

基于置信区域内多级动态层表达的类型序蒙特卡洛视觉跟踪方法

宋涛, 李鸥, 刘广怡

(信息工程大学信息工程学院, 河南郑州 450002)

摘要: 视觉跟踪是智能监控、机器人和视觉导航等领域的核心技术. 针对现有类型序蒙特卡洛跟踪方法复杂度高、实时性差的问题, 本文提出了一种融合置信区域内多级动态层表达的跟踪方法, 采用更加可靠、有效的粒子模拟状态后验概率. 该方法利用检测模块得到目标可能出现的置信区域, 根据真实目标尺寸给出一种粒子采样策略, 每个粒子代表一级动态层表达, 并为每个粒子建立双层运动模型; 构建 Mean-Shift 分块观测模型以引入空间和外观信息, 同时根据子块的匹配程度计算粒子权值、检测目标遮挡状态并提出模型更新策略. 在公开视频序列上同经典粒子滤波和 Mean-Shift 等算法的实验对比结果证明了本文算法具有较优的跟踪准确度和实时性.

关键词: 视觉跟踪; 置信区域; 双层多级运动模型; 分块观测模型; 模型更新

中图分类号: TP319 **文献标识码:** A **文章编号:** 0372-2112 (2016)06-1355-07

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2016.06.014

Quasi-Sequential Monte Carlo Visual Tracking Based on Multilevel Dynamic Layer Representations in Confidence Region

SONG Tao, LI Ou, LIU Guang-yi

(Institute of Information System Engineering, Information Engineering University, Zhengzhou, Henan 450002, China)

Abstract: Visual tracking is a core technology for the application domains of intelligent monitoring, robotics and vision navigation, etc. Aiming at the problem of high complexity and poor real-time performance in the existing quasi-sequential Monte Carlo tracking algorithms, this paper presents a method based on multilevel dynamic layer representations, which simulates the posteriori probability of a state using more reliable and effective particles. Then a sampling strategy is proposed in confidence areas derived from the detector, in which each particle represents a dynamic representation and has a two-layer motion model. The observation model based on parted-mean-shift is constructed for space and appearance information. Depending on the degree of matching sub-blocks, the weight of particles is calculated and a way to detect the occlusion state of an object is put forward for realtime model update. Experimental results using challenging public video sequences show better accuracy and efficiency of the proposed method, compared with the classical particle filter and mean-shift algorithms, etc.

Key words: visual tracking; confidence region; two-layer multilevel motion model; block-based observation model; model update

1 引言

目标跟踪是计算机视觉领域的关键底层技术, 是一项正在不断创新、发展和进步的热点研究方向, 越来越多的算法被开发以满足不同的应用需求 (比如, 智能机器人、自动监控系统和人机交互等). 但是, 在复杂环境下构建性能优越且兼顾实时性的视觉跟踪算法仍是

一个极具挑战性的课题. 当前主流跟踪方法大致可以分为两类: 基于数据驱动的确定性方法和基于模型驱动的概率性方法. 概率性跟踪方法将视觉跟踪转化为求解目标状态最大后验概率密度估计的问题, 通过观测模型和动态模型将上述两类跟踪方法融入跟踪框架, 具有良好的跟踪性能和扩展性, 其最具代表性的方法是基于粒子滤波的视觉跟踪方法.

粒子滤波算法也被称为贯序蒙特卡洛方法 (SMC, Sequential Monte Carlo), 能够处理非线性、非高斯和多模态问题, 被广泛应用于视觉跟踪领域. 近些年来, 针对粒子滤波的改进方法主要包括三个方面: (1) 增加粒子数目, 尽可能覆盖目标可能出现的位置; (2) 设计有效的分配方案使粒子更加接近目标后验概率的模式点; (3) 建立更加鲁棒的观测模型, 提高粒子权值的可信度. 文献[1]利用交互多模结构 (IMM, Interacting Multiple Model) 为动态模型引入多种运动模式, 可以有效处理机动目标, 但计算量较大. 为了提高动态模型的预测精度同时减少跟踪所需的粒子数, 文献[2]在粒子滤波的框架中提出双层运动模型来寻找更加可靠的预测点, 但无法解决目标不规则运动问题. 文献[3,4]结合 Mean-Shift^[5,6] 算法构造跟踪方案, 将其应用于每个粒子, 通过迭代把粒子移动到局部极值点, 然后利用最小均方误差得到最终的目标状态. 文献[7]提出了基于马尔科夫链蒙特卡洛粒子滤波方法, 通过增加粒子数来提高跟踪精度, 不适合实际应用. 文献[8]建立三维观测模型, 使粒子滤波的更新阶段更加可靠, 在复杂跟踪环境中有很好的跟踪效果, 但是三维模型的计算量大, 增加了算法的复杂度.

