

文章编号 1004-924X(2010)01-0234-06

鲁棒的基于改进 Mean-shift 的目标跟踪

薛 陈^{1,2}, 朱 明¹, 陈爱华^{1,2}

(1. 中国科学院 长春光学精密机械与物理研究所, 吉林 长春 130033;
2. 中国科学院 研究生院, 北京 100039)

摘要: 为了克服传统 Mean-shift 算法在跟踪运动目标时由于背景像素造成的定位偏差和由于遮挡造成的跟踪失效, 提出了相应的改进措施。其一, 根据初始帧目标和背景在颜色分布上的差异, 建立对数似然图(log-likelihood image), 筛选出目标中与背景可区分性好的颜色特征建立目标模型, 并以同样的方法在后续帧建立候选模型, 从而有效地减小背景像素的影响。另外, 将候选区域划分为若干重叠的子块, 分别利用 Mean-shift 算法对各个子块进行迭代, 以与目标区域相应子块最为匹配的子块的所在位置对整个目标重新定位, 由此很好地实现了目标部分遮挡情况下的稳定跟踪。当目标被严重遮挡时, 则采用简单的线性预测, 估计下一帧目标可能出现的位置。实验结果表明: 提出的改进算法可以准确地进行目标跟踪, 对部分遮挡和严重遮挡都有较强的鲁棒性。

关键词: 目标跟踪; Mean-shift; 对数似然图; 遮挡

中图分类号: TP301.6; TP391 **文献标识码:** A

Robust object tracking based on improved Mean-shift algorithm

XUE Chen^{1,2}, ZHU Ming¹, CHEN Ai-hua^{1,2}

(1. *Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China;*
2. *Graduate University of Chinese Academy of Sciences, Beijing 100039, China*)

Abstract: To overcome the shortcomings of the traditional Mean-shift algorithm for object tracking such as the localization error caused by background pixels and the tracking failure from the object occlusion, an improved Mean-shift algorithm is proposed. Firstly, according to the difference of color distribution between the object and the background in the initial frame, a log-likelihood image is set up to select the discriminative color features for object modeling, and then the candidate modeling is established by the same way. By above operation, the effect of background pixels on the image has reduced greatly. Secondly, the whole candidate region is separated into several overlapped fragments, and every fragment is iterated by the Mean-shift. Then, the object localization is reset by the location of fragment in the candidate region, which matches mostly to the corresponding fragment in the object region. Experimental results show that the fragment based on the Mean-shift is very robust to partial occlusion. Furthermore, when object is severely occluded, the linear prediction can be used to estimate the probable location of the object in the next frame. These results prove that the tracking using

收稿日期: 2008-12-08; 修订日期: 2009-01-16.

基金项目: 国家 863 高技术研究发展计划资助项目(No. 2005AA778032)

the improved Mean-shift algorithms has good localization precision and is robust to partial and severe occlusions.

Key words: object tracking; Mean-shift; log-likelihood image; occlusion

1 引 言

目标跟踪是机器视觉的一个重要研究方向,广泛应用于视频监控、运动分析、军事国防等众多领域。基于 Mean-shift^[1-2] 的跟踪算法因计算复杂度低、跟踪效果好、实时性强等优点受到了普遍的关注。

传统的 Mean-shift 算法存在一些缺陷,如跟踪窗内除目标像素外,不可避免会引入背景像素,而背景像素的存在会造成目标的定位偏差。另外,虽然 Mean-shift 是一种中心加权算法,中心像素具有较高权值,对小幅遮挡有较强鲁棒性,但是当目标发生部分遮挡或者严重遮挡时,Mean-shift 会失效。

