

doi:10.3788/gzxb20184706.0610001

# 联合稀疏特性和邻域相似度量的高光谱图像分类

刘嘉敏,张丽梅,石光耀,黄鸿

(重庆大学 光电工程学院 光电技术与系统教育部重点实验室,重庆 400044)

**摘 要:**传统的稀疏表示分类方法仅利用图像数据的稀疏特性分类,未利用高光谱图像的邻域信息,为此提出了一种联合稀疏特性和邻域相似度量的分类方法.该方法首先利用稀疏表示揭示出数据的稀疏特性,然后计算在各类样本中的稀疏相似性,并结合邻域特性,构建数据在各类样本中的稀疏-邻域联合相似关系,最后根据联合相似性大小判断数据类别.在利用数据的稀疏特性的同时结合像元的邻域信息,增强各种地物类别间的区分性,提升分类效果.在 Indian Pines 和 PaviaU 高光谱数据集上的实验表明:本文算法的分类精度高于其他方法,总体分类精度分别达到了 81.69% 和 86.59%,能得到具有更多同质区域的分类结果图,拥有更好的总体分类精度、平均分类精度和 Kappa 系数.

**关键词:**高光谱图像;相似度量;稀疏表示;联合相似性;邻域信息

中图分类号:TP751

文献标识码:A

文章编号:1004-4213(2018)06-0610001-10

## Hyperspectral Image Classification with Combination of Sparse Characteristic and Neighborhood Similarity Metrics

LIU Jia-min, ZHANG Li-mei, SHI Guang-yao, HUANG Hong

(Key Laboratory of Optoelectronic Technique System of the Ministry of Education, College of Optoelectronic Engineering, Chongqing University, Chongqing 400044, China)

**Abstract:** The traditional sparse representation classification methods only exploit the sparse property while they ignore the neighborhood similarity information in hyperspectral image. To address this problem, a novel sparsity-neighborhood metric classification method was proposed in this paper. Firstly, the proposed algorithm utilizes sparse representation to reveal the sparse properties of data, the following sparse similarity can be calculated in each class of samples. Then, according to neighborhood information, the method constructs the sparsity-neighborhood similarity relationship in each class of samples. Finally, the land cover types can be obtained with the federated sparsity-neighborhood similarity. The proposed algorithm possesses sparse property and neighborhood information, which can enhance the discrimination among different land cover classes to improve the classification performance. The experiments were performed on the Indian Pines and PaviaU hyperspectral data sets. Experimental results demonstrate that the proposed algorithm has better classification accuracy than other algorithms, the overall classification accuracies reach 81.69% and 86.59%, respectively. The proposed algorithm can obtain more homogeneous regions and improve classification accuracy and Kappa coefficient.

**Key words:** Hyperspectral Image; Similarity Metrics; Sparse Representation; Federated Similarity; Neighborhood Information

**OCIS Codes:** 100.4145; 110.4234; 100.5010; 100.2960

## 0 引言

高光谱遥感是现代遥感领域中最前沿的应用技术之一<sup>[1]</sup>.高光谱遥感图像是由几十个乃至数百个连续

基金项目:国家自然科学基金(No.41371338)和中央高校校长专项基地平台项目(No.106112017CDJPT120001)资助

第一作者:刘嘉敏(1973-),男,副教授,博士,主要研究方向为信息获取与处理技术和图像处理技术. Email: liujm@cqu.edu.cn

收稿日期:2017-11-13;录用日期:2018-04-02

<http://www.photon.ac.cn>

波段图像组成的数据立方体,蕴含丰富的地物信息,使得高光谱分类技术被广泛应用于资源探索、环境监测、精细农业、目标识别等领域<sup>[2-4]</sup>.高光谱图像具有数据量大、波段多、光谱分辨率高的特点<sup>[5-7]</sup>,不同的光谱曲线图呈现出不同的地物类别特性,利用该类别特性可实现地物的精细识别与分类.

高光谱图像分类作为高光谱遥感影像应用的一个重要研究方向,国内外学者积极开展了相关研究,并且取得了不错的分类效果.其中最近邻(Nearest Neighbor, NN)算法<sup>[8]</sup>利用训练样本集的邻域信息判断未知像元的类别.光谱角制图(Spectral Angle Mapping, SAM)<sup>[9]</sup>是一种波谱间相似性度量算法,它通过计算光谱曲线间的角度对两个样本间的相似程度进行判断.随机森林(Random Forest, RF)<sup>[10]</sup>是一种由一系列决策树组成的分类器,每一个决策树对像元进行分类,像元的最终类别由投票决定.支持向量机(Support Vector Machine, SVM)<sup>[11]</sup>是一种核分类方法,可以有效地解决分类中的小样本和非线性问题.虽然这几种传统算法都被广泛地应用于高光谱图像的地物分类中,但是这些算法要取得较好的分类效果需要大量的标记样本,而标记样本的获取往往需要耗费巨大的人力和物力.同时,上述分类算法也仅利用了图像中的光谱信息,并未考虑图像数据的其他特性.

近年来,随着稀疏表示在数据表征方面的广泛应用,研究者将稀疏模型引入到高光谱图像分类中.稀疏表示分类(Sparse Representation Classifier, SRC)法<sup>[12]</sup>就是在由各类训练样本构成的子字典中选取少数几个原子对待测像元进行线性表示,并通过计算测试样本与各类训练样本的最小残差判定测试样本的类别. Tropp 等<sup>[13]</sup>提出正交匹配追踪(Orthogonal Matching Pursuit, OMP)算法,该算法通过不断的迭代过程找到最匹配的重构信号.文献<sup>[14]</sup>提出了一种稀疏邻域分类(Sparse Neighborhood Classifier, SNC)算法,该算法利用稀疏系数构造相似性权值,最终根据权值大小判断类别.文献<sup>[15]</sup>提出了一种融合稀疏表示和协同表示(Fused Representation-based Classification, FRC)的分类算法,该算法通过融合残差值大小确定地物类别.文献<sup>[16]</sup>提出了一种基于空-谱的残差融合分类方法(Fused Representation Classification based Spatial-Spectral, FRC-SS),该算法有效利用图像数据的光谱信息和空间信息.文献<sup>[17]</sup>提出了一种基于类别关系的稀疏表示分类(Class-dependent Sparse Representation Classifier, CDSRC)方法,该方法将像元间的相关性和欧式距离关系有效结合,提高地物分类效果.以上稀疏表示分类算法利用像元的本质特征进行分类,广泛地应用于高光谱图像的分类中.但是仅仅依赖单个样本数据用于稀疏表示的线性叠加,并没有考虑到样本数据的邻域信息,会影响分类效果.因此,在高光谱图像的分类应用中,如果将图像数据的稀疏特性和邻域信息进行有效融合,会为地物的精细分类带来新的机遇.

