

基于相关滤波的目标重检测跟踪

姜 珊, 张 超, 韩 成, 底晓强*

(长春理工大学 计算机科学技术学院, 吉林 长春 130022)

摘 要: 近年来, 相关滤波方法由于具备运算速度快, 鲁棒性强的优势, 在目标跟踪领域发展迅速。然而, 面对复杂场景时, 现有模型难以满足实际需求。针对背景感知相关滤波方法 (BACF) 在目标发生自身旋转、尺度变换、运动出视野等挑战下, 相关滤波器最大响应值减弱, 造成跟踪精度下降的问题, 提出了一种基于相关滤波的目标重检测跟踪方法。在原有背景感知相关滤波方法的基础上, 引入滤波器响应检测机制, 当判定到相关滤波跟踪结果不可信时, 利用粒子滤波采样策略生成大量粒子, 感知目标状态, 重新确定目标中心位置。在此基础上, 利用自适应尺度估计机制重新计算目标尺度信息, 从而实现目标的重新跟踪。为了验证改进算法的有效性, 实验选取了 OTB2013、OTB2015、VOT2016 共 3 个公开数据集进行测试, 同时与相关滤波及深度学习方法进行对比, 从视频属性、跟踪精确度、算法鲁棒性等角度展示所有算法的性能。实验结果表明: 基于相关滤波的目标重检测跟踪方法在 3 个公开数据集中取得较好的实验结果, 并在目标发生旋转, 尺度变换及运动超出视野的情况下, 有效提高了 BACF 的准确率和成功率。

关键词: 目标跟踪; 相关滤波; 粒子滤波

中图分类号: TP391.4 **文献标志码:** A **DOI:** 10.3788/IRLA20200182

Target redetection method for object tracking based on correlation filter

Jiang Shan, Zhang Chao, Han Cheng, Di Xiaoqiang*

(College of Computer Science and Technology, Changchun University of Science and Technology, Changchun 130022, China)

Abstract: In recent years, due to the advantages of fast speed and strong robustness, correlation filter based methods have been developed rapidly in the tracking community. However, when the existing models are used to deal with complex scenes, it is difficult to meet the requirements of practical application. The background aware correlation filter (BACF) suffers from the maximum response weakening problem when handling the challenging scenes, such as rotation of the target appearance, scale variation and out of view, thus result in inaccurate tracking result. In order to tackle these problems, a target redetection method for visual tracking based on correlation filter was proposed. On the basis of the background aware correlation filter, a correlation response detection mechanism was introduced to judge the quality of the tracking result generated by the correlation filter. After detecting the tracking result was not credible, a particle filter resampling strategy was exploited to generate abundant particles which was beneficial to perceive the state of the target, and the center of the target could be redetected. On this foundation, an adaptive scale estimation mechanism was adopted to calculate the size information for the target, by which the final tracking result could be obtained. To validate the effectiveness of the improved algorithm, the extensive experiments on three public datasets: OTB2013, OTB2015 and VOT2016 were conducted, meanwhile, several state-of-the-art trackers: correlation filter and deep learning based trackers were also chosen as

收稿日期: 2020-10-15; 修订日期: 2020-12-08

基金项目: 国家自然科学基金青年科学基金 (61702051, 61602058)

comparison, and the performance of all the compared trackers was shown from the aspects of annotated video attributes, tracking accuracy, and robustness of the algorithms. Experimental results demonstrate that the proposed target redetection tracker achieve a favorable performance on these three datasets, meanwhile, it effectively improves the accuracy and success rate of the BACF when handling the challenging situations of target rotation, scale variation, and out of view.

Key words: object tracking; correlation filters; particle filters

0 引言

目标跟踪是视频内容分析的重要手段,通过对特定目标的跟踪、定位,实现对目标行为的分析、理解,并为更高阶的任务提供技术支撑,因此在视频监控,人机交互,虚拟现实,自动驾驶等领域中得到广泛的应用。

目标跟踪可分为生成类方法和判别类方法。生成类方法对目标外观模型进行建模,寻找与该模型最为相似的区域作为跟踪结果。判别类方法将目标跟踪视为二元分类问题,利用目标和背景训练分类器,将目标从背景中分离。

粒子滤波属于生成类方法。通过搜索具有最小重构误差的图像区域来寻找潜在的目标位置。Mei^[1]提出的 L1 tracker 将稀疏表示理论与粒子滤波方法相结合,利用目标模板和琐碎模板重构目标外观表示模型,取得了很大的成功,但由于稀疏求解过程计算量庞大,随后, C.bao 提出了加速求解过程的 L1APG^[2]方法,张天柱^[3]提出循环稀疏结构,在频域完成计算。此外, J.Zhan^[4]利用背景信息模板模拟外观模型的变化, Zhang^[5]提出遮挡情况下的模板更新方案。尽管上述工作对于粒子滤波性能有一定幅度的提升,但目标发生严重遮挡,背景与目标相似度较高时,对目标外观模型干扰较大,跟踪精度仍有待提高。

相关滤波属于判别类方法。根据前一帧采集到的图像训练获取相关滤波器,在当前帧中选择使相关滤波器响应值达到最大的区域作为跟踪结果。Bolme 提出的相关滤波 MOSSE^[6]跟踪速度可达 700 fps。随后许多工作从多个角度开展对相关滤波跟踪的探索^[7-12]。Danelljan^[13]提出 SRDCF 方法,在更大的空间区域内搜索目标,有效提升跟踪精度。CCOT 方法^[14]利用 VGG 网络提取目标特征,获取了多重分辨率的特征图,从而提升了跟踪器的性能。后续的 ECO 方法^[15]从特征复杂度、训练集大小、更新方式

等 3 个角度完善 CCOT 模型,显著地提升了时间效率和空间效率。虽然相关滤波类方法在目标跟踪领域中取得了瞩目的成就。但是,在训练和检测阶段所采用的循环移位操作减少了有效样本,引起“边界效应”,影响方法的判别能力。为了解决这一问题, Galoogahi H K^[16]提出了 BACF 方法,创新性地使用了二值矩阵和滤波器相乘的方式,对提取特征区域实现特征选择,保留有效部分,进而实现对背景信息的抑制,比传统滤波器使用余弦窗抑制背景的方式更为有效。然而由于只对滤波器作简单预设限定,未考虑到整个特征的多样性和冗余性,跟踪目标发生旋转、尺度变化或出视野时,相关滤波跟踪器极易丢失跟踪目标或再次跟踪困难。Xu^[17]改进了 BACF 的特征选择方式,即在原有的二值矩阵的基础上,对特征采用稀疏化表达,抑制无效背景信息干扰。Dai^[18]构造了自适应空域正则滤波器,引入了自适应空间正则化项,取代 BACF 中的固定空间约束形式,可根据目标外观变化自适应地调整滤波参数,解决了滤波模型退化问题。Sun^[19]提出了一种可靠性学习方案,将滤波器分解为多个基滤波器,对应图像中的各个子区域,利用局部响应一致性约束,获取基滤波器中的可靠的特征。

