

一种空间划分与谱间预测相结合的超光谱图像压缩新方法

高文^{1,2}, 朱明¹, 田润澜³, 汤洋⁴

1. 中国科学院长春光学精密机械与物理研究所, 吉林 长春 130033

2. 中国科学院研究生院, 北京 100049

3. 空军航空大学, 吉林 长春 130000

4. 吉林大学电子科学与工程学院, 吉林 长春 130000

摘要 超光谱图像因其纹理丰富复杂、波段多、光谱分辨率高等特点, 在军事、海洋、农业等多方面都应用广泛, 但也因其数据量大在数据传输时受到很多限制, 研究一种高效超光谱图像压缩方法有着重要的意义。目前光谱图像的压缩方法主要分为基于变换、基于预测和基于矢量量化三大方面, 其中基于预测的方法具有易于实现、压缩比高等优势应用广泛。文章首先对超光谱图像特性进行了详细分析, 对谱间预测算法进行了改进, 并结合超光谱图像谱间相关性高、各谱谱间相关性不完全相同的特点, 提出自适应划分子空间多种谱间预测方法相结合的新方法, 通过实验证明了该方法的有效性。

关键词 超光谱; 图像压缩; 谱间预测; 误差补偿; 子空间划分

中图分类号: TP751 文献标识码: A DOI: 10.3964/j.issn.1000-0593(2011)08-2287-07

引言

超光谱遥感是20世纪80年代兴起的新型对地观测技术, 超光谱图像是利用超光谱传感器对同一对象获得的多个窄光谱波段范围上的图像^[1]。超光谱图像记录了成像地物对多个光谱波段的响应特性, 可有效用于地物类型像素级甚至亚像素级识别, 因纹理丰富、波段多、光谱分辨率高等特点, 其在军事、海洋等多方面均有很高应用价值。但其较大数据量和较高数据维给传输、存储都带来较大困难, 因此图像压缩^[2]是极其必要的。

目前, 超光谱图像的压缩^[3-6] (包括有损压缩和无损压缩) 主要分三大类: 基于变换技术^[7-9]、基于矢量量化^[10,11]和基于预测的方法^[12,13], 其中基于预测的方法不需变换, 直接探索像素间的相关性和波段间的相关性, 特别适合高光谱图像压缩, 得到了广泛研究。1994年由Memon提出的基于预测树模型的多光谱图像压缩方法^[14]是一种有效的超光谱图像无损压缩方法, 随后将其推广到3D-CALID(3D context-based adaptive lossless image codec)。张荣等对预测树的定义做了修改, 提出一种侧邻域最小绝对权值(SNMAW)预测树方法^[15], 大大简化了预测树构造算法, 降低了计算复杂度。此后, 吴铮等结合自适应线性预测方法^[16], 利用谱间的局部统计冗余和结构冗余建立自适应预测器, 提出了误差补偿预

测树方法, 并取得较好的效果。

本文对SNMAW算法进行了改进, 结合超光谱图像谱间相关性, 提出一种自适应划分子空间与多种谱间预测方式相结合的压缩方法, 实验证明本文算法有效、可行。

1 超光谱图像的特性分析

美国JPL实验室研制的AVIRIS(airborne visible and infrared imaging spectrometer)曾测试过超光谱图像^[17], 每一维图像均为 512×614 , 像素深度: 16 bit, 共224波段, 其光谱范围为369.85~2506.81 nm, 光谱分辨率约10 nm, 数据组织形式为波段像间隔(band interleaved by pixel, BIP), 见图1。

对于超光谱图像的特性分析将从空间相关性和谱间相关性两方面进行。

1.1 空间相关性

对于超光谱图像, 空间相关性是指同一个波段内某一像素与其相邻像素之间的相似性。其经过归一化和离散化处理后的图像相关性公式为

$$R(l, k) = \frac{\sum_{x=1}^M \sum_{y=1}^N [f(x, y) - u_f][f(x+k, y+l) - u_f]}{\sum_{x=1}^M \sum_{y=1}^N [f(x, y) - u_f]^2} \quad (1)$$

