

高光谱图像的JM变换自适应降维

康孝岩^{1,2}, 张爱武^{1,2}, 胡少兴³, 肖青⁴, 柴沙驼⁵

1. 首都师范大学 三维信息获取与应用教育部重点实验室, 北京 100048;

2. 首都师范大学 空间信息技术教育部工程研究中心, 北京 100048;

3. 北京航空航天大学 机械工程及自动化学院, 北京 100191;

4. 中国科学院遥感与数字地球研究所 遥感科学国家重点实验室, 北京 100101;

5. 青海大学 畜牧兽医科学院, 西宁 810016

摘要: 在无需先验标签样本的情况下, 非监督降维可以有效简化高光谱图像的特征空间, 避免目标分类中的霍夫效应。本文提出JM非线性变换优化的自适应降维模型来研究面向图像目标分类的高光谱波段选择问题。该方法考虑波段的信息量和独立性等两个重要因子, 针对其测度方法的差异性, 引入JM变换函数进行规范化优化。选用线阵高光谱和面阵显微光谱等两个图像数据集, 在 k 最近邻和随机森林分类器下, 进行了多组监督分类实验, 结果表明, 在Kappa系数、总体分类精度和平均分类精度上, 本文方法均优于3种非监督方法MABS、InfFS和LSFS。说明本文提出的JM变换的自适应降维模型能够有效降低特征维度, 满足高光谱图像分类的高精度要求。

关键词: 遥感, JM变换, 规范化, 自适应降维, 非监督波段选择, 高光谱图像

引用格式: 康孝岩, 张爱武, 胡少兴, 肖青, 柴沙驼. 2020. 高光谱图像的JM变换自适应降维. 遥感学报, 24(1): 67-75

Kang X Y, Zhang A W, Hu S X, Xiao Q and Chai S T. 2020. Hyperspectral images adaptive dimensionality reduction optimized by JM transform. Journal of Remote Sensing (Chinese), 24(1): 67-75 [DOI: 10.11834/jrs.20208178]

1 引言

目标光谱特征是遥感图像目标属性识别的重要基础, 不同目标均有其特有的电磁波辐射特征(童庆禧等, 2016)。传统的遥感技术仅在若干宽波段上获取图像, 这样便会丢失大量的光谱信息, 对目标的检测、识别和分类等带来较大的不确定性(张康等, 2018)。而高光谱目标感知技术则致力于在光谱维度上, 对目标进行精细刻画, 以描绘其连续的辐射光谱特征曲线; 与此同时, 兼容传统遥感对目标空间维的表达能力, 高光谱遥感具有了“图谱合一”的显著特点(余旭初等, 2013)。然而, 鉴于其波段之间往往存在较高的相关性, 高光谱图像的原始数据空间具有高度的冗余信息, 在处理时容易出现“维数灾难”问题; 故而, 有效地降低数据空间维度即降维, 有助于更好地理

解和认知高光谱图像, 也更有利于图像目标分类(孙伟伟等, 2018)。

作为自适应降维的主要方式之一, 非监督特征(波段)选择以“高信息量”和“强独立性”为准来寻求特征空间的简化, 按照算法复杂程度不同, 可将其大致分为两类: 一类需要迭代计算、字典学习和微分几何求解等操作, 以损失一定时间为代价, 来有效提升算法精度, 常用方法有遗传算法(王立国和魏芳洁, 2013)、稀疏学习(Li和Qi, 2011)、图论方法(Roffo等, 2015; Roffo和Melzi, 2017)和流形理论(He等, 2005; Sun等, 2015)等; 而另一类具有较低的算法时间复杂度, 往往运用数理统计学中的经典结论, 来实现即时或实时的算法效率, 但其算法精度有待进一步提升。其中, 后者方法为本文关注的重点。

秦方普等(2015)提出了一种基于互信息统计量

收稿日期: 2018-04-10; 预印本: 2018-05-31

基金项目: 国家自然科学基金(编号:41571369); 国家重点研发计划(编号:2016YFB0502500); 北京市自然科学基金(编号:4162034); 青海省科技计划(编号:2016-NK-138)

第一作者简介: 康孝岩, 1989年生, 男, 博士研究生, 研究方向为计算机视觉、机器学习、高维遥感数据处理与应用。E-mail: xykang@cnu.edu.cn; xy.muap.kxy@gmail.com

通信作者简介: 张爱武, 1972年生, 女, 教授, 研究方向为空间信息获取与处理、三维激光成像、高光谱成像等。E-mail: zhangaw98@163.com

的谱聚类方法, 来对高光谱影像进行维数约简, 但该方法中互信息矩阵构建的非规范化 Laplacian 矩阵存在非正定的可能, 在降维维度较大时会出现偏差。Cao 等(2016)提出一种两次降维的组合算法来获取高光谱图像的低维数据, 首先, 以波段 Canny 边缘结构特征图与全波段结构均值特征图的 Pearson 相关系数来定义波段信息量, 通过设定阈值来剔除低信息量波段; 然后, 在二次降维时分别进行了强分类器交叉验证精度最大化试验、相似波段聚类 and 类内距离均值最小波段搜索等解算步骤, 整个过程较为繁杂。Shukla 和 Nanda(2018)利用香农熵和一阶光谱导数来分别度量波段信息量和冗余量, 并将降维问题转换为优化问题, 提出了一种二元社会蜘蛛群的元启发式算法, 其算法复杂度主要集中在迭代和搜索过程, 效率有待提升。以上 3 种方法均是以典型统计量为基础的降维方法, 但都引入了复杂度较大的分类、空间结构提取或迭代优化方法, 使得算法整体上效率较低。运用 Pearson 相关系数, 张爱武等(2017)构建了一种信息相邻相关系数来表达波段的相关性, 继而提出了一种改进的自适应特征选择方法 MABS (Modified Adaptive Band Selection), 该方法较好地解决了波段独立性度量问题, 所选特征的信息量、独立性以及波谱范围相比于原算法 ABS (Adaptive Band Selection) 均有较大提升, 且算法效率较高, 然而, 其对信息量和独立性的二元指标建模问题的理解有待加深。

