

一种基于 ZYNQ 的视频采集处理系统

魏洪健¹, 徐琦琳¹, 张 瑛¹, 孙科学^{1,2*}

(1. 南京邮电大学 电子与光学工程学院, 江苏 南京 210023;

2. 射频集成与微组装技术国家地方联合工程实验室, 江苏 南京 210023)

摘 要:针对当前基于 ARM 与 DSP 的图像处理框架存在的带宽低、速度慢的问题,提出了一种 ARM+FPGA 异构的 ZYNQ-7020 芯片作为主控的视频实时处理框架。该系统由 500 万像素级别 CMOS 摄像头 OV5640 采集 1 280×720 分辨率下 60 帧图像数据,处理后通过 HDMI 高清接口进行输出。系统搭建了三种图像滤波处理,图像边缘检测及增强,形态学腐蚀与膨胀等图像预处理算法模块,均采用流水结构最大化处理速度以保证视频信息的实时性,封装为通用视频接口 IP 核,有较好的可移植性与扩展性。同时系统集成了 RTL8211E 路由芯片,可通过 UDP/IP 协议对内部算法模块进行控制,切换实时输出的算法通道。搭建 Vivado 与 Modelsim 联合仿真环境,对系统的实时性及算法有效性进行验证。结果表明,系统可完成对图像的实时采集、处理、输出功能,可很好地抑制摄像头前端采集时产生的噪点,相比传统 ARM 端图像处理速度有较大的提升。

关键词:ZYNQ; ARM 处理器; 图像预处理; 实时视频处理; 现场可编程门阵列

中图分类号: TP274+.2

文献标识码: A

文章编号: 1673-629X(2022)11-0024-06

doi: 10.3969/j.issn.1673-629X.2022.11.004

A Video Acquisition and Processing System Based on ZYNQ

WEI Hong-jian¹, XU Qi-lin¹, ZHANG Ying¹, SUN Ke-xue^{1,2*}

(1. School of Electronic and Optical Engineering, Nanjing University of Posts and Telecommunications,
Nanjing 210023, China;

2. Nation-Local Joint Project Engineering Lab of RF Integration & Micropackage, Nanjing 210023, China)

Abstract: The current image processing framework based on ARM and DSP has the disadvantages of low bandwidth and slow speed. To address this problem, a video real-time processing framework based on ARM+FPGA heterogeneous Zynq-7020 chip is proposed. The 5 million pixel level CMOS camera OV5640 collects 60 frames of image data at a resolution of 1 280×720, then outputs the video stream after being processed through the HDMI high-definition interface. The system integrates lots of image preprocessing algorithm modules such as three kinds of image filtering processing, image edge detection, sharpen, morphological corrosion and dilation. All the above adopt a pipeline structure to maximize the processing speed to ensure the real-time performance of video information, packaged as a general video interface IP core to gain better performance on flexibility and scalability. The system also integrates the RTL8211E routing chip, which can control the internal algorithm module through the UDP/IP protocol and switch the output algorithm channel. To verify the real-time performance of the system and the effectiveness of the algorithm through setting up a joint simulation environment of Vivado and Modelsim, it is concluded that the system can complete the functions of real-time image acquisition, processing and output, and can well suppress the noise generated when the camera front acquisition, and the image processing speed is greatly improved compared with the traditional ARM end.

Key words: ZYNQ; ARM processor; image preprocessing; real-time video processing; field programmable gate array

0 引 言

实时视频处理技术在工业生产、缺陷检测、自动驾驶等领域有不可替代的作用,同时各个领域也对图像处理带宽、实时性及清晰度有了更高的要求。以往的

基于 ARM(Advanced RISC Machine)数字图像处理系统^[1-2]虽然能满足复杂的图像处理需求,但因为其处理器取样率低,难以得到带宽及实时性的保障。而基于 DSP(Digital Signal Processing)的处理框架^[3]能满

收稿日期: 2021-11-28

修回日期: 2022-03-30

基金项目: 江苏省研究生科研创新计划(SJCX21_0279, KYCX21_0711)

作者简介: 魏洪健(1999-),男,硕士研究生,研究方向为智能信号处理;通信作者: 孙科学(1981-),男,博士,教授,硕导,研究方向为智能信号处理与通信软件设计。

