - \mu_{L_i})^2}{2\sigma_{L_i}^2}\right)$$



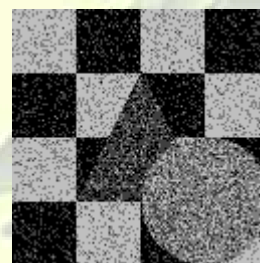
Y: 图像

μ_1 μ_2 μ_3 μ_4

(运行程序观察参数)



直方图



$X^{(0)}$: 初始解

最优化方法

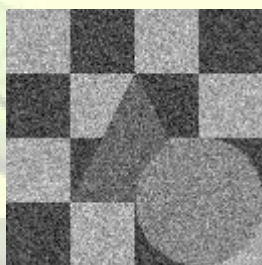
- 什么是概率推理？已知观测 Y ，以及MRF模型的参数（如均值 $\mu_1=76.1$ ），推理出令联合概率公式最大的 X 的值。也叫做MAP推理（Maximum A Posteriori）。换句话说，在可能解空间中，搜索使联合概率最高的解。
- 如果随机变量 X 拥有离散的样本空间，MAP就是一个组合优化问题。如果随机变量是连续的样本空间，MAP就是一个数值优化问题。
- 可是，MRF推理的困难在于：非线性；解析不可解；仅利用梯度信息，容易陷入局部极值；

概率推理

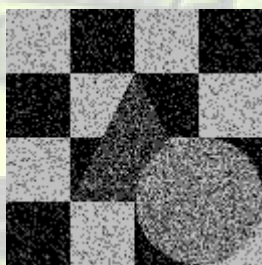
- 基于MRF的概率推理分为精确推理和近似推理两类。后者使用取样方法来近似求解。
 - Gibbs sampling, simulated annealing
 - Iterated conditional modes (ICM)
 - Metropolis sampling
 - Variational methods
 - Belief propagation
 - Graph cuts

取样方法

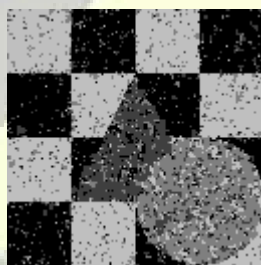
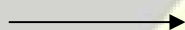
- 取样方法：通过某种算法产生大量样本，取后验概率最大的样本为解。



Y: 图像

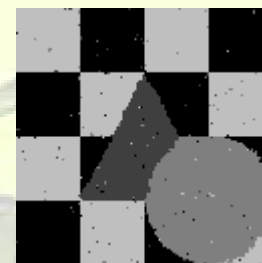


$X^{(0)}$



$X^{(1)}$

.....



$X^{(n)}$

取样方法

- 在取样时，考虑像素之间的依赖关系。

$$P(X, Y) = \frac{1}{Z} \prod_{i,j} \varphi(x_i, x_j) \prod_i \phi(x_i, y_i)$$

$$\varphi(x_i, x_j) \begin{cases} \alpha & \text{if } x_i = x_j \\ \beta & \text{if } x_i \neq x_j \end{cases}$$

例如 $\alpha = 0.8$ $\beta = 0.2$

取样方法

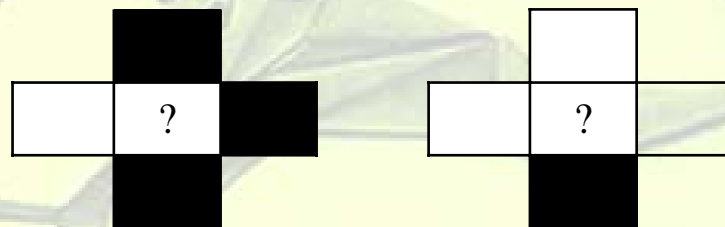
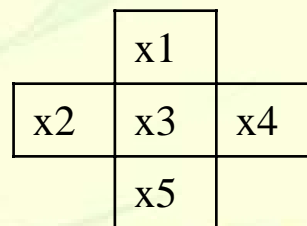
- 吉布斯取样（Gibbs sampling）方法用于产生样本。直观解释：根据邻域 N_3 的变量的值来产生样本。模拟退火（simulated annealing）用于寻找全局最优值。

$$x_1^{(t+1)} \sim \pi(x_1 | x_2^{(t)}, x_3^{(t)}, \dots, x_K^{(t)})$$

$$x_2^{(t+1)} \sim \pi(x_2 | x_1^{(t+1)}, x_3^{(t)}, \dots, x_K^{(t)})$$

⋮

$$x_K^{(t+1)} \sim \pi(x_K | x_1^{(t+1)}, \dots, x_{K-1}^{(t+1)})$$



取样方法

- 教程与源代码:
- UGM: A Matlab toolbox for probabilistic undirected graphical models
- 网址: <http://www.cs.ubc.ca/~schmidtm/>



蒙特卡洛仿真

- 假设生成了N个样本。如前文，可用于求MAP:

$$\hat{x} = \arg \max_{x^{(i)}; i=1, \dots, N} p(x^{(i)})$$

- 如果N趋向于无穷，可以求积分:

$$I_N(f) = \frac{1}{N} \sum_{i=1}^N f(x^{(i)}) \xrightarrow[N \rightarrow \infty]{a.s.} I(f) = \int_{\mathcal{X}} f(x)p(x)dx.$$

