

基于双模型竞争机制的目标跟踪算法

林 彬^{1,2}, 王华通³, 封全喜^{1,2}

(1. 桂林理工大学理学院, 广西桂林 541004; 2. 广西高校应用统计重点实验室, 广西桂林 541004;
3. 广东外语外贸大学信息科学与技术学院, 广东广州 510006)

摘要: 为解决背景感知相关滤波器存在的特征表达能力不足和模型漂移问题, 本文提出了一种基于双模型竞争机制的目标跟踪算法. 一方面, 本文基于颜色和梯度信息设计了一种简单高效的特征描述子, 以实现更鲁棒的目标表现建模. 另一方面, 本文分别构建初始模型和变化模型作用于目标搜索区域, 并根据两者的跟踪响应图置信度来决定跟踪结果. 跟踪过程中, 随着双模型主导地位不断地动态切换, 变化模型也被赋予了可逆向学习的能力, 从而达到缓解模型漂移的效果. 实验结果表明, 相比于基准算法, 本文算法在 OTB2015、TinyTLP 和 UAV20L 三个数据集的跟踪精度分别提升 5.0%、1.3% 和 4.1%, 跟踪成功率分别提升 3.8%、2.8% 和 1.7%, 且在对不同跟踪场景实现稳定跟踪的同时能够保持 25.5 fps 的实时跟踪速度.

关键词: 目标跟踪; 相关滤波; 双模型竞争机制; 特征描述子; 跟踪置信度; 模型漂移

基金项目: 国家自然科学基金(No.62166015); 广西自然科学基金(No.2019GXNSFBA245056)

中图分类号: TP391.4

文献标识码: A

文章编号: 0372-2112(2023)05-1381-07

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20221375

Object Tracking Algorithm Based on Dual-Model Competition Mechanism

LIN Bin^{1,2}, WANG Hua-tong³, FENG Quan-xi^{1,2}

(1. College of Science, Guilin University of Technology, Guilin, Guangxi 541004, China;

2. Guangxi Colleges and Universities Key Laboratory of Applied Statistics, Guilin, Guangxi 541004, China;

3. School of Information Science and Technology, Guangdong University of Foreign Studies, Guangzhou, Guangdong 510006, China)

Abstract: To solve the problems of insufficient feature expression ability and model drift in the background-aware correlation filters, this paper proposes an object tracking algorithm based on a dual-model competition mechanism. On the one hand, a simple and efficient feature descriptor that integrates color and gradient information is designed to achieve more robust target appearance modeling. On the other hand, we construct two filter models to describe the object's initial appearance and its variations, and then apply them to the target searching area respectively. The tracking results are determined by the confidence of the tracking response maps corresponding to these two models. During the tracking process, with the dynamic switching of the dominant position of the two models, the filter model for adapting to object variations is also endowed with the ability of reversible learning to alleviate the model drift. The experimental results show that, compared with the baseline tracker, the tracking precision of the proposed algorithm on OTB2015, TinyTLP and UAV20L datasets is improved by 5.0%, 1.3% and 4.1%, and the tracking success rate is improved by 3.8%, 2.8% and 1.7%. The proposed algorithm can also achieve stable tracking performance for different tracking scenarios while maintaining a running speed of 25.5 fps.

Key words: object tracking; correlation filter; dual-model competition mechanism; feature descriptor; tracking confidence; model drift

Foundation Item(s): National Natural Science Foundation of China (No.62166015); Guangxi Natural Science Foundation (No.2019GXNSFBA245056)

1 引言

目标跟踪是计算机视觉领域的重要研究内容,在视频监控、自动驾驶、智能机器人、人机交互等方面具有广泛的应用价值^[1]. 经过近十余年的快速发展,目标跟踪技术已经取得了长足的进步,但由于受到光照变化、尺度变化、遮挡等多种干扰因素的影响,实现鲁棒的在线跟踪仍然是一项极具挑战性的任务.

自 2010 年 Bolme 等^[2]提出 MOSSE (Minimum Output Sum of Squared Error) 算法以来,相关滤波类方法凭借其性能优势逐渐成为了领域内的主导性方法. 在 MOSSE 的基础上, Henriques 等^[3]提出的 KCF (Kernelized Correlation Filter) 算法将原有的灰度特征替换为多通道的方向梯度直方图 (Histogram of Oriented Gradient, HOG) 特征以提升模型的特征表达能力,并利用循环矩阵构造虚拟负样本的方式在目标搜索区域进行密集采样,以实现训练样本的扩充. 然而,由中心图像块经过循环移位得到的虚拟样本会导致边界效应问题,影响滤波器对目标和背景的判别能力^[4]. 为此, Galoogahi 等^[5]提出了背景感知相关滤波器,即 BACF (Background-Aware Correlation Filter) 算法,通过引入一个掩膜矩阵在更大的搜索区域内对样本进行裁剪,利用在循环采样时获取到的真实背景来训练滤波器. 虽然 BACF 取得了令人瞩目的跟踪性能,却依然存在以下两点缺陷: (1) 与 KCF 类似, BACF 仅使用单一的 HOG 特征对目标进行表观建模,特征表达能力十分有限. 近

