全天实时跟踪无人机目标的多正则化相关滤波算法

王法胜[1] 李 富[1] 尹双双[1] 王 星[1] 孙福明[1] 朱 兵[2]

摘 要 相关滤波算法 (Correlation filter, CF) 已广泛应用于无人机目标跟踪. 然而, 受无人机 (Unmanned aerial vehicle, UAV) 平台本身计算性能的制约, 现有的无人机相关滤波跟踪算法大都仅采用手工特征来描述目标的外观, 难以获得目标的全面语义信息. 并且这些跟踪算法仅能较好地进行光照条件良好场景下的跟踪, 而在跟踪夜间场景下的目标时性能严重下降. 此外, 相关滤波跟踪器采用余弦窗口来抑制循环移位产生的边界效应, 缩小了样本提取区域, 产生了训练样本污染的问题, 这不可避免地降低了跟踪器的性能. 针对以上问题, 提出全天实时多正则化相关滤波算法 (All-day and real-time multi-regularized correlation filter, AMRCF) 跟踪无人机目标. 首先, 引入一个自适应图像增强模块, 在不影响图像各通道颜色比例的前提下, 对获得的图像进行增强, 以提高夜间目标跟踪性能. 其次, 引入一个轻量型的深度网络来提取目标的深度特征, 并与手工特征一起来表示目标的语义信息. 此外, 在算法框架中嵌入高斯形状掩膜, 在抑制边界效应的同时, 有效避免训练样本污染. 最后, 在 5 个公开的无人机基准数据集上进行充分的实验. 实验结果表明, 所提出的算法与多个先进的相关滤波跟踪器相比, 取得了有竞争力的结果, 且算法的实时速度约为 25 fps, 能够胜任无人机的目标跟踪任务.

关键词 无人机目标跟踪, 相关滤波, 自适应图像增强模块, 轻量型深度网络, 高斯形状掩膜

引用格式 王法胜, 李富, 尹双双, 王星, 孙福明, 朱兵. 全天实时跟踪无人机目标的多正则化相关滤波算法. 自动化学报, 2023, **49**(11): 2409−2425

DOI 10.16383/j.aas.c220424

All-day and Real-time Multi-regularized Correlation Filter for UAV Object Tracking

WANG Fa-Sheng[1] LI Fu[1] YIN Shuang-Shuang[1] WANG Xing[1] SUN Fu-Ming[1] ZHU Bing[2]

Abstract Correlation filter (CF) has been widely used in unmanned aerial vehicle (UAV) object tracking. Due to the computational limitation of the UAV platform, the existing UAV tracking algorithms rely heavily on the hand-crafted features, which cannot obtain all the semantic information of the target. Meanwhile, the existing tracking algorithms focus on the tracking of daytime targets, and ignore the tracking problem of nighttime targets. In addition, when the correlation filter-based tracker uses the cosine window to suppress the boundary effect caused by the cyclic shift, the sampling area is shrunk, which causes the contamination of training samples and inevitably deteriorates the performance of the tracker. Aiming at the above problems, we propose an all-day and real-time multi-regularized correlation filter (AMRCF) for UAV object tracking. Firstly, an adaptive image enhancement module is introduced to enhance the image without affecting its color ratio of each channel and improve the nighttime tracking performance. Secondly, a lightweight deep network is introduced to extract the deep features of the target and represent the semantic information together with the hand-crafted features. Thirdly, a Gaussian-shaped mask is embedded in the correlation filter framework, which can effectively avoid the contamination of training samples while suppressing the boundary effect. Finally, extensive experiments are conducted on five publicly available UAV benchmark datasets. The experimental results show that the proposed algorithm achieves competitive result compared with several state-of-the-art correlation filter-based trackers, and the real-time speed is about 25 fps, which makes it competent for UAV object tracking task.

Key words Unmanned aerial vehicle (UAV) object tracking, correlation filter (CF), adaptive image enhancement module, lightweight deep network, Gaussian-shaped mask

Citation Wang Fa-Sheng, Li Fu, Yin Shuang-Shuang, Wang Xing, Sun Fu-Ming, Zhu Bing. All-day and real-time multi-regularized correlation filter for UAV object tracking. *Acta Automatica Sinica*, 2023, **49**(11): 2409−2425

收稿日期 2022-05-23 录用日期 2023-02-10
Manuscript received May 23, 2022; accepted February 10, 2023
国家自然科学基金 (61972068, 61976042), 国家重点研发计划 (2021YFC3320300), 兴辽英才计划 (XLYC2007023), 辽宁省高等学校创新人才支持计划 (LR2019020) 资助
Supported by National Natural Science Foundation of China (61972068, 61976042), National Key Research and Development Program of China (2021YFC3320300), Liaoning Revitalization Talents Program (XLYC2007023), and Innovative Talents Program for Liaoning Universities (LR2019020)
本文责任编委 丛杨
Recommended by Associate Editor CONG Yang
1. 大连民族大学信息与通信工程学院 大连 116600 2. 哈尔滨工业大学电子与信息工程学院 哈尔滨 150001
1. School of Information and Communication Engineering, Dalian Minzu University, Dalian 116600 2. School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin 150001

近年来, 无人机 (Unmanned aerial vehicle, UAV) 目标跟踪技术在道路规划、交通监控等问题中得到广泛应用[1]. 目前, 有两类目标跟踪算法得到研究人员的高度关注: 基于相关滤波 (Correlation filter, CF) 的跟踪方法和基于深度学习 (Deep learning, DL) 的跟踪方法[2]. 后者使用一个或多个深度网络对目标进行跟踪, 大大提高了跟踪的准确性和鲁棒性. 但是这类方法需要较高的算力, 导致其在无人机平台上难以满足实时跟踪目标的要求. 而 CF 跟踪方法因具有较低的计算成本和较高的跟踪精度, 在无人机目标跟踪领域得到了广泛的应用[3]. 这类方法通过空间域中的循环移位生成训练样本[4], 但循环移位操作会导致出现边界效应问题, 从而降低滤波器的判别力. 为此, 研究人员提出在跟踪框架中加入余弦窗口来抑制边界效应[4−5], 并在此基础上结合其他策略来进一步增强算法的判别能力. 例如, 具有尺度和旋转自适应的长时目标跟踪算法[6], 提出将傅里叶−梅林变换与核相关滤波方法结合, 以提高滤波器对尺度和旋转参数估计的鲁棒性和准确性. 融合显著性与运动信息的相关滤波跟踪算法[7], 将目标的像素级概率性表征模型与相关滤波算法融合, 并且提出一种基于显著性的观测模型, 提高算法面对复杂跟踪场景时识别目标的能力. 另有研究人员提出空间正则化[5, 8−10] 和上下文学习[11−14] 等策略提高相关滤波的性能. 但这些算法均使用余弦窗口, 而余弦窗口的引入对训练样本造成了污染, 降低了算法的跟踪性能.

CF 跟踪算法中采用的手工特征主要包括灰度特征、方向梯度直方图 (Histogram of oriented gradient, HOG) 和颜色名称 (Color names, CN) 等. Bolme 等[15] 首次提出了仅使用灰度特征的 CF 跟踪算法. 后续的一些工作[4, 16−17] 为了学习具有多通道的滤波器, 使用了灰度特征、HOG、CN 等多种特征的组合. 但手工特征无法表述目标全部的语义信息. 随着 ResNet、VGG (Visual geometry group) 等提取深度特征的卷积神经网络 (Convolutional neural network, CNN) 在诸如图像分类和图像识别等具有挑战性视觉任务上取得巨大成功, 研究人员开始将深度特征加入到 CF 跟踪框架中[18−19], 显著提高了 CF 跟踪算法的鲁棒性. 但 CNN 特征的计算负担较高, 严重影响跟踪器的实时性. 受到无人机平台本身计算性能的限制 (仅搭载 CPU), 现有的无人机目标跟踪算法[20−22] 大都仅使用手工特征来描述目标的外观. 此外, 这些跟踪算法仅能较好地处理光照条件良好的跟踪场景, 而在跟踪夜间场景下的目标时则遇到性能严重下降的挑战.

为了解决以上问题, 本文提出一种全天实时多正则化相关滤波算法 (All-day and real-time multi-regularized correlation filter, AMRCF) 跟踪无人机目标, 即, 以多正则化的相关滤波器为基础框架[21], 引入自适应图像增强模块、轻量型深度网络和高斯形状掩膜. 首先, 引入一个自适应图像增强模块, 在不影响图像各通道颜色比例的前提下, 对获得的图像进行增强, 以提高夜间目标的跟踪性能. 其次, 引入一个轻量型的深度网络来提取目标的深度特征, 并与手工特征一起来表示目标的语义信息. 此外, 在算法框架中嵌入高斯形状掩膜, 在抑制边界效应的同时, 有效避免训练样本污染. AMRCF 与其他先进的目标跟踪算法在 DTB70 基准数据集[23] 上的总体比较如图 1 所示.

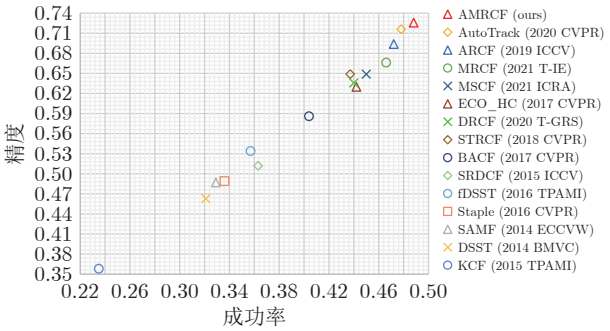


图 1 AMRCF 与其他算法在 DTB70 上的总体性能比较

Fig. 1 Overall performance of AMRCF compared with other algorithms on DTB70

本文主要贡献如下:

1) 以多正则化的相关滤波器为基础框架, 引入一个自适应图像增强模块, 自动判别无人机目标场景的光照条件, 并在不改变图像各通道颜色比例的前提下, 根据判别结果对图像进行相应的处理.

2) 引入轻量型的深度网络[24] 来提取目标的深度特征, 并与手工特征一起表示目标的语义信息. 该网络的计算量小, 使用 CPU 平台便可以进行特征的提取, 并且不会大幅降低算法的跟踪速度.

3) 在 CF 跟踪框架中嵌入高斯形状掩膜, 建立相关滤波跟踪算法模型, 并使用交替方向乘子方法优化求解, 在抑制边界效应的同时, 减少样本污染问题.

4) 在 5 个公开的无人机基准数据集 DTB70[23]、UAVTrack112[25]、UAVDark135[26]、VisDrone-SOT2018[27] 和 UAV123[28] 上进行综合实验. 结果表明, 所提出的算法具有较高的精确度和鲁棒性. 本文算法代码及结果发布于 https://gitee.com/he_bing_DLMZ/amrcf.

1 相关工作

本节简要介绍与本文内容相关的跟踪算法, 包

括基于相关滤波的跟踪算法以及无人机目标跟踪算法.

1.1 基于相关滤波的跟踪算法

基于 CF 的跟踪算法将跟踪任务描述为岭回归问题, 目的是训练滤波器来区分目标与背景区域. 首个 CF 跟踪器 MOSSE[15] 被提出后, Henriques 等[4, 29] 将核技巧 (Kernel trick) 引入到相关滤波公式中, 并利用循环矩阵的特性将求解过程转换到傅里叶域, 简化了滤波器优化过程. 此后, 研究人员又从很多方面提出不同方法来提高 CF 跟踪器的性能, 包括尺度估计[30−31]、边界效应抑制[5, 8−10, 32−33]、时间退化解决方案[34−36]、特征表示[19, 37]、注意力机制[38−40] 以及特征去冗余[14, 41−42].

