

多特征融合的车辆识别技术

程 全¹, 樊 宇², 刘玉春¹, 王志良³

- (1. 周口师范学院 机械与电气工程学院, 河南 周口 466001;
2. 周口师范学院 网络工程学院, 河南 周口 466001;
3. 北京科技大学 计算机与通信工程学院, 北京 100083)

摘 要: 针对运动车辆目标识别问题提出了一种自然场景下车辆识别方法。首先采用图像差分技术对目标车辆的显著特征进行统计学习, 并将学习所得目标局部特征以及图像进行编码, 根据以上两个信息实现目标车辆的显著性检测。其次针对车辆运动的复杂性, 采用分块投影匹配方法进行全局运动估计和补偿, 并利用差分技术进行运动特征检测。然后将目标车辆的显著性特征与运动特征进行融合, 从而获得更精确的候选目标区域。最后对候选区域进一步使用视觉显著特征进行目标判别。实验表明该方法具有较好的目标判别性能, 能较好地解决自然场景下运动车辆的识别问题。

关键词: 特征融合; 车辆识别; 运动特征; 统计学习

中图分类号: TP391 文献标志码: A DOI: 10.3788/IRLA201847.0726003

Multi-feature fusion vehicle identification technology

Cheng Quan¹, Fan Yu², Liu Yuchun¹, Wang Zhiliang³

- (1. Mechanical and Electrical Engineering, Zhoukou Normal University, Zhoukou 466001, China;
2. College of Network Engineering, Zhoukou Normal University, Zhoukou 466001, China;
3. School of Computer & Communication Engineering, University of Science & Technology Beijing, Beijing 100083, China)

Abstract: A method of vehicle identification in natural scene was proposed for the target recognition of moving vehicles. Firstly, the image difference technique was used to study the significant characteristics of the target vehicle, and encode the local features and the image of the learning target, then the salient detection of the target vehicle was realized according to the above two information. Secondly, aiming at the complexity of vehicle movement, a block projection matching method was used for global motion estimation and compensation, and the difference technique was used to detect the motion feature. Then the target vehicle's saliency features and the motion features were fused to obtain more accurate candidate target areas. Finally, the candidate region was further used to discriminate the target by using the visual features. The experiment shows that the method has good performance of target discrimination and can solve the problem of vehicle motion recognition in natural scene.

Key words: feature fusion; vehicle identification; motion feature; statistic study

收稿日期: 2018-02-05; 修订日期: 2018-03-03

基金项目: 国家自然科学基金(61401526); 河南省科技厅科技攻关项目(182102210151, 182102310761)

作者简介: 程全(1978-), 男, 副教授, 硕士, 主要从事智能控制方面的研究。Email: quan8888@126.com

0 引言

目标识别技术是利用图像处理和大规模集成电路技术,对目标进行快速识别的技术。由于目标所处场景和本身信息的缺损或模糊等诸多因素,使得对目标的识别成为一个很复杂的过程。随着微电子技术的飞速发展以及生物视觉、机器视觉、图像处理、模式识别、机器学习、人工智能等学科理论研究的深入开展,目标识别技术研究取得了较大进展,已成为当今机器视觉领域的热门研究方向。近年来,国内外若干高校和研究机构在该领域开展了较为系统深入的研究工作^[1-4],推出了一些具有一定实用价值的目标识别系统,但在不同的目标类型、运动模式、成像条件及场景下,目标识别系统的性能会受到显著影响^[5]。迄今为止,尚没有各种复杂场景下通用的目标识别方法。

近几年来,国际上有关稀疏编码、字典学习、深度学习等特征学习技术研究取得了重要进展,获得了一些有重要意义的理论研究成果,部分理论与算法已在机器视觉领域成功用于对特定视觉目标的检测与目标的类别识别^[6-7],如采用稀疏编码表示与识别人脸^[8]、采用字典学习表示与识别多类视觉目标^[9]。国际上,已有多个研究机构与大学构建了若干专用图像数据库如 Caltech101、Caltech256、ImageNet、Norb、LabelMe 等用于目标识别研究,并举办了若干年度比赛会议如 Pascal2 来衡量有关目标识别算法的性能,从而大力促进视觉目标识别研究的发展。军事上,在自动目标识别领域,有关目标特征学习的研究才刚刚起步,美国国防部先进研究计划局(DARPA)于 2010 年初发布了有关征求开展深度特征学习技术研究的项目申请指南,旨在促进该领域的技术研究^[10]。

