

# 基于 GF-6 卫星影像多特征优选的酿酒葡萄精准识别

李文杰<sup>1,2</sup>, 郭晓雷<sup>3</sup>, 杨玲波<sup>1,2</sup>, 闫 鸣<sup>3</sup>, 邹晨曦<sup>4</sup>, 方亚华<sup>4</sup>,  
孙 涵<sup>5</sup>, 黄敬峰<sup>1,2\*</sup>

(1. 浙江大学农业遥感与信息技术应用研究所, 杭州 310058; 2. 浙江省农业遥感与信息技术重点研究实验室, 杭州 310058;  
3. 银川市气象局, 银川 750000; 4. 北京思湃德信息技术有限公司, 北京 100086; 5. 内蒙古新天元防灾减灾研究院,  
呼和浩特 010051)

**摘 要:** 多源遥感信息和特征优选是提高农作物识别精度的重要支撑, 高分六号 (GF-6) 卫星作为首次引入红边波段的国产卫星, 其丰富的光谱信息为作物识别提供了新的思路 and 解决途径。该研究基于宁夏回族自治区银川市永宁县 2018 年 6 月—2019 年 3 月的 GF-6 数据, 充分利用红边优势提取光谱特征、纹理特征和植被指数特征, 构建多种特征组合方案, 并根据随机森林算法对特征重要性进行度量, 选取最优特征组合对酿酒葡萄进行精准识别。结果表明, 与单一特征相比, 多源遥感特征的增加显著改善了酿酒葡萄分类效果, 其中, 植被指数贡献程度最大, 光谱特征次之; 基于随机森林的优选特征组合分类效果最佳, 其中, 总体分类精度为 94.15%, 酿酒葡萄用户精度为 94.23%, 制图精度为 92.59%; 以实地调查的 4 个酒庄为验证区, 将酿酒葡萄提取结果与统计数据进行对比, 面积相对精度均在 70% 以上, 其中优选特征结果相对精度在 90% 以上, 研究结果将为国产卫星红边波段在植被分类和识别方面的应用提供数据参考。

**关键词:** 遥感; 图像处理; 高分六号; 酿酒葡萄; 特征优选; 随机森林; 红边波段; 精准识别

doi: 10.11975/j.issn.1002-6819.2020.18.020

中图分类号: TP79; S127

文献标志码: A

文章编号: 1002-6819(2020)-18-0165-09

李文杰, 郭晓雷, 杨玲波, 等. 基于 GF-6 卫星影像多特征优选的酿酒葡萄精准识别[J]. 农业工程学报, 2020, 36(18): 165-173. doi: 10.11975/j.issn.1002-6819.2020.18.020 http://www.tcsae.org

Li Wenjie, Guo Xiaolei, Yang Lingbo, et al. Accurate recognition of wine grapes using multi-feature optimization based on GF-6 satellite images[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2020, 36(18): 165-173. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.2020.18.020 http://www.tcsae.org

## 0 引 言

农作物信息的快速获取和准确识别是开展精准农业的基础, 从依赖国外遥感数据到自主研发国产卫星, 中国农业遥感技术从深度和广度上都取得了显著进展<sup>[1]</sup>。随着高分六号 (GF-6) 卫星正式投入使用, 中国首次增加了有效反映作物特有光谱特性的“红边”波段, 实现了时间分辨率和光谱分辨率的优化组合, 打破了国外卫星对红边波段遥感影像的垄断。与中国高分六号卫星相比, 欧美卫星对红边波段的应用起步较早, 如 2008 年, 德国发射了全球第一颗具有“红边”波段的商业卫星——RapidEye; 2009 年, 美国成功发射 Worldview-2 卫星, 除 4 个业内标准谱段外, 该卫星还提供了包括红边波段在内的 4 个额外谱段; 2015 年, 欧航局成功发射 Sentinel-2 卫星, Sentinel-2 是唯一一个在红边范围内含有 3 个波段的光学卫星。

增加敏感波段和采用多源特征是农作物精准识别的 2 个重要手段<sup>[2]</sup>。红边波段 (Red-edge) 介于红光波段和红外波段之间, 对作物冠层的微小变化敏感, 是作物识别<sup>[3-4]</sup>和指示绿色植物生长状况<sup>[5-7]</sup>的特征波段之一。如刘佳等<sup>[8]</sup>采用监督分类方法对玉米、大豆等农作物进行分类, 计算了 RapidEye 影像有无红边参与条件下的分类精度及景观破碎度等指标, 研究表明, 红边波段的引入后, 总体精度提高了 6.7%, 同时降低了景观破碎度; Adelabu 等<sup>[9]</sup>采用支持向量机和随机森林 2 种算法对 RapidEye 影像做分类研究, 有红边参与条件下, 总体精度分别提高了 19% 和 21%; Tigges 等<sup>[10]</sup>利用不同波段组合方式的多时相 RapidEye 影像对柏林城市森林进行了品种识别, 加入红边波段后具有更高的样本分离度和 Kappa 系数; Sertel 等<sup>[11]</sup>建立了一套规则集, 采用多级分割的面向对象方法从 Worldview-2 影像中提取地物类型信息, 分类精度达 94.87%。以上研究表明, 红边波段在农业遥感应用中已取得很大进展, 但数据来源主要是国外卫星数据, 中国卫星红边波段的应用研究尚处于起步阶段, 且研究对象多为小麦、水稻等大宗粮食作物, 对经济作物的研究相对较少。

丰富的多源遥感特征可以弥补单一类型特征关键信息不足、分类结果破碎度高等缺陷, 改善分类效果, 提升植被分析的精度。然而, 特征过量易出现信息冗余、

收稿日期: 2020-06-02 修订日期: 2020-08-27

基金项目: 国家科技重大专项: 高分辨率对地观测系统重大专项 (09-Y20A05-9001-17/18); 银川市科技计划项目 (2018-ZY-18017)

作者简介: 李文杰, 主要从事卫星图像处理与农业遥感方面的研究。

Email: liwenjie@zju.edu.cn

\*通信作者: 黄敬峰, 博士, 教授, 主要从事农业遥感与信息技术方面的研究。Email: hjf@zju.edu.cn

处理效率低等问题, 严重时会发生“维数灾难”现象<sup>[12]</sup>, 不仅增加了处理成本, 还会降低分类精度, 因此, 合理选择有效特征对农作物识别具有重要意义<sup>[13-14]</sup>。程希萌等<sup>[15]</sup>采用最小冗余最大相关 (minimal Redundancy Maximal Relevance, mRMR) 算法对特征进行优选, 证明了 mRMR 特征优选算法在分类过程中的有效性; Atkinson 等<sup>[16]</sup>结合支持向量机与特征递归消除法在 272 个特征集中选取 17 个最优特征, 对南非东部金缕草的分布进行了有效监测。除此之外, 不少学者采用随机森林算法对特征进行优选, 如张磊等<sup>[17]</sup>利用光谱特征、水体指数等信息, 采用随机森林算法对 Sentinel 影像进行了多特征优选, 实现了对黄河三角洲实地信息的提取, 总体精度达 90.93%; 蒙莉娜等<sup>[18]</sup>通过随机森林模型获得 16 个环境辅助变量的相对重要性, 研究了绿洲土壤盐分的空间分布与环境变量间的关系。目前, 基于遥感信息开展的识别研究中, 不足之处在于特征较单一, 如光谱特征仅利用可见光波段及近红外波段, 植被指数特征多使用归一化差值植被指数 (Normalized Difference Vegetation Index, NDVI) 和增强型植被指数 (Enhanced Vegetation Index, EVI) 等, 很少使用红边指数特征。

