

基于核相关滤波的 TLD 跟踪算法的研究

董春燕, 刘 怀, 梁秦嘉

(南京师范大学电气与自动化工程学院, 江苏 南京 210023)

[摘要] 对 TLD 跟踪算法进行改进,以提高在跟踪目标发生尺度变化或被遮挡时的跟踪性能. 首先使用 KCF 跟踪器替代 TLD 算法中原有的中值光流跟踪器,并在特征提取时增加目标的 Lab 颜色特征,在寻找目标位置时引入尺度估计,在模型更新阶段引入跟踪状态判别机制,通过设定跟踪器中输出响应最大值阈值、APCE 阈值及检测器中随机森林分类器阈值来判断跟踪器跟踪结果的可靠性,改善跟踪器在尺度变化、出现遮挡、光照变化等情况下的跟踪效果. 针对 TLD 算法中的检测器,为了减少大量无意义的窗口,提升算法在存在遮挡时的精确性,在检测之前使用 Kalman 滤波预估出目标位置,在预估位置周围使用改进的级联分类器更精准地定位目标,改进的级联分类器的前两级仍采用方差分类器和随机森林分类器,第三级则采用改进的 KCF 跟踪器. 在 OTB-50 数据集上的实验结果分析表明,该算法跟踪性能优于其他算法,能够满足实时性.

[关键词] 目标跟踪,核相关滤波,特征融合,遮挡检测

[中图分类号] TP391 **[文献标志码]** A **[文章编号]** 1672-1292(2020)04-0037-07

TLD Object Tracking Algorithm Based on Kernelized Correlation Filter

Dong Chunyan, Liu Huai, Liang Qinjia

(School of Electrical and Automation Engineering, Nanjing Normal University, Nanjing 210023, China)

Abstract: TLD tracker is improved in this paper in order to increase the tracking performance when the tracked object scale is changing or in occlusion. For the tracker, KCF tracker is used to replace the original tracker of TLD algorithm, and KCF tracker is improved as follows: the Lab color feature of the object is added in feature extraction, scale estimation is introduced in search for the object location, in the model updating stage, and tracking state judgment mechanism is adopted to determine the reliability of the tracking results by setting the threshold value of maximum output response, the threshold value of APCE and the threshold value of the random fern classifier, so as to improve the tracking effect of the tracker under the conditions of scale variation, occlusion and illumination variation. For the detector, in order to reduce a large number of meaningless windows before detection, Kalman filter is used to estimate the object position, and then an improved cascade classifier is used to locate the object more accurately around the estimated position. Variance classifier and random fern classifier are still used in the first two levels of the improved cascade classifier, and the third level is classified by the improved KCF tracker. Qualitative and quantitative analysis of experimental results on the OTB-50 dataset show that the tracking performance of the proposed algorithm is superior to that of other tracking algorithms and that it can ensure the real-time performance.

Key words: object tracking, kernelized correlation filter, feature fusion, occlusion detection

目标跟踪是视觉分析和处理领域的一个重要研究内容,在智能交通管理、视频监控和军事制导等领域有广泛的应用^[1]. 目标跟踪的任务是确定目标在视频序列中每一帧的位置和大小,并构建其运动轨迹. 然而,目标在运动过程中可能会发生形变、尺度变化以及相互遮挡等干扰情况,会造成目标丢失或跟踪精确度降低. 因此,需要研究一种在干扰存在的情况下具有先验不确定性的实时目标跟踪方法. 早期的运动目标跟踪借鉴了图像配准的思想^[2-3],主要有 KLT(kanade-lucas-tomasi)跟踪、Meanshift 跟踪、粒子滤波跟踪算法等. 随着机器学习技术在目标跟踪领域的迅速发展和应用,研究人员提出了一种更强大的跟踪框架,

收稿日期:2020-03-06.

基金项目:国家自然科学基金项目(61603194).

