

引用格式:Luo Renbo, Pi Youguo, Liao Wenzhi. Research on Supervised LPP Feature Extraction for Hyperspectral Image[J]. Remote Sensing Technology and Application, 2012, 27(6): 850-856. [骆仁波, 皮佑国, 廖文志. 超光谱遥感图像有监督 LPP 特征提取研究[J]. 遥感技术与应用, 2012, 27(6): 850-856.]

超光谱遥感图像有监督 LPP 特征提取研究

骆仁波¹, 皮佑国¹, 廖文志^{1,2}

(1. 华南理工大学自动化学院自主系统与网络控制教育部重点实验室, 广东 广州 510640;

2. 根特大学通讯与信息学院, 比利时 根特 B-9000)

摘要:在局部保留投影(LPP)特征提取算法的基础上,利用样本标签信息提出了一种有监督的局部保留投影算法(SPLPP),该算法的邻接图的权值不仅考虑了 LPP 算法中的相似性权值,而且加入了监督类的相关权值。SPLPP 算法主要步骤是先用 PCA 去除高维超光谱遥感图像的冗余信息,再把监督机制引入到 LPP 中,实现图像的特征提取,将高维超光谱遥感图像投影到低维空间中,利于分类。应用 SPLPP 算法对高维的遥感原始超光谱图像进行特征提取后,利用支持向量机(SVM)和最近邻分类器(KNN)对降维后的遥感图像数据进行分类;并与 PCA、LPP、LDA 等特征提取算法进行了比较实验。实验表明:结合了 LPP 局部信息保留能力和全域标签信息的 SPLPP 算法,有更好的局部信息保留能力和类判别能力,使分类器分类精度更高,分类效果更好。

关 键 词:有监督局部保留投影;特征提取;超光谱遥感图像;分类

中图分类号:TP 751 **文献标志码:**A **文章编号:**1004-0323(2012)06-0850-07

1 引 言

超光谱遥感图像信息丰富,广泛应用于地表目标识别领域,如环境监测^[1]、土地农作物分析^[2]、气象分析^[3]等。由于超光谱遥感图像光谱频带波段数量多、图像维数高,处理中数据量庞大,计算量大,处理时间长,给常规的低维数据处理技术带来了新的挑战。常规的分类算法易产生“Hughes 现象^[4]”等,因而高维图像特征提取引起了广泛的兴趣并获得了快速发展,如主成分分析 PCA(Principal Component Analysis)^[5]、线性判别分析 LDA(Linear Discriminant Analysis)^[6-7]、局部保留投影 LPP(Locality Preserving Projection)^[8]等等。

PCA 通过分析原始数据的协方差矩阵进行特征提取,协方差矩阵的特征值代表了不同特征所含的信息量,已广泛应用于对超光谱图像的特征提取。LDA 主要以类可分离性来寻求投影方向,其目的是

把高维数据投影到低维空间,且使不同类的像素点间的距离尽可能远,而同一类内的数据节点尽可能近,从而使得不同的类得以区分。然而,PCA 与 LDA 都只考虑了遥感图像的全局特征信息,而忽略了局部信息,不利于分辨率越来越高的超光谱遥感图像的分类。LPP 能够保留像素点与其相邻点的重要的空间局部结构信息,有很好的局部判别能力,因此在人脸模式识别和遥感图像特征提取方面得到应用^[9-10]。但是 LPP 本质上只是一种无监督的降维方法,当两个类靠得较近甚至部分重合时,由于保留局部结构的特性,LPP 将会把两个不同的类投影到一起,得到不理想的结果^[11]。为此申中华等^[11]利用标签信息提出了有监督局部保留投影算法,其思想是在近邻域内进行搜索,构造类内相似矩阵和类间相似矩阵,使近邻内数据点投影到低维时同类数据点离的近,不同类数据点离得远,但其在某种程度上破坏了局部保留特性。本文提出一种有监督的局

收稿日期:2012-04-10;修订日期:2012-10-15

基金项目:广东省教育部产学研结合项目(2009B090300269)。

作者简介:骆仁波(1986—),男,湖北阳新人,博士研究生,主要从事超光谱遥感图像处理研究。E-mail:luo_rb@mail.scut.edu.cn。

通讯作者:皮佑国(1953—),男,重庆人,研究员,主要从事伺服电机控制、自动控制、认知模式识别研究。E-mail:auygpi@scut.edu.cn。

部保留投影算法 SPLPP (Supervised Principal Locality Preserving Projection), 基本步骤是先用 PCA 去掉原始数据冗余信息, 再利用训练样本全域的标签信息, 引入监督机制, 对拥有局部信息保留能力和判别能力的特征提取算法 LPP 进行改进。

