

8-15-2022

## An Efficient Tracker via Multi-feature Adaptive Correlation Filter

Sixian Zhang

*1.State Key Laboratory for Strength and Vibration of Mechanical Structure, Xi'an Jiaotong University, Xi'an 710049, China;2.School of Aerospace, Xi'an Jiaotong University, Xi'an 710049, China,;*  
touchmeteor@stu.xjtu.edu.cn

Yi Yang

*1.State Key Laboratory for Strength and Vibration of Mechanical Structure, Xi'an Jiaotong University, Xi'an 710049, China;2.School of Aerospace, Xi'an Jiaotong University, Xi'an 710049, China,;*  
jiafeiyy@mail.xjtu.edu.cn

Meng Zhang

*1.State Key Laboratory for Strength and Vibration of Mechanical Structure, Xi'an Jiaotong University, Xi'an 710049, China;2.School of Aerospace, Xi'an Jiaotong University, Xi'an 710049, China;*

Pengbo Mi

*1.State Key Laboratory for Strength and Vibration of Mechanical Structure, Xi'an Jiaotong University, Xi'an 710049, China;2.School of Aerospace, Xi'an Jiaotong University, Xi'an 710049, China;*

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the [Artificial Intelligence and Robotics Commons](#), [Computer Engineering Commons](#), [Numerical Analysis and Scientific Computing Commons](#), [Operations Research](#), [Systems Engineering and Industrial Engineering Commons](#), and the [Systems Science Commons](#)

---

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

---

## An Efficient Tracker via Multi-feature Adaptive Correlation Filter

### Abstract

**Abstract:** Aiming at the low tracking effect of the correlation filters tracker based on manual features in challenging scenes of rapid deformation and background clutter, a new correlation filter tracker based on Staple tracker is proposed. *An appearance model based on HOG features and color-naming features is built to enhance the robustness to the challenging scenes of rapid deformation and background clutter. A self-adjust evaluation function is designed to merge the two kinds of feature information and a more discriminative feature is obtained. The novel online update strategies to reduce the training over-fitting and model drift for different features are proposed.* The tracker shows excellent performance in accuracy and real-time capability on OTB2015 benchmark.

### Keywords

object tracking, correlation filter, self-adjust evaluation function, color-naming feature, online update

### Recommended Citation

Sixian Zhang, Yi Yang, Meng Zhang, Pengbo Mi. An Efficient Tracker via Multi-feature Adaptive Correlation Filter[J]. Journal of System Simulation, 2022, 34(8): 1864-1873.

# 高效的多特征自适应相关滤波跟踪器

张思贤<sup>1,2</sup>, 杨艺<sup>1,2\*</sup>, 张猛<sup>1,2</sup>, 米鹏博<sup>1,2</sup>

(1. 西安交通大学 机械结构强度与振动国家重点实验室, 陕西 西安 710049; 2. 西安交通大学 航天航空学院, 陕西 西安 710049)

**摘要:** 针对基于手工特征的相关滤波跟踪器在快速变形、背景杂乱等挑战性跟踪场景效果不佳的问题, 在Staple跟踪器的基础上提出了一种新型的相关滤波跟踪器。结合方向梯度直方图特征与颜色命名特征构建了目标外观模型, 增强其对快速变形以及背景杂乱等场景的鲁棒性; 设计了自适应评分函数对2种特征进行融合, 得到更具有鉴别性的特征; 针对不同的特征分别提出了在线更新策略以减小训练过拟合与模型漂移。实验结果表明: 该跟踪器在跟踪的准确性与实时性上均有着优良表现。

**关键词:** 目标跟踪; 相关滤波; 自适应评分函数; 颜色命名特征; 在线更新

中图分类号: TP391.4

文献标志码: A

文章编号: 1004-731X(2022)08-1864-10

DOI: 10.16182/j.issn1004731x.joss.21-0274

## An Efficient Tracker via Multi-feature Adaptive Correlation Filter

Zhang Sixian<sup>1,2</sup>, Yang Yi<sup>1,2\*</sup>, Zhang Meng<sup>1,2</sup>, Mi Pengbo<sup>1,2</sup>

(1. State Key Laboratory for Strength and Vibration of Mechanical Structure, Xi'an Jiaotong University, Xi'an 710049, China;

2. School of Aerospace, Xi'an Jiaotong University, Xi'an 710049, China)

**Abstract:** Aiming at the low tracking effect of the correlation filters tracker based on manual features in challenging scenes of rapid deformation and background clutter, a new correlation filter tracker based on Staple tracker is proposed. An appearance model based on HOG features and color-naming features is built to enhance the robustness to the challenging scenes of rapid deformation and background clutter. A self-adjust evaluation function is designed to merge the two kinds of feature information and a more discriminative feature is obtained. The novel online update strategies to reduce the training over-fitting and model drift for different features are proposed. The tracker shows excellent performance in accuracy and real-time capability on OTB2015 benchmark.

**Keywords:** object tracking; correlation filter; self-adjust evaluation function; color-naming feature; online update

## 引言

目标跟踪是计算机视觉中的一个重要研究方向。在给定第一帧目标的情况下, 目标跟踪就在接下来一系列的视频帧中估计目标位置。对于运

动目标而言, 其运动的场景非常复杂, 甚至目标本身也在不断变化, 如何在复杂场景中识别并跟踪不断变化的目标就成为一个挑战性的任务。主要挑战因素有快速变形、亮度变化、快速运动、背景干扰、尺度变化等。

收稿日期: 2021-03-31 修回日期: 2021-08-12

基金项目: 国家自然科学基金(61671370); 中国博士后科学基金(2016M592790); 中央高校基本科研业务费专项资金(xjj2016066); 陕西省博士后基金(2016BSHEDZZ46)

第一作者: 张思贤(1999-), 男, 博士生, 研究方向为多元信息融合等。E-mail: touchmeteor@stu.xjtu.edu.cn

通讯作者: 杨艺(1980-), 女, 博士, 博士生导师, 副教授, 研究方向为信息融合, 遥感图像处理, 不确定性建模与推理, 飞行器控制, 组合导航等。E-mail: jiafeiyi@mail.xjtu.edu.cn