足高实时性的需求,但在开发高复杂度的集成系统时周期长,且系统并行度低。在目标识别、机器验布、流水线产品缺陷检测等领域^[4-6],图像处理系统需要采集产品实时的高清画面,并且同步处理视频特征信息,图像的有效信息通常只会存在 1~2 帧,这就要求系统要具有极高的实时性。此外对不同产品检测时算法的选择也不同,这便要求处理系统不仅需要具有高实时性的特点,也要具有一定的灵活性。而 ARM+FPGA (Field Programmable Gate Array) 的异构处理框架可以很好地兼顾两者。FPGA 在处理高带宽视频数据时能充分展现其强大的并行计算能力^[7],对已经相对成熟的图像预处理算法可以用 Verilog HDL 硬件描述语言来实现硬件的加速。因此以 FPGA 来实现数字图像处理算法,用 ARM 进行控制的系统框架可以满足对视频处理有较高实时性需求的领域。

该文设计了一种基于 Xilinx 公司推出的 ZYNQ 系列 ARM+FPGA 异构结构作为处理核心的视频数据实时采集、处理、控制、输出的框架,可支持高清 1 280×720 分辨率下 60 帧的实时处理。结合了双核 ARM Cortex-A9 处理器强大的驱动能力、高灵活性、可配置性与 FPGA 并行、高速、再编程的特点,实现了软硬均可重复编程的结构。基于 Verilog HDL 硬件描述

语言实现了图像的均值模糊,中值滤波,高斯模糊、边缘特征锐化、二值化、色彩空间转化、sobel 边缘检测、形态学腐蚀与膨胀等经典图像预处理算法,均采用流水处理的形式最大化处理速度来保障系统的实时性^[8],并且可以通过上位机来实现不同预处理算法的组合与实时切换,所有算法均封装为 IP 核来增加算法的灵活性与可移植性。主要由软件处理系统 PS (Processing System) 端与硬件可编程逻辑 PL (Programmable Logic) 端的软硬协同方式进行工作。

1 系统整体框架

系统采用 ZYNQ 系列的 XC7Z020-CLG400 芯片作为主控,OV5640 摄像头作为视频数据源,另外还包括一台支持 HDMI 高清接口的显示器。搭建了视频数据的输入、输出通路以及硬件算法模块,整体结构框图如图 1 所示。PL 硬件部分负责完成视频输入、硬件算法设计、视频输出等部分。PS 处理器部分实现对整个系统的控制,包括硬件初始化流程,本地网络初始化以及 UDP 网络数据的接收与判断,并且通过 AXI4 总线对图像处理算法模块进行控制。PL 与 PS 通过 AXI (Advanced eXtensible Interface) 协议进行数据互通,实现软硬件的协同工作。

