

## 基于深度特征与局部约束掩膜的相关滤波跟踪算法

王译萱 吴小俊

(江南大学物联网工程学院, 无锡, 214122)

**摘要:** 为提升相关滤波算法在目标遮挡、快速运动以及背景杂乱等情况下跟踪结果的精确度和鲁棒性, 提出了一种基于深度特征与局部约束掩膜(Local constrained mask, LCM)的相关滤波跟踪算法。在鉴别性相关滤波跟踪算法的基础上, 利用学习得到的二值矩阵作为LCM对滤波器的能量分布进行裁剪, 对模板边缘与测试图像之间产生的响应值进行抑制, 实现扩大目标搜索区域的同时降低边界效应对跟踪结果的影响; 将深度特征引入到特征提取过程中, 通过对目标样本进行旋转、翻折和高斯模糊等处理, 扩充训练样本数量, 使模板学习到更为丰富的目标信息。与主流算法进行对比实验, 验证了本文算法在处理目标遮挡、背景嘈杂以及光照变化等干扰时的鲁棒性。

**关键词:** 相关滤波; 特征提取; 掩膜; 背景干扰

**中图分类号:** TP391.4      **文献标志码:** A

### Correlation Filter Tracking Algorithm Based on Deep Features with Local Constrained Mask

Wang Yixuan, Wu Xiaojun

(School of Internet of Things Engineering, Jiangnan University, Wuxi, 214122, China)

**Abstract:** To improve the tracking accuracy and robustness of correlation filter based trackers, especially with challenging factors such as occlusion, fast motion, background clutter and so forth, an improved correlation filter tracking algorithm based on deep features with local constrained mask is proposed. Based on discriminative correlation filter tracking algorithms, the learned binary matrix is proposed as a local constrained mask to achieve pruning the filter energy, which suppresses the response map generated by the template edge and the testing images. This allows the proposed method to expand the target search region and effectively alleviates the boundary effects caused by circular shifted training samples. Deep features are introduced in the process of feature extraction. By exploiting rotation, flipping, and Gaussian blur operations, the training sample set is expanded, which makes feature templates learn more target information. Compared the robustness of our algorithm with the mainstream methods under distractions like occlusion, background clutter and illumination changes.

**Key words:** correlation filter; feature extraction; mask; background distractor

## 引言

目标跟踪作为计算机视觉中的重要研究方向,在安防监控、无人驾驶和智慧城市等方向有着广泛的应用<sup>[1-3]</sup>。跟踪的主要目的是在连续的视频序列中建立和更新目标的动态表观模型,以此为依据获得运动目标的运动参数和轨迹等,从而满足军用以及民用需求<sup>[4]</sup>。

历经30年的研究与发展,跟踪算法的性能有了广泛的提高,目前算法主要分为生成学习和鉴别学习两大类<sup>[5]</sup>。生成学习跟踪算法对当前帧目标区域建模,在下一帧采取寻找与模型最为相似区域,以此确定目标位置<sup>[6]</sup>。早期的相关工作主要有粒子滤波算法<sup>[7]</sup>、均值漂移跟踪算法<sup>[8]</sup>和稀疏表示等跟踪算法<sup>[9]</sup>等。生成学习类算法的主要局限在于过分关注目标本身的外观信息,忽略了目标外观具有动态性和随机性等特点,因此不能准确把握目标外观动态信息,无法实现鲁棒跟踪。为了在跟踪过程中引入背景信息,研究人员设计出了鉴别学习类跟踪算法,此类算法将跟踪问题转为分类或回归问题,将目标信息作为正样本,背景信息为负样本训练分类器,将目标从图像背景中分离出来,以此实现跟踪。近年来,基于鉴别学习类的相关滤波跟踪算法成为了研究热点并受到了广泛关注,该方法的设计初衷是学习得到一个对目标响应高而对背景响应低的滤波器。通过滤波器模板和搜索区域的响应值,判断目标位置。Bolme等<sup>[10]</sup>首先提出了输出误差最小平方和跟踪算法(Minimum output sum of squared error, MOSSE),通过频率域特性缓解了训练过程中样本数量不足的问题。随后,Henriques等<sup>[11]</sup>在MOSSE算法的基础上利用循环结构扩充了训练样本集,提出了循环结构跟踪算法(Circulant structure kernel, CSK)。在此基础上,Henriques等<sup>[12]</sup>进一步将原有的单通道灰度特征(Gray scale)扩展为多通道的方向梯度直方图特征(Histogram of oriented gradient, HoG),同时将核计巧(Kernel trick)加入到滤波器求解过程中,提出核化相关滤波跟踪算法(Kernelized correlation filter, KCF)。为了解决跟踪过程中目标的尺度变化问题,Danelljan等<sup>[13]</sup>提出了多尺度相关滤波跟踪算法(Discriminative scale space tracking, DSST),该方法针对目标存在的尺度变化问题,将单独的尺度滤波器加入到算法中,通过尺度采样和尺度回归的方式计算目标的尺度变化。为了将目标的颜色信息引入跟踪过程中,Danelljan和Bertinetto等分别提出了使用颜色名称特征的自适应颜色属性实时跟踪算法(Adaptive color accurate tracking, ACAT)<sup>[14]</sup>以及基于颜色直方图独立模板的相关滤波跟踪算法(Complementary learners for real-time tracking, Staple)<sup>[15]</sup>。颜色特征(Color name, CN)在目标发生形变、遮挡时具有良好的鲁棒性,因此后续的相关滤波算法在特征提取阶段多数都采用了HoG特征结合CN的方式增加跟踪结果的可靠性。近年来大量优秀算法将深度特征引入相关滤波跟踪框架中<sup>[16-18]</sup>,与传统特征相比,深度特征包含更加丰富的语义信息,对目标产生的旋转、平移和缩放等具有较好的鲁棒性<sup>[19-20]</sup>。尽管深度特征的提取过程较为复杂,但随着计算机硬件的不断发展,目前已经可以通过图形处理器(Graphics processing unit, GPU)实现深度特征的运算,因此对基于深度特征的目标跟踪算法还有着重要的研究意义和价值。