基于判别式相关滤波的目标跟踪方法在各类数据集上都表现出了较好的准确性与鲁棒性。从 Blome 等<sup>[1]</sup>提出基于灰度特征的误差最小平方和滤波器(minimum output sum of squared error, MOSSE)以来, 相关滤波类方法一直在不断发展进步。利用粒子滤波采样的相关滤波跟踪器在减小计算量的同时一定程度上提升了跟踪的性能<sup>[2-4]</sup>。利用循环结构的核相关滤波器<sup>[5]</sup>(exploiting the circulant structure of tracking-by-detection with kernels, CSK)通过循环矩阵的性质, 从第一帧图像中提取大量样本进行训练, 并在傅里叶域中加速求解损失函数, 解决了目标跟踪训练样本缺乏的问题。尽管通过循环采样能够得到大量的训练样本, 在傅里叶域中快速运算, 但是该措施会产生边界效应。为减小边界效应的影响, 一系列的算法相继被提出。其中空间正则化判别式滤波器<sup>[6]</sup>(spatially regularized discriminative correlation filters, SRDCF)引入了一个权重项对样本进行空间正则化, 减小了边界附近特征对训练模型的影响, 取得了不错的效果。

此后, 学者们基于不同特征得到了一系列的跟踪器。其中, 高速核相关滤波器(kernelized correlation filters, KCF)基于多通道的方向梯度直方图特征(histograms of oriented gradients, HOG), 在目标跟踪中大放异彩, 展示出了优异的性能<sup>[7-10]</sup>。判别式尺度空间相关滤波跟踪器<sup>[11]</sup>(discriminative scale space tracking, DSST)运用一维尺度滤波器, 结合基于 HOG 特征训练得到的滤波器, 进行判别式的目标跟踪, 解决了尺度自适应问题。为了使跟踪器对于快速变形的目标具有鲁棒性, M. Danelljan<sup>[12]</sup>利用主成分分析对颜色命名特征降维, 并基于该特征训练得到了自适应颜色特征跟踪器(adaptive color attributes for real-time visual tracking, CN)。

由于单一特征的局限性, 学者们开始着手于多种特征的融合, 以获得更具有鉴别性的目标外观模型<sup>[13]</sup>。Staple 等<sup>[14]</sup>(sum of template and pixel-wise LEarners)算法结合了 HOG 特征以及颜色直方图特

征, 基于岭回归框架学习得到一个对亮度变化和快速变形都具有鲁棒性的模型, 对 2 个独立岭回归问题分别求解, 利用每种特征表示的内在结构加速运算, 具有很高的运算速度, 但在部分挑战性场景下跟踪准确性有待提高。此外, 结合深度特征的相关滤波器<sup>[15-18]</sup>得益于深度特征优异的鉴别性, 在跟踪性能上有着显著的提升。利用端到端的孪生网络的跟踪算法获得了较快的跟踪速度<sup>[19-21]</sup>, 基于复杂学习模型的算法<sup>[22-23]</sup>也有着不错的表现。

随着相关滤波类方法的不断推陈出新, 目标跟踪的准确率在逐步上升, 然而, 其独特的高效性却因为越来越庞大的计算量而丧失了。例如, 结合深度特征的高效卷积算子相关滤波跟踪器<sup>[24]</sup>(efficient convolution operators for tracking, ECO)算法虽然成功率较 MOSSE 提升了一倍, 但其运算速度只有后者的 1%。

在快速变形、背景杂乱等挑战性跟踪场景下, 为了兼顾跟踪器的准确性与实时性, 本文提出了一种基于多特征融合的自适应相关滤波跟踪器(sum of template and color-naming learners, STACL)。该跟踪器结合 HOG 特征与颜色命名特征构建鲁棒的目标外观模型, 提出了一种自适应评分函数; 使用 HOG 相关响应图判据与颜色特征得分衰减系数, 计算 2 种特征融合时对应的权重, 得到更具鉴别性的目标特征; 针对 2 种特征模型提出了不同的模型更新策略, 减小训练过拟合与模型漂移。

## 1 Staple 跟踪器

Staple 跟踪器采用检测跟踪的模式, 在第  $t$  帧时, 为了确定图像  $x_t$  中目标的位置从区域  $S_t$  内选择矩形  $p_t$  使以下得分最大化:

$$p_t = \arg \max_{p \in S_t} f(T(x_t, p); \theta_{t-1}) \quad (1)$$

式中: 函数  $T$  是一种图像变换, 使得  $f(T(x_t, p); \theta_{t-1})$  为在图像  $x_t$  中模型参数为  $\theta$  时矩形区域  $p$  的得分。为了获得准确可靠的目标位置, Staple 提出了一种对 HOG 特征得分(模板得分)与

颜色直方图特征得分线性组合的评分函数:

$$f(x_t) = \lambda_{\text{tmpl}} f_{\text{tmpl}}(x_t) + \lambda_{\text{hist}} f_{\text{hist}}(x_t) \quad (2)$$

其中, 原文使用的模板得分权重  $\lambda_{\text{tmpl}} = 0.7$ , 颜色直方图得分权重  $\lambda_{\text{hist}} = 0.3$ 。

模板得分  $f_{\text{tmpl}}(x_t)$  是  $x_t$  的  $K$  通道 HOG 特征图像  $\phi_x: T \rightarrow R^K$  的线性函数, 定义在有限网格  $H \subset Z^2$  上:

$$f_{\text{tmpl}}(x_t; \mathbf{h}) = \sum_{u \in T} \mathbf{h}[u]^T \phi_x[u] \quad (3)$$

其中, 模板  $\mathbf{h}$  对应的是训练得到的  $K$  通道的特征图像。

对于颜色直方图得分而言, 其是从  $x_t$  的  $M$  通道的颜色特征图像  $\psi_x: H \rightarrow R^M$  中计算出来的, 定义在有限网格  $H \subset Z^2$  上:

$$f_{\text{hist}}(x_t; \boldsymbol{\beta}) = \boldsymbol{\beta}^T \left( \frac{1}{|H|} \sum_{u \in H} \psi_x[u] \right) \quad (4)$$

式中:  $\boldsymbol{\beta}$  为训练得到的颜色特征的权重矩阵;  $|H|$  为区域  $H$  内总像素数。

Staple 对 2 种特征分别建立了 2 个独立的岭回归问题来学习目标外观模型, 在保证相关滤波器的速度和效率的同时, 对 HOG 特征与颜色直方图特征包含的信息进行处理。

