

文章编号 1004-924X(2014)02-0434-09

基于半监督稀疏鉴别嵌入的高光谱遥感影像分类

黄 鸿^{1,2*}, 曲焕鹏¹

- (1. 重庆大学 光电技术及系统教育部重点实验室, 重庆 400044;
2. 重庆川仪自动化股份有限公司技术中心, 重庆 401121)

摘要: 为了有效利用已标记与未标记样本提高高光谱遥感影像分类精度, 提出一种新的半监督流形学习方法——半监督稀疏鉴别嵌入算法(SSDE)。该算法结合了近邻流形结构及稀疏性的优点, 不仅保留样本间的稀疏重构关系, 而且通过引入少量有标记的训练样本以及大量无标记训练样本来获得高维数据的内在属性以及低维流形结构, 实现鉴别特征提取, 提高分类精度。在 Washington DC Mall 和 Indian Pine 数据集上的分类识别实验表明, 该算法能够较为有效地发现高维空间中数据的内蕴结构, 分类性能比其他算法有明显的提升。在随机选取 8 个有类别标记和 60 个无类别标记的数据作为训练样本的情况下, 本文提出的 SSDE 算法在上述两个数据集上的分类精度分别达到了 77.36% 和 97.85%。

关键词: 高光谱遥感影像; 影像分类; 维数约简; 稀疏表示; 流形学习; 半监督稀疏鉴别嵌入

中图分类号: TP751 **文献标识码:** A **doi:** 10.3788/OPE.20142202.0434

Hyperspectral remote sensing image classification based on SSDE

HUANG Hong^{1,2*}, QU Huan-peng¹

(1. Key Laboratory of Opto-electronic Technique of the Ministry of Education, Chongqing University, Chongqing 400044, China;

2. The Technical Center of Chongqing Chuanyi Automation Co., Ltd., Chongqing 401121, China)

* Corresponding author, E-mail: hhuang.cqu. @gmail.com

Abstract: To improve the classification accuracy of hyperspectral remote sensing images by utilizing labeled and unlabeled samples, a new semi-supervised manifold learning method called Semi-supervised Sparse Discriminant Embedding (SSDE) is proposed. By combining the advantages of manifold structure among classes and sparsity, the algorithm not only preserves the sparse reconstruction relationship between the samples, but also gets the intrinsic attribute of high dimensional data and the manifold structure of low dimensional data by introducing a few labeled training samples and a large number of unlabeled training samples. So, it extracts the discriminant feature of data and improves the classification accuracy. The classification experiments in Washington DC Mall and Indian Pine data set show that the method is a more effective way to find the internal structure of data in a high dimensional space. Compared to other methods, the SSDE obviously improves the classification perform-

收稿日期: 2013-06-21; **修订日期:** 2013-08-15.

基金项目: 国家自然科学基金资助项目 (No. 61101168, No. 41371338); 中国博士后科学基金资助项目 (No. 2012M511906, No. 2013T60837); 重庆市基础与前沿研究计划资助项目 (No. cstc2013jcyjA40005); 重庆市博士后科研资金特别资助项目 (No. XM2012001); 中央高校基本科研业务费专项资金资助项目 (No. 1061120131204, No. 106112013CDJZR125501)

ance. By taking randomly selected 8 training samples with classification labels and 60 ones without classification labels as examples, the highest classification precision of SSDE respectively reach 77.36% in Indian Pine and 97.85% in Washington DC Mall data set.

Key words: hyperspectral remote sensing image; image classification; dimensionality reduction; sparsity representation; manifold learning; semi-supervised sparsity discriminant embedding

1 引 言

高光谱遥感技术自 20 世纪 80 年代以来发展迅速,其影像记录了地物目标的连续光谱,包含的信息更丰富。由于具有更高精度的目标分类能力可以识别更多种类的地物目标^[1-2]。但由于高光谱数据是由大量波段构成的高维特征空间,而大多数算法的复杂度随维数呈指数关系增长,因此处理这类数据,计算量更大。另外,其波段间具有高度相关性和冗余性,同时存在维数很高,分类时易因 Hughes 现象^[3]而无法获得理想结果等问题。科学家通过研究发现:高光谱数据可被描述为在低维嵌入空间上的流形(Manifold)^[4],即:高维空间的点是在少数独立变量的共同作用下在观测空间张成的一个流形,如果能有效地发现其内在的主要变量,便能更好地了解高维数据的本质属性和特点。维数约减是解决以上问题的有效办法,通过维数约减能够降低数据的维数,得到高维数据有意义的低维表示,以便理解其内在的结构及后续处理。

迄今为止,在高维数据处理的研究领域中,国内外学者已经提出了一系列经典算法,其中,应用较广泛的降维算法主要包括:主成分分析(Principal Component Analysis, PCA)^[5]和线性判别分析(Linear Discriminant Analysis, LDA)^[6]等子空间学习方法,而 PCA 和 LDA 均基于高维数据空间的嵌入子空间是线性的这一假设,这样,隐藏在低维数据中的内在属性较难被发现,因此无法揭示高光谱数据的低维流形结构。局部保持投影(Local Preserving Projection, LPP)^[7]及邻域保持嵌入算法(Neighborhood Preserving Embedding, NPE)^[8]等局部方法通过近邻图保留样本的局部近邻结构,在一定程度上保持了原始数据的非线性流形,但是这两种算法依赖于人工预定义的近邻图,要获得好的性能往往需要更多的训练

样本,存在如近邻参数选择困难(如近邻数 k 、核宽 σ)、噪声敏感等问题,分类效果受到限制。

稀疏保持投影(Sparsity Preserving Projections, SPP)^[9]是最近提出的一种基于稀疏表示理论的无监督降维算法。该算法不同于传统方法的图构建方式(如 k 近邻),其利用样本间的稀疏重构关系建图,是全局性的稀疏构图算法,并归结为 l_1 -范数最小化问题。SPP 算法不仅利用稀疏表示的自然判别能力,而且无需人为选择近邻参数值,一定程度上缓解了近邻参数选择的困难。

