

基于图论的复杂交通环境下车辆检测方法

苏 帅^{1,2},袁 雪¹,张立平²,李寒松²

(1.北京交通大学 电子信息工程学院,北京 100044;2.北京华航无线电测量研究所,北京 100013)

摘 要:目前现有的基于图像的车辆检测系统大多数是利用滑动窗口法来确定车辆候选区域.为了提高车辆检测的速度并减少计算量,提出了一种新的基于图论的车辆检测方法.该方法针对每幅图像通过简单线性迭代聚类(SLIC)算法得到含有若干个超像素节点的图像,分析节点间的相互关系最终确定车辆候选区域.在检测阶段,本文把大量不同视角的车辆图片作为正样本进行训练,得到多视角的分类器;基于候选区域的几何信息,选择适当的多视角分类器进行检测.由公共交通分析数据集(KITTI)检测结果表明:与目前最新的、具有相同提取特征和分类器的算法相比,本文的方法具有更好的检测精度,在复杂的背景下也能取得很好的检测结果.

关键词:信息处理;车辆检测;车辆候选区域;多视角分类器

中图分类号:TP391 文献标志码:A

Graph-based vehicle detection in complex traffic environment

SU Shuai^{1,2}, YUAN Xue¹, ZHANG Liping², LI Hansong²

(1.School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, China;

2.Beijing Huahang Radio Measurement Institute, Beijing 100013, China)

Abstract: The majority of the existing graph-based vehicle-detection systems make use of sliding-window paradigm for vehicle-candidate regions location. In order to improve the speed of vehicle detection and reduce the computational complexity, a new vehicle detection method based on graph theory is proposed in this paper. The algorithm uses Simple Linear Iterative Clustering (SLIC) algorithm to obtain images with several super-pixel nodes for each image, and analyzes the relationship among the nodes to determine the vehicle candidate region finally. In the detection stage, multi-view detectors are established by training the vehicle images which seen as the positive samples and collected on each distinct view. Based on the geometrical information of the bounding boxes, the suitable viewpoint detectors are selected from the multi-view detectors. The results of the public traffic analysis dataset (KITTI) show that the proposed approach leads to better performances when compared with the current state-of-the-art methods with the same feature extraction and classifier algorithms. Moreover, it can also yield better results under the complex background.

Keywords: information processing; vehicle detection; vehicle candidate location; multi-view classifiers

收稿日期:2016-11-15
基金项目:国家自然科学基金项目(61301186, 61673047);北京市科委重大研究专项(SX2016-04)
Foundation items: National Natural Science Foundation of China(61301186, 61673047); Beijing Municipal Science & Technology Commission Major Program(SX2016-04)
第一作者:苏帅(1990—),男,河南开封人,硕士.研究方向为图像处理和模式识别. email:14120280@bjtu.edu.cn.
通信作者:袁雪(1977—),女,辽宁沈阳人,副教授,博士. email:xyuan@bjtu.edu.cn.
引用格式:苏帅,袁雪,张立平,等. 基于图论的复杂交通环境下车辆检测方法[J].北京交通大学学报, 2017, 41(5):66-72.
SU Shuai, YUAN Xue, ZHANG Liping, et al. Graph-based vehicle detection in complex traffic environment [J]. Journal of Beijing Jiaotong University, 2017, 41(5):66-72. (in Chinese)

车辆检测是自动驾驶系统和辅助驾驶系统的重要组成部分。自动驾驶系统通过车辆检测系统对当前道路信息的收集,获取车辆周边其他车辆的信息,从而做出车速调节、车距保持和并线等各种操作。辅助驾驶系统的应用可以极大降低交通事故的发生率,使交通运输更加高效。现代车辆检测技术主要分类两类:基于传感器的车辆检测和基于计算机视觉的车辆检测。比较流行的基于传感器的车辆检测主要用到了雷达(如毫米波雷达)、激光(如激光雷达)及声音传感器。然而,基于传感器的车辆检测还存在一些缺点,如低空间分辨率、扫描速度较慢和较高的成本。