以上工作在 BACF 方法的基础上,通过添加正则项约束滤波器或背景,提升了跟踪精度,但额外引入多维空间约束^[20],增加了大量的计算代价。此外,由于相关滤波器更新过程依据当前跟踪结果,不可信的跟踪结果会使得误差累积,最终导致目标丢失。

为了解决上述问题,提出一种基于相关滤波的目标重检测跟踪方法,以 BACF 为基准方法,借助粒子滤波样本生成方式的优势,弥补 BACF 方法由于误差累积引起跟踪失败的不足,完成对不可信跟踪结果的重新检测。融合了判别方法和生成方法的各自优势,提高了基准方法在旋转、尺度变化、出视野等复杂情况下的跟踪精度。使用相关滤波进行跟踪,若检测到相关滤波最大响应值小于预设阈值,采用粒子滤波重

新检测。以当前帧的上一帧跟踪结果为中心,依照高斯分布生成大量粒子,分别计算对应图像块的特征与相关滤波器的响应值,选择具有最大响应值的图像块作为重新跟踪后的目标中心位置。结合自适应尺度机制为目标分配合理尺度信息,完成对目标的重新检测,并利用该结果更新相关滤波器。

1 方法描述

基于目标重检测机制的相关滤波方法主要分为

4 个步骤:(1) 利用视频初始帧中目标信息,进行特征提取并训练相关滤波器;(2) 使用相关滤波器计算图像特征响应值,相关滤波器最大响应值与预设阈值 θ 作比较,若不低于 θ , 取对应位置作为跟踪结果;(3) 若相关滤波器最大响应值低于 θ , 采用粒子滤波对目标重新定位,并进行尺度估计;(4) 根据粒子滤波跟踪结果重新训练相关滤波器。文中方法流程图如图 1 所示。

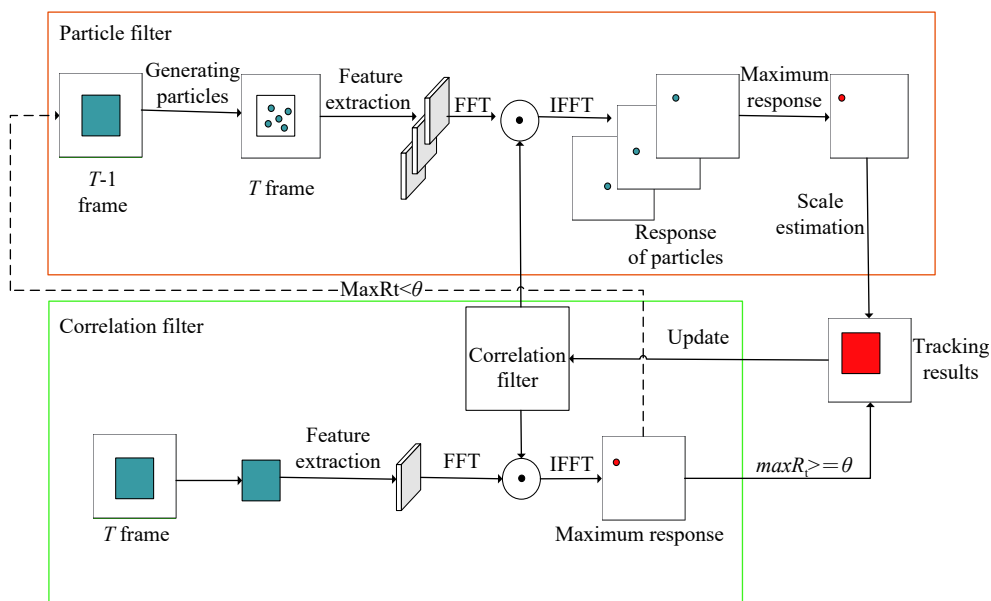


图 1 基于相关滤波的目标重检测跟踪方法流程图

Fig.1 Flowchart of target redetection method for object tracking based on correlation filter

1.1 基于背景感知的相关滤波跟踪

为了解决循环移位操作引起的边界效应,增强滤波器的判别能力, BACF 方法选择了更大的图像区域训练相关滤波器。此外,为了避免正样本中包含过多的背景信息,对正样本区域进行“背景剔除”。这一过程可以描述为:

$$E(h) = \frac{1}{2} \sum_{j=1}^D \left\| y - \sum_{k=1}^K \alpha_k^T \mathbf{P} x_k [\Delta \tau_j] \right\|_2^2 + \frac{\lambda}{2} \sum_{k=1}^K \|\alpha_k\|_2^2 \quad (1)$$

式中: y 为期望输出; α_k 为第 k 个通道的滤波器; x_k 为第 k 个通道的样本; \mathbf{P} 为一个由 0 和 1 构成,大小为 $D \times T$ 的矩阵; $[\Delta \tau_j]$ 为循环移位算子; λ 为惩罚滤波器的正则项系数。矩阵 \mathbf{P} 的引入,能够实现对正样本的准确提取,从而提高滤波器的判别能力。在视频序列初始帧中,搜索窗的大小为 $M \times N$ 。

为了提升方法效率,计算过程选择在频率域内进行,公式 (1) 可转化为:

$$\begin{cases} E(\alpha, \hat{g}) = \frac{1}{2} \|\hat{y} - \hat{X} \hat{g}\|_2^2 + \frac{\lambda}{2} \|\alpha\|_2^2 \\ s.t. \hat{g} = \sqrt{S} (F P^T \otimes I_D) \alpha \end{cases} \quad (2)$$

式中: \hat{X} 为 $S \times DS$ 的特征矩阵; I_D 为一个 $D \times D$ 的单位矩阵; F 为频域计算得到的基本正交基,大小为 $S \times S$; 辅助变量 g 经过傅里叶变换后表示为 \hat{g} 。

为了获取全局最优解,可以采用增广拉格朗日 (ALM) 方法将公式 (2) 进一步转化为:

$$\begin{aligned} L(h, \alpha, \hat{g}) = & \frac{1}{2} \|\hat{y} - \hat{X} \hat{g}\|_2^2 + \frac{\lambda}{2} \|\alpha\|_2^2 + \\ & \hat{\xi}^T (\hat{g} - \sqrt{S} (F P^T \otimes I_D) \alpha) + \\ & \frac{\mu}{2} \|\hat{g} - \sqrt{S} (F P^T \otimes I_D) \alpha\|_2^2 \end{aligned} \quad (3)$$

