

基于自适应响应融合的相关滤波红外目标跟踪

房胜男,谷小婧,顾幸生

(华东理工大学 化工过程先进控制与优化技术教育部重点实验室,上海 200237)

摘要: 红外目标跟踪在军事和民用视频监控领域有重要的研究意义,但受热成像原理限制,红外目标分辨率低、对比度低、纹理信息缺失。针对红外目标特征信息量少导致跟踪性能较低的问题,提出一种基于自适应响应融合的相关滤波跟踪算法。该算法基于连续卷积运算的相关滤波跟踪框架,通过构造视觉显著性特征来增强目标外观描述,并结合对冲决策理论对由不同特征计算得到的多个滤波响应进行自适应融合,最终根据融合响应预测目标中心位置。此外,通过尺度滤波器来实现目标的尺度预测,得到完整的跟踪结果。在公开的红外视频数据集 VOT-TIR2016 进行测试,实验结果表明:与同类算法相比,该算法表现出更高的跟踪精确度和鲁棒性。

关键词: 红外目标跟踪; 相关滤波; 响应融合; 显著性特征; 对冲算法

中图分类号: TP391.4 **文献标志码:** A **DOI:** 10.3788/IRLA201948.0626003

Infrared target tracking with correlation filter based on adaptive fusion of responses

Fang Shengnan, Gu Xiaojing, Gu Xingsheng

(Key Laboratory of Advanced Control and Optimization for Chemical Process, Ministry of Education, East China University of Science and Technology, Shanghai 200237, China)

Abstract: Infrared target tracking is of great significance to the research in military and civil video surveillance. Due to the special thermal-imaging mechanism, infrared targets are often with low resolution, low contrast and in the lack of textures. Aiming at the deterioration of tracking performance caused by insufficient common features of infrared targets, a novel tracking algorithm was proposed based on adaptive fusion of correlation filter responses. The algorithm explored the framework of the correlation filter with continuous convolution operators. The saliency feature was comprised to enhance the object appearance description. The center location of a target was predicted by the fused responses that were calculated from an adaptive fusion of multiple correlation responses with Hedge decision-theoretic. Additionally, the final tracking result was obtained after multi-scale estimation based on scale filters. The experimental results show that the algorithm has better performance in tracking accuracy and robustness compared with other tracking methods on the public infrared video dataset VOT-TIR2016.

Key words: infrared target tracking; correlation filter; response fusion; saliency feature; Hedge algorithm

收稿日期:2019-01-25; 修订日期:2019-02-13

基金项目:国家自然科学基金(61775058,61573144,61502293)

作者简介:房胜男(1994-),女,硕士生,主要从事红外目标检测和跟踪方面的研究。Email:fsn506@163.com

导师简介:顾幸生(1960-),男,教授,博士生导师,主要从事生产计划与调度、工业过程建模控制与优化、故障诊断和检测方面的研究。

Email:xsgu@ecust.edu.cn

0 引言

红外图像由于其不受光照条件影响、抗干扰能力强、全天候工作等特点,在导弹制导、交通监控、医疗诊断、气象分析等军事和民用领域得到了广泛应用。基于红外热像视频的目标跟踪是这些应用中的一项关键技术,也是计算机视觉领域的研究热点。

红外目标跟踪与可见光目标跟踪相比更具挑战性,因为除了需要解决外观形变、运动模糊、遮挡干扰、平面内外旋转等通用性跟踪问题,它还需考虑红外成像固有的低分辨率、纹理缺失、噪声大、对比度低等显著缺陷。为了提高红外目标跟踪性能,学者们进行了相关研究,例如结合红外行人的运动特性和边缘信息,基于模式匹配原理设计的改进 Mean-shift 跟踪算法^[1];将卡尔曼滤波和均值漂移相结合的红外点目标跟踪算法^[2]。参考文献[3]结合粒子滤波跟踪框架,利用稀疏表示模型来减少红外目标区域的噪声。

随着相关滤波理论的广泛发展,它的优越性能吸引了学者们将其用于红外目标跟踪。参考文献[4]将多种流行的跟踪算法同时用于配准的可见光和红外视频数据集进行测试,通过分析可见光跟踪算法直接应用于红外视频的性能损失,得出相关滤波比其他算法更适用于红外目标跟踪的结论。因为相关滤波器对不同光谱的可见光和红外图像并不敏感,所以跟踪性能损失较小。

目前,引入相关滤波进行红外目标跟踪的算法^[5-7]

