

基于 Haar-like 特征与 Adaboost 算法的 前方车辆辨识技术研究*

朱志明¹ 乔洁²

(1. 华东交通大学 南昌 330013; 2. 长安大学汽车学院 西安 710064)

摘要: 为了提高前方车辆的辨识效能,提出一种融合 Haar-like 特征与 Adaboost 算法的前方车辆辨识方法,基于海量车辆样本集进行离线训练,提取有效车辆轮廓与纹理特征,以 Haar-like 特征作为目标描述方法,采用 Adaboost 机器学习算法训练分类器,并构建特征样本级联分类器,对测试对象进行车辆存在性检测。试验结果表明,提出的融合 Haar-like 与 Adaboost 的车辆辨识算法检测准确率为 91% 以上,平均检测速率 28 ms,对车辆类型和环境干扰等非确定因素具有较强的自适应能力,提高了前方车辆纵向检测的鲁棒性,满足了车辆纵向维度的安全行驶应用需求。

关键词: Haar-like 特征; Adaboost; 训练样本集; 辨识

中图分类号: TN081 **文献标识码:** A **国家标准学科分类代码:** 510.4050

Research of preceding vehicle identification based on Haar-like features and Adaboost algorithm

Zhu Zhiming¹ Qiao Jie²

(1. East China Jiaotong University, Nanchang 330013, China; 2. School of Automobile, Chang'an University, Xi'an 710064, China)

Abstract: In order to improve the efficiency of the preceding vehicle identification, a preceding vehicle identification algorithm combined Haar-like features and Adaboost algorithm was proposed. Based on massive amounts of offline training sample set, effective vehicle contour and texture characteristics was extracted, Haar-like characteristics was used to describe the goal. Adaboost machine learning algorithms was used to trained classifier, the sample characteristics of cascade classifier was built, and the test object was used to detect the vehicle existence. The experimental result shows that the algorithm based on Haar-like features and Adaboost algorithm in the paper has a high detection rate of above 91%, and average detection speed of 28 ms, it can well adapt to the uncertain factors such as environmental interference and vehicle type, improve the robustness of the preceding vehicle detection, also meet the requirement of the safe driving in the longitudinal dimensionality.

Keywords: Haar-like feature; Adaboost; training sample set; identification

1 引言

随着社会经济的快速发展及我国车辆保有量的快速增加,我国的道路交通事故数量一直居高不下。据道路交通管理部门统计分析,车辆纵向追尾事故是重大道路交通事故的主要形态之一。因此,采用先进辅助驾驶系统,实时检测前方车辆纵向位置,对提高车辆在途行驶安全性具有重大现实意义。

常见的车辆识别方法有基于车辆阴影^[1]、基于车尾数学模型^[2]、基于车辆轮廓识别^[3]等,基于车辆阴影的识别方

法受外界光线因素影响较大,基于车尾数学模型方法通过对应点标定来获取图像的深度信息,但受制于器材限制及标定客观原因,无法得到较高精度的坐标系之间相互转变的转换矩阵,适用性受到限制,基于 Adaboost 机器学习方法^[4-7]进行车辆轮廓识别的方法对外界环境干扰具有较强的免疫能力、鲁棒性好,受到广泛关注,但此方法样本训练时间特别长、虚警率高。因此,本文提出一种新的前方车辆辨识方法,充分挖掘目标车辆图像信息,寻找目标的 Haar-like 特征,结合机器学习算法进行基本分类器的训练,利用特征样本级联分类器进行车辆存在性检测,提高前方车辆

辨识效能。

2 车辆矩形特征提取

机器学习的功能是利用已知的训练样本对未知的数据进行检测,级联分类器的构建在车辆目标识别检测过程中起到重要作用,通过对海量样本集进行离线训练一方面提取对分类识别起重要作用的特征,另一方面构建车辆目标识别的级联分类器。作为训练样本的目标特征提取在目标识别中起着重要作用,相对于采用像素值进行目标提取而言,Haar-like 矩形特征^[8-9]是一种快速、简单提取车辆轮廓与纹理特征的有效提取方式,主要因为利用 Haar-like 矩形特征可以减少同类样本之间的差异,同时其计算速度远远大于基于像素值的计算速度。在使用 Haar-like 矩形特征进行图像扫描时,以这些基础区域像素大小为单位进行遍历搜索,因此,当扫描一副图片时,会产生大量的特征,导致计算速度下降。

