

# 可重构视频阵列处理器簇内存储结构设计与实现

郭佳乐, 蒋 林, 山 蕊, 崔朋飞, 武 鑫  
(西安邮电大学 电子工程学院, 陕西 西安 710121)

**摘 要:** 提出了一种簇内高效并行访问存储结构, 该结构采用“逻辑共享、物理分布”多个存储块并行存储的方法, 实现了  $4 \times 4$  视频阵列处理器的并行访问。实验结果表明, 在无冲突情况下, 该结构支持 16 个轻核处理元的同时读/写操作, 最高频率 200 MHz, 访存峰值带宽 6.25 GB/s。最后对  $8 \times 8$  二维离散余弦变换算法进行映射实现和性能比较发现, 簇内存储结构能够为该算法提供 312.2 Msamples/s 的数据访存带宽, 相较于同类型阵列结构, 执行周期数降低了 31.67%, 工作频率提高了一倍, 访存带宽增加了 192.60%。

**关键词:** 阵列处理器; 存储结构; 并行存储; 视频编解码

**中图分类号:** TP302

**文献标识码:** A

**文章编号:** 1000—7180(2017)09—0116—05

## Design of Cluster Memory Structure for Reconfigurable Cideo Array Processor

GUO Jia-le, JIANG Lin, SHAN Rui, CUI Peng-fei, WU Xin

(School of Electronic Engineering, Xi'an University of Posts and Telecommunications, Xi'an 710121, China)

**Abstract:** A high efficient and parallel access memory structure is proposed. The architecture adopts the method of "logical sharing, physical distribution" and parallel storage of multiple memory blocks, which realizes the parallel access of  $4 \times 4$  video array processors. The experimental results show that the proposed architecture can support simultaneous read/write operations of 16 light nuclear processing elements, the highest frequency is 200 MHz, access to the peak bandwidth of 6.25 GB/s. Finally, the  $8 \times 8$  two-dimensional discrete cosine transform algorithm is mapped and compared. It is found that the cluster memory structure can provide data storage bandwidth of 312.2 Msamples/s. Compared with the same type of array structure, the number of execution cycles decreased 31.67%, frequency doubled, memory bandwidth is increased by 192.60%.

**Key words:** array processor; memory structure; parallel storage; video codec

### 1 引言

随着视频阵列处理器内部计算核数目增多, 集成的功能日益复杂, 对主存的数据访问需求也逐渐增加, 且处理速度与存取速度失配所引发的“存储墙”问题<sup>[1]</sup>也随着工艺的进步而日益严重, 成为制约处理器性能提高的重要因素。Godson-T 众核处理器<sup>[2]</sup>、嵌入式众核流处理器 TILE64 结构<sup>[3]</sup>, 以及 Cell 处理器<sup>[4]</sup>中, 均采用多级 Cache 技术来缓解不

断恶化的“存储墙”问题。然而, 多级 Cache 结构需要附加电路才能完成地址映射, 保证数据的一致性。且随着集成电路工艺的不断进步, 片上集成的处理器核数显著增加, 附加电路也会随之增多, 采用多级 Cache 存储结构不仅会增加电路的复杂度, 而且会消耗更多的能量。文献[5]提出的 V-RCA 结构, 虽然不包含任何缓存机制, 在一定程度上减少了多级 Cache 技术带来的电路复杂问题, 但该结构仅采用一块本地数据存储, 这不仅减小了数据访存带宽, 加

收稿日期: 2016—12—02; 修回日期: 2017—01—28

基金项目: 国家自然科学基金资助项目(61272120, 61634004, 61602377); 陕西省自然科学基金资助项目(2015JM6326); 陕西省科技统筹创新工程项目(2016KTZDGY02—04—02)

剧了数据访存难度和复杂度,而且降低了处理器的访存能力.为此,研究阵列处理器簇内存储结构,满足视频算法对存储带宽的要求,提高访存过程的并行性,并在此基础上缓解处理器与存储器之间的“存储墙”问题,解决多级 Cache 技术带来的附加电路增多、存储架构复杂,无 Cache 结构访问复杂度高、处理器访存灵活性差等问题变得日益迫切.

