

采用多相关滤波策略的鲁棒长时自适应目标跟踪



谭建豪 殷旺 刘力铭 王耀南

湖南大学电气与信息工程学院 长沙 410082

机器人视觉感知与控制技术国家工程实验室 长沙 410082

(tanjianhao@hnu.edu.com)

摘要 传统相关滤波方法在目标运动模糊和光照变化上取得了一定的鲁棒效果,但当目标存在形变、颜色变化、重度遮挡等干扰因素时难以实现跟踪,鲁棒性差,且当目标丢失后不能再恢复,无法实现长时间跟踪。因此,文中提出了一种鲁棒长时自适应目标跟踪算法。首先,提出了一种特征互补策略,将方向梯度直方图和全局颜色直方图的特征响应线性加权,学习对颜色变化和形变都具有鲁棒性的相关滤波模型,用以估计目标位移;然后,仅提取目标前景 HOG 特征,学习一个判别滤波器,用以保持对目标外观的长期记忆,使用该长期滤波器的输出响应来判别是否出现遮挡或跟踪失败,采用在线 SVM 分类器对丢失目标进行再检测,从而能够跟踪已丢失目标,以实现长期跟踪;其次,学习了以目标位置为中心的特征金字塔模型以预测尺度变化,防止目标框漂移;最后,在 OTB 目标跟踪基准数据集上对算法进行实验,并与目前较为流行的目标跟踪算法进行对比,进一步验证了所提算法的鲁棒性、准确性和优越性。

关键词: 长时目标跟踪;颜色直方图;相关滤波;SVM 再检测器;尺度自适应

中图分类号 TP391.41

Robust Long-term Adaptive Object Tracking Based on Multi-correlation Filtering Strategy

TAN Jian-hao, YIN Wang, LIU Li-ming and WANG Yao-nan

College of Electrical and Information Engineering, Hunan University, Changsha 410082, China

National Engineering Laboratory for Robot Visual Perception and Control Technology, Changsha 410082, China

Abstract The traditional correlation filtering methods have recently achieved excellent performance and shown great robustness to exhibiting motion blur and illumination changes. However, it is difficult to achieve tracking when the object has interference factors such as deformation, color change, and heavy occlusion. It shows poor robustness when the object is lost and cannot be recovered to achieve long-term tracking. Therefore, this paper proposes a robust long-term object tracking algorithm. First, a feature complementation strategy is proposed, which linearly weights the feature responses of the directional gradient histogram and the global color histogram, and learns a correlation filtering model that is robust to color changes and deformations to estimate the target displacement. Then, the object features are taken to learn a discriminant correlation filter to maintain long-term memory of object appearance. We use the output responses of this model to determine if tracking failure occurs. We use the online SVM classifier to re-detect the lost object and retrack the lost target which can effectively recover the tracking target from failure to achieve long-term tracking. In addition, this paper learns a correlation filter over a feature pyramid centered at the estimated object position for predicting scale changes and further enhance robustness and accuracy. Finally, this paper compares the proposed algorithm with the state-of-the-art performance tracking algorithms on the online object tracking benchmark. The result shows that the proposed algorithm performs great robustness and accuracy.

Keywords Long-term object tracking, Color histogram, Correlation filter, SVM re-detector, Scale adaptation

目标跟踪是计算机视觉领域最基本、最有挑战性的热点研究问题之一。基于视觉的运动目标跟踪已被广泛应用于监控系统、无人机视觉系统、行为理解、人机交互以及无人驾驶车辆导航等领域^[1]。近年来,国内外的专家学者对其进行了大量深入的研究,提出了大量的算法^[2-3],并且吸引了大量优

秀学者参加全球视觉跟踪算法挑战赛^[4-7]。目前目标视觉跟踪方法主要分为两大类^[8-9]:生成(generative)模型类方法和判别(discriminative)模型类方法。

近年来,基于相关滤波和深度学习的方法在目标跟踪上得到了广泛的研究,相关滤波表现出了非常优秀的跟踪性能。

收稿日期:2019-10-04 返修日期:2020-03-07 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。

基金项目:国家自然科学基金(61433016)

This work was supported by the National Natural Science Foundation of China (61433016).

通信作者:殷旺(yinwang@hnu.edu.cn)

该方法通过学习一个滤波模板,将下一帧的图像和滤波模板进行卷积运算,根据输出响应来预测目标位置,在实际计算中使用FFT将图像的卷积运算转换为频域中的点乘,大大降低了计算复杂度。Bolme等于2010年提出了MOSSE(Visual object tracking using adaptive correlation filters)算法^[10],第一次使用相关滤波来解决视觉目标跟踪,其使用灰度特征平均帧率可达几百,表现出了强大的速度优越性,在光照和部分形变变化时具有一定的鲁棒性,但是其没有考虑目标的尺度变化和长时间跟踪。Heriques等提出了CSK(Exploiting the circulant structure of tracking-by-detection with kernels)算法^[11],解决了跟踪过程中训练样本太少的问题,并且加入了小技巧来增强算法的可靠性。在此基础上,Henriques对CSK的单通道特征做出了改进,使用多个通道的HOG特征融合,于2014年提出了KCF跟踪算法^[12],该算法取得了较好的跟踪效果,但是同样没有考虑跟踪过程中尺度变化引起的目标框漂移的问题。针对尺度变化,文献[13]利用目标对象的颜色属性,将多通道特征映射到高斯核空间来学习自适应相关滤波器。Danelljan等提出的DSST(Accurate scale estimation for robust visual tracking)算法^[14]在CSK框架中加入尺度滤波器,通过使用HOG特征来学习自适应多尺度相关滤波器,以应对跟踪过程中的尺度变化。2016年Bertinetto等提出Staple算法^[15],引入颜色统计特征,在较高准确度的前提下,最终达到80fps的优越性能,在VOT16中排名第5^[16]。然而,以上方法无法解决有关在线模型更新的关键问题,因此这些相关滤波器容易出现漂移,并且在处理长期遮挡或者视野外的目标时效果较差。Ma等提出的LCT(Long-term Correlation Tracking)长时跟踪算法^[17]加入了检测模块,在尺度相关滤波和平移相关滤波上设置目标置信度,实现了较好的长时跟踪,但是其在特征上未结合颜色特征,对于颜色变化的跟踪不具备鲁棒性。2017年,由Danelljan提出的ECO算法^[18]结合了传统的HOG特征、CN特征以及CNN特征,特征全面且训练样本具有多样性,减少了冗余,但该方法同样只适用于短时间跟踪。

