

引用格式:温小乐,钟奥,胡秀娟.基于随机森林特征选择的绿化乔木树种分类[J].地球信息科学学报,2018,20(12):1777-1786. [Wen X L, Zhong A, Hu X J. The classification of urban greening tree species based on feature selection of random forest[J]. Journal of Geo-information Science, 2018,20(12):1777-1786.] DOI:10.12082/dqxxkx.2018.180310

基于随机森林特征选择的绿化乔木树种分类

温小乐^{1,2,3}, 钟 奥^{1,2}, 胡秀娟^{1,2,3*}

1. 福州大学环境与资源学院,福州 350116;
2. 福州大学遥感信息工程研究所,福州 350116;
3. 福建省水土流失遥感监测评估与灾害防治重点实验室,福州 350116

The Classification of Urban Greening Tree Species Based on Feature Selection of Random Forest

WEN Xiaole^{1,2,3}, ZHONG Ao^{1,2}, HU Xiujuan^{1,2,3*}

1. College of Environment and Resources, Fuzhou University, Fuzhou 350116, China; 2. Institute of remote Sensing Information Engineering, Fuzhou University, Fuzhou 350116, China; 3. Fujian Provincial Key Laboratory of Remote Sensing of Soil Erosion and Disaster Protection, Fuzhou University, Fuzhou 350116, China

Abstract: Since Urban forests played important roles in improving air, water and land quality, absorbing and mitigating carbon dioxide and many pollutants, mitigating urban heat island and reducing storm water runoff, its monitoring is a major issue for urban planners. It is of great significance to obtain the tree species timely and precisely in urban planning and green space management. At present, urban forest tree species mapping has benefitted from advances in remote sensing techniques. Using an object-oriented method combining spectral, textural, indicial and geometric features from high-resolution WorldView-2 imagery, this paper aimed to carry out the classification of seven main tree species in Fuzhou university, including Banyan (*Ficus microcarpa*), Mango(*Mangifera indica* L.), Camphor tree (*Cinnamomum camphora*), Bishop wood (*Bischofia polycarpa*), Chinese orchid tree(*Bauhinia purpurea* L.), Weeping fig (*Ficus benjamina* L.), and Kapok tree (*Bombax malabaricum* DC.). A random forest method was employed to determine the feature selection in this study. When eliminating 20 percent of the total features, the in situ validation results showed that the overall accuracy reached a highest value of 74.95% with Kappa coefficient of 0.67 when using 34 features for classification, which including 15 spectral features, 6 textural features, 8 indicial features and 5 geometric features, and the feature of mean spectral was the most significant, however, the standard deviation of each band is less important. The results also revealed that the feature selection of random forest could reduce or avoid the data redundancy and Hughes phenomenon, and thus could improve the classification accuracy of same type tree species. Moreover, the four additional bands of WorldView-2 imagery, especially the yellow and red edge band, and their composite indexes showed a higher importance in classification, which also indicates that these bands have great application prospects in vegetation remote sensing, especially in tree species classification.

Key words: WorldView-2; object-oriented; random forest; feature selection; tree species classification

***Corresponding author:** HU Xiujuan, E-mail: huxiujuan@fzu.edu.cn

收稿日期:2018-07-03;修回日期:2018-09-28.

基金项目:福建省教育厅重点科技项目(JA15044);福建省教育厅中青年教师教育科研项目(JAT170059). [**Foundation items:** Key Technologies Program of Fujian Educational Committee, No.JA15044; Educational Research Project for Young and Middle-aged Teachers of Fujian Educational Committee, No.JAT170059.]

作者简介:温小乐(1976-),女,博士,副教授,研究方向为环境资源遥感与环境评价. E-mail: wenxiaole@sina.com

***通讯作者:**胡秀娟(1982-),女,博士生,讲师,研究方向为环境资源遥感与景观规划. E-mail: huxiujuan@fzu.edu.cn

摘要:城市绿化在改善空气、水和土壤质量,吸收和减少二氧化碳及各种污染物,缓解城市热岛和减少雨水径流等方面发挥着重要作用。及时准确地获取树种信息是城市规划与绿化管理的先决条件,对进一步改善城市生态环境 also 具有重要意义。基于遥感技术,使用高空间分辨率的 WorldView-2 卫星影像,采用光谱、纹理、指数以及几何等多种特征相结合的面向对象方法,并通过随机森林进行特征选择,对福州大学旗山校区北部的榕树、杧果、香樟、重阳木、羊蹄甲、垂叶榕以及木棉 7 种主要绿化乔木进行树种分类。实地验证结果表明:通过特征选择可以减少或规避数据冗余以及休斯效应的产生,该方法可以提高现有同类型树种分类的精度,当淘汰全部特征的 20%,利用 34 个特征(包括 15 个光谱特征、6 个纹理特征、8 个指数特征和 5 个几何特征)进行分类时,总精度最高,可达 74.95%,Kappa 系数为 0.67。其中,光谱平均值的特征重要性最高,而各波段的标准差的重要性较低。WorldView-2 卫星影像的 4 个新增波段,特别是黄光和红边波段及其构建的指数特征重要性较高,也说明这些波段在植被遥感,特别是树种分类中极具应用前景。

关键词:WorldView-2;面向对象;随机森林;特征选择;树种分类

1 引言

城市绿化在减缓城市热岛效应及提高城市环境效益方面作用突出,其中树种多样性也是描述生态环境质量的重要参数之一^[1]。随着对城市系统认识的逐步深入,人们越来越意识到树种分类信息对研究城市气候、制定城市规划具有不可或缺的作用。树种信息也为城市的规划管理者提供了监测、控制与评估城市绿化的主要依据^[2-3]。及时、准确地获取城市树种信息对于改善城市环境、推进城市可持续发展至关重要^[4]。

