

# 主-从 PTZ 视频检测软件模块化设计及架构

张志杰<sup>1</sup>, 杨 硕<sup>2</sup>

(1. 西南民族大学 计算机科学与技术学院, 四川 成都 610000;  
2. 沈阳化工大学 计算机科学与技术学院, 辽宁 沈阳 110142)

**摘要:**前方路面车辆的检测和跟踪是智能交通系统中的重要问题。使用单定焦相机的系统,对处于复杂环境下的车辆视野有限,远处目标的特征细节很难捕获,较宽视野区域中的特征分辨率较低;而要获得较高分辨率,就只能缩小监控区域的视野,因此检测效果较差。为解决单定焦相机视野与分辨率之间的冲突,文中引入主-从 PTZ 系统加以改进。其中,主系统采用固定的广角镜头以获得较宽的视野,同时采用高清晰度的 PTZ 相机作为从系统,提升较宽的视野中特定目标区域的分辨率。文中提出并实现了一个主-从 PTZ 视觉检测软件的模块化解决方案。该方案利用主系统进行广角搜索,利用从 PTZ 增强特征细节。测试结果表明,该系统能够实现较宽视野中特定目标的较快检测,提高目标图像细节分辨率,改进目标特征细节检测率,同时具有简单、快速、高效、易于二次开发等优点。

**关键词:**视觉检测;主-从 PTZ;系统框架;智能交通;设计模式

中图分类号:TP31

文献标识码:A

文章编号:1673-629X(2016)09-0114-05

doi:10.3969/j.issn.1673-629X.2016.09.026

## Modular Design and Architecture of Master-slave PTZ Video Detection Software

ZHANG Zhi-jie<sup>1</sup>, YANG Shuo<sup>2</sup>

(1. College of Computer Science & Technology, Southwest University for Nationalities,  
Chengdu 610000, China;

2. College of Computer Science & Technology, Shenyang University of Chemical Technology,  
Shenyang 110142, China)

**Abstract:** The detecting and tracking of vehicles on the road is an important problem in intelligent transportation system. To the vehicle in complicated environment, only using single fixed-focus camera system, the details of remote environment around the vehicle are hard to capture for limited vision, and the effect of detection is poorer for same reason. For single fixed-focus camera system, the common method that could improve the details and higher resolution of remote object is focusing on a smaller vision coverage area. In order to solve the conflict between vision coverage area and resolution, the master-slave Pan/Tilt/Zoom camera system (M-S PTZ) is introduced. The master camera with fixed wide-angle lens is used to achieve a wider field of vision, and the slave one used PTZ lens to enhance the detail of the image and improve the detection rate. A modular video detection system of master-slave PTZ is presented and implemented in this paper. The system of M-S PTZ as solution has many advantages such as faster detection for specific object under field environment, higher resolution for specific target image, and more feature details of given target. At the same time, it is simple, rapid, efficient and easy to conduct secondary development.

**Key words:** vision detection; master-slave PTZ; system framework; intelligent transportation; modular design

## 0 引言

智能交通系统(Intelligent Transportation System, ITS)的目的是实现全方位、大范围内交通运输管理的实时、准确、高效。为此,智能交通系统正在将数据通

讯传输技术、电子传感技术、自动控制技术、数据挖掘技术等进行交叉、集成,力图将先进的信息技术有效地集成运用于整个交通管理系统,实现立体管理<sup>[1]</sup>。

美、欧、日等发达国家基本上完成了 ITS 体系框

收稿日期:2015-11-17

修回日期:2016-03-10

网络出版时间:2016-06-22

基金项目:四川省科技支撑计划(2014GZ0150)

作者简介:张志杰(1972-),男,博士,高级实验师,研究方向为算法设计、信号与图像处理与物联网。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20160622.0845.062.html>

架,现阶段的重点是在发展领域进行大规模的应用。ITS 的应用发展表明,ITS 能够更加高效地利用现有交通设施、保证交通安全、提高运输效率、减少交通负荷和环境污染<sup>[2]</sup>。

ITS 的一个重要特点是与物联网技术关系密切,其具体应用均围绕信息的收集、处理、发布、交换、分析、利用等内容,动态交通信息流是 ITS 研发与应用的主线。

高效、智能的视频检测技术,是智能控制也是物联网服务的研发目标之一,具有良好的市场应用前景,且经济效益与社会效益良好。

通过对主-从 PTZ 视觉检测系统的研究,再辅以信息技术、通信技术、控制技术、传感技术和系统综合等主要技术手段,以智能道路、智能路口、智能监控和智能车辆等为主要研究内容,可以进一步降低交通事故率,其应用前景广阔。

