

引用格式:马方,赵丽娜,何磊,等. 基于潜在低秩图判别分析的高光谱分类[J]. 北京化工大学学报(自然科学版), 2019, 46(4): 116–121.

MA Fang, ZHAO LiNa, HE Lei, et al. Hyperspectral image classification based on latent low-rank graphs[J]. Journal of Beijing University of Chemical Technology (Natural Science), 2019, 46(4): 116–121.

# 基于潜在低秩图判别分析的高光谱分类

马方<sup>1</sup> 赵丽娜<sup>1\*</sup> 何磊<sup>1</sup> 杨宏伟<sup>2</sup>

(北京化工大学 1. 理学院; 2. 信息中心, 北京 100029)

**摘要:**提出一种基于潜在低秩图判别分析(LatLGDA)算法,利用数据的自表示对数据的列表示系数矩阵和行表示系数矩阵同时施加低秩约束,得到保留数据结构的亲和矩阵,再与图嵌入模型相结合实现高光谱图像的流形降维并进行分类。与其他基于稀疏图或稀疏低秩图的高光谱特征提取算法相比,LatLGDA可利用数据的行信息弥补列信息的不足或缺失,对噪音的抗干扰能力更强;在真实数据集上的实验结果表明,LatLGDA算法具有较高的分类精度和运算效率,应用前景广阔。

**关键词:**稀疏图;稀疏低秩图;高光谱分类

**中图分类号:** TP75 **DOI:** 10.13543/j.bhxbzr.2019.04.017

## 引言

近年来,随着数据的爆炸式增长,数据量之大、维度之高和结构之复杂,都为数据分析带来了挑战。高光谱遥感是一种多维数据信息获取技术,它将成像技术和光谱技术相结合,在航空、卫星和飞机等多种平台上搭建传感器,通过远距离采集地面物体的电磁波反射数据来辨别地面物体的大小和位置信息。高光谱遥感在探测物体二维空间信息的同时还能探测其光谱信息,得到的高光谱图像具有立方体结构,因此可用 $x$ 轴和 $y$ 轴组成的平面表示空间信息, $z$ 轴表示光谱信息。由于包含信息丰富,高光谱图像在环境灾害检测、大气污染、农业生产及天文观测等领域都有着广泛的应用<sup>[1-2]</sup>。

高光谱分辨率高、波段数足够多,能够获取地物连续的光谱曲线。由于不同的物体在光谱曲线上存在显著差异,可以依据不同地物的光谱维特征来实现地物类别分类和目标检测。高光谱遥感图像所具有的大量光谱波段提供了物体的丰富信息,一般的

高光谱遥感图像都包含几百个波段,数据量为单波段的几百倍。高光谱遥感图像的光谱维数据分布在高达几百维的光谱特征空间中,相邻波段之间的高度相关性所带来的数据冗余增加了计算的复杂度,因此有必要对光谱维进行特征提取以得到低维数据,这无论在数据压缩还是数据分类方面都有着重要的意义<sup>[3]</sup>。

压缩感知和稀疏表示是高光谱图像处理领域的热点研究问题。Ly等<sup>[4]</sup>最先提出稀疏图判别分析(SGDA)算法,在图嵌入模型上引入稀疏表示,同时保留降维后的低维空间样本间表示的稀疏特性。由于SGDA算法未能在低维流形空间中保留数据的全局特性,Li等<sup>[5]</sup>提出了基于稀疏和低秩图嵌入(SLRGE)算法以及基于稀疏和低秩图判别分析(SLGDA)算法,通过加入核范数约束来保留具有全局约束的样本间表示的低秩特性。但SGDA和SLGDA算法仅考虑了列空间,还不能有效应对数据样本不足的情形。本文结合潜在低秩表示<sup>[6]</sup>提出潜在低秩图判别分析(latent low-rank graph discrimination analysis, LatLGDA)算法,首先通过不精确的增广拉格朗日乘法(IALM)<sup>[7]</sup>求解出最优的列表示系数矩阵,再与图嵌入模型相结合实现高光谱数据的特征提取,最终利用支持向量机(SVM)得到分类结果。LatLGDA算法从行空间和列空间重构数据,利用行信息弥补列信息的不足,并在低维空间里保留数据