本文在经典相关滤波算法基础上,提出了基于深度特征与局部约束掩膜(Local constrained mask, LCM)的相关滤波跟踪算法。在滤波器的学习过程中,利用二值矩阵作为局部约束掩膜对滤波器边缘能量分布进行裁剪,对模板边缘与测试图像之间产生的响应进行抑制,提升了滤波器的鉴别能力;将通过ResNet50深度网络提取到的深度特征加入跟踪算法中,利用深度特征学习目标表观模型时,将目标图像通过旋转、翻折和高斯模糊3种方式扩充样本数量,提升目标模板的表达能力,最终采用深度特征结合传统特征的组合形式,提高了复杂背景下跟踪结果的准确性和鲁棒性。

## 1 多通道相关滤波算法

相关滤波跟踪算法通过在目标周围建立采样区域,采集训练样本实现滤波器的训练。通过滤波器

和测试图像之间产生的响应值判断目标的当前位置,响应值最高的位置通常为候选目标位置<sup>[5,11]</sup>。在采样的过程中,相关滤波算法对目标样本进行循环移位操作快速生成大量训练样本,进而通过离散傅里叶变换(Discrete Fourier transform, DFT),将滤波器的求解过程转换到频率域进行,其求解过程可视为优化函数 $\epsilon$ 和求解滤波器 $h$ ,即

$$\epsilon = \left( \sum_{d=1}^{N_d} \|f_d * h_d - g\|^2 + \lambda \sum_{d=1}^{N_d} \|h_d\|^2 \right) \quad (1)$$

式中: $h_d$ 为待求解的滤波器,\*表示循环卷积操作, $f_d$ 为提取到的目标特征, $N_d$ 为特征通道总和, $g$ 为对应的高斯响应,其峰值位于中心, $\lambda$ 为正则项系数,防止模型过拟合。对式(1)进行离散傅里叶变换,可将循环卷积操作变为频域中的点乘,减少卷积耗时,即

$$\epsilon = \left( \sum_{d=1}^{N_d} \|\hat{h}_d^H \text{diag}(\hat{f}_d) - \hat{g}\|^2 + \lambda \sum_{d=1}^{N_d} \|\hat{h}_d\|^2 \right) \quad (2)$$

式中: $\hat{f}_d$ 和 $\hat{h}_d$ 分别为 $f_d$ 和 $h_d$ 的傅里叶变换, $\hat{h}_d^H$ 代表 $\hat{h}_d$ 的复数共轭形式,对式(2)的求解为标准的岭回归问题,因此相关滤波器 $h_d$ 在频域内有闭环解(Close-form solution),即

$$\hat{h}_d = \frac{\hat{g}^H \odot \hat{f}_d}{\sum_{d=1}^{N_d} \hat{f}_d^H \odot \hat{f}_d + \lambda} \quad (3)$$

式中: $\odot$ 表示频率域中的点乘操作,式(3)的详细推导过程请参阅文献[5]。

考虑到目标在跟踪过程中的外观变化,学习得到的滤波器需要在线更新,更新过程通常采用滑动平均的方式,因此在第 $t$ 帧视频序列中相关滤波器 $\hat{h}$ 的更新公式为

$$\hat{h} = (1 - \eta) \hat{h}^{t-1} + \eta \hat{h}^t \quad (4)$$

式中: $\eta$ 为模板更新速率, $t$ 为当前帧数。

在新一帧的目标检测过程中,根据上一帧目标的位置,提取测试图像 $z$ 和学习得到的滤波器 $\hat{h}$ 进行相关操作,得到当前帧的目标响应,响应值最高的位置为目标位置,即

$$R(z) = F^{-1} \{ \hat{h} \odot \hat{z} \} \quad (5)$$

式中: $R(z)$ 为当前帧的响应值, $F^{-1}$ 表示傅里叶逆变换。

## 2 多通道相关滤波算法的改进

在多通道相关滤波算法的基础上,本文提出基于深度特征与LCM的相关滤波跟踪算法,利用学习得到的二值矩阵作为LCM对滤波器进行处理,消除滤波器模板的边缘信息,实现扩大目标搜索区域的同时,降低循环样本带来的边界效应,获得更高的跟踪精度。