收稿日期: 2010-06-19, 修订日期: 2010-09-20

基金项目: 国家高技术研究发展计划(863计划)项目(2005AA778032)资助

作者简介: 高文, 女, 1985年生, 中国科学院长春光学精密机械与物理研究所博士研究生 e-mail: cwenzhi@126.com

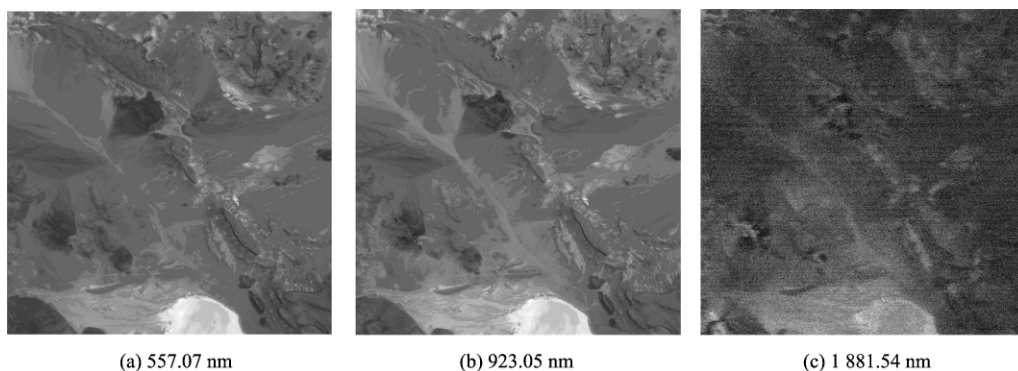


Fig 1 Single-band image of the AVIRIS data

(a): 557.07 nm; (b): 923.05 nm; (c): 1881.54 nm

其中 $M \times N$ 为图像的大小, u_f 为图像灰度的平均值, 即 $u_f =$

$\frac{1}{M \times N} \sum_{x=1}^M \sum_{y=1}^N f(x, y)$ 。图像是二维的信号, 具有行和列二个方向, 当 $l = 0, k = 1$ 时式(1)为列自相关公式; $l = 1, k = 0$ 时, 式(1)为行自相关公式。

从图 2 和图 3 可以看出, 各波段图像自相关函数曲线形状基本相同, 表 1 为算得的超光谱图像部分波段的自相关系数, 可以看出行自相关系数与列自相关系数无明显差异, 最小值均出现在第 160 波段, 最大值均在第 20 波段, 这些都表明行、列相关性基本一致。自相关函数曲线随 k 的增加下降很快, 说明随着像素间隔增大图像的相关性迅速下降。像素与其相邻一、两个像素的相关性较高, 在 0.9 以上; 当间隔大于 8 个像素时, 相关系数已降到 0.7 左右, 表明相关性已经不大。在图像压缩过程中, 应考虑这一点。

Table 1 Auto correlation of several bands of hyperspectral image

波段	行自相关系数	列自相关系数	波段	行自相关系数	列自相关系数
10	0.965 8	0.978 7	110	0.749 1	0.763 4
20	0.969 8	0.981 4	120	0.966 3	0.977 5
30	0.969 7	0.980 9	130	0.963 3	0.975 1
40	0.967 0	0.979 0	140	0.962 9	0.974 8
50	0.967 0	0.978 7	150	0.963 7	0.975 7
60	0.966 9	0.978 5	160	0.585 4	0.609 0
70	0.965 6	0.977 3	170	0.961 7	0.972 4
80	0.965 9	0.977 4	180	0.963 0	0.973 8
90	0.964 1	0.975 9	190	0.963 1	0.973 8
100	0.963 0	0.975 0	200	0.963 1	0.973 5