CF 跟踪算法通过空间域的循环移位操作来生成更多的训练样本, 导致出现边界效应问题. Danelljan 等[5] 提出空间正则化相关滤波跟踪算法 (Spatially regularized discriminative correlation filter, SRDCF), 该算法加入空间正则化组件, 根据相关滤波器的空间位置, 对滤波器系数进行惩罚来抑制边界效应. 蒲磊等[10] 结合 SRDCF 与 CNN, 提出了基于深度空间正则化的 CF 算法. 该算法选取 VGG 网络第 5 层卷积特征提取目标的空间可靠区域, 将该区域信息用于对样本进行裁剪并引入目标函数, 从而建立空间约束模型. Xu 等[35] 在时空正则化相关滤波算法[34] (Spatial-temporal regularized correlation filter, STRCF) 的基础上, 通过将多特征通道与滤波器的结构进行关联, 以在空间维度上进行组特征选择来解决边界效应问题. 王科平等[43] 以跟踪目标的邻域信息为基础, 引入掩膜矩阵, 提取目标的空间信息, 并通过在目标函数中构建时空感知约束项, 强化相关滤波器对时空二元信息的学习能力, 增强滤波器对干扰信息的鲁棒性.

上述方法在抑制边界效应上存在一个共同点, 即使用余弦窗口. 余弦窗口与采样区域相乘之后, 边界附近区域被强制为零, 使得获取的训练样本极易受到噪声的污染, 进而导致滤波器性能下降. 为此, Li 等[44] 提出在背景感知相关滤波算法[8] (Background aware correlation filter, BACF) 中用高斯掩膜替换余弦窗口以解决样本污染问题. 与使用空间域循环移位产生的样本来训练滤波器不同, 文献 [45] 通过传统的滑动窗口对一组真实且密集的块进行采样, 用采样后的样本来训练滤波器, 进而从根本上抑制边界效应.

1.2 无人机目标跟踪算法

基于 CF 的跟踪器在无人机目标跟踪中得到了广泛的应用, 但平面内外旋转、光照不足、遮挡、形变等给无人机目标跟踪带来了极大的挑战. 为了更好地应对这些挑战, 研究人员进行了大量的研究工作以改进无人机目标跟踪算法. Li 等[22] 提出一种自适应时空正则化无人机跟踪算法, 通过超参数对正则化项进行自动调整来抑制边界效应, 同时利用隐藏在响应图中的局部和全局信息平滑时间. Fu 等[46] 提出一种具有双重正则化策略的无人机跟踪器, 通过对涉及相关运算的滤波器直接正则化来抑制边界效应, 并利用正则化器识别目标区域来减少背景噪声. 为了应对无人机目标的外观突变, 文献 [20] 在 CF 跟踪框架中嵌入自适应混合标签以增强算法对外观突变的抗性, 并将理想标签与相关滤波器进行联合优化来保持时间一致性. 传统的 CF 无人机跟踪算法没有考虑历史样本信息与不同通道之间的差异性, Ye 等[21] 通过规范响应的偏差对响应变化进行平滑, 同时通过通道的可靠性来实现自适应通道权重分布, 从而更好地适应对象外观变化, 以此增强跟踪器从背景中区分目标的能力. 针对目前没有专用于夜间跟踪的算法, Li 等[26] 在目标跟踪框架中加入自适应光照强度模块, 利用该模块改善夜间目标序列的质量, 以此提高算法的夜间跟踪效果.

2 本文算法

本节介绍所提出的跟踪算法 AMRCF, 算法的整体框架如图 2 所示. AMRCF 由三部分组成, 第一部分是自适应图像增强模块, 第二部分是特征提取模块, 第三部分是嵌入高斯形状掩膜的多正则化相关滤波 (Multi-regularized correlation filter, MRCF) 模块. 对于输入的视频序列, 首先将序列的第 1 帧取出做光照条件判别, 并根据判别结果对后续图像做不同的处理. 随后, 在每一帧中, 计算高斯形状掩膜, 并提取样本区域中的手工特征和深度特征, 将两种特征进行线性加权. 最后, 将优化的滤波器 h_k, 通道权重分布 β_k 与样本特征进行相关运算以得到响应图, 响应图中的最大值位置即为目标位置.

2.1 自适应图像增强模块

为了提高算法的夜间跟踪能力, 本文引入一种自适应图像增强模块. 该模块首先使用光照强度表达式, 将图像复杂的光照信息转换成一个常数, 并利用该常数构建光照条件判别式. 然后在像素级别上根据判别结果对图像采取不同的处理措施, 以达到图像增强的目的.

对于一幅大小为 $a \times b$ 的 RGB 图像 $\mathcal{I} \in \mathbf{R}^{a\times b\times 3}$, 利用式 (1) 计算出它的光照强度 $\mathcal{L}$

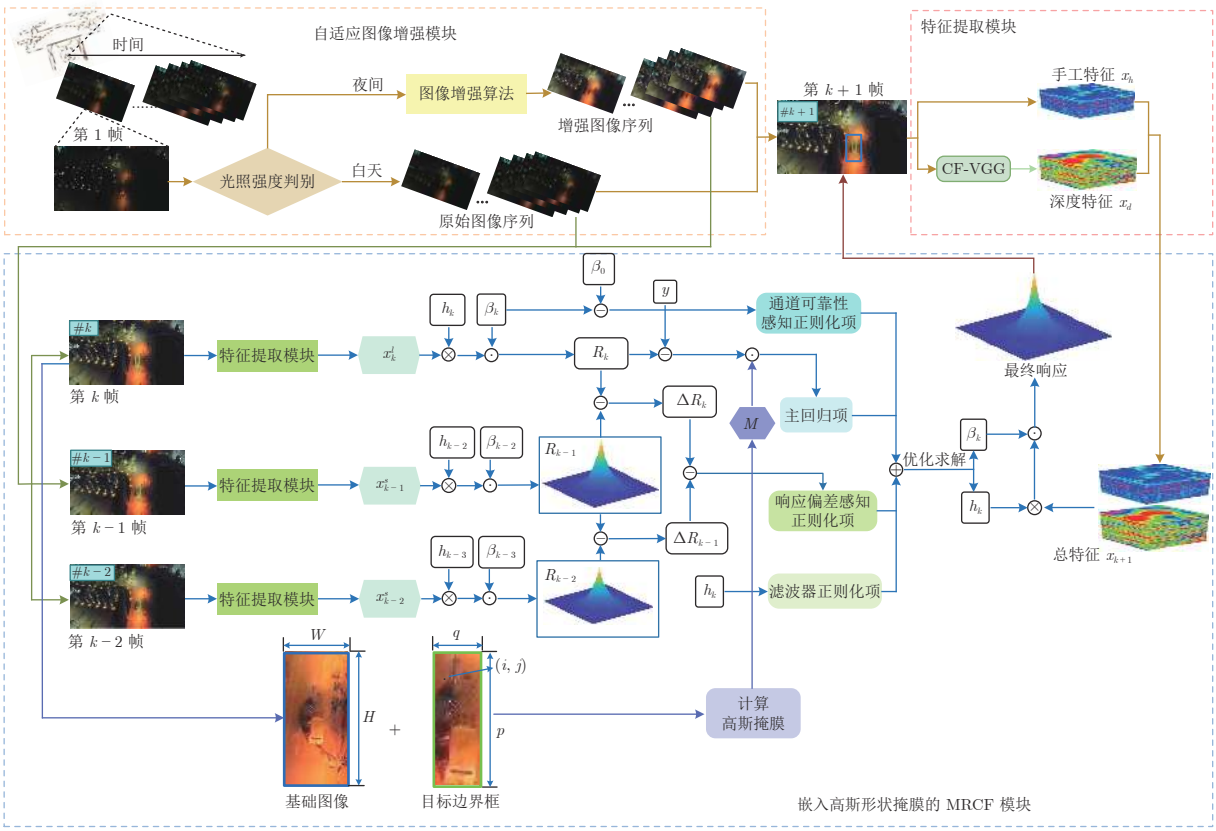


图 2 AMRCF 跟踪算法框架图

Fig. 2 Framework of the proposed AMRCF algorithm

$$\mathcal{L}(i,\ j,\ \mathcal{I}) = \sum \alpha_m \psi_m(\mathcal{I}(i,\ j)),\ \ m \in \{\mathrm{R},\ \mathrm{G},\ \mathrm{B}\} \tag{1}$$

其中, $\mathcal{L}(i,\ j,\ \mathcal{I})$ 表示图像 $\mathcal{I}$ 在坐标 $(i,\ j)$ 处的光照强度, $\psi_m(\mathcal{I}(i,\ j))$ 表示图像 $\mathcal{I}$ 在颜色通道 m 的坐标 $(i,\ j)$ 处的光照强度值. 具体来说, $\psi_{\mathrm{R}}(\mathcal{I}(i,\ j))$ 表示红色通道中对应位置的光照强度值, 且颜色通道系数 α_{R}, α_{G}, α_{B} 满足 $\alpha_{\mathrm{R}} + \alpha_{\mathrm{G}} + \alpha_{\mathrm{B}} = 1$.

将式 (1) 代入式 (2) 中, 得到图像 $\mathcal{I}$ 的光照强度对数平均值 $\tilde{\mathcal{L}}(\mathcal{I})$

$$\tilde{\mathcal{L}}(\mathcal{I}) = \exp\left(\frac{1}{ab}\sum_{i,\ j}\ln(\delta + \mathcal{L}(i,\ j,\ \mathcal{I}))\right) \tag{2}$$

其中, δ 是一个很小的非零数. 至此, 便完成了由复杂光照信息到简单常数的转换.

文献 [26] 针对式 (2) 有效性的实验结果表明, 光照强度的对数平均值可以有效地表达图像的光照强度信息. 基于此, 构建式 (3) 所示的判别式

$$S(\mathcal{I}) = \begin{cases} 1, & \tilde{\mathcal{L}}(\mathcal{I}) < \tau \\ 0, & \tilde{\mathcal{L}}(\mathcal{I}) \geq \tau \end{cases} \tag{3}$$

式 (3) 是一个夜间场景判决器, 其中 τ 为给定阈值, 也就是说, $S(\mathcal{I}) = 1$ 表明图像 $\mathcal{I}$ 是一幅夜间图像, 反之则是白天图像.

通过以上公式, 便可以判别图像所处的光照场景, 但自适应图像增强模块的最终目的是对图像进行增强. 为此, 采取下面的方案对图像进行增强处理.

首先, 使用 $\mathcal{L}(i,\ j,\ \mathcal{I})$ 和 $\tilde{\mathcal{L}}(\mathcal{I})$ 计算出式 (4) 所示的全局自适应因子 $\mathcal{L}_{gb}(i,\ j,\ \mathcal{I})$

$$\mathcal{L}_{gb}(i,\ j,\ \mathcal{I}) = \frac{\ln\left(\frac{\mathcal{L}(i,\ j,\ \mathcal{I})}{\tilde{\mathcal{L}}(\mathcal{I})} + 1\right)}{\ln\left(\frac{\mathcal{L}_{\max}(\mathcal{I})}{\tilde{\mathcal{L}}(\mathcal{I})} + 1\right)} \tag{4}$$

其中, $\mathcal{L}_{\max}(\mathcal{I}) = \max(\mathcal{L}(i,\ j,\ \mathcal{I}))$.

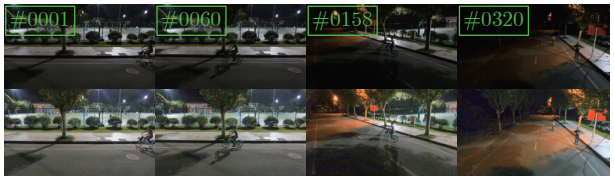
然后用全局自适应因子 $\mathcal{L}_{gb}(i,\ j,\ \mathcal{I})$ 对图像 $\mathcal{I}$ 进行像素级别的增强处理

$$\psi_m(\mathcal{I}_e(i,\ j)) = \psi_m(\mathcal{I}(i,\ j)) \times \frac{\mathcal{L}_{gb}(i,\ j,\ \mathcal{I})}{\mathcal{L}(i,\ j,\ \mathcal{I})} \tag{5}$$

其中, $\mathcal{I}_e$ 为增强后的图像.

通过式 (1) ~ (5), 实现对图像的自适应增强处理, 具体的实验效果对比如图 3 所示. 与黑暗场景下的原始图像相比, 经过自适应图像增强模块处理后, 图像的质量得到了明显提升, 体现在光照强度变化和图像颜色变化上, 这有利于特征提取模块提

(a) UAVDark_bike3



(b) UAVDark_bike8

(c) UAVDark_boat1

(d) UAVDark_girl1

图 3　原始图像与增强图像的对比 (每组图像的第一行是原始图像, 第二行是相应的增强图像)

Fig. 3　Comparison of original image and enhanced image (The first row of each image set is the original image, and the second row is the corresponding enhanced image)

取目标特征.

