

采用局部-全局区域重检测机制的 无人机长期跟踪算法

黄鹤^{1,2}, 马浩然^{1,2}, 刘国权¹, 王会峰², 高涛³, 张科^{1,2}

(1. 长安大学西安市智慧高速公路信息融合与控制重点实验室, 710064, 西安; 2. 长安大学
电子与控制工程学院, 710064, 西安; 3. 长安大学数据科学与人工智能研究院, 710064, 西安)

摘要: 为解决基础跟踪器面对遮挡和移出视野等长期跟踪场景时易出现跟踪失败等问题,提出了一种基于局部-全局区域重检测的无人机长期跟踪算法。设计了基础滤波器,将高置信度样本与其结合,并融入自适应时空正则化,解决了滤波器退化问题,提高了模型鲁棒性以及复杂场景下的性能;优化了滤波器更新策略,通过评价跟踪结果进行自适应更新;设计快速尺度滤波器,解决了跟踪过程中的尺度变化问题;设计了局部-全局区域重检测机制,跟踪失败时启动重检测器恢复跟踪目标,先完成局部区域重检测,若恢复跟踪失败,再利用全局区域重检测器继续恢复目标跟踪状态。实验结果表明:所提算法在 UAV 20L 数据集上的精确度和准确率分别可达 0.724 和 0.621,与基于时空正则化相关滤波器的跟踪算法(STRCF)相比分别提升了 25.9%和 20.6%,与同类主流算法相比,跟踪效果得到提升,证明了算法的有效性。

关键词: 无人机;长期跟踪;相关滤波器;重检测器;快速尺度滤波;高置信度

中图分类号: TP391.4 **文献标志码:** A

DOI: 10.7652/xjtuxb202406001 **文章编号:** 0253-987X(2024)06-0001-13

A Long-Term Unmanned Aerial Vehicle Tracking Algorithm Using Local-Global Region Redetection Mechanism

HUANG He^{1,2}, MA Haoran^{1,2}, LIU Guoquan¹, WANG Huifeng², GAO Tao³, ZHANG Ke^{1,2}

(1. Xi'an Key Laboratory of Intelligent Expressway Information Fusion and Control, Chang'an University, Xi'an 710064, China; 2. School of Electronic and Control Engineering, Chang'an University, Xi'an 710064, China; 3. Institute of Data Science and Artificial Intelligence, Chang'an University, Xi'an 710064, China)

Abstract: In order to solve the problem that the basic tracker is prone to tracking failure in long-term tracking scenarios such as occlusion and out of view, a long-term tracking algorithm for unmanned aerial vehicle (UAV) based on local-global region redetection is proposed. The basic filter is designed, and the high-confidence samples are combined with it, and the adaptive spatio-temporal regularization is integrated to solve the filter degradation problem and improve the robustness of the model and its performance in complex scenarios. The filter update strategy is optimized, and the adaptive update is performed by evaluating the tracking results. A fast scale filter is designed to solve the problem of scale change in the tracking process. A local-global

收稿日期: 2023-10-11。 作者简介: 黄鹤(1979—),男,教授,博士生导师;刘国权(通信作者),男,副教授,硕士生导师。

基金项目: 国家自然科学基金资助项目(62341301);西安市智慧高速公路信息融合与控制重点实验室(长安大学)开放基金资助项目(300102323502);中央高校基本科研业务费资助项目(300102324501)。

网络出版时间: 2024-01-24

网络出版地址: <https://link.cnki.net/urlid/61.1069.T.20240124.1355.006>

region redetection mechanism is designed. When the tracking fails, the re-detector is started to recover the tracking target and the local region re-detection is completed first. If the tracking recovery fails, the global region re-detector is used to continue to recover the target tracking state. The experimental results show that the precision and accuracy of the proposed algorithm on the UAV20L data set can reach 0.724 and 0.621 respectively, representing improvement of 25.9% and 20.6% respectively compared with the STRCF algorithm. Compared with the similar mainstream algorithms, the tracking effect of the algorithm is improved, which proves its effectiveness.

Keywords: unmanned aerial vehicle; long-term tracking; correlation filter; heavy detector; fast scale filtering; high confidence

相关滤波目标跟踪算法在无人机测控领域作为一种基础性方法,具有广泛的应用前景^[1-4]。Bolme等^[5]在单目标跟踪中引入傅里叶域快速处理方法,提出了最小误差平方和滤波器,跟踪速度有所提升。Henriques等^[6]提出了带内核的循环检测(CSK)跟踪算法,基于循环矩阵理论来获得循环唯一的近似密集采样,解决稀疏采样造成的样本冗余问题。Henriques等^[7]在CSK基础上提出了基于核相关滤波(KCF)算法,对多通道特征进行有效组合。Danelljan等^[8]提出在CSK基础上融入图像颜色名特称描述行为,通过主成分分析法降低特征维度,提升跟踪整体性能。为适应目标的尺度变化, Danelljan等^[9]提出了判别尺度空间跟踪器(DSST),通过33个尺度,使用位置滤波器确定目标位置,以目标位置为中心确定尺度大小。Danelljan等^[10]在DSST基础上,提出了快速判别尺度空间跟踪器(FDSST),将33个尺度减少至17个,显著提高了跟踪性能。Galoogahi等^[11]提出了背景感知相关滤波器来缓解边界效应,并使用交替方向乘子法(ADMM)优化求解滤波器。2018年,Li等^[12]提出了时空正则化相关滤波器,缓解滤波器退化问题。2020年,Li等^[13]又提出了一种在线自动自适应学习时空正则化项的新方法AutoTrack。目前,长期跟踪成为跟踪领域的研究新的热点和难点。长期跟踪序列帧数较多,需要对目标进行长时间连续跟踪,期间可能会因为部分遮挡、完全遮挡等场景导致跟踪丢失。Kalal等^[14]提出了一种“跟踪-学习-检测”框架,使用前向-反向光流匹配来追踪目标,并利用集成学习进行目标重检测,同时还引入了P-N(positive-negative)学习模块来帮助实现跟踪器和检测器完成在线更新。Ma等^[15]提出了长期相关性跟踪(LCT)算法,在相关滤波器基础上增加了重新检测机制,通过在线随机森林分类器,在目标遮挡或者快速运动导致丢失后可以成功找回目标,但由于检

测器比较单一,跟踪器鲁棒性较差,导致跟踪效果不佳。2021年,Yan等^[16]提出了一种新颖、灵活和准确的细化模块Alpha-Refine,该模块可以在低延迟的情况下,显著提升基础跟踪器边界框的估计质量。在无人机长期跟踪中,目标可能会遭受外观严重变化和其他挑战,需要提升目标丢失后重新检测的能力;同时,面对移出视野、遮挡等复杂场景,长期跟踪器需要进一步提升鲁棒性。

因此,本文提出了一种基于局部-全局区域重检测的长期跟踪算法,设计了跟踪状态判断模块和局部-全局区域重检测机制,能够稳定恢复跟踪丢失的目标,双重检测机制下鲁棒性较高。

1 基础跟踪器设计

以相关滤波算法为基准,设计自适应空间正则化方法,惩罚对背景的学习,解决循环采样导致的边界效应问题;设计基于高置信度样本的时间正则化方法和优化更新策略,在限制滤波器的学习速度的同时,使用高置信度样本约束滤波器的更新方向(其中高置信度样本由其队列加权得到),解决滤波器退化问题;设计快速尺度滤波器,在保证速度的前提下,确定目标的尺度,解决无人机航拍过程面临目标尺度变化的问题。将方向梯度直方图特征称为FHOG特征^[17],颜色名特征称为CN特征^[18],基础跟踪器设计框图如图1所示。

1.1 高置信度样本队列

目标定位通过相关的滤波器和图像操作来实现。滤波器在不断更新中存在退化问题,本文通过构建高置信度样本队列用于后续缓解滤波器退化,提高滤波器的鲁棒性。引入跟踪状态判断因子

$$q_t = P_{\text{SR}}(\mathbf{R}_t) \max(\mathbf{R}_t) \quad (1)$$

式中: P_{SR} 为峰值旁瓣比; $\mathbf{R}_t$ 为第 t 帧的响应图,将峰值旁瓣比与响应最大值相乘,可以放大其反应跟踪效果的能力。设 $\bar{q}_t$ 为最近 N_q 连续帧上计算得到的

均值, τ_q 为阈值, 若满足

$$\bar{q}_t / q_t > \tau_q \quad (2)$$

则认为跟踪不确定。

通过式(2)来判断当前帧的跟踪状态, 将不满足的帧视为高置信度样本, 将其对应的相关滤波器加入到队列中, 构成高置信度样本队列, 无人机长期跟踪过程中, 目标在时间上具有较大变化, 考虑到误差影响以及算法响应速度, 设置其最大容量为 15, 队列通过判断当前帧是否满足式(2)进行更新判断。

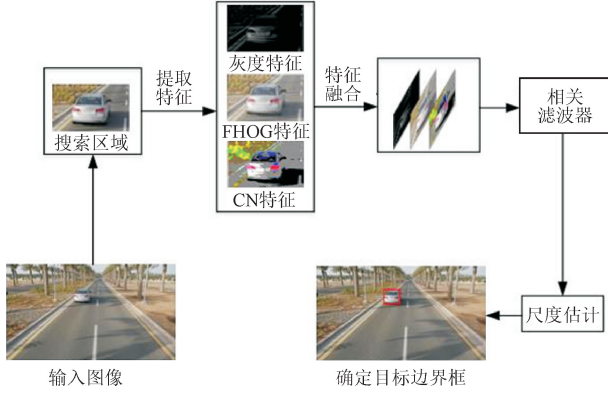


图1 基于自适应时空正则化的相关滤波跟踪算法框图

Fig.1 Block diagram of correlation filter tracking algorithm based on adaptive spatiotemporal regularization

1.2 自适应空间正则化

设计自适应空间正则化^[19]可以使模型更好地适应目标外观的变化, 增加模型的鲁棒性。基于自适应空间正则化的目标函数如下

$$E(\mathbf{H}_t, \mathbf{w}) = \frac{1}{2} \left\| \mathbf{y} - \sum_{k=1}^K \mathbf{h}_{t,k} \otimes \boldsymbol{\phi}_{t,k} \right\|_2^2 + \frac{\lambda_1}{2} \sum_{k=1}^K \left\| \mathbf{w} \odot \mathbf{h}_{t,k} \right\|_2^2 + \frac{\lambda_2}{2} \sum_{k=1}^K \left\| \mathbf{w} - \mathbf{w}_r \right\|_2^2 \quad (3)$$