可以预见,基于大量目标图像,通过特征学习自动提取与选择目标特征、建立有效的目标视觉显著特征模型,并与目标的运动特征进行融合处理,可实现复杂场景下高精度的目标识别。因此,针对自动目标识别开展稀疏编码、字典学习等特征学习理论与方法研究,无论在军用还是民用领域,都具有极其重要的理论意义与广阔的应用前景。

经过多年研究,目前已提出许多目标识别方法,

主要有基于统计模式识别的方法、基于知识的方法、基于模型的方法、基于信息融合的方法等。实际应用证明,对于复杂场景下的运动目标识别问题,由于目标运动造成姿态改变而使目标外观变化很大,外加遮挡与地面杂波干扰的影响,使得传统的目标识别方法难以奏效^[11]。为此文中提出了基于多关系模板匹配的车辆识别技术,该方法根据特定应用的任务需求,运用统计学习技术自动提取目标的视觉显著特征,并将之与目标的运动特征进行融合处理,提高目标识别模型的泛化能力。对有限复杂场景下的目标识别准确率较高,解决了依靠经验人为事先选定的目标特征对变化场景的适应能力不足造成的影响,是一种切实可行的车辆目标识别方法。

1 车辆目标检测方法

从车辆视频图像中正确识别运动车辆目标的前提是可靠提取目标所在区域,即准确检测出运动车辆目标。由于自然场景相当复杂,图像信息极为丰富,要正确检测车辆目标,首先要准确描述目标,即采用合适的特征表示目标^[12]。考虑到目标运动与背景运动存在差异,可以通过背景运动补偿和历史运动累积检测序列图像中的运动显著区域作为候选车辆目标。但仅仅依靠运动信息检测车辆目标还不够准确,车辆目标可能会由于其相对背景做低速运动致使历史运动累积不够显著而造成漏检,因此还需要检测对车辆目标而言的视觉特征显著区域^[13]。针对自然场景下的运动目标识别,文中采用的处理流程如图 1 所示。

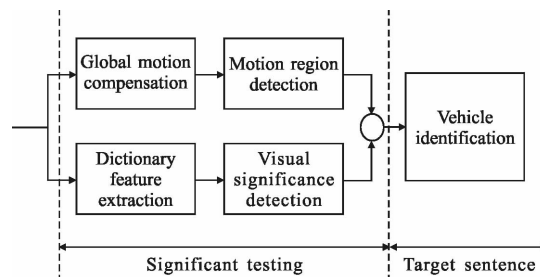


图 1 车辆目标识别处理流程

Fig.1 Vehicle target recognition process

由图 1 可见,车辆目标识别过程总体上可分为两个阶段,即显著检测和目标判决。显著检测阶段基于视觉选择注意机理,根据目标的运动特征与视觉

显著特征对序列图像进行逐块分析,提取可能存在目标的局部区域,该过程仅利用少量的区分能力强的特征对序列图像进行分析,是目标检测的粗分析过程,计算快速性是关键指标之一。目标判决阶段基于大量目标图像,通过统计学习建立视觉词典,构建目标的稀疏特征表示,进一步采用集成学习方法逐步选择区分能力强的视觉显著特征,构建级联分类器,实现目标与非目标的准确识别,是目标检测的精细分析过程,计算精确性是关键指标之一。

1.1 运动区域检测

针对运动目标检测对实时性的要求,提出一种基于分块投影匹配的快速全局运动估计方法。首先采用分块投影匹配进行全局运动估计和补偿,然后采用图像差分技术进行运动检测。

基于分块匹配思想的全局运动估计匹配精度高,特别适用于对连续图像序列的全局运动估计,但常见的分块匹配方法通常是基于某种匹配准则(如绝对误差和准则、均方误差准则、归一化互相关函数准则等),在给定搜索范围内寻找最相似的图像块,虽然可以通过对搜索策略进行优化而提高运算速度,但仍然相当耗时;灰度投影方法利用图像灰度总体分布的变化规律估计图像运动矢量,虽然不能适用于旋转、缩放等运动情形,但该方法简单、运算速度快。为了提高算法的实时性,将灰度投影算法与分块匹配法相结合,首先通过灰度投影匹配求取图像子块的平移运动矢量,在此基础上采用 WRANSAC 方法估计全局运动。

