

GMS 卫星红外云图数据压缩

吴建华 占传杰 黎 鹰

王顺长

(南昌大学信息科学学院, 南昌 330029)

(江西省防汛抗旱总指挥部, 南昌 330046)

摘 要 依据 GMS 卫星红外气象云图的统计特性, 提出了一种“亚抽样+修改的 JPEG+图象放大”的静止图象编码方法, 使有失真压缩比在 50:1 至 100:1 时, 恢复图象峰值信噪比在 35dB~40dB 之间, 主观质量无视觉可觉察到的失真。对于信息保持型 GMS 卫星红外云图数据压缩, 选取训练集求取最小均方误差意义下的最佳线性预测器, 并作为固定预测器应用于每幅云图, 性能优于目前文献报导的固定预测器。

关键词 JPEG 修改的 JPEG 数据压缩 气象云图

0 引 言

卫星接收机接收到的 GMS 卫星红外云图经压缩后在防汛网上传输到各终端, 解压缩后供防汛、抗旱、气象及有关决策部门使用。由于卫星云图的数据量太大使得没有数据压缩而直接传输原始数据不能实用化。防汛抗旱等部门利用卫星云图定性观察天气形势是可以容许一些失真的, 因此可以以视觉觉察不到的失真为代价以得到尽可能大的压缩比。但某些应用部门例如气象部门需要进行定量分析, 则不希望出现失真, 况且卫星数据资源是很宝贵的, 8bit 红外云图每个量化级对应大约 0.5℃ 的温度变化。本文依据红外云图的统计特性, 提出“亚抽样+修改的 JPEG+图象放大”的静止图象编码方案, 得到了 50:1 至 100:1 的有损压缩比, 主观质量无明显失真。对于信息保持型卫星云图数据压缩, 本文采用的预测器得到了大约 2.3 的无损压缩比, 与文献[1]介绍的诸固定预测器相比, 性能最好, 其次是 Sawada 和徐孟侠的固定帧内预测器。

1 GMS 卫星红外云图的统计特性

反映随机信号相关特性的自协方差矩阵 C 是随

机信号的一个重要统计特征。在对 GMS 卫星云图数据相关特性进行分析的过程中, 我们得到了下列统计结果, 该结果是利用几十幅大小为 1 024 × 1 024 × 8bit 的 GMS 红外云图得到的。作为对照, 将张春田给出的自然图象的自协方差矩阵一并列出^[2]。所列自协方差矩阵中的元素已进行了归一化处理, 实际上是相关系数。

卫星云图自协方差矩阵:

1.000	0.996	0.989	0.981	0.975	0.969	0.963	0.958
0.996	1.000	0.996	0.989	0.982	0.975	0.969	0.963
0.989	0.996	1.000	0.996	0.989	0.982	0.975	0.969
0.981	0.989	0.996	1.000	0.996	0.989	0.982	0.975
0.975	0.982	0.989	0.996	1.000	0.996	0.989	0.982
0.969	0.975	0.982	0.989	0.996	1.000	0.996	0.989
0.963	0.969	0.975	0.982	0.989	0.996	1.000	0.996
0.958	0.963	0.969	0.975	0.982	0.989	0.996	1.000

自然图象自协方差矩阵:

1.00	0.99	0.97	0.96	0.94	0.93	0.91	0.89
0.99	1.00	0.99	0.97	0.96	0.94	0.93	0.91
0.97	0.99	1.00	0.99	0.97	0.96	0.94	0.93
0.96	0.97	0.99	1.00	0.99	0.97	0.96	0.94
0.94	0.96	0.97	0.99	1.00	0.99	0.97	0.96
0.93	0.94	0.96	0.97	0.99	1.00	0.99	0.97
0.91	0.93	0.94	0.96	0.97	0.99	1.00	0.98
0.89	0.91	0.92	0.94	0.96	0.97	0.98	1.00