式中: K 为特征维数; l 为序列长度; $\mathbf{h}_{t,k} \in \mathbf{R}^{l \times 1}$ 代表在第 t 帧 k 通道中的滤波器的滤波函数; $\otimes$ 表示循环相关运算符; $\odot$ 表示矩阵的哈达玛积; $\mathbf{H}_t = [\mathbf{h}_{t,1}, \mathbf{h}_{t,2}, \dots, \mathbf{h}_{t,K}]$; $\mathbf{y} \in \mathbf{R}^{l \times 1}$ 是高斯响应输出; $\boldsymbol{\phi}_{t,k} \in \mathbf{R}^{l \times 1} (k = 1, 2, 3, \dots, K)$ 表示第 t 帧提取的特征; $\mathbf{w}$ 为空间正则化权重; $\mathbf{w}_r$ 为参考权重; λ_1 和 λ_2 是正则化参数; $\frac{1}{2} \left\| \mathbf{y} - \sum_{k=1}^K \mathbf{h}_{t,k} \otimes \boldsymbol{\phi}_{t,k} \right\|_2^2$ 为基础滤波器; $\frac{\lambda_1}{2} \sum_{k=1}^K \left\| \mathbf{w} \odot \mathbf{h}_{t,k} \right\|_2^2$ 为自适应空间正则项, 可以有效适应跟踪过程中的目标的特征外观变化情况, 减少非目标区域的响应; $\frac{\lambda_2}{2} \sum_{k=1}^K \left\| \mathbf{w} - \mathbf{w}_r \right\|_2^2$ 使空间正则化权重尽可能接近于参考权重。参考权重更新方式如下

$$\mathbf{w}_{r,k} = (1 - l_r) \mathbf{w}_{hca,k} + l_r \mathbf{w}_t \quad (4)$$

式中: $\mathbf{w}$ 初始化为负高斯形状的空间权重矩阵; l_r 为学习率; $\mathbf{w}_{hca,k}$ 为最近的高置信度样本帧对应的空间参考权重, 高置信度样本帧通过式(2)来确定。

1.3 基于高置信度样本的自适应时空正则化

在无人机航拍过程中跟踪器面临尺度变化、快速运动、低分辨率等复杂场景, 导致滤波器退化, 跟踪效果降低。针对无人机航拍中复杂的场景变化, 通过设计基于高置信度样本的时间正则化方法, 限制滤波器学习速度, 并利用高置信度样本修正滤波器更新方向, 提高复杂场景下模型的性能和鲁棒性。定义参考值 $\bar{\theta}$, 根据相邻两帧之间的响应变化优化固定的超参数。参考值计算如下

$$\bar{\theta} = \omega \ln(-v \left\| \mathbf{R}_t[\Psi_\Delta] - \mathbf{R}_{t-1} \right\|_2^2 + 2), \quad \left\| \mathbf{R}_t[\Psi_\Delta] - \mathbf{R}_{t-1} \right\|_2^2 \leq \phi \quad (5)$$

式中: ω 和 v 为超参数; $[\Psi_\Delta]$ 是使得两个响应图 $\mathbf{R}_t$ 和 $\mathbf{R}_{t-1}$ 的峰值重合的移位算子; $\mathbf{R}_t[\Psi_\Delta]$ 为当前样本生成的响应图。当 $\left\| \mathbf{R}_t[\Psi_\Delta] - \mathbf{R}_{t-1} \right\|_2^2$ 高于阈值 ϕ 时, 表示前后两帧响应图畸变较大, 停止滤波器学习; 当 $\left\| \mathbf{R}_t[\Psi_\Delta] - \mathbf{R}_{t-1} \right\|_2^2$ 低于阈值 ϕ 时, 前后两帧响应图相似。参考值 $\bar{\theta}$ 较大, 使得时间正则项对模型的约束能力较强, 因此应进行更新滤波器的更新, 避免模型发生突变。参考 UAV 数据集集中的评价指标, 选取了多个满足指标的帧, 以这些帧与上一帧的 $\left\| \mathbf{R}_t[\Psi_\Delta] - \mathbf{R}_{t-1} \right\|_2^2$ 均值为参考, 可将 ϕ 设置为 3 000。

结合 1.2 节设计的基于自适应空间正则化的目标函数, 融入自适应时间正则化项, 得到基于自适应时空正则化的目标函数如下

$$E(\mathbf{H}_t, \mathbf{w}, \theta_t) = \frac{1}{2} \left\| \mathbf{y} - \sum_{k=1}^K \mathbf{h}_{t,k} \otimes \boldsymbol{\phi}_{t,k} \right\|_2^2 + \frac{\lambda_1}{2} \sum_{k=1}^K \left\| \mathbf{w} \odot \mathbf{h}_{t,k} \right\|_2^2 + \frac{1}{2} \left\| \theta_t - \bar{\theta} \right\|_2^2 + \frac{\lambda_2}{2} \sum_{k=1}^K \left\| \mathbf{w} - \mathbf{w}_r \right\|_2^2 + \frac{\lambda_3}{2} \sum_{k=1}^K \left\| \mathbf{h}_{t,k} - \mathbf{h}_{hca,k} \right\|_2^2 + \frac{\theta_t}{2} \sum_{k=1}^K \left\| \mathbf{h}_{t,k} - \mathbf{h}_{t-1,k} \right\|_2^2 \quad (6)$$

式中: $\bar{\theta}$ 表示参考值; θ_t 和 λ_3 为时间正则化参数; $\mathbf{h}_{hca,k}$ 代表历史高置信度样本加权得到的滤波器, 通过对高置信度样本队列中相关滤波器加权求和得到。在时间正则化中通过 $\frac{\theta_t}{2} \sum_{k=1}^K \left\| \mathbf{h}_{t,k} - \mathbf{h}_{t-1,k} \right\|_2^2$ 和 $\frac{\lambda_3}{2} \sum_{k=1}^K \left\| \mathbf{h}_{t,k} - \mathbf{h}_{hca,k} \right\|_2^2$ 分别通过惩罚当前滤波器与前一帧滤波器之间的差异和惩罚当前滤波器与高置信度样本加权得到的滤波器来限制滤波器的学习速度, 避免对当前的数据进行过度拟合, 使得滤波器更

谨慎的对参数进行更新,并通过高置信度样本修正滤波器的更新方向。 $\frac{1}{2} \|\theta_t - \bar{\theta}\|_2^2$ 表示让时间正则化参数 θ_t 尽可能接近参考值 $\bar{\theta}$ 。

1.4 交替方向乘子法求解

1.4.1 滤波器的函数和时间正则化参数求解

利用交替方向乘子法(ADMM)^[20]对时间正则化参数和滤波器函数进行求解,引入辅助变量 $\mathbf{g}_t = \mathbf{h}_t (\mathbf{G} = [\mathbf{g}_{t,1}, \mathbf{g}_{t,2}, \mathbf{g}_{t,3}, \dots, \mathbf{g}_{t,k}])$,将目标函数写为增广拉格朗日形式

$$L(\mathbf{w}, \mathbf{G}_t, \mathbf{M}_t, \theta_t) = \frac{1}{2} \|\mathbf{y} - \sum_{k=1}^K \mathbf{h}_{t,k} \otimes \boldsymbol{\phi}_{t,k}\|_2^2 + \frac{\lambda_1}{2} \sum_{k=1}^K \|\mathbf{w} \odot \mathbf{g}_{t,k}\|_2^2 + \frac{\lambda_2}{2} \|\mathbf{w} - \mathbf{w}_r\|_2^2 + \frac{\lambda_3}{2} \sum_{k=1}^K \|\mathbf{h}_{t,k} - \mathbf{h}_{\text{hcaa},k}\|_2^2 + \frac{\theta_t}{2} \sum_{k=1}^K \|\mathbf{h}_{t,k} - \mathbf{h}_{t-1,k}\|_2^2 + \frac{1}{2} \|\theta_t - \bar{\theta}\|_2^2 + \frac{\gamma}{2} \sum_{k=1}^K \|\mathbf{h}_{t,k} - \mathbf{g}_{t,k}\|_2^2 + \sum_{k=1}^K (\mathbf{h}_{t,k} - \mathbf{g}_{t,k})^T \mathbf{m}_{t,k} \quad (7)$$

式中: $\hat{\mathbf{M}}_t = [\hat{\mathbf{m}}_1, \hat{\mathbf{m}}_2, \dots, \hat{\mathbf{m}}_K] \in \mathbf{R}^{L \times K}$ 为拉格朗日乘子; γ 为步进参数,令 $\mathbf{v}_{t,k} = \mathbf{m}_{t,k} / \gamma$, $\mathbf{v}_{t,k} \in \mathbf{V}_{t,k} = [\mathbf{v}_{t,1}, \mathbf{v}_{t,2}, \dots, \mathbf{v}_{t,k}]$,则式(7)可简化为

$$L(\mathbf{w}, \mathbf{G}_t, \mathbf{V}_t, \theta_t) = \frac{1}{2} \|\mathbf{y} - \sum_{k=1}^K \mathbf{h}_{t,k} \otimes \boldsymbol{\phi}_{t,k}\|_2^2 + \frac{\lambda_1}{2} \sum_{k=1}^K \|\mathbf{w} \odot \mathbf{g}_{t,k}\|_2^2 + \frac{\lambda_2}{2} \|\mathbf{w} - \mathbf{w}_r\|_2^2 + \frac{\lambda_3}{2} \sum_{k=1}^K \|\mathbf{h}_{t,k} - \mathbf{h}_{\text{hcaa},k}\|_2^2 + \frac{\theta_t}{2} \sum_{k=1}^K \|\mathbf{h}_{t,k} - \mathbf{h}_{t-1,k}\|_2^2 + \frac{1}{2} \|\theta_t - \bar{\theta}\|_2^2 + \frac{\gamma}{2} \sum_{k=1}^K \|\mathbf{h}_{t,k} - \mathbf{g}_{t,k} + \mathbf{v}_{t,k}\|_2^2 \quad (8)$$

