

文章编号 1004-924X(2017)01-0263-11

加权空谱局部保持投影的高光谱图像特征提取

何 芳, 王 榕*, 于 强, 贾维敏

(火箭军工程大学 信息与通信工程系, 陕西 西安 710025)

摘要: 为了提高高光谱图像的分类精度, 有效利用高光谱图像的空间信息和光谱信息对高光谱图像进行预处理, 本文提出了一种新的空谱联合特征提取方法, 加权空-谱局部保持投影算法(WSSLPP)。该算法结合高光谱图像的物理特性对高光谱图像进行重构, 降低了图像中奇异点的干扰; 然后对局部像素近邻保持嵌入(LPNPE)和局部保持投影(LPP)的目标函数进行加权求和, 有效融合高光谱图像空间维和光谱维的信息来构建投影矩阵。WSSLPP 不仅保留了高光谱图像在空间维上像素间的近邻关系, 而且保持了在光谱维上样本的固有结构, 有利于高光谱图像的分类。在 Indian Pines 和 PaviU 数据库上对该算法进行验证分析, 结果表明: 基于 WSSLPP 算法得到的分类精度明显高于其他算法, 总体分类精度的最大值分别为 99.00%, 99.50%, 有效提高了高光谱图像的分类精度。

关键词: 空间信息; 光谱信息; 加权空-谱局部保持投影; 高光谱图像; 特征提取

中图分类号: TP751 文献标识码: A doi: 10.3788/OPE.20172501.0263

Feature Extraction of Hyperspectral Images of Weighted Spatial and Spectral Locality Preserving Projection (WSSLPP)

HE Fang, WANG Rong*, YU Qiang, JIA Wei-min

(Department of Information and Communication Engineering,
Rocket Force Engineering University, Xi'an 710025, China)

* Corresponding author, E-mail: wangrong07@tsinghua.org.cn

Abstract: In order to improve the classification accuracy of Hyperspectral images (HSI) and preprocess HSI by effectively using the spatial and spectral information of HIS, a new spatial-spectral feature extraction method, Weighted Spatial and Spectral Locality Preserving Projection (WSSLPP) is proposed in this paper. The HSI was reconstructed combining the physical characters of HSI to avoid the interference of singular point; then the target functions of locality pixel neighbor preserving embedding (LPNPE) and locality preserving projection (LPP) were weighted and summed, thus the spatial and spectral dimension information of HSI was effectively fused to construct the projection matrix. WSSLPP not only keeps the pixel neighborhood in spatial domain, but also keeps the implicit structure of samples in spectral domain, which helps for the HIS classification. The benchmark verification on Indian Pines and PaviU database show that the classification accuracy resulted from WSSLPP algorithm is significantly higher than that from other algorithms, and the overall classification

收稿日期: 2016-08-04; 修订日期: 2016-10-01.

基金项目: 国家自然科学基金资助项目(No. 61401471); 中国博士后基金资助项目(No. 2014M562636).

accuracy is 99.00% and 99.50% respectively, effectively improving the HSI classification accuracy.

Key words: spatial information; spectral information; Weighted Spatial and Spectral Locality Preserving Projection (WSSLPP); Hyperspectral images (HSI); feature extraction

1 引言

高光谱遥感图像的空间分辨率高,地物光谱曲线连续^[1],具有较强的地物分类和识别能力^[2],在地质勘查、医疗检测、生命科学、司法鉴别、军事侦察、环境监测、精细农业等领域得到了广泛的应用^[3-5]。高光谱图像是由成像光谱仪获取的遥感图像^[6],随着成像光谱仪空间和谱间分辨率指标的提高,同时具有二维空间和一维光谱信息的高光谱遥感图像数据量呈指数增加^[7]。数据量的急剧膨胀不但给数据存储与传输带来了巨大的困难,同时也加剧了数据处理过程的复杂性,降低了数据处理效率^[8],甚至会出现维数灾难问题^[9-10]。降维可以为高维数据寻找一个包含样本重要信息的低维表示方法,是高光谱图像分类预处理的重要环节^[8]。降维不仅可以保存数据的有用信息,并且能够大大减少数据量,而且可以有效避免维数灾难^[1],使数据表示更加简单清晰,有利于后续分类^[11]。

近年来,应用较广泛的降维方法主要有线性判别分析(Linear Discriminant Analysis, LDA)^[12]和非参数加权特征提取^[13](Nonparametric Weighted Feature Extraction, NWFE)、主成分分析^[14-16](Principle Component Analysis, PCA)、局部保持投影(Locality Preserving Projection, LPP)、近邻保持嵌入(Neighborhood Preserving Embedding, NPE)等。LDA 和 NWFE 均通过构造合适的散度矩阵来有效测量类之间的可分性。PCA 是基于最小均方误差准则和二阶统计方法^[17]寻找一个投影方向,使投影后数据间的方差最小。但它属于全局降维方法,隐藏在低维数据中的内在属性较难被发现,故无法揭示高光谱数据的低维流形结构。LPP 和 NPE 等局部方法通过近邻图保持了原始数据的局部近邻结构^[18-19]。但是,以上降维算法只单纯利用了高光谱遥感数据的光谱信息而忽略了空间信息。对于高光谱图像而言,基于光谱相似性的衡量方法并不能有效揭示不同样本间的

内在关系。由于两个光谱值相近的样本可能空间距离较大,而高光谱图像样本的分布在空间上具有连续性,故仅基于光谱相似性的投影和分类可能会导致错分或漏分现象^[8]。因此,在衡量样本间的相似性和学习判别投影时应该考虑高光谱图像元间的空间相互关系。

高光谱图像空间信息方法可以分为空间滤波和特征提取两种。空间滤波的方法包括空间滤波预处理方法和空间滤波后处理方法。空间滤波预处理是在分类前对高光谱图像空间连续区域进行滤波预处理^[20-21]。空间滤波后处理是指分类完成后对高光谱图像的像素进行处理^[22]。特征提取方法包括基于均值漂移的目标提取方法^[23],结构特征集方法^[24]和在高光谱图像分类中用于提取图像主要成分的轮廓信息的形态学变换方法^[25]。例如:马尔科夫随机场通过建立空间近邻像素间的相互关系将高光谱图像的空间信息放入概率论的框架中^[26-27]。基于复合核的支持向量机(Support Vector Machine, SVM)方法,将高光谱图像的空间特征表示为空间近邻像素的均值或标准差向量,用于构造 SVM 核。多尺度模型在高光谱图像的分类中通过集成高光谱图像的多重特征,包括光谱特征,结构特征和语义特征,来建立 SVM 分类器^[28]。