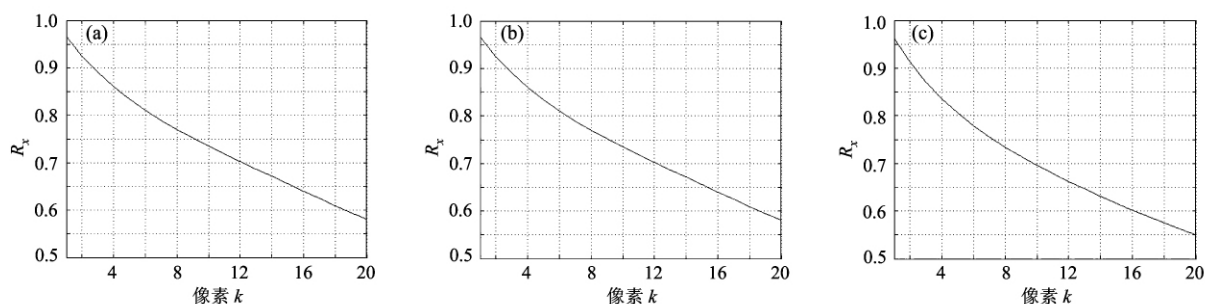


Fig 2 Row auto correlation of several bands of hyperspectral image

(a): Twentieth; (b): Fortieth; (c): Hundredth

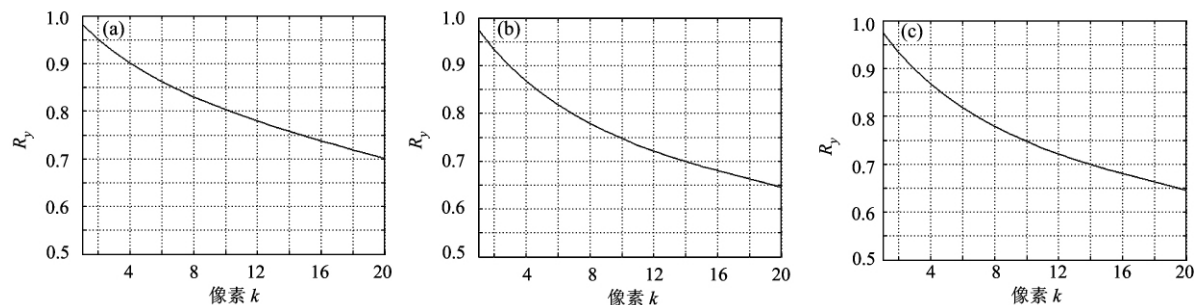


Fig 3 Column auto correlation of several bands of hyperspectral image

(a): Twentieth; (b): Fortieth; (c): Hundredth

为进一步对超光谱图像特性进行分析,选取国际标准图像 Lena 进行对比测试,结果见图 4(虚线代表 Lena 图像相关函数曲线,实线代表超光谱图像相关函数曲线)可见超光谱图像的空间自相关性低于普通图像,且随相隔像素数增加相关系数下降得更快,说明超光谱遥感图像的相关范围比 Lena 图像小,且随间隔像素增加,相关性下降迅速。

1.2 谱间相关性

$$h(l, k) = \frac{\sum_{x=1}^M \sum_{y=1}^N [f(x+l, y+k) - u_f][g(x, y) - u_g]}{\sqrt{(\sum_{x=1}^M \sum_{y=1}^N [f(x, y) - u_f]^2)(\sum_{x=1}^M \sum_{y=1}^N [g(x, y) - u_g]^2)}} \quad (2)$$

其中, $f(x, y)$ 和 $g(x, y)$ 分别代表相邻波段点 (x, y) 的像素灰度值, l 和 k 代表行、列位置变化。 $l = k = 0$ 时,称 $h(0, 0)$

谱间相关性包括统计相关性和结构相关性。统计相关性即各波段图像灰度分布的相关性,其大小很大程度上由光谱分辨率决定,光谱分辨率越高统计相关性越高;结构相关性即不同波段同一空间位置像素与其相邻像素的关系的相似性。为分析谱间相关性,首先引入互相关函数概念,经归一化和离散化的互相关函数公式为