为有效融合高光谱数据集的稀疏特性和邻域信息,增强各类地物间的区分性,本文提出一种联合稀疏特性和邻域相似度的分类算法(Sparsity-Neighborhood Metric Classifier, SNMC).该方法首先利用稀疏表示揭示出图像数据的稀疏特性,进而获得各类样本的稀疏相似性,同时结合样本的邻域信息,构建数据在各类样本中的稀疏-邻域联合相似关系,最后根据联合相似性大小判断样本数据的类别.在 Indian Pines 和 PaviaU 高光谱图像上的实验结果表明:本文方法提高了分类效果,能得到具有更多同质区域的分类结果图.

## 1 算法介绍

### 1.1 稀疏表示分类方法

稀疏表示是近年来图像处理和模式识别研究领域的热点<sup>[18]</sup>.根据高光谱像元分解观点可知<sup>[19]</sup>,稀疏表示是将每一个未知类别像元看成字典原子的线性组合,并通过计算测试样本与各类训练样本的最小残差判定测试样本的类别.为有效地提升地物类别区分性,在高光谱图像的稀疏表示应用中,字典由不同类别的训练样本构成不同的子字典.假设在高光谱数据集中共有  $c$  个不同类别地物和  $n$  个有标记像元.其中像元  $x_{ij} \in R^d$  表示训练样本的任意一个像元,给定训练样本集  $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_c]$ ,  $\mathbf{X}_i = [x_{i1}, x_{i2}, \dots, x_{in_i}]$  表示各类子字典,其中  $x_{in_i} (1, 2, \dots, n) \in R^d$ .由此,可以对待测像元  $y \in R^d$  建立稀疏表示数学模型, $y$  可以看成是子字典中少量训练样本的线性组合,具体可表示为

$$\mathbf{y} \approx \alpha_{i1} x_{i1} + \alpha_{i2} x_{i2} + \dots + \alpha_{in_i} x_{in_i} = [x_{i1}, x_{i2}, \dots, x_{in_i}] [\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{in_i}]^T = \mathbf{X}_i \alpha_i \quad (1)$$

式中  $\alpha_i$  为对应训练样本子字典中求解的稀疏系数向量.若要利用全部训练样本信息,待测样本  $y$  可由全部训练样本中的已有标记像元线性表示

$$\mathbf{y} = \mathbf{X} \alpha \quad (2)$$

式中  $\alpha = [\alpha_{11}, \alpha_{22}, \dots, \alpha_{m_c}]$  为对应的全部有标记像元的稀疏向量.

为了获得稀疏系数,需求解优化问题,即

$$\begin{cases} \hat{\alpha} = \operatorname{argmin}_{\alpha} \|\alpha\|_0 \\ \text{st. } X\alpha = y \end{cases} \quad (3)$$

式中零范数为稀疏向量  $\hat{\alpha}$  中非零系数的个数,可以表示为稀疏度  $S$ ,主要控制稀疏向量  $\hat{\alpha}$  的稀疏程度.

稀疏系数可以通过正交匹配追踪(OMP)方法进行求解,OMP算法的核心思想是在每次迭代中选择最为匹配的原子进行稀疏重构,对选取的原子集合进行施密特正交化,之后计算各类样本的残差值,通过不断的迭代过程不断更新字典得到残差值最小的原子集,进而求解稀疏系数. $y$  的类别可通过稀疏向量和各类子字典之间的最小残差值确定,即

$$\text{Label}(y) = \operatorname{argmin}_{i=1,2,\dots,c} \|y - X_i \hat{\alpha}_i\|_2 \quad (4)$$

式中  $\hat{\alpha}_i$  为第  $i$  类训练样本对应的稀疏系数向量.这种分类算法仅利用了像元的稀疏特性进行分类,对选取的字典数据依赖性程度高,从而导致分类效果不佳.

### 1.2 联合稀疏特性和邻域结构度量的分类方法

在高光谱图像分类应用中,为提升地物分类性能,可将样本数据的稀疏特性和邻域信息相融合.基于此,本文提出了一种联合稀疏特性和邻域相似度的分类(SNMC)算法.该算法的核心思想是:将像元的稀疏特性和邻域信息有效结合起来,利用权重因子平衡像元的稀疏相似性和邻域相似性,通过最终相似性值大小判断待测样本的类别标签.SNMC算法如图1所示.