Rainer Lienhart 等提出了一种扩展的 Haar-like 矩形特征^[10],并针对扩展的矩形特征提出了一种快速计算方法。扩展的 Haar-like 矩形特征如图 1 所示,共有 4 大类 15 种特征来描述车辆的边缘及结构特征。

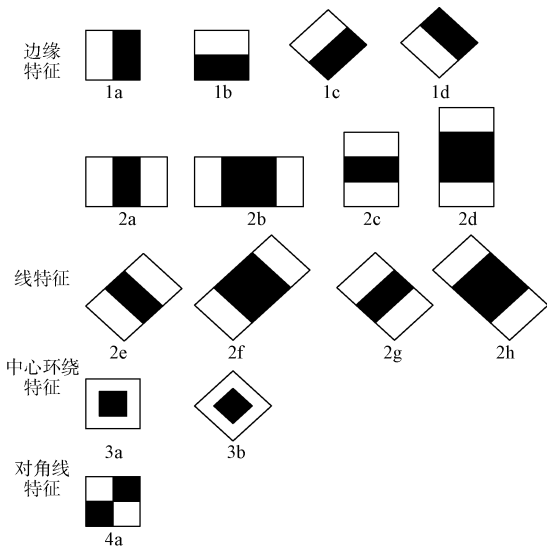


图 1 扩展的 Haar-like 矩形特征图

对于 32×32 像素的车辆样本而言,仅传统的矩形特征至少几万个,如果每个矩形特征都需要计算,这样产生的特征量太大,因此,本文提高算法运行效率采用的是积分图方法。在实际使用中,第一步要转化像素图像,转化后每个点表示图像左上角起点到该点矩形区域内的像素和,接着进行积分图像提取。所用的积分图计算方法如下式所示:

$$i(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y') \quad (1)$$

式中: $i(x, y)$ 为 (x, y) 在积分图像上的像素积分值, $i(x', y')$ 为原图中 (x, y) 点的像素值。将图像遍历一次即可计算

出所有区域的积分图。

3 基于离线训练的车辆级联分类器构建

3.1 基于 Adaboost 算法的分类器训练

通过积分图计算获得的 Haar-like 特征集相当混乱,因此必须通过组建强分类器进行目标识别,本文进行强分类器训练的方法是采用 Adaboost 算法。Adaboost 分类器^[11-12]通过每次循环提取一个相应的有效特征,每个有效特征随之产生一个对应的弱分类器,最后许多弱分类器叠加变成一个强分类器。

本所建立的强分类器训练步骤如下:

1)构建样本集: $\{(X_1, Y_1) \cdots (X_i, Y_i) \cdots (X_n, Y_n)\}$,负样本为 Y 取 0 时,正样本为 Y 取 1 时,所有包含目标车辆的图像归为正样本,所有含有常见车辆周围环境的图片归结为负样本。对于待训练的每幅图片,其特征值都存储在特征向量 X 内,即 $X = \{x_1, x_2, \dots, x_j, \dots, x_n\}$,特征向量 X 中每个元素代表该幅图片的一个特征值。

2)生成弱分类器。对于每个 Haar-like 特征,都可以生成一个弱分类器,具体表现形式如下:

$$h_j(x) = \begin{cases} 1, & p_j f_j < p_j \theta_j \\ 0, & \text{其他} \end{cases} \quad (2)$$

其中: h_j 表示弱分类器, f_j 表示第 j 个特征值, θ_j 为阈值, p_j 为奇偶性验证因子,Adaboost 可以自动选择阈值使正负样本最大程度的分开。

