

DOI:10.19651/j.cnki.emt.1802234

支持行人检测的智能车载监控终端

王颖¹ 金若辰² 金志刚³

(1.61660部队 北京 100089; 2.东北大学 材料科学与工程学院 沈阳 110819; 3.天津大学 电气自动化与信息工程学院 天津 300072)

摘要:大量交通事故的分析表明行人是交通的弱势参与者。为了达到保护行人安全的目的,需要提高车辆的智能水平,通过具备智能监控能力的车载终端来协助发现处于事故危险中的行人,及时进行预警。针对车载智能监控的需要,基于聚合积分通道模型,设计了具有行人检测能力的智能移动视频监控终端。面向TI DM365芯片组与TVP5158组成的多路视频处理平台,设计了基于聚合积分通道模型的行人检测算法,并且进行了典型场景实验。实验结果表明,该终端可以通过移动网络实现远程视频监控,并且有效实现行人检测,提高了车辆智能水平与交通安全性。

关键词:行人检测;视频监控;聚合积分;移动通信

中图分类号: TP391.4 **文献标识码:** A **国家标准学科分类代码:** 510.1050

Vehicle video surveillance terminal with pedestrian detection

Wang Ying¹ Jin Ruochen² Jin Zhigang³

(1. Troops 61660, Beijing 100089, China; 2. School of Material Science and Engineering, Northeast University, Shenyang 110819, China; 3. School of Electrical Engineering, Tianjin University, Tianjin 300072, China)

Abstract: Analysis of a large number of traffic accidents showed that pedestrians are the vulnerable participants in traffic. In order to protect the safety of pedestrians, it is necessary to improve the intelligent level of vehicles, and assist the detection of pedestrians in danger of accidents through on-board terminals equipped with intelligent monitoring capabilities, so as to timely give early warning. The intelligent mobile video monitoring terminal with pedestrian detection capability has been designed based on the aggregated integral channel model to meet the needs of vehicle-mounted intelligent monitoring. A pedestrian detection algorithm based on the polymer integral channel model has been designed for multichannel video processing platform composed of TI DM365 chipset and TVP5158, and typical scene experiments were conducted. Experimental results showed that the terminal can implement remote video monitoring via mobile network, and effectively complete pedestrian detection, which improves the level of vehicle intelligence and traffic safety.

Keywords: pedestrian detection; video surveillance; aggregated channel; mobile communication

0 引言

据统计,2016年全国发生了800多万起交通事故,其中涉及人员伤亡的事故有20多万起。交通事故导致伤亡人员中约1/3是行人,迫切需要车辆增加行人检测装置来减少事故伤害与损失。

机器视觉算法与视频监控技术的进步为机动车采用智能监控系统实现行人检测与事故预警提供了可能^[1]。为保证行人安全,不仅需要车辆行驶过程中可能发生事故的行人和事件进行检测及跟踪,提取有用信息报告给驾驶人进行处理,还需要实时进行网络视频监控,远程保存现场情

况与关键数据,为事故分析提供依据。

1) 车载视频监控现状

安防厂商主要关注固定场所安防的视频系统,少数厂商推出了移动视频监控系统。如深圳锐明视讯公司面向公交车车厢内部乘客安全,开发了“公交智能调度视频监控系统”。但是安防厂商设计的车载视频系统存在智能水平低、成本高、部署复杂等问题^[2]。从技术角度分析,主要问题在于:

(1)车辆一般处于运动状态,只能采用无线网络与后台服务器通信,需要车载终端具有自适应网络的编码能力和低视频码率;

(2)多通道的视频采集与行人检测之类的视频分析需要较高的分辨率,需要高码率;

(3)多功能的前提下,基于比较成熟的硬件核心设计车载终端,以平衡低成本与高可靠性。

2)行人检测研究现状

行人检测采用机器视觉与人工智能技术,在获取的交通监控视频中及时提取出行人信息,广泛应用于无人驾驶、智能交通等领域^[3]。行人检测包括背景检测、感兴趣区域提取等关键步骤。在快速、高效行人检测基础上,需要实时进行预警。

