

DOI: 10.20103/j.stxb.202310092169

马金超, 郭振, 许昊, 宁焕杉, 沈家伟, 张志卫. 森林结构复杂度对单木分割精度的影响——以田横岛为例. 生态学报, 2024, 44(11): 4770-4781.

Ma J C, Guo Z, Xu H, Ning H S, Shen J W, Zhang Z W. Effect of forest structure complexity on single wood segmentation accuracy: a case study of Tianheng Island. Acta Ecologica Sinica, 2024, 44(11): 4770-4781.

森林结构复杂度对单木分割精度的影响 ——以田横岛为例

马金超^{1,2}, 郭振^{1,2,*}, 许昊^{1,2}, 宁焕杉¹, 沈家伟², 张志卫²

¹ 山东科技大学测绘与空间信息学院, 青岛 266590

² 自然资源部第一海洋研究所海岸带科学与海洋发展战略研究中心, 青岛 266061

摘要: 单木分割对于森林资源调查具有重要的意义, 不同结构复杂度的森林单木分割算法的选择以及分割参数的选取对分割精度有着很大的影响。以山东田横岛为研究区, 基于无人机正射影像与激光雷达数据, 首先提取海岛森林典型植被二维与三维特征, 然后利用随机森林算法对不同树种的树木进行分类, 最后基于分类后的点云数据, 选取不同结构复杂度的森林样地, 对比分析聚类算法、层堆叠算法、分水岭算法在不同复杂度林区的适用性。结果表明: (1) 随机森林算法结合单木二维、三维特征可有效对混交林树木进行分类, 模型总体的精度为 94.51%, Kappa 系数为 0.9038; (2) 聚类算法对结构简单的林区具有更高的分割精度 ($F=96.41$), 但依赖于分割参数的选取; 面对复杂单木集群, 分水岭算法总体得分波动最小 ($\Delta F=14.56$), 表现出较强的稳定性; (3) 混交林预先进行树种分类可有效改善单木分割环境, 相比于直接进行单木分割, 聚类算法、层堆叠算法、分水岭算法的分割精度均得到不同程度的提升 ($\Delta F_1=10.06, \Delta F_2=9.51, \Delta F_3=12.6$)。

关键词: 海岛森林; 森林结构复杂度; 树种分类; 随机森林; 单木分割

Effect of forest structure complexity on single wood segmentation accuracy: a case study of Tianheng Island

MA Jinchao^{1,2}, GUO Zhen^{1,2,*}, XU Hao^{1,2}, NING Huanshan¹, SHEN Jiawei², ZHANG Zhiwei²

¹ College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao 266590, China

² Coastal Science and Marine Policy Center, First Institute of Oceanography, Ministry of Natural Resources, Qingdao 266061, China

Abstract: The individual tree segmentation is of great significance for forest resource surveys. The accuracy of segmentation is profoundly influenced by the choice of the forest single-wood segmentation algorithm and the parameters associated with different structural complexities. This research introduced drone orthophoto and laser radar data from Tianheng Island. Initially, 2D and 3D characteristics of typical forest vegetation were extracted. Subsequently, the random forest algorithm was applied to classify different tree species. With the classification of point cloud data, sampling plots with varying structural complexities were selected to conduct comparative analysis encompassing clustering algorithms, stacking algorithms, and the watershed algorithm, in order to enhance segmentation accuracy. The findings reveal that: (1) the random forest algorithm, combined with 2D and 3D features of a single wood, effectively classifies the mixed forest trees, achieving an impressive overall accuracy of 94.51% and a Kappa coefficient of 0.9038. (2) The clustering algorithm shows the highest segmentation accuracy for forest areas with simple structures ($F=96.41$), while depending upon the selection of segmentation parameters. In the case of complex single wood clusters, the watershed algorithm displays the least fluctuation

基金项目: 国家自然科学基金项目 (42171292, 42376228); 外交部亚洲专项资金项目 (WJ0923011); 中国海洋发展基金会国际合作项目 (B19029)

收稿日期: 2023-10-09; **网络出版日期:** 2024-03-16

* 通讯作者 Corresponding author. E-mail: guozhen@fio.org.cn

($\Delta F = 14.56$), indicating its superior stability. (3) Pre-classification of tree species in mixed forests effectively improves the single wood segmentation environment. Compared to direct single wood segmentation, the clustering, stacking, and watershed algorithms yield increased segmentation accuracy to varying degrees ($\Delta F_1 = 10.06$, $\Delta F_2 = 9.51$, $\Delta F_3 = 12.6$).

