



基于特征融合及自适应模型更新的相关滤波目标跟踪算法

刘威, 靳宝, 周璇, 付杰, 王薪予, 郭直清, 牛英杰

引用本文:

刘威, 靳宝, 周璇, 等. 基于特征融合及自适应模型更新的相关滤波目标跟踪算法[J]. 智能系统学报, 2020, 15(4): 714–721.

LIU Wei, JIN Bao, ZHOU Xuan, et al. Correlation filter target tracking algorithm based on feature fusion and adaptive model updating[J]. *CAAI Transactions on Intelligent Systems*, 2020, 15(4): 714–721.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.201803036>

您可能感兴趣的其他文章

区域损失函数的孪生网络目标跟踪

Regional loss function based siamese network for object tracking

智能系统学报. 2020, 15(4): 722–731 <https://dx.doi.org/10.11992/tis.201910005>

基于LiDAR/INS的野外移动机器人组合导航方法

Integrated navigation approach for the field mobile robot based on LiDAR/INS

智能系统学报. 2020, 15(4): 804–810 <https://dx.doi.org/10.11992/tis.202008026>

无人机群多目标协同主动感知的自组织映射方法

Self-organizing feature map method for multi-target active perception of unmanned aerial vehicle systems

智能系统学报. 2020, 15(3): 609–614 <https://dx.doi.org/10.11992/tis.201908022>

基于反卷积和特征融合的SSD小目标检测算法

SSD small target detection algorithm based on deconvolution and feature fusion

智能系统学报. 2020, 15(2): 310–316 <https://dx.doi.org/10.11992/tis.201905035>

注意力机制和Faster RCNN相结合的绝缘子识别

Insulator recognition based on attention mechanism and Faster RCNN

智能系统学报. 2020, 15(1): 92–98 <https://dx.doi.org/10.11992/tis.201907023>

压缩感知W-HOG特征的运动手势跟踪

Motion gesture tracking based on compressed sensing W-HOG features

智能系统学报. 2016, 11(01): 124–128 <https://dx.doi.org/10.11992/tis.201507005>

微信公众平台



关注微信公众号, 获取更多资讯信息

DOI: 10.11992/tis.201803036

基于特征融合及自适应模型更新的相关滤波目标跟踪算法

刘威^{1,2,3}, 靳宝^{1,2,3}, 周璇^{1,2,3}, 付杰^{1,2,3}, 王薪予^{1,2,3}, 郭直清^{1,2,3}, 牛英杰^{1,2,3}

(1. 辽宁工程技术大学 理学院, 辽宁 阜新 123000; 2. 辽宁工程技术大学 智能工程与数学研究院, 辽宁 阜新 123000; 3. 辽宁工程技术大学 数学与系统科学研究所, 辽宁 阜新 123000)

摘要: 针对单一特征目标跟踪算法因背景干扰、目标遮挡造成的跟踪失败问题, 以及跟踪过程中每帧进行模型更新容易造成错误更新和实时性差的问题, 提出了一种基于特征融合及自适应模型更新策略的相关滤波目标跟踪算法——多特征自适应相关滤波目标跟踪算法。该算法在特征提取阶段将边缘特征及 HOG 特征加权融合作为目标特征, 加强对边缘特征的学习; 在模型更新阶段通过计算预测区域与真实区域的奇异值特征向量相似度, 并结合设定的阈值判断是否需要模型更新, 通过自适应更新的方式减少模型的更新次数。在标准测试视频集下验证所提算法, 并与两种经典相关滤波算法进行比较, 结果表明该算法能够较好地适应背景干扰及目标遮挡问题, 跟踪目标的平均中心误差减少了 9.05 像素, 平均距离精度提高 12.2%, 平均重叠率提高 4.53%。
关键词: 目标跟踪; 相关滤波; 特征融合; 模型更新; 目标遮挡; 背景干扰; 计算机视觉; 奇异值分解

中图分类号: TP301 文献标志码: A 文章编号: 1673-4785(2020)04-0714-08

中文引用格式: 刘威, 靳宝, 周璇, 等. 基于特征融合及自适应模型更新的相关滤波目标跟踪算法 [J]. 智能系统学报, 2020, 15(4): 714-721.

英文引用格式: LIU Wei, JIN Bao, ZHOU Xuan, et al. Correlation filter target tracking algorithm based on feature fusion and adaptive model updating[J]. CAAI transactions on intelligent systems, 2020, 15(4): 714-721.

