

基于二阶池化网络的鲁棒视觉跟踪算法

蒲 磊¹,冯新喜²,侯志强³,余旺盛²

(1. 空军工程大学研究生院,陕西西安 710077; 2. 空军工程大学信息与导航学院,陕西西安 710077;
3. 西安邮电大学计算机学院,陕西西安 710121)

摘 要: 针对低分辨率、遮挡以及相似物体干扰等复杂场景下目标易丢失的问题,本文提出了基于二阶池化网络的视觉跟踪算法.已有的方法大多采用一阶池化网络,使得对低分辨目标和相似目标间的区分性不足.对此,本文首先在 VGG16 网络结构的基础上,将网络最后的一阶池化层替换为二阶协方差池化层,接着在 ImageNet 和 CUB200-2011 数据集上对网络进行重新训练.在跟踪阶段,为了减少运算负担,仅提取预训练网络的第四层卷积特征作为目标的外观表征.最后将提取的特征与已有的相关滤波算法进行结合.实验结果表明,本文算法在跟踪精度和成功率上均取得了优异的性能表现.

关键词: 视觉跟踪;二阶池化网络;深度特征;相关滤波

中图分类号: TP391.4 **文献标识码:** A **文章编号:** 0372-2112 (2020)08-1472-07

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2020.08.003

Robust Visual Tracking Based on Second Order Pooling Network

PU Lei¹, FENG Xin-xi², HOU Zhi-qiang³, YU Wang-sheng²

(1. Graduate College, Air Force Engineering University, Xi'an, Shaanxi 710077, China;

2. Institute of Information and Navigation, Air Force Engineering University, Xi'an, Shaanxi 710077, China;

3. School of Computer Science and Technology, Xian University of Posts and Telecommunications, Xi'an, Shaanxi 710121, China)

Abstract: Aiming at the problem that the target is easy to lose in the complex scene such as low resolution, occlusion, the interference of similar objects, this paper proposes a visual tracking algorithm based on second-order pooling network. Most of the existing methods use the first-order pooling network, which makes the difference between similar targets insufficient. In this paper, based on the VGG16 network structure, the last first-order pooling layer is replaced by the second-order covariance pooling layer, and then the network is retrained on ImageNet and CUB200-2011 image data sets. In order to reduce the computational burden, only the fourth convolution feature of the pre-training network is extracted as the appearance representation of the target. Finally, the extracted features are combined with the existing correlation filtering algorithm. The experimental results show that the algorithm achieves excellent performance in tracking accuracy and success rate.

Key words: visual tracking; second-order pooling network; deep features; correlation filter

1 引言

视觉跟踪是计算机视觉领域的基础和难点问题,在实际生活中有着广泛的应用,比如视频监控、运动分析、自动驾驶和人机交互等领域.视觉跟踪一般指的是单目标跟踪,在只给定首帧关于目标的位置及尺度信息的条件下,在后续帧中估计出目标的当前状态信息^[1].虽然经过了三十多年的发展,实现鲁棒的视觉跟踪系统依然是一项富有挑战性的任务.该任务的难点

主要集中在两个方面,一方面是目标自身的变化,包括尺度变化、形变、快速运动等,另一方面是外在的复杂环境变化,包括光照变化、背景杂波、严重遮挡等场景.研究人员已经证明,良好的目标特征表达在应对上述复杂场景中其中起着非常关键的作用^[2].为了获得对目标的良好特征表示,研究人员先后设计了大量优秀的手工特征,如 HOG、SURF 等.但是这些特征仅对部分简单场景具有良好的表现,对复杂场景难以适用.

近年来,深度学习在计算机视觉的方方面面都得

到了很好的应用,在诸如图像分类^[3]、目标检测^[4]、图像分割^[5]、目标识别^[6]等任务上都取得了大幅度的性能提升.由于深度特征具有良好的特征表达能力和泛化性,将其迁移到跟踪任务便是一个很值得研究的课题.目前将深度特征和跟踪任务进行结合的研究主要可分为三个方向,一是直接采用离线预训练的网络进行特征提取^[7-9];二是离线训练加在线微调,代表性的工作是2015年提出的MDNet算法^[10],采用对每一个跟踪对象分别训练一个分支网络的方式取得了VOT2015的冠军;三是尝试直接训练出一个端到端的跟踪网络,主要的工作是基于孪生网络的系列算法,比如早期的SiamFC算法^[11]将特征表达和分类器联合训练,在速度和精度上都取得了很好的表现.

