

文章编号 1004-924X(2018)02-0492-11

利用最佳伙伴相似性的改进空间 正则化判别相关滤波目标跟踪

杨德东, 毛 宁*, 杨福才, 李雪晴

(河北工业大学 控制科学与工程学院, 天津 300130)

摘要:针对空间正则化判别相关滤波跟踪算法(SRDCF)在目标发生遮挡、尺度变化和形变情况下的跟踪失败问题,提出利用最佳伙伴相似性的改进 SRDCF 目标跟踪算法。首先,以 SRDCF 算法为基础,利用双层搜索策略解决目标跟踪中的目标定位问题和尺度估计问题;然后,利用一种新颖的鲁棒模板匹配技术,通过融合空间权重、相关滤波得分和最佳伙伴相似性得分来估计候选目标位置,解决遮挡情况下的目标重定位问题;最后,采用自适应模板更新策略解决遮挡情况下模板漂移问题。本文采用 OTB-2013 数据集评估本文算法的性能,同时与 34 种流行算法进行比较,结果表明本文算法的精确度得分和成功率得分分别为 0.853 和 0.648,相比传统的 SRDCF 算法分别提高 1.79% 和 3.51%。本文算法能很好地解决目标遮挡、尺度变化和形变情况下的目标跟踪问题,具有一定研究价值。

关键词:目标跟踪;最佳伙伴相似性;判别式相关滤波;空间正则化

中图分类号:TP391 **文献标识码:**A **doi:**10.3788/OPE.20182602.0492

Improved SRDCF object tracking via the Best-Buddies Similarity

YANG De-dong, MAO Ning*, YANG Fu-cai, LI Xue-qing

(College of Control Science and Engineering, Hebei University of Technology, Tianjin 300130, China)

* Corresponding author, E-mail: maon316@163.com

Abstract: Aiming at the failure of tracking via spatially regularized discriminant correlation filter (SRDCF) algorithm caused by occlusion, scale change and deformation, an improved SRDCF algorithm based on Best-Buddies Similarity was proposed. Firstly, the proposed algorithm based on SRDCF, locating target and estimating scale in the process of object tracking were complemented by using bi-level search strategy. Secondly, a novel robust template matching technique was used to estimate the candidate object position by integrating the spatial weights, the correlation filter score and the Best-Buddies Similarity score, thus the problem of target relocation in the occlusion was resolved. Finally, the adaptive template updating strategy was employed to mitigate the template drift problem in the case of occlusion. The performance of the proposed algorithm was evaluated on OTB-2013 datasets and was compared with 34 popular algorithms. The results show that the accuracy and the success rate of the proposed algorithm are 0.853 and 0.648, which are 1.79% and 3.51% higher than the traditional SRDCF algorithm, respectively. The proposed algorithm can deal with the matter of occlu-

收稿日期:2017-06-16;修订日期:2017-08-14.

基金项目:国家自然科学基金资助项目(No. 61203076);河北省自然科学基金资助项目(No. F2017202009)

sion, scale change and deformation effectively, and has some value of research.

Key words: object tracking; Best-Buddies Similarity; discriminative correlation filter; spatial regularization

1 引言

在线目标跟踪作为机器视觉中的重要角色已经被广泛应用到自动驾驶、智能监控和人机交互等^[1]领域。在目标发生遮挡、形变和尺度变化等情况下,目标外观模型发生剧烈变化,此时很难解决复杂场景下的精确目标跟踪问题。因此如何设计一种鲁棒的目标跟踪算法来解决复杂场景下的跟踪问题仍是非常艰巨的任务。

近年来为了解决复杂情况下的目标跟踪问题,涌现出了大量优秀的算法^[2-7]。Kalal 等^[2]提出的跟踪学习检测算法(Tracking Learning Detection, TLD)是一种新颖的单目标长期跟踪算法。TLD 在传统跟踪算法的基础上引入在线学习模块和检测模块,通过在线学习模块和检测模块解决目标发生形变和遮挡情况下的样本自适应学习问题和目标重定位问题,从而实现稳定和可靠的实时目标跟踪;Jia 等^[4]提出基于自适应结构化局部稀疏外观模型的跟踪算法(Adaptive Structural Local Sparse Appearance, ASLA),其以粒子滤波算法为基础,在传统稀疏表示的基础上构造结构化局部稀疏模型,通过利用目标的稀疏结构化信息和空间信息解决遮挡和相似目标混淆情况下的跟踪问题;Zhong 等^[5]提出利用稀疏协同模型的鲁棒目标跟踪算法(Sparsity-based Collaborative Model, SCM),其通过构造一个稀疏判别模型和一个稀疏生成模型,解决目标外观模型剧烈变化情况下的目标跟踪问题,降低此情况下目标跟踪漂移现象。

然而上面的方法大多以生成式方法为主,需要构造复杂的目标外观模型,计算复杂度较高,不适合在线实时目标跟踪。而最近以相关滤波算法为主的判别式跟踪算法^[8-15]引起科研人员的注意,相关滤波算法结构简单,不需要构造复杂外观模型,计算复杂度小,非常适合解决在线目标跟踪问题。Henriques 等^[8]提出高速核化相关滤波算法(Kernelized Correlation Filter, KCF),其通过利用循环结构构造大量训练样本,以提高分类器

的分类能力,然后利用循环结构的性质和快速傅里叶变换将时域内的计算转换到频域,极大地降低计算复杂度,实现了在线实时目标跟踪。Danelljan 等^[11]提出精确尺度估计的鲁棒目标跟踪算法(Accurate Scale Estimation for Robust Visual Tracking, DSST),其通过利用尺度池策略,解决相关滤波跟踪算法中的目标尺度估计问题,提高了跟踪的精确性。Danelljan 等^[12]又提出稀疏正则化判别式相关滤波算法(Spatially Regularized Correlation Filters, SRDCF),通过引入空间正则化权重来降低背景中样本所占权重,提高目标中心区域样本的空间权重,降低相关滤波算法中存在的边界效应,显著提高了相关滤波算法的跟踪精确性。

但是上述算法不能很好地解决目标剧烈尺度变化、形变和遮挡情况下的跟踪问题,本文在 SRDCF 算法基础上提出利用最佳伙伴相似性的改进 SRDCF 目标跟踪算法。为了解决目标遮挡和形变情况下的跟踪问题,本文通过结合一种鲁棒的目标模板匹配技术^[16]设计出在线重检测模块,解决遮挡和剧烈形变情况下的重定位问题。然后,通过利用双层搜索策略和自适应更新策略,解决目标精确位置估计和尺度估计。为了更好地验证本文算法的性能,利用 OTB-2013 数据集^[17]中的评价标准和 50 组视频序列测试本文算法,同时与 34 种流行算法进行性能比较。