3)初始化权重。当 y_i 取 1 为正样本时,权重因子为:

$$\omega_{i,j} = \frac{1}{2l} \quad (3)$$

当 y_i 取 0 为负样本时,权重因子为:

$$\omega_{i,j} = \frac{1}{2m} \quad (4)$$

对于每个循环体 $t=1, 2, \dots, T$ (T 为迭代次数,决定弱分类器的个数),归一化权重为:

$$\omega_{i,j} = \frac{\omega_{i,j}}{\sum_{j=1}^n \omega_{i,j}} \quad (5)$$

4)对于第 j 个特征,计算生成的弱分类器 h_j ,推导其相对于当前权重的误差:

$$\epsilon_j = \sum_{i=1}^n \omega_i |h_j(x_i) - y_i| \quad (6)$$

把有较小误差 ϵ_j 的弱分类器放到 h_j 的强分类器中。

5)更新单个样本权重:

$$\omega_{i+1,i} = \omega_{i,j} \beta^{1-\epsilon_i} \quad (7)$$

若 x_i 分类错误, $\epsilon_i=1$,若 x_i 分类正确, $\epsilon_i=0$,并且 $\beta_i = \epsilon_i / (1 - \epsilon_i)$ 。

则最终的强分类器为:

$$h(x) = \begin{cases} 1, & \sum_{i=1}^T \log(1/\beta_i) h_i \geq \frac{1}{2} \sum_{i=1}^T \log(1/\beta_i) \\ 0, & \text{其他} \end{cases} \quad (8)$$

当使用 Adaboost 算法进行样本分类时,如果错样本分类,则增加其对应的比例系数,如果进行正样本分类,则降低其对应的比例系数,把弱分类器依据错样本比例系数重新加权,最后能训练成一个强分类器。

3.2 特征样本的级联分类器构建

通常待检测图像中大部分区域都不包含目标车辆,若直接用一个高性能的分类器进行检测,大量的时间会浪费在非车辆区域,因此,本文采用级联分类器快速排除非车辆区域,提高目标检测速度。

如图 2 所示,本文采用经典的瀑布型分类器进行级联^[13],每个层为用 Adaboost 算法训练得到的 Adaboost 分类器,每个 Adaboost 分类器又包含若干个弱分类器,待识别的样本图被级联分类器一层一层进行检测,如果在任何一层被判为负样本,后面的强分类器都通不过,这样最终能提高正样本的识别效率。

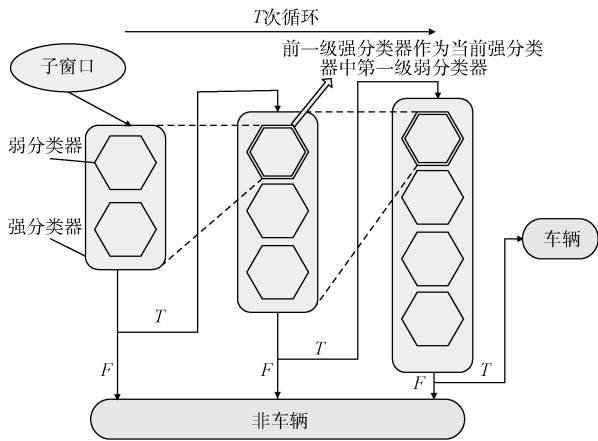


图 2 级联分类器训练过程图

K 层的级联分类器总特征数如下所示:

$$N = n_1 + \sum_{j=2}^K n_j \prod_{i < j} p_i \quad (9)$$

其中,公式中 p_i 代表每一层的通过率, n_i 代表分类器的特征数,则由式(9)可以看出:

$$N \leq \sum_{i=1}^K n_i \quad (10)$$

说明级联分类器的窗口检测速度随着特征数目的减少而增加。

使用级联分类器时,希望检测率越高越好,虚警率越小越好,只要知道最终期望的检测率和虚警率,就可以反推每层分类器的检测率和虚警率。设级联分类器的虚警率为:

$$F = \prod_{i=1}^K f_i \quad (11)$$

式中: F 为总的虚警率, K 为分类器层数, f_i 为每一层的虚警率。

检测率为:

$$D = \prod_{i=1}^K d_i \quad (12)$$

式中: D 为总的检测率, K 为分类器层数, d_i 为每一层的检测率。

考虑到车辆目标检测的可信度及算法的实时性,本文所用级联分类器采用 8 层分类器,最终的检测率在 0.9 以上,每一层的误检率为 0.5,则每一层检测率在 0.99 以上。

