

自适应分块多线索融合粒子滤波跟踪方法

孙晓燕^{1,2}, 常发亮¹

(1. 山东大学 控制科学与工程学院, 济南 250061; 2. 山东建筑大学 计算机科学与技术学院, 济南 250101)

摘要: 多线索融合是解决复杂情况下跟踪问题的有效手段, 为此提出一种基于自适应分块目标模型的多线索融合粒子滤波跟踪方法. 根据目标颜色分布自适应分块建立目标描述模型, 可提高对目标初始描述的适应性; 采用多线索融合粒子滤波跟踪, 在跟踪过程中能根据子块可靠程度动态调整权重, 提高对剧烈光照变化、目标姿态变化、遮挡等复杂情况的适应性. 实验结果表明, 所提出的跟踪方法在多种复杂情况下能准确有效地跟踪目标.

关键词: 目标跟踪; 粒子滤波; 自适应分块; 多线索融合

中图分类号: TP391

文献标志码: A

Adaptive fragment multi-cue fusion particle filter tracking method

SUN Xiao-yan^{1,2}, CHANG Fa-liang¹

(1. School of Control Science and Engineering, Shandong University, Ji'nan 250061, China; 2. School of Computer Science and Technology, Shandong Jianzhu University, Ji'nan 250101, China. Correspondent: CHANG Fa-liang, E-mail: flchang@sdu.edu.cn)

Abstract: Multi-cue fusion is an effective means to solve the tracking problem under many complex conditions. Therefore, a multi-cue fusion particle filter tracking method is proposed, in which the target can be adaptive fragmented. The target model is built based on the color distribution of the target, which can improve the adaptability of the initial description. During the tracking process, the multi-cue fusion particle filter can dynamically adjust the weights of fragments according to the reliability of them, which can improve the flexibility of the tracker to deal with some complex circumstances, such as dramatic light change, target posture change and occluded by other objects. Experimental results show that the proposed method can accurately track the target effectively on many kinds of complex circumstances.

Key words: object tracking; particle filter; adaptive fragment; multi-cue fusion

0 引言

目标跟踪是计算机视觉领域重要的研究课题, 有着广泛的应用领域. 由于目标所处背景的复杂性、目标表现特性变化以及遮挡等情况的存在, 自然条件下使用单一线索通常难以鲁棒地跟踪目标^[1]. 融合多种线索, 使线索之间取长补短是解决这些复杂情况下跟踪的有效手段. 目前多线索的组成有两种方法: 一种是采用多种特征描述目标组成多线索, 国内外很多学者在这方面做了大量工作^[2-4]. 另一种是用同一特征描述目标的多个子块组成多线索^[5-9]. Elgammal等^[5]将人体垂直分为3部分: 头部, 躯干和腿部; Mittal等^[6]将目标垂直分为 h 块, 每一块由一个高斯核密度函数

建模; Adam等^[7]提出FragTrack方法, 将目标框进行垂直和水平方向固定分块, 提取子块颜色直方图组成多线索进行模板匹配跟踪. Erdem等^[8]采用FragTrack中的固定分块方法, 将原来各子块固定权重改为根据子块的可靠性动态调整权重.

采用固定个数子块组成多线索描述目标时, 若目标颜色分布较单一, 势必造成线索的冗余; 而若目标颜色分布较复杂, 又无法完全表达目标颜色的变异性. 为此, 本文提出一种根据目标颜色分布自适应分块, 将子块自适应调整权重, 组成多线索进行粒子滤波跟踪的方法. 自适应分块方法根据目标颜色分布确定子块个数, 提高了对目标初始描述的适应性; 粒子滤波

收稿日期: 2013-05-24; 修回日期: 2013-10-24.

基金项目: 国家自然科学基金项目(60975025, 61273277); 教育部留学回国人员科研启动基金项目(20101174); 山东省自然科学基金项目(ZR2011FM032); 济南市高校自主创新计划项目(201004002); 高等学校博士学科专项科研基金项目(20130131110038).