## 1 视觉检测

视觉检测,是指利用摄像机作为记录设备,通过对目标图像、目标视频进行预处理,然后再进行后继分析,以获得动态信息流的检测手段<sup>[3]</sup>。

视觉检测能够应用于许多领域:工业自动化中的工程控制,比如工业机器人;医学图像与信号处理,比如超声波视频检测、CT 图像检测等;监控视频的自动筛选,比如人数统计与特定目标的识别;军用民用的各类交通工具、运载武器的自动识别,比如精确制导导弹的目标识别与战场感知。

此外,视觉检测快速发展的一个应用领域是各类无人机(Unmanned Aerial Vehicle, UAV),此类应用不仅包括无人机,另有无人潜水艇、无人车辆、特种作业机器人等,均依赖于基于计算机视觉的视觉检测系统,实现机器自主导航(Machine Navigation or Robot Navigation)、障碍物检测(Obstacle Detection)、机器路线规划(Route Planning)、环境生成(Simultaneous Localization And Mapping, SLAM 或 Concurrent Mapping and Localization, CML)、检测特定的事件等功能。该类视觉检测技术,尤其在 NASA 的太空探索计划中,得到大量的应用,例如:探月车、火星探测器等,其技术水平正在不断进步与提升中。

具体到 ITS 的信息采集方式,主要是自动采集。而现阶段 ITS 自动采集的技术手段,又可以分为磁性检测器、光学检测器、微波检测器、路况传感器及测重传感器等。在 ITS 中,视觉检测的优势在于其形象、直观、易于获取,作为第一手的信息,其可靠性、信息包含的内容都具有明显优势。

而在现有繁忙的 ITS 中,视觉检测存在的主要问

题在于:

(1)精度与价格相关。其可靠性、精确性依赖采集设备的物理特性,且往往与设备的价格成正比例关系,高精度的设备价格也更高。

(2)数据分析、处理的工作量大。作为摄像头采集图像、视频信息,其采集具有自动、全方位、大范围、多角度等特点,数据量与采集设备工作时间成正比;采集的图像、视频信息中,存在大量的冗余信息,需要进行筛选。现阶段,筛选工作仍然在很大程度上依赖于人力,其工作效率与机器数据采集的速度,存在很大差异。

(3)复杂环境下的精确度降低。视频检测的精度、速度,在很多情况下会随着环境的复杂性变化而受到很大的影响。比如,晴朗天气下视频的单独车辆车牌识别,就比雨天、雪天、雾天、复杂背景、快速变化、遮挡阻拦等情况下的精度、速度要高;换而言之,复杂环境下的视频检测还有待改进。

(4)算法设计的局限性。由于复杂环境下的视频检测算法在实际应用中存在诸多局限,因此,ITS 作为视频检测的应用背景,彰显出若干视频检测算法的若干局限性,所有这些局限性为视频检测算法的发展提出了新的要求<sup>[4]</sup>。

## 2 主-从 PTZ 解决方案

视频检测,作为动态信息流采集的重要渠道,一直是研究的热点。而前方路面车辆的检测和跟踪则是 ITS 中动态信息交通流的重要问题,因而,日益受到各国的重视。

因此,解决好复杂环境下周围车辆的检测和跟踪问题是降低事故发生率的一个有效手段。复杂环境分为:自然环境的复杂性(如雨天、雾天、能见度低等)和获取的视频图像的复杂性(如复杂的图像背景、剧烈的光照度变化、遮挡、运动平台下相机的抖动等)。

在现有文献中,基于视频的车辆检测算法可以分为两个步骤:

(1)假设生成:获得图像或视频帧中可能包含车辆的候选区域;

(2)假设检验:用分类器或类似方法测试候选区域是否真实包含车辆。

假设生成一般使用简单的特征和快速的搜索算法将整个图像区域内可能包含车辆的候选提取出来,使用的方法可以分为:基于先验知识的、基于立体视觉的和基于运动信息的。假设检验用于进一步验证候选的正确性,可以分为两种:基于模板(知识)模型的和基于外观模型的<sup>[5]</sup>。

上述经典方法在特定环境下(如白天、高速公路、

简单路况、无遮挡)取得了良好的检测效果,但在复杂环境下,由于可供提取的车辆特征较少并且容易受到干扰,使得实际的检测效果受到了限制。在课题组前期的研究中,主要针对复杂环境并对局部方法、稀疏表达和在线训练等方法做了一些研究,虽然取得了一定成果,但在实际环境下的效果依然不理想。

