

基于改进背景加权的 Mean-Shift 目标跟踪算法

刘翔, 侯志强, 余旺盛, 李武, 黄安奇

(空军工程大学信息与导航学院, 西安 710077)

摘要: 针对采用核加权直方图的方法计算目标模板与候选区域目标特征无法实现对运动目标的准确定位这一问题, 提出了一种利用改进背景加权增强直方图显著性的鲁棒 Mean-Shift 跟踪算法。在传统 Mean-Shift 的框架下, 通过计算目标和背景特征直方图 bin 值, 得到两者特征显著性大小, 将其代入传统相似性度量中, 定义新的背景加权系数, 进而更好地提高目标与背景的区分度, 减少背景信息对目标定位的干扰。通过算法改进前后的实验结果以及跟踪误差和正确跟踪率的比较发现, 跟踪效果得到明显改善。

关键词: 目标跟踪; Mean-Shift 算法; 特征显著性; 特征相似性

中图分类号: TP391 **文献标志码:** A **文章编号:** 1671-637X(2014)10-0001-04

Mean-Shift Object Tracking Algorithm Based on Improved Weighted Background

LIU Xiang, HOU Zhi-qiang, YU Wang-sheng, LI Wu, HUANG An-qi

(Institute of Information and Navigation, Air Force Engineering University, Xi'an 710077, China)

Abstract: Considering that it is difficult to locate the moving objects accurately from the clutter background by using the weighted kernel based color histogram to compute the feature of object template and candidate regions, we proposed a robust Mean-Shift object tracking algorithm based on improved weighted background to enhance histogram saliency. The saliencies of the object and of the background were calculated out from their bin of histogram, which was incorporated into the traditional similarity measurement for defining an improved weighted background coefficient. Therefore, the discriminability of the object from the background was increased, and the effect of background information on object locating was reduced. Experiment result shows that the tracking effect is improved apparently.

Key words: target tracking; Mean-Shift algorithm; feature saliency; feature similarity

0 引言

对视频序列中感兴趣的运动目标进行有效跟踪, 是计算机视觉技术研究的一个重要部分, 广泛应用于军事及民用领域。1995年, 文献[1]将 Mean-Shift 算法^[2]运用到模式分类与图像分割中, 此后, Mean-Shift 被广泛用于模式分类、图像分割以及目标跟踪等方面。文献[3-4]提出了一种加权直方图来描述目标特征, 分别建立目标模型和候选目标模型, 通过 Bhattacharyya 系数的大小判别两种模型的匹配度, 选择匹配度最好的候

选目标模型作为下一帧目标模型, 获得了良好的跟踪效果。

然而在实际应用中, 若目标被部分遮挡或有背景因素干扰, 则 Mean-Shift 算法的跟踪精度会降低, 最终导致目标跟踪丢失。文献[5]提出利用目标的显著性权值计算颜色直方图栅权值, 从而通过加权方法对复杂背景中的目标进行跟踪。但由于其只考虑初始帧目标显著性权值, 在跟踪过程中随着目标颜色直方图发生变化, 跟踪效果有所下降。文献[6]证明了 Comaniciu 的背景加权策略并没有提高目标与背景的区分度, 并进一步提出了仅在目标直方图上引入背景加权的方法, 最终提高了目标跟踪的稳健性。本文从 Mean-Shift 算法的数学模型入手, 通过对传统 Bhattacharyya 相似性度量加入目标实时更新的特征显著性权值, 从而更好地描述目标与背景的显著性区别, 对背景进行有效抑制, 提

收稿日期: 2013-12-02

修回日期: 2013-12-31

基金项目: 国家自然科学基金(61175029, 61473309); 陕西省自然科学基金(2011JM8015)

作者简介: 刘翔(1990—), 男, 陕西西安人, 硕士生, 研究方向为视觉目标跟踪。

高跟踪的准确性。

1 传统 Mean-Shift 算法

1.1 算法原理

对于 Mean-Shift 目标跟踪,目标模型和候选目标模型往往是对目标区域的图像特征进行统计而得到的离散概率密度函数,即归一化的直方图向量^[4]。若目标区域中心的图像坐标为 x_0 ,假设目标区域中有 n 个像素,用 $\{x_i, i=1, 2, \dots, n\}$ 表示,特征值 bin 的个数为 m ,则目标模型的概率密度为