大多数是基于 KCF(Kernelized Correlation Filters)^[8]跟踪框架。相比可见光图像,红外目标缺失颜色、纹理等特征,边缘模糊且对比度低,因此,总体信息量相对较少,许多学者考虑从多信息融合的角度来改善跟踪性能。参考文献[7]将二维灰度特征与显著性特征按页方式进行拼接扩展成三维特征输入 KCF 框架中。参考文献[9]构建权值函数来融合空间、运动和卷积三种特征,改进 KCF 实现红外目标跟踪。以上从特征层面进行信息融合的方法,会由于不同帧之间的目标特征差异造成信息融合不稳定。参考文献[10]根据不同卷积层特征预测得到的多个目标位置进行加权融合,这种直接的终端位置融合方法不能很好地利用不同特征在跟踪过程中的互补信息。

文中提出了一种基于自适应响应融合的相关滤波跟踪算法,不同于直接的特征融合和位置融合,该算法为了充分且稳定地利用多信息的互补性,对由多种特征计算得到的滤波响应,结合对冲算法^[11]自适应调整响应权值,用融合响应预测目标中心位置。其中,文中的跟踪框架为基于连续卷积运算的相关滤波器 ECO (Efficient Convolution Operators)^[12],与 KCF 不同,它可以处理不同分辨率的特征图,且通过简化模型参数、多样化训练样本提升了跟踪性能。此外,对于多信息的选择,文中针对红外目标特性,采用方向梯度直方图特征(HOG)^[13]和强度通道特征(IC)^[14],并引入视觉显著性机制^[15]构建有效的显著性特征(Sal)来增强红外目标信息。最后结合尺度滤波器实现目标的尺度预测,完成对红外目标的稳健跟踪。图 1 为文中算法流程图。

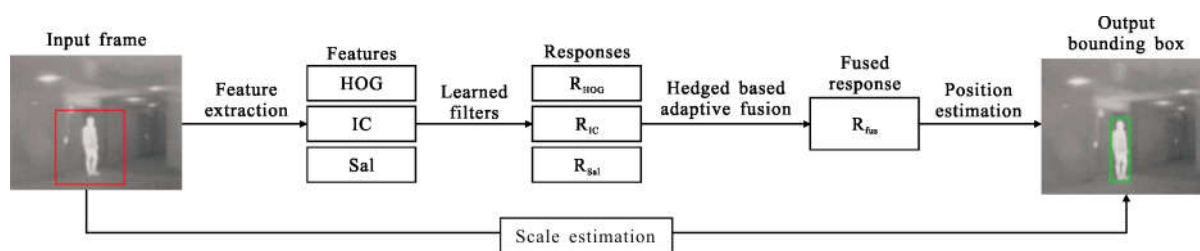


图 1 算法流程图

Fig.1 Flow chart of algorithm

1 相关滤波基本原理

1.1 基于连续卷积运算的 ECO 跟踪器

不同于传统相关滤波器只能使用单一分辨率的

特征,ECO 跟踪器^[12]使用隐式插值模型将相关滤波器的学习过程转换到连续空间域,这样可以处理分辨率不同的特征图,而且这种连续性相关滤波模型能实现更精确的像素级定位,适用于缺乏足够信息

的红外目标跟踪。

对于输入的目标区域样本 x_j , 对应提取的多通道特征为 x_j^d , 其中 d 为特征通道数, 总特征通道个数为 D 。设 N_d 为样本特征的空间分辨率, 那么特征 $x_j^d \in \mathbb{R}^{N_d}$ 可以看作由离散空间变量 $n \in \{0, \dots, N_d-1\}$ 确定的函数 $x_j^d[n]$ 。为了引入连续空间域, 将特征图定义在连续区间 $[0, T) \subset \mathbb{R}$ 中, 对特征的插值操作为:

$$J_d\{x^d\}(t) = \sum_{n=0}^{N_d-1} x^d[n] b_d\left(t - \frac{T}{N_d} n\right) \quad (1)$$

式中: b_d 为插值函数, 这里取三次样条插值法。那么, 所有的插值特征图为 $J\{x\}$ 。不同通道特征对应学习的相关滤波器为 $f=(f^1, \dots, f^D)$, 对应的滤波响应即为:

$$S_f\{x\} = f^* J\{x\} = \sum_{d=1}^D f^{d*} J_d\{x^d\} \quad (2)$$

式中: $*$ 为循环卷积操作。相关滤波器的在线训练是公式(3)目标函数的最小化过程:

$$E(f) = \sum_{j=1}^M \alpha_j \|S_f\{x_j\} - y_j\|^2 + \sum_{d=1}^D \|\omega f^d\|^2 \quad (3)$$