但是这些方法所采用的网络结构都是基于一阶池化操作,并没有对特征的二阶统计信息进行进一步的分析.近年来,高阶统计建模在各种计算机视觉任务中引起了极大的关注,如目标识别^[12]、语义分割^[13]、行人重识别^[14]以及细粒度图像分类^[15]等.在分类任务中,Li等人^[15]通过大量实验证明了协方差池化在很大程度上优于传统的一阶池化网络.对此,本文对采用的特征提取网络进行适当的改进,将原来的一阶均值池化改为二阶协方差池化方法,使得网络可以提取更多关于目标特征的高阶统计量,提高模型的判别力.同时,本文也在ImageNet^[16]以及细粒度图像CUB200-2011数据集上对网络进行了重新训练,进一步提高了模型对相似物体的区分能力.

本文从提高目标特征的表达力入手,提出了基于高阶池化网络的视觉跟踪算法.首先通过大量数据集对构建的高阶池化网络进行训练,使得学习到的特征更加具有判别能力.接着和相关滤波算法进行结合,在大规模跟踪数据集上进行测试,取得了非常好的性能表现.据我们所知,本文算法是首次将高阶池化网络运用于视觉跟踪任务.

2 本文算法

本文提出了基于高阶池化网络的鲁棒视觉跟踪算法.首先对传统的多通道相关滤波算法进行简单的介绍,接着介绍本文提出的高阶池化网络特征训练和提取方法.最后将其与现有的STRCF^[17]进行结合,在未做任何参数调整的情况下,取得了较大幅度的性能提升.

2.1 多通道相关滤波算法

相关滤波算法的核心就是一个岭回归模型,通过对基础样本进行循环移位得到大量的训练样本,具体的模型如下所示:

$$E(f) = \left\| \sum_{l=1}^d \mathbf{x}^l * f^l - \mathbf{y} \right\|_2^2 + \lambda \sum_{l=1}^d \left\| f^l \right\|_2^2 \quad (1)$$

其中 \mathbf{x} 是循环样本, f 是滤波器, d 是通道数, \mathbf{y} 是期望输出的响应标签值,一般由二维高斯函数产生, λ 是正则化系数,*表示循环相关操作.

在定位阶段,为了降低运算量,在傅里叶域进行求解,最后经过逆变换得到最终的响应图如下所示:

$$\mathbf{g} = \mathbf{F}^{-1} \left(\sum_{l=1}^d \hat{\mathbf{z}}^l \odot \hat{\mathbf{f}}^l \right) \quad (2)$$

2.2 高阶池化网络

卷积神经网络通过对卷积层、非线性激活层以及池化层的堆叠,构建了许多非常优秀的网络架构,并被广泛应用于计算机视觉领域的方方面面^[6].目前大部分基于网络的改进更多的是通过对网络进行加宽加深来提升网络的性能,鲜有关于从高阶信息角度来提高特征表达能力的研究工作.为了增强特征的判别力,本文采用高阶池化网络作为目标的特征提取器.对于一帧输入图像,该网络输出一个规范化的协方差矩阵作为目标的表示,接下来本文将简要介绍该结构涉及到的前向与后向传播公式.

2.2.1 前向传播推导

首先需要对最后一层卷积层的输出进行处理,将处理后的特征矩阵表示为 $\mathbf{X} \in \mathbf{R}^{d \times N}$,其中 d 为特征通道数, $N = h \times w$,其中 h, w 为最后一层卷积层的特征图大小.接着计算特征图的协方差矩阵:

$$\mathbf{P} = \mathbf{X} \mathbf{X}^T \quad (3)$$

其中 $\mathbf{I} = \frac{1}{N} \left(\mathbf{I} - \frac{1}{N} \mathbf{i} \mathbf{i}^T \right)$, $\mathbf{i} = [1, \dots, 1]^T$ 为一个 N 维的向量.接着采用特征值分解对得到的协方差矩阵进行处理,得到特征值和特征向量:

$$\mathbf{P} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T \quad (4)$$

其中 $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_d)$ 是一个对角矩阵,且 λ_i 为特征值. $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_d]$, \mathbf{u}_i 为对应的特征向量.通过上述的特征值分解我们可以将矩阵的幂转换为对特征值的幂进行求解:

$$\mathbf{Q} \triangleq \mathbf{P}^\alpha = \mathbf{U} \mathbf{F}(\mathbf{\Lambda}) \mathbf{U}^T \quad (5)$$

在本文中, $\alpha = 0.5$. $\mathbf{F}(\mathbf{\Lambda}) = \text{diag}(f(\lambda_1), \dots, f(\lambda_d))$,其中 $f(\lambda_i)$ 表示对特征值进行幂运算:

$$f(\lambda_i) = \lambda_i^\alpha \quad (6)$$

到此,协方差池化层的前向传播已经推导结束.

