

基于主成分分析与局部二值模式的高光谱图像分类

叶珍, 白麟

长安大学电子与控制工程学院, 陕西 西安 710064

摘要 提出了两种基于主成分分析与局部二值模式的高光谱图像分类算法。利用主成分分析去除高光谱图像的谱间冗余信息,对降维后的图像利用局部二值模式进行空间纹理特征分析,采用稀疏表示分类和支持向量机分别对提取的特征进行分类。其通过将主成分分析与局部二值模式相结合对高光谱图像进行特征提取,保证了高光谱图像的谱间冗余的有效去除,同时保护了高光谱图像的空间局部邻域信息,因此,此类算法不但能充分挖掘高光谱图像的谱间-空间特征,在较大程度上提高分类精度和 Kappa 系数,而且在高斯噪声环境中和小样本情况下也具有良好的分类性能。

关键词 图像处理; 高光谱图像分类; 主成分分析; 局部二值模式; 特征提取

中图分类号 TP751.1 **文献标识码** A

doi: 10.3788/LOP54.111006

Hyperspectral Image Classification Based on Principal Component Analysis and Local Binary Patterns

Ye Zhen, Bai Lin

School of Electronics and Control Engineering, Chang'an University, Xi'an, Shaanxi 710064, China

Abstract Two kinds of hyperspectral image classification algorithms based on principal component analysis and local binary patterns are proposed. The principal component analysis is employed to reduce the redundant information in spectral domain. Following that, the local binary patterns are studied to analyze the spatial texture features. And the sparse presentation classification and support vector machine are used for a classification of extracted results, respectively. Combining the principal component analysis with the local binary patterns for extracting the features of hyperspectral image, we ensure that the spectral redundant information is reduced effectively, and the spatial local neighborhood information is protected. Hence, the proposed algorithms can not only sufficiently excavate spectral-spatial features of hyperspectral image for improving classification accuracy and Kappa coefficient, but also have outstanding classification performance in Gaussian noise environments and small-sample-size condition.

Key words image processing; hyperspectral image classification; principal component analysis; local binary patterns; feature extraction

OCIS codes 100.3008; 100.5010; 120.0280

1 引言

随着遥感成像技术的发展,高光谱成像光谱仪的光谱分辨率越来越高,高光谱图像获得的波段数越来越多。虽然数据量的增加能更加细致地对遥感地物分类,但是高维数据中也包含了大量的冗余信息,不但增加了计算成本,而且降低了分类精度^[1-2]。因此,常常需要将高光谱遥感图像先投影到更低维的子空间,再进行后续处理。传统的降维方法分为两类:1) 无监督方法,如主成分分析(PCA)法^[3];2) 有监督方法,如线性判别式分析(LDA)法^[4]。

收稿日期: 2017-05-16; **收到修改稿日期:** 2017-06-22

基金项目: 国家自然科学基金(41601344,61601059,51407012)、中央高校基本科研业务费专项资金(310832171006,310832163402,310832161004)

作者简介: 叶珍(1983—),女,博士,讲师,主要从事高光谱图像处理方面的研究。E-mail: yezhen525@126.com

然而,单纯通过光谱特征对高光谱图像进行分类,会忽略空间特征导致分类效果不理想^[5]。近年来,越来越多的学者致力于研究高光谱图像的空间特征提取技术,如离散小波变换(DWT)^[6]、离散 Gabor 变换(DGT)^[7]、马尔科夫随机场(MRF)^[8]等。MRF 通过贝叶斯最大后验概率(MAP)框架来获取高光谱图像的空间上下文信息,与支持向量机(SVM)^[9]结合的 SVM-MRF^[10]已被证明是一种具有较低复杂度的高光谱图像分类算法。局部保护判别分析(LFDA)和局部保护非负矩阵分离(LPNMF)是两种适用于高光谱图像的降维方法。将 LFDA 与高斯混合模型(GMM)结合的 LFDA-GMM 算法^[11]能有效地对高光谱图像分类。进一步地,通过将 LFDA-GMM 算法^[11]与 MRF 结合,文献[12]提出的 LFDA-G-MRF 和 LPNMF-G-MRF 谱间-空间分类算法,相比传统的高光谱图像分类算法,能更有效地提高分类性能。文献[13]提出的 Gabor-LFDA-GMM 和 Gabor-LPNMF-GMM 是基于 PCA 降维和 Gabor 特征提取的高光谱图像分类算法,不但能得到较高的分类精度和 Kappa 系数,而且在噪声环境中具有良好的稳健性。

对于高光谱图像分类应用,一方面希望充分挖掘谱间-空间的有用信息来提高分类精度,另一方面又希望降低算法的复杂度。因此,研究低复杂度的分类算法具有实际的应用价值。通过研究发现,局部二值模式(LBP)^[14]是一种适用于高光谱图像的空间特征提取技术,它具有灰度和旋转不变性的优点,可以在角度空间的圆形邻域内和任意空间分辨率下挖掘高光谱图像的局部纹理特征。在不同主成分的投影子空间,利用 LBP 高效地提取高光谱图像的局部空间纹理特征,结合稀疏表达分类(SRC)^[15]和 SVM,本文提出了两种基于 PCA 与 LBP 的高光谱图像分类算法,分别命名为 PCA-LBP-SRC 和 PCA-LBP-SVM。实验结果表明,与传统的 SRC、SVM 以及近期提出的 SVM-MRF、LFDA-GMM、LFDA-G-MRF、LPNMF-G-MRF、Gabor-LFDA-GMM 和 Gabor-LPNMF-GMM 算法相比,本文两种算法具有更优的分类性能。其中,PCA-LBP-SVM 算法不但分类精度最高,而且具有相对较低的计算复杂度。

