

文章编号:1007-2780(2019)02-0177-11

基于多特征融合相关滤波的红外目标跟踪

韩亚君, 杨德东*, 李 勇, 李雪晴

(河北工业大学 人工智能与数据科学学院, 天津 300130)

摘要:针对红外目标分辨率低、对比度差、信噪比低、纹理信息缺失等特点,提出一种融合多特征的红外目标跟踪算法。利用背景感知相关滤波器生成大量真实样本,对红外目标提取 HOG 特征和运动特征,通过线性求和方式进行特征融合,更好地发挥各自特征优势,实现对红外目标运动的精准跟踪。另外,提出使用空间加权窗代替传统相关滤波器中的余弦窗,可以更加突出目标的中心位置,同时也能很好地抑制边缘效应。采用 VOT-TIR 2016 数据集对算法性能进行评估,同时和 15 种流行算法进行比较。结果表明,本文算法在精确度和成功率上的得分分别为 0.751 和 0.697,在精确度和成功率指标方面分别提高了 8.8% 和 15.4%,具有一定的研究价值。

关键词:红外目标跟踪;背景感知;特征融合;空间窗加权

中图分类号:TP391.4 **文献标识码:**A **doi:**10.3788/YJYXS20193402.0177

Infrared target tracking based on correlation filter with multi-features fusion

HAN Ya-jun, YANG De-dong*, LI Yong, LI Xue-qing

(School of Artificial Intelligence, Hebei University of Technology, Tianjin 300130, China)

Abstract: An infrared target tracking algorithm with multi-features was proposed in consideration of low resolution, poor contrast, low signal-to-noise ratio and lack of texture information of infrared target. The background perceptual correlation filter was used to generate a large number of real samples, and the HOG feature and motion feature were extracted for the infrared target. The feature fusion was performed by linear interpolation, and the advantages of the respective feature were well utilized to achieve accurate tracking of the infrared target motion. In addition, it was proposed to adopt the spatial weighting window instead of the cosine window in the traditional correlation filter to highlight the center position of the target and suppressed the edge effect. The VOT-TIR 2016 dataset was utilized to evaluate algorithm performance in comparison it with 15 popular algorithms. Simulation results show that the algorithm's scores on accuracy and success rate are 0.751 and 0.697 respectively. Furthermore, it is 8.8% and 15.4% higher than the second-ranking algorithm, which shows that the proposed algorithm has certain research value.

Key words: infrared target tracking; background-aware; feature fusion; spatial window weighting

收稿日期:2018-07-17;修订日期:2018-10-13.

基金项目:河北省自然科学基金(F2017202009)

Supported by Natural Science Foundation of Hebei Province(No.F2017202009)

* 通信联系人, E-mail: ydd12677@163.com

1 引言

红外目标跟踪在生活以及军事方面有着广泛的应用,具有研究价值。因其自身所具有的分辨率低、对比度差、信噪比低、纹理信息缺失等特点,同时在目标跟踪过程中又存在着目标形变、遮挡、快速运动、旋转等需要解决的问题,从而使得红外目标跟踪成为一项较难的挑战。

红外目标跟踪算法最初是由可见光目标跟踪算法发展而来的,而在可见光目标跟踪算法中,核相关滤波器^[1](KCF)、空间正则化判别式相关滤波器^[2](SRDCF)和上下文感知滤波器^[3](CA-CF)是相关滤波算法中兼备较好的实时性和有效性的跟踪算法,因此,有许多学者在此基础上进行了算法改进并应用于红外目标跟踪,从而解决红外目标跟踪过程中的问题。赵东^[4]等提出一种基于图像引导滤波和核相关滤波的红外弱小目标跟踪算法,可以有效区分背景边缘和红外弱小目标。Asha^[5]等提出了一种基于判别法和生成法相结合的红外图像目标跟踪方法。利用空间特征的核相关滤波器和像素强度特征的 AdaBoost 分类器两种互补方法进行并行运算,在通过判别方法获得优化位置后,通过线性搜索方法来确定最佳目标位置。刘先红^[6]等将引导滤波、高斯低通滤波与非下采样方向滤波器组相结合,提出一种新的图像融合方法。杨福才^[7]等提出基于稀疏编码直方图(HSC)特征和扰动感知模型(DAM)的红外目标跟踪方法,利用扰动感知模型来减轻背景干扰影响。Gao 等^[8]提出了一种结合背景自相关特征和目标在时空域的共性特征的红外目标协同检测模型,通过对 Logistic Sigmoid 函数构造非线性权重和在局部加权熵设置入口权重方法来抑制背景,然后对单个帧的密集和稀疏重构映射进行位置关联,提高区分目标和背景能力。

目标跟踪算法跟踪成功的关键之处在于特征的提取。方向梯度直方图(HOG)特征利用梯度信息,可以很好地描述图像局部的形状信息,对平移、旋转和光照变化鲁棒性强。运动特征信息能够描述目标的大致运动方向,是描述目标的一个重要视觉特征^[9]。位移光流法对大目标的运动变化鲁棒性强。

受上述算法和背景感知相关滤波器跟踪算

法^[10]的启发,本文提出一种基于多特征融合的稳健红外目标算法,并利用空间加权窗代替余弦窗来抑制边缘效应。为了更好地验证本文算法的性能,在 VOT-TIR2016 数据集上进行测试,同时与 15 种流行算法进行比较。实验结果表明本文提出算法在精确度和成功率上的得分都得到了提升,具有一定的研究价值。

2 本文算法

针对红外目标跟踪任务中存在的问题,提出多特征融合的方法对红外目标进行跟踪,使用线性求和方式对 HOG 特征和运动特征的响应图进行综合,并利用空间窗加权方法代替余弦窗,更好地突出跟踪目标中心的特征,抑制边缘效应。

2.1 跟踪算法

在空间域中,多通道背景感知滤波器模板^[10]可通过最小化以下目标函数来实现:

$$E(\mathbf{h}) = \frac{1}{2} \sum_{i=1}^Q \left\| \mathbf{y}(i) - \sum_{k=1}^K \mathbf{h}_k^T \mathbf{P} \mathbf{x}_k [\Delta \tau_i] \right\|_2^2 + \frac{\lambda}{2} \sum_{k=1}^K \left\| \mathbf{h}_k \right\|_2^2, \quad (1)$$

其中: Q 为训练样本数, K 是特征通道数, $\mathbf{y} \in \mathbf{R}^T$ 是期望响应值, T 是滤波器模板维数, $\mathbf{x}_k \in \mathbf{R}^T$ 和 $\mathbf{h}_k \in \mathbf{R}^D$ 分别为第 k 个通道的目标特征和滤波器 ($T \gg D$), $\mathbf{P}$ 是 $D \times T$ 维的二进制矩阵,用于对目标特征 $\mathbf{x}$ 进行裁剪, τ 为矩阵转置符号, λ 为正则化系数。

为了提高计算效率,可将上式变换到频域中进行相关计算。在频域内,目标函数可表示为:

$$E(\mathbf{h}, \hat{\mathbf{g}}) = \frac{1}{2} \left\| \hat{\mathbf{y}} - \hat{\mathbf{X}} \hat{\mathbf{g}} \right\|_2^2 + \frac{\lambda}{2} \left\| \mathbf{h} \right\|_2^2$$

$$s.t. \hat{\mathbf{g}} = \sqrt{T} (\mathbf{F} \mathbf{P}^T \otimes \mathbf{I}_K) \mathbf{h}, \quad (2)$$

