

基于响应图置信度的并行相关滤波跟踪算法

宋奇奇, 李晓丽*, 左伟, 顾立鹏

东华大学信息科学与技术学院, 上海 201620

摘要 相比传统目标跟踪算法, 基于相关滤波的目标跟踪算法跟踪精度更高、实时性更好, 但当目标被遮挡或超出视野时, 难以提取目标特征、正确定位检测窗口, 容易导致目标跟踪失败或漂移。因此, 提出了一种基于置信度的并行相关滤波跟踪算法。首先, 使用一种新的置信度评价方法判断目标是否被遮挡或存在异常情况。其次, 在置信度的基础上, 用组合权重融合两种不同的跟踪器, 构建并行相关滤波跟踪算法, 以提高算法的跟踪精度和鲁棒性。最后, 为了防止模型污染, 对两种滤波器模型采取自适应权重更新策略。在 OTB-2013 和 OTB-2015 数据集上进行实验, 结果表明, 相比传统算法, 本算法在跟踪精度和成功率上均有显著提升。

关键词 图像处理; 相关滤波; 目标跟踪; 置信度; 自适应学习率; 并行跟踪

中图分类号 TP391

文献标志码 A

doi: 10.3788/LOP57.201017

Parallel Correlation Filter Tracking Algorithm Based on Response Map Confidence

Song Qiqi, Li Xiaoli*, Zuo Wei, Gu Lipeng

School of Information Science and Technology, Donghua University, Shanghai 201620, China

Abstract Compared with traditional target tracking algorithms, the target tracking algorithm based on correlation filter has great advantages in tracking accuracy and real-time performance. However, when the target is blocked or out of view, it is difficult to extract target features and correctly locate the detection window, which may easily lead to target tracking failure or drift. Therefore, in this paper, a parallel correlation filter tracking algorithm based on confidence is proposed. First, the paper proposes a new confidence evaluation method to determine whether the target is blocked or out of view. Second, on the basis of confidence, two different trackers are fused with the combination weight of confidence to construct a parallel correlation filter tracking algorithm in order to improve the tracking accuracy and robustness. Finally, in order to prevent model pollution, an adaptive weight update strategy is adopted for the two filter models. Experiments on OTB-2013 and OTB-2015 datasets show that compared with traditional algorithms, this algorithm has significantly improved tracking accuracy and success rate.

Key words image processing; correlation filter; target tracking; confidence; adaptive learning rate; parallel tracking

OCIS codes 100.2000; 150.1135

1 引言

在计算机视觉领域, 目标跟踪是重要的研究分支, 广泛应用于视频监控、运动特征分析、智能交通、无人驾驶等领域^[1]。单目标跟踪的本质是在给定视频序列初始帧目标位置和尺寸的情况下, 搜索目标在后续视频中的位置和尺寸。随着目标跟踪

算法的不断发展和计算机硬件性能的提升, 目标跟踪算法取得了明显的突破。但在真实环境中, 目标往往存在遮挡、尺寸变化、光照变化、背景变化、外观变化、快速移动、短暂消失等问题, 这对高精度、高成功率且鲁棒性较好的跟踪算法提出了更大的挑战^[2]。

相关滤波跟踪算法利用最小化误差的思想训

收稿日期: 2020-01-15; 修回日期: 2020-02-11; 录用日期: 2020-03-19

基金项目: 上海市自然科学基金(16ZR1446700, 17ZR1400100)、上海市浦江人才计划(18PJ1400100)

* E-mail: xlli@dhu.edu.cn;

练分类器,在线学习相关滤波器模型,将提取的目标信息与相关滤波器进行相关运算,得到一组目标可能所在位置的响应值,并将响应值最大的位置作为目标的中心位置。其中,输入信息、滤波器模型、输出信息均为尺寸相同的矩阵。为了提升矩阵的运算速度,用快速傅里叶变换将矩阵变换至频域计算,以保证算法的实时性,滤波器模型的作用是衡量输入信息与输出信息的相似度。

Bolme 等^[3]提出最小输出误差平方和(MOSSE)滤波算法,将相关滤波算法应用到目标跟踪领域,对初始帧目标进行随机仿射变换并提取灰度特征训练分类器。在此基础上,Henriques 等^[4]提出了一种基于循环结构核(CKS)的检测跟踪算法,该算法将目标区域作为基样本进行循环移位,构建了大量的训练样本训练分类器,使分类器的鲁棒性更强;并用同样的方式对待检测区域进行密集采样,得到较大的候选区域,但该算法的实时性较差。Henriques 等^[5]在 CSK 算法的基础上,利用循环矩阵在傅里叶空间可对角化的性质将矩阵运算转化为向量的 Hadamard 积,提出了核相关滤波器(KCF)算法,大大提升了算法的运算速度。Danelljan 等^[6]在 CSK 算法的基础上引入颜色属性,提出了 Color Name(CN)算法,将 3 维 RGB(Red, Green, Blue)特征映射为 11 维 CN 特征,提升了算法在彩色图像序列上的鲁棒性。

针对目标在复杂场景下容易受到干扰的问题, Bertinetto 等^[7]提出了 Staple 算法,将方向梯度直方图(HOG)特征和颜色直方图特征融合,增加了算法的适用范围。李聪等^[8]将 RGB 颜色空间映射到 Lab(Luminosity, a, b)颜色空间,然后结合 HOG 特征作为目标的特征输入。沈秋等^[9-10]提出用卷积神经网络(CNN)提取的目标特征作为相关滤波器的输入,但算法的跟踪速度不能满足在线跟踪的要求。虽然可以将更复杂的目标特征作为输入以提升算法的跟踪效果,但目标被遮挡时,滤波器模型会引入遮挡物信息或背景信息,导致目标跟踪漂移或失败。

为了提升复杂场景下跟踪算法的成功率与精确度,并改善目标短暂遮挡和超出视野的问题,本文提出了一种基于置信度的并行相关滤波跟踪算法,可解决跟踪过程中目标被短暂遮挡或消失的问题。通过跟踪过程中的响应图计算次峰均值与主峰的比值(SMR)和峰值旁瓣比(PSR)^[3],融合 SMR 和 PSR 作为新的置信度评价机制,提出了融合响应