2 基于 PCA 与 LBP 的高光谱图像分类

2.1 PCA 与 LBP 特征提取

设 $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_P\}$ 为高光谱图像的谱向矢量集,其中 $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{iN})^T$ 为任一样本像元的谱向矢量, $i = 1, 2, \dots, P$, P 为矢量个数, N 为波段数。对 \mathbf{X} 的协方差矩阵 $\mathbf{R}_{N \times N}$ 进行特征值分解

$$\mathbf{R}_{N \times N} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T, \quad (1)$$

式中 $\mathbf{U} = [\mathbf{u}_1 \mathbf{u}_2 \dots \mathbf{u}_N]$ 为正交特征矢量矩阵, $\mathbf{\Lambda} = \text{diag}[\lambda_1 \lambda_2 \dots \lambda_N]$ 为 $\mathbf{R}_{N \times N}$ 的特征值矩阵,且 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N \geq 0$ 。对谱向矢量进行 PCA 变换,得到

$$\mathbf{y}_i = \mathbf{U}^T (\mathbf{x}_i - \boldsymbol{\mu}), \quad (2)$$

式中 $\mathbf{y}_i = (y_{i1}, y_{i2}, \dots, y_{iN})^T$, $i = 1, 2, \dots, P$; $\boldsymbol{\mu}$ 为样本的均值;相应的谱向矢量集为 $\mathbf{Y} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_P\}$ 。经过 PCA 变换得到的各主成分分量彼此不相关,随着主成分编号的增加,各个分量所包含的信息量在逐渐减少。可将高光谱图像在谱向空间视为一个奇异矩阵,利用 PCA 去除其中的零空间,以达到谱向降维的目的。通过 PCA 变换可以避免在谱向对冗余特征进行 LBP 分析,不但能有效提高分类精度,而且大幅度降低了计算成本。在实验部分,将通过分类精度和运算时间的比较来确定主成分的个数。

LBP 是基于灰度和旋转不变性的空间纹理特征分析法,可在角度空间的圆形邻域内和任意空间分辨率下计算图像块(即局部图像,如大小为 13 pixel \times 13 pixel 的区域)的模式。LBP 通过计算中心像素 g_c 和其邻域像素 $\{g_i\}_{i=0}^{B-1}$ 的灰度差值来描绘图像块的空间纹理特征,其中 B 是以 R 为半径的圆形邻域内的成员个数。可将 B 和 R 视为两个参数,分别控制并决定角度空间内的成员个数和空间分辨率的大小。对于同一图像块,设置不同的 (B, R) 参数,将得到不同的邻域集,而各邻域集内的成员均具有对称性^[14]。当给定中心像素后,可给邻域内的每个成员分配一个二值标记,如 1 或 0,取决于邻域成员的灰度值是否大于中心像素的灰度值。LBP 策略可定义为

$$P_{\text{LBPB}, R} = \sum_{i=0}^{B-1} s(g_i - g_c) 2^i, \quad s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}, \quad (3)$$

式中 $s(g_i - g_c)$ 为符号函数。由于半径为 R 的圆形邻域内包含 2^B 种 LBP,随着邻域集内成员个数的增加, LBP 的种类也会急剧增加,从而增加了分类的难度。根据 LBP 的旋转不变性,可以采用等价模式:

$$P_{LBPB,R}^{riu2} = \begin{cases} \sum_{i=0}^{B-1} s(g_i - g_c), & \text{if } U(P_{LBPB,R}) \leq 2, \\ B + 1, & \text{otherwise} \end{cases}, \quad (4)$$

式中 $U(P_{LBPB,R})$ 为从 1 到 0 或从 0 到 1 的跳变次数。以 LBP 的种类作为横轴, 每个种类出现的次数作为纵轴, 则可建立 LBP 统计直方图。由于在实际高光谱图像中, 绝大多数 LBP 的 1/0 或 0/1 跳变次数都小于或等于 2, 因此, 根据等价模式, LBP 统计直方图的个数由原来的 2^B 减小为 $B(B-1)+3$, 这使得特征向量的维数大幅度降低。本文将所有图像块的 LBP 统计直方图连接成为一个特征向量, 作为高光谱图像的 LBP 空间纹理特征。在实验部分, 通过比较分类结果和运算时间来确定 LBP 中参数 B 和 R 的大小。

2.2 SRC

对于高光谱图像来说, 每个像元可能包含几种地物, 可以通过几个端元的线性组合来表示^[16]。设高光谱图像有 C 个不同的类, 测试样本为 $\mathbf{x} \in \mathbb{R}^N$, 其中 N 代表波段总数。在 SRC 中, 属于同一类的像素位于同一低维子空间^[17], 未知的测试样本 \mathbf{x} 可以看作是训练样本的稀疏线性结合

$$\mathbf{x} = \mathbf{A}_1 \boldsymbol{\alpha}_1 + \mathbf{A}_2 \boldsymbol{\alpha}_2 + \cdots + \mathbf{A}_C \boldsymbol{\alpha}_C = [\mathbf{A}_1 \cdots \mathbf{A}_C] \begin{bmatrix} \boldsymbol{\alpha}_1 \\ \vdots \\ \boldsymbol{\alpha}_C \end{bmatrix} = \mathbf{A} \boldsymbol{\alpha}, \quad (5)$$

