

可编程逻辑器件逆向工程研究^{*})

李清宝 张 平 赵荣彩 曾光裕

(信息工程大学信息工程学院 郑州 450002)

摘 要 逆向工程在科技发展和商业用途上都发挥着重要作用。本文针对在电子、电机和信息领域广泛应用的可编程逻辑器件进行逆向工程研究,首先根据设计中可实现的逻辑结构特征对可编程逻辑器件进行分类,分析论证使用非破坏性方法对可编程逻辑器件进行逆向还原、再造的可能性,然后理论证明可编程逻辑器件逆向工程的可行性,最后给出逆向工程分析中所涉及的关键技术。

关键词 可编程逻辑器件,逆向工程,组合逻辑,时序逻辑,反馈

Research on the Reverse Engineering of Programmable Logic Devices

LI Qing-bao ZHANG Ping ZHAO Rong-cai ZENG Guang-yu

(Information and Engineering University, Zhengzhou 450002, China)

Abstract Reverse engineering technology is very important in the development of science and technology and commercial applications. Techniques of reverse engineering of programmable logic devices (PLD) are discussed which are commonly used in electronic, electrical and information industries. First, programmable logic devices are divided into different types according to the designed logic structure. Second, a methodology is developed to demonstrate the feasibility of non-destructive technique for reverse engineering PLDs. Finally, some key techniques in the reverse analysis are presented.

Keywords Programmable logic device, Reverse engineering, Combination logic, Timing logic, Feedback

1 引言

可编程逻辑器件 PLD(Programmable Logic Device)是以 And-Or 逻辑阵列为基础结构,由用户自行编程实现随机逻辑的各类半定制 ASIC 电路^[1],在电子、信息和通信等领域应用十分广泛,发挥着重要的作用。

绝大多数 PLD 器件都可以加密,内部设有保密位。写入正常功能逻辑后,将保密位加密编程,用户便无法用常规方法获取其内部功能逻辑码点内容。PLD 的逆向工程是指以合法目的为前提,通过某种技术,分析器件的内部逻辑设计,复制出功能等价的器件。PLD 的逆向工程对于维护现有仪器设备,更好地消化国外的先进科学技术具有重要意义。

一般可编程逻辑器件的逆向工程分为破坏性和非破坏性两种方法。破坏性方法通过打开芯片外封装,恢复其保密位的方法读取芯片内部资料,可以复制内容完全相同的芯片。这种方法需要完善的场地和设备,代价很高。非破坏性方法则是利用芯片的构造和编程特性使用还原和解析两种方法来获得芯片的内部逻辑。

还原法主要是针对以电可擦除(electrically erasable)技术制造的可编程逻辑阵列芯片,例如 22v10 等,可以利用其制造的特点,在读取时序中的适当时机,针对某一引脚加入高电压,还原其保密位,让原来有编程保护的 PLD 变成没有保护的状态,然后直接读取芯片内的编程码点内容。还原法的缺

点是只能适用于部分 Double E 类型的芯片,而且必须针对特定厂商的特定型号,了解其编程或烧录时序,然后编写相应的读取程序,不具有通用性。另外,可能在施加高电压时损坏芯片,具有一定的危险性。

本文主要讨论另一种 PLD 逆向工程方法:解析法。解析法是一种安全、通用的方法,适用于各种工艺的 PLD 器件,对芯片没有任何损害。解析法通过采集 PLD 工作数据,分析数据以获知其逻辑功能。所谓逻辑功能,实际上就是 PLD 输入端和输出端响应信号之间的对应关系。如果能够测出 PLD 的所有输入与输出之间的对应关系,就等于知道了这片 PLD 的逻辑功能。解析法工作的基本过程为:分析编程类型-采集数据-逻辑综合,其关键在分析编程类型和采集数据。本文所讨论的逆向工程方法适用于下面四种目前市场上广泛应用的主流元件:PAL(programmable array logic),EPLD(erasable programmable logic device),GAL(generic array logic)和 PEEL(programmable electrically erasable logic)。

2 PLD 逆向工程理论

2.1 PLD 基本结构

PLD 器件是一种由用户配置的半定制逻辑器件,基本结构如图 1 所示。它对输入数据执行一定的操作,以产生所需的输出数据。把一系列的电压脉冲加到 PLD 器件上后,便能完成对它的配置。