2 SRDCF 跟踪算法

SRDCF 跟踪算法^[12]是传统相关滤波算法的拓展,由于传统相关滤波算法采用循环偏移结构产生训练样本,使得相关滤波算法中存在边界效应,其严重影响相关滤波算法跟踪的精确性。而为了降低边界效应的影响,SRDCF 算法在传统相关滤波算法的基础上引入稀疏正则化空间权重。

传统的相关滤波算法将目标跟踪问题转换为在线分类问题,即通过求解最小均方误差下的岭回归问题:

$$\epsilon_t(h) = \sum_{k=1}^t \alpha_k \|f(x_k) - y_k\|^2 + \lambda \sum_{l=1}^d \|h^l\|^2, \quad (1)$$

其中: t 表示第 t 帧, y_k 表示目标期望响应, 其服从高斯分布, $f(x)$ 表示跟踪目标的输出响应, d 表示特征空间的维数, α_k 表示样本学习的权重因子, x_k 表示第 k 帧的训练样本, λ 是正则化因子, h 表示滤波器参数。

为了降低边界效应, SRDCF 算法中引入稀疏空间正则化权重因子 ω 。使得目标区域的正则化权重要大于背景部分的正则化权重, 则公式(1)变为:

$$\epsilon_t(h) = \sum_{k=1}^t \alpha_k \|f(x_k) - y_k\|^2 + \sum_{l=1}^d \|\omega \cdot h^l\|^2, \quad (2)$$

其中 \cdot 表示对应位置元素的点积。

对公式(2)应用帕萨瓦尔定理, 并向量化则有:

$$\tilde{\epsilon}(\tilde{\mathbf{h}}^1 \cdots \tilde{\mathbf{h}}^d) = \sum_{k=1}^t \alpha_k \left\| \sum_{l=1}^d D_k \tilde{\mathbf{h}}^l - \tilde{\mathbf{y}}_k \right\|^2 + \sum_{l=1}^d \|\mathbf{C} \tilde{\mathbf{h}}^l\|^2, \quad (3)$$

其中: $\tilde{\mathbf{h}}^l = \mathbf{B} \hat{\mathbf{h}}^l$, 上标 \cdot 表示离散傅立叶变换, 粗体表示相应函数的向量化, \mathbf{B} 是构造的傅立叶传递变换矩阵, 其保证 $\tilde{\mathbf{h}}^l$ 是关于 $\hat{\mathbf{h}}^l$ 的实数, $\tilde{\mathbf{y}}_k = \mathbf{B} \hat{\mathbf{y}}_k$, $\mathbf{C} = (1/MN) \mathbf{B} \mathbf{C}(\hat{\mathbf{w}}) \mathbf{B}^H$, $\mathbf{D}_k^l = \mathbf{B} \mathbf{D}(\hat{\mathbf{x}}_k^l) \mathbf{B}^H$, 这里 H 表示共轭转置操作, D 表示对角化操作, C 表示 2 维循环卷积操作, MN 表示当前帧输入训练样本的数目。

对公式(3)进行简化, 则有:

$$\tilde{\epsilon}(\tilde{\mathbf{h}}) = \sum_{k=1}^t \alpha_k \|\mathbf{D}_k \tilde{\mathbf{h}} - \tilde{\mathbf{y}}_k\|^2 + \sum_{l=1}^d \|\mathbf{W} \tilde{\mathbf{h}}\|^2, \quad (4)$$

其中: $\tilde{\mathbf{h}} = ((\tilde{\mathbf{h}}^1)^T \cdots (\tilde{\mathbf{h}}^d)^T)^T$, $\mathbf{D} = (\mathbf{D}_k^1 \cdots \mathbf{D}_k^d)$, \mathbf{W} 是关于 \mathbf{C} 的 $dMN \times dMN$ 维块对角矩阵。

使公式(4)的损失误差最小化, 相当于求解正规矩阵 $\mathbf{A}_t \tilde{\mathbf{h}} = \tilde{\mathbf{b}}_t$, 有:

$$\mathbf{A}_t = \sum_{k=1}^t \alpha_k \mathbf{D}_k^T \mathbf{D}_k + \mathbf{W}^T \mathbf{W}, \quad (5)$$

$$\tilde{\mathbf{b}}_t = \sum_{k=1}^t \alpha_k \mathbf{D}_k^T \tilde{\mathbf{y}}_k, \quad (6)$$

其中: $\tilde{\mathbf{h}}^l$ 就是获取的最佳滤波参数, $\tilde{\mathbf{b}}_t$ 和 \mathbf{A}_t 分别表示是滤波器分式的分子与分母。

3 本文算法

为了解决 SRDCF 算法在目标发生遮挡、剧

烈尺度变化和形变情况下容易跟踪失败的问题, 本文提出利用最佳伙伴相似性的改进 SRDCF 跟踪算法, 通过在线重检测模块解决遮挡情况下的跟踪问题, 提升了跟踪算法的鲁棒性和可靠性。

3.1 双层搜索策略的尺度自适应目标跟踪

为了更好地解决目标精确定位和尺度估计问题, 本文提出在线双层搜索策略的尺度自适应目标跟踪, 设计两个相关滤波器, 一个用于目标位置估计, 一个用于目标尺度估计。

目标位置估计: 在进行目标位置估计时, 本文以 SRDCF 跟踪器为基础, 通过扩大跟踪搜索区域来增加分类器的训练样本, 提升分类器的分类能力, 提升跟踪算法的鲁棒性。这里直接利用上一帧训练的 SRDCF 跟踪器 \tilde{h}_{l_0} , 输入搜索区域 z (目标大小的 4^2 倍搜索区域), 计算搜索区域的滤波响应有:

$$f_{l_0}(z) = F^{-1} \left\{ \sum_{l=1}^d \tilde{z}^l \cdot \tilde{h}_{l_0}^l \right\}, \quad (7)$$

其中: F^{-1} 表示傅立叶逆变换, $f_{l_0}(z)$ 是定位模块计算得到目标搜索区域的时域响应, 通过寻找响应图中最大响应所在位置即估计目标位置。

尺度估计: 为了更精确地对目标尺度进行估计, 本文单独设计一个尺度估计器 \tilde{h}_s , 训练尺度估计器时, 为了保证尺度估计的精确性, 只利用目标样本信息进行训练分类器, 而文献[11]中训练分类器时利用的是目标信息和背景信息。同时训练的尺度估计还有一个作用是用于重检测模块中候选目标响应置信度计算, 此部分作用将在 3.3 节介绍。

