

## 基于矩阵分解的高光谱数据特征提取

魏峰, 何明一, 冯燕, 李晓会

(西北工业大学 电子信息学院, 陕西 西安 710129)

**摘要:**利用有限的标记样本,将其作为硬性约束加入矩阵分解中;同时构建局部邻域 graph,挖掘数据的流形结构并保持局部的不变特性,提出一种基于矩阵分解的高光谱数据特征提取(FEMF)方法.经过矩阵分解,使得原始高维光谱特征空间中相近的数据在低维空间中仍然相近,而相同类别的标记数据则被投影到同一个位置.这样的低维表示具有更强的判别性能,从而得到更好的分类和聚类效果.该方法的求解过程是非凸规划问题,同时给出了一个乘性更新规则获得局部优化解.最后,对真实高光谱数据进行特征提取验证了该方法的有效性.

**关键词:**高光谱;遥感;流形;特征提取;矩阵分解

中图分类号:TP701 文献标识码:A

## Feature extraction on matrix factorization for hyperspectral data

WEI Feng, HE Ming-Yi, FENG Yan, LI Xiao-Hui

(School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710129, China)

**Abstract:** Limited labeled samples is used adequately as a hard constraint into matrix factorization. Meanwhile, the local graph of data is constructed to exploit the manifold structure and maintain the local invariance. As a result, feature extraction on matrix factorization for hyperspectral data(FEMF) was proposed. Through matrix decompose process, the nearby data in original space is still close after dimensionality reduction, and the congeneric labeled data is projected into the same position. Such low-dimensional representation has more powerful discriminate performance for classification and clustering. This issue is a non-convex programming problem, and an iterative multiplicative updates algorithm is introduced to achieving the local optimization. The efficiency of the proposed method was verified in real hyperspectral data feature extraction.

**Key words:** hyperspectral, remote sensing, manifold, feature extraction, matrix factorization

**PACS:** 42.30.Va, 07.07.Df

### 引言

高光谱遥感技术基于电磁波谱理论,在可见光,近红外,中红外和热红外波段的范围内,采集许多非常窄的光谱的影像数据<sup>[1]</sup>,其光谱分辨率达到纳米数量级,光谱通道数多达数十甚至数百.著名的高光谱传感器如美国航空航天局研制的航空机载可见光-红外成像光谱仪(Airborne Visible/Infrared Imaging Spectrometer, AVIRIS),其空间分辨率为20 m,光谱分辨率为10 nm,有224个波段,可广泛用于精细农业、资源调查、大气观测等方面的研究<sup>[2]</sup>.

随着高光谱传感器技术的发展,高光谱数据的

光谱分辨率日益提高.然而巨大的数据量以及光谱波段间的强相关性也为数据的处理造成了困难<sup>[1]</sup>.

在模式识别领域,由于标记样本有限,数据维数较高往往导致分类性能下降<sup>[2]</sup>.因此,利用特征提取技术通过空间映射,去除波段间大量冗余信息实现数据降维,成为高光谱数据解译过程中重要的预处理部分.

近年来,矩阵分解技术作为高维数据特征提取的工具受到重要关注<sup>[3]</sup>.然而,大多数方法并不能够保证经过降维后的高维数据仍然保留非负特性,在很多实际应用中没有物理意义.基于此,非负矩阵分解(Nonnegative Matrix Factorization, NMF)<sup>[4]</sup>和概

收稿日期:2014-04-11,修回日期:2014-10-07

基金项目:国家自然科学基金(61171154,61420106007)

Foundation items: Supported by National Natural Science Foundation of China (61171154, 61420106007)

作者简介(Biography):魏峰(1987-),男,博士,陕西咸阳人,主要研究方向为高光谱遥感信息获取与处理,模式识别和信号处理.