2 逻辑共享、物理分布簇内存储结构

本文采用的可重构视频阵列处理器,在实现数据级并行计算和指令级并行计算的同时,解决了数据并行算法和非数据并行算法的计算高效性和编程灵活性问题<sup>[6]</sup>.阵列处理器中每  $4 \times 4$  个 PE 组成一个阵列处理器簇 (Processing Element Group, PEG),每个 PE 的访存数据位宽为 16 位,可以应用于多种视编解码标准<sup>[7]</sup>.簇内存储结构如图 1 所示,其上层为  $4 \times 4$  个 PE 组成的 PEG,下层为  $4 \times 4$  个  $512 \times 16$  bit 大小的 RAM 块,簇内存储结构在可重构视频阵列处理器和分布式存储间形成一个高速数据交互通路,实现数据的高效访问.从 PE 角度看,每个 PE 能够直接访问片上所有存储资源,是片上共享存储结构;物理实现上,采用多个独立的 RAM 块,通过簇内存储结构实现对本地存储单元优先访问,簇内远程存储单元次之的优先级策略,同时利用多个存储块的并行存储技术,实现了  $4 \times 4$  个 PE 的并行访问.

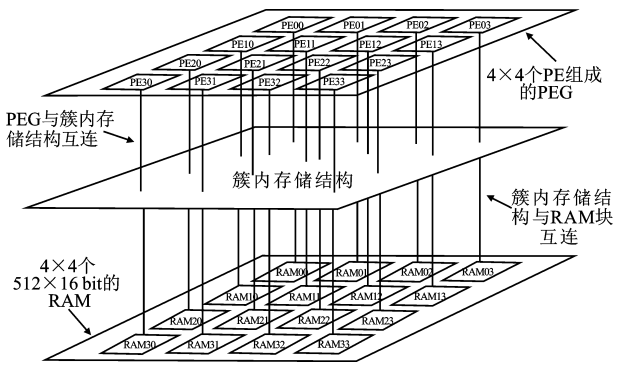


图 1 逻辑共享、物理分布簇内存储结构

当 PE 访问存储单元时,簇内存储结构接收来自 PE 的读写请求,根据地址判断是对本地存储单元访问还是对簇内远程存储单元访问.当 PE 访问本地存储单元时,由于其优先级最高,可无阻塞直接访问;当 PE 访问簇内远程存储单元时,簇内存储结构会根据访问的地址判断出所要访问的具体 RAM 块,并选出对应的访问路径进行数据传输.由于优先

级策略的存在,远程访问的响应不一定及时,此时会先将接收到的读写访问操作进行保持,待响应后再进行释放.当簇内存储结构还没有完成对当前 PE 发送的访问操作处理时,若 PE 再发送一组数据则可能覆盖前一组数据,发生数据丢失问题.为了很好地解决这个问题,簇内存储结构通过增加读写反馈信号,巧妙的将当前存储器的状态告知 PE,具体过程:若存储结构在当前时刻正在对 PE 发来的读写请求进行处理,则该 PE 对应的读写反馈信号为高,此时的数据为有效数据,下一拍 PE 可重新发送一组数据进行访问;若当前存储结构正在对其他 PE 发送的读写请求进行处理,则该 PE 对应的读写反馈信号为低,此时数据无效,下一拍 PE 不能发送新的数据进行访问.

3 簇内存储结构的设计实现

在设计可重构视频阵列处理器簇内存储结构时,考虑到簇内存储结构的主要功能是完成可重构视频阵列处理器与分布式存储之间的数据交互,如:访问本地存储单元、访问簇内远程存储单元、无冲突访问时对不同存储单元的同时访问,以及冲突访问时本地 PE 访问本地存储单元优先于簇内远程 PE 的访问等.将顶层结构划分为三大电路模块,如图 2 所示分别为:分配模块、仲裁模块和总线模块.总线模块又分为 4 个行总线和 4 个列总线,4 个行总线分别与对应的分配模块相连,4 个列总线分别和对应的仲裁模块相连,并且将行、列总线进行互连,最终完成整个结构的搭建,实现过程如下.