式中: M 为训练样本个数; α_j 为样本的重要性权值; y_j 代表目标的高斯标签函数; ω 为处理边缘效应的空间正则化项。和传统相关滤波器处理方法类似, 为了快速训练滤波器 f , ECO 结合傅里叶变换理论处理相关函数, 并使用共轭梯度法优化变换后的目标函数来求解滤波器参数。预测的目标中心位置是连续函数 $S_f\{x\}$ 最大值对应的样本位置。

除了以上结合连续卷积运算的基本框架之外, 为了解决训练过拟合和计算复杂度的问题, ECO^[12] 还采用分解卷积操作对特征模型进行降维, 结合高斯混合模型对训练样本集进行多样化处理。

1.2 多尺度估计

针对目标运动过程中的尺度变化问题, 文中结合快速多尺度跟踪器(fDSST)^[16]算法, 通过学习一个独立的一维相关滤波器来检测尺度变化。主要原理是基于尺度池的方法, 以目标位置为中心, 在其周围区域采集一系列不同尺度的图像块, 对这些图像块提取特征来训练尺度相关滤波器, 最后选定合适的尺度模型用于尺度预测。

2 自适应滤波响应融合

红外图像缺少颜色、纹理、边缘等特征, 对于不

同场景下非特定类别的红外目标跟踪, 在跟踪框架中如何选择并利用目标的有限信息显得尤为重要。首先针对红外目标特性, 在 HOG 和 IC 两种特征的基础上, 引入视觉注意机制构建有效的显著性特征来增强目标信息。对于如何充分利用构建的多特征信息, 文中在对每种特征独立训练相关滤波器之后, 再采用基于对冲决策的自适应权值调整机制将计算得到的不同滤波响应进行融合。

2.1 特征构建

文中使用手工特征描述红外目标, 未采用 ECO 中结合卷积神经网络(CNN)提取特征的方式。原因之一在于复杂的计算量会大大降低跟踪效率; 其次, 由于目前的卷积网络大多是由大型可见光图像数据集训练得到, 缺乏基于通用红外目标数据集训练的深度网络, 导致一般的 CNN 对红外图像的特征表达不全面。因此, 首先选择两种在跟踪领域应用较广的 HOG 特征^[13]和强度通道特征^[14]。

HOG 特征通过计算和统计图像局部区域的梯度方向直方图来构建多通道梯度特征, 这种密集特征表示能较好地反映局部像素之间的关联性且充分表达了图像信息, 对几何和光学形变鲁棒, 但也存在对噪声敏感、对快速形变表现不好的缺陷。强度通道特征基于分布场描述算子, 它是通过估计数据的概率密度分布进行图像信息分层量化而构成的一组特征。分布场描述算子的统计特性服从平均移位直方图理论, 包含目标的空间结构信息, 能较好地适应跟踪环境的变化, 对噪声比较鲁棒。

为了充分突出红外目标, 文中引入基于频域的视觉显著性机制^[15], 从图像的频谱特性入手, 通过分析输入图像的对数谱, 提取频域中的谱残差, 高效地构建显著性特征来增强红外目标信息, 抑制背景干扰。

对输入的目标图像 I , 首先通过傅里叶变换转化到频域, 计算幅度谱 $A(f)=r(F(I))$ 和相位谱 $P(f)=\varphi(F(I))$, 其中, $F(\cdot)$ 表示傅里叶变换。然后, 将幅度谱转化为对数谱 $L(f)=\log(A(f))$, 与线性空间滤波器相乘得到平均频谱 $V(f)=h_n(f)L(f)$, 其中 $h_n(f)$ 是均值滤波器。谱残差 $R(f)$ 为对数谱和平均频谱之间的差值:

$$R(f) = L(f) - V(f) \quad (4)$$

最后, 显著图通过公式(5)计算得到:

$$S = g_n F^{-1}(\exp(R(f) + P(f)))^2 \quad (5)$$

式中: g_n 为高斯滤波器; $F^{-1}(\cdot)$ 为傅里叶逆变换。

为了提高显著图质量,假设目标运动较平滑且相邻帧之间的目标位置波动不大,结合参考文献[17]的中心强化-四周弱化改进机制,构建增强后的显著性特征 Sal 为:

$$Sal = \frac{\sqrt{W^2+H^2} - \sqrt{(m-x)^2+(n-y)^2}}{\sqrt{W^2+H^2}} S \quad (6)$$