由上面2个矩阵可以看出,如果将图象近似看成一阶马尔可夫过程,则 GMS 卫星云图的相关系数 ρ 比自然图象的相关系数 ρ 要大,即 GMS 卫星云图象素间的相关性(或说象素间多余度)要大于自然图象。

2 GMS 卫星云图有损压缩

静止图象压缩的国际标准 JPEG 算法^[3]是对一般的静止图象提出来的,对于象素间多余度更大的卫星云图,我们可以从以下2个途径在保证恢复图象主客观质量的前提下,进一步提高压缩比。

(1)采用类似 JPEG 基本系统^[3]推荐的方案,但图象子块的大小取 16×16 。JPEG 基本系统包括 DCT、量化及符号编码3个步骤,而在进行 DCT 前,首先将图象分为 8×8 的子块。子块大小的选取应使得相邻子块的相关性减小到某个可以接受的水平。从前面自然图象与卫星云图自协方差矩阵的对比来

看,如以一阶马尔可夫过程来近似,则可大致估计自然图象某个象素与相邻子块对应位置的象素之间的相关系数为 $\rho^8 = 0.99^8 = 0.923$;当子块大小为 16×16 时,卫星云图的某个象素与相邻子块对应位置的象素之间的相关系数为 $\rho^{16} = 0.996^{16} = 0.937 (> 0.923)$ 。因此,卫星云图可以选取 16×16 的子块,以利于提高压缩比,代价是增加了 DCT 计算的复杂性。

(2)“亚抽样+修改的 JPEG+图象放大”由于象素间有大于自然图象的相关性,则从频谱的观点来看,它的频宽要小于自然图象的频宽,或者说,卫星云图信号的能量更多地集中在低频端,因此可以对原图象进行隔点抽样(行列方向取样率减小一倍),再取子块大小 8×8 进行压缩。在解压缩时用频域补零的方法进行图象放大。由于亚抽样的过程已经完成了 4:1 的压缩,则量化阶段可以较小的量化阶距进行量化,从而保证恢复图象质量。(图1)

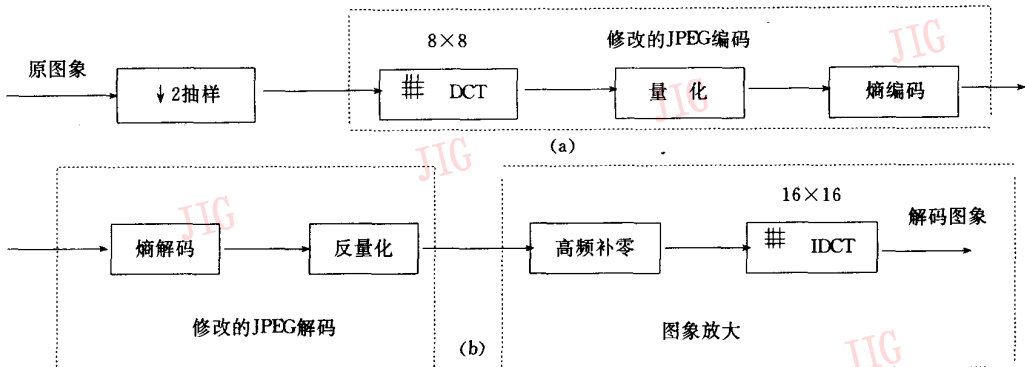


图1 卫星云图有损压缩模型: (a)编码器 (b)解码器

实验结果表明,上述2个方案在恢复图象峰值信噪比相当的情况下得到大致相同的压缩比,但前者压缩时间较长(子块增大,DCT 及 IDCT 运算复杂性增大),且在压缩比较大时容易出现方块效应,故实际应用中首选“亚抽样+改进的 JPEG+图象放大”方案,它的实施细节有2点值得说明。

