

结合自适应空间权重的改进型 时空正则项跟踪算法

王鹏^{1,2}, 孙梦宇¹, 王海燕³, 李晓艳¹, 吕志刚¹

(1. 西安工业大学电子信息工程学院, 710021, 西安; 2. 西北工业大学航海学院, 710072, 西安;
3. 陕西科技大学电子信息与人工智能学院, 710021, 西安)

摘要: 为解决时空正则项的相关滤波视觉跟踪算法在目标部分遮挡时存在的模型漂移和尺度估计不准确问题,提出了结合自适应空间权重的改进型时空正则项跟踪算法。采用平均特征能量比将无法准确表达目标或过多表达背景信息的特征通道裁剪掉,以提高跟踪精度。在滤波器训练时加入空间权重正则项,利用时间正则项在目标遮挡时被动更新滤波器,使得在空间权重更新时更为准确,以此着重学习目标未被遮挡部分,获取可靠的相关滤波器系数。将滤波器求解划分为2个子问题,分别采用交替方向乘子法进行优化计算,保证算法运算速率。在牛顿迭代法中设置精度阈值,在保证定位精度的同时减少迭代次数。实验结果表明:在 OTB-100 数据集上所选择的6个视频序列中,所提算法较 STRCF 算法的平均中心位置误差降低了12.3像素,平均重叠率增加了7%,运算帧率可达19.25帧/s;在 OTB2015 遮挡视频序列中,所提算法较 STRCF 算法的成功率曲线下的积分面积(S_{AUC})增加了0.7%,使用深度特征的所提算法较 DeepSTRCF 和 ASRCF 算法的 S_{AUC} 分别提升了3.9%与0.9%。

关键词: 视觉跟踪;相关滤波;时间正则项;牛顿迭代法;空间权重正则项

中图分类号: TP391 **文献标志码:** A

DOI: 10.7652/xjtuxb202005021 **文章编号:** 0253-987X(2020)05-0158-12



OSID 码

An Improved Spatial-Temporal Regularization Term Tracking Algorithm Combining Adaptive Spatial Weights

WANG Peng^{1,2}, SUN Mengyu¹, WANG Haiyan³, LI Xiaoyan¹, LÜ Zhigang¹

(1. School of Electronics and Information Engineering, Xi'an Technological University, Xi'an 710021, China;

2. School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an 710072, China; 3. School of Electronic Information and Artificial Intelligence, Shaanxi University of Science & Technology, Xi'an 710021, China)

Abstract: To solve the problem of model drifting and inaccurate scale estimation for partially occluded object in spatial-temporal regularized correlation filter visual tracking algorithm, an improved spatial-temporal regularization term tracking algorithm combining adaptive spatial weights is proposed. For cutting out the feature channel that unables to express the background information more accurately, the average feature energy ratio channel is chosen to improve the tracking accuracy. Then the spatial weight regularization term is added to the filter training, and the time regularization term is used to passively update the filter when the target is occluded, so that the spatial weight update gets more accurate. Focusing on learning the non-occluded part of the target, the reliable correlation filter coefficient is obtained. The filter solution is divided into

收稿日期: 2019-08-12。 作者简介: 王鹏(1978—),男,教授。 基金项目: 国家自然科学基金资助项目(61271362);陕西省科技厅重点研发计划资助项目(2019GY-022,2019GY-066)。

网络出版时间: 2019-12-26

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1069.T.20191226.0913.002.html>

two subproblems, the alternating direction multiplier method is used for optimization calculation to ensure the algorithm operation rate. The accuracy threshold is set in Newton iteration method to reduce the number of iterations as ensuring the positioning accuracy. In the selected video sequences of OTB-100 dataset, the proposed algorithm reduces the center position error by 12.3 pixels, the average overlap rate increases by 7%, and the frames per second reaches up to 19.25, compared with those of the algorithm STRCF. And the integral area under the success curve of STRCF in the occlusion video sequence of OTB2015 dataset is increased by 0.7%, while the integral area of DeepSTRCF and ASRCF is increased by 3.9% and 0.9%, respectively, by the deep feature version of the proposed algorithm.

Keywords: visual tracking; correlation filter; temporal regularization term; Newton iteration method; spatial-weight regularization term

目标跟踪是机器视觉领域重要的研究方向之一^[1]。在目标跟踪过程中,只需在视频的第一帧图像内确定目标,即可在接下来的图像序列中自动实现对目标的定位和尺度的估计。由于目标在跟踪过程中可能存在遮挡^[2-3]、尺度变化^[4-5]、形变^[6-7]等多种情况,因此跟踪算法的精度受到了一定程度影响^[8]。

近年来,由于相关滤波跟踪算法在具有高速率的同时也具有高精度,因此受到了广泛的研究关注。2014年,Li Y等提出了具有特征积分的尺度自适应核相关滤波器跟踪算法(SAMF),在平移滤波器估计位置之后,从多尺度缩放的图像块上进行目标检测,取响应值最大的位置对应的位移和尺度作为目标估计状态,但出现遮挡时会跟踪失败^[9]。2015年,Danelljan等提出了空间正则项相关滤波跟踪算法(SRDCF),利用空间正则项解决边界问题,但运行速率较慢^[10-11]。2017年,Danelljan等提出了基于快速连续卷积的相关滤波跟踪算法(ECO),通过主成分分析法降低特征维度,采用样本融合的多帧样本进行训练,但出现遮挡时会发生模型漂移的现象^[12]。2018年,Li F等提出了基于空间-时间正则项的相关滤波跟踪算法(STRCF),该算法在离散模型中加入时间-空间正则项,保证模型在时间上的连续性,但出现遮挡较为严重时依然会产生模型漂移的现象^[13]。2019年,Dai等提出了自适应空间正则项的跟踪算法(ASRCF),采用空间权重自适应正则项使滤波器能够适应目标遮挡与形变,但遇到遮挡较为严重时跟踪误差较大^[14]。

基于上述问题,本文提出了结合自适应空间权重的改进型时空正则项跟踪算法。通过计算各个特征通道的平均特征能量比,将未能准确表达目标信息的特征通道裁剪掉。在模型训练过程中加入自适应空间权重正则项,通过更新空间权重使得滤波器

准确学习未被遮挡的目标信息。采用自适应牛顿迭代法估计目标状态,减少多余的迭代次数,实现对目标的精确跟踪。

1 时空正则项相关滤波跟踪

相关滤波算法应用于跟踪,本质上是提取目标多种特征,利用岭回归的方式训练滤波器,通过循环矩阵扩展样本,利用空间权重解决边界效应,从而增加算法定位跟踪^[10],并引入时间正则项来增加算法鲁棒性^[13]。

设在当前帧图像中提取到的样本为 $\mathbf{x}^d$, d 表示特征通道索引且 $d=1,\dots,D$, D 表示特征通道总数, $\mathbf{y}$ 为高斯型的理想目标响应函数, $\mathbf{f}$ 为本帧的滤波器系数, $\mathbf{f}_{t-1}$ 为上一帧滤波器系数, $\mathbf{w}$ 为空间权重,则滤波器系数 $\mathbf{f}$ 的求解公式为

$$\mathbf{f} = \arg \min_{\mathbf{f}} \left(\frac{1}{2} \left\| \sum_{d=1}^D \mathbf{x}^d * \mathbf{f}^d - \mathbf{y} \right\|_2^2 + \frac{1}{2} \sum_{d=1}^D \left\| \mathbf{w} \mathbf{f}^d \right\|_2^2 + \frac{\mu}{2} \left\| \mathbf{f} - \mathbf{f}_{t-1} \right\|_2^2 \right) \quad (1)$$