图的并行相关滤波跟踪算法,并对滤波器模型采取自适应权重更新策略。

2 相关滤波跟踪算法

2.1 核相关滤波跟踪算法

相关滤波跟踪算法的实质是通过函数 $f(\mathbf{z}) = \mathbf{w}^T \mathbf{z}$ 求得目标特征的实际输出,使误差函数最小,其中, \mathbf{z} 为输入特征, $f(\mathbf{z})$ 为在权重 \mathbf{w}^T 作用下的输出,可表示为

$$\min_{\mathbf{w}} \sum_i [f(\mathbf{x}_i) - \mathbf{y}_i]^2 + \lambda \|\mathbf{w}\|^2, \quad (1)$$

式中, \mathbf{w} 为分类器的待求参数,即相关滤波器模型, i 为向量的行数, \mathbf{x}_i 为目标的输入特征, \mathbf{y}_i 为理想输出结果, $f(\mathbf{x}_i)$ 为实际输出, $\lambda \|\mathbf{w}\|^2$ 为防止分类器过拟合的正则化项, λ 为正则化系数, $\|\mathbf{w}\|^2$ 为 \mathbf{w} 的二范数。将误差平方作为损失函数,得到

$$\mathbf{w} = (\mathbf{X}^T \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^T \mathbf{y}, \quad (2)$$

式中, \mathbf{X} 为样本矩阵,每行 \mathbf{x}_i 表示一个样本, \mathbf{I} 为与 \mathbf{X} 同维数的单位矩阵, \mathbf{y} 中的每个 \mathbf{y}_i 对应输入 \mathbf{x}_i 在滤波器作用下的响应值。对于 KCF 算法,其输入由初始目标样本循环移位得到,其训练样本集 \mathbf{X} 具有循环结构。利用循环矩阵的离散傅里叶变换(DFT)性质,(2)式在频域中可表示为

$$\hat{\mathbf{w}} = \frac{\hat{\mathbf{x}}^* \odot \hat{\mathbf{y}}}{\hat{\mathbf{x}}^* \odot \hat{\mathbf{x}} + \lambda'}, \quad (3)$$

式中, \odot 为点乘, \mathbf{x} 、 \mathbf{y} 为矩阵 \mathbf{X} 及向量 \mathbf{y} 中元素的统一表示, $\hat{\mathbf{x}}$ 、 $\hat{\mathbf{y}}$ 和 $\hat{\mathbf{w}}$ 分别为 \mathbf{x} 、 \mathbf{y} 和 \mathbf{w} 的离散傅里叶变换, $\hat{\mathbf{x}}^*$ 为 $\hat{\mathbf{x}}$ 的复共轭, λ' 为频域中的正则化系数。通过 KCF 算法将 \mathbf{x} 映射到高维特征空间中, \mathbf{w} 在其对偶空间中可表示为 $\mathbf{w} = \sum_i \alpha \varphi(\mathbf{x}_i)$, 其中, $\varphi(\mathbf{x}_i)$ 为 \mathbf{x}_i 的非线性映射函数。此时将 \mathbf{w} 的求解问题转换为求 α , α 在频域中可表示为

$$\hat{\alpha} = \frac{\hat{\mathbf{y}}}{\hat{\kappa}^{xx} + \lambda'}, \quad (4)$$

式中, $\hat{\kappa}^{xx} = \langle \varphi(\mathbf{x}), \varphi(\mathbf{x}) \rangle$, $\langle \rangle$ 为元素的线性核函数, $\varphi(\mathbf{x})$ 为 \mathbf{x} 的非线性函数映射,输入特征 \mathbf{z} 的输出响应可表示为

$$f(\mathbf{z}) = \mathcal{F}^{-1}[\hat{f}(\mathbf{z})] = \mathcal{F}^{-1}(\hat{\kappa}^{xz} \odot \hat{\alpha}), \quad (5)$$

式中, $\hat{\kappa}^{xz} = \langle \varphi(\mathbf{x}), \varphi(\mathbf{z}) \rangle$, \mathcal{F}^{-1} 为快速傅里叶逆变换(IFFT), $\hat{f}(\mathbf{z})$ 为频域中的响应值。经 IFFT 后, $f(\mathbf{z})$ 最大值所在位置就是目标的中心位置。

2.2 快速判别尺寸空间跟踪算法

与 KCF 算法的原理类似,快速判别尺寸空间跟

踪(FDSST)算法^[11]将输入信号 f 设置为 d 维特征向量,通过建立最小化代价函数构造最优相关滤波器 h ,可表示为

$$\epsilon = \left\| \sum_{l=1}^d h^l * f^l - g \right\|^2 + \lambda \sum_{l=1}^d \|h^l\|^2, \quad (6)$$

式中, ϵ 为代价函数, l 为特征的某一维度,正则化项系数 λ 可消除 f 频谱中的零频分量,避免解的分子为零, g 为理想的输出响应。(6)式为线性最小二乘问题,可通过快速傅里叶变换(FFT)在频域中求解。令(6)式的导数为零,得到

$$\mathbf{H}^l = \frac{\mathbf{G} * \mathbf{F}^l}{\sum_{k=1}^d \mathbf{F}^k * \mathbf{F}^k + \lambda} = \frac{\mathbf{A}_t^l}{B_t}, \quad (7)$$

式中, \mathbf{G} 、 \mathbf{F} 分别为 g 、 f 的 FFT, \mathbf{H} 为频域中的相关滤波器, \mathbf{A}_t^l 为滤波器的分子, B_t 为滤波器的分母, t 为当前帧的时间, k 为输入的第 k 个样本。

为了提高计算效率,相关滤波器的尺寸大小与搜索区域尺寸相同,利用从当前帧提取的新样本特征更新 \mathbf{A}_t^l 和 B_t ,可表示为

$$\mathbf{A}_t^l = (1 - \eta)\mathbf{A}_{t-1}^l + \eta\mathbf{G}_t * \mathbf{F}_t^l, \quad (8)$$

$$B_t = (1 - \eta)B_{t-1} + \eta \sum_{k=1}^d \mathbf{F}_t^k * \mathbf{F}_t^k, \quad (9)$$

式中, η 为学习率, \mathbf{A}_{t-1}^l 、 B_{t-1} 分别为前一帧滤波器的模型参数, \mathbf{F}_t^l 、 \mathbf{G}_t 分别为当前帧输入与输出的傅里叶变换。

