

相关滤波目标跟踪算法鲁棒性提升研究

王 奇, 王录涛, 江 山, 文成江
(成都信息工程大学 计算机学院, 成都 610225)

摘要: ECO算法虽然在跟踪效果上有了很大的提升, 但是它在复杂场景中表现不好甚至会丢失目标, 即鲁棒性不高; 对此, 文章探讨了怎样在 ECO算法中利用深度特征处理深层语义的能力和浅层特征处理纹理颜色信息的能力来提升算法的鲁棒性, 同时对比了深度特征和浅层特征在目标跟踪的不同作用, 并因此提出了改进方法, 首先在深度网络上选择了具有更深层次的 ResNet-101 网络; 其次修改了适宜此网络的参数; 算法在 OTB-2015 进行的实验也取得了比较好的结果, 在低分辨率、背景杂波、光照变化及尺度变化 4 个挑战因素的成功率分别领先基准算法 ECO 为 0.135, 0.034, 0.031, 0.024。

关键词: 目标跟踪; 深度特征; 浅层特征; 鲁棒性; 复杂场景

Research on Improving the Robustness of Correlation Filtering Target Tracking Algorithm

WANG Qi, WANG Lutao, JIANG Shan, WEN Chengjiang

(College of Computer Sciences, Chengdu University of Information Technology, Chengdu 610225, China)

Abstract: Although ECO algorithm has greatly improved the tracking effect, there are some shortcomings in complex scenes and even loses the target, and its robustness is not high. In this regard, how to improve the robustness of ECO algorithm is discussed by using the ability of deep feature to process deep semantics and the ability of shallow feature to process texture color information. At the same time, the different functions of depth feature and shallow feature are compared in target tracking, and an improved method is put forwarded, Firstly, the ResNet-101 network with deeper level is selected on the deep network; Secondly, the parameters suitable for this network are modified. The experiment of the algorithm in OTB-2015 also achieved good results, The success rates of low resolution, background clutter, illumination change and scale change are 0.135, 0.034, 0.031 and 0.024 respectively.

Keywords: visual tracking; depth features; shallow features; robustness; complex scene

0 引言

目标跟踪是一个只给予初始状态来评估一个目标在一段视频中轨迹位置的问题, 在现实生活中有很多应用, 比如汽车自动驾驶, 军事上的反导系统等^[1], 但是这个问题十分复杂, 因为在线学习目标外观模型时训练数据获得受到限制 (仅有第一帧的目标信息)^[2], 现有的方法几乎都依赖特征表现来处理这个问题, 而在深度特征广泛应用之前, 手工特征 (浅层特征) 的特征表现是处理上述问题的主流方法^[3], 虽然手工特征在准确性上表现优秀, 但其在复杂场景中的鲁棒性较差。随着深度神经网络 (convolutional neural networks) CNN 的发展, 研究人员发现深度特征相较于浅层特征具有编码高层信息的能力, 所以现今深度特征也被用于处理上述问题并且提升跟踪算法鲁棒性。

我们通过对本文章基准算法 ECO 及 ECO-HC (手工特征 ECO 算法) 代码和结果的分析, 在用同一环境下运行具有各

种挑战因素的数据集, 采用深度特征的 ECO 并不比采用手工特征的 ECO-HC 效果好, 原因可能是网络模型的深度不够并且没有选择适合特征表现的参数。如果我们能够在原有浅层特征和参数的基础上加入合适的深度网络和参数, 那么算法的鲁棒性能获得一定提升, 这是本文所要探讨的内容。

1 研究现状

目标跟踪算法现今大致可分为三类, 基于生成式模型的方法, 基于判别式模型的方法和基于深度学习的方法^[4]。生成式模型提取目标特征构建表现模型, 在图像中搜索与模型最匹配的区域作为跟踪结果; 与生成式模型不同的是, 判别式模型同时考虑了目标和背景信息。判别式模型将跟踪问题看做分类或者回归问题, 目的是寻找一个判别函数, 将目标从背景中分离出来, 从而实现对目标的跟踪; 基于深度学习的目标跟踪方法主要是利用深度特征强大的表征能力来实现跟踪。按照利用深度特征的方式, 可分为基于

收稿日期: 2021-12-03; 修回日期: 2021-12-22。

基金项目: 四川省科技厅重点研发项目 (2020YFG0442)。

作者简介: 王 奇 (1996-), 男, 四川万源人, 硕士研究生, 主要从事目标跟踪与嵌入式平台开发方向的研究。

通讯作者: 王录涛 (1979-), 男, 山东定陶人, 博士, 硕士生导师, 副教授, 主要从事计算机视觉、目标检测与跟踪及智能信号处理方向的研究。

引用格式: 王 奇, 王录涛, 江 山, 等. 相关滤波目标跟踪算法鲁棒性提升研究[J]. 计算机测量与控制, 2022, 30(6): 210-215.