Key Words: island forest; forest structural complexities; classification of different tree species; random forest; single-wood segmentation

海岛森林是维系海岛生态系统物质能量循环以及生物多样性的重要组成部分,相较内陆森林生态系统,海岛陆生植被的生态脆弱性通常更为显著,群落演替过程易受人类活动与陆海多重环境要素的干扰,其群落垂直结构与水平结构的时空动态特征往往更加复杂^[1]。另外,海岛面积相对狭小,在借助遥感手段开展海岛植被群落解译与判读时,传统的卫星遥感影像由于分辨率过低和缺失高度信息,导致不同树种之间常存在“同物异谱”、“异物同谱”以及“冠层重叠”等问题,难以精准地对海岛森林树种进行定位与区分,往往无法满足海岛植被辅助调查和制图的要求,一定程度上影响了我国海岛植被多样性评价以及海岛生态承载力评估工作的开展^[2]。

单木分割是指利用激光雷达点云数据和其他传感器数据来提取单株树木信息的技术,通常首先借助激光雷达点云数据提取树木三维信息,然后对树木点云进行单体化处理。目前,基于 3S 技术的单木分割方法在森林资源调查以及生物量估算等方面发挥着重要作用。通过获取的单木模型可以快速判定各单株树木的生长状况和生长潜力,并可判定采用林分密度控制措施后保留树木的生长状况,这些信息对于林分的集约经营以及森林生态系统健康度具有科研和实用价值。常用的单木分割算法有基于距离判别的聚类算法、层堆叠算法和分水岭算法^[3-5]。基于距离判别的聚类算法通过确立全局最大值为种子点,并以此作为树顶,根据设定的距离阈值逐步区分比种子点低的点,该方法分割高度与距离阈值的设定决定着分割的效果^[6]。层堆叠算法利用归一化点云按照设定的阈值进行分层并对每层进行 K-Means 聚类,迭代重复计算确定种子点的位置,然后根据集群外缓冲区确定最佳冠幅,最终通过平滑堆叠图确定局部最大值进行单木分割,该方法堆叠图分辨率、参数的选取是影响分割效果的主要因素^[7]。分水岭算法则是基于拓扑理论数学形态学原理的一种分割算法,该算法将冠层高度模型(CHM)每个像元的灰度值视作该点的冠层高度,基于每个像元的灰度值确定局域最小值的位置,即集水域。并在集水域的交汇处建立大坝,通过不断迭代建立分水岭实现单木的分割,该方法 CHM 的分辨率影响冠层像元间局域最小值和分水岭的确定^[8]。

相对而言,地形简单树种单一的人工林其单木分割效果较好。而立地环境越复杂,树种越多样,树冠遮挡越严重,单木分割面临的挑战越大。结合人工智能算法,发展广泛适用的单木分割方法以及针对不同林种研发定制化的分割方法是目前该领域的热点研究方向^[9-10]。在数据获取方面,近些年无人机技术与激光雷达遥感技术的飞速发展为海岛森林资源调查提供了新的思路^[11]。无人机搭载激光雷达与多光谱相机绕岛飞行可便捷地获取森林冠层亚米级的光谱影像与雷达数据^[12]。通过正射影像提取林木的二维光谱信息,结合点云数据提取可靠的三维数据与森林冠层数据,使得在海岛狭小且复杂多变的森林类型中提取单木信息成为可能^[13]。

本研究以山东田横岛为研究区,基于无人机倾斜摄影和激光雷达数据构建数字表面模型(Digital Surface Model, DSM),采用随机森林(Random Forest, RF)方法结合二维与三维信息对海岛森林地区不同的树种进行分类^[14],并使用距离判别的聚类方法、层堆叠算法以及分水岭算法对田横岛典型森林样地进行单木分割,探讨海岛森林结构复杂度对单木分割精度的影响以及上述 3 种算法对海岛地区森林的适用性。

1 研究区与数据

1.1 研究区概况

田横岛位于山东青岛即墨东部海域的横门湾中(海岛中心位置 $36^{\circ}25'08''N$, $120^{\circ}57'32''E$),总面积约

1.46 km², 岸线长约 8 km, 海岛东西长度约 3.2 km, 南北约 1.0 km。该地区植被丰富, 气候宜人, 为典型的海岛型气候, 距青岛湾约 68 km。岛上乔木以人工植被为主, 人工栽培主要树种为黑松 (*Pinus thunbergii*) 和刺槐 (*Robinia pseudoacacia*)。