### 2.1 LCM的相关滤波跟踪算法

#### 2.1.1 局部掩膜

鉴别学习类跟踪算法的主要难点在于训练分类器时的训练样本过少,传统相关滤波算法为了增加训练样本的数量,对目标周围建立一个搜索区域,以区域内的图像作为原始样本,在空间域内采用循环移位的方式生成大量虚拟样本,搜索区域越大,获得的训练样本数越多<sup>[20]</sup>。尽管该方式为算法提供了更多样本,但这些虚拟样本因受边界效应的影响,弱化了分类器的鉴别能力,对算法的跟踪精度造成影响,因此传统方法只能将目标搜索区域控制在相对较小的范围内。为了有效解决上述问题,近年来的改进算法,如空间正则化相关滤波器(Spatially regularized correlation filters, SRDCF)<sup>[21]</sup>和鉴别性空间通道可靠性相关滤波跟踪算法(Channel spatial reliability discriminative correlation filter, CSRDCF)<sup>[22]</sup>通过抑制滤波器模板的边缘信息,实现了目标搜索区域的扩大,有效提升了跟踪结果的准确性和鲁棒性。

在此基础上,本文提出了一种新型的掩膜操作,通过二值矩阵作为LCM,在频域与滤波器进行点乘操作,即

$$\mathbf{h}_m = \mathbf{m} \odot \mathbf{h} \quad (6)$$

式中, $\mathbf{m}$ 为二值矩阵,其大小与与目标搜索区域大小相同,训练时可由测试视频序列中第1帧给定的目标位置和大小求得, $\mathbf{h}_m$ 则表示与 $\mathbf{m}$ 点乘后的滤波器,点乘的过程中对目标区域取1,保留目标区域模板信息,模板边缘区域取0,消除滤波器模板边缘的能量,达到抑制边界效应的目的,在后续的跟踪过程中,为保证 $\mathbf{m}$ 的准确性,其大小始终与第一帧(Ground-truth)得到的结果保持不变。此外,由于本文提出的掩膜操作作用于滤波器模板,而并非直接作用于原始图像,因此在训练阶段不会因目标运动产生干扰,造成跟踪结果下降。与使用颜色直方图前景掩膜的CSRDCF算法相比<sup>[21]</sup>,本文提出的二值矩阵掩膜无需获取输入图像的颜色直方图特征,因此减少了算法耗时。此外,本文根据目标区域大小生成的掩膜具有更大的空间范围,在跟踪的过程中可以有效减少因使用感受野较大的深度特征而产生的信息丢失问题,提升算法的跟踪精度。本文提出的二值矩阵掩膜示意图如图1所示。



图1 二值矩阵掩膜示意图

Fig.1 Schematic diagram of binary matrix mask

### 2.1.2 优化过程

将式(6)作为约束条件加入损失函数式(1)中,即

$$\varepsilon = \left( \sum_{d=1}^{N_d} \|\mathbf{f}_d * \mathbf{h}_d - \mathbf{g}\|^2 + \lambda \sum_{d=1}^{N_d} \|\mathbf{m} \odot \mathbf{h}_d\|^2 \right) \quad (7)$$

采用增广拉格朗日乘法(Augmented Lagrange method, ALM)<sup>[23]</sup>对式(7)进行优化求解,优化过程中需引入松弛变量 $\mathbf{h}'_d = \mathbf{h}_d$ ,即

$$\ell = \left( \sum_{d=1}^{N_d} \|\mathbf{f}_d \cdot \mathbf{h}_d - \mathbf{g}\|^2 + \lambda \sum_{d=1}^{N_d} \|\mathbf{m} \odot \mathbf{h}_d\|^2 + \frac{\mu}{2} \sum_{d=1}^{N_d} \left\| \mathbf{h}_d - \mathbf{h}'_d + \frac{\mathbf{y}_d}{\mu} \right\|^2 \right) \quad (8)$$

式中: $\mathbf{y}_d$ 为拉格朗日乘子, $\mu$ 为惩罚系数。式(7,8)的详细推导过程请参阅文献[22]。对式(8)采用交替方向乘法(Alternating direction method of multipliers, ADMM)<sup>[24]</sup>,将式(8)分为2个子问题,交替求解,即

$$\mathbf{h}_d = \min \left( \sum_{d=1}^{N_d} \|\mathbf{f}_d * \mathbf{h}_d - \mathbf{g}\|^2 + \frac{\mu}{2} \sum_{d=1}^{N_d} \left\| \mathbf{h}_d - \mathbf{h}'_d + \frac{\mathbf{y}_d}{\mu} \right\|^2 \right) \quad (9)$$

$$\mathbf{h}'_d = \min \left( \lambda \sum_{d=1}^{N_d} \|\mathbf{m} \odot \mathbf{h}_d\|^2 + \frac{\mu}{2} \sum_{d=1}^{N_d} \left\| \mathbf{h}_d - \mathbf{h}'_d + \frac{\mathbf{y}_d}{\mu} \right\|^2 \right) \quad (10)$$

式(9)需要利用帕萨瓦尔定理(Parseval's theorem)<sup>[18]</sup>转换到频域内求解,即

$$\hat{\mathbf{h}}_d = \min \left( \sum_{d=1}^{N_d} \|\hat{\mathbf{f}}_d * \hat{\mathbf{h}}_d - \hat{\mathbf{g}}\|^2 + \frac{\mu}{2} \sum_{d=1}^{N_d} \left\| \hat{\mathbf{h}}_d - \hat{\mathbf{h}}'_d + \frac{\hat{\mathbf{y}}_d}{\mu} \right\|^2 \right) \quad (11)$$