鉴于此, 本研究挖掘出模式识别经典方法 JM (Jeffries-Matusita) 非线性变换, 来对 MABS 中的二元自变量—信息量、独立性—进行规范化处理, 提出 JM 变换优化的自适应降维方法 JM²ABS (JM Modified Adaptive Band Selection), 选用线阵和面阵高光谱图像公共数据集进行图像分类实验, 并通过与典型方法做比较, 以验证 JM²ABS 的有效性。

2 JM²ABS 方法

基于概率统计学的经典统计量(如均值、标准差、协方差、香农熵等)及其基础之上构造的复合统计工具(如变异系数(张康等, 2018)、主成分分析、Pearson 相关系数(Cao 等, 2016; 张爱武等, 2017)、互熵(秦方普等, 2015)等)构建的非监督自适应降维方法, 往往有着复杂度低、简单易

行和效率高的优点。

2.1 信息相邻相关系数

高光谱图像波段之间的相关性很强, 用于表达波段间相关性的指标有互信息、Pearson 线性相关系数、一二阶光谱导数、对比指数、光谱比值和关联系数等(Bajcsy 和 Groves, 2004)。一直以来, 易于被理解和获取的相关性指标常被用来间接度量高光谱波段的独立性(秦方普等, 2015; Cao 等, 2016; 张爱武等, 2017)。其中, 张爱武等(2017)通过机理分析和实例验证, 发现刘春红等(2005)提出的波段相邻相关系数并不能有效表达波段独立性, 进而提出了一种信息相邻相关系数来度量波段独立性

$$\text{MABS}_{\text{indep}}^i = \frac{1}{\left(\left| r_{j-1,j}^{\text{sort}(\text{infor})} \right| + \left| r_{j+1,j}^{\text{sort}(\text{infor})} \right| \right) / 2} \quad (1)$$

假设 $A_n: a_1, a_2, a_3, \dots, a_i, \dots, a_n$ 为原始波段的序列; 按照波段信息量 infor 排序后的序列为 $B_n^\sigma: b_1^\sigma, b_2^\sigma, \dots, b_i^\sigma, \dots, b_n^\sigma$; 波段 b_j^σ 为波段 $i(a_i)$ 在 B_n^σ 中的对应波段; 则式中, $\text{MABS}_{\text{indep}}^i$ 为 MABS 中波段 i 的独立性因子; $r_{j-1,j}^{\text{sort}(\text{infor})}$ 和 $r_{j+1,j}^{\text{sort}(\text{infor})}$ 分别为 B_n^σ 中的波段 $j(b_j^\sigma)$ 与前、后波段的相关系数。实验验证了 $\text{MABS}_{\text{indep}}^i$ 对波段独立性的度量能力。

2.2 信息量—独立性的二元复合建模

假设 (X, Y) 为 2 维连续型随机变量, 则其常见的二元复合构型有如下两种

$$Z = aX + bY \quad (2)$$

$$Z = X^\alpha + Y^\beta \quad (3)$$

式(2)中, Z 表示随机变量 X 和 Y 的线性组合, a 、 b 为权重系数(常数); 式(3)中, Z 表示为 X 和 Y 的非线性函数, α 、 β 分别为 X 和 Y 的幂次(常数)。通过调节权重系数或幂次大小, 可以对随机变量的重要程度进行定量调整, 可称之为比例因子, 可根据先验知识或监督训练来确定比例因子大小。

研究发现, 上述独立性度量指标结合信息量对高光谱降维模型的构建属于非线性函数构型(式(3)), 并且均默认了比例因子为 1, 即认为信息量与独立性在对评价波段优劣中同等重要; 然而, 这些方法并未考虑获取的信息量和独立性指标是否在同一规范下。MABS 算法中波段信息量以 DN 值标准差 σ 来表示

$$\sigma_i = \sqrt{\frac{\sum_{p=1}^{nl} \sum_{q=1}^{ns} (\text{DN}(p, q) - \mu_i)^2}{nl \cdot ns}} \quad (4)$$

式中, σ_i 为波段*i*的标准差; nl 、 ns 分别为波段图像的行数和列数; $\text{DN}(p, q)$ 为波段*i*中

行q列像素的DN值, μ_i 为波段*i*的DN值均值。由此可见, 信息量与波段DN值大小有着紧密关系; 进一步讲, 标准差与高光谱的数据位深关系紧密; 可以推断, 对于相同的观测目标, 传感器的数据位深越大, 则获取图像的标准差也将越大, 并且波段标准差的值域为 $[0, F]$, 其中 $F \ll +\infty$ 为正的常数。而MABS的独立性指标本质上是由相关系数得到的, Pearson相关系数的值域大小是明确的 $[-1, 1]$, 仅与波段之间的共变关系有关, 而与单一波段的DN值大小无关; $\text{MABS}_{\text{indep}}^i$ 值域为 $[1, +\infty]$ 。其他几种类似的降维方法存在上述相同的问题。