本文根据上述(2)、(3)的基本思想, 提出了一种融合多级动态层表达的类贯序蒙特卡洛跟踪方法: 在无需任何先验知识和建议分布的前提下, 通过建立双层多级动态模型预测目标行为, 为 Mean-Shift 提供准确的初始搜索位置; 利用检测模块得到目标可能出现的区域, 根据真实目标尺寸给出一种粒子采样策略, 每个粒子独立分布并具有唯一性, 代表一级动态层表达, 即用最少的粒子覆盖目标可能出现的范围; 利用分块思想提出一种融合像素点空间信息的 Mean-Shift 迭代准则和粒子权值计算方法, 提高观测模型的可靠性.

2 构建双层多级运动模型和基于置信区域的采样策略

2.1 双层多级运动模型

感兴趣目标的运动模式主要有两个关键要素: 速度和方向, 大致可以分为两种类型: (1) 目标匀速运动或阶段性速度变化较为缓慢, 且其动态方程可以近似表达为过去 K 个已修正状态的线性组合; (2) 目标不规则运动, 其加速度、方向等易发生突然变化. 本文根据上述两种运动类型建立双层运动模型.

第一层运动模型, 引入目标运动的平滑机制和限制条件: 假设目标速度和方位没有发生突变, 且当前帧的状态满足前 K 帧运动规律. 其本质是作为一个“稳定的粒子”来规范第二层动态模型的预测结果, 防止目标在全遮挡、相似性背景或其他目标干扰时发生漂移. 假

设运动目标在半秒钟内不会发生大的突变, 本文按照 25 frame/s 的帧率取值 $K = 25/2$.

假设, $\phi_{t-k:t-1} = \{\phi_i\}_{i=t-k}^{t-1}$ 表示目标在前 K 帧中经过校正的位置, 按照帧间距离越近相关性越大的原则, 定义 $\pi_{j=t-1}^{(i=t-k, \dots, t-2)} = \exp\left[-\frac{(i-j)^2}{2\sigma^2}\right]$ 为第 j 帧与其他帧之间的关联程度, 第一层动态方程如式(1)所示:

$$\begin{cases} \hat{\phi}_{t-1} = \left[\frac{\sum_{i=t-k}^{t-2} \|\phi_{i+1} - \phi_i\|^2 \pi_{i-1}^{(i+1)}}{\sum_{i=t-k}^{t-2} \pi_{i-1}^{(i+1)}} \right]^{\frac{1}{2}} \\ \hat{x}_t = \hat{v}_{t-1} \cdot t + \varepsilon_t \end{cases} \quad (1)$$

式中, \hat{x}_t 和 \hat{v}_{t-1} 分别表示目标在第 t 帧的预测位置和速度; ε_t 是随机扰动, 由真实值 ϕ_{t-1} 和估计值 \hat{x}_{t-1} 之间的相似度决定.

第二层运动模型也称为多级动态层表达, 主要处理跟踪过程中目标运动的不确定性, 如发生突然加速、减速以及方向突变等不规则运动. 本文通过 2.2 节提出的置信区域采样构建第二层运动模型.

2.2 置信区域及采样策略

本文定义置信区域为目标在下一帧中可能出现的区域, 根据文献[9]提出的三帧差分检测算法来构建目标的活动区域. 基于置信区域的采样策略: 根据图像分割方法得到的真实目标尺寸 (详见 3.1 节) 和置信区域, 利用算法 1 采样粒子, 得到粒子的个数 Num_t 和位置坐标 $\{\hat{x}_t^j\}_{j=1}^{\text{Num}_t}$.

算法 1 基于置信区域的粒子采样方法

- 步骤 1 利用检测模块得到目标可能出现的置信区域 $\{\text{Area}_D^i\}_{i=1, \dots, n}$.
- 步骤 2 根据最近邻域法来关联目标与其对应的置信区域.
其中, $\text{Area}_D^i = \text{argmin}_i \|\hat{x}_t - C_D^i\|$.
- 步骤 3 根据置信区域的大小采取不同的采样方案. 如果 $\text{Area}_D^i \leq \text{Area}_T$, 真实目标可以一次覆盖置信区域, 所以置信区域内仅有一个采样点 C_D^i .
此时, $\text{Num}_t = 2, \{\hat{x}_t^j\}_{j=1}^{\text{Num}_t} = \{\hat{x}_t, C_D^i\}$.
- 步骤 4 如果 $\text{Area}_D^i > \text{Area}_T$, 用真实目标的大小不重复地遍历置信区域, 则粒子数如下式所示. 式中的两个 1 分别代表 \hat{x}_t 和 C_D^i ; N_i 是其余采样点.

