

基于搜索最优双预测波段的超光谱遥感图像无损压缩

霍承富 张荣

(中国科学技术大学电子工程与信息科学系 合肥 230027)

摘要: 该文针对超光谱各波段成像图像之间的相关性强弱程度互不相同这一特点,提出了基于搜索最优双预测波段的超光谱图像无损压缩算法。该算法通过建立一个搜索最优双预测波段的二叉树模型,搜索与每一波段相关性最强的两个波段,并用这两个波段对该波段进行预测。实验表明,与目前研究的其它超光谱图像压缩算法相比,该算法有着很好的压缩性能。

关键词: 超光谱图像; 最优双预测波段; 二叉树

中图分类号: TP751.1

文献标识码: A

文章编号: 1009-5896(2009)05-1144-04

Lossless Compression Based on Searching the Optimal Couple Prediction Wavebands for Hyperspectral Image

Huo Cheng-fu Zhang Rong

(Department of Electronic Engineering and Information Science,
University of Science and Technology of China, Hefei 230027, China)

Abstract: The relativity between hyperspectral images of different waveband is differ from each other, contraposing this characteristic, this paper presents a lossless compression algorithm of hyperspectral image which is based on searching the optimal couple prediction wavebands. Through a searching model which is built by consulting the binary tree, two prediction wavebands are found, which have the most relativity to each basal waveband, and then the basal waveband is predicted using the two prediction wavebands. Experimental reault shows that, the compression algorithm has a primely compression performace compared with other compression algorithms present recently.

Key words: Hyperspectral image; Optimal couple prediction wavebands; Binary tree

1 引言

超光谱图像是同一地物在上百个电磁波谱段上的成像,它在普通二维图像的基础上增加了一维光谱信息。光谱维揭示了每一像素的光谱曲线特征,实现了遥感数据图像维与光谱维的有机结合,使得超光谱图像所含有的信息丰富,有利于地物的识别和分类。然而,超光谱图像数据海量,对传输、存储和处理带来了极大的负担。因此,研究超光谱图像的有效压缩算法,对推广超光谱图像的应用有着十分重要的意义。

超光谱图像可以视为一种三维数据,它既具有谱间相关性又具有空间相关性,压缩超光谱图像时应充分去除谱间冗余和空间冗余,使得压缩效率尽可能地提高。常用的算法有基于三维小波变换的超光谱图像压缩^[1],基于矢量量化的超光谱图像压缩^[2],以及基于预测的超光谱图像压缩^[3-5]。其中基于预测的压缩是目前权衡计算复杂度和压缩性能最好的算法,也是人们研究最多的算法。LPVQ^[6]是目前压缩性能最优的基于矢量量化的超光谱图像压缩算法,然而由于它

所需的计算量大,不适合运用于实时性要求较强的场合。SLSQ^[7]算法利用线性预测去除超光谱图像的谱间冗余,并对预测残差进行熵编码,其压缩性能相比于 JPEG2000 等压缩标准有着一定的提高。M-CALIC^[8]算法通过构建精确的谱间预测模型和预测系数以实现很好的压缩性能。CAAP^[9]是一种基于上下文条件平均的快速预测算法,能够很好地实现实时处理。LUT^[10]是一种基于查找表的超光谱遥感图像无损压缩算法,其基本思想与预测树算法相似,通过建立查找表以达到简化计算复杂性的目的。

上述提到的各种算法大多是基于单波段预测的,它们并没有充分利用双波段或多波段预测带来的优势。本文在分析了超光谱图像谱间和空间相关性特点的基础上,提出了基于搜索最优双预测波段的超光谱图像无损压缩算法。该算法的基本思想是:通过建立一个搜索最优双预测波段的二叉树模型,搜索与每一波段相关性最强的两个波段,并用搜索到的这两个波段对该波段进行预测。对于得到的预测残差图像,采用 JPEG-LS 压缩标准^[11]进行编码。