收稿日期: 2018-12-13

基金项目: 国家自然科学基金(11301021/11571031)

第一作者: 女, 1993年生, 硕士生

\* 通信联系人

E-mail: zhaoln@mail.buct.edu.cn

的全局低秩特性。特别是在处理被噪音污染的数据或样本数量不足的数据时, LatLGDA 算法能够提取出更加丰富的数据信息, 从而提升数据子空间表示能力。

## 1 相关算法

### 1.1 稀疏低秩表示

基于大多数样本可以被某些带有重要信息的样本稀疏表示的特性, SLRGE 算法及 SLGDA 算法通过对列表示系数矩阵  $\mathbf{W}$  的  $l_1$  范数构建稀疏性来保留数据的局部结构, 通过核范数构造低秩特性来保留数据的整体低秩结构。

给定数据集  $\mathbf{X} \in \mathbf{R}^{D \times N}$ ,  $D$  为光谱特征维数,  $N$  为样本数目, 稀疏和低秩表示模型如式(1)所示

$$\begin{cases} \arg \min_{\mathbf{W}} \|\mathbf{W}\|_1 + \lambda \|\mathbf{W}\|_* \\ \text{s. t. } \mathbf{X} = \mathbf{X}\mathbf{W}, \mathbf{W}_{ii} = 0 \end{cases} \quad (1)$$

式中,  $\|\cdot\|_1$  是列表示系数矩阵  $\mathbf{W}$  的  $l_1$  范数, 其值等于矩阵中每个非零元素的绝对值之和;  $\|\cdot\|_*$  是  $\mathbf{W}$  的核范数, 其值为该矩阵所有奇异值之和;  $\lambda$  是正则化参数;  $\mathbf{W}_{ii} = 0, i = 1, 2, \dots, N$  是用于去掉平凡解、防止数据自表示的约束条件。式(1)中求解得到的  $\mathbf{W}$  是  $N \times N$  维矩阵, 其中每个  $N \times 1$  维列向量  $\mathbf{W}_i$  是其余样本点对第  $i$  个点的稀疏和低秩表示的系数。

### 1.2 图嵌入模型

设数据集  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbf{R}^{d \times n}$  有  $n$  个数据样本点, 每个样本点分布于  $d$  维空间中。图嵌入的模型是指找到  $\mathbf{X}$  在某个低维空间中的表示,  $\mathbf{X}$  的图模型有  $n$  个点, 每个点可以写成  $d$  维向量, 其中矩阵  $\mathbf{W}$  的每个元素  $\mathbf{W}_{ij}$  代表数据点  $\mathbf{x}_i$  与  $\mathbf{x}_j$  之间的权重, 则  $\mathbf{W}$  包含了  $\mathbf{X}$  内所有的数据样本点之间的关联度大小和数据的流形结构<sup>[8]</sup>。对于一个线性映射  $\mathbf{P}^T \in \mathbf{R}^{k \times d}, k < d$ , 将数据集  $\mathbf{X}$  映射至一个新的低维特征空间, 即  $\mathbf{Y} = \mathbf{P}^T \mathbf{X}, \mathbf{y}_i = \mathbf{P}^T \mathbf{x}_i$ ,  $\mathbf{X}$  的特征维可从  $d$  维降至  $k$  维, 则最优的投影方式由目标函数式(2)表示

$$\min_{\mathbf{P}} \sum_{i,j} (\mathbf{y}_i - \mathbf{y}_j)^2 \mathbf{W}_{ij} \quad (2)$$