式中 $\mathbf{A} \in \mathbb{R}^{N \times M}$ 是由各类别字典 $\{\mathbf{A}^i\}_{i=1, \dots, C}$ 组成的整体结构字典, M 代表训练样本的个数。 $\boldsymbol{\alpha} \in \mathbb{R}^M$ 由各类别的稀疏向量 $\{\boldsymbol{\alpha}^i\}_{i=1, \dots, C}$ 组成。当训练样本的字典 \mathbf{A} 已知时, 求解稀疏向量 $\boldsymbol{\alpha}$ 可通过最优化以下问题来解决

$$\min_{\boldsymbol{\alpha}} \frac{1}{2} \|\mathbf{x} - \mathbf{A} \boldsymbol{\alpha}\|_F^2 + \lambda \|\boldsymbol{\alpha}\|_1, \quad (6)$$

式中 $\min_{\boldsymbol{\alpha}} \frac{1}{2} \|\mathbf{x} - \mathbf{A} \boldsymbol{\alpha}\|_F^2$ 为重构误差(测试样本和重建图像间的误差), $\lambda \|\boldsymbol{\alpha}\|_1$ 为惩罚系数。当字典 \mathbf{A} 和稀疏向量 $\boldsymbol{\alpha}$ 不固定时, 上述问题通常采用迭代法求解, 得到的解 $\hat{\boldsymbol{\alpha}}$ 是一个只含有少数不为零元素的稀疏向量^[18]。测试样本 \mathbf{x} 的类别可以直接由被恢复的稀疏向量 $\hat{\boldsymbol{\alpha}}$ 来判定, $\hat{\boldsymbol{\alpha}}_i$ 代表第 i 类已恢复的稀疏向量。那么, 测试样本的类别标记就可以通过下式来判定

$$\text{Class}(z) = \arg \min_{i=1, \dots, C} \|\mathbf{x} - \mathbf{A}_i \hat{\boldsymbol{\alpha}}_i\|_F. \quad (7)$$

由于高光谱图像的相邻像元间具有空间相关性, 单纯的利用谱间稀疏性对其进行分类, 在一定程度上会影响分类精度。本文算法 PCA-LBP-SRC 将 LBP 空间邻域特征用于对高光谱图像稀疏表示分类, 对 LBP 空间每个成员的光谱维向量通过稀疏表达构建测试样本, 再根据邻域相关性和最小重构误差约束稀疏向量, 实现了利用 LBP 空间特征辅助表达光谱稀疏性, 有效改善了分类性能。

2.3 SVM

假设 $\mathbf{x}'_i \in \mathbb{R}^d$ 为来自于 d 维 LBP 特征空间的特征向量, 定义训练样本 $\mathbf{X}' = \{\mathbf{x}'_i\}_{i=1}^n$, 类标记 $y'_i \in \{-1, +1\}$ 和非线性核映射 $\phi(\cdot)$, 设最优超平面为 $\phi^T(\mathbf{x}'_i) \cdot \mathbf{w}_i = 0$, 则权值向量 \mathbf{w}_i 和偏置 b 必须满足以下约束

$$y'_i [\phi^T(\mathbf{x}'_i) \cdot \mathbf{w}_i + b] \geq 1 - \xi_i, \quad (8)$$

式中 ξ_i 为松弛变量。SVM 找到一个使训练数据平均错误分类误差最小的超平面, 可推导出以下优化问题

$$\min_{\mathbf{w}, \xi, b} \frac{1}{2} \left\{ \|\mathbf{w}\|_2^2 + \delta \sum_{i=1}^n \xi_i \right\}, \quad (9)$$

式中 δ 为一指定的正参数(惩罚系数), 表示 SVM 对错分样本的惩罚程度。根据拉格朗日乘法准则, 最优分类超平面的求解可以转化为以下约束优化问题

$$\max_{\mathbf{a}} \left\{ \sum_{i=1}^n a_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n a_i a_j y'_i y'_j K(\mathbf{x}'_i, \mathbf{x}'_j) \right\}, \quad (10)$$

式中 $\{a_i\}_{i=1}^n$ 为非零拉格朗日乘子, 其对应的样本称为支持向量, 且满足条件 $0 < a_i < \delta$ 和 $\sum_i a_i y'_i = 0$ 。定义决策函数为

$$f(\mathbf{x}') = \text{sgn} \left[\sum_{i=1}^n y_i a_i K(\mathbf{x}', \mathbf{x}'_i) + b \right], \quad (11)$$

式中 $K(\mathbf{x}'_i, \mathbf{x}'_j)$ 为满足 Mercer 定理的核函数。由于径向基函数(RBF)具有较好的分类性能,且只有一个参数 σ ,因此采用 RBF 作为 SVM 的核函数,定义为

$$K(\mathbf{x}'_i, \mathbf{x}'_j) = \exp\left(-\frac{\|\mathbf{x}'_i - \mathbf{x}'_j\|^2}{2\sigma^2}\right), \sigma \in \mathbb{R}. \quad (12)$$

2.4 算法步骤

研究了基于 PCA 和 LBP 的谱间-空间特征提取方法,并通过结合 SRC 和 SVM 分类器,提出了两种高光谱图像分类算法,即 PCA-LBP-SRC 和 PCA-LBP-SVM。具体步骤为:1) 对原始高光谱图像利用 PCA 降维,得到前 N_{PC} 个最大主成分;2) 对降维后的图像在空间利用检测窗口划分成图像块,如印第安松林数据采用 $13 \text{ pixel} \times 13 \text{ pixel}$ 的小区域;3) 将各个图像块的中心像素的灰度值与邻域像素的灰度值进行比较,得到中心像素的 LBP 值;4) 计算每个图像块内 LBP 的直方图(即每个图像块的中心像素 LBP 值出现的频率),将所有图像块的 LBP 统计直方图连接成为一个特征向量,作为降维后高光谱图像的 LBP 空间纹理特征;5) 分别结合 SRC 分类器和 SVM 分类器对提取的特征进行分类。

