基于局部保持线性判别嵌入特征提取的光谱图像分类

温金环[1]，田 铮[1]，赵永强[2]，延伟东[1]

(1. 西北工业大学 理学院应用数学系，陕西 西安 710072; 2. 西北工业大学 自动化学院，陕西 西安 710072)

摘 要: 针对难于获得足够多的高光谱图像训练样本的问题，基于流形学习标准、Fisher 标准和最大边缘标准，提出了一种适用于高光谱图像小样本问题的局部保持线性判别嵌入(LPLDE)监督线性流形学习特征提取方法。LPLDE 方法利用类内近邻图和类间近邻图描述类内的紧性和类间的可分性，有效地避免了因类内离散度矩阵奇异导致的小样本问题，具有更好的判别性能，更适合于分类问题。高光谱数据的实验结果表明了该方法的有效性。

关 键 词: 特征提取，降维，流形学习，小样本问题，高光谱图像分类
中图分类号: TP751.1 **文献标识码**: A **文章编号**: 1000-2758(2013)02-0323-06

高光谱影像往往有着数十甚至数百的波段，且地物类别丰富，这使得训练样本点的选取变得困难。研究者已经提出了许多方法以减轻高维和小样本问题(Small Sample Size，SSS)对高光谱数据信息提取的影响，如非参权特征提取[1] (Nonparametric Weighted Feature Extraction，NWFE)及其改进算法[2,3]等。

近年来兴起并在模式识别领域得到广泛应用的流形学习技术，如局部线性嵌入(Locally Linear Embedding，LLE)[4]、等度规划映射(Isometric Mapping，Isomap)[5]和拉普拉斯特征映射(Laplacian Eigenmaps，LE)[6]等已在高光谱数据上得到了初步应用[7-10]。由于流形学习算法对计算机的存储要求较高，使得流形学习方法难于处理 70×70 以上的高光谱图像[7-10]。流形学习方法适合描述非线性数据，但仅在训练集上产生映射，如何把一新的测试采样投影到低维空间并不清楚，因此，这些方法不能简单地应用到分类问题上。为了克服此缺点，一些线性流形学习方法，如局部保持投影(Locally Preserving Projection，LPP)[11]和近邻保持嵌入(Neighborhood Preserving Embedding，NPE)[12]被提出直接处理新的测试采样。但由于它们的目的主要是保留局部性(或相似性)，这些算法宜用于恢复或聚类而不是分类。仅有少数流形学习方法考虑类内和类间信息来分类，如局部判别嵌入(Local Discriminant Embedding，LDE)[13]，边缘 Fisher 分析(Marginal Fisher Analysis，MFA)[14]，但这些方法为了避免小采样问题，均先用 PCA 降维，这样可能丢掉一些重要的判别信息。

基于以上分析，本文基于流形学习标准、Fisher 标准和最大边缘标准(Maximum Margin Criterion，MMC)[15]，提出了一种适用于高光谱图像小样本问题的局部保持线性判别嵌入(Locally Preserving Linear Discriminant Embedding，LPLDE)监督线性流形学习特征提取算法，并结合 NN 分类器对高光谱图像分类。该方法利用类内近邻图和类间近邻图描述类内的紧性和类间的可分性，有效地避免了因类内离散度矩阵奇异导致的小样本问题，具有更好的判别性能，更适合于分类问题。实验表明该方法有效降低了高光谱数据的维数，提高了分类精度，尤其在病态分类情形下也能取得较高的分类精度。

1 局部保持线性判别嵌入方法

局部保持线性判别嵌入(LPLDE)方法主要有如下 3 步:

收稿日期: 2012-05-26 **基金项目**: 国家自然科学基金(61201323、11202161)与西北工业大学基础研究基金(JC201053)资助
作者简介: 温金环(1974—)，女，西北工业大学讲师，主要从事流形学习、非负矩阵分解及高光谱图像处理研究。