感兴趣区域提取算法的主要性能判据是漏检率与误检率^[4]。传统行人检测方法一般只使用了单视点图像信息,通过提取颜色、纹理和运动等特征来区分前景与背景^[5-7]。为了更有效进行特征提取,文献[8]提出了采用双视点图像的视差信息的行人检测新机制,并设计了神经网络分类模型。Keller等^[9]在稠密视差图基础上建立了更细化的道路模型,设计了基于HOG特征的分类器进行行人判定。双目图像信息更好地突出了行人的深度及位置特征,更适合进行交通分析与预警^[10]。

通过文献分析,发现融合双目视觉进行行人检测的车载网络视频监控系统的论文较少。本文提出支持无线网络传输、实现多路视频监控并具有行人检测能力的车载视频监控终端与行人检测算法。针对车辆网络视频监控与行人预警等需要,基于TI DaVinci平台^[11],提出了具有行人检测能力的车载智能监控终端设计方案。终端的硬件方案采用的主处理器是支持H.264编解码的TMS320DM365,并配置了视频协处理芯片,用于实现行人检测算法。为了实现多路视频采集编码和基于双视点的行人检测,设计了基于TVP5158视频解码芯片的多路采集版与配套软件;为了对车载终端数据进行远程网络备份与管理,设计了基于RTP协议的H.264码流打包软件^[12],实现了事故现场数据的可靠网络存储。为了保证复杂场景下行人检测的有效性,在双视点视频数据与协处理芯片支持下,提出了基于汇聚积分通道的行人检测模型,实现了车载终端硬件加速的行人检测。利用协处理器输出的图像金字塔简化了行人最小尺度信息,大大降低了潜在兴趣区发现的计算量,达到预期的行人检测率。

1 车载视频监控终端设计

为了实现车载本地视频采集与编码、网络视频传输与保存、车载行人检测等功能,设计了包含车载终端、流媒体服务器和中心管理平台组成车载智能视频监控系统整体方案。本文研究重点为车载智能终端设计与行人检测算法。车载终端是系统的核心,其硬件平台设计决定了整体性能。该终端硬件包括核心板与扩展板,如图1所示。终端核心板基于TMS320DM365处理器,扩展了128MB的内存,并且采用AEMIF接口扩展了512MBFlash存储。多路视

频、无线网络和多路USB等外设通过核心板连接扩展板来实现。

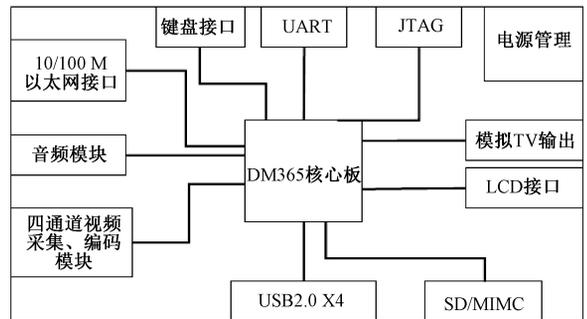


图1 车载终端硬件框图

1)四通道视频采集编码模块。采用TVP5158音视频编解码芯片支持4个摄像头同时进行视频信息采集编码。

2)USB集线器模块。实现4路USB并行工作,同时接入移动存储、无线网卡等设备。

3)LCD接口电路。使用低压差分方式(LVDS),外接LCD显示器,实现本地视频的播放。

另外还包含以太网接口、JTAG、UART、SD/MMC等支持电路,以进行电路板开发与调试。

TMS320DM365主芯片属于ARM+VICP处理器。要让整个硬件系统正常启动,必须遵守该处理器的启动过程,并设计适合自主开发硬件的驱动程序和定制化操作系统。根据TI的参考设计,需要分步烧写UBL、U-BOOT和内核。UBL、U-BOOT可以通过CCS和仿真器烧写到Flash中。Linux内核和文件系统的烧写通过U-BOOT指令完成。