首先, 在目标位置被确定的情况下, 从当前估计位置处获取不同尺度的图像块, 每个图像块大小为 $(s_i \omega, s_i h)$, 其中 s_i 为尺度因子, $s_i \in S$, S 为尺度池。然后对获取的图像进行插值处理, 使其大小变为 (ω, h) 。最后利用尺度滤波器, 计算响应最大的图像所在尺度即为当前目标尺度:

$$\hat{s} = \arg \max_{s_i} (f_{\max}^{s_1}, f_{\max}^{s_2}, \dots, f_{\max}^{s_S}), \quad (8)$$

其中, $f_{\max}^{s_i}$ 表示 s_i 尺度时尺度响应图中的最大响应值, 通过 \hat{s} 可以知道当前目标尺度变化情况。

3.2 利用最佳伙伴相似度的鲁棒模板匹配

为了更好地解决目标遮挡和形变等情况下的目标跟踪问题, 本文利用了一种新颖的鲁棒模板匹配技术, 下面简单介绍其原理, 其详细介绍和推导请参考文献[16]。

利用最佳伙伴相似度(BBS)能很好地解决遮挡、形变和背景嘈杂情况下的鲁棒模板匹配,BBS是计算两组点相似性的方法,为了利用BBS方法,首先,需要将图像分割成一个个 $l \times l$ 大小的子图像块,然后将其转化成大小为 l^2 的向量。则此时目标模板和匹配图像分别变成两组点集 $T = \{t_i\}_{i=1}^H$ 和 $O = \{o_j\}_{j=1}^J$,然后计算两组点之间的最佳伙伴对(Best-Buddies Pairs, BBP),BBP的定义为:

$$bb(t_i, o_j, T, O) = \begin{cases} 1 & \text{NN}(t_i, O) = o_j \wedge \text{NN}(o_j, T) = t_i \\ 0 & \text{otherwise} \end{cases}, \quad (9)$$

其中: $\text{NN}(t_i, O) = \arg \min_{o \in O} d(t_i, o)$,其表示点集 O 中 t_i 的最近邻, $d(\cdot)$ 表示距离函数,其计算两个点之间的距离。

BBS是与BBP相关的,其定义为:

$$\text{BBS}(T, O) = \frac{1}{\min\{H, J\}} \sum_{i=1}^H \sum_{j=1}^J bb(t_i, o_j, T, O), \quad (10)$$

其中: H 和 J 分别点集 T 和 O 的大小, $\text{BBS}(\cdot)$ 是最终计算得到的搜索区域的目标最佳伙伴似然概率图,其似然概率图中得分的高低表示目标在此位置的似然概率大小。

3.3 最佳伙伴相似性与目标重检测

本文利用BBS方法构建目标重检测模块,当目标位置和尺度被确定时,判断跟踪目标是否可靠,是否发生严重遮挡或者跟踪失败,其目标置信度判别为:

$$P_t^{\text{similarly}} = \exp(-\eta \times \|x - x_{t-1}\|^2), \quad (11)$$

$$P_t^{\text{PSR}} = \frac{\max(f(x)) - \mu_\varphi}{\sigma_\varphi(f(x))}, \quad (12)$$

其中: $P_t^{\text{similarly}}$ 和 P_t^{PSR} 分别表示目标外观模型置信度和响应置信度, η 表示惩罚权重, x 为当前帧采样目标外观模型, x_{t-1} 为第 $t-1$ 帧目标外观模型, μ_φ 和 σ_φ 分别表示指定范围内的滤波响应均值与标准差, φ 表示指定的目标峰值响应区域窗口大小, $\|\cdot\|$ 表示计算目标外观模型之间的欧式距离:

$$\gamma = \begin{cases} 1, & P_t^{\text{similarly}} > P_{\text{threshold}}^{\text{similarly}} \ \& \ P_t^{\text{PSR}} > P_{\text{threshold}}^{\text{PSR}} \\ 0, & \text{otherwise} \end{cases}, \quad (13)$$

其中: γ 表示跟踪目标是否可靠,当为1时表示跟踪目标可靠,为0表示跟踪目标发生遮挡或跟踪

失败, $P_{\text{threshold}}^{\text{similarly}}$ 和 $P_{\text{threshold}}^{\text{PSR}}$ 分别表示跟踪目标的外观模型置信度阈值和响应置信度阈值。

当目标跟踪失败,启动重检测模块,首先计算目标的最佳伙伴似然概率图,通过似然概率的大小确定得分最大的 k 个候选目标 $\{z_1^i, z_2^i, \dots, z_k^i\}$,结合候选目标最佳伙伴似然概率得分、空间距离权重得分和滤波响应图,对候选目标进行重排名,获取得分最大的候选目标:

$$z_i^i = \arg \max_{z_i^i} \{f_d(z_i^i) f_{\text{bbs}}(z_i^i) f_{\text{h}_{\text{sc}}}(z_i^i), \dots, f_d(z_i^k) f_{\text{bbs}}(z_i^k) f_{\text{h}_{\text{sc}}}(z_i^k)\}, \quad (14)$$

其中: $f_d(z)$, $f_{\text{bbs}}(z)$ 和 $f_{\text{h}_{\text{sc}}}(z)$ 分别表示目标的空间距离权重函数、最佳伙伴似然函数和滤波响应函数,其分别定义为: $f_d(z_i^i) = \exp(-(d(z_i^i, z_{t-1}^{\text{object}})^2 / S_{\text{area}}))$,这里 $d(z_i^i, z_{t-1}^{\text{object}})$ 表示当前候选目标 z_i^i 与上一帧跟踪目标 z_{t-1}^{object} 之间的欧式距离, S_{area} 表示输入图像搜索区域大小, $f_{\text{bbs}}(z_i^i) = \text{BBS}(T, C(z_i^i))$, $f_{\text{h}_{\text{sc}}}(z_i^i) = f(z_i^i)$,最终获取的最佳候选目标位置即为估计目标位置。

利用最佳伙伴相似性进行目标重检测流程如图1所示。

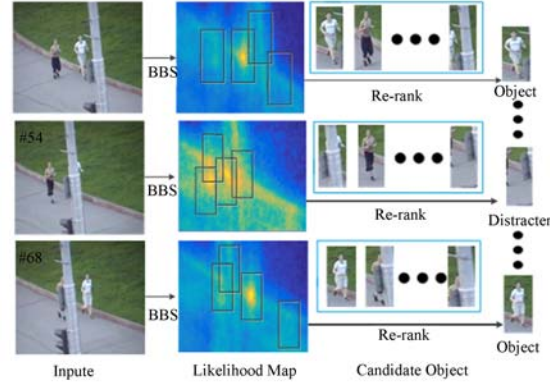


图1 利用最佳伙伴相似性进行目标重检测流程

Fig.1 Object re-detection processing with Best-Buddies Similarly

3.4 自适应模板更新

为解决遮挡情况下的跟踪漂移问题,本文采用自适应模板更新策略,通过引入自适应权重因子,降低遮挡情况下的模板漂移,提高模板的可靠性。

$$\rho_t = \begin{cases} \exp[-(0.02(R_t - R_{\text{max}}))^2] & \text{if } \gamma = 1 \\ 0 & \text{if } \gamma = 0 \end{cases}, \quad (15)$$