式中,  $\mathbf{P}$  为要求解的最优投影矩阵。

常见的刻画数据点间关系的方法是令  $\mathbf{W}_{ij}$  等于两个数据点欧式距离的倒数, 这样数据点相距越远, 相应权重值就越小; 反之亦然。

若数据点  $\mathbf{x}_i$  与  $\mathbf{x}_j$  之间的权重越大, 则在降维时对目标函数式(2)的惩罚越大, 经过  $\mathbf{P}$  投影后的对

应低维空间中的点  $\mathbf{y}_i$  与  $\mathbf{y}_j$  相距越近; 反之若数据点之间的权重越小, 则投影后低维空间中两点间的距离越远。所以找到一个合适的矩阵  $\mathbf{W}$  来较好地表示数据点之间的关系, 使得在投影后保持这种流形结构, 是图嵌入模型中最关键的步骤。

## 2 潜在低秩图判别分析算法

将数据本身作为字典, 从列空间和行空间共同重构原始数据, 同时考虑列信息和行信息, 以便恢复出对噪音更加鲁棒的原始数据。对列表示系数矩阵  $\mathbf{W}$  和行表示系数矩阵  $\mathbf{G}$  同时施加核范数正则化约束, 最终将观测数据  $\mathbf{X}$  分解为 3 部分: 低秩部分  $\mathbf{X}\mathbf{W}$ 、显著特征部分  $\mathbf{G}\mathbf{X}$  和误差项  $\mathbf{E}$ 。目标函数如式(3)

$$\begin{cases} \min_{\mathbf{W}, \mathbf{G}, \mathbf{E}} \|\mathbf{W}\|_* + \|\mathbf{G}\|_* + \lambda \|\mathbf{E}\|_{2,1} \\ \text{s. t. } \mathbf{X} = \mathbf{X}\mathbf{W} + \mathbf{G}\mathbf{X} + \mathbf{E} \end{cases} \quad (3)$$

式中,  $\|\mathbf{W}\|_* = \sum_i \sigma_i$  是  $\mathbf{W}$  的核范数,  $\sigma_i$  是  $\mathbf{W}$  的奇异值;  $\|\mathbf{E}\|_{2,1} = \sum_i \sqrt{\sum_j \mathbf{E}_{ij}^2}$  是  $\mathbf{E}$  的  $l_{2,1}$  范数, 用于刻画数据的误差项, 其中  $\mathbf{E}_{ij}$  是矩阵的非零元素。

式(3)为凸优化问题, 可通过 IALM 算法求解。加入两个辅助矩阵  $\mathbf{J}, \mathbf{K}$ , 则可将式(3)转化为如下同等问题

$$\begin{cases} \min_{\mathbf{J}, \mathbf{K}, \mathbf{W}, \mathbf{G}, \mathbf{E}} \|\mathbf{J}\|_* + \|\mathbf{K}\|_* + \lambda \|\mathbf{E}\|_{2,1} \\ \text{s. t. } \mathbf{X} = \mathbf{X}\mathbf{W} + \mathbf{G}\mathbf{X} + \mathbf{E}, \mathbf{W} = \mathbf{J}, \mathbf{G} = \mathbf{K} \end{cases} \quad (4)$$

构造式(4)的增广拉格朗日函数如式(5)所示<sup>[7]</sup>

$$\begin{aligned} L(\mathbf{J}, \mathbf{K}, \mathbf{W}, \mathbf{G}, \mathbf{E}) = & \|\mathbf{J}\|_* + \|\mathbf{K}\|_* + \lambda \|\mathbf{E}\|_{2,1} + \langle \mathbf{Y}_2, \mathbf{W} - \mathbf{J} \rangle + \langle \mathbf{Y}_3, \mathbf{G} - \mathbf{K} \rangle + \langle \mathbf{Y}_1, \mathbf{X} - \\ & \mathbf{X}\mathbf{W} - \mathbf{G}\mathbf{X} - \mathbf{E} \rangle + \frac{\mu}{2} \|\mathbf{W} - \mathbf{J}\|_F^2 + \frac{\mu}{2} \|\mathbf{G} - \mathbf{K}\|_F^2 + \frac{\mu}{2} \\ & \|\mathbf{X} - \mathbf{X}\mathbf{W} - \mathbf{G}\mathbf{X} - \mathbf{E}\|_F^2 \end{aligned} \quad (5)$$