UBL、U-BOOT和Linux内核都是与硬件电路板的内存和外设情况相关,需要生成符合本系统的程序才能让硬件系统运行。由于核心板的内存设计与DM365的参考设计一致,所以UBL和U-BOOT可以直接使用DVSDK提供的版本。但是扩展外设与参考设计有较大的差别,所以需要根据系统情况裁减内核和添加必须的设备驱动支持。对于本车载终端来说,最主要就是需要自主设计程序以驱动4路视频采集进行编码的TVP5158芯片组。

为了解决车载终端本地视频分析对高清视频的需求和网络传输对码率的带宽限制之间的矛盾,采用双码流设计。高码率流用于本地存储和行人检测,低码率码流用于网络传输。TVP51584路视频采集中采用4路D1,形成第1路码流,用于本地显示、行人检测等处理。在系统得到上传指令后,对第1路码流可以通过VPFE的数据处理硬件模块IPIPE Resizer,对原始码流能完成缩放为CIF,得到第2路码流,压缩打包后用于网络传输。车载终端的码流框图如图2所示。

软件系统的主要模块有视频采集、视频编码、视频传

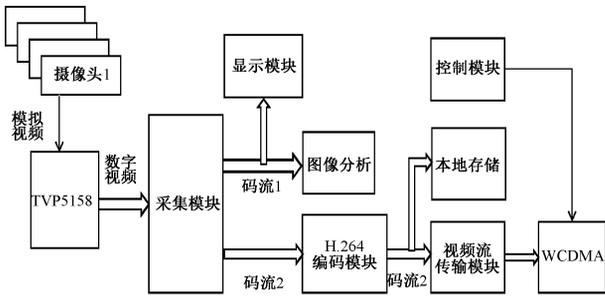


图 2 车载终端码流框图

输、本地显示、本地存储和行人检测模块。

1) 视频采集模块

视频采集模块基于 MCVIP 技术,对 4 路 D1 视频数据进行采集。采集的码流可分为 3 路:1 路用于本地显示,1 路用于本地存储,1 路根据需要可以发送到网络传输模块。

2) 视频编码模块

视频编码模块采用基于 DM365 协处理器的 H.264 编码方式,对感兴趣的区域采用更加精细的编码结构,对背景等节省字节的编码。

3) 视频传输模块

视频传输模块,负责 H.264 码流的打包和无线网络传输,可根据网络情况动态调节帧率。

4) 本地显示模块

本地显示模块主要是为了司乘人员查看车内情况,支持单独 1 路的查看和 4 路同时显示。

5) 本地存储模块

本地存储是将编码后的文件保存到本地硬盘,便于司法机关日后取证,为了方便管理,需要对存储的视频标记上时间戳和摄像头 ID。

6) 行人检测模块

行人检测模块利用 DM365 协处理器的硬件处理能力对感兴趣区信息及时检测行人对象。

2 多通道视频采集驱动的开发

在 Linux 系统中,应用程序通过特定函数库向内核提出操作请求,而内核则依靠设备驱动程序来与各种不同的设备进行交互。设备驱动程序是硬件和应用软件之间纽带,使得应用程序只需要调用系统的 API 就可以让硬件完成相应的工作。硬件驱动程序的开发是应用程序工作的前提。本设计采用 DVSDK 开发,其中内核包含了对 DM365 评估版的驱动程序。对比本设计,需要完全重新添加的驱动主要是支持 TVP5158 多通道视频采集的驱动和对无线网络拨号支持的驱动。

TVP5158 是 4 通道高质量 NTSC/PAL 视频解码器, TVP5158 对达芬奇视频处理器提供无缝的视频和音频接口。可通过 I²C 串行总线对 TVP5158 进行编程控制,其内