- 用N个样本来表示概率密度分布:

$$p_N(x) = \frac{1}{N} \sum_{i=1}^N \delta_{x^{(i)}}(x)$$

蒙特卡洛仿真

- 已知 Y ，可以用 N 个样本描述 $P(X|Y)$ 的分布。

$$\{x_{t-1}^r\}_{r=1}^N \sim P(x_{t-1} | y_{t-1})$$

- 在前文中 X 表示像素，现在表示目标的坐标。目标跟踪的任务，就是求解已知图像 Y ，和已知上一时刻的目标位置的条件下，目标 X 的值。

$$P(x_t | y_t) = \eta P(y_t | x_t) \int P(x_t | x_{t-1}) P(x_{t-1} | y_{t-1}) dx_{t-1}$$

当前状态

观测方程

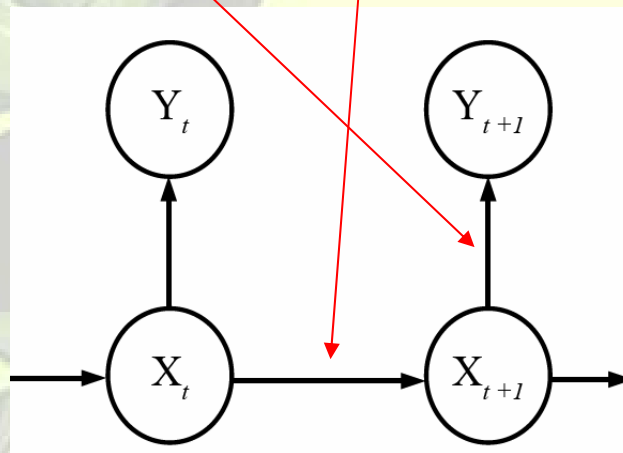
动态方程

上一时刻状态

粒子滤波器

- 动态方程 Specifies how objects move between frames
- 观测方程 Specifies the likelihood of an object being in a specific state (i.e. at a specific location)

$$P(x_t | y_t) = \eta P(y_t | x_t) \int P(x_t | x_{t-1}) P(x_{t-1} | y_{t-1}) dx_{t-1}$$