目前获取准确的树种信息有一定的难度,现有的树种调查方法大致分为实地调查和遥感解译 2 种。基于实地调查的树种分类只适用于较小的区域范围,需要投入大量的人力物力且耗时巨大^[5]。近年来,遥感解译识别树种信息正成为研究热点,遥感数据源多为多光谱、高光谱、LiDAR 以及 SAR 等。其中,有部分学者利用雷达以及高光谱数据进行树种分类的研究^[6-8],但成本高昂以及应用的地域局限性,限制了其应用的前景。多光谱数据(如 IKONOS 及 Sentinel-2A 等)都在树种分类研究中得到了一定的应用^[9-10],但由于 IKONOS 波段数量较少、Sentinel-2A 空间分辨率较低等原因,在城市树种分类的研究中也有一定的局限性。而 WorldView-2 数据在保持高空间分辨率的同时加入了 4 个新增波段,在满足空间分辨率需求的同时也满足光谱分辨率的需求,因此本文选用该影像进行研究。

现有的树种分类研究多针对农田以及森林地区,而针对城市地区的树种分类研究相对较少。由于景观美化的需求以及植物配置多样化的原因,城市绿化是复杂多变的,相对较高的局部空间变化为遥感自动分类研究带来了一定的困难^[11]。仅依靠光谱曲线无法区分不同的绿化树种,而面向对象的方法可以将影像分割成均匀区域并且通过一系列

特征如光谱、纹理和几何等来提取该区域的信息^[12],从而解决这个问题。如 Immitzer 等^[13]对奥地利布尔根兰州的 10 种乔木树种进行分类,使用基于像元的方法分类总精度为 69%,当使用面向对象的方法时,精度得到提高。机器学习算法在进行大量且复杂的遥感数据分析时表现出卓越的性能,如 Pu 等^[14]利用 LDA 和 CART 方法对美国坦帕市的 7 种乔木树种进行分类,总精度分别为 65.61% 和 67.22%,表明非参数机器学习算法在树种分类上的性能要优于参数机器学习算法。随机森林(Random Forest, RF)是一种基于决策树分类的非参数机器学习算法,只需较少的训练数据即可获得较高的分类精度,已有部分学者将其应用于城市树种分类的研究当中,并取得了较好的分类结果,如 Naidoo 等^[15]结合 LiDAR 与高光谱数据对克鲁格国家公园 8 种热带稀树草原树种进行研究,得出 RF 是在异质性较高的环境中,对树种分类最适用的方法;李丹等^[16]对首都师范大学和北京师范大学周边地区分别进行优势乔木树种分类,利用 RF 算法分类精度分别为 75.8% 和 65.3%。由于大量的特征数据集不但会造成数据的冗余,还容易产生休斯效应^[17],因此利用 RF 中的特征重要性分析,对相关特征进行筛选,并分析得出适合的特征以提高分类精度。

有鉴于此,本文通过面向对象的方法对城市绿化树种进行特征提取,通过 RF 中的特征重要性分析对提取的光谱、纹理、指数和几何等特征进行排序、筛选与分析,以达到提高分类精度的目的,并评估不同特征对分类结果的影响。

2 研究区概况

本文以福建省福州市福州大学旗山校区为研究区(图 1),其地理坐标为 26°03'~26°05' N, 119°11'

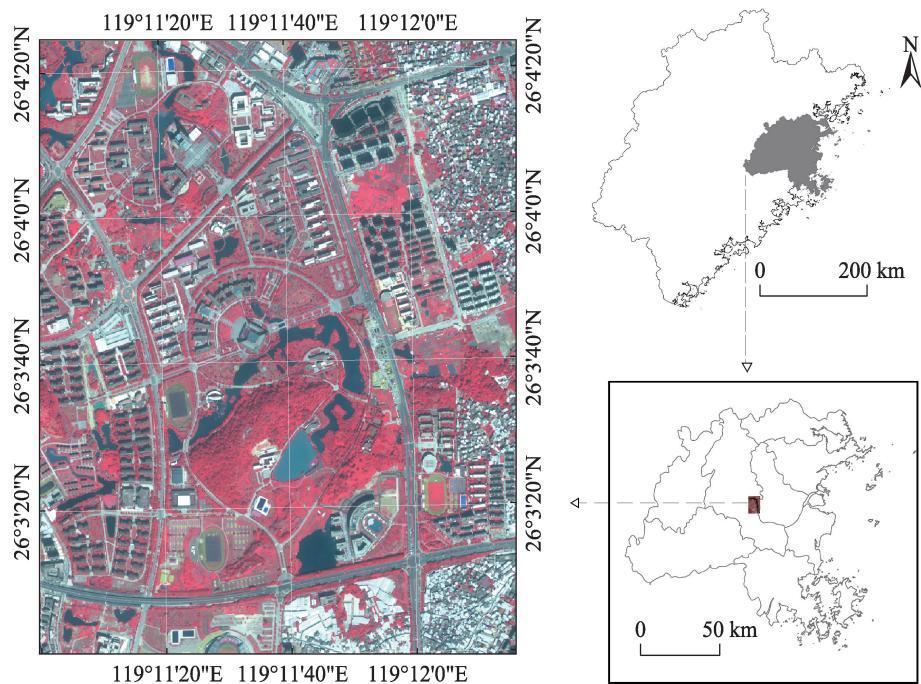


图1 研究区在福州市的地理位置
Fig. 1 A location map of the study area in Fuzhou

~119°13' E,属于典型的亚热带季风气候,年平均气温为20~25 ℃。选取校区内人工绿化较为完善的地块(面积约0.6 km²)作为实验区(图1中黑色多边形区域)进行主要绿化乔木的判别与分类。