年来,颜色名 (Color Names, CN)^[6,7]、颜色直方图^[8]等颜色特征和利用卷积神经网络 (Convolutional Neural Networks, CNN) 提取的深度特征^[9-12]陆续被引入到相关滤波跟踪算法中,但如何在保证算法实时性的前提下进行特征选择和特征融合,仍是一个开放性的问题. (2) BACF 在每一帧采用固定方式进行模型更新,在应对复杂环境或长时跟踪场景时易由遮挡、目标运动出视野等因素导致模型漂移. 一些算法^[13,14]以 PSR (Peak-to-Sidelobe Ratio)、APCE (Average Peak-to-Correlation Energy) 等响应图评价指标为依据,当指标得分低于某一固定阈值时阻止模型更新、防止模型漂移,但由于模型始终处于单向的更新模式 (更新过程不可逆),因此一旦发生漂移则模型无法从错误状态中恢复.

为解决上述问题,本文提出了一种基于双模型竞争机制的目标跟踪算法,通过设计新的特征描述子和可逆向学习策略,高效地提升模型的特征表达能力并达到抑制模型漂移的效果.

2 本文算法

本节首先介绍本文所提出的特征描述子,再介绍背景感知相关滤波器的训练方式,最后着重阐述双模型竞争机制的具体内容. 本文算法的总体流程如图 1 所示,图中 PCA (Principal Components Analysis) 表示主成分分析技术,FFT (Fast Fourier Transform) 和 IFFT (Inverse FFT) 表示傅里叶变换及其逆变换,⊗ 表示相关滤波操作,频域内为哈达玛 (Hadamard) 积.

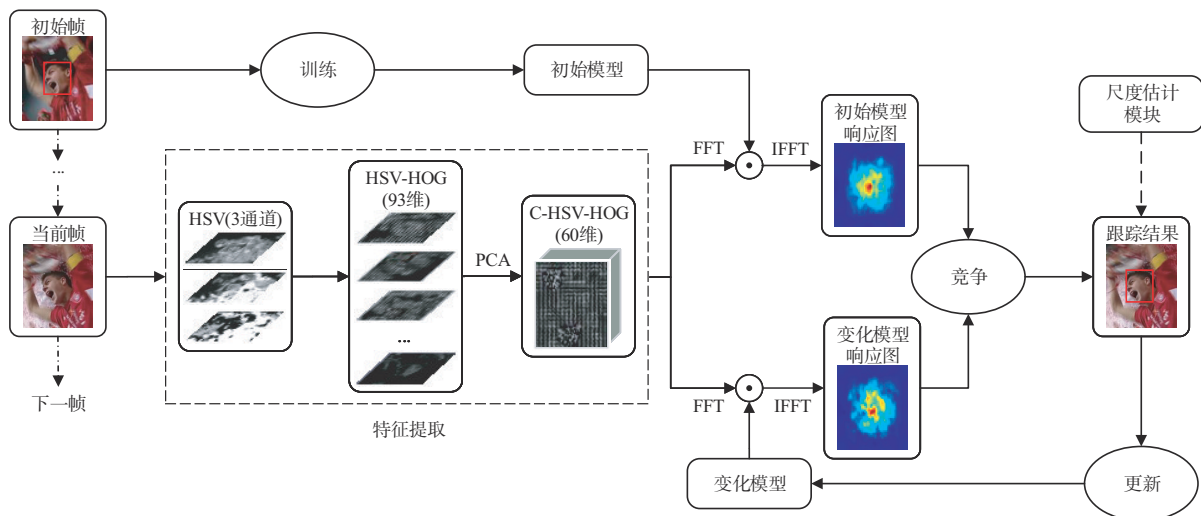


图1 算法总体流程

2.1 C-HSV-HOG 特征描述子

特征提取部分是整个目标跟踪框架中的关键组件之一,恰当的特征选择和特征设计可以显著提升算法的跟踪性能. 现有工作已经验证了梯度、颜色等人工特征以及从卷积神经网络中提取的深度特征对相关滤波

跟踪算法的有效性. 考虑到引入深度特征对跟踪速度的影响,本文基于颜色和梯度信息的互补特性来设计特征提取方案. 首先,将输入的RGB图像转换到HSV (Hue, Saturation, Value) 颜色空间. 相比于RGB空间,HSV空间更接近于人眼视觉的感知方式. 紧接着,分别

对 HSV 空间的每个通道提取 HOG 特征,并将得到的所有 HOG 特征拼接起来,组成一个 93 维的特征向量.最后,通过 PCA 进行特征压缩以去除特征向量中的冗余成分,压缩后的特征向量(经实验,固定为 60 维)可称之为 C-HSV-HOG (Compressed HSV-HOG) 描述子.由此得到的融合特征除具备颜色特征对目标图像的尺寸、方向、视角变化的不敏感性外,也具有 HOG 特征对光照变化及目标几何形变的鲁棒性优势,且计算较为简单高效.