作者简介: 孙晓燕(1978-), 女, 讲师, 博士生, 从事计算机视觉、运动目标跟踪等研究; 常发亮(1965-), 男, 教授, 博士生导师, 从事模式识别、计算机视觉等研究.

跟踪根据各子块的可靠性动态调整各子块的权重, 提高了跟踪过程中对目标姿态变化、遮挡等情况的适应性.

1 自适应分块目标模型建立

固定数量分块无法适应目标的多样性, 而目标的颜色分布往往在局部小范围内具有较高的相似性, 故考虑从颜色分布上直接对目标矩形区域进行自适应分块, 无需提取目标掩膜.

设目标矩形区域为 $O(M, N)$, M, N 为目标的高度和宽度; $C(i, j)$ 为目标点 (i, j) 处的颜色值 (灰度图像为灰度值, 彩色图像为 HSV 空间的 H 分量). 提取目标框内所有像素颜色的水平投影 $C^H = \{C_j^h = \frac{1}{N} \sum_{i=1}^N C(i, j), j=1, 2, \dots, M\}$ 和垂直投影 $C^V = \{C_i^v = \frac{1}{M} \sum_{j=1}^M C(i, j), i=1, 2, \dots, N\}$. 选取阈值 T_C , 对 C^H 和 C^V 进行如下判断: 若 $|C_{j+1}^h - C_j^h| < T_C$, 则 C_{j+1}^h 和 C_j^h 划入同一子块; 若 $|C_{i+1}^v - C_i^v| < T_C$, 则 C_{i+1}^v 和 C_i^v 划入同一子块.

对所有的 $C_j^h (j=1, 2, \dots, M)$ 和 $C_i^v (i=1, 2, \dots, N)$ 进行如上扫描, 得到水平和垂直两个方向的分块 P^H 和 P^V , H 和 V 分别为两个方向的分块数. 为了增强目标表示能力, 参考文献 [7] 和 [8] 中的分块方法, 再将每个分块分为左右或上下两块, 组成最终的目标子块集 $P = \{P^{2H}, P^{2V}\}$. 每一子块均表示为 $P_j = \{dx, dy, w, h\}$, (dx, dy) 为当前子块距离中心的偏移量, w 和 h 为子块的宽度和高度. 图 1 为自适应分块方法的示意图. 根据自适应分块结果可建立目标模型 $P^T = \{p_j^T, w_{p_j}\}_{j=1, 2, \dots, 2H+2V}$, 其中: p_j^T 为第 j 个子块的直方图, w_{p_j} 为子块对应权值.

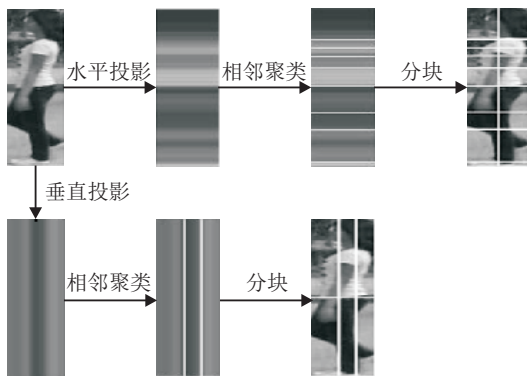


图 1 自适应分块方法示意图

2 基于多线索融合和空间分布的粒子滤波

粒子滤波是一种通过蒙特卡罗方法来实现递归贝叶斯滤波的算法, 其适用于非线性非高斯的特性非常适合于目标跟踪领域, 在近年来应用广泛 [1]. 采用自适应分块组成多线索之后, 基本粒子滤波中计算粒

子匹配度的方法要做相应的调整; 为了提高跟踪的适应性, 还要根据各子块的可靠程度和粒子的空间分布动态调整子块权重, 并进行子块的分裂和合并.

