

# 基于纵横比自适应的相关滤波跟踪算法

钟钰彬, 杨鹏, 窦磊\*

(南京理工大学瞬态物理国家重点实验室, 江苏南京 210094)

**摘要:** 由于跟踪过程目标不规则形变的影响, 采用固定纵横比的尺度模型无法精确地估计目标的尺度. 为解决该问题, 本文提出基于纵横比自适应的相关滤波跟踪算法. 基于fDSST(fast Discriminative Scale Space Tracking)算法, 训练学习纵横比模型, 更新目标的纵横比, 获取更精确的目标尺度. 在此基础上, 本文设计了平滑修正方案以及学习率自适应机制, 可以有效地缓解因目标出现遮挡导致的模型漂移问题. 在OTB100、VOT2016和VOT2018数据集上与其他跟踪算法进行对比实验, 结果表明本文算法改善了基准算法的性能, 特别是在OTB100上的总体准确率和成功率比fDSST提高了9.6%和6.2%.

**关键词:** 目标跟踪; 相关滤波; 纵横比; 尺度估计; 平滑修正; 学习率自适应

**基金项目:** 国家自然科学基金(No.60904085); 瞬态物理国家重点实验室基金项目

**中图分类号:** TP319.41

**文献标识码:** A

**文章编号:** 0372-2112(2024)06-2112-11

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.12263/DZXB.20230162

## Correlation Filtering Tracking Algorithm Based on Adaptive Aspect-Ratio

ZHONG Yu-bin, YANG Peng, DOU Lei\*

(National Key Laboratory of Transient Physics, Nanjing University of Science and Technology, Nanjing, Jiangsu 210094, China)

**Abstract:** Due to the irregular deformation of target in the tracking process, it is unable to accurately estimate the target scale, while using the scale model with fixed aspect ratio. In this paper, we propose an aspect-ratio-based correlation filtering tracking algorithm to address this problem. Based on the fDSST (fast Discriminative Scale Space Tracking) algorithm, first train and learn an aspect-ratio model to update the aspect ratio of the target, which could help to obtain a more accurate target scale. On this basis, this paper designs a smoothing correction scheme and an adaptive learning rate mechanism to alleviate the model drift and achieve more accurate tracking. The results of comparative experiments on OTB100, VOT2016 and VOT2018 datasets show that the proposed algorithm improves the performance of the baseline algorithm. Especially, the overall precision and success rate of the proposed algorithm on OTB100 are 9.6% and 6.2% higher than those of fDSST.

**Key words:** object tracking; correlation filter; aspect ratio; scale estimation; smoothing correction; adaptive learning rate

**Foundation Item(s):** National Science Foundation of China (No.60904085); Foundation of National Key Laboratory of Transient Physics

## 1 引言

作为计算机视觉的重要研究方向, 目标跟踪受到了国内外众多研究人员的关注. 近几十年来, 研究者开发了大量目标跟踪算法, 但由于只利用初始帧信息导致跟踪信息的不足以及现实场景中存在的各种影响因素, 例如光照变化、运动模糊、姿态变化、遮挡、形变等, 目标跟踪仍是一个具有挑战性的热点问题.

在目标跟踪领域, 根据是否利用背景信息, 目标跟踪算法大致分为生成式模型和判别式模型. 早期的工作主要集中在生成式模型算法的研究中<sup>[1]</sup>, 如粒子滤波<sup>[2]</sup>、卡尔曼滤波<sup>[3]</sup>、Camshift算法<sup>[3]</sup>、光流法<sup>[4]</sup>和Mean-shift算法<sup>[5]</sup>等. 生成式模型根据对外观模型的相似性进行目标定位, 选择与目标相似性最高的候选区域作为目标的跟踪结果. 但生成式模型未利用图像的背景

信息,且单一模型描述目标的状态信息存在一定的局限性,面临光照变化、旋转、形变等挑战表现不佳.判别式模型充分利用背景信息、目标信息作为正样本,背景信息作为负样本,对比前景信息和背景信息的特征差异,将目标区域提取出来.因其鲁棒性强和速度快的特点,引起了判别式模型目标跟踪的研究热潮<sup>[6]</sup>.2010年,David<sup>[7]</sup>提出最小输出误差平方和(Minimum Output Sum of Square Error, MOSSE)跟踪算法,首次将信号领域的相关滤波应用于图像跟踪,采用快速傅里叶变换优化计算过程,跟踪速度提高了20倍.2012年,Henriques<sup>[8]</sup>提出循环结构的检测跟踪(Circulant Structure of tracking-by-detection with Kernels, CSK)算法,利用循环矩阵密集采样,获取大量样本,有效解决了过拟合的问题.虽然循环矩阵提高了跟踪精度,但也增加了运算成本,为了降低计算复杂度,引入核结构,提高了跟踪器的运行效率.2015年,Henriques<sup>[9]</sup>又提出核相关滤波(Kernelized Correlation Filters, KCF)跟踪算法,将目标的状态求解问题抽象转化为前景和背景的分类问题<sup>[10]</sup>.利用方向梯度直方图(Histogram of Oriented Gradient, HOG)将单通道变为多通道,并采用高斯核函数将线性岭回归问题映射到非线性空间,在非线性空间中依据傅里叶空间形成对角化原理简化计算,解决该分类问题<sup>[9]</sup>.

近年,学者们为抑制背景等干扰进一步改善相关滤波算法<sup>[11-20]</sup>.如李豪等人<sup>[11]</sup>利用空间加权融合特征,以及结合目标分割技术,提升了算法性能.王叶等人<sup>[12]</sup>通过增加畸变抑制项和自适应空间正则化项改善模型,提高跟踪的鲁棒性.谢青松等人<sup>[13]</sup>使用前景优化框架,利用前景交并比极大策略选择最优跟踪框角度.田昊东等人<sup>[14]</sup>设计空间弹性网络约束缓解干扰并添加时间正则化项抑制畸变,增强目标在复杂情况下跟踪效果.

以上相关滤波算法定位快速精准,但忽视了目标的尺度信息.针对目标的尺度变化问题,主流的改进方案是基于尺度池的方法.Yang<sup>[21]</sup>提出多特征自适应尺度跟踪算法,使用7个固定尺度池对目标进行尺度缩放,相关得响应图,选择响应数值最大的尺度作为目标的尺度.Danelljan<sup>[22]</sup>提出尺度空间判别跟踪(Discriminative Scale Space Tracking, DSST)算法,独立训练位置滤波器和尺度滤波器,利用位置滤波器进行定位,并构造33个尺度金字塔进行尺度检测,提高跟踪准确度.由于尺度金字塔导致跟踪效率降低,Danelljan<sup>[23]</sup>进一步提出快速尺度空间判别跟踪(fast Discriminative Scale Space Tracking, fDSST)算法,利用主成分分析的方法提高了算法的速度.同时,fDSST相比于其他同类算法,解耦了位置关系和尺度关系,提高了跟踪精度;采用三角插值,减少位置样本和尺度样本数量,提高运

行速度.但以上基于尺度池的跟踪算法都采用固定的尺度纵横比,存在一定的局限性,即当目标发生形变或平面内外旋转等导致纵横比变化时,无法精确地估计目标的尺度.

鉴于此,本文提出了基于纵横比自适应的相关滤波跟踪算法.该算法训练纵横比滤波器,用于目标尺度的精确跟踪;同时采用高置信度综合判据确定目标的遮挡情况,并设计平滑修正方案对遮挡目标进行位置补偿;最后自适应更新学习率,防止模型退化.

## 2 fDSST 算法

相关滤波的概念最早出现在信号领域中,用于描述两个信号之间的相关性或相似性,而卷积是求系统输出响应(零状态响应)的重要方法<sup>[24]</sup>.由相关和卷积的计算公式,相关和卷积具有相同的效果,时域内的卷积可以转换为频域内的点乘,从而减少计算量.