其中, $\hat{\mathbf{g}}$ 是一个辅助变量, $\hat{\mathbf{X}} = \left[\text{diag}(\hat{\mathbf{x}}_1)^T, \dots, \text{diag}(\hat{\mathbf{x}}_K)^T \right]$ 表示为 $T \times KT$ 维的矩阵, $\mathbf{h} = [\mathbf{h}_1^T, \dots, \mathbf{h}_K^T]^T$ 为 $KD \times 1$ 维的矩阵, $\hat{\mathbf{g}} = [\hat{\mathbf{g}}_1^T, \dots, \hat{\mathbf{g}}_K^T]^T$ 为 $KT \times 1$ 维的矩阵, $\mathbf{I}_K$ 为 $K \times K$ 维的单位矩阵。 $\otimes$ 为 Kronecker 积运算, $\hat{*} = \sqrt{T} \mathbf{F} *$ 表示离散傅里叶变换(DFT),其中 $\mathbf{F}$ 是标准正交 $T \times T$ 维矩阵,可把任何 T 维向量化信号映射到傅里叶域中。

为了求解公式(2),可用增广拉格朗日函数进行扩展,可得:

$$L(\hat{\mathbf{g}}, \hat{\mathbf{h}}, \hat{\boldsymbol{\zeta}}) = \frac{1}{2} \left\| \hat{\mathbf{y}} - \hat{\mathbf{X}} \hat{\mathbf{g}} \right\|_2^2 + \frac{\lambda}{2} \left\| \hat{\mathbf{h}} \right\|_2^2 + \hat{\boldsymbol{\zeta}}^T (\hat{\mathbf{g}} - \sqrt{T} (\mathbf{F} \mathbf{P}^T \otimes \mathbf{I}_K) \hat{\mathbf{h}}) + \frac{\mu}{2} \left\| \hat{\mathbf{g}} - \sqrt{T} (\mathbf{F} \mathbf{P}^T \otimes \mathbf{I}_K) \hat{\mathbf{h}} \right\|_2^2, \quad (3)$$

其中: μ 为惩罚因子, $\hat{\boldsymbol{\zeta}} = [\hat{\zeta}_1^T, \dots, \hat{\zeta}_K^T]^T$ 是频域内 $KT \times 1$ 维的拉格朗日系数矩阵。式(3)可由 ADMM 迭代优化求解,得到 $\hat{\mathbf{h}}^*$ 和 $\hat{\mathbf{g}}^*$ 两个子问题,继续用迭代优化算法对子问题求解,可得:

$$\mathbf{h} = \left(\mu + \frac{\lambda}{\sqrt{T}} \right)^{-1} (\mu \mathbf{g} + \boldsymbol{\zeta}), \quad (4)$$

$$\hat{\mathbf{g}}(t)^* = \frac{1}{\mu} (T \hat{\mathbf{y}}(t) \hat{\mathbf{x}}(t) - \hat{\boldsymbol{\zeta}}(t) + \mu \hat{\mathbf{h}}(t)) - \frac{\hat{\mathbf{x}}(t)}{\mu b} (T \hat{\mathbf{y}}(t) \hat{\mathbf{s}}_x(t) - \hat{\mathbf{s}}_z(t) + \mu \hat{\mathbf{s}}_h(t)), \quad (5)$$

其中 $\mathbf{g} = \frac{1}{\sqrt{T}} (\mathbf{P} \mathbf{F}^T \otimes \mathbf{I}_K) \hat{\mathbf{g}}, \boldsymbol{\zeta} = \frac{1}{\sqrt{T}} (\mathbf{P} \mathbf{F}^T \otimes \mathbf{I}_K) \hat{\boldsymbol{\zeta}}, \hat{\mathbf{s}}_x(t) = \hat{\mathbf{x}}(t)^T \hat{\mathbf{x}}, \hat{\mathbf{s}}_z(t) = \hat{\mathbf{x}}(t)^T \hat{\boldsymbol{\zeta}}, \hat{\mathbf{s}}_h(t) = \hat{\mathbf{x}}(t)^T \hat{\mathbf{h}}, b = \hat{\mathbf{s}}_x(t) + T\mu$ 是标量。

采用式(6)对拉格朗日系数进行更新:

$$\hat{\boldsymbol{\zeta}}^{(i+1)} \leftarrow \hat{\boldsymbol{\zeta}}^i + \mu (\hat{\mathbf{g}}^{(i+1)} - \hat{\mathbf{h}}^{(i+1)}), \quad (6)$$

本文算法原理图如图1所示,主要跟踪步骤如表1所示。下面将对本文算法关键部分进行阐述。

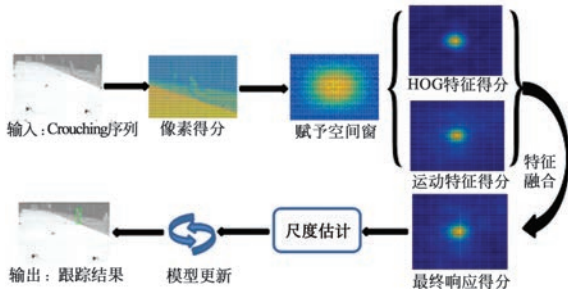


图1 本文算法原理图

Fig.1 Block diagram of the algorithm

表1 基于多特征融合相关滤波的红外目标跟踪算法

Tab.1 Infrared target tracking based on multi-feature fusion correlation filtering

输入: 图像序列, 初始化跟踪目标位置及目标尺度

输出: 预测目标位置和尺度

步骤1: 计算当前帧的目标像素值, 并赋予空间窗加权, 抑制边缘效应;

步骤2: 对目标进行 HOG 特征提取, 得到 HOG 特征响应图;

步骤3: 对目标进行运动特征提取, 得到运动特征响应图;

步骤4: 进行特征融合, 得到目标最终响应图, 响应图的最大值即为目标预测位置;

步骤5: 对目标进行尺度估计;

步骤6: 对滤波器模板使用 ADMM 优化方法进行更新。

2.2 特征融合

本文算法通过融合 HOG 特征和运动特征来增强算法在复杂背景环境中的判别力, 从而提高算法跟踪红外目标的鲁棒性。

2.2.1 HOG 特征

HOG 特征最早是由 Nvaneet 和 Bill Triggs 提出的, 之后就成为了机器视觉中广泛使用的手工特征。提取 HOG 特征首先将图像灰度化, 通过 gamma 校正法调整图像整体亮度; 其次计算梯度, 对 x, y 的梯度计算定义为:

$$G_x(x, y) = I(x+1, y) - I(x-1, y), \quad (7)$$

$$G_y(x, y) = I(x, y+1) - I(x, y-1), \quad (8)$$

其中: G_x, G_y 分别是图像 I 的水平和垂直梯度, 计算每个像素点的梯度。(x, y) 点的梯度值和方向分别为下式:

$$G_{x,y} = \sqrt{G_x(x, y)^2 + G_y(x, y)^2}, \quad (9)$$

$$\theta(x, y) = \arctan\left(\frac{G_y(x, y)}{G_x(x, y)}\right), \quad (10)$$

梯度计算完成后, 将图像划分为若干个 8×8 像素的胞元(cell), 统计各个胞元的梯度方向直方图, 将 4×4 大小的胞元串联成一个块(block), 对块内的梯度方向直方图进行归一化就得到了目标图像的 HOG 特征。