式中  $\mathbf{Y}_1, \mathbf{Y}_2, \mathbf{Y}_3$  为拉格朗日乘子,  $\mu > 0$  为惩罚参数。

对式(5)最小化来交替更新变量, 即逐一更新每个变量时固定其他变量, 则每个变量的更新表达式如下。

固定  $\mathbf{K}, \mathbf{W}, \mathbf{G}, \mathbf{E}$ , 更新  $\mathbf{J}$ , 目标函数转化为式(6)

$$\begin{aligned} \mathbf{J} = \arg \min_{\mathbf{J}} \frac{1}{\mu} \|\mathbf{J}\|_* + \frac{1}{2} \left\| \mathbf{J} - \left( \mathbf{W} + \frac{\mathbf{Y}_2}{\mu} \right) \right\|_F^2 = \\ \Theta_{\frac{1}{\mu}} \left( \mathbf{W} + \frac{\mathbf{Y}_2}{\mu} \right) \end{aligned} \quad (6)$$

固定  $J, W, G, E$ , 更新  $K$ , 目标函数转化为式 (7)

$$K = \arg \min_K \frac{1}{\mu} \|K\|_* + \frac{1}{2} \left\| K - \left( G + \frac{Y_3}{\mu} \right) \right\|_F^2 = \Theta_{\frac{1}{\mu}} \left( G + \frac{Y_3}{\mu} \right) \quad (7)$$

式中  $\Theta_{\frac{1}{\mu}}(\cdot)$  为奇异值阈值算子<sup>[9]</sup>。

固定  $J, K, G, E$ , 更新  $W$ , 目标函数转化为式 (8)

$$L(W) = \frac{\mu}{2} \left( \left\| W - J + \frac{Y_2}{\mu} \right\|_F^2 + \left\| X - XW - GX - E + \frac{Y_1}{\mu} \right\|_F^2 \right) \quad (8)$$

式(8)中只包含矩阵的 Frobenius 范数, 因此最小化时可直接对  $W$  求导, 整理可得

$$W = (X^T X + I)^{-1} \left( X^T (X - GX - E) + J + \frac{X^T Y_1 - Y_2}{\mu} \right) \quad (9)$$

固定  $J, K, W, E$ , 更新  $G$ , 目标函数转化为式 (10)

$$L(G) = \frac{\mu}{2} \left\| G - K + \frac{Y_3}{\mu} \right\|_F^2 + \frac{\mu}{2} \left\| X - XW - GX - E + \frac{Y_1}{\mu} \right\|_F^2 \quad (10)$$

最小化式(10)时可直接对  $G$  求导, 整理可得更新表达式如式(11)

$$G = \left( (X - XW - E) X^T + K + \frac{Y_1 X^T - Y_3}{\mu} \right) (XX^T + I)^{-1} \quad (11)$$

固定  $J, K, W, G$ , 更新  $E$ , 目标函数转化为式 (12)

$$E = \arg \min_E \frac{\lambda}{\mu} \|E\|_{2,1} + \frac{1}{2} \left\| E - \left( X - XW - GX + \frac{Y_1}{\mu} \right) \right\|_F^2 \quad (12)$$

