

# 基于谱聚类 and 稀疏表示的高光谱图像分类算法

董安国<sup>1</sup>, 李佳逊<sup>1\*</sup>, 张 蓓<sup>1</sup>, 梁苗苗<sup>2</sup>

<sup>1</sup>长安大学理学院, 陕西 西安 710064;

<sup>2</sup>西安电子科技大学电子工程学院, 陕西 西安 710071

**摘要** 为了增强高光谱遥感图像的分类效果, 提出基于谱聚类和稀疏表示的两级分类算法。利用谱聚类将待分类的像元及其邻域内所有的像元分成两类, 利用联合稀疏表示模型确定按规则选取的其中一类的具体类别, 并以该类别作为像元的类。该算法充分利用高光谱图像的光谱及空间信息, 两级分类过程均考虑了噪声及区域边界对分类效果的影响。进一步利用空间信息对分类算法进行修正, 即关联邻近像元的类别, 平滑分类结果。数值实验表明, 该算法的分类精度高、稳定性好、抗噪性强。

**关键词** 遥感; 高光谱遥感图像; 遥感图像分类; 联合稀疏表示; 谱聚类

**中图分类号** TP751.1 **文献标识码** A

**doi:** 10.3788/AOS201737.0828005

## Hyperspectral Image Classification Algorithm Based on Spectral Clustering and Sparse Representation

Dong Anguo<sup>1</sup>, Li Jiaxun<sup>1</sup>, Zhang Bei<sup>1</sup>, Liang Miaomiao<sup>2</sup>

<sup>1</sup>School of Science, Chang'an University, Xi'an, Shaanxi 710064, China;

<sup>2</sup>School of Electronic Engineering, Xidian University, Xi'an, Shaanxi 710071, China

**Abstract** In order to improve classification effect of hyperspectral image, a classification algorithm with two levels is proposed based on spectral clustering and sparse representation. Pixels to be classified and its neighborhood pixels are divided into two parts by spectral clustering. The class of selected pixels is identified by the joint sparse representation model. This algorithm makes full use of hyperspectral image spectral and spatial information of hyperspectral images, and both of the two levels. Finally, the proposed algorithm is corrected with the spatial information, namely, neighboring pixels' class is associated and classification results is smoothed. Numerical experiments demonstrate that this algorithm has high classification accuracy, good stability and anti-noise performance.

**Key words** remote sensing; hyperspectral remote sensing image; remote sensing image classification; joint sparse representation; spectral clustering

**OCIS codes** 280.4788; 100.4145; 100.3008; 100.2960

## 1 引 言

高光谱遥感技术目前在资源勘探、环境监测、精细农业、灾害评估以及目标识别等领域发挥着重要的作用。高光谱遥感图像<sup>[1]</sup>包含从可见光到红外光谱区域的数百个狭窄且连续波段的光谱信息和空间信息。高光谱图像中的每个像元均包含几十维甚至几百维的光谱向量, 不同的地物在不同波段下的光谱响应值不同, 进而呈现出不同的光谱向量, 因此可以根据高光谱图像中的光谱向量确定每个像元的具体类别。

近年来, 多种图像处理技术广泛应用于高光谱遥感图像的分类。在现有的分类算法中, 均值聚类算

**收稿日期:** 2017-03-29; **收到修改稿日期:** 2017-04-18

**基金项目:** 国家自然科学基金(41601437, 41571346, 11201038)

**作者简介:** 董安国(1964—), 男, 教授, 硕士生导师, 主要从事数值代数、数字图像处理方面的研究。

E-mail: donganguo@chd.edu.cn

\* **通信联系人.** E-mail: 15637793688@163.com

法<sup>[2-3]</sup>是一种无监督的分类算法,该算法的不足之处在于初始聚类中心和最佳聚类数对结果影响很大;支持向量机(SVM)<sup>[4-5]</sup>作为一种高维有监督的分类算法,凭借其不受 Hughes 现象影响的优势,取得了很好的分类效果;被大量研究的神经网络算法<sup>[6]</sup>则通过特征选择或特征变换等进行降维,然后通过合适的神经网络类型、神经网络层次、各层神经元数目、传输函数以及学习算法进行分类,可以得到具有高维性、自组织性和冗余性等优势的分类效果;而稀疏表示算法<sup>[7]</sup>是将每一个待分类测试样本作为训练样本的线性组合,然后求解稀疏系数,并通过计算该测试样本与各类训练样本的最小残差确定其类别。然而,这些算法都只应用了高光谱图像的光谱信息,一般会造成分类结果的不平滑性<sup>[8]</sup>。