在高光谱影像分类中,面对的常是大量的未标记数据以及相对较少的有标记数据。虽然 SPP 算法无需对训练样本进行标记,但是没有有效利用标记样本中所提供的鉴别信息。因此,如何在利用少量标记样本中的鉴别信息的同时,进一步充分利用大量无标记数据来帮助估计数据中的流形结构,以发现数据的整体分布信息,进而提高算法的性能,这在高光谱图像分类领域中尤为重要。

本文提出一种半监督稀疏鉴别嵌入算法(Semi-supervised Sparsity Discriminant Embedding, SSDE),同时将其应用于高光谱遥感影像分类。该方法结合了近邻流形结构及稀疏性^[9]的优点,不仅保留样本间的稀疏重构关系,而且通过利用少量有标记训练样本以及大量无标记训练本来发现蕴藏在高维数据内的内在属性以及低维流形结构,最终提高算法分类精度。

2 稀疏保持投影

2.1 稀疏表示

稀疏表示(Sparsity Representation, SR)^[11-12]最初用于信号的压缩和表示,它对于有残缺和有强干扰的数据具有良好的鲁棒性,现已成功应用于信号处理、统计学及模式识别等领域。其核心思想是将信号 $\mathbf{x} \in \mathbf{R}^m$ 分解为一系列基信号 $\mathbf{X} = \{\mathbf{x}_i \in \mathbf{R}^M\}_{i=1}^n$ 的线性组合 $\mathbf{x} \approx \sum_{i=1}^n s_i \mathbf{x}_i$,并希望尽

可能多的 s_i 系数为零(即向量 $\mathbf{S}=[s_1, s_2, \dots, s_n]^T$ 尽可能稀疏),其中,具有非零系数的信号揭示了原信号的主要特征和内在结构。数学模型如下:

$$\begin{aligned} & \min \|s_i\|_0 \\ & s. t. \quad \mathbf{x}_i = \mathbf{X}\mathbf{s}_i, \end{aligned} \quad (1)$$

式中: $\mathbf{X}=[x_1, x_2, \dots, x_n] \in \mathbf{R}^{m \times n}$ 为基信号矩阵, $\mathbf{s}_i \in \mathbf{R}^n$ 是用 \mathbf{X} 中除 \mathbf{x}_i 剩下的样本进行稀疏表示的重建系数向量, $\|s_i\|_0$ 表示 s_i 的 l_0 范数,用于衡量 s_i 的稀疏性。其中:向量的 l_0 范数表示向量中非零元素的数目。

2.2 稀疏重构

由于最小化 l_0 范数问题是一个 NP-hard 问题,求解困难。在系数足够稀疏的条件下, l_0 范数最小化问题可以转化为 l_1 范数最小化问题:

$$\begin{aligned} & \min \|s_i\|_1 \\ & s. t. \quad \mathbf{x}_i = \mathbf{X}\mathbf{s}_i \quad 1 = 1^T \mathbf{s}_i. \end{aligned} \quad (2)$$

式中: $\|s_i\|_1$ 表示 s_i 的 l_1 范数。向量 s_i 的 l_1 范数可表示为: $\sum_{i=1}^n |s_i|$ 。 $\mathbf{s}_i = [s_{i,1}, \dots, s_{i,i-1}, 0, s_{i,i+1}, \dots, s_{i,n}]^T \in \mathbf{R}^n$, $\mathbf{1} \in \mathbf{R}^n$ 表示全 1 向量。 s_{ij} 表示样本 \mathbf{x}_i 的重构系数,即:

$$\mathbf{x}_i = s_{i,1} \mathbf{x}_1 + \dots + s_{i,i-1} \mathbf{x}_{i-1} + s_{i,i+1} \mathbf{x}_{i+1} + \dots + s_{i,n} \mathbf{x}_n. \quad (3)$$

对于每一个 $\mathbf{x}_i \in \mathbf{X}$, 计算相应的 s_i , 就可以得到训练样本的稀疏重构矩阵 $\mathbf{s}=[s_1, s_2, \dots, s_n]$ 。

2.3 稀疏保持投影(SPP)

目前大多数降维算法可统一于图嵌入框架^[13]下,但其近邻图的构建依赖于人为选择的近邻参数,通常很难找到合适的参数值^[14]。SPP 通过稀疏重构处理,保留了样本稀疏重构信息的同时也保留了样本全局近邻结构。这样可从原始样本中提取出具备一定鉴别性的内在特征,而且不需要人为选择近邻参数值^[15-16]。

为了方便介绍以下算法,首先进行相关约定:假设 $\{\mathbf{x}_i, i=1, 2, \dots, n\}$ 为一组训练样本点,其中, $\mathbf{x}_i \in \mathbf{R}^m$, 令 $\mathbf{X}=[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbf{R}^{m \times n}$ 为训练样本集矩阵,对于某一样本 $\mathbf{x}_i \in \mathbf{X}$, $l_i \in (1, 2, \dots, c)$ 表示训练样本 \mathbf{x}_i 的类别标签。

为获得最优的投影向量, SPP 定义如下目标函数:

$$J(\mathbf{V}) = \min \sum_{i=1}^n \|\mathbf{V}^T \mathbf{x}_i - \mathbf{V}^T \mathbf{X}\mathbf{s}_i\|^2. \quad (4)$$