通讯作者:刘怀,副教授,研究方向:数字图像处理、实时控制系统. E-mail:liuhuai@njnu.edu.cn

该框架具有较强的泛化能力,被称为判别类跟踪^[4-5]。近几年,从世界顶级的视觉期刊和会议上来看,视觉跟踪的研究重点已转向基于相关滤波的目标跟踪算法以及深度学习的目标跟踪算法^[6],如 Choi 等^[7]提出的 RDT(real-time deep tracking)算法,Gordon 等^[8]提出的 Re3(real-time recurrent regression networks)算法以及 Wang 等^[9]提出的 SiamMask 算法等。但目前深度学习的跟踪算法对于硬件性能的要求较高,普遍存在实时性较差的问题,准确性和实时性难以兼得。

在判别类算法中,核相关滤波(kernel correlation filter,KCF)跟踪算法^[10]能够兼顾跟踪精度和速度,但在目标被遮挡或发生形变的情况下,易丢失目标。TLD 算法将跟踪器与检测器结合起来以解决此问题,然而该算法耗时量大,且跟踪器性能不稳定。为了提高 TLD 算法的性能,许多学者对其进行了改进。在对其中的跟踪器进行改进时,张惊雷等^[11]使用尺度自适应均值偏移(scale-adaptive mean-shift for tracking,ASMS)跟踪器代替原始 TLD 跟踪算法中的中值光流跟踪器。毛晓波等^[12]采用 FAST 特征点代替中值光流跟踪器中均匀分布的特征点,一定程度上提高了算法的准确性,但基于特征点的中值光流法跟踪算法耗时量大,且光流法假设目标颜色不变,在发生光照变化时会导致目标丢失。在对其中的检测器进行改进时,王皎尧等^[13]使用光流法对检测区域进行粗定位以缩小检测范围,但在目标发生遮挡时,特征点数量减少甚至完全消失,不能准确预测目标位置。常立博等^[14]使用基于 ViBe 算法的检测器代替原算法的方差分类器,虽然一定程度上提高了算法的跟踪速度,但在背景运动时,不能获得前景目标,会导致目标丢失。

本文借鉴以上算法的设计思路,对 TLD 算法进行了改进。考虑到 TLD 算法中的中值光流跟踪器在亮度发生变化时不够鲁棒且计算量大,而 KCF 跟踪器对光照变化鲁棒且运算速度快,故而本文使用 KCF 跟踪器替代原有的中值光流跟踪器。为了进一步提高跟踪器的精度,本文又对其进行改进,在特征提取阶段增加了目标特征表达,在寻找目标位置时加入了尺度估计,在模型更新阶段对更新策略进行了优化;针对其中的检测器,在检测之前引入了卡尔曼滤波对目标位置进行预测,既确保了算法存在遮挡时的精确性,又利用了目标的运动信息,极大地减少了检测窗口的数目以及相似目标的影响。同时,对检测器中的级联分类器进行改进,级联分类器的前两级保持不变,第三级则采用上述改进的 KCF 跟踪器进行分类。在 OTB-50 数据集上的实验结果分析表明,本文算法跟踪性能优于其他算法。

1 相关算法基础

1.1 TLD 算法

TLD 算法包括跟踪器、检测器、综合模块与学习器。跟踪器通过中值光流法执行帧到帧的目标跟踪;检测器采用级联分类器,逐帧扫描以检测目标位置;综合模块根据跟踪器和检测器的结果得出最终的预测位置;学习器根据此预测位置,对检测器有选择地进行更新。TLD 算法为遮挡跟踪和长时间跟踪提供了可行的方法与框架,但当 TLD 算法的中值光流跟踪器在当前帧的图像亮度与前一帧不同时,会导致跟踪失败,且该算法计算量大。检测器需要检测大量无意义的窗口,这些无意义的窗口造成了内存消耗和时间消耗的问题。学习器中主要是对检测器进行学习,缺乏必要的跟踪器模型的保存与更新,从而使跟踪器不能适应目标的变化。同时,TLD 算法未利用图像的颜色信息,当目标外观变化时易累积模型误差。

1.2 KCF 算法

KCF 跟踪算法在判别类算法中具有较高的精度,其将跟踪转化为脊回归的问题,并利用循环矩阵的性质,通过在频域中训练出的滤波器模型,来寻找相关响应值最大的图像位置,即为后续帧中的目标位置。由于是在线训练,为了适应目标的变化,需要使用线性插值对目标样本特征和滤波器模型进行更新,从而实现对目标的学习与检测。由于相关运算可在频域内有效地计算,因此 KCF 跟踪算法具有很高的运行速度。但 KCF 算法中缺乏遮挡判别机制,当目标被遮挡时,模型中会引入错误的目标信息,导致跟踪漂移甚至目标丢失。

