引文格式：李倩楠,彭菲,虞瑶. 基于多特征随机最近正则化子空间的高光谱影像分类算法[J]. 地理空间信息, 2025, 23(4):25-29. doi:10. 3969/j. issn. 1672-4623.2025.04.006

基于多特征随机最近正则化子空间的高光谱影像分类算法

李倩楠[1]，彭　菲[1]，虞　瑶[1]
（1. 江苏省基础地理信息中心，江苏 南京 210013）

摘　要：基于单一特征的随机子空间的最近正则化子空间集成算法（random eubspace based nearest regularized subspace，RSNRS）精度受限，因此设计一种基于多特征的随机最近正则化子空间（multi-meature random subspace-based nearest regularized Subspace，MFRSNRS）的高光谱遥感算法。提取高光谱遥感影像的Gabor特征、局部二值模式特征（local binary pattern，LBP）特征、拓展形态学多属性剖面（extended multi-attribute profiles，EMAPs）特征和光谱特征，利用随机子空间的集成方式构建特征子空间，基于特征子空间训练基分类器最近正则化子空间（nearest regularized subspace，NRS），并将基于分类器的输出结果进行集成得到分类结果。MFRSNRS算法在数据集Indian Pines总体精度达到92.88%，与RSNRS算法相比，总体精度提高了17.05%。在数据集Pavia University上总体精度达到78.88%，与RSNRS算法相比，总体精度提高7.73%。实验结果表明，多特征的使用具有一定的提升效果。

关键词：多特征；最近规则子空间；高光谱影像分类；随机子空间
中图分类号：P237　　**文献标志码：**B　　**文章编号：**1672-4623（2025）04-0025-05

Hyperspectral Image Classification Algorithm Based on Multi-feature and Random Subspace Nearest Regularized Subspace

LI Qiannan[1], PENG Fei[1], YU Yao[1]
(1. Basic Geographic Information Center of Jiangsu Province, Nanjing 210013, China)

Abstract: The random subspaces nearest regularized subspace (RSNRS) based on a single feature has a limited accuracy, we proposed a new hyperspectral remote sensing algorithm named multi-feature random subspace nearest regularized subspace (MFRSNRS). Firstly, we extracted Gabor features, LBP features, EMAPs features and spectral features by hyperspectral remote sensing images. Then, we selected the nearest regularized subspace (NRS) as the base classifier, and used the random subspace method to construct the feature subspace. Finally, we trained the base classifiers based on the feature subspace, and integrated the results of base classifiers to obtain the final classification results. The OA of MFRSNRS algorithm reached 92.88% in the Indian Pines data set, which is 17.05% higher than RSNRS algorithm. The OA of MFRSNRS algorithm reached 78.88% in the Pavia University data set, which is 7.73% higher than RSNRS algorithm. The experimental result indicates that the use of multiple features has shown a certain improvement in performance.

Key words: multi-feature, NRS, hyperspectral image classification, random subspace

成像光谱技术不断发展和进步，高光谱遥感影像的光谱分辨率随之不断提高，数据维度也相应不断增加，这就直接导致了高光谱遥感数据处理中面临的“维度灾难”问题[1-3]。因此，在高光谱遥感数据的分类研究中，如何有效地解决“维度灾难”问题成为一大研究热点。

随机子空间集成学习算法可以有效的解决“维度灾难”问题。该集成学习算法核心思想为从原始特征空间中随机选取部分特征得到新的特征子集，不仅有效缓解维度灾难问题，更增加了特征子集间的差异性，从而有效提升基分类器的分类表现。其中，基于随机子空间的最近正则化子空间集成算法（random subspace based nearest regularized subspace classification，RSNRS）中选择最近正则化子空间（nearest regularized subspace，NRS）为基分类器，利用随机子空间的集成方式取得较优的分类表现[4]。但该算法未考虑单一的光谱特征，只从某个特定的角度描述地物目标，而多特征可以有效地提供地物目标的多重信息。

