

引用格式:胡穹,梅雪,卞佳丽.使用自适应背景约束改进的颜色属性跟踪算法[J].电光与控制,2018,25(10):33-38,51. HU Q, MEI X, BIAN J L. An improved color attribute tracker with adaptive context constraints[J]. Electronics Optics & Control, 2018, 25(10):33-38, 51.

使用自适应背景约束改进的颜色属性跟踪算法

胡穹, 梅雪, 卞佳丽
(南京工业大学电气工程与控制科学学院, 南京 211800)

摘要:颜色属性跟踪算法使用的核脊回归分类器对于背景信息的利用过少,导致算法在目标快速运动、局部遮挡和背景相似物干扰等情形下容易发生漂移。针对此问题,首先,通过计算响应矩阵峰值旁瓣比确定干扰峰所在位置;然后,在核脊回归分类器上引入对应的背景约束项来加强分类器对背景信息的利用;最后,使用构建尺度空间滤波器的方法实现对多尺度变化的支持。在 Visual Tracker Benchmark 数据集上的对比实验结果表明,该方法在不过多损失算法速度的前提下,能够有较高的跟踪精度。

关键词:目标跟踪;颜色属性跟踪;自适应背景约束;峰值旁瓣比;核脊回归

中图分类号: TP391.41 **文献标志码:** A **doi:**10.3969/j.issn.1671-637X.2018.10.007

An Improved Color Attribute Tracker with Adaptive Context Constraints

HU Qiong, MEI Xue, BIAN Jia-li

(College of Electrical Engineering and Control Science, Nanjing Technique University, Nanjing 211800, China)

Abstract: The kernel ridge regression classifier used in adaptive color attribute tracker makes little use of information about the context, which may easily result in drifting in cases of fast target motion, partial occlusion or background clutter. To solve the problem, the interference peak position is determined at first by calculating the peak side lobe ratio of the response matrix. Then, the corresponding context constraints are introduced on the kernel ridge regression classifier to enhance the background utilization. Finally, a scale space filter is established for supporting multi-scale variation. The experimental results based on the Visual Tracker Benchmark dataset show that this method can achieve high tracking accuracy with a mild loss of frame rates.

Key words: target tracking; color attribute tracking; adaptive context constraint; peak side lobe ratio; kernel ridge regression

0 引言

目标跟踪是近年来计算机视觉领域一个热门的研究课题,已广泛地应用于智能人机交互、智能监控、社会安保、居家安防、交通控制、工业生产、军事侦测等诸多领域,也得到了越来越多学者和研究人员的关注^[1]。尽管近年来涌现出了一大批新颖的算法,但由于在实际测试和应用中存在着复杂背景相似物干扰、局部遮挡、目标姿态变化等诸多因素的干扰,仍然没有一种单一的方法能够成功适应所有场景^[2]。文献[3]首次将相关滤波技术应用用于目标跟踪领域,该方法将图像卷

积操作通过快速傅里叶变换转换成频域中的点乘操作,极大地减少了计算量,保证了算法的高效性,由于该算法有极高的运算效率和较强的拓展性,因而迅速成为目标跟踪领域的热点;文献[4]在此基础上引入脊回归作为分类器,使用循环移位样本近似替代密集采样训练样本,提出循环结构检测(CSK)算法,但该算法由于仅使用单通道的灰度特征,因此跟踪鲁棒性还不够好;文献[5]则在其基础上将单通道的灰度特征拓展为多通道的HOG特征,提高了算法的鲁棒性,但由于算法并未使用颜色特征,因此当目标出现快速移动或者快速形变时,容易发生跟踪漂移;而文献[6]则尝试使用了多通道的颜色属性特征拓展CSK算法,并取得了较好的效果,但对于在局部遮挡、目标快速运动和背景相似物干扰等情形下的鲁棒性仍然不够;文献[7]则在文献[5]的基础上构造了图像金字塔,实现了

收稿日期:2017-11-15 修回日期:2017-12-07

基金项目:江苏省研究生培养创新工程项目(KYCX17_0925);江苏省“六大人才高峰”项目(XXRJ-012)

作者简介:胡穹(1992—),男,湖北红安人,硕士生,研究方向为图像处理、运动目标跟踪。

算法对目标尺度的支持,并使用 HOG 特征和颜色特征作为组合特征,增强了目标特征的表现力,但由于算法是在 7 个尺度图像上分别进行检测,且要多次进行复合特征的提取计算,所以运算速度较慢且对尺度支持不够精准;而文献[8-9]则通过在线性核脊回归跟踪算法的基础上额外构造一个尺度空间滤波器实现了算法对多尺度较为精确的估计,但其仅使用单一的 HOG 特征且用线性核函数代替高斯核函数,故在目标快速运动和发生形变的情况下跟踪效果也不够好。与此同时,相关滤波跟踪的方法为了保持算法的高速运行,其本身的检测区域一般设置较小,并对训练样本进行了加窗操作,故而对背景信息的利用很少,因而对目标快速运动、遮挡、背景相似物干扰等情况容易造成核滤波响应出现多峰的情况,进而影响跟踪效果^[10-12]。