1) 构造类内近邻图 G 和类间近邻图 G': 设 G 和 G' 分别表示包含 m 个顶点的图,第 i 个节点对应采样 $\mathbf{x}_i$。为了构造类内近邻图 G,考虑来自同一类的采样 $\mathbf{x}_i$ 和 $\mathbf{x}_j$,若 $\mathbf{x}_i$ 是 $\mathbf{x}_j$ 的 k 最近邻或者 $\mathbf{x}_j$ 是 $\mathbf{x}_i$ 的 k_1 最近邻,则在 $\mathbf{x}_i$ 与 $\mathbf{x}_j$ 间连接一条边。对于类间近邻图 G',则考虑来自不同类的采样 $\mathbf{x}_i$ 和 $\mathbf{x}_j$,同样,如果 $\mathbf{x}_i$ 是 $\mathbf{x}_j$ 的 k_2 近邻或者 $\mathbf{x}_j$ 是 $\mathbf{x}_i$ 的 k_2 近邻,则在 $\mathbf{x}_i$ 与 $\mathbf{x}_j$ 间连接一条边。

2) 选择类内近邻图 G 和类间近邻图 G' 的权重: 为了保留类内近邻几何,假设每一个采样 $\mathbf{x}_j^i$($\mathbf{x}_j^i$ 表示第 i 类的第 j 个样本) 可以由第 i 类的其它样本线性重构。令 $\mathbf{W}_{w,ij}$ 表示图 G 类内亲和权,则 $\mathbf{W}_{w,ij}$ 可通过最小化下面的重构误差得到

$$\min \sum_i \left\| \mathbf{x}_i - \sum_{j \in N_{k_1}(i)} \mathbf{W}_{w,ij} \mathbf{x}_j \right\|^2 \tag{1}$$

且满足下面 2 个约束:

a) 如果图 G 的 2 个节点 i 与 j 间无边,则 $\mathbf{W}_{w,ij} = 0$;

b) $\sum_j \mathbf{W}_{ij} = 1, j = 1,2,\cdots,k$,即 $\mathbf{x}_i$ 可以表示为其近邻的凸组合。这里,集合 $N_{k_1}(i)$ 由与 $\mathbf{x}_i$ 同类的 k_1 个最近邻的指标组成。

注: 这里类内亲和权重矩阵 $\mathbf{W}_{w,ij}$ 的选择和 *LLE* 算法中的权重矩阵 W 的选择不同,*LLE* 算法中的权重矩阵 W 的选择没有考虑样本的类标号信息,即这些近邻点也许属于同类,也许属于不同类。这里在选择类内亲和权重矩阵 $\mathbf{W}_{w,ij}$ 时考虑了样本的类标号信息,即点 $\mathbf{x}_i$ 的 k_1 个最近邻仅由与其同类的 k_1 个最近邻点组成。

设 $\mathbf{W}_{b,ij}$ 表示类间近邻图 G' 的类间亲和权,可以定义 $\mathbf{W}_{b,ij}$ 为热核或简单的盲源方式[3]。为简单起见,本文定义类间亲和权 $\mathbf{W}_{b,ij}$ 为简单的盲源方式,即

$$\mathbf{W}_{ij} = \begin{cases} 1 & \mathbf{x}_i \text{ 与 } \mathbf{x}_j \text{ 相连} \\ 0 & \text{否则} \end{cases} \tag{2}$$

3) 计算嵌入向量: 找下面特征值问题的前 d 个最大特征值所对应的特征向量 $\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_d$

$$\mathbf{X}(\mathbf{L}_b - \mathbf{M}_w)\mathbf{X}^T\mathbf{a}_k = \lambda_k \mathbf{a}_k \tag{3}$$

