

文章编号: 2095-2163(2020)03-0044-07

中图分类号: TP391.41

文献标志码: A

基于卡尔曼滤波的 SiamRPN 目标跟踪方法

张子龙, 王永雄

(上海理工大学 光电信息与计算机工程学院, 上海 200093)

摘要: 基于深度学习的视觉跟踪方法在多个基准数据库上取得了很好的跟踪性能,特别是基于 Siamese 框架的目标跟踪方法取得了突破性的进展。为了提高跟踪效果,有效解决跟踪过程中干扰和遮挡问题,本文提出了一种基于卡尔曼滤波的 SiamRPN(Siamese+RPN)目标跟踪方法。首先,利用训练好的 SiamRPN 跟踪算法和卡尔曼滤波跟踪模型分别对目标物体进行跟踪,得到 2 种跟踪算法跟踪结果的置信度,然后,基于置信度加权融合模型得到最后的跟踪框。卡尔曼滤波器可预测目标在一定遮挡干扰等情况下的位置,SiamRPN 算法利用区域候选网络 RPN 将每一帧的跟踪转换为一次局部检测任务,快速准确地得到跟踪框的位置和尺度,提出的算法避免了使用常规的低效费时的多尺度自适应方法,融合了 2 种优秀跟踪算法的优点,不仅跟踪速度较快,而且抗干扰和遮挡能力明显提高。在经典数据库上的实验验证了提出的算法明显提高了目标运动较快、干扰较强和有遮挡情况下的跟踪效果,在速度没有明显下降的前提下,成功率和精度等多个性能指标均有较大的提升。

关键词: 目标跟踪;卡尔曼滤波;孪生网络;加权融合

SiamRP network for object tracking based on Kalman filter

ZHANG Zilong, WANG Yongxiong

(School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China)

[Abstract] The visual learning method based on deep learning has achieved good tracking performance on multiple benchmark databases. Especially, the object tracking method based on Siamese framework is a breakthrough. In order to improve the tracking effect and solve effectively the interference and occlusion problems in tracking, a SiamRPN (Siamese+RPN) object tracking method based on Kalman filtering is proposed. Firstly, the trained SiamRPN tracking algorithm and the Kalman filter tracking model are used to track the object respectively, and the confidence of the results of the two tracking algorithms is obtained. Then, the final tracking frame is obtained based on the confidence weighted fusion model. The Kalman filter can predict the position of the object under certain occlusion interference. The regional candidate network RPN in SiamRPN algorithm is used to convert the tracking of each frame into a local detection task, and obtain the position and scale of the tracking frame both quickly and accurately. The conventional inefficient and time-consuming multi-scale test and online fine-tuning is abandoned. The new method includes the advantages of two excellent tracking algorithms. Thus, not only the tracking speed is fast, but also the anti-interference and occlusion capabilities are significantly improved. Experimental results on the classical database verify that the proposed algorithm significantly improves the tracking effect in the condition of fast object motion, strong interference and occlusion. The performances such as success rate and accuracy are greatly achieved without decreasing the tracking speed obviously.

[Key words] object tracking; Kalman filter; SiamRPN network; weighted fusion;

0 引言

视觉目标跟踪是最近几年计算机视觉领域的主要研究方向之一,也是计算机视觉分析、自动驾驶和姿态估计等领域的一个基本问题。跟踪的主要难点是在具有遮挡、干扰、变形、背景杂波等变化的复杂场景中准确地检测和定位目标。目前主流的跟踪方法分为 2 类。一类基于相关滤波的跟踪算法^[1-3],该方法因为速度快、效果好吸引了众多研究者的目光。相关滤波器通过将输入特征回归为目标高斯分布来训练滤波器,Bolme 等人^[2]提出用相关滤波器进行目标跟踪,并优化傅里叶域中误差平方和的输出,使跟踪速度达到了每秒 669 帧。Henriques 等

人^[4]通过引入循环移位矩阵和和技巧(Exploiting the Circulant Structure of Tracking-by-detection with Kernels, CSK)可以在基本不损失计算速度的基础上极大地扩充训练样本,但该算法的目标框大小是固定的,对发生尺度变化的目标不具有鲁棒性。Henriques 等人^[5]在后续工作中提出了核相关滤波器(High-Speed Tracking with Kernelized Correlation Filters, KCF),该滤波器在 CSK 的基础上提出用快速梯度直方图特征(Fast Histogram of Oriented Gradiend, FHOG)提高跟踪精度。针对目标运动中造成的尺度变化问题,Danelljian 等人^[6]提出了平移滤波加尺度自适应的方法(Discriminative Scale