为两图像的互相关系数,第 i 波段的谱间相关系数定义为第 i 波段与第 $i + l$ 波段的互相关系数。

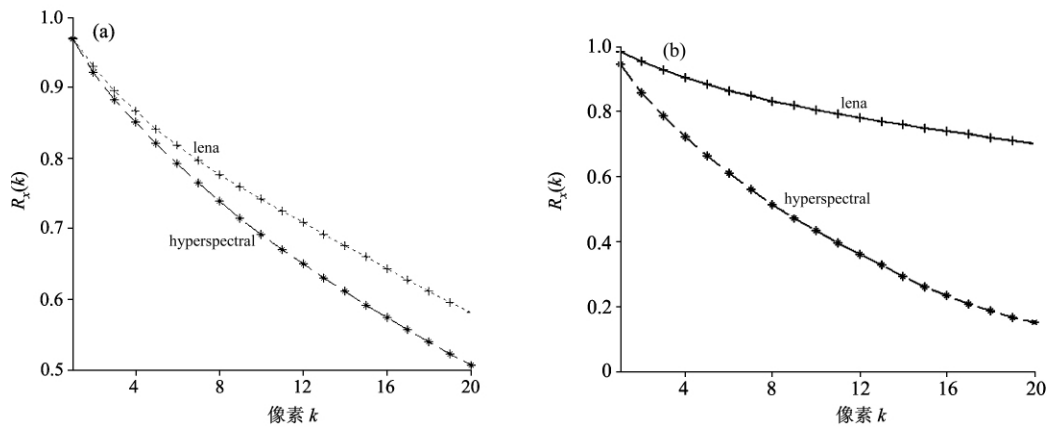


Fig 4 Auto correlation's contrast between hyperspectral image and RGB image

(a): Contrast of row auto correlation; (b): Contrast of column auto correlation

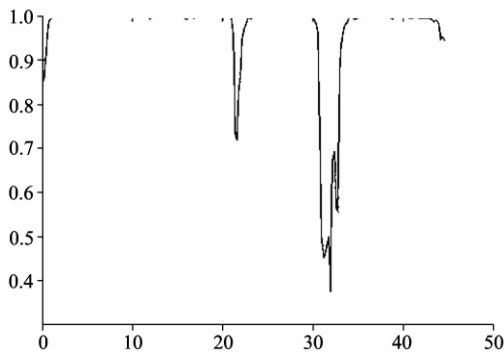


Fig 5 Graph of cross correlation

由图 5 和表 2 可以看出,除波段 1~10, 100~120, 150~175, 其余波段谱间相关系数均大于 0.95, 有的甚至高达 0.999 9, 这些都说明超光谱图像相邻波段间的相似度非常高,充分利用这点,可得到很好的压缩效果。

下面分析超光谱图像的统计相关性,通过对相邻波段图像直方图研究发现,相邻波段直方图形状十分接近(如图 6),说明其灰度分布有很高相关性。和普通 RGB 图像对比,实验表明普通彩色图像各波段直方图形状相差很大,基本不具备相似性且无规律。

通过上面的对比可得出:超光谱图像具有很强的谱间统计相关性,大大高于普通的彩色图像,产生这种相似性的原因有两点:(1)光谱图像的每个波段图像的像素值是相同区

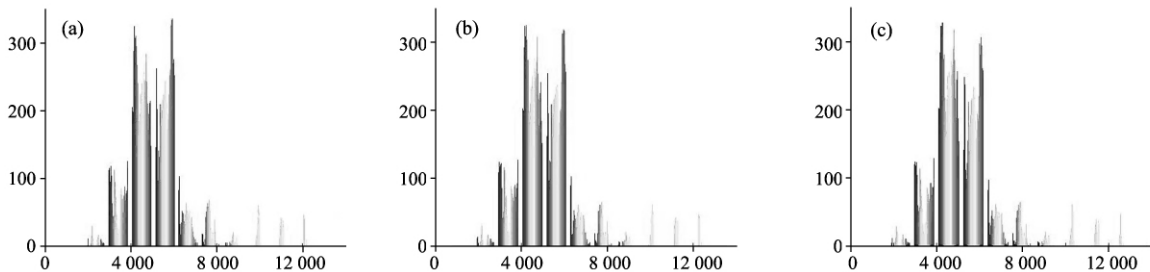


Fig 6 Histograms of 20~22 bands

(a): Twentieth; (b): Twenty-first; (c): Twenty-second

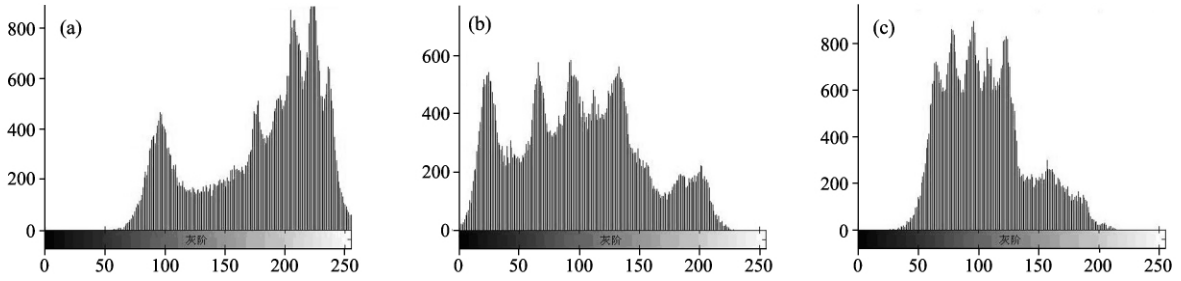


Fig 7 Histograms of R, G, B bands

(a): R band; (b): G band; (c): B band

域地物对各个波段光的反射强度值, 相邻波地物反射率相近, 由此产生了一定相关性。(2) 不同波段图像所涉及的地面目标相同, 故它们具有相同的空间拓扑结构。

Table 2 Inter-band correlation coefficients of hyperspectral

波段	谱间相关系数	波段	谱间相关系数
10	0.999 4	110	0.865 0
20	0.999 8	120	0.996 2
30	0.999 9	130	0.999 8
40	0.999 7	140	0.999 8
50	0.999 7	150	0.997 4
60	0.994 7	160	0.375 1
70	0.999 8	170	0.997 2
80	0.992 5	180	0.999 1
90	0.999 8	190	0.999 0
100	0.999 6	200	0.998 8