2 基于核相关滤波的 TLD 跟踪算法

本文基于 KCF 算法对 TLD 算法进行改进,改进的算法框图如图 1 所示。该算法不仅采用改进的 KCF 跟踪器替代原有的中值光流跟踪器,而且在检测器前加入卡尔曼(Kalman)滤波对目标位置进行预测,并使用改进的 KCF 跟踪器代替检测器中耗时的最近邻分类器,有效地提高了算法的速度与跟踪精度。

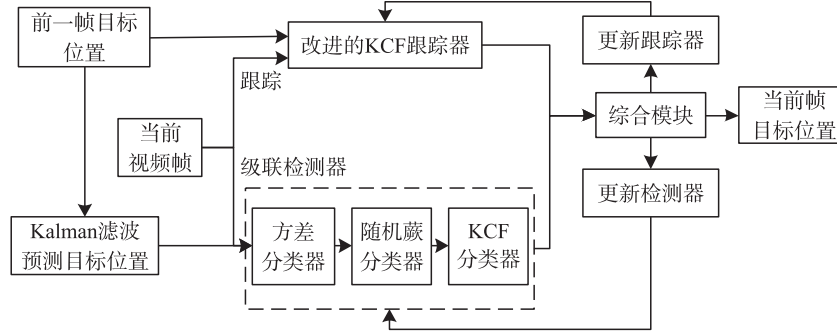


图1 本文算法框图

Fig. 1 Block diagram of the proposed algorithm

2.1 多维特征及尺度自适应的改进 KCF 跟踪算法

传统 KCF 算法采用单一 HOG 特征对目标建模,对目标的描述有限. HOG 特征描述了目标的形状,对颜色和光照变化具有较强的鲁棒性^[15],但当目标快速变形时,无法确定目标位置. Lab 特征在 Lab 颜色空间求每个像素的 Lab 值和 15 组特定的 Lab 值的欧氏距离的平方以此描述目标颜色,对目标形变鲁棒,然而当发生颜色变化时,无法确定目标位置. HOG 特征和 Lab 特征分别描述了目标的形状和颜色,具有一定的互补性. 因此,本文算法在提取 HOG 特征的基础上进一步提取 Lab 特征以丰富特征表达.

传统的 KCF 跟踪算法不能适应目标大小的变化,在目标变大或变小时,可能会出现目标丢失或跟踪精度降低的问题,本文受文献[16]的启发,通过在同一帧图像上提取不同尺度下的候选目标样本特征,来计算目标的最优尺度. 为了不显著降低运行速度,本文只选取长宽分别为前一帧目标搜索区域的 0.95 倍、1.0 倍以及 1.05 倍的 3 个尺度作为检测尺度,如图 2 所示.

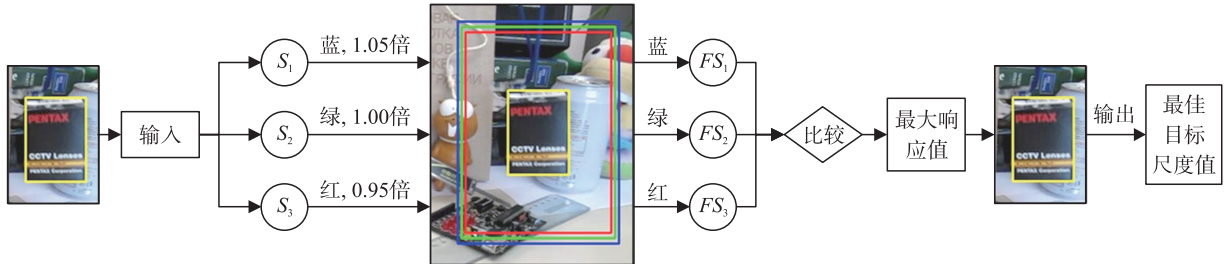


图2 尺度检测原理图

Fig. 2 Block diagram of scale detection

2.2 应对遮挡的改进措施

TLD 算法为遮挡跟踪和长时间跟踪提供了可行的方法与框架,但 TLD 算法中的检测器需要检测大量无意义的窗口,这些无意义的窗口造成了内存消耗和时间消耗.

