

# 应用LTRNet卷积特征的ECO目标跟踪算法改进

陈志旺<sup>1,2</sup>, 王莹<sup>1†</sup>, 宋娟<sup>3</sup>, 姚权允<sup>1</sup>, 彭勇<sup>4</sup>

(1. 燕山大学 智能控制系统与智能装备教育部工程研究中心, 河北 秦皇岛 066004;

2. 燕山大学 工业计算机控制工程河北省重点实验室, 河北 秦皇岛 066004;

3. 国网黑龙江省电力有限公司佳木斯供电公司, 黑龙江 佳木斯 154002; 4. 燕山大学 电气工程学院, 河北 秦皇岛 066004)

**摘要:** 本文提出了一种基于应用高效卷积算子(ECO)改进的LRECT跟踪算法。首先, 为了增强网络所提取特征的认识能力, 堆叠线性两步(LT)残差结构设计具有32层的线性两步方法性质的残差网络(LTRNet), 并且融合该网络浅层与深层卷积特征信息形成跟踪算法的特征提取模块; 其次, 采用投影矩阵压缩LTRNet提取的高维特征, 将压缩特征通过插值处理后, 与当前滤波器在傅里叶域进行卷积定位确定目标位置; 最后, 使用高斯牛顿算法和共轭梯度算法求解以响应误差和惩罚项之和为优化目标的优化问题, 实现滤波器和投影矩阵的更新。在OTB2015标准数据集上进行测试实验, 结果表明本文所提算法可以实现较高精度的稳健性跟踪。

**关键词:** 目标跟踪; 应用高效卷积算子(ECO); 具有线性两步方法性质的残差网络(LTRNet); 特征压缩; 目标优化

**引用格式:** 陈志旺, 王莹, 宋娟, 等. 应用LTRNet卷积特征的ECO目标跟踪算法改进. 控制理论与应用, 2020, 37(12): 2601 – 2610

DOI: 10.7641/CTA.2020.00213

## The application of LTRNet convolution features in the improvement of ECO

CHEN Zhi-wang<sup>1,2</sup>, WANG Ying<sup>1†</sup>, SONG Juan<sup>3</sup>, YAO Quan-yun<sup>1</sup>, PENG Yong<sup>4</sup>

(1. Engineering Research Center of the Ministry of Education for Intelligent Control System and Intelligent Equipment, Yanshan University, Qinhuangdao Hebei 066004, China;

2. Key Laboratory of Industrial Computer Control Engineering of Hebei Province, Yanshan University, Qinhuangdao Hebei 066004, China;

3. Jiamusi Electric Power Company, State Grid Heilongjiang Electric Power Co., Ltd, Jiamusi Heilongjiang 154002, China;

4. School of Electrical Engineering, Yanshan University, Qinhuangdao Hebei 066004, China)

**Abstract:** An improved LRECT tracking algorithm based on ECO (efficient convolution operators for tracking) is proposed in this paper. Firstly, in order to enhance the extracted features for object recognition, a 32-convolution layers LTRNet (linear two-step residual network) composed of LT (linear-two step) residual structures is designed as the features extraction module of LRECT, and its shallow and deep features information are fused. Secondly, the projection matrix is used to compress the high-dimensional features extracted by LTRNet, and the compressed features are interpolated, then they are convoluted with the current filter to get the object location in the Fourier domain. Finally, filter and projection matrix are updated by Gauss-Newton and conjugate gradient methods which are applied to solve the optimization problem, and the optimization goal is the sum of response errors and penalties. The results of experiments on the OTB2015 dataset show that the proposed algorithm can achieve robust tracking with a higher accuracy.

**Key words:** object tracking; ECO; LTRNet; feature compression; objective optimization

**Citation:** CHEN Zhiwang, WANG Ying, SONG Juan, et al. The application of LTRNet convolution features in the improvement of ECO. *Control Theory & Applications*, 2020, 37(12): 2601 – 2610

## 1 引言

目标跟踪是机器视觉当前研究的热点方向, 也是一项非常具有挑战性的工作。目前, 目标跟踪技术在

高级人机交互、安全监控和行为分析等方面具有潜在的经济价值和广泛的应用前景<sup>[1]</sup>。因此, 需深入研究、设计高性能的目标跟踪算法。

收稿日期: 2020-04-25; 录用日期: 2020-08-10.

†通信作者. E-mail: 1718819591@qq.com; Tel.: +86 17333501243.

本文责任编辑: 杨辰光.

国家自然科学基金项目(61573305)资助.

Supported by the National Natural Science Foundation of China (61573305).

近年来随着深度学习的不断发展,卷积神经网络强大的特征学习能力引起了国内外专家学者的广泛关注<sup>[2]</sup>. 基于深度学习的跟踪方法因精度高、性能稳定逐步成为了目标跟踪发展的主流方向. 2015年, VGGNet<sup>[3]</sup>网络模型的出现将跟踪算法的精度提升到了一个新高度. Danelljan等提出的连续卷积跟踪算法(learning continuous convolution operators for visual tracking, C-COT)<sup>[4]</sup>, 采用多层VGGNet卷积特征, 通过连续空间域插值转换操作, 可以将不同分辨率的特征输入滤波器. 2017年, Danelljan等以C-COT为基础, 提出了一种具有高效卷积特性的跟踪算法(efficient convolution operators for tracking, ECO)<sup>[5]</sup>, 该算法从模型大小、训练集大小以及更新策略3个方面进行改进. 在OTB2015<sup>[6]</sup>上的测试结果很长一段时间排名第一. 随后, He等基于ECO提出了具有加权卷积响应的(correlation filters with weighted convolution responses, CFWCR)<sup>[7]</sup>改进算法, 此算法使用VGGNet卷积特征的多尺度相关滤波跟踪方案, 通过加权处理深、浅卷积特征的卷积响应结果提升了跟踪精度. 2018年, Goutam等将ECO中的VGG16<sup>[3]</sup>换成了ResNet50<sup>[8]</sup>, 并加入数据增强处理提取目标特征, 提出了可以有效利用深度卷积网络进行跟踪的(unveiling the power of deep tracking, UPDT)<sup>[9]</sup>跟踪算法, 先对不同层次特征进行分治, 然后, 自适应计算两种特征的响应加权系数并加权融合, 有效提高了深度特征下的跟踪算法性能. 此外, 在VOT2018<sup>[10]</sup>短时跟踪公开测试集挑战中, 具有多分辨率融合特性的(multi-solution fusion for visual tracking, MFT)算法, 通过针对不同分辨率特征独立求解, 并根据跟踪视频的难易程度自适应优化融合多个解来预测目标位置, 使算法鲁棒性排名第一.

以上算法除C-COT是ECO算法的基础外, 其他都可看作是ECO算法的改进. 这些算法主要利用深度卷积网络提取强表征能力的目标特征提高跟踪精度, 且网络越深, 精度越高. 但随着卷积网络层数的增加, 训练网络愈加复杂, 目标特征提取效率逐渐降低, 加上要对高维数据进行计算, 开销较大, 最终导致跟踪实时性差.

因此, 本文在文献[5]的基础上, 从深度卷积网络结构出发, 提出了改进的ECO算法LRECT(LTRNet for efficient convolution tracker). 首先, 引入了Dong等人<sup>[11]</sup>因受常微分方程(ordinary differential equation, ODE)<sup>[12]</sup>线性多步数值方法启发提出的LT残差结构, 堆叠LT形成了利于目标跟踪的LTRNet(LT-ResNet)深度卷积网络; 其次, 利用该网络提取目标特征, 并通过投影矩阵压缩特征, 插值处理提高特征的亚像素精度. 之后, 在傅里叶域将处理后的特征与当前滤波器进行卷积定位计算. 最后, 利用高斯牛顿算法<sup>[13]</sup>和共轭梯度算法<sup>[13]</sup>联合求解优化目标更新滤波器和投影矩阵.