预训练深度特征的跟踪和基于离线训练特征的跟踪。本文是相关滤波类的跟踪方法, 是典型的基于回归判别模型的跟踪算法。

相关滤波类的跟踪算法已经发展到比较成熟的阶段, 最初的 MOSSE^[5] 的基本思想是: 越是相似的两个目标相关值越大, 也就是视频帧中与初始化目标越相似, 得到的响应也就越大, MOSSE 算法的跟踪效果不好, 但是开辟了新的道路来研究目标跟踪; 后来极为成功的 C-COT^[6] 提出了将不同分辨率的特征图通过插值操作扩展到同一周期的连续空间域, 再应用 Hessian 矩阵可以求得亚像素精度的目标位置, 还解决了在连续空间域进行训练的问题, 但是由于计算量增大, C-COT 算法的运行速度十分缓慢; C-COT 的后作, ECO 旨在解决前身的速度慢的问题^[1], 在解决速度问题的同时兼顾提升了跟踪效果, ECO 算法没有区别地使用浅层特征和深度特征, 在复杂环境中不能稳定的跟踪目标; 之后的 UPDT^[7] 算法提出了区别的处理深度特征和浅层特征来提升跟踪效果, 但是 UPDT 算法采用的网络模型深度不够, 虽然在跟踪效果相比 ECO 有了提升, 但是在复杂场景的跟踪鲁棒性未能完全领先^[8-9]; 同年的 MFT^[10] 算法采用了和 UPDT 类似的方法, 并且引用了两个深度网络来提升跟踪效果, MFT 和 UPDT 存在同样的问题, 网络深度不够导致跟踪鲁棒性未能更进一步。怎样提高算法鲁棒性, 这是目前目标跟踪领域研究的热点。

2 对比测试及改进方法

2.1 深度特征和浅层特征

目标跟踪中很重要的两个标准, 即准确性 (accuracy) 和鲁棒性 (robustness), 准确性就是测量在跟踪过程中对目标定位的精度, 鲁棒性则是跟踪器在挑战性场景中的故障恢复能力, 换言之, 鲁棒性就是测量目标被准确定位的频率。在过往的目标跟踪发展中, 浅度特征被认为在处理跟踪准确性上有很大作用, 而深度特征则在鲁棒性上作用明显, 这由它们的特性决定。

深度特征主要是 CNN 的高层特征^[11], 优点是包含高层语义, 对旋转和变形等外观变化具有不变性, 基本是任何时间任何空间都能找到目标, 也就是上文提到的鲁棒性很好; 但缺点也很明显, 即空间分辨率低, 对平移和尺度存在不变性, 无法精确定位目标, 会造成目标漂移和跟踪失败, 即准确性很差^[12]。

浅层特征主要是颜色特征 (RGB, red-green-blue), 方向梯度直方图 (HOG, histogram of oriented gradient), 颜色名 (CN, color name), 和 CNN 的底层特征。优点主要是包含纹理信息和颜色信息, 空间分辨率高, 适合高精度定位目标, 即准确性很高; 但是不变性很差, 目标只要稍微形变就不能识别, 尤其对旋转十分敏感, 也就是鲁棒性很差。

