

基于改进 SSD 的电力检修多目标人员追踪方法^①



沈茂东¹, 高宏¹, 付新阳¹, 周伟¹, 张俊岭², 公凡奎², 冯志珍³

¹(国网山东省电力公司, 济南 250001)

²(山东鲁能软件技术有限公司, 济南 250001)

³(中国石油大学 计算机科学与技术学院, 青岛 266580)

通讯作者: 冯志珍, E-mail: fengzhizhenabc@163.com

摘要: 随着计算机人工智能领域的日益飞速发展, 对摄像头数量要求日益增加, 视频数据量也在迅速增加, 视频下的人形轨迹安全监控跟踪, 是大规模智能监控系统的重要研究方向. 考虑到安全管控现实场景中不同摄像头的光照亮暗程度和每帧图片的人形角度、尺寸等的差异, 会影响人形追踪的准确度, 为此提出具有快速优势的 CSSD 网络 (Correct Single Shot multi-box Detector, CSSD) 和关联分析应用于人形目标追踪. 本文基于行人多目标追踪技术基础, 提出了一种 CSSD 网络来进行模型的检测, 并使用简单的卡尔曼滤波方法跟踪预测目标的位置状态, 预测检测框位置, 使用 IOU 方法和匈牙利算法来解决视频连续前后帧目标匹配问题. 经验证, 该方法可以有效地提高人形目标准确率, 缓解目标之间的部分遮挡和位置突变问题, 能最大程度的适应目标的尺寸、远近和角度改变等.

关键词: 目标追踪; 行人检测; 目标识别; 神经网络; 卡尔曼滤波

引用格式: 沈茂东, 高宏, 付新阳, 周伟, 张俊岭, 公凡奎, 冯志珍. 基于改进 SSD 的电力检修多目标人员追踪方法. 计算机系统应用, 2020, 29(8): 152-157. <http://www.c-s-a.org.cn/1003-3254/7497.html>

Multi-Object Personnel Tracking Method for Electric Power Maintenance Based on Improved SSD

SHEN Mao-Dong¹, GAO Hong¹, FU Xin-Yang¹, ZHOU Wei¹, ZHANG Jun-Ling², GONG Fan-Kui², FENG Zhi-Zhen³

¹(State Grid Shandong Electric Power Company, Jinan 250001, China)

²(Shandong Luneng Software Technology Co. Ltd, Jinan 250001, China)

³(College of Computer Science and Technology, China University of Petroleum, Qingdao 266580, China)

Abstract: With the rapid development of computer artificial intelligence, the number of cameras is increasing, and the amount of video data is also increasing rapidly. The security monitoring and tracking of humanoid trajectory in video is an important research direction of large-scale intelligent monitoring system. Considering that the difference of illumination and darkness of different cameras in different security control scenarios and the human angle and size of each frame will affect the accuracy of human tracking, Correct Single Shot multibox Detector (CSSD) network with advantage of fastness and associated analysis are proposed for human tracking. Based on the pedestrian multi-object tracking technology, this study proposes a CSSD network for model detection, and uses ordinary Kalman filter to track and predict the position of the target, predicts the position of the detection box, and uses IOU method and Hungarian algorithm to solve the problem of video frame target matching before and after. It has been proved that this method can effectively improve the accuracy of humanoid targets, alleviate the large changes caused by epigenetic mutation or partial occlusion, and adapt to the size, distance, and angle changes of targets to the greatest extent.

Key words: target tracking; pedestrian detection; target recognition; neural network; Kalman filtering

① 基金项目: 山东省自然科学基金 (ZR2014FM038, ZR2019MF049)

Foundation item: Natural Science Foundation of Shandong Province (ZR2014FM038, ZR2019MF049)

收稿时间: 2019-12-13; 修改时间: 2020-01-07; 采用时间: 2020-01-14; csa 在线出版时间: 2020-07-29

1 研究背景

电力复杂场景下人员的检测跟踪,对于预防人员违规及保证检修人员的人身安全有极为重要的研究意义。目前流行的单目标追踪主要是 Siamese 系列,在孪生网络的基础上通过增加改进全连接层、区域候选网络、Mask 分支来改进目标检测的准确度。本论文针对多目标人员跟踪现阶段算法进行了深刻的研究,对于电力场景下的人员检测跟踪方法相对较少,本论文针对电力检修特殊场景下进行行人追踪,有效地提高人员目标准确率,缓解目标之间的部分遮挡和位置突变问题。不同的视频跟踪算法,模型的建立与目标的更新方式都不一样。判定模型法对后验概率建模,对物体的跟踪效果较好,但其运算较复杂,判别式算法一般是通过分类器从图像中找出与目标外观最为相似的区域^[1]。摄像头下视频跟踪算法层出不穷,大致可分为生成式和判别式两类。

