

自助重要性采样用于实时多目标视觉跟踪

沈乐君^{1,2} 游志胜^{1,3} 李晓峰^{1,3}

摘要 多目标视觉跟踪的主要困难来自于多个目标交互(部分或完全遮挡)导致的歧义性。马尔可夫随机场(Markov random field, MRF)可以消除这种歧义性且无需显式的数据关联。但是,通用概率推理算法的计算代价很高。针对上述问题,本文做出了 3 点贡献:1) 设计了新的具有“分散-集中-分散”结构的递归贝叶斯跟踪框架——自助重要性采样粒子滤波器,它使用融入当前时刻观测的重要性密度函数解决维数灾难问题,将计算复杂度从指数增长变为线性增长;2) 提出了新的蒙特卡洛策略——自助重要性采样,利用 MRF 的因子分解性质进行重要性采样,并使用自助法产生低成本高质量的样本、降低似然度计算次数和维持多模式分布;3) 采用了新的边缘化技术——使用辅助变量采样进行边缘化,使用自助直方图对边缘后验分布进行密度估计。实验结果表明,本文提出的算法能够对大量目标进行实时跟踪,能够处理目标间复杂的交互,能够在目标消失后维持多模式分布。

关键词 多目标跟踪, 视觉跟踪, 粒子滤波, 马尔可夫随机场, 自助法, 重要性采样

引用格式 沈乐君, 游志胜, 李晓峰. 自助重要性采样用于实时多目标视觉跟踪. 自动化学报, 2012, 38(10): 1663–1670

DOI 10.3724/SP.J.1004.2012.01663

Real-time Visual Tracking of Multiple Targets Using Bootstrap Importance Sampling

SHEN Le-Jun^{1,2} YOU Zhi-Sheng^{1,3} LI Xiao-Feng^{1,3}

Abstract Ambiguity is the major difficulty in multi-object tracking problem due to the interactions of multiple targets (partial or complete occlusion). This ambiguity can be resolved by Markov random field (MRF) without explicit data association. However, the computational cost of general probabilistic inference algorithms of MRF is expensive. This paper presents a novel approach to this problem. Firstly, a new recursive Bayesian estimation framework, bootstrap importance sampling particle filter (BIS-PF), is devised, which has a “distributed-central-distributed” structure. The core of this framework is a suboptimal importance density which uses the observation at present time. So, it does not suffer from the curse of dimensionality. Secondly, a new Monte Carlo strategy is proposed, which uses bootstrap sampling to generate low-cost and high-quality samples, maintains multi-modality and decreases the number of likelihood computations. Thirdly, a new marginalization technology is presented, which uses an auxiliary variable sampler to obtain marginal samples and bootstrap based histogram for density estimation. The experiments show that the proposed method can track multiple targets in real-time, handle the complex interaction and maintain multi-modalities even the objects disappear.

Key words Multi-object tracking, visual tracking, particle filter, Markov random field (MRF), bootstrap, importance sampling

Citation Shen Le-Jun, You Zhi-Sheng, Li Xiao-Feng. Real-time visual tracking of multiple targets using bootstrap importance sampling. *Acta Automatica Sinica*, 2012, 38(10): 1663–1670

收稿日期 2011-08-26 录用日期 2012-03-23
Manuscript received August 26, 2011; accepted March 23, 2012
国家高技术研究发展计划(863 计划)(2012AA011804), 国家自然科学基金(61001195), 四川省科技支撑计划(2010JY0078) 资助
Supported by National High Technology Research and Development Program of China (863 Program) (2012AA011804), National Natural Science Foundation of China (61001195), and Technology Research and Development Program of Sichuan Province of China (2010JY0078)
本文责任编辑 章毓晋
Recommended by Associate Editor ZHANG Yu-Jin
1. 四川大学视觉合成图形图像技术国防重点实验室 成都 610064
2. 成都体育学院信息技术中心 成都 610041 3. 四川大学计算机学院 成都 610064
1. State Key Laboratory of Fundamental Science on Synthetic Vision, Sichuan University, Chengdu 610064 2. Department of Information, Chengdu Sport University, Chengdu 610041 3. College of Computer Science, Sichuan University, Chengdu 610064

多目标跟踪是计算机视觉中具有挑战性的问题。它的应用领域广泛,所以多年来一直是研究的热点。对于外观相似目标交互(部分或完全遮挡)导致的跟踪失败问题,目前主要有两大类解决思路:1) 显式数据关联方法,利用事件互斥规则将状态与观测关联,如联合概率数据关联(Joint probability data association, JPDA)^[1]与衍生方法^[2]、先聚类再数据关联的方法^[3–4]、先目标检测再数据关联的方法^[5–7];2) 隐式数据关联方法,它没有观测与状态间的关联模型,但有互斥规则^[8–11],其中,隐式数据关联具有实时性的潜力。因此,本文主要讨论后者。

马尔可夫随机场(Markov random field, MRF)广泛应用于图像处理相关领域^[12]。它具有马尔可夫

性质,联合概率公式可分解为多个因子的乘积,因此,可以降低数据关联的复杂度.它有许多通用概率推理算法可用于求解多目标跟踪问题,如马尔可夫链蒙特卡洛(Markov chain Monte Carlo, MCMC)^[8]、均值场^[9]、信念传播^[13]等.但是,基于MRF的隐式数据关联算法有两个缺点:1)没有针对视觉跟踪的特点来设计算法,导致计算效率低下、实时性差;2)在追求实时性的过程中,容易牺牲维持多模式分布的能力,导致样本多样性变差.这两个缺点阻碍了MRF在多目标跟踪领域的发展与应用.

