

基于多特征融合的红外目标跟踪方法研究

侯正阳, 赵春阳

(沈阳理工大学 信息科学与工程学院, 沈阳 110159)

摘要: 针对复杂场景下因红外图像成像质量差、纹理模糊、易受亮目标干扰等而导致红外目标跟踪效果不佳的问题, 提出一种基于多特征自适应融合的红外目标跟踪方法。设计一种大核小波网络 (large kernel & wavelet network, LKWNet) 用于提取深度特征, 设计自适应方向梯度直方图算法用于提取手工特征, 并将两种特征相结合, 提高红外图像特征表达能力。实验结果表明, 基于本文特征提取方法设计的红外目标跟踪器 ECOIR 在 LSOTB-TIR 数据集上的 Succ. AUC 和 Prec@20 分别为 79.0% 和 65.0%, 在 PTB-TIR 数据集上的 Succ. AUC 和 Prec@20 分别为 81.9% 和 63.2%, 表现出较好的鲁棒性和跟踪精度。

关键词: 红外目标跟踪; 自适应融合; 深度特征; 方向梯度直方图

中图分类号: TP391.4 **文献标志码:** A **DOI:** 10.3969/j.issn.1003-1251.2026.04.003

Research on Infrared Target Tracking Method Based on Multi-feature Fusion

HOU Zhengyang, ZHAO Chunyang

(Shenyang Ligong University, Shenyang 110159, China)

Abstract: To solve the problem of poor infrared target tracking performance caused by poor imaging quality, fuzzy texture and susceptibility to interference from bright targets in complex scenes, an infrared target tracking method based on adaptive fusion of multiple features is proposed. A large kernel & wavelet network (LKWNet) is designed to extract deep features, and an adaptive histogram of oriented gradients method is designed to extract hand-crafted features. The two features are combined to improve the feature expression ability of infrared images. The experimental results show that the Succ. AUC and Prec@20 of the infrared target tracker ECOIR designed based on the proposed features extracting method are 79% and 65.0% on the LSOTB-TIR dataset, and 81.9% and 63.2% on the PTB-TIR dataset, respectively. It shows good robustness and tracking accuracy.

Key words: infrared target tracking; adaptive fusion; deep features; histogram of oriented gradients

红外目标跟踪不受光照变化影响, 可在一些极端天气下进行^[1-2], 广泛用于军事侦察、辅助驾驶等领域。但在复杂场景下, 红外图像存在成像质量差^[3]、目标纹理特征不清晰^[4]、易受亮目标干扰^[5]等问题, 影响目标跟踪效果。

相关滤波算法具有在线训练的优势, 鲁棒性较好, 在红外目标跟踪任务中发挥了重要作用。

Fan 等^[6]提出一种红外目标跟踪框架, 采用传统核化相关滤波 (KCF) 跟踪器和轻量级卷积神经网络 (CNN), 可更准确地获得小目标的位置和大小。Ding 等^[7]针对背景感知相关滤波 (BACF) 跟踪框架中直接引入卷积特征导致跟踪性能下降的问题, 提出一种手工特征和深度特征融合的方法, 实现了红外场景下的单行人目标跟踪。Zhang

等^[8]提出一种基于水平集方法的高效卷积算子(ECO),设计了热红外跟踪器 ECO_LS,跟踪效果较好,但其采用并行结构,无法实时运算。孙梦宇等^[9]提出一种自适应信息选择的变尺度相关滤波红外目标跟踪算法,刘刚阳等^[10]设计了一种自适应模型更新策略来改进 ECO-HC 跟踪器,但这两种算法只考虑了手工特征,未采用深度特征,特征表达不够丰富。针对海上目标跟踪,王宁等^[11]提出一种多特征融合连续卷积算子跟踪算法,其性能优于传统的连续卷积算子跟踪算法,但在复杂场景下跟踪效果有限。此外,宫华等^[12]设计了一种融合跟踪模型,采用通道注意力机制和双向特征融合策略,实现了红外多目标跟踪,熊偌炎等^[13]设计了一种多尺度特征响应融合算法来应对现有红外目标跟踪算法在复杂场景下性能退化的问题,但上述两种算法复杂度较高,需要强大的算力支撑。

针对上述问题,本文设计一种大核小波网络(large kernel & wavelet network, LKWNNet)提取深度特征,设计自适应方向梯度直方图(HOG)^[14]算

法提取手工特征,采用多特征自适应融合方法融合手工特征和深度特征,提高红外图像特征表达能力。采用上述特征提取方法设计红外目标跟踪器 ECOIR (ECO infrared),考察其在复杂场景下的红外目标跟踪效果。

1 ECOIR 红外目标跟踪器

ECOIR 跟踪算法结构如图 1 所示。首先,输入第一帧图片以初始化滤波器,对输入的第二帧及后续图片分别采用 LKWNNet 和自适应 HOG (adaptive HOG, AHOG) 算法提取深度特征和手工特征;然后,将提取的多特征和滤波器输出送入 ECO 位置跟踪模块获得响应图,再进一步经过多特征自适应融合得到目标位置,并根据浅层特征响应进行尺度响应获得最优尺度;最后,根据目标位置采用丢失判断模块判断目标是否丢失,若丢失则不做更新,若未丢失则结合最优尺度得到跟踪结果,并更新相关滤波器用于下一帧图片的响应。