针对以上问题, 本文选取酿酒葡萄作为研究对象, 基于高分六号卫星影像对研究区酿酒葡萄进行精准识别, 通过多种特征组合比较分类结果差异, 全面客观地评价特征优选对酿酒葡萄精准识别的影响, 为国产卫星红边波段在农作物识别上的性能提升以及农业遥感领域的后续应用提供参考。

## 1 研究区与数据

### 1.1 研究区概况

永宁县隶属于宁夏回族自治区银川市, 地处 105°49′~106°22′E、38°7′~38°29′N 之间, 地势西高东低, 面积为 934 km<sup>2</sup>。该区属中温带干旱气候区, 年平均温度 8.7℃, 无霜期平均 167 d, 年均降水量 201.4 mm, 全年平均日照 2 866 h, 光能资源丰富。永宁县地处宁夏贺兰山东麓酿酒葡萄核心种植区, 日温差大, 有利于有机物质的合成和积累, 适宜优质酿酒葡萄生长, 是世界公认的葡萄种植“黄金”地带<sup>[19]</sup>, 葡萄酒产业是永宁县的支柱产业之一。除酿酒葡萄外, 研究区主要作物还有小麦、玉米等。

### 1.2 数据及其预处理

#### 1.2.1 GF-6 数据

高分六号低轨光学遥感卫星是中国首颗精准农业观测的高分卫星, 搭载了观测幅宽为 800 km 的宽视场 (Wide Field of View, WFOV) 传感器和观测幅宽为 90 km 的全色多光谱传感器 (Panchromatic Multispectral Sensor, PMS), 具有高分辨率、宽覆盖、高质量成像等特点<sup>[20]</sup>。其中, WFOV 传感器空间分辨率为 16 m, 其多光谱数据共包含 8 个波段, 本文选取 GF-6/WFOV 的蓝光 (0.45~0.52 μm)、绿光 (0.52~0.59 μm)、红光 (0.63~0.69 μm)、近红外 (0.77~0.89 μm)、红边波段 1 (0.69~0.73 μm) 和红

边波段 2 (0.73~0.77 μm) 作为试验波段, 分别标记为 1~6 波段。

本文所用到的经过系统几何和辐射校正的 GF-6 影像 1 级产品数据均由中国资源卫星应用中心提供 (http://www.cresda.com)。在此基础上, 对所有数据进行辐射定标、大气校正和几何校正, 并以同期 10 m 分辨率的 Sentinel-2 数据作为参考影像进行几何精校正。预处理中使用的数字高程模型 (Digital Elevation Model, DEM) 采用分辨率为 90 m 的 SRTM DEM 数据。

综合考虑研究区内作物的发育时期以及样本类型, 本文选取了 2018 年 6 月—2019 年 3 月之间的 10 景影像进行识别, 成像时间分别为: 2018 年 6 月 13 日、2018 年 6 月 22 日、2018 年 7 月 12 日、2018 年 7 月 16 日、2018 年 9 月 8 日、2018 年 10 月 2 日、2018 年 11 月 12 日、2019 年 2 月 10 日、2019 年 3 月 3 日、2019 年 3 月 23 日。该时期包含了开花坐果期、浆果生长期以及成熟采收期等酿酒葡萄的主要物候期, 有利于酿酒葡萄的识别和提取。经实地调查, 研究区内葡萄种植以酿酒葡萄为主, 鲜食葡萄对本研究的影响可忽略不计。

#### 1.2.2 样本数据

本文通过实地调查研究和 Google Earth 高分辨率影像目视解译, 共获取 20 个酒庄样点数据和 865 个分类样本数据: 酿酒葡萄 (172 个)、耕地 (158 个)、林地 (165 个)、草地 (79 个)、裸地 (56 个)、建筑 (134 个)、水体 (101 个), 总面积达 34.02 km<sup>2</sup>。其中, 本文将玉米、小麦等作物统分为耕地类型; 将果树、行道树以及山林等统分为林地; 将房屋、道路、工业园区等统分为建筑; 将池塘、河流等统分为水体。样本均匀分布在整个研究区内, 其中, 60% 用作训练样本, 40% 用作验证样本, 样本分布情况如图 1 所示。