作者简介: 张子龙(1993-),男,硕士研究生,主要研究方向:目标跟踪;王永雄(1970-),男,博士,教授,主要研究方向:智能机器人及视觉。

收稿日期: 2019-12-25

Space Tracking, DSST), 并引入多特征融合机制, 显著地提高了跟踪精度。另一类方法是基于强大的深度学习的方法, 随着深度学习的发展, 其强大的学习能力和优秀的特征表达能力在视觉跟踪领域展现了巨大的潜力。学者们开始用鲁棒性更强的卷积神经网络来提取目标特征。Ma 等人^[7]提出了 HCF (Hierarchical Convolutional Features for visual tracking) 算法, 将传统的 HOG 特征替换为深度卷积特征, 由于低层特征有较高的分辨率能够对目标进行精准的定位, 高层特征包含更多的语义信息, 故而提出用分层卷积特征对目标进行学习。Danelljan 等人^[8]提出了连续卷积算子 (Continuous Convolution Operators for visual Tracking, CCOT), 该算法在特征融合的基础上, 通过在连续空间域的学习, 创建时域连续的相关滤波器, 并将不同分辨率的特征图输入滤波器, 使传统特征和深度特征相结合, 显著地提高了跟踪的实时性和鲁棒性。Danelljan 等人^[9]针对 CCOT 处理高维特征需要处理多个滤波器的问题, 为了提高时间和空间效率, 提出了高效的卷积算子 ECO (Efficient Convolution Operators for tracking), 利用因式分解构造一组更小的滤波器, 降低了模型的复杂度, 防止过高的维度而导致的低效率和过拟合问题, 同时利用高斯混合模型表示不同的目标外观, 降低模型的更新频率, 防止过拟合。

近年来, 采用基于 Siamese (孪生) 网络框架进行目标跟踪的孪生网络以其优良的性能在视觉跟踪领域受到了广泛的关注。经典的 SiamFC 方法首先学习构建先验的孪生相似性函数^[10], 然后利用这个相似性函数比较目标模板和搜索区域的相似度, 最后得到搜索区域的得分图 (score map)。该跟踪算法在速度方面取得不错的效果。SiamRPN^[11]是在 Siamese 网络的基础上引入了一个区域候选网络 (RPN), 将每一帧的跟踪转换为一次局部检测任务, SiamRPN 网络强大的检测功能使得跟踪方法获得了较高的精度和速度。然而, 当目标运动速度较快或者目标周围有干扰物的时候, 跟丢目标的概率依然很高。为此, Morimitsu^[12]提出用多种上下文特征来扩展 SiamFC 跟踪器, 从非常深的网络中提取多个上下文和语义级别的特征, 获取跟踪目标的更加完备特性描述, 从而提高跟踪效果。该方法能够有效地从不同层中提取出互补特征, 并在融合时显著提高了模型的抗干扰性能。同样为了提高模型的抗干扰能力和泛化能力, Wang 等人^[13]提出了 RASNet (Residual Attentional Siamese Network for

high performance online visual tracking), 该算法在 Siamese 网络的基础上, 通过引入 3 种注意力机制 (Residual Attention、General Attention 和 Channel Attention), 精确描述目标的外观轮廓, 优先选择更加强有力的特征通道, 再对其进行加权融合, 并把这项注意力机制作为一个 layer 嵌入 Siamese 网络中, 提升了模型的判别能力和抗干扰性能。但是, 在长时跟踪中, Siamese 类跟踪器不能很好地应对目标全遮挡、目标出画面等挑战。

为了提高模型抗干扰性和解决短时全遮挡问题, 本文提出了一种基于卡尔曼滤波的 SiamRPN 目标跟踪方法, 当视频输入时, SiamRPN 和卡尔曼滤波模块同时进行跟踪, SiamRPN 模型分支首先利用卷积神经网络对模板帧和检测帧进行特征提取, 然后经过 RPN 网络回归得到跟踪目标的位置坐标和尺度。模型的另一分支利用卡尔曼滤波对检测帧的目标进行预测, 最后通过置信度加权融合 2 个子模型的跟踪结果, 确定最优跟踪框。SiamRPN 中的区域候选网络 RPN 可以快速准确地计算出跟踪框的位置和尺度, 避免了使用常规的低效费时的多尺度自适应方法。当目标受到严重干扰或遮挡时, 卡尔曼预测模块依靠优秀的预测位置能力, 相对准确地跟踪到目标的位置, 这是 Siamese 类跟踪器不具有的性能。因此, 本文的方法不仅跟踪速度快, 而且在抗干扰和解决遮挡问题上表现出明显的优势。

1 基于卡尔曼滤波的孪生网络

1.1 基于 Siamese 的跟踪网络

近年来, 基于 Siamese 网络的跟踪器因其良好的跟踪精度和效率而受到广泛关注^[14-16]。这些跟踪器将目标跟踪表示为通过学习目标模板和搜索区域的特征表示之间的相互关联来学习一般的相似性映射, 并从具有孪生网络结构的深度模型中学习跟踪相似度, 一个分支用于学习目标特征表示, 另一个分支用于学习搜索区域的特征表示。为了从 2 个分支的相互关系中生成相似映射, 研究中训练了一个全卷积神经网络, 将 2 个网络分支连接起来, 一个用于对象模板, 另一个用于搜索区域。目标补丁 (实例) 通常是在第一帧的序列, 可以被视为一个范例 z 。研究的目的是在搜索区域 x 找到与范例 z 最相似的目标对象。