一般来说,前一帧跟踪到的目标位置对当前帧目标位置具有一定的指示性,此外 Kalman 滤波能够根据之前帧的目标位置来得到当前帧的预测位置. 为了比较选择前一帧跟踪到的目标位置 (KCFprevFrame) 作为预测位置与选择 Kalman 滤波得到的当前帧的预测位置的准确性,本文在 OTB-50 数据集上进行测试,选取中心位置误差这一评价指标对跟踪误差进行分析. 中心位置误差 (center location error, CLE) 计算公式为:

$$d_{CLE} = \sqrt{(x_T - x_C)^2 + (y_T - y_C)^2}, \quad (1)$$

式中, (x_T, y_T) 是跟踪到的目标中心位置坐标; (x_C, y_C) 是真实的目标中心位置坐标.

表 1 为上述两种预测算法在 OTB-50 数据集上的 5 个视频序列上运行得到的中心位置误差. 可以看出,前一帧跟踪到的目标位置与当前帧目标真实位置的中心位置误差较大,而采用 Kalman 滤波得到的预测位置与目标真实位置的中心位置误差较小.

表 1 不同算法的平均中心位置误差
Table 1 Average center location errors of different algorithms

算法	视频序列				
	Deer	BlurFace	Jumping	Basketball	Dudek
Kalman	14.83	25.33	11.26	6.95	11.66
KCFprevFrame	20.88	28.35	11.75	7.04	11.73

因此,本文以 Kalman 滤波预测的目标位置为中心,以长宽分别为前一帧目标区域的两倍范围来确定搜索区域,使用级联分类器实现精确的目标定位. 由于 TLD 算法中检测器的前两级分类器的运行速度较快,而第三级分类器非常耗时,故本文中检测器的前两级分类器保持不变,为了在第三级分类中高效准确地搜索目标,本文采用与跟踪器相同的 KCF 算法进行分类,即将通过前两级的样本依次输入 KCF 跟踪器中,通过输出响应的最大值来进行样本筛选:

$$f_{\max} = \max(F^{-1}(F(k^x) \odot F(\alpha))), \quad (2)$$

式中, k^x 为目标样本特征 x 和检测样本特征 z 的核互相关函数; \odot 表示点乘.

设定输出响应最大值的阈值,阈值初始设定为 $0.6 \times f_{\max}$,后续会进行自动的学习更新. 将满足判定跟踪状态条件且最大值大于阈值的样本作为检测器的输出结果. 若所有的剩余样本均不满足判断条件,则判定当前帧没有检测到目标. 当连续 5 帧级联分类器无输出时进入全局搜索,这样显著减少了大量无意义的窗口,提高了实时性.

同时引入跟踪状态判别机制,将跟踪器响应图的峰值 f_{\max} 以及描述响应图波动程度的平均峰值相关能量 APCE^[17] 作为判断跟踪状态的依据,如下所示:

$$E_{\text{APCE}_t} = \frac{|f_{\max,t} - f_{\min,t}|^2}{\text{mean}\left(\sum_{w,h} (f_{w,h,t} - f_{\min,t})^2\right)}, \quad (3)$$

式中, f_t 表示第 t 帧的响应图; $f_{\max,t}$, $f_{\min,t}$ 及 $f_{w,h,t}$ 分别表示第 t 帧响应图的峰值、最小值以及坐标为 (w, h) 位置处的值.

此外,级联检测器中对样本进行分类的依据能够将目标样本和非目标样本区分开,因此也能够对跟踪状态进行判断. 因此本文选取随机森林分类器中的平均后验概率作为判断跟踪状态的依据:

$$p_{\text{mean}}(y|x) = \frac{1}{n} \sum_{i=1}^n p_i(y|x). \quad (4)$$

综上所述,判定跟踪状态的条件为:

$$f_{\max} > \tau_{\max} \&\& E_{\text{APCE}} > \tau_{\text{APCE}} \&\& p_{\text{mean}}(y|x) > \tau_{\text{mean}}, \quad (5)$$

式中, 阈值 $\tau_{\max} = 0.5 \times f_{\text{mean}}$; $\tau_{\text{APCE}} = 0.4 \times E_{\text{APCE mean}}$; $\tau_{\text{mean}} = 0.6 \times p_{\text{allmean}}$; f_{mean} , $E_{\text{APCE mean}}$ 及 p_{allmean} 分别为判断跟踪状态正常帧的最大值历史均值、APCE 值历史均值及平均后验概率值历史均值.

