

doi:10.6041/j.issn.1000-1298.2018.06.025

面向空间自相关信息的高光谱图像分类方法

廖建尚¹ 王立国²

(1. 广东交通职业技术学院轨道交通学院, 广州 510650; 2. 哈尔滨工程大学信息与通信工程学院, 哈尔滨 150001)

摘要: 空间滤波器在提取高光谱图像纹理信息过程中容易丢失空间自相关信息, 导致植被分类精度不高。针对当前方法的不足, 提出一种空间自相关信息的高光谱图像分类算法 (Classification of hyperspectral image based on spatial autocorrelation information, CHISCI)。该方法先用域转换线性插值卷积滤波 (Domain transform filter of interpolated convolution, DTFOIC) 对高光谱全波段图像提取空间自相关信息, 然后对高光谱数据进行主成分分析 (Principal component analysis, PCA) 降维后的前部分主成分提取空间自相关信息, 两种空间自相关信息线性融合后交由支持向量机 (Support vector machine, SVM) 完成分类。试验表明, 相比使用光谱信息、高光谱降维、空谱结合的 SVM 分类方法和边缘保持滤波以及递归滤波的方法, 所提出的 CHISCI 方法对高光谱图像的植被分类精度有较大提高, 在训练样本仅为 6% 和 1% 的情况下, 对印第安农林和萨里斯山谷数据集分类的总体分类精度分别达到 96.16% 和 98.67%, 比其他算法高出 2~16 个百分点, 验证了该方法的有效性。

关键词: 高光谱图像; 空间自相关信息; 域转换线性插值卷积滤波; 植被分类

中图分类号: TP753 **文献标识码:** A **文章编号:** 1000-1298(2018)06-0215-10

Classification Method of Hyperspectral Image Based on Spatial Autocorrelation Information

LIAO Jianshang¹ WANG Ligu²

(1. School of Rail Transit, Guangdong Communication Polytechnic, Guangzhou 510650, China

2. College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China)

Abstract: Spatial autocorrelation information is easily lost in the process of traditional texture information extraction methods of hyperspectral image, leading to low accuracy of vegetation classification. An improved scheme was put forward aiming at the shortcoming of existent methods to form a new classification algorithm (CHISCI) based on spatial autocorrelation information. Firstly, one kind of spatial autocorrelation information of hyperspectral image was extracted by domain transform filter of interpolated convolution (DTFOIC). Secondly, another kind of spatial autocorrelation information was obtained by the same filter on dimensionality reduced hyperspectral data. Finally, the two kinds of spatial information were combined and then classified by SVM which was not sensitive to high-dimensional data, forming CHISCI classification algorithm of hyperspectral image by spatial autocorrelation information. The CHISCI classification method was implemented on the hyperspectral data of Indian Pines and Salinas Valley. The following results were obtained. In the first place, the overall accuracy (OA) of Indian Pines was 96.16% and the Salinas Valley was 98.67%, which were 12~16 percentage points higher than those of SVM and PCA-SVM, and 4~16 percentage points higher than those of SGB-SVM, SBL-SVM and SGD-SVM by spatial-spectral information, and 4~6 percentage points higher than that of EPF, and 2~3 percentage points higher than that of IFRF. Furthermore, the average accuracy (AA) and Kappa of the CHISCI were also increased substantially, showing very good performance in hyperspectral classification. In the second place, although the training samples were only 6% of Indian Pines and 1% of Salinas Valley, the OA of both can reach 96.16% and 98.67%, which can remove salt

收稿日期: 2017-08-15 修回日期: 2017-10-17

基金项目: 国家自然科学基金项目 (61275010、61675051)、广东高校省级科研项目 (2017GKTSX021)、广东省科技计划项目 (2017ZC0358)、广州市科技计划项目 (201804010262)、国家星火计划项目 (2014GA780056) 和广东交通职业技术学院重点科研项目 (2017-1-001)

作者简介: 廖建尚 (1982—), 男, 副教授, 主要从事高光谱图像处理研究, E-mail: liaojianshang@126.com

and pepper noise in the classification map obviously. When the training samples were reduced to 3% and 0.3%, the OA can be over 90% and 95%, respectively. The effectiveness of CHISCI was fully verified in the hyperspectral classification. In the last place, the classification of some methods for grapes_untrained and vinyard_untrained in Salinas Valley were bad. The reason was that the spectral reflectances of the two vegetables for all bands were very close. However, the classification for the two vegetables of CHISCI can still reach 98.38% and 99.17%. It was showed that the CHISCI had excellent performance on the vegetable classification with close spectra. The experiments showed that the CHISCI algorithm was better than original SVM with pure spectrum information, the dimensionality reduction-based methods, the spatial-spectral information-based methods, and the methods based on edge-preserving filtering and recursive filtering. With the spatial autocorrelation information extracted by the DTFOIC, the performance of the classification of hyperspectral image with CHISCI algorithm was greatly improved, and the effectiveness of CHISCI was fully verified in the classification of hyperspectral vegetables, especially of those with close spectra. The method can be applied to the field of crop growing, diseases and pests monitoring, accurate classification and identification. It would also have potential significance for precision agriculture and agricultural modernization.

Key words: hyperspectral image; spatial autocorrelation information; domain transform filter of interpolated convolution; vegetable classification

0 引言

通过成像光谱仪获得的高光谱遥感图像能达到几百个波段的光谱,在遥感高光谱分类中,地物(特别是植被)种类丰富且光谱相似^[1],利用高光谱图像可以有效对农作物精准分类与识别,在农业灾害和产量评估方面已有广泛应用^[2-3]。有学者将高光谱遥感应用于植被分类^[4-7],例如有一些学者采用光谱特征和模式识别等方法提高植被分类精度^[8-12],对高光谱遥感植被分类进行研究并取得了一定成果,但仅仅关注光谱信息特征挖掘,对于光谱信息接近的植被分类效果并不好;也有一部分学者提出利用空间信息来实现植被分类^[13-15],对空间信息的挖掘进行了初探,但是空间信息的挖掘并不够充分,导致分类性能不高。近年来,越来越多学者利用空间纹理信息结合光谱信息改善高光谱图像地物分类性能,主要有形态滤波特征提取方法^[16]、马尔科夫随机场特征提取^[17-18]和图像分割特征提取方法^[19-20],其中用滤波器提取图像纹理信息来辅助光谱信息进行有效分类更是一个研究热点,例如不少学者用 Gabor 滤波器^[21]提取纹理信息来辅助高光谱分类^[22-23];还有一些学者用双边滤波器^[24-25]来提取高光谱空间纹理信息^[26-29],用于辅助 SVM 分类。植被分布有较好的统一性,具有较强的空间相关性,以上学者通过挖掘空间纹理信息来提高分类性能有了很大进步,但在提取空间纹理信息的同时,忽略了空间相关性,因此存在一定的局限性。

