

目标跟踪中基于光流映射的模板更新算法

张静, 郝志晖, 刘婧*

天津大学电气自动化与信息工程学院, 天津 300072

摘要 基于深层孪生网络的目标跟踪算法普遍缺乏目标模板在线更新方法, 从而在某些复杂应用场景中适应能力较差。针对这一问题, 提出一种基于光流映射的目标模板在线更新算法, 该算法能够在保证实时性的前提下有效提高对复杂场景的适应能力。首先在跟踪过程中计算模板帧之间的光流信息; 其次由光流映射和残差计算获取目标的运动变化信息。除此以外, 还提出一种基于奇异值分解的由初始帧生成的修正项以修正目标位置偏差的方法。在 OTB100 和 VOT2016 数据集上对所提算法进行测试评估, 结果显示, 所提算法可以较好地优化新生成的目标模板, 增强算法的鲁棒性, 且与现有的跟踪算法相比, 所提算法结果更佳。

关键词 机器视觉; 目标跟踪; 孪生网络; 光流映射; 模板更新

中图分类号 TP391.4 **文献标志码** A

doi: 10.3788/LOP57.221507

Template-Updating Algorithm Based on Optical Flow Mapping in Object Tracking

Zhang Jing, Hao Zhihui, Liu Jing*

School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China

Abstract The tracking algorithm based on deep-level siamese network generally lacks the capability to update the target template online; hence, it exhibits poor adaptability in some complex application environments. Aiming to resolve this problem, a target template online-updating algorithm based on optical flow mapping is proposed herein. On the premise of ensuring real-time operation, the proposed algorithm can efficiently improve its adaptability in complex circumstances. First, the optical-flow information between the template frames is calculated in the tracking process. Then, the information of motion change is generated via optical flow mapping and residual calculation. Furthermore, based on singular value decomposition, a method that creates a correction term via the initial frame, which modifies the target-position deviation, is proposed herein. Finally, the proposed algorithm is tested on OTB100 and VOT2016 datasets. The results show that the proposed algorithm can optimize the new target template to enhance the robustness and can achieve the best results compared with existing tracking algorithms.

Key words machine vision; object tracking; siamese network; optical flow mapping; template update

OCIS codes 150.1135; 100.4996; 100.4999

1 引言

基于视频图像的计算机视觉在当今社会中的地位越来越高, 目标跟踪作为其中一个非常重要的研究方向, 近年来受到了国内外研究者的广泛关注, 也取得了一系列令人瞩目的研究成果, 目前已广泛应用于视频监控、自动驾驶及人机交互等领域^[1]。简要来讲, 目标跟踪的基本任务是: 在视频的

第一帧中人为给定初始目标后, 在后续帧中逐帧计算目标的位置和尺寸大小。由于任务场景的复杂性和多样性, 目标跟踪算法仍然面临很多具有挑战性的问题, 如背景杂乱、有相似干扰物、遮挡、形变、目标平面内和外旋转等^[2]。虽然目标跟踪技术种类繁多且发展迅猛, 但至今仍然没有一个通用的跟踪算法能够完全解决这些挑战性问题。

目标跟踪算法按照跟踪方法主要分为两类, 分

收稿日期: 2020-03-06; 修回日期: 2020-04-08; 录用日期: 2020-04-30

基金项目: 国家自然科学基金(61701341)、天津大学自主创新基金(2020XRG-0021)

* E-mail: jliu_tju@tju.edu.cn

别是相关滤波类目标跟踪算法和深度学习类目标跟踪算法。其中,相关滤波类目标跟踪算法的核心思想是通过使用手工标注或深层网络提取出目标特征来训练滤波器,计算目标模板与输入图像的相关性,依靠相关图判断目标位置和大小。在这类算法中,一些工作会通过强化滤波器的训练过程、优化滤波器模板的求解方法来使其具有良好的二分类性能,主要包括:Danelljan 等^[3]提出的 C-COT 跟踪算法,该算法利用多尺度特征训练滤波器,获得亚像素精度的目标位置;Danelljan 等^[4]提出的 ECO 跟踪算法,该算法在 C-COT^[3]的基础上进一步提升了时间效率和空间效率,调整了模型大小、训练集大小及更新机制;孟球等^[5]提出的双特征模型核相关滤波目标跟踪算法,该算法采用深度特征和手工特征联合的方式进行训练,提升了滤波器性能;Danelljan 等^[6]提出的 SRDCF 跟踪算法,该算法利用 Gauss-Seidel 函数求解滤波器,同时加入空间正则项来解决边界效应;王殿伟等^[7]提出的基于多特征融合的核相关滤波目标跟踪算法,该算法通过融合梯度方向梯度直方图(HOG)特征和颜色特征来进行多特征训练及计算;Danelljan 等^[8]提出的 SRDCFdecon 跟踪算法,该算法在 SRDCF^[6]的基础上增强训练样本质量,动态管理训练集,提升了滤波器的训练质量;Galoogahi 等^[9]提出的 BACF 跟踪算法,该算法通过利用多通道特征来求解滤波器最小化岭回归问题,进而求解滤波器模板。而另一些工作则通过多特征融合、多尺度计算、改进滤波器模板更新机制等方法来增强滤波器模板的适应能力,以应对跟踪任务中出现的挑战性问题,主要包括:王凯宇等^[10]提出的基于相关滤波器的目标抗遮挡算法,该算法通过保留原目标模板和优化小范围预测,缓解了滤波器对于目标被遮挡时性能下降的问题;杨亚光等^[11]提出的基于相关滤波融合卷积残差学习的目标跟踪算法,该算法融合卷积残差学习,将相关滤波器构建为神经网络中的一层进行端到端的训练和跟踪,缓解了复杂场景下传统单一手工特征表达能力的不足;成悦等^[12]提出的基于模型与尺度更新的相关滤波跟踪算法,该算法通过定性分析改变了在线更新机制,同时采用了快捷的尺度更新方法;虞跃洋等^[13]提出的基于前景感知的时空相关滤波跟踪算法,该算法通过增加具有时空正则化功能的滤波器和重新定义目标重检测器的激活机制,增强了滤波算法的适应能力;李健鹏等^[14]提出的融合多层卷积特征的相关滤波运动目标跟踪