子问题 $\mathbf{h}$ 、 $\mathbf{g}$ 和 θ_t 通过 ADMM 迭代求解

$$\begin{cases} \mathbf{h}^{(i+1)} = \arg \min_{\mathbf{h}} (\|\mathbf{y} - \sum_{k=1}^K \mathbf{h}_{t,k} \otimes \boldsymbol{\phi}_{t,k}\|_2^2 + \lambda_3 \sum_{k=1}^K \|\mathbf{h}_{t,k} - \mathbf{h}_{\text{hcaa},k}\|_2^2 + \theta_t \sum_{k=1}^K \|\mathbf{h}_{t,k} - \mathbf{h}_{t-1,k}\|_2^2 + \gamma \sum_{k=1}^K \|\mathbf{h}_{t,k} - \mathbf{g}_{t,k} + \mathbf{v}_{t,k}\|_2^2) \\ \mathbf{g}^{(i+1)} = \arg \min_{\mathbf{g}} (\sum_{k=1}^K \|\mathbf{w} \odot \mathbf{g}_{t,k}\|_2^2 + \gamma \sum_{k=1}^K \|\mathbf{h}_{t,k} - \mathbf{g}_{t,k} + \mathbf{v}_{t,k}\|_2^2) \\ \theta_t^{(i+1)} = \arg \min_{\theta_t} (\theta_t \sum_{k=1}^K \|\mathbf{h}_{t,k} - \mathbf{h}_{t-1,k}\|_2^2 + \|\theta_t - \bar{\theta}\|_2^2) \\ \mathbf{v}^{(i+1)} = \mathbf{v}^{(i)} + \mathbf{h}^{(i+1)} - \mathbf{g}^{(i+1)} \end{cases} \quad (9)$$

对子问题进行求解,具体过程如下。

(1)子问题 $\mathbf{h}$ 。根据帕塞瓦尔定理,通过式(9)中的第一项从时域转换到频域来降低算法复杂度,第一个子问题转换如下

$$\arg \min_{\mathbf{h}} (\|\hat{\mathbf{y}} - \sum_{k=1}^K \hat{\mathbf{h}}_{t,k} \odot \hat{\boldsymbol{\phi}}_{t,k}\|_2^2 + \lambda_3 \sum_{k=1}^K \|\hat{\mathbf{h}}_{t,k} - \hat{\mathbf{h}}_{\text{hcaa},k}\|_2^2 + \theta_t \sum_{k=1}^K \|\hat{\mathbf{h}}_{t,k} - \hat{\mathbf{h}}_{t-1,k}\|_2^2 + \gamma \sum_{k=1}^K \|\hat{\mathbf{h}}_{t,k} - \hat{\mathbf{g}}_{t,k} + \hat{\mathbf{v}}_{t,k}\|_2^2) \quad (10)$$

式中:符号 $\wedge$ 表示变量的离散傅里叶变换。 $\hat{\mathbf{y}}$ 的元素相关于 $\hat{\mathbf{h}}_{t,k}$ 和 $\hat{\boldsymbol{\phi}}_{t,k}$ 上 K 通道的对应元素,为提高运算效率,式(10)可以化简为

$$\arg \min_{\mathbf{G}_i(\hat{\mathbf{h}}_t)} (\|\hat{\mathbf{y}}_i - \mathbf{G}_i(\hat{\boldsymbol{\phi}})_L \mathbf{G}_i(\hat{\mathbf{h}}_t)\|_2^2 + \lambda_3 \|\mathbf{G}_i(\hat{\mathbf{h}}_t) - \mathbf{G}_i(\hat{\mathbf{h}}_{\text{hcaa}})\|_2^2 + \theta_t \|\mathbf{G}_i(\hat{\mathbf{h}}_t) - \mathbf{G}_i(\hat{\mathbf{h}}_{t-1})\|_2^2 + \gamma \|\mathbf{G}_i(\hat{\mathbf{h}}_t) - \mathbf{G}_i(\hat{\mathbf{g}}_t) + \mathbf{G}_i(\hat{\mathbf{v}}_t)\|_2^2) \quad (11)$$

式中: $\mathbf{G}_i(\hat{\boldsymbol{\phi}})_L \in \mathbf{C}^{K \times 1}$ 表示包含像素 $i(i=1, 2, \dots, L)$ 上 $\hat{\boldsymbol{\phi}}$ 的所有 K 个通道的值的向量。

对式(11)进行求导,整理可得 $\mathbf{G}_i(\hat{\mathbf{h}}_t)$ 的封闭解为

$$\mathbf{G}_i(\hat{\mathbf{h}}_t) = [\mathbf{G}_i(\hat{\boldsymbol{\phi}})_L \mathbf{G}_i(\hat{\boldsymbol{\phi}})_L + (\theta_t + \gamma + \lambda_3) \mathbf{I}]^{-1} \cdot [\mathbf{G}_i(\hat{\boldsymbol{\phi}})_L \hat{\mathbf{y}}_i + \gamma \mathbf{G}_i(\hat{\mathbf{g}}_t) - \gamma \mathbf{G}_i(\hat{\mathbf{v}}_t) + \theta_t \mathbf{G}_i(\hat{\mathbf{h}}_{t-1}) + \lambda_3 \mathbf{G}_i(\hat{\mathbf{h}}_{\text{hcaa}})] \quad (12)$$

由于式(12)存在大量的矩阵求逆步骤,会增加算法时长,因此根据谢尔曼-莫里森公式可得

$$\mathbf{G}_i(\hat{\mathbf{h}}_t) = \frac{1}{(\theta_t + \gamma + \lambda_3)} \left(\mathbf{I} - \frac{\mathbf{G}_i(\hat{\boldsymbol{\phi}})_L \mathbf{G}_i(\hat{\boldsymbol{\phi}})_L}{\theta_t + \gamma + \lambda_3 + \mathbf{G}_i(\hat{\boldsymbol{\phi}})_L \mathbf{G}_i(\hat{\boldsymbol{\phi}})_L} \right) \cdot [\mathbf{G}_i(\hat{\boldsymbol{\phi}})_L \hat{\mathbf{y}}_i + \gamma \mathbf{G}_i(\hat{\mathbf{g}}_t) - \gamma \mathbf{G}_i(\hat{\mathbf{v}}_t) + \theta_t \mathbf{G}_i(\hat{\mathbf{h}}_{t-1}) + \lambda_3 \mathbf{G}_i(\hat{\mathbf{h}}_{\text{hcaa}})] \quad (13)$$

(2)子问题 $\mathbf{g}$ 。式(9)的第二个子问题在时域进行求解,可得

$$\mathbf{g} = (\mathbf{W}^T \mathbf{W} + \gamma \mathbf{I})^{-1} \gamma (\mathbf{h}_t + \mathbf{v}_t) \quad (14)$$

式中: $\mathbf{W} = \text{diag}(\mathbf{w}) \in \mathbf{R}^{T \times T}$ 表示 $T \times T$ 维对角矩阵。

(3)子问题 θ_t 。 θ_t 的求解公式如下

$$\theta_t = \bar{\theta} - \frac{\sum_{k=1}^K \|\hat{\mathbf{h}}_{t,k} - \hat{\mathbf{h}}_{t-1,k}\|_2^2}{2} \quad (15)$$

更新步进参数 γ

$$\gamma^{(i+1)} = \min(\gamma^{(\max)}, \alpha \gamma^{(i)}) \quad (16)$$

式中: α 为时间正则化固定参数。

1.4.2 空间正则化权重求解

使用 ADMM 对空间正则化权重求解,由式(8)可得

$$\arg \min_{\mathbf{w}} (\lambda_1 \sum_{k=1}^K \|\mathbf{w} \otimes \mathbf{g}_{t,k}\|_2^2 + \lambda_2 \|\mathbf{w} - \mathbf{w}_r\|_2^2) \quad (17)$$

引入辅助变量 $\mathbf{q}=\mathbf{w}$, 式(17)的增广拉格朗日形式如下

$$L(\mathbf{w}, \mathbf{q}, \mathbf{s}) = \frac{\lambda_1}{2} \sum_{k=1}^K \|\mathbf{w} \odot \mathbf{g}_{t,k}\|_2^2 + \frac{\lambda_2}{2} \|\mathbf{q} - \mathbf{w}_r\|_2^2 + \frac{\xi}{2} \|\mathbf{w} - \mathbf{q} + \mathbf{s}\|_2^2 \quad (18)$$

式中: ξ 为步进参数; $\mathbf{s}$ 为拉格朗日乘子。将式(18)分为 3 个子问题, 对每个子问题进行求解。

(1) 子问题 $\mathbf{w}$ 。由式(14)可求得

$$\mathbf{w} = (\mathbf{q} - \mathbf{s}) \left(\mathbf{I} + \frac{\lambda_1}{\xi} \mathbf{G}^T \mathbf{G} \right)^{-1} \quad (19)$$

式中: $\mathbf{G}=\text{diag}(\mathbf{g})$ 为对角矩阵。

(2) 子问题 $\mathbf{q}$ 。求解形式如下

$$\mathbf{q} = \frac{\lambda_2 \mathbf{w}_r + \xi(\mathbf{w} + \mathbf{s})}{\lambda_2 + \xi} \quad (20)$$

(3) 子问题 $\mathbf{s}$ 。在求得 $\mathbf{w}$ 和 $\mathbf{q}$ 的值后, 得到

$$\mathbf{s}^{(i+1)} = \mathbf{s}^{(i)} + \mathbf{w}^{(i+1)} - \mathbf{q}^{(i+1)} \quad (21)$$

更新步进参数

$$\xi^{(i+1)} = \min(\xi^{(\max)}, \mu \xi^{(i)}) \quad (22)$$

式中: $\xi^{(\max)}$ 为步进参数 ξ 的最大值; μ 为空间正则化固定参数。

最后, 确定第 t 帧目标位置

$$\mathbf{R}_t = \text{IFT} \left(\sum_{k=1}^K (\hat{\mathbf{z}}_{t,k} \odot \hat{\mathbf{h}}_{t-1,k}) \right) \quad (23)$$