针对以上问题,相关学者提出很多改进的算法,比如基于水平集非对称核函数^[3-4]的算法、基于空间直方图(Spatiogram)^[5-6]的算法、基于 Mean-shift 与粒子滤波相结合^[7-8]的算法以及基于 Mean-shift 与 Kalman 滤波相结合^[9-10]的算法。考虑上述算法的优缺点,本文提出两点新的改进算法:(1)根据初始帧目标和背景在颜色分布上的差异,建立对数似然图(log-likelihood image),筛选出与背景可区分性好的颜色特征进行目标建模,并以同样的方法在后续帧建立候选模型,该方法有效地减小了背景像素的影响,使得目标定位更加准确;(2)针对部分遮挡,将候选区域划分成若干个重叠的子块,分别利用 Mean-shift 对各个子块进行迭代,以与目标区域相应子块最为匹配的子块的所在位置对整个目标重新定位,对于全部遮挡,采用简单的线性预测,估计目标在下一帧可能出现的位置。

2 传统 Mean-shift 算法

Mean-shift 算法是一种基于核密度估计的非参数模式匹配算法。对在 n 维欧氏空间 X 中的有限数据集 $S, x(x \in X)$ 的样本均值定义为:

$$m(x) = \frac{\sum_{s \in S} K(s-x)\omega(s)s}{\sum_{s \in S} K(s-x)\omega(s)}, s \in S, \quad (1)$$

式中: K 为核函数; ω 为样本权值函数, $m(x) - x$ 差值称为 Mean-shift 矢量。Mean-shift 算法反复将数据点朝着 Mean-shift 矢量方向移动直至收敛。

基于 Mean-shift 的跟踪算法将颜色核直方图作为目标特征的描述,目标模型表示为:

$$q_u = C \sum_{i=1}^n k\left(\left\|\frac{x_i - x_0}{h}\right\|^2\right) \delta[b(x_i) - u], \quad (2)$$

式中: u 是直方图的第 u 个分量, $\{x_i\}_{i=1, \dots, n}$ 是跟踪窗内像素的位置, k 是核函数 K 的轮廓函数, δ 是 Kronecker 函数, $b: R^2 \rightarrow \{1, \dots, m\}$ 为像素点到像素特征的映射, C 是归一化常数使得 $\sum q_u = 1$ 。同样的,候选模型表示为:

$$p_u(y) = C_h \sum_{i=1}^n k\left(\left\|\frac{y - x_i}{h}\right\|^2\right) \delta[b(x_i) - u], \quad (3)$$

其中, C_h 是归一化常数使得 $\sum p_u = 1$ 。

目标模型 q_u 与候选模型 $p_u(y)$ 的相似性采用 Bhattacharyya 系数来计算:

$$\rho[p(y), q] = \sum_{u=1}^M \sqrt{p_u(y)q_u}. \quad (4)$$

通过对 Bhattacharyya 系数的 Taylor 展开可以得到 Mean-shift 下一次迭代的目标中心位置为:

$$y_1 = \frac{\sum_{i=1}^n x_i \omega_i g\left(\left\|\frac{y - x_i}{h}\right\|^2\right)}{\sum_{i=1}^n \omega_i g\left(\left\|\frac{y - x_i}{h}\right\|^2\right)}, \quad (5)$$

式中:

$$g(x) = -k'(x), \quad (6)$$

$$\omega_i = \sum_{u=1}^m \sqrt{\frac{q_u}{p_u(y)}} \delta[b(x_i) - u]. \quad (7)$$

3 基于对数似然图的 Mean-shift 算法

传统 Mean-shift 算法采用全局颜色直方图进行建模,所有像素(包括目标像素和背景像素)对直方图分布都产生影响。当目标模型中包含的背景像素较少时,Mean-shift 算法可以获得较好的跟踪效果。但是大量的背景像素会造成目标定