式中: W, H 为输入图像区域的宽度和高度; (m, n) 为像素点坐标; (x, y) 为目标上一帧得到的中心位置结果。

2.2 响应融合

文中针对红外目标选择的三种手工特征在不同的跟踪场景中具有不同的表达能力,因此它们对目标位置预测的作用会有一定差异。为了深入利用多信息的互补性,文中没有从特征层面进行直接融合,而是引入对冲(Hedge)算法[10]的在线决策原理,设计一种自适应权值分配方法来融合不同的相关滤波响应。而在 ECO 算法中,最终的滤波响应是由多个响应直接相加得到,可以看作是固定权值为 1 的融合方式,这种权值不变的融合方法忽略了视频不同场景和时刻的差异,会使跟踪性能受到影响。

文中在 1.1 节 ECO 跟踪器的基础上,定义第 t 帧时刻第 k 个特征的滤波响应为 $R_{k,t}, k=1, \dots, N$, 其中 N 为特征类别总数,等价于 1.1 节中滤波响应 S_f 经过傅里叶变换后的变量,表现为矩阵形式。最终的融合响应通过各响应的权值 $\omega_{k,t}$ 自适应加权得到:

$$R_{fus,t} = \sum_{k=1}^N \omega_{k,t} R_{k,t} \quad (7)$$

那么,问题即为权值 $\omega_{k,t}$ 的求解。对于融合后的响应 $R_{fus,t}$,设置每个特征对应响应 $R_{k,t}$ 的损失变量为 $l_{k,t}$,文中采用向量化后的响应 \mathbf{R} ,结合皮尔森相关系数定义损失为:

$$l_{k,t} = \frac{\langle \mathbf{R}_{k,t} - \bar{\mathbf{R}}_{k,t}, \mathbf{R}_{fus,t} - \bar{\mathbf{R}}_{fus,t} \rangle}{\|\mathbf{R}_{k,t} - \bar{\mathbf{R}}_{k,t}\| \|\mathbf{R}_{fus,t} - \bar{\mathbf{R}}_{fus,t}\|} \quad (8)$$

皮尔森相关系数是将数据进行归一化之后的余弦相似度计算量,值在 $[-1, 1]$ 之间,绝对值越大,相关性越强,具有较好的平移不变性和尺度不变性,这样,使用不同响应之间的相似性作为损失变量比较合理。然后,定义期望损失为:

$$l_{A,t} = \sum_{k=1}^N \omega_{k,t} l_{k,t} \quad (9)$$

引入后悔变量 $reg_{k,t} = l_{A,t} - l_{k,t}$,累积后悔值为:

$$Reg_{k,t} = \sum_{\tau=1}^t reg_{k,\tau} = Reg_{k,t-1} + (l_{A,t} - l_{k,t}) \quad (10)$$

求解权值 $\omega_{k,t}$ 就是优化 $Reg_{k,t}$ 的过程,结合 Hedge 算法的推导结论,权值结果为:

$$\omega_{k,t} = \frac{[Reg_{k,t-1}]_+}{c_{t-1}} \exp\left(\frac{[Reg_{k,t-1}]_+^2}{2c_{t-1}}\right) \quad (11)$$

式中: $[Reg_{k,t-1}]_+$ 为 $Reg_{k,t-1}$ 和 0 相比的较大值; c_{t-1} 为尺度因子,由公式(12)计算得到:

$$\frac{1}{N} \sum_{k=1}^N \exp\left(\frac{[Reg_{k,t-1}]_+^2}{2c_{t-1}}\right) = e \quad (12)$$

式中: e 为自然常数。

根据上述方法,除了在视频第一帧需要手动设置初始权值,后续帧中,多特征对应的滤波响应通过上述流程可以自适应调整融合权值,以获取最终用于预测目标位置的融合响应,它比固定权值融合的方法更适用于红外视频的场景变化。

3 实验结果与分析

3.1 实验环境与参数设置

文中实验的计算机配置为 Intel Xeon CPU E5-2640,主频为 2.4 GHz,内存为 64 GB,开发平台为 MATLAB R2017a。实验中与 ECO 基本参数设置不同之处在于,采用三种手工特征 HOG、IC、Sal 来表示目标外观,其中决定特征分辨率的 cell size 分别设置为 6、4、1。此外,对于三个滤波响应的初始权值,文中使用单一特征分别得到各自的跟踪结果,据此确定初始权值,经过试验计算,初始权值分别设定为 0.30、0.55、0.15。

3.2 数据库及评价指标

文中对公开的红外视频数据集 VOT-TIR2016 进行实验,该数据集一共有 25 个红外视频,平均视频长度为 740 帧,分辨率在 $305 \times 225 \sim 1920 \times 480$ 之间。视频从不同的室内外环境中采集得到,跟踪的红外目标包括不同场景下的行人、汽车、飞行器、动物、船只等,且目标具有诸多难以处理的属性,如遮挡、平面内外旋转、出视野、尺度变化、快速运动等。