Correlation filter target tracking algorithm based on feature fusion and adaptive model updating

LIU Wei^{1,2,3}, JIN Bao^{1,2,3}, ZHOU Xuan^{1,2,3}, FU Jie^{1,2,3}, WANG Xinyu^{1,2,3},
GUO Zhiqing^{1,2,3}, NIU Yingjie^{1,2,3}

(1. School of Sciences, Liaoning Technical University, Fuxin 123000, China; 2. Institute of Intelligence Engineering and Mathematics, Liaoning Technical University, Fuxin 123000, China; 3. Institute of Mathematics and Systems Science, Liaoning Technical University, Fuxin 123000, China)

Abstract: For the problem of object tracking failure caused by background interference, object occlusion in object tracking algorithm based on a single feature, and the problem of error updating and poor real-time performance caused by model updating for each frame during tracking, a correlation filtering object tracking algorithm based on feature fusion and adaptive model updating is proposed in this paper. In the feature extraction phase, the edge feature and HOG feature are weighted together as the object features to enhance the learning of edge features. During the updating phase of the model, calculating the similarity of singular value eigenvector between the predicted region and the real region, and according to the set threshold to determine whether the model needs to be updated or not, and reducing the number of updates to the model by adaptively updating. The algorithm is verified under the standard test video set, and compared with two typical correlation filtering algorithms. The results show that the algorithm can better adapt to background interference and object occlusion problem. The average center location error is reduced by 9.05 pixels, the average distance precision is increased by 12.2%, and the average overlapping precision is increased by 4.53%.

Keywords: object tracking; correlation filter; feature fusion; model updating; object occlusion; background interference; computer vision; singular value decomposition

收稿日期: 2018-03-22.

基金项目: 国家自然科学基金项目 (51974144, 71771111).

通信作者: 刘威. E-mail: lv8218218@126.com.

视频目标跟踪 (video object tracking, VOT) 是计算机视觉领域的重要研究方向, 综合了图像分

析与处理、模式识别、人工智能、自动控制等领域的先进技术,在视觉导航、机器人、智能交通、视频监控等方面具有广泛应用^[1-2]。但在实际应用中仍面临诸多困难与挑战,主要包括光照变化、尺度变化、目标遮挡、背景干扰、非刚体形变、旋转、运动模糊等。因此近年来一些目标跟踪算法被提出^[3-6],其中基于相关滤波的目标跟踪采用密集采样策略,利用频域中的快速傅里叶变换(fast fourier transformation, FFT)代替空间卷积运算,进行相关滤波的训练,通过对后续帧进行相关滤波快速定位目标,因其跟踪精度及实时性显著提升而被广泛关注。

最小输出平方误差和(minimum output sum of squared error, MOSSE)^[7]滤波器最早将相关滤波用于目标跟踪,该方法采用单一的灰度特征,速度远高于其他算法,但准确度一般。之后在MOSSE算法基础上,一系列改进算法被提出^[8-13]。其中CSK^[8]采用单一的灰度特征,易受背景复杂、目标和背景颜色相似的影响,KCF^[9]采用方向梯度直方图HOG特征,CN^[10]中采用颜色特征,运用自适应降维方法将11维颜色特征降为2维用于目标跟踪,均得到了较好的跟踪效果。文献[14]将原始图像的灰度特征、颜色特征及HOG特征融合,提高了跟踪性能。随着深度学习的发展,许多算法将深度学习提取的特征同现有的特征进行融合^[15-17],同样提高了跟踪性能。

虽然上述算法取得了较好的跟踪效果,但是存在严重背景干扰及目标遮挡等情况下,仍容易发生跟踪失败的情况。为了在严重背景干扰及目标遮挡情况下更好地完成跟踪,考虑到图像的边缘特征作为图像的基本特征,携带了原始图像的绝大部分信息,而HOG特征是分块提取图像的局部特征,两个特征具有一定的互补性,从而提出一种基于特征融合方法及自适应模型更新策略的相关滤波目标跟踪算法——多特征自适应相关滤波目标跟踪算法,通过选择边缘特征和HOG特征的线性加权融合结果作为目标特征,加强对边缘特征的学习;采用基于奇异值特征向量相似度的自适应模型更新策略,避免了每一帧都更新模型造成的重复计算,降低了错误更新的概率,更好适应了背景干扰及目标遮挡等问题。

1 算法描述

相关滤波跟踪算法的总体框架是根据视频第一帧中给定的目标位置提取图像块,训练、学习得到相关滤波器;对后续的每一帧,从上一帧目

标位置的区域提取图像块用来检测;提取图像块的特征,同时加余弦窗口平滑边缘;利用离散傅里叶变换进行相关滤波操作;通过傅里叶变换得到响应图谱,具有最大响应值的位置就是要预测的目标位置;提取估计位置的目标外观,训练和更新相关滤波器。

1.1 相关滤波目标跟踪算法

将一张图片以 d 维特征图表示。设 f 是从这个特征图中提取出来的目标的矩形块。用 f^l 来表示 f 的特征维数 $l \in \{1, 2, \dots, d\}$ 。相关滤波算法目的是找到一个最佳的相关滤波器 h ,该滤波器由每个特征维度的滤波器 h^l 组成。通过最小化代价函数式(1)来实现:

$$\varepsilon = \left\| \sum_{l=1}^d h^l * f^l - g \right\|^2 + \lambda \sum_{l=1}^d \|h^l\|^2 \quad (1)$$