2.2.2 后向传播推导

接下来本文对协方差池化层的后向传播过程进行简要的推导.首先通过全连接层反传得到 $\mathbf{K} = \{K_{ij}\}$, $K_{ij} = \begin{cases} 1/(\lambda_i - \lambda_j), & i \neq j \\ 0, & i = j \end{cases}$,关于 \mathbf{U} 和 $\mathbf{\Lambda}$ 的链式传递过程描述如下:

$$\text{tr} \left(\left(\frac{\partial l}{\partial \mathbf{U}} \right)^T d\mathbf{U} + \left(\frac{\partial l}{\partial \mathbf{\Lambda}} \right)^T d\mathbf{\Lambda} \right) = \text{tr} \left(\left(\frac{\partial l}{\partial \mathbf{Q}} \right)^T d\mathbf{Q} \right) \quad (7)$$

其中 l 表示损失函数,根据式(7),可得:

$$\frac{\partial l}{\partial \mathbf{P}} = \left(\frac{\partial l}{\partial \mathbf{Q}} + \left(\frac{\partial l}{\partial \mathbf{Q}} \right)^T \right) \mathbf{U} \mathbf{F} \quad (8)$$

$$\frac{\partial l}{\partial \mathbf{X}} = \alpha \left(\text{diag}(\lambda_1^{\alpha-1}, \dots, \lambda_d^{\alpha-1}) \mathbf{U}^T \frac{\partial l}{\partial \mathbf{Q}} \right)_{\text{diag}}$$

依照链式求导法则,响应的 $\frac{\partial l}{\partial \mathbf{P}}$ 和 $\frac{\partial l}{\partial \mathbf{X}}$ 求解如下:

$$\frac{\partial l}{\partial \mathbf{P}} = \mathbf{U} \left(\left(\mathbf{K}^T \circ \left(\mathbf{U}^T \frac{\partial l}{\partial \mathbf{U}} \right) \right) + \left(\frac{\partial l}{\partial \mathbf{Q}} \right)_{\text{diag}} \right) \mathbf{U}^T \quad (9)$$

$$\frac{\partial l}{\partial \mathbf{X}} = \bar{\mathbf{I}} \mathbf{X} \left(\frac{\partial l}{\partial \mathbf{P}} + \left(\frac{\partial l}{\partial \mathbf{P}} \right)^T \right)$$

其中 \circ 表示矩阵 Kronecker 积, $\mathbf{K} = \{K_{ij}\}$, K_{ij} 表示为:

$$K_{ij} = \begin{cases} 1/(\lambda_i - \lambda_j), & i \neq j \\ 0, & i = j \end{cases} \quad (10)$$

到此,协方差池化层的后向传播已经推导结束.

2.3 与 STRCF 结合

为了验证经过二阶池化处理后的深度特征在跟踪任务上的有效性,本文将其与已有的 STRCF 算法进行结合. STRCF 在原有的相关滤波框架下增加了时间正则和空间正则项,有效的缓解了相关滤波的边界效应问题,目标函数如下所示:

$$E(f) = \left\| \sum_{l=1}^d \mathbf{x}^l * \mathbf{f}^l - \mathbf{y} \right\|_2^2 + \lambda \sum_{l=1}^d \|\mathbf{w} \mathbf{f}^l\|^2 + \mu \|\mathbf{f} - \mathbf{f}_{i-1}\|_2^2 \quad (11)$$

其中,增加的 \mathbf{w} 是表示空间正则项,用来约束滤波器系数的响应范围. μ 是时间正则化系数, \mathbf{f}_{i-1} 表示上一帧的滤波器模板,作为时间约束控制滤波器的更新程度,在遮挡、快速变化等场景下可以有效应对.

2.4 算法流程

本文跟踪算法的主要流程如图 1 所示. 首先通过 ImageNet 和大量细粒度图像数据集对构建的网络进行训练,训练好后仅保留前四层网络参数作为特征提取

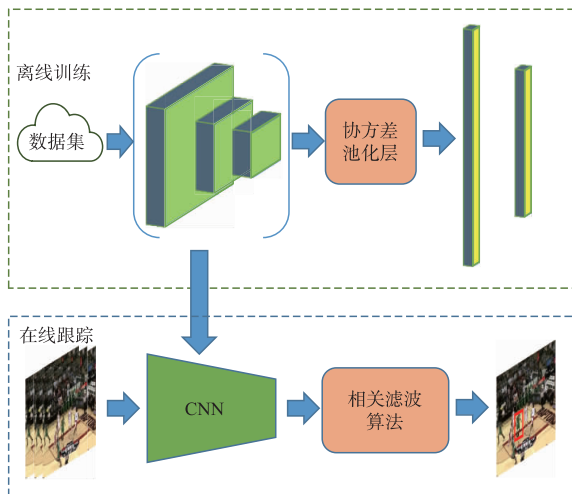


图1 算法流程图

器. 在跟踪过程中,为了减轻运算负担,仅采用第四层的特征作为目标的外观表示,通过和 STRCF 算法进行结合,在大规模数据集上取得了非常好的性能表现. 在尺度估计上,本文算法通过多尺度搜索的方式确定目标当前的最佳尺度. 由于在模型中添加了时间正则项,所以不用再进行模型的滑动平均更新.