现有的大多数方法是在高光谱图像的分类过程中使用高光谱图像的空间信息,本文利用高光谱图像“图谱合一”的特点^[29],将其空间信息和光谱信息同时用于高光谱图像降维,提出了一种新的降维算法——加权空-谱局部保持投影(Weighted Spatial and Spectral Locality Preserving Projection, WSSLPP)算法。该方法利用高光谱图像的物理特性,通过引入近邻窗口和光谱因子两个参数对高光谱图像进行重构,以降低奇异点的干扰,实现高光谱图像的平滑处理;然后对近邻保持矩阵和 LPP 的投影矩阵进行加权求和来构建最终的投影矩阵,由于充分利用高光谱图像的空间信息和光谱信息,从而提高了分类精度。

2 加权空-谱算法

在高光谱图像中,设样本点 \mathbf{x}_i 的像素坐标为 (p_i, q_i) ,则像元 \mathbf{x}_i 的近邻空间为:

$$N(\mathbf{x}_i) = \{x \triangle (p, q)\},$$

$$p \in [p_i - a, p_i + a], q \in [q_i - a, q_i + a], \quad (1)$$

其中: $a = (\omega - 1)/2$, ω 为 \mathbf{x}_i 的近邻窗口的宽度,即尺度。近邻空间 $N(\mathbf{x}_i)$ 中的像素点定义为 $\mathbf{x}_i, \mathbf{x}_{i_1}, \mathbf{x}_{i_2}, \dots, \mathbf{x}_{i_s}$, 则 \mathbf{x}_i 的近邻点个数为 $s = \omega^2 - 1$ 。

采用如下规则对像素点 \mathbf{x}_i 进行重构:

$$\hat{\mathbf{x}}_i = \frac{\sum_{\mathbf{x}_j \in N(\mathbf{x}_i)} \nu_j \mathbf{x}_j}{\sum_{\mathbf{x}_j \in N(\mathbf{x}_i)} \nu_j} = \frac{\mathbf{x}_i + \sum_{k=1}^{\omega^2-1} \nu_k \mathbf{x}_{ik}}{1 + \sum_{k=1}^{\omega^2-1} \nu_k}, \quad (2)$$

其中: $\nu_k = \exp\{-\gamma_0 \|\mathbf{x}_i - \mathbf{x}_{ik}\|^2\}$ 为近邻空间 $N(\mathbf{x}_i)$

中任一像元 \mathbf{x}_{ik} 到中心像元 \mathbf{x}_i 的权重。参数 γ_0 为光谱因子,衡量同一近邻空间中不同像元间的相互影响程度^[30]。同时,光谱因子通过为近邻空间中与中心像元差异较大的点赋予较小的权重来降低奇异点的干扰,实现对高光谱图像的平滑处理。

近邻窗口尺度 ω 可以调节近邻空间的大小,光谱因子 γ_0 可以调整像元间的相互影响程度。因此,参数 ω 和 γ_0 的选取直接影响高光谱图像处理的结果,本文通过实验分析的方法进行参数选取。

对于位于图像边缘或角落的像元,采用图 1 方法进行预处理。图中,每一个正方形格子代表高光谱图像中的一个像元,浅灰色格子为中心像元,深灰色格子为填补方式。对于位于四周的像素点,直接取与其近邻的像素补充,4 个角的像素用中心像元进行填补。

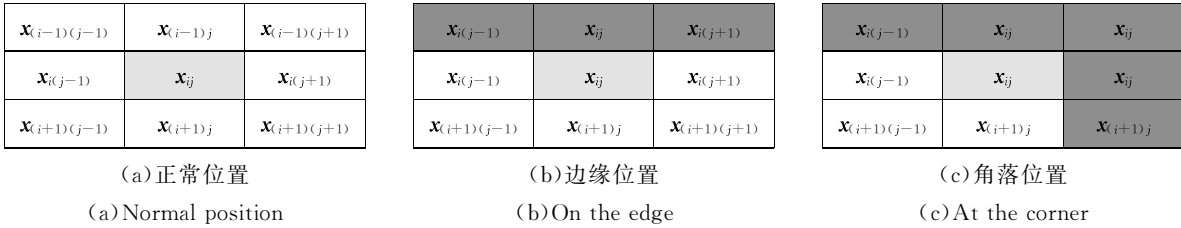


图 1 不同情况下像元点的近邻空间

Fig. 1 Neighbor space of pixel under different conditions

3 加权空-谱局部保持投影算法

3.1 局部像素近邻保持嵌入算法

根据高光谱图像空间分布连续性的特点,局部像素近邻保持嵌入(Local Pixel Neighborhood Preserving Projection, LPNPE)算法^[7]利用样本的空间信息得到判别映射,使原始空间中的近邻像素点经过 LPNPE 映射后仍然能保持近邻关系。

基于空间的 LPNPE 在高光谱图像分类中过程如下:高光谱图像可以表示为 $p \times (m \times n)$ 的矩阵 \mathbf{X} ,其中 p 为波段数, m, n 分别为每个波段上的图像的行列数,局部像素块 $P_i = \{\mathbf{X}_i, \mathbf{X}_{i_1}, \mathbf{X}_{i_2}, \dots, \mathbf{X}_{i_s}\}$ 由训练样本 \mathbf{X}_i 及其近邻像素点 $N(\mathbf{X}_i)$ 组成,则局部像素块的距离散度矩

阵为:

$$\mathbf{h}_i = \sum_{k=1}^s \frac{\nu_k}{\sum_{j=1}^s \nu_j} (\mathbf{X}_i - \mathbf{X}_{ik}) (\mathbf{X}_i - \mathbf{X}_{ik})^T. \quad (3)$$

因此,所有样本的局部像素近邻保持矩阵为:

$$\mathbf{H} = \sum_{i=1}^m \mathbf{h}_i = \sum_{i=1}^m \sum_{k=1}^s \frac{\nu_k}{\sum_{j=1}^s \nu_j} (\mathbf{X}_i - \mathbf{X}_{ik}) (\mathbf{X}_i - \mathbf{X}_{ik})^T, \quad (4)$$

其中: m 为训练样本的个数, s 为 \mathbf{x}_i 近邻点的个数, ν_j 为近邻空间 $N(\mathbf{x}_i)$ 中任一像元 \mathbf{x}_{ik} 到中心像元 \mathbf{x}_i 的权重。

定义总体散度矩阵为:

$$\mathbf{S} = \sum_{i=1}^m (\mathbf{X}_i - \mathbf{m}) (\mathbf{X}_i - \mathbf{m})^T = \overline{\mathbf{X}\mathbf{X}^T}, \quad (5)$$