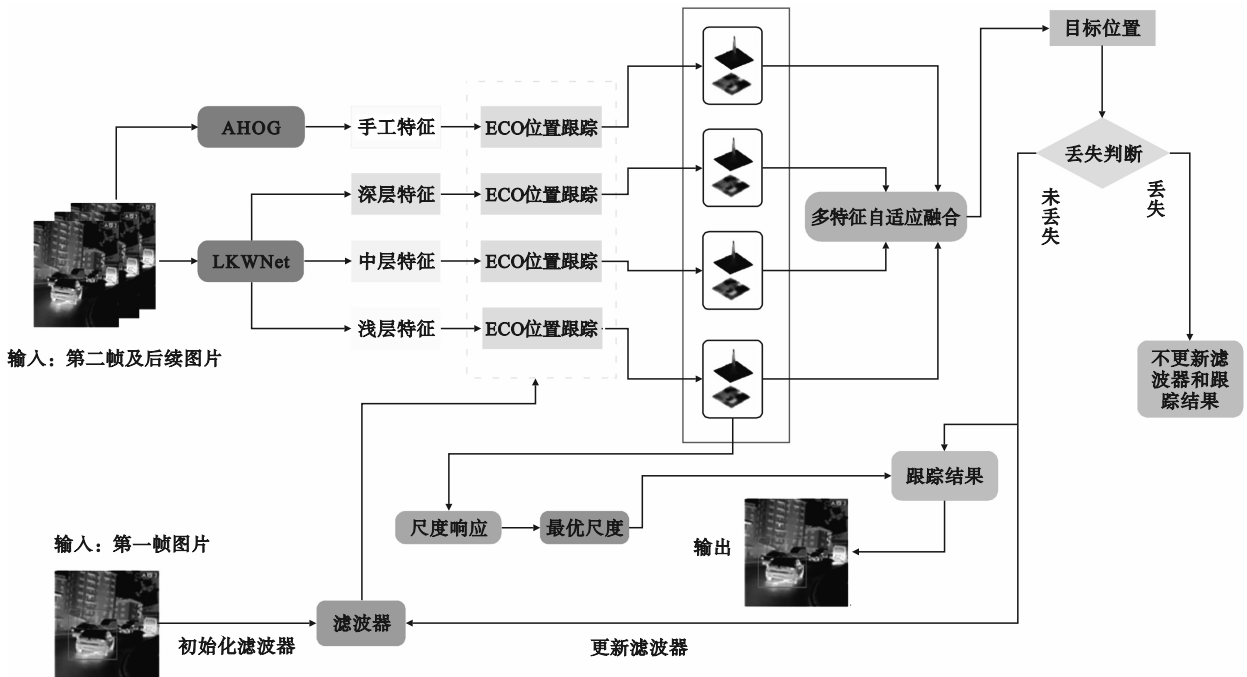


图 1 ECOIR 跟踪算法结构图

Fig. 1 The structure diagram of ECOIR tracking algorithm

1.1 LKWNNet 深度特征提取

LKWNNet 的总体结构如图 2 所示。图像序列首先经过卷积核大小为 4×4 、步长为 4 的卷积层(Conv),然后进行层归一化(Layer Normal),最后通过 LKW-Block 提取图像特征,图中 LKW-Block0、LKW-Block1 和 LKW-Block2 分别用来提

取浅层特征、中层特征和深层特征。

LKW-Block 结构如图 3 所示,其上部分为 Loop-Block,下部分为下采样(DownSample)模块。在 Loop-Block 中,首先采用 Depthwise(DW)卷积分别对各通道特征图(图中 H 、 W 、 D_{in} 分别表示特征图的高度、宽度和维度)进行卷积操作,卷