2.2 背景感知相关滤波器

对输入图像提取 C-HSV-HOG 特征,将得到的特征向量记为 \mathbf{x} ,并用 \mathbf{y} 表示预期的相关滤波响应输出(预定义为高斯函数分布),则可根据文献[5]中公式,通过求解最小化目标函数来训练背景感知相关滤波器 \mathbf{h} ,即:

$$E(\mathbf{h}) = \frac{1}{2} \sum_{j=1}^L \|\mathbf{y}_j - \sum_{k=1}^K \mathbf{h}_k^T \mathbf{P} \text{circ}(\mathbf{x}_k)\|_2^2 + \frac{\lambda}{2} \sum_{k=1}^K \|\mathbf{h}_k\|_2^2 \quad (1)$$

其中, \mathbf{h}_k 和 \mathbf{x}_k 分别表示第 k 个通道的滤波器和特征向量, $\mathbf{h}_k \in \mathbb{R}^D$ 、 $\mathbf{x}_k \in \mathbb{R}^L$, L 和 D 分别对应目标搜索区域和滤波器的大小,且 $L \gg D$, K 为特征向量的通道数, \mathbf{y}_j 表示 \mathbf{y} 中第 j 个元素, $\mathbf{y} \in \mathbb{R}^L$, $\text{circ}(\cdot)$ 表示循环移位操作, \mathbf{P} 表示一个 $D \times L$ 的二值矩阵,用于在 L 维的特征向量中裁剪出 D 维的真实负样本, $\mathbf{P} \text{circ}(\mathbf{x}_k)$ 可生成所有循环移位样本的第 k 维特征, λ 为正则化参数,上标“ T ”表示向量或矩阵的转置.为简化计算,式(1)可转换到频域进行快速求解.引入辅助变量 $\hat{\mathbf{g}} = (\hat{\mathbf{g}}_1^H, \dots, \hat{\mathbf{g}}_K^H)^H$,上标“ \cdot ”表示频域形式,上标“ H ”表示复向量或矩阵的共轭转置,则可将式(1)改写为:

$$E(\mathbf{h}, \hat{\mathbf{g}}) = \frac{1}{2} \|\hat{\mathbf{y}} - \hat{\mathbf{X}}\hat{\mathbf{g}}\|_2^2 + \frac{\lambda}{2} \|\mathbf{h}\|_2^2 \quad (2)$$

$$\text{s.t. } \hat{\mathbf{g}} = \sqrt{T} (\mathbf{F}\mathbf{P}^T \otimes \mathbf{I}_K) \mathbf{h}$$

其中, $\hat{\mathbf{X}} = (\text{diag}(\hat{\mathbf{x}}_1)^H, \dots, \text{diag}(\hat{\mathbf{x}}_K)^H)$, 函数 $\text{diag}(\cdot)$ 表示向量或矩阵对角化, $\mathbf{h} = (\mathbf{h}_1^T, \dots, \mathbf{h}_K^T)^T$, \mathbf{F} 为 $L \times L$ 的正交矩阵,用于将任意的 L 维向量或矩阵映射到其频域, \mathbf{I}_K 为 $K \times K$ 的单位矩阵, \otimes 表示克罗内克(Kronecker)积.式(2)可通过建立增广拉格朗日方程并利用交替方向乘法(Alternating Direction Method of Multipliers, ADMM)[5]进行迭代求解.后续操作中,使用 $\hat{\mathbf{g}}$ 作为滤波器用于目标的中心定位和尺度估计.

2.3 双模型竞争机制

目标跟踪问题的本质可以认为是一个在线学习问题. BACF 等相关滤波跟踪算法对每一帧进行模型更新操作,不断将目标模型向新的目标外观方向上进行微调.但是,如跟踪过程中出现遮挡、目标运动出视野等干扰因素,这种固定的单向学习模式容易引入背景噪声并产生跟踪误差,并且随着时间的推移,累积误差将导致模型漂移.此时,第一帧的初始模型反而具有更准