因此, 本研究认为, MABS降维方法虽然认为波段信息量与独立性的等同作用, 但是在模型构建时并未对两者进行规范化处理, 造成模型表达的物理含义含混和不明确, 既不能表达信息量与独立性的等同作用, 又不能反映出两者的轻重关系。

2.3 JM非线性变换

非线性规范化处理是消除不同随机变量在量纲、值域和尺度等方面差异的有效手段之一, 高斯近似是较为典型的一种, 其常被用于二元非正态分布随机变量在计算Pearson相关系数前的预处理, 近来卷积神经网络研究中为解决“梯度弥散”问题而提出的BN(Batch Normalization)规范化(Ioffe和Szegedy, 2015)亦用到标准高斯近似。

与高斯近似规范化类似, 模式识别领域有一种经典的非线性变换方法, 即JM变换。JM概率距离由Matusita(1955)提出, 一直以来, 其被用于表达两个概率分布之间的概率距离

$$\text{JM} = \left\| F_1 - F_2 \right\| = \sqrt{\int_{\mathbf{R}} (\sqrt{p_1(x)} - \sqrt{p_2(x)})^2 dm} \quad (5)$$

式中, \mathbf{R} 为样本空间; 同分布函数 F_1 、 F_2 的概率密度函数分别为 $p_1(x)$ 、 $p_2(x)$, 即

$$F_1 = \int_{\mathbf{R}} p_1(x) dm \quad (6)$$

$$F_2 = \int_{\mathbf{R}} p_2(x) dm \quad (7)$$

Swain和Davis(1978)将JM概率距离简化为一种对巴氏(Bhattacharyya)距离的非线性变换形式, 用于表征两种目标的类间可分性

$$\text{JM} = \sqrt{2 \times (1 - \exp(-B))} \quad (8)$$

式中, B 为巴氏距离(Richards, 2013)。通过JM变换将之规范化到 $[0, 2)$ 。

在分析了JM距离和分类精度与归一化距离关系的基础上, 童庆禧等(2006)指出了JM距离与分类精度的一致性关系, 即当归一化距离达到一定值时, JM距离和分类精度均达到极值, 再增大归一化距离对JM距离和分类精度的提升是没有意义的。因此, JM非线性变换的实质是将随机变量的作用限制在一定范围内。

2.4 JM变换的自适应降维

基于上文对类间距离和分类精度的关系的分析, 本研究作出合理推断和假设: 波段信息量(或独立性)与波段分类性能之间是非线性的正相关关系, 即随着波段信息量(或独立性)的增大, 波段分类性能快速提升, 当信息量(或独立性)达到一定值后, 分类性能提升变缓至稳定状态。这与ABS和MABS等方法“信息量(或独立性)越大, 分类性能越强”的假设有所不同。

由上文可知, JM变换是一种简单有效的非线性规范化工具, 且通过JM变换优化的类间距离与分类精度具有一致性关系; MABS二元复合降维模型中信息量和独立性缺乏必要的规范化处理。由此, 本研究拟通过JM变换对MABS进行优化改进, 构建高光谱自适应降维方法JM²ABS(算法1):

算法1: JM²ABS降维

输入:

含有*n*个波段的高光谱影像 $H = \{H_1, H_2, H_3, \dots, H_n\}$;

欲选出的波段数量*k*

输出:

选出的*k*个波段的集合

1: 根据式(4)计算波段的信息量 $\text{Infor}(\sigma_i, i = 1, 2, \dots, n)$;

2: 根据式(8)计算波段信息量的规范化 Infor_{JM} ;

3: 根据式(1)计算波段独立性 Indep ;

4: 根据式(8)计算波段独立性的规范化 Indep_{JM} ;

5: 设置比例因子超参数 α 和 β , 计算降维的参考值目标函数:

$$\begin{cases} Z_i = (\text{Infor}_{\text{JM}}^i)^\alpha \times (\text{Indep}_{\text{JM}}^i)^\beta \\ \text{s. t. } \alpha + \beta = 1 \end{cases} \quad (9)$$

6: 对全波段的 Z_i 值进行降序排列, 返回前*k*个值对应的波段。

如上文所述,比例因子可以通过先验知识或监督训练获取,为便于与MABS作比较,本文将比例因子设置为1,默认信息量与独立性同等作用。

3 实验与分析

为了客观地探讨算法的有效性,除了选用改进前的降维方法MABS外,本研究选用了基于拉普拉斯映射和局部保持投影的流形学习方法LSFS(Laplacian Score Feature Selection)(He等,2005)和基于带权有向图的降维方法InfFS(Infinite Feature Selection)(Roffo等,2015;Roffo和Melzi,2017)等两种非监督方法进行对比实验。其中,MABS以波段信息量和独立性为决策目标;LSFS可以有效选出体现高维数据潜在流形结构的特征,是一种高频的主流对比方法;而InfFS则从分类的角度,利用信息量和秩相关系数来揭示波段重要性。相对而言,后两种对比算法时间复杂度较高,以选取更优质波段来提升分类精度。

3.1 实验数据集

(1)样例一:数据为AVIRIS传感器探测的美国印第安纳州西北部的Indian Pines农场影像([http://www.ehu.es/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes\[2018-03-20\]](http://www.ehu.es/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes[2018-03-20])),样本数据大小为145 pixel × 145 pixel,数据位深为14 bit,波长范围为0.4—2.5 μm,在剔除水吸收严重的[104,105,106,107,108],[150,151,152,153,154,155,156,157,158,159,160,161,162,163],220等20个波段后,剩余200个波段,样例中共有16种覆被类型(图1)。