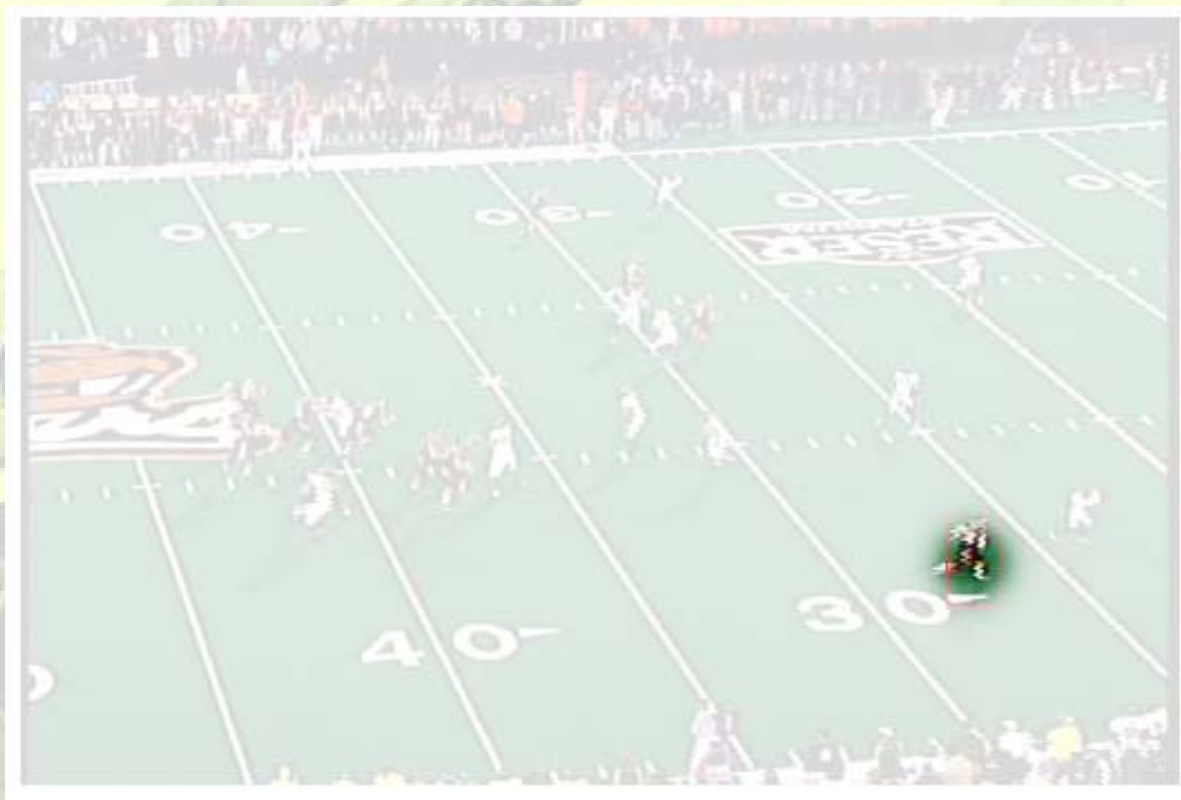
粒子滤波器

- 初始状态 $P(x_0)$: 目标检测或手工选中



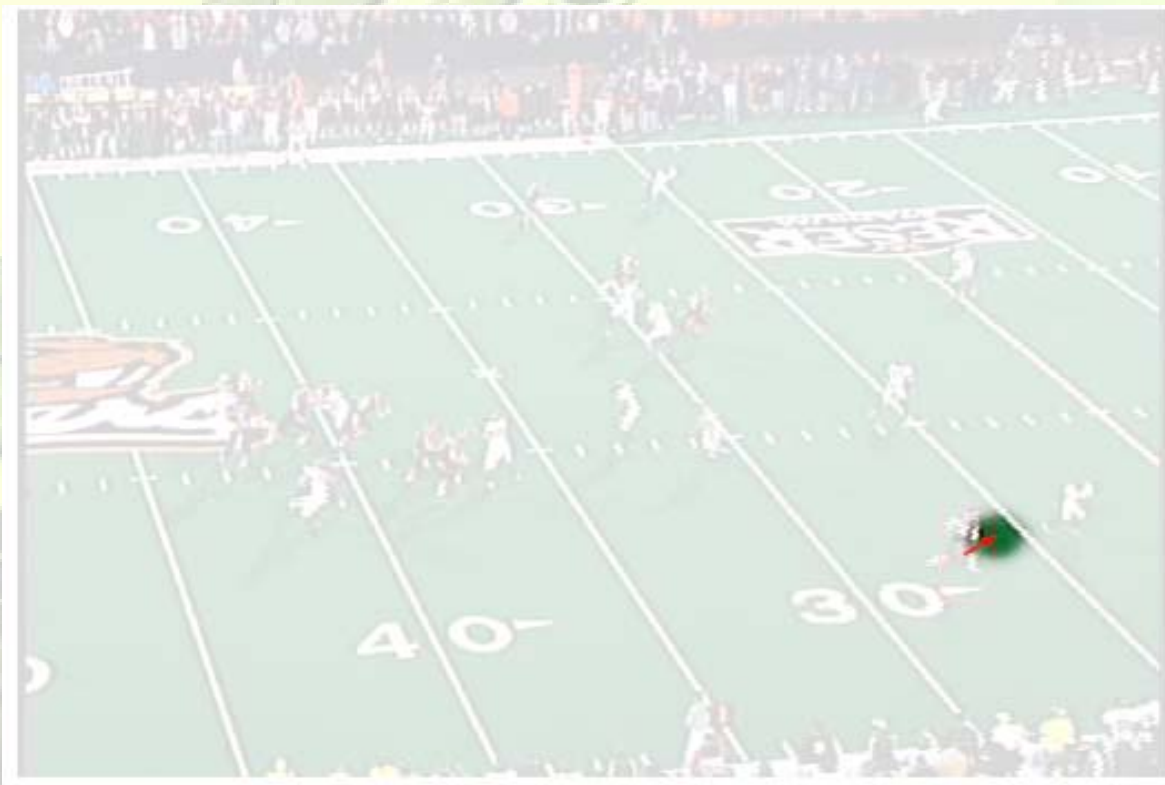
粒子滤波器

- 最简单的动态方程是高斯窗口



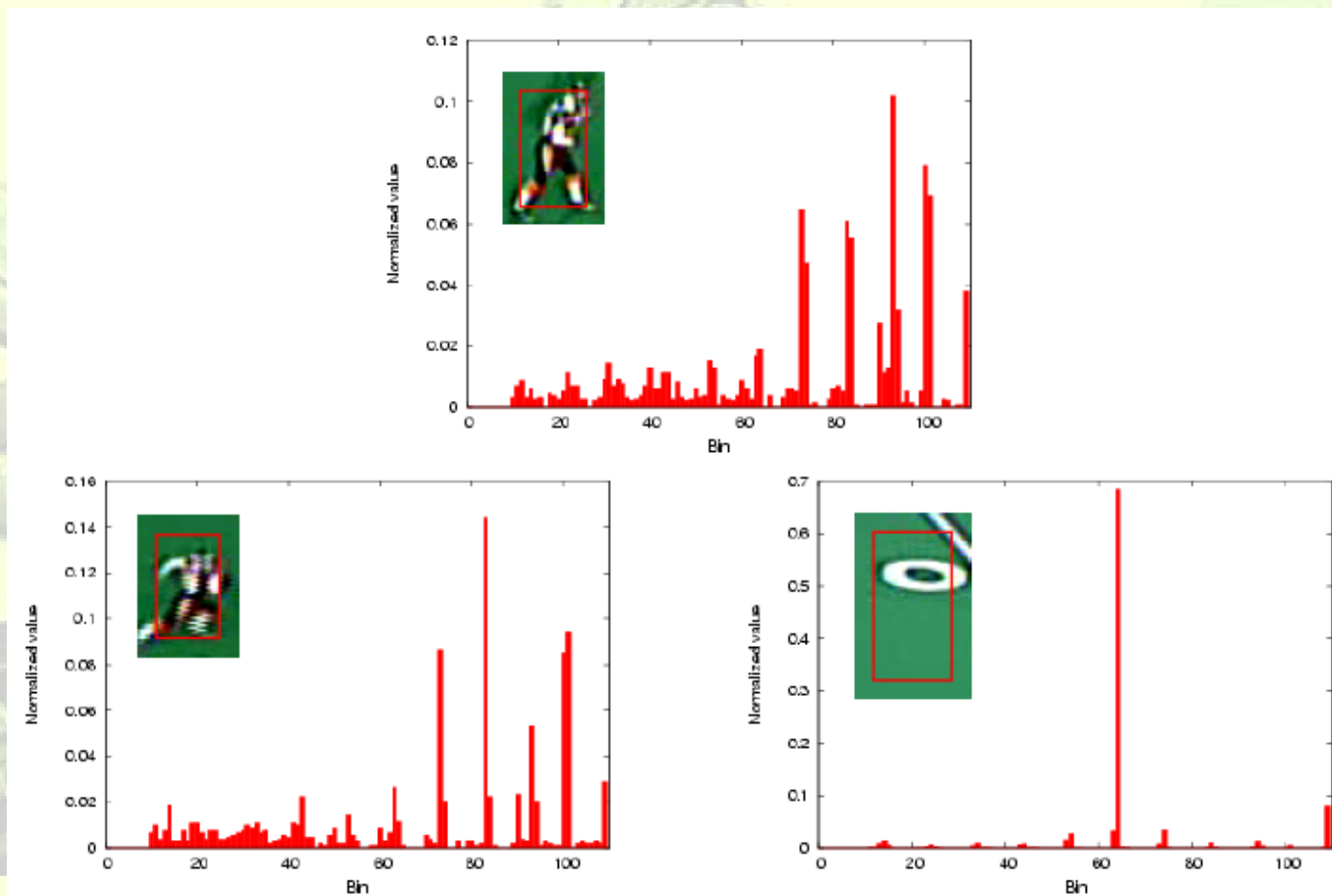
粒子滤波器