假设图像中的景物运动是由摄像机的运动而引起,令对应像素的图像坐标在运动前后分别为 x, y , 则以六参数仿射运动模型表示的摄像机全局运动可由公式(1)描述:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} \quad (1)$$

式中: h_1, h_2, h_4, h_5 表示旋转、缩放运动; h_3, h_6 表示平移运动,如果已知图像中若干像素的对应关系,那么全局运动模型参数估计即转化为根据对应关系求 H 矩阵。

首先将图像划分为 $M \times N$ 块,通过投影匹配计算相邻两帧图像对应子块间的平移矢量,将其作为该子图像中心点的位移矢量,然后提出一种评估函数

以估计平移运动矢量的准确性,最后采用稳健估计方法计算全局运动模型参数,并根据参数对 F_{k-1} 进行运动补偿,得到 F_{k-1}' ,如图 2 所示。

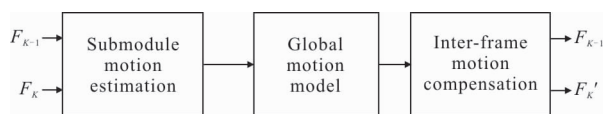


图 2 全局运动补偿

Fig.2 Global motion compensation

对于动态场景的图像序列,如果对其进行全局运动补偿,则静态背景下的运动目标检测方法也能适用。

1.2 显著性检测

显著检测即根据目标的运动特征与视觉显著特征进行候选目标区域判决,其关键是特征的高效提取。根据目标的视觉显著特征进行候选目标区域判决的前提是建立目标的视觉显著特征字典,文中研究针对预定类别的目标图像,提取图像局部兴趣区,采用 SURF 特征、类 Harr 特征、面片图像特征描述目标的局部兴趣区,获得目标的稀疏特征表示,进而采用通用的像素级显著检测算法,实现特定的目标级显著检测,可明显提高检测精度。

文中采用 K 均值聚类学习车辆目标的视觉特征字典。首先在车辆目标上提取大量面片图像以供学习。在各种光照、姿态、种类的车辆目标图像上,随机选取 $s \times s$ 像素大小的面片图像,面片图像的数量为 N ,将每个面片图像保存为一个 M 维向量 $x_n, M = s \times s$,所有面片图像形成一个 $M \times N$ 的数据集 $X = [x_1, x_2, \dots, x_n]$ 。

为了使聚类算法易于获得 M 维向量上的局部特性,首先把 x_n 标准化为均值为 0、标准差为 1 的向量,并对整个数据集做白化处理:

$$\tilde{X} = E\Lambda^{-1/2}E^T X \quad (2)$$

式中: Λ 与 E 来自对协方差矩阵的特征值分解:

$$E\{XX^T\} = E\Lambda E^T \quad (3)$$

$\Lambda^{-1/2}$ 是对角阵,对角线上为特征值矩阵 Λ 中对角元平方根的倒数:

$$\Lambda^{-1/2} = \text{diag}(\lambda_1^{-1/2}, \lambda_2^{-1/2}, \dots, \lambda_M^{-1/2}) \quad (4)$$

经白化处理后,协方差矩阵变为单位阵:

$$E\{\tilde{X}\tilde{X}^T\} = I \quad (5)$$

经预处理后,采用 K 均值聚类算法进行无监督学习,设定 K 个聚类中心,将计算后得到的 K 个聚

类中心作为视觉特征字典 $D=[d^{(1)}, d^{(2)}, \dots, d^{(K)}]$ 。

特征学习所得特征主要由不同宽度、方向、长度的边条构成,也包括稍小的点与稍大的斑,其中部分特征的方向、长度完全相同,只是位置略有偏移,说明特征存在冗余,但这种冗余并不影响目标识别的准确性。