本文第二部分在对 LPP 进行简述的基础上提出了超光谱遥感图像的特征提取算法 SPLPP, 并详细叙述了其原理及实现步骤。第三部分以两幅超光谱遥感图像为实验对象, 利用 SPLPP 对其进行特征提取, 然后采用支持向量机 SVM (Support Vector Machine) 和 K-最近邻 KNN (K-Nearest Neighbors) 两种分类器分别进行分类实验, 并与 PCA、LDA、LPP 等特征提取方法进行特征提取和采用相同分类器进行图像分类进行了比较研究。

2 算法描述

2.1 LPP 简介

设 $X = [x_1, x_2, \dots, x_N]$ 为一组高维原始数据向量集, 其中 $\{x_1, x_2, \dots, x_N\} \in R^M$, $Y = [y_1, y_2, \dots, y_N]$ 表示与高维原始数据的对应的低维空间, $\{y_1, y_2, \dots, y_N\} \in R_m$, ($m \ll M$), M 表示遥感图像光谱波段数, m 表示特征提取后特征空间的维数。特征提取的目标是找到一个 $M \times m$ 的投影矩阵 W , 使每一个原始数据点 x 都能通过 $y = W^T x$, 映射到另一特征空间, 并且保留高维数据的绝大部分信息。

作为一种局部线性算法, LPP 旨在找到一个最佳投影矩阵, 使得高维空间 R^M 中相邻样本在低维特征空间 R^m 中有基本相同或相似的投影, 保持样本固有的局部流形结构不变。因此, LPP 算法在低维空间表征中能够保留原始数据中重要的空间信息。

LPP 的目标函数定义为:

$$\min \sum_{ij} \|y_i - y_j\|^2 S_{ij} \quad (1)$$

约束条件为: $y_i^T D y_j = I$

其中:

$$S_{ij} = \begin{cases} e^{-\frac{\|x_i - x_j\|^2}{t}}, & \|x_i - x_j\|^2 \leq \epsilon \\ 0, & \|x_i - x_j\|^2 > \epsilon \end{cases}$$

$$\frac{1}{2} \sum_{ij} \|y_i - y_j\|^2 (P_{ij} + S_{ij}) = \frac{1}{2} \sum_{ij} (a^T x_i - a^T x_j)^2 (P_{ij} + S_{ij}) =$$

$$\sum_i a^T x_i D_{ii} x_i^T a - \sum_{ij} a^T x_i (P_{ij} + S_{ij}) x_j^T a = a^T X (D - P - S) X^T a = a^T X L X a$$

其中: $X = [x_1, x_2, \dots, x_N]$, D 是对角矩阵, 其元素是 $(P + S)$ 的列(或行)之和, $D_{ii} = \sum_j (P_{ji} + S_{ji})$,

其中: S_{ij} 为定义在像素点上的相似性矩阵, ϵ 为 x_i 局部邻域阈值。

设最优的投影矩阵 $W_{LPP} = (a_1, a_2, \dots, a_k)$, 可以通过求解式(2)最小化问题得到:

$$W_{LPP} = \arg \min_w \frac{W^T X L X^T W}{W^T X D X^T W} = \arg \min_w \frac{W^T S_L W}{W^T S_D W} \quad (2)$$

其中: $S_D = X D X^T$, $S_L = X L X^T$, D 是对角矩阵, 其元素是 S 的列(或行)之和: $D_{ii} = \sum_j S_{ji}$, $L = D - S$ 是拉氏矩阵, 投影矩阵 $W_{LPP} = (a_1, a_2, \dots, a_k)$ 由矩阵 $S_D^{-1} S_L$ 的 k 个特征值 $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_k$ 所对应的 k 个特征向量构成。

