

文章编号:1007-2780(2023)08-1118-10

基于改进 ECO-HC 的目标跟踪方法

刘刚阳^{1,2*}, 胡 博³, 王宇庆¹

(1. 中国科学院 长春光学精密机械与物理研究所, 吉林 长春 130033;

2. 中国科学院大学, 北京 100049;

3. 中国人民解放军第 32035 部队, 陕西 西安 710000)

摘要:基于相关滤波的跟踪方法帧率高但容易发生跟踪漂移导致目标丢失。针对此问题,本文以相关滤波方法 ECO-HC 为基线,提出了一个用于描述图像场景变化幅度的参数 D ,根据实时计算该参数的值并以不同的间隔数更新模型,实现自适应更新模型在使算法以较高帧率运行的同时兼具良好的准确率,此效果在图像光照变化复杂、目标变形或被遮挡的情况下更加明显。实验结果表明,ECO-HC 嵌入本文方法后,在光照变化的情况下成功率提高了 1.6%,准确率提高了 2.2%;在目标离开视场的情况下成功率提高了 3.7%,准确率提高了 3.4%。在本文实验环境中平均帧率达到 60 FPS。

关键词:目标跟踪;相关滤波;相关度;模型更新策略

中图分类号:TP391.4 文献标识码:A doi:10.37188/CJLCD.2022-0290

Target tracking method based on improved ECO-HC

LIU Gang-yang^{1,2*}, HU Bo³, WANG Yu-qing¹

(1. Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences,
Changchun 130033, China;

2. University of the Chinese Academy of Sciences, Beijing 100049, China;

3. The 32035th Unit of the Chinese People's Liberation Army, Xi'an 710000, China)

Abstract: The tracking method based on correlation filtering has fast frame rate but is prone to tracking drift, resulting in target loss. Aiming to this problem, taking the correlation filtering method ECO-HC as the baseline, this paper proposes parameter D to describe the change amplitude of the image scene. According to real-time calculation of the value for this parameter and updating the model with different intervals, the adaptive updating model is realized, so that the algorithm runs at a higher frame rate and has a good accuracy. This effect is more obvious in the case of complex image illumination changing, target deformation or occlusion. Experimental results show that success rate and accuracy rate of ECO-HC embedded in this method are improved by 1.6% and 2.2% under the condition of changing illumination. When the target leaves the field of view, the success rate is improved by 3.7%, and the accuracy is improved by 3.4%. In this experiment environment, the average frame rate reaches 60 FPS.

Key words: target tracking; correlation filtering; correlation; model update strategy

收稿日期:2022-09-01;修订日期:2022-10-25.

*通信联系人, E-mail: liugangyang20@mails.ucas.ac.cn

1 引言

目标跟踪是计算机视觉领域的一个重要研究课题。视觉目标跟踪一般是给定初始帧图像中目标的位置和矩形框尺寸,利用程序自动搜寻以后帧中目标的位置和矩形框尺寸。目标跟踪技术应用广泛,其在智能交通、自动驾驶、安防监控、机器人、军事等众多领域内已不可或缺^[1]。目标跟踪根据目标个数不同可分为单目标跟踪和多目标跟踪,根据其技术发展可分为经典的跟踪方法、基于相关滤波的跟踪方法和基于深度学习的跟踪方法^[2]。近些年比较热门的目标跟踪方法主要是基于相关滤波的算法和基于深度学习的算法,两者都属于判别式跟踪方法。MOSSE方法^[3]首次在跟踪算法中引入相关滤波的概念,使运算速度提高了20倍。随着深度学习的快速发展,应用了深度学习的目标跟踪技术取得了前所未有的成功,它和基于相关滤波的跟踪方法各有优劣。基于相关滤波的跟踪方法对硬件平台要求较低且运行速度快,而基于深度学习的跟踪方法运算复杂度高、消耗内存大。基于深度学习的目标跟踪多为基于检测的跟踪技术,其跟踪性能依赖检测阶段的准确性。由于卷积神经网络在计算过程中参数较多,为提高运算速度,算法需运行在性能较好的GPU平台上。一般地,单目标跟踪青睐基于相关滤波的方法,多目标跟踪则使用深度学习能获得更佳的指标。

能否有效跟踪的一个关键因素是提取恰当数量且具有鉴别性的特征^[4-5]。深度学习使用卷积神经网络提取深度特征,相关滤波提取方向梯度直方图(HOG)、灰度或纹理等作为特征,相关滤波必须考虑到模型更新的问题^[6]。在跟踪过程中往往出现目标变形或被遮挡等情况,跟踪模型会随之退化,导致跟踪漂移^[7]。以往大多数算法均采用连续更新策略,即在每一帧中都要进行模型更新,而后来有人使用Siamese^[8]网络没有进行模型更新,仍然取得了不错的效果。MEEM方法^[9]根据最小熵准则选择在需要时恢复当前跟踪器以减少不必要的模型更新,利用预测算法中的在线SVM和显式特征映射方法实现了高效的模型更新和推理。Nam^[10]等人用预训练卷积神经网络中的共享层和在线更新的新二元分类层相结

合构建网络,通过评估围绕先前目标状态随机采样的候选窗口执行在线更新。SRDCFdecon方法^[11]引入样本权重对作用较小的样本进行抑制。ECO方法^[12](Efficient Convolution Operators for Tracking)采用了每隔5帧更新一次模型的稀疏更新方法,效果显著,作者把这种效果归因于稀疏更新减少了与最新样本的过拟合。