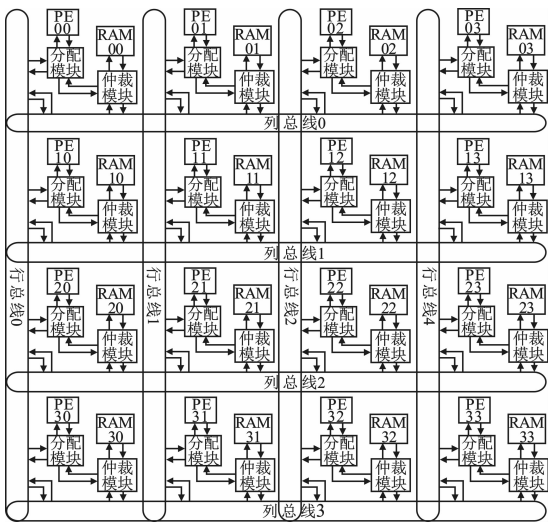


图 2 簇内存储结构图

(1)分配模块

为了实现 PE 对簇内任意存储单元的数据交互访问,区分是对本地存储单元进行访问还是对簇内远程存储单元进行访问,并解决后续模块能够实现对本地存储单元进行优先访问的问题. 如图 3 所示,分配模块接收来自 PE 发送的读写访问信号并根据读写地址  $addr$  的 9 到 12 位进行判断(高两位表示所要访问存储单元的行地址,低两位表示列地址),将读写操作分配到仲裁模块或者总线模块. 其中,若和所传参数  $local\_id$  相同表明是对本地存储单元进行访问,则将信号分配到仲裁模块,反之则分配到总线模块.

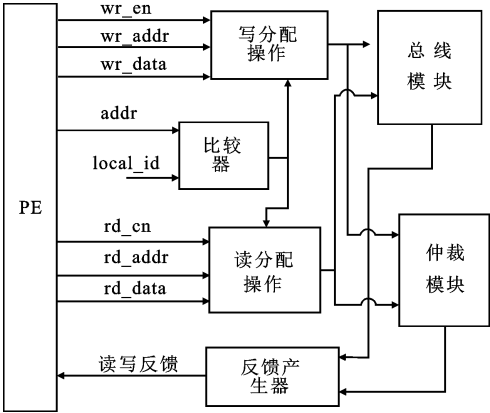


图 3 分配模块结构框图

(2)总线模块

考虑到该模块主要是在无冲突访问时能够正确处理分配模块发送的远程访问信号,以及冲突访问时的冲突避免与处理,所以将总线模块分为行总线和列总线. 如图 4 所示为行总线模块结构框图(列总线设计思路与之相同),其中,由于 1 个行总线与 4 个列总线相连,且每个 PE 都能对簇内的任意存储单元进行访问,所以行总线接收分配模块发送的读写请求信号时,会将读写使能进行列选择,选择出具体访问的列总线,而在同一时刻行总线可能会接收到来自 4 个分配模块发送的读写访问信号,这样就会发生冲突. 为此,使能仲裁模块采用轮询机制(Round Robin)<sup>[8]</sup>实现 4 路读写使能信号的仲裁选择,并产生一个 4 位使能信号,通过高有效位来标识哪一路信号为有效信号(同一时刻有且仅有一个高有效位). 为了能够正确选择 4 路访存数据的列总线输出,在通路选择时,采用 4 选 1 多路选择器通过使能仲裁模块仲裁出的 4 位使能信号  $wr\_en/rd\_en$  作为多路选择器的控制端,进行读/写使能信号、读/写地址信号、写数据信号的选择. 在处理过程中,由于存储器需要将读数据发送回给 PE,所以在对读数据

进行选择时,需同时采用 4 路读使能信号来作为选择条件进行数据选择.

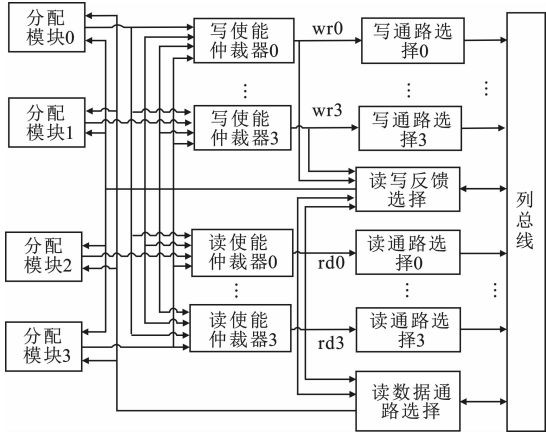


图 4 行总线模块结构框图

(3)仲裁模块

由于分配模块的读写使能来自本地 PE,而总线模块的读写使能来自簇内远程 PE,所以分配模块写使能优先级高于总线模块,当分配模块读写使能有效时,立即对其进行响应,而总线模块的读写使能则必须在分配模块读写使能无效时才能得到响应. 为了使分配模块的优先级高于总线模块,并完成两路数据进入 RAM 前的数据选择与优先级判断,仲裁模块通过使能产生器生成进入 RAM 的读写使能信号. 如图 5 所示,当分配模块有数据请求时,优先响应,阻塞总线模块的数据请求;当分配模块无数据请求且总线模块有数据请求时,响应总线模块的数据请求.