基于最小二乘正则化的模板得分训练损失为

$$l_{\text{tmpl}}(x_t, p_t, \mathbf{h}) = \left\| \sum_{k=1}^K \mathbf{h}^k \star \phi^k - \mathbf{y} \right\|^2 + \lambda \sum_{k=1}^K \|\mathbf{h}^k\|^2 \quad (5)$$

式中:  $\star$  为循环相关操作;  $\mathbf{h}^k$  为模板  $\mathbf{h}$  的第  $k$  通道分量,  $k = 1, 2, \dots, K$ ;  $\phi$  为矩形区域  $p_t$  特征图像  $\phi_{T(x, p)}$  的缩写;  $\mathbf{y}$  为期望的响应(通常为高斯函数);  $\lambda$  为防止过拟合而设置的正则化参数。求解式(5)即可得到模板  $\mathbf{h}$  的封闭解。

在颜色直方图得分求解的过程中, 对样本内目标区域  $O \subset Z^2$  与背景区域  $B \subset Z^2$  分别应用线性回归, 并引入 one-hot 假设, 将目标颜色特征分解为每个特征维度的独立项, 得到其损失函数

$$l_{\text{hist}}(x_t, p_t, \boldsymbol{\beta}) = \sum_{j=1}^M \left[ \frac{N^j(O)}{|O|} \cdot (\boldsymbol{\beta}^j - 1)^2 + \frac{N^j(B)}{|B|} \cdot (\boldsymbol{\beta}^j)^2 \right] \quad (6)$$

式中:  $|\cdot|$  为该区域内像素点总个数;  $N^j(A) = |\{u \in A: k[u] = j\}|$  为特征图像区域  $A$  内特征值  $k[u] = j$  的像素数;  $\boldsymbol{\beta}^j$  为  $\boldsymbol{\beta}$  在第  $j$  维特征上的分量,  $j = 1, 2, \dots, M$ 。

## 2 STACL 跟踪器

尽管 Staple 跟踪器在保证实时性的同时相较于传统单一特征的跟踪器在跟踪精度上有着不错的提升, 然而, 仍然存在训练过拟合以及在部分挑战性场景下表现不佳等问题<sup>[11]</sup>。本文从目标外观模型优化、自适应评分函数设计和目标外观模型更新策略 3 个方面着手, 设计高速且高效的跟踪器。图 1 是算法训练与检测过程框图。

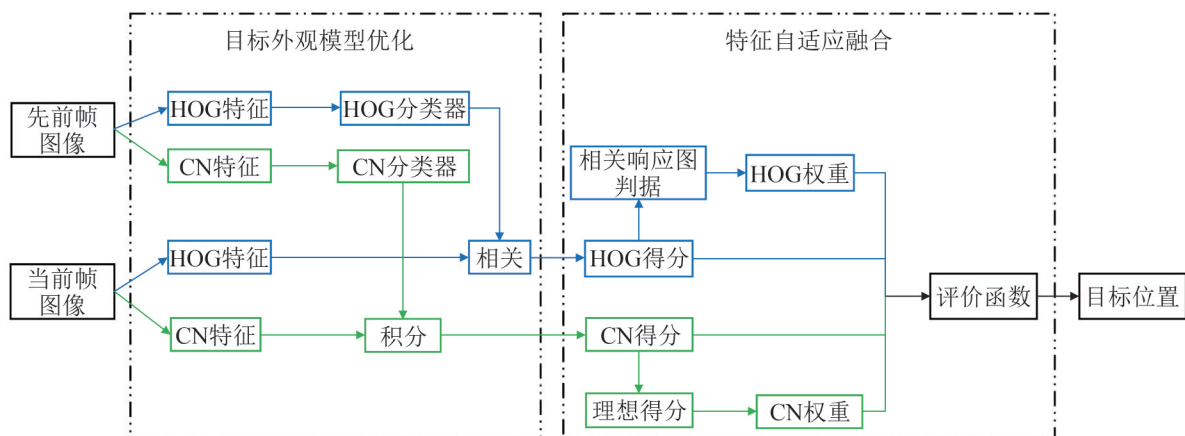


图1 STACL 算法框图

Fig. 1 Block diagram of STACL

## 2.1 目标外观模型优化

目标特征的选择对跟踪器的性能至关重要, 选取合适的特征信息往往能达到事半功倍的效果。HOG特征通过计算和统计图像局部区域的梯度方向直方图来构成特征。由于它只在图像的局部方格单元上操作, 所以对图像几何和光学的形变都能保持很好的不变性, 在目标跟踪领域有着广泛的应用。而颜色命名特征(color naming, CN)<sup>[10-17]</sup>将所有色彩分为11个基础类别, 是一种全局特征, 对于目标的快速运动以及变形都具有不错的鲁棒性。尽管利用颜色直方图积分快速求解岭回归是Staple算法的一大亮点, 然而, 式(6)仅使用图像的RGB值来生成相关滤波器, 训练得到的滤波器并不具备很好的鉴别性, 在光照变化和背景杂乱的场景下表现不佳。为了更有效地提取目标特征信息, 利用HOG特征与图像RGB值生成的CN特征这一对互补的特征, 保留它们对快速变形以及光照变化的不变性, 构建鲁棒的目标外观模型, 并计算相应的特征得分, 以此得到目标的准确位置。

### 2.1.1 HOG特征得分

HOG特征得分就是基于HOG特征训练得到的滤波器与样本相关后得到的响应图, 计算HOG特征得分的关键是对滤波器 $h$ 进行求解。

定义 $\Phi$ 为特征图像 $\phi_{T(x,p)}$ 通过循环移位采样形成的循环矩阵。式(5)是一个线性回归问题, 利用帕萨瓦尔公式将其变换到傅里叶域中可以对其快速求解。式(5)的封闭解可以在高速核相关滤波器<sup>[4]</sup>中得到:

$$h = (\Phi^H \Phi + \lambda I)^{-1} \Phi^H y \quad (7)$$

式中:  $\Phi^H$ 为 $\Phi$ 的厄米共轭转置, 这意味着 $\Phi^H = (\Phi^*)^T$ , 且 $\Phi^*$ 是 $\Phi$ 的复共轭矩阵;  $y$ 为期望的响应(通常为高斯函数)。定义 $\hat{x}$ 为矩阵 $x$ 的快速傅里叶变换 $F_x$ ,  $X$ 是由矩阵 $x$ 循环移位采样得到的循环矩阵, 常数矩阵 $F$ 为离散傅里叶变换矩阵:

$$X = F \text{diag}(\hat{x}) F^H \quad (8)$$

将式(8)代入式(7)可得模板 $h$ 的傅里叶变换:

$$\hat{h}^l = \frac{\hat{\phi}^l \odot \hat{y}^*}{\sum_{i=1}^K \hat{\phi}^{i*} \odot \hat{\phi}^i + \lambda}, l = 1, 2, \dots, K \quad (9)$$