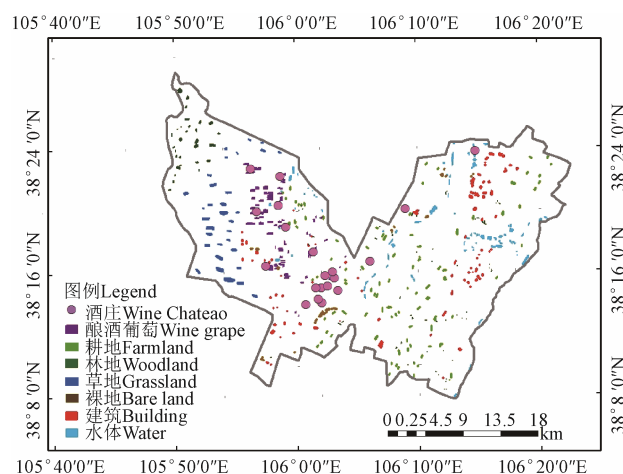


图 1 研究区样本分布

Fig.1 Distribution of samples in study area

## 2 研究方法

### 2.1 技术路线

本研究酿酒葡萄精准识别过程包括数据预处理、样本选择、特征提取和优选、分类及精度验证等内容 (图 2)。

本研究充分利用 GF-6 卫星红边优势提取光谱特征、植被指数特征和纹理特征，使用基于随机森林基尼指数的特征优选算法探索酿酒葡萄识别中的最优特征组合并进行分类。

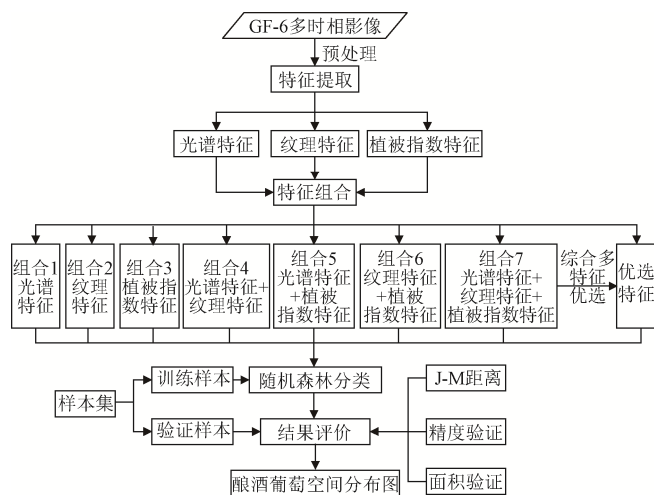


图 2 本研究技术路线  
Fig.2 Flow chart of this study

## 2.2 特征提取

### 2.2.1 光谱特征

经调查研究，9—10 月份为酿酒葡萄成熟期，植株茂盛，光谱特征突出，与其他地物的物候差异较大，是识别酿酒葡萄的最佳时间段，因此选取 2018 年 9 月 8 日影像作为最佳时相数据，取其 1~6 波段作为光谱特征。图 3 为 2018 年 9 月 8 日 GF-6 影像的光谱反射率平均值曲线，蓝光波段、绿光波段和红光波段的酿酒葡萄-林地两类地物光谱反射率接近，识别难度较大；近红外波段光谱反射率平均值明显高于其他波段，但地物混淆度高，“异物同谱”现象严重；红边波段 1 是较易区分酿酒葡萄的波段，裸地和水体等非植被类型与其他地物分层明显，酿酒葡萄-耕地-林地-草地 4 类植被类型也较易区分；在红边波段 2 中，酿酒葡萄和草地的反射率曲线趋于重叠，但其他地物的光谱可分度明显增大，如酿酒葡萄-耕地、酿酒葡萄-林地和裸地-草地等地物，酿酒葡萄-草地可结合红边波段 1 进行区分。

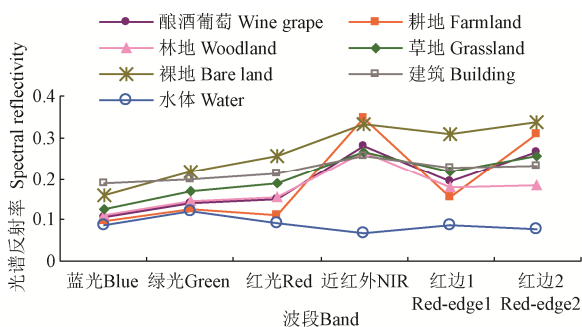


图 3 2018-09-08 典型地物 GF-6 平均光谱反射率曲线  
Fig.3 Mean spectral reflectivity of typical ground object based on GF-6 images on 2018-09-08

### 2.2.2 植被指数特征

归一化差值红边指数 (Normalized Difference

Red-Edge Index, NDRE) 可充分发挥红边波段在植被识别方面的优越性，相比于其他宽带绿度指数，NDRE 具有更高的灵敏度，广泛应用于精细农业、森林监测、植被胁迫性探测等；归一化差值植被指数 (Normalized Difference Vegetation Index, NDVI) 能够反映植被在可见光、近红外波段反射与土壤背景之间的差异，适用于植被早、中期生长阶段的动态监测。季节性是植被最显著的特征之一<sup>[21]</sup>，NDVI 随时间变化与植被的物候信息呈现一定的规律性，可监测植被生长状态和植被覆盖度<sup>[22]</sup>，是区分地物的重要特征；相比于 NDVI，增强型植被指数 (Enhanced Vegetation Index, EVI) 对植被密集区的冠层结构变化更加敏感<sup>[23]</sup>，在研究植被遥感时可以与 NDVI 相互补充。本文所用植被指数介绍如表 1 所示。

表 1 植被指数  
Table 1 Vegetation index

植被指数名称 Name of vegetation index	公式 Formula	参考文献 Reference
归一化差值红边指数 Normalized Difference Red-Edge Index (NDRE)	$NDRE = (NIR - RE) / (NIR + RE)$	Gitelson et al <sup>[25]</sup>
归一化差值植被指数 Normalized Difference Vegetation Index (NDVI)	$NDVI = (NIR - RED) / (NIR + RED)$	Tucker et al <sup>[26]</sup>
增强型植被指数 Enhanced Vegetation Index (EVI)	$EVI = 2.5 \times (NIR - RED) / (NIR + 6 \times RED - 7.5 \times BLUE + 1)$	Huete et al <sup>[27]</sup>

本研究应用多时相遥感数据，利用 GF-6 数据分别合成 NDRE、NDVI 和 EVI 时间序列，在植被指数时间序列构建过程中，时相选择以酿酒葡萄物候期为主要依据，同时结合耕地（以小麦、玉米为主）、林地等地物类别的物候特点，10 景影像满足全覆盖、无云、可见度高等条件<sup>[24]</sup>。

### 2.2.3 纹理特征

纹理特征借助像素点的空间结构相似性对像元表面粗糙度、均匀性等情况进行度量，是图像分类的重要特征之一<sup>[28]</sup>。丰富的纹理信息比单一光谱信息具有更好的分类效果，对解决错分问题提高分类精度具有重要意义，因此常将纹理特征作为光谱特征和植被指数特征的重要补充来参与图像分类。

本文选用普遍认可的灰度共生矩阵 (Grey-Level Co-occurrence Matrix, GLCM) 方法对纹理信息进行提取，首先对 2018 年 9 月 8 日 GF-6 影像 2 个红边波段进行主成分分析，提取第一和第二主成分分量，分别计算 8 个纹理特征，即同质性 (Homogeneity)、方差 (Variance)、对比度 (Contrast)、信息熵 (Entropy)、均值 (Mean)、二阶距 (Second moment)、相关性 (Correlation) 和相异性 (Dissimilarity)，取主成分分量特征的平均值作为最终纹理特征值。

## 2.3 特征组合

基于光谱波段、植被指数和纹理特征构建特征集，选取植被指数序列的 40 景影像作为植被指数特征，其中由红边波段 1 计算得到 NDRE 指数标记为 1NDRE，由红边波段 2 计算得到 NDRE 指数标记为 2NDRE；选取 2018



年 9 月 8 日 GF-6 影像 1~6 波段作为光谱特征; 选取由红边波段计算得到的 8 个纹理信息作为纹理特征, 用于分类的特征集信息如表 2 所示。

表 2 多源特征集描述  
Table 2 Description of multi-source feature corpus

特征类型 Feature type	序号 Order number	日期 Time	数目 Number	特征变量简称 Abbreviation of feature variable
植被指数特征 Vegetation index feature	1~10	2018-06-13 至 2019-03-23	40	1NDRE1~1NDRE10、2NDRE1~2NDRE10、NDVI1~NDVI 10、EVI1~EVI10
光谱特征 Spectral feature	5	2018-09-08	6	B_Blue、B_Green、B_Red、B_Nir、1B_Rededge、2B_Rededge
纹理特征 Texture feature	5	2018-09-08	8	Mean、Dissimilarity、Second、Correlation、Entropy、Homogeneity、Variance、Contrast

注: 序号“1~10”依次为 2018-06-13 至 2019-03-23 的 10 个时相; 序号“5”为最佳时相 2018-09-08。

Note: Order number “1~10” is from 2018-06-13 to 2019-03-23; Order number “5” is the optimal data 2018-09-08.