(1)图象放大^[4] 图象解码并不完全采用编码的逆过程,即完成 IDCT 解码后对图象进行内插处理,而是直接在 DCT 变换域的高频部分向外延拓补零,将 8×8 变换域子块变为 16×16 子块,经逆 DCT 后即得到放大的了图象子块。考虑到 16×16 变换域高频部分为零(左上角四分之一部分不为零),在用快速算法实现 IDCT 时,还可节省乘、加次数^[5]。

(2)霍夫曼码表 JPEG 推荐的标准霍夫曼码表对于某些特殊类图象并不是最佳的。在修改的 JPEG 方案中,本文随机选取1996年元月至1997年5月的35幅红外云图进行分块 DCT 和量化,经 Zigzag 排序得到不同长度的零系数游程,据此统计出零系数游程出现的频度编出霍夫曼码表。经实验统计出,使用这样的霍夫曼码表比 JPEG 推荐的标准霍夫曼码表压缩比平均提高约5%。

采用“亚抽样+修改的 JPEG+图象放大”方案的卫星云图压缩已应用于江西省防汛计算机网络系统,部分数据如表1。其中主观质量是由防汛部门专家、课题组成员、非图象处理技术人员等对14”彩色计算机监视器上的伪彩色云图进行观察而得到的综合评价。解压缩图象主观上和原图象没有可觉察到

的差别。象素值乘以4的误差图象,由于地球边缘的 两边温差太大,故恢复图象局部方块效应略为严重。

表1 卫星云图有损压缩部分实验数据

云图日期时间	压缩比	恢复图象峰值信噪比(dB)	主观质量	云图日期时间	压缩比	恢复图象峰值信噪比(dB)	主观质量
1996:12:10:04	52.04	38.77	无失真	1996:01:18:12	65.20	35.19	无失真
	72.55	37.66	无失真	1997:05:04:03	60.80	35.08	无失真
	97.93	36.52	略有方块效应	1997:05:04:11	55.06	33.56	无失真
	111.39	36.21	略有方块效应	1997:05:04:21	62.88	34.60	略有方块效应
				1997:05:05:05	62.92	35.31	略有方块效应

3 信息保持型卫星云图数据压缩

GMS 卫星云图信息保持型图象压缩使用简单的 DPCM 方法,如图2。除邻域及其预测器的选择

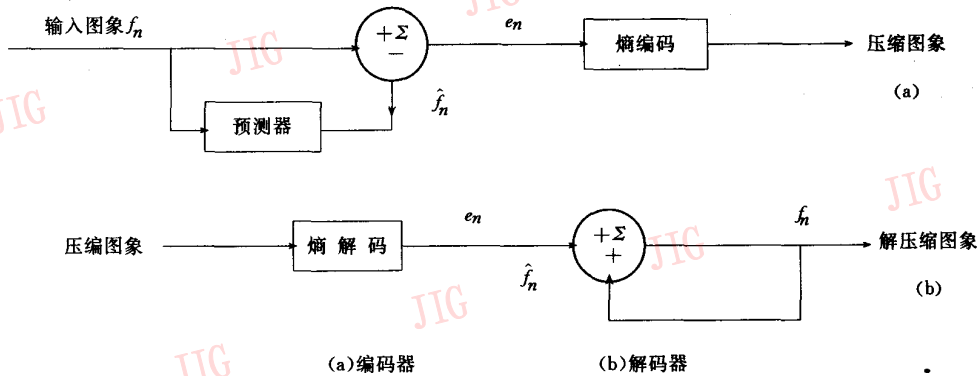


图2 无失真预测编码模型

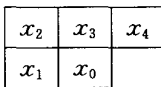


图3 4点预测邻域

由图2可以看出,无失真编码器的性能很大程度上取决于预测器的性能。许多文献指出(例如[6]),为使总的预测误差最小,预测器

$$x_0 = \sum_{i=1}^4 \alpha_i x_i \quad (1)$$

应使用最小均方误差预测器,预测系数向量 α 是方程

$$R\alpha = r \quad (2)$$

的解,其中 R 是 4×4 自相关矩阵, α 和 r 是4维列向量

$$\alpha = [\alpha_1 \ \alpha_2 \ \alpha_3 \ \alpha_4]^T, \\ r = [E(x_0 x_1) \ E(x_0 x_2) \ E(x_0 x_3) \ E(x_0 x_4)]^T$$