式中: $\hat{\xi}^T$ 为经过傅里叶变换的拉格朗日向量; μ 为用来控制 ALM 方法收敛速度的惩罚因子。公式 (3) 最优解的过程由交替求解方法^[21](ADMM) 迭代方法完成, 通过将全局优化问题拆成两个子问题, 降低计算规模。根据公式 (4) 和公式 (5) 分别求解 α^* 和 \hat{g}^* 。

$$\alpha^* = \arg \min_{\alpha} \left\{ \frac{\lambda}{2} \|\alpha\|_2^2 + \hat{\xi}^T (\hat{g} - \sqrt{S} (FP^T \otimes I_K) \alpha) + \frac{\mu}{2} \|\hat{g} - \sqrt{S} (FP^T \otimes I_D) \alpha\|_2^2 \right\} = \left(\mu + \frac{\lambda}{\sqrt{S}} \right)^{-1} (\mu \hat{g} + \varsigma) \quad (4)$$

$$\hat{g}^* = \arg \min_{\hat{g}} \left\{ \frac{1}{2} \|\hat{y} - \hat{X} \hat{g}\|_2^2 + \hat{\xi}^T (\hat{g} - \sqrt{S} (FP^T \otimes I_K) \alpha) + \frac{\mu}{2} \|\hat{g} - \sqrt{S} (FP^T \otimes I_K) \alpha\|_2^2 \right\} \quad (5)$$

简化公式 (5) 中的矩阵求逆运算, 变形成得到公式 (6), 最终可以求得最优解。利用公式 (7) 更新相关滤波器, η 代表预先设置的滤波器的学习率。至此, 可以根据相关滤波器定位目标所在位置。

$$\hat{g}(s)^* = (\hat{I}(s) \hat{I}(s)^T + S \mu I_K)^{-1} (\hat{y}(s) \hat{I}(s) - S \hat{\xi}(s) + \mu \hat{\alpha}(t)) = \frac{1}{\mu} (S \hat{y}(s) \hat{I}(s) - \hat{\xi}(s) + \mu \hat{\alpha}(t)) - \frac{\hat{I}(s)}{\mu (\hat{I}(s) + S \mu)} (S \hat{y}(s) \hat{I}(s) \hat{I}(s)^T - \hat{I}(s)^T \hat{\xi} + \mu \hat{I}(s)^T \alpha) \quad (6)$$

$$\hat{x}_{\text{model}}^{(t)} = (1 - \eta) \hat{x}_{\text{model}}^{(t-1)} + \eta \hat{x}^{(t)} \quad (7)$$

1.2 粒子滤波重检测跟踪

粒子滤波利用大量带有权重的粒子可以实现对目标状态的动态拟合, 得益于这一特性, 将粒子滤波采样机制引入至 BACF 方法, 以缓解目标发生自身运动、尺度变化等情况时, 造成滤波器最大响应值弱化造成的跟踪性能下降甚至丢失目标的现象。

跟踪任务中, 目标在 t 时刻的状态向量表示为 x_t , 观测向量表示为 z_t , 目标跟踪问题可以视为根据目标在 $t-1$ 时刻的状态信息预测目标在 t 时刻的状态信息, 这一过程可以表示为:

$$x_t = \arg \max p(x_t | z_{1:t}) = \arg \max \int p(x_t | z_{1:t}) p(x_{t-1} | z_{1:t-1}) dx_{t-1} \quad (8)$$

根据贝叶斯推论, 公式 (8) 进一步可以表示为:

$$p(x_t | z_{1:t}) = \frac{p(z_t | x_t) p(x_{t-1} | z_{1:t-1})}{p(z_t | z_{1:t-1})} \quad (9)$$

粒子滤波方法利用了 n 个具有权重的粒子采样获得先验状态分布 $p(x_t | z_{1:t})$, 所有粒子权重的总和为 1。粒子权重的更新方式如下:

$$w_t^i = w_{t-1}^i \frac{p(z_t | x_t)}{q(x_t^i | x_{0:t-1}^i, z_{1:t})} \quad (10)$$

在实际应用中, 通常采用 $q(x_{t-1} | x_t)$ 作为建议分布代替 $q(x_t^i | x_{0:t-1}^i, z_{1:t})$ 简化计算, 因此公式 (10) 可以简化为:

$$w_t^i = w_{t-1}^i p(z_t | x_t^i) \quad (11)$$

相关滤波方法定位目标依赖于相关滤波器最大响应, 目标在运动过程中受到外界干扰引起外观产生不可知变化时, 相关滤波最大响应值也会受到影响, 此时, 需要比较最大响应值与阈值 θ 的关系, 若相关滤波器最大响应值大于等于 θ , 继续使用相关滤波器跟踪, 若小于 θ , 则采用离子滤波器生成可靠粒子。

利用粒子滤波重新检测目标 (如图 1 中橘黄色矩形框中流程), 由于相邻帧间目标变化程度比较轻微, 因此, 以上一帧跟踪结果所在位置为中心, 在当前帧中依照高斯分布生成 M 个粒子, 对应 M 个与相关滤波器中搜索窗等大的图像块。提取每个图像块对应的 HOG 特征, 并在频域中计算其与相关滤波器的相关响应值, 最后选取具有最大响应值的粒子作为预测目标中心。粒子滤波的引入, 可以实现在更大范围内搜索最优候选样本, 为实现准确跟踪提供了更多可能性。

1.3 目标尺度估计与更新

粒子滤波获取目标中心后, 需要对目标的尺度信息进行估计。目标在初始帧的尺寸大小表示为 $size_1 = (h_1, w_1)$, 由于视频序列具有时空连续性, 同一目标在连续两帧之间的运动幅度较小, 可根据目标在前一帧中的尺度信息进行计算:

$$d_t = \frac{\max R_t}{\max R_{t-1}} - \frac{\max R_{t-1}}{\max R_{t-2}} \quad (12)$$

式中: $\max R_t$ 表示在第 t 帧中相关滤波器的最大响应值。在粒子滤波重检测阶段, 引入两个尺度阈值 ϕ 、 ψ 、判别目标尺度的变化趋势。若 $d_t > \phi$, 表示目标尺度小于前一帧尺寸; 若 $d_t < \psi$, 那么则说明目标尺度逐渐变大。尺度变化的表达方式如下:

$$size_t = size_t \cdot s_t, \quad s_t = \begin{cases} 0.95, & d_t > \phi \\ 1.05, & d_t < \psi \\ 1, & \text{其他} \end{cases} \quad (13)$$