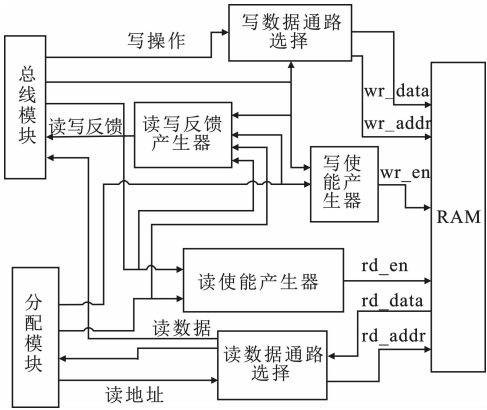


图 5 仲裁模块结构框图

在同一时刻簇内存储结构可能会接收来自不同 PE 发送的读写请求信号,且在有冲突情况下,由于优先级的存在,响应可能不能及时完成,为了防止数据丢失,簇内存储结构通过反馈模块根据当前所处理的数据通路进行读写反馈,若反馈信号为高则说

明读写请求得到响应,并将请求使能拉低,相反,则告知 PE 将对应的请求信号继续保持直到得到响应为止。

4 原型开发及算法实现

4.1 簇内存储结构硬件实现

采用 Verilog HDL 硬件描述语言,在 Questa-sim 10. 1d 工具下进行功能仿真验证,选用 Xilinx 公司的 ZYNQ 系列芯片 XC7Z045 FFG900-2 FPGA 进行综合测试。其中芯片资源使用情况如表 1 所示,测试结果表明,簇内存储结构在无冲突情况下,支持 16 个 PE 的同时读/写操作,读操作 2 个周期完成,写操作 1 个周期完成;冲突情况下,支持冲突的正确避免与轮询访存。最后通过 Synopsys 公司 Design Compiler 在 SMIC 0. 13 μm CMOS 工艺标准单元库下综合,频率可达 200 MHz。

表 1 FPGA 芯片资源使用报告表

| 逻辑器件            | 资源利用率             |
|-----------------|-------------------|
| Slice Registers | 1 401/437 200(0%) |
| Slice LUTs      | 6 464/218 600(2%) |
| LUT-FF pairs    | 1 300/6 565(19%)  |

4.2 视频算法的 FPGA 实现与测试

为了验证簇内存储结构的可行性和正确性,将该存储结构与可重构阵列结构进行互连,同时,选取了 MPEG-2 解码中 8×8 2D-DCT 算法进行算法映射与分析,具体过程为

$$QF''[v][u] = (((QF[v][u] \times 2) + \text{Sign}(QF[v][u] \\ [u])) \times W[w][v][u] \times \text{quantiser\_scale}) / 32 \tag{1}$$

