

增强目标模型鲁棒性的 Mean-shift 算法

郭运艳, 陈 愚*, 宋康康, 刘 哲, 黄小霞

(宁波大学 信息科学与工程学院, 浙江 宁波 315211)

摘要: 针对传统 Mean-shift 算法中颜色核函数直方图对目标特征描述较弱的缺点, 提出了一种联合目标特征点的二维结构信息和颜色信息的 Mean-shift 改进算法. 改进算法细化了 Harris 检测算子的角点响应阈值, 提取出更多的目标特征点计算其方向分布, 并以方向与部分颜色特征的直方图构建目标模型, 该模型能显著区分目标与背景. 实验对不同算法进行了仿真及性能比较, 结果表明: 提出的改进算法在一定的复杂场景中提高了跟踪精度, 且具有较好的鲁棒性.

关键词: Mean-shift; Harris 检测算子; 旋转不变性; 联合直方图

中图分类号: TN919.8

文献标识码: A

文章编号: 1001-5132 (2012) 04-0025-04

Mean-shift 算法是一种非参的概率密度估计算法^[1], Comaniciu 等人^[2]利用 Bhattacharyya 系数作为目标模型与候选目标模型之间的相似度量标准, 用 Mean-shift 算法搜索最优候选目标, 获得了很好的跟踪效果. 但 Mean-shift 算法采用的颜色直方图是一种较弱的目标特征描述, 当背景或者背景中存在于目标特征类似的物体时, 将增加目标跟踪的难度, 从而导致跟踪过程不稳定. 为解决该问题, 很多研究者对 Mean-shift 算法进行了改进和优化. 文献[3]引入混合直方图并对直方图重新量化, 再通过减少背景像素在概率密度函数 PDF 中的权重来对背景进行优化, 从而降低背景区域对跟踪的影响. 文献[4]根据 2 个区域中统计特征数值的大小关系将模板特征分为目标类特征和交融类特征, 利用后验概率的图像匹配相似性搜索区域的统计特征, 有效抑制了待匹配区域特征中背景因素的影响, 同时突出了目标特征的权重. 这些算法虽然在一定程度上凸显了目标特征, 但是同时也增大了跟踪的复杂度.

目标特征^[5]是目标跟踪的一个关键因素, 并且目标特征选取的好坏直接影响到跟踪性能. 而角点是一种非常重要的图像点特征, 包含了图像中

比较丰富的二维结构信息, 它不仅能有效地抑制背景成分对目标的影响, 而且能加速 Mean-shift 算法收敛. 文献[6]已给出了理论与实验证明, 但文献[6]中依然单纯地运用角点的颜色信息来构建核函数直方图, 抗干扰性和鲁棒性并不稳定. 对此, 笔者在文献[6]的基础上提出了一种联合目标特征点的二维结构信息的 Mean-shift 改进算法.

1 Mean-shift 运动目标跟踪分析

传统 Mean-shift 算法应用于跟踪领域^[7]是一种基于目标颜色直方图^[8]作为目标特征描述的跟踪算法, 能够对机动目标进行较好的实时跟踪. 设 y_0 为跟踪窗口的中心位置, h 为带宽, $y_0, y_1, y_2, \dots, y_i$ 为核窗口内的所有采样点, 目标模型表示为:

$$\tilde{q}_u = C \sum_{i=1}^n k \left(\left\| \frac{y_0 - y_i}{h} \right\|^2 \right) \delta [b(y_i) - u], \quad (1)$$

其中, $k(x)$ 是一个非负、非增、分段连续的核函数, 且满足 $\int k(x) dx = 1$, n 为特征值 bin 的个数; $b(y_i)$ 为像素 y_i 对应的特征值; $\delta(x)$ 为 Kronecker 函数, 判断 y_i 处的颜色信息是否属于特征值 u ; C 为归一化系数.

相应的候选目标的特征分布为:

收稿日期: 2012-06-02.

