

引用格式: ZHANG Minghua, LUO Hongling, SONG Wei, et al. Feature Extraction of Hyperspectral Image Based on Sparse Representation and Learning Graph Regularity[J]. Acta Photonica Sinica, 2021, 50(4):0410002

张明华, 罗红玲, 宋巍, 等. 基于稀疏表示和学习图正则的高光谱图像特征提取[J]. 光子学报, 2021, 50(4):0410002

基于稀疏表示和学习图正则的高光谱图像特征提取

张明华¹, 罗红玲¹, 宋巍¹, 黄冬梅^{1,2}, 贺琪¹, 苏诚³

(1 上海海洋大学 信息学院, 上海 201306)

(2 上海电力大学, 上海 200090)

(3 自然资源部东海预报中心, 上海 200136)

摘 要: 针对传统局部特征提取算法难以确定邻域参数, 以及仅考虑数据间的单一结构而漏掉重要信息的问题, 提出一种基于稀疏表示和学习图正则的局部判别与全局稀疏保持投影算法。该算法首先对稀疏表示模型施加基于学习的图正则器, 用该改进的稀疏表示模型自适应揭示样本数据间的局部线性结构, 通过局部判别模型全局集成算法来提取局部线性结构中的判别信息; 利用基于学习图正则稀疏表示模型构建的新型稀疏图来揭示数据间的全局稀疏结构; 使得数据的局部判别结构和全局稀疏结构在低维特征空间得以保持。通过 1-近邻和支持向量机分类器对实验结果进行评估, 在 PaviaU 和 Indian Pines 两个高光谱公共数据集上的实验显示, 提出的局部判别与全局稀疏保持投影算法较对比算法取得了最好的性能, 由于提取了全局和局部的判别信息, 有效提升了高光谱图像的地物分类精度。

关键词: 高光谱图像; 特征提取; 稀疏表示; 局部判别信息; 学习图正则

中图分类号: TP751

文献标识码: A

doi: 10.3788/gzxb20215004.0410002

Feature Extraction of Hyperspectral Image Based on Sparse Representation and Learning Graph Regularity

ZHANG Minghua¹, LUO Hongling¹, SONG Wei¹, HUANG Dongmei^{1,2}, HE Qi¹, SU Cheng³

(1 College of Information Technology, Shanghai Ocean University, Shanghai 201306, China)

(2 Shanghai University of Electric Power, Shanghai 200090, China)

(3 East China Sea Forecast Center, Ministry of Natural Resources, Shanghai 200136, China)

Abstract: Traditional local feature extraction algorithms are difficult to determine neighborhood parameters, and they only consider the single structure information of the data, which ignore important information. To solve the above problems, a Local Discrimination and Global Sparse Preservation Projection Algorithm (LDGSPP) based on sparse representation and learning graph regularity is proposed. The algorithm firstly applies a learning-based graph regularizer to the sparse representation model. Then the improved sparse representation model is used to reveal the local linear structure of the sample data adaptively. The local discriminant model global integration algorithm is used to extract the discriminant information of the local linear structure. The new sparse graph constructed by the improved sparse representation model is used to reveal the global sparse structure of data. The local discriminant structure and the global sparse structure of the data are preserved in the low dimensional feature space. 1-nearest

基金项目: 国家自然科学基金 (Nos. 61972240, 41906179), 上海市科委部分地方高校能力建设项目 (No. 20050501900)

第一作者: 张明华 (1977—), 女, 副教授, 博士, 主要研究方向为海洋大数据管理及遥感图像处理。Email: mhzhang@shou.edu.cn

通讯作者: 宋巍 (1977—), 女, 教授, 博士, 主要研究方向为海洋遥感图像分析、海洋大数据认知和理解。Email: wsong@shou.edu.cn

收稿日期: 2020-10-16; 录用日期: 2021-02-01

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

neighbors and support vector machine classifier are used to evaluate the experimental results. The experiments on PaviaU and Indian Pines show that LDGSPP achieves the best performance compared with the comparison algorithm. As global and local discriminant information is extracted, the ground object classification accuracy of hyperspectral images is effectively improved.

Key words: Hyperspectral image; Feature extraction; Sparse representation; Local discriminant information; Learning graph regularization

OCIS Codes: 100.4145; 110.4234; 100.5010; 100.2960; 100.3008

0 引言

高光谱图像(Hyperspectral Image, HSI)包含从可见光到红外光谱的数百个光谱波段,将每个高光谱像元根据其丰富光谱特征进行分类已经成为遥感领域的一个重要课题^[1],在地质勘查、医疗检测、生命科学、司法鉴别、军事勘察、环境监测、精细农业等领域得到广泛的应用^[2-3]。但是纳米级的谱分辨率造成了严重的谱相关和谱冗余,成百上千个波段对于数据的传输、存储和处理带来了巨大挑战^[4]。同时,光谱分辨率的提高也引起了“维数灾难”现象,即在有限样本的前提下,较高的特征维度反而会降低分类精度^[5-6]。针对这些问题,特征提取是有效的解决途径之一。通过特征提取将高光谱数据映射到低维子空间,减少冗余信息,同时保持最高可能的光谱分辨力,以提高高光谱图像分类性能。

主成分分析(Principal Component Analysis, PCA)^[7],最小噪声分数(Minimum Noise Fraction, MNF)^[8]为两种经典的特征提取方法。但是PCA和MNF只考虑数据的统计特征,忽视了数据的内在结构,而实际的高光谱图像数据中包含了固有的流形结构。近些年来大量基于流形学习的特征提取算法被提出,常见的有局部保持投影(Locality Preserving Projections, LPP)^[9]和邻域保持嵌入(Neighborhood Preserving Embedding, NPE)^[10],这两种算法均利用近邻图来揭示数据的局部结构,在低维嵌入时保持数据的局部结构。大量研究已证明局部判别信息相当重要^[11]。为了同时保留数据局部几何结构和局部判别信息,局部判别模型全局集成^[12](Local Discriminant Models and Global Integration, LDMGI)利用局部判别分数为每个局部流形构建局部判别模型。基于流形学习的特征提取算法有效提取了高光谱图像的局部信息,但邻域参数的选择依赖于专家经验,难以确定。另外,上述方法忽略了数据间的全局流形结构。

近些年,由于稀疏表示可获得天然的判别信息而被广泛应用在信号处理等应用中。通过稀疏表示可自适应揭示数据的局部线性结构,且利用稀疏表示构建的稀疏图是一种重要的信息图,所捕获的数据全局稀疏结构保留了有利于分类任务的判别信息。许多基于稀疏表示的算法如稀疏保持投影算法(Sparsity Preservation Projection, SPP)被提出,SPP利用稀疏表示构建图,应用稀疏子空间理论在高维空间中寻找保持数据全局稀疏结构所在低维子空间^[13]。稀疏保持分析(Sparsity Preserving Analysis, SPA)应用稀疏表示的自适应性来构造一个揭示数据稀疏特性的图^[14]。多结构流形嵌入(Multi-structure Manifold Embedding, MSME)通过稀疏表示设计稀疏图和稀疏超图,并结合这些图揭示高维数据中的固有稀疏结构^[6]。但是基于L1范式的稀疏表示模型由于稀疏分解时潜在的不稳定性,产生不稳定的稀疏表示。

为了缓解这种不稳定性,一些改进的稀疏表示模型被应用在高光谱图像特征提取中。约束稀疏图(LE-based Sparse Representation, LESR)在稀疏优化中加入流形正则化项来平滑稀疏分解^[15]。基于局部流形的稀疏判别学习(Local Manifold-based Sparse Discriminant Learning, LMSDL)设计了一种新的稀疏优化模型,利用基于欧式距离的相似性权重对稀疏表示模型施加约束,以揭示基于局部流形的数据稀疏结构^[16]。虽然上述稀疏表示模型通过施加流形约束使得稀疏分解更加稳定,但所施加的流形约束依赖于预定义相似度量矩阵的质量,而通过高斯核函数预定义的相似度量矩阵通常受噪声等因素的影响,设计一个合适的预定义相似度量矩阵往往是有困难的。