空间纹理信息提取用于高光谱图像分类的研究中取得了一定成效,但也存在一些不足:①光谱信息接近的植被分类性能有待提高。②过去的高光谱空

谱结合分类方法更多关注提取空间纹理信息,忽略了高光谱空间自相关信息的提取。本文通过挖掘高光谱图像的空间自相关信息来提高分类性能,提出用域转换线性插值卷积滤波^[30]来提取高光谱图像空间自相关信息,构建基于空间自相关信息的高光谱图像分类算法(Classification of hyperspectral image based on spatial autocorrelation information, CHISCI),以期有效去除高光谱分类中的椒盐现象,且利用空间自相关信息有效对光谱特性相似的植被实现高精度分类,弥补用光谱信息进行分类的不足。

1 方法

本文实现了基于空间自相关信息的高光谱图像分类算法(CHISCI),对高光谱全波段和和高光谱数据 PCA 降维后的前部分主成分进行滤波,获得有效的空间自相关信息,融合后将其交由 SVM 实现分类。

1.1 空间自相关信息的提取方法与评价指标

1.1.1 域转换线性插值卷积滤波提取空间自相关信息

域转换线性插值卷积滤波是 2011 年由 GASTAL 等^[30]提出的图像特征提取算法,可以将二维的滤波转换为二维的特征提取滤波,有较好的离散噪声去除效果,对有 k 个波段的高光谱图像,第 i 个波段的域转换线性插值卷积滤波 $T_i(u)$ 在高光谱的定义为

$$T_i(u) = \int_{\Omega_w} G_w(x) H(f(u), x) dx$$

$$(i = 1, 2, \dots, k; u \in \Omega_w) \quad (1)$$

其中 $H(f(u), x) = \frac{1}{2r} \delta\{|f(u) - x| \leq r\}$ (2)

$$\delta(F) = \begin{cases} 1 & (F \text{ 为真}) \\ 0 & (\text{其他}) \end{cases} \quad (3)$$

式中 $G_w(x)$ ——通过滤波集 Ω_w 中进行线性插值重建的图像

$H(f(u), x)$ ——滤波核, 通过一个布尔函数表示, 其意义为像元邻域可认为是同种地物, 因此该滤波过程中保留了高光谱图像的空间自相关性

将式(2)、(3)代入式(1), 可得

$$T_i(u) = \frac{1}{2r} \int_{f(u)-r}^{f(u)+r} G_w(x) dx \quad (4)$$

其中 $f(u) = \int_0^u \left(1 + \frac{\sigma_s}{\sigma_r} \sum_{l=1}^c I'_l(x) \right) dx \quad (5)$

$$\sigma_r = \sqrt{3}\sigma_j \quad (6)$$

$$\sigma_{J_t} = \sigma_s \sqrt{3} \frac{2^{N-t}}{\sqrt{4^N - 1}} \quad (7)$$

式中 r ——滤波半径

σ_s ——空间标准偏差

σ_r ——范围标准偏差

σ_{J_t} ——第 t 次迭代的偏差

N ——总迭代次数

高光谱数据集波段多、数据量大, 先对高光谱数据进行归一化, 再对每个波段图像进行域转换线性插值卷积滤波提取空间信息, 图 1 为域转换卷积滤波提取高光谱图像空间信息的流程图。

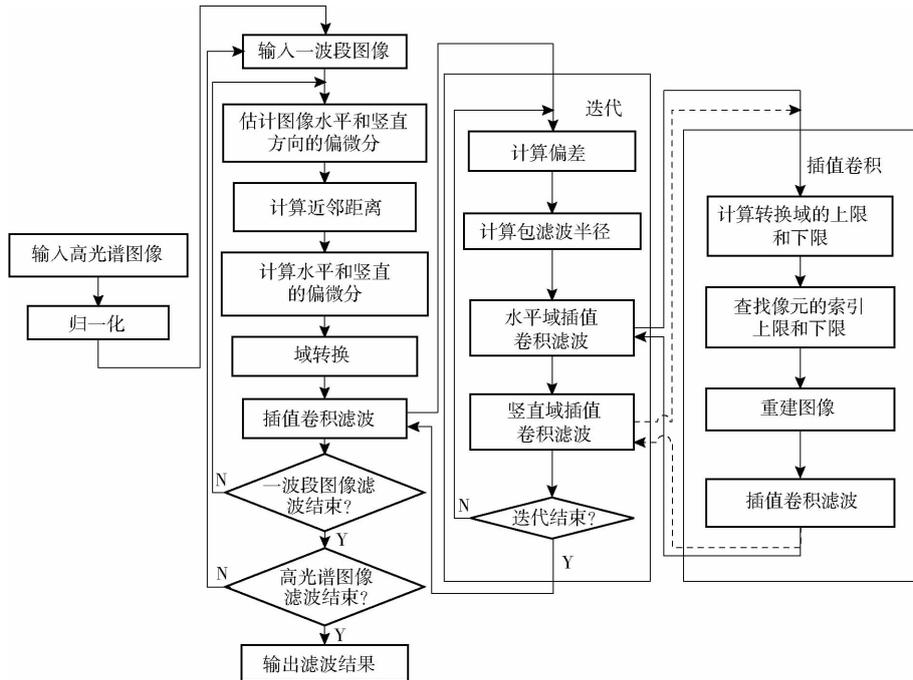


图 1 域转换线性插值卷积滤波流程图

Fig. 1 Flowchart of DTFOIC

为了确定空间偏差系数 σ_s 、范围偏差系数 σ_r 对分类的影响, 选用印第安农林图像作为试验数据集, 其中 8% 作为训练样本和 92% 作为测试样本进行验证, 用控制变量的交叉验证法, 先取 $\sigma_r = 0.3$,

$\sigma_s = 10 \sim 400$ 分别验证分类精度, 当 $\sigma_s = 30$ 时, 总体分类精度 OA、平均分类精度 AA 和 Kappa 系数综合分类性能较好; 然后取 $\sigma_s = 30, \sigma_r = 0.1 \sim 1.0$, 当 $\sigma_r = 0.3$ 时, 综合分类性能较好, 分类折线如图 2 所

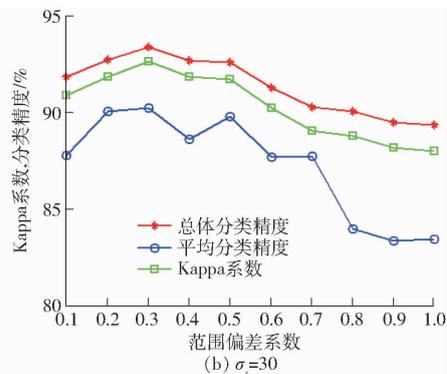
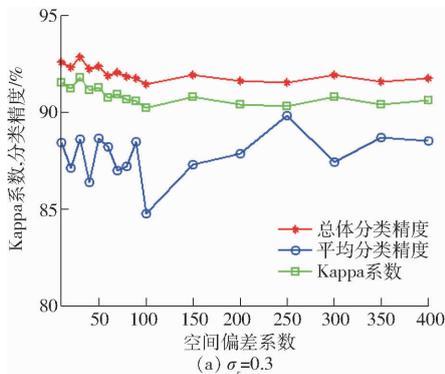