部结构功能如图 3 所示。

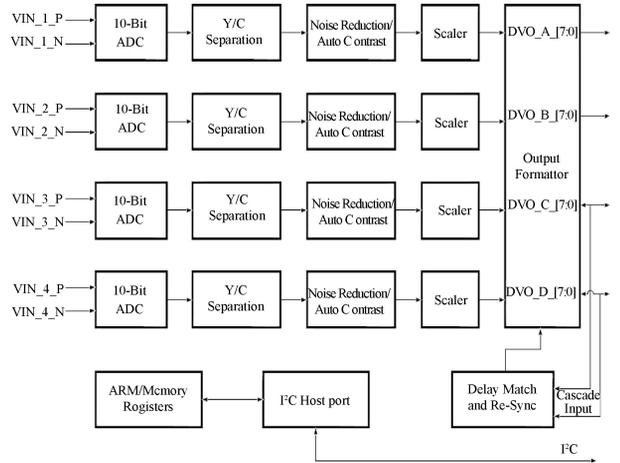


图 3 TVP5158 功能框图

TVP5158 视频输出端口支持 8 位 BT656 和 16 位 YCbCr4;2:2 信号输出。TVP5158 支持无交织输出模式、像素交织输出模式和行交织输出模式。在交织模式,来自各个解码通道的视频输出数据被复用在一起,然后从一个 8 位或 16 位的端口输出。交织数据帧使用 SAV 作为帧的开始符号,EAV 作为帧的结束符号。按照上述结构通过 TVP5158 产生的包含多路视频数据的帧被称为“超帧”。为了能顺利的解出“超帧”中各路数据,在每一行 SAV 的后面添加 4 字节的开始码(SC3;SC0)。