图 1 是 OTB-2015 视频序列 CarScale 运行是的截图, 我们选取的是几个代表性场景帧, 93 帧是目标在相对良好的

背景下运动的状态, 而 161 帧是在复杂场景运动的状态, 190 帧脱离复杂场景的状态。



图 1 CarScale 视频序列在场景中效果框图

对于图 1, 我们用更加直观的特征通道响应值图 (由于各个特征分辨率不同, 现在的算法都将不同分辨率插值处理, 表现为不同的分辨率对应不同的通道, 深层特征和浅层特征处于不同通道, 可用 Matlab 提取作图) 来反映深度特征 (图 2 第一列) 和浅层特征 (图 2 第二列) 的在此过程的受影响而产生的变化^[13]。首先, 图 2 (a) 对应图 1 第 93 帧, 图中目标还没有进入复杂背景的区域, 深层特征最大波峰显著高于其余波峰, 周围环境噪声虽然也产生影响, 但不明显; 而 (b) 图 (也为 93 帧的时候) 显示浅层特征波峰明显, 周围几乎没有噪声影响; (c) 图对应图 1 第 161 帧, 当目标进入复杂背景区域后, 深层特征周围噪声明显增加, 但是目标产生的波峰依旧十分明显; (d) (也为 161 帧) 图显示此时浅层特征明显受环境干扰, 目标波峰周围出现明显的噪声; (e) 图 (对应图 1 第 190 帧) 显示目标脱离背景复杂区域后, 从坐标轴数据看深层特征波峰恢复并且超越之前的水平, (f) 图对应浅层特征峰值则并未恢复到之前水平。对此我们得出结论, 浅层特征在背景环境复杂时受影响较大, 深层特征受影响则相对较小; 深层特征恢复能力较强, 但是浅层特征在无明显干扰的环境下精度较高, 即深层特征鲁棒性较高 (稳健性), 浅层特征准确性较高^[14]。

图 3 (a) 是视频序列 Basketball 的第一帧图像, 我们把它作为原始图像来进行分析, (b) 图对应 HOG 特征图, HOG 特征对应的特征图纹理比较明显, 图像中目标的轮廓大致清晰, 这也对应了上面我们提到的浅层特征空间分辨率高, 包含纹理信息和颜色信息, 对目标定位的精度高, 准确性好。(c) 图对应的是 CNN 的高层特征图, 在第一行单位已经到了千分之一的级别, 目标在图中已经凝聚成块状, 轮廓不明显, 但是它将目标的范围缩小了, 虽然分辨率不高, 但是不易受环境干扰而且层次更深, 这是鲁棒性高的表现。(d) 图对应的是 CNN 的底层特征, 与 CNN 的高层特征相比, 底层特征图纹理更加明显, 目标反映成点状, 精度更高; 但是和 HOG 特征相比, 其没有 HOG 特征在处理光照, 纹理信息方面优秀^[15]。

2.2 具体方法

ECO 采用了 VGGNet^[16] 网络模型, 虽然取得了不错的跟踪效果, 但是由于深度的限制无法对复杂的场景进行更进一步的分析, 所以在面对一些复杂的场景时会出现效果变差或者丢失目标的情况, 也即鲁棒性 (稳健性) 不够。

