

应用监督近邻重构分析的高光谱遥感数据特征提取

方敏, 王君, 王红艳, 李天涯

(四川理工学院理学院, 四川自贡 643000)

摘要: 针对高光谱遥感数据特征提取方法的研究, 提出了一种新的监督近邻重构分析(Supervised Neighbor Reconstruction Analysis, SNRA) 算法。该方法首先利用同一类别的近邻数据点对各数据点进行重构; 然后在低维嵌入空间中保持该重构关系不变, 尽可能地分离非同类数据点, 并利用总体散度矩阵来约束数据间的相关性; 最后求解得到一个最佳投影矩阵, 进而提取出鉴别特征。SNRA 算法不仅保持了同类数据的局部结构而且增强了非同类数据的可分性, 同时减少了数据的冗余信息。在 Indian Pine 和 KSC 高光谱遥感数据集上的实验结果表明: 提出的方法能更好地揭示出高光谱遥感数据的内在特性, 提取出更有效的鉴别特征, 改善分类效果。

关键词: 高光谱遥感数据分类; 特征提取; 监督学习; 邻域重构; 总体散度矩阵

中图分类号: TP751 **文献标志码:** A **DOI:** 10.3788/IRLA201645.1028003

Feature extraction of hyperspectral remote sensing data using supervised neighbor reconstruction analysis

Fang Min, Wang Jun, Wang Hongyan, Li Tianya

(College of Science, Sichuan University of Science & Engineering, Zigong 643000, China)

Abstract: For the feature extraction methods of hyperspectral remote sensing data, a new method, called supervised neighbor reconstruction analysis (SNRA), was proposed. First, this method reconstructs each point with neighbor points from the same class. Then, it preserves the reconstruction relationship and separates the data points from different classes as far as possible in low-dimension embedding space. And a total scatter matrix is used to constrain the correlation between data points. Finally, it obtains an optimized projection matrix and extracts the discriminating feature. SNRA not only preserves the local structures of intraclass data but also enhances the separability of interclass data. And it reduces the redundant information. The experiments on Indian Pine and KSC hyperspectral remote data sets show that the proposed method can better reveal the intrinsic property of hyperspectral remote sensing data and effectively extract the discriminating feature to improve the classification result.

Key words: hyperspectral remote sensing data classification; feature extraction; supervised learning; neighbor reconstruction; total scatter matrix

收稿日期: 2016-02-11; 修订日期: 2016-03-05

基金项目: 国家自然科学基金(理论物理专项)(11547196); 四川省教育厅重点资助(15ZA0229); 四川理工学院人才引进资助(2013RC07)

作者简介: 方敏(1963-), 女, 副教授, 硕士, 主要从事光学信息处理、高光谱遥感信息处理、光谱分析等方面的研究。

Email: fangmin631114@163.com

0 引言

遥感是从远处由传感器对地物反射或发射的电磁波信息进行探测,并揭示不同地物特性的一门技术^[1-2]。随着空间技术、信息技术和电子计算机技术的高速发展,遥感逐步从全色遥感、彩色遥感、多光谱遥感发展到高光谱遥感^[1]。高光谱遥感数据拥有数十数百个电磁波波段,所获取的地物光谱是连续的,能够精细表达不同地物的特征,但高光谱遥感数据不同波段间的相关性强,存在冗余信息,且具有丰富的隐含信息^[3-4]。因此,如何有效地提取出高光谱遥感数据中的地物特征是实现地物准确分类的关键。