针对以上问题,本文提出基于稀疏表示和学习图正则的局部判别与全局稀疏保持投影算法(Local Discriminant and Global Sparse Preserving Projection, LDGSPP)。该算法基于流形正则化项提出学习图正则约束的稀疏表示模型,自适应获取局部线性结构,解决了LDMGI算法中使用 k 近邻图时邻域参数选择困难问题。同时,该算法还可以利用基于学习图正则的稀疏表示模型来提取数据中的全局稀疏结构。通过将局部判别信息和全局稀疏结构相结合,提取即包含局部又包含全局的判别性特征,从而提高最终的分类

性能。

1 局部判别与全局稀疏保持投影算法

本节主要描述所提特征提取算法,首先介绍了高光谱图像的预处理算法高斯加权空-谱算法(Gaussian Weighted Spatial Spectral, GWSS),以融入样本的空间上下文信息。然后介绍了基于学习图正则的稀疏表示(Sparse Representation based on Learning Graph Regularization, SR-LGR)。最后,结合基于学习图正则的稀疏表示,介绍 LDGSPP 算法。并给出了基于 ADMM 的优化器的求解方式。图 1 为 LDGSPP 算法的流程。

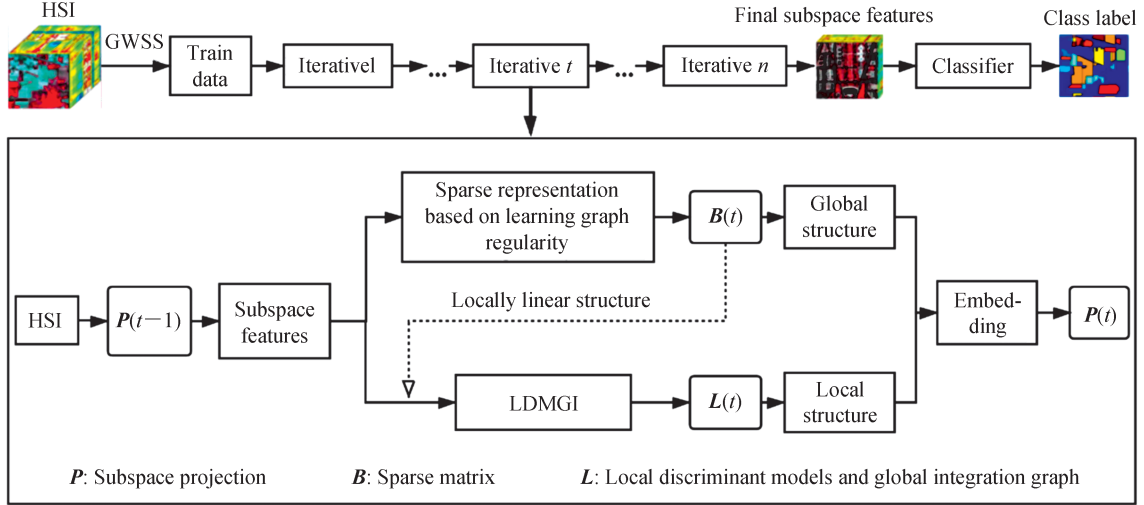


图 1 LDGSPP 算法流程

Fig1 Flowchart of algorithm LDGSPP

算法首先使用 GWSS 融入空间上下文信息,然后经过多次迭代求得最终特征提取后的特征空间。图 1 实线框内为一次迭代过程,矩阵 $B(t)$ 通过施加了学习图正则约束的稀疏表示模型得到的新型稀疏矩阵。一方面用稀疏矩阵 $B(t)$ 捕获数据的全局稀疏结构;另一方面用稀疏矩阵 $B(t)$ 自适应揭示每个样本的局部邻域,用 LDMGI 算法捕获局部块中的判别信息,求取所有局部块的判别信息得到局部判别模型的全局集成图 $L(t)$, $L(t)$ 捕获了数据的局部判别结构。最终将两种结构进行低维嵌入得到本次迭代的映射矩阵 $P(t)$ 。

1.1 高斯加权空-谱算法

在高光谱图像中,采用式(1)规则重构像素

$$\hat{x}_i = \frac{\sum_{x_j \in N(x_i)} \sigma_j x_j}{\sum_{x_j \in N(x_i)} \sigma_j} \quad (1)$$

式中, $N(x_i)$ 为样本点 x_i 的 $\omega \times \omega$ 近邻空间, σ_j 是像素点 x_i 与它的近邻空间像素 x_j 的权重值,通过高斯核函数定义。本文算法通过这种规则来融合像素点 x_i 的空间上下文信息。

1.2 基于学习图正则的稀疏表示

假设数据集 $X = \{x_1, x_2, \dots, x_n\}$, $x_i \in \mathbb{R}^{1 \times d}$ ($1 \leq i \leq n$), d 为维度。通过求解式(2)得到稀疏图矩阵 B

$$B = \arg \min_B \{ \|X - XB\|_F^2 + \lambda \|B\|_1 \} \quad (2)$$

式中, $B = [b_1, b_2, \dots, b_n]^T \in \mathbb{R}^{n \times n}$ 。

本文对稀疏图施加基于学习的图正则,构建了一种新型的稀疏图。该图正则遵循相似的样本所得编码向量是相似的这一准则,可表示为

$$\sum_{i,j} \|b_i - b_j\|^2 W_{ij} = \text{Tr}(B(D - W)B^T) = \text{Tr}(BLB^T) \quad (3)$$

式中, W_{ij} 为样本点 x_i 和 x_j 间的权重系数, L 是图矩 W 的拉普拉斯矩阵, $L = D - W$, $D_{ii} = \sum_j W_{ij}$, $\text{Tr}(S)$ 是

矩阵 S 的迹。

式(4)同时集成稀疏表示和图正则,求新型稀疏图矩阵 B

$$B = \arg \min_B \|X - XB\|_F^2 + \lambda \|B\|_1 + \rho \text{Tr}(BL_B B^T) \quad (4)$$

式中, L_B 是 B 的拉普拉斯矩阵, λ 和 ρ 为两个正则参数。 L_B 连续, 式(4)即施加学习图正则的 L1 范数求解, 采用递归策略求解 B 。具体参考节本节 3) 中算法 1 描述。

1) 初始化图矩阵 W , 可得图拉普拉斯 $L_B = D - W$ 。

2) 固定变量 L_B , 求解式(4), 引入辅助变量 C , 使得 $B = C$, 可得

$$\min v \left[\|X - XB\|_F^2 + \lambda \|C\|_1 + \rho \text{Tr}(BL_B B^T) \right], s.t. B = C \quad (5)$$

式(5)的增广拉格朗日乘式形式为

$$Q(B, C, M) = \|X - XB\|_F^2 + \lambda \|C\|_1 + \rho \text{Tr}(BL_B B^T) + \frac{\gamma}{2} \|B - C\|^2 + \text{Tr}[M^T (B - C)] \quad (6)$$

式中, γ 为正则参数, M 为拉格朗日乘子矩阵。分别对三个变量求偏导

$$\frac{\partial Q(B, C, M)}{\partial B} = X^T XB - X^T X + 2\rho BL + \gamma(B - C) + M \quad (7)$$

$$\frac{\partial Q(B, C, M)}{\partial C} = \gamma(B - C) + \lambda \text{sgn}(C) + M \quad (8)$$

$$\frac{\partial Q(B, C, M)}{\partial M} = B - C \quad (9)$$

利用 ADMM 求解变量 B, C, M 。

a) 固定变量 C, M 求解 B 。

$$(X^T X + \gamma I)B + 2\rho BL = X^T X + \gamma C - M \quad (10)$$

b) 固定变量 B, M , 求解 C 。

$$C = S(B + M/\gamma, \lambda/\gamma) \quad (11)$$

c) 固定变量 B, C , 用梯度上升法更新 M 。

$$M_{j+1} = M_j + \gamma \left. \frac{\partial Q(B, C, M)}{\partial M} \right|_{B=B_{j+1}, C=C_{j+1}, M=M_j} \quad (12)$$

