

# 多层次快速聚类的遥感图象无损压缩

王朝晖

(中国科学技术大学电子科学与技术系, 合肥 230026)

**摘要** 由于 K-means 聚类要求每个像素要和所有聚类中心求欧氏距离, 因此当聚类数很多时, 这是一个相当耗时的工作。改进后的 K-means 聚类算法使类内像素只通过和相邻的聚类中心进行距离计算来聚类, 由于随着算法的迭代进行, 大量类的状态基本固定, 因此使得聚类速度不断加快。多层次聚类无损压缩就是利用改进的 K-means 聚类算法具有快速收敛的特点, 和利用分层次去冗余的方法来聚类, 因此可最大限度消除残差冗余。基于 SP 整数小波变换的多层次聚类由于其不仅能消除空间冗余、结构冗余, 还能进一步对残差数据去冗余, 因而实现了多光谱遥感图象无损压缩的突破。最后通过不同算法对 TM 图象进行压缩的比较和参数分析, 论证了多层次聚类无损压缩的高效及合理性。

**关键词** 图象处理(510·4050) 多光谱图象 无损压缩 整数小波变换 聚类

中图法分类号: P751.1 TN919.81 文献标识码: A 文章编号: 1006-8961(2003)07-0843-06

## Fast Multi-level Clustering Lossless Compression Algorithm for Remotely Sensed Images

WANG Zhao-hui

(Department of Electrical Science and Technology, University of Science and Technology of China, Hefei 230026)

**Abstract** Every pixel in the super space is required by K-means algorithm to calculate Euclidean distance for clustering. When there are much many class centers, this is a much hard work. In this paper, an improved K-means clustering algorithm is presented to accelerate clustering process with more and more classes becoming stable by judging with neighbor centers nearest to the pixel. The inter-spectral redundancy and intra-spectral redundancy can be eliminated mostly by multilevel clustering algorithm with quickly convergent K-means classification and the method clearing redundancy at step through enhancing the intra-class pixel redundancy. The multi-level clustering process with initial S+P (Sequential transform + Prediction) integer wavelet transformation can not only remove the spatial and structural redundancy, but also delete the residual data redundancy realizing the breakthrough of lossless compression for multi-spectral images. Furthermore, the comparison with other lossless compression algorithm and the parameter analysis of the TM (Landsat Thematic Mapper) images show that this multilevel clustering compression algorithm is more reasonable and efficient.

**Keywords** Image Processing, Multispectral images, Lossless compression, Integer wavelet transform, Clustering

## 0 引言

多光谱成像技术的发展有力地推动了人类社会对自然资源的利用和开发, 但由于遥感数据量急剧增加, 利用十分困难, 因此如何实现快速有效的压缩始终是遥感信息处理的一个重要课题。如今已提出了一些压缩算法, 其中无损压缩主要采用聚类矢量

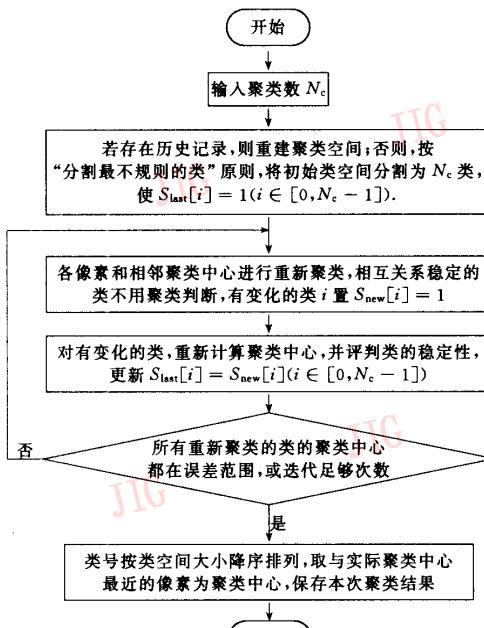
量化<sup>[1]</sup>、预测树<sup>[2]</sup>、及差分脉冲编码 DPCM (Differential pulse code modulation)<sup>[3]</sup>等方法, 并通过设计预测器来实现去冗余; 而近无损压缩则利用 KL 变换<sup>[2,4]</sup>来实现数据维的减少, 但产生变换矩阵比较费时。文献[5]采用整数小波变换和聚类的方法, 虽也可有效去除空间冗余和结构冗余, 但由于以上算法只是注意到原始数据的冗余处理, 因此压缩效率受到限制, 实际上, 残差数据的冗余更大。本文