近年来,流形学习在高维数据特征提取中应用十分广泛^[5],主要方法有等距映射(Isomap, Isometric Mapping)、局部保持嵌入(LLE, Locally Linear Embedding)和拉普拉斯特征映射(LE, Laplacian Eigenmap)。该类方法的主要思想是假设每个样本数据点和其邻域点位于一个线性或近似线性的流形上,采用非线性映射方式,在低维嵌入空间中保持数据间的内在结构关系不变。通过流形学习方法学习样本数据就能有效地揭示数据中的内在属性,进而得到隐含鉴别特征。在高光谱遥感分类的应用中该类方法能很好地反映出高光谱遥感数据的内在流形结构,得到有用的鉴别特征^[6]。但是这些方法只在所学习的样本中有定义,对外来的新样本无法直接进行投影,使其泛化能力受到限制。为此,He等人提出LE和LLE的线性近似方法,分别为局部保持投影^[7](LPP, Locality Preserving Projection)和邻域保持嵌入^[8](NPE, Neighborhood Preserving Embedding)。其中NPE通过投影矩阵把LLE中隐含的非线性映射转换为线性映射,因而能够通过投影矩阵得到新样本数据点在嵌入空间中的映射。虽然NPE本质上为线性方法,但同传统的主成分分析(PCA, Principal Component Analysis)、线性判别分析(LDA, Linear Discriminant Analysis)等子空间方法相比,NPE算法能够在嵌入空间中保持数据原始空间的局部几何结构,进而揭示出高维空间中的内在低维流形结构。但NPE表示数据点与其近邻重构间的关系,忽略了数据点之间的相似关系;且NPE为无监督方法,限制了其鉴别性能。为利用先验类别信息,

学者提出了监督NPE^[8](SNPE, Supervised NPE),但该方法只在选取近邻时考虑了类别信息,使得先验信息不能被充分利用,其特征的鉴别力受限。

针对上述问题,文中提出了一种监督近邻重构分析(SNRA, Supervised Neighbor Reconstruction Analysis)算法,该方法充分利用了数据点的类别标记,对各数据点采用与其同类的近邻点进行重构,在低维空间中保持重构关系不变,并抑制非同类数据点间的相似性表征,使非同类数据尽可能地分离;另外,利用总体散度矩阵来约束数据间的相关性,减少数据间的冗余。在公开的高光谱数据集Indian Pine和KSC上进行了实验,其结果验证了该方法的有效性。

1 NPE 算法

假设原始数据为 $X=[x_1, x_2, \dots, x_n] \in R^{m \times n}$,其中 m 为维数,样本 $x_i(i=1, 2, \dots, n)$ 的类别标记为 $l(x_i) \in \{1, 2, \dots, c\}$,其中 c 为类别数,映射到低维嵌入空间中对数据为 $Y=[y_1, y_2, \dots, y_n] \in R^{d \times n}$,其中 d 为嵌入维度,相应的投影矩阵为 $V^T \in R^{d \times m}$,则有 $Y=V^T X$ 。

NPE算法本质上为LLE的线性逼近,其主要思想假设数据点与其近邻数据点位于或近似位于某一个内在流形上,并寻求一个投影矩阵得到数据的低维嵌入,即在原始空间中各数据点由其近邻进行重构,在低维嵌入空间中保持数据的局部重构关系不变。

首先,在原始空间中,通过近邻来重构各数据点,最小化重构误差,得到重构权重。其目标函数为:

$$\min_w \sum_i \|x_i - \sum_j w_{ij} x_j\|^2 \quad (1)$$

式中: w_{ij} 表示 x_i 与 x_j 之间的重构权值;若 x_i 与 x_j 为近邻关系,则 $w_{ij} \neq 0$;否则, $w_{ij}=0$;并且 $\sum_j w_{ij}=1$,权重矩阵 $W=[w_{ij}]_{i,j=1}^n$ 。

其次,在低维嵌入空间中,保持各数据在原始空间中的重构权重不变,计算投影矩阵,使各数据点满足 $y_i=V^T x_i$,构建目标函数为:

$$\min_v \sum_i \|V^T x_i - \sum_j w_{ij} V^T x_j\|^2 \quad (2)$$

为消除尺度因子的影响,增加一约束为:

$$V^T X X^T V = I \quad (3)$$

式中: I 为单位矩阵。

2 监督近邻重构分析

SNRA是根据样本的类别标记,各数据点由与

其同类的近邻进行重构,在低维嵌入空间中,保持数据间的重构关系不变,即不改变数据间的重构权重,最大化各数据点与其非同类数据间的距离,尽可能分离不同类别的数据;同时,通过总体散度矩阵来约束各数据间的相关性,消除数据间的相关性,进一步减少数据间的冗余信息。该方法不仅保持了数据间的局部结构而且增强了数据间的可分性,并通过约束数据间的相关性来减少冗余信息,能更好地表征数据的内在流形结构,提取出更有效的鉴别特征,改善分类效果。

