

一种多线索融合的均值偏移跟踪算法

马加庆, 韩崇昭

(西安交通大学电子与信息工程学院, 710049, 西安)

摘要: 针对传统基于的均值偏移(Mean Shift, MS)跟踪算法不能对运动目标准确跟踪的缺点, 提出了一种融合多视觉线索的 MS 跟踪算法. 首先根据目标在前一帧的估计位置, 将目标搜索区域划分为目标区域和背景区域, 其次在区域划分的基础上定义了一种新的颜色、运动线索直方图模型, 能有效地抑制目标相邻背景的混乱干扰, 最后基于 MS 理论框架提出了一种融合目标颜色、运动线索的跟踪算法, 其颜色、运动线索可在跟踪过程中互补. 实验表明, 在目标快速运动、姿态发生较大变化或被遮挡的情况下, 算法能够获得更为准确、鲁棒的跟踪结果.

关键词: 视觉跟踪; 颜色线索; 运动线索

中图分类号: TP391 **文献标志码:** A **文章编号:** 0253-987X(2009)10-0042-05

A Mean Shift Tracking Algorithm Based on Multi-Cue Fusion

MA Jiaqing, HAN Chongzhao

(School of Electronics and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: The traditional color-based mean shift algorithm is unable to accurately track the object in some challenging scenarios. A mean shift tracker that fuses multiple visual cues is proposed to address the problem. The search region of the object is partitioned into one target region and some background regions according to the estimated location of the object in the previous frame. Then based on the region partition, a novel histogram model is presented for characterizing the color and motion cues. The model can effectively suppress the clutter of adjacent backgrounds. Finally the visual tracker is proposed within the mean shift framework to fuse the color and motion cues so that both. The visual cues can be complemented each other and the accuracy and robustness of tracking is improved. Experiments show that the proposed tracker can obtain more accurate and robust tracking results in some challenging scenarios such as rapid motion, large pose variation or occlusion of objects.

Keywords: visual tracking; color cue; motion cue

传统的均值偏移(Mean Shift, MS) 算法计算速度快, 跟踪效果好, 因而吸引了众多研究者的关注^[1-3], 但在一些复杂的场景中, 基于单个视觉线索的 MS 跟踪算法往往不能取得较为准确、鲁棒的跟踪结果. 如果能在实际场景中将目标的多个视觉线索融合在跟踪方法中, 则可以利用线索之间的互补性更好地解决真实场景下跟踪的精度和鲁棒性问题^[4-6].

基于上述考虑和在摄像机静止的假定条件下,

本文依据传统的 MS 理论框架提出了一种融合目标颜色和运动线索的视觉跟踪算法, 融合了描述目标外观的颜色线索和描述目标运动的运动线索, 从而提高了算法的跟踪准确性和鲁棒性.

1 跟踪算法描述

1.1 颜色线索描述

目标参考颜色直方图 $q^c = \{q_u^c\}_{u=1 \dots B_c}$, 其中 B_c

为直方图颜色区间的个数. 假设上一帧目标的估计中心位置为 $\mathbf{y}$, 则在当前帧对目标所在的搜索区域进行划分的方法如图 1 所示. 为简单起见, 本文假设每一块背景区域均与目标候选区域的面积相同, 且中心位置由目标候选区域中心位置 $\mathbf{y}$ 的相应坐标加上(或减去)椭圆的长半轴或短半轴的长度获得; 又设 4 个椭圆所围背景区域的中心位置为 $\{\mathbf{y}_j\}_{j=1 \dots 4}$, 背景区域的参考颜色直方图 $\mathbf{b}_j^c = \{b_{ju}^c\}_{u=1 \dots B_c}$, $j=1 \dots 4$. 在实际操作中, 背景区域的参考颜色直方图可以在手工初始化跟踪目标时采用核密度估计方法来计算. 考虑到视频序列时域的连续性, 在跟踪过程中, 当前帧的参考颜色直方图始终可用上一帧观察到的背景颜色直方图来替代. 如果用 $\{\mathbf{x}_i\}_{i=1 \dots n}$ 表示中心位置为 $\mathbf{y}$ 的目标候选区域的像素位置, 则目标候选区域的颜色直方图 $\mathbf{p}^c(\mathbf{y} = \{p_u^c(\mathbf{y})\}_{u=1 \dots B_c})$, 其中