目前,虽然基于计算机视觉的车辆检测系统受到了越来越多的关注,但其仍存在着一些具有挑战性的问题。在实际应用中,基于计算机视觉的车辆检测系统还需要考虑到光照变化、部分遮挡检测、相机角度变化和尺寸变化等情况。在城市道路驾驶过程中,频繁的遮挡,伴随着各种车辆方向和尺寸的变化,使得车辆检测更加困难。此外,由于车载摄像机的高速移动,车辆检测系统需要实时地将检测结果输出。因此,提高车辆检测系统的检测速率和精度成为主要问题。

计算机视觉检测中的光学相机往往能更有效地解决这些问题如下:1)检测范围大。光学相机能提供比传感器更高的分辨率和更广泛的视角;2)视觉检测立体化。在大量相关应用中,视觉信息是非常重要的。如交通标志的识别,以及车道线、行人和障碍物

的检测;3)安装调试方便。光学相机的安装不需要修改任何道路基础设施。

基于图论的图像分割方法是近年来的研究新热点之一,许多学者都将研究重点立足与如何准确快速地定位目标区域,以便后续的识别或其他处理。现存在的图像检测算法非常多,文献[1]提出了基于图论的显著性区域检测的算法,将图像每一个像素看作一个节点,网络图对应为一个马尔科夫链,并将马尔科夫链达到稳态时的特征向量作为测量视觉显著性的指标。文献[2]提出了一种基于图论的等级排序方法用于显著性区域的检测,根据图中各节点与图像边界的相似性及节点之间的连续性建立目标函数,目标函数最小化时各节点的解与各节点的显著性程度相对应。当图像较大时,以像素为单位计算往往耗时较多,计算复杂度相应增加^[3]。

本文作者提出了一种新的基于图论的车辆检测方法来提高检测速率和精度。1)提出了自适应边界框生成算法来确定车辆候选区域,可以得到更少和更准确的搜索窗口进行检测;2)提出了一种基于图论的算法计算边界框窗口中车辆可能性,由数值表示。用非极大值抑制(Non-Maximum Suppression, NMS)算法将这些带有车辆可能分数的边界框进一步筛选;3)根据每个边界框的几何信息,选择合适的多视角分类器进行检测。