- 好一点的动态方程，根据上一时刻的速度预测高斯窗口的中心位置。



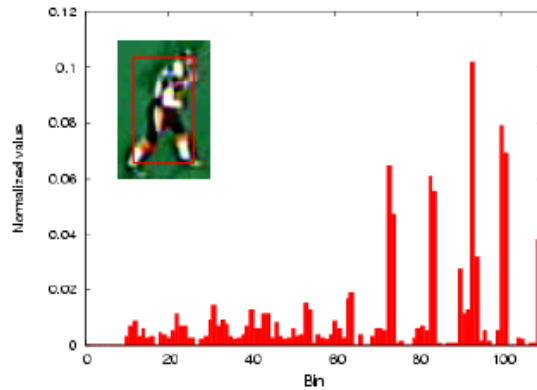
粒子滤波器

- 观测方程基于HSV色彩直方图模型

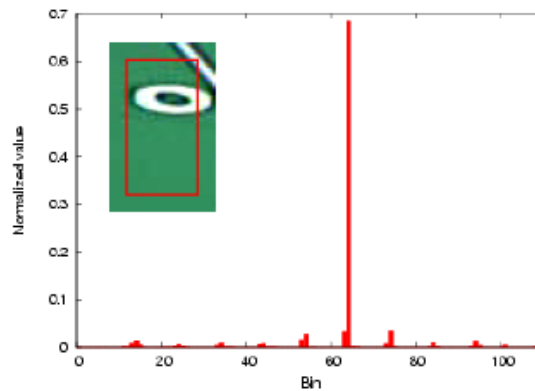
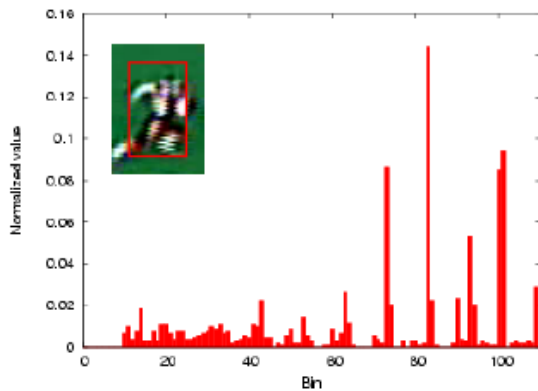


粒子滤波器

- 观测方程的似然度计算的是两个直方图之间的巴氏距离:

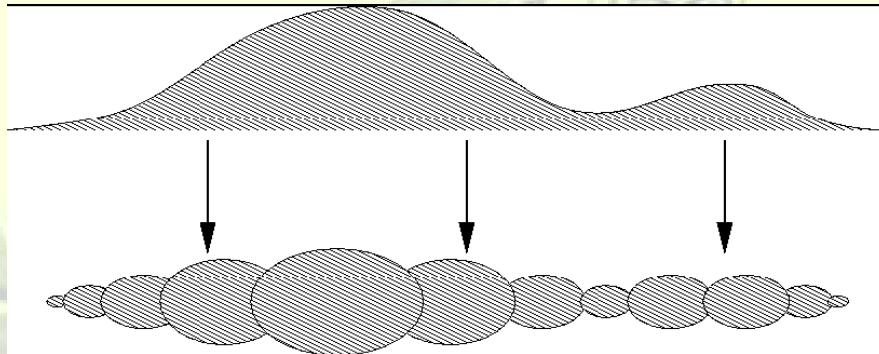


$$p(y_t|x_t) \propto e^{-\lambda D^2[h_0, h(x_t)]}$$



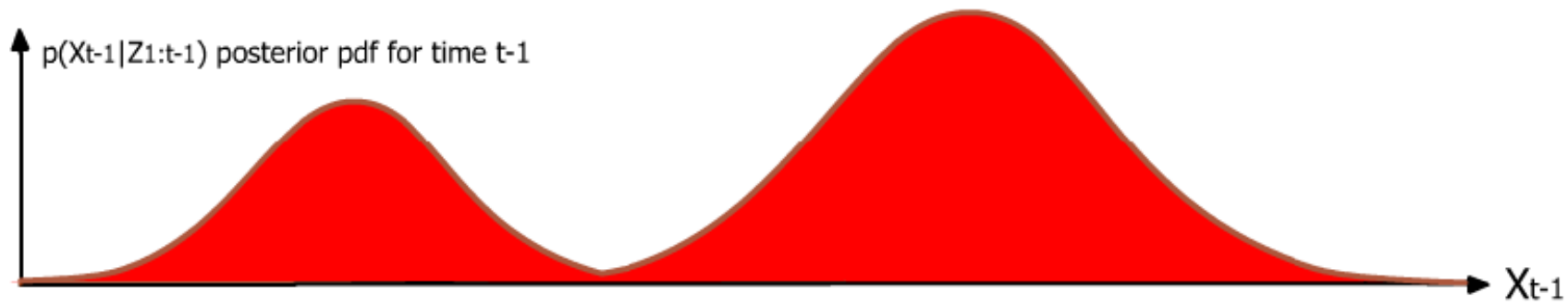
粒子滤波器

- 粒子滤波器的核心思想 a set of weighted particles approximates the filtering distribution

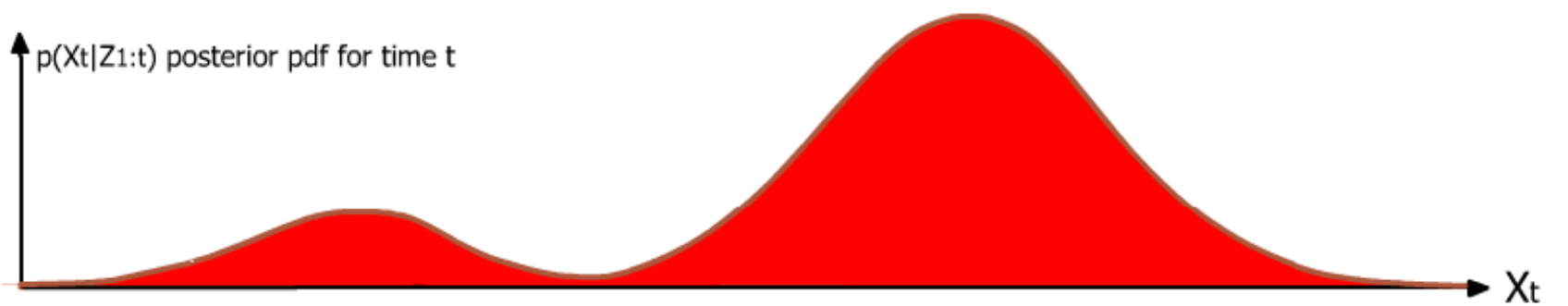
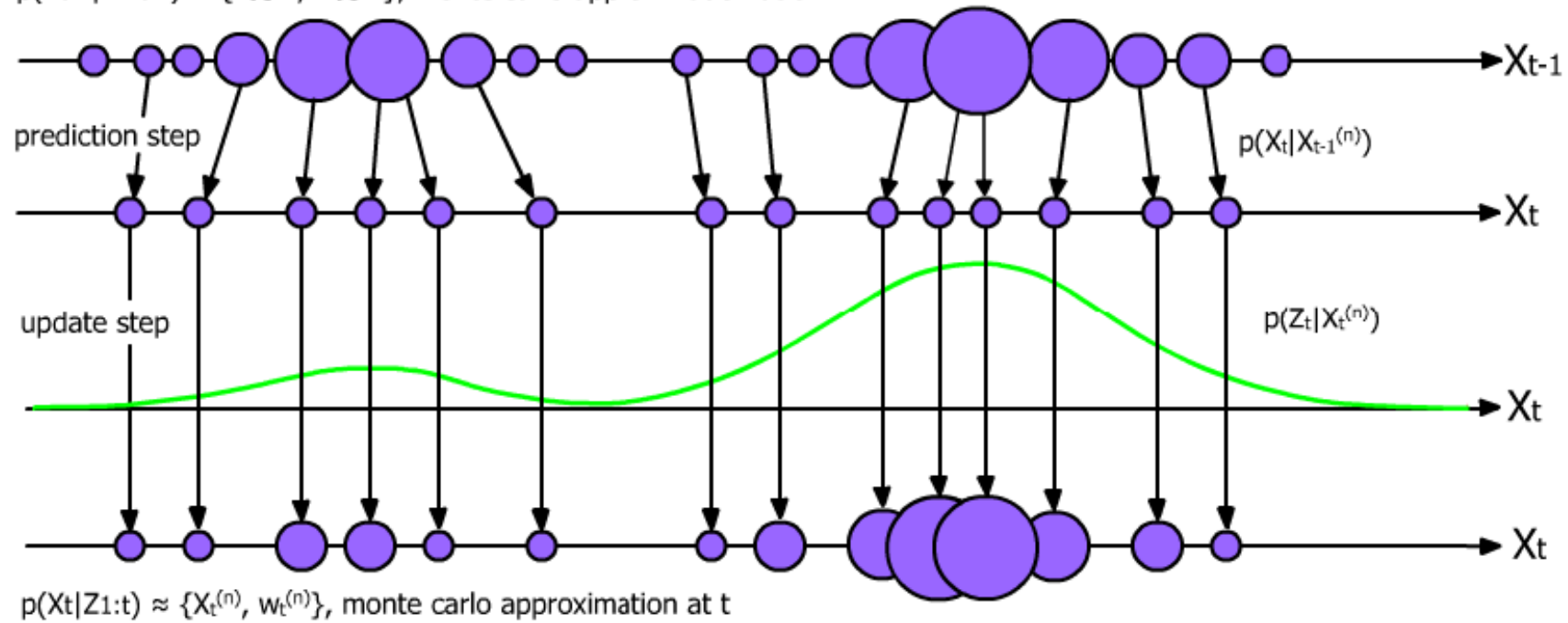


$$P(x_t | y_t) \approx P(y_t | x_t) \int P(x_t | x_{t-1}) P(x_{t-1} | y_{t-1}) dx_{t-1}$$

$$P(x_t | y_t) \approx P(y_t | x_t) \sum_r P(x_t | x_{t-1}^r)$$



$p(X_{t-1}|Z_{1:t-1}) \approx \{X_{t-1}^{(n)}, w_{t-1}^{(n)}\}$, monte carlo approximation at t-1