(2)样例二:数据为美国哥伦比亚大学计算机

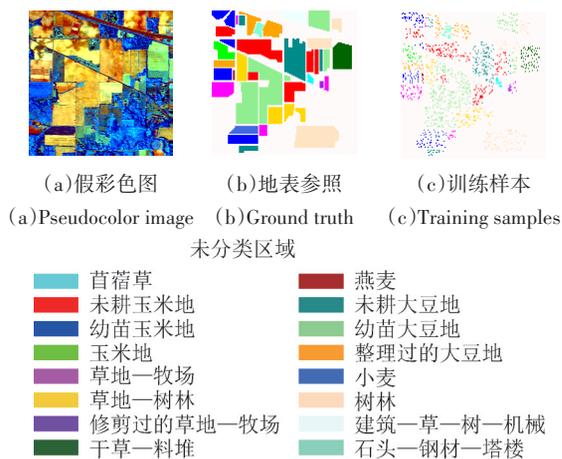


图1 样例一
Fig.1 Sample 1

科学系的计算机视觉实验室CAVE([http://www.cs.columbia.edu/CAVE/databases/multispectral/\[2018-03-20\]](http://www.cs.columbia.edu/CAVE/databases/multispectral/[2018-03-20]))的面阵显微高光谱相机(Apogee Alta U260)采集的Superballs数据集,相机CCD大小为512 pixel × 512 pixel,数据位深为16 bit,波长范围为0.4—0.7 μm,光谱分辨率为10 nm,共31个波段。截取了大小为400 pixel × 400 pixel的范围作为实验样本,Superballs数据集共分类为7种目标(图2)。

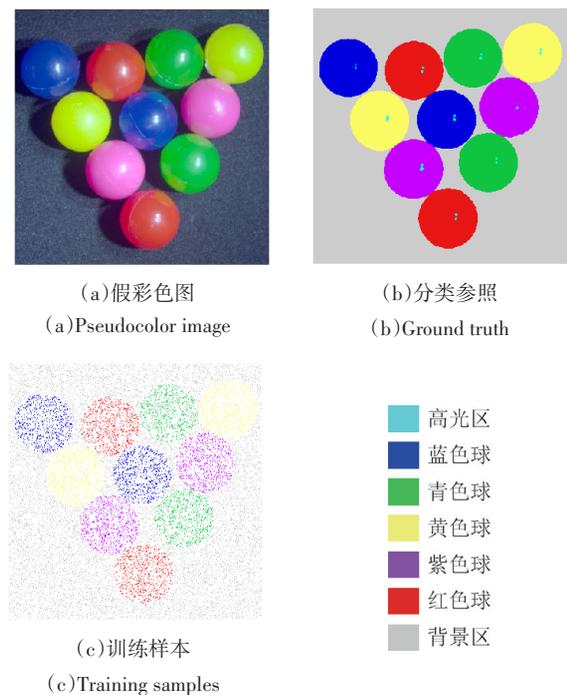


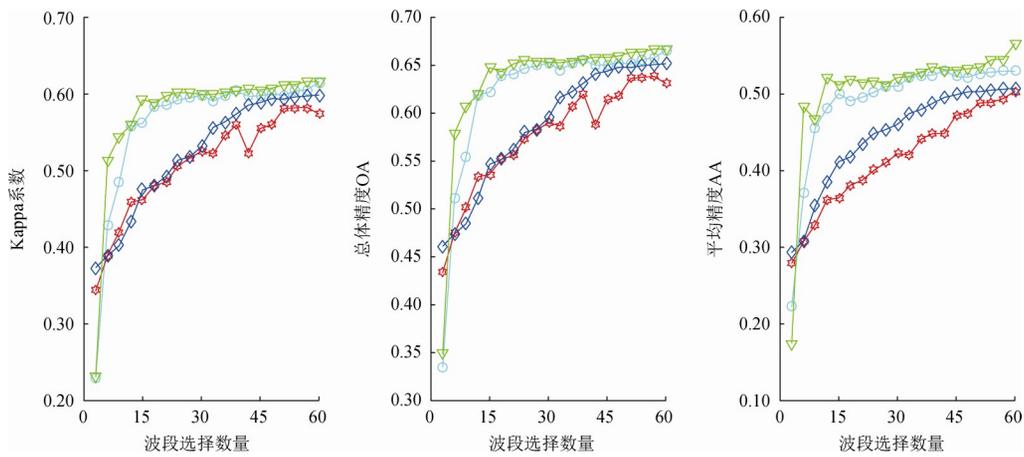
图2 样例二
Fig.2 Sample 2

3.2 降维及监督分类实验

采用两组数据集为实验样例,各选择20组子集(样例一:以3个波段为步长,3—60;样例二:以1个波段为步长,1—20);分别选取 k 最邻近分类器KNNC(k -Nearest Neighbor Classifier, k 设为9(Sui等,2015))和随机森林分类器RFC(Random Forest Classifier,决策树数量为100(Sui等,2015))对每个子集进行了监督分类(随机选择15%的真实数据作为训练样本(样例一如图1(c)所示;样例二如图2(c)所示),剩余85%作为测试样本),并分别统计了Kappa系数、总体分类精度OA(Overall Accuracy)和平均分类精度AA(Average Accuracy),以便对分类性能进行定量评价。