其中 \mathbf{m} 为所有训练样本的均值。

LPNPE 的目标是寻找一个线性投影矩阵,

使投影后样本的局部像素的近邻保持散度最小, 总体散度矩阵最大。因此, 最佳的投影矩阵为 $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_l]$, 其可通过如下广义特征值分解进行求解:

$$\mathbf{S}\mathbf{v} = \lambda\mathbf{H}\mathbf{v}. \quad (6)$$

LPNPE 仅利用了高光谱图像的空间信息来保留空间局部像素的近邻结构, 并没有使用有标签的光谱信息, 实际上是一种无监督的降维方法。如果在光谱域中两个已知标签的样本在同一均匀区域中的像素距离较大, 则 LPNPE 也不能将这两点直接联系起来。因此, 有标签的样本信息推动了光谱间的近邻联系, 有助于提高分类能力。

3.2 局部保持投影算法

LPP 算法的目的是寻找最优的投影矩阵 \mathbf{A} , 将存在于高维空间中的数据投影到低维空间中。原始高维空间中的数据点 x_i 和 x_j 为近邻点, 经过 LPP 算法映射在低维空间后, 对应的低维数据点 y_i 和 y_j 同样保持先前的近邻关系, 其目标函数为:

$$\begin{cases} \arg \min \mathbf{a}^T \mathbf{X}\mathbf{L}\mathbf{X}^T \mathbf{a} \\ \text{s. t. } \mathbf{a}^T \mathbf{X}\mathbf{D}\mathbf{X}^T \mathbf{a} = \mathbf{I} \end{cases}, \quad (7)$$

其中: \mathbf{D} 为对角矩阵, 且 $D_{ii} = \sum W_{ij}$:

$$W_{ij} = \begin{cases} e^{-\frac{\|x_i - x_j\|^2}{\sigma}}, & \text{if } x_i \in N_k(x_j) \text{ or } x_j \in N_k(x_i) \\ 0, & \text{otherwise} \end{cases},$$

设 \mathbf{L} 为拉普拉斯矩阵, $\mathbf{L} = \mathbf{D} - \mathbf{W}$ 。LPP 算法对其目标函数的求解可转换为如下广义特征值的求解:

$$\mathbf{X}\mathbf{L}\mathbf{X}^T \mathbf{a} = \lambda \mathbf{X}\mathbf{D}\mathbf{X}^T \mathbf{a}. \quad (8)$$

将求得特征值按照从小到大的顺序进行排序, 选取最小的前 m 个特征值对应的特征向量构成特征向量组 $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_l]^T$, 对原始高维数据集 \mathbf{X} 进行低维投影后, 得到低维坐标 $\mathbf{Y} = \mathbf{A}^T \mathbf{X}$ 。

LPP 算法既具有线性降维算法的优点又能克服非线性降维算法的缺点, 但是 LPP 算法只考虑了高光谱图像的光谱维信息, 忽略了高光谱样本在空间上的连续性, 是一种基于光谱维的降维算法。

3.3 加权空-谱局部保持投影算法

近年来, 联合利用高光谱数据的空间信息和光谱信息来实现更有效的降维与分类已经得到众多国内外学者的广泛关注^[10]。本文提出一种加权空-谱局部保持投影 (Weighted Spatial and Spectral Locality Preserving Projection, WSSLPP) 算法。首先采用加权空谱算法对高光谱图像进行重构, 得到新的数据集 $\hat{\mathbf{X}}$, 然后在构造判别投影矩阵时, 既使用已知光谱域的标签信息, 又考虑空间中像素间的相互关系, 将 LPP 和 LPNPE 进行融合, 得到空谱局部近邻的保持散度矩阵:

$$\begin{cases} \arg \min \mathbf{A}^T (r\hat{\mathbf{X}}\mathbf{L}\hat{\mathbf{X}}^T + (1-r)\mathbf{H})\mathbf{A} \\ \text{s. t. } \mathbf{A}^T \hat{\mathbf{X}}\mathbf{D}\hat{\mathbf{X}}^T \mathbf{A} = \mathbf{I} \end{cases}, \quad (9)$$

式中 r 为规则化参数, 可通过实验分析的方法进行选取。

最佳的投影矩阵为 $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_l]$ 可通过如下广义特征值分解进行求解:

$$\hat{\mathbf{X}}\mathbf{D}\hat{\mathbf{X}}^T \mathbf{v} = \lambda (r\hat{\mathbf{X}}\mathbf{L}\hat{\mathbf{X}}^T + (1-r)\mathbf{H}) \mathbf{v}. \quad (10)$$

4 实验验证与结果分析

4.1 实验数据选择

本文选择具有代表性的 Indian Pines^[31] 和 PaviaU^[30] 图像数据库验证本文算法的有效性。

Indian Pines 图像数据库是 1992 年 6 月由 AVIRIS 传感器在美国印第安纳州的一块印度松树测试地获取的。该图像由 $145 \text{ pixel} \times 145 \text{ pixel}$ 组成, 去除水的吸收及噪声影响, 将剩下的 200 个波段用作实验。每个波段图像含有 16 种不同类型作物的像元, 共 10 249 个样本点。图 2(a) (彩图见期刊电子版) 为选取 Indian Pines 的第 50、27、17 波段合成的 RGB 图像, 图 2(b) (彩图见期刊电子版) 为其地面真实数据及相应的图例。

PaviaU 图像数据库是 2001 年由 ROSIS 传感器拍摄 Pavia 大学获得的, 该影像共有 115 个波段, 光谱范围为 $0.43 \sim 0.86 \mu\text{m}$, 空间分辨率为 1.3 m 。PaviaU 图像包括 $610 \text{ pixel} \times 340 \text{ pixel}$, 去除 12 个噪声影响最大的波段, 剩余 103 个波段用于实验分析。该图像共包括 9 个类别的地物。图 3(a) (彩图见期刊电子版) 为选取 PaviaU 的第 50、27、17 波段合成的 RGB 图像, 图 3(b) (彩图见期刊电子版) 为地面真实数据及相应的图例。