E-mail: weifeng@163.com

Received date: 2014-04-11, revised date: 2014-10-07

念分解 (Concept Factorization, CF)<sup>[5]</sup> 被提了出来. NMF 和 CF 在提取数据光谱特征的过程中没有应用标记信息<sup>[6-7]</sup>, 然而, 在高光谱技术应用中, 尽管对采集的光谱数据全部进行标记是一项复杂且代价高的事情, 但对少量样本进行标记却是一个相对容易的过程. 在大量无标记样本中加入少量标记样本, 能够提取到光谱间的主要信息, 得到具有更强泛化性能的模式<sup>[8]</sup>. 此外, 研究表明, 由于高光谱数据在其高分辨率光谱波段上的光谱反射值呈现为一条近似连续的曲线, 因此高光谱数据可以看作采样自一个嵌入在原始光谱特征空间中的非线性低维流形结构<sup>[9]</sup>. 所以, 高光谱数据基于流形结构的半监督降维具有重要的应用价值, 如何在该模式下应用矩阵分解方法提取光谱之间的主要特征是一个有意义的研究问题.

本文充分利用高光谱技术应用中有限标记样本的类别信息, 将其作为硬性约束<sup>[10-11]</sup> 加入矩阵分解的过程中; 同时根据图谱理论和流形学习思想<sup>[12-13]</sup>, 通过构建一个局部邻域 graph 结构来描述数据分布的几何结构, 提出基于矩阵分解的高光谱数据特征提取 (Feature Extraction on Matrix Factorization, FEMF) 方法. 本文主要思想是, 对于原始高维光谱特征空间中的数据, 经过矩阵分解, 相同类别的标记样本具有相同的低维表示. 与此同时, 通过构建局部 graph, 挖掘数据的潜在流形结构, 使得在新的低维空间中, 原来相近的数据点仍然彼此相近. 综合而述, 原始高光谱数据经过该思想进行矩阵分解之后, 最终获得的低维表示具有更强的判别性能, 空间结构更加优化, 从而有利于分类或聚类. 该方法的求解过程是非凸规划问题, 同时给出一个迭代的乘性更新规则, 通过替代的凸规划问题实现高光谱数据的光谱特征提取.

## 1 相关理论简介

给定一组高光谱数据矩阵  $X = [x_1, \dots, x_N] \in \mathbb{R}^{M \times N}$ , 每一个列向量代表具有  $M$  个光谱波段的原始数据点, NMF 旨在寻找两个非负的矩阵  $U = [u_r] \in \mathbb{R}^{M \times R}$  和  $V = [v_j] \in \mathbb{R}^{N \times R}$ , 使得它们的乘积可以近似表示  $X$ , 取  $R \ll M$ , 则  $v_i$  为高光谱数据  $x_i$  在基向量集  $\{u_r | r = 1, \dots, R\}$  上的低维表示, 从而通过矩阵重构实现光谱特征提取的过程.

很多高维数据处在高度非线性分布的情况, 为了解决这个问题, 概念分解作为 NMF 的改进提了出来. 在概念分解中, 任意一个高维数据  $x_i$  被表示为

一组基向量的非负线性组合, 与此同时任意一个基向量  $u_r$  也被表示为数据集的非负线性组合:

$$x_i \cong \sum_{r=1}^R u_r v_r, u_j \cong \sum_{j=1}^N x_j s_j, \quad (1)$$

其中,  $s_j \geq 0$ , 设  $S = \{s_j\} \in \mathbb{R}^{N \times R}$ , 因此其原理思想可以由数学方程表示如下:

$$X \cong UV^T = XSV^T. \quad (2)$$

使用欧氏距离衡量重构误差, 则需要通过最小化下列的目标函数得到优化结果:

$$\begin{aligned} \Phi_{CF} &= \|X - XSV^T\|_F^2 \\ &= \text{Tr}(XX^T) - 2\text{Tr}(VS^TK) + \text{Tr}(VS^TKSV^T) \\ &= \text{Tr}(K) - 2\text{Tr}(VS^TK) + \text{Tr}(VS^TKSV^T) \\ &\text{s. t. } S \geq 0, Z \geq 0. \end{aligned} \quad (3)$$

其中  $\text{Tr}(\cdot)$  表示矩阵的迹, 文献[6]给出了迭代的更新方法获得该目标函数的优化解:

$$s_j \leftarrow s_j \frac{(KV)_j}{(KSV^TV)_j}, v_j \leftarrow v_j \frac{(KS)_j}{(VS^TKS)_j}, \quad (4)$$

其中  $K = X^T X$ , 为了解决这个优化问题, 需要在原始空间中计算内积, 可以引入核方法的思想.