2.2　轻量型深度网络

VGG、ResNet 等深度网络大大提升了目标跟踪算法的性能, 但大幅降低了跟踪的实时性, 难以满足无人机目标跟踪需求. Wang 等[24] 提出一种轻量型的深度网络, 该网络的计算量较小, 在提高算法精度的同时, 能达到实时的跟踪速度.

文献 [24] 将 VGG-M 作为教师网络, 使用知识蒸馏[47] 来生成轻量型的学生网络 CF-VGG. 教师网络和学生网络的具体架构如图 4 所示. 在蒸馏的过程中, 使用式 (6) 所示的语义保真度损失 $\mathcal{L}_{\text{fidelity}}$ 来确保学生网络能够保持与教师网络相似的特征表示能力

$$\mathcal{L}_{\text{fidelity}} = \mathcal{L}_{\text{target}} + \mathcal{L}_{\text{search}} = ||\varphi(t) - \varsigma(t)||^2 + ||\varphi(s) - \varsigma(s)||^2 \quad (6)$$

其中, t 和 s 分别表示目标区域和搜索区域, $\mathcal{L}_{\text{target}}$, $\mathcal{L}_{\text{search}}$ 是目标区域损失和搜索区域损失. $\varphi(\cdot)$ 表示

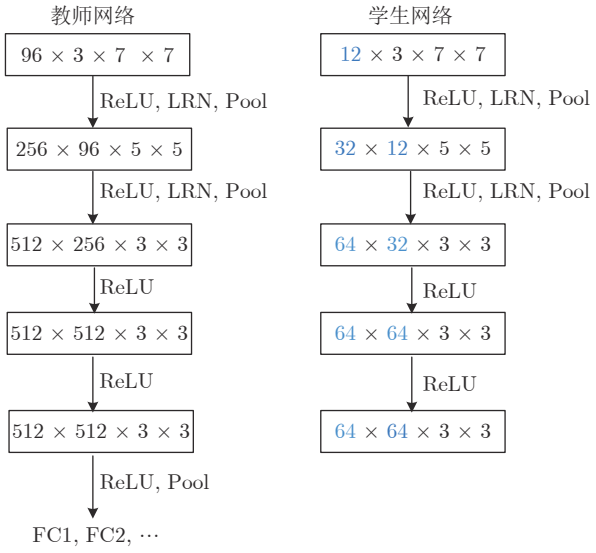


图 4　教师网络和学生网络的架构

Fig. 4　Architectures of the teacher network and the student network

学生网络可训练的特征嵌入, $\varsigma(\cdot)$ 表示教师网络的固定嵌入.

学生网络的离线提取过程在主要面向目标识别任务的 ImageNet 数据集[48] 上进行. 为此, 文献 [24] 提出式 (7) 所示的相关跟踪损失 $\mathcal{L}_{\text{tracking}}$, 将学生网络的任务从目标识别转移到目标跟踪

$$\mathcal{L}_{\text{tracking}} = \sum_c ||r_c - v_c||^2$$
$$\text{s.t. } r_c = \mathcal{F}^{-1}(\hat{h}_c^* \odot \hat{\varphi}_c(s)),\ c \in \{\text{low, middle, high}\} \quad (7)$$

其中, $\odot$ 表示哈达玛积 (Hadamard product), $\mathcal{F}^{-1}$ 表示逆离散傅里叶变换, $c \in \{\text{low, middle, high}\}$ 表征低中高卷积层的特征表示的索引, $\hat{h}_c^*$ 是相应卷积层滤波器的复共轭, $\hat{\varphi}_c(s)$ 表示相应层特征嵌入的复共轭, r_c, v_c 分别是相关响应图和真实值标签, 相关跟踪损失是两者之间的欧氏距离. 相关跟踪损失将分类的源目标转换为回归目标进行跟踪, 同时通过学习相似度 (或模板匹配) 来评估目标对象在帧之间的微小外观变化.

本文将 CF-VGG 引入到 AMRCF 框架中来提取目标的深度特征, 与手工提取的特征一起来描述目标. 具体而言, 对于第 $k+1$ 帧, 以上一帧 k 的目标位置为中心, 以自适应的方式裁剪样本区域, 对于该区域, CF-VGG 使用 $stride = 2$ 的步长进行下采样, 并用 ReLU 函数对更有助于精确定位[18] 的第 3 卷积层上的 CNN 特征进行处理, 以描述目标的语义信息.

2.3 AMRCF 算法

2.3.1 基线算法

MRCF 的目标函数如式 (8) 所示

$$\varepsilon(h_k,\,\beta_k)=\frac{1}{2}\sum_{d=1}^{D}\left|\left|y-\beta_k^d x_k^d\otimes(P^{\mathrm{T}}h_k^d)\right|\right|_2^2+\frac{\kappa}{2}\sum_{d=1}^{D}\left|\left|h_k^d\right|\right|_2^2+\varepsilon_1+\varepsilon_2 \tag{8}$$

$$\varepsilon_1=\frac{\lambda}{2}\sum_{d=1}^{D}\left|\left|\Delta R_k^d-\Delta R_{k-1}^d\right|\right|_2^2 \tag{9}$$

$$\varepsilon_2=\frac{\gamma}{2}\left|\left|\beta_k-\beta_0\right|\right|_2^2 \tag{10}$$

其中, T 表示矩阵的转置, $y\in\mathbf{R}^N$ 是期望响应, x_k^d, $h_k^d\in\mathbf{R}^N$ 分别是第 k 帧的第 d 通道所对应的矢量化特征和滤波器. $\otimes$ 代表空间域的卷积操作, 而 P 是用于减轻边界效应的二进制矩阵. κ 是预定义的系数, ε_1 和 ε_2 分别是响应偏差感知正则化项 (Response deviation-aware regularization) 和通道可靠性感知正则化项 (Channel reliability-aware regularization), 表达式分别如式 (9)、式 (10) 所示. λ 是控制响应偏差感知正则化权重的超参数, γ 是预定义的常数. β_k, β_0 分别是模型训练过程中的通道权重分布和初始权重分布, $\Delta R_k^d\in\mathbf{R}^N$ 表征两个连续帧之间的响应差异, 其具体表达式见式 (11)

$$\Delta R_k^d=S(R_k^d)-S(R_{k-1}^d) \tag{11}$$

其中, R_k^d 表示第 k 帧第 d 通道上的响应值.

2.3.2 高斯形状掩膜

CF 跟踪器通过空间域中的循环移位操作来获取训练样本, 但没有考虑到采样区域的边界而导致训练样本出现边界不连续性, 从而引起边界效应. 为抑制边界效应, 研究人员在模型的训练过程中加入余弦窗口. 但余弦窗口更强调基础图像的中心区域, 与采样区域相乘之后, 边界附近区域被强制为零, 从而在模型训练过程中造成样本污染. 为此, 本文将高斯形状掩膜嵌入跟踪器的目标函数中, 在抑制边界效应的同时, 减少训练样本污染问题. 高斯掩膜 M 的定义如下所示

$$M(i,\,j)=\begin{cases}\exp\left(-\left(\dfrac{i}{p\upsilon}\right)^2-\left(\dfrac{j}{q\upsilon}\right)^2\right), & |i|\le\dfrac{H}{2}-\dfrac{p}{2},\ |j|\le\dfrac{W}{2}-\dfrac{q}{2}\\ 0, & \text{otherwise}\end{cases} \tag{12}$$

其中, $p\times q$ 和 $H\times W$ 分别是目标边界框和基础图像的尺寸, $(i,\,j)$ 是样本区域中的坐标点. υ 是用来控制训练样本权重衰减速度的参数.

图 5 为相关的可视化图. 图 5(a) 中的红框是基础图像区域, 用于获取训练样本, 绿框是目标区域, 其大小分别是 $H\times W$ 和 $p\times q$. 训练样本附带的边界效应是基础图像循环移位而使样本边界不连续造成的, 图中紫色框所代表的样本中包含了目标区域及其周围的背景区域, 其内部区域为连续区域, 不包含图像边界. 而黄色框所代表的样本中则包含了图像边界. 图 5(b) 是循环样本与余弦窗口相乘后的可视化图, 样本的边界效应得到抑制, 但也抑制了目标边界框之外的像素. 黄色框中的样本存在大量的黑色区域, 使得该样本被黑色区域污染. 因此, 引入余弦窗口后存在使远离目标中心的负样本质量降低的风险. 图 5(c) 是高斯掩膜的可视化图像, 从图像中心到四周, 权值呈递减趋势, 以此突出中心样本, 并降低边缘样本的重要性, 达到抑制边界效应和防止样本污染的目的.

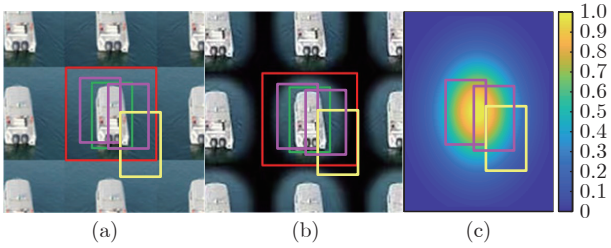


图 5 边界效应抑制可视化图 ((a) 空间域循环移位产生的带有边界效应的训练样本; (b) 加入余弦窗口后, 带有污染的训练样本; (c) 高斯形状掩膜可视化图, 样本的中心距离图像中心越近, 其权重越大, 重要性越高, 反之则越低)

Fig. 5 Visualization of boundary effect suppression ((a) Training samples with boundary effect generated by cyclic shifts in the spatial domain; (b) Training samples with contamination after adding cosine window; (c) Gaussian-shaped mask visualization, the closer the center of the sample is to the center of the image, the higher the weight and importance, and vice versa)

2.3.3 AMRCF 算法的建模和优化

在目标函数中嵌入高斯形状掩膜, 得到本文算法 AMRCF 的目标函数, 如式 (13) 所示

$$\varepsilon(h_k,\,\beta_k)=\frac{1}{2}\sum_{d=1}^{D}\left|\left|\left(y-\beta_k^d x_k^d\otimes\left(P^{\mathrm{T}}h_k^d\right)\right)\odot M\right|\right|_2^2+\frac{\kappa}{2}\sum_{d=1}^{D}\left|\left|h_k^d\right|\right|_2^2+\varepsilon_1+\varepsilon_2 \tag{13}$$

定义辅助变量 $g_k^d=P^{\mathrm{T}}h_k^d$, 将 ε_1 和 ε_2 的表达式

(9) 和式 (10) 代入式 (13), 得到更新后的函数表达式

$$\begin{aligned}\varepsilon(h_k,\, g_k,\, \beta_k) = \frac{1}{2}\sum_{d=1}^{D}\Big(\big|\big|\big(y-\beta_k^d x_k^d\otimes g_k^d\big)\odot M\big|\big|_2^2 + \\ \kappa\left|\left|h_k^d\right|\right|_2^2 + \gamma\left|\left|\beta_k^d-\beta_0^d\right|\right|_2^2 + \\ \lambda\left|\left|\Delta R_k^d-\Delta R_{k-1}^d\right|\right|_2^2\Big)\end{aligned} \tag{14}$$

令 $R_k^d=\beta_k^d x_k^d\otimes g_k^d$, 其表示第 k 帧的第 d 通道的响应. 应用帕萨瓦尔定理将上述时域表达式 (14) 转换到频域中, 得到式 (15)

$$\begin{aligned}&\varepsilon(h_k,\, \hat{g}_k,\, \beta_k) = \\ &\frac{1}{2}\sum_{d=1}^{D}\left(\frac{1}{N}\left|\left|\big(\hat{y}-\beta_k^d\hat{x}_k^d\odot\hat{g}_k^d\big)\odot M\right|\right|_2^2 + \right. \\ &\kappa\left|\left|h_k^d\right|\right|_2^2 + \gamma\left|\left|\beta_k^d-\beta_0^d\right|\right|_2^2 + \\ &\left.\frac{\lambda}{N}\left|\left|\beta_k^d\hat{x}_k^d\odot\hat{g}_k^d-2\widehat{S(R_{k-1}^d)}+\widehat{S(R_{k-2}^d)}\right|\right|_2^2\right) \\ &\text{s.t. } \hat{g}_k^d=\sqrt{N}FP^{\mathrm{T}}h_k^d\end{aligned} \tag{15}$$

其中, $F\in\mathbf{C}^{N\times N}$ 是离散傅里叶变换矩阵, $\hat{\cdot}$ 表示离散傅里叶变换操作, 即 $\hat{\rho}=\sqrt{N}F\rho$.