主-从 PTZ 视频检测方法 (Master-Slave PTZ),是目前较好的解决复杂环境中目标视觉检测的方法。清华大学在 2013 年度国家自然科学基金的研究课题中,就进行过这方面的探索与研究:Cui Zhigao 等的《Master-slave tracking algorithm using dual PTZ cameras based on ground plane constraint》、《A Master-slave Tracking Algorithm Using Two PTZ Cameras》。其主要研究内容,即为一主一从的主-从 PTZ。

主-从 PTZ 视频检测方法,从各类相关文献与自主软硬件研发的结果中,被发现可以从物理结构上一定程度地解决复杂背景、快速变化、遮挡阻拦等复杂环境带来的问题,同时结合专门的去雾、去雪、去雨等算法的设计,可以较好地解决若干视频检测算法的局限性<sup>[6-8]</sup>。

通过分析发现,原有算法效果不理想的主要原因之一是:为了增加检测范围,实验采用了较宽视野的广角镜头,使图像细节丢失较多,不利于特征的提取和分类,从而降低了车辆检测率。因此,在前期研究基础上,文中提出使用主-从-PTZ 相机构建超分辨率图像的方法,以提高复杂环境下的车辆检测率。其中主相机具有固定的广角镜头可以获得较宽的视野,而通过 PTZ 相机可以将图像、视频的具体细节进行保留,以便于特征的提取与分类。

### 3 系统整体框架

计算机视觉系统实现的方法、结构、模块、接口等,很大程度上依赖并取决于其具体应用方向、技术要求等。标准的视觉检测系统一般由图像获取、预处理、特征提取、检测分割、后期处理等功能模块组成。其中,后期处理模块往往与机器学习、数据挖掘、人工智能等内容密切相关,其具体的目的就是要对获得的视频信息进行分类与识别,其典型算法的代表即为 2010 年 1 月《IEEE Transactions on Pattern Analysis and Machine Intelligence》上,Zdenek Kalal 的《Tracking Learning Detection》。

作为一个系统的框架设计,还需要考虑到系统的框架、模块的功能以及划分、接口的定义与数据的统一格式、算法库的架构等因素,同时,算法库要保持较好的可扩展性,易于二次开发。其中,比较有挑战性的内容为算法库中不同视觉检测算法的实现、效果比较、分

析与改进。

具体到该主-从 PTZ 视觉检测系统,从现阶段功能需求进行划分,主要包含如下模块:运动目标检测模块、运动目标分类模块、车辆跟踪模块、车牌检测模块、车牌字符分割模块、车牌字符识别模块以及其他辅助功能模块。

其中,运动目标检测模块主要进行运动目标的捕获,采用尽可能简单的算法排除掉尽可能多的非目标对象。运动目标分类模块主要进行运动目标的分类,其分类的依据主要是若干常用的图像预处理、特征提取,由于视频的工作量,仍然优先选择简单高速的算法。车辆跟踪模块与检测模块主要进行基于不同特征的车辆检测分类,作为系统的分类检测依据,主要是进行特征的综合提取。车辆字符分割与字符识别模块,是对需要捕获车牌信息的特定检测对象进行车牌的识别等处理。

综上所述,该主-从 PTZ 视觉检测系统主要结构如图 1 所示。

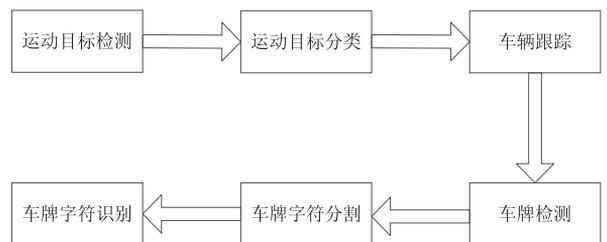


图 1 主-从 PTZ 视觉检测系统主要结构示意图

其中,运动目标检测模块主要实现的方法为 Visual Background Extractor,以下简称为 VIBE<sup>[9-11]</sup>。

VIBE 是一种像素级的视频背景检测算法,其算法简单、实用,对于硬件内存的占用较少,因此,可以达到较高的处理速度和理想的处理结果。其具体实现如下:首先用第一帧初始化 VIBE 算法,该函数只调用一次,其中 input 是函数输入,为视频的第一帧,24 位 RGB 彩色图像。

在视频的随后帧中,每帧或隔帧(跳帧)调用处理函数,完成背景差分,并对运动目标进行检测。

运动目标分类模块主要实现的方法如下:检测该目标底部是否有阴影(该阴影就是本影)。因为车辆在白天时车底必出现本影,不管是阴天还是晴天,而且车底本影的长度与车宽一致<sup>[12-14]</sup>。