^{*})国家 863 高科技发展计划资助项目(2006AA01Z404)。李清宝 博士,教授,研究方向为数字电路设计与分析、计算机系统结构;张 平 博士,副教授,研究方向为并行计算;赵荣彩 教授,博导,研究方向为并行计算、软件逆向工程和网络安全;曾光裕 硕士,副教授,研究方向为数字电路设计与分析。

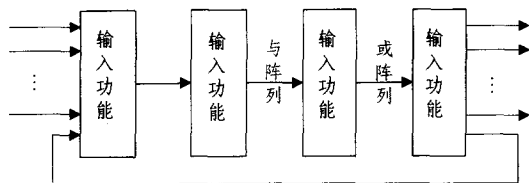


图1 PLD器件的基本结构

2.2 PLD逻辑结构分类

从PLD设计的角度出发,通常根据是否需要寄存器将PLD的逻辑结构分为组合型和寄存器型两大类^[2]。组合型逻辑设计的应用主要分为五个方面:①译码器和编码器;②多路复用器;③比较器;④加法和算术逻辑;⑤锁存器。在组合逻辑电路中,其输出完全与任何时钟无关,而寄存器型逻辑设计为时序电路,主要用于数据处理与控制。寄存器型设计可分为两类:同步时序和异步时序。对于同步设计,所有寄存器的输入时钟都连到一个公共的时钟上;对于异步设计,寄存器的输入时钟可以不连在一起,时钟可以被选通或由其他寄存器驱动^[3]。

下面我们从PLD逆向工程的角度出发,将组合型和寄存器性两类PLD逻辑结构进一步细化,针对不同逻辑结构特点,分析论证采用解析法逆向分析PLD器件的可行性以及逆向工程过程中必须解决的关键问题。

2.3 各类逻辑结构逆向分析讨论

2.3.1 基本组合逻辑型

基本组合型PLD的内部电路结构如图2所示,其特征可以用3元组 $P=\langle X, Y, F_y \rangle$,其中:

(1) 输入端为一个有限的输入向量(一般PLD的每个and项设有一定的限制),输入集合定义为

$$X = \{x_1, x_2, \dots, x_m\}, x_i \in L, i=1, 2, \dots, m, L = \{0, 1\} \quad (1)$$

(2) 输出端为一个有限的输出向量,其集合定义为:

$$Y = \{y_1, y_2, \dots, y_n\}, y_i \in L_z, i=1, 2, \dots, n, L_z = \{0, 1, z\} \quad (2)$$

其中 z 表示输出为高阻状态。

(3) 一个输入输出映射函数:

$$F_y: X \rightarrow Y, X \in L^m, Y \in L_z^n, \quad (3)$$

因此,基本组合逻辑型PLD逆向工程必须解决的问题包括:

(1) 判别芯片输出端集合和输入端集合,这是对各种结构PLD器件逆向分析首先必须解决的问题,我们将在下节中集中讨论;

(2) 判别输出信号的电平: Lo, Hi 或高阻,这个问题由接口硬件电路来实现;

(3) 求解输入输出映射函数。

在实际设计应用的芯片中,任意的 $X \in L^m, Y$ 都有对应的函数值,即 F_y 也是可确定的。因此,只要顺序输入所有 X 值,并相应地读取 Y 的值,即可完整求得 $F_y: X \rightarrow Y$ 的映射函数。

由(3)式可知 $Y \in L_z^n$ 的值由 $X \in L^m$ 决定, $X \in L^m$ 有 2^m 个取值,因此依次施加 2^m 个输入值,采集对应的 Y 输出,进一步通过逻辑综合,便可完成映射函数的逆向求解工作。

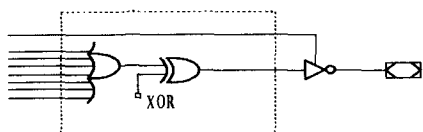


图2 基本组合逻辑型PLD内部结构图

2.3.2 带反馈组合逻辑型

带反馈组合逻辑型PLD在基本组合逻辑型的基础上,某些输出引脚同时可作为输入,影响其他输出或其自身的输出值,其内部电路结构如图3所示。特征可以用4元组 P 来描述: $P=\langle X, Y, B, F_y \rangle$,其中:

(1) 一个有限的输入变量,其集合定义形式为

$$X = \{x_1, x_2, \dots, x_m\}, x_i \in L, i=1, 2, \dots, m, L = \{0, 1\} \quad (4)$$

(2) 一个有限的输出向量,其集合的定义形式为

$$Y = \{y_1, y_2, \dots, y_n\}, y_i \in L_z, i=1, 2, \dots, n, L_z = \{0, 1, z\} \quad (5)$$

(3) 一个有限的反馈变量,其集合的定义形式为

$$B = \{b_1, b_2, \dots, b_k\}, b_i \in L, i=1, 2, \dots, k \quad (6)$$

(4) 一个输入输出映射函数:

$$F_y: X * B \rightarrow Y, X \in L^m, B \in L^k, Y \in L_z^n \quad (7)$$

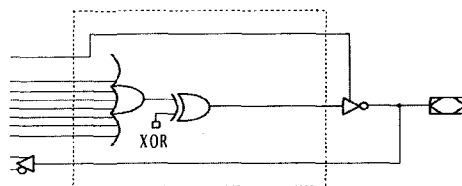


图3 带反馈组合逻辑型PLD内部结构图

带反馈组合逻辑型PLD的逆向工程在基本组合型处理过程的基础上需要进一步考虑下面问题:

(1) 反馈变量值的输入利用硬件接口电路输出,将任意的 $B \in L^k$ 加载到反馈引脚上。

(2) 采集数据,求解输出映射函数。

扫描并完整地给出 $(X, B) \in L^m * L^k$ 的 2^{m+k} 个取值,采集相对应的 Y 输出值,建立输入输出真值表,通过逻辑综合便可完成此函数的逆向求解工作。

2.3.3 同步时序寄存器型

同步时序PLD以全局统一的时钟驱动所有寄存器输出,是数据处理和控制中最常用的一种编程形式,其特征可以用5元组 P 来描述: $P=\langle X, Y, S, F_y, F_s \rangle$,其中:

(1) 一个有限的输入变量,其集合定义形式为

$$X = \{x_1, x_2, \dots, x_m\}, x_i \in L, i=1, 2, \dots, m, L = \{0, 1\} \quad (8)$$

(2) 一个有限的输出向量,其集合的定义形式为

$$Y = \{y_1, y_2, \dots, y_n\}, y_i \in L_z, i=1, 2, \dots, n, L_z = \{0, 1, z\} \quad (9)$$

(3) 一个有限的、有效状态集合如下:

$$S = \{s_1, s_2, \dots, s_r\}, r \leq 2^k \quad (10)$$

k 为寄存器输出总数。

(4) 一个输出映射函数:

$$F_y: X * S \rightarrow Y \quad (11)$$

(5) 一个状态转换映射函数:

$$F_s: X * S \rightarrow S \quad (12)$$

由于输出状态是由每个寄存器的状态组成,而每个寄存器又都可以输出,所以式(11)和式(12)中的两个函数仅用其中的任一函数取代即可。在逆向分析过程中,取 $F_y: X * S \rightarrow Y$ 。

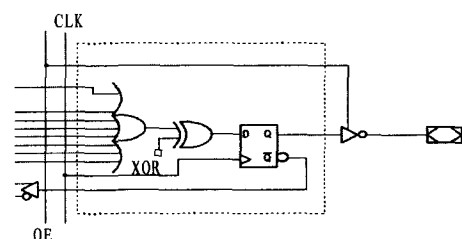


图4 同步时序寄存器型PLD内部结构图

同步时序寄存器型 PLD 逆向工程的关键在于同步寄存器输出端的资料要详尽完整,即要采集到有效状态下的所有数据,使得所有状态达到饱和。

当被分析元件处于某确定状态 $s \in S$ 时,可取任意的 $x \in L^m$,求得 $F_y(x, s)$ 。但数据采集电路只能任意提供 $x \in L^m$,不能任意提供 $s \in S$ 。因此,对于同步时序寄存器型 PLD 的逆向工程,必须解决如何获得任意有效状态 s 的问题。

一个实际应用的同步时序型 PLD,有下面定理:

定理 1 同步时序的 PLD 满足:对任意的 $q \in Q$ 和任意的 $s \in S$,可以通过操作将某一状态 q 驱动至另一状态 s ,其中 $Q = \{q | q \text{ 为芯片上电后的可能的初始状态}\}$ 。

定理 1 的正确性是非常直接的,因为实际应用的 PLD 上电后必然达到某一确定的初始状态。如果从初始状态 $q \in Q$,无法通过操作达到某一状态 $s \in S$,则 s 在芯片实际工作时也无法达到,即 s 是无效状态。所以对于一个有效状态 s ,可以通过操作由初始状态 q 驱动至 s 。

根据定理 1,可以得出定理 2。

定理 2 同步时序寄存器型 PLD 可经由有限步操作而达到逆向工程的要求,得到完整而且涵盖原始逻辑功能的数据集合。

证明:

- (1) 假设芯片上电后的初始状态为 s_0 ;
- (2) 取 s_0 的未作用输入条件 $X \in L^m$ 采集 $F_y(x, s_0)$,存储结果,利用接口电路设置 s_0 为新的状态。
- (3) 判断 s_0 是否饱和:若饱和,转(4);若不饱和,则返回(2);
- (4) 判断是否所有的 $s \in S$ 饱和:若饱和,求解结束;若不饱和,则继续向下执行;
- (5) 待解析的 PLD 器件重新上电,通过数据采集电路设置 s_0 是否饱和:若饱和,则往下执行;若不饱和,则转(2)继续取值;
- (6) 对不饱和的有效状态 $s \in S$,寻找将待解析 PLD 从状态 s_0 驱动至状态 s 的途径,并依照寻找的结果通过接口电路驱动 PLD 芯片,从 s_0 驱动至 s ,然后返回(2),再次采集。在上述过程中:

a) $X \in L^m$ 是状态 $s \in S$ 的已作用输入条件,指的是 $F_y(x, s)$ 已经由数据采集电路求得,否则称为未作用的输入条件。

b) 如果任意的 $X \in L^m$, X 为 s 的已作用的输入条件,即 s 为饱和状态:否则为不饱和状态, s_0 则为待解析 PLD 芯片的当前状态。

上述证明的关键在第(6)步,由定理 1 可知,该步一定可以在有限步内完成。因此,求解算法一定可以在有限步数内求得全部有效数据。

定理 2 的证明过程实际上也是同步时序型 PLD 数据采集算法的执行过程。在这个过程中,如何记录并计算从 s_0 到任意状态 s 的最短路径,直接影响着数据采集算法的性能。

2.3.4 异步时序寄存器型

异步时序寄存器型 PLD 的寄存器的输入时钟可以不连在一起,时钟可以被选通或由其他寄存器驱动。其内部结构如图 5,特征可以由 5 元组 P 来描述: $P = \langle X, y, SW, S, F_y \rangle$, 其中:

- (1) 一个有限的输入变量,其集合定义形式为

$$X = \{x_1, x_2, \dots, x_m\}, x_i \in L, i = 1, 2, \dots, m \quad (13)$$

- (2) 一个输出变量定义为

$$y, y \in L_z, L_z = \{0, 1, z\} \quad (14)$$

其中 z 表示输出为高阻状态。

- (3) 一个有限的状态变量:

$$SW = \{x_1, x_2, \dots, x_m, y\} \quad (15)$$

- (4) 一个有限的有效状态集,其状态集合定义为

$$S = \{s_1, s_2, \dots, s_r\}, r \leq 2^e \quad (16)$$

e 为内部结构中的输出正反器数目。

- (5) 一个特征函数:

$$F_y: X * S \rightarrow Y \quad (17)$$

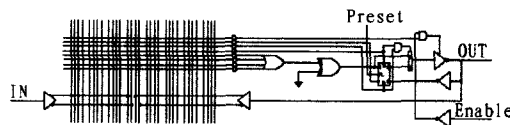


图 5 异步时序寄存器型 PLD 内部结构图

对于任意的 $s_1, s_2 \in S$ 且 $s_1 \neq s_2$ 存在一个输入系列,待解析元件的输入端 s_1 经由内部非同步时序作用后到达输出端,表现为状态 s_2 。这样逆向的可行性证明包含于第 3 类情况。