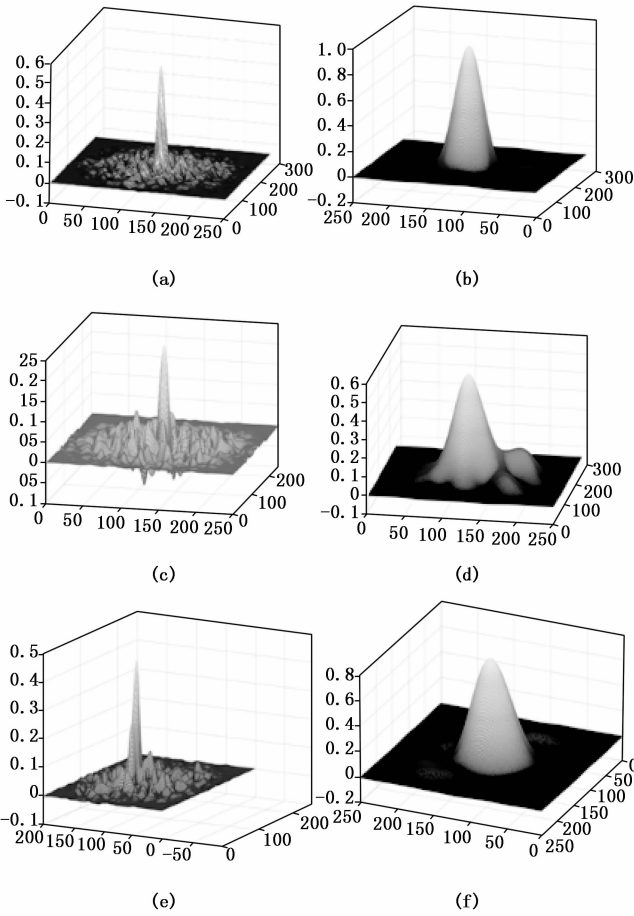


图 2 特征通道响应值对比图

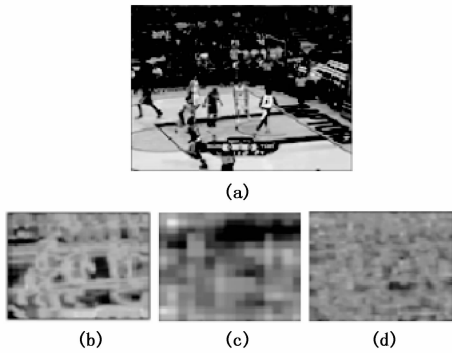


图 3 原始图和 ECO 特征可视化图

对于此问题我们首先想到了层次更深的网络模型 ResNet-50^[17]和 ResNet-101^[17]。表 1 显示 3 个网络模型各自的层数和官方验证错误率。

表 1 3 个网络模型对比表

	VGGNet	ResNet-50	ResNet-101
Layer	16	50	101
Top-1 err/%	28.07	22.85	21.75
Top-5 err/%	9.33	6.71	6.05

ECO 之后的跟踪算法普遍都采用了层次更深的 Res-

Net-50 网络模型来提升跟踪效果，我们则测试了更深层次的 ResNet-101 网络模型，当我们单纯把 ResNet-101 应用到 ECO 算法上时发现效果不佳甚至不如之前的表现，图 4 是单纯将不同网络 (ResNet-50, ResNet-101) 应用到算法中的效果图 (图 4、图 5 都采用目标跟踪基准数据库 OTB-2015 的 SRE 指标来做验证, SRE 是空间鲁棒性评估, 满足我们鲁棒性评估的要求, 我们选用的视频序列都是第三章中 OTB-2015 满足 4 种挑战的视频序列)。

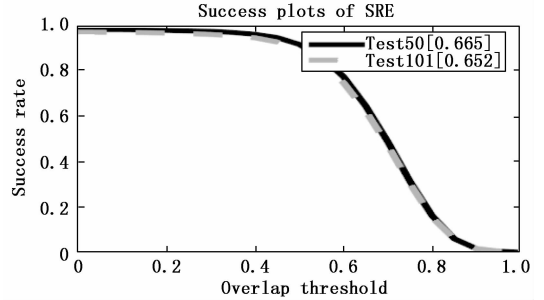


图 4 只修改两种网络效果图

另外, ECO 采用高斯函数, 用参数 σ 表示标准差, 这个标准差设置会对效果有很大影响, 但是这些影响建立在网络层次差异基础之上, 这和深度特征和浅层特征在目标跟踪中的不同作用有关。如果在 ResNet-50 的基础上, 我们设置 $\sigma = \frac{1}{4}$, 那么深度特征会发挥最好的作用; 相反的, 采用手工特征的方法则需要 $\sigma = \frac{1}{16}$ 才可以发挥较好的作用, ECO-HC 原本采用了这个数字并且有很好的效果, 但是 ECO 采用的 $\sigma = \frac{1}{12}$ 这个值, 出发点为了适应 VGGNet 网络模型, 如果我们在原本的 ECO 方法单纯改为 $\sigma = \frac{1}{4}$, 效果不但没有变好, 反而变差了。

图 5 是我们只设置 $\sigma = \frac{1}{4}$ (即 ECOS) 与 ECO 原本设置的 $\sigma = \frac{1}{12}$ 的对比效果, 可以看到单独修改 σ 的 ECOS 相比 ECO 效果有所下降。

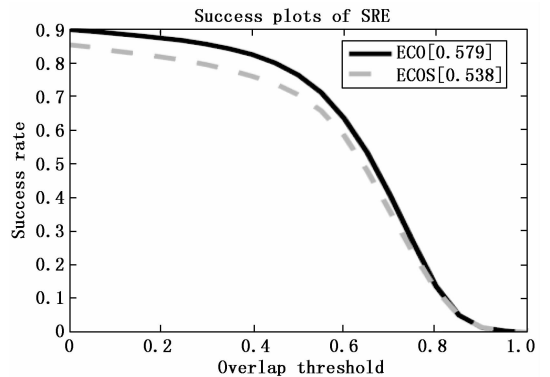


图 5 修改 σ 后和之前对比图

表 2 单纯对比更改网络和参数在 OTB-2015 中的 AUC (area under curve) 分值 (AUC 为成功率图的曲线下面积, 分值越高效果越好), 它证实了单纯修改网络和单独修改参数 σ 无法提高跟踪效果, 必须从以上两个方面进行改进。