## 2.4 特征优选

多源综合特征充分利用了影像的光谱和时空信息, 极大地改善了地物区分度低以及易混淆现象, 但特征数量并非越多越好, 特征过量易出现处理效率低、信息冗余等问题, 甚至使得分类精度下降。因此, 如何降低特征维数以及正确选择特征变量对农作物精准识别具有极其重要的意义。本试验采用基于基尼 (Gini) 指数的随机森林算法评估特征重要性并对特征进行优选, 并根据投票机制确定样本的类别<sup>[29]</sup>。基尼指数通过评估特征变量在随机森林每棵树上的分类贡献, 将所有特征重要性得分进行归一化处理比较特征重要性权重。假设有特征变量  $X_1, X_2, \dots, X_f$ , 基于基尼指数的特征重要性评估模型为

$$\text{Gini}_m = \sum_{k=1}^K p_{mk} (1 - p_{mk}) = 1 - \sum_{k=1}^K p_{mk}^2 \quad (1)$$

式中  $\text{Gini}_m$  表示树节点  $m$  的基尼指数,  $k$  表示  $k$  个类别,  $p_{mk}$  表示树节点  $m$  样本属于第  $k$  类的概率。

特征  $X_j$  在树节点  $m$  的重要性  $\text{VIM}_{jm}^{\text{Gini}}$ , 即分支前后的基尼指数变化量为

$$\text{VIM}_{jm}^{\text{Gini}} = \text{Gini}_m - \text{Gini}_l - \text{Gini}_r \quad (2)$$

式中  $\text{Gini}_l$  和  $\text{Gini}_r$  分别表示新节点  $l$  和  $r$  的基尼指数, 两者差值越大说明当前特征重要性越高。

若特征  $X_j$  在第  $i$  棵树中出现  $M$  次, 则特征  $X_j$  在第  $i$  棵树中的重要性得分  $\text{VIM}_{ij}^{\text{Gini}}$  为

$$\text{VIM}_{ij}^{\text{Gini}} = \sum_{m=1}^M \text{VIM}_{jm}^{\text{Gini}} \quad (3)$$

假设随机森林中有  $n$  棵树, 则特征  $X_j$  在随机森林中的重要性评分为

$$\text{VIM}_j^{\text{Gini}} = \sum_{i=1}^n \text{VIM}_{ij}^{\text{Gini}} \quad (4)$$

最后, 将所有树上的特征得分进行归一化处理得到

最终特征重要性得分。

各特征重要性排序图如图 4 所示, 特征数目与总体精度和 Kappa 系数的关系如图 5 所示, 特征描述方式如表 2 所示。

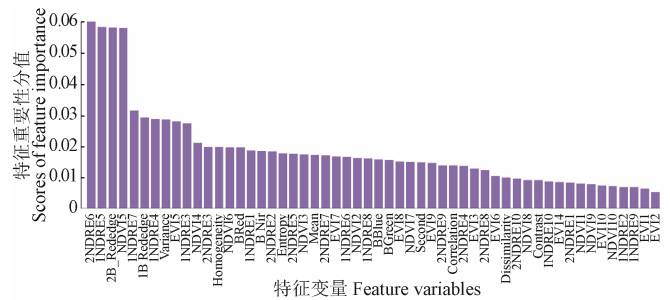


图 4 特征重要性分析

Fig.4 Analysis of the features importance

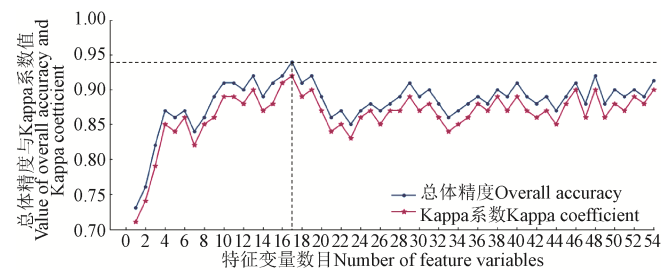


图 5 特征变量数目与分类精度关系

Fig.5 Relationship between number of feature variables and classification accuracy

由图 4 可知, 54 个特征变量权重差异较大, 其中, 2NDRE6 的重要性得分最高, EVI2 的重要性得分最低, 2 个红边波段特征重要性排名分别为第 3 和第 6。由图 5 可知, 当特征数目为 1 时, 总体精度值和 Kappa 系数值最低, 均低于 0.75, 随后总体精度值和 Kappa 系数值随特征数目增多而变大; 特征数目从 4 开始, 总体精度值和 Kappa 系数不再与特征数目呈正相关变化, 参数值呈波动状态, 但 2 个参数总体变化趋势相近, 且数值基本保持在 0.80 以上; 当特征数目取 17 时, 总体精度值和 Kappa 系数值达到最高, 因此将重要性得分排名前 17 的特征作为优选特征参与分类。前 17 个优选特征中, 包含 4 个光谱特征、11 个植被指数特征 (7 个 NDRE 指数、3 个 NDVI 指数、1 个 EVI 指数) 和 2 个纹理特征, 特征贡献度大小依次为植被指数特征、光谱特征、纹理特征, 重要时相主要分布在 7、9、10 和 11 月。

## 2.5 随机森林分类器

随机森林 (Random Forest, RF) 算法是一种采用多棵决策树对样本进行训练并预测的机器学习算法, 最早由 Breiman<sup>[30]</sup>提出。该方法对训练样本和待选特征变量进行双重随机采样, 可有效解决过拟合现象, 并通过集成学习的思想构建多棵决策树, 每颗决策树在不做任何修剪的前提下最大限度地生长, 通过内部评估在随机森林生成过程中对误差进行无偏估计, 最终预测结果和分类结果由多棵树预测值的均值和分类器投票决定。

随机森林法对于遥感影像分类具有很好的抗噪声性能<sup>[31]</sup>，在当前分类算法中，具有较高的分类精度；处理具有高维特征的样本集时，不需要做降维处理，便捷又高效；同时，随机森林算法可以对大数据集进行有效运行，并能较好地评估各特征变量在分类过程中的重要性，在数据挖掘和视觉学习领域具有广泛的应用前景<sup>[32]</sup>。

3 结果与分析

3.1 Jeffries-Matusita 距离

Jeffries-Matusita (J-M) 距离是基于条件概率论的地表光谱可分性判定指标，作为研究不同波段组合对地物特征识别能力的重要参数，其实质是计算不同地物间的概率密度之差<sup>[33]</sup>，公式如下

$$J_{i,j} = \int_X \left\{ \left( p(X|\omega_i) \right)^{1/2} - \left( p(X|\omega_j) \right)^{1/2} \right\}^2 dX \quad (5)$$

式中 $\left( p(X|\omega_i) \right)^{1/2}$ 表示第*i*个像元属于第 $\omega_i$ 个类别的条件概率密度； $\left( p(X|\omega_j) \right)^{1/2}$ 表示第*j*个像元属于第 $\omega_j$ 个类别的条件概率密度。

为量化分析 7 种地物的可分性，本文利用 Jeffries-Matusita 算法计算了 7 种地物类型的统计距离，表 3 列出酿酒葡萄与另外 6 种地物的 Jeffries-Matusita 距离。由表 3 可知，酿酒葡萄与裸地、建筑和水体等非植被类型地物的可分性明显高于 3 种植被类地物；包含 2 种及 2 种以上特征组合的地物可分性高于单一特征的可分性，比如，与单一光谱特征组合相比，优选特征组合中酿酒葡萄-耕地的可分性由 1.57 提高到 1.99，酿酒葡萄-林地的可分性由 1.40 提高到 1.97，这表明多源信息更有助于混合像元中酿酒葡萄样本的准确识别，但包含优选特征组合在内的 5 种多源特征组合之间可分性测度差异不大。