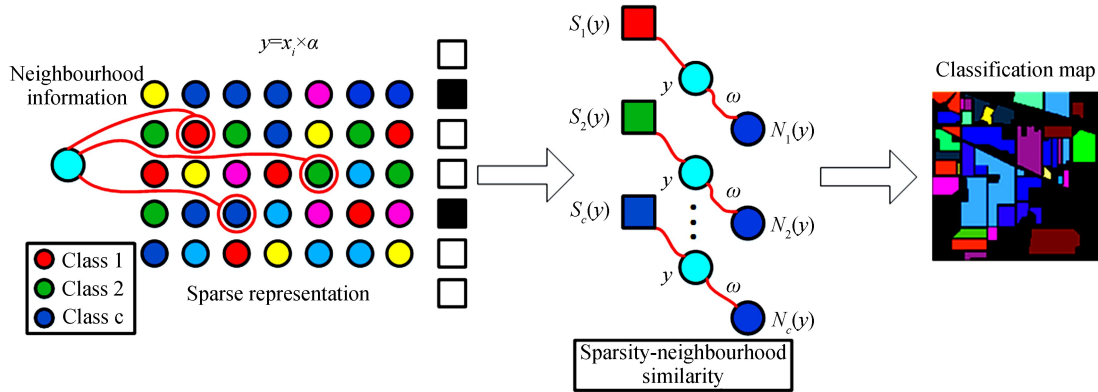


图1 SNMC算法

Fig.1 The schematic diagram of SNMC algorithm

图1中,每一个圆圈代表每一个像素点,圆圈集合代表训练样本字典.首先从训练样本字典中选取原子对待测像元进行稀疏表示,三个方框分别表示在各个子字典中选取的与待测像元稀疏相似性最大的像素点,其中稀疏系数通过正交匹配追踪方法获得,从而揭示各类别像元的稀疏相似性.其次,在一定邻域范围内,待测像元与其邻近像元具有邻域相似性关系,3个圆形框表示待测像元的邻近像元.由图1可知,可将待测像元的稀疏特性与邻域信息结合起来,最终可通过引入权重因子  $\omega$ ,构建像元的稀疏-邻域联合相似性关系,共同决定待测像元的类别.

在高光谱图像中,为衡量待测像元与各类别的相似性程度,即  $y$  在各个类别中的稀疏相似性定义为

$$S_i(y) = \exp\left(-\frac{\|y - X_i \hat{\alpha}_i\|_2^2}{2t_s^2}\right) \quad (i=1,2,\dots,c) \quad (5)$$

式中  $t_s = \frac{1}{c} \sum_{i=1}^c \|y - X_i \hat{\alpha}_i\|_2$  为热核参数.其中,  $S_i(y)$  表征测试样本  $y$  与训练样本的相似性程度,其值越大表示与真实地物类别越接近.

在高光谱分类应用中,如果仅仅利用像元的稀疏特性进行分类,则只关注到像元间的相关性程度,容易造成分类结果中的离散错分情况.因此,将像元的光谱邻域信息引入到分类中能够提高分类结果的精度.对于待测样本  $y$ ,其邻域相似性可定义为

$$N_i(y) = \max \left\{ \left[ \exp \left( - \frac{\|y - x_j^i\|^2}{2t_{N_i}^2} \right) \right]_{j=1}^n \right\} \quad (6)$$

式中,  $N_i(y)$  为  $y$  在各类中的邻域相似性,  $t_{N_i} = \frac{1}{n_i} \sum_{j=1}^{n_i} \|y - x_j^i\|$  为权重系数, 不同类别拥有不同的系数. 式(6)通过计算未知像元  $y$  与各类训练样本间光谱特征的相似性, 最终计算出待测像元在各类别中的邻域相似值  $N_i(y)$ ,  $N_i(y)$  值越大, 说明待测像元所属的地物类别与相对应的训练样本类别相近.

为了同时利用图像数据稀疏特性和邻域信息, 提升各类地物的区分性, 本文提出一种联合稀疏特性和邻域相似度量的方法. 待测像元  $y$  的标签可以由稀疏相似性和邻域相似性共同确定,

$$\text{Label}(y) = \underset{i=1,2,\dots,c}{\operatorname{argmax}} [S_i(y) + \omega N_i(y)] \quad (7)$$

式中:  $\omega$  为权重因子, 主要作用为平衡像元稀疏相似性和邻域相似性对地物类别的影响程度. 式(7)通过引入权重因子  $\omega$  计算待测像元的各类稀疏-邻域相似值, 根据各类联合相似值的大小判断未知像元  $y$  的类别标签, 联合相似值最大的类别属于未知像元的地物类别.

由式(7)可以看出, 新的类别度量方式通过引入参数稀疏度  $S$  来控制稀疏向量的稀疏程度, 进而控制像元间稀疏相似性大小, 引入像元的稀疏特性; 同时, 通过权重因子  $\omega$  来调整像元的稀疏相似性和邻域相似性的关系, 以此平衡像元的稀疏特性和邻域信息. 因此, 可以将像元的稀疏特性和邻域信息有效结合起来, 将数据的稀疏相似性矩阵和邻域相似性矩阵进行加权, 通过最终联合相似性大小判断待测样本的类别标签. 此方法可以有效避免待测像元的错分现象, 从而提高分类精度.

### 1.3 SNMC 算法步骤

提出的联合稀疏特性和邻域相似度量的分类算法(SNMC)步骤如下:

输入: 训练样本集  $X \in R^{d \times n}$ , 测试样本集  $y \in R^d$ , 稀疏度  $S$ , 权重因子  $\omega$ .

输出: 测试样本的类别.

步骤 1: 根据训练样本比例, 从高光谱数据集中随机选取训练样本和测试样本.

步骤 2: 由各个训练样本集构造各个类别的子字典, 通过 OMP 算法求解各个类别的稀疏系数, 进而获得数据稀疏相似性矩阵  $S_i(y)$ .

步骤 3: 根据式(6), 计算训练样本点和测试样本点间的邻域相似值, 并构造邻域相似性矩阵  $N_i(y)$ .

步骤 4: 根据式(7), 引入权重因子, 将未知像元的稀疏特性和邻域信息进行融合, 并且根据稀疏-邻域联合相似值大小确定待测像元  $y$  的类别.

## 2 实验结果及分析

### 2.1 实验数据集

实验中使用 Indian Pines 高光谱遥感数据集和 PaviaU 高光谱遥感数据集进行分类结果验证.

1) Indian Pines 数据集: 该数据集是由 AVIRIS 传感器采集于美国西北部印第安纳松林. 图像大小为  $145 \text{ pixel} \times 145 \text{ pixel}$ , 空间分辨率为  $20 \text{ m}$ , 光谱分辨率为  $5 \text{ nm}$ , 共包含 220 个波段, 除去受到水汽和噪声影响的 20 个波段, 余下的 200 个波段用于分类实验. 其伪彩色图和真实地物类别信息如图 2 所示.