针对以上问题,为了在保持算法具有较高跟踪效率的同时,对于目标快速运动、遮挡、背景相似物干扰等情况容易造成核滤波多峰响应进行抑制,本文在颜色属性跟踪(CN)算法^[6]的基础上,通过计算响应矩阵上的峰值旁瓣比(PSR)来判断是否出现干扰和干扰峰所在的位置,进而在核化相关滤波器中引入对应的背景约束项来加强分类器对背景信息的学习利用,从而提高算法的鲁棒性。

1 算法原理

1.1 颜色属性跟踪算法

颜色属性跟踪算法是在 CSK 算法的基础上,将单通道的灰度特征拓展到多通道的颜色属性特征^[13-14],从而增强算法的鲁棒性,并通过使用 PCA 主成分分析法将原 11 维颜色特征降维到 2 维来降低拓展颜色特征对算法速度的影响。该算法所采用的核脊回归分类器,是利用训练样本找到一个使样本 \mathbf{X} 对回归目标 \mathbf{y} 的平方差最小的函数 $f(\mathbf{z}) = \mathbf{w}^T \mathbf{z}$,则该最优化问题的解等价于

$$\min_{\mathbf{w}} \sum \|(X\mathbf{w} - \mathbf{y})\|_2^2 + \lambda \|\mathbf{w}\|_2^2 \quad (1)$$

式中: \mathbf{X} 是原图像经特征提取得到的特征图像进行循环平移之后的所有图像块构成的循环样本矩阵; λ 为正则化参数; $\|\mathbf{w}\|_2^2$ 是正则化项。可得到脊回归解为

$$\mathbf{w} = (\mathbf{X}^T \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^T \mathbf{y} \quad (2)$$

根据矩阵求逆引理,可得其对偶解为

$$\mathbf{w} = \mathbf{X}^T (\mathbf{X} \mathbf{X}^T + \lambda \mathbf{I})^{-1} \mathbf{y} \quad (3)$$

将 \mathbf{w} 映射到高维空间用多项式 $\mathbf{w} = \sum_i \alpha_i \mathbf{x}_i$ ^[15] 来表示,同时引入核方法将内积 $\mathbf{X} \mathbf{X}^T$ 用核矩阵 \mathbf{K} 代替,则核脊回归的解可化为

$$\boldsymbol{\alpha} = (\mathbf{K} + \lambda \mathbf{I})^{-1} \mathbf{y} \quad (4)$$

式中, \mathbf{K} 是由循环移位样本生成的核矩阵,文献[4]证

明了 \mathbf{K} 也是循环矩阵,利用循环矩阵的对角化性质,可以求得非线性分类器

$$\hat{\boldsymbol{\alpha}} = \frac{\hat{\mathbf{y}}}{\hat{\mathbf{k}}^{xx} + \lambda} \quad (5)$$

式中: $\hat{\mathbf{y}}$ 表示对 \mathbf{Y} 进行傅里叶变换; $\hat{\mathbf{k}}^{xx}$ 表示样本 \mathbf{X} 的核自相关矩阵。

在检测阶段,对于新输入的采样图像片 Z ,利用非线性分类器 $\boldsymbol{\alpha}$,可得出其分类器响应

$$\hat{f}(\mathbf{z}) = \hat{\mathbf{z}} \cdot \hat{\mathbf{w}} = \hat{\mathbf{z}} \cdot \hat{\mathbf{x}}_0^* \cdot \hat{\boldsymbol{\alpha}} = \hat{\mathbf{k}}^{zx} \cdot \hat{\boldsymbol{\alpha}} \quad (6)$$

式中: $\hat{\mathbf{x}}_0^*$ 是目标模型的共轭转置; $\hat{\mathbf{k}}^{zx}$ 表示采样图像片 Z 与目标模型的互相关核矩阵; \cdot 表示点乘。对响应结果矩阵 $\hat{f}(\mathbf{z})$ 进行傅里叶逆变换归一化之后,其元素值表示目标位于该元素位置的概率,求得响应矩阵中最大值的位置即求得输入图像 Z 中目标所在的位置。模型 $\hat{\boldsymbol{\alpha}}$ 和 $\hat{\mathbf{x}}_0$ 的更新策略是线性加权更新

$$\hat{\boldsymbol{\alpha}}_t = (1 - \eta) \hat{\boldsymbol{\alpha}}_{t-1} + \eta \hat{\boldsymbol{\alpha}} \quad (7)$$

$$\hat{\mathbf{x}}_{0,t} = (1 - \eta) \hat{\mathbf{x}}_{0,t-1} + \zeta \hat{\mathbf{x}}_0 \quad (8)$$

式中: $\hat{\boldsymbol{\alpha}}$ 和 $\hat{\boldsymbol{\alpha}}_{t-1}$ 分别是当前帧和上一帧的分类器模型的傅里叶变换; $\hat{\mathbf{x}}_0$ 和 $\hat{\mathbf{x}}_{0,t-1}$ 分别是本帧和上一帧的目标模型的傅里叶变换; η 和 ζ 是学习参数。