$$\text{Num}_t = \text{cvFloor}\left(\frac{2W_D^i}{W_T}\right) + \text{cvFloor}\left(\frac{2H_D^i}{H_T}\right) + 1 + 1 = N_i + 2$$

其中, $\{\text{Area}_D^i, C_D^i, W_D^i, H_D^i\}_{i=1, \dots, n}$ 分别表示第 t 帧检测得到的 n 个运动区域、区域几何中心、及其框选窗口的宽和高; $\{\text{Area}_T, W_T, H_T\}$ 表示真实目标的尺寸 (即真实目标区域内像素点个数之和)、及其框选真实目标窗口的宽和高. 算法 1 得到的粒子均具有唯一性 (根据真

实目标尺寸和置信区域范围,用最少的粒子来覆盖目标可能出现的活动范围,且每个采样粒子所处的位置不会重复出现),每个粒子都代表目标在下一帧中可能出现的某种运动类型.结合蒙特卡洛估计框架,算法 1 也是本文第 2.1 节所提双层多级运动模型的第二层模型,每个粒子代表一级动态表达.

3 基于分块的 Mean-Shift 和粒子权值计算方法

3.1 构建真实目标区域的 KGHM 模型

在实际应用中,跟踪窗口常会引入大量的背景信息,影响目标模型的精确表达.文献[10]采用局部背景加权的直方图模型去除窗口内的背景像素,该算法对

复杂背景变化有良好的跟踪性能,但额外背景信息的引入增加了算法复杂度.本文利用图像分割技术得到精确目标区域来建立核高斯混合模型(KGHM, Kernel Gaussian Hybrid Model)^[11],在无需引入额外背景信息的条件下,去除局部背景像素,分割效果如图 1 所示.

在初始帧检测感兴趣目标,如图 1(a)所示,白色窗口为所要跟踪的目标;利用 Canny 算子检测目标的边缘信息,然后对目标边缘进行角点检测以近似确定目标真实轮廓,并根据目标的轮廓信息分割目标,如图 1(c)所示;通过对比目标分割前、后的直方图可以看出,分割之后的目标区域很大程度上消除了局部背景干扰,其对应直方图也更加紧凑,如图 1(b)和(d)所示.

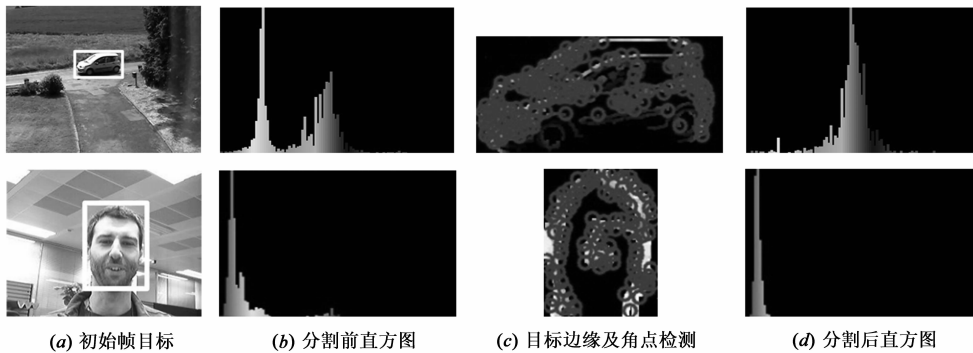


图1 真实目标分割示意图(第一行为刚性目标,第二行为非刚性目标)

本文采用无参核密度估计(KDE, Kernel Density Estimation)建立真实目标的概率密度模型:

$$p(f|o) = \frac{1}{n} \sum_{i=1}^{C_n} k_{hc}(c - c_i) \left[\sum_{j=1}^{S_n} k_{hs} \left(\left\| \frac{s - s_j}{h_s} \right\|^2 \right) \right] \quad (2)$$

其中, $\{f_i\}_{i=1, \dots, n}$ 表示真实目标区域 o 内的 n 个像素点; c_i 表示像素点 i 的色度值; s_j 表示像素点 j 的位置坐标; C_n 是目标区域内不同色度值的个数; S_n 是目标区域内色度值为 i 的像素点个数. $k(\cdot)$ 是核函数. 式(2)遍历区域所有像素,计算量较大. 本文采用分块统计的思想将目标区域分为若干子块,利用空间核估计遍历各子块中心点来近似全局像素点估计,式(2)转化为:

$$p(f|o) \approx \frac{1}{n} \sum_{i=1}^{C_n} k_{hc}(c - c_i) \left[\sum_{j=1}^{B_n} N_{ij} k_{hs} \left(\left\| \frac{s - l_j}{h_s} \right\|^2 \right) \right] \quad (3)$$

