

## 基于 FPGA 的大视场图像实时拼接技术的研究与实现

杨磊,任龙,刘庆,王华,周祚峰,曹剑中

(中国科学院西安光学精密机械研究所,陕西西安 710119)

**摘要:** 鉴于目前普遍采用软件方法获得大视场视频图像操作不便且实时性受限的缺陷,研究并设计了一种基于 FPGA 的可编程技术来实现多个摄像头视频数据实时拼接的大视场成像系统。系统通过 APTINA 公司的彩色 CMOS 图像传感器 MT9M034 获取原始视频图像信息,以 Xilinx 公司的 Virtex-5 系列 FPGA 为核心完成视频数据的实时采集、缓存、拼接及传输。图像拼接部分首先对原始图像进行亮度差异自动调节的预处理以提高整体的拼接效果,然后利用相位相关法完成相对平移量信息检测,对原始图像进行配准,最后采用线性加权融合算法对相邻两幅图像的重合区域进行融合处理,使拼接之后的大视场图像达到渐进渐出平滑过渡的效果。实验结果表明,该成像系统简单可靠,有效地增大了可观视场,经过拼接处理之后的大视场视频图像清晰度高、实时性强,具有一定的代表性和实用性。

**关键词:** FPGA; 大视场成像; 实时拼接; 相位相关; 线性加权融合

**中图分类号:** TP274; TP391 **文献标志码:** A **文章编号:** 1007-2276(2015)06-1929-07

## Research and implementation of large field image real-time mosaic technology based on FPGA

Yang Lei, Ren Long, Liu Qing, Wang Hua, Zhou Zuofeng, Cao Jianzhong

(Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, China)

**Abstract:** As a general rule, the software method is used to obtain large field images, which is not timely and convenient. In order to tackle the disadvantage of this method, based on FPGA, a kind of programmable technologies, a large field of view imaging system had put forward and achieved, which can fulfill the real time stitching of the data from multiple cameras. Through the APTINA's color CMOS image sensor MT9M034, the original image information had gained and then the real time data collection, data cache, stitching and transmission had accomplished centering on the Xilinx's Virtex-5 FPGA. Firstly, the automatic adjustment of brightness differences of the original images was preprocessed in order to improve the overall mosaic effect. Secondly, information detection of relative shift amount was completed by the use of phase correlation method to register the original images. Finally, the two adjacent images' overlap area was fused by the use of linear weighted fusion algorithm to make the mosaic image achieve a smoothly fading in and out transitional effect. Experimental results show that the imaging system is simple and reliable, and can effectively increase the field of view of observation. The stitched large field images are of high-definition and real time, with a certain degree of

收稿日期:2014-10-08; 修订日期:2014-11-10

基金项目:国家自然科学基金(61201376);陕西省自然科学基金(2012JQ8003)

作者简介:杨磊(1987-),男,硕士,主要从事图像与视频信号处理方面的研究。Email: lei\_best888@163.com

representativeness and practicality.

**Key words:** FPGA; large field of view imaging; real-time mosaic; phase correlation; linear weighted fusion

## 0 引言

视觉感知系统作为人类获取信息,感知外部世界的主要途径,具有举足轻重的作用,是获取目标信息的主要技术手段。随着航空、航天等领域的不断发展,大多数应用都需要一次性获取较大的视场信息,然而传统的视觉系统视场角较小,均不能满足这种需求<sup>[1]</sup>。目前的大视场成像技术还主要是基于后期的软件制作,像现在已经商业化的拼接软件: autostitch、妙手接图、全景拼接等软件,虽然拼接精度已经很高,但是都必须基于前期先拍摄原始图像,然后再后期载入软件进行拼接处理,可以说是没有实时性可言,达不到实际工程应用的目的,而且该类软件对输入组相邻两幅图像的重合度均有一定的要求(一般为不低于 30%),如果重合度较低则可能影响拼接精度,甚至会错拼,同时过高的重合比例也使得同样的视场需要采集更多的原始图像。基于硬件实施的研究方法中,参考文献[2]最后的图像融合部分没有说明具体是如何实现的,拼接部分融合效果欠佳,缝隙比较明显。