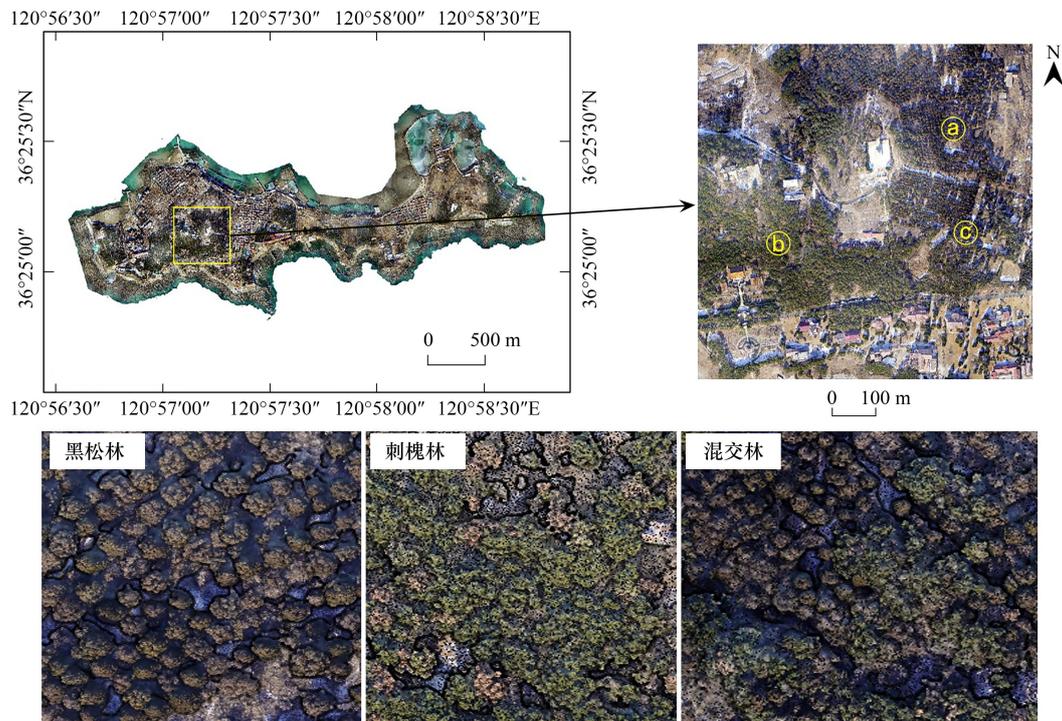


图 1 田横岛不同样地正射影像

Fig.1 The digital orthophoto map of the sampling plots in Tianheng Island

a: 黑松林 Pinus forest; b: 刺槐林 Robinia forest; c: 混交林 Mixed forest

1.2 数据获取

1.2.1 无人机数据

无人机倾斜摄影测量数据与点云数据为 2021 年 1 月于田横岛上由无人机搭载的倾斜摄影相机和激光雷达获得, 飞行高度 100 m, 飞行姿态以五镜头倾角 45° 环绕飞行。点云数据通过密集匹配算法获取, 数据密度为 27 点/m²。正射影像利用 Pix4Dmapper 软件生成, 地面分辨率为 0.1 m, 包含红、绿、蓝三个波段, 数值范围为 0—255。无人机相机具体参数如表 1 所示。

表 1 无人机航拍相机主要参数信息

Table 1 The technical specification of the drone camera

	倾斜摄影 Oblique photography	激光雷达 Lidar
传感器 Sensor	共享 5 镜头斜相机	机载激光移动测量系统
测量范围 Measuring range	25.16 百万像素	1500 m
频率 Revisiting period	0.5/次	200 万点/s

1.2.2 样地数据

研究区地理位置位于 120°57'20"—120°57'35"E 至 36°25'01"—36°25'16"N 范围之内, 为 400 m×400 m 的矩形区域, 区域内地形最大坡度 14°。在研究区内选取群落的组成、结构和生长状况比较典型的森林区域布设样方, 包括黑松、刺槐、黑松-刺槐混交 3 种森林类型, 样地具体参数如表 2 所示。分别对黑松、刺槐两个纯林与黑松-刺槐混交林分类前后的点云数据进行提取, 并按前后顺序依次编号, 记为 L1—L4。

表 2 田横岛森林样地调查信息

Table 2 Survey information of forest sample plots in Tianheng Island

	面积 Area/m ²	单木数量 Counts	最小高度 Minimum tree height/m	最大高度 Maximum tree height/m	平均树高 Mean tree height/m	平均冠幅 Mean crown width/m
黑松林 Pinus forest	1720	85	2.47	7.73	5.29	5.60
刺槐林 Robinia forest	3120	145	2.17	10.06	7.83	6.20
混交林 Mixed forest	2860	176	2.29	7.22	4.79	4.73

2 研究方法

2.1 数据预处理

基于 MicroStation 点云处理平台实现点云数据的去噪、滤波、裁剪等预处理操作,通过渐进加密三角网滤波算法(Improved Progressive TIN Densification, IPTD)分离地面点^[15],获取区域数字高程模型(Digital Elevation Model, DEM)。为去除地形、建筑物以及灌木等地物对后期乔木参数提取的影响,通过计算每个点距离地面点的高度,基于高度阈值提取林地与地面点云。正射影像校正包括数字表面模型生成和图像正射校正。通过调整拼接线、替换区域中的图像和边缘裁剪等手动修改操作^[16],生成地面分辨率为 0.1 m 的正射影像。

森林区域 CHM 由 DSM 与 DEM 做差运算获取。高质量的 CHM 模型对于后期的单木分割的精度具有很大的影响,为减少误差产生,本研究使用 Matlab R2021a 软件在图像中形成密闭的冠层区域对 CHM 表面进行了形态学“开”运算。即在 CHM 中将每个栅格单元值替换为其邻域内激光点云高度最大值,得到树冠最大模型(Crown Max Model, CMM),并通过高斯滤波对 CMM 进行平滑处理以降低模型噪点引起树冠顶点的误判率。

2.2 特征提取指标体系构建

基于无人机正射影像和激光雷达数据选取黑松、刺槐两种北方海岛典型单木冠幅各 200 个样本,利用 ENVI 5.6 波段运算及图像滤波功能计算获取各树种单木的二维特征,通过 Lidar 360 统计森林参数模块计算获取各树种单木三维特征^[17]。具体内容如下。

2.2.1 二维特征提取

结合不同树种光谱曲线中的特定波段反射特征,参考当前广泛用于 UAV 生物量检索的植被指数^[18-21],选取归一化绿红差异指数(NGRDI)、归一化绿蓝差异指数(NGBDI)、可见大气阻力指数(VARI)、可见差异植被指数(VDVI)和过剩绿色指数(EXG)共 5 个可见光植被指数(表 3)。