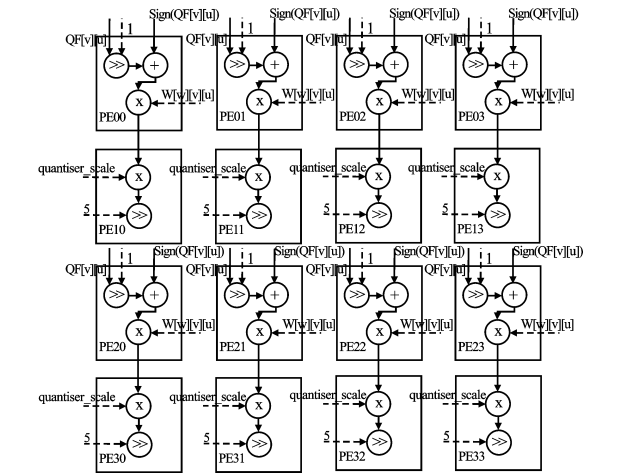


图 6 MPEG-2 解码中 2D-DCT 算法的映射结构图

选取系数最多的 macroblock\_non\_intra AC 系数情况(见式 1)映射到该可重构阵列结构上,将量化比例因子 quantiser\_scale 和权系数 W[w][v][u] 作为立即数输入,其算法映射数据流图如图 6 所示,其中常数 1 和 5,均从各自 PE 对应的本地存储中进行加载,Sign(QF[v][u])、量化比例因子 quantiser\_scale 和权系数 W[w][v][u] 分别从 PE00、PE11 和 PE22 所对应的本地存储中加载,而所需要的 QF[v][u] 则从 PE11 所对应的本地存储中进行加载,并通过邻接互连传给相邻 PE01 和 PE21,PE00 通过 PE01 进行邻接互连传递,PE20 通过 PE21 进行邻接互连传递,剩余的 PE02、PE03、PE22、PE23 均通过簇内存储结构进行数据加载。

测试结果表明,簇内存储结构能够正确处理该算法的数据访问,在无冲突的情况下,支持 16 个 PE 的同时读/写操作,达到读操作 2 个周期完成,写操作 1 个周期完成;在有冲突的情况下,可正确发送反馈信号告知相应 PE,使其进行访问数据的保留,避免了数据的丢失。于此同时,该阵列结构可一次并行计算 8 个系数,并且可以通过流水的方式循环 8 次来完成一个 8×8 宏块的处理,最终在算法的正确处理下,执行周期数为 41。

5 性能分析和对比

5.1 性能分析

簇内存储结构带宽计算为  
带宽=工作频率×存储器数据总线位宽/8×无冲突概率  $\tag{2}$

其中,存储器数据总线位宽=单 RAM 块位宽×当前发送访问的 PE 数目;无冲突概率=同一时刻处理的 PE 访问数目/当前发送访问的 PE 数目。

根据簇内存储结构特性可以将 PE 访问存储单元的情况分为以下几种典型情况来讨论:

(1)4×4 个 PE 同时对本地存储单元进行访问,由于 PE 对本地存储单元访问优先级最高,所以,同一时刻簇内存储结构所能处理的 PE 数目为 16 个,访问不会发生冲突,无冲突概率为 100%,此时带宽达到峰值 6. 25 GB/s。

(2)多个 PE 对同一个存储单元进行访问,由于 RAM 块同一时刻只能处理一组访问操作,当出现最坏情况即 16 个 PE 同时对同一个 RAM 进行读写操作时,此时出现全冲突访问,无冲突概率为 1/16,此时带宽最小为 400 MB/s。

(3)多个 PE 对不同的存储单元进行访问,由于

簇内存存储结构可实现冲突避免以及对本地存储单元优先访问的功能,假设当前发送访问的 PE 数目为 M,簇内存存储结构同一时刻所能处理的 PE 数目为 N,则无冲突概率为  $N/M$ ,此时带宽为  $N \times 400 \text{ MB/s}$ .

为了更好地理解上述分析过程,以本文所映射的 MPEG-2 解码中  $8 \times 8$  的 2D-DCT 算法为例,对一个  $8 \times 8$  的处理块而言,每个像素的位宽为 16 位,共处理 64 个像素,因此所需数据量为 128 Byte,当对所需要的  $QF[v][u]$  进行数据加载时,同时发送访问请求的 PE 数目为 6,簇内存存储结构同一时刻所能处理的 PE 数目为 1,则无冲突概率为  $1/6$ ,访存带宽为 400 MB/s;当对常数 1 和 5 进行数据加载时,16 个 PE 同时访问本地存储单元,此时簇内存存储结构所能处理的 PE 数目为 16 个,访问不会发生冲突,无冲突概率为 100%,带宽达到峰值为 6.25 GB/s.而在处理过程中当对  $Sign(QF[v][u])$ 、量化比例因子  $quantiser\_scale$  和权系数  $W[w][v][u]$  进行数据加载时,出现了 16 个 PE 交叉访问不同存储的情况,此时簇内存存储结构所能处理的 PE 数目为 7,则无冲突概率为  $7/16$ ,访存带宽为 2.73 GB/s.