3 数据源及技术方法

3.1 遥感数据源

本文选用的WorldView-2卫星影像获取时间为2017年9月25日10时58分,影像无云且质量好,平均侧摆角为11.7°。WorldView-2卫星由Digital-Globe公司发射,是具有8个多光谱波段的商用高分辨率卫星(传感器),幅宽为16.4 km,重访周期平均为1.1 d。详细的光谱和空间信息见表1。

3.2 遥感数据预处理

对影像进行辐射定标以及FLAASH大气校正等,将其亮度值(DN)转换为传感器处反射率(At-sensor Reflectance),用来消除大气引起的影响。影像融合可以使丰富的纹理和色彩信息相结合,充分利用高分辨率遥感影像的信息内容。本次使用Gram-Schmidt Spectral Sharpening方法对影像进行融合,该方法不受波段数量的限制并且可以较好地保持原影像的光谱和纹理信息^[18]。

表1 WorldView-2影像的光谱和空间信息
Tab. 1 The spectral and spatial information of WorldView-2 imagery

波段名	波长/nm	空间分辨率/m
海岸波段	400~450	2.00
蓝光波段	450~510	2.00
绿光波段	510~580	2.00
黄光波段	585~625	2.00
红光波段	630~690	2.00
红边波段	705~745	2.00
近红外1波段	770~895	2.00
近红外2波段	860~1040	2.00
全色波段	450~800	0.50

3.3 技术方法

3.3.1 样本数据选取

样本数据的获取均通过实地调查进行,沿着研究区主要道路进行绿化树种的记录,排除数量较少以及被其他冠层遮挡的树种最终将研究树种分为7类,分别为榕树(*Ficus microcarpa*)、杧果(*Mangifera indica* L.)、香樟(*Cinnamomum camphora*)、重阳木(*Bischofia polycarpa*)、羊蹄甲(*Bauhinia purpurea* L.)、垂叶榕(*Ficus benjamina* L.)以及木棉(*Bombax malabaricum* DC.),其中榕树包括黄葛树(*Ficus virens* Aiton)和高山榕(*Ficus altissima*)。各树种训练样本如图2所示。

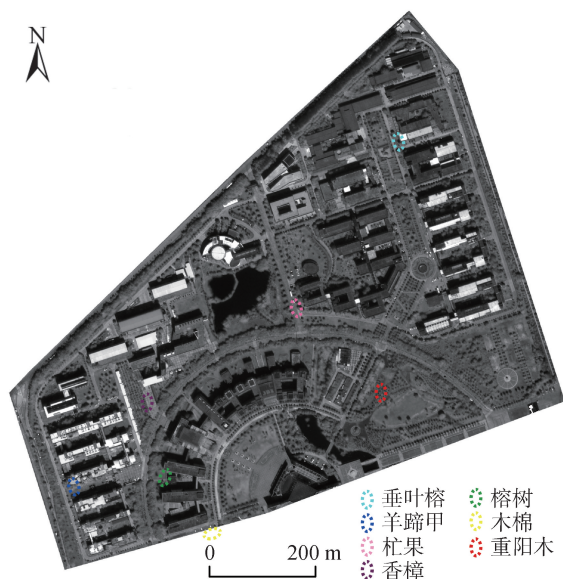


图2 各树种的训练样本位置

Fig. 2 Locations of training data for different tree species

3.3.2 影像分割

影像分割使用eCognition软件中的多尺度分割算法与光谱差异分割算法。多尺度分割由对象的尺度、形状、颜色、紧凑性以及平滑性所决定,颜色代表的是光谱的均匀性,而形状则考虑物体的几何地貌特征。光谱差异分割通常应用在多尺度分割之后,通过合并对象来增强分割结果。在分割时加入土壤调节植被指数(SAVI)^[19]有助于植被的研究^[20]。分割时将SAVI指数正规化后保持与8个多光谱波段值域相同,其公式表示为式(1)。

$$NSAVI_i = (SAVI_i - SAVI_{\min}) / (SAVI_{\max} - SAVI_{\min}) \quad (1)$$

式中: $NSAVI_i$ 为正规化后的SAVI值; $SAVI_i$ 为SAVI在像元 i 的值; $SAVI_{\max}$ 和 $SAVI_{\min}$ 分别为SAVI的最大值和最小值。

3.3.3 随机森林

Leo Breiman于2001年提出了随机森林算法^[21],是一种基于分类和回归树的多决策树分类器。每一棵决策树执行Bootstrap抽样(一种有放回的抽样算法)之后基于样本的袋外数据误差 OOB(out-of-bag) error 来估计计算误差。RF不考虑每个节点上的所有变量来确定最佳的分割阈值,而是使用原始特征集的随机子集,因此产生了大量的非相关决策树。RF具有不需要变量遵循特定的统计分布、训练样本少、对过度拟合不太敏感以及可以对特征重要性进行排序等优点,适用于城市树种分类的研究。本次研究中决策树的数量设为500,每个节点的特征数使用默认数值即特征总数的平方根。

3.3.4 特征提取与选择

对于融合后的WorldView-2影像而言,高空间分辨率以及8个波段可以提供更加多样的信息。结合面向对象的方法对各对象进行特征的提取(表2),本研究主要使用光谱、纹理、指数以及几何这4类特征。