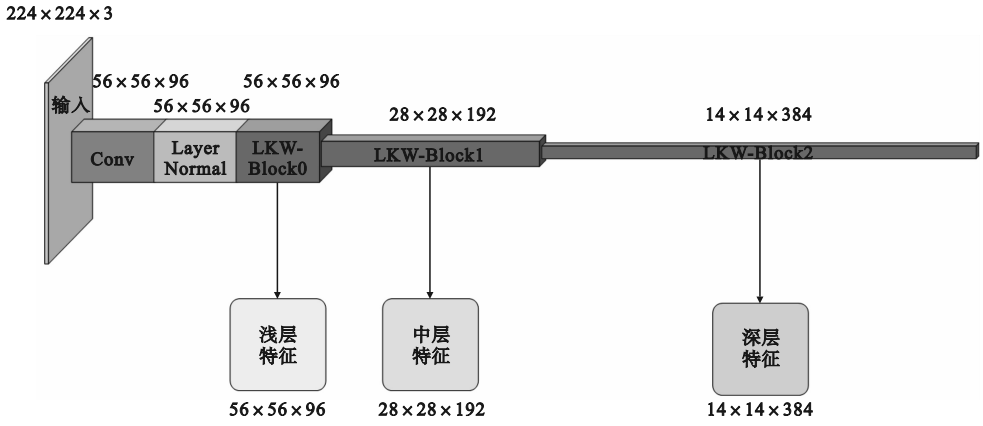


图2 LKWNet 总体结构

Fig. 2 The overall structure of LKWNet

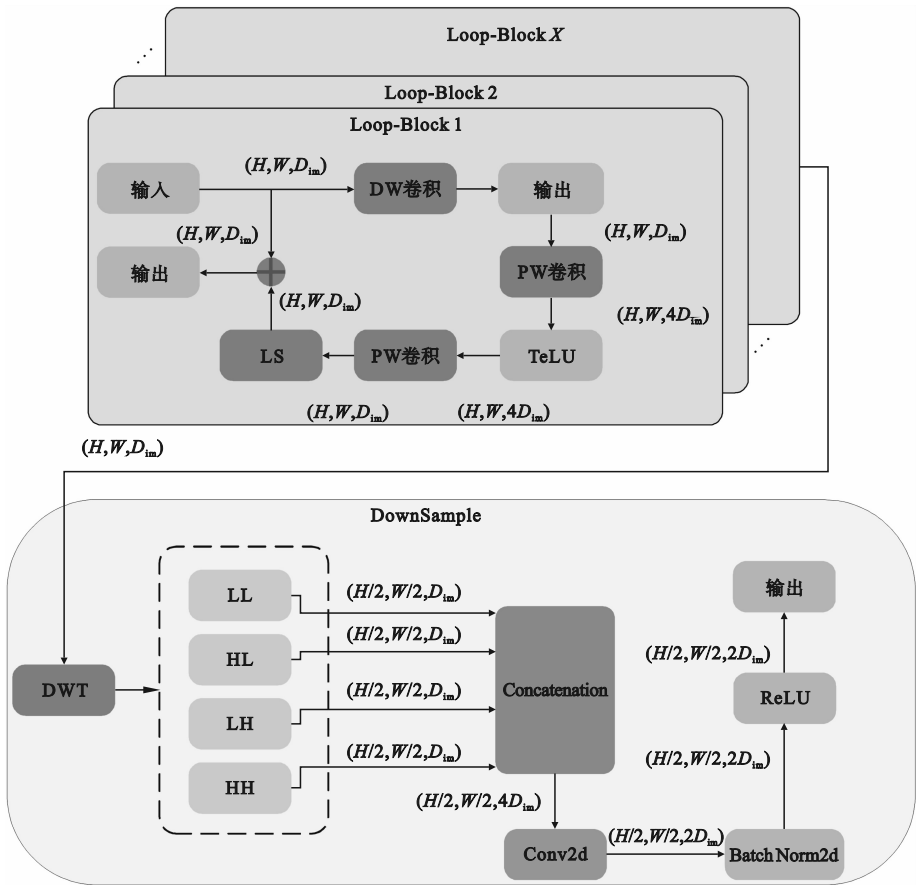


图3 LKW-Block 结构图

Fig. 3 Structure diagram of LKW-Block

积核大小为 7×7 。该设计可有效减少网络浮点运算数,同时通过大卷积核显著扩大局部感受野,增强模型对上下文信息的捕捉能力,提升局部特征的判别能力。由于DW卷积未融合通道信息,故对经过层归一化后的特征进行Pointwise(PW)卷积。第一个PW卷积采用 $4D_{in}$ 个 1×1 卷积核,以充分融合多个通道特征;第二个PW卷积采用 D_{in} 个 1×1 卷积核,目的是恢复通道数与输入一致。

在两个PW卷积之间使用TeLU激活函数^[15],用于提升模型的泛化能力和训练稳定性;缩放层(LS)作用于残差连接前的分支输入。将初始输入和LS层输出进行残差连接,便得到一个Loop-Block的输出。Loop-Block多次堆叠构成LKW-Block,LKW-Block0、LKW-Block1和LKW-Block2中Loop-Block的堆叠次数分别为3、3和9,最后一个Loop-Block的输出作为DownSample模块的

输入。

DownSample 模块使用 Haar 小波下采样^[16], 不仅可降低空间分辨率, 且可保留更多信息, 更好地适应复杂场景。在 DownSample 模块中, 首先对输入图像(尺寸为 $H \times W \times D_{im}$) 进行离散小波变换(DWT), 得到一个低频子带(LL)和三个高频子带(HL、LH 和 HH), 然后将四个子带进行合并操作(Concatenation), 再对合并后的数据进行二维卷积(Conv2d)处理, 该卷积输入维度为 $4D_{im}$, 输出维度为 $2D_{im}$, 卷积核大小设定为 1×1 , 步长为 1。完成卷积后, 对数据进行批归一化(Batch Norm2d)处理, 再通过 ReLU 激活函数后输出。此时, LKW-Block 输出图像的高和宽较其输入图像减半, 而通道数则增加了一倍。

1.2 AHOG 特征提取

本文提出一种 AHOG 特征提取方法, 用于增强红外目标特征表达能力。首先对输入的红外图像进行预处理, 以增强局部对比度并减少噪声, 预处理操作表示为

$$I_p = \text{CLAHE}(I, c_1, t_g) \quad (1)$$

式中: I_p 为预处理后的图像; $\text{CLAHE}(\cdot)$ 表示限制对比度自适应直方图均衡化, 该操作先将图像分割成多个小块, 再对每个小块分别进行直方图均衡化, 并通过限制对比度来避免噪声放大; I 为输入的红外图像; c_1 为对比度限制阈值; t_g 为分割块规格。

然后计算梯度幅值和方向, 计算表达式为

$$(M, O) = \text{gradMag}(G_\sigma(I_p)) \quad (2)$$

式中: $\text{gradMag}(\cdot)$ 表示梯度计算; M 表示计算出的梯度幅值; O 表示梯度方向; $G_\sigma(\cdot)$ 表示标准差为 σ 的高斯滤波过程。

HOG 算法中通过一个固定特征裁剪阈值进行归一化, 本文则采用自适应特征裁剪阈值, 该值根据图像的局部对比度进行调整, 以更好地适应红外图像特征提取。本文动态特征裁剪阈值计算过程表示为

$$C = \frac{\text{std}(G_\sigma(I_p))}{\text{mean}(G_\sigma(I_p))} \quad (3)$$

$$d_{\text{clip}} = \text{const}(d_{\text{clip}} \cdot (1 + \alpha \cdot C), d_{\text{limit}}) \quad (4)$$

式中: C 为预处理后图像的标准差与均值的比值; $\text{std}(\cdot)$ 和 $\text{mean}(\cdot)$ 分别表示计算标准差和均值的函数; d_{clip} 表示特征裁剪阈值; d_{clip} 为自适应特征裁剪阈值; $\text{const}(\cdot, \cdot)$ 表示限制最终特征裁剪阈值范围的函数; d_{limit} 表示特征裁剪阈值限制范围; α 为权重系数。