式中: 符号 $\odot$ 为元素乘;  $\hat{\phi}^i$ 为 $K$ 通道特征图像 $\phi_{T(x,p)}$ 的第 $i$ 维。

对于包含 $m$ 个像素与 $K$ 个特征通道的样本而言, 计算其特征得分以及训练滤波器的时间复杂度为 $O(Km \log m)$ 。

### 2.1.2 CN特征得分

在人类所涉及的各种视觉任务中, 颜色命名是最普遍的一种。目前被普遍接受的框架是把色彩命名看作一个模糊过程, 任意的颜色刺激对每个基础颜色类别都有一个0-1之间的隶属值。使用文献[10]中的映射函数对像素值进行处理, 任意RGB值对每个颜色类别都有一个0-1之间的隶属值, 而最大隶属值对应的颜色即为该RGB值的CN特征。对图像进行进一步处理而得到其CN特征, 目的是深度挖掘图像中包含的特征信息以优化目标外观模型, 使训练得到的滤波器更具鉴别性, 提高目标跟踪的性能。

在第 $t$ 帧时, 图像 $x_t$ 内矩形框 $p_t$ 的 $N$ 通道( $N=11$ )CN特征图像表示为

$$\Omega(x_t, p_t) = \text{TS}(x_t, p_t) \quad (10)$$

式中: TS为将RGB值变换为CN特征的映射函数。

对于CN特征图像而言, 其特征得分计算式为

$$f_{\text{hist}}(x; y) = y^T \left( \frac{1}{|H|} \sum_{u \in H} \Omega[u] \right) \quad (11)$$

式中:  $y$ 为CN特征权重矩阵。

由式(6)可得, 最小化损失函数 $l_{\text{hist}}(x_t, p_t, y)$ 的封闭解为

$$y_i^j = \frac{\rho^j(O)}{\rho^j(O) + \rho^j(B) + \lambda}, j = 1, 2, \dots, N \quad (12)$$

式中:  $\rho^j(A) = N^j(A)/|A|$ , 表示在区域 $A$ 内CN特征 $\Omega[u]=j$ 的像素点个数与区域像素点总数之比;  $y_i^j$ 为 $(x_t, p_t)$ 的CN特征权重矩阵 $y$ 的第 $j$ 通道分量。

每种特征得分最大值处即为基于该特征得到的当前目标中心位置。尽管大部分情况下HOG特征与CN特征均具有良好的性能，但是由于所选取的手工特征的模型局限性，不同情境下的HOG特征与CN特征表现并不相同，它们独立得到的目标中心位置自然也是不同的。因此，需要对这一对互补特征进行特征融合，保留它们各自的不变性，得到更具有鉴别性的特征。

## 2.2 自适应评分函数设计

实验表明，HOG特征在目标表现发生明显变化时跟踪性能表现不佳，特别是目标在平面内旋转或是大幅度变形时，训练所得相关滤波器跟踪精度大大降低<sup>[25]</sup>；对于CN特征而言，当目标经过与其颜色相似的大面积的背景时，跟踪器往往锁定在背景物体上，导致后续跟踪的失败。倘若使用固定的得分权重计算目标位置，当某一种特征对于准确的目标跟踪几乎毫无贡献时，其特征得分仍会对最终结果产生很大影响。

为了提高跟踪器在不同跟踪场景下的鲁棒性，本文设计了一种自适应评分函数用于特征的融合，使得不同特征的得分权重随着目标跟踪过程中特征滤波器表现的优劣而变化。在式(2)中引入自适应权重，定义 $g_{\text{tmpl}}(x)$ 与 $g_{\text{hist}}(x)$ 分别为HOG特征得分与CN特征得分的权重，新的得分函数为

$$f(x) = g_{\text{tmpl}}(x)f_{\text{tmpl}}(x) + g_{\text{hist}}(x)f_{\text{hist}}(x) \quad (13)$$

为了评估HOG特征滤波器的跟踪效果，需要对相关滤波后的响应图进行分析。然而，单纯通过相关响应图峰值的大小并不能很好地评估目标跟踪效果，因为在实际跟踪过程中目标外观往往是无规律变化的，即使是较小的形变也可能导致相关响应图峰值显著减小或是出现多个峰，仅仅使用一个最大值点不利于跟踪效果的评估，故需要对原始相关响应图进行处理以得到有效的评价指标。为此引入文献[25]中提出的平均峰值相关能量(average peak-to-correlation energy, APCE)作为HOG相关响应图判据：

$$APCE = \frac{|F_{\max} - F_{\min}|^2}{\text{mean}\left(\sum_{w,h} (F_{w,h} - F_{\min})^2\right)} \quad (14)$$

式中： $F_{\max}$ 为HOG特征相关响应图 $F_{w,h}$ 中最大值； $F_{\min}$ 是相关响应图中最小值。图像的APCE值反映了相关响应图的全局振荡程度，当APCE值急剧减小时，说明相关响应图较为平缓，此时跟踪器的跟踪效果不佳。

设置目标跟踪的阈值 $threshold$ ，若HOG特征响应图的值小于该阈值，则基于该特征得到的目标位置置信度不高，应选用较小的特征权重；反之则说明基于该特征得到的目标位置置信度较高，应选用较大的特征权重：

$$g_{\text{tmpl}}^0(APCE) = \begin{cases} \alpha_1, & APCE < threshold \\ \alpha_2, & \text{else} \end{cases} \quad (15)$$

式中： $\alpha_1$ 、 $\alpha_2$ 为对应的特征权重；实验中设置 $threshold = 5$ ； $g_{\text{tmpl}}^0$ 为初始HOG特征权重，有 $g_{\text{tmpl}}^0 + g_{\text{hist}}^0 = 1$ ， $g_{\text{hist}}^0$ 为初始CN特征得分权重。