为了检测目标位置和尺寸的变化,FDSST 算法首先学习用于位置估计的 2 维判别相关滤波器(DCF),然后学习用于尺寸估计的 1 维 DCF,当前帧的响应结果可表示为

$$\mathbf{y} = \mathcal{F}^{-1} \left(\frac{\sum_{l=1}^d \overline{\mathbf{A}_t^l} \mathbf{Z}^l}{B_t + \lambda} \right), \quad (10)$$

式中, \mathbf{Z}^l 为从待检测帧中提取的 31 维 HOG 特征和 1 维 Gray 特征, $\overline{\mathbf{A}_t^l}$ 为矩阵的共轭, \mathbf{y} 最大值所在的位置是当前帧目标的中心位置。

综上所述,KCF 算法的优势在于保证跟踪速度的同时,在目标周围进行循环采样,以增加训练样本、提升算法的性能,但目标尺寸变化较大时表现较差。FDSST 算法先确定位置再搜寻尺寸的方法可以应对目标尺寸变化的情况,但滤波器与检测框进行相关运算时,默认对检测框边缘进行就近填充,且训练样本单一,容易导致跟踪漂移现象。

3 本文算法

3.1 响应图置信度评价方法

在相关滤波跟踪算法中,PSR^[3]表示相关滤波响应峰值的尖锐度,可以衡量目标跟踪的置信度,相关滤波响应在峰值位置 x 处的 PSR 可表示为

$$X_{\text{PSR}}(x) = \frac{\max(x) - \mu(x)}{\sigma(x)}, \quad (11)$$

式中, $\max(x)$ 为响应图中的峰值, $\mu(x)$ 和 $\sigma(x)$ 分别为响应的均值和标准差。PSR 越大,表示响应图的置信度越高,反之则表示响应图的置信度越低。在目标被遮挡或跟踪过程中出现相似物体干扰时,PSR 不能及时反映目标的异常情况,容易导致跟踪漂移甚至失败。

文献[12]提出用平均峰值相关能量(APCE)衡量目标跟踪的置信度,基于此,实验用响应图的 SMR 衡量响应图的遮挡或异常情况,主要步骤如下。

1) 求取当前帧响应图的所有峰值,并将峰值按降序排序。

2) 将最大值作为主峰 X_m , 选取剩下峰值的前 50% 作为次峰 X_s , 对于次峰数目较少的情况用零值填充。

3) 相关滤波响应的 SMR 可表示为

$$X_{\text{SMR}} = \frac{\text{mean}(X_s)}{X_m}. \quad (12)$$

4) 确定每帧响应图的 SMR 后,当前帧前若干帧的 SMR 均值可表示为 MSMR,成功跟踪时,特定序列的 SMR 和 MSMR 会稳定在一定范围内;当跟踪失败或发生漂移时,SMR 变化明显,MSMR 则会保持在相对稳定的区间内。

在 MSMR 的基础上设置合理的阈值,当某一帧的 SMR 超过该阈值时,认为目标存在异常,即目标可能被局部遮挡或存在较大干扰。图 1 为跟踪失败序列的图像与对应的响应图,可以发现,当目标未被遮挡或干扰较小时,响应图置信度较大且只有单个主峰;当目标被遮挡时,响应图会出现较大的次峰干扰。图 2 为序列 Human4 相关指标的变化图,可以发现,在 130 帧左右目标被局部遮挡,SMR 超过设定的阈值,表明目标被干扰或跟踪发生异常。此时模型如果继续更新会被污染,导致跟踪失败,阈值也会出现较大波动。

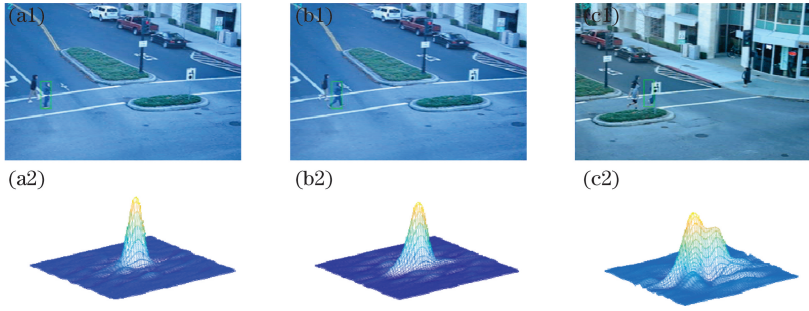


图1 跟踪失败时的结果。(a1)~(c1)不同帧的图像;(a2)~(c2)对应的响应图

Fig. 1 Result when the trace fails. (a1)–(c1) Images of different frames; (a2)–(c2) corresponding response images

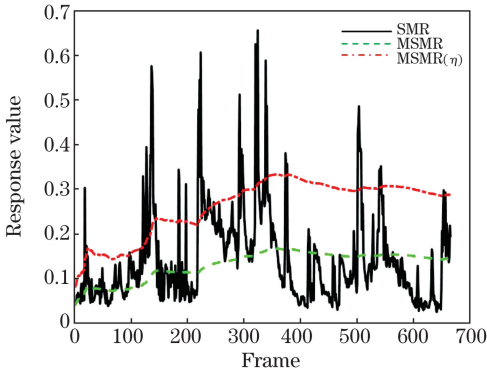


图2 Human4序列三个指标的变化曲线

Fig. 2 Change curves of the three indicators of the Human4 sequence

3.2 并行相关滤波跟踪算法

PSR可以评价目标跟踪响应图的置信度,SMR可从响应图的峰值角度判定目标是否出现异常。为了充分利用响应图的有效信息,在(11)式、(12)式的基础上,融合了两个维度的信息,提出了新的置信度评价函数 c ,可表示为

$$c = X_{\text{PSR}} + \frac{1}{X_{\text{SMR}}} = \frac{\max(x) - \mu(x)}{\sigma(x)} + \frac{X_m}{\text{mean}(X_s)}. \quad (13)$$

基于(13)式的置信度评价函数,结合FDSST与KCF算法,提出了一种并行相关滤波跟踪算法,具体步骤如下。

1) 特征提取。KCF算法提取31维的HOG特征、1维的Gray特征和10维的CN特征;FDSST算法提取31维的HOG特征和1维的Gray特征。

2) 采样方式。KCF算法在提取特征环节和检测环节均采用循环采样方式增加样本数量,FDSST算法则用1.5倍初始框大小的提取框提取目标特征;不同的采样方式可以弥补两种算法在特征处理上的不足,防止欠拟合或过拟合;不同的初始化信