式中: $\sum_{d=1}^D \left\| \mathbf{w} \mathbf{f}^d \right\|_2^2$ 为空间正则项; $\left\| \mathbf{f} - \mathbf{f}_{t-1} \right\|_2^2$ 为时间正则项; μ 为时间正则项系数; $*$ 为相关运算。

采用交替方向乘子法^[15]对式(1)进行求解。引入变量 $\mathbf{g}$ 和参数 γ ,且令 $\mathbf{g} = \mathbf{f}$,因此式(1)可改写为

$$\mathbf{f}, \mathbf{g}, \mathbf{s} = \arg \min_{\mathbf{f}, \mathbf{g}, \mathbf{s}} \left(\frac{1}{2} \left\| \sum_{d=1}^D \mathbf{x}^d * \mathbf{f}^d - \mathbf{y} \right\|_2^2 + \frac{1}{2} \sum_{d=1}^D \left\| \mathbf{w} \mathbf{g}^d \right\|_2^2 + \sum_{d=1}^D (\mathbf{f}^d - \mathbf{g}^d)^T \mathbf{s}^d + \frac{\gamma}{2} \sum_{d=1}^D \left\| \mathbf{f}^d - \mathbf{g}^d \right\|_2^2 + \frac{\mu}{2} \left\| \mathbf{f} - \mathbf{f}_{t-1} \right\|_2^2 \right) \quad (2)$$

式中 $\mathbf{s}$ 和 γ 分别为拉格朗日乘子和惩罚系数。令 $\mathbf{h} = \mathbf{s}/\gamma$,式(2)可改写为

$$f, g, h = \arg \min_{f, g, h} \left(\frac{1}{2} \left\| \sum_{d=1}^D x^d * f^d - y \right\|_2^2 + \right. \\ \left. \frac{1}{2} \sum_{d=1}^D \|w g^d\|_2^2 + \frac{\gamma}{2} \sum_{d=1}^D \|f^d - g^d + h^d\|_2^2 + \frac{\mu}{2} \|f - f_{t-1}\|_2^2 \right) \quad (3)$$

将式(3)划分为3个子问题迭代求解,即

$$\left. \begin{aligned} f^{i+1} &= \arg \min_f \left(\left\| \sum_{d=1}^D x^d * f^d - y \right\|_2^2 + \right. \\ &\quad \left. \gamma \|f - g + h\|_2^2 + \mu \|f - f_{t-1}\|_2^2 \right) \\ g^{i+1} &= \arg \min_g \left(\sum_{d=1}^D \|w g^d\|_2^2 + \gamma \|f - g + h\|_2^2 \right) \\ h^{i+1} &= h^i + f^{i+1} - g^{i+1} \end{aligned} \right\} \quad (4)$$

将式(4)进行快速傅里叶变换,根据交替方向乘子法对其求导,求得闭合形式的解,并循环2次,取逆傅里叶变换即可求解得到滤波器系数 f 。利用频域的滤波器系数计算得到响应图

$$\hat{s} = \sum_{d=1}^D \hat{x}^d \hat{f}^d \quad (5)$$

式中 $\hat{\cdot}$ 形式的各变量为原变量傅里叶变换的结果,即原变量的频域形式。在响应图中求解最大值位置作为目标估计位置,最大值对应的尺度作为目标估计尺度。

2 改进型时空正则项跟踪算法

本文在 STRCF 的基础上,提出基于自适应时空正则项的跟踪算法。在目标区域提取灰度、梯度直方图(HOG)^[16]和颜色属性(CN)^[17]共3种特征作为样本,利用平均特征能量比对特征通道进行裁剪,通过加入自适应空间正则项的目标函数计算滤波器系数,在下一帧图像输入时,提取不同尺度的目标特征并与滤波器系数做相关运算计算响应图,以此获取目标估计位置及尺度估计。本文算法框如图

1所示。

2.1 特征通道裁剪

本文采用灰度、HOG及CN共3种特征,其中灰度特征通道数为1,HOG特征通道数为31,CN特征通道数为10,所有特征的通道数为42。但是,并不是每一通道特征均能准确表示目标,以图1中的CN特征为例,可以看到第一通道特征表示的均是背景信息,并未表达目标。因此,为避免包含大量背景信息或没有表达任何信息的特征对滤波器训练带来的负面影响,采用平均特征能量比 R^d 去评价各个特征通道对目标的表达程度,其公式为

$$R^d = \frac{F^d(\mathbf{O})}{1 + F^d(\mathbf{S})} \quad (6)$$

式中: $\mathbf{O}$ 表示跟踪目标区域图像; $\mathbf{S}$ 表示整个检测块图像; $F^d(\mathbf{A})$ 表示第 d 个特征通道中图像块 $\mathbf{A}$ 的平均特征能量,其公式^[18]为

$$F^d(\mathbf{A}) = \frac{\sum_{i,j} P(i,j)}{IJ} \quad (7)$$

其中, I 和 J 分别表示图像块 $\mathbf{A}$ 的长和宽, i 和 j 表示像素点位置坐标, $P(i,j)$ 表示在位置 (i,j) 处经过卷积计算后得到的能量。

R^d 越大,则此通道表示的目标特征越多; R^d 越小,则表示未能学习到目标信息或包含较多的背景信息。因此,本文设定阈值,将小于阈值的特征通道裁剪掉,增加全部通道特征对目标的表达能力,同时运算速率也得到了提升。

由图1可以看出,在 Soccer 序列中,通过计算的平均特征能量比 R^d ,HOG特征从31维降低至14维,CN特征从10维降低至2维,灰度特征由于仅有1维,因此并未计算 R^d ,直接保留,不做更改。总维度从42维降低至17维。

为展示利用 R^d 选择特征的有效性,将部分特征进行了可视化,结果如图2所示。

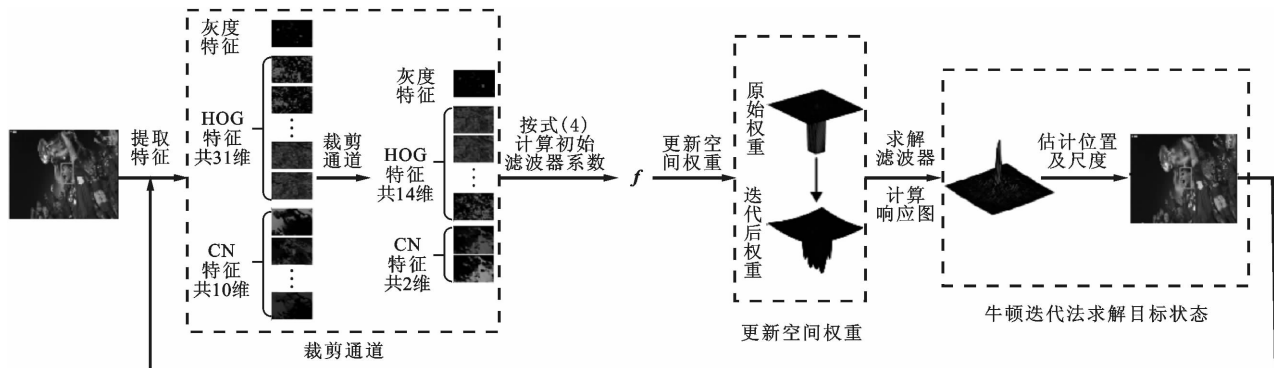
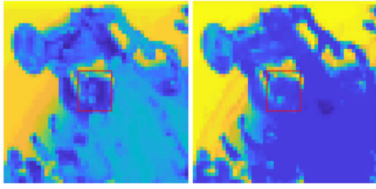
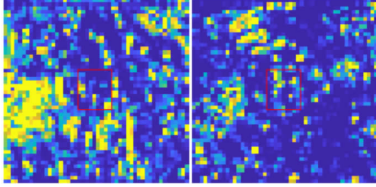