文献[15,18]虽然结合了相关颜色特征,在运动模糊、光照和颜色变化上有一定的效果,但只是在特征融合后训练一个相关滤波器,鲁棒性较差,且并未考虑遮挡或目标丢失的长时间跟踪;文献[17,19]考虑了目标遮挡或丢失后的再检测以达到长时跟踪的目的,但未考虑颜色特征,对于颜色变化等干扰不能实现强鲁棒性跟踪。因此,本文考虑了目标跟踪过程中存在运动模糊、光照变化、形变、颜色变化以及严重遮挡或目标丢失的干扰,采用一种特征响应互补策略来提取目标前景和背景的FHOG特征,学习一个相关滤波器 CF_a ,用于处理位移过程中的运动模糊和光照颜色变化;提取目标和目标周围背景的全局颜色直方图(GCH)特征,通过积分图计算得到其特征响应,为了方便表达,本文将此特征响应训练过程称为相关滤波,用于处理位移过程中的形变,将两个相关滤波的输出响应进行线性加权,得到对颜色变化和形变等都具有鲁棒性的位移估计模型;为了应对目标遮挡或丢失的影响,仅考虑目标HOG特征,不考虑目标周围的背景信息,学习一个判

别滤波器 CF_b ,用于判断是否出现遮挡或丢失以便激活再检测模块,对丢失目标实现恢复再跟踪,该滤波器还用于判断是否更新训练样本及更新模型;本文采用在线SVM分类器对丢失目标进行再检测,能够有效地将跟踪目标从失败中恢复,以实现长期跟踪;另外,本文还将判别滤波器 CF_a 用于尺度估计,以防止目标框漂移,进一步增强鲁棒性和准确性。因此,本文共训练了3个相关滤波器和1个在线分类器。

1 基于相关滤波特征响应加权的目標搜索与位移估计

基于HOG特征的相关滤波跟踪器在运动模糊和光照变化上表现出了稳健性,但是它们对目标变形非常敏感,因此,本文提出一种特征响应互补策略用以训练一种对颜色和形变等都具有鲁棒性的模型。

假设目标在第 t 帧的图像为 x_t ,根据TLD的流程可以定义第 t 帧目标位置在对应图像 x_t 中的最大响应矩形框。

$$Q_t = \arg \max_{Q \in C_t} f(T(x_t, Q); \omega_{t-1}) \quad (1)$$

其中, C 表示所有预测框的集合; T 是一种图像变换,表示在图像 x_t 中提取矩形框 Q 的特征,它使得 f 根据模型参数 ω 将响应分配给图像 x 中的每个矩形框 Q ; f 表示求响应的过程,最大响应所对应的矩形框即可视为跟踪目标的位置。定义损失函数 $L(\omega; x_t)$,模型参数设为 ω ,即:

$$\omega_t = \arg \min_{\omega \in S} (L(\omega, x_t) + \lambda g(\omega)) \quad (2)$$

其中, S 表示模型参数的数据空间,通过加入有相对影响因子 λ 的正则项 $g(\omega)$ 来防止出现过度拟合。

考虑到算法的实时性,本文采用一种特征响应线性加权的互补策略,它是梯度特征响应和颜色直观图特征响应的线性组合,模型如下:

$$f(x) = \mu_{CF_a} f_{CF_a}(x) + \mu_{CF_b} f_{CF_b}(x) \quad (3)$$

其中, μ_{CF_a} 和 μ_{CF_b} 为线性加权参数, $f_{CF_a}(x)$ 为梯度特征的响应,通过提取 $t-1$ 帧图像中目标前景与背景的梯度特征(HOG),根据相关滤波的规则,学习得到一个相关滤波模型 CF_a ,使用下一帧图像特征可得相关滤波模型响应。

$$f_{CF_a}(x, d) = \sum_{o \in K} (d[o])^T \psi_x[o]$$

其中, T 表示二维多通道特征图, o 表示图像块中的某个像素, K 表示图像的像素集合, ψ_x 表示求 x 图像中的HOG特征函数, $d[o]$ 表示HOG梯度模板。 f_{CF_b} 为颜色直观图特征的积分图响应 $f_{CF_b}(x, \beta) = \beta^T (\frac{1}{|H|} \sum_{o \in H} \gamma_x[o])$ 。 $\frac{1}{|H|} \sum_{o \in H} \gamma_x[o]$ 表示积分图的计算,可解释为标量值得分图的平均值; β 表示某一点属于前景的概率,对于三通道图像,它实际是一个向量。

(d, β) 为整个模型的更新参数,因此定义 $\omega = (d, \beta)$,假设损失为训练单个图像损失的线性组合,则每个图像的损失和为:

$$L(\omega, x_t) = \sum_{t=1}^T kl(x_t, Q, \omega)$$

其中,每张图像的损失定义为:

$$l(x, Q, \omega) = d(Q, \arg \max_{Q \in C_t} f(T(x, q), \omega))$$

其中, $d(Q, q)$ 表示当 Q 为正确目标框时选择目标框 q 的损失值^[17-18],该损失是一个非凸函数,若优化该问题,则计算复杂且限制了特征的个数和训练样本数,因此为了保持相关滤波

器的速度和效率,同时又能保证获取直方图得分相对应的特征响应,通过两个独立的岭回归来求解该问题^[20-21],即把 d 和 β 分开求解,模型如下:

$$\begin{cases} d_t = \arg \min_d \{L(d, x_t) + 0.5\lambda \|d\|^2\} \\ \beta_t = \arg \min_\beta \{L(\beta, x_t) + 0.5\lambda \|\beta\|^2\} \end{cases} \quad (4)$$

参考相关滤波的最小二乘方式求解,则存在矩阵 A_t 和向量 b_t ,使得:

$$\begin{cases} L(\omega; x_t) + \lambda \|\omega\|^2 = \frac{1}{2} \omega(A_t + \lambda I) \omega + b_t \omega \\ \omega_t = (A_t + \lambda I)^{-1} b_t \end{cases}$$

其中, λ 表示正则化参数, I 表示单位向量。这是一种递归的求解方式,可得:

$$\begin{cases} A_t = (1-\alpha)A_{t-1} + \alpha A_t' \\ b_t = (1-\alpha)b_{t-1} + \alpha b_t' \end{cases}$$