提出的多层次聚类方法,可彻底消除这种冗余,但压缩数据仍主要受分类地图的影响。

本文根据像素的空间几何关系,对传统 K-means 聚类算法进行了改进,继而提出基于多层次快速聚类的无损压缩算法(Multilevel clustering lossless compression algorithm,即 MCLCA),该算法还可和整数小波变换 SP(Sequential transform+Prediction)进行有机结合(即 SP+MCLCA),即首先对各波段数据用整数小波变换去除空间冗余,再聚成 4 类,即可去除结构冗余<sup>[5]</sup>;然后用多层次聚类的方法去除残差数据冗余,使压缩率取得突破;最后取 TM 的 6 个波段数据进行参数分析和算法比较,实验结果表明,基于聚类的 SPMCLCA 算法较其他算法优越,无损压缩率可达 3.84 倍以上。

## 1 改进的 K-means 聚类算法(IKA)

K-means 聚类要求每个像素要和所有聚类中心求欧氏距离,并根据该距离来确定像素的归属。但当聚类数很多时,这是一个相当耗时的工作。据分析,实际上,有很多和像素相距较远的类,根本无须判断,而且很多类的状态稳定之后,就不用重新聚类。由于



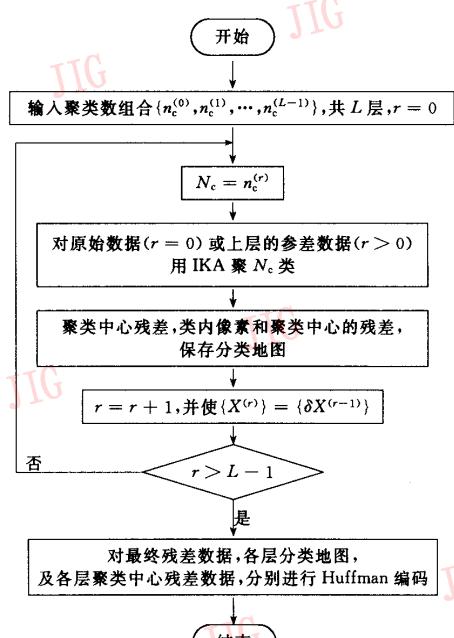
(a) IKA 算法

IKA 的初始聚类,对类空间分布最不规则的类是按空间伸展方向进行分割,且可基本保证类的均衡,从而为聚类迭代过程的快速收敛创造了条件。

### 1.1 IKA 聚类

设  $N_s$  是参加聚类的像素数;  $b$  表波段数;  $n_c$  为当前类分割数;  $C_i$  为类  $i$  像素集合;  $N_{C_i}$  表示类  $i$  集合  $C_i$  中的像素个数;  $e_p^{(i)}$  为类  $i$  中第  $p$  个像素(即  $e_p^{(i)} \in C_i$ );  $e_{\max}^{(i)}$  为类  $i$  中偏离中心最远的像素;  $c_{\text{last}}^{(i)}$  为上次迭代类  $i$  的聚类中心;  $C_p^{(i)}$  表示像素  $e_p^{(i)}$  的需重新聚类判断的类集合;  $D_{e_p^{(i)}, c_i}$  表示类内像素  $e_p^{(i)} \in C_i$  到聚类中心  $c_i$  的距离;  $d_{e_p^{(i)}, c_k}$  表示类  $i$  中第  $p$  个像素  $e_p^{(i)}$  到类  $k$  的聚类中心  $c_k$  的距离;  $d_{c_i, c_j}$  表示类  $i$  和类  $j$  的聚类中心  $c_i, c_j$  间的距离;  $d_{\max}[i]$  和  $\overline{d}_{e_p^{(i)}, c_i}$  分别是  $\{d_{e_p^{(i)}, c_i} | e_p^{(i)} \in C_i\}$  的最大值和均值(即数组  $d_{\max}[i] (i \in [0, n_c - 1])$  是类  $i$  内像素偏离聚类中心的最大值,并记该像素  $e_{\max}^{(i)}$  为  $P_i$ );  $S_{\text{last}}[i]$  是类  $i$  的上次迭代修改标志,用于判断类间稳定关系;  $S_{\text{new}}[i]$  是类  $i$  本次迭代修改标志,用于判定聚类收敛性;  $\epsilon$  是聚类中心最大迭代误差。聚  $N_c$  个类的 IKA 算法步骤如下(图 1(a)):

(1) 类空间初始分割  $N_c$  个类,产生初始聚类中



(b) MCLCA 算法

图 1 流程图

心  $c_i (i \in [0, N_c - 1])$ .

