

基于相关滤波的红外目标跟踪抗遮挡处理

张晋, 王元余, 林丹丹, 杜欣悦, 林宇, 兰戈

(昆明物理研究所, 云南昆明 650223)

摘要: 针对传统的相关滤波算法在红外目标跟踪过程中, 目标被完全遮挡后跟踪失效的问题, 提出一种结合了多尺度滤波跟踪器和基于深度学习检测器的目标实时跟踪抗遮挡算法。首先使用跟踪器跟踪目标, 计算目标的峰值响应强度并比较峰值响应强度与经验阈值的大小以判断目标是否被遮挡或跟踪丢失。然后当目标被遮挡或跟踪丢失时, 停止更新跟踪器, 由于目标被遮挡后再次进入画面的位置可能会发生巨大变化, 并且跟踪器在整个图像中搜索目标的速度特别慢, 因此在不降低跟踪精度和速度的情况下, 后续帧中采用了检测器检测目标并得到多个目标框。利用检测器得到的目标框, 分别利用跟踪器进行相关滤波, 针对每个目标框得到一个峰值响应强度, 其中峰值响应强度最大且超过经验阈值的目标框即为重新进入画面的目标。通过与多尺度相关滤波算法比较, 所提算法在满足实时跟踪的情况下, 能有效地解决红外目标被遮挡的问题, 具有更高的鲁棒性和精确度。

关键词: 红外目标跟踪; 相关滤波; 多尺度滤波; 深度学习; 遮挡处理; 实时跟踪

中图分类号: TN219 **文献标识码:** A **文章编号:** 1001-8891(2022)03-0277-09

Anti-occlusion Process of Infrared Target Tracking Based on Correlation Filters

ZHANG Jin, WANG Yuanyu, LIN Dandan, DU Xinyue, LIN Yu, LAN Ge

(Kunming Institute of Physics, Kunming 650223, China)

Abstract: Focusing on the issue that traditional correlation filters have poor performance in infrared target tracking with occlusion, an anti-occlusion and real-time target-tracking algorithm based on a multi-scale filter tracker and a deep learning detector is proposed. First, the peak response value is calculated using the tracker; if the peak value is less than the threshold, the target is occluded or tracking is lost. Second, the detector stops updating when the target is occluded or tracking is lost. The position of the target changes significantly when it comes in frame again after occlusion, and the speed of target searching with the tracker will be very slow. At this time, a detector is employed to detect the targets in the subsequent frames without loss of tracking accuracy and speed. The peak values are calculated for each target box that is detected by the detector, and the target with a maximum peak value larger than the threshold is tracked. The results of the experiment compared with the multi-scale correlation filter show that the proposed real-time tracking algorithm can not only effectively solve infrared target occlusion, but also has higher tracking robustness and accuracy.

Key words: infrared target tracking, deep learning, correlation filters, multi-scale filter, occlusion process, real-time tracking

0 引言

红外目标跟踪是指对红外视频中的目标进行检测和跟踪的技术, 由于红外具有全天候监控的特征, 其广泛应用于国防军事和民用方面, 在国防军事上,

如地海空目标监控和预警系统以及精确制导等, 在民用上, 如安防监控、人机交互和无人机对地面目标的跟踪等^[1-4]。目标跟踪技术近年来受到了人们的广泛关注, 也是当今计算机视觉领域最具挑战性的研究领域之一。在实际应用中, 由于实际场景复杂, 视频中目

收稿日期: 2021-12-13; 修订日期: 2022-01-28.

作者简介: 张晋 (1986-), 男, 云南保山人, 硕士研究生, 主要研究方向为视频目标识别与跟踪。E-mail: zhangjin_211@163.com.

通信作者: 王元余 (1989-), 男, 湖南衡阳人, 高级工程师, 主要从事红外图像处理的研究。E-mail: wxyjin232425@163.com.

标的形变、光照变化、旋转、快速运动和遮挡等因素都会影响跟踪的准确度，提出一种实时、鲁棒性强、跟踪精度高的跟踪算法仍具有较大的难度，为解决上述因素的影响，近年来目标跟踪领域也出现了许多算法。目标跟踪算法根据表观模型可以分为两类：生成式和判别式模型。其中生成式模型主要是在视频图像中寻找与给定目标最相似的图像窗口，而判别式模型则是通过学习分类器来对图像进行目标和背景的二分类。

近年来在判别式跟踪算法中，由于相关滤波的跟踪算法具有高速和高效的特点，得到了人们的广泛关注和迅速的发展^[5]。相关滤波算法最初用于信号处理领域中比较两个信号的相似度，随后 Bolme 等人^[6]首次基于相关滤波提出了最小误差输出平方和跟踪算法 (minimum output sum of squared error filter, MOSSE)，利用傅里叶变换将时域上的卷积转换为频域上的点积，极大地减少了计算量，提高了目标跟踪速度，跟踪速度超过了 600 fps，完全满足了实际应用中目标跟踪的实时性要求。由于该算法在目标跟踪上的优异表现，此后众多基于此算法框架的相关滤波改进算法相继被提出，其中最具代表性的有 Henriques 等人^[7-8]提出的循环结构跟踪算法 (circulant structure Kernel, CSK) 以及在此基础上用方向梯度直方图 (histogram of oriented gradient, HOG) 特征替代 CSK 算法中灰度特征的核相关滤波跟踪算法 (Kernel correlation filter, KCF)，极大地提高了跟踪准确度。随后 Danelljan 等人^[9-10]基于循环结构先后提出了判别尺度空间跟踪算法 (discriminative scale space tracker, DSST) 和快速判别尺度空间跟踪算法 (fast discriminative scale space tracker, fDSST)，能在快速跟踪目标的同时进行尺度更新，较好地解决了目标跟踪尺度变化的问题。