表 2 各种修改后测试的 AUC 分值

	VGGnet	VGGnet+ σ	ResNet	ResNet+ σ
AUC	57.0	53.8	65.2	68.0

参数 σ 在我们的方法中十分重要, 要理解我们修改它的原因就必须了解它的作用。首先相关滤波类的算法训练样本来自目标扩展区域, ground truth 是高斯分布, 中心部分接近 1 是正样本, 边缘部分接近 0 是负样本, 正负样本比例通过高斯函数的标准差 σ 控制, 标准差越大正样本越多, 而这样可以得到更接近于目标中心的参数, 这会利于深度特征, 但也并不是越大越好, 如果太接近中心则会很容易产生漂移; 相反如果标准差选的小, 对应边界的轮廓会相对清晰, 这有利于浅层特征, 但是也不能太小, 太小则不能定位目标区域。ECO 之后很多算法选择了 $\sigma = \frac{1}{4}$ 作为深层特征的参数, 但是我们在实验过程中发现, 在应用比 ResNet-50 更深层次的网络时, 我们取 $\sigma = \frac{1}{3}$ 效果会稍好一点, 这可以归结为此时的正样本相对较多, 深度特征受益较大, 而在一般的变换中深度特征具有不变性, 也即我们提到深度特征的鲁棒性, 即不易受影响改变的特性。浅层特征的首选 σ 参数 ECO-HC 已经给出了标准, 我们可以直接引用。

表 3 中, 我们设置 $\sigma = \frac{1}{3}$ 为 σ_1 , $\sigma = \frac{1}{4}$ 为 σ_2 , 对比 ResNet-50 和 ResNet-101 在不同参数的下最终的 AUC 分值, 以此分值来确定我们最终要使用的网络模型和参数。

表 3 各种修改后测试的 AUC 分值

	ResNet50+ σ_2	ResNet50+ σ_1	ResNet101+ σ_2	ResNet101+ σ_1
AUC	68.0	65.2	67.1	70.0

以上实验我们得出结论, 单纯修改网络模型或者 σ 参数不但没有提高鲁棒性, 还会降低算法原有的跟踪效果。选用不同的网络模型和其对应的参数会有不同的 AUC 值, 使用网络模型 ResNet-101 和参数 $\sigma = \frac{1}{3}$ 可以得到最高 AUC 值, 也即最佳效果。

我们将原本的 σ 由单个参数修改为一个参数集, 这个参数集既适宜深度特征发挥作用也能够使原本的浅层特征发挥作用。

$$\sigma = \frac{1}{12}$$

$$\sigma = \frac{1}{16}$$

变换为:

$$\sigma = \left[\frac{1}{3}, \frac{1}{16} \right]$$

计算过程中在利用手工特征时, 会调用 σ 数组的第二个元素; 在运用深度特征时, 会调用 σ_2 数组的第一个元素, 达到最好的鲁棒性效果。

ECO 采用高斯函数中的高斯混合模型 (GMM), 这也是我们修改参数 σ 所起作用的地方。单高斯模型中, 一维数据:

$$p(x | \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-a)^2/2\sigma^2} \quad (1)$$

样本多维时, \sum 为协方差, u 为数据均值, D 为数据维度, σ 在参与了协方差的计算, 用来给定正负样本比例:

$$p(x | \theta) = \frac{1}{(2\pi)^{D/2} \sum^{1/2}} e^{-(x-u)^T \sum^{-1} (x-u)/2} \quad (2)$$

混合模型中:

$$p(x | \theta) = \sum_{k=1}^k \alpha_k \phi(x | \theta_k) \quad (3)$$

在此参数 θ 由每个子模型的期望、方差 (或协方差)、在混合模型中发生的概率组成, 我们修改的参数 σ 表示为协方差。在 ECO 中则是表现在对式 (4) y 值的影响, 即对特征通道响应值的影响, 式 (4) 中 σ 是我们修改的参数集, $u \in [0, t]$ 是目标估计的位置, y 是以 u 为中心的高斯函数 $e^{-(t-u)^2/2\sigma^2}$ 的周期重复。

$$y = \frac{\sqrt{2\pi}\sigma^2}{T} e^{(-2\sigma^2(\pi k/T)^2 - 2\pi/Tuk)} \quad (4)$$

而 y 在 ECO 中参与期望 E 的运算。

$$E(f) = \epsilon \{ \|S_f(X) - y\|_{L^2}^2 \} + \sum_{\sigma=1}^{\sigma} \|wf^d\|_{L^2}^2 \quad (5)$$