其具体的实现如下:阴影像素的灰度值远低于车身灰度值,充分利用车底阴影作为特征进行特征提取,就可将车辆和行人(或自行车)分开。

车辆跟踪模块主要实现的方法如下:包围框覆盖法。

由于车辆属于刚性目标(本身不会形变),且速度不会发生突变,在屏幕中的视角一般也不变,因此,对

车辆的跟踪就可以使用带自适应参数的包围框覆盖法<sup>[15-16]</sup>。

其中,自适应参数主要解决某些干扰会导致包围框偏离正常位置,或者过小、过大,从而导致跟踪失败的问题。

车牌检测模块主要实现的方法如下:定位、特征点检测与特征点提取。其中,定位主要是为了提升特征提取的角度;而特征提取车牌主要的问题是车辆高速行驶时造成的车牌图像、视频模糊。

为尽可能排除干扰,采用综合特征提取方法,结合了颜色特征、梯度特征、形状特征、几何特征等,提高特征点提取的精确度。

车牌字符分割模块主要实现的方法为滤波器字符分割。

由于车牌的有用信息要包括0到9的数字、A到Z的字母,外加代表省的字符,有用字符一共为7个,所以,车牌检测的主要功能就是在包含车牌的图像中找到车牌的位置并将7个字符分割出来<sup>[17-18]</sup>。

这里采用一个8条竖线的“栅栏状”滤波器,对投影图进行滤波。通过改变“栅栏线”的间距和位置可得到最大响应,根据得到的最大响应就可以将字符彼

此分开。

车牌字符识别模块主要实现的方法为模板匹配。即将待识别字符与事先获取的字符模板库(10个数字、24个英文字母和31个汉字)逐一对比,找到最近的作为识别结果<sup>[19-20]</sup>。

由于类别较少,且现代计算机的处理速度很高,模板匹配的实时性很好,车牌字符识别模块完全适用于视频连续识别。同时,模板匹配在各类图像上的识别率表现比较稳定,因此模板匹配仍然具有很高的使用和研究价值。

另外,该系统还包括车型识别、驾驶员异常行为检测、特定目标跟踪等功能。

#### 4 实验结果与分析

该主-从 PTZ 视觉检测系统中,主相机具有较宽的视野,从相机可以获得图像的细节。经过初步研究发现,该系统可以在一定程度上提高现有算法的检测率,但具体的效果的改进评估,则还需要进一步的数学证明、实验和论证。

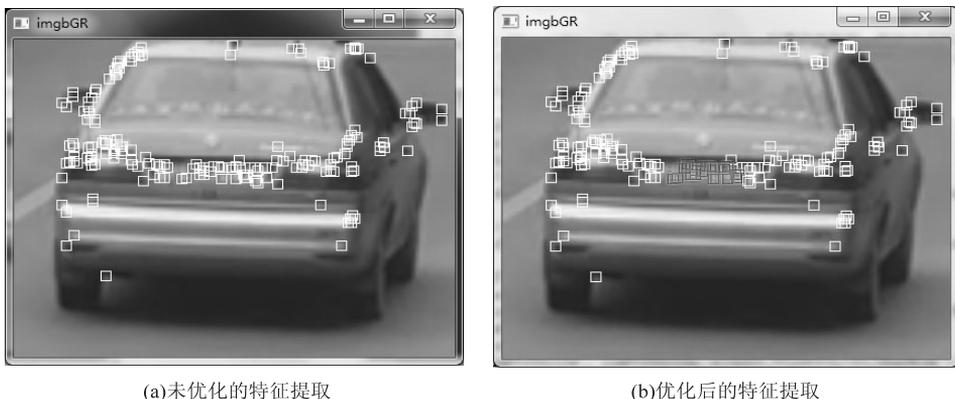
该主-从 PTZ 视觉检测系统部分运行结果见图2。



图2 运动目标检测结果示意图

对运动目标检测之后,还需要进行本影检测,因此,良好、可靠、抗干扰的运动目标检测就成为先决条件,其质量的优劣将直接影响到后面的运算结果。

特征提取的关键在于根据特征的特点,将干扰特征点进行过滤。从图3可见,该优化特征提取方法效率相当高。



(a)未优化的特征提取

(b)优化后的特征提取

图3 车牌特征检测与提取示意图

## 5 结束语

该主-从 PTZ 视频检测系统主要依靠实时弱标定算法,基于局部表达模型对车辆检测局部特征的检测效果进行改进。实验结果表明,该系统能够进行基于局部特征的车辆视觉效果的检测,同时进行有效特征的检测与分析,减少了算法的运算量,其运算速度较快,对于复杂环境的抗干扰性有所提高,且其运算的精准度也有所提升。下一步将继续进行主-从 PTZ 视频检测算法的相关研究。