式中: $size_t$ 表示目标在第 t 帧的尺寸; s_t 表示尺度缩放因子。文中 ϕ 的值设为 0.1, ψ 设为 -0.1。通过上述设置, 完善了目标重新定位后的尺度估计, 有利于提高跟踪方法的成功率。

2 实验结果与分析

为了验证文中方法的有效性, 使用 OTB2013^[22] 和 OTB2015^[23] 以及 VOT2016^[24] 3 个标准数据集对文中方法进行评估。文中所有实验是在 Windows10 以及 Matlab2017b 的软件环境, 以及 Intel i7-8700k CPU (3.7 GHz), 32 G 内存的硬件环境下完成的。

2.1 实验数据集

OTB2013 数据集包含 50 个视频图像序列, 在此基础上, OTB2015 数据集将数目扩充至 100 个。这两个数据集利用 11 种属性标注视频序列, 包括低分辨率 (LR), 超出平面旋转 (OPR), 尺度变化 (SV), 遮挡 (OCC), 外观形变 (DEF), 运动模糊 (MR), 快速运动 (FM), 平面内旋转 (IPR), 出视野 (OV), 背景干扰 (BC), 光照变化 (IV)。每个视频可以包含多重属性。OTB2013 和 OTB2015 数据集准确率和成功率两个指标完成对 6 种方法的综合评估。准确率是方法跟踪结果的中心像素与真实值中心像素的欧氏距离, 反映跟踪框与真实框之间的偏离程度。成功率使用跟踪框与真实框交集与并集之间的比值, 描述跟踪框对跟

踪目标的覆盖程度。

VOT2016 数据集包含 60 组视频序列, 并对以 6 种属性对所有序列加以标注, 分别为: 相机抖动、光照变化、目标尺寸变化、目标位移模糊和未退化。VOT2016 数据集主要采用的评价指标有: 精确度 A (Accuracy) 反应预测框与真实标注框之间的覆盖情况, 鲁棒性 R (Robustness) 统计跟踪算法丢失目标的次数。预期平均覆盖率 EAO (Excepted average overlap) 计算跟踪算法在同一视频序列非重置重叠区域的期望。此外, VOT2016 数据库在评测算法时, 采用了跟丢重新启动机制, 如果某一帧中检测到跟踪器完全覆盖不到目标, 5 帧之后将利用真实值重新初始化算法。

2.2 基准算法对比分析

文中方法是基于 BACF 的改进工作, 因此, 通过对比两种算法的实验结果来验证改进方法的有效性, 并对选取比较有代表性的序列进行分析, 如图 2 所示, 红色矩形框表示文中方法, 绿色表示基准 BACF 方法。在视频序列 ironman 中, 目标发生了较为明显的尺度及旋转变化, 基准算法在第 11 帧时可以成功跟踪目标, 57 帧后, 基准算法完全丢失目标, 直至序列截止, 基准方法并未重新找到目标正确位置。这说明文中重检测机制预判跟踪结果的有效性, 通过预设阈值提前判断跟踪策略, 有效避免了由于相关滤波跟踪结果不可信导致的目标持续丢失问题。

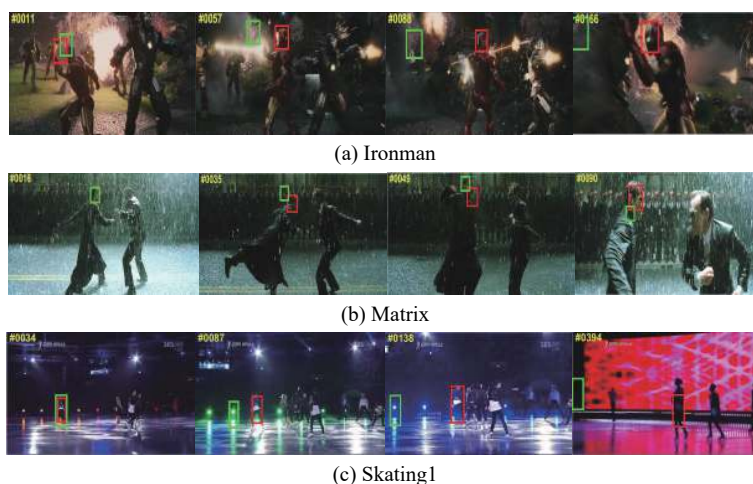


图 2 文中方法 (红色) 与基准方法 (绿色) 对比

Fig.2 Comparison results of the proposed method (red) and the baseline (green)

Matrix 序列展示了目标在光照变化, 运动模糊, 面内旋转等挑战下的直观效果。目标外观发生丰富多样的变化, 对算法捕捉模型变化能力具有一定的要求, 在该视频序列的 49 帧, 文中方法出现了轻微偏离, 但这一情况在后续帧中得到了改善, 证明了粒子滤波采样机制在跟踪任务中的优势。

Skating 序列中目标受到了相似背景信息及光照变化等干扰, 目标在这一序列中, 跟踪目标发生的较为明显的肢体变化, 目标从远处逐步进入到人群之后, 随后离开人群。文中方法可以有效应对源自不同外观变化, 从而得到较为稳定的跟踪结果。

2.3 同类方法对比分析

2.3.1 定量分析

(1) OTB2013 数据集跟踪结果

OTB2013 数据集上选取了 5 种经典方法作为对比, 分别为: ECO^[15]、BACF^[16]、SRDCF^[13]、DSST^[10]、KCF^[25]。根据实验结果绘制的准确率曲线与成功率曲线如图 3(a)、3(b) 所示。文中方法在准确率和覆盖率上均取得了最高的分数。准确率上, 文中方法以 0.3% 的细微优势领先于 ECO 方法, 对其他方法的领先差距按照 BACF、SRDCF、DSST、KCF 方法的次序依次增大。在成功率上, 文中方法领先 BACF 方法及 ECO 方法 0.5%, 较其余三种方法优势明显。各个属性的跟踪结果如表 1 所示。表中的每个单元包含两个数据, 前者为准确率, 后者为成功率。下划线标志表示在所在某属性下指标最好的方法, 加粗字体表示指标

次好的方法。在超出平面旋转 (OPR) 属性上准确率超出 BACF 方法 4.8%, 成功率领先 4.3%。文中方法在目标出视野 (OV) 属性准确率高于 BACF 方法 1.9%, 成功率领先 3%。验证了引入粒子滤波重新检测机制的有效性。