式中 $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_m]$, $\mathbf{M}_w = (\mathbf{I} - \mathbf{W}_w)^T(\mathbf{I} - \mathbf{W}_w)$, $\mathbf{I} = diag(1,1,\cdots,1)$, $\mathbf{L}_b = \mathbf{D}_b - \mathbf{W}_b$ 是图 G' 的 *Laplacian* 矩阵,$\mathbf{D}_b$ 是对角矩阵,其对角线元素为类间亲和权矩阵 $\mathbf{W}_b$ 的列和或行和,即 $\mathbf{D}_{b,ii} = \sum_j \mathbf{W}_{b,ij}$。

下面给出局部保持线性判别嵌入(*LPLDE*) 方法的详细推导过程。

为了使投影向量保持局部几何,类似于 NPE 算法,最小化下面的代价函数

$$\begin{aligned} \Phi_1(\mathbf{A}) &= \sum_i \left\| \mathbf{y}_i - \sum_j \mathbf{W}_{w,ij}\mathbf{y}_j \right\|_2^2 = \| \mathbf{Y}(\mathbf{I} - \mathbf{W}_w) \|_F^2 \\ &= Tr(\mathbf{Y}(\mathbf{I} - \mathbf{W}_w)^T(\mathbf{I} - \mathbf{W}_w)\mathbf{Y}^T) \\ &= Tr(\mathbf{A}^T\mathbf{X}\mathbf{M}_w\mathbf{X}^T\mathbf{A}) \end{aligned} \tag{4}$$

式中 "$\| \cdot \|_F$" 表示矩阵的 Frobenius 范数,"Tr" 表示矩阵的迹,$\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_m]$, $\mathbf{y}_i = \mathbf{A}^T\mathbf{x}_i$, $\mathbf{y}_i \in \mathbf{R}^d$($d \ll m$) 是 $\mathbf{x}_i$ 的低维 "表示", $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_d] \in \mathbf{R}^{n \times d}$ 是变换矩阵。

为了使不同类的投影向量远离,类似于 *LPP* 算法,最大化下面的代价函数

$$\begin{aligned} \Phi_2(\mathbf{A}) &= \sum_{ij} \| \mathbf{y}_i - \mathbf{y}_j \|_2^2 \mathbf{W}_{b,ij} \\ &= \sum_{ij} \| \mathbf{A}^T\mathbf{x}_i - \mathbf{A}^T\mathbf{x}_j \|_2^2 \mathbf{W}_{b,ij} \\ &= 2Tr(\mathbf{A}^T\mathbf{X}(\mathbf{D}_b - \mathbf{W}_b)\mathbf{X}^T\mathbf{A}) \propto Tr(\mathbf{A}^T\mathbf{X}\mathbf{L}_b\mathbf{X}^T\mathbf{A}) \end{aligned} \tag{5}$$

为了找到一个既能保持数据局部几何又能将不同类数据区分得更开的线性子空间,意味着不仅要最大化 $\Phi_2(\mathbf{A})$ 而且要最小化 $\Phi_1(\mathbf{A})$,即最大化下式

$$\Phi_{LPLDE}(\mathbf{A}_{opt}) = \underset{A}{argmax} \frac{Tr(\mathbf{A}^T\mathbf{S}_{BN}\mathbf{A})}{Tr(\mathbf{A}^T\mathbf{S}_{WN}\mathbf{A})} \tag{6}$$

式中 $\mathbf{S}_{BN} = \mathbf{X}\mathbf{L}_b\mathbf{X}^T$ 是类间近邻散度矩阵,$\mathbf{S}_{WN} = \mathbf{X}\mathbf{M}_w\mathbf{X}^T$ 是类内近邻散度矩阵。由于小样本问题,$n \gg m$,由 $rank(\mathbf{X}) \leqslant m$, $rank(\mathbf{X}^T) \leqslant m$, $rank(\mathbf{M}) \leqslant m$ 可知 $rank(\mathbf{S}_{WN}) \leqslant m$。矩阵 $\mathbf{S}_{WN}$ 是 $n \times n$ 的矩阵意味着在大多数情况下 $\mathbf{S}_{WN}$ 是奇异矩阵,因此利用公式(6) 不能计算投影变换矩阵 A。