式中: IFT 为傅里叶变换逆变换; $\hat{\mathbf{z}}_{t,k}$ 表示在第 t 帧提取特征的傅里叶形式。

1.5 优化更新策略

无人机对航拍目标进行跟踪时, 跟踪目标会经历尺度变化、长宽比变化、遮挡等复杂场景, 因此逐帧的滤波器更新会导致模型发生漂移, 滤波器退化。这里通过评价跟踪结果自适应更新滤波器。传统按帧逐次更新模型的策略耗时长, 且目标一旦出现检测不准确、严重遮挡或在当前帧中完全缺失, 易导致模型发生漂移。因此, 本文基于平均峰值相关能量 (average peak correlation energy, APCE) 提出优化更新策略, 通过跟踪结果来确定模型的更新情况, 实现目标跟踪稳定时滤波器的更新。令 E_{APCE} 表示响应图的稳定程度, 计算式为

$$E_{\text{APCE}} = \frac{|F_{\max} - F_{\min}|^2}{\text{mean}(\sum_{w,h} (F_{w,h} - F_{\min})^2)} \quad (24)$$

式中: $F_{\max}$ 、 $F_{\min}$ 分别为响应值的最大值、最小值; $F_{w,h}$ 为 w 行 h 列的响应值; mean 为求取均值函数。

$F_{\max}$ 的样本队列为 $\mathbf{q}_F$, 样本队列的均值为 $\bar{q}_F$; APCE 的样本队列为 $\mathbf{q}_{\text{APCE}}$, 样本队列的均值为 $\bar{q}_{\text{APCE}}$; a_1 、 a_2 、 a_3 为判定跟踪结果的比例判别因

子。队列容量均为 50, 队列满足先进先出, 当前帧不满足式(2)时, 对队列进行更新。当 $F_{\max} > a_1 \bar{q}_F$ 、 $E_{\text{APCE}} > a_2 \bar{q}_{\text{APCE}}$ 时, 表明当前跟踪结果具有高置信度, 以学习率 η_1 更新模型; 当 $a_3 \bar{q}_F < F_{\max} < a_1 \bar{q}_F$ 、 $E_{\text{APCE}} > a_2 \bar{q}_{\text{APCE}}$ 时, 表明当前跟踪效果一般, 以学习率 η_2 更新模型。否则, 模型不进行更新。滤波器的更新方式如下

$$\mathbf{h}_{t+1}^{\text{ST}} = (1 - \eta) \mathbf{h}_t^{\text{ST}} + \eta \mathbf{h}_t^{\text{ST}*} \quad (25)$$

式中: $\mathbf{h}_{t+1}^{\text{ST}}$ 、 $\mathbf{h}_t^{\text{ST}}$ 和 $\mathbf{h}_t^{\text{ST}*}$ 分别为更新后的短期位置相关滤波器、第 t 帧用来定位的短期位置相关滤波器和求出的短期位置相关滤波器; η 为学习率, 由 $F_{\max}$ 和 E_{APCE} 确定。

1.6 快速尺度滤波器的设计

无人机在航拍过程中目标面临复杂的场景变化。为适应目标在跟踪过程中的尺度变化, 在位置滤波器确定目标所在位置之后, 设计了一种自适应更新的快速尺度估计滤波器。设 $M \times N$ 为当前帧目标尺寸大小, 位置滤波器得到当前帧目标的位置后, 以其为中心, 提取尺寸为 $anM \times anN$ 的 s 个训练样本, $a=1.02$ 为尺度系数, $n \in \{(-s-1)/2, \dots, [(s-1)/2]\}$, 并将样本整合为相同尺寸。

设计的尺度滤波器的参数 $s=33$, 尺度范围 $a'' \in \{1.02^{-16}, \dots, 1.02^{16}\}$ 。在无人机场景下, 为了满足实时性, 对尺度滤波器的尺度范围进行改进, 尺度范围用 S_{range} 表示。其中: $S_{\text{range}} = \{0.70, 0.75, 0.80, 0.85, 0.90, 0.95, 1.0, 1.05, 1.1, 1.15, 1.20, 1.25, 1.30, 1.35, 1.40\}$, 共 15 个尺度, 尺度范围在 $[0.70, 1.40]$ 之间, 由此可以保证相邻帧之间的尺度变化都可以被本文设计的尺度滤波器精确估计。

此外, 针对无人机航拍过程中面临复杂的场景变化, 尺度模型需要适应尺度变化, 设计的尺度滤波器遵循优化后的更新策略, 当 $F_{\max} > a_1 \bar{q}_F$ 、 $E_{\text{APCE}} > a_2 \bar{q}_{\text{APCE}}$ 时, 表明当前跟踪结果具有高置信度, 以学习率 η_3 更新尺度模型。当 $a_3 \bar{q}_F < F_{\max} < a_1 \bar{q}_F$ 、 $E_{\text{APCE}} > a_2 \bar{q}_{\text{APCE}}$ 时, 表明当前跟踪效果一般, 以学习率 η_4 更新尺度模型。否则尺度模型不更新。

1.7 算法流程

本文算法流程如图 2 所示。算法的主要步骤如下:

- (1) 输入当前序列第一帧图像和目标的真实边界框, 求出初始的短期跟踪器, 完成模型初始化;
- (2) 从第二帧开始在目标搜索区域提取并融合多种特征, 通过位置相关滤波器得到响应图, 根据响应图得到目标中心位置;

(3)使用快速尺度滤波器对目标尺度大小进行估计,得到最佳尺度;

(4)根据优化更新策略,对位置滤波器和快速尺度滤波器组成的短期跟踪器进行自适应模型更新,同时对高置信度样本队列进行自适应更新;

(5)判断当前序列是否跟踪结束,是则停止跟踪,未结束则读取当前序列下一帧的图像,重复上述操作,直至跟踪结束。

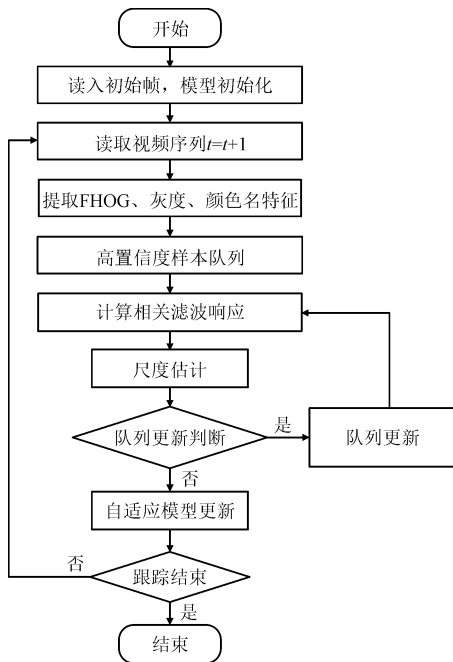


图2 本文算法流程

Fig. 2 Flow chart of proposed algorithm

2 局部-全局区域重检测机制

2.1 长期跟踪算法

无人机航拍目标长期跟踪中需要跟踪器在面临遮挡、移出视野等复杂场景导致跟踪失败后具备对目标进行恢复跟踪的能力,本文在提出的基础跟踪器的基础上,加入跟踪状态判断模块和基于局部-全局区域重检测机制的重检测器,提出了一种基于局部-全局区域重检测机制的无人机长期跟踪算法。局部-全局区域重检测机制流程如图3所示。通过基于跟踪状态判断因子和颜色直方图的跟踪状态判断模块判断基础跟踪器是否跟踪失败,跟踪失败的情况下启动重检测器。长期跟踪需要在短期跟踪失败的情况下,搜索目标可能的存在区域,并在目标重新出现的时候重新定位目标,本文算法通过重检测器完成目标搜索和重新定位。重检测器由局部区域重检测和全局区域重检测组成,其中局部区域重检

测通过局部区域搜索模块在目标的局部区域范围进行搜索实现局部区域重检测,并通过跟踪状态判断模块判断其是否跟踪成功,跟踪失败的情况下启用全局区域重检测器,通过全局区域重检测完成对目标跟踪状态的恢复。

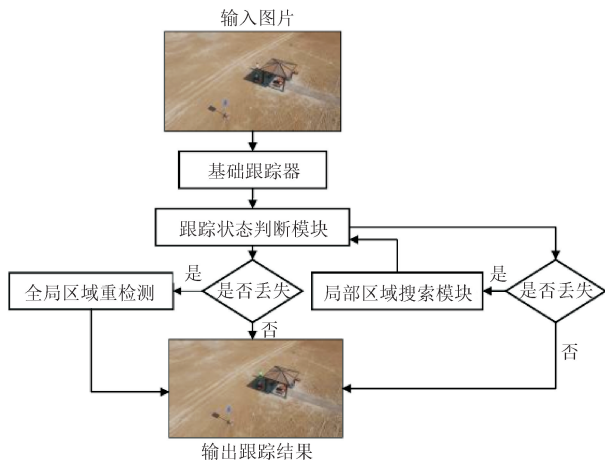


图3 局部-全局区域重检测机制流程图

Fig. 3 Flow chart of local-global region re-detection mechanism

2.2 跟踪状态判断

2.2.1 基于跟踪状态判断因子的跟踪状态判断

在无人机长期航拍过程中,由于会面临部分遮挡、完全遮挡、移出视野等场景导致基础跟踪器跟踪丢失,如果不对跟踪状态进行判断,继续使用没有重检测能力的跟踪器,会导致持续的跟踪失败。因此在无人机长期跟踪中,需要对跟踪状态进行判断,使得当跟踪失败的时候启用重检测器对目标进行找回,本文根据式(1)和式(2)的跟踪状态判断因子 q_t ,结合式(6)中前后两帧响应图的变化情况,定义跟踪状态的判断指标为 T_t ,公式为

$$T_t = \begin{cases} 1, & \text{其他} \\ 0, & \bar{q}_t/q_t > \tau_q; \| \mathbf{R}_t[\Psi_\Delta] - \mathbf{R}_{t-1} \|_2^2 > \phi \end{cases} \quad (26)$$

式中: $T_t=1$ 表示跟踪稳定,可以继续使用基础跟踪器对目标进行跟踪; $T_t=0$ 表示跟踪失败,需要启用重检测器对目标进行恢复,即 $\| \mathbf{R}_t[\Psi_\Delta] - \mathbf{R}_{t-1} \|_2^2$ 高于阈值 ϕ ,当前帧与上一帧的响应变化剧烈;当 $\bar{q}_t/q_t > \tau_q$ 时,认为目标跟踪失败。