(1)光谱特征主要指8个波段的平均值以及标准差,共产生16种光谱特征。

(2)纹理特征使用Haralick等^[22]于1973年提出的灰度共生矩阵来描述。灰度共生矩阵(Gray-Level Co-occurrence Matrix, GLCM)提供了影像中像元与整体影像及像元与像元之间的空间关系,用2个位置像元的联合条件概率密度来表示纹理。灰度差异矢量(Gray-Level Difference Vector, GLDV)是GLCM对角线的总和,它计算相邻像元间的绝对差异。共产生12个纹理特征。如果对8个波段分别计算则产生96个特征,计算量巨大并产生信息冗余。主成份分析(PCA)可以将多个变量通过多维数据压缩技术选出少数重要变量,将多维的信息集中到少数的几个特征分量上,一般主要信息集中在前2个分量中^[23]。因此,通过PCA对影像进行分析,选用占据了90.46%特征贡献率的第一主成分影像来提取纹理特征,共8个GLCM特征和4个GLDV特征。

(3)指数特征是指利用指数对影像进行特征提取。比值型指数通常选取反射差异较大的两个波段,反射强的波段置于分子,通过比值运算强化差异。利用WorldView-2影像新增的4个波段,将树木反射率差异较大的波段选出构建比值型指数并进行归一化处理。因此,本文创建了6个比值型指数(Normalized Difference Index, NDI)进行研究,分别用 NDI_{61} 、 NDI_{65} 、 NDI_{84} 、 NDI_{74} 、 NDI_{85} 以及 NDI_{86} 表示,其中 NDI_{mn} 中 m 代表反射率较高波段, n 代表反射率较低波段。此外,本文还选用SAVI以及归一化差值植被指数(NDVI)进行提取,共得到8个指数特征。

(4)几何特征基于对象的形状获取信息,根据构成对象的像元计算特征值。共选取6个与绿化乔木树种相关的几何特征^[14,20]。各特征描述或计算公式见表2。

特征选择选取了特征变量中重要性较高的特征变量。特征选取的方法是利用RF中的特征重要性分析,通过大量的决策树构建RF,计算每一棵决策树的OOB误差,之后对所有OOB数据中某特征值加入干扰并再次计算OOB误差,如果OOB误差

表2 从 WorldView-2 影像对象中提取的特征
Tab. 2 Image-object (IO) features extracted from WorldView-2 imagery

特征类型	特征名称	描述或公式
光谱特征	平均值	1-8 波段的光谱平均值
	标准差	1-8 波段的光谱标准差
指数特征	NDI ₆₁	$\frac{\rho_{\text{Red-Edge}} - \rho_{\text{Coastal}}}{\rho_{\text{Red-Edge}} + \rho_{\text{Coastal}}}$
	NDI ₆₄	$\frac{\rho_{\text{NIR2}} - \rho_{\text{Yellow}}}{\rho_{\text{NIR2}} + \rho_{\text{Yellow}}}$
	NDI ₈₆	$\frac{\rho_{\text{NIR2}} - \rho_{\text{Red-Edge}}}{\rho_{\text{NIR2}} + \rho_{\text{Red-Edge}}}$
	NDI ₆₅	$\frac{\rho_{\text{Red-Edge}} - \rho_{\text{Red}}}{\rho_{\text{Red-Edge}} + \rho_{\text{Red}}}$
	NDI ₇₄	$\frac{\rho_{\text{NIR1}} - \rho_{\text{Yellow}}}{\rho_{\text{NIR1}} + \rho_{\text{Yellow}}}$
	NDI ₈₅	$\frac{\rho_{\text{NIR2}} - \rho_{\text{Red}}}{\rho_{\text{NIR2}} + \rho_{\text{Red}}}$
	NDVI	$\frac{\rho_{\text{NIR1}} - \rho_{\text{Red}}}{\rho_{\text{NIR1}} + \rho_{\text{Red}}}$
	SAVI	$\frac{(\rho_{\text{NIR1}} - \rho_{\text{Red}})(1 + L)}{\rho_{\text{NIR1}} + \rho_{\text{Red}} + L}, \text{ 其中 } L=0.5$
	GLCM Mean	$\frac{\sum_{i,j=0}^{N-1} P_{i,j}}{N^2}$
	GLCM Std.dev	$\sqrt{\sum_{i,j=0}^{N-1} P_{i,j} (i,j - \mu_{i,j})}$
纹理特征	GLCM Homogeneity	$\sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1 + (i - j)^2}$
	GLCM Contrast	$\sum_{i,j=0}^{N-1} P_{i,j} (i - j)^2$
	GLCM Dissimilarity	$\sum_{i,j=0}^{N-1} P_{i,j} i - j $
	GLCM Entropy	$\sum_{i,j=0}^{N-1} P_{i,j} (-\ln P_{i,j})$
	GLCM Angular second moment	$\sum_{i,j=0}^{N-1} (P_{i,j}) \left[\frac{(i, \mu_i)(j - \mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}} \right]$
	GLCM Correlation	$\sum_{i,j=0}^{N-1} (P_{i,j})^2$
	GLDV Entropy	$\sum_{k=0}^{N-1} V_k (-\ln V_k)$
	GLDV Mean	$\sum_{k=0}^{N-1} k(V_k)$
	GLDV Contrast	$\sum_{k=0}^{N-1} V_k k^2$
	GLDV Angular second moment	$\sum_{k=0}^{N-1} V_k^2$
几何特征	Compactness P	$\frac{4 \times \pi \times \text{面积}}{\text{周长}}$
	Compactness	对象的紧致程度
	Shape index	对象的光滑程度
	Roundness	对象与椭圆的相似程度
	Border index	对象的不规则程度
	Number of edges	对象边的数量

