

## 面向高光谱图像分类的半监督 Laplace 鉴别嵌入

李志敏<sup>①</sup> 张杰<sup>\*①</sup> 黄鸿<sup>①</sup> 马泽忠<sup>②</sup>

<sup>①</sup>(重庆大学光电技术与系统教育部重点实验室 重庆 400044)

<sup>②</sup>(重庆市国土资源和房屋勘测规划院 重庆 400020)

**摘要:** 为有效提取出高光谱遥感图像数据的鉴别特征, 该文阐述一种融合标记样本中鉴别信息和无标记样本中局部结构信息的半监督 Laplace 鉴别嵌入(SSLDE)算法。该算法利用标记样本的类别信息来保持样本集的可分性, 并通过构建标记样本和无标记样本的 Laplace 矩阵来发现样本集中局部流形结构, 实现半监督的流形鉴别。在 KSC 和 Urban 数据集上的实验结果说明: 该算法具有更高的分类精度, 可以有效地提取出鉴别特征信息。在总体分类精度上, 该算法比半监督最大边界准则(SSMMC)算法提升了 6.3%~7.4%, 比半监督流形保持嵌入(SSSMPE)算法提升了 1.6%~4.4%。

**关键词:** 图像处理; 高光谱遥感图像; 鉴别特征; Laplace 矩阵; 半监督 Laplace 鉴别嵌入

中图分类号: TP751.1

文献标识码: A

文章编号: 1009-5896(2015)04-0995-07

DOI: 10.11999/JEIT140600

## Semi-supervised Laplace Discriminant Embedding for Hyperspectral Image Classification

Li Zhi-min<sup>①</sup> Zhang Jie<sup>①</sup> Huang Hong<sup>①</sup> Ma Ze-zhong<sup>②</sup>

<sup>①</sup>(Key Laboratory of Optoelectronic Technique and Systems of the Ministry of Education, Chongqing University, Chongqing 400044, China)

<sup>②</sup>(Chongqing Institute of Surveying and Planning for Land Resources and Houses, Chongqing 400020, China)

**Abstract:** In order to extract effectively the discriminant characteristics of hyperspectral remote sensing image data, this paper presents a Semi-Supervised Laplace Discriminant Embedding (SSLDE) algorithm based on the discriminant information of labeled samples and the local structural information of unlabeled samples. The proposed algorithm makes use of the class information of labeled samples to maintain the separability of sample set, and discovers the local manifold structure in sample set by constructing Laplace matrix of labeled and unlabeled samples, which can achieve semi-supervised manifold discriminant. The experimental results on KSC and Urban database show that the algorithm has higher classification accuracy and can effectively extract the information of discriminant characteristics. In the overall classification accuracy, this algorithm is improved by 6.3%~7.4% compared with Semi-Supervised Maximum Margin Criterion (SSMMC) algorithm and increased by 1.6%~4.4% compared with Semi-Supervised Sub-Manifold Preserving Embedding (SSSMPE) algorithm.

**Key words:** Image processing; Hyperspectral remote sensing image; Discriminant characteristics; Laplace matrix; Semi-Supervised Laplace Discriminant Embedding (SSLDE)

### 1 引言

高光谱遥感作为一种高精尖端技术, 它的出现无疑是现代科技领域里的一次履新革面的发展。高光谱遥感图像具有丰富的空间地物和光谱信息, 可

应用于资源探测、海洋及大气监测、军事国防等诸多现代化领域<sup>[1,2]</sup>。然而, 由于高光谱遥感图像具有数据量偏大、维数过高、冗余性较强等特点, 致使高光谱图像的处理过程复杂化, 并且分类精度和分类效率明显降低, 严重制约着高光谱遥感技术的广泛应用<sup>[3,4]</sup>。因此, 如何提取出有用的鉴别特征信息来提升高光谱图像的分类性能已成为一个亟待深入探究的关键问题。

目前, 研究者们已经提出了一系列的有关高维数据降维及特征提取的理论和算法。如传统的主成分分析(Principal Component Analysis, PCA)<sup>[5]</sup>、线

2014-05-08 收到, 2014-07-04 改回

国家自然科学基金(61101168, 41371338), 中国博士后科学基金(2012M511906, 2013T60837), 重庆市基础与前沿研究计划项目(cstc2013jcyjA40005), 重庆市国土房管局科技计划项目(CQGT-KJ-2012028)和博士后科研计划项目(2012M511906, 2013T60837, XM2012001)资助课题