外,这个方法与 JPEG 建议的无失真编码系统极其相似。

本文预测器的预测邻域选用徐孟侠^[1]提出的邻域象素,并如图3那样记为 x_1, x_2, x_3, x_4 ,当前象素记为 x_0 。

随机选取1996、1997年的数十幅云图作为训练集,求出自相关函数并解方程(2),并进而求出以1/8阶距进行量化的预测系数为

$$\alpha_1 = 0.75, \alpha_2 = -0.375, \alpha_3 = 0.25, \alpha_4 = 0.375$$

以此预测器作为固定预测器,对 GMS 卫星云图压缩的实验得到了约2.1~2.4的压缩倍数。并同文献[1]中选取2种较好的预测器的压缩结果以及磁盘文件压缩程序 ARJ 的压缩结果作一比较,实验数据见表2。从中可以看出,对于 GMS 卫星云图的压缩,本文预测器的误差熵最小,压缩倍数最大,其次是 Sawada (1978, $\alpha_1 = 0.875, \alpha_2 = -0.5, \alpha_3 = 0.5, \alpha_4 = 0.125$) 和徐孟侠 ($\alpha_1 = 0.5, \alpha_2 = 0, \alpha_3 = 0.25, \alpha_4 = 0.25$) 的固定预测器。由于系数是0.125的整数倍,而不是0.25的整数倍,故硬件实现和 Sawada 固定预测器难度相同,但比徐孟侠固定预测器略复杂。文献[1]列出的其他固定预测器均次于上述预测器,

故实验数据也不在表中列出。本文熵编码采用吴乐南^[7]提出的反射型熵编码器,编码效率在60~70%之间。

表2 卫星云图无损压缩实验数据

卫星云图日期	Sawada		徐孟侠		本文		最小均方误差		ARJ
	误差熵	压缩比	误差熵	压缩比	误差熵	压缩比	误差熵	压缩比	压缩比
1996:01:18:12	2.163	2.228	2.26	2.174	2.147	2.232	2.159	2.233	1.819
1996:12:10:04	1.879	2.354	1.98	2.308	1.857	2.369	1.879	2.364	2.12
1997:05:04:03	2.157	2.229	2.269	2.166	2.114	2.245	2.111	2.248	1.90
1997:05:04:10	2.303	2.156	2.401	2.099	2.270	2.166	2.265	2.170	1.80
1997:05:04:19	2.278	2.168	2.39	2.104	2.241	2.181	2.23	2.28	1.81

4 小结及讨论

(1)根据卫星云图相关特性大于自然图象的结论,本文提出对图象进行亚抽样,实现了4:1的预先压缩。然后在应用 JPEG 算法时,又根据卫星云图的统计结果获得相应的霍夫曼码表,并适当选择量化矩阵的比例系数以得到12:1至25:1的压缩倍数,两个步骤共实现了约50:1至100:1的压缩倍数。解压缩时利用数据已经是 DCT 变换域的数据这一事实,在频域补零再进行逆 DCT 既实现了图象放大,又保持了图象的全部低频分量,从而保证了高的主观质量。本成果已成功应用于江西省防汛计算网络系统二期工程的卫星云图压缩和传输。

(2)由于本文的预测器是依据足够的云图的统计特性求出的,理论和实验数据说明它应用于信息保持型卫星云图数据压缩是有效的。信息保持型编码的最佳预测器许多文献都是基于最小均方误差准则而导出线性预测器,但从压缩编码的观点来看,应以误差熵最小来导出最佳预测器,且应是非线性的。虽然在某些假设下,例如典型差值信号 e 的概率分布以拉普拉斯分布近似时^[6],均方误差最小与误差

