

# 无失真并行数据压缩的脉动阵列 ASIC 设计

顾 静<sup>1</sup>, 帅典勋<sup>1</sup>, 顾 清<sup>2</sup>

(1. 华东理工大学计算机科学与工程系, 上海 200237; 2. 清华大学智能技术和系统国家重点实验室, 北京 100084)

**摘 要:** 本文提出适用于无失真并行数据压缩的超大规模 ASIC 的逻辑电路设计. 与其他传统的串行或小规模并行无失真数据压缩的硬件或软件方法相比, 本文的 Systolic 阵列结构有更好的并行性、实时性和普适性. 对 ASIC 的时序和功能进行的模拟验证, 证明了逻辑和电路设计的正确性和有效性.

**关键词:** 数据压缩; 遗传算法; 功能模拟; 时序模拟

**中图分类号:** TP18 **文献标识码:** A **文章编号:** 0372-2112 (2000) 09-0135-02

## A Systolic ASIC Scheme for Undistorted Parallel Data Compression

GU Jing<sup>1</sup>, SHUAI Dian-xun<sup>1</sup>, GU Qing<sup>2</sup>

(1. Department of Computer Science and Engineering, East China University of Science and Technology, Shanghai 200237;

2. State Key Laboratory of Intelligence Technology and System, Tsinghua University, Beijing 100084, China)

**Abstract:** This paper presents a new systolic ASIC scheme to implement the undistorted parallel data compression algorithm based on genetic algorithm (GA). In comparison with other traditional sequential or small-scale parallel methods of data compression, the proposed systolic scheme has much higher parallelism, real-time performance and more suitability. The simulation of timing performance and functionality for the ASIC scheme verifies its correctness and effectiveness.

**Key words:** data compression; genetic algorithm; timing simulation; functional simulation; ASIC (Application Specific Integrated Circuit)

### 1 引言

数据压缩是信息论中信源编码的主要问题之一. 通常的串行和小规模并行数据压缩方法的实时性差, Systolic 阵列的结构复杂. 神经网络方法虽然已广泛应用于不可逆数据压缩, 但是对于可逆数据压缩仍然存在困难. 随着超大规模集成电路 (VLSI) 技术的发展, 历来由软件实现的许多算法可以在 VLSI 单片上实现. 但是, 计算机上的高效算法不一定适合于 VLSI. 适合于硬件化的 Systolic 算法应该具备以下性质: 算法可以由种类不多的微处理细胞来实现; 算法的数据流和控制流应该单纯而规则; 算法便于采用流水线或阵列处理方式<sup>[1~6]</sup>.

本文给出一种 Systolic 结构的超大规模 ASIC 逻辑设计, 实现无失真并行数据压缩的细胞自动机<sup>[7~10]</sup>, 可对数字和任何符号序列实时地进行无失真压缩. 仿真结果表明了该 ASIC 逻辑设计的正确性和有效性.

文[1, 2]提出基于遗传算法的元胞级并行无失真数据压缩方法, 其中, 一条染色体由两行代码组成, 第一行表示压缩规则 (染色体), 即由局部函数  $f$  的像组成的代码, 第二行表示压缩时的计数规则 (染色体). 对于一阶细胞自动机有:

\* 001 \* 101 \* \* 001 \* 101 \* \* \* \* \* 01 \* (染色体)

0 0 0 0 0 0 1 1 0 0 0 0 0 0 1 1 0 0 0 0 0 0 1 1 0 (染色体)

与染色体 对应的压缩规则集为:

000 \* , 001 0, 00 \* , 0, 010 1, 011 \* , 01 \* , 1, 0 \* 0  
0, 0 \* 1 1, 0 \* \* , \* , 100 \* , 101 0, 10 \* , 0, 110 1,  
111 \* , 11 \* , 1, 1 \* 0 0, 1 \* 1 1, 1 \* \* , \* , \* 00 \* , \*  
01 \* , \* 0 \* , \* , \* 10 \* , \* 11 \* , \* 1 \* , \* , \* \* 0  
0, \* \* 1 1, \* \* \* \*

其中每条规则表示待压缩数据根据其自身及左右细胞的状态决定其下一时刻的状态. 对每位数据可以并行操作.

对于任何有限符号集上的符号序列, 得到如下的文本压缩规则集:

令  $a, b \in S, a \neq b, * \in \{a, b\}$  则

$aaa \rightarrow *, aba \rightarrow a, aa * \rightarrow a, aba \rightarrow b, abb \rightarrow *, ab * \rightarrow b,$   
 $a * a \rightarrow a, a * b \rightarrow b, a * * \rightarrow *, * aa \rightarrow *, * ab \rightarrow *, * a *$   
 $\rightarrow *, * * a \rightarrow a, * * * \rightarrow *$

### 2 细胞结构和逻辑电路

实现无失真并行数据压缩细胞自动机的硬件框图如图 1 所示. 串-并转换寄存器按处理单元的多少, 依次将数据并行送入细胞阵列. 每个细胞中不需要任何计数器、堆栈寄存器,

以及与被压缩数据长度  $n$  有关的逻辑部件. 细胞内与细胞之间的信息线、控制线的数目均与  $n$  无关. 各细胞的结构都一样. 被压缩后的数据由并-串转换寄存器送至译码器, 得到由原符号集中的符号所表示的数据.