对式(10)中 $\mathbf{h}'_d$ 和式(11)中的 $\hat{\mathbf{h}}_d$ 分别求偏导数,得到2个变量的闭环解,即

$$\mathbf{h}'_d = (1 - \mathbf{m}) \odot \frac{\mu \mathbf{h}_d + \mathbf{y}_d}{2\lambda + \mu} \quad (12)$$

$$\hat{h}_d = \frac{\hat{f}_d \odot \hat{g}^H + \mu \hat{h}'_d - \hat{y}_d}{\mu + \hat{f}_d \odot \hat{f}_d^H} \quad (13)$$

利用式(12,13)进行迭代求解,随着迭代次数的增加,松弛变量  $h'_d$  的值会无限逼近于  $h_d$ ,因此本文将选取  $\hat{h}_d$  作为滤波器最终训练模板。

$y_d$  和  $\mu$  在迭代的过程中需要更新,更新公式为

$$y = y + \mu(h - h') \quad (14)$$

$$\mu = \min(\beta\mu, \mu_{\max}) \quad (15)$$

## 2.2 深度特征的引入以及训练样本扩充

目前大部分针对相关滤波跟踪的改进算法在特征提取阶段均采用 HoG 结合 CN 的形式,这种传统特征组合在跟踪速度上具有一定优势,但在目标发生遮挡、尺度变化、背景杂乱和形变以及复杂背景下表现较差<sup>[14-20]</sup>。为了进一步提升算法性能,本文在 HoG 和 CN 的基础上<sup>[21]</sup>,将通过深度网络提取到的 ResNet50 深度特征引入跟踪算法中,深度特征是网络通过大数据自动学习得到的特征,其中包含大量参数。近些年来,深度特征被广泛应用在图片检索和分类等任务中<sup>[19]</sup>,与传统特征相比,深度特征具有较大的感受野,鲁棒性较好<sup>[25]</sup>。为了获取更为鲁棒的深度特征,本文选择 ResNet50 网络中第 4 个卷积块的 res4ex 层的输出作为输入特征,该输出共包含 1 024 个特征图谱。与浅层的网络输出相比,该层输出的特征图谱感受野较大,并包含较为丰富的语义信息,具有很强的鲁棒性,因此可有效提升本文算法的跟踪精度和鲁棒性。为了进一步增强深度特征模板的表达能力,本文利用深度特征具有丰富语义信息且对目标产旋转、缩放和平移具有不变性等优势,将输入目标图像进行旋转、翻折和高斯模糊 3 种方式进行处理,扩充样本数量,使深度特征模板学习到更为丰富的目标信息,样本扩充的示意图如图 2 所示。

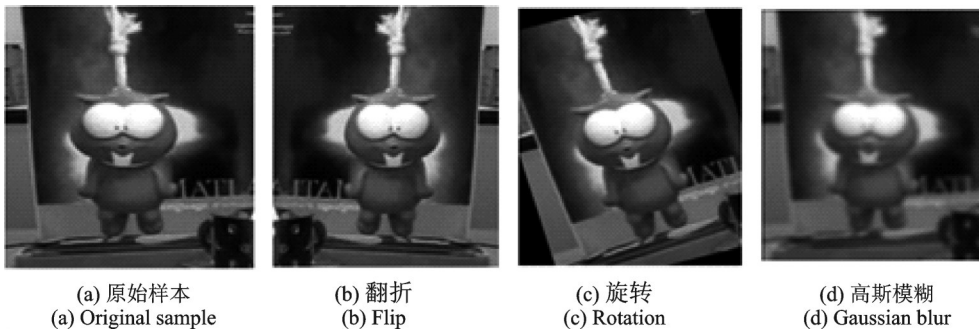


图2 训练样本扩充意图

Fig.2 The diagram of the training sample augmentation.

## 2.3 尺度估计

针对目标的尺度变化问题,本文采用文献[2,13]中的尺度估计法估算目标尺度大小,根据估测出的目标位置,利用尺度滤波器在尺度空间进行搜索,获取目标的尺度响应,进而实现准确的目标尺度估计。

## 2.4 算法流程图

本文算法的完整流程图如图 3 所示,算法首先根据初始帧给出的目标位置和大小获取用于采集训练样本的目标区域,并根据目标区域生成本文提出的二值矩阵掩膜,其次,对采集到的训练样本进行特

征提取,并通过固定迭代次数的增广拉格朗日乘法求解滤波器模板,最后将训练好的滤波器模板与测试图像进行相关操作,获取相应的目标响应,根据响应值的大小判断目标的当前位置,并通过尺度滤波器获取目标的尺度信息。