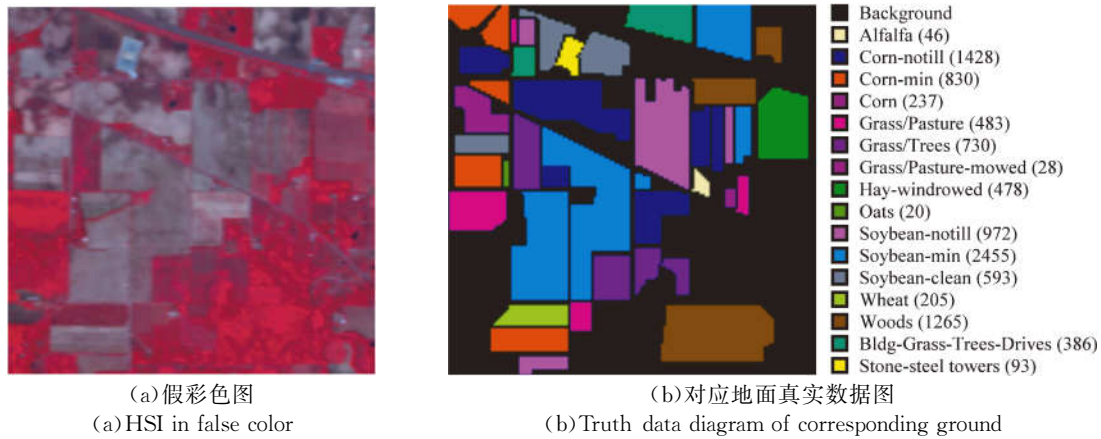


图 2 Indian-pines 图像示意图

Fig. 2 Schematic of Indian-pines hyperspectral image

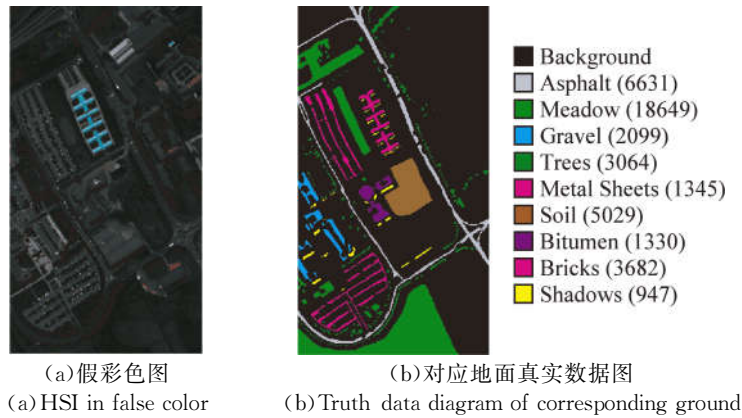


图 3 PaviaU 图像示意图

Fig. 3 Schematic of PaviaU hyperspectral image

4.2 实验设置

本文首先采用 WSSLPP, PCA, LPP 以及 LPNPE 算法对高光谱图像进行降维,然后采用最近邻分类器对降维后的数据进行分类,同时,将不做降维处理直接进行分类的结果作为基准线 Baseline。通过实验分析的方法选取 WSSLPP 的最佳参数,采用热核法构造 LPP 的权重矩阵 \mathbf{W} 。常用的高光谱图像分类精度评价指标有:误差矩阵 (Error Matrix)、总体精度 (Overall Accuracy, OA)、平均精度 (Average Accuracy, AA) 和 Kappa 系数。本文主要选取 OA、AA 和 Kappa 系数用于衡量算法的有效性,OA、AA 和 Kappa 系数越高表明该方法对高光谱图像的分类效果越好。

误差矩阵是衡量高光谱图像分类性能好坏的基本标准之一,是在分类完成后将分类结果图与相应的真实地物标记图进行对比得出的,其公式如下:

$$\mathbf{M} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{bmatrix}, \quad (11)$$

其中: x_{ij} 表示第 j 类的元素被错误地划分到第 i 类元素的总个数, n 为图像中所有地物的类别总数。

OA 指被正确分类的像元数占总像元数百分比的总分类精度, AA 为所有地物的平均分类精度, Kappa 系数反映的是分类后的遥感图像与真实地物标记图的相似程度,计算公式为:

$$Kappa = \frac{N \sum_{i=1}^n x_{ii} - \sum_{i=1}^n (x_{i+} x_{+i})}{N^2 - \sum_{i=1}^n (x_{i+} x_{+i})}, \quad (12)$$

其中: N 为参与分类的总像元, x_{+i} 为误差矩阵中第 i 列之和, x_{i+} 为误差矩阵中第 i 行之和。

4.3 基于 Indian Pines 数据库的高光谱图像分类

WSSLPP 的 3 个主要参数为近邻窗口尺度 ω , 光谱因子 γ_0 和规则化参数 r 。本文通过实验分析不同的 ω, γ_0 和 r 对实验结果的影响, 从中选取最佳的参数。从 Indian Pines 数据库的每类样本中随机选取样本数的 10% (样本数不足 100 时选取 10 个) 作为训练样本, 其余的样本作为测试样本。图 4(a)、(b) 给出了在维数为 16 时, 不同的近邻窗口尺度 ω 和光谱因子 γ_0 对分类精度的影响。图 4(c) 给出了不同的规则化参数 r 对分类

精度 OA 和 Kappa 系数的影响。当分析 ω 对分类精度的影响时, γ_0 设置为 0.2, r 设置为 0.1; 同样地, 分析 γ_0 时, ω 设置为 11, r 设置为 0.1; 分析 r 时, γ_0 设置为 0.5, ω 设置为 11。

可以看出, 近邻窗口尺度、光谱因子和规则化参数的大小对算法的分类性能会产生一定的影响, 它们的取值过大或过小都会降低算法的分类性能。根据图 4 结果, 设置 Indian Pines 数据库上的最佳参数为: $\omega=11, \gamma_0=0.5, r=0.1$ 。

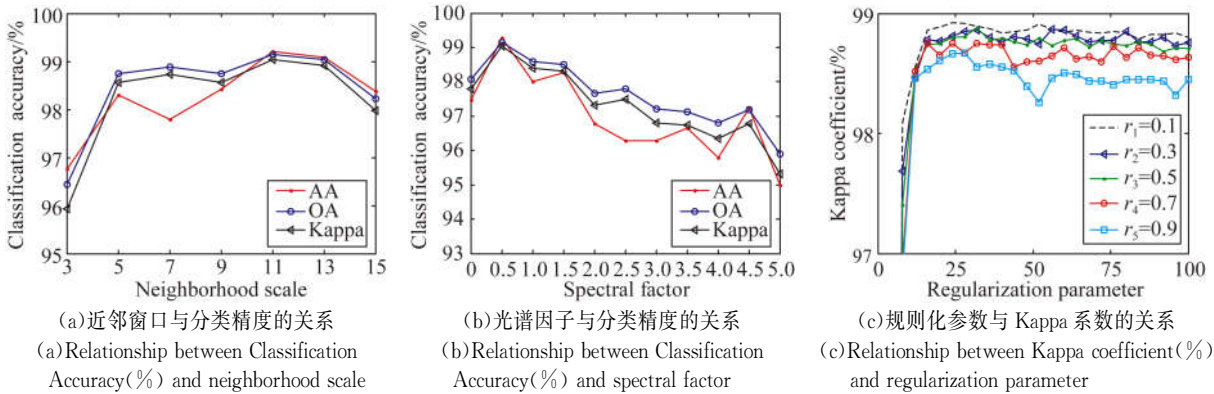


图 4 在 Indian Pines 数据库上分类精度与不同参数的关系

Fig. 4 Relationship between Classification Accuracy(%) and different parameters on Indian Pines dataset