2.2 SPLPP 算法

LPP 算法本质上是无监督特征提取算法, 不能很好地利用样本的标签信息, 由此引入监督机制, 使 SPLPP 算法邻接图的权值不仅考虑了 LPP 算法中的相似性权值, 而且加入了监督类的相关权值。SPLPP 算法的思想是找到一个最佳投影矩阵, 使得高维空间中相邻样本在低维特征空间中有基本相同或相似的投影, 保持样本固有的局部流形结构不变。同时, 利用全域的标签信息, 使同类样本点产生联系, 投影到低维空间时尽量离得近。SPLPP 的目标函数定义为:

$$\min \sum_{ij} \|y_i - y_j\|^2 (P_{ij} + S_{ij}) \quad (3)$$

其中: P 为像素点相似性矩阵, S 为有监督类相关矩阵:

$$P_{ij} = \begin{cases} e^{-\frac{\|x_i - x_j\|^2}{t}}, & \|x_i - x_j\|^2 \leq \epsilon \\ 0, & \|x_i - x_j\|^2 > \epsilon \end{cases}$$

其中: ϵ 为 x_i 局部邻域阈值。

$$S_{ij} = \begin{cases} 1 & x_i, x_j \in w_l \\ 0 & x_i \in w_l, x_j \in w_k \end{cases} \quad (4)$$

其中: w_l 为训练样本的第 l 类, $l, k \in (1, 2, \dots, C)$, C 为类别数。

假设 a 为所求的投影向量, 就有 $y^T = a^T X$, 其中 y 的第 i 行与 X 的第 i 列分别为 y_i 和 x_i , 经过简单变换之后, 目标函数变为:

$L = D - P - S$ 。为了避免 $a = 0$ 这个解, 引入约束条件:

$$y^T D y = 1 \Rightarrow a^T X D X^T a = 1$$

上述问题转化为一个广义特征值问题,最小化目标函数的解可以通过求解式(5)中最小化问题得:

$$W_{\text{SPLPP}} = \arg \min_w \frac{W^T X L X^T W}{W^T X D X^T W} = \arg \min_w \frac{W^T S_L W}{W^T S_D W} \quad (5)$$

其中: $W_{\text{SPLPP}} = (a_1, a_2, \dots, a_k)$ 为最优的投影矩阵, $S_D = X D X^T$, $S_L = X L X^T$ 。

2.3 SPLPP 算法实现步骤

(1) PCA 投影。通过去除与零特征值相对应的特征成分将遥感图像投影到 PCA 子空间中,通常这些去除的特征成分所代表的是噪声信息,它们对图像的分类会产生干扰。用 W_{PCA} 表示 PCA 的变换矩阵。

(2) 构造邻接图。 G 表示有 n 个样本点的图。第 i 个样本与遥感图像像素点 x_i 相对应。如果 x_i 和 x_j 是接近的,那么用一条细边将 x_i 和 x_j 连起来。如果 x_i 和 x_j 属于同一类,那么用一条粗边将 x_i 和 x_j 连起来。

(3) 选择权值。如果 x_i 和 x_j 是细边相连的,设置 $P_{ij} = e^{-\frac{\|x_i - x_j\|^2}{t}}$ (本文 t 取 1), 否则, $P_{ij} = 0$; 如果 x_i 和 x_j 属于同一类,设置 $S_{ij} = 1$, 否则 $S_{ij} = 0$ 。 G 的权值矩阵 $(P + S)$ 模拟了遥感图像的局部结构信息和有监督类相关信息。

(4) 计算 W_{SPLPP} 。通过求解式(6)最小化问题得:

$$W_{\text{SPLPP}} = \arg \min_w \frac{W^T X L X^T W}{W^T X D X^T W} = \arg \min_w \frac{W^T S_L W}{W^T S_D W} \quad (6)$$

$W_{\text{SPLPP}} = (a_1, a_2, \dots, a_k)$ 由矩阵 $S_D^{-1} S_L$ 的 k 个特征值 $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_k$ 所对应的 k 个特征向量构成。

(5) SPLPP 嵌入。假设:

$$W_{\text{SPLPP}} = (a_1, a_2, \dots, a_k), \text{嵌入表示如下:} \\ x \rightarrow y = W^T x, W = W_{\text{PCA}} W_{\text{SPLPP}} \quad (7)$$

其中: y 是遥感图像 x 的 k 维表达, W 是投影矩阵。

3 实验

3.1 实验方案

利用 SPLPP 特征提取算法对超光谱图像进行特征提取(降维),并与目前较流行的 PCA、LPP 以及有监督 LDA 算法进行比较。实验对象为 India Pine 和 DC mall 两幅光谱图像,并选取支持向量机(SVM)分类器和 K-最近邻分类器两种分类器进行实验。除了特征提取算法外,试验中的对象、参数等环境完全相同,每次实验重复 10,取平均值。