3 实验与分析

3.1 实验数据

实验数据为机载高光谱遥感数据。实验数据一为来自 AVIRIS 传感器在美国印第安那州获取的 220 个波段的印第安松林高光谱遥感数据。该数据的空间分辨率为 20 m,每个波段图像的大小为 $145 \text{ pixel} \times 145 \text{ pixel}$ 。原始印第安松林数据有 16 个类别,选取其中像素较多的 8 个类别作为主要分析对象,如图 1 所示。实验数据二为来自 AVIRIS 传感器在美国加利福尼亚州获取的 224 个波段的萨利纳斯山谷高光谱遥感数据。该数据的空间分辨率为 3.7 m,每个波段图像的大小为 $512 \text{ pixel} \times 217 \text{ pixel}$,共包括 16 种地物,如图 2 所示。



图 1 印第安松林数据。(a)假彩色图像;(b)真实地物标记

Fig. 1 Indian Pines dataset. (a) False-color image; (b) ground truth

3.2 参数设置

所有的实验结果都是通过 Matlab7.6.0 仿真软件,在 2.8 GHz 主频和 7.89 GB 内存的计算机上实施得到的。为了保证实验的客观性,对各个类别随机选择训练样本,所有实验均重复 20 次。表 1 为 20 次实验得到的平均总体分类精度(OA),简称分类精度。括号内数据为其标准差,其值越小,说明分类结果越稳定;反之,亦然。对算法中的几个主要参数进行设置,包括主成分个数 N_{PC} 、LBP 中的邻域半径 R 和邻域成员个数 B 。对于印第安松林高光谱数据,对每个类别随机选择 50 pixel 作为训练样本,对主成分个数 N_{PC} 进行测试

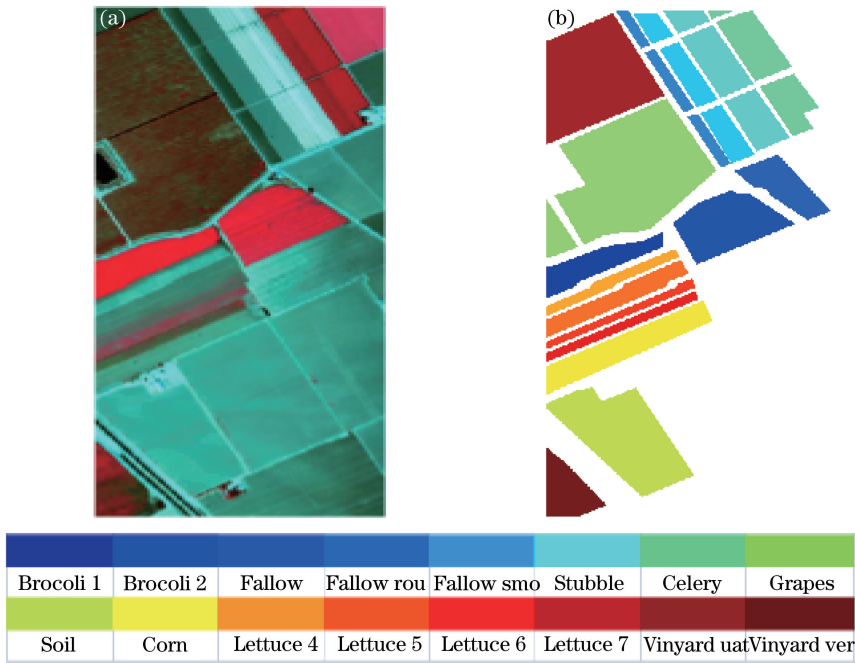


图2 萨利纳斯山谷数据。(a)假彩色图像;(b)真实地物标记

Fig. 2 Salinas Valley dataset. (a) False-color image; (b) ground truth

与调节。如表1所示,当主成分个数 $N_{PC}=5$ 时,PCA-LBP-SRC 算法的 OA 最高;当主成分个数 $N_{PC}=6$ 时,PCA-LBP-SVM 算法的 OA 最高,标准差也达到了最小值,说明此时分类性能最稳定。表1中的数据也反映了不同主成分个数 N_{PC} 对运算时间的影响。对于本文两种算法,当主成分个数 N_{PC} 增加时,计算成本持续增大,而 OA 并没有持续升高。这是因为增加主成分个数 N_{PC} 导致光谱特征的信息量增大,但是到一定程度时会引入冗余特征。另外,通过 LBP 对高光谱图像进行空间特征提取时,需要对空间像素进行区域划分,该区域的大小直接决定了所含信息量(即参与的像素个数)的多少。因此,在 PCA 降维后,需要通过调整检测窗口的大小在空间划分图像块来提取 LBP 特征。以印第安松林数据为例,取大小为 $13 \text{ pixel} \times 13 \text{ pixel}$ 的检测窗口来划分图像块。通过分析表2中的数据可知,本文两种算法对参数 R 不敏感,且均在邻域成员个数 $B=10$ 时达到最优 OA 和最小标准差。与 PCA 类似,随着邻域成员个数的增加,OA 先升高后降低,但变化的幅度不大,而计算成本持续增加。因此,在对程序运行时间有一定要求的情况下,可适当减少主成分个数和邻域成员个数,来控制计算成本。后续实验对萨利纳斯山谷数据采用相同的方法来进行参数设置。

表1 参数 N_{PC} 对印第安松林数据分类精度和运算时间的影响Table 1 Classification accuracy and execution time versus N_{PC} for the Indian Pines dataset