1.2 使用自适应背景约束改进

为了保持运算速度,常规的核相关滤波类的算法设置的搜索区域较小,一般是上一帧目标大小的 1~2 倍。同时,为了提高循环位移样本的真实性,算法会对取到的样本进行加余弦窗操作。以上两点会过滤掉大量分类器本来需要学习的背景信息,导致分类器判别能力下降。而引入背景约束项,使其在核相关滤波器的损失函数中回归到 0,增加分类器对目标周围背景块的学习,则可以提高分类器的判别能力。但过多的引入背景约束项会极大减缓跟踪算法的速度,因此通过检测响应矩阵中的峰值旁瓣比(PSR)值来确定是否出现干扰,通过计算极大值和次极大值之间的向量角来确定干扰物出现的背景块。PSR 的定义如下

$$P_{\text{PSR}} = \frac{g_{\text{max}} - \mu_{s1}}{\sigma_{s1}} \quad (9)$$

式中: g_{max} 是响应矩阵中的极大峰值; μ_{s1} 和 σ_{s1} 是旁瓣的均值和标准差。当 PSR 满足下式时,则认定出现多峰响应,即

$$P_t < P_{\text{mean}} \quad (10)$$

式中: P_t 表示当前帧的响应矩阵的 PSR 值; P_{mean} 表示 PSR 的平均值。

出现多峰响应之后,通过极值检测可以在响应矩阵中确定最大峰值的坐标 (m_0, n_0) 和次极大峰值的坐标 (m_1, n_1) 。然后取以 (m_1, n_1) 为中心、大小与检测目标大小相同的背景块作为背景约束块。

之后引入对应的背景约束项,原优化问题最优解式(1)变为

$$\min_{\mathbf{w}} \sum \| \mathbf{X}_0 \mathbf{w} - \mathbf{y} \|^2 + \lambda_1 \| \mathbf{w} \|^2 + \lambda_2 \| \mathbf{X}_B \mathbf{w} \|^2 \quad (11)$$

式中: \mathbf{X}_0 代表原目标样本进行循环位移之后构成的训练样本的循环位移矩阵; \mathbf{X}_B 表示对应背景块经过循环位移之后形成的循环矩阵。

将背景块 \mathbf{X}_B 的回归目标定为 0,则式(11)可以统一为

$$f(\mathbf{w}, \mathbf{A}) = \| \mathbf{A} \mathbf{w} - \bar{\mathbf{y}} \|^2 + \lambda_1 \| \mathbf{w} \|^2 \quad (12)$$

$$\text{式中: } \mathbf{A} = \begin{bmatrix} \mathbf{X}_0 \\ \sqrt{\lambda_2} \mathbf{X}_B \end{bmatrix}; \bar{\mathbf{y}} = \begin{bmatrix} \mathbf{y} \\ 0 \end{bmatrix}.$$

类比之前的推导,可得出问题的解为

$$\mathbf{w} = (\mathbf{A}^T \mathbf{A} + \lambda_1 \mathbf{I})^{-1} \mathbf{A}^T \bar{\mathbf{y}}. \quad (13)$$

转换到复数域,则

$$\hat{\mathbf{w}} = \frac{\hat{\mathbf{x}}_0^* \cdot \mathbf{y}}{\hat{\mathbf{x}}_0^* \cdot \hat{\mathbf{x}}_0 + \lambda_1 + \lambda_2 (\hat{\mathbf{x}}_B^* \cdot \hat{\mathbf{x}}_B)} = \hat{\mathbf{x}}_0^* \cdot \hat{\alpha}_B. \quad (14)$$

在检测阶段,对于新输入的采样图像片 \mathbf{Z} ,可得出其分类器响应为

$$\hat{f}(\mathbf{z}) = \hat{\mathbf{z}} \cdot \hat{\mathbf{w}} = \hat{\mathbf{z}} \cdot \hat{\mathbf{x}}_0^* \cdot \frac{\hat{\mathbf{y}}}{\hat{\mathbf{x}}_0^* \cdot \hat{\mathbf{x}}_0 + \lambda_1 + \lambda_2 (\hat{\mathbf{x}}_B^* \cdot \hat{\mathbf{x}}_B)} = \frac{\hat{\mathbf{k}}^{z x_0} \cdot \hat{\mathbf{y}}}{\hat{\mathbf{k}}^{x_0 x_0} + \lambda_1 + \lambda_2 \hat{\mathbf{k}}^{x_B x_B}} = \hat{\mathbf{k}}^{z x_0} \cdot \hat{\alpha}_B \quad (15)$$

式中: $\hat{\mathbf{k}}^{z x_0}$ 表示采样图片 \mathbf{Z} 与目标模型 \mathbf{x}_0 的核相关矩阵; $\hat{\mathbf{k}}^{x_0 x_0}$ 表示目标模型 \mathbf{x}_0 的自相关核矩阵; $\hat{\mathbf{k}}^{x_B x_B}$ 表示对应背景块 \mathbf{x}_B 的自相关核矩阵。本文算法选用的核函数是高斯核函数。

在跟踪每帧并计算出待检区域的响应矩阵之后,计算其 PSR 值,当 PSR 值满足式(10)时,则认定当前跟踪区域出现了目标快速运动、遮挡、背景相似物干扰等情况,此时将式(8)中的目标模型学习参数 ζ 调整为 0,同时按照式(14)来计算下一帧的分类器模型 $\hat{\alpha}_B$;若计算出的 PSR 值不满足式(10),则认定当前跟踪区域无较大的干扰,直接按照式(5)计算下一帧的分类器模型 $\hat{\alpha}_B$,同时将学习参数 ζ 调整为初始值(一般初始值的范围设定为 0.3 ~ 0.75,不同测试视频最佳学习参数不同)。