3.2 分类器设计

3.2.1 SVM 分类器

SVM^[12] 具有小样本学习、抗噪声性能优、学习

效率高、推广性好等优点,并广泛应用于光谱图像分类中^[13]。设计分类器为:

公式表述:

$$\varphi(W) = \min_{W, \zeta_l} \left\{ \frac{1}{2} W^T \cdot W + C \sum_{l=1}^n \zeta_l \right\} \quad (8)$$

约束条件:

$$y_l((W \cdot x_l) + b) + \zeta_l \geq 1, \zeta_l \geq 0, l = 1, \dots, n$$

选择的核函数为径向基函数:

$$K(x, x_j) = \exp(-g \|x - x_i\|^2), g > 0 \quad (9)$$

式(8)表示的是 C-SVC 类型 SVM 分类器, C 为惩罚参数,式(9)表示径向基函数。式(8)和式(9)中的参数 C 和 g 构成参数对 (C, g) ,通过 5 折交叉验证及网格寻优搜索技术搜索最优参数对,并设定: C 与 g 的搜索范围都为 $[2^{-8} \ 2^8]$,步长都为 2^1 。

3.2.2 K-最近邻分类器

K-最近邻分类算法(KNN)在 N 个已知样本中,找出样本 x 的 k 个近邻。

定义判别函数: $g_i = k_i, i = 1, 2, \dots, C$
决策规则为:若 $g_j(x) = \max_i k_i$,则决策 $x \in w_j$ 。本文中 k 都取值为 1。

3.3 实验对象

实验对象图像如图 1 所示。

图 1(a)为 India Pine 拍摄对象是印第安 India Pine 的西北部部分区域的光谱图像^[15],其大小是 145×145 ,此图含有 220 个波段,通过去除噪声通道频带后保留 185 个波段^[14],覆盖的波长范围是 $0.4 \sim 2.4 \mu\text{m}$ 。此区域所含类别数为 16,每类的实验样本数量见表 1,实验样本总数为 10 366。试验样本真实地物图如图 2(a),每一种颜色代表一类。

图 1(b)为 DC mall 拍摄对象是美国华盛顿城市的一片区域^[16],大小是 $1\ 280 \times 307$,去掉噪声通道频带后保留 191 个波段,覆盖的波长范围是 $0.4 \sim$

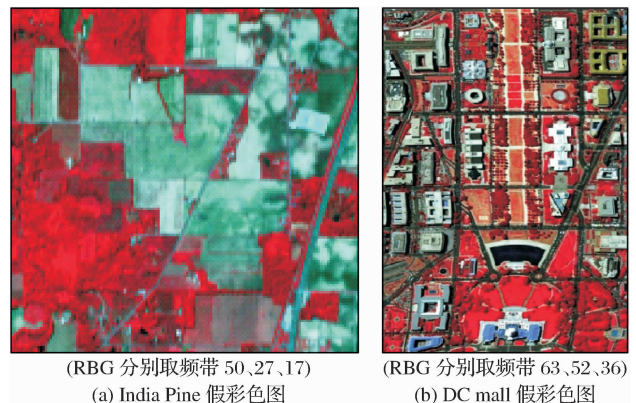


图 1 实验图像

Fig. 1 Images for experiment

表 1 实验样本类别及其数量

Table 1 Classes and corresponding numbers of experimental samples

图像 India Pine						图像 DC mall		
类别	名称	数量	类别	名称	数量	类别	名称	数量
1	玉米 1	1 434	9	大豆 1	986	1	屋顶	2 685
2	玉米 2	834	10	大豆 2	2 468	2	街道	416
3	玉米 3	234	11	大豆 3	614	3	小路	73
4	苜蓿	54	12	燕麦	20	4	草地	418
5	牧场 1	497	13	小麦	212	5	树木	405
6	牧场 2	747	14	乔木	1 294	6	水域	449
7	牧场 3	26	15	木屋	380	7	阴影	47
8	干草	489	16	石塔	95			