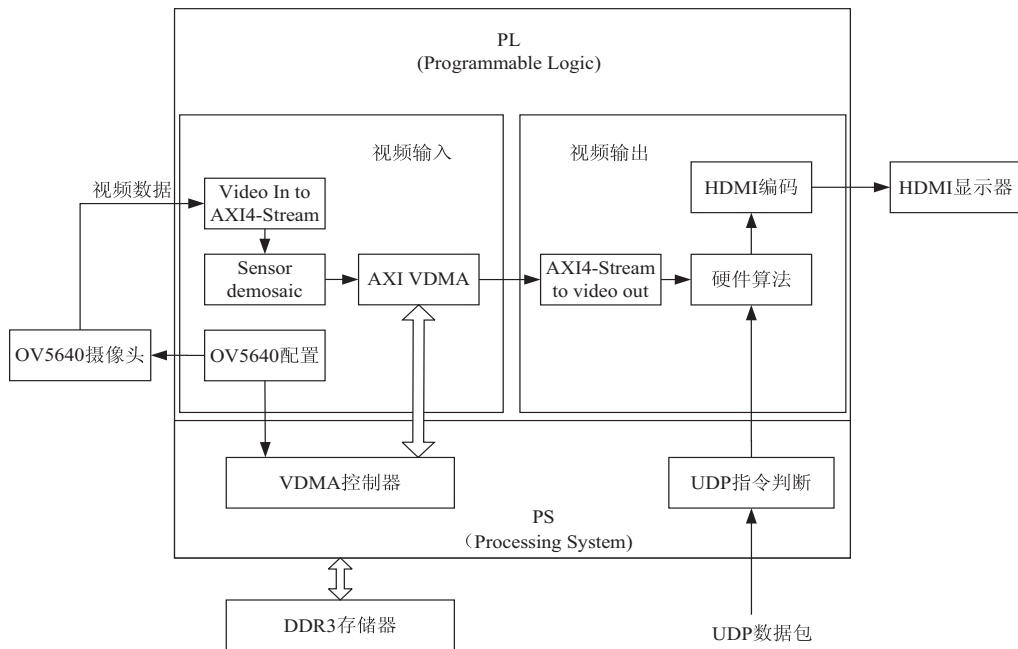


图 1 系统框图

从视频数据采集并在系统中的传输处理直至显示角度,可将系统按功能划分为如下主要模块:

(1) 视频数据采集模块:系统通过 500 万像素 CMOS 摄像头采集外界视频数据,输入至 PL 硬件部分完成数据采集。摄像头初始化由 PL 部分 IIC 时序控制器完成,在上电后 2 秒内即可完成初始化,正常输出视频流数据。

(2) 视频输入模块:摄像头 RAW 数据传入后,通过 Video in to AXI4StreamIP 将数据格式转化为 AXI-stream 格式进入 demosaic 模块,对视频色彩进行插值还原,同时将单个像素数据由 RAW 转化为 RGB888 格式,最后再经 VDMA (Video Direct Memory Access) 缓存至 DDR3 中,完成视频输入通路搭建。

(3) 视频输出模块:首先通过 AXI4Stream to

Video out 将图像数据按行从 DDR 存储器中读出,再经过硬件算法处理模块,对视频数据进行流水线式处理,实时输出处理后的视频数据,在经过 HDMI 编码器,将标准 VGA 时序协议视频编码为 HDMI 差分信号进行输出。

2 系统关键部分实现

2.1 摄像头初始化

系统的图像源由 OmniVision 公司的 OV5640 摄像头进行采集,其拥有 $2\,592 \times 1\,944$ (500 万像素) 的感光阵列,最大输出能达到 QSVGA ($2\,592 \times 1\,944$) 下的 15 帧输出,支持 SCCB (Serial Camera Control Bus) 协议控制。SCCB 协议是一种向上兼容 IIC 协议的摄像头寄存器配置协议,在主机以单字节发送数据至从机时完全兼容 IIC 协议。系统选择了 $1\,280 \times 720$ (92 万像素) 下 60 FPS 格式的视频流输出。由硬件 PL 端完成 IIC 时序控制器对摄像头内部控制寄存器以单字节写的格式进行初始化设置,表 1 列出了部分关键寄存器地址、功能及配置值。系统总共对 OV5640 内部寄存器进行 253 次 SCCB 读写配置。

表 1 OV5640 部分关键寄存器配置

寄存器地址	寄存器功能	配置值
0x3037	摄像头锁相环设定	0x13
0x3108	摄像头时钟设定	0x01
0x3808	输出图像水平宽度高 4 位	0x05
0x3809	输出图像水平宽度低 8 位	0x00
0x380A	输出图像垂直高度高 3 位	0x20
0x380B	输出图像垂直高度低 8 位	0xd0
0x501F	输出格式设定	0x03

输出图像的水平宽度为 0x0500 为 1 280, 垂直高度为 0x02d0 为 720, 即设定为 $1\,280 \times 720$ 的输出分辨率。此时可计算出输出像素的时钟为 $1\,650 \times 750 \times 60 = 74.25\text{ MHz}$ 。

2.2 摄像头数据 Demosaic

CMOS 摄像头是由覆盖在硅衬底上的滤色片测量不同波长的光强度来采集图像数据,且每种滤色片只能对应一种颜色,即每个像素原始数据只包含一种颜色信息,要通过临近像素的可用信息来还原出原像素的三种颜色分量,而人眼对绿色的敏感度最高,因此往往 CFA (Color Filter Array) 阵列由 $1/2$ 的绿色分量与 $1/4$ 的红色与蓝色分量组成, Bayer 阵列共有四种排列模式,如图 2 所示。