式中: $\mathbf{V} \in \mathbf{R}^{m \times n}$ 为投影矩阵, \mathbf{s}_i 为使 l_1 范数最小的

最优稀疏系数解。

对目标函数进行简单的代数运算,可得到:

$$\begin{aligned} J(\mathbf{V}) &= \min \text{tr}\{\mathbf{V}^T [\sum_{i=1}^n (\mathbf{x}_i - \mathbf{X}\mathbf{s}_i)(\mathbf{x}_i - \mathbf{X}\mathbf{s}_i)^T] \mathbf{V}\} = \\ & \min \text{tr}\{\mathbf{V}^T [\sum_{i=1}^n (\mathbf{X}\mathbf{e}_i - \mathbf{X}\mathbf{s}_i)(\mathbf{X}\mathbf{e}_i - \mathbf{X}\mathbf{s}_i)^T] \mathbf{V}\} = \\ & \min \text{tr}\{\mathbf{V}^T \mathbf{X} [\sum_{i=1}^n (\mathbf{e}_i - \mathbf{s}_i)(\mathbf{e}_i - \mathbf{s}_i)^T] \mathbf{X}^T \mathbf{V}\} = \\ & \min \text{tr}\{\mathbf{V}^T \mathbf{X} [\sum_{i=1}^n (\mathbf{e}_i \mathbf{e}_i^T - \mathbf{s}_i \mathbf{e}_i^T - \mathbf{e}_i \mathbf{s}_i^T + \mathbf{s}_i \mathbf{s}_i^T)] \mathbf{X}^T \mathbf{V}\} = \\ & \min \text{tr}[\mathbf{V}^T \mathbf{X} (\mathbf{I} - \mathbf{S} - \mathbf{S}^T - \mathbf{S}^T \mathbf{S}) \mathbf{X}^T \mathbf{V}], \end{aligned} \quad (5)$$

式中: \mathbf{e}_i 为 n 维向量(其中,除第 i 个元素为 1 外,其余都为零)。

为了避免出现退化解,增加约束 $\text{tr}(\mathbf{V}^T \mathbf{X}\mathbf{X}^T \mathbf{V}) = 1$ 。从而,目标函数可以表示为以下最优化问题:

$$J(\mathbf{V}) = \min_{\mathbf{V}} \frac{\text{tr}[\mathbf{V}^T \mathbf{X}\mathbf{S}_{\beta} \mathbf{X}^T \mathbf{V}]}{\text{tr}(\mathbf{V}^T \mathbf{X}\mathbf{X}^T \mathbf{V})}. \quad (6)$$

其中: $\mathbf{S}_{\beta} = \mathbf{I} - \mathbf{S} - \mathbf{S}^T + \mathbf{S}^T \mathbf{S}$ 。

通过求解广义特征值问题 $\mathbf{X}\mathbf{S}_{\beta} \mathbf{X}^T \mathbf{V} = \lambda \mathbf{X}\mathbf{X}^T \mathbf{V}$, 获得最优投影矩阵 $\mathbf{V} = [v_1, v_2, \dots, v_d]$, 其中: $v_i (i=1, \dots, d)$ 为前 d 个最小特征值对应的特征向量。

3 半监督稀疏鉴别嵌入(SSDE)

针对如何从少量标记数据以及大量未标记数据中提取稀疏鉴别特征的问题,本文在 SPP 算法的基础上提出一种新的半监督稀疏鉴别嵌入算法(SSDE),下面对其进行介绍。

3.1 SSDE 算法原理

SSDE 算法不仅保留了样本间的稀疏重构关系,而且可发现数据的内在流形结构,实现鉴别特征提取。

SSDE 算法首先对数据进行稀疏表示,因高光谱遥感数据中可能存在噪声,式(2)的条件不一定满足,因此,本文采用下式进行稀疏表示:

$$\begin{aligned} & \min \| \tilde{\mathbf{s}}_i \|_1 \\ & s. t. \quad \| \mathbf{x}_i - \mathbf{X}_j \tilde{\mathbf{s}}_i \| \leq \epsilon \quad 1 = 1^T \tilde{\mathbf{s}}_i. \end{aligned} \quad (8)$$

式中: \mathbf{X}_j 表示除 \mathbf{x}_i 外剩下的训练样本集, ϵ 为算法容许误差。

同时,为了发现数据中的流形结构信息,构建近邻图 G_w 。具体定义为:若样本点 x_i 和 x_j 属于

相同类别或者近邻,则在 G_w 中用一条边连接 x_i 和 x_j 两点,为前者赋予权值 β ,为后者赋予权值 1,否则 x_i 和 x_j 不相连。然后,计算 G_w 的权重矩阵 ω_w ,如下:

$$\omega_{w,ij} = \begin{cases} \beta, & x_i \text{ 与 } x_j \text{ 属于相同的类别} \\ 1, & x_i \text{ 与 } x_j \text{ 近邻} \\ 0, & \text{其他} \end{cases} \quad (9)$$

式中: $\beta > 1$,以进一步突出有标记样本在发现局部流形的作用,实现有区别地利用有标记与无标记样本的目的。

对于分类而言,为保持数据的局部特性,便于分类,希望同类和近邻样本点在低维嵌入空间中保持紧凑。因此,SSDE 目标函数可以定义为:

$$J(\mathbf{V}) = \min \sum_{ij} \| \mathbf{V}^T \mathbf{x}_i - \mathbf{V}^T \tilde{\mathbf{X}} \tilde{\mathbf{s}}_j \|^2 \omega_{w,ij}, \quad (10)$$