设拉格朗日乘子为 $\zeta\in\mathbf{R}^{N\times D}$, $\zeta=[\zeta^1,\, \zeta^2,\, \cdots,\, \zeta^D]$, 惩罚因子为 μ, 将式 (15) 重新表述为

$$\begin{aligned}\mathcal{L}(h_k,\, \hat{g}_k,\, \beta_k,\, \hat{\zeta}) = \varepsilon(h_k,\, \hat{g}_k,\, \beta_k) + \\ \frac{\mu}{2}\sum_{d=1}^{D}\left\|\hat{g}_k^d-\sqrt{N}FP^{\mathrm{T}}h_k^d+\frac{\hat{\zeta}_k^d}{\mu}\right\|_2^2\end{aligned} \tag{16}$$

为了最大限度地提高求解速度, 应用交替方向乘子法 (Alternative direction multiplier method, ADMM), 通过迭代求解以下子问题来优化式 (16).

1) 子问题 h_k^*

假设 $\hat{g}$, $\hat{\zeta}$, β 为常数, 则可通过式 (17) 得到 h^*

$$\begin{aligned}h_k^{d*(i+1)} = \arg\min_{h_k}\left(\frac{1}{2}\sum_{d=1}^{D}\left(\kappa\left\|h_k^{d(i)}\right\|_2^2 + \right.\right. \\ \left.\left.\mu\left\|\hat{g}_k^{d(i)}-\sqrt{N}FP^{\mathrm{T}}h_k^{d(i)}+\frac{\hat{\zeta}_k^{d(i)}}{\mu}\right\|_2^2\right)\right) = \\ \frac{N\left(\mu g_k^d+\zeta_k^d\right)}{\kappa+\mu N}\end{aligned} \tag{17}$$

其中, 上标 i 和 $*$ 分别表示第 i 次迭代和共轭运算. 为了使表达式更加简洁, 在式 (17) 的最终结果里省略了上标 i. 此外, $g_k^d=\frac{1}{\sqrt{N}}PF^{\mathrm{T}}\hat{g}_k^d$ 和 $\zeta_k^d=\frac{1}{\sqrt{N}}PF^{\mathrm{T}}\hat{\zeta}_k^d$ 分别是 $\hat{g}_k^d$ 和 $\hat{\zeta}_k^d$ 的逆离散傅里叶变换表达式.

2) 子问题 g_k^*

在给定 h, $\hat{\zeta}$, β 的情况下, 可以通过求解式 (18) 来得到 $\hat{g}^*$

$$\begin{aligned}&\hat{g}_k^{*(i+1)} = \\ &\arg\min_{\hat{g}_k}\left(\frac{1}{2}\sum_{d=1}^{D}\left(\frac{1}{N}\left\|\big(\hat{y}-\beta_k^d\hat{x}_k^d\odot\hat{g}_k^d\big)\odot M\right\|_2^2+\right.\right. \\ &\frac{\lambda}{N}\left\|\beta_k^d\hat{x}_k^d\odot\hat{g}_k^d-2\widehat{S(R_{k-1}^d)}+\widehat{S(R_{k-2}^d)}\right\|_2^2+ \\ &\left.\left.\mu\left\|\hat{g}_k^d-\sqrt{N}FP^{\mathrm{T}}h_k^d+\frac{\hat{\zeta}_k^d}{\mu}\right\|_2^2\right)\right)\end{aligned} \tag{18}$$

因 $\hat{y}$ 依赖于 $\hat{x}(n)=[\hat{x}_k^1(n),\, \hat{x}_k^2(n),\, \cdots,\, \hat{x}_k^D(n)]^{\mathrm{T}}$, 并且 $\hat{g}(n)=\left[\hat{g}_k^1(n),\, \hat{g}_k^2(n),\, \cdots,\, \hat{g}_k^D(n)\right]^{\mathrm{T}}$, 式 (18) 可以分成 N 个独立的子表达式进行单独求解, 独立子表达式如式 (19) 所示

$$\begin{aligned}&\hat{g}_k^{*(i+1)}(n) = \\ &\arg\min_{\hat{g}_k(n)}\left(\frac{1}{2}\sum_{d=1}^{D}\left(\frac{1}{N}\left\|(\hat{y}(n)-\hat{x}_k^{\mathrm{T}}(n)\beta_k\hat{g}_k(n))\odot\right.\right.\right. \\ &\left.M\right\|_2^2+\frac{\lambda}{N}\left\|\hat{x}_k^{\mathrm{T}}(n)\beta_k\hat{g}_k(n)-2\widehat{S(R_{k-1}^d)}(n)+\right. \\ &\left.\widehat{S(R_{k-2}^d)}(n)\right\|_2^2+\mu\left\|\hat{g}_k(n)-\sqrt{N}FP^{\mathrm{T}}h_k(n)+\right. \\ &\left.\left.\left.\frac{\hat{\zeta}_k(n)}{\mu}\right\|_2^2\right)\right)\end{aligned} \tag{19}$$

式 (19) 对 $\hat{g}_k(n)$ 求导并令导数为零, 得到 $\hat{g}_k^{*(i+1)}(n)$ 的表达式如下

$$\begin{aligned}&\hat{g}_k^{*(i+1)}(n) = \\ &\big(\mathrm{diag}\,\{M\odot M+\lambda\boldsymbol{I}\}\times \\ &(\beta_k^{\mathrm{T}}\hat{x}_k(n)\hat{x}_k^{\mathrm{T}}(n)\beta_k)+\mu N\boldsymbol{I}\big)^{-1}\times \\ &\Big(\beta_k^{\mathrm{T}}\hat{x}_k(n)\hat{y}(n)+2\lambda\beta_k^{\mathrm{T}}\hat{x}_k(n)\widehat{S(R_{k-1})}(n)- \\ &\lambda\beta_k^{\mathrm{T}}\hat{x}_k(n)\widehat{S(R_{k-2})}(n)-N\hat{\zeta}(n)+\mu N\hat{h}_k(n)\Big)\end{aligned} \tag{20}$$

其中, $\boldsymbol{I}$ 是单位矩阵, $\mathrm{diag}\{\cdot\}$ 表示构造的对角矩阵. 式 (20) 中的矩阵求逆过程会影响计算速度, 为此, 应用 Sherman-Morrison 定理, 即 $(A+uv^{\mathrm{T}})^{-1}=A^{-1}-A^{-1}u(\boldsymbol{I}+v^{\mathrm{T}}A^{-1}u)^{-1}v^{\mathrm{T}}A^{-1}$, 进一步推导式 (20), 得到 $\hat{g}_k^{*(i+1)}(n)$ 的最终表达式为

$$\hat{g}_k^{*(i+1)}(n)=\frac{1}{uN}\left(\boldsymbol{I}-\frac{\beta_k\hat{x}_k(n)\hat{x}_k^{\mathrm{T}}(n)\beta_k\tilde{M}M}{uN+\lambda\hat{x}_k^{\mathrm{T}}(n)\beta_k\beta_k\hat{x}_k(n)}\right)\times\left(\beta_k\hat{x}_k(n)\hat{y}_k(n)\tilde{M}M+2\lambda\beta_k\hat{x}_k(n)\widehat{S(R_{k-1})}(n)-\lambda\beta_k\hat{x}_k(n)\widehat{S(R_{k-2})}(n)-N\hat{\zeta}(n)+\mu N\hat{h}_k(n)\right)\tag{21}$$

其中, $\tilde{M}$ 是矩阵 M 的复共轭矩阵.

3) 子问题 β

假设 h, $\hat{\zeta}$, $\hat{g}$ 为常数, 则可以通过求 β 的偏导数, 并令偏导数为零来得到 β 的最终表达式

$$\beta_k^{d(i+1)}=\arg\min_{\beta_k^d}\left(\frac{1}{2N}\sum_{d=1}^{D}\left\|\left(\hat{y}-\beta_k^d\hat{x}_k^d\odot\hat{g}_k^d\right)\odot M\right\|_2^2+\frac{\lambda}{2N}\sum_{d=1}^{D}\left\|\beta_k^d\hat{x}_k^d\odot\hat{g}_k^d-2\widehat{S(R_{k-1}^d)}+\widehat{S(R_{k-2}^d)}\right\|_2^2+\frac{\gamma}{2}\sum_{d=1}^{D}\left\|\beta_k^d-\beta_0^d\right\|_2^2\right)=\frac{(2\lambda\widehat{S(R_{k-1}^d)}+\tilde{M}M\hat{y}-\lambda\widehat{S(R_{k-2}^d)})(\hat{x}_k^d\odot\hat{g}_k^d)^{\mathrm{T}}+\gamma N\beta_0^d}{(\lambda\boldsymbol{I}+\tilde{M}M)(\hat{x}_k^d\odot\hat{g}_k^d)^{\mathrm{T}}(\hat{x}_k^d\odot\hat{g}_k^d)+\gamma N}\tag{22}$$

4) 更新拉格朗日乘子

与文献 [49] 一致, 本文分别使用式 (23) 和式 (24) 来更新拉格朗日乘子和惩罚因子

$$\hat{\zeta}_k^{(i+1)}=\hat{\zeta}_k^{(i)}+\mu\left(\hat{g}_k^{*(i+1)}-\hat{h}_k^{*(i+1)}\right)\tag{23}$$

$$\mu^{i+1}=\min(\mu_{\max},\sigma\mu^{(i)})\tag{24}$$

2.3.4 模型更新

本文采用线性插值方法对目标的外观模型 $\hat{x}_k^{\mathrm{model}}$ 进行更新, 更新后的表达式为

$$\hat{x}_k^{\mathrm{model}}=(1-\eta)\hat{x}_{k-1}^{\mathrm{model}}+\eta\hat{x}_k^l\tag{25}$$

其中, η 是模型更新率, $\hat{x}_k^l$ 表示从傅里叶域的学习区域中提取的目标特征, 并使用 $\hat{x}_k^{\mathrm{model}}$ 求解 g_k^* 和 β.

2.3.5 跟踪框架

跟踪过程中, 对于每一个视频序列, 首先用式 (3) 来判别序列第 1 帧图像的光照条件, 并根据判别结果使用自适应图像增强模块对后续帧进行不同的处理. 随后, 对于每一帧 $k+1$, 以前一帧 k 中的目标位置为中心来裁剪相应尺度的样本区域, 以提取其手工特征和深度特征, 并采用式 (26) 将两种特征进行线性求和以得到最终的目标特征

$$x_{\mathrm{total}}=\vartheta x_h+(1-\vartheta)x_d\tag{26}$$

其中, x_h 和 x_d 分别是手工特征和深度特征, 而 x_{total} 是总特征, ϑ 是平衡两种特征的参数.

最后, 将每个通道的响应与优化的通道权重进行元素乘积, 则最终的响应函数 R_{k+1} 为

$$R_{k+1}=\mathcal{F}^{-1}\left(\sum_{d=1}^{D}\beta_k^d\odot\hat{x}_{k+1}^{d,\,s}\odot\hat{g}_k^d\right)\tag{27}$$

其中, $\hat{x}_{k+1}^{d,\,s}$ 是从傅里叶域的搜索区域中提取的目标特征. 响应的最大值位置即为目标的估计位置, 并利用 DSST[30] 估计目标尺度的策略来估计目标大小.

3 实验分析

本节采用 5 个公开的无人机目标跟踪基准 DTB70、UAVTrack112、UAVDark135、VisDroneSOT2018 和 UAV123, 与 14 个先进的相关滤波跟踪算法进行对比实验, 包括 ARCF[50]、AutoTrack[22]、MRCF[21]、MSCF[20]、DRCF[46]、SRDCF[5]、STRCF[34]、Staple[51]、SAMF[52]、ECO_HC[37]、DSST[30]、fDSST[30]、KCF[4] 和 BACF[8].

