

卫星海洋遥感数据的无损压缩及解压算法

付东洋, 李宇隆, 侯骏雄, 肖志强

(广东海洋大学信息学院 海洋遥感与信息技术实验室, 广东 广州 524088)

摘要:在分析卫星海洋遥感单轨三级数据产品特征的基础上, 针对游程编码无损压缩算法特征, 提出了优化游程编码压缩算法, 运用该算法实现了海洋遥感三级数据产品的无损压缩, 并与流行的 WinRAR 压缩软件进行了比较测试, 结果表明上述算法对海洋遥感数据均有较高的压缩率, 优化游程编码压缩算法在空间复杂度及时间复杂度上均有明显的比较优势, 可极大节约海洋遥感数据的存储与共享发布空间。

关键词:海洋遥感数据; 无损压缩; 解压缩; 优化游程编码; 压缩比

中图分类号: TP751.1 **文献标志码:** A **文章编号:** 1003-2029(2012)02-0045-04

潘德炉^[1]指出, 随着卫星遥感技术向高时间分辨率、高空间分辨率和高辐射分辨率方向的快速发展, 空间遥感与信息技术已经发展成为满足人类对海洋资源和环境不同尺度和不同层次连续、动态的信息需求的必要手段, 海洋光学遥感在实施大范围海面瞬间信息监测、数年至几十年长序列全球海洋数据采集等方面, 表现出了极大的优势。但是, 在遥感数字化图像分发与处理中, 普通城市的高分辨率遥感镶嵌图可达几十 GB 的数据量, 超大的数据量给卫星链路的带宽、图像数据的传输、存储、共享、浏览和网络发布带来了时间和网络资源上的巨大压力。以我校海洋遥感与信息技术实验室为例, 该室每天可接收和处理 9 颗卫星(如 MODIS 系列、NOAA 系列、FY、HY、MISAT)资料, 包括原始数据以及 1 级、2 级、3 级等各种产品的数据和图像, 同时为保证数据资源的长期连续性和可靠性, 这些图像及数据均需实施双备份, 这样每天的数据量可超过 30 GB, 因此, 如何将海量的遥感数据进行有效的压缩与存储已成为人们日益关注的问题。赫华颖^[2]比较了基于不同小波基的压缩算法对不同地貌结构遥感图像压缩效率的高低, 并发现 D(9, 7)小波基的应用效果最好; 李强^[3]根据遥感图像局部相关性较弱、纹理复杂丰富的特点, 提出了基于小波分析理论的自适应标量、矢量混合量化压缩方法, 实现了遥感图像纹理的高保真压缩; 周孝宽^[4]基于空间重采样压缩算法实现遥感图像的低运算复杂度的压缩。上述压缩算法的共同特点是适合于陆地遥感影像的有损压缩, 遥感数据属于科学数据, 应尽可能地保证数据的真实性与可靠性, 因而无损压缩对于数据的存储与发布具有十分重要的意义。张晓玲^[5]提出一种基于感知器模型的遥感图像无损压缩

编码方法, 其压缩效果明显优于无损 JPEG 的最优预测模型; 耿则勋^[6]提出的水平与垂直预测相结合的无损压缩方法可使压缩比在 2 倍左右。但这些压缩算法基本上针对陆地航空或卫星遥感领域, 很少有学者涉及海洋遥感图像数据的压缩存储与发布问题的研究, 本文针对海量的卫星海洋遥感图像数据在存储与发布中遇到的问题, 并根据我国自主研究的海洋遥感数据产品的特点, 比较研究了不同压缩算法对海表温度、叶绿素 a 浓度等不同遥感产品在压缩比方面的差异, 提出的优化游程编码压缩算法对于海洋遥感 L3A 产品压缩具有较大的优势。

1 卫星海洋遥感产品处理流程及三级产品特征

1.1 卫星海洋遥感图像数据处理流程

海洋遥感数据接收与处理系统的一般流程如图 1 所示, 其单轨 L3A 级水色水温产品具有以下特征:

(1) 其空间分辨率较陆地遥感而言小得多, 目前可达 1 km, 但通常每轨影像可覆盖 20~30 个经纬度的空间范围, 相应的数据文件可达 600 万个像素以上, 因而数据量巨大。

(2) 海洋遥感研究的是海洋科学领域的问题, 通常是海表光学、水温及水色特征的反应, 因而不涉及陆地部分的信息, 一般在数据处理过程中, 将陆地信息用 0 掩模掉。

(3) 由于受天气(云、雨、雾、风沙)等原因的影响, 很多时候数据缺失严重, 如图 2(海表温度)、图 3(叶绿素浓度)所示, 可实现图像数据的压缩存储、共享及发布。

收稿日期 2011-10-21

基金项目 国家海洋公益专项资助项目(200805028); 广东海洋大学自然科学基金资助项目(1012152); 博时基金资助项目(1112332); 卫星海洋环境动力学国家重点实验室开放基金资助项目(SOED1202); 中国科学院南海海洋研究所热带海洋环境国家重点实验室开放基金资助项目(LTO1105)

作者简介 付东洋(1969-), 男, 博士, 副教授, 研究方向为海洋遥感。Email fdy163@163.com

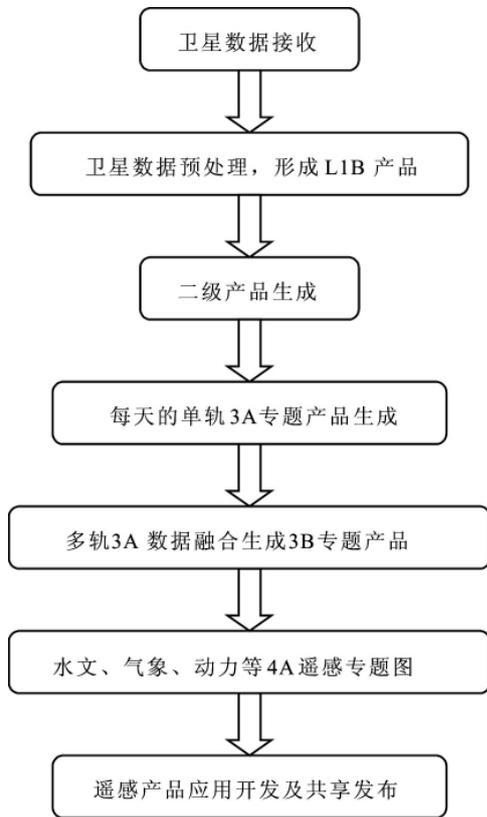


图1 卫星海洋遥感数据接收与处理流程

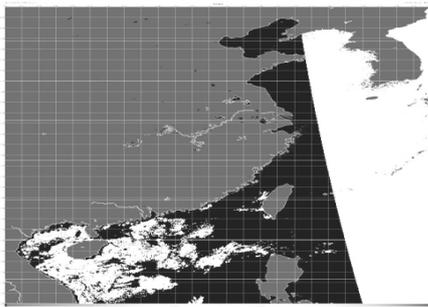


图2 2010年2月3日遥感海表温度—NOAA

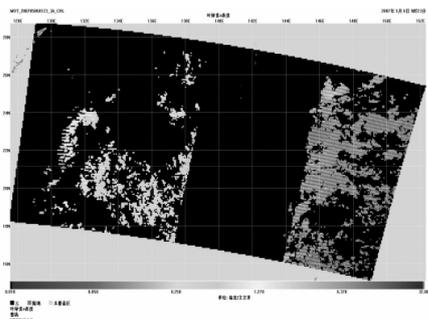


图3 2007年5月6日遥感叶绿素浓度—MODIS

(4) 海表叶绿素浓度、透明度、悬浮物含量等水色要素均为非负值,对我国及东南亚海域,海表温度一般亦为非负值,即有效数据通常具有非负特征。

2 卫星海洋遥感数据的优化游程编码无损压缩及解压算法

游程编码(RLE)是压缩文件最简单的方法之一。它是把一系列的重复值用一个单独的值再加上一个计数值来取代。该算法对于具有长重复值的串的压缩很有效。其算法流程图如图4a所示,该算法压缩比主要是取决于数据本身的特点,当海洋遥感图像数据中的陆地面积越大,获得的压缩比越高。反之,该算法就显得力不从心,同一行的海洋水色水温数值往往不具有连续性,而连续几行具有相同数值的情况更少。