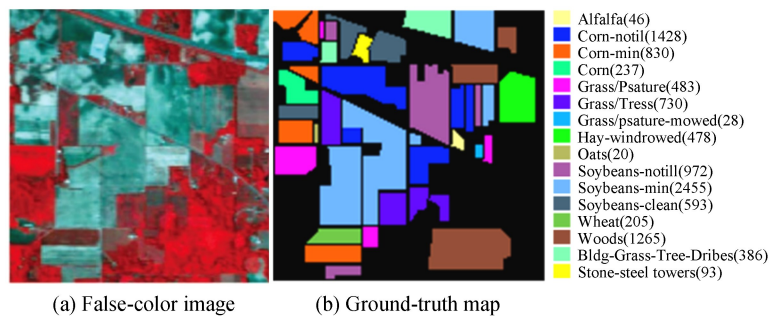


图 2 Indian Pines 高光谱图像  
Fig.2 Indian Pines hyperspectral image

2)PaviaU 数据集:该数据集是由 ROSIS 传感器采集于意大利帕维亚大学校园.该遥感图像大小为 610 pixel×340 pixel,空间分辨率为 1.3 m,共包含 115 个波段,除去受水汽影响的 12 个波段,余下的 103 个波段用于分类实验.其伪彩色图和真实地物图类别信息如图 3 所示.

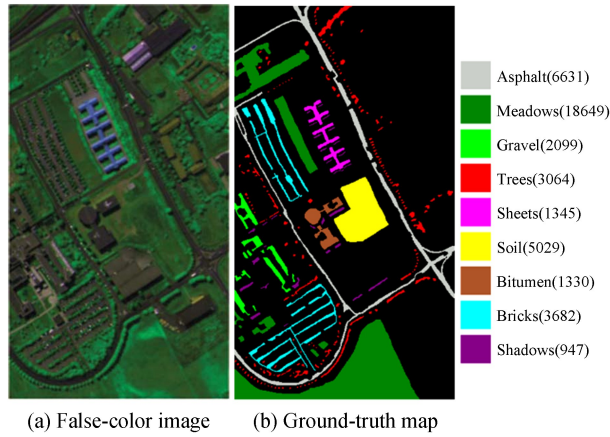


图 3 PaviaU 高光谱图像  
Fig.3 PaviaU hyperspectral image

### 2.2 实验设置

使用总分类精度(Overall Accuracy, OA)、平均分类精度(Average Accuracy, AA)和 Kappa 系数三个指标评估各分类算法性能.实验中,选取 NN、SAM、OMP、SRC、RF、SNC、NSC 和 FRC 算法进行对比,其中 NSC 为本文中提出的仅利用像元间邻域相似性(Neighborhood Similarity Classifier, NSC)的分类算法.为验证本文算法的有效性,实验进行交叉验证,将各种对比算法的参数都调整到最佳,其中 NN、SAM 和 NSC 算法没有参数,OMP 算法调参范围为 1~50,稀疏度选择为 15;SRC 算法调参范围为 0.1~0.9,误差参数选择为 0.9;RF 算法调参范围为 1~10,决策树个数选择为 5;SNC 算法的误差参数选择为 0.001;FRC 算法的平衡参数选择为 0.5.每次实验随机选取一定比例的训练样本,其余为测试样本,每种算法的试验次数设置为 10 次,将 10 次实验结果的平均值作为最终结果.

### 2.3 Indian Pines 实验结果

本文提出的算法最主要的参数是稀疏度  $S$  和权重因子  $\omega$ .为确定稀疏度和权重因子的最优值,从 Indian Pines 数据集中选取 10% 的标记像元作为训练样本,其余为测试样本,当训练样本数目不足 5 个时,统一选取 5 个训练样本.图 4 显示了本文算法中稀疏度  $S$  和权重因子  $\omega$  对分类精度的影响.

由图 4 可知,本文算法的分类精度会受到稀疏度  $S$  和权重因子  $\omega$  的影响.随着  $S$  和  $\omega$  的增加,算法分类精度也先随之增大,而后被削弱.由于  $\omega$  起到一个平衡稀疏相似性和邻域相似性的作用,当  $\omega$  过大时,最终地物类别主要是由像元的邻域信息决定;当  $\omega$  过小时,地物类别主要由像元的稀疏相似性决定,因此无论  $\omega$  的值过大还是过小,都会导致像元在稀疏相似性和邻域相似性的比例失调,使得分类效果不佳.为寻求在稀疏相似性和邻域相似性的平衡,因此  $S$  选择为 15,  $\omega$  为 50 时,分类效果达到最佳.

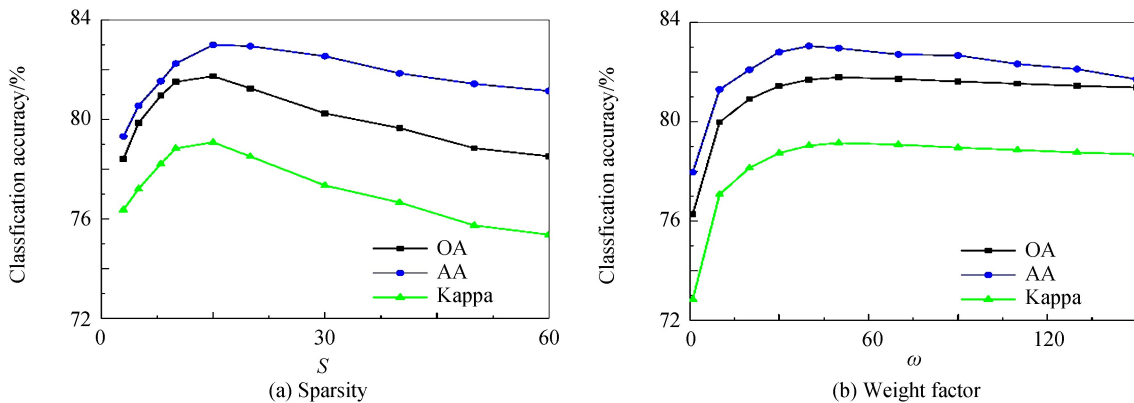


图 4 不同参数  $S$ 、 $\omega$  对分类精度的影响  
Fig.4 The influence of different parameters  $S$  and  $\omega$  on the classification accuracy

为了验证提出算法的有效性,从数据集中选取 2%、4%、6%、8%、10% 的像元作为训练样本,其余为测试样本.表 1 表示在选取不同比例的训练样本的情况下不同算法的总分类精度.