根据文献[8-9],本文方法还额外构造一个一维尺度空间滤波器实现算法对多尺度的支持。

2 本文算法流程

结合以上论述,本文算法步骤如下。

1) 根据目标初始状态,提取其颜色属性特征,并使用 PCA 将原 11 维特征降维到 2 维,然后与原图像的灰度特征组合成 3 维特征作为算法输入。同时根据式(5)初始化位置分类器模型 $\hat{\alpha}_B$,根据文献[8-9]初始化一维尺度空间滤波器模型。

2) 根据式(6)计算待检区域的响应矩阵,同时计算该响应矩阵的 PSR 值,并确定当前帧中目标的中心位置,然后根据文献[8-9],以目标中心位置为中心取 33 个不同尺度大小的待检区域,利用一维尺度空间滤波器确定目标的真实尺度。

3) 根据上一步中计算的 PSR 值,按照式(10)进行多峰响应判断。若判定产生了多峰响应,则将目标模型的学习因子 ζ 置 0,同时以响应矩阵中次极大峰值的坐标为中心,构建一个和目标大小相同的背景块作为约束背景块,进行同样的特征提取和降维操作,然后按照式(14)计算新的分类器模型 $\hat{\alpha}_B$;若判定没有产生多峰响应,则将目标模型的学习因子 ζ 重置为初始值,按照式(5)计算新分类器模型 $\hat{\alpha}_B$ 。

4) 跟踪新的跟踪结果和分类器模型,按照式(7)和式(8)对各模型进行线性加权更新,继续跟踪下一帧。

3 实验结果与分析

本文实验所使用的测试视频均来自 Visual Tracker Benchmark,所选用的视频有 David, carscale, surf, girl1, girl2 和 dragonbaby 等。这些视频包含了目标快速运动、目标较大姿态变化、尺度变化、光照变化、摄像机模糊、目标遮挡、背景相似物干扰等一些具有挑战性的影响因素。大量实验表明,目标模型正则化参数 λ_1 取 0.001 时跟踪效果最好,背景约束模型正则化参数 λ_2 ,学习因子 η 和 ζ 在不同测试视频中的最佳值不同。采用的对比算法包括 DSST^[9], CN^[6], STC^[16], SAMF^[5],均使用文献作者主页上提供的源码。实验环境是 Matlab2016b,实验计算机配置是 I7-6700HQ 的 CPU,8 GB 内存。

3.1 定性分析

图 1 是各算法的部分实验结果对比,不同算法用不同颜色的实线框表示。图 1a 是 surf 视频的部分实验结果。由于原始的 CN 算法未利用背景信息,故而在第 17 帧时会因为身体肤色干扰产生多峰响应,导致一定程度的漂移。而 STC 所使用的是单一的灰度图像特征,对于身体肤色干扰的鲁棒性也不够,从第 141 帧开始漂移。DSST 算法使用单一的 HOG 特征,在目标发生旋转和快速形变时容易产生响应多峰,造成误差的累积,故而在第 376 帧跟错目标,而本文算法和 SAMF 算法则能对目标进行较好的跟踪。图 1b 的 carscale 视频主要测试算法对尺度变化目标的跟踪性能。由于 STC 算法是通过每 5 帧更新一次权值函数中的方差 θ 来实现对多尺度的支持,故在目标尺度发生连续性变化时不能产生较好的估计。而 SAMF 算法则是在 7 层的图像金字塔上检测平移位置的,因而对于尺度的估计比较粗糙,而原始的 CN 算法没有对多尺度的支持,仅本文算法和

DSST 算法能对目标尺度的变化进行较为精确的估计。图 1c 是 David 视频的部分实验结果,此视频存在目标尺度变化干扰和人物姿态变化干扰,由于干扰变化不是很大,各算法均能较好地跟踪,但由于跟踪误差的累积,在后面几帧中,CN,STC,SAMF 算法对于尺度支持不够好,造成在分类器模型中引入了较多的错误背景信息,导致跟踪有了一定程度的偏移,而本文算法和 DSST 算法仍然可以进行较好的跟踪。图 1d 的 girl1 视频跟踪难度较大,存在着人物姿态的多次变化、小幅度的目标尺度变化和干扰物遮挡。从实验结果可以看出,在第 103 帧人物转身时,目标特征发生了较大的变化,STC 和 DSST 算法均产生了一定的跟踪误差,而在第 252 帧目标第二次转身回来,CN 算法没有对周围背

景块的响应抑制,导致其模型也被污染,产生了一定程度的跟踪漂移,而第 439 帧的目标局部遮挡则使 CN 模型完全被污染,彻底跟丢,仅 SMAF 算法和本文算法对于目标特征改变和局部遮挡有较好的鲁棒性。图 1e 的 girl2 视频则存在着人物姿态变化和尺度变化,由于颜色属性特征和 HOG 特征均能较好地适应目标小幅度的姿态变化,因此 DSST,SMAF 和本文算法的跟踪效果都比较好,而 CN 算法在目标变小时学习到了过多错误的背景信息而导致跟丢,STC 算法则由于跟踪误差的累积逐步发生漂移。图 1f 的 dragonbaby 视频跟踪难度极大,视频中不仅存在目标转身导致的特征变化与尺度变化,还存在着目标快速运动导致的画面模糊和目标平移度过大,以及干扰物的局部遮挡。