高光谱图像在空间分布上具有一致性,即局部区域内的像素通常具有相同的材质和光谱特征,因此,光谱信息和空间信息相融合的分类算法受到越来越多的关注。在原有的光谱信息的基础上考虑空间邻域的信息,Chen 等<sup>[9]</sup>提出了联合稀疏表示(JSR)模型,将邻域内的所有像元利用相同训练样本表示,这些像元将被标记为同样的标签来保证图像的空间平滑性;文献<sup>[10]</sup>中,多尺度自适应稀疏表示(MASR)能够有效地利用空间信息,在多尺度空间通过自适应的稀疏矩阵对高光谱图像进行分类;Camps-Valls 等<sup>[11]</sup>在 SVM<sup>[12]</sup>基础上进行了复合内核的改进,提出基于复合内核的支持向量机(CK-SVM)算法,研究表明,采用光谱信息和空间信息相融合的算法,能有效提高高光谱图像的分类精度。

然而,由于高光谱图像地物种类的复杂性,空间分布的一致性并非普遍存在。针对该问题,本文提出一种基于谱聚类和稀疏表示的高光谱图像两级分类算法。首先,通过谱聚类的算法将当前像元及其邻域内所有的像元分成两部分;然后,在一定的规则下选择其中一部分像元进行稀疏矩阵的求解;最后,通过计算该部分像元与各类训练样本的最小残差,确定其具体类别。数值实验结果表明,该算法不仅能够提高高光谱图像分类的各项评价指标,而且能够提高分类结果的平滑性。

## 2 稀疏表示理论

稀疏表示是一种针对遥感图像分类非常有效的手段<sup>[13]</sup>,能够取得较好的分类效果。给定训练样本字典  $D$ ,可以将求解像元  $x$  被字典  $D$  表示的稀疏系数  $\alpha$  的问题转化为稀疏表示<sup>[14]</sup>的优化模型,即有

$$\hat{\alpha} = \arg \min_{\alpha} \|\alpha\|_0, \quad \text{s.t.} \quad \|x - D\alpha\|_2 \leq \epsilon, \quad (1)$$

式中  $x$  为测试样本像元,  $\epsilon$  为容错误差。

求出(1)式的解  $\hat{\alpha}$  后,即可得到各像元的类别<sup>[7]</sup>。

在高光谱图像中,存在着同类异谱、异类同谱的现象,如图 1 所示。图 1(a)为地物的实际类别;图 1(b)为利用正交匹配追踪(OMP)算法得到的分类结果;图 1(c)为点 A、B、C 的光谱曲线图。相同类别的点 A 和 B,其光谱曲线差异较大,即同类异谱;而不同类别的点 B 和 C 的光谱曲线很接近,属异类同谱。

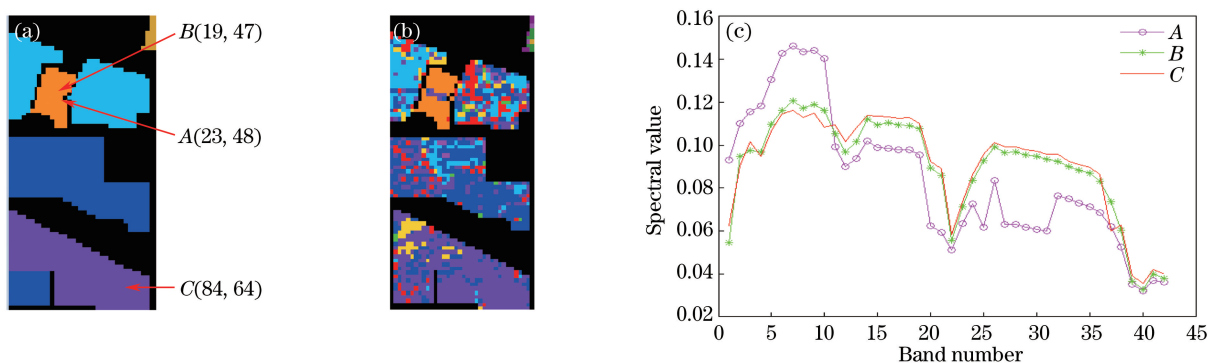


图 1 高光谱图像。(a)实际地物类别;(b)OMP 算法的分类结果;(c)光谱曲线图

Fig. 1 Hyperspectral images. (a) Category of real ground; (b) classification result of OMP algorithm; (c) spectral curves