用习得的视觉特征字典 $D=[d^{(1)}, d^{(2)}, \dots, d^{(K)}]$ 对原始图像进行编码,得到 K 个特征响应 $z_k(k=1, 2, \dots, K)$,再对各响应计算其显著性。为快速计算,对原始图像编码时直接采用二维卷积:

$$z_k = d^{(k)} * I \quad (6)$$

式中: I 表示原始图像; $*$ 为二维卷积运算,得到的响应 z_k 和原始图像尺寸一样,将其视为一张图像,进行显著性检测,并将 K 个显著图像累加合并:

$$s_{\text{conv}} = \sum_{k=1}^K \text{SRD}(z_k) \quad (7)$$

式中: SRD 代表某种显著性检测算法; s_{conv} 为合并后的结果,在 s_{conv} 上响应强的区域即为待识目标区域。最后,采用谱残差法(spectral residual-SR)进行显著性检测^[14]。

2 目标判别

目标判别性能的好坏,很大程度上依赖于参与识别的视觉显著特征的类别区分能力,因此进行恰当的特征选择尤为重要。对目标判别问题,文中采用集成学习方法,如 AdaBoost、GentleBoost 的二类版本,逐步选择特征,构建级联分类器,实现目标/非目标的二类识别,该级联式结构具有由粗到精的分类能力,在保证目标检测精度的前提下可显著提高计算速度^[15]。

对车辆的判别是将待识别目标区域分为一张待分类的图像,然后将各个面片特征响应形成一张图像,对应一个特征向量。对待识别目标图像的各个面片进行编码表示。若图像 I 大小为 $w \times w$ 像素,以 1 像素为步长取 $s \times s$ 像素的面片,从 I 中取 $(w-s+1) \times (w-s+1)$ 个面片。每个面片记为 x_i ,采用基元数为 K 的特征字典进行编码,形成 $F(x_i)=[f_1(x_i), f_2(x_i), \dots, f_K(x_i)]$ 的 K 维表示。

对于目标识别,通过 K 均值聚类可得到 K 个基元的字典 D ,同样可计算聚类归属的 0~1 编码:

$$f_k(x_i) = \begin{cases} 1 & \text{if } k = \text{argmin}_j \|d^{(j)} - x_i\|_2^2 \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

可以采用公式(8)的一些编码方式:

$$f_k(x_i) = \max\{0, \mu(z) - z_k\} \quad (9)$$

式中: $z_k = \|x - d^{(k)}\|_2^2$, $\mu(z)$ 是 $z=[z_1, z_2, \dots, z_K]$ 的均值。

3 实验结果与分析

自然场景下地面车辆目标识别的关键是建立目标视觉显著特征字典,这需要采集大量目标图像,考虑到真实目标与目标模型的高度相似性,可以基于多种车辆目标模型,采集大量目标图像,或者利用国际上相关研究机构公开的目标识别图像数据库如 Caltech101、Caltech256、ImageNet、Norb、LabelMe 等,通过统计学习构建预定目标类别的视觉显著特征字典如图 3 所示,可以看出文中所提出的算法能够有效抑制噪声和空洞现象,效果更好。