3 仿真实验

为验证本文算法的有效性,在 Windows10 操作系统下,采用 MATLAB 和 C++ 混合编程实现了本文算法,并采用了 MatConvNet 工具箱^[18]构建改进的基于协方差池化的 VGG16 网络^[19],在传统的 VGG16 网络结构基础上,将网络的最后一个一阶池化层修改为协方差池化层,其他层结构和参数保持不变. 本文采用图像分类的 ImageNet2012 分类数据集以及细粒度图像分类领域广泛使用的 CUB200-2011 数据集作为网络的训练集. 在 Intel Xeon 2.4 GHz 的处理器上对本文算法进行训练和测试,并利用 GPU (TITANXp) 进行加速.

测试参数设置如下:在网络训练阶段,针对最后一层卷积层维度过高的问题,本文采用 1×1 卷积的方式将最后一层卷积层的特征维度从 512 维降到 256 维,接着再送入协方差池化层进行处理. 在算法测试阶段,选取网络的第四层特征作为目标的外观表示,提取的特征经过余弦窗处理以消除边界的不连续性. 目标搜索框大小取以目标为中心,边长为 $\sqrt{5WH}$ 的正方形区域,其中 W 和 H 分别表示目标的长和宽. 对于式(11)中的正则化参数 $\lambda = 1, \mu = 16$. 尺度估计采用和文献[17]相同的参数设置.

本文在 OTB2015 数据集^[20]上将所提算法与其他 10 种主流且相关的跟踪算法进行比较,这些方法包括:

- (1) 基于预训练深度特征与相关滤波算结合的方法: MCPF^[21], DeepSTRCF^[17], DeepSRDCF^[9], HDT^[8], HCF^[7].
- (2) 基于在线微调深度特征与相关滤波结合的方法: MDNet^[10].
- (3) 基于端到端学习的跟踪方法: GradNet^[22], Siam-RPN^[23], SiamFC^[11].
- (4) 基于传统手工特征的相关滤波方法: STRCF^[17].

3.1 算法整体性能

图 2 显示了本文算法和对比算法的跟踪精度和成功率的对比曲线图. 对比算法均是近些年取得最好性能的跟踪算法,其中 MDNet 采用离线训练加在线微调的方式,在多个数据集上均取得了最好的性能表现,但是 MDNet 不到 1FPS 的跟踪速度使得应用极为受限. 本文算法在精度上与 MDNet 持平,在成功率上

获得了 1.5% 的性能提升,同时速度是 MDNet 的 10 倍左右. 本文算法相比于基于预训练一阶网络的 HCF、HDT、DeepSRDCF、MCPF 等算法有着明显的性能优势,且速度与之相当. 对比算法中的 GradNet 和 SiamRPN 是近两年提出的基于孪生网络跟踪算法,在速度

和精度上都有着很好的表现,虽然本文算速度稍慢,但是在精度和成功率上依然有近 4% 的性能优势. 最后和本文算法所依托的 STRCF 算法相比,在精度上提升了 3.8%,在成功率上提升了 3.4%,均表现出了明显的性能提升.