3) 固定变量 B , 根据 $W = \frac{1}{2} (|B| + |B^T|)$ 更新图 W 。

算法 1: 基于学习图正则的稀疏表示算法 (SR-LGR)

输入: 数据 X , 映射矩阵 P , 参数 λ, ρ , 阈值 τ_1, τ_2 , 最大迭代次数 I ;

输出: B ;

初始化: W_0 和 C_0 是随机产生的矩阵, M_0 为零矩阵, τ_0 ;

1: While ($i < I$ or $\text{obj}1 > \tau_1$);

2: While ($j < I$ or $\text{obj}2 > \tau_2$);

3: $(X^T X + \gamma I)B_{j+1} + 2\rho B_{j+1}L = X^T X + \gamma C_j - M_j$, calculate B_{j+1} ;

4: $C_{j+1} = S(B_{j+1} + M_j/\gamma, \lambda/\gamma)$;

5: $M_{j+1} = M_j + \gamma(B_{j+1} - C_{j+1})$;

6: $\gamma_{j+1} = \max(5, 1.1\gamma_j)$;

7: $\text{obj}2 = \|B_j - C_j\|_F^2$;

8: $j++$;

9: $W_{i+1} = \frac{1}{2} (|B_{i+1}| + |B_{i+1}^T|)$;

10: $\text{obj}1 = \|B_{i+1} - B_i\|_F^2$;

11: $i++$

算法1中西尔维斯特方程($AX+XB=C$)采用文献[17]中的方法解决;软阈值函数的定义为 $S(x,y)=\text{sign}(x)*(\max|x|-y,0)$,作用于向量或矩阵的每个元素。

1.3 局部判别与全局稀疏保持投影

利用上述稀疏图 B 自适应揭示的局部线性结构可表示为 $\mathbf{Z}(x_i), \mathbf{Z}(x_i)(1 \leq i \leq n)$ 中所有像元为训练样本集 $\mathbf{X}, \mathbf{Z}(x_i)$ 内样本记为 \mathbf{X}_i 。将式(4)改进稀疏表示模型和文献[12]中LDMGI模型同时低维嵌入,得到本文所提算法LDGSP目标函数,即

$$\min v \left[\left\| \mathbf{P}^T \mathbf{X} - \mathbf{P}^T \mathbf{X} \mathbf{B} \right\|_F^2 + \lambda \|\mathbf{B}\|_1 + \rho \text{Tr}(\mathbf{B} \mathbf{L}_B \mathbf{B}^T) \right] + \eta \left\| \mathbf{X} - \mathbf{P} \mathbf{P}^T \mathbf{X} \right\|_F^2 + \text{Tr}(\mathbf{P}^T \mathbf{X} \mathbf{L} \mathbf{X}^T \mathbf{P}), \text{ s.t. } \mathbf{P}^T \mathbf{X} \mathbf{X}^T \mathbf{P} = \mathbf{I} \quad (13)$$

式中, v, η 均是正则参数。 $\mathbf{L} = \sum_{i=1}^n \mathbf{S}_i \mathbf{L}_i \mathbf{S}_i^T = [\mathbf{S}_1, \dots, \mathbf{S}_n] \begin{bmatrix} L_1 & & \\ & \dots & \\ & & L_n \end{bmatrix} [\mathbf{S}_1, \dots, \mathbf{S}_n]^T, \mathbf{S}_i \in \mathbb{R}^{n \times k}$ 为选择矩阵, $F_i = \{i_0, i_1, \dots, i_{k-1}\}$ 为样本 x_i 的邻域内每个样本的索引集, $p = F_i\{q\}$ 时, $(\mathbf{S}_i)_{pq} = 1$,否则 $(\mathbf{S}_i)_{pq} = 0, k$ 表示第 i 个样本的局部流形结构中包含的像元个数。 $\mathbf{L}_i = \mathbf{H}_k (\bar{\mathbf{X}}_i^T \bar{\mathbf{X}}_i + u\mathbf{I})^{-1} \mathbf{H}_k$,中心化矩阵记为 $\mathbf{H}_n = \mathbf{I}_n - \frac{1}{n} \mathbf{I}_n \mathbf{I}_n^T, \bar{\mathbf{X}} = \mathbf{X} \mathbf{H}_n$ 为中心化后的样本集, u 为正则参数。

式(13)第一项为稀疏表示。第二项表示数据集 \mathbf{X} 在映射矩阵 \mathbf{P} 的作用下,子空间能更好地被重构。第三项为局部判别模型的全局集成。 \mathbf{L} 是一种新型的拉普拉斯矩阵,即基于稀疏图的局部判别模型的全局集成图,可用来捕获数据间的局部聚类结构。

以往基于图嵌入的高光谱特征提取框架中,低维嵌入均在解决广义特征值分解(Generalized Eigenvalues Decomposition, GED)问题^[18]。与之不同,所提算法目标函数为一个非凸函数,故采用交替迭代策略优化凸优化子问题,即分别求解 \mathbf{B}, \mathbf{P} 的局部或全局最小值,代替采用分离变量的策略直接求解非凸优化策略。具体参考本节2)中算法2描述。

1)固定变量 \mathbf{P} ,求解式(13)转换为求解子问题

$$\min v \left[\left\| \mathbf{P}^T \mathbf{X} - \mathbf{P}^T \mathbf{X} \mathbf{B} \right\|_F^2 + \lambda \|\mathbf{B}\|_1 + \rho \text{Tr}(\mathbf{B} \mathbf{L}_B \mathbf{B}^T) \right] \quad (14)$$

2)固定 \mathbf{B} ,求解式(13)转换为求解子问题

$$\begin{aligned} & \min_{\text{s.t. } \mathbf{P}^T \mathbf{X} \mathbf{X}^T \mathbf{P} = \mathbf{I}} v \left[\left\| \mathbf{P}^T \mathbf{X} - \mathbf{P}^T \mathbf{X} \mathbf{B} \right\|_F^2 + \eta \left\| \mathbf{X} - \mathbf{P} \mathbf{P}^T \mathbf{X} \right\|_F^2 + \text{Tr}(\mathbf{P}^T \mathbf{X} \mathbf{L} \mathbf{X}^T \mathbf{P}) \right] \\ & \Rightarrow \min_{\text{s.t. } \mathbf{P}^T \mathbf{X} \mathbf{X}^T \mathbf{P} = \mathbf{I}} \text{Tr}(\mathbf{P}^T (v \mathbf{\Gamma} \mathbf{\Gamma}^T - \eta \mathbf{X} \mathbf{X}^T + \mathbf{X} \mathbf{L} \mathbf{X}^T) \mathbf{P}) \end{aligned} \quad (15)$$

式中, $\mathbf{\Gamma} = [\chi_1, \chi_2, \dots, \chi_n], \chi_i = x_i - \mathbf{X} b_i$,通过式(16)可求得 \mathbf{P}

$$(v \mathbf{\Gamma} \mathbf{\Gamma}^T - \eta \mathbf{X} \mathbf{X}^T + \mathbf{X} \mathbf{L} \mathbf{X}^T) \zeta = \lambda \mathbf{X} \mathbf{X}^T \zeta \quad (16)$$

投影矩阵 \mathbf{P} 由式(16)中 d 个最小的广义特征值对应的特征向量 $\zeta_1, \zeta_2, \dots, \zeta_d$ 组成。

算法2:局部判别全局稀疏保持算法(LDGSP)

输入:数据 \mathbf{X} ,参数 λ, ρ, η, v ,阈值 τ_1, τ_2, τ_3 ,最大迭代次数 I ;

输出: \mathbf{P} ;