在高光谱图像中同类异谱、异类同谱的现象普遍存在,对单个像元进行稀疏表示时,只考虑光谱信息,分类结果会出现图 1(b)所示的现象,因此可以考虑在光谱信息的基础上加入空间信息。为此,利用联合稀疏表示模型对当前像元及其整个邻域进行稀疏表示,来确定当前像元的类别。

设地物类别数为  $M$ , 训练样本数量为  $N$ , 记  $S = \{1, \dots, N\}$ ,  $S = S_1 + S_2 + \dots + S_M$ , 其中  $S_m$  为第  $m$  类训练样本 ( $m = 1, 2, \dots, M$ ), 设  $\mathbf{d}_i$  为第  $i$  个训练样本的光谱向量 ( $i \in S$ ), 字典矩阵  $\mathbf{D} = (\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_N)$ 。设  $X = \{x_1, x_2, \dots, x_t\}$  为当前待分类像元邻域内的像元集 ( $t$  为邻域内像元总数), 此时, 关于  $X$  的联合稀疏表示模型为

$$\hat{\mathbf{E}} = \arg \min_{\mathbf{E}} \|\mathbf{X} - \mathbf{D}\mathbf{E}\|_F, \quad \text{s.t.} \quad \|\mathbf{E}\|_{\text{row},0} \leq K, \quad (2)$$

式中  $\mathbf{E} = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_t]$ ,  $\boldsymbol{\alpha}_j$  为  $X$  中第  $j$  个像元的稀疏系数向量 ( $j = 1, 2, \dots, t$ ),  $\|\mathbf{E}\|_{\text{row},0}$  表示稀疏矩阵中非零行的个数。

联合稀疏表示模型<sup>[8]</sup>可用同步正交匹配追踪(SOMP)算法实现<sup>[9]</sup>。

输入: 训练样本字典  $\mathbf{D}$ , 邻域内像元集  $X$ , 支撑索引集初始值  $\Lambda_0 = \emptyset$ , 迭代次数(稀疏度)  $K$ , 迭代初始值  $k = 1$ , 初始残差  $R_0 = X$ ;

输出: 稀疏矩阵  $\hat{\mathbf{E}}$ 。

对每一个测试像元, 进行如下步骤:

- 1) 如果迭代次数  $k < K$ , 转步骤 2); 否则, 转步骤 5);
- 2) 确定与残差最接近的原子  $\lambda_k = \arg \min_{i=1,2,\dots,N} \|\mathbf{R}_{k-1}^T \mathbf{d}_i\|_F$ , 更新索引集  $\Lambda_k = \Lambda_{k-1} \cup \lambda_k$ ;
- 3) 更新残差  $R_k = X - \mathbf{D}_{\Lambda_k} [(\mathbf{D}_{\Lambda_k}^T \mathbf{D}_{\Lambda_k})^{-1} \mathbf{D}_{\Lambda_k}^T X]$ ,  $\mathbf{D}_{\Lambda_k}$  为字典中索引集所在位置的训练样本;
- 4) 更新迭代次数  $k = k + 1$ , 转步骤 1);
- 5) 求得稀疏矩阵  $\hat{\mathbf{E}} = \mathbf{D}_{\Lambda_k} [(\mathbf{D}_{\Lambda_k}^T \mathbf{D}_{\Lambda_k})^{-1} \mathbf{D}_{\Lambda_k}^T X]$ 。

将整个邻域作为一个整体并代替之前的单个像元, 加入空间信息后更符合遥感图像的实际情况。若测试样本集中有噪声点, 可以减弱噪声点造成的影响, 降低噪声点处分错的频率, 进而提高整个图像的分类精度。若测试样本集包含边界, 那么邻域内的像元有多种类别, 鉴于该情况, 提出一种与谱聚类相结合的稀疏表示算法对地物进行分类。

### 3 谱聚类和稀疏表示相融合的高光谱图像分类算法

邻域内地物有两类或两类以上时, 若考虑对整个邻域的分类, 会影响噪声点和边界处的分类效果。因此, 提出基于谱聚类和稀疏表示的两级分类算法。首先, 考虑以当前像元为中心的邻域内各像元间的相关性, 并通过谱聚类算法将像元分为两类, 再根据具体规则选择其中一类像元, 即第一级分类; 然后, 建立联合稀疏表示模型对所选像元进行第二级分类。