最近,文献[15] 提出的 MMC 算法可以解决因类内散度矩阵奇异导致的 SSS 问题。利用 MMC 算法,局部保持线性判别嵌入(LPLDE) 方法的目标函数可转换为如下形式

$$J_{LPLDE}(\mathbf{A}_{opt}) = \underset{A}{argmax} Tr(\mathbf{A}^T(\mathbf{S}_{BN} - \mathbf{S}_{WN})\mathbf{A}) \tag{7}$$

为了解此优化问题,对 A 加一附加约束,即要求 A 是正交矩阵,这样可得到下面的约束优化问题

$$\begin{aligned} J_{LPLDE}(\mathbf{A}_{opt}) &= argmax \sum_{k=1}^{d} \mathbf{a}_k^T(\mathbf{S}_{BN} - \mathbf{S}_{WN})\mathbf{a}_k \\ s.t. \quad & \mathbf{a}_k^T\mathbf{a}_k - 1 = 0 \end{aligned} \tag{8}$$

上述约束优化问题的解可通过 Lagrange 乘数法求得。方程(8) 所对应的 Lagrange 函数为

$$L(\mathbf{a}_k,\lambda_k) = \sum_{k=1}^{d}\mathbf{a}_k^T(\mathbf{S}_{BN}-\mathbf{S}_{WN})\mathbf{a}_k - \lambda_k(\mathbf{a}_k^T\mathbf{a}_k - 1) \tag{9}$$

令 $\frac{\partial L(\ ,\lambda_k)}{\partial \mathbf{a}_k} = 0$，可得

$$(\mathbf{S}_{BN}-\mathbf{S}_{WN})\mathbf{a}_k = \lambda_k\mathbf{a}_k \quad k = 1,2,\cdots,d \tag{10}$$

式中 λ_k 是矩阵 $\mathbf{S}_{BN}-\mathbf{S}_{WN}$ 的特征值，即矩阵 $\mathbf{X}(\mathbf{L}_b-\mathbf{M}_w)\mathbf{X}^T$ 的特征值，$\mathbf{a}_k$ 是 λ_k 所对应的特征向量。利用(10)式，代价函数(8)变为

$$\begin{aligned} J_{LPLDE}(\mathbf{A}_{opt}) &= argmax\sum_{k=1}^{d}\mathbf{a}_k^T(\mathbf{S}_{BN}-\mathbf{S}_{WN})\mathbf{a}_k \\ &= argmax\sum_{k=1}^{d}\mathbf{a}_k^T\lambda_k\mathbf{a}_k \\ &= argmax\sum_{k=1}^{d}\lambda_k \end{aligned} \tag{11}$$

由(11)式可以看出，如果 A 的列是由矩阵 $\mathbf{X}(\mathbf{L}_b-\mathbf{M}_w)\mathbf{X}^T$ 的前 d 个最大特征值所对应的特征向量组成的，则目标函数 $J_{LPLDE}(\mathbf{A}_{opt})$ 达到最大化。这里，不需要计算类内近邻散度矩阵 $\mathbf{S}_{WN}$ 的逆矩阵，简单有效地避免了小样本问题。

2　实验结果与分析

为了验证 LPLDE 算法的性能，本文通过 AVIRIS 和 Washington DC Mall 2 个高光谱遥感数据的分类实验比较了 LPLDE 和其它著名的特征提取方法，如 PCA、LDA、LPP、NPE 和 MFA 的性能。