研究中采用了一种关于输入图像 x 的全卷积 Siamese 架构。如果能够对输入图像进行平移转换, 则这个函数是完全卷积。引入 L_τ 来表示平移算子 ($L_\tau x$) [u] = $x[u - \tau]$, 对于任意平移 τ 与整数步长

k , 定义如下:

$$h(L_{kT}x) = L_T h(x), \quad (1)$$

全卷积网络与候选图像大小无关, 能够计算所有转换子窗口 x 与范例 z 的相似性。通过使用卷积嵌入函数 φ 得到范例图像 z 和搜索图像 x 的相似性, 对应公式可表示为:

$$f(z, x) = \varphi(z) * \varphi(x) + b_1. \quad (2)$$

其中, b_1 表示在每个位置的取值 ($b \in R$); $*$ 表示互相关操作符; f 是用来比较范例图像 z 和搜索图像 x 的相似性函数。

1.2 SiamRPN 网络框架

1.2.1 SiamRPN 介绍

SiamRPN 第一次是在 Faster RCNN 中使用^[17], 专门用来提取候选框, 在 RCNN 和 Fast RCNN 等物体检测架构中, 提取候选框的方法通常是 Selective Search, 是传统的方法之一, 而且比较耗时, 使用 CPU 需要花费 2 s 计算一张图。为提高计算速度, 提出了利用 RPN 提取候选框, 一方面 RPN 耗时少, 另一方面 RPN 可以很容易结合到 Fast RCNN 中, 合成为一个整体。

在视觉跟踪中, 区域候选网络 RPN 的引入使得网络在跟踪的时候不需要多尺度测试, 通过网络回归可以直接得到更加准确的目标位置, 不需要插值等附加计算。在网络训练中, 研究引入了大规模的视频数据集 Youtube-BB 进行训练, 相比较于原来的 VID 数据集, Youtube-BB 在视频数量上有大约 50 倍的提升, 这保证了网络能够得到更为充分的训练。

本文采用的 SiamRPN 网络框架如图 1 所示。由图 1 可知, 左边是用于特征提取的孪生网络, 上下支路的网络结构和参数完全相同, 上面是输入第一帧的 bounding box, 靠此信息检测候选区域中的目标, 即模板帧。下面是检测帧, 显然, 检测帧的搜索区域比模板帧的区域大。中间是区域提取网络, 包含分类分支和回归分支, 这里采用两两相关的卷积运算得到 2 个分支的输出。在分类分支, 输出的特征图包含 $2k$ 个通道, 表示 k 个锚点的前景与背景的概率, 在回归分支上, 输出的特征图包含 $4k$ 个通道, 分别表示 k 个锚点坐标偏移的预测值。右侧是输出结果。图 1 中, $*$ 表示相关运算。

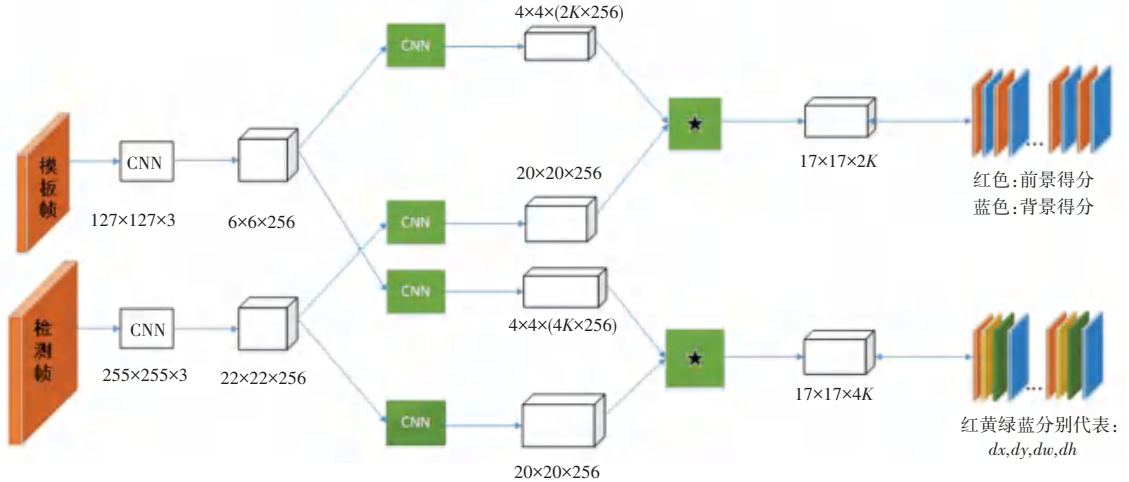


图 1 SiamRPN 网络架构

Fig. 1 SiamRPN network architecture

分类分支和回归分支分别对模板帧和检测帧的特征进行卷积运算, 由此推得:

$$A_{w \times h \times 2k}^{cls} = [\varphi(x)]_{cls} * [\varphi(z)]_{cls}, \quad (3)$$

$$A_{w \times h \times 4k}^{reg} = [\varphi(x)]_{reg} * [\varphi(z)]_{reg}, \quad (4)$$