在上述硬件平台和多路视频编码配置下,针对该开发板设计了相应的驱动程序,并且调试成功。该驱动程序的主要流程如图 4 所示。

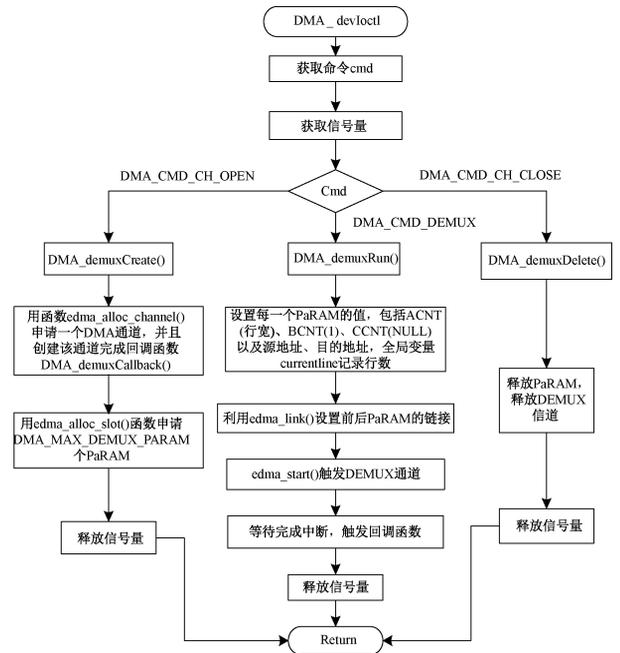


图 4 TVP5158 驱动主要流程

3 行人检测算法

为了充分利用 DM365 的硬件视频处理能力,提出了基于双目视点图像的车载视频行人检测算法,算法主要流程如图 5 所示。该算法由训练阶段与检测阶段组成。

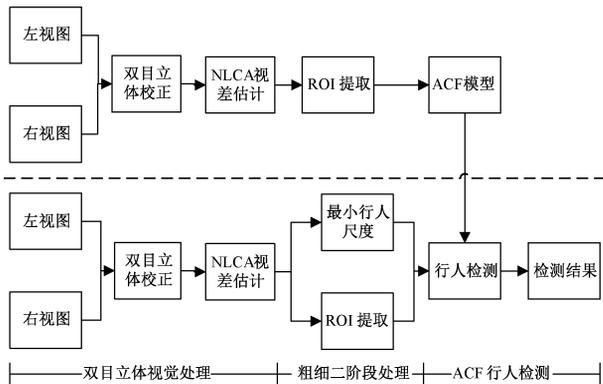


图 5 双目视觉行人检测算法框图

1) 训练阶段

基于双目视点图像进行立体校正与视差估计,获取感兴趣区域、提取聚合积分通道特征,训练弱分类器级联组成的 ACF 分类模型。

2) 检测阶段

基于不同区域生长阈值分别获取 ROIs 与最小目标行人尺度,设定图像特征金字塔结构。基于聚合积分通道特征分类模型,进行行人检测。

下文给出基于聚合积分通道特征的行人检测方法。

对于已知的行人潜在窗口,积分通道是其线性与非线性变换的映射,如灰度通道、LUV 色彩通道、梯度幅值与梯度直方图通道等,记作 $C = \Omega(I)$ 。ACF 计算上述通道图像上固定尺寸矩形窗口内像素值之和,作为分类特征。基于上述特征构建候选特征池,采用 Adaboost 算法筛选并组合最具区分度的特征,构成强分类器^[13]。为加快训练速度,在每个弱分类器后添加门限阈值进行判断,在输出检测系数低于阈值的情况下,提前终止某些负测试样本的检测过程。

传统行人检测方式的特征提取多采用 sliding window^[14-16]方法,在整幅图像上移动所有可能尺寸的检测窗口,输出多种尺度构建的图像特征金字塔。在训练分类模型与测试样本特征的计算上,首先,通过大量实验权衡高低分辨率下区分度最好的最佳尺度,训练最佳尺度下图像的单一模型,随后,在检测阶段,需要将输入图像调整 N 次,并分别计算 N 种尺度下的输入特征。由于自然图像的分形统计性质,可基于某一尺度估测相邻尺度下的特征,即将测试样本窗口尺度调整 N/K 次,仅计算 N/K 种尺度下的图像特征,并通过对上述尺度下的特征采样,就近逼近其他 $N-N/K$ 种特征。逼近方法依据下式:

$$C_s = \Omega(I_s) \tag{1}$$

$$C_{s'} \approx R(C_s, s'/s) \times (s'/s)^{-\lambda_\Omega} \tag{2}$$

式中: $C_s, C_{s'}$ 分别表示图像 I 在 Ω 通道中、 s 尺度与 s' 尺度下的积分通道特征; s' 为待逼近尺度; $R(C, s)$ 表示 C_s 在 s'/s 尺度下的采样结果; λ_Ω 为通道 Ω 对应的逼近系数。

为了进一步减小行人检测的搜寻区域与尺度范围,不再在整个图像区域进行滑窗,而是基于双目视点有效提取的 ROI 区域,构建多层图像特征金字塔。结合根滤波器尺度 s_0 , 区域 i 构建的图像特征金字塔层数可确定为:

$$n_i = 2 \log \left(\frac{s_0}{s_{m_i}} \right) \tag{3}$$

此时,每一 ROI 区域输出的图像金字塔层数均不同,从而减少了不必要的尺度计算。

4 实验及结果分析

为验证所提出方法的性能,选用 Ess 等^[14]提供的行人检测数据库作为实验样本。选择 Bahnhof、Crossing、Sunnyday 序列共约 2 000 帧行人图片,图像分辨率为 640×480 。实验过程如下:主要衡量新算法行人检测的准确度和鲁棒性。与单目视点下基于聚合积分通道特征的传统行人检测方法相比较,记录查准率(true positive rate, TPR)和误检率(false positive rate, FPR)两项性能指标。其中,查准率表示被检测出的行人数量占所有行人数量的比例,而误检率表示被误检测为行人的干扰物占所有干扰物的比例。

表 1 所示为基于 ACF 模型的行人检测方法与本文优化方法在 TPR、FPR 指标的差异。

表 1 单双目视点行人检测结果 (%)

项目	Bahnhof		Sunny Day		Crossing	
	TPR	FDR	TPR	FDR	TPR	FDR
单目视点	96.27	5.67	96.18	6.39	95.65	6.30
双目视点	95.85	1.20	95.69	2.01	95.65	1.65