Classification algorithm		$N_{PC}=2$	$N_{PC}=3$	$N_{PC}=4$	$N_{PC}=5$	$N_{PC}=6$	$N_{PC}=7$
PCA-LBP-SRC	OA / %	88.82(± 2.30)	88.85(± 2.02)	89.33(± 2.26)	90.66(± 1.41)	90.28(± 1.89)	90.28(± 2.25)
	Time / s	10.67	14.67	18.46	23.97	27.92	31.69
PCA-LBP-SVM	OA / %	89.46(± 2.33)	89.40(± 1.95)	89.55(± 2.06)	90.17(± 1.58)	91.36(± 1.05)	90.10(± 2.67)
	Time / s	4.96	8.30	13.37	20.15	25.15	30.32

3.3 实验结果与分析

对印第安松林数据中的每个类别随机选择 50 个非零像素作为训练样本,将所有非零像素作为测试样本,测试样本总数为 8624 个,实验数据为 20 次实验结果的平均值。表3对本文算法和比较算法的 OA、Kappa 系数及运算时间进行了比较,所有算法的分类结果均在各自的最优参数下获得。当 $N_{PC}=5$ 和 $(B, R)=(10, 1)$ 时,PCA-LBP-SRC 算法的分类精度和 Kappa 系数最高,分别达到 96.20% 和 0.9546,比传统的 SRC 算法的分类精度高出了约 17%;当 $N_{PC}=6$ 和 $(B, R)=(10, 1)$ 时,PCA-LBP-SVM 算法的分类精度和 Kappa 系数分别达到了 96.13% 和 0.9536,比传统的 SVM^[9] 算法的分类精度提高了将近 20%,比 SVM-

MRF^[10]算法的分类精度提高了8%以上。同时,本文两种算法相比文献[12]LPNMF-G-MRF算法的分类精度提高了8%左右,比文献[13]Gabor-LFDA-GMM算法和Gabor-LPNMF-GMM算法的分类精度也提高了4%以上。其中,PCA-LBP-SVM算法不但能得到较高的分类精度,而且运算成本相对较低。另外,本文算法的分类精度标准差最小,说明分类性能更稳定。图3为上述几种算法对印第安松林高光谱数据的分类图。

表2 参数B和R对印第安松林数据分类精度和运算时间的影响

Table 2 Classification accuracy and execution time versus B and R for the Indian Pines dataset

Classification algorithm			B=4	B=6	B=8	B=10	B=12
PCA-LBP-SRC	OA / %	R=1	88.10(±2.18)	88.69(±1.78)	89.41(±1.63)	90.66(±1.42)	89.90(±2.24)
		R=2	88.77(±1.44)	88.98(±1.87)	89.79(±1.50)	89.25(±1.52)	89.70(±1.67)
		R=3	88.29(±1.98)	89.30(±1.83)	89.74(±2.0)	89.22(±2.09)	89.58(±1.13)
	Time / s		12.38	14.62	18.20	22.10	28.39
PCA-LBP-SVM	OA / %	R=1	89.14(±2.07)	89.20(±2.33)	89.79(±2.32)	90.79(±1.14)	89.45(±1.28)
		R=2	88.70(±2.58)	89.71(±2.23)	89.53(±3.34)	89.60(±1.55)	90.10(±1.09)
		R=1	89.01(±1.56)	89.79(±1.96)	89.04(±2.15)	88.87(±2.40)	89.00(±1.28)
	Time / s		10.06	10.74	14.72	24.40	42.74

表3 印第安松林数据的分类结果

Table 3 Classification results for the Indian Pines dataset

Classification algorithm	OA / %	Kappa	Time / s
LFDA-GMM	64.17(±2.04)	0.5767(±0.0218)	14.82
LFDA-G-MRF	82.09(±2.89)	0.7840(±0.0351)	21.91
LPNMF-G-MRF	88.17(±2.59)	0.8591(±0.0303)	28.76
Gabor-LFDA-GMM	87.08(±1.71)	0.8439(±0.0210)	43.36
Gabor-LPNMF-GMM	91.88(±1.15)	0.9026(±0.0136)	60.19
SRC	79.23(±1.00)	0.7553(±0.0112)	35.51
PCA-LBP-SRC	96.20(±0.57)	0.9546(±0.0068)	78.14
SVM	77.46(±1.71)	0.7337(±0.0195)	3.46
SVM-MRF	88.04(±1.14)	0.8578(±0.0135)	4.04
PCA-LBP-SVM	96.13(±0.85)	0.9536(±0.0102)	27.00

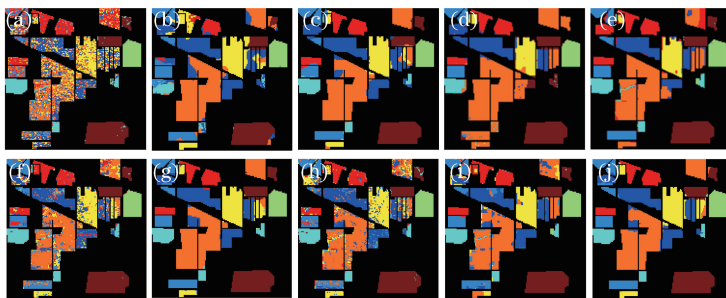


图3 印第安松林数据的分类图。(a) LFDA-GMM; (b) Gabor-LFDA-GMM; (c) Gabor-LPNMF-GMM; (d) LFDA-G-MRF; (e) LPNMF-G-MRF; (f) SRC; (g) PCA-LBP-SRC; (h) SVM; (i) SVM-MRF; (j) PCA-LBP-SVM