① 若存在历史记录(聚类数  $\tilde{N}_c$  小于  $N_c$ ), 则读出并重新构造上次聚类结果, 并取  $n_c = \tilde{N}_c$ ; 否则, 所有像素归为一类, 使  $N_c = 1$ , 则第 1 个类空间大小为  $N_s$  个像素, 并求聚类中心  $c_0$ , 计算类内像素  $e_p^{(0)} \in C_0$  ( $p \in [0, N_s - 1]$ ) 到聚类中心的距离  $d_{e_p^{(0)}, c_0}$ , 取  $\{d_{e_p^{(0)}, c_0} | p \in [0, N_s - 1]\}$  最大值为  $d_{\max}[0]$ , 相对应的点为  $P_0$ , 使  $S_{\text{last}}[0] = 1$ .

② 根据图 2(a) 和式(1), 对与数组  $d_{\max}[j] (j \in [0, n_c - 1])$  中最大值相应的类  $i$ (该类中偏离中心最远的像素  $e_{\max}^{(i)}$  记为  $P_i$ ), 按方向  $\overrightarrow{P_i C_i} = e_{\max}^{(i)} - c_i$ , 再用通过该类中心  $c_i$ , 并按方向  $\overrightarrow{P_i C_i}$  反向延伸  $\overline{d_{e_p^{(i)}, c_i}}$  处的点和最远点  $e_{\max}^{(i)}$  之间的中点的垂直平面来分割该类, 以产生两子类  $\tilde{C}_i$  ( $\tilde{C}_i$  取代  $C_i$  的位置, 使  $C_i = \tilde{C}_i$ ) 和  $C_{n_c}$ , 置  $S_{\text{last}}[i] = S_{\text{last}}[n_c] = 1$ , 每个子类分配空间大小为  $N_{C_i}$  和  $N_{C_{n_c}}$ , 并使  $n_c = n_c + 1$ , 再对新生的两子类分别进行如下计算(以新类  $C_i$  为例): 首先求类  $C_i$  的聚类中心  $c_i$ (式(2)), 然后计算类内像素  $e_p^{(i)} \in C_i$  到聚类中心  $c_i$  的距离  $d_{e_p^{(i)}, c_i}$ 、 $d_{\max}[i]$  和  $\overline{d_{e_p^{(i)}, c_i}}$ .

$$I = \frac{\overrightarrow{P_i C_i}}{\|\overrightarrow{P_i C_i}\|} \quad (1)$$

$\forall e_p^{(i)} \in C_i, e_p^{(i)} \neq e_{\max}^{(i)}$ , if

$$d_{e_p^{(i)}, c_i} < d_{e_{\max}^{(i)}, c_i} - \frac{d_{e_{\max}^{(i)}, c_i} + \overline{d_{e_p^{(i)}, c_i}}}{2} \Rightarrow$$

$\tilde{C}_i = C_i \cup \{e_p^{(i)}\}$  else if

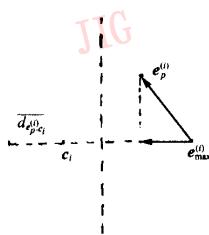
$$|(e_p^{(i)} - e_{\max}^{(i)}) \cdot I| < \frac{d_{e_{\max}^{(i)}, c_i} + \overline{d_{e_p^{(i)}, c_i}}}{2} \Rightarrow$$

$C_{n_c} = C_{n_c} \cup \{e_p^{(i)}\}$  else  $\tilde{C}_i = \tilde{C}_i \cup \{e_p^{(i)}\}$

$$c_i = \frac{1}{N_{C_i}} \sum_{p=1}^{N_{C_i}} e_p^{(i)}, e_p^{(i)} \in C_i \quad (2)$$

$$d_{e_p^{(i)}, c_i} = \|e_p^{(i)} - c_i\|$$

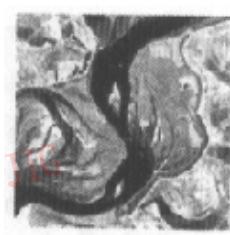
③ 若获得指定的类数  $n_c = N_c$  时, 则停止分割,



(a)  $C_i$  类分割



(b) 类 1 内像素  $e_{1,i}$  的重新聚类



(c) 灰度均衡调整后的 TM5 图象

图 2 密西西比河地区 TM 图象聚类分割

转步骤(2); 否则转步骤(1)第②步, 继续分割操作.