以计算得到的梯度幅值和方向、自适应特征

裁剪阈值作为参数, 通过 HOG 算法获得本文 AHOG 算法提取的特征, 表示为

$$H = \text{hog}(M, O, d_{\text{clip}}, v_{\text{size}}, q) \quad (5)$$

式中: H 为提取的 AHOG 特征; $\text{hog}(\cdot)$ 表示 HOG 特征提取方法; v_{size} 为空间量化步长; q 为方向通道数。

1.3 多特征自适应融合

多特征自适应融合算法的总体结构如图 4 所示。给定特征响应后, 计算其各自权重。如果计算出的权重在限定范围内, 进行加权融合, 否则采用二次规划方法重新调整权重再加权融合。加权融合后得到目标位置作为输出。

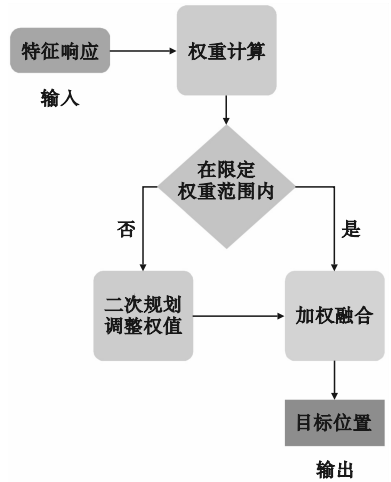


图4 多特征自适应融合算法总体结构图

Fig.4 Overall structure diagram of multi-feature adaptive fusion algorithm

本文使用 LKWNet 提取深层特征、中层特征、浅层特征, 使用 AHOG 提取手工特征, 对每个特征响应赋予不同的权重, 并进行特征融合, 融合过程表示为

$$(x_j^*, y_j^*) = \sum_{n=1}^N [w_j^n \cdot (x_j^n, y_j^n)] \quad (6)$$

式中: (x_j^*, y_j^*) 为特征融合后第 j 帧图片的目标中心位置坐标; (x_j^n, y_j^n) 为第 j 帧图片在第 n 个特征响应层中最大特征响应值对应的位置坐标; N 为特征响应层总数; w_j^n 为 (x_j^n, y_j^n) 对应的权重, $\sum_{n=1}^N w_j^n = 1$ 。

特征响应损失计算式为

$$L_j^n = \max(\mathbf{R}_j^n) - \text{res}_j^n(x_j^*, y_j^*) \quad (7)$$

式中: L_j^n 为第 j 帧图片在第 n 层的特征响应损失; \mathbf{R}_j^n 为第 j 帧图片在第 n 层的特征响应值集合; $\max(\cdot)$ 表示取最大值函数; $\text{res}_j^n(\cdot)$ 表示求解第 j 帧

图片在第 n 层某位置对应的响应值。第 j 帧图片的特征响应加权平均损失 \bar{L}_j 计算式为

$$\bar{L}_j = \sum_{n=1}^N w_j^n L_j^n \quad (8)$$

后悔度 r_j^n 定义式为

$$r_j^n = \bar{L}_j - L_j^n \quad (9)$$

第 j 帧图片在第 n 层的特征响应损失在帧间隔 Δj 内的均值 μ_j^n 为

$$\mu_j^n = \frac{1}{\Delta j} \sum_{\rho=j-\Delta j+1}^j L_\rho^n \quad (10)$$

在 Δj 内的标准差 σ_j^n 为

$$\sigma_j^n = \sqrt{\frac{1}{\Delta j - 1} \sum_{\rho=j-\Delta j+1}^j (L_\rho^n - \mu_j^n)^2} \quad (11)$$

对损失进行归一化为

$$g_j^n = \frac{|L_j^n - \mu_j^n|}{\sigma_j^n} \quad (12)$$

式中 g_j^n 为第 j 帧图片在第 n 层的特征响应损失归一化后的值,用来衡量特征响应在 Δj 内的稳定性。 g_j^n 越小,该层的目标特征响应越稳定。

自适应累积后悔度计算过程表示为

$$R_j^n = (1 - \beta_j^n) R_j^n - 1 + \beta_j^n r_j^n \quad (13)$$

$$\beta_j^n = \min(r_g, e^{-\theta g_j^n}) \quad (14)$$

式中: R_j^n 为第 j 帧图片在第 n 层的自适应累积后悔度; β_j^n 为一个系数,可使 g_j^n 较小的响应层拥有较大的权值; r_g 表示当前后悔的最大比率; θ 为比例因子; $\min(\cdot)$ 表示最小值函数。

第 $j+1$ 帧图片在第 n 层的特征响应权重 w_{j+1}^n 计算表达式为

$$w_{j+1}^n \propto \frac{\max(0, R_j^n + A)}{c_j} \cdot \exp\left[\frac{(\max(0, R_j^n + A))^2}{2c_j}\right] \quad (15)$$

式中: \propto 表示比例符号; $\max(\cdot)$ 表示求最大值; A 为一个平滑正则项,用来防止更新过程中权重趋于零,避免某些特征响应层因为短期表现不佳而在后续帧图片中被完全丢弃,提升跟踪器在长期序列和外观快速变化场景下的鲁棒性; c_j 为尺度参数,其求解约束表示为

$$\frac{1}{N} \sum_{n=1}^N \exp\left[\frac{(\max(0, R_j^n + A))^2}{2c_j}\right] = e \quad (16)$$

如式(16)所示,本文首先通过指数缩放确定搜索空间,再通过二分法求解在区间内收敛至满足精度的 c_j 值。

鉴于各层特征及手工特征在红外图像中对目标定位的贡献不同,本文设定浅层特征、中层特

征、深层特征和手工特征响应最终目标位置影响的权重可行域。当求得的权重 w_{j+1}^n 落在可行域内时, w_{j+1}^n 即为最终权重。设计权重可行域约束,可防止权重退化或者主导,增强算法的鲁棒性和容错能力。