Fig. 3 Classification maps for the Indian Pines dataset. (a) LFDA-GMM; (b) Gabor-LFDA-GMM; (c) Gabor-LPNMF-GMM; (d) LFDA-G-MRF; (e) LPNMF-G-MRF; (f) SRC; (g) PCA-LBP-SRC; (h) SVM; (i) SVM-MRF; (j) PCA-LBP-SVM

对萨利纳斯山谷数据以相同的方法进行实验测试,每个类别随机选择50个非零像素作为训练样本,测试样本总数为54129个。根据实验数据分析,取大小为17 pixel×17 pixel的检测窗口来划分对该数据划分图像块,能得到较好的分类结果。通过分类精度、Kappa系数和运算时间将本文算法同其他算法进行了比较,如表4所示,所有算法的分类结果均在各自的最优参数下获得。PCA-LBP-SVM算法的参数设置为

$N_{PC}=4$ 和 $(B,R)=(8,1)$,分类精度和 Kappa 系数分别达到 97.82%和 0.9751,比传统的 SRC 算法的分类精度高出了 9%; PCA-LBP-SVM 算法的参数设置为 $N_{PC}=4$ 和 $(B,R)=(10,1)$ 时,分类精度和 Kappa 系数分别达到了 98.42%和 0.9825,比传统的 SVM 算法^[9]的分类精度高出了 9%,比 SVM-MRF 算法^[10]的分类精度高出了 5%以上。本文两种算法比文献[12]提出的 LPNMF-G-MRF 算法的分类精度提高了约 5%,比文献[13]提出的 Gabor-LPNMF-GMM 算法的分类精度也提高了 2%以上。对于萨利纳斯山谷数据来说,本文算法的分类精度标准差依然是最小的。与印第安松林数据类似,PCA-LBP-SVM 算法不但能得到最优的分类精度和 Kappa 系数,而且大大降低了运算成本。图 4 为本文算法和其他比较算法对萨利纳斯山谷数据的分类图。

表 4 萨利纳斯山谷数据的分类结果
Table 4 Classification results for the Salinas Valley dataset

Classification algorithm	OA /%	Kappa	Time /s
LFDA-GMM	86.52(± 0.94)	0.8503(± 0.0103)	169.13
LFDA-G-MRF	88.33(± 1.85)	0.8705(± 0.0207)	336.72
LPNMF-G-MRF	93.37(± 0.96)	0.9265(± 0.0109)	202.98
Gabor-LFDA-GMM	93.48(± 1.42)	0.9271(± 0.0160)	648.60
Gabor-LPNMF-GMM	95.66(± 0.61)	0.9517(± 0.0068)	483.81
SRC	88.35(± 0.67)	0.8705(± 0.0074)	824.12
PCA-LBP-SRC	97.82(± 0.35)	0.9751(± 0.0039)	455.29
SVM	89.30(± 1.09)	0.8812(± 0.0120)	10.16
SVM-MRF	93.93(± 0.98)	0.9324(± 0.0110)	12.39
PCA-LBP-SVM	98.42(± 0.37)	0.9825(± 0.0041)	90.26

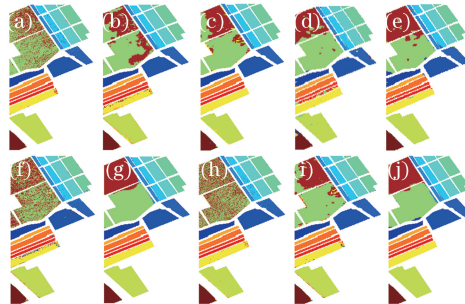


图 4 萨利纳斯山谷数据的分类图。(a) LFDA-GMM; (b) Gabor-LFDA-GMM; (c) Gabor-LPNMF-GMM; (d) LFDA-G-MRF; (e) LPNMF-G-MRF; (f) SRC; (g) PCA-LBP-SRC; (h) SVM; (i) SVM-MRF; (j) PCA-LBP-SVM

Fig. 4 Classification maps for the Salinas Valley dataset. (a) LFDA-GMM; (b) Gabor-LFDA-GMM; (c) Gabor-LPNMF-GMM; (d) LFDA-G-MRF; (e) LPNMF-G-MRF; (f) SRC; (g) PCA-LBP-SRC; (h) SVM; (i) SVM-MRF; (j) PCA-LBP-SVM

在噪声环境中的分类性能,常常作为遥感图像分类算法的衡量标准之一。在光学遥感成像模型中,一般采用加性高斯噪声来模拟遥感图像中的噪声^[19]。在实验中,对原始高光谱遥感图像添加不同幅度的高斯噪声,得到具有不同信噪比(SNR)的带噪图像。印第安松林数据添加噪声后的信噪比范围为 9.30 dB~36.36 dB,萨利纳斯山谷数据添加噪声后的信噪比范围为 6.38 dB~32.69 dB。由于添加的高斯噪声为随机噪声,在柱形图中给出了各个算法分类结果的误差线,且该实验中的所有误差线均在 95%的置信区间下获得。如图 5(a)所示,对于印第安松林数据来说,随着噪声强度的增加,传统的 SRC、SVM 和 SVM-MRF 算法^[10]的分类精度快速下降,而 Gabor-LPNMF-GMM 算法^[12]和本文两种算法的分类精度受其影响较小,其中本文算法的分类精度略高于 Gabor-LPNMF-GMM 算法。由图 5(b)可知,对于萨利纳斯山谷数据来说,本文两种分类算法的抗噪声性能依然优于其他比较算法,其中 PCA-LBP-SVM 算法得到的分类精度始终最高。通过该实验分析可知,所提出的基于 PCA 与 LBP 的分类算法具有良好的抗噪声性能。