相关滤波的核心思想是两个目标越相似,其相关值越大;在视频图像序列中,候选区域与目标区域的相关值越大,表征该候选区域和初始目标越相似.因此,需要训练滤波器使目标模板和候选区域互相关的响应分数最大化.

对于二维图像,设 $f$ 为候选区域图像, $h$ 为训练的滤波器, $g$ 为互相关结果, $\otimes$ 表示卷积操作,则在时域有

$$g = f \otimes h \quad (1)$$

经过傅里叶变换后有(其中 $G, F, H$ 分别为 $g, f, h$ 对应的傅里叶变换结果, $\odot$ 为矩阵点乘符号):

$$G = F \odot H^* \quad (2)$$

从而得到理想的目标模板:

$$H^* = \frac{G}{F} \quad (3)$$

目标的位置信息和尺度信息相互耦合、相互联系.位置信息确定目标运动模型,有利于估计尺度;尺度信息有利于保护位置滤波器不受背景污染和陷入局部最优,保证跟踪的精确性和鲁棒性.MOSSE算法<sup>[7]</sup>采用单一的尺度跟踪目标,当目标发生形变时,跟踪尺度过大,学习大量背景特征,滤波器受到污染而跟踪失败;跟踪尺度过小,只提取了目标局部特征,发生跟踪漂移.

因此,DSST算法<sup>[22]</sup>采用相互独立的滤波器对目标的位置和尺度进行跟踪.在给定目标位置周围的矩形区域,构造尺度金字塔,设 $W \times Z$ 表示当前帧中目标的大小, $S$ 为尺度数量, $a$ 为尺度因子,金字塔的构造方式为

$$a^n W \times a^n Z, n \in \left\{ \left\lfloor -\frac{S-1}{2} \right\rfloor, \dots, \left\lfloor \frac{S-1}{2} \right\rfloor \right\} \quad (4)$$

利用这些尺度获取多个样本转化为响应值,根据

响应分数选择相应的尺度作为目标的最佳尺度. 但是 DSST 由于特征维度过高导致计算成本昂贵, fDSST 通过三角多项式插值和矩阵分解对特征进行降维, 将 33 个内插尺度池降为 17 个尺度池, 从而获得  $17 \times 17$  维的特征矩阵.

### 3 基于纵横比自适应的相关滤波跟踪算法

面对目标纵横比变化的挑战, 基于固定纵横比的多尺度跟踪器, 往往只能提取目标的局部信息或过多地提取到背景信息, 从而导致跟踪失败. 如图 1 所示, 绿框为真值框, 红框和蓝框是固定纵横比跟踪结果. 当目标发生形变或平面内外旋转等导致纵横比变化时, 使用固定纵横比跟踪目标, 目标框只包含部分目标(如红框)或包含背景(如蓝框), 都不利于目标的跟踪. 此外, 当发生遮挡或跟踪漂移时, 使用固定学习率更新模板, 滤波器会提取到错误的特征信息, 从而影响目标跟踪结果; 而当发生形变时, 过小的学习率又会使模型更新失败, 导致目标丢失.

因此, 本文以 fDSST 算法为基础框架, 提出基于纵横比自适应的相关滤波跟踪算法, 在线学习纵横比模型 (Aspect-Ratio Model, ARM) 提高尺度估计精度; 以及提出高置信度综合判据和学习率自适应更新机制, 判断目标状态, 自适应调整学习率; 此外, 位置更新部分, 提出遮挡平滑修正模型 (Occlusion Smoothing Correction Model, OSCS), 解决遮挡导致的相关滤波位置误判, 提高模型识别能力.

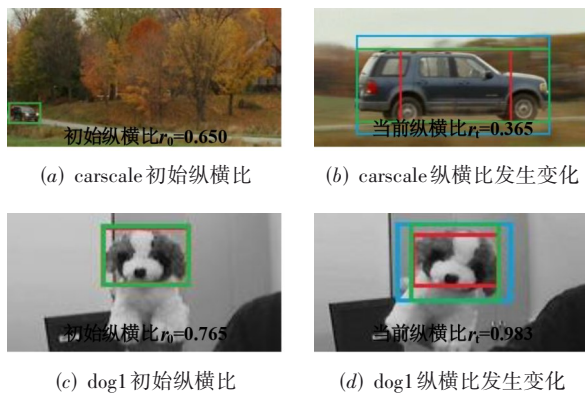


图1 固定纵横比跟踪

#### 3.1 纵横比滤波器

目标的状态信息为  $(x, y, w, z)$ , 分别表示中心位置和尺度大小, 基于固定纵横比的相关滤波状态方程如下:

$$\begin{cases} x_t = x_{t-1} + dx \\ y_t = y_{t-1} + dy \\ w_t = w_{t-1} \times S_t^* \\ z_t = z_{t-1} \times S_t^* \end{cases} \quad (5)$$

其中,  $t$  为图像帧数,  $(dx, dy)$  为位置滤波器获取的中心位置偏移量,  $S_t^*$  为尺度滤波器获取的最优尺度变化量.

在训练阶段, 通过最小化训练图像的误差平方和<sup>[7]</sup>来构造相关滤波器, 目标是设计最优的滤波器  $H^*$  最小化代价函数:

$$\min_{H^*} \sum_i |F_i H^* - G_i|^2 \quad (6)$$

求解式(6)得

$$H^* = \frac{\sum_i G_i \odot F_i^*}{\sum_i F_i \odot F_i^*} \quad (7)$$

在检测阶段, 通过式(2)获取目标候选区域的相关滤波响应, 根据响应分数确定目标的位置.

在更新阶段, 引入学习率  $\alpha$  更新模板:

$$H_t^* = \frac{A_t}{B_t} \quad (8)$$

$$A_t = \alpha G_t \odot F_t^* + (1 - \alpha) A_{t-1} \quad (9)$$

$$B_t = \alpha F_t \odot F_t^* + (1 - \alpha) B_{t-1} \quad (10)$$

其中,  $t$  为图像的帧数,  $A_t$  和  $B_t$  分别是第  $t$  帧模板  $H_t^*$  的分子和分母.

式(5)中尺度方程也可以表示为

$$\begin{cases} w_t = w_0 \times s_t \\ z_t = z_0 \times s_t \\ s_t = s_{t-1} \times S_t^* \end{cases} \quad (11)$$

方程(11)中,  $(w_0, z_0)$  为初始帧尺度大小,  $s_t$  为第  $t$  帧的总尺度变化量. 由式(11)可以看出, 穷举搜索策略基于初始帧的尺度信息进行跟踪, 其尺度纵横比  $r_0$  为固定值, 即

$$r_0 = z_0/w_0 = z_t/w_t \quad (12)$$

根据图 1, 固定纵横比尺度估计, 不能同时满足水平和垂直方向上的尺度和目标尺度一致, 因此需要训练两个相互独立的纵横比跟踪器在两个方向上对目标纵横比进行跟踪.

在位置和尺度估计的基础上进行纵横比的自适应调整, 即利用获取的位置中心、尺度信息和上一帧的纵横比确定感兴趣区域, 分别在水平方向和垂直方向上以非线性的指数缩放感兴趣区域采集样本, 这些样本和纵横比模板进行相关响应, 由响应峰值和纵横比更新机制自适应确定新一帧的纵横比信息. 如图 2 所示, 红色框是固定纵横比跟踪结果, 保持高度不变, 增减纵横比因子, 获取不同宽度的候选区域; 同理保持高度不变, 改变纵横比因子, 获取不同高度的候选区域. 对这些候选区域进行相关滤波, 选择响应得分最高的纵横比因子进行纵横比更新, 如选择绿色框作为当前帧的尺度. 跟踪模型如图 3 所示, 纵横比跟踪器进行训练、检测和更新.