样本矩形框 $(x_t, p_t)$  CN特征得分可根据式(12)得到的CN特征权重矩阵 $\gamma$ 计算：

$$f_{\text{hist}}(x_t; \gamma_t) = \frac{\sum_{j=1}^N \gamma_t^j N^j(p_t)}{\sum_{i=1}^N N^i(p_t)} \quad (16)$$

从式(16)可以看出，当具有 $\gamma_t^j$ 最大值对应颜色类别的像素点越多时，CN特征得分越大，这意味着即便某一样本与训练样本完全一致，该样本的CN特征得分也不一定是最大的，这将会导致不准确的目标跟踪。定义上一帧目标样本在当前帧的CN特征得分为理想得分：

$$score_{\text{hist}}(x_t; \gamma_t) = \frac{\sum_{j=1}^N \gamma_t^j N^j(O_{t-1})}{\sum_{i=1}^N N^i(O_{t-1})} \quad (17)$$

不妨设 $r = f_{\text{hist}}(x_t; \gamma_t) / score_{\text{hist}}(x_t; \gamma_t)$ ，定义CN特征得分衰减系数：

$$\mu = e^{-(r-1)^2} \quad (18)$$

当  $0 < r < 1$  时, 随着  $r$  的增大, 衰减系数  $\mu$  增大, 说明当前帧样本与上一帧目标样本 CN 特征相似度增大; 当  $r = 1$  时, 说明当前帧样本与上一帧目标样本 CN 特征完全匹配, 此时衰减系数  $\mu$  取最大值 1; 当  $r > 1$  时, 说明当前帧跟踪结果受到背景颜色影响较大, 应适当降低 CN 特征得分权重, 随着  $r$  的增大, 衰减系数  $\mu$  减小。最终 2 类特征得分权重由式(19)确定:

$$\begin{aligned} g_{\text{templ}}(x) &= \frac{g_{\text{templ}}^0(x)}{g_{\text{templ}}^0(x) + \mu g_{\text{hist}}^0(x)} \\ g_{\text{hist}}(x) &= \frac{\mu g_{\text{templ}}^0(x)}{g_{\text{templ}}^0(x) + \mu g_{\text{hist}}^0(x)} \end{aligned} \quad (19)$$

### 2.3 目标外观模型更新

视频中目标外观会随着时间发生无规律变化, 仅仅从第一帧获取的特征得到的目标外观模型, 在后续的跟踪过程中表现往往差强人意。现如今使用最多的更新策略是基于跟踪器在每一帧中预测得到的目标位置提取特征, 并采用固定的学习参数进行模型的更新。其不足之处在于一旦跟踪器的预测位置产生误差, 从该位置学习得到的模型也将不准确, 误差的逐帧积累会导致模型漂移。据此, 利用 2.2 节中相关响应图评估判据来判定是否需要更新当前模型, 减小训练过拟合。由于利用积分方法计算 CN 特征得分, 所以背景区域特征与目标区域特征可以被分开处理, 以此减少背景干扰对目标跟踪的影响, 提高跟踪的准确性。

在目标跟踪的过程中, 为了更新模型, 需要考虑到所有跟踪完成的视频帧, 对于一系列的视频序列  $x_1, x_2, \dots, x_T$  而言, 其训练的损失函数应当是每一帧图像损失的线性加权:

$$L(\theta) = \sum_{t=1}^T \omega_t l(x_t, p_t, \theta) \quad (20)$$

式中:  $\omega_t > 0$  为第  $t$  帧的图像权重;  $l(x_t, p_t, \theta)$  为第  $t$  帧图像的损失函数。

对于 HOG 特征模型而言, 该最小化损失函数的解为

$$\hat{h}^l = \frac{\sum_{t=1}^T \omega_t \hat{y}_t^* \hat{\phi}_t^l}{\sum_{t=1}^T \sum_{i=1}^K \omega_t \hat{\phi}_t^{i*} \hat{\phi}_t^i + \lambda}, l = 1, 2, \dots, K \quad (21)$$

设式(21)中  $\hat{h}^l$  的分子为  $\hat{h}_N$ , 分母为  $\hat{h}_D$ , 为避免产生次优解<sup>[17]</sup>, 对  $\hat{h}_N$  与  $\hat{h}_D$  分别进行更新, 同时, 这一措施减小了模型更新的计算量并保证了目标跟踪的速度不会因模型更新而显著变化。式(21)中的权重  $\omega_t$  用学习速率参数  $\eta$  设置, 最终的模型更新策略为

$$\begin{aligned} \hat{h}_{Nt}^l &= (1 - \eta) \hat{h}_{Nt-1}^l + \eta \hat{y}_t^* \hat{\phi}_t^l, l = 1, 2, \dots, K \\ \hat{h}_{Dt} &= (1 - \eta) \hat{h}_{Dt-1} + \eta \sum_{i=1}^K \omega_t \hat{\phi}_t^{i*} \hat{\phi}_t^i \end{aligned} \quad (22)$$

得到第  $t$  帧的 HOG 特征滤波器之后, 提取候选样本  $\phi(x_{t+1}, p_{t+1})$  并计算其 HOG 特征得分:

$$\hat{z}_{t+1} = \frac{\sum_{i=1}^K \hat{h}_{Nt}^{i*} \hat{\phi}_{t+1}^i}{\hat{h}_{Dt} + \lambda} \quad (23)$$

式中:  $\hat{h}_{Nt}^l$  与  $\hat{h}_{Dt}$  为第  $t$  帧更新所得滤波器的分子与分母。对  $\hat{z}_{t+1}$  作快速傅里叶逆变换得到相关响应图。只有当 HOG 相关响应图的 APCE 值大于设定阈值时, 才进行模型的更新。在发生遮挡或是目标消失时, 目标跟踪效果不佳, 但是用于目标跟踪的滤波器参数并未改变, 那么当目标重新出现后, 跟踪器仍然能够准确地定位目标, 防止模型漂移。