其中, $A_{w \times h \times 2k}^{cls}$ 包含 $2k$ 个通道数量, 每个点表示正负激励, 通过计算 softmax 来求得分类损失; $A_{w \times h \times 4k}^{reg}$ 包含 $4k$ 个通道, 每个点表示锚点框和真实框之间的 dx , dy , dw , dh , 并通过 $smooth_{L_1}$ 来计算损失, 损失函数定义如下:

$$L_{reg} = \sum_{i=0}^3 smooth_{L_1}(\delta[i], \sigma), \quad (5)$$

$$\delta[0] = \frac{T_x - A_x}{A_w}, \quad \delta[1] = \frac{T_y - A_y}{A_h}, \quad \delta[2] = \ln \frac{T_w}{A_w},$$

$$\delta[3] = \ln \frac{T_h}{A_h}, \quad (6)$$

其中, A_x, A_y, A_w 和 A_h 是锚点框的中心点坐标和长宽, T_x, T_y, T_w 和 T_h 是真实框的中心点坐标和长宽, 因为不同图片的尺寸存在差异, 需要先做正则化处理。

研究求得, 公式(5)中的 $smooth_{L_1}$ 损失函数定义如下:

$$smooth_{L_1}(x, \sigma) = \begin{cases} 0.5\sigma^2 x^2, & |x| < \frac{1}{\sigma^2}; \\ |x| - \frac{1}{2\sigma^2}, & |x| \geq \frac{1}{\sigma^2}. \end{cases} \quad (7)$$

最终的损失函数为:

$$loss = L_{cls} + \lambda L_{reg}. \quad (8)$$

其中, λ 是平衡分类损失和回归损失的超参数, L_{cls} 是交叉熵损失。

1.2.2 RPN 的训练

研究采用端到端的方式训练 RPN。训练过程中, 首先从 Youtube-BB 和 ILSVRC 选择随机间隔的样本, 在同一视频中随机提取模板帧和检测帧进行预训练, 然后, 用随机梯度下降法 (SGD) 对孪生 RPN 进行端到端的训练。在训练数据增强方面, 采用了一些常规的数据增强的方法, 例如仿射变换。此外, 因为同一目标在相邻两帧之间一般变化不大, 所以需要锚的数量比较少。同时, 也采用了尺度为 5 种的不同比例的锚 [0.33, 0.5, 1, 2, 3]。在训练时, 识别锚框属于正样本还是负样本是一个重要依据, 研究中运用指标 IOU (锚点框和真实值框的交并比) 判断正负样本, 这里规定 $IOU > 0.6$ 定义为正样本, 当 $IOU < 0.3$ 时定义为负样本。

1.3 卡尔曼滤波模型

1.3.1 卡尔曼滤波

卡尔曼滤波是一种最优化的回归算法^[18-19], 通过最小化均方误差实现线性滤波, 并利用观测方程和状态方程, 迭代地对跟踪目标的状态进行预测和更新。在目标跟踪中, 当目标受到干扰, 遮挡等问题, 会导致目标跟丢, 此时卡尔曼滤波器可以解决这一问题。卡尔曼滤波算法可以预测目标的状态, 并对存在遮挡或干扰的目标预测位置进行标记, 以便在下一帧中仍然可以跟踪到被遮挡或干扰的目标。状态转移方程和观测方程是卡尔曼滤波的基础, 其公式如下:

$$\hat{x}_t = F_t \hat{x}_{t-1} + w_{t-1}, \quad (9)$$

$$y_t = H[x_t] + v_t, \quad (10)$$

其中, \hat{x}_t 和 \hat{x}_{t-1} 分别表示 t 时刻和 $t-1$ 时刻系统的状态; y_t 表示 t 时刻系统的观测状态; F_t 表示状态转移矩阵, 对应目标的运动方式; H 是观测矩阵, 表示目标真实状态值和观测值的对应关系; w_{t-1} 表示 $t-1$ 时刻的系统噪声, 满足正态分布, 引入协方差矩阵是 Q 。卡尔曼滤波器的预测和更新公式如下:

$$\hat{x}_t = F_t \hat{x}_{t-1} + w_{t-1}, \quad (11)$$

$$\Sigma_t^- = F_t \Sigma_{t-1} F_t^T + Q, \quad (12)$$

$$K_t = \Sigma_t^- H^T (H \Sigma_t^- H^T + R)^{-1}, \quad (13)$$

$$\hat{x}_t = \hat{x}_t^- + K_t (y_t - H \hat{x}_t^-), \quad (14)$$

$$\Sigma_t = (I - K_t H) \Sigma_t^-, \quad (15)$$

分析可知, 公式 (11) 为状态预测方程, 公式 (12) 表示误差协方差的预测公式, Σ_{t-1} 表示 $t-1$ 时刻误差协方差矩阵的更新值, Σ_t^- 表示误差协方差的预测值, 协方差矩阵 Q 表示预测模型本身的噪声。公式 (13) 表示卡尔曼滤波增益阵, K_t 代表 t 时刻的卡尔曼增益值, 滤波增益阵首先权衡预测状态协方差矩阵 Σ 和观测值矩阵 R 的大小, 并以此判断是更倾向于预测模型还是观测模型。如果认为预测模型更可靠, 那么这个残差的权重就会小一点。反之亦然。不仅如此, 滤波增益阵还负责把残差的表现形式从观测域转换到了状态域。公式 (14) 表示目标状态优化方式, \hat{x}_t^- 是根据上一状态推测而来的, 那么其与最优估计值之间的差距就是 $K_t (y_t - H \hat{x}_t^-)$ 。

$y_t - H \hat{x}_t^-$ 表示实际观察值与预估的观测值之间的残差。 $K_t (y_t - H \hat{x}_t^-)$ 就可以用来对估计值进行修正。公式 (15) 表示对误差协方差的优化和更新方式。

1.3.2 基于卡尔曼滤波的位置估算

当要预测目标的状态时, 首先要建立起目标的初始状态表示, 本文中, 目标的状态可以表示成 $x = (p, v)$, 其中 p 为目标的位置, v 为目标的速度, 目标的状态用向量表示成如下形式:

$$x = (p_x, p_y, v_x, v_y, \theta)^T, \quad (16)$$

目标的观测方程表示为:

$$y(t) = (x, y, \theta)^T \quad (17)$$

其中, θ 是目标转动的角度, 跟踪的基本过程如下:

(1) 初始化阶段, 在参考图像定义目标观测 z_{pos} 和目标的初始状态 $x(0)$ 。采用卡尔曼滤波的过程模型预测目标的可能出现的区域:

$$\hat{x}_t = F_t \hat{x}_{t-1} + w_{t-1}, \quad (18)$$

其中,

$$F_t = \begin{bmatrix} \hat{e}^1 & 0 & \Delta t & 0 & 0 & \hat{y} \\ \hat{e}^0 & 1 & 0 & \Delta t & 0 & \hat{u} \\ \hat{e}^0 & 0 & 1 & 0 & 0 & \hat{u} \\ \hat{e}^0 & 0 & 0 & 1 & 0 & \hat{u} \\ \hat{e}^0 & 0 & 0 & 0 & 1 & \hat{y} \end{bmatrix}, \quad Q = \begin{bmatrix} \hat{e}^1 & 0 & 0 & 0 & 0 & \hat{y} \\ \hat{e}^0 & 1 & 0 & 0 & 0 & \hat{u} \\ \hat{e}^0 & 0 & 1 & 0 & 0 & \hat{u} \\ \hat{e}^0 & 0 & 0 & 1 & 0 & \hat{u} \\ \hat{e}^0 & 0 & 0 & 0 & 1 & \hat{y} \end{bmatrix}. \quad (19)$$

分析后可知,式(18)根据 \hat{x}_{t-1} 预测目标在下一时刻的位置 \hat{x}_t 。

(2)初始化卡尔曼滤波器之后,可以根据后续的目标跟踪结果在一个迭代过程中对目标的状态进行预测和最优化估计,该迭代过程即对应目标跟踪过程。

1.4 卡尔曼滤波与 SiamRPN 的融合

针对目标被干扰和遮挡的问题,本文提出卡尔曼滤波与 SiamRPN 融合的算法,算法的流程图如图 2 所示。

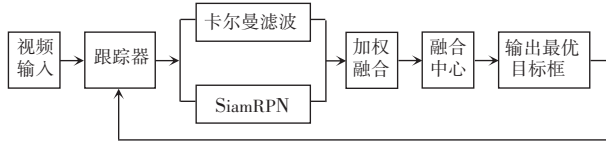


图 2 本文算法总流程图

Fig. 2 The general flow of the algorithm

在算法流程中,研究的目标跟踪总体过程如下:首先,初始化系统,将初始框的中心和位置赋值给 SiamRPN 和卡尔曼滤波器,同时修改卡尔曼滤波器的系统状态的初始值和初始预测。然后, SiamRPN 通过特征提取、分类回归和分支回归,最终得到一个置信度值,卡尔曼滤波器通过对下一帧进行预测,得到置信度值。最后,通过对 2 种跟踪算法进行自动加权融合得到最后的跟踪框。