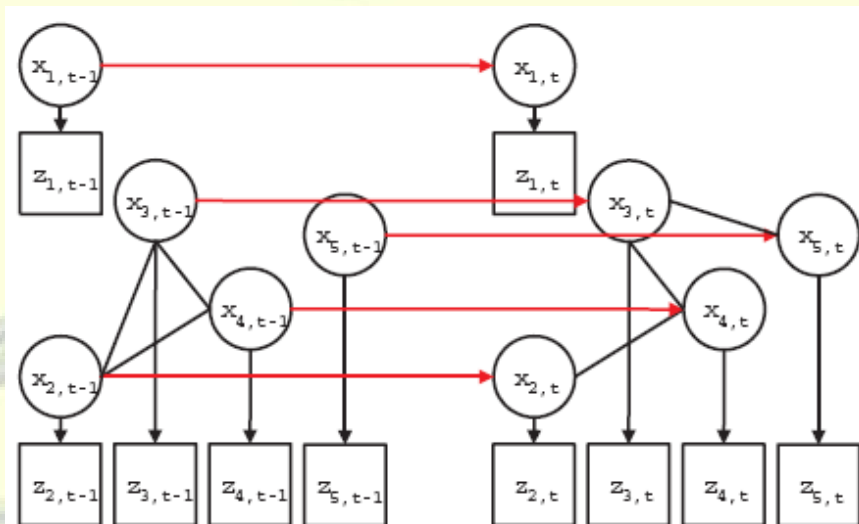


基于MRF的粒子滤波器

- 基于MRF的粒子滤波器还要考虑目标间的关系。
(n 是目标个数, E 目标间的遮挡关系)

$$P(x_t | y_t) \approx P(y_t | x_t) \sum P(x_t | x_{t-1}^r)$$

$$P(X_t | Y_t) \approx P(Y_t | X_t) \prod_{i,j \in E} \varphi(x_{it}, x_{jt}) \sum_r \prod_{i=1}^n P(x_{it} | x_{i(t-1)}^r)$$