## 2 高光谱数据基于矩阵分解的特征提取

在高光谱数据实际处理中, 通常存在部分标记样本, 因此希望能够利用这些有价值的先验信息. 本文考虑标记数据的类别信息, 并将其作为硬性约束加入到概念分解理论中; 同时根据高光谱数据几何分布的特点, 通过构建数据的局部 graph, 挖掘其流形结构并保持局部不变. 提出基于矩阵分解的高光谱数据特征提取 (Hyperspectral data Feature Extraction on Matrix Factorization, FEMF) 方法.

### 2.1 高光谱数据标记信息的硬性约束

考虑标记样本的类别信息, 并将其作为一个硬性约束, 该思想能够保证具有相同类别属性的标记数据, 投影到低维空间中的同一位置.

给定  $M$  个光谱波段的数据集  $X \in \mathbb{R}^{M \times N}$ ,  $N$  为数据点的个数. 假设前  $l$  个数据点  $x_1, \dots, x_l$  具有标记信息, 剩下  $N-l$  个数据点为无标记样本. 该数据集中有  $c$  个类别, 每一个标记数据归属于其中的某个类.

首先, 对于标记数据, 构建  $l \times c$  维的指示矩阵  $C$ , 当且仅当数据点  $x_i$  属于第  $j$  类的时候,  $C_{ij} = 1$ , 否则为 0. 由该指示矩阵, 建立约束矩阵  $A$  如下<sup>[11]</sup>:

$$A = \begin{pmatrix} C_{l \times c} & 0 \\ 0 & I_{N-l} \end{pmatrix}, \quad (5)$$

其中  $I_{N-l}$  为  $(N-l) \times (N-l)$  维的单位矩阵. 列举一

例,假设在标记样本中, $x_1$  属于第一类, $x_2, x_3$  属于第二类, $x_4$  属于第三类,其余的  $N-4$  个数据点是无标记的,此时的约束矩阵  $A$  为:

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & I_{N-4} \end{pmatrix}, \quad (6)$$

然后将约束矩阵  $A$  通过一个辅助矩阵  $Z$  引入到降维后的数据矩阵  $V$  中:

$$V = AZ \quad (7)$$

从上面的约束方程可以看出,如果  $x_i$  和  $x_j$  为同类别的标记样本,则  $v_i = v_j$ . 有了这个标记约束,结合式(2),本文方法可以通过寻找两个矩阵因子  $S$  和  $Z$ , 获得一个半监督学习的有效模式. 此时,原始数据矩阵  $X$  的重构问题就转化为:

$$X \cong XS(AZ)^T \quad (8)$$

## 2.2 高光谱数据的流形正则化

真实高光谱数据可以看作嵌入在高维空间中的一个低维流形<sup>[9]</sup>, 希望在光谱特征的提取过程中充分挖掘数据的几何结构, 一个自然的假设是, 在原高维光谱特征空间中, 如果数据点  $x_i$  和  $x_j$  距离很近, 那么在经过特征提取后的低维空间中, 其表示  $y_i$  和  $y_j$  也彼此相近. 该假设被称为局部一致性<sup>[14]</sup>, 近年来在模式识别领域产生了重要的影响.

设  $\Omega(\cdot)$  为一个投影函数, 将原始高维光谱空间中的数据点  $x_i$  映射到轴  $u_r$  上, 有  $\Omega_r(x_i) = v_{ir}$ , 使用  $\|\Omega_r\|_M^2$  来测量  $\Omega_r$  沿着数据内在几何结构的测地线方向的平滑度. 考虑一个例子, 该组数据是一个紧致的子流形  $M \in \mathbb{R}^m$ , 对于  $\|\Omega_r\|_M^2$  一个自然的定义为:

$$\|\Omega_r\|_M^2 = \int_{x \in M} \|\nabla_M \Omega_r\|^2 dP_X(x) \quad (9)$$

上式用来计算在分布概率密度函数  $P_X$  上的积分, 其中,  $\nabla_M$  是  $\Omega_r$  沿着流形测地线方向的梯度. 在实际处理中, 通过在数据点上构建一个最近邻 graph 结构来离散地近似表示  $\|\Omega_r\|_M^2$ .