## 2 基础理论

### 2.1 ECO跟踪算法

ECO<sup>[5]</sup>是一种具有高效卷积特性的跟踪算法, 其提取目标的VGGNet<sup>[3]</sup>, HOG (histogram of oriented gradients), CN (color-names)特征用于相关滤波框架实现目标定位和滤波器更新.

首先, 沿用C-COT<sup>[4]</sup>的插值方法对检测样本的目标搜索区域特征 $x$ 进行如下式所示的插值运算:

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

式中:  $x^d$ 表示 $x$ 的 $d$ 通道特征( $d \in [1, D]$ );  $J_d\{x^d\}(t)$ 是一个关于 $t \in [0, T)$ 的函数, 表示 $x^d$ 的插值运算结果;  $x^d[n] \in \mathbb{R}^{N_d}$ 是关于 $n \in \{0, \dots, N_d - 1\}$ 的函数;  $N_d$ 是其分辨率;  $b_d(t - \frac{T}{N_d}n)$ 是 $d$ 通道的插值函数. 用 $J\{x\}(t) \in \mathbb{R}^D$ 表示的1至 $D$ 通道的插值结果, 简记为 $J\{x\}$ . 其次, 使用主成分分析法<sup>[14]</sup>(principal component analysis, PCA)简化滤波器, 然后与 $J\{x\}$ 进行卷积计算求响应得分 $S_{Pf}\{x\}$ :

$$S_{Pf}\{x\} = Pf * J\{x\} = f * P^T J\{x\}, \quad (2)$$

式中:  $f$ 是通道数为 $D$ 的滤波器, “\*”是卷积运算符,  $P$ 是 $D$ 行 $C$ 列的投影矩阵,  $P^T$ 是其转置矩阵. 对 $S_{Pf}\{x\}$ 使用高斯牛顿算法<sup>[13]</sup>进行优化, 可找到最大得分值的位置(目标的新位置). 最后, 先使用高斯混合模型(Gaussian mixture model, GMM)<sup>[15]</sup>方法压缩训练集. 之后, 对训练样本和当前滤波器 $f$ 的卷积响应得分 $S_{Pf}\{\mu_m\}$ 与训练样本高斯标签 $y_0$ 的误差取 $L^2$ 范数, 并添加惩罚项构造如下所示的损失函数:

$$E(f) = \sum_{m=1}^M \pi_m \|S_{Pf}\{\mu_m\} - y_0\|_{L^2}^2 + \sum_{c=1}^C \|\omega f^c\|_{L^2}^2, \quad (3)$$

式中:  $\mu_m$ 和 $\pi_m$ 分别是训练样本的均值和权重;  $M$ 是训练样本总个数;  $\omega$ 是 $f$ 的惩罚项, 此处使用的 $P$ 仅在第1帧中计算确定, 以后使用共轭梯度算法<sup>[13]</sup>每隔6帧求解一次式(3)更新 $f$ 时,  $P$ 保持不变.

综上, ECO通过减小滤波器、训练集以及滤波器更新频率降低了算法复杂度, 提升了跟踪速度. 但其忽略了深度VGGNet提取高维目标特征所消耗的时间. 对此, 本文从卷积网络结构出发, 进一步研究提升ECO跟踪性能的方法.

### 2.2 LT残差结构

针对深度卷积网络的可解释性问题, Dong Bin等人<sup>[11]</sup>从数学的角度将ResNet<sup>[8]</sup>等深度卷积网络理解为常微分方程ODE<sup>[12]</sup>显式欧拉的数值离散近似. 结合图1(a)可形象化地将ResNet残差结构Basic block写

为

$$u_{n+1} = u_n + f(t_n, u_n), \quad (4)$$

式中:  $u_n$ 和 $u_{n+1}$ 分别表示第 $n+1$ 个Basic block残差结构的输入和输出,  $f(t_n, u_n)$ 表示由 $t_n$ 和 $u_n$ 参数化了的两个卷积层(Conv).  $u'(t) = f(t, u(t))$ 的显式欧拉离散如下:

$$u_{n+1} - u_n = \Delta t f(t_n, u_n), \quad (5)$$

式中:  $u_n$ 和 $u_{n+1}$ 表示ODE在 $t_n$ 和 $t_{n+1}$ 处的数值解,  $\Delta t$ 表示步长,  $f(t_n, u_n) \approx u'(t_n)$ ,  $u'(t_n)$ 是函数 $u(t)$ 在 $t_n$ 处的导数. 对比式(4)–(5)可知, 式(4)是式(5)步长 $\Delta t = 1$ 时的一个特例.

然后, 为利用ODE的线性多步数值方法(式(6))可以提高当前数值解 $u_{n+1}$ 对精确解 $u(t_{n+1})$ 的近似精度的特性, 提出了LT残差结构(式(7)):

$$u_{n+k} = \sum_{i=0}^{k-1} \alpha_i u_{n+i} + \Delta t \sum_{i=0}^k \beta_i f(t_{n+i}, u_{n+i}), \quad (6)$$

$$u_{n+1} = (1 - K_n)u_n + K_n u_{n-1} + f(t_n, u_n), \quad (7)$$

式(6)中:  $k$ 表示步数,  $\alpha_i$ 和 $\beta_i$ 为标量参数,  $u_{n+k}$ 表示第 $k$ 步的预测数值解,  $f(t_{n+i}, u_{n+i})$ 表示 $u(t)$ 在 $t_{n+i}$ 处的导数近似, 且 $t_{n+i} = t_n + i\Delta t$ . 取式(6)中的参数

$$k = 2, \Delta t = 1, n = n - 1, \alpha_0 = K_n,$$

$$\alpha_1 = 1 - K_n, \beta_0 = 0, \beta_1 = 1, \beta_2 = 0,$$

可得到式(7). 式(7)中 $K_n \in \mathbb{R}$ 表示第 $n$ 个LT残差结构的可训练参数,  $u_{n-1}$ 表示第 $n$ 个Basic block残差结构的输入, 其他变量含义同式(4). 由 $u_{n-1}$ 的含义解释可知, LT残差结构可用于任何类似于ResNet的深度卷积网络, 本文称这一新网络为LTRNet.

图1(a)–1(b)分别是Basic block和LT残差结构原理图, 对比可知, 1个LT包含2个Basic block和1个训练参数 $K_n$ . 在求解 $u_{n+1}$ 时, 图1(b)较图1(a)多引入一历史信息项 $u_{n-1}$ , 且 $u_{n-1}$ 和 $u_n$ 分别配有 $K_n$ 和 $1-K_n$ 这样相互制约的训练权重, 使LT能够选择保留 $u_{n-1}$ 和 $u_n$ 的有效特征信息融入到 $u_{n+1}$ . 对比仅保留 $u_n$ 这一历史信息项的Basic block, LT使历史信息遗忘较慢, 更能得到充分利用.