注: 其中 i 是行号; j 是列号; P_{ij} 是单元格 i, j 中的归一化值; N 是行数或者列数; $\mu_{i,j}$ 是 GLCM 的平均值; $\sigma_{i,j}$ 是 GLCM 的标准差; V_k 是单元格 i, j 矩阵中的值, $k=1, 2, \dots, n$

增加幅度越大即精准度下降越多则说明其重要性越高。通过多次迭代,在每一次迭代中排除特征重要性排名靠后的20%左右的特征并重复这个过程。

4 结果与分析

4.1 影像分割结果

对影像设置不同的分割尺度(50–100)进行多尺度分割。根据研究表明对颜色设置较高的权重有助于优化分割结果,因此将颜色权重设置较高(0.6、0.7、0.8、0.9),形状与颜色权重之和为1。紧致度以及光滑度的权重各自为0.5。之后对影像进行光谱差异分割,可以提高分类效率^[20]。最后设定设置分割尺度为75,形状权重为0.1,光谱差异分割的阈值为200,图3为不同分割尺度下的结果,其中图3(a)为原始影像(RGB:753),图3(e)为最终选择的分割影像。图3(d)–(f)分别为50、75以及100分割尺度下的分割结果,可以很明显的看出100分割尺度下绿化乔木连成一片,对比矩形区域可以看出50分割尺度下对象过于细碎。图3(b)、(e)为分割尺度75时,形状权重分别为0.2和0.1的分割结果,

椭圆区域显示权重为0.2时有部分乔木连成一片。图3(e)、(c)为加入SAVI指数和不加入SAVI指数的分割结果,可以看出加入SAVI指数后分割效果明显加强。

4.2 特征选择结果

特征数量与分类总精度的关系如图4所示,为了验证特征选择的有效性,本文采用两种非参数机器学习算法,RF与支持向量机(Support Vector Machine, SVM)同时进行实验。结果表明,并非特征数量越多精度越高,无论是RF还是SVM,特征选择都有助于分类精度的提高,且RF在本次研究中的表现总是优于SVM。

对特征选择中淘汰的特征以及精度最高的一次分类中的特征重要性的排名(图5)进行分析可知,黄光波段的光谱平均值特征在特征重要性中排名第一,这是由于研究区内存在叶色黄绿的羊蹄甲和香樟,2类共占研究区乔木树种的37.71%,叶绿素含量降低和类胡萝卜素的增多会导致叶色发黄,在光谱变化上表现为黄光附近反射率升高^[24]。因此,树种在黄光波段的可分性增强;几乎所有的指