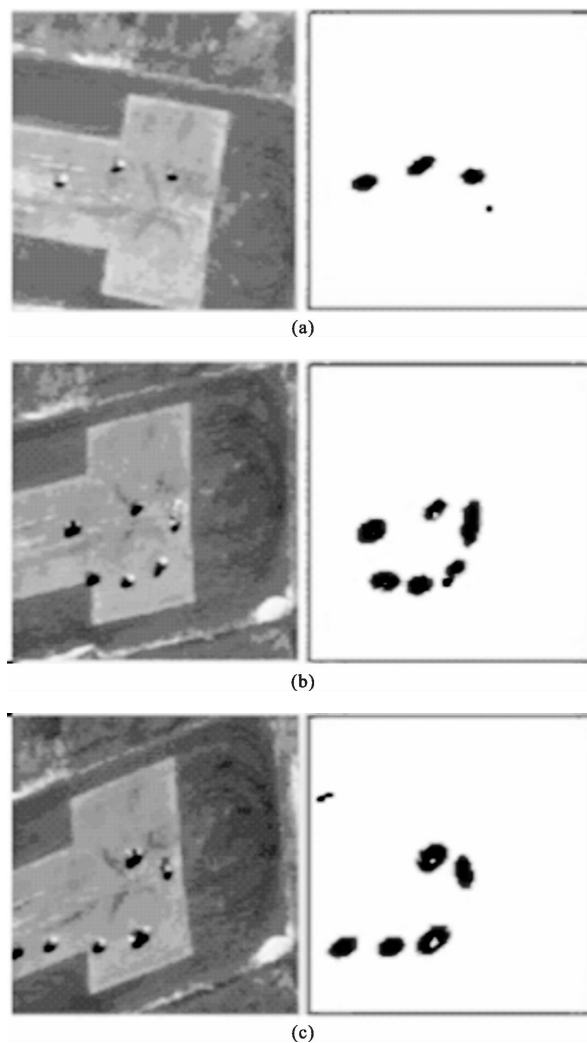
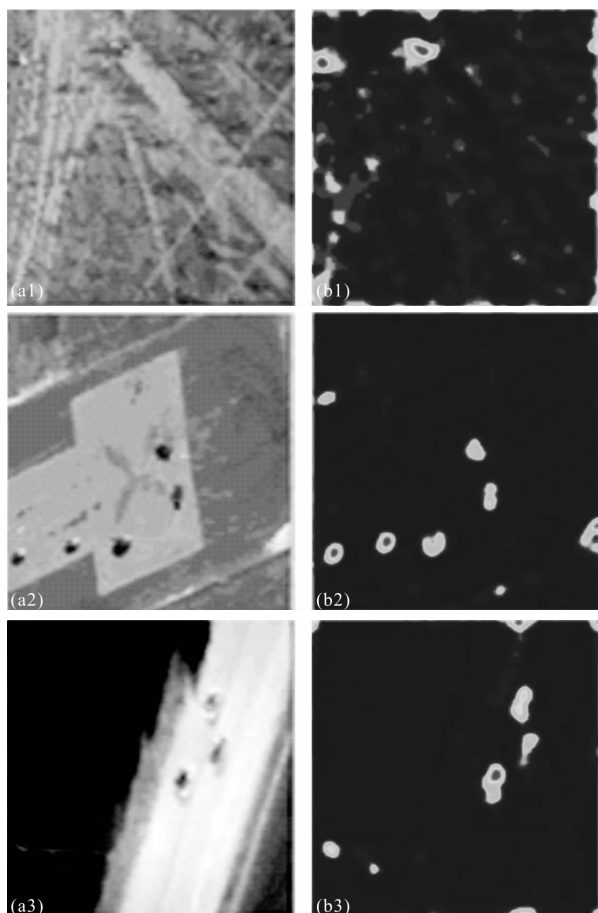


图 3 运动目标检测结果图

Fig.3 Moving target test result graph

为了验证文中所提目标判别算法的性能,采用公开的航拍视频数据进行测试如图 4 所示,这类数据能反映出目标判别算法的实际能力,验证相关算法在不同成像条件与不同复杂度背景下的性能参数,体现出对运动目标在复杂场景下的鲁棒性。



(a) 输入 (b) 输出
(a) Input (b) Output

图 4 统计学习特征响应结果图

Fig.4 Statistical learning feature response chart

文中多特征融合的车辆目标检测结果如图 5 所示,方框为正确检测,圆形框为漏检,三角框为误检。经验证在场景复杂多变、图像对比度低、目标响应弱、图像质量退化和干扰物遮挡等多因素影响的情况下,文中的识别方法具有较好的稳健性。

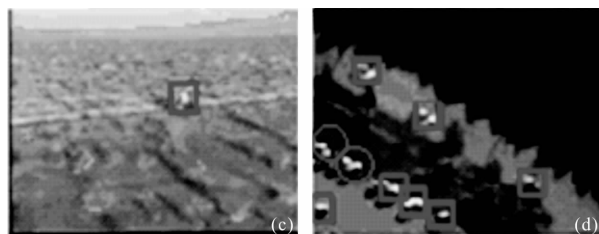
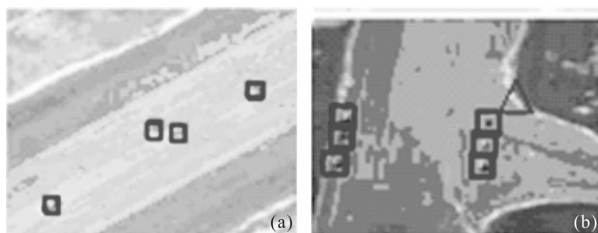