算法,该算法通过丰富特征信息、融合多特征来估计目标的位置和最佳尺度。

虽然上述方法改进了相关滤波类算法,缓解了目标跟踪中的一些问题,但受相关滤波方法本身的局限,在应对挑战性问题时,该类方法的综合性能仍然低于深度学习类方法。深度学习类目标跟踪算法主要分为基于目标检测的跟踪算法和基于孪生网络的跟踪算法。其中,基于目标检测的跟踪算法主要利用多次采样分类正负样本,将跟踪问题转换为前景和背景的二分类问题,主要算法包括:Wang 等^[15]提出的 LMCFF 跟踪算法,该算法利用结构化支持向量机(SVM)分类器分类候选框;Danelljan 等^[16]提出的 ATOM 算法,该算法经过 IoU-Net 多次迭代求解目标位置。这类算法精度较高,但速度较慢,难以满足实时性要求。基于孪生网络的目标跟踪算法主要采用匹配初始化帧目标模板特征和搜索区域特征的方法,将初始化目标模板当作特征分类器的分类前景和背景,该类算法主要包括:Bertinetto 等^[17]提出的 Siam-FC 跟踪算法,该算法通过目标模板特征和搜索区域特征的匹配及多尺度搜索机制确定目标位置和大小;任珈民等^[18]提出的一种改进的基于孪生卷积神经网络的目标跟踪算法,该算法通过构建特征通道之间的相互关系来增强特征表达能力;Li 等^[19]提出的 SiameseRPN++ 跟踪算法,该算法利用多层级联的区域候选网络来提高跟踪精度;周迪雅等^[20]提出的基于孪生网络与注意力机制的目标跟踪方法,该方法通过利用注意力机制来增强网络对目标前景的提取和搜索能力,实现跟踪器性能的提升。这类算法较好地平衡了跟踪算法的速度和精度,既满足实时性要求,又有较好的精度和鲁棒性。但是,这类算法通常缺乏在线更新机制,在一些目标物体运动变化快或被遮挡的场景中,跟踪结果会发生较为明显的偏移。受相关滤波类方法在线更新滤波器模板的启发,本文在孪生网络框架中增加基于光流映射的模板更新机制,既保证了算法的速度和精度,又增强了算法对困难场景的适应能力。

本文提出一种基于孪生网络结构且包含目标模板在线更新的跟踪算法,通过光流映射和逐帧更新目标模板实现对算法性能的提升。利用目标跟踪任务中视频序列和目标空间变化的连续性,通过视频序列两帧之间的光流信息指导生成包含目标变化的新模板;对深层神经网络进行特征提取,有关目标变化的信息也在网络中完成了特征编码;当

搜索区域出现遮挡或目标变化较大等情况时,跟踪器能通过目标变化信息的特征及时捕捉目标运动、适应目标运动变化;此外,还利用初始化帧目标模板生成修正项,修正光流映射产生的位置偏差,抑制跟踪过程中的误差累积。

2 基本原理

2.1 基于光流映射的目标运动表示

当视频序列中的某一物体进行空间运动时,视频序列中每一帧每个像素点的瞬时速度都可以用光流表示,且由于视频序列中相邻两帧之间具有连续性,目标物体运动变化位置上对应的光流变化可以较为准确地描述该物体的运动变化。在目标跟踪任务中,由于视频序列前后帧中的光照强度变化幅度小,目标的运动变化通常在时域和空间域上都具有较强的连贯性,因此视频序列基本满足运用光流计算物体运动变化的前提条件,同时,目标跟踪算法也可将光流信息作指导,以获取目标的运动变化信息。

FlowNetC^[21]是一种基于卷积神经网络(CNN)预测视频序列某两帧或某两幅图像之间光流的网络框架,近年来一直被广泛应用于提取视频序列相邻两帧之间的光流信息中。在堆叠前后两帧图像后,该网络首先通过CNN计算前后两帧图像的特征,再通过互相关层计算前后两帧图像特征之间的关系,最后通过解码层计算前后两帧图像之间的光流。本文利用FlowNet^[21]中的FlowNetC模块,计算目标从起始位置运动到当前位置时目标模板之间的光流,再通过基于双线性插值的映射函数和此光流信息将目标当前帧模板映射到目标初始化帧模板,该模板即为由光流信息指导的目标初始化帧模板,表达式为

$$\mathbf{T}_{0,i}^{\text{wrap}} = W[\mathbf{T}_i, F(\mathbf{T}_0, \mathbf{T}_i)], \quad (1)$$

式中: \mathbf{T}_0 为目标初始化帧模板; \mathbf{T}_i 为跟踪到第*i*帧的目标当前帧模板; $F(\cdot)$ 为FlowNetC^[21]网络; $W[\cdot]$ 为FlowNet^[21]中的映射函数; $\mathbf{T}_{0,i}^{\text{wrap}}$ 为由初始帧到第*i*帧的光流信息指导生成的目标初始化帧模板。受Memtrack算法^[22]中更新目标模板方式的启发,将视频序列中相邻帧的残差表示目标运动的变化情况,以此来更新目标模板。计算物体运动变化信息的表达式为

$$\mathbf{R}_i(x, y) = \begin{cases} \mathbf{T}_i(x, y) - \mathbf{T}_{0,i}^{\text{wrap}}(x, y), & \mathbf{T}_i(x, y) > \mathbf{T}_{0,i}^{\text{wrap}}(x, y) \\ 0, & \mathbf{T}_i(x, y) \leq \mathbf{T}_{0,i}^{\text{wrap}}(x, y) \end{cases}, \quad (2)$$