1 自适应边界框生成算法

本文提出的车辆检测流程图如图 1 所示。

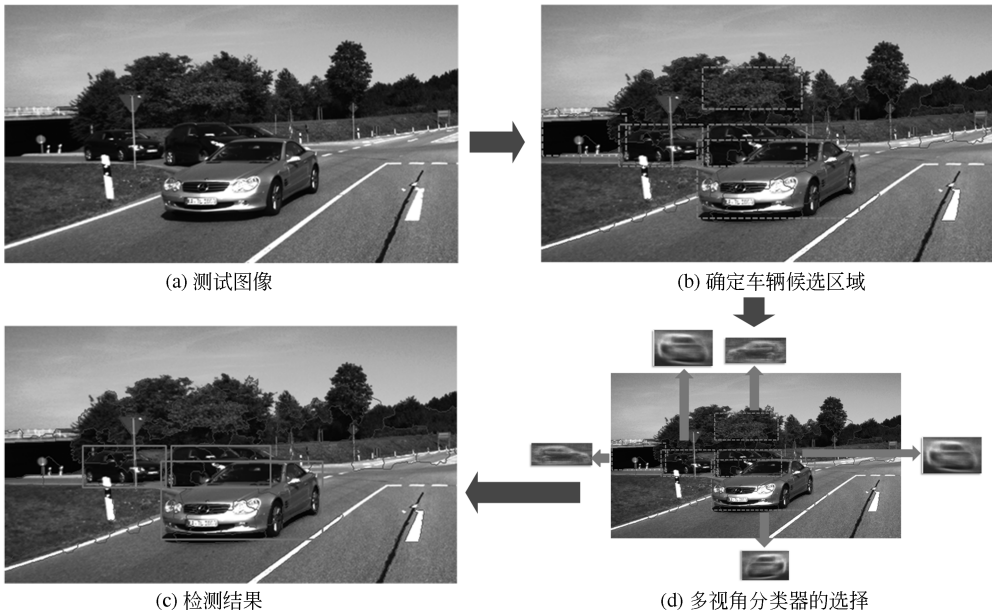


图 1 本文车辆检测算法流程图

Fig.1 Flow chart of vehicle detection algorithm in this paper

目前现有的基于图像的车辆检测系统大多数是利用滑动窗口法来确定车辆候选区域的.滑动窗口法用不同尺寸和形状的窗口遍历每幅图像,模板窗口的形状(宽与高的比例)会随着车辆的形状改变,从而生成大量的车辆候选区域.如对于一个大小为 $1\,242 \times 375$ 像素的图像,当图像变换尺度的种类设为 36,模板窗口的宽高比例种类设为 25,步长设为 1 像素时,将会产生超过 3×10^6 个车辆候选区域.滑动窗口法产生的大量候选区域将会导致巨大的检测运算.因此,本文提出了一种产生更少和更精确自适应边界框生成方法.

1.1 图像分割与融合

本文通过简单线性迭代聚类(SLIC)算法^[4]和基于密度的聚类(Density-based Spatial Clustering of Applications with Noise, DBSCAN)算法^[5]对每幅图像处理得到超像素图(Superpixels).超像素图的一个重要属性是保护物体边界:某一超像素块中所有像素都属于同一物体(理想状况下)^[6].分割结果的超像素块的边界如图 2 所示.



图 2 超像素示意图

Fig.2 Schematic diagram of superpixels

1.2 边界框起始位置的确定

本文把每一个超像素块作为一个节点.对于每个节点,一系列的边界框是基于以下算法确定的.图 3 为基于某一节点 K 确定边界框位置的示意图.

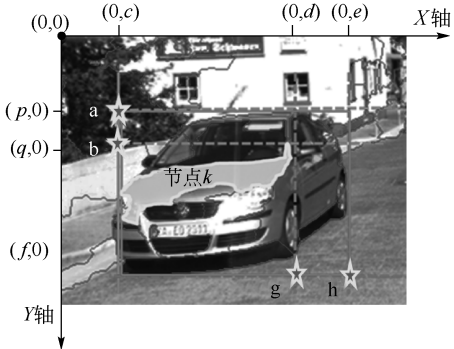


图 3 基于某一个节点确定边界框的示意图

Fig.3 Schematic diagram of locating the bounding-boxes based on a node

对于节点 K ,确定候选边界框起始位置的步骤

如下:1) 边界框的左边是由节点 K 中 X 轴坐标最小的像素点所在的纵轴决定(如图 3 中直线 $y=c$);2) 边界框的上边是由与节点 K 相邻的节点(包含节点 K)中 Y 轴坐标最小的像素点所在的横轴决定(如图 3 中直线 $x=p, x=q$).图 3 中所得到的左边和上边相交的点 a 和点 b ,便是候选边界框的起始点.而边界框终点的选取由节点 K 或其相邻节点的 X 坐标最大像素点所在纵轴与 Y 坐标最大像素点所在横轴的交点决定,如点 g 或点 h .如图 3 所示,同一节点可得到多个候选边界框.

1.3 候选边界框位置确定

一旦边界框的起始点确定(如 a 点的坐标为 (p,c)),则可以确定边界框的左缘(如图 3 中直线 $x=c$)和上缘(如图 3 中直线 $y=p$).对于每一个起始点,必须满足以下准则才能确定一系列边界框:1) 边界框的右缘是由与节点 K 相邻的两层节点(包含节点 K)中 X 轴坐标最大的像素点所在的纵轴决定(如图 3 中的垂直虚线);2) 边界框的底边是由与节点 K 相邻的两层节点(包含节点 K)中 Y 轴坐标最大的像素点所在的横轴决定(如图 3 中的水平虚线);3) 边界框宽与高的比例应在车辆宽高比的范围内.例如图 3 中边界框(实线矩形框)是由起始点图 3 的 a 点确定的边界框之一.