表 1 不同算法在不同训练样本比例下的总体精度

Algorithm	Percent of training samples/%				
	2	4	6	8	10
NN	65.66(1.10)	70.51(0.66)	72.41(0.52)	74.58(0.54)	75.16(0.66)
SAM	65.74(0.95)	70.50(0.75)	73.23(0.37)	74.48(0.40)	76.24(0.78)
OMP	65.18(1.00)	69.81(0.80)	72.55(0.27)	74.00(0.53)	74.87(0.31)
SRC	64.79(1.72)	70.86(0.56)	73.52(0.69)	74.68(0.63)	76.70(0.77)
RF	68.70(0.87)	73.65(0.39)	76.46(0.27)	77.70(0.60)	79.42(0.24)
SNC	56.68(0.71)	58.75(0.80)	64.40(1.21)	66.59(1.04)	68.72(0.67)
NSC	66.31(1.24)	70.92(0.73)	73.34(0.76)	75.13(0.63)	76.24(0.47)
FRC	59.58(1.69)	71.91(1.19)	74.57(0.60)	76.46(0.60)	77.10(0.57)
SNMC	<b>72.12(1.19)</b>	<b>78.01(0.89)</b>	<b>79.29(0.78)</b>	<b>80.76(0.45)</b>	<b>81.69(0.49)</b>

表 1 实验结果表明,随着训练样本数目的增多,各种算法的分类精度呈增长趋势.NN 算法和 SAM 算法仅考虑了图像中蕴含的光谱信息,NSC 算法仅考虑邻域信息,所以分类效果较差.SRC、SNC 和 OMP 算法都是基于单个像元的稀疏表示分类算法,仅考虑了单个像元与各类训练样本之间的相关性程度,未考虑待测像元的邻域信息,故分类效果不佳.RF 算法采用统计学的方法进行分类,通过一系列的决策树分类器对未知像元类别投票判决,其分类效果优于稀疏表示和光谱方法.FRC 算法同时考虑了稀疏表示与协同表示的残差信息,故分类效果比其他几种算法好.本文提出的 SNMC 算法的分类精度要优于其它算法,这是由于其不仅考虑了像元间的邻域信息,还较好地利用了待测像元与其它已有标记像元的稀疏关联程度,因此减小了待测像元被错分的可能性,有效提高了分类精度.

为进一步分析各个分类方法在每一个地物类别上的分类精度,实验中选取 10% 的训练样本,其余为测试样本进行测试.在 Indian Pines 数据集中每个地物类别的分类精度、总分类精度(OA)、平均分类精度(AA)和 Kappa 系数如表 2 所示.

表 2 Indian Pines 数据集中各类地物在不同算法下的分类精度

Class	Class name	Train	Test	NN	SAM	OMP	SRC	RF	SNC	NSC	FRC	SNMC
1	Alfalfa	5	49	72.67	74.18	33.21	76.88	76.00	<b>100</b>	81.24	79.58	92.86
2	Corn-notill	143	1291	62.08	64.53	68.53	66.02	<b>75.37</b>	57.29	64.17	64.74	70.95
3	Corn-min	83	751	63.62	64.80	61.21	64.66	73.49	<b>76.58</b>	63.95	66.11	70.01
4	Corn	23	211	53.28	48.39	58.98	53.87	<b>63.93</b>	52.17	54.09	57.75	58.76
5	Grass/Pasture	50	447	89.82	90.26	87.82	91.34	90.39	<b>93.17</b>	89.28	90.60	92.22
6	Grass/Trees	75	672	87.94	86.75	86.66	88.51	<b>92.83</b>	80.68	86.35	87.13	91.71
7	Grass/Pasture-mowed	5	21	69.17	75.66	47.87	73.21	<b>83.33</b>	36.36	76.53	77.50	81.82
8	Hay-windrowed	49	440	95.61	96.14	91.03	96.25	95.95	84.15	96.93	95.65	<b>97.97</b>
9	Oats	5	15	48.92	53.42	41.19	45.19	80.00	<b>100</b>	44.59	43.03	52.38
10	Soybeans-notill	97	871	63.47	67.38	67.19	68.85	70.51	66.61	67.64	69.09	<b>79.67</b>
11	Soybeans-min	247	2221	73.41	75.35	71.59	75.00	75.21	57.10	75.63	75.50	<b>76.92</b>
12	Soybean-clean	61	553	67.23	72.65	72.92	69.92	66.85	58.14	70.11	73.69	<b>86.47</b>
13	Wheat	21	191	87.65	90.15	93.00	91.22	91.11	85.08	91.50	89.92	<b>95.00</b>
14	Woods	129	1165	92.98	92.05	90.95	93.11	<b>93.83</b>	85.11	92.84	93.12	93.46
15	Building-Grass-Trees-Drives	38	342	57.76	62.42	65.98	61.65	68.83	68.27	59.30	62.98	<b>72.70</b>
16	Stone-steel Towers	10	85	98.48	96.93	92.40	<b>98.74</b>	94.44	94.59	95.51	97.91	97.53
OA				74.67	76.24	74.87	76.70	79.45	68.52	76.24	77.10	<b>81.69</b>
AA				74.01	75.69	70.66	75.90	80.75	74.70	76.91	77.84	<b>83.34</b>
Kappa				0.711	0.729	0.713	0.734	0.735	0.631	0.732	0.740	<b>0.743</b>

由表 2 可知,本文算法在大多数类别上分类精度较高,尤其显著的是 Alfalfa、Grass/Pasture-mowed、Soybean-clean 等类别,主要因为在稀疏表示算法中,待测像元和该类训练样本的相似性低,而其他类别相似性高,导致像元的错分现象,当引入待测像元的邻域信息时,可以在很大程度上避免错分现象.同时,SNMC 算法的 OA、AA 和 Kappa 系数均高于其它几种利用图像单一性信息的分类算法,这是因为本文算法考虑了像元的邻域信息和稀疏特性,提升了分类的鉴别性能,减小了错分现象.