5.2 结果对比

根据前文对簇内存存储结构的描述,结合该存储结构“逻辑共享、物理分布”和多个存储块并行存储的特点,以及 MPEG-2 解码中  $8 \times 8$  2D-DCT 算法在基于该簇内存存储结构的可重构视频阵列结构上的 FPGA 实现与性能分析,如表 2 所示,本文对比了同类型阵列结构中 MPEG-2 解码  $8 \times 8$  2D-DCT 视频算法的实现性能.

| 表 2 性能比较 |            |            |         |
|----------|------------|------------|---------|
|          | V-RCA 结构   | 本文         | 性能提升    |
| 核数       | 16         | 16         | —       |
| 工作频率     | 100 MHz    | 200 MHz    | 100.00% |
| 周期数      | 60         | 41         | 31.67%  |
| 执行时间     | 600 ns     | 205 ns     | 65.83%  |
| 访存带宽     | 106.7      | 312.2      | 192.60% |
|          | Msamples/s | Msamples/s |         |

文献[5]中 V-RCA 结构不包含缓存机制,仅采用一块本地数据存储单元,每一次数据访问都必须经过数据读取和分发、输出数据拼接和写入才能完成数据的访存,并且一次读取最多只有 128 位数据,这使得访存带宽降低的同时也加剧了 PE 的数据访存难度.如表 3 所示,相较于此,本文所采用的存储结构可同时提供 16 个 16 位数据的访问,省去了数据的拆分和拼接过程,在实现片上共享存储的同时又为  $4 \times 4$  个 PE 提供了本地存储单元的优先访问,使

得读操作 2 个周期完成,写操作 1 个周期完成,减小了 PE 的数据访存难度和复杂度,最终在执行周期数降低 31.67%的同时,工作频率提高了一倍,访存带宽增加了 192.60%.

6 结束语

针对视频阵列处理器对数据的访存问题设计了一种基于可重构视频阵列处理器簇内存存储结构,该结构去除了多级 Cache,采用统一寻址方式的分布式共享存储架构,使存储结构在逻辑上共享、物理上分布,同时利用多个存储块的并行存储技术,实现了  $4 \times 4$  个 PE 同时对本地存储单元的访问,无冲突情况下对簇内远程存储单元的同时访问,有冲突情况下对冲突的避免以及对本地存储单元的优先访问等功能.分析表明,该存储结构在算法正确处理的情况下,执行周期数降低了 31.67%,工作频率提高了 100.00%,访存带宽增加了 192.60%,实现了执行周期数少、访存带宽高的特点.

参考文献:

[1] Wulf W A, Mckee S A. Hitting the memory wall: implications of the obvious[J]. Acm Sigarch Computer Architecture News, 1995, 23(1):20-24.

[2] Fan D, Zhang H, Wang D, et al. Godson-T: An efficient many-core processor exploring thread-level parallelism[J]. Micro IEEE, 2012, 32(2):38-47.

[3] Berezecki M, Frachtenberg E, Paleczny M, et al. Power and performance evaluation of Memcached on the TILEPro64 architecture[J]. Sustainable Computing Informatics & Systems, 2012, 2(2):81-90.

[4] Kahle J A, Day M N, Hofstee H P, et al. Introduction to the cell multiprocessor[J]. Ibm Journal of Research & Development, 2005, 49(45):589-604.

[5] 王浩. 基于视频压缩算法的硬件模板设计和可重构阵列架构研究[D]. 上海:上海交通大学, 2011.

[6] 孙璐, 沈绪榜. FPGA 阵列芯片的自然统一研究[J]. 微电子学与计算机, 2013, 30(1):1-4.

[7] 吕青, 蒋林, 邓军勇, 等. 面向对数与指数函数的可重构阵列结构[J]. 微电子学与计算机, 2016, 33(10):28-31.

[8] Paul A, Khan M H, Rahman M M, et al. Reconfigurable parallel architecture of high speed round robin arbiter[C]//Electrical, Electronics, Signals, Communication and Optimization (EESCO), 2015 International Conference on. IEEE, 2015: 1-5.