$$p_u^c(\mathbf{y}) = \frac{1}{C_1} \sum_{i=1}^n k\left(\left\|\frac{\mathbf{y} - \mathbf{x}_i}{h}\right\|^2\right) \delta[b_f(\mathbf{x}_i) - u] \cdot q_u^c(\mathbf{x}_i) k\left(\left\|\frac{\mathbf{y} - \mathbf{x}_i}{h}\right\|^2\right) / [q_u^c(\mathbf{x}_i) k\left(\left\|\frac{\mathbf{y} - \mathbf{x}_i}{h}\right\|^2\right) + \sum_{j=1}^4 b_{ju}^c(\mathbf{x}_i) k\left(\left\|\frac{\mathbf{y}_j - \mathbf{x}_i}{h}\right\|^2\right)] \quad (1)$$

式中: $k(\cdot)$ 表示空间加权函数; $q_u^c(\mathbf{x}_i)$ 、 $b_{ju}^c(\mathbf{x}_i)$ 分别表示 $\mathbf{x}_i$ 位置像素在目标参考颜色直方图和背景参考颜色直方图上的颜色区间直方图值; $b_f(\cdot)$ 表示将给定位置的像素映射到直方图颜色区间的映射函数; δ 表示 Kronecker 函数; C_1 表示颜色直方图归一化常数^[7]. 相对于文献[7], 式(1)所示的颜色直方图对于目标中和背景具有共同颜色特征的像素赋予了较小的权值, 而对于目标独有颜色特征的像素赋予了较大的权值, 因此式(1)给出的颜色直方图能有效地抑制邻近相似背景的干扰.

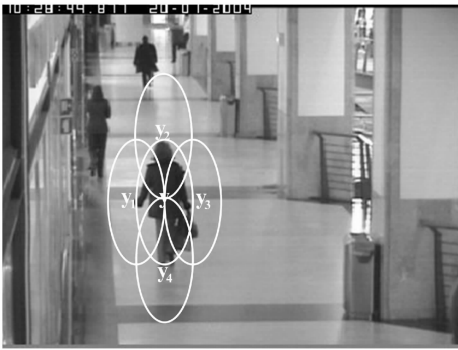


图 1 目标搜索区域划分方法

1.2 运动线索描述

本文用相邻帧差绝对值图像描述目标的运动线

索. 为了便于描述, 文中将帧差绝对值图像的颜色直方图作为运动直方图. 一般来说, 当某一区域静止时, 帧差绝对值图像在该区域的像素值会落入运动直方图的低颜色值区间; 当某一区域存在运动时, 帧差绝对值图像在该区域的像素值会随机落入直方图的每个颜色区间. 此时, 均匀分布的帧差绝对值图像颜色直方图将作为参考运动直方图^[8]. 然而, 当运动目标邻近区域存在其他干扰时, 均匀分布的颜色直方图作为参考运动直方图往往无法从运动干扰中分辨出所要跟踪的目标. 为了解决这一问题, 在考虑目标运动持续性的情况下, 应将上一帧中根据目标估计位置计算出的目标运动直方图 $\mathbf{p}_{t-1}^m$ 设为均匀分布的情况. 运动参考直方图定义为

$$\mathbf{q}^m = \{q_u^m\}_{u=1 \dots B_m} = (1 - \alpha) \mathbf{p}_{t-1}^m + \alpha \bar{\mathbf{q}}^m \quad (2)$$

式中: $\bar{\mathbf{q}}^m = \{\bar{q}_u^m\}_{u=1 \dots B_m}$ 表示均匀分布直方图, 即 $\bar{q}_u^m = 1/B_m$, $u=1 \dots B_m$, 其中 B_m 表示直方图颜色区间的个数; $\alpha \in (0, 1)$ 为权值. 一般来说 α 的选取与运动线索的可靠性有关, 当跟踪运动目标接近未干扰目标时, 目标所在帧差绝对值图像区域仅仅包含目标的运动信息, 而这一区域的运动直方图可以近似为均匀分布^[8]的情况, 所以 α 取值较大, 反之 α 取值较小.