针对现有图像拼接方法在应用时的种种不足,文中结合可编程器件 FPGA 的广泛应用性和灵活性,研究并实现了一套完整地大视场实时成像系统。首先通过 3 块基于彩色 CMOS 图像传感器的成像板同时获取 3 路原始视频图像,并进行亮度差异自动校正调节,然后通过相位相关法对获取的图像进行配准检测,确定相邻两幅图像的重叠区域,得到水平方向的位移量信息;然后在可编程器件 FPGA 内部利用检测到的平移量信息,对相邻两幅有重叠部分的原始图像进行匹配处理,使 3 幅图像拼合成一幅完整的大视场图像;最后对配准之后图像的重合部分采用线性加权融合算法进行融合处理,达到渐进出平滑过渡的效果。

## 1 系统的总体设计

前端 3 个成像板上的镜头均选用 CCTV 的标准

C 接口手动变焦镜头,焦距  $f=6\text{ mm}$ ,光圈  $F=1.2$ ,  $1/3''F$ ,通过 CMOS 的成像靶面计算出水平视场角为  $30^\circ$ ,所以三路经过拼接之后的水平视场角应接近  $90^\circ$ 。

系统后端的控制处理模块以 Xilinx 公司的 Virtex-5 系列 XC5VLX110 为核心,它具有高达 17 280 个 Slice(一个 Virtex-5 Slice 具有 4 个 LUT 和 4 个触发器),含有 64 个 DSP48E Slice,多达 1 120 kb 的分布式 RAM 和 4 608 kb 的块 RAM,最大用户 I/O 为 800 个,可以满足各种复杂算法的设计与实现。

系统的总体结构如图 1 所示。

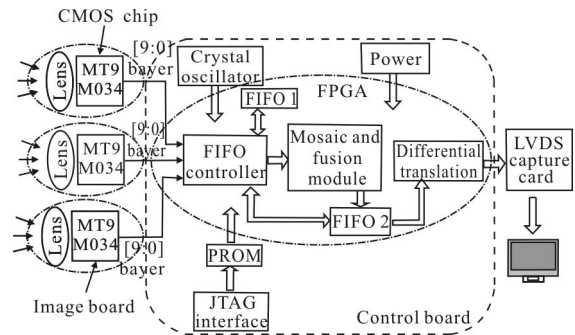


图 1 大视场成像系统的总体结构图

Fig.1 Structure of large field of view imaging system

前端采集到的原始视频信息同时送入 FPGA 内部,经过预处理、匹配、融合等操作之后通过 LVDS 接口经采集卡由显示器输出显示<sup>[3]</sup>。

## 2 系统硬件电路设计

大视场成像系统的硬件结构共分为电源、上电配置、数据采集、数据缓存、数据处理、输出接口等部分。

### 2.1 图像采集模块电路

图像采集模块是整个设计的一个重点,为后续的处理工作提供了“原材料”,它输出的数据质量将对后端的处理结果产生直接影响。在该模块设计中,图像传感器选用 APTINA 公司新一代基于高动态技术的彩色高清 CMOS 图像传感器 MT9M034,其与 FPGA 的控制电路连接关系如图 2 所示。

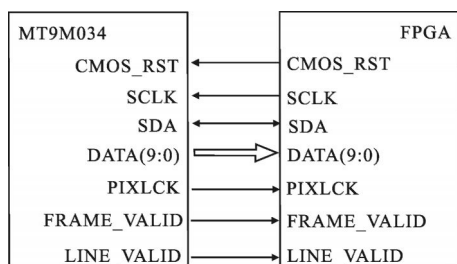


图2 MT9M034 控制电路连接框图

Fig.2 Diagram for MT9M034 control circuit connection

FPGA 芯片通过 I<sup>2</sup>C 数据总线对前端的 MT9M034 图像传感器的寄存器进行配置(包括芯片的行消隐、列消隐、起始位置、窗口宽度、高度等)<sup>[4]</sup>。在 FPGA 内部产生的时钟信号、应答信号和复位信号通过 I<sup>2</sup>C 接口送入 MT9M034 中,MT9M034 在输入时序的控制下将外部的光信号转换成电信号,以像素同步时序 PIXCLK 的速率将电信号通过数据端口输出至 FPGA,同时输出的还有帧同步信号 FRAME\_VALID 和行同步信号 LINE\_VALID。