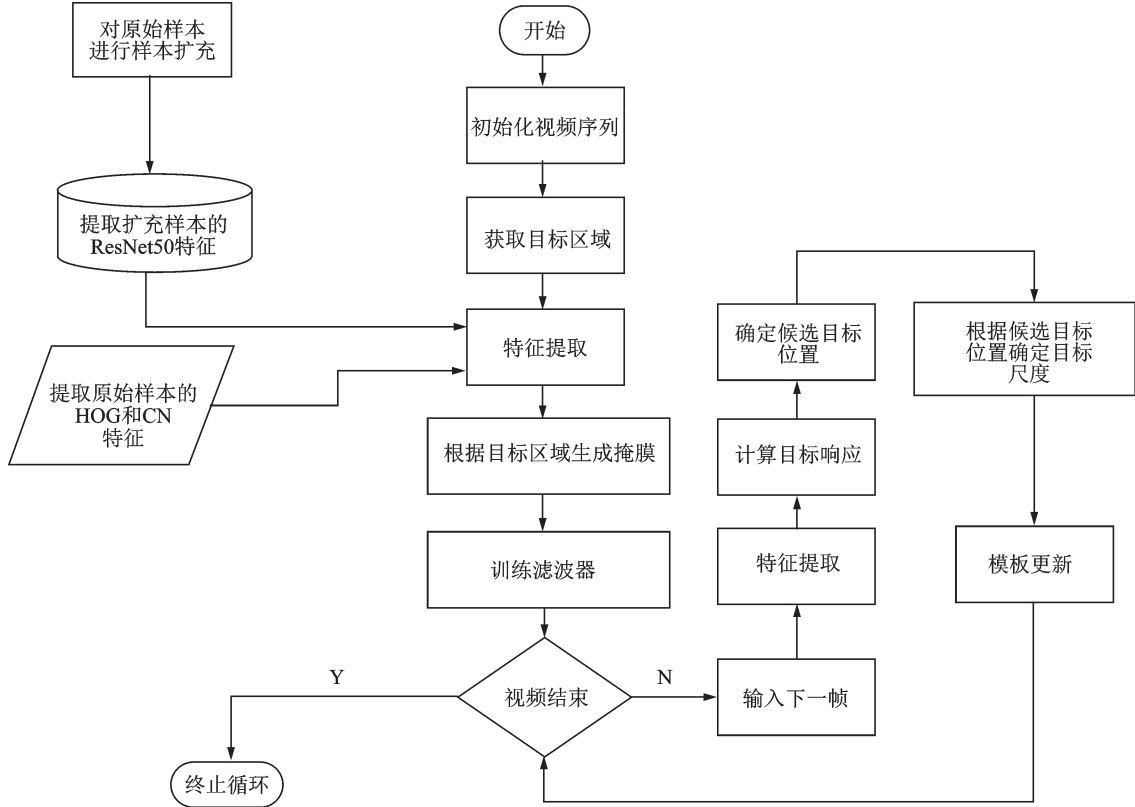


图3 算法流程图

Fig.3 Flow chart of tracking algorithm

### 3 实验结果与分析

#### 3.1 测试数据集和评价指标

实验测试采用的数据集为 OTB-100 (Object Tracking Benchmark 100, OTB100),共包含 100 个视频序列,采用一次通过评估模式(One pass evaluation, OPE),评估标准采用成功率曲线面积(Area under curve, AUC)、距离精度(Distance precision, DP)评价指标。成功率曲线面积 AUC 计算重叠率大于某个阈值的帧数占总帧数的百分比,其结果用曲线下面积 AUC 表示,DP 为中心误差小于 20 个像素的帧数占帧数总体的百分比<sup>[4]</sup>。

#### 3.2 总体性能对比实验

本节主要在 OTB 测试数据集中 100 组具有复杂属性的视频序列上,利用 OPE 模式,将所提算法与 5 种主流算法进行对比测试,得到成功率曲线和距离精度曲线的比较结果。实验结果如图 4 所示。

从图 4 中的实验对比结果可以看出,本文采用 HoG+CN 特征组合的跟踪算法的平均 AUC 值为 61.0%,比 CSRDCF 算法提升约 3.3%,比 SRDCF 算法提升约 1.2%,因此在不使用深度特征的情况下,