计算效率低下的原因与解决办法如下:1)没有对视觉跟踪进行优化.与雷达不同,视觉跟踪的特点是“计算似然度(提取图像特征)是跟踪算法中计算量最大的操作”.因此,减少似然度的计算次数,或者降低粒子个数,就能够提高计算效率.Khan等使用特殊的提案分布降低似然度的计算次数^[8];Chang等使用均值平移法降低粒子个数^[4];Yu等使用变分法降低粒子个数^[9]等.本文使用自助法(Bootstrap)^[14-15]来降低似然度的计算次数.2)维数灾难.为了处理目标交互导致的跟踪失败问题,多目标跟踪器必须建立高维模型.无论JPDA模型中数据关联假设的个数,还是MRF模型中样本的个数,都会随着目标个数的增加呈指数级增长.当目标个数大于4时,MRF-JPF(Joint particle filters)几乎无法使用^[8].蒙特卡洛仿真适合解决高维问题^[16],使用少量高质量样本就可以得到满意的结果.本文解决维数灾难的思路是:使用边缘后验分布^[11]减少样本尺寸,使用融入当前时刻观测信息^[17]的提案分布产生高质量样本,使用重要性采样纠正偏差.3)分布式模型缺乏高效的实现方法.与集中式方法^[8]相比,分布式方法^[9-11]更加高效.Lanz提出的HJS(Hybrid joint-separable)模型^[11]与本文的信念传播结构类似.前者的联合分离机制更加复杂,并使用均值场来计算边缘分布.本文使用更简单的辅助变量采样法^[16]来计算边缘分布,使用自助直方图完成密度估计.

维持多模式分布对于多目标跟踪问题至关重要.外观相似目标交互导致似然度分布呈现多模式分布^[3](或非高斯分布^[18]).多目标跟踪器需要根据运动模型、观测模型、目标检测响应、数据关联规则或MRF模型等,为多个模式赋予正确的ID.但是,并非所有的目标都存在对应的模式.当目标被完全遮挡后,该目标在图像中消失,因此它没有对应的模式.JPDA模型将消失目标与第0个观测关联.MRF模型通过多模式后验分布,维持消失目标与似然度分布中多个模式之间的多重假设(关联).粒子滤波器具有维持多重假设的能力.因此,基于MRF的粒子滤波器能够处理完全遮挡导致的目标消失.

可是,现有方法在提高计算效率的同时,容易牺

牲(或忽视)维持多模式分布的能力.例如,Khan等提出的MCMC方法^[8]提高了计算效率.但是因为MCMC仅传播10个样本到下一时刻,这些样本很难覆盖后验分布的全部模式,所以当消失目标再次出现时算法出错.Qu等提出的DMOT(Distributed multi-object tracking)方法^[10]实时性很强.但是,DMOT在多次迭代后,后验分布会收敛到单个模式附近,算法出错.本文利用自助法不改变总体分布的特性,确保维持多模式分布的能力不丧失.实验证明,具备这种能力的跟踪器比MCMC和DMOT的稳定性更好.

本文的主要贡献有三点:1)提出了一种新的多目标跟踪框架:自助重要性采样粒子滤波器,其具有“分散-集中-分散”的信念传播结构;2)提出了一种新的蒙特卡洛策略:自助重要性采样,其将计算复杂度从指数级增长变为线性增长;3)提出了一种新的维持多模式分布的思路,它使用自助法维持多模式分布,并使用自助直方图进行密度估计.

1 问题描述

本文将多目标跟踪看作递归贝叶斯估计问题.设目标个数为 M ,粒子个数为 N .记 $x_{i,t}$ 为第 i 个目标在 t 时刻的状态, $x_{i,t} \in S$. S 表示状态空间. M 个目标的联合状态为 $\mathbf{X}_t = \{x_{1,t}, x_{2,t}, \dots, x_{M,t}\}$, $\mathbf{X} \in S^M$. S^M 表示配置空间.目标 $x_{i,t}$ 的图像观测为 $y_{i,t}$.联合状态的观测为 $\mathbf{Y}_t = \{y_{1,t}, y_{2,t}, \dots, y_{M,t}\}$.记马尔可夫随机场为 $G = \{V, E\}$.其中, V 表示节点,每个节点对应一个目标,共有 M 个随机变量; E 表示无向边,即随机变量在空间上的邻域关系.如果 $x_{i,t}$ 与 $x_{j,t}$ 之间存在遮挡,则在 E 中存在一条边连接 $x_{i,t}$ 与 $x_{j,t}$.