经过简单的代数运算,式(10)可化简为:

$$\begin{aligned} & \sum_{ij} \| \mathbf{V}^T \mathbf{x}_i - \mathbf{V}^T \tilde{\mathbf{X}} \tilde{\mathbf{s}}_j \|^2 \omega_{w,ij} = \\ & \sum_{ij} (\mathbf{V}^T \mathbf{x}_i - \mathbf{V}^T \tilde{\mathbf{X}} \tilde{\mathbf{s}}_j) (\mathbf{V}^T \mathbf{x}_i - \mathbf{V}^T \tilde{\mathbf{X}} \tilde{\mathbf{s}}_j)^T \omega_{w,ij} = \\ & \mathbf{V}^T \left(\sum_{ij} (\mathbf{x}_i - \tilde{\mathbf{X}} \tilde{\mathbf{s}}_j) \omega_{w,ij} (\mathbf{x}_i - \tilde{\mathbf{X}} \tilde{\mathbf{s}}_j)^T \right) \mathbf{V} = \\ & \mathbf{V}^T \left(\sum_{ij} (\mathbf{X} \mathbf{e}_i - \tilde{\mathbf{X}} \tilde{\mathbf{s}}_j) \omega_{w,ij} (\mathbf{X} \mathbf{e}_i - \tilde{\mathbf{X}} \tilde{\mathbf{s}}_j)^T \right) \mathbf{V} = \\ & \mathbf{V}^T \mathbf{X} \left(\sum_{ij} (\mathbf{e}_i \omega_{w,ij} \mathbf{e}_i^T - \mathbf{e}_i \omega_{w,ij} \tilde{\mathbf{s}}_j^T - \tilde{\mathbf{s}}_j \omega_{w,ij} \mathbf{e}_i^T + \tilde{\mathbf{s}}_j \omega_{w,ij} \tilde{\mathbf{s}}_j^T) \right) \mathbf{X}^T \mathbf{V} = \\ & \mathbf{V}^T \mathbf{X} \left(\sum_{ij} \mathbf{e}_i \omega_{w,ij} \mathbf{e}_i^T - \sum_{ij} \mathbf{e}_i \omega_{w,ij} \tilde{\mathbf{s}}_j^T - \sum_{ij} \tilde{\mathbf{s}}_j \omega_{w,ij} \mathbf{e}_i^T + \sum_{ij} \tilde{\mathbf{s}}_j \omega_{w,ij} \tilde{\mathbf{s}}_j^T \right) \mathbf{X}^T \mathbf{V} = \\ & \mathbf{V}^T \mathbf{X} (\mathbf{D}_w - \mathbf{W}_w \tilde{\mathbf{S}}^T - \mathbf{S} \mathbf{W}_w + \tilde{\mathbf{S}} \mathbf{D}_w \tilde{\mathbf{S}}^T) \mathbf{X}^T \mathbf{V} = \\ & \mathbf{V}^T \mathbf{X} \mathbf{S}_\gamma \mathbf{X}^T \mathbf{V}, \quad (11) \end{aligned}$$

式中: $\mathbf{S}_\gamma = \mathbf{D}_w - \mathbf{W}_w \tilde{\mathbf{S}}^T - \mathbf{S} \mathbf{W}_w + \tilde{\mathbf{S}} \mathbf{D}_w \tilde{\mathbf{S}}^T$, \mathbf{D}_w 为对角矩阵,对角线元素为 $D_{w,ii} = \sum_j \omega_{w,ij}$, $\tilde{\mathbf{S}} = [\tilde{s}_1, \tilde{s}_2, \dots, \tilde{s}_n]$ 。

然后,添加约束 $\text{tr}(\mathbf{V}^T \mathbf{X} \mathbf{X}^T \mathbf{V}) = 1$,则 SSDE 的目标函数可表达为:

$$J(\mathbf{V}) = \arg \min \mathbf{V}^T \mathbf{X} \mathbf{S}_\gamma \mathbf{X}^T \mathbf{V}, \quad (12)$$

$$s. t. \quad \mathbf{V}^T \mathbf{X} \mathbf{X}^T \mathbf{V} = \mathbf{I}$$

对式(12)运用拉格朗日乘法求解,即令 $J(\mathbf{V})$ 对 \mathbf{V} 的偏导数为零:

$$\frac{\partial}{\partial \mathbf{V}} (\mathbf{V}^T \mathbf{X} \mathbf{S}_\gamma \mathbf{X}^T \mathbf{V} - \lambda (\mathbf{V}^T \mathbf{X} \mathbf{X}^T \mathbf{V} - \mathbf{I})) = 0, \quad (13)$$

则式(13)可表示为以下的广义特征值求解问题:

$$\mathbf{X} \mathbf{S}_\gamma \mathbf{X}^T \mathbf{V} = \lambda \mathbf{X} \mathbf{X}^T \mathbf{V}. \quad (14)$$

通过上式可得到最优投影矩阵 $\mathbf{V} = [v_1, v_2, \dots, v_d]$,其中: $v_i (i=1, \dots, d)$ 为前 d 个最大特征值对应的特征向量。

3.2 SSDE 算法步骤

本文所提出的 SSDE 算法的步骤具体如下:

表 1 SSDE 算法步骤

Tab. 1 Steps of SSDE algorithm

输入: N 个训练样本的数据 $\mathbf{X} = \{(x_1, l_1), (x_2, l_2), \dots, (x_i, l_i), x_{i+1}, \dots, x_n\}$, $x_i \in \mathbf{R}^m$, 嵌入特征维数 d
输出: 投影矩阵 \mathbf{V} , 嵌入特征 \mathbf{Y}
Step1: 根据式(8)构造稀疏系数矩阵 $\tilde{\mathbf{S}}$
Step2: 构造近邻图 G_w , 根据式(9)计算权重矩阵 ω_w
Step3: 计算 $\mathbf{S}_\gamma = \mathbf{D}_w - \omega_w \tilde{\mathbf{S}}^T - \mathbf{S} \omega_w + \tilde{\mathbf{S}} \mathbf{D}_w \tilde{\mathbf{S}}^T$
Step4: 通过广义特征值问题 $\mathbf{X} \mathbf{S}_\gamma \mathbf{X}^T \mathbf{V} = \lambda \mathbf{X} \mathbf{X}^T \mathbf{V}$ 计算特征向量, 获得投影矩阵 \mathbf{V} 。