为了验证本文算法在小样本情况下的分类性能,对印第安松林数据和萨利纳斯山谷数据的每个类从 10~50 选择训练样本的个数进行测试。如图 6 所示,在不同训练样本个数的情况下,PCA-LBP-SRC 算法始

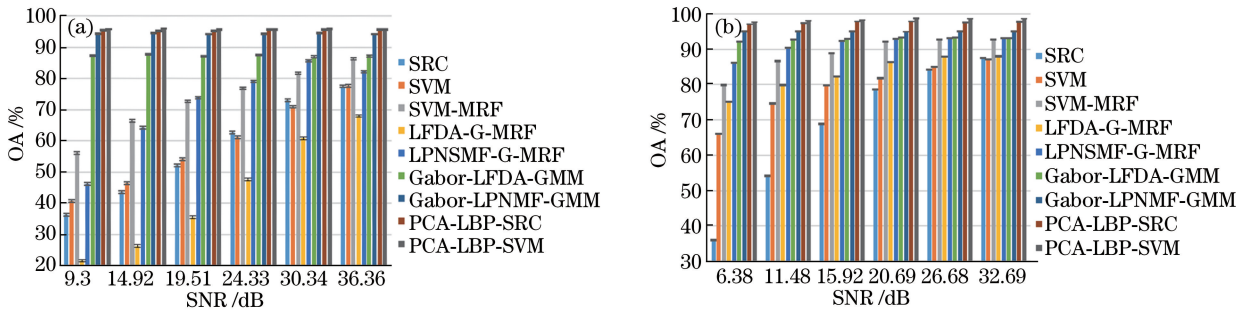


图 5 在噪声环境中的分类结果。(a)印第安松林数据；(b)萨利纳斯山谷数据

Fig. 5 Classification results in a noisy environment. (a) Indian Pines dataset; (b) Salinas Valley dataset

终比传统的 SRC 算法的分类精度有较大幅度的提高。如图 6(a) 所示,即使在各类训练样本为 10 时,PCA-LBP-SRC 算法对印第安松林数据的分类精度能达到 82.5%;如图 6(b) 所示,PCA-LBP-SRC 算法对萨利纳斯山谷数据的分类精度能达到 90.3%。由图 7 可知,在不同训练样本个数的情况下,PCA-LBP-SVM 算法也比传统的 SVM 算法的分类精度有较大幅度的提高。对比图 6 和图 7 的分类结果可知,在小样本情况下,本文算法对印第安松林数据的分类优势尤为突出,且具有较好的稳定性,当各类训练样本数为 10 时,比传统的 SRC 算法、SVM 算法的分类精度均高了约 18%。

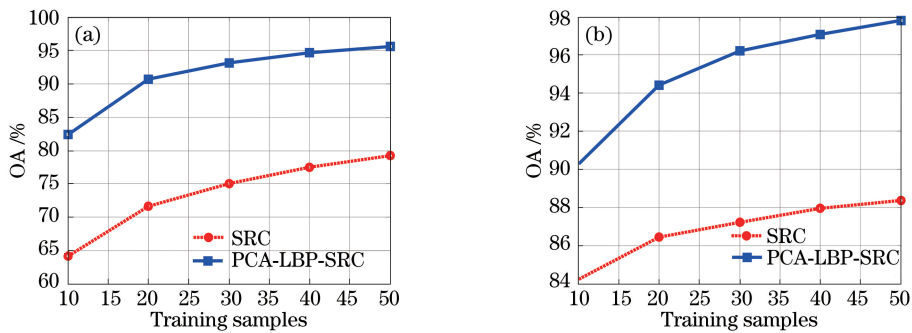


图 6 基于 SRC 的分类算法在不同训练样本个数下的分类性能。(a)印第安松林数据；(b)萨利纳斯山谷数据

Fig. 6 Classification performances of SRC-based algorithms with different numbers of training sample.

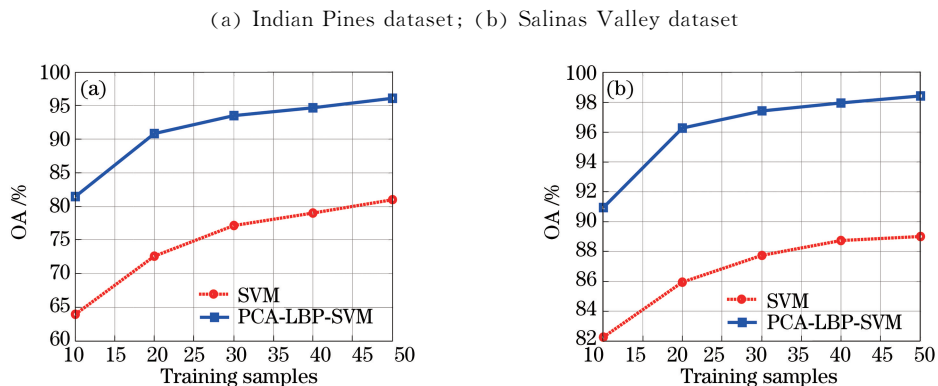


图 7 基于 SVM 的分类算法在不同训练样本个数下的分类性能。(a)印第安松林数据；(b)萨利纳斯山谷数据

Fig. 7 Classification performances of SVM-based algorithm with different numbers of training sample.