多特征与单一特征相比，一般提供3种及以上特征参与高光谱的影像分类，一般包含光谱特征、纹理特征（如Gabor特征）和形状特征（如扩展属性剖面）等。研究表明多特征的融合可在高光谱遥感表示分类中体现明显的优势[5-7]。多特征可以从光谱、纹理、形状等多个角度互补地定量描述地物目标，从而

项目来源：江苏自然资源智库2023年度开放合作项目（ZK23002）。
第一作者简介：李倩楠（1991—），硕士，工程师，从事自然资源调查监测数据分析与应用研究工作，E-mail：yaogan1126@163.com 。

提升分类效果[8-10]。

多特征融合虽然提升了特征间的互补性，但易造成波段数目多和计算复杂度高的问题，因此本文提出了一种基于多特征的最近正则化子空间(multi-feature random subspace based nearest regularized subspace，MFRSNRS)的高光谱影像分类方法。基于高光谱遥感影像提取的Gabor特征、LBP特征、拓展形态学多属性剖面（extended multi-attribute profiles，EMAPs）和光谱特征，以堆栈的方式融合多特征。采用随机子空间的集成方式随机选取部分特征形成多个特征子集，利用特征子集构建特征子空间，在每个特征空间内选择基分类器NRS进行分类，并将所有基分类器的输出结果进行集合。实验部分使用AVIRIS传感器采集的Indian Pines数据和ROSIS传感器采集的Pavia University数据对提出的MFRSNRS算法进行了验证。

1 多特征与RSNRS算法

1.1 多特征

1.1.1 Gabor特征

Gabor滤波具有光照变化适应能力强等特点。在计算机视觉和图像分析等领域，Gabor滤波在纹理特征提取发挥重要作用。常见的二维Gabor滤波器表达式为：

$$g(a,b,\delta,\theta,\psi,\sigma,\gamma)=\exp(-\frac{a'^2+\gamma^2 b'^2}{2\sigma^2})\times \exp(j(2\pi\frac{a'}{\delta}+\psi))$$
$$a'=a\cos\theta+b\sin\theta \quad (1)$$
$$b'=-a\cos\theta+b\sin\theta$$

式中，δ 为正弦波函数中的波长因子；θ 为Gabor小波的方向；ψ 为正弦波函数中的相位差；γ 为空间比率（一般取值设为0.5）；（a，b）为原始二维坐标系中的空间位置坐标；（a'，b'）为变换后坐标系中的空间位置坐标；σ 为高斯分布的标准差，可以决定Gabor滤波器的尺寸和形状。该数值与波长因子 δ 和空间尺度频率波段宽度 bw 相关，具体计算公式如下：

$$\sigma=\frac{\delta}{\pi}\sqrt{\frac{\ln 2}{2}}\left(\frac{2^{bw}+1}{2^{bw}-1}\right) \quad (2)$$

1.1.2 LBP特征

局部二值模式LBP也是一种常见的图像纹理特征，在灰度不变性和旋转不变性等方面具有突出的优势。该算法的原理为根据每个像素点与其邻域像素间的灰度差异，在邻域区域内的不同位置进行二值编码处理，从而确定该像素点所处区域的纹理特征。假设邻域区域内范围大小设置3×3，其中心像素点的灰度值设为 I_c，邻域8个像素的灰度值设为 I_n，$n=1,2,\dots,8$。依次将8个邻域像素的灰度值 I_n 与中心像素灰度值 I_c 进行对比。如果邻域像素的灰度值大于中心像素灰度值，即表达式 $I_n-I_c\geq 0$，则二值函数 $s(I_n-I_c)=1$，反之表达式 $I_n-I_c<0$，则二值函数 $s(I_n-I_c)=0$。邻域8个像素点经过与中心像素比较之后可以得到8位二进制数，然后将其转化为十进制数，可得到该中心像素点的LBP编码值。最终的LBP编码值可用来表达该像素点所在位置的局部纹理结构信息。计算公式如下：

$$\mathrm{LBP}_{N,R}=\sum_{n=0}^{N-1}s(I_n-I_c)2^n \text{ [10]}$$
$$s(x)=\begin{cases}1 & x\geq 0\\ 0 & x<0\end{cases} \quad (3)$$