2.2.2 运动特征

利用 LDOF^[11] 计算出相邻两帧像素之间的光流, 然后对其进行运动特征提取, 最后得到运动特征响应值。根据[11], 我们可假设对应的点应该具有相同灰度值或颜色, 可用能量表示为:

$$E_{\text{color}}(l) = \int_{\Omega} \Psi(|\mathbf{I}_2(\mathbf{x} + l(x)) - \mathbf{I}_1(\mathbf{x})|^2) d\mathbf{x}, \quad (11)$$

其中: $\mathbf{I}_1, \mathbf{I}_2: (\Omega \subset \mathbb{R}^2) \rightarrow \mathbb{R}^d$ 分别为当前帧和前一帧图像, $d=1$ 是灰度图像, $d=3$ 是彩色图

像, $\mathbf{x} := (x, y)^T$ 为图像域 Ω 中的一点, $\mathbf{l} := (u, v)^T$ 为光流域。

由于光照变化影响, 匹配颜色或灰度值并不总是可靠的, 可通过对梯度和光流域总变化的约束对式(11)进行补充, 表达式分别为:

$$E_{grad}(\mathbf{l}) = \int_{\Omega} \Psi(|\nabla I_2(x + \mathbf{l}(x)) - I_1(x)|^2) dx, \quad (12)$$

$$E_{smooth}(\mathbf{l}) = \int_{\Omega} \Psi(|\nabla u(x)|^2 + |\nabla v(x)|^2) dx, \quad (13)$$

将上述 3 个约束结合在一起, 可建立模型为:

$$E(\mathbf{l}) = E_{color} + \gamma E_{gradient} + \alpha E_{smooth}, \quad (14)$$

LDOF 对上式提出了优化方案, 在此模型上进行了补充, 增加了两项约束, 分别为:

$$E_{match}(\mathbf{l}) = \int \delta(x) \rho(x) \Psi(|\mathbf{l}(x) - \mathbf{l}_1(x)|^2) dx, \quad (15)$$

$$E_{desc}(\mathbf{l}_1) = \int \delta(x) |f_2(x + \mathbf{l}_1(x)) - f_1(x)|^2 dx, \quad (16)$$

其中: f_1 和 f_2 分别表示第 1 帧和第 2 帧特征向量的稀疏区域, 代入所有约束项, 可得到优化模型:

$$E(\mathbf{l}) = E_{color}(\mathbf{l}) + \gamma E_{gradient}(\mathbf{l}) + \alpha E_{smooth}(\mathbf{l}) + \beta E_{match}(\mathbf{l}, \mathbf{l}_1) + E_{desc}(\mathbf{l}_1), \quad (17)$$

其中: α, β, γ 为 $E(\mathbf{l})$ 自动估计视频数据的可调参数。利用上述光流模型, 可得到视频帧间的光流信息。对得到的光流信息进行运动特征提取, 得到运动特征响应值。

2.3 空间窗加权

MOSSE^[12] 算法提出通过对目标原始像素值赋予余弦窗权重来突出靠近中心位置的目标, 但对目标的边缘信息缺乏鲁棒性, 导致跟踪性能下降, 因此本文算法提出使用空间加权窗^[13]来替代余弦窗, 更加能突出目标中心位置的特征, 同时也能抑制边界效应。

对于第 t 帧的特征 $\mathbf{x}'$, 经过相关滤波算法输出期望响应 $\hat{\mathbf{y}} = \mathbf{h}^t * \mathbf{x}'$ 。然而, 当目标经历局部遮挡、外观变化和快速运动时, $\hat{\mathbf{y}}$ 可能会有多个波峰或错误波峰, 这将会导致在下一帧目标位置发生漂移, 从而跟踪失败。为此, 提出窗估计法来解决上述不利影响, 通过最小化式, 可增加 $\hat{\mathbf{y}}$ 和 $\mathbf{y}$ 的相似性:

$$\varepsilon(\mathbf{w}) = \|\mathbf{h}^t * (\mathbf{w} \odot \mathbf{x}') - \mathbf{y}\|^2, \quad (18)$$

其中: $\mathbf{w}$ 为空间窗, 式(18)最小化时的 $\mathbf{w}$ 就是理想的空间窗。可利用梯度下降法求式的最小化解, $\mathbf{w}_i \leftarrow \mathbf{w}_i - \mu \partial \varepsilon(\mathbf{w}) / \partial \mathbf{w}_i$ 。为了提高计算效率, 提出一种针对空间窗 $\mathbf{w}$ 中每一个元素进行梯度计算的方法。为了证明该方法的有效性, 考虑一维信号。信号 $\mathbf{x}_n$ 和 $\mathbf{y}_n$ 的相关输出 $\mathbf{z}_n$ 定义为: $\mathbf{z}_n = \sum_i \mathbf{x}_i \mathbf{y}_{i+n}$, 式(18)的最小化可写为:

$$\varepsilon(\mathbf{w}) = \sum_n \left(\sum_i h_i \mathbf{w}_{i+n} \mathbf{x}_{i+n} - \mathbf{y}_n \right)^2, \quad (19)$$

求 $\varepsilon(\mathbf{w})$ 关于 $\mathbf{w}_m$ 的偏导数, 并让其为 0, 即:

$$\frac{\partial \varepsilon(\mathbf{w})}{\partial \mathbf{w}_m} = \sum_n \left(\sum_i h_i \mathbf{w}_{i+n} \mathbf{x}_{i+n} - \mathbf{y}_n \right) \mathbf{h}_{m-n} \mathbf{x}_m, \quad (20)$$

$$\left(\sum_n \mathbf{h}_{m-n} \mathbf{x}_m \right) \mathbf{x}_m = s_m \mathbf{x}_m, \quad (21)$$

其中: $s_m \triangleq \sum_n k_n \mathbf{h}_{m-n}$, $k_n \triangleq \sum_i h_i \mathbf{w}_{i+n} \mathbf{x}_{i+n} - \mathbf{y}_n$ 。因为中间信号 s 和 k 都是通过信号关联计算的, 所以可以利用卷积定理在频域内进行计算, 表达如下:

$$\mathbf{k} = \mathbf{T}^{-1} \{ \mathbf{T} \{ \mathbf{w} \odot \mathbf{x} \} \mathbf{H}^* - \mathbf{G} \}, \quad (22)$$

$$\mathbf{s} = \mathbf{F}^{-1} \{ \mathbf{K} \odot \mathbf{H} \}, \quad (23)$$

利用 $\mathbf{k}_m$ 和 s_m 可以高效的计算偏导数:

$$\frac{\partial \varepsilon(\mathbf{w})}{\partial \mathbf{w}_m} = s_m \mathbf{x}_m. \quad (24)$$