从 CMOS 摄像头输出的 RAW 格式的数据,通过 Video in to AXI4Stream 转化为 AXI-stream 格式的数据进入 Sensor Demosaic 电路,利用 5×5 矩阵通过色度

方差来抑制在还原像素时产生的颜色混叠效应,提升传感器输出图像色彩质量。同时由于 CMOS 摄像头 OV5640 数据输出引脚只有 10 位,实际上摄像头的输出为每 2 个 pix clock 输出一个 RGB565 格式的像素数据, Sensor Demosaic 同时也对按字节接收到的 RAW 数据进行格式重构,转换 RAW 格式为 24 bit 的 RGB888 格式。Sensor Demosaic 偏移地址 0x28 寄存器 [1:0] 的值对应了全部四种 Bayer 阵列格式。可以在 PS 端通过 AXI 总线来设定像素的阵列模式,色彩还原后仍然以 AXI4-Stream 格式输出到 DDR 中进行视频缓存。

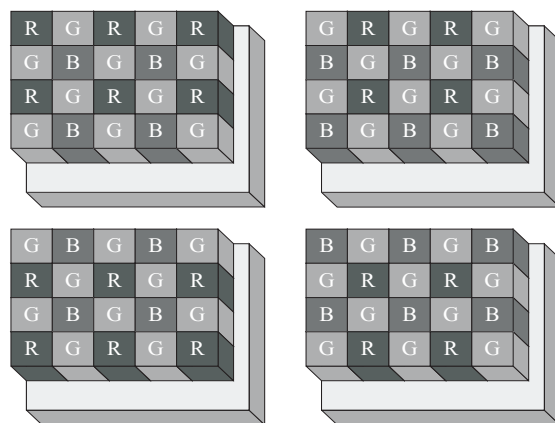


图 2 四种 CFA 阵列左上角 Bayer 组合

2.3 图像处理 IP 封装

图像卷积运算作为图像预处理中应用最广泛的方式,由 FPGA 实现流水线运算时,能达到最高的实时性^[9],这是因为 FPGA 在处理图像进行运算时以行为单位,而 CPU、GPU、DSP 在处理时是以帧为单位。如图 3 (左) 所示,图像的传输是从左上角的第一个像素开始,向右输出 1 整行 1 280 个像素点后,向下移动一行开始传输第二行的像素数据,直至 720 行扫描结束后完成整帧的输出。通过 ShiftRegister 缓存图像的前两行数据,同时接受当前数据作为第三行,就实现了三行数据的同时缓存。每个时钟的上升沿到来时, ShiftRegister 中每个数据会向后移动一位,移位寄存器将最后一个数据作为当前的输出。图 3 (右) 结构中 ROW1 ~ 3 即为当前的三行像素数据输出,通过对三行数据的一个缓存打拍处理,就可得到图 3 (左) 中原图像左上角像素的 3×3 矩阵窗口,此时每经过一个时钟周期,就相当于该 3×3 矩阵窗口在像素阵列中向右移动一格,即实现步长为 1 的图像卷积运算。

得到图像的 3×3 像素矩阵后,通过与不同的卷积模板进行运算后再替换原像素点数据,如此遍历整幅图像之后即可实现图像的滤波操作。常见的处理有用 9 个数据平均值代替原像素点值的均值滤波;用 9 个值中的中值来代替原像素点的中值滤波。