位的偏差。因此,减少或者消除背景像素对目标模型的影响对于目标的准确跟踪有非常重要的作用。

3.1 建立对数似然图

采用中心-周围的方法^[11]对目标和背景像素进行采样。选择一个覆盖目标的像素矩形集来描述目标模型,而选择目标外围的矩形环内像素来描述背景。

设有特征集 f , H_{obj} 和 $H_{\text{bg}}(i)$ 分别为目标和背景中特征值为 i 的像素总个数,定义其概率为 $p(i)$ 和 $q(i)$:

$$p(i) = H_{\text{obj}}(i) / n_{\text{obj}}, \quad (8)$$

$$q(i) = H_{\text{bg}}(i) / n_{\text{bg}}, \quad (9)$$

其中 n_{obj} 和 n_{bg} 分别是目标区域和背景区域的像素个数。对于特征值为 i 的对数似然函数定义为:

$$L(i) = \log \frac{\max(p(i), \epsilon)}{\max(q(i), \epsilon)}, \quad (10)$$

其中, ϵ 是一个极小值(文章中取值为 0.000 1)以避免公式无意义。该非线性的对数似然函数将大部分目标区域像素出现的概率映射为正值,而将大部分背景区域像素出现的概率映射为负值。原



(a) 目标 & 背景 (b) 对数似然图

(a) Object & background (b) "Log-likelihood" image

图 1 对数似然图

Fig. 1 "Log-likelihood" image

图像中每个像素都对应一个对数似然值,称此图为“对数似然图(log-likelihood image)”,为了获得更加可靠的目标像素,对求得的似然图按式(11)进行阈值分割,其中 tho 是决定更可靠像素的阈值,只有对数似然值大于该阈值的像素才被选用作为目标建模。为了消除图像内部空洞对核初始化的不利影响,利用数学形态学对似然图像进行处理,使得算法更加稳健。

$$T(x_u) = \begin{cases} 1, & L(i) > \text{tho} \\ 0, & \text{otherwise} \end{cases}. \quad (11)$$

3.2 基于对数似然图的 Mean-shift 跟踪

式(8)中定义的对数似然值 $L(i)$ 可在负值与

正值的一个很大范围内变化,文中采用 sigmoid 函数将 $L(i)$ 映射到 $(0, 1)$ 。对于特征值 u , 其权重系数计算如下:

$$\lambda_u = \max\left(1 - \frac{1}{1 + \exp(L_u)}, 0.1\right). \quad (12)$$

根据式(2)和式(3),重新定义目标模型和候选模型为:

$$q_u = \frac{1}{C} \sum_{i=1}^n k(\|x_i\|^2) \lambda_u \delta[b(x_i) - u], \quad (13)$$

和

$$p_u(y) = \frac{1}{C_h} \sum_{i=1}^n k\left(\left\|\frac{y - x_i}{h}\right\|^2\right) \lambda_{u_h} \delta[b(x_i) - u], \quad (14)$$

其中 C 和 C_h 分别为归一化常数使得 $\sum q_u = 1$ 和 $\sum p_u = 1$ 。其它计算与传统 Mean-shift 算法一致。

4 遮挡处理

Mean-shift 是一种中心加权算法,中心位置像素权值较高,边缘位置像素权值较低,该算法对目标小幅遮挡有较强的鲁棒性。但是当目标被部分遮挡或者严重遮挡时,Mean-shift 算法会失效。本文提出采用分块方法来处理部分遮挡,采用线性预测处理全部遮挡。

将目标区域划分为若干个重叠的子块,分别用 Mean-shift 对各子块进行迭代,找出匹配度最高(即 Bhattacharyya 系数最大)的子块,并以该子块位置作为目标的最终定位。子块越多,定位越准确,但计算量也就越大;同时,过小的子块划分会丢失目标原有的特征,从而造成定位偏差。文中将目标区域划分为 6 个子块。第 1 个子块为目标整个矩形区域,反映目标的整体特征;第 2 个子块为与第一部分同心但尺寸减小的矩形区域,反映目标的中心特征;第 3 到第 6 个子块分别为目标的上下左右区域,反映目标的上下左右特征,如图 2 所示。