若满足上述判定条件,则判定当前跟踪状态正常,更新检测器中的正负样本及 KCF 滤波器模型. 否则,判定跟踪状态异常,不更新检测器中的正负样本及 KCF 滤波器模型.

3 实验结果及分析

本文实验基于 VS2010 开发平台,使用 Opencv2.4.9 的库编程实现,PC 机配置为 Intel(R) Core(TM) i5-3337U CPU,主频 1.80 GHz,6GB 内存. 为了评估本文算法的跟踪性能,采用文献[18]中提出的精确度(distance precision, DP)和成功率(success rate, SR)两个评价指标对跟踪性能进行分析. 其中,DP 是指中心位置误差小于阈值的帧数占整个视频序列帧数的百分比;SR 是指重叠精度(overlap precision, OP)小于阈值的帧数占整个视频序列帧数的百分比. OP 的计算公式为:

$$S_{\text{op}} = \frac{|r_T \cap r_G|}{|r_T \cup r_G|}, \quad (6)$$

式中, r_T 为跟踪到的目标区域面积; r_G 为目标真实区域面积.

本文评估跟踪器运行速度的指标是每秒帧率(frame per second, FPS),在计算 FPS 时,数据集集中的每

个视频都赋予了权值,以求出不同目标尺寸和不同图像分辨率下的平均 FPS.

算法实验分析选取了整个 OTB-50 数据集的 49 组视频序列进行了两组实验. 实验 1 将本文算法和采用单一 HOG 特征的 KCF 算法、采用 HOG 和 Lab 两种特征的 KCF 算法、采用两种特征并引入尺度估计的 KCF 算法进行对比,验证本文特征融合和尺度适应算法的有效性. 实验 2 将本文算法与 ASMS 算法^[2]、TLD 算法^[4]、KCF_TLD 算法、KCF 算法^[9]以及 ASMS_TLD 算法^[10]进行对比分析,评估跟踪算法的整体性能,并选取 Box、Jogging2、subway 以及 Lemming 4 组视频序列进行分析,评估跟踪算法在遮挡情况下的性能.

实验 1 图 3 是本文算法和采用 HOG 特征的 KCF 算法、采用 HOG 和 Lab 两种特征的改进 KCF 算法、采用两种特征并引入尺度估计的改进 KCF 算法(图中标记为 KCF+)跟踪得到的精确度与成功率曲线图,其中本文算法在曲线图中标记为 Proposed. 从图 3 可以看出,使用多个特征的 KCF 跟踪算法比使用单一特征跟踪的效果更好,引入尺度估计的改进 KCF 算法比未引入的效果更好,且本文算法基于改进 KCF 算法对 TLD 算法进行改进,使跟踪结果优于其他算法.

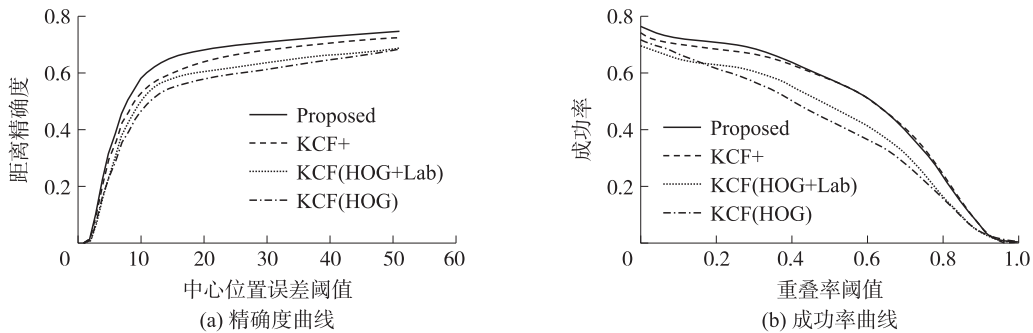


图 3 OTB-50 数据集上不同特征跟踪的精确度与成功率曲线

Fig. 3 The accuracy curve and success rate curve of different feature tracking in OTB-50 dataset

实验 2 图 4 为本文算法与 ASMS 算法^[2]、TLD 算法^[4]、KCF_TLD 算法、KCF 算法^[9]以及 ASMS_TLD 算法^[10]在整个 OTB-50 数据集的精确度与成功率曲线图. 其中本文算法在曲线图中的标记为 Proposed. 从图 4 可以看出,本文算法比其他 5 种算法具有更高的精确度和成功率.