息可以保证滤波器学习到不同的信息。

3) 位置预测。

$$\epsilon_K = \min \sum_i [f(\mathbf{x}) - \mathbf{y}]^2 + \lambda_K \|\mathbf{w}\|^2, \quad (14)$$

$$\epsilon_F = \left\| \sum_{l=1}^d \mathbf{h}^l * \mathbf{f}^l - \mathbf{g} \right\|^2 + \lambda_F \sum_{l=1}^d \|\mathbf{h}^l\|^2, \quad (15)$$

式中, ϵ_K, λ_K 和 ϵ_F, λ_F 分别为KCF算法与FDSST算法的误差函数与正则化系数。通过傅里叶变换得到不同算法在当前帧的响应值,可表示为

$$\mathbf{y}_K = \mathcal{F}^{-1}(\hat{\kappa}^{xz} \odot \hat{\alpha}), \quad (16)$$

$$\mathbf{y}_F = \mathcal{F}^{-1} \left\{ \frac{\sum_{l=1}^d \overline{\mathbf{A}}_p^l \mathbf{Z}^l}{B_p + \lambda} \right\}, \quad (17)$$

式中, $\mathbf{y}_K, \mathbf{y}_F$ 分别为KCF和FDSST算法的响应值, \mathbf{A}_p, B_p 分别为用于位置预测的滤波器模型分子与分母。当某个算法的响应值置信度较大,表明其在当前帧中的预测结果相对准确,定义当前帧的实际响应值为

$$\mathbf{y}_p = \mathcal{F}^{-1} \left\{ \frac{c_K}{c_K + c_F} \times \hat{\kappa}^{xz} \odot \hat{\alpha} + \frac{c_F}{c_K + c_F} \times \frac{\sum_{l=1}^d \overline{\mathbf{A}}_p^l \mathbf{Z}^l}{B_p + \lambda_F} \right\}, \quad (18)$$

式中, c_K, c_F 为KCF和FDSST算法在当前帧的置信度。这种方式可对两种跟踪算法的响应值权重进行自适应调整,并根据最终响应 \mathbf{y}_p 确定目标在当前帧的位置。

4) 尺寸预测。在步骤3)确定的目标位置基础上,提取31维的HOG特征构成1维输入特征,选取33个尺寸,每个特征为 $1 \times S$ 的向量,输入特征 \mathbf{Z} 的输出结果 \mathbf{y}_s 可表示为

$$\mathbf{y}_s = \mathcal{F}^{-1} \left\{ \frac{\sum_{l=1}^d \overline{\mathbf{A}}_s^l \mathbf{Z}^l}{B_s + \lambda_F} \right\}, \quad (19)$$

式中, \mathbf{A}_s, B_s 分别为尺寸预测滤波器模型的分子与分母。

3.3 模型更新策略

目标跟踪过程中,可能会受到严重的背景干扰或遮挡,如果固定在起始帧获得跟踪模型,不能适应目标外观的轻微变化。因此,相关滤波算法在跟踪过程中对跟踪模型进行更新,以适应目标变化。但如果对跟踪模型逐帧更新,在目标被遮挡或干扰较大时容易引入噪声。文献[13]中的更新频率根据视频中跟踪目标的运动速度进行动态调整,以提升算法的速度,但当目标被遮挡时仍会引入噪声。因此,实验根据 SMR 设置阈值,若当前帧的 SMR 小于 MSMR 时,表示跟踪过程不存在干扰或干扰较小,滤波器模型按初始化学学习率逐帧更新模型;当 SMR 大于 MSMR 且小于阈值时,目标可能存在轻微干扰,此时用更小的学习率更新模型;当 SMR 大于阈值时,目标存在较大干扰,模型不更新。

在本算法中,有两个位置跟踪模型,分别是 KCF 和 FDSST 算法用于确定目标位置的模型,独立的更新策略能保证每个跟踪器互不干扰,更新方式可表示为

$$\begin{cases} \hat{\mathbf{a}}^t = (1 - \eta_K) \hat{\mathbf{a}}^{t-1} + \eta_K \frac{\hat{\mathbf{y}}}{\hat{\kappa}^{xx} + \lambda'_K}, \\ \hat{\mathbf{x}}^t = (1 - \eta_K) \hat{\mathbf{x}}^{t-1} + \eta_K \hat{\mathbf{x}}' \\ \eta_K = \begin{cases} \eta_{K1}, X_{SMR} < X_{MSMR} \\ \eta_{K2}, X_{SMR} < \beta \times X_{MSMR} \\ 0, X_{SMR} \geq \beta \times X_{MSMR} \end{cases}, \quad (20) \\ \mathbf{A}_{tp}^t = (1 - \eta_F) \mathbf{A}_{(t-1)p}^t + \eta_F \mathbf{G}_t * \mathbf{F}_t^t \\ \mathbf{B}_{tp}^t = (1 - \eta_F) \mathbf{B}_{(t-1)p}^t + \eta_F \sum_{k=1}^d \mathbf{F}_t^k * \mathbf{F}_t^t \\ \eta_F = \begin{cases} \eta_{F1}, X_{SMR} < X_{MSMR} \\ \eta_{F2}, X_{SMR} < \beta \times X_{MSMR} \\ 0, X_{SMR} \geq \beta \times X_{MSMR} \end{cases}, \quad (21) \end{cases}$$

式中, η_K, η_F 分别为 KCF、FDSST 的位置模型学习率, β 为阈值系数, $\hat{\mathbf{a}}^t$ 为当前帧的 KCF 滤波器模型因子, $\mathbf{A}_{tp}^t, \mathbf{B}_{tp}^t$ 为当前帧 FDSST 的位置与尺寸滤波器模型因子, $\hat{\mathbf{a}}^{t-1}, \mathbf{A}_{(t-1)p}^t, \mathbf{B}_{(t-1)p}^t$ 为上一帧的滤波器模型因子。由于位置变化往往大于尺寸变化,因此,实验在进行尺寸跟踪时,只采取两种学习率进行更新,可表示为

$$\begin{cases} \mathbf{A}_{ts}^t = (1 - \eta_s) \mathbf{A}_{(t-1)s}^t + \eta_s \mathbf{G}_t * \mathbf{F}_t^t \\ \mathbf{B}_{ts}^t = (1 - \eta_s) \mathbf{B}_{(t-1)s}^t + \eta_s \sum_{k=1}^d \mathbf{F}_t^k * \mathbf{F}_t^t \\ \eta_s = \begin{cases} \eta_s, X_{SMR} < \beta \times X_{MSMR} \\ 0, X_{SMR} \geq \beta \times X_{MSMR} \end{cases}, \quad (22) \end{cases}$$