确的目标描述能力.因此,如能赋予跟踪算法可逆向的学习能力,使得目标模型可以自适应地根据跟踪结果在时间维度上进行正向或逆向更新,将能够在一定程度上有效地缓解模型漂移问题.基于此思路,本文提出了双模型竞争机制.首先,分别构建初始模型(即初始滤波器 $\hat{\mathbf{g}}_i$) 和变化模型(包括变化滤波器 $\hat{\mathbf{g}}_c$ 及其外观表示 $\hat{\mathbf{x}}_c$), 初始滤波器和变化滤波器在第一帧时相同,均由求解式(2)得到,但前者在整个跟踪过程中保持不变,后者则逐帧更新.将当前帧记为第 t 帧,对目标搜索区域对应的图像进行特征提取,用 $\hat{\mathbf{x}}_k^{(t)}$ 表示当前帧特征向量的频域形式,则可按式(3)计算当前帧初始模型和变化模型的相关滤波响应图 $\mathbf{r}_i^{(t)}$ 和 $\mathbf{r}_c^{(t)}$:

$$\begin{cases} \mathbf{r}_i^{(t)} = \mathcal{F}^{-1} \left(\sum_{k=1}^K \hat{\mathbf{x}}_k^{(t)} \odot \hat{\mathbf{g}}_{i,k} \right) \\ \mathbf{r}_c^{(t)} = \mathcal{F}^{-1} \left(\sum_{k=1}^K \hat{\mathbf{x}}_k^{(t)} \odot \hat{\mathbf{g}}_{c,k}^{(t-1)} \right) \end{cases} \quad (3)$$

其中, \mathcal{F}^{-1} 表示傅里叶逆变换, $\hat{\mathbf{g}}_{c,k}^{(t-1)}$ 表示上一帧第 k 个通道的变化滤波器.

紧接着,利用文献[14]提出的响应图评价指标 APCE 对 $\mathbf{r}_i^{(t)}$ 和 $\mathbf{r}_c^{(t)}$ 分别计算其响应图置信度,如式(4)所示:

$$s = \frac{|\max(\mathbf{r}) - \min(\mathbf{r})|^2}{\text{mean} \left(\sum_{m,n} (\mathbf{r}(m,n) - \min(\mathbf{r})) \right)} \quad (4)$$

其中,函数 $\max(\cdot)$ 、 $\min(\cdot)$ 和 $\text{mean}(\cdot)$ 分别用于求最大值、最小值和平均值, $\mathbf{r}(m,n)$ 为响应图 \mathbf{r} 中第 m 行、第 n 列的元素值.将当前帧计算得到的初始模型和变化模型的响应图置信度记为 $s_i^{(t)}$ 和 $s_c^{(t)}$,当式(5)条件成立时,变化模型取得主导权:

$$s_c^{(t)} \geq \alpha \cdot s_i^{(t)} \quad (5)$$

其中, α 为平衡因子.反之,当不满足式(5)条件时,则认为初始模型对当前帧的跟踪结果更为可靠,即初始模型取得主导权.取得主导权的模型响应图中,取值最大的坐标位置即作为目标的预测中心.

最后,无论变化模型是否取得当前帧主导权,均使用式(6)更新其外观表示:

$$\hat{\mathbf{x}}_c^{(t)} = (1 - \beta) \hat{\mathbf{x}}_c^{(t-1)} + \beta \hat{\mathbf{x}}_{\text{new}}^{(t)} \quad (6)$$

其中, β 为学习率, $\hat{\mathbf{x}}_c^{(t-1)}$ 为上一帧变化模型的外观表示, $\hat{\mathbf{x}}_{\text{new}}^{(t)}$ 为根据目标预测位置重新计算的频域上的特征向量.将 $\hat{\mathbf{x}}_c^{(t)}$ 代入到式(2)中替换 $\hat{\mathbf{x}}$,并再次利用 ADMM 算法求解,可得更新后的变化滤波器 $\hat{\mathbf{g}}_c^{(t)}$,用于在下一帧中计算变化模型的响应图.跟踪过程中,如当前帧由初始模型主导跟踪结果,则 $\hat{\mathbf{x}}_{\text{new}}^{(t)}$ 将使得变化模型朝初始模型方向更新,即逆向更新.随着初始模型和变化模型交替在竞争中起主导作用,变化模型实现了自适应的正向

或逆向更新过程. 本文算法的尺度估计模块与原BACF算法一致, 具体细节可参阅文献[5].

3 实验结果与分析

本文算法在 Matlab R2017a 平台编程实现, 硬件环境为一颗 Intel Core i7-8700 3.2 GHz CPU (6 核心 12 线程) 和 16 GB RAM. 算法的主要参数设置如下: 正则化参数 λ 设置为 0.001, 双模型竞争机制中的平衡因子 α 为 0.84, 变化模型的学习率 β 为 0.012 5. 为保证比较的公平性, 所有参数值在不同数据集的测试中保持不变. 本文算法的源代码下载地址为 <https://github.com/huatog/DMCF>.