其中, B_n 表示目标区域内子块的个数; l_j 是子块 j 的中心点坐标; N_{ij} 表示子块 j 中色度值为 i 的像素点个数. 本

文提出以分块为主线来指导目标建模,式(3)转化为:

$$p(f|o) \approx \sum_{j=1}^{B_n} \frac{N_j}{n} k_{hs} \left(\left\| \frac{s - \gamma_j}{h_s} \right\|^2 \right) \frac{1}{N_j} \sum_{z=1}^{N_j} k_{hc}(c - c_z) \quad (4)$$

其中, N_j 表示子块 j 中像素点的个数, c_z 表示子块 j 中像素点 z 的色度值. 式(4)更加直观地体现了由目标区域过渡到目标子块,再过渡到子块内像素点的建模思想,并利用核函数融合了各子块之间的空间信息.

3.2 建立基于分块的 Mean-Shift 向量

在实际跟踪应用场景中,多个完全一模一样目标出现的概率极低.大多数情况下,跟踪过程中出现所谓的相似或伪目标是指全局特征相似,而局部特征各有不同.如图 2 所示的六个目标,它们拥有一样的直方图特征,但其局部特征却各不相同;利用分块的匹配策略完全可以区分彼此.

本文采用最基本的分块方案,按照几何区域将目标均匀分为 m 乘 n 个互不重叠的子块.以色度 u_1 为

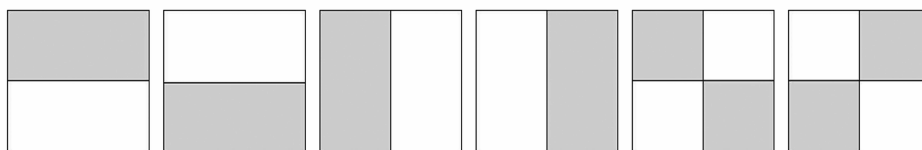


图2 同一直方图特征的不同空间信息

例,对目标的整体区域和每个子块进行核函数加权的直方图建模:

$$\begin{cases} q_{ul} = C_h \cdot \sum_{i=1}^N k\left(\left\|\frac{x_0 - x_i}{h}\right\|^2\right) \cdot \delta[b(x_i) - u1] \\ \quad \cdot p(f_i | o) \\ q_{ul}^{(b)} = C_{h_b} \cdot \sum_{i=1}^{N_b} k\left(\left\|\frac{x_0^{(b)} - x_i^{(b)}}{h_b}\right\|^2\right) \\ \quad \cdot \delta[b(x_i^{(b)}) - u1] \cdot p(f_i^{(b)} | o) \end{cases} \quad (5)$$

其中, $p(f_i | o)$ 表示像素点 i 为真实目标的概率,降低了背景干扰; q_{ul} 表示目标的整体特性;而建立 $q_{ul}^{(b)}$ 的目的:(1)引入空间信息;(2)突出局部特征的独立性;(3)提高模型对遮挡的敏感性和推理遮挡出现的可靠性.同理,可根据式(5)的方法计算候选模板 $p_u(y)$ 和 $p_u^{(b)}(y)$.

本文选择 Bhattacharyya 系数来度量目标模板 q_u 、 $q_u^{(b)}$ 和候选模板 $p_u(y)$ 、 $p_u^{(b)}(y)$ 之间的相似程度:

$$\begin{cases} \rho_{MS}(y) = \rho[p_u(y), q_u] = \sum_{u=1}^m \sqrt{p_u(y) \cdot q_u} = W_C \\ \rho_{ms}(y) = v^{(1)} \rho^{(1)} + \dots + v^{(b)} \rho^{(b)} + \dots + v^{(B)} \rho^{(B)} \\ = \sum_{b=1}^B v^{(b)} \rho^{(b)} = W_L \end{cases} \quad (6)$$

其中, $\rho_{MS}(y)$ 为全局巴氏系数,也是粒子的全局权值; $\rho_{ms}(y)$ 为各子块加权融合巴氏系数,即粒子的局部权值; $b=1, \dots, B$ 表示子块的数目, $v^{(b)}$ 决定了每个子块的贡献度/重要程度(距离中心点越近,可靠性越高).