式中: $*$ 表示循环相关(h^l 的复共轭与 f^l 对应位置元素乘积), g 是与训练样本 f 相关的期望相关输出。 $\lambda \geq 0$ 是正则化参数。式中只考虑一个训练样本,求解得:

$$H^l = \frac{\bar{G}F^l}{\sum_{k=1}^d \bar{F}^k F^k + \lambda} \quad (2)$$

式(1)中正则化参数缓解了 f 频谱中零频分量的问题,避免了除数为零。通过最小化所有训练块的输出误差来获得最优滤波器。但需要对每个像素的 $d \times d$ 线性系统求解,这对于在线学习计算资源的消耗是高昂的。为了获得一个鲁棒稳定的近似,通过式(3)、(4)更新式(2)中的相关滤波器 H_t^l 的分子 A_t^l 和分母 B_t :

$$A_t^l = (1 - \eta)A_{t-1}^l + \eta \bar{G}_t F_t^l \quad (3)$$

$$B_t = (1 - \eta)B_{t-1} + \eta \sum_{k=1}^d \bar{F}_t^k F_t^k \quad (4)$$

式中: η 是学习率,使用式(5)计算特征图的矩形区域 z 处的相关得分 y ,然后通过最大化分数 y 找到新的目标状态。

$$y = F^{-1} \left\{ \sum_{l=1}^d \bar{A}^l Z^l / (B + \lambda) \right\} \quad (5)$$

1.2 边缘特征和HOG特征的融合策略

图像中的边界信息携带了原始图像的绝大部分信息,描述了图像的整体边缘特征,可以通过边缘线就能识别出一个物体,因此对图像边缘特征的提取与学习一定程度上影响了对后续处理的精度;HOG特征是分块提取图像的局部特征;因此通过将两种特征进行融合,可以更好地掌握图

像的特征,为目标跟踪的后续工作提供更加精确的特征表示。

在特征提取阶段采用线性加权方式融合边缘特征和HOG特征作为目标特征,具体流程如下:

1) 分别提取目标区域的边缘特征 F_{edge} 和 HOG 特征 F_{HOG} ;

2) 初始化融合参数 δ ;

3) 计算边缘特征与 HOG 特征融合后的目标特征: $F = \delta F_{\text{edge}} + (1 - \delta) F_{\text{HOG}}$ 。

其中 $\delta \in [0, 1]$, 当 $\delta=0$ 时, 表示跟踪只使用 HOG 特征; 当 $\delta=1$ 时, 表示跟踪只使用边缘特征。实验中根据所需的对象特征适当选取 δ 的值。

1.3 基于奇异值特征向量相似性的自适应模型更新策略

同离线跟踪相比,在线跟踪方式的优势在于可以实时对模型进行更新,增加了跟踪的鲁棒性,目前主要的更新方式有以下3种:

1) 每一帧都进行更新^[18]。这种更新方式过于频繁,增加了错误更新的概率,同时增加了模型漂移的可能。

2) 当响应分数低于一定阈值才进行模型更新^[6]。减少了更新频率,增加了对模型的判断。

3) 计算正负样本的响应分数,当二者的差值低于一定阈值时更新模型^[19]。

本文提出了一种基于奇异值特征向量相似性的模型自适应更新策略: 首先将预测区域图像矩阵与真实区域图像矩阵进行奇异值分解,然后计算分解后得到的奇异值特征向量之间的余弦距离,来表征预测区域和真实区域的相似程度,并与给定阈值比较以决定模型是否更新。

矩阵奇异值分解(singular value decomposition, SVD)是矩阵特征提取的一种主要方法,它通过将复杂的矩阵转换为代表矩阵特征的更小、更简单的几个子矩阵相乘来实现矩阵特征提取。图像的奇异值特征向量具有正交变换、旋转等代数和几何上的不变性,同时具有稳定性和抗噪性,因此在模式识别、图像分析等领域中应用广泛^[20]。

定义1 设 $A \in C_r^{m \times n} (r > 0)$, $A^H A$ 的特征值为 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r \geq \lambda_{r+1} = \dots = \lambda_n = 0$, 则称 $\sigma_i = \sqrt{\lambda_i} (i = 1, 2, \dots, n)$ 是 A 的奇异值; 规定零矩阵的奇异值都是0。

定理1 设 $A \in C_r^{m \times n} (r > 0)$, 则存在 m 阶酉矩阵 U 和 n 阶酉矩阵 V , 使得

$$A = U \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} V^H \quad (6)$$

式中: 矩阵 $\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$, 而 $\sigma_1, \sigma_2, \dots, \sigma_r$ 是矩阵 A 的所有非零奇异值。称式(6)是矩阵 A 的