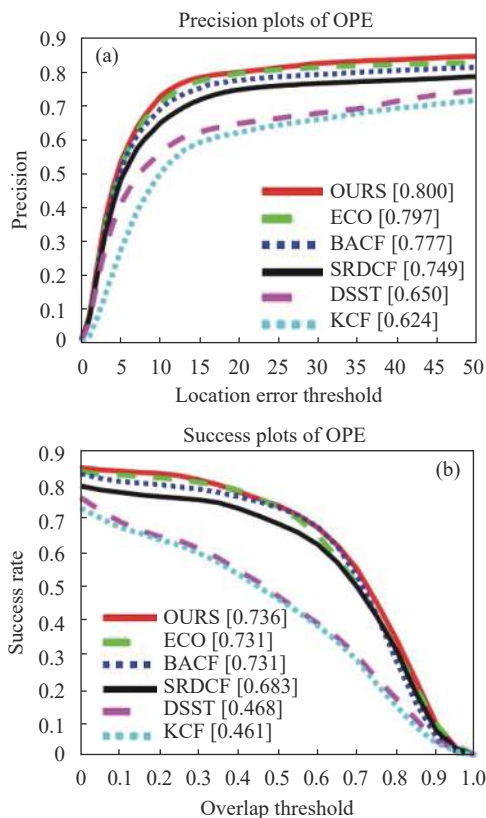


图 3 OTB2013 数据集的准确率及成功率曲线

Fig.3 Overall precision and success plots on OTB2013

表 1 OTB2013 数据集各属性对比

Tab.1 Attributes comparisons on OTB2013 dataset

Attribute/Name	Proposed method	ECO	DSST	SRDCF	KCF	BACF
LR	75.5/68.9	<u>76.0/66.2</u>	54.8/28.8	63.8/60.4	54.6/30.1	70.8/63.9
OPR	77.1/71.1	75.9/69.6	59.5/46.6	69.5/62.8	60.2/47.3	72.3/66.8
SV	78.5/72.4	78.0/71.0	61.6/38.8	72.4/65.8	57.7/37.2	73.8/69.2
OCC	72.2/64.3	<u>77.0/69.3</u>	62.8/43.3	69.2/59.7	59.2/45.1	72.0/64.2
DEF	82.3/70.7	76.5/67.9	62.3/45.9	71.4/62.9	60.5/45.8	76.9/70.5
MB	71.5/68.1	71.0/65.8	52.0/38.2	<u>73.7/66.7</u>	55.4/42.3	69.1/66.2
FM	73.1/68.4	<u>76.1/70.3</u>	49.9/36.1	74.3/68.8	53.9/39.6	73.7/70.3
IPR	<u>73.3/67.7</u>	67.3/59.8	61.6/49.1	62.3/56.6	58.9/46.0	68.5/62.6
OV	71.1/62.0	<u>73.1/61.7</u>	41.6/32.3	57.9/51.9	44.1/37.4	69.2/59.0
BC	74.4/69.1	<u>76.6/71.7</u>	68.9/54.6	73.6/63.6	62.5/50.0	71.5/66.1
IV	<u>76.5/70.0</u>	71.6/66.7	70.8/49.0	74.4/66.4	67.0/46.4	73.3/69.6

(2) OTB2015 数据集跟踪结果

图 4 展示了文中方法与 ECO^[15]、BACF^[16]、SRDCF^[13]、DSST^[10]、KCF^[25] 等方法在 OTB2015 数据集上的准确度曲线与成功率曲线，文中方法在 OTB2015 数据集上准确度曲线逊色于 ECO 方法，覆盖率曲线表现良好。

6 种方法在 OTB2015 数据集中的 11 种属性的实验结果如表 2 所示。下划线标志表示在所在某属性下指标最好的方法，加粗字体表示指标次好的方法。文中方法在 11 个属性上的准确率测试中，拥有 6 个最优，3 个次优；在成功覆盖率测试中，6 个属性达到最优，4 个属性表现次优。处理尺度变化 (SV) 属性的视频序列时，文中方法准确率先 BACF 方法 4.6%，成功率领先 3.4%。平面内旋转 (IPR) 属性成功率领先 BACF 方法 4.2%，成功率领先 4.8%。实验数据证明了基于相关滤波的目标重检测跟踪方法具有很好的鲁棒性。

(3) VOT2016 数据集跟踪结果

图 5(a) 展示了文中方法和 BACF、SRDCF、DSST、ECO、KCF 以及 MOSSE 共 7 种方法在 VOT2016 数据集中 60 组序列的 EAO 曲线对比结果，图 5(b) 呈现了 EAO 数值排名情况。表 3 和表 4 分别记录了 7 种算法在 VOT2016 数据集中对应 6 种视频属性的精确度得分和鲁棒性得分，每种属性得分第一名用下划线标注，第二名用粗体标注。

观察图 5(a) 中 EAO 曲线，可以发现文中方法 (红色) 明显优于基准 BACF 方法 (绿色)，验证了粒子滤波重新检测机制的有效性。图 5(b) 给出了各个算法

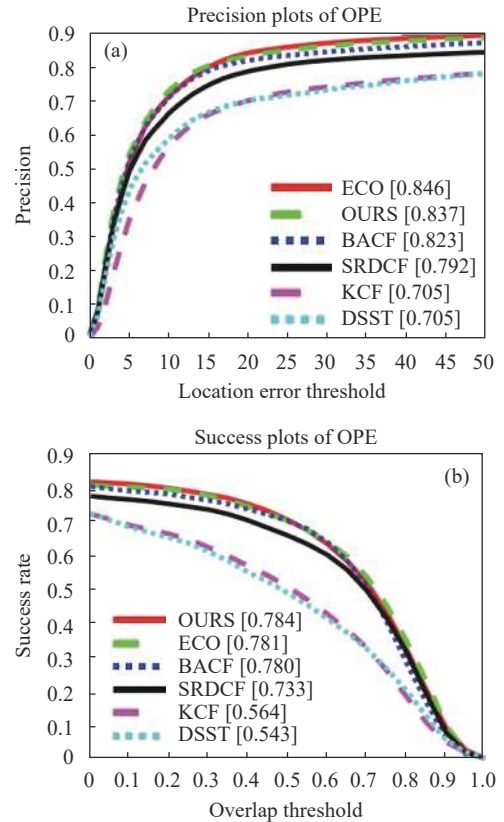


图 4 OTB2015 数据集的中心误差及覆盖率曲线

Fig.4 Overall precision and success plots on OTB2015

EAO 得分排序，文中方法的表现仅次于使用深度特征的 ECO 方法，较其他使用手工特征的相关滤波方法仍具有一定优势。此外，表 3 及表 4 中记录数据表明，在精确度及鲁棒性测试中，文中方法均有良好的表现，可以很好应对视频序列中相机抖动、目标抖动及尺寸变化等挑战。