2.4 计算最大响应位置

分别求出 HOG 特征响应值 $f_{HOG}(\mathbf{I})$ 和运动特征的响应值 $f_{MO}(\mathbf{I})$ 后, 利用线性求和的方法, 对其进行融合, 可得到最终响应值为:

$$f(\mathbf{I}) = \gamma f_{HOG}(\mathbf{I}) + (\mathbf{I} - \gamma) f_{MO}(\mathbf{I}), \quad (25)$$

得到最终响应图后, 通过寻找响应图的最大值来确定目标中心位置, 目标预测位置为:

$$\mathbf{p}_t = \operatorname{argmax} f(\mathbf{I}). \quad (26)$$

2.5 模型更新

本文算法采用传统 CF 的线性插值模型更新策略, 可对形变、尺度和光照变化鲁棒。第 t 帧的模型更新可用下式表示:

$$\hat{\mathbf{x}}_{\text{model}}^{(t)} = (1 - \eta) \hat{\mathbf{x}}_{\text{model}}^{(t-1)} + \eta \hat{\mathbf{x}}^{(t)}, \quad (27)$$

其中 η 为更新率, 并利用 $\hat{\mathbf{x}}_{\text{model}}^{(t)}$ 计算 $\hat{\mathbf{h}}(t)^*$ 、 $\hat{s}_x(t)$ 、 $\hat{s}_y(t)$ 和 $\hat{s}_h$ 。

2.6 尺度估计

本文算法通过利用第 $t-1$ 帧更新的滤波器 $\hat{\mathbf{h}}^{t-1}$ 来检测第 t 帧的目标位置。根据文献[2], 可

在多个分辨率下将滤波器应用于搜索区域来估计尺度变化。搜索区域设置为和滤波器 $\hat{h}$ 大小一致。以先前目标位置为中心并且对当前目标尺度 a^r 提取样本 $\{z_r\}$, 其中 $r \in \{\lfloor \frac{1-S}{2} \rfloor, \dots, \lfloor \frac{S-1}{2} \rfloor\}$, a 是尺度因子, S 是尺度数, 然后返回 S 个相关输出。本文算法采用文献[2]中的插值策略, 对每个相关输出的检测分数进行最大化, 然后用最大相关输出的尺度来更新目标位置和尺度。

3 实验结果与分析

3.1 数据集与评价标准

本文算法是在 MATLAB 2014b, Intel Core i3, 主频为 2.5 GHz, 4G 内存, 64 位 WIN10 操作系统上进行所有实验。为了保证实验结果的客观性, 使用 VOT-TIR2016 红外目标跟踪数据集^[14]对算法进行测试。该数据集共包括 25 个视频序列和跟踪目标所对应的 11 种属性, 平均序列长度为 740 帧, 分辨率从 305×225 到 1920×480 像素不等。VOT-TIR 2016 数据集与 VOT-TIR 2015 数据集相比, 删除了 Crossing, Horse 和 Rhino_behind_tree 序列, 引入了更困难的 Bird, Boat1, Boat2, Car2, Dog, Excavator, Ragged, 和 Trees2 序列。为了证明本文算法的有效性, 提供本文算法和其他 15 种算法进行比较, 分别是 SRDCF^[2], Staple^[15], Staple_CA^[3], TGPR^[16], SAMF^[17], DSST^[18], ACT^[19], KCF^[1], DST^[20], KCF_DP^[21], RPT^[22], HRP^[23], KCF_MTSA^[24], ASLA^[25], PDMFD^[26]。

本文使用精确度 (precision) 和成功率 (success rate) 两种评价标准对本文算法的性能进行评价。精确度指的是小于中心位置误差阈值的视频帧数占总帧数的百分比, 而中心位置误差定义为算法得到的跟踪目标框中心位置和标定的真实目标框的中心位置之间的欧式距离。成功率指的是跟踪目标的重叠率大于重叠率阈值的视频帧数占总帧数的百分比, 重叠率定义为跟踪目标框和标定的真实目标框之间交集与并集比值^[14]。

3.2 参数设置

本文通过大量的试验来确定该算法最优的参数设置, 实验参数设置如下: 选择 31 维 HOG 特

征, cell 大小设置为 4×4 并乘空间窗, 正则化系数设置为 $\lambda = 0.001$, 尺度池设置为 5, 步长为 1.01。ADMM 优化算法中, 设置迭代次数和惩罚因子分别为 2 和 1, 第 $i+1$ 次的惩罚因子用 $\mu^{i+1} = \min(\mu_{\max}, \beta\mu^{(i)})$ 更新, 其中 $\beta = 10$, $\mu = 10^{-3}$, 特征融合系数设置为 $\gamma_{\text{HOG}} = 0.7$ 和 $\gamma_{\text{MO}} = 0.3$, 更新率设置为 $\eta = 0.013$ 。

3.3 定量分析

本文使用 OPE(one-pass evaluation)方法^[27]在所有视频序列上进行本文算法的测试。设置中心位置误差阈值为 20 来评估算法的精确度, 设置重叠率为 0.5 来评估算法的成功率。为显示清晰, 图 2 和图 3 分别给出结果排名前 10 算法的精确度图和成功率图。由图可知, 本文算法与排名第二的算法相比, 精确度提高了 8.8%, 成功率提高了 15.4%, 两项指标均有显著的提升, 证明了本文算法的有效性和鲁棒性。

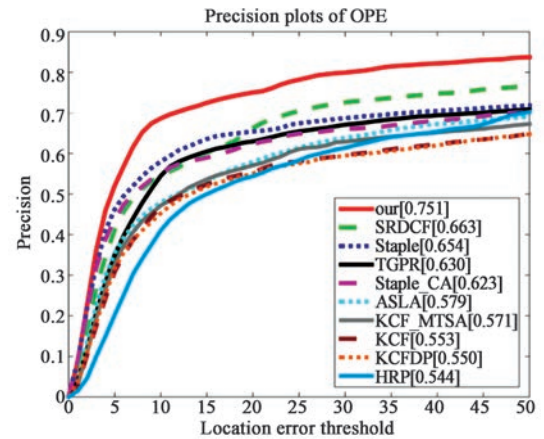


图2 25个视频序列上得到的不同算法的精确度
Fig.2 Accuracy of different algorithms tested on 25 sequences

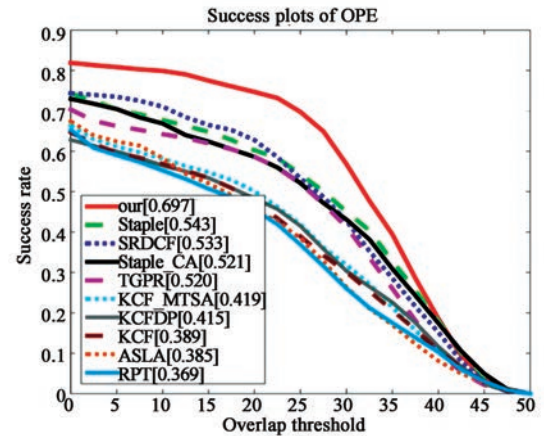


图3 25个视频序列上得到的不同算法的成功率
Fig.3 Success rate of different algorithms tested on 25 sequences