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

式中: C 为使 $\sum_{u=1}^m q_u = 1$ 的归一化常数; $k(x)$ 为一个各向同性核函数,带宽为 h ; δ 为 Kronecker delta 函数。相似地,假设候选目标区域中心的图像坐标为 y ,则位于位置 y 的候选目标模型的概率分布为

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

式中, C_h 为使目标候选模型归一化的常数因子。一般采用 Bhattacharyya 系数 ρ 表示目标模型和候选目标模型之间的相似性程度,即

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

Mean-Shift 目标跟踪算法的目的就是找到位置 y , 即使式(3)取最大值时的候选目标模型的位置。假定目标在当前帧的初始位置为 y_0 ,通过对 Bhattacharyya 系数的 Taylor 展开可以得到 Mean-Shift 算法下一次迭代的目标中心位置,为

$$y_{\text{new}} = \frac{\sum_{i=1}^n g\left(\left\|\frac{y_0 - x_i}{h}\right\|^2\right) w_i x_i}{\sum_{i=1}^n g\left(\left\|\frac{y_0 - x_i}{h}\right\|^2\right) w_i} \quad (4)$$

式中: $g(x)$ 为核函数 $k(x)$ 的负倒数; $w_i = \sum_{u=1}^m \sqrt{\hat{q}_u / \hat{p}_u(y_0)} \cdot \delta(b(x_i) - u)$ 。当 $\|y_{\text{new}} - y_0\| < \varepsilon$ 时,停止本帧跟踪,该点就为当前帧目标位置;否则设定 $y_0 \leftarrow y_{\text{new}}$,继续迭代。

1.2 传统 Mean-Shift 算法的不足

从以上对算法的分析可以看出,在描述目标方面, Mean-Shift 算法主要是根据目标颜色直方图进行跟踪,尽管核函数加权可以在一定程度上减弱背景干扰,增加算法的跟踪鲁棒性,但单一的颜色特征直方图在复杂场景中不足以彻底分辨目标,特别是背景和颜色分布比较接近时,往往会导致目标丢失。另外,在针对形状不规范的目标时,背景像素往往被赋予较大权

值,从而导致跟踪效果不佳。

2 本文改进算法

2.1 显著性加权

在初始跟踪时,目标区域往往会包含背景信息,而目标本身也有可能含有与背景相同的颜色特征,这些都会导致在计算目标模板时包含属于背景的颜色特征。

视觉显著性是由多种视觉敏感特征引起的一种局部反差,在各类视频图像中,与背景差异越大的目标局部特征对于鲁棒跟踪的贡献越大,反之越小^[5,7]。针对传统 Mean-Shift 算法的不足,本文提出了一种将背景直方图的显著性加权引入到传统的 Bhattacharyya 系数相似性度量中的改进方法,更多地关注目标的显著性区域,对显著性高的区域赋予较大的权值,显著性低的区域赋予较小的权值,将背景区域的权值赋为零,从而达到稳健的跟踪效果。

定义 $H_{o,u}$ 和 $H_{b,u}$ 分别表示目标和背景特征直方图第 u 个 bin 的值, $p_f(O_{\text{obj}} | u)$ 为特征直方图第 u 个 bin 中的目标模型特征直方图的概率,即目标的显著性大小,为

$$p_f(O_{\text{obj}} | u) = \begin{cases} (H_{o,u} - H_{b,u}) / (H_{o,u} + n_{\text{eps}}), & H_{o,u} \geq H_{b,u} \\ 0, & H_{o,u} < H_{b,u} \end{cases} \quad (5)$$

式中: n_{eps} 表示任意小的非零常数,防止上式出现无意义的情况;为方便起见, $H_{o,u}$ 为目标模板的特征直方图, $H_{b,u}$ 则随着每一次 Mean-Shift 算法的迭代而改变。把 $p_f(O_{\text{obj}} | u)$ 代入传统的相似性度量式(3)中

$$\tilde{\rho}(y) = \tilde{\rho}[\rho(y), q] = \sum_y \sqrt{p_f(O_{\text{obj}} | u)} \sum_{u=1}^m \sqrt{\hat{p}_u(y) \hat{q}_u} \quad (6)$$