此外, 由原理图还可知, 使用Basic block或LT堆叠而成的深度卷积网络训练的实质是进行残差学习. 分别将式(4)和式(7)在保留 $\Delta t$ 的情况下写成残差形式

$$\frac{u_{n+1} - u_n}{\Delta t} = f(t_n, u_n),$$

$$\frac{u_{n+1} - (1 - K_n)u_n - K_n u_{n-1}}{\Delta t} = f(t_n, u_n).$$

对上两残差等式的左侧部分在 $t_n$ 处做泰勒展开, 分别得到式(4)和式(7)忽略高阶项的残差式(8)–(9):

$$f(t_n, u_n) = u'_n, \quad (8)$$

$$f(t_n, u_n) = (1 + K_n)u'_n + (1 - K_n)\frac{\Delta t}{2}u''_n. \quad (9)$$

对比以上两式可知, 式(9)对残差 $f(t_n, u_n)$ 的逼近程度较式(8)提高了一个级别的精度. 综上可得: 相比于Basic block, 使用LT可以得到更高的深度卷积网络分类精度.

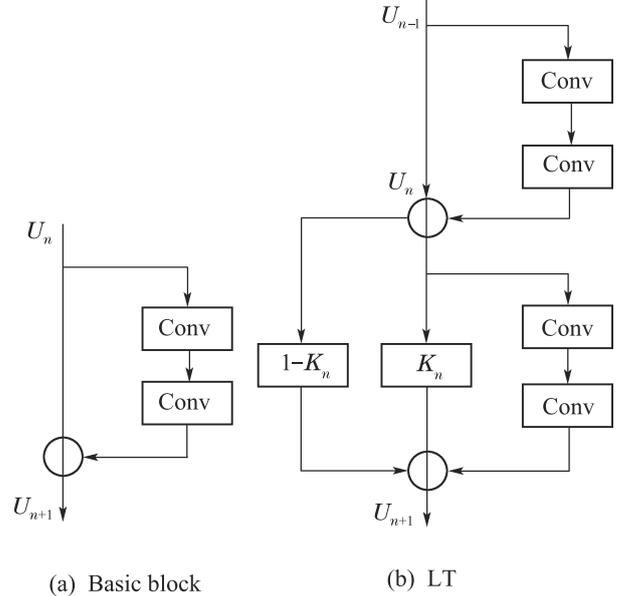


图 1 残差结构原理图

Fig. 1 Schematics of residual structures

### 3 LRECT跟踪算法

本文提出的是一种应用LTRNet卷积特征对ECO<sup>[5]</sup>进行改进的目标跟踪算法, 主要由特征提取模块、目标定位模块和滤波器更新模块3部分构成, 跟踪过程如图2所示(图中 $\otimes$ 表示卷积操作).

#### 3.1 特征提取模块

##### 3.1.1 LTRNet32网络

根据第2.2节中的理论分析和第5.3节的实验结果, 设计如图3所示的LTRNet32作为本文跟踪算法的特征提取模块. 由图可知该网络由1个卷积层(Conv), 4个卷积块(Layer 1, Layer 2, Layer 3, Layer 4), 3个最大池化层(Pool 1, Pool 2, Pool 3), 1个全局平均池化层(Gavp), 1个全连接层(Fc)和1个Softmax函数层组成. 若将含有2个卷积层的结构视为1个Block, Layer 1, Layer 2, Layer 3和Layer 4分别包含Block个数为1, 2, 8, 4. 考虑到新搭建的网络在使用前需要先进行训练, 对于Conv, 首先对输入依次进行BN归一化处理、ReLU激活和2D卷积操作的模式来解决网络训练初期误差下降缓慢的问题. 其次, 采用 $3 \times 3$ 的小卷积核, 这样所需训练参数少, 进行堆叠操作还能增加特征的多样性. 此外, 为了进一步减少网络训练参数, 在Layer之间使用1个 $3 \times 3$ 的最大池化层, 保留显著特征信息, 降低特征图维度; 对Layer 4的输出先进行全局平均池化, 再执行全连接操作. 最后加上Softmax函数形成了一个完整的LTRNet32图像分类网络.