3.1 定量比较

为充分评估本文算法的跟踪性能, 在 OTB2015^[15]、TinyTLP^[16]、UAV20L^[17] 三个数据集上与 10 个性能优越的相关滤波跟踪算法进行了整体性能的定量比较. 对比算法具体包括: SAMF (Scale Adaptive with Multiple

Features tracker)^[7]、STAPLE (Sum of Template And Pixel-wise Learners)^[8]、ACS-DCF (Discriminative Correlation Filter with Adaptive Channel Selection)^[12]、RRDT (Reliable Re-Detection Tracker)^[13]、FDSST (Fast Discriminative Scale Space Tracker)^[18]、CCT (Collaborative Correlation Tracker)^[19]、STRCF (Spatial-Temporal Regularized Correlation Filter)^[20]、DRCF (Dual Regularized Correlation Filter)^[21]、ARCF (Aberrance Repressed Correlation Filter)^[22], 以及基准算法 BACF^[5]. 评价指标采用领域内广泛使用的跟踪精度和成功率.

OTB2015 和 TinyTLP 均为短时跟踪数据集, 分别包含 100 组和 50 组视频序列, UAV20L 为无人机跟踪数据集, 包含 20 组更具挑战性的长时视频序列. 图 2 和图 3 分别显示了所有参与比较的算法在这三个数据集上的跟踪精度和成功率曲线. 由图可见, 本文算法的整体性能表现优异, 跟踪精度和成功率结果在三个数据集上均排名第一, 即使在应对长时跟踪场景时也依然表现稳定.

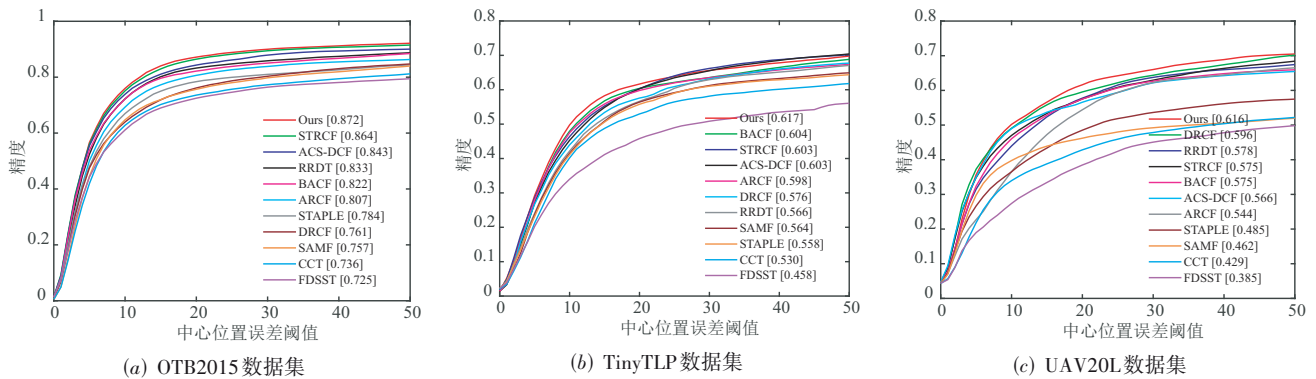


图2 各算法在三个数据集上的跟踪精度曲线

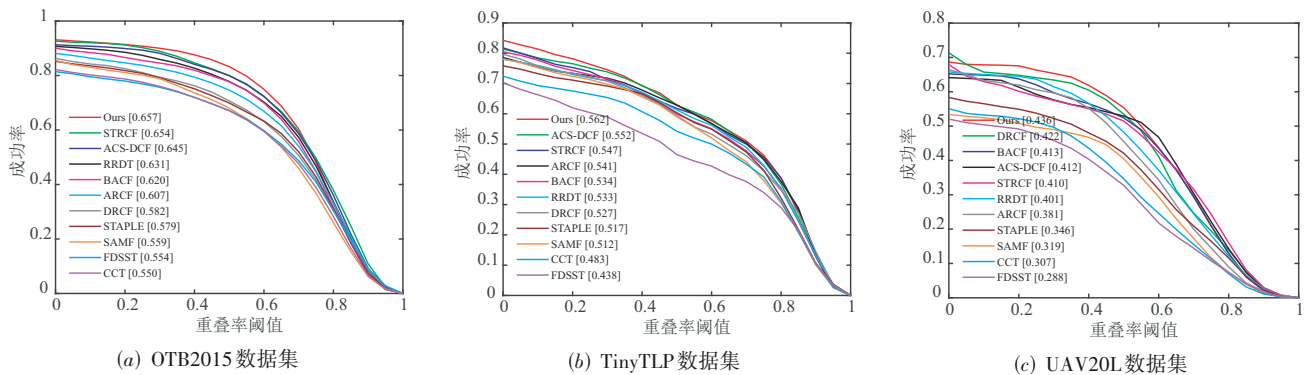


图3 各算法在三个数据集上的跟踪成功率曲线

为评估本文算法在应对不同干扰因素时的跟踪鲁棒性, 在 OTB2015 数据集上对 11 种不同属性测试算法的性能, 将各算法的跟踪精度结果归纳为表 1,

表中每种属性最优的结果加粗表示. 干扰属性具体包括: 光照变化 (Illumination Variation, IV)、尺度变化 (Scale Variation, SV)、遮挡 (Occlusion, OCC)、形变