2.1 粒子滤波跟踪

设 x_k 表示目标在 k 时刻的状态向量, z_k 表示 k 时刻的观测值, 从贝叶斯估计角度看, 跟踪问题就是根据给定观测序列 $\{z_{1:k}\}$ 递归估计目标状态 $\{x_k\}$, 即构造状态的后验概率分布 $p(x_k|z_{1:k})$ 的过程, 而粒子滤波则用一组带有权重的随机样本 (粒子) 集 $\{x_t^{(i)}, \omega_t^{(i)}\}_{i=1}^N$ 来逼近此后验概率 [9].

用 $p(z_k|x_k^{(i)})$ 表示 k 时刻粒子 $x_k^{(i)}$ 的匹配度, 基本粒子滤波中权重更新过程为

$$\omega_k^{(i)} \propto \omega_{k-1}^{(i)} p(z_k|x_k^{(i)}), \quad (1)$$

后续时刻的状态估计为

$$\hat{x}_k = \sum_{i=1}^N \omega_k^{(i)} x_k^{(i)}. \quad (2)$$

本文的粒子滤波中, 选取状态向量 $x_k = \{c_x, c_y, s_x, s_y\}$, c_x, c_y 为矩形框中心坐标, s_x, s_y 分别为 x, y 方向的尺度缩放比例. 选取一阶自回归模型作为状态转移模型

$$x_k - x_{k-1} = x_{k-1} - x_{k-2} + W_{k-1}. \quad (3)$$

其中 W_{k-1} 是噪声矩阵, $W_{k-1} \sim N(0, G)$, $G = \text{diag}(\sigma_x^2, \sigma_y^2, \sigma_{s_x}^2, \sigma_{s_y}^2)$.

2.2 多线索融合的粒子匹配度

采用自适应分块方法之后, 每个粒子也要进行与模板相同的自适应分块, 粒子总的匹配度应由所有子块匹配度融合得到.

设 $x_k^{(i)} = \{x^{(i)}, y^{(i)}, s_x^{(i)}, s_y^{(i)}\}$ 表示 k 时刻的第 i 个粒子, 它是当前目标状态的一个假设. 结合初始模型分块方法可得到对应于该粒子的所有子块为

$$P_j^{(i)} = \{x^{(i)} + s_x^{(i)} dx, y^{(i)} + s_y^{(i)} dy, s_x^{(i)} w, s_y^{(i)} h\},$$

$$j = 1, 2, \dots, 2H + 2V.$$

通过积分直方图快速提取 $P_j^{(i)}$ 的直方图, 并计算与目标模型中对应子块 P_j^T 直方图之间的 Bhattacharyya 距离

$$d(H_1, H_2) = \left(1 - \sum_{i=1}^{\text{bin}} \sqrt{h_{i,1} h_{i,2}}\right)^{1/2}, \quad (4)$$

可得到粒子 $x_k^{(i)}$ 由当前子块 $P_j^{(i)}$ 决定的与模板的匹配度. 如何将所有子块的匹配度融合成当前粒子总的匹配度是一个关键问题, 采用如下融合方式:

$$p(z_k|x_k^{(i)}) = \exp\left(-\frac{\sum_{P_j^T \in P^T} w_{P_j} d(P_j^{(i)}, P_j^T)^2}{2\sigma^2}\right), \quad (5)$$

其中 w_{P_j} 为目标模型中每个子块的权值, 代表着子块的可靠性, 且 $\sum_{P_j^T \in P^T} w_{P_j} = 1$. 在跟踪过程中, 若某些

子块因为遮挡或光照等原因变得不可靠,可将其权值变小,降低遮挡或光照变化等复杂情况对整个跟踪结果的影响.子块权值的调整在下文 2.3 节中进行说明.

