

基于 JPEG2000 的地震数据无损压缩

孙寰宇¹ 武文波² 杨志高² 秦前清²

(武汉大学电子信息学院 ,武汉 430079)

(武汉大学遥感测绘信息工程国家重点实验室 ,武汉 430079)

E-mail :whuuniverse@yahoo.com.cn

摘 要 在实验基础上 ,针对地震信号的特性 ,基于 JPEG2000 标准文章提出了一套新的地震数据压缩混和编码方案 ,该方案不仅取得了较大的压缩率 ,而且大大提高了网络传输、实时处理方面的性能。

关键词 JPEG2000 混合编码 提升小波变换

文章编号 1002-8331-(2005)16-0185-04 文献标识码 A 中图分类号 TP39

Quasi-Lossless Compression of Seismic Data Based on JPEG2000

Sun Huanyu¹ Wu Wenbo² Yang Zhigao² Qin Qianqing²

(School of Electronic Information ,Wuhan University ,Wuhan 430079)

(State Key Laboratory of Information Engineering in Surveying ,Mapping and Remote Sensing ,Wuhan University ,Wuhan 430079)

Abstract : According to the special features of seismic data a new mixed coding method of seismic data compression based on JPEG2000 is proposed on experiment. This coding method makes high performance not only in seismic data compression but also in coding transmission and real-time analysis.

Keywords : JPEG2000 combined coding lifting wavelet transform

1 引言

近几年来 ,地震勘探正向多维、多分量、多参数、高分辨率方向发展 ,地震数据量不断增多 ,给存储器的存储容量、远程网络通信信道的带宽以及计算机的处理速度带来了极大的压力。必须对地震数据进行有效的、高精度的压缩 ,以节省大量的存储空间 ,提高实时处理速度 ,改善网络传输性能。

地震数据不同于一般的图像数据 ,其数据压缩更有挑战性。与一般的图像数据相比 ,地震数据的特性突出表现在以下几个方面 (1)地震波形是紧支集的 ,空间相关性和频谱间相关性都很强 ,函数的转换需要在空间域和频率域都要有良好的局部性 (2)地震信号分频带特性强 ,在二维谱上常表现出很强的条带性。在水平方向和垂直方向频谱特性差异很大 (3)地震信号的相位、幅度都包含着丰富的地层信息且动态范围大 (4)数据格式多样 ,数据结构复杂。地震数据最常用的格式是 SEG-Y 格式 ,其中包含着丰富的卷头信息和道头信息 ,涉及参数多 ,必须进行可逆的无损压缩 (5)压缩保真度要求高。地震信号最显著的特点之一 ,限制物理信息或信号特征的丢失 ,特别是高频信息。因为需要根据这些信息判断地层结构、油气藏情况等。

小波变换是一种具有较强时、频局部分析功能的非平稳信号分析方法 ,广泛用于地震数据压缩。由于地震信号的相位对于地层信息解译十分重要 ,有学者提出采用自适应小波变换可以更好地保留原始信号的相位信息 ,^[1]根据地震信号的分频带

特性 ,国内有学者提出了基于多进小波的压缩算法 ;^[2]为减少重建信号的畸变 ,有人提出在压缩算法中采用软阈值。^[3]但是传统的小波变换依赖于傅立叶变换 ,有大量的卷积运算 ,运算速度较慢 ,硬件实现复杂 ,内存占用大 ,更无法适应互联网时代网络传输的要求。

JPEG2000 标准采用第二代小波变换技术即提升小波变换技术 ,它继承了传统小波变换的优点 ,又因为它是一种整数到整数的非线性变换 ,所有的运算都在空间域进行 ,算法简单 ,运算速度快 ,量化误差小等优点更适合于地震数据的压缩。JPEG2000 标准又因其通用性强、扩展性好同时支持无损压缩、多分辨率分析以及渐进传输等优点更适应了地震勘探发展数据的压缩要求。