图 4 和图 5 为不同算法在 11 种属性上的精确度和成功率图,为显示清晰,给出排名前十算

法的结果,11 种属性分别为变形(deformation)、光照变化(illumination variation)、平面内旋转(in-

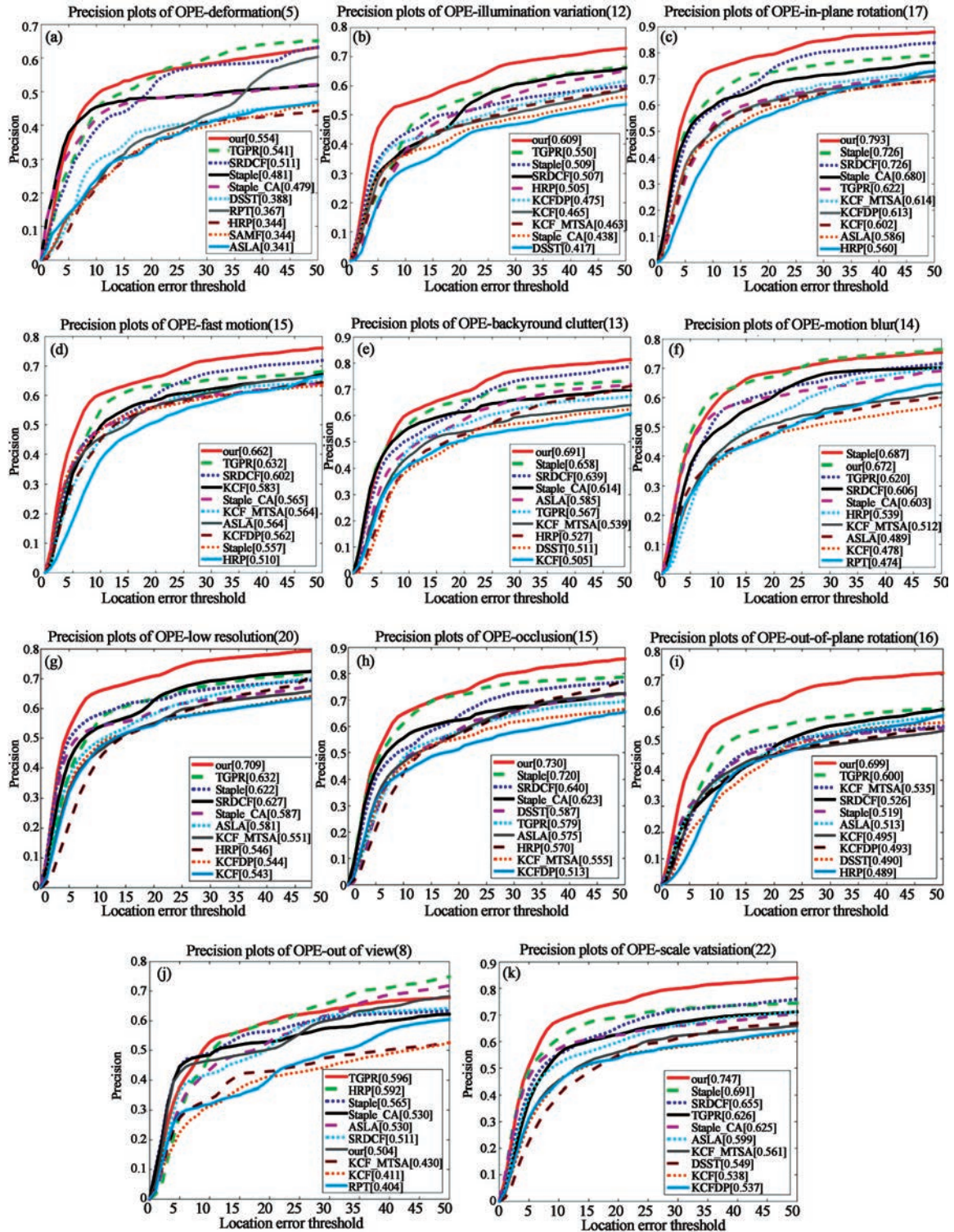


图 4 在 11 种属性上不同算法的精确度。(a)形变;(b)光照变化;(c)平面内旋转;(d)快速运动;(e)背景杂乱;(f)运动模糊;(g)低分辨率;(h)遮挡;(i)平面外旋转;(j)出视野;(k)尺度变化。

Fig.4 Precision plots of different algorithms on 11 attributes. (a) Deformation; (b) Illumination variation; (c) In-plane rotation; (d) Fast motion; (e) Background clutter; (f) Motion blur; (g) Low resolution; (h) Occlusion; (i) Out-of-plane rotation; (j) Out of view; (k) Scale variation.