3.2 精度验证

通过实地调查研究和 Google Earth 高分辨率影像目视解译选取验证样本并采用混淆矩阵方法对分类结果进行精度评价，分类精度如表 4 所示。单一特征组合中，光谱特征和纹理特征各精度指标在所有组合中均较低；多时相植被指数特征组合在总体精度、酿酒葡萄的制图精度和用户精度方面得到较大提升；多源特征中，优选特征总体精度最高，达到 94.15%，酿酒葡萄用户精度为 94.23%，制图精度为 92.59%；与分类效果最差的纹理特征组合相比，优选特征总体精度提高了 19.62%，Kappa 系数值增加了 0.24，酿酒葡萄用户精度提高了 16.40%，制图精度提高了 28.65%；与包含光谱、纹理和植被指数在内的交叉特征组合相比，优选特征总体精度提高了 2.85%，酿酒葡萄用户精度和制图精度分别提高了 2.77% 和 2.24%，Kappa 系数值增加了 0.03。结合结果数据分析，得到以下几个结论：优选特征一定程度上减少了冗余信息的干扰，对有效信息的利用率增加，有利于提高分类精度；多源遥感特征分类效果优于单一特征，多时相数

据比单时相数据更适用于分类。

表 3 不同地物 Jeffries-Matusita 距离结果  
Table 3 Jeffries-Matusita distance result of different ground objects

特征组合 Feature combination	酿酒葡萄-耕地 Wine grape-Farmland	酿酒葡萄-林地 Wine grape-Woodland	酿酒葡萄-草地 Wine grape-Grassland	酿酒葡萄-裸地 Wine grape-Bareland	酿酒葡萄-建筑 Wine grape-Building	酿酒葡萄-水体 Wine grape-Water
光谱特征 Spectral feature	1.57	1.40	1.17	1.59	1.88	1.96
纹理特征 Texture feature	1.71	1.71	1.24	1.56	1.98	1.58
植被指数特征 Vegetable index feature	1.95	1.98	1.93	1.95	1.93	1.99
光谱+纹理 Spectrum + texture	1.93	1.91	1.84	1.84	1.99	1.99
光谱+植被指数 Spectrum + vegetable index	1.98	1.98	1.96	1.98	1.91	1.99
纹理+植被指数 Texture + vegetable index	1.97	1.98	1.96	1.95	1.91	1.99
光谱+纹理+植被指数 Spectrum + texture + vegetable index	1.97	1.99	1.98	1.97	1.98	2.00
优选特征 Optimal feature	1.99	1.97	1.97	1.99	1.99	1.99

表 4 分类精度统计  
Table 4 Classification accuracy statistics

特征组合 Feature combination	酿酒葡萄 Wine grape		OA /%	Kappa
	UA/%	PA/%		
光谱特征 Spectral feature	80.47	76.73	84.54	0.81
纹理特征 Texture feature	77.83	63.94	74.53	0.69
植被指数特征 Vegetable index feature	89.55	81.00	87.27	0.85
光谱+纹理 Spectrum + texture	81.05	79.64	85.82	0.83
光谱+植被指数 Spectrum + vegetable index	89.41	85.93	89.68	0.88
纹理+植被指数 Texture + vegetable index	88.78	85.76	90.80	0.89
光谱+纹理+植被指数 Spectrum + texture + vegetable index	91.46	90.35	91.30	0.90
优选特征 Optimal feature	94.23	92.59	94.15	0.93

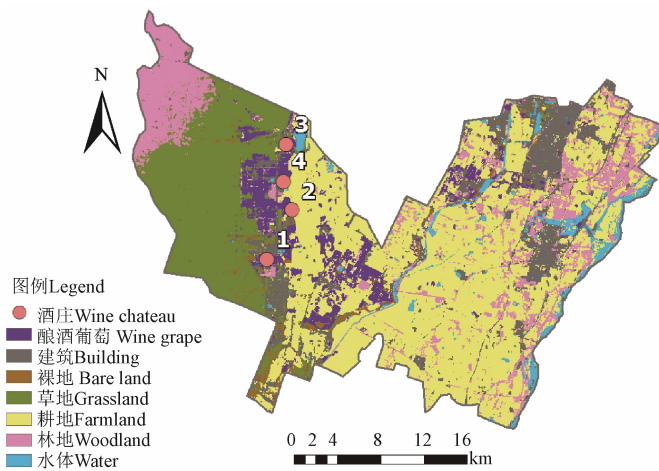
注：UA 为用户精度；PA 为制图精度，即生产者精度；OA 为总体精度；Kappa 为 Kappa 系数。  
Note: UA is user's accuracy, PA is mapping accuracy, that is, producer's accuracy, OA is overall accuracy, Kappa is Kappa coefficient.

### 3.3 面积验证

本文选取实地调查的 4 个葡萄酒庄为验证区,酒庄位置及优选特征分类结果如图 6 所示。

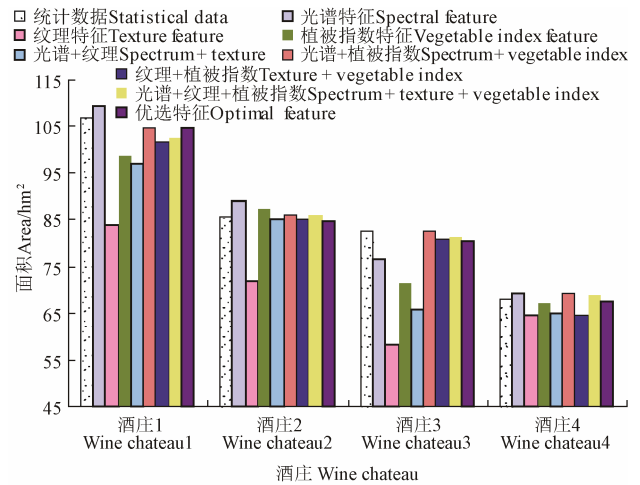
对各组合方案识别出的酿酒葡萄进行面积统计,将

统计结果与统计局(统计年鉴)提供的酿酒葡萄种植面积统计数据进行比较,计算其相对精度。结果显示各酒庄面积精度均在 70%以上,其中,优选特征面积精度在 90%以上。



a. 优选特征分类结果及 4 个酒庄位置示意

a. Classification result of optimal features and position of four wine chateaus



b. 不同特征组合酒庄面积提取结果对比

b. Comparison of wine chateaus area extraction results in different feature combinations

图 6 不同特征组合面积验证图

Fig.6 Diagram of area verification result in different feature combinations

本研究得到的永宁县酿酒葡萄分布图显示,永宁县的酿酒葡萄主要分布在西部地区,贺兰山东麓分布最为密集;各特征组合中酿酒葡萄分布的细节信息差异较大,光谱、纹理和植被指数等单一特征组合中错分现象严重;

相比于单一特征组合,多源特征分类得到的景观破碎度更低,地物间误判面积小,整体识别准确度较高。从目视解译和数据对比角度观察,优选特征以及其他多源特征组合的酿酒葡萄识别效果更好。