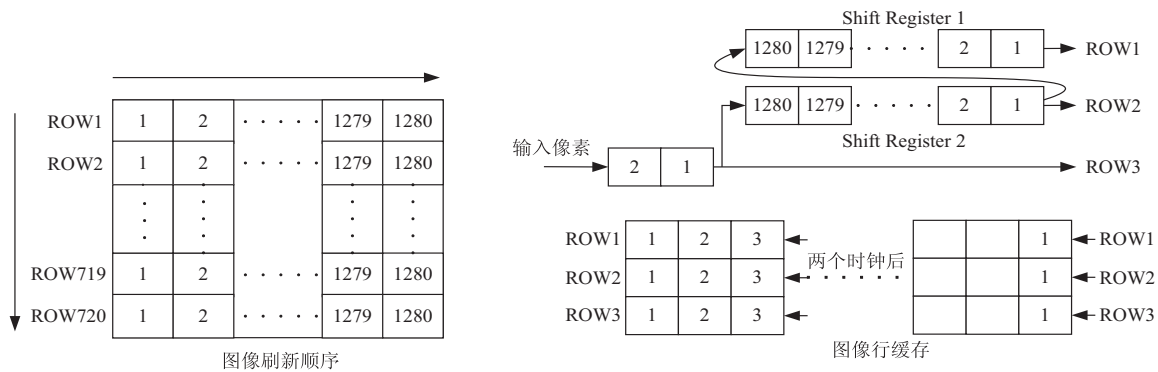


图 3 像素阵列及扫描方向

以 sobel 边缘检测^[10]为例,其原理是根据像素灰度值与其周围的离散差值来近似计算出该点不同方向上的梯度。根据式(1)求出原像素矩阵 I 与两个 3×3 核的卷积结果水平积分因子 G_x 与竖直积分因子 G_y , 再通过式(2)求出近似梯度 G , 判断 G 与边界阈值 S 大小即可判断出该点是否为边缘点。

$$G_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * I \quad (1)$$

$$G_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * I$$

$$G = \sqrt{G_x^2 + G_y^2} \quad (2)$$

$$\text{pix_data} = \begin{cases} 0, G < S \\ 255, G \geq S \end{cases} \quad (3)$$

为验证硬件算法的实时性,通过 Vivado 与 Modelsim 进行联合仿真,编写测试文件模拟 VGA 时序驱动,每帧读取不同的图片,经过硬件处理后再保存为新的图片,共进行三帧的仿真即处理三张不同的图片,第四帧开始时结束仿真。结果表明,所述设计能满足视频处理实时性的要求。

2.4 硬件算法模块实现

现代数字图像处理的模型一般分为多级的形式,由视频采集前端、图像预处理以及图像信息提取三个部分构成。其中视频采集前端往往是由采集端对原生图像数据进行最基本的处理,不会对原始图像数据的信息特征产生较大影响。本系统中 OV5640 摄像头内部已集成了 ISP 图像信号处理单元,其中包括 AEC (自动曝光控制)、AWB (自动白平衡)、gamma 矫正、BLC (黑电平矫正) 等图像补偿功能。

该系统针对图像预处理以及图像信息提取给出了不同的方案,读者可根据实际情况选择或者增删图像处理模块。系统硬件算法内部结构如图 4 所示,其中图像处理算法包括三种预处理滤波、边缘检测以及基于拉普拉斯算子的图像边缘增强、图像形态学腐蚀与膨胀。均采用 PL 端硬件流水的形式完成图像处理,封装为通用视频接口的 IP 核,保证了系统的可扩展性与灵活性。视频选择器为封装了 AXI 接口的数据接收自订 IP 核,接收 PS 端的控制,可从输入的视频数据中选择一路输出。PS 端通过对视频选择器的控制即可实时调整系统硬件算法的组合。

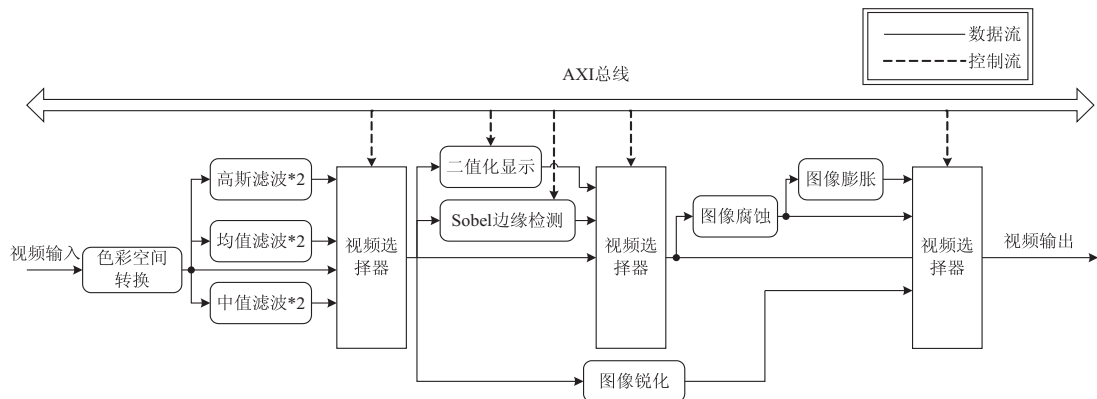


图 4 硬件算法内部结构

2.5 PS 端实现

PS 软件部分^[11]主要实现的内容包括开启 DDR 视频缓存通道^[12],初始化硬件算法模块,设置开发板

IP 地址以及对 UDP 网络数据的处理。SDK 中处理的流程图如图 5 所示。系统上电后,首先 PS 端将通过 AXI 总线对硬件算法模块中的 3 个视频选择器进行初