式中, 期望 E 是对联合样本的值分布的评估, $S_f \{X\}$ 是对目标检测的分值, Y 是对样本的标记检测分值。 W 表示权重, f^d 表示多通道卷积滤波器, L 表示高斯组件的数量。

3 运行结果对比及分析

基于目标跟踪基准数据库 OTB-2015, 采用 SRE (空间鲁棒评估) 的指标进行评测^[18], 由于验证算法鲁棒性的需要, 我们选取 Liquor, Dudek, Lemming 这几个测试视频序列, 这几个视频序列包含我们用于评测的 4 种挑战因素 (背景杂波 (BC), 尺度变化 (SV), 光照变化 (IV), 低分辨率 (LR))。我们选取的都是各个视频序列具有代表性的帧, 它们可以很好的代表我们想测试的 4 种挑战因素。参与比较的 3 种算法分别是 ECO, ECO-HC, 我们改进的算法 (命名为 OURS), 分别采用不同的曲线代表, 即直线, 虚线和点线, 这可以很好的区分图中的 3 种算法, 使比较结果更加清晰。ECO 是我们的基准算法, ECO-HC 使手工特征的 ECO 版本, 和他们对比可以体现出我们改进算法在效果方面的提升。

由图 6 的几张效果评测图可以看出, 在以上 4 种挑战因素的场景中, 我们的算法始终能定位目标, 与我们改进算法相对比的基准算法出现了不能稳定跟踪甚至丢失目标的情况, 具体来说, 3 种算法的定位框清晰的显示, 在上图所展示的挑战因素中, 我们的改进算法的定位框始终能将目标大体包裹, 而其他两种对比算法出现了包裹部分减小甚



图 6 算法定性效果评测

至完全未包裹目标，这在实际应用中就已经丢失了目标。在场景不太复杂的情况下，如图 6 第二行两幅图，虽然另外两种算法也可以对目标进行有效的跟踪，定位框也基本能包裹目标，但是我们改进的算法定位更加精准，表现在定位框能够完全的包裹目标。此定性分析说明我们改进的算法对比基准算法鲁棒性有所提升，在几种挑战因素的场景中比基准算法的跟踪效果和稳定性更好。

OTB-2015 拥有 100 个视频序列，包含 12 个挑战因素^[19]，每个视频序列可以用于验证一个或者多个挑战因素。OTB-2015 是目标跟踪领域应用广泛的基准数据库，几乎所有最新提出的算法都会放在该数据库上进行评测，评测结果会作为算法效果的重要参照，OTB-2015 也是十分科学的平台，在该平台的评测数据基本能够反映一个算法的可行性，对此我们改进算法选择在 OTB-2015 标准数据库评测。图 8 在低分辨率，背景杂波，光照变化，尺度变化用成功率 (Success plots) 曲线作了对比，表 4 列举了除以上我们演示的挑战因素外 OTB-2015 中在各个挑战因素具有代表性的 4 个视频序列，这也是用于验证我们算法的一部分视频序列。

表 4 各种挑战因素验证视频序列

	Dataset1	Dataset2	Dataset3	Dataset4
LR	BoxCar1	Car1	Ironman	Surfer
BC	Basketball	Dudek	Bolt2	Football1
IV	Singer1	Trans	Tiger	Man
SV	Human	Twinnings	Vase	Skater