式中: $\mathbf{R}_i(x, y)$ 为 $\mathbf{T}_i(x, y)$ 和 $\mathbf{T}_{0,i}^{\text{wrap}}(x, y)$ 对应位置的残差,表示初始帧 \mathbf{T}_0 运动到第*i*帧时目标每个像素点的运动变化, $\mathbf{R}_i(x, y)$ 中每个位置值反映目标在该位置的运动情况。由于运动变化本身的连续性,前一帧目标的运动信息对后一帧的目标跟踪具有一定的指导作用,通过(2)式,目标的当前运动信息被富集在初始化帧目标模板上,使得目标模板中包含目标当前的运动变化信息。借助这些信息,算法对目标运动变化的适应能力得到了有效提升。

2.2 基于奇异值分解(SVD)的修正项计算

在跟踪过程中,跟踪结果会随着跟踪过程中的误差积累而发生不同程度的偏移,这些偏移将使光流信息产生偏差,进而导致由光流信息指导生成的目标运动变化信息产生偏差,且由于跟踪算法本身对跟踪结果没有判断能力,这些偏差会逐渐累积并表现为目标位置上的偏差。为缓解上述问题,提出一种基于奇异值分解生成修正项的方法来修正所述目标位置偏差,通过重构初始化帧目标模板及线性叠加其与目标运动变化信息,实现对位置偏差的修正,抑制位置偏差带来的误差累积。首先对目标初始化帧模板进行奇异值分解:

$$\mathbf{T}_0 = \mathbf{u}_0 \cdot \mathbf{S}_0 \cdot \mathbf{v}_0, \quad (3)$$

式中: \mathbf{u}_0 、 \mathbf{v}_0 分别为初始化帧目标模板 \mathbf{T}_0 经过奇异值分解得到的左、右奇异向量; \mathbf{S}_0 为初始化帧目标模板经过奇异值分解得到的奇异值矩阵。由于目标模板中的前景信息对应数值较大的奇异值,因此使用 \mathbf{S}_0 中数值较大的前*k*个奇异值重构图像:

$$\mathbf{T}_0^k = \mathbf{u}_0^k \cdot \mathbf{S}_0^k \cdot \mathbf{v}_0^k, \quad (4)$$

式中: \mathbf{u}_0^k 、 \mathbf{v}_0^k 分别为 \mathbf{u}_0 、 \mathbf{v}_0 前*k*行、前*k*列由 \mathbf{S}_0^k 构造的前*k*个奇异值的对角矩阵; \mathbf{T}_0^k 为重构后的初始化帧目标模板。基于奇异值分解的修正项的表达式为

$$\mathbf{N} = \alpha \mathbf{T}_0^k, \quad (5)$$

式中: α 为超参数,用于调整重构模板; \mathbf{N} 为由初始化帧目标模板生成的修正项。这种方法不仅在一定程度上增强了目标模板的前景信息,提升了特征匹配的准确性,还增加了目标模板前景信息在网络提取特征时的数值占比。当跟踪结果发生偏移时,由于 \mathbf{T}_0^k 主要重构的是 \mathbf{T}_0 的前景信息,而 \mathbf{T}_0 是给定的初始化帧目标模板,不存在位置偏差,对由 \mathbf{T}_0^k 生成的修正项与模板进行线性叠加,新生成的目标模板中的前景信息会被强调,进而在特征提取及匹配时,前景有关的特征会被网络注意,由此抑制光流映射偏差导致的目标位置误差。当跟踪结果较为准确时,由于基于奇异值分解的修正项不影响运动

变化信息的生成,目标模板仍然可以富集目标运动变化信息,所以该修正项可以在一定程度上抑制目标模板中的背景,进而有效抑制跟踪过程中位置偏差带来的误差累积,有效提升跟踪过程中跟踪准确时的精度,降低跟踪偏移时的误差,从而提升跟踪算法的综合性能。

2.3 目标跟踪算法中的模板更新策略

采用 SiameseRPN++^[19] 作为目标跟踪的基础网络框架。与实现该网络框架的目标跟踪方法不同,所提算法在目标模板分支处增加了在线更新机制,根据从初始帧到前一帧目标模板的变化,逐帧更新后一帧的目标模板,表达式为

$$T_{i+1} = T_0 + \lambda R_i + N, \quad (6)$$

式中: R_i 为目标从第 0 帧到第 i 帧的运动变化; T_{i+1} 为第 $i+1$ 帧的目标模板; λ 为超参数,用于调整目标变化信息的占比。

所提跟踪算法的框架如图 1 所示,基础的跟踪框架包含两个分支,分别为目标模板分支和搜索区域分支。首先,将目标模板和搜索区域输入到

ResNet50 网络中提取对应的特征,这些特征包含目标变化信息的特征和目标变化区域所在图像块的特征,在特征图上主要表现为目标变化区域具有较高响应;再将 ResNet50 网络中的后三个 block 输出的三个特征分别输入到三个级联的区域候选网络(RPN)中;最终计算出跟踪结果^[19]。在跟踪过程中,所提算法增加了目标模板在线更新策略,该策略将基础网络框架输出的结果作为阶段性结果,将这一结果和初始化帧目标模板共同输入到目标运动信息生成模块中;同时用修正项生成模块获取基于初始化帧目标模板的修正项;最后将两模块的输出线性叠加,获得新的目标模板,实现目标模板的在线更新。该方法主要影响 RPN 中的特征匹配阶段,使网络对存在目标变化的候选框(proposal)更加敏感,且由于有修正项修正目标位置和尺寸偏差,富集在新目标模板中的目标变化信息也较为准确。总之,所提算法可以有效地捕捉存在目标变化的区域,较好地适应目标变化,在一定程度上提升跟踪算法的精度和鲁棒性。