本文基于相关滤波算法ECO的手工特征版本ECO-HC提出了一种以图像场景变化为依据、自适应调节模型更新间隔的策略,有效提高了算法在图像背景光照复杂、目标外形改变或被遮挡等情况下的跟踪准确率,并保持了较高帧率的优点。

2 用于跟踪的高效卷积算子

C-COT^[13]是一种基于判别相关滤波器(Discriminative Correlation Filter)的目标跟踪方法,该算法提出了一种连续卷积算子的理论框架,将特征信息从离散空间变换到连续空间,如公式(1)所示:

$$J_d\{x^d\}(t) = \sum_{n=0}^{N_d-1} x^d[n] b_d\left(t - \frac{T}{N_d}n\right), \quad (1)$$

其中: $J_d\{x^d\}(t)$ 表示经插值计算后的第 d 个通道的值, b_d 是一个周期为 T 的3次插值函数。根据此公式,既可以通过在连续域上执行卷积实现不同分辨率下特征的独立表达,又能利用连续函数得到目标预测的分数,实现更准确的亚像素定位。

ECO在C-COT的基础上提出了因式分解卷积算子(Factorized Convolution Operator)和紧凑的生成式样本空间模型(Generative Sample Space Model),并改进了模型更新策略(Model Update Strategy)。因式分解卷积算子通过一个 d 行 c 列的系数矩阵 P 将原本C-COT的 d 个滤波器缩减为 c 个, $c < d$,达到了减少模型参数的目的,如公式(2)所示:

$$S_{PF}\{x\} = Pf * J\{x\} = \sum_{c,d} p_{d,c} f^c * J_d\{x^d\} = f * P^T J\{x\}, \quad (2)$$

其中: f 表示滤波器, $J\{x\}$ 为特征化后的图像, $J_d\{x^d\}$ 为第 d 维滤波器, P^T 为 P 的转置矩阵。

因为相邻帧图像中目标大概率外形相似,所以连续每帧都存储新样本会导致样本数据冗余。紧凑的生成式样本空间模型不再使用原本的连

续训练集,而是将训练数据用混合高斯模型重新建模分类,其各个分量分别表示一类特征与其他样本不同的样本数据,可以提高样本多样性并减少数据冗余。

大多数基于相关滤波的目标跟踪方法都是逐帧更新模型,而 ECO 每隔 5 帧更新一次,大幅减少了计算量,提高了跟踪速度,并发现降低滤波器更新频率可以提高学习过程的鲁棒性。

3 算法改进

本文尝试使算法自适应确定模型更新间隔并以不等间隔更新模型,使算法在确保运算速度的同时提高跟踪准确率。在确定具体的模型更新间隔数前,需要先计算出相邻两帧图像的差异,并以此作为判断图像中场景变化幅度的依据。我们定义了一个参数 D ,用它来表示相邻两帧图像间的变化幅度。有必要说明的是,并不是每两帧图像都要计算一次差异值。如果将模型更新间隔数记为 train_gap ,那么模型将是每间隔 train_gap 帧更新一次,即平均 $\text{train_gap}+1$ 帧更新一次。

3.1 相邻帧图像相关度计算

先将当前帧彩色图像转为灰阶图像,然后对像素值做归一化处理,把 m 行 n 列的图像转化成一个 $m \times n$ 维的列向量 X_1 ,上一帧图像也如此操作得到 X_2 。合并 X_1 和 X_2 ,即:

$$X = [X_1, X_2]. \quad (3)$$

分别计算 X 各列元素的均值,即:

$$\bar{X} = [\bar{X}_1, \bar{X}_2]. \quad (4)$$

计算 X_1 与 X_2 的协方差:

$$\text{cov}(X) = \frac{(X - \bar{X})^T (X - \bar{X})}{m \times n - 1}, \quad (5)$$

其中 $\bar{X}_c$ 每行都等于 $\bar{X}$ 的 $m \times n$ 维矩阵。结果 $\text{cov}(X)$ 是一个 2 行 2 列的矩阵,若记:

$$\text{cov}(X) = \begin{bmatrix} a & b \\ c & d \end{bmatrix}. \quad (6)$$

单个样本序列的样本方差为:

$$a = \frac{1}{m \times n - 1} \sum_{i=1}^{m \times n} (x_{1i} - \bar{x}_1)^2, \quad (7)$$

$$d = \frac{1}{m \times n - 1} \sum_{i=1}^{m \times n} (x_{2i} - \bar{x}_2)^2. \quad (8)$$

则令相关度为:

$$\text{correlation} = \frac{b + c}{2\sqrt{ad}}. \quad (9)$$

当 X 列数为 2 时, $\text{cov}(X)$ 的副对角元素 b 和 c 相等,则:

$$\text{correlation} = \frac{b}{\sqrt{ad}}. \quad (10)$$

3.2 基于 HOG 计算两图像中目标区域的相关度

方向梯度直方图(HOG)能简化图像并表达出图像中的物体轮廓信息。我们根据第一帧图像中给出的目标框预测下一帧的目标框位置和大小,在每次需要更新模型时不只是计算全图的相关度,还需要利用图像中目标框内的 HOG 计算基于 HOG 的相关度,避免了背景信息的干扰。其中根据 HOG 计算出的特征向量可近似看作是目标的特征信息,过程如图 1 所示。