在计算 CN 特征得分时, 候选样本框内背景区域以及目标区域独立地应用了线性回归。因此, 式(12)中的  $\rho^j(O)$  与  $\rho^j(B)$  可以分别进行在线更新。对于目标区域而言, 目标过去的外观信息十分重要, 在更新时需要利用所有先前帧的特征信息; 然而, 背景区域也在不断变化, 由于 CN 特征得分的特殊性, 先前帧中的背景区域信息对目标跟踪的结果影响不大, 其特征信息对目标跟踪而言具有时效性, 也就是说, 只有当前帧的背景信息才是有效的。类似 HOG 模型更新, 引入学习参数  $\eta$ , 最终 CN 特征模型更新策略为

$$\begin{aligned} \rho_t(O) &= (1 - \eta) \rho_{t-1}(O) + \eta \rho_t'(O) \\ \rho_t(B) &= \rho_t'(B) \end{aligned} \quad (24)$$

式中:  $\rho_t'(\cdot)$  为由  $\rho^j(\cdot), j = 1, 2, \dots, N$  组成的向量。



### 3 实验结果

本文选取了OTB2015<sup>[26]</sup>作为测试基准，将本文算法与其他较为先进的目标跟踪算法相比较。为了保证实验结果的准确性，在Matlab上使用以上待测试算法原作者提供的源码进行测试。本文算法所使用的主要参数的值如表1所示。

主要参数	量值	主要参数	量值
学习率 $\eta$	0.01	融合参数 $\alpha_2$	0.6
颜色特征	Color Naming	检测区域	150×150
融合参数 $\alpha_1$	0.3	Hog单元尺寸	4×4

本文从经典的OTB2015<sup>[26]</sup>数据集中选出了有关快速变形与背景杂乱的30个彩色序列。在选取的视频序列中，往往还包含多种目标跟踪的挑战因素，即背景干扰(background clutter)、平面旋转(in plane rotation)、变形(deformation)、尺度变换(scale variation)、光照变化(illumination variation)等。对于OTB2015数据集而言，它主要通过准确率与成功率曲线，对跟踪器的准确率与鲁棒性行评测。统计每次视频序列的目标位置(bounding box)的中心点与人工标注的目标的中心点的距离，计算这两者的距离小于给定阈值的视频帧，占视频序列总数的百分比来绘制准确率图。定义重合率得分  $OS = |a \cap b| / |a \cup b|$ 。其中  $a$  为视频序列

目标位置的边界框； $b$  为人工标注的边界框； $|\cdot|$  为区域像素总数目。当某一帧的  $OS$  大于设定的阈值时，则该视频帧被视为追踪成功，追踪成功的帧数占视频序列总数的百分比即为成功率。本文使用了OTB 1.0工具箱生成结果，并结合近年来部分性能良好的追踪器进行对比评估，分别为ECO\_HC<sup>[24]</sup>、Staple<sup>[14]</sup>、SRDCF<sup>[6]</sup>、DSST<sup>[11]</sup>、CN<sup>[12]</sup>、KCF<sup>[9]</sup>，以及工具箱自带的基准跟踪器CSK<sup>[5]</sup>。其中，ECO\_HC是在ECO框架上基于HOG特征和颜色特征训练得到的跟踪器。图2列出了本文跟踪器与数种跟踪器在skiing、dragonbaby 2个场景下的跟踪定性结果图。图3和图4分别为所有测试算法的准确率图和成功率图。显而易见，对于本文算法而言，无论是目标跟踪的平均准确率或是平均成功率都位居首位，相对于基准算法CSK分别提升了79.5%与81.7%，超过了基于HOG特征的相关滤波跟踪器DSST、KCF，基于颜色特征的CN，基于空间正则化的SRDCF，以及同样使用了多通道特征融合的Staple与ECO\_HC，而且算法的运行速度亦名列前茅，如表2所示。特别的，本文算法在图3与图4所有挑战因素下目标跟踪的准确率与成功率几乎均为第一。这是因为本文利用了一对互补的特征得到了更具有鉴别性目标外观模型，能够进一步提升目标跟踪的精度。除此以外，模型的更新策略也是目标跟踪精度提升的原因之一。

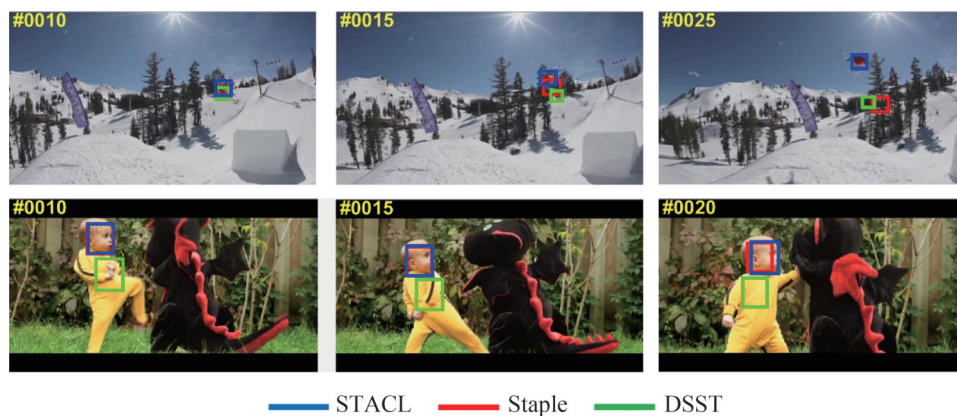


图2 3种跟踪器在2个示例视频中定性结果

Fig. 2 Qualitative result of three trackers on two example sequences.