考虑到 w_{j+1}^n 不在可行域的情况,本文构造一个二次规划问题如下。

$$\begin{cases} \min \| \mathbf{W} - \mathbf{W}_0 \|^2 + \lambda \| \mathbf{W} \|^2 \\ \text{s. t. } \text{sum}(\mathbf{W}) = 1 \\ \mathbf{W}_{\text{lb}} \leq \mathbf{W} \leq \mathbf{W}_{\text{ub}} \end{cases} \quad (17)$$

式中: $\mathbf{W}_0 = \{w_{j+1}^1, \dots, w_{j+1}^N\}$; \mathbf{W} 为新权重矩阵; $\text{sum}(\mathbf{W})$ 表示计算 \mathbf{W} 中所有元素的和; \mathbf{W}_{lb} 和 \mathbf{W}_{ub} 分别为可行域下限和上限所构成的矩阵; λ 为正则化参数,旨在控制解的复杂度。式(17)中对权重向量 \mathbf{W} 施加 L2 范数惩罚,可有效抑制权重的剧烈波动,防止模型因追求完美拟合当前帧的响应而产生过拟合,在满足约束条件的前提下,控制最终权重 \mathbf{W} 与原始权重 \mathbf{W}_0 之间的权衡关系。通过求解二次规划问题,可得到一个新的落在可行域内的 \mathbf{W} ,将其作为融合时各特征的权重。

权重计算过程如图5所示。首先,输入深层、中层、浅层和手工特征响应和响应值,计算损失;然后,使用上一帧生成的权重对损失进行加权求和(从第二帧开始计算损失,其权重采用第一帧初始化的权重),再计算各层响应在一个时间段内所有损失的均值和标准差,用于衡量特征层的稳定性。特征层稳定性衡量结果用于后悔度计算(从第二帧开始计算后悔度,第二帧的后悔度采用第一帧初始化的后悔度),使用后悔度和参数 A 更新权重,更新后的权重作为加权融合模块的输出权重,并用于下一次加权损失计算。

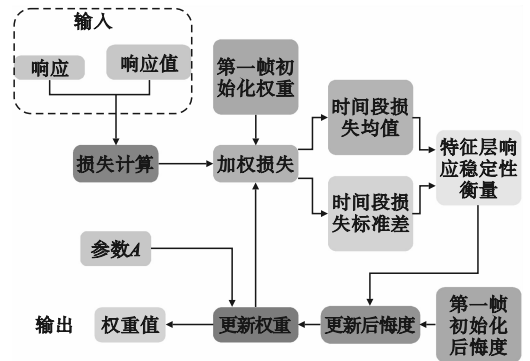


图5 权重计算过程

Fig. 5 Weight calculation process

由于复杂场景下易出现遮挡后跟错目标的情况,本文采用平均峰值相关能量(average peak-to-

correlation energy, APCE) 和峰值旁瓣比 (peak-to-sidelobe ratio, PSR) 组合判定目标是否丢失。APCE 和 PSR 的计算式分别为

$$S_{\text{apce}} = \frac{|f_{\text{max}} - f_{\text{min}}|^2}{\frac{1}{B} \sum_{b=1}^B (f_b - f_{\text{min}})^2} \quad (18)$$

$$S_{\text{psr}} = \frac{f_{\text{max}} - \text{mean}(\mathbf{F})}{\text{std}(\mathbf{F})} \quad (19)$$

式中: S_{apce} 和 S_{psr} 分别为平均峰值相关能量和峰值旁瓣比; \mathbf{F} 为特征融合后的响应; f_{max} 和 f_{min} 分别为响应的最大值和最小值; B 为特征融合后响应的响应值总数; f_b 表示第 b 个响应值。

由于单帧图片的 APCE 与 PSR 值易受噪声影响, 本文采用指数加权移动平均策略对其进行平滑处理, 以反映目标跟踪状态的长期趋势。具体而言, 若当前帧的 APCE 和 PSR 值低于其历史均值的给定阈值倍数, 则判定目标丢失, 暂停模型更新, 避免引入错误样本污染滤波器。反之, 则更新滤波器, 并使用当前帧图片的 APCE 和 PSR 值以学习率 α 在线更新历史均值, 即

$$\overline{S_{\text{apce}}} = S_{\text{apce}} \cdot \alpha + \overline{S_{\text{apce}}} \cdot (1 - \alpha) \quad (20)$$

$$\overline{S_{\text{psr}}} = S_{\text{psr}} \cdot \alpha + \overline{S_{\text{psr}}} \cdot (1 - \alpha) \quad (21)$$

式中: $\overline{S_{\text{apce}}}$ 为 S_{apce} 的历史均值; $\overline{S_{\text{psr}}}$ 为 S_{psr} 的历史均值。该机制可有效平衡响应的即时性与历史稳定性, 增强跟踪器对目标在长时遮挡和外观剧烈变化下的恢复能力。

2 实验结果及分析

2.1 实验环境

本文实验环境配置如表 1 所示。

表 1 实验环境

Table 1 Experimental environments

实验环境	参数配置
操作系统	Ubuntu 18.04.6
GPU	NVIDIA GeForce RTX 4090
CPU	Intel(R) Xeon(R) Silver 4314 CPU @ 2.40 GHz
深度学习框架	torch1.12.0 + cu113 torchvision0.13.0 + cu113 torchaudio0.12.0
编程环境	python3.8