表 3 二维与三维特征指数选取

Table 3 The 2D and 3D feature indices

指数类型 Type of indicator	指数名称 Indicators	计算公式 Formula
光谱特征指数 Spectroscopic feature index	归一化绿红差异指数(NGRDI)	$NGRDI = (DNG - DNR) / (DNG + DNR)$
	归一化绿蓝差异指数(NGBDI)	$NGBDI = (DNG - DNB) / (DNG + DNB)$
	可见大气阻力指数(VARI)	$VARI = (DNG - DNR) / (DNG + DNR - DNB)$
	可见差异植被指数(VDVI)	$VDVI = (2DNG - DNR - DNB) / (2DNG + DNR + DNB)$
	过剩绿色指数(EXG)	$EXG = (2DNG - DNR) / (DNG + DNR + DNB)$
纹理特征指数 Texture feature index	均值(TM)	$TM_i = \sum_i P[i, j]$
	均匀度(Hom)	$Hom = \sum_i \sum_j \frac{P[i, j]}{1 + (i - j)^2}$
	异质性(Dis)	$Dis = \sum_i \sum_j P[i, j]$

续表

指数类型 Type of indicator	指数名称 Indicators	计算公式 Formula
	相关性(Cor)	$\text{Cor} = \frac{\sum_i \sum_j ijP[i,j] - TM_i TM_j}{\text{Var}_i \text{Var}_j}$
	方差(Var)	$\text{Var}_i^2 = \sum_i i^2 P[i,j] - TM_i^2$
	熵(Ent)	$\text{Ent} = - \sum_i \sum_j P[i,j] (\ln P[i,j])$
	对比度(Con)	$\text{Con} = \sum_i \sum_j P(i-j)^2 P[i,j]$
高度特征指数	高度百分位(H_P)	$H_P = \left(\frac{P}{100}\right) \times (n+1)$
Highly characteristic index	点云高度峰度(H_K)	$H_K = \frac{\frac{1}{n-1} \sum_{i=1}^n (Z_i - \bar{Z})^4}{\sigma^4}$
	点云高度偏度(H_S)	$H_S = \frac{\frac{1}{n-1} \sum_{i=1}^n (Z_i - \bar{Z})^3}{\sigma^3}$
	点云高度均值(H_{mean})	$H_{\text{mean}} = \text{Mean}(Z_i)$
	点云高度最大值(H_{max})	$H_{\text{max}} = \text{Max}(Z_i)$
	冠层起伏率(H_{Crr})	$H_{\text{Crr}} = \frac{\text{Mean} - \text{Min}}{\text{Max} - \text{Min}}$

NGRDI:归一化绿红差异指数 Normalized green-red difference index; NGBDI:归一化绿蓝差异指数 Normalized green-blue difference index; VARI:可见大气阻力指数 Visible atmospherically resistant index; VDVI:可见差异植被指数 Visible differential vegetation index; EXG:过剩绿色指数 Excess green index; TM :均值 Textural Mean; Hom:均匀度 Homogeneity; Dis:异质性 Dissimilarity; Cor:相关性 Correlation; Var:方差 Variance; Ent:熵 Entropy; Con:对比度 Contrast; H_P :高度百分位 Height percentile; H_K :点云高度峰度 Kurtosis of point cloud height; H_S :点云高度偏度 Skewness of point cloud height; H_{mean} :点云高度均值 Mean of point cloud height; H_{max} :点云高度最大值 Maximum of point cloud height; H_{Crr} :冠层起伏率 Canopy relief ratio; DNG:红波段像元值 Red band digital number; DNB:蓝波段像元值 Blue band digital number; DNG:绿波段像元值 Green band digital number; P_{ij} :归一化共生矩阵 The normalized symbiosis matrix; P :点云高度值 Point cloud height value; n :统计单元内点云总个数 Statistics of the total number of point clouds in the unit; σ :统计单元内点云高度分布标准差 Standard deviation of point cloud height distribution within a statistical unit; Z_i :统计单元内第 i 个点的高度值 Height values of the i -th point within the statistical unit; \bar{Z} :统计单元内所有点的平均高度 Mean height of all points within the statistical unit; Mean:高度均值 Height mean; Max:高度最大值 Height maximum; Min:高度最小值 Height minimum

灰度共生矩阵(GLCM)方法因其旋转不变性和多尺度特性而被广泛应用于提取图像纹理信息^[22]。本研究采用灰度共生矩阵方法,对样地的均值(TM)、均匀度(Hom)、异质性(Dis)、相关性(Cor)、方差(Var)、熵(Ent)和对比度(Con)共7个纹理特征指标进行了提取计算(表3)。

2.2.3 三维特征提取

在忽略冠层重叠影响的情况下,点云数据相比于正射影像,能更精细地重建树木的三维结构,并对不同树种的群落结构进行统计性描述^[23]。本研究参考基本的点云结构指标,选取包括点云高度百分位(H_{P50} 、 H_{P75} 、 H_{P95} 、 H_{P99})、25%—75%累计高度百分位(H_{IQ})、高度峰度(H_K)、高度偏度(H_S)、高度均值(H_{mean})、高度最大值(H_{max})、冠层起伏率(H_{Crr})共10个高度特征指标^[24](表3)。

