

基于深度学习的粒子滤波视频目标跟踪算法

李敏, 吴莎

(长安大学信息工程学院, 陕西西安 710000)

摘要:在视频目标跟踪中,由于环境以及目标形变等因素的影响,会导致跟踪算法出现鲁棒性差的问题,针对该问题提出了一种基于预训练卷积神经网络,在粒子滤波框架下将深度特征和手工特征相结合的视觉目标跟踪算法。该算法通过卷积神经网络对数万张通用目标图像进行离线预训练,得到可以对通用目标进行表示的从简单到复杂的结构性特征,再在粒子滤波跟踪框架下将深度特征和手工特征相结合用于目标跟踪。同时,该算法以一种懒惰的方式更新跟踪模型,避免了模板频繁更新而导致的昂贵计算代价。实验结果表明,与已有的传统粒子滤波跟踪方法相比,该方法在现有的跟踪基准测试中显示出优越的性能,在复杂背景、光照以及目标形变等恶劣条件影响下仍能稳定地跟踪目标,具有更强的鲁棒性。

关键词:目标跟踪;粒子滤波;卷积神经网络;深度特征;手工特征

中图分类号:TP391.41

文献标识码:A

文章编号:1673-629X(2020)06-0023-06

doi:10.3969/j.issn.1673-629X.2020.06.005

Particle Filter Video Target Tracking Algorithm Based on Deep Learning

LI Min, WU Sha

(School of Information Engineering, Chang'an University, Xi'an 710000, China)

Abstract: In video target tracking, the tracking algorithm has poor robustness due to the influence of changes in the environment and its objects and other factors. A target tracking algorithm based on pre-trained convolutional neural network and combining depth features with manual features under the framework of particle filter is proposed to solve this problem. In this algorithm, tens of thousands of generic targets images are pre-trained offline by convolutional neural network, and structural features ranging from simple to complex that can represent generic targets can be obtained. Then deep representations and hand-crafted features are combined for target tracking under the particle filter framework. At the same time, the algorithm updates the tracking model in a lazy way and avoids the expensive calculation cost caused by frequent update of templates. The experiment shows that compared with the traditional particle filter tracking methods, this method has superior performance in the existing tracking benchmark test, and it can still track the target stably under the influence of complex background, illumination, target deformation and other harsh conditions, with stronger robustness.

Key words: target tracking; particle filter; CNN; depth feature; manual feature

0 引言

目标跟踪是一个具有挑战性的计算机视觉问题,特别是存在遮挡、目标变形、光照变化等情况下。粒子滤波器由于能够有效地确定目标的位置和大小,多年来一直被应用于视觉跟踪^[1-2]。例如,文献[3]提出了一种基于自顶向下视觉注意力计算模型的粒子滤波视觉跟踪器。该算法是检测目标相关的突出区域,然后将显著区域发送到粒子滤波器,确定目标的位置。文献[4]提出了一种名为萤火虫的优化算法,该算法通过动态增加有意义粒子的数量来提高跟踪性能。文献[5]提出基于一种新的迭代粒子滤波器(IPF),该滤波

器在收缩的搜索范围内对粒子进行迭代采样。此外,文献[6]中将粒子滤波与基于高斯混合的目标外观模型结合使用。虽然基于粒子滤波器的视觉跟踪方法,或者基于相关滤波器的视觉跟踪方法^[7]的跟踪结果都可以接受,但是它们不如基于卷积神经网络(CNN)产生具有高度识别性的目标特征^[8]的跟踪器有效。

卷积神经网络在成功应用于对象检测任务^[9]之后,又被应用于视觉跟踪,到目前为止,已经提出了几种不同的CNN结构。文献[10]提出的是基于3个卷积层和若干降采样层的CNN对当前帧和上一帧进行采样,获取目标和背景的空间、时间特征。采用了两个

收稿日期:2019-10-10

修回日期:2020-02-12

基金项目:国家自然科学基金(61771075)