## 2.2 输出接口设计电路

系统中采用 LVDS 接口进行输出显示。如图 3 所示,发送器将 TTL 电平信号送入 LVDS 模型中,在内部经过差分变换转换成一组正、负的差分对信号,然后通过互联器传输至接收器,在接收端通过相减消除共模干扰信号,得到真正的数据信号,需要注意的是每组差分对之间需要进行阻抗匹配。

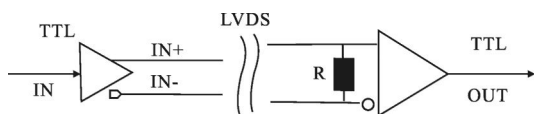


图3 LVDS 接口原理图

Fig.3 LVDS interface schematics

LVDS 既可以做输入接口,又可以做输出接口。做输入接口时需要在每对差分信号的正、负端之间加 100~120 Ω 的电阻进行阻抗匹配,在此系统中作为输出接口,所以不用外加电阻。输出信号包括:数据时钟同步信号(PIXCLK)、帧同步信号(FV)、行同步信号(LV)、数据信号(DATA0~DATA9),分别在 FPGA 内部做差分变换,所以输出共有 26 根数据线。

## 2.3 FPGA 外围连接电路

系统外围电路主要包括:电源模块、晶振电路、

JTAG 接口链路等。

在设计电源电路时分别对前端成像板和后端控制处理板进行分开供电,以确保电源的稳定性,输入电压均为直流+5 V;对 FPGA 核电压采取单独的开关电源进行供电,以减小波动。系统输入时钟频率为 74.25 MHz,在 FPGA 内部进行分、倍频操作,得到不同速率的时钟供其他模块使用,同时为了消除时钟传输时产生的反射波跳变现象,在晶振的输出管脚需要串接一个 33 Ω 的电阻。设计中 FPGA 的配置采用主串模式,即将模式控制端 M1、M2、M3 均接地置为 000 状态,选用 Xilinx 的 XCF32PFSG48C 配置芯片,核电压为 1.8 V,封装为 FSG48,容量为 32 M。

## 3 系统软件设计

整个拼接系统的软件设计主要包括:MT9M034 时序控制模块、基于 I<sup>2</sup>C 接口的 CMOS 配置模块、图像拼接模块、LVDS 接口控制模块,其中 I<sup>2</sup>C 配置模块又分为寄存器配置和 I<sup>2</sup>C 接口控制模块,各模块功能都在 Xilinx Virtex-5 FPGA 芯片内部完成。

### 3.1 成像模块驱动时序设计

MT9M034 的复位控制模块输入信号为系统时钟 24 MHz,输出信号有传感器外部时钟 ext\_clk,系统复位信号 reset 和传感器复位控制信号 cmos\_reset。

MT9M034 通过 SCLK 和 SDATA 两个数据,使用 I<sup>2</sup>C 接口传输数据来实现对片内寄存器的读写控制<sup>[4]</sup>。由于 I<sup>2</sup>C 总线频率较低,一般不会超过 400 kHz,所以在系统设计时对 24 MHz 时钟进行 480 倍分频得到 50 kHz 的时钟。

### 3.2 数据缓存模块

由于数据采集、数据处理和数据输出三者是跨时钟域的,所以必须用内部缓存来进行隔离,设计中在 FPGA 内部加入了两组大容量的 FIFO 控制模块。第 1 组有 3 个 FIFO 模块,用来同时缓存采集到的 3 路原始视频数据,当内部缓存两行数据以后,通过控制模块由第 1 组 FIFO 将数据读出进行拼接处理,处理后的数据再送入第 2 组 FIFO 内部,然后再通过内部的 FIFO 控制器输出至 LVDS 接口,经采集卡由上位机进行显示,工作模式如图 4 所示。