在高维空间中,对数据点 x_i 选取来自相同类别的近邻样本数据进行重构,使重构误差最小,构建目标函数,可得到重构权重。

$$\min_w \sum_{i,l} \|x_i^l - \sum_j w_{ij} x_j^l\|^2 \quad (4)$$

式中: x_j^l 为 x_i^l 来自同一类别 l 的近邻点, w_{ij} 表示 x_i^l 与 x_j^l 间的权值,需满足条件(1) 与 x_i^l 的同类非近邻点和非同类点对应的权重 $w_{ij}=0$; (2) 各数据点的所有权重和为 1,即 $\sum_j w_{ij}=1$ 。

在低维嵌入空间中,保持数据的局部重构关系不变,即不改变重构权重矩阵 $W=[w_{ij}]_{i,j=1}^n$ 的值,使投影后数据的局部结构关系被保持;同时,使投影后不同类别数据间的间隔最大化,进而分离不同类别的数据。通过构建一类间关系矩阵 $S=[s_{ij}]_{i,j=1}^n$ 来抑制不同数据间的相似性,其中 s_{ij} 表示数据点 x_i 与 x_j 间的权值,定义为:

$$s_{ij} = \begin{cases} 10, & x_i \text{ 和 } x_j \text{ 为非同类数据} \\ 0, & \text{其他} \end{cases} \quad (5)$$

在低维空间中,为尽可能地分离非同类数据点,构建目标函数为:

$$\min_v \sum_{i,l} \|V^T x_i - V^T \hat{x}_j\|^2 s_{ij} \quad (6)$$

式中: \hat{x}_j 为数据点 x_j 的同类近邻重构点,即 $\hat{x}_j = \sum_k w_{jk} x_k = XW_j$,且 $W_j=[w_{j1}, w_{j2}, \dots, w_{jn}]$ 。

对目标函数进行化解,可得:

$$\begin{aligned} \sum_{i,j} \|V^T x_i - V^T \hat{x}_j\|^2 s_{ij} &= \sum_{i,j} \|V^T x_i - V^T XW_j\|^2 s_{ij} = \\ &= \sum_{i,j} (V^T x_i - V^T XW_j) s_{ij} (V^T x_i - V^T XW_j)^T = \\ &= \sum_{i,j} \{ V^T (x_i s_{ij} x_i^T - XW_j s_{ij} x_i^T - x_i s_{ij} W_j^T X^T + XW_j s_{ij} W_j^T X^T) V \} = \\ &= \text{tr}(V^T X H X^T V) \end{aligned} \quad (7)$$

式中: $H=D_s-WS-SW^T+WD_sW^T$, D_s 为对角矩阵,且 $D_{sij}=\sum_l s_{ij}$ 。

由于投影后的特征数据间仍然具有一定的相关性,不同特征之间仍存在一些信息冗余,这些重叠的信息会导致不同数据间的差异性减弱,进而影响分类结果^[9]。为进一步消除特征数据间的相关性,减少冗余信息,提取出更有效的鉴别特征,因而对投影后的数据增加不相关约束。

为使投影后的数据不相关,数据间的协方差 $E((y_i-E(y_i))(y_j-E(y_j)))=v_i^T S v_j$ 应满足 $v_i^T S v_j=0, i \neq j$ 和 $v_i^T S v_j=1, i=j$,则有:

$$V^T S V = I \quad (8)$$

式中: $S_i = \sum_l (x_i - m)(x_i - m)^T = X \left(I - \frac{1}{n} e e^T \right) X^T$ 为总体散度矩阵,且 $m = \frac{1}{n} \sum x_i, e = (1, \dots, 1)^T$ 。

根据约束条件,目标函数可以转化为:

$$\begin{cases} \max \text{tr}(V^T X H X^T V) \\ \text{s.t. } V^T X G X^T V = I \end{cases} \quad (9)$$

其中: $G = I - \frac{1}{n} e e^T$ 。

由拉格朗日乘子法,最优解可表示为:

$$\frac{\partial}{\partial V} \text{tr}(V^T X H X^T V + \lambda (V^T X G X^T V - I)) = 0 \quad (10)$$