2.2.2 基于颜色直方图的跟踪状态判断

无人机航拍目标跟踪中面临尺度变化、长宽比变化等挑战,而颜色信息对于变形和旋转非常鲁棒,通过基于颜色直方图的贝叶斯分类器可以区分输入图像中的目标像素与周围背景像素。本文通过联合跟踪状态判断因子和颜色直方图,设计目标跟踪状态判断指标 T_t ,基于颜色直方图的跟踪状态判断方

法介绍如下。

基于颜色直方图计算汽车目标的似然概率,结果如图 4 所示。通过在输入图像上使用基于颜色直方图的贝叶斯分类器将属于目标的像素与周围背景像素区分开。

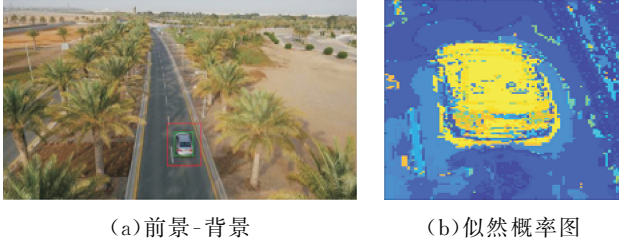


图 4 目标似然概率计算结果图

Fig. 4 Calculate the target likelihood probability graph

像素点 X 属于目标的概率如下

$$P(X \in \mathbf{O} | \mathbf{O}, \mathbf{B}, b_X) \approx \frac{P(b_X | X \in \mathbf{O})P(X \in \mathbf{O})}{\sum_{\mathbf{O} \in (\mathbf{O}, \mathbf{B})} P(b_X | X \in \mathbf{O})P(X \in \mathbf{O})} \quad (27)$$

式中: $P(b_X | X \in \mathbf{O}) \approx \mathbf{H}_\mathbf{O}(b_X) / |\mathbf{O}|$, $P(b_X | X \in \mathbf{B}) \approx \mathbf{H}_\mathbf{B}(b_X) / |\mathbf{B}|$, b_X 为像素点 X 的值,位于直方图的第 b 个区间中, $\mathbf{H}_\mathbf{O}$ 和 $\mathbf{H}_\mathbf{B}$ 分别为目标前景及其背景颜色直方图。在图 4 中, $\mathbf{O}$ 为绿色框内区域, $\mathbf{B}$ 为红色框内除去绿色框区域的部分。先验概率由 $P(X \in \mathbf{O}) = |\mathbf{O}| / (|\mathbf{O}| + |\mathbf{B}|)$ 得到,式(27)可以化简为

$$P(X \in \mathbf{O} | \mathbf{O}, \mathbf{B}, b_X) \approx \frac{\mathbf{H}_\mathbf{O}(b_X)}{\mathbf{H}_\mathbf{O}(b_X) + \mathbf{H}_\mathbf{B}(b_X)} \quad (28)$$

为防止第 t 帧目标前景直方图和背景直方图更新时只受到当前帧影响,增加了对上一帧的学习,使公式更具鲁棒性,更新后的目标前景直方图和背景直方图为

$$\mathbf{H}_{\mathbf{O},t} = (1 - \eta_{\text{hist}}) \mathbf{H}_{\mathbf{O},t-1} + \eta_{\text{hist}} \tilde{\mathbf{H}}_{\mathbf{O},t} \quad (29)$$

$$\mathbf{H}_{\mathbf{B},t} = (1 - \eta_{\text{hist}}) \mathbf{H}_{\mathbf{B},t-1} + \eta_{\text{hist}} \tilde{\mathbf{H}}_{\mathbf{B},t} \quad (30)$$

式中: $\tilde{\mathbf{H}}_{\mathbf{O},t}$ 和 $\tilde{\mathbf{H}}_{\mathbf{B},t}$ 表示第 t 帧计算得到的目标前景直方图和背景直方图; η_{hist} 为学习率。

本文基于颜色直方图计算颜色分数,根据颜色分数来确定跟踪可靠性。根据上述关于颜色直方图的阐述,可以得出像素 u 的颜色直方图权重向量 $\mathbf{B}_t(u)$ 如下

$$\mathbf{B}_t(u) = \frac{\mathbf{H}_\mathbf{O}(b_u)}{\mathbf{H}_\mathbf{O}(b_u) + \mathbf{H}_\mathbf{B}(b_u) + \lambda} \quad (31)$$

式中: λ 为一个正常数,用于调整权重占比。对于第 t 帧图像,从中提取目标区域 z_t ,对每个像素计算响

应得到分图为 $\sum_{v \in \mathbf{B}_u} \mathbf{B}_t(v) \boldsymbol{\varphi}_z(v)$, 并定义颜色分数为

$$S_t = \max_{u \in z^t} \sum_{v \in \mathbf{B}_u} \mathbf{B}_t(v) \boldsymbol{\varphi}_z(v) \quad (32)$$

式中: $\boldsymbol{\varphi}_z(v)$ 表示在图像块 z 中像素 v 的颜色特征; $\mathbf{B}_u$ 表示将目标的边界框移动到 u 的中心而获得的边界框。

本文构造了一个可靠颜色分数样本队列 $\mathbf{Q}_c$, 通过将检测结果可靠帧的颜色分数值加入到样本队列 $\mathbf{Q}_c$ 中, $\mathbf{Q}_c$ 满足先进先出, $\mathbf{Q}_c$ 最大容量为 30, 并将 $\mathbf{Q}_c$ 的平均值记为 A_c 。当 S_t 满足

$$S_t < \tau_1 A_c \quad (33)$$

认为检测结果不可靠。

当 S_t 满足

$$S_t > \tau_2 A_c \quad (34)$$

认为检测结果是可靠的,并更新队列 $\mathbf{Q}_c$ 。

综上所述,目标跟踪状态判断如下

$$T_t = \begin{cases} 1, & \text{其他} \\ 0, & \bar{q}_t/q_t > \tau_q; \|\mathbf{R}_t[\Psi_\Delta] - \mathbf{R}_{t-1}\|_2 > \phi; S_t < \tau_1 A_c \end{cases} \quad (35)$$

当 T_t 为 1 时,表明当前跟踪结果是可靠的,不需要进行目标的重检测。当 $T_t = 0$ 时,表明当前跟踪结果不可靠,需要通过重检测器对目标进行恢复。

2.3 局部区域重检测

目标丢失后,首先通过局部区域搜索模块完成对目标附近潜在出现区域的局部重检测,通过对目标丢失后潜在位置的所在区域使用基础跟踪器进行搜索,找出响应值最大值,来确定目标所在位置的坐标。局部区域搜索如图 5 所示,目标被遮挡后暂时消失在视野里,目标重新出现在视野里时通常在消失位置的附近区域,本文使用目标区域位置扩展搜索的策略,以上一帧目标位置为中心,在 8 个潜在目标位置区域进行目标搜索,具体做法是计算潜在目标位置区域的响应,响应值最大值即为目标的所在区域,为 8 个潜在目标位置的坐标。(x, y)

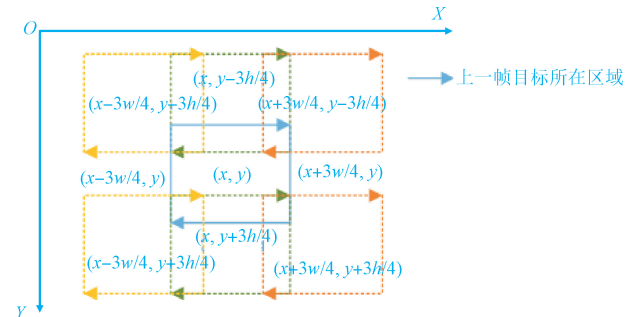


图 5 目标局部区域搜索图

Fig. 5 Target local area search map

为上一帧目标所在位置区域的中心位置坐标, w 为目标搜索区域的宽, h 为目标搜索区域的高。对 8 个目标位置潜在区域使用基础跟踪器中的位置相关滤波器进行目标位置搜索。

2.4 全局区域重检测

通过跟踪状态判断模块判断出短期跟踪失败后, 启用重检测器恢复目标。首先使用局部区域重检测, 然后使用跟踪状态判断模块进行跟踪状态判断, 判断局部区域重检测得到跟踪结果的可靠性, 如果跟踪成功则不启用全局区域重检测, 如果跟踪失败, 则使用全局区域重检测对目标跟踪状态进行恢复。

对目标全局区域重检测, 恢复丢失后的目标。为了实现目标的重新定位, 采用多个长期滤波器构成重检测器。重检测器需要维护多个长期滤波器 $\mathbf{H}^{\text{LT}} = \{\mathbf{h}_i^{\text{LT}}\}_{i \in (1, 2, 3, 4, 5, 6)}$, 其中 $\mathbf{h}_i^{\text{LT}}$ 表示第 i 个长期滤波器。启用重检测器时, 从当前帧中提取特征, 并从 $\mathbf{H}^{\text{LT}}$ 中取出一个滤波器, 通过相关滤波器得到相关响应 $\mathbf{R}^{\text{LT}}$ 。由于目标丢失后, 重新出现的位置是在上一个可靠位置的附近, 上一个可靠位置通过跟踪状态判断模块确定, 所以添加一个目标区域运动先验 $\mathbf{V}(\mathbf{X}_t)$, 将相关响应 $\mathbf{R}^{\text{LT}}$ 与运动先验 $\mathbf{V}(\mathbf{X}_t)$ 通过相乘进行融合, 最终取融合后响应的最大值作为检测目标位置 $\mathbf{X}_t^{\text{LT}}$ 。目标位置和运动先验分别如下

$$\mathbf{X}_t^{\text{LT}} = \max(\mathbf{V}(\mathbf{X}_t) \odot \mathbf{R}^{\text{LT}}) \quad (36)$$

$$\mathbf{V}(\mathbf{X}_t) =$$

$$\frac{1}{2\rho\alpha_w\alpha_s^{\Delta_t}x_h\alpha_s^{\Delta_t}} \exp\left(-\frac{1}{2}(\mathbf{X}_t - \mathbf{X}_c)^{\text{T}} \sum_t \frac{1}{(\mathbf{X}_t - \mathbf{X}_c)}\right) \quad (37)$$