VOT-TIR2016 使用标准的 VOT toolkit 工具包进行跟踪性能测试,跟踪器对数据集中的每一个视频都重复运行多次,过程中若出现跟踪失败的情况,系统会在 5 帧之后自动重新初始化跟踪器参数以继续完成任务。

实验中涉及的跟踪性能指标包括准确度(A)、鲁棒性(R)、期望平均重叠率(EAO)。A 用来评价跟踪的准确度,表示跟踪成功状态下,预测目标边界框和真值之间的平均重叠率,数值越大准确度越高。R 用来评价跟踪的稳定性,表示平均跟踪失败次数,数值越大鲁棒性越差。EAO 用来综合评价目标跟踪性能,考虑了鲁棒性和准确度的结合,表示固定帧数的一段视频中,跟踪器在非重置情况下边界框重叠率的期望值,数值越大性能越好。

对于算法的速度评价指标,使用等效滤波运算(EFO)指标,即跟踪算法处理每帧图像的运行时间除以标准滤波操作时间得到的一个归一化指标值,EFO 值越大,速度越大。

3.3 结果与分析

3.3.1 多特征响应融合结果分析

为了评价不同特征对于红外目标跟踪性能的影响,文中在原始 ECO 方法框架的基础上,进行六组实验。分别为单一特征测试(记为 ECO_HOG、ECO_IC、ECO_Sal)、除去显著性特征的二特征滤波响应直接相加(记为 ECO_2+)、三特征滤波响应直接相加(记为 ECO_3+)、文中的三特征滤波响应自适应融合跟踪。表 1 列出了六组实验的跟踪结果。

表 1 不同特征的跟踪结果

Tab.1 Tracking results of different features

Methods	A	R	EAO
ECO_HOG	0.44	2.01	0.231
ECO_IC	0.52	1.79	0.296
ECO_Sal	0.38	1.21	0.197
ECO_2+	0.60	1.54	0.349
ECO_3+	0.61	1.42	0.356
Proposed	0.63	1.08	0.382

由表 1 可见,基于 ECO 框架的多响应融合方法对比单特征方法,跟踪性能得到很大提升。由单一特征跟踪结果分析,IC 特征由于其空间结构信息的丰富性和对噪声的鲁棒性,在红外目标跟踪中 ECO_IC 表现最好。而 ECO_Sal 仅用显著性特征进行跟踪准确度较低,这是由于文中使用的显著性检测方法完全从图像频谱特性出发,不考虑图像的具体细节信

息,所以单一模型的判别力不强。但文中构造该特征的初衷是借助视觉显著性来增强红外目标信息,抑制背景干扰,通过和其他基本特征的互补来提升跟踪性能。因此,加入显著性特征 Sal 的 ECO_3+对比使用 HOG、IC 两种基本特征的 ECO_2+结果有一定提升。文中结合对冲算法设计的自适应滤波响应融合方法对比固定权值融合的方法 ECO_3+,跟踪准确度和鲁棒性都得到加强。

3.3.2 算法对比分析

为了对比不同跟踪方法的性能,选取 VOT-TIR2016 Challenge 提供的排名靠前的六组跟踪器 EBT、SRDCFir、TCNN、Staple-TIR、SHCT、MDNet-N 作为对比算法。其中,EBT、SHCT 是基于分块模型的方法,TCNN、MDNet-N 是基于深度学习的算法,SRDCFir、Staple-TIR 是基于相关滤波器的算法,具体方法介绍参见参考文献[18]。表 2 列出了文中算法和六个方法的评价指标结果,图 2 是不同算法的期望平均重叠率 EAO 在不同视频序列长度下的变化曲线,图例中显示了按从小到大排序的 EAO 数值。

表 2 不同算法跟踪结果

Tab.2 Tracking results of different algorithms

Methods	A	R	EAO	EFO(speed)
EBT	0.43	0.88	0.368	3.60
SRDCFir	0.61	1.12	0.364	3.30
TCNN	0.61	2.05	0.287	0.76
Staple-TIR	0.64	2.80	0.264	14.25
SHCT	0.60	2.76	0.263	0.91
MDNet-N	0.64	2.58	0.243	0.61
Proposed	0.63	1.08	0.382	13.35

结合表 2 和图 2 的实验结果,文中算法相比 VOT-TIR2016 排名前六的跟踪器,在主要的评价指标 EAO 上领先第二名算法 3.8%,精度 A 为 0.63 和第一名 0.64 差距很小,鲁棒性 R 为 1.08 比第一名的 0.88 稍有逊色。对比其他算法的高精度低鲁棒性或者高鲁棒性低精度,文中算法在这两方面达到较好的平衡,整体跟踪性能达到最优。从跟踪速度上看,文中算法的 EFO 指标值仅次于 Staple-TIR 法,对比深度学习算法有明显的速度优势。

