

DOI: 10.11830/ISSN.1000-5013.202010035



# 采用通用公共无线接口协议的 有损 I/Q 数据压缩方案

段辉鹏, 凌朝东, 罗继亮

(华侨大学 信息科学与工程学院, 福建 厦门 361021)

**摘要:** 为了降低光纤链路数据传输量, 提出一种采用通用公共无线接口(CPRI)协议的有损 I/Q 数据压缩方案. 采用分组压缩的方式对原始数据进行压缩, 用 Matlab 软件验证该压缩方案的可行性, 并用 Verilog HDL 语言设计压缩模块, 在现场可编程门阵列(FPGA)上进行硬件验证. 结果表明: 文中压缩方案能够有效地压缩数据并恢复数据, 降低光纤链路数据传输量.

**关键词:** 通用公共无线接口; 数据压缩; 数据传输量; 现场可编程门阵列

**中图分类号:** TN 911.25      **文献标志码:** A      **文章编号:** 1000-5013(2021)06-0825-06

## Lossy I/Q Data Compression Scheme Using Common Public Radio Interface Protocol

DUAN Huipeng, LING Chaodong, LUO Jiliang

(College of Information Science and Engineer, Huaqiao University, Xiamen 361021, China)

**Abstract:** In order to reduce the data transmission volume of optical fiber link, a lossy I/Q data compression scheme base on the common public radio interface (CPRI) protocol was proposed. The original data is compressed by the form of grouping compression, and the feasibility of the compression scheme is verified by Matlab software, and the compression module is designed with Verilog HDL language, the hardware verification is carried out on field programmable gate array (FPGA). The results show that the proposed compression scheme can effectively compress and recover data, reduce the data transmission volume of optical fiber link.

**Keywords:** common public radio interface; data compression; data transmission volume; field programmable gate array

未来移动通信系统大容量、低延时、高速率等应用场景正推动着第 5 代移动通信技术(5G)的蓬勃发展, 以低成本网络提供高质量的通信业务将是影响 5G 未来发展的关键因素<sup>[1]</sup>. 以云化无线接入网(C-RAN)为代表的集中化无线接入网架构是实现低成本接入网的重要技术之一<sup>[2]</sup>, 它将基站的基带部分和射频部分分离, 基带处理部分集中于室内基带处理单元(BBU), 而远端基站退化为射频拉远单元(RRU), 连接 BBU 和 RRU 之间的通信链路称为前传链路<sup>[3]</sup>. 5G 传输的数据速率是 4G 的 100~1 000 倍<sup>[4]</sup>, 为实现网络的高速接入, 天线端将采用大规模多天线阵列技术, 前传传输的带宽也会急剧增加. 因此, 采用通用公共无线接口(CPRI)协议的光纤接入前传承载方案非常具有优势<sup>[5]</sup>. 然而, 采用 CPRI 协

**收稿日期:** 2020-10-24

**通信作者:** 凌朝东(1964-), 男, 教授, 博士, 主要从事专用集成电路与系统设计的研究. E-mail: edac@hqu.edu.cn.

**基金项目:** 国家自然科学基金资助项目(61973130); 福建省自然科学基金资助项目(2017J01116); 华侨大学研究生科研创新基金资助项目(18013082019)

议的 5G 单站点带宽需求将达到百吉比特每秒量级<sup>[6]</sup>, 这对光纤接入的成本而言是一大挑战, 为了降低数据传输量, 本文提出一种采用通用公共无线接口协议的有损 I/Q 数据压缩方案<sup>[7]</sup>.

1 CPRI 协议

通用公共无线接口 (CPRI) 联盟制定了开放的接口标准, 定义了两个基本单元, 即无线设备控制中心 (REC) 和无线设备 (RE). CPRI 协议涉及射频、基带、控制、传输等模块之间的接口规范, 主要是针对物理层和数据链路层. CPRI 协议的整体架构, 如图 1 所示<sup>[8]</sup>.