$\tilde{\rho}$ 即本文提出的将显著性权值与巴氏系数相结合的相似性度量方法。定义新的权系数为 $\tilde{w}(x_i) = \sqrt{p_f(O_{\text{obj}} | u)} w(x_i)$,式(4)可以写成

$$y_{\text{new}} = \frac{\sum_{i=1}^n g\left(\left\|\frac{y_0 - x_i}{h}\right\|^2\right) \tilde{w}_i x_i}{\sum_{i=1}^n g\left(\left\|\frac{y_0 - x_i}{h}\right\|^2\right) \tilde{w}_i} \quad (7)$$

如以上分析所述,跟踪过程将更加倾向于那些目标中与背景差别大的特征,也即跟踪在复杂背景中将更加鲁棒。

在本文中,设背景区域是一个为 3 倍目标候选矩形窗大小的矩形区域(在图像边界处进行了边界处理),暂不考虑目标尺度变化情况,因此将跟踪窗设为一个固定值。

2.2 算法步骤

1) 手动选取初始帧的跟踪目标,根据初始跟踪位置 y_0 确定目标和背景区域,并建立目标模型直方图;

2) 将式(5)中计算目标的显著性权值引入到传统 Bhattacharyya 系数的相似性度量中,得到 $\tilde{\rho}$,从而计算出新的权系数 \tilde{w}_i ;

3) 根据式(7)计算偏移向量,找到目标下一次迭代的新位置 y_{new} ;

4) 如果 $\|y_{new} - y_0\| < \varepsilon$,停止本帧跟踪,当前位置 y_{new} 就是目标跟踪的下一位置;

5) 更新目标模板,得到新的显著性权值;

6) 返回 2) 进行下一帧跟踪。

3 实验及结果分析

为验证算法的有效性,通过多段视频序列对传统 Mean-Shift 算法、CBWH^[5] 算法、BCKBT^[8] 算法以及本文提出的改进算法进行分析对比。实验使用 Matlab 2009b 仿真软件进行仿真,均采用 RGB 颜色模型下,量化空间为 $16 \times 16 \times 16$ 的视频序列,目标初始位置由手动给定,本文首先从视觉上对比了 4 种算法,而后为进一步对比,将跟踪结果与真实结果(人工标注得到)进行了定量分析。设第 $j=1, 2, \dots, N$ 帧跟踪框和实际目标的位置为 $D_j(x_j, y_j)$ 、 $D_1(x_{j_0}, y_{j_0})$,则跟踪框和手动标注框的面积重叠比为

$$S_o = \frac{D_1 \cap D_2}{D_1 \cup D_2} \quad (8)$$

可得 S_o 的取值范围为 $[0, 1]$, S_o 取值越大表明跟踪准确度越高。这里认为, $S_o > 0.5$ 时,此帧为有效跟踪。设正确跟踪率为 A_t ,有效跟踪帧数为 F_s ,视频总帧数为 N ,则

$$A_t = F_s / N \quad (9)$$

首先,针对场景不变的视频序列 Market 和 Right Hand 进行实验分析,结果如图 1 和图 2 所示。实验 1 中,商场中行走的顾客,其衣服颜色和背景颜色相近。当目标在复杂背景中时,目标周围的背景对跟踪影响很大,导致传统 Mean-Shift 算法极易丢失目标;在视频序列的最后阶段,目标与背景中物体相似时,CBWH 和 BCKBT 算法易丢失目标;而本文提出的方法可以持续地跟踪目标,效果良好。