式(12)的最小化可由引理 1 得到。

**引理 1**<sup>[10]</sup> 当式(13)最优化问题中的  $H$  为已知矩阵时, 最优解  $E^*$  的第  $i$  列元素值如式(14)所示

$$\arg \min_E \alpha \|E\|_{2,1} + \frac{1}{2} \|E - H\|_F^2 \quad (13)$$

$$[E^*]_{:,i} = \begin{cases} \frac{\|H_{:,i}\|_2 - \alpha}{\|H_{:,i}\|_2} H_{:,i}, & \|H_{:,i}\|_2 > \alpha \\ 0, & \text{其他} \end{cases} \quad (14)$$

通过式(15)和式(16)更新  $Y_1, Y_2, Y_3$  和  $\mu$

$$\begin{cases} Y_1 = Y_1 + \mu(X - XW - GX - E) \\ Y_2 = Y_2 + \mu(W - J) \\ Y_3 = Y_3 + \mu(G - K) \end{cases} \quad (15)$$

$$\mu = \min(\rho\mu, \max_{\mu}) \quad (16)$$

式(16)中  $\rho > 1$  是标量, 用于更新  $\mu$  的值, 根据 IALM 算法,  $\rho$  取 1.5。

对目标函数式(3)采用 IALM 算法求解的具体流程如算法 1 所示。

**算法 1** 利用 IALM 求解式(3)

- 1) 输入  $X \in \mathbf{R}^{d \times n}, \lambda > 0$
- 2) 初始化  $W_0, J_0, K_0, G_0, E_0 = 0, \mu_0 = 10^{-6}$   
 $\max_{\mu} = 10^6, \varepsilon = 10^{-6}, \rho = 1.5$
- 3) 迭代 ( $t = 0, 1, 2, \dots$ )

固定其他变量, 通过式(6)更新  $J$

固定其他变量, 通过式(7)更新  $K$

固定其他变量, 通过式(9)更新  $W$

固定其他变量, 通过式(11)更新  $G$

固定其他变量, 通过式(12)更新  $E$

由式(15)更新  $Y_1, Y_2, Y_3$

由式(16)更新  $\mu$

判断迭代结束条件为

$$\begin{cases} \|X - XW - GX - E\|_{\infty} < \varepsilon \\ \|W - J\|_{\infty} < \varepsilon, \|G - K\|_{\infty} < \varepsilon \end{cases}$$

- 4) 输出 最优解  $W, G, E$

应用算法 1 求解出  $W$  之后, 通过基于图嵌入模型找到一个最优的投影矩阵  $P$ , 经过投影变换后的低维数据表示为  $Y = P^T X$ 。为了在低维空间中保留数据的原有流形结构, 根据 1.1 节的图嵌入模型式(2)得到目标函数

$$\begin{cases} \arg \min_P \sum_{i=1}^n \left\| P^T x_i - \sum_{j=1}^n w_{ij} P^T x_j \right\|^2 = \\ \arg \min_P \text{tr}(P^T X L X^T P) \\ \text{s. t. } P^T X L_P X^T P = I \end{cases} \quad (17)$$

式中,  $L = D - W$ , 其中  $n \times n$  维对角矩阵  $D$  是  $W$  的度矩阵,  $L$  是  $W$  的拉普拉斯矩阵;  $I = L_P$  为单位矩阵;  $P^T X L_P X^T P = I$  为拉格朗日约束条件。

式(17)的求解过程是广义特征值分解问题<sup>[11]</sup>, 即  $X L X^T P = V X L_P X^T P$ , 其中  $V$  为对角矩阵, 对角线上每一个元素对应一个广义特征值, 投影矩阵  $P$  由对应的特征值向量组成。

综上, 本文提出的 LatLGDA 算法用于高光谱图

像分类的具体步骤如下。

1)数据预处理。输入高光谱数据集,先对数据进行归一化运算,再将归一化后的三维高光谱数据转化为二维数据矩阵,矩阵的行数等于光谱的波段数,列数等于样本点数,则矩阵每列表示每个样本点的光谱特征。