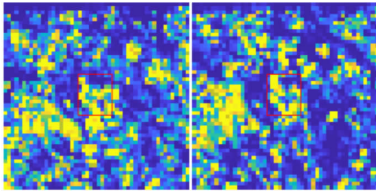
图1 本文算法框图



(a)无法准确表达目标的特征



(b)干扰过多的特征



(c)本文选择的特征

图2 特征可视化结果

图2中方框为跟踪目标区域,颜色越深表示特征值越低。可以看出,虽然图2c中的特征周围干扰依然很大,但相较于图2b中的特征,图2c中的特征对目标的表达能力更好。图2b中的特征对目标的表达能力较低,且周围具有高特征值的干扰项,对目标定位带来的干扰较大,图2c选择的特征相较于其他特征能够对目标进行很好的表达,相较于未被选择的特征对目标定位精度贡献更高。

2.2 模型构建及求解

目标在运动过程中可能存在被部分遮挡的情况,而STRCF仅利用单帧样本对滤波器进行更新,滤波器会学习到遮挡物的信息,因此无法准确跟踪目标。虽然加入的时间正则项保持了滤波器系数在时间上的连续性,但在连续遮挡或遮挡较为严重的情况下,目标跟踪依然会受到影响。因此,在式(1)的基础上加入了空间权重正则项,通过在迭代求解的过程中更新空间权重,使得滤波器系数在学习过程中着重学习未被遮挡或遮挡情况较小的目标信息,以此提高模型的表达能力。构建的滤波器系数求解公式为

$$f = \arg \min_f \left(\frac{1}{2} \left\| \sum_{d=1}^D x^d * f^d - y \right\|_2^2 + \frac{1}{2} \sum_{d=1}^D \|w f^d\|_2^2 + \frac{\mu}{2} \|f - f_{t-1}\|_2^2 + \right.$$

$$\left. \frac{\lambda}{2} \|w - w^r\|_2^2 \right) \quad (8)$$

式中 w^r 为预设的空间权重,将其设置为高斯型函数,表示图像块中各个特征点的重要程度。

式(8)可以划分为两个子问题迭代求解,即

$$\left. \begin{aligned} f &= \arg \min_f \left(\frac{1}{2} \left\| \sum_{d=1}^D x^d * f^d - y \right\|_2^2 + \frac{1}{2} \sum_{d=1}^D \|w f^d\|_2^2 + \frac{\mu}{2} \|f - f_{t-1}\|_2^2 \right) \\ w^i &= \arg \min_{w^i} \left(\frac{1}{2} \sum_{d=1}^D \|w^i f^d\|_2^2 + \frac{\lambda}{2} \|w^i - w^r\|_2^2 \right) \end{aligned} \right\} \quad (9)$$

本文滤波器系数 f 的求解过程与STRCF的相同,在此不做过多描述,但迭代过程中需要先求解滤波器系数 f 用于更新空间权重,因此利用交替方向乘法(ADMM)对其进行求解。滤波器系数 f 初始迭代次数设为1,在更新空间权重后,利用ADMM求解滤波器系数,迭代次数为2。在初始化滤波器系数的过程中, $w^1 = w$ 。

采用ADMM对空间权重 w 进行求解。引入辅助变量 g ,令 $g = w$ 。令 s 和 γ 分别为拉格朗日乘子和惩罚系数,可以得到

$$w, g, s = \arg \min_{w, g, s} \left(\frac{1}{2} \sum_{d=1}^D \|w f^d\|_2^2 + \frac{\lambda}{2} \|g - w^r\|_2^2 + (w - g)s + \frac{\gamma}{2} \|g - w\|_2^2 \right) \quad (10)$$

令 $h = s/\gamma$,则式(10)改写为

$$w, g, h = \arg \min_{w, g, h} \left(\frac{1}{2} \sum_{d=1}^D \|w f^d\|_2^2 + \frac{\lambda}{2} \|g - w^r\|_2^2 + \frac{\gamma}{2} \|w - g + h\|_2^2 \right) \quad (11)$$

将式(11)分为3个子式进行迭代求解,即

$$\left. \begin{aligned} w^{i+1} &= \arg \min_w \left(\frac{1}{2} \sum_{d=1}^D \|w f^d\|_2^2 + \frac{\gamma}{2} \|w - g + h\|_2^2 \right) \\ g^{i+1} &= \arg \min_g \left(\frac{\lambda}{2} \|g - w^r\|_2^2 + \frac{\gamma}{2} \|w - g + h\|_2^2 \right) \\ h^{i+1} &= h^i + f^{i+1} - g^{i+1} \end{aligned} \right\} \quad (12)$$

(1) w 的求解过程。将 w 转换至频域,即

$$\arg \min_w (\|W \hat{f}\|_2^2 + \gamma \|W - G + H\|_2^2) \quad (13)$$

式中: W 为 w 的频域矩阵形式, W 为分块矩阵, $W = W_1 \oplus W_2 \oplus \dots \oplus W_D$, $\oplus$ 表示将参加运算的矩阵放置

在生成矩阵的对角线上, $\mathbf{W}_d$ 为托普利兹矩阵, $\mathbf{W}_d \hat{\mathbf{f}}^d = (\text{vec} \hat{\mathbf{w}}) * \hat{\mathbf{f}}^d$, $\text{vec} \hat{\mathbf{w}}$ 表示将 $\hat{\mathbf{w}}$ 向量化; 矩阵 $\hat{\mathbf{f}} = [(\hat{\mathbf{f}}^1)^\top, (\hat{\mathbf{f}}^2)^\top, \dots, (\hat{\mathbf{f}}^D)^\top]^\top$; 矩阵 $\mathbf{G}$ 和 $\mathbf{H}$ 与 $\mathbf{W}$ 相似, 仅是将对角矩阵全部替换为 $\hat{\mathbf{g}}$ 与 $\hat{\mathbf{h}}$ 。对式(13)进行求导并进行整理, 得到

$$\mathbf{W} = \frac{\gamma(\mathbf{G} - \mathbf{H})}{\lambda \hat{\mathbf{f}}^\top \hat{\mathbf{f}} + \gamma} \quad (14)$$

至此, $\mathbf{w}$ 的频域矩阵形式 $\mathbf{W}$ 求解完毕。

(2) $\mathbf{g}$ 的求解过程。将 $\mathbf{g}$ 转换至频域, 即

$$\arg \min_{\mathbf{g}} (\lambda \|\hat{\mathbf{g}} - \hat{\mathbf{w}}^r\|_2^2 + \gamma \|\hat{\mathbf{w}}^{i+1} - \hat{\mathbf{g}} + \hat{\mathbf{h}}\|_2^2) \quad (15)$$

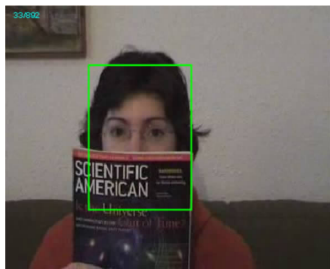
对式(15)进行求导并进行整理, 得到

$$\hat{\mathbf{g}} = \frac{\lambda \hat{\mathbf{w}}^r + \gamma(\hat{\mathbf{w}}^{i+1} + \hat{\mathbf{h}})}{\lambda + \gamma} \quad (16)$$