2 超光谱图像相关特性分析

分析超光谱图像可知,每一波段图像的相邻像素之间存

在着空间相关性，不同波段图像之间具有很强的结构相似性。图 1 为 AVIRIS 拍摄的 Cuprite 场景分别在 1 波段，120 波段和 220 波段处的成像图像，从图中可以看出，各波段之间具有极强的结构相似性，存在着很大的谱间冗余。

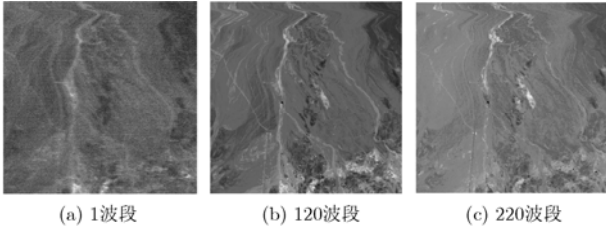


图 1 Cuprite 场景各波段遥感图像(相似性强)

然而，由于大气和成像地物对不同波段电磁波的吸收和反射能力不同，使得不同波段成像图像的像素值有很大差异，而且各波段图像之间的相关性强弱程度也不一样。图 2 为 AVIRIS 拍摄的 Cuprite 场景分别在 150 波段，156 波段和 161 波段处的遥感图像。

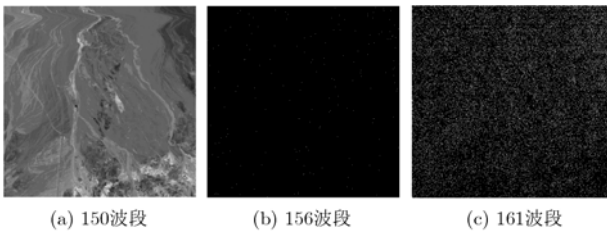


图 2 Cuprite 场景各波段遥感图像(相似性弱)

从图 2 可以知道，由于大气对有些波段的衰减极为严重，使得噪声成为了这些波段成像图像的主要成分。此时，无论是图像的空间相关性还是谱间相关性都很差，对这些波段图像如果还采用统一的预测处理，将不会达到很好的效果。因此，为了对超光谱图像的压缩达到一个理想的效果，设计的压缩算法应该能够根据不同波段图像本身的一些特征，自适应地选择不同的处理方式。

3 压缩算法

本文提出的基于搜索最优双预测波段的超光谱图像无损压缩算法分 4 个步骤：(1)提出一种二叉树搜索模型，并基于这一模型，搜索到与基波段(子基波段)相关性最强的两个波段，同时计算它们之间的相关系数。其中，基波段是指每一棵二叉树根节点对应的波段，子基波段是指每一棵二叉树非叶子节点对应的波段；(2)利用搜索到的这两个最相关波段按照最小均方误差准则预测基波段(子基波段)，以去除谱间预测；(3)对预测残差图像利用 JPEG-LS 标准进行压缩，进一步去除冗余；(4)对于没有成为二叉树节点的各孤立波段以及二叉树叶子节点所对应的各波段，算法直接采用 JPEG-LS

标准对其进行压缩，以去除空间冗余。

3.1 二叉树搜索模型

为了实现很好的预测效果，本文提出超光谱图像的最优双波段预测搜索模型，图 3 为基于完全二叉树结构的搜索模型。根据图 3(a)所示的完全二叉树基搜索模型，搜索步骤如下所述：

(1)从所有波段中选择一个作为基波段(A)，并设定阈值 T；

(2)按式(1)计算波段 A 与其余波段的相关系数；

$$\text{Cor}_{r,c} = \frac{\sum_{i=1}^M \sum_{j=1}^N (x_{i,j,r} - \bar{x}_r)(x_{i,j,c} - \bar{x}_c)}{\sqrt{\sum_{i=1}^M \sum_{j=1}^N (x_{i,j,r} - \bar{x}_r)^2 \sum_{i=1}^M \sum_{j=1}^N (x_{i,j,c} - \bar{x}_c)^2}} \quad (1)$$