式中: $\mathbf{V}(\mathbf{X}_t)$ 为运动先验, 在本文无人机航拍目标跟踪的场景下中使用运动先验 $\mathbf{V}(\mathbf{X}_t)$ 可以更好地定位目标; $\sum_t (\mathbf{X}_t - \mathbf{X}_c) = \text{diag}[(x_w\alpha_s^{\Delta_t})^2, (x_h\alpha_s^{\Delta_t})^2]$, x_h 表示目标的高, x_w 表示目标的宽; ρ 为运动先验估计参数; $\mathbf{X}_c$ 为通过跟踪状态判断模块确定的最近一个可靠目标位置的坐标; α_s 是尺度变换参数; Δ_t 代表当前帧与通过跟踪状态判断模块确定的可靠跟踪帧的差值。

具体地在启用重检测器时, 跟踪算法顺序从重检测器 $\mathbf{H}^{\text{LT}} = \{\mathbf{h}_i^{\text{LT}}\}_{i \in (1, 2, 3, 4, 5, 6)}$ 取出滤波器计算相关响应, 从尺度比例池 $\mathbf{S}^{\text{LT}}$ 取出比例因子并计算运动先验, 将得到的相关响应和运动先验通过相乘进行融合, 得到融合后的响应用来确定目标所在位置, 图 6 为目标全局区域重检测的过程。

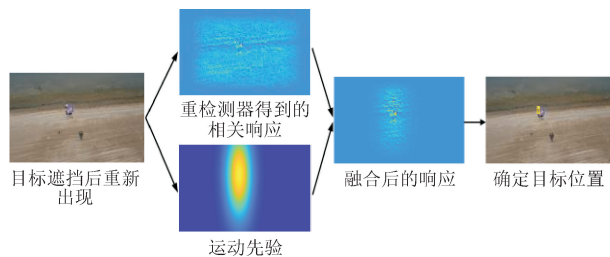


图6 目标全局区域重检测

Fig. 6 Target global region re-detection

对于重检测器 $\mathbf{H}^{\text{LT}} = \{\mathbf{h}_i^{\text{LT}}\}_{i \in (1, 2, 3, 4, 5, 6)}$ 中的滤波器选择不同的更新频率可以使重检测器应对不同的复杂场景, 提升重检测器的鲁棒性。因此重检测器中每个滤波器的更新频率不同, 更新频率为每 f_i^{LT} 帧更新滤波器, $F_i^{\text{LT}} = \{f_i^{\text{LT}}\}_{i \in (1, 2, 3, 4, 5, 6)} = \{0, 1, 5, 25, 50, 200\}$ 。同时, $\mathbf{h}_i^{\text{LT}}$ 根据优化更新策略进行自适应更新, 更新方式如下

$$\mathbf{h}_i^{\text{LT}} = (1 - \eta_5)\mathbf{h}_{i-1}^{\text{LT}} + \eta_5\mathbf{h}_i^{\text{ST}} \quad (38)$$

式中: $\mathbf{h}_i^{\text{LT}}$ 为重检测器中的第 i 个检测滤波器; $\mathbf{h}_i^{\text{ST}}$ 为在第 t 帧求出的短期跟踪滤波器; η_5 为学习率, 与短期位置相关滤波器中的学习率 η 取值一致, $\eta_5 = \eta$ 。通过比较基础跟踪器的跟踪质量 q_i^{ST} 和重检测器的跟踪质量 q_i^{LT} , 确定最终的目标所在位置

$$(\mathbf{X}_t, q_t) = \begin{cases} (\mathbf{X}_t^{\text{ST}}, q_t^{\text{ST}}), & q_t^{\text{ST}} \geq q_t^{\text{LT}} \\ (\mathbf{X}_t^{\text{LT}}, q_t^{\text{LT}}), & \text{其他} \end{cases} \quad (39)$$

式中: $\mathbf{X}_t^{\text{ST}}$ 为基础跟踪器确定的目标位置; $\mathbf{X}_t^{\text{LT}}$ 为重检测器确定的目标位置; $\mathbf{X}_t$ 为最终确定的目标位置。

基于局部-全局区域重检测的长期跟踪算法流程如图 7 所示, 算法流程主要步骤如下。

(1) 输入当前序列第一帧图像和目标的真实边界框, 求出初始的位置相关滤波器和尺度滤波器, 完成模型初始化。

(2) 从第二帧开始通过基础跟踪器定位目标, 利用跟踪状态判断模块进行跟踪可靠判断, 如果计算的目标中心位置可靠, 则继续后续短期跟踪操作, 其中短期跟踪操作与基础跟踪器的算法流程一致; 跟踪不可靠则启用重检测器, 进行局部区域搜索。

(3) 如果局部区域搜索得到的目标中心位置不可靠, 则进行全局区域重检测, 通过比较组合局部区域搜索和全局区域重检测器得到的检测位置, 以及基础跟踪器得到目标位置之间的跟踪质量, 确定最终目标所在位置。

(4) 判断当前序列是否跟踪结束, 如果结束, 则停止跟踪; 若未结束, 则读取当前序列下一帧图像, 重复上述操作, 直至当前序列跟踪结束。

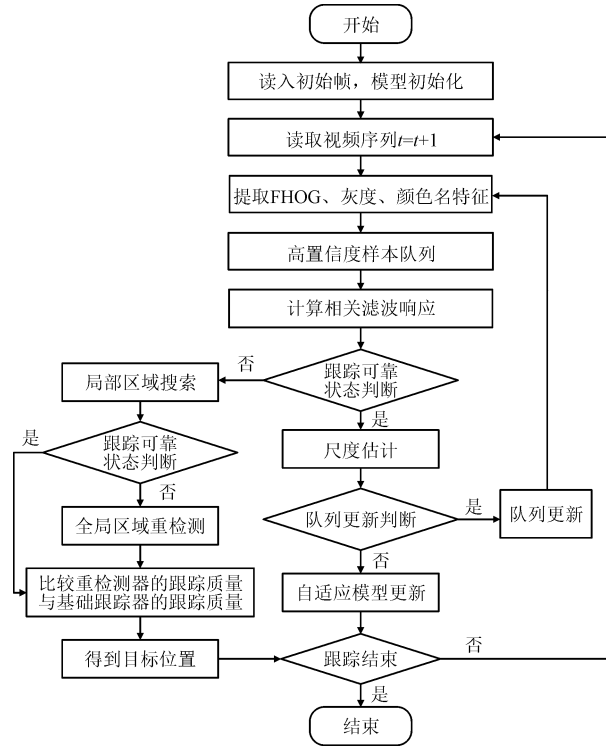


图7 基于局部-全局区域重检测的长期跟踪算法流程

Fig. 7 Flow chart of long-term tracking algorithm based on local-global region re-detection

3 实验与分析

3.1 实验环境与参数设置

3.1.1 实验环境

实验环境为地面站环境,地面站为计算机,算法软件平台为 Matlab2020b。实验硬件配置为:CPU 为 Intel(R) Core(TM) i7-8700;GPU 为 NVIDIA RTX-2080Ti;RAM 为 32 GB。

3.1.2 参数设置

(1)基础跟踪器参数设置如下。相关滤波器中,最近帧 $N_q = 90$, $\tau_q = 2.7$,参数 $\lambda_1 = 1.0$, $\lambda_2 = 0.98$, $\lambda_3 = 16$, $\omega = 20$, $v = 2 \times 10^{-5}$, $\phi = 3\ 000$,学习率 $l_r = 0.2$,ADMM 迭代次数设为 2,步进参数 γ 初始值为 1,最大值 $\gamma^{\max} = 10$, $\alpha = 0.1$,步进参数 ξ 初始值为 1,最大值 $\xi^{\max} = 100$, $\mu = 10$ 。优化更新策略中, $a_1 = 0.6$, $a_2 = 0.45$, $a_3 = 0.43$ 。位置滤波器学习率 $\eta_1 = 0.02$, $\eta_2 = 0.01$,尺度滤波器学习率 $\eta_3 = 0.025$, $\eta_4 = 0.015$ 。

(2)跟踪状态判断模块与重检测器的参数设置如下。在跟踪状态判断模块中, $\tau_1 = 0.6$, $\tau_2 = 0.71$,颜色直方图的学习率 $\eta_{\text{hist}} = 0.04$,颜色直方图划分为 32 个区间, $\lambda = 10^{-3}$ 。重检测器中,参数 $\rho = 0.25\pi$,尺度变换参数 $\alpha_s = 1.02$,尺度比例池 $\mathbf{S}^{\text{LT}} = \{0.5, 0.8, 1.1, 1.3, 1.6, 2\}$ 。

3.2 数据集及其评价标准

实验数据集采用 UAV20L 数据集^[21],利用一次性评估算法(OPE),通过精确度图和成功率图评价算法的跟踪性能,通过算法每秒处理帧数评价实时性能。

精确度。中心点误差 C_{LE} 可以评估跟踪算法性能,用于衡量跟踪目标的边界框中心位置与目标真实边界框中心位置之间的距离, C_{LE} 越小定位目标更准确,具有更高的精度和鲁棒性,公式如下

$$C_{\text{LE}} = \sqrt{(x_t - x_t^{\text{truth}})^2 + (y_t - y_t^{\text{truth}})^2} \quad (40)$$

式中: (x_t, y_t) 表示计算的目标中心点坐标; $(x_t^{\text{truth}}, y_t^{\text{truth}})$ 表示目标的真实中心点坐标。精确度通过 C_{LE} 来计算,具体为数据集视频序列中小于某一阈值帧数与总帧数的比值。通过精确度图可以直观清晰地看出不同算法在同一阈值下的精确度差异,精确度图的坐标横轴和纵轴分别为阈值和精确度。本文设定阈值为 20 像素。

成功率。重叠率 O_R 是跟踪边界框与真实目标边界框之间的重叠程度,用于评估跟踪算法的准确性,重叠率越接近于 1 表明跟踪越准确。成功率通过 O_R 计算,具体为数据集视频序列中大于某阈值的帧数与总帧数的比值。 O_R 定义如下

$$O_R = \frac{|r_t \cap r_t^{\text{truth}}|}{|r_t \cup r_t^{\text{truth}}|} \quad (41)$$