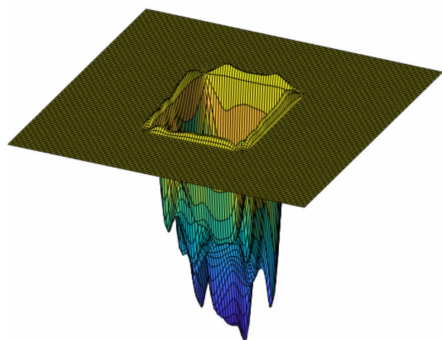
至此, $\mathbf{g}$ 的频域形式 $\hat{\mathbf{g}}$ 求解完毕。

在目标发生遮挡, 尤其是部分遮挡时, 更新空间权重可使得滤波器能够更好地学习到跟踪目标信息, 而不是过多的遮挡物信息。

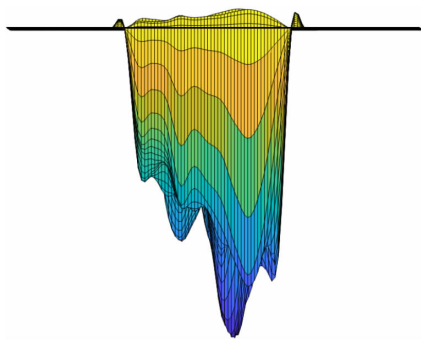
图3是 Faceoccl 序列中第33帧跟踪结果与空间权重可视化结果。在图3a中, 目标下半部分产生了遮挡, 所以空间权重应产生相应的改变。图3b为改变后的空间权重可视化结果。为能够更好地展示空间权重的变化, 将图3b旋转, 得到图3c, 原始空间权重图像已在图1中展示。在图3c中, 左边较大的值对应于目标下部分, 右边较小的值对应目标上部分, 可以看到空间权重更新后将下半部分的权重增加, 上半部分的权重降低, 即惩罚了下半部分, 更



(a) 跟踪结果



(b) 空间权重可视化结果



(c) 旋转后的空间权重可视化结果

图3 Faceoccl 序列第33帧跟踪结果与空间权重可视化结果

关注于上半部分, 使得滤波器能够更好地学习到目标信息, 避免了过量学习遮挡物特征而导致模型衰退。

在滤波器系数初始化后, 进行1次迭代求解, 将求解结果代入式(14)中求解空间权重。实验结果表明, 求解空间权重的迭代次数设置为2即可。与ASRCF不同, 本文在第一帧求解初始滤波器系数, 将目标位置空间权重设置较低, 而周围背景权重设置较高, 而ASRCF^[6]在第一帧求解初始滤波器系数时采用的是目标内外区域均为同一参数。

2.3 响应图求解

与STRCF相同, 本文依然采用牛顿迭代法求解目标最大值位置与尺度。不同点在于, 本文将尺度设置为7, 步长设置为1.01, 且在牛顿迭代法求解的过程中设置了阈值, 将前后两次迭代求解的结果做差, 与阈值相对比。如果差值小于阈值则表明结果已经满足要求, 无需再进行迭代; 如果差值大于阈值, 则继续迭代求解。最大迭代次数与STRCF相同, 设置为5次。

3 实验

测试环境为 Intel(R) Core(TM) i5-8400 CPU @3.20 GHz 的台式电脑, 操作系统为 Windows 10 64 bit, 实验软件为 Matlab 2018a。测试数据集为 OTB-2015^[19], 采用一次性测试方式进行测试, 利用成功率曲线下的积分面积 S_{AUC} 对算法进行评价。

选择 SAMF、SRDCF、ECO-HC、STRCF、ASRCF 共5种算法进行对比。由于ASRCF仅提供了加入深度特征的程序及测试结果, 为保证公平, 设定实验方案如下: ①仅采用手工特征的本文算法, 即仅采用了灰度、HOG 和 CN 这3种特征的本文算法, 与前4种算法进行对比; ②为证明本文算法对于目

标遮挡处理的有效性,选择 6 个包含遮挡及尺度变化的视频序列,采用定性与定量的方式对比前 4 种算法与仅使用手工特征的本文算法,包含遮挡及尺度变化的视频序列及其属性如表 1 所示;③在手工特征的基础上加入了深度特征的本文算法,与 AS-RCF、STRCF、ECO 这 3 种使用深度特征的算法进行对比。

表 1 包含遮挡及尺度变化的视频序列及其属性

视频序列名称	视频序列属性
Carscale	遮挡、尺度变化、快速运动、旋转
Soccer	遮挡、尺度变化、光照变化、运动模糊、快速运动、背景干扰、旋转
Woman	遮挡、尺度变化、光照变化、形变、运动模糊、快速运动、旋转
KiteSurf	遮挡、光照变化、旋转
Football	遮挡、旋转、背景干扰
Dudek	遮挡、尺度变化、形变、快速运动、超出视野、旋转、背景干扰

3.1 手工特征算法对比

全部视频序列上的成功率曲线如图 4 所示。分析可知:按照 S_{AUC} 从高到低,5 种算法分别是本文算法(0.653)、STRCF(0.649)、ECO-HC(0.631)、SRDCF(0.580)、SAMF(0.556);本文算法的 S_{AUC} 较 STRCF 的提高了 0.4%,提升幅度较小,这是因为本文算法更注重目标遮挡的处理,并且特征通道裁剪的方法对于手工特征选择效果一般。

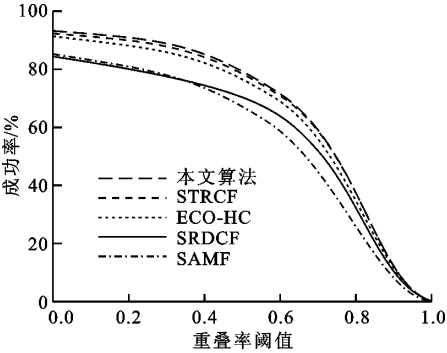


图 4 全部视频序列上的成功率曲线

具有遮挡属性的视频序列上的成功率曲线如图 5 所示。分析可知:按照 S_{AUC} 从高到低,5 种算法分别是本文算法(0.612)、STRCF(0.605)、ECO-HC(0.604)、SAMF(0.536)、SRDCF(0.522);本文算法的 S_{AUC} 较 STRCF 的提高了 0.7%;与全部视频序列的测试结果相比,本文算法在遮挡视频序列中

的提升效果更为明显,本文算法对遮挡处理的有效性得到了验证。

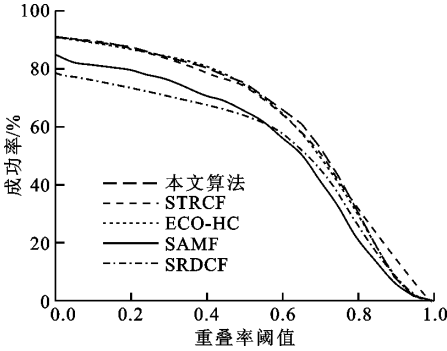


图 5 遮挡视频序列上的成功率曲线

3.2 包含遮挡及尺度变化视频序列上的手工特征法定性对比

5 种算法在 6 个视频序列上的跟踪结果如图 6 所示。

从图 6a 和图 6f 可以看出:Carscale 和 Dudek 序列中目标遮挡程度较低,尺度变化过程较慢,全部 5 种算法均能跟踪目标,但本文算法尺度估计更为准确。