图 2 σ_s 和 σ_r 的分类验证

Fig. 2 Verification of classification with σ_s and σ_r

示,本文域转换线性插值卷积滤波系数取 $\sigma_s = 30$ 和 $\sigma_r = 0.3$ 。

1.1.2 空间自相关信息的评价指标

引入 MORAN^[31-32]于1950年提出的莫兰指数 (Moran's I) 来衡量高光谱中每个波段图像的空间自相关性, I 表达式为

$$I = \frac{n}{\sum_{i=1}^n \sum_{j=1}^n w_{ij}} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} z_i z_j}{\sum_{i=1}^n z_i^2} \quad (8)$$

其中

$$\begin{cases} z_i = X_i - \bar{X} \\ z_j = X_j - \bar{X} \end{cases} \quad (9)$$

式中 X_i ——高光谱某像元的反射强度

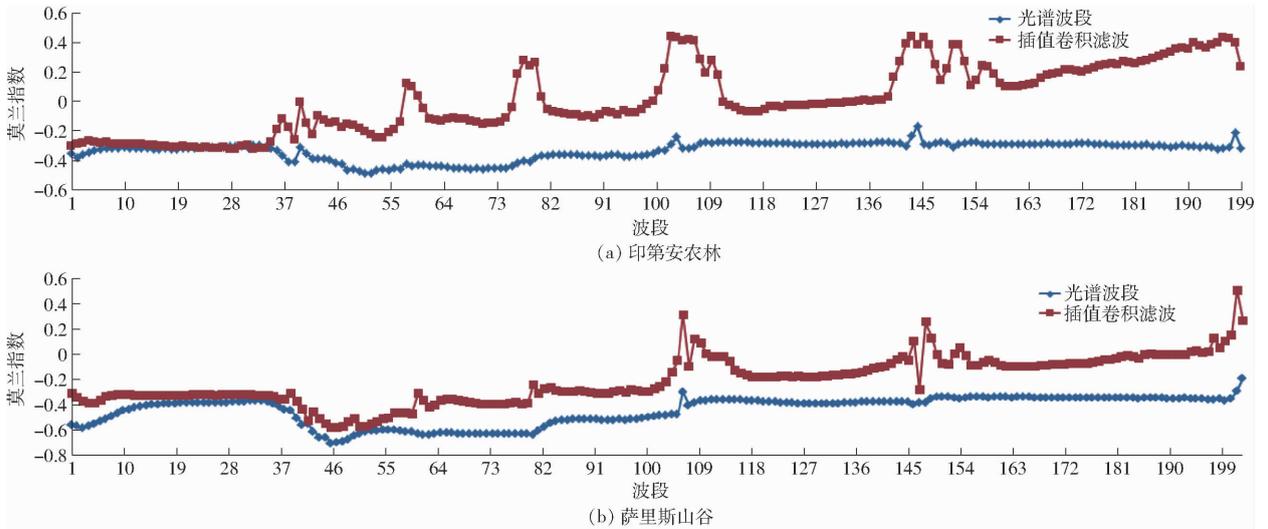


图3 原光谱和域转换线性插值卷积滤波的莫兰指数的均值对比

Fig.3 Comparison of Moran's I for spectral and DTFOIC

1.2 基于空间自相关信息的高光谱分类

首先,本文采用了 SVM 分类器对高光谱图像进行分类研究, SVM 的分类原理: 寻找一个分类超平面, 使得训练样本中的两类样本点能被分开, 并且距离该平面尽可能地远, 即间隔最大化; 而对线性不可分的问题, 通过核函数将低维输入空间的数据映射到高维空间, 将原低维空间的线性不可分问题转换为高维空间上的线性可分问题。SVM 为映射特征空间上间隔最大的线性分类器, 假设对原高光谱图像进行分类, 如果地物的光谱相近, 容易产生相互误分现象, 分类效果差。其次, 植被分布有较好的统一性, 因此具有较强的空间相关性, 因此提取空间自相关性信息将大大有利于 SVM 对高光谱植被进行分类。

过去高光谱图像空谱结合的分类方法在空间信息中更多关注地物的纹理信息提取, 滤波器往往能提取较好的纹理信息, 但纹理信息的提取过程往往容易丢失地物的空间相关性信息, 从 1.1 节可以看

X_j ——其他像元的反射强度

n ——像元数量 \bar{X} —— X 的均值

w_{ij} ——空间比重权重, 将指数 I 归一化为 $[-1, 1]$, 值越大空间相关性越强, 值越小空间差异性越大

为了体现域转换线性插值卷积滤波对空间自相关信息的保持特性, 先用域转换线性插值卷积滤波对印第安农林和萨里纳斯山谷数据集的每个波段逐个滤波, 然后计算每个波段的 Moran's I 指数并求均值, 如图 3 所示。由图 3 可以看出, 2 个数据集的域转换线性插值卷积滤波提取的空间相关性比原光谱信息好, 验证了域转换线性插值卷积滤波有较好的空间自相关性保持特性。

出, 与原光谱信息相比, 域转换线性插值卷积滤波后的高光谱图像能保持良好的空间自相关性。

利用域转换线性插值卷积滤波特性, 本文设计了一种基于空间自相关信息的高光谱图像植被分类方法, 首先用域转换线性插值卷积滤波对高光谱图像的光谱信息提取空间自相关信息, 其次为了从多角度获取更丰富的空间自相关信息, 用 PCA 主成分分析方法对光谱数据进行降维, 考虑降维后的大部分信息集中在前面的主成分, 用域转换线性插值卷积滤波提取前主成分的空间自相关信息, 两种信息线性融合后, 由对高维数据不敏感的 SVM 分类器完成分类, 构建了基于空间自相关信息的高光谱图像植被分类算法 (CHISCI), 实现过程如下:

(1) 归一化: 高光谱像元的反射强度较大, 利用

$$G = \frac{R - R_{\min}}{R_{\max} - R_{\min}} \quad (10)$$

对波段数为 l 的高光谱数据集作归一化处理, R_{\min} 指的是某个波段的最小值, R_{\max} 为最大值, 得到信息量

重新分布的高光谱图像数据集 R 。

(2) 高光谱 PCA 降维: 对于有 l 个波段的高光谱数据集 R 进行 PCA 降维, 为了充分利用 PCA 成分来进行分类, 本文选择前 10% 的 PCA 成分组成新的数据集

$$K = \text{PCA}(R) \quad (11)$$

其中选择前 10% 的 PCA 成分将在 3.2 节中进行验证分析。

(3) 提取空间自相关信息 1: 用域转换线性插值卷积滤波对全部光谱信息 K 提取空间信息

$$R_{ds}^i = T(R^i) \quad (i=1, 2, \dots, l) \quad (12)$$

(4) 提取空间自相关信息 2: 用域转换线性插值

卷积滤波对数据集 K 提取空间信息

$$K_{ds}^j = T(K^j) \quad (j=1, 2, \dots, n) \quad (13)$$

(5) 空间信息融合: 将空间自相关信息 R_{ds} 和 K_{ds} 线性融合, 形成新的数据集

$$H = R_{ds} + K_{ds} \quad (14)$$

(6) 随机从空间信息数据集 H 以一定比例随机抽取训练集 H_s , 其余部分作为训练集 H_t 。

(7) 用径向基函数支持的 SVM 对 H_s 进行训练, 获取训练模型。

(8) 获取模型后, 用径向基函数支持的 SVM 对测试集 H_t 进行分类, 得出分类结果。

CHISCI 算法流程如图 4 所示。

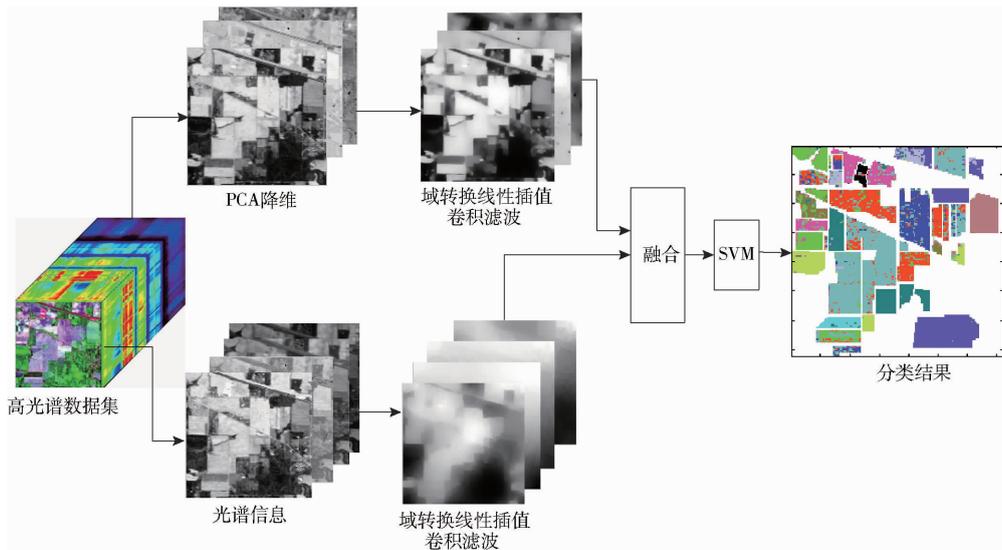


图 4 CHISCI 算法流程图

Fig. 4 Flowchart of CHISCI

2 试验数据集

为了验证本文提出算法的有效性, 采用印第安农林和萨里斯山谷 2 个高光谱数据集进行验证, 数据集特征如下:

(1) 印第安农林: 来自机载可见光红外成像光谱仪, 1992 年在印第安纳州西北部印第安农林收集到的高光谱遥感图像, 有 20 m 的空间分辨率, 包含 145×145 个像元, 220 个波段, 其中波长范围在 $4 \times 10^{-7} \sim 2.5 \times 10^{-6}$ m, 由于噪声和水吸收等因素除去其中的 20 个波段, 包括 [104 ~ 108]、[150 ~ 163] 和 220, 剩余 200 个波段, 包含 16 类地物, 具体地物类别和数量见表 1。

(2) 萨里斯山谷: 来自机载可见光红外成像光谱仪, 是 1992 年在美国加利福尼亚州萨里斯山谷收集到的图像, 具有 3.7 m 的空间分辨率, 其包含 512×217 个像元, 224 个波段, 由于噪声和水吸收等因素除去其中的 20 个波段, 包括 [108 ~ 112]、[154 ~ 167] 和 224, 剩余 204 个波段, 包含 16 类地物, 具体

地物类别和数量见表 2。

3 试验

3.1 试验方法与评价指标

为了验证 CHISCI 算法在高光谱植被分类的优越性, 本文采用 6 种方法进行比较, 分别为: 方法 1: 利用文献 [33] 提出的 SVM; 方法 2: 利用文献 [34] 提出的高光谱图像 PCA 降维后用 SVM 进行分类; 方法 3: 用 Gabor 滤波器、双边滤波器和导向滤波器^[35]分别对高光谱数据用 PCA 降维后的前 20 个主成分提取空间信息, 并将获取的空间信息和光谱信息线性结合后, 用 SVM 进行分类, 形成 SGB-SVM、SBL-SVM、SGD-SVM; 方法 4: 采用文献 [26] 提出的 EPF 算法对高光谱图像进行分类, 有 EPF-B-c 和 EPF-G-c; 方法 5: 采用文献 [28] 提出的域转换递归滤波的方法 IFRF; 方法 6: 本文提出的 CHISCI。

为了更好地进行对比, 印第安农林数据集图像选取全部 16 种类别, 每类随机选取 6% 样本组成有

表1 印第安农林数据集图像分类数据统计
Tab.1 Classification statistics of Indian Pines

地物	总数	SVM	PCA-SVM	SGB-SVM	SBL-SVM	SGD-SVM	EPF-B-c	EPF-G-c	IFRF	CHISCI
苜蓿	54	79.25	80.01	75.44	72.63	76.98	94.79	94.65	90.54	93.58
免耕玉米	1434	73.94	70.78	71.06	84.37	87.97	92.21	90.25	90.49	94.16
少耕玉米	834	68.03	67.81	69.80	86.54	86.39	83.27	86.94	87.49	96.96
玉米	234	48.44	41.64	43.71	52.21	70.55	56.67	58.76	89.43	90.23
草地/牧草	497	90.38	89.01	86.69	92.80	93.21	95.60	93.97	92.87	94.67
草地/树木	747	95.43	96.13	95.14	97.35	96.63	99.90	99.91	98.03	98.22
修剪的牧草	26	59.71	73.40	78.38	59.54	48.39	62.11	79.47	74.86	94.66
干草料堆	489	94.43	95.98	96.12	98.36	95.20	99.83	99.94	99.59	99.55
燕麦	20	74.99	56.36	70.10	42.42	51.31	39.64	41.75	49.81	59.57
免耕大豆	968	68.35	64.74	66.57	82.85	84.69	85.07	83.03	86.59	91.03
少耕大豆	2468	84.03	85.16	84.09	91.49	93.06	97.85	97.40	96.41	98.67
纯净大豆	614	79.09	76.96	77.94	86.32	88.59	96.20	96.79	97.78	93.52
小麦	212	92.11	95.49	97.95	98.17	97.44	94.15	99.60	99.10	99.20
森林	1294	97.57	97.10	97.04	98.08	98.52	99.72	99.61	99.15	98.84
建筑-草-树	380	37.91	44.12	46.07	63.99	69.88	58.44	59.42	85.86	95.08
石钢塔	95	75.75	75.11	82.25	77.69	83.00	92.84	91.10	85.44	81.68
OA/%		80.38	79.85	80.10	88.48	90.20	92.28	92.14	93.71	96.16
AA/%		76.21	75.61	77.40	80.30	82.61	84.27	85.79	88.97	92.48
Kappa		77.48	76.82	77.17	86.83	88.80	91.14	90.97	92.82	95.62