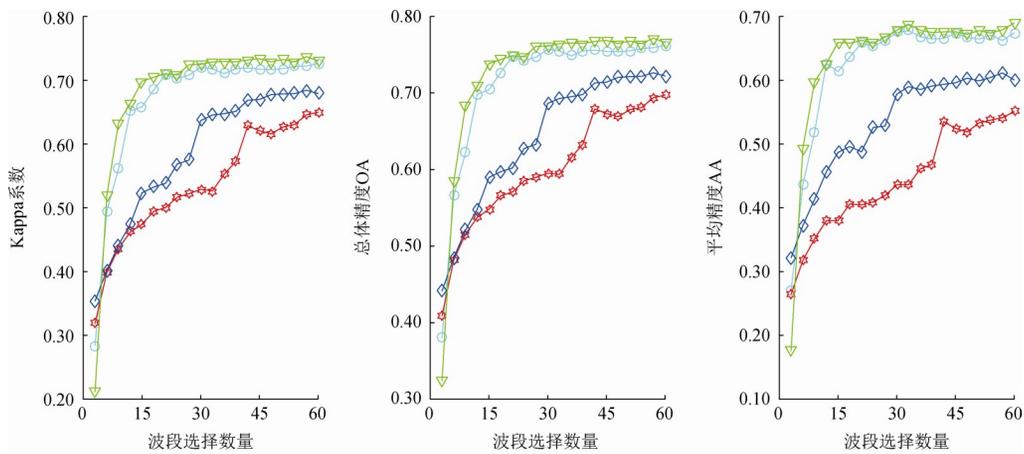
3.2.1 样例一

样例一的监督分类的精度对比如图3所示。



(a) k 最邻近分类器

(a) k -Nearest neighbor classifier



(b) 随机森林分类器

(b) Random forest classifier

—○— MABS —★— LSFS —◇— InfFS —▽— JM²ABS

图3 样例一的分类精度结果

Fig.3 Classification accuracy results of sample 1

样例一在两种分类器下, 均显示出本文方法 JM²ABS 的优势, 在 Kappa 系数、OA 和 AA 上, JM²ABS 均总体高于 MABS、InfFS 和 LSBS 等 3 种对比方法。在 KNNC 下: (1)除第 1 组子集(3 个波段)外, JM²ABS 在其他各组子集中均优于 3 种对比算法, 并且在少量波段子集上便达到了较高精度(其中, Kappa 和 OA 均在第 5 子集上达到较高精度; 而 AA 在第 4 子集上达到高值), 随后趋于稳定且有缓慢爬升; 与次优的 MABS 相比, JM²ABS 个别子集(第 5、11 和 13 子集)上的优势较弱, 而在其他多个子集上均有明显优势。(2)MABS 的变化趋势与 JM²ABS 相似, 其 Kappa 和 OA 均在第 8 子集上达到较高精度, 而 AA 则第 11 子集(33 个波段)达到高值。(3)LSFS 和 InfFS 的性能低于前两者,

但 InfFS 总体优于 LSFS。而随机森林分类器下, 4 种算法的分类性能均有 0.1 左右的大幅提升, 而 4 者之间的对比情况则与 KNNC 情况相似: (1)与 LSFS 和 InfFS 相比, JM²ABS 优势仍较明显, 而与 MABS 相比, 总体来看, JM²ABS 性能略占优。(2)LSFS 和 InfFS 的性能相对较差, 但 InfFS 明显优于 LSFS。

表 1 进一步展示了各方法的多组子集分类性能的平均情况, 可以看出两种分类器下, 较之 3 种对比方法, JM²ABS 均表现出优异的性能: JM²ABS 分别比次优高 0.0142(KNNC Kappa)、0.0127(KNNC OA)、0.0158(KNNC AA)、0.0123(RFC Kappa)、0.0109(RFC OA)、0.0113(RFC AA)。

表1 样例一多组波段子集的分类性能均值

Table 1 The average classification performance of 20 subsets in sample 1

Method	KNNC			RFC		
	Kappa	OA	AA	Kappa	OA	AA
MABS	<u>0.5615</u>	<u>0.6204</u>	<u>0.4901</u>	<u>0.6682</u>	<u>0.7124</u>	<u>0.6218</u>
LSFS	0.5087	0.5755	0.4163	0.5362	0.6007	0.4438
InfFS	0.5230	0.5877	0.4465	0.5865	0.6428	0.5321
JM ² ABS	0.5757	0.6331	0.5058	0.6805	0.7233	0.6331

注: 粗体表示最优, 下划线表示次优。

3.2.2 样例二

样例二的监督分类结果如图4所示。从图4中可以看出, 与其他3种对比方法相比, JM²ABS的优势明显, 总体而言, 在Kappa系数、OA和AA

上, JM²ABS均明显高于MABS、LSFS和LSBS等3种非监督对比方法。从KNN分类结果看: (1)JM²ABS在少量波段子集上达到较高的分类精度(Kappa、OA和AA3种指标均在第2子集上达到0.90及以上); 次优的LSFS方法的3种评估指标(根据上下文, Kappa、OA、AA)首次达到0.90精度是在第5、5、6子集上; 而MABS和LSFS则分别需要6个和10个波段才能达到较高精度。(2)总体上, MABS与LSFS表现出相似的性能, 后者略优于前者; 除了前5个子集外, 两者均优于InfFS。从RFC看: (1)4种方法的精度均随子集波段数量的增加而提高, 总体而言, JM²ABS最优, LSFS和MABS次之, InfFS最弱; (2)除第1子集外, 其他各子集上, JM²ABS均为最优。