同时,为了将压缩数据与普通数据区分,该算法本身包含有标志性数据,特别是当出现类似1X1Y1Z情况时,压缩后反而会使数据串长度增加一倍,这是一种“病态”情况。由于海洋水色数据非负,且对于东南亚海域,海洋水温数据亦具有非负的特点,所以可取消游程编中区分压缩数据的最高标志位并将非连续重复的数据取反,即变为负数,这样就可以通过正副来区分哪些是压缩数据哪些是非压缩数据。以此达到优化游程编码的目的。而且当出现类似上述XYZ的“病态”情况时,压缩时写入(-X)(-Y)(-Z),还原时判断到负值再取绝对值即还原为XYZ。这样就实现了原算法的优化。优化游程编码(DRLE)压缩算法如图4b所示。

3 测试实验与分析讨论

为了比较验证优化游程编码与常规的三元组压缩算法、Huffman 编码压缩算法、游程编码及常见的压缩工具 WinRAR 在海洋遥感数据文件压缩效率上的优劣程度,我们选取了广东海洋大学海洋遥感与信息技术实验室提供的2011年5月的NOAA(National Oceanic and Atmospheric Administration)系列卫星海表温度(SST)3A数据文件92个及MODIS(The Moderate Resolution Imaging Spectroradiometer)卫星叶绿素a浓度(CHL)数据文件70个,并与经典的WinRAR压缩工具进行了对比实验,实验及测试结果如表1所示,相应的92个海表温度文件(原始大小为1045167830字节)的压缩情况如图6所示,70个叶绿素a浓度(原始大小为410593008字节)压缩文件结果如图5所示。

3.1 实验结果

(1) 与WinRAR压缩工具一样,不论是三元组、哈夫曼编码、游程编码还是本文改进的优化游程编码压缩算法均能高效地压缩海洋遥感三级产品数据,且解压结果均可完整正确复原。

(2) 对于70个叶绿素浓度数据文件,就空间复杂度压缩率最高的是WinRAR,其压缩率达4.11%,其次是哈夫曼编码,优化游程编码也接近WinRAR算法,效果相对较差的是三元组压缩算法,但从时间复杂度来看,最佳的是游程压缩和优化游程压缩,用时不到三元组算法及WinRAR算法的1/4。