下面是本文的核心公式:

预测与更新

$$P(x_{i,t}|y_{i,t}) = \frac{1}{Z_i} P(y_{i,t}|x_{i,t}) \times \int P(x_{i,t}|x_{i,t-1}) P(x_{i,t-1}|\mathbf{Y}_{t-1}) dx_{i,t-1} \quad (1)$$

联合更新

$$P(\mathbf{X}_t|\mathbf{Y}_t) = \frac{1}{Z_G} \prod_{ij \in E} \varphi(x_{i,t}, x_{j,t}) \underbrace{\prod_{j \in V} P(x_{j,t}|y_{j,t})}_{Q(\mathbf{X}_t|\mathbf{Y}_t)} \quad (2)$$

边缘化

$$P(x_{i,t}|\mathbf{Y}_t) = \int P(\mathbf{X}_t|\mathbf{Y}_t) dx_{-i,t} \quad (3)$$

式(1)根据先验 $P(x_{i,t-1}|\mathbf{Y}_{t-1})$ 和运动模型 $P(x_{i,t}|x_{i,t-1})$ 对第 i 个目标的状态进行预测,然后,

根据观测模型 $P(y_{i,t}|x_{i,t})$ 得到后验分布 $P(x_{i,t}|y_{i,t})$ (亦被称为局部证据). 本文使用自助粒子滤波来实现式 (1). 式 (2) 使用点对马尔科夫随机场模型对目标的空间关系进行建模. 其核心思想是惩罚重叠^[8]. 它的作用是对联合状态的后验概率分布 $P(\mathbf{X}_t|\mathbf{Y}_t)$ 进行联合更新. 本文使用自助重要性采样来实现式 (2), 其重要性函数 (提案分布) 是 $Q(\mathbf{X}_t|\mathbf{Y}_t)$. $\varphi(x_{i,t}, x_{j,t})$ 是二阶簇势函数 (Potential). 式 (3) 的作用是计算边缘后验分布 $P(x_{i,t}|\mathbf{Y}_t)$. 它将全局高维联合分布 $P(\mathbf{X}_t|\mathbf{Y}_t)$ 映射到局部低维状态空间. 本文使用辅助变量采样与自助直方图来实现式 (3). 其中, $x_{-i,t}$ 表示联合状态 \mathbf{X}_t 中移除 $x_{i,t}$ 后剩余的状态向量. Z_i 和 Z_G 是归一化系数. 当前时刻的边缘后验分布 $P(x_{i,t}|\mathbf{Y}_t)$ 是下一时刻的先验.

自助重要性采样是本文最主要的创新, 因此, 将上述跟踪框架称为: 自助重要性采样粒子滤波器 (Bootstrap importance sampling particle filter, BIS-PF). 下一节将详细阐述其实现细节.

2 自助重要性采样粒子滤波器

2.1 自助重要性采样

自助重要性采样源于一个事实: 在视觉跟踪中, 耗时最多的操作是计算似然度 (提取图像特征). 这是图像传感器不同于雷达、声纳等传感器的一大特点, 也是缺点. 因此, 减少似然度的计算次数, 就能提高计算效率.

自助 (Bootstrap) 采样法, 是一种被广泛使用的基于数据集的采样方法. Gordon 等^[15] 将自助法用于 SIS (Sequential importance sampling) 滤波器, 提出使用重采样解决退化问题, 引发了粒子滤波器的研究热潮. 由于自助采样的计算代价远远小于直接从似然度分布中采样, 因此, 本文使用自助法来降低似然度的计算次数. 其次, MRF 模型的因子分解性质使得重要性采样^[19] 易于使用. 虽然 Khan 等的实验结果显示: 由于维数灾难的原因, 很难将重要性采样直接用于高维 MRF 模型^[8]. 但是, 如果重要性函数 (提案分布) 设计合理, 重要性采样依然可以用于高维模型. 为此, 本文在式 (2) 中设计了新的提案分布 Q . 因为 Q 融入了当前时刻的观测 $y_{i,t}$, 所以它的外形比 $P(\mathbf{X}_t|\mathbf{X}_{t-1})$ ^[8] 更接近 $P(\mathbf{X}_t|\mathbf{Y}_t)$. 它能够产生质量更高的样本, 故所需样本数更少, 能显著地缩短计算时间. 而且, 自助法使得从 Q 中采样非常容易. 具体而言, 本文提出分步实施的自助重要性采样策略: 1) 使用重要性采样在局部低维状态空间进行初步探索 (采样 N 次); 2) 运用自助法根据初步探索得到的数据集生成提案分布; 3) 使用重要性采样在全局高维配置空间中探索 (采样 K 次). 下面介绍算法细节.

首先, 对 M 个目标分别独立运行自助粒子滤波器, 得到近似表示式 (1) 的粒子集合:

$$P(x_{i,t}|y_{i,t}) \approx \left\{ x_{i,t}^{(n)}, \pi_{i,t}^{(n)} = P(y_{i,t}|x_{i,t}^{(n)}) \right\}_{n=1}^N \quad (4)$$

然后, 对 M 个目标各自独立地以 $\pi_{i,t}$ 为概率有放回地选出 K 个无权重自助样本:

$$P(x_{i,t}|y_{i,t}) \approx \left\{ x_{i,t}^{(k)*} \right\}_{k=1}^K, \quad i = 1, \dots, M \quad (5)$$

其中, 上标 “*” 表示自助样本, k 是序号. 接着, 将 k 相同的 M 个自助样本组合成为联合状态样本:

$$\mathbf{X}_t^{(k)*} = \left(x_{1,t}^{(k)*}, \dots, x_{M,t}^{(k)*} \right), \quad k = 1, \dots, K \quad (6)$$

由于式 (5) 中各个目标的自助采样相互独立, 故式 (6) 符合式 (2) 中的提案分布 Q :

$$\mathbf{X}_t^{(k)*} \sim Q(\mathbf{X}_t|\mathbf{Y}_t) = \prod_{j \in V} P(x_{j,t}|y_{j,t}) \quad (7)$$