根据经典 Mean-Shift 算法的理论框架计算得到基于目标区域分块的 Mean-Shift 向量,如式(7)所示:

$$\begin{aligned} MS &= MS(y) - y = \hat{y} - y \\ &= \frac{\sum_{b=1}^B \sum_{l=1}^L \sum_{i=1}^{N_i} w_{l,i}^{(b)} \cdot (x_i^{(b)} - z^{(b)}) \cdot \prod_{j \neq l} \rho_{uj}^{(b)}(y_0)}{\sum_{b=1}^B \sum_{l=1}^L \sum_{i=1}^{N_i} w_{l,i}^{(b)} \cdot \prod_{j \neq l} \rho_{uj}^{(b)}(y_0)} - y \end{aligned} \quad (7)$$

其中, $l=1, \dots, L$ 表示像素点的颜色分量;定义:

$$\begin{aligned} w_{l,i}^{(b)} &= \lambda^{(b)} \cdot g\left(\left\|\frac{x_i^{(b)} - y - z^{(b)}}{h^{(b)}}\right\|^2\right) \cdot \sum_{u=1}^{ml} \sqrt{\frac{\hat{q}_{ul}^{(b)}}{p_{ul}^{(b)}(y_0)}} \\ &\quad \cdot \delta[\beta(x_i^{(b)}) - ul]; \end{aligned}$$

$z^{(b)}$ 表示子块 b 的中心点和整个目标中心点之间的矢量;函数 $\beta(x_i^{(b)})$ 决定像素点的色度值.根据 Mean-Shift 向量可以得到新的迭代准则.

4 遮挡处理和模型更新

4.1 遮挡检测

遮挡推理是决定目标是否处于遮挡状态的辅助机制,指导系统及时调整外观模型和跟踪策略.文献[12]

构建自遮挡概率图来判断非刚性目标的遮挡情况,但该方法比较复杂、计算量大,不适用于实时跟踪.本文利用分块模型和目标当前帧状态提出一种简单的推理策略:(1)边缘像素最先受到遮挡的影响,当边缘子块的度量函数开始变小而核心子块正常时,说明遮挡开始出现;(2)多目标跟踪的情况下,通过目标的状态和相似性度量函数来判断遮挡目标和被遮挡目标;(3)当大部分/所有子块的度量函数都变小时,说明全遮挡出现,仅通过动态函数来继续跟踪目标.

4.2 遮挡处理及更新机制

外观模型更新的基本准则:(1)当出现半遮挡/部分遮挡时,只更新没有受到遮挡的子块;(2)当出现全遮挡时,终止外观模型的匹配计算,利用检测模块或动态模型来寻找目标,直到目标重新出现,如果全遮挡状态超过时间阈值 T_{time} ,则判断目标消失;(3)如果 n 帧之内没有遮挡情况发生,则对目标的整个区域进行更新 $q_u^t = (1-a)p_u^t + q_u^{t-1}$ 并重新计算 $p(f_i | o)$.

5 本文跟踪方法

本文以贯序蒙特卡洛方法为框架,将双层多级动态模型、Mean-Shift 分块观测模型、遮挡处理机制和目标模型更新方法相结合建立视觉跟踪方法,如算法 2 所示.

算法 2 置信区域内多级动态层表达的类型序蒙特卡洛跟踪方法

- 步骤 1 初始化:手动选择或利用检测模块得到感兴趣目标,按照 3.2 节的方法划分目标区域,并分别根据式(4)和(5)计算目标的 KGHM 模型、全局模型 q_{ul} 和局部模型 $q_{ul}^{(b)}$.
- 步骤 2 在下一帧 t 时刻,根据算法 1 得到采样粒子 $\{\tilde{x}_i\}_{i=1}^{Num_t}$.
- 步骤 3 根据式(5)计算粒子的候选全局模型 p_{ul} 和局部模型 $p_{ul}^{(b)}$;根据式(6)计算粒子的局部权值 W_L^i 和子块系数 $\rho^{(b)}$.
- 步骤 4 根据 4.1 节方法判断目标的遮挡状态.如果没有遮挡或出现部分遮挡,直接利用式(7)进行 Mean-Shift 迭代,得到新的位置 \tilde{x}_i^t 和权值 W_L^i ;如果出现全遮挡,仅利用第一层运动模型预测目标可能出现的位置,直到目标出现再进行 Mean-Shift 迭代.
- 步骤 5 计算目标的最终状态: $\phi_t = \sum_{i=1}^{Num_t} \tilde{x}_i^t \cdot W_L^i$.
- 步骤 6 根据 4.2 节的更新机制调整目标模型.并在下一帧循环步骤 2.