2.3 基于随机森林的树种分类方法

随机森林是一种由多个决策树模型集成的分类器,对异常值和噪声有很强的容忍度,能够有效地预测分析高维、非线性或具有交互作用的数据^[25]。分类问题中最常使用的结合策略是投票法(Voting),其基于重抽样对每个样本进行决策树建模并组合多棵决策树的预测结果,最终采用投票的方式选取票数最多类别的作为最终的预测结果。投票法计算原理如公式(1)所示:

$$H(x) = \begin{cases} c_j, & \sum_{i=1}^T h_i^j(x) > 0.5 \sum_{k=1}^N \sum_{i=1}^T h_i^k(x) \\ \text{reject, 其他} & \end{cases} \quad (1)$$

式中, $H(x)$ 表示预测类别, $\{c_1, c_2, \dots, c_N\}$ 为类别集合, h_i 在样本 x 上的预测输出表示为 N 维向量 $(h_i^1(x), h_i^2(x), \dots, h_i^N(x))^T$, $h_i^j(x)$ 表示 h_i 在类别 c_j 上的输出, reject 表示拒绝预测。即若某个标记得票过半数, 则预测为该类别, 否则拒绝预测。

针对过多的特征易导致 Hughes 现象产生, 致使树种分类精度下降的问题, 本研究使用袋外误差法 (Out-of-bag error, OBB) 对提取的特征进行重要性排序, 剔除重要性低的特征, 并根据筛选后的特征和袋外误差选取最优的特征组合。袋外误差法计算如公式 (2) 所示:

$$I_i = \sum (\text{err}_{\text{ooB}_2} - \text{err}_{\text{ooB}_1}) / N \quad (2)$$

式中, I_i 为第 i 个特征的重要程度, $\text{err}_{\text{ooB}_1}$ 为特征 i 的袋外误差, $\text{err}_{\text{ooB}_2}$ 为特征 i 加入高斯噪声或随机重排后计算得到的袋外误差^[26], N 为随机森林中树的个数。

2.4 单木分割及精度评估方法

聚类算法分割精度主要与设置的距离阈值以及分割高度有关, 不同的参数设置会影响分割的精度。为探究森林结构复杂度对该算法精度的影响, 本研究 L1—L4 样地距离阈值统一设置为 1 m, 最小树高设置为 2 m。

层堆叠算法精度与栅格分辨率以及分层厚度的选取有较大的关系, 经对比分析, 栅格分辨率过低, 会出现明显的过分割情况, 随着栅格分辨率的增长, 识别精度得到了较大的改善。研究最终选取 1.5 m×1.5 m 分辨率的栅格, 层厚度设置为 1 m, 最小树间距设置为 1 m, 并采用半径为 5、平滑因子为 1.5 的高斯滤波对样地进行分割。

分水岭算法主要基于冠层高度模型的精度, 分辨率越高, 模型越精确, 但分辨率太高往往会产生大量的无效值, 导致大量错误识别的种子点^[27]。但相对于其他两种算法, 该算法可对低矮的植被进行较准确的识别。经反复实验, 本研究最终选用 1.5 m×1.5 m 分辨率的栅格影像, 利用半径为 5, 平滑因子为 2.5 的高斯滤波对样地进行分割。

本研究利用 Lidar 360 中的机载林业模块实现聚类算法、层堆叠算法及分水岭算法的参数选取及样地的单木分割。并采用式 (3) 中的三个指标衡量分割精度:

$$\begin{cases} Ad = Nc/Nd \times 100\% \\ Ar = Nc/Nr \times 100\% \\ F = 2AdAr / (Ar + Ad) \end{cases} \quad (3)$$

式中: Ad 表示准确率, Ar 表示识别率, F 表示 F 测度, Nc 表示被正确分割的树冠个数, Nr 表示参考图中树冠个数总数, Nd 代表自动分割结果中树冠个数的总数。

3 结果与分析

3.1 特征指标重要性排序与树种分类

本研究中随机森林分类及特征选取采用 Python 3.7 中的 sklearn 模块进行实现。对提取的 23 种特征进行重要性排序 (图 2), 选取排名前 10 位的特征对研究区内森林样地进行分类。基于选取后的特征, 随机筛选 80% 各树种单木样地样本点进行训练, 剩余 20% 样本作为验证集, 基于验证数据建立混淆矩阵对点云的分类结果进行评价。最终模型总体的分类精度为 94.51%, Kappa 系数为 0.9038, 可将大部分不同树种的树木进行分离。针对冠幅的遮挡以及 CHM 噪声所致的树木分类错误进行人工修正, 最终分类效果如图 3 所示。

3.2 3 种算法单木分割结果比较

为分别统计单木分割算法 3 种算法在不同样地的正确分割的树冠个数 Nc , 样地实际树冠个数 Nr , 自动分割结果中树冠个数的总数 Nd ^[28], 以探究不同算法在不同复杂度海岛林地的适用性, 最终结果如图 3 所示。