2.1　数据集

AVIRIS 高光谱图像采用航空可见光、红外成像光谱仪 AVIRIS 于 1992 年 6 月所测高光谱数据 92AV3C。该数据采集自美国印第安纳州北部 Indian Pine 测试区农场和森林交杂的地带，有 220 个光谱带，光谱范围 0.4 ~ 2.45μm，空间分辨率为 20 m。实验时只考虑了 Corn-min、Corn-notill、Grass/Pasture、Grass/Tree、Hay-windrowed、Soybeans-min、Soybeans-clean、Soybeans-notill、Woods 和 Wheat 这 10 类地物。

Washington DC Mall 高光谱遥感数据是美国华盛顿地区 HYDICE 航空高光谱数据。该数据的获取时间为 1995 年 8 月 23 日，有 220 个光谱带，光谱范围 0.4 ~ 2.4μm，空间分辨率为 4 m。图像有 280 × 307 共 85 960 个像元。实验中去除了水蒸气吸收波段，剩余 191 个波段用于分类。该数据的背景主要有 7 个地物类别: 屋顶(Roofs)、道沂(Road)、小径(Trail)、草地(Grass)、树木(Tree)、水(Water)和阴影(Shadow)。

本文设计了 2 组实验:

1) 第 1 组实验是为了探索训练样本数对分类准确率的影响。令 N_i 表示每类的训练样本数，N 表示总的训练样本数，m 表示高光谱数据的波段数，类似于文献[13]，考虑如下 3 种情形的分类问题: $N_i = 20 < N < m$(情形 1)、$N_i = 40 < m < N$(情形 2)和 $N_i = 100 < m < N$(情形 3)。对于 Washington，DC Mall，实验时每类随机选择 20、40 和 100 个标号样本作为训练样本，100 个标号样本作为测试样本。对于 AVIRIS 数据，每类随机选取 20、40 和 100 个标号样本作为训练数据，剩下的作为测试数据。

2) 第 2 组实验是为了比较利用 LPLDE、PCA、LDA、LPP、NPE 和 MFA 特征组合 NN 分类器及最大似然分类器(ML)分类的分类性能。对于 AVIRIS 数据，每类随机选取标号样本的 10% 作为训练数据，剩下的 90% 作为测试数据。实验时先用不同的特征提取算法对 220 波段的 AVIRIS 数据降维，然后对投影后的低维数据利用 NN 和 ML 分类器分类。

2.2　实验结果与分析

表 1、表 2 分别给出了在情形 1 ~ 3 时测试数据的分类精度(10 次随机选择的训练样本所得结果的平均)，括号内的数据是获得最大识别率时的特征维数。图 1 ~ 图 3 是第 2 组实验的实验结果，图 1 是用不同的特征提取方法降维，用 NN 分类器分类 AVIRIS 图像的分类结果。图 2 是对降维后的 AVIRIS 图像数据利用 NN 和 ML 分类器分类的分类结果，图 3 给出了对降维后的低维 AVIRIS 图像数据利用 ML 分类器分类的分类主题图。

从表 1 ~ 表 2 和图 1 ~ 图 3 可以看出:

1) 从表 1 可以看出，当训练样本为 20 和 40 时，LPLDE 对于 AVIRIS 数据的分类明显优于其它的特征提取方法。在情形 1 ~ 3 下，先用 LPLDE 特征降维，再用 NN 分类器分类所得到的分类精度高于直接用 NN 分类器分类的结果，分类精度分别提高了 10.62%、6.07% 和 8.85%。实验结果充分证明了用 LPLDE 作为分类前的预处理非常必要。

2) 对于 Washington DC Mall 数据，在情形 1 ~ 2 时，LPLDE 特征与 NN 分类器组合的辨识率稍好于直接利用 NN 分类器分类的分类精度。当训练样本为 100 时，LPLDE 性能略差于 NPE 算法。