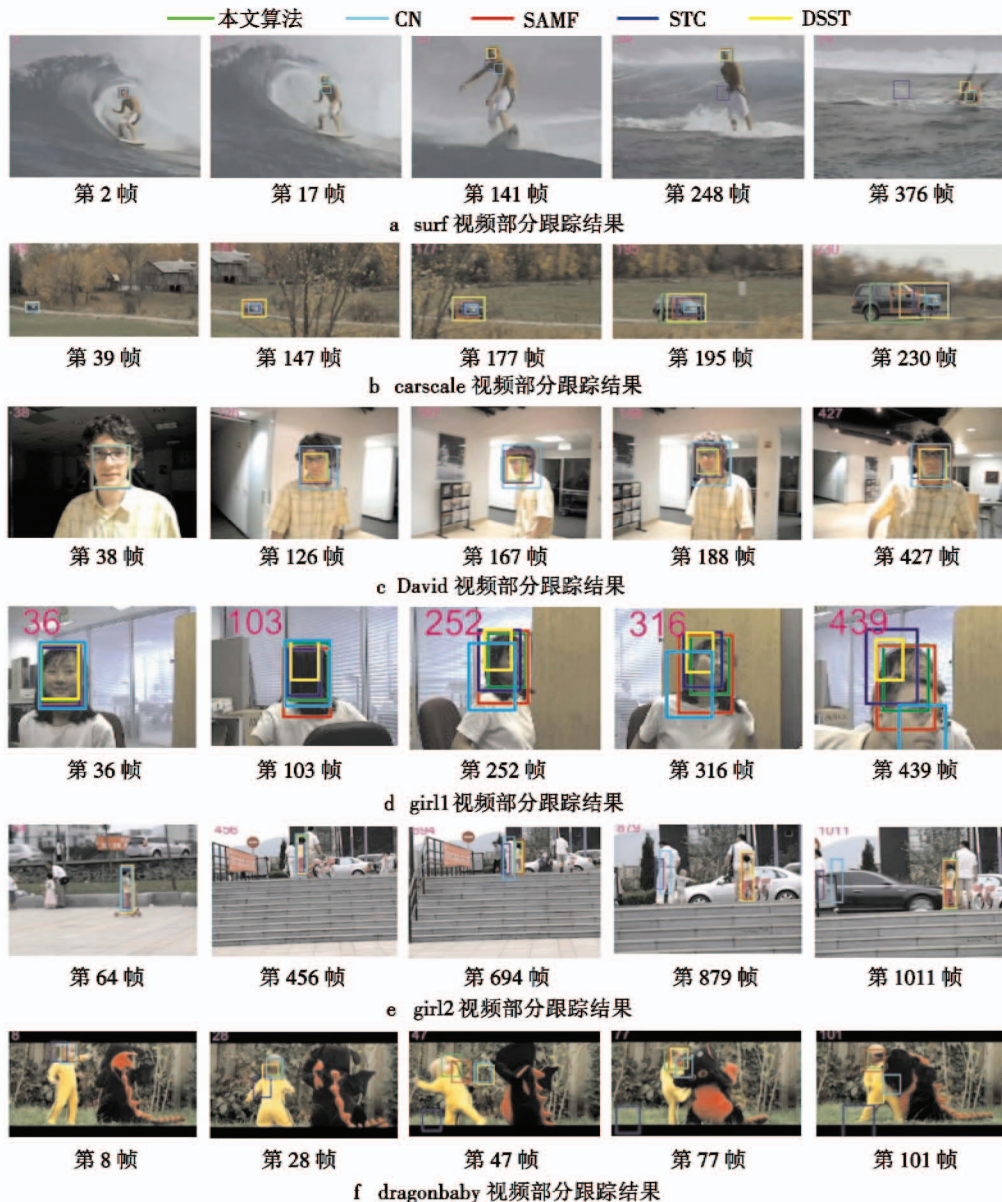


图 1 5 种算法部分实验结果定性对比

Fig. 1 Qualitative comparison of the experimental results using the five algorithms

从部分实验结果可以看到,当目标转身时,使用 HOG 特征的 DSST 算法和单一灰度特征的 STC 算法都产生了较大的漂移,而随着目标的快速跳开,由于 SAMF 算法和原始 CN 算法没有对周围背景信息学习的能力,因此当目标移动至检测区域边缘时,样本余弦窗的遮盖会导致训练的正样本质量大幅下降而产生漂移,因为 SAMF 算法使用的是复合特征,因此模型受污染程度不大,后续不会产生过大的跟踪漂移,而本文算法因为添加了特定的背景约束项,因此在正样本质量下降时可以通过学习背景信息来抑制周围错误信息,从而达到较高的跟踪精度。

