

FPGA 实现滑动平均滤波算法和 LZW 压缩算法

陈世海, 裴东兴, 张琦

(中北大学 仪器科学与动态测试教育部重点实验室, 山西 太原 030051)

摘要: 针对数据采集系统高速长时间的采样和后端数据传输及存储能力有限的问题, 提出基于 FPGA 的数据压缩解决方案。同时为平滑采集数据中的量化噪声, 在进行数据压缩前采用滤波的预处理技术。介绍 LZW 算法和滑动滤波算法的基本理论, 详细阐述用单片 FPGA 实现两种算法的方法。最终测试结果表明, 该设计方案能够有效滤除数据中的高频噪声, 同时也可获得较好的压缩比和压缩速度, 具有一定实用价值。

关键词: 数据采集; 滑动平均滤波算法; LZW; FPGA

中图分类号: TN43

文献标识码: A

文章编号: 1674-6236(2010)02-0067-03

Implementation of moving-average filter algorithm and LZW algorithm based on FPGA

CHEN Shi-hai, PEI Dong-xing, ZHANG Qi

(Key Laboratory of Instrumentation Science & Dynamic Measurement, Ministry of Education, North University of China, Taiyuan 030051)

Abstract: Considering the problems between data acquisition system for a long period of high-speed sampling and the limited data transmission and storage capacity, a solution for data compress based on FPGA is proposed. At the same time, in order to smooth the quantizing noise, before data compression using a filter as the pre-treatment technology. This paper introduces the basic theory of LZW algorithm and moving-average filter algorithm, describes in detail the methods to achieve two algorithms by FPGA. The experimental results show that the design scheme can effectively filter out the high-frequency noise of data, meanwhile, it obtains a good compression ratio and processing speed, and has a certain practical value.

Key words: data acquisition; moving-average filter algorithm; LZW; FPGA

在工业生产、科学研究过程中常常为观察一个信号, 需要数据采集系统对信号进行长时间高速采样, 因此, 庞大的数据量给数据传输、存储和处理带来极大的挑战。通过数据压缩技术可以缓解数据传输和存储系统的压力。数据压缩技术分为有损压缩和无损压缩两大类。无损压缩是利用数据的冗余度压缩, 压缩后的数据进行重构后与原来的数据完全相同; 有损压缩压缩后数据进行重构与原来的数据有所不同。多数的数据采集系统因被测对象的不确定性则需要采用无损数据压缩^[1]。由于 LZW 无损压缩算法具有自适应特性, 在对信号统计特性不明确的情况下, 仍然有较好的压缩效果。结合现场可编程门阵列(FPGA)的高集成度、低功耗、灵活、及能进行并行运算的特性^[2], 该设计采用 FPGA 以硬件的方式实现 LZW 算法, 以提高系统的实时压缩能力。同时, 考虑到 LZW 算法自身的特点, 为提高压缩比和减轻上位机对数据进行软件滤波的运算量, 采用在对数据进行压缩前进行滑动平均数字滤波的预处理方法。

1 滑动平均数字滤波

为消除采样系统中的小幅度高频 A/D 量化噪声, 平滑采

收稿日期: 2009-08-20

稿件编号: 200908040

作者简介: 陈世海(1983—), 男, 安徽旌德人, 硕士研究生。研究方向: 数据压缩。

集数据曲线, 提高后续压缩的效能, 在系统中采用具有低通特性的滑动平均数字滤波器。滑动平均滤波法是把 N 个采集数据看成一个队列, 队列的长度固定为 N, 每进行一次新的采样, 把新的采样结果放入队尾, 而去掉原来队首的一个数据, 把队列中的 N 个数据进行算术平均运算, 其结果就为新的滤波结果^[3]。滤波算法的数学模型描述如下:

$$\bar{X}_n = \frac{1}{N} \sum_{i=0}^{N-1} X_{n-i} \quad (1)$$

式中, \bar{X}_n 为第 n 次采样经滤波后的结果; N 为滤波滑动平均的点数; X_{n-i} 为未经过滤波的第(n-i)次采样值。

具体实现时, 综合考虑采样频率、信号的截止频率以及硬件实现的规模和难易程度, 设定 N 值为 8。模块采用 8 级移位寄存器和 7 个加法器来实现对输入数据的滑动求和。8 级移位寄存器即组成一个数据队列, 其中的数据在每个时钟沿的推动下移动一次, 形成新旧数据的移入移出操作。另外用一个固定数值为 8 的除法器实现对滑动和求平均的运算。