\*通信作者: 张杰 zhangjie\_fly@126.com

性判别分析(Linear Discriminant Analysis, LDA)<sup>[6]</sup>、极大边界准则(Maximum Margin Criterion, MMC)<sup>[7]</sup>等。这些算法具有坚实的理论基础,易于执行和分析,应用广泛。但它们均为全局假设的线性算法,不能发现隐藏在高光光谱遥感图像数据中的局部流形特征。近年来,学者们提出了各种基于邻域流形结构关系的非线性特征提取技术,如局部线性嵌入(Locally Linear Embedding, LLE)<sup>[8]</sup>、等距特征映射(ISOmetric feature MAPping, ISOMAP)<sup>[9]</sup>、拉普拉斯特征映射(Laplacian Eigenmap, LE)<sup>[10]</sup>等。这些算法无法解决对于新样本的学习问题。因而,文献[11,12]分别提出了局部保持投影(Locality Preserving Projection, LPP)算法和邻域保持嵌入(Neighborhood Preserving Embedding, NPE)算法,其非线性流形学习能力较好,但忽略了标记样本的类别信息,容易破坏数据集在特征空间中的可分性。

为了有效利用标记样本和未标记样本的鉴别特征信息以及节约人工标记样本所需的高额成本,研究者提出了半监督的学习方法。例如,文献[13]提出半监督稀疏鉴别嵌入(Semi-supervised Sparse Discriminant Embedding, SSDE)算法。文献[14,15]分别提出半监督最大边界准则(Semi-Supervised Maximum Margin Criterion, SSMMC)算法和半监督流形保持嵌入(Semi-Supervised Sub-Manifold Preserving Embedding, SSSMPE)算法,用于图像识别时取得了较好的识别效果,但存在参数的优化选择困难,识别效率较低,噪声敏感等问题,导致算法稳定可靠性较差。

由于传统的全局线性特征提取算法存在不能发现数据集的局部流形结构特征,以及单纯的局部流形学习方法忽略了标记样本的类别信息,容易破坏数据集的可分性问题。因而,本文结合上述各算法中的优缺点,提出了半监督 Laplace 鉴别嵌入(Semi-Supervised Laplace Discriminant Embedding, SSLDE)算法,该算法利用标记样本的类别信息来保持样本数据集在特征空间中的可分性,并通过构建标记样本和无标记样本的 Laplace 矩阵来发现样本数据集在特征空间中局部流形结构信息。特别地,该算法融入了无标记样本数据的局部结构信息,实现半监督的鉴别特征提取,进一步提升了高光光谱图像的分类精度,并且具有较好的稳定可靠性,以及参数优化选择简单等优点。

## 2 相关理论

### 2.1 特征提取

特征提取是指利用某种方法将隐藏在原始高维

数据中的特征信息提取出来,在低维嵌入空间中简洁、有效地表征出原始数据中的鉴别特征信息。针对高光光谱遥感图像数据的特征提取问题,就是寻找一个最佳鉴别的投影矩阵,将原始高维数据集映射到一个恰当的低维嵌入空间,此低维嵌入空间保留了高维数据集中的鉴别信息特征,如局部流形特征,从而在低维嵌入空间中进行高光光谱遥感数据的分类,可有效提升分类性能。具体描述:首先将高光光谱遥感图像数据表示成  $N \times b$  的数据集  $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ , 其中  $N$  为像素点数,  $b$  为波段数,然后根据相关理论寻找一个低维特征空间  $\mathbf{Y}$  来表征高维数据  $\mathbf{X}$  的有用鉴别特征信息。

### 2.2 谱图理论

谱图理论的基本思想是把数据空间中各个数据点之间的局部近邻结构关系看作是一个流形上的近邻关系图  $\mathbf{G}$ 。首先,将流形上的数据点用  $V$  表示,各个数据点之间的近邻关系用  $E$  表示,从而可将整个数据空间形式化为一个基于局部流形结构关系的近邻关系图  $\mathbf{G}=(V, E)$ 。其中,利用近邻数据点之间的某种距离或者相似性来权重近邻关系  $E$ ,用  $w_{ij}$  表示数据点  $\mathbf{x}_i$  与  $\mathbf{x}_j$  之间  $E$  的权重值。于是可以得到近邻关系图  $\mathbf{G}$  的权重矩阵  $\mathbf{W} = [w_{ij}]_{N \times N}$ , 其表达式为

$$\mathbf{W}_{ij} = \begin{cases} w_{ij}, & \mathbf{x}_i \in k(\mathbf{x}_j) \text{ 或 } \mathbf{x}_j \in k(\mathbf{x}_i) \\ 0, & \text{其它} \end{cases} \quad (1)$$

