

基于 FPGA 的智能视频加速检索系统

赵博^{1,2} 杨明² 汤志伟² 蔡玉鑫²

1 上海大学通信与信息工程学院 上海 200444

2 公安部第三研究所 上海 201204

(95130880@qq.com)

摘要 根据视频监控对处理、检索视频的准确性、速度以及功耗等新的需求,文中研究设计一套基于 FPGA 的智能视频检索系统,可同时实现视频预处理、人脸检测、智能背景去除、视频结构化描述等功能,并实现硬件加速的智能视频加速检索,以达到快速分析、处理监控视频,并迅速得到理想的处理结果的目的。该系统通过 FPGA 集成 CPU(ARM9)、DDR、视频采集模块、各种外设(UART,GPIO 等)可实现对实时视频的检测分析。对于 1280 * 720 分辨率的视频,所有时钟都采用的 280MHz 同步时钟,系统帧率大约为 6.6fps,和 FPGA 上结论一致,将该系统 IP 集成到 SOC 芯片中时,处理速度可达到 30 fps。

关键词: FPGA;背景去除;人脸检测;结构化描述;视频检索

中图法分类号 TP391

Intelligent Video Surveillance Systems Based on FPGA

ZHAO Bo^{1,2}, YANG Ming², TANG Zhi-wei² and CAI Yu-xin²

1 School of Communication and Information Engineering, Shanghai University, Shanghai 200444, China

2 The Third Research Institute of the Ministry of Public Security, Shanghai 201204, China

Abstract According to the video monitoring new requirements, this paper designed a set of intelligent video surveillance and retrieval system based on FPGA. The system can perform video preprocessing, face detection, intelligent background removal and video structural description simultaneously, and realize intelligent video acceleration retrieval relies on hardware acceleration, so that to achieve rapid analysis and processing of surveillance video, and quickly get the ideal processing results. By FPGA integration (ARM9) CPU, DDR, video acquisition module, a variety of peripherals (UART, GPIO, etc.), the system can realize the detection and analysis of real-time video. For videos with a resolution of 1280 * 720, 280MHz synchronous clock is adopted, and the system frame rate is about 6.6fps. When these Intellectual Property cores are integrated into the SOC chip, the processing speed can reach 30 fps.

Keywords FPGA, Background removal, Face detection, Structured description, Video retrieval

1 引言

本文研究设计一套可实现视频预处理、智能分析、视频结构化描述等功能,并支持基于内容的高速检索,支持 96 dB 宽动态视频数据处理的智能视频处理系统^[1]。

本文采用 FPGA 搭建设计硬件系统,实现 3 个硬件加速功能模块,其各自具有独立的配置寻址空间。当智能视频加速引擎 IP 集成到 FPGA 时,通过软件,可以根据不同需求配置智能加速引擎,调用这 3 个硬件加速模块,完成需要的功能。

2 智能视频加速检索系统

本文设计的 3 个独立硬件加速功能模块,可以互不依赖地同时工作,如图 1 所示。该系统根据不同需求配置智能加速引擎,调用这 3 个硬件加速模块,完成视频的处理检索功能。

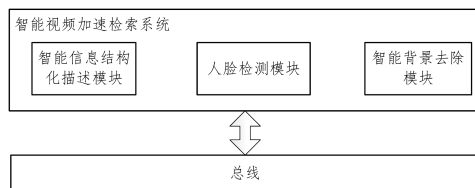


图 1 智能视频加速检索系统示意图

Fig. 1 Schematic diagram of intelligent video acceleration retrieval system

2.1 智能背景去除

常见的背景去除算法有很多,多数算法还停留在实验室的理论研究状态,应用于工程实践的主要有:背景相减法、均值滤波法、帧间差分法、单高斯背景算法和混合高斯背景法。本文主要从系统的抗干扰性和执行效率等角度考虑,选择混合高斯建模的方法进行背景去除。混合高斯算法是一种较为复杂的背景建模方法,采用多个高斯分布来描述背景,它克服

了单高斯模型对背景稳定性的要求,但是由于算法复杂,实现过程也较为复杂^[2-7]。

包括以建立高斯分布、初始化高斯模型和更新高斯模型,动态适应当前环境变化,参数的设置等等。设每个像素用 K 个高斯分布表示,其中第 k 个高斯分布为 $\eta_k(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$:

$$f_k(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^2 |\boldsymbol{\Sigma}_k|}} \exp\left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1} (\mathbf{x} - \boldsymbol{\mu}_k)\right] \quad (1)$$