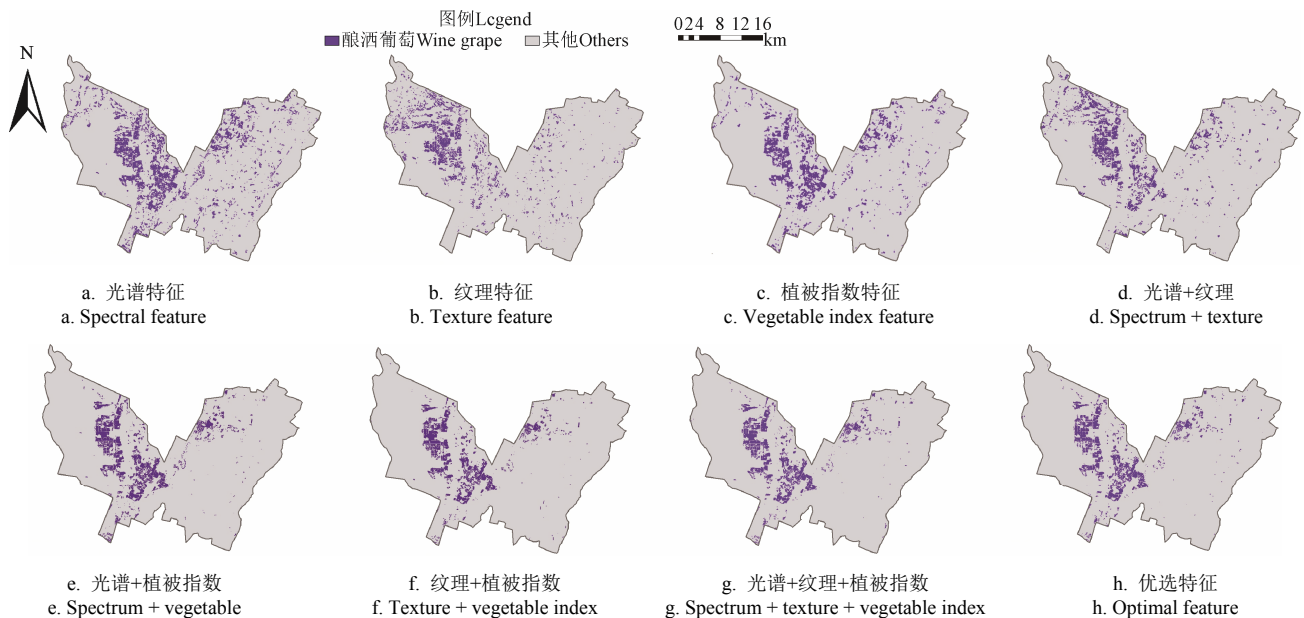


图 7 不同特征组合酿酒葡萄分布图

Fig.7 Distribution map of wine grape in different feature combinations

## 4 讨论

本研究针对当前农作物遥感监测中特征信息有效性弱、针对性不足以及特征单一等问题,共设计了 8 组特征组合方案,包含单一特征组合、多源特征组合以及优选特征组合,通过计算特征重要性得分,获取最佳特

征种类和数目,从而实现酿酒葡萄的精准识别,主要创新点在于:

首次在酿酒葡萄识别中引入 GF-6 卫星红边波段和红边指数。传统可见光波段和近红外波段中,不同地物间光谱反射率相近,尤其酿酒葡萄、耕地、林地、草地等存在严重的“异物同谱”现象;而在 2 个红边波段中,

“异物同谱”现象明显减弱,酿酒葡萄、耕地、林地、草地等地物分类精度得到提升,因此,可通过 2 个红边波段优势互补实现多地物复杂背景下酿酒葡萄的识别和提取。GF-6 卫星是中国首颗具有红边波段的卫星,本文利用 GF-6 卫星实现了对酿酒葡萄的精准识别,可为农业资源监测、防灾减灾等精准农业的发展提供数据支撑,推动国产卫星在农业领域的应用,同时也为国产卫星红边波段的性能改进和提升提供参考。

利用随机森林算法对多源特征集进行优选。通过随机森林算法计算特征的分类贡献度,获得特征重要性排名以及最佳特征数目,从而获取最佳特征组合。最佳特征组合中包含 4 个光谱特征、11 个植被指数特征和 2 个纹理特征,重要时相集中在 7 月、9 月、10 月和 11 月,原因在于此时酿酒葡萄与研究区内其他作物如玉米、小麦等物候特征差异较大,易于识别。优选特征组合与其余组合相比,分类精度最高,酿酒葡萄、耕地、林地等易混淆地物的误判面积显著减小,且地块破碎度低,更加符合实际情况。优选特征既结合了多源特征信息丰富的优势,又避免了特征冗余,最大程度利用有效信息。

以贺兰山东麓酿酒葡萄核心种植区永宁县为研究区对酿酒葡萄进行遥感识别,以点及面,为后续大尺度遥感监测奠定基础。酿酒葡萄种植业是贺兰山东麓的支柱产业,对发展当地特色农业及加强生态保护具有重要意义。本研究以核心区永宁县为切入点提取酿酒葡萄空间分布信息,为今后贺兰山东麓酿酒葡萄动态大尺度监测及管理提供了技术支撑和实践基础,具有重要的社会价值和生态价值。

以上是本文主要创新点,但与以往研究相比,还存在一些不足,比如本文植被指数特征采用时间序列,而光谱和纹理特征采用单时相数据,未深入研究特征变量时相不同带来的影响;本文重点研究了酿酒葡萄的分类情况,对于耕地、草地、林地等地物划分不够细致,易造成错分,应对其进行更详细的划分,比如耕地细分为玉米、小麦等,草地细分为草原、草甸等;国外对红边波段的研究多采用更高空间分辨率的 Sentinel 数据,对于酿酒葡萄而言,以地块尺度的酒庄种植为主,更适合从高分辨率角度开展研究。

## 5 结 论

本文基于 GF-6 遥感数据,以酿酒葡萄为主要研究对象,提取光谱特征、植被指数特征和纹理特征进行特征组合,通过随机森林算法对多源特征进行优选及分类,得到以下结论:

特征优选对合理配置资源、提高分类精度具有重要意义。通过计算特征重要性得分,本研究对特征变量进行降维,根据特征重要性得分在 54 个特征中选取 17 个有效特征。与引入全部特征相比,优选特征组合的总体精度提高了 2.85%,酿酒葡萄用户精度和制图精度分别提高了 2.77%和 2.24%。

GF-6 卫星红边波段的引入可以减弱“异物同谱”现

象,提高酿酒葡萄与其他地物的光谱响应差异。红边波段是绿色植被的敏感波段,在增大酿酒葡萄可分性和区分地物上优于近红外、绿光和蓝光等可见光波段,2 个红边波段特征重要性排名分别为第 3 和第 6,在所有光谱特征中具有更高的重要性得分。

针对研究不足之处,未来将在以下几个方面进行深入探索:优化分类算法,尝试其他机器学习算法并根据研究对象的特点对算法进行优化,实现大尺度遥感监测;尝试构建完整的对比模型,将 GF-6 与高分系列卫星的性能和效果进行对比,从定性和定量两方面深入探究红边波段对农作物识别的影响。