4 实验与分析

为了验证 SSDE 算法的有效性,本文在人工数据集、Indian Pine^[17] 和 Washington DC Mall^[17] 高光谱遥感数据集上进行实验,并将 SSDE 与 PCA、LDA、NPE 及 SPP 算法进行比较。在实验中,通过调整参数,使各个算法均达到最佳效果。具体设置为:PCA 的维数降到 50 维, LDA 的维数降到类别数-1 维, NPE 近邻 $k=5$, Baseline 将原始数据直接用最近邻分类器(1-NN)进行分类。

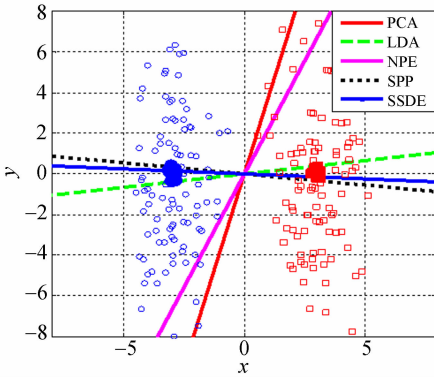
4.1 人工数据集(Synthetic datasets)

图 1 是 2 个人工数据集,图中红色和蓝色的点分别表示 2 类数据,空心和实心点则表示无标记样本和有标记样本(彩图见期刊电子版)。在图 1(a)的数据集中,2 类数据分别呈高斯分布。图中的线表示几种算法将数据样本投影到一维空间上得到的投影面。由图 1(a)可以看出,与无监督算法 PCA 相比,监督算法 LDA 能更好地利用样本的类别信息,得到的投影面更有利于数据点的分离。SPP 的投影面比 PCA、LDA 及 NPE 算法的都好,但本文提出的 SSDE 算法得到的投影面最优,能更有效地将不同数据分开。

图 1(b)中,第一类数据呈单高斯分布,第二

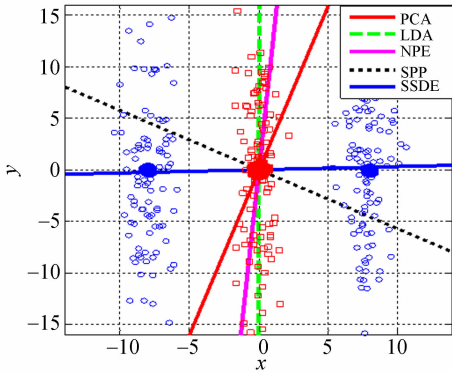
类数据呈 2 个分开的高斯分布。从图 1(b)可以看出,PCA、LDA 无法有效解决样本多模分布问题,得到的分类投影面效果较差。SPP 的投影面比 NPE、LDA 及 PCA 算法的都好,这是因为 SPP 有效利用稀疏表示自然的判别能力。但由于其本身是无监督算法,不能有效利用训练样本提供的类别信息,因此,得到的投影面也不是特别理想。而本文提出的 SSDE 算法在稀疏重构的基础上,能够更好地发现数据的内在流形结构,提取出鉴别特征,获得更好的投影方向,更有利于分类。

在 2 个合成数据集上实验验证了 SSDE 算法的有效性。



(a) 人工数据集 1

(a) Synthetic data set 1



(b) 人工数据集 2

(b) Synthetic data set 2

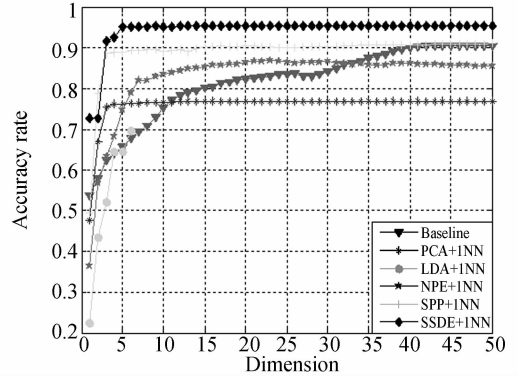
图 1 不同算法在人工数据集上的二维嵌入图

Fig. 1 Two-dimensional embedded graphs of different algorithms on synthetic data sets

4.2 Washington DC Mall 数据集

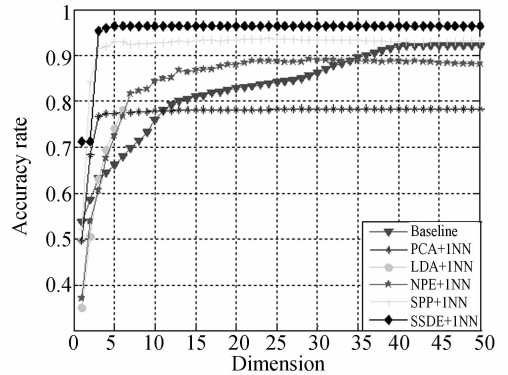
实验中采用的 Washington DC Mall 高光谱遥感影像^[17]为华盛顿特区国家广场的局部区域。该数据集有 191 个波段,波段间隔为 10 nm,空间

分辨率为 3 m。已知的地物类别包括“建筑”、“树林”、“(石径)小路”、“道路”、“草坪”、“湖泊”和“阴影”7 类。在实验中,先分别从每类地物中随机选取 2, 4, 6, 8 个数据点作为有类别点,然后从剩余训练样本中随机选取 60 个无类别标签的数据作为无标记样本,组成训练样本集,分别记做 2-lab、4-lab、6-lab、8-lab,剩下的全部数据点作为测试样本,每种方法重复做 10 次实验,最终分类精度取其平均值。