1.1 生成式跟踪方法

对于生成式跟踪方法,跟踪成为了在目标物体附近搜索最相似的区域。生成式可理解为先提取多目标的特征,学习出代表相应目标的外观模型,通过它依次搜索图像区域进行模式目标匹配,在图像中找到和模型最匹配的区域,即判别为目标^[2-4]。生成式算法大致流程:输入图像序列并初始化目标状态,提取目标特征生成外观模型,目标表示,建立目标模型,通过相邻帧间的时空特性,定位目标最可能出现的区域,在该候选区域提取目标特征,与待跟踪目标进行匹配,输出目标状态^[5-10]。

传统的跟踪算法均属于生成式跟踪算法。卡尔曼滤波跟踪方法是充分利用目标在相邻帧间的空间位置特性对目标进行快速预测定位。粒子滤波跟踪方法是通过非参数化的蒙特卡罗模拟 (Monte Carlo) 方法实现递推贝叶斯滤波,该类方法适用于非线性系统。CamShift 算法是在 MeanShift 基础上改进的可以很好的适应多尺度变换的行人多目标跟踪算法。在目标匹配问题上,当目标匹配不准确时此方法容易引起检测漂移^[10-17]。对于生成模型法,其跟踪更新策略较简单,容易产生物体特征漂移等情况,跟踪效果相对较差。

1.2 判别式跟踪方法

对于判别式跟踪方法,通过训练分类器准确区分检测目标与背景,训练出一个分类器从背景中区分出检测目标物体。目前大部分的深度学习方法普遍使用的分类器主要有 KNN、SVM、AdaBoost 等分类器。近年来,基于相关滤波的跟踪方法由于其运行速度快而颇受欢

迎,相关滤波类方法主要是通过将输入特征回归为高斯分布来对滤波器进行训练,并且在后续的目标跟踪中寻找预测分布中的响应峰值来定位目标的位置。相关滤波器类算法巧妙地应用快速傅里叶变换算法使其速度大幅提升^[18-21]。由于相关滤波算法在单目标跟踪领域的成功应用,提出核相关滤波算法把实数域的数据处理过程转为频率域处理,在处理速度上得到大大提升。

目前多目标运动跟踪算法大多是判别式方法,主要是从提升准确性和跟踪速度两个方面提升目标跟踪。提升准确性的算法有马尔科夫决策模型^[22](MDP)、卷积网络的在线多目标追踪 (AP_RCNN)^[23]、多特征融合的相关滤波追踪 (MAP-RF)^[24]等;提升跟踪速度的算法有交并比 (IOU-tracker)^[20,25](SORT)、深度关联简单在线实时追踪 (DeepSORT)^[26]等。

对于多目标追踪算法,一般准确率和跟踪速度是呈反比的,如果仅利用目标的位置状态信息建立目标运动模型追踪速度相对较快,但实际追踪效果的准确性较差;当考虑外观模型进行追踪处理时,追踪准确性虽会提升,但又会使处理速度相对减慢^[5]。某些方法虽然对网络进行集成,提高了对遗漏和遮挡的鲁棒性,但多目标在实时运动过程中远小近大而产生的尺度大小变化现象并没有实际解决。本论文提出了基于改进 SSD 的电力检修多目标人员跟踪算法,根据目标检测、预测及匈牙利关联算法确定当前帧多目标行人跟踪位置。