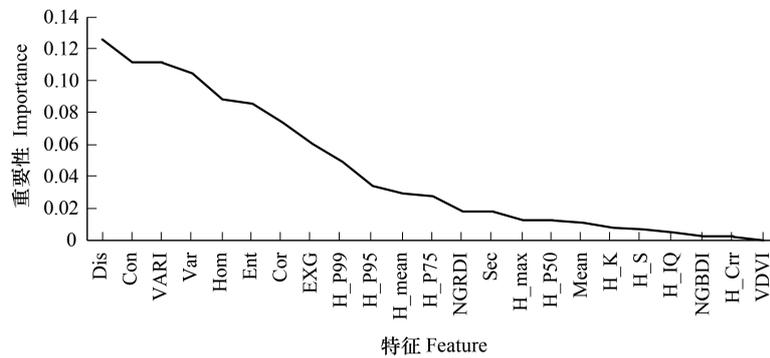


图 2 23 种特征重要性排序

Fig.2 23 kinds of feature importance ranking

NGRDI:归一化绿红差异指数;NGBDI:归一化绿蓝差异指数;VARI:可见大气阻力指数;VDVI:可见差异植被指数;EXG:过剩绿色指数;TM:均值;Hom:均匀度;Dis:异质性;Cor:相关性;Var:方差;Ent:熵;Con:对比度;Sec:二阶矩;H_P:高度百分位;H_K:点云高度峰度;H_S:点云高度偏度;H_mean:点云高度均值;H_max:点云高度最大值;H_Crr:冠层起伏率;H_IQ:累计高度百分位

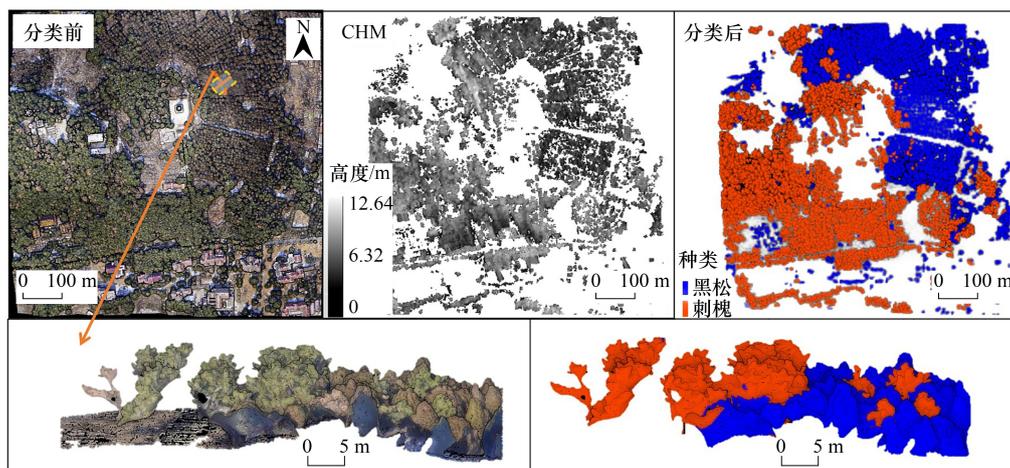


图 3 研究区树种分类流程图

Fig.3 Flow chart of tree species classification in the study area

CHM:冠层高度模型 Crown height modal

如图 4 所示,通过聚类算法,黑松林样地中,85 棵黑松中有 79 棵被正确识别,存在少数错误分割的情况主要是由于部分树木矮小未被识别到,总体识别率为 92.94%。层堆叠算法识别到 70 棵黑松,正确识别 69 棵,分割准确度较高,但识别率略低于聚类算法,为 81.18%,部分树木由于距离太近而出现欠分割的情况,分割效果可通过减小距离参数进行改善;分水岭算法识别过程中,由于灌木和噪声的影响,出现了过分割的情况,部分单木被识别成多颗树木,导致 85 棵单木最终识别到 93 棵树木,但该算法对矮小树木的识别效果较好,总体识别率可达到 90.59%,后期可通过提高 CHM 的模型质量来提高分割的精度。

在刺槐林中,聚类算法 145 棵树木识别到 101 棵单木,但正确识别仅 83 棵,由于树冠之间的相互遮挡,树顶点很难确定,导致欠分割的情况,识别率仅为 57.24%;层堆叠算法在刺槐林样地识别中表现出与聚类算法同样的问题,仍存在大量错误识别现象。但由于其分层聚类的特点,树顶点的识别情况优于聚类算法,面对复杂的林区表现出更强的稳定性,134 棵单木中有 103 棵被正确识别,识别率为 71.03%;分水岭算法相较于其他两种算法,对树顶点要求较低,该算法根据 CHM 像元灰度值确定局部最小值生成分水岭分离单木,145 棵树