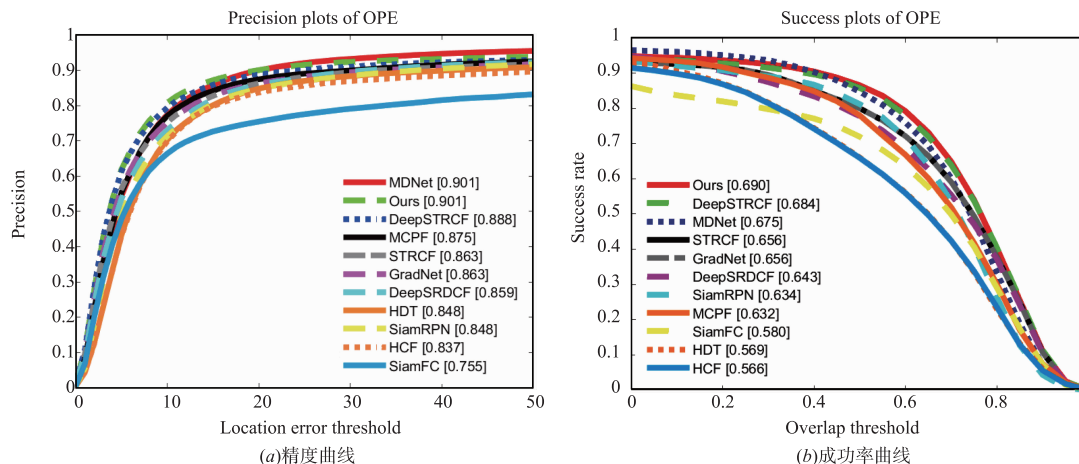
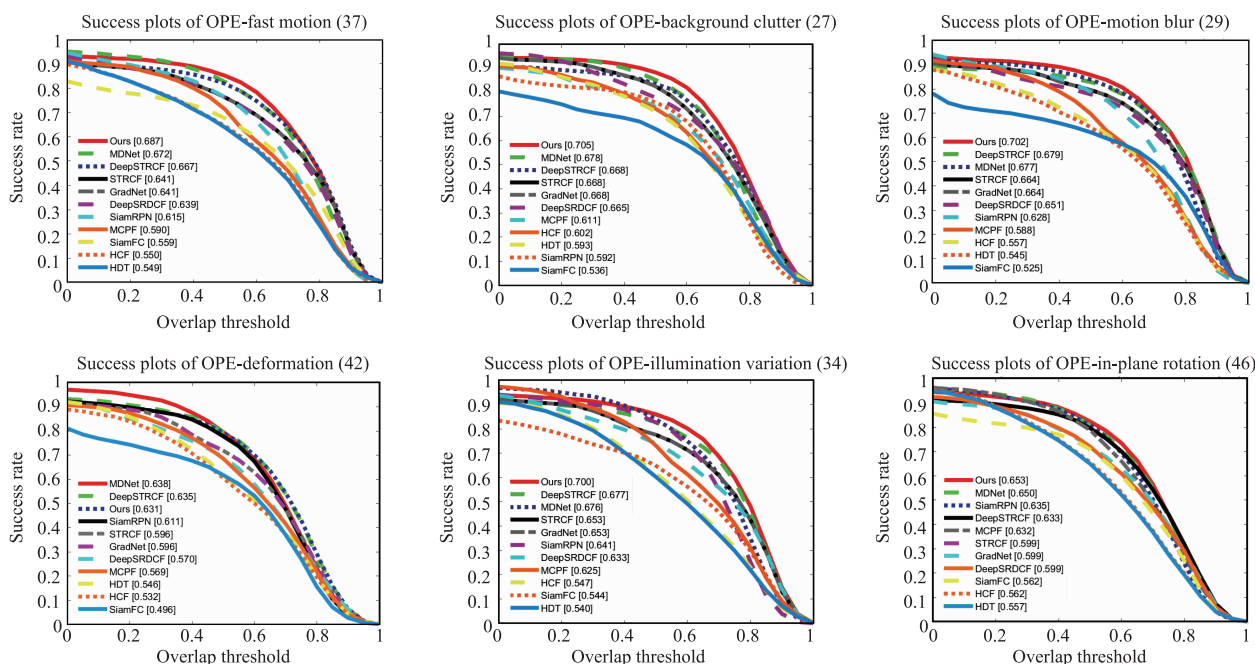


图2 OTB2015测试结果的精度曲线和成功率曲线

3.2 算法各属性性能分析

为了更加细致地分析本文算法在各类跟踪场景下的性能表现,采用 OTB2015 数据集中的 11 个标注属性对本文算法与对比算法进行分析,图 3 列出了各属性下各个算法的成功率对比图,红色代表最优结果,绿色代表次优结果. 从图 3 中可以看出,本文算法在几乎所有属性上均取得了最优或次优的跟踪结果. 尤其在应对背景杂波 (Background Clutters, BC) 方面,本文算法在

成功率上达到了 0.705,远优于类似的基于深度特征的跟踪算法 (HCF、HDT 等). 这进一步验证了本文所采用的二阶池化网络对相似物体干扰也具有较好的区分能力. 另外,在低分辨率目标的跟踪上,本文算法在成功率上为 0.702,主要得益于本文的特征提取网络在细粒度图像数据集上进行了进一步的训练,提升了网络的表达能力. 在诸如光照变化、旋转、运动模糊等复杂场景下,本文算法均取得了最优的跟踪精度和成功率.