其中, $k=0, 1, \dots, K$, $\boldsymbol{\mu}_k$ 是高斯分布的均值, $\boldsymbol{\Sigma}_k$ 是高斯分布的方差,并且 K 个高斯分布的权重分别为 $\omega_1, \omega_2, \dots, \omega_K$ 。初始化阶段:以当前读入像素值 \mathbf{x}_t 为均值,初始方差 $\boldsymbol{\Sigma}_0$ 建立高斯分布,分配初始权重 ω_0 。

当读入新的当前像素值 \mathbf{x}_t 后,计算它与当前高斯分布的马氏距离:

$$L_k = \sqrt{(\mathbf{x} - \boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1} (\mathbf{x} - \boldsymbol{\mu}_k)} \quad (2)$$

如果计算出的最小马氏距离大于阈值 θ_L ,则当前像素被判定为前景。如果计算出的最小马氏距离不大于阈值 θ_L ,则将马氏距离最小的高斯分布的下标记为 k_{\min} 。对当前已有的高斯分布进行排序,按权重值 ω 由大到小排列,从第一个分布的权重值开始累加,直到权重累加值大于阈值 θ_w 。如果 k_{\min} 对应的高斯分布属于被累加的分布,则当前像素为背景,否则当前像素为前景。

如果计算出的最小的马氏距离大于阈值 θ_L :如果当前有效的高斯分布数量不足 K 个,则新增一个高斯分布;如果当前有效的高斯分布数量达到 K 个,则删除权重最小的高斯分布。所建立的高斯分布,以 \mathbf{x}_t 为均值, $\boldsymbol{\Sigma}_0$ 为初始方差建立高斯分布,分配初始权重 ω_0 。

其他的高斯分布的权重按如下规则进行更新: $\omega_k = (1 - \alpha)\omega_k$; 然后对 K 个高斯分布的权重进行归一化处理。如果计算出的最小马氏距离不大于阈值 θ_L ,将马氏距离最小的高斯分布的下标记为 k_{\min} 。更新现有高斯分布的权重:

$$\omega_k = \begin{cases} (1 - \alpha)\omega_k + \alpha, & k = k_{\min} \\ (1 - \alpha)\omega_k, & k \neq k_{\min} \end{cases} \quad (3)$$

更新下标为 k_{\min} 的高斯分布:

$$\boldsymbol{\mu}_k = (1 - \rho)\boldsymbol{\mu}_k + \rho\mathbf{x}_k \quad (4)$$

$$\boldsymbol{\Sigma}_k = (1 - \rho)\boldsymbol{\Sigma}_k + \rho(\mathbf{x} - \boldsymbol{\mu})^T (\mathbf{x} - \boldsymbol{\mu})$$

图2为背景去除实现的流程图。

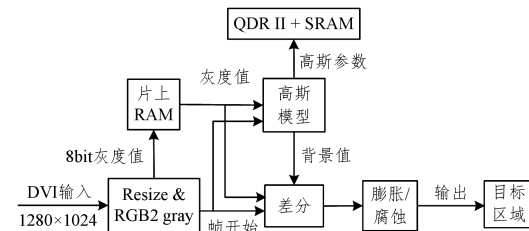


图2 背景去除流程图

Fig. 2 Background removal flow chart

本文分别从C语言层级、CPU层级和算法结构进行优化,以减小运算量为目标,开展了汇编优化、编译器优化、内联函数优化、库优化、流水线优化等工作,在保证算法效果的同时对算法的执行速度进行了大规模优化,并根据定点芯片的要求进行了定浮点转换,使系统的效率提高了30倍,算法在FPGA上的执行速率可达到30fps,CPU占用率约为25%。

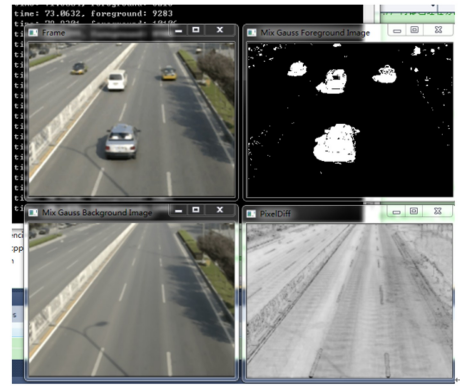


图3 智能背景去除结果

Fig. 3 Intelligent background removal results

2.2 实时人脸检测

本文需要实现的人脸检测功能模块,必须在性能优异的软件算法的基础上提出一种适合硬件并行实现的算法架构,以兼顾系统资源消耗、能源消耗限制和实时性要求。本文采用的 Haar-like 特征结合 Adaboost+Cascade 算法^[8-11]实现人脸检测功能。