从图 6b 可以看出:在 Soccer 序列的第 61 帧中,目标在不断跳动,画面模糊,此时仅有本文算法及 SRDCF 跟踪较为准确,其他算法均有较大的偏差;在第 104 帧中,目标受到大幅度的遮挡,ECO-HC 与 SRDCF 跟踪失败;此后由于目标受到持续不断的相似物体的严重遮挡,全部算法跟踪失败,但在第 128 帧中,目标受到遮挡较少时,本文算法与 STRCF 能够重新跟踪目标,而在目标重新受到相同遮挡时,如在第 180 帧中,STRCF 受到遮挡物影响,跟踪失败,但在后续序列中能够重新跟踪目标;从第 272 帧开始,仅有本文算法和 STRCF 算法能够在后续序列中成功跟踪目标。

从图 6c 可以看出:虽然 Woman 序列中的目标受到车辆、树木等小幅度的遮挡,但 5 种算法均未受到干扰;在视频序列的后期,目标尺度变化剧烈,此时 SAMF 跟踪失败,ECO-HC 与 STRCF 无法应对快速的尺度变化,尺度估计偏差较大,而本文算法和 SRDCF 尺度估计最为准确。

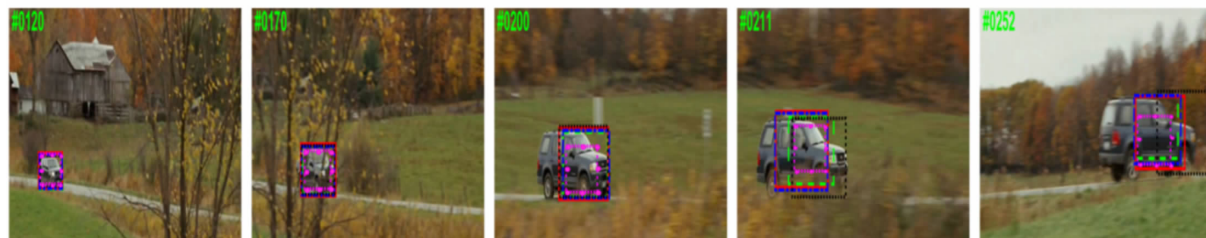
从图 6d 可以看出:在 KiteSurf 序列的第 34 帧中,目标向上跃起,产生遮挡,此时 SRDCF 跟踪偏差较为明显;在第 40 帧中,目标从遮挡物后出现,仅有本文算法能够跟踪目标,在后续序列中其他算法也均未能重新跟踪目标。

从图 6e 可以看出:在 Football 序列的第 151 帧

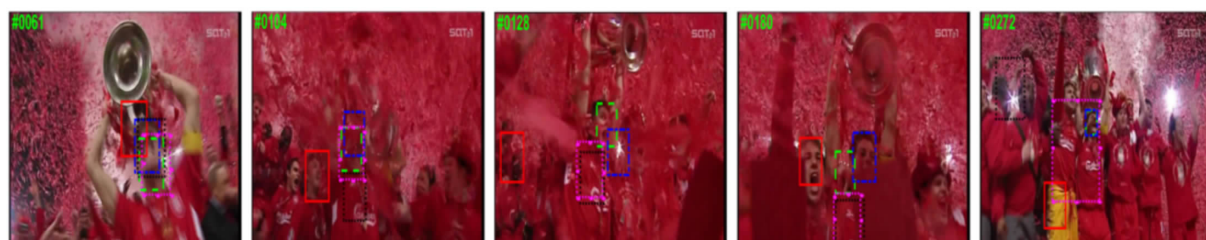
中,目标发生小幅度旋转,SAMF 的定位出现微小偏差,但在后续序列中恢复了对目标的精确跟踪;在第 183 帧中,目标旋转较为明显,受益于时间正则项,本文算法与 STRCF 的位置估计较为准确,其他算法均有偏差;在第 292 帧中,目标与相似物体发生

碰撞,相似物体对跟踪目标产生遮挡,SAMF 跟踪失败;在第 362 帧中,目标再次被相似物体部分遮挡,STRCF 与 SAMF 均跟踪失败,虽然 ECO-HC 与 SRDCF 能够跟踪目标,但跟踪结果与本文算法相比不够准确。

--- 本文算法 — ECO-HC - - - STRCF — SRDCF - - - SAMF



(a)Carscale



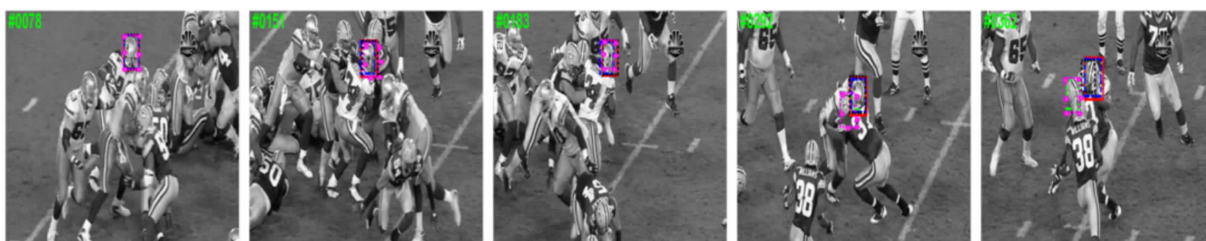
(b)Soccer



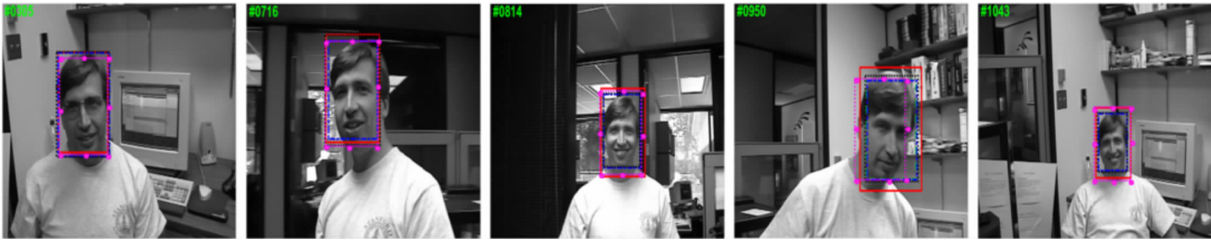
(c)Woman



(d)KiteSurf



(e)Football



(f)Dudek

图 6 5 种算法在 6 个视频序列下的跟踪结果

由于本文算法共设计了 7 个尺度,且步长较小,因此可以快速应对缓慢的尺度变化,获取更为准确的尺度估计。加入的空间权重正则项,能够在后续遮挡物不多的情况下使得滤波器准确学习跟踪目标信息,因此在重新跟踪目标后不会再跟踪失败,且在目标与遮挡物特征差距较为明显时,通道裁剪将不必要的特征通道裁剪后能够更好地表达目标信息,从而提高了跟踪精度。

3.3 包含遮挡及尺度变化视频序列上的

手工特征算法定量对比

采用中心位置误差、重叠率以及运算帧率共 3 个指标对 5 种算法进行对比。

(1)中心位置误差。中心位置误差公式^[20]为

$$D_1 = \sqrt{\|\mathbf{O}_1 - \mathbf{O}_t\|^2} \tag{17}$$

式中: $\mathbf{O}_1$ 表示跟踪算法估计的目标中心位置; $\mathbf{O}_t$ 表示由数据集提供的真实目标中心位置。中心位置误差越小表示算法性能越好。