跟踪过程中,找出 6 个子块中最大的 Bhattacharyya 系数,并以该系数作为目标是否严重遮挡的判据。当目标被严重遮挡时, Bhattacharyya

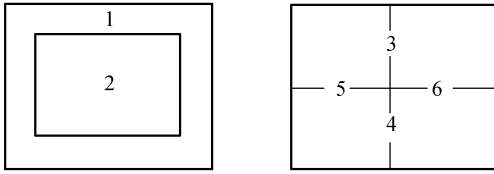


图 2 目标区域分块表示

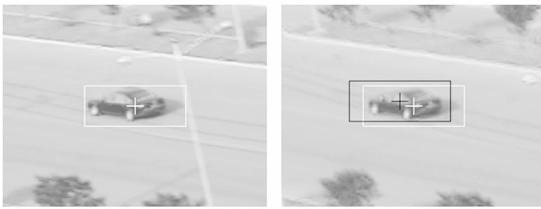
Fig. 2 Fragments of object region

系数减小到某个阈值 T_{occ} , 停止 Mean-shift 迭代, 并假定目标做匀速直线运动, 改用线性预测估计目标下一帧的可能位置。当目标重新出现时, Bhattacharyya 系数上升到某个阈值 T_{app} , 停止线性预测, 重新开启 Mean-shift 迭代。

5 实验结果

根据提出的改进方法, 给出实验结果。

图 3 是汽车跟踪效果图。图中黑色矩形是传统 Mean-shift 跟踪效果图, 白色矩形是改进 Mean-shift 跟踪效果图, 矩形中的十字是目标中心。传统的 Mean-shift 算法中, 背景像素以一定的权值(如式 2 所示)参与为目标定位的计算, 大量的背景会造成定位的偏差, 而改进 Mean-shift 算法通过建立对数似然图, 排除了背景像素的干扰, 实现了目标的精确定位。



(a) 初始帧

(b) 第 18 帧

(a) Frame 1

(b) Frame 18

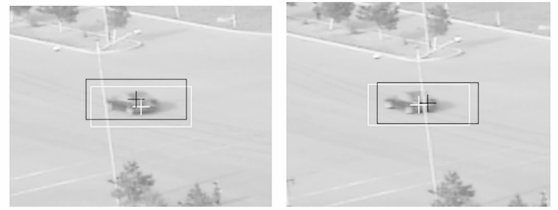


(c) 第 44 帧

(d) 第 70 帧

(c) Frame 44

(d) Frame 70



(e) 第 98 帧

(f) 第 106 帧

(e) Frame 98

(f) Frame 106

图 3 汽车跟踪示例

Fig. 3 Examples of car tracking

图 4 给出了整个跟踪过程中两种算法定位偏差的曲线图。定位偏差 $error_i$ 计算如下:

$$error_i = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2} \quad (15)$$

式中, (x_i, y_i) 为算法计算所得的目标中心位置, (x_c, y_c) 为通过手工逐帧获取的目标中心位置。手工获取目标中心的方法带有很强的主观性, 会带来一定的误差, 但整体上不会对比较结果产生太大的影响。图 4 表明改进的 Mean-shift 算法具有更高的跟踪精度。