2 研究方法

2.1 特征提取网络

电力检修视频下的人员安全监控跟踪,是大规模智能电力检修监控系统的重要研究方向。基于 CSSD 的电力检修多目标人员跟踪算法引入模块的出发点针对现阶段基于摄像头的目标尺寸、远近和角度改变的特点,通过增大感受野来加强网络的特征提取性能,在结构上增加了特征图的融合,注重目标的边缘特征,使提取的特征细粒化,提高了多目标检测的准确率和回归率。网络在前 5 层中加入特征融合模块,前 5 层直接采用 VGG 网络,第 6、7 层由全连接层改为卷积层,利于参数共享,便于计算。Conv1-2 采用 64 个 $3 \times 3 \times 3$ 卷积核,边缘填充 $padding=2$,步长为 $s=1$,得 $300 \times 300 \times 64$;最大池化 $max\ pooling$ 为 2×2 , $s=2$,得 $150 \times 150 \times 64$; Conv2-3, Conv3-4, Conv4-5 同样采用 3×3 卷积,卷积核数量分别为 128、256、512,网络结构各个层级的卷积数量以及参数依次如图 1 所示。

本文基于行人多目标追踪技术为基础,提出了一种 CSSD 网络来进行模型的检测,并使用卡尔曼滤波跟踪预测目标位置状态,基于预测检测框的位置和 IOU 的匈牙利算法来解决视频前后帧目标匹配问题. 本论文利用 CSSD 目标检测网络对图像序列进行特征提取并进行目标检测,在 Conv6(19×19)层之前进行细

化特征映射,该网络主要在 SSD 网络^[27]中引入 1×1 Conv、2×2 pool、4×4 pool 3 个层进行处理,将卷积层得到的特征层进行 2 次池化应用于网络前端的特征映射中,再将得到的特征图融合在一起,将此模块集成到 SSD 网络结构的前端卷积层进行卷积池化训练,在控制计算损失的情况下准确度相对提高.

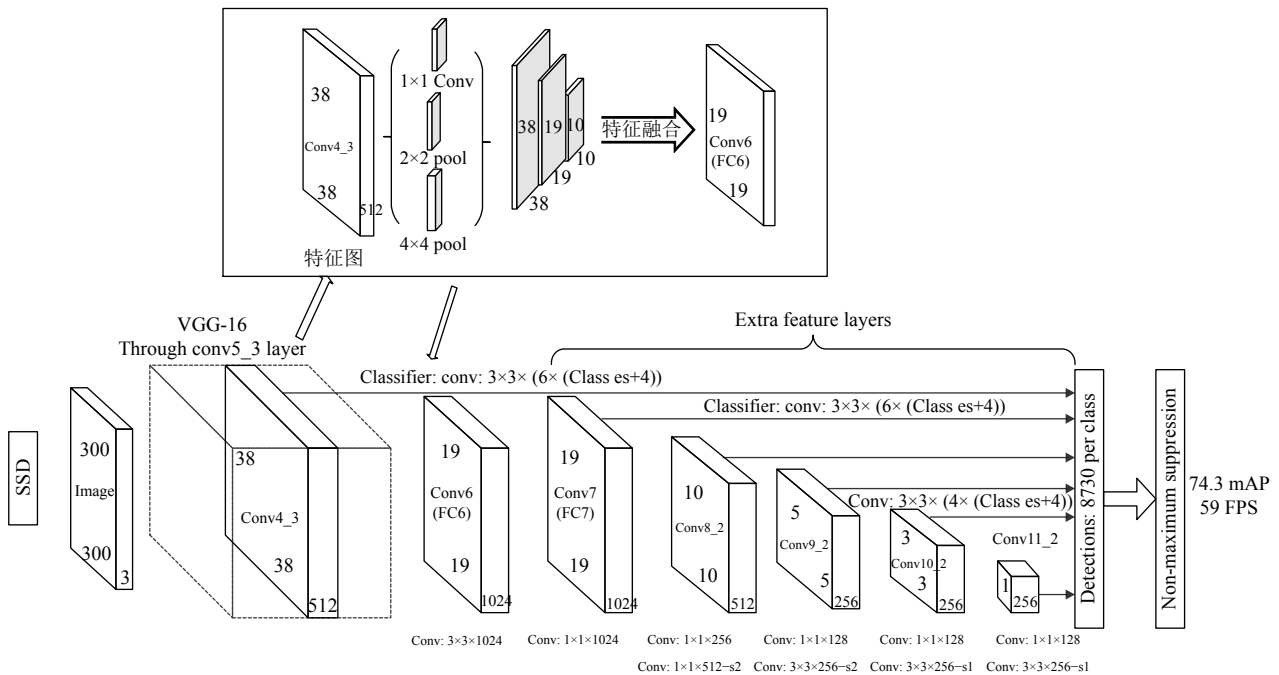


图 1 CSSD 网络结构图

目标定位损失采用如式 (1), 避免梯度爆炸.