## [参 考 文 献]

- [1] 陈仲新, 任建强, 唐华俊, 等. 农业遥感研究应用进展与展望[J]. 遥感学报, 2016, 20(5): 748-767.  
Chen Zhongxin, Ren Jianqiang, Tang Huajun, et al. Progress and perspectives on agricultural remote sensing research and applications in China[J]. Journal of Remote Sensing, 2016, 20(5): 748-767. (in Chinese with English abstract)
- [2] 赵春江. 农业遥感研究与应用进展[J]. 农业机械学报, 2014, 45(12): 277-293.  
Zhao Chunjiang. Advances of research and application in remote sensing for agriculture[J]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(12): 277-293. (in Chinese with English abstract)
- [3] 余宝, 黄敬峰, 石晶晶, 等. 基于红边位置变化特征的油菜种植区域提取[J]. 农业工程学报, 2013, 29(15): 145-152.  
She Bao, Huang Jingfeng, Shi Jingjing, et al. Extracting oilseed rape growing regions based on variation characteristics of red edge position[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2013, 29(15): 145-152. (in Chinese with English abstract)
- [4] 邹红玉, 郑红平. 浅述植被“红边”效应及其定量分析方法[J]. 遥感信息, 2010(4): 112-116.  
Zou Hongyu, Zheng Hongping. The effect and method of quantitative analysis of “red edge” of vegetation[J]. Remote Sensing Information, 2010(4): 112-116. (in Chinese with English abstract)
- [5] 苏伟, 侯宁, 李琪, 等. 基于 Sentinel-2 遥感影像的玉米冠层叶面积指数反演[J]. 农业机械学报, 2018, 49(1): 151-156.  
Su Wei, Hou Ning, Li Qi, et al. Retrieving leaf area index of corn or canopy Sentinel-2 Remote Sensing based on image[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(1): 151-156. (in Chinese with English abstract)
- [6] 黄敬峰, 王渊, 王福民, 等. 油菜红边特征及其叶面积指数的高光谱估算模型[J]. 农业工程学报, 2006, 22(8): 22-26.  
Huang Jingfeng, Wang Yuan, Wang Fumin, et al. Red edge characteristics and leaf area index estimation model using hyperspectral data for rape[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2006, 22(8): 22-26. (in Chinese with English abstract)
- [7] Huang Wenjiang, Guan Qingsong, Luo Juhua, et al. New optimized spectral indices for identifying and monitoring winter wheat diseases[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2014, 7(6):

- 2516-2524.
- [8] 刘佳, 王利民, 滕飞, 等. RapidEye 卫星红边波段对农作物面积提取精度的影响[J]. 农业工程学报, 2016, 32(13): 140-148.  
Liu Jia, Wang Limin, Teng Fei, et al. Impact of red-edge waveband of RapidEye satellite on estimation accuracy of crop planting area[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2016, 32(13): 140-148. (in Chinese with English abstract)
- [9] Adelabu S, Mutanga O, Adam E. Evaluating the impact of red-edge band from Rapideye image for classifying insect defoliation levels[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2014, 95: 34-41.
- [10] Tigges J, Lakes T, Hostert P. Urban vegetation classification: Benefits of multitemporal RapidEye satellite data[J]. Remote Sensing of Environment, 2013, 136: 66-75.
- [11] Sertel E, Yay I. Vineyard parcel identification from Worldview-2 images using object-based classification model[J]. Journal of Applied Remote Sensing, 2014, 8: 1-17.
- [12] Hughes G F. On the mean accuracy of statistical pattern recognizers[J]. IEEE Transactions on Information Theory, 1968, 14(1): 55-63.
- [13] 王娜, 李强子, 杜鑫, 等. 单变量特征选择的苏北地区主要农作物遥感识别[J]. 遥感学报, 2017, 21(4): 519-530.  
Wang Na, Li Qiangzi, Du Xin, et al. Identification of main crops based on the univariate feature selection in Subei[J]. Journal of Remote Sensing, 2017, 21(4): 519-530. (in Chinese with English abstract)
- [14] 王文静, 张霞, 赵银娣, 等. 综合多特征的 Landsat 8 时序遥感图像棉花分类方法[J]. 遥感学报, 2017, 21(1): 115-124.  
Wang Wenjing, Zhang Xia, Zhao Yindi, et al. Cotton extraction method of integrated multi-features based on multitemporal Landsat 8 images[J]. Journal of Remote Sensing, 2017, 21(1): 115-124. (in Chinese with English abstract)
- [15] 程希萌, 沈占锋, 邢廷炎, 等. 基于 mRMR 特征优选算法的多光谱遥感影像分类效率精度分析[J]. 地球信息科学学报, 2016, 18(6): 815-823.  
Cheng Ximeng, Shen Zhanfeng, Xing Tingyan, et al. Efficiency and accuracy analysis of multispectral image classification based on mRMR feature selection method[J]. Journal of Geo-information Science, 2016, 18(6): 815-823. (in Chinese with English abstract)
- [16] Atkinson J T, Ismail R, Robertson M. Mapping bugweed (*Solanum mauritianum*) infestations in pinus patula plantations using hyperspectral imagery and support vector machines[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2014, 7(1): 17-28.
- [17] 张磊, 宫兆宁, 王启为, 等. Sentinel-2 影像多特征优选的黄河三角洲湿地信息提取[J]. 遥感学报, 2019, 23(2): 313-326.  
Zhang Lei, Gong Zhaoning, Wang Qiwei, et al. Wetland mapping of Yellow River Delta wetlands based on multi-feature optimization of Sentinel-2 images[J]. Journal of Remote Sensing, 2019, 23(2): 313-326. (in Chinese with English abstract)
- [18] 蒙莉娜, 丁建丽, 王敬哲, 等. 基于环境变量的渭干河-库车河绿洲土壤盐分空间分布[J]. 农业工程学报, 2020, 36(1): 175-181.  
Meng Lina, Ding Jianli, Wang Jingzhe, et al. Spatial distribution of soil salinity in Ugan-Kuqa River delta oasis based on environmental variables[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2020, 36(1): 175-181. (in Chinese with English abstract)
- [19] 韩颖娟, 张磊, 卫建国, 等. 宁夏酿酒葡萄生育期气象条件及管理措施综述[J]. 中国农业气象, 2011, 32(增刊): 108-112.  
Han Yingjuan, Zhang Lei, Wei Jianguo, et al. Review on research of meteorological conditions and managing measure for wine grape growth in Ningxia[J]. Chinese Journal of Agrometeorology, 2011, 32(Suppl.): 108-112. (in Chinese with English abstract)
- [20] 郑利娟. 基于高分一/六号卫星影像特征的农作物分类研究[D]. 北京: 中国科学院大学, 2017.  
Zheng Lijuan. Crop Classification Using Multi-features of Chinese Gaofen-1/6 Satellite Remote Sensing Images[D]. Beijing: University of Chinese Academy of Sciences, 2017. (in Chinese with English abstract)
- [21] Zhong Liheng, Hu Lina, Zhou Hang. Deep learning based multi-temporal crop classification[J]. Remote Sensing of Environment, 2019, 221: 430-443.
- [22] 王福民, 黄敬峰, 王秀珍, 等. 波段位置和宽度对不同生育期水稻 NDVI 影响研究[J]. 遥感学报, 2008, 12(4): 626-632.  
Wang Fumin, Huang Jingfeng, Wang Xiuzhen, et al. Effects of band position and band width on NDVI measurements of rice at different growth stages[J]. Journal of Remote Sensing, 2008, 12(4): 626-632. (in Chinese with English abstract)
- [23] Bolton D K, Friedl M A. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics[J]. Agricultural and Forest Meteorology, 2013, 173: 74-84.
- [24] 张晶, 占玉林, 李如仁. 高分一号归一化植被指数时间序列用于冬小麦识别[J]. 遥感信息, 2017, 32(1): 50-56.  
Zhang Jing, Zhan Yulin, Li Ruren. Application of GF-1 NDVI time series in winter wheat identification[J]. Remote Sensing Information, 2017, 32(1): 50-56. (in Chinese with English abstract)
- [25] Gitelson A, Merzlyak M N. Spectral reflectance changes associated with autumn senescence of *aesculus hippocastanum* L. and *acer platanoides* L. leaves. Spectral features and relation to chlorophyll estimation[J]. Journal of Plant Physiology, 1994, 143(3): 286-292.
- [26] Tucker Compton J. Red and photographic infrared linear combinations for monitoring vegetation[J]. Remote Sensing of Environment, 1979, 8(2): 127-150.
- [27] Huete A, Didan K, Miura T, et al. Overview of the radiometric and biophysical performance of the MODIS vegetation indices[J]. Remote Sensing of Environment, 2002, 83(1): 195-213.
- [28] Kim H O, Yeom J M. Effect of red-edge and texture features for object-based paddy rice crop classification using RapidEye multi-spectral satellite image data[J]. International Journal of Remote Sensing, 2014, 35(19): 7046-7068.
- [29] 刘明, 李忠任, 张海涛, 等. 基于二分搜索结合修剪随机森林的特征选择算法在近红外光谱分类中的应用[J]. 激光与光电子学进展, 2017, 54(10): 455-462.  
Liu Ming, Li Zhongren, Zhang Haitao, et al. Feature selection algorithm application in near-infrared spectroscopy classification based on binary search combined with random forest pruning[J]. Laser & Optoelectronics Progress, 2017, 54(10): 455-462. (in Chinese with English abstract)