最后, 根据联合概率式 (2) 计算重要性权重, 得到近似表示式 (2) 的样本集合:

$$P(\mathbf{X}_t|\mathbf{Y}_t) \approx \left\{ \mathbf{X}_t^{(k)*}, w_t^{(k)} = \prod_{ij \in E} \varphi(x_{i,t}^{(k)*}, x_{j,t}^{(k)*}) \right\}_{k=1}^K \quad (8)$$

自助重要性采样有 3 个优点: 1) 计算时间短. 由于 K 大于 N , 且自助采样的计算代价远远小于直接从似然度分布中采样 (例如, $K = 4000$, $N = 200$, 直接采样 0.15 毫秒, 自助采样 0.004 毫秒). 因此, BIS 法抽取 $N + K$ 个样本的计算时间 ($200 \times 0.15 + 4000 \times 0.004 = 46$ 毫秒) 远远小于直接抽取 K 个样本的计算时间 ($4000 \times 0.15 = 600$ 毫秒). 2) 样本质量高. 因为重要性采样式 (4) 在低维状态空间中很高效, 且提案分布 Q 融入了当前时刻的观测信息, 所以式 (8) 的样本质量很高^[17]. 3) 维持多模式分布的能力强. 这是本文与其他实时多目标跟踪算法相比最主要的区别, 后文将详细叙述.

2.2 边缘后验分布与自助直方图

自助重要性采样粒子滤波器在贝叶斯滤波器的基础之上, 增加了描述目标空间关系的 MRF 模型和联合更新步骤. MRF 模型的维度高 (与目标个数 M 成正比). 虽然自助重要性采样提高了联合更新的计算效率, 但是表示联合分布 $P(\mathbf{X}_t|\mathbf{Y}_t)$ 的样本集式 (8) 依然是高维的, 且无法直接传播到下一时刻.

相比而言, 边缘后验分布 $P(x_{i,t}|\mathbf{Y}_t)$ 的维度低, 所需粒子个数与目标个数呈线性关系, 因此, 它

更紧凑和高效^[11]. 将 $P(\mathbf{X}_t|\mathbf{Y}_t)$ 边缘化即可得到 $P(x_{i,t}|\mathbf{Y}_t)$. 信念传播与均值场是计算边缘分布的有效方法, 并被广泛应用于图像处理领域. 但是, 它们的计算代价较大. 本文采用更简单的方法: 辅助变量采样法^[16]. 对 $x_{i,t}$ 而言, 联合状态向量 \mathbf{X}_t 中除 $x_{i,t}$ 之外的 $M-1$ 个分量 (即 $x_{\rightarrow i,t}$) 被视为辅助变量. 因此, 丢弃式 (8) 中的辅助变量分量, 就可以得到符合边缘后验分布 $P(x_{i,t}|\mathbf{Y}_t)$ 的样本:

$$P(x_{i,t}|\mathbf{Y}_t) \approx \left\{ x_{i,t}^{(k)*}, w_t^{(k)} \right\}_{k=1}^K \quad (9)$$

为了将式 (9) 传播到下一个时刻, 本文设计了自助直方图 (Bootstrap based histogram). 在式 (5) 中的每个样本都是从式 (4) 中依概率 π 选出的. 令 $B_i(k)$ 表示式 (5) 与式 (4) 之间的映射关系, 即式 (5) 中第 k 个样本对应式 (4) 中第 $B_i(k)$ 个样本. 当自助采样实施后, 映射关系 $B_i(k)$ 就确定下来. 由于在目标跟踪问题中, 感兴趣的是目标状态的期望, 将 $B_i(k)$ 相同的项合并, 并不改变期望的最终结果. 因此, 可以将 $B_i(k)$ 相同的样本权重累加起来, 形成以 N 个样本为支撑的自助直方图:

$$P(x_{i,t}|\mathbf{Y}_t) \approx \left\{ x_{i,t}^{(n)}, W_{i,t}^{(n)} = \sum_{B_i(k)=n} w_t^{(k)} \right\}_{n=1}^N \quad (10)$$

自助直方图的优点有 3 方面: 1) 适合序列数据. 生成直方图式 (10) 后, 式 (9) 中 K 个序列数据全部可以丢弃. 由于无需保存 K 个样本, 因此, 本文对 K 的大小无限制. 自助法的计算代价极低, 所以提高 K 值后, BIS-PF 既能模拟复杂的模型与分布, 又具有实时性. 2) 紧凑高效. 注意式 (10) 与式 (4) 的目标状态部分完全相同. 因此, 只需增加 1 个浮点数的存储空间, 就可以将式 (10) 与式 (4) 存储在一起. 从期望的角度看, 式 (9) 与式 (10) 等价, 但后者

更紧凑. 3) 信息可直接传播. 式 (10) 可以直接作为下一个时刻的先验. 总之, 自助直方图简洁、紧凑、高效, 是将自助重要性采样用于多目标跟踪的关键.

综上所述, 自助重要性采样粒子滤波器具有“分散-集中-分散”的信念传播结构. 图 1 概括了它的工作流程和各个模块的相互关系.

2.3 算法实现

隐式数据关联可以用 MRF 模型的势函数来表达^[8-11]. 下面是本文所用的势函数:

$$\varphi(x_i, x_j) = \exp \left(-\beta \left(\frac{\text{overlap}(x_i, x_j)^2}{\text{area}(x_i)\text{area}(x_j)} \right)^2 \right) \quad (11)$$

其中, β 是互斥系数, area 是粒子的矩形面积, overlap 是两个矩形的重叠面积 (单位: 像素).