作者简介:李敏(1992-),女,硕士研究生,研究方向为视频中的目标跟踪。

采样对,分别输入两个 CNN 网络,得到 4 张关键点的概率图,从而提高了跟踪的精确度。文献[11]是离线训练多层的卷积神经网络,学习得到可以反映图像一般性特征的卷积核后,通过迁移学习在线特征提取,使用粒子滤波算法进行目标跟踪。文献[12]采用较小的网络来进行特征提取并在全连接层进行综合,得到输出向量,最后根据该向量来判别样本是目标还是背景,从而估计目标位置。虽然近几年基于深度学习的跟踪算法大幅提升了跟踪算法性能,对目标和背景信息进行较详细的提取,具有很好的判别能力,但在应对目标表观发生变化的情况下,缺乏目标全局的高级语义信息,所以跟踪算法性能还有待提升。

文中将卷积粒子滤波技术与 CNN 相结合,以提高其在变形、光照变化、快速运动等情况下的性能。该算法使用预先训练的网络来帮助 PF 建模跟踪目标对象,卷积粒子滤波通过贝叶斯最优估计来选择候选目标并定位目标,避免了人工干预。深层特征和手工特征融合带来了更稳定的观测模型,使得 PF 在保证鲁棒跟踪的同时,能够选择少量的粒子,减轻了计算负担。这些操作克服了跟踪过程中存在较大的遮挡或快速运动,目标形变等问题,实现了鲁棒跟踪。

综上所述,针对单一特征带来的跟踪不足问题,将基于目标的颜色特征和基于 CNN 的特征方法^[13]结合

对目标进行特征表述,同时集合了两种算法的优点,利用深度学习和粒子滤波的特点提出了一种新的跟踪方法,该方法使用预先训练的网络来帮助粒子滤波跟踪来建模目标对象。粒子滤波是指通过贝叶斯最优估计来选择候选目标并定位目标,避免了人工干预,实现了自动跟踪。同时,深层特征和手工特征融合带来了更稳定的观测模型,使得 PF 在保证鲁棒跟踪的同时,能够选择少量的粒子,减轻了计算负担。

1 基于 CNN 的粒子滤波跟踪算法

1.1 深度学习网络的体系结构

在卷积神经网络训练数据时,空间信息会随着网络的加深而稀释,因此会导致图像中目标定位不精准的情况^[14]。经研究表明,在卷积神经网络结构中,第三个卷积层的卷积特征保留了目标更多的空间细节信息^[15],如目标的位置信息等。由于图像经过卷积神经网络的三个卷积层得到的特征保留了目标的空间位置信息,而且考虑到卷积运算的复杂耗时,跟踪算法必须要满足实时性,因此文中设计一个包括 3 个卷积结构的深度网络,每个卷积结构包含一个卷积层和一个池化层,加上一个完全连接层由此构建了一个 9 层的卷积神经网络进行特征提取。整体框架如图 1 所示。

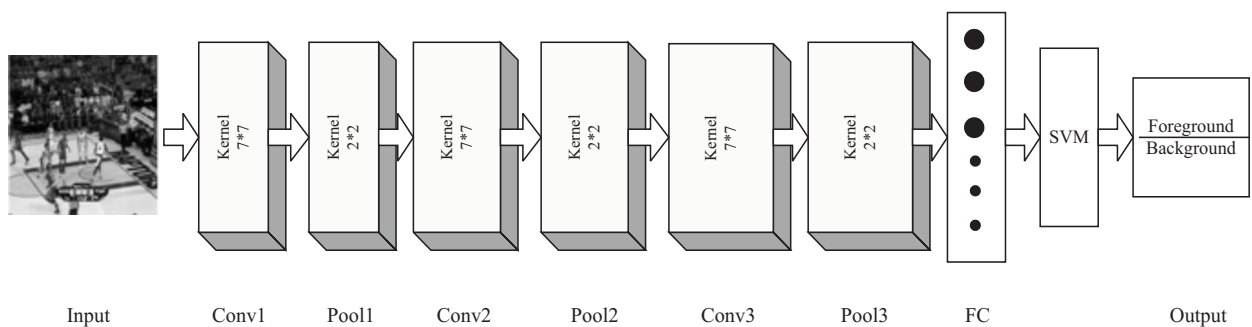


图 1 CNN 网络结构

输入层,向网络输入 $X = 58 \times 58$ 的目标图像。

卷积层,卷积层的卷积核对输入的图像进行卷积操作,从而得到输入图像的特征图。过程:Conv1 层使用 8 个 7×7 的卷积核,得到 8 个 52×52 大小的特征图;Conv2 层使用 16 个 7×7 的卷积核,得到 16 个 20×20 大小的特征图;Conv3 层使用 32 个 7×7 的卷积核,得到 32 个 4×4 大小的特征图。卷积槽的激活函数为 sigmoid 函数:

$$H_x = \text{sigmoid}(f_x \otimes H_{x-1} + b_x) = \frac{1}{1 + e^{-(f_x \otimes H_{x-1} + b_x)}} \quad (1)$$

$(x = 1, 2, 3)$

其中, f_1, f_2 和 f_3 分别是三个卷积层的卷积核大小,分别用于生成 Conv1、Conv2 和 Conv3 三层特性, \otimes 为卷积操作, H_0 为原始的图像输入 X , b_x 是一个偏移值。

池化层,通常情况是在卷积层之后,根据一定的规则对卷积层生成的特征图进行池化。池化层的功能主要有两点:(1)对特征图进行降维;(2)在一定程度上保持特征图特征的尺度不变性。

$$p(i, j) = \frac{\text{sum}}{a_1 + a_2 + a_3 + a_4} \quad (2)$$

其中, a_1, a_2, a_3, a_4 为四个相邻的卷积核的子采样值, sum 由以下公式得出:

$$\text{sum} = a_1 p(i, j) + a_2 p(i, j + 1) + a_3 p(i + 1, j) + a_4 p(i + 1, j + 1) \quad (3)$$

其中 $p(i, j), p(i, j + 1), p(i + 1, j), p(i + 1, j + 1)$ 表示卷积特征映射中的四个邻居值。这些简单的操作不仅减少了数据数量,而且提取了有用的深层信息。

全连接层,32 个 2×2 的卷积核对第三个池化层输

出的特征图进行卷积操作,得到1个128维的目标特征矩阵。

输出层,在完全连接层之后,支持向量机有助于区分正样本和负样本,通过支持向量机对目标特征进行分类,将输出层分为目标和背景。

由于深度学习网络需要大量的训练样本,而这在视觉跟踪中经常是不可用的,因为从初始帧中提取的可靠正实例数量很少。在文献[16]中指出对使用大规模数据集训练卷积神经网络模型而提取到的特征,对其他的任务也具有普适性。因此这里把这个网络看成是用于在线跟踪应用程序的预定义数据集。数据集由视觉跟踪器基准^[3]提供,并且包括大量的正和负样本。该数据集具有光照变化、旋转、尺度变化等多种数据增强的挑战性。在训练过程中,随机梯度法(SGD)优化了模型中的参数。

1.2 在线粒子滤波跟踪

在跟踪过程中,传统的粒子滤波以颜色直方图作为外观模型,它不能有效地处理突然的变化,因为它把传播模型看成是一个简单的零均值高斯过程,其位置的方差是一致的。而将卷积神经网络的层参数固定,仅对全连接层和输出层参数进行更新,得到跟踪目标的深度特征,再在粒子滤波框架下得到跟踪目标的颜色直方图特征。而将深层特征与手工制作的特征融合在一起,用于视频目标跟踪,提高跟踪目标的精度,得到持续和鲁棒的跟踪。

在线粒子滤波跟踪算法包括四个步骤:

步骤1:初始化粒子为 $X_0^n \sim p(X_0)$, 权重为 $w_0^n = \frac{1}{n}$, 其中 $n = 1, 2, \dots, N$, 其中 n 是粒子数。状态向量定义为 $X = [s \ x \ y \ \dot{x} \ \dot{y} \ c]$ 。其中 x 和 y 是矩形的中心垂直位置和水平位置, \dot{x} 为水平速度, \dot{y} 为垂直速度, s 为围绕希望跟踪的目标以 (x, y) 为中心的矩形的比例, $c = [c_1 \ c_2]$ 是一个一维的特征向量,包含两个部分: $c_1 \in R^{1 \times 128}$ 是预训练学习网络的深层特征向量。 $c_2 = [h_1, h_2, \dots, h_U] \in R^{1 \times U}$ 为颜色直方图, U 为直方图箱数,可以通过三通道融合^[17]的计算公式得到。

c_2 中每个元素 h_U 的计算公式为:

$$h_U = K \sum_{i=1}^N \delta[f(x_i) - u], u = 1, 2, \dots, U \quad (4)$$

其中, K 为归一化常数, N 为目标区域像素总个数, u 为直方图箱数索引值, $f(x_i)$ 为 x_i 所在直方图段数的函数, δ 为 Delta 函数。