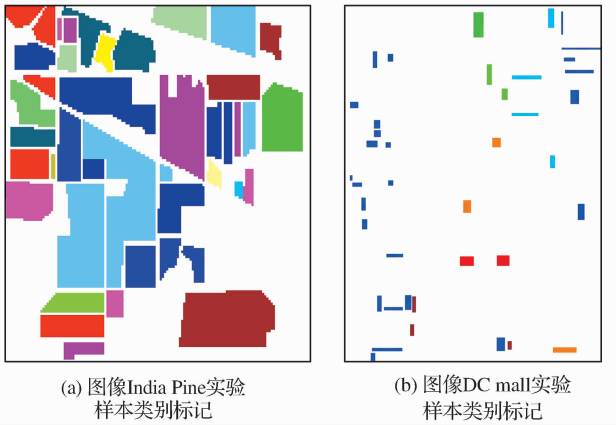


图 2 实验样本真实地物标记图

Fig. 2 Ground truth for experimental samples

2.4 μm 。本文选取了 550×307 大小的区域进行实

验,这个区域所含类别数为 7,每类的实验样本数量见表 1,实验样本总数为 4 493。试验样本真实地物图如图 2,每一种颜色代表一类。

3.4 实验及结果分析

3.4.1 实验过程及结果

对如图 2 所示的实验对象先进行波段提取,去掉噪声通道波段。然后选取相应的训练样本,用上述 4 种不同的特征提取算法进行特征提取,得到各自的投影矩阵 W ,通过 W 将高维的实验样本投影到低维特征空间中。接着用低维特征空间的训练样本训练分类器,最后用训练好的分类器对测试样本进行分类,得出分类精度。

从图像 India Pine 和 DC mall 每类样本中随机抽取 10% 的样本作为训练样本时的实验结果(图 3)。

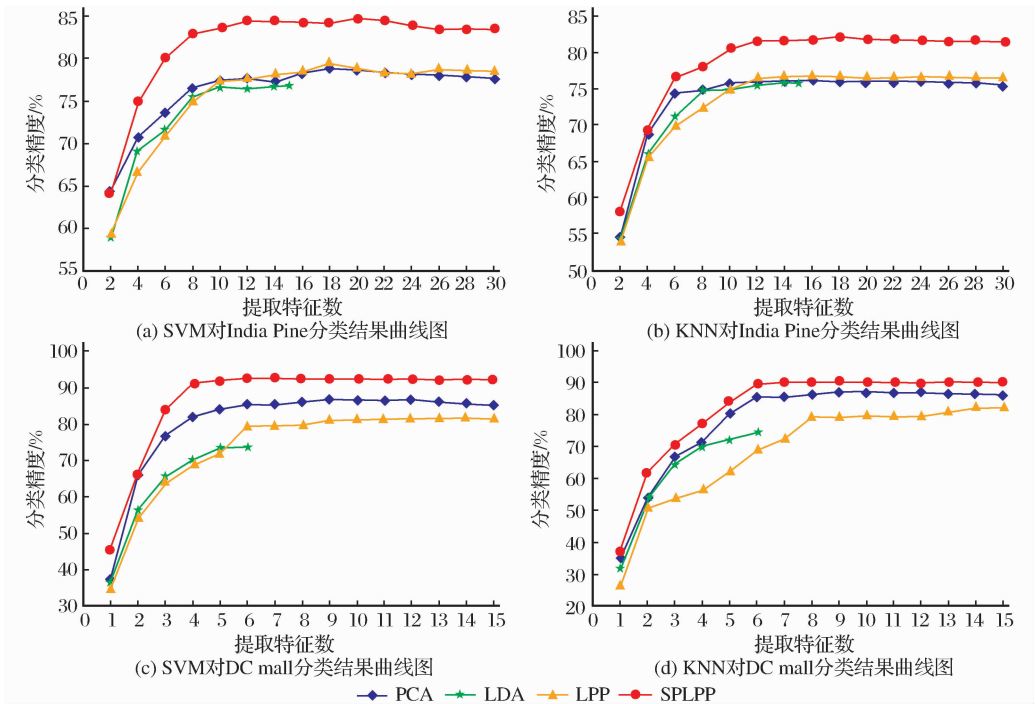


图 3 平均分类精度曲线图

Fig. 3 Curves of average classification accuracies

其中图 3(a)为 SVM 对 India Pine 的分类结果曲线图(横坐标为提取特征数,纵坐标为分类精度),图 3 (b)为 KNN 对 India Pine 的分类结果曲线图,图 3 (c)为 SVM 对 DC mall 的分类结果曲线图,图 3(d)

为 SVM 对 DC mall 的分类结果曲线图。提取特征数为 14 时,SVM 对 India Pine 的分类效果如图 4 所示,特征提取数为 6 时,SVM 对 DC mall 分类后的效果图如图 5 所示。