其中, A_t 和 b_t 表示从第 1 帧到第 t 帧的参数估计,而 A_t' 和 b_t' 表示仅从第 t 帧获取的参数估计, α 表示参数的更新速率。

每张图像的具体损失函数的形式为: $l(x, Q, d) = \|\sum d^k * \psi^k - y\|^2$, d^k 表示多通道图像 d 的第 k 个通道, $*$ 表示相关运算, ψ 表示特征函数, y 表示实际标签响应值。目标函数为 $l(x, Q, d) + \lambda \|d\|^2$, 对其进行最小二乘求解得到:

$$d = \frac{\hat{\delta}}{\hat{\xi} + \lambda}$$

对于 HOG 特征,在第 $t-1$ 帧时,提取目标图像块的

HOG 特征用于更新模型 \hat{d}_t 中的分子 $\hat{\delta}_t$ 和分母 $\hat{\xi}_t$,假设学习率为 α ,则模型中的参数更新过程为:

$$\begin{cases} \hat{\delta}_t = (1-\alpha_{cfa})\hat{\delta}_{t-1} + \alpha_{cfa}\hat{\delta}_t \\ \hat{\xi}_t = (1-\alpha_{cfa})\hat{\xi}_{t-1} + \alpha_{cfa}\hat{\xi}_t \end{cases} \quad (5)$$

在第 t 帧搜索目标位置时,从第 $t-1$ 帧中目标位置的周围提取待检测 HOG 特征,并与模板 $\hat{d}[o]$ 进行相关滤波运算,获得梯度特征的响应 f_{cfa} 。

对于颜色特征,其得分图响应模型为:

$$\beta_t = \rho(M) / (\rho(M) + \rho(B) + \lambda)$$

其中, $\rho(M)$ 为目标块像素的前景概率, $\rho(B)$ 为像素背景区域概率, M, B 无实际含义,仅用于区分前景、背景。在第 t 帧时,提取目标块前景与背景区域颜色特征用于更新颜色 $\rho(M)$ 和 $\rho(B)$ 的概率,并用于计算与更新权重 β_t ,模型的更新过程为:

$$\begin{cases} \rho_t(M) = (1-\alpha_{cfb})\rho_{t-1}(M) + \alpha_{cfb}\rho_t(M) \\ \rho_t(B) = (1-\alpha_{cfb})\rho_{t-1}(B) + \alpha_{cfb}\rho_t(B) \end{cases} \quad (6)$$

在第 t 帧搜索目标位置时,以第 $t-1$ 帧目标位置为中心区域,计算图像中每个像素的得分,以此获取相应图像直方图响应 f_{cfb} ,通过式(3)将响应进行线性加权实现目标位移估计。

基于相关滤波的特征响应线性加重的过程和思路如图 1 所示。

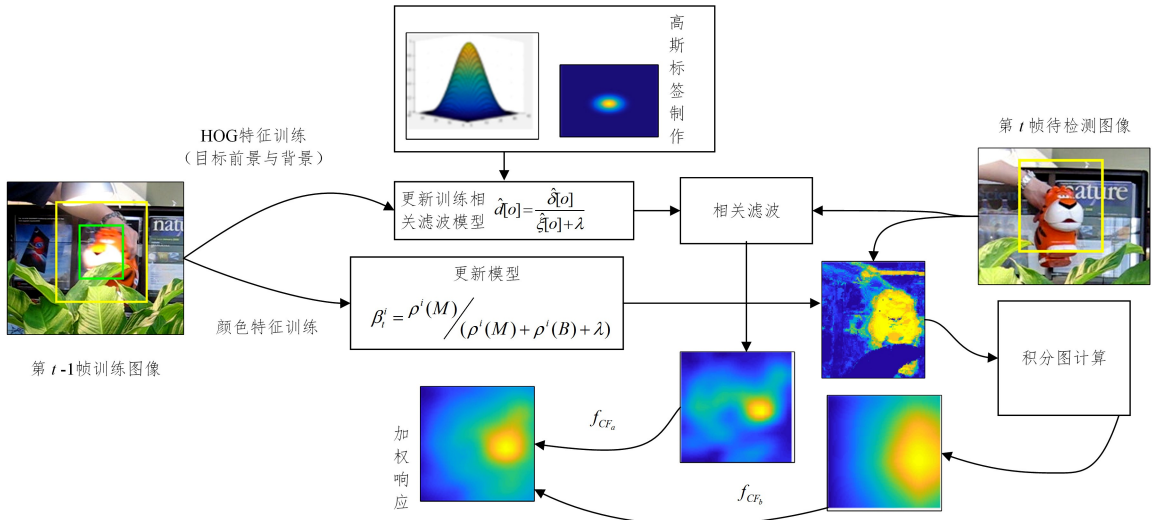


图 1 基于相关滤波的特征响应加权模型与其更新过程

Fig. 1 Correlation filter based characteristics weighted model and update process

2 基于核相关滤波的尺度回归和跟踪置信度判别模型

目标跟踪过程中除了有光照变化、运动模糊等情形外,还会出现的旋转尺度变化以及遮挡等问题。在跟踪过程中,随着目标模型的更新,相关滤波模板会出现漂移现象,导致长时间的目标跟踪失败。因此,本文学习一个基于核相关滤波模型 CF_c 来执行目标运动过程中的尺度估计,另外设置了目标预测阈值。该滤波模型还用于检验目标跟踪的置信度,即目标预测是否可信。

2.1 核相关滤波

相关滤波方法是根据序列在当前帧的信息和之前所在帧的信息,来学习一个相关滤波器,将下一帧新的输入与滤波模板进行相关计算,将得到的最大输出响应作为预测的跟踪目标结果。而核相关滤波是一种基于核函数岭回归的跟踪器,在实际应用中,线性空间中的岭回归显然不能保证样本存在线性可分关系,因此需要使用更为普遍的非线性回归方法,利用核技巧将非线性样本映射到更高维的空间中进行快速内积运算。