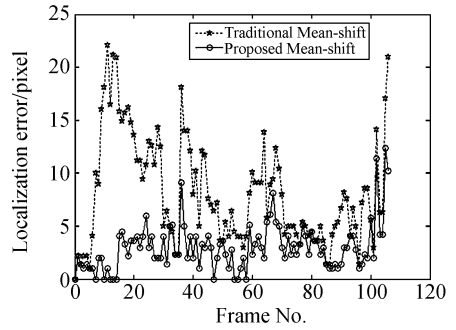


图 4 定位偏差

Fig. 4 Localization errors

图 5 是汽车穿越遮挡物的跟踪效果图。当汽车从右侧进入遮挡区域时, 如第 35 帧, 右边子块 Bhattacharyya 系数最高, 因此以右边子块位置作为目标的最最终定位; 当汽车在遮挡区内时, 如第 43 帧和第 51 帧, 所有子块 Bhattacharyya 系数均低于阈值 T_{occ} , 因此该区间内采用线性预测, 估计目标下一帧的可能位置; 当汽车逐渐离开遮挡区域时, 如第 55 帧, 左边子块 Bhattacharyya 系数最高, 且高于阈值 T_{app} , 目标被重新捕获。图 6 是跟踪过程中由每一帧各子块最大的 Bhattacharyya 系数组成的曲线图。

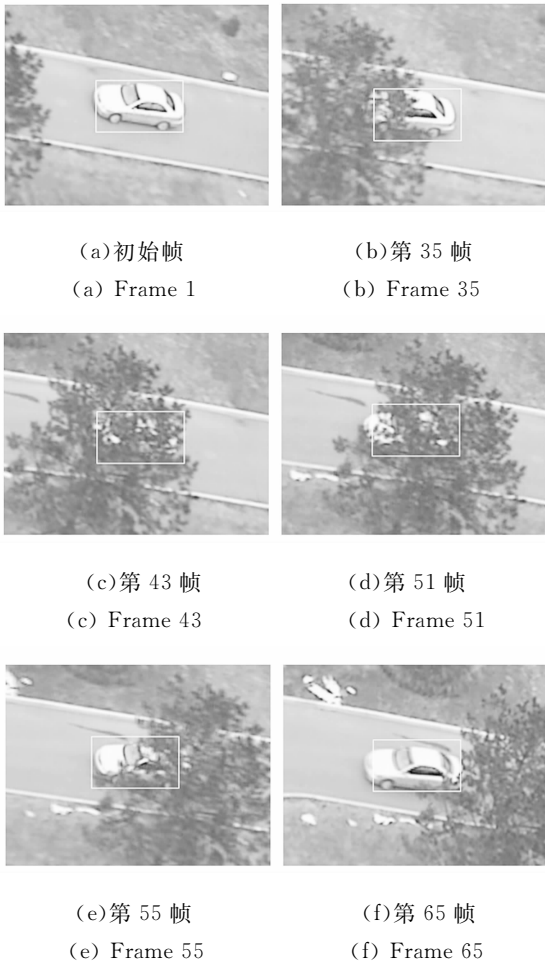


图 5 汽车穿越遮挡示例

Fig. 5 Car tracking through occlusion

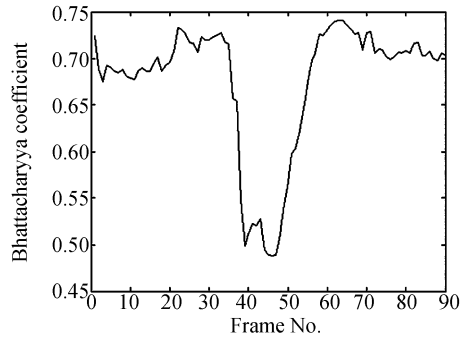


图 6 图 5 中的 Bhattacharyya 系数

Fig. 6 Bhattacharyya coefficient for Fig. 5

6 结 论

本文针对传统 Mean-shift 算法存在的两点缺陷,提出了基于对数似然图的目标跟踪和分块的目标跟踪,很好地解决了跟踪过程中的定位偏差和部分遮挡问题。当目标被严重遮挡时,采用线性预测进行目标跟踪。实验结果表明,改进的 Mean-shift 算法能准确地对目标进行跟踪,对于部分遮挡和严重遮挡都有较强的鲁棒性。

参考文献:

- [1] CHENG Y. Mean shift, mode seeking, and clustering [J]. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1995,17(8):790-799.
 - [2] COMANICIU D, RAMESH V, MEER P. Kernel-based object tracking [J]. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2003,25(5):564-577.
 - [3] YIMAZ A. Object tracking by asymmetric kernel mean-shift with automatic scale and orientation selection [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, 2007:1-6.
 - [4] CHEN X P, YU SH SH, MA ZH L. An improved mean-shift algorithm for object tracking [C]. *Proceedings of the 7th World Congress on Intelligent Control and Automation*, IEEE, 2008:5111-5114.
 - [5] BIRCHFIELD S T, RANGARAJAN S. Spatio-grams versus histograms for region based tracking [C]. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE, 2005:1158-1163.
 - [6] O'CONAIRE C, O'CONNOR N, SMEATON A F. An improved spatiogram similarity measure for robust object localization [C]. *ICASSP*, IEEE, 2007:1069-1072.
 - [7] NUMMIARO K, KOLLER-MEIER E, VAN G L. Color features for tracking non-rigid object [J]. *Special Issue on Vision Surveillance*, 2003,29(3):345-355.
 - [8] 孟勃,朱明. MSMC 跟踪算法在目标跟踪中的应用 [J]. *光学 精密工程*, 2008,16(1):122-127.
- MENG B, ZHU M. Application of MSMC algorithm to visual tracking [J]. *Opt. Precision Eng.*,

2008,16(1):122-127. (in Chinese)

- [9] 孙中森,孙俊喜,宋建中,等.一种抗遮挡的运动目标跟踪算法[J].光学精密工程,2007,15(2):268-272.

SUN ZH S, SUN J X, SONG J ZH, *et al.*. Anti-occlusion arithmetic for moving object tracking [J]. *Opt. Precision Eng.*, 2007, 15 (2): 268-272. (in Chinese)

- [10] WANG ZH Q, FAN Y F, ZHANG G L, *et al.*. Robust face tracking algorithm with occlusions [J]. *SPIE*, 2007,67861:67861X1-67861X10.

- [11] COLLINS R T, LIU Y X. On-line selection of discriminative tracking features [C]. *Proceedings of the Ninth IEEE International Conference on Computer Vision, IEEE*,2003,346-352.

作者简介:



薛 陈(1983—),男,四川成都人,博士研究生,主要从事图像处理、目标跟踪等方面的研究。E-mail: achen5225@sina.com

陈爱华(1981—),女,山东泰安人,博士研究生,主要从事图像采集与处理、目标跟踪、模式识别等方面的研究。E-mail: chen_1216@163.com

导师简介:

朱 明(1964—),男,江西南昌人,研究员,博士生导师,主要从事图像处理、电视跟踪和自动目标识别技术的研究。E-mail: zhu_mingca@163.com

●下期预告

光切法形貌测量中光条中心的亚像素提取

贾倩倩,王伯雄,罗秀芝

(清华大学精密仪器与机械学系精密测试技术及仪器国家重点实验室,北京 100084)

为了在光切法三维形貌测量中快速准确提取光条的中心位置,保证测量精度,研究了一种光条中心的亚像素提取方法。将图像的灰度视为关于坐标的二维函数,在 CCD 拍摄的激光光条图像中求取灰度的偏导数和梯度,图像中每一点的梯度代表此点的灰度变化方向,将此方向简化为光条上的点的法线方向。对梯度的大小设定阈值并结合该点的灰度信息,可提取图像灰度变化最激烈的部分,即光条饱和带的像素位置。对光条饱和带的每个像素,按照不重复计算的原则在梯度方向上和规定邻域内求灰度中心,可以得到精确到亚像素级的光条中心位置。此方法将光条曲线的法线方向简化为灰度的梯度方向,省去了求海赛矩阵以确定法向的步骤。实验证明,同等条件下,方法的计算耗时为常规亚像素提取方法的 10%。利用该算法进行实际测量,200 mm×200 mm 测量范围下,误差(3 σ)为 0.057 mm,满足了测量对精确性和实时性的要求。