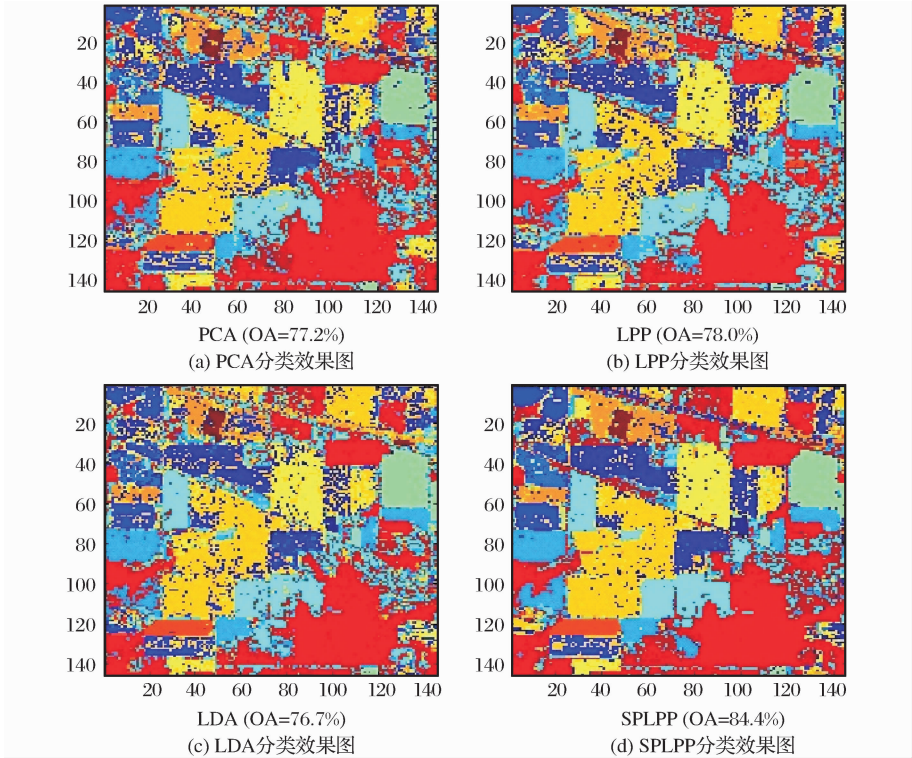


图 4 利用 SVM 对 India Pine 分类后的效果图,提取特征数为 14
Fig. 4 Classified maps using SVM classifier for India Pine,14 features

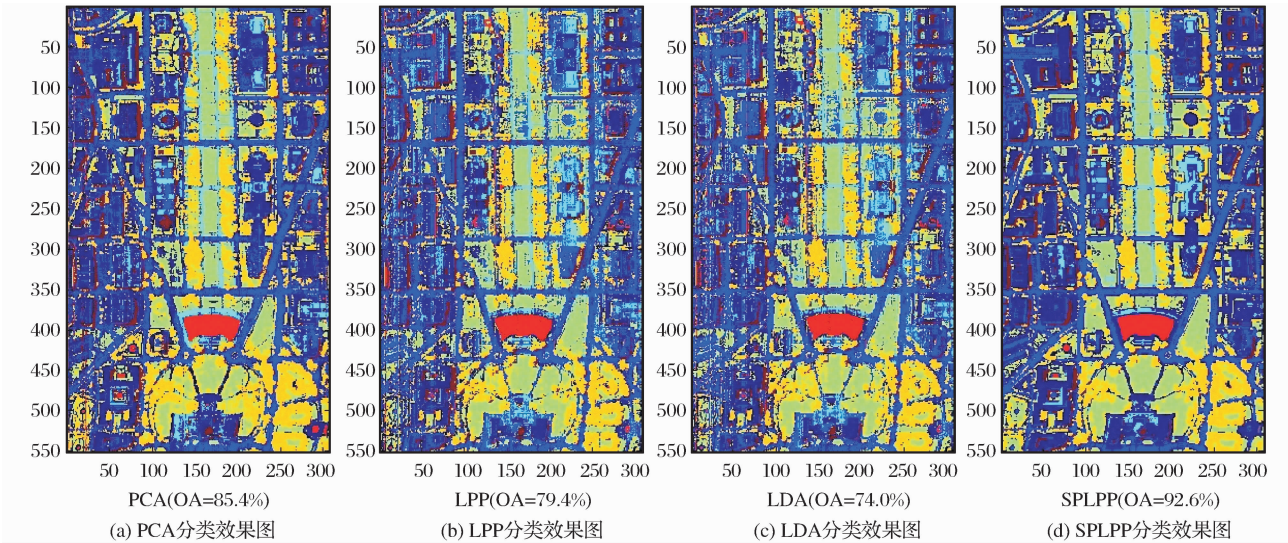


图 5 利用 SVM 对 DC mall 分类后的效果图,提取特征数为 6
Fig. 5 Classified maps using SVM classifier for DC mall,6 features