表1 用不同的特征提取方法降维，用NN分类器分类，AVIRIS测试数据的分类结果

方法	$N_i=20$		$N_i=40$		$N_i=100$	
	OA/%	κ	OA/%	κ	OA/%	κ
NN	63.23(185)	0.611 3	71.34(185)	0.691 3	76.18(185)	0.735 7
LDA	44.37(9)	0.411 5	71.53(8)	0.692 0	81.79(8)	0.796 8
PCA	65.75(13)	0.619 2	70.54(26)	0.678 6	76.69(33)	0.738 6
NPE	40.15(17)	0.361 2	71.97(32)	0.700 3	82.34(35)	0.803 2
MFA	64.89(24)	0.614 6	58.15(24)	0.560 1	60.57(17)	0.573 5
LPP	47.53(2)	0.430 0	53.28(8)	0.511 5	81.12(8)	0.792 3
LPLDE	73.85(14)	0.705 7	77.41(27)	0.753 4	85.03(28)	0.832 7

表2 用不同的特征提取方法降维，用NN分类器分类，Washington DC Mall测试数据的分类结果

Methods	$N_i=20$		$N_i=40$		$N_i=100$	
	OA/%	κ	OA/%	κ	OA/%	κ
NN	98.77(191)	0.985 6	98.92(191)	0.987 3	99.51(191)	0.994 2
LDA	96.46(6)	0.958 7	99.23(6)	0.991 5	99.67(6)	0.995 9
PCA	98.96(10)	0.986 3	98.49(10)	0.983 9	99.54(8)	0.994 6
NPE	98.73(21)	0.985 2	99.43(8)	0.993 8	99.89(19)	0.998 0
MFA	96.35(18)	0.954 0	94.86(18)	0.947 7	98.59(18)	0.984 6
LPP	96.48(6)	0.962 1	99.57(6)	0.994 6	99.60(6)	0.995 3
LPLDE	99.90(6)	0.998 7	99.63(8)	0.995 6	99.72(10)	0.9963

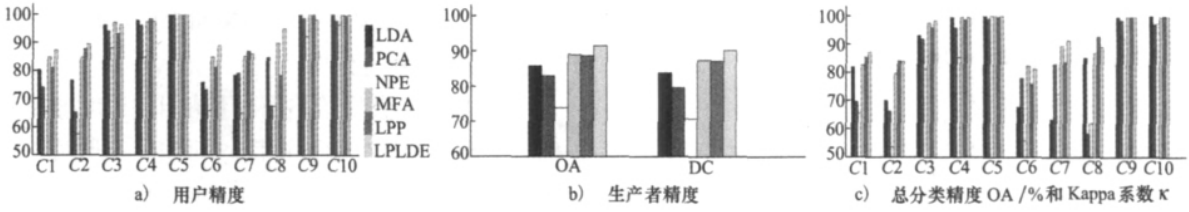


图1 用不同的特征提取方法降维，用NN分类器分类AVIRIS图像的分类结果

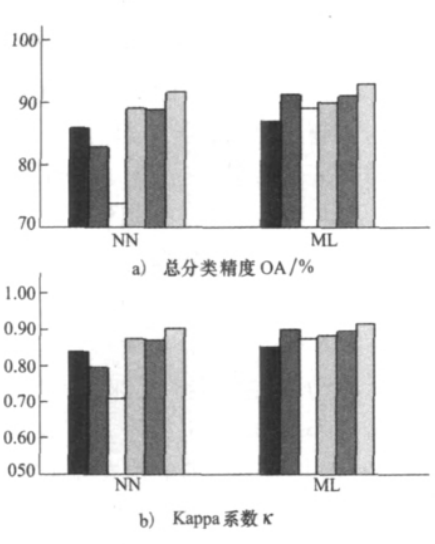


图2 用不同的特征提取方法降维，用NN和ML分类器分类AVIRIS图像的分类结果

3) 从图1a) ~图1b)可以看出LPLDE在类C1、C2、C6和C8上的用户精度要明显优于其它算法，在类C1、C3和C7上的生产者精度比其它方法高。从图1也可以看出，各种特征提取方法在C3、C4、C5、C9和C10上的分类精度较高，而对C1、C2、C6、C7和C8的分类较差，这主要是由于图片拍摄于六月份，corn和soybean正处于生长的早期，它们所反应的光谱特性极其相似，极易误分，且它们有特别的子类：notill和min。