式中, η_s 为尺寸模型的学习率, $\mathbf{A}_{ts}^t, \mathbf{B}_{ts}^t$ 为当前帧 FDSST 的位置与尺寸滤波器模型因子, $\mathbf{A}_{(t-1)s}^t, \mathbf{B}_{(t-1)s}^t$ 为上一帧的滤波器模型因子。

3.4 算法流程

本算法的具体流程如下。

1) 训练滤波器:在视频序列的第一帧图像中,从初始化信息中获取目标的位置坐标和尺寸,并设置目标候选框的尺寸;设计两个 2 维的相关滤波器用于预测目标的位置以及一个 1 维的相关滤波器用于预测目标的尺寸;用 2 维和 1 维高斯分布函数生成理想响应图作为第一帧图像的输出;由输入特征和理想输出得到对应的最初相关滤波器模型,用于后续帧的目标检测与跟踪和模型更新。

2) 计算置信度:从候选框中提取输入特征,与对应滤波器模型进行相关运算得到当前帧的输出响应图;找出响应图中的主峰和所有次峰。用计算得到的 SMR 作为决定更新策略的依据;将 SMR 和 PSR 的组合值 c 作为响应图的置信度评价标准。

3) 融合相关滤波器的响应图:为了进一步提高目标跟踪的精度和鲁棒性,将两个具有不同采样方式和特征的滤波器模型进行融合,将最终响应图中最大值所在的位置作为目标的中心位置。

4) 更新模型:为了避免目标在被遮挡或受到较大干扰时更新模型造成的滤波器污染,根据 SMR 设置不同的区间;赋予滤波器模型不同的学习率,制定自适应的模型更新策略,以缓解滤波器污染的问题。

4 实验和结果分析

所有实验均在 Matlab R2016b 中实现,硬件配置:计算机的内存为 8 GB,处理器为 INTEL i5-5257U。FDSST 算法初始化的位置滤波器学习率 $\eta_p = 0.026$,尺寸滤波器的学习率 $\eta_s = 0.025$ 。用 3 个实验对算法进行评估,实验 1 的消融实验用于分析本算法的有效性;实验 2 将本算法与当前主流跟踪算法在 OTB-2013 和 OTB-2015 数据集上进行对比定性分析;实验 3 对本算法进行定量分析。

4.1 数据集以及评估方法

选取 OTB-2013^[14] 和 OTB-2015^[15] 数据集公开的测试集视频序列进行跟踪性能测试,该测试集按照目标跟踪中常见的难点将视频序列中的属性分成 11 类,包括平面旋转(IPR)、遮挡(OCC)、目标形变(DEF)、尺寸变化(SV)、运动模糊(MB)、目标消失(OV)、快速运动(FM)、光照变化(IV)、非平面旋

转(OPR)、背景杂乱(BC)和低分辨率(LR)。测试集中的每个视频序列至少包含其中一种属性,采用一次性通过评估(OPE)测试本算法的目标跟踪性能。

4.2 消融实验

为了验证不同算法的性能,设置了5组对比实验,其中,ours_KF为并行相关滤波算法,ours_UP为在基础算法上加入模型更新策略的算法,ours为所提出的并行相关滤波跟踪算法。不同算法在OTB-2015数据集得到的成功率和精度如图3所示

示,可以发现,相比FDSST算法,ours,ours_KF,ours_UP的跟踪成功率分别提升了1.5%,0.7%,0.1%,跟踪精度分别提升了3.6%,2.4%,1.9%。为了进一步表明本算法的有效性,在OTB-2015数据集中选取了两组存在遮挡的视频序列,不同算法的跟踪效果如图4所示,图例后的数值分别表示成功率或跟踪精度。可以发现,当目标存在短暂的遮挡或消失时,只有本算法跟踪成功,原因是本算法使用响应图置信度评价方法。

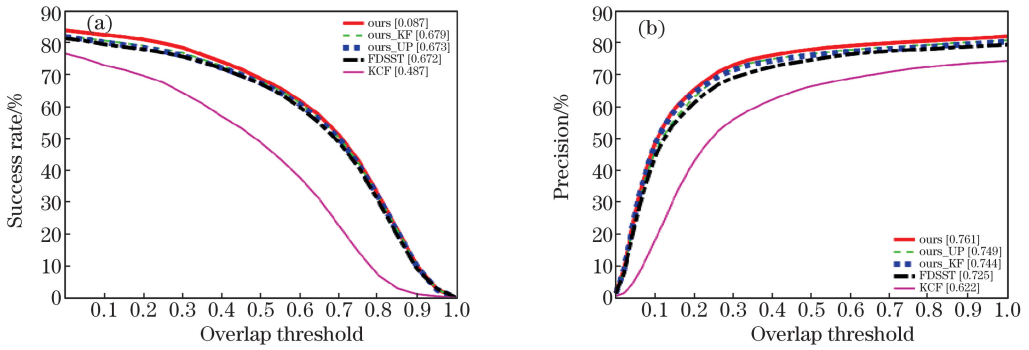


图3 不同算法在 OTB-2015 数据集上的 OPE。(a) 成功率;(b) 精度

Fig. 3 OPE of different algorithms on the OTB-2015 dataset. (a) Success rate; (b) precision

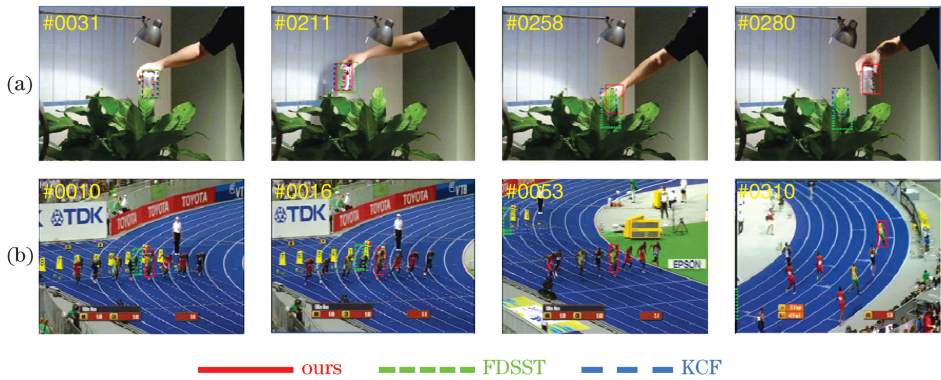


图4 不同算法在 OTB-2015 数据集上的跟踪效果。(a) Coke;(b) bolt

Fig. 4 Tracking effect of different algorithms on the OTB-2015 dataset. (a) Coke; (b) bolt