1.1.3 EMAPs特征

基于多属性滤波的拓展形态学多属性剖面EMAPs是一种经典的图像结构特征提取方法。该方法通过使用不同尺度和方向上的滤波器，对图像进行滤波操作，从而提取出图像中的纹理信息。属性剖面（attribute profiles，AP）可以作为粗化和细化操作的组合[11]，具体公式如下，其中 f 表示输入图像：

$$AP(f)=\prod_i:\begin{cases}\prod_i=\prod_{\Phi\ T_\lambda'},\lambda=(n-1+i)\ \forall\lambda\in[1,n]\\ \prod_i=\prod_{\gamma\ T_\lambda'},\lambda=(i-n-1)\ \forall\lambda\in[n+1,2n+1]\end{cases} \quad (4)$$

式中，$\prod_{\Phi\ T'}$ 为细化操作；$\prod_{\gamma\ T'}$ 为粗化操作。高光谱遥感影像波段数目一般达到100以上，采用主成分分析方法（principal component analysis，PCA）对遥感影像进行降维，得到前 m 个主成分。对前 m 个主成分（如FR_1，FR_2，…，FR_m）采用形态学属性剖面操作获得AP。一系列AP堆栈而成得到扩展属性剖面（extended attribute profiles，EAP）。可表示为[11]：

$$\mathrm{EAP}=\{\mathrm{AP}(\mathrm{FR}_1),\mathrm{AP}(\mathrm{FR}_2),\dots,\mathrm{AP}(\mathrm{FR}_m)\} \quad (5)$$

每一种属性滤波可以从某个角度得到图像的结构特征。本文使用4种属性滤波，分别为面积、标准差、对角量和惯性矩。基于不同属性滤波提取的EAP进行叠加，最终形成EMAP特征。如由 q 个EAP组成的EMAP特征可表示为：

$$\mathrm{EMAP}=\{\mathrm{EAP}_{a_1},\mathrm{EAP}'_{a_2},\dots,\mathrm{EAP}'_{a_q}\} \quad (6)$$

式中，a_i 为一个属性滤波：

$$\mathrm{EMAP}'=\mathrm{EAP}\{\mathrm{FR}_1,\mathrm{FR}_2,\dots\mathrm{FR}_m\} \quad (7)$$

1.2 RSNRS算法

在高光谱遥感数据中，特征间存在较强的相关性

和冗余度，这会导致算法模型过于复杂、容易造成过拟合等问题。利用随机子空间的方法可以明显降低特征间的相关性，减少冗余信息，提高算法的可解释性和鲁棒性[12]。文献[13]设计的RSNRS算法的具体步骤如下：假设高光谱遥感影像训练集设置为 $X=\{x_i\}_{i=1}^{n}$ 的特征维度（波段数目）为 N，n 为总的训练样本数量，x_i 为第 i 个训练样本。从训练集 X 的全部波段中按照一定方式随机选取若干波段，从而组成新的特征子集 S_t，$t=1,2,\ldots,T$，T 为集成次数，特征子集 S_t 的维度大小为 $N'=N\times d$，d 表示特征比例。T 次集成后，得到特征子集的集合 $S=\{S_1,S_2,\ldots,S_T\}$。在第 t 次集成过程中，特征子集构建的新字典表示为 $\hat{D}=[D_{1t},D_{2t},\ldots,D_{Ct}]$，样本类别为 $l\in\{1,2,\ldots,C\}$，C 为全部的地物类别数目。第1类的样本字典表示为 $D_{l,t}$，样本数量为 n_l，则 $\sum_{l=1}^{C}n_l=n$。在每个特征子集中 S_t 利用NRS算法对测试样本 y 进行分类，测试样本可由训练样本线性组合表示，第 t 次集成中的测试样本 y_t，某类测试样本的表示估计值 $\tilde{y}_{l,t}$ 可表示为：