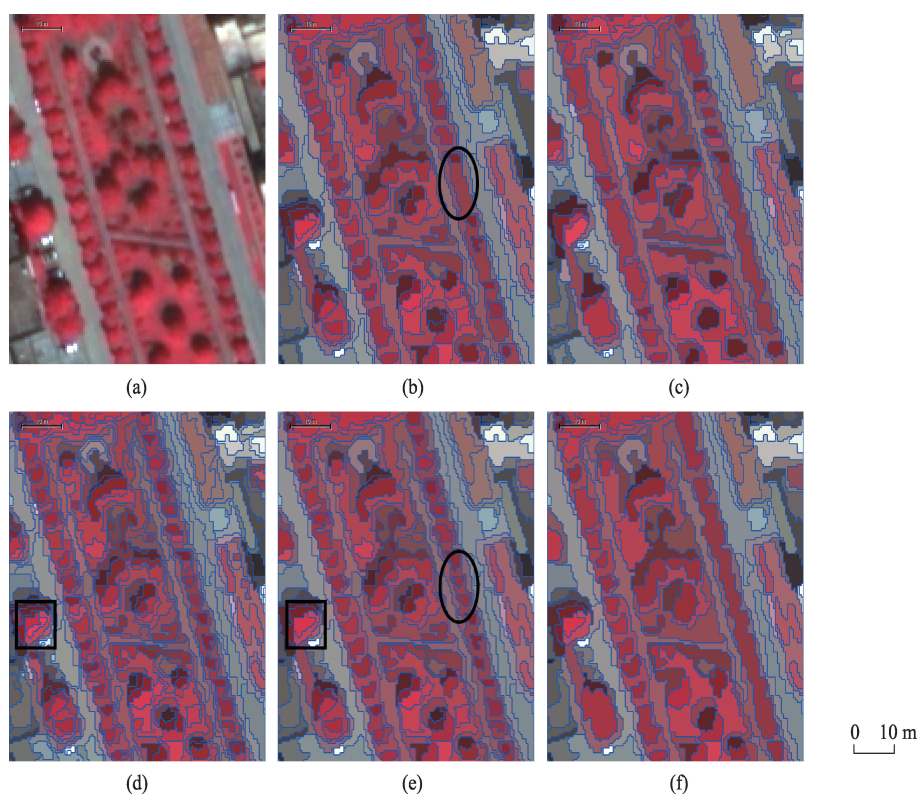


图3 影像分割的结果

Fig. 3 Segmentation results

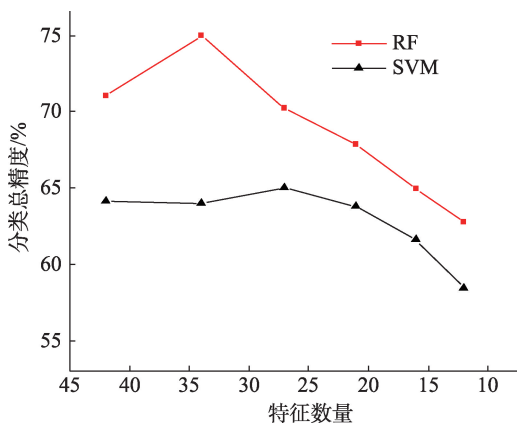


图4 特征数量与分类总精度的关系

Fig. 4 Relationship between the number of features and overall accuracy

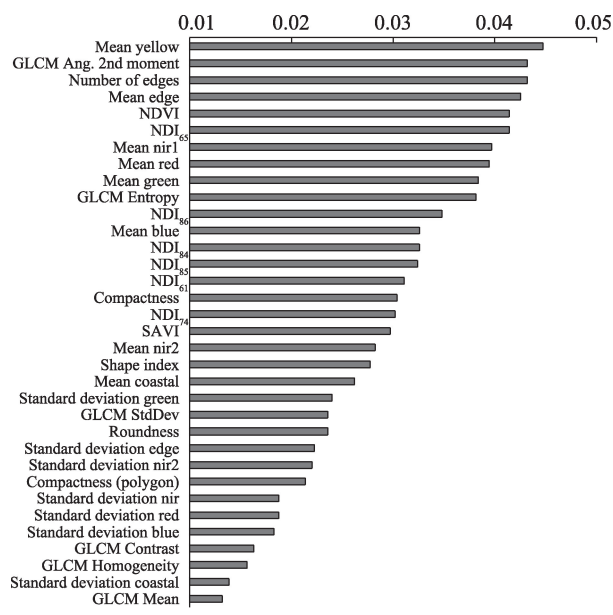


图5 选取的特征重要性排序

Fig. 5 Importance Ranking of Seleted Features

数特征和光谱平均值特征都在前50%,说明这2个特征对城市绿化乔木树种分类有着重要的作用。由WorldView-2新增的4个波段构成的指数特征的重要性也大部分位于前40%,并且优于SAVI,这表明海岸、黄、红边和近红外2波段对树种分类的研究具有帮助。GLCM Angular second moment、Number of edges 分别位于第2、3名,这表明纹理和几何特征对树种的分类研究也有着一定的帮助。各波段的标准差的重要性较低,排名基本都在后20%。

被淘汰的8个特征分别为4个GLDV、1个标准差特征(Standard Deviation Yellow)、1个几何特征(Border Index)以及2个GLCM特征(GLCM Dis-

similarity 和 GLCM Correlation),其中,GLDV的重要性最低,在第一次的选择中就全部被淘汰。对于GLCM和几何特征而言,特征内重要性差异较大。例如,GLCM Angular second moment 与 GLCM Entropy 是排名前两位的GLCM特征,二者反映了图像中纹理的均匀或复杂程度,即图像灰度分布的均匀性,若灰度共生矩阵内所有值越相近,则越均匀,而淘汰的GLCM Dissimilarity 与 GLCM Correlation 则表示纹理的差异性或一致性,当灰度共生矩阵行或列元素值之间越相近,则一致性越高。这表明对于城市绿化乔木树种分类的纹理特征选取,应选用与均匀性相关的特征。被淘汰的几何特征为 Border index,体现了分割对象的不规则程度,该值越大,对象多边形越扭曲,即与真实树冠的形状相差越大。

4.3 分类的结果

首先利用NIR1波段对水体和阴影进行掩膜,其次通过SAVI对植被进行提取,由于草地和灌木在影像分割中的面积都相对较大,因此内部像元数量较多,利用对象内像元数量(Number of pixels)对其进行掩膜,最终得到绿化乔木占总面积的32.23%,其中各树种占其面积比例见表3。

表3 树种分类的面积和比例统计		
Tab. 3 The statistics of areas and proportion of tree species		
树种	面积/m ²	百分比/%
香樟	57 047.10	30.00
木棉	6693.53	3.52
重阳木	3232.67	1.70
垂叶榕	15 649.92	8.23
杧果	59 462.09	31.27
榕树	33 410.58	17.57
羊蹄甲	14 661.10	7.71

分类结果如图6所示,使用约70%的样本点数据(511个样点)进行精度验证(表4)。由图4可知,当使用34个特征时分类精度最高,总精度为74.95%,kappa系数为0.67。

木棉的分类精度最高,这是由于研究区内木棉长势良好,与其他树种相比树冠明显并且旁边基本没有其他乔木生长。垂叶榕的生产者精度较低,有一部分被误分为杧果,原因是由于研究区内存在大量杧果,占研究区乔木的31.27%,RF会对样本多的类别有一定的倾向性,从而导致分类结果的增多,尽

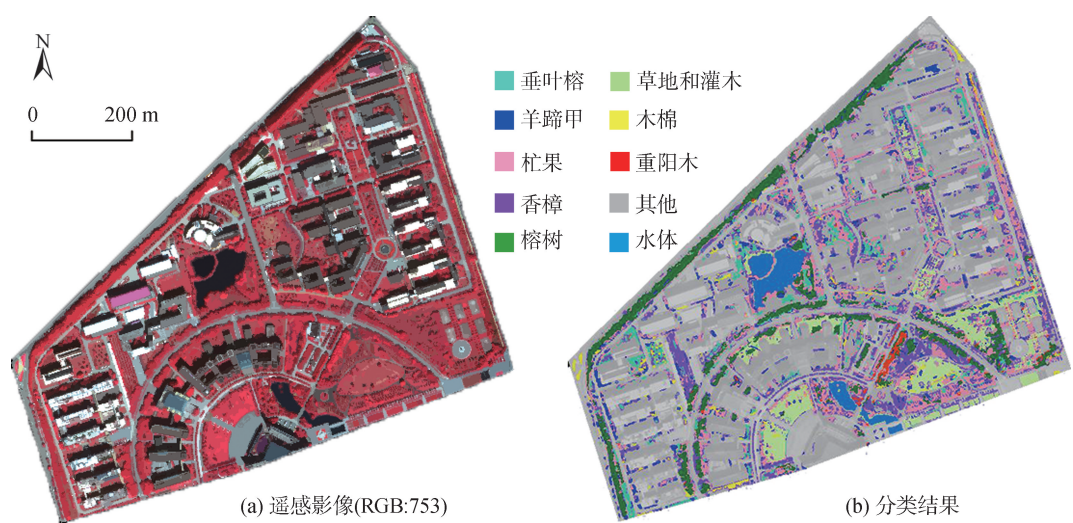


图6 遥感影像与分类结果
Fig. 6 The false color composite image and mapping results of the study area