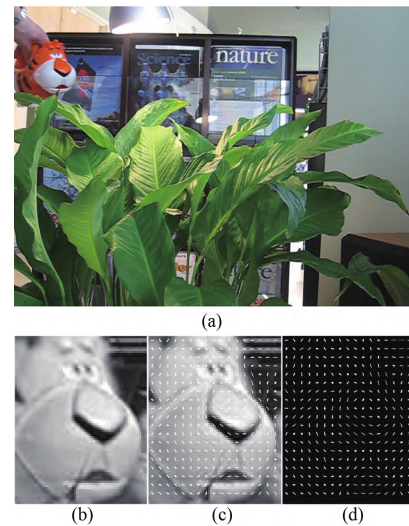


图 1 (a) Tiger2 (OTB2015) 的一张原图; (b) 截取目标所在区域并将其灰度化; (c) 计算 HOG 特征; (d) 显示梯度直方图。

Fig. 1 (a) An original image of Tiger2 (OTB2015); (b) Capturing the target area and graying it; (c) Calculating of the HOG feature; (d) Display of the gradient histogram.

当前帧与上一帧图像的 HOG 特征向量分别是 H_1 和 H_2 , 根据上一节求相关度的方法计算 H_1 与 H_2 的相关度, 可以更好地表达相邻两帧图像中目标的变化情况。

3.3 场景变化联合表示

用计算得到的相邻两图像的相关度 correlation_1 和基于 HOG 计算的目标区域相关度

correlation_2 联合表示场景变化情况,记为 D ,如公式(11)所示:

$$D = \lambda \text{correlation}_1 + (1 - \lambda) \text{correlation}_2, \quad (11)$$

其中 λ 为权重系数。不能单一使用 correlation_2 作为判断场景变化的原因是跟踪过程中可能发生目标跟踪错误或被遮挡等情况。

3.4 自适应模型更新策略

将计算出的场景变化幅度 D 作为依据,选择不同的模型更新间隔数 train_gap ,替代大多数相关滤波算法逐帧更新或等间隔更新的方法,形成一种以判断图像场景变化幅度来自适应调整模型更新间隔的模型更新方法,可以在较少损失计算速度的情况下有效提高跟踪准确率且能避免因多次更新导致的过拟合发生。

我们选择适当的阈值 threshold1 和 threshold2 ,确定最终的模型更新间隔数方案。其中第一帧时的模型是必须要更新的,第二帧时开始计算 D ,将结果与两阈值做比较,确定模型更新的间隔数 train_gap ,然后在当前帧 seq.frame 的基础上,停止更新 train_gap 次,第 $\text{seq.frame} + \text{train_gap} + 1$ 时再次更新并计算新的 D ,确定新的模型更新间隔数,过程如图2所示。

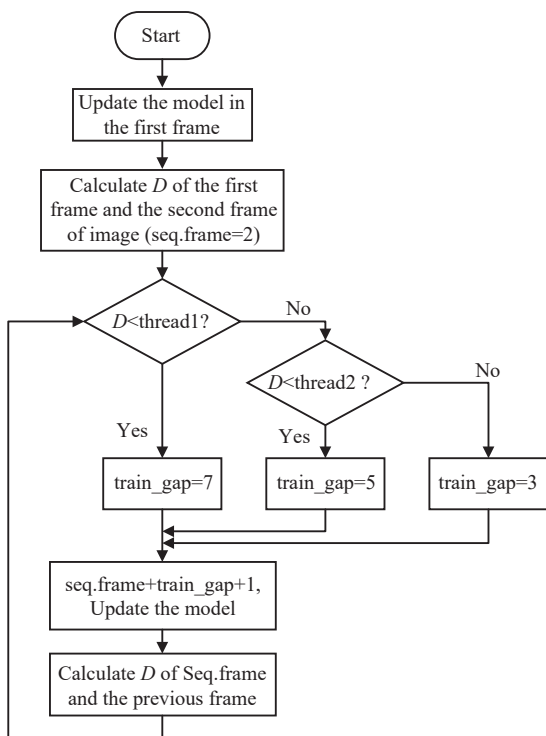


图2 本文策略流程图

Fig. 2 Strategy flow chart of this paper

4 实验与结果分析

4.1 评价指标

评价单目标跟踪算法的指标有很多种,本文主要用到了跟踪准确率(Precision)、成功率(Success rate)、AUC(Area under curve)、重叠率(Overlap)、平均重叠期望(Expected Average Overlap, EAO)、速度指标(Equivalent Filter Operations, EFO)和失败次数(Failures),这些指标中失败次数和跟踪准确率呈负相关,其余均为正相关。

准确率计算公式为:

$$\text{Precision} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}, \quad (12)$$

即计算预测框中心点与真值中心点间的欧氏距离,若该值小于阈值(一般设为20个像素)则跟踪成功。

计算成功率需先计算预测框与真值框的交并比(IoU),若该值大于阈值(一般设为0.5)则称在该帧图像上跟踪成功,在整个视频序列上跟踪成功的帧数除以视频全部帧数即得该序列上的成功率。阈值与成功率的大小呈负相关,设置阈值范围为0~1,画出阈值与成功率的曲线,曲线下面积即为AUC(Area under curve)分数,它也用来反映跟踪成功效果。