2 边界框的选择算法

基于图像分割得到的超像素图,本文提出了一种图论模型 $G = \langle V, E \rangle$ (见图 4),其中 V 代表的是节点集, E 代表的是无向边集.在本文的研究中,每个节点都是由分割算法生成的超像素块.每个节点(图 4 中 a 点)不仅和相邻节点(图 4 中 b 点)相连,而且还和与相邻节点有共同边界的节点(图 4 中 c 点)相连.每一对相连的节点 i 和 j 之间的权重被定义为

$$w(i,j) = \exp(-\frac{||\bar{g}_i - \bar{g}_j||}{\sigma^2}) \tag{1}$$

式中: \bar{g}_i 和 \bar{g}_j 表示节点 i 和节点 j 在某颜色空间中的平均值; σ 是控制权重强度的常量.

通过深入研究车辆数据集超像素图的每一个超像素块,发现所有包含车辆的边界框拥有一些共同的特征:1) 在一个包含完整汽车的边界框中,有许多封闭的超像素块;2) 属于同一辆车的超像素块之间具有较高的相似度.

根据以上提到的特性,本文提出了一种基于图论的方法来计算每个边界框中车辆可能存在分数.

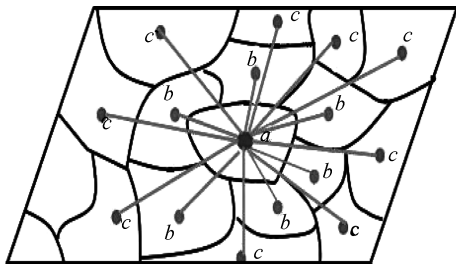
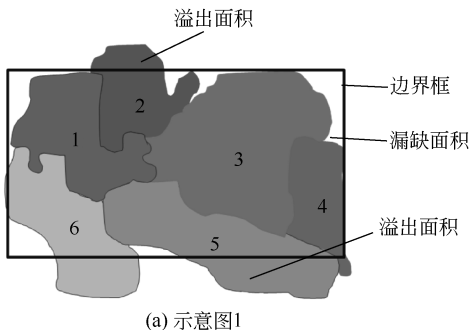
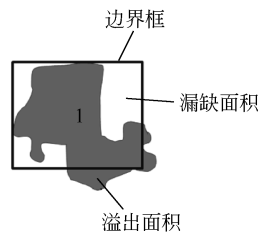


图 4 本文提出的图论模型示意图

Fig.4 Schematic diagram of graph theory model proposed in this paper



(a) 示意图1



(b) 示意图2

图 5 边界框中溢出面积和漏缺面积示意图

Fig.5 Schematic diagram of the overflow area and the missing area in a bounding box

溢出率和漏缺率计算公式为: 向量 $\mathbf{S} = [S_1, S_2, \dots, S_N]^T \in \mathbf{F}$, 其中 T 表示矩阵的转置, \mathbf{S} 为重心在边界框 \mathbf{F} 内的超像素块, N 为重心在边界框 \mathbf{F} 内超像素块的数量. 则溢出率 Ψ 和漏缺率 Ξ 由以下公式计算

$$\Psi = \sum_{i=1}^N O_{S_i} / A_{\text{box}} \quad (2)$$

$$\Xi = \sum_{i=1}^N M_{S_i} / A_{\text{box}} \quad (3)$$

式中: O_{S_i} 表示第 i 个超像素块的溢出面积; M_{S_i} 表示第 i 块超像素块的漏缺面积; A_{box} 表示限定框的面积.

2.2 边界框的相似度

边界框的相似度代表了该框中超像素块属于同一物体的可能性. 计算边界框 b 的相似度得

$$\zeta = \sum_{i=1}^N \sum_{j=1}^N w(i, j), \quad i \neq j \quad (4)$$

式中: $w(i, j)$ 为超像素块 i 与 j 之间的相似度, 即权值.

2.3 车辆可能存在分数

任意一边界框 b 的车辆可能存在分数 (Vehicle Proposal Score, VPS) 计算得

$$V_b = \frac{\zeta_b}{(t \cdot \Psi_b + \Xi_b) \cdot (b_w + b_h)^k} \quad (5)$$

式中: b_w 是边界框 b 的宽; b_h 是边界框 b 的高; Ψ_b

计算该分数分为 3 个步骤: 1) 计算每个边界框中溢出率和漏缺率; 2) 计算每个边界框中超像素块的相似度; 3) 计算每个边界框中车辆可能存在分数.