其中,  $k(\mathbf{x}_i)$  表示数据点  $\mathbf{x}_i$  的  $k$  个最近邻数据点。Laplace 矩阵  $\mathbf{L} = \mathbf{D} - \mathbf{W}$ ,  $\mathbf{D}$  为对角矩阵,且  $D_{ii} = \sum_{j=1}^N \mathbf{W}_{ij}$ 。Laplace 矩阵  $\mathbf{L}$  包含了数据点的局部近邻信息,其特征向量与流形结构上 Laplace-Beltrami 算子的特征函数可看作为是一种近似离散关系。因而, Laplace 矩阵  $\mathbf{L}$  应用在高维数据的特征提取时,能够使得高维数据空间中流形上近邻数据点在低维空间中也可保持局部近邻结构关系不变。

## 3 半监督 Laplace 鉴别嵌入算法

### 3.1 SSLDE 算法原理

基于保护样本集的可分性原则及谱图理论的邻域流形学习思想,本文提出一种融合标记样本中鉴别信息和无标记样本中局部结构信息的半监督 Laplace 鉴别嵌入(SSLDE)算法。该算法利用标记样本的类别信息来保持样本集在特征空间中的可分性,并通过构建标记样本和无标记样本的 Laplace 矩阵来发现数据集在特征空间中的局部流形结构信息,实现半监督的流形鉴别特征提取。该算法不需要调整过多的参数就可获得更加理想的分类效果,且具有较高的稳定可靠性。

(1)利用标记样本的类别信息,构建同类数据的类内散度矩阵  $S_w$  和异类数据的类间散度矩阵  $S_b$ , 寻求一个最佳鉴别矢量的目标路径,使异类数据的类间散度矩阵极大化、同类数据的类内散度矩阵极小化来保持样本集在特征空间中的可分性。其中,  $S_w$  和  $S_b$  的表达式分别为

$$S_w = \sum_{i=1}^c \sum_{j=1}^{n_i} (\mathbf{x}_j^i - \mathbf{m}_i)(\mathbf{x}_j^i - \mathbf{m}_i)^T \quad (2)$$

$$S_b = \sum_{i=1}^c n_i (\mathbf{m}_i - \mathbf{m})(\mathbf{m}_i - \mathbf{m})^T \quad (3)$$

其中,  $c$  是类别数,  $n_i$  表示第  $i$  类样本数,  $\mathbf{x}_j^i$  表示第  $i$  类的第  $j$  个样本,  $\mathbf{m}_i$  是第  $i$  类样本的均值向量,  $\mathbf{m}$  是总样本集的均值向量。

(2)利用标记样本和无标记样本的Laplace矩阵保持高维数据集中的局部流形结构信息,使在原始高维空间中流形上相距较远的样本数据点映射到低维特征空间中仍保持较远距离,流形上相距较近的样本数据点映射到低维特征空间中仍保持较近距离,从而保持高光谱遥感图像数据集的局部流形结构信息不变。

步骤1 利用式(1)描述的近邻法来构建标记样本和无标记样本的近邻关系图  $G$ , 并设置权值矩阵  $W = [w_{ij}]_{N \times N}$  的表达式为

$$W_{ij} = \begin{cases} e^{-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|}{\sigma^2}}, & \mathbf{x}_i \in k(\mathbf{x}_j) \text{ 或 } \mathbf{x}_j \in k(\mathbf{x}_i) \\ 0, & \text{其它} \end{cases} \quad (4)$$

步骤2 计算局部流形结构上的Laplace矩阵  $L = D - W$ ;

步骤3 利用步骤2中所得的Laplace矩阵  $L$ , 计算  $d$  维嵌入空间  $Y$ 。为保持高光谱遥感图像数据集从原始高维数据空间映射到低维嵌入空间后邻域流形特征不变,构造变换函数为

$$\begin{aligned} \min & \frac{1}{2} \sum_{i,j=1}^N (\mathbf{y}_i - \mathbf{y}_j)^2 w_{ij} \\ & = \frac{1}{2} \sum_{i,j=1}^N (\mathbf{y}_i w_{ij} \mathbf{y}_i^T + \mathbf{y}_j w_{ij} \mathbf{y}_j^T - \mathbf{y}_i w_{ij} \mathbf{y}_j^T - \mathbf{y}_j w_{ij} \mathbf{y}_i^T) \\ & = Y D Y^T - Y W Y^T \\ & = \text{tr}(Y L Y^T) \\ & = \text{tr}(A^T X L X^T A) \end{aligned} \quad (5)$$