4 融合 Haar-like 特征与 Adaboost 算法的前方车辆识别

本文提出的融合 Haar-like 特征与 Adaboost 的前方车辆识别算法总共分为两个步骤:1)离线训练,建立样本集,利用上节所述的方法对大量车辆样本进行 Haar-like 特征提取,基于 Adaboost 算法进行级联分类器的构建;2)在线识别,对样本库的样本逐一提取特征,进行车辆存在性检测,所采用的方法为将提取的 Haar-like 特征输入到 Adaboost 级联分类器。算法结构如图 3 所示。

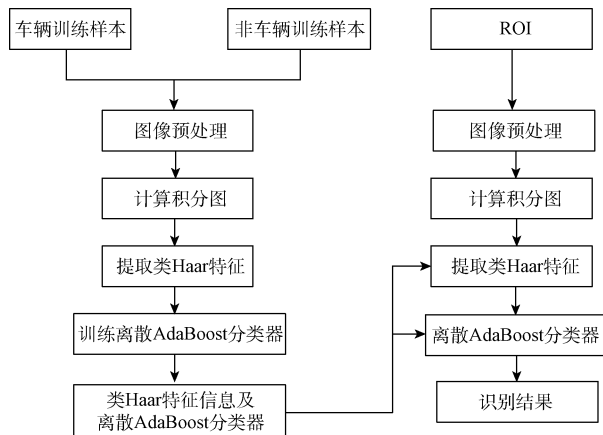


图 3 基于 Haar-like 特征和 Adaboost 的车辆识别算法结构

训练样本分为正样本和负样本,正样本指包含车辆样本图片,选择大小不同、位置不同的包含目标车辆图片;其余只包含背景区域的图片归结为负样本,但这些图片应为车辆周围常出现的环境,但不能包含车辆的任何特征。本文训练样本库收集了 500 张正样本,1 500 张负样本,如图 4、5 所示。



图 4 正样本集部分例图



图 5 负样本集部分例图

为了确保特征矩阵的数目一致,必须利用不变性的矩对图像归一化处理到设定尺寸,由该矩推出变换函数,此函数可归一化样本库图像,本文将样本归一化 30×30 像素大小的样本集,如图 6 所示。



图 6 归一化的样本集图

也就意味着在实际车辆识别过程中,车辆级联分类器只能对图像中 30×30 像素大小的目标车辆进行识别,因此,在识别过程中,本文使用高斯图像金字塔将不同距离下的车辆缩放到级联分类器可识别范围内。

如图 3 所示,基于 Haar-like 特征和 Adaboost 的车辆识别过程首先对待识别图像的感兴趣区域(ROD)进行图像预处理和计算积分图,根据摄像头的固定位置,本文的 ROI 区域为图像的下半部分,图像预处理及计算积分图同训练过程一致;其次,用训练过程选择的 Haar-like 特征信息包含结构、位置、类型等来提取测试样本的 Haar-like 特征值,组成特征向量;最后,采用通过海量离线训练得到的 Adaboost 级联分类器对感兴趣区域进行车辆存在性检测。

5 实验应用

为了验证文中提出的车辆识别算法的有效性和可行性,本文采用离线实验的方法进行测试,选取高速公路车载视频的 100 帧图像作为测试样本,部分帧目标车辆识别实验结果如图 7 所示。实验结果表明,本文提出的基于 Haar-like 特征和 Adaboost 的车辆识别算法的检测准确率为 91% 以上,误检率为 9% 左右,相对于传统基于 Adaboost

算法的车辆识别准确率提高了 10%,平均检测速率 28 ms,在保证算法实时性的同时,能够有效地避免行车道两侧道路阴影、其他非本车道内的车辆等干扰因素的影响,提高系统的检测鲁棒性。



图 7 前方目标车辆识别结果

6 结 论

本文基于海量车辆样本集进行离线训练,提取有效车辆轮廓与纹理特征,结合 Adaboost 机器学习算法训练分类器,进而组建特征样本级联分类器,对测试对象进行车辆存在性辨识。实验结果表明:本文提出的融合 Haar-like 与 Adaboost 的车辆辨识算法检测准确率为 91% 以上,误检率为 9% 左右,平均检测速率 28 ms,在保证了算法实时性的同时,能够有效地避免行车道两侧道路阴影、其它非本车道内的车辆等干扰因素的影响,提高了系统的检测鲁棒性,满足了结构化道路上车辆辅助驾驶系统的纵向应用需求。