图 7 是我们改进算法在 OTB-2015 中选取包括表 4 中验证序列的运行结果。

我们将数据以表格的形式展示 (只对比 OURS, ECO, ECO-HC)，从表 5 我们可以更直观地看到，我们改进的算

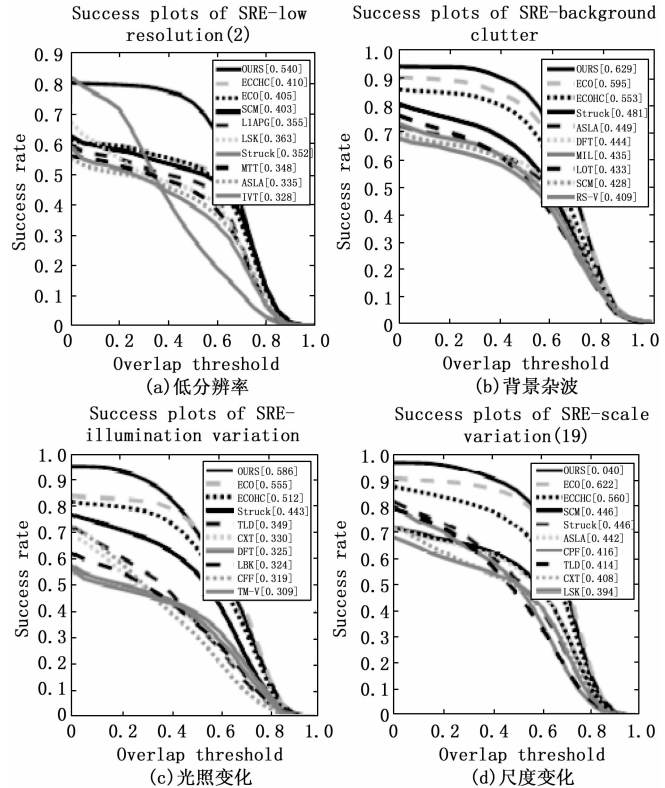


图 7 算法定量分析效果评测

法在这 4 种挑战因素中表现比基准算法有所提高。

表 5 实验结果定量分析表格

	ECO	ECO-HC	OURS
LR	0.405	0.410	0.540
BC	0.595	0.553	0.629
IV	0.555	0.512	0.586
SV	0.622	0.560	0.646

我们的方法在这几种挑战因素的鲁棒性测试中略优于 ECO 与 ECO-HC，对此我们对这几种挑战因素^[20]的视频序列分析发现：

1) 低分辨率 (LR)^[19]的视频序列目标都比较小而且图像相对模糊，这种情况底层特征的作用就会被减弱，层次不够深的网络模型也会处理困难，但是我们的方法涉及的网络模型层次足够，这也是为什么我们的方法在评分上领先 ECO-HC 和 ECO 大约 0.1 的原因。

2) 背景杂波 (BC)^[19]的视频序列大多都有相同物体的干扰或者图片的背景信息丰富，这会导致浅层特征在纹理，背景信息处理出现困难，但是深层特征可以很好的应对这种状况。在数据上，我们的方法领先 ECO 大约 0.03 分值，领先 ECO-HC 大约 0.075。

3) 测试光照变化 (IV)^[19]的视频序列基本都是图像光线较暗或者黑夜环境，我们对底层特征参数的应用对这种情况起到了作用，底层特征对光流信息比较敏感，但是对比 ECO-HC 我们还多了 CNN 的底层特征，这是我们改进

算法优于 ECO 和 ECO-HC 的原因。

4) 测试尺度变化 (SV)^[19] 场景的视频序列大多有一个共同点, 那就是目标会由远及近, 尺寸会存在变化, 这是深度特征和浅层特征共同起作用的地方, 目标尺寸小的时候会很模糊, 这就依赖于深度特征来实现定位目标, 在目标尺寸变大的时候, 浅层特征可以较好处理目标的边界和形状信息。所以在数据上我们的方法领先 ECO 大约 0.02。

4 结束语

本文旨在提升跟踪算法的鲁棒性 (稳健性) 用以提升跟踪算法在复杂环境中跟踪目标的能力, 同时我们对算法原有的准确性也有一定提升, 我们不仅采用了目前深度较深的 ResNet-101 网络模型, 而且也设置了适宜此网络模型的 σ 参数, 我们还保持了 ECO-HC 原有的 σ 参数来延续浅层特征跟踪过程中对准确性所起的作用。我们的方法在 OTB-2015 基准数据库中效果相比原算法有所进步。深层次网络训练需要采用数据增强, 数据增强可以在不采用巨大样本集的前提下提供深度网络更多的样本, 通过翻转, 平移等方法。在深层特征的选用上, 更深层次的网络已经存在, 如果运用 ResNet-152 或者更深层次的 Inception-ResNetV2 网络模型可能会取得更好的效果, 即随着网络层次深度变化, 跟踪算法或许会取得更加优秀的成果, 但是随着网络模型深度的增加, 跟踪算法的运行时间也会随之增加, 这是因为深度越深的网络模型越需要大量的样本, 而从大量样本中提取特征会增加运行时间。如何把运行速度提高以缩短运行时间, 这是我们今后的研究要解决的问题。