2.3 考虑粒子空间分布的子块权值更新

根据目标的最优状态估计 \hat{x}_k , 可得到其相应的子块 \hat{P}_j 及其与目标模型对应子块 P_j^T 的匹配度

$$p_{P_j} = \exp\left(-\frac{d(\hat{P}_j, P_j^T)^2}{2\sigma^2}\right). \quad (6)$$

每个子块的匹配度代表着这一子块与模板对应子块的相似程度,可直接据此调整各子块的权值 w_{P_j} . 但是,一旦经最优状态估计得到的 \hat{x}_k 存在误差,则据此得到的匹配度也会不准确,若直接根据此匹配度调整子块权值 w_{P_j} , 并继续进行粒子采样及贝叶斯递推,则势必造成误差的累积,导致跟踪的漂移.为了使得基于自适应分块的粒子滤波更加稳定,在计算子块适应度并调整权值时,不只考虑该子块与模板子块的相似性,还应考虑粒子的空间分布情况,即采用如下方法计算各子块的适应度^[10]:

$$q_{P_j} = \frac{\sum_{i=1,2,\dots,N} \frac{1}{N} |x_k^{(i)} - \hat{x}_k|}{\sum_{i=1,2,\dots,N} w_{P_j}^{(i)} |x_k^{(i)} - \hat{x}_k|} p_{P_j}(z|\hat{x}_k). \quad (7)$$

其中: $w_{P_j}^{(i)}$ 为对应于当前子块 \hat{P}_j 每个粒子的适应度, $|x_k^{(i)} - \hat{x}_k|$ 为当前粒子位置与最优状态估计目标位置的欧式距离.得到各个子块的适应度后,采用民主融合方式^[11]进行每个子块权值的调整:

$$w_k^b = (1 - \alpha)w_{k-1}^b + \alpha \frac{q_{P_j}}{\sum_{j=1}^{2H+2V} q_{P_j}}, \quad (8)$$

其中 α 为更新因子, α 越大更新速率越快.

2.4 子块的分裂与合并

跟踪过程中目标的表现特性或姿态会发生变化,若一直采用初始帧的分块方法,即使调整子块权值,仍可能无法适应该变化,因此后续帧考虑进行子块的分裂或合并.在得到目标的最优状态估计 \hat{x}_k 及相应分块,并进行了直方图的更新之后,考察垂直和水平两个方向上 P^H 和 P^V 相邻分块的灰度均值相似性和方差来确定是否进行子块的合并或分裂.

根据更新后子块 P_k^T , 计算 P^H 和 P^V 的灰度均值 Cm^n 和方差 σ^n ($n = H$ 或 V). 若 $|Cm_{j+1}^n - Cm_j^n| < T_C$ (T_C 与前相同), 则应将相邻两块 j 和 $j+1$ 合并为一块,子块集合 $P = \{P^{2H}, P^{2V}\}$ 中相应的 4 块合并为两块,相应的模板和权重直接合并;若当前分块的方差 $\sigma^n > T_\sigma$, 则将该分块分为两块,相应的 $P = \{P^{2H}, P^{2V}\}$ 中将对应分块的两个子块分裂为 4 块,相应的模板和权重作调整.在分裂时要保证子块最小宽度或高度不能小于 10 个像素.

2.5 算法步骤

基于自适应分块的多线索融合粒子滤波算法分为初始化、预测、度量、更新 4 个部分.

初始化: 初始帧划定目标,按照颜色分布自适应分块,建立目标模型 $P^T = \{p_j^T, w_{p_j}\}, j = 1, 2, \dots, 2H + 2V$, 初始子块权值 w_{p_j} 为等权值;将目标中心位置作为粒子初始位置,初始权重为等权重.

读取后续帧进行如下 Step 1 到 Step 3 的循环处理.

Step 1: 预测(包含粒子的重采样).

- 1) 粒子重采样,防止粒子退化;
- 2) 按式(3)进行粒子运动.

Step 2: 度量.

1) 根据粒子新位置,按式(4)和(5)计算每个粒子的匹配度,按式(1)进行粒子权值的更新,并按式(2)估计最优目标位置 \hat{x}_k ;

2) 根据 \hat{x}_k 得到对应子块,按式(6)和(7)计算每个子块的适应度.