定义核函数为 $\phi(x)$,将样本空间进行非线性变换,核回归模型^[22]为:

$$f(z) = \sum_{i=1}^n \omega_i \phi(z, x_i), x_i = 1, 2, \dots, n \quad (7)$$

通过求解可得到^[23]:

$$\omega = (\mathbf{K} + \lambda \mathbf{I})^{-1} \mathbf{y} \quad (8)$$

其中, $\mathbf{K}_{ij} = \phi(x_i, x_j)$ 表示核矩阵, 为保证矩阵能够进行循环移位操作, 即保证 \mathbf{K} 为循环矩阵, 需满足存在任意偏移矩阵 \mathbf{W} , 都有 $k(\mathbf{x}, \mathbf{x}') = k(\mathbf{W}\mathbf{x}, \mathbf{W}\mathbf{x}')$, $\mathbf{K}_{ij} = k(x_i, x_j)$, \mathbf{x} 表示特征向量, 详情请参考文献[16]。给定循环数据 $C(x)$, 对于任意置换矩阵 \mathbf{M} , 核函数需满足 $\phi(\mathbf{x}, \mathbf{x}') = \phi(\mathbf{M}\mathbf{x}, \mathbf{M}\mathbf{x}')$, 再转换到傅里叶域求解式(1)可得:

$$\hat{\omega} = \left(\frac{1}{\hat{\mathbf{k}}^{xx} + \lambda} \right) * \odot \hat{\mathbf{y}} = \frac{\hat{\mathbf{y}}}{\hat{\mathbf{k}}^{xx} + \lambda} \quad (9)$$

其中, 符号 $\hat{\cdot}$ 表示经过傅里叶转换后的结果, \mathbf{k}^{xx} 表示循环矩阵的第一行向量, $\mathbf{k}^{xx} = [k(\mathbf{x}, \mathbf{x}), k(\mathbf{x}, P\mathbf{x}), \dots, k(\mathbf{x}, P^{n-1}\mathbf{x})]$, 可推广到一般 $\mathbf{k}^{xx'}$, 即 \mathbf{x} 与 \mathbf{x}' 不相等时, $k_i^{xx'} = k(\mathbf{x}, P^{n-i}\mathbf{x}')$ 。

满足上述循环矩阵要求的核函数有高斯核函数, 其形式为:

$$\phi(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{1}{\sigma^2} \|\mathbf{x} - \mathbf{x}'\|^2\right) \quad (10)$$

此时模型(2)中的求解过程为:

$$\mathbf{k}^{xx'} = \exp\left(-\frac{1}{\sigma^2} (\|\mathbf{x}\|^2 + \|\mathbf{x}'\|^2 - 2\mathcal{F}^{-1}(\hat{\mathbf{x}} * \odot \hat{\mathbf{x}}'))\right) \quad (11)$$

除此之外, 满足定理 1^[16] 的核函数还有多项式核函数、线性核函数以及指数核函数等, 本文主要用到高斯核函数, 因此其余核函数不在此详细说明。

在目标跟踪中, 训练样本提取目标区域图像块(目标区域和背景区域)的特征向量 \mathbf{x}_i 以及它的循环移位样本特征矩阵 $\mathbf{x}_i (i=2, 3, \dots, n)$, 样本的标签制作也是基于此的, 越靠近目标区域即循环移位距离越小, 则为正样本的可能性就越大, 本文通过高斯回归模型来生成响应的标签。同理, 对待检测的图像块 \mathbf{z} 进行循环移位, 生成候选样本集, 结合回归模型及循环矩阵对角化的特性, 得到相关滤波模型:

$$\hat{f}(z) = \hat{\mathbf{k}}^{xz} \odot \hat{\omega} \quad (12)$$

其中, $f(z)$ 是一个由待检测样本的输出所组成的一个向量, 向量中的每一个元素代表候选区域与目标的相似概率, 或者称其为置信度, 因此 $f(z)$ 中最大元素所对应的位置即可作为跟踪目标的位置。

2.2 跟踪置信度判别与尺度估计

本文基于核相关滤波训练模型 CF_c 进行尺度估计与跟踪置信度判别, 尺度估计能够防止出现跟踪漂移进而导致跟踪失败, 跟踪置信度能够判别目标位置预测是否可信, 据此更新模型防止出现跟踪漂移。当置信度低于设置阈值时, 将启用在线分类检测以重新恢复跟踪。

与第 1 节中的目标特征提取方式不同, 这里仅提取目标所在位置的 HOG 特征, 不考虑目标背景及其颜色特征。假设图像大小为 (M, N) , 对特征样本进行循环移位可以得到 $x_{(m,n)} (m, n) \in \{0, 1, \dots, M-1\} \times \{0, 1, \dots, N-1\}$ 。对于每个样本通过高斯函数生成对应的高斯标签 $y(m, n)$, 结合式(7), 核函数岭回归模型为 $f(z) = \sum_{m,n} \omega_{(m,n)} \phi(z, x_{(m,n)})$, ϕ 函数见

式(10), 由此定义训练模型 CF_c 的损失函数为:

$$L = \arg \min_{\omega} \sum_{m,n} |\phi(x_{(m,n)}) \times \omega - y(m, n)|^2 + \lambda |\omega|^2 \quad (13)$$

与训练样本集的生成方式相同, 对待检测特征样本进行循环移位得到 \mathbf{z} , 利用 2.1 节中的核相关滤波的方法求解运算式(12), 得到核相关滤波模型的最大响应:

$$\hat{y} = \mathcal{F}^{-1}(f(z)) = \mathcal{F}^{-1}(\hat{\mathbf{k}}^{xz} \odot \hat{\omega}) \quad (14)$$

其中, \odot 表示 Hadamard 乘积, 模型 CF_c 的训练更新流程如图 2 所示。

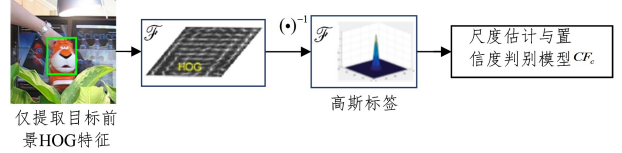


图 2 基于核相关滤波的置信度判别与尺度估计

Fig. 2 Confidence discrimination and scale estimation based on kernel correlation filtering

本文定义置信度判别阈值为 TS_a , 仅当式(14)中的最大响应 \hat{y} 大于阈值, 即 $\max(\hat{y}) > TS_a$ 时才更新模型 CF_c , 当目标出现遮挡时, 不更新该模型, 以保证目标在跟踪过程中不会出现模型的漂移。

通过第 1 节中所述位移估计模型 CF_c 获得第 t 帧目标的预测位置 (x_t, y_t) , 在 (x_t, y_t) 位置附近构造目标金字塔以进行尺度估计, 令目标大小为 $P \times Q$, 金字塔层数为 N , 则有 $S = \{a_n | n = [-\frac{N-1}{2}, [-\frac{N-3}{2}, \dots, [\frac{N-1}{2}]]\}$ 。对于 $s \in S$, 选取预测位置中心的 $sP \times sQ$ 区域块, 并将其重新调整到 $P \times Q$ 大小, 构建尺度特征金字塔, 然后提取 HOG 特征得到其响应, 目标的最佳尺寸为:

$$\hat{s} = \arg \max_s \{ \max(\hat{y}_1), \max(\hat{y}_2), \dots, \max(\hat{y}_s) \} \quad (15)$$