然而目标遮挡作为目标跟踪领域一种常见的情况，相关滤波跟踪算法无法有效的解决。由于相关滤波大多采用循环矩阵的方式，如果采用全图像搜索目标，则会极大地降低跟踪速度，无法满足实际应用的实时性要求，因此相关滤波的大部分算法均采用在目标附近的图像窗口内进行计算响应强度，但是在实际目标跟踪过程中，当目标被遮挡或出画面后再次进入画面时，目标很可能不会出现在之前的位置，这将会导致相关滤波跟踪算法无法有效地重新跟踪上目标。

针对以上情况，本文结合快速判别尺度空间滤波跟踪算法^[10]和深度学习中的 YOLO (you only look once) 目标识别算法^[11]进行改进并提出 YOLO-fDSST 跟踪算法 (YfDSST)，当目标跟踪丢失后，先采用

YOLO 算法快速检测出图像中的目标，随后利用 fDSST 算法在各检测出来的目标框内进行峰值响应强度计算，得到包含最大峰值响应强度且峰值大于经验阈值的目标框即为待跟踪的目标，利用该目标更新相关滤波模型并进行后续视频帧的跟踪。对比实验结果表明，本文改进的算法在满足实时性的情况下，能有效解决跟踪过程中目标被遮挡和出画面的情况，具有更高的鲁棒性和精确度。

1 目标跟踪器

本文所提的跟踪器算法由跟踪模块和目标再检测模块构成，其中跟踪模块使用快速判别尺度空间滤波算法，其引入了融合方向梯度直方图 (fusion of HOG, FHOG) 特征^[12]的，再检测模块使用基于深度学习的 YOLO 算法实现对目标位置和大小快速检测。

1.1 快速判别尺度空间滤波器

fDSST 采用了两个独立的相关滤波器，分别实现目标的位置跟踪和尺度变换，并分别定义为位置滤波器和尺度滤波器。在连续的两帧视频图像中，通常位置的变化大于尺度的变化，因此该算法先采用位置滤波器确定目标位置信息，在目标位置的基础上再使用尺度滤波器确定其尺度信息。

在位置滤波过程中，提取目标所在图像块的 FHOG 特征，并通过式(1)构造最优滤波器：

$$E = \left\| \sum_{l=1}^d \mathbf{h}^l \otimes \mathbf{f}^l - \mathbf{g} \right\|^2 + \lambda \sum_{l=1}^d \|\mathbf{h}^l\|^2 \quad (1)$$

式中： \mathbf{h}^l 、 \mathbf{f}^l 、 \mathbf{g} 均为 $M \times N$ 的矩阵，其中 M 和 N 分别为图像块的高度和宽度大小， l 表示特征维数且 $l \in \{1, 2, \dots, d\}$ ； λ 为正则项系数； \otimes 代表循环相关。

由于循环相关的矩阵点积运算的计算量巨大，利用傅里叶变换可以将时域内的矩阵卷积运算转换为频域上的点积运算，从而极大地减少了计算量，加快了计算速度^[7]。傅里叶变换后的最优相关滤波器如式(2)所示：

$$H^l = \overline{GF^l} / \left(\sum_{k=1}^d \overline{F^k F^k} + \lambda \right) \quad (2)$$

式中：大写字母代表相应的离散傅里叶变换，上划线表示共轭，特征维数取值范围 $k \in \{1, 2, \dots, d\}$ 。

在实际工程应用中，通常把相关滤波器 H^l 拆分成分子 A_l 和分母 B_l 的形式分别进行迭代更新操作，如式(3)所示：

$$\begin{cases} A_i^l = (1-\eta)A_{i-1}^l + \eta \overline{G}_i^l F_i^l \\ B_i = (1-\eta)B_{i-1} + \eta \sum_{k=1}^d \overline{F}_i^k F_i^k \end{cases} \quad (3)$$

式中: η 表示学习率。

对于新的一帧视频图像, 利用相关滤波器和高斯函数构建相应的图像块的输出响应得分 y , 其中输出响应得分最高的地方则为当前视频帧的目标位置, 如式(4)所示:

$$y = F^{-1} \left\{ \sum_{l=1}^d \overline{A}^l Z^l / (B + \lambda) \right\} \quad (4)$$

在尺度滤波过程中, 为了确定视频帧中目标的最佳尺度大小, 跟踪算法采用了一个一维的相关滤波器去估计目标在图像中的尺度, 假设在当前帧中目标大小为 $P \times Q$, 尺度大小设置为 S , 在目标中心位置提取尺度等级为 n 的训练样本图像块 J_n , 并提取特征 f 作为图像块的 d 维特征描述子, 其中图像块的大小为 $a_n P \times a_n R$, a 表示尺度因子, 尺度等级 n 的选取范围为 $n \in \{[-(S-1)/2], \dots, [(S-1)/2]\}$ 。