(Deformation, DEF)、运动模糊(Motion Blur, MB)、快速运动(Fast Motion, FM)、平面内旋转(In-Plane Rotation, IPR)、平面外旋转(Out-Of-Plane Rotation, OPR)、运动出视野(Out-of-View, OV)、背景杂乱(Background

Clutters, BC)、低分辨率(Low Resolution, LR)。由表 1 可见,本文算法在除 DEF(排名第二)和 OV(排名第三)外的 9 项属性测试中排名第一,展现了较好的跟踪鲁棒性。

表 1 不同干扰属性下各算法的跟踪精度结果

算法	IV	SV	OCC	DEF	MB	FM	IPR	OPR	OV	BC	LR
SAMF	0.728	0.717	0.737	0.688	0.692	0.669	0.718	0.718	0.747	0.704	0.709
STAPLE	0.782	0.731	0.728	0.751	0.719	0.709	0.767	0.739	0.668	0.749	0.591
ACS-DCF	0.802	0.849	0.832	0.790	0.844	0.800	0.818	0.848	0.790	0.827	0.770
RRDT	0.818	0.785	0.769	0.800	0.804	0.750	0.775	0.794	0.728	0.822	0.605
FDSST	0.751	0.670	0.636	0.610	0.711	0.698	0.734	0.668	0.577	0.780	0.617
CCT	0.742	0.695	0.681	0.657	0.666	0.668	0.710	0.702	0.591	0.749	0.650
STRCF	0.841	0.844	0.814	0.844	0.837	0.793	0.812	0.848	0.766	0.872	0.756
DRCF	0.744	0.704	0.647	0.765	0.729	0.742	0.703	0.701	0.597	0.735	0.599
ARCF	0.766	0.775	0.740	0.769	0.772	0.763	0.785	0.770	0.671	0.760	0.714
BACF	0.820	0.771	0.740	0.802	0.774	0.781	0.784	0.785	0.714	0.858	0.711
Ours	0.884	0.857	0.833	0.826	0.859	0.827	0.843	0.859	0.762	0.924	0.777

3.2 定性比较

图 4 展示了本文算法与 4 个代表性算法(基准算法 BACF 及在定量比较中表现突出的 STRCF、ACS-DCF 和 DRCF)在部分视频上的定性比较结果。图中按照从左到右、从上到下的次序,6 个视频依次为选自 OTB2015 中的 Rubik、Box、Ironman, TinyTLP 中的 Sam、DriftCar2, 和 UAV20L 中的 Person19。由图可见, Rubik 视频中存在目标旋转、尺度变化等干扰因素,只有本文算法能够较为精准的预测目标位置和尺度; Box 视频中目标在运动过程中被遮挡, STRCF、ACS-DCF 和本文算法能够较好地应对,而 BACF 和 DRCF 则发生了模型漂移; Ironman 视频中存在目标快速运动、运动模糊、背景杂乱、低分

辨率及剧烈光照变化等多种干扰因素,跟踪难度极大,在所有对比算法均彻底丢失目标的情况下本文算法仍然能够跟踪成功;受目标形变、尺度变化和背景干扰的影响,除本文算法外,其余算法也无法完成对 Sam 视频中目标的准确定位; DriftCar2 视频中目标始终快速运动且呈现出较大程度的尺度变化, BACF、STRCF 和 DRCF 跟踪失败, ACS-DCF 在中途也曾偏离目标,而本文算法的中心定位和尺度估计结果最为稳定; Person19 视频中目标逐渐运动出摄像机视野外,当目标重新出现时,所有对比算法由于已经发生模型漂移,均无法从错误状态中恢复,只有本文算法依靠初始模型的作用再次成功定位到目标。