物理层支持电气接口和光接口, 并且支持多种线比特率和 8B/10B, 64B/66B 两种编码方式. 不同的线比特率旨在满足 REC 和 RE 之间传输的灵活配置, 可以节约成本、提高效率<sup>[9]</sup>. 数据链路层支持 I/Q 数据、厂商特定信息、以太网协议或高级链路控制协议、L1 带内协议. 根据物理层和数据链路层的协议, 数据都采用多路复用的方式进行传输<sup>[10]</sup>. CPRI 协议传输 3 种类型的数据流, 分别为同步数据流、控制管理数据流和用户数据流, 同步数据流主要负责 REC 和 RE、RE 和 RE、REC 和 REC 之间的数据同步, 控制管理数据流主要负责对 CPRI 链路的控制和管理, 用户数据流即 I/Q 数据流, 主要是基站和移动设备之间传输的数据, 即文中的压缩对象.

2 系统整体设计

光纤基带传输系统简化框图, 如图 2 所示. 图 2 中: FPGA 为现场可编程门阵列. 由于上行链路和下行链路的工作方式基本一致, 故以下行链路为例进行说明.

信号由天线接收进入混频器, 混频器主要完成射频到中频的处理, 经混频器处理后, 信号进入模数转换器 (ADC), 将模拟信号转变为数字信号, 输出两路信号; 为了降低后续光模块的数据传输量, 先将信号送入一个数据压缩模块, 以降低数据的吞吐量, 将压缩后的信号送入 CPRI 协议接口模块, 并进行相应的组帧; 经串行收发器将信号通过光纤发送到远端, 远端通过串行收发器恢复出压缩后的信号; 利用 CPRI 协议对信号进行解帧, 并将数据输出到解压缩模块, 进行解压处理; 将信号送给后续的数模转换器 (DAC), 完成数字信号到模拟信号的转换, 经混频器完成中频到射频的上变频处理并输出.

3 数据压缩方案

ADC 的采样率为 122.88 MHz, 输出为 16 bit 二进制有符号数, 为了降低光纤链路的数据传输量, 降低开发成本, 根据原始数据绝对值的大小, 对数据截断处理后进行压缩, 截断后的数据为 8 bit 二进制有符号数. 在压缩过程中, 先将经过 ADC 之后的数据进行分组<sup>[11]</sup>, 为了简化压缩过程的复杂度, 以及便于 FPGA 实现, 需要将每组数据样本点数都设置为相同. 为了保证压缩精度, 每组数据样本点数不能太多, 但每组数据样本点数太少又会导致压缩效率的降低, 因此, 在确定每组数据样本点数时, 需要进行折中考虑, 既要保证一定的压缩精度, 又要保证一定的压缩效率<sup>[12]</sup>, 最终压缩方案的每组数据样本点数确定为 32 个. 在压缩操作前, 需要对原始数据样本点绝对值进行逐一比较, 找出每组数据样本点绝对值中的最大值, 通过最大绝对值确定每组数据样本点的压缩因子  $K$ , 最大绝对值越大, 压缩因子  $K$  就越大. 压缩方案的压缩因子为 4 bit 无符号数, 但压缩因子  $K$  的范围只为 4~8, 而剩余的数并未用到. 在对原始数据绝对值进行比较时, 首先, 把每组第 1 个数据的绝对值当作这组的最大绝对值. 然后, 再对后续的 31 个样本点绝对值进行逐一比较, 直到比较 31 次后, 得到本组样本点中的最大绝对值. 在进行绝对值

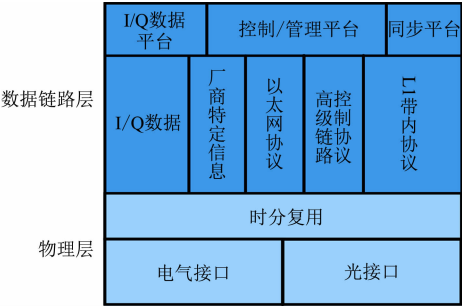


图 1 CPRI 协议的整体架构  
Fig. 1 Overall framework of CPRI protocol

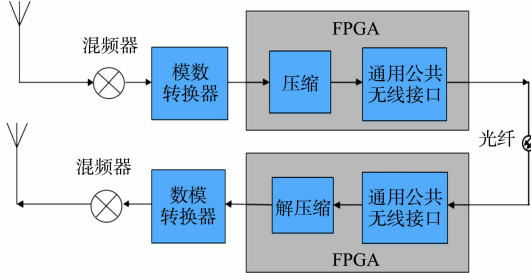


图 2 光纤基带传输系统简化框图  
Fig. 2 Simplified block diagram of optical fiber baseband transmission system