2.2 数据集和评价指标

本文使用 LSOTB-TIR^[17] 和 PTB-TIR 作为实

验基准数据集。评价指标采用 Prec@20 和 Succ. AUC; Prec@20 用来衡量模型的精确率, Succ. AUC 用来衡量跟踪器在整个跟踪序列中预测的目标包围盒与真实目标包围盒的重叠质量。Prec@20 计算式为

$$d_j = \sqrt{(x_j^* - x_j)^2 + (y_j^* - y_j)^2} \quad (22)$$

$$\text{Prec@}\tau = \frac{J_\tau}{J} \quad (23)$$

式中: d_j 为第 j 帧图像中预测目标中心位置 (x_j^*, y_j^*) 与目标真实中心位置 (x_j, y_j) 之间的距离, 一般认为 d_j 不超过 20 个像素值时即表明该帧目标跟踪成功; τ 为精确度的指标值, $\tau \in \{0, 0.5, 1.0, \dots, 49.5, 50.0\}$; Prec@ τ 表示 τ 对应的精确率; J_τ 为跟踪序列中满足 $d_j < \tau$ 的帧数; J 为跟踪序列的总帧数。本文采用 $\tau = 20$ 的 Prec@ τ 作为目标跟踪的精确率, 表示为 Prec@20。

Succ. AUC 计算式为

$$\text{Succ. AUC} = \int_0^1 \frac{J_\zeta}{J} d\zeta \quad (24)$$

式中: ζ 为交并比的阈值指标值, $\zeta \in \{0, 0.01, 0.02, \dots, 1.00\}$; J_ζ 为跟踪序列中交并比大于 ζ 的帧数。

2.3 对比实验

通过 THOP 对本文设计的 LKWNet 进行测试, 并与 FasterNet-S 算法^[18] 进行对比。测试得到的参数量 (Params) 和计算量 (用浮点运算数 GFLOPs 表示) 如表 2 所示, LKWNet 在复杂度方面优于 FasterNet-S, 其参数量与计算量均低于 FasterNet-S。

表 2 测试结果对比

Table 2 Comparison of test results

网络	GFLOPs	Params/10 ⁶
LKWNet	3.70	12.3
FasterNet-S	4.55	31.1

将本文设计的 ECOIR 跟踪器与其他常见跟踪器在 LSOTB-TIR 和 PTB-TIR 数据集上进行对比实验, 结果如表 3 所示。与 ECO 跟踪器相比较, 本文设计的 ECOIR 跟踪器在 LSOTB-TIR 数据集上的 Prec@20 和 Succ. AUC 分别提升了 5.47% 和 5.69%, 在 PTB-TIR 数据集上的 Prec@20 和 Succ. AUC 分别提升了 0.24% 和 1.44%。本文 ECOIR 跟踪器在两个基准数据集上的执行速度 (用帧率 (FPS) 表示) 相比于 ECO 跟踪器有所下降, 但仍在 25 以上。与 EASATrack 相比较, 本文 ECOIR 跟踪器在 LSOTB-TIR 和 PTB-TIR 数据集

上的 Prec@20/Succ. AUC 分别提高了 6.61%/4.67% 和 3.15%/2.43%。与 TFF-Siam 相比较,本文 ECOIR 跟踪器在 LSOTB-TIR 和 PTB-TIR 数据集上

的 Prec@20/Succ. AUC 分别提高了 3.13%/1.56% 和 4.73%/6.76%。本文 ECOIR 跟踪器在 Prec@20 和 Succ. AUC 两项指标上均优于其他对比跟踪器。

表3 在 LSOTB-TIR 和 PTB-TIR 基准数据集上的对比实验结果

Table 3 Results of comparative experiments on LSOTB-TIR and PTB-TIR benchmarks

跟踪器	LSOTB-TIR			LSOTB-TIR		
	Prec@20	Succ. AUC	FPS	Prec@20	Succ. AUC	FPS
ECO	0.749	0.615	57.5	0.817	0.623	51.3
EASATrack ^[19]	0.741	0.621		0.794	0.617	
TFF-Siam ^[20]	0.766	0.640		0.782	0.592	
ECOIR	0.790	0.650	27.4	0.819	0.632	27.8

2.4 参数实验与消融实验

为探究不同特征层在不同权重范围设置下的性能表现,确定最优的权重配置策略,采用本文 ECOIR 跟踪器对浅层特征、中层特征、深层特征和 AHOG 特征融合的权重范围进行评估,结果如表 4 所示,表中上角标注 a、b、c、d 分别表示浅层特征、中层特征、深层特征和 AHOG 特征。由表 4 可见:当浅层特征、中层特征、深层特征、AHOG 特征的融合权重范围分别为 (0.2, 0.6)、(0.15,

0.5)、(0.15, 0.4)、(0.05, 0.2) 时,本文 ECOIR 跟踪器在 LSOTB-TIR 数据集上表现最好;当上述四种特征的融合权重范围分别为 (0.4, 0.8)、(0.2, 0.5)、(0.15, 0.5)、(0.15, 0.2) 时,本文 ECOIR 跟踪器在 PTB-TIR 数据集上可获得较高的 Prec@20;当上述四种特征的融合权重范围分别为 (0.3, 0.8)、(0.2, 0.5)、(0.15, 0.5)、(0.05, 0.2) 时,ECOIR 跟踪器在 PTB-TIR 数据集上可获得较高的 Succ. AUC。