表2 萨里斯山谷数据集图像分类数据统计
Tab.2 Classification statistics of Salinas Valley

地物	总数	SVM	PCA-SVM	SGB-SVM	SBL-SVM	SGD-SVM	EPF-B-c	EPF-G-c	IFRF	CHISCI
椰菜绿色杂草1	2009	97.78	99.52	99.46	99.82	99.80	99.59	99.89	96.63	100
椰菜绿色杂草2	3726	98.65	97.71	95.29	99.09	98.87	100	99.96	97.72	99.61
休耕地	1976	83.07	86.79	93.68	89.93	94.33	88.97	93.47	99.08	100
粗糙的休耕地	1394	94.97	98.48	98.01	98.30	97.00	99.86	99.70	83.60	96.41
平滑的休耕地	2678	98.12	97.59	96.23	97.67	96.17	99.52	99.46	96.03	99.27
残株	3959	99.61	99.53	99.42	99.47	99.55	100	99.92	98.50	99.52
芹菜	3579	99.41	99.64	99.64	99.78	99.62	100	99.97	99.16	99.80
未结果实的葡萄	11271	88.49	88.87	88.58	91.12	91.43	94.88	94.85	96.56	98.38
正在开发的葡萄园土地	6203	99.02	99.79	99.88	99.86	99.82	99.99	99.89	99.01	99.97
开始衰老的玉米	3278	86.72	87.27	90.86	91.12	92.92	95.88	92.70	96.82	97.23
长叶莴苣4wk	1068	72.43	84.72	62.60	84.27	77.37	82.75	79.48	87.51	97.13
长叶莴苣5wk	1927	96.76	98.28	97.86	98.15	98.56	99.90	99.46	97.77	98.66
长叶莴苣6wk	916	95.22	95.99	91.04	84.97	98.10	100	99.86	67.49	95.56
长叶莴苣7wk	1070	91.21	88.32	92.11	89.09	91.29	99.38	97.06	95.71	96.39
未结果实的葡萄园	7268	59.10	54.77	60.78	69.82	73.40	60.80	60.19	98.96	99.17
葡萄园架	1807	89.93	91.59	92.16	97.11	95.74	91.12	90.44	95.88	95.14
OA/%		88.86	88.92	89.44	91.90	92.68	92.32	92.03	96.63	98.67
AA/%		90.66	91.80	91.10	93.10	94.00	94.54	94.14	94.15	98.23
Kappa		87.55	87.61	88.20	90.95	91.83	91.41	91.09	96.25	98.52

标签训练集,其余94%作为测试集;萨里斯山谷数据集图像选取全部16种类别,每类随机选取1%样本组成有标签训练集,其余99%作为测试集。

用总体分类精度(Overall accuracy, OA)、平均分类精度(Average accuracy, AA)以及Kappa统计系数(Kappa statistic, Kappa)来衡量分类算法的精

度,为了避免随机偏差,每个方法重复10次试验并统计平均结果,验证平台采用6GB RAM的试验平台。

3.2 PCA降维和分类试验与分析

提出的CHISCI算法采用了PCA方法,考虑降维后的大部分信息集中在前面的主成分,为了确定

分类中使用 PCA 主成分的数量比例,从而达到较优分类结果,本文对 2 个数据集进行以下分类验证。

印第安农林数据集和萨里斯山谷数据集分别有 200 和 204 个波段图像,先对 2 个数据集进行 PCA 降维,印第安林分别选择前 15% (1 ~ 30) 主成分进行分类试验,分类结果如图 5a 所示,从试验结果可以看出,利用前 5% (10) 主成分进行分类的 OA 可以达到 89%,前 8% (16) 主成分的 OA 可以稳定在 90%。同理,萨里斯山谷数据集选择前 15% (1 ~ 31) 主成分进行分类试验,分类结果如图 5b 所示,从试验结果可以得知,利用前 4.4% (9) 主成分进行分类的 OA 可以达到 95%,前 8.3% (17) 主成分的 OA 可以稳定在 96% 以上。因此为了更充分利用 PCA 成分来进行分类,本文选择对高光谱数据 PCA 降维后的前 10% 作为滤波以及分类成分。

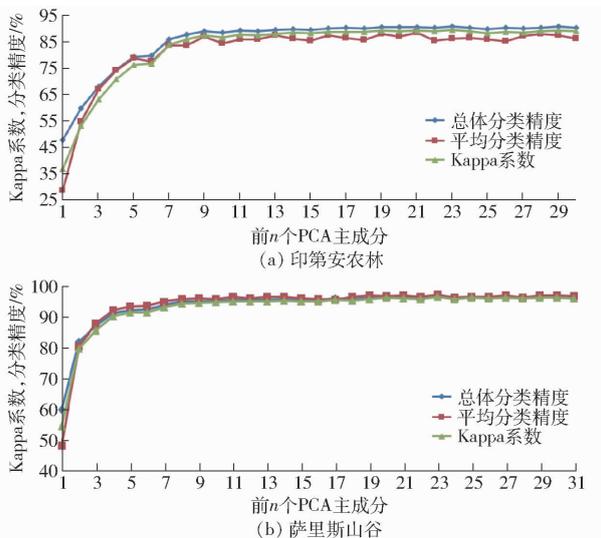


图 5 PCA 成分分类结果

Fig. 5 Classification results for PCA

3.3 试验与分析

为了验证本文提出方法对分类性能的有效性,用 6 种方法分别对印第安林和帕维亚大学数据集进行分类对比验证,得到表 1 (训练数占总数 6%,测试数占总数 94%) 和表 2 (训练数占总数 1%,测试数占总数 99%) 的分类结果,分类效果如图 6 和图 7 所示,对试验分析如下:

(1) CHISCI 实现对两类数据集的较优分类,其中印第安农林数据集的总体分类精度 OA 为 96.16%,萨里斯山谷数据集为 98.67%,比 SVM 和 PCA-SVM 的总体分类精度高出 12~16 个百分点,比空谱结合方法 SGB-SVM、SBL-SVM、SGD-SVM 高出 4~16 个百分点,比 EPF 算法高出 4~6 个百分点,比 IFRF 算法高出 2~3 个百分点,另外相对其他方法,CHISCI 算法对两类数据集的平均分类精度 AA 和 Kappa 系数同样有大幅度的提高,充分验证了 CHISCI 算法在高光谱图像植被分类的有效性。

(2) 从印第安农林数据集分类试验可知,在训练样本仅为 6%,总体分类精度 OA 达到了 96.16%,说明对植被多、分布复杂的高光谱数据有较好的分类效果,有效去除椒盐现象,用方框标注的地方去除效果更为明显;萨里斯山谷数据集分类训练样本仅为 1%,CHISCI 也能达到较优的分类效果,OA、AA 均超 98%,Kappa 系数也超过 98,有效地去除椒盐现象,用方框标注的地方效果较为突出,验证了 CHISCI 对高光谱图像的植被有较好的分类性能。

(3) 从萨里斯山谷试验来看,前面几个方法对未结果实的葡萄和未结果实的葡萄园分类效果比较差,两者间错分的比较多,从地物中选择休耕地、未

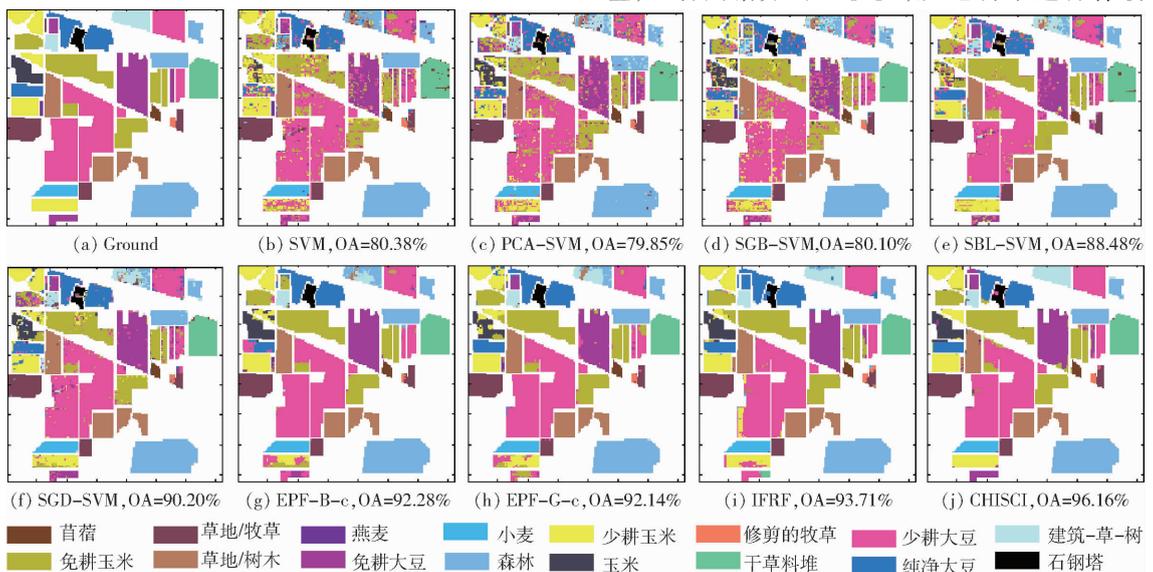


图 6 印第安农林数据集分类

Fig. 6 Classifications of Indian Pines

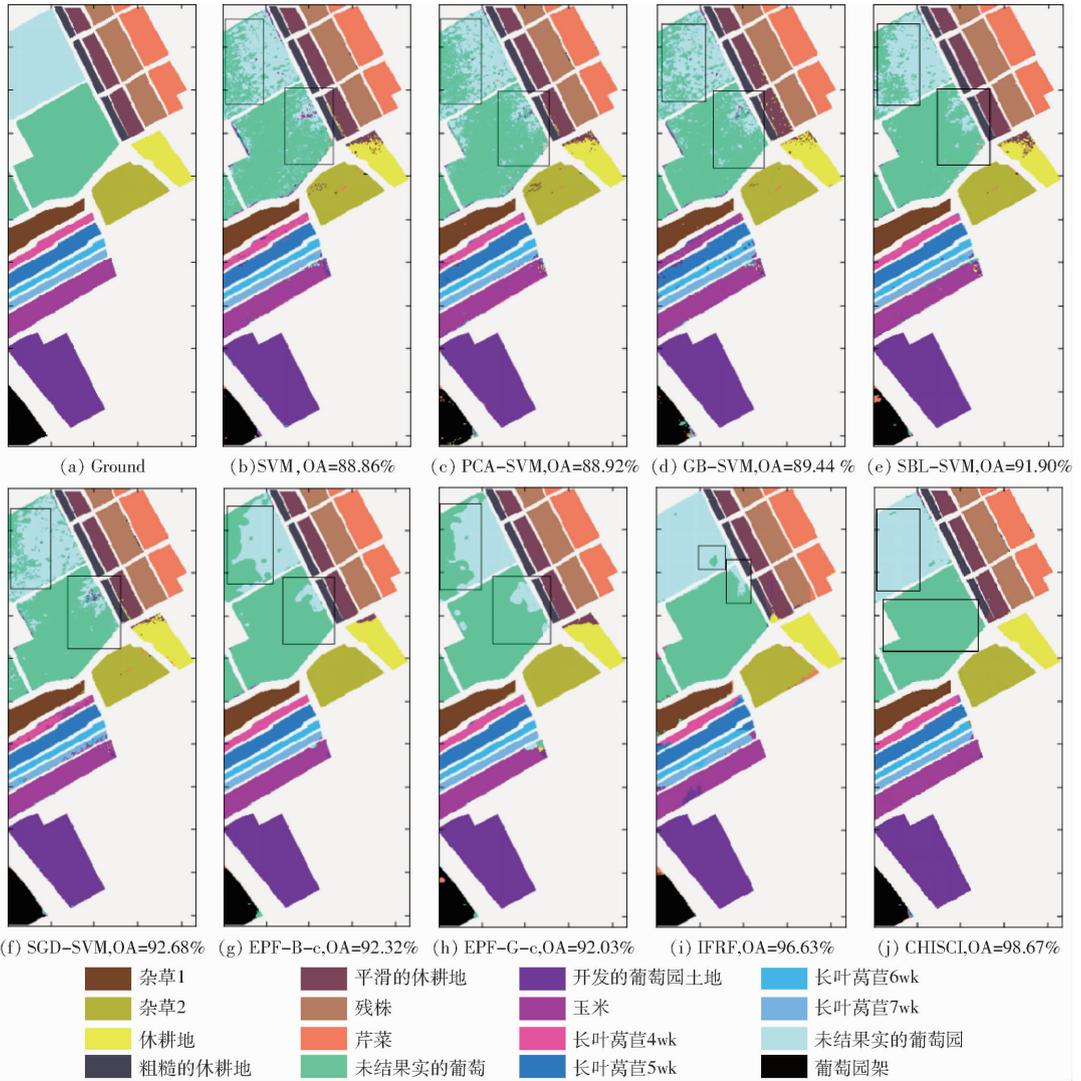


图7 萨里斯山谷数据集分类

Fig. 7 Classifications of Salinas Valley data set

结果实的葡萄、玉米、长叶莴苣 5wk 和未结果实的葡萄园 5 个植被中的一个像元,并画出全波段光谱图,如图 8 箭头所示,可知葡萄和葡萄园各波段的光谱反射率非常接近,导致分类器用光谱信息来分类的效果比较差,尤其是采用 SVM 对原光谱数据分类,由表 2 可知,两者的 OA 仅为 88.49% 和 59.10%,从图 7b 可以看出,两者相互误分的情况比较严重(用方框圈出),而 CHISCI 对两种植被的 OA 分别达到了 98.38% 和 99.17%,比其他分类器的分类精度都高,两者相互误分的现象大大改善,验证了 CHISCI 算法对光谱特性接近的植被有较好的分类性能。