2 JPEG2000 的基本原理

JPEG2000 的核心部分是图像编码系统 ,其编码器和解码器的框图分别如图 1 和图 2 所示。

JPEG2000 图像编码系统基于最佳截断嵌入式码块编码算法 (Embedded Block Coding with Optimized Truncation ,EBCOT) ,使用小波变换 ,对压缩位流分层组织 ,不仅可以取得更好的压缩比 ,而且使得压缩码流具有很大的灵活性。编码时 ,对输入图像进行离散小波变换 ,根据变换后的小波系数特点进行量化。将量化后的小波系数划分成小的数据单元码块 ,对每个

码块进行独立的算术编码,将得到的所有码块的嵌入式位流,按照率失真最优原则分层组织,形成不同质量的层,对每一层,按照一定的码流格式打包,输出压缩码流。

解码过程相对比较简单,根据压缩码流中存储的参数,对应于编码器各部分进行逆向操作,输出重构图像数据。

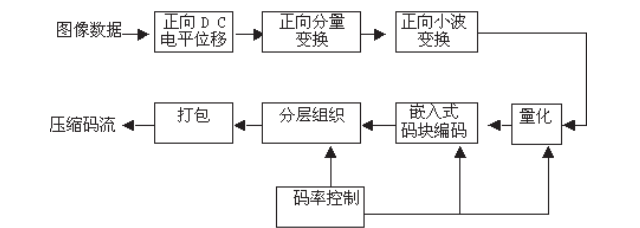


图1 JPEG2000 编码系统框图

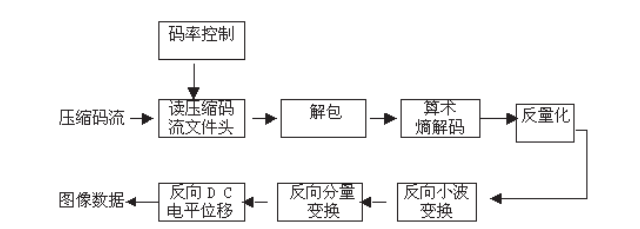


图2 JPEG2000 解码系统框图

3 地震数据压缩

为了针对地震数据的特性进行较大比率的压缩而采用混和编码方案,其技术思路是预处理后的地震数据先进行变换编码,最大限度地消除数据空间的相关性,再进行常规压缩,采用适当的量化和熵编码,进一步压缩。国际比较流行的混和编码方法采用小波变换+线性预测+Huffman 编码的混和编码方法,取得了良好的压缩效果^[12]。但算法复杂,设计困难,影响了数据处理的速度且无法满足网络传输的要求。为了获得更好的实际应用效果,必须选择更快速更先进的算法去完成数据压缩的过程。采用 JPEG2000 标准压缩地震数据可以达到理想的效果。

3.1 系统结构

图3给出了地震数据压缩系统基本结构框图,其中以 JPEG2000 编解码系统为核心。首先读取地震数据并输入到编码缓冲区,然后进行 JPEG2000 编码,输出压缩码流,进行存贮或传输,再进行解压缩过程,和上述压缩过程刚好相反,最后输出重建信号,并对重建信号质量评价。

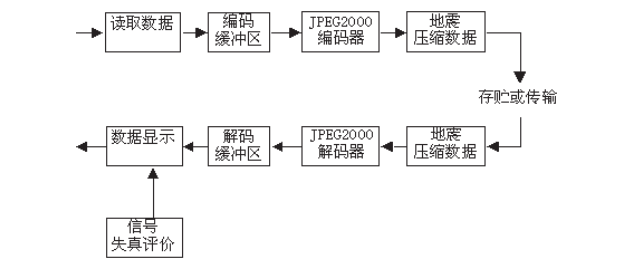


图3 系统基本结构框图

3.2 算法描述

3.2.1 地震数据读取

地震数据格式多样,最常用的格式是 SEG-Y 格式,SEG-Y 格式的地震数据结构复杂,其中卷头信息(REEL HEAD)和道头信息(TRACK HEAD)数据达到 280bytes,涉及到的参数多