3.1 实验细节

为公平起见, 受测跟踪算法的参数均来自官方版本, 均在同一平台进行.

1) 实验平台. 本文所有实验都在配备 Intel(R) Xeon(R) E-2224 CPU @ 3.4 GHz 16 GB 内存的 Windows10 系统上进行, 使用的软件是 Matlab R2018a.

2) 实验特征. AMRCF 采用 HOG、CN、灰度特征和深度特征的组合来表示目标对象. 其他算法采用的目标特征与原文保持一致, 实验时未作修改.

3) 实验参数. 颜色通道强度系数 α_{R}, α_{G}, α_{B} 分别取 0.299、0.587、0.114, 夜间场景指示器的判决阈值 τ 取 0.148, 光照强度对数平均值式 (2) 中 δ 取 0.001. 高斯掩膜中用来控制训练样本权重衰减速度的参数 $\upsilon=4$. 在目标函数求解中, 设置 ADMM 的迭代次数为 3, 学习率是 $\eta=[0.0199,\ 0.0199]$, 正则化系数分别是 $\kappa=0.01$, $\lambda=0.004$, $\gamma=10$, 并用 $\mu=1$, $\mu_{\max}=100$, $\sigma=10$ 来更新拉格朗日乘子, 用 $\vartheta=0.7$ 来平衡手工特征与深度特征.

4) 评估指标. 所有评估均基于一次性评估 (One-pass evaluation, OPE), 其包括两个指标, 即精确度和成功率[53]. 通过估计边界框与真实边界框之间的中心位置误差 (Center location error, CLE) 来

衡量精确度, 用精确度图 (Precision plot, PP) 中 CLE 低于 20 像素的帧的百分比对跟踪器进行排名. 通过计算估计边界框与真实边界框之间的交并比 (Intersection over union, IoU) 来衡量成功率, 并使用成功图 (Success plot, SP) 的曲线下面积 (Area under curve, AUC) 对跟踪器进行排名. 此外, 使用帧率对算法的跟踪速度进行排名.

3.2　消融实验

3.2.1　跟踪性能对比

为验证 AMRCF 算法中自适应图像增强模块、轻量型深度网络和高斯掩膜对本文算法性能的影响, 在 5 个无人机目标跟踪基准数据集上进行了消融实验, 实验结果如表 1 所示.

其中, 与基线算法相比, 双下划线表示性能增强, 单下划线表示性能减弱. Baseline 是基线算法 MRCF, Baseline+Ada、Baseline+CF-VGG、Baseline+M、Baseline+Ada+CF-VGG 分别表示在 MRCF 框架中加入自适应图像增强模块、轻量型深度网络、高斯掩膜、自适应图像增强模块和轻量型深度网络. Baseline+Ada+CF-VGG+M 则是加了所有模块的本文算法 AMRCF.

在由白天序列构成的 DTB70 数据集上, 根据式 (4) 的判别结果, 不进行图像增强处理, 所以只添加自适应图像增强模块的算法结果与基线算法结果持平. 只添加高斯掩膜的算法结果低于基线算法结果, 当某一帧的估计边界框发生漂移时, 下一帧由基础图像循环移位产生的训练样本因不包含目标而使高斯掩膜无效, 致使跟踪性能降低. 而同时添加 3 个模块的算法达到了最好的实验效果.

在由夜间序列构成的 UAVDark135 基准数据集上, Baseline+Ada、Baseline+CF-VGG、Baseline+M 和 Baseline+Ada+CF-VGG 算法的性能都超越了基线算法, 且在只添加单个模块的算法中, Baseline+Ada 获得了最好的实验结果, 与基线算法相比, 在精确度和成功率上分别获得 0.9% 和 0.8% 的提升. 这表明, 自适应图像增强模块的引入能显著提升算法的夜间跟踪效果. 添加 3 个模块的 Baseline+Ada+CF-VGG+M 提升最大, 分别提升 1.7% 和 0.9%.

在 UAV123 数据集上, 除 Baseline+Ada 和 Baseline+Ada+CF-VGG 外, 其他改进算法的成功率均低于基线算法. UAV123 包含动画游戏等虚拟视频序列和较多快速移动序列, 本文算法对于此类序列难以获得提升. 在精确度方面, Baseline+Ada+CF-VGG 以及加入 3 个模块的算法均比基线算法的精确度高, 表明深度特征与其他模块共同作用能够在一定程度上提高算法精确度.

综合上述的分析, 同时添加自适应图像增强模块、轻量型深度网络和高斯形状掩膜后, 本文算法在 5 个基准数据集、共 10 组指标数据上, 有 8 组结果得到提升, 而减弱的指标数据与基线算法的对应指标相比, 其降幅较小, 从而验证了 AMRCF 算法中各模块的有效性.

3.2.2　跟踪速度对比

为比较不同模块对计算复杂度及跟踪速度的影响, 本节在 5 个数据集上, 对包含不同模块的算法进行帧率对比, 实验结果如表 2 所示. Ada 代表自适应图像增强模块, 它根据序列首帧的光照情况对整个序列采取不同的处理策略, 对夜间序列需要采取逐像素重计算的方式来进行增强, 因此该模块的计算复杂度取决于图像帧的大小. UAVDark135 和 UAVTrack112 数据集中均含有夜间序列, Baseline+Ada 对此类序列进行增强处理, 相较于基线算

表 1　消融实验结果对比

Table 1　Comparison of ablation experiment results

基准数据集	评价指标	Baseline	Baseline+Ada	Baseline+CF-VGG	Baseline+M	Baseline+Ada+CF-VGG	Baseline+Ada+CF-VGG+M
DTB70	精确度	0.666	0.666	0.689	0.656	0.689	**0.726**
	成功率	0.466	0.466	0.479	0.457	0.479	**0.488**
UAVDark135	精确度	0.593	0.602	0.598	0.590	0.608	**0.610**
	成功率	0.455	0.463	0.462	0.458	0.461	**0.464**
UAVTrack112	精确度	0.681	0.672	0.692	0.678	0.713	0.712
	成功率	0.467	0.463	0.469	0.466	0.481	**0.484**
UAV123	精确度	0.693	0.693	0.689	0.674	0.708	0.694
	成功率	0.485	0.485	0.471	0.475	0.487	0.478
VisDrone-SOT2018	精确度	0.812	0.811	0.802	0.795	0.811	**0.816**
	成功率	0.600	0.600	0.591	0.584	0.596	0.598

表 2 消融实验帧率对比
Table 2 Frame rate comparison of ablation experiment

基准数据集	Baseline	Baseline+Ada	Baseline+CF-VGG	Baseline+M	Baseline+Ada+CF-VGG+M
DTB70	34.2662	34.5329	27.2841	32.7516	**26.56060**
UAVDark135	25.9653	21.3622	22.5572	23.4383	**14.59290**
UAVTrack112	36.5963	31.9615	30.6200	34.6642	**28.57300**
UAV123	35.0628	34.3915	28.0824	34.6348	**27.30980**
VisDrone-SOT2018	33.3641	32.9790	27.1486	31.8434	**25.77150**
平均	33.0509	31.0454	27.1385	31.4665	**24.56156**

法, 计算复杂度增加, 帧率都下降了约 4.6 fps. DTB70、UAV123 和 VisDrone-SOT2018 不含夜间序列, 自适应图像增强模块判断每序列首帧的光照条件, 不做增强处理, 帧率下降可忽略不计.

本文使用轻量级深度网络 CF-VGG 第 10 层的输出作为深度特征来描述目标的语义信息, 该层数据输出的计算量大, 耗费时间长, 对帧率影响明显, 在 5 个数据集上的平均帧率下降了约 5.9 fps. 高斯掩膜的计算复杂度取决于目标边界框和基础图像的尺寸, 嵌入跟踪框架后, 算法整体复杂度略有上升, 在 5 个数据集上的平均帧率下降了约 1.6 fps. 引入 3 个模块的 AMRCF 算法计算复杂度最高, 在 UAVDark135 和 UAVTrack112 上分别为 14.5929 fps 和 28.573 fps, 平均帧率约为 25 fps, 降低约 8.5 fps, 但整体上仍然可满足实时性的要求. 需要强调的是, 本文实验均是在仅配备有单 CPU 的计算机上进行, 以最大程度模拟无人机运行环境.

3.2.3 不同深度网络的对比

本节将 CF-VGG 与其教师网络 VGG 进行对比消融实验. 在本文算法框架中, 将 CF-VGG 替换成其教师网络, 得到 Baseline+Ada+M+VGG, 实验结果如表 3 所示. 由表 3 可知, 与基线算法相比, 采用 VGG 深度网络的特征能够显著提高跟踪的精确度和成功率, 但算法帧率却急剧下降, 3 个数据集分别下降约 20.7 fps、16.7 fps 和 21.8 fps. 这表明, VGG 网络结构复杂, 特征提取过程耗费大量时间. 而本文算法 Baseline+Ada+M+CF-VGG 在 3 个数据集上的精确度和成功率均高于采用 VGG 网络的算法, 且帧率方面大大超过了采用 VGG 网络的算法. 该结果充分说明, 相较于教师网络 VGG, 使用 CF-VGG 网络能够大幅提高跟踪帧率, 而其在精确度和成功率方面的提升则归因于在生成 CF-VGG 过程中采用了相关跟踪损失.

3.3 定量分析

消融实验证明了 AMRCF 算法中各模块的有效性, 本节为得到更加准确的实验结果, 在 5 个无人机目标跟踪数据集上, 与 14 种 CF 跟踪算法进行实验对比, 结果如图 6 所示.

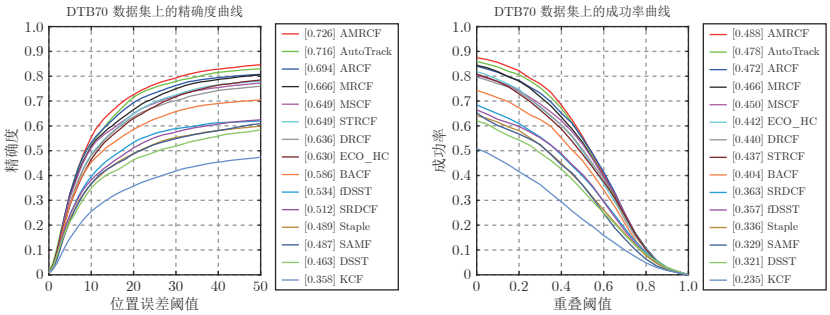
表 3 CF-VGG 与教师网络 VGG 性能对比
Table 3 Comparison of performance between CF-VGG and teacher network VGG

基准数据集	评价指标	Baseline	Baseline+Ada+M+VGG	Baseline+Ada+M+CF-VGG
DTB70	精确度	0.6660	0.7080	**0.7260**
	成功率	0.4660	0.4670	**0.4880**
	帧率	34.2662	13.5239	26.5606
UAVDark135	精确度	0.5930	0.6040	**0.6100**
	成功率	0.4550	0.4630	**0.4640**
	帧率	25.9653	9.2928	14.5929
UAVTrack112	精确度	0.6810	0.7110	**0.7120**
	成功率	0.4670	0.4820	**0.4840**
	帧率	36.5963	14.7906	28.5730

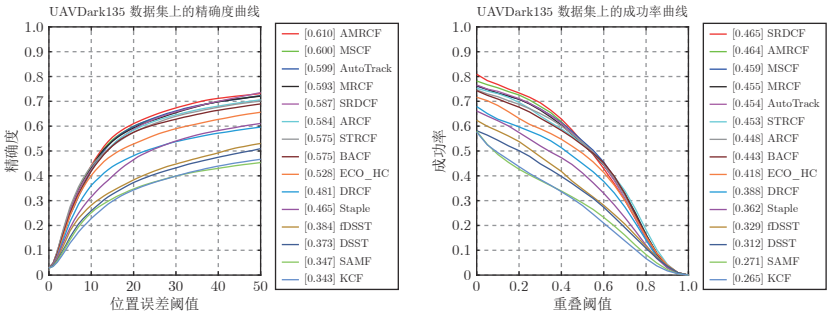
3.3.1 整体性能评估

1) 白天跟踪性能. 所使用的 5 个基准数据集中, DTB70、UAV123 和 VisDrone-SOT2018 均由白天序列构成, 使用这 3 个数据集进行白天跟踪性能的评估. 在 DTB70 上, AMRCF 以 72.6% 的精确度和 48.8% 的成功率位居第 1, 两项指标均比排名第 2 的 AutoTrack 高 1.0%, 而且比基线算法 MRCF 分别高 6.0% 和 2.2%. 在 UAV123 上, 本文算法虽未获得最好的性能, 但以 69.4% 的精确度 (比 MRCF 高 0.1%) 和 47.8% 的成功率位于前列. 在 VisDrone-SOT2018 上, 本文算法分别以 81.6% 的精确度位居第 1 和 59.8% 的成功率位居第 2, 在成功率上比 MRCF 低 0.2%, 但精确度要高 0.4%.