因此,在上述基础上针对智能视频监控的一些底层需求,结合视频结构化描述的特征需要和FPGA开发的具体硬件环境进行算法改进,采用多层分类器协处理器并行结构来实现人脸检测功能模块。图4为该算法级联结构图。

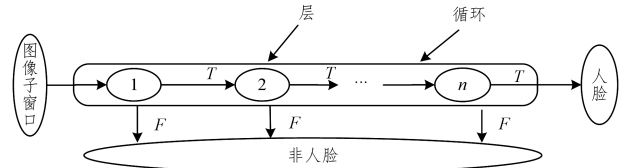


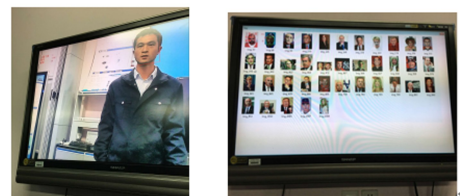
图4 人脸检测算法级联结构图

Fig. 4 Cascade structure of face detection algorithm

分类器中的特征值可以通过一种被称为积分图像的中间图像快速计算出来。在点 (x, y) 处的积分图像值定义为该点左上角所有像素灰度值之和,计算公式为:

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

根据 Adaboost 算法,一幅图像中一个子区域被判定为 人脸区域,当且仅当该子区域能通过算法中级联的所有分类器,而一个子区域能通过一个分类器,当且仅当该子区域在该分类器所有特征上的特征值之和大于该分类器的阈值。最终通过所有分类器的区域,就被认为是 人脸区域。



(a) 单人脸

(b) 多人脸

图5 实时人脸检测结果

Fig. 5 Real-time face detection results

2.3 智能信息化描述

视频结构化描述作为新一代智能监控系统的基础,可支持面向海量异构信息融合建模的需求,即能够提供符合统一

标准的、对监控内容具有语义描述功能的结构化元数据。图像匹配和更高级的语义描述,必须先对这部分区域的底层特征,如颜色、纹理等进行提取^[12-15]。图 6 为信息化描述流程图。

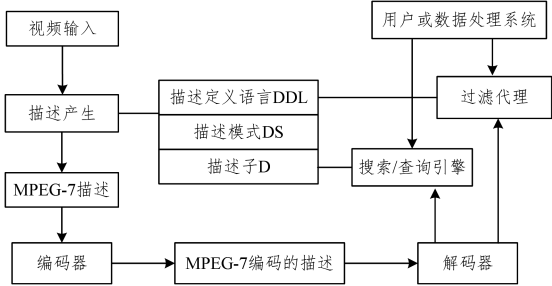


图 6 信息化描述流程

Fig. 6 Flow chart of information description

本模块将这些底层特征分为主要颜色特征、颜色分布特征、颜色结构特征和边缘直方图特征 4 种。在“智能背景去除”“人脸检测”的处理结果的基础上进行信息结构化描述及封装,以支持后端海量数据的检索、异构数据融合、语义建模及基于语义模型的推理等,更加侧重后端海量异构信息的融合建模和对多模式快速检索、基于语义模型的智能推理和辅助决策支持^[16-18]。实时结果如图 7 所示。

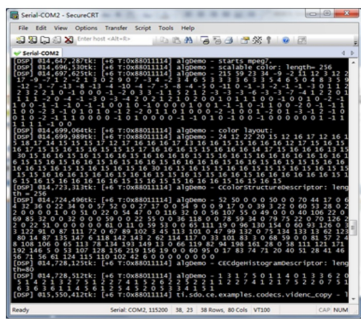


图 7 信息结构化描述结果

Fig. 7 Results of structured information description

3 智能视频加速检索硬件系统

智能视频加速检索硬件系统,基于 FPGA 主板集成了 CPU(ARM9)、DDR、视频采集模块、各种外设(UART,GPIO 等)等外围设置。通过该系统的编程,配合外部显示处理的视频通路子板,搭建一个比较直观的、可以直接看到智能视频加速引擎效果。图 8 为系统示意图,图 9 为系统硬件平台实物图。