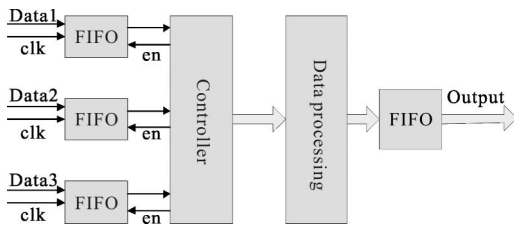


图 4 数据缓存模块功能图

Fig. 4 Functional diagram for data cache module

### 3.3 图像拼接算法

#### 3.3.1 亮度差异调整

在对原始图像做配准之前先进行亮度差异调整,以使得前后两幅图像的亮度趋于一致。具体的调整过程如下<sup>[5]</sup>:

(1) 设重叠区域像素均值分别为  $\bar{I}_1$  和  $\bar{I}_2$ , 计算两者的公共均值  $\bar{I}$ :

$$\bar{I} = \frac{\bar{I}_1 + \bar{I}_2}{2} \quad (1)$$

(2) 利用上面得到的公共均值对  $I_1$  和  $I_2$  进行亮度调节,得到拼接所需图像  $I_1^*$  和  $I_2^*$ ,即:

$$\begin{cases} I_1^* = I_1 - (\bar{I}_1 - \bar{I}) \\ I_2^* = I_2 - (\bar{I}_2 - \bar{I}) \end{cases} \quad (2)$$

在 FPGA 内部具体应用时,首先分别计算第 1 幅和第 2 幅图像重叠区域的平均灰度值,然后根据公式(1)计算二者的公共均值,利用该公共均值通过公式(2)进一步求得待拼接图像的平均灰度值,最后根据原始图像和待拼接图像的平均灰度关系,给每幅图像乘以相应的比例系数。

这样待拼接图像整体亮度将保持一致,不会明显偏亮或偏暗,既能提高图像的拼接精度,而且也能改善拼接痕迹,达到更好的融合效果。

#### 3.3.2 图像配准

系统中前端的 3 个成像板可以通过机械结构定位,保证垂直方向处于同一水平面,也就是 3 个成像板相互之间只存在水平方向的位移,所以采用效率较高的相位相关法<sup>[6]</sup>对原始图像进行匹配处理。相位相关法是借助于傅里叶变换将相邻两幅待拼接的图像映射到频域,根据傅里叶变换的移位性质:空域的平移对应频域的相移,然后通过计算相邻两幅图像的互功率谱密度求出相邻两幅图像间的平移量。

其基本原理如下<sup>[7]</sup>:假设相邻两幅图像分别用  $I_1(x, y)$  和  $I_2(x, y)$  表示,若两者之间除了平移变换以外,不存在其他仿射变换,即  $I_1(x, y)$  可用  $I_2(x, y)$  表示为:

$$I_1(x, y) = I_2(x - \Delta x, y - \Delta y) \quad (3)$$

若用  $\hat{I}_1(u, v)$ ,  $\hat{I}_2(u, v)$  分别表示  $I_1(x, y)$  和  $I_2(x, y)$  的傅氏变换,则对上式进行傅里叶变换即可得到:

$$\hat{I}_1(u, v) = \exp[-j2\pi(u\Delta x + v\Delta y)] \hat{I}_2(u, v) \quad (4)$$

计算  $\hat{I}_1(u, v)$  和  $\hat{I}_2(u, v)$  的归一化互功率谱密度为:

$$f_{\text{corr}}(u, v) = \frac{\hat{I}_1(u, v) \hat{I}_2^*(u, v)}{|\hat{I}_1(u, v) \hat{I}_2^*(u, v)|} e^{-j2\pi(u\Delta x + v\Delta y)} \quad (5)$$

式中:  $\hat{I}_2^*$  表示  $\hat{I}_2$  的复共轭。通过分析求得的互功率谱,可以发现其相位等同于两幅图像的相位差。

将上述所得互功率谱密度再做逆傅里叶变换,就可得到一个冲激函数:

$$f_{\text{corr}}(x, y) = F^{-1}[e^{-j2\pi(u\Delta x + v\Delta y)}] = \delta(x - \Delta x, y - \Delta y) \quad (6)$$

可以看出,该冲激函数在  $(\Delta x, \Delta y)$  处取得最大值,其它地方接近于零,而相邻两幅图像的相对位移就是  $(\Delta x, \Delta y)$ 。因此只要找到该冲激函数取得最大值时对应的位置,就可以确定出相对位移量  $(\Delta x, \Delta y)$ <sup>[7]</sup>。由于相位相关法中存在频域计算,考虑到硬件实现的局限性,所以利用相位相关法对图像的配准处理放在 MATLAB 软件中实施。