6 实验结果与分析

为了验证所提算法的有效性,本文针对 Mean-Shift^[6]、粒子滤波^[13]和基于核贝叶斯框架的文献[14]进行了多组对比实验,包括阴影干扰、目标复杂运动和全遮挡三个场景.本文选择 PETS 提供的公共标准视频序列进行测试,且所有实验都是在 Intel 双核 2.2GHz CPU、2.0GB 内存的平台上,利用 C/C++ 和 OpenCV 函数库实现的.

本文通过质化评价(跟踪效果图)和量化评价(跟

跟踪误差 d_t 、目标覆盖率 r_t 两个方面来论述算法的有效性;根据耗时(实时性)来评价算法的运行效率.

$$\begin{cases} d_t = \sqrt{(x_{\text{True}}^t - x_{\text{Tracking}}^t)^2 + (y_{\text{True}}^t - y_{\text{Tracking}}^t)^2} \\ r_t = \frac{\text{Area}(R_{\text{True}}^t \cap R_{\text{Tracking}}^t)}{\text{Area}(R_{\text{True}}^t \cup R_{\text{Tracking}}^t)} \end{cases} \quad (8)$$

其中, $(x_{\text{True}}^t, y_{\text{True}}^t)$ 和 $(x_{\text{Tracking}}^t, y_{\text{Tracking}}^t)$ 分别表示 t 时刻目标真实位置和算法得到的估计位置; R_{True}^t 和 R_{Tracking}^t 分别表示 t 时刻目标真实区域和跟踪算法得到的区域.

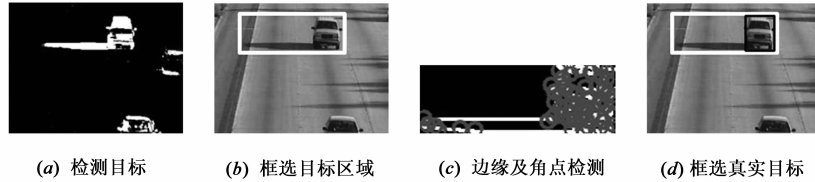


图3 阴影干扰处理示意图

图 4 显示了本文算法在目标复杂运动场景中与文献[6]、[13]和[14]的跟踪效果对比图,其跟踪窗口分别由白色实线矩形框、黑色实线矩形框、白色虚线矩形框和黑色虚线矩形框表示. 场景解析:感兴趣目标为图像上方的轿车,目标在第 850 帧向西驶入丁字路口,并开始发生转向;从第 878 帧至 932 帧,目标先后经历了树木的局部遮挡和草坪的大面积局部遮挡. 从图 4 可以看出,文献[13]无法处理目标复杂的运动形式,其状态预测阶段的误差逐渐积累,从第 878 帧开始丢失目标;跟踪过程中,由于目标发生了快速转向运动,其帧间移动范围超过了核函数的半径,导致文献[6]在第

图 3 显示了本文算法去除阴影干扰的有效性. 首先,根据目标检测模块捕捉感兴趣目标,如图 3(a) 和 (b) 所示;框选目标可能出现的区域,从图 3(b) 中可以看出,白色矩形框中存在大量的阴影;利用 3.1 节目标分割方法去除阴影并近似确定目标真实轮廓,从图 3(c) 中可以看出,通过角点的聚集分类可以快速去除孤立角点(阴影部分)并确定真实目标范围,如图 3(d) 中的黑色矩形框.

878 帧搜索不到感兴趣目标,跟踪失败;文献[14]利用 Mean-Shift 迭代得到目标的预测位置,通过增加粒子数弥补预测误差,虽然 Mean-Shift 算法在第 878 帧没有提供有效预测,但通过增加粒子数仍能找到目标,而且文献[14]的模型更新策略可处理半遮挡问题. 本文算法根据置信区域采样所得到的粒子可以有效捕捉到机动目标,然后利用分块 Mean-Shift 算法寻找目标的精确位置,同时通过遮挡检测和模型更新策略来解决遮挡问题,如图 4 所示,本文算法在整个跟踪过程中都可以正确跟踪感兴趣目标.



图4 本文算法和文献[6]、[13]、[14]的跟踪结果对比图

表 1 本文算法与文献[6]、[13]、[14]的跟踪效率

跟踪算法	粒子数	耗时 (ms)
本文算法	5	28.7
文献[6]	0	22.9
文献[13]	100	61.8
文献[14]	18	39.3

图 5 量化反映了本文算法和文献[6]、[13]、[14]在目标复杂场景下的跟踪性能;表 1 根据各算法所需的粒子数和运行耗时来评价算法执行效率. 一般情况下,视频帧率为 25 帧/s,显示一帧图像的时间为 40ms,因此,算法的运行耗时小于 40ms 就可满足实时跟踪. 文献[6]和[13]在跟踪过程中丢失目标,失去了量化分析的意义;文献[14]通过多线索融合使跟踪窗

口逐渐接近真实目标,如图 5 所示:该方法中心点误差和覆盖率均优于其他方法,但当目标发生不规则运动和遮挡时,其需要增加粒子来扩散目标搜索范围,导致算法的执行效率下降(如表 1 所示);综合图 4、5 和表 1 可以看出,本文算法不但可以有效跟踪目标且具有较好的实时性.