步骤2:粒子传播采用的是一个动态模型:

$$x_t^n = Fx_{t-1}^n + q_t^n \quad (5)$$

其中, x_t^n 为第 n 个粒子在 $t = 1, 2, \dots, t$ 时的状态, q_t^n 为

零均值高斯噪声,对于每个粒子 $q_t^n \sim N(0, Q)$, F 为线性运动模型。

$$F = \begin{bmatrix} 1 & T & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & T & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$Q = \begin{bmatrix} \sigma^2 & 0 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & 0 & \sigma_s^2 \end{bmatrix}$$

其中, T 为相邻两帧之间的周期, σ_s^2 为尺度的方差, σ^2 为位置和速度的方差。

步骤3:通过观测模型对粒子进行加权,每个粒子的权重为:

$$w_t^n = e^{-\lambda(D(n))^2}, n = 1, 2, \dots, N \quad (6)$$

其中, λ 为设计参数, $D(n)$ 为 Bhattacharya 距离:

$$D(n) = \sqrt{1 - \sum_{u=1}^{128+U} \sqrt{m(u)c_1^n(u)}} \quad (7)$$

其中, $m(u)$ 是用户初始化并及时更新的当前动态模型的特征向量; $c_1^n(u)$ 是第 n 个粒子的特征向量,包含两个部分:预训练学习网络的深层特征向量 $c_1^n(u)$ 和颜色直方图特征向量 $c_2^n(u)$ 。因为研究发现,HSV 对光照变化具有较强的鲁棒性^[18],故该算法选择了 HSV 颜色直方图。

步骤4:采用归一化方法,保证 $\sum_{n=1}^N w_t^n = 1$, 预测目标的位置为:

$$(x_t, y_t) = \sum_{n=1}^N w_t^n (x_t^n, y_t^n) \quad (8)$$

1.3 模型更新

遮挡、背景干扰、目标变形以及摄像机的参数都会影响跟踪器的鲁棒性。然而,在物体的外观不是总在变化的情况下,良好的跟踪模型可以长期保持。此外,过度更新容易导致移位,带来昂贵的计算代价。在这些条件影响下,文中以一种懒惰的方式更新跟踪模型。假设目标在当前帧中的权重为 w_t , 并且之前十个连续帧中的权重为 w_{t-10} 到 w_{t-1} 。比较平均值 $\bar{w} = \frac{\sum_{i=1}^{10} w_{t-i}}{10}$ 和 w_t :

如果 $|w_t - \bar{w}| < T$ (阈值),观测模型将不会更新;否则,模型将获得新的外观特性。即从当前帧中选取权重最大的 $N/5$ 个粒子,计算其均值特征 \bar{C} , 然后更

新模型的新特征向量为:

$$C_n = a \times \bar{C} + (1 - a) \times C_0 \quad (9)$$

其中, a 为权重参数, C_0 为更新前模型的特征向量。通过这种方式,模型可以保持以前的外观和当前的功能,阈值和加权平均控制了更新程度。该方法使跟踪算法能够处理运动、变形和遮挡等挑战。

2 实验

实验计算机配置为 Windows8 系统内配置 Ubuntu16.04 虚拟机,以及 CUDA8.0、CUDNN5.1、NVIDIA GTX1070TI。该算法在 Matlab2016a 上实现,对于离线学习,是在 Ubuntu16.04 虚拟机下进行,从最近发布的跟踪基准数据集^[3]中收集 9 075 个正样本和 17 535 个负样本,实验选用目前主流的跟踪平台 OTB-100 和 OTB-50 数据集中具有实质性变化的全注释视频中的 10 个视频序列进行实验验证,这些视频序列在跟踪问题上都存在不同场景和不同程度的跟踪困难因素:遮挡、快速运动、形变、光照变化、背景相等等。