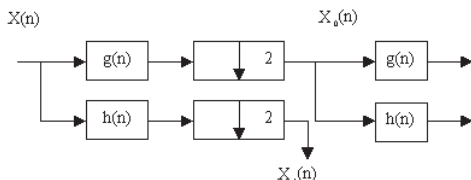
达 113 个,其中大部分参数直接影响对地震数据压缩算法的设计,对于数据正确的解读、分析是保证数据压缩质量的关键。

3.2.2 数据压缩

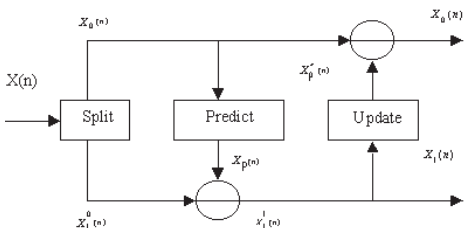
(1)快速提升小波变换算法

选用 JPEG2000 标准的快速提升小波变换算法来压缩地震数据。该算法运算占用内存小,无论分解或重构都不涉及复杂的内积运算,简单快速,小波分解后的系数全为整数,正、逆变换简单,量化误差小,可以实现无损压缩。

快速提升小波变换算法的基本思想是将 Mallat 算法中的每一级滤波运算分解为分裂(split)、预测(predict)和更新(update)3 个过程,如图4所示。



(a) Mallat 算法



(b) Lifting 算法

图4 Mallat 算法与 Lifting 算法关系示意图

其中,分裂是将输入信号根据其序数的奇偶性分为以下2

组:

$$\begin{cases} X_0^0(n) = x(2n) \\ X_1^0(n) = x(2n+1) \end{cases}$$

预测是将滤波器 p 作用于偶信号 $X_0^0(n)$

得到奇信号的预测值 $X_1^0(n)$,再将该预测值与奇信号 $X_1^0(n)$ 相减以得到奇信号的预测误差 $X_1^1(n)$,因此:

$$\begin{cases} X_1^0(n) = \sum_k p_k X_0^0(n-k) \\ X_1^1(n) = X_1^0(n) - X_1^0(n) \end{cases}$$

更新是将滤波器 u 作用于奇信号的预测误差 $X_1^1(n)$,得到偶信号 $X_0^0(n)$ 的预测值 $X_0^p(n)$,然后用来对偶信号进行校正。

$$\begin{cases} X_0^p(n) = \sum_k u_k X_1^1(n-k) \\ X_0^1(n) = X_0^0(n) + X_0^p(n) \end{cases}$$

式(1)与(2)中 k 的取值范围分别对应于滤波器 P 和 U 的阶数。

Mallat 算法任一组小波滤波器($h(n)$, $g(n)$)均可以通过因式分解得到相应的一组或多组预测滤波器(P)和更新滤波器(U),并可在分裂以后进行多次的预测和更新操作。

图4(b)仅表示了一组滤波器,其最后结果为:

$$\begin{cases} X_0(n) = X_0^1(n) \\ X_1(n) = X_1^1(n) \end{cases}$$

(2)小波基的选取

JPEG2000 标准中,小波基的选取对数据的压缩性能影响很大,小波基的选取要考虑两个因素:正交性和规则性。^[3]地震数据无损压缩部分如道头信息和卷头信息部分,这里采用 Daubenchies(2+2,2)双正交小波滤波器。测试地震数据的无损压缩性能对比实验结果如表 1。地震数据做有损压缩时由于数据水平和垂直方向的相关性并不相同,不宜选用相同的小波基。根据试验测试选取两对双正交小波滤波器 Daubenchies9/7 和 Villasenor13/11 分别用于水平和垂直方向的变换。测试地震数据有损压缩性能对比实验结果如表 2 所示。