表4 不同特征层的融合权重范围评估结果

Table 4 Evaluation results for fusion weight ranges of different feature layers

		不同特征层的融合权重范围评估结果					
数据集	指标	(0.25,0.7) ^a	(0.2,0.6) ^a	(0.3,0.9) ^a	(0.2,0.8) ^a	(0.3,0.8) ^a	(0.4,0.8) ^a
		(0.2,0.6) ^b	(0.15,0.5) ^b	(0.2,0.5) ^b	(0.15,0.5) ^b	(0.2,0.5) ^b	(0.2,0.5) ^b
		(0.15,0.5) ^c	(0.15,0.4) ^c	(0.1,0.3) ^c	(0.15,0.5) ^c	(0.15,0.5) ^c	(0.15,0.5) ^c
		(0.1,0.2) ^d	(0.05,0.2) ^d	(0.1,0.3) ^d	(0.01,0.1) ^d	(0.05,0.2) ^d	(0.15,0.2) ^d
LSOTB-TIR	Prec@20	0.761	0.787	0.739	0.771	0.750	0.752
	Succ. AUC	0.627	0.646	0.621	0.639	0.627	0.626
	FPS	29.0	29.0	29.1	28.9	28.9	29.0
PTB-TIR	Prec@20	0.792	0.802	0.805	0.798	0.805	0.811
	Succ. AUC	0.614	0.621	0.617	0.614	0.624	0.622
	FPS	27.0	27.0	27.1	27.1	27.1	27.1

为防止特征权重因短期跟踪性能波动而过度衰减,本文引入正则项 A 来提升跟踪器的鲁棒性。不同 A 值下 ECOIR 跟踪器在 LSOTB-TIR 数据集

上的实验结果如表 5 所示,A 值设置为 0.10 时,ECOIR 的性能最好。

表5 不同 A 值下 ECOIR 在 LSOTB-TIR 数据集上的实验结果

Table 5 Experimental results of ECOIR with different values of parameter A on LSOTB-TIR

指标	A=0.01	A=0.05	A=0.10	A=0.12	A=0.14	A=0.16
Prec@20	0.748	0.778	0.785	0.757	0.739	0.738
Succ. AUC	0.620	0.638	0.640	0.631	0.615	0.615
FPS	28.7	28.7	28.2	28.3	28.5	28.3

分别采用深度特征提取模块 LKWNet、手工特征提取模块 AHOG 及同时采用上述两种特征提取模块构建红外目标跟踪器,进行消融实验,结果如表 6 所示。

表 6 消融实验结果

Table 6 Results of ablation experiments

特征提取模块	LSOTB-TIR		PTB-TIR	
	Prec@20	Succ. AUC	Prec@20	Succ. AUC
LKWNet	0.754	0.624	0.818	0.625
AHOG	0.660	0.544	0.738	0.600
LKWNet + AHOG	0.790	0.650	0.819	0.632

由表 6 可见:单独采用 AHOG 进行红外目标跟踪时的性能明显弱于单独采用 LKWNet 时的跟踪性能;相比于单独采用 LKWNet 进行红外目标跟踪,采用融合特征的跟踪器在两个数据集上的 Prec@20 和 Succ. AUC 均有提升,其中 LSOTB-TIR

数据集上的提升更明显,Prec@20 和 Succ. AUC 分别提升了 4.77% 和 4.17%,表明采用融合策略能有效增强特征表达能力。

使用 LKWNet 各层特征、AHOG 特征以及融合 AHOG 与 LKWNet 各层特征的 ECOIR 跟踪器在不同序列上的跟踪结果如图 6 所示。可见:使用浅层特征和 AHOG 特征的跟踪器擅长捕捉细节和边缘,但易受背景干扰,在背景复杂的场景中容易偏离目标,如图 6(a)、6(b)所示,或者被相似目标干扰,如图 6(d)、6(f)所示;深层特征表达高层语义信息,具有较强的抗干扰能力,但其经过多次下采样,空间尺寸大幅缩小,导致边界模糊、细节丢失,跟踪器难以精确框定目标,如图 6(c)所示,且对于远距离目标,采用深层特征的跟踪器难以有效识别,如图 6(a)、6(e)所示;采用融合特征的跟踪器可根据场景特点调整各层特征的权值,充分发挥各层特征的优势,具有较好的跟踪效果。



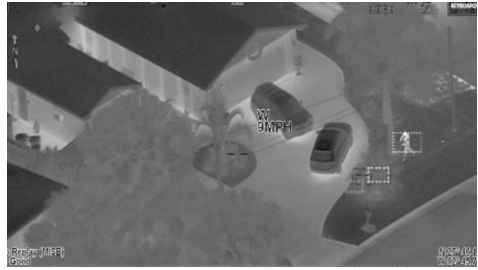
(a) car_S_003序列(LSOTB-TIR)



(b) street_S_005序列(LSOTB-TIR)



(c) Person_D_009序列(LSOTB-TIR)



(d) dog_D_002序列(LSOTB-TIR)



(e) crowd1序列(PTB-TIR)



(f) Distractor2序列(PTB-TIR)

真实跟踪结果
 使用浅层特征跟踪结果
 使用中层特征跟踪结果
 使用深层特征跟踪结果
 使用AHOG特征跟踪结果
 融合LKWNet和AHOG特征的跟踪结果

图 6 采用不同特征的 ECOIR 跟踪器在 LSOTB-TIR 和 PTB-TIR 上的跟踪情况

Fig. 6 Tracking of ECOIR trackers using different feature methods on LSOTB-TIR and PTB-TIR

3 结论

红外图像存在成像质量差、目标纹理特征不清晰、易受亮目标干扰等问题,在复杂场景下红外目标跟踪效果不佳。本文设计了基于 DW 大卷积核的特征提取模块,通过扩大感受野增强了上下文信息的捕获能力,同时结合 Haar 小波下采样构建了 LKWNNet 红外特征提取网络;基于局部增强和自适应特征裁剪阈值设计了 AHOG 特征提取方法,提升了弱边缘与弱纹理的可检测性;使用二次规划方法实现了多特征的自适应融合权重计算,提高了融合策略在复杂环境下的适应性;采用 APCE 和 PSR 对红外目标进行丢失判断,增强了跟踪器在背景干扰、目标遮挡及形变情况下的鲁棒性。实验结果表明,本文基于多特征融合的红外目标跟踪方法在复杂场景下具有良好的适用性。