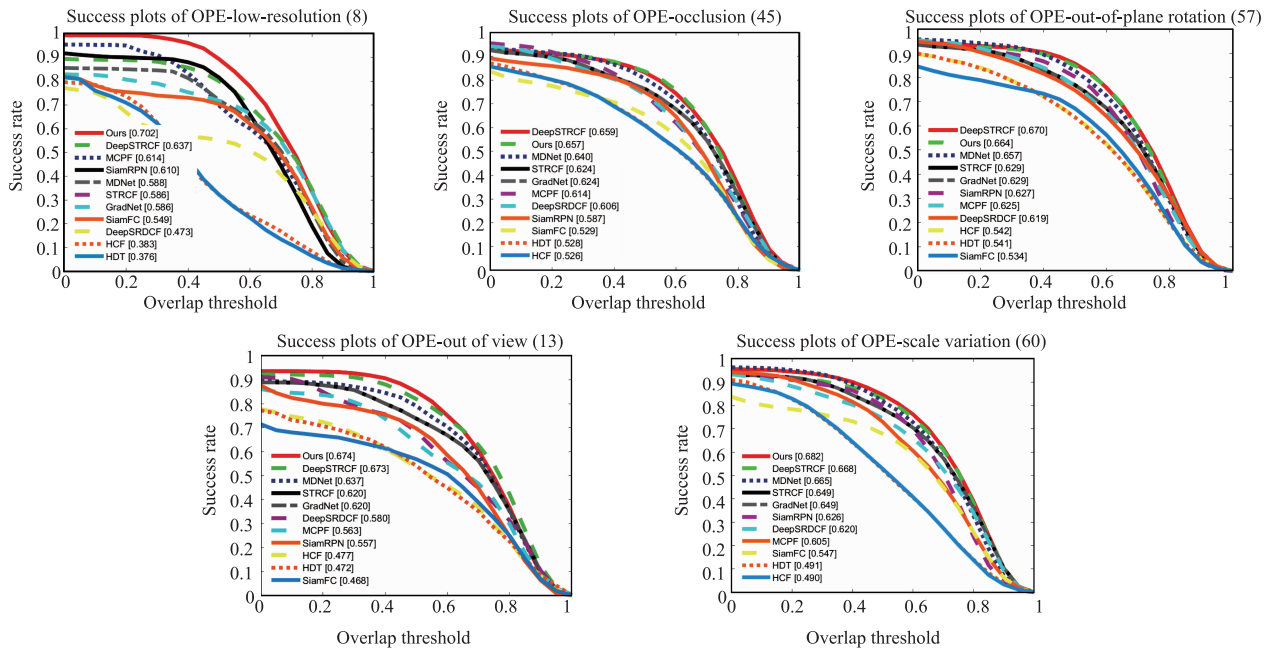


图3 不同属性下算法的跟踪成功率对比曲线

3.3 不同卷积层特征对跟踪性能的影响

为了进一步分析高阶池化网络不同卷积层特征对跟踪性能的影响,本文提取各层特征在 OTB2015 数据集上做了 5 组对比实验,实验结果如表 1 所示。

从表 1 中可以看出,对跟踪性能最好的是第四层特征,最差的是第二层特征,性能下降较多。另外,为了运算效率,本文并未采用和 HCF 与 HDT 类似的多层特征融合的方法。

表 1 不同卷积层特征对跟踪性能影响对比实验

	Conv1	Conv2	Conv3	Conv4	Conv5
成功率	0.651	0.645	0.663	0.690	0.679
跟踪精度	0.864	0.857	0.875	0.901	0.892

3.4 与一阶池化网络的性能对比

为了进一步验证二阶池化网络的性能优势,本文在 STRCF 算法基础上分别采用二阶池化网络和一阶池化网络进行测试,实验结果如图 4 所示。如图所示,相比于采用一阶池化网络的 DeepSTRCF 算法,本文算法在跟踪精度和成功率上分别提高了 1.3% 和 0.6%,进一步证明了二阶网络有着更好的跟踪性能。

3.5 在 TC128 数据集上的性能测试

为了充分验证算法的跟踪性能,在 Temple-Color128 数据集^[24]上对本文算法进行评测。TempleColor128 包括 128 组不同跟踪场景下的彩色视频序列,本文采用和 OTB 数据集相同的评测指标对算法的性能进行比较。图 5 显示了 6 个跟踪算法的精度和成功率曲线。在这些对比算法中,本文算法在两个指标上都获得

