

文章编号: 2095-2163(2019)04-0263-05

中图分类号: TP391

文献标志码: A

结合 R-DAD 和 KCF 的行人目标跟踪改进算法

许喜斌

(广东工程职业技术学院, 广州 510520)

摘要: 行人目标跟踪技术在现实生活中具有较强应用性,但存在抗遮挡性弱等问题。本文利用 Jetson TK1 平台的便捷性结合 R-DAD(Region Decomposition and Assembly Detector)思想改进了 KCF(Kernelized Correlation Filters)跟踪算法,实现对行人目标跟踪。实验结果表明,改进的行人目标跟踪算法运算量少、性能高、速度快,具有良好的抗遮挡性和精确性。

关键词: Jetson TK1 平台; R-DAD; KCF; 目标跟踪; 抗遮挡性

Improved pedestrian target tracking algorithm combined with R-DAD and KCF

XU Xibin

(Guangdong Engineering Polytechnic, Guangzhou 510520, China)

[Abstract] Pedestrian target tracking technology has strong applicability in real life, but there are problems such as weak anti-blocking. Therefore, using the convenience of the Jetson TK1 platform combined with the R-DAD (Region Decomposition and Assembly Detector) idea, the KCF (Kernelized Correlation Filters) tracking algorithm is improved to achieve the tracking of pedestrian targets. Experiments show that the improved pedestrian target tracking system not only has less computation, high performance, fast speed, but also has good anti-blocking and accuracy.

[Key words] JetsonTK1 platform; R-DAD; KCF; target tracking; anti-blocking

0 引言

随着人工智能的快速发展,图像检测识别和目标跟踪技术在多个领域被广泛使用,例如商业、小区安全、军事保密和法律取证等方面的智能视频监控;基于 5G 网络的智能医疗诊断;解决交通道路拥堵的智能交通监管系统及自动驾驶系统。

目标检测算法主要有基于背景建模、基于轮廓模板、基于光流法和基于底层特征等 4 类。基于背景建模算法是先建立当前前景的背景模型,再与背景图像模型做差分,提取运动前景,建立临时块模型,结合更新的行人检测的背景模型实现行人检测^[1],该算法对动态变化背景较为敏感,不适用复杂的动态场景。基于轮廓模板的方法是通过构建图像目标物体的边缘轮廓、灰度、纹理等信息模板,再匹配模板的方法来检测目标^[2],该方法简单易行,但需要构建大量的模板才能取得较好的效果。基于光流法是赋予图像各个像素点一个速度矢量形成图像运动场,再根据像素点的速度矢量特征进行图像分析,获取物体在图像中的运动位置,实现目标检测^[3],该方法适合分析多目标的运动,可以解决运动目标检测中的遮挡、重合等问题,但存在实时性差、鲁棒性弱、计算方法复杂等不足。基于底层特征

的算法是利用积分图技术进行快速计算,但对行人目标的表达能力不足,判别力较差。

行人检测技术主要有基于特征提取和深度学习(Deep Learning)2 种。其中对于特征提取的方法,为了更好地获取行人图像特征,Dalal N 等人^[4]采用梯度方向直方图(HOG)特征描述子结合线性分类器支持向量机(SVM)的方法,在行人检测方面取得了巨大的突破,但在处理遮挡问题的时候,HOG 特征并不能很有效地解决遮挡问题。基于深度学习方面,随着近年来的理论和相关技术的高速发展,深度学习的行人检测算法取得很大的突破^[5-8],其检测精度也有较大地提高,深度学习具备的自主学习特点可以避免复杂的特征提取和数据建模过程,很好地描述检测目标特性。本文基于 KCF 算法,结合区域分解集成的思想,改进 KCF 算法提高目标跟踪的抗遮挡性,在 NVIDIA Jetson TK1 平台上具有良好的抗遮挡性和实时性。

1 Jetson TK1 平台

NVIDIA Jetson TK1 平台是 NVIDIA 针对快速开发和部署面向机器人技术、计算机视觉、医疗等领域的计算密集型系统。该平台以 NVIDIA Tegra® K1 SoC 为基础构建,包含 NVIDIA Kepler™ 计算核心,

作者简介: 许喜斌(1987-),男,硕士,高级工程师,主要研究方向:计算机视觉、智能嵌入式处理。

收稿日期: 2019-05-08

且具有 192 个 CUDA 核心,非常适合算法的并行运算,具体参数见表 1。

表 1 Jetson TK1 主要参数

Tab. 1 Jetson TK1 main parameters

平台型号	NVIDIA JetsonTK1
处理器	四核 2.3GHz ARM Cortex-A15
GPU	Tegra K1
内存	2GB
操作系统	Ubuntu14.04.5LTS
摄像头	300 万像素(1920x1080,USB)