1.2 目标再检测

相关滤波器进行目标跟踪的原理是以视频图像与滤波器模板作用后的输出响应最大的地方作为目标的位置, 但是图像在进行每一次滤波的位置是上一帧中目标的位置, 如果目标被遮挡或目标出画面导致目标跟踪丢失, 当目标再次进入画面时, 目标可能不会出现在丢失前的位置, 而滤波器模板的滤波位置仍然停留在丢失前的位置, 此时将无法跟踪重新进入画面的目标。由于相关滤波器一般会采用循环矩阵的方式来增加训练样本数量, 如果使相关滤波器进行全图像搜索, 则会导致计算量急剧上升使得跟踪速度无法满足实时性要求。为了能够在跟踪目标丢失后, 能够重新找到目标并持续对其进行跟踪, 有必要引入再检测机制。不同于传统的目标检测中针对每一帧视频都进行目标检测, 在本文所提的再检测机制中, 只有当图像块中目标的峰值响应强度低于经验阈值后, 即目标跟踪丢失后才进行目标再检测, 从而不会显著地影响跟踪速度。

由于目标再检测算法的准确度和速度会对目标跟踪准确度和速度造成较大影响, 因此对目标再检测算法提出了较高的要求。本文的再检测机制采用的基于深度学习的 YOLOv4 算法^[11]对目标进行检测, 在目标检测的准确性和速度上均能较好的满足要求。

YOLOv4 是 YOLO 系列最新的网络, 其骨干网络

主要 CSPDarknet53 构成, 相对于其它骨干网络, CSPDarknet53 具有参数量少和传输速度快的优势, 并且在 CSPDarknet53 网络中采用了空间金字塔池化模块, 能有效地改善感受野尺寸, 在不降低网络处理速度的情况下将图像中最重要的上下位特征提取出来。同时 YOLOv4 网络还利用路径聚合网络 (path aggregation network, PAN) 对特征进行多通道融合, PAN 网络的应用可以实现特征的重用、提高检测准确率以及降低计算瓶颈和减少内存开销。

网络训练过程中的损失函数如式(5)所示:

$$\begin{cases} \text{Loss} = L_{\text{cywh}} + L_{\text{confidence}} + L_{\text{classes}} \\ L_{\text{cywh}} = \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} \left[(x_i - x_i')^2 + (y_i - y_i')^2 \right] \\ \quad + \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} \left[(\sqrt{w_i} - \sqrt{w_i'})^2 + (\sqrt{h_i} - \sqrt{h_i'})^2 \right] \\ L_{\text{confidence}} = -\lambda_{\text{obj}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} \left[c_i' \log(c_i) + (1-c_i') \log(1-c_i) \right] \\ \quad - \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{noobj}} \left[c_i' \log(c_i) + (1-c_i') \log(1-c_i) \right] \\ L_{\text{classes}} = \sum_{i=0}^{S^2} I_{ij}^{\text{obj}} \sum_{c \in \text{classes}} \left[p_i'(c) \log(p_i(c)) + (1-p_i'(c)) \log(1-p_i(c)) \right] \end{cases} \quad (5)$$

式中: L_{cywh} 为目标预测框与真实框的中心点和宽高误差之和; λ_{coord} 为坐标系数; $L_{\text{confidence}}$ 为目标置信度误差, 包含了格子中有物体和无物体的两项置信度误差; λ_{obj} 和 λ_{noobj} 分别为格子中有物体和无物体的置信度系数; L_{classes} 为目标分类损失; I_{objij} 表示为第 i 个格子的第 j 个锚框的匹配情况; c_i 表示第 i 个预测框的置信度; p_i 表示第 i 个预测框包含某类目标的概率。

在检测过程中, 该算法将图片划分成 $S \times S$ 个网格, 将每个格子作为先验锚框位置, 并在格子内对物体置信度、类别置信度和位置偏移量进行拟合, 最终经过非极大值抑制后得到目标的边界框坐标和类别信息。

1.3 跟踪流程

本文算法是在 fDSST 算法^[10]的基础上引入目标再检测机制来实现目标跟踪的抗遮挡处理, 首先利用 fDSST 对目标进行跟踪, 当目标被遮挡或出画面时, 当前视频帧中目标的峰值响应强度会低于经验阈值, 此时停止相关滤波器模型更新并会启用目标再检测机制。

本文算法针对场景中存在一个或多个目标的情形均是适用的, 如果场景中存在多个目标, 目标再检

测机制会在目标跟踪丢失后,在后续视频帧中检测出所有目标,从而得到所有目标的位置和大小信息,但是待跟踪的目标可能仅仅是检测出的所有目标中的一个,因此必须从所有被检测出来的目标中将待跟踪的目标筛选出来,在实际光电系统应用中,会进一步结合伺服系统返回的跟踪目标速度大小和方向信息,排除相似干扰目标的影响,提高目标重捕获的成功率。

针对检测出来的每一个目标,使用相关滤波器计算目标的峰值响应强度,选取峰值响应强度最大且大于经验阈值的目标即为待跟踪目标。如果最大峰值响应强度小于经验阈值,则说明在检测出来的目标中可能不包含待跟踪目标,则需要依据后续视频帧的检测结果继续进行筛选。

在筛选出待跟踪目标后,以式(3)对相关滤波器模型进行更新,并以经过相关滤波器筛选后的目标峰值响应强度位置初始化跟踪目标位置,考虑到目标在被遮挡或出画面后重新进入画面时,目标的尺寸大小可能会发生变化,因此需要依据检测目标的尺寸大小对跟踪目标进行尺度更新,更新方式如式(6)所示:

$$\begin{cases} w_t = (1 - \beta)w_{t-1} + \beta w_b \\ h_t = (1 - \beta)h_{t-1} + \beta h_b \end{cases} \quad (6)$$