#### 3.1 谱聚类

利用谱聚类算法进行聚类时, 将待聚类的对象集作为节点集构建赋权图, 然后通过分析与节点相关的相似度矩阵特征值和特征向量得到聚类结果。

邻域内像元  $\{x_1, x_2, \dots, x_t\}$  作为赋权图中的节点集  $V$ , 将两像元  $x_i, x_j$  ( $i, j = 1, 2, \dots, t$ ) 之间的相似度  $g_{i,j}$  作为两节点之间的赋权重, 定义相似度  $g_{i,j} = [\text{cov}(x_i, x_j)] / [\sqrt{D(x_i)D(x_j)}]$ , 记  $\mathbf{G} = (g_{i,j})_{t \times t}$ , 此时可以得到一个无向图  $(V, \mathbf{G})$ , 邻域内像元分为两类的问题已转化为无向图的最小割问题, 即求得相似度矩阵  $\mathbf{G}$  的特征值和特征向量的问题。通过分析特征向量, 即可解决像元的分类问题。

邻域内各像元的光谱向量按列排放, 计算中间像元与其他像元间的相似度, 根据相似程度, 分为两种情况。

第一种情况, 当邻域内只有一类地物时, 中间像元和其他像元的相似度均较高, 当最小相似度高于参数  $\delta$  时, 可以认为邻域内的像元为同一类, 因此可以省略谱聚类步骤, 直接对整个邻域进行稀疏表示。

第二种情况, 当部分相似度低于参数  $\delta$  时, 认为邻域内的像元来自不同的类别, 因此需要将邻域内所有像元分为两类。像元的类别和谱(特征值)的联系较大, 可以对所得的特征向量空间数据进行  $K$ -mean 聚类, 那么像元的聚类结果与此时的聚类结果相同。具体算法为: 首先求得相似度矩阵  $\mathbf{G}$  的前两个特征值及相对应的特征向量(特征值按照从大到小的顺序排列); 然后将这两个特征向量按列排放构成一个  $t \times 2$  的矩阵, 将该矩阵的每一行作为一个新的数据点, 并对这  $t$  个数据点进行  $K$ -mean 聚类; 最终将邻域内的像元分为两类。

### 3.2 基于谱聚类和稀疏表示的两级分类算法

将像元分为两类之后,依据两类个数的差值,按照以下规则选出其中一类像元:

1) 两类个数相差较大时,除去邻域内噪声点后,认为其余像元为同一类,因此选择个数多的一类进行联合稀疏表示可以避免噪声点对最终分类结果的影响;

2) 两类个数相差较小时,邻域内的地物来自不同类别,因此选取当前像元所在的类别进行联合稀疏表示。

图 2 以 Indian Pines 地区为实例,列举以上两种情况的聚类结果。

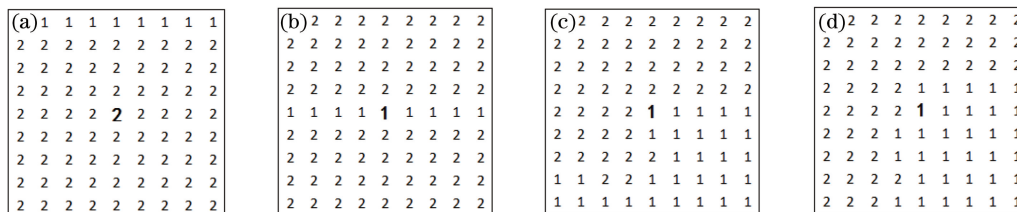


图 2 邻域内地物的聚类结果。(a) (129,35);(b) (96,39);(c) (38,52);(d) (100,58)

Fig. 2 Clustering results of ground in the neighborhood. (a) (129,35); (b) (96,39); (c) (38,52); (d) (100,58)

图 2(a)~(d)分别给出了 4 个像元(129,35)、(96,39)、(38,52)、(100,58)的邻域。由图 2 可知,谱聚类的分类结果与实际情况一致,选择像元时,图 2(a)、(b)选择个数较多的一类像元,图 2(c)、(d)选择与当前像元同类的像元,作为下一步稀疏表示的测试样本集。

谱聚类和稀疏表示相融合的高光谱图像分类算法具体的算法流程如下( $\delta, \beta$  为参数):

1) 记  $U(x)$  为像元  $x$  的邻域,并计算相似度矩阵  $G$ ;

2) 若  $G$  中的所有元素均大于  $\delta$ ,则  $T=U(x)$ ,转步骤 4),否则利用谱聚类将  $U(x)$  分为两类,记  $U_1$  和  $U_2$  (其中  $x \in U_1$ );