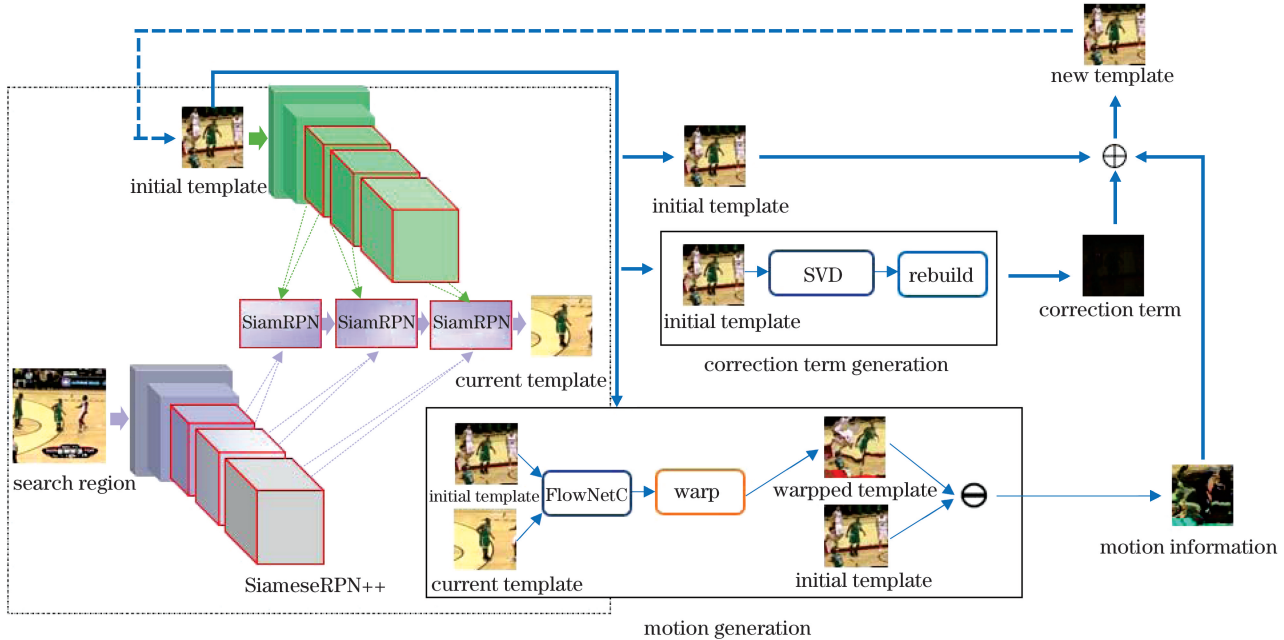


图 1 所提跟踪算法框架示意图

Fig. 1 Schematic of proposed tracking algorithm framework

所提算法运行过程如下:首先经初始化获取初始化帧目标模板及第二帧搜索区域的中心位置和尺寸大小;再将目标模板和搜索区域同时输入到 SiameseRPN++ 网络^[19],获取第二帧的跟踪模板及第三帧搜索区域的中心位置和大小,经过(1)、

(2)式生成由光流信息指导的目标变化信息;再经过(3)~(5)式生成基于初始化帧目标模板的修正项,最后经过(6)式生成第三帧所需的新目标模板;重复上述步骤,逐帧实现目标模板在线更新。

3 实验分析

3.1 实验设置

所提算法采用的网络框架基于 Python、Pytorch^[23]实现,并使用 CPU Intel(R) Core(TM) i7-7800X 和 GPU GeForce GTX 2080Ti 在计算机上实验。采用 SiameseRPN++^[19]作为基础跟踪网络框架,通过光流映射和残差计算获取目标的运动变化信息, $\lambda = 0.01$, 有关修正项 \mathbf{N} 的系数 $k = 100$, $\alpha = 0.0002$ 。算法速度大约为 32.0 frame/s。

3.2 数据集和衡量指标

在标准数据集 OTB100、VOT2016 上分别进行实验。其中 OTB100 数据集包含有 100 个不同的视频序列,其中 25 个为灰度序列。该数据集中每个视频序列都有特定的标签,包含目标跟踪任务中的各种挑战,例如遮挡、尺度变化、形变及光照变化等。对 OTB100 数据集,采用的评价指标为一次通过评估(OPE)。VOT2016 数据集中包含 60 个不同的视频序列,均为彩色序列。对 VOT2016 数据集,采用的评价指标为准确率(Accuracy)、鲁棒性(Robustness)、期望平均重叠率(EAO)。

在 OTB100 数据集上测试算法时,由 OPE 计算精确率图和成功率图。前者通过计算目标预测位置的中心点和真实值之间的距离,计算距离小于设定阈值的帧数占序列总帧数的比例,不同的阈值设定得到的结果不同。后者通过计算目标回归框的预测值和真实值之间的交并比(IoU)并得到重合

率得分(overlap score),计算得分大于设定阈值的帧数占序列总帧数的比例,不同的阈值设定得到的结果也不同。在 VOT2016 数据集上,测试指标中的准确率用于计算视频序列中目标回归框的预测值和真实值之间的平均 IoU,准确率值越大,表明算法的预测结果越接近真实值;鲁棒性值根据在某一序列上跟踪失败(即 IoU 为 0)的总帧数来评估算法稳定性,鲁棒性值越小,表明算法越稳定;EAO 值用于衡量算法在不同特定总帧数的短时视频序列片段上非重置重叠率的平均期望值,由于 EAO 值的计算过程综合评估了准确率和鲁棒性两个指标,因此通常把 EAO 值作为衡量跟踪算法综合性能的重要指标,EAO 值越大,表明跟踪算法的综合性能越好。