$$SmoothL1(x) = \begin{cases} 0.5x^2, & \text{if } |x| < 1 \\ |x| - 0.5, & \text{otherwise} \end{cases} \quad (1)$$

置信度损失采用式 (2) Softmax 损失函数:

$$\begin{cases} L_{conf}(x, c) = - \sum_{i \in Pos} x_{ij}^p \log(\hat{c}_i^p) - \sum_{i \in Neg} \log(\hat{c}_i^0) \\ \text{where, } \hat{c}_i^p = \frac{\exp(\hat{c}_i^p)}{\sum_p \exp(\hat{c}_i^p)} \end{cases} \quad (2)$$

目标定位损失和置信度损失之和, 式 (3) 为 CSSD 网络总损失函数:

$$L(x, c, l, g) = \frac{1}{N} (L_{conf}(x, c) + \partial L_{loc}(x, l, g)) \quad (3)$$

在结构上 CSSD 网络整体上是基于 SSD 网络进行改进, 主要是在 SSD 的基础上加入了卷积池化层, 注重目标的边缘特征, 加强网络的特征提取性能, 使提取的特征细粒化, 使用轻量级网络来实现其高准确性

和高速率, 改进后的网络检测速度比较快, 通过增加网络的感受野加强网络的特征提取能力有效提高对于小目标的检测识别, 同时对于多目标的检测速度和识别精度也有一定的提高. 网络结构类似于金字塔结构, 网络全部采用卷积层, 加入的特征融合部分可以提高网络结构的感受野, 网络特征提取全部采用 3×3 的卷积核, 而且通过特征图的降采样方法, 小的特征图对应小目标, 大大特征图对应大目标, 随着层数的加深, 输出特征图尺寸变小, 加强网络特征提取性能, 最后直接用卷积神经网络来识别检测目标. 轻量级模型部署简单便于优化, 有利于电力检修复杂场景应用.

2.2 视频帧关联分析

基于 CSSD 的电力检修多目标匹配匈牙利算法解决人员多目标跟踪中的前后帧数据关联问题, 求得摄像头视频前后帧目标匹配问题的组合优化算法, 实现人员的最准确匹配问题. 将置信度较高的目标输入匈牙利算法进行匹配, 这样才能得到较好的结果. 为进一

步提高跟踪准确率,跟踪器也使用了多种方法对运动多目标的外观特征状态进行建模。

匈牙利算法用来匹配前后两帧目标进行最大匹配,求出来的匹配不一定是饱和匹配,但却是最优关联匹配。如图2, $t-1$ 帧的 $\text{object}\{1, 2, 3, 4, 5, 6\}$ 和 t 帧 $\text{object}\{1, 2, 3, 4, 5\}$ 的人员匹配,匹配后 object4 离开第 t 帧视频。跟踪人员目标的过程中,因人员的运动速度不是恒定的,即人员目标运动加减速,会产生相应的噪声。针对因噪音产生的检测边界框不准确和检测模型的非线性问题,本文采用卡尔曼滤波原理进行处理,处理完噪音后,再进行预测估计检测框位置,采用简单的卡尔曼滤波原理结合改进的 SSD 网络模块,调整行人检测的边界框,有效提高目标检测框定位的准确性。