图 5 文中研究方法的检测结果

Fig.5 Test results of the proposed method

4 结论

文中针对自然场景下车辆运动目标识别问题,采用了一种由粗到精的分析策略,旨在保证目标识别性能的同时提高计算效率;同时为提高目标识别性能,采用特征学习建立视觉显著特征词典,然后运用集成学习方法选择判别特征,实现对目标的准确分类和判别。实验表明该方法具有较好的目标判别性能,能较好地解决自然场景下运动车辆的识别问题。

参考文献:

- [1] Shindler Luca. Using optical flow equation for particle detection and velocity prediction in particle tracking [J]. *Applied Mathematics and Computation*, 2012, (218): 8684–8689.
- [2] Shi Jie, Li Yinya, Qi Guoqing, et al. Machine vision based passive tracking algorithm with intermittent observations [J]. *Journal of Huazhong University of Science and Technology (Natural Science Edition)*, 2017, 45(6): 33–37. (in Chinese)
- [3] Shan Zhongde, Zhang Fei, Ren Yongxin, et al. On line detection technology of the hardness of cast iron parts based on machine vision [J]. *Journal of Mechanical Engineering*, 2017, 53(1): 158–164. (in Chinese)
- [4] Chen Junjun, Xu Bing. Research on machine vision image clearness under fog and haze weather conditions [J]. *Computer Engineering*, 2017, 43(2): 280–285, 292. (in Chinese)
- [5] Ding Peng, Zhang Ye, Jia Ping, et al. Ship detection an sea surface based on multi-feature and multi-scale visual attention [J]. *Optics and Precision Engineering*, 2017, 25(9): 2462–2468. (in Chinese)
- [6] Fu Wenlong, Zhang Mengjie. Low-level feature extraction for edge detection using genetic programming [J]. *IEEE Transactions Cybernetics*, 2014, 44: 1459–1472.
- [7] Chen Mingsheng, Liang Guangming, Sun Jixiang, et al. Fast moving object detection method using temporal spatial back

- ground model[J]. *Journal of Image and Graphics*, 2011, 16(6): 1002–1007. (in Chinese)
- [8] Song Yanming, Pan Zhikang, Meng Xiaochen, et al. Multispectral visual image processing method based on adaptive regulation of humanoid eye [J]. *Infrared and Laser Engineering*, 2017, 46(9): 0910001. (in Chinese)
- [9] Yang Hang, Wu Xiaotian, Wang Yuqing. Image restoration approach based on structure dictionary learning [J]. *Chinese Optics*, 2017, 10(2): 207–218. (in Chinese)
- [10] Sun Qian, Feng Hao, Zeng Zhoumo. Recognition of optical fiber pre-warning system based on image processing [J]. *Optics and Precision Engineering*, 2015, 23(2): 334–341.
- [11] Zhang Zhi, Lin Xuling, He Hongyan. Filtering method for remote sensing image based on quantum mechanics [J]. *Infrared and Laser Engineering*, 2016, 45(12): 1204005. (in Chinese)
- [12] Zheng W M, Zhou X Y, Zou C R, et al. Facial expression recognition using kernel canonical correlation analysis [J]. *IEEE Transactions on Neural Networks*, 2014, 17(1): 233–238.
- [13] Chang C C, Lin C J. LIBSVM: A library for support vector machines [J]. *Transactions on Intelligent Systems and Technology*, 2011, 2(27): 1–27.
- [14] Hou X, Zhang L. Saliency detection: A spectral residual approach [C]//IEEE International Conference on Computer Vision and Pattern Recognition, 2007: 1–8.
- [15] Chen Ying, Zhu Ming. Multiple sub-histogram equalization low light level image enhancement and realization on FPGA [J]. *Chinese Optics*, 2014, 7(2): 225–233. (in Chinese)