初始化: $\mathbf{P} = \mathbf{P}_{\text{PCA}}$;

1: While($k < I$ or $\text{obj3} > \tau_3$);

2: $\mathbf{B}_{k+1} = \text{SR} - \text{LGR}(\mathbf{X}, \lambda, \rho)$;

3: \mathbf{P} can be obtained from Eq.(16);

4: $\text{obj3} = \left\| \mathbf{P}_{k+1} - \mathbf{P}_k \right\|_F^2$;

5: $k++$;

2 实验

2.1 数据集描述

本文选择了PaviaU和Indian Pines两个高光谱遥感图像数据集来检测所提LDGSP算法的性能。

1) PaviaU数据集: 图像大小为 610×340 像素, 空间分辨率为 1.3 m; $0.43 \sim 0.86 \mu\text{m}$ 光谱范围内共 115 个波段, 选择其中 103 个无严重噪声影响的波段用于实验; 数据集包含 9 类地物, 可参考样本数据共 42 776 个; 假彩色图像和真实地物图如图 2。

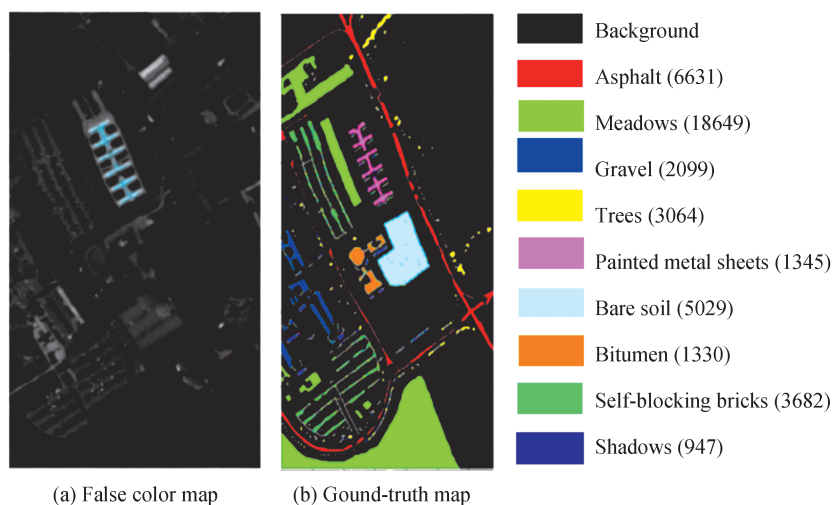


图 2 PaviaU 数据集
Fig. 2 PaviaU dataset

2) Indian Pines数据集: 图像大小为 145×145 像素, 空间分辨率为 20 m; $0.4 \sim 2.5 \mu\text{m}$ 光谱范围内共 220 个波段, 去除噪声影响严重的波段后可用波段为 200 个; 数据集包含 16 类地物, 可参考样本数据共 10 249 个; 假彩色图像和真实地物图如图 3。

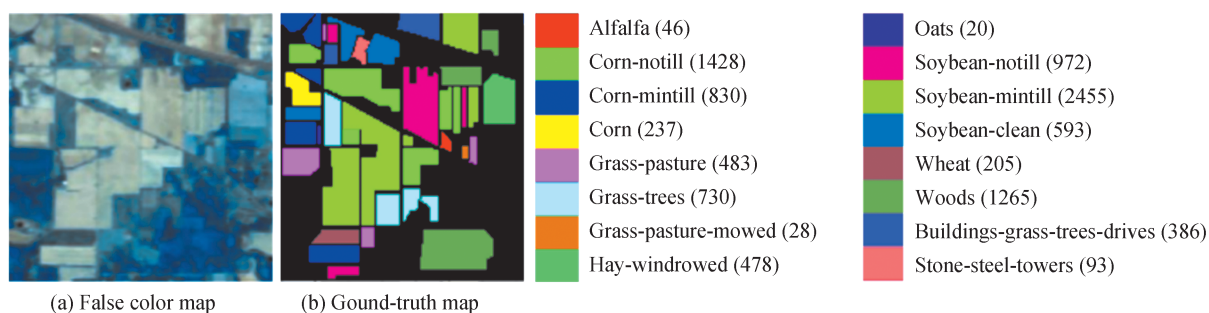


图 3 Indian Pines 数据集
Fig. 3 Indian Pines dataset

2.2 实验设置

本文随机从每个类中选取 60 个样本作为训练样本, 其中样本总数不足 100 个的类选取 15 个样本作为训练样本, 其余样本作为测试集。使用最近邻分类器(1-NN)和支持向量机分类器(SVM)对低维特征进行逐像素分类, 并使用总体精度(OA)、平均精度(AA)和 Kappa 系数(KA)对特征提取算法进行评估。所有实验重复进行 10 次求取平均值。

将所有原始特征(Original spectral feature, OSF)直接用于分类作为基准, 通过对比 PCA、LPP、LDMGI、SPA、MSME、LESR、LMSDL 来评估 LDGSP 的性能。在实验中, 采取交叉验证的方式来选取最优参数。其中 LDGSP 中正则参数 η 的选择范围 $\{0.2, 0.5, \dots, 2.3, 2.6, 2.9\}$, 参数 v 的选择范围 $\{0.5, 1, \dots, 4, 4.5, 5\}$, 参数 λ, ρ 在 $\{10^{-7}, 10^{-6}, \dots, 10^0, 10^1, 10^2\}$ 中选择, 参数 u 在 $\{10^{-8}, 10^{-6}, \dots, 10^4, 10^6, 10^8\}$ 中选择。PaviaU 数据集上各方法嵌入维度统一为 20, 由于 Indian Pines 数据集原始波段数远大于 PaviaU 数据集, 故根据经验统一设置嵌入维度为 $30^{[6]}$ 。

2.3 实验结果与分析

2.3.1 PaviaU数据集分类精度分析

表1为各方法取最佳参数的实验结果,图4为各方法使用SVM分类器得到的分类图。当使用1-NN分类器时,本文所提算法OA较LDMGI提高了14.1%,说明通过稀疏表示自适应选取局部邻域的优势。LDGSP较LESR、LMSDL分类精度分别提高了13.6%、7.5%,体现了同时保留局部判别结构和全局稀疏结构的作用,同时表明本文所提出的学习图正则的有效性。而LESR较SPA和MSME分类精度分别提高了1.5%、3.6%,表明通过施加流形正则项可在一定程度缓解稀疏表示的不稳定性问题。使用SVM分类器有同样的趋势。从图4的分类结果可看出LDGSP算法得到的分类识别图错误点最少。

表1 PaviaU数据集在不同方法上的分类性能(%)

Table 1 Classification performance of PaviaU dataset with various methods(%)