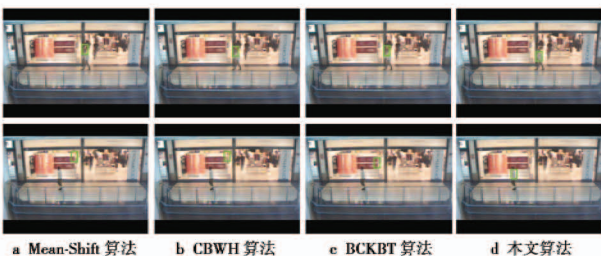


图 1 Market 序列 73、108 帧跟踪结果

Fig.1 Tracking result of two frames from Market sequence

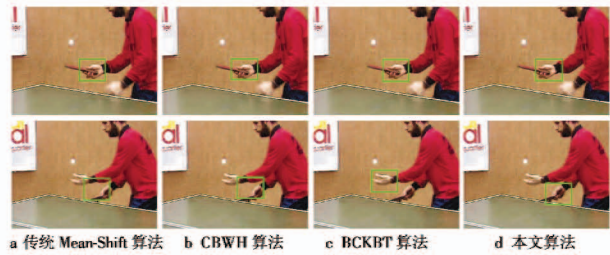


图 2 Right Hand 序列 44、52 帧跟踪结果

Fig.2 Tracking result of two frames from Right Hand sequence

实验 2 中,选取 Right Hand 为跟踪目标,在左手的交叉干扰下,传统 Mean-Shift 算法、BCKBT 算法的抗干扰能力较差,算法发生漂移并最终导致跟踪错误。单一的颜色特征很容易受背景中相似颜色的干扰,BCKBT 算法采用非归一化的背景直方图对目标直方图进行加权,由于所选的背景区域的大小会对加权结果造成一定影响,所以在特征相似时,出现了跟踪偏差。CBWH 算法和本文算法采用归一化背景直方图对目标直方图进行显著性加权,保持了目标和背景的公平性,适应能力也更强。

为测试算法适应跟踪场景变化的能力,分别从 Takeoff 序列和 Airplane 序列截取了一段视频,实验结果如图 3 和图 4 所示。

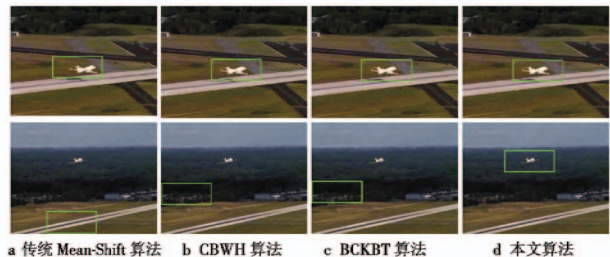


图 3 Takeoff 序列 111、260 帧跟踪结果

Fig.3 Tracking result of two frames from Takeoff sequence

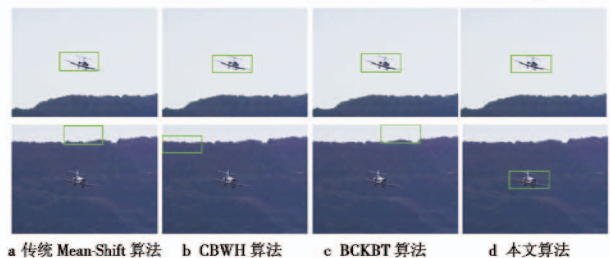


图 4 Airplane 序列 12、145 帧跟踪结果

Fig.4 Tracking result of two frames from Airplane sequence

由于传统的 Mean-Shift 算法、CBWH 算法、BCKBT 算法对目标特征的描述能力有限,或者同时忽略背景信息更新,或者同时更新目标和背景信息的模板更新方式并不能真实反映目标和背景的实际情况,本文中的改进算法依据不同的相似性度量分别对目标和背景模板进行必要的加权,在场景变化的目标跟踪过程中,不仅降低累积误差,而且有效保证了目标特征的显著

性。实验3和实验4中视频序列跟踪场景变化复杂,在视频序列前部分,目标与背景区分度高,跟踪框尺度与目标大小较吻合。跟踪框中背景信息不断增多,进而“淹没”了目标主体的信息,传统 Mean-Shift 算法、CBWH 与 BCKBT 算法会很快漂移到背景中去。本文提出的改进方法突出了目标的显著性信息,增强了算法的鲁棒性。

图5利用跟踪框中心位置与目标实际位置间的欧氏距离来对比本文提出的算法与其他3种算法的跟踪准确度,相比其他3种跟踪算法,本文算法具有较高的目标跟踪精度。为了进一步评估改进算法的有效性,通过各算法的跟踪误差、正确跟踪率、算法跟踪时间等方面,深入分析了算法的跟踪准确性。