2.1 参数设置

λ 值设置为仅用于传统的 PF 跟踪的 0.2,用不同数量的粒子对几个视频进行测试,得到粒子数为 40 时可以带来良好的效率。通过整体比较,子采样 a_1 、 a_2 、

a_3 与 a_4 的核参数设置成 0.25。在训练和跟踪过程中,它们保持相同的值不变。

2.2 定量比较

从两个方面(跟踪精度、跟踪成功率)来评价该跟踪算法的跟踪结果,并比较各个算法性能。为了定量地突出深层特征和手工制作特征的跟踪效果,将该算法与当前多种主流跟踪算法进行对比分析,包括 CPF、LOT、VTD、DFT、CSK、MIL 跟踪算法。为了比较公平,使用作者提供的代码在测试序列中从第一帧的基础真值位置初始化它们进行实验。

为了评估这些跟踪器,在测试序列中从第一帧的基础真值位置初始化它们,并报告单程评估(OPE)值^[3]。重叠率定义为 JACARD 系数:

$$S = \frac{r_t \cap r_m}{r_t \cup r_m} \quad (10)$$

其中, r_t 和 r_m 分别为地面真实边框和跟踪边框。一个视频序列的重叠率是所有 Jaccard 系数的平均值。若 $s > 0.5$,则认为该帧目标被正确跟踪^[19]。平均跟踪成功率的结果见表 1,其中粗体表示该跟踪序列中成功率最高的值。由表 1 中可知,文中算法在这些视频序列的跟踪性能总体上优于其他主流跟踪算法。

表 1 跟踪成功率

| 序列 | Ours | VTD | CPF | CSK | MIL | LOT | DFT |
|----------|----------------|--------------|--------------|---------|-------|--------------|--------------|
| Bolt | 0.873 | 0.579 | 0.438 | 0.349 | 0.206 | 0.347 | 0.338 |
| Singer1 | 0.752 | 0.359 | 0.374 | 0.342 | 0.426 | 0.513 | 0.353 |
| Car4 | 0.572 | 0.344 | 0.338 | 0.239 | 0.342 | 0.320 | 0.316 |
| Jogging | 0.867 | 0.527 | 0.399 | 0.348 | 0.505 | 0.261 | 0.297 |
| Lemming | 0.859 | 0.339 | 0.489 | 0.664 | 0.319 | 0.358 | 0.318 |
| Dudek | 0.735 | 0.608 | 0.736 | 0.297 | 0.232 | 0.435 | 0.410 |
| Soccer | 0.429 | 0.568 | 0.368 | 0.337 | 0.251 | 0.533 | 0.459 |
| Girl | 0.531 | 0.607 | 0.637 | 0.548 | 0.544 | 0.397 | 0.637 |
| Skating1 | 0.583 | 0.540 | 0.416 | 0.419 | 0.401 | 0.643 | 0.525 |
| Tiger1 | 0.759 | 0.559 | 0.485 | 0.407 | 0.297 | 0.351 | 0.328 |
| Average | 0.687 7 | 0.503 3 | 0.468 8 | 0.395 5 | 0.352 | 0.416. | 0.398 8 |