下面是 BIS-PF 的伪代码:

```

输入.  $t-1$  时刻的粒子集合  $\{x_{i,t-1}^{(n)}, W_{i,t-1}^{(n)}, \pi_{i,t-1}^{(n)}\}_{n=1}^N$ 
步骤 1. 运行  $M$  个独立的粒子滤波器
For  $i = 1 : M$ 
    1) 重采样: 从  $\{x_{i,t-1}^{(n)}, W_{i,t-1}^{(n)}\}_{n=1}^N$  中根据  $W_{i,t-1}^{(n)}$  进行重采样, 得到新的  $N$  个样本  $\{x_{i,t-1}^{(n)}\}_{n=1}^N$ ;
    2) 预测与更新: 计算  $P(x_{i,t}|x_{i,t-1}^{(n)})$  和  $\pi_{i,t}^{(n)} = P(y_{i,t}|x_{i,t}^{(n)})$ , For  $n = 1, \dots, N$ ;
    3) 清零:  $W_{i,t}^{(n)} = 0$ , For  $n = 1, \dots, N$ 
End
步骤 2. 自助重要性采样
For  $k = 1 : K$ 
    1)  $w = 1.0$ ;
    2) 自助采样: 从  $\{x_{i,t}^{(n)}\}_{n=1}^N$  中根据概率  $\pi_{i,t}^{(n)}$  选中样本, 将其序号保存到  $B(i)$ , for  $i = 1, \dots, M$ ;
For  $i = 1 : M$ 
    For  $j = i+1 : M$ 
        3)  $o = \text{overlap}(x_{i,t}^{(B(i))}, x_{j,t}^{(B(j))})$ ,
    
```

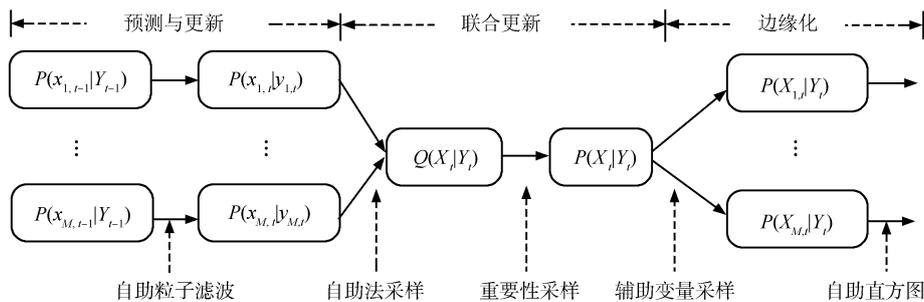


图 1 自助重要性采样粒子滤波器的信念传播结构图

Fig. 1 Architecture of bootstrap importance sampling particle filter

$$a_j = \text{area}(x_{j,t}^{(B(j))}), a_i = \text{area}(x_{i,t}^{(B(i))})$$

4) 重要性权重: If $o > 0$, Then

$$w = w \times \exp\left(-\beta \left(\frac{o^2}{a_i a_j}\right)^2\right)$$

End

End

5) 边缘化与自助直方图累加: $W_{i,t}^{(B(i))} = W_{i,t}^{(B(i))} + w$, for $i = 1, \dots, M$

End

3 实验与结果

本文的实验环境是 Pentium 4 1.4GHz CPU, 512MB 内存. 参与评估的算法有 MRF-JPF^[8]、MCMC^[8]、MFMC (Mean field Monte Carlo)^[9]、DMOT^[10]、多个独立粒子滤波器 (Multiple independent paricle filters, MIPF) 和 BIS-PF. 全部算法用 VS 2005 开发, 并对 SSE 指令优化. 观测模型是色彩直方图的 Bhattacharyya 距离, 运动模型是恒速度方程^[20]. 本文的缺省实验参数是: 互斥系数 $\beta = 4$; MRF-JPF 的 $N = 6000$; MCMC 的 $N = 6000$, 传播 10 个最佳样本到下一时刻, 预烧期抛弃 25% 的样本; MFMC 迭代次数 5, $N = 400$; DMOT 迭代次数 6, $N = 400$; BIS-PF 的 $K = 4000$, $N = 400$.

3.1 维持多模式分布能力的实验

为了测试多目标跟踪算法是否具有维持多模式分布的能力, 本文设计了新的实验.

实验 1 (“3123”实验). 如图 2 所示, 三个外观相同目标之间的相互遮挡, 导致图像中可见目标的个数不断变化 (依次为 3-1-2-3). 在 51 帧, 目标 2 和 3 被目标 1 完全遮挡后消失; 从 51 帧到 93 帧, 目标 3 不可见; 在 93 帧, 目标 1 分裂为目标 1 和 3. 从 51 帧到 93 帧期间, 目标 3 可能被目标 1 遮挡 (假设 1), 也有可能被目标 2 遮挡 (假设 2). 所以, 跟踪器必须维持 2 模式分布, 才能在目标 3 出现时正确跟踪它. 实验证明: MCMC^[8] 和 DMOT^[10] 只能维持单模式, 跟踪失败. BIS-PF、MRF-JPF^[8] 和 MFMC^[9] 能维持 2 个模式, 跟踪成功.

