

一种混合前缀编码的测试数据压缩方法

谈恩民, 李 贞

(桂林电子科技大学 电子工程与自动化学院, 广西 桂林 541004)

摘 要: SOC 测试中的测试数据具有数据量大、游程长等特征, 对此提出一种新的测试数据压缩方法——混合前缀游程编码。通过改进传统双游程编码思想, 在每一个码组内添加带有特殊前缀的新编码对测试集中的任意长度的 0、1 游程同时进行从变长到变长编码, 达到仅折损一位从而缩小整体编码长度的目的, 同时给出了基于有限状态机的解码器方案。理论分析和 ISCAS89 电路的实际测试结果验证了该编码方法有着良好的测试压缩效果。

关键词: 双游程编码; 测试数据压缩; 解码

中图分类号: TN4

文献标识码: A

文章编号: 1000—7180(2016)09—0089—04

A Hybrid Prefix Encoding Method of Test Data Compression

TAN En-min, LI Zhen

(College of Electronic Engineering and Automation, Guilin University of Electronic Technology, Guilin 541004, China)

Abstract: As the test data in the SOC test has the characteristics of large amount of data and its length, put forward a new testmethod for data compression - hybridprefix-run-length encoding. By improving the thought of raditional double run-length coding thought, in each new code added in the group with a special prefix codes to test set of arbitrary length 0 and 1, run simultaneously from longer to variable length coding. By only wreck a code, so as to reduce overall length coding. At the same time decoder scheme based on the finite state machine is presented. The theory analysis and ISCAS89 actual test results show that the coding method has good test compression effect.

Key words: dual run length; test data compression; decode

1 引言

测试数据规模过大是当前集成电路所面临的众多挑战之一。日益增大的测试集规模会给整个测试仪器的带宽和容量等带来巨大的影响, 更直接导致了测试时间过长、测试功耗过大等问题, 而集成电路的测试时间直接影响测试费用, 所以使用高效的方法直接压缩测试集能较好地解决上述问题。

测试集的压缩包括测试激励和测试响应的压缩。测试压缩贯穿整个电路测试过程, 从测试向量被产生之后开始, 测试压缩方案即开始实施: 测试激励的压缩主要是通过适当的压缩策略对测试激励数据本身进行直接编码压缩或者通过测试结构优化和其他混合方案间接使得测试激励的规模变小; 而测试

响应的压缩跟前者类似, 均是通过一定的压缩方案进行对数据集的规模进行压缩, 主要包括空间和时间两种压缩方案^[1]。

本文主要探讨的是对测试激励数据集进行编码压缩的具体策略。编码压缩策略指的是通过压缩程序对给定的测试集进行编码^[2-4], 以 FDR 码^[5]等为代表的游程编码是编码压缩中的一大类, 均是对连续的 0 或 1 数据的长度, 也就是游程的长度进行直接无损编码。但解码器中必须添加循环扫描寄存器来解压差分向量, 硬件成本显然较其他方案更高。

在测试过程中往往会发现, 0、1 两种游程出现的频率同样重要, 为使得编码压缩率更好, 传统双游程编码^[6]采用了同时对 0 和 1 游程进行编码的方案, 取得了不错的压缩效果; 而本文提出的混合前缀编码同

样是基于传统双游程编码^[6]的改进,主要思想是通过仅仅折损一个码字,对整体游程编码每个码组的长度都进行扩展,达到较经典方案更好的压缩效果,而且对双游程同时编码使得该方案具有很好的适用性,能避免测试集对单一游程的概率要求所需要进行的优化差分过程,减少硬件负担,降低测试成本.

2 混合前缀编码原理

游程编码,是指在由任意个 0 和 1 组成的测试集中,对连续的相同的 0 或 1 的个数进行统计、编码,力求将无规律的、重复的测试数据压缩成无损的、有规律的较短码字.

混合前缀的基本思想是改进经典的双游程编码方案,将游程编码中后缀全为 0 的码字,统一变成下一组编码的前缀,后缀则保持为两位,达到与下一组码字长度相同的效果. 编码规律有以下几点:

(1)对测试集内连续的 0 和 1 同时编码,直到游程遇到相异值结束,例如 00001 视为一个长度为 4 的 0 游程,编码码字为 010000;

(2)除第一组外的每组编码前四个特殊前缀皆为前一组的后缀全为 0 的码字,而该码字已从前一组剔除,该四个码字的后缀则规定为 00、01、10、11;余下前缀与后缀等长,其长度为组数 K 加 1,码字长度为 $2K+2$;

(3)除第一组外每组包含的游程个数为 $2^{(K+1)}+3$.