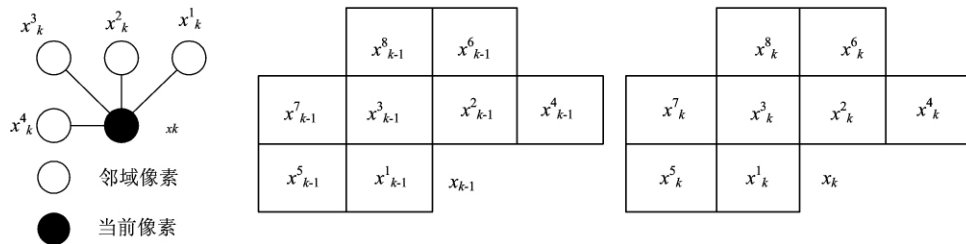


Fig 8 Neighbor domain and neighboring pixels

(a): Schematic of side neighbors domain; (b): Neighboring-pixels used by calculating error compensation coefficient

补偿系数 α 由前一波段相邻的 8 个像素及当前波段相邻的 8 个像素(具体位置定义如图 8(b))在最小均方误差准则下计算得到。

$$\alpha = \frac{\sum_{n=1}^8 x_{k-1}^n x_k^n - \sum_{n=1}^8 x_{k-1}^n \sum_{n=1}^8 x_k^n}{8 \sum_{n=1}^8 (x_{k-1}^n)^2 - \left(\sum_{n=1}^8 x_{k-1}^n \right)^2} \quad (5)$$

由于对每一个像素计算一次补偿系数, 计算量庞大, 对补偿系数的计算使用了两个波段共 16 点的邻域, 考虑对于局部稳定区域可以使用同一个补偿系数, 无需重新计算, 依据文献[18]中提到的判断局部稳定区域方法, 对当前像素点如果其当前波段侧四邻域四个像素的预测误差和小于某一阈

2 压缩算法

2.1 自适应误差补偿的侧四邻域预测算法

图 8(a)为侧四邻域的示意图^[6], 利用其进行谱间预测的公式为

$$\bar{x}_k = x_k^n \quad (3)$$

式中 \bar{x}_k 由与 k 波段当前像素对应的 $k-1$ 波段图像像素 x_{k-1} 距离最近的邻近像素位置号确定, 即满足 $|x_{k-1} - x_{k-1}^n| = \min_{n=1, \dots, 4} |x_{k-1} - x_{k-1}^n|$, 侧四邻域预测算法利用超光谱图像谱间结构相关性和空间相关性, 但未利用谱间统计相关性, 因此考虑利用前一波段预测误差对当前预测进行补偿, 如式(4)。

$$\hat{x}_k = x_k^n + \alpha(x_{k-1} - \bar{x}_{k-1}) \quad (4)$$

值, 则认为当前像素在局部稳定区域内(如式(6))。

$$e_{\text{sum}} = \sum_{n=1}^4 |x_k^n - \hat{x}_k| = \sum_{n=1}^4 |e^n| < T \quad (6)$$

2.2 多波段预测算法

对像素 $p(i, j, k)$, 其线性预测关系可表示为

$$x(i, j, k) = \hat{x}(i, j, k) + e(i, j, k) =$$

$$\sum_m \sum_n \sum_b a_{m,n,b} x(i-m, j-n, k-b) + e(i, j, k) \quad (7)$$

式中, $\hat{x}(i, j, k)$ 表示预测值, $e(i, j, k)$ 表示预测残差。考虑到超光谱数据具有较强谱间相关性及相对较弱空间相关性的特点, 提出自适应选取单空间邻点及多波段预测模型。邻点的选择采用对侧四邻域像素点的四选一方法, 即分别以多个相邻波段当前空间位置的左、左上、上、右上四个方向空间邻