表 2 OTB2015 数据集各属性对比

Tab.2 Attributes comparisons on OTB2015 dataset

Attribute/Name	Proposed method	ECO	DSST	SRDCF	KCF	BACF
LR	78.4/70.3	<u>80.1/66.2</u>	56.5/32.1	66.8/62.8	57.7/36.5	71.8/65.9
OPR	<u>82.8/76.0</u>	81.7/74.4	66.3/51.2	75.0/67.6	67.6/53.7	78.0/71.8
SV	<u>82.6/75.3</u>	80.8/73.4	66.8/43.1	75.1/67.5	64.6/42.9	78.0/71.9
OCC	75.9/71.5	<u>79.9/74.8</u>	62.8/49.0	72.2/66.8	63.5/53.2	72.9/69.5
DEF	<u>83.1/73.9</u>	81.2/73.9	59.8/46.7	74.8/67.7	63.0/52.0	78.9/71.3
MB	<u>79.5/78.1</u>	77.9/75.2	59.1/50.1	77.5/73.7	60.5/52.4	74.8/73.7
FM	78.3/73.7	<u>80.3/75.1</u>	59.0/46.7	77.3/72.0	62.9/50.4	79.5/76.1
IPR	<u>79.5/72.6</u>	75.3/66.5	71.2/56.3	72.2/64.4	71.0/57.2	75.3/67.8
OV	73.0/66.9	<u>76.3/67.0</u>	45.7/38.0	59.0/53.5	48.7/42.7	72.4/64.6
BC	81.0/76.8	<u>83.6/78.5</u>	72.0/57.2	78.4/70.3	71.6/60.2	77.6/73.5
IV	<u>81.5/77.7</u>	78.7/75.4	73.3/55.3	78.1/73.7	71.3/53.8	80.3/77.9

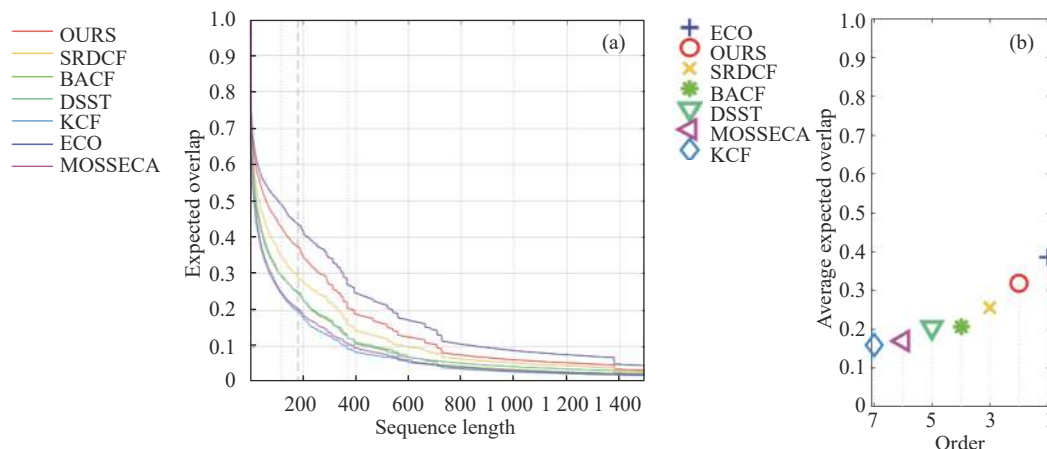


图 5 文中方法与同类方法在 VOT2016 数据集上的对比结果

Fig.5 Comparisons results of the proposed method against correlation filter based methods on VOT2016

表 3 VOT2016 数据集精确度得分

Tab.3 Accuracy scores on VOT2016 dataset

Name/Attribute	Camera motion	Illumination change	Motion change	Occlusion	Size change	Mean	Weighted mean
Proposed method	0.5900	0.6627	<u>0.5147</u>	<u>0.4886</u>	<u>0.5215</u>	<u>0.5641</u>	<u>0.5727</u>
BACF	0.4986	<u>0.6924</u>	0.4412	0.4413	0.4586	0.5265	0.4976
SRDCF	<u>0.5909</u>	0.6872	0.4900	0.4206	0.5053	0.5415	0.5377
DSST	0.5544	0.6765	0.4896	0.403	0.5194	0.383	0.5367
KCF	0.5034	0.4540	0.4219	0.4638	0.3625	0.4517	0.4610
ECO	0.5858	0.6597	0.4918	0.4254	0.5151	0.5465	0.5503
MOSSECA	0.4708	0.3911	0.3691	0.3639	0.3392	0.4086	0.4287

表 4 VOT2016 数据集鲁棒性得分

Tab.4 Robustness scores on VOT2016 dataset

Name/Attribute	Camera motion	Illumination change	Motion change	Occlusion	Size change	Mean	Weighted mean
Proposed method	18.00	2.00	22.00	<u>16.00</u>	10.00	13.6667	15.6584
BACF	42.00	9.00	8.00	49.00	18.00	28.00	32.5734
SRDCF	34.00	8.00	31.00	20.00	20.00	21.1667	24.1220
DSST	49.00	6.00	<u>6.00</u>	50.00	18.00	28.00	28.6667
KCF	54.00	8.00	56.00	24.00	28.00	34.00	40.9333
ECO	<u>15.00</u>	<u>0.00</u>	12.00	17.00	<u>7.00</u>	<u>9.1667</u>	<u>10.0788</u>
MOSSECA	55.00	11.00	52.00	20.00	36.00	32.3333	38.1698

2.3.2 定性分析

图 6 展示了 OTB2013 和 OTB2015 数据集的部分跟踪结果, 利用颜色对不同方法加以区分。

Box3 序列测试是相似背景干扰和目标出视野时算法的跟踪效果。在该组序列运行一段时间后, 只有文中方法可以准确跟踪目标, 其他方法跟踪框出现了一定程度的偏离, 直至丢失目标。Jumping 序列检测算法在快速运动、运动模糊等挑战下的跟踪结果。KCF 方法、DSST 方法在该序列表现欠佳。Human3、

Human9 序列测试方法在目标发生形变和尺度变化情况下的性能。文中方法在这两组序列上均实现了对目标的准确跟踪。Dragonbaby 序列中, 目标在与玩偶的互动中具有平面内旋转、平面外旋转、尺度变化等测试属性, 文中方法可以很好地捕捉跟踪目标旋转时的特征变换。Bird1 序列中候鸟在穿越云层时被完全遮挡, 目标出视野, 后续序列中目标又重新出现, 文中方法实现了对候鸟的准确跟踪, 再次验证了重新检测机制的有效性^[26]。Tiger2 序列具有遮挡、形变等测试属