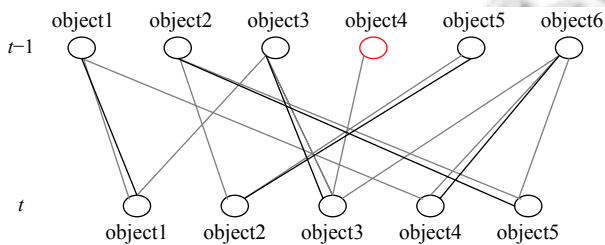


图2 匈牙利算法应用

对于 SORT 算法,直接采用匈牙利算法进行目标关联分析,使用马氏距离计算检测物体和物体跟踪之间的距离。忽略了被检测物体的表面特征,且不能很好

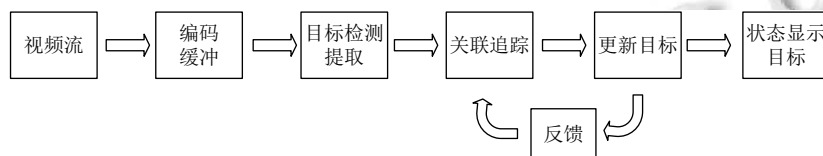


图3 目标追踪整体流程图

3 实验效果

3.1 实验配置

训练模型时使用的软硬件设备为显卡: NVIDIA GeForce GTX 1080TI; CPU 版本 Intel Core i5-6 500 @3.2 GHZ×4; 显存: 16 GB; Ubuntu 16.04 LTS 64 位操作系统; CUDA8.0 版本。

3.2 评价指标

(1) 多目标跟踪准确度 MOTA (Multiple Object Tracking Accuracy):

$$R_{\text{MOTA}} = 1 - \frac{\sum_t (FN_t + FP_t + IDS W_t)}{\sum_t GT_t} \quad (4)$$

的解决长时间被遮掩的物体关联性问题^[20]。在 DeepSORT 中,采用卷积神经网络进行人员大规模数据集提取特征和训练,使用最近邻匹配算法,融合目标运动信息和目标特征信息的测量指标,有效改善了目标追踪过程中的遮掩问题^[25]。基于改进 SSD 的电力检修多目标人员跟踪在使用简单的卡尔曼滤波处理逐帧数据的关联性的基础上使用匈牙利算法进行关联度量,在电力检修高帧速率视频中获得了良好的多目标人员追踪性能。

2.3 算法整体流程

本文中使用了 CSSD 网络来进行模型的检测,并使用卡尔曼滤波方法跟踪预测目标的位置状态,预测检测框位置,使用 IOU 方法和匈牙利算法联合来解决视频连续前后帧目标匹配问题。使得算法具有很高的效率。目标追踪流程如图3所示。

由于 Kalman 滤波的自身特点,使得目标追踪具有很好的准确性和处理数据的快速性,本论文采用简单的卡尔曼滤波原理结合改进的 SSD 网络模块,调整行人检测的边界框,利用 Kalman 滤波建立目标的外观模型,并将其融合到多目标追踪中,针对多个目标建立多个追踪器进行目标训练和检测。用计算目标相似性距离解决目标重识别问题,求目标图像的平均像素,当两个目标平均像素误差小于特定阈值时,认为是同一个目标。对于追踪过程中目标消失问题,本论文设定一个最大丢失时间,当目标丢失时间超过此阈值则认为目标消失,即追踪结束。

MOTA 度量算法中, FN 为 False Negative, FP 为 False Positive, $IDS W$ 为 ID Switch, GT 为 Ground Truth 物体的数量。 FN_t 、 FP_t 、 $IDS W_t$ 分别是缺失数、误判数和误配数。MOTA 考虑了追踪中所有帧中目标位置匹配错误。MOTA 给出了一个非常直观的衡量跟踪器在检测物体和保持轨迹时的性能,与物体位置的估计精度无关。MOTA 以及 MOTP 是计算所有帧的相关指标再进行平均。

(2) 多目标跟踪精度 MOTP (Multiple Object Tracking Precision):

$$R_{\text{MOTP}} = \frac{\sum_{t,i} d_{t,i}}{\sum_t c_t} \quad (5)$$

MOTP 主要量化检测器的定位精度,是关于位置误差的评判指标. $d_{t,i}$ 表示第 t 帧下目标和它配对假设位置之间的距离大小. MOTP 精度高低的的好坏主要取决于度量距离 d 的定义方式,本论文中多目标算法精度越大越好,且不包含与跟踪器实际性能相关的信息. c 为在当前帧匹配成功的数目.

(3) MT (Mostly Tracked): 正确跟踪轨迹占 80% 以上的百分比,在所有跟踪目标中所占的比例.

(4) ML (Mostly Lost): 目标跟丢的轨迹占 20% 以下的百分比,在所有跟踪目标中所占的比例.

(5) Recall: 召回率是指正确匹配的检测目标数占总正例的比重.

(6) Precision: 精确度是指追踪时正确匹配的检测目标数/检测出的总目标数.