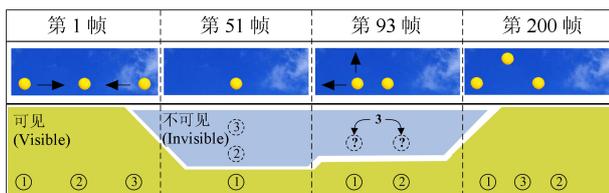


图 2 测试多模式分布能力的“3123”实验

Fig. 2 “3123” experiment: testing the ability of maintaining the multi-modality

3.2 同时跟踪多个目标的实验

本文的实验数据中既有真实视频, 也有合成视频. 其中, 合成视频实时生成, 目标是 20 像素 × 20 像素的网球, 目标运动是恒速度运动方程, 并受高斯噪声 (方差 = 0.4 像素) 的扰动而表现为随机漫步, 最大速度限制在 2 像素/帧. 与通常的测试视频 (如匀速前进的行人或车辆) 不同, 随机漫步合成视频存在难以预测的复杂运动 (如目标在遮挡前后改变运动方向) 和严重遮挡 (如 3 个以上目标长时间混叠). 这些都给多目标跟踪算法带来了挑战.

实验 2. 合成视频, 同时跟踪 16 个目标. 实验证明: 1) 自助重要性采样的样本质量高, 100 个样本就具备同时跟踪 16 个目标的基本能力. 2) BIS-PF 比 MCMC、MRF-JPF 和 MFMC 的计算时间更短; 与 MCMC 和 DMOT 相比, 由于具备维持多模式的能力, BIS-PF 更稳定.

实验 3. 合成视频, 同时跟踪 36 个目标. 由于目标个数多, 容易出现严重遮挡 (图 3). 实验证明: BIS-PF 能够实时跟踪大量密集目标 (5.52 fps).

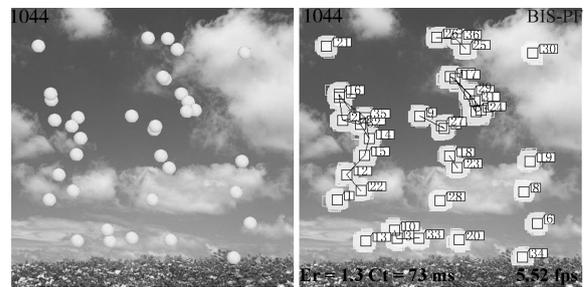


图 3 实时跟踪 36 个目标

Fig. 3 Real-time tracking of 36-objects

实验 4 (足球比赛视频). 该视频存在严重遮挡 (图 4). MRF-JPF 需要的粒子个数太多, 无法接受; MCMC 运算时间较长, 误差较大; MFMC 与 DMOT 不太稳定; BIS-PF 误差小且稳定.

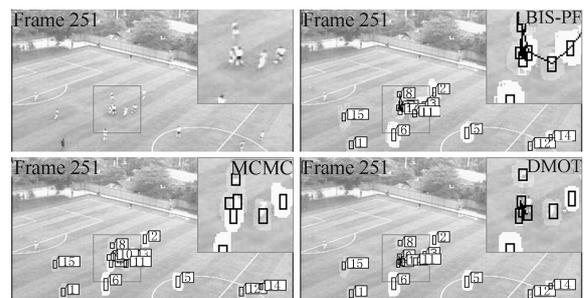


图 4 在 251 帧, 3 名运动员因身体对抗发生严重遮挡
Fig. 4 The severe occlusion of 3-players at Frame 251

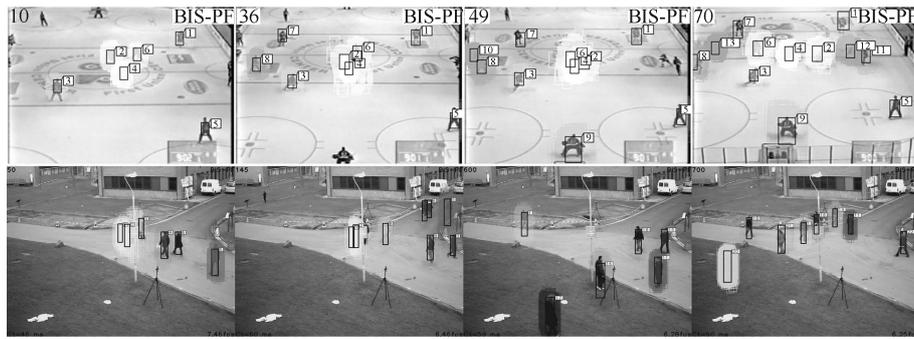


图 5 曲棍球视频^[6] (第 1 行) 与 PETS'09 视频^[21] (第 2 行) 的跟踪结果

Fig. 5 Tracking results of hockey sequence^[6] (the first row) and PETS'09 sequence^[21] (the second row)

实验 5 (曲棍球视频)^[6]. 该视频存在目标的出现、消失和交互 (图 5). 由于它没有如图 4 的严重遮挡, 故 DMOT 与 BIS-PF 的跟踪性能几乎相同. 若不进行在线目标检测, BIS-PF 的处理速度约 15.4 fps.

实验 6 (PETS'09 视频)^[21]. 该视频存在行人的出现、消失、严重遮挡和被前景 (标志牌) 遮挡. 在 BIS-PF 中, 使用加性融合目标检测信息^[6-7, 22]. 若不进行在线行人检测, BIS-PF 的处理速度约 7.1 fps.