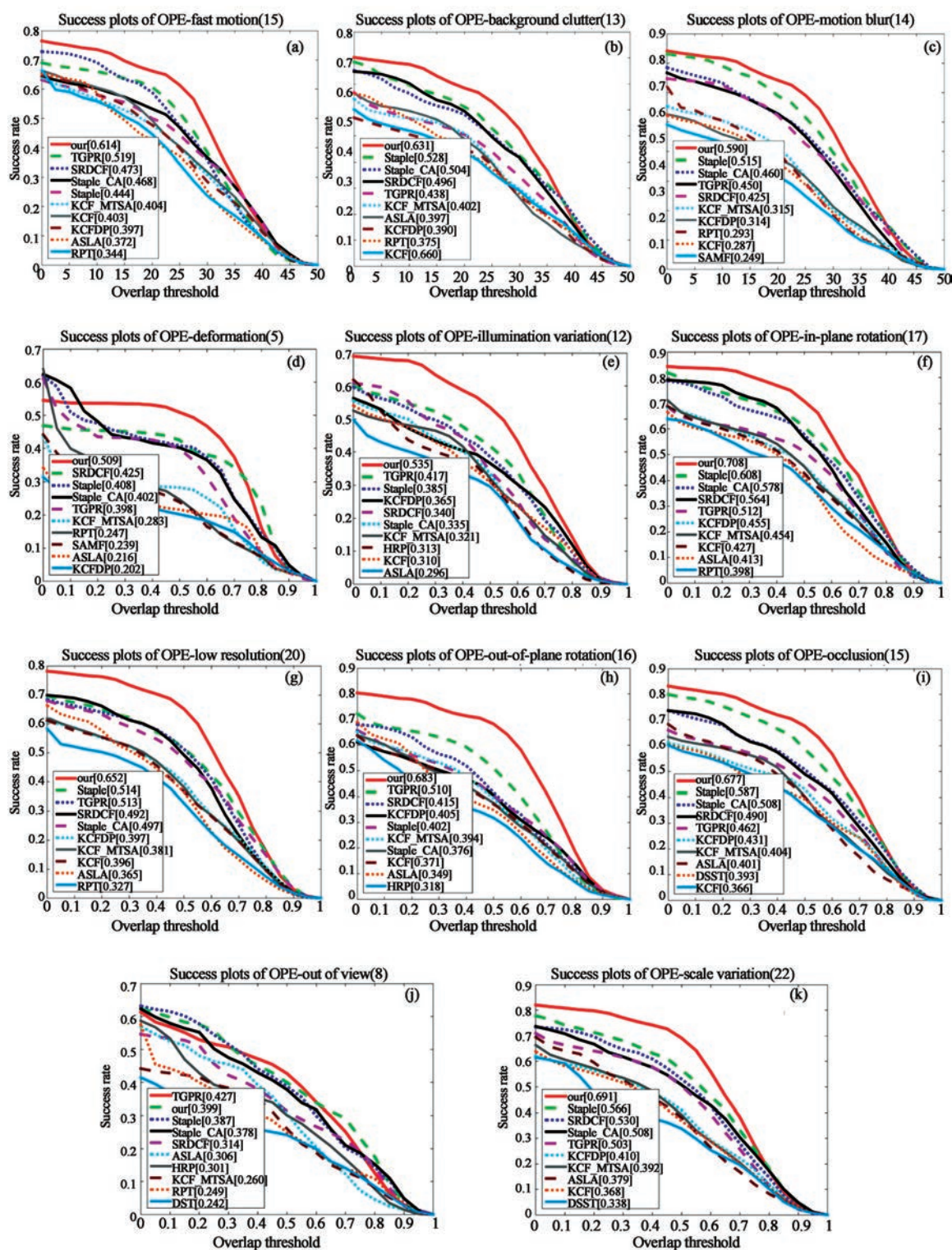


图5 在11种属性上不同算法的成功率。(a)形变;(b)光照变化;(c)平面内旋转;(d)快速运动;(e)背景杂乱;(f)运动模糊;(g)低分辨率;(h)遮挡;(i)平面外旋转;(j)出视野;(k)尺度变化。

Fig.5 Success rate plots of different algorithms on 11 attributes. (a) Deformation; (b) Illumination variation; (c) In-plane rotation; (d) Fast motion; (e) Background clutter; (f) Motion blur; (g) Low resolution; (h) Occlusion; (i) Out-of-plane rotation; (j) Out of view; (k) Scale variation.

plane rotation)、快速运动(fast motion)、背景杂乱(background clutter)、运动模糊(motion blur)、低分辨率(low resolution)、遮挡(occlusion)、平面外旋转(out-plane rotation)、出视野(out of view)、尺度变化(scale variation)^[28]。从图 4 中可看出,本文算法在运动模糊和出视野两种属性上精度排名第二,在其余 9 种属性上均表现出良好效果。可证明本文算法可以很好的处理红外目标尺度变化、变形、快速运动、平面内外旋转等问题。从图 5 中可看出,对于在出视野之外的属性,本文算法比排名第二的算法均有显著提高,说明融合特征以及空间窗的引入可以极大的提高算法的精确度和成功率,证明了本文算法的有效性。

3.4 定性分析

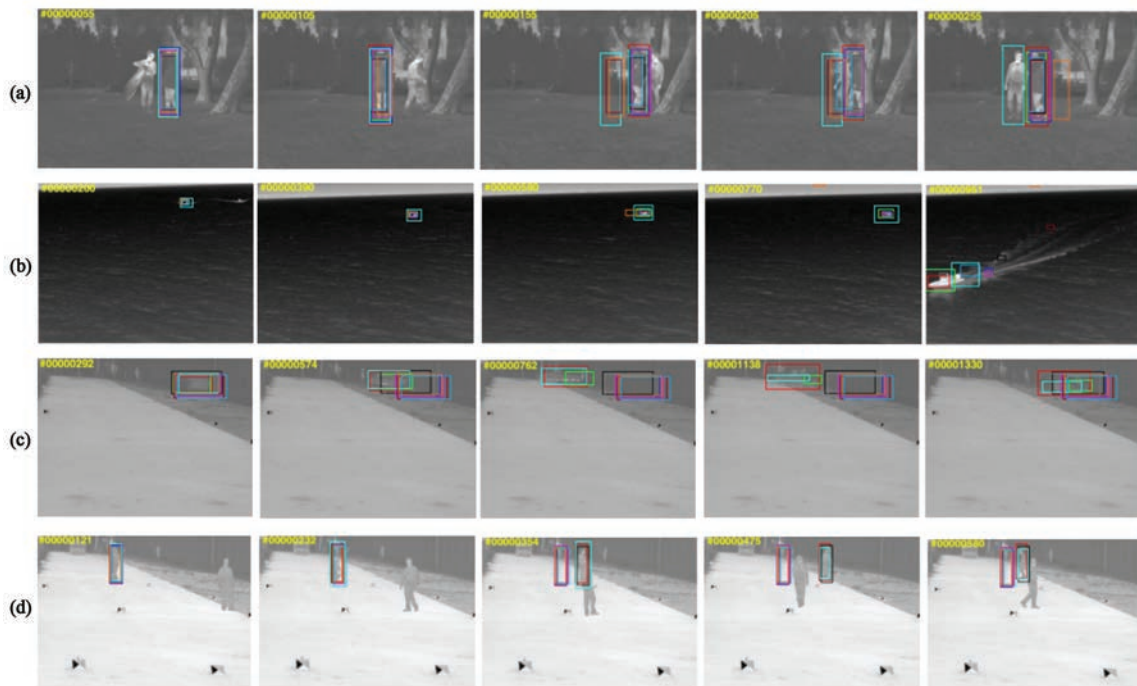
为了更好的分析本文算法的性能,给出本文算法和其他 9 种算法在 25 个视频序列上的目标跟踪结果,如图 6 所示。当跟踪目标经历不同变化时,针对相应的属性进行分析和比较。

(1)形变。在 birds 和 crouching 视频序列中,当目标发生形变时,只有少数算法能够利用有效的红外信息提取关键特征,实现精准目标跟踪,本文算法为其中之一。在 birds 视频序列中,当目标在第 105 帧和第 225 帧发生形变时,只有本文算法和 SRDCF、Staple、Staple_CA 和 TGPR 算法能正确跟踪。在 crouching 视频序列中,当

目标在第 232 帧下蹲之后又在第 354 帧站起来行走时,发生了较大的形变,只有本文算法和 TGPR、ASLA 算法没有发生漂移,能正确跟踪,说明了本文算法较好的鲁棒性。

(2)遮挡。遮挡视频序列有 hiding 和 trees1。其中,hiding 视频序列属于完全遮挡,而 trees2 视频序列属于部分遮挡。在 hiding 视频序列中,当第 173 帧目标完全被遮挡,第 248 帧目标再出现的时候,只有本文算法和 HRP、ASLA 算法能准确跟踪,并且在后续跟踪中,本文算法没有发生漂移。在 trees1 视频序列中,当目标经历遮挡后可从第 272 帧和第 403 帧看出,有 4 种算法跟踪失败,而本文算法能够正确跟踪,证明本文算法有很好的处理遮挡能力。

(3)平面外旋转。25 个视频序列中有很多视频都具有平面外旋转属性,其中包括 birds、car1、crouching、depthwise_crossing 和 jacket 等。下面将针对 car1 和 jacket 视频序列进行该属性分析。在 car1 视频序列中,当目标在第 574 帧、第 762 帧、第 1 138 帧发生平面外旋转时,只有本文算法未发生漂移,能够准确跟踪目标,其他算法都跟踪失败。在 jacket 视频序列中,当目标在第 586 帧发生平面外旋转(转身)时,只有本文算法和 SRDCF、Staple 算法没有跟踪失败,得益于本文算法的特征融合,能够对红外目标正确跟踪。