(2) 求各类中心之间的距离, 产生聚类中心间距离矩阵  $D = \{d_{c_i, c_j}, i, j \in [0, N_c - 1]\}$ , 记录上次聚类中心坐标值  $c_{\text{last}}^{(i)} = c_i$ , 并置当前类修改标志  $S_{\text{new}}[i] = 0 (i \in [0, N_c - 1])$ .

$$d_{c_i, c_j} = \|c_i - c_j\|$$

(3) 各类中的像素和相邻聚类中心进行重新聚类(图 2(b)):

① 对每一类  $i \in [0, N_c - 1]$  进行如下处理:  $d_{c_i, c_j} (j \in [0, N_c - 1], j \neq i)$  按由小到大排序, 以确定其他类  $j$  和当前类  $i$  的远近程度.

② 对类  $i$  中每一像素  $e_p^{(i)} \in C_i$  进行如下处理:

1) 按由近到远的顺序, 确定搜索范围  $C_{i,p}$ (对  $j \in C_{i,p}$ , 有  $j \neq i, j \in [0, N_c - 1], d_{e_p^{(i)}, c_j} > 0.5d_{c_i, c_j}$ ).

2) 若上次修改标志  $S_{\text{last}}[i] = 0$ , 且  $S_{\text{last}}[k] = 0 (k \in C_{i,p})$ , 则类  $k$  和类  $i$  的相对关系已经确定, 不用进行聚类判断, 可将  $k$  从  $C_{i,p}$  中删除.

3) 求  $e_p^{(i)}$  与所有邻近类  $j \in C_{i,p}$  的聚类中心  $c_j$  的距离  $d_{e_p^{(i)}, c_j}$ , 若与  $\{d_{e_p^{(i)}, c_j} | j \in C_{i,p}, j \neq i\} \cup \{d_{e_p^{(i)}, c_i}\}$  中最小值对应的类  $k \neq i$ , 则置类修改标志  $S_{\text{new}}[i] = 1, S_{\text{net}}[k] = 1$ , 并修改类集合:

$$\begin{aligned} C_i &= C_i \setminus \{e_p^{(i)}\}, C_k = C_k \cup \{e_p^{(i)}\} \\ d_{e_p^{(i)}, c_k} &= \|e_p^{(i)} - c_k\|, k \in C_{i,p} \end{aligned}$$

(4) 对所有修改标志  $S_{\text{new}}[i]$  为 1 的类  $i \in [0, N_c - 1]$ , 重新计算  $c_i$  及  $d_{e_p^{(i)}, c_i} (e_p^{(i)} \in C_i)$ , 若  $|c_{\text{last}}^{(i)} - c_i| < \epsilon$ , 则该类状态稳定, 置  $S_{\text{new}}[i] = 0$ .

(5) 更新上次类修改标志  $S_{\text{last}}[i] = S_{\text{new}}[i] (i \in [0, N_c - 1])$ .

(6) 若所有  $S_{\text{new}}[i] = 0 (i \in [0, N_c - 1])$ , 或已迭代足够的次数, 则聚类结束, 转步骤(7); 否则转步骤(2).

(7) 类号按类空间大小降序排列, 取与实际聚类中心最近的像素为聚类中心, 保存本次聚类结果.

## 1.2 IKA 聚类的特点

(1) 初始聚类时,对类空间分布最不规则的类,如按空间伸展方向进行分割,可基本保证类的均衡;而且随着类规模的降低,初始类分割的速度加快;

(2) 由于只对类之间相对关系有变化的类进行重新聚类,并且类内像素只和相邻的聚类中心通过计算距离来进行聚类,随着算法的迭代进行,大量类的状态基本固定,因此使得聚类过程不断加快.

## 2 多层次聚类无损压缩算法(MCLCA)

MCLCA 根据改进的 K-means 聚类算法(IKA)来实现较高聚类数时的快速均衡聚类,虽可最大限度提高类内像素冗余度,但与聚类中心的残差又存在更强的冗余,这样通过多次聚类去冗余,可实现最大限度的压缩.由于超谱图象相邻波段的相关性最强,其聚类中心可以根据相邻波段的残差来消除谱间冗余;最后的编码数据由于主要是各层聚类产生的分类地图,而此时残差数据基本在零值附近,因此编码数很少.