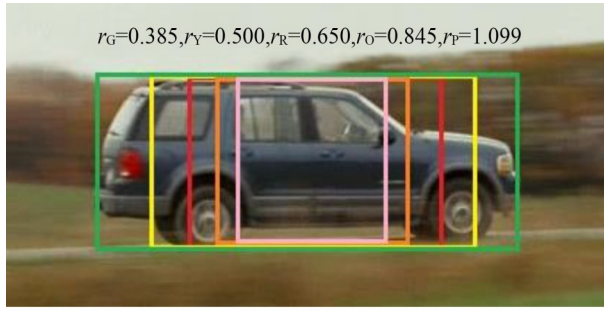


图2 纵横比跟踪器

在训练阶段,以目标位置、尺度和初始粗精度纵横比尺度池构造特征金字塔,提取特征 $f$ ,使用理想高斯信号作为期望的相关输出 $g$ ,最小化损失函数(9)求得初始模板 $H_0^w$ 和 $H_0^z$ ,分解得初始分子分母 $A_0^w$ 、 $B_0^w$ 、 $A_0^z$ 、 $B_0^z$ .

设置初始粗精度纵横比金字塔为

$$\begin{cases} b^m w_0^s \times z_0^s, m \in \left\{ \left\lfloor -\frac{S_r-1}{2} \right\rfloor, \dots, \left\lfloor \frac{S_r-1}{2} \right\rfloor \right\} \\ w_0^s \times b^m z_0^s, m \in \left\{ \left\lfloor -\frac{S_r-1}{2} \right\rfloor, \dots, \left\lfloor \frac{S_r-1}{2} \right\rfloor \right\} \end{cases} \quad (13)$$

其中, $S_r$ 为粗精度金字塔的个数, $b$ 为步长因子, $w_0^s \times z_0^s$ 是初始尺度.

在检测阶段,围绕先前估计的目标位置 $(x_t, y_t)$ 、尺度 $(w_t^s, z_t^s)$ 和粗精度纵横比尺度池构造特征金字塔,提取特征 $F_t^w$ 和 $F_t^z$ ,与模板 $H_{t-1}^w$ 和 $H_{t-1}^z$ 利用式(2)计算响应图 $R_t^w$ 和 $R_t^z$ ,在内插金字塔中由响应峰值 $resw$ 和 $resz$ 确定纵横比变化因子 $w_t^*$ 和 $z_t^*$ .内插金字塔如下:

$$\begin{cases} b^n r_{t-1}^w w_t^s \times r_{t-1}^z z_t^s, n \in \left\{ \left\lfloor -\frac{S_R-1}{2} \right\rfloor, \dots, \left\lfloor \frac{S_R-1}{2} \right\rfloor \right\} \\ r_{t-1}^w w_t^s \times b^n r_{t-1}^z z_t^s, n \in \left\{ \left\lfloor -\frac{S_R-1}{2} \right\rfloor, \dots, \left\lfloor \frac{S_R-1}{2} \right\rfloor \right\} \end{cases} \quad (14)$$

其中, $w_t^s$ 和 $z_t^s$ 是第 $t$ 帧的尺度信息, $r_{t-1}^w$ 和 $r_{t-1}^z$ 是第 $(t-1)$ 帧的纵横比因子,初始纵横比因子 $r_0^w = r_0^z = 1$ , $S_R$ 是内插金字塔的个数.

设计纵横比更新机制.响应分数表征目标置信度,尺度响应峰值 $res$ 和纵横比响应峰值 $resw$ 、 $resz$ 表征该尺度下目标最大置信度,选择最大置信度下的尺度和纵横比作为当前帧目标的最佳尺度信息,即

$$\begin{cases} r_t^w = r_{t-1}^w \times w_t^*, r_t^z = r_{t-1}^z, \text{if } resw > resz, resw > res \\ r_t^w = r_{t-1}^w, r_t^z = r_{t-1}^z \times z_t^*, \text{if } resz > resw, resz > res \\ r_t^w = r_{t-1}^w, r_t^z = r_{t-1}^z, \text{else} \end{cases} \quad (15)$$

在更新阶段,重新提取特征金字塔,转化为模板 $H_t^w$ 和 $H_t^z$ .通过矩阵分解得 $N_t^w$ 、 $D_t^w$ 、 $N_t^z$ 、 $D_t^z$ ,再引入学习率 $\beta$ 进行更新:

$$H_t^w = \frac{A_t^w}{B_t^w} \quad (16)$$

$$A_t^w = \beta N_t^w + (1 - \beta) A_{t-1}^w \quad (17)$$

$$B_t^w = \beta D_t^w + (1 - \beta) B_{t-1}^w \quad (18)$$

$$H_t^z = \frac{A_t^z}{B_t^z} \quad (19)$$

$$A_t^z = \beta N_t^z + (1 - \beta) A_{t-1}^z \quad (20)$$

$$B_t^z = \beta D_t^z + (1 - \beta) B_{t-1}^z \quad (21)$$

此时,尺度状态方程(11)改进为

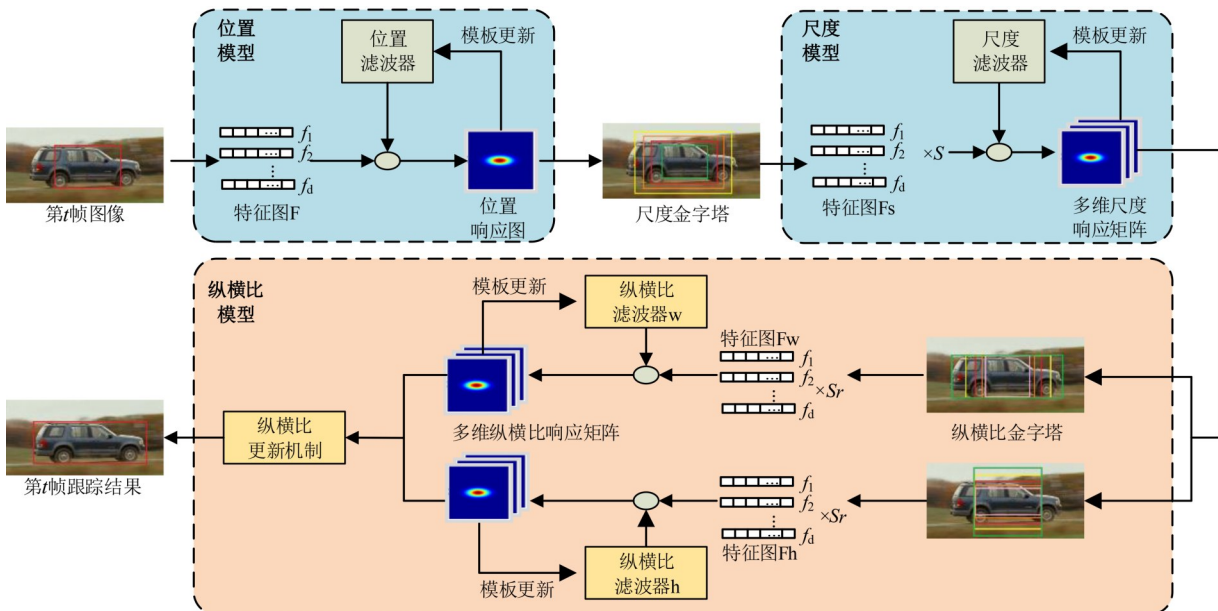


图3 纵横比跟踪模型

$$\begin{cases} w_t = w_0 \times r_t^w \times S_t \\ z_t = z_0 \times r_t^z \times S_t \\ s_t = s_{t-1} \times S_t^* \end{cases} \quad (22)$$

其中,总纵横比变化量  $r_t^w$  和  $r_t^z$  更新见式(15).