2 LZW 算法及其 FPGA 的实现

2.1 LZW 算法的基本原理

LZW 算法是一种基于字典的压缩算法, 该算法的关键是

在数据的压缩过程中会根据输入的数据动态地建立一个字典。巧妙的是该字典不需要与压缩数据流一道进行传输和存储。在对数据进行解压时也能够通过压缩数据流重新建立一个字典,来完成解压缩^[4]。LZW 算法描述如下^[5]:

```

Initialize Table;
String=Input character;
while((Character==Input character)≠EOF)
{if(Table[String+Character]≠Empty)
String=String+Character;
else
{Output String;
Table[String+Character]=String+Character;
String=Character;
}
}
Output String;
Output End_flag;

```

算法首先进行字典的初始化,然后输入第 1 个字符数据赋给变量 String (S), 数据逐个输入压缩器中并赋给变量 Character(C), S 和 C 生成一个索引与字典中的词条数据进行匹配,如果匹配成功就将 S 和 C 的编码值赋给 S,继续进行下一轮匹配。如果匹配失败,则将 S 输出,将 S 和 C 的编码值存入字典,并且将 C 的值赋给 S,这一过程一直进行直到结束。最后,输出 S,输出结束标志。由算法描述可见,LZW 算法并不复杂,有利于硬件实现且能够实现对数据进行实时的压缩。仔细分析该算法特性可知,对于数据流中连续重复出现的数据,LZW 压缩算法具有很高的压缩比。在前端数据滤波的过程,其不仅能够起到平滑数据曲线的作用,同时也能产生一些连续重复出现的数据,而这对于算法实现高压缩比是非常有利的。

2.2 LZW 算法的 FPGA 实现

硬件模块框图如图 1 所示。为获得较高的实时压缩速度,采用 FPGA 内部的 RAM 块来构造 LZW 字典。字典中的词条由编码值(Code)、字符串(String)及字符(Character) 3 部分构成,其中 Code(CO)和 String(S)所占空间分别为 9 bit,而 Character(C)为 8 bit。每个词条所占用的存储空间为 9+9+8=26 bit。可以构造位宽为 26 位的存储器作为字典。字典的查找技术是 LZW 算法中的一个关键部分,关系到每进行一次数据匹配过程的效率。为提高字典的地址计算速度,采用哈希函数建立字典索引地址与输入数据之间的关系。具体的哈希函数计数公式如下^[6]:

$$\text{Index} = (\text{C} \ll 1) \text{ XOR } (\text{S}) \quad (2)$$

$$\text{Index} = \text{Index} - \text{Offset} \quad (3)$$

$$\text{Index} = \text{Index} - \text{Offset} + \text{Tab_size} \quad (4)$$

$$\text{Offset} = \text{Tab_size} - \text{Index} \quad (5)$$

Index 为字典索引地址;Tab_size 为哈希表的容量大小,

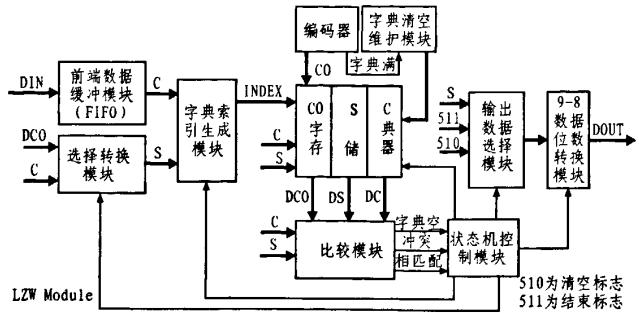


图 1 压缩模块内部结构框图

考虑 FPGA 内的 RAM 容量有限,这里选为 1 021;Offset 为偏移量,是为解决哈希函数冲突时用到的一个中间量。字典地址等于当前输入的字符数据左移一位后与字符串编码相异或的值。当哈希表中地址产生冲突时,则字典地址用式(3)计算,如果式(3)计算的字典地址小于 0,则字典地址用式(4)计算。随着压缩过程的进行,字典将被词条逐步填满,考虑到数据的局部相关性,为提高压缩比,字典清空维护模块将对存储满的字典进行清空操作,使字典能够自适应新输入的数据。状态机控制模块是整个压缩模块的时序控制中心,它是根据其他模块反馈给它的信号,判断压缩的运行状态,从而控制其他各个模块的运行。