给定一个拥有  $N$  个节点的 graph, 其上的每个节点对应一个数据点, 节点之间通过权值进行连接, 则可以定义权值矩阵  $W$  为:

$$W_{ij} = \begin{cases} 1 & \text{if } x_j \in N_i \text{ or if } x_i \in N_j \\ 0 & \text{otherwise} \end{cases}, \quad (10)$$

其中,  $N_i$  表示  $x_i$  的  $K$  个最近邻点集, 权值矩阵  $W$  是一个对称矩阵,  $W_{ij}$  可以衡量两个样本点  $x_i$  和  $x_j$  的相近程度. 定义  $D$  为对角线矩阵, 其中每一个值为  $W$

的列和(或者行和, 因为  $W$  为对称矩阵).  $L = D - W$  被称为 graph Laplacian<sup>[15]</sup>. 这样,  $\|\Omega_r\|_M^2$  可以由下面的式子近似得到:

$$\begin{aligned} \mathcal{R}_r &= \frac{1}{2} \sum_{i,j=1}^N (\Omega_r(x_i) - \Omega_r(x_j))^2 W_{ij} \\ &= \sum_{i=1}^N \Omega_r(x_i)^2 D_{ii} - \sum_{i,j=1}^N \Omega_r(x_i) \Omega_r(x_j) W_{ij} \\ &= \sum_{i=1}^N v_{ir}^2 D_{ii} - \sum_{i,j=1}^N v_{ir} v_{jr} W_{ij} \quad (11) \\ &= v_r^T L v_r \end{aligned}$$

$$\begin{aligned} \mathcal{R} &= \sum_{r=1}^R \mathcal{R}_r = \text{Tr}(V^T L V) \\ &= \text{Tr}((AZ)^T L (AZ)) = \text{Tr}(Z^T A^T L A Z) \end{aligned}$$

$\mathcal{R}$  可以测量映射函数  $\Omega$  沿着数据内在几何结构的测地线方向的平滑度. 通过最小化目标  $\mathcal{R}$ , 可以使得原始高维光谱空间中距离相近的数据点  $x_i$  和  $x_j$  的低维表达  $\Omega(x_i)$  和  $\Omega(x_j)$ , 彼此仍然相近.

## 2.3 矩阵重构实现光谱的特征提取

通过上面分析, 本文需要最小化目标函数:

$$\begin{aligned} \Phi &= \|X - XSV^T\|^2 \\ &= \|X - XS(AZ)^T\|^2 + \mu \mathcal{R} \\ &= \text{Tr}(XX^T) - 2\text{Tr}(AZS^T K) \\ &\quad + \text{Tr}(AZS^T K S Z^T A^T) + \mu \text{Tr}(Z^T A^T L A Z) \quad (12) \\ &= \text{Tr}(K) - 2\text{Tr}(AZS^T K) \\ &\quad + \text{Tr}(AZS^T K S Z^T A^T) + \mu \text{Tr}(Z^T A^T L A Z) \end{aligned}$$

s. t.  $S \geq 0, Z \geq 0$ .

其中  $K = X^T X$  为数据矩阵的内积, 因此能够引入核函数. 由于该目标函数单独对  $S$  或  $Z$  是凸函数, 而对两个变量同时存在时则是非凸函数, 难以获得一个理想的全局最优解. 因此, 本文给出一个迭代的乘性更新规则得到该目标函数的局部优化解.