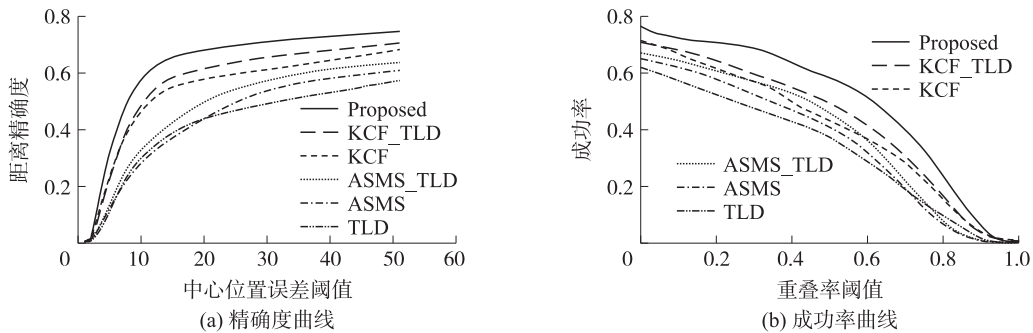


图 4 OTB-50 数据集上 6 种算法的精确度与成功率曲线

Fig. 4 The accuracy curve and success rate curve of six algorithms in OTB-50 dataset

图 5 为本文算法与 ASMS 算法^[2]、TLD 算法^[4]、KCF_TLD 算法、KCF 算法^[9]以及 ASMS_TLD 算法^[10]在 Box、Jogging2、subway 以及 Lemming 4 组目标发生遮挡的视频序列上的精确度与成功率曲线图,其中本文算法在曲线图中标记为 Proposed. 从图 5 可以看出,在遮挡情况下本文算法比其他 5 种算法具有更高的精确度和成功率.

在速度方面,由于本文算法采用 KCF 跟踪器替代 TLD 算法中较耗时的中值光流跟踪器,并采用 Kalman 滤波算法,在不影响跟踪性能的前提下,缩小了检测区域,提高了算法的实时性,使得跟踪速度略高于经典 TLD 算法. 但由于 TLD 算法中的检测器比较耗时,故本文算法的速度与高速的 KCF 算法和 ASMS 算法相比仍存在较大差距,因此提高算法的速度是今后工作的重点. 本文算法与其他 5 种跟踪算法在 OTB-50 数据集的 49 组视频序列中运行的平均速度(mean FPS)如表 2 所示.

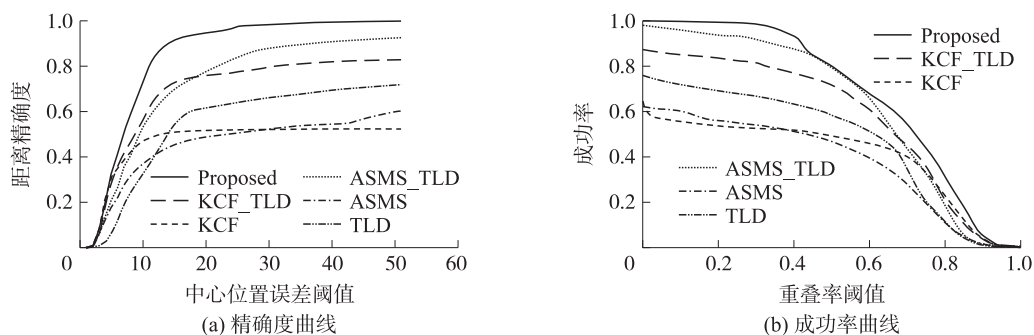


图 5 4 个视频上 6 种算法的精确度与成功率曲线

Fig. 5 The accuracy curve and success rate curve of six algorithms in four sequences

表 2 6 种跟踪算法的平均速度比较

Table 2 Comparison of mean FPS of six tracking algorithms

算法	本文算法	KCF_TLD	ASMS_TLD	TLD	ASMS	KCF
Mean FPS	24.6	24.5	24.6	24.5	125.1	58.35

4 结论

本文在 TLD 算法框架下,提出了一种基于核相关滤波的 TLD 跟踪算法. 该算法对 TLD 算法的跟踪器和检测器两部分进行了改进,提高了算法在发生尺度变化或遮挡情况下的跟踪性能.