选取 3 组不同比例数量的训练样本,其中 India Pine 中训练样本分别从各类中随机选取 5%、10%、15%,特征提取数 2~30;DC mall 中训练样本分别

从各类中随机选取 2%、5%、10%,特征提取数 1~15。如表 2,数据为出现的最高分类精度,括号内的数据表示出现最高分类精度时的特征提取数。

3.4.2 结果分析

由图 3 和表 2 可见,超光谱图像 India Pine 和 DC mall,对于两种分类器,采用 SPLPP 作为特征提取算法,提取特征数大于 2 以后的分类效果优于其它特征提取方法,且提取特征数为 6 时,分类器的分类精度就已经较高了;表明本文提出的 SPLPP 特征提取算法比 PCA、LPP、LDA 更适用于超光谱图像

的特征提取。同时结果也显示了 SPLPP 算法的邻接图的权值不仅考虑了 LPP 算法中的相似性权值,而且加入了监督类的相关权值,较 LPP 更具判别能力。SPLPP 算法不但能够保留像素点重要的局部结构信息还能充分利用样本的全局标签信息,提取出超光谱遥感图像的重要特征,有利于分类器的分类。

表 2 最高分类精度(提取特征数)
Table 2 Highest classification accuracies(features)

图像	分类器	样本/每类	PCA	LPP	LDA	SPLPP
India Pine	KNN 分类器	5%	64.5%(18)	61.4%(16)	60.3%(15)	74.3%(14)
		10%	76.2%(16)	76.7%(18)	75.9%(14)	82.2%(18)
		15%	82.1%(20)	79.8%(20)	78.4%(14)	86.4%(16)
	SVM 分类器	5%	62.8%(16)	63.5%(14)	61.4%(14)	75.6%(16)
		10%	78.9%(18)	79.2%(18)	76.8%(15)	84.7%(20)
		15%	85.0%(14)	84.4%(18)	82.8%(15)	86.5%(16)
DC mall	KNN 分类器	2%	82.2%(10)	78.4%(7)	72.9%(6)	85.1%(9)
		5%	87.0%(9)	80.6%(13)	75.2%(6)	90.2%(13)
		10%	88.8%(10)	85.6%(11)	80.7%(5)	92.4%(12)
	SVM 分类器	2%	84.1%(10)	79.9%(12)	73.3%(5)	87.8%(9)
		5%	86.9%(9)	81.7%(14)	74.0%(6)	92.6%(11)
		10%	89.5%(8)	87.9%(13)	81.1%(6)	93.9%(8)

4 结 语

本文在 LPP 基础特征提取算法上,结合训练样本的标签信息,提出了一种有监督局部保留投影算法(SPLPP),并用于超光谱遥感图像特征的提取。实验选取主成分分析(PCA)、邻域保留投影(LPP)和线性判别分析(LDA)3 种特征提取算法进行对比,在相同分类器的条件下对两幅超光谱图像进行分类对比研究。超光谱遥感图像分类对比实验表明:SPLPP 特征提取算法能充分利用训练样本的全局标签信息,保留样本固有的局部流形结构不变,整体与部分的结合,使分类器分类效果得到的提升,分类效果好于其他 3 种特征提取算法的分类效果。

参考文献(References):

[1] Agyemang T, Heblinski K, Schmieder K, *et al.* Accuracy Assessment of Supervised Classification of Submersed Macrophytes; The Case of the Gavaraget Region of Lake Sevan, Armenia[J]. *Hydrobiologia*, 2011, 661(1): 85-96 .
[2] Tracol Y, Gutierre J R, Squeo F A. Plant Area Index and Microclimate Underneath Shrub Species from a Chilean Semiarid Community[J]. *Journal of Arid Environments*, 2011, 75(1): 1-6.
[3] Tomlinson C J, Chapman L. Remote Sensing Land Surface Te-