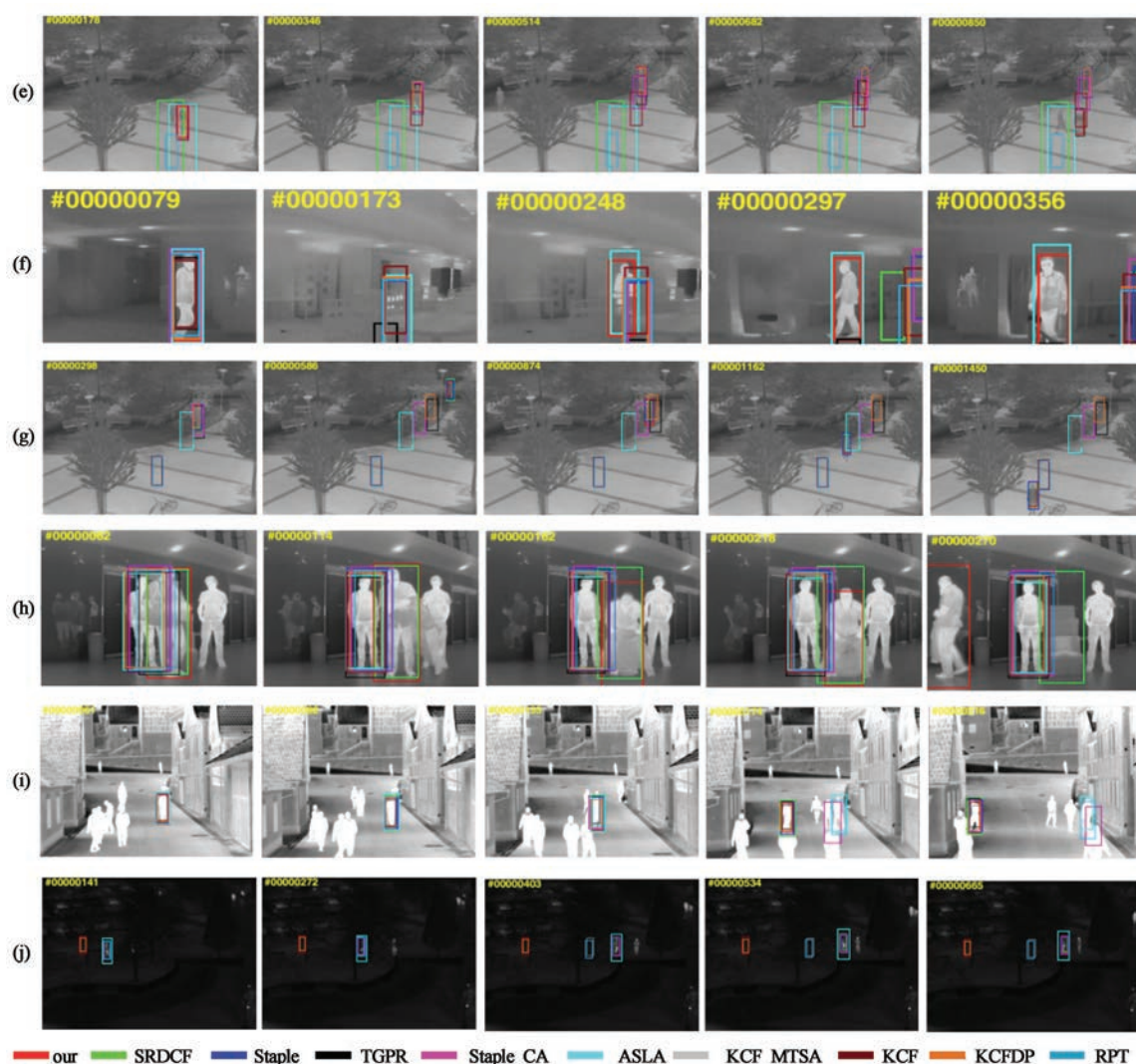


图6 本文算法与其他9种算法在10个视频序列上的跟踪结果。(a) Birds; (b) Boat2; (c) Car1; (d) Crouching; (e) Depthwise_crossing; (f) Hiding; (g) Jacket; (h) Mixed_distractors; (i) Saturated; (j) Trees1。
Fig.6 Tracking results of the algorithm and other 9 algorithms on 10 video sequences. (a) Birds; (b) Boat2; (c) Car1; (d) Crouching; (e) Depthwise_crossing; (f) Hiding; (g) Jacket; (h) Mixed_distractors; (i) Saturated; (j) Trees1.

(4)背景杂乱。由于红外目标缺少颜色、纹理等信息,当目标处在背景杂乱的情况下,目标跟踪十分困难。而在 mixed_distractors 和 saturated 视频序列中都存在背景杂乱的情况。下面将逐一进行分析。在 mixed_distractors 视频序列中,由于背景有很多相似物体,使得目标跟踪算法很容易发生漂移。在视频的第114帧,当目标经过第一个相似物体时,只有本文算法会和 SRDCF 算法未发生漂移,其他算法都跟踪失败,但是在后续跟踪中,SRDCF 算法发生了漂移,导致跟踪失败,而本文算法可以一直很好的跟踪。在 saturated 视频序列中,我们可看出背景中有很多和目标相

似的物体,具有很大的干扰性,而本文算法可以很好的跟踪目标。证明了本文算法适用于背景杂乱的场景,有很好的鲁棒性。

(5)尺度变化。尺度变化是目标跟踪中比较常见和基础的问题,当目标缩小时,滤波器就会学习到大量背景信息,当目标扩大时,滤波器就会跟着目标局部纹理走丢。视频序列 boat2 和 mixed_distractors 都存在尺度变化情况。在 boat2 视频序列中,从第951帧可看出,当目标越来越远时,目标尺度也是越来越大,只有本文算法能有效的对变化的目标进行尺度估计,其余算法都跟踪失败。在 mixed_distractors 视频序列中,目标在第

218 帧坐下时,本文算法可对目标尺度变化进行自适应估计,调整为相应大小的目标框,减少了背景信息的引入,提高了跟踪算法的精度和成功率。由此可说明本文算法尺度估计的有效性。

4 结 论

本文提出基于多特征融合的红外目标跟踪算法,在背景感知相关滤波器的框架下,对红外目标进行 HOG 特征和运动特征提取,通过线性求和方式对特征进行融合,使得跟踪算法对目标的形

变、快速运动、遮挡等属性能有较好的鲁棒效果。同时采用空间加权窗代替传统相关滤波器中的余弦窗,更加突出目标特征,也能抑制边界效应。本文算法采用 VOT TIR2016 数据集进行性能测试,并和 15 种流行算法进行了分析和比较,结果表明,本文算法在精确度和成功率上的得分分别为 0.751 和 0.697,在精确度和成功率方面分别提高了 8.8% 和 15.4%,证明了本文算法的有效性和鲁棒性,具有一定的研究价值。目前本文算法不适用于红外低分辨率目标跟踪任务,对光照变化情况处理不是很好,这将是未来改进的主要方向。

参 考 文 献:

- [1] HENRIQUES J F, CASEIRO R, MARTINS P, *et al.* High-speed tracking with kernelized correlation filters [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(3): 583-596.
- [2] DANELLJAN M, HÄGER G, SHAHBAZ KHAN F, *et al.* Learning spatially regularized correlation filters for visual tracking [C]// *IEEE International Conference on Computer Vision*. Santiago, Chile: IEEE, 2015: 4310-4318.
- [3] MUELLER M, SMITH N, GHANEM B. Context-aware correlation filter tracking [C]// *IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, HI, USA: IEEE, 2017: 1387-1395.
- [4] 赵东,周慧鑫,秦翰林,等.基于引导滤波和核相关滤波的红外弱小目标跟踪[J].*光学学报*,2018,38(2):204004.
ZHAO D, ZHOU H X, QIN H L, *et al.* Infrared dim-small target tracking based on guided image filtering and kernelized correlation filtering [J]. *Acta Optica Sinica*, 2018, 38(2): 204004. (in Chinese)
- [5] ASHA C S, NARASIMHADHAN A V. Robust infrared target tracking using discriminative and generative approaches [J]. *Infrared Physics & Technology*, 2017, 85: 114-127.
- [6] 刘先红,陈志斌,秦梦泽.结合引导滤波和卷积稀疏表示的红外与可见光图像融合[J].*光学精密工程*,2018,26(5): 1242-1253.
LIU X H, CHEN Z B, QIN M Z. Infrared and visible image fusion using guided filter and convolutional sparse representation [J]. *Optical Precision Engineering*, 2018, 26(5): 1242-1253. (in Chinese)
- [7] 杨福才,杨德东,毛宁,等.基于稀疏编码直方图的稳健红外目标跟踪[J].*光学学报*,2017,37(11):1115002.
YANG F C, YANG D D, MAO N, *et al.* Robust infrared target tracking based on histograms of sparse coding [J]. *Acta Optica Sinica*, 2017, 37(11): 1115002. (in Chinese)
- [8] GAO J L, WEN C L, LIU M Q. Robust small target co-detection from airborne infrared image sequences [J]. *Sensors*, 2017, 17(10): 2242.
- [9] 张春宇.基于运动特征的视频目标跟踪算法研究[D].深圳:深圳大学,2016:1-73.
ZHANG C Y. Research on video target tracking algorithm based on motion feature [D]. Shenzhen: Shenzhen University, 2016: 1-73. (in Chinese)
- [10] GALOOGAHI H K, FAGG A, LUCEY S. Learning background-aware correlation filters for visual tracking [C]// *IEEE Conference on Computer Vision and Pattern Recognition*. Venice, Italy: IEEE, 2017: 21-26.
- [11] BROX T, MALIK J. Large displacement optical flow: descriptor matching invarianational motion estimation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(3): 500-513.
- [12] BOLME D S, BEVERIDGE J R, DRAPER B A, *et al.* Visual object tracking using adaptive correlation filters [C]// *IEEE Conference on Computer Vision and Pattern Recognition*. San Francisco, CA, USA: IEEE, 2010: 2544-2550.
- [13] GUNDOGDU E, ALATAN AA. Spatial windowing for correlation filter based visual tracking [C]// *IEEE International Conference on Image Processing*. Phoenix, AZ, USA: IEEE, 2016: 1684-1688.
- [14] FELSBURG M, KRISTAN M, MATAS J, *et al.* The thermal infrared visual object tracking VOT-TIR2016 challenge results [C]// *Computer Vision-ECCV 2016 Workshops*. Cham: Springer, 2016: 639-651.

- [15] BERTINETTO L, VALMADRE J, GOLODETZ S, *et al.* Staple: complementary learners for real-time tracking [C]//*IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV, USA: IEEE, 2016: 1401-1409.
- [16] GAO J, LING H B, HU W M, *et al.* Transfer learning based visual tracking with Gaussian processes regression [C]//*Proceedings of the 13th European Conference on Computer Vision*. Zurich, Switzerland: Springer, 2014: 188-203.
- [17] LI Y, ZHU J K. A scale adaptive kernel correlation filter tracker with feature integration [C]//*European Conference on Computer Vision*. Zurich, Switzerland: Springer, 2014: 254-265.
- [18] DANELLJAN M, HÄGER G, KHAN F S, *et al.* Accurate scale estimation for robust visual tracking [C]//*British Machine Vision Conference*. Nottingham, UK: BMVA Press, 2014: 1-11.
- [19] LIU Q S, YANG J, ZHANG K H, *et al.* Adaptive compressive tracking via online vector boosting feature selection [J]. *IEEE Transactions on Cybernetics*, 2017, 47(12): 4289-4301.
- [20] XIAO J J, QIAO L B, STOLKIN R, *et al.* Distractor-supported single target tracking in extremely cluttered scenes [C]//*Proceedings of the 14th European Conference on Computer Vision*. Amsterdam, The Netherlands: Springer, 2016: 121-136.
- [21] HUANG D F, LUO L, WEN M, *et al.* Enable scale and aspect ratio adaptability in visual tracking with detection proposals [C]//*British Machine Vision Conference*. Swansea, UK: BMVA Press, 2015: 1.
- [22] LI Y, ZHU J K, HOI S C H. Reliable patch trackers: robust visual tracking by exploiting reliable patches [C]//*IEEE Conference on Computer Vision and Pattern Recognition*. Boston, MA, USA: IEEE, 2015: 353-361.
- [23] WANG N Y, SHI J P, YEUNG D Y, *et al.* Understanding and diagnosing visual tracking systems [C]//*IEEE International Conference on Computer Vision*. Santiago, Chile: IEEE, 2015: 3101-3109.
- [24] BIBI A, GHANEM B. Multi-template scale-adaptive kernelized correlation filters [C]//*IEEE International Conference on Computer Vision Workshops*. Santiago, Chile: IEEE, 2015: 50-57.
- [25] JIA X, LU H C, YANG M H. Visual tracking via adaptive structural local sparse appearance model [C]//*IEEE Conference on Computer Vision and Pattern Recognition*. Providence, RI, USA: IEEE, 2012: 1822-1829.
- [26] KARANAM S, LI Y, RADKE R J. Particle dynamics and multi-channel feature dictionaries for robust visual tracking [C]//*British Machine Vision Conference*. Swansea, UK: BMVA Press, 2015: 1-12.
- [27] WU Y, LIM J, YANG M H. Online object tracking: A benchmark [C]//*IEEE Conference on Computer Vision and Pattern Recognition*. Portland, OR, USA: IEEE, 2013: 2411-2418.
- [28] DANELLJAN M, BHAT G, KHAN F S, *et al.* ECO: Efficient convolution operators for tracking [C]//*IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, HI, USA: IEEE, 2017: 21-26.

作者简介:



韩亚君(1993—),女,山西阳泉人,硕士研究生,2017年于西安工程大学获得学士学位,主要从事机器视觉学习、目标跟踪方面的研究。E-mail: 1109272402@qq.com



杨德东(1978—),男,辽宁阜新人,博士后,副教授,2007年于东北大学获得博士学位,2010年于北京航空航天大学博士后出站,主要从事智能感知与控制方面的研究。E-mail: ydd12677@163.com