Method	1-NN								SVM											
	OSF	PCAL	PPL	DMG	ISP	AMS	ELES	RMSD	LLD	GS	POS	FPCAL	PPL	DMG	ISP	AMS	ELES	RMSD	LLD	LDGSP
Class1	70.2	70.4	74.2	73.8	71.0	69.2	74.4	81.8	78.2	75.8	78.2	81.5	78.9	77.8	74.6	79.3	89.7	92.3		
Class2	70.1	70.2	72.0	70.9	70.8	67.6	71.8	78.1	91.6	79.0	79.5	70.5	69.6	80.9	70.4	85.2	82.1	94.9		
Class3	68.3	68.2	73.0	65.0	70.1	63.8	67.5	78.1	80.9	84.4	82.5	80.9	73.7	83.7	82.1	84.7	84.6	78.4		
Class4	90.9	90.9	90.6	93.4	91.5	91.0	92.3	92.5	93.1	92.4	90.1	90.8	96.0	92.6	93.2	94.2	96.2	96.3		
Class5	98.9	98.9	99.8	99.9	98.9	98.4	99.5	100	100	99.5	99.5	100	99.9	99.6	99.5	99.6	100	100		
Class6	70.0	69.9	75.1	70.3	67.4	67.7	71.1	78.7	94.8	80.7	82.9	83.2	89.0	83.7	79.8	85.5	84.6	97.7		
Class7	90.5	90.2	83.2	89.8	89.4	87.2	88.7	92.5	97.8	92.2	92.0	81.1	91.8	93.1	91.7	92.9	91.8	98.5		
Class8	70.5	70.5	69.0	72.1	69.2	67.5	70.4	77.0	77.1	76.7	76.6	75.8	77.9	75.8	75.8	80.7	83.2	88.6		
Class9	99.8	99.9	99.9	99.8	99.8	100	99.8	99.7	100	99.5	99.5	99.7	99.8	99.7	99.9	99.7	99.8	100		
AA(%)	81.0	81.0	81.9	81.7	80.8	79.2	81.7	86.5	90.4	86.7	86.8	84.8	86.3	87.4	85.2	89.1	90.2	94.1		
OA(%)	73.6	73.7	75.6	74.7	73.7	71.6	75.2	81.3	88.8	81.2	81.8	77.9	78.3	82.6	76.9	85.5	86.0	94.0		
KA(%)	66.5	66.6	68.9	60.8	66.6	64.0	68.4	75.9	85.4	75.8	76.5	72.1	72.7	77.6	70.8	81.2	81.9	92.0		

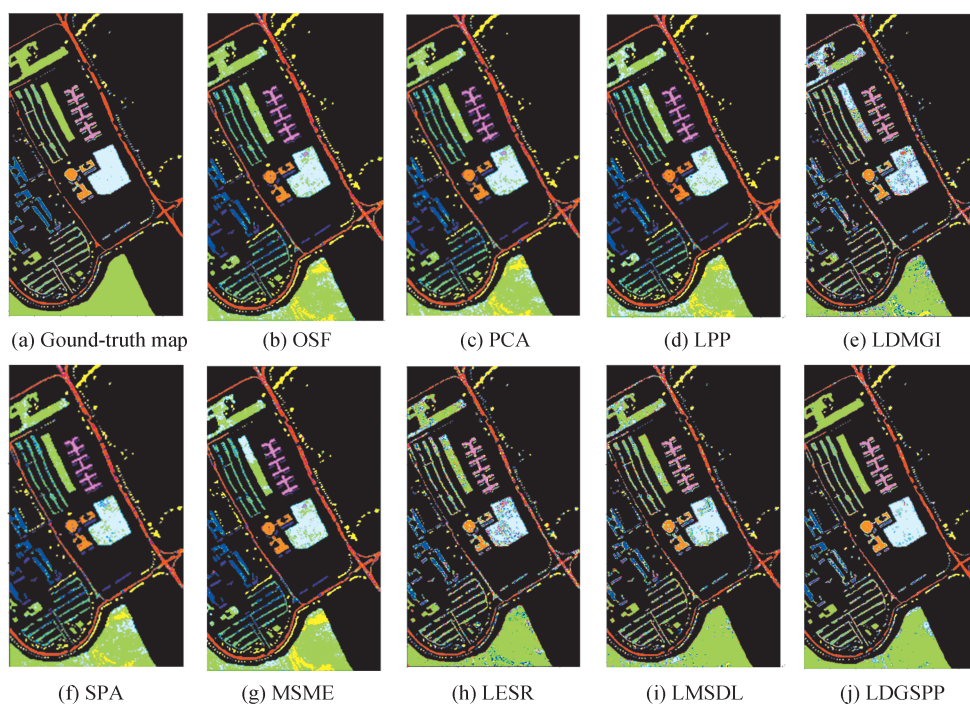


图4 在PaviaU数据集上各算法的分类结果

Fig. 4 Classification results of PaviaU dataset with different algorithms

对降维后的前三维主成分进行可视化,得到图5三维空间投影图,图5(a)为原始光谱前三个波段的可视化图,图5(b)~(i)为不同特征提取算法特征提取后前三主成分的可视化。可以明显看出所有特征提取算法投影图在可分性上均优于OSF。由于稀疏表示的天然判别性,SPA、MSME、LESR、LMSDL在图中均表现出一定的可分性。而同时考虑了局部判别信息和全局稀疏结构的LDGSPP表现出很强的可分性。

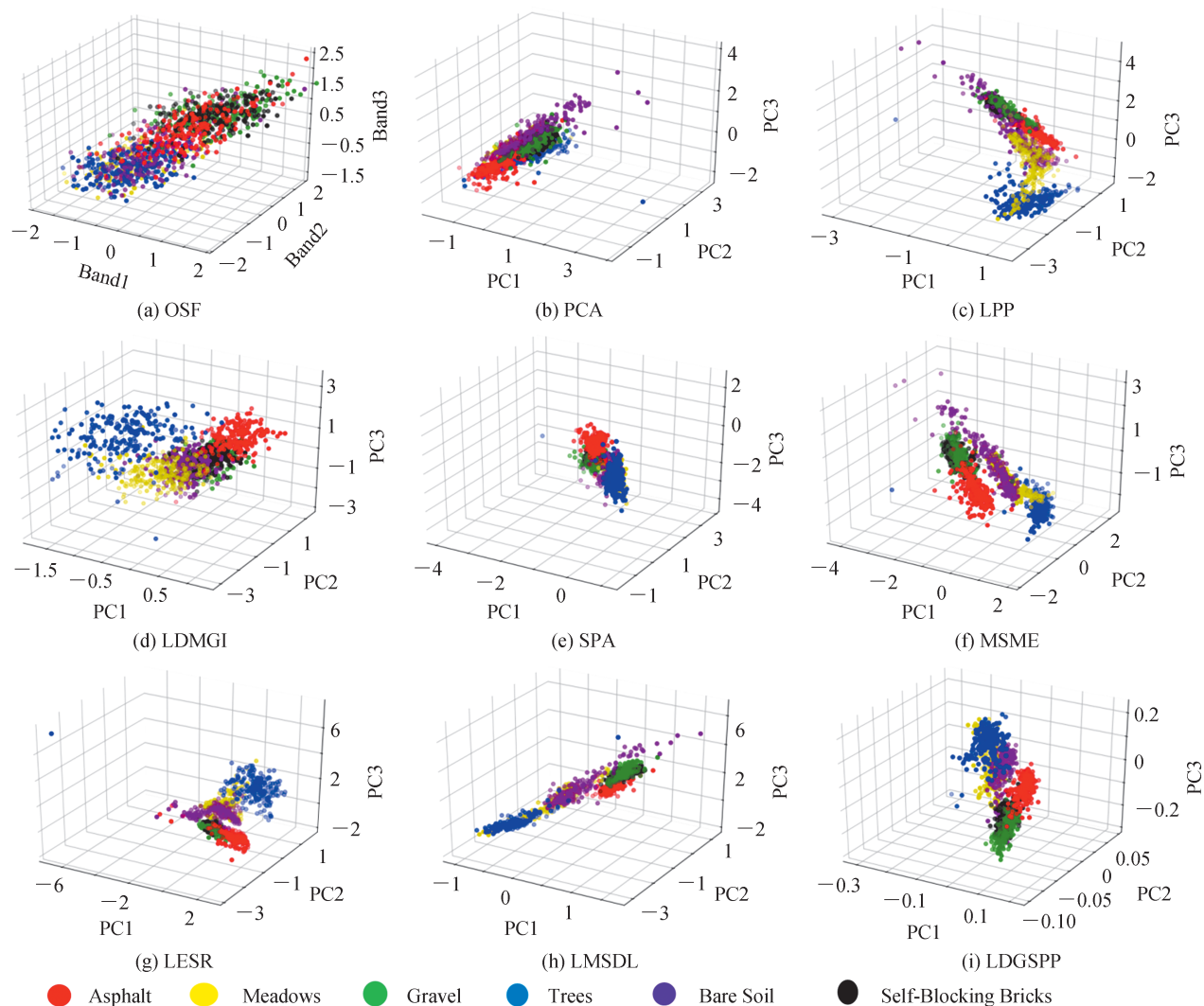


图5 在PaviaU数据集上各算法3D投影图

Fig. 5 3D projection graph of PaviaU dataset with different algorithms

2.3.2 Indian Pines 数据集分类精度分析

表2为各方法取最佳参数的实验结果,图6为各方法使用SVM分类器得到的分类图。从实验结果可以直观看出,当使用1-NN分类器时,LDGSPP较所有对比方法,取得了最高精度,相比LDMGI算法,分类精度提高幅度达到了14.8%,说明了LDGSPP能有效地提取更多的判别信息,为分类作出贡献。使用SVM分类器也得到相似的结论。