3 逆向分析的关键技术

从上面 PLD 逆向工程可行性分析过程中,可以得出逆向工程必须解决以下问题:

- (1) 判别引脚类型,确定芯片的编程类型;
- (2) 采集数据,建立真值表;
- (3) 对真值表进行逻辑综合,求得逻辑功能的布尔方程描述。

其中,第(2)点数据采集问题已在第 2 节理论分析过程中讨论。而有关第(3)点逻辑综合的内容,是 IC 设计中的一项重要研究内容,已有许多研究成果及相关工具^[4,5],本文不再赘述。

对 PLD 加密器件引脚类型进行判别时,需要判别芯片引脚为输入、输出还是空闲引脚。输入引脚包括逻辑输入引脚 I、时钟输入引脚 CLK;输出类引脚又分为组合输出引脚 C、寄存器输出引脚 R 等,根据芯片引脚所具有的类型,即可确定芯片的编程类型^[6]。

芯片引脚类型判别过程如下:

(1) PLD 器件引脚类型为下面三种情况之一:确定为输入、确定为输出、可编程为输入或输出。确定为输入或输出的引脚可以通过芯片型号直接去确定,而可编程的引脚需要确定在本设计中的实现方式,判别芯片引脚编程为输入、输出还是空闲引脚;通过向分析芯片施加伪随机序列码,判别电路监测每一引脚的变化情况,即可自动分离出输入、输出和空闲引脚。

(2) 进行时钟引脚判别。输出引脚判出后,还需继续判别输出是寄存器输出还是组合输出、寄存器输出是单时钟控制还是多时钟控制输出等,从而确定芯片的编程类型。如果所有的输出引脚都是组合,则芯片的编程类型为组合型,否则为寄存器型。判别单时钟时序输出相对简单。对于编程为多时钟时序电路 PLD 器件来说,每一时序输出引脚均独立受控于某一时钟引脚。因此,可通过模拟时钟的独立作用过程判别出时序输出的控制时钟。

(3) 进行器件反馈机制的判别。对于组合输出类(C类)引脚,有两种方法可测试其是否为反馈引脚:一是当被测 C 类引脚输出为逻辑 1 时,从测试电路给它输入逻辑 0。被测

(下转第 235 页)

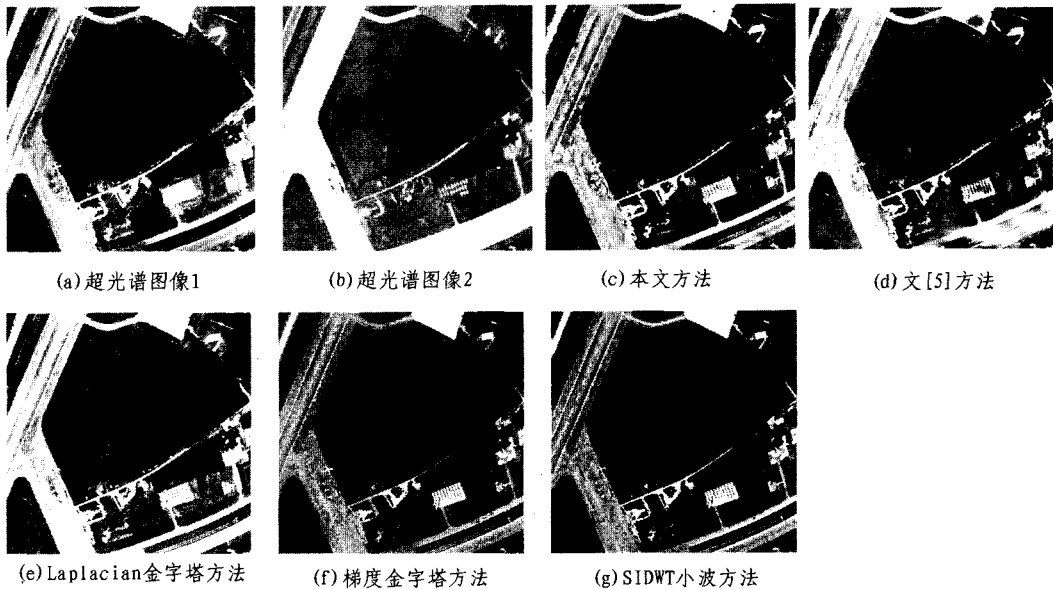


图7 超光谱图像的融合

表3 不同情况下的超光谱图像融合客观性能评价指标

	清晰度	熵值	平均交叉熵	均方根交叉熵
12.4371	7.2923	1.2579	1.2581	
文献[5]方法	11.2790	7.2264	1.8880	2.4872
Laplacian金字塔方法	11.9885	7.2535	1.4429	1.4572
梯度金字塔方法	10.0341	7.1143	1.4306	1.4687
SIDWT小波方法	11.7657	7.2569	1.3930	1.4177