参考文献 (References):

- [1] Yuan D, Zhang H P, Shu X, et al. Thermal infrared target tracking: a comprehensive review [J]. IEEE Transactions on Instrumentation and Measurement, 2023, 73: 5000419.
- [2] 李清忠. 基于两阶段时空加权特征的红外目标跟踪算法 [J]. 红外技术, 2025, 47(4): 437 - 444.
Li Q Z. Infrared object tracking algorithm based on two-stage spatiotemporal weighted features [J]. Infrared Technology, 2025, 47(4): 437 - 444. (in Chinese)
- [3] Zhang W X, Huang B R, Meng S J, et al. Infrared fast-maneuvering target tracking based on robust exact differentiator with improved particle filter [J]. Journal of Robotics, 2022, 2022(1): 6406672.
- [4] Wu F Y, Liu A R, Zhang T F, et al. Saliency at the helm: steering infrared small target detection with learnable kernels [J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 63: 5000514.
- [5] Wang H, Hu Y, Wang Y, et al. Infrared small target detection based on weighted improved double local contrast measure [J]. Remote Sensing, 2024, 16(21): 4030.
- [6] Fan J, Wei J B, Huang H, et al. IRSDT: a framework for infrared small target tracking with enhanced detection [J]. Sensors, 2023, 23(9): 4240.
- [7] Ding M, Chen W H, Cao Y F. Thermal infrared single-pedestrian tracking for advanced driver assistance system [J]. IEEE Transactions on Intelligent Vehicles, 2023, 8(1): 814 - 824.
- [8] Zhang H, Yin Z Y, Zhang H L. Thermal infrared object tracking using correlation filters improved by level set [J]. Signal, Image and Video Processing, 2023, 17(3): 791 - 797.
- [9] 孙梦宇, 王鹏, 徐均琪, 等. 自适应信息选择的变尺度相关滤波红外目标跟踪算法 [J]. 光子学报, 2023, 52(12): 226 - 238.
Sun M Y, Wang P, Xu J Q, et al. Adaptive information selec-

- tion for infrared object tracking with variable scale correlation filter [J]. Acta Photonica Sinica, 2023, 52(12): 226 - 238. (in Chinese)
- [10] 刘刚阳, 胡博, 王宇庆. 基于改进 ECO-HC 的目标跟踪方法 [J]. 液晶与显示, 2023, 38(8): 1118 - 1127.
Liu G Y, Hu B, Wang Y Q. Target tracking method based on improved ECO-HC [J]. Chinese Journal of Liquid Crystals and Displays, 2023, 38(8): 1118 - 1127. (in Chinese)
- [11] 王宁, 吴伟, 王元元, 等. 多特征融合的无人艇视觉小目标鲁棒跟踪 [J]. 中国舰船研究, 2024, 19(5): 65 - 78.
Wang N, Wu W, Wang Y Y, et al. Multi-feature fusion-based robust tracking of small targets in unmanned surface vehicle vision [J]. Chinese Journal of Ship Research, 2024, 19(5): 65 - 78. (in Chinese)
- [12] 宫华, 张众垚, 胡雨桐, 等. 基于改进 YOLOv7-DeepSort 的红外视频多目标跟踪 [J]. 沈阳理工大学学报, 2024, 43(6): 20 - 27.
Gong H, Zhang Z Y, Hu Y T, et al. Infrared video multi target tracking based on improved YOLOv7-DeepSort [J]. Journal of Shenyang Ligong University, 2024, 43(6): 20 - 27. (in Chinese)
- [13] 熊熊炎, 张上, 张岳. 多尺度特征响应融合的复杂场景红外目标跟踪 [J]. 计算机工程与应用, 2025, 61(12): 243 - 257.
Xiong R Y, Zhang S, Zhang Y. Multi-scale feature response fusion for complex scene infrared target tracking [J]. Computer Engineering and Applications, 2025, 61(12): 243 - 257. (in Chinese)
- [14] Dalal N, Triggs B. Histograms of oriented gradients for human detection [C] // 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR2005). San Diego, CA, USA: IEEE, 2005: 886 - 893.
- [15] Fernandez A, Mali A. TeLU activation function for fast and stable deep learning [PP/OL]. arXiv (2024 - 12 - 28) [2025 - 07 - 10]. <https://doi.org/10.48550/arXiv.2412.20269>.
- [16] Xu G P, Liao W T, Zhang X, et al. Haar wavelet downsampling: a simple but effective downsampling module for semantic segmentation [J]. Pattern Recognition, 2023, 143: 109819.
- [17] Liu Q, Li X, Yuan D, et al. LSOTB-TIR: a large-scale high-diversity thermal infrared single object tracking benchmark [J]. IEEE Transactions on Neural Networks and Learning Systems, 2024, 35(7): 9844 - 9857.
- [18] Chen J, Kao S H, He H, et al. Run, don't walk: chasing higher FLOPS for faster neural networks [C] // 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, BC, Canada: IEEE, 2023: 12021 - 12031.
- [19] Wang B R, Yuan J, Zhang X B, et al. EASATrack: infrared object tracking based on edge activation and scale adaptive cross-correlation [C] // 2023 5th International Conference on Robotics, Intelligent Control and Artificial Intelligence (RIC-AI). Hangzhou, China: IEEE, 2023: 801 - 805.
- [20] Song J F, Xu H, Ren G D, et al. Infrared object-tracking algorithm with temporal features fusion Siamese network [J]. IEEE Transactions on Instrumentation and Measurement, 2024, 73: 1 - 13.

(责任编辑: 宋颖韬)