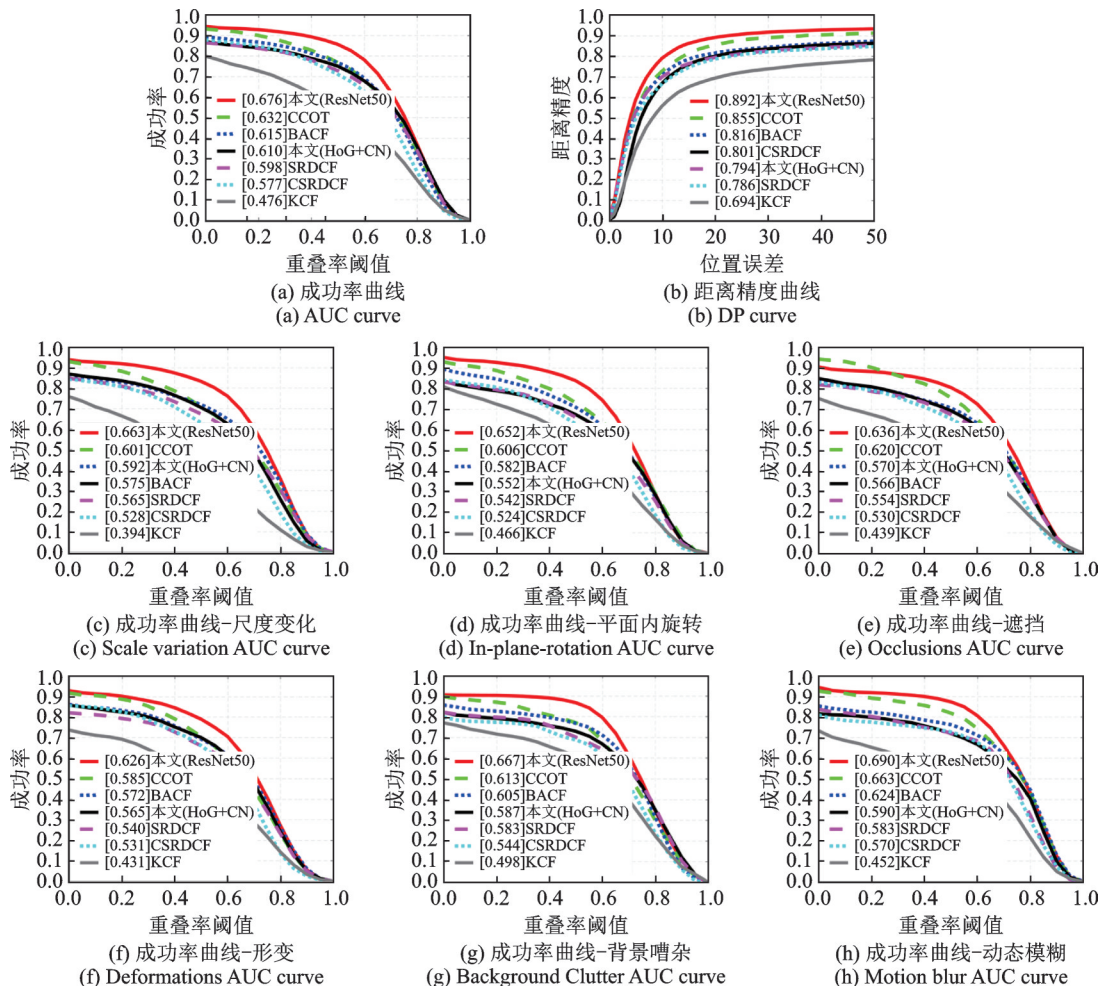


图4 7种跟踪算法的成功率曲线和距离精度曲线对比图

Fig.4 Precision plots and success rate plots of seven trackers on OTB-100 benchmark

本文算法仍然具有一定改进效果;本文在 HoG 和 CN 基础上加入 ResNet50 深度特征的算法 Ours (ResNet50)在距离精度和成功率上明显优于其他算法,平均 AUC 值为 67.6%,平均 DP 值为 89.2%,比次佳效果算法 CCOT 提升大约 5%,改进效果较为明显。主要原因为:(1)本文算法在二值矩阵掩膜约束的将目标搜索区域扩大为目标大小的 4 倍,因此训练阶段获取了数量更多的训练样本,同时由于 LCM 对滤波器的约束,缓解了搜索窗口扩大后的边界效应问题,不会因搜索窗口扩大导致训练样本质量降低,保证了训练样本的质量,提升了分类器的鉴别能力;(2)ResNet50 深度特征在复杂背景下具有很强的鲁棒性,一定程度上提升了算法整体的跟踪精度。

### 3.3 样本扩充对比实验

为了提升深度特征的目标模板表达能力,将输入目标图像进行旋转、翻折以及高斯模糊等处理,扩充样本数量,使深度特征模板学习到更为丰富的目标语义信息。为了说明样本扩充实验的有效性,本文在该部分实验去除了 3.2 节中算法使用的 HoG 和 CN 特征,仅测试单纯使用 ResNet50 深度特征时,进行样本扩充和不作任何处理 2 种方式在 OTB-100 测试集中针对遮挡 (Occlusion, OCC)、背景杂乱

(Background clutter, BC)、形变(Deformation, DEF)、运动模糊(Motion blur, MB)、平面内旋转(In-plane rotation, IPR)、尺度变化(Scale variations, SV)6种挑战下的平均AUC和DP,实验结果如表1,2所示。通过表1,2中显示的AUC和DP值可以看出,使用样本扩充方式得到的AUC和DP值总体上普遍高于不作处理的方式,该实验结果验证了对深度特征采取样本扩充可以一定程度上提升跟踪结果的精度和鲁棒性。

表1 样本扩充实验6种属性成功率结果

Tab. 1 Six attributes AUC results of training set augmentation

样本类型	OCC	BC	DEF	MB	IPR	SV
样本扩充	0.584	0.667	0.626	0.690	0.652	0.663
不作处理	0.576	0.632	0.564	0.647	0.632	0.641

表2 样本扩充实6种属性距离精度结果

Tab. 2 Six attributes DP results of training set augmentation

样本类型	OCC	BC	DEF	MB	IPR	SV
样本扩充	0.851	0.887	0.850	0.881	0.892	0.873
不作处理	0.824	0.867	0.804	0.872	0.893	0.886

### 3.4 复杂背景条件下实验

为了充分验证本文算法效果,本文选取了包括复杂背景、目标遮挡、目标旋转、光照变化等干扰视频序列,将本文采用HoG、CN和ResNet50深度特征的跟踪算法与5种相关滤波跟踪算法进行对比测试,测试视频序列分别为MotorRolling、BikerSkiing、Skating1、Matrix以及Lemming,其相关属性如表3所示,本文算法与5种算法在上述视频序列中的跟踪效果图如图5所示。

表3 测试视频属性

Tab. 3 Video sequence attributes

测试视频	帧数 / 帧	目标大小 / 像素 × 像素	主要干扰
MotorRolling	164	122 × 125	IPR、BC、SV、MB
Biker	81	29 × 26	IPR、SV
Skating1	400	34 × 84	OCC、DEF、BC
Matrix	100	38 × 42	BC、DEF、SV
Lemming	1 336	103 × 61	OCC、IPR、SV