其中, \hat{s} 为最优尺度, 当置信度 $\max(\hat{y}_{\hat{s}}) \geq TS_a$ 时, 更新模型 CF_c ; 当置信度低于给定的阈值 TS_b 时, 即 $\max(\hat{y}_{\hat{s}}) < TS_b$, 将激活再检测模块, 恢复目标以实现长时间跟踪。

3 目标跟踪在线检测器

在目标出现严重遮挡或丢失时会引起跟踪失败, 对于每个目标, 本文使用核相关滤波器 CF_c 计算置信度, 当置信度的得分低于给定阈值时, 将会激活检测器。设置阈值的目的是减小在每帧中滤波操作的计算量, 从而提高跟踪效率。

本文使用一种在线 SVM 分类器作为检测器, 通过在预估位置和不同尺度框位置附近提取大量的训练本来逐步训练 SVM 分类器, 并根据类似于重叠比例的方式^[18] 为这些样本分配二进制标签, 以颜色直观图的量化关系表示图像特征。其中 CIE LAB 空间中的每个通道被量化成为 4 个区间, 为了提高对于光照变化的鲁棒性, 将非参数局部秩^[23] 应用于 L 通道。给定训练集 $\{(v_i, c_i) | i = 1, 2, \dots, N\}$, 该帧中有 N 个样本, v_i 表示第 i 个样本生成的特征向量, $c_i \in \{+1, -1\}$ 表示类的标签, 则 SVM 检测器求解超平面的目标函数为^[24]:

$$\begin{cases} \min_h \frac{\lambda}{2} \|h\|^2 + \frac{1}{N} \sum_i l(h; (\mathbf{v}_i, c_i)) \\ l(h; (\mathbf{v}, c)) = \max\{0, 1 - c \langle h, \mathbf{v} \rangle\} \end{cases} \quad (16)$$

其中, $\langle h, \mathbf{v} \rangle$ 表示内积运算。本文采用一种被动选择的算法^[25]来有效更新超平面参数:

$$h = h - \frac{l(h; (\mathbf{v}, c))}{\|\nabla_h l(h; (\mathbf{v}, c))\|^2 + \frac{1}{2\tau}} \nabla_h l(h; (\mathbf{v}, c)) \quad (17)$$

其中, $\nabla_h l(h; (\mathbf{v}, c))$ 表示损失函数的梯度; $\tau \in (0, +\infty)$ 表示更新 h 的学习速率参数。类似于相关滤波器 CF_i , 仅当其置信度得分低于设置阈值 TS_b 时, 才使用式(17)更新超平面参数。

4 基于相关滤波的长时间自适应目标跟踪算法

在目标跟踪过程中, 为了达到长时间的鲁棒跟踪效果, 本文训练了 3 个相关滤波器和 1 个目标丢失在线检测器, 其中相关滤波器 CF_a 和 CF_b 用于得到目标运动过程中的位移估计, 预测目标位置; 相关滤波器 CF_c 用于估计目标的最佳尺

度, 并且判断是否出现遮挡或丢失并激活在线检测器; 在线检测器采用 SVM 在线分类器将目标从丢失状态中恢复, 实现长时间跟踪。

本文提出的基于相关滤波长时间自适应目标跟踪算法如图 3 所示, 算法的具体流程如下:

Step1 对于第 t 帧图像, 根据 (x_{t-1}, y_{t-1}) 在第 t 帧图像中选取目标框前景与背景图像块, 提取其 HOG 特征训练相关滤波器 CF_a , 提取颜色直方图特征训练 CF_b ;

Step2 将特征响应 f_{CF_a} 和 f_{CF_b} 进行线性加权, 使用位移估计模型 CF_w 计算响应, 并估计当前帧的位置 (x_t, y_t) ;

Step3 在 (x_t, y_t) 位置附近构造目标金字塔以进行尺度估计, 使用模型 CF_c 得到最佳尺度 \hat{s} , 得到第 t 帧的状态 (x_t, y_t, s_t) ;

Step4 进行跟踪结果置信度判别, 当 $\max(\hat{y}_s) < TS_b$ 时, 启用 SVM 在线检测器, 更新超平面参数 h ;

Step5 当 $\max(\hat{y}_s) \geq TS_a$ 时, 更新模型 CF_c ;

Step6 循环上述步骤, 当图像序列结束时, 算法终止。

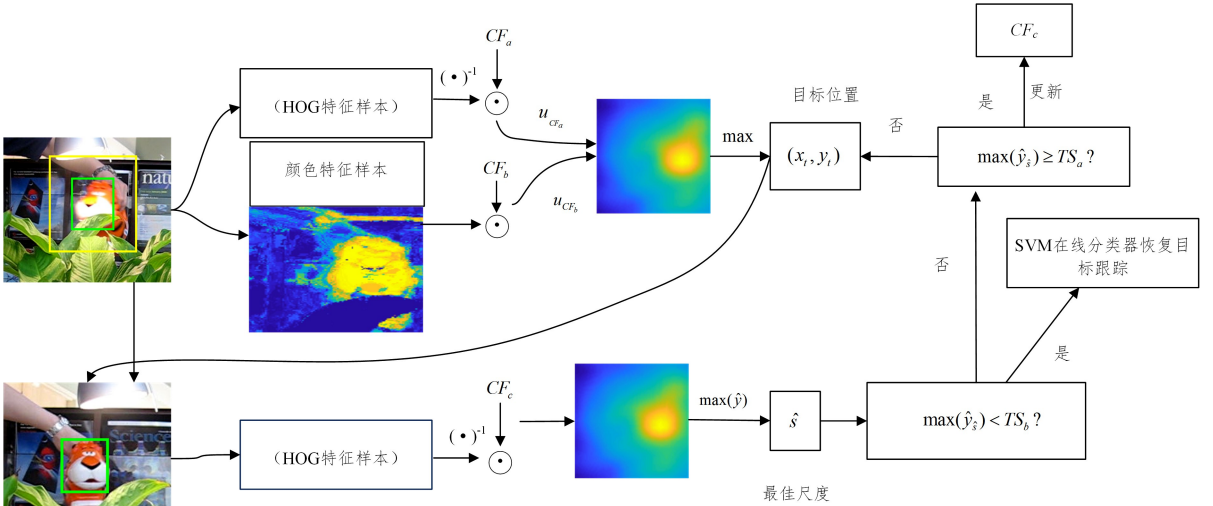


图 3 基于相关滤波长时间自适应目标跟踪算法框图

Fig. 3 Block diagram of long-term adaptive target tracking algorithm based on correlation filtering