$$\tilde{y}_{l,t}=D_{l,t}\alpha_{l,t} \tag{8}$$

通过上式获取每一类的表示估计值 $\tilde{y}_l$，测试样本 y 的标签值由表示残差最小的准则确定。

$$\begin{cases}\text{class}(y)=\arg\min\limits_{l=1,2\ldots C}(r_{l,t})\\ r_{l,t}=\left\|\tilde{y}_{l,t}-y_t\right\|_2^2\end{cases} \tag{9}$$

在NRS算法中，表示系数 α_l 求解可以转化为优化问题：

$$\alpha_{l,t}=\arg\min_{\alpha^*}\left\|y_t-D_{l,t}\alpha^*\right\|_2^2+\lambda\left\|\Gamma_{l,y,t}\alpha^*\right\|_2^2 \tag{10}$$

集成 T 次后，T 个特征子集中对测试样本有 T 个分类结果，采用多数投票法的方法最终确定测试样本的标签值。

2　基于多特征的随机正则化子空间算法

研究表明，针对遥感影像的复杂多变的地理环境而言，单一特征很难解决所有的分类问题。多特征中所包含的互补性可有效解决多类别分类问题。因此提出基于多特征的随机正则化子空间MFRSNRS算法。参考Su[10]拓展的非线性多特征字典学习协同表示分类模型（MMKCRC-DL），选择光谱特征、Gabor特征、LBP特征和EMAPs特征。

首先采用主成分分析的方法，得到前3个主成分，基于前3个主成分提取影像的Gabor特征、LBP特征、EMAPs特征和原始的光谱特征，多种特征通过堆栈的方式级联。多种特征通过不同角度提取影像的各种信息，虽在一定程度上弥补了多种特征，但不可避免地造成空间维度高的问题。因此考虑随机子空间的集成方式，集成过程中通过随机选取波段的方式构造多个特征空间，基于构建的特征空间，采用基分类器NRS对影像进行分类。每个特征空间中的NRS模型参与测试样本标签值的决策，最后采用多数投票法的方式决定测试样本最终的分类结果。

3　实验与分析

3.1　实验设置

为了验证MFRSNRS算法的有效性，选择AVIRIS和ROSIS 2种传感器采集的高光谱图像分类数据集。两组实验采用协同表示正则化分类器（classification representation classification with tikhonov，CRT）、最近正则化子空间（nearest regularized subspace，NRS）和随机子空间最近正则化子空间作为对比算法。分类性能评价指标是评估分类算法效果的重要依据，实验选用总体分类精度（overall accuracy，OA）、平均分类精度（average accuracy，AA）和Kappa系数3种评价指标。Indian Pines和Pavia University两组数据集中4种分类算法设置的最优参数如表1所示，其中 T 为集成次数，λ 为正则化参数，d 为特征比例。将所有实验重复5次取平均值，利用交叉验证的方式确定算法参数。

表1　各算法最佳参数

数据集		Indian Pines	Pavia University
CRT	λ	1e−2	1e−2
NRS	λ	1e+0	1e−1
RSNRS	λ	1e+0	1e−1
	d	0.1	0.1
	T	20	10
MFRSNRS	λ	1e−1	1e−1
	d	0.9	0.3
	T	20	10

3.2　Indian Pines高光谱影像实验

第一组实验数据为AVIRIS传感器采集的位于美国印第安纳西北部地区的Indian Pines高光谱影像数据，截取大小为145×145的图像并进行标注，包含16类地物，一共10 249个样本。该数据去除部分波段后保留200个波段用于实验，波长范围区间为400~2 500 nm。空间分辨率较低，大约为20 m[14]。本文实验中共选择9类地物，主要包含牧草、乔木和干草等地类。实验过程