图 5 为各个分类算法的分类结果图,从图中可以看出,其他方法的分类结果图上出现了不同程度的样本错分点,而考虑稀疏相关性信息和邻域信息的 SNMC 算法在分类结果性能上优于其他几种算法,分类结果图更为平滑,尤其在 Alfalfa、Grass/Pasture、Wheat 等类别更为明显.

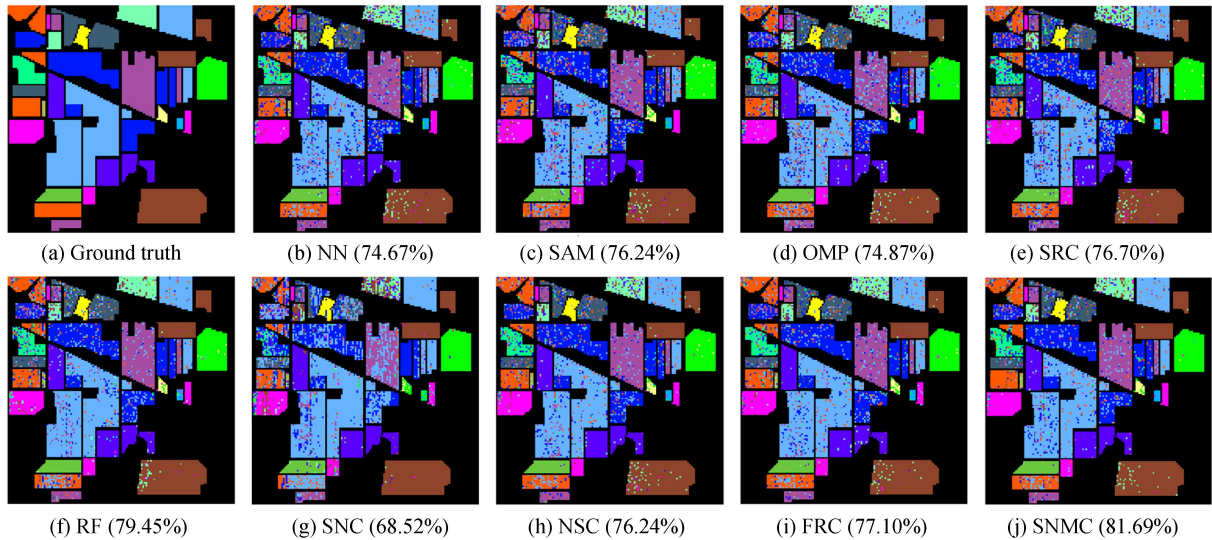


图 5 各算法在 Indian Pines 数据集上的分类结果

Fig.5 Classification results of different algorithms on Indian Pines data set

## 2.4 PaviaU 实验结果

实验中,从 PaviaU 数据集中选取 10% 的标记像元作为训练样本,其余为测试样本,使用和 Indian Pines 相同的参数选取方法,确定稀疏度  $S$  和权重因子  $\omega$  分别为: $S=15, \omega=50$ ,如图 6 所示.

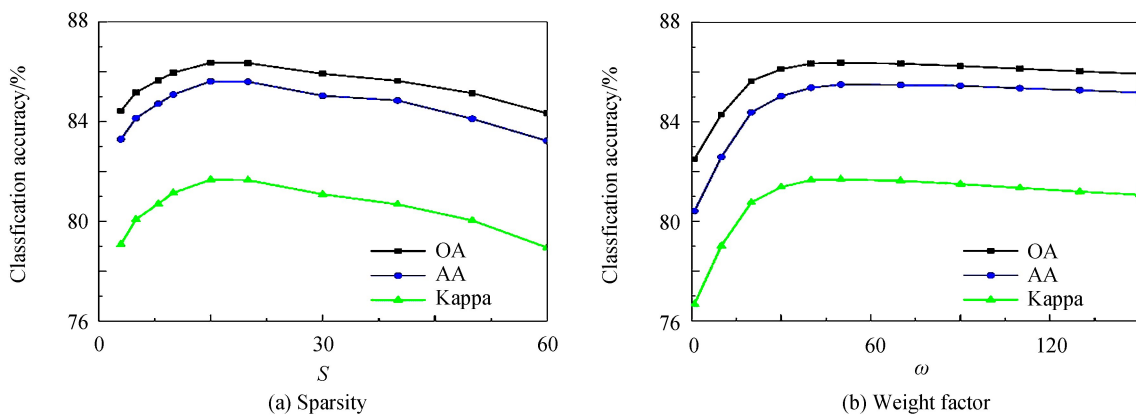


图 6 不同参数  $S, \omega$  对分类精度的影响

Fig.6 The influence of different parameters  $S$  and  $\omega$  on the classification accuracy

表 3 为不同算法在不同训练样本比例下的总分类精度.如表 3 所示,随着训练样本比例的增加,各种算法的分类精度呈增长趋势.在小样本训练样本的条件下,SNMC 算法分类效果明显提升,主要是因为待测像元所属真实地物类别的邻域相似值和稀疏相似值较大,而且在各种训练比例条件下本文算法的分类精度要优于其它算法.

表 3 不同算法在不同训练样本比例下的总体精度

Table 3 Overall accuracy of different algorithms in different proportion of training samples

Algorithm	Percent of training samples/%				
	2	4	6	8	10
NN	79.84(0.33)	82.05(0.39)	82.47(0.23)	83.27(0.08)	83.54(0.29)
SAM	80.12(0.38)	81.77(0.20)	82.33(0.19)	83.01(0.14)	83.48(0.29)
OMP	77.53(0.42)	79.76(0.42)	80.82(0.16)	81.26(0.20)	81.74(0.26)
SRC	79.80(0.96)	81.86(0.24)	83.12(0.26)	83.28(0.26)	83.67(0.23)
RF	80.89(0.26)	82.97(0.33)	84.30(0.37)	85.43(0.33)	85.91(0.45)
SNC	80.40(0.67)	82.44(0.38)	84.14(0.29)	85.20(0.17)	85.63(0.15)
NSC	79.68(0.51)	81.70(0.11)	82.45(0.25)	83.14(0.29)	83.70(0.20)
FRC	80.97(0.67)	82.46(0.46)	83.25(0.35)	83.76(0.12)	84.22(0.27)
SNMC	<b>83.92(0.51)</b>	<b>85.11(0.29)</b>	<b>85.82(0.21)</b>	<b>86.39(0.23)</b>	<b>86.59(0.09)</b>