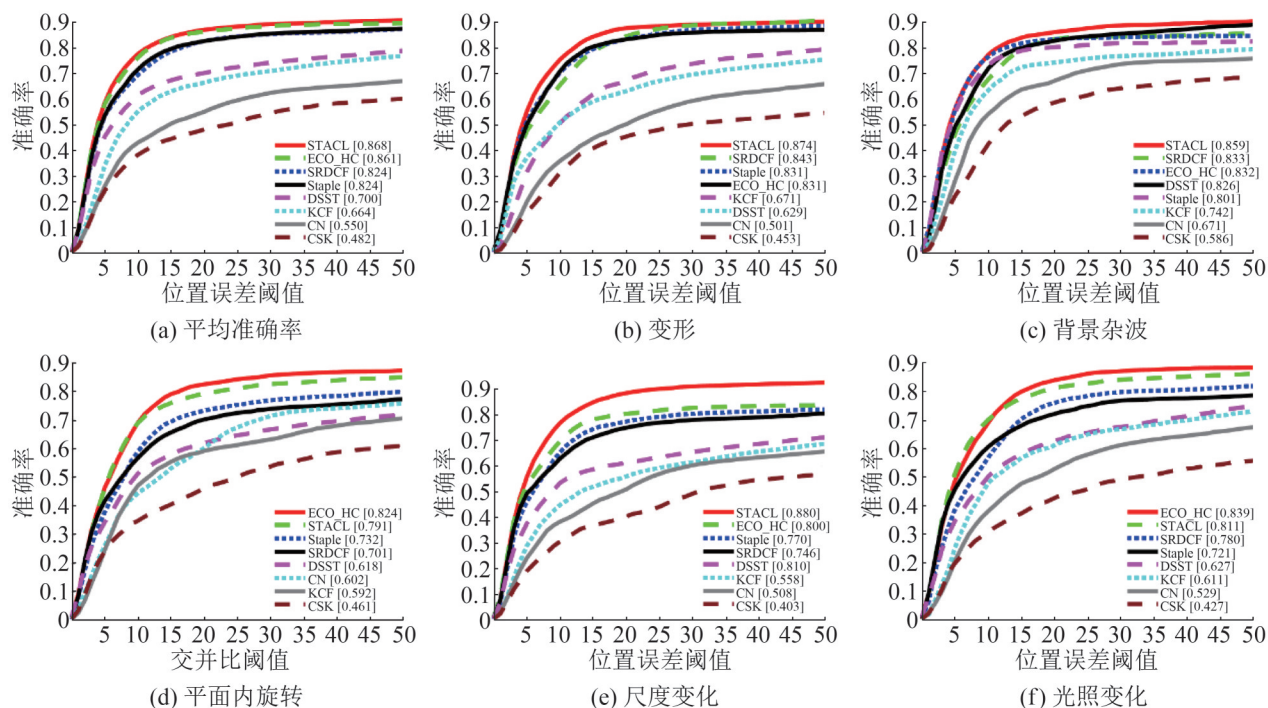


图3 OTB2015数据集上跟踪器OPE(one pass evaluation)准确率  
Fig. 3 Precision plots for OPE on OTB2015 benchmark

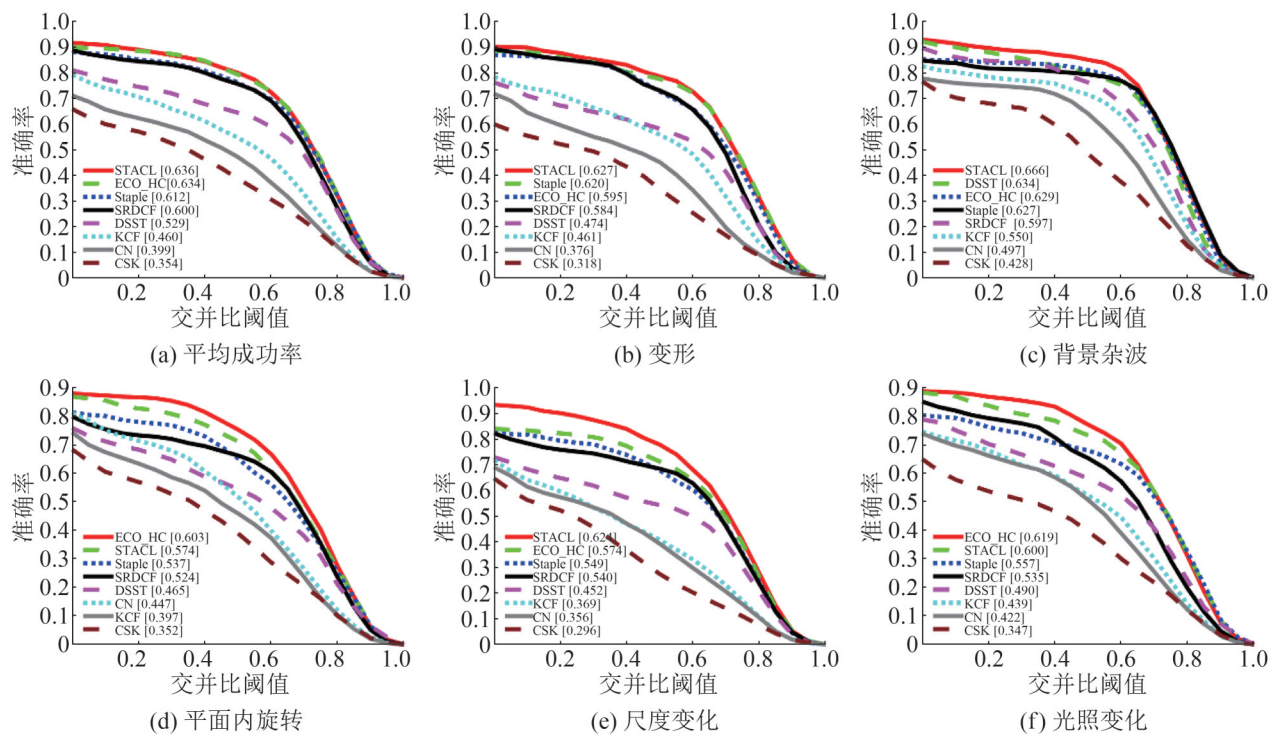


图4 OTB2015数据集上跟踪器OPE成功率  
Fig. 4 Success plots for OPE on OTB2015 benchmark

表2 跟踪器运行帧数  
Table 2 Trackers running frames

跟踪器	帧数	跟踪器	帧数
CSK	275.0	ECO_HC	37.5
KCF	120.5	DSST	33.7
Staple	55.1	CN	13.5
STACL	50.3	SRDCF	9.8

由表3可得, 本文算法相对于 Staple 算法在除遮挡外所有的挑战因素下都有着更强的鲁棒性。在快速变形的情况下, 目标跟踪成功率相较 Staple 算法成功率提升了 3.73%; 在光照变化的情况下, 相较提升了 5.87%; 在背景干扰的情况下, 提高了 9.47%; 在尺度变化的情况下, 则提升了 11.1%。

表3 STACL 与 Staple 在不同因素下成功率对比  
Table 3 Comparisons of success rate of STACL and Staple under different factors

challenges	STACL	Staple
变形	0.640	0.617
背景杂波	0.682	0.623
光照变化	0.631	0.596
面外旋转	0.594	0.561
面内旋转	0.589	0.549
尺度变化	0.613	0.552
遮挡	<b>0.593</b>	<b>0.621</b>
抖动	0.631	0.625
快速移动	0.657	0.651
低分辨率	0.570	0.569