中选择有标签的样本，一部分作为训练样本，剩余的标签样本作为测试样本用于验证分类精度。对于Indian Pines数据，每类别选取3%的样本作为训练样本，剩余样本全部作为测试样本，具体样本数量见表2。

表2 Indian Pines分类精度

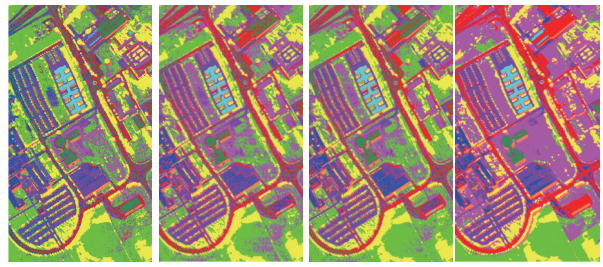
类别	训练样本	测试样本	CRT	NRS	RSNRS	MFRSNRS
1	43	1 385	67.50	58.32	69.77	88.28
2	25	805	54.74	57.25	74.03	94.05
3	15	468	86.15	80.85	91.06	98.14
4	22	708	90.76	92.30	93.00	97.67
5	15	463	94.30	95.68	97.73	99.91
6	30	942	57.08	57.60	73.23	94.00
7	74	2 381	65.10	63.90	64.77	89.00
8	18	575	54.98	44.11	70.70	91.00
9	38	1 227	98.15	96.59	97.46	99.72
	OA/%		72.36	69.43	75.83	92.88
	AA/%		74.31	71.84	81.31	94.64
	Kappa		0.674 5	0.643 3	0.712 9	0.916 6

CRT、NRS、RSNRS和MFRSNRS 4种算法详细的分类结果见表2。本文提出的MFRSNRS算法*OA*值、*AA*值和Kappa系数均最高。分别比RSNRS方法的*OA*值高17.05%，*AA*值高13.33%，Kappa值高0.203 7。MFRSNRS算法在9类地物类别均取得了最高的精度，其中在第5类别上精度达到99.91%，在第9类别上精度达到99.72%。4种算法得到的分类结果，与3种对比方法相比，MFRSNRS方法分类效果更好，可以更为精准地刻画地物分布格局。

3.3 Pavia University高光谱影像实验

第二组实验数据由ROSIS传感器采集的意大利帕维亚大学（Pavia University）高光谱遥感数据。它截取大小为610×340的图像并进行标注，空间分辨率为1.3 m分辨率较高，达到米级分辨率。波谱范围区间为430~860 nm。该数据去除部分波段后保留103个波段用于实验。数据共有9种地物类型，主要包含柏油路面（Asphalt）、草地（Meadows）、树木（Tree）和金属屋顶建筑（Painted metal sheets）等地物[15]。实验过程中训练样本为每类选取5个样本，其余样本作为测试样本以验证精度。

CRT、NRS、RSNRS和MFRSNRS4种算法在Pavia University数据集详细的分类结果见图1和表3。MFRSNRS方法的*OA*值、*AA*值和Kappa系数均最高。比RSNRS方法的*OA*值高7.73%，*AA*值高11.01%，Kappa值高0.117 1。MFRSNRS在9类中均取得最高精度，其中在第5类别精度达到97.81%，在第9类别精度达到99.98%。