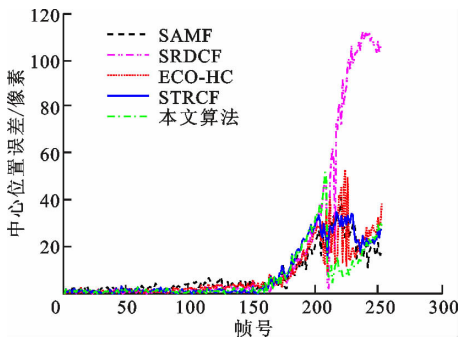
5 种算法在 6 个视频序列下的中心位置误差对比如图 7 所示。可以看出,本文算法的中心位置误差在 6 个视频序列下均是较小的。为更为直观地对比,计算了 5 种算法在 6 个视频序列中的平均中心位置误差,结果如表 2 所示。

表 2 平均中心位置误差对比

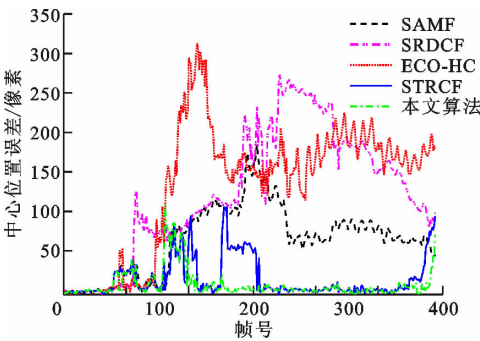
序列名称	平均中心位置误差/像素				
	SAMF	SRD-CF	ECO-HC	STR-CF	本文算法
Carscale	8.41	19.68	7.66	8.65	6.99
Soccer	68.48	126.28	136.12	19.56	13.50
Woman	10.79	4.02	2.96	2.50	2.17
KiteSurf	59.26	59.18	63.83	66.73	3.36
Football	13.37	5.88	5.18	6.08	4.24
Dudek	8.81	12.30	11.08	10.91	10.39

从表 2 可以看出,除了在 Dudek 序列中,本文算法的平均中心位置误差略大于 SAMF 的,在其他

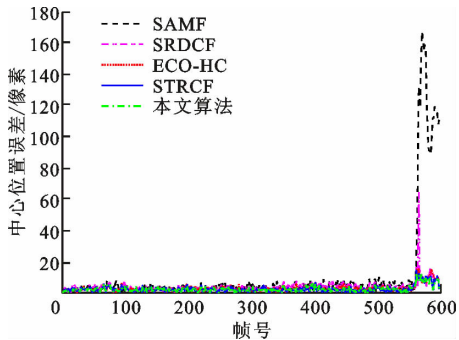
视频序列中,本文算法均获得了最小的平均中心位置误差。但是,即使在 Dudek 序列中,本文算法较基础算法 STRCF 依然有所提升。在全部视频序列中,本文算法较 STRCF 在平均中心位置误差上降低了 12.3 像素,较 SAMF 降低了 21.41 像素。



(a)Carscale



(b)Soccer



(c)Woman

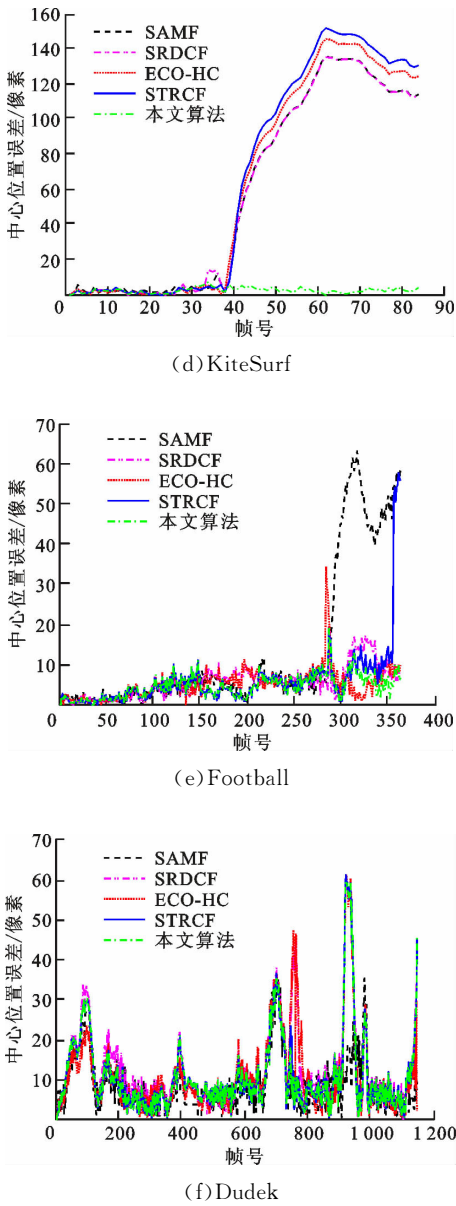


图 7 6 个视频序列的中心位置误差对比

(2)重叠率。重叠率的公式^[20]为

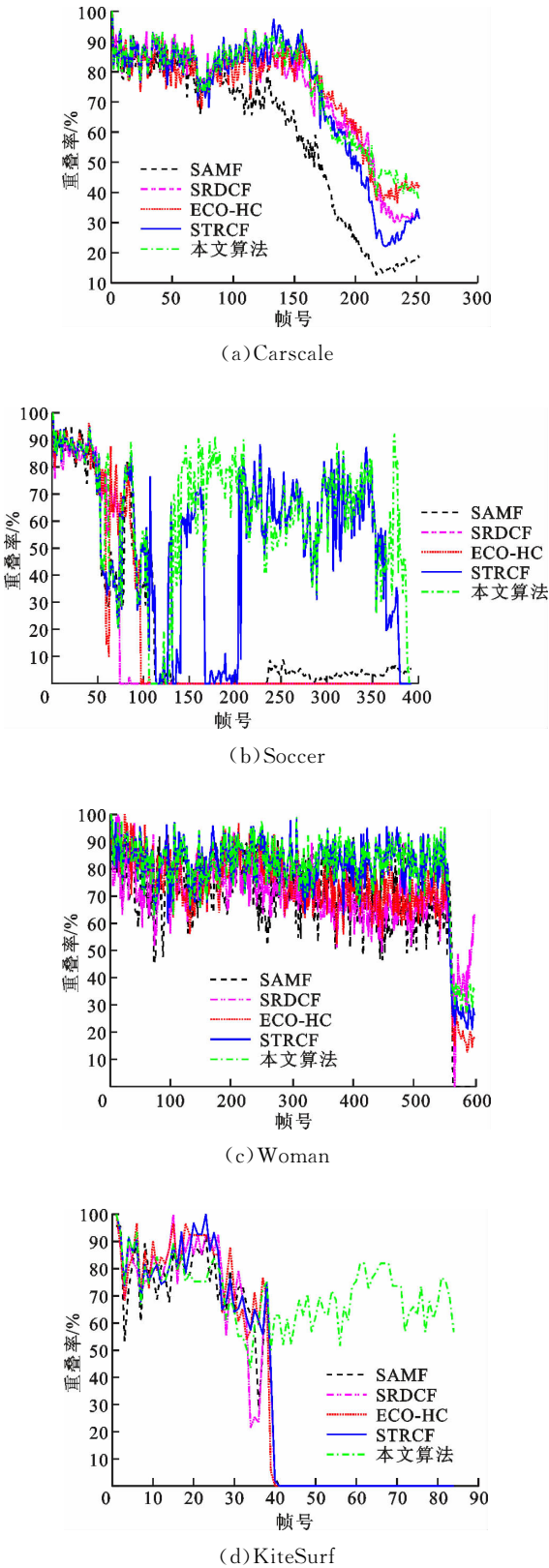
$$L = \frac{B_1 \cap B_0}{B_1 \cup B_0} \tag{18}$$