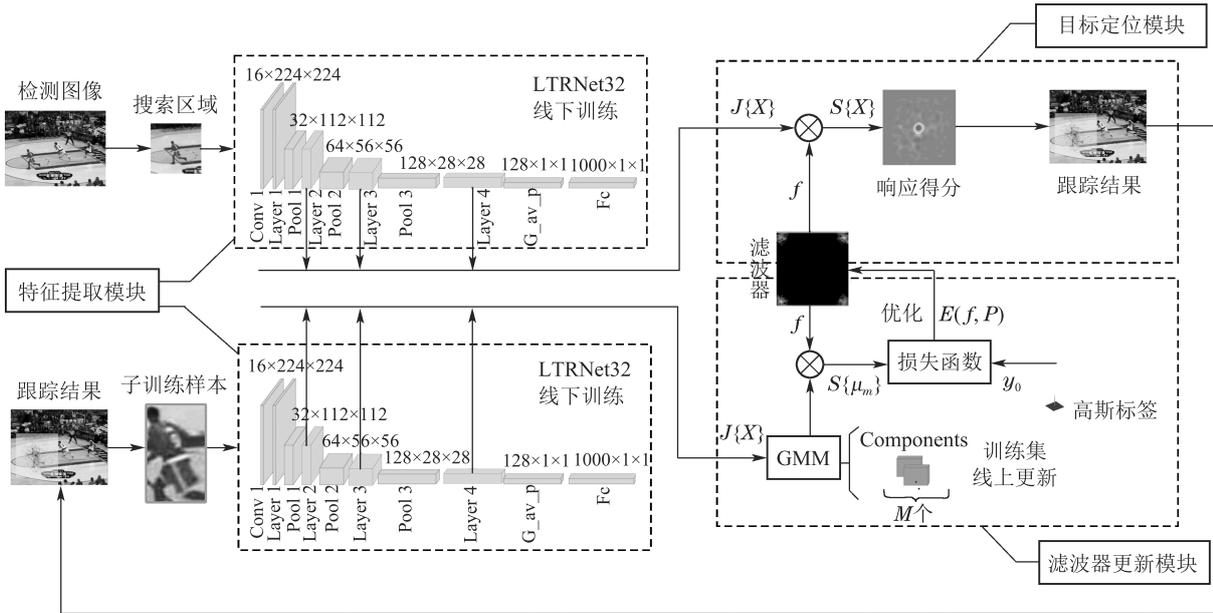


图2 LRECT跟踪算法框架图

Fig. 2 The framework of LRECT tracking algorithm

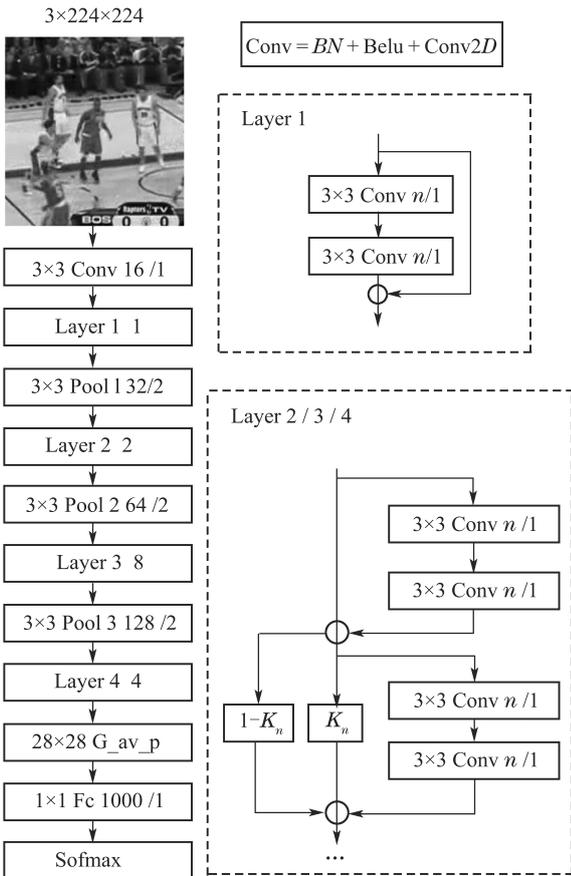


图3 LTRNet32网络结构

Fig. 3 Network structure of LTRNet32

3.1.2 卷积特征的选取

深度卷积网络不同层所提取的卷积特征侧重点不同, 浅层卷积特征保留了更多细粒度的纹理信息

以及空间信息; 深层卷积特征含有丰富的语义信息. 为了分析、选取有效的目标特征, 对LTRNet32的不同Layer层进行如图4所示的可视化输出, 图4(a)是检测图像中的目标搜索区域, 图4(b)–4(e)分别是Layer 1, Layer 2, Layer 3, Layer 4对图4(a)卷积输出的特征可视化图. 观察图像可知, 图4(b)中包含较多原始图像的轮廓和位置信息, 目标的有效特征信息占比少; 图4(c)包含较为细致的轮廓信息; 图4(d)包含高细粒度的纹理信息和部分语义信息; 图4(e)包含丰富的高级语义信息, 可以明显区分目标(人)和背景. 考虑到视频跟踪任务需要利用目标的空间信息进行定位, 高级语义信息应对目标外观变化, 本文舍弃Layer 1, 选取Layer 2, Layer 3, Layer 4分别提取32, 64, 128维的卷积特征, 用于后续目标定位和滤波器更新.

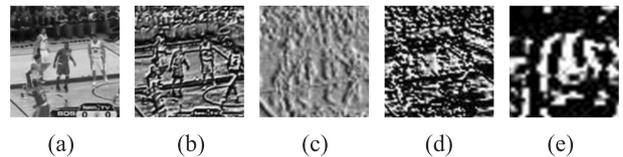


图4 LTRNet32网络Layers卷积输出可视化图  
Fig. 4 Convolution output visualizations of Layers in LTRNet32

3.2 目标定位模块

目标定位是实现目标跟踪的表现形式. 图2中的定位模块缺乏实现细节. 对此, 本节将从特征压缩、特征处理和卷积定位这三个方面对该模块进行

详细论述, 其实现框架如图5所示.

1) 特征压缩.

由于不同卷积核对图像的敏感方向不同, 所以不是每个通道提取的特征都是有用的. 图6是Layer 4 输出特征的128个通道的可视化图, 观察可知, 1通道和127通道提取了较多的背景信息; 65通道对图像不敏感, 几乎没有提取到有效信息; 128通道较2通道可以提取更为详细的目标信息. 所以, 可将一些类似于1通道和65通道等对目标定位贡献不大的

卷积通道舍去, 之后再用于目标定位, 这样可以减少算法计算量, 也不会影响所提取特征的代表能力. 假设 $x$ 是LTRNet32针对某帧检测图像的目标搜索区域提取的 $D$ 维卷积特征, 受式(2)启发, 使用投影矩阵 $P$ 对其进行压缩:

$$x^* = Px, \tag{10}$$

式中:  $x^*$ 表示压缩后的 $C$ 维特征, 即 $x^*$ 有 $C$ 个特征通道.  $P$ 是一个 $C$ 行 $D$ 列矩阵( $C \leq D$ ), 随滤波器一起更新计算, 其具体求解见第3.3节内容, 计算见式(17).

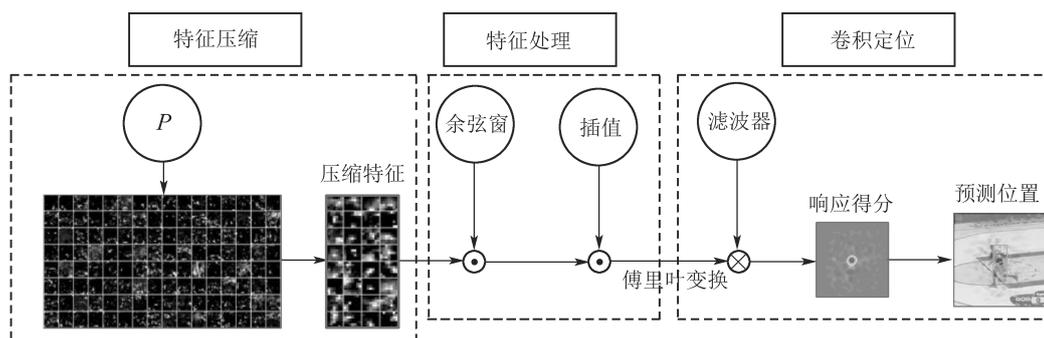


图 5 目标定位框架图

Fig. 5 Framework of object localization

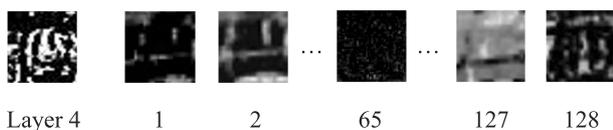


图 4 128个通道的特征可视化图

Fig. 6 Feature visualizations of 128 channels

2) 特征处理.

首先, 对特征压缩所得的 $x^*$ 添加余弦窗 $W_{\cos}$ 解决边界效应问题:

$$X = x^* \times W_{\cos}, \tag{11}$$

式中 $X = (X^1, \dots, X^C)$ . 取 $X$ 通项 $X^c (c \in [1, C])$ , 参考式(1), 对其进行如下式所示的可提高亚像素精度的逐通道插值运算:

$$J\{X\}(t) = \sum_{c=1}^C \sum_{n=0}^{N_c-1} X^c[n] b_c(t - \frac{T}{N_c}n), \tag{12}$$

式中:  $J\{X\}(t) \in \mathbb{R}^C$ 表示 $X$ 的 $C$ 个通道插值结果的叠加;  $X^c[n] \in \mathbb{R}^{N_c}$ 随离散变量 $n \in \{0, \dots, N_c-1\}$ 变化而变化,  $N_c$ 表示 $X^c$ 的分辨率;  $b_c(t - \frac{T}{N_c}n)$ 表示 $c$ 通道的插值核函数.

3) 卷积定位.

目标定位主要是将处理后的低维压缩特征 $J\{X\}$ 同当前滤波器 $f = (f^1, \dots, f^C)$ 进行卷积计算. 为了便于计算, 先利用傅里叶变换将 $J\{X\}$ 和 $f$ 转化到频域, 记为 $\hat{z}$ 和 $\hat{f}$ , 然后通过下式得到频域的

响应得分 $\hat{S}\{\hat{z}\}$ ,

$$\hat{S}\{\hat{z}\} = \hat{z} \hat{f} = \sum_{c=1}^C \hat{z}_c \hat{f}^c, \tag{13}$$

式中: “ $\wedge$ ”表示频域, 时域中的卷积计算可参考式(2). 取得分响应图中最大得分值的位置作为预测的目标位置, 定位结果如图5中预测位置所示.