类似于颜色线索描述, 对于运动线索, 假设给定背景区域的参考运动直方图 $\mathbf{b}_j^m = \{b_{ju}^m\}_{u=1 \dots B_m}$, $j=1 \dots 4$, 则目标候选区域的运动直方图 $\mathbf{p}^m(\mathbf{y} = \{p_u^m(\mathbf{y})\}_{u=1 \dots B_m})$, 其中

$$p_u^m(\mathbf{y}) = \frac{1}{C_2} \sum_{i=1}^n k\left(\left\|\frac{\mathbf{y} - \mathbf{x}_i}{h}\right\|^2\right) \delta[b_f(\mathbf{x}_i) - u] \cdot q_u^m(\mathbf{x}_i) k\left(\left\|\frac{\mathbf{y} - \mathbf{x}_i}{h}\right\|^2\right) / [q_u^m(\mathbf{x}_i) k\left(\left\|\frac{\mathbf{y} - \mathbf{x}_i}{h}\right\|^2\right) + \sum_{j=1}^4 b_{ju}^m(\mathbf{x}_i) k\left(\left\|\frac{\mathbf{y}_j - \mathbf{x}_i}{h}\right\|^2\right)] \quad (3)$$

式中: $q_u^m(\mathbf{x}_i)$ 和 $b_{ju}^m(\mathbf{x}_i)$ 分别表示帧差绝对值图像的 $\mathbf{x}_i$ 位置像素在目标和背景参考运动柱状图上的颜色区间的直方图值; C_2 为运动直方图归一化常数. 类似地, 式(1)、式(3)给出的运动直方图能有效地抑制候选目标邻近区域中运动目标的干扰.

1.3 MS 搜索

给定目标在上一帧的估计位置 $\hat{\mathbf{y}}_0$, 则目标在当前帧的位置 $\hat{\mathbf{y}}_1$ 可由最大化目标函数获得, 即

$$D(\mathbf{y}) = \rho(\mathbf{q}^c, \mathbf{p}^c(\mathbf{y})) + \rho(\mathbf{q}^m, \mathbf{p}^m(\mathbf{y})) \quad (4)$$

式中: $\rho(\cdot)$ 表示 Bhattacharyya 系数^[7]. 将式(4)在 $\mathbf{p}^c(\hat{\mathbf{y}}_0)$ 和 $\mathbf{p}^m(\hat{\mathbf{y}}_0)$ 处进行 Taylor 展开, 得

$$D(\mathbf{y}) \approx L_c(\mathbf{y}) + L_m(\mathbf{y}) \quad (5)$$

式中

$$L_c(\mathbf{y}) = c_1 + \frac{1}{2C_1} \sum_{i=1}^n (\omega_i^c k(\|\frac{\mathbf{y}-\mathbf{x}_i}{h}\|^2)) \cdot (1 + \sum_{j=1}^4 \frac{b_{ju}^c(\mathbf{x}_i)}{q_u^c(\mathbf{x}_i)} k(\|\frac{\mathbf{y}_j-\mathbf{x}_i}{h}\|^2))^{-1} \quad (6)$$

$$L_m(\mathbf{y}) = c_2 + \frac{1}{2C_2} \sum_{i=1}^n (\omega_i^m k(\|\frac{\mathbf{y}-\mathbf{x}_i}{h}\|^2)) \cdot (1 + \sum_{j=1}^4 \frac{b_{ju}^m(\mathbf{x}_i)}{q_u^m(\mathbf{x}_i)} k(\|\frac{\mathbf{y}_j-\mathbf{x}_i}{h}\|^2))^{-1} \quad (7)$$