### 2.1 MCLCA 算法

设  $H, W, b$  为遥感图象的长、宽和波段数,  $X_{i,p,l}^{(r)}$  为第  $r$  层聚类后第  $i$  类中第  $p$  个样本的第  $l$  波段的数据,  $\{X^{(r)}\}$  为第  $r$  层聚类的像素矢量集合,  $I_{i,l}^{(r)}$  为第  $r$  层聚类产生的第  $i$  聚类中心的第  $l$  波段数据,  $N_i$  为第  $i$  类的样本数,  $N_c$  为当前聚类数,  $s_{i,j}$  为像素  $(i, j)$  的类序号,  $S^{(r)}$  是第  $r$  层聚类产生的分类地图,并设  $L$  层聚类的各层类数为  $\{n_c^{(0)}, n_c^{(1)}, \dots, n_c^{(L-1)}\}$ , 则 MCLCA 具体步骤(图 1(b))如下:

(1) 输入聚类数组合  $\{n_c^{(0)}, n_c^{(1)}, \dots, n_c^{(L-1)}\}$ , 原始遥感数据为  $\{X^{(0)}\}$ , 取  $r=0$ .

(2) 对聚类数  $N_c = n_c^{(r)}$  进行如下处理:

① 对数据  $\{X^{(r)}\}$  进行聚类数为  $N_c$  的 IKA 聚类,产生  $N_c$  个聚类中心  $I_{i,l}^{(r)}$ ,  $i \in [0, N_c - 1]$  和分类地图  $S^{(r)} = \{s_{i,j} \in [0, N_c - 1], i \in [0, H - 1], j \in [0, W - 1]\}$ .

② 求聚类中心残差

1) 求每个聚类中心的谱向残差

$$\delta I_{i,l}^{(r)} = I_{i,l}^{(r)} - I_{i,l-1}^{(r)}, i \in [0, N_c - 1], l \in [1, b - 1]$$

2) 求各类第 1 维数据的残差

$$\delta I_{i,0}^{(r)} = I_{i,0}^{(r)} - I_{i,-1,0}^{(r)}, i \in [0, N_c - 1]$$

③ 类内像素和聚类中心的残差为

$$\delta X_{i,p,l}^{(r)} = X_{i,p,l}^{(r)} - I_{i,p,l}^{(r)}$$

$$i \in [0, N_c - 1], p \in [0, N_i - 1], l \in [0, b - 1]$$

(3) 取下一聚类数  $r = r + 1$ , 并使  $\{X^{(r)}\} = \{\delta X^{(r-1)}\}$ , 转第(2)步对残差数据进一步去冗余,直到聚类数组合完成为止,或残差的标准差足够小为止.

(4) 编码:

① 对各层聚类的分类地图  $S^{(r)}$  ( $r \in [0, L - 1]$ ) 进行 Huffman 统计编码.

② 对各层聚类中心的残差数据  $\delta I_{i,l}^{(r)}$ ,  $i \in [1, b - 1], r \in [0, L - 1]$  和  $I_{i,0}^{(r)}$  ( $r \in [1, L - 1]$ ) 进行 Huffman 统计编码,  $I_{i,0}^{(r)}$  单独保存.

③ 对最后的残差数据  $\{\delta X^{(L-1)}\}$  进行 Huffman 统计编码.

### 2.2 算法特点

(1) 由于 IKA 聚类具有收敛速度快的特点,因此每一层对上次产生的残差数据可进一步通过 IKA 聚类去冗余;

(2) 编码长度主要由残差数据和分类地图彼此的消长所决定,而聚类中心编码则很少;

(3) 随层次增加,大量残差数据处于零值附近,冗余度急剧上升,由于该算法可使残差数据被聚在少数的几个类中,并使分类地图中类序号小的数据占大多数,因此可有效提高 Huffman 编码效率,这不仅使分类地图增加的速度趋缓,也使残差数据编码下降幅度趋于收敛.