2)选择训练集和测试集。从归一化后的数据矩阵的每一类别中随机选取一定数量数据组成训练样本集,剩余作为测试样本集。

3)通过算法 1 求解最优列表示系数矩阵  $W$ 。

4)求解最优投影矩阵  $P$ 。通过  $W$  计算出拉普拉斯矩阵  $L$ ,再根据广义特征值求解最优投影矩阵  $P$ ,则训练集降维后的数据变成  $Y = P^T X$ 。

5)高光谱图像分类。利用 SVM 对降维后的训练集和测试集进行分类。

### 3 实验分析

为了说明本文提出算法的有效性,选取真实的高光谱图像数据集 University of Pavia<sup>[12]</sup> 进行高光谱分类实验,并与不同图判别分析方法的结果进行对比。计算测试样本每一类别的分类准确率和总体分类准确率作为各个算法的评价指标。仿真实验环境为 Matlab R2016b 版,机器配置为 ThinkCentre M8400t-N000,内存 4 GB,频率 3.4 GHz,操作系统 Windows 7。

University of Pavia 实验数据集由 ROSIS 传感器采集,图像场景是位于意大利 Pavia 城的一所大学。此高光谱图像共有  $610 \times 340$  个像元,波段数 103,空间分辨率 1.3 m,其中有标签的像元为 42 776 个,分别属于 9 个类别。在分类实验中,首先选取每一类的 5% 作为训练集,9 个类别的样本信息如表 1 所示。为保证实验公平性,不同算法都先通过同样的训练样本数据求解各自的最优投影矩阵进行降维,得到降维后的训练数据集与测试数据集,再通过 SVM 分类来比较这些算法特征提取的效果。

SVM 工具选用台湾大学林智仁开发的 LIBSVM 工具包。惩罚系数  $c$  取值范围  $0.000\ 1 \sim 10\ 000$ ,  $c$  越大对错误例惩罚程度就越大,本文实验中所有算法均取  $c = 10\ 000$ ;选择核函数类型为高斯径向基核函数(RBF)时的参数  $\sigma$ ,分类时  $\sigma$  从  $\{0.01, 0.05, 0.5, 1, 5, 10, 50, 100\}$  中依次循环取值进行计算,最终的分类结果取 8 个值中最高者。

表 1 University of Pavia 数据集选择的训练集和测试集  
Table 1 Training and test samples selected from University of Pavia data set

类号	类目	训练样本数	测试样本数
1	Asphalt	332	6 299
2	Meadows	934	17 715
3	Gravel	105	1 994
4	Trees	154	2 910
5	Painted metal sheets	68	1 277
6	Bare soil	252	4 777
7	Bitumen	67	1 263
8	Self-blocking bricks	184	3 498
9	Shadows	48	899
合计		2 144	40 632

图 1 为降维数变化时,LatLGDA 算法与已有算法的分类精确度对比。从图中可以看出,随着最终降维数从 10 增加到 50,本文算法的分类精确度始终高于 SGDA、SLRGE 和 SLGDA。然后统一将光谱特征维度降至 22,得到不同方法对每一类别和总体的分类精确度如表 2 所示。可以看出本文算法的类别和总体分类精确度均为最高,其中总体分类精确度分别比 SGDA、SLRGE 和 SLGDA 方法高 8.9%, 5.12% 和 3.58%。

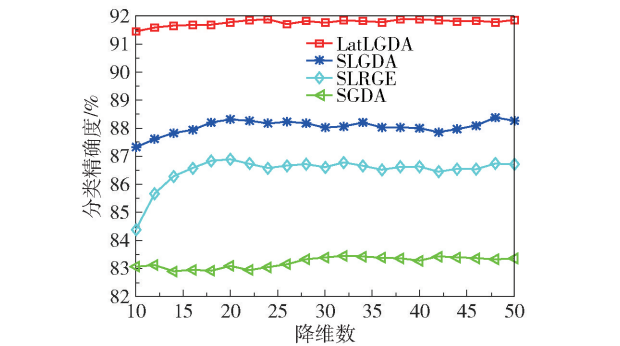


图 1 降维后不同维度对分类精确度的影响  
Fig.1 The effect of dimensionality on classification accuracy