Jetson TK1 平台使得计算机视觉、深度学习的计算更加高性能、低能耗,是嵌入式系统设计的理想平台,如图 1 所示。



图 1 Jetson TK1 开发平台

Fig. 1 Jetson TK1 development platform

2 跟踪算法

2.1 KCF 算法

KCF(Kernelized Correlation Filters),即核相关滤波器方法。首先,采用循环偏移图像矩阵,构建出分类器的训练样本,其中中间部分为正样本,其它周边为负样本,使数据矩阵转换成循环矩阵,然后,简化循环矩阵计算,进而求解基于循环矩阵特性的问题。该方法建立在相关滤波跟踪算法的框架上,将 FFT 快速傅里叶变换转换到频域,从而避免矩阵求逆过程,极大地降低算法复杂度和提高跟踪精度^[9]。下面给出 KCF 算法的具体过程:

2.1.1 训练分类器

KCF 算法的实质就是求解岭回归问题,其中目标函数为: $f(z) = w^T z$; 岭回归函数为: $f(x_i) = w^T \varphi(x_i)$; 假定映射函数为: $\varphi(x)$, 可得权重系数为:

$$w = \min_w \|\varphi(x)w - y\|^2 + \lambda \|w\|^2, \quad (1)$$

用样本线性组合表示 w 为: $w = \sum_i \alpha_i \varphi(x_i)$, 故式(1)为:

$$\alpha = \min_{\alpha} \|\varphi(x) \varphi(x)^T \alpha - y\|^2 + \lambda \|\varphi(x)^T \alpha\|^2, \quad (2)$$

对式(2)进行求导,且令 α 导数为 0,可得:

$$\alpha = (\varphi(x) \varphi(x)^T + \lambda I)^{-1} y, \quad (3)$$

其中, $\varphi(x) \varphi(x)^T$ 为核矩阵,用 K 表示核空间的核矩阵,即 $K = \varphi(x) \varphi(x)^T$, 则 $\alpha = (K + \lambda I)^{-1} y$, 由于 K 为循环矩阵,故 $K = C(k^{xx}) = F \text{diag}(\hat{k}^{xx}) F^H$, 可得:

$$\hat{\alpha} = \frac{\hat{y}}{\hat{k}^{xx} + \lambda}. \quad (4)$$

2.1.2 目标检测

目标检测的计算响应公式为:

$$f(z) = w^T z = \sum_{i=1}^n \alpha_i k(z, x_i), \quad (5)$$

其中, $k(z, x_i)$ 为核函数,令候选图像块和训练样本之间的核矩阵为 K^Z , 由文献[10]中的定理 1 可知 K^Z 为循环矩阵,即 $K^Z = C(k^{xz})$, 根据循环矩阵特性^[11], 将 $K = F \text{diag}(\hat{k}^{xx}) F^H$ 带入式子(5)中,可得:

$$\hat{f}(z) = \hat{k}^{xz} \odot \hat{\alpha}. \quad (6)$$

KCF 算法目标跟踪流程如图 2 所示。

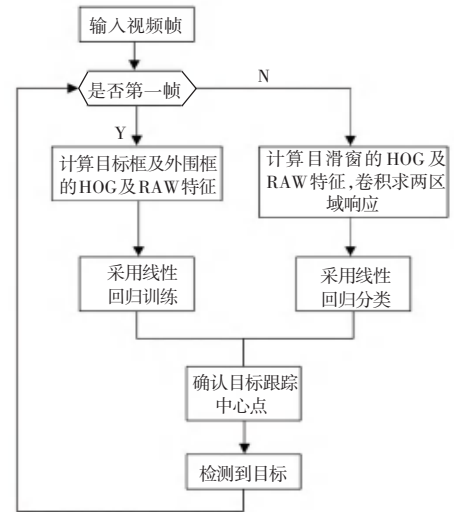


图 2 KCF 算法目标跟踪流程图

Fig. 2 KCF algorithm target tracking flowchart

2.2 R-DAD 思想

目前主流的目标检测算法中以 Faster-RCNN 为代表,由 RPN(Region Proposals Network)生成感兴趣区域 RoI(Region of Interests),该算法会因目标被遮挡或不精确的候选区域(Region Proposals)导致目标检测不准确,因此,Seung-Hwan Bae 提出 R-DAD(Region Decomposition and Assembly Detector),即区域分解组装检测器,来改善生成的候选区域^[12]。图 3 示出了 R-DAD 架构。