## 4 结论

基于相关滤波的目标跟踪方法是计算机视觉领域的热点问题。然而, 复杂的跟踪场景给准确的目标跟踪带来了巨大的挑战。因此, 为在快速变形及背景杂乱等挑战性场景下目标跟踪的性能, 本文提出了一种基于多特征融合相关滤波的实时滤波器。该滤波器结合了 HOG 特征与 CN 特征, 并设计了自适应评分函数以训练鲁棒的外观模型, 此外, 探讨了一种新颖的外观模型更新策略对跟踪结果的影响。在 OTB2015 数据集上的实验结果显示, 得益于多特征滤波器的自适应融合, 本文提出的跟踪器效果优于其余仅使用单一

特征的跟踪器, 而且依据相关响应图的更新策略相较于传统的逐帧更新策略更能减小模型漂移发生的概率。然而, 由于缺少重检测模块, 一旦在某一帧跟踪器跟踪失败, 在后续帧中也很难重新跟踪上目标。所以, 进一步的研究应着重于跟踪的重检测模块的设计, 一旦跟踪失败, 能够迅速重定位到目标位置。

## 参考文献:

- [1] D S Bolme, J Ross Beveridge, Bruce A Draper, et al. Visual Object Tracking Using Adaptive Correlation Filters[C]//2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Francisco, CA: IEEE, 2010: 2544-2550.
- [2] Zhang T, Liu S, Xu C, et al. Correlation Particle Filter for Visual Tracking[J]. IEEE Transactions on Image Processing(S1057-7149), 2018, 27(6): 2676-2687.
- [3] Zhang T, Xu C, Yang M. Learning Multi-Task Correlation Particle Filters for Visual Tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence(S0162-8828), 2019, 41(2): 365-378.
- [4] Li S, Zhao S, Cheng B, et al. Robust Visual Tracking Via Hierarchical Particle Filter and Ensemble Deep Features [J]. IEEE Transactions on Circuits and Systems for Video Technology(S1051-8215), 2020, 30(1): 179-191.
- [5] Henriques J, Caseiro R, Martins P, et al. Exploiting the Circulant Structure of Tracking-by-Detection with Kernels[C]//12th European Conference on Computer Vision. Berlin, Heidelberg: Springer, 2012: 702-715.
- [6] Danelljan M, Häger G, Khan F S, et al. Learning Spatially Regularized Correlation Filters for Visual Tracking[C]//2015 IEEE International Conference on Computer Vision. Santiago: IEEE, 2015: 4310-4318.
- [7] Song Z, Sun J, Duan B. Collaborative Correlation Filter Tracking with Online Re-detection[C]//2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC). Chengdu: IEEE, 2019: 1303-1313.
- [8] Zhang S, Lu W, Xing W, et al. Learning Scale-Adaptive Tight Correlation Filter for Object Tracking[J]. IEEE Transactions on Cybernetics(S2168-2267), 2020, 50(1): 270-283.
- [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 (S0162-8828), 2015, 37(3): 583-596.

- [10] Lukežič A, Zajc L, Čehovin, Kristan M. Deformable Parts Correlation Filters for Robust Visual Tracking[J]. IEEE Transactions on Cybernetics(S2168-2267), 2018, 48(6): 1849-1861.
- [11] Danelljan Martin, Häger Gustav, Khan Shahbaz Fahad, et al. Discriminative Scale Space Tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence (S0162-8828), 2017, 39(8): 1561-1575.
- [12] Danelljan M, Khan F S, Felsberg M, et al. Adaptive Color Attributes for Real-Time Visual Tracking[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH: IEEE, 2014: 1090-1097.
- [13] Gundogdu E, Alatan A A. Good Features to Correlate for Visual Tracking[J]. IEEE Transactions on Image Processing(S1057-7149), 2018, 27(5): 2526-2540.
- [14] Bertinetto L, Valmadre J, Golodetz S, et al. Staple: Complementary Learners for Real-Time Tracking[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV: IEEE, 2016: 1401-1409.
- [15] Valmadre J, Bertinetto L, Henriques J, et al. End-to-End Representation Learning for Correlation Filter Based Tracking[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI: IEEE, 2017: 5000-5008.
- [16] Feng J, Pu S, Zhao K, et al. Enhanced Initialization with Multi-Stage Learning for Robust Visual Tracking[C]//2019 IEEE Visual Communications and Image Processing. Sydney, NSW: IEEE, 2019: 1-4.
- [17] Bhat G, Danelljan M, Gool L Van, et al. Learning Discriminative Model Prediction for Tracking[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea(South): IEEE, 2019: 6181-6190.
- [18] Danelljan M, Bhat G, Khan F S, et al. ATOM: Accurate Tracking by Overlap Maximization[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, CA: IEEE, 2019: 4655-4664.
- [19] Guo D, Wang J, Cui Y, et al. SiamCAR: Siamese Fully Convolutional Classification and Regression for Visual Tracking[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, WA: IEEE, 2020: 6268-6276.
- [20] Voigtlaender P, Luiten J, Torr P H S, et al. SiamR-CNN: Visual Tracking by Re-Detection[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, WA: IEEE, 2020: 6577-6587.
- [21] Li B, Wu W, Wang Q, et al. SiamRPN++: Evolution of Siamese Visual Tracking with Very Deep Networks[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, CA: IEEE, 2019: 4277-4286.
- [22] Hu J, Shen L, Albanie S, et al. Squeeze-and-Excitation Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence(S0162-8828), 2020, 42(8): 2011-2023.
- [23] Danelljan M, Gool L Van, Timofte R. Probabilistic Regression for Visual Tracking[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, WA: IEEE, 2020: 7181-7190.
- [24] Danelljan Martin, Bhat Goutam, Khan Fahad Shahbaz, et al. ECO: Efficient Convolution Operators for Tracking[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI: IEEE, 2017: 6931-6939.
- [25] Wang M, Liu Y, Huang Z. Large Margin Object Tracking with Circulant Feature Maps [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI: IEEE, 2017: 4021-4029.
- [26] Wu Y, Lim J, Yang M H. Object Tracking Benchmark[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence (S0162-8828), 2015, 37(9): 1834-1848.