结束语 Contourlet 变换能够解决小波变换所存在的不能处理二维或更高维奇异性的问题,也非常适合于二维图像的处理。本文提出了基于 Contourlet 变换方法进行图像融合的基本架构,并对变换系数的设置与融合过程的处理进行了详细的说明。最后通过仿真实验,并与其它一些基于金字塔变换和基于小波变换的图像融合方法的仿真图像进行了主观比较和客观性能指标比较分析,证明了使用 Contourlet 变换进行图像融合的正确性和优越性。

(上接第 200 页)

芯片的这一外部环境的变化,如果能引起芯片其他输出引脚的输出状态发生变迁,则说明该引脚从逻辑 1 到 0 的变化是一次成功判别,也就是说该引脚是一条反馈引脚。二是不管被测的 C 类引脚输出状态如何,直接给被测引脚强制施加逻辑 0 或 1 的输入,以判定该引脚是否具备输入特性,即是否为反馈引脚。对于寄存器输出类引脚是否反馈的问题,可通过不同的状态输入相同的激励,然后取其结果进行比较。若所有不同状态在相同输入情况下其输出均相同,则没有寄存器反馈输出,否则含有寄存器反馈输出。根据结果可进一步搜索找出具体哪个引脚是寄存器反馈输出。

结束语 随着电子技术的发展,可编程逻辑器件的功能正变得日益强大,其结构复杂度也日益提高,应用涉及的领域更加广泛。如何对涉及 PLD 的电路系统进行维护、测试诊断、系统升级换代等工作,都需要 PLD 逆向工程强有力的支持,PLD 逆向工程的作用也越来越引起我们的重视。本文所讨论的逆向分析技术在实际工作中完成了上百枚加密逻辑器

参考文献

[1] 苗启广.多传感器图像融合方法研究[D].学位论文.西安电子科技大学,2005

[2] Do M N, Vetterli M. The Contourlet Transform: An Efficient Directional Multiresolution Image Representation. IEEE Transactions on Image Processing, 2004;1-16

[3] Do M N, Vetterli M. Contourlets[A]. Academic Press,2002;1-27

[4] 喻汉龙,余胜生,周敬利,等.一种基于改进的 Contourlet 变换的图像压缩算法.计算机工程与应用,2005(14):40-43

[5] Li S T. Multifocus Image Fusion Using Contourlet [C] // ISTM/2005;6th International Symposium on Test and Measurement. Vols1-9, Conference Proceedings;6358-6361

[6] Burt P J, Adelson E H. The Laplacian Pyramid as a Compact Image Code[J]. IEEE Transactions on Communications, 1983, 31(4);532-540

[7] Do M N, Vetterli M. Framing Pyramids[J]. IEEE Trans on Signal Processing, 2003, 51(9);2329-2342

件的解密,为先进电子设备设计技术的消化吸收、系统恢复等工作作出了重要贡献,但是对含有异步复位与置位的特殊器件还有一定的局限性,还有待进一步研究。

参考文献

[1] 边计年,薛宏熙,等.数字系统设计自动化.北京:清华大学出版社,2005

[2] 简弘伦.精通 Verilog HDL:IC设计核心技术实例详解.北京:电子工业出版社,2005

[3] Macrovote A B. Introduction to Logic Design. Second Edition. US:The McGraw-Hill Company. Inc,2005

[4] 翟献军,肖志祥.一种改进的快速逻辑综合算法.计算机工程与应用,2002(1)

[5] 王伟,李清宝.一种快速逻辑综合算法.计算机工程与应用,2000(6)

[6] 张平,李清宝.加密可编程逻辑阵列芯片引脚的判别.电子技术,2002,29(1);5-6