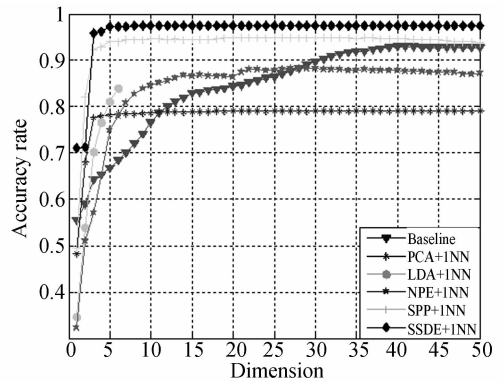
(a) 2 个标记样本

(a) 2-lab



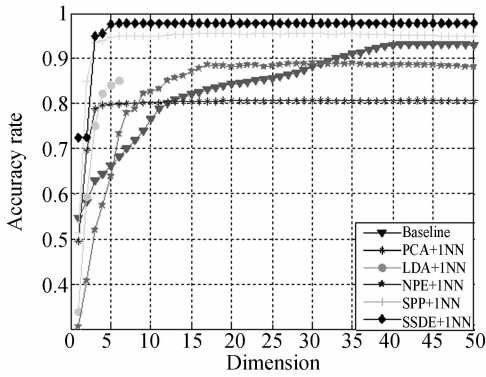
(b) 4 个标记样本

(b) 4-lab



(c) 6 个标记样本

(c) 6-lab



(d) 8 个标记样本

(d) 8-lab

图 2 不同算法在 Washington DC Mall 数据集上的平均分类精度

Fig. 2 Average classification accuracy of different algorithms on Washington DC Mall data set

在 Washington DC Mall 数据集进行实验,分别用 PCA、LDA、NPE、SPP 以及 SSDE 算法对其进行降维,利用最近邻分类器(1-NN)^[18]进行分类。图 2 给出了各种方法在不同维数嵌入特征上的平均分类精度。表 2 列出了不同方法在 Washington DC Mall 数据集上的最高总体精度。

表 2 各种方法在 Washington DC Mall 数据集上的最高识别率 (mean±std)

Tab. 2 The highest recognition rate of different methods on Washington DC Mall data set(mean±std)

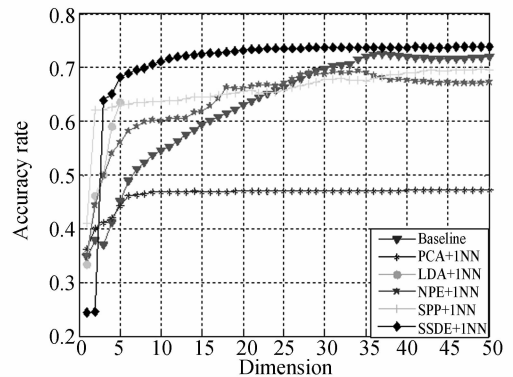
算法	最高总体分类精度 (%)			
	2-lab	4-lab	6-lab	8-lab
Baseline	90.59±3.39	92.22±3.16	92.95±2.56	93.18±2.65
PCA+1NN	77.05±2.8	78.38±2.03	79.06±2.88	80.68±2.13
LDA+1NN	69.62±4.26	78.2±4.8	83.93±4.18	77.05±3.28
NPE+1NN	87.08±4.53	89.32±2.04	88.2±3.98	88.98±4.78
SPP+1NN	91.12±4.16	93.1±2.94	94.88±1.8	95.55±1.32
SSDE+1NN	95.45±2.83	96.51±1.32	97.46±1.03	97.85±0.5

由图 2 和表 2 可以看出,训练样本越多,从样本中获得的先验信息就越全面,结果就能更有效地反映数据的低维流形结构,各种算法的总体分类精度也随之升高。LDA 是监督算法,利用有标记样本的类别信息提取鉴别特征,因此,分类精度比 PCA 高。NPE 和 SPP 虽然都是利用了训练数据的局部近邻结构,但是 SPP 保留样本稀疏重构信息的同时从训练样本中提取具备一定鉴别性的

低维特征,所以分类效果要优于 NPE。SSDE 算法在利用训练样本的稀疏重构关系建图的同时,能够充分利用少量有类别标记和大量无标记的训练样本,发现隐藏在低光谱数据中的内在几何结构,相比其他方法,分类精度有不同程度的提高。而且从图 2 可得到,SSDE 算法能够在较低维数获得最优分类精度,更有利于降维。

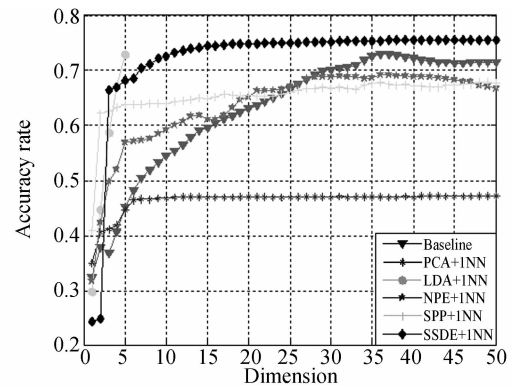
4.3 Indian Pine 数据集

实验中采用的 Indian Pine 高光谱遥感影像为覆盖了美国 Indiana 州西北的一块农业区域^[17]。该数据集的大小为 145×145,共有 220 个波段以及 17 类已知地物类别。考虑到噪声的影响,本实验从中选取了 200 个波段,并从数据点较多的地物类别中随机选择 6 类进行实验,分别为“Hay_windrowed”,“Soybeans_min”,“Woods”,“Corn_notill”,“Grass_pasture”,“Grass_trees”。在实验中,先分别从每类地物中随机选取 2,4,6,8 个有类别标签的数据点,然后从剩余训练样本中随机选取 60 个无类别标签的数据作为无标记样本组成训练样本集,分别记做 2-lab、4-lab、6-