3.3 滤波器更新模块

滤波器更新模块包含生成训练集和优化损失函数两部分, 其完整实现过程见图2下半部分. 训练集的生成策略: 将处理后的目标特征通过高斯混合模型(Gaussian mixture model, GMM)<sup>[15]</sup>聚类的方法合并入训练集进行训练样本在线更新(图2中GMM表示训练集, Components表示训练样本). 该方法区别于“单纯添加新训练样本, 盲目丢弃旧训练样本”的传统训练集更新策略, 将空间内相距较近, 且符合高斯分布的一组卷积特征视为一个训练样本, 通过在线更新这些卷积特征组实现训练集的更新.

优化损失函数就是更新滤波器 $f$ 和投影矩阵 $P$ . 为了降低因目标被遮挡滤波器发生漂移的概率, 本文沿用文献[5]的稀疏更新策略: 在使用第1帧检测图像给出的目标初始信息计算 $f$ 和 $P$ 之后, 每隔 $N_s$ 帧检测图像, 使用当前的Components训练样本进行一次 $f$ 和 $P$ 更新. 参考式(3)构建如下所示的损失函数:

$$E(f, P) = \sum_{m=1}^M \pi_m \|S\{\mu_m\} - y_0\|_{L^2}^2 + \sum_{c=1}^C \|\omega f^c\|^2 + \lambda \|P\|_F^2, \quad (14)$$

式中:  $M, y_0, C, \omega$  含义同式(3), 分别表示训练样本的总个数、高斯标签、滤波器通道数和  $f$  的惩罚项;  $\mu_m, \pi_m, S\{\mu_m\}$  分别表示第  $m$  个训练样本 Component 的均值和权重, 以及其与  $f$  的卷积响应得分;  $\|P\|_F^2$  表示  $P$  的  $F$  范数, 是  $P$  的惩罚项;  $\lambda$  为其权重系数。

优化求解式(14)完成  $f$  和  $P$  更新, 常用的主流方法有标准梯度下降法<sup>[13]</sup>或随机梯度下降法<sup>[13]</sup>等, 但这些算法收敛速度慢, 针对这一问题, 本文采用高斯牛顿算法<sup>[13]</sup>与共轭梯度算法<sup>[13]</sup>联合优化式(14)。

首先, 定义对应  $f = (f^1, \dots, f^C)$  和  $P$  的残差:

$$r_m(f, P) = \sqrt{\pi_m} (S\{\mu_m\} - y_0),$$

$$r(f, P) = \sum_m r_m(f, P).$$

其次, 引入高斯牛顿算法, 对  $r(f, P)$  进行一阶泰勒展开:

$$r(f + \Delta f, P + \Delta P) \approx r_{f,P} + J_{f,P} \Delta(f, P), \quad (15)$$

式中:

$$r_{f,P} = r(f, P), \quad J_{f,P} = \frac{\partial r}{\partial (f, P)}$$

是  $r$  在当前  $(f, P)$  下的雅可比式,  $\Delta(f, P)$  是关于滤波器增量  $\Delta f$  和投影矩阵增量  $\Delta P$  的表达式. 将式(15)代入式(14)得  $E(f, P)$  二阶高斯牛顿近似表达式:

$$\begin{aligned} \tilde{E}_{f,P}(\Delta f, \Delta P) = & \|r_{f,P} + J_{f,P} \Delta(f, P)\|_{L^2}^2 + \\ & \sum_{c=1}^C \|\omega(f^c + \Delta f^c)\|_{L^2}^2 + \lambda \|(P + \Delta P)\|_F^2. \end{aligned} \quad (16)$$

最后, 使用共轭梯度算法迭代优化式(16), 先得到  $\Delta f$  和  $\Delta P$ , 再通过下式更新  $f$  和  $P$ .

$$f = f + \Delta f, \quad P = P + \Delta P. \quad (17)$$

为了直观理解上述  $f$  和  $P$  的更新过程, 给出如下联合优化的伪代码, 其中  $N_{GN}$  代表高斯牛顿算法的迭代次数,  $u$  代表和  $r$  大小相同的向量,  $N_{CG}$  代表共轭梯度算法的迭代次数,  $p$  和  $\alpha$  分别代表每步优化的方向和步长,  $\beta$  代表更新动量。

联合优化伪代码:  $f$  和  $P$  的优化过程。

输入:  $f, P$ , 误差函数  $r(f, P)$ ,  $N_{GN}$ ,  $N_{CG}$ .

for  $i = 1, \dots, N_{GN}$  do

$r \leftarrow r(f, P), u \leftarrow r;$

$h \leftarrow r^T u;$

$g \leftarrow -h, p \leftarrow 0, \rho \leftarrow 1, \Delta f \leftarrow 0, \Delta P \leftarrow 0;$

for  $i = 1, \dots, N_{CG}$  do

$\rho_1 \leftarrow \rho, \rho \leftarrow g^T g, \beta \leftarrow \frac{\rho}{\rho_1};$

$p \leftarrow g + \beta p;$

$q \leftarrow J_{f,P}^T J_{f,P} p;$

$pq \leftarrow p^T q;$

$\alpha \leftarrow \frac{\rho}{pq};$

$g \leftarrow g - \alpha q;$

$\Delta f \leftarrow \Delta f + \alpha p, \Delta P \leftarrow \Delta P + \alpha p;$

end for

$f \leftarrow f + \Delta f, P \leftarrow P + \Delta P;$

end for

综合以上内容, 本文算法 LRECT 舍弃了 ECO 中的传统手工特征 HOG 和 CN, 仅依靠深度卷积网络 LTRNet32 提取目标特征. 为了使这一特征得以充分利用, 本文算法选择采用投影矩阵随滤波器一起进行稀疏更新的策略. 这是基于当前帧目标特征与临近帧中的目标特征相似性大这一基础理论, 稀疏更新投影矩阵能够使其适应视频序列不同帧中目标的不同状态, 保持较高效、准确的压缩特性. 使用这样压缩后的目标特征作为训练样本, 有利于学习得到更加稳定的滤波器。

## 4 算法步骤

LRECT 跟踪算法具体过程如下:

**步骤 1** 初始化算法参数: 初始化投影矩阵  $P$ , 滤波器  $f$ , 训练样本权重  $\pi_m$  等; 创建余弦窗口  $W_{\cos}$ ; 构建高斯标签函数  $y_0$ ; 设置特征压缩后维数  $C$  以及高斯牛顿和共轭梯度算法的迭代次数  $N_{GN}$  和  $N_{CG}$  等参数。

**步骤 2** 读取视频序列的第一帧检测图像, 使用其提供的目标初始信息裁剪出子训练样本。

**步骤 3** 将子训练样本输入 LTRNet32 网络提取卷积特征  $x$ . 先后利用式(10)–(12)对  $x$  进行压缩、加窗、插值处理得  $J\{X\}$ .

**步骤 4** 使用  $J\{X\}$  初始化训练集 GMM 中的第 1 个 Component。

**步骤 5** 采用高斯牛顿算法和共轭梯度算法联合优化式(14)得增量  $\Delta f$  和  $\Delta P$ , 利用式(17)处理两增量得到新滤波器  $f$  和投影矩阵  $P$ . 记录保存当前目标信息。

**步骤 6** 读取下一帧检测图像, 使用LTRNet 32提取其搜索区域的卷积特征 $x$ , 之后对 $x$ 做同步步骤3一样的压缩、加窗、插值处理得到 $J\{X\}$ .

**步骤 7** 使用傅里叶变换将 $J\{X\}$ 和当前滤波器 $f$ 转化到频域. 之后利用式(13)完成频域中的卷积计算, 得到卷积响应得分 $\hat{S}\{\hat{z}\}$ , 找到图中最高响应得分值的位置, 将其作为当前帧中目标位置进行保存.