奇异值分解。

基于奇异值特征向量相似性的模型自适应更新算法流程如下:

1) 分别对预测区域与真实区域图像矩阵进行奇异值分解,得到预测区域与真实区域图像的奇异值特征向量 M_1 和 M_2 ;

2) 利用余弦距离计算 M_1 和 M_2 的相似度:

$$\lambda = \frac{M_1 \cdot M_2}{|M_1| \cdot |M_2|} \quad (7)$$

3) 若 $\lambda \geq 0.5$, 表示预测区域与真实区域相似,则通过式(3)、(4)进行模型更新,同时设置较低的学习率 $l = 0.025$, 以保持较好的跟踪; 若 $\lambda < 0.5$, 则认为预测区域与真实区域相差很大,不进行模型更新,同时设置较大的学习率 $l = 1$ 。

通过对大量数值实验结果的对比分析可知,当 λ 取值较大或较小时,目标跟踪易发生偏移,鲁棒性较差,跟踪准确率低,因此最终实验确定 λ 的取值为0.5。而学习率表征了对目标表观的学习能力,一般由经验值来确定,学习率越大说明学习速度越快,对目标表观变化大的复杂环境跟踪良好,学习率越小说明学习速度越慢,对目标表观变化小的环境跟踪良好。

1.4 多特征自适应相关滤波目标跟踪算法

基于边缘特征和HOG特征的融合策略和奇异值特征向量相似性的自适应模型更新策略,本文提出了一种新的相关滤波目标跟踪算法——多特征自适应相关滤波目标跟踪算法。算法流程如下:

1) 根据给定的视频第1帧目标位置,提取边缘特征及HOG特征;

2) 根据视频内容设定特征融合参数,计算得到融合后的特征值;

3) 训练、学习相关滤波器;

4) 提取下一帧目标位置特征,进行相关滤波操作,得到预测区域;

5) 计算预测区域与真实区域的奇异值特征向量相似度;

6) 若相似度大于设定的阈值,则进行相关滤波器更新,若相似度小于设定的阈值,则不更新;

7) 重复4)~6)。

2 实验结果与分析

2.1 数据集说明

为了验证基于特征融合及自适应模型更新的相关滤波目标跟踪算法的有效性,选取视频跟踪

标准数据集 OTB-50 和 OTB-100 中存在背景干扰问题的 18 组视频序列进行数值实验,其中 9 组同

时存在目标遮挡问题,各组视频序列的帧数、目标大小以及存在的问题信息见表 1。

表 1 测试视频信息

Table 1 Test video information

测试视频	帧数	目标大小/(像素×像素)	主要问题	测试视频	帧数	目标大小/(像素×像素)	主要问题
Basketball	725	81×34	OCC、BC、IV	Liquor	1 741	210×73	OCC、BC、SV、IV
CarDark	393	23×29	BC、IV	Matrix	100	42×38	OCC、BC、SV、IV
Coke	291	80×48	OCC、BC、IV	Mhyang	1 490	70×62	BC、IV
Crossing	120	50×17	BC、SV	MountainBike	228	56×67	BC
David3	252	78×64	OCC、BC	Shaking	365	71×61	BC、SV、IV
Deer	71	65×95	BC、MB	Singer2	366	122×67	BC、IV
Dudek	1 145	36×51	OCC、BC、SV	Skating1	400	84×34	OCC、BC、SV
Football	362	50×39	OCC、BC	Subway	175	51×19	OCC、BC
Football1	74	43×26	BC	Trellis	569	101×68	BC、SV、IV

注: OCC—遮挡, BC—背景干扰, SV—尺度变化, IV—照明变化, MB—运动模糊

2.2 评估方式

为了评估目标跟踪算法的性能,本文使用如下评估指标^[11]: 中心位置误差 (center location error, CLE)、区域重叠准确率 (region overlap precision, OP)、距离精度 (distance precision, DP)、跟踪速度 (frames per second, FPS)、跟踪长度 (tracking length, TL)。其中 CLE 为检测到的目标中心与真实目标中心间的平均欧氏距离, OP 为检测区域与真实区域重叠面积超过一定阈值的视频帧数占视频总帧数的百分比, DP 为中心位置误差小于一定阈值的视频帧数占总帧数的百分比, FPS 为每秒处理的视频帧数, TL 为从跟踪开始到中心位置误差下降到一定阈值以下的帧数。CLE 的值越小, OP、DP、FPS、TL 的值越大, 跟踪效果越好。

2.3 实验分析

本文共设计 3 组实验, 第 1 组实验在 18 组测试视频上对现有的两种算法 (基于多维特征的 DSST^[12] 算法、基于 HOG 特征的 SRDCF 算法^[13]) 和本文的算法的跟踪效果进行对比; 第 2 组实验重点测试算法在同时具有背景干扰及目标遮挡问题的视频序列上的跟踪效果; 第 3 组实验对比 3 种跟踪算法的跟踪速度。