3 系统综合、仿真及调试

将滤波和压缩两个功能模块整合到一个工程里,以 Altera 公司的 CycloneII 系列中的 EP2C5T144I8 作为目标器件,经过 QuartusII 软件编译综合,得到工程的综合报告。表 1 是整个工程的 FPGA 资源占用情况。

表 1 FPGA 资源占用情况

资源类型	总数量	使用量	占用百分比
LE	4 608	734	16%
Memory Bits	119 808	33 280	28%
PLL	2	1	50%
PIN	89	22	25%

由综合结果可知,系统能够稳定运行的最高工作频率是 102 MHz,根据设计特点,平均每 10 个时钟周期压缩 1 个数据,数据输入的位宽为 8 bit,所以整个系统能够有效的处理能力为:81.6 Mb/s。

图 2 是 LZW 模块的压缩仿真图,图中的 OUTPUT 是 9 位的输出,Dataout 是最终的 8 位结果,可见 LZW 模块仿真的结果无误。

系统调试时,由信号发生器产生相应的信号,经 A/D 转换后数据分为两路,其中一路传送给 FPGA,FPGA 对数据进行滤波和压缩,然后将处理过的数据传给电脑并保存为数据文件。另一路数据不经过 FPGA 处理直接传给电脑,在电脑上用软件实现对数据进行滤波和压缩运算,将运算结果同样保存为文件。2 个文件经对比后,其中的数据内容完全一致,说明整个硬件滤波和压缩电路能够正确无误的实现功能。同

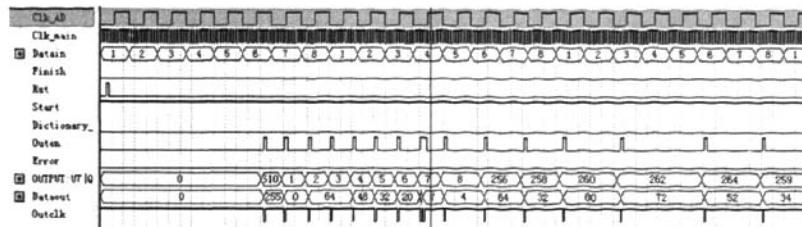


图2 压缩模块的仿真结果

时,用专用软件打开滤波前后的数据,观察波形后可知信号经过滤波后衰减很小,波形中的高频A/D量化噪声被有效的抑制,曲线的平滑度较好,这也说明滤波器的设计达到预期的效果。

为了评估LZW系统的压缩效果,选取Canterbury文件集

合进行测试^[6],表2是测试的部分结果。由于在FPGA内部实现LZW字典,其容量有限,造成对文本类的数据压缩效果较差。相比文本数据,对传真图像类的数据有着较好的压缩效果。对实际由A/D转换回来的数据进行测试,压缩率约为25%,说明本设计可以应用于实际的数据采集系统中。

表2 Canterbury文件测试部分结果

文件描述	文件名称	文件大小/Byte	压缩后大小/Byte	压缩率/%
英文文本(Alice in Wonderland)	alice29.txt	152 089	113 518	74.6
HTML文档	cp.html	24 603	20 161	81.9
电子表格文件	kennedy.xls	1 029 744	279 142	27.1
传真图像	ptt5	513 216	83 127	16.2
SPARC可执行文件	sum	38 240	22 619	59.2
E.Coli 细菌的完整基因	E.coli	4 638 690	1 799 757	38.8

4 结论

本设计通过FPGA从硬件上实现了滑动滤波算法和LZW无损压缩算法。经过实际测试验证,该系统能够正确无误的对数据进行滤波和压缩运算。压缩效果较好,能够有效缓解采集系统中的数据传输及存储的压力。同时,81.6 Mb/s的压缩速度可以满足大部分数据采集系统的实时要求。若用外部RAM存储器实现LZW字典,将能大幅度地扩大字典容量,显著改善数据的压缩效果。

参考文献:

[1] 新 鸿,裴东兴,祖 静.无损数据压缩在数据采集系统中的

应用探讨[J].华北工学院学报,2003,24(1):69-71.

- [2] 李雷定,马铁华.利用FPGA实现LZW算法的研究[J].电子测量技术,2008,31(10):170-172.
- [3] 黄凯明.滑动平均数字滤波参数研究[J].集美大学学报,2006,11(4):381-384.
- [4] 朱 钢,臧亚琴,刘 方.压缩技术在数据采集系统中的应用[J].数据采集与处理,1997,12(1):69-71.
- [5] Nelson Mark.LZW data compression[EB/OL].1989.http://marknelsonus/1989/1001/lzw-data-compression/
- [6] David salomon.数据压缩原理与应用[M].吴乐南,等译.北京:电子工业出版社,2003.