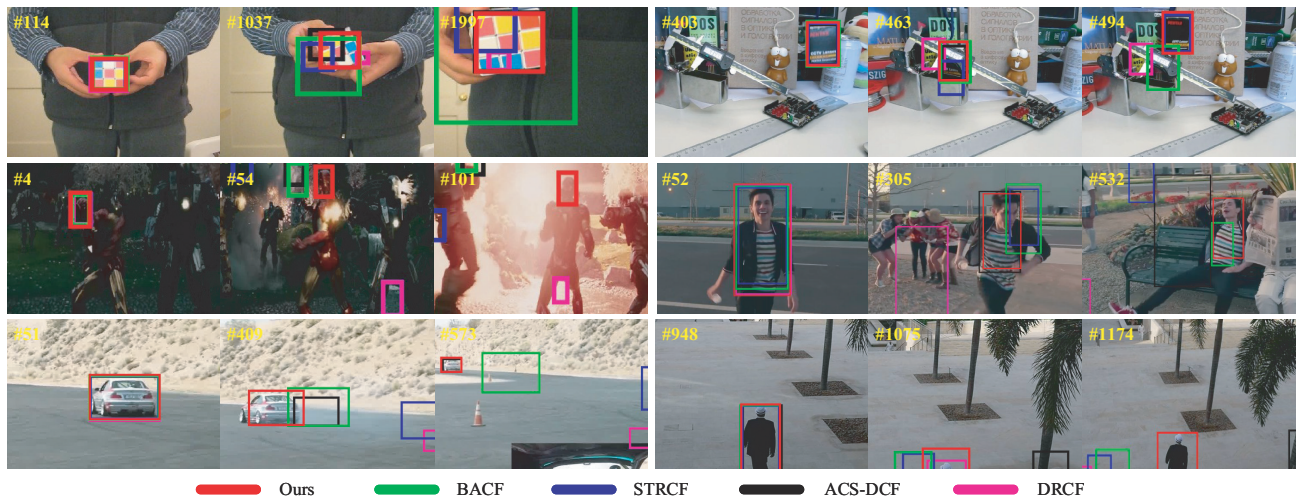


图 4 部分视频的定性比较结果

3.3 特征改进的有效性分析

为进一步评估本文特征改进方案对算法性能的影响,

将基准算法 BACF 结合不同的特征描述子在 OTB2015 数据集中的 74 组彩色视频序列上进行测试,

结果如表2所示.表中,HOGCN表示HOG和CN特征进行串联融合,融合后为41维的特征向量;HSV-HOG表示先转换到HSV空间,然后分别对每个通道提取HOG特征再合并融合,融合后为93维的特征向量,但不进行特征压缩;C-HSV-HOG即为特征压缩后60维的特征向量.表3说明使用HSV-HOG是优于HOGCN的,特征压缩后可进一步提高跟踪性能.图5是HSV-HOG压缩至不同维度时生成的跟踪精度折线图,由图可见该融合特征降到60维时可得到最优结果.

表2 不同特征改进方案在OTB彩色视频数据集的测试结果

算法	精度	成功率
BACF	0.782	0.595
BACF + HOGCN	0.804	0.601
BACF + HSV-HOG	0.807	0.619
BACF + C-HSV-HOG	0.819	0.625

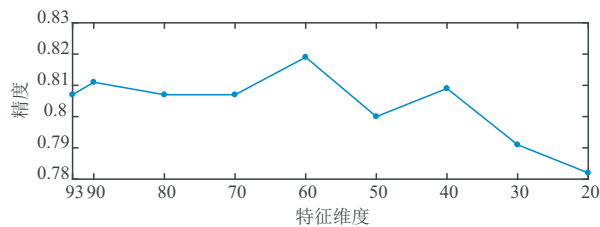


图5 OTB彩色视频数据集的特征压缩维度测试结果

3.4 消融性分析

表3为本文算法在OTB2015完整数据集上的消融性测试结果(算法在处理灰度视频时退化为HOG单特征),表中FI(Feature Improvement)表示特征改进,DM(Dual-Model)表示双模型竞争机制.由表3可见,特征改进可帮助基准算法在跟踪精度和成功率方面分别提升3.0%和2.3%;继续叠加双模型竞争机制后,算法的跟踪精度和成功率能够再次提升2.0%和1.5%;实时性方面,本文算法在CPU环境下的平均跟踪速度为25.5 fps,能够满足在线跟踪需求.此外,本文提出的C-HSV-HOG特征和双模型竞争机制具有较好的泛化性,可与其它相关滤波跟踪算法相结合以进一步提高跟踪性能.

表3 OTB2015数据集的消融性测试结果

算法	精度	成功率	速度/fps
BACF	0.822	0.620	47.2
BACF + FI	0.852	0.643	26.6
BACF + FI + DM	0.872	0.658	25.5

4 结论

本文以背景感知相关滤波跟踪框架为基础,提出了基于双模型竞争机制的目标跟踪算法.算法通过引入双模型竞争机制,将相关滤波类算法普遍采用的单

向更新模式巧妙地转化为自适应的可逆向更新模式,能够有效地抑制由遮挡等因素引发的模型漂移问题;结合本文所设计的C-HSV-HOG特征描述子,则进一步提升了算法在应对光照变化、目标形变、尺度变化等环境时的鲁棒性. OTB2015、TinyTLP、UAV20L三个数据集上的定量、定性实验结果充分地验证了本文算法的先进性.并且,本文算法能够在单CPU支持下满足在线跟踪任务的实时性需求,非常适合于部署在无人机等轻负载平台,具有较高的实际应用价值.