表 1 几种常用小波基的地震数据有损压缩率(PSNR=32)

小波滤波器	Daubenchies(3,1)	Daubenchies(4,2)	6/14	S	Daubenchies5/3	Daubenchies(2+2,2)
压缩比(CR)	2.89	2.80	2.98	2.87	2.85	3.89

表 2 几种常用小波基的地震数据无损压缩率

小波滤波器	Villasenor13/11	Daubenchies5/3	Daubenchies9/7	Adelson9	Harr	Daubenchies8
压缩比(水平方向)	36.21	34.58	38.97	38.88	35.82	33.79
压缩比(垂直方向)	33.24	28.77	31.56	29.98	28.56	30.83

(3)地震数据的小波分解

在对地震数据进行提升小波变换时,首先对地震数据的每道进行 7 层小波分解,形成(3×7+1=22)个分频子带,滤掉大于 12.5Hz 的成分和小于 0.39Hz 的成份(即第七层的细节分量)。当地震数据实现了小波的多分辨分解时,并不意味着地震数据已实现了压缩。实际上,小波变换只是给地震数据压缩提供了好的数据表示形式,而存储地震数据所用的空间并没改变;由于远距离测得的地震信号,其能量绝大部分集中在 0.5Hz 到 10Hz 内,而原始采样的地震数据,由于数字化引起的高频混淆,即波谱高频成分增强,造成了对有用信号的干扰,因此,在进一步处理之前,要对原始采样得到的地震信号进行低通滤波。另外,由于地震探测器在记录震动信号时,其记录线与水平线不完全垂直,采样时又把无震时的记录基线作为零线,这将带来低频误差(一般在 0.5Hz 以内),因此,在进一步处理之前,也要对原始采样得到的地震信号进行高通滤波,删除不合理的低频部分。又由于地震信号的进一步处理需要信号的变化趋势及其信号的频谱特性,因而,做任何预处理都应尽量保证信号的波形变化趋势和频谱特性失真小。试验结果表明:滤波后数据量减少 3 倍。

(4)量化

对变换后的系数进行量化是有损压缩中的重要环节。量化的关键是根据变换后的图像的特征,重构图像的质量要求等因素设定合理的量化步长。对地震数据进行了 7 级小波分解后,可以得到 22 个子带。每一个子带系数都反映了图像不同频域的特征,因此对每个子带都采用不同的量化步长进行量化。在地震数据无损压缩部分,这里采用量化步长 1。

(5)熵编码

JPEG2000 基本编码方法是在最佳截断嵌入式码块编码(EBCOT)算法基础上开发的。^[13]EBCOT 编码的主要思想是把嵌入码块编码方法与码块位流的最佳截断力法结合在一起,使重构图像的失真最小。EBCOT 编码算法引入了一个“质量层(Quality Layers)”的概念,图像的最终码块位流以质量层的形式组织,从而使压缩码流具有了渐进传输、随机访问等卓越的特性。

3.2.3 数据重建

地震数据的重建(即解压)为 JPEG2000 编码的逆过程,即算术熵解码、反量化、利用小波逆变换重建地震数据。

3.2.4 数据显示与质量评价

地震数据显示是对地震数据进行分析处理的基础。该系统采用灰度显示和波形显示两种显示方式可供用户根据实际进行选择。

系统的另外一个主要功能就是在压缩/解压的基础上进行质量评估。采用的指标包括:压缩比(CR)、峰值信噪比(PSNR)、压缩前后文件大小、原信号与重建信号的方差(RMSE)等。