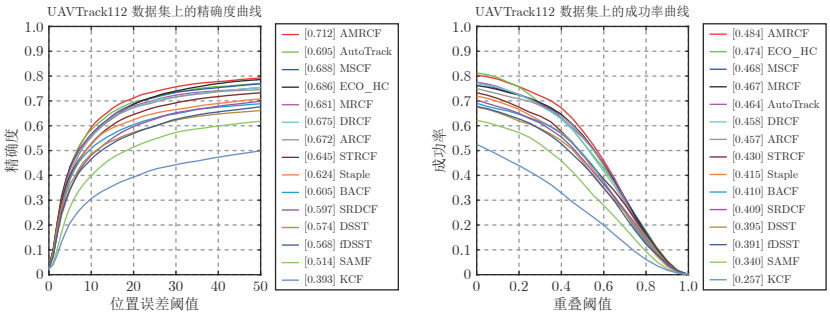
2) 夜间跟踪性能. UAVDark135 是包含 135 个夜间序列的无人机目标跟踪数据集. 为了验证 AMRCF 的夜间跟踪性能, 本文在该基准数据集上进行了相关实验. 实验结果表明, 本文算法能对夜间目标进行精确跟踪, 并以 61.0% 的精确度位居第 1, 以 46.4% 的成功率位居第 2. 在成功率上, 仅比排名第 1 的 SRDCF (46.5%) 低 0.1%, 但在精确度上, 比排名第 2 的 MSCF 高 1.0%. 且本文算法在两组



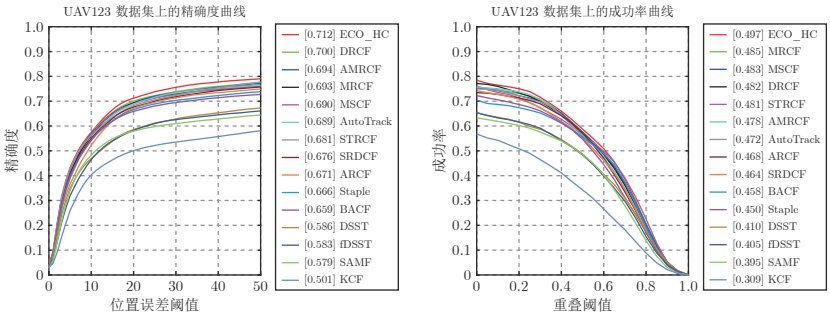
(a) DTB70



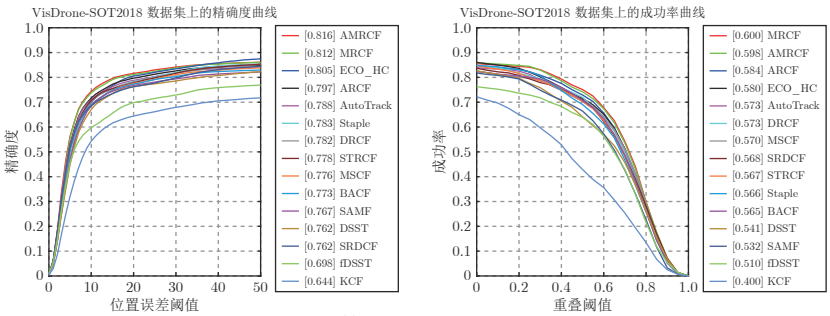
(b) UAVDark135



(c) UAVTrack112



(d) UAV123



(e) VisDrone-SOT2018

图 6　AMRCF 与其他算法在 5 个无人机目标跟踪基准上的实验结果

Fig. 6　Experimental results of AMRCF and other algorithms on five UAV target tracking benchmarks

指标上分别比基线算法高 1.7% 和 0.9%.

3) 全天跟踪性能. 本文提出的跟踪器能够进行全天目标跟踪, 为此, 在同时含有白天序列和夜间序列的 UAVTrack112 上进行验证. 由图 6(c) 可知, AMRCF 分别以 71.2% 的精确度和 48.4% 的成功率位居第 1. 在两项指标上, 比排名第 2 的跟踪器分别高 1.7% 和 1.0%, 且分别比基线算法高 3.1% 和 1.7%. 而在 UAV123 数据集上排名第 1 的 ECO_HC, 在 UAVTrack112 的精确度图上仅排名第 4, 比 AMRCF 低 2.6%.

4) 跟踪速度. 对无人机目标跟踪来说, 帧率是至关重要的评价指标, 它表征算法每秒内跟踪的序列帧数. 各算法在 5 个无人机基准数据集上的帧率如表 4 所示. 其中, 粗体、双下划线、单下划线分别表示排名前 3 的算法. 在夜间数据集 UAVDark135 上, AMRCF 对每帧图像都要进行光照增强处理, 会耗费大量时间, 故帧率较低, 约为 14 fps. 在其他 4 个基准数据集上的帧率都超过 25 fps, 并且在 5 个数据集上达到 24.6 fps 的平均帧率, 能够满足实时跟踪的需求.

3.3.2 与无人机目标跟踪算法的比较

AMRCF 是面向无人机目标的跟踪算法, 为此, 将 AMRCF 与 5 个先进的无人机目标跟踪算法在 5 个基准数据集上的性能进行比较, 结果如表 5 所示. 其中, 粗体、双下划线、单下划线分别表示排名前 3 的算法.

AMRCF 在 5 个基准数据集上, 有 9 组指标排在前 3 名, 且有 7 组指标数据是第 1 名, 与排名第

表 4 各跟踪算法在无人机目标跟踪基准上的帧率比较

Table 4 Comparison of frame rates of various tracking algorithms on UAV target tracking benchmarks

	DTB70	UAVDark135	UAVTrack112	VisDrone-SOT2018	UAV123	平均
MRCF	34.2662	25.9653	36.5963	33.3641	35.0628	33.0509
MSCF	24.4487	20.0218	23.1446	24.8189	23.3937	23.1655
BACF	34.6232	23.7905	42.4856	37.0889	39.5027	35.4982
DSST	59.5299	25.2672	95.0388	61.7457	84.6391	65.2441
ECO_HC	50.2676	46.2619	60.8356	58.2079	59.6936	55.0533
fDSST	137.9438	165.0210	178.2769	146.0930	164.2684	158.3206
KCF	**546.8368**	**291.7405**	**970.2931**	**619.1967**	**894.7070**	**664.5548**
SAMF	8.0580	5.5381	7.6382	7.5533	9.9559	7.7487
SRDCF	8.0273	5.5125	10.4828	8.0745	10.2923	8.4779
STRCF	20.4708	15.3309	21.9033	20.9129	21.2095	19.9655
AutoTrack	38.9684	33.1091	41.0309	41.8124	40.8652	39.1572
ARCF	21.8288	17.4036	22.7775	21.1463	22.4400	21.1192
DRCF	29.9372	22.0180	31.8315	30.2080	31.9774	29.1944
Staple	76.8257	59.0852	77.7228	90.2850	79.1470	76.6131
AMRCF	26.5606	14.5929	28.5730	25.7715	27.3098	24.5616

表 5 AMRCF 与无人机目标跟踪算法的性能比较

Table 5 Performance comparison of AMRCF and UAV target tracking algorithms

基准数据集	评价指标	MRCF	MSCF	AutoTrack	DRCF	ARCF	AMRCF
DTB70	精确度	0.666	0.649	0.716	0.636	0.694	**0.726**
	成功率	0.466	0.450	0.478	0.440	0.472	**0.488**
UAVDark135	精确度	0.593	0.600	0.599	0.481	0.584	**0.610**
	成功率	0.455	0.459	0.454	0.388	0.448	**0.464**
UAVTrack112	精确度	0.681	0.688	0.695	0.675	0.672	**0.712**
	成功率	0.467	0.468	0.464	0.458	0.457	**0.484**
UAV123	精确度	0.693	0.690	0.689	**0.700**	0.671	0.694
	成功率	**0.485**	0.483	0.472	0.482	0.468	0.478
VisDrone-SOT2018	精确度	0.812	0.776	0.788	0.782	0.797	**0.816**
	成功率	**0.600**	0.570	0.573	0.573	0.584	0.598

2 的算法相比, 领先 0.4% ~ 1.7%. 而基线算法 MRCF 有两组指标排在第 1, 与排名第 2 的算法相比, 仅有 0.2% 的提升. 在 UAV123 上, AMRCF 以 69.4% 的精确度排第 2, 比排名第 1 的 DRCF低 0.6%. 在 VisDrone-SOT2018 上, AMRCF 以 59.8% 的成功率排第 2, 仅比基线算法 MRCF 低 0.2%, 比排名第 3 的 ARCF 高 1.4%. 这表明 AMRCF 在无人机目标跟踪中的出色性能.

3.3.3 针对挑战属性的评估

为评估 AMRCF 在不同挑战属性上的表现, 本节将跟踪算法在 DTB70 数据集不同挑战属性上的性能进行了汇总, 如表 6 所示. 其中, 粗体、双下划线、单下划线分别表示排名前 3 的算法. SV、ARV、OCC、DEF、FCM、IR、OR、OV、BC 和 SOA 代表的挑战属性分别是尺度变化、纵横比变化、遮挡、形变、相机快速移动、平面内旋转、平面外旋转、移出视野、背景杂波和相似物体.

本文算法在 SV、ARV、DEF、IR、OR、BC 上获得最好的性能, 在精确度上分别比排名第 2 的跟踪器高 2.8%、5.3%、7.8%、1.2%、7.3%、0.5%, 在成功率上高 2.9%、3.7%、4.0%、1.2%、2.6%、0.1%. AMRCF 在 OV 上位居第 2 名, 在 OCC、FCM 和 SOA 属性上没有排进前 3 名, 但在评测算法中依然名列前茅. 由此表明, 本文算法能够很好地应对无人机目标跟踪中的各种挑战.

3.4 定性分析

3.4.1 跟踪成功案例分析

为更加直观地展示 AMRCF 相对于其他无人机目标跟踪算法的优势, 本节对以上 6 个无人机算法在具有平面内外旋转、遮挡、光照不足、形变等特定场景下的跟踪结果进行了可视化, 可视化结果如图 7 所示 (完整的视频在该站点中可以访问 https://gitee.com/he_bing_DLMZ/amrcf/tree/master/AMRCF/results/vis-video).

1) 平面内外旋转. 在视频序列 DTB70_BMX5 中, 相对于拍摄平面, 序列中的目标做极限动作时, 在第 47、77、158 和 200 帧图像中发生平面内外旋转. 其他几个算法在前几帧对目标进行了跟踪, 但在后续帧中跟踪框发生漂移而丢失目标. AMRCF 算法能够持续捕获目标, 未发生目标丢失情况. 在视频序列 UAV123_wakeboard5 中, 目标在第 343、581 和 643 帧中发生旋转, AMRCF 算法对目标进行了精准的定位. 这表明, 通过引入高斯形状掩膜并用 CF-VGG 提取目标的深度特征, 有效解决了因目标旋转而导致跟踪失败的问题.