3.3 消融实验

为了证明基于光流映射的目标模板在线更新策略和基于奇异值分解的修正项生成方法对目标模板适应目标运动变化有效,在 OTB100、VOT2016 数据集上分别设计了消融实验。在 OTB100 数据集上测试的消融实验结果如图 2 所示,ours 表示在更新过程中只使用目标变化信息,不使用修正项的实验结果;ours+ 表示在更新过程中既使用目标变化信息,又使用修正项的实验结果。可以看出,在使用光流映射指导生成的目标运动变化信息更新模板的同时,使用修正项修正位置偏差可以有效提升目标跟踪算法在各视频序列上的平均精度。

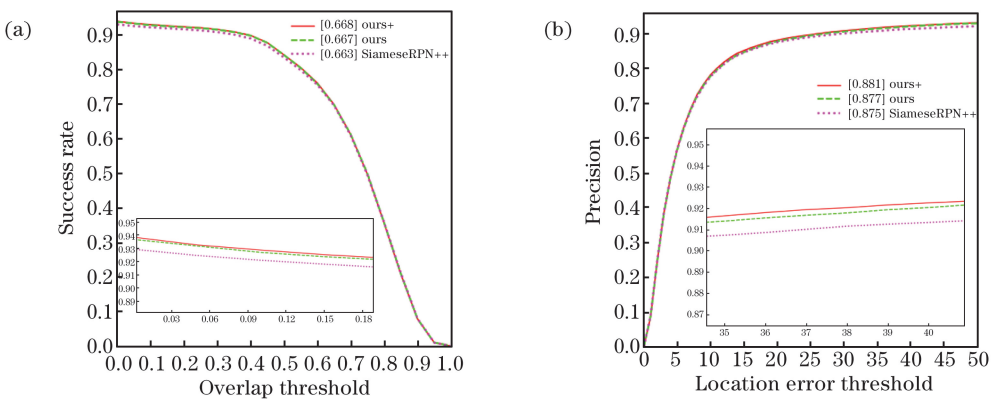


图 2 在 OTB100 数据集上的消融实验结果。(a)成功率;(b)精确率

Fig. 2 Results of ablation experiments on OTB100 dataset. (a) Success rate; (b) precision

在 VOT2016 数据集上的测试结果如表 1 所示。在该数据集上,设置了 3 组消融实验以检验通过光流映射生成的目标模板变化信息和基于奇异

值分解生成的修正项对指导模板更新、修正目标位置偏移是否有效。其中,B(baseline)表示使用基础网络框架,B+W(ours)表示在基础网络框架中增加

模板更新策略, B+W+C(ours+)表示在逐帧更新模板的基础上增加修正项。表1中, EAO表示跟踪算法在不同特定帧数视频片段上的非重置重叠率

的平均期望值, Lost number表示跟踪算法在进行跟踪时跟踪失败的总帧数。

表1 在VOT2016数据集上的消融实验结果

Table 1 Results of ablation experiments on VOT2016 dataset

Network	EAO \uparrow	Robustness \downarrow	Lost number \downarrow	Accuracy \uparrow	Speed/(frame \cdot s $^{-1}$) \uparrow
B (baseline)	0.464	0.196	42.0	0.642	32.3
B+W (ours)	0.474	0.191	41.0	0.642	31.6
B+W+C (ours+)	0.478	0.172	37.0	0.640	32.0

从表1可知:在VOT2016数据集上,三种算法的准确率几乎一致,但基础算法的鲁棒性值为0.196, EAO值为0.464;仅使用目标运动变化信息更新目标模板时,算法的鲁棒性值为0.191, EAO值为0.474,相比于基础网络,鲁棒性值上升了约2.5%, EAO值上升了约2.2%;既使用目标运动变化信息更新模板,又使用修正项修正时,算法的鲁棒性值为0.172, EAO值为0.478,相比于基础网络,鲁棒性值上升了约12.2%, EAO值上升了约3.0%。由上述客观结果分析,目标变化信息的富集对算法综合能力具有一定的提升作用,使用修正项修正位置偏差可以降低跟踪算法的丢失概率,从

而保证算法的稳定性。既使用目标变化信息更新目标模板,又使用修正项抑制误差累积时,算法的综合性能更佳。

此外,为了直观说明基于光流映射的在线更新策略和基于奇异值分解的修正项有效,分别在3个具有挑战性的序列上进行主观结果的比较,并计算在该序列上的平均精度。图3~5依次为Coke、Dudek、Doll三个序列,这些序列分别包含遮挡、平面内旋转及尺度变化三种情况。由图3~5可以看出,当视频序列中出现目标物体被遮挡、旋转或尺度变化时,使用模板在线更新机制并增加修正项的实验结果更佳。



图3 Coke序列中消融实验的主、客观(平均精度)结果

Fig. 3 Subjective and objective (average precision) results of ablation experiments in Coke sequence



图4 Dudek序列中消融实验的主、客观(平均精度)结果

Fig. 4 Subjective and objective (average precision) results of ablation experiments in Dudek sequence