4.3 与主流算法的对比实验

选取9种主流的相关滤波跟踪算法与本算法进行对比,包括SAMF(Scale adaptive kernel correlationfilter tracker with feature integration)^[16], FDSST^[11], Staple^[7], LMCF(Large margin object tracking with circulant feature maps)^[12], LCT(Long-term correlation tracking)^[17], DSST(Accurate scale estimation for robust visual tracking)^[18], KCF^[5], CSK^[4], DFT

(Distribution fields for tracking)^[19]。图5为10种算法在OTB-2013和OTB-2015数据集上的测试结果,可以发现,在OTB-2013数据集上,相比FDSST和KCF算法,本算法的成功率分别提升了1.1%和21.2%,精度分别提升了1.0%和15.7%;在OTB-2015数据集上,本算法的成功率为68.7%,精度为76.1%,在所有算法中的表现最好。

表1、表2为10种算法对11个属性在跟踪成功率和精度方面的得分情况。其中,曲线下面积

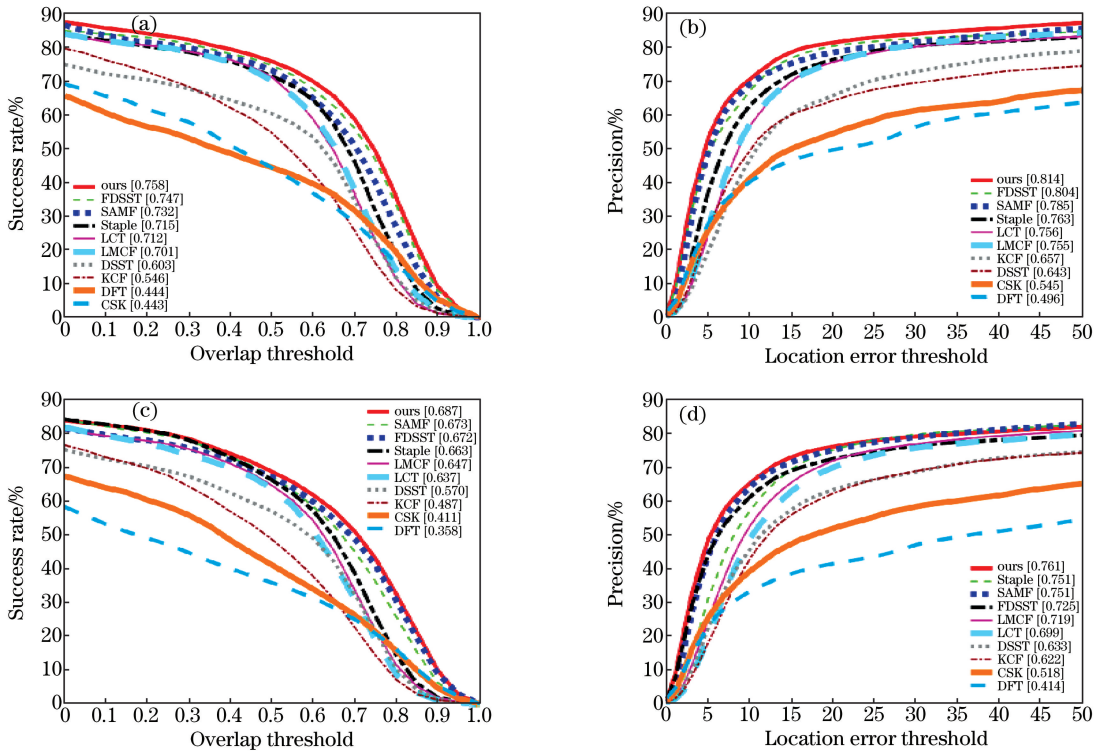


图 5 不同算法的测试结果。(a) OPE成功率(OTB-2013数据集);(b) OPE精度(OTB-2013数据集);(c) OPE成功率(OTB-2015数据集);(d) OPE精度(OTB-2015数据集)

Fig. 5 Test results of different algorithms . (a) OPE success rate (OTB- 2 0 1 3 dataset) ; (b) OPE precision (OTB- 2013 dataset) ; (c) OPE success rate (OTB-2015 dataset) ; (d) OPE precision (OTB-2015) dataset)

(AUC)为受试者工作特征(ROC)曲线与坐标轴围成的面积,表示最终的成功率;AP为跟踪的平均精度。可以发现,本算法在保证跟踪精度和成功率的同时,每秒传输帧数(FPS)可达 53.8 frame。且本

算法在跟踪成功率上,有 5 种属性的表现最好,4 种属性的表现第二,有 7 种属性的表现优于 FDSST 算法;在跟踪精度方面,有 6 种属性表现最好,3 种属性表现第二,有 9 种属性的表现优于 FDSST 算法。

表 1 10 种算法在 OTB-2015 数据集的成功率
Table 1 Success rate of 10 algorithms in the OTB-2015 dataset

Tracker	SV	IV	OPR	OCC	BC	DEF	MB	FM	IPR	OV	LR	AUC
Ours	61.1	73.4	64.2	62.3	72.7	59.5	65.9	64.5	67.2	59.5	61.9	68.7
SAMF	58.7	64.0	66.2	66.7	63.9	60.9	64.1	59.8	64.1	55.1	51.5	67.3
Staple	58.9	70.4	63.4	63.3	65.4	62.5	61.2	60.8	64.0	54.1	43.2	67.2
FDSST	59.8	72.0	59.5	59.4	72.8	57.3	66.4	67.1	66.1	54.8	58.9	66.3
LMCF	57.4	66.7	60.1	61.0	67.1	58.3	64.7	58.8	58.2	58.0	51.2	64.7
LCT	54.3	64.1	61.3	55.8	65.5	56.9	57.5	57.2	61.0	47.5	36.9	63.7
DSST	52.2	57.5	50.3	49.9	57.4	44.3	51.6	51.9	55.1	47.8	44.3	57.0
KCF	37.3	46.4	44.4	41.4	52.9	43.7	47.6	47.8	49.0	39.3	28.4	48.7
CSK	30.9	39.9	37.8	35.3	46.5	35.7	34.5	36.8	41.3	27.4	27.7	41.1
DFT	27.0	37.8	36.9	37.1	43.1	38.2	28.3	29.3	36.2	31.3	19.6	35.8