在 ECO 算法中我们就发现传统的跟踪算法很多都是只在第一帧学习, 然后将学习的结果在后面许多帧保持并利用, 这种方式在复杂环境中很难保证跟踪质量^[21]。如果在复杂的环境中除了目标还有与目标类型相似的物体, 那么跟踪效果往往会很差甚至丢失目标。现在已有研究者提出了能够部分处理此问题的方法, 但是效果还未在正式的平台展示, 将他们的研究和我们的发现结合, 我们大致得出结论, 如果我们在第一帧学习并且逐帧采集背景信息不断向后面帧传播我们收集的信息或许对解决上述问题有所帮助^[22-23], 这也是我们今后进一步研究的另一个方向。

参考文献:

- [1] DANELLJAN M, BHAT G, KHAN F S, et al. ECO: efficient convolution operators for tracking [C] // IEEE Computer Society. IEEE Computer Society, 2016.
- [2] 张 微, 康宝生. 相关滤波目标跟踪进展综述 [J]. 中国图象图形学报, 2017, 22 (8): 1017-1033.
- [3] 陈 旭, 孟朝晖. 基于深度学习的目标视频跟踪算法综述 [J]. 计算机系统应用, 2019, 28 (1): 1-9.
- [4] 李 玺, 查宇飞, 张天柱, 等. 深度学习的目标跟踪算法综述 [J]. 中国图象图形学报, 2019, 24 (12): 2057-2080.
- [5] BOLME D S, BEVERIDGE J R, DRAPER B A, et al. Visual object tracking using adaptive correlation filters [C] // The

- Twenty-Third IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2010, San Francisco, CA, USA, IEEE, 2010.
- [6] DANELLJAN M, ROBINSON A, KHAN F S, et al. Beyond Correlation Filters: Learning Continuous Convolution Operators for Visual Tracking [C] // Springer International Publishing. Springer International Publishing, 2016.
- [7] BHAT G, JOHNANDER J, DANELLJAN M, et al. Unveiling the Power of Deep Tracking [C] // European Conference on Computer Vision. Springer, Cham, 2018.
- [8] 何 容, 赖际舟, 吕 品, 等. 复杂环境下一种基于改进核相关滤波的视觉鲁棒目标跟踪方法 [J]. 导航定位与授时, 2020, 7 (6): 71-78.
- [9] 刘 航. 复杂背景下基于相关滤波的目标跟踪算法研究 [D]. 湘潭: 湘潭大学, 2020.
- [10] BAI S, HE Z, DONG Y, et al. Multi-hierarchical independent correlation filters for visual tracking [C] // 2020 IEEE International Conference on Multimedia and Expo (ICME), IEEE, 2020.
- [11] 周飞燕, 金林鹏, 董 军. 卷积神经网络研究综述 [J]. 计算机学报, 2017, 40 (6): 1229-1251.
- [12] 卢宏涛, 张秦川. 深度卷积神经网络在计算机视觉中的应用研究综述 [J]. 数据采集与处理, 2016, 31 (1): 1-17.
- [13] 谢 悦. 基于相关滤波的抗遮挡目标跟踪算法研究 [D]. 长沙: 湖南大学, 2020.
- [14] 张艳琳, 钱小燕, 张 森, 等. 自适应多特征融合相关滤波目标跟踪 [J]. 中国图象图形学报, 2020, 25 (6): 1160-1170.
- [15] 许倩倩. 基于卷积神经网络多特征融合的相关滤波目标跟踪算法研究 [D]. 杭州: 浙江工业大学, 2020.
- [16] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. arXiv preprint arXiv: 1409.1556, 2014.
- [17] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [18] YI W, LIM J, YANG M H. Online object tracking: a benchmark [C] // Computer Vision & Pattern Recognition, IEEE, 2013.
- [19] 姜继朋. 基于相关滤波的复杂场景下目标跟踪方法研究及实现 [D]. 重庆: 重庆邮电大学, 2020.
- [20] 熊昌镇, 赵璐璐, 郭芬红. 自适应特征融合的核相关滤波跟踪算法 [J]. 计算机辅助设计与图形学学报, 2017, 29 (6): 1068-1074.
- [21] BHAT G, DANELLJAN M, VAN GOOL L, et al. Know your surroundings: exploiting scene information for object tracking [C] // European Conference on Computer Vision. Springer, Cham, 2020: 205-221.
- [22] 陈万敏. 结合时空上下文信息的相关滤波运动目标跟踪方法研究 [D]. 昆明: 昆明理工大学, 2020.
- [23] 孙 欣, 何 宁. 结合背景空间信息的相关滤波目标跟踪算法研究 [J]. 计算机应用与软件, 2020, 37 (2): 130-133, 176.