其中: $R_t = P_t^{\text{PSR}}$,表示当前目标响应置信度, R_{max}

表示初始帧的目标响应置信度, ρ_t 表示自适应权重因子, 用于决定模板更新的速率。

目标定位模块模板更新定义为:

$$A_t^l = (1 - \rho_t \beta) A_{t-1}^l + \rho_t \beta ((D_t^l)^T D_t^l + W^T W), \quad (16)$$

$$\tilde{b}_t^l = (1 - \rho_t \beta) \tilde{b}_{t-1}^l + \rho_t \beta (D_t^l)^T \tilde{y}_t^l, \quad (17)$$

其中: β 表示目标模板学习速率, A_t^l 和 \tilde{b}_t^l 分别表示第 t 帧位置分类器 h^l 的分母与分子参数, A_{t-1}^l 和 \tilde{b}_{t-1}^l 分别表示第 $t-1$ 帧位置分类器 h^l 的分母与分子参数, \tilde{y}_t^l 表示位置估计模块中输出期望响应。

而由于本文利用的尺度估计模块, 并未利用背景信息, 因此不需要增加稀疏空间正则化权重, 其对应的模板更新新定义为:

$$A_t^s = (1 - \rho_t \beta) A_t^s + \rho_t \beta ((D_t^s)^T D_t^s + \lambda), \quad (18)$$

$$\tilde{b}_t^s = (1 - \rho_t \beta) \tilde{b}_t^s + \rho_t \beta (D_t^s)^T \tilde{y}_t^s, \quad (19)$$

其中: A_t^s 和 \tilde{b}_t^s 分别表示第 t 帧尺度分类器 h^s 的分母与分子参数, A_{t-1}^s 和 \tilde{b}_{t-1}^s 分别表示第 $t-1$ 帧尺度分类器 h^s 的分母与分子参数, \tilde{y}_t^s 表示尺度估计模块中输出期望响应。

在 BBS 方法中(重检测模块中), 目标匹配模板更新方法为:

$$\hat{T}_t = (1 - \rho_t \beta) \hat{T}_{t-1} + \rho_t \beta T_t, \quad (20)$$

其中: T_t 表示当前跟踪目标的外观模型, \hat{T}_{t-1} 表示上一帧目标的匹配模板, \hat{T}_t 表示当前帧目标的匹配模板。

4 实验结果与分析

本文所有实验均在 MATLAB2014b, Intel i5-4200 CPU, 主频 2.50 GHz, 内存 4 G 的电脑上进行。本文算法的实验参数为: 本文算法利用与 SRDCF 算法相同的 HOG 特征, 正则化因子设置为 0.0001, 目标定位模块的搜索窗口为 4^2 倍目标大小, 惩罚因子 η 为 0.05, φ 表示 11×11 的窗口, 置信度阈值 $P_{\text{threshold}}^{\text{similarity}}$ 和 $P_{\text{threshold}}^{\text{PSR}}$ 分别设置为 0.5 和 16, 尺度池为 $S = \{1.03, 1.02, 1.01, 1, 0.99, 0.98, 0.97\}$, 候选目标个数 k 大小设置为 10, 模板学习速率 β 设置为 0.025, 对于本文所有实验算法参数均固定不变。

同时利用 OTB-2013 数据集测试本文算法, 其共包括 50 组视频序列和 11 种视频属性, 经测试本文算法在 50 组视频上的平均速度为 3.98F/s。为了更好的评估本文算法性能, 本文

算法与 34 种流行算法进行比较, 其中 29 种算法由 OTB-2013 数据集提供, 另 5 种算法分别为: SRDCF^[12], KCF^[8], DSST^[11], SAMF^[19] 和 TG-PR^[18]。

4.1 定量分析

为了更好地分析本文算法的性能, 本文采用 OTB-2013 数据集中的方法测试本文算法的性能, 用成功率图和精确度图评价本文算法的性能。成功率指跟踪目标重叠率大于指定阈值的帧数占总帧数的百分比, 其中重叠率指跟踪目标框与真实标定目标框的重叠面积与两者面积和的比值。为保证评价标准的公平性, 在成功率图中利用成功率曲线下的面积 (Area under curve, AUC) 大小对算法进行排名。精确度指表示跟踪目标中心位置误差 (Center location error, CLE) 小于指定阈值的帧数占总帧数的百分比, 这里中心位置误差指跟踪目标中心位置与标定真实目标中心位置之间的欧式距离。在精确度图中利用精确度得分对各个算法进行排名, 用指定阈值 (中心位置误差) 下的精确度表示算法的精确度得分, 这里中心位置误差阈值设置为 20 pixel, 对应于精确度图中横坐标的 20。

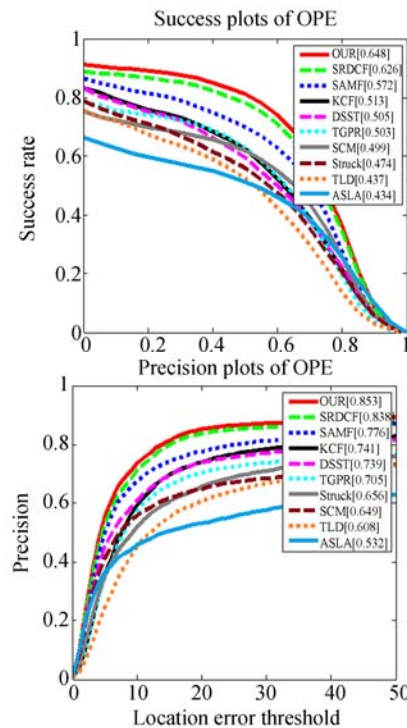


图 2 排名靠前的 10 种算法 OPE 成功率图和精确度图
Fig. 2 Success plots and precision plots of OPE for top ten trackers