表4 精度验证
Tab. 4 Accuracy assessment of results

	验证数据							行合计	使用者精度/%
	木棉	榕树	羊蹄甲	香樟	杧果	重阳木	垂叶榕		
木棉	16	0	0	1	0	0	0	17	94.11
榕树	0	53	0	4	2	2	0	61	86.88
羊蹄甲	2	0	38	4	7	0	1	52	73.07
香樟	0	3	5	49	19	3	5	84	58.33
杧果	0	6	6	17	176	3	18	226	77.87
重阳木	0	1	0	0	5	18	0	24	75.00
垂叶榕	0	1	1	4	8	0	33	47	70.21
列合计	18	64	50	79	217	26	57		
生产者精度/%	88.89	82.81	76.00	62.02	81.11	69.23	57.89		
总精度/%	74.95								
Kappa 系数	0.67								

管香樟所占比例也有30%,但由于之前提到的其叶色黄绿,因此,光谱可分性较高,只有少量的垂叶榕被误分为香樟。香樟的生产者精度和使用者精度都比较低,这是因为树木的光谱特征由大量的生化和结构参数所决定,相似的特征可以通过不同的参数组合所产生^[13],且研究区内香樟树龄差异大,由此产生了混淆,对分类结果造成了影响。

5 结论

本研究通过面向对象的方法,对光谱信息丰富的 WorldView-2 卫星影像进行特征提取,对福州大学旗山校区北部片区内种植的榕树、杧果、香樟、重

阳木、羊蹄甲、垂叶榕以及木棉7种主要绿化乔木树种进行分类。通过RF的特征重要性分析对42种特征进行特征选择,对大量特征的重要性进行量化并排序,去除不重要的特征,分析分类精度随不重要特征减少时的变化情况,并研究不同的特征对于分类的重要程度。结果表明,通过特征选择可以简化特征空间,从而明显提高同类型遥感数据树种分类的精度,获取的城市绿化树种信息,可为城市规划和绿化管理提供新的方法和科学依据。

分类结果表明:①利用RF算法进行特征选择明显有助于提高分类精度,当使用34个特征对城市绿化乔木树种进行分类时,精度达到最高的74.95%,Kappa系数为0.67;②特征重要性分析表明,黄光

波段的光谱平均值、GLCM Angular second moment、Number of edges 为重要性排名前三的特征,各波段的标准差的光谱特征排名均在后 20%,GLDV 提取的纹理特征重要性最低,全部被淘汰;③ World-View-2 影像的黄光、红边波段以及新增 4 个波段所构成的比值型指数重要性排名靠前,在植被遥感领域,特别是树种分类的研究中极具应用前景;④ 基于 GLCM 的纹理特征和几何特征存在较大的内部差异,反映纹理均匀或复杂程度的特征重要性较高,但反映一致或差异性的特征重要性却很低,对于城市绿化乔木树种的研究,应选用表征均匀或复杂程度的特征。然而,各种提取特征对树种分类精度的影响机理,还有待于进一步深入探索。

对于高空间分辨率卫星影像的树种分类而言,仍存在着不小的挑战。本文的分类方法对于树龄差异不大的乔木,分类精度较高,而对于树龄相差较大的树种,分类精度偏低。今后的研究中可以考虑进一步按树龄细分训练样本,如幼林和成熟林,以提高分类精度,或加入 LiDAR 数据以获取树木的结构参数特征,如胸径和树高等,用来训练更加精确的估算模型。后续研究中也可以结合实地测量的高光谱数据,完善分类模型,并有助于解析各类光谱特征对分类精度的影响。近年来,深度学习和主动学习的多种算法逐渐应用于遥感领域之中,今后可尝试结合多源遥感数据以及新算法,在更大区域的城市环境中进行绿化树种分类的研究,以获得更大尺度、更高精度的分类方法,作进一步的研究与探讨。

参考文献(References):