宁波大学学报(理工版) 网址: <http://3xb.nbu.edu.cn>

基金项目: 国家自然科学基金(61071120); 浙江省教育厅科研项目(Y200907622); 宁波市自然科学基金(210A610109).

第一作者: 郭运艳(1987-), 女, 山东临沂人, 在读硕士研究生, 主要研究方向: 视频目标跟踪及图像处理. E-mail: yunyan_509@163.com

*通讯作者: 陈 愚(1962-), 男, 重庆人, 博士/副教授, 主要研究方向: 图像与视频分析处理及智能控制. E-mail: chenken@nbu.edu.cn

$$\hat{p}_u = C_h \sum_{i=1}^{n_h} k \left(\left\| \frac{y_i - y}{h} \right\|^2 \right) \delta[b(x_i) - u], \quad (2)$$

式中, C_h 为归一化系数; n_h 为候选目标区域中的像素总数. 候选模型与目标模型间的相似度采用巴氏系数来衡量, 定义为:

$$\rho[\hat{p}(y), \hat{q}] = \sum_{u=1}^m \sqrt{\hat{p}_u(y) \hat{q}_u}, \quad (3)$$

(3)式在 $p(y_0)$ 处进行泰勒展开, 保留一阶近似:

$$\rho[\hat{p}(y), \hat{q}] \approx \frac{1}{2} \sum_{u=1}^m \sqrt{\hat{p}_u(\hat{y}_0) \hat{q}_u} + \frac{C}{2} \sum_{i=1}^{n_c} w_i k \left(\left\| \frac{y - y_i}{h} \right\|^2 \right), \quad (4)$$

其中,

$$w_i = \sum_{u=1}^m \sqrt{\frac{\hat{q}_u}{\hat{p}_u(y_0)}} \delta[b(y_i) - u].$$

(4)式中的第一项与 y 无关, 要使巴氏系数最大, 只需用 Mean-shift 算法的迭代式(5)取得使巴氏系数第二项最大(或达到规定的迭代次数)的 y 为止即可. 迭代公式如下:

$$y_1 = \frac{\sum_{i=1}^{n_c} y_i w_i g \left(\left\| \frac{\hat{y}_0 - y_i}{h} \right\|^2 \right)}{\sum_{i=1}^{n_c} w_i g \left(\left\| \frac{\hat{y}_0 - y_i}{h} \right\|^2 \right)}. \quad (5)$$

在运动目标跟踪中, 目标颜色与背景颜色的相似程度、背景的稳定程度等因素影响到目标特征对目标模型描述的准确性, 进而加大了目标跟踪难度. 由此, 运动目标跟踪的特征选择也是关键.

2 特征点提取及度量

2.1 特征点的提取

Harris 等人^[9]提出的 Harris 算子计算简单、提取的角点特征均匀合理, 其处理过程表示如下:

$$c(x, y, \Delta x, \Delta y) = \sum_{(u, v) \in W(x, y)} w(u, v) (I(u, v) - I(u + \Delta x, v + \Delta y))^2, \quad (6)$$

其中, $w(x, y)$ 是以点 (x, y) 为中心的窗口, $w(x, y)$

为高斯加权函数: $\exp \frac{-(u-x)^2 - (v-y)^2}{2\sigma^2}$, $\sigma = 2$, (6)式根据泰勒展开, 并取一阶近似, 可表示为:

$$c(x, y, \Delta x, \Delta y) \approx [\Delta x, \Delta y] M(x, y) \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}, \quad (7)$$

$$M(x, y) = \sum_w \begin{bmatrix} I_x(x, y)^2 & I_x(x, y)I_y(x, y) \\ I_x(x, y)I_y(x, y) & I_y(x, y)^2 \end{bmatrix} = \begin{bmatrix} A & C \\ C & B \end{bmatrix}, \quad (8)$$

其中, I_x, I_y 是图像 $I(x, y)$ 的一阶偏导数.