前端的 3 个成像模块定位好之后,分别将采集到的 3 幅图像保存下来,然后通过 MATLAB 软件利用相位相关法对相邻两幅图像进行检测,确定出彼此的相对平移量。在 FPGA 内部直接利用该相对平移量信息,通过系统的数据同步时钟 PIX\_CLK 和复位信号 SYS\_RST 产生相邻两幅图像拼接位置的控制信号,同时为后端的大视场图像采集提供所需的帧同步信号 FV 和行同步信号 LV。这是整个配准模块所需要实现的功能,如图 5 所示。

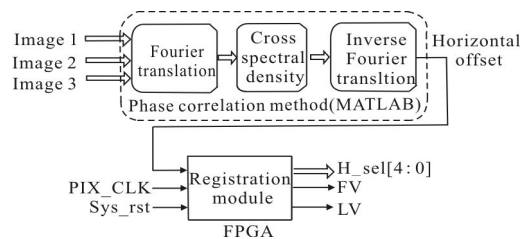


图 5 图像配准模块功能图

Fig. 5 Functional diagram for image registration module

### 3.3.3 图像融合

为了消除接缝达到更好的视觉效果,还需要对拼接位置进行融合处理。综合考虑算法的复杂度、有效性以及硬件实现的难易程度,该系统中视频融合模块使用的是线性加权融合算法<sup>[9]</sup>。应用时通过结构定位,相邻两路视频均有 40 列像素的重叠区域,将这 40 列像素每 4 列一组,由前到后共分为 10 组,每组乘以不同的线性权值进行融合处理。

假设  $M$ 、 $N$  表示待拼接两幅图像的重叠区域,  $I$  表示融合之后的图像信息。将  $M$ 、 $N$ 、 $I$  分别等分为 10 份,用  $M_0, M_1 \dots M_9, N_0, N_1 \dots N_9$  和  $I_0, I_1 \dots I_9$  表示。则线性加权之后:

$$I = d_1 \times M + d_2 \times N \quad (7)$$

这里,  $d_1$  由 8/8(即 1)渐变至 0/8(即 0),  $d_2$  由 0/8(即 0)渐变至 8/8(即 1),所以具体计算结果如下:

$$\begin{cases} I_0 = M_0 \times 8/8 + N_0 \times 0/8 \\ I_1 = M_1 \times 7/8 + N_1 \times 1/8 \\ I_2 = M_2 \times 6/8 + N_2 \times 2/8 \\ I_3 = M_3 \times 5/8 + N_3 \times 3/8 \\ I_4 = M_4 \times 4/8 + N_4 \times 4/8 \\ I_5 = M_5 \times 4/8 + N_5 \times 4/8 \\ I_6 = M_6 \times 3/8 + N_6 \times 5/8 \\ I_7 = M_7 \times 2/8 + N_7 \times 6/8 \\ I_8 = M_8 \times 1/8 + N_8 \times 7/8 \\ I_9 = M_9 \times 0/8 + N_9 \times 8/8 \end{cases} \quad (8)$$

可以看出在重叠区域由左图到右图变化时,左图的权值逐渐线性减小,同时右图的权值逐渐线性增大,达到一个渐进渐出的效果<sup>[9]</sup>。

3 幅图像在进行融合处理时整体可以被分为 5 个部分:第 1 幅的未重合区域、第 1 幅和第 2 幅的公共重合区域、第 2 幅的未重合区域、第 2 幅和第 3 幅的公共重合区域、第 3 幅的未重合区域。在进行数据读取时,每行数据都是按照该顺序进行操作,未重合区域直接读取亮度调整之后的数据,而相邻的重合区域则按照上述的线性加权算法进行处理。如图 6 所