式中: λ 为拉格朗日常数。

进一步简化,公式(10)可表示为求解广义特征值问题,即:

$$V^T X H X^T V = \lambda V^T X G X^T V \quad (11)$$

通过求解公式(11)最大 d 个特征值对应的特征向量,便得到最优投影矩阵 $V=[v_1, v_2, \dots, v_d]$,于是低维嵌入数据可表示为 $Y=V^T X$ 。

3 实验结果与分析

为验证 SNRA 算法的有效性,选取了 Indian Pine 和 KSC 高光谱遥感数据集进行实验,并同 Baseline、PCA、NPE、LPP、SNPE、监督 LPP(SLPP)^[7]和边界费希尔分析(Marginal Fisher Analysis, MFA)^[10]算法进行对比,其中 Baseline 表示直接用原始空间数据进行分类, MFA 为监督方法,同时考虑了类内近邻的聚集性和类间近邻的分离性,但受数据间的边界差异影响较大。

实验中, 分别从各数据集的每类样本中随机选取一定数据量的训练样本, 剩余样本作为测试样本。首先由训练样本对各特征提取算法进行学习, 得到投影矩阵, 进而提取出各样本的鉴别特征, 然后采用最近邻分类器(NN, Nearest Neighbor)对测试样本的特征数据进行分类, 最后通过总体分类精度(OA, Overall Accuracy)、Kappa 系数和平均分类精度(AA, Average Accuracy) 对分类结果进行评价。在每种情况下, 分别进行 10 次重复实验, 并把其平均值和标准差(std, standard deviation)作为最终结果。

3.1 Indian Pine 数据集

Indian Pine 数据集是于 1992 年 6 月 12 日由美国宇航局通过成像光谱仪(AVIRIS)获取的高光谱影像, 该影像拍摄了美国印第安纳州西北 100 km² 区域的地物, 其大小为 145×145 像素, 包含了 0.4~2.5 μm 的光谱范围, 波段数为 220 个, 剔除因大气吸收影响的波段后, 剩余 200 个波段可用于实验研究, 根据各地物的真实空间分布, 对每类地物赋予一个 RGB 颜色值, 可得图 1 所示的地物类别信息分布, 图中括号

内的数字表示像元数。



图 1 Indian Pine 的真实地图

Fig.1 Ground truth of Indian Pine

为比较不同训练样本下各算法的性能, 分别从每类中随机选取 30、40、50、80 个像元作为训练样本, 分别表示为 30-train、40-train、50-train、80-train, 图 2 给出了各方法在不同维度下的总体分类精度曲线, 表 1 给出了各方法的最高分类精度、标准差及其最佳维度。

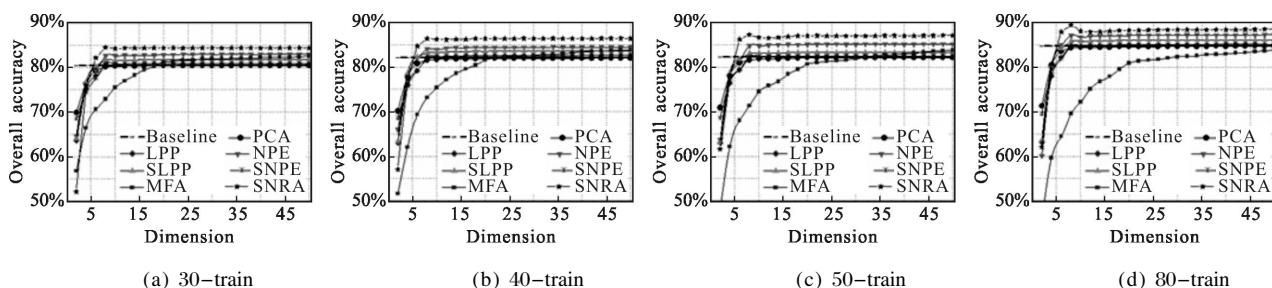


图 2 不同训练样本下各方法的总体分类精度

Fig.2 Overall accuracy of different methods under different number of training samples

表 1 不同方法的最高总体分类精度(OA±std(%))(dimension)

Tab.1 Highest overall accuracy of different methods(OA±std (%))(dimension)