图5(a)所示的MotoRolling视频中,目标在运动的过程中发生了旋转变化,除本文算法外其余算法均产生跟踪失败,主要原因为目标旋转时传统特征的鲁棒性较差,对旋转变化较为敏感,本文算法采用了较为鲁棒的深度特征,因此在目标运动发生变化时,依然可以保持鲁棒跟踪;图5(b)所示的Skating1视频中,当目标发生遮挡时和光照变化时,只有本文算法、CCOT以及CSRDCF可以较为准确地跟踪目标,在视频将要结束时,本文算法对目标尺度估计较为精确,效果上优于其他算法;图5(c)所示的Matrix视频中,目标的运动过程较为剧烈,背景与目标相似度极高,较为杂乱,在此环境下,只有本文算法以及SRDCF算法可以跟踪到目标,同时本文算法的跟踪准确度略微优于SRDCF;图5(d)所示的Lemming视频中,目标发生了长时间遮挡,本文算法依靠较大的目标搜索和采样区域,依然可以准确地追踪到目标;图5(e)所示的Biker视频中,跟踪的主要难点在于目标发生的平面外旋转,测试算法中除本文方法外的其余算法均发生跟踪漂移或失败,在视频结束时,只有本文算法可以较为准确地跟踪目标。综上所述,本文算法在遮挡、背景杂乱、光照变化、目标形变以及旋转等干扰下,具有较强的鲁棒性。

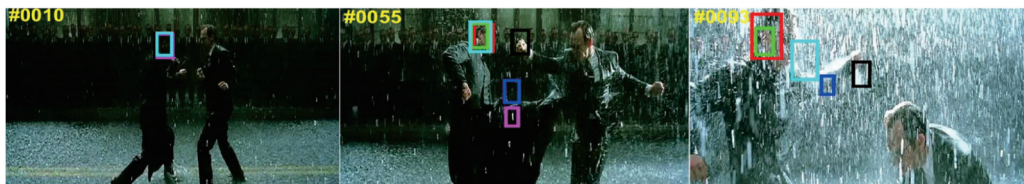




(a) MotorRolling



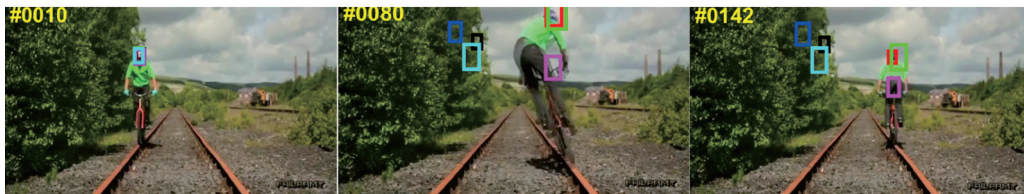
(b) Skating1



(c) Matrix



(d) Lemming



(e) Biker

— Ours    — CCOT    — SRDCF    — BACF    — CSRDCF    — KCF

图5 6种跟踪算法在5个视频序列的跟踪结果比较

Fig.5 Tracking results of six algorithms on five video sequences (Ours, CCOT, SRDCF, BACF, CSRDCF, KCF)

#### 4 结束语

本文在相关滤波跟踪算法的基础上,提出一种基于深度特征与LCM的相关滤波算法。该方法首先利用二值矩阵作为局部掩膜消除滤波器模板边缘的能量分布,降低循环样本带来的边界效应,以此为基础扩大目标的搜索区域,进而获取更多的训练样本;在经典算法中的HoG和CN特征的基础上,加入ResNet50深度特征,利用深度特征鲁棒性较强且具有丰富语义信息等特点,通过旋转、翻折和高斯模糊3种方式对目标样本进行处理,扩充了训练样本集。与其他算法的对比实验结果表明,本文算法在跟