图 6 的第一行序列是本文算法、文献[6]和文献[14]在全遮挡场景下的跟踪效果图,其跟踪窗口分别由白色实线框、黑色实线框和黑色虚线框表示;第二行序列是本文算法对应的检测结果. 场景解析:目标为地铁站内由南向北行走的行人,目标从第 3469 帧开始向东转向出口处,行走过程中目标将遇到立柱的完全遮挡. 由图 4 可知,传统粒子滤波方法不能处理目标的复杂运动,因此本文算法只与文献[6]和[14]进行对比.

从图 6 第一行的第 3511 帧和第 3522 帧可以看出,文献 [6]和[14]的特征匹配策略在目标消失的情况下完全失效,导致目标丢失并发生漂移.而本文算法根据第一层动态模型提供的运动信息和目标检测信息依然可以在目标重新出现之后找到并继续跟踪目标,如图 6 第一行的第 3522 帧和第 3534 帧所示.如图 6 第二行的第 3431 帧所示,检测模块捕捉到两个相互叠加的运动目标(白色虚线框),本文算法通过 Mean-Shift 迭代找到感兴趣目标(白色实线框);当目标于第 3511 帧丢失时,在目标可能出现的附近检测不到任何运动信息;当目标在第 3522 帧和第 3534 帧出现时,检测模块能重新捕捉到目标,并通过 Mean-Shift 迭代确定其最终位置.

表 2 本文算法与文献[6]、[14]的跟踪效率

跟踪算法	粒子数	耗时(ms)
本文算法	4	26.3
文献[6]	0	20.5
文献[14]	13	35.8

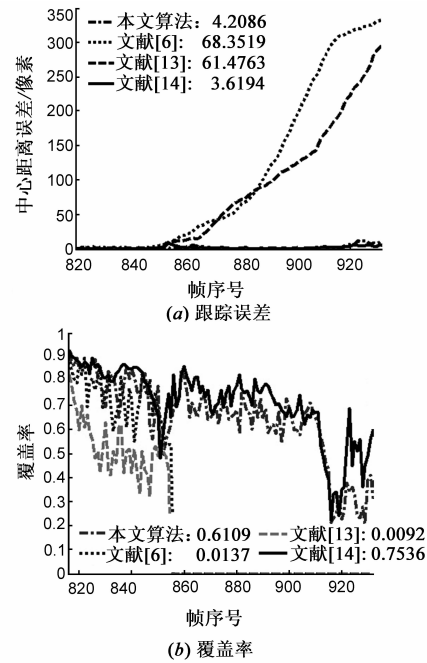


图 5 本文算法与文献[6]、[13]、[14]的定量分析图



图 6 本文算法和文献 [6] 、 [14] 的跟踪结果对比图

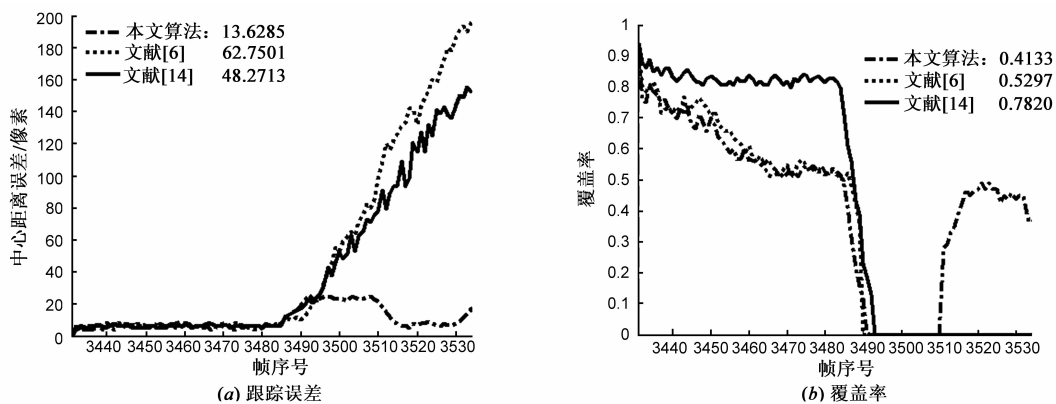


图 7 本文算法与文献[6]、[14]的定量分析图

图 7 和表 2 量化反映了本文算法和文献[6]、[14]在全遮挡场景下的跟踪性能和执行效率.从图 7 可以看出,当全遮挡出现时,文献[6]、[14]的跟踪性能迅速

恶化,而本文算法始终保持较好的跟踪性能.如表 2 所示,三种算法均具有良好的执行效率(运行耗时均小于 40ms),都可以满足实时性要求较高的应用场景.