3) 若  $||U_1| - |U_2|| < \beta$ ,则  $T=U_1$ ,若  $|U_1| - |U_2| \geq \beta$ ,则  $T=U_1$ ,若  $|U_1| - |U_2| \leq -\beta$ ,则  $T=U_2$ ;

4) 以  $T$  为整体进行联合稀疏表示(见 SOMP 算法),求得稀疏矩阵  $\hat{E}$ ;

5) 计算  $T$  被第  $m$  类字典表示的残差

$$r^m(T) = \| T - D_{S_m} \hat{E}_{S_m} \|_F, m = 1, 2, \dots, M, \quad (3)$$

式中  $D_{S_m}$  和  $\hat{E}_{S_m}$  分别为  $D$  和  $\hat{E}$  中取下标集  $S_m$  列和  $S_m$  行所得的矩阵,即  $D_{S_m} = D(:, S_m), \hat{E}_{S_m} = \hat{E}(S_m, :)$ ,  $m = 1, 2, \dots, M$ ;

6)  $T$  的地物类别为  $\text{class}(T) = \arg \min_{m=1, \dots, M} r^m(T)$ ,并将其作为当前像元的类别。

## 4 算法修正

上述算法对同谱异类、异谱同类问题有很大改善,但又不能完全避免。因此在所有像元分类结束后,还需利用空间信息对图像中邻近像元的类别进行关联,使得图像分类结果趋于平滑。具体算法为:在像元  $x$  的邻域  $U(x)$  内搜索所有与像元  $x$  谱聚类结果相同的点,记为  $C(x)$ ,并统计  $C(x)$  中属于第  $m$  类地物的像元个数,记为  $\text{num}(x, m)$ ,由于  $C(x)$  中的像元光谱特征相似,因此当前像元  $x$  的最终地物类别可表示为

$$L(x) = \arg \max_{m=1, \dots, M} [\text{num}(x, m)]. \quad (4)$$

图 3 给出了算法修正前后的对比图,修正后的图像离散错分样本数明显减少,如图 3(b)所示,边界的轮廓更加清晰,比修正前的图像图 3(a)更平滑。



图 3 算法修正前后的对比图。(a)修正前的图像;(b)修正后的图像

Fig. 3 Contrast figures before and after algorithm correction. (a) Before correction; (b) after correction



## 5 实验验证及分析

为了验证本文算法的有效性,分别采用 SVM、CK-SVM<sup>[15]</sup>、OMP、SOMP、MASR 以及本文算法 SC-SOMP 对三组高光谱遥感图像数据集进行分析和对比。主要通过总分类精度(OA)和 Kappa 系数两个指标进行分类效果评价。每组实验均进行 10 次并对结果取平均值,三组实验所用数据都是将原始的高光谱图像数据归一化后得到。

### 5.1 实验数据

#### 5.1.1 实验 1

实验 1 中遥感图像的大小为 610 pixel×340 pixel,空间分辨率为 1.3 m,谱段范围为 0.43~0.86  $\mu\text{m}$ 。除去原始的 115 个波段中的 12 个低信噪比波段后,选取剩余的 103 个波段作为研究对象。整个图像包含 9 类地物,各类样本个数相差较大,若按比例抽取,样本少的地物会被样本多的地物表示,为此,实验中随机抽取每类样本的 250 个像元作为训练样本,其余的样本作为测试样本,但总的训练样本个数仍不超过总样本的 10%,邻域大小为  $9\times 9$ ,参数  $\delta=0.99$ , $\beta=0.375$ 。表 1 给出了 9 类地物的实验数据和分类精度,最优结果用粗体表示。图 4 为分别利用 6 种算法得到的 Pavia University 地区的高光谱遥感图像分类结果。

表 1 Pavia University 地区实验数据和分类精度

Table 1 Experimental data and classification accuracies of the Pavia University dataset

Class	Sample		Classification algorithm					
	Train	Test	SVM	CK-SVM	OMP	SOMP	MASR	SC-SOMP
Asphalt	250	6381	80.50	<b>97.90</b>	49.72	66.27	77.26	91.87
Meadows	250	18339	84.48	98.95	62.36	92.32	96.62	<b>99.11</b>
Gravel	250	1849	78.91	93.77	63.00	96.76	99.18	<b>99.78</b>
Trees	250	2814	96.24	98.96	84.49	94.94	96.91	<b>98.33</b>
Painted metal sheets	250	1095	99.74	<b>100.00</b>	99.12	99.13	<b>100.00</b>	99.36
Bare soil	250	4779	83.96	97.06	54.12	92.73	98.74	<b>100.00</b>
Bitumen	250	1080	91.39	99.56	83.46	99.16	<b>99.99</b>	99.91
Self-blocking bricks	250	3432	81.27	96.44	62.27	90.06	96.18	<b>98.08</b>
Shadows	250	697	98.44	<b>99.87</b>	95.18	85.02	83.59	89.67
OA / %			84.98	<b>98.16</b>	63.55	88.77	93.86	<b>98.03</b>
Kappa			0.80	<b>0.98</b>	0.54	0.85	0.92	<b>0.97</b>