图 5 是一实际原始地震数据的局部,图 6 是在 CR=82, PSNR=34 下的压缩重建结果。可以看出重建信号较好地保留了地质信息,显示了良好的解压缩效果。

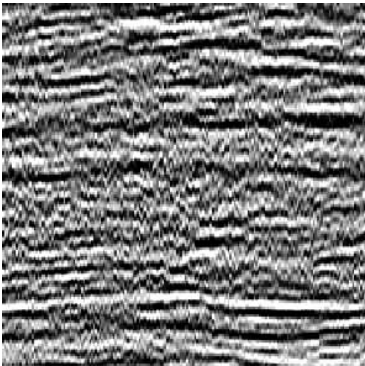


图 5 原始地震数据剖面图

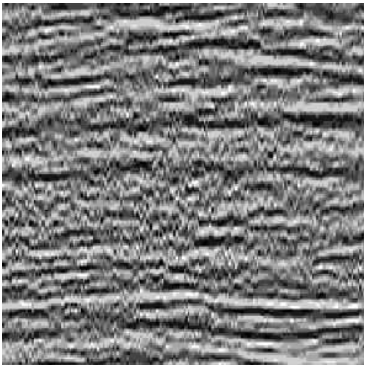


图 6 原始地震数据压缩重建剖面图

为了分析 JPEG2000 对地震数据的压缩效率,这里选取了某油田实际地震数据进行了测试。表 3 给出了压缩实验结果,表 4 是传统小波变换压缩实验结果。

表 3 基于 JPEG2000 的地震数据压缩实验结果

压缩比(CR)	PSNR	RMSE
97.838	40.7942	0.29172
69.626	40.94776	0.22073
58.453	41.59861	0.17609
49.655	42.24995	0.13764
43.396	42.84428	0.11725
33.292	44.08511	0.12574
30.458	44.7851	0.10256

实验结果表明,基于 JPEG2000 地震数据编码方法较传统压缩方法压缩性能有了很大的提高,不仅可以取得较大的压缩比,而且在相同压缩率下地震数据质量更高。同时 JPEG2000

更大的优势在于它是一种小波变换编码,算法快速,且 JPEG2000 压缩后的数据更利于累进传输、随机访问等。

为了进一步分析 JPEG2000 对于地震数据压缩的有效性,该文测试了 50 幅地震数据剖面图,在保证压缩质量的前提下,获得了平均 44.19 的压缩率。实验发现地层结构对地震数据的压缩率影响最大,地层结构越简单压缩率越高,反之越小。对于地层平坦,构造简单的图幅可以获得 200 甚至更高的压缩率。为了获得普遍的统计分析,本实验排除了压缩率大于 100 小于 10 地震数据图幅。实验结果如图 7 所示。

表 4 基于传统小波变换地震数据压缩实验结果

压缩比(CR)	PSNR	RMSE
62.774	37.26237	0.41379
40.478	36.71341	0.29626
31.081	37.82132	0.22956
20.586	38.86485	0.18052
19.341	39.4621	0.15733
16.742	40.01729	0.13845
12.992	42.21221	0.10515

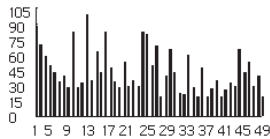


图 7 50 幅地震数据剖面的压缩率

这一实验结果表明,基于 JPEG2000 混和编码方法对地震数据压缩是十分有效的。

4 结论

基于 JPEG2000 标准压缩地震数据,针对地震数据的特性,采用合理的编码方案可以达到理想的压缩效果。实验表明,该文设计的编码方法十分有效,与传统的方法相比较,它至少具有下面三个方面的优点 (1) JPEG2000 标准可以大大改善地震数据压缩性能 (2) JPEG2000 的提升小波算法是一种整数小

波,算法简单、运算快速、易于软硬件实现 (3) JPEG2000 具有渐进传输等特性,能很好地满足网络环境下的地震数据处理需求。因此,对于地震数据的存储或远程传输问题, JPEG2000 提供了很好的解决途径,具有一定的实用价值。

(收稿日期:2004 年 1 月)