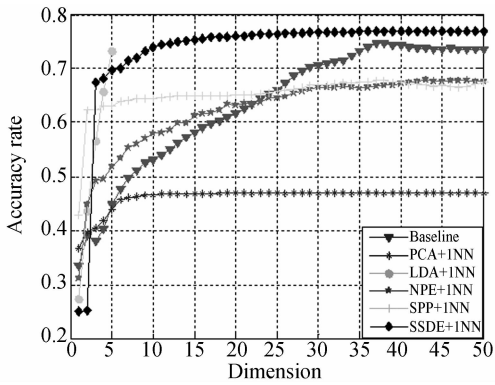
(a) 2 个标记样本

(a) 2-lab



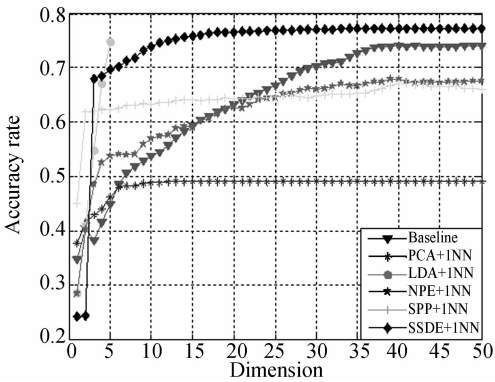
(b) 4 个标记样本

(b) 4-lab



(c) 6 个标记样本

(c) 6-lab



(d) 8 个标记样本

(d) 8-lab

图 3 不同算法在 Indian Pine 数据集上的平均分类精度

Fig. 3 Average classification accuracy of different algorithms on Indian Pine data set

lab、8-lab, 剩下全部数据点作为测试样本集。每种方法重复实验 10 次, 最终分类精度取其平均值。图 3 给出了各种算法在不同维数低维嵌入特征用最近邻分类器进行分类的平均总体精度。表

表 3 各种方法在 Indian Pine 数据集上的最高识别率 (mean±std)

Tab. 3 The highest recognition of different methods on Indian Pine data set (mean±std)

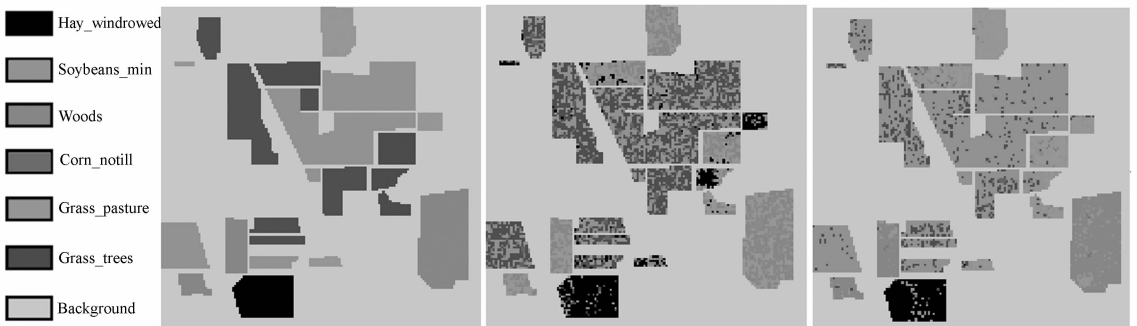
算法	最高总体分类精度 (%)			
	2-lab	4-lab	6-lab	8-lab
Baseline	72.48±2.8	72.94±3.13	74.67±2.42	74.17±1.29
PCA+1NN	47.1±2.81	47.03±2.69	46.94±2.71	49.12±1.42
LDA+1NN	63.48±3.7	72.87±4.7	73.22±4.9	74.8±3.73
NPE+1NN	69.34±3.6	69.17±3.64	67.94±3.59	67.96±2.39
SPP+1NN	69.66±2.92	67.8±3.11	67.89±2.5	67.18±1.81
SSDE+1NN	73.85±3.46	75.55±2.34	74.61±1.83	77.36±1.75

3 列出了各种不同方法在 Indian Pine 数据集上的最高总体精度。

由图 3 和表 3 可以得出, SSDE 在有标记样本比较少情况下, 不仅可以利用稀疏表示的自然鉴别能力, 而且可以利用少量标记和大量无标记样本来发现数据中的流形结构, 以提高分类精度。

本文算法在 2、4、6、8 个有标记训练样本的情况下, 其最高分类精度比 SPP 算法分别提高了 4.19%、7.75%、6.72% 和 10.18%, 而且能够更快地达到较高的分类率。

图 4 显示的是不同算法在 Indian Pine 数据集 6 种真实地物类别的地面分布情况以及实验分类效果图。从图 4 可以得出, SPP 相比 PCA、LDA、NPE 等算法, 分类精度有一定程度的提高, 这体现出利用稀疏表示进行自然判别能力的优势, 但由于其本质仍然是无监督算法, 不能有效利用标记样本提供的鉴别信息, 所以, 分类效果提升相对有限。而本文提出的 SSDE 算法在保留样本



(a) 图例

(b) 真实地物类别

(c) PCA 算法结果 (48.45%)

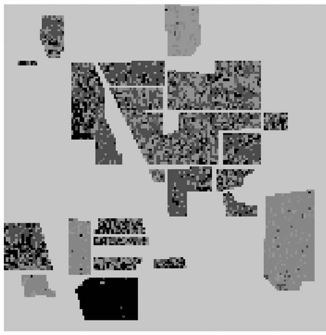
(d) LDA 算法结果 (67.51%)

(a) Legend

(b) True feature classification

(c) Results of PCA algorithm (48.45%)

(d) Results of LDA algorithm (67.51%)



(e) NPE 算法结果(62.83%)
(e) Results of NPE algorithm
(62.83%)



(f) SPP 算法结果(70.73%)
(f) Results of SPP algorithm
(70.73%)



(g) SSDE 算法结果(77.5%)
(g) Results of SSDE algorithm
(77.5%)

图4 不同算法在 Indian Pine 数据集上的分类图(括号内为分类精度)

Fig. 4 Classification graphs of different algorithms on Indian Pine data and their classification accuracies