(7) FM (Frag Mentation): 每当轨迹将其状态从跟踪状态改变为未跟踪状态,并且在稍后的时间点跟踪相同的轨迹时,就会对 FM 进行计数. 跳变数是指跟踪轨迹从“跟踪”到“不跟踪”状态的变化数. FM 计算的是跟踪有多少次被打断(即 Ground Truth 的 track 没有被匹配上).

3.3 算法性能评估

本文对帧差法、GMM、ViBe、SORT、DeepSORT 等算法在 MOT2017 数据集上进行了测试,测试结果如下,经过试验对比表明基于 CSSD 网络的人形多目标追踪算法在追踪准确率上表现最好,如表 1.

表 1 各算法性能对比

算法	MOTA↑	MOTP↑	MT↑(%)	ML↓(%)	Recall↑	Precision↑	FM↓
帧差法	48.7	73.3	25.3	20.6	0.81	0.72	1873
GMM	49.9	75.6	30.8	24.7	0.28	0.94	1680
ViBe	52.4	76.5	27.5	22.8	0.81	0.97	1832
SORT	59.8	78.6	25.4	22.7	0.83	0.95	2343
DeepSORT	61.4	78.1	32.8	18.2	0.89	0.92	1798
本文	62.3	78.4	28.4	17.6	0.86	0.96	1743

对比了其它方法的差距,基于 CSSD 网络的多目标追踪方法在公开行人视频数据集上测试取得了优化,特别是在 MOTA 和 MOTP 测试上算法满足可靠鲁棒性. 经实验测试表明,本方法在电力设备场景下取得了人员跟踪检测在准确率和速度两方面最优,跟踪时可有效处理视频帧中人员遮挡问题. 以下是 SORT、Deep SORT 及 CSSD 算法截取连续帧的实验效果图,如图 4.



图 4 算法对比效果图

图 4 中绿色代表 ViBe 算法的追踪;紫色代表 DeepSORT 算法的追踪;红色是本文方法,图 4(a)、图 4(b) 为 MOT2017 数据集上的测试结果,图 4(c)、图 4(d) 为电力现场数据集上的测试结果. 实验分析表明本论文

算法目标跟踪平均速度 22 fps 接近实时,能够有效缓解追踪时因为表观突变或者部分遮挡导致的准确率下降问题,小目标跟踪精度和准确率都大大提升,能最大程度的适应目标的尺寸、远近和角度改变等.

4 结论与展望

本文基于改进 SSD 网络的电力检修多目标人员跟踪算法, 构建目标追踪外观模型, 利用图形检测建模, 构建目标追踪的运动模型, 利用位置信息追踪, 结合空间和时间多策略的匹配对电力检修特定场景的目标追踪有了很大提升. 摄像头采集到大量视频数据, 主要任务为检测图片中是否存在人员目标, 可在各种复杂场景中快速实时的检测到图片中的工作人员, 是一种鲁棒性的目标检测追踪算法. 本论文结合实际项目需要, 具体项目已部署于电力场地, 为促进国民企业发展提供强有力的理论依据和技术支撑.