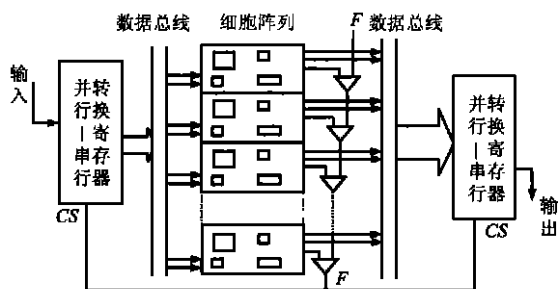


图1 数据压缩细胞自动机的硬件实现框图

细胞结构及组织框图如图2所示. 可编程逻辑阵列 PLA 用于存储细胞自动机中的细胞状态更新规则集, 即文本压缩规则集. 各细胞按外部同步单相时钟的时序, 完成以下工作:

(1) 在时钟上升沿, 输入符号或  $N$  态寄存器的状态进入当前状态寄存器  $C$ ; (2) 在时钟的, PLA 收到三组  $C$  态寄存器的信息, 并产生输出; (3) PLA 的输出结果, 在时钟下降沿送入  $N$  态寄存器; (4) 在时钟期间,  $N$  态和  $C$  态寄存器的状态送入比较器比较. 若比较器结果不相等 (输出为 1), 在上升沿将  $N$  态寄存器的信息复写入  $C$  态寄存器. 若细胞阵列中所有的比较器的输出均为 0, 则表示压缩完毕.

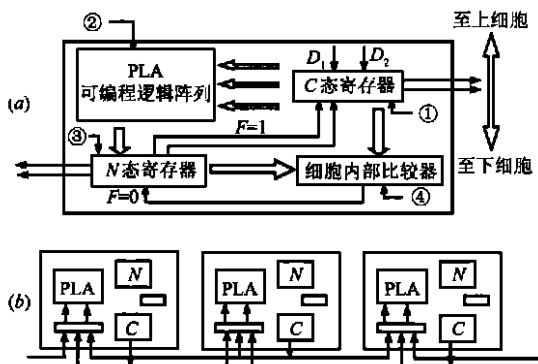


图3 细胞的结构与组织框图. (a) 一个细胞的内部结构; (b) 相邻细胞的互联

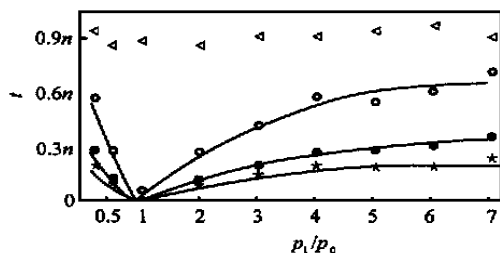


图4 对不同输入数据进行压缩时处理时间的比较 ( , , 和 分别表示一阶、二阶、三阶细胞自动机和经典游程码算法的结果) (其中  $p_0$ ,  $p_1$  分别表示符号为 0 和 1 的概率)

### 3 模拟验证及结论

设计的 ASIC 逻辑电路的计算机模拟验证结果, 如图 4 所示. 正确地实现了无失真并行数据压缩的细胞自动机, 与其他传统的串行或小规模并行无失真数据压缩的硬件相比, 此电路的特点主要是: 完全符合 Systolic 结构的要求; 具有可扩展性, 可对数字或任何符号序列实时地进行无失真并行压缩.

上述的数据压缩一维细胞自动机 ASIC 脉动阵列可以推广到二维和更高维的细胞自动机以及压缩任何有限字符集中的符号组成的数据. 将给定的冗余字符当作规则集中的  $*$ , 且将规则  $aaa \rightarrow *$  和  $abb \rightarrow *$  分别改为  $aaa \rightarrow a$  和  $aaa \rightarrow b$ , 其中  $a, b$  为任何非冗余符号.

### 参考文献:

- [1] C. Shannon. Bell Sys. Tech. J., 1948, 27(3): 379 - 465.
- [2] 吴乐南. 数据压缩的原理与应用 [M]. 北京: 电子工业出版社, 1995.
- [3] 高文. 多媒体数据压缩技术 [M]. 北京: 电子工业出版社, 1994.
- [4] 沈绪榜. 超大规模集成电路设计 [M]. 北京: 科学出版社, 1991.
- [5] B. Randell. VLSI Architecture [M]. Prentice-Hall International, 1983: 348 - 380.
- [6] H. Kung. Why Systolic Architecture? [M]. Computer, 1982: 37 - 46.
- [7] 帅典勋, 顾静. 更快的高阶细胞自动机超并行数据压缩方法 [J]. 计算机学报, 2000, 23(1): 39 - 46.
- [8] 帅典勋, 顾静. 计算机学报, 1999, 22(8): 797 - 803.
- [9] 顾静, 帅典勋. The faster high-order cellular automaton for hyper-parallel undistorted data compression [J]. Journal of Computer Sci. and Tech., Accepted, to appear in 2000, 15(2).
- [10] 顾静, 帅典勋. Journal of Computer Sci. and Tech., 1999, 14(6): 572 - 579.

### 作者简介:



顾 静 1977 年出生, 1998 年于华东理工大学计算机科学与技术系本科毕业, 现为硕士研究生. 主要从事人工智能、知识工程、人工胚胎等理论及应用研究.



帅典勋 1941 年出生, 华东理工大学计算机科学与技术系教授, 博士生导师. 1962 年毕业于华中理工大学电机工程系, 后获清华大学计算机专业工学博士学位. 主要从事分布并行计算、人工智能、智能控制、知识工程、计算机系统组织与结构、人工生命等理论及应用研究, 在国内外发表论文 100 多篇.