(a) Indian Pines dataset; (b) Salinas Valley dataset

4 结 论

近年来,基于谱间-空间特征的分类算法渐渐成为高光谱图像分类领域的研究热点之一,该类算法通常比基于谱间特征的传统算法能得到更好的分类结果。但是,对高光谱图像在空间进行特征挖掘,往往会增加

算法的复杂度,具有较高的计算成本。针对该问题,研究了一种模型简单、复杂度低的谱间-空间特征提取方法,具有一定的理论与实际应用价值。该特征提取方法是利用 PCA 对高光谱图像进行谱间降维,在不同主成分的投影子空间中,利用 LBP 在圆形邻域内和任意空间分辨率下高效的挖掘局部空间纹理特征。分别结合 SRC 和 SVM 分类器,提出了两种基于 PCA 和 LBP 的高光谱图像分类算法,命名为 PCA-LBP-SRC 和 PCA-LBP-SVM。实验结果表明,本文算法不但能有效提高分类精度和 Kappa 系数,而且在高斯噪声环境中和小样本情况下也具有较好的分类性能。尤其 PCA-LBP-SVM 算法是一种计算复杂度低且性能优良的高光谱图像分类算法。

参 考 文 献

- [1] Liu Dawei, Han Ling, Han Xiaoyong. High spatial resolution remote sensing image classification based on deep learning[J]. *Acta Optica Sinica*, 2016, 36(4): 0428001.
刘大伟, 韩玲, 韩晓勇. 基于深度学习的高分辨率遥感影像分类研究[J]. *光学学报*, 2016, 36(4): 0428001.
- [2] Li Tie, Sun Jinguang, Zhang Xinjun, *et al.* Research of hyperspectral image classification based on hierarchical sparse representation feature learning[J]. *Laser & Optoelectronics Progress*, 2016, 53(9): 091001.
李铁, 孙劲光, 张新君, 等. 基于分层稀疏表示特征学习的高光谱图像分类研究[J]. *激光与光电子学进展*, 2016, 53(9): 091001.
- [3] Prasad S, Bruce L M. Limitations of principal components analysis for hyperspectral target recognition[J]. *IEEE Geoscience and Remote Sensing Letters*, 2008, 5(4): 625-629.
- [4] Duda R O, Hart P E, Stork D G. *Pattern Classification*[M]. 2nd ed. New York: John Wiley & Sons, Inc., 2001.
- [5] Xue Z, Li J, Cheng L, *et al.* Spectral-spatial classification of hyperspectral data via morphological component analysis-based image separation[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2015, 53(1): 70-84.
- [6] Ye Z, Prasad S, Li W, *et al.* Classification based on 3D DWT and decision fusion for hyperspectral image analysis[J]. *IEEE Geoscience and Remote Sensing Letters*, 2014, 11(1): 173-177.
- [7] Zhang L, Zhang L, Tao D, *et al.* On combining multiple features for hyperspectral remote sensing image classification [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2012, 50(3): 879-893.
- [8] Sun L, Wu Z, Liu J, *et al.* Supervised spectral-spatial hypersepectral image classification with weighted Markov random fields[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2015, 53(3): 1490-1503.
- [9] Camps-Valls G, Bruzzone L. Kernel-based methods for hyperspectral image classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2005, 43(6): 1351-1362.
- [10] Tarabalka Y, Fauvel M, Chanussot J, *et al.* SVM- and MRF-based method for accurate classification of hyperspectral images[J]. *IEEE Geoscience and Remote Sensing Letters*, 2010, 7(4): 736-740.
- [11] Li W, Prasad S, Fowler J E, *et al.* Locality-preserving dimensionality reduction and classification for hyperspectral image analysis[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2012, 50(4): 1185-1198.
- [12] Li W, Prasad S, J. E. Fowler. Hyperspectral image classification using Gaussian mixture models and Markov random fields[J]. *IEEE Geoscience and Remote Sensing Letters*, 2014, 11(1): 153-157.
- [13] Ye Zhen, Bai Lin, Nian Yongjian. Hyperspectral image classification based on Gabor features and locality-preserving dimensionality reduction[J]. *Acta Optica Sinica*, 2016, 36(10): 1028003.
叶珍, 白麟, 粘永健. 基于 Gabor 特征与局部保护降维的高光谱图像分类[J]. *光学学报*, 2016, 36(10): 1028003.
- [14] Ojala T, Pietikainen M, Maenpaa T T. Multiresolution gray-scale and rotation invariant texture classification with local binary pattern[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2002, 24(7): 971-987.
- [15] Chen Y, Nasrabadi N M, Tran T D. Hyperspectral image classification using dictionary-based sparse representation [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2011, 49(10): 3973-3985.
- [16] Cheng Baozhi, Zhao Chunhui, Zhang Lili, *et al.* Joint spatial preprocessing and spectral clustering based collaborative sparsity anomaly detection for hyperspectral images[J]. *Acta Optica Sinica*, 2017, 37(4): 0428001.
成宝芝, 赵春晖, 张丽丽, 等. 联合空间预处理与谱聚类的协同稀疏高光谱异常检测 [J]. *光学学报*, 2017, 37(4): 0428001.
- [17] Li C, Ma Y, Mei X, *et al.* Hyperspectral image classification with robust sparse representation[J]. *IEEE Geoscience*

and Remote Sensing Letters. 2016, 13(5): 641-645.

- [18] Song Xiangfa, Jiao Licheng. Classification of hyperspectral remote sensing image based on sparse representation and spectral information[J]. Journal of Electronics & Information Technology, 2012, 34(2): 268-272.
宋相法, 焦李成. 基于稀疏表示及光谱信息的高光谱遥感图像分类[J]. 电子与信息学报, 2012, 34(2): 268-272.
- [19] Pu H Y, Chen Z, Wang B, *et al.* A novel spatial-spectral similarity measure for dimensionality reduction and classification of hyperspectral imagery[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52(11): 7008-7022.