mperature for Meteorology and Climatology: A Review[J]. *Meteorological Applications*, 2011, 18(3): 296-306.
[4] Hughes G F. On the Mean Accuracy of Statistical Pattern Recognizers[J]. *IEEE Transactions on Information Theory*, 1968, 14(1): 55-63.
[5] Fauvel M, Chanussot J, Benediktsson J A. Kernel Principal Component Analysis for the Classification of Hyperspectral Remote-sensing Data over Urban Areas[J]. *Eurasip Journal on Advances in Signal Processing*, 2009, 10(1): 1155-1168.
[6] Chang C I, Ren H. An Experiment-based Quantitative and Comparative Analysis of Target Detection and Image Classification Algorithms for Hyperspectral Imagery[J]. *IEEE Transactions on Geoscience Remote Sensing*, 2000, 38(2): 1044-1063.
[7] Yang Jinnmin, Yu Paota, Kuo Borchon. A Nonparametric Feature Extraction and Its Application to Nearest Neighbor Classification for Hyperspectral Image Data[J]. *IEEE Transactions on Geoscience Remote Sensing*, 2010, 48(3): 1279-1293.
[8] He X F, Niyogi P. Locality Preserving Projections[C]// *Advances in Neural Information Processing Systems 16 (NIPS)*, Vancouver, Canada, 2003: 153-160.
[9] Yu C, Xu X H, Lai J H. Optimal Locality Preserving Projection for Face Recognition[J]. *Neurocomputing*, 2011, 74(18): 3941-3945.
[10] Chen G Y, Qian S E. Dimensionality Reduction of Hyperspectral Imagery Using Improved Locally Linear Embedding[J].

Journal of Applied Remote Sensing,2007,1(1):1-10.

[11] Shen Zhonghua, Pan Yonghui, Wang Shitong. A Supervised Locality Preserving Projection Algorithm for Dimensionality Reduction[J]. Pattern Recognition and Artificial Intelligence, 2008,21(2):233-239. [申中华, 潘永惠, 王士同. 有监督的局部保留投影降维算法[J]. 模式识别与人工智能, 2008,21(2):233-239.]

[12] Cristianini N, Shawe T J. An Introduction to Support Vector Machines [M]. Cambridge: Cambridge University Press, 2000.

[13] Tang Jingtian, Hu Dan, Gong Zhimin. Study of Classification by Support Vector Machine on Synthetic Aperture Radar Image[J]. Remote Sensing Technology and Application, 2008,22(3):341-345. [汤井田, 胡丹, 龚智敏. 基于 SVM 的 SAR 图像分类研究[J]. 遥感技术与应用, 2008,22(3):341-345.]

[14] Mojaradi B. Dimensionality Reduction of Hyperspectral Data via Spectral Feature Extraction[J]. IEEE Transactions on Geoscience Remote Sensing, 2009,47(7):2091-2105.

[15] AVIRIS NW Indiana's Indian Pines 1992 Data Set[DB/OL]. ftp.ecn.purdue.edu /biehl/MultiSpec/92AV3C (Original Files) and ftp://ftp.ecn.purdue.edu/biehl/PC_MultiSpec/ThyFiles.zip (ground truth) .

[16] Landgrebe D A. Signal Theory Methods in Multispectral Remote Sensing. Hoboken[M]. New Jersey: Wiley, 2003.

Research on Supervised LPP Feature Extraction for Hyperspectral Image

Luo Renbo¹, Pi Youguo¹, Liao Wenzhi^{1,2}

(1. Key Laboratory of Autonomous Systems and Network Control Ministry of Education College of Automation Science and Engineering, South China University of Technology, Guangzhou 510640, China;
2. Department of TELIN-IPI-IBBT, Ghent University, Ghent, Belgium)

Abstract: For the classification among different land-cover types in a hyperspectral image, particularly in the small-sample-size problem, a feature extraction method is an approach for reducing the dimensionality and increasing the classification accuracy. A Supervised Principal Locality Preserving Projection (SPLPP) feature extraction algorithms, which uses the label information of training sample in Locality Preserving Projection (LPP), was proposed in this paper. Three main steps are involved in the proposed SLPP: firstly, use PCA to remove redundant information, and then combine the label information in LPP, finally, SPLPP projects high-dimensional hyperspectral image into a low-dimensional space. Last but not least, SPLPP uses the extracted features as inputs of classifiers (e. g. Support Vector Machine (SVM) and K-Nearest Neighbors (KNN)) to do classification. The experimental results show that the proposed SPLPP has better local information retention ability and class discrimination ability, compared with PCA, LPP, LDA.

Key words: Supervised principal locality preserving projection; Feature extraction; Hyperspectral remote sensing image; Classification