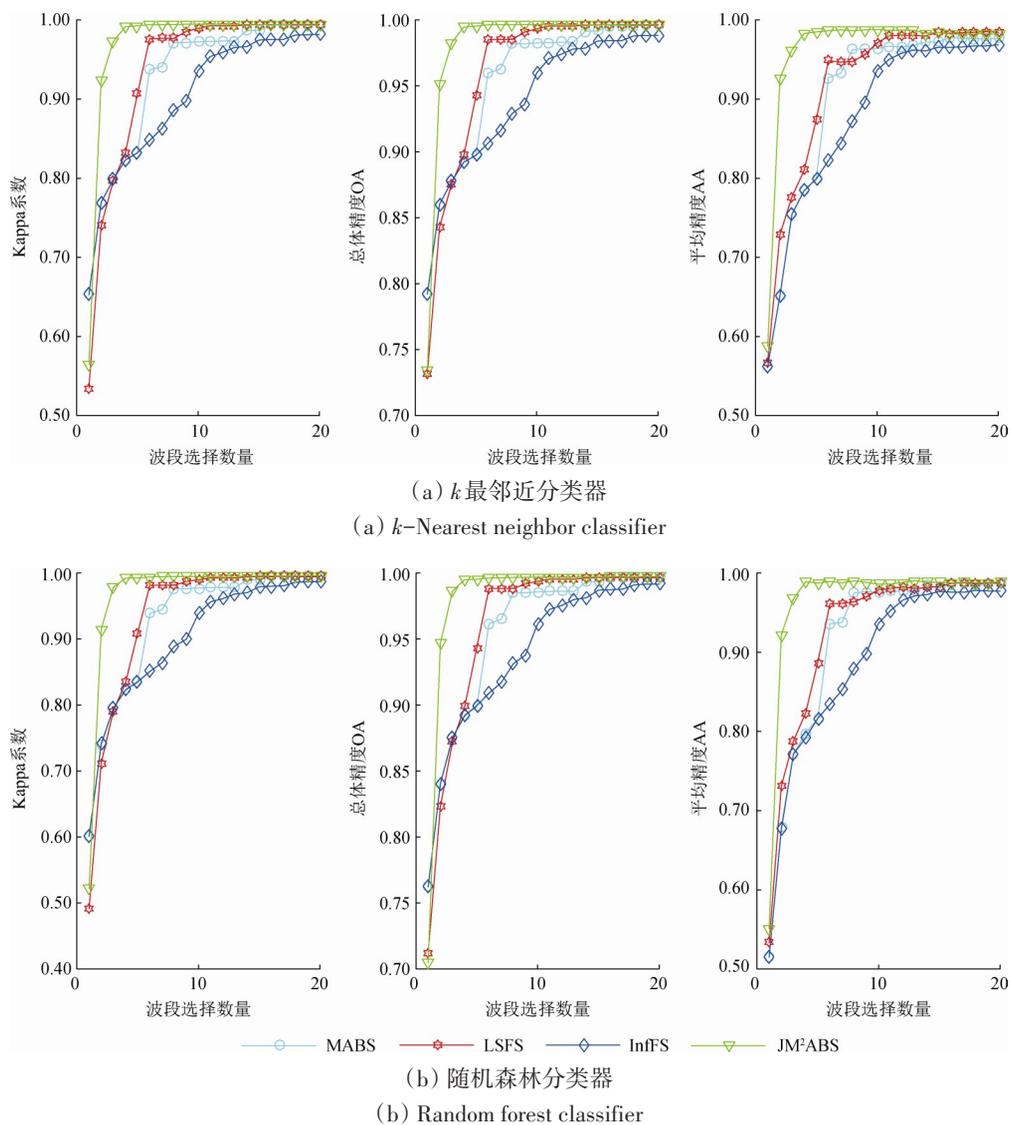


图4 样例二的分类精度结果

Fig. 4 Classification accuracy results of sample 2

从表2的平均分类性能上看,与样例一相比, JM²ABS的精度性能优势更为明显,各指标均值均为最优,在KNNC下,分别比次优高0.0356(Kappa)、0.0208(OA)、0.0434(AA);在RFC下,分别比次优高0.0364(Kappa)、0.0210(OA)、0.0399(AA)。

表2 样例二多组波段子集的分类性能均值

Table 2 The average classification performance of 20 subsets in sample 2

Method	KNNC			RFC		
	Kappa	OA	AA	Kappa	OA	AA
MABS	0.9264	0.9550	0.9015	0.9248	0.9540	0.9105
LSFS	<u>0.9326</u>	<u>0.9596</u>	<u>0.9177</u>	<u>0.9299</u>	<u>0.9581</u>	<u>0.9221</u>
InfFS	0.9010	0.9393	0.8778	0.8996	0.9386	0.8848
JM ² ABS	0.9681	0.9804	0.9611	0.9664	0.9791	0.9620

注:粗体表示最优,下划线表示次优。

3.3 结果分析

两组高光谱图像公共数据集(线阵高光谱和面阵显微高光谱),各20组子集的分类结果表明,本研究提出的JM²ABS算法优于MABS、InfFS和LSFS等3种对比方法,验证了JM²ABS的有效性。

选用经典的简单分类器KNNC和强分类器RFC等两种分类方法进行算法之间降维结果分类性能的对比,验证了JM²ABS算法性能的健壮性;多组子集的实验显示,较之3种对比方法,在选择较少波段时, JM²ABS更易表现出最优的分类性能,且性能相对稳定。