对于自适应加权融合过程, SiamRPN 和卡尔曼滤波器得到的跟踪结果可以看作相互独立的变量,对于自适应加权融合计算过程可描述如下:假设由 SiamRPN 跟踪算法得到的跟踪框置信度为 k_1 , 卡尔曼滤波器得到的跟踪框置信度为 k_2 。 SiamRPN 和卡尔曼滤波器的加权因子分别为 w_1 和 w_2 。因为这 2 个跟踪算法相互独立,而且加权融合思想的核心是兼顾每个跟踪算法的优势,所以加权因子的取值应满足以下关系:

$$w_1 = \frac{k_1}{k_1 + k_2}, \quad w_2 = \frac{k_2}{k_1 + k_2}, \quad (20)$$

设 SiamRPN 和卡尔曼滤波器得到跟踪框的值分别为 x_1 和 x_2 , 则加权融合后的最优跟踪框的值为:

$$x = w_1 x_1 + w_2 x_2. \quad (21)$$

2 实验结果分析

实验中,拟在公认的具有挑战性的跟踪数据集上进行: OTB2015^[20], 其中 OTB2015 数据集包含 100 个视频。涉及了目标跟踪常见的干扰问题,例如背景相似干扰、快速运动、光照变化、运动模糊等

问题,该数据集能够全面反映算法的跟踪性能,所有的跟踪结果都使用报告的结果,以确保同等条件下进行比较。

2.1 实验环境及参数设置

本文实验环节使用基于 Linux 系统的硬件平台,工作站配置为 2.6 GHz, 32 G 内存,显卡为 GTX1080TiGPU。软件工具为 Pycharm2018。实验使用修改后的 AlexNet 网络,由 ImageNet 数据集预训练的前三层卷积层的参数是不变的,只微调后面两个卷积层的参数。这些参数是通过损失函数进行优化,通过在 VID 和 Youtube-BB 中选择帧间隔小于 100 的图像对进行进一步的裁剪工作。如果目标边界框的大小表示为 (w, h) , 则裁剪模板大小为 $A * A$ 。相应的定义方式如下:

$$(w + p) \times (h + p) = A^2. \quad (22)$$

其中, p 的大小为 $\frac{w+h}{2}$, 把尺寸调整为

127 × 127, 研究中用同样的方法把检测帧裁剪为模板帧的 2 倍大小,并把其尺寸调整为 256 × 256 的大小。

2.2 定性分析

选取 OTB2015 中 3 组典型的序列 Soccer、Human6 与 Bolt2, 其中, Soccer 序列主要具有平面内旋转、快速运动、遮挡等干扰, Human6 和 Bolt2 序列主要具有遮挡与运动模糊等干扰。采用本文算法、SiamRPN 和 SianFC_3s 在 3 组序列上的跟踪结果如图 3 所示。在 OTB2015 测试序列的 Human6 序列中,当目标在序列的前 350 帧之前,即目标没有经过红绿灯杆时,3 种算法均能有效跟踪到行人,并且差异不大,当目标遇到遮挡时,即在 350 ~ 370 帧之间,由于目标部分被遮挡, SianFC_3s 算法很快丢失目标, SiamRPN 算法虽未完全丢失目标,但跟踪效果也已大幅度下降。通过实验验证了本文算法在严重遮挡的情况下,依然能准确跟踪到目标。在图 3(b) 中,采用 OTB2015 测试序列中的 Soccer 序列,该视频集具有运动模糊、背景干扰以及遮挡等问题,在前面的序列中,目标并未受到大的干扰和遮挡问题,此时实验采用的 3 种算法均能很好地跟踪到目标,在第 38 帧之后,由于目标快速运动,期间伴有部分遮挡,导致 SiamRPN 和 SianFC_3s 两种算法很快跟丢目标。在图 3(c) 中,由于目标运动员在快速运动中周围含有相似干扰运动员,当第 81 帧时,由于相似干扰运动员与目标太近,导致 SiamRPN 和 SianFC_3s 算法跟丢。而本文所采用的抗干扰和遮挡的算法很

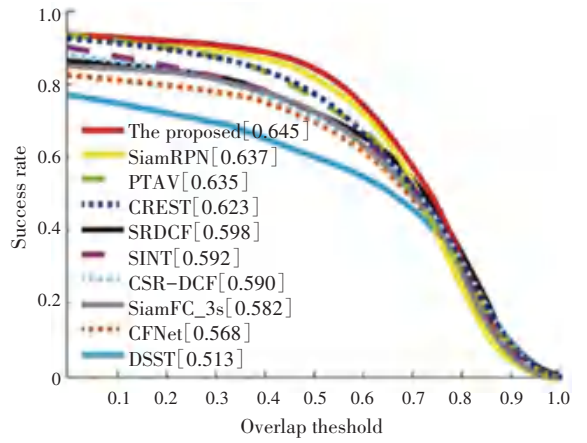
好地解决了这个问题, 当目标遇到遮挡时, 由于 RPN 模块不能应对该问题, 本文提出的卡尔曼滤波可以很好地预测目标的位置。综上分析可知在图 3 中, (a) 是遮挡问题, (b) 和 (c) 是严重的干扰和遮挡问题, 从实验结果可以看出, 相比于 SiamRPN 等算法, 本文所提出的基于卡尔曼滤波的 SiamRPN 目标跟踪方法在面对背景干扰问题上具有明显的优势。