为比较各个算法的性能, 每种算法重复进行 5 次分类实验, 然后对 5 次实验的分类精度求平均值得到最终的 AA、OA 和 Kappa 系数。图 5 给出了在不同算法下, AA、OA 和 Kappa 系数与前 100 维

特征空间维数的关系。表 1 给出了在训练样本相同的情况下, 在 Indian Pines 数据库上由不同算法得到的各类地物的最高分类精度, 平均精度 AA, 总体精度 OA 和 Kappa 系数。

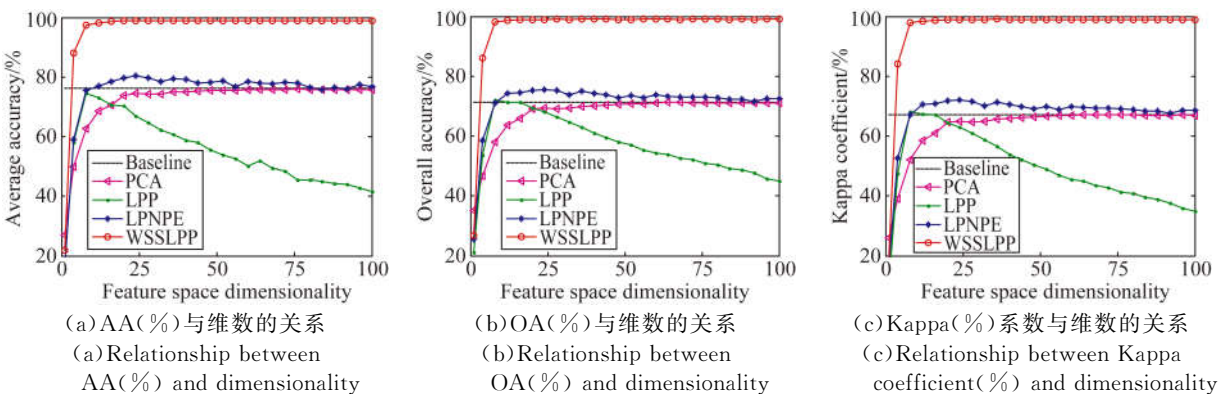


图 5 在 Indian Pines 图像上的分类精度评价指标(%)与维数的关系

Fig. 5 Relationship between classification accuracy evaluation indexes(%) and dimensionality on Indian Pines dataset

由图 5 可知, WSSLPP 算法的分类性能远远优于其他算法。由 PCA、LPP 得到的分类结果均低于 Baseline, LPNPE 的结果略高于 Baseline, 而当维数超过 4 维时, WSSLPP 的分类结果便远远超出了基准线。

由表 1 可知, 由 WSSLPP 算法得到的 OA 最大值达到了 99%, Kappa 系数最大值达到了 0.988 6, 分别较 Baseline 提高了 27.78% 和 0.317 2。这是由于 WSSLPP 算法有效融合了高光谱图像的空间信息和光谱信息对高光谱图像的

像素进行重构,使重构后的像素点类内相似性和类间差异性变大,从而增强了图像的可分性;对近邻散度矩阵和 LPP 的投影矩阵进行加权求和构建判别投影矩阵,从而增强了判别能力;采用实验分析的方法选取了最佳的近邻窗口尺度、光谱因子和规则化参数,从而有效地提高了分类精度。

图 6(彩图见期刊电子版)直观地给出了 Indian Pines 数据库的真实地表,训练样本,测试样本的分布图以及各算法的分类结果,从图中可以看出,与其他算法相比,由 WSSLPP 算法得到的分类结果较好,错分或漏分现象较少。

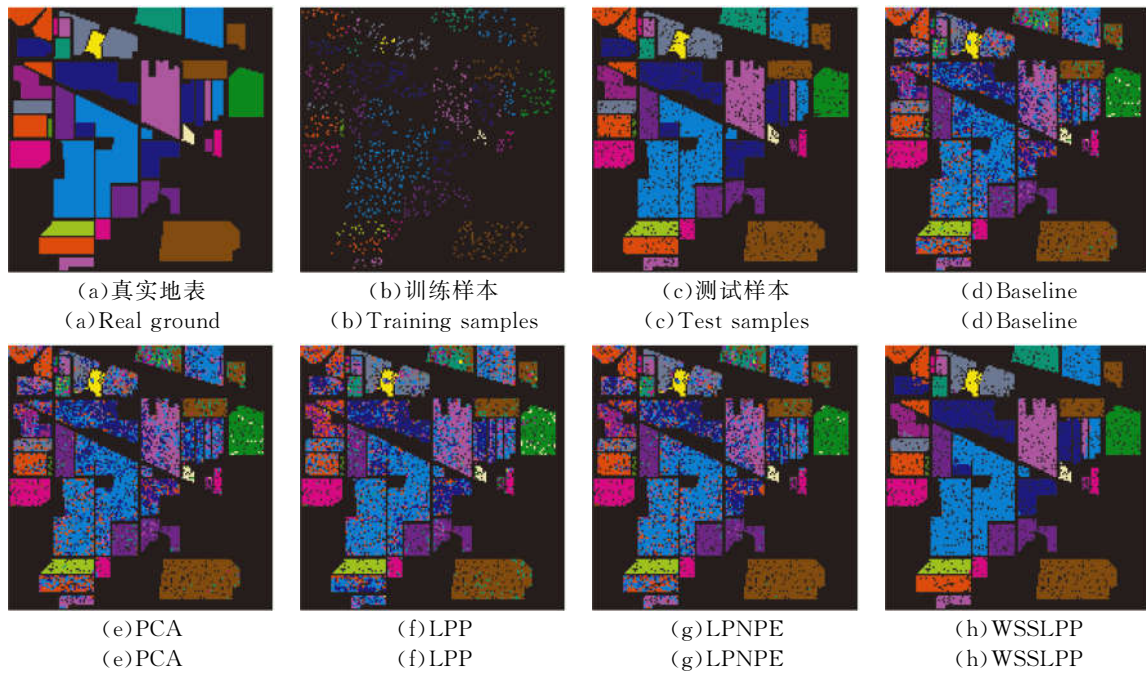


图 6 各算法在 Indian Pines 数据库上的分类结果图

Fig. 6 Classification results of different algorithms in Indian Pines data set

表 1 Indian Pines 数据库上各类地物在不同算法下的最高分类精度

Tab. 1 Classification accuracy of different types of features in Indian Pines data set by different algorithms (%)