式中: r_t 由算法求出的目标边界框; r_t^{truth} 为目标的真实边界框。通过成功率图可以直观清晰地看出不同算法在同一阈值下的成功率差异,成功率图的坐标横轴和纵轴分别为阈值和成功率。实验采用 0.5 作为阈值。

3.3 定量分析

在无人机长时数据集 UAV20L 上进行实验,通过精确度图和成功率图将本文算法 Our_LT 与基础跟踪器 8 种算法:本文 Our 算法、高效卷积算子跟踪算法(ECO_HC)^[22]、AutoTrack 算法、畸变学习抑制相关滤波器跟踪算法(ARCF)^[23]、STRCF 算法、空间正则化相关滤波器学习跟踪算法(SRDCF)^[24]、实时目标跟踪的互补学习算法(Staple)^[25]和 fDSST 算法进行客观评价。其中,Our 算法表示不包含重检测模块的本文算法,Our_LT 算法表示包含重检测模块的本文算法。

3.3.1 总体性能评估

将本文 Our_LT 算法与其他 8 种算法在 UAV20L 数据集上进行定量评估,所得精确度图和成功率图如图 8 和图 9 所示。

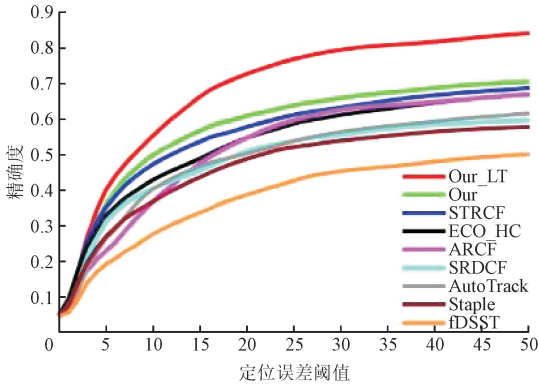


图 8 9 种算法精确度比较

Fig. 8 Accuracy comparison of 9 algorithms

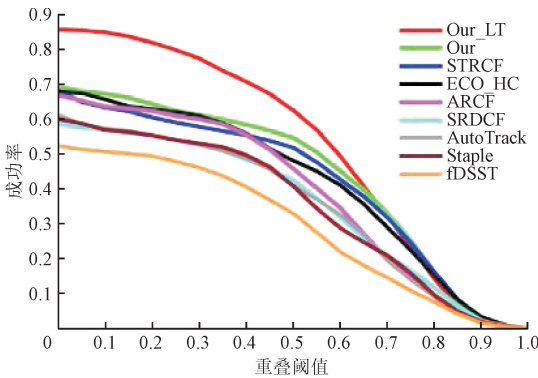


图 9 9 种算法成功率比较

Fig. 9 Comparison of success rates of 9 algorithms

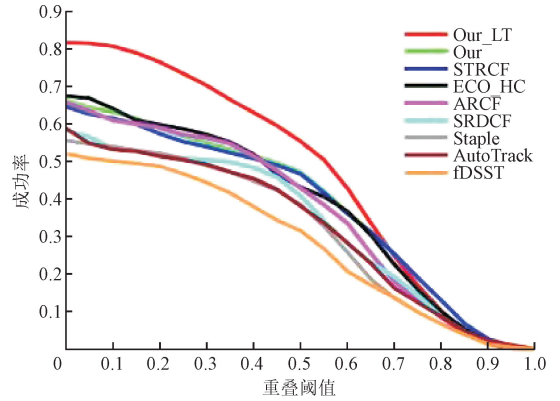
从图 8 和图 9 可以看出:本文 Our_LT 算法的精确度为 0.724,成功率为 0.621,在 9 种算法中排名第一;相比 Our、ECO_HC、ARCF、AutoTrack、STRCF 和 SRDCF 算法,Our_LT 算法的精确度分别提高了 19.5%、33.1%、33.1%、44.5%、25.9% 和 42.8%;Our_LT 算法的成功率分别提高了 14.4%、29.9%、36.5%、52.6%、20.6% 和 53.3%。此外,Our_LT 的在 UAV20L 数据集上的平均运行速度为 34.3FPS,相比于 AutoTrack、ARCF、STRCF 算法分别提高了 20.3%、27.9%、57.6%,满足实时性。综合来看,本文 Our_LT 算法与其余 9 种算法相比,其成功率和精确度得到了较大的提升,在无人机航拍场景下取得了较好的跟踪效果。

3.3.2 属性性能评估

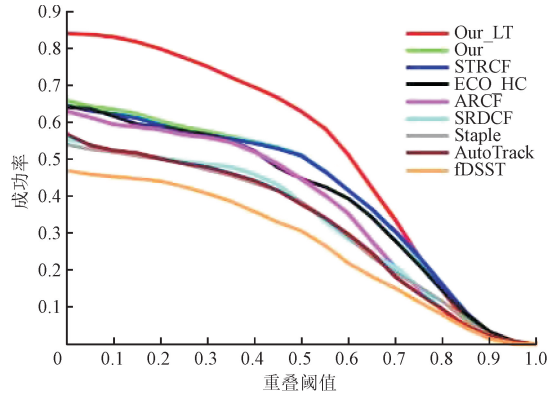
将 Our_LT 算法与其他 8 种算法在 UAV20L 数据集 6 种长期跟踪代表性属性下的成功率进行评估,如图 10 示。其中,阈值取 0.5 时,成功率如表 1 所示。

从图 10 可以看出,在相机运动、移出视野、部分

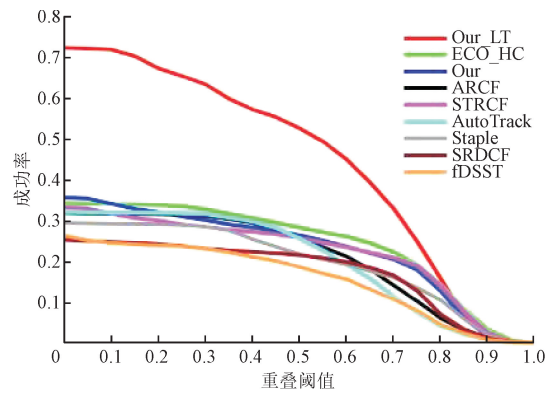
遮挡、完全遮挡、长宽比变化、尺度变化 6 种复杂场景上,Our_LT 算法的成功率在 9 种算法中排名第一。



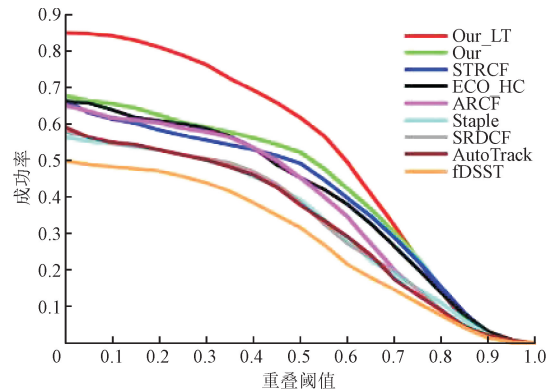
(a) 移出视野



(b) 部分遮挡



(c) 完全遮挡



(d) 相机运动

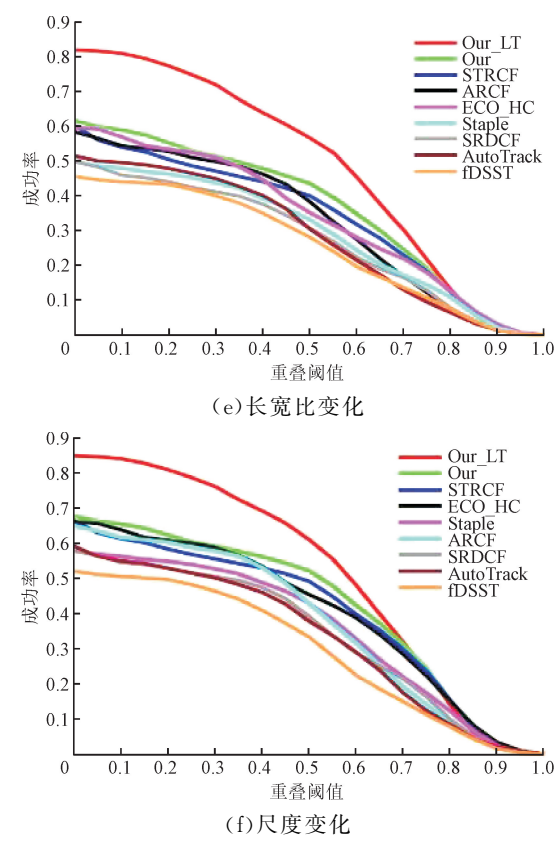


图 10 9 种算法 6 种属性下的成功率比较
Fig. 10 The success rate diagram of 9 algorithms under 6 attributes

表 1 对比了 9 种算法在 6 种属性下的成功率。可以看出,本文 Our_LT 算法在相机运动、移出视野、部分遮挡、完全遮挡、长宽比变化、尺度变化等代表性属性的成功率大于 0.5,具备较好的跟踪性能。Our_LT 算法在长期跟踪代表性属性完全遮挡下,成功率为 0.523,而 Our、ECO_HC、AutoTrack、ARCF、STRCF、SRDCF、Staple、fDSST 等算法的成功率均小于 0.3,所以 Our_LT 算法在完全遮挡属性下跟踪性能优异。

表 1 9 种算法在 6 种属性下的成功率
Table 1 The success rate of 9 algorithms under 6 attributes

算法	成功率					
	OV	POC	FOC	CM	ARC	SV
Our_LT	0.551	0.626	0.523	0.613	0.563	0.607
Our	0.470	0.510	0.263	0.519	0.433	0.519
ECO_HC	0.430	0.446	0.281	0.451	0.349	0.452
AutoTrack	0.379	0.376	0.254	0.376	0.303	0.376
ARCF	0.427	0.444	0.257	0.450	0.380	0.427
STRCF	0.465	0.507	0.256	0.489	0.397	0.489
SRDCF	0.407	0.383	0.214	0.384	0.305	0.390
Staple	0.383	0.383	0.216	0.390	0.330	0.427
fDSST	0.314	0.306	0.185	0.315	0.279	0.332