参考文献

1.赵爱华.基于提升小波的自适应无损图像压缩.2004
2.李飞鹏,梅天灿,秦前清.雷达数据的准无损压缩[J].武汉:武汉大学学报(理学版),2004
3.于超,王志军,冯坚.JPEG2000 中不同小波基的图像压缩性能分析[J].小型微型计算机系统,2003
4.周宁,汤晓军,徐维朴.JPEG2000 图像压缩标准及其关键算法[J].现代电子技术,2002
5.JPEG2000 Image Coding system.JPEG2000 Final Committee Draft Version 1.0.2000
6.邓集锋.第二代小波变换及其遥感图象应用[J].中国图象图形学报,1998
7.张军华,仝兆岐.用小波变换法定量压缩地震数据[J].东营:石油大学学报,2003
8.唐向宏,何振华,黄德济.利用二维 M 带小波变换进行地震数据压缩[J].石油地球物理勘探,1999
9.Amir Z Averbuch ,F Meyer ,J-O Strömberg et al.Low Bit-Rate Efficient Compression for Seismic Data[J].IEEE Transactions on Image Processing,2001 ;10(12)
10.J D Villasenor ,R A Ergas ,P L Donoho.Seismic Data Compression Using High-Dimensional Wavelet Transforms.IEEE 1068-0314/96 ,1996
11.H Guo ,C S Burrus.Phase-preserving Compression of Seismic Data using the Self-adjusting Wavelet Transform[R].Technical report ,Rice University Houston ,1996
12.Giridhar Mandyam ,Neeraj Magotra ,Wes McCoy.Lossless Seismic Data Compression using Adaptive Linear Prediction.IEEE 0-7803-3068-4/96 ,1996
13.D Taubman.High Performance Scalable Image Compression With EBCOT[J].IEEE Trans Image Processing 2000 9(7) :1158~1170

(上接 115 页)

表 2 转发引擎转发处理能力测试结果

报文长度	40	60	64	65	128	256	512	1024	1280	1500
分组转发率	24450516	17332584	16364012	15793544	8703472	4487556	2283572	1151276	922556	787992
丢包率	0	0	0	0	0	0	0	0	0	0
分组转发时延		3.2475	2.97	2.15	3.39	2.98	4.2875	6.435	7.4425	8.195

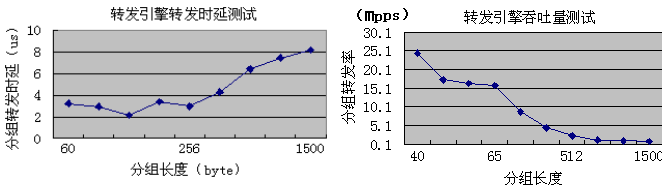


图 6 转发引擎吞吐量和时延测试

4 结论

线路接口和转发引擎之间、转发引擎和交换网络一般都采用同轴缆或者光缆的柔性连接,其信号传输格式多采用 8B10B 编码,因此可以将超高速转发引擎设计成线速转发能力由最高速率向下兼容的 ASIC,从而研究和设计我国具有自主知识产权的超高速转发引擎,全面提高我国核心节点设备的竞争力。

(收稿日期:2005 年 1 月)

参考文献

1.许恪,熊勇强,吴建平.宽带 IP 路由器的体系结构分析[J].软件学报,2000 (3) :1~8
2.R Hinden.IP Version 6 Addressing Architecture[S].RFC 3513,2003
3.S Deering.Internet Protocol ,Version 6 (IPv6)Specification[S].RFC2460 ,1998
4.D Estrin.Protocol Independent Multicast-Sparse Mode(PIM-SM) : Protocol Specification[S].RFC 2362,1998
5.A Conta.Internet Control Message Protocol(ICMPv6) for the Internet Protocol Version 6 (IPv6)Specification[S].RFC 2463,1998
6.白建军,卢泽新.路由器原理与设计[M].人民邮电出版社,2002