式中 $x_{i,j,r}$ 和 $x_{i,j,c}$ 分别表示预测参考波段图像和当前波段图像中 (i, j) 位置处的像素值， \bar{x}_r 和 \bar{x}_c 分别表示预测参考波段图像和当前波段图像的像素均值，M 和 N 分别表示每一波段图像的长和宽。

(3)根据式(2)和式(3)，选择最大的两个相关系数，并比较它们与阈值 T 的大小；

$$\text{Cor}_{r1,c} = \left\{ \text{Cor}_{r,c} \mid r1 = \arg \left\{ \max_{r=1,2,\dots,s} (\text{Cor}_{r,c}) \right\} \right\} \quad (2)$$

$$\text{Cor}_{r2,c} = \left\{ \text{Cor}_{r,c} \mid r2 = \arg \left\{ \max_{\substack{r=1,2,\dots,s \\ r \neq r1}} (\text{Cor}_{r,c}) \right\} \right\} \quad (3)$$

式(2)和式(3)中 s 为当前的波段数。假设 $\text{Cor}_{r1,c}$ 和 $\text{Cor}_{r2,c}$ 对应的两个波段分别是 B 和 C，如果 $\text{Cor}_{r1,c}$ 和 $\text{Cor}_{r2,c}$ 都大于阈值 T，则分别以 B 和 C 为子基波段，重复步骤(2)和步骤(3)；否则，二叉树停止生长，搜索过程结束，并从剩余的波段中选择一个作为新的基波段，重复步骤(1)，步骤(2)和步骤(3)。当搜索不到两个波段与所选择的一个新基波段的相关系数大于 T 时，则认为该基波段为孤立波段，直接采用 JPEG-LS 模式进行压缩。

(4)当以基波段 A 为根的一棵树生成后，对所有叶子节点所对应的波段按图 3(b)所示的子搜索模型进行处理；

分析图 3(a)所示的基搜索模型可知，如果对生成的二叉树不再作后续处理，那么只有非叶子节点所对应的各波段可以通过双波段预测得到很好的预测效果，而大量的叶子节点所对应的各波段只能采用 JPEG-LS 模式直接压缩，这样会

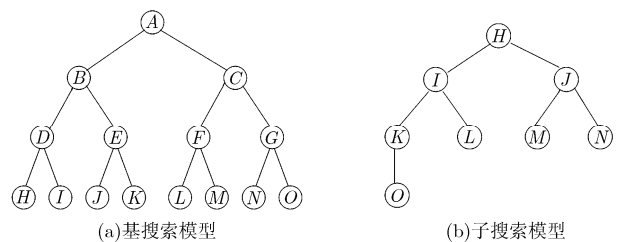


图 3 完全二叉树搜索模型

严重影响算法整体的压缩性能。为了充分利用超光谱图像的谱间相关性，我们将所有叶子节点视为一个子波段集合，在这个子波段集合中选择一个作为基波段，并按图3(b)所示的子搜索模型重复步骤(1)，步骤(2)，步骤(3)和步骤(4)。

算法真正的实现过程是搜索过程的逆过程，即最先得到所有叶子所对应的各波段，并对其进行编码，再按二叉树结构每两个子节点预测一个父节点，得到一幅对应的残差图像，最后对残差图像按 JPEG-LS 模式压缩。

仔细分析图3(a)所示的完全二叉树基搜索模型会发现，随着二叉树深度的增加，叶子节点所对应的各波段之间的相关性将越来越弱，这导致运用图3(b)所示的完全二叉树子搜索模型处理所有叶子节点构成的子波段集合时，并不能达到很好的预测效果。为了克服这一缺点，我们对搜索模型进行了改进，即如图4所示的非完全二叉树搜索模型。