由表 1 可以看出,采用本文方法的行人检测方法在误检率方面得到了明显改善,较为有效地排除了虚假目标与障碍遮挡的干扰。由于视差估计及 ROI 提取在目标较远或遮挡情况下存在一定的误差,前两个序列的查准率出现 1% 左右的细微下降。总体上讲,本文方法的行人检测效果得到了较大改善。在行人运行速度适中的街道、空地等公共区域,可满足实时监测的要求。

5 结 论

本文针对车载智能视频监控的需要,基于 TI 公司的 Davinci 技术采用 DM365 核心处理器以及多路视频芯片组 TVP5158 设计并实现了车载视频终端硬件开发板。在该平台基础上,针对行人检测的需要,设计了基于自适应的优

化聚合积分通道分类模型,以实现行人检测。通过简单、有效的区域提取与分割,减小潜在行人区域搜索范围,通过预判断潜在行人尺度,大大降低特征金字塔的计算冗余度。在大量视频数据与实际场景下的实验结果表明,该车载终端配合行人检测算法对复杂场景、多种行人姿态分布均能达到良好的检测效果。

随着 TPU 等技术的进展与新的人工智能芯片平台的发展,未来可以提出基于 AI 芯片的车载智能监控方案,并且实现云服务检测及预警,更好的保证交通安全。

参考文献

- [1] 齐美彬,李佶,蒋建国,等.改进特征与 GPU 加速的行人检测[J].中国图象图形学报,2018,23(8):1171-1180.
- [2] 常欢,谢永斌.多模车载视频监控系统中负载均衡算法设计[J].电视技术,2018,42(1):90-94.
- [3] 陈飞玲,陈湘军,郁建桥,等.移动视频监控系统设计[J].电子测量技术,2014,37(4):103-107.
- [4] 钱基德,陈斌,钱基业,等.基于感兴趣区域模型的车道线快速检测算法[J].电子科技大学学报,2018,47(3):356-361.
- [5] LLORCA D F, SOTELO M A, HELLÍN A M, et al. Stereo regions-of-interest selection for pedestrian protection: A survey [J]. Transportation Research Part C Emerging Technologies, 2012, 25(4):226-237.
- [6] BENENSON R, OMRAN M, HOSANG J, et al. Ten years of pedestrian detection, what have we learned? [C]. European Conference on Computer Vision, IEEE, 2014.
- [7] LAI C Q, TEOH S S. A review on pedestrian detection techniques based on Histogram of Oriented gradient feature [C]. Research and Development, IEEE, 2014:1-6.
- [8] MAJI S, BERG A C, MALIK J. Efficient classification for additive kernel SVMs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(1):66-77.
- [9] KELLER C G, ENZWEILER M, ROHRBACH M, et al. The Benefits of dense stereo for pedestrian detection[J]. IEEE Transactions on Intelligent Transportation Systems, 2011, 12(4):1096-1106.
- [10] 陈双玉,王海晖,吴云韬,等.基于双目立体视觉和 SVM 算法行人检测方法[J].华中科技大学学报(自然科学版),2015,43(S1):140-143.
- [11] 杨峰,陈建春.基于达芬奇技术的音视频压缩传输系统[J].电子科技,2017,30(8):96-98,101.
- [12] 赵太飞,孙孝彬,娄俊鹏.基于 H.264 的多传输模式 IP 网络视频监控系统[J].计算机系统应用,2015,24(11):81-87.
- [13] ZHAO L, THORPE C E. Stereo-and neural network-based pedestrian detection[J]. IEEE Transactions on Intelligent Transportation Systems, 2000, 1(3):148-154.
- [14] ESS A, LEIBE B, SCHINDLER K, et al. A Mobile Vision System for Robust Multi-Person Tracking[C]. 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2008), USA, IEEE, 2008.
- [15] 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.
- [16] ZHANG C, VIOLA P A. Multiple-instance pruning for learning efficient cascade detectors[C]. Advances in Neural Information Processing Systems 20, Proceedings of the Twenty-First Annual Conference on Neural Information Processing Systems, Vancouver, British Columbia, Canada, Curran Associates Inc. 2007.

作者简介

王颖,博士、工程师,主要研究方向网络信息安全、智能视频处理等。

E-mail: 925202126@qq.com