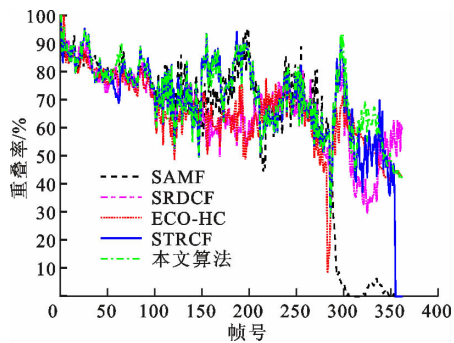
式中： B_1 为算法估计的目标区域； B_0 为数据集提供的目标实际覆盖区域，重叠率越高表示算法性能越好。5 种算法在 6 个视频序列下的重叠率对比如图 8 所示。

从图 8 可以看出，本文算法在 6 个视频序列中的重叠率在大部分时间内高于其他 4 种算法的。为更直观地对比，计算了 5 种算法在 6 个视频序列下的平均重叠率，结果如表 3 所示。

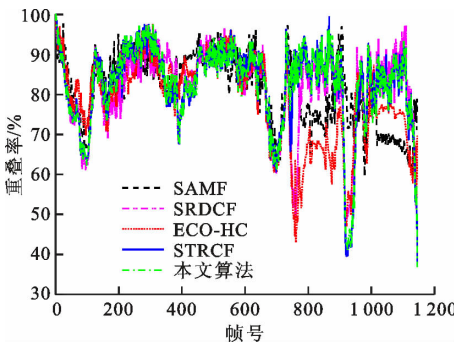
从表 3 可以看出，本文算法在全部 6 个视频序列中获得了最高的平均重叠率，仅在 Dudek 序列下本文算法和 STRCF 获得了相同的平均重叠率。本

文算法的总平均重叠率较 STRCF 的增加了 7%。在 Carscale 和 Woman 序列中，本文算法与 STRCF 均能成功跟踪目标，但由于目标尺度变化较大，所以本文算法的平均重叠率提升幅度较大。





(e)Football



(f)Dudek

图 8 6 个视频序列的重叠率对比

表 3 平均重叠率对比

序列名称	平均重叠率/%				
	SAMF	SRD-CF	ECO-HC	STR-CF	本文算法
Carscale	58	72	73	71	75
Soccer	20	15	18	51	58
Woman	67	70	72	79	82
KiteSurf	34	35	37	36	67
Football	60	66	66	70	71
Dudek	81	82	79	83	83

(3)运算帧率。运算帧率公式为

$$v = \frac{s}{t}$$

(19)

式中: s 表示视频总帧数; t 表示算法在此视频序列上的运算时间。运算帧率结果如表 4 所示。

从表 4 可以看出,ECO-HC 在全部视频序列中运行速率最高,本文算法较 STRCF 仅在 Carscale 序列中获得了更高的运算帧率。这是因为本文采用通道裁剪以及在牛顿迭代法中设置阈值来降低迭代次数,从而提升算法运算速率,但是由于跟踪目标以及所处背景不同,通道裁剪后的通道数并不一致,且不同帧中实际迭代次数也不相同,因此本文算法仅

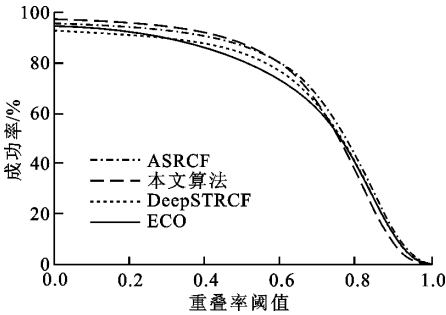
在部分序列中获得比 STRCF 更好或相似的运算帧率,尤其在 KiteSurf 序列中,本文算法的运算帧率下降较大。在 6 个序列中,本文算法的平均运算帧率为 19.25 帧/s,但其他指标均优于 STRCF 的。

表 4 运算帧率对比

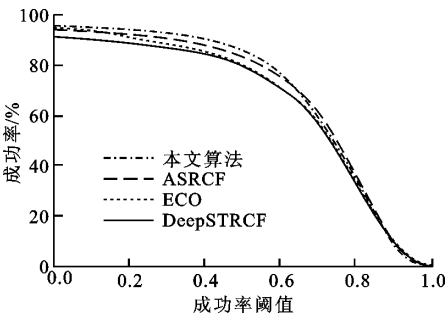
序列名称	运算帧率/(帧·s ⁻¹)				
	SAMF	SRD-CF	ECO-HC	STR-CF	本文算法
Carscale	19.36	6.75	45.70	11.21	18.40
Soccer	11.83	6.02	47.18	15.96	15.04
Woman	27.55	8.46	62.54	19.75	15.34
KiteSurf	46.49	12.22	62.68	34.82	13.85
Football	40.67	7.93	65.54	44.96	30.54
Dudek	29.19	6.21	59.80	29.49	22.35

3.4 深度特征算法对比

由于本文算法是在 STRCF 的基础上改进而来的,因此与 STRCF 相同,采用 VGG-M 网络的第一层(Conv-1)与最后一层(Conv-5)卷积层的输出作为算法的深度特征,而 ASRCF 采用的则是 VGG-M 中的 Conv-1 与 VGG-16 的 Conv4-3 的输出作为深度特征。VGG-16 的 Conv4-3 较 VGG-M 的 Conv-5 而言,网络深度更深,包含更高级的语义信息。本文在加入深度特征后与 ASRCF、ECO、DeepSTRCF 在 OTB-2015 数据集进行对比,结果如图 9 所示。



(a)数据集整体测试结果



(b)遮挡视频序列测试结果

图 9 深度特征算法对比

从图 9a 可以看出:按照 S_{AUC} 从高到低,4 种算法分别是 ASRCF(0.696)、本文算法(0.692)、DeepSTRCF(0.669)、ECO(0.665);使用更深层次的语义信息特征的 ASRCF 表现更佳,其 S_{AUC} 较本文算法的增加了 0.4%。从图 9b 可以看出:按照 S_{AUC} 从高到低,4 种算法分别是本文算法(0.679)、ASRCF(0.670)、ECO(0.654)、DeepSTRCF(0.640);本文算法的 S_{AUC} 较 ASRCF 的提高了 0.9%。这是因为 ASRCF 虽然使用了更深层次的语义信息特征,但在目标产生遮挡时,模型受到污染,产生漂移,导致跟踪精度下降。本文算法采用平均特征能量比裁减了不能准确表达目标的特征,尤其在使用深度特征时,这种裁剪更为明显,Conv-5 输出的特征第 3 维度为 512,而经裁剪后维度通常会下降至 100 以下,与通道裁剪的视觉目标跟踪^[18]结论相同。深度特征的维度较大,但实际能够真实反映目标特征的特征通道将会提高目标在定位时的精度。在遮挡的情况下,由于遮挡物与目标特征具有一定区别,因此保留能够准确表达目标的特征能够更好地从遮挡中恢复对目标的跟踪。本文算法较 ASRCF 加入了时间正则项,能够使模型在目标产生遮挡后受到较小的影响,及时恢复对目标的跟踪。

由于深度特征对目标的表达能力较高,所以在选择的测试视频序列中,除 DeepSTRCF 无法在 Soccer 序列中成功跟踪目标外,其他算法均可在全部视频序列中准确跟踪目标。深度特征算法在选择的测试视频序列中表现相差不大,因此不再做单独视频序列测试。