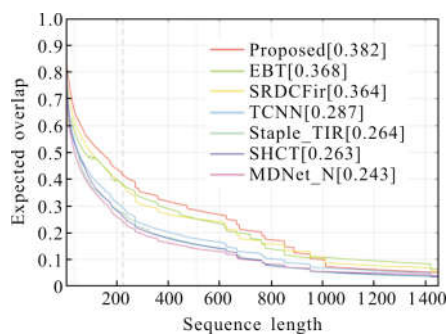


图 2 不同算法的 EAO 变化曲线

Fig.2 EAO curves of different algorithms

VOT-TIR2016 数据库对所有视频涉及的属性类别都进行标注,包括相机移动、遮挡、尺寸变化、运

动变化、动态变化(将红外图像从 16 位转为 8 位造成的动态范围改变)、没有特殊属性,这些属性是目标跟踪所要解决的通用性问题。针对不同属性的实验结果如图 3 所示,除了动态变化属性之外,文中算法的平均重叠率都处于前三名的水平,说明对于视频的不同变化场景,该算法有较优的跟踪性能。其中,文中算法对于快速运动的变化属性表现最优,因为目标在快速移动时会出现运动模糊,而文中算法构建了合理的目标特征且进行了在线多信息融合,即使目标模糊、特征突变,也会根据鲁棒的自适应响应融合机制加强目标的鉴别力,所以相比其他算法表现最好。