如表 1 所示,混合前缀编码(下文统称为混编)的编码表可以看出,在游程为 4 的时候,混编码字折损两位,但从第 2 组开始,每组都增加了 4 个有着特殊前缀的编码,同时剔除一个后缀全为零的码字,作为下一组的特殊前缀. 游程长度从 5 到 14,混编依然保持码字长度为 5. 随 K 值的增大,混编进入 $K+1$ 组的游程长度较经典游程编码就会越来越大,达到了仅仅折损两位码字而整体上增大了码组长度的目的.

3 解码器

混合前缀编码的重点是同组内同时采用不同长度的前缀编码,所以基于有限状态机的解码器需要较经典编码例如 FDR 码^[5]更多的计数器和稍复杂的解码步骤. 表 1 中,设非特殊前缀的编码长度为 $k+1$,则该编码所在组数为 k ;可推得第 k 组的第一位编码的游程长度为 $2(k+1)+3k-10$,例如 $k=2$ 时,第二组的第一位编码游程长度为 $8+6-10=4$,符合编码表规则.

表 1 混合前缀编码

游程	K	前缀码	后缀码	0 游程 编码	1 游程 编码
1	1	01(10)	01	0101	1001
2			10	0110	1010
3			11	0111	0111
4	2	0100(1000)	00	010000	100000
5			01	010001	100001
6			10	010010	100010
7			11	010011	100011
8	2	001(110)	001	001001	110001
9			010	001100	110010
10			011	001011	110011
11	3	(110000)	100	001100	110100
12			101	001101	110101
13			110	001110	110110
14			111	001111	110111
15	3	001000	00	00100000	11000000
16			01	00100001	11000001
17			10	00100010	11000010
18			11	00100011	11000011
19	...	0001(1110)	0001	00010001	11100001
...			...	...	...
...			...	...	...

故设 $m=\log_2(2(k+1)+3k)$,解码器见图 1,由一个有限状态机 FSM、两个 m 位计数器、两个 $3\log_2(m)$ 位计数器和一个异或门组成. 压缩完成的数据如果需要解码,当使能端 en 输出为 1 时,待解压的数据便从数据输入端 b_in 移入此时准备就绪的 FSM 中. 具体解码器信号对应功能如表 2 所示.

表 2 解码器信号功能

信号	功能
b_in	数据输入端,待解码的数据通过此端移入 FSM 中
en	使能端,en 为 1 表示解码器准备就绪,等待接收数据
c_in1/2	FSM 向 m 位计数器输入编码数据的通道 0010001111000011
shift1/2	控制编码数据移入 m 位计数器
dec1/2/3/4	控制计数器作减 1 操作
inc1/2/3/4	控制计数器作加 1 操作
rs1/2/3/4	计数器的复位标志
flag	标志位,为 1 表示待解压的为 1 游程,反之为 0 游程
v	输出控制信号,为 1 表示输出数据有效
data	原始输出数据
out	完整输出数据