注:OV 表示移出视野;POC 表示部分遮挡;FOC 表示完全遮挡;CM 表示相机运动;ARC 表示长宽比变化;SV 表示尺度变化;表中加粗数据为最优值。

综上所述可知,本文 Our_LT 算法在无人机长期跟踪的代表性属性下跟踪效果较好。

3.4 定性分析

为了直观比较 9 种跟踪算法在 UAV20L 数据集上的性能表现,使用 2 组具有代表性的视频序列定性分析,图 11、12 为 9 种算法在 2 组序列上的跟踪结果。

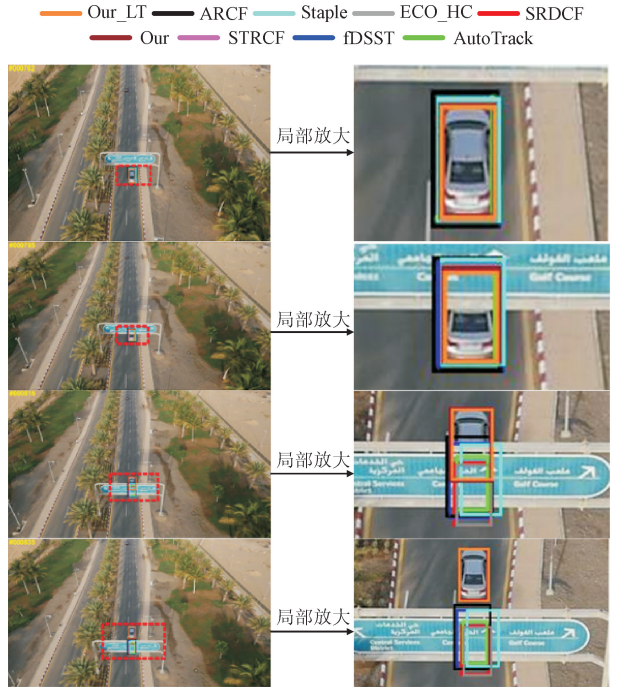


图 11 9 种算法在序列 car9 上的跟踪结果
Fig. 11 The tracking results of 9 algorithms on sequence car9

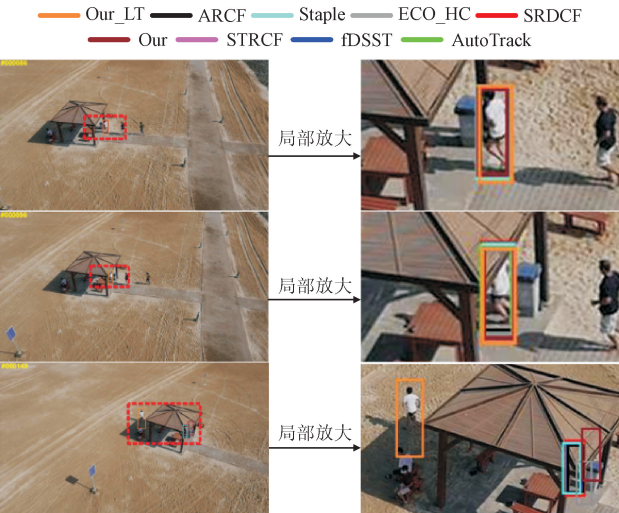


图 12 9 种算法在序列 group2 上的跟踪结果
Fig. 12 The tracking results of 9 algorithms on sequence group2

由图 11 可见,在序列 car9,从开始至第 762 帧,未出现任何目标被遮挡情况,9 种算法均稳定跟踪

目标;在 785 帧时目标经历了部分遮挡,此时 9 种算法均能成功跟踪目标;在第 785 帧至第 816 帧时,目标经历了完全遮挡后重新出现在视野里但仍部分遮挡,此时只有 Our_LT 算法可以成功定位目标,并且得到的目标边界框与目标真实边界框最为接近;在第 835 帧时,目标重新完整出现在视野里,Our_LT 算法可以成功跟踪目标,对目标恢复跟踪,且跟踪框具备较高的精确度和成功率。可见,Our_LT 算法在面对部分遮挡、完全遮挡等场景时具备稳健的鲁棒性,可以在目标丢失后重新出现在视野里时,完成目标的恢复跟踪。

由图 12 可见:在序列 group2 起始到第 86 帧,未出现任何目标被遮挡情况,9 种算法均能稳定跟踪目标;在第 96 帧时,目标出现了部分遮挡的情况,此时 9 种算法仍能成功跟踪目标;在第 96 帧到第 148 帧之间,目标发生完全遮挡,当第 149 帧目标重新完全出现在视野中,此时,只有 Our_LT 算法能够成功重新找回目标,其余算法均跟踪失败。由 group2 序列可以看出,本文 Our_LT 算法在遮挡等复杂场景中具有更好的鲁棒性,可以在目标丢失后重新出现在视野里时,完成目标的恢复跟踪。

4 结 论

无人机在航拍目标长期跟踪中,需要跟踪器在面临遮挡、移出视野等复杂场景跟踪失败后,具备对目标进行恢复跟踪的能力,对鲁棒性要求很高。因此,本文提出了一种基于局部-全局区域重检测的无人机长期跟踪框,在基础跟踪器的基础上,设计跟踪状态判断模块和局部-全局区域的重检测机制,能够在判断目标跟踪状态后,利用局部-全局区域重检测器来恢复跟踪丢失的目标。将本文算法与同类型的 8 种算法在 UAV20L 数据集上进行验证评估,实验结果表明,本文算法在保持实时性的同时,可以较好地应对无人机航拍目标时遮挡和移出视野等长期跟踪复杂场景。

参考文献:

- [1] 黄鹤,李文龙,吴琨,等. 基于 ALCE-SSA 优化的三维无人机低空突防 [J]. 南京大学学报(自然科学版), 2022, 58(3): 448-459.
HUANG He, LI Wenlong, WU Kun, et al. 3D UAV low altitude penetration optimization based on ALCE-SSA [J]. Journal of Nanjing University(Natural Science), 2022, 58(3): 448-459.
- [2] 韩瑞泽,冯伟,郭青,等. 视频单目标跟踪研究进展

综述 [J]. 计算机学报, 2022, 45(9): 1877-1907.

HAN Ruize, FENG Wei, GUO Qing, et al. Single object tracking research: a survey [J]. Chinese Journal of Computers, 2022, 45(9): 1877-1907.

- [3] 李行. 基于多传感器的无人机室内自主导航研究 [D]. 南昌: 南昌大学, 2022.
- [4] 闫超,涂良辉,王丰豪,等. 无人机在我国民用领域应用综述 [J]. 飞行力学, 2022, 40(3): 1-6.
YAN Chao, TU Lianghui, WANG Yuhao, et al. Application of unmanned aerial vehicle in civil field in China [J]. Flight Dynamics, 2022, 40(3): 1-6.
- [5] 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. Piscataway, NJ, USA: IEEE, 2010: 2544-2550.
- [6] HENRIQUES J F, CASEIRO R, MARTINS P, et al. Exploiting the circulant structure of tracking-by-detection with kernels [C]//Computer Vision-ECCV 2012. Berlin, Germany: Springer, 2012: 702-715.
- [7] 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.
- [8] 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. Piscataway, NJ, USA: IEEE, 2014: 1090-1097.
- [9] DANELLJAN M, HÄGER G, KHAN F, et al. Accurate scale estimation for robust visual tracking [C]//British Machine Vision Conference 2014. Durham, UK: BMVA Press, 2014: 1-11.
- [10] DANELLJAN M, HÄGER G, KHAN F S, et al. Discriminative scale space tracking [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(8): 1561-1575.
- [11] GALOOGAHI H K, FAGG A, LUCEY S. Learning background-aware correlation filters for visual tracking [C]//2017 IEEE International Conference on Computer Vision (ICCV). Piscataway, NJ, USA: IEEE, 2017: 1144-1152.
- [12] LI Feng, TIAN Cheng, ZUO Wangmeng, et al. Learning spatial-temporal regularized correlation filters for visual tracking [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2018: 4904-4913.

- [13] LI Yiming, FU Changhong, DING Fangqiang, et al. AutoTrack: towards high-performance visual tracking for UAV with automatic spatio-temporal regularization [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA; IEEE, 2020: 11920-11929.
- [14] KALAL Z, MIKOLAJCZYK K, MATAS J. Tracking-learning-detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(7): 1409-1422.
- [15] MA Chao, YANG Xiaokang, ZHANG Chongyang, et al. Long-term correlation tracking [C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA; IEEE, 2015: 5388-5396.
- [16] YAN Bin, ZHANG Xinyu, WANG Dong, et al. Alpha-refine: boosting tracking performance by precise bounding box estimation [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA; IEEE, 2021: 5285-5294.
- [17] FELZENSZWALB P F, GIRSHICK R B, MCALL-ESTER D, et al. Object detection with discriminatively trained part-based models [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(9): 1627-1645.
- [18] VAN DE WEIJER J, SCHMID C, VERBEEK J, et al. Learning color names for real-world applications [J]. IEEE Transactions on Image Processing, 2009, 18(7): 1512-1523.
- [19] 黄鹤, 李文龙, 吴琨, 等. 动态自适应特征融合的MFOPA跟踪器 [J]. 电子学报, 2023, 51(5): 1350-1358.
- HUANG He, LI Wenlong, WU Kun, et al. MFOPA tracker with dynamic adaptive feature fusion [J]. Acta Electronica Sinica, 2023, 51(5): 1350-1358.
- [20] DAI Kenan, WANG Dong, LU Huchuan, et al. Visual tracking via adaptive spatially-regularized correlation filters [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA; IEEE, 2019: 4665-4674.
- [21] MUELLER M, SMITH N, GHANEM B. A benchmark and simulator for UAV tracking [C]//Computer Vision-ECCV 2016. Cham, Germany; Springer International Publishing, 2016: 445-461.
- [22] DANELLJAN M, BHAT G, KHAN F S, et al. ECO: efficient convolution operators for tracking [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA; IEEE, 2017: 6931-6939.
- [23] HUANG Ziyuan, FU Changhong, LI Yiming, et al. Learning aberrance repressed correlation filters for real-time UAV tracking [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway, NJ, USA; IEEE, 2019: 2891-2900.
- [24] 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 (ICCV). Piscataway, NJ, USA; IEEE, 2015: 4310-4318.
- [25] 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 (CVPR). Piscataway, NJ, USA; IEEE, 2016: 1401-1409.

(编辑 刘杨 陶晴)