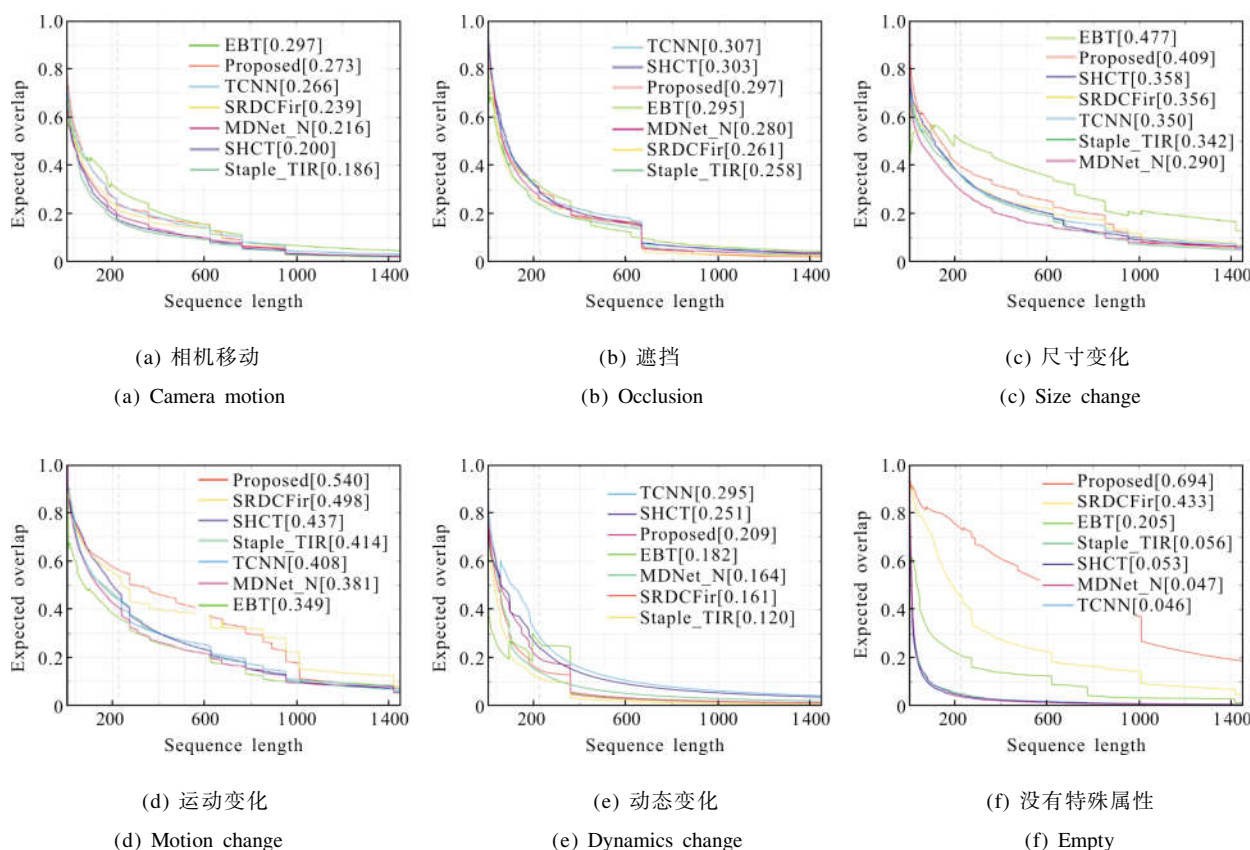


图 3 不同属性条件下对比算法的 EAO 变化曲线

Fig.3 EAO curves of different algorithms in different attributes

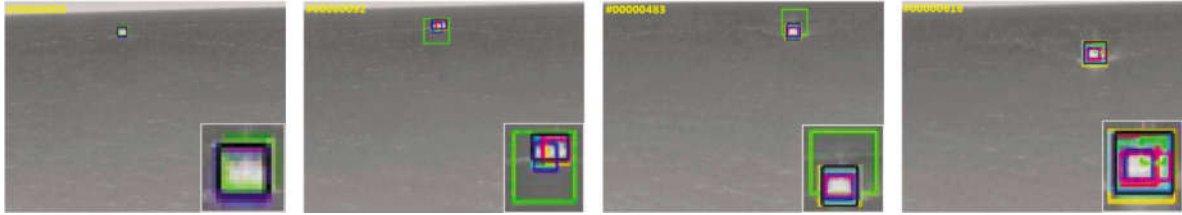
图 4 是不同算法在部分视频某些时刻的跟踪框显示图。在 birds 序列中,人体目标会发生形状改变,如正面背面弯腰等动作,MDNet_N、EBT 等算法出现了跟踪不准现象。对于尺寸较小且运动迅速的船体目标即 boat1 序列,文中算法表现最好。当目标尺度发生较大变化,如 car1 序列,一些算法如Staple_TIR

因为不能迅速调整目标尺寸而造成跟踪精度变低。在有相似背景干扰的 quadcopter 序列中,飞行器目标在背景亮度相似的窗口运动时,多个跟踪器出现明显漂移。在 street 序列中,当其他行人运动到目标行人位置时对其造成了遮挡,文中算法对此情况同样可以保持较好的跟踪性能。



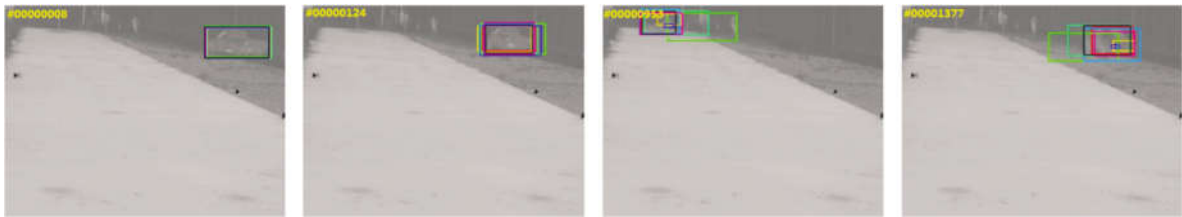
(a) Birds (变形)

(a) Birds (deformation)



(b) Boat1 (快速移动)

(b) Boat1 (fast motion)



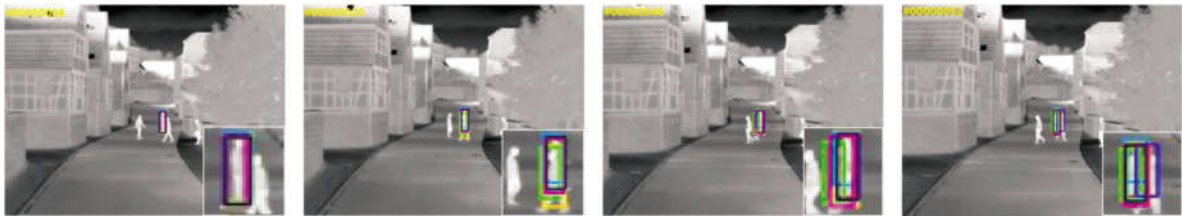
(c) Car1 (尺寸变化)

(c) Car1 (scale variation)



(d) Quadcopter (背景干扰)

(d) Quadcopter (background clutter)



(e) Street (遮挡)

(e) Street (occlusion)

— Proposed
 — EBT
 — SRDCFir
 — TCNN
 — Staple_TIR
 — SHCT
 — MDNet_N
 — Ground truth

图 4 不同算法的部分视频帧的跟踪框显示图

Fig.4 Different algorithms' tracking results of bounding box in some video frames