与PaviaU数据集不同的是,Indian Pines数据集明显存在着样本不平衡以及可用于实验的样本数极少的挑战,其中“Alfalfa”、“Grass-pasture-mowed”、“Oats”等地物可用于实验的样本数均不足60个,十分有限的样本个数使得分类性能受限。但1-NN分类器下所得各降维方法分类结果显示,这三类地物使用LDGSPP特征提取算法所得分类精度最高,较OSF分别提高了29.1%、7.7%、20%。这说明了本文提出的算法在有限样本上可以提取更多有效的信息。

表2 Indian Pines数据集在不同方法上的分类性能(%)
Table 2 Classification performance of Indian Pines dataset with various methods(%)

Method	1-NN									SVM								
	OSF	PCALP	PLDMG	ISPAMS	MELES	RMSD	LLDG	SPOS	FP	PCALP	PLDMG	ISPAMS	MELES	RMSD	LLDG	SPP		
Class1	48.3	41.9	48.4	41.9	64.5	48.4	64.5	64.5	77.4	35.4	48.4	58.1	64.5	80.6	61.3	80.6	64.5	87.1
Class2	43.1	41.8	43.9	35.2	43.8	41.5	43.5	59.4	60.8	52.1	51.6	60.3	47.2	61.5	54.9	60.5	75.4	85.2
Class3	53.3	52.9	53.4	51.8	51.4	52.9	55.9	62.7	66.6	59.8	58.7	56.6	59.3	60.0	53.1	68.9	78.1	83.2
Class4	66.6	70.1	65.5	61.0	66.1	63.3	62.1	76.2	92.6	77.9	75.1	73.4	74.5	75.1	74.0	75.1	87.0	96.0
Class5	90.3	90.3	89.8	84.1	91.3	88.9	81.0	93.8	95.9	90.3	91.5	94.3	90.5	92.4	92.0	87.9	95.2	95.9
Class6	91.6	89.0	92.4	85.5	88.1	90.0	89.7	92.5	98.5	91.9	91.2	90.0	88.6	91.9	90.7	92.8	97.4	99.4
Class7	92.3	84.6	92.3	84.6	92.3	92.3	61.5	84.6	100	92.3	92.3	92.3	92.3	92.3	92.3	84.6	84.6	100
Class8	88.7	87.3	84.9	87.0	86.6	84.9	92.1	93.5	94.7	95.2	94.7	88.0	89.2	95.0	95.5	96.6	98.3	99.7
Class9	80	80.0	80.0	100	80.0	60.0	100	100	100	100	100	100	100	100	100	100	100	100
Class10	64.4	63.2	60.0	58.9	62.7	62.6	65.4	76.6	70.1	67.5	68.9	61.4	72.0	65.1	63.7	75.5	80.0	83.0
Class11	47.2	46.9	46.7	46.6	47.4	46.2	47.7	55.7	56.2	55.7	55.4	54.2	53.6	52.9	56.2	62.5	64.1	72.5
Class12	48.9	46.9	47.8	39.7	51.2	46.5	53.1	59.6	73.3	55.9	52.5	60.0	55.7	61.2	60.6	69.6	80.4	81.6
Class13	97.9	95.9	96.6	97.2	96.6	95.2	96.5	99.3	98.6	98.6	98.6	97.9	98.6	97.9	98.6	96.5	99.3	98.6
Class14	75.4	75.7	73.8	84.4	74.9	76.2	77.5	85.6	86.5	76.5	76.7	80.7	88.2	81.8	78.7	84.4	93.2	89.8
Class15	48.7	46.9	52.8	41.7	47.8	48.2	49.6	74.5	87.1	64.7	58.6	58.9	65.3	69.6	62.6	74.8	86.8	86.5
Class16	85.9	85.9	85.9	84.6	88.5	84.6	85.9	85.9	90.1	62.8	73.1	89.7	83.3	88.5	89.7	84.6	85.9	98.7
AA(%)	70.2	68.7	69.6	67.8	70.8	67.6	70.4	79.0	84.3	73.5	74.2	76.0	76.4	79.1	76.5	80.9	85.6	91.1
OA(%)	60.9	60.1	60.2	58.4	60.6	59.7	61.5	70.8	73.2	66.6	66.2	67.1	67.0	68.6	66.9	73.7	79.1	84.7
KA(%)	56.0	55.1	55.2	53.5	55.6	54.7	56.7	67.1	69.8	62.3	61.9	62.8	62.7	64.5	62.6	69.9	77.6	82.6

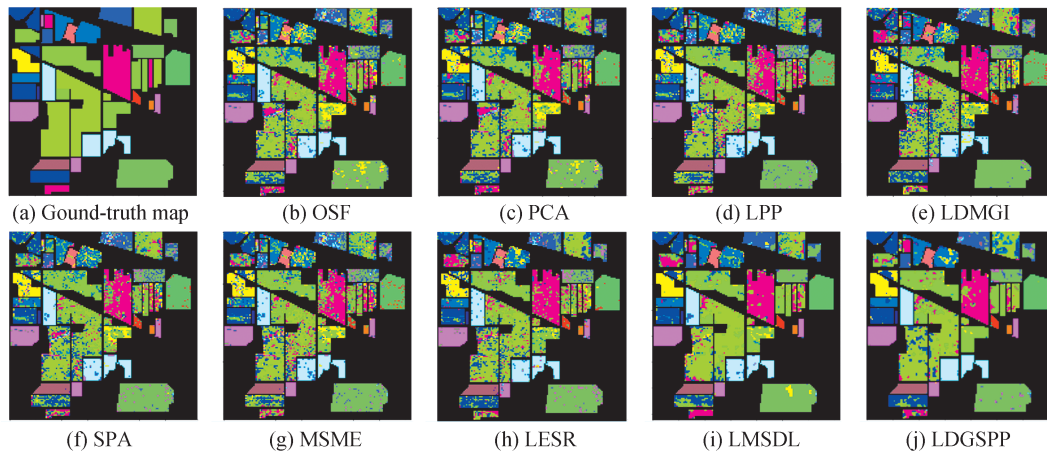


图6 在Indian Pines数据集上各算法的分类结果
Fig. 6 Classification results of Indian Pines dataset with different algorithms

2.3.3 执行时间

算法1用两重循环求解稀疏图B,外循环需要次数为M,内循环次数为N,内层循环中的第一步求解西尔维斯特方程的计算复杂度为 $O(n^3)$,其中n为样本个数,内层循环后两步可忽略不计,故算法1的总复杂度为 $O(MNn^3)$ 。本文所提LDGSP见算法2,需T次迭代,迭代第一步需 $O(MNn^3)$,第二步特征值分解花费 $O(d^3)$,故LDGSP所需总消耗为 $O(T(d^3 + MNn^3))$ 。

以上所有实验的硬件平台均为Intel(R) Core(TM) i5-6200U 2.30GHz,8GB随机存取存储器(RAM)。为了对比各算法的运行效率,表3中列出了在两个数据集上使用不同特征提取算法的运行时间。

对于不同数据集,由于IndianPines数据集用于实验的类别数和波段数远高于PaviaU数据集,故在实验的八种算法中均显示IndianPines训练所消耗的时间成本高于PaviaU。对于各特征提取算法,在两个数据集