参考文献

- 刘鹏举. 基于判别学习的单目标跟踪系统研究与实现 [硕士学位论文]. 哈尔滨: 哈尔滨工业大学, 2016.
- 王守义, 周海英, 杨阳. 基于卷积特征的核相关自适应目标跟踪. 中国图象图形学报, 2017, 22(9): 1230–1239. [doi: 10.11834/jig.170009]
- 朱文青, 刘艳, 卞乐, 等. 基于生成式模型的目标跟踪方法综述. 微处理机, 2017, 38(1): 41–47. [doi: 10.3969/j.issn.1002-2279.2017.01.011]
- 陆星家, 郭璘, 陈志荣, 等. 基于外观和运动的车辆检测和追踪算法研究. 计算机工程, 2014, 40(8): 152–157. [doi: 10.3969/j.issn.1000-3428.2014.08.029]
- 刘欢, 李春庚, 安居白, 等. 基于核相关滤波的多目标追踪. 激光与光电子学进展, 2019, 56(12): 154–161.
- 王保宪. 复杂背景下的视频目标跟踪算法研究 [博士学位论文]. 北京: 北京理工大学, 2016.
- 李冰寒, 酒锐波, 刘玉婷. 基于卡尔曼预测的快速目标跟踪算法. 中国新通信, 2019, 21(8): 130. [doi: 10.3969/j.issn.1673-4866.2019.08.104]
- 樊庆宇. 稀疏表示在目标追踪中的应用研究 [硕士学位论文]. 成都: 电子科技大学, 2017.
- 梅立雪, 汪兆栋, 张浦哲. 一种邻帧匹配与卡尔曼滤波相结合的多目标跟踪算法. 科学技术与工程, 2019, 19(15): 204–210. [doi: 10.3969/j.issn.1671-1815.2019.15.033]
- 曹凯悦. 基于多 Agent 的多摄像头目标轨迹追踪 [硕士学位论文]. 北京: 北京交通大学, 2018.
- 陈小娟, 余二永. 基于粒子滤波和 Mean-shift 的自适应目标跟踪方法. 科学技术与工程, 2013, 13(33): 10013–10016, 10031. [doi: 10.3969/j.issn.1671-1815.2013.33.041]
- 林庆, 徐小丁, 廖定安, 等. 稀疏表示因子模糊粒子滤波的目标追踪. 小型微型计算机系统, 2014, 35(1): 181–184. [doi: 10.3969/j.issn.1000-1220.2014.01.037]
- 岳昊恩, 袁亮, 吕凯. 结合帧差法与 Mean Shift 的抗遮挡跟踪算法. 现代电子技术, 2019, 42(12): 180–182, 186.
- Wojke N, Bewley A, Paulus D. Simple online and realtime tracking with a deep association metric. Proceedings of 2017 IEEE International Conference on Image Processing. Beijing, China. 2017.3645–3649.
- Fan ZY, Weng SQ, Jiang J, et al. Particle filter object tracking algorithm based on sparse representation and nonlinear resampling. Journal of Beijing Institute of Technology, 2018, 27(1): 51–57.
- Tian MC, Bo YM, Chen ZM, et al. Multi-target tracking method based on improved firefly algorithm optimized particle filter. Neurocomputing, 2019, 359: 438–448. [doi: 10.1016/j.neucom.2019.06.003]
- Jakob EM, Long SM, Harland DP, et al. Lateral eyes direct principal eyes as jumping spiders track objects. Current Biology, 2018, 28(18): R1092–R1093. [doi: 10.1016/j.cub.2018.07.065]
- 汤文. 复杂背景下红外运动小目标的检测与跟踪技术研究 [硕士学位论文]. 长沙: 国防科学技术大学, 2016.
- 娄涵. 融合时空信息的连续相关滤波用于目标跟踪 [硕士学位论文]. 北京: 北京邮电大学, 2018.
- Zhou HY, Yang Y, Wang SY. Multiple object tracking algorithm based on kernel correlation filter. Laser & Optoelectronics Progress, 2018, 55(9): 091502.
- 贺潇. 基于深度学习检测与核相关滤波器的追踪算法 [硕士学位论文]. 北京: 北京邮电大学, 2018.
- Xiang Y, Alahi A, Savarese S. Learning to track: online multi-object tracking by decision making. Proceedings of 2015 IEEE International Conference on Computer Vision. Santiago, Chile. 2015.4705–4713.
- Chen L, Ai HZ, Shang C, et al. Online multi-object tracking with convolutional neural networks. Proceedings of 2017 IEEE International Conference on Image Processing. Beijing, China. 2017.645–649.
- 陈智, 柳培忠, 骆炎民, 等. 自适应特征融合的多尺度相关滤波目标跟踪算法. 计算机辅助设计与图形学学报, 2018, 30(11): 2063–2073.
- Bochinski E, Eiselein V, Sikora T. High-speed tracking-by-detection without using image information. Proceedings of the 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance. Lecce, Italy. 2017.1–6.
- Bewley A, Ge ZY, Ott L, et al. Simple online and realtime tracking. Proceedings of 2016 IEEE International Conference on Image Processing. Phoenix, AZ, USA. 2016.3464–3468.
- Liu W, Anguelov D, Erhan D, et al. SSD: Single shot MultiBox detector. Proceedings of the 14th European Conference on Computer Vision. Amsterdam, the Netherland. 2016.21–37.