其中  $A$  为投影矩阵,  $A^T X = Y$ , 且  $A^T A = I$ 。

SSLDE算法思想是在保持数据集在特征空间中可分性和局部流形特征的前提下将高维数据集  $X$  投影到低维嵌入空间  $Y$ , 因此可以通过优化方法求得最佳投影矩阵  $A$  以及低维嵌入空间  $Y$ 。

$$\left. \begin{aligned} \max & \text{tr}(A^T (S_b - S_w) A) \\ \min & \text{tr}(A^T X L X^T A) \end{aligned} \right\} \quad (6)$$

由式(6)可定义SSLDE算法的目标函数为

$$J(A) = \max \text{tr} \frac{A^T (S_b - S_w) A}{A^T X L X^T A} \quad (7)$$

采用熟悉的Lagrange乘法,可以将式(7)的最优化问题转换为特征方程的广义特征值求解:

$$(S_b - S_w) A = \lambda X L X^T A \quad (8)$$

解得式(8)中最大的  $d$  个特征值  $\lambda_1, \lambda_2, \dots, \lambda_d$  和与之相应的  $d$  个特征向量  $[a_1, a_2, \dots, a_d]$ , 即可求得从高维空间投影到低维嵌入空间的投影矩阵  $A = [a_1, a_2, \dots, a_d]$ 。

### 3.2 SSLDE 算法步骤

本文所提出的SSLDE算法的具体步骤为:

步骤1 对训练样本数据集进行主成分变换并求取变换矩阵  $A_{\text{PCA}}$  及低维数据  $Y_{\text{PCA}} = A_{\text{PCA}}^T X$ ;

步骤2 利用步骤1所得数据集  $Y_{\text{PCA}}$  中标记样本数据点求得类内散度矩阵  $S_w$  和类间散度矩阵  $S_b$ ;

步骤3 构建  $Y_{\text{PCA}}$  的近邻关系图  $G$ , 利用式(4)计算权值矩阵  $W = [w_{ij}]_{N \times N}$ , 求Laplace矩阵  $L = D - W$ ;

步骤4 计算式(8)的  $d$  个最大特征值及其相应的  $d$  个特征向量  $[a_1, a_2, \dots, a_d]$ , 且  $A_Z = [a_1, a_2, \dots, a_d]$ ;

步骤5 计算最佳投影矩阵  $A = A_{\text{PCA}} A_Z$ , 即可获得低维嵌入空间  $Y = A^T X$ 。

## 4 实验与分析

为了评估本文所提出的 SSLDE 算法在高光谱遥感图像中的分类性能,分别对具有代表性的 KSC 和 Urban 高光谱遥感数据集进行分类实验。首先利用各种算法对同一训练样本集计算得出在特征空间中各自相应的投影矩阵,然后利用所得的投影矩阵对相同的测试样本集进行投影降维,得到测试样本集在低维嵌入空间中的特征表示,最后用 K 最邻近(K-Nearest Neighbor, K-NN)结点算法进行遥感数据的分类处理。在满足各种算法参数选取最优的条件下,将 PCA, LDA, LPP, NPE, MMC, SSMC 以及 SSMPE 与 SSLDE 算法的分类结果进行对比分析。

### 4.1 KSC 数据集

KSC 数据集是位于美国 Florida 的肯尼迪航天中心(Kennedy Space Center, KSC)的高光谱遥感图像,如图 1 所示。已知该遥感图像的地面空间分辨

率为 18 m, 共有 614×512 个像素点, 光谱范围为 400~2500 nm, 波段数为 224, 经过预处理后将剩余的 176 个波段进行分类实验研究。该数据集中含有 13 类已知地物, 其中样本点共有 5211 个, 本实验选取其中含地物数据样本点最多的 7 个类别数据进行分类实验, 其类别信息如表 1 所示。

表 1 KSC 数据集的类别信息

类别	类名	数目
C1	水(Water)	927
C2	灌丛(Scrub)	761
C3	米草沼泽(Spartina marsh)	520
C4	泥滩(Mud flats)	503
C5	禾草沼泽(Graminoid marsh)	431
C6	盐沼(Salt marsh)	419
C7	香蒲沼泽(Cattail marsh)	404

本实验分别从表 1 中的每类已知地物中随机选取 10, 20, 30 个标记样本和 30 个无标记样本作为训练样本, 100 个样本点作为测试样本。由于 KSC 数据集中每类地物数据量大, 而实验选取的测试样本数目相对较少, 为了提升实验的精确性及可信度,