因为  $S \geq 0, Z \geq 0$ , 这两个约束的拉格朗日乘子定义为  $\alpha = [\alpha_{ij}]$  和  $\beta = [\beta_{ij}]$ , 则拉格朗日函数为<sup>[16]</sup>:

$$\begin{aligned} \ell(S, Z) &= \text{Tr}(K) - 2\text{Tr}(AZS^T K) + \text{Tr}(AZS^T K S Z^T A^T) \\ &\quad + \mu \text{Tr}(Z^T A^T L A Z) + \text{Tr}(\alpha S^T) + \text{Tr}(\beta Z^T) \quad (13) \end{aligned}$$

通过设定:

$$\frac{\partial \ell(S, Z)}{\partial S} = 0, \quad \frac{\partial \ell(S, Z)}{\partial Z} = 0 \quad (14)$$

从而得到:

$$\begin{aligned} \alpha &= 2KAZ - 2KSZ^T A^T A Z \quad (15) \\ \beta &= 2A^T K S - 2S^T K S Z^T A^T A - 2\mu A^T L A Z \end{aligned}$$

根据 KKT 条件<sup>[16]</sup>, 则有  $\alpha_{ij} u_{ij} = 0$  和  $\beta_{ij} z_{ij} = 0$ , 从而推导出矩阵重构的更新规则:

$$\begin{cases} s_{ij} \leftarrow s_{ij} \frac{(KAZ)_{ij}}{(KSZ^T A^T AZ)_{ij}} \\ z_{ij} \leftarrow z_{ij} \frac{(A^T KS)_{ij}}{(S^T KSZ^T A^T A + \mu A^T LAZ)_{ij}} \end{cases} \quad (16)$$

因为该更新规则是非增的,且目标函数有下界0,所以能够实现收敛.对于该类问题初始值的选取,可以参考文献[3]和[7],然后通过式(16)进行迭代更新实现收敛,得到一组局部优化解.在迭代的过程中,通过最小化目标函数(12),引入矩阵 $L$ 和 $A$ ,将局部一致性思想和标记数据的先验信息嵌入特征提取的过程.最终保证原始高维特征空间中相近的数据点在低维空间中仍然彼此相近,而相同类别的标记样本则被投影到同一个位置,从而使高光谱数据的空间结构更加优化.

#### 2.4 方法计算复杂度分析

这一部分比较 FEMF 与标准 NMF 方法的计算复杂度.在每一步迭代的更新过程中,NMF 和 FEMF 的计算复杂度如表 1 所示.考虑  $K$  为一个较小的常数,FEMF 需要  $O(N^2M)$  次运算来构建原始高维数据的局部邻域结构,假设经过  $t$  次迭代可以收敛,得到优化结果,则标准 NMF 方法总的计算复杂度为:

$$O(tMNR)$$

本文所提 FEMF 方法的总的计算复杂度为:

$$O(tMNR + N^2M)$$

表 1 每次迭代 NMF 和 FEMF 的计算复杂度

Table 1 Computational complexity of NMF and FEMF in each iteration

	NMF	FEMF
浮点加	$2MNR + 2(M+N)R^2$	$2MNR + 2(M+N)R^2 + N(K+3)R$
浮点乘	$2MNR + 2(M+N)R^2 + (M+N)R$	$2MNR + 2(M+N)R^2 + (M+N)R + N(K+1)R$
浮点除	$(M+N)R$	$(M+N)R$
总计	$O(MNR)$	$O(MNR)$

### 3 高光谱数据实验结果与分析

#### 3.1 实验数据说明

采用两组 224 波段 AVIRIS 高光谱遥感数据进行试验,对不同的特征提取算法进行比较.这两组数据的光谱波段范围为  $0.4 \mu\text{m} : 2.5 \mu\text{m}$  之间,其空间分辨率达到  $3.7 \text{ m}$ .具有完备的地理信息.

##### 3.1.1 数据集一

图 1 所示数据取自美国 Indian Pine 测试区,实验去除受水汽和臭氧影响比较严重的光谱波段,保留 180 个信噪比较高的波段构成数据集.