Harris 提出角点判别方法无需计算具体的特征值, 而用角点响应值 R 来判断角点. R 计算公式为:

$$R = \det M - \alpha(\text{trace} M)^2, \quad (9)$$

其中, $\det M = \lambda_1 \lambda_2 = AC - B^2$ 为矩阵 M 的行列式; $\text{trace} M = \lambda_1 + \lambda_2 = A + C$ 为矩阵 M 的直迹; α 为经验常数, 取值范围是 0.04~0.07.

图 1 即为响应值 R 与点类型之间的关系.

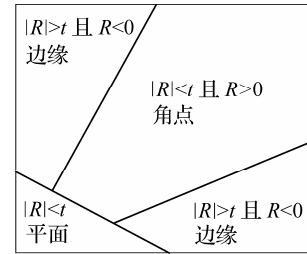
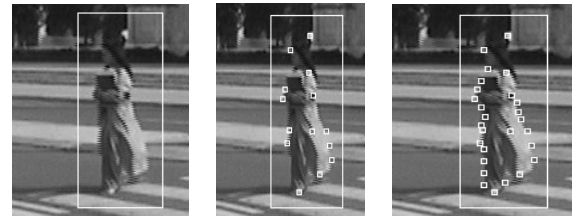


图 1 R 与点的类型关系 (t 为设定的角点响应阈值)

那么调节阈值大小, 同时判定响应值 R 的正负, 将边缘点与角点均作为目标的特征点提取出来, 而细化后获得的特征点如图 2 所示.



(a) 目标区域

(b) 目标角点

(c) 边缘角点

图 2 阈值

2.2 特征点的量化及度量

Harris 角点检测算子具有旋转不变性, 对复杂环境具有一定鲁棒性. 可采用方向直方图来表征特征点的信息. 点 (x, y) 处的梯度强度可表示为:

$$D(x, y) = \sqrt{I_x(x, y)^2 + I_y(x, y)^2}, \quad (10)$$

该像素点的方向可由下式进行计算:

$$\theta(x, y) = \arctan(D_y(x, y) / D_x(x, y)). \quad (11)$$

角度(即方向)平均分布在 $(0 \sim 180^\circ)$ 之间的无向角度范围内, 文中量化为 9 个 bin, 每个特征点都需要为某个 bin 提供权值. 由于检测到的特征点均是感兴趣点, 故权值的确定方法不同于颜色直方图

加权. 方向的权值是根据该特征点的梯度幅度计算出来, 实验测试表明, 使用幅值来表示某个方向的权值能获得最佳效果, 则图 2(c)的特征点方向加权直方图如图 3 所示.

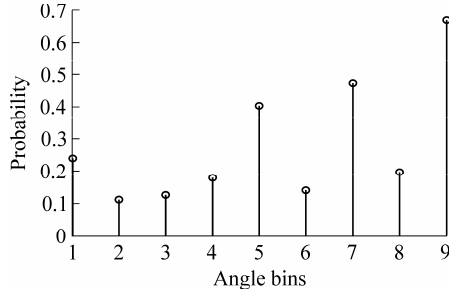


图3 图 2(c)中特征点的方向直方图

2.3 联合直方图

人眼对于相同强度单色光主观亮度感觉^[10]不同, 人的主观感觉是绿光仅次于白光, 是三原色中最亮的, 红光次之, 亮度约占绿光的一半, 蓝光最弱, 亮度约占红光的 1/3. 为尽量不增加跟踪过程的复杂度, 笔者仅采用红、绿 2 个颜色通道, 用量化的方向信息代替蓝色通道, 建立联合直方图 $\text{Hist}(\text{R}, \text{G}, \text{A})$, 大小为 $16 \times 16 \times 9$.

则目标区域的联合直方图计算方法如下: (1) 依据 2.1 节和 2.2 节的原理, 提取出目标区域内的特征点; (2) 分别计算颜色(R, G)的加权分布直方图和加权方向直方图, 得到相应的特征向量; (3) 归一化目标直方图, 得到初始目标模型.