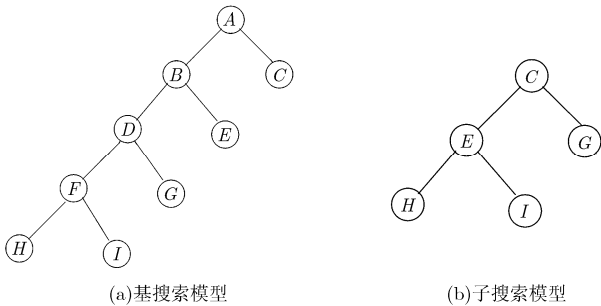


图4 非完全二叉树搜索模型

按图4(a)所示的非完全二叉树基搜索模型，对上述搜索过程稍加改进，即只选择每一个基波段(子基波段)所对应的两个子节点中的一个作为后续搜索的父节点(子基波段)。经如此改进，各叶子节点之间的相关性并不会随着二叉树深度的增加而明显减弱。这样，运用图4(b)所示的子搜索模型处理所有叶子节点所对应的波段时就可以达到更好的预测效果。

3.2 双波段预测

分析搜索模型可以知道，对于所有非叶子节点对应的各波段，算法实现的是双波段预测，它可以达到比单波段预测更好的预测效果。

假设 x_c 表示当前被预测波段， x_{r1} 和 x_{r2} 分别是 x_c 的两个子节点所对应的预测波段，则 x_c 中任意一点的像素 $x_{i,j,c}$ 可以由式(4)所示的最优线性预测^[12,13]公式进行预测，预测残差可以由式(5)计算得到。

$$\hat{x}_{i,j,c} = \alpha x_{i,j,r1} + \beta x_{i,j,r2} \tag{4}$$

$$\varepsilon_{i,j} = x_{i,j,c} - \hat{x}_{i,j,c} \tag{5}$$

式中 α 和 β 为预测系数。求得每一个像素点的预测残差后，整幅图像的预测均方误差可由式(6)计算得到。

$$\varepsilon^2 = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N \varepsilon_{i,j}^2 \tag{6}$$

为使 ε^2 的值最小，将 ε^2 分别对 α 和 β 求偏导的结果须都为 0，如此便可求解出 α 和 β ，求解结果如式(7)所示。

$$\left. \begin{aligned} \alpha &= \left[\sum_{i=1}^M \sum_{j=1}^N x_{i,j,c} x_{i,j,r2} \sum_{i=1}^M \sum_{j=1}^N x_{i,j,r1} x_{i,j,r2} - \sum_{i=1}^M \sum_{j=1}^N x_{i,j,c} x_{i,j,r1} \right. \\ &\quad \left. \cdot \sum_{i=1}^M \sum_{j=1}^N (x_{i,j,r2})^2 \right] / \varepsilon' \\ \beta &= \left[\sum_{i=1}^M \sum_{j=1}^N x_{i,j,c} x_{i,j,r1} \sum_{i=1}^M \sum_{j=1}^N x_{i,j,r1} x_{i,j,r2} - \sum_{i=1}^M \sum_{j=1}^N x_{i,j,c} x_{i,j,r2} \right. \\ &\quad \left. \cdot \sum_{i=1}^M \sum_{j=1}^N (x_{i,j,r1})^2 \right] / \varepsilon' \end{aligned} \right\} \tag{7}$$

$$\varepsilon' = \left(\sum_{i=1}^M \sum_{j=1}^N x_{i,j,r1} x_{i,j,r2} \right)^2 - \sum_{i=1}^M \sum_{j=1}^N (x_{i,j,r1})^2 \sum_{i=1}^M \sum_{j=1}^N (x_{i,j,r2})^2$$

3.3 JPEG-LS 模式压缩

对于双波段预测残差图像、所有叶子节点所对应的各波段图像以及未构成二叉树的各孤立波段图像，本文统一采用 JPEG-LS 模式进行压缩，这样在减少计算量的同时，也可以进一步提高压缩效果。

JPEG-LS 是 ISO/ITU 新批准的图像无损压缩标准^[14]，该算法具有复杂度低、高无损压缩率的特点。图5是 JPEG-LS 无损编码流程示意图，对于输入的当前像素，算法首先分析其领域信息，然后根据分析结果选择不同的编码方式：游程编码或 Golomb 编码。实验结果表明，JPEG-LS 所采用的核心算法可以获得比自适应算术编码更高的压缩性能。