1) 跟踪效果对比实验

对 18 组存在背景干扰的视频进行实验, 为确定每组视频的最优融合参数 δ , 以步长为 0.1, 从 0~1 对融合参数 δ 进行采样测试, 选择测试中最优的 δ 如表 2 所示, 验证本文算法和现有的两种相关滤波跟踪算法的跟踪效果。所有实验中均采

用固定学习率 $\gamma=0.025$, 均采用同 DSST 相同的尺度估计, 除本文算法采用自适应模型更新策略外, 其他算法均采用每帧更新策略。实验结果如表 2 所示。可以看出, 本文算法对 18 个视频的 CLE、DP 和 OP 指标均表现良好, 跟踪长度也略优于其他对比算法。本文算法得到的 18 组视频的平均中心位置误差 CLE 为 8.91, 平均中心距离精度 DP 为 92.92%, 平均重叠面积精度 OP 为 81.65%, 同其他两种算法表现最优的结果相比, 本文算法的平均中心位置误差 CLE 减少 9.05, 平均中心距离精度 DP 增加 12.2%, 平均重叠面积精度 OP 增加 4.53%。

2) 背景干扰及目标遮挡下的跟踪实验

对存在严重背景干扰及目标遮挡的 3 组视频 Basketball、David3 和 Liquor 序列进行实验, 图 1 所示为 3 组视频的跟踪曲线。从图 1 中 (a)~(c) 可以看出, 在 Basketball 数据集中, 当期期望重叠率为 0.5 时, 本章算法的重叠面积精度高于 DSST 算法、SRDCF 算法; 在 David3 数据集中, 本文算法的重叠面积精度略低于 SRDCF 算法, 优于 DSST 算法; 在 Liquor 数据集中, 本文算法的重叠率精度取得了同 SRDCF 算法相近的结果, 优于 DSST 算法。由图 (d)~(f) 的距离精度可以看出, 当期期望像素误差为 20 时, 在 Basketball 数据集中, 本文算法的距离精度远远高于 SRDCF 算法和 DSST 算法; 在 David3 数据集中, 本文算法取得同 SRDCF 算法, 但都远远高于 DSST 算法; 在 Liquor 数据集中, 本文算法的距离精度, 略高于 SRDCF 算法, 远高于 DSST 算法。当距离精度为

0.8 时, Basketball、David3、Liquor 3 个数据集中, 本文算法的误差像素均低于 SRDCF 和算法 DSST 算法。由图 (g)~(i) 的跟踪长度可以看出, 当期望像素误差为 20 时, 在 Basketball 数据集中, 本文算法的跟踪长度大于 SRDCF 算法及 DSST 算法; 在

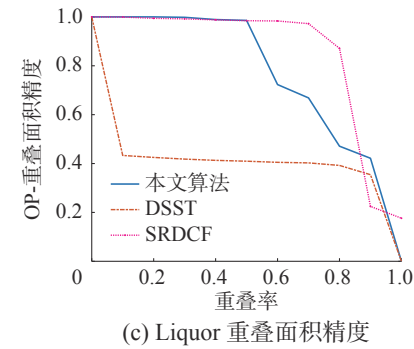
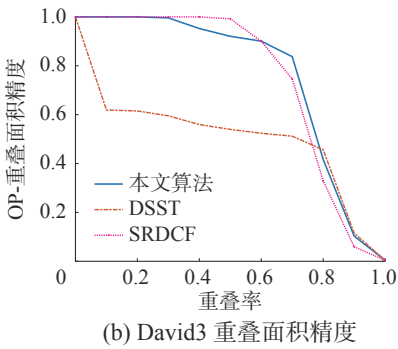
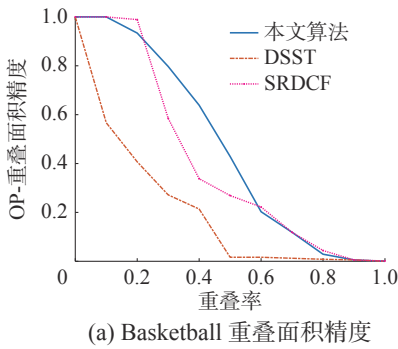
David3 数据集中, 本文算法和 SRDCF 算法具有相同的跟踪长度, 但都大于 DSST 算法; 在 Liquor 数据集中, 本文算法略大于 SRDCF 算法和 DSST 算法。综合下来, 同其他两种算法相比, 本文算法的跟踪效果更好。

表 2 3 种算法在 18 组视频序列上的跟踪效果
Table 2 The tracking effect of 3 algorithms on 18 groups of video sequences