4 结 论

由于 STRCF 在目标遮挡严重时会跟踪失败,且尺度估计同样会受到影响,因此本文提出了结合自适应空间权重的改进型时空正则项跟踪算法。本文获得的结论如下。

(1)在视频序列第一帧,对样本块进行提取,并计算不同特征通道的平均能量比,如能量比较低则表明该通道无法准确表示目标,将其裁剪可提高目标定位精度。

(2)在相关滤波器训练公式中加入空间权重正则项,在目标被部分遮挡时可自适应调整滤波器学习的重点区域,提高目标在部分遮挡时的跟踪成功率。由于时间正则项在目标遮挡时被动更新的方式依然能够准确表示目标,因此在更新空间权重时准

确率更高。

(3)采用自适应牛顿迭代法可减少求解目标位置时的迭代次数,增加算法运算速率。为证明算法性能,分别使用了手工特征和手工特征+深度特征在 OTB 数据集中进行测试,结果表明,本文算法可在目标遮挡时保证跟踪精确度以及尺度估计的准确性。

参考文献:

- [1] 王旭东,王屹炜,闫贺.背景抑制直方图模型的连续自适应均值漂移跟踪算法[J].电子与信息学报,2019,41(6):1480-1487.
WANG Xudong, WANG Yiwei, YAN He. Continuously adaptive mean-shift tracking algorithm with suppressed background histogram mode[J]. Journal of Electronics & Information Technology, 2019, 41(6): 1480-1487.
- [2] 张博,江沸波,刘刚.利用感知模型的长期目标跟踪[J].中国图象图形学报,2019,24(11):1906-1917.
ZHANG Bo, JIANG Feibo, LIU Gang. Long-term target tracking based on perceptual model[J]. Journal of Image and Graphics, 2019, 24(11): 1906-1917.
- [3] LI H M, SHI L P. Robust event-based object tracking combining correlation filter and CNN representation[J]. Frontiers in Neurorobotics, 2019, 13: 00082.
- [4] 胡云层,路红,杨晨,等.融合 DSST 和 KCF 的尺度自适应跟踪算法[J].计算机工程与设计,2019,40(12):3563-3568.
HU Yunceng, LU Hong, YANG Chen, et al. Scale adaptive tracking algorithm combining DSST and KCF[J]. Computer Engineering and Design, 2019, 40(11): 3563-3568.
- [5] MEMARMOGHADAM A, MOALLEM P. Size-aware visual object tracking via dynamic fusion of correlation filter-based part regressors[J]. Signal Processing, 2019, 164: 84-98.
- [6] ISWANTO I A, CHOA T W, LI B. Object tracking based on meanshift and particle-Kalman filter algorithm with multi features[J]. Procedia Computer Science, 2019, 157: 521-529.
- [7] 李牧子.基于图像识别的目标检测与跟踪系统的设计与实现[D].北京:北京邮电大学,2019:40-41.
- [8] 杨源,库涛,查宇飞,等.快速多特征金字塔的尺度目标跟踪方法[J].西安交通大学学报,2016,50(10):49-56.
YANG Yuan, KU Tao, ZHA Yufei, et al. Fast multi-feature pyramid scale target tracking method[J]. Journal of Xi'an Jiaotong University, 2016, 50(10):

- 49-56.
- [9] LI Y, ZHU J K. A scale adaptive kernel correlation filter tracker with feature integration [C]//Proceedings of the 13th European Conference on Computer Vision. Berlin, Germany: Springer, 2015: 254-265.
- [10] DANELLJAN M, HAGER G, KHAN F S, et al. Learning spatially regularized correlation filters for visual tracking [C]//Proceedings of the 2015 IEEE International Conference on Computer Vision. Piscataway, NJ, USA: IEEE, 2015: 4310-4318.
- [11] DANELLJAN M, HAGER G, KHAN F S, et al. Convolutional features for correlation filter based visual tracking [C]//Proceedings of the 2015 IEEE International Conference on Computer Vision Workshop. Piscataway, NJ, USA: IEEE, 2015: 7406433.
- [12] DANELLJAN M, BHAT G, KHAN F S, et al. ECO: efficient convolution operators for tracking [C]//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2017: 6931-6939.
- [13] LI F, TIAN C, ZUO W M, et al. Learning spatial-temporal regularized correlation filters for visual tracking [C]//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2018: 4904-4913.
- [14] DAI Kenan, WANG Dong, LU Huchuan, et al. Visual tracking via adaptive spatially-regularized correlation filters [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2019: 4670-4679.
- [15] RAMDAS A, TIBSHIRANI R J. Fast and flexible ADMM algorithms for trend filtering [J]. Journal of Computational and Graphical Statistics, 2016, 25(3): 839-858.
- [16] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection [C]//Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE: 886-893.
- [17] DANELLJAN M, KHAN F S, FELSBERG M, et al. Adaptive color attributes for real-time visual tracking [C]//Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2014: 1090-1097.
- [18] CHE M Q, WANG R L, LU Y, et al. Channel pruning for visual tracking [C]//Proceeding of the 15th European Conference on Computer Vision. Berlin, Germany: Springer, 2019: 70-82.
- [19] WU Y, LIM J, YANG M H. Object tracking benchmark [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1834-1848.
- [20] WU Y, LIM J, YANG M H. Online object tracking: a benchmark [C]//Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2013: 2411-2418.
- [本刊相关文献链接]
- 吕明明,侯润民,柯于峰,等.光电跟踪平台脱靶量滞后补偿方法.2019,53(11):141-147.[doi:10.7652/xjtuxb201911020]
- 廖加文,齐春,曹剑中,等.结合自适应特征选择和蕨类分类器的相关滤波跟踪算法.2019,53(6):101-108.[doi:10.7652/xjtuxb201906014]
- 侯利明,连峰,王伟. δ -广义标记多伯努利滤波器的非线性应用扩展.2019,53(6):109-116.[doi:10.7652/xjtuxb201906015]
- 王树亮,毕大平,张奎,等.认知雷达波形和检测门限自适应跟踪算法.2019,53(4):65-71.[doi:10.7652/xjtuxb201904010]
- 钟德星,杨元,刘瑞玲,等.基于单目视觉的装配机器人研究及应用.2018,52(5):81-87.[doi:10.7652/xjtuxb201805012]
- 彭华甫,黄高明,田威,等.标签多伯努利机动目标跟踪与分类算法.2019,53(2):157-162.[doi:10.7652/xjtuxb201902021]
- 宋喜玉,任修坤,郑娜娥,等.多目标跟踪下的分布式MIMO雷达资源联合优化算法.2018,52(10):110-115.[doi:10.7652/xjtuxb201810015]
- 杜明洋,毕大平,王树亮,等.一种混合的机动群目标分离检测跟踪算法.2018,52(10):116-123.[doi:10.7652/xjtuxb201810016]
- 隗寒冰,陈尧,贾志杰,等.融合历史轨迹的智能汽车城市复杂环境多目标检测与跟踪算法.2018,52(10):132-140.[doi:10.7652/xjtuxb201810018]
- 初红霞,谢忠玉,王科俊.一种结合颜色纹理直方图的改进型Camshift目标跟踪算法.2018,52(3):145-152.[doi:10.7652/xjtuxb201803020]
- 苗雨,宋骥平,姬红兵.箱粒子广义标签多伯努利滤波的目标跟踪算法.2017,51(10):107-112.[doi:10.7652/xjtuxb201710018]
- 庞策,黄树彩,刘锦昌,等.多传感器交叉提示技术在传感器联盟中的应用.2017,51(7):148-155.[doi:10.7652/xjtuxb201707021]

(编辑 陶晴)