始化配置,设置其默认输出原始视频数据。随后通过 AXI 总线向 PL 端的 IIC 时序控制器发送启动信号,开始对 OV5640 摄像头内部寄存器进行配置,共发送 253 个 SCCB 协议数据。PL 端寄存器配置数据完成后,硬件端通过 ZYNQ 中 IRQ_F2P 接口发送硬件的中断信号。IRQ_F2P 为 ZYNQ 提供的硬件中断接口,硬件中断号 ID 有 [91:84],[68:61] 共 16 个,可以设置边沿触发或电平触发,本系统中使用上升边沿触发模式进入中断处理函数。通过硬件中断功能软件端可以在硬件信号触发后最快地做出响应,即最优先处理中断处理函数中的程序部分。中断处理函数中包括 PS 端对 PL 端 VDMA 进行初始化配置。视频数据的 DDR 读

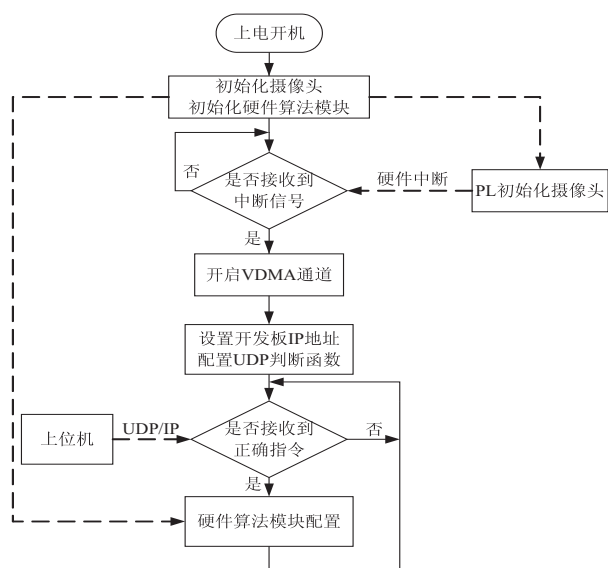


图 5 SDK 处理流程

写地址为 0x01100000,缓存三张图像。则从 0x01100000 开始的 $1\,280 \times 720 \times 3$ 个地址内会按 BGR 顺序存放第一帧的图像数据,从地址 0x013A3000 (0x01100000+1 280×720×3) 开始会存放第二帧数据,以此类推。

物联网技术已被广泛应用在 SOC 领域,本系统基于 RTL8211E 千兆路由芯片^[13-14],PS 端部署 lwip (light weight IP) 轻量化 IP 协议实现网络通信。通过 UDP 协议完成与上位机的数据交互,实现对系统的远程控制。PS 端不断读取上位机发送的 UDP 网络数据包,根据网络指令通过 AXI 总线协议对图 4 中的 3 个视频选择器的输出进行切换,也可以改变二值化阈值与 sobel 边缘检测阈值。

3 测试结果与分析

在 vivado 平台与 modelsim 仿真器完成系统的设计与仿真后,进行工程的综合、实现、生成比特流文件,导出 hardware 进入 SDK 环境进行 ps 端程序开发。完成后创建 FSBL (First Stage Boot Loader) 工程对 SDK 中 C 程序进行固化。

开启系统等待初始化流程结束后显示器通过 HDMI 线接收来自系统的 $1\,280 \times 720$ 分辨率下 60 帧视频数据,图 6 为在正光源照射下的书本文字显示结果,其中左上至右下依次为原图(a)、高斯滤波(b)、图像锐化(c)、sobel 边缘检测(d)、中值滤波+sobel 边缘检测(e)、中值滤波+sobel 边缘检测+图像腐蚀(f)的处理效果图。