测试视频	融合参数	CLE/像素			DP/%			OP/%			TL/帧			总帧数
		DSST	SRDCF	本文	DSST	SRDCF	本文	DSST	SRDCF	本文	DSST	SRDCF	本文	
Basketball	0.6	83.4	25.1	12.9	45.1	32	77.9	1.66	26.9	42.9	12	8	243	725
CarDark	0.5	1.03	1.31	1.01	100	100	100	100	100	100	393	393	393	393
Coke	0.2	17.5	20	17.5	86.9	80.4	86.6	79	71.1	78.4	101	91	101	291
Crossing	0.5	1.5	1.88	1.37	100	100	100	100	100	100	120	120	120	120
David3	0.5	88	5.23	3.95	61.1	100	100	54	99.2	92.1	154	252	252	252
Deer	0.2	10.3	5.12	9.54	88.7	100	90.1	88.7	100	84.5	24	71	24	71
Dudek	0.7	13.6	14.2	11.8	81.4	79.6	82.9	98.9	96.1	93	48	46	46	1145
Football	0.5	16.7	6.82	9.09	79.8	100	98.6	63	84.8	50.8	289	362	225	362
Football1	0.9	32.1	21.5	8.64	51.4	78.4	97.3	35.1	54.1	39.2	38	58	63	74
Liquor	0.3	130	7.66	3.76	40.5	98.2	98.1	41	98.4	98.6	385	385	385	1741
Matrix	0.3	66.4	68.8	42.5	29	37	42	20	36	27	18	36	33	100
Mhyang	0.3	2.1	5.06	2.02	100	99.8	100	100	95.8	100	1490	1380	1490	1490
MountainBike	0.6	7.73	5.15	7.86	100	100	100	100	100	100	228	228	228	228
Shaking	0.2	8.37	86.9	8.17	100	1.64	100	100	1.1	100	365	4	365	365
Singer2	0.3	7.66	15.2	8.16	100	70.8	99.5	100	81.1	100	366	242	361	366
Skating1	0.6	10.3	23.6	7.6	89	64	99.5	65.5	49.8	63.2	309	177	309	400
Subway	0.5	2.46	2.97	2.38	100	100	100	100	98.3	100	175	175	175	175
Trellis	0.5	2.66	6.82	2.11	100	100	100	97.5	95.4	100	569	569	569	569
平均值		27.88	17.96	8.91	80.72	80.10	92.92	74.69	77.12	81.65				

图 1 中 3 组视频的实验数据如表 2 中第 1、5、10 行所示, 结合其他测试视频的对比结果, 本文所提算法在多数测试视频上的性能好于 DSST 算法。其中, 在 Coke 和 Singer2 等测试视频中的

性能也达到与 DSST 算法相近的性能, 在存在遮挡问题的视频中各项评价指标均好于 DSST, 与 SRDCF 相比, 本文提出的算法平均 CLE 降低 5.79, DP 提高 15.27%, OP 提高 3.04%。