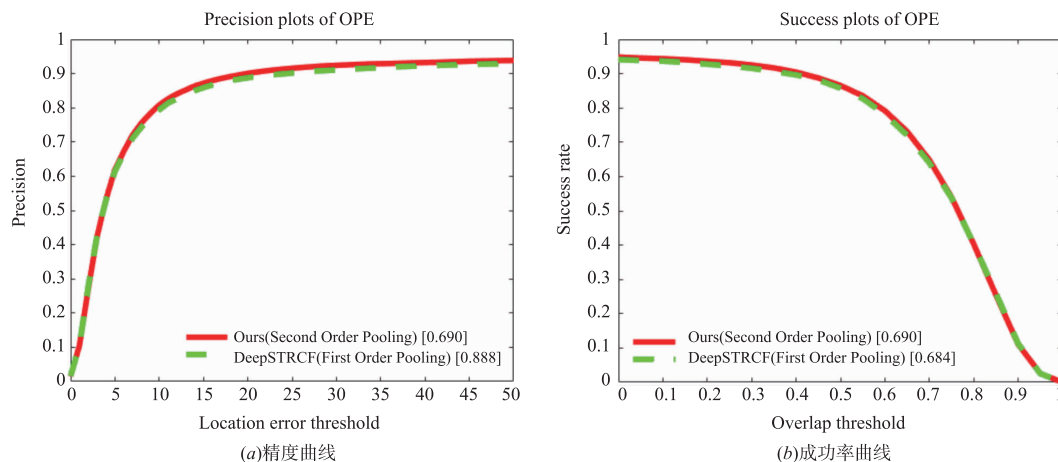


图4 二阶池化网络和一阶均值池化的性能对比

了最好结果 (0.785, 0.584), 优于 C-COT 算法^[25]. 同时, 与 DeepSRDCF 算法相比, 本文算法在精度上获得了 4.7% 的性能提升, 成功率上获得了 5% 的提升, 并且

显著优于 HCF 算法. 总体而言, 本文所提算法在 TempleColor128 数据集上与最先进的跟踪算法相比, 依然具有良好的性能竞争力.

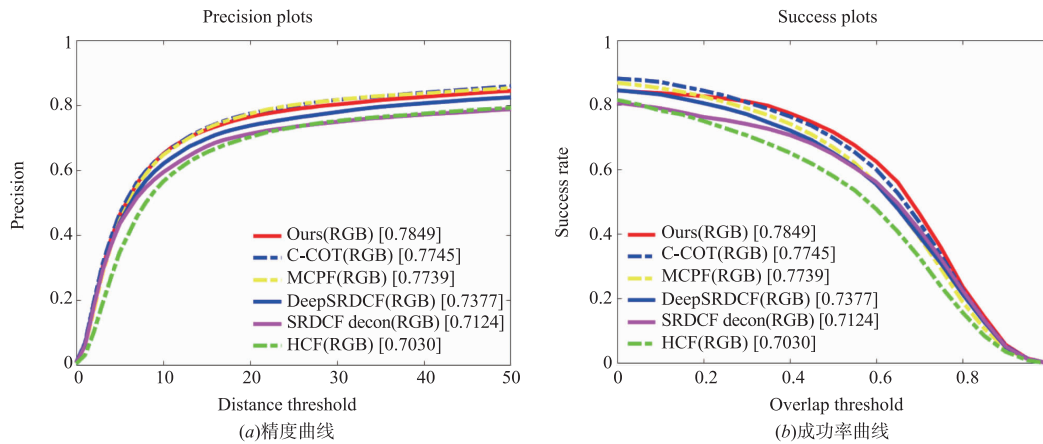


图5 TempleColor128测试结果的精度曲线和成功率曲线

4 结论

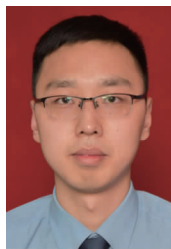
本文提出了一种基于高阶池化网络的鲁棒视觉跟踪算法. 通过协方差池化层和 VGG 网络的结合, 提高了预训练网络的判别能力, 并首次将高阶网络用于视觉跟踪任务. 实验结果也证明了本文算法的有效性, 在大规模数据集上取得了最好的性能表现. 由于端到端的思想在跟踪任务上的成功应用, 后续将继续研究如何在孪生网络结构中对深度特征进行高阶建模, 以增强网络的特征判别能力.