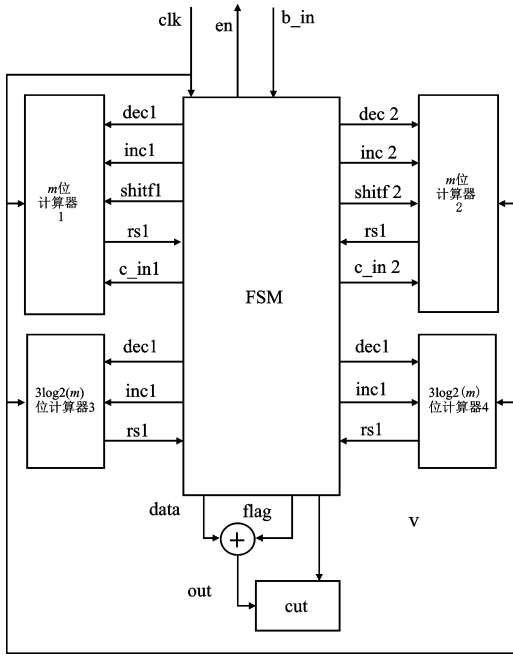


图1 解码器模型

该解码器主要工作原理如下所示。

(1)当解码开始时,FSM发出使能信号en为1,进而读取数据。由于混编的特殊性,首先统一按照翻转结束的规律读取前缀,前缀数据由b_in端移入FSM,读取第一位数据时将flag赋予同样的值,代表了该游程的类型;同时shiftf1为1,数据移入m位计数器1中;inc3和inc4则也为高电平,即 $3\log_2(m)$ 位计数器3、4t同时进行加1计数,当b_in为0时,此过程持续至b_in为1时停止,反之则b_in为0时,前缀数据输入结束,在此过程中, $3\log_2(m)$ 位计数器3、4记录了前缀数据的位数,设前缀k位。

(2) $3\log_2(m)$ 位计数器3作减1操作,同时从b_in读入后缀数据,shiftf2为1,m位计数器2移入后缀信息。当 $3\log_2(m)$ 位计数器3为0时,后缀移入结束,此时若m位计数器2中为0,则代表该后缀全为0,即表明该前缀和后缀组成特殊前缀,由编码表规则可知,组数为k,此时 $3\log_2(m)$ 位计数器4为k。步骤跳转至4);若后缀不全为0,当k=2时,FSM向data输出1位0,为1,结束后v恢复低电平,跳转至(6),否则跳转至(3)。

(3)此时前缀编码为k位,代表组数为k-1;则m位计数器1做减1操作,当rs1为0时,FSM先向计数器1移入1位1,然后计数器4开始做减1操作,进行1次之后,计数器3开始作加1操作,同时FSM向计数器1移入0,直至计数器4为0时结束,此时计数器3为k-1;然后计数器3进行减1操

作,计数器1、计数器4同时作加1操作至计数器3为0时结束;重复计数器1加1操作,计数器3、4相互加减计数的步骤,直至计数器1为 $2k+3k-3$ 时为止;

(4)此时前缀编码所在组为k,则m位计数器1作减1操作,rs1为0后,开始重复(3)的操作,直至计数器3、4相互加减计数三次,计数器1内为 $2(k+1)+3k$ 为止;en为1,shiftf2为1,直至FSM往计数器2内读入两位结束,然后计数器1作减1操作,同时v为1,表示输出有效,FSM开始对data输出0,直到m位计数器1为10时,前缀解码结束。然后跳转至(6)。

(5)计数器1作减1操作,同时v为1,表示输出有效,FSM开始对data输出0,直到m位计数器1为10时,前缀解码结束。FSM向data输出4位0,输入完毕后v为0,表示输出无效。

(6)后缀数据已经移入了计数器2中,计数器2先作加1操作,然后计数器2开始作减1操作,同时v为1,表示输出有效,FSM开始对data输出0,直至计数器2为0,FSM再向data移入1位1,解码完毕。

4 实验结果

首先推导混编压缩效率的公式。根据编码表可知,混合前缀编码的组数为k时($k \geq 2$),则该组第一位码字的游程长度为 $2(k+1)+3k-10$,而末尾码字的游程长度则为 $2(k+2)+3k-8$,故对于k组,其码字长度均为 $2k+2$,所代表的游程范围为 $[2(k+1)+3k-10, 2(k+2)+3k-8]$ 。由于是双游程编码,故设在整个测试集中0出现的概率为 $p(0 \leq p \leq 1)$,1出现的概率为 $1-p$ 。由编码规则可得,当0游程长度为m时,其出现的概率为 $pm(1-p)$,同理可得同样长的1游程出现概率为 $p(1-p)m$,故任意游程其所在组k出现的概率P为

$$p(m, k) = \sum_{m=2^{k+1}+3k-10}^{2^{k+1}+3k-8} [p^m(1-p) + p(1-p)^m].$$

整个混合编码的平均游程长度为

$$\bar{C} = \sum_{i=1}^3 3[p^i(1-p) + (1-p)^i p] + \sum_{k=2}^{\infty} (2k+2)p(m, k).$$