2.1 溢出率和漏缺率

有 6 个超像素块的重心在同一个边界框内, 如图 5(a) 所示. 图中超像素块 2、4、5、6 溢出边界框, 溢出边界框部分的面积称作溢出面积. 除此之外, 该边界框中不属于这 6 个超像素块的区域面积成为漏缺面积. 图 5(b) 为溢出面积和漏缺面积更直观的示意图.

和 Ξ_b 分别是边界框 b 的溢出率和漏缺率; ζ_b 是边界框 b 的相似度, t 和 k 是常量.

非极大值抑制 NMS 方法的功能是找出具有局部最大值的框. 计算出每个边界框的 VPS 并排序后, 用 NMS 方法去除局部分数较低的边界框, 筛选出具有局部最大值的边界框.

3 多视角分类器的选择

本文采用多通道特征 (Aggregate Channel Features, ACF)^[7], 其综合了颜色和梯度直方图特征. 分类器部分采用了 Adaboost 分类器.

多视角分类器模型是通过不同视角的正样本训练得到的. 多视角模型的提出对目前最先进的检测技术进行了改进. 在文献[8]Subcat 方法中提到设定的多视角分类器分为 25 个角度; 同时, 该设定的遮挡情况分为 3 类: 不遮挡、部分遮挡和严重遮挡. 因此, 通过训练不同视角和不同遮挡情况正样本, 共得到 75 个多视角分类器模型.

在本文中根据文献[8]Subcat 方法中提出的方法训练了一个检测器. 该检测器中每一个子检测器模型是通过不同视角和不同遮挡情况正样本单独训练的. 每一个子检测器模型 $i \{i=1, 2, \dots, 75\}$ 对应一个宽高比的平均值为 α_i .

对于每一个车辆候选区域边界框, 根据其宽高比可选出几个与其相适应的子模型. 根据子模型的

大小来调整车辆候选区域边界框的大小,通过分类器来判断边界框内是否含有车.对于同一个边界框,不同的分类器输出的检测分数不同.由于每个分类器对应不同的车辆视角,所以选择分数最高的分类器作为该边界框中车辆的行驶角度.

4 实验结果与分析

本文采用 KITTI 来评价所提出的方法.KITTI 数据集^[9]包含了 7 481 张训练图像和 7 518 张测试图像.图 6 为该数据集中一些图像,这些图像中的车辆是在不同角度和光照拍摄的.从图 6 可以看出,该数据集中的车辆往往处于不同的复杂背景下,如有些车辆遮挡比较严重,有些车辆处于较强光照或者阴影下,还有些车辆处于与自身颜色相近的建筑群



图 6 KITTI 数据集示意图
Fig.6 Schematic diagram of KITTI dataset

4.1 实验参数设置

在本文中,为了优化实验结果,对不同实验参数得到结果进行了分析.

1) 超像素块的数量:设置超像素块的数量为 800、1 000、1 500 和 2 000.如图 7(a)所示,当超像素块的数量为 1 000 时, F 最高.因此,把超像素块的数量设置为 1 000.

2) 计算权重:分别在 LAB (Lightness-A-B)、

中,这些复杂情况使得车辆检测变得更加困难.该数据集的官方网站上给出了结果评价标准:根据车辆的遮挡程度和可检测出的车辆高度将车辆分:简单、中等和困难 3 类;精确度 P ,召回率 R ,检测综合评价指标 F 分别由下式计算而得

$$P = TP / (TP + FP) \tag{6}$$

$$R = TP / (TP + FN) \tag{7}$$

$$F = \frac{2 \times PR}{P + R} \tag{8}$$

式中: TP 为正确的检测数量; FN 目标区域被判定为非目标区域; FP 非目标区域判定为目标区域; F 越高说明检测精度越高.如果检测结果框与标定框的重叠率达到了 70% 以上,则判定为正确的检测.

HIS (Hue, Intensity and Saturation) 和 LUV (CIE 1976 L^*, u^*, v^*) 的 3 种颜色空间下进行实验.实验结果如图 7(b)所示,在 LAB 颜色空间下的 F 最高.因此,选用 LAB 颜色空间.