## 3 S+P 变换多层次聚类压缩

### (SPMCLCA)

由于 S+P(Sequential transform+Prediction)整数小波变换则只是简单的移位和加法操作,速度很快,因此比一般小波变换更适于消除超光谱遥感数据谱间冗余和空间冗余. 设数据长度为  $N, j \in [0, \lfloor \log_2 N \rfloor]$  为当前分解层数,  $s_{j,l}$  为第  $j$  层未分解的第  $l$  个数据,  $s_{j-1,l}$  为第  $j$  层分解后的第  $l$  个低频分量,  $d_{j-1,l}^{(1)}$  为第  $j$  层分解后的第  $l$  个高频分量,  $d_{j-1,l}$  为  $d_{j-1,l}^{(1)}$  的修正值,则提升方法的整数小波变换(S+P 变换)步骤<sup>[5]</sup>如下:

(1) 当数据长度为偶数即  $s_{j,i}$  ( $i \in [0, 2n - 1]$ ) 时,变换后数据区  $[0, n - 1]$  为  $s_{j-1}, [n, 2n - 1]$  为  $d_{j-1}$

$$d_{j-1,l}^{(1)} = s_{j,2l+1} - s_{j,2l}, l \in [0, n - 1]$$

$$s_{j-1,l} = s_{j,2l} + \lfloor d_{j-1,l}^{(1)} / 2 \rfloor, l \in [0, n - 1]$$

(2) 当数据长度为奇数,即  $s_{j,i}$  ( $i \in [0, 2n]$ ) 时,变换后数据区  $[0, n]$  为  $s_{j-1}$ ,  $[n+1, 2n]$  为  $d_{j-1}$ ;

$$\begin{aligned} d_{j-1,l}^{(1)} &= s_{j,2l+1} - s_{j,2l}, l \in [0, n-1] \\ \left\{ \begin{array}{l} s_{j-1,l} = s_{j,2l} + \lfloor d_{j-1,l}^{(1)}/2 \rfloor \\ s_{j-1,n} = s_{j,2n} \end{array} \right. \end{aligned}$$

(3)  $d_{j-1}$  修正;

$$\begin{aligned} d_{j-1,l} &= d_{j-1,l}^{(1)} + \lfloor \alpha_{-1}(s_{j-1,l-2} - s_{j-1,l-1}) + \\ &\quad \alpha_0(s_{j-1,l-1} - s_{j-1,l}) + \alpha_1(s_{j-1,l} - s_{j-1,l+1}) - \\ &\quad \beta_1 d_{j-1,l+1}^{(1)} \rfloor \end{aligned}$$

$$A: \alpha_{-1} = \beta_1 = 0, \alpha_0 = \alpha_1 = 0.25$$

由于小波变换具有多分辨率特性,因此变换后的图象数据能够保持在各种分辨率下原始的精细结构。对于多光谱图象,在变换后的系数图象中,存在谱间结构相关性,即各波段平坦区域的系数都较小,而细节丰富区域系数都较大。文献[5]指出变换后波段矢量可分为 4 类,并可对各类分别构造谱间预测器。SPMCLCA 算法可对 S+P 聚类压缩算法<sup>[5]</sup>产生的残差数据进行 MCLCA 迭代处理。实验表明,SPMCLCA 算法压缩效率更高。其算法具体步骤如下:

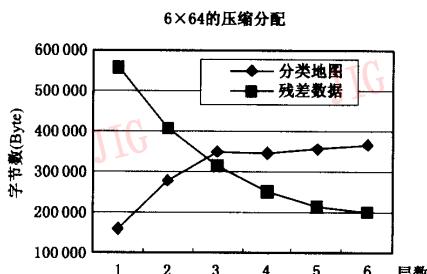
(1) 首先对多光谱各波段数据进行二维 S+P 整数小波变换,以消除空间冗余。小波分解层数为  $\lceil \log_2 N \rceil$  (其中  $N$  为一维数据长度);

(2) 取 MCLCA 第 1 层聚类数为 4,可有效消除结构冗余;

(3) 提高聚类层次,进一步消除残差数据的冗余。

## 4 实验

选取  $512 \times 512$  (8bit) 密西西比河地区 TM 遥感图象(网上获取,图 2(c)),删去受大气影响比较大的第 6 波段数据,用其余的 6 个波段数据进行算法