Method	30-train	40-train	50-train	80-train
Baseline	80.50±1.74(200)	82.08±1.65(200)	82.22±1.29(200)	84.72±1.11(200)
PCA	80.48±1.77(46)	82.04±1.62(50)	82.16±1.31(32)	84.67±1.09(32)
LPP	80.91±1.60(8)	82.66±1.43(42)	82.47±1.51(40)	85.09±1.11(42)
NPE	82.98±1.33(44)	84.47±1.09(34)	85.15±1.11(26)	87.22±1.30(32)
SLPP	81.80±1.71(42)	83.64±1.51(44)	83.37±1.48(46)	86.12±1.03(24)
SNPE	80.46±1.72(46)	82.24±1.59(32)	82.43±1.47(46)	84.85±1.15(44)
MFA	82.32±1.90(44)	83.73±1.46(46)	83.78±2.05(50)	83.49±1.03(46)
SNRA	84.49±1.36(8)	86.45±1.35(46)	87.27±1.12(8)	89.48±0.55(8)

由图 2 和表 1 可知,随着低维嵌入维度的增加,分类效果也不断增加并达到平稳,这是因为随着嵌入维度的增加,提取的特征更丰富,更能区分各类数据,得到的分类效果更好,当达到一定的嵌入维度后,得到的信息足以能够反应数据的特征,分类结果将保持稳定。当训练样本增加时,分类精度基本呈增加趋势,这是由于训练样本越多可利用信息越丰富,越能更好地反应各类地物的本质属性,进而得到更好的分类效果。在不同的训练样本下,SNRA 算法比其余算法的分类效果都好,因为 SNRA 不仅保留了数据的局部结构而且增强了数据的可分性,同时通过总体散度矩阵来消除数据间的相关性,进而能更

好地揭示数据的本质属性,提取更有效地鉴别特征,改善分类效果。

研究在平衡训练样本的情况下对分类结果的影响,从每类样本中随机选取 15%的像元作为训练样本,剩余样本作为测试样本,表 2 为对应各类地物的分类精度、总体分类精度、平均分类精度及其 Kappa 系数。

根据表 2 可知,SNRA 算法对每类地物的分类结果大部分都比其余各算法好,而且其总体分类精度、平均分类精度和 Kappa 系数都比其他方法好,表明该方法能更好地揭示高光谱数据的内在特征,提高分类精度。图 3 为各算法的分类结果图,图中

表 2 各算法对每类地物的分类结果

Tab.2 Classification results of each land cover

Class	Training	Test	Baseline	PCA	LPP	NPE	SLPP	SNPE	MFA	SNRA
Hay-windrowed	73	416	98.56%	98.56%	99.04%	99.28%	99.04%	99.28%	99.04%	99.76%
Soybeans-min	370	2 098	88.47%	88.08%	87.61%	88.18%	88.85%	88.61%	82.89%	91.71%
Woods	194	1 100	97.00%	96.82%	97.55%	97.55%	97.64%	97.64%	93.91%	99.45%
Corn-notill	215	1 219	76.62%	76.54%	75.55%	76.37%	77.36%	80.07%	61.03%	81.46%
Grass/Pasture	75	422	94.79%	95.02%	92.89%	94.08%	93.13%	95.50%	71.80%	95.50%
Grass/Trees	112	635	99.06%	98.74%	97.95%	98.58%	97.95%	98.58%	89.45%	97.80%
Average classification accuracy			92.41%	92.29%	91.76%	92.34%	92.33%	93.28%	83.02%	94.28%
Overall classification accuracy			89.92%	89.71%	89.27%	89.81%	90.12%	90.85%	81.48%	92.53%
Kappa coefficient			0.869 6	0.867 0	0.861 2	0.868 3	0.872 2	0.881 8	0.759 5	0.903 2