点为预测数据训练预测系数。对每个空间位置，从四组预测系数中选择其一作为预测器。预测器选取准则为，该空间位置在各个波段的误差绝对值之和最小。

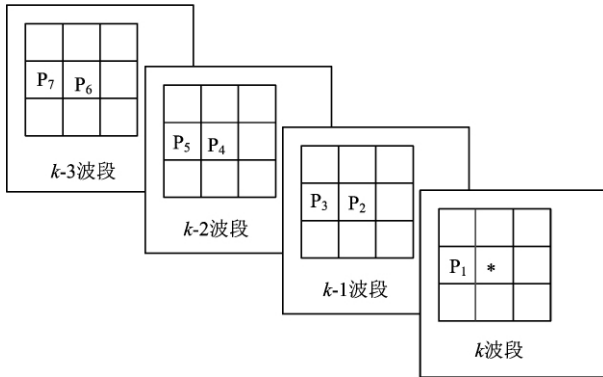


Fig 9 Schematic of predicted points

图 9 给出了采用左方向空间位置、三个预测波段、共七个预测点的示意图，图中标有“*”者为当前点。以残差平方和最小为准则，可以对各样本点训练得到预测系数。假定采用 k 点预测，设当前点为 p_0 ，预测点为 p_1, p_2, \dots, p_k ，预测系数表示成 $A = [a_1 \ a_2 \ \dots \ a_k]$ ，选择 L 个点训练，整理预测矩阵

$$\phi = \begin{bmatrix} p_{1,1} & p_{2,1} & \dots & p_{k,1} \\ p_{1,2} & p_{2,2} & \dots & p_{k,2} \\ \vdots & \vdots & \dots & \vdots \\ p_{1,L-1} & p_{2,L-1} & \dots & p_{k,L-1} \\ p_{1,L} & p_{2,L} & \dots & p_{k,L} \end{bmatrix} \quad (8)$$

设 $G = [p_{0,1} \ p_{0,2} \ \dots \ p_{0,L-1} \ p_{0,L}]^T$ ，则预测值 $\hat{G} = \phi \cdot A$ ，使 $(\hat{G} - G)^T (\hat{G} - G)$ 最小化，运算得

$$A = (\phi^T \phi)^{-1} \phi^T G \quad (9)$$

对不同样本点分别得到四组预测系数表示， k 表示波段序号， d 表示邻域像素的方向标识。利用式(7)，将应用到各相应像素得到其预测值，与真值相减，得到残差图像。具体算法如下：

(1) 选择起始波段 k ；起始波段 k 逐行逐列对各像素分别计算三个波段的四个方向单邻点作为预测数据时的预测系数以及对应预测值(若像素处于边缘，则只计算存在的侧四邻域像素点对应的预测系数和预测值)；

(2) 计算该像素各方向的均方误差，取均方误差最小的方向所计算出来的预测值作为该像素点的预测值，同时记录该波段各像素点所选取的方向；

(3) 对其他所有波段各像素点采用与起始波段对应像素点所选取方向对应的邻点作为单邻点，计算预测值，压缩完毕。

自适应误差补偿侧四邻域预测算法以及多波段预测算法的效果都与谱间相关性有很大关系，下面取谱间相关性较大的 12~59 波段和谱间相关性较小的 149~178 波段分别用两种方法进行实验，图 10 和图 11 给出了原图熵和残差图像熵对比结果。

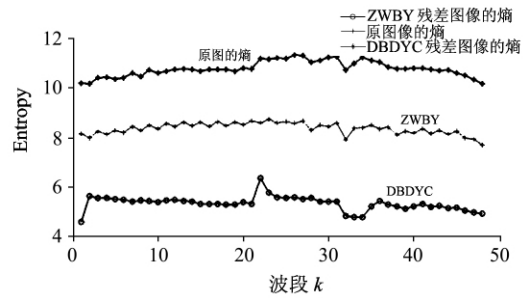


Fig 10 Entropy comparison from band 12 to 59

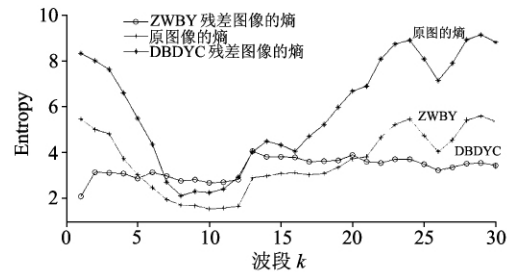


Fig 11 Entropy comparison from band 149 to 178

图中 ZWBY 代表自适应误差补偿侧四邻域预测算法，DBDYC 代表多波段预测算法。从图 10 可以看出对谱间相关性较高的波段，两种方法都有很好的效果，所得残差熵均降低很多，从图中明显看出利用多波段预测的方法得出的结果更好，残差图像的熵更小。由图 11 可得对于谱间相关性相对较低的波段来说，多波段预测虽然在有的波段，比如 149~153 信息熵压缩效果也很好，但从整体来看意义并不大，而自适应误差补偿预测方法则表现非常出色，在整体上均起到了熵压缩的目的，并且压缩的幅度令人满意。