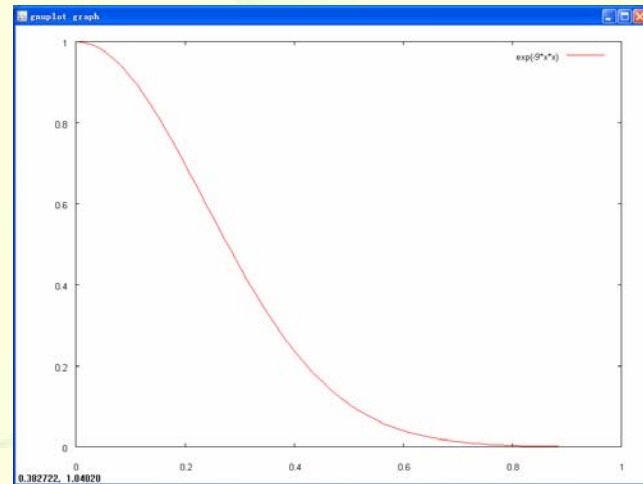
基于MRF的粒子滤波器

- 边势函数 Ψ 表达了目标间的关系，是MRF设计的关键。惩罚完全重叠的目标！

$$\psi_{ij}(x_i, x_j) \propto 1 - e^{-d(x_i, x_j)^T \Sigma^{-1} d(x_i, x_j)}$$

$$\psi(X_{it}, X_{jt}) \propto \exp(-g(X_{it}, X_{jt}))$$

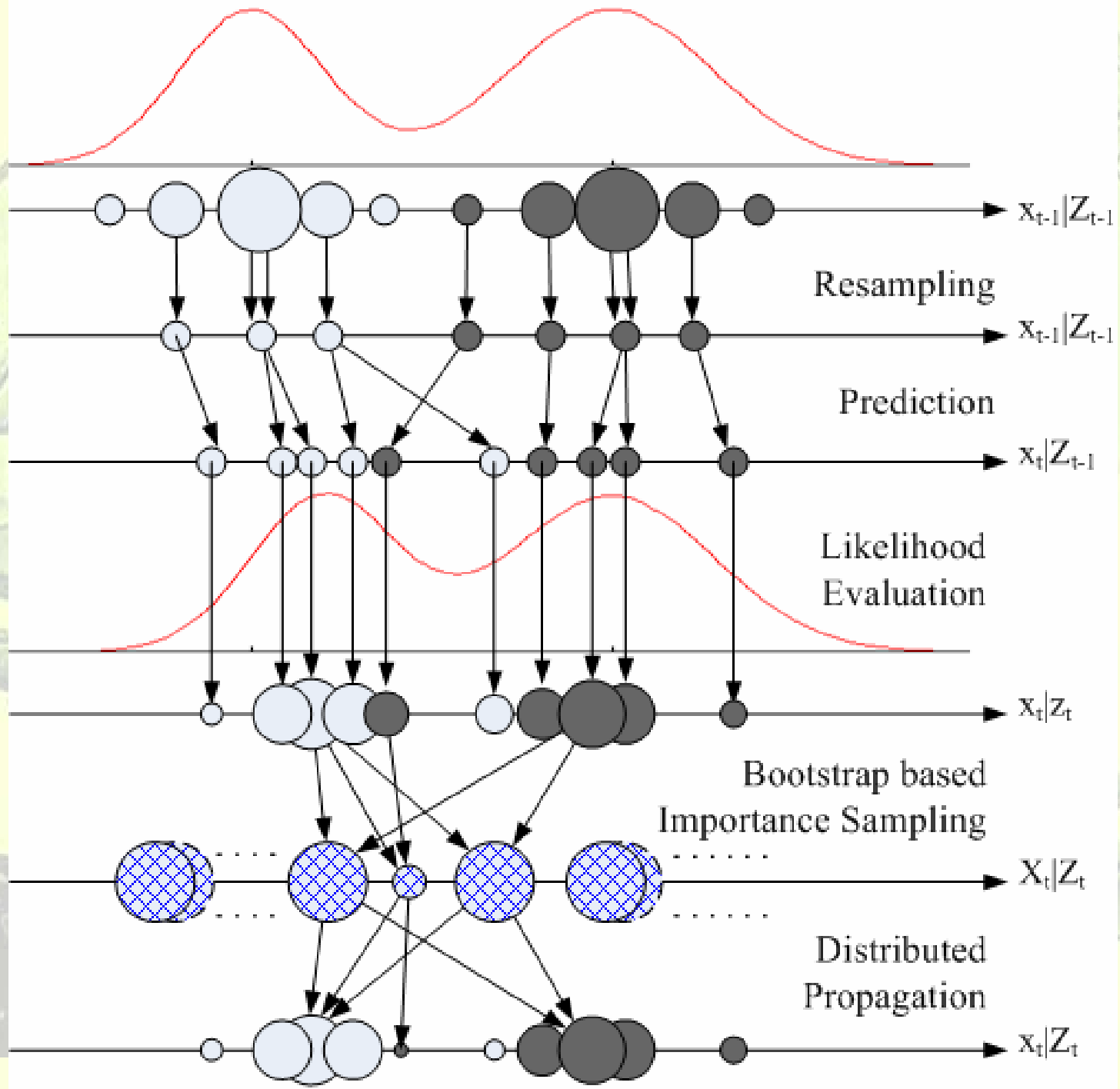
$$\varphi_{ij}(x_i, x_j) = \exp\left(-9 \times \left(\frac{\text{overlap}(x_i, x_j)}{\text{area}(x_i)}\right)^2\right)$$



[1] Z. Khan, T. Balch, and F. Dellaert. MCMC-based particle filtering for tracking a variable number of interacting targets. PAMI, 27(11):1805–1918, 2005.

[2] T. Yu and Y. Wu. Collaborative tracking of multiple targets. In IEEE CVPR, 2004.

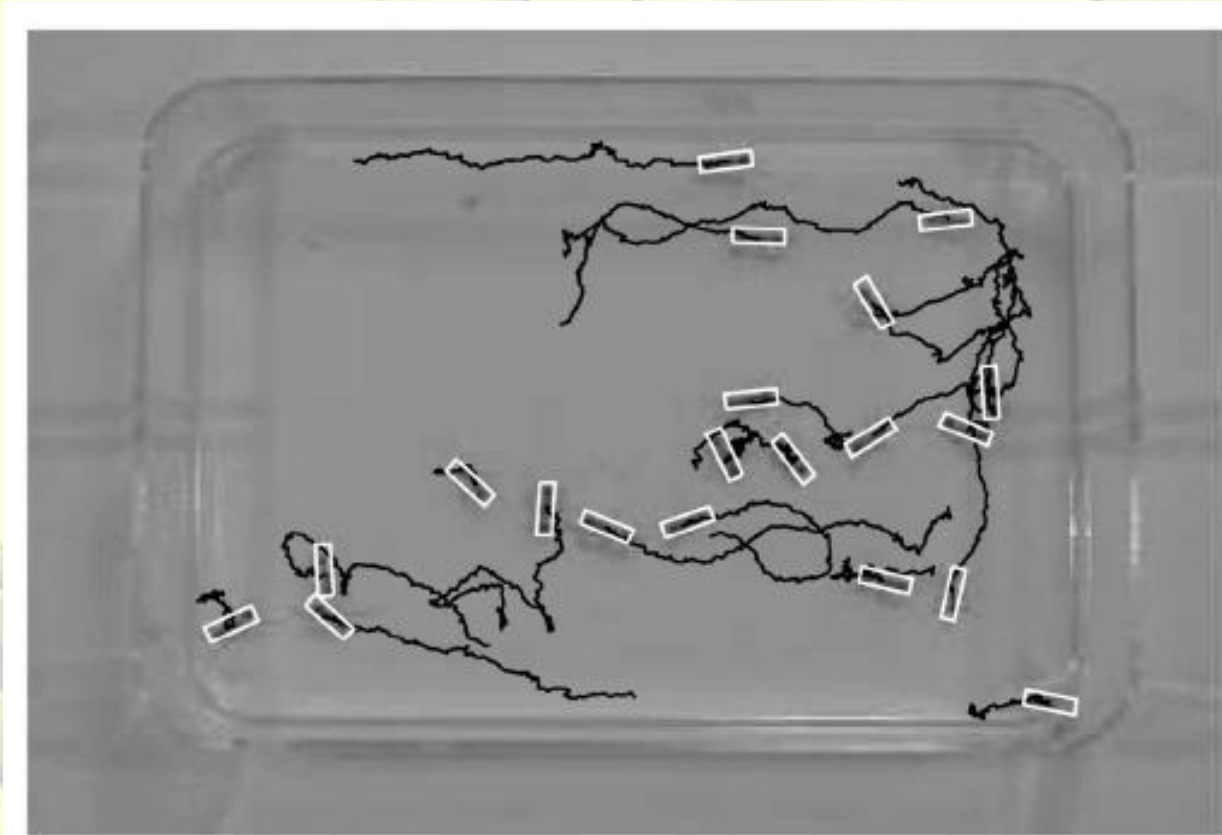
[3] Lejun Shen. (appear soon!)