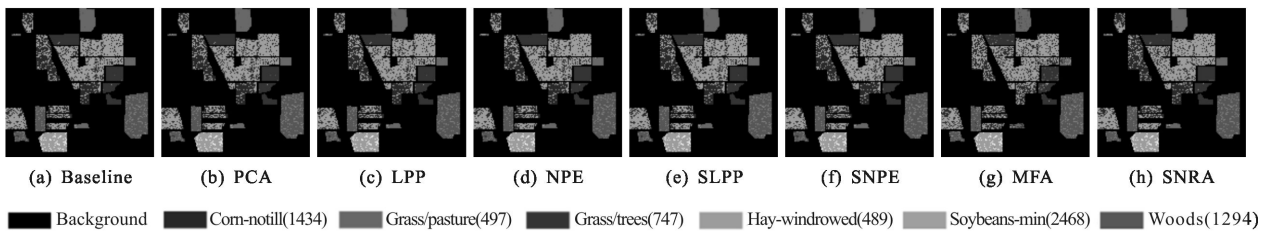


图 3 各方法的分类结果图

Fig.3 Classification maps of different methods

Soybeans-min、Woods 和 Corn-notill 地物区域,由 SNRA 算法得到的结果比其他方法更光滑,总体分类结果更好。

3.2 KSC 数据集

KSC 数据集是于 1996 年 3 月 23 日由美国宇航

局通过成像光谱仪(AVIRIS)获取的高光谱影像,该影像在佛罗里达州肯尼迪空间中心附近上空 20 km 左右航空飞行拍摄,其大小为 614×512 像素,地面分辨率达到 18m,包含了 0.4~2.5 μm 的光谱范围,一共有 224 波段数,去除大气水分吸收及低信噪比的波段

后,把剩余的 176 个波段用于实验研究,根据地物真实空间分布赋予每类地物一个 RGB 值,可得到图 4 所示的地物分布,图中括号内的数字表示像元数。

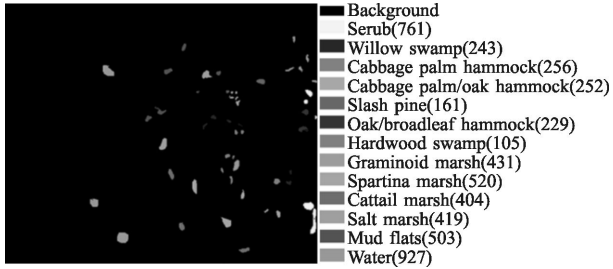


图 4 KSC 的真实地物图
Fig.4 Ground truth of KSC

研究不同训练样本数对各算法性能的影响,图 5 为各方法随嵌入维度变化的总体分类精度,表 3 为各方法的最高分类精度、标准差及其最佳维度。

由图 5 和表 3 可知,在不同训练样本数下,SNPE 和 SLPP 算法分别比 NPE 和 LPP 的分类结果更好,由于 SNPE 和 SLPP 充分利用了已知类别信息,提取出更好的特征,从而改善分类结果。然而,SNRA 算法的精度比其他算法都好,因为 SNRA 算法得到的特征更能揭示数据的本质属性,从而改善分类效果。

在平衡训练样本的情况下,从每类样本中随机选取 15% 的像元作为训练样本,剩余样本作为测试

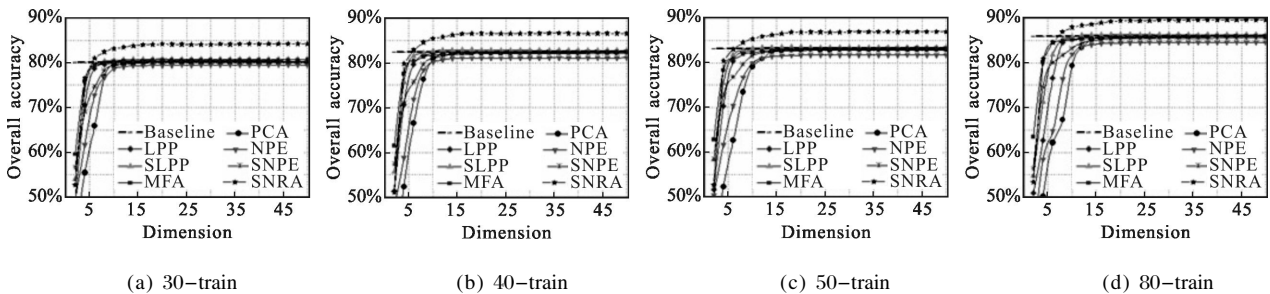


图 5 不同训练样本数下各方法的总体分类精度

Fig.5 Overall accuracy of different methods under different number of training samples

表 3 不同方法的最高总体分类精度(OA±std(%))(dimension)

Tab.3 Highest overall accuracy of different methods (OA±std(%) (dimension))