实验结果与评价. 本文的实验评价指标有 3 个. 1) 跟踪算法的运行时间 (单位: 毫秒/帧). 运行时间不包括视频生成 (或解码), RGB 到 HSV 变换, 结果分析与可视化等其他步骤. 2) 跟踪误差, 即粒子集均值与真值间的距离 (单位: 像素). 真实视频的真值依靠人工逐帧标记. 3) 稳定性. 运行 10 次实验, 如果没有出错 (实验 2 和实验 3 在 1000 帧内) 为好, 出错 2 次以上为中, 出错 5 次以上为差.

表 1 跟踪算法的平均运行时间 (毫秒/帧)

Table 1 Average computation time (ms/frame)

实验	BIS-PF	MCMC	MFMC	DMOT
实验 1	5.36	25.61	352.88	5.66
实验 2	24.84	64.43	2145.99	21.88
实验 3	72.02	-	-	-
实验 4	18.56	52.80	438.47	16.02
实验 5	30.65	72.22	1146.63	29.73
实验 6	54.37	-	-	-

表 2 跟踪算法的平均误差 (像素)

Table 2 Average error per target (pixel)

实验	BIS-PF	MCMC	MFMC	DMOT
实验 2	0.91	(出错)	2.72	0.88
实验 4	1.96	2.29	1.99	2.80

表 3 跟踪算法的稳定性

Table 1 Stability of algorithms

实验	BIS-PF	MCMC	MFMC	DMOT
实验 1	好	差	好	差
实验 2	好	差	好	好
实验 4	好	中	中	中
实验 5	好	差	差	好

4 讨论与结论

本文提出了一种多目标跟踪算法: 自助重要性采样粒子滤波器 (BIS-PF). 1) 它使用自助法降低似然度的计算次数, 使用边缘分布降低维度和样本尺寸, 使用融入当前时刻观测的提案分布获得高质量样本. 因此, 该算法实时性强. 2) 它使用 MRF 模型对多个目标的空间关系进行建模. 它具有马尔可夫性质, 能够降低计算的复杂度; 使用空间互斥势函数避免显式数据关联, 降低了问题的难度^[23]; 具有可因子分解性质, 能够使用重要性采样. 与信念传播、均值场、MCMC 等相比, 重要性采样更简单高效. 3) 自助法不改变总体的分布, 因此, 它维持多模式分布的能力强. 4) 它使用蒙特卡洛仿真来实现全部核心公式, 是完全非参数的, 可兼容任意的分布和模型. 因此, 该算法的可扩展性强. 5) 自助采样、重要性采样、辅助变量采样和自助直方图, 都比较简单. 因此, 该算法简单高效、容易实现.

与显式数据关联相比, 隐式数据关联算法的缺点是: 外观相似目标完全遮挡会导致 ID 交换错误^[10]. 试验发现, 互斥系数 β 的大小影响着跟踪器的行为. β 越小, BIS-PF 越可能出现绑架错误^[8], 但没有 ID 交换错误^[10]; 反之, β 越大越可能导致 ID 交换错误, 但没有绑架错误. 本文中, β 是固定经验值, 所以动态调节 β 有助于避免这两类错误. 因此, 建议根据合并分裂动态改变互斥系数 β 来消除上述错误. 1) 在步骤 1 结束后, 为 E 中每条边 (连接 $x_{i,t}$ 与 $x_{j,t}$ 的边) 计算出各边不同的互斥系数 β_{ij} , 若后

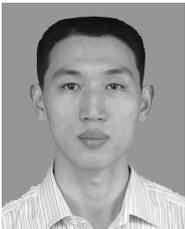
验分布大部分重叠且仅有 1 个模式 (Merge), 则 β 采用很低的值, 允许合并; 反之 (Split) 采用很高的值, 强化惩罚. 2) 在步骤 2 中, 根据 β_{ij} (而不是 β) 来计算权重. 该方法的优点是不会干扰运动方程的预测.

实时性与遮挡处理能力^[24] 是一对矛盾. 自助重要性采样在这对矛盾中取得了平衡. 研究能力更强更复杂, 且具有实时性的多目标跟踪算法将是我们下一步工作的目标.