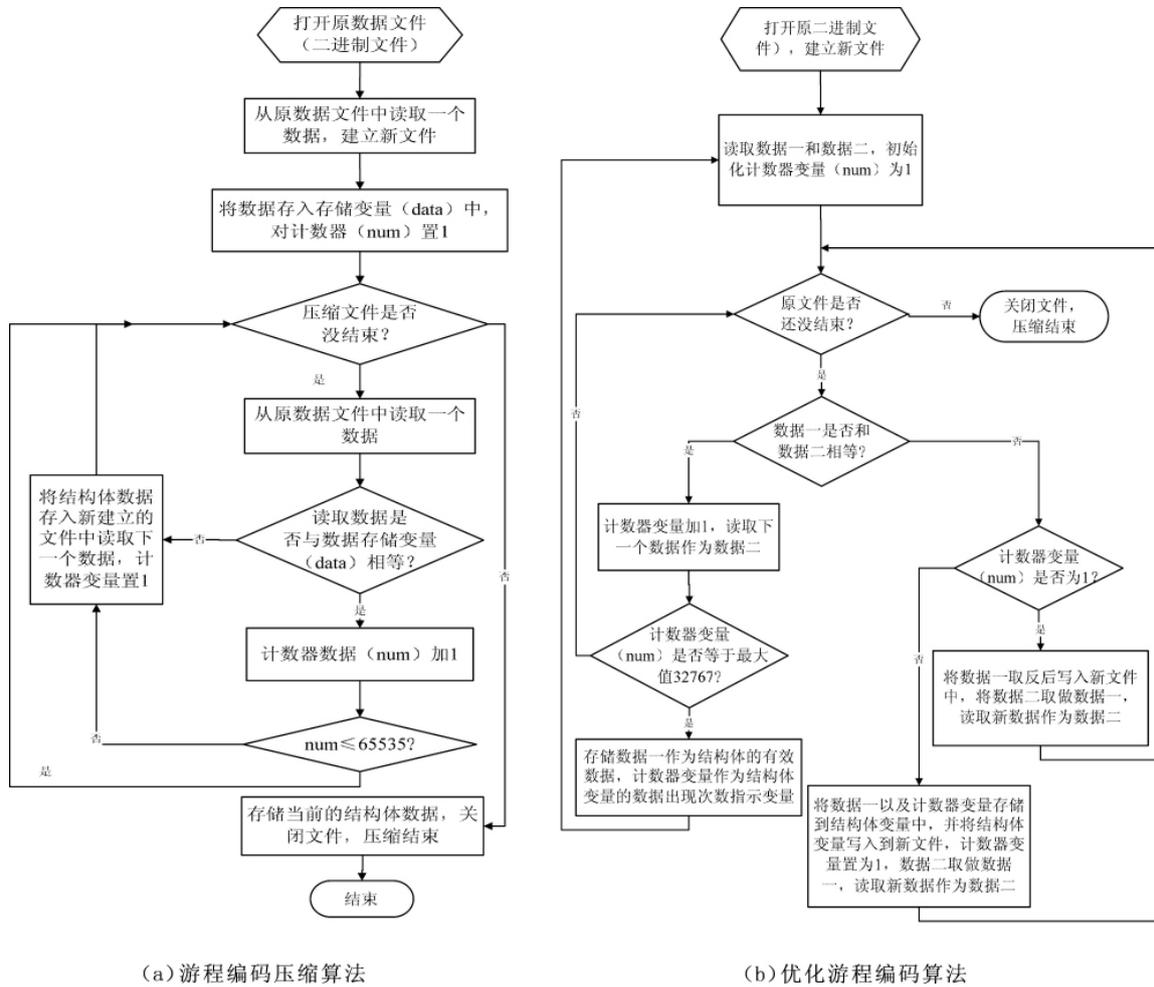


图 4 压缩算法流程

表 1 海洋遥感之海表温度及叶绿素 a 浓度 L3A 数据压缩统计表

| 压缩算法 | 压缩文件大小 70CHL/byte | 压缩率 /% | 压缩时间 /s | 压缩文件大小 92 SST/byte | 压缩率 /% | 压缩时间 /s |
|--------|----------------------|-----------|------------|-----------------------|-----------|------------|
| 三元组压缩 | 68 444 660 | 16.67 | 8.83 | 136 789 586 | 13.09 | 20.02 |
| 哈夫曼压缩 | 17 739 310 | 4.32 | 39.85 | 160 727 137 | 15.38 | 138.69 |
| 游程压缩 | 42 425 056 | 10.33 | 7.90 | 57 910 638 | 5.54 | 19.14 |
| 优化游程压缩 | 25 497 534 | 5.39 | 8.09 | 37 219 520 | 3.56 | 19.37 |
| WinRAR | 16 874 386 | 4.11 | 36.00 | 24 673 779 | 2.36 | 60.00 |

(3) 对于 92 个海表温度数据文件,就空间复杂度,压缩率最高的是 WinRAR 工具,其压缩率达 2.36%,其次是优化游程编码,达 3.56%,非常接近 WinRAR 算法,效果相对较差的是三元组和游程编码压缩算法,但从时间复杂度来看,最佳的是游程编码,优化游程编码和三元组算法,它们均约 WinRAR 算法的 1/3,是哈夫曼编码算法的 1/6。

3.2 分析讨论

(1) 从评价压缩算法的空间复杂度来看,对于水色水温数据,WinRAR 压缩工具都最好的,最接近该方法的是本文

改进的优化游程编码。

(2) 从评价压缩算法的时间复杂度来看,对于水色水温数据,游程编码和优化游程编码是最好,而 WinRAR 压缩工具都在该两算法的 3 倍时长以上,时间复杂度也最高,该算法不适宜于实时性要求较高的场合。