7 结论

本文提出一种融合置信区域内多级动态层表达的类贝叶斯蒙特卡洛视觉跟踪方法,通过置信区域采样和第一层动态方程得到有效粒子在下一帧不同的预测位置;构建基于分块模型的 Mean-Shift 迭代算法,计算粒子修正位置和权值;并依据子块的似然程度提出遮挡检测和模型更新策略.实验结果表明,该方法可以有效处理阴影干扰、目标复杂运动及全遮挡等问题,极大减少了跟踪过程所需要的粒子数量和 Mean-Shift 迭代次数.同时,该方法具有很好的扩展性:在粒子的权值计算阶段,可以根据应用场景的不同需求来平衡跟踪精度与实时性之间的权重比例,构建适当的外观模型,如:颜色、纹理、梯度或融合各类特征的混合模型等.

参考文献

- [1] Xue J, Zheng N, Geng J. Tracking multiple visual targets via particle-based belief propagation[J]. IEEE Transactions on SMC, Part B, 2008, 38(1): 196 – 209.
- [2] Matej K, Stanislav K. A two-stage dynamic model for visual tracking[J]. IEEE Transactions on SMC, Part B, 2010, 40(6): 1505 – 1520.
- [3] Han H, Ding Y S, Hao K R. An evolutionary particle filter with immune genetic algorithm for intelligent video target tracking[J]. Computer & Mathematics with Applications, 2011, 62(7): 2685 – 2695.
- [4] Chang C, Ansari R. Kernel particle filter for visual tracking[J]. IEEE Signal Processing Letters, 2005, 12: 242 – 245.
- [5] Comaniciu D, Meer P. Mean shift: A robust approach toward feature space analysis[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(5): 603 – 619.
- [6] Comaniciu D, Ramesh V, Meer P. Kernel-based object tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(5): 564 – 577.
- [7] Cong D T, Septier F, Garnier C, et al. Robust visual tracking via MCMC-based particle filter[A]. IEEE International Conference on Acoustics, Speech and Signal Processing [C]. Kyoto: IEEE Press, 2012. 1493 – 1496.
- [8] Brown J A, Capson D W, Member S. A framework for 3D model-based visual tracking using a GPU-accelerated particle filter[J]. IEEE Transactions on Visualization and Computer Graphics, 2012, 18(1): 68 – 80.
- [9] Anderson C, Bert P, Vander W. Change detection and tracking using pyramids transformation techniques[A]. Proceedings of SPIE Conference on Intelligent Robots and Computer Vision [C]. Cambridge MA: Elsevier Press, 1985. 72 – 78.
- [10] Wang L F, Yan H P, Wu H Y. Forward-backward mean-shift for visual tracking with local-background-weighted histogram[J]. IEEE Transactions on Intelligent Transportation Systems, 2013, 14(3): 1480 – 1489.
- [11] Zhou L, Huang K, Tan T. Foreground object detection using top-down information based on EM framework[J]. IEEE Transactions on Image Processing, 2012, 21(9): 4204 – 4217.
- [12] Gay-Bellile V, Bartoli A, Sayd P. Direct estimation of non-rigid registrations with image-based self-occlusion reasoning[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(1): 87 – 104.
- [13] Isard M, Blake A. Condensation-conditional density propagation for visual tracking[J]. International Journal of Computer Vision, 1998, 29(1): 5 – 28.
- [14] Zhang X, Hu W, Bao H. Robust head tracking based on multiple cues fusion in the kernel-Bayesian framework[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2013, 23(7): 1197 – 1208.

作者简介



宋 涛 男, 1983 年 9 月出生于河南省焦作市. 现为解放军信息工程大学信息工程学院博士研究生. 主要研究方向为计算机视觉和目标跟踪技术.
E-mail: taosong_1983@126.com



李 鸥 男, 1961 年出生于河南省郑州市. 现为解放军信息工程大学信息工程学院教授、博士生导师. 主要研究方向为无线网络、信息融合和目标跟踪技术.
E-mail: zzliou@126.com