比较的同时,需要将接收的基带 I/Q 数据存入缓存,以便和压缩因子  $K$  进行时序对齐,保证对数据的正确压缩.最后,进行压缩操作,将 16 bit 的基带 I/Q 数据进行相应的截位压缩至 8 bit,根据压缩因子  $K$  的大小,截掉经缓存后数据的低  $K$  位.由于有符号数的高位通常是符号位的扩展,除了最低一位符号位,其他都可看作冗余位,在截位过程中也会把部分符号位截掉,剩下中间的 8 bit,即第  $K$  到第  $K+7$  bit 为最终压缩的数据(最低位为从右至左第 0 bit).

为保证接收端能够正确解压数据,4 bit 压缩因子  $K$  总是在 32 个已压缩数据之前传输,然后,按照 CPRI 协议进行组帧,通过光纤传输给远端.压缩模块,如图 3 所示.接收端先将数据按照 CPRI 协议进行相应的解帧,再对数据进行解压处理.解压过程相对简单,解压端接收到压缩因子  $K$ ,并将其保持 32 个时钟周期.同时,根据压缩因子  $K$  对紧接着接收到的 32 个 8 bit 已压缩数据进行解压处理,恢复出 16 bit 数据,其中,低位填补 0,高位填充符号位.

文中压缩方案能够将原来需要传输的 16 bit 数据压缩到 8 bit,实现了 2 : 1 的压缩,可以节约一半的光纤资源带宽.例如,单天线基带带宽为 20 MHz 的 LTE-FDD 信号,采样速率  $F_s$  为 30.72 MHz,数据位宽  $W$  为 16 bit 时,不对数据进行压缩,CPRI 所需线比特率(Mbit · s<sup>-1</sup>)为

$$F_s \times W \times 2 \times 16 / 15 \times 10 / 8 = 30.72 \times 16 \times 2 \times 16 / 15 \times 10 / 8 = 1\,310.72. \tag{1}$$

式(1)中:2 表示基带 I/Q 两路信号;16/15 表示 CPRI 协议本身有控制字开销,需要占一定的传输资源;10/8 表示采用 8B/10B 编码<sup>[13]</sup>.

根据压缩方案,数据位宽  $W$  被压缩至 8 bit,此时,CPRI 所需线比特率(Mbit · s<sup>-1</sup>)为

$$F_s \times W \times 2 \times 16 / 15 \times 10 / 8 = 30.72 \times 8 \times 2 \times 16 / 15 \times 10 / 8 = 655.36. \tag{2}$$

## 4 Matlab 仿真分析和性能分析

### 4.1 Matlab 仿真分析

为验证压缩方案的有效性,通过 Matlab 软件对 LTE-FDD 信号进行仿真分析.取经过 ADC 之后的信号,采样率为 122.88 MHz,每组数据固定为 32 个样本.为提高压缩精度,对压缩后的数据进行四舍五入处理.

原始信号和压缩解压后信号的时域波形,如图 4 所示.4 096 个 I/Q 样本值的原始信号和压缩解压后信号的功率谱密度(PSD),如图 5,6 所示.图 5,6 中: $f$  为信号频率.由图 4~6 可知:无论是时域波形,还是功率谱密度,压缩解压后的信号都能够较好地逼近原始信号.

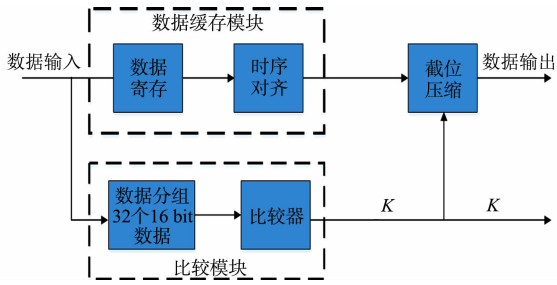


图 3 压缩模块

Fig. 3 Compression module

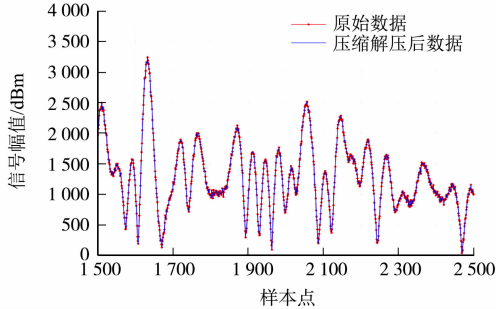


图 4 原始信号和压缩解压后信号的时域波形  
Fig. 4 Time domain waveforms of original signal and signal after compression and decompression