5 目标跟踪的实验结果与对比分析

本文在目标跟踪数据集 OTB50 上对算法进行测试, 并与目前较为主流的目标跟踪算法进行比较, 为方便表示, 本文算法在以下部分均用 ours 算法表示。OTB50 数据集共包含 50 个视频图像测试序列, 这些序列包含目标跟踪过程中可能出现的快速运动、光照变化、运动模糊、遮挡以及尺度变化等一系列问题。OTB 数据集能够用于评估算法的性能, 国内外主流目标跟踪算法均在此数据集上进行性能测试, 测试结果具备说服力。本文提出的基于相关滤波的长时间自适应目标跟踪算法以及所有用于对比分析的主流算法均在 2.5 GHz 的 Intel i5-7300HQ CPU 以及 8.0 GB RAM 的计算机上使用 MATLAB 2017a 运行实现。

5.1 算法参数设置

本文提出的基于相关滤波的长时间自适应目标跟踪算法共设计了 3 个相关滤波器和 1 个目标丢失在线检测器。其

中, 位移估计的模型 CF_c 涉及到线性加权参数 u_{CF_a} 和 u_{CF_b} , 满足 $u_{CF_a} + u_{CF_b} = 1$, 当 $u_{CF_a} = 0.7$ 时, 算法表现出了最佳性能。在 CF_a 模型中, 模型参数更新学习率 α_{CF_a} 取 0.01; 在 CF_b 模型中, 参数更新学习率 α_{CF_b} 取 0.04。在尺度估计与置信度判别模型中, 用于更新位移估计模型的阈值 TS_a 设置为 0.5, 用于激活 SVM 再检测模块的阈值 TS_b 设置为 0.23。

5.2 目标跟踪算法的性能评估

5.2.1 定量分析

本节给出本文提出的基于相关滤波的长时间自适应目标跟踪算法和主流算法的性能比较结果。在空间鲁棒性 SRE 评估中, ours 算法与其他 9 种算法相比性能更具优势, 事实证明, ours 算法虽然训练了较多的相关滤波模型, 但是在目标跟踪过程中, 仍然能够达到实时跟踪的目的。OPE 是指 OTB 序列第一帧被标注的真实位置进行初始化后, 得到的跟踪算法的平均精度和成功率, 该指标只通过运行一次跟踪算法而得到。从图 4 可以看到, 本文算法的一次成功率和平均精度均处于领先地位。