4) 从图1c)和图2可以看出LPLDE的辨识率最高，LPLDE、PCA、LDA、LPP、NPE和MFA与ML分类器组合的辨识率要好于它们与NN分类器组合的辨识率，这说明特征提取方法与分类器的不同组合对分类精度有极大的影响，但是在实际应用中我们很难找到特征提取方法与分类器的最佳组合。尽管

"特征提取 + ML 分类器"组合所得到的分类结果要好于"特征提取 + NN 分类器"组合所得到的结果(除了"MFA + NN 分类器"组合效果差外),"特征提取 + NN 分类器"组合也取得了很好的结果。

5) 从图 3 的可视化结果来看,我们方法的误分点明显较少,尤其对于 Corn-min 和 Soybeans-min 这两个类别的错分现象明显减少。

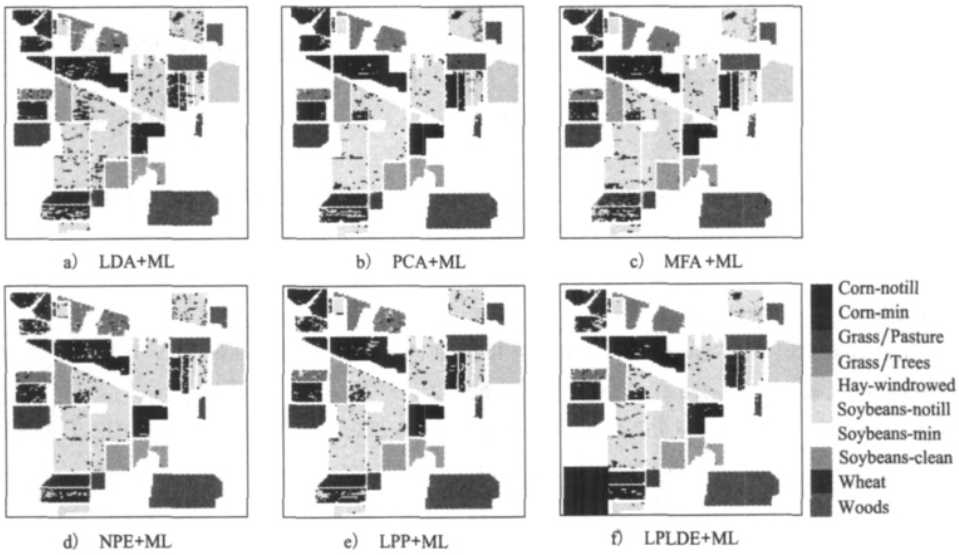


图 3　用不同的特征提取方法降维,用 ML 分类器分类 AVIRIS 图像的分类主题图

3　结　论

针对难于获得足够多高光谱数据训练样本的问题,本文以线性流形学习、Fisher 标准和最大边缘标准理论为基础,提出了一种适用于高光谱小样本问题的局部保持线性判别嵌入(LPLDE)监督线性流形学习特征提取算法,该方法利用类内近邻图和类间近邻图描述类内的紧性和类间的可分性,有效地避免了因类内离散度矩阵奇异导致的小样本问题,具有更好的判别性能,更适合于分类。实验表明,LPLDE 方法在病态分类情形下也能取得较高的分类精度。

参考文献:

[1] Kuo B C, Landgrebe D A. Nonparametric Weighted Feature Extraction for Classification. IEEE Trans on Geoscience and Remote Sensing, 2004, 42(5): 1096-1105

[2] Kuo B C, Li C H, Yang J M. Kernel Nonparametric Weighted Feature Extraction for Hyperspectral Image Classification. IEEE Trans on Geoscience and Remote Sensing, 2009, 47(4): 1139-1155