**步骤 8** 根据上步得到的目标信息在当前帧中裁出子训练样本, 并对其执行同步步骤3的操作.

**步骤 9** 对 $J\{X\}$ 使用GMM方法更新训练集.

**步骤 10** 判断是否需要更新 $f$ 和 $P$ . 如果需要, 执行同步步骤5的操作.

**步骤 11** 判断是否跟踪完视频序列, 如果没有, 跳转至步骤6; 若跟踪完毕, 输出视频所有帧的目标信息, 保存结果.

## 5 实验结果与分析

### 5.1 实验平台

本文所做实验均在一台装有 1 张 Nvidia GTX 1080ti GPU 的计算机上进行, 处理器为 Intel core (TM) i7-8700K, 主频为 3.70 GHz, 内存为 32 GB, 操作系统为 64 位 Ubuntu 16.04, 编程环境为 python 3.6, 深度学习框架为 PyTorch.

### 5.2 实验参数设置

1) 网络训练参数设置: 超参数设置如表 1 所示.

表 1 超参数值

Table 1 Values of hyper-parameters

| 训练数据集<br>迭代次数 | 优化方法                   | 最小批训练<br>数据大小 |
|---------------|------------------------|---------------|
| 90            | 随机梯度下降 <sup>[13]</sup> | 128           |
| 学习率           | 权重更新动量                 | 权重衰减因子        |
| 0.01          | 0.9                    | 0.0001        |

2) LRECT跟踪算法参数设置: 使用训练好的 LTRNet32 作为特征提取模块, 设置其输入图像大小为  $224 \times 224$ , 选择 Layer 2, Layer 3, Layer 4 作为特征提取层, 设置特征压缩后维数为 8, 116, 32; 使用零初始化滤波器 $f$ , 随机方式初始化投影矩阵 $P$ ; 对于第 1 帧检测图像, 设置高斯牛顿算法和共轭梯度算法的优化迭代次数:  $N_{GN} = 6$ ,  $N_{CG} = 15$ . 以后每隔  $N_s = 6$  帧检测图像, 取  $N_{GN} = 1$ ,  $N_{CG} = 10$ . 此处未提及的需要初始化的参数均沿用文献[5]中的设置.

### 5.3 网络性能对比分析

为验证第 2.2 节中引入的 LT 残差结构和第 3.1.1 节中提取模块网络层数设计的有效性, 本节将通过堆叠 LT 残差结构形成的 LTRNet 与通过堆叠 Basic block 残差结构形成的 ResNet<sup>[8]</sup>, 先后在数据集 CIFAR10<sup>[16]</sup> 的训练集和验证集上进行训练与评估, 训练超参数除训练数据集迭代次数为 200 外, 其他设置同表 1. 实验细节: 首先, 采用 Lee 等人<sup>[17]</sup> 的方法对 CIFAR10 的训练集图像做预处理, 然后随机剪裁出  $32 \times 32$  大小的图像块用于网络训练. 当网络遍历学习一次训练集图像后, 以验证集图像的初始目标中心为中心裁剪出  $32 \times 32$  大小的图像块用于网络验证. 此外, LTRNet 中的参数  $k_n$  在集合  $[-0.1, 0.0]$  中随机取值进行初始化, 其他参数和 ResNet 网络参数同文献[11], 使用 PyTorch 框架下的随机方式进行初始化. 最终测试评估结果见表 2.

表 2 卷积网络在 CIFAR10 验证集上的错误率

Table 2 Error rates of convolution networks on CIFAR10 validation set

| 卷积层<br>个数 | 训练参<br>数量/兆 | 错误率/%  |        |
|-----------|-------------|--------|--------|
|           |             | ResNet | LTRNet |
| 20        | 0.27        | 12.21  | 11.56  |
| 32        | 0.46        | 10.97  | 10.41  |
| 44        | 0.66        | 10.63  | 9.88   |
| 56        | 0.85        | 10.42  | 9.75   |
| 110       | 1.7         | 9.53   | 8.82   |

由表 2 数据可知, LTRNet 和 ResNet 随着网络层数的加深, 训练参数量逐渐增加, 分类错误率逐渐降低; 在两者同层情况下, 参数量相同, 但 LTRNet 的错误率比 ResNet 平均要低 0.7% 左右; LTRNet56 几乎可以达到 ResNet110 的分类效果. 以上实验结果说明, LTRNet 是一种较 ResNet 拥有更高精度、更高效率的卷积网络模型. 此外, 对比 LTRNet 不同卷积层数的训练参数量和错误率, 发现当卷积层数从 20 到 32 时, 错误率下降了 1.15%, 之后随着网络层数的加深, 错误率下降并不显著. 结合这些数据并考虑跟踪的时效性, 本文算法选择 LTRNet32 作为特征提取模块.

### 5.4 LTRNet32 网络训练与特征选取结果分析

为了利用 LTRNet32 卷积网络提取强表征能力的目标特征, 先利用 ImageNet<sup>[18]</sup> 数据集对其进行训练与验证评估, 其训练超参数设置见表 1, 但学习率为 0.01 是初始值, 当训练数据集迭代次数为 30 和 60 时, 会相继缩减为 0.001 和 0.0001. 实验细节: 首先,

随机剪裁训练集图像为 $224 \times 224$ 大小的图像块. 然后, 采用同文献[18]的预处理方式处理图像块. 当网络遍历一次训练集图像后, 以验证集图像初始目标中心为中心裁出一个 $224 \times 224$ 大小的图像块用于网络评估, 评估结果见表3. LTRNet32网络的参数初始化同第5.3节实验设置. 之后, 将训练好的LTRNet32的不同Layer的卷积输出进行组合用于本文算法LRECT, 在OTB2015<sup>[6]</sup>数据集上进行分组测试, 得到如表3所示的测试结果(LRECT的参数设置见第5.2节).

表3 不同特征组合在OTB2015上的测试结果

Table 3 Test results of different feature combinations on OTB2015

| 特征组合                    | 跟踪精度/% | 跟踪成功率/% |
|-------------------------|--------|---------|
| Layer 2+Layer 4         | 88.6   | 67.2    |
| Layer 3+Layer 4         | 89.8   | 69.1    |
| Layer 2+Layer 3+Layer 4 | 92.1   | 69.8    |

由表4数据可知, 使用LTRNet32对一张图像进行分类时, 由Softmax函数统计出1000个概率, 其中, 最大概率为正确识别答案的概率(Top-1)为82.0%, 比VGG16<sup>[3]</sup>和ResNet50<sup>[8]</sup>分类网络的Top-1值分别高出6.7%和2.7%; 排名前5的概率中包含正确识别答案的概率(Top-5)为94.5%, 与ResNet50的Top-5仅差0.3%. 综上可知, LTRNet32的综合性能优于VGG16和ResNet50.

表4 ImageNet验证集上的正确率

Table 4 Accuracy rates of validation set on ImageNet

| 卷积网络     | Top-1/% | Top-5/% |
|----------|---------|---------|
| LTRNet32 | 82      | 94.5    |
| VGG16    | 75.3    | 92.5    |
| ResNet50 | 79.3    | 94.8    |

表3中Layer 2, Layer 3, Layer 4表示其卷积输出的特征图, 跟踪精度<sup>[6]</sup>指算法估计的目标中心点与人工标注的目标中心点两者的距离小于给定阈值的视频帧的百分比; 跟踪成功率<sup>[6]</sup>指算法估计的目标边界框与人工标注的目标边界框的交并比大于给定阈值的视频帧的百分比. 因Layer 1相对于目标包含较多背景信息, 对表征目标和实现精确定位干扰性强, 所以, 此处未考虑Layer1层的卷积输出. 另外, Layer 4作为LTRNet32的最后一个卷积块的输出, 包含最为丰富的目标高级语义信息, 在第3.1.2节的可视化图显示能够明显区分目标和背景, 所以, 将其列为特征组合中的一员, 但不参与以下具体的数据对比分析.