参考文献

- [1] 韩瑞泽, 冯伟, 郭青, 等. 视频单目标跟踪研究进展综述[J]. 计算机学报, 2022, 45(9): 1877-1907.
HAN R Z, FENG W, GUO Q, et al. Single object tracking research: A survey[J]. Chinese Journal of Computers, 2022, 45(9): 1877-1907. (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. Piscataway: IEEE, 2010: 2544-2550.
- [3] HENRIQUES J F, CASEIRO R, MARTINS P, et al. High-speed tracking with kernelized correlation filters[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(3): 583-596.
- [4] 仇祝令, 查宇飞, 吴敏, 等. 基于注意力学习的正则化相关滤波跟踪算法[J]. 电子学报, 2020, 48(9): 1762-1768.
QIU Z L, ZHA Y F, WU M, et al. Learning attentional regularized correlation filter for visual tracking[J]. Acta Electronica Sinica, 2020, 48(9): 1762-1768. (in Chinese)
- [5] GALOOGAHI H K, FAGG A, LUCEY S. Learning background-aware correlation filters for visual tracking[C]// 2017 IEEE International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2017: 1144-1152
- [6] DANELLJAN M, KHAN F S, FELSBERG M, et al. Adaptive color attributes for real-time visual tracking[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2014: 1090-1097.
- [7] LI Y, ZHU J K. A scale adaptive kernel correlation filter tracker with feature integration[C]//European Conference on Computer Vision Workshops. Cham: Springer, 2015: 254-265.
- [8] BERTINETTO L, VALMADRE J, GOLODETZ S, et al. Staple: complementary learners for real-time tracking[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2016: 1401-1409.

- [9] MA C, HUANG J B, YANG X K, et al. Hierarchical convolutional features for visual tracking[C]//2015 IEEE International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2016: 3074-3082.
- [10] 蒲磊, 冯新喜, 侯志强, 等. 基于二阶池化网络的鲁棒视觉跟踪算法[J]. 电子学报, 2020, 48(8): 1472-1478.
PU L, FENG X X, HOU Z Q, et al. Robust visual tracking based on second order pooling network[J]. Acta Electronica Sinica, 2020, 48(8): 1472-1478. (in Chinese)
- [11] 蒲磊, 冯新喜, 侯志强, 等. 基于深度空间正则化的相关滤波跟踪算法[J]. 电子学报, 2020, 48(10): 2025-2032.
PU L, FENG X X, HOU Z Q, et al. Correlation filter tracking based on deep spatial regularization[J]. Acta Electronica Sinica, 2020, 48(10): 2025-2032. (in Chinese)
- [12] XU T Y, FENG Z H, WU X J, et al. Adaptive channel selection for robust visual object tracking with discriminative correlation filters[J]. International Journal of Computer Vision, 2021, 129(5): 1359-1375.
- [13] WANG N, ZHOU W G, LI H Q. Reliable re-detection for long-term tracking[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2019, 29(3): 730-743.
- [14] WANG M M, LIU Y, HUANG Z Y. Large margin object tracking with circulant feature maps[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2017: 4800-4808.
- [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] MOUDGIL A, GANDHI V. Long-term visual object tracking benchmark[C]//Asian Conference on Computer Vision. Cham: Springer, 2019: 629-645.
- [17] MUELLER M, SMITH N, GHANEM B. A benchmark and simulator for UAV tracking[C]//European Conference on Computer Vision. Cham: Springer, 2016: 445-461.
- [18] DANELLJAN M, HÄGER G, KHAN F S, et al. Discriminative scale space tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(8): 1561-1575.
- [19] ZHU G B, WANG J Q, WU Y, et al. Collaborative correlation tracking[C]//Proceedings of the British Machine Vision Conference (BMVC). Durham: BMVA, 2015: 184.1-184.12.
- [20] LI F, TIAN C, ZUO W M, et al. Learning spatial-temporal regularized correlation filters for visual tracking[C]//2018 IEEE/CVF Conference on Computer Vision and Pat-

tern Recognition. Piscataway: IEEE, 2018: 4904-4913.

- [21] FU C H, XU J T, LIN F L, et al. Object saliency-aware dual regularized correlation filter for real-time aerial tracking[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 58(12): 8940-8951.
- [22] HUANG Z Y, FU C H, LI Y M, et al. Learning aberrance repressed correlation filters for real-time UAV tracking [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2020: 2891-2900.

作者简介



林 彬 男, 1985年4月出生于广西壮族自治区桂林市. 现为桂林理工大学理学院副教授. 主要研究方向为计算机视觉.
E-mail: linbin@glut.edu.cn



王华通 男, 1997年7月出生于广西壮族自治区玉林市. 现为广东外语外贸大学硕士研究生. 主要研究方向为计算机视觉.
E-mail: 20211050021@gdufs.edu.cn



封全喜(通讯作者) 男, 1980年2月出生于湖南省衡阳市. 现为桂林理工大学理学院教授. 主要研究方向为机器学习及其应用.
E-mail: fqx9904@163.com