每种算法重复进行 10 次分类实验, 然后对 10 次实验的分类精度求平均值得到最终的总体分类精度。由于嵌入在高维特征空间中流形结构的本征维数是未知、不可确定的, 因而本实验在不同的低维嵌入空间维数下进行了分类实验, 同时为了对比分析不同算法在不同训练样本数目情况下的分类性能, 图 2 给出了不同算法在不同训练样本数、不同维数下进行分类实验的总体分类精度, 表 2 给出了不同算法在不同训练样本数情况下的最高总体分类精度值及其标准偏差和所在低维嵌入空间的维数。



图 1 KSC 高光谱遥感图像

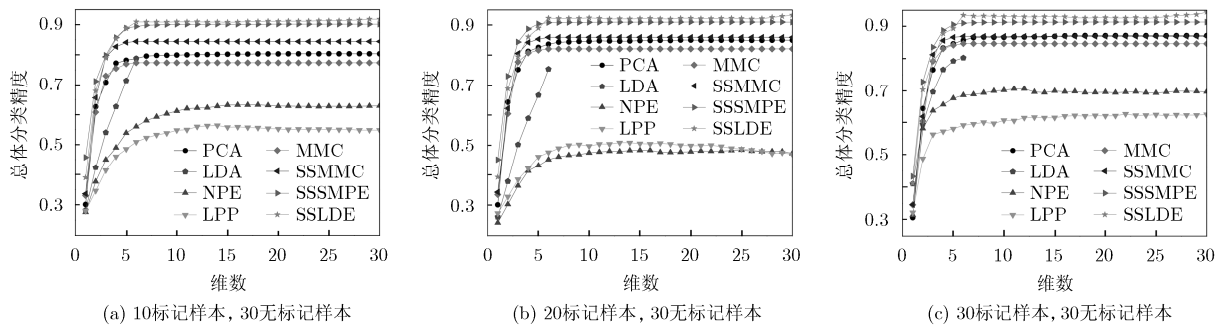


图 2 不同算法在不同训练样本数、不同维数下的总体分类精度

表 2 不同算法的最高总体分类精度(平均值±标准偏差(%)(维数))

算法	10 标记样本 30 无标记样本	20 标记样本 30 无标记样本	30 标记样本 30 无标记样本
PCA	80.40±0.17(28)	85.01±0.21(29)	87.20±0.15(27)
LDA	78.83±0.39(6)	75.38±0.55(6)	80.29±0.49(6)
NPE	63.26±0.47(18)	48.35±0.54(16)	70.59±0.31(12)
LPP	56.38±0.35(14)	50.58±0.46(13)	62.55±0.37(22)
MMC	77.39±0.29(24)	82.11±0.14(22)	84.68±0.18(20)
SSMMC	84.47±0.15(26)	86.94±0.17(23)	86.96±0.15(24)
SSSMPE	90.18±0.21(20)	90.98±0.20(19)	91.34±0.25(16)
SSLDE	<b>91.84±0.11(29)</b>	<b>93.26±0.10(30)</b>	<b>94.03±0.10(30)</b>

由图 2 和表 2 可知,本文 SSLDE 算法在分类性能上较优于其它几种特征提取算法。其中 PCA, MMC, SSMC, SSSMPE 以及 SSLDE 算法的分类精度随训练样本数的增加而增加; NPE, LPP 算法是基于局部结构关系的流形学习方法,其分类精度主要受邻域结构关系和最近邻点数  $k$  的取值影响; LDA 在线性二分类问题中应用效果较好,但高光谱遥感数据属于多类非线性问题。

PCA 算法是在寻求全局最佳逼近的情况下提取特征信息,对于全局结构为线性属性的数据集能够很好地去其冗余性,有效保留数据集中的主成分特征信息,但高光谱遥感数据是存在非线性属性的,因而 PCA 算法并不能较好地原始遥感数据中提取出鉴别特征信息,导致总体分类精度并不理想; LDA 和 MMC 算法利用了标记样本的类别信息,保持了样本集在特征空间中的可分性,但忽略了高光谱遥感数据集的局部流形结构特征,因而通过 LDA 和 MMC 算法对高光谱遥感图像数据进行特征提取后的分类效果并不突出; NPE 和 LPP 算法仅保持数据集在特征空间中的邻域流形结构信息,却没有考虑到标记样本中最重要的类别信息,破坏了样本集的可分性,导致算法的鉴别性差,总体分类精度低; SSMC 和 SSSMPE 算法的参数优化复杂,分类性能受到影响;本文 SSLDE 算法既保持了样本集的局部流形结构特征,也没有破坏样本集在特征空间中的可分性,分类精度最高,且该算法参数优化选择简单,分类效果稳定可靠。