实验中选取 5% 的训练集时,样本数小于高光谱维度的第 5、7 和 9 类别出现小样本问题,这与文献[4]中选取 8% 训练集时 SLGDA 算法的分类结果相似,即第 7 类别的分类精确度是 9 个类别中最低的,这是由于训练样本少造成的。表 2 中本文算法关于第 7 类的分类精确度最低,为 65.64%,但仍比 SGDA 和 SLGDA 方法分别高出 29.95% 和 9.47%,说明本文方法利用行信息弥补列信息的不足,能更好地应对小样本问题。



表2 不同算法的分类精确度

Table 2		Classification accuracy of different algorithms			
类号	类目	分类精确度/%			
		SGDA	SLRGE	SLGDA	LatLGDA
1	Asphalt	87.17	91.86	90.56	95.13
2	Meadows	92.92	94.58	95.09	96.98
3	Gravel	56.98	45.84	59.79	63.89
4	Trees	85.28	89.20	91.45	94.13
5	Painted metal sheets	99.48	99.78	99.48	99.85
6	Bare soil	64.41	78.66	80.99	86.67
7	Bitumen	35.69	46.77	56.17	65.64
8	Self-blocking bricks	68.06	76.75	78.52	82.54
9	Shadows	99.58	98.20	96.73	99.68
总体分类精确度/%		82.95	86.73	88.27	91.85

为验证本文算法在训练样本数目很少时的分类效果,分别选取每一类的 1%、2%、3%、4%、5% 作为训练集,分类结果如图 2 所示。可以看出随着训练集的增大,SGDA、SLRGE、SLGDA 和 LatLGDA 的分类精确度都有所提升,其中 LatLGDA 分类效果最好,SLRGE 和 SLGDA 次之,SGDA 最差。

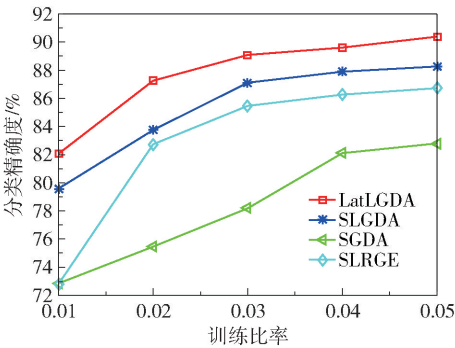


图2 训练样本数目变化对分类精确度的影响  
Fig. 2 The influence of the training ratio on the classification

为了从计算时间和效率方面比较不同算法的差异,在训练样本占比 1%、2%、3%、4%、5% 条件下,记录不同算法的运算时间如图 3 所示。可以看出,随着训练比率的提高,各算法的计算时间都随之增加,但比较来看,本文算法计算时间最短、效率最高,更适用于真实高光谱数据分类。

4 结束语

将潜在低秩表示和图嵌入模型相结合提出一种高光谱图像光谱维特征提取算法,该算法同时考虑

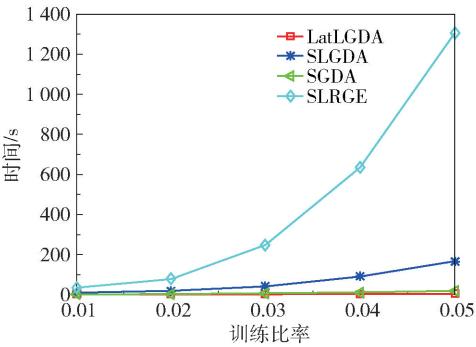


图3 不同算法的运算时间比较  
Fig. 3 Computation time of different algorithms

行信息和列信息,能够恢复出对噪音更加鲁棒的原  
始数据;与其他基于图判别分析的高光谱分类算法  
相比,本文算法对噪音更加鲁棒,分类精度更高,计  
算效率更快,并且在训练样本较少时仍具有较高的  
分类准确率。