重叠率的计算方法为重叠区域面积/(预测框的面积+真值框的面积-重叠区域面积),其值等于两矩形的交并比。

平均重叠期望是VOT^[14]中最主要的指标之一,它的计算过程如下:先将所有视频序列按长度分类,再将跟踪器在长度为 N_s 的序列上测试求其每一帧的精度 $\Phi(t)$,计算平均值得到该序列的精度:

$$\Phi_{N_s} = \frac{1}{N_s} \sum_{i=1}^{N_s} \Phi(t). \quad (13)$$

计算全部序列的平均值即得该跟踪器在长度为 N_s 序列上的平均重叠期望(EAO)值:

$$\text{EAO} = \frac{1}{N_s} \sum_{i=1}^{N_s} \Phi_{N_s}. \quad (14)$$

速度指标(EFO)的全称为等效滤波操作(Equivalent Filter Operations)。先测试当前机器用 30×30 的滤波器在 600×600 图像上的滤波时间,再将实际跟踪时间除此滤波时间即得EFO值,它相对常用的速度指标帧率(FPS)而言减少了硬件差异对数值造成的差异,对算法速度评价更具普适性。

当预测框与真值框的重叠率为0时认为在当前帧中目标跟踪丢失,在跟踪丢失发生后的第5帧开始对跟踪器重新初始化。定义第*i*个跟踪器在第*k*次重复中跟踪丢失的次数为*F(i, k)*,则其平均失败次数如公式(15)所示,*N_{rep}*为重复次数:

$$\text{Failures} = \frac{1}{N_{\text{rep}}} \sum_{k=1}^{N_{\text{rep}}} F(i, k).$$

(15)

失败次数(Failures)反映的是跟踪器的鲁棒性,其值越大,跟踪器鲁棒性越差。

4.2 实验详解

本文实验硬件平台为 Intel(R)Core(TM)i7-9 700K CPU@3.60 GHz、32.0 GB RAM,运行环

境为 Windows 10+MATLAB 2018a。

我们统计了 OTB2015^[15]中所有视频序列的相邻两帧图像的相关度,并对每个序列结果各自求平均以作为该序列的平均相关度。类似地,我们依据真值将各图像中的目标区域裁剪出并计算其 HOG,进而求出基于 HOG 的相邻帧目标区域的平均相关度。系数λ为一常数,它负责调节全图的相关度和基于 HOG 的目标所在区域的相关度间的关系,在本实验中将其赋值 0.5,即在本文中简单认为计算全图的变化幅度和仅计算目标的变化幅度同等重要,由公式(11)计算联合相关度,如图 3 所示,部分结果如表 1 所示。

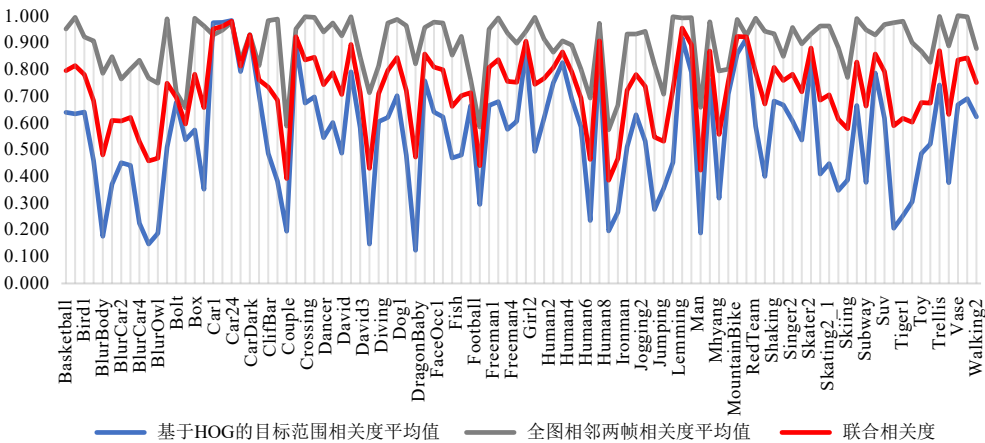


图3 OTB2015各视频序列相邻帧的相关度统计

Fig. 3 Correlation statistics of adjacent frames of each video sequence in OTB2015

表 1 相邻帧的相关度统计表

Tab. 1 Correlation statistics of adjacent frames

视频序列	平均相关度	基于 HOG 的目标 区域平均相关度	联合 相关度
Matrix	0.188	0.657	0.423
Soccer	0.385	0.767	0.576
Shaking	0.399	0.939	0.669
Ironman	0.266	0.664	0.465

将 100 个序列按联合相关度从小到大排序,取第 17 个序列的值 0.588 作为 threshold1,第 84 个序列的值 0.855 作为 threshold2,对数据集按 1:4:1 的比例切分。

ECO 根据其提取特征的方式可以分为两个版本:一个利用卷积神经网络结合 HOG 和颜色特征;另一个仅用到 HOG 与颜色特征,即 ECO-

HC。本文以 ECO-HC 作为基线,嵌入本文提出的策略进行实验,设置牛顿迭代和共轭迭代数均等于模型更新间隔数,搜索区域缩放比设为 4.5,最小和最大样本尺寸分别设为 200×200 和 250×250,共轭方向遗忘率为 75。本文模型更新策略是从第一帧图像输入即开始生效,而 ECO-HC 从第 10 帧输入后才开始每隔 5 帧更新一次。

以 OTB2015 中的 Biker 为例,它由 142 帧连续的图像组成,图 4 为本文算法对 Biker 进行实验时的模型更新间隔与帧号的变化图。从图 4 可以看出,应用本文方法后,跟踪过程共更新模型 25 次,改进前是 31 次。过程中有 3 段更新间隔为 7,6,段间隔为 3,其余和改进前一致,均为 5。图 5(a)和图 5(c)分别是对 Biker 跟踪的第 13 帧和第 100 帧图像,在这些时刻,目标缓慢移动且背景不变,场景变化幅度较低,模型更新间隔较大。