图 2 为 7 种跟踪算法对 10 组测试视频序列的整体跟踪精度和跟踪成功率的 OPE 曲线,可以看出,文

中跟踪算法的跟踪成功率和跟踪精度对于这 10 组测试视频中的目标的平均值最高。

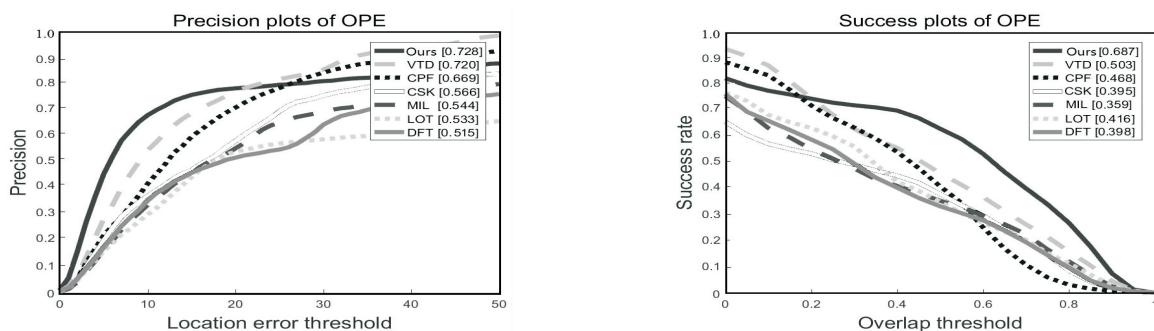


图 2 10 个测试序列的跟踪精度和跟踪成功率曲线

2.3 定向比较

图 3 是 7 种跟踪算法 10 个视频序列的部分序列跟踪结果,左上角是该图像序列的当前帧数。

(1) 目标遮挡, Jogging, Lemming, Dudek, Soccer 序列中,在运动过程中被跟踪目标出现被不同物体不同程度的遮挡。Jogging 在第 79 帧时目标被完全遮

挡,直到第 86 帧目标又重新出现,而其他跟踪算法皆丢失目标,只有文中算法能很好地跟踪到目标;Lemming 在第 317 帧时,目标被部分遮挡时,除了文中算法和 CSK 和 VTD,其他算法都出现了跟踪漂移,不难看出文中算法在出现遮挡情况下始终正确跟踪目标,对目标遮挡有较好的鲁棒性。



图 3 7 种跟踪算法 10 个视频序列的部分序列图

(2) 光照变化, Singer1, Car4, Skating, Tiger1 视频序列中,在运动过程中被跟踪目标出现不同程度的光照变化。其中 Singer1 序列在 79-142 帧光照出现了不同程度的强弱变化,VTD 彻底丢失目标,而除了文中算法,其他算法都出现了跟踪漂移。在 Tiger1 视频序列中,在 175 帧和 291 帧的时候光照突然变强,除了文中算法和 CPF、DFT,其他算法都发生了跟踪漂移。不难看出文中算法对出现光照变化时始终正确跟踪目标,对光照变化有较好的鲁棒性。

踪算法都出现了跟踪漂移的情况。在 Tiger1 视频跟踪序列中,90 帧的时候出现快速运动的情况,其中 VTD、LOT 跟踪丢失,其他算法出现跟踪漂移的现象,只有文中算法保持了良好的跟踪现象。不难看出文中算法对在出现目标快速运动时始终正确跟踪目标,有较好的鲁棒性。

(3) 快速运动, Lemming, Dudek, Soccer, Tiger1 视频序列中,被跟踪目标在运动过程中都存在快速运动的情况。其中 Lemming 序列中,在 535 帧到 583 帧中存在快速运动的情况,除了文中算法和 CSK,其他跟

踪算法都出现了跟踪漂移的情况。在 Bolt 序列中,整个被跟踪目标都是处于形变的状态,而除了文中算法和 VTD 外,其他算法均跟踪失败,而在 137 帧后,VTD 出现了跟踪漂移的现象。在 Tiger1 序列中,在 166 帧和 289 帧时出现目标形变的现象,在 166 帧时 VTD 和 LOT 跟踪失败,在 289 帧

时,除文中跟踪算法外,其他算法均出现了跟踪漂移的现象。所以文中算法对目标形变的情况也表现出了较好的鲁棒性,基本上能始终正确跟踪目标。

(5)相似背景, Dudek, Soccer, Skating1 视频序列中,被跟踪目标在运动整个视频序列中都出现了相似的背景情况。在 Dudek 序列中,整个跟踪过程都处于灰色基调下,在 173 到 200 帧时,出现了背景杂斑现象,这时 CSK 彻底丢失目标,而除了文中算法,其他的跟踪算法虽然也完成了整个过程的跟踪,但是中间出现了跟踪漂移的现象。Soccer 序列和 Dudek 序列相似,整个过程背景与目标是同一个颜色,但是该序列中,从 117 帧开始,除了文中算法、DFT 和 LOT,其他算法均跟踪失败。而在 Skating1 序列中,从 99 帧开始出现了相似的目标,文中算法虽然中间出现过跟踪漂移,但是最终还是完成了目标跟踪。所以文中算法对存在相似背景的情况也表现出了较好的鲁棒性,基本上能始终正确跟踪目标。