性,所有方法均可以成功跟踪,但跟踪框的覆盖率仍有提升空间。Skiing 序列中,滑雪运动员在光照强烈

变化的背景下高速地旋转运动,所有方法该序列上均发生漂移,甚至完全丢失目标。

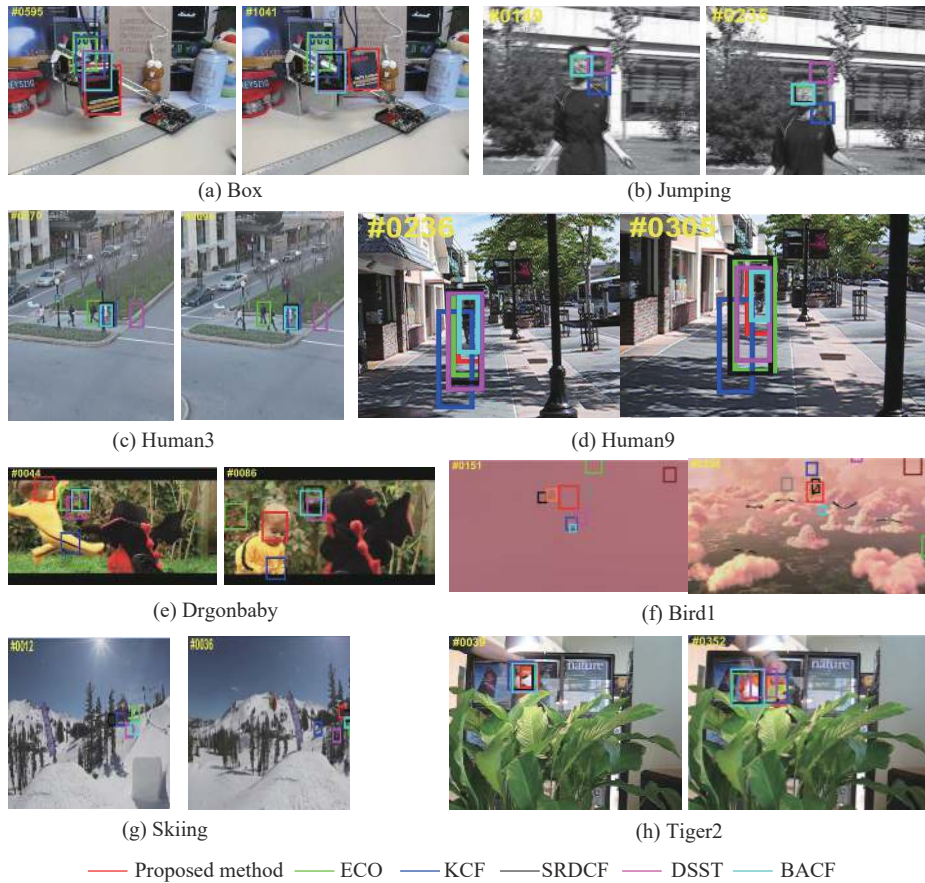


图 6 OTB 数据集中 8 组跟踪序列上的跟踪结果

Fig.6 Comparisons results of eight sequences on OTB dataset

2.4 与其他方法对比分析

图 7 展示了文中方法与基准方法 BACF、深度学习方 法 SimaFC^[27]、SiamRPN^[28]、SiamVGG^[29]、Up-

dateNet^[30], 以及 MCPF^[31] 方法在 VOT2016 数据集的对比结果,依照图 7(b) 中的排序,排在第一位的方法是 SiamRPN,文中方法位于第三位,相较于其他方法

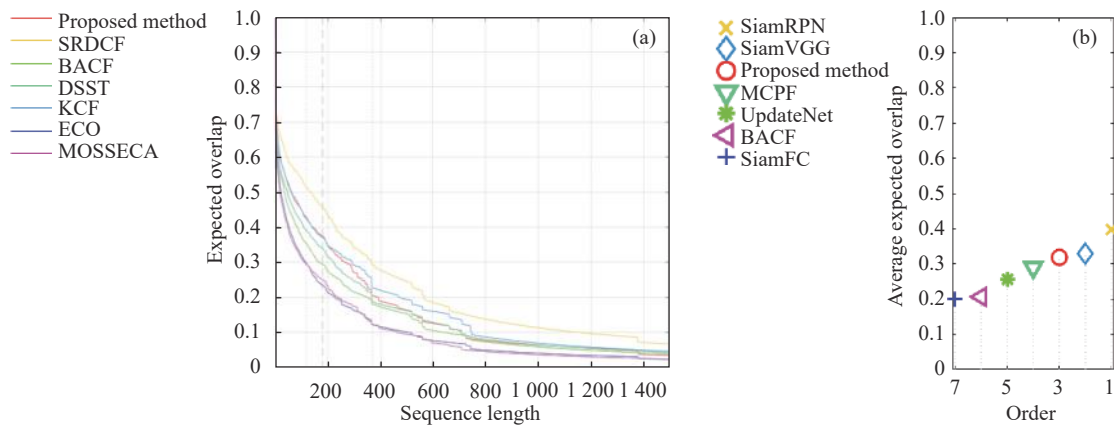


图 7 文中方法与其他方法在 VOT2016 数据集的对比结果

Fig.7 Comparison results of the proposed method and the others on VOT2016 dataset

具备一定优势。

2.5 粒子数目对方法速度的评估

为了评估粒子数目对算法速度的影响,选取了不同粒子数进行实验,并利用方法在 OTB2015 数据集上的平均每秒帧数 (FPS) 来评估算法速度,实验结果如表 5 所示。

表 5 粒子数目对算法速度的影响

Tab.5 Effect of particle numbers on speed performance

Particle numbers	20	30	40	50
Speed/FPS	31.2	28.3	26.8	19.9

2.6 方法局限性分析

当目标遇到与相似背景干扰或被遮挡时,文中方法未能很好的实现准确跟踪,如图 8 所示,造成上述问题的主要有两点原因:第一是相关滤波采用的手工特征并不能为目标提供多层次的特征表达,影响了模型的判别能力。第二是遮挡物信息被视为目标的一部分,更新滤波器的过程中传导了偏差信息,导致跟踪失败。未来的研究工作可通过结合深度模型中的注意力机制更多关注目标自身的特征,以及通过引入未被污染的跟踪模板作为参考,弥补文中方法的不足之处。



图 8 跟踪失败样例

Fig.8 Failure tracking cases

3 结论

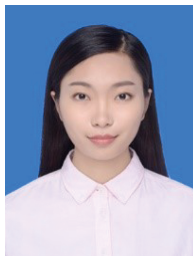
针对 BACF 跟踪中目标由于快速运动、自身旋转等原因引起的跟踪精度下降问题,提出一种基于相关滤波的粒子滤波重新检测跟踪方法。利用相关滤波跟踪目标,相关滤波响应低于一定阈值,使用粒子滤波对目标重新检测,获取目标中心位置,结合自适应尺度估计机制,为目标分配尺度信息,得到重新跟踪结果并根据该结果更新相关滤波。为了验证方法的有效性,在 OTB2013、OTB2015 及 VOT2016 数据集