对比Layer 3+Layer 4, Layer 2+Layer 3+Layer 4两组数据可知, 后者较前者在跟踪精度和成功率方面, 分别提升了2.3%和0.7%, 这是因为Layer 2卷积特征包含了较Layer 3和Layer 4更为全面的目标边缘和位置信息, 对目标定位贡献较大, 所以提升跟踪精度比较明显.

对比Layer 2+Layer 4, Layer 2+Layer 3+Layer 4可知, 后者较前者精度和成功率分别提高了3.5%和2.6%, 这是因为Layer 3属于LTRNet32的中间层, 含有非常细致的目标轮廓信息, 和比较丰富的目标语义信息, 使其具有了利于目标定位和表征目标的双重优势.

### 5.5 跟踪算法性能对比分析

为了验证本文提出的LRECT跟踪算法的有效性, 引入近几年一些主流跟踪算法, 同本文算法一起在OTB2015<sup>[6]</sup>数据集上进行对比测试, 并采用一次跟踪评估(one-pass evaluation, OPE)<sup>[6]</sup>中的跟踪精度和跟踪成功率作为算法测试的评价指标. 测试结果见图7.

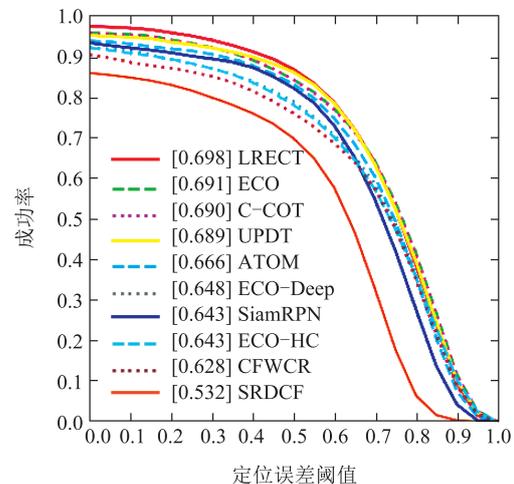
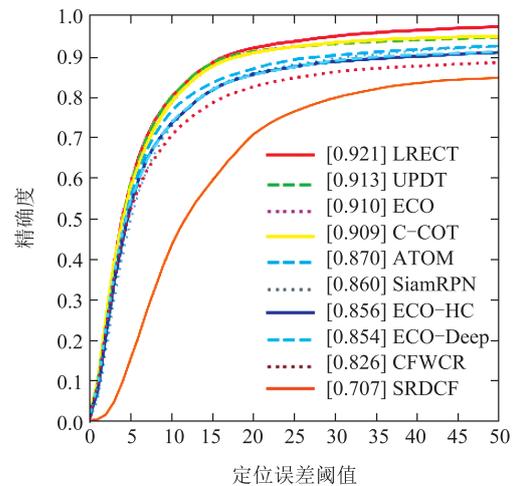


图7 算法在OTB2015上的对比测试结果

Fig. 7 Comparative test results of the algorithms on OTB2015

由图7可知, 本文算法LRECT获得了0.921的跟踪精度和0.698的成功率, 在10种算法中跟踪效果最好. 将其他算法按照目标使用特征类型进行分类, 与本文算法在精度方面相比较, 得到如表5所示的LRECT 百分点提升结果, 由表中数据可知, 一般情况下卷积特征的表征能力优于传统手工特征, 越深的卷积网络提取的卷积特征表征能力越强; 本文算法使用LTRNet32提取的目标特征比使用ResNet50<sup>[8]</sup>的UPDT算法跟踪精度高了0.8%, 这有效证明了LTRNet32网络的有效性. 在成功率方面, 本文算法较其他算法平均高出了7.1%, 这可能得益于稀疏更新投影矩阵的策略, 保留了适应目标状态的有效特征信息, 提高了跟踪成功率.

图8是ECO<sup>[5]</sup>, CFWCR<sup>[7]</sup>, UPDT<sup>[9]</sup>算法与本文

算法在OTB2015<sup>[6]</sup>数据集的11种属性上进行测试得到的精确度得分. 横轴的11种属性的相关缩写英文含义见文献 [6]. 由图 8 可知, 本文算法LRECT在SV, LR, BC, DEF, OPR这5个属性上均优于其他算法, 尤其在LR, BC这两种属性上相比基础算法ECO分别提升了10%和13%左右, 在DEF属性上比UPDT提升了3%左右. 这有效证明了LTRNet32网络的强泛化能力, 以及其提取卷积特征的强表征能力. 同时, 易知在OV这一属性上本文算法较ECO降低了7%左右, 这是因为属于OV属性的图像无法为LTRNet32提供有效的目标信息, 在跟踪过程中, 投影矩阵的更新会对后续帧的表征能力造成一定程度的负面影响, 严重时(当小目标超出视野时)可能导致跟踪失败.

表 5 LRECT较其他算法提升的精度百分点

Table 5 The improved accuracy percentages of LRECT compared with other algorithms

| LRECT跟踪精度百分点提升值 |       |              |       |          |       |              |      |         |
|-----------------|-------|--------------|-------|----------|-------|--------------|------|---------|
| 使用传统手工特征        |       | 使用VGGNet卷积特征 |       |          |       | 使用ResNet卷积特征 |      |         |
| ECO-HC          | SRDCF | ECO          | C-COT | ECO-Deep | CFWCR | UPDT         | ATOM | SiamRPN |
| [5]             | [19]  | [5]          | [4]   | [5]      | [7]   | [9]          | [20] | [21]    |
| 6.5             | 21.4  | 1.1          | 1.2   | 6.7      | 9.5   | 0.8          | 5.1  | 6.1     |

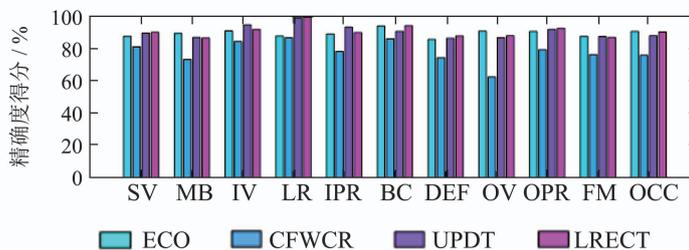


图 8 算法在OTB2015数据集的11种属性上的精确度得分

Fig. 8 Accuracy scores of algorithms on 11 attributes of the OTB2015

为了保证实验结果的客观性, 测试对比本文算法与ECO<sup>[5]</sup>算法在OTB2015<sup>[6]</sup>数据集中的100个视频序列上的表现. 跟踪精度测试结果: 本文算法跟踪Basketball, Jump, Skiing, Trans视频序列的精度结果较基础算法ECO提升了10%左右, 在跟踪Soccer, Tiger 1, Tiger 2时, 却减少了20%左右, 表明本文跟踪算法所提取的目标特征并不有利于所有的视频序列, 当遇到像Tiger 1这样涉及快速移动和遮挡等跟踪难点的序列时, 算法所提取的特征缺乏高级语义信息, 继续进行压缩操作导致其表征能力大大减弱, 最终造成不同程度的精度损失. 跟踪成功率测试结果: 本文算法跟踪David时, 由于高强度、高频率的光照变化, 使其成功率较ECO低了25%左右, 但成功率值还在60%以上. 并且在100个视频序列

中有84个的跟踪成功率在60%以上, 整体比ECO提高了7%. 以上结果表明, 本文算法不能有效应对一个视频序列中所包含的所有跟踪难点, 但与基础算法ECO相比, 训练的滤波器跟踪性能更加稳定, 验证了更新投影矩阵的有效性. 此外, 测得平均帧率如表6所示.