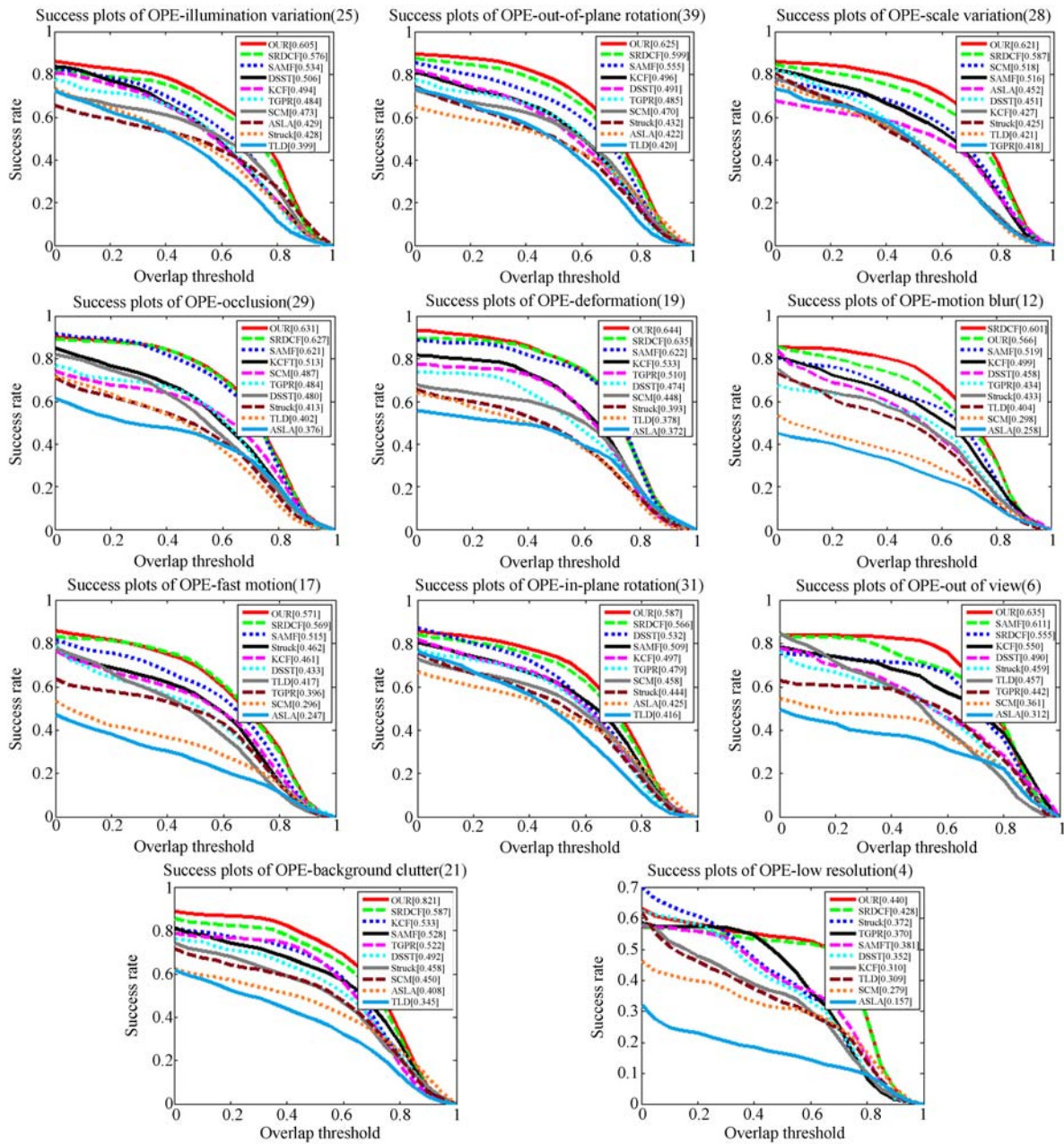


图 3 排名靠前的 10 种算法在 11 种属性上的成功率图(标题上数字表示此种属性的数据集数量)

Fig. 3 Success plots of videos with eleven different attributes for top ten trackers. (The number in the title indicates the number of sequences)

图 2 是 10 种排名靠前的算法在 50 组视频上总体精确度图和成功率图,从图中看出本文算法的精确度得分为 0.853,相对于 SRDCF 的 0.838 提高了 1.79%,本文算法在 34 种算中精确度排名第一。同时从图 2 得出,本文算法的成功率得分为 0.648,相对于 SRDCF 提高 3.51%,这说明本文算法引入的重检测模块能很好的解决目标遮挡、形变情况下的跟踪失败问题,本文算法具有更好的鲁棒性。

为了更好地分析本文算法在不同跟踪情况下的可靠性,记录 10 种算法在 11 种不同属性视频上精确度图和成功率图。图 3 和图 4 反映算法在不同属性上的跟踪得分情况,从图中看出本文算法在 11 种属性中的 10 种属性上成功率排名第一,且成功率在 11 种属性上均排名在前两位,这说明本文算法能很好地解决目标遮挡、形变、尺度变化和其他复杂场景下的目标跟踪失败问题,提