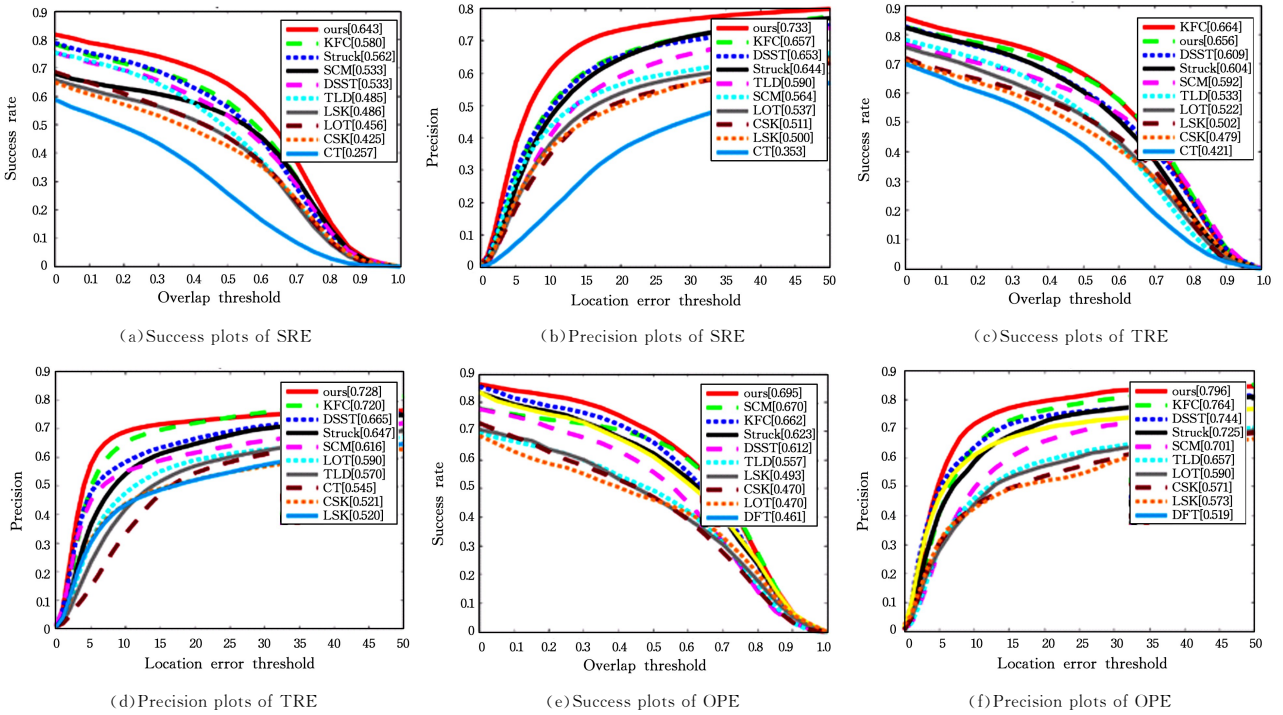


图4 目标跟踪算法的精度和成功率

Fig. 4 Accuracy and success rate of target tracking algorithms

如图5所示,本文针对目标跟踪的具体难点(形变、遮挡以及超出平面内旋转),通过OPE估算标准得到ours算法与主流算法比较的成功率和精度。从图中可以看出,本文算法

在处理遮挡、变形以及超出平面内的旋转时具有较好的效果,验证了本文的尺度估计置信度判别模型的准确性和优越性,以及在处理遮挡问题时采用在线SVM分类检测器的合理性。

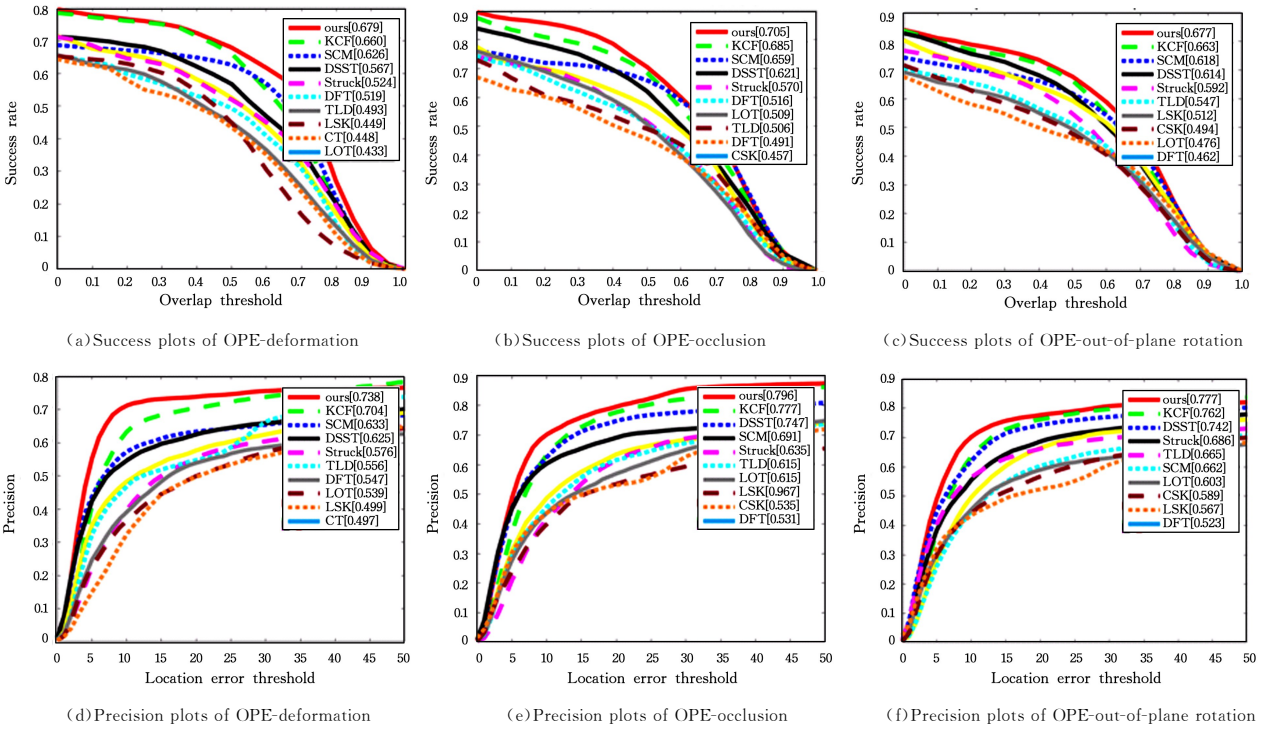


图5 OPE评估标准下本文算法与主流算法对于形变、遮挡、超出平面内旋转跟踪的对比

Fig. 5 Comparison of ours algorithm and mainstream algorithms for deformation, occlusion, out-of-plane rotation tracking under OPE evaluation standard

表1列出了本文算法与其他算法在测试数据集OTB50上的性能对比,其在距离准确度、中心误差等指标上有一定的

性能提升,说明本文算法的长时自适应目标跟踪优于其他算法。

表1 在50个序列上不同跟踪算法的性能对比

Table 1 Performance comparison of different tracking algorithms on 50 sequences

定性指标	ours	CSK	KCF	TLD	Struck	DSST	LCT	SCM
距离准确度/%	82.4	52.5	72.8	55.7	63.4	73.2	81.9	62.1
重叠率/%	74.7	43.2	61.3	51.6	54.9	62.6	74.9	60.4
中心误差	27.8	89.7	38.7	47.9	50.1	35.3	27.4	54.2
速度(FPS)	21	219	78	16	6	46	26	0.2

5.2.2 定性分析

为了直观地展现算法的性能,本文在 OTB50 上选取部分序列结果进行展示,并与主流算法进行比较。OTB 数据集中含有光照变化、快速运动、运动模糊、遮挡丢失等几大视觉挑战序列,其序列结果能够反映出每种算法应对视觉挑战的性能。

本文算法与主流算法 LCT,DSST,KCF,TLD,Struck 在 OTB 中的 7 个典型序列上进行实际跟踪的对比效果如图 6 所示,序列名称从上至下依次为 woman,football,tiger2,jogging2,football1,carDark,soccer。

根据图 6,将本文算法与其他算法在几个跟踪难点上进

行对比,可以看出本文算法具有如下优势:

(1)运动模糊,颜色干扰。LCT 算法将 HOG 特征和 HOI 特征结合来训练相关滤波器,同时还训练了一个在线的随机的 fern 分类器用于目标丢失再检测,但其没有结合颜色特征,当图像序列出现颜色干扰、分辨率低等问题时,跟踪效果不太理想,如在 carDark 序列中本文算法表现良好。

(2)遮挡、目标丢失。序列 jogging-2,tiger2 上存在目标遮挡,从而导致 CSK,DSST 算法出现目标跟丢的问题,而本文算法在出现遮挡甚至丢失目标之后能够快速进行再检测,恢复跟踪,实现长时跟踪。TLD 算法具有长时检测模块,但是跟踪模块采用传统的光流法,其在众多具有挑战的跟踪图像序列上的总体性能不太理想,即使通过检测子模块再次实现跟踪,依然会出现丢失许多帧的情况。

(3)快速移动,尺度变化。从序列 football 和 football1 中可看出,Struck 算法^[26]在尺度变化、快速运动、背景复杂等情况下的跟踪效果都不太理想,其在遮挡处理上有一定的鲁棒性。本文算法结合了 HOG 特征和颜色直方图特征,对快速运动和颜色变化都具有较好的鲁棒性。



图6 本文算法与主流算法在 OTB 中的 7 个典型序列上进行实际跟踪的对比效果

Fig. 6 Comparison among ours algorithm and mainstream algorithms on actual tracking of 7 typical sequences in OTB

虽然 KCF 算法融合了多通道的 HOG 特征来训练核相关滤波器,但其没有考虑跟踪过程中的遮挡以及尺度变化,当出现遮挡、尺度变化、快速运动时,KCF 的跟踪效果较差,如序列 jogging-2,carDark。DSST 将算法分为位置滤波器和尺度滤波器,在自适应尺度变化上取得了一定的效果,且能够和其他没有考虑尺度的算法相结合,但其也未考虑目标遮挡丢失等问题,难以实现长时间的目标跟踪,如序列 jogging-2。