3.2 定量分析

本文采用每秒处理帧数(FPS)衡量算法的跟踪效率(见表 1),采用文献[17]中平均中心位置误差(CLE)(见表 2)、距离精度(DP)、重叠精度(OP)(见表 3)这 3 个性能指标来衡量算法的跟踪精度。中心位置误差表示跟踪目标状态矩形框中心和真实目标状态矩形框中心之间的欧氏距离,其值越小,表示算法对目标位置的跟踪越精确。表 1~表 3 中,最优的结果用加粗字体表示,次优的结果用斜体表示。

表 1 平均每秒帧数(FPS)

Table 1 Average frames per second(FPS) 帧/s

视频	DSST	SAMF	STC	CN	本文算法
surf	63.0	18.4	<i>151.0</i>	191.0	92.0
carscale	56.8	17.7	63.9	208.0	<i>65.4</i>
David	21.2	13.7	<i>125.0</i>	135.0	65.7
girl1	65.7	34.1	187.0	<i>146.0</i>	80.6
girl2	15.7	7.7	<i>51.3</i>	91.1	45.9
dragonbaby	13.1	9.6	87.3	<i>73.8</i>	40.3

表 2 平均中心位置误差(CLE)

Table 2 Average center location error(CLE) 像素

视频	DSST	SAMF	STC	CN	本文算法
surf	7.2	3.9	41.7	10.5	<i>4.5</i>
carscale	19.1	8.3	27.3	26.1	7.0
David	4.2	4.8	4.8	8.1	<i>4.6</i>
girl1	10.6	7.2	7.6	20.3	4.9
girl2	9.8	10.5	146.0	119.0	<i>10.0</i>
dragonbaby	53.7	25.8	158.0	58.8	24.3

表 3 平均重叠率

Table 3 Average overlap rate %

视频	DSST	SAMF	STC	CN	本文算法
surf	47.3	68.1	5.92	34.6	<i>67.1</i>
carscale	74.7	58.8	44.5	44.5	75.6
David	79.9	69.9	67.4	62.2	80.3
girl1	38.9	<i>64.8</i>	55.1	39.2	70.7
girl2	<i>68.4</i>	67.4	25.1	48.9	68.6
dragonbaby	37.6	54.1	11.5	32.7	<i>53.4</i>

从表 2 可以看到,由于本文算法在原始的 CN 算法

基础上引入了自适应的背景约束项来提高算法分类器对背景信息的利用,达到了较高的跟踪精度,所以在各测试图像序列中,本文算法的 CLE 均排在前列,相比原始的 CN 算法,除去 CN 算法在 girl2 中跟丢的结果,总的平均 CLE 从 24.76 像素提高到了 9.06 像素。

3.2.1 算法跟踪效率

每秒处理帧数表示算法每秒所能处理的帧数,其值越大,表示算法的跟踪效率越高。表 1 为在各测试视频中,本文算法与对比算法的平均每秒处理帧数。基于核相关滤波的算法,其精度、速度与检测区域有较大的关系,因此在不同测试视频中速度表现不一样。在 girl2, dragonbaby 等目标移动较快的视频中,所取的检测区域面积是目标大小的 2 倍,其余测试视频检测区域面积是目标大小的 1 倍,因此速度差异较大。其中, SAMF 算法是将 HOG 特征和 CN 特征相结合,在 7 层图像金字塔上进行检测,因此速度最慢,平均速度仅 16.89 帧/s。DSST 算法使用的 31 维 HOG 特征,在特征提取和处理阶段比使用 2 维的 CN 颜色特征的算法要慢很多,平均速度只有 39.25 帧/s。STC 算法利用的是低阶灰度特征,因此相比之下速度很快,平均速度可以达到 100 多帧/s。原始 CN 算法利用循环结构构造训练样本,利用 PCA 将特征降至 2 维,极大地节省了运算开支,其平均速度达到了 140.82 帧/s,是对比算法中最快的。而本文是在 CN 算法基础上添加了背景约束项,并增加了一维尺度滤波器,因此相比原始的 CN 算法损失了一半的速度,平均速度有 63.85 帧/s,但还是可以很好地满足实时性要求。

3.2.2 算法跟踪精度

距离精度表示 CLE 小于某一个阈值时的帧数占总测试序列帧数的百分比,图 2 表示各测试视频中各算法的 DP 曲线,可以看到,本文算法基本可以包围其他算法的曲线,完全包围原始 CN 算法的曲线,对于目标位置有较高的跟踪精度。综合考虑算法对目标位置和目标尺度的跟踪精度,还可以用重叠精度来衡量。重叠率可表示为

$$r = \frac{S_T \cap S_R}{S_T + S_R - (S_T \cup S_R)} \quad (16)$$

式中, S_T 表示跟踪目标状态矩形框, S_R 表示真实目标状态矩形框, $S_T \cap S_R$ 表示这二者的重叠部分。而重叠精度表示重叠率大于某一个阈值的帧数占测试图像序列总帧数的比率。表 3 表示各算法在各个测试图像序列的平均重叠率,本文算法的 OP 均值为 69.28%,相比原始 CN 算法 43.68% 的 OP 均值,提升了 25.6%,相比次优的 SAMF 算法的 63.83%,提升了 5.45%。从图 2 中各视频的实验结果可以较为直观地比较各算法在各测试图像序列中的 OP 值。