Step 3: 更新.

- 1) 按式(8)更新各子块权值.
- 2) 根据 \hat{x}_k 和得到的分块情况,按下式进行每个子块直方图模型的更新:

$$P_k^T = (1 - \beta)P_{k-1}^T + \beta\hat{P}. \quad (9)$$

其中: \hat{P} 为根据 \hat{x}_k 得到的各子块直方图; β 为更新因子, β 越大更新越快.

3) 根据当前分块情况和更新后的模型 P_k^T , 按 2.4 节进行子块的分裂或合并.

3 实验结果及分析

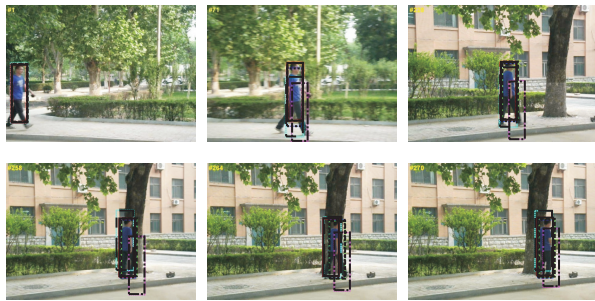
3.1 实验环境和参数说明

实验环境为 CPU Intel Core 2, 2.93 GHz, 内存为 3 GB, 软件采用 Matlab R2010a, 其中计算量较大的积分直方图部分采用 C 与 Matlab 的接口函数实现. 实验中所用部分参数如下: 粒子个数 $N = 500$; 直方图分块数 bins = 16; 子块权值更新率 $\alpha = 0.01$; 模板更新率 $\beta = 0.1$; 粒子匹配度标准差 $\sigma = 0.1$; 子块分裂时阈值 $T_\sigma = 60$; 状态噪声标准差 $\sigma_x = \sigma_y = 5, \sigma_{sx} = \sigma_{sy} = 7.5 \times 10^{-3}$. 分块阈值 T_C 决定着子块的数目,根据不同的视频需进行不同的设置.

3.2 与 AMFTrack、FragTrack 和 PF 的比较

实验中采用自己拍摄的行人视频序列,分别采用本文提出的自适应分块的多线索融合粒子滤波(ABMFTrack)、文献[7]提出的FragTrack、文献[8]提出的AMFTrack和不进行任何分块直接提取整体直方图作为目标模型的基本粒子滤波算法(PFTrack)进行比较.视频中背景较复杂,行人目标也经历了从正

侧面到背面的姿态变化. 实验时 ABMFTrack 分块阈值 $T_C = 20$, 得到的初始分块个数为 $H = 9, V = 6$. PFTTrack 采用直方图分块数为 3×10 的 HSV 空间直方图, 粒子个数等其他参数与 ABMFTrack 中一致, FragTrack 和 AMFTrack 均采用原作者提供的参数. 实验结果的定性比较如图 2 所示, 通过计算跟踪结果中心点坐标与人工标定真理图的欧式距离, 得到的误差曲线如图 3 所示.



— ABMFTrack - - - FragTrack GroundTruth - · - · - AMFTrack
- - - PFTTrack

图 2 blue 视频跟踪结果比较

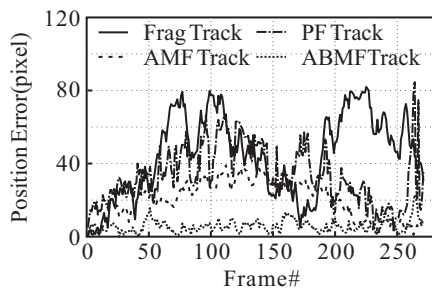


图 3 blue 视频 3 种跟踪方法误差曲线

可以看出, ABMFTrack 用更少的子块 $2 \times (9 + 6) = 30$ 块 (FragTrack 和 AMFTrack 固定采用 36 个子块), 得到比 FragTrack 和 AMFTrack 更稳定的跟踪结果.