- [1] Shang X, Chisholm L A. Classification of Australian native forest species using hyperspectral remote sensing and machine-learning classification algorithms[J]. IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing, 2014,7(6):2481-2489.
- [2] Jensen R R, Hardin P J, Hardin A J. Classification of urban tree species using hyperspectral imagery[J]. Geocarto International, 2012,27(5):443-458.
- [3] 黄慧萍,吴炳方,李苗苗,等.高分辨率影像城市绿地快速提取技术与应用[J].遥感学报,2004,8(1):68-74. [Huang H P, Wu B F, Li M M, et al. Detecting urban vegetation efficiently with high resolution remote sensing data[J]. Journal of Remote Sensing, 2004,8(1):68-74.]
- [4] 李成范,尹京苑,赵俊娟.一种面向对象的遥感影像城市绿地提取方法[J].测绘科学,2011,36(5):112-114. [Li C F, Yin J Y, Zhao J J. An extraction algorithm of urban vegetation from remote sensing image based on object-oriented approach[J]. Science of Surveying and Mapping, 2011, 36(5):112-114.]
- [5] 李登秋,居为民,郑光,等.基于生态过程模型和森林清查数据的森林生长量估算对比研究[J].生态环境学报,2013,22(10):1647-1657. [Li D Q, Ju W M, Zheng G, et al. Comparison of estimated forest biomass increment rate based on a process-based ecological model and forest inventory data[J]. Ecology and Environmental Sciences, 2013,22(10):1647-1657.]
- [6] Korpela I, Ørka H O, Maltamo M, et al. Tree species classification using airborne LiDAR-effects of stand and tree parameters downsizing of training set intensity normalization and Sensor Type[J]. Silva Fennica, 2010,44(2):319-339.
- [7] 宫鹏,浦瑞良,郁彬.不同季相针叶树种高光谱数据识别分析[J].遥感学报,1998,2(3):211-217. [Gong P, Pu R L, Yu B. Conifer species recognition with seasonal hyperspectral data[J]. Journal of Remote Sensing, 1998,2(3): 211-217.]
- [8] 王芳,黎夏,卓莉,等.基于 Hyperion 高光谱数据的城市植被胁迫评价[J].应用生态学报,2007,18(6):1286-1292. [Wang F, Li X, Zhuo L, et al. A stressed level of urban vegetation: Its assessment based on Hyperion hyperspectral data[J]. Chinese Journal of Applied Ecology, 2007,18 (6):1286-1292.]
- [9] Agarwal S, Vailshery L S, Jaganmohan M, et al. Mapping urban tree species using very high resolution satellite imagery: Comparing pixel-based and object-based approaches [J]. ISPRS International Journal of Geo-information, 2013, 2(1):220-236.
- [10] 毕恺艺,牛铮,黄妮,等.基于 Sentinel-2A 时序数据和面向对象决策树方法的植被识别[J].地理与地理信息科学,2017,33 (5):16-27. [Bi K Y, Niu Z, Huang N, et al. Identifying vegetation with decision tree model based on object-oriented method using multi-temporal Sentinel-2A images[J]. Geography and Geo-information Science, 2017,33(5):16-27.]
- [11] 关舒婧,韩鹏鹏,王月如,等.华南地区典型种植园地遥感分类研究[J].地球信息科学学报,2017,19(11):1538-1546. [Guan S J, Han P P, Wang Y R, et al. Study on the classification of typical plantations in South China[J]. Journal of Geo-information Science, 2017,19(11):1538-1546.]
- [12] Dian Y Y, Li Z, Pang Y. Spectral and texture features combined for forest tree species classification with airborne hyperspectral imagery[J]. Journal of the Indian Society of Remote sensing, 2015,43(1):101-107.
- [13] Immitzer M, Atzberger C, Koukal T. Tree species classification with random forest using very high spatial resolution

- tion 8-band WorldView-2 satellite data[J]. *Remote Sensing*, 2012,4(9):2661-2693.
- [14] Pu R L, Landry S. A comparative analysis of high spatial resolution IKONOS and WorldView-2 imagery for mapping urban tree species[J]. *Remote Sensing of Environment*, 2012,124:516-533.
- [15] Naidoo L, Cho M A, Mathieu R, et al. Classification of savanna tree species in the Greater Kruger National Park region by integrating hyperspectral and LiDAR data in a random forest datamining environment[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2012,69(3):167-179.
- [16] 李丹,柯樱海,宫辉力,等.基于高分辨率遥感影像的城市典型乔木树种分类研究[J].*地理与地理信息科学*,2016,32(1):84-89. [Li D, Ke Y H, Gong H L, et al. Urban tree species classification with machine learning classifier using WorldView-2 imagery[J]. *Geography and Geo-information Science*, 2016,32(1):84-89.]
- [17] 刘怀鹏,安慧君,王冰,等.基于递归纹理特征消除的WorldView-2树种分类[J].*北京林业大学学报*,2015,37(8):53-59. [Liu H P, An H J, Wang B, et al. Tree species classification using WorldView-2 images based on recursive texture feature elimination[J]. *Journal of Beijing Forestry University*, 2015,37(8):53-59.]
- [18] Pu R L, Landry S, Yu Q. Object-based urban detailed land cover classification with high spatial resolution IKONOS imagery[J]. *International Journal of Remote Sensing*, 2011,32(12):3285-3308.
- [19] Huete A R. A soil-adjusted vegetation index (SAVI) [J]. *Remote Sensing of Environment*, 1988,25(3):295-309.
- [20] Puissant A, Rougier S, Stumpf A. Object-oriented mapping of urban trees using random forest classifiers[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2014,26(1):235-245.
- [21] Leo Breiman. Random forests[J]. *Machine Learning*, 2001,45(1):5-32.
- [22] 刘凯,汤国安,黄晓力,等.面向地形特征的DEM与影像纹理差异分析[J].*地球信息科学学报*,2016,18(3):386-395. [Liu K, Tang G A, Huang X L, et al. Research on the difference between textures derived from DEM and remote-sensing image for topographic analysis[J]. *Journal of Geo-information Science*, 2016,18(3):386-395.]
- [23] 徐涵秋.城市遥感生态指数的创建及其应用[J].*生态学报*,2013,33(24):7854-7862. [Xu H Q. A remote sensing urban ecological index and its application[J]. *Acta Ecologica Sinica*, 2013,33(24):7854-7862.]
- [24] Blackburn G A. Quantifying chlorophylls and carotenoids at leaf and canopy scales: An evaluation of some hyperspectral approaches[J]. *Remote Sensing of Environment*, 1998,66(3):273-285.