[3] Yang J M, Yu P T, Kuo B C. A Nonparametric Feature Extraction and Its Application to Nearest Neighbour Classification for Hyperspectral Image Data. IEEE Trans on Geoscience and Remote Sensing, 2010, 48(3): 1279-1293

[4] Roweis S T, Saul L K. Nonlinear Dimensionality Reduction by Locally Linear Embedding. Science, 2000, 290(5000): 2323-2326

[5] Tenenbaum J B, de Silva V, Langford J C. A Global Geometric Framework for Nonlinear Dimensionality Reduction. Science, 2000, 290(5500): 2319-2323

[6] Belkin M, Niyogi P. Laplacian Eigenmaps and Spectral Techniques for Embedding and Clustering. In Advances in Neural Information Processing System, 14, Cambridge, MA, USA, 2002, 585-591

[7] Bachmann C M, Ainsworth T L, Fusina R A. Exploiting Manifold Geometry in Hyperspectral Imagery. IEEE Trans on Geoscience and Remote Sensing, 2005, 43(3): 441-454

[8] Han T, Goodenough G. Nonlinear Feature Extraction of Hyperspectral Data Based on Locally Linear Embedding (LLE). IEEE International Geoscience and Remote Sensing Symposium, 2005, 1237-1240

[9] Qian S E, Chen G Y. A New Nonlinear Dimensionality Reduction Method with Application to Hyperspectral Image Analysis. IEEE International Geoscience and Remote Sensing Symposium, 2007, 270-273

[10] Mohan A, Sapiro G, Bosch E. Spatially Coherent Nonlinear Dimensionality Reduction and Segmentation of Hyperspectral Images. IEEE Geoscience and Remote Sensing Letters, 2007, 4(2): 206-210

[11] He X F, Niyogi P. Locality Preserving Projections. Advances in Neural Information Processing System 16, MIT Press, Cambridge, MA, 2004, 153-160

[12] He X F, Cai D. Neighborhood Preserving Embedding. Proceedings of the 10th IEEE International Conference on Computer Vision, 2005, 2: 1208-1213

[13] Chen H, Chang H, Liu T. Local Discriminant Embedding and Its Variants. Proc 2005 International Conference on Computer Vision and Pattern Recognition, 2005, 2: 846-853

[14] Yan S C, Xu D, Zhang B Y, et al. Graph Embedding: A General Framework for Dimensionality Reduction. Proceedings of Conference on Vision and Pattern Recognition, 2005, 29: 830-837

[15] Li H, Jiang T, Zhang K. Efficient and Robust Feature Extraction by Maximum Margin Criterion. IEEE Trans on Neural Networks, 2006, 17: 157-165

LPLDE (Locally Preserving Linear Discriminant Embedding) Method of Feature Extraction for Hyperspectral Image Classification

Wen Jinhuan[1], Tian Zheng[1], Zhao Yongqiang[2], Yan Weidong[1]

(1. Department of Applied Mathematics, Northwestern Polytechnical University, Xi'an 710072, China
2. Department of Automatic Control, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: In view of the difficulty to get enough training samples in hyperspectral image, this paper presents a novel supervised linear manifold learning feature extraction method based on the manifold learning, Fisher criterion and Maximum Margin Criterion, named LPLDE by us, for hyperspectral image classification with nearest neighbor (NN) classifier. The intraclass compactness and interclass separability of hyperspectral data are respectively characterized by within-class neighboring graph and between-class neighboring graphs via embedding. The LPLDE method which efficiently avoids the within-class scatter matrix singularity caused by small-sample-size problem has better discriminative performance and is more suitable for classification. Experimental results on hyperspectral datasets and their analysis demonstrate preliminarily the efficiency of our LPLDE method as compared with other existing methods.

Key words: efficiency, experiments, feature extraction, image classification; dimensionality reduction, hyperspectral image classification, manifold learning, small-sample-size problem