3.3 遮挡情况下的跟踪结果及权重变化情况

实验选用文献 [7] 中提供的 face 视频对人体头部进行跟踪, 目标姿态变化较小, 但中间过程中出现多次遮挡甚至严重遮挡的情况. 实验时, ABMFTrack 分块阈值 $T_C = 29$, 得到的初始分块个数为 $H = 3, V = 4$, 且跟踪过程中未进行子块个数的调整. 实验结果如图 4 所示 (实线框为 ABMFTrack 跟踪结果, 虚线框为 AMFTrack 跟踪结果), 误差曲线如图 5 所示.



图 4 face 视频跟踪结果比较

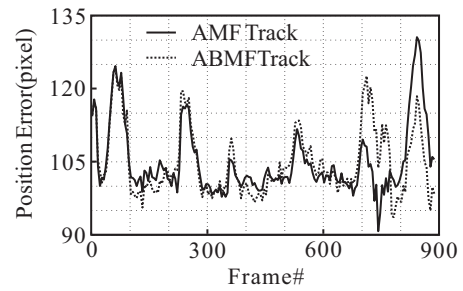


图 5 face 视频跟踪误差比较

从图 5 可以看出, ABMFTrack 使用较少数量的子块 ($2 \times (3 + 4) = 14$ 个), 就可达到文献 [8] 中 AMFTrack 基本相同的跟踪精度 (子块数 36), 且能够对遮挡进行较好的处理.

图 6 为初始分块时水平和垂直的分块情况, 从中选取水平分块的左上、右上、左下和右下, 子块分别编号为 1、2、5、6. 图 7 记录了整个跟踪过程中这 4 个子块的权重变化情况. 在目标被由左侧进入的遮挡物遮挡时 (图 4 中 125 帧所示), 子块 1 和子块 5 由于遮挡导致可靠性降低, 权重出现较大下降, 而未被遮挡的子块 2 和子块 6 权重则有较大提高; 而在右侧出现遮挡物对目标进行遮挡时 (图 4 中 194 帧左右), 子块权重则出现相反的变化过程. 图 4 中 310 帧和 855 帧出现下部进入的遮挡物时, 子块 1 和子块 2 的权重增大, 相应的子块 5 和子块 6 的权重降低. 实验结果说明自适应调整权重的多线索融合粒子滤波方法能够根据子块的可靠性进行权值的调整, 使得跟踪能够较好地处理遮挡情况.



图 6 face 视频跟踪中自适应分块结果

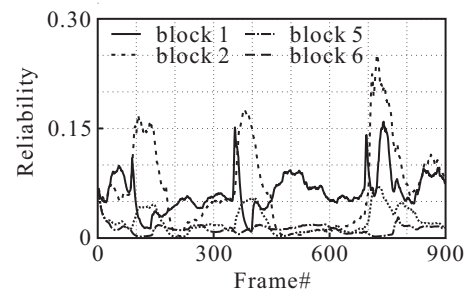


图 7 face 视频跟踪中子块权重变化

3.4 光照剧烈变化和目标表现变化情况下的跟踪

实验采用文献 [12] 提供的 David 视频序列对目标头部进行跟踪. 该视频序列一开始光照较弱, 后来越来越强, 中间过程中目标出现了多种姿态变化和尺

度变化,还存在与目标有较大相似性的背景物体干扰. 实验中,分块阈值 $T_C = 15$, 得到的初始分块数 $H = 8$, $V = 6$, 定性实验结果如图 8 所示.