上述实验结果表明,本文提出的基于相关滤波的鲁棒自适应长时目标跟踪算法在运动模糊、光照变化、尺度估计以及

遮挡丢失等问题上具有准确性和鲁棒性,相比其他算法有明显优势。

结束语 本文提出了一种基于相关滤波的鲁棒长时自适应目标跟踪算法。该算法主要分为 3 个部分:1)位移估计,采用一种特征响应互补策略,提取目标前景及背景的 HOG 特征和全局颜色直方图特征,将训练的两个模型的输出响应进行线性加权,得到对颜色变化和形变等都具有鲁棒性的位移估计模型;2)尺度估计与置信度判别,仅提取目标前景的 HOG 特征,训练一个用于尺度估计和置信度判别的相关滤波

模型,该模型能保持对目标外观的长期记忆,使用该长期滤波器的输出响应来判别是否出现遮挡或跟踪失败并且激活再检测模块,构建以估计的目标位置为中心的特征金字塔模型来预测尺度变化,以防止目标框漂移;3)SVM再检测分类器,通过在预估位置和不同尺度框位置附近提取大量的训练样本来逐步训练SVM分类器,通过置信度判别中的阈值来确定是否激活该分类器,将跟踪从失败状态中恢复。最后,本文在目标跟踪基准数据集 OTB50 上对算法进行验证,并与目前的主流算法进行对比分析,实验结果进一步验证了本文算法的鲁棒性和优越性。

参 考 文 献

- [1] LIU D Q, LIU W J, FEI B W. Anti-jamming Matched Target Tracking Method with Foreground Constraints[J]. *Journal of Automation*, 2018, 44(6): 1138-1152.
- [2] WANG Y N, LUO Q H, CHEN Y J. Multi-aircraft Visual-tracking and Location System and Method for Rotor Flying Robot[J]. *Chinese Journal of Scientific Instrument*, 2018(2): 1-10.
- [3] YUAN J, YANG L, DONG X L. Mobile Robot Target Tracking Based on Online Classification of Motion Patterns[J]. *Chinese Journal of Scientific Instrument*, 2017, 38(3): 568-577.
- [4] WU Y, LIM J, YANG M H. Online object tracking: A benchmark[C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2013: 2411-2418.
- [5] WU Y, LIM J, YANG M H. Online object tracking: A benchmark[C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2013: 2411-2418.
- [6] KRISTAN M, MATAS J, LEONARDIS A, et al. The visual object tracking vot2015 challenge results[C] // *Proceedings of the IEEE International Conference on Computer Vision Workshops*. 2015: 1-23.
- [7] KRISTAN M, MATAS J, LEONARDIS A, et al. The visual object tracking VOT2015 challenge results[C] // *In Proceedings of the 2015 IEEE International Conference on Computer Vision Workshops*. Santiago, Chile: IEEE, 2015: 564-586.
- [8] LIU C, ZHAO W, LIU P, et al. Selection, tracking and updating of auxiliary targets in target tracking[J]. *Journal of Automation*, 2018, 44(7): 1195-1211.
- [9] XIONG D, LU H M, XIAO J H, et al. Long time target tracking with scale and rotation adaptability[J]. *Journal of Automation*, 2019, 45(2): 289-304.
- [10] BERTINETTO L, VALMADRE J, GOLODETZ S, et al. Staple: Complementary learners for real-time tracking[C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016: 1401-1409.
- [11] HADFIELD S J, BOWDEN R, LEBEDA K. The visual object tracking VOT2016 challenge results[J]. *Lecture Notes in Computer Science*, 2016, 9914: 777-823.
- [12] MA C, YANG X, ZHANG C, et al. Long-term correlation tracking[C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015: 5388-5396.
- [13] DANELLJAN M, BHAT G, SHAHBAZ K F, et al. Eco: Efficient convolution operators for tracking[C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017: 6638-6646.
- [14] DANELLJAN M, HÄGER G, KHAN F, et al. Accurate scale estimation for robust visual tracking[C] // *British Machine Vision Conference*. Nottingham, 2014: 1-5.
- [15] BLASCHKO M B, LAMPERT C H. Learning to localize objects with structured output regression[C] // *European Conference on Computer Vision*. Springer, Berlin, Heidelberg, 2008: 2-15.
- [16] 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*, 2014, 37(3): 583-596.
- [17] HARE S, GOLODETZ S, SAFFARI A, et al. Struck: Structured output tracking with kernels[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 38(10): 2096-2109.
- [18] ZHANG J, MA S, SCLAROFF S. MEEM: robust tracking via multiple experts using entropy minimization [C] // *European Conference on Computer Vision*. Springer, Cham, 2014: 188-203.
- [19] DANELLJAN M, HÄGER G, KHAN F, et al. Accurate scale estimation for robust visual tracking[C] // *British Machine Vision Conference* Nottingham. BMVA Press, 2014: 2014.
- [20] 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*. IEEE, 2010: 2544-2550.
- [21] HENRIQUES J F, CASEIRO R, MARTINS P, et al. Exploiting the circulant structure of tracking-by-detection with kernels [C] // *European Conference on Computer Vision*. Springer, Berlin, Heidelberg, 2012: 702-715.
- [22] DANELLJAN M, SHAHBAZ K F, FELSBERG M, et al. Adaptive color attributes for real-time visual tracking[C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2014: 1090-1097.
- [23] ZABIH R, WOODFILL J. Non-parametric local transforms for computing visual correspondence[C] // *European Conference on Computer Vision*. Springer, Berlin, Heidelberg, 1994: 151-158.
- [24] MA C, HUANG J B, YANG X, et al. Adaptive correlation filters with long-term and short-term memory for object tracking[J]. *International Journal of Computer Vision*, 2018, 126(8): 771-796.
- [25] CRAMMER K, DEKEL O, KESHET J, et al. Online passive-aggressive algorithms[J]. *Journal of Machine Learning Research*, 2006, 7(Mar): 551-585.
- [26] HARE S, GOLODETZ S, SAFFARI A, et al. Struck: Structured output tracking with kernels[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 38(10): 2096-2109.



TAN Jian-hao, born in 1962, Ph.D, professor. His research interests include intelligent robot, data mining, pattern recognition, system identification and image processing.



YIN Wang, born in 1995, master. His main research interests include machine vision, robot technology, etc.