上显示,LDGSP运行消耗的时间均高于对比算法,这是由于LDGSP使用递归策略求未知稀疏编码和低维空间映射矩阵。但事实证明,LDGSP算法效果极佳且所消耗的时间成本是可以接受的。

表3 各算法的运行时间(s)
Table 3 Running time of different algorithms (s)

Dataset	PCA	LPP	LDMGI	SPA	MSME	LESR	LMSDL	LDGSP
PaviaU	0.02	0.12	2.50	2.24	3.39	46.55	57.47	81.01
Indian Pines	0.03	0.23	2.92	9.62	10.72	101.11	143.38	168.23

2.4 参数敏感性分析

2.4.1 正则参数

分析不同参数设置下用SVM分类器得到的OA。图7(a)中显示随着 η 的变化,OA变化趋势不大;图7(b)中显示随着 v 值增加,OA逐渐增加,表明全局稀疏结构包含了有利于分类任务的判别信息;图7(c)中稀疏正则参数 λ 在经验范围内取值,能使得算法取得较好的效果。图7(d)中随着正则参数 ρ 增加,分类精度先不变,然后大幅减少,这是因为适度加入正则能提高模型性能,而过度正则会导致精度下降。PaviaU数据集最终选择的参数组合为($\eta=0.8, v=2.5, \lambda=0.001, \rho=0.1$); Indian Pines数据集选择的参数组合为($\eta=0.2, v=2, \lambda=0.001, \rho=0.01$)。实验结果显示参数设置与3.2节中通过交叉验证方法获取的最优参数设置保持一致。这表明了交叉验证方法能够有效地进行模型参数选择。

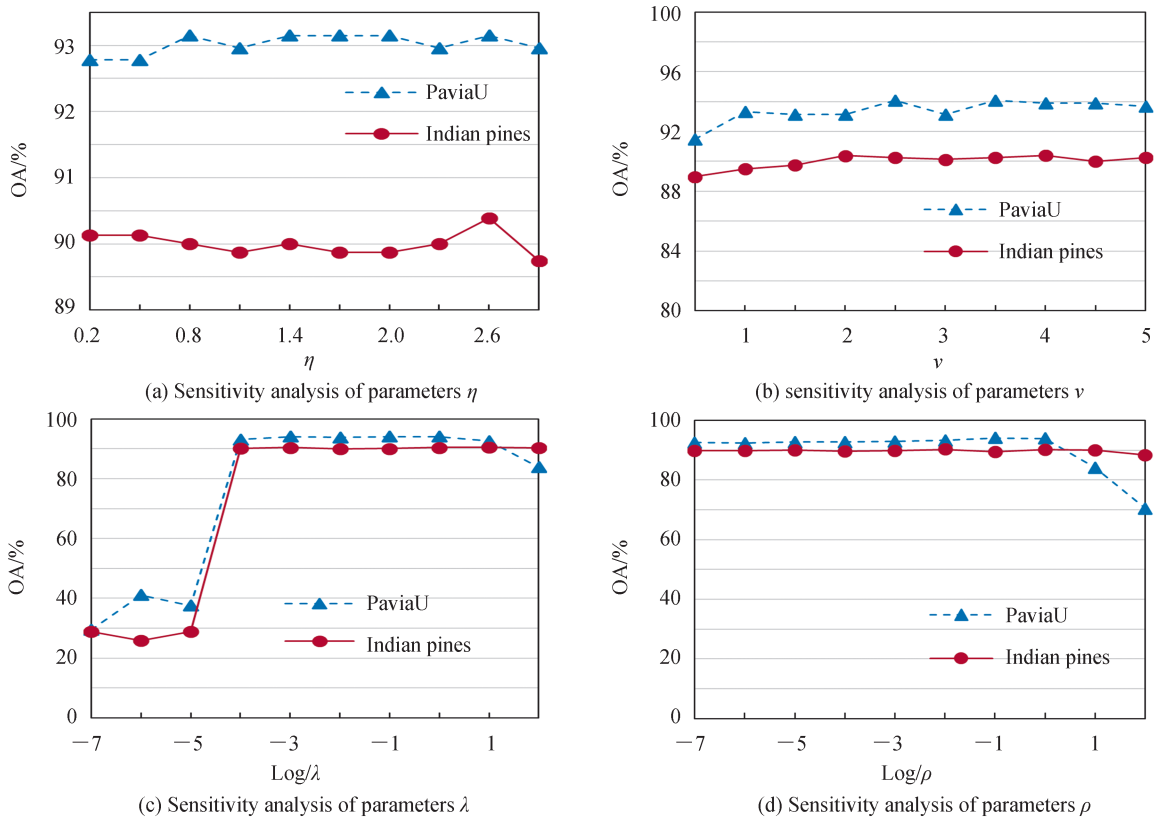


图7 正则参数 ρ 的敏感性分析

Fig. 7 Sensitivity analysis of regular parameters ρ

2.4.2 维度

合适的子空间应该拥有更多的判别能力和更少的冗余信息。实验中通过SVM分类器和OA来衡量所得子空间,同时对八种特征提取方法进行对比,如图8所示,在PaviaU和Indian Pines两个数据集上的实验显示,所有方法随着嵌入维度的增加,精度均快速增长,随后达到较稳定的数值;LDGSPP得到的分类精度远高于所有对比方法,这说明LDGSPP更加有效地揭示了数据间的内在结构,提取了更有利于分类的判别信息。

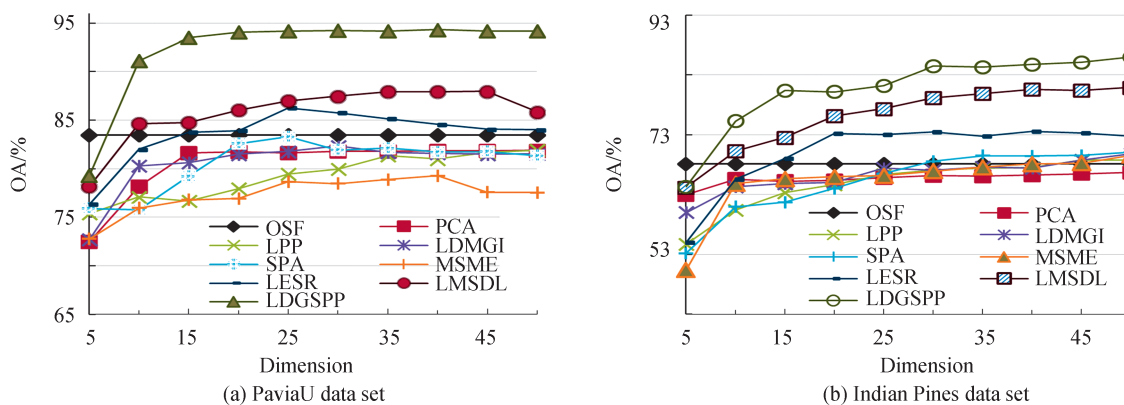


图8 子空间维度的敏感性分析
Fig. 8 Sensitivity analysis of subspace dimension

2.4.3 训练集大小

使用SVM分类器,从图9两幅图均可得出随着训练样本数的增加,所有方法的分类精度均在提高,这说明越多的训练样本包含着更多有助于分类的信息;同时也可看出本文方法在前期增长速度快,随着训练样本数目的增加增长速度变缓,且在每类样本数只有30的情况下在PaviaU数据集和Indian Pines数据集上分类精度分别达到了88.97%和77.34%,均高于其他几种对比方法,这说明本文的特征提取模型不依赖于更多的训练样本且所提取的特征极具判别性,有助于最终的分类性能提升。因此,所提的特征提取模型适用于样本不平衡或者样本数量少的数据集。