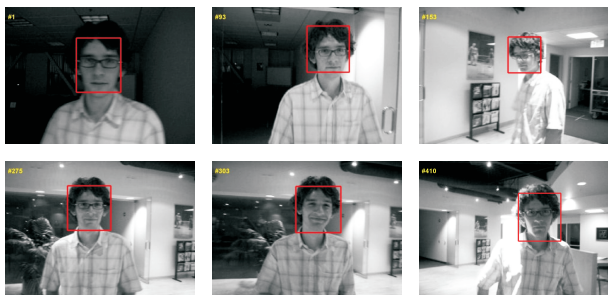


图 8 David 视频跟踪结果

从图 8 可以看出,自适应分块多线索融合粒子滤波跟踪方法对于光照剧烈变化和 target 姿态剧烈变化的情况,也可以得到较稳定的跟踪结果.

4 结 论

本文提出了一种基于自适应分块目标模型的多线索融合粒子滤波跟踪方法. 根据目标的颜色分布进行水平投影和垂直投影,找到垂直和水平方向相邻范围内相近的颜色块进行自适应分块,这种分块方法可以使初始的目标描述更加准确有效;提取子块的颜色直方图组成多线索,用粒子滤波进行目标跟踪,跟踪过程中子块的权重和数目进行自适应调整,模板也进行相应更新. 多种复杂情况下的实验结果表明,本文所提出的方法跟踪准确有效. 目前,该方法只使用了一种灰度直方图特征来描述多个子块组成多线索,下一步可以考虑用多种特征描述多个子块作为多线索建立目标模型. 如何将该方法应用到文献 [13] 中的大视域多摄像机跟踪系统中也是未来的研究方向.

参考文献(References)

- [1] Yilmaz A, Javed O, Shah M. Object tracking: A survey[J]. ACM Computer Survey, 2006, 38(4): 1-45.
- [2] Ning J F. Robust object tracking using joint color-texture histogram[J]. Int J of Pattern Recognition and Artificial Intelligence, 2009, 23(7): 1245-1263.
- [3] Wang D, Lu H C, Chen Y W. Object tracking by multi-cues spatial pyramid matching[C]. Proc of IEEE 17th Int Conf on Image Processing. Hong Kong: IEEE, 2010: 3957-3960.
- [4] 顾鑫, 王海涛, 汪凌峰, 等. 基于不确定性度量的多特征融合跟踪[J]. 自动化学报, 2011, 37(5): 550-559. (Gu X, Wang H T, Wang L F, et al. Fusing multiple features for object tracking based on uncertainty measurement[J]. Acta Automatica Sinica, 2011, 37(5): 550-559.)
- [5] Elgammal A, Duraiswami R, Davis L S. Efficient kernel density estimation using the fast gauss transform with applications to color modeling and tracking[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2003, 25(11): 1499-1504.
- [6] Mittal A, Davis L S. M2tracker: A multi-view approach to segmenting and tracking people in a cluttered scene[J]. Int J of Computer Vision, 2002, 51(3): 189-203.
- [7] Adam A, Rivlin E, Shimshoni I. Robust fragments-based tracking using the integral histogram[C]. IEEE Computer Society Conf on Computer Vision and Pattern Recognition. New York: IEEE, 2006: 798-805.
- [8] Erdem E, Dubuisson S, Bloch I. Fragments based tracking with adaptive cue integration[J]. Computer Vision and Image Understanding, 2012, 116(7): 827-841.
- [9] Isard M, Blake A. Condensation conditional density propagation for visual tracking[J]. Int J of Computer Vision, 1998, 29(1): 5-28.
- [10] Nickel K, Stiefelhagen R. Dynamic integration of generalized cues for person tracking[C]. Proc of Europe Conf Computer Vision. Marseille: Springer, 2008: 514-526.
- [11] Spengler M, Schiele B. Towards robust multi-cue integration for visual tracking[J]. Machine Vision and Applications, 2003, 14(1): 50-58.
- [12] Ross D A, Lim J, Lin R S, et al. Incremental learning for robust visual tracking[J]. Int J of Computer Vision, 2008, 77(1): 125-141.
- [13] Chang F L, Dong W H, Li J B. Scheduling strategy for network-based multi-camera relay tracking in large scene[J]. J of Computational Information System, 2013, 9(1): 161-169.

(责任编辑: 齐 霖)