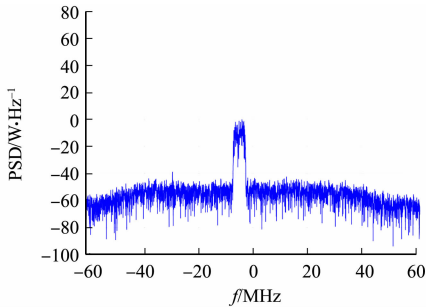


图 5 原始信号的功率谱密度

Fig. 5 Power spectral density of original signal

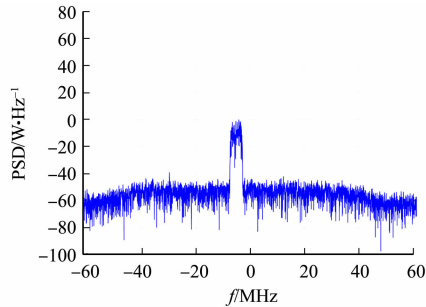


图 6 压缩解压后信号的功率谱密度

Fig. 6 Power spectral density of signal after compression and decompression

4.2 性能分析

文中压缩方案为有损压缩,压缩解压后的信号相对于原始信号的误差主要来自以下 3 个方面. 1) 单个样本点绝对值较大的原始数据压缩解压后数据的低位直接填补 0,这会造成较大绝对误差,但相对误差较小且原始信号本身质量较好,压缩解压后不会影响接收端的正确接收,所以相对于原始数据,压缩解压后的相对误差可以忽略. 2) 当一组数据中既有绝对值很大的数据,也有绝对值很小的数据,这样一组数据经过文中压缩方案压缩解压后,绝对值很小的数据与原始数据的相对误差较大,这是分组压缩固有的缺点. 这种压缩解压后的精度损失无法在解压端进行补偿,但此时原始信号与压缩解压后的信号的绝对误差较小,且原始信号本身携带的信息量较小,压缩解压后的相对误差不会影响接收端的正确接收. 3) 误差和原始信号的大小无直接关系,但最大绝对误差和压缩因子成正比关系,即压缩因子越大,最大绝对误差就越大. 由于 I/Q 两路同时分别进行压缩,I/Q 两路产生的最大绝对误差  $E_{\max}$  与压缩因子  $K$  的关系为  $E_{\max}=2^{K-1}$ .

由此可知,压缩因子  $K$  的最大值为 8,那么理论上最大绝对误差的最大值可达到 128,但由仿真的时域波形可知,原始数据的最大绝对值不会超过 4 000,故压缩因子  $K$  也不会超过 6. I/Q 两路原始信号与压缩解压后的信号的绝对误差,如图 7 所示. 图 7 中: $n$  为样本点数; $E_I$  为同向链路绝对误差; $E_Q$  为正交链路绝对误差.

误差矢量幅度 EVM 常被用于评估压缩解压对信号造成的压缩损伤,EVM 越高,说明压缩对原始信号造成的损伤越大. EVM 的计算公式为

$$EVM=\sqrt{E[|decompressed\_data-source\_data|^2]}/E[|source\_data|^2]}\times100\%.$$
 (3)

式(3)中:source\_data 为原始信号;decompressed\_data 为压缩解压后的信号.

根据国际电信联盟 3GPP 规范文档要求,在 QPSK,16QAM,64QAM,256QAM 等 4 种调制方式下,端到端的 EVM 分别为 17.5%,12.5%,8.0%,3.5%. 压缩造成的 EVM 是不同调制方式端到端 EVM 的一部分<sup>[14-15]</sup>. 文中压缩方案实现 2 : 1 压缩比造成的 EVM 为 0.58%,在 3GPP 规范文档的要求范围之内.