实验证明, 通过归一化后能对光照变化和阴影获得更好的效果, 候选目标模型的计算方法同上.

3 实验结果及分析

3.1 基于联合直方图的跟踪算法

算法的流程如下:

(1) 在第 1 帧图像中手动选择跟踪目标, 计算联合直方图 $\text{Hist}(\text{R}, \text{G}, \text{A})$ 并归一化, 得初始化目标模型 $\tilde{q}_u$;

(2) 计算目标在当前帧中的中心位置 y_0 , 初始化迭代次数 $m = 0$;

(3) 在当前帧的目标中心处, 计算得候选目标模型 $\hat{p}_u$;

(4) 根据 $w_i = \sum_{u=1}^m \sqrt{\frac{\hat{q}_u}{\hat{p}_u(y_0)}} \delta[b(y_i) - u]$ 计算权值;

(5) 得到目标的新中心 y_1 , 迭代次数 $m = m + 1$;

(6) $y_1 \leftarrow y_0$, 若 $\|y_1 - y_0\|$ 足够小或者迭代次数 $m \leq N$, 则停止迭代, 否则返回(3), 继续执行.

3.2 实验结果比较

为验证算法在复杂背景下的有效性和鲁棒性, 在 2.00 GHz CPU, 512MB 内存 PC 上以 Matlab 7.1 为开发工具, 通过对多段图像序列进行跟踪测试.

实验一采用视频序列 “Laboratory.avi”, 每帧图像的尺寸为 320×240 , 跟踪对象是人. 人与背景具有非常相似的颜色, 并且目标的灰度变化不明显, 所以在整个视频序列中能够提取的目标的角点信息不足, 则角点的 RGB 颜色直方图就不能准确的表示目标了. 图 4 和图 5 分别是运用文献[6]的方法与本文算法的跟踪效果图, 可以看出, 文献[6]算法从 99 帧的位置开始不能定位到目标.

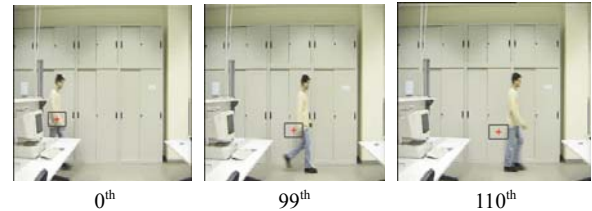


图4 文献[6]算法的跟踪结果

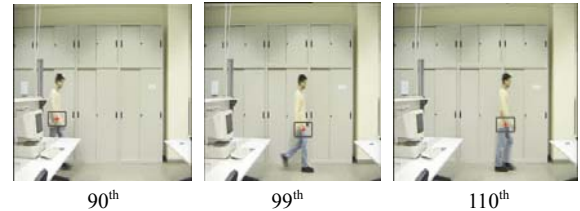


图5 本文提出算法的跟踪效果

表 1 则是 2 种算法的匹配程度均值与方差比较数据, 与文献[6]中算法相比, 本文算法的匹配程度有了很大提高, 跟踪偏差减少, 使得目标跟踪的鲁棒性明显增强. 跟踪偏差减少主要是因为本文算法中的联合直方图增大了候选目标模型与目标模型之间的差异, 减少了背景像素的影响, 从而使定位更精确, 说明笔者提出的算法是有效性.

表 1 2 种算法的匹配程度均值与方差

算法	匹配程度均值	匹配程度方差
文献[6]跟踪算法	0.062 2	0.107 5
提出的算法	0.183 7	0.031 4

实验二则测试算法的跟踪精度及优越性. 采用的视频序列是来自 CVRR-ATON 标准的 “tennis.avi”, 跟踪目标是与背景有明显的颜色区别的头部,

从图 6 可见, 2 种算法均能准确跟踪到目标.