本文使用 KCF 跟踪器替代了 TLD 中的光流跟踪器,并对 KCF 跟踪器进行了改进:在特征提取时增加了目标的 Lab 颜色特征,在寻找目标位置时引入尺度估计,在模型更新阶段引入跟踪状态判别机制,不仅使用检测器判定跟踪器的跟踪结果的可靠性,且设定跟踪器中输出响应最大值及 APCE 值的阈值,根据判定结果来选择是否更新滤波器模型. 对于其中的检测器,为了减少大量无意义的窗口,并提升算法在存在遮挡时的精确性,使用 Kalman 滤波预估出目标位置,然后在预估位置周围使用级联分类器更精准地定位目标. 本文改进的级联分类器的前两级保持不变,第三级采用上述改进的 KCF 跟踪器进行分类.

为了验证本文算法的有效性,在 OTB-50 数据集上进行了对比实验. 结果表明,本文算法在发生遮挡、形变和亮度变化的情况下跟踪性能优于其他算法.

[参考文献] (References)

- [1] 卢湖川,李佩霞,王栋. 目标跟踪算法综述[J]. 模式识别与人工智能,2018,31(1):61-76.
- [2] VOJIR T,NOSKOVA J,MATAS J. Robust scale-adaptive mean-shift for tracking[J]. Pattern Recognition Letters,2014,49:250-258.
- [3] ZHANG T Z,XU C S,YANG M H. Learning multi-task correlation particle filters for visual tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2019,41(2):365-378.
- [4] KALAL Z,MIKOLAJCZYK K,MATAS J. Tracking-learning-detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2012,34(7):1409-1422.
- [5] HARE S,GOLODETZ S,SAFFARI A,et al. Struck:structured output tracking with kernels[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2016,38(10):2096-2109.
- [6] LI H,LI Y,PORIKLI F. DeepTrack:learning discriminative feature representations online for robust visual tracking[J]. IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society,2016,25(4):1834-1848.
- [7] CHOI J,KWON J,LEE K M. Real-time visual tracking by deep reinforced decision making[J]. Computer Vision and Image Understanding,2018,171:10-19.
- [8] GORDON D,FARHADI A,FOX D. Re3:real-time recurrent regression networks for visual tracking of generic objects[J]. IEEE Robotics and Automation Letters,2018,3(2):788-795.
- [9] WANG Q,ZHANG L,BERTINETTO L,et al. Fast online object tracking and segmentation:a unifying approach[C]//IEEE Conference on Computer Vision and Pattern Recognition. Los Angeles,USA:IEEE Computer Society,2019:1328-1338.

- [10] 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.
- [11] 张惊雷, 时鹏, 温显斌. 基于尺度自适应均值偏移优化的 TLD 跟踪算法[J]. 控制与决策, 2019, 34(1): 144–150.
- [12] 毛晓波, 周晓东, 刘艳红. 基于 FAST 特征点改进的 TLD 目标跟踪算法[J]. 郑州大学学报(工学版), 2018, 39(2): 1–5, 17.
- [13] 王姣尧, 侯志强, 余旺盛, 等. 采用核相关滤波的快速 TLD 视觉目标跟踪[J]. 中国图象图形学报, 2018, 23(11): 1686–1696.
- [14] 常立博, 杜慧敏, 毛智礼, 等. 面向视频监控的 TLD 改进目标跟踪算法研究[J]. 计算机工程与应用, 2018, 54(14): 191–198.
- [15] SINGH G, CHHABRA I. Effective and fast face recognition system using complementary OC-LBP and HOG feature descriptors with SVM classifier[J]. Journal of Information Technology Research, 2018, 11(1): 91–110.
- [16] BERTINETTO L, VALMADRE J, GOLODETZ S, et al. Staple: complementary learners for real-time tracking[C]//IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE Computer Society, 2016: 1401–1409.
- [17] WANG M, LIU Y, HUANG Z. Large margin object tracking with circulant feature maps[C]//IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE Computer Society, 2017: 4800–4808.
- [18] WU Y, LIM J, YANG M H. Object tracking benchmark[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1834–1848.

[责任编辑: 严海琳]