2) 遮挡. 遮挡是无人机目标跟踪中占重要比例的挑战因素, 在视频序列 DTB70_BMX5 和 UAV-123_wakeboard5 中都存在遮挡现象. 在 DTB70_BMX5 的第 96 帧, 目标被部分遮挡时, 其他算法都丢失了目标, 而 AMRCF 和 DRCF 成功捕获目标. 不同的是, 本文算法对目标进行了精准的定位, 而 DRCF 边界框发生漂移, 导致仅捕获到部分目标. 在第 105 帧, 目标发生完全遮挡后重现时, DRCF 彻底失去目标, 而本文算法对目标进行了精准捕获. 在视频序列 UAV123_wakeboard5 的第 365 帧, 目标被海浪遮挡, 在第 373 帧再次出现时, 仅本文算法、AutoTrack 和 DRCF 定位到了目标. 引入高斯形状掩膜来抑制边界效应和防止样本污染, 使得本文算法的定位精度要高于 AutoTrack 和 DRCF.

3) 光照不足. UAVDark135_bike2 和 UAVTrack112_dark car1-n 是夜间拍摄的, 光照不足是它们共有的挑战. 在视频序列 UAVDark135_bike2 的前 30 帧, 6 个算法都对目标进行了精准定位, 但在第 30 帧后, 其他算法无法应对黑暗场景而纷纷丢失目标. 在第 88 帧后, 除 AMRCF 外, 其他 5 个算法彻底丢失目标. 在视频序列 UAVTrack112_dark car1-n 的前 197 帧, 测试算法都定位到目标, 但 AutoTrack 发生小幅度的漂移而未精准定位. 在第 227 帧到 259 帧, 目标从相对较亮的区域移至黑暗区域时, AutoTrack、ARCF 和 DRCF 丢失目标, AMRCF 引入的自适应图像增强模块, 可以在不改变图像各通道颜色比例的前提下对图像进行增强处理, 而未丢失目标, 且对目标进行精准定位.

4) 形变. 在视频序列 UAV123_wakeboard5 的跟踪过程中, 目标要通过不同的动作来维持自身平衡而发生大幅度形变, 致使目标特征变化较大. 在第 373 帧和 432 帧, 随着目标形态的变化, MRCF、MSCF、ARCF 算法相继失去对目标的跟踪能力, 且 AutoTrack 在目标尺度估计上出现较大偏差. 而本文算法使用 CF-VGG 来提取深度特征, 并结合 CN 等手工特征来描述目标的语义信息, 实现了对目标的鲁棒跟踪.

通过以上分析和可视化结果可以看出, 本文将 MRCF 与自适应图像增强模块、CF-VGG 和高斯形状掩膜结合而构建的 AMRCF, 可以轻松应对诸如平面内外旋转、遮挡、光照不足、形变等挑战.

3.4.2 跟踪失败案例分析

本文所提出的 AMRCF 跟踪器在 5 个无人机数据集上的总体跟踪效果优于当前其他先进的无人机目标跟踪算法, 但在无人机目标的跟踪过程中仍存在局限性. 如图 8 所示, 当场景中出现完全遮挡、相似物体、背景杂波以及低分辨率等因素时, 会导

表 6 各算法在 DTB70 基准不同属性上的性能比较

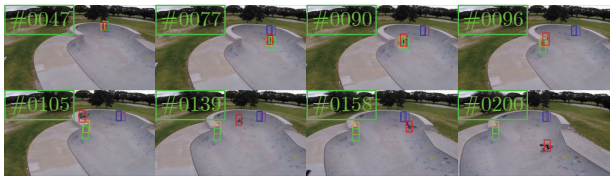
Table 6 Performance comparison of each algorithm on different attributes of the DTB70 benchmark

属性	评价指标	SRDCF	STRCF	Staple	SAMF	ECO_HC	DSST	fDSST
SV	精确度	0.462	0.568	0.489	0.456	0.538	0.473	0.543
	成功率	0.359	0.417	0.349	0.339	0.431	0.361	0.378
ARV	精确度	0.343	0.492	0.430	0.375	0.487	0.366	0.405
	成功率	0.268	0.347	0.314	0.289	0.361	0.290	0.275
OCC	精确度	0.478	0.617	0.528	0.500	**0.639**	0.423	0.467
	成功率	0.310	0.400	0.349	0.321	0.430	0.275	0.310
DEF	精确度	0.289	0.554	0.419	0.387	0.561	0.339	0.390
	成功率	0.208	0.390	0.283	0.293	0.391	0.261	0.243
FCM	精确度	0.554	0.713	0.494	0.499	0.678	0.485	0.587
	成功率	0.398	0.467	0.331	0.333	0.461	0.322	0.388
IR	精确度	0.419	0.586	0.457	0.409	0.548	0.400	0.489
	成功率	0.312	0.393	0.318	0.301	0.393	0.282	0.326
OR	精确度	0.220	0.385	0.371	0.225	0.390	0.237	0.226
	成功率	0.204	0.257	0.283	0.218	0.271	0.255	0.153
OV	精确度	0.552	0.652	0.420	0.537	0.515	0.474	0.564
	成功率	0.378	0.424	0.278	0.324	0.380	0.309	0.381
BC	精确度	0.393	0.611	0.393	0.352	0.553	0.312	0.356
	成功率	0.256	0.369	0.231	0.221	0.335	0.198	0.225
SOA	精确度	0.569	0.677	0.529	0.499	0.654	0.528	0.562
	成功率	0.379	0.447	0.346	0.319	0.445	0.334	0.351

属性	评价指标	BACF	ARCF	AutoTrack	DRCF	MSCF	MRCF	AMRCF
SV	精确度	0.545	0.707	0.688	0.574	0.610	0.613	**0.735**
	成功率	0.398	0.487	0.493	0.446	0.450	0.451	**0.522**
ARV	精确度	0.370	0.608	0.605	0.526	0.542	0.556	**0.661**
	成功率	0.274	0.393	0.405	0.377	0.377	0.384	**0.442**
OCC	精确度	0.556	0.638	0.631	0.585	0.585	0.587	0.624
	成功率	0.368	**0.446**	0.415	0.414	0.412	0.418	0.406
DEF	精确度	0.410	0.654	0.670	0.559	0.610	0.628	**0.748**
	成功率	0.282	0.426	0.452	0.392	0.419	0.416	**0.492**
FCM	精确度	0.640	0.742	**0.744**	0.699	0.694	0.719	0.738
	成功率	0.434	**0.496**	**0.496**	0.472	0.471	**0.496**	0.494
IR	精确度	0.537	0.636	0.684	0.594	0.603	0.615	**0.696**
	成功率	0.374	0.429	0.454	0.413	0.416	0.429	**0.466**
OR	精确度	0.270	0.451	0.439	0.346	0.385	0.412	**0.524**
	成功率	0.223	0.321	0.343	0.263	0.297	0.323	**0.369**
OV	精确度	0.575	0.642	**0.690**	0.611	0.629	0.675	0.680
	成功率	0.383	0.427	0.407	0.430	0.414	**0.448**	0.435
BC	精确度	0.491	0.585	0.635	0.594	0.610	0.616	**0.640**
	成功率	0.311	0.377	0.394	0.365	0.376	0.399	**0.400**
SOA	精确度	0.641	0.730	**0.731**	0.700	0.678	0.695	0.694
	成功率	0.421	**0.484**	0.473	0.453	0.452	0.471	0.447

致 AMRCF 目标框发生漂移甚至丢失目标.

如图 8 第 1 行所示, 在目标车辆未被遮挡前, AMRCF 能够对目标进行精确跟踪, 未产生漂移或目标丢失现象. 在第 818 帧, 车辆被树木遮挡致使提取的目标特征减少而产生漂移; 在第 920 帧, 目标被完全遮挡, 提取的深度特征及手工特征中不包



(a) DTB70_BMX5

(b) UAV123_wakeboard5

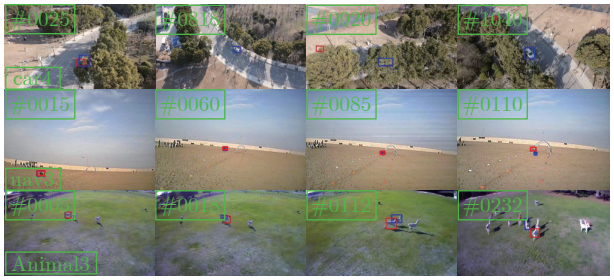
(c) UAVDark135_bike2

(d) UAVTrack112_dark car1-n

AMRCF　MRCF　MSCF
AutoTrack　ARCF　DRCF

图 7　6 个无人机目标跟踪算法的跟踪结果在部分序列上的可视化对比

Fig. 7　Visual comparison of tracking results of six UAV tracking algorithms on selected sequences



AMRCF　Ground truth

图 8　失败案例的可视化对比

Fig. 8　Visual comparison of failure cases

含目标信息, 目标丢失; 在第 1 040 帧, 目标在长期完全遮挡重现后, AMRCF 算法缺少重检测策略而未能重新捕获目标. 在第 2 行中, 目标的快速移动使其与拍摄设备之间的距离逐渐增大, 在视野中的分辨率变小并受到背景杂波的干扰, 提取的特征中目标特征占据极小的比重而无法继续定位并跟踪目标.

图 8 第 3 行是 AMRCF 算法跟踪相似物体的失败情形. 跟踪前期, 跟踪目标未与相似目标产生重叠, 目标框始终定位于跟踪目标而未产生漂移; 跟踪到第 18 帧时, 跟踪目标与其他目标重叠, 目标框发生漂移; 第 232 帧中, 两目标之间不再重叠, 跟踪框未重新找回目标, 定位于错误目标直至跟踪结束. 由于 AMRCF 没有重检测模块, 无法检测到目标何时消失及重现而导致完全遮挡挑战下的跟踪失败, 未来可以通过为算法添加重检测机制来进一步提高其性能.

4　结论

本文提出一种全天实时跟踪无人机目标的多正则化相关滤波算法 AMRCF, 在不改变图像各通道颜色比例的前提下, 使用自适应图像增强模块对夜间图像进行光照增强处理, 以应对夜间无人机目标跟踪的挑战; 同时, 在保证实时性的前提下, 引入一个轻量型深度网络 CF-VGG 来提取目标的深度特征, 联合手工特征一起描述目标的语义信息, 提高跟踪器的鲁棒性; 通过嵌入高斯形状掩膜, 在抑制边界效应的同时, 解决训练样本污染问题, 提高滤波器的训练质量. 在 5 个公开的无人机基准数据集上的实验结果表明, 与先进的无人机目标跟踪算法相比, AMRCF 算法表现出较强的竞争力, 并且具有较好的实时性. 在下一步工作中, 将考虑在本文算法框架中加入重检测模块, 进一步提升算法的性能.