5 FPGA 的验证

在 FPGA 验证中,I/Q 数据将分开进行压缩,由于 I/Q 压缩模块是相同的,故只对其中一个压缩模块进行说明. FPGA 采用 Intel 公司的 Cyclone IV 系列芯片 EP4CE115F29C7,输入时钟频率为 122.88 MHz,与之配合的软件开发平台选用 Quartus II 13.0. 首先,将每组数据与压缩因子进行时序对齐,此处时序对齐并未选择先入先出队列(FIFO)缓存数据,而是使用 32 个 16 bit 移位寄存器进行延迟,以达到对数据和压缩因子  $K$  进行时序对齐的目的;然后,根据压缩因子  $K$ ,右移  $K$  位;最后,截取右移数据的低 8 bit. 整个压缩模块采用 Verilog HDL 语言,其寄存器传输级(RTL)结构图,如图 8 所示.

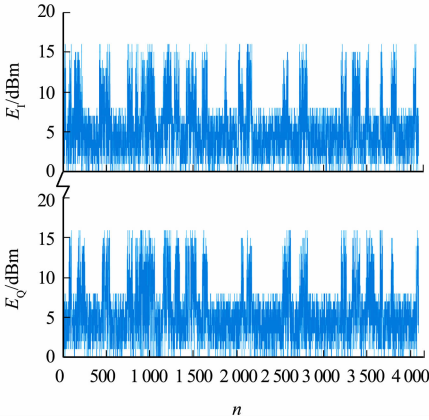


图 7 I/Q 两路数据的绝对误差

Fig. 7 Absolute error of I/Q two channel data

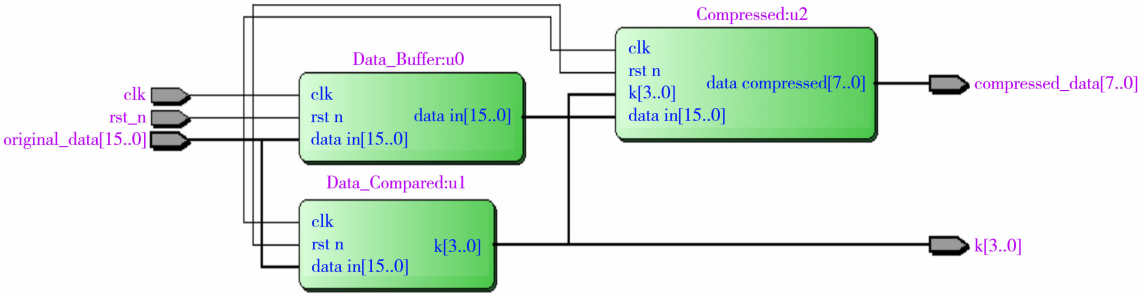


图 8 压缩模块 RTL 结构图

Fig. 8 RTL structure diagram of compression module



Data\_Compared 模块完成对原始数据的分组和比较, 分组实际上采用了模为 32 的计数器, 便于对 32 个原始数据进行比较, 得到对应的压缩因子  $K$ , 计数器从 0 开始计数, 每计数到 31 时就可比较出每组数据样本点的最大绝对值; 然后, 根据最大绝对值, 得出这组数据对应的压缩因子  $K$ , Data\_Buffer 模块实现对原始数据的缓存, 以便和压缩因子  $K$  进行时序对齐; 最后, 将压缩因子  $K$  和缓存后的数据同时送入 Compressed 模块进行压缩。

为了提高压缩精度, 对压缩后的数据进行四舍五入处理, 四舍五入会根据原始数据的正负, 分别判断截位后的数据是否需要进位。记 carry\_bit 为进位, 原始数据为正数时, 判断截去的最高位是否为 1, 如果为 1, 那么 carry\_bit 就等于 1。此时, 截位后的数据应该进位, 否则, carry\_bit 为 0, 截位后的数据就不会进位。和正数相反, 负数不进位要加 1, 当原始数据为负数时, 除了要判断截去的最高位, 还要判断截去的其他任意位是否为 1, 如果截去的最高位为 1, 且其他位任意一位上还有 1, 那么 carry\_bit 为 1, 否则, carry\_bit 等于 0。考虑到进位相加可能会有溢出, 所以需要将四舍五入的数据扩展 1 bit, 防止溢出。记 compressed\_round[8:0] 为四舍五入后的数据, 其最高位为符号扩展位。记 compressed\_temp[7:0] 为四舍五入前的数据, 是否需要进位可以表示为

$$\text{compressed\_round} = \{\text{compressed\_temp}[7], \text{compressed\_temp}\} + \text{carry\_bit}.$$
(4)