熵最小是等价的,但这毕竟是一种近似,实际上并不总是这样。表2中有的云图本文预测器的误差熵略小于最小均方误差预测器的误差熵即能说明这个问题。笔者以为,误差熵最小准则下的非线性预测器是个值得研究的课题。

致谢 解楠、徐厚城、李田等同学完成了部分软件的调试工作,谨致谢意。北京大学徐孟侠教授对本文的修改提出了指导性意见,谨表示衷心的感谢。

参考文献

- 1 徐孟侠.多光谱扫描仪(MSS)卫星遥感图象原始数据的信息保持型压缩编码.电子学报,1993,21(7):104~106
- 2 张春田.运动补偿视频编码中 DCT 编码效率的研究.电子学报,1996,24(1):1~5
- 3 ISO 10918, Digital compression and coding of continuous-tone still images, 1994. 2
- 4 崔屹.数字图象处理技术与应用,北京:电子工业出版社,1997. 3
- 5 解楠,吴建华.基于卫星云图相关特性的 DCT 快速算法.全国第八届信号处理学组委员会联合学术会议论文集,北京:航空工业出版社,1997. 8
- 6 Rafael C. Gonzales, Digital Image Processing, Addison-Wesley, Reading, Mass. 1993
- 7 吴乐南.数据压缩的原理与应用,北京:电子工业出版社,1994.



吴建华 南昌大学信息科学工程学院副教授,主要研究方向为数字图象处理,模式识别。



占传杰 南昌大学信息科学工程学院副教授,主要研究方向为计算机网络及图象处理。

黎 鹰 南昌大学信息科学工程学院工程师,主要研究方向为计算机网络及通信技术。



王顺长 江西省防汛抗旱总指挥部办公室工程师,主要从事应用软件开发工作。

Data Compression of Meteorological Satellite Cloud-Charts

Wu Jianhua, Li Ying, Zhan Chuanjie

(Information & Sci. Eng. College, Nanchang University, Nanchang 330029)

Wang Shunchang

(Jiangxi Provincial Headquarters of Anti-flood & Draught, Nanchang 330046)

Abstract On the basis of the statistical characteristics of satellite meteorological infrared cloud-charts this paper presents a coding algorithm of "Sub-sampling+Modified JPEG+Image enlarging", realizing a compression ratio of 50:1~100:1, with the PSNR of 35~40dB and visually unperceptible distortion. For error-free data compression of satellite cloud-charts, training set is selected to seek the minimum mean-square-error predictor, which is used as the fixed predictor and applied to the compression of any cloud-chart image, with a better performance than any other fixed predictors reported so far.

Keywords JPEG, Modified JPEG, Data compression, Satellite cloud-charts

关于举办“首届全国 电脑动画设计和制作大赛”的通知

各地方协会、各分会、各电视台、广告公司、计算机公司、各有关单位:

为了推广普及电脑动画设计和制作的应用知识、促进国内与国外电脑动画技术的信息交流、发展培育优秀人才、推动我国电脑动画设计和制作技术的发展,经原电子工业部计算机与信息化推进司[电算(1998)03号]文批准,中国计算机用户协会、中国电视艺术家协会、中央电视台、中国计算机报社研究决定,联合举办“首届全国电脑动画设计和制作大赛”。

电脑动画是融电脑技术、艺术、美术和创意于一体的全新的表现形式。本届大赛以歌颂祖国悠久文化,赞美祖国锦绣山河,展现改革开放以来现代化建设所取得的辉煌成就和社会主义市场经济的繁荣景象为主题,是一次全国性的计算机技术应用成果的交流活动的,是电脑动画设计与制作应用领域的重要赛事。

现将通知以及有关参赛办法发给你们,希望能够积极参加并支持此次活动,使这项活动取得圆满成功。

中国计算机用户协会
中央电视台

中国电视艺术家协会
中国计算机报社

(编者注:有关参赛办法将在本刊99'第2期刊发)