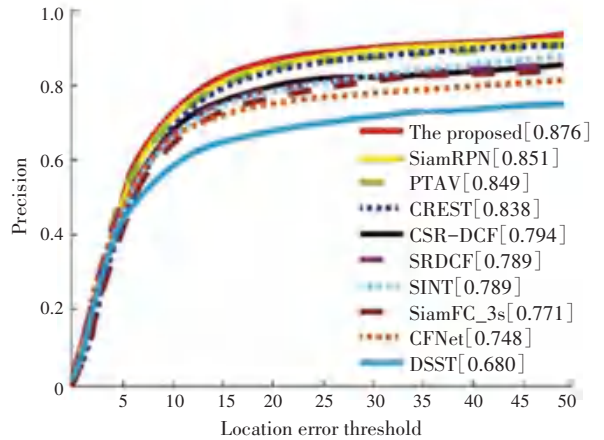
图 3 不同算法在 Human6, Soccer 与 Bolt2 序列上的跟踪结果图
Fig. 3 Tracking results of the different algorithms on Human6, Soccer and Bolt2 sequences

2.3 定量分析

OTB2015 包含从常用跟踪序列中收集的 100 个序列。评估基于 2 个指标: 精度图和成功率图。精度图显示目标跟踪结果与目标距离在 20 像素以内的帧数百分比。成功率图显示了阈值在 0~1 之间变化时成功帧的比例, 其中成功帧表示其重叠大于给定阈值。曲线下面积 (AUC) 的成功图用于排序需要对比的跟踪算法。在本实验中, 是将本文提出的方法与数个代表性的跟踪器进行了比较, 包括 SiamRPN、PTAV、CREST、SRDCF、SINT、CSR-DCF、SianFC_3s、CFNet、DSST。不同跟踪算法的成功率图与精度图见图 4。从图 4 中可以看出, 相比于传统算法 SianFC_3s, 加入了区域候选网络 RPN 后的 SiamRPN 其成功率和精度分别提高了 5.5% 和 8%。本文算法在 SiamRPN 上引入了卡尔曼滤波模块, 成功率和精度分别提高了 0.8% 和 2.1%, 通过定量分析表明, 本文提出的算法具有更好的跟踪性能。



(a) 成功率图
(a) Success rate graph



(b) 精度图
(b) Accuracy map

图 4 不同跟踪算法的成功率图与精度图

Fig. 4 Success rate graph and accuracy map of different tracking algorithms

3 结束语

本文提出了一种基于卡尔曼滤波的 SiamRPN 目标跟踪方法, 在 SiamRPN 算法的基础上, 融合了卡尔曼滤波, 该算法不仅跟踪速度快, 而且有效解决了跟踪过程中的干扰和遮挡问题, 实现了跟踪目标在运动较快、干扰较强和有遮挡情况下的高性能跟踪。其中, SiamRPN 算法中的区域候选网络 RPN 可以快速准确地计算出跟踪框的位置和尺度, 避免了使用低效率时的多尺度测试来确定跟踪框的尺度。通过对比实验可以看出本文提出的基于卡尔曼滤波的 SiamRPN 目标跟踪方法, 在面对目标遮挡、背景干扰等传统目标跟踪问题时, 具有更好的跟踪性能。虽然本文算法与对比算法相比能够取得较好的效果, 但是增加卡尔曼滤波模块的同时也降低了跟踪速度。本文算法针对短期的干扰具有优良的性能, 但在面对长期的遮挡和干扰的跟踪效果仍不够理