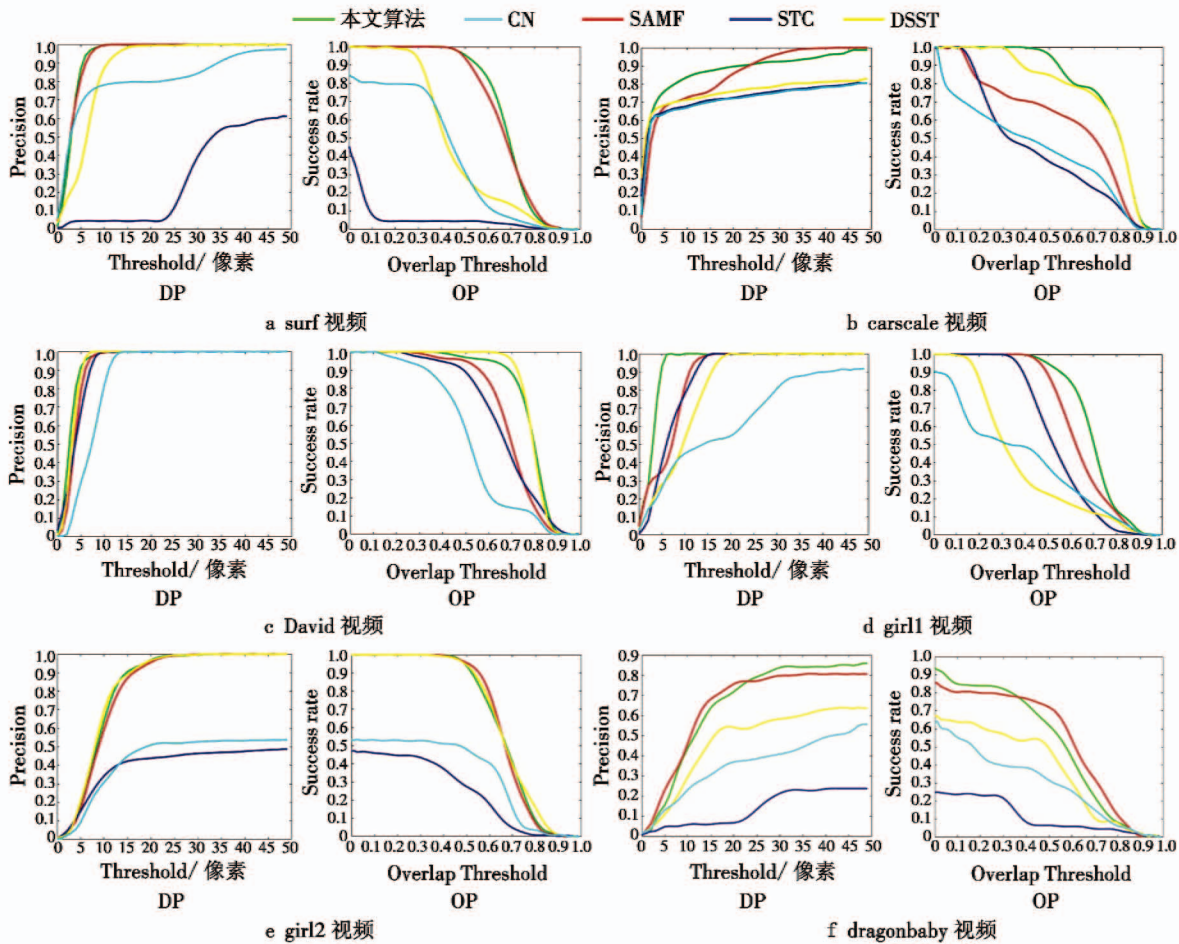


图2 各视频的实验结果

Fig. 2 Experimental result of each video

4 结论

针对颜色属性跟踪算法对背景信息利用过少,在目标被局部遮挡、快速形变和快速运动时会产生漂移的问题,提出了一种使用自适应背景约束改进的颜色属性跟踪算法。该算法基于颜色属性跟踪算法,在脊回归分类器目标函数中引入特定的背景约束来抑制干扰响应峰,同时使用一维尺度滤波器对目标尺度进行估计。

为保证算法有较高的跟踪效率,本文算法仅使用单一的CN特征,且仅对影响最大的背景干扰块引入背景约束来进行抑制。在目标非快速变色的情况下,本文算法对目标姿态变化、目标局部遮挡、运动模糊、背景相似物等干扰有较强的鲁棒性,可以满足一般情况下的实时应用。

参考文献

- [1] YANG H, SHAO L, ZHENG F, et al. Recent advances and trends in visual tracking: a review[J]. *Neurocomputing*, 2011, 74(18):3823-3831.
- [2] KRISTAN M, LEONARDIS A, MATAS J, et al. The visual

- object tracking VOT2016 challenge results[C]//European Conference on Computer Vision, 2016:777-823.
- [3] BOLME D S, BEVERIDGE J R, DRAPER B A, et al. Visual object tracking using adaptive correlation filters[C]//Conference on Computer Vision and Pattern Recognition, IEEE, 2010:2544-2550.
- [4] HENRIQUES J F, CASEIRO R, MARTINS P, et al. Exploiting the circulant structure of tracking-by-detection with kernels[C]//European Conference on Computer Vision. Springer, Berlin, Heidelberg, 2012:702-715.
- [5] HENRIQUES J F, CASEIRO R, MARTINS P, et al. High-speed tracking with kernelized correlation filters[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(3):583-596.
- [6] DANELLJAN M, KHAN F S, FELSBERG M, et al. Adaptive color attributes for real-time visual tracking[C]//Conference on Computer Vision and Pattern Recognition, IEEE, 2014:1090-1097.