表 6 100个视频序列的平均帧率

Table 6 Average frame rate of 100 video sequences

| 跟踪算法  | 平均帧率(FPS) |
|-------|-----------|
| ECO   | 10.7      |
| LRECT | 14.3      |

由表6可知, 本文算法平均每秒跟踪帧数比ECO多4帧, 验证了使用特征压缩和高斯牛顿算法、共轭

梯度算法联合求解损失函数策略的有效性.

## 6 结论

本文提出的LRECT跟踪算法利用LTRNet深度卷积网络提升了算法跟踪性能. 首先, 利用LT残差结构设计高效率、高精度的LTRNet32作为特征提取模块, 通过观察、分析实验结果选取LTRNet32的Layer 2, Layer 3, Layer 4卷积块的输出表征目标. 其次, 通过分析卷积特征特点, 使用投影矩阵压缩目标高维特征, 在不影响特征表征能力的同时, 降低了计算量. 最后, 先使用高斯牛顿算法构造损失函数的二阶高斯牛顿近似表达式, 再使用共轭梯度算法对该表达式迭代优化, 同时实现了滤波器和投影矩阵的更新, 改善了在线学习滤波器收敛速度慢的问题, 使投影矩阵保持了较高效、准确的压缩特性. 最后, 实验结果表明本文跟踪算法具有较好的抗非刚性形变、低分辨率、背景混乱等能力, 以及拥有较高的精度和抗干扰能力. 此外, 帧率虽有所提升, 但仍未达到实时性要求. 对此, 会考虑采用具有高效特性的端到端网络实现目标边界框的回归等策略实现算法加速.

## 参考文献:

- [1] LIU Qiaoyuan, WANG Yuru, ZHANG Jinling, et al. Research progress of visual tracking methods based on correlation filter. *Acta Automatica Sinica*, 2019, 45(2): 265 – 275.  
(刘巧元, 王玉茹, 张金玲, 等. 基于相关滤波器的视频跟踪方法研究进展. *自动化学报*, 2019, 45(2): 265 – 275.)
- [2] LIN Jingdong, WU Xinyi, CHAI Yi, et al. Structure optimization of convolutional neural networks: A survey. *Acta Automatica Sinica*, 2020, 46(1): 24 – 37.  
(林景栋, 吴欣怡, 柴毅, 等. 卷积神经网络结构优化综述. *自动化学报*, 2020, 46(1): 24 – 37.)
- [3] KAREN S, ANDREW Z. Very deep convolutional networks for large-scale image recognition. *ArXiv Preprint ArXiv 1409.1556* (2014).
- [4] DANELLJAN M, ROBINSON A, KHAN F S, et al. Beyond correlation filters: Learning continuous convolution operators for visual tracking. *Proceedings of the 14th Computer Vision*. Amsterdam, Netherlands: Springer, 2016: 472 – 488.
- [5] DANELLJAN M, BHAT G, KHAN F S, et al. ECO: Efficient convolution operators for tracking. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. New York, USA: IEEE, 2017: 6638 – 6646.
- [6] WU Y, LIM J, YANG M H. Object tracking benchmark. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(9): 1834 – 1848.
- [7] HE Z, FAN Y, ZHUANG J, et al. Correlation filters with weighted convolution responses. *2017 IEEE International Conference on Computer Vision Workshop*. New York, USA: IEEE, 2017: 1992 – 2000.
- [8] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition. *Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*. New York, USA: IEEE, 2016: 770 – 778.
- [9] BHAT G, JOHNANDER J, DANELLJAN M, et al. Unveiling the power of deep tracking. *Proceedings of the European Conference on Computer Vision*. Munich, Germany: Springer, 2018: 483 – 498.
- [10] KRISTAN M, LEONARDIS A, MATAS J, et al. The sixth visual object tracking vot2018 challenge results. *Proceedings of the European Conference on Computer Vision*. Munich, Germany: Springer, 2018: 3 – 53.
- [11] DONG B, LU Y P, ZHONG A X, et al. Beyond finite layer neural networks: Bridging deep architectures and numerical differential equations. *ArXiv Preprint ArXiv 1710.10121*(2017).
- [12] E W N. A proposal on machine learning via dynamical systems. *Communications in Mathematics and Statistics*, 2017, 5(1): 1 – 11.
- [13] NOCEDAL J, WRIGHT S J. *Numerical Optimization*. 2nd edition. New York, USA: Springer, 2006.
- [14] ZHOU Chenghao. *Research on design and mid-frequency information improving for optical sparse aperture system*. Harbin: Harbin Institute of Technology, 2019.  
(周程灏. 光学稀疏孔径系统设计与中频信息改善方法研究. 哈尔滨: 哈尔滨工业大学, 2019.)
- [15] DECLERCQ A, PIATER J H. Online learning of gaussian mixture models-a two-level approach. *Proceedings of the Third International Conference on Computer Vision Theory and Applications*. Funchal, Madeira, Portugal: INSTICC, 2008: 605 – 611.
- [16] KRIZHEVSKY A, HINTON G. *Learning multiple layers of features from tiny images*. Toronto: University of Toronto, 2009: 1 – 60.
- [17] LEE C Y, XIE S N, ZHANG Z Y, et al. Deeply supervised nets. *Proceedings of the Eighteenth International Conference on Artificial Intelligence and Statistics*. San Diego, California, USA: [s.n.], 2015: 562 – 570.
- [18] KRIZHEVSKY A, SUTSKEVER I, HINTON G. ImageNet classification with deep convolutional neural networks. *Proceedings of the 25th International Conference on Neural Information Processing Systems*. Lake Tahoe, Nevada, USA: Curran Associates Inc, 2012: 1097 – 1105.
- [19] DANELLJAN M, HAGER G, KHAN F S, et al. Learning spatially regularized correlation filters for visual tracking. *IEEE International Conference on Computer Vision*. Santiago, Chile: IEEE, 2015: 4310 – 4318.
- [20] DANELLJAN M, BHAT G, KHAN F S, et al. ATOM: Accurate tracking by overlap maximization. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. New York, USA: IEEE, 2019: 4660 – 4669.
- [21] LI B, YAN J J, WU W, et al. High performance visual tracking with siamese region proposal network. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. New York, USA: IEEE, 2018: 8971 – 8980.

## 作者简介:

陈志旺 副教授, 博士, 目前研究方向为运动物体目标检测与跟踪, E-mail: czwaaron@ysu.edu.cn;

王莹 硕士研究生, 目前研究方向为单目标跟踪, E-mail: 1718819591@qq.com;

宋娟 工程师, 目前研究方向为无人机电力系统巡线, E-mail: 1138812341@qq.com;

姚权允 硕士研究生, 目前研究飞行器组合导航, E-mail: 1103706199@qq.com;

彭勇 教授, 博士生导师, 目前研究方向为脑电应用科学, E-mail: PY81@sina.com.