### 3.2 遮挡平滑修正模型

在高置信度时更新跟踪模型,可避免模型污染,减少更新次数,提高跟踪速度. 峰值旁瓣比(Peak to Sidelobe Ratio, PSR)<sup>[25]</sup>可揭示跟踪结果的置信水平,设响应图信号为 Res,则第  $t$  帧 PSR 计算公式为

$$P_t = \frac{F_{\max} - \mu_{\text{Res}}}{\sigma_{\text{Res}}} \quad (23)$$

其中,  $\mu_{\text{Res}}$  是旁瓣均值,  $\sigma_{\text{Res}}$  是旁瓣标准差,旁瓣是响应峰值周围  $11 \times 11$  像素空间的信号.

图4为采用 fDSST 算法不同遮挡情况下的输出响应图,其中三维图表示候选区域中目标位置的二维高斯分布. 当目标发生遮挡或跟踪失败相较于跟踪成功,  $F_{\max}$  变化不大,但 PSR 明显变小,因此,可通过 PSR 判断目标的遮挡情况.

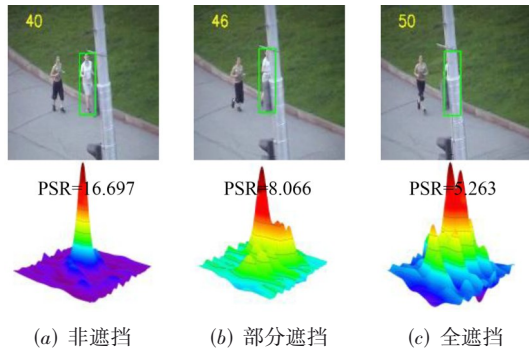


图4 不同遮挡情况下输出响应

由于  $F_{\max}$  可直接反映跟踪结果的好坏,结合 PSR 和  $F_{\max}$  提出高置信度综合判据,设计学习率自适应更新机制. 本文的自适应更新机制如下:若  $P_t$  小于阈值  $P_T$ ,则赋予较小的学习率  $\varepsilon$ ;若  $P_t$  大于阈值  $P_T$  且最大响应分数  $F_{\max}$  小于阈值  $F_T$ ,学习自适应调整为  $\rho F_{\max}^\delta$ ;若  $P_t$  和  $F_{\max}$  都大于相应阈值,则赋予较大的学习率  $\gamma_{\max}$ ,初始学习率设为  $\gamma_0$ .

此时位置跟踪器更新方程如下

$$H_t^* = \frac{A_t}{B_t} \quad (24)$$

$$A_t = \gamma G_t \odot F_t^* + (1 - \gamma) A_{t-1} \quad (25)$$

$$B_t = \gamma F_t \odot F_t^* + (1 - \gamma) B_{t-1} \quad (26)$$

目标发生遮挡时,仅利用响应分数去判断位置会出现偏差,本文在指数平滑模型<sup>[26]</sup>上进行改进,提出遮挡平滑修正模型,以实现目标的重新定位,具体公式如下:

$$x_t = x_{t-1} + dx_t \quad (27)$$

$$y_t = y_{t-1} + dy_t \quad (28)$$

$$dx_t = \sum_k \theta_{t-k} dx_{t-k} \quad (29)$$

$$dy_t = \sum_k \theta_{t-k} dy_{t-k} \quad (30)$$

$$\theta_{t-k} = \frac{P_{t-k}}{\sum_k P_{t-k}} \quad (31)$$

上式中,  $k$  表示修正步数,  $dx_{t-k}$  和  $dy_{t-k}$  表示第  $(t-k)$  帧水平和垂直方向上的位置偏移量,  $\theta_{t-k}$  是第  $(t-k)$  帧偏移量的权重,  $P_{t-k}$  是第  $(t-k)$  帧的峰值旁瓣比 PSR.

### 3.3 基于纵横比自适应的跟踪算法

在本文算法中,位置滤波器、尺度滤波器和纵横比滤波器串行地对目标的位置、尺度和纵横比进行估计,并且位置、尺度和纵横比是相互耦合、相互联系的. 本文算法的基本步骤如算法1所示.

#### 算法1 基于纵横比自适应的目标跟踪算法

输入: 视频第  $(t-1)$  帧图像、第  $(t-1)$  帧目标信息、第  $(t-k)$  帧到第  $(t-1)$  帧的位置偏移量以及纵横比跟踪模型.

输出: 视频第  $t$  帧目标信息以及第  $t$  帧位置偏移量.

1. 使用位置跟踪器获取位置偏移量  $(dx_{t-k}, dy_{t-k})$ .
2. 根据高置信度综合判据对目标进行判断:若目标跟踪正常,则通过式(5)定位;若发生遮挡,通过遮挡平滑修正模型补偿校正位置信息  $(x_t, y_t)$  并更新位置偏移量  $(dx_t, dy_t)$ .
3. 使用尺度滤波器获取尺度变化因子  $S_t^*$ .
4. 使用纵横比跟踪模型确定纵横比变化因子,利用式(15)对纵横比进行更新.
5. 根据学习率自适应更新机制对位置学习率  $\gamma$  进行更新.
6. 由式(24)-(26)更新位置跟踪器.
7. 由式(8)-(10)更新尺度跟踪器.
8. 由式(16)-(21)更新纵横比滤波器.

## 4 实验结果和分析

为了验证跟踪算法的有效性,本文在3个基准数据集 OTB100<sup>[27]</sup>, VOT2016<sup>[28]</sup> 以及 VOT2018<sup>[29]</sup> 上与19种先进跟踪算法进行对比实验. 其中基于相关滤波的跟踪算法包括 DSST<sup>[22]</sup>, fDSST<sup>[23]</sup>, LDES<sup>[30]</sup> (Large Displacement Estimation of Similarity transformation), CSRDCF<sup>[31]</sup> (Discriminative Correlation Filter with Channel and Spatial Reliability), MKCFup<sup>[32]</sup> (Multiple Kernelized Correlation Filters up), SITUP<sup>[33]</sup> (Scale Invariant Tracking using Average Peak-to-Correlation), STRCF<sup>[34]</sup> (Spatial-Temporal Regularized Correlation Filters) 和 WSCFst<sup>[35]</sup> (Weighted Sample based CF tracking); 基于深度学习的跟踪算法包括 CFnet2-tri<sup>[36]</sup> (Triplet loss in 2 convolutional layers with CF layer in Siamese network), SiamFC-

tri<sup>[36]</sup> (Triplet loss in Fully Convolutional Siamese network), SiamImp-tri<sup>[36]</sup> (Triplet loss in improved Siamese network), USOT<sup>[37]</sup> (Unsupervised Single Object Tracker), USOT\*<sup>[37]</sup> (Unsupervised Single Object Tracker star), LUDT<sup>[38]</sup> (Learning Unsupervised Deep Tracking), SiamSNN<sup>[39]</sup> (Siamese Spiking Neural Networks), Spiking SiamFC++<sup>[40]</sup> (Spiking Fully Convolutional Siamese tracker++), RLS-RTMDNet<sup>[41]</sup> (Recursive LSE-aided Real-Time Multi-Domain Convolutional Neural Networks), DOTCL<sup>[42]</sup> (Deep Object Tracking model with Circular Loss Function) 和 MFST<sup>[43]</sup> (Multiple Features Siamese Tracker).

#### 4.1 实验配置

本文实验环境为 64 位 Windows 10 操作系统, CPU: Inter Core i5-12400F, 主频为 2.5 GHz, Python3.8.1 实验. 参数设置: 粗精度纵横比金字塔个数  $S_r=17$ , 内插纵横比金字塔个数  $S_R=33$ , 步长因子  $b=1.025$ , 学习率  $\beta=0.025$ ,  $P_T=10$ ,  $F_T=0.5$ , 修正步数  $k=5$ , 初始学习率  $\gamma_0=0.025$ , 低置信度学习率  $\varepsilon=0$ , 高置信度学习率  $\gamma_{\max}=0.025$ , 更新系数  $\rho=0.3$ , 更新指数  $\delta=0.5$ , 其余参数和 fDSST 算法一致.