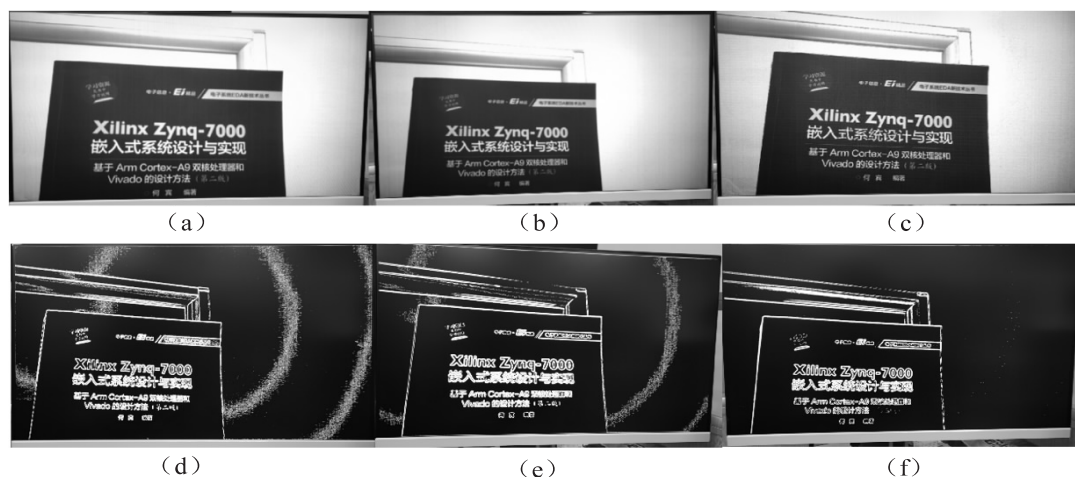


图 6 书本文字显示处理结果

图 7 为拍摄带有浮雕表面二维码图像的处理结果,其中左上至右下依次为原图(a)、图像锐化(b)、sobel 边缘检测(c)、高斯滤波+sobel 边缘检测(d)、高斯滤波+sobel 边缘检测+图像腐蚀(e)、二值化显示(f)的处理结果。

上述两种不同物体的实时图像采集结果很好地证

明了所述系统功能完备性以及算法有效性,通过不同的算法组合最终均可获得较好的图像特征信息,并且能直观地观察到每个中间级算法处理的结果。

以硬件流水线实现图像处理算法时,其延时是可以计算出的^[15]。该文所述设计中,处理算法使用的 3×3 缓存矩阵会使图像同步时钟延迟一行来进行数据同

步,1 280 p 分辨率下则会延迟 1 560 个延时,后续则在
进行不同流水线算法时处理所用的延时不同。表 2 中
给出了该系统中通过硬件算法与在 ZYNQ 芯片中

ARM 端调用 OpenCV 库^[16] 处理一帧图像的延时
对比。

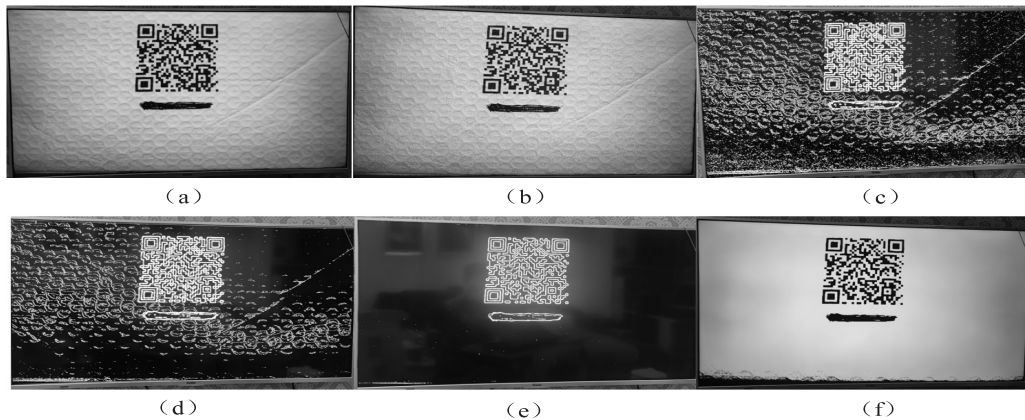


图 7 二维码图像显示处理结果

表 2 不同算法处理延迟时间

算法名称	硬件延迟时间/ μs	Opencv 处理/ μs
中值滤波	13 122.73	28 835.77
高斯滤波	13 122.73	64 670.32
均值滤波	13 122.90	64 187.05
图像锐化	13 122.73	71 566.10
腐蚀	13 122.71	46 859.26
膨胀	13 122.71	51 250.69
sobel 边缘检测	13 122.89	200 263.02