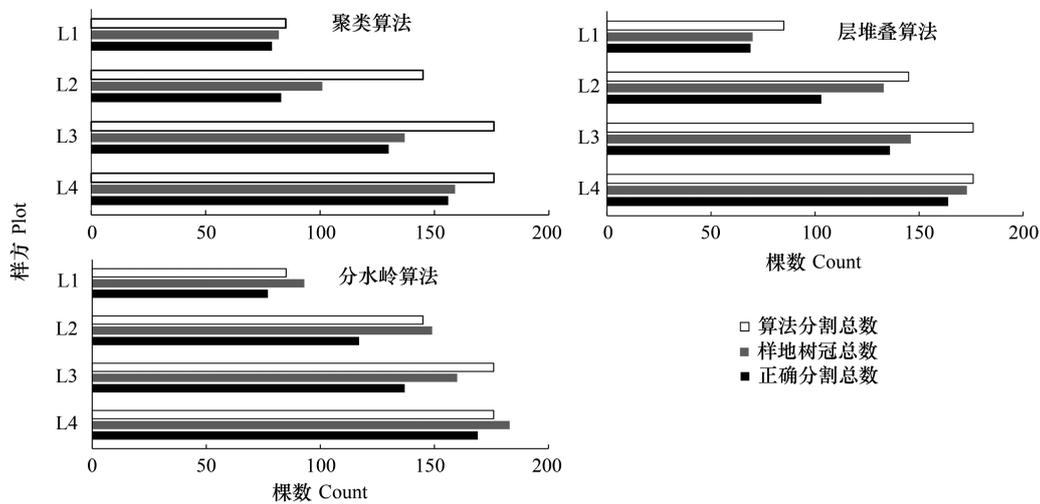


图 4 3 种单木算法分割结果比较

Fig.4 Comparison of Results of Three Single Wood Segmentation Algorithms

L1:黑松林 Pinus forest;L2:刺槐林 Robinia forest;L3:混交林(分类前) Mixed forest (Before classification);L4:混交林(分类后) Mixed forest (After Classification)

木被分割成 149 棵,由于模型噪点的存在导致少量的过分割情况。纯刺槐样地的识别率可为 80.69%,相较于聚类算法和层堆叠算法识别率分别提升 23.45%、9.66%。

混交林直接进行单木分割,3 种算法差异并不明显,176 棵树中均有约 130 棵可以正确识别,但不同算法对黑松与刺槐识别数量各不相同。如聚类算法对于混交林中黑松的识别率相较于在纯黑松样地下降 3.21%,但对于混交林中刺槐的识别率较于在纯刺槐样地提升了 9.52%;层堆叠算法与分水岭算法对于混交林中黑松的识别率分别为 84.33% 和 82.09%,相较于纯黑松林波动比较小,但刺槐识别率均下降了 20%左右,聚类算法、层堆叠算法和分水岭算法对黑松、刺槐的总体识别率分别为 73.86%、77.27% 和 77.84%。优先对混交林树木进行分类可有效的避免树冠遮挡与参数选取的问题,分割前预先进行树种分类,聚类算法、层堆叠算法和分水岭算法识别数相比于直接分割分别增长 26、28、32 棵,最终总体识别率均可达到 85% 以上。聚类算法对 4 种样地单木分割的结果如图 5 所示。

3.3 3 种分割算法精度评估

3 种算法在不同样地分割精度如表 4 所示。可以看出 3 种单木分割算法对于黑松样地,分割准确度相差不大,集中在 0.8—0.95 之间浮动,总体上,聚类算法相较于其他两种算法精度更高,但对于不同复杂度的森林样地,总体得分波动较大;分水岭算法在简单林区分割中总体得分均为 80 左右,略低于其他两种算法,但面对复杂的森林环境(L2)时表现出更强的稳定性($F=79.59$),对低矮的树木也有较好的识别效果;层堆叠算法对于结构特征明显的黑松林具有最高的分割准确度($Ad=98.57%$),但识别率较低($Ar=81.18%$),总体表现介于聚类算法与分水岭算法之间。

4 讨论

黑松属常绿针叶乔木,外部特征明显且林分相对简单,较少存在冠幅遮挡情况,在本研究中森林复杂度最低,3 种分割算法在黑松林中总体识别精度均在 80% 以上(图 6)。刺槐林属落叶阔叶乔木,该样地类型虽然树种单一,但由于枝叶较密集,树冠间遮挡相对较重,无人机航拍无法获取树冠下详细结构,森林结构相较于针叶林更加复杂。聚类算法虽然识别精度可以达到 80%,但存在大量错分漏分的情况,总体得分仅有 67.5。分水岭算法对于低矮的植被有更好的识别效果,总体得分最高为 79.59。