式(4)中: {} 表示拼接符。

当 compressed\_temp[7:0] 为 01111111 时, 如果 carry\_bit 等于 1, 根据式(4), compressed\_round 等于 01000000, 因为需要传输的数据为 8 bit 有符号数, 能表示最大的正数为 127, 所以需要进行饱和处理。记 compressed\_data[7:0] 为最终的压缩数据, 根据 compressed\_round 最高两位是否相等来判断是否有溢出。如果 compressed\_round 最高两位相等, 表示没有溢出, 最终压缩后的数据为四舍五入后数据的低 8 bit, 即

$$\text{compressed\_data}[7:0] = \text{compressed\_round}[7:0].$$
(5)

如果 compressed\_round 最高两位不相等, 表示有溢出, 此时要进行饱和处理, 最终压缩后的数据为

$$\text{compressed\_data}[7:0] = \{\text{compressed\_round}[8], 7\{!\text{compressed\_round}[8]\}\}.$$
(6)

解压过程相对简单, 解压端先提取压缩因子  $K$ , 并将  $K$  保持 32 个时钟周期, 以保证对后续 32 个数据进行正确解压; 然后, 根据压缩因子  $K$ , 将压缩后的 8 bit 有符号恢复为 16 bit 有符号数。文中压缩方案在综合布局布线后的资源消耗情况, 如图 9 所示。由图 9 可知: 直到第 32 个周期才比较出每一组数据中的最大值; 第 33 个周期才得出压缩因子  $K$ ; 根据压缩因子得到压缩后的值, 每一个数据都相应地向后延迟了 32 个周期; 虽然压缩方案采用四舍五入消耗了更多的逻辑资源, 但提高了压缩精度。在慢速 1.2 V, 85 °C 模式下, 时钟频率达到 182.62 MHz, 满足系统时钟频率要求的 122.88 MHz。

ModelSim 软件的仿真效果图及仿真误差, 如图 10, 11 所示。由图 10, 11 可知: 文中压缩方案能够有效地压缩数据并恢复数据。

Flow Summary	
Flow Status	Successful - Wed Sep 16 22:13:50 2020
Quartus II 64-Bit Version	13.0.0 Build 156 04/24/2013 SJ Full Version
Revision Name	data_compress
Top-level Entity Name	co2de_data
Family	Cyclone IV E
Device	EP4CE115F29C7
Timing Models	Final
Total logic elements	447 / 114,480 (< 1 %)
Total combinational functions	430 / 114,480 (< 1 %)
Dedicated logic registers	56 / 114,480 (< 1 %)
Total registers	56
Total pins	66 / 529 (12 %)
Total virtual pins	0
Total memory bits	928 / 3,981,312 (< 1 %)
Embedded Multiplier 9-bit elements	0 / 532 (0 %)
Total PLLs	1 / 4 (25 %)