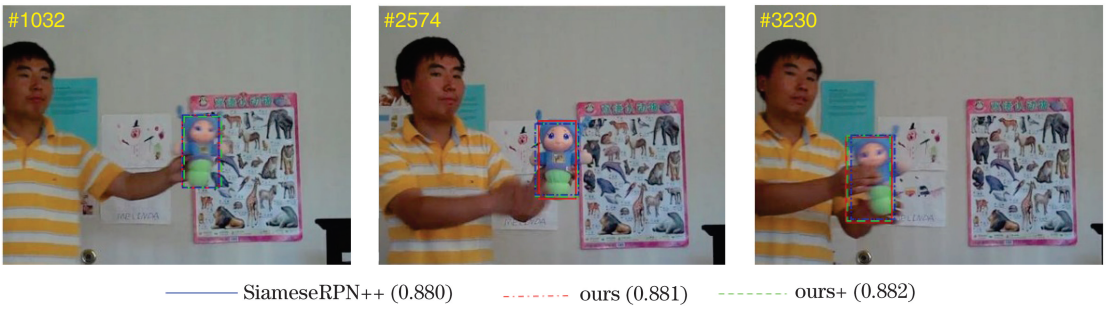


图 5 Doll 序列中消融实验的主、客观(平均精度)结果

Fig. 5 Subjective and objective (average precision) results of ablation experiments in Doll sequence

综上所述,所提目标模板在线更新策略在 OTB100 和 VOT2016 数据集上都具有良好的表现。实验表明,这种在线更新方法为目标模板富集了可靠的目标运动变化信息,能够使目标模板具有一定适应目标变化的能力,从而在跟踪过程中适应后续帧中目标的形态和动作变化。由于跟踪过程本身存在误差累积效应,目标的运动变化信息会随着跟踪过程的跟进而出现偏移、尺度信息失真等情况。所提基于初始化帧目标模板生成的修正项可以在一定程度上抑制这些误差带来的负面影响,降低跟踪失败的概率,提升跟踪算法的稳定性。因此,所提算法可以很好地弥补基础网络框架中目标

模板单一、适应能力差的缺陷,从而改善了跟踪算法的适应能力,提高了跟踪算法的综合性能。

3.4 对比实验

在 OTB100 数据集上对所提算法和 9 个现有算法进行比较,这些算法不仅包括目前性能较好的相关滤波类算法 Staple^[24]、BACF^[14]、LMCF^[15]、SRDCF^[10]、SRDCFdecon^[12] 及 ECO-HC^[8],还包括使用深层特征提取网络的孪生网络类算法 Memtrack^[22]、SiamFCRes22^[25] 及 SiamRPNRes22^[25]。使用 OPE 计算精确率图和成功率图来衡量算法性能。各算法整体性能如图 6 所示。

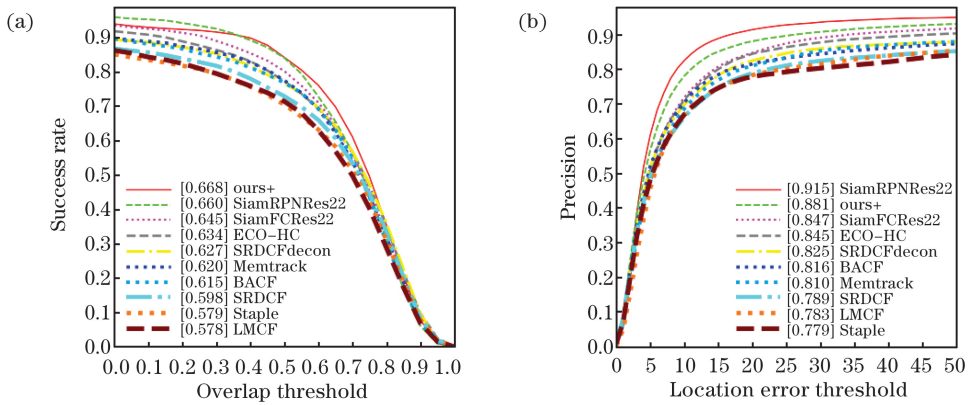


图 6 在 OTB100 数据集上不同算法的实验结果。(a)成功率;(b)精确率

Fig. 6 Experimental results of different algorithms on OTB100 dataset. (a) Success rate; (b) precision

由于增加了目标模板在线更新策略,所提算法能够更好地应对跟踪中目标遮挡、快速变化等挑战性问题,图 7 为 OTB100 数据集中,当视频序列中的目标出现快速运动时算法性能的对比如,可以看出,所提算法可以更好地应对这一难点。

在 VOT2016 数据集上对所提算法和 9 个现有算法进行比较,分别是相关滤波类算法 SRDCF^[10]、C-COT^[4] 及 ECO-HC^[8],深度学习类算法 Siam-

FC^[17]、SiamFCRes22^[25]、SiamRPNRes22^[25]、TADT^[26]、C-RPN^[27] 及 SPM^[28]。采用 EAO 值、鲁棒性值和准确率值三个指标衡量算法性能,结果如表 2 所示。可以看出:较相关滤波类算法和深度学习类算法,所提算法的准确率值有明显提升;与深度学习类算法相比,鲁棒性值略有提升,与相关滤波类算法相比提升明显;EAO 值总体提升较大,表明所提算法性能较好。