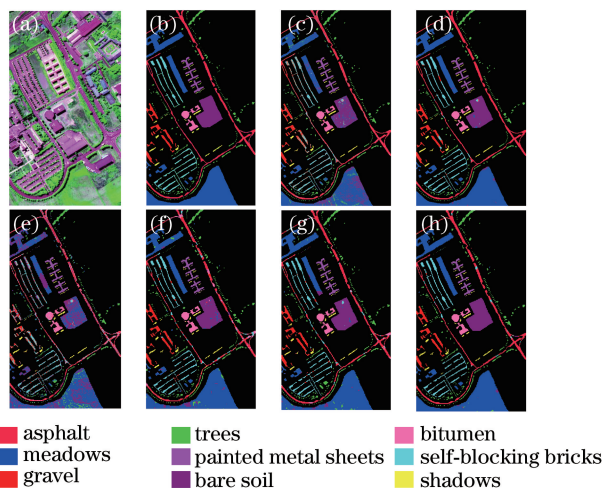


图 4 Pavia University 地区各算法的分类结果。(a)原图;(b)实际地物;(c) SVM 算法;(d) CK-SVM 算法;(e) OMP 算法;(f) SOMP 算法;(g) MASR 算法;(h) SC-SOMP 算法

Fig. 4 Classification results of Pavia University dataset obtained by different algorithms.

(a) Original image; (b) real ground; (c) SVM algorithm; (d) CK-SVM algorithm; (e) OMP algorithm; (f) SOMP algorithm; (g) MASR algorithm; (h) SC-SOMP algorithm

5.1.2 实验 2

实验 2 中高光谱遥感图像的大小为 145 pixel×145 pixel,空间分辨率为 20 m,谱间分辨率为 10 nm,谱段范围为0.4~2.5 μm。除去原始 220 个波段中的 20 个水汽吸收和低信噪比波段后,选取剩余的 200 个波段作为研究对象。整个图像包含 16 类地物,实验中随机抽取每类样本的 10%作为训练样本,其余的 90%样本作为测试样本,邻域大小为 5×5,参数  $\delta=0.99, \beta=0.375$ 。表 2 给出了 6 种算法的分类精度。图 5 为分别利用 6 种算法得到的 Indian Pines 地区的高光谱遥感图像分类结果。

表 2 Indian Pines 地区各算法分类精度  
Table 2 Classification accuracies of Indian Pines dataset obtained by different algorithms

Class	Classification algorithm					
	SVM	CK-SVM	OMP	SOMP	MASK	SC-SOMP
OA /%	77.64	94.86	61.01	90.88	98.41	98.37
Kappa	0.74	0.94	0.66	0.90	0.98	0.97



图 5 Indian Pines 地区各算法的分类结果。(a)原图;(b)实际地物;(c)SVM 算法;(d) CK-SVM 算法;  
(e) OMP 算法;(f) SOMP 算法;(g) MASR 算法;(h) SC-SOMP 算法

Fig. 5 Classification results of Indian Pines dataset obtained by different algorithms. (a) Original image; (b) real ground; (c) SVM algorithm; (d) CK-SVM algorithm; (e) OMP algorithm; (f) SOMP algorithm; (g) MASR algorithm; (h) SC-SOMP algorithm

5.1.3 实验 3

实验 3 的高光谱遥感图像大小为 512 pixel×217 pixel,高空间分辨率为 3.7 m,图像除去原始波段中的 20 个水汽波段后,选取剩余的 204 个波段作为研究对象。整个图像包含 16 类地物,实验中随机抽取每类样本的 1%作为训练样本,其余的 99%样本作为测试样本,邻域大小为 9×9,参数  $\delta=0.99, \beta=0.46$ 。表 3 给出了 6 种算法的分类精度。图 6 为分别利用 6 种算法得到的 Salinas Valley 地区的高光谱遥感图像分类结果。

表 3 Salinas Valley 地区各算法的分类精度  
Table 3 Classification accuracies of Salinas Valley dataset obtained by different algorithms