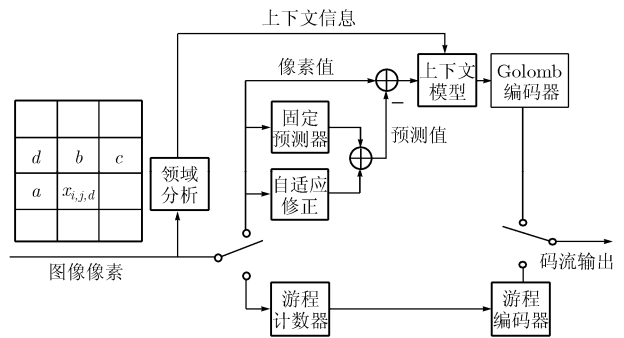


图5 JPEG-LS 编码流程示意图

4 实验结果

采用 AVIRIS 拍摄的 Cuprite 和 Jasper Ridge 场景 224 波段超光谱图像作为实验数据，其波长范围为 380nm 到 2500nm，相邻两波段间的波长间隔为 10nm，每幅图像大小为 512×614 像素，阈值 T 可以根据实验效果选取不同的值，本文选定阈值 $T = 0.92$ 。

由于超光谱遥感图像一般都成像于机载或星载设备，实时性传输是必须考虑的一个重要因素。因此，在对超光谱图像进行压缩处理时必须综合考虑压缩效果和算法的复杂度。分析本文提出的搜索过程可知，从所有剩余波段中找到最优的两个预测波段，所需的计算量很大，这将会影响算法

的实时性。为了避免这一点, 本文对搜索过程作了两点改进: (1)由经验可知, 任意波段与它附近的波段之间的相关性一般强于与其余波段的相关性。基于此, 在搜索两个最优预测波段时, 可以只搜索该基波段(子基波段)前后相邻的 10 个或更少波段即可; (2)在搜索过程中计算每两个波段图像之间相关系数时, 可以不直接对原图像进行计算, 而是先按 4:1 的比例对原图像进行采样, 然后计算采样得到的两图像间的相关系数, 并用它表示原图像间的相关性。这样, 计算量就会明显减少, 同时也不会影响预测效果。通过实验对算法进行仿真, 实验结果如表 1 所示。

表 1 压缩效果比较

| | 压缩比 | | | | | |
|---------|-------|-------|---------|------------|-----------|-------|
| | LPVQ | SLSQ | M-CALIC | CAAP (ACC) | CAAP (GR) | 本文算法 |
| Cuprite | 3.18 | 3.15 | 3.21 | 3.25 | 3.16 | 3.27 |
| Jasper | 2.88 | 3.15 | 3.16 | 3.23 | 3.15 | 3.22 |
| 平均 | 3.030 | 3.150 | 3.185 | 3.240 | 3.155 | 3.245 |

谱间冗余是超光谱图像独有的特征, 为了利用这一特征达到压缩超光谱图像的目的, 人们提出了很多基于波段分类^[3]和波段重排序^[4]的算法, 本文提出的算法综合了各类算法的优点, 通过二叉树模型搜索, 能够同时达到波段分类和波段重排序的效果。从表 1 中可以看到, 本文提出的算法压缩性能明显优于目前提出的大多数压缩算法, 有着很好的实际可操作性以及应用前景。

5 结束语

本文根据超光谱遥感图像自身的特点, 提出了基于搜索最优双预测波段的超光谱图像无损压缩算法。该算法复杂度不大, 压缩性能与其它算法相比有着一定的优势, 是一种切实可行的超光谱图像压缩方案。

值得注意的是, 每一棵二叉树的根节点所对应的基波段的选择会影响算法的整体压缩性能。对于同一组超光谱图像, 选择不同的基波段作为根节点, 压缩效果也会稍有不同。以后的工作可以致力于设计简单有效的算法确定最优基波段(根节点), 以使得压缩性能始终保持最优。