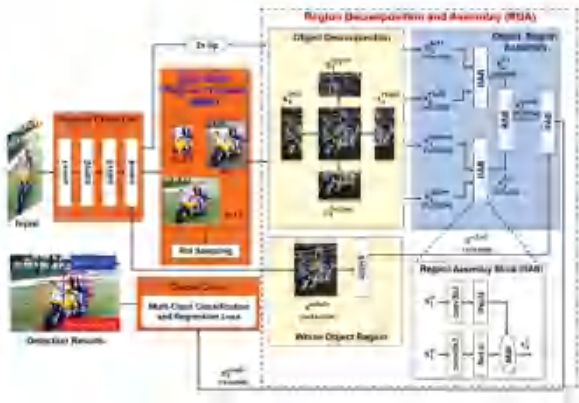


图 3 R-DAD 架构

Fig. 3 R-DAD architecture

在 R-DAD 架构中,重新对 MRP 网络进行调整,对目标进行缩放形成多个不同尺度的候选区域,再对其分解成几个部分,如上、下、左、右,其中,设计了一个带有 3×3 的卷积滤波器区域组装块,让分解的部分两两进行卷积和使用 ReLU 函数,比较得出最大单元模块,再继续使用 RAB (Region Assembly Block),接着对分解部分的强响应进行结合,然后再学习整个对象和基于部件的特征之间的语义关系, R-DAD 网络架构主要分成 MRP 和 RDA 2 个模块。

2.2.1 MRP (Multi-Scale Region Proposal) 模块

MRP 模块主要是改善 RPN 生成的候选区域的准确率。是生成多尺度的候选区域,首先用传统的 RPN 生成一些建议框,然后用不同的缩放因子对生成的候选区域进行不同比例的缩小放大,从而提高候选区域的多样性。如图 4 所示,框内分别对应 $S = 0.7, 1, 1.2$ 的候选区域,不同尺度的区域,会导致部分局部大于目标本身,也使得候选区域数量过多,导致无法完全利用,因此需要添加 RoI 的采样层,对分数低和跟 ground truth 重叠率低的进行筛选。因此,MRP 网络生成的各种候选区域,可以进一步适应目标之间因为空间变化所导致的特征变化,进而提高结构的鲁棒性。



图 4 MRP 模块

Fig. 4 MRP module

2.2.2 RDA (Region Decomposition and Assembly) 模块

RDA 模块同时描述物体的全局及局部外观,将目标物体分为目标区域集成和目标分解 2 部分,其中,目标分解如图 5 所示,将目标分为上、下、左、右 4 个部分,其中图左边为不同尺度的候选区域,图右边为物体目标的分解区域。



图 5 RDA 模块

Fig. 5 RDA module

对目标进行分解时,先用线性插值 2 倍以上采样之后再分解,会达到更好的效果。因为左右刚好是特征图的左右一半,上下也是一样,然后再送入 RAB 模块,其中 RAB 模块如图 6 所示。

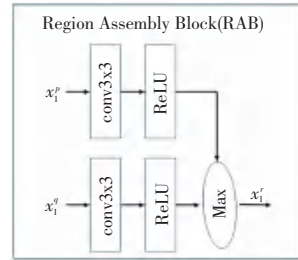


图 6 RAB 模块

Fig. 6 RAB module

RDA 模块主要由以下 2 个函数表示:

$$X_{j,l}^p = f\left(\sum_{i=1}^{k_l} X_{i,l-1}^p * W_{ij}^l + b_j^l\right), l = 2, 3, 4, \quad (7)$$

$$X_l^r = \max(X_l^p, X_l^q). \quad (8)$$

其中, p 为上下左右的一部分或组合后的部分,如左 - 右 (l/r)、下 - 上 (b/u)、 $comb(l/r$ 和 b/u 的组合); $*$ 是卷积操作; $f()$ 是 ReLU 单元。为了融合 X_l^p 和 X_l^q 的信息,生成同样大小的 X_l^r ,需要再取两者的 \max ; X^{whole} 是全局信息 scale 为 1 生成的候选区域,然后再继续使用 RAB 模块,这样就可以捕捉到局部信息,且兼顾全局信息。RAB 模块是类似 maxout 单元,可以逼近任何连续的函数,因此,可以配置不同的分层地组合 RAB 模块来表示各种目标的特征。

R-DAD 可以描述全局特征和局部特征的语义信息,在较为复杂的场景下,若目标对象被遮挡,通过左、右、上、下模板筛选出来的特征会更加符合真实场景,使得候选区域更加可信,增强目标跟踪的抗遮挡性。