Method	30-train	40-train	50-train	80-train
Baseline	80.19±0.64(176)	82.38±1.02(176)	83.12±0.74(176)	85.94±0.54(176)
PCA	80.21±0.66(46)	82.40±0.98(34)	83.08±0.72(50)	85.95±0.51(32)
LPP	80.28±0.83(30)	82.52±0.97(16)	83.27±0.66(16)	86.03±0.46(38)
NPE	79.54±0.84(40)	81.11±0.96(30)	81.78±1.12(42)	84.57±0.88(42)
SLPP	80.81±0.85(20)	82.91±1.02(16)	83.52±0.64(16)	86.34±0.49(20)
SNPE	80.15±0.72(40)	82.11±0.76(34)	82.85±0.96(36)	85.56±0.55(42)
MFA	80.70±0.69(50)	82.40±1.06(50)	83.23±0.86(48)	85.96±0.47(50)
SNRA	84.26±0.87(48)	86.65±0.91(34)	86.96±0.72(48)	89.60±0.87(40)

样本,表 4 为其对应各类地物的分类精度、总体分类精度、平均分类精度及其 Kappa 系数,图 6 为对应各方法的分类结果图。

由表 4 和图 6 可知,在平衡样本条件下,SNRA 算法在大部分类别中都能够取得更好的分类结果,并且平均分类精度、总体分类精度和 Kappa 系数都

表 4 各算法对每类地物的分类结果

Tab.4 Classification results of each ground cover

Class	Training	Test	Baseline	PCA	LPP	NPE	SLPP	SNPE	MFA	SNRA
Scrub	114	647	88.87%	88.87%	89.34%	87.64%	90.95%	89.49%	89.34%	90.73%
Willow swamp	36	207	73.43%	73.91%	73.91%	71.98%	74.43%	72.95%	75.36%	84.06%
Cabbage palm hammock	38	218	85.32%	85.32%	84.40%	76.61%	85.86%	82.118%	86.70%	89.45%
Cabbage oak hammock	38	214	50.47%	49.07%	49.07%	47.20%	50.07%	49.53%	52.34%	58.41%
Slash pine	24	137	58.39%	58.39%	56.93%	49.64%	57.20%	52.55%	59.12%	62.04%
Oak/broadleaf hammock	34	195	44.10%	44.62%	45.64%	42.56%	47.15%	46.67%	46.15%	46.67%
Hardwood swamp	16	89	74.16%	75.28%	74.16%	75.28%	76.28%	78.65%	78.65%	91.01%
Graminoid marsh	65	366	83.61%	83.61%	84.15%	82.79%	84.88%	83.88%	84.70%	89.34%
Spartina marsh	78	442	92.76%	92.53%	92.53%	87.56%	93.76%	92.08%	93.89%	95.93%
Cattail marsh	61	343	93.59%	93.00%	93.29%	88.05%	94.29%	93.00%	94.46%	96.50%
Salt marsh	63	356	89.89%	89.89%	89.89%	90.45%	90.89%	90.45%	90.17%	93.54%
Mud flats	75	428	80.61%	80.84%	80.84%	80.37%	83.48%	80.14%	80.61%	89.95%
Water	139	788	99.11%	99.11%	99.11%	98.98%	100.00%	99.11%	99.24%	99.24%
Average classification accuracy			78.02%	78.03%	77.94%	75.32%	79.18%	77.74%	79.29%	83.61%
Overall classification accuracy			84.33%	84.29%	84.36%	82.17%	85.63%	84.13%	85.17%	88.49%
Kappa coefficient			0.825 5	0.825 0	0.825 8	0.801 4	1.828 8	0.823 2	0.834 9	0.871 8