4 结论

针对红外图像特征信息量不足,文中提出一种结合多特征的自适应响应融合的相关滤波红外目标跟踪算法。通过方向梯度直方图特征、强度通道特征、显著性特征来描述目标模型,基于连续卷积运算的相关滤波器跟踪框架,考虑不同特征的跟踪性能贡献,引入对冲算法对得到的相关滤波响应进行自适应权值融合,实现红外目标的稳健跟踪。公开的红外视频数据集的实验测试结果表明:对于不同场景不同属性下的目标,该算法达到了较好的跟踪鲁棒性和准确度。

参考文献:

- [1] Wang X, Liu L, Tang Z. Infrared human tracking with improved mean shift algorithm based on multicue fusion[J]. *Applied Optics*, 2009, 48(21): 4201-4212.
- [2] Liu R, Li X, Han L, et al. Track infrared point targets based on projection coefficient templates and non-linear correlation combined with Kalman prediction [J]. *Infrared Physics and Technology*, 2013, 57(2): 68-75.
- [3] Tang Zhengyuan, Zhao Jiajia, Yang Jie, et al. Infrared target tracking algorithm based on sparse representation model[J]. *Infrared and Laser Engineering*, 2012, 41(5): 1389-1395. (in Chinese)
唐峥远, 赵佳佳, 杨杰, 等. 基于稀疏表示模型的红外目标跟踪算法[J]. *红外与激光工程*, 2012, 41(5): 1389-1395.
- [4] Gundogdu E, Ozkan H, Demir H S, et al. Comparison of infrared and visible imagery for object tracking: Toward trackers with superior IR performance [C]//Computer Vision and Pattern Recognition Workshops, IEEE, 2015:1-9.
- [5] Liu Q, Lu X, He Z, et al. Deep convolutional neural networks for thermal Infrared object tracking[J]. *Knowledge-Based Systems*, 2017, 134: 189-198.
- [6] Qian K, Zhou H, Rong S, et al. Infrared dim-small target tracking via singular value decomposition and improved Kernelized correlation filter [J]. *Infrared Physics and Technology*, 2017, 82: 18-27.
- [7] Zheng Wuxing, Wang Chunping, Fu Qiang, et al. Aerial infrared target tracking based on gray and saliency features fusion[J]. *Laser and Infrared*, 2018, 48(3): 338-342. (in Chinese)
郑武兴, 王春平, 付强, 等. 融合灰度与显著性特征的空中红外目标跟踪[J]. *激光与红外*, 2018, 48(3): 338-342.
- [8] Henriques J F, Rui C, Martins P, et al. High-speed tracking with kernelized correlation filters [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, 37(3): 583-596.
- [9] He Y J, Li M, Zhang J L, et al. Infrared target tracking via weighted correlation filter [J]. *Infrared Physics and Technology*, 2015, 73: 103-114.
- [10] Qi Y, Zhang S, Qin L, et al. Hedged Deep Tracking[C]//Computer Vision and Pattern Recognition. IEEE, 2016: 4303-4311.
- [11] Chaudhuri K, Freund Y, Hsu D. A parameter-free hedging algorithm [C]//Advances in Neural Information processing systems, 2009: 297-305.
- [12] Danelljan M, Bhat G, Khan F S, et al. ECO: efficient convolution operators for tracking [C]//Computer Vision and Pattern Recognition, IEEE, 2017: 6931-6939.
- [13] Dalal N, Triggs B. Histograms of oriented gradients for human detection [C]//Computer Vision and Pattern Recognition Workshops, IEEE, 2005: 886-893.
- [14] Felsberg M. Enhanced distribution field tracking using channel representations [C]//International Conference on Computer Vision Workshops, IEEE, 2013:121-128.
- [15] Hou X, Zhang L. Saliency detection: a spectral residual approach [C]//Computer Vision and Pattern Recognition, IEEE, 2007: 1-8.
- [16] Danelljan M, Hager G, Khan F S, et al. Discriminative scale space tracking [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(8): 1561-1575.
- [17] Wu Shidong, Bao Hua, Zhang Chenbin, et al. Particle filter tracking based on visual saliency feature [J]. *Journal of University of Science and Technology of China*, 2015, 45(11): 934-942. (in Chinese)
吴世东, 鲍华, 张陈斌, 等. 基于视觉显著性特征的粒子滤波跟踪算法 [J]. *中国科学技术大学学报*, 2015, 45(11): 934-942.
- [18] Felsberg M, Kristan M, Matas J, et al. The thermal infrared visual object tracking VOT-TIR2016 challenge results[C]//Computer Vision-ECCV 2016 Workshops, 2016: 824-849.