图5(b)为第70帧的跟踪效果,目标运动姿态变化较大,此时基于HOG的目标区域相关度减小导致 D 值变小,故减小模型更新间隔为3,使跟踪器能更快适应目标的改变,提高跟踪效果。再结合图6可以发现,场景变化幅度指标 D 随帧号变化的规律与模型更新间隔数随帧号变化的规律(图4)很相似,同升同降。

OTB2015数据集中jogging和Skating2各有两个目标,在实验前需将其拆分。David、BlurCar1、BlurCar3和BlurCar4的真值不是从第一帧开始的,Board的帧号编码是5位,其他序列都是4位。图7为ECO-HC和在其基础上应用了本文方法(OURS)在OTB2013^[16]数据集上与其他8种算法ASLA^[17]、DFT^[18]、SCM^[19]、L1APG^[20]、IVT^[21]、MTT^[22]、CT^[23]、ORIA^[24]的部分对比结果。本文方法相比于原方法,准确率提高了1.5%,在目标

遮挡的场景中准确率提高了3.4%,在图像变形的场景中成功率提高了1.9%。

在OTB2015数据集上的实验结果显示,本文方法相比于ECO-HC,在光照变化的情况下,成功率提高了1.6%,准确率提高了2.2%;在目标离开视场的情况下,成功率提高了3.7%,准确率提高了3.4%,如图8所示。从以上实验结果可以看出,改进后的算法在图像背景复杂、图像模糊、目标被遮挡或某几帧中离开视场的情况下相比于原算法跟踪效果均有提升。

图9为本文算法与ECO-HC、C-COT、DFT^[18]、EBT^[25]和SRBT^[26]的对比实验,6种算法中只有C-COT用了卷积神经网络,其预训练模型为imagenet-vgg-m-2048。选用视频序列依次是Matrix、Soccer、Shaking和Ironman,这4个视频的光照均变化剧烈。Matrix、Soccer和Ironman中的目标运动快速且外观变化较多,跟踪难度大。在视频后半段,各对比算法发生跟踪漂移,丢失真实目标,而本文算法表现良好,各算法对比效果明显。即在图像场景变化复杂的情况下,应提高模型更新的速率,使算法能及时调整并适应目标或背景的改变,避免跟踪漂移。而在场景变化幅度不大时,减少模型更新既能降低过拟合的可能,又能减少运算并提高跟踪速度,本文算法在本文实验环境中平均速度可达60 FPS。

在VOT2016^[16]上的对比实验结果如表2所示,第一列为几种跟踪算法名称,中间3列为基线下的结果,最右两列为无监督情况下的结果。从表2可以看出,本文算法与基线ECO-HC相比,提高了重叠率(Overlap)、平均重叠期望(EAO)、AUC(Area under curve)和速度指标(EFO),减小了失败次数(Failures),提高了鲁棒性。从对比实验结果中可以看出C-COT的指标也很优秀,说明基于深度学习的目标跟踪方法能获得很好的跟踪效果,但与其改进型ECO^[14]的手工特征版本ECO-HC与本文方法相比,后者的速度是其30倍。ECO-HC是一种基于相关滤波的跟踪方法,本文方法是在其基础上改进而来的。C-COT在应用卷积神经网络提取特征的过程中产生了大量的参数使得运行速度变慢且占内存。同样应用深度学习的ECO使用因式分解卷积算子在数学上减少了滤波器个数,提高了计算效率,又有间隔的