表 4 PaviaU 数据集中各类地物在不同算法下的分类精度

Table 4 Classification accuracy of different types of features in PaviaU data set by different algorithms

Class	Class name	Train	Test	NN	SAM	OMP	SRC	RF	SNC	NSC	FRC	SNMC
1	Asphalt	663	5968	91.17	92.16	81.64	91.88	83.79	85.28	<b>92.25</b>	90.93	91.10
2	Corn-notill	1865	16784	87.49	87.33	88.92	87.27	89.27	85.93	87.50	86.94	<b>89.87</b>
3	Corn-min	210	1889	64.57	63.91	60.71	66.80	69.56	<b>82.84</b>	62.81	67.54	69.79
4	Corn	306	2758	94.36	94.66	91.89	94.96	92.33	95.38	96.04	95.66	<b>96.78</b>
5	Grass/Pasture	135	1210	99.00	99.21	99.17	<b>99.74</b>	99.67	97.24	99.59	99.51	99.42
6	Grass/Trees	503	4526	71.40	70.92	68.81	71.74	77.95	<b>87.35</b>	72.34	76.07	77.75
7	Grass/Pasture-mowed	133	1197	68.19	68.39	67.02	69.90	82.26	<b>96.75</b>	69.91	71.27	76.27
8	Hay-windrowed	368	3314	65.46	65.35	64.18	64.70	<b>72.99</b>	69.29	65.90	65.73	70.53
9	Oats	95	852	99.55	99.76	89.40	99.61	96.70	<b>100</b>	99.63	99.69	99.76
OA				83.54	83.48	81.74	83.67	85.58	85.54	83.70	84.22	<b>86.59</b>
AA				82.35	82.41	79.08	82.96	84.95	<b>88.89</b>	83.85	83.58	85.70
Kappa				0.779	0.778	0.757	0.780	0.806	0.802	0.783	0.786	<b>0.820</b>

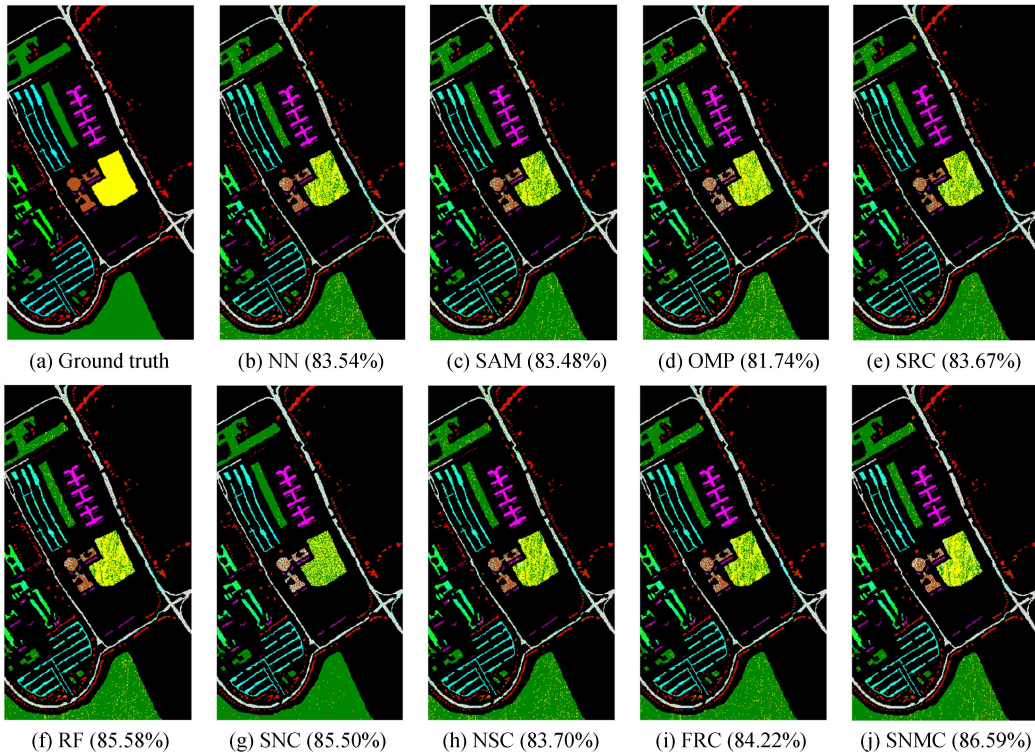


图 7 各算法在 PaviaU 数据集上的分类结果图

Fig.7 Classification results of different algorithms on PaviaU data set



表4为PaviaU数据集在不同地物类别下不同算法的分类精度,图7为相对应的分类图.由此可得,本文算法的OA、AA、Kappa系数相对于其它几种分类算法均有提升,在其中几种地物类型中效果较为显著,且得到的分类图中本文算法的图像分布较为平滑,分类效果较好.由此更加验证了本文算法的有效性和可行性.

### 3 结论

针对传统的稀疏表示分类算法仅利用数据的稀疏相似性,未考虑将数据的邻域信息利用到分类过程中的问题,本文提出了一种新的基于稀疏特性和邻域相似度量的分类方法.该方法结合高光谱图像数据的稀疏特性和邻域信息,通过引入权重因子获得较好地物分类效果.在真实高光谱数据集上的实验表明,提出的SNMC算法与相关分类算法相比,能得到更好的分类性能,在Indian Pines数据集上分类精度提升约7%,在PaviaU数据集上分类精度提升约5%.

但是本文算法在求解邻域相似性时存在引入异类地物的风险,对于分布过于密集的区域,存在较多的异类地物,并且其光谱距离相近,此时邻域信息的引入易造成错误分类.基于此,后续研究工作将考虑:1)引入空间信息,进一步提升分类效果;2)寻找更加合适的相似度平衡方法,减少参数个数,为实现多种特性融合分类打下基础.