a CRT b NRS c RSNRS d MFRSNRS

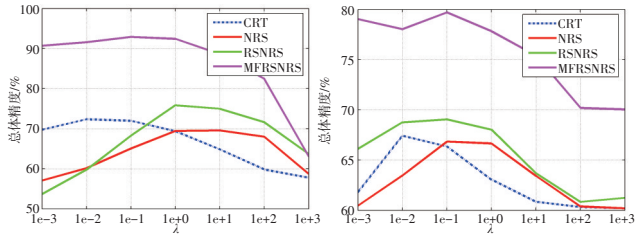
图1 Pavia University数据不同算法分类图

表3 Pavia University分类精度

类别	训练样本	测试样本	CRT	NRS	RSNRS	MFRSNRS
1	5	6 626	93.39	87.86	89.67	98.78
2	5	18 644	87.12	86.05	86.14	95.92
3	5	2 094	46.22	56.67	57.31	58.66
4	5	3 059	63.14	68.41	68.40	72.28
5	5	1 340	95.83	95.93	95.11	97.81
6	5	5 024	45.19	37.92	42.24	49.90
7	5	1 325	32.87	44.80	48.04	93.95
8	5	3 677	56.77	60.03	64.55	83.25
9	5	942	100	99.83	99.98	99.98
	OA/%		67.42	66.84	71.15	78.88
	AA/%		68.95	70.83	72.38	83.39
	Kappa		0.597 9	0.592 3	0.628 9	0.746 0

3.4 参数分析

1）正则化参数λ。本节研究λ对两组数据集总体精度的影响，λ对两组数据集的总体精度变化如图2所示。横坐标表示为正则化参数，纵坐标表示总体精度，其余参数设置见表1。在Indian Pines数据上，CRT算法的最优正则化参数为1e−2，NRS算法最优正则化参数为1e+0，RSNRS算法正则化参数为1e+0，MFRSNRS算法最优正则化参数为1e−1。在Pavia University数据上，CRT算法的最优正则化参数为1e−2，NRS算法的最优正则化参数为1e−1，RSNRS算法的最优正则化参数为1e+0，MFRSNRS算法的最优正则化参数为1e−1。

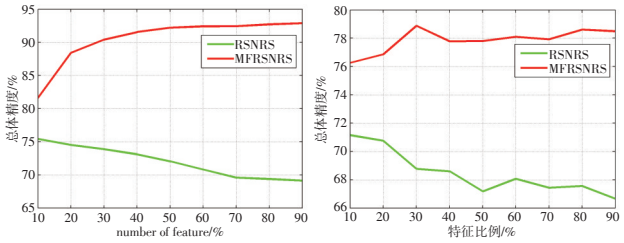


a λ对Indian Pines数据的影响 b λ对Pavia University数据的影响

图2 λ对两组数据集总体精度的影响

2）特征比例d。本节研究d值对2组数据集的影响，不同d值2组数据集总体精度变化图如图3所示。的取值范围区间为{0.1，0.2，0.3，0.4，0.5，0.6，0.7，0.8，0.9}，其余参数设置如表1所示。

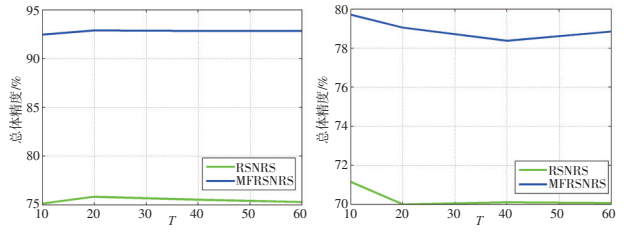
在 Indian Pines 数据上，RSNRS 算法的总体精度随着特征比例的不断增加总体精度随着下降，在 d 值为 10%时，总体分类精度达到最大值。MFRSNRS 算法的总体精度随着特征比例的不断增加总体精度不断上升，在 d 值为 90%时，总体分类精度达到最大值。在 Pavia University 数据上，RSNRS 算法的分类总体精度在 66%~71%之间，最佳特征比例参数值为 10%。MFRSNRS 算法获得的总体分类精度在 76%~78%之间，最佳特征比例参数值为 30%。总体来说，MFRSNRS 算法优于 RSNRS 算法。