3) 式(5)中 k 的值:本文中设置 k 值为 0.5、1.0、1.5 和 2.0.实验结果如图 7(c)所示,当 k 为 1.0 时, F 最高.因此,将 k 值设置为 1.0.

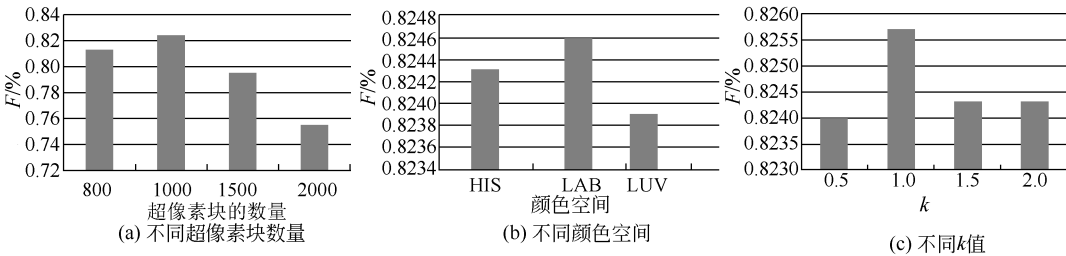


图 7 不同参数对实验结果的影响

Fig.7 Impact of different parameters on the experimental results

4.2 不同车辆检测系统的比较

表 1 显示了不同车辆检测在 KITTI 数据集中检测结果,尤其是 SubCat^[8]和 ACF 使用的图像特征和检测框架与本文相同,其他当前最先进的方法使用不同的图像特征,如 CNN^[10]、Bag-of-Words (mBoW)^[11]和 spatially pooled features (spLBP)^[12].本文提出的方法检测性能与 ACF 方法相比,在表 1 中可知,在简单、中等和困难的 3 类中分别提高了 23.99%、13.04%和 14.71%.

表 1 KITTI 中现有检测方法的车辆检测结果

Tab.1 Vehicle detection results of approaches in KITTI %

方法	简单	中等	困难
本文	79.88	67.78	57.69
SubCat	81.94	66.32	51.10
OC-DPM ^[13]	74.94	65.95	53.86
DPM-VOC+VP ^[14]	74.95	64.71	48.76
ACF	55.89	54.74	42.98
mBoW	36.02	23.76	18.44

4.3 不同似物性窗口生成算法的比较

选择了两个目前比较流行的似物性窗口生成算法 Selective-Search^[15]和 Edge Boxes^[16]与本文的方法做比较,其中,Selective-Search 算法也是基于图论的目标窗口生成算法.分别从 3 种方法取 2 000 个窗口进行检测,检测结果如表 2 所示.由表 2 可知,本文提出的方法检测结果在简单、中等和困难的 3 类中都是最高的.

表 2 KITTI 中现有似物性窗口生成方法的车辆检测结果

Tab.2 Vehicle detection results of object proposals windows generation methods in KITTI %

方法	简单	中等	困难
本文	63.14	54.32	46.24
Edge Boxes	44.62	40.74	30.93
Selective-Search	24.89	17.08	16.61

4.4 算法复杂度分析

为了确保实验的公平性,本文的实验都是在操作环境 2.00 GHz Intel Xeon CPU 和 Matlab 2014 的仿真平台和相同的实验参数下(参照 4.1 节)进行的.不仅统计了算法检测部分的计算耗时,而且计算了算法的时间复杂度.在实验中,测试图像的大小为 375 × 1 242,多视角分类器的数量为 75 个,测试图像尺寸变换次数为 36.如表 3 所示.

表 3 中, $O(m \times n)$ 代表算法复杂度(时间复杂度), m 代表每次问题的语句执行数, n 代表问题的

规模循环次数.

由于 SubCat 方法使用的是滑动窗口法,检测耗时为 1.683 s;而本文提出的检测方法,系统检测耗时为 0.781 s;虽然 ACF 检测耗时较短(见表 1),但是检测性能较差.