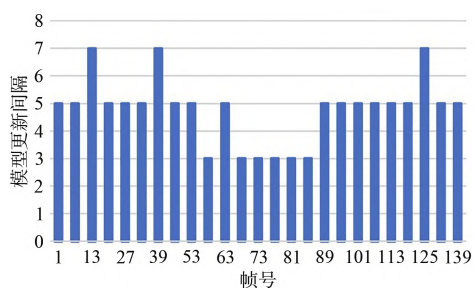


图4 Biker中模型更新间隔与帧号的关系

Fig. 4 Relationship between model update interval and frame number in Biker



图5 Biker中部分跟踪效果

Fig. 5 Partial tracking effects in Biker

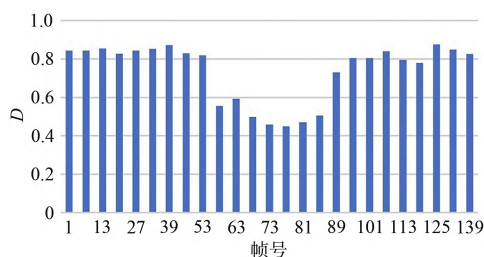


图6 Biker中 D 与帧号的关系

Fig. 6 Relationship between D and frame number in Biker

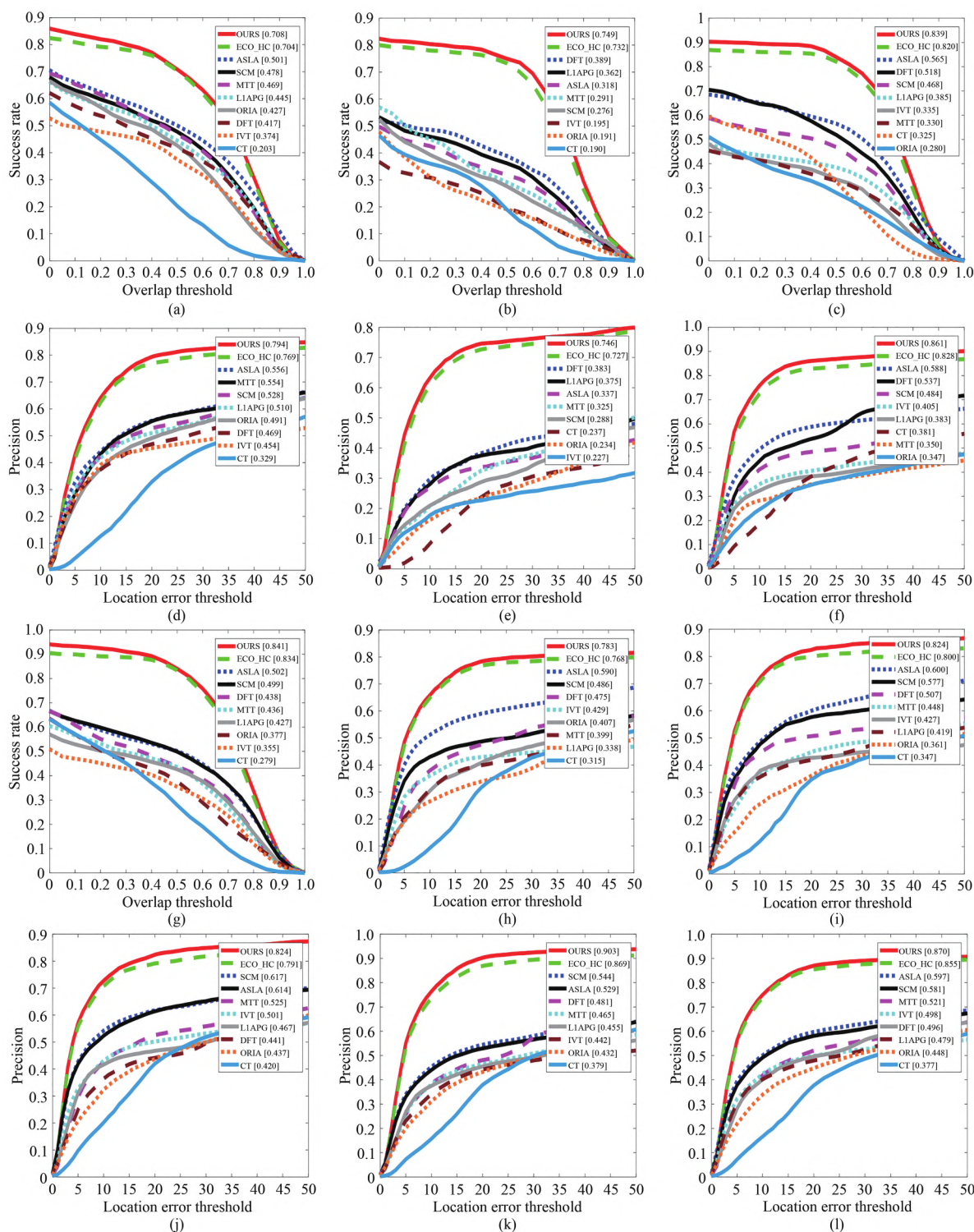


图 7 在 OTB2013 上的部分结果。(a) 面内旋转成功率图；(b) 运动模糊成功率图；(c) 目标变形成功率图；(d) 平面内旋转精度图；(e) 运动模糊精度图；(f) 目标变形精度图；(g) 目标被遮挡成功率图；(h) 光照变化精度图；(i) 背景杂乱精度图；(j) 尺度变化精度图；(k) 目标被遮挡精度图；(l) OPE^[15] 的精度图。

Fig. 7 Partial results on OTB2013. (a) Success plot of in-plane rotation; (b) Success plot of motion blur; (c) Success plot of deformation; (d) Precision plot of in-plane rotation; (e) Precision plot of motion blur; (f) Precision plot of deformation; (g) Success plot of occlusion; (h) Precision plot of illumination variation; (i) Precision plot of background clutter; (j) Precision plot of scale variation; (k) Precision plot of occlusion; (l) Precision plot of OPE^[15].