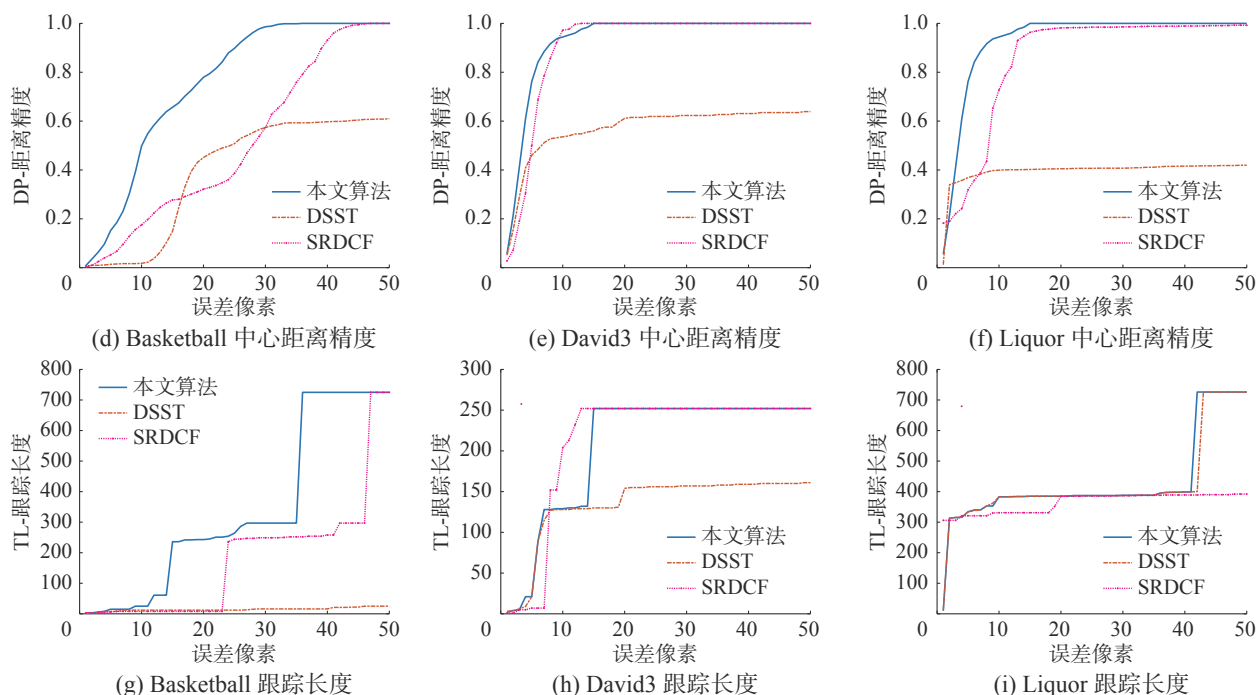


图1 3组背景干扰及目标遮挡下跟踪曲线

Fig. 1 Tracking curves of 3 groups of background interference and target occlusion

图2所示为遇到严重背景干扰及目标遮挡到遮挡结束时的跟踪结果,可以看出,视频Basketball、David3、Liquor 分别在第16帧、第80帧、第722帧发生严重遮挡,分别在第32帧、第90帧、

第736帧遮挡结束,在遮挡结束时本文算法和SRDCF都可以准确地继续跟踪目标,DSST发生跟丢或偏移。结果表明,在发生严重背景干扰及目标遮挡情况下本文算法的跟踪性能更好。



图2 3组背景干扰及目标遮挡下的跟踪结果

Fig. 2 Tracking results of 3 groups of background interference and target occlusion

3) 跟踪速度对比实验

视频跟踪算法在保证跟踪性能的前提下,同时还要考虑跟踪速度的问题,本文算法同两种对

比算法在18组视频序列上的跟踪速度如表3所示,相比于DSST算法平均19 f/s的跟踪速度,本文算法在特征提取阶段加入边缘特征的融合,模

型更新阶段加入区域相似性计算,增加了计算量,从而导致跟踪速度有所下降,平均跟踪速度为 8.95 f/s。但相比于 DSST 算法的另一个改进

SRDCF 算法的平均 4.26 f/s 的跟踪速度,本文算法的速度提高了 1 倍,同时在准确率及跟踪的鲁棒性上有所提高。

表 3 3 种算法跟踪速度对比
Table 3 Tracking speed contrast of 3 algorithms

测试视频	FPS			总帧数
	DSST	SRDCF	本文	
Basketball	19.1	3.08	10.1	725
CarDark	43.5	6.8	17.9	393
Coke	17.9	3.18	6.94	291
Crossing	40.4	8.7	19.4	120
David3	8.17	3.26	5.77	252
Deer	8.74	3.2	5.25	71
Dudek	2.78	2.62	1.86	1 145
Football	30.7	4.16	15.5	362
Football1	35.2	7.36	12.7	74

测试视频	FPS			总帧数
	DSST	SRDCF	本文	
Liquor	3.64	2.98	2.36	1 741
Matrix	30.6	4.89	11.4	100
Mhyang	12.4	3.21	7.34	1 490
MountainBike	12.2	3.03	6.72	228
Shaking	9.13	3.03	5.43	365
Singer2	4.91	3.03	2.89	366
Skating1	17.3	3.14	7.74	400
Subway	38.8	7.76	17	175
Trellis	6.64	3.23	4.8	569

3 结束语

本文对 DSST 相关滤波目标跟踪算法进行了改进,提出了融合边缘特征和 HOG 特征的特征提取策略和基于奇异值特征向量相似度的模型自适应更新策略,通过两种策略的结合实现了一种新的相关滤波目标跟踪算法——多特征自适应相关滤波目标跟踪算法,实验表明本文所提出的算法在存在严重背景干扰和目标遮挡的情况下仍能很好地进行跟踪,提高了算法精度。在严重背景干扰和目标遮挡的情况下仍能准确进行跟踪,提高了算法精度和鲁棒性。

由于特征自适应相关滤波目标跟踪算法采用了边缘特征和 HOG 特征的融合,并在模型更新阶段加入了相似性计算,因此加大了计算量,速度相比于 DSST 有所下降,但仍高于 SRDCF,在准确率和鲁棒性上有所提高,后续研究工作可围绕进一步解决计算上所带来的速度下降问题、进一步改进特征融合的方式,以及将深度学习所提取的特征与现有特征进行自适应融合等方面展开,使改进的算法能够同时兼顾实时性、鲁棒性和准确性的要求。

参考文献:

- [1] CHEN Zhe, HONG Zhibin, TAO Dacheng. An experimental survey on correlation filter-based tracking[J]. arXiv: 1509.05520, 2015.
- [2] 张微, 康宝生. 相关滤波目标跟踪进展综述 [J]. 中国图象

图形学报, 2017, 22(8): 1017–1033.