(下转第51页)

在运行每一个视频跟踪程序时,都输出了每帧匹配的时间,根据输出的数据可以统计出每一种算法在每个视频中的匹配时间,如表 2 所示。

表 2 跟踪平均匹配时间

Table 2 The average matching time for tracking s/帧

	本文算法	基于 DCT 的压缩感知哈希算法	传统哈希算法
David	0.561	0.548	0.354
Girl	0.583	0.524	0.338
CarScale	0.535	0.543	0.367

由表 1 可以清楚地看到,本文算法比其他两种哈希算法在跟踪成功率方面有提高;由表 2 可知,本文算法的匹配速度一般,但是基本上能满足视频跟踪要求。

5 结束语

本文提出了基于类间方差和离散余弦变换的模板匹配增强哈希目标跟踪算法,两种算法提取了目标的两种特征,优势互补,克服了其他哈希算法在复杂环境下目标发生形变、有遮挡物、光照发生变化等跟踪障碍,并在实验中取得了较好的效果。本文算法的综合性能较好,鲁棒性较强,提高了基于 DCT 的压缩感知哈希跟踪算法的成功率。但本文算法的跟踪框不能达到自适应,还有待改善。

参考文献

- [1] 李雨佳,李京兵. 基于 DCT 和感知哈希的纹理防伪标鉴别算法[J]. 计算机应用研究, 2014, 31(12): 3734-3737.
- [2] 孙锐, 闫晓星, 丁志中. 一种用于图像认证的感知哈希方法[J]. 系统仿真学报, 2010, 22(2): 483-486.
- [3] YANG B, GU F, NIU X. Block mean value based image perceptual hashing[C]//International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2006:167-172.
- [4] FEI M, LI J, SHAO L, et al. Robust visual tracking based on improved perceptual hashing for robot vision[C]//Intelligent Robotics and Applications, 2015:331-340.
- [5] 齐苏敏, 隋煜舜. 基于改进 Adaboost 特征检测的感知哈希跟踪算法[J]. 通信技术, 2017, 50(3): 430-435.
- [6] BARIS C, BULENT S, NASIR M, et al. Spatio temporal transform based video hashing[J]. IEEE Transactions on Multimedia, 2006, 8(6): 1190-1208.
- [7] FRID J, GOL M. Robust hash functions for digital watermarking[C]//IEEE Proceedings of the International Conference on Information Technology: Coding and Computing, 2000:178-183.
- [8] 徐珩, 贺飞越. 模板匹配的哈希增强算法[J]. 计算机应用与软件, 2016, 33(7): 167-171.
- [9] FEI M, LI J, LIU H H, et al. Visual tracking based on improved foreground detection and perceptual hashing[J]. Neuro Computing, 2015, 152(10): 413-428.
- [10] WU Y, LIM J W, YANG M H. Object tracking benchmark[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1834-1848.
- [11] DANELLJAN M, HÄGER G, KHAN F S, et al. Accurate scale estimation for robust visual tracking[C]//British Machine Vision Conference, Nottingham, 2014:1-11.
- [12] DANELLJAN M, HÄGER G, KHAN F S, et al. Discriminative scale space tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(8): 1561-1575.
- [13] BIBI A, MUELLER M, GHANEM B. Target response adaptation for correlation filter tracking[C]//European Conference on Computer Vision, 2016:419-433.
- [14] WANG M, LIU Y, HUANG Z. Large margin object tracking with circulant feature maps[J]. Computer Vision and Pattern Recognition, 2017. doi:10.1109/CVPR.2017.510.
- [15] LUKEŽIČ A, VOJÍŘ T, ČEHOVIN L, et al. Discriminative correlation filter with channel and spatial reliability[J]. Computer Vision and Pattern Recognition, 2016. doi:10.1007/s11263-017-1061-3.
- [16] VAN DE WEIJER J, SCHMID C, VERBEEK J. Learning color names from real-world images[C]//IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2007:1-8.
- [17] VAN DE WEIJER J, SCHMID C, VERBEEK J, et al. Learning color names for real-world applications[J]. IEEE Transactions on Image Processing, 2009, 18(7): 1512-1523.
- [18] SCHÖLKOPF B, SMOLA A J. Learning with kernels: support vector machines, regularization, optimization, and beyond[M]. Cambridge: MIT Press, 2002.
- [19] ZHANG K, ZHANG L, LIU Q, et al. Fast visual tracking via dense spatio-temporal context learning[C]//European Conference on Computer Vision, 2014:127-141.
- [20] WU Y, LIM J, YANG M H. Object tracking benchmark[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1834-1848.

(上接第 38 页)