踪精度和鲁棒性提升上具有一定的改进。算法目前存在的主要缺陷为无法满足实时性,在下一步的研究工作中可以考虑对提取得到的深度特征实现自适应降维,改善算法的实时性。

#### 参考文献:

- [1] Yilmaz A, Javed O, Shah M. Object tracking: A survey[J]. *ACM Computing Survey*, 2006, 38(4): 81-93.
- [2] 谢柳, 尚振宏, 刘辉. 基于相关滤波融合多特征的运动目标跟踪方法[J]. *数据采集与处理*, 2019, 34(1): 122-134.  
Xie Liu, Shang Zhenhong, Liu Hui. Object tracking combining multiple features based on correlation filter[J]. *Journal of Data Acquisition and Processing*, 2019, 34(1): 122-134.
- [3] 徐国军, 笄良龙, 李玉阳, 等. 基于 Gabor 滤波的引导源目标定位[J]. *数据采集与处理*, 2011, 26(1): 31-35.  
Xu Guojun, Da Lianglong, Li Yuyang, et al. Target localization algorithm using guide source based on Gabor filters[J]. *Journal of Data Acquisition and Processing*, 2011, 26(1): 31-35.
- [4] 冯斐, 吴小俊, 徐天阳. 基于子空间和直方图的多记忆自适应相关滤波目标跟踪算法[J]. *模式识别与人工智能*, 2018, 31(7): 612-624.  
Feng Fei, Wu Xiaojun, Xu Tianyang. Object tracking with multiple memory learning and adaptive correlation filter based on subspace and histogram[J]. *Pattern Recognition and Artificial Intelligence*, 2018, 31(7): 612-624.
- [5] Smeulders A W M, Chu D M, Cucchiara R, et al. Visual tracking: An experimental survey[J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2014, 36(7): 1442-1468.
- [6] Adam A, Rivlin E, Shimshoni I. Robust fragments-based tracking using the integral histogram[C]// *Computer vision and pattern recognition*, 2006 IEEE Computer Society Conference on. [S.l.]: IEEE, 2006, 1: 798-805.
- [7] Nummiaro K, Koller-Meier E, Gool L V. An adaptive color-based particle filter[J]. *Image & Vision Computing*, 2003, 21(1): 99-110.
- [8] Comaniciu D, Ramesh V, Meer P. Real-time tracking of non-rigid objects using mean shift[C]// *Proceedings IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2000)*. [S.l.]: IEEE, 2002.
- [9] Mei X, Ling H. Robust visual tracking and vehicle classification via sparse representation[J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2011, 33(11): 2259-2272.
- [10] Bolme D S, Beveridge J R, Draper B A, et al. Visual object tracking using adaptive correlation filters[C]// *Proceeding of 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.]: IEEE, 2010: 2544-2550.
- [11] Henriques J F, Caseiro R, Martins P, et al. Exploiting the circulant structure of tracking-by-detection with kernels[C]// *Computer Vision—ECCV 2012. Lecture notes in computer science*. Berlin, Heidelberg: Springer, 2012: 702-715.
- [12] Henriques J F, Caseiro R, Martins P, et al. High-speed tracking with kernelized correlation filters[J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2015, 37(3): 583-596.
- [13] Danelljan M, Häger G, Shahbaz K F, et al. Accurate scale estimation for robust visual tracking[C]// *Proceedings of the British Machine Vision Conference 2014*. Nottingham, UK: BMVA Press, 2014.
- [14] Danelljan M, Shahbaz Khan F, Felsberg M, et al. Adaptive color attributes for real-time visual tracking[C]// *Proc. of 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Columbus, Ohio, USA: CVPR, 2014: 1090-1097.
- [15] Bertinetto L, Valmadre J, Golodetz S, et al. Staple: Complementary learners for real-time tracking[C]// *International Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.]: IEEE, 2016.
- [16] Danelljan M, Hager G, Khan F S, et al. Convolutional features for correlation filter based visual tracking[C]// *2015 IEEE International Conference on Computer Vision Workshop (ICCVW)*. [S.l.]: IEEE Computer Society, 2015.
- [17] Danelljan M, Robinson A, Khan F S, et al. Beyond correlation filters: Learning continuous convolution operators for visual tracking[C]// *European Conference on Computer Vision*. [S.l.]: Springer International Publishing, 2016.
- [18] Danelljan M, Bhat G, Khan F S, et al. ECO: efficient convolution operators for tracking[C]// *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI, USA: IEEE, 2017: 6931-6939.
- [19] 余丹, 吴小俊. 一种卷积神经网络和极限学习机相结合的人脸识别方法[J]. *数据采集与处理*, 2016, 31(5): 996-1003.  
Yu Dan, Wu Xiaojun. Face recognition algorithm based on combination of convolutional neural networks and extreme learning

machine[J]. *Journal of Data Acquisition and Processing*, 2016, 31(5): 996-1003.

- [20] 王译萱, 吴小俊, 徐天阳. 基于时序一致和空间剪裁的多特征相关滤波跟踪算法[J]. *激光与光电子学进展*, 2019, 56(22): 221502.  
Wang Yixuan, Wu Xiaojun, Xu Tianyang. Tracking algorithm via temporal consistency and spatial pruning based correlation filter with multiple features[J]. *Laser & Optoelectronics Progress*, 2019, 56(22): 221502.
- [21] Danelljan M, Häger G, Khan F S, et al. Learning spatially regularized correlation filters for visual tracking[C]// 2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile: IEEE, 2015: 4310-4318.
- [22] Lukezic A, Vojir T, Zajc L C, et al. Discriminative correlation filter with channel and spatial reliability[C]// IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE Computer Society, 2017: 4847-4856.
- [23] Lin Z C, Liu R S, Su Z X. Linearized alternating direction method with adaptive penalty for low-rank representation[C]// NIPS 2011: Neural Information Processing Systems Conference. Granada, Spain: NIPS, 2011: 612-620.
- [24] Boyd S, Parikh N, Chu E., et al. Distributed optimization and statistical learning via the alternating direction method of multipliers[J]. *Foundations and Trends in Machine Learning*, 2011, 3(1): 1-122.
- [25] Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks[C]// NIPS 2012: Neural Information Processing Systems Conference. Lake Tahoe, Nevada, USA: NIPS, 2012: 1106-1114.

#### 作者简介:



王译萱(1992-):男,硕士研究生,研究方向:目标跟踪,  
E-mail:yixuan\_wang\_jun@163.com。



吴小俊(1967-):男,教授,博士生导师,研究方向:模式识别、计算机视觉和人工智能, E-mail: xiaojun\_wu\_jnu@163.com。

(编辑:张彤)