#### 4.2 评价指标

OTB100 数据集评价指标包括准确率和成功率, 其中准确率是中心位置误差在给定阈值  $s_0$  以内的视频帧数占总帧数的百分比, 一般阈值  $s_0$  为 20 个像素; 成功率是交并比  $\text{IOU} = \frac{S_T \cap S_G}{S_T \cup S_G}$  大于给定阈值  $i_0$  的视频帧数占总帧数的百分比, 通常阈值  $i_0$  取 50%, 其中  $S_T$  表示跟踪算法对应的矩形框面积,  $S_G$  表示 ground-truth 对应的矩形框面积.

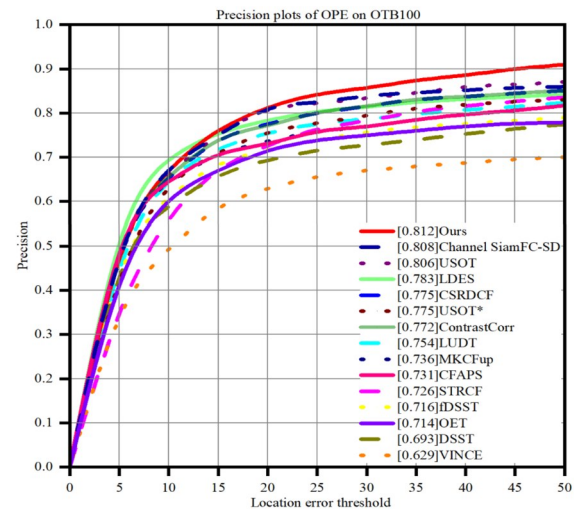
VOT2016 和 VOT2018 数据集评价指标包括平均重叠期望、准确性和鲁棒性, 分别用 EAO (Expected Average Overlap)、A (Accuracy) 和 R (Robustness) 表示. 平均重叠期望是从测试序列截取短序列进行不重新初始化跟踪的有效帧平均重叠率, 可同时反映跟踪器的准确性和鲁棒性; 准确性表示预测框和真实框的平均重叠率; 鲁棒性定义为跟踪失败 (交并比为 0 判定失败) 的平均次数, 用以评价跟踪目标的稳定性, 其数值越大, 稳定性越差.

#### 4.3 实验结果分析

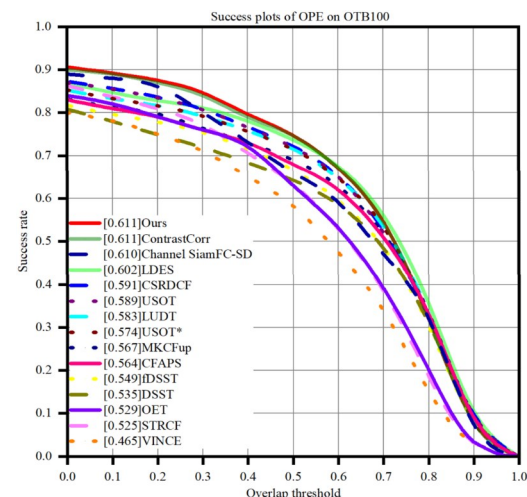
##### 4.3.1 OTB100 实验结果分析

在本节中, 将本文算法与 DSST<sup>[22]</sup>, fDSST<sup>[23]</sup>, LDES<sup>[30]</sup>, CSRDCF<sup>[31]</sup>, MKCFup<sup>[32]</sup>, SITUP<sup>[33]</sup>, STRCF<sup>[34]</sup>, WSCFst<sup>[35]</sup>, CFnet2-tri<sup>[36]</sup>, SiamFC-tri<sup>[36]</sup>, SiamImp-tri<sup>[36]</sup>, USOT<sup>[37]</sup>, USOT\*<sup>[37]</sup> 和 LUDT<sup>[38]</sup> 在 OTB 数据集上进行对比.

图 5 展示了这些算法在 OTB100 上的总体准确率和总体成功率, 本文算法在准确率和成功率方面都排第一. 在准确率图中, 本文算法总体的准确率得分是 0.812, 相比于 fDSST 提高了 9.6%; 在成功率图中, 本文算法总体的成功率得分是 0.611, 相比于 fDSST 提高了 6.2%. 为进一步验证算法在各种挑战下的鲁棒性, 这些算法在 11 种挑战下进行对比实验, 表 1、表 2 和表 3 分别是挑战序列上的准确率、成功率和速度, 红色、绿色和蓝色分别表示前三名, 本文算法准确率和成功率基本取得前三的成绩, 表明本文算法在各种挑战下均取得鲁棒性结果, 速度也满足了实时性要求.



(a) 总体准确率图



(b) 总体成功率图

图 5 OTB 总体准确率和总体成功率

实验的部分跟踪视频截图如图 6 所示, 左为尺度挑战, 右为遮挡挑战. 下面针对跟踪视频序列进行实验结果分析.

(1)CarScale:在初始帧确定目标的初始纵横比,当目标纵横比变化不大时,如第50帧,所有算法都跟踪良好;第190帧目标纵横比变化较大,基于固定纵横比的尺度估计算法对于目标的尺度跟踪效果不佳,而本文算法采用纵横比自适应模型,准确跟踪;第240帧,大部分算法尺度估计误差较大使得模型发生漂移,而本文算法始终准确估计目标位置和尺度,展现了本文算法纵横比模型在尺度变化上的有效性以及算法的优越性.

(2)Skating2-1:目标在运动过程姿态变化导致尺度变化,第10帧fDSST算法就已经发生了跟踪漂移;到

140帧,累积的错误尺度估计使得大部分算法发生了跟踪漂移,尺度估计器也深受污染;在第200帧,除了本算法,所有跟踪算法都跟踪失败,尺度估计完全失效,验证了纵横比模型的先进性.

(3)Human2:目标在第250帧发生了遮挡,本文算法准确检测并跟踪目标,部分算法发生了跟踪漂移以及尺度估计不佳,在第700帧目标运动导致尺度变化以及第960帧目标形态变化时,由于本文算法使用自适应纵横比,都很好定位和跟踪上目标,fDSST跟踪失败,其余算法发生了跟踪漂移以及尺度估计错误,展现了本文算法在尺度变化挑战下的优势.

表1 15种算法在11种挑战序列上的准确率

算法	BC	DEF	FM	IPR	IV	LR	MB	OCC	OPR	OV	SV
Ours	0.805	0.731	0.729	0.832	0.846	0.835	0.733	0.783	0.813	0.750	0.795
LDES	0.681	0.709	0.750	0.776	0.716	0.866	0.707	0.748	0.766	0.623	0.757
CSRDCF	0.752	0.746	0.782	0.728	0.781	0.796	0.774	0.700	0.720	0.745	0.720
LUDT	0.749	0.658	0.714	0.751	0.685	0.699	0.683	0.694	0.738	0.648	0.709
MKCFup	0.667	0.661	0.656	0.705	0.718	0.778	0.642	0.718	0.709	0.613	0.663
fDSST	0.780	0.591	0.675	0.717	0.728	0.670	0.660	0.618	0.652	0.574	0.664
DSST	0.677	0.556	0.572	0.729	0.709	0.662	0.570	0.594	0.662	0.506	0.662
STRCF	0.737	0.685	0.676	0.714	0.690	0.839	0.655	0.657	0.715	0.680	0.721
USOT	0.757	0.771	0.836	0.831	0.828	0.963	0.790	0.674	0.797	0.717	0.802
USOT*	0.691	0.737	0.818	0.802	0.735	0.808	0.799	0.683	0.781	0.682	0.790
SITUP	0.776	0.703	0.701	0.739	0.741	0.675	0.705	0.742	0.765	0.665	0.748
WSCFst	0.799	0.735	0.722	0.813	0.821	0.761	0.736	0.731	0.765	0.720	0.751
CFnet2-tri	0.791	0.713	0.717	0.746	0.757	0.834	0.695	0.740	0.741	0.731	0.759
SiamFC-tri	0.715	0.680	0.763	0.774	0.746	0.900	0.727	0.726	0.763	0.723	0.748
SiamImp-tri	0.737	0.711	0.739	0.801	0.730	0.843	0.654	0.732	0.765	0.611	0.730