其中

$$\omega_i^c = \sum_{u=1}^{B_c} (\frac{q_u^c}{p_u^c(\hat{\mathbf{y}}_0)})^{1/2} \delta[b_f(\mathbf{x}_i) - u] \quad (8)$$

$$\omega_i^m = \sum_{u=1}^{B_m} (\frac{q_u^m}{p_u^m(\hat{\mathbf{y}}_0)})^{1/2} \delta[b_f(\mathbf{x}_i) - u] \quad (9)$$

c_1, c_2 为常数^[7]. 将式(6)、式(7)代入式(5), 得

$$D(\mathbf{y}) \approx c_3 + \frac{1}{2} \sum_{i=1}^n (\omega_i^c + \omega_i^m) k(\|\frac{\mathbf{y}-\mathbf{x}_i}{h}\|^2) \quad (10)$$

式中

$$\omega_i^c = \frac{\omega_i^c}{C_1} (1 + \sum_{j=1}^4 \frac{b_{ju}^c(\mathbf{x}_i)}{q_u^c(\mathbf{x}_i)} k(\|\frac{\mathbf{y}_j-\mathbf{x}_i}{h}\|^2))^{-1} \quad (11)$$

$$\omega_i^m = \frac{\omega_i^m}{C_2} (1 + \sum_{j=1}^4 \frac{b_{ju}^m(\mathbf{x}_i)}{q_u^m(\mathbf{x}_i)} k(\|\frac{\mathbf{y}_j-\mathbf{x}_i}{h}\|^2))^{-1} \quad (12)$$

$c_3 = c_1 + c_2$ 为常数. 应用下式进行 MS 迭代^[1], 当核窗口由位置 $\hat{\mathbf{y}}_0$ 移动到 $\hat{\mathbf{y}}_1$ 时, 式(10)右端将会增加项

$$\hat{\mathbf{y}}_1 = \sum_{i=1}^n \mathbf{x}_i \omega_i g(\|\frac{\mathbf{y}_0-\mathbf{x}_i}{h}\|^2) \cdot (\sum_{i=1}^n \omega_i g(\|\frac{\mathbf{y}_0-\mathbf{x}_i}{h}\|^2))^{-1} \quad (13)$$

式中: $g() = -k'()$; $\omega_i = \omega_i^c + \omega_i^m$. 令 $\hat{\mathbf{y}}_0 = \hat{\mathbf{y}}_1$, 继续迭代, 直到 $\|\hat{\mathbf{y}}_1 - \hat{\mathbf{y}}_0\| < \epsilon$ 或者达到最大迭代次数时为止, 这里 ϵ 是预先给定的较小的正数. 为了适应目标尺度的变化, 加权核函数的半径以 $\pm 10\%$ 变化, 并在新的尺度上执行跟踪算法, 直到收敛, 最后选择可使式(10)右端取得极大值的尺度.

2 实验结果

为验证算法的有效性, 本文选取了 3 组图像序列进行实验, 结果见图 2~图 4. 为便于叙述, 将文献[4]中的基于多特征分布的 MS 算法称之为 MF-DMS 算法. 将本文算法与 MFDMS 算法、传统 MS 算法进行对比, 在每一组图像序列实验中, 目标在第一帧的位置均用手工标定, 目标初始尺度相同.

实验 1 在图像序列 1 描述的场景中, 行人从图像靠近底部位置走到中部位置, 然后从中部位置经过短暂停留又返回到底部位置. 在这一过程中, 由于相



(a)传统 MS 算法



(b)MFDMS 算法



(c)本文算法

图 2 图像序列 1 跟踪结果



(a)传统 MS 算法



(b)MFDMS 算法



(c)本文算法

图 3 图像序列 2 跟踪结果



(a)传统 MS 算法



(b)MFDMS 算法



(c)本文算法

图 4 图像序列 3 跟踪结果

机视角发生变化,所以行人的姿态发生了较大的变化. 算法的跟踪结果如图 2 所示,可见传统 MS 算法虽然没有丢失目标,但对目标的跟踪定位误差比较大. MFDMS 算法在行人从图像底部向中部走动的过程中完全丢失了目标,而在行人重新回到图像底

部的过程中又重新捕获到目标,显然跟踪偏差较大. 本文算法对于目标独有的颜色特征像素,其颜色线索赋予了较大的权值,而对于目标中的和背景共有的颜色特征像素,其颜色线索赋予了较小的权值. 基于帧差的运动线索反映了目标的轮廓和边缘位置像

素的运动信息,这种运动线索使得位于目标轮廓和边缘位置的像素具有较大的权值,在一定程度上弥补了颜色线索对目标定位不够准确的缺陷.所以,本文算法在整个跟踪过程中能较为准确地定位运动目标.