式中: β 为尺度更新的学习率; w_b 和 h_b 分别为检测器检测到的目标宽高大小; w_{t-1} 和 h_{t-1} 分别为前一帧中目标的宽高大小。在得到重新进入画面的目标后继续以 fDSST 算法对目标进行跟踪从而实现目标的抗遮挡处理。图 1 为基于相关滤波的目标跟踪抗遮挡处理的流程图。

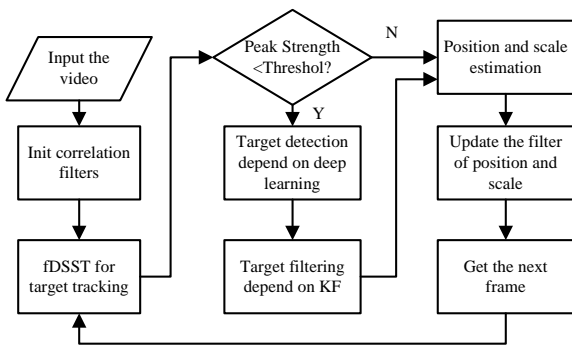


图 1 基于相关滤波的目标跟踪抗遮挡处理流程图

Fig.1 The schematic illustration of anti-occlusion process of target tracking based on correlation filters

2 实验结果与分析

实验中目标检测器的训练数据采用红外仿真软

件获得超过 7 万张的红外图像,图像包含坦克、悍马、直升机、舰艇、战斗机、卡车和大巴车等在内的多个目标类别,人工标注后将数据按 9:1 的比例分成训练集和测试集两部分,并采用变学习率的方式在训练服务器上训练 10 万步后得到目标检测器模型。

目标跟踪测试视频采用红外仿真软件获得两段目标被完全遮挡又重新进入画面的视频并进行人工标注,视频帧数超过 500 帧,帧频为 30 Hz,宽高为 640×512 像素。并在移植该跟踪算法到实际光电系统后,针对实际场景进行了两次在线测试,在线测试的视频采集帧率为 40 Hz,图像大小为 640×512 像素。

2.1 实验环境与参数

考虑到实际嵌入式系统应用,测试采用 C++ 语言,所用硬件采用 Nvidia Jetson TX2 核心板, GPU 为 Pascal 包含 256 颗 CUDA 核心, CPU 为 HMP Dual Denver 2 和 Quad ARM A57。显存和内存共享大小为 8 GB 的主存。相关滤波器学习率 η 为 0.01,尺度更新学习率 β 为 0.1,尺度因子 a 为 1.04,峰值响应强度阈值为 0.6。检测器训练时学习率采用变学习率策略,初始学习率为 0.00025,在 4 万步和 7 万步时学习率缩小 10 倍。经过 10 万步迭代后损失降低到 0.5 左右并保持基本稳定。检测器的测试结果如表 1 所示,测试时检测阈值为 0.5,从表 1 可以看出对所有的目标类型检测准确率和召回率均超过 97%,在该跟踪算法中,检测器的检测阈值采用 0.3,以提高目标检测召回率,从而提高跟踪目标丢失重捕获的成功率。

表 1 检测器测试结果

Table 1 The test results of detector

Target	Image number	Precision/%	Recall/%
Fighter	864	100	99.3
Helicopter	864	99.5	99.7
Patrol	864	100	99.8
Freighter	864	99.6	100
Bus	560	99.4	98.7
Hummer	560	99.3	98.7
Tank	560	97.7	97.3
Truck	560	99.6	98.7

2.2 性能评估方法

为评估本文所提算法在目标跟踪抗遮挡处理的性能,实验采用了距离精度 (Distance Precision, DP) 和成功率 (success rate, SR) 作为评估标准^[13]。距离精度是在跟踪过程中,跟踪算法所跟踪的目标中心位置与人工标注的目标中心位置的欧氏距离小于某一距

离阈值的视频帧数占整个视频中包含目标的视频帧数(剔除目标遮挡的帧数)的百分比。如果在相同的距离阈值下,距离精度越大则算法的跟踪效果越好。成功率则是指跟踪算法所跟踪的目标矩形框与人工标注的目标矩形框的交并比(intersection over union, IoU)大于某一IoU阈值的视频帧数占整个视频中包含目标的视频帧数的百分比。如果在相同的IoU阈值