(3) 综合比较 5 种算法,不论在空间复杂度还是时间复杂度上,本文提出的针对海洋遥感数据文件的优化游程编码算法都有明显的比较优势。

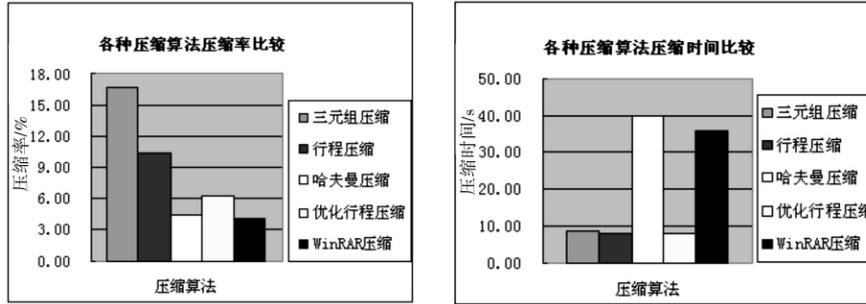


图 5 5 种不同压缩算法对 70 个叶绿素浓度数据文件压缩结果比较

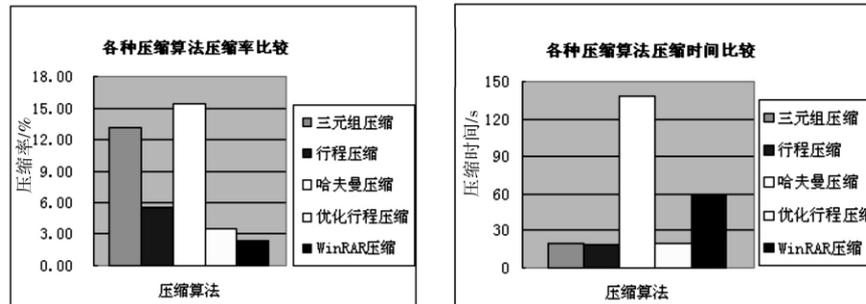


图 6 5 种不同压缩算法对 92 个海表温度数据文件压缩结果比较

参考文献：

- [1] 潘德炉,王迪峰.我国海洋光学遥感应用科学研究的新进展[J].地球科学进展,2004,19(4):506-512.
- [2] 赫华颖,陆书宁.几种小波基在遥感图像压缩中的应用效果比较[J].国土资源遥感,2008,77(3) 27-31.
- [3] 李强,王正志.基于小波理论的遥感图像高保真压缩方法研究[J].遥感学报,1999,3(1):31-37.
- [4] 周孝宽,程子敬.基于空间重采样的遥感图像压缩[J].北京航空航天大学学报,2001,27(5):611-614.
- [5] 张晓玲,毋立芳.基于感知器的遥感图像无损压缩编码[J].电子与信息学报,2001,23(7):712-715.
- [6] 耿则勋,钱曾波.遥感影像无损压缩编码的实验分析与改进[J].测绘学报,1996,25(4):262-265.

Lossless Compression and Decompression Algorithms for Satellite Ocean Remote Sensing Data

FU Dong-yang, LI Yu-long, HOU Jun-xiong, XIAO Zhi-qiang

(Lab of Ocean Remote Sensing & Information Technology, Information College, Guangdong Ocean University, Zhanjiang Guangdong 524088, China)

Abstract: Based on analysis of the satellite ocean remote sensing L3A data products and the features of RLE (run-length encoding). An optimization algorithm to run-length encoding (ORLE) compression was put forward, which was used for lossless compression to ocean remote sensing L3A data. The lossless compression effectiveness has been tested and compared with WinRAR. The results show that all the algorithms for ocean remote sensing data have a higher compression ratio, ORLE algorithm has advantages in the space and time complexity. A significant reduction of storage, sharing and publishing space in ocean remote sensing scientific data is realized.

Key words: ocean remote sensing data; lossless compression; decompression; optimization run-length encoding; compression ratio