4 结论

自适应波段选择ABS及改进的MABS等降维方法均是基于信息量因子与独立性因子作用等同的假设而进行建模的,但是不同测度理论下的两种因子在量纲和值域范围等方面均有较大的差异,而使得模型并不契合假设; JM变换常结合巴氏距离来表达模式识别中的类间差异程度, JM距离本质上是对巴氏概率距离的JM分布近似规范化。鉴于此,本文试用JM非线性变换对波段信息量及其独立性进行优化,构建了JM²ABS自适应降维模型,以同等地兼顾独立性与信息量。详实的监督分类实验表明:与MABS、InfFS和LSFS等3种方法相比, JM²ABS降维结果的目标分类性能总体上最优,有较好的可行性。为了与MABS作比较,本

文JM²ABS默认了波段信息量与独立性的同等作用,即设置比例因子为1;在下一步的研究中,将着手调整比例因子,探讨波段信息量与独立性的最优关系。

参考文献(References)

- Bajcsy P and Groves P. 2004. Methodology for hyperspectral band selection. *Photogrammetric Engineering and Remote Sensing*, 70(7): 793-802 [DOI: 10.14358/PERS.70.7.793]
- Cao X H, Wu B, Tao D C and Jiao L C. 2016. Automatic band selection using spatial-structure information and classifier-based clustering. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(9): 4352-4360 [DOI: 10.1109/JSTARS.2015.2509461]
- He X F, Cai D and Niyogi P. 2005. Laplacian score for feature selection//*Proceedings of the 18th International Conference on Neural Information Processing Systems*. Vancouver, British Columbia, Canada: MIT Press: 507-514
- Ioffe S and Szegedy C. 2015. Batch normalization: accelerating deep network training by reducing internal covariate shift//*Proceedings of the 32nd International Conference on Machine Learning*. Lille, France: 448-456
- Li S J and Qi H R. 2011. Sparse representation based band selection for hyperspectral images//*Proceedings of the 18th IEEE International Conference on Image Processing*. Brussels, Belgium: IEEE: 2693-2696 [DOI: 10.1109/ICIP.2011.6116223]
- Liu C H, Zhao C H and Zhang L Y. 2005. A new method of hyperspectral remote sensing image dimensional reduction. *Journal of Image and Graphics*, 10(2): 218-222(刘春红, 赵春晖, 张凌雁. 2005. 一种新的高光谱遥感图像降维方法. *中国图象图形学报*, 10(2): 218-222) [DOI: 10.3969/j.issn.1006-8961.2005.02.015]
- Matusita K. 1955. Decision rules, based on the distance, for problems of fit, two samples, and estimation. *The Annals of Mathematical Statistics*, 26(4): 631-640 [DOI: 10.1214/aoms/1177728422]
- Qin F P, Zhang A W, Wang S M, Meng X G, Hu S X and Sun W D. 2015. Hyperspectral band selection based on spectral clustering and inter-class separability factor. *Spectroscopy and Spectral Analysis*, 35(5): 1357-1364(秦方普, 张爱武, 王书民, 孟宪刚, 胡少兴, 孙卫东. 2015. 基于谱聚类与类间可分性因子的高光谱波段选择. *光谱学与光谱分析*, 35(5): 1357-1364) [DOI: 10.3964/j.issn.1000-0593(2015)05-1357-08]
- Richards J A. 2013. *Remote Sensing Digital Image Analysis*. 5th ed. Berlin, Heidelberg: Springer: 350-352
- Roffo G, Melzi S and Cristani M. 2015. Infinite feature selection//*IEEE International Conference on Computer Vision*. Santiago, Chile: IEEE: 4202-4210 [DOI: 10.1109/ICCV.2015.478]
- Roffo G and Melzi S. 2017. Ranking to learn: feature ranking and selection via eigenvector centrality//Appice A, Ceci M, Loglisci C, Masciari E and Raš Z, eds. *New Frontiers in Mining Complex Patterns*. Berlin Heidelberg: Springer: 19-35 [DOI: 10.1007/978-3-