ZHANG Wei, KANG Baosheng. Recent advances in correlation filter-based object tracking: a review[J]. *Journal of image and graphics*, 2017, 22(8): 1017–1033.

- [3] VALMADRE J, BERTINETTO L, HENRIQUES J F, et al. End-to-end tracking[J]. arXiv: 1704.06036, 2017.

- [4] 刘新卉. 基于深度信息的核相关滤波目标跟踪算法研究 [D]. 深圳: 哈尔滨工业大学, 2017.

LIU Xinhui. Research on kernel correlation filter object tracking algorithm based on depth information[D]. Shenzhen: Harbin Institute of Technology, 2017

- [5] 董艳梅. 基于相关滤波器的目标跟踪技术 [D]. 北京: 北京理工大学, 2015.

DONG Yanmei. Visual object tracking based on correlation filters[D]. Beijing: Beijing Institute of Technology, 2015.

- [6] 朱明敏, 胡茂海. 基于相关滤波器的长时视觉目标跟踪方法 [J]. 计算机应用, 2017, 37(5): 1466–1470.

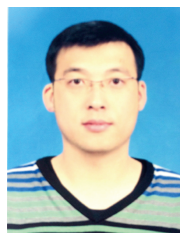
ZHU Mingmin, HU Maohai. Long-term visual object tracking algorithm based on correlation filter[J]. *Journal of computer applications*, 2017, 37(5): 1466–1470.

- [7] BOLME D S, BEVERIDGE J R, DRAPER B A, et al. Visual object representation learning for correlation filter based tracking using adaptive correlation filters[C]//2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Francisco, USA, 2010: 2544–2550.

- [8] 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. Berlin, Germany, 2012: 702–715.
- [9] 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.
- [10] DANELLJAN M, SHAHBAZ K F, FELSBURG M, et al. Adaptive color attributes for real-time visual tracking[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA, 2014: 1090–1097.
- [11] WU Y, LIM J, YANG M H. Online object tracking: A benchmark[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Portland, USA, 2013: 2411–2418.
- [12] DANELLJAN M, HÄGER G, KHAN F, et al. Accurate scale estimation for robust visual tracking[C]//British Machine Vision Conference. Nottingham, UK, 2014:1-11.
- [13] DANELLJAN M, HAGER G, SHAHBAZ K F, et al. Learning spatially regularized correlation filters for visual tracking[C]//Proceedings of the IEEE International Conference on Computer Vision. Santiago, Chile, 2015: 4310–4318.
- [14] LI Y, ZHU J. A scale adaptive kernel correlation filter tracker with feature integration[C]//European Conference on Computer Vision. Zurich, Switzerland, 2014: 254–265.
- [15] DANELLJAN M, HAGER G, KHAN F S, et al. Convolutional features for correlation filter based visual tracking[C]//IEEE International Conference on Computer Vision Workshop. Santiago, Chile, 2015: 621–629.
- [16] 白冰, 钟必能, 欧阳谷. 融合分层卷积特征和尺度自适应核相关滤波器的目标跟踪 [J]. *小型微型计算机系统*, 2017, 38(9): 2062–2066.
- BAI Bing, ZHONG Bineng, OUYANG Gu. Object tracking based on hierarchical convolution feature and scale adaptive kernel correlation filter[J]. *Journal of chinese computer systems*, 2017, 38(9): 2062–2066.
- [17] MA C, HUANG J B, YANG X, et al. Hierarchical convolutional features for visual tracking[C]//IEEE International Conference on Computer Vision. Santiago, Chile, 2015: 3074–3082.
- [18] KALAL Z, MIKOLAJCZYK K, MATAS J. Tracking-learning-detection[J]. *IEEE Transactions on software engineering*, 2011, 34(7): 1409–1422.
- [19] BERTINETTO L, VALMADRE J, HENRIQUES J F, et al. Fully-convolutional siamese networks for object tracking[C]//European Conference on Computer Vision. Amsterdam, Netherlands, 2016: 850–865.
- [20] 洪子泉, 杨静宇. 用于图像识别的图像代数特征抽取 [J]. *自动化学报*, 1992, 18(2): 233–238.
- HONG Ziquan, YANG Jingyu. Algebraic feature extraction of images for recognition[J]. *Acta automatica sinica*, 1992, 18(2): 233–238.

作者简介:



刘威, 副教授, 博士, 中国人工智能学会会员, 中国计算机学会会员, 主要研究方向为深度神经网络、机器学习、矿业系统工程。



靳宝, 硕士研究生, 主要研究方向为强化学习、机器学习。



周璇, 硕士研究生, 中国计算机学会会员, 主要研究方向为深度神经网络、机器学习。