征 稿 启 事

—您的成功需要我们成就

《电子设计工程》(月刊)以其创新性、实用性、前瞻性,全面提升自身的学术水平和国际化程度。主要栏目为:测控与仪器仪表;网络与通信工程;图像分析与多媒体;数字信号处理;嵌入式系统;计算机应用;工业自动化;传感器技术;电源技术应用;信息安全;电力控制;汽车电子;集成电路;移动商务技术;消费电子;新特器件应用等。《电子设计工程》质优、实用、时效性强。敬请广大业界人士积极关注,踊跃投稿!

投稿信箱: editor@ieechina.com

在线投稿: http://mag.ieechina.com

《电子设计工程》杂志社

FPGA实现滑动平均滤波算法和LZW压缩算法

作者: 陈世海, 裴东兴, 张琦
 作者单位: 中北大学仪器科学与动态测试教育部重点实验室, 山西, 太原, 030051
 刊名: 电子设计工程 [STIC]
 英文刊名: ELECTRONIC DESIGN ENGINEERING
 年, 卷(期): 2010, 18(2)
 被引用次数: 0次

参考文献(6条)

1. 靳鸿, 裴东兴, 祖静. 无损数据压缩在数据采集系统中的应用探讨. 2003(1)
2. 李雷定, 马铁华. 利用FPGA实现LZW算法的研究. 2008(10)
3. 黄凯明. 滑动平均数字滤波参数研究. 2006(4)
4. 朱钢, 臧亚琴, 刘方. 压缩技术在数据采集系统中的应用. 1997(1)
5. Nelson Mark. LZW data compression. 1989
6. David salomon. 吴乐南. 数据压缩原理与应用. 2003

相似文献(3条)

1. 学位论文 孙王强 多路微弱电流采集系统设计 2007

本文详细介绍了多路微弱光电流采集系统的设计方案。传统的电流采集由于反馈电阻参数的分散性及信号放大处理电路中各通道的不均匀性误差对测量结果的准确性有很大影响, 基于上述原因, 本文提出了电流直接切换方案。通过大量试验选择合适的模拟开关, 选择某一路电流后利用精密I/V转换电路, 得到相应电压, 此方案可从原理上消除I/V转换电路带来的分散性误差。由于模拟开关各通道间寄生电容的存在, 在I/V转换时转换电压并不能立即稳定, 出现振荡现象, 本文分析这一现象, 提出电路模型并作出相应的解释。由于上述振荡现象的出现, 使每一路信号的采集周期变长, 为达到系统指标, 利用并行处理思想, 用4个I/V转换模块同时工作以降低整个电路的采集时间, 使用模拟开关CD4052选择即将达到稳定的电压信号进入采样保持器, 在前向通道中采用“流水线”式作业, 通过编码电路与可编程控制放大器PGA203实现自动量程转换, 不仅提高了模数转换的精度和速度, 而且还提高了电路的动态范围和处理速度。

比较了89C51与MSP430单片机的优缺点, 由于MSP430内部集成了模数转换、两通道支持异步或同步串行通信接口、支持16位运算并且超低功耗, 因此选用以TI公司生产的MSP430F149单片机作为数据采集系统的核心器件, 在单片机内部采用滑动平均滤波算法对采集数据作数字滤波处理, 利用RS-232实现单片机和PC机的串行通讯。在整个PCB板设计中前向通道与MSP430的高电平电压并不兼容, 在产品设计时要转换电平。最后简要说明制作PCB板时的注意事项。

上位机使用功能强大的图形化编程工具Labview为数据处理平台, 对采集数据作各种指标分析, 包括: 均值、偏差以及方差计算、线性度分析、均匀性分析、稳定性分析。测试结果表明, 该方案能提高系统性能。