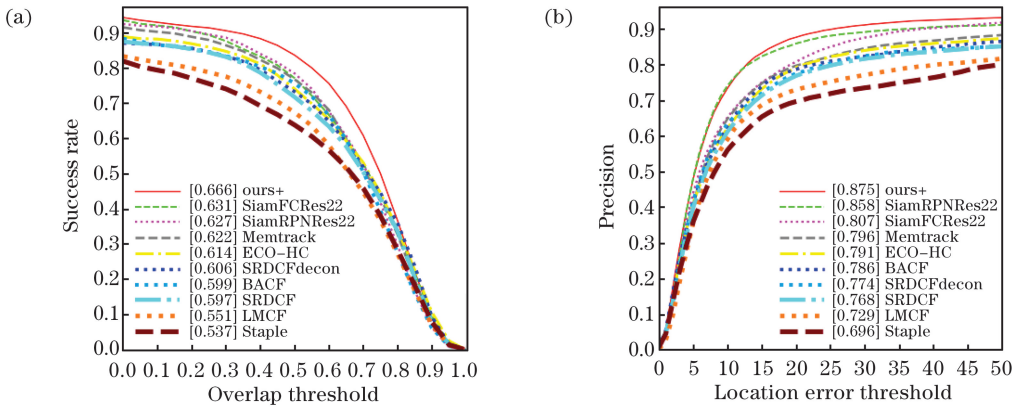


图7 在 OTB100 数据集上不同算法在目标出现快速运动时的对比结果。(a)成功率;(b)精确率

Fig. 7 Comparison results of different algorithms when target moves rapidly on OTB100 dataset. (a) Success rate; (b) precision

表2 在 VOT2016 数据集上不同算法的实验结果

Table 2 Experimental results of different algorithms on VOT2016 dataset

Network	EAO \uparrow	Robustness \downarrow	Accuracy \uparrow
SRDCF	0.338	0.24	0.51
C-COT	0.331	0.24	0.51
ECO-HC	0.322	0.30	0.54
Siam-FC	0.235	0.46	0.53
SiamFCRes22	0.303	0.38	0.54
SiamRPNRes22	0.376	0.24	0.58
TADT	0.299	0.31	0.55
C-RPN	0.363	0.27	0.59
SPM	0.434	0.21	0.62
ours+	0.478	0.17	0.64

综合对比情况可得,所提算法依靠计算目标的形态变化和位置修正实现目标模板在线更新,提升了跟踪算法在目标出现被遮挡、形变、快速变化等情况时的稳定性和准确性,降低了跟踪结果的误差。

图8~10 分别表示所提算法、SRDCF^[10]、BACF^[14]在3个具有挑战性的序列上的结果比较。Coke序列如图8所示,该序列包含目标被遮挡的情况。所提算法在该序列的主观结果上明显优于BACF^[14],较优于SRDCF^[10]。在对应的客观结果中,所提算法在该序列上的平均精度为0.689,SRDCF在该序列上的平均精度为0.513,BACF在该序列上的平均精度为0.575。



图8 Coke序列中各算法的主、客观(平均精度)结果

Fig. 8 Subjective and objective (average precision) results of each algorithm in Coke sequence

在 Dudek 序列上各算法的结果如图9所示,该序列包含目标在平面内旋转的情况。从主观结果上看,所提算法在目标位置和尺度上较优于对比算法,当视频序列中的目标在平面内旋转时,对这种

旋转变化的适应能力更强。在对应的客观结果中,所提算法在该序列上的平均精度为0.882,SRDCF在该序列上的平均精度为0.817,BACF在该序列上的平均精度为0.835。

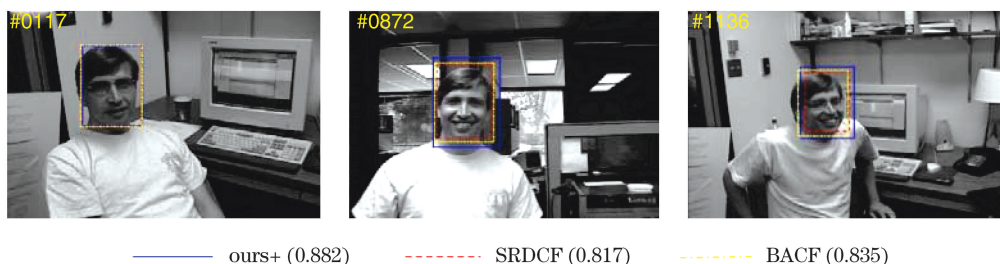


图9 Dudek序列中各算法的主、客观(平均精度)结果

Fig. 9 Subjective and objective (average precision) results of each algorithm in Dudek sequence

各算法在 Doll 序列上的结果如图 10 所示,该序列包含目标尺度变化的情况。从该序列的主观结果上看,所提算法在能够更好地适应目标的尺度变化时,既不会丢失前景信息,也不会包含目标以外的冗余背景信息;在对应的客观结果中,所提算

法在该序列上的平均精度为 0.860,SRDCF 在该序列上的平均精度为 0.837,BACF 在该序列上的平均精度为 0.841。所提算法的精度比 SRDCF 提升了约 2.7%,比 BACF 提升了约 2.3%。



图10 Doll序列中各算法的主、客观(平均精度)结果

Fig. 10 Subjective and objective (average precision) results of each algorithm in Doll sequence

4 结 论

基于深度孪生网络的目标跟踪框架,提出一种运用光流映射实现模板在线更新策略的目标跟踪算法。该算法通过计算模板间的残差来表示目标变化信息,在目标模板上富集目标运动变化信息,增强目标模板的适应能力。除此之外,还利用基于奇异值分解的方法计算针对目标运动位置偏移的修正项,提高跟踪算法对目标运动变化的适应能力,增强跟踪算法的稳定性和准确性。分别在 OTB100 数据集和 VOT2016 数据集上进行测试,与现有算法相比,所提算法取得了更好的结果。

参 考 文 献

- [1] Chen Y F, Wu Y, Zhang W. Survey of target tracking algorithm based on siamese network structure [J]. Computer Engineering and Applications, 2020, 56(6): 10-18.
陈云芳, 吴懿, 张伟. 基于孪生网络结构的目标跟踪算法综述[J]. 计算机工程与应用, 2020, 56(6):