a d 值对 Indian Pines 数据的影响　b d 值对 Pavia University 数据的影响

图3　不同 d 值下两组数据集总体精度变化图

3）集成次数 T。本节研究 T 值对两组数据集的影响，不同 T 值两组数据集总体精度变化图如图 4 所示。横坐标为 T 值，纵坐标为总体精度，其余参数设置如表 1 所示。在 Indian Pines 数据上，MFRSNRS 和 RSNRS 2 种算法随着集成次数 T 的增加，总体分类精度先增加后下降，最佳 T 值为 20。在 Pavia University 数据上，MFRSNRS 和 RSNRS 两种算法的最佳 T 值均为 10。



a T 值对 Indian Pines 数据的影响　b T 值对 Pavia University 数据的影响

图4　不同 T 值下两组数据集总体精度变化图

4　结　语

本文针对随机子空间的最近正则化子空间集成算法（RSNRS）精度受限等问题，设计了一种基于多特征的最近规则子空间的高光谱遥感影像分类 MFRSNRS 方法。首先基于遥感影像提取 Gabor 特征、LBP 特征和 EMAPs 特征，组合光谱信息形成多特征集；然后利用随机子空间集成方式随机选取若干大小的特征子集；最后利用 NRS 作为基分类器，采用集成方式得到分类结果。实验结果表明，多特征的使用确实有助于提升分类表现。后续研究方向可考虑子空间的自动划分和并行运算等。

参考文献

[1] 童庆禧,张兵,郑兰芬. 高光谱遥感[M]. 北京:高等教育出版社,2006

[2] 赵英时. 遥感应用分析原理与方法[M]. 北京:科学出版社,2003

[3] 张良培,张立福. 高光谱遥感[M]. 北京:测绘出版社,2011

[4] Li W, Tramel E W, Prasad S, et al. Nearest Regularized Subspace for Hyperspectral Classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52(1): 477-489

[5] Chen C, Li W, Su H, et al. Spectral-spatial Cassification of Hyperspectral Image Based on Kernel Extreme Learning Machine[J]. Remote Sensing, 2014, 6(6): 5 795-5 814

[6] Gan L, Xia J, Du P, et al. Multiple Feature Kernel Sparse Representation Classifier for Hyperspectral Imagery[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(9): 5 343-5 356

[7] 纪磊. 基于空谱联合的高光谱图像分类方法研究[D]. 贵阳: 贵州大学

[8] Li W, Du Q. Gabor-fltering-based Nearest Regularized Subspace for Hyperspectral Image Classification[J]. IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing, 2014, 7(4):1 012-1 022

[9] Peng B, Li W, Xie X, et al. Weighted-fusion-based Representation Classifiers for Hyperspectral Imagery[J]. Remote Sensing, 2015, 7(11): 14 806-14 826

[10] Su H, Zhao B, Du Q, et al. Multi-feature Dictionary Learning for Collaborative Representation Classification of Hyperspectral Imagery[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018(10):1-18

[11] 阿茹罕,原晓斌,慕晓冬,等. 结合扩展多属性剖面与快速局部RX算法的高光谱异常目标检测[J]. 光子学报,2021,50(9):289-299

[12] 程圆娥,吕志慧、袁春琦. 结合随机子空间与SSAE-LR的高光谱图像分类[J]. 地理空间信息,2020,18(9):7-12

[13] 虞瑶,杨锦,邱洁. 随机子空间和MRF的高光谱遥感集成学习算法[J]. 测绘科学,2022,47(10):161-167

[14] 虞瑶,苏红军,陶旸. 联合空间信息的高光谱遥感协同表示动态集成分类算法[J]. 遥感学报,2024,28(1):187-202

[15] 杜培军,张伟,张鹏,等. 一种联合空谱特征的高光谱影像分类胶囊网络[J]. 测绘学报,2023,52(7):1 090-1 104