2. 学位论文 史步海 城市供水混凝土沉淀大时滞过程建模及控制研究 2009

水是人类生活和生产过程中必不可少的重要资源, 特别是现代社会水资源遭受环境的破坏越来越严重, 而人们对水的需求量更是与日俱增, 且对水质要求也越来越大。伴随着人类社会的进步和城市的发展, 人们对自来水净化处理技术的研究特别是混凝沉淀过程的研究一直都没有停止。人们通过各种先进技术创新供水质量和生产自动化水平, 降低生产能耗、药耗、提高经济效益。混凝沉淀过程是城市供水处理过程中的一个重要工艺环节, 其混凝剂投加的自动控制是实现自来水生产自动化和保障水质品质的重要一环。混凝沉淀过程是一个大时滞的化学和物理反应过程, 其反应机理非常复杂, 不但与原水浊度、温度、流量、pH值、水中有机物有关, 还与水的混合方式、反应池环境、水力条件等反应条件有关。所以混凝沉淀过程是一个多变量、大时滞、多扰动、非线性过程。多年来对混凝沉淀过程以及自动化控制的研究一直是自来水和自动控制行业人们长期以来最为关心和研究的一个热点。

对于城市自来水混凝沉淀过程的自动化控制, 人们不断地尝试着采用各种的方法进行自动化控制方面的努力。特别是20世纪70年代以后随着自动化技术和电子计算机技术的发展而得到了更加迅猛的发展, 使得自来水混凝沉淀过程的自动控制技术和理论有了飞跃式现代的发展, 产生许多种的智能控制方法。智能控制是自动控制发展的高级阶段, 是人工智能、控制论、系统论和信息论等多种学科的高度综合与集成, 它主要包括神经网络控制、模糊控制、预测控制和专家系统控制等。

本文在分析介绍了混凝沉淀过程的机理以及前人对混凝沉淀过程进行的各种自动控制方法研究基础上, 为了更加高效和高精度针对多变量、大时滞、时变的混凝沉淀过程进行自动控制, 主要从智能控制策略方面展开了深入研究, 主要内容包含有: 现场数据采集的滤波及数据归一化预处理方面; 神经网络结构与训练算法方面; 模糊神经网络结合及控制算法方面; 神经网络多步预测自适应控制方面。

本文通过系统和深入的研究, 获得了以下有一定创新性成果:

1. 提出和实现了一种神经网络样本数据预处理中数字滤波的新算法, 即多限幅滤波加滑动平均滤波复合算法。新的方法采用多种限幅值的方式, 有效的去除了粗大误差和非正常数据输出, 保存了在一定噪声限幅内的实际采样值, 并与滑动平均滤波算法一起实现了样本数据的去伪存真和平滑滤波。本算法直观性强、实用有效, 在较好的去除了粗大误差和超标噪声信号的同时, 也很好的提高了算法实现的效率。

2. 改进和实现了一种神经网络样本数据预处理中数字归一化的新算法。改进归一化算法是在无量纲极值法基础上改进而来, 此方法根据混凝沉淀过程中河水温度变量对混凝沉淀的影响方式主要是全年范围上的温度值变化, 而非这一时期采样时间段的温度变化, 故提出了改进的全范围数据归一化处理方法。此归一化方法既有效的显示出在全范围的温度值大小也显示出温度值的变化趋势。即描述了变量值的大小还描述了变量的变化量趋势, 从而能够更加准确的反映此变量对控制过程变化的影响和函数关系, 是一种在全数据范围内有效的数字归一化方法。

3. 提出和实现了变步长 θ 的改进LMBP算法。本文在分析标准LMBP算法优缺点过程中, 知道 μ 值的动态合理变化量是LMBP算法在梯度下降法和高斯-牛顿法之间快速转换和收敛的重要参数之一。为了进一步加快LMBP算法的收敛速度, 同时为了避免 μ 值变化过程中发生“小步长”情况, 而阻碍LMBP算法的收敛速度, 提出了变步长 θ 的改进LMBP算法。仿真研究表明, 改进LMBP算法的收敛速度总是优于标准LMBP算法。此算法比标准的LMBP算法的收敛次数平均减少了约70%。

4. 提出和实现了基于先验知识约束和少量样本数据的Elman动态递归神经网络新的算法。本文通过对混凝沉淀过程机理分析与总结, 提出了混凝沉淀过程的两个先验知识约束条件, 并在基于Elman动态递归神经网络性能函数中引入了先验知识的约束条件, 从而使得能够在用较少数据样本的神经网络训练中获得较为完备的神经网络模型, 增强了神经网络模型的有效性和可靠性。本文并对此神经网络的算法进行了全面的推导, 并对此方法进行了仿真实验, 实验表明了所述方法的正确和有效的。