表2 10种算法在OTB-2015数据集上的精度
Table 2 Precision of 10 algorithms on the OTB-2015 dataset

Tracker	SV	IV	OPR	OCC	BC	DEF	MB	FM	IPR	OV	LR	AP
Ours	71.6	78.8	73.3	68.4	79.4	66.7	69.1	69.5	77.1	64.8	68.5	76.1
SAMF	70.2	70.8	73.9	72.2	68.9	68.0	65.5	65.5	72.1	62.8	68.5	75.1
Staple	72.5	77.8	73.0	70.9	71.4	70.5	63.9	66.6	75.3	59.3	60.9	75.1
FDSST	67.2	76.5	66.6	64.6	78.0	62.0	69.1	69.8	73.4	57.7	64.1	72.5
LMCF	67.7	73.8	68.7	68.2	72.0	65.2	66.5	63.5	67.5	57.2	67.9	71.9
LCT	63.0	69.7	69.0	60.3	66.3	64.8	54.9	57.2	70.7	43.6	50.5	69.9
DSST	62.5	63.2	57.1	54.4	62.2	50.3	51.4	51.9	64.1	50.1	64.3	63.3
KCF	58.0	63.2	59.5	52.6	62.2	55.9	50.5	54.8	63.4	37.1	54.6	62.2
CSK	45.2	49.0	48.9	43.3	57.4	45.7	35.5	39.7	51.4	27.6	42.2	51.8
DFT	35.5	41.6	44.0	42.0	45.5	41.7	27.2	30.5	43.4	35.3	29.0	41.4

4.4 定性分析实验

为了更直观地评估本算法的跟踪性能,选择定量分析中排名前6的算法进行实验。为了便于区分,用不同的矩形框标注不同算法的跟踪结果,共

展示了四组存在不同干扰信息的视频序列,图6(a)为运动模糊的情况,图6(b)为目标变形的情况,图6(c)为光照变化的情况,图6(d)为背景杂乱的情况,左上角的数字为视频序列的帧数。

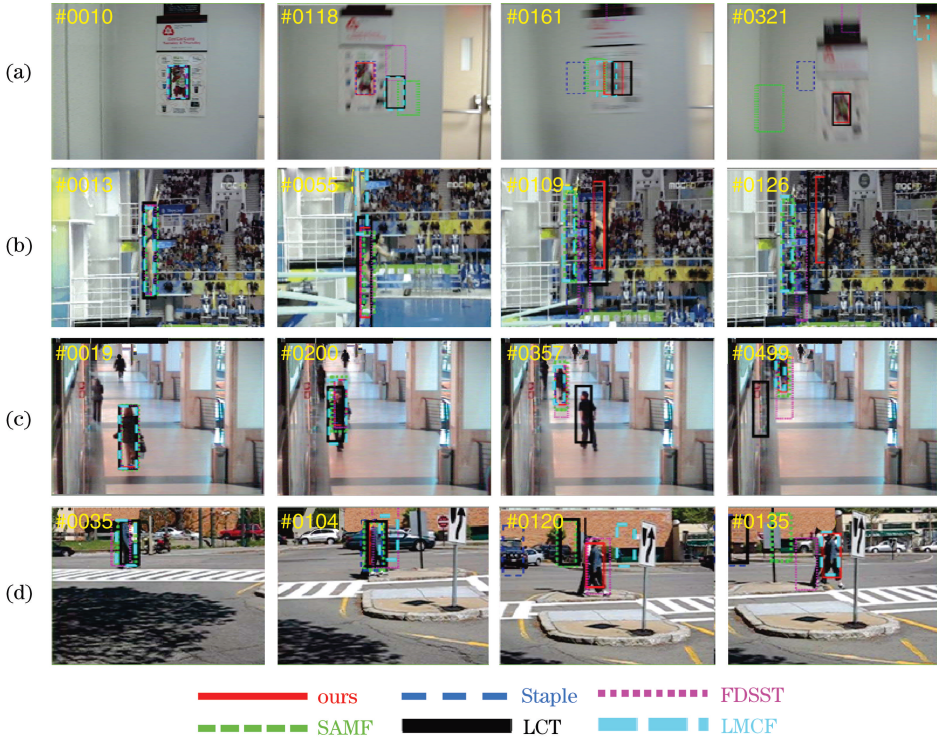


图6 不同算法的跟踪结果。(a) BlurOwl; (b) Diving; (c) Walking2; (d) Couple

Fig. 6 Tracking results of different algorithms. (a) BlurOwl; (b) Diving; (c) Walking2; (d) Couple

可以发现,图6(a)存在严重的运动模糊的情况,在161帧后于目标因背景剧烈振动而变模糊,其

他跟踪算法均出现跟踪漂移或丢失,但本算法能够成功跟踪。图6(b)在109帧目标发生变形,LMCF

和 FDSST 算法出现跟踪漂移;在 126 帧目标出现明显的姿态变化,只有本算法和 LCT 算法成功跟踪。图 6(c)存在严重的光照变化和遮挡,且视频序列伴随轻微的抖动,除本算法外,其他算法均出现不同程度的跟踪漂移。图 6(d)中背景信息多变,在 120 帧后,只有本算法和 LCT 算法成功跟踪,但本算法的跟踪精度更高。

4.5 失败案例分析

图 7 为本算法跟踪失败的情况,在图 7(a)