下,成功率越高则算法的跟踪效果越好。

2.3 抗遮挡处理性能比较

2.3.1 定性比较

图2和图3分别为在红外中波和长波下不同算法的跟踪结果,下面结合图2和图3对本文所提的算法进行分析。

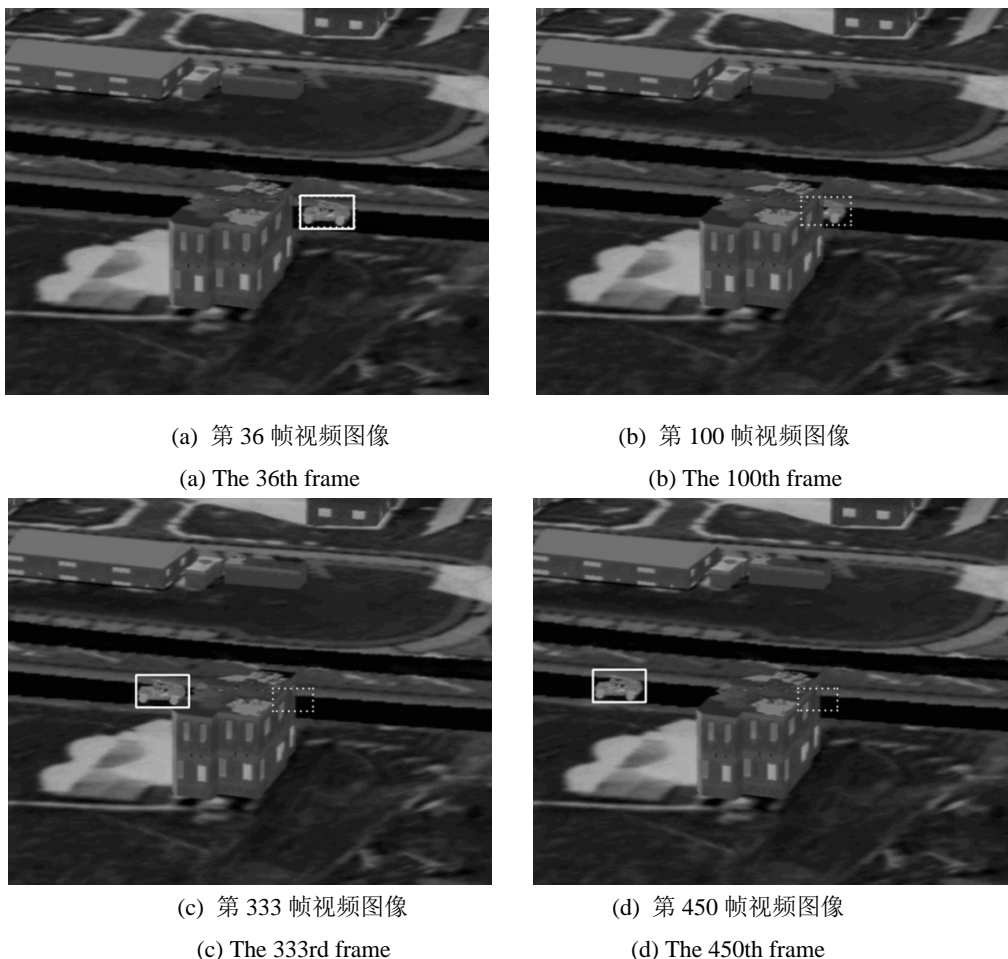
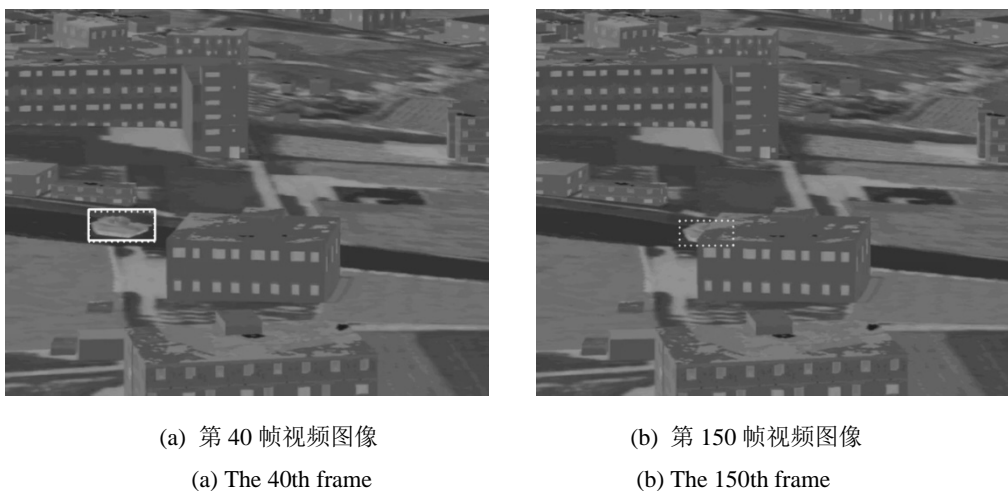


图2 中波红外下目标跟踪结果(实线:本文算法,虚线:fDSST)

Fig.2 The target tracking results of middle-wave infrared images (solid line: proposed algorithm, dotted line: fDSST)





(c) 第 471 帧视频图像
(c) The 471st frame



(d) 第 600 帧视频图像
(d) The 600th frame

图 3 长波红外下目标跟踪结果 (实线: 本文算法, 虚线: fDSST)

Fig.3 The target tracking results of long-wave infrared images(solid line: proposed algorithm, dotted line: fDSST)

从图 2 中可以看出, 针对中波红外波段, 在目标未被遮挡之前, 两种算法均能较好地跟踪目标, 与实际目标位置偏差较小。但是在目标被部分遮挡时, fDSST 算法仍然在更新滤波器模型, 导致滤波器模型加入较多的背景信息, 从而在第 333 帧视频图像之后, 当目标重新进入画面并运动时, 该算法的目标框仍然停留在之前更新学习的图像背景位置, 即图 2(c)中虚线目标框所示, 无法有效实现目标的抗遮挡处理。而本文所提算法在目标被部分遮挡时, 能及时地停止更新相关滤波器模型, 以及在目标重新进入画面后的短时间内, 准确地重捕获目标并具有较高的时效性, 如图 2(c)中实线目标框所示, 并在后续视频帧中稳定准确地对目标持续跟踪。在图 3 的长波红外波段下, 同样可以看出 fDSST 跟踪算法无法有效地对红外目标在完全被遮挡后进行重捕获, 而本文所提算法能有效、快速地对长波红外目标在被遮挡后进行再检测跟踪, 实现目标的持续稳定跟踪。