图 1 原始 Indian Pine 高光谱数据  
Fig. 1 Original Indian Pine hyperspectral data

该组数据集中共有 16 类地物目标,试验选取样本数较多的 10 个类别作为样本集,所选取的样本集和地物分类信息分别如图 2 所示.

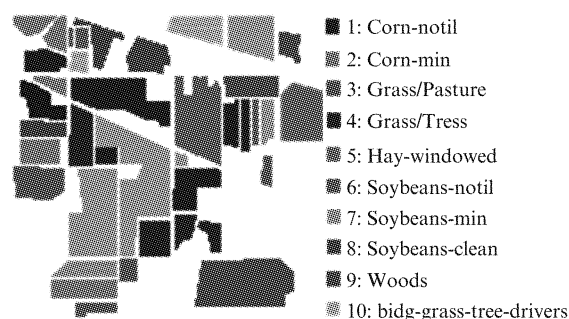


图 2 Indian Pine 地物分类信息  
Fig. 2 Indian Pine ground truth

##### 3.1.2 数据集二

图 3 所示数据取自美国 Salinas Valley 测试区.实验保留 200 个信噪比较高的光谱波段构成数据集.

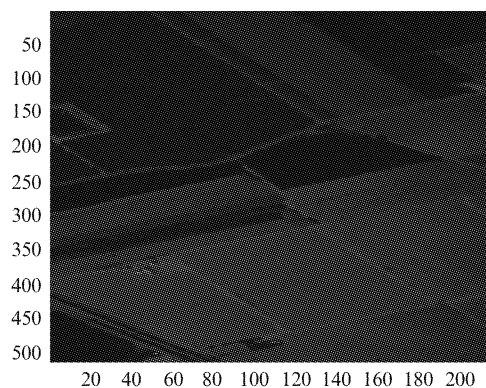


图 3 原始 Salinas Valley 高光谱数据  
Fig. 3 Original Salinas Valley hyperspectral data

该组数据集共有 111 104 ( $512 \times 217$ ) 个数据点,分别归属 16 类地物,地物分类真实信息如图 4 所示. 试验选取具有代表性的 8 个类别作为样本测试集,所选取的样本集如表 2 所示.

表 2 Salinas Valley 样本集

Table 2 Salinas Valley sample set

category	sample size	category	sample size
1	2 099	12	1 927
3	1 976	13	916
4	1 394	14	1 070
11	1 068	16	1 807

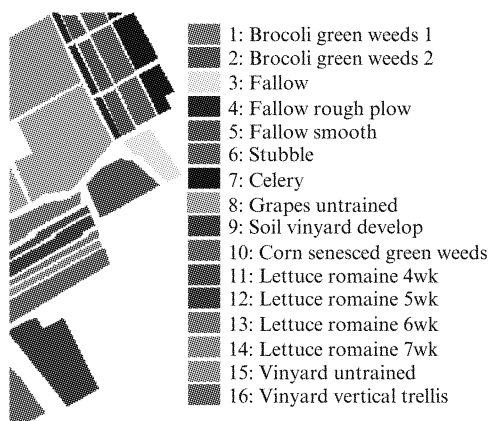


图 4 Salinas Valley 地物分类信息  
Fig. 4 Salinas Valley ground truth

### 3.2 实验参数选择

在验证本文方法的实验过程中,随机地选取表 2 和 3 中每类样本的 30% 作为标记样本,剩下的作为无标记样本,通过矩阵分解得到高光谱数据特征提取后的低维表示.

该方法需要确定每个样本数据点  $x_i$  局部邻域  $N_i$  内的样本个数  $K$ . 局部邻域内的样本点范围通常由欧式距离和向量角余弦值来确定,本文采用向量角余弦值来确定局部邻域内的样本点. 实验随机选取每类样本的 15% 作为训练样本集,其余样本作为测试集,在提取相同特征个数的前提下,采用 KNN 分类器对矩阵分解后的数据进行分类. 实验结果如图 5 所示: 当  $K$  值较小和较大时,分类精度均有所下降.