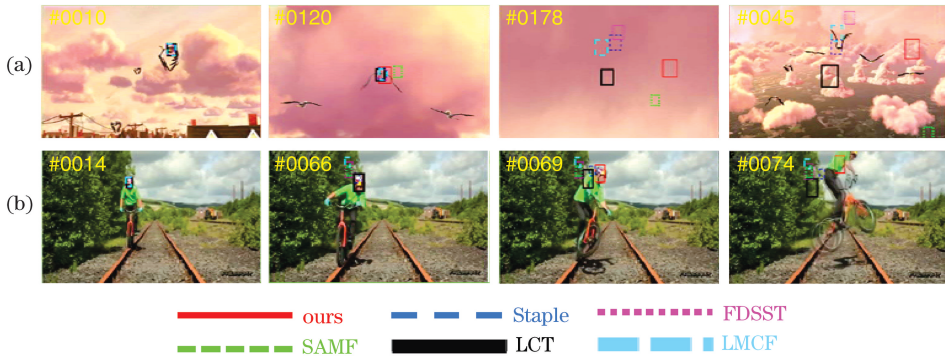


图 7 目标跟踪失败的情形。(a) Bird1; (b) Biker

Fig. 7 Case of target tracking failure. (a) Bird1; (b) Biker

5 结 论

在相关滤波的基础上,提出了一种基于置信度的并行相关滤波目标跟踪算法。首先,提出结合 PSR 的置信度评价方法,用于评价当前目标的状态;然后,融合两种滤波器的响应图,构建并行跟踪以增强跟踪器的鲁棒性。最后,根据置信度对跟踪算法采取自适应学习率的跟踪策略。实验结果表明,本算法能解决目标在跟踪过程中存在的光照变化、目标形变和遮挡等问题,有效提高了目标跟踪的成功率,整体性能优于其他对比算法。在实验中发现,如果跟踪过程中出现目标快速运动或长期完全消失的情况,本算法容易出现跟踪漂移甚至失败的问题,下一步还需对目标完全消失或跟踪失败后重新定位目标的问题进行研究。

参 考 文 献

- [1] Zhang H L, Hu S Q, Yang G S. Video object tracking based on appearance models learning [J]. Journal of Computer Research and Development, 2015, 52(1): 177.
- [2] Luo H B, Xu L Y, Hui B, et al. Status and prospect of target tracking based on deep learning [J]. Infrared

and Laser Engineering, 2017, 46(5): 0502002. 罗海波, 许凌云, 惠斌, 等. 基于深度学习的目标跟踪方法研究现状与展望 [J]. 红外与激光工程, 2017, 46(5): 0502002.

中,目标被长时间完全遮挡,导致跟踪算法发生漂移,只有 LCT 算法因目标出现后的重新检测机制重新搜索到目标。在图 7(b)中,由于被跟踪的目标框尺寸较小,提取的输入特征不足,且同时存在多种因素干扰,所有的跟踪算法均跟踪失败。因此,当目标同时存在多种因素干扰或目标因长期被完全遮挡而导致提取特征困难时,检测框无法正确定位搜索区域,极易导致目标跟踪失败。

- [3] Bolme D S, Beveridge J R, Draper B A, et al. Visual object tracking using adaptive correlation filters [C]// 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 13-18, 2010, San Francisco, CA, USA. New York: IEEE, 2010: 2544-2550.
- [4] Henriques J F, Caseiro R, Martins P, et al. Exploiting the circulant structure of tracking-by-detection with kernels [M]// Fitzgibbon A, Lazebnik S, Perona P, et al. Computer Vision-ECCV 2012. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2012, 7575: 702-715.
- [5] 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.
- [6] Danelljan M, Khan F S, Felsberg M, et al. Adaptive color attributes for real-time visual tracking [C]// 2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE, 2014: 1090-1097.

- [7] Bertinetto L, Valmadre J, Golodetz S, et al. Staple: complementary learners for real-time tracking [C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 1401-1409.
- [8] Li C, Lu C Y, Zhao X, et al. Scale adaptive correlation filtering tracking algorithm based on feature fusion [J]. *Acta Optica Sinica*, 2018, 38 (5): 0515001.
李聪, 鹿存跃, 赵珣, 等. 特征融合的尺寸自适应相关滤波跟踪算法 [J]. *光学学报*, 2018, 38 (5): 0515001.
- [9] Shen Q, Yan X L, Liu L F, et al. Multi-scale correlation filtering tracker based on adaptive feature selection [J]. *Acta Optica Sinica*, 2017, 37 (5): 0515001.
沈秋, 严小乐, 刘霖枫, 等. 基于自适应特征选择的多尺度相关滤波跟踪 [J]. *光学学报*, 2017, 37 (5): 0515001.
- [10] Zhao G P, Shen Y P, Wang J Y. Adaptive feature fusion object tracking based on circulant structure with kernel [J]. *Acta Optica Sinica*, 2017, 37 (8): 0815001.
赵高鹏, 沈玉鹏, 王建宇. 基于核循环结构的自适应特征融合目标跟踪 [J]. *光学学报*, 2017, 37 (8): 0815001.
- [11] Danelljan M, Häger G, Khan F S, et al. Discriminative scale space tracking [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(8): 1561-1575.
- [12] Wang M M, Liu Y, Huang Z Y. Large margin object tracking with circulant feature maps [C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 4800-4808.
- [13] Danelljan M, Robinson A, Khan F S, et al. Beyond correlation filters: learning continuous convolution operators for visual tracking [M]// Leibe B, Matas J, Sebe N, et al. *Computer Vision-ECCV 2016. Lecture Notes in Computer Science*. Cham: Springer, 2016, 9909: 472-488.
- [15] Wu Y, Lim J, Yang M H. Object tracking benchmark [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37 (9): 1834-1848.
- [16] Li Y, Zhu J K. A scale adaptive kernel correlation filter tracker with feature integration [M]// Agapito L, Bronstein M, Rother C, et al. *Computer Vision-ECCV 2015. Lecture Notes in Computer Science*. Cham: Springer, 2015, 8926: 254-265.
- [17] Ma C, Yang X, Zhang C, et al. Long-term correlation tracking [C]// 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE, 2015: 5388-5396.
- [18] Danelljan M, Hager G, Khan F S, et al. Accurate scale estimation for robust visual tracking [C]// *British Machine Vision Conference*, 2014, September 1-5, 2014, Nottingham, UK. Nottingham: BMVA, 2014: 65.
- [19] Sevilla-Lara L, Learned-Miller E. Distribution fields for tracking [C]// 2012 IEEE Conference on Computer Vision and Pattern Recognition, June 16-21, 2012, Providence, RI, USA. New York: IEEE, 2012: 1910-1917.