间的稀疏重构关系的同时,通过利用少量有标记和大量无标记训练样本提取鉴别特征,相比其他算法的分类效果更好,分类精度也有较明显的提升。

5 结 论

本文提出一种新的半监督稀疏鉴别嵌入算法(SSDE)。该算法结合了近邻流形结构及稀疏性的优点,不仅保留了样本间的稀疏重构关系,而且通过引入少量有标记的训练样本以及大量无标记训练样本来获得数据的内在流形结构,实现鉴别

特征提取,更有利于分类。在 Indian pine 和 Washington DC Mall 数据集上的实验结果表明,本文提出的方法能够利用较少的有标记训练样本发现高光谱遥感影像中的内蕴结构。在随机选取 2,4,6,8 个有类别标记和 60 个无类别标记的数据作为训练样本的情况下,本文提出的 SSDE 算法最高分类精度分别达到了 77.36% 和 97.85%。

但是本文提出的半监督稀疏鉴别嵌入算法只是通过构建近邻图将高维数据嵌入在低维子空间中,使近邻和同类的样本保持紧凑,如何通过建图使具有不同类别的样本点保持远离,以得到更好的分类效果是下一步研究的内容。

参考文献:

- [1] 陈进. 高光谱图像分类方法研究[D]. 长沙:国防科技大学,2010.
CHEN J. *Hyperspectral image classification method*[D]. Changsha: National University of Defense Technology,2010. (in Chinese)
- [2] 苏红军,杜培军,盛业华. 一种基于分形维数的高光谱遥感波段选择算法研究[J]. 测绘通报,2007(3): 23-26.
SU H J, DU P J, SHENG Y H. A study of band selection algorithm of hyperspectral RS based on fractal dimensions [J]. *Bulletin of Surveying and Mapping*, 2007(3): 23-26. (in Chinese)
- [3] 黄鸿,秦高峰,冯海亮. 半监督流形学习及其在遥感影像分类中的应用[J]. 光学精密工程,2011,19(12):3025-3033.

- HUANG H, QIN G F, FENG H L. Semi-supervised manifold learning and its application to remote sensing image classification [J]. *Opt. Precision Eng.*, 2011,19(12):3025-3033. (in Chinese)
- [4] BACHMANN C M, AINSWORTH T L, FUSINA R A. Exploiting manifold geometry in hyperspectral imagery [J]. *IEEE Trans. On Geosciences and Remote Sensing*, 2005,43(3):884-897.
- [5] EDWARD J J. *A Use's Guide To Principal Components*[M]. New York: A Wiley-Interscience Publication,1992.
- [6] BELHUMEUR P N, HESPANHA J P, KRIGEMAN D. Eigenfaces vs. Fisherfaces: recognition using class specific linear projection [J]. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1997,19(7):711-720.
- [7] HE X F, YAN SH CH, HU Y X, et al.. Face rec-

- ognition using laplacianfaces [J]. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2005, 27(3):328-340.
- [8] HE X F, CAI D, YAN SH CH, *et al.*. Neighborhood preserving embedding [C]. *Proceedings of the 10th IEEE Int'l Conf. Computer Vision (ICCV'2005)*, Beijing, 2005, 2:1208-1213.
- [9] QIAO L SH, CHEN S C, TAN X Y. Sparsity preserving projections with applications to face recognition[J]. *Pattern Recognition*, 2010, 43 (1): 331-341.
- [10] WRIGHT J, YANG A Y, GANESHA, *et al.*. Robust face recognition via sparse representation [J]. *IEEE Transactions on Pattern Analysis Machine Intelligence*, 2009, 31(2):210-227.
- [11] INABA F K, SALLES E O T. Face recognition based on sparse representation and joint sparsity model with matrix completion [J]. *IEEE Latin America Transactions*, 2012, 10(1):1344-1351.
- [12] 黄勇. 稀疏保留投影及在表情识别中的应用[J]. *计算机应用*, 2012, 30(2):100-102.
HUANG Y. Sparse preserving projection and its application to expression recognition[J]. *Journal of Computer Applications*, 2012, 30(2):100-102. (in Chinese)
- [13] 黄鸿. 图嵌入框架下流形学习理论及应用研究 [D]. 重庆:重庆大学, 2008.
HUANG H. *Research on manifold learning theories and applications under the framework of graph embedding* [D]. Chongqing: Chongqing University, 2008. (in Chinese)
- [14] ROWEIS S T, SAUL L K. Nonlinear dimensionality reduction by locally linear embedding [J]. *Science*, 2000, 290(5500):2323-2326.
- [15] DEERWESTER S C, DUMAIS S T, LANDAUER T K, *et al.*. Indexing by latent semantic analysis [J]. *Journal of the American Society of Information Science*, 1990, 41(6):391-407.
- [16] FU Y, YAN S C, HUANG T S. Classification and feature extraction by simplexization [J]. *IEEE Trans. on Information Forensics and Security*, 2008, 3(1):91-100.
- [17] 王立志, 黄鸿, 冯海亮. 基于 SSMFA 与 kNNS 算法的高光谱遥感影像分类[J]. *电子学报*, 2012, 40(4):780-787.
WANG L ZH, HUANG H, FENG H L. Hyperspectral remote sensing image classification based on SSMFA and kNNS [J]. *Acta Electronica Sinica*, 2012, 40(4):780-787. (in Chinese)
- [18] GREGORY S, TREVOR D, PIOTR I. *Nearest-neighbor Methods in Learning and Vision: Theory and Practice* [M]. London: MIT Press, 2006.

作者简介:



黄 鸿(1980—),男,湖南新宁人,博士,副教授,硕士生导师,2003、2005、2008年于重庆大学分别获得学士、硕士和博士学位,主要从事流形学习、模式识别、遥感影像智能化处理等方面的研究。E-mail: hhuang. cq@ gmail. com



曲焕鹏(1989—),男,山东烟台人,硕士研究生,2012年于山东理工大学获得工学学士学位,主要从事图像处理、模式识别、遥感影像分类等方面的研究。E-mail: zqbhp2008@163.com

(版权所有 未经许可 不得转载)