高了跟踪算法的可靠性。而在精确度上,本文算法在 6 种属性的视频上精确度得分排名第一,在

11 种属性视频上精确度得分均排名前三位,这说明本文提出的算法能实现精确的目标跟踪任务。

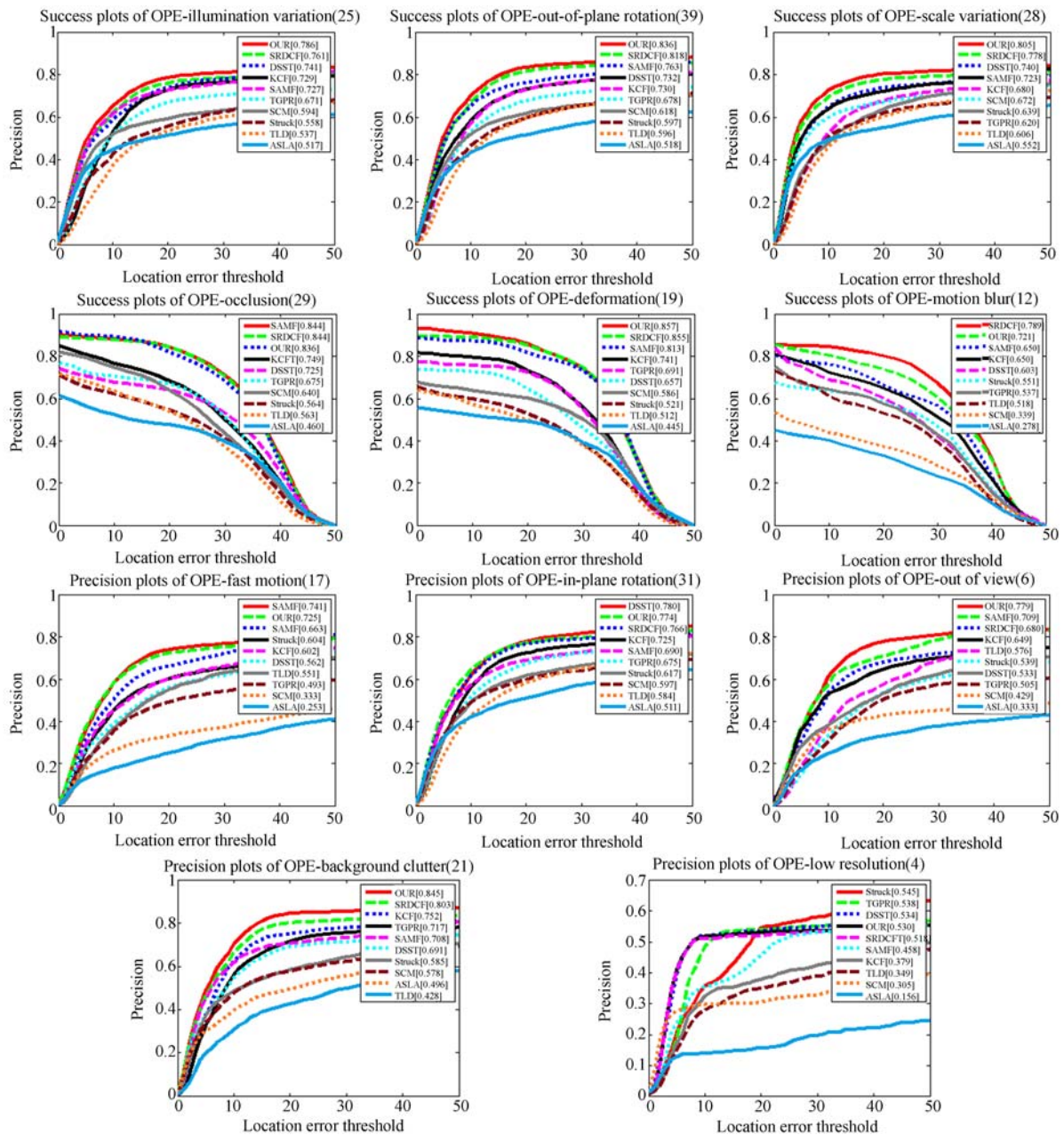


图 4 排名靠前的 10 种算法在 11 种属性上的精确度图(标题上数字表示此种属性的数据集数量)

Fig. 4 Precision plots of videos with eleven different attributes for top ten trackers. (The number in the title indicates the number of sequences)

4.2 定性分析

为了更好地分析本文算法的性能,记录本文算法在不同属性视频序列上的跟踪结果。图 5 和图 6 是本文算法与其他 9 种算法在目标发生遮挡、形变、尺度变化、快速运动、光照变化、背景嘈杂和旋转情况下的跟踪情况。

(1) 遮挡情况下性能分析。

在 lemming, jogging-1, faceoccl, walking2 序列中跟踪目标发生部分或者完全遮挡。其中在 lemming 和 jogging-1 序列中目标发生完全遮挡,在此情况下容易出现跟踪失败问题。例如,在 lemming 序列的第 341 帧中,目标被遮挡物完全

遮挡,到第 395 帧时只有本文算法、SAMF、TLD 和 Struck[20]算法能正确跟踪指定目标,SRDCF 因为不具有重定位模块,在目标再次出现时,SRDCF 算法已经无法重新定位目标位置,出现跟踪失败问题;在 faceoccl 和 walking2 视频序列中,目标发生部分遮挡,此时由于遮挡造成模板误

更新,容易造成跟踪漂移问题,甚至造成跟踪失败,影响跟踪的精确性。例如,在 faceoccl 序列中颅面被书部分遮挡,在第 520 帧 ASLA 算法已经出现重大跟踪漂移现象,TLD 算法出现了错误的尺度更新,而本文算法由于采用自适应模板更新策略,很好地解决了模板漂移现象。

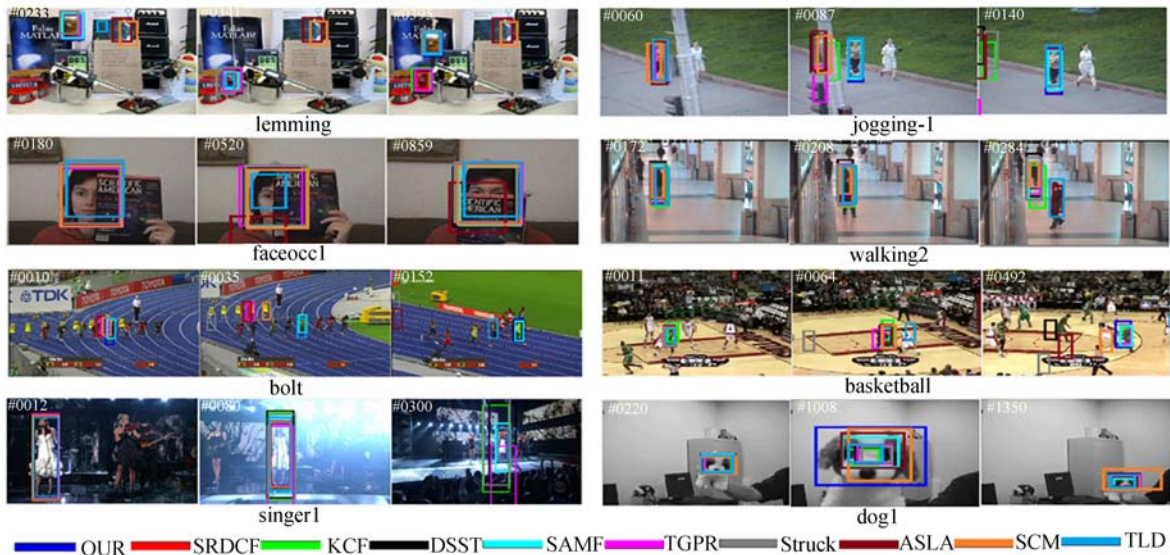


图 5 10 种跟踪算法在遮挡、形变和尺度变化情况下的跟踪结果

Fig. 5 The tracking results of 10 algorithms in the occlusion, deformation and scale variation case