(4)为了验证监督数据对算法的影响,用不同比例的训练样本来测试算法,如图 9 所示,印第安农林数据集总体分类精度 OA 在训练样本为 3% 时就超过 90%,萨里斯山谷数据集总体分类精度 OA 在训练样本为 0.3% 就超过了 95%,验证了 CHISCI 算法在高光谱图像分类中的优越性。

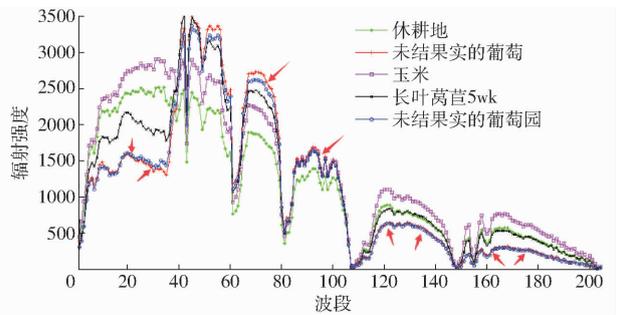


图8 植被光谱图

Fig. 8 Vegetation spectrum

4 结论

(1)CHISCI 算法对高光谱像元分类精度高。试验结果表明,CHISCI 算法比单纯光谱信息 SVM 算法、空谱结合分类算法(SGB-SVM、SBL-SVM 和 SGD-SVM)、边缘保持分类算法(EPF)以及 IFRF 算法相比,OA、AA 和 Kappa 系数都有较大幅度提高,印第安农林数据集的训练比例为 6%,OA、AA

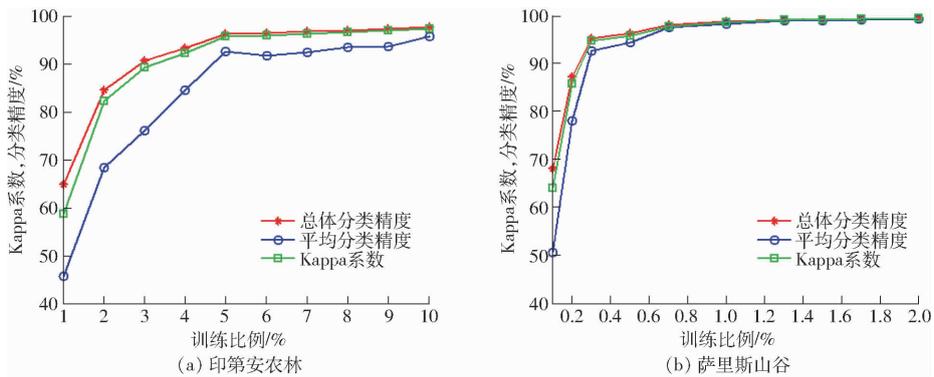


图9 不同训练样本比例分类性能

Fig. 9 Classifications of different training samples

和 Kappa 系数分别达到了 96.16%、92.48% 和 95.62, 而萨里斯山谷数据集的训练比例仅为 1%, 3 种分类精度系数分别达到了 98.67%、98.23% 和 98.52, 充分说明 CHISCI 算法有效地用域转换线性插值卷积滤波提取空间自相关信息, 并基于自相关信息来提高高光谱图像的分类精度。

(2) CHISCI 算法有效去除高光谱图像分类的椒盐现象。域转换线性插值卷积滤波提取高光谱图像

的空间自相关信息, 并利用空间自相关信息进行高光谱植被分类, 能在有限的训练标签中获得较高的分类精度, 且有效去除椒盐现象。

(3) CHISCI 算法对光谱特性接近的植被有较好的分类性能。植被具有丰富的光谱特性, 且光谱相似, 试验结果表明, CHISCI 可以利用空间自相关信息有效对光谱特性相似的植被实现高精度分类, 弥补了用光谱信息进行分类的不足。