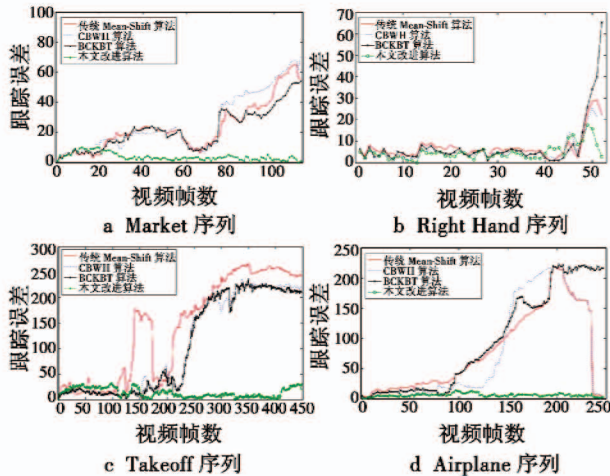


图5 4组序列的跟踪误差比较

Fig. 5 Tracking errors of the four sequences

各算法性能比较如表1所示。

表1 算法性能比较

Table 1 Performance contrast of different algorithms

实验	算法	E_x	E_y	$A_t/\%$	T/s
Market	传统 Mean-Shift 算法	20.89	8.86	26.8	0.244
	CBWH 算法	23.68	9.31	33.1	0.229
	BCKBT 算法	18.41	8.91	36.6	0.241
	本文算法	2.92	1.89	99.1	0.219
Right Hand	传统 Mean-Shift 算法	2.56	6.33	96.2	0.471
	CBWH 算法	3.55	3.90	98.1	0.460
	BCKBT 算法	4.25	5.48	92.3	0.457
	本文算法	3.19	3.44	100	0.409
Takeoff	传统 Mean-Shift 算法	126.78	46.90	35.4	2.203
	CBWH 算法	85.04	55.54	49.3	2.516
	BCKBT 算法	82.04	57.79	44.6	2.630
	本文算法	5.36	11.64	100	2.436
Airplane	传统 Mean-Shift 算法	56.19	40.94	66.9	1.592
	CBWH 算法	40.98	62.25	72.7	1.422
	BCKBT 算法	75.14	50.11	65.5	1.271
	本文算法	3.44	4.70	100	1.388

表1中: E_x, E_y 分别表示中心坐标在水平和垂直方

向的平均绝对误差(像素); A_t 为目标正确跟踪率; T 为跟踪平均时间。从表1中数据比较可以看出,本文改进算法的跟踪误差明显低于其他3种算法,特别是当运动目标的运动场景发生变化时。因为4种算法在复杂性上基本相同,所以在跟踪实时性上,4种算法没有太大差异。综上所述,本文算法不仅有较高的跟踪精度,而且跟踪过程相对稳定。

4 结论

针对传统 Mean-Shift 算法中核函数加权直方图对目标的描述能力有限的缺陷,本文将背景直方图显著性加权加入目标描述中,利用特征的显著性和相似性对背景与目标的区分度进行量化分析,通过将目标特征和背景的显著性大小作为权值进一步提高目标和背景的区分度,增强了算法的抗干扰性。最后通过实验验证了本文算法的有效性和跟踪准确性。

参考文献

- [1] CHENG Y. Mean shift, mode seeking, and clustering[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995, 17(8):790-799.
- [2] FUKUNAGA K, HOSTELLER L D. The estimation of the gradient of a density function, with applications in pattern recognition[J]. IEEE Transactions on Information Theory, 1975, 21(1):32-40.
- [3] COMANICIU D, RAMESH V, MEER P. Real-time tracking of non-rigid objects using mean shift[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2000: 142-149.
- [4] COMANICIU D, RAMESH V, MEER P. Kernel-based object tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(5):564-577.
- [5] 张恒,李由,李立春,等.基于显著性加权的 mean shift 跟踪方法[J].光学技术,2008,34(3):404-407. ZHANG H, LI Y, LI L C, et al. A new significance weighted mean shift tracking method[J]. Optical Technique, 2008, 34(3):404-407.
- [6] NING J, ZHANG L, ZHANG D, et al. Robust mean-shift tracking with corrected background-weighted histogram[J]. IET Computer Vision, 2012, 6(1):62-69.
- [7] 张巧荣,景丽,肖会敏,等.利用视觉显著性的图像分割方法[J].中国图象图形学报,2011,16(5):767-772. ZHANG Q R, JING L, XIAO H M, et al. Image segmentation based on visual saliency[J]. Journal of Image and Graphics, 2011, 16(5):767-772.
- [8] LIU R L, JING Z L. Robust kernel-based tracking algorithm with background contrasting[J]. Chinese Optics Letters, 2012, 10(2):1-3.