参 考 文 献

- [1] 陈林杰, 刘学斌, 刘庆飞. 基于直方图变换的多光谱 3D SPIHT 压缩编码算法. 光学技术, 2007, 33(1): 137-140.
Chen Ling-je, Liu Xue-bin, and Liu Qing-fei. 3D SPIHT algorithm of the multispectral image based on histogram modify. *Optical Technique*, 2007, 33(1): 137-140.
- [2] Ryan M J and Arnold J F. The lossless compression of AVIRIS images by vector quantization. *IEEE Trans. on Geoscience and Remote Sensing*, 1997, 35(3): 546-550.
- [3] Zhou Zheng, Tan Yihua, and Liu Jian. Satellite hyperspectral

- imagery compression algorithm based on adaptive band regrouping. *Wireless Communications, Networking and Mobile Computing*, Wuhan, China, 2006: 1-4.
- [4] Zhang Jing and Liu Guizhong. An efficient reordering prediction-based lossless compression algorithm for hyperspectral images. *IEEE Geoscience and Remote Sensing Letters*, 2007, 4(2): 283-287.
- [5] Jain S K and Adjeroh D A. Edge-based prediction for lossless compression of hyperspectral images. *Data Compression Conference. Snowbird, Utah, USA, 2007.DCC'07 27-29 March. 2007: 153-162.*
- [6] Motta G, Rizzo F, and Storer J A. Compression of hyperspectral imagery. *Data Compression Conference, 2003, Snowbird, Utah, USA, Proceedings. DCC 2003, 25-27. March 2003: 333-342.*
- [7] Rizzo F, Carpentieri B, Motta G, and Storer J A. High performance compression of hyperspectral imagery with reduced search complexity in the compressed domain. *Data Compression Conference, 2004. Proceedings. DCC 2004. Snowbird, Utah, USA, 2004: 479-488.*
- [8] Magli E, Olmo G, and Quacchio E. Optimized onboard lossless and near lossless compression of hyperspectral data using CALIC. *IEEE Geoscience and Remote Sensing Letters*, 2004, 1(1): 21-25.
- [9] Wang Hongqiang, Babacan S D, and Sayod K. Lossless hyperspectral-image compression using context-based conditional average. *IEEE Trans. on Geoscience and Remote Sensing*, 2007, 45(2): 4187-4193.
- [10] Mielikainen Jarno. Lossless compression of hyperspectral images using lookup tables. *IEEE Signal Processing Letters*, 2006, 13(3): 157-160.
- [11] Weinberger M, Seroussi G, and Sapiro G. The LOCO-I lossless image compression algorithm: Principle and standardization into JPEG-LS. *IEEE Trans. on Image Processing*, 2000, 9(8): 1309-1324.
- [12] 吴铮, 何明一, 冯燕, 贾应彪. 基于误差补偿预测树的多光谱遥感图像无损压缩方法. *遥感学报*, 2005, 9(2): 143-147.
Wu Zheng, He Ming-yi, Feng Yan, and Jia Ying-biao. Lossless compression of multispectral Imagery by error compensated prediction tree. *Journal of Remote Sensing*, 2005, 9(2): 143-147.
- [13] 孙蕾, 罗建书. 基于最佳线性预测器的高光谱图像无损压缩算法. *遥感信息*. 2005, (6): 7-9.
Sun Lei and Luo Jian-shu. Hyperspectral image lossless compression based on optimal linear prediction. *Remote Sensing Information*, 2005, (6): 7-9.
- [14] HP Labs LOCO-I/JPEG-LS Home Page. <http://www.hpl.hp.com/loco>.

霍承富: 男, 1986 年生, 硕士生, 研究方向为超光谱图像压缩。
张 荣: 女, 1968 年生, 博士, 副教授, 研究方向为数字图像处理、计算机图形学、超光谱图像压缩。