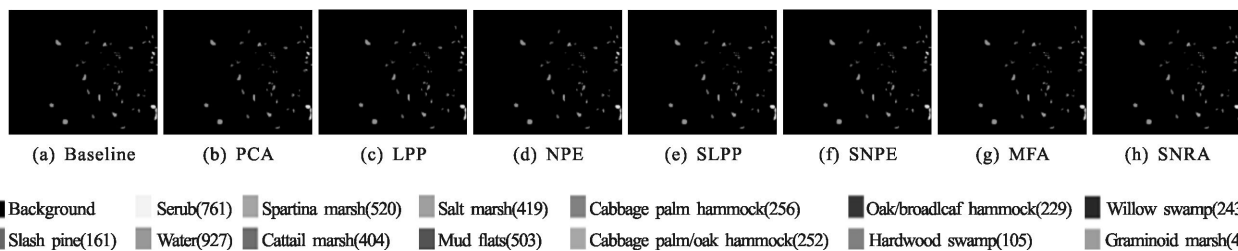


图 6 各方法的分类结果图

Fig.6 Classification maps of different methods

比其他方法好,其分类结果图更光滑。

4 结论

文中在高光谱遥感数据特征提取的研究中,提出了一种新的监督近邻重构分析算法,该方法充分利用已知类别信息在保持数据局部结构的同时尽可能分离非同类数据,并由总体散度矩阵来约束数据间的相关性。该方法能更好地体现数据原有的本

征属性,进而更有效地提取出鉴别特征,改善分类精度。在 Indian Pine 和 KSC 两个真实高光谱遥感数据集上进行了实验,验证了文中提出的方法能够较好地改善高光谱遥感数据地物分类的效果。在文中的研究方法中,需要标记高光谱遥感数据全部训练样本的类别信息,而实际应用中,标记类别信息十分困难,同时需要耗费大量的人力、财力。因此,应用少量的标记样本和大量的无标记本来构建半监督近邻重构分析算法是文中下一步研究的重点。

参考文献:

- [1] Tang Y Y, Yuan H L, Li L Q. Manifold-based sparse representation for hyperspectral image classification [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2014, 52(12): 7606–7618.
- [2] Wang Xiaofei, Hou Chuanlong, Yan Qiuqing, et al. Noise estimation algorithm based on relevance vector machine for hyperspectral imagery [J]. *Infrared and Laser Engineering*, 2014, 43(12): 4159–4163. (in Chinese)
王晓飞, 侯传龙, 阎秋静, 等. 基于相关向量机的高光谱图像噪声评估算法 [J]. 红外与激光工程, 2014, 43(12): 4159–4163.
- [3] Cai Hui, Li Na, Zhao Huijie. Features extraction method based on intrinsic mode function for hyperspectral data [J]. *Infrared and Laser Engineering*, 2013, 42(12): 3475–3480. (in Chinese)
蔡辉, 李娜, 赵慧洁. 基于本征模函数的高光谱数据特征提取方法[J]. 红外与激光工程, 2013, 42(12): 3475–3480.
- [4] Pu Hanye, Wang Bin, Zhang Liming. New dimensionality reduction algorithms for hyperspectral imagery based on manifold learning [J]. *Infrared and Laser Engineering*, 2014, 43(1): 232–237. (in Chinese)
普哈晔, 王斌, 张立明. 基于流形学习的新高光谱图像降维算法[J]. 红外与激光工程, 2014, 43(1): 232–237.
- [5] Ding L, Tang P, Li H Y. Subspace feature analysis of local manifold learning for hyperspectral remote sensing images classification [J]. *Applied Mathematics & Information Sciences*, 2014, 8(4): 1987–1995.
- [6] Bachmann C M, Ainsworth T L, Fusina R A. Exploiting manifold geometry in hyperspectral imagery [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2005, 43(3): 884–897.
- [7] He X, Niyogi P. Locality preserving projections [C]// *Proceedings of the 17th Annual Conference on Neural Information Processing Systems*, 2004, 16: 153–160.
- [8] He X F, Cai D, Yan S C, et al. Neighborhood preserving embedding [C]// *Proceedings of the 10th IEEE International Conference on Computer Vision*, 2005: 1208–1213.
- [9] Jin Zhong, Yang Jingyu, Lu Jianfeng. An optimal set of uncorrelated discriminant features [J]. *Chinese Journal of Computers*, 1999, 22(10): 1105–1108. (in Chinese)
金忠, 杨静宇, 陆建峰. 一种具有统计不相关性的最佳鉴别矢量集[J]. 计算机学报, 1999, 22(10): 1105–1108.
- [10] Yan S, Xu D, Zhang B, et al. Graph embedding and extensions: a general framework for dimensionality reduction [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, 29(1): 40–51.