参考文献

- [1] ASAIIDI H, AARAB A, BELLOUKI M. Shadow elimination and vehicles classification approaches in traffic video surveillance context [J]. Journal of Visual Languages & Computing, 2014, 25 (4): 333-345.
- [2] CHEN Y, CHEN Y. Calculation method of preceding vehicle distance based on monocular vision [J]. Journal of Mechanical & Electrical Engineering, 2014, 31(11): 1499-1504.
- [3] 张楚金, 王耀南, 卢笑, 等. 基于假设验证和改进 HOG 特征的前车检测算法[J]. 电子测量与仪器学报, 2015, 29(2): 165-171.
- [4] 娄青青. 基于 Adaboost 算法的车辆识别系统研究与实现[D]. 沈阳:东北大学, 2009.
- [5] 张丽红, 李林. 基于 Adaboost 算法选取和组合 SVM 的行人检测算法[J]. 测试技术学报, 2013, 27(1): 73-78.
- [6] 胡玉琛, 刘一清. 基于 FPGA 车牌识别系统研究与实现[J]. 电子测量技术, 2015, 38(12): 54-58.
- [7] 张晓娜, 何仁, 陈士安, 等. 基于主动学习 Adaboost 算法与颜色特征的车牌定位[J]. 交通运输工程学

- 报, 2013, 13(1): 121-126.
- [8] 张亮修. 基于 Haar-like 特征的实时道路车辆识别方法研究[D]. 青岛: 青岛大学, 2009.
- [9] 王庆伟, 应自炉. 一种基于 Haar-like T 特征的人脸检测算法[J]. 模式识别与人工智能, 2015, 28(1): 35-41.
- [10] 甘玲, 朱江, 苗东. 扩展 Haar 特征检测人眼的方法[J]. 电子科技大学学报, 2010, 39(2): 247-250.
- [11] 朱文球, 刘强. 融合 AdaBoost 和启发式特征搜索的人脸性别分类[J]. 计算机工程, 2007, 33(2): 171-173.
- [12] 张艳阳, 顾明. 基于 AdaBoost 分类器和特征脸的人脸检测与识别[J]. 漳州师范学院学报: 自然科学版, 2010(2): 43-47.
- [13] 李芳. 基于 AdaBoost 算法的人脸检测研究与实现[D]. 武汉: 武汉理工大学, 2013.

作者简介

朱志明, 1999 年于华东交通大学获得学士学位, 2002 年于华东交通大学获得硕士学位, 现为华东交通大学助理工程师。研究方向为图像处理、计算机控制。

E-mail: 514366153@qq.com

乔洁(通讯作者), 2011 年于长安大学获得学士学位, 2014 年于长安大学获得硕士学位, 现为长安大学汽车学院博士研究生, 研究方向为图像处理、车辆安全。

E-mail: liuyongtao86@163.com

(上接第 167 页)

- [6] 郭斌, 朱戈, 陆艺, 等. 智能化上料系统中工件视觉定位的研究[J]. 计算机测量与控制, 2016, 24(2): 232-234.
- [7] SINGH S, SAINI A K, SAINI R, et al. A novel real-time resource efficient implementation of Sobel operator-based edge detection on FPGA [J]. International Journal of Electronics, 2014, 101(12): 1705-1715.
- [8] SARVAIYA J N. Automatic image registration using mexican hat wavelet, invariant moment, and radon transform [J]. International Journal of Advanced Computer Sciences and Applications, 2011, 32(12): 1245-1249.
- [9] 韩嘉隆, 毛征, 王宁. 基于二维 OTSU 的海天分界线提取算法[J]. 国外电子测量技术, 2016, 12(8): 67-70.
- [10] YU G. An algorithm for fully affine invariant workiece comparison[J]. Image Processing on Line, 2013, 8(3): 28-35.

作者简介

吴鹏 1982 年出生, 硕士, 讲师, 研究方向为计算机应用、图像处理。

E-mail: wujuan9652xzy@sina.com