由表 2 可得,在 10, 20, 30 个标记训练样本和 30 个无标记训练样本的情况下,本文 SSLDE 算法的最高总体分类精度比 SSMC 分别提升了约 7.37%, 6.32%, 7.07%; 相对于 SSSMPE 分别提升了约 1.66%, 2.28%, 2.69%。

### 4.2 Urban 数据集

Urban 数据集是位于某城郊住宿区的高光谱图遥感图像,如图 3 所示。已知该遥感图像的地面空间分辨率为 3 m,共有  $307 \times 307$  个像素点,波谱范围为 400~2500 nm,波段数为 210,经过预处理后将剩下的 162 个波段进行分类实验研究。该遥感数据集中有道路、建筑、树木和草地 4 类已知的地物种类,其中阴影部分为未知地物。

本实验分别从每类已知地物中随机选取 10, 20, 30 个标记样本点和 30 个无标记样本点作为训练样本,800 个样本点作为测试样本进行实验。由于 Urban 数据集中每类地物数据量大,而实验选取的测试样本数目相对较少,为了提升实验的精确性及可信度,每种算法重复进行 10 次分类实验,然后对 10 次实验的分类精度求平均值得到最终的总体分类精度。由于嵌入在高维特征空间中流形结构的本征维数是未知、不可确定的,因而本实验在不同的低维嵌入空间维数下进行了分类实验,同时为了对比分析不同算法在不同训练样本数目情况下的分类性能,图 4 给出了不同算法在不同训练样本数、不同维数下进行分类实验的总体分类精度,表 3 给出了不同算法在不同训练样本数情况下的最高总体分类精度值及其标准偏差和所在低维嵌入空间的维数。

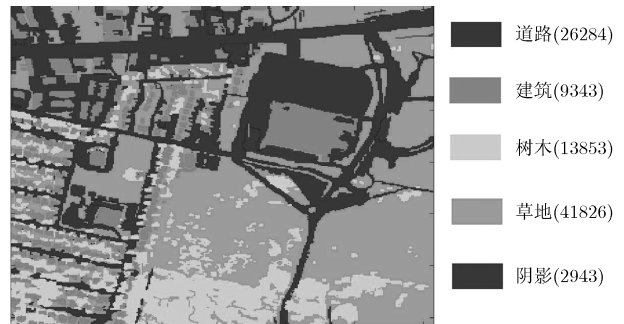


图 3 Urban 高光谱遥感图像

表 3 不同算法的最高总体分类精度(平均值±标准偏差(%)(维数))

算法	10 标记样本 30 无标记样本	20 标记样本 30 无标记样本	30 标记样本 30 无标记样本
PCA	70.29±0.29(14)	72.84±0.26(18)	74.75±0.16(30)
LDA	70.73±0.39(3)	66.91±0.30(3)	57.95±0.42(3)
NPE	69.46±0.51(30)	68.96±0.32(30)	63.07±0.34(15)
LPP	70.02±0.40(30)	66.17±0.32(28)	61.60±0.25(21)
MMC	68.65±0.44(15)	69.74±0.35(10)	71.80±0.29(16)
SSMMC	71.02±0.27(14)	71.59±0.21(10)	72.54±0.25(25)
SSSMPE	73.42±0.19(19)	74.89±0.19(25)	75.64±0.21(27)
SSLDE	<b>77.82±0.13(29)</b>	<b>78.74±0.14(28)</b>	<b>78.88±0.12(29)</b>