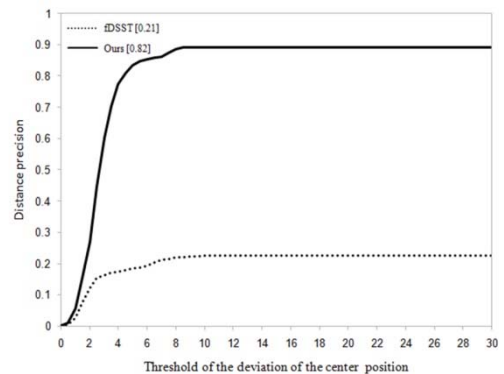
通过以上分析可知本文所提算法针对红外目标, 能在目标被遮挡时及时停止更新相关滤波器模型, 避免模型带入背景噪声, 并在引入再检测机制后, 能快速准确地对跟踪目标进行再捕获, 并实现持续稳定跟踪。

2.3.2 定量比较

由于距离精度和成功率是跟踪算法中广泛使用的评估标准, 因此本实验选用这两个标准来评估算法的整体性能^[13]。

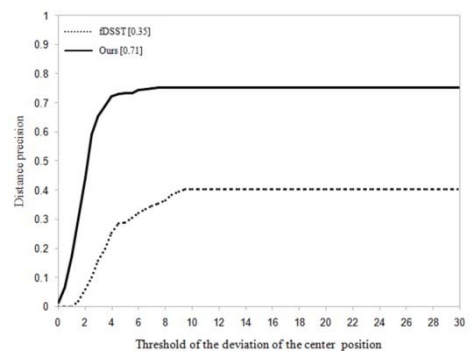
图 4 分别是中波和长波红外下, 不同算法的距离精度随目标中心位置偏差阈值变化的实验结果。从图中可以看出本文算法在中波和长波红外下的距离精度均值分别为 82%和 71%, 相比于 fDSST 算法分别提高了 61%和 36%, 在中心位置偏差阈值为 10 个像

素时, 本文算法在中波和长波红外下的距离精度分别达到了 90%和 75%, 明显优于相同阈值下 fDSST 算法的距离精度。



(a) 中波红外不同距离偏差阈值下的距离精度

(a) The distance precision of middle-wave infrared images at various thresholds of deviation



(b) 长波红外不同距离偏差阈值下的距离精度

(b) The distance precision of long-wave infrared images at various thresholds of deviation

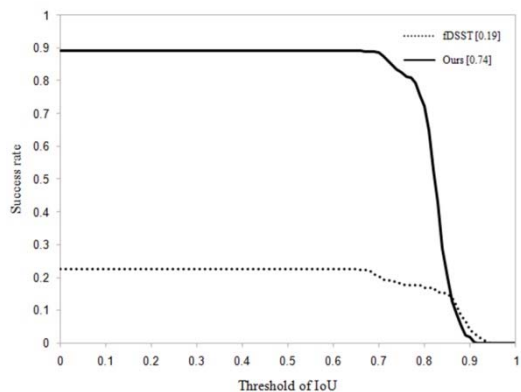
图 4 中波和长波红外下的平均距离精度

Fig.4 The average distance precision of middle-wave and long-wave infrared images

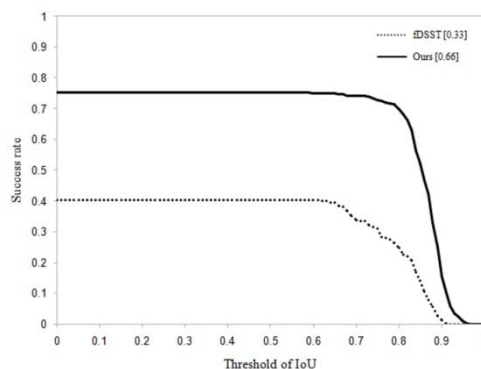
图5分别是中波和长波红外下,不同算法的跟踪成功率随交并比阈值变化的实验结果。在交并比阈值为0.5的情况下,本文算法的成功率在中波和长波红外下分别达到了89%和75%,分别比fDSST算法高出了66%和34%,并且平均成功率无论是在中波红外

还是长波红外情况下均明显优于fDSST算法。

综上实验结果可知,本文在针对红外目标跟踪的抗遮挡处理时,基于相关滤波引入目标再检测机制能够较好地提高跟踪效果,并能很好对解决目标被全部遮挡的问题。



(a) 中波红外不同交并比阈值下的成功率
(a) The success rate of middle-wave images at various thresholds of IoU



(b) 长波红外不同交并比阈值下的成功率
(b) The success rate of long-wave images at various thresholds of IoU

图5 中波和长波红外下的平均成功率

Fig.5 The average success rate of middle-wave and long-wave infrared images

在跟踪效率上,表2为跟踪算法运行速度的对比结果,从对比结果中可以看出,本文提出的算法与fDSST算法在运行速度上相差不大,均在100 fps左右,具有较高的运行效率,能够完全满足嵌入式系统中实时跟踪的需求。