### 3.3 各种光谱特征提取方法的性能比较

图 6 为 Salinas Valley 数据光谱波段之间的相关系数矩阵图,其亮度大小反映了光谱波段间的相关程度,白色表示最强,黑色表示最弱. 可以很清楚地看出相关性系数矩阵沿着对角线方向呈块状分布,且对角线附近的亮度最高,这意味着高光谱数据相

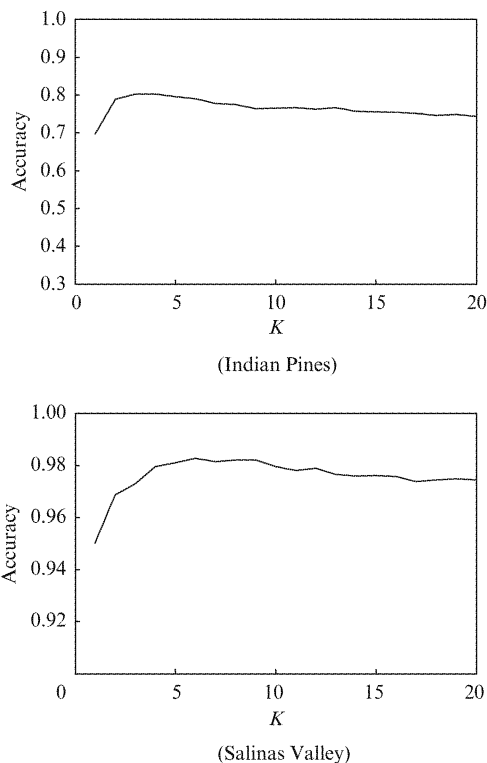


图 5  $K$  值的影响  
Fig. 5 Classification accuracy under different  $K$

邻光谱之间存在很强的相关性,因此提取光谱波段的主要特征实现数据的低维表示很有必要.

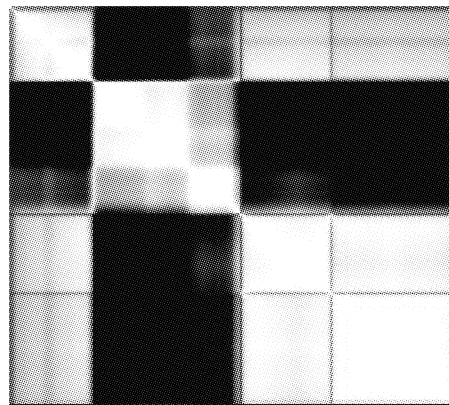


图 6 Salinas Valley 数据光谱波段的相关系数矩阵(白色为 1 或 -1;黑色为 0)  
Fig. 6 Spectral bands' correlation coefficient matrix of Salinas Valley dataset (white = 1 or -1; black = 0).

为了验证本文方法的有效性,通过对高光谱数据提取不同数目的特征,采用 KNN 分类器得到数据经过降维后的分类精度,并与其它特征提取方法, NMF、Constrained Nonnegative Matrix Factorization (CNMF)<sup>[11]</sup>、Discriminative Concept Factorization

(DCF)<sup>[6]</sup> 和 LocallyConsistent Concept Factorization (LCCF)<sup>[7]</sup> 进行性能比较. 实验结果如图 7 所示:

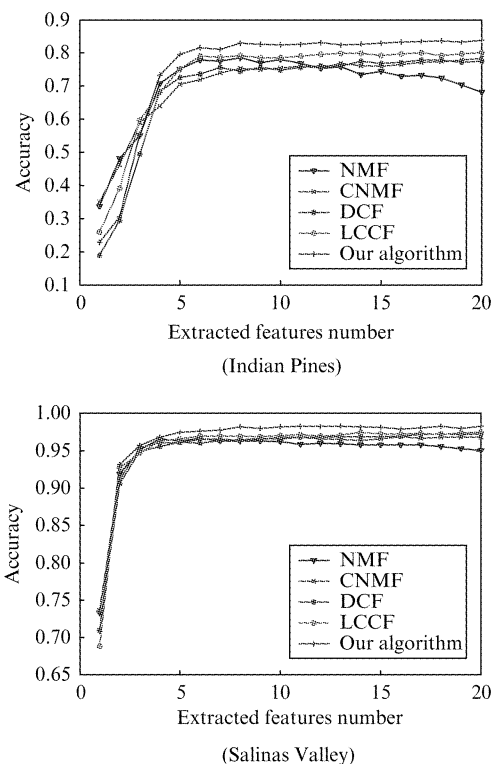


图 7 各种方法的性能比较

Fig. 7 The performance comparison of different methods

从分类结果可以看出,在提取较少特征数时,各种特征提取方法的分类效果均不理想. 随着特征个数的增加,各种方法的性能大为增强,分类精度有了显著提高. 与其他方法相比,本文方法在高光谱数据实际处理中,大量无标记数据和少量有标记数据并存的情况下,可以利用宝贵的标记数据提升特征提取模型的泛化性能;方法基于数据分布的实际结构而设计,嵌入一种具有判别性质的思想,能够提取到高光谱数据的主要特征,从而优化数据的空间结构,有利于识别和解译.

#### 4 结论

利用高光谱数据有限的标记样本,将其作为硬性约束加入标准的概念分解方法中;同时构建局部邻域 graph 结构来描述数据分布的几何结构,提出一种基于矩阵分解的高光谱数据特征提取 (FEMF) 方法. 经过矩阵分解,使得原始高维光谱特征空间中相近的数据点在低维空间中仍然彼此相近,而相同

类别的标记数据点则被投影到同一个位置,最终获得的低维数据具有更好的空间分布结构. 最后,对真实高光谱数据进行特征提取验证了该方法的优越性.

#### References

- [1] Jia X, Kuo B C, Crawford M M. Feature mining for hyperspectral image classification[J]. *Proceedings of the IEEE*, 2013, **101**(3): 676–697.
- [2] Plaza A, Benediktsson J A, Boardman J W, et al. Recent advances in techniques for hyperspectral image processing [J]. *Remote Sensing of Environment*, 2009, **113**: S110–S122.
- [3] Cai D, He X, Han J, et al. Graph regularized nonnegative matrix factorization for data representation[J]. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 2011, **33**(8): 1548–1560.
- [4] Lee D D, Seung H S. Learning the parts of objects by non-negative matrix factorization [J]. *Nature*, 1999, **401**(6755): 788–791.
- [5] Xu W, Gong Y. Document clustering by concept factorization[C]. In Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval. ACM, 2004: 202–209.
- [6] Hua W, He X. Discriminative concept factorization for data representation[J]. *Neurocomputing*, 2011, **74**(18): 3800–3807.
- [7] Cai D, He X, Han J. Locally consistent concept factorization for document clustering[J]. *Knowledge and Data Engineering, IEEE Transactions on*, 2011, **23**(6): 902–913.
- [8] Camps-Valls G, BandosMarshveva T, Zhou D. Semi-supervised graph-based hyperspectral image classification [J]. *Geoscience and Remote Sensing, IEEE Transactions on*, 2007, **45**(10): 3044–3054.
- [9] 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): 441–454.
- [10] He X, Ji M, Bao H. Graph Embedding with Constraints [C]. In IJCAI. 2009, **9**: 1065–1070.
- [11] Liu H, Wu Z, Li X, et al. Constrained nonnegative matrix factorization for image representation[J]. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 2012, **34**(7): 1299–1311.
- [12] Chung F R K. *Spectral graph theory*[M]. AMS Bookstore, 1997.
- [13] Belkin M, Niyogi P. Laplacian eigenmaps for dimensionality reduction and data representation[J]. *Neural computation*, 2003, **15**(6): 1373–1396.
- [14] He X, Yan S, Hu Y, et al. Face recognition using laplacian faces[J]. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 2005, **27**(3): 328–340.
- [15] Spielman D. Spectral graph theory [J]. *Lecture Notes, Yale University*, 2009: 740–0776.
- [16] Candès E J, Recht B. Exact matrix completion via convex optimization[J]. *Foundations of Computational mathematics*, 2009, **9**(6): 717–772.