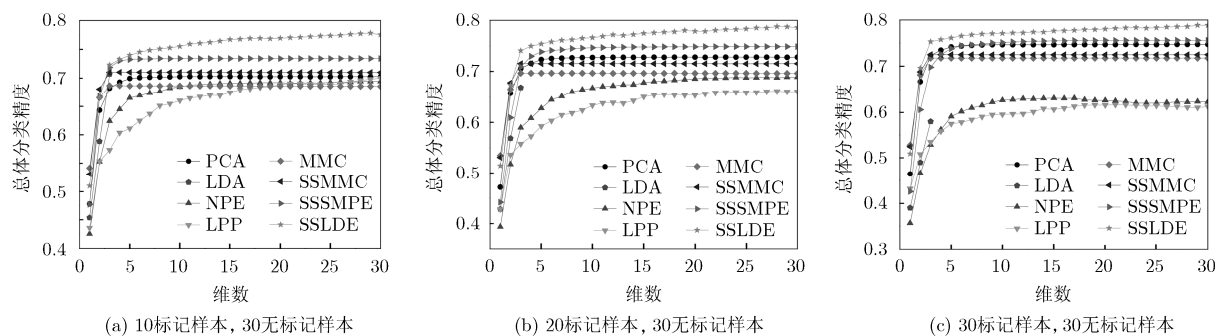


图 4 不同算法在不同训练样本数、不同维数下的总体分类精度

由图 4 可得, 针对 Urban 高光谱遥感数据集进行特征提取及分类时, 本文 SSLDE 算法的分类精度明显高于其它几种特征提取算法。因为 SSLDE 算法不仅利用标记样本的类别信息保持样本集在特征空间中的可分性, 还利用标记样本和无标记样本的 Laplace 矩阵保持样本集在特征空间中的局部流形结构信息, 有效提取出高光谱遥感图像的鉴别特征信息, 具有更高的分类精度。

PCA 算法降低了高光谱遥感数据的冗余性, 在一定程度上提升了分类精度; LDA 和 MMC 算法虽利用了标记样本的类别信息, 但忽略数据集中邻域流形特征, 所以分类精度不够突出; NPE 和 LPP 算法仅仅保持了样本集的邻域流形结构关系, 忽略了标记样本中最重要的类别信息, 破坏了高光谱遥感数据集在特征空间中的可分性, 结果导致算法分类精度较低; SSMMC 和 SSSMPE 算法为半监督学习算法, 利用了无标记样本的流形特征信息, 在总体分类精度上有一定的提高。本文 SSLDE 算法同样也作为一种半监督的特征提取算法, 可是相对于 SSMMC 和 SSSMPE 算法的参数优化选择较为困难及分类性能易受所选参数影响的缺点, SSLDE 算法不仅无复杂的参数优化过程, 并且分类精度更高。

从表 3 可以得出, 在 10, 20, 30 个标记训练样本和 30 个无标记训练样本的情况下, 本文 SSLDE 算法的最高总体分类精度比 SSMMC 分别提升了约 6.80%, 7.15%, 6.34%; 相对于 SSSMPE 分别提升了约 4.40%, 3.85%, 3.24%。同时, SSLDE 算法的标准偏差值也最小, 可说明该算法的稳定性更好。

### 4.3 算法复杂度比较

为了比较算法的复杂度与实时性, 通过记录各种算法的运行时间来反映算法复杂度。本实验是在 Intel(R) Core(TM)2 T6570 2.10 Hz 的 CPU 和 2.00 GB 的 RAM 平台上进行的。在低维嵌入空间维数  $d=30$  (LDA 例外) 的情况下, 各种算法分别重复进行 10 次实验, 记录每次实验的运行时间, 再求得平均值得到最终结果如表 4 所示。

由表 4 可得 SSMMC, SSSMPE, SSLDE 算法的运行时间明显比其他算法更长, 这是因为半监督算法考虑了大量无标记样本的特征信息, 所以算法复杂度更高, 运行时间更长。在实际应用中, 牺牲一定的实时性来提高分类精度也是可取的。

## 5 结束语

本文通过对比分析高光谱遥感图像处理中常用的降维及特征提取算法, 提出一种融合标记样本中鉴别信息和无标记样本中局部结构信息的半监督 Laplace 鉴别嵌入 (SSLDE) 算法。该算法充分利用标记样本的类别信息和样本集中局部流形结构信息将高维数据投影到一个恰当的低维特征空间中, 简洁、有效地刻画出数据集的鉴别特征信息。特别地, 利用无标记样本的 Laplace 矩阵来引入半监督的流形鉴别嵌入方法, 进一步提升了高光谱图像数据的分类性能。

在 KSC 和 Urban 数据集上的实验结果证明: 本文 SSLDE 算法比其它常见的方法具有更高的分类精度, 可以有效提取出高光谱遥感图像数据的鉴别特征信息。在总体分类精度上, 本文算法比

表 4 各种算法的运行时间(s)

算法	PCA	LDA	NPE	LPP	MMC	SSMMC	SSSMPE	SSLDE
KSC数据集运行时间	1.2736	1.0912	1.2845	1.3071	1.3154	2.5603	2.5073	2.5521
Urban数据集运行时间	0.5614	0.4540	0.5751	0.5528	0.5604	1.1306	1.1264	1.1362

SSMMC 提升了约 6.3%~7.4%, 比 SSSMPE 提升了 1.6%~4.4%。该算法具有参数优化选择简单, 分类精度较高, 稳定可靠性好等优点。但本文的研究工作还存在某些不足或者有待深入的地方, 比如如何进一步提高运算效率, 如何结合遥感图像中的空间纹理结构信息来提高分类性能, 以及进一步研究多流形或者从流形结构的特征提取问题。