上同相关滤波方法及深度学习方法进行对比,实验证明文中方法对旋转、尺度变换、运动出视野等挑战具有良好的鲁棒性。下一步工作计划采用深度网络丰富外观特征表达,提升跟踪精度,并引入时间空间信息以应对长序列跟踪。

参考文献:

- [1] Mei X, Ling H. Robust visual tracking using L1 minimization[C]//IEEE, International Conference on Computer Vision. DBLP, 2009: 1436-1443.
- [2] Bao C, Wu Y, Ling H, et al. Real time robust L1 tracker using accelerated proximal gradient approach[C]//IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2012: 1830-1837.
- [3] Zhang T, Bibi A, Ghanem B. In defense of sparse tracking: Circulant sparse tracker[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 3880-3888.
- [4] Zhan J, Wu H, Zhang H, et al. Cascaded probabilistic tracking with supervised dictionary learning [J]. *Signal Processing: Image Communication*, 2015, 39: 212-225.
- [5] Zhang T, Jia K, Xu C, et al. Partial occlusion handling for visual tracking via robust part matching[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014: 1258-1265.
- [6] Bolme D S, Beveridge J R, Draper B A, et al. Visual object tracking using adaptive correlation filters[C]//2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2010: 2544-2550.
- [7] Danelljan M, Shahbaz Khan F, Felsberg M, et al. Adaptive color attributes for real-time visual tracking[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014: 1090-1097.
- [8] Ma C, Huang J B, Yang X, et al. Hierarchical convolutional features for visual tracking[C]//Proceedings of the IEEE International Conference on Computer Vision, 2015: 3074-3082.
- [9] Li Y, Zhu J. A scale adaptive kernel correlation filter tracker with feature integration[C]//European Conference on Computer Vision, 2014: 254-265.
- [10] Danelljan M, Häger G, Khan F, et al. Accurate scale estimation for robust visual tracking[C]//British Machine Vision Con-

- ference, 2014.
- [11] Hong Z, Chen Z, Wang C, et al. Multi-store tracker (muster): A cognitive psychology inspired approach to object tracking[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015: 749-758.
- [12] Ma C, Yang X, Zhang C, et al. Long-term correlation tracking[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 5388-5396.
- [13] Danelljan M, Hager G, Shahbaz Khan F, et al. Learning spatially regularized correlation filters for visual tracking[C]//Proceedings of the IEEE International Conference on Computer Vision, 2015: 4310-4318.
- [14] Danelljan M, Robinson A, Khan F S, et al. Beyond correlation filters: Learning continuous convolution operators for visual tracking[C]//European Conference on Computer Vision, 2016: 472-488.
- [15] Danelljan M, Bhat G, Khan F S, et al. ECO: Efficient Convolution Operators for Tracking[C]//CVPR, 2017, 1(2): 3.
- [16] Galoogahi H K, Fagg A, Lucey S. Learning background-aware correlation filters for visual tracking[C]//ICCV, 2017: 1144-1152.
- [17] Xu T, Feng Z H, Wu X J, et al. Learning adaptive discriminative correlation filters via temporal consistency preserving spatial feature selection for robust visual object tracking[J]. *IEEE Transactions on Image Processing*, 2019, 28(11): 5596-5609. DOI:10.1109/TIP.2019.2919201.
- [18] Dai K, Wang D, Lu H, et al. Visual tracking via adaptive spatially-regularized correlation filters[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019: 4670-4679.
- [19] 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.
- [20] Mueller M, Smith N, Ghanem B. Context-aware correlation filter tracking[C]//Proc of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, 2(3): 6.
- [21] Wu Y, Lim J, Yang M H. Object tracking benchmark[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(9): 1834-1848.
- [22] Wu Y, Lim J, Yang M H. Online object tracking: A benchmark[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2013: 2411-2418.
- [23] Wu Y, Lim J, Yang M H. Object tracking benchmark [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(9): 1834-1848.
- [24] Xu T, Feng Z H, Wu X J, et al. Learning adaptive discriminative correlation filters via temporal consistency preserving spatial feature selection for robust visual object tracking[J]. *IEEE Transactions on Image Processing*, 2019, 28(11): 5596-5609.
- [25] 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.
- [26] Ma C, Yang X, Zhang C, et al. Long-term correlation tracking[C]//Proceedings of the IEEE conference on computer vision and pattern recognition, 2015: 5388-5396.
- [27] Bertinetto L, Valmadre J, Henriques J F, et al. Fully-convolutional siamese networks for object tracking[C]//European Conference on Computer Vision, 2016: 850-865.
- [28] Li B, Yan J, Wu W, et al. High performance visual tracking with siamese region proposal network[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8971-8980.
- [29] Li Y, Zhang X. SiamVGG: Visual tracking using deeper siamese networks[J]. arXiv: Computer Vision and Pattern Recognition, 2019.
- [30] Zhang T, Xu C, Yang M H. Multi-task correlation particle filter for robust object tracking[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 4335-4343.
- [31] Zhang L, Gonzalez-Garcia A, Weijer J, et al. Learning the model update for siamese trackers[C]//Proceedings of the IEEE International Conference on Computer Vision, 2019: 4010-4019.



第一作者简介：姜珊 (1991-), 女, 2012 年获得长春理工大学软件工程学术学位, 2016 年获得长春理工大学计算机科学技术硕士学位。2016 年至今正在攻读长春理工大学大学计算机科学技术博士学位。主要研究方向包括计算机视觉、目标跟踪。研究生期间已发表学术论文 10 篇, 参与了国家自然科学基金青年科学基金、吉林省重大科技公关项目、吉林省科技攻关计划项目等多个科研项目, 参与申请国家发明专利 2 项。Email: sakina_js@163.com



通讯作者简介：底晓强 (1978-), 男, 2002 年获得长春理工大学计算机科学与技术学士学位, 2007 年和 2014 年分别获得长春理工大学通信与信息系统硕士学位和博士学位。2012 年 8 月至 2013 年 8 月, 任挪威科技大学访问学者。现任长春理工大学教授、博士生导师, 信息化中心主任。主要研究方向包括网络信息安全、综合网络和模式识别。获国防科学技术进步二等奖 1 次、兵器科学技术进步一等奖 1 次、省科技进步二等奖和三等奖各 1 次、国家多媒体课件二等奖一次、吉林省教育技术成果一等奖两次。Email: dixiaoqiang@cust.edu.cn