基于MRF的粒子滤波器

- 如何利用粒子滤波器进行MRF推理呢？
 - JPF（联合状态粒子滤波）
 - MCMC（马尔科夫链蒙特卡洛）法
 - Mean-Field（均值场变分法）
 - DMOT（分布式法）
 - BISP（作者:沈乐君。它是一种基于MRF的实时多目标跟踪算法，尚未发表）

展示MCMC方法



* Khan et al., "MCMC-Based Particle Filtering for Tracking a Variable Number of Interacting Targets", PAMI, 2005.

展示MFMC方法

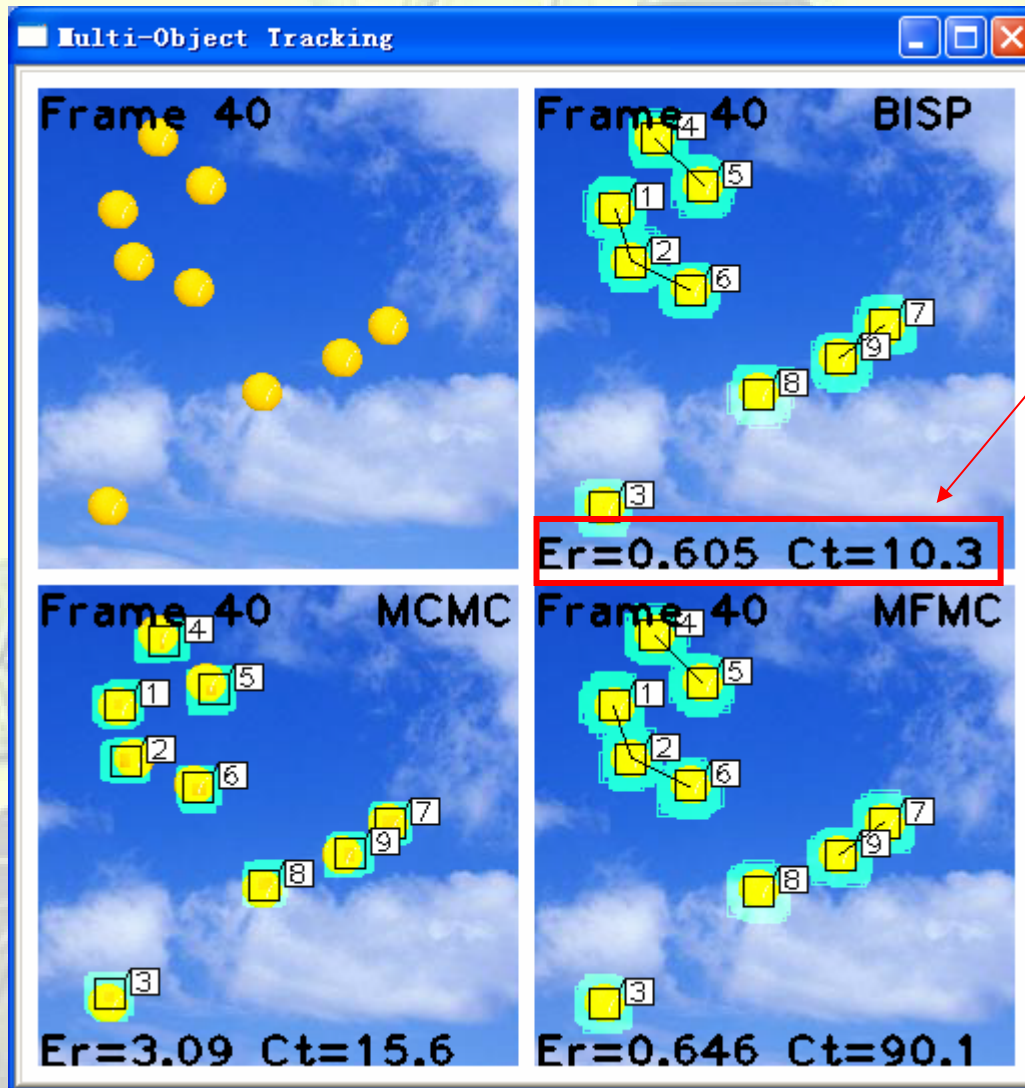


Figure 4: MFMC tracker: 5 tennis in a synthetic video. The blue links among the targets illustrate the structure of the *ad hoc* Markov network. Details please see Tennis1_MFMC.avi.




Figure 5: M.i.T. tracker: 5 tennis in a synthetic video. Details please see Tennis1_MiT.avi.

展示BISP方法



Er: 误差

Ct: 运算时间



谢谢!