表 3 算法复杂度和检测耗时的比较

Tab.3 Comparison of computational complexity and detection time consumption

方法	算法复杂度	检测耗时/s
本文	$O(1)$	0.781
SubCat ^[7]	$O(m \times n)$	1.683
OC-DPM	$O(m \times n)$	10.000
ACF	$O(1)$	0.200

5 结论

1)本文作者提出了一种新的基于图论的车辆检测算法来确定车辆候选区域,通过该方法获得的车辆候选区域数量更少和位置更准确.在检测阶段,通过训练不同视角的车辆正样本来获得多视角检测器,可以有效地解决复杂背景下的车辆检测问题.

2)采用的公共交通分析数据集(KITTI)是由汽车在城市环境中通过常规交通行驶拍摄的,该数据集中大部分车辆处于强光照或者遮挡等复杂背景条件下,满足验证检测模块有效性的需求.通过实验仿真并与目前常用的几种确定目标候选区域的方法(具有相同提取特征)相比较,本文提出的车辆检测算法的检测精度高于其他算法 10%以上.

参考文献(References):

[1] HAREL J, KOCH C, PERONA P. Graph-based visual saliency[C]// The Neural Information Processing Systems Conference, 2006.

[2] YANG C, ZHANG L, LU H, et al. Saliency detection via graph-based manifold ranking[C]// IEEE Conference on Computer Vision and Pattern Recognition, 2013, 9(4): 3166—3173.

[3] 刘建龙. 基于图论的图像分割算法研究[D]. 哈尔滨: 哈尔滨工业大学, 2006.

LIU Jianlong. Image segmentation algorithm based on graph theory [D]. Harbin: Harbin Institute of Technology, 2006. (in Chinese)

[4] ACHANTA R, SHAJI A, SMITH K, et al. SLIC superpixels compared to state-of-the-art superpixel methods[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(11): 2274—2281.

[5] CAMPELLO R J G B, MOULAVI D, SANDER J.

- Density-based clustering based on hierarchical density estimates[C]// 17th Pacific-Asia Conference on Knowledge Discovery and Data Mining, 2013:160—172.
- [6] 许新征, 丁世飞, 史忠植, 等. 图像分割的新理论和新方法[J]. 电子学报, 2010, 38(2A): 76—82.
- XU Xinzhen, DING Shifei, SHI Zhongzhi, et al. New theories and methods of image segmentation[J]. ACTA Electronica Sinica, 2010, 38(2A): 76—82. (in Chinese)
- [7] DOLLAR P, APPEL R, BELONGIE S, et al. Fast feature pyramids for object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36(8): 1532—1545.
- [8] OHN-BAR E, TRIVEDI M M. Learning to detect vehicles by clustering appearance patterns[J]. IEEE Transactions on Intelligent Transportation Systems, 2015, 16(5): 2511—2521.
- [9] GEIGER A, LENZ P, URTASUN R. Are we ready for autonomous driving the KITTI vision benchmark suite [C]// CVPR, 2012:3354—3361.
- [10] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]//CVPR, 2014: 580—587.
- [11] BEHLEY J, STEINHAGE V, CREMERS A B. Laser-based segment classification using a mixture of bag-of-words[C]// 2013 IEEE/RSJ Intelligent Robots and Systems(IROS), 2013:4195—4200.
- [12] HU Q, PAISITKRIANGKRAI S, SHEN C, et al. Fast detection of multiple objects in traffic scenes with a common detection framework[J]. IEEE Transactions on Intelligent Transportation Systems, 2016, 17(4): 1002—1014.
- [13] PEPIK B, STARK M, GEHLER P, et al. Occlusion patterns for object class detection[C]// IEEE Conference on Computer Vision and Pattern Recognition, 2013, 9(4):3286—3293.
- [14] PEPIK B, STARK M, GEHLER P. Multi-view and 3D deformable part models[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(11):2232—2245.
- [15] UIJLINGS J R, SANDE K E, GEVERS T, et al. Selective search for object recognition[J]. International Journal of Computer Vision, 2013, 104(2):154—171.
- [16] ZITNICK C L, DOLLAR P. Edge boxes: locating object proposals from edges [C]//European Conference on Computer Vision, 2014: 391—405.