3 结束语

提出了一种基于简化 CNN 的视觉跟踪算法,该算法将预处理网络的全连接层的深度特征输出与 HSV 颜色空间中颜色直方图的手工特征相连接,将两种特征相结合,用于在线粒子滤波的目标跟踪。深层特征和手工特征的融合更有效地代表了目标,在具有背景杂波、光照变化、变形和遮挡等属性的视频中,相比其他跟踪方法,该方法有较强的鲁棒性。

参考文献:

- [1] BABENKO B, YANG M, BELONGIE S. Robust object tracking with online multiple instance learning [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(8): 1619–1632.
- [2] 李天成, 范红旗, 孙树栋. 粒子滤波理论、方法及其在多目标跟踪中的应用 [J]. 自动化学报, 2015, 41(12): 1981–2002.
- [3] WU Y, LIM J, YANG M. Online object tracking: a benchmark [C]//2013 IEEE conference on computer vision and pattern recognition. Portland, OR: IEEE, 2013: 2411–2418.
- [4] 佟威, 和箫, 卢英. 引入视觉显著性的多特征融合跟踪 [J]. 计算机科学与探索, 2017, 11(3): 438–449.
- [5] HARE S, SAFFARI A, TORR P H S. Structured output tracking with kernels [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2016, 38(10): 2096–2109.
- [6] ROSS D A, LIM J, LIN R S, et al. Incremental learning for robust visual tracking [J]. International Journal of Computer Vision, 2008, 77(1–3): 125–141.
- [7] PADFIELD D, RITTSCHER J, ROYSAM B. Coupled minimum-cost flow cell tracking for high-throughput quantitative analysis [J]. Medical Image Analysis, 2011, 15(4): 650–668.
- [8] HENRIQUES J F, RUI C, 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.
- [9] FAN J, XU W, WU Y, et al. Human tracking using convolutional neural networks [J]. IEEE Transactions on Neural Networks, 2010, 21(10): 1610–1623.
- [10] HONG S, YOU T, KWAK S, et al. Online tracking by learning discriminative saliency map with convolutional neural network [C]//Proceedings of the 32th international conference on machine learning. Lille, France: JMLR, 2015: 597–606.
- [11] LI H, LI Y, PORIKLI F. Robust online visual tracking with a single convolutional neural network [C]//2014 Asian conference on computer vision (ACCV). Singapore: [s. n.], 2014: 194–209.
- [12] MA C, HUANG J, YANG X, et al. Hierarchical convolutional features for visual tracking [C]//2015 IEEE international conference on computer vision (ICCV). Santiago: IEEE, 2015: 3074–3082.
- [13] 刘志成, 祝永新, 汪辉, 等. 基于卷积神经网络的多目标实时检测 [J]. 计算机工程与设计, 2019, 40(4): 1085–1090.
- [14] NAM H, HAN B. Learning multi-domain convolutional neural networks for visual tracking [C]//Proc of the 2016 IEEE conference on computer vision and pattern recognition. Las Vegas, NV, USA: IEEE, 2016: 4293–4302.
- [15] CIMPOI M, MAJI S, VEDALDI A. Deep filter banks for texture recognition and segmentation [C]//2015 IEEE conference on computer vision and pattern recognition (CVPR). Boston, MA: IEEE, 2015: 3828–3836.
- [16] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]//Proceedings of IEEE conference on computer vision and pattern recognition. Columbus, OH, USA: IEEE, 2014: 269–275.
- [17] SONG Y, LI Q, SUN F. Shannon entropy-based adaptive fusion particle filter for visual tracking [C]//2009 Chinese conference on pattern recognition. Nanjing: [s. n.], 2009: 1–5.
- [18] QIAN X, WANG Y, HAN L. An object tracking method based on guided filter for night fusion image [J]. Infrared Physics & Technology, 2015, 74: 38–43.
- [19] 刘芳, 黄光伟, 路丽霞, 等. 自适应模板更新的鲁棒目标跟踪算法 [J]. 计算机科学与探索, 2019, 13(1): 83–96.