#### 参考文献

- [1] DU Pei-jun, XIA Jun-shi, XUE Chao-hui, *et al.* Progress of hyperspectral remote sensing image classification[J]. *Journal of Remote Sensing*, 2016, **20**(2): 236-256.  
杜培军, 夏俊士, 薛朝辉, 等. 高光谱遥感影像分类研究进展[J]. 遥感学报, 2016, **20**(2): 236-256.
- [2] DU Bo, HUANG Zhi-qiang, WANG Nan, *et al.* Joint weighted nuclear norm and total variation regularization for hyperspectral image denoising[J]. *International Journal of Remote Sensing*, 2018, **39**(2):334-355.
- [3] LI Xue, ZHANG Liang-pei, DU Bo, *et al.* Iterative reweighting heterogeneous transfer learning framework for supervised remote sensing image classification[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2017, **10**(5):2022-2035.
- [4] SHI Qian, DU Bo, ZHANG Liang-pei. Spatial coherence-based batch-mode active learning for remote sensing image classification[J].*IEEE Transactions on Image Processing*, 2015, **24**(7):2037.
- [5] DU Bo, ZHANG Meng-fei, ZHANG Le-fei, *et al.* PLTD: Patch-based low-rank tensor decomposition for hyperspectral images[J]. *IEEE Transactions on Multimedia*, 2017, **19**(1):67-79.
- [6] HUANG Hong, ZHENG Xin-lei. Hyperspectral image classification with combination of weighted spatial-spectral and KNN[J]. *Optics and Precision Engineering*, 2016, **24**(4): 873-880.  
黄鸿, 郑新磊. 加权空-谱与KNN相结合的高光谱图像分类[J]. 光学精密工程, 2016, **24**(4): 873-880.
- [7] ZHAO Wen-zhi, GUO Zhou, YUE Jun, *et al.* On combining multiscale deep learning features for the classification of hyperspectral remote sensing imagery[J]. *International Journal of Remote Sensing*, 2015, **36**(13): 3368-3379.
- [8] HUANG Hong, ZHENG Xin-lei, LUO Fu-lin. Hyperspectral image classification with combination of weighted mean filter and manifold reconstruction preserving embedding[J]. *Acta Photonica Sinica*, 2016, **45**(10): 146-154.  
黄鸿, 郑新磊, 罗甫林. 融合加权均值滤波与流形重构保持嵌入的高光谱影像分类[J]. 光子学报, 2016, **45**(10): 146-154.
- [9] TONG Qing-xi, ZHANG Bing, ZHANG Li-fu, *et al.* Current progress of hyperspectral remote sensing in China[J]. *Journal of remote sensing*, 2016, **20**(5): 689-707.  
童庆禧, 张兵, 张立福. 中国高光谱遥感的前沿进展[J]. 遥感学报, 2016, **20**(5): 689-707.
- [10] HU Wei, HUANG Yang-yu, WEI Li, *et al.* Deep convolutional neural networks for hyperspectral image classification [J]. *Journal of Sensors*, 2015, **2015**(2):1-12.
- [11] BO Chun-juan, LU Hu-chuan, DONG Wang. Weighted generalized nearest neighbor for hyperspectral image classification[J].*IEEE Access*, 2017, PP(99): 1-1.
- [12] CAMPS-VALL G. Kernel spectral angle mapper[J]. *Electronics Letters*, 2016, **52**(14): 1218-1220.
- [13] LI Da-wei, YANG Feng-bao, WANG Xiao-xia. Multisource remote sensing classification based on random forest and adaptive weighted D-S evidence synthesis[J].*Laser and Optoelectronics Progress*, 2016, **53**(3): 031001.
- [14] BANERJEE A, BURLINA P, DIEHL C. A support vector method for anomaly detection in hyperspectral imagery[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2006, **44**(8): 2282-2291.
- [15] WRIGHT J, YANG A Y, GANESH A, *et al.* Robust face recognition via sparse representation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, **31**(2):210.
- [16] SPARRER S, FISCHER R F H. MMSE-based version of OMP for recovery of discrete-valued sparse signals[J]. *Electronics Letters*, 2015, **52**(1): 75-77.

- [17] LUO Fu-lin, HUANG Hong, LIU Jia-min, *et al.* Fusion of graph embedding and sparse representation for feature extraction and classification of hyperspectral imagery[J]. *Photogrammetric Engineering and Remote Sensing*, 2017, **83**(1):37-46.
- [18] LI Wei, DU Qian, ZHANG Fan, *et al.* Hyperspectral image classification by fusing collaborative and sparse representations[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2016, **9**(9):4178-4187.
- [19] YANG Jing-hui, WANG Li-guo, QIAN Jin-xi. A new residual fusion classification method for hyperspectral images[J]. *International Journal of Remote Sensing*, 2016, **37**(4):745-769.
- [20] CUI M, PRASAD S. Class-Dependent sparse representation classifier for robust hyperspectral image classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2015, **53**(5):2683-2695.
- [21] WEI Q, BIOUCAS-DIAS J, DOBIGEON N, *et al.* Hyperspectral and multispectral image fusion based on a sparse representation[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2015, **53**(7): 3658-3668.
- [22] ZHANG Fei-fei, SUN Xu, XUE Liang-yong, *et al.* Hyperspectral mixed pixel decomposition policy merging simple linear iterative clustering[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2015, **31**(17):199-206. 张飞飞, 孙旭, 薛良勇, 等. 融合简单线性迭代聚类的高光谱混合像元分解策略[J]. *农业工程学报*, 2015, **31**(17):199-206.

---

**Foundation item:** The National Natural Science Foundation of China (No.41371338) and Central University President Special Base Platform Project (No.106112017CDJPT120001)

引用格式: LIU Jia-min, ZHANG Li-mei, SHI Guang-yao, *et al.* Hyperspectral Image Classification with Combination of Sparse Characteristic and Neighborhood Similarity Metrics[J]. *Acta Photonica Sinica*, 2018, **47**(6):0610001  
刘嘉敏, 张丽梅, 石光耀, 等. 联合稀疏特性和邻域相似度量的高光谱图像分类[J]. *光子学报*, 2018, **47**(6):0610001