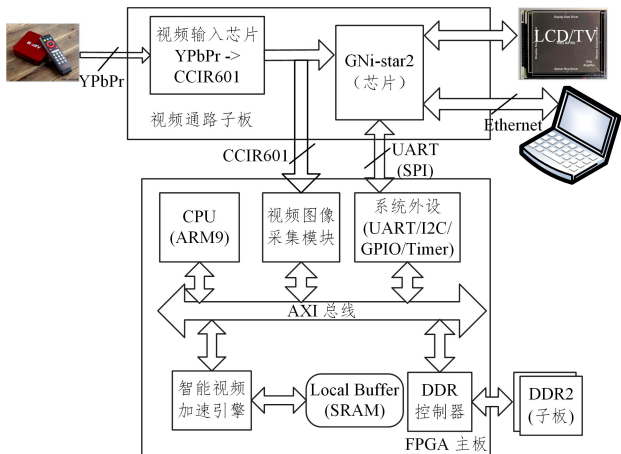


图 8 智能视频检索系统示意图

Fig. 8 Schematic diagram of intelligent video retrieval system

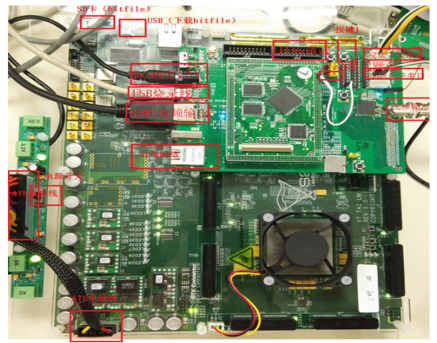


图 9 基于 FPGA 的智能视频检索加速系统

Fig. 9 Intelligent video retrieval system based on FPGA

4 测试结果

本文在实测的配置情况下,对于 1280 * 720 分辨率的视频,所有时钟都采用的 280MHz 同步时钟,系统帧率大约为 6.6 fps,和 FPGA 上结论一致,智能视频加速引擎 IP 集成到 SOC 芯片中时,实时处理 720P 视频,处理速度可达到 30 fps。进一步的检测中可以通过统计参数结合自学习来调整相关配置以满足检测效果和检查速率的要求。由于各个 Casdet 是完全并发的,相互没有依赖关系。根据实时视频的处理结果,可在数据库中快速检索匹配信息,结果如图 10 所示。

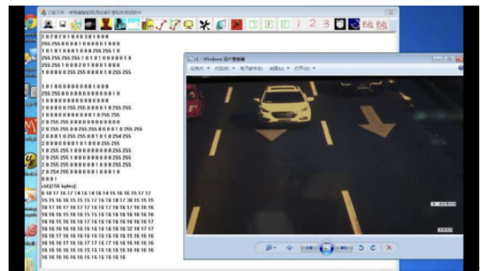


图 10 实时视频检索结果

Fig. 10 Retrieval results of real-time image

结束语 本文研究基于 FPGA 平台的智能视频加速检索系统,实现智能监控功能的硬件平台及软件的设计与优化,利用该平台代替服务器等进行算法处理,实现对视频结构化描述等技术的硬件加速功能,支持基于内容的高速检索。该系统优化了图像处理算法在内存空间有限,但具有功耗低且便携的平台上的资源开销。下一步工作可进一步开发、实现支持国家标准 SVAC 的视音频编解码、图像预处理、智能分析和码流加密功能的视频监控 SoC 芯片,开发基于 SoC 芯片的高清网络摄像机等视频处理类产品。

参考文献

- [1] 骆云志,刘治红. 视频监控技术发展综述[J]. 兵工自动化, 2009 (1):1-3,11.
- [2] 徐晓夏,陈泉林. 智能交通监控系统中的自适应背景更新算法研究[J]. 上海大学学报(自然科学版), 2003(5):35-36.
- [3] 王少斌,苏淑靖,袁财源. 基于 FPGA 的高清视频采集系统设计[J]. 电子技术应用, 2019, 45(7):63-66, 71.
- [4] 王鹏,王慧,孔凡宁,等. 基于 FPGA 的视频监控摔倒检测报警系统[J]. 电机与控制学报, 2019(8):122-128.
- [5] 杨家成,刘杰. 智能视频监控技术存在的问题及应对措施浅议[J]. 中国新通信, 2019, 21(7):60-61.