参考文献:

[1] 赵春晖,王立国,齐滨. 高光谱遥感图像处理方法及应用[M]. 北京:电子工业出版社,2016:369-381.  
ZHAO C H, WANG L G, QI B. Hyperspectral remote sensing image processing methods and applications[M]. Beijing: Electronics Industry Press, 2016: 369-381. (in Chinese)

[2] CHIVASA W, MUTANGA O, BIRADAR C. Application of remote sensing in estimating maize grain yield in heterogeneous African agricultural landscapes: a review[J]. International Journal of Remote Sensing, 2017, 38(23): 6816-6845.

[3] YANG H, DU Q, CHEN G S. Particle swarm optimization-based hyperspectral dimensionality reduction for urban land cover classification[J]. IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing, 2012, 5(2): 544-554.

[4] LY N H, DU Q, FOWLER J E. Sparse graph-based discriminant analysis for hyperspectral imagery[J]. IEEE Transactions on Geoscience & Remote Sensing, 2014, 52(7): 3872-3884.

[5] LI W, LIU J B, DU Q. Sparse and low-rank graph for discriminant analysis of hyperspectral imagery[J]. IEEE Transactions on Geoscience & Remote Sensing, 2016, 54(7): 4094-4105.

[6] LIU G C, YAN S C. Latent low-rank representation for subspace segmentation and feature extraction[C] // Proceedings of the 2011 IEEE International Conference on

- Computer Vision. Barcelona, 2011: 1615–1622.
- [7] LIN Z C, CHEN M M, MA Y. The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices[J/OL]. (2013–10–18)[2018–11–02]. <https://arxiv.org/abs/1009.5055>.
- [8] YAN S C, XU D, ZHANG B Y, 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.
- [9] CAI J F, CANDES E J, SHEN Z. A singular value thresholding algorithm for matrix completion[J]. SIAM Journal on Optimization, 2010, 20(4): 1956–1982.
- [10] YANG J F, YIN W T, ZHANG Y, et al. A fast algorithm for edge-preserving variational multichannel image restoration[J]. SIAM Journal on Imaging Sciences, 2009, 2(2): 569–592.
- [11] GUATTERY S, MILLER G L. Graph embeddings and laplacian eigenvalues[J]. SIAM Journal on Matrix Analysis and Applications, 2000, 21(3): 703–723.
- [12] PENG B, LI W, XIE X M, et al. Weighted-fusion-based representation classifiers for hyperspectral imagery[J]. Remote Sensing, 2015, 7(11): 14806–14826.

## Hyperspectral image classification based on latent low-rank graphs

MA Fang<sup>1</sup> ZHAO LiNa<sup>1\*</sup> HE Lei<sup>1</sup> YANG HongWei<sup>2</sup>

(1. Faculty of Science; 2. Center for Information Technology, Beijing University of Chemical Technology, Beijing 100029, China)

**Abstract:** In this paper, latent low-rank graph discrimination analysis (LatLGDA) is proposed. Our algorithm uses self-representation of the data to apply low-rank constraints to the column and row representation coefficient matrix in order to obtain the affinity matrix of the retained data structure. Combined with a graph embedding model, both manifold dimension reduction and classification of hyperspectral images can be realized. Compared with other hyperspectral feature extraction algorithms based on principles such as sparse graphs or sparse and low-rank graph discrimination analysis, LatLGDA can use the row information data to compensate for the lack of column information and has better resistance to interference from noise. Experiments on a real hyperspectral data set from the University of Pavia demonstrate that LatLGDA has the advantages of high classification accuracy, fast operation efficiency and broad application prospects.

**Key words:** sparse graph; sparse and low-rank graph; hyperspectral image classification

(责任编辑:汪 琴)