- [30] Breiman L. Random forests[J]. *Machine Learning*, 2001, 45: 5-32.
- [31] 王利民, 刘佳, 杨玲波, 等. 随机森林方法在玉米-大豆精细识别中的应用[J]. *作物学报*, 2018, 44(4): 569-580.  
Wang Limin, Liu Jia, Yang Lingbo, et al. Application of random forest method in maize-soybean accurate identification[J]. *Acta Agronomica Sinica*, 2018, 44(4): 569-580. (in Chinese with English abstract)
- [32] Yang Lingbo, Mansaray Lamin, Huang Jingfeng, et al. Optimal segmentation scale parameter, feature subset and classification algorithm for geographic object-based crop recognition using multisource satellite imagery[J]. *Remote Sensing*, 2019, 11(5): 1-23.
- [33] 李静, 杨敏华, 吴孔江. 基于波段选择的高光谱遥感影像分类[J]. *测绘信息与工程*, 2012, 37(2): 41-44.  
Li Jing, Yang Minhua, Wu Kongjiang. Band selection based hyperspectral remote sensing image classification[J]. *Journal of Geomatics*, 2012, 37(2): 41-44. (in Chinese with English abstract)

## Accurate recognition of wine grapes using multi-feature optimization based on GF-6 satellite images

Li Wenjie<sup>1,2</sup>, Guo Xiaolei<sup>3</sup>, Yang Lingbo<sup>1,2</sup>, Yan Ming<sup>3</sup>, Zou Chenxi<sup>4</sup>, Fang Yahua<sup>4</sup>, Sun Han<sup>5</sup>, Huang Jingfeng<sup>1,2\*</sup>

(1. *Institute of Agricultural Remote Sensing and Information Technology Application, Zhejiang University, Hangzhou 310058, China*;  
2. *Key Laboratory of Agricultural Remote Sensing and Information System, Zhejiang University, Hangzhou 310058, China*;  
3. *The Yinchuan Meteorological Bureau, Yinchuan 750000, China*; 4. *Spaider Information Technology Co. Ltd, Beijing 100086, China*;  
5. *Xintianyuan Institute of Disaster Prevention and Reduction, Inner Mongolia, Hohhot 010051, China*)

**Abstract:** Multi-source remote sensing information and feature optimization have become important supports to improve the accuracy of crop recognition. As the first Chinese satellite to introduce red-edge bands, the rich spectral information of GF-6 satellite provides new ideas for crop recognition. However, the use of crop features is confine to one single source, and previous studies on cash crop is relatively lacking. In this study, an available recognition method was proposed for the wine grape on multi-features using GF-6 satellite images. This paper first introduced the red-edge bands of GF-6 to the multi-source features in the study of accurate recognition for wine grape. Based on the GF-6 satellite data from June 2018 to March 2019 of Yongning County, Yinchuan City, Ningxia Hui Autonomous Region, two red-edge bands were selected to extract the spectral, texture, and temporal vegetation index features, including Normalized Difference Red-Edge Index (NDRE), Normalized Difference Vegetation Index (NDVI), and Enhanced Vegetation Index (EVI). A Random Forest algorithm with Gini index was used to choose the optimal number of features according to the measured importance scores, thereby to construct the optimal feature combination. Seven types samples of ground objects were selected to accurately recognize wine grape. Seven combinations of comparative features were designed, including three single-source and four multi-source feature combinations. Training samples and verification samples were obtained by the field investigation and visual interpretation with the Google Earth. The results showed that, compared with single-source features, the multi-source remote sensing features significantly improved the recognition effect of wine grape, where the vegetation index features contributed the most, then followed by the spectral and texture features. The accuracy of user and producer for the wine grape were 94.23% and 92.59%, respectively, the overall classification accuracy was 94.15%, and the Kappa coefficient was 0.93 in the optimal feature combination. Compared with spectral feature combination, the Jeffries-Matusita distance between wine grape-farmland and wine grape-woodland were improved from 1.57 to 1.99 and 1.40 to 1.97 respectively in the optimal feature combination. Compared with the combination 7, which includes all fifty-four features, the overall accuracy was improved by 2.85% in the optimal feature combination with only seventeen features. Taking four wine chateaus by field survey as the verification area, the results of wine grape extraction were compared with the statistical data, where the relative area accuracy of eight feature combinations were all above 70%, and that of the optimal feature combination was above 90%. Compared with other seven feature combinations, the optimal feature combination improved the separable measure of different ground objects and reduced field fragment, indicating more conformable with actual situation. In addition, the operation time of classification model was shortened, and the reasonable allocation of resources was realized by feature optimization. The successful launch of GF-6 enriched the existing satellite data sources, including red-edge bands, (such as RapidEye of Germany and Sentinel-2 of Europe). The findings can contribute to the large-scale remote sensing monitoring of wine grape and popularize the application of red-edge bands of Chinese satellite in agriculture, and also provide a sound reference to improve the performance of red-edge bands of Chinese satellite.

**Keywords:** remote sensing; image processing; GF-6 satellite; wine grape; feature optimization; Random Forest; red-edge band; accurate recognition