表2 跟踪算法运行速度对比

Table 2 Running speed comparison of tracking algorithms

Algorithm	Frames Per Second
fDSST	105
Ours	94

2.3.3 在线测试

在基于仿真视频针对跟踪算法进行相关性能测试的同时,也将跟踪算法移植到实际光电转塔中,并基于光电转塔,在两个实际场景中跟踪算法进行了在线测试。

在在线测试场景一中,正常被跟踪的目标车辆在第42帧被广告牌遮挡后跟踪丢失,并在第44帧重新出现在画面中,从图6中可以看出,在第44帧中目标被光电转塔重新捕获跟踪。在第42帧和第43帧的画面中,目标车辆跟踪丢失后,画面中存在其它干扰车辆,然而跟踪器并未重捕获其它干扰目标车辆,而是保持目标跟踪丢失状态直到真实目标出现并重捕获。



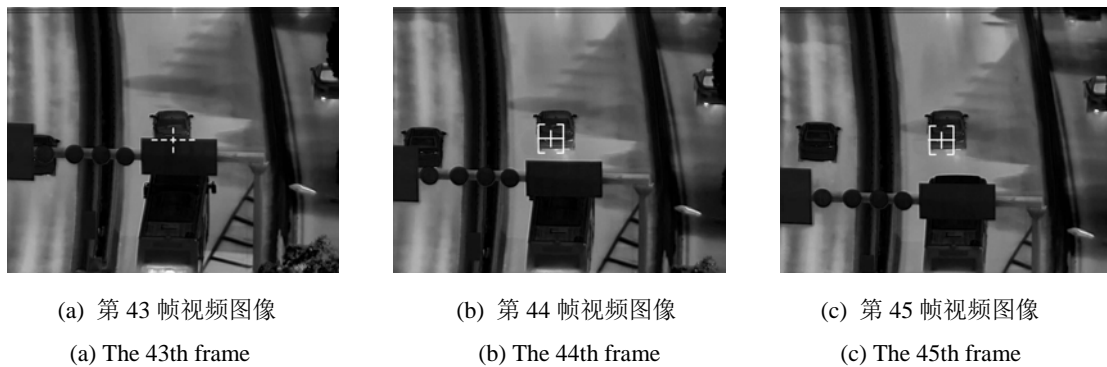
(a) 第39帧视频图像
(a) The 39th frame



(b) 第41帧视频图像
(b) The 41th frame



(c) 第42帧视频图像
(c) The 42th frame



(a) 第43帧视频图像
(a) The 43th frame

(b) 第44帧视频图像
(b) The 44th frame

(c) 第45帧视频图像
(c) The 45th frame

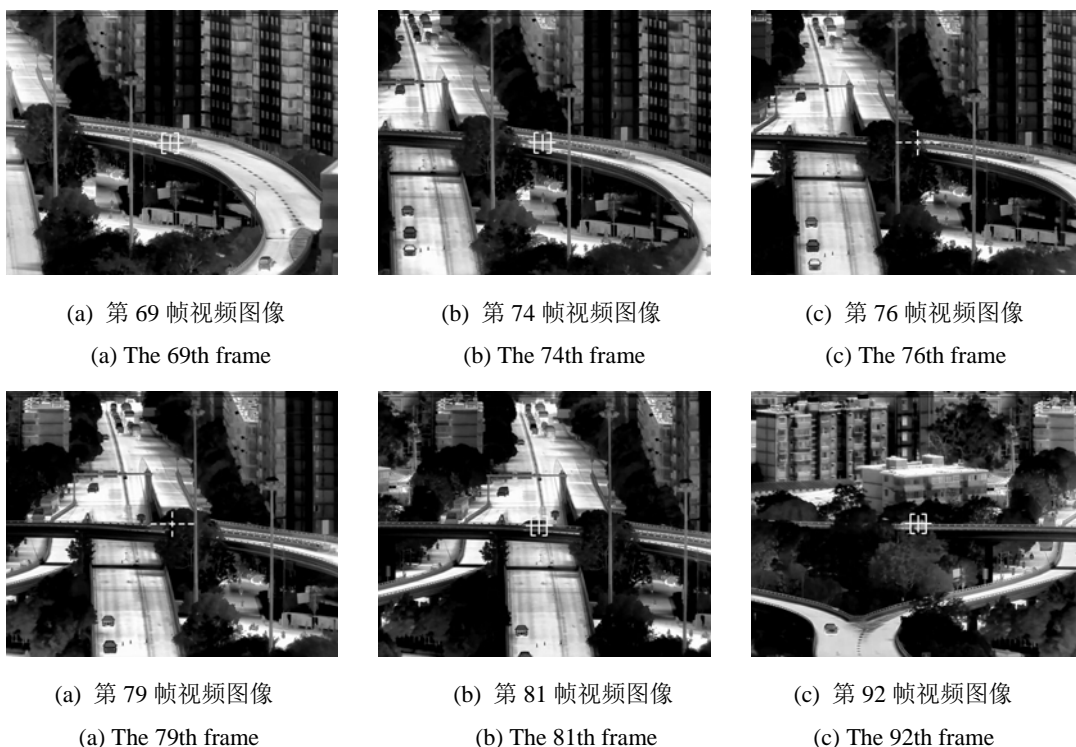
图6 在线测试场景一跟踪结果(从左至右、从上至下分别为第39、41、42、43、44、45帧)