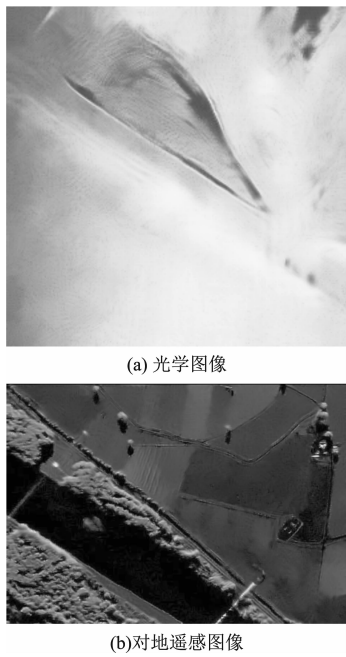


图 4 图像增强结果

目标快速融合识别的应用需求。

表 1 识别准确率对比

| 样本数  | 本文方法  | Gabor 方法 | LBP 方法 | % |
|------|-------|----------|--------|---|
| 100  | 99.85 | 92.11    | 90.08  |   |
| 500  | 97.63 | 92.14    | 89.32  |   |
| 1000 | 98.44 | 95.43    | 93.21  |   |
| 1500 | 99.08 | 89.32    | 91.09  |   |
| 2000 | 99.36 | 90.34    | 94.77  |   |

## 5 结束语

本文研究了遥感图像的融合和优化识别问题,提出一种基于分块 Radon 尺度变换信息增强的图像融合算法,对采集的模糊遥感目标图像进行小波降噪处理,对降噪输出的图像进行自适应模板匹配,结合图像分割方法将目标图像进行分块,利用 Radon 尺度变换的几何特征不变性对目标的关键特征点进行信息增强,实现特征点优化提取和目标准确辨识。实验分析得知,采用该方法进行图像融合和目标识别的准确性较好,对大样本的目标图像融合识别的准确性较高,实时性较好,具有较高的应用价值。

## 参考文献:

[1] 思宇,雷印杰,周新志. 针对单样本人脸的三维部分人脸识别[J]. 计算机工程, 2016, 42(9): 144-150.

[2] 蒋政,程春玲. 基于 Haar 特性的改进 HOG 的人脸特征提取算法[J]. 计算机科学, 2017, 44(1): 303-307.

[3] 邱玉,赵杰煜,汪燕芳. 结合运动时序性的人脸表情识别方法[J]. 电子学报, 2016, 44(6): 1307-1313.

[4] 李昕昕,龚勋. 三维人脸建模及在跨姿态人脸匹配中的有效性验证[J]. 计算机应用, 2017, 37(1): 262-267.

[5] 黄森,王刘涛,张海朝. 基于 Gabor 小波变换与 K-L 高斯黎曼流形判别的人脸识别[J]. 计算机工程, 2016, 42(9): 208-213.

[6] 刘婷,程建. 基于训练字典的遥感图像融合[J]. 计算机工程与应用, 2013, 49(19): 135-140.

[7] 冯彦超,冯华君,徐之海,等. 基于分色校正的大像差光学成像系统图像处理技术[J]. 光子学报, 2015, 44(6): 33-38.

[8] 沈学举,刘旭敏,许芹祖,等. 球面波照明下傅里叶变换全息多图像加密方法研究[J]. 激光技术, 2014, 38(4): 488-493.

[9] El-MEZOUAR M C, KPALMA K, TALEB N, et al. A pan-sharpening based on the non-subsampled contourlet transform: application to worldview-2 imagery [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2014, 7(5): 1806-1815.

[10] Moreno-Salinas D, Pascoal A M, Aranda J. Optimal sensor placement for multiple target positioning with range-only measurements in two-dimensional scenarios [J]. Sensors, 2013, 13(8): 10674-10710.

[11] 宋涛,李鸥,刘广怡. 基于空时多线索融合的超像素运动目标检测方法[J]. 电子与信息学报, 2016, 38(6): 1503-1511.

## 作者简介:

张荆沙 男, (1983-), 硕士研究生, 讲师. 研究方向为自动化. E-mail: zhang\_zjsha@163.com.

(上接第 120 页)

## 作者简介:

郭佳乐 女, (1992-), 硕士研究生. 研究方向为集成电路系统设计. E-mail: 315762794@qq.com.

蒋林 男, (1970-), 教授. 研究方向为专用集成电路设计.

山蕊 女, (1986-), 讲师. 研究方向为 VLSI 系统与计算机

辅助设计.

崔朋飞 男, (1990-), 硕士研究生. 研究方向为集成电路系统设计.

武鑫 女, (1992-), 硕士研究生. 研究方向为集成电路系统设计.