References

- 1 Bar-Shalom Y, Fortmann T E. *Tracking and Data Association*. San Diego: Academic Press Professional, 1988
- 2 Oh S, Russell S, Sastry S. Markov chain Monte Carlo data association for multi-target tracking. *IEEE Transactions on Automatic Control*, 2009, **54**(3): 481–497
- 3 Vermaak J, Doucet A, Pérez P. Maintaining multi-modality through mixture tracking. In: Proceedings of the 9th IEEE International Conference on Computer Vision. Nice, France: IEEE, 2003. 1110–1116
- 4 Chang C, Ansari R, Khokhar A. Multiple object tracking with kernel particle filter. In: Proceedings of the 2005 IEEE Conference on Computer Vision and Pattern Recognition. San Diego, USA: IEEE, 2005. 566–573
- 5 Li Y, Ai H Z, Yamashita T, Lao S H, Kawade M. Tracking in low frame rate video: a cascade particle filter with discriminative observers of different life spans. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, **30**(10): 1728–1740
- 6 Okuma K, Taleghani A, Freitas N D, Little J J, Lowe D G. A Boosted Particle filter: multitarget detection and tracking. In: Proceedings of European Conference on Computer Vision. Prague, Czech Republic: Springer, 2004. 28–39
- 7 Breitenstein M D, Reichlin F, Leibe B, Koller-Meier E, van Gool L. Online multiperson tracking-by-detection from a single, uncalibrated camera. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, **33**(9): 1820–1833
- 8 Khan Z, Balch T, Dellaert F. MCMC-based particle filtering for tracking a variable number of interacting targets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, **27**(11): 1805–1819
- 9 Yu T, Wu Y. Collaborative tracking of multiple targets. In: Proceedings of the 2004 IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE, 2004. 834–841
- 10 Qu W, Schonfeld D, Mohamed M. Real-time distributed multi-object tracking using multiple interactive trackers and a magnetic-inertia potential model. *IEEE Transactions on Multimedia*, 2007, **9**(3): 511–519
- 11 Lanz O. Approximate Bayesian multibody tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, **28**(9): 1436–1449
- 12 Li S Z. *Markov Random Field Modeling in Image Analysis*. Berlin: Springer, 2009
- 13 Xue J R, Zheng N N, Zhong X P. Sequential stratified sampling belief propagation for multiple targets tracking. *Science in China Series F: Information Sciences*, 2006, **49**(1): 48–62
- 14 Efron B, Tibshirani R J. *An Introduction to the Bootstrap*. New York: Chapman and Hall, 1993
- 15 Gordon N, Salmond D J, Smith A F M. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEEE Proceedings F: Radar and Signal Processing*, 1993, **140**(2): 107–113
- 16 Andrieu C, de Freitas N, Doucet A, Jordan M I. An introduction to MCMC for machine learning. *Machine Learning*, 2003, **50**(1–2): 5–43
- 17 Pitt M K, Shepard N. Filtering via simulation: auxiliary particle filters. *Journal of the American Statistical Association*, 1999, **94**(446): 590–599
- 18 Isard M, Blake A. Condensation-conditional density propagation for visual tracking. *International Journal of Computer Vision*, 1998, **29**(1): 5–28
- 19 Liu J S. *Monte Carlo Strategies in Scientific Computing*. New York: Springer, 2001
- 20 Shen Le-Jun, Ke Zun-Yu, Cheng Xiao-Ping. The real time tracking on athlete during the competition of ball games. *China Sport Science*, 2007, **27**(1): 64–67
(沈乐君, 柯遵渝, 程小平. 球类比赛中运动员的实时跟踪. *体育科学*, 2007, **27**(1): 64–67)

- 21 PETS 2009 Benchmark Data [Online], available: <http://www.cvg.rdg.ac.uk/WINTERPETS09/>, April 1, 2012
- 22 Gu Xin, Wang Hai-Tao, Wang Ling-Feng, Wang Ying, Chen Ru-Bing, Pan Chun-Hong. Fusing multiple features for object tracking based on uncertainty measurement. *Acta Automatica Sinica*, 2011, **37**(5): 550–559
(顾鑫, 王海涛, 汪凌峰, 王颖, 陈如冰, 潘春洪. 基于不确定性度量的多特征融合跟踪. 自动化学报, 2011, **37**(5): 550–559)
- 23 Yan Xiao-Xi, Han Chong-Zhao. Multiple target tracking algorithm based on online estimation of target birth intensity. *Acta Automatica Sinica*, 2011, **37**(8): 963–972
(闫小喜, 韩崇昭. 基于目标出生强度在线估计的多目标跟踪算法. 自动化学报, 2011, **37**(8): 963–972)
- 24 Yang Tao, Li Jing, Pan Quan, Zhang Yan-Ning. A greedy searching algorithm for multiple object tracking and occlusion handling. *Acta Automatica Sinica*, 2010, **36**(3): 375–384
(杨涛, 李静, 潘泉, 张艳宁. 一种基于贪心搜索的实时多目标遮挡处理算法. 自动化学报, 2010, **36**(3): 375–384)



沈乐君 四川大学博士研究生, 成都体育学院副教授. 主要研究方向为图像处理, 模式识别, 计算机视觉在运动训练中的应用. 本文通信作者.
E-mail: lejunshen@gmail.com
(**SHEN Le-Jun** Ph.D. candidate at Sichuan University and associate

professor at Chengdu Sport University. His research interest covers image processing, computer vision, and pattern recognition and their applications in sport science. Corresponding author of this paper.)



游志胜 四川大学计算机学院教授. 1981 年获得四川大学无线电系硕士学位, 1981 至 1983 年在美国密执安州立大学做访问学者. 主要研究方向为图像处理, 模式识别, 视觉合成, 实时软件工程.
E-mail: youzhisheng@scu.edu.cn

(**YOU Zhi-Sheng** Professor at the College of Computer Science, Sichuan University. He received his master degree from Sichuan University in 1981. He was a visiting scholar at Michigan State University, USA from 1981 to 1983. His research interest covers image processing, pattern recognition, synthetic vision, and real-time software engineering.)



李晓峰 四川大学计算机学院副教授. 1999 年获得清华大学精密仪器与机械学系硕士学位, 2005 年获得美国普度大学工程技术学院硕士学位. 主要研究方向为图像处理, 模式识别, 视觉合成.
E-mail: lixiaofeng2009@scu.edu.cn

(**LI Xiao-Feng** Associate professor at the College of Computer Science, Sichuan University. He received his master degrees from Tsinghua University in 1999, and School of Engineering and Technology (Indianapolis) of Purdue University, USA in 2005. His research interest covers image processing, pattern recognition, and synthetic vision.)