测试数据未编码前的平均游程长度为

$$\lambda = 1 + \sum_{i=1}^{\infty} ipi(1-p) + 1 + \sum_{i=1}^{\infty} ip(1-p)i = \frac{1}{p(1-p)}.$$

故可得混编的压缩效率为

$$\beta = \frac{\lambda}{C} = \frac{1}{p(1-p)[\sum_{i=1}^{\infty}(2k+2)]p(m,k) + \sum_{i=2}^3 3[p^i(1-p) + (1-p)^i p]}$$

为了验证该编码的有效性和优越性,本文使用 MATLAB7.0 对其压缩效率当经典双游程编码以及 EFDR 码进行对比,使用 1 000 项近似为无限逼近,如图 2、图 3 所示。

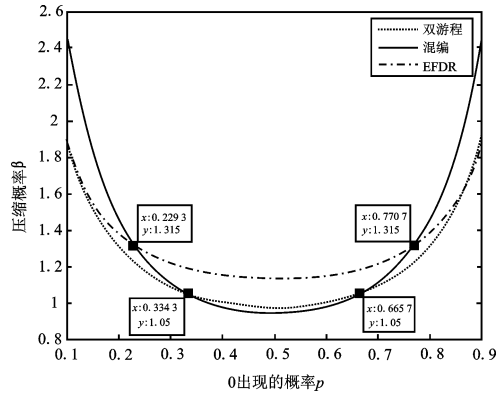


图 2 混合前缀编码与经典双游程以及 EFDR 比较

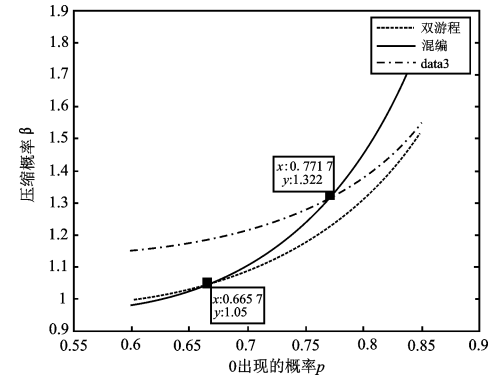


图 3 (0.6≤p≤0.85)时三者压缩率比较

由图 2 所示,三种双游程编码均在 $p=0.5$ 时压缩率最为低下,随着单游程概率的升高,压缩比也随之增大.图 3 明显地可以看出在 $p>0.66$ 之后,混编的压缩效率要高于经典双游程编码而低于 EFDR,但在 $p>0.7$ 之后,混编的效率不光超过前两者,爬升的速率也要高于两者,这源于混编每一组都较其他编码方式多了 3 位有效编码,随着单一游程的概率增大,折损一个码字换来的收益将越来越高,最终获得很高的压缩率。

最后采用美国 Duke 大学提供的 mintest AT-PG 测试向量生成工具产生的 TSCAS'89 国际时序基准电路测试集,经过无关位填充后作为待编码数

据,选取其中几个有代表性的电路,采用 Visual C++ 2012 进行软件编程,进行测试集的编码工作.设 TD 为原始的测试集,而 TE 则为压缩后的编码数据,压缩率的计算方法为

$$\text{压缩率} = \frac{T_D - T_E}{T_D} \times 100\%.$$

实验结果如表 3 所示。

表 3 混合前缀编码与其他编码方式压缩比较 %

电路	Golomb 码 ^[2] 压缩率	FDR 码 ^[5] 压缩率	交替游程编 码 ^[1] 压缩率	混合前缀编码 压缩率
S5378	37.11	47.98	49.74	51.29
S9234	45.25	43.61	47.69	48.03
S13207	79.74	81.30	82.95	82.78
S15850	62.82	66.21	68.04	68.46
S38417	28.37	43.37	62.48	62.26
S38584	57.17	60.93	63.94	64.31
AVG	51.74	57.23	62.47	62.86

由表 3 可知,一共 6 种基准电路.在平均编码压缩率上,本方法比 Golomb 编码^[2]高出 11.12%,比 FDR^[5]码高出 5.63%,虽然与交替游程编码^[1]差距为 0.39%,但是测试集中 0,1 的出现概率显然没有达到本方案最优的概率区间中,除了 s3841 电路本方法略低于交替游程编码^[1]之外,其他所有电路,本方法都优于前人方法,具有优越的稳定性和适应性.综合考虑在大游程上的理想表现以及实验结果,混合前缀编码是一种有前景有实用价值的数据压缩方法。