表2 15种算法在11种挑战序列上的成功率

算法	BC	DEF	FM	IPR	IV	LR	MB	OCC	OPR	OV	SV
Ours	0.601	0.536	0.565	0.602	0.640	0.578	0.585	0.582	0.595	0.567	0.595
LDES	0.535	0.523	0.589	0.577	0.556	0.592	0.574	0.575	0.577	0.504	0.572
CSRDCF	0.570	0.552	0.600	0.535	0.605	0.506	0.600	0.548	0.537	0.540	0.548
LUDT	0.571	0.504	0.578	0.564	0.541	0.502	0.560	0.538	0.560	0.511	0.549
MKCFup	0.526	0.505	0.520	0.529	0.569	0.537	0.514	0.547	0.532	0.472	0.513
fDSST	0.586	0.457	0.543	0.541	0.556	0.429	0.531	0.474	0.494	0.458	0.505
DSST	0.529	0.431	0.480	0.538	0.566	0.520	0.481	0.468	0.496	0.415	0.501
STRCF	0.554	0.492	0.513	0.497	0.525	0.518	0.494	0.487	0.505	0.493	0.509
USOT	0.544	0.545	0.629	0.595	0.604	0.615	0.593	0.496	0.575	0.507	0.583
USOT*	0.511	0.533	0.623	0.575	0.553	0.548	0.609	0.513	0.571	0.496	0.588
SITUP	0.572	0.507	0.542	0.540	0.558	0.395	0.574	0.558	0.558	0.520	0.533
WSCFst	0.593	0.552	0.560	0.592	0.622	0.441	0.566	0.554	0.564	0.483	0.553
CFnet2-tri	0.595	0.528	0.566	0.555	0.581	0.546	0.568	0.559	0.552	0.537	0.578
SiamFC-tri	0.542	0.501	0.585	0.580	0.579	0.615	0.567	0.549	0.563	0.543	0.563
SiamImp-tri	0.565	0.521	0.575	0.587	0.565	0.596	0.528	0.561	0.560	0.468	0.553

表 3 15种算法在11种挑战序列上的速度

算法	BC/fps	DEF/fps	FM/fps	IPR/fps	IV/fps	LR/fps	MB/fps	OCC/fps	OPR/fps	OV/fps	SV/fps
Ours	46.56	42.15	35.48	53.43	40.68	131.99	29.86	47.26	49.87	49.45	49.47
LDES	12.38	10.91	9.94	11.44	10.74	11.23	8.63	10.85	11.05	11.05	10.72
CSRDCF	22.29	21.19	19.69	22.36	21.50	18.69	19.42	19.79	21.29	19.99	20.92
LU DT	69.19	68.15	66.90	71.28	68.85	77.55	65.15	67.77	69.77	68.69	70.18
MKCFup	51.25	49.97	44.21	52.83	48.73	73.51	39.56	49.31	51.10	48.22	50.44
fDSST	63.66	56.30	49.74	82.78	56.41	238.84	40.36	69.10	73.87	72.25	76.33
DSST	24.82	21.53	20.86	32.13	27.64	67.47	18.20	24.86	28.41	28.64	28.65
STRCF	34.63	34.17	33.14	34.85	34.36	42.37	32.58	34.01	34.60	34.89	34.47
USOT	74.13	73.02	71.75	76.38	73.77	83.09	69.81	72.62	74.75	73.59	75.19
USOT*	74.20	73.03	71.67	75.62	72.29	83.26	68.41	70.44	77.98	71.38	74.43
SITUP	35.61	30.98	31.26	45.82	33.26	89.18	26.96	34.30	41.33	32.45	41.52
WSCFst	57.98	57.81	53.92	60.91	58.76	71.08	50.95	56.27	59.05	54.31	58.11
CFnet2-tri	56.19	54.74	54.20	55.45	55.24	57.39	53.79	54.91	54.97	54.54	55.06
SiamFC-tri	84.51	80.48	80.37	82.89	83.43	87.55	79.87	82.08	82.63	80.62	80.83
SiamImp-tri	53.40	51.04	51.54	52.74	52.72	55.82	51.15	52.24	52.66	51.68	51.14

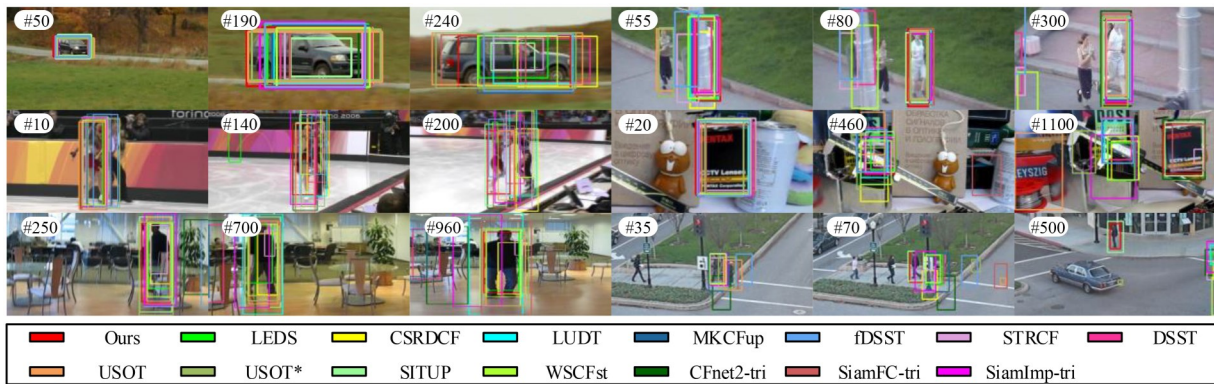


图 6 跟踪视频部分截图

(4) Jogging-2: 当目标经过电线杆(第 55 帧)时经历部分遮挡和完全遮挡, fDSST 发生了跟踪漂移, 本文算法由于高置信度综合判据和遮挡平滑修正模型对位置进行预测; 当第 80 帧和第 300 帧目标重新出现时, 由于遮挡期间自适应更新学习率, 保护了位置跟踪器, 本文算法跟踪定位准确, 表明了高置信度综合判据、遮挡平滑修正模型以及学习率自适应更新机制的有效性。

(5) Box: 目标在第 460 帧发生局部遮挡, 本文算法通过高置信度综合判据以及学习率自适应机制使遮挡期间保持低学习率更新, 防止跟踪器污染, 从而一直到 1 100 帧都只有本文算法跟踪上目标, 说明了本文算法遮挡平滑修正模型的优越性。

(6) Human3: 目标分别在 35 帧和 70 帧前经历两次遮挡后, 只有本文算法和 SITUP 算法能够准确跟踪目标, 其余算法都跟踪失败, 验证了本文算法在遮挡挑战下的鲁棒性。

### 4.3.2 VOT 实验结果分析

在本节中, 选取了 fDSST<sup>[23]</sup>、Spiking SiamFC++<sup>[40]</sup>等

性能优异的跟踪算法与本文算法进行对比与分析。实验结果如表 4 和表 5 所示, 其中红色、绿色和蓝色分别表示前三名, 本文算法各项指标都取得前三的成绩。

表 4 9种算法在 VOT2016 上性能对比

算法	EAO ↑	A ↑	R ↓
Ours	0.344	0.590	0.308
fDSST	0.177	0.531	0.732
Spiking SiamFC++	0.302	0.600	0.359
LU DT	0.231	0.540	0.466
SiamSNN	0.210	0.497	0.630
RLS-RTMDNet	0.336	0.573	0.255
LDES	0.266	0.586	0.424
MKCFup	0.266	0.533	0.704
STRCF	0.256	0.520	0.359