参 考 文 献

- 童庆禧, 张兵, 张立福. 中国高光谱遥感的前沿进展[J]. 遥感学报, 2016, 20(5): 689-707.
TONG Qingxi, ZHANG Bing, ZHANG Lifu. Current progress of hyperspectral remote sensing in China[J]. Journal of Remote Sensing, 2016, 20(5): 689-707. (in Chinese)
- 王俊淑, 江南, 张国明, 等. 高光谱遥感图像 DE-self-training 半监督分类算法[J/OL]. 农业机械学报, 2015, 46(5): 239-244. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20150534&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2015.05.034.
WANG Junshu, JIANG Nan, ZHANG Guoming, et al. Semi-supervised classification algorithm for hyperspectral remote sensing image based on DE-self-training [J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(5): 239-244. (in Chinese)
- IMANI M, GHASSEMIAN H. Edge patch image-based morphological profiles for classification of multispectral and hyperspectral data[J]. IET Image Processing, 2017, 11(3): 164-172.
- 邵军勇, 潘泉. 高光谱遥感在植被精细分类中的应用[J]. 微电子学与计算机, 2005(10): 12-13, 19.
SHAO Junyong, PAN Quan. The application of hyperspectral remote-sense to plant fine-classification [J]. Microelectronics & Computer, 2005(10): 12-13, 19. (in Chinese)
- 陶秋香, 张连蓬, 李红梅. 植被高光谱遥感分类中训练样本的选择方法[J]. 国土资源遥感, 2005, 17(2): 33-35, 81.
TAO Qiuxiang, ZHANG Lianpeng, LI Hongmei. The methods for selecting training samples in vegetation classification based on hyperspectral remote sensing [J]. Remote Sensing for Land & Resources, 2005, 17(2): 33-35, 81. (in Chinese)
- 陶秋香. 植被高光谱遥感分类方法研究[J]. 山东科技大学学报: 自然科学版, 2007, 26(5): 61-65.
TAO Qiuxiang. Vegetation classification methods based on hyperspectral remote sensing [J]. Journal of Shandong University of Science and Technology: Natural Science, 2007, 26(5): 61-65. (in Chinese)
- HU B, FREEMANTLE J, MILLER J, et al. Vegetation classification using hyperspectral and multi-angular remote sensing data [C]//2007 IEEE International Geoscience and Remote Sensing Symposium, 2007: 1749-1750.
- 董连英, 邢立新, 潘军, 等. 高光谱图像植被类型的 CART 决策树分类[J]. 吉林大学学报: 信息科学版, 2013, 31(1): 83-89.
DONG L Y, XING L X, PAN J, et al. Vegetation classification in hyperspectral image with CART decision tree [J]. Journal of Jilin University: Information Science Edition, 2013, 31(1): 83-89. (in Chinese)
- 马心璐, 任志远, 王永丽. 支持向量机在高光谱遥感图像植被分类中的应用[J]. 农业系统科学与综合研究, 2009(2): 204-207.
MA Xinlu, REN Zhiyuan, WANG Yongli. Research on hyperspectral remote sensing image classification based on SAM [J]. System Sciences and Comprehensive Studies in Agriculture, 2009(2): 204-207. (in Chinese)
- FILIPPI A M, JENSEN J R. Effect of continuum removal on hyperspectral coastal vegetation classification using a fuzzy learning vector quantizer[J]. IEEE Transactions on Geoscience and Remote Sensing, 2007, 45(6): 1857-1869.
- WANG M, ZHANG L, CHEN S, et al. The analysis about factors influencing the supervised classification accuracy for vegetation hyperspectral remote sensing imagery[C]//2011 4th International Congress on Image and Signal Processing (CISP), 2011, 3:

- 1685 - 1689.
- 12 ZHANG X, SUN Y, SHANG K, et al. Crop classification based on feature band set construction and object-oriented approach using hyperspectral images[J]. *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, 2016, 9(9): 4117 - 4128.
- 13 吴见, 彭道黎. 基于空间信息的高光谱遥感植被分类技术[J]. *农业工程学报*, 2012, 28(5): 150 - 153.
WU Jian, PENG Daoli. Vegetation classification technology of hyperspectral remote sensing based on spatial information[J]. *Transactions of CSAE*, 2012, 28(5): 150 - 153. (in Chinese)
- 14 CUI M, PRASAD S, BRUCE L M, et al. Robust spatial-spectral hyperspectral image classification for vegetation stress detection [C]//2012 IEEE International Geoscience and Remote Sensing Symposium, 2012: 5486 - 5489.
- 15 SHANG K, ZHANG X, SUN Y, et al. Sophisticated vegetation classification based on feature band set using hyperspectral image [J]. *Spectroscopy and Spectral Analysis*, 2015, 35(6): 1669 - 1676.
- 16 MURA M D, BENEDIKTSSON J A, WASKE B, et al. Morphological attribute profiles for the analysis of very high resolution images[J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2010, 48(10): 3747 - 3762.
- 17 GHAMISI P, BENEDIKTSSON J A, ULFARSSON M O. Spectral-spatial classification of hyperspectral images based on hidden Markov random fields[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2014, 52(5): 2565 - 2574.
- 18 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.
- 19 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.
- 20 SONG H, WANG Y. A spectral-spatial classification of hyperspectral images based on the algebraic multigrid method and hierarchical segmentation algorithm[J]. *Remote Sensing*, 2016, 8(4): 296.
- 21 SHEN L, BAI L. Mutual boost learning for selecting Gabor features for face recognition[J]. *Pattern Recognition Letters*, 2006, 27(15): 1758 - 1767.
- 22 WANG L, HAO S, WANG Y, et al. Spatial-spectral information-based semisupervised classification algorithm for hyperspectral imagery[J]. *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, 2014, 7(8): 3577 - 3585.
- 23 RAJADELL O, GARCÍA-SEVILLA P, PLA F. Spectral-spatial pixel characterization using Gabor filters for hyperspectral image classification[J]. *IEEE Geoscience and Remote Sensing Letters*, 2013, 10(4): 860 - 864.
- 24 TOMASI C, MANDUCHI R. Bilateral filtering for gray and color images[C]//1998 Sixth International Conference on Computer Vision, 1998: 839 - 846.
- 25 KOTWAL K, CHAUDHURI S. Visualization of hyperspectral images using bilateral filtering [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2010, 48(5): 2308 - 2316.
- 26 WANG K, HUANG R, SONG Q. Spectral-spatial hyperspectral image classification using extended multi attribute profiles and guided bilateral filter[C]//2015 International Conference on Computer Science and Mechanical Automation (CSMA), 2015: 235 - 239.
- 27 KANG X, LI S, 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.
- 28 廖建尚, 王立国, 郝思媛. 基于双边滤波和空间邻域信息的高光谱图像分类方法[J/OL]. *农业机械学报*, 2017, 48(8): 140 - 146, 211. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20170815&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2017.08.015.
LIAO Jianshang, WANG Liguang, HAO Siyuan. Hyperspectral image classification method combined with bilateral filtering and pixel neighborhood information[J/OL]. *Transactions of the Chinese Society for Agricultural Machinery*, 2017, 48(8): 140 - 146, 211. (in Chinese)
- 29 KANG X, LI S, BENEDIKTSSON J A. Feature extraction of hyperspectral images with image fusion and recursive filtering[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2014, 52(6): 3742 - 3752.
- 30 GASTAL E S L, OLIVEIRA M M. Domain transform for edge-aware image and video processing[J]. *ACM Transactions on Graphics (TOG)*, 2011, 30(4): 69.
- 31 MORAN P A P. The interpretation of statistical maps[J]. *Journal of the Royal Statistical Society. Series B (Methodological)*, 1948, 10(2): 243 - 251.
- 32 MORAN P A P. Notes on continuous stochastic phenomena[J]. *Biometrika*, 1950, 37(1-2): 17 - 23.
- 33 MELGANI F, BRUZZONE L. Classification of hyperspectral remote sensing images with support vector machines[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2004, 42(8): 1778 - 1790.
- 34 CAI S, DU Q, MOORHEAD R, et al. Noise-adjusted principal component analysis for hyperspectral remotely sensed imagery visualization[C]//16th IEEE Visualization Conference (Vis2005), 2005: 119 - 120.
- 35 HE K, SUN J, TANG X. Guided image filtering[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2013, 35(6): 1397 - 1409.