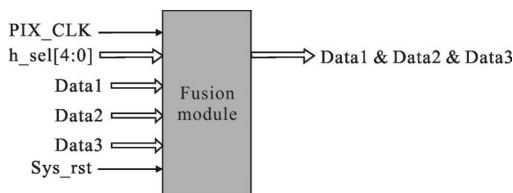


图 6 图像融合模块功能图

Fig. 6 Functional diagram for image fusion module

示,经过融合处理之后的大视场图像将同时包含前端输入的三路视频图像的所有信息。

## 4 实验结果及分析

将系统前端的 3 个成像板分别通过排线与后端的控制处理板相连,时钟和同步信号等关键信号线均与地线双绞。3 个成像板分别通过 1 个金属小支架固定在底座上,如图 7 所示,小支架可以沿水平方向转动,以使 3 个板之间相互成一定的角度。

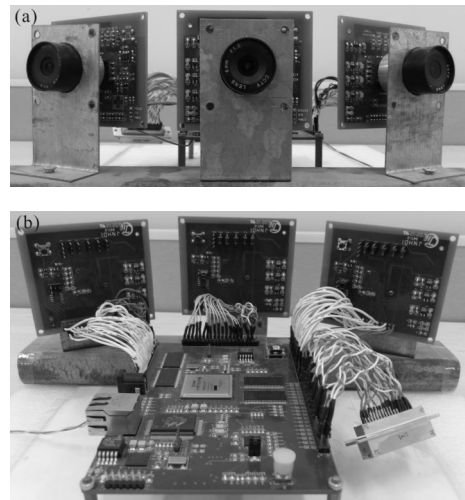


图 7 大视场成像系统工作实物图

Fig. 7 Large field imaging system

图 8 给出了前端 3 个 CMOS 摄像头分别由 Matrox II 图像采集卡采集到的视频图像,大小均为 720×560。可以看出 3 个摄像头成像质量良好,相邻两幅图像均有一定的重合区域。



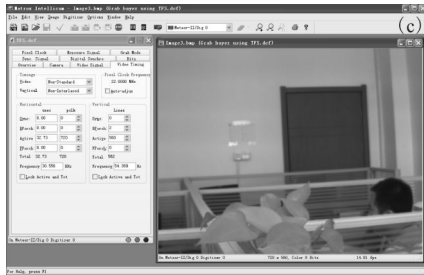


图 8 前端 3 个 CMOS 摄像头分别采集到的视频图像  
Fig.8 Images collected by three CMOS cameras from front end

经过实时拼接系统处理之后,采集到的视频图像如图 9 所示,大小为 1 920×560,帧频为 10 fps,成像清晰,较拼接之前视频质量基本没有降低。

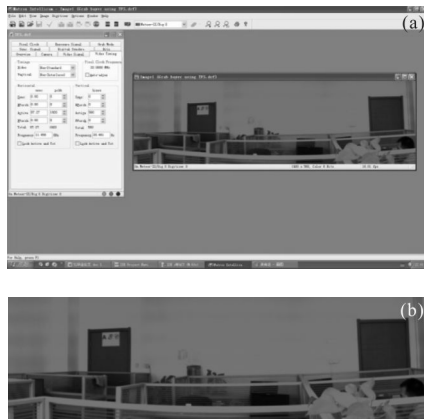


图 9 拼接处理之后的大视场图像  
Fig.9 Large field mosaic image

表 1 是拼接前后的参数对比,前端每个摄像机水平视场为 30°,经过拼接处理之后水平视场扩大到了 80°,所以相邻两路相机的光轴夹角为 25°。可以看出经过大视场成像系统处理之后的图像尺寸和水平视场角均较拼接之前有很大提高,并且无明显拼接痕迹,具有良好的大视场实时成像效果,达到了预期的设计目标。

表 1 拼接前后参数对比

Tab.1 Parameters before and after stitching

Parameter	Before stitching	After stitching
Pixels	720×560	1 920×560
Horizontal angle/(°)	30	80
Vertical angle/(°)	24	24
Clock frequency/MHz	8	24
Frame frequency/fps	10	10

图 10 给出了利用精度更高的 Harris 算法通过

Matlab 软件进行匹配拼接的效果,分别从标准差、空间频率和拼接时间 3 方面,对该大视场实时成像系统和 Harris 算法得到的图像进行对比,比较结果如表 2 所示。



图 10 利用 Harris 算法得到的大视场图像  
Fig.10 Large field images by the use of the Harris algorithm