### 参考文献:

- [1] Wang F Y. Scanning the issue and beyond: five transportations in one—a new direction for ITS from qingdao [J]. IEEE Transactions on Intelligent Transportation Systems, 2015, 16 (5): 2310–2317.
- [2] Rimmer P J. Port dynamics since 1965: past patterns, current conditions and future directions [J]. Journal of International Logistics and Trade, 2007, 5(1): 75–97.
- [3] Bertozzi M, Broggi A. GOLD: a parallel real-time stereo vision system for generic obstacle and lane detection [J]. IEEE Transactions on Image Processing, 1998, 7(1): 62–81.
- [4] Sivaraman S, Trivedi M M. Looking at vehicles on the road: a survey of vision-based vehicle detection, tracking, and behavior analysis [J]. IEEE Transactions on Intelligent Transportation Systems, 2013, 14(4): 1773–1795.
- [5] Bertozzi M, Broggi A, Cellario M, et al. Artificial vision in road vehicles [J]. Proceedings of the IEEE, 2002, 90(7): 1258–1271.
- [6] Everts I, Sebe N, Jones G A. Cooperative object tracking with multiple PTZ cameras [C]//Proc of the international conference on image analysis and processing. [s. l.]: [s. n.], 2007: 323–330.
- [7] Kumar P, Dick A, Sheng T S. Real time target tracking with pan tilt zoom camera [C]//Proc of digital image computing: techniques and applications. [s. l.]: IEEE, 2009: 492–497.
- [8] Liao H C, Chen W Y. A dual-PTZ-camera system for visual tracking of a moving target in an open area [C]//Proc of 11th international conference on advanced communication technology. [s. l.]: IEEE, 2009: 440–443.
- [9] Barnich O, van Droogenbroeck M. ViBe: a powerful random technique to estimate the background in video sequences [C]//Proc of IEEE International conference on acoustics, speech and signal processing. [s. l.]: IEEE, 2009: 945–948.
- [10] Barnich O, van Droogenbroeck M. ViBe: a universal background subtraction algorithm for video sequences [J]. IEEE Transactions on Image Processing, 2011, 20(6): 1709–1724.
- [11] Feng W, Liu R, Jia B, et al. An efficient pixel-wise method for moving object detection in complex scenes [C]//Proc of 10th IEEE international conference on advanced video and signal based surveillance. [s. l.]: IEEE, 2013: 389–394.
- [12] Johansson B, Wiklund J, Forsén P E, et al. Combining shadow detection and simulation for estimation of vehicle size and position [J]. Pattern Recognition Letters, 2009, 30(8): 751–759.
- [13] Song X, Nevatia R. Detection and tracking of moving vehicles in crowded scenes [C]//Proc of IEEE workshop on motion and video computing. [s. l.]: IEEE, 2007.
- [14] Messelodi S, Modena C M, Zanin M. A computer vision system for the detection and classification of vehicles at urban road intersections [J]. Pattern Analysis and Applications, 2005, 8(1): 17–31.
- [15] Lempitsky V, Kohli P, Rother C, et al. Image segmentation with a bounding box prior [C]//Proc of 12th international conference on computer vision. [s. l.]: IEEE, 2009: 277–284.
- [16] Assarsson U, Moller T. Optimized view frustum culling algorithms for bounding boxes [J]. Journal of Graphics Tools, 2000, 5(1): 9–22.
- [17] Casey R G, Lecolinet E. A survey of methods and strategies in character segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996, 18(7): 690–706.
- [18] Fujisawa H, Nakano Y, Kurino K. Segmentation methods for character recognition: from segmentation to document structure analysis [J]. Proceedings of the IEEE, 1992, 80(7): 1079–1092.
- [19] Lalonde M, Beaulieu M, Gagnon L. Fast and robust optic disc detection using pyramidal decomposition and Hausdorff-based template matching [J]. IEEE Transactions on Medical Imaging, 2001, 20(11): 1193–1200.
- [20] Battiato S, Farinella G M, Furnari A, et al. Vehicle tracking based on customized template matching [C]//Proc of international workshop on ultra wide context and content aware imaging. [s. l.]: [s. n.], 2014: 755–760.