表 4 为 VOT2016 上的不同算法性能对比, 和基线算法 fDSST 相比, 本文算法在 EAO 指标上提升了 16.7%, 在准确性指标上提高了 5.9%, 在鲁棒性指标上降低了 42.4%; 表 5 为 VOT2018 上的不同算法性能对

表5 9种算法在VOT2018上性能对比

算法	EAO ↑	A ↑	R ↓
Ours	0.209	0.562	0.581
fDSST	0.104	0.496	1.218
LUDT	0.154	0.460	0.698
SiamSNN	0.176	0.460	0.860
DOTCL	0.396	0.533	0.151
MFST	0.200	0.497	0.438
LDES	0.179	0.533	0.660
MKCFup	0.110	0.499	1.138
STRCF	0.162	0.480	0.613

比,本文算法在EAO指标上提升了10.5%,在准确性指标上提高了6.6%,在鲁棒性指标上降低了63.7%。以上表明本文算法在准确性和鲁棒性上得到了增强。

### 4.3.3 消融实验分析

为测试本文算法所提出的两个策略对跟踪性能的影响,采用准确率、成功率和速度作为评价指标在OTB100上进行消融实验,表6为OTB100的消融实验结果,其中ARM和OSCS分别表示纵横比模型和遮挡平滑修正模型。

由表6可知:单独加入纵横比模型后,总体准确率和成功率分别提高了3.2%和3.0%,运行效率降低了29.0%;单独加入遮挡平滑修正模型后,总体准确率和成功率提高了2.4%和1.4%,运行效率基本不变。以上表明了纵横比模型和遮挡平滑修正模型的有效性,而纵横比模型由于训练了纵横比跟踪器导致运行效率降低。

表6 消融实验结果

算法	总体准确率	总体成功率	速度/fps	SV准确率	SV成功率
基线算法	0.716	0.549	70.83	0.664	0.505
+ARM	0.748	0.579	50.28	0.730	0.559
+OSCS	0.740	0.563	69.73	0.715	0.539
Ours(+ARM+OSCS)	0.812	0.611	48.33	0.795	0.595

结合两种策略后,准确率和成功率都优于单一改进策略算法,表明了共同作用的改进策略可协同提高跟踪效果。相较于fDSST算法,本文算法总体准确率和成功率上分别提高了9.6%和6.2%,运行速度满足实时性要求;在尺度变换挑战上准确率和成功率分别提高了13.1%和9.0%,验证了本文算法在尺度挑战上的鲁棒性。

## 5 结论

相关滤波跟踪算法使用固定纵横比模型,会导致目标尺度估计不准确。针对这一问题,本文改进fDSST算法,提出了基于纵横比自适应的相关滤波跟踪算法,在线学习纵横比模型,实现更精确的目标尺度估计。此外,本文针对目标的遮挡问题,提出高置信度综合判据

确定目标的遮挡情况,并设计平滑修正方案对遮挡目标进行位置补偿,最后自适应更新学习率保护模型和提高识别能力。在OTB100,VOT2016以及VOT2018基准数据集上的对比评估验证了本文算法,取得了优越的鲁棒性结果。

### 参考文献

- [1] 孟璟, 杨旭. 目标跟踪算法综述[J]. 自动化学报, 2019(7): 17.  
MENG L, YANG X. A survey of object tracking algorithms[J]. Acta Automatica Sinica, 2019, 45(7): 1244-1260. (in Chinese)
- [2] 文志强, 朱艳辉, 彭召意. 粒子滤波目标跟踪中的有效粒子数控制方法[J]. 控制与决策, 2013, 28(9): 1349-1354, 1360.  
WEN Z Q, ZHU Y H, PENG Z Y. Control method of effective particle number in particle filter object tracking[J]. Control and Decision, 2013, 28(9): 1349-1354, 1360. (in Chinese)
- [3] 吴良健, 况璐, 邓庆林, 等. 基于Camshift和Kalman滤波结合的改进多目标跟踪算法[J]. 现代科学仪器, 2010, 129(1): 29-33,38.  
WU L J, KUANG L, DENG Q L, et al. Improved tracking algorithm for multiple targets based on camshift algorithm combined with Kalman filter [J]. Modern Scientific Instruments, 2010, 129(1): 29-33, 38. (in Chinese)
- [4] 刘向前, 闫娟, 杨慧斌, 等. 基于改进光流法的目标跟踪技术研究[J]. 上海工程技术大学学报, 2021, 35(3): 237-242.  
LIU X Q, YAN J, YANG H B, et al. Research on target tracking based on improved optical flow method[J]. Journal Of Shanghai University Of Engineering Science, 2021, 35(3): 237-242. (in Chinese)
- [5] ZHANG C, ZHANG Y F, GAO X P, et al. An improved meanshift tracking algorithm using adaptive quantization step in color space[C]//2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM). Piscataway: IEEE, 2019: 224-230.
- [6] 黄月平, 李小锋, 杨小冈, 等. 基于相关滤波的视觉目标跟踪算法新进展[J]. 系统工程与电子技术, 2021, 43(8): 2051-2065.  
HUANG Y P, LI X F, YANG X G, et al. Advances in visual object tracking algorithm based on correlation filter[J]. Systems Engineering and Electronics, 2021, 43(8): 2051-2065. (in Chinese)
- [7] BOLME D S, BEVERIDGE J R, DRAPER B A, et al. Visual object tracking using adaptive correlation filters[C]//The Twenty-Third IEEE Conference on Computer Vision