- 10-18.
[2] Yilmaz A, Javed O, Shah M. Object tracking: a survey[J]. ACM Computing Surveys, 2006, 38(4): 13-58.
[3] Danelljan M, Robinson A, Shahbaz Khan F, et al. Beyond correlation filters: learning continuous convolution operators for visual tracking[M]//Leibe B, Matas J, Sbebe N, et al. Computer vision—ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9909: 472-488.
[4] 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), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 6931-6939.
[5] Meng L, Li C X. Kernel correlation filtering algorithm based on a dual-feature model[J]. Journal of Image and Graphics, 2019(12): 2183-2199.
孟球, 李诚新. 双特征模型核相关滤波目标跟踪算法[J]. 中国图象图形学报, 2019(12): 2183-2199.
[6] 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), December 7-13, 2015, Santiago, Chile. New York: IEEE, 2015: 4310-4318.
- [7] Wang D W, Xu C X, Liu Y. Kernelized correlation filter for target tracking with multi-feature fusion[J]. Computer Engineering and Design, 2019, 40(12): 3463-3468.
王殿伟, 许春香, 刘颖. 基于多特征融合的核相关滤波目标跟踪算法[J]. 计算机工程与设计, 2019, 40(12): 3463-3468.
- [8] Danelljan M, Häger G, Khan F S, et al. Adaptive decontamination of the training set: a unified formulation for discriminative visual tracking [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 1430-1438.
- [9] Galoogahi H K, Fagg A, Lucey S. Learning background-aware correlation filters for visual tracking[C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 1144-1152.
- [10] Wang K Y, Chen Z G, Fu Y. Target anti-occlusion algorithm based on correlation filter [J]. Laser & Optoelectronics Progress, 2019, 56(3): 030401.
王凯宇, 陈志国, 傅毅. 基于相关滤波器的目标抗遮挡算法[J]. 激光与光电子学进展, 2019, 56(3): 030401.
- [11] Yang Y G, Shang Z H. Object Tracking Algorithm Based on Correlation Filtering and Convolution Residuals Learning [J]. Laser & Optoelectronics Progress, 2020, 57(12): 121012.
杨亚光, 尚振宏. 相关滤波融合卷积残差学习的目标跟踪算法[J]. 激光与光电子学进展, 2020, 57(12): 121012.
- [12] Cheng Y, Li J Z, Zhu L N, et al. Correlation filter tracking algorithm based on model and scale updating [J]. Laser & Optoelectronics Progress, 2018, 55(12): 121015.
成悦, 李建增, 褚丽娜, 等. 基于模型与尺度更新的相关滤波跟踪算法[J]. 激光与光电子学进展, 2018, 55(12): 121015.
- [13] Yu Y Y, Shi Z L, Liu Y P. Foreground-aware based spatiotemporal correlation filter tracking algorithm [J]. Laser & Optoelectronics Progress, 2019, 56(22): 221503.
虞跃洋, 史泽林, 刘云鹏. 基于前景感知的时空相关滤波跟踪算法[J]. 激光与光电子学进展, 2019, 56(22): 221503.
- [14] Li J P, Shang Z H, Liu H. Visual object tracking algorithm based on correlation filters with hierarchical convolutional features[J]. Computer Science, 2019, 46(7): 252-257.
李健鹏, 尚振宏, 刘辉. 融合多层卷积特征的相关滤波运动目标跟踪算法[J]. 计算机科学, 2019, 46(7): 252-257.
- [15] 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), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 4800-4808.
- [16] Danelljan M, Bhat G, Khan F S, et al. ATOM: accurate tracking by overlap maximization[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE, 2019: 4655-4664.
- [17] Bertinetto L, Valmadre J, Henriques J F, et al. Fully-convolutional siamese networks for object tracking[M]//Hua G, Jégou H. Computer vision-ECCV 2016 workshops. Lecture notes in computer science. Cham: Springer, 2016, 9914: 850-865.
- [18] Ren J M, Gong N S, Han Z Y. Improved target tracking algorithm based on siamese convolution neural network [J]. Journal of Chinese Computer Systems, 2019, 40(12): 2686-2690.
任珈民, 宫宁生, 韩镇阳. 一种改进的基于孪生卷积神经网络的目标跟踪算法[J]. 小型微型计算机系统, 2019, 40(12): 2686-2690.
- [19] Li B, Wu W, Wang Q, et al. SiamRPN++: evolution of Siamese visual tracking with very deep networks [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE, 2019: 4277-4286.
- [20] Zhou D Y, Duan X P. Target tracking method based on siamese network and attention mechanism [J]. Information & Communications, 2019(12): 61-63.
周迪雅, 段喜萍. 基于孪生网络与注意力机制的目标跟踪方法[J]. 信息通信, 2019(12): 61-63.
- [21] Ilg E, Mayer N, Saikia T, et al. FlowNet 2.0: evolution of optical flow estimation with deep networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26,

- 2017, Honolulu, HI, USA. New York: IEEE, 2017: 1647-1655.
- [22] Yang T, Chan A B. Learning dynamic memory networks for object tracking[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision—ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11213: 153-169.
- [23] Paszke A, Gross S, Massa F, et al. PyTorch: an imperative style, high-performance deep learning library [EB/OL]. (2019-12-03) [2020-03-05]. <https://arxiv.org/abs/1912.01703>.
- [24] 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), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 1401-1409.
- [25] Zhang Z P, Peng H W. Deeper and wider Siamese networks for real-time visual tracking [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE, 2019: 4586-4595.
- [26] Li X, Ma C, Wu B Y, et al. Target-aware deep tracking [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE, 2019: 1369-1378.
- [27] Fan H, Ling H B. Siamese cascaded region proposal networks for real-time visual tracking [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE, 2019: 7944-7953.
- [28] Wang G T, Luo C, Xiong Z W, et al. SPM-tracker: series-parallel matching for real-time visual object tracking [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE, 2019: 3638-3647.