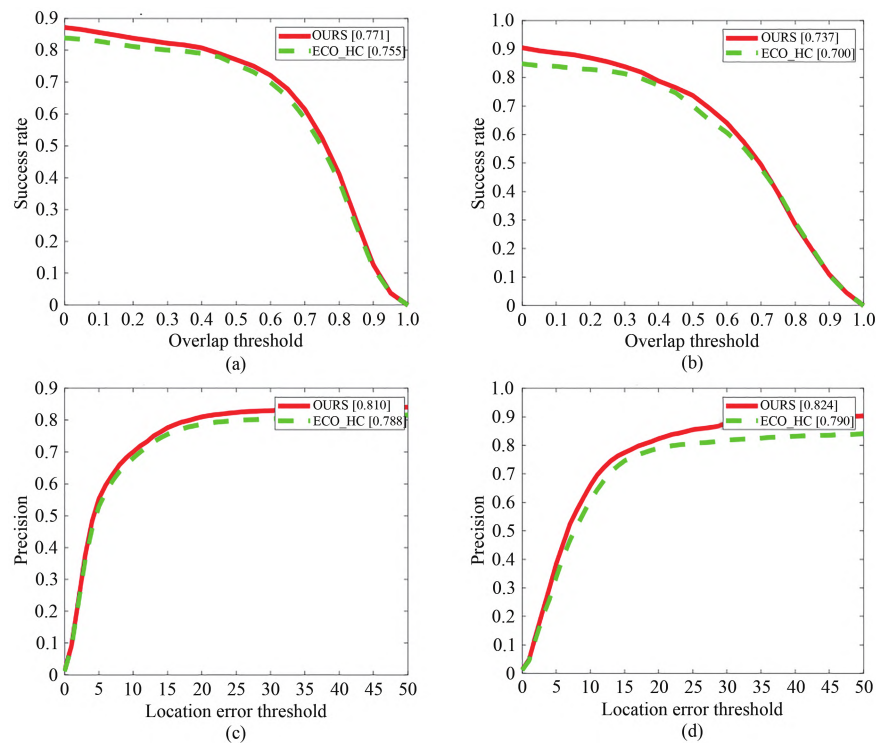


图 8 在 OTB2015 上的部分结果。(a)光照变化成功率图;(b)目标离开视场的成功率图;(c)光照变化精度图;(d)目标离开视场的精度图。

Fig. 8 Partial results on OTB2015. (a) Success plot of illumination variation; (b) Success plot of out of view; (c) Precision plot of illumination variation; (d) Precision plot of out of view.

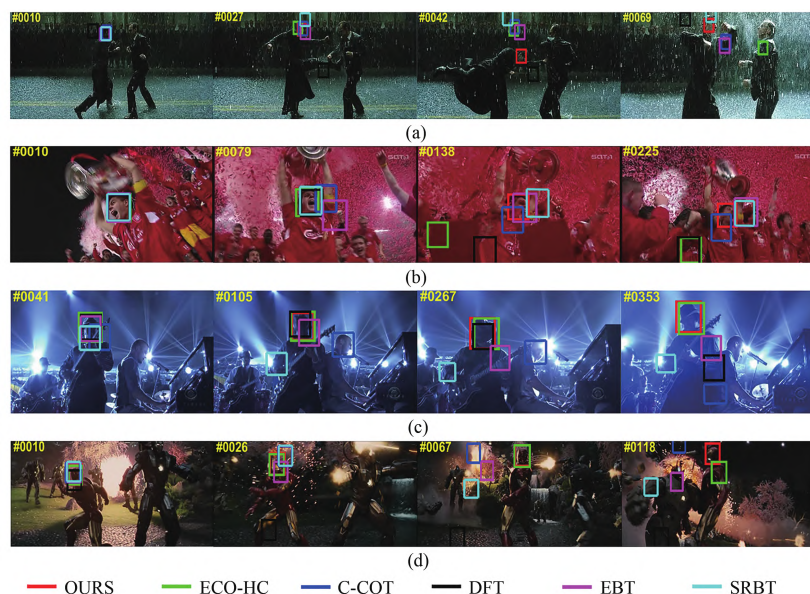


图 9 本文算法与其基线 ECO-HC 及其他 4 种算法的对比实验。(a)Matrix;(b)Soccer;(c)Shaking;(d)Ironman。

Fig. 9 Comparison between the algorithm Ours in this paper and its baseline ECO-HC and other four algorithms. (a) Matrix; (b) Soccer; (c) Shaking; (d) Ironman.

模型更新方式减少了模型更新次数,使其速度达到了 C-COT 的 8 倍。而 ECO-HC 更快的原因是

其未用卷积神经网络,仅用了梯度特征(HOG)和颜色特征(CN),运算复杂度更小。本文方法

表 2 在 VOT2016 上的对比实验
Tab. 2 Comparative experiment on VOT2016

	Overlap	Failures	EAO	AUC	EFO
OURS	0.530	21.195	0.288	0.404	15.416
ECO-HC	0.519	22.676	0.274	0.397	15.322
C-COT	0.515	16.582	0.329	0.460	0.507
DFT	0.444	59.612	0.139	0.209	3.330
EBT	0.447	15.194	0.287	0.365	3.011
SRBT	0.484	21.325	0.290	0.334	3.688

OURS除了具有ECO-HC的优点外,采用自适应模型更新策略使精度、速度和鲁棒性都得到了进一步提高。

5 结 论

本文提出了一种基于相关滤波的自适应模型更新策略,即摒弃常用的逐帧更新方法和每隔固定帧数更新一次的策略,自适应地以适当的更新间隔更新模型。本文定义了一个用于表述相邻两帧图像间差异的差异值 D ,以此来反映视频序列中的场景在此时刻的变化情况。当场景变化较剧烈时,减小跟踪模型相邻两次更新的间隔数,使之加速适应目标变化,提高跟踪准确率;当场景变化较和缓时,增大更新的间隔数,减少运算以提高运行速度,并避免过拟合。这样的模型更新策略更加灵活,可以在保持跟踪精度的同时提高跟踪效率,可应用于更多变的场景。在OTB2015上的实验结果显示,相比于基线ECO-HC,本文方法在光照变化的情况下成功率提高了1.6%,准确率提高了2.2%;在目标离开视场的情况下成功率提高了3.7%,准确率提高了3.4%。在本文实验环境中平均帧率可达60 FPS。

景变化较剧烈时,减小跟踪模型相邻两次更新的间隔数,使之加速适应目标变化,提高跟踪准确率;当场景变化较和缓时,增大更新的间隔数,减少运算以提高运行速度,并避免过拟合。这样的模型更新策略更加灵活,可以在保持跟踪精度的同时提高跟踪效率,可应用于更多变的场景。在OTB2015上的实验结果显示,相比于基线ECO-HC,本文方法在光照变化的情况下成功率提高了1.6%,准确率提高了2.2%;在目标离开视场的情况下成功率提高了3.7%,准确率提高了3.4%。在本文实验环境中平均帧率可达60 FPS。

参 考 文 献:

[1] 崔艺涵,陈涛,陈宝刚. 基于DPM和KCF的十字靶标检测与跟踪[J]. 液晶与显示,2018,33(12):1026-1032.
CUI Y H, CHEN T, CHEN B G. Detection and tracking method of cross target based on DPM detector and KCF tracker [J]. *Chinese Journal of Liquid Crystals and Displays*, 2018, 33(12): 1026-1032. (in Chinese)

[2] 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. San Francisco: IEEE, 2010: 2544-2550.