想,还需要进一步的研究工作完善该算法。

参考文献

- [1] BIBI A, MUELLER M, GHANEM B. Target response adaptation for correlation filter tracking[M]// LEIBE B, MATAS J, SEBE N, et al. Computer Vision - ECCV 2016. Lecture Notes in Computer Science. Cham:Springer, 2010,9910:2544.
- [2] BOLME D S, BEVERIDGE J R, DRAPER B A, et al. Visual object tracking using adaptive correlation filters [C]//IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2010:2544.
- [3] 李麒麟, 李磊民, 黄玉清. 尺度自适应的核相关滤波跟踪器[J]. 计算机应用, 2016, 36(12):3385.
- [4] HENRIQUES J F, CASEIRO R, MARTINS P, et al. Exploiting the circulant structure of tracking-by-detection with kernels[C]// European Conference on Computer Vision. Berlin/ Heidelberg: Springer, 2012: 702.
- [5] 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, 2014, 37(3): 583.
- [6] DANELLJAN M, HÄGER G, KHAN F S, et al. Discriminative scale space tracking [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(8): 1561.
- [7] MA Chao, HUANG Jiabin, YANG Xiaokang, et al. Hierarchical convolutional features for visual tracking[C]//Proceedings of the IEEE International Conference on Computer Vision. Santiago, Chile:IEEE,2015: 3074.
- [8] DANELLJAN M, ROBINSON A, KHAN F S, et al. Beyond correlation filters; Learning continuous convolution operators for visual tracking[M]// LEIBE B, et al. Proceedings of the 14th European Conference on Computer Vision. Amsterdam, The Netherlands:Springer, 2016: 472.
- [9] DANELLJAN M, BHAT G, KHAN F S, et al. ECO: Efficient convolution operators for tracking [C]// Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA:IEEE, 2017:6931.
- [10] 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.
- [11] LI B, YAN J, WU W, et al. High performance visual tracking with siamese region proposal network [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City:IEEE,2018: 8971.
- [12] MORIMITSU H. Multiple context features in siamese networks for visual object tracking [C]//Proceedings of the European Conference on Computer Vision (ECCV). Munich, Germany: Diamond,2018; 116.
- [13] WANG Q, TENG Z, XING J, et al. Learning attentions: Residual attentional siamese network for high performance online visual tracking [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City:IEEE, 2018: 4854.
- [14] ZHU Z, WANG Q, LI B, et al. Distractor-aware siamese networks for visual object tracking [C]//Proceedings of the European Conference on Computer Vision (ECCV). Munich, Germany:Diamond, 2018; 101.
- [15] HE Anfeng, LUO Chong, TIAN Xinmei, et al. A twofold Siamese network for real-time object tracking [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City:IEEE, 2018; 4834.
- [16] SUN C, WANG D, LU H, et al. Learning spatial-aware regressions for visual tracking [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City:IEEE,2018; 8962.
- [17] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [C]// International Conference on Neural Information Processing Systems. Cambridge:MIT Press, 2015: 91.
- [18] 胡波, 陈恩, 徐建瑜, 等. 基于 Kalman 预测和 Mean-shift 算法的视频目标跟踪[J]. 光电子. 激光, 2009 (11):1157.
- [19] 赵广辉, 卓松, 徐晓龙. 基于卡尔曼滤波的多目标跟踪方法[J]. 计算机科学, 2018, 45(8):253.
- [20] WU Y, LIM J, YANG M H. Object tracking benchmark [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015,37(9):1834.

(上接第43页)

化示意图对锂电池 RUL 预测的每一步都给出说明。最后分别利用 RVM 与 ICS+RVM 模型对 RW12 组锂电池进行预测实验,通过计算出的 *MSE* 与 *RMSE* 误差指标验证了优化后模型的预测优越性。本文创新之处在于通过双参数动态调整改进 CS 算法,提高其寻优能力,从而使得最终预测准确度更高。

参考文献

- [1] 刘大同,周建宝,郭力萌,等. 锂离子电池健康评估和寿命预测综述[J]. 仪器仪表学报,2015,36(1):1.
- [2] 张吉宣. 锂离子电池剩余寿命预测方法研究[D]. 太原:中北大学,2018.
- [3] 林娅,陈则王. 锂离子电池剩余寿命预测研究综述[J]. 电子测量技术,2018,41(4):29.
- [4] 王玉斐. 一种基于 ARI 模型和 SRCKF 的融合型算法的锂电池剩余寿命预测方法[J]. 舰船电子对抗,2019,42(4):117.
- [5] 朱晓栋,陈则王. 基于维纳过程的电池剩余使用寿命预测[J]. 机械制造与自动化,2018,47(4):197.
- [6] 王飞. 改进初值 Π 隐马尔可夫模型预测电池剩余寿命研究[D]. 哈尔滨:哈尔滨理工大学,2018.
- [7] 何成,刘长春,武洋,等. 基于改进神经网络算法的医疗锂电池 PHM 系统设计[J]. 计算机测量与控制,2018,26(12):72.
- [8] 陈雄姿,于劲松,唐菽音,等. 基于贝叶斯 LS-SVR 的锂电池剩余寿命概率性预测[J]. 航空学报,2013,34(9):2219.
- [9] 王腾蛟,郭建胜,慕容政,等. 一种预测锂电池剩余寿命的改进粒子滤波算法[J]. 空军工程大学学报(自然科学版),2018,19(5):47.
- [10] BOLE B, KULKARNI C S, DAIGLE M. Adaptation of an electrochemistry-based li-Ion battery model to account for deterioration observed under randomized use [C]// Annual Conference of the Prognostics and Health Management Society (PHM 2014). Forth worth, TX:[s.n.], 2014.