Class	Classification algorithm					
	SVM	CK-SVM	OMP	SOMP	MASR	SC-SOMP
OA /%	86.29	94.56	82.17	88.81	88.04	98.34
Kappa	0.85	0.94	0.8	0.88	0.87	0.97

5.2 实验结果及分析

实验结果表明,利用仅考虑光谱信息的分类算法,结果中出现了较多的错误离散样本点,如 OMP、SVM 分类结果图;加入空间信息后,分类精度明显比仅考虑光谱信息的 OMP 算法提高了 16%以上,比基于内核的 SVM 算法的总分类精度提高了 10%;CK-SVM 算法在 SVM 算法的基础上对内核进行改进,大幅提高了

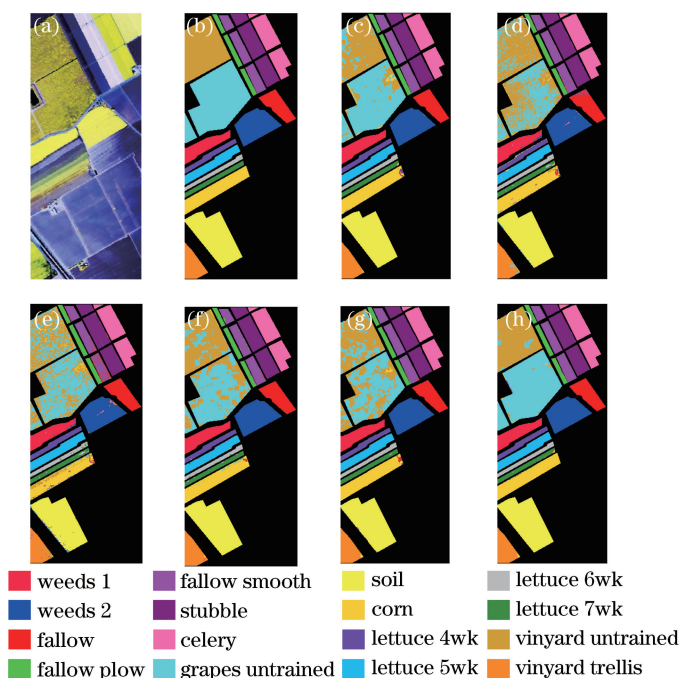


图 6 Salinas Valley 地区各算法的分类结果。(a)原图;(b)实际地物;(c) SVM 算法;(d) CK-SVM 算法;(e) OMP 算法;(f) SOMP 算法;(g) MASR 算法;(h) SC-SOMP 算法

Fig. 6 Classification results of Salinas Valley dataset obtained by different algorithms.

(a) Original image; (b) real ground; (c) SVM algorithm; (d) CK-SVM algorithm; (e) OMP algorithm; (f) SOMP algorithm; (g) MASR algorithm; (h) SC-SOMP algorithm

高光谱遥感图像的分类精度;在实验 2 和实验 3 中,SC-SOMP 算法比 CK-SVM 算法的精度均提高 4% 左右,实验 1 中 CK-SVM 算法的分类精度稍高于 SC-SOMP 算法,但 SC-SOMP 算法各类的分类精度优于 CK-SVM 算法各类的分类精度。综上所述,相比于其他分类算法,本文提出的 SC-SOMP 算法的总分类精度和 Kappa 系数以及各类地物的精度均有提高,该算法有效减少了内部分错离散样本数,边界处地物的分类平滑性也有较大改善。

为测试训练样本个数对分类结果的影响,分别比较各算法在不同训练样本个数条件下的分类精度。结果表明,随着训练样本数的增加,各算法的分类精度都随之提高,但基于空间信息和光谱信息的分类算法优于仅考虑光谱信息的分类算法。图 7 所示为训练样本个数对 OA 影响的实验结果。