[3] DANELLJAN M, HÄGER G, SHAHBAZ KHAN F, *et al.* Convolutional features for correlation filter based visual tracking [C]//*Proceedings of the IEEE International Conference on Computer Vision Workshops*. Santiago: IEEE, 2015: 621-629.

[4] WANG F, WANG C L, CHEN M L, *et al.* Far-field super-resolution ghost imaging with a deep neural network constraint [J]. *Light: Science & Applications*, 2022, 11(1): 1.

[5] 梁慧慧,何秋生,贾伟振,等. 一种多特征融合的目标跟踪算法[J]. 液晶与显示,2020,35(6):583-594.
LIANG H H, HE Q S, JIA W Z, *et al.* Multi-feature fusion target tracking algorithm [J]. *Chinese Journal of Liquid Crystals and Displays*, 2020, 35(6): 583-594. (in Chinese)

[6] WANG J J, YANG H R, XU N, *et al.* Long-term target tracking combined with re-detection [J]. *EURASIP Journal on Advances in Signal Processing*, 2021, 2021(1): 2.

[7] DANELLJAN M, HÄGER G, SHAHBAZ KHAN F, *et al.* Adaptive decontamination of the training set: a unified formulation for discriminative visual tracking [C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016: 1430-1438.

[8] KOCH G, ZEMEL R, SALAKHUTDINOV R. Siamese neural networks for one-shot image recognition [C]//*Proceedings of the 32nd International Conference on Machine Learning*. Lille: JMLR, 2015.

[9] ZHANG J M, MA S G, SCLAROFF S. MEEM: robust tracking *via* multiple experts using entropy minimization [C].

- 13th European Conference on Computer Vision. Zurich: Springer, 2014: 188-203.
- [10] NAM H, HAN B. Learning multi-domain convolutional neural networks for visual tracking [C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016: 4293-4302.
- [11] DANELLJAN M, HÄGER G, SHAHBAZ KHAN F, *et al.* Learning spatially regularized correlation filters for visual tracking [C]//*Proceedings of the IEEE International Conference on Computer Vision*. Santiago: IEEE, 2015: 4310-4318.
- [12] DANELLJAN M, BHAT G, SHAHBAZ KHAN F, *et al.* ECO: efficient convolution operators for tracking [C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017: 6931-6939.
- [13] DANELLJAN M, ROBINSON A, SHAHBAZ KHAN F, *et al.* Beyond correlation filters: learning continuous convolution operators for visual tracking [C]. 14th European Conference on Computer Vision. Amsterdam: Springer, 2016: 472-488.
- [14] KRISTAN M, LEONARDIS A, MATAS J, *et al.* The sixth visual object tracking VOT2018 challenge results [C]//*Proceedings of the European Conference on Computer Vision (ECCV) Workshops*. Munich: Springer, 2019: 3-53.
- [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] WU Y, LIM J, YANG M H. Online object tracking: a benchmark [C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Portland: IEEE, 2013: 2411-2418.
- [17] JIA X, LU H C, YANG M H. Visual tracking via adaptive structural local sparse appearance model [C]. 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence: IEEE, 2012: 1822-1829.
- [18] SEVILLA-LARA L, LEARNED-MILLER E. Distribution fields for tracking [C]. 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence: IEEE, 2012: 1910-1917.
- [19] ZHONG W, LU H C, YANG M H. Robust object tracking via sparsity-based collaborative model [C]. 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence: IEEE, 2012: 1838-1845.
- [20] BAO C L, WU Y, LING H B, *et al.* Real time robust L1 tracker using accelerated proximal gradient approach [C]. 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence: IEEE, 2012: 1830-1837.
- [21] ROSS D A, LIM J, LIN R S, *et al.* Incremental learning for robust visual tracking [J]. *International Journal of Computer Vision*, 2008, 77(1/3): 125-141.
- [22] ZHANG T Z, GHANEM B, LIU S, *et al.* Robust visual tracking via structured multi-task sparse learning [J]. *International Journal of Computer Vision*, 2013, 101(2): 367-383.
- [23] ZHANG K H, ZHANG L, YANG M H. Real-time compressive tracking [C]. 12th European Conference on Computer Vision. Florence: Springer, 2012: 864-877.
- [24] WU Y, SHEN B, LING H B. Online robust image alignment via iterative convex optimization [C]. 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence: IEEE, 2012: 1808-1814.
- [25] ZHU G, PORIKLI F, LI H D. Beyond local search: tracking objects everywhere with instance-specific proposals [C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016: 943-951.
- [26] LEE H, KIM D. Salient region-based online object tracking [C]. 2018 IEEE Winter Conference on Applications of Computer Vision (WACV). Lake Tahoe: IEEE, 2018: 1170-1177.

作者简介:



刘刚阳(1996—),男,甘肃天水人,硕士研究生,2019年于兰州交通大学获得学士学位,主要从事图像处理与视觉目标跟踪方面的研究。E-mail:liugangyang20@mails.ucas.ac.cn