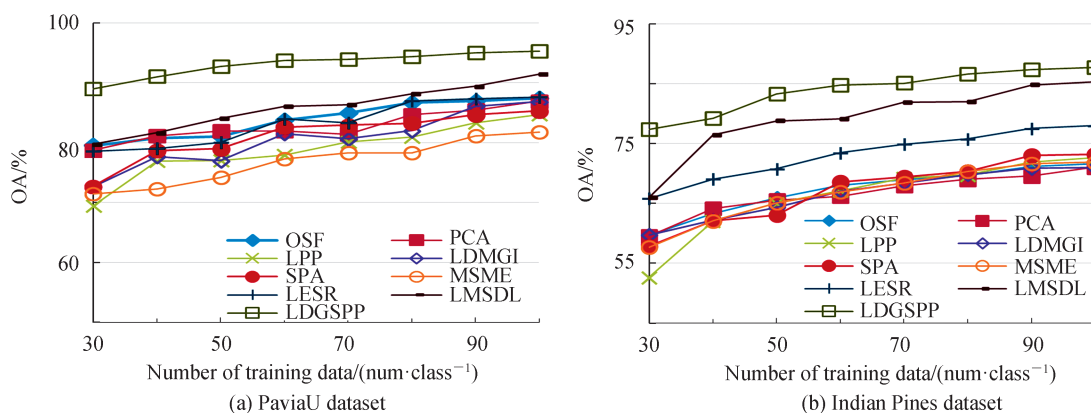


图9 训练集大小的敏感性分析
Fig. 9 Sensitivity analysis of training set size

2.4.4 迭代次数

在PaviaU、Indian Pines两个数据集上训练LDGSPP,分别迭代50次得到图10所示的损失曲线。从图中可以观察到本文的模型很快可以达到收敛,在两个数据集上不超过20次迭代均可达到收敛。

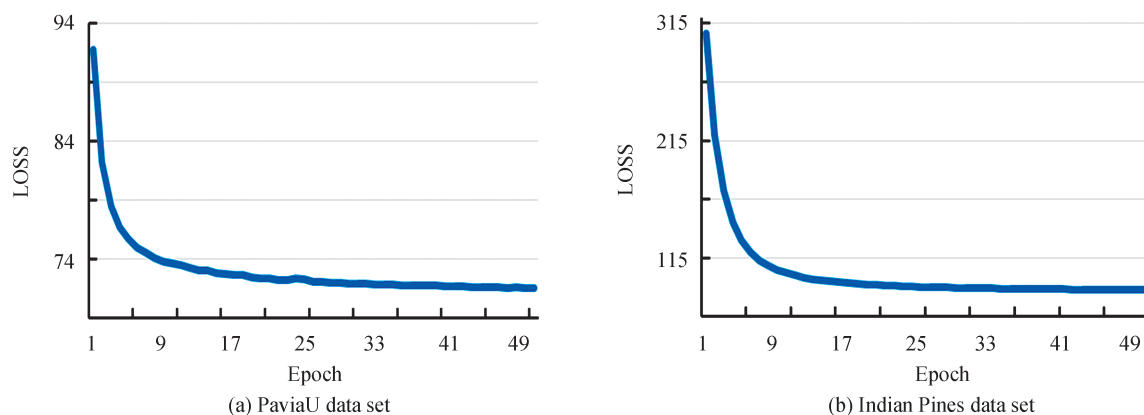


图10 收敛曲线
Fig. 10 Convergence curve

3 结论

本文提出一种基于稀疏表示和学习图正则的局部判别与全局稀疏保持投影算法。该方法基于LESR算法,提出施加学习图正则约束的稀疏表示模型,自适应获取局部邻域,利用LDMGI提取获取的局部邻域中的局部判别信息,并利用基于学习图正则的稀疏表示模型来提取数据的全局稀疏结构。该算法有效揭示了数据间的多种判别结构,从而提升了分类精度。实验结果显示在PaivaU和Indian Pines两个数据集上所得到的分类结果均优于所对比的方法。

参考文献

- [1] GHAMISI P, YOKOYA N, LI Jun, et al. Advances in hyperspectral image and signal processing: a comprehensive overview of the state of the art[J]. IEEE Geoscience and Remote Sensing Magazine, 2017, 5(4): 37-78.
- [2] ZHANG Chengye, QIN Qiming, CHEN Li, et al. Research and development of mineral identification utilizing hyperspectral remote sensing[J]. Optics and Precision Engineering, 2015, 23(8): 2407-2418.
张成业, 秦其明, 陈理, 等. 高光谱遥感岩矿识别的研究进展[J]. 光学精密工程, 2015, 23(8): 2407-2418.
- [3] TANG Zhongqi, FU Guangyuan, CHEN Jin, et al. Multiscale segmentation based sparse coding for hyperspectral image classification[J]. Optics and Precision Engineering, 2015, 23(9): 2708-2714.
唐中奇, 付光远, 陈进, 等. 基于多尺度分割的高光谱图像稀疏表示与分类[J]. 光学精密工程. 2015, 23(9): 2708-2714.
- [4] LI Wei, FENG Fubiao, LI Hengchao, et al. Discriminant analysis-based dimension reduction for hyperspectral image Classification: a survey of the most recent advances and an experimental comparison of different techniques [J]. IEEE Geoscience and Remote Sensing Magazine, 2018, 6(1): 15-34.
- [5] SUN Weiwei, YANG Gang, DU Bo, et al. A sparse and low-Rank near-Isometric linear embedding method for feature extraction in hyperspectral imagery classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(7): 4032-4046.
- [6] GAN Yuhang, LUO Fulin, LIU Juhua, et al. Feature extraction based multi-structure manifold embedding for hyperspectral remote sensing image classification[J]. IEEE Access, 2017, PP: 25069-25080.
- [7] CHENG Xuemei, CHEN Yuren, TAO Yang, et al. A novel integrated PCA and FLD method on hyperspectral image feature extraction for cucumber chilling damage inspection[J]. Transactions of the ASAE, 2004, 47(4): 1313-1320.
- [8] GUAN Lixin, XIE Weixin, PEI Jihong. Segmented minimum noise fraction transformation for efficient feature extraction of hyperspectral images[J]. Pattern Recognition, 2015, 48(10): 3216-3226.
- [9] HE Xiaofei, NIYOGI P. Locality preserving projection[M]. Neural Information Processing Systems, 2004: 153-160.
- [10] HE Xiaofei, CAI Deng, YAN Shuicheng, et al. Neighborhood preserving embedding[J]. IEEE International Conference on Computer Vision, 2005: 1208-1213.
- [11] DU Xingzhong, YAN Yan, PAN Pingbo, et al. Multiple graph unsupervised feature selection [J]. Signal Processing, 2016: 754-760.
- [12] YANG Yi, XU Dong, NIE Feiping, et al. Image clustering using local discriminant models and global integration[J]. IEEE Transactions on Image Processing, 2010, 19(10): 2761-2773.
- [13] QIAO Lishan, CHEN Songcan, TAN Xiaoyang. Sparsity preserving projections with applications to face recognition[J]. Pattern Recognition, 2010, 43(1): 331-341.

- [14] LUO Fulin, HUANG Hong, LIU Jiamin, 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.
- [15] TANG Yuanyan, YUAN Haoliang, LI Luoqing. Manifold-based sparse representation for hyperspectral image classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2014, 52(12):7606-7618.
- [16] DUAN Yule, HUANG Hong, LI Zhengying, et al. Local Manifold-based sparse discriminant learning for feature extraction of hyperspectral image[J]. *IEEE Transactions on Cybernetics*, 2020, PP:1-14.
- [17] CALVETTI D, REICHEL L. Application of ADI iterative methods to the restoration of noisy images[J]. *SIAM Journal on Matrix Analysis and Applications*, 1996, 17(11):165-186.
- [18] HUANG Hong, SHI Guangyao, HE Haibo, et al. Dimensionality reduction of hyperspectral imagery based on spatial-spectral manifold learning[J]. *IEEE Transactions on Cybernetics*, 2020, 50(6): 2604-2616.