(2) 目标形变和尺度变化情况下性能分析。

在 bolt 和 basketball 序列中目标发生剧烈形变,此情况下目标外观模型发生变化,容易出现跟踪失败。例如,在 bolt 序列的第 10,36,152 帧中运动员发生剧烈形变,在第 152 帧只有本文算法、DSST、SAMF 和 KCF 算法能正确跟踪目标,SRDCF 算法出现跟踪失败,由于 SRDCF 跟踪引入过多的背景信息(搜索窗口为目标 4^2 大)使得分类器对目标形变更加敏感,容易跟踪失败,而本文算法因为带重检测模块,因此能解决目标形变情况下的跟踪任务;在 singer1 和 dog1 序列中目标发生剧烈尺度变化,例如,从 dog1 序列中看出,本文提出的尺度估计策略能解决剧烈尺度变化问题,SRDCF、DSST 和 SAMF 算法虽然具有尺度估计策略(利用目标信息与背景信息训练尺度分类器),但是其尺度估计策略不够鲁棒,容易产生尺度漂移。

(3) 快速运动和光照变化情况下性能分析。

在 jumping 和 deer 序列中,跟踪目标主要面临的挑战是目标快速运动,此时由于受跟踪目标

搜索区域大小的影响,往往容易造成跟踪失败。例如,在 deer 序列的第 29 帧中,目标发生快速运动,本文算法、SRDCF、Struck 和 TGPR 算法能准确跟踪指定目标,本文算法与 SRDCF 都利用了更大的目标搜索区域(目标大小的 4^2 倍搜索区域),以此解决快速运动情况下跟踪失败问题。在 trellis 和 shaking 序列中,目标发生剧烈光照变化,此时目标外观模型也发生剧烈变化,跟踪器很难区分背景与目标,造成跟踪失败,例如,在 shaking 序列的第 273 帧,受目标光照变化的影响,只有本文算法和 DSST 算法能正确跟踪目标,而 SRDCF 跟踪算法完全跟踪失败,这说明,本文算法的重检测模块是有效的,能显著提高光照变化情况下的跟踪鲁棒性。

(4) 背景嘈杂和旋转情况下性能分析。

在 soccer 和 freeman4 序列中,跟踪目标处于嘈杂的环境中,易受背景干扰,此时分类器很难从嘈杂的背景中区分出目标。例如,在 soccer 序列中,运动员处于非常复杂的环境中,这些运动员颜

色非常相近,跟踪目标易受到相似目标的干扰,从而造成跟踪失败。在 motorRolling 和 football 序列中,目标发生平面内旋转和平面外旋转,此时目标外观模型也发生剧烈变化,而传统的相关滤波算法因为不具有旋转不变性,因此很难解决旋转

情况下跟踪失败问题,例如,在 motorRolling 序列中,目标发生剧烈的旋转,从第 83 帧看出,传统的相关滤波算法 SRDCF、KCF、DSST 和 SAMF 都跟踪失败,而本文算法通过引入重检测模块,能很好地解决旋转情况下的跟踪问题。

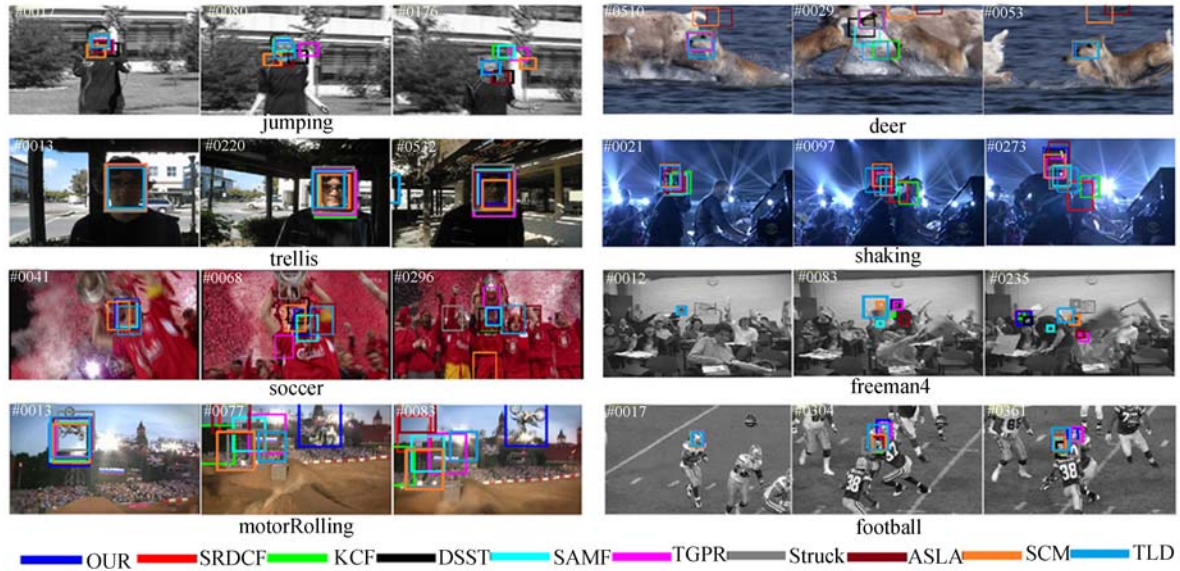


图 6 10 种跟踪算法在目标发生快速运动、光照变化、背景嘈杂和旋转情况下的跟踪结果

Fig. 6 Tracking results of 10 algorithms in the fast motion, illumination variation and rotation case

4.3 子系统定量分析

为更好分析各个模块对算法整体贡献度,分别记录 OUR(SRD+Scale+BBS)、OUR(SRD+Scale)、OUR(SRD+NoScale)和 SRDCF 算法在 OTB-50 数据集上的跟踪重叠率精度(Scale 指使用本文算法的尺度估计模块,BBS 指使用本文算法的重检测模块,NoScale 指不使用尺度估计策略,SRD 表示使用空间正则化权重,本文算法默认均利用自适应模板更新策略)。跟踪重叠率精度(Overlap Precision, OP)指跟踪目标重叠率大于 0.5 时的帧数占总帧数的百分比,其反映了算法的跟踪可靠性与鲁棒性。

表 1 不同情况下的重叠率精度对比

Tab. 1 Comparison of overlap precision in different situations

算法属性	跟踪重叠率(OP)
OUR(SRD+Scale+BBS)	0.810
OUR(SRD+Scale)	0.787
OUR(SRD+NoScale)	0.743
SRDCF	0.781

从表 1 中看出本文算法在使用尺度估计模块和重检测模块时的重叠率精度为 0.810,明显优于 SRDCF 跟踪算法的 0.781,同时从表 1 中看出本文提出的尺度估计策略和重检测策略对算法鲁棒性提升具有明显增益。