参考文献

- [1] Smeulders A W M, CHu D M, Cucchiara R, et al. Visual tracking: An experimental survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 36(7): 1442 – 1468.
- [2] Wang Naiyan, SHi Jianping, Yeung Dityan, et al. Understanding and diagnosing visual tracking systems[A]. IEEE International Conference on Computer Vision[C]. Santiago, Chile; 2015. 3101 – 3109.
- [3] Rawat W, Wang Z. Deep convolutional neural networks for image classification: A comprehensive review[J]. Neural computation, 2017, 29(9): 2352 – 2449.
- [4] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[A]. Proceedings of the IEEE conference on computer vision and pattern recognition. Columbus[C]. America; 2014. 580 – 587.
- [5] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[A]. Proceedings of the IEEE conference on computer vision and pattern recognition[C]. Boston, America; 2015. 3431 – 3440.
- [6] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[A]. IEEE Conference on Computer Vision and Pattern Recognition[C]. Las Vegas, America; 2016. 770 – 778.
- [7] Ma Chao, Huang Jiabin, YANG Xiaokang, et al. Hierarchical convolutional features for visual tracking[A]. IEEE International Conference on Computer Vision[C]. Santiago, Chile; 2015. 3074 – 3082.
- [8] Qi Yuankai, Zhang Shenging, Qin Lei, et al. Hedged deep tracking[A]. Proceedings of the IEEE conference on computer vision and pattern recognition[C]. Las Vegas, America; 2016. 4303 – 4311.
- [9] Danelljan M, Hager G, Shahbaz Khan F, et al. Convolutional features for correlation filter based visual tracking[A]. Proceedings of the IEEE International Conference on Computer Vision Workshops[C]. Santiago, Chile; 2015. 58 – 66.
- [10] Nam H, Han B. Learning multi-domain convolutional neural networks for visual tracking[A]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition[C]. 2016. 4293 – 4302.
- [11] Bertinetto L, Valmadre J, Henriques J F, et al. Fully-convolutional siamese networks for object tracking[A]. European conference on computer vision[C]. Springer, Cham; 2016. 850 – 865.
- [12] Gao Z, Xie J, Wang Q, et al. Global second-order pooling convolutional networks[A]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition[C]. Los Angeles, USA; 2019. 3024 – 3033.
- [13] Carreira J, Caseiro R, Batista J, et al. Semantic segmentation with second-order pooling[A]. European Conference

- on Computer Vision [C]. Berlin, Heidelberg: Springer, 2012. 430 – 443.
- [14] Chen B, Deng W, Hu J. Mixed high-order attention network for person re-identification [A]. Proceedings of the IEEE International Conference on Computer Vision [C]. Los Angeles, USA: 2019. 371 – 381.
- [15] Li P, Xie J, Wang Q, et al. Is second-order information helpful for large-scale visual recognition? [A]. In Proceedings of the IEEE International Conference on Computer Vision [C]. Venice: 2017. 2070 – 2078.
- [16] Krizhevsky A, Sutskever I, Hinton G. Imagenet classification with deep convolutional neural networks [A]. International Conference on Neural Information Processing Systems [C]. Lake Tahoe, Spain: 2012, 25 (2): 809 – 817.
- [17] Li Feng, Tian Cheng, Zuo Wangmeng, et al. Learning spatial-temporal regularized correlation filters for visual tracking [A]. Computer Vision and Pattern Recognition [C]. Salt Lake City, USA: 2018. 4904 – 4913.
- [18] Vedaldi A, Lenc K. Matconvnet: Convolutional neural networks for matlab [A]. Proceedings of the 23rd ACM international conference on Multimedia. Brisbane [C]. Australia: 2015. 689 – 692.
- [19] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [A]. International Conference on Learning Representations [C]. San Diego, USA: 2015. 563 – 567.
- [20] Wu Y, Lim J, Yang M H. Object tracking benchmark [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37 (9): 1834 – 1848.
- [21] T Zhang, C Xu, M Yang. Learning multi-task correlation particle filters for visual tracking [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 41 (2): 365 – 378.
- [22] LI P, CHEN B, OUYANG W, et al. GradNet: Gradient-guided network for visual object tracking [A]. Proceedings of the IEEE International Conference on Computer Vision [C]. Venice: 2019. 6162 – 6171.
- [23] Li B, Yan J, Wu W, et al. High performance visual tracking with siamese region proposal network [A]. Computer Vision and Pattern Recognition [C]. Salt Lake City, USA: 2018. 8971 – 8980.
- [24] Liang Pengpeng, Erik B, Ling Haibin. Encoding color information for visual tracking: Algorithms and benchmark [J]. IEEE Transactions on Image Processing, 2015, 24 (12): 5630 – 5644.
- [25] Danelljan M, Robinson A, Khan F S, et al. Beyond correlation filters: Learning continuous convolution operators for visual tracking [A]. European Conference on Computer Vision. Springer [C]. Cham: 2016. 472 – 488.

作者简介



蒲 磊 (通信作者) 男, 1991 年 9 月出生, 四川遂宁人, 分别于 2014 年和 2017 年获空军工程大学信息与导航学院信息与通信工程专业学士学位和硕士学位, 现于空军工程大学研究生院攻读博士学位, 主要研究方向为计算机视觉与目标跟踪。

E-mail: warmstoner@163.com



冯新喜 男, 1964 年 10 月出生, 陕西富平人 1991 年获西北工业大学博士学位, 现为空军工程大学信息与导航学院教授、博士研究生导师, 主要研究领域为信息融合, 信号处理, 目标跟踪等。

E-mail: fengxinxi2005@aliyun.com