- [J]. 浙江大学学报(工学版),2019(3):1-7.
- [8] 刘鹏,常振杰.基于 PCIe 的无线电数据传输接口实现[J]. 西安邮电大学学报,2018,23(6):68-73.
- [9] 金锐.一种基于 PCIe 接口的高速安全芯片设计[C]//数字中国能源互联——2018 电力行业信息化年会论文集.中国电机工程学会电力信息化专业委员会:人民邮电出版社电信科学编辑部,2018:4.
- [10] 许少榕.基于 DPM 的图像视觉目标识别改进算法[J]. 济宁学院学报,2018,39(2):31-38.
- [11] 孙营,王波涛.基于可变形部件改进模型的夜间车辆检测方法[J]. 计算机工程,2019,45(3):202-206.
- [12] 杨亚涛,张松涛,李子臣,等.基于 Zynq 平台 PCIE 高速数据接口的设计与实现[J]. 电子科技大学学报,2017,46(3):522-528.
- [13] 侯宁.嵌入式多核处理器设计与实现关键技术研究[D]. 合肥:合肥工业大学,2012.
- [14] 杨川.MPCore 多核处理器并行计算方法的研究与实现[D]. 成都:西南交通大学,2014.
- [15] 政越伟.嵌入式多通道监测系统通信技术研究与实现[D]. 北京:北方工业大学,2013.
- [16] 王庆.面向嵌入式多核系统的并行程序优化技术研究[D]. 哈尔滨:哈尔滨工业大学,2013.
- [17] 张沁馨.基于 PXIe 高速接口的高速数据传输平台的设计与实现[D]. 成都:电子科技大学,2018.
- [18] 黄丹萍.基于 PCIe 的高速数据采集模块的软件设计与实现[D]. 成都:电子科技大学,2018.



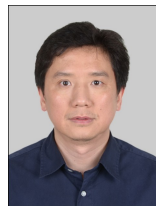
CAI Yu-xin, born in 1989, master, senior engineer. Her main research interests include mobile police and image processing.

(上接第 611 页)

- [6] WINDYG A, PIOTR S. Fast impulsive noise removal[J]. IEEE Transactions on Image Processing, 2001, 10(1):173-179.
- [7] YOU Y, BADAL S P, SHELLEY J T. Automatic Analyte-Ion Recognition and Background Removal for Ambient Mass-Spectrometric Data Based on Cross-Correlation[J]. Journal of the American Society for Mass Spectrometry, 2019, 30:1720-1732.
- [8] HONG S, FENG S. A Real-Time Algorithm for Moving Objects Detection in Video Images[C]// Fifth World Congress on Intelligent Control and Automation, 2004.
- [9] 闫河,王鹏,杨晓龙,等.一种新的 Ada Boost 人脸检测方法[J]. 重庆理工大学学报(自然科学),2019,33(2):106-110,161.
- [10] LONG X Z, LU H T, PENG Y, et al. Graph regularized discriminative non-negative matrix factorization for face recognition [J]. Multimedia Tools and Applications, 2014, 72(3):2679-2699.
- [11] 李天煌.基于级联结构的人脸检测若干方法研究[D]. 南京:东南大学,2018.
- [12] SUNG M K, SOO J P, CHEE S W. Image Retrieval via Query-by-Layout Using MPEG-7 Visual Descriptors[J]. ETRI Journal, 2007.
- [13] 徐检.浅议 MPEG-7 标准及应用[J]. 有线电视技术, 2005(3):39-43.
- [14] SARNO R, RIDOAN J A, KOM S, et al. Classification of Music Mood Using MPEG-7 Audio Features and SVM with Confidence Interval[J]. International Journal of Artificial Intelligence Tools, 2018, 27.
- [15] 卢成浪,吴宗大,李桂玲.基于 MPEG-7 视频检索系统的设计与实现[J]. 西北大学学报(自然科学版),2018,48(3):369-375.
- [16] DING S T, QU S R, XI Y L, et al. A long video caption generation algorithm for big video data retrieval[J]. Future Generation Computer Systems, 2019, 93:583-595.
- [17] 王治,韩祥.视频结构化解析技术在公安警务实战中的建设与应用[J]. 警察技术,2018(5):63-66.
- [18] SURYONO S, PUTRO S P, WIDOWATI, et al. System on chip (SOC) wi-fi microcontroller for multistation measurement of water surface level using ultrasonic sensor[J]. Journal of Physics: Conference Series, 2018, 27(5).



ZHAO Bo, Ph.D. Her main research interests include signal and information processing and artificial intelligence, IoT technology.



YANG Ming, associate professor. His main research interests include IoT security and artificial intelligence, IoT technology.