5. 提出和实现了一种基于混凝沉淀过程的自适应模糊神经网络控制改进算法。本文在阐述了模糊神经网络系统相融合的理论基础上, 根据第二章中所

述的影响混凝沉淀过程的主要变量并对这些变量具体的采样数据进行了模糊化分析。文中在论述了传统的自适应模糊神经网络ANFIS算法基础上，为提高模糊神经网络的收敛速度，提出了采用了本文所述的改进LMBP+LES的改进算法，进一步加快了模糊神经网络的收敛速度。仿真数据表明文中所述方法是可行的，控制精度基本满足要求。

6. 提出和实现了一种针对多变量大时滞系统的神经网络多步预测神经网络自适应控制的新方法。本文根据具体的混凝沉淀多变量大时滞系统大时滞、多变量和时变的特点，结合本文在第三章和第四章所提出改进LMBP算法和基于先验知识的Elman网络算法基础上，提出了针对多变量大时滞系统的神经网络多步预测神经网络自适应控制方法。在此方法中，为了进一步加强了时滞系统对扰动量的动态预测输出能力，在多步预测神经网络模型的输入端，增加扰动量的一阶导数为输入元。在神经网络NNC控制器中并联一个辅助控制器AC组成了复合控制器方式，提高了控制器的抗干扰能力稳定性，加快了对于大扰动量干扰时，神经网络控制器NNC的快速导入和收敛速度。分别采用了基于改进LMBP的神经网络算法和基于先验知识的Elman神经网络算法对神经网络多步预测自适应控制系统中的NNP模型和INNC模型进行了仿真实验。通过仿真研究表明，二者都能够快速的响应和收敛，方法是可行和有效的。

3. 学位论文 刘维东 多通道微弱电流采集电路的检测与改进 2008

本文介绍了一种适用于激光粒度测试仪的多路微弱光电流采集电路的设计方案以及对采集电路进行性能评价和故障诊断的设计方案。课题中完成的任务共分四部分：设计了一个128路光电流采集电路；制作了一个8路程控电流源；制作了一个采集电路性能评价系统；初步研制了一个采集电路故障诊断系统。程控电流源作为可控标准信号源向采集电路提供电流信号，采集电路完成采集后把数据发送给上位机，性能评价系统完成采集电路各项指标的评价。故障诊断系统通过软件诊断和硬件诊断相结合的方式快速准确地定位故障点。

传统的多路电流采集电路设计方案中，先进行I/V转换，再模拟开关切换。I/V转换中反馈电阻参数的分散性及信号放大处理电路中各通道的不均匀误差对测量结果的准确性有很大影响，基于上述原因，本文提出了电流直接切换方案。电流先通过模拟开关进行切换，然后再利用精密I/V转换电路，得到相应的电压，此方案可从原理上消除先I/V转换后模拟开关切换带来的分散性误差。在试验当中发现模拟开关切换后电流有不稳定现象。本文分析了这一现象，并提出改进的措施。在信号滤波环节，制作了一个50HZ的陷波器，用以滤出工频电对采集电路的干扰。通过电平比较编码电路与可编程放大器PGA203实现自动量程放大，提高了电路的动态范围和处理速度。鉴于采集电路中控制端口多和实时性要求高的特点，主控器采用FPGA，型号为EP2C8Q208C8。FPGA完成采集电路前向通道中模拟开关、采样保持器和模数转换器的控制。在FPGA内部采用滑动平均滤波算法对采集数据作数字滤波处理，利用RS232实现FPGA与PC机的通信。

八路程控电流源以AT89S51单片机为控制核心，通过数模转换器控制压控电流源实现电流的给定，同时利用信号调理电路与模数转换器采集生成的电流信号，在单片机内部使用PID数字调节器实现电流闭环调节，在此基础上还实现了串行通信功能，可用PC机代替键盘对电流进行设定与监控。

采集电路性能评价系统根据数据采集系统的检测方法及计量检测标准，对采集数据的线性度、均匀一致性、准确度、稳定性等指标进行了测试，然后用字符和曲线图将测试结果显示出来。

故障诊断系统通过软件诊断与硬件诊断相结合的方式，对128路光电流采集电路中的故障进行快速准确的定位。按照采集电路各部件由后到前的顺序，依次完成串口、FPGA、RAM、程控放大、采保、模拟开关等器件的故障搜索与定位。

本文链接：http://d.g.wanfangdata.com.cn/Periodical_dzsjgc201002024.aspx

授权使用：陕西理工学院(sxlgxy)，授权号：69d45995-a4cc-4f2d-8419-9df20102c257

下载时间：2010年9月15日