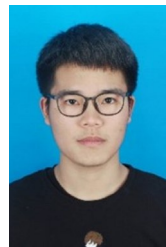
- and Pattern Recognition. Piscataway: IEEE, 2010: 13-18.
- [8] HENRIQUES J F, RUI C, MARTINS P, et al. Exploiting the circulant structure of tracking-by-detection with kernels[C]//Proceedings of the 12th European Conference on Computer Vision. Heidelberg: Springer, 2012:1-12.
- [9] HENRIQUES J F, CASEIRO 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] 刘艺, 李蒙蒙, 郑奇斌, 等. 视频目标跟踪算法综述[J]. 计算机科学与探索, 2022, 16(7): 1504-1515.  
LIU Y, LI M M, ZHENG Q B, et al. Survey on video object tracking algorithms[J]. Journal of Frontiers of Computer Science and Technology, 2022, 16(7): 1504-1515. (in Chinese)
- [11] 李豪, 袁广林, 秦晓燕, 等. 基于空间加权对数似然比相关滤波与 Deep Snake 的目标轮廓跟踪[J]. 电子学报, 2023, 51(1): 105-116.  
LI H, YUAN G L, QIN X Y, et al. Object contour tracking based on correlation filters with spatially-weighted logarithm likelihood ratio and deep snake[J]. Acta Electronica Sinica, 2023, 51(1): 105-116. (in Chinese)
- [12] 王叶, 刘强, 卿粼波, 等. 基于自适应空间正则化和畸变抑制的相关滤波跟踪[J]. 光电工程, 2021, 48(1): 27-37.  
WANG Y, LIU Q, QIN L B, et al. Learning adaptive spatial regularization and aberrance repression correlation filters for visual tracking[J]. Opto-Electronic Engineering, 2021, 48(1): 200068. (in Chinese)
- [13] 谢青松, 刘晓庆, 安志勇, 等. 基于前景优化的视觉目标跟踪算法[J]. 电子学报, 2022, 50(7): 1558-1566.  
XIE Q S, LIU X Q, AN Z Y, et al. Visual object tracking algorithm based on foreground optimization[J]. Acta Electronica Sinica, 2022, 50(7): 1558-1566. (in Chinese)
- [14] 田昊东, 张津浦, 王岳环. 稀疏约束的时空正则相关滤波无人机视觉跟踪[J]. 中国图象图形学报, 2023, 28(2): 458-470.  
TIAN H D, ZHANG J P, WANG Y H. Sparse constraint and spatial-temporal regularized correlation filter for UAV tracking[J]. Journal of Image and Graphics, 2023, 28(2): 458-470. (in Chinese)
- [15] 王法胜, 贺冰, 孙福明, 等. 自适应内容感知空间正则化相关滤波跟踪算法[J]. 吉林大学学报(工学版), 2023, 2: 1-15.  
WANG F S, HE B, SUN F M, et al. Adaptive content aware spatially-regularized correlation filter for object tracking[J]. Journal of Jilin University (Engineering and Technology Edition), 2023, 2: 1-15. (in Chinese)
- [16] ZHANG J, HE Y, WANG S. Learning adaptive sparse spatially-regularized correlation filters for visual tracking[J]. IEEE Signal Processing Letters, 2023, 30: 11-15.
- [17] XIAO L, NIE F, SHAO J, et al. The correlation filter with adaptive spatial and temporal regularization for inland ship tracking[C]//2022 8th International Conference on Systems and Informatics (ICSAI). Piscataway: IEEE, 2022: 1-6.
- [18] ZHOU Z, SUN Q, LI H, et al. Regression-selective feature-adaptive tracker for visual object tracking[J]. IEEE Transactions on Multimedia, 2023, 25: 5444-5457.
- [19] 陶洋, 唐函, 欧双江, 等. 稀疏约束与时间一致的背景感知相关滤波目标跟踪[J]. 小型微型计算机系统, 2024, 3: 1-8.  
TAO Y, TANG H, OU S J, et al. Sparse constrained and time-consistent background-aware correlation filtered target tracking[J]. Journal of Chinese Computer Systems, 2024, 3:1-8. (in Chinese)
- [20] 姜文涛, 孟庆姣. 自适应时空正则化的相关滤波目标跟踪[J]. 智能系统学报, 2023, 18(4): 754-763.  
JIANG W T, MENG Q J. Correlation filter tracking for adaptive spatiotemporal regularization[J]. CAAI Transactions on Intelligent Systems, 2023, 18(4): 754-763. (in Chinese)
- [21] LI Y, ZHU J K. A Scale Adaptive Kernel Correlation Filter Tracker with Feature Integration[C]//Computer Vision - ECCV 2014 Workshops. Cham: Springer International Publishing, 2015: 254-265.
- [22] DANELLJAN M, HÄGER G, SHAHBAZ KHAN F, et al. Accurate scale estimation for robust visual tracking[C]//Proceedings of the British Machine Vision Conference 2014. London: British Machine Vision Association, 2014: 1-11.
- [23] DANELLJAN M, HAGER G, KHAN F S, et al. Discriminative scale space tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(8): 1561-1575.
- [24] 程佩青. 数字信号处理教程: 经典版[M]. 北京: 清华大学出版社, 2015.  
CHENG P Q. Digital Signal Processing Tutorial: Classic edition[M]. Beijing: Tsinghua University Press, 2015. (in Chinese)
- [25] LI Y, ZHU J K, HOI S C H. Reliable patch trackers: Robust visual tracking by exploiting reliable patches[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2015: 1-11.
- [26] 胡显东, 陈伯孝, 王俊. 基于指数平滑的动态规划检测前跟踪算法[J]. 电波科学学报, 2016, 31(3): 468-472, 478.

- HU X D, CHEN B X, WANG J. Dynamic programming track-before-detect based on exponential smoothing method[J]. Chinese Journal Of Radio Science, 2016, 31(3): 468-472, 478. (in Chinese)
- [27] WU Y, LIM J, YANG M H. Object tracking benchmark[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 37(9): 1834-1848.
- [28] KRISTAN M. The visual object tracking VOT2015 challenge results[C]//2015 IEEE International Conference on Computer Vision Workshop (ICCVW). Piscataway: IEEE, 2015: 777-823.
- [29] KRISTAN M. The sixth visual object tracking VOT2018 challenge results[C]//Proceedings of the European Conference on Computer Vision (ECCV). Piscataway: IEEE, 2018: 3-53.
- [30] LI Y, ZHU J, HOI S, et al. Robust estimation of similarity transformation for visual object tracking: Association for the advancement of artificial intelligence (AAAI), 10.1609/AAAI.V33I01.33018666[P]. 2019.
- [31] LUKEZIC A, VOJIR T, ZAJC L C, et al. Discriminative correlation filter with channel and spatial reliability[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2017: 1-12.
- [32] ZHOU Y, WANG T, HU R H, et al. Multiple kernelized correlation filters (MKCF) for extended object tracking using X-band marine radar data[J]. IEEE Transactions on Signal Processing, 2019, 67(14): 3676-3688.
- [33] MA H Y, ACTON S T, LIN Z L. SITUP: Scale invariant tracking using average peak-to-correlation energy[J]. IEEE Transactions on Image Processing, 2020, 29: 3546-3557.
- [34] LI F, TIAN C, ZUO W M, et al. Learning spatial-temporal regularized correlation filters for visual tracking[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 4904-4913.
- [35] HAN R Z, FENG W, WANG S. Fast learning of spatially regularized and content aware correlation filter for visual tracking[J]. IEEE Transactions on Image Processing, 2020, 29: 7128-7140.
- [36] DONG X P, SHEN J B. Triplet loss in siamese network for object tracking[C]//Proceedings of the European Conference on Computer Vision (EC-CV). Piscataway: IEEE, 2018: 472-488.
- [37] ZHENG J L, MA C, PENG H W, et al. Learning to track objects from unlabeled videos[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2021: 13526-13535.
- [38] WANG N, ZHOU W G, SONG Y B, et al. Unsupervised deep representation learning for real-time tracking[J]. International Journal of Computer Vision, 2021, 129(2): 400-418.
- [39] LUO Y H, XU M, YUAN C H, et al. SiamSNN: Siamese spiking neural networks for energy-efficient object tracking[M]//Lecture Notes in Computer Science. Cham: Springer International Publishing, 2021: 182-194.
- [40] XIANG S Y, ZHANG T, JIANG S Q, et al. Spiking Siam-FC++: Deep spiking neural network for object tracking[J]. Nonlinear Dynamics, 2024, 112(10): 8417-8429.
- [41] GAO J, HU W M, LU Y. Recursive least-squares estimator-aided online learning for visual tracking[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2020: 1-13.
- [42] BELYAEV V, MALYSHEVA A, SHPILMAN A. End-to-end deep object tracking with circular loss function for rotated bounding box[C]//2019 XVI International Symposium "Problems of Redundancy in Information and Control Systems" (REDUNDANCY). Piscataway: IEEE, 2019: 1-12.
- [43] LI Z X, BILODEAU G A, BOUACHIR W. Multiple convolutional features in Siamese networks for object tracking[J]. Machine Vision and Applications, 2021, 32(3): 59.

#### 作者简介



钟钰彬 男,1999年出生,江西赣州人. 硕士研究生,主要研究方向为计算机视觉、目标跟踪.  
E-mail: 121121023652@njjust.edu.cn



杨鹏 男,1998年出生,安徽六安人. 博士研究生,主要研究方向为人工智能、计算机视觉.  
E-mail: pyang\_15@163.com



窦磊 男,1974年出生,江苏南京人. 博士研究生,南京理工大学博士生导师,主要研究方向为人工智能、导航制导与控制.  
E-mail: douleinj@163.com