实验 2 在视频序列 2 描述的场景中,人脸从左向右移动时比较缓慢,从右向左移动时比较快,视频的帧采样率为 10 帧/s.算法的跟踪结果如图 3 所示.当人脸从左向右缓慢移动时,传统 MS 算法还算奏效,但当人脸从右向左快速移动时,传统 MS 算法在目标尺度不断扩大时丢失了目标. MFDMS 算法可以跟踪目标,但跟踪的偏差比较大.本文算法对目标运动的像素赋予了较大的权值,使得无论人脸移动速度是快、是慢,都能较准确地跟踪目标.

实验 3 视频图像序列 3 是移动摄像机拍摄的场景,本文从中选取了背景相对静止、人脸被遮挡的 25 帧.算法的跟踪结果如图 4 所示.当移动的纸张遮挡在人脸时,传统 MS 算法很快丢失了目标, MFDMS 算法可以跟踪运动目标,但目标定位不够准确,而本文算法能够有效地解决遮挡问题,并获得了较为准确的跟踪结果.

3 结束语

本文基于传统 MS 跟踪算法的理论框架,在摄像头静态的情况下,提出了一种融合颜色线索和运动线索的视觉跟踪算法.这 2 种视觉线索可以在跟踪过程中互相弥补彼此间的不足,并能够有效地抑制邻近背景颜色或运动目标的干扰,从而使得本文算法能够更为准确地跟踪运动目标.复杂场景中的多组实验对比结果也说明了本文算法是有效的.

参考文献:

[1] COLLINS R T. Mean-shift blob tracking through scale space [C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA; IEEE, 2003: 234-240.

[2] PENG Ningsong, YANG Jie, LIU Zhi. Mean shift blob tracking with kernel histogram filtering and hypothesis testing [J]. Pattern Recognition Letters, 2005, 26(5): 605-614.

[3] YANG C, DURAISWAMI R, DAVIS L S. Efficient mean-shift tracking via a new similarity measure [C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA; IEEE, 2005: 176-183.

[4] POLAT E, OZDEN M. A nonparametric adaptive tracking algorithm based on multiple feature distributions [J]. IEEE Trans on Multimedia, 2006, 8(6): 1156-1163.

[5] WANG Junqiu, YAGI Y. Integrating color and shape-texture features for adaptive real-time object tracking [J]. IEEE Trans on Image Processing, 2008, 17(2): 235-240.

[6] JEYAKAR J, BABU V R. Robust object tracking with background-weighted local kernels [J]. Computer Vision and Image Understanding, 2008, 112(3): 296-309.

[7] COMANICIU D, MEER P. Kernel-based object tracking [J]. IEEE Trans on PAMI, 2003, 25(5): 564-577.

[8] PÉREZ P, VERMAAK J, BLAKE A. Data fusion for visual tracking with particles [J]. Proceedings of the IEEE, 2004, 92(3): 495-513.

(编辑 苗凌)

西安交通大学 4 教授入选全国博士后管委会专家组

目前,国家人力资源和社会保障部、全国博士后管委会公布了全国博士后管委会第七届专家组成员名单,西安交通大学卢秉恒院士为工学一组成员(含力学、机械工程、动力工程及工程热物理、林业工程、农业工程),卢黎歌教授为法学组成员(含法学、政治学、社会学、民族学、马克思主义理论),任巍教授为工学五组成员(含电子科学与技术、信息与通信工程、计算机科学与技术),郑崇勋教授为工学九组成员(含光学工程、仪器科学与技术、生物医学工程)。

全国博士后管委会专家组每届任期两年,是全国博士后管委会的学术性咨询、评议组织,主要在博士后政策咨询、博士后科研流动站评审、博士后基金评审等方面开展工作。

(来源:交大新闻网)