2.3 改进 KCF 算法

KCF 算法在各种场景的应用中,在跟踪效果和

跟踪速度上都表现的比较优秀。但是,该算法存在无法较好解决跟踪过程中目标被遮挡的问题,一旦跟踪目标发生遮挡就容易导致跟踪失败。

本文结合 R-DAD 的思想,对行人目标区域进行全局和局部特征提取,并进行训练,获得行人目标的局部特征,再对局部区域使用 KCF 跟踪器进行初始化并输出响应值,利用响应值定位目标局部区域的关联响应,进而将部分区域响应值组合形成整体的行人目标响应值。局部特征提取如图 7 所示。

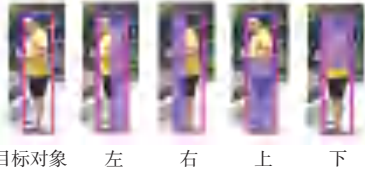


图 7 局部特征提取

Fig. 7 Local feature extraction

基于 R-DAD 思想的 KCF 目标跟踪改进算法描述如下:

步骤 1 先将行人目标整体检测出来,然后对行人整体目标进行划分,例如,假设整体行人目标为 X ,将其等分为上、下、左、右 4 部分: X^l 、 X^r 、 X^u 、 X^d ;

步骤 2 对每部分都进行特征提取,并形成相对应的特征值: E^l 、 E^r 、 E^u 、 E^d ;

步骤 3 将整体目标跟踪化为局部跟踪,先跟踪局部部分,即先跟踪 E^l ,以 E^l 特征为目标 X^l 作为行人目标所在的跟踪对象进行跟踪;

步骤 4 当局部目标 X^l 被遮挡时,跟踪另一个局部目标局部块 X^r ,使用 E^r 作为目标跟踪特征进行跟踪,重复步骤 3;

步骤 5 当局部目标特征失效或整体目标被完全遮挡,则跟踪失败;

步骤 6 算法结束。

3 实验结果

为了提高识别的速度,提升实验效果,本文在 JetsonTK1 平台上调用 GPU 进行行人识别运算,对同一段行人跟踪视频分别用 KCF 算法及改进的 KCF 算法进行目标跟踪,对比两者的召回率和准确率,分析两者之间的性能区别,见表 2。

3.1 召回率

$$R = \frac{TP}{TP + TN} \quad (9)$$

其中, R 为召回率; TP 为目标跟踪正确的图像帧总数; TN 为含有目标的图像帧及跟踪失败的总

数。召回率可以评估目标跟踪算法的全面性衡量及跟踪质量,进而反映出算法在所有目标出现的图像中的成功跟踪目标的能力。

3.2 准确率

$$P = \frac{TP}{TP + FP} \quad (10)$$

其中, P 为准确率; TP 为目标跟踪正确的图像帧总数; FP 为目标跟踪错误及未出现目标的图像帧总数。准确率可以衡量目标跟踪的准确性。

见表 2, KCF 目标跟踪算法的准确率为 28.55%,改进后的 KCF 目标跟踪算法的准确率为 48.69%,改进后的 KCF 算法准确率提升 20.14%。KCF 算法召回率为 28.11%,改进后的 KCF 算法召回率为 46.67%,改进后的 KCF 算法召回率提升 18.56%。实验结果表明,基于 R-DAD 思想改进的 KCF 目标跟踪算法的跟踪性能更佳、抗遮挡性更强、跟踪效果更好。

表 2 基于 Jetson TK1 平台 KCF 算法改进前后跟踪效果对比

Tab. 2 Comparison of tracking effects before and after improvement of KCF algorithm based on Jetson TK1 platform

方法	准确率/%	召回率/%
JetsonTK1+KCF	28.55	28.11
JetsonTK1+改进 KCF	48.69	46.67

4 结束语

本文基于 JetsonTK1 平台,利用该平台性能高、便捷、小巧的特点,设计了行人目标跟踪算法,采用 KCF 跟踪算法,并利用 R-DAD 思想对 KCF 算法进行改进,在保持 KCF 算法高正确率、高运算速度、少计算量的同时,提升了行人目标跟踪算法的抗遮挡性和跟踪精度。

参考文献

- [1] WREN C, AZARBAYEJANI A, DARRELL T, et al. Real-time tracking of the human Body [C]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7): 781-785.
- [2] 张春风,宋加涛,王万良.行人检测技术研究综述[J].电视技术, 2014, 38(3): 157-162.
- [3] Yalin X, Steven A S. Moment and Hyper Geometric Filter for High Precision Computation of Focus, Stereo and Optical Flow [J]. International Journal of Computer Vision, 1997, 22(1): 25-29.
- [4] Dalal N, Triggs B. Histograms of oriented gradients for human detection [C] // 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005, IEEE, 2005: 886-893.