图 9 压缩方案资源消耗情况

Fig. 9 Resource consumption situation of compression scheme

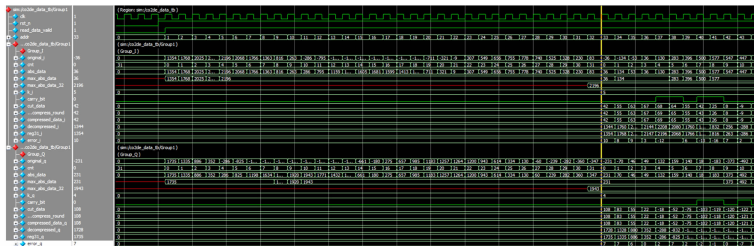


图 10 ModelSim 软件的仿真效果图

Fig. 10 Simulation effect diagram of ModelSim software

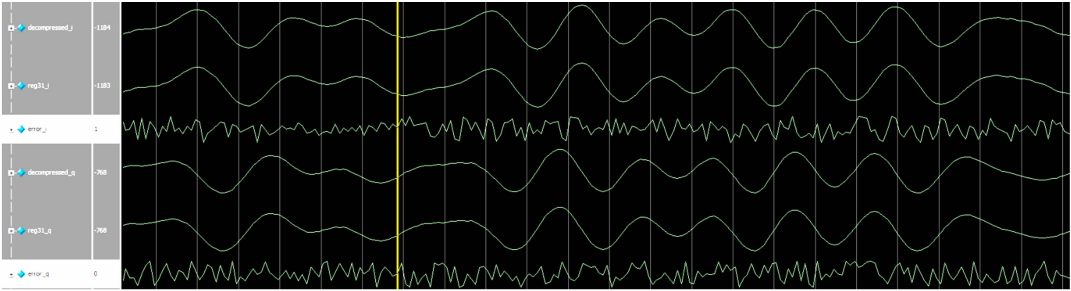


图 11 ModelSim 软件的仿真误差  
Fig. 11 Simulation error of ModelSim software

6 结 束 语

为了降低光纤链路数据传输量,提出一种采用 CPRI 协议的有损 I/Q 数据压缩方案,并对其进行仿真。仿真结果表明:文中方案能够有效地压缩数据并恢复数据,降低光纤链路数据传输量,从而减少光纤资源的消耗量,降低开发成本。

参考文献:

[1] 李隆胜,胡卫生. 面向移动前传的大容量光接入网技术[J]. 中兴通讯技术,2019,25(5):31-37. DOI:10. 12142/ZTETJ. 201905005.

[2] 中国移动通信有限公司研究院. C-RAN 白皮书:无线接入网绿色演进 V3. 0[R]. 北京:中国移动通信有限公司研究院,2014.

[3] 邹辉辉. 基于 C-RAN 的前传压缩技术研究[D]. 南京:东南大学,2018.

[4] SI H,NG B L,RAHMAN M S,*et al*. A vector quantization based compression algorithm for CPRI link[C]//IEEE Global Communications Conference. San Diego:IEEE Press,2015:1-6. DOI:10. 1109/GLOCOM. 2015. 7417796.

[5] 常圆亮. 基于 FPGA 的多路 CPRI 测试系统设计与实现[D]. 哈尔滨:哈尔滨工程大学,2018.

[6] CPRI Group. Common public radio interface: CPRI interface specification V7. 0[EB/OL]. (2015-10-02)[2019-07-20]. [http://www.cpri.info/downloads/CPRI\\_v\\_7\\_0\\_2015-10-09.pdf](http://www.cpri.info/downloads/CPRI_v_7_0_2015-10-09.pdf).

[7] 郭彬,曹伟,陶安. LTE 和 LTE-Advanced 系统中 CPRI 压缩算法研究[J]. 信息通信技术,2013,7(2):64-69.

[8] 袁行猛,陈亮,徐兰天. 基于 CPRI 协议的 5G 基带数据传输技术的研究与实现[J]. 电子产品世界,2019,26(4):41-45,61.

[9] 赵小雪. 基于 Arria 10 芯片的 CPRI 实现[D]. 西安:西安电子科技大学,2019.

[10] 马志刚. CPRI 原理及测试解决方案(一)[J]. 电信网技术,2010(5):68-70.

[11] 张小波. 10 Gbps 以太网 CPRI 分组传输关键技术研究[D]. 成都:电子科技大学,2013.

[12] 胡塘,周涛,郭跃挺. 一种基于 CPRI 协议的 I/Q 数据压缩方案[J]. 光通信技术,2015,39(7):57-59. DOI:10. 13921/j.cnki. issn1002-5561. 2015. 07. 018.

[13] 曾素华. 5G 无线网基站带宽需求研究[J]. 湖南邮电职业技术学院学报,2019,18(4):6-8.

[14] GUO Bin,CAO Wei,TAO An,*et al*. LTE/LTE-A signal compression on the CPRI interface[J]. Bell Labs Technical Journal,2013,18(2):117-133.

[15] 3GPP Organizational Partners' Publications Offices. 3rd generation partnership project; technical specification group radio access network; evolved universal terrestrial radio access (E-UTRA); user equipment (UE) radio-transmission and reception (Release 14)[EB/OL]. (2018-03-12)[2019-07-20]. [https://www.3gpp.org/ftp/Specs/archive/36\\_series/36.101](https://www.3gpp.org/ftp/Specs/archive/36_series/36.101).

(责任编辑:钱筠 英文审校:吴逢铁)