由此得由于超光谱图像谱间相关性的差异性以及各预测方法的局限性对整幅超光谱图像采用一种压缩方法效果并非最优。

3 总体预测压缩方案

基于上述分析本文总体方案如图 12。

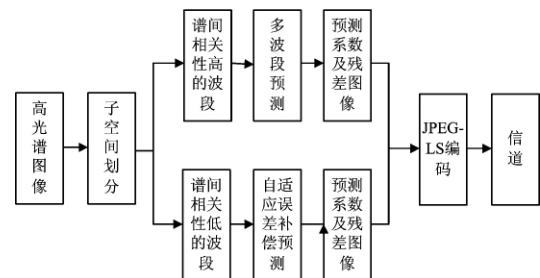


Fig 12 Block diagram of algorithm proposed

(1) 首先根据超光谱图像谱间相关性，将超光谱图像分为若干个子空间，根据每个子空间的特点进行相应的压缩。

(2) 对谱间相关性高的子空间利用多波段预测进行谱间

压缩;对谱间相关性低的子空间利用自适应误差补偿预测的方法进行压缩。分别得到残差图像和预测系数。

(3) 将(2)得到的残差图像和预测系数用 JPEG-LS 进行压缩编码,即得压缩图像。

4 实验结果

本文对 AVIRIS 超光谱图像在 Intel 迅驰主频为 1.66 GHz 的处理器、1G 内存的电脑上进行了实验,为便于计算和比较结果,把各波段图像切割成 512×512 的图像进行处理。首先根据式(2)计算原始 AVIRIS 数据空间各谱间相关系数,然后根据各相关系数值在预先给定的门限下进行子空间分解。各子空间内谱带间相关性接近。通过实验得门限值取为 0.999 时较为合适。此时,224 个谱带空间被分解为 9 个子空间(如表 3),据谱间相关性大小对子空间 1、4、6 和 9 采用自适应误差补偿预测方法进行谱间压缩,对子空间 2、3、5、7 和 8 采用多波段预测方法。

Table 3 Partition results of subspaces

划分子空间	谱带范围	数据维数	谱带间压缩的方法
子空间 1	1~8	8	自适应误差补偿预测
子空间 2	9~59	51	多波段预测
子空间 3	60~103	44	多波段预测
子空间 4	104~121	18	自适应误差补偿预测
子空间 5	122~148	27	多波段预测
子空间 6	149~178	30	自适应误差补偿预测
子空间 7	179~186	8	多波段预测
子空间 8	187~216	30	多波段预测
子空间 9	217~224	8	自适应误差补偿预测

由于数据信息熵决定了压缩可获得的压缩比,图 13 给出了原图像熵及采用本文的谱间压缩方案处理后的残差图像熵。

将图 13 与图 5 联系起来分析可得对谱间相关性大的子空间其信息熵压缩效果很好,而对谱间相关性较小的子空间信息熵没有其他波段高,熵的压缩效果也没有其他波段好,但较单独使用一种预测算法进行谱间压缩的效果和速度都提高很多,稳定很多。

仿真过程中,多波段预测算法采用三个相邻波段、七个

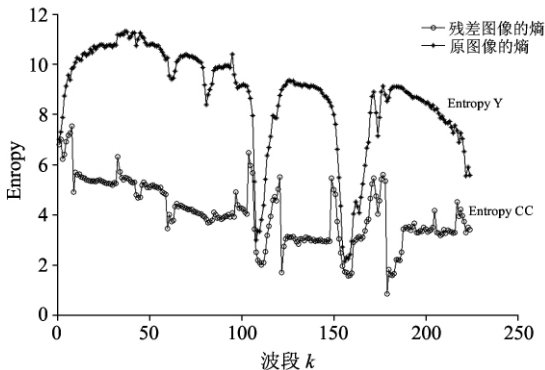


Fig 13 Entropy contrast curve of 224-band AVIRIS image and residual image

点来进行预测。对预测系数的选取采用的预测误差绝对值之和最小准则是对子空间各波段而非对整个波段的,并且对各子空间只计算该空间首波段上各点预测点的方向并记录下来,在该子空间中其他波段均采用该对应预测点方向,大大简化了算法,提高了运算效率而效果几乎一样。

表 4 给出了各算法得到的压缩比,从表中可看出相较预测树算法,自适应误差补偿预测算法的压缩比得到一定程度上的提高,而单独使用自适应误差补偿预测和多波段预测算法都没有结合起来的本文算法的压缩比高,这充分说明了本文算法的有效性。

Table 4 Contrast of compression ratio

算法	预测树	自适应误差补偿预测	多波段预测	JPEG-LS	本文算法
AVIRIS	2.75	2.86	2.92	2.43	3.10

5 结 论

本文首先分析了超光谱图像的特性,其谱间相关性很高但不是所有波段都一致不变,而是在一定范围内各不相同。自适应误差补偿算法和多波段预测算法都有其局限性,基于此本文提出子空间划分思想针对不同情况使用不同算法、扬长避短充分利用超光谱图像特性,通过实验证明了其较单独使用一种方法更有效性,压缩比更高。

References

[1] YAN Jing-wen, YU Jian, QU Xiao-bo, et al(闫敬文,余见,屈小波,等). Optics and Precision Engineering(光学精密工程), 2007, 10: 1622.