地物类别	训练样本个数	测试样本个数	Baseline	PCA	LPP	LPNPE	WSSLPP
Alfalfa	10	36	86.11	86.11	83.33	91.67	100
Corn-notill	143	1 285	53.46	53.23	68.95	59.14	97.51
Corn-min	83	747	64.93	65.19	46.32	66.00	100
Corn	24	213	40.85	40.85	39.44	59.15	96.24
Grass/Pasture	48	435	90.80	90.80	91.49	95.17	100
Grass/Tress	73	657	92.54	92.54	87.82	96.35	99.70
Grass/Pasture-mowed	10	18	83.33	83.33	66.67	83.33	94.44
Hay-windrowed	48	430	95.58	95.58	90.47	97.91	100
Oats	10	10	100	100	90.00	100	100
Soybeans-notill	97	875	65.03	65.03	61.71	65.60	97.37
Soybeans-min	246	2 209	68.49	68.54	68.40	73.20	99.59
Soybeans-clean	59	534	49.25	49.44	64.42	60.86	97.38
Wheat	21	184	97.83	97.83	99.46	99.46	98.91
Woods	127	1 138	94.29	94.29	92.18	96.49	100
Bldg-Grass-Tree-Drives	39	347	42.65	42.36	55.04	54.76	99.71
Stone-steel towers	10	83	93.98	93.98	87.95	92.77	100
平均精度(AA)			76.20	76.19	74.60	80.74	98.80
总体精度(OA)			71.22	71.22	71.97	75.74	99.00
Kappa 系数			0.671 4	0.671 4	0.680 4	0.722 9	0.988 6

4.4 基于 PaviaU 数据库的高光谱图像分类

本实验从 PaviaU 数据库中选取每一类样本数量的 5% 作为训练样本,剩下的所有样本作为测试样本。采用与 Indian Pines 相同的参数检验

方法,得到在 PaviaU 数据库上的最佳参数取值为: $\omega=15, \gamma_0=1.5, r=0.1$ 。

图 7 给出了在不同算法下,AA、OA 和 Kappa 系数与前 40 维特征空间维数的关系。

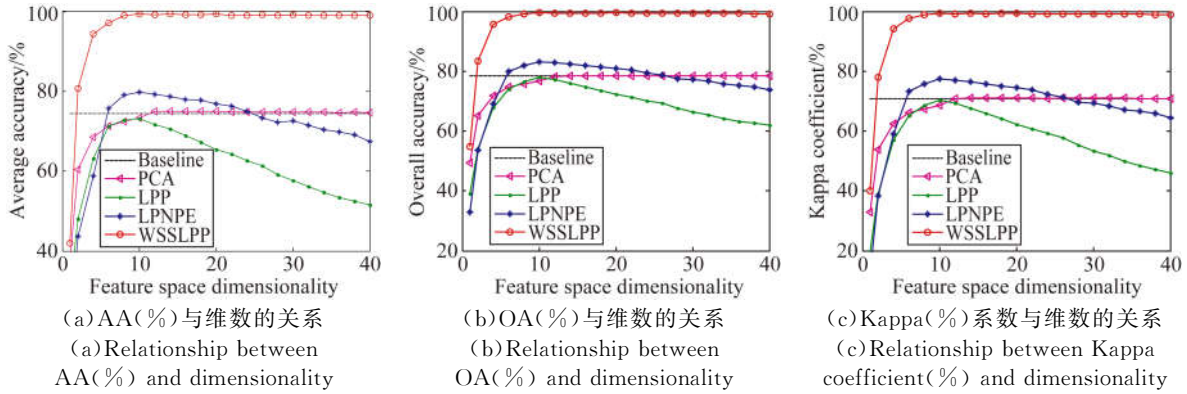


图 7 在 PaviaU 图像上的分类精度评价指标(%)与维数的关系

Fig. 7 Relationship between classification accuracy evaluation indexes(%) and dimensions on PaviaU dataset

表 2 给出了在训练样本相同的情况下,由不同算法得到的各类地物的分类精度。由图 7 可知,WSSLPP 算法的分类性能明显优于其他算法。表 2 给出了各种算法的分类结果,其对应的分类图如图 8(彩图见期刊电子版)所示。由图 8 和表 2 可知,本文方法对绝大部分地物的分类精

度、总体精度、平均精度和 Kappa 系数都较其他算法有显著提高,且得到的分类图中错分点最少,所得结果中地物分布更加平滑,这再次证明了本文算法的有效性。

由 WSSLPP 算法得到的分类结果较好,错分或漏分现象较少。

表 2 PaviU 数据库上各类地物在不同算法下的最高分类精度

Tab. 2 Classification accuracy of different types of features in PaviU data set by different algorithms (%)

地物类别	训练样本个数	测试样本个数	Baseline	PCA	LPP	LPNPE	WSSLPP
Asphalt	332	6 299	69.79	71.74	68.88	70.76	99.59
Meadows	932	17 717	92.54	91.65	91.09	94.47	100
Gravel	105	1 994	59.83	59.73	60.93	73.02	99.85
Trees	153	2 911	76.26	76.57	83.55	86.02	95.71
Sheets	67	1 278	99.61	99.61	99.92	99.84	99.69
Soil	251	4 778	51.70	50.86	51.23	69.00	99.81
Bitumen	67	1 263	53.92	59.38	59.86	68.33	100
Bricks	184	3 498	60.89	63.58	58.72	65.49	98.97
Shadows	47	900	94.11	92.11	86.33	94.11	100
平均精度(AA)	97	875	73.18	73.91	73.39	80.12	99.29
总体精度(OA)	246	2 209	77.77	77.96	77.36	83.00	99.50
Kappa 系数	59	534	0.700 1	0.703 3	0.696 4	0.772 6	0.993 4