- 319-61461-8_2]
- Shukla U P and Nanda S J. 2018. A Binary Social Spider Optimization algorithm for unsupervised band selection in compressed hyperspectral images. *Expert Systems with Applications*, 97: 336-356 [DOI: 10.1016/j.eswa.2017.12.034]
- Sui C H, Tian Y, Xu Y P and Xie Y. 2015. Unsupervised band selection by integrating the overall accuracy and redundancy. *IEEE Geoscience and Remote Sensing Letters*, 12(1):185-189 [DOI: 10.1109/LGRS.2014.2331674]
- Sun W W, Zhang D F, Yang G and Li W Y. 2018. Band selection for hyperspectral imagery based on weighted probabilistic archetypal analysis. *Journal of Remote Sensing*, 22(1): 110-118(孙伟伟, 张殿发, 杨刚, 李巍岳. 2018. 加权概率原型分析的高光谱影像波段选择. *遥感学报*, 22(1): 110-118) [DOI: 10.11834/jrs.20186446]
- Sun W W, Zhang L P, Du B, Li W Y and Lai Y M. 2015. Band selection using improved sparse subspace clustering for hyperspectral imagery classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(6): 2784-2797 [DOI: 10.1109/JSTARS.2015.2417156]
- Swain P H and Davis S M. 1978. *Remote Sensing: the Quantitative Approach*. New York: McGraw-Hill
- Tong Q X, Zhang B and Zhang L F. 2016. Current progress of hyperspectral remote sensing in China. *Journal of Remote Sensing*, 20(5): 689-707(童庆禧, 张兵, 张立福. 2016. 中国高光谱遥感的前沿进展. *遥感学报*, 20(5): 689-707) [DOI: 10.11834/jrs.20166264]
- Tong Q X, Zhang B and Zheng L F. 2006. *Hyperspectral Remote Sensing: the Principle, Technology and Application*. Beijing: Higher Education Press: 135-136(童庆禧, 张兵, 郑兰芬. 2006. 高光谱遥感: 原理、技术与应用. 北京: 高等教育出版社: 135-136)
- Wang L G and Wei F J. 2013. Band selection for hyperspectral imagery based on combination of genetic algorithm and ant colony algorithm. *Journal of Image and Graphics*, 18(2): 235-242(王立国, 魏芳洁. 2013. 结合遗传算法和蚁群算法的高光谱图像波段选择. *中国图象图形学报*, 18(2): 235-242)
- Yu X C, Feng W F, Yang G P and Chen W. 2013. *Analysis and Application for Hyperspectral Imagery*. Beijing: Science Press: 1-6(余旭初, 冯伍法, 杨国鹏, 陈伟. 2013. 高光谱影像分析与应用. 北京: 科学出版社: 1-6)
- Zhang A W, Du N, Kang X Y and Guo C F. 2017. Hyperspectral adaptive band selection method through nonlinear transform and information adjacency correlation. *Infrared and Laser Engineering*, 46(5): 213-221(张爱武, 杜楠, 康孝岩, 郭超凡. 2017. 非线性变换和信息相邻相关的高光谱自适应波段选择. *红外与激光工程*, 46(5): 213-221) [DOI: 10.3788/IRLA201746.0538001]
- Zhang K, Hei B Q, Zhou Z and Li S Y. 2018. CNN with coefficient of variation-based dimensionality reduction for hyperspectral remote sensing images classification. *Journal of Remote Sensing*, 22(1): 87-96(张康, 黑保琴, 周壮, 李盛阳. 2018. 变异系数降维的CNN高光谱遥感图像分类. *遥感学报*, 22(1): 87-96) [DOI: 10.11834/jrs.20187075]

Hyperspectral images adaptive dimensionality reduction optimized by JM transform

KANG Xiaoyan^{1,2}, ZHANG Aiwu^{1,2}, HU Shaoxing³, XIAO Qing⁴, CHAI Shatuo⁵

1. Key Laboratory of 3D Information Acquisition and Application, Ministry of Education, Capital Normal University, Beijing 100048, China;

2. Engineering Research Center of Spatial Information Technology, Ministry of Education, Capital Normal University, Beijing 100048, China;

3. School of Mechanical Engineering & Automation, Beihang University, Beijing 100191, China;

4. State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China;

5. Academy of Animal and Veterinary Sciences, Qinghai University, Xining, 810016, China

Abstract: Hyperspectral remote sensing images, which collect rich spectral and spatial information of observed targets, usually contain dozens to hundreds of narrow bands with wavelengths ranging from the visible light region to the near-infrared spectra. With such an abundant number of spectral features, hyperspectral images (HSI) allow us to distinguish different of objects or targets by rule and line. Unfortunately, such high-dimensionality data pose a challenge in data transmission, storage, and processing. Specifically, those HSIs with high redundancy information and strong correlation are prone to a Hughes phenomenon during the image classification process. Therefore, dimensionality reduction is necessary for targets classification. Moreover, without using prior label samples, unsupervised dimensionality reduction can effectively simplify the HSI feature space, and prevent the Hughes phenomenon in the targets classification.

In this paper, the Jeffries – Matusita (JM) modified adaptive band selection (JM²ABS) method is proposed to extract proper features from HSI datasets. Generally speaking, a band that contains many information and demonstrates strong independence is a very important feature that helps unsupervised band selection methods to classify targets. The JM²ABS method considers both the information capacity and independence of HSI bands. Given the significant differences in the measurements of a band's information capacity and its independence, we introduce the JM transform function to normalize the distributions of the information capacity and the independence of HSI data. Thus

JM²ABS shows that both the information capacity and the independence are equally important in unsupervised dimensionality reduction.

We also compare our proposed JM²ABS method against three typical methods, namely, the modified adaptive band selection method, the Laplacian score feature selection method, and the infinite feature selection method. By using random training samples, we perform supervised classification experiments on two kinds of HSI public datasets (linear and planar arrays). The results demonstrate that JM²ABS outperforms the other three typical methods in terms of Kappa value, overall classification accuracy, and average classification accuracy. Moreover, under a small number of bands, JM²ABS can reach a high and stable level regardless of the different datasets and different classifiers used.

The proposed JM²ABS can select the proper features of HSI datasets for their classification application. The JM transform function is a kind of nonlinear distribution function that can standardize variables from different scales to the same. To demonstrate the feasibility of JM transform optimization, we set the same weight for the information content and the independence. In our future work, we will explore the similarities and differences between the information capacity and the independence in dimensionality reduction.

Key words: remote sensing, Jeffries-Matusita transform, normalization, adaptive dimensionality reduction, unsupervised band selection, hyperspectral image

Supported by National Natural Science Foundation of China (No. 41571369); National Key Research and Development Program of China (No. 2016YFB0502500); Beijing Municipal Natural Science Foundation (No. 4162034); Science and Technology R&D Program of Qinghai Province (No. 2016-NK-138).