[2] LUO Xin, GUO Lei, LIU Zhen(罗欣,郭雷,刘震). J. Infrared Million. Waves(红外与毫米波学报), 2007, 6(26): 443.

[3] SUN Lei, LUO Jian-shu, GU De-feng(孙蕾,罗建书,谷德峰). Optics and Precision Engineering(光学精密工程), 2008, 4(16): 752.

[4] ZHAN Yong-jian, XIN Qin, SUN Lei, et al(粘永健,辛勤,孙蕾,等). Optics and Precision Engineering(光学精密工程), 2008, 6(16): 1146.

[5] HU Xing-tang, ZHANG Bing, ZHANG Xia, et al. Science in China Series E: Technological Sciences, 2006, (49): 48.

[6] WU Ying-qian, FANG Tao, SHI Peng-fei(吴颖谦,方涛,施鹏飞). J. Infrared Million. Waves(红外与毫米波学报), 2004, 5(23): 349.

- [7] Memon K Sayood, Magliveras S S. IEEE Trans. Geosci. Remote Sensing, 1994, 32(2): 282.
- [8] Wu X, Memon N D. IEEE Trans. on Image Processing, 2000, 9(6): 994.
- [9] ZHANG Rong, YAN Qing, LIU Zheng-kai(张 荣, 阎 青, 刘政凯). Journal of Remote Sensing(遥感学报), 1998, 2(3): 171.
- [10] WU Zheng, HE Ming-yi, FENG Yan, et al(吴 铮, 何明一, 冯 燕, 等). Journal of Remote Sensing(遥感学报), 2005, (3): 143.
- [11] <http://aviris.jpl.nasa.gov/>.
- [12] Michael J Ryan, John F Arnold. Remote Sensing of Environment, 1997, 61(3): 419.
- [13] YU Guo-xia, Tanya Vladimirova, Martin N Sweeting. Acta Astronautica, 2009, 64(9-10): 988.
- [14] LI Chun-xin, WANG Xiao-tong(李春鑫, 王孝通). Optics and Precision Engineering(光学精密工程), 2009, 9(17): 2321.
- [15] Liu G, Zhao F. IET Image Processing, 2008, 2(2): 72.
- [16] SUN Lei, GU De-feng, LUO Jian-shu(孙 蕾, 谷德峰, 罗建书). Optics and Precision Engineering(光学精密工程), 2009, 11(17): 2864.
- [17] Li X, Orchard M T. IEEE Trans. Image Processing, 2001, 10(6): 813.
- [18] DONG Chao, ZHAO Hui-jie, WANG Wei, et al(董 超, 赵慧洁, 王 维, 等). Optics and Precision Engineering(光学精密工程), 2009, 8(17): 2004.

A New Hyperspectral Image Compression Method Combined with Subspace Partition and Multi-Inter-Spectral Prediction

GAO Wen^{1,2}, ZHU Ming¹, TIAN Run-lan³, TANG Yang⁴

1. Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China
2. Graduate University of Chinese Academy of Sciences, Beijing 100049, China
3. University of Aeronautics and Astronautics, Changchun 130000, China
4. College of Electronic Science & Engineering, Jilin University, Changchun 130000, China

Abstract Because of their rich and complex texture, multi-band and high spectral resolution, hyperspectral images are widely used in military, marine, agriculture and other aspects; however, their huge data also brings many troubles. Currently, there're almost three types of the hyper-spectral image compression methods: the method based on transform, prediction and vector quantization, among which, the method based on prediction has the advantages such as easy implement and higher compression ratio, then it's widely used. In the present paper, features of hyperspectral images are analyzed first in detail, then inter-spectral prediction algorithm is improved, and a new scheme which consists of subspace partition and multi-inter-spectral prediction is proposed, and finally, the experiments show that the program is effective.

Keywords Hyperspectral image; Compression; Multi-inter-spectral prediction; Error compensation; Subspace partition

(Received Jun. 19, 2010; accepted Sep. 20, 2010)