5 结束语

为了解决日益增大的数字电路测试的成本、存储等问题,本文提出一种基于传统双游程编码的新方法,利用上一组后缀为 0 的码字为下一组增加 4 位同长度码字,有效地扩大了每组编码的游程范围,通过与同类方法的理论和实验对比,本文方法表现出了相当大的优越性和实用性。

参考文献:

[1] 程一飞,詹文法.一种游程交替编码的测试数据压缩方法[J]. 计算机科学,2014,41(11):22-55.
[2] Chandra A,Chakrabarty K. System-on-a-chip test-data compression and decompression architectures based on-Golomb codes[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2001, 20(3):355-368.

验证,经过不断的添加测试用例和测试序列,修改约束以及随机种子,最终得到了 100%的功能覆盖率,并且记分板中的数据匹配性为 100%,没有逻辑功能错误,预期的功能验证目标全部达到. 具体的测试功能点和覆盖率统计如表 1 所示. 表中所列的是对标准帧和扩展帧共同统计的结果.

表 1 工作模式及功能覆盖率

工作模式	功能点	覆盖率/%
初始化模式	主要寄存器的读操作	100
	主要寄存器的写操作	100
正常工作模式	发送功能	100
	接收功能	100
	应答功能	100
	发送中断	100
	接收中断	100
	仲裁机制	100
	接收过滤功能	100
监听模式	监听功能	100
自检模式	自检功能	100
	位错误	100
	格式错误	100
	填充错误	100
	应答错误	100
	CRC 错误	100

6 结束语

UVM 和传统的验证平台相比,前期的验证平台的搭建会比较复杂,但是进入回归测试之后以及在可重用性方面,优势很明显. 本文所讨论的 CAN 模块的自验证方法,由于没有利用 CAN 总线模型,验证平台的搭建比较简单,具有很好的可重用性. 同时,这种自验证的方法在其他总线上也有很好的参考和利用价值,能够大大的提高验证的效率.

参考文献:

[1] 张强. UVM 实战[M]. 北京:机械工业出版社,2014.
[2] 赵强,田泽,楼小强,等. CAN 总线控制器 IP 核设计实现[J]. 计算机技术与发展,2013,23(8): 230-233.
[3] BOSCH CAN. Specification V2. 0[R]. Robert Bosch GmbH,1991.
[4] 吕易俗,田文杰,李翔宇,等. 基于 Verilog 的 CAN 总线协议验证模型库的建立[J]. 微电子学与计算机, 2015,32(3) : 122-126.
[5] 崔瑾. SoC 芯片中 CAN 总线控制器的研究和验证 [D]. 西安:西安电子科技大学,2013.

作者简介:

熊 涛 男,(1989-),硕士研究生. 研究方向为数字集成电路物理设计与验证. E-mail:xiongtao@ime. ac. cn.
蒋见花 女,(1976 -),硕士生导师. 研究方向为数字集成电路低功耗设计.

(上接第 92 页)

[3] Nourani M, Tehranipour M. RL-huffman encoding for test compression and power reduction in scan application[J]. ACM Trans Des Autom Electron Syst, 2005, 10(1): 91-115.
[4] 韩银和,李晓维. 测试数据压缩和测试耗协同优化技术 [J]. 计算机辅助设计图形学学报, 2005, 17 (6): 1308-1311.
[5] Chandra A,Chakrabarty K. Test data compression and test resource partitioningforsystem-on-a-chip using frequency-directedrun-length (FDR) coeds [J]. IEEETransactionson Computer Aided Design of Inte-

grated CircuitsandSystems,2003,52(8):1076-1088.
[6] 商进,张礼勇. 一种双游程编码的测试压缩方案[J]. 哈尔滨理工大学学报,2010,15(4):19-22.

作者简介:

谈恩民 男,(1966-),教授. 研究方向为计算机辅助测试和集成电路可测试性设计.
李 贞(通讯作者) 男,(1990-),硕士研究生. 研究方向为计算机辅助测试和集成电路可测试性设计.
E-mail:1078300120@qq. com.