Fig.6 The tracking result of online test in the first scene(the frame index are 39、41、42、43、44、45 respectively from left to right and from top to bottom)

在在线测试场景二中,跟踪的部分结果如图7所示,在第74帧及以前,虽然画面中存在着较多其它车辆,但是目标车辆仍被较好地持续跟踪,在第76帧时,目标开始被树木遮挡,并持续遮挡至第79帧,在被遮挡期间,跟踪器处于目标丢失状态,从画面中可以看出,目标跟踪丢失期间,场景中存在较多的其它车辆甚至是极为相似的车辆,但是跟踪器并未重捕获至其它干扰车辆造成误跟踪情况。从图7中可以看出

在第81帧目标车辆重新出现在画面后跟踪器迅速对目标进行重捕获并进行持续准确的实时跟踪。

综上基于光电转塔的在线测试结果可以看出,在实际场景存在多个目标的情况下,当跟踪目标被遮挡导致跟踪丢失后,跟踪器能较好地避免重捕获其它相似目标而导致误跟踪,并在真实目标重新出现在画面中时迅速对其进行重捕获并持续、准确地进行实时跟踪。



(a) 第69帧视频图像
(a) The 69th frame

(b) 第74帧视频图像
(b) The 74th frame

(c) 第76帧视频图像
(c) The 76th frame

(a) 第79帧视频图像
(a) The 79th frame

(b) 第81帧视频图像
(b) The 81th frame

(c) 第92帧视频图像
(c) The 92th frame

图7 在线测试场景二跟踪结果(从左至右、从上至下分别为第69、74、76、79、81、92帧)

Fig.7 The tracking result of online test in the second scene(the frame index are 69、74、76、79、81、92 respectively from left to right and from top to bottom)

3 结论

针对相关滤波跟踪器在红外目标被遮挡时无法有效持续跟踪目标的问题, 本文在快速判别尺度空间滤波器跟踪器算法的基础上引入基于深度学习的目标再检测机制, 提出了一种红外目标跟踪抗遮挡处理的跟踪算法, 在通过大量的红外图像数据集对目标检测器进行训练后, 该算法在目标被遮挡后又重新进入画面时, 能够快速准确地对目标进行重捕获并进行持续跟踪。相比与原始的 fDSST 算法, 本文算法在中波和长波红外下的距离精度和跟踪成功率均有较大地提升, 在具有较好的跟踪准确度的同时具有较高的跟踪速率, 能满足嵌入式系统中实时跟踪的要求, 具有一定的实际应用价值。尽管提出的方法在红外目标跟踪上取得了良好的跟踪效果, 但是还需要在更加复杂实际红外场景中进行测试和改进。

参考文献:

- [1] 蔡毅. 红外系统中的扫描型和凝视型FPA[J]. *红外技术*, 2001, **23**(1): 1-5,18.
CAI Yi. The scanning and starring FPA in infrared systems[J]. *Infrared Technology*, 2001, **23**(1): 1-5, 18.
- [2] 范晋祥, 杨建宇. 军用红外成像探测跟踪技术发展趋势与新技术研究进展[C]//第十届全国光电技术学术交流会议论文集, 2012: 125-134.
FAN Jinxiang, YANG Jianyu. Development trend and research progress of military infrared imaging detection and tracking technology[C]//*Proceedings of the 10th National Symposium on Optoelectronic Technology*, 2012: 125-134.
- [3] 王云萍. 美国天基红外导弹预警技术分析[J]. *光电技术应用*, 2019, **34**(3): 1-7.
WANG Yunping. Analysis of space-based infrared missile warning system in America[J]. *Electro-Optic Technology Application*, 2019, **34**(3): 1-7.
- [4] FU Z, CHANG Z, YONG H, et al. Foreground gated network for surveillance object detection[C]//2018 *IEEE Fourth International Conference on Multimedia Big Data (BigMM)*, 2018: 1-7.
- [5] Danelljan M, Hager G, Khan F S, et al. Learning spatially regularized correlation filters for visual tracking[C]// *IEEE International Conference on Computer Vision*, 2015: 4310-4318.
- [6] 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, CVPR*, 2010: 13-18.
- [7] 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*, 2012: 702-715.
- [8] Henriques J F, Caseiro R, Martins P, et al. High-speed tracking with kernelized correlation filters[C]//*IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2015, **37**: 583-596.
- [9] Danelljan M, Häger G, Khan F S, et al. Accurate scale estimation for robust visual tracking[C]//*British Machine Vision Conference*, 2014(65): 1-11.
- [10] Danelljan M, Häger G, Khan F S, et al. Discriminative scale space tracking[C]//*IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, **39**(8): 1561-1575.
- [11] Bochkovskiy A, WANG C Y, LIAO H. YOLOv4: optimal speed and accuracy of object detection[J/OL]. *Computer Vision and Pattern Recognition*, arXiv:2004.10934. <https://arxiv.org/abs/2004.10934>.
- [12] Felzenszwalb P F, Girshick R B, Mc Allester D, et al. Object detection with discriminatively trained part-based models[C]//*IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, **32**: 1627-1645.
- [13] WU Y, LIM J, YANG M H. Object tracking benchmark[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2015, **37**: 1834-1848.