References

1　Zhang F, Ma S P, Yu L X, Zhang Y L, Qiu Z L, Li Z Y. Learning future-aware correlation filters for efficient UAV tracking. *Remote Sensing*, 2021, **13**(20): Article No. 4111

2　Meng Lu, Yang Xu. A survey of object tracking algorithms. *Acta Automatica Sinica*, 2019, **45**(7): 1244−1260
(孟琭, 杨旭. 目标跟踪算法综述. 自动化学报, 2019, **45**(7): 1244−1260)

3　Xu L B, Li Q L, Jiang J, Zou G F, Liu Z, Gao M L. Adaptive spatio-temporal regularized correlation filters for UAV-based tracking. In: Proceedings of the 15th Asian Conference on Computer Vision (ACCV). Kyoto, Japan: Springer, 2020. 391−404

4　Henriques J F, Caseiro R, Martins P, Batista J. High-speed tracking with kernelized correlation filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, **37**(3): 583−596

5　Danelljan M, Hager G, Khan F S, Felsberg M. Learning spatially regularized correlation filters for visual tracking. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV). Santiago, Chile: IEEE, 2015. 4310−4318

6　Xiong Dan, Lu Hui-Min, Xiao Jun-Hao, Zheng Zhi-Qiang. Robust long-term object tracking with adaptive scale and rotation estimation. *Acta Automatica Sinica*, 2019, **45**(2): 289−304
(熊丹, 卢惠民, 肖军浩, 郑志强. 具有尺度和旋转适应性的长时间目标跟踪. 自动化学报, 2019, **45**(2): 289−304)

7　Zhang Wei-Jun, Zhong Sheng, Xu Wen-Hui, Wu Ying. Correlation filter based visual tracking integrating saliency and motion cues. *Acta Automatica Sinica*, 2021, **47**(7): 1572−1588
(张伟俊, 钟胜, 徐文辉, WU Ying. 融合显著性与运动信息的相关

滤波跟踪算法. 自动化学报, 2021, **47**(7): 1572−1588)

8 Galoogahi H K, Fagg A, Lucey S. Learning background-aware correlation filters for visual tracking. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017. 1144−1152

9 Wu X D, Xu J, Zhu Z Y, Wang Y N, Zhang Q, Tang S M, et al. Correlation filter tracking algorithm based on spatial-temporal regularization and context awareness. *Applied Intelligence*, 2022, **52**: 17772−17783

10 Pu Lei, Feng Xin-Xi, Hou Zhi-Qiang, Zha Yu-Fei, Yu Wang-Sheng. Correlation filter tracking based on deep spatial regularization. *Acta Electronica Sinica*, 2020, **48**(10): 2025−2032
(蒲磊, 冯新喜, 侯志强, 查宇飞, 余旺盛. 基于深度空间正则化的相关滤波跟踪算法. 电子学报, 2020, **48**(10): 2025−2032)

11 Wang F S, Yin S S, Mbelwa J T, Sun F M. Context and saliency aware correlation filter for visual tracking. *Multimedia Tools and Applications*, 2022, **81**: 27879−27893

12 Yuan D, Chang X J, Li Z H, He Z Y. Learning adaptive spatial-temporal context-aware correlation filters for UAV tracking. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 2022, **18**(3): Article No. 70

13 Wang L, Li J A, Huang B, Chen J J, Li X M, Wang J H, et al. Auto-perceiving correlation filter for UAV tracking. *IEEE Transactions on Circuits and Systems for Video Technology*, 2022, **32**(9): 5748−5761

14 Ding Xin-Yao, Zhang Xin. Visual tracking with salient features and selective mechanism. *Acta Electronica Sinica*, 2020, **48**(1): 118−123
(丁新尧, 张鑫. 基于显著性特征的选择性目标跟踪算法. 电子学报, 2020, **48**(1): 118−123)

15 Bolme D S, Beveridge J R, Draper B A, Lui Y M. Visual object tracking using adaptive correlation filters. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Francisco, USA: IEEE, 2010. 2544−2550

16 Wang Y B, Wang F S, Wang C, Sun F M, He J J. Learning saliency-aware correlation filters for visual tracking. *The Computer Journal*, 2022, **65**(7): 1846−1859

17 Pu Lei, Feng Xin-Xi, Hou Zhi-Qiang, Yu Wang-Sheng. Robust visual tracking based on second order pooling network. *Acta Electronica Sinica*, 2020, **48**(8): 1472−1478
(蒲磊, 冯新喜, 侯志强, 余旺盛. 基于二阶池化网络的鲁棒视觉跟踪算法. 电子学报, 2020, **48**(8): 1472−1478)

18 Ma C, Huang J B, Yang X K, Yang M H. Hierarchical convolutional features for visual tracking. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV). Santiago, Chile: IEEE, 2015. 3074−3082

19 Javed S, Danelljan M, Khan F S, Khan M H, Felsberg M, Matas J. Visual object tracking with discriminative filters and Siamese networks: A survey and outlook. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, **45**(5): 6552−6574

20 Zheng G Z, Fu C H, Ye J J, Lin F L, Ding F Q. Mutation sensitive correlation filter for real-time UAV tracking with adaptive hybrid label. In: Proceedings of the IEEE International Conference on Robotics and Automation (ICRA). Xi′an, China: IEEE, 2021. 503−509

21 Ye J J, Fu C H, Lin F L, Ding F Q, An S, Lu G. Multi-regularized correlation filter for UAV tracking and self-localization. *IEEE Transactions on Industrial Electronics*, 2022, **69**(6): 6004−6014

22 Li Y M, Fu C H, Ding F Q, Huang Z Y, Lu G. AutoTrack: Towards high-performance visual tracking for UAV with automatic spatio-temporal regularization. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE, 2020. 11920−11929

23 Li S Y, Yeung D Y. Visual object tracking for unmanned aerial vehicles: A benchmark and new motion models. In: Proceedings of the Thirty-first AAAI Conference on Artificial Intelligence. San Francisco, California, USA: AAAI, 2017. 4140−4146

24 Wang N, Zhou W G, Song Y B, Ma C, Li H Q. Real-time correlation tracking via joint model compression and transfer. *IEEE Transactions on Image Processing*, 2020, **29**: 6123−6135

25 Fu C H, Cao Z A, Li Y M, Ye J J, Feng C. Onboard real-time aerial tracking with efficient Siamese anchor proposal network. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, **60**: Article No. 5606913

26 Li B W, Fu C H, Ding F Q, Ye J J, Lin F L. All-day object tracking for unmanned aerial vehicle. *IEEE Transactions on Mobile Computing*, 2023, **22**(8): 4515−4529

27 Wen L Y, Zhu P F, Du D W, Bian X, Ling H B, Hu Q H, et al. VisDrone-SOT2018: The vision meets drone single-object tracking challenge results. In: Proceedings of the European Conference on Computer Vision Workshops. Munich, Germany: Springer, 2018. 469−495

28 Mueller M, Smith N, Ghanem B. A benchmark and simulator for UAV tracking. In: Proceedings of the European Conference on Computer Vision. Amsterdam, The Netherlands: Springer, 2016. 445−461

29 Henriques J F, Caseiro R, Martins P, Batista J. Exploiting the circulant structure of tracking-by-detection with kernels. In: Proceedings of the European Conference on Computer Vision. Florence, Italy: Springer, 2012. 702−715

30 Danelljan M, Hager G, Khan F S, Felsberg M. Discriminative scale space tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, **39**(8): 1561−1575

31 Wang C, Zhang L, Xie L H, Yuan J S. Kernel cross-correlator. In: Proceedings of the Thirty-second AAAI Conference on Artificial Intelligence. New Orleans, USA: AAAI, 2018. 4179−4186

32 Zhang Y, Liu G X, Zhang H Y, Huang H L. Robust visual tracker combining temporal consistent constraint and adaptive spatial regularization. *Neural Computing and Applications*, 2021, **33**(14): 8355−8374

33 Deng C W, He S C, Han Y Q, Zhao B Y. Learning dynamic spatial-temporal regularization for UAV object tracking. *IEEE Signal Processing Letters*, 2021, **28**: 1230−1234

34 Li F, Tian C, Zuo W M, Zhang L, Yang M H. Learning spatial-temporal regularized correlation filters for visual tracking. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City, USA: IEEE, 2018. 4904−4913

35 Xu T Y, Feng Z H, Wu X J, Kittler J. Joint group feature selection and discriminative filter learning for robust visual object tracking. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, South Korea: IEEE, 2019. 7950−7960

36 Feng S, Hu K L, Fan E, Zhao L P, Wu C D. Kalman filter for spatial-temporal regularized correlation filters. *IEEE Transactions on Image Processing*, 2021, **30**: 3263−3278

37 Danelljan M, Bhat G, Khan F S, Felsberg M. ECO: Efficient convolution operators for tracking. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Hawaii, USA: IEEE, 2017. 6931−6939

38 Zhao J W, Wei F Y, Chen N N, Zhou Z H. Spatial and long-short temporal attention correlation filters for visual tracking. *IET Image Processing*, 2022, **16**(11): 3011−3024

39 Hua X, Wang X Q, Rui T, Shao F M, Wang D. Light-weight UAV object tracking network based on strategy gradient and attention mechanism. *Knowledge-Based Systems*, 2021, **224**: Article No. 107071

40 Zhang Z P, Liu Y F, Li B, Hu W M, Peng H W. Toward accurate pixelwise object tracking via attention retrieval. *IEEE Transactions on Image Processing*, 2021, **30**: 8553−8566

41 Xu T Y, Feng Z H, Wu X J, Kittler J. Adaptive channel selection for robust visual object tracking with discriminative correlation filters. *International Journal of Computer Vision*, 2021, **129**:

1359−1375

42 Zhu X F, Wu X J, Xu T Y, Feng Z H, Kittler J. Robust visual object tracking via adaptive attribute-aware discriminative correlation filters. *IEEE Transactions on Multimedia*, 2022, **24**: 301−312

43 Wang Ke-Ping, Zhu Peng-Fei, Yang Yi, Fei Shu-Min. Target tracking algorithm based on multi-time-space perception correlation filters fusion. *Journal of Computer-Aided Design & Computer Graphics*, 2020, **32**(11): 1840−1852
(王科平, 朱朋飞, 杨艺, 费树岷. 多时空感知相关滤波融合的目标跟踪算法. 计算机辅助设计与图形学学报, 2020, **32**(11): 1840−1852)

44 Li F, Wu X H, Zuo W M, Zhang D, Zhang L. Remove cosine window from correlation filter-based visual trackers: When and how. *IEEE Transactions on Image Processing*, 2020, **29**: 7045−7060

45 Tang M, Zheng L Y, Yu B, Wang J Q. Fast kernelized correlation filter without boundary effect. In: Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV). Waikoloa, HI, USA: IEEE, 2021. 2998−3007

46 Fu C H, Xu J T, Lin F L, Guo F Y, Liu T C, Zhang Z J. Object saliency-aware dual regularized correlation filter for real-time aerial tracking. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, **58**(12): 8940−8951

47 Hinton G, Vinyals O, Dean J. Distilling the knowledge in a neural network. arXiv preprint arXiv: 1503.02531, 2015.

48 Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, et al. ImageNet large scale visual recognition challenge. *International Journal of Computer Vision*, 2015, **115**: 211−252

49 Boyd S, Parikh N, Chu E, Peleato B, Eckstein J. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 2011, **3**(1): 1−122

50 Huang Z Y, Fu C H, Li Y M, Lin F L, Lu P. Learning aberrance repressed correlation filters for real-time UAV tracking. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, South Korea: IEEE, 2019. 2891−2900

51 Bertinetto L, Valmadre J, Golodetz S, Miksik O, Torr P H S. Staple: Complementary learners for real-time tracking. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA: IEEE, 2016. 1401−1409

52 Li Y, Zhu J K. A scale adaptive kernel correlation filter tracker with feature integration. In: Proceedings of the European Conference on Computer Vision. Zurich, Switzerlan: Springer, 2014. 254−265

53 Wu Y, Lim J, Yang M H. Online object tracking: A benchmark. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Portland, USA: IEEE, 2013. 2411−2418

王法胜 大连民族大学信息与通信工程学院教授. 主要研究方向为计算机视觉, 模式识别.
E-mail: wangfasheng@dlnu.edu.cn
(**WANG Fa-Sheng** Professor at the School of Information and Communication Engineering, Dalian Minzu University. His research interest covers computer vision and pattern recognition.)

李 富 大连民族大学信息与通信工程学院硕士研究生. 主要研究方向为计算机视觉.
E-mail: fuliytu@163.com
(**LI Fu** Master student at the School of Information and Communication Engineering, Dalian Minzu University. His main research interest is computer vision.)

尹双双 大连民族大学信息与通信工程学院硕士研究生. 主要研究方向为计算机视觉.
E-mail: yss460229828@163.com
(**YIN Shuang-Shuang** Master student at the School of Information and Communication Engineering, Dalian Minzu University. Her main research interest is computer vision.)

王 星 大连民族大学信息与通信工程学院硕士研究生. 主要研究方向为计算机视觉.
E-mail: dlnuwangxing@gmail.com
(**WANG Xing** Master student at the School of Information and Communication Engineering, Dalian Minzu University. His main research interest is computer vision.)

孙福明 大连民族大学信息与通信工程学院教授. 主要研究方向为计算机视觉, 多媒体技术. 本文通信作者.
E-mail: sunfuming@dlnu.edu.cn
(**SUN Fu-Ming** Professor at the School of Information and Communication Engineering, Dalian Minzu University. His research interest covers computer vision and multimedia technology. Corresponding author of this paper.)

朱 兵 哈尔滨工业大学电子与信息工程学院副研究员. 主要研究方向为图像处理, 模式识别.
E-mail: zhubing@hit.edu.cn
(**ZHU Bing** Associate research fellow at the School of Electronics and Information Engineering, Harbin Institute of Technology. His research interest covers image processing and pattern recognition.)