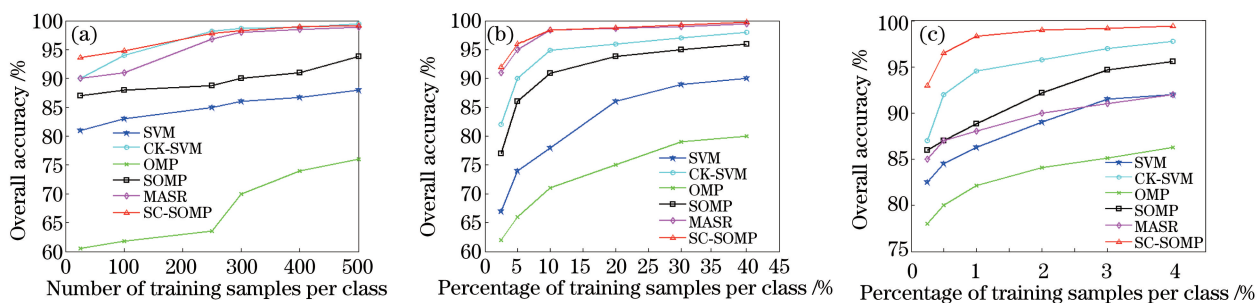


图 7 训练样本个数的影响。(a) Pavia University;(b) Indian Pines;(c) Salinas Valley

Fig. 7 Effect of the number of training samples. (a) Pavia University; (b) Indian Pines; (c) Salinas Valley

## 6 结 论

针对高光谱遥感图像的分类问题,提出一种新的谱聚类和稀疏表示相结合的分类算法,该算法充分考虑高光谱图像光谱特征,同时对局部邻域内的像元进行谱聚类算法分类,在空间信息上加入谱聚类后,对空间

邻域内的像元进行进一步判断,选取其中一类像元联合进行稀疏表示。数值实验结果表明,与其他算法相比,该算法不仅提高了分类精度,而且减少了离散错分样本数,使分类结果更平滑。

## 参 考 文 献

- [1] Liu Dawei, Han Ling, Han Xiaoyong. High spatial resolution remote sensing image classification based on deep learning[J]. *Acta Optica Sinica*, 2016, 36(4): 0428001.  
刘大伟, 韩 玲, 韩晓勇. 基于深度学习的高分辨率遥感影像分类研究[J]. *光学学报*, 2016, 36(4): 0428001.
- [2] Liang J Y, Zhao X W, Li D Y, *et al.* Determining the number of clusters using information entropy for mixed data[J]. *Pattern Recognition*, 2012, 45(6): 2251-2265.
- [3] Ishidoshiro N, Yamaguchi Y, Noda S, *et al.* Geological mapping by combining spectral unmixing and cluster analysis for hyperspectral data [J]. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2016, XLI-B8: 431-435.
- [4] Xu Jie, Zhang Jianqi, Liu Delian, *et al.* Classification of hyperspectral image with FSVM[J]. *Optical Technique*, 2008, 34(s1): 138-140.  
徐 杰, 张建奇, 刘连, 等. 基于模糊支持向量机的高光谱图像分类[J]. *光学技术*, 2008, 34(s1): 138-140.
- [5] Gao L R, Li J, Khodadadzadeh M, *et al.* Subspace-based support vector machines for hyperspectral image classification [J]. *IEEE Geoscience & Remote Sensing Letters*, 2015, 12(12): 349-353.
- [6] Ratle F, Camps-Valls G, Weston J. Semisupervised neural networks for efficient hyperspectral image classification[J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2010, 48(5): 2271-2282.
- [7] Fang L Y, Li S T, Kang X D, *et al.* Spectral-spatial hyperspectral image classification via multiscale adaptive sparse representation[J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2014, 52(12): 7738-7749.
- [8] Zhang H Y, Li J Y, Huang Y C, *et al.* A nonlocal weighted joint sparse representation classification method for hyperspectral imagery[J]. *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, 2014, 7(6): 2056-2065.
- [9] Chen Y, Nasrabadi N M, Tran T D. Hyperspectral image classification using dictionary-based sparse representation [J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2011, 49(10): 3973-3985.
- [10] Fang L Y, Li S T, Kang X D, *et al.* Spectral-spatial classification of hyperspectral images with a superpixel-based discriminative sparse model[J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2015, 53(8): 4186-4201.
- [11] Camps-Valls G, Bruzzone L. Kernel-based methods for hyperspectral image classification[J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2005, 43(6): 1351-1362.
- [12] Xie J Y, Hone K, Xie W X, *et al.* Extending twin support vector machine classifier for multi-category classification problems[J]. *Intelligent Data Analysis*, 2013, 17(4): 649-664.
- [13] Chen Y, Nasrabadi N M, Tran T D. Hyperspectral image classification via kernel sparse representation [J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2013, 51(1): 217-231.
- [14] Sun X X, Qu Q, Nasrabadi N M, *et al.* Structured priors for sparse-representation-based hyperspectral image classification[J]. *IEEE Geoscience & Remote Sensing Letters*, 2014, 11(7): 1235-1239.
- [15] Peng J T, Zhou Y C, Chen C L P. Region-kernel-based support vector machines for hyperspectral image classification [J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2015, 53(9): 4810-4824.