对比软硬件在图像处理速度上的差异,该系统在
进行各种图像的处理时均有更快的速度,且实现的算
法复杂度越高,则硬件流水处理的提升越大。

4 结束语

该文提出并实现了一种基于 ZYNQ-7020 芯片的
视频实时采集、处理、显示平台,充分结合了 ARM 的
灵活性与 FPGA 在图像处理领域的高实时性。平台集
成了多种经典的图像预处理算法,延时低、易于替换与
扩展,为复杂图像处理算法实现提供良好的基础平台。
经测试所述系统实时性高、性能稳定,可应用于动态捕
捉、缺陷检测、视频前端处理等领域。

参考文献:

- [1] 王 柯,汪贵华,徐 元,等. 基于 ARM 和 CMOS 图像处
理的地下管线检测技术研究[J]. 电子设计工程,2018,26
(9):158-162.
- [2] 杨成禹,庄晓奇. 基于 ARM 的图像采集与传输系统设
计[J]. 电子测量技术,2015,38(12):59-62.
- [3] 白 杰,孟令军,张慧慧. 基于 DSP 的视频处理及传输系
统设计[J]. 实验室研究与探索,2017,36(3):111-115.
- [4] JI Qingbo,DAI Chong,HOU Changbo,et al. Real-time em-
bedded object detection and tracking system in Zynq SoC

- [J]. EURASIP Journal on Image and Video Processing,
2021,2021(1):1713-1718.
- [5] 李 斌,许华杰,吴朝晖. 基于 FPGA 的多人脸实时检测系
统设计与实现[J]. 微电子学与计算机,2021,38(4):57-
62.
- [6] 李向阳. 基于 ZYNQ 的车载目标检测系统设计与实现
[J]. 机械设计,2020,37(S1):35-38.
- [7] ZHANG J,QU G. Recent attacks and defenses on FPGA-
based systems [J]. ACM Transactions on Reconfigurable
Technology and Systems,2019,12(3):1-24.
- [8] 张振利,韩凌峰. 基于 FPGA 的 HDMI 视频显示系统设计
[J]. 现代电子技术,2021,44(16):35-39.
- [9] 赵广州,张天序,王新赛,等. 基于 DSP 和 FPGA 的模块
化实时图像处理系统设计[J]. 华中科技大学学报:自然科学
版,2004,32(10):4-6.
- [10] 沈德海,侯 建,鄂 旭. 基于改进的 Sobel 算子边缘检测
算法[J]. 计算机技术与发展,2013,23(11):22-25.
- [11] 李 朗,张索非,杨 浩. 基于 Zynq-7000 的视频处理系统
框架设计[J]. 计算机技术与发展,2017,27(5):192-195.
- [12] 王春江,李 鹏. 基于 ZYNQ 的运动目标检测系统设计
[J]. 电子科技,2020,33(5):82-86.
- [13] BILAVARN S,GOGNIAT G,PHILIPPE J,et al. Design
space pruning through early estimations of area/delay
tradeoffs for FPGA implementations [J]. IEEE Transactions
on Computer-Aided Design of Integrated Circuits and Sys-
tems,2006,25(10):1950-1968.
- [14] 杨振雷,刘承敏,青先国,等. 基于 Zynq-7000 的千兆以太
网传输系统设计与实现[J]. 核技术,2021,44(2):35-42.
- [15] 何 宾,张艳辉. Xilinx FPGA 数字信号处理系统设计指
南:从 HDL、simulink 到 HLS 的实现[M]. 北京:电子工业
出版社,2019.
- [16] 刘培军,马明栋,王得玉. 基于 OpenCV 图像处理系统的开
发与实现[J]. 计算机技术与发展,2019,29(3):127-131.