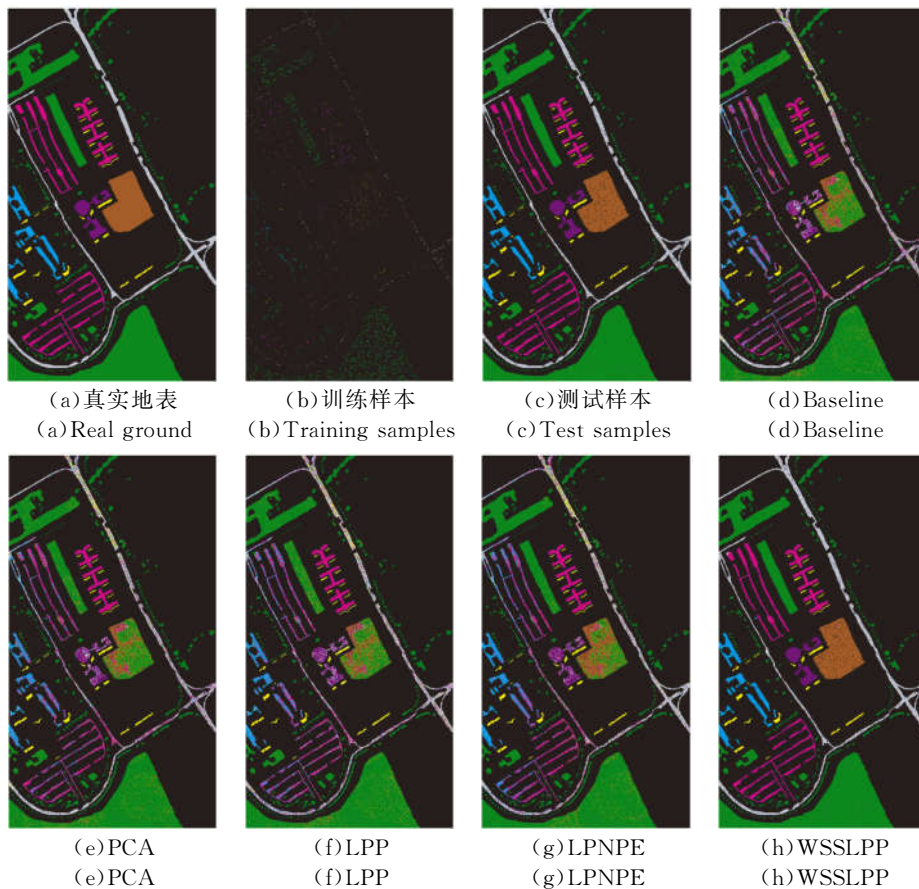


图 8 各算法在 PaviaU 数据库上的分类结果图

Fig. 8 Classification results of different algorithms in PaviaU data set

5 结 论

LPP 算法能够表征数据的固有结构,保持数据的局部信息,LPNPE 算法能够保持高光谱数据在空间上的局部像素间的近邻关系,但是单纯地利用光谱信息或者空间信息都不利于高光谱图像分类精度的提高。本文提出了一种新的融合高光谱图像空间信息和光谱信息的降维算法——加权空-谱局部保持投影(WSSLPP)算法。WSSLPP 算法结合高光谱图像的物理特性,其不仅能够保持高光谱图像谱域中像素的固有结构和局部信息,而且利用了空域中近邻像素的空间关系,故能够有效提高分类精度。

在 Indian Pines 和 PaviaU 数据库上的实验结果表明:相比基于光谱维或者空间维的降维算法,由本文方法得到的分类结果更好,对地物分类

精度提高的幅度更大。在 Indian Pines 数据库上随机选取每类样本的 10% 作为训练样本时,由本文方法得到的最大分类精度 OA 达到了 99%,较 Baseline 提高了 27.78%,较 PCA 提高了 27.78%,较 LPP 提高了 27.03%,较 LPNPE 提高了 23.26%; Kappa 系数达到了 0.988 6,较 Baseline 提高了 0.317 2,较 PCA 提高了 0.317 2,较 LPP 提高了 0.308 2,较 LPNPE 提高了 0.265 7。在 PaviaU 数据库上每类随机选取样本数的 5% 作为训练样本时,由 WSSLPP 得到的最大分类精度 OA 为 99.50%,较 Baseline 提高了 21.73%,较 PCA 提高了 21.54%,较 LPP 提高了 22.14%,较 LPNPE 提高了 16.5%; Kappa 系数为 0.993 4,较 Baseline 提高了 0.293 3,较 PCA 提高了 0.290 1,较 LPP 提高了 0.297,较 LPNPE 提高了 0.220 8; 有效地提高了高光谱图像的分类精度。

参考文献:

- [1] 黄鸿,杨媚,张满菊. 基于稀疏鉴别嵌入的高光谱遥感影像分类[J]. 光学精密工程, 2013, 21(11): 2922-2930.
HUANG H, YANG M, ZHANG M J. Hyperspectral image classification via sparse discriminant manifold embedding [J]. *Opt. Precision Eng.*, 2013, 21(11):2922-2930. (in Chinese)
- [2] 王忠良,冯燕,王丽. 推扫式高光谱谱间压缩感知成像与重构[J]. 光学精密工程, 2014, 22(11): 3129-3135.
WANG ZH L, FENG Y, WANG L. Compressive sensing imaging and reconstruction of pushbroom hyperspectra [J]. *Opt. Precision Eng.*, 2014, 22(11):3129-3135. (in Chinese)
- [3] 张成业,秦其明,陈理,等. 高光谱遥感岩矿识别的研究进展[J]. 光学精密工程, 2015, 23(8): 2407-2418.
ZHANG CH Y, QIN Q M, CHEN L, *et al.*. Research and development of mineral identification utilizing hyperspectral remote sensing[J]. *Opt. Precision Eng.*, 2015, 23(8):2407-2418. (in Chinese)
- [4] 宋金伟,张忠伟,陈晓敏. 利用线性预测与查表法的高光谱图像压缩[J]. 光学精密工程, 2013, 21(8): 2201-2208.
SONG J W, ZHANG ZH W, CHEN X M. Hyperspectral imagery compression via linear prediction and lookup tables [J]. *Opt. Precision Eng.*, 2013, 21(8):2201-2208. (in Chinese)
- [5] 唐中奇,付光远,陈进,等. 基于多尺度分割的高光谱图像稀疏表示与分类[J]. 光学精密工程, 2015, 23(9):2708-2714.
TANG ZH Q, FU G Y, CHEN J, *et al.*. Multi-scale segmentation-based sparse coding for hyperspectral image classification [J]. *Opt. Precision Eng.*, 2015, 23(9):2708-2714. (in Chinese)
- [6] 王晓飞,侯传龙,阎秋静,等. 基于相关向量机的高光谱图像噪声评估算法[J]. 红外与激光工程, 2014, 43(12):4159-4163.
WANG X F, HOU CH L, YAN Q J, *et al.*. Noise estimation algorithm based on relevance vector machine for hyperspectral imagery [J]. *Infrared and Laser Engineering*, 2014, 43(12):4159-4163. (in Chinese)
- [7] 李昌国,郭科. 应用自适应预测器排序的三阶预测高光谱图像无损压缩[J]. 光学精密工程, 2014, 22(3):760-769.
LI CH G, GUO K. Lossless compression of hyperspectral images using three-stage prediction based on adaptive predictor reordering [J]. *Opt. Precision Eng.*, 2014, 22(3):760-769. (in Chinese)
- [8] ZHOU Y C, PENG J T, CHEN C L P. Dimension reduction using spatial and spectral regularized local discriminant embedding for hyperspectral image classification [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2015, 53(2):1082-1095.
- [9] 黄鸿,曲焕鹏. 基于半监督稀疏鉴别嵌入的高光谱遥感影像分类[J]. 光学精密工程, 2014, 22(2): 434-442.
HUANG H, QU H P. Hyperspectral remote sensing image classification based on SSDE [J]. *Opt. Precision Eng.*, 2014, 22(2):434-442. (in Chinese)
- [10] 刘嘉敏,罗甫林,黄鸿,等. 应用相关近邻局部线性嵌入算法的高光谱遥感影像分类[J]. 光学精密工程, 2014, 22(6):1668-1676.
LIU J M, LUO F L, HUANG H, *et al.*. Hyperspectral image classification by related neighbor locally linear embedding algorithm [J]. *Opt. Precision Eng.*, 2014, 22(6):1668-1676. (in Chinese)
- [11] 金鹏磊. 空-谱联合高光谱数据降维与分类方法研究[D]. 西安电子科技大学, 2014.
JIN P L. *Researches on spatial-spectral based dimensionality reduction and classification of hyperspectral data* [D]. Xi'an: Xidian University, 2014. (in Chinese)
- [12] BANDOS T V, BRUZZONE L, CAMPS-VALLS G. Classification of hyperspectral image with regularized linear discriminant analysis [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2009, 47(3):862-873.
- [13] KUO B C, LANDGREBE D A. Nonparametric weighted feature extraction for classification [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2004, 42(5):1096-1105.
- [14] LU X, LI X. Multiresolution imaging [J]. *IEEE Transactions on Cybernetics*, 2014, 44(1):149-160.
- [15] XIAO Y, ZHU Z, ZHAO Y, *et al.*. Topographic NMF for data representation [J]. *IEEE Transactions on Cybernetics*, 2014, 44(10):1762-1771
- [16] WANG R, NIE F P, YANG X J, *et al.*. Robust 2DPCA with non-greedy-norm maximization for image analysis [J]. *IEEE Transactions on Cybernetics*, 2015, 45(5):1108-1112.
- [17] 练秋生,石保顺,陈书贞. 字典学习模型、算法及其应用研究进展[J]. 自动化学报, 2015, 41(2): 240-260.
LIAN Q SH, SHI B SH, CHEN SH ZH. Research advances on dictionary learning models, algorithms and applications [J]. *Acta Automatica Sinica*, 2015, 41(2):240-260. (in Chinese)
- [18] HE X, YAN S, HU Y, *et al.*. Face recognition using laplacian faces [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, 27(3):328-340.
- [19] HE X, CAI D, YAN S, *et al.*. Neighborhood

- preserving embedding [C]. *Proceedings of the 10th IEEE Conference on Computer Vision*, 2005, 2:1208-1213.
- [20] VELASCO-FORERO S, MANIAN V. Improving hyperspectral image classification using spatial pre-processing [J]. *IEEE Geoscience and Remote Sensing Letters*, 2009, 6(2):297-301.
- [21] PHILLIPS R D, BLINN C E, WATSON L T, *et al.*. An adaptive noise-filtering algorithm for AVIRIS data with implications for classification accuracy [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2009, 47(9):3168-3179.
- [22] KANG X D, LI S T, BENEDIKTSSON J A. Spectral-Spatial hyperspectral image classification with edge-preserving filtering [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2014, 52(5):2666-2677.
- [23] HUANG X, ZHANG L. An adaptive mean-shift analysis approach for object extraction and classification from urban hyperspectral imagery [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2008, 46(12):4173-4185.
- [24] HUANG X, ZHANG L, LI P. Classification and extraction of spatial features in urban areas using high-resolution multispectral imagery [J]. *IEEE Geoscience and Remote Sensing Letters*, 2007, 4(2):260-264.
- [25] FAUVEL M, BENEDIKTSSON J A, CHANUS-SOT J, *et al.*. Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2008, 46(11):3804-3814.
- [26] MOSER G, SERPICO S B. Combining support vector machines and Markov random fields in an integrated framework for contextual image classification [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2013, 51(5):2734-2752.
- [27] LI J, BIOUCAS-DIAS J M, PLAZA A. Spectral-Spatial hyperspectral image segmentation using subspace multinomial logistic regression and Markov random fields [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2012, 50(3):809-823.
- [28] HUANG X, ZHANG L. An SVM ensemble approach combining spectral, structural, semantic features for the classification of high-resolution remotely sensed imagery [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2013, 51(1):257-272.
- [29] 邓承志, 张绍泉, 汪胜前, 等. L1 稀疏正则化的高光谱混合像元分解算法比较[J]. *红外与激光工程*, 2015, 44(3):1092-1097.
- DENG CH ZH, ZHANG SH Q, WANG SH Q, *et al.*. Hyperspectral unmixing algorithm based on L1 regularization [J]. *Infrared and Laser Engineering*, 2015, 44(3):1092-1097. (in Chinese)
- [30] 黄鸿, 郑新磊. 加权空-谱与最近邻分类器相结合的高光谱图像分类[J]. *光学精密工程*, 2016, 24(4):873-881.
- HUANG H, ZHENG X L. Hyperspectral image classification with combination of weighted spatial-spectral and KNN [J]. *Opt. Precision Eng.*, 2016, 24(4):873-881. (in Chinese)
- [31] PU H Y, CHEN ZH, WANG B. 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.

作者简介:



何 芳(1991—),女,湖北武汉人,博士研究生,2014年于第二炮兵工程大学获得学士学位,主要从事信号与信息处理,机器学习方面的研究。E-mail: hefangsi@163.com

导师简介:



贾维敏(1971—),女,河北保定人,教授,博士生导师,2007年于第二炮兵工程大学获得博士学位,主要从事信号与信息处理,阵列天线信号处理,宽带移动通信方面的研究。E-mail:jwm602@163.com



王 榕(1983—),男,山西长治人,讲师,硕士生导师,2013年于第二炮兵工程大学获得博士学位,主要从事信号与信息处理、机器学习方面的研究。E-mail:wangrong07@tsinghua.org.cn