### 参考文献

- [1] 张达, 郑玉权. 高光谱遥感的发展与应用[J]. 光学与光电技术, 2013, 11(3): 67-73.  
Zhang Da and Zheng Yu-quan. The development and application of hyperspectral remote sensing[J]. *Optical and Photoelectric Technology*, 2013, 11(3): 67-73.
  - [2] 宋相法, 焦李成. 基于稀疏表示及光谱信息的高光谱遥感图像分类[J]. 电子与信息学报, 2012, 34(2): 268-272.  
Song Xiang-fa and Jiao Li-cheng. Hyperspectral remote sensing image classification based on the sparse representation and spectral information[J]. *Journal of Electronics & Information Technology*, 2012, 34(2): 268-272.
  - [3] 陈宏达, 普晗晔, 王斌, 等. 基于图像欧氏距离的高光谱图像流形降维算法[J]. 红外与毫米波学报, 2013, 32(5): 450-455.  
Chen Hong-da, Pu Han-ye, Wang Bin, et al. Based on Euclidean distance of hyperspectral image manifold dimensionality reduction algorithm[J]. *Journal of Infrared and Millimeter Wave*, 2013, 32(5): 450-455.
  - [4] 杜博, 张乐飞, 张良培, 等. 高光谱图像降维的判别流形学习方法[J]. 光子学报, 2013, 42(3): 320-325.  
Du Bo, Zhang Le-fei, Zhang Liang-pei, et al. Hyperspectral image discriminant manifold learning method of dimension reduction[J]. *Acta Photonica Sinica*, 2013, 42(3): 320-325.
  - [5] Abdi H and Williams L J. Principal component analysis[J]. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2010, 2(4): 433-459.
  - [6] Moulin C, Largeton C, Ducottet C, et al. Fisher linear discriminant analysis for text-image combination in multimedia information retrieval[J]. *Pattern Recognition*, 2014, 47(1): 260-269.
  - [7] Wan M, Lai Z, and Jin Z. Feature extraction using two-dimensional local graph embedding based on maximum margin criterion[J]. *Applied Mathematics and Computation*, 2011, 217(23): 9659-9668.
  - [8] Zhao X and Zhang S. Facial expression recognition using local binary patterns and discriminant kernel locally linear embedding[J]. *EURASIP Journal on Advances in Signal Processing*, 2012, 2012(1): 1-9.
  - [9] Rosman G, Bronstein M M, Bronstein A M, et al. Nonlinear dimensionality reduction by topologically constrained isometric embedding[J]. *International Journal of Computer Vision*, 2010, 89(1): 56-68.
  - [10] Luo W. Face recognition based on Laplacian Eigenmaps[C]. Proceedings of the 2011 IEEE International Conference on Computer Science and Service System (CSSS), Nanjing, China, 2011: 416-419.
  - [11] He X and Niyogi P. Locality preserving projections[C]. Proceedings of the Advances in Neural Information Processing Systems, Vancouver, BC, Canada, 2004: 234-241.
  - [12] He X, Cai D, Yan S, et al. Neighborhood preserving embedding[C]//Proceedings of the 2005 IEEE International Conference on Computer Vision (ICCV), Beijing, China, 2005: 1208-1213.
  - [13] 黄鸿, 曲焕鹏. 基于半监督稀疏鉴别嵌入的高光谱遥感影像分类[J]. 光学精密工程, 2014, 22(2): 434-442.  
Huang Hong and Qu Huan-peng. Hyperspectral remote sensing image classification based on SSDE[J]. *Optics and Precision Engineering*, 2014, 22(2): 434-442.
  - [14] Song Y Q, Nie F P, Zhang C S, et al. A unified framework for semi-supervised dimensionality reduction[J]. *Pattern Recognition*, 2008, 41(9): 2789-2799.
  - [15] Song Y Q, Nie F P, and Zhang C S. Semi-supervised sub-manifold discriminant analysis[J]. *Pattern Recognition Letters*, 2008, 29(13): 1806-1813.
- 李志敏: 男, 1955年生, 副教授, 硕士生导师, 研究方向为数字图像处理、计算机网络、嵌入式系统。  
张杰: 男, 1988年生, 硕士生, 研究方向为数字图像处理、遥感图像分类、嵌入式系统等。  
黄鸿: 男, 1980年生, 副教授, 硕士生导师, 研究方向为流形学习、模式识别、遥感图像处理等。  
马泽忠: 男, 1972年生, 教授级高级工程师, 硕士生导师, 研究方向为空间数据信息处理与系统开发、虚拟现实与三维空间数据制作体系理论、定量。