5 结 论

本文提出利用最佳伙伴相似性的改进 SRDCF 目标跟踪算法,通过双层搜索策略解决目标位置估计和尺度估计,在训练目标定位分类器时,引入更多的背景信息,提高分类器的分类能力,在训练目标尺度估计模块时,仅利用目标信息,提高目标尺度估计的准确性。然后利用最佳伙伴相似性的构造目标重检测模块,通过融合空间权重、相关滤波得分和最佳伙伴相似性得分来估计候选目标位置,实现目标重定位。最后,利用自适应模板更新策略,解决遮挡情况下目标模板和分类器模板更新问题,提高了跟踪精确度。利用 OTB-2013 标准数据集测试本文算法,通过实验得到本文算法的精确度得分和成功率得分分别为 0.853 和

0.648,其在35种算法中均排名第一。通过实验说明相对于SRDCF算法,本文算法能更好的解

决目标遮挡、尺度变化和形变情况下的目标跟踪任务,具有更高的精确性和可靠性。

参考文献:

- [1] 管皓,薛向阳,安志勇.在线单目标视频跟踪算法综述[J].小型微型计算机系统,2017,38(1):147-153.
GUAN H, XUE X R, AN ZH Y. Survey of video object tracking[J]. *Journal of Chinese Computer Systems*, 2017, 38(1): 147-153. (in Chinese)
- [2] KALAL Z, MIKOLAJCZYK K, MATAS J. Tracking-learning-detection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(7): 1409-1422.
- [3] 孙锐,张东东,高隽.基于分层极限学习机和局部稀疏模型的视觉跟踪算法[J].模式识别与人工智能,2017,30(4):302-313.
SUN R, ZHANG D D, GAO J. Visual tracking via hierarchical extreme learning machine and local sparse model[J]. *Pattern Recognition and Artificial Intelligence*, 2017, 30(4): 302-313. (in Chinese)
- [4] JIA X, LU HCH, YANG M H. Visual tracking via adaptive structural local sparse appearance model [C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2012: 1822-1829.
- [5] ZHONG W, LU HCH, YANG M H. Robust object tracking via sparsity-based collaborative model [C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2012: 1838-1845.
- [6] 黄丹丹,孙怡.基于稀疏子空间选择的在线目标跟踪[J].自动化学报,2016,42(7):1077-1089.
HAUNG D D, SUN Y. Online object tracking via sparse subspace selection[J]. *Laser & Acta Automatica Sinica*, 2016, 42(7): 1077-1089. (in Chinese)
- [7] 刘文琢,袁广林,薛模根.基于两阶段稀疏表示的稳健快速视觉跟踪[J].光学学报,2016,36(12):183-189.
LIU W ZH, YUAN G L, XUE M G. Robust fast visual tracking based on two-stage sparse representation[J]. *Acta Optica Sinica*, 2016, 36(12): 183-189. (in Chinese)
- [8] HENRIQUES J F, CASEIRO R, MARTINS P, *et al.*. High-speed tracking with kernelized correlation filter[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(3): 583-596.
- [9] 张雷,王延杰,孙宏海,等.采用核相关滤波器的自适应尺度目标跟踪[J].光学精密工程,2016,24(2):448-459.
ZHANG L, WANG Y J, SUN H M, *et al.*. Adaptive scale object tracking with kernelized correlation filters[J]. *Opt. Precision Eng.*, 2016, 24(2): 448-459. (in Chinese)
- [10] 杨德东,蔡玉柱,毛宁,等.采用核相关滤波器的长期目标跟踪[J].光学精密工程,2016,24(8):2037-2049.
YANG D D, CAI Y ZH, MAO N, *et al.*. Long-term object tracking based on kernelized correlation filters[J]. *Opt. Precision Eng.*, 2016, 24(8): 2037-2049. (in Chinese)
- [11] DANELLJAN M, H? GER G, KHAN F, *et al.*. Accurate scale estimation for robust visual tracking [C]. *British Machine Vision Conference (BMVC)*, BMVC, 2014: 1-5.
- [12] DANELLJAN M, H? GER G, KHAN F S, *et al.*. Learning spatially regularized correlation filters for visual tracking [C]. *IEEE International Conference on Computer Vision (ICCV)*, IEEE, 2015: 4310-4318.
- [13] 毛宁,杨德东,杨福才,等.基于分层卷积特征的自适应目标跟踪[J].激光与光电子学进展,2016,53(12):201-212.
MAO N, YANG D D, YANG F C, *et al.*. Adaptive object tracking based on hierarchical convolution features [J]. *Laser & Optoelectronics Progress*, 2016, 53(12): 201-212. (in Chinese)
- [14] 王暉,王春平,李军,等.特征融合和模型自适应更新相结合的相关滤波目标跟踪[J].光学精密工程,2016,24(8):2059-2066.
WANG W, WANG CH P, LI J, *et al.*. Correlation filter tracking based on feature fusing and model adaptive updating[J]. *Opt. Precision Eng.*, 2016, 24(8): 2059-2066. (in Chinese)
- [15] 潘振福,朱永利.多尺度估计的核相关滤波器目标跟踪方法[J].激光与光电子学进展,2016,53(10):199-205.
PAN ZH G, ZHU Y L. Kernelized correlation filters object tracking method with multi-scale esti-

- mation [J]. *Laser & Optoelectronics Progress*, 2016, 53(10): 199-205. (in Chinese)
- [16] DEKEL T, ORON S, RUBINSTEIN M, *et al.*. Best-Buddies Similarity for robust template matching [C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2015: 2021-2029.
- [17] WU Y, LIM J, YANG M H. Online object tracking: A benchmark [C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2013: 2411-2418.
- [18] GAO J, LING H B, HU W M, *et al.*. Transfer learning based visual tracking with gaussian processes regression [C]. *European Conference on Computer Vision (ECCV)*, Springer, 2014: 188-203.
- [19] LI Y, ZHU J K. A scale adaptive kernel correlation filter tracker with feature integration [C]. *European Conference on Computer Vision (ECCV)*, Springer, 2014: 254-265.
- [20] HARE S, SAFFARI A, TORR P H S. Structured output tracking with kernels [C]. *IEEE International Conference on Computer Vision (ICCV)*, IEEE, 2011: 263-270.

作者简介:



杨德东(1977—),男,辽宁阜新人,副教授,硕士生导师,2000年、2003年于大连铁道学院分别获得学士学位、硕士学位,2007年于东北大学获得博士学位,主要从事智能感知与控制、目标检测与跟踪等方面的研究。E-mail: ydd12677@163.com



毛宁(1992—),男,河南周口人,硕士研究生,2015年于河南城建学院获得学士学位,主要从事数字图像处理、目标跟踪等方面的研究。E-mail: maon316@163.com