表 2 两种方法性能对比

Tab.2 Performance contrast of both methods

	Standard deviation	Spatial frequency	Mosaic time/s
Harris algorithm	42.785 0	4.617 5	16.99
Proposed system	39.758 3	4.421 4	0.10

通过比较可以看出,该系统输出的大视场图像,在标准差和空间频率方面与 Harris 算法通过 Matlab 软件得到的图像基本一致,无明显差别,但是在整个图像拼接的运算时间上,该系统占据了很大优势,可以达到实时动态拼接的效果,具有非常重要的应用价值。

### 5 结 论

文中研究并设计了一种基于可编程器件 FPGA 的大视场实时成像系统。系统中设计的两个独立缓存模块,有效地解决了数据采集、数据处理以及数据传输之间的同步问题,图像拼接部分充分利用了相位相关法的稳健性,在检测平移量信息之前先对图像进行亮度差异自动调节,既保证了图像的拼接精度,又进一步提高了图像的融合效果<sup>[10]</sup>。图像融合部分合理的将相邻图像的重叠区域进行分组处理,满足整体融合效果的同时又很大程度上减小了计算量。实验结果表明,文中的大视场成像系统拼接精度高、图像重合部分融合效果好,而且实时性很强,达到了动态拼接的效果。

### 参考文献:

[1] Huang Qiongdan, Qiu Yuehong, Tian Xiaoping. Image automatic mosaic method based on feature points and optimization theory[J]. *Acta Photonica Sinica*, 2009, 38(8):

- 2139–2143. (in Chinese)  
黄琼丹, 邱跃洪, 田小平. 基于特征点及优化理论的图像自动拼接方法[J]. 光子学报, 2009, 38(8): 2139–2143.
- [2] Peng Bo, He Bin. Application and realization of FPGA in video mosaicing [J]. *Computer Engineering and Design*, 2013, 34(5): 1635–1639. (in Chinese)  
彭勃, 何宾. FPGA在视频拼接中的应用与实现 [J]. 计算机工程与设计, 2013, 34(5): 1635–1639.
- [3] Dong Linli. Design of digital image mosaic system based on FPGA [D]. Chengdu: Chengdu University of Technology, 2012. (in Chinese)  
董林粒. 基于FPGA的数字图像拼接器系统设计 [D]. 成都: 成都理工大学, 2012.
- [4] MT9M034 Data Sheet. Aptina Imaging[M]. 2010.
- [5] An Jianni, Liu Guixi. Image mosaic algorithm based on feature points matching and Automatic transform parameters identifying [J]. *Infrared and Laser Engineering*, 2011, 40(3): 564–569. (in Chinese)  
安建妮, 刘贵喜. 利用特征点配准和变换参数自动识别的图像拼接算法[J]. 红外与激光工程, 2011, 40(3): 564–569.
- [6] Kuglin C, Hines D. The phase correlation image alignment method [C]//Conference on Cybernetics and Society, IEEE, 1975: 163–165.
- [7] Wu Xianxiang, Guo Baolong, Wang Juan. An improved automatic image sequencing algorithm [J]. *Journal of Optoelectronics Laser*, 2009, 20(8): 1114–1117. (in Chinese)  
吴宪祥, 郭宝龙, 王娟. 一种改进的序列图像自动排序算法[J]. 光电子激光, 2009, 20(8): 1114–1117.
- [8] Yang Lei, Cao Jianzhong, Wang Hao, et al. Multifocus image fusion algorithm based on the contourlet transform[C] //SPIE, AOMATT 2012, 84201F: 1–6.
- [9] Shang Mingshu. A fast image mosaic algorithm based on improved SIFT feature matching [J]. *Microelectronics & Computer*, 2014, 31(1): 64–67. (in Chinese)  
尚明妹. 基于改进SIFT特征匹配的快速图像拼接算法[J]. 微电子学与计算机, 2014, 31(1): 64–67.
- [10] Li Liang, Gu Guohua, Qian Weixian, et al. Infrared image sequence mosaic based on feature points and Poisson fusion [J]. *Infrared and Laser Engineering*, 2013, 42(9): 2584–2588. (in Chinese)  
李亮, 顾国华, 钱惟贤, 等. 基于特征点和泊松融合的红外序列图像拼接[J]. 红外与激光工程, 2013, 42(9): 2584–2588.