(a) MCLCA 残差数据和分类地图

实验. 实验分析中,取 IKA 迭代 20 次,  $\epsilon = 0.02$ , MCLCA/SPMCLCA 分 6 层聚类,压缩比  $P$  定义为

$$P = \frac{N_0}{N_1} (\text{倍})$$

其中,  $N_0$  为原数据量,  $N_1$  为压缩后数据量。

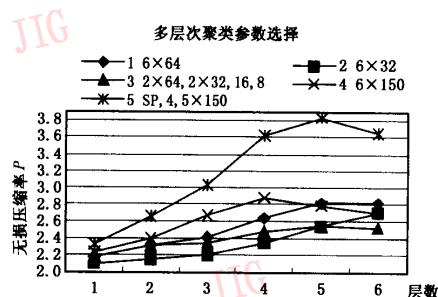
取 MCLCA 6 层聚类数为  $6 \times 64$  (即  $64, 64, 64, 64, 64$ ),经 IKA 快速聚类,由于各层产生的残差标准差  $\sigma$  为  $1.722\ 023, 0.994\ 301, 0.678\ 550, 0.438\ 375, 0.232\ 927, 0.113\ 205$ ,因此分层聚类可最大限度消除遥感数据的冗余。由图 3(a)可看出,残差数据和分类地图的编码数随层数的上升趋于收敛,第 4 层的无损压缩字节数为 557 215,无损压缩率为 2.822 7。

由不同 MCLCA 聚类数组合(图 3(b)中曲线 1~4)可看出:各层聚类数愈高,压缩效率变化愈快,且存在极值。如 3 层 150 类的压缩率为 2.882,但此后继续分层聚类,分类地图增加的速度比残差编码数据减少得要快,且压缩率下降。图 3(b)中曲线 5 是经 SPMCLCA 6 层聚类数为 {4, 150, 150, 150, 150, 150} 压缩的结果,各层标准差为 2.835 850、1.154 751、0.697 969、0.392 355, 0.156 175、0.049 571, 第 5 层压缩率达到 3.842,用时 547s。

表 1 是 SPMCLCA、MCLCA 和 DPCM<sup>[6]</sup>、S+P<sup>[5]</sup>压缩算法相比较的结果,由表 1 可见,本文提出的算法是非常合理的。该压缩方案为遥感图象、视频图象处理提供了一个新的思路,但具体的参数(如各层类数)选择还有待进一步深入的研究。

表 1 无损压缩结果比较

算法	DPCM	S+P	MCLCA (4×150)	SPMCLCA (4,4×150)
无损压缩率 $P$	1.478	2.484	2.882	3.842



(b) 聚类数组合的比较(MCLCA 与 SPMCLCA)

图 3 Mississipi TM 图象压缩

## 5 结 论

基于多层次聚类的无损压缩算法具有如下优点:采用改进的 K-means 聚类算法来进行聚类,由于其充分利用了空间几何关系,像素只和近邻聚类中心进行聚类判断,因此明显提高了迭代收敛速度;多层次聚类方法可完全消除残差数据的冗余,故可分为提高压缩率创造条件;每层聚类数愈多,压缩率收敛速度愈快,并存在极值情况。另外,基于 S+P 小波变换的多层次聚类,不仅能消除空间相关、结构相关,而且能清除残差数据的冗余,从而实现了遥感图象无损压缩率的突破。实验表明,基于 IKA 和 S+P 整数小波变换的多层次聚类有着实现简单、收敛速度快、压缩率高等优点,在遥感图象压缩方面有着良好的应用前景,但该压缩方案在一些参数选择上还有待进一步深入的研究。

### 参 考 文 献

- 1 倪林. 基于地物反射特性的多波段遥感图象无损压缩方法[J]. 遥感技术与应用, 1999, 14(4): 53~58.

- 2 张荣, 刘政凯. 一种多光谱遥感图象的近无损压缩方法[J]. 中国图象图形学报, 1998, 3(10): 823~826.
- 3 张荣, 刘政凯, 李厚强. 基于分类的多波段遥感图象无损压缩方法[J]. 中国图象图形学报, 1998, 3(2): 106~109.
- 4 倪林. 基于分类 K-L 变换的多波段遥感图象近无损压缩方法[J]. 遥感学报, 2001, 5(3): 205~213.
- 5 张荣, 俞能海, 刘政凯. 基于整数小波变换的多光谱图象无损压缩[J]. 遥感学报, 2001, 5(1): 29~33.
- 6 王学良. 基于 HRIS 光谱图像帧序列相关性的 D2PCM 无损压缩方法[J]. 遥感学报, 2001, 5(2): 119~121.



**王朝晖** 1970 年生, 1993 年获山东大学电子学与信息系统专业学士学位, 2002 年获中国科学技术大学电路与系统专业硕士学位。主要从事图象压缩与编码、人工智能、多媒体通讯等方面研究。发表论文 5 篇。