图 6 序列二中的跟踪目标

图 7 为在 2 种算法下, 每帧定位到目标所需要的迭代次数, 表 2 统计了具体数字. 从表 2 可知, 改进算法的平均迭代次数提高了 27.3%, 平均每帧耗时减少了 23.8%.

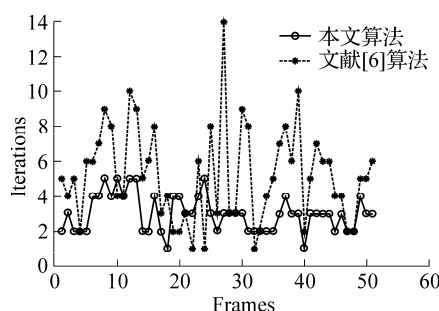


图 7 2 种算法迭代次数比较

表 2 2 种算法对 Tennis 运动员头部的跟踪性能比较

算法	平均迭代数/次	平均每帧耗时/s
文献[6]跟踪算法	3.3	0.056 1
提出的算法	2.4	0.042 7

4 结语

深入研究 Harris 检测算子, 对角点响应阈值细化, 得到了更多具有旋转不变性的特征点(角点和

边缘)的方向信息, 成功地嵌入了一种基于角点的 Mean-shift 算法中. 通过实验证明, 在复杂的背景环境中, 改进算法明显深化了目标与背景的差异, 降低了跟踪位置误差, 具有更好的鲁棒性.

参考文献:

- [1] Li P. An adaptive binning color model for mean shift tracking[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2008, 18(9):1293-1299.
- [2] Comaniciu D, Ramesh V, Meer P. Kernel-based object tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(5):564-577.
- [3] 柳伟, 罗以宁, 孙南. 基于背景优化的 Mean-shift 目标跟踪算法[J]. 计算机应用, 2009, 29(4):1015-1017.
- [4] 冯祖仁, 吕娜, 李良福. 基于最大后验概率的图像匹配相似性指标研究[J]. 自动化学报, 2007, 33(1):1-8.
- [5] Wang Yongming, Wang Guijin. Image local invariant features and descriptors[M]. Beijing: National Defense Industry Press, 2010.
- [6] Ning Jifeng, Jiang Guang, Li Peifei. Mean shift tracking algorithm based on corner points[J]. Application Research of Computers, 2009, 26(1):4348-4350.
- [7] Shen Huan, Li Shunming, Bo Fangchao, et al. On road vehicles detection approach using multi-cues fuse[J]. Journal of Optoelectronics Laser, 2010, 21(1):74-77.
- [8] 郑玉凤, 马秀荣. 基于颜色与边缘特征的均值迁移目标跟踪算法[J]. 光电子激光, 2011, 22(8):1231-1235.
- [9] Harris C, Stephens M J. A combined corner and edge detector[C]. Alley Vision Conference, 1988:147-152.
- [10] 荆其诚. 色度学[M]. 北京: 科学出版社, 1991.

Mean-shift Tracking with Enhanced Robustness of Target Model

GUO Yun-yan, CHEN Ken*, SONG Kang-kang, LIU Zhe, HUANG Xiao-xia

(Faculty of Information Science and Technology, Ningbo University, Ningbo 315211, China)

Abstract: The traditional Mean-shift cannot represent accurately the color distribution of the object. To tackle this problem, a target motion-tracking algorithm is presented based on two-dimensional structural information of the target feature points and color histogram. The improved algorithm refines the Harris operator threshold to seek more feature points, with which their orientation distribution can also be sought. The target model built by combining the orientation histogram and the partial color histogram is found to be able to separate target from background to a satisfying extent. The application results from various video sequences suggest good localization precision in object tracking, and algorithmic robustness in complex scenes.

Key words: Mean-shift; Harris; rotation-invariance; color and gradient orientation histogram distribution

(责任编辑 章践立)