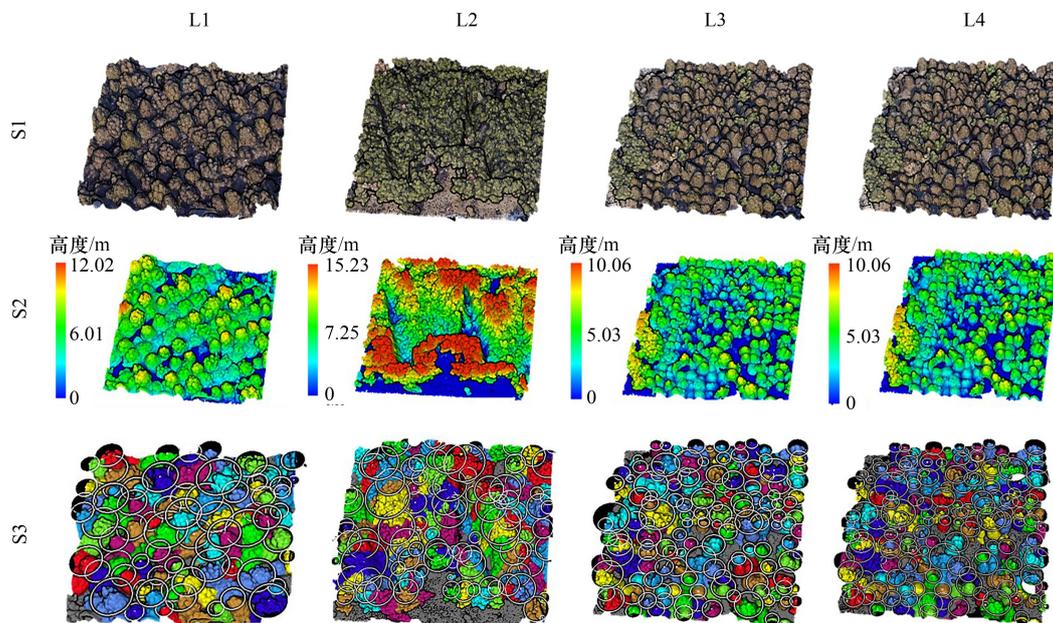


图5 聚类算法4种样地单木分割结果

Fig.5 Single wood segmentation results of 4 plots of clustering algorithm

S1:原始数据 Raw data;S2:高度模型 Height model;S3:分割结果 Split the results

表4 3种分割算法精度对比

Table 4 Comparison of the accuracy of the three segmentation algorithms

	L1	L2	L3	L4
聚类算法 Clustering Algorithm	$Ad=96.34\%$ $Ar=92.94\%$ $F=96.41$	$Ad=82.18\%$ $Ar=57.24\%$ $F=67.50$	$Ad=94.89\%$ $Ar=73.86\%$ $F=83.07$	$Ad=98.11\%$ $Ar=88.64\%$ $F=93.13$
层堆叠算法 Layer Stacking	$Ad=98.57\%$ $Ar=81.18\%$ $F=89.03$	$Ad=77.44\%$ $Ar=71.03\%$ $F=74.10$	$Ad=93.15\%$ $Ar=77.27\%$ $F=84.47$	$Ad=94.80\%$ $Ar=93.18\%$ $F=93.98$
分水岭算法 Watershed Algorithm	$Ad=82.80\%$ $Ar=90.59\%$ $F=86.51$	$Ad=78.52\%$ $Ar=80.69\%$ $F=79.59$	$Ad=85.63\%$ $Ar=77.84\%$ $F=81.55$	$Ad=92.35\%$ $Ar=96.02\%$ $F=94.15$

 Ad : 准确率; Ar : 识别率; F : 总体得分

混交林样地中包括黑松、刺槐两种植被群落,但黑松的数量明显高于刺槐,属针叶林为主的针阔混交林,虽然部分区域存在遮挡的情况,但单木轮廓相对明显,森林复杂度介于针叶林与阔叶林之间。由于存在两种树种的混杂,若直接对混交林进行单木分割,分割参数的选取相较纯林难度更高。研究表明,混交林直接进行单木分割,3种算法总体得分均在80左右,且刺槐的识别率低于黑松的识别率,可能是由于分割参数的设置更适合黑松所致。而预先对混交林进行树种分类,并根据黑松与刺槐不同的结构选取不同的分割参数,聚类算法、层堆叠算法和分水岭算法3种算法的总体得分分别增长了10.06、9.51、12.6,总体得分均达到90以上。

分别统计各样地黑松刺槐的数量并计算总体得分情况(图7),可以发现分类后的混交林黑松的识别率基本与纯林持平,相比于直接进行分割3种算法总体得分分别增长10.67、5.51、9.51。纯林中聚类算法对刺槐的识别精度最低,但由于混交林中刺槐树木分布相比于纯林更加分散,直接进行单木分割聚类算法对刺槐的识别率得到了提升,且高于其他两种算法,说明聚类算法对边界明显的树木具有更好的识别效果。树种分类后的混交林刺槐的总体得分有较大的提升,3种算法分别增长7.67、23.61、21.99。经比较,聚类算法对黑松总体识别率最高,但层堆叠算法更为稳定,分水岭算法对刺槐的识别情况最好。

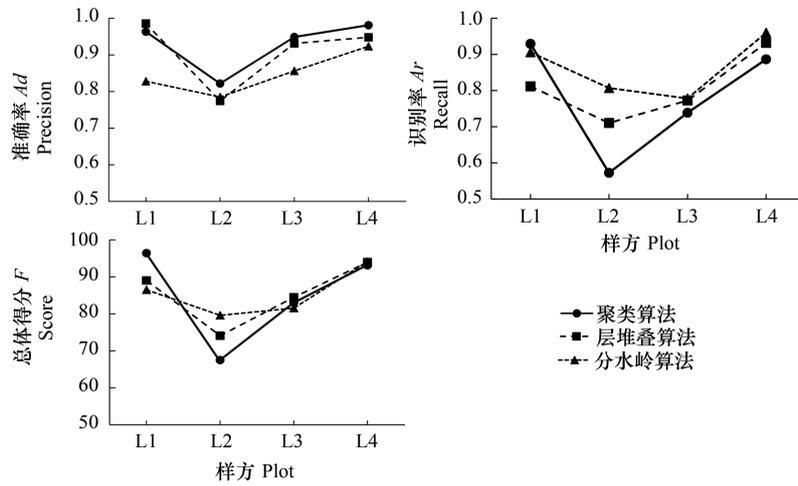


图 6 Ad、Ar、F 折线图

Fig.6 Ad, Ar and F line plots

由上述分析可知,四种森林样地的结构复杂度 $L2 > L3 > L4 > L1$, 随森林复杂度的提高,单木分割精度会受到较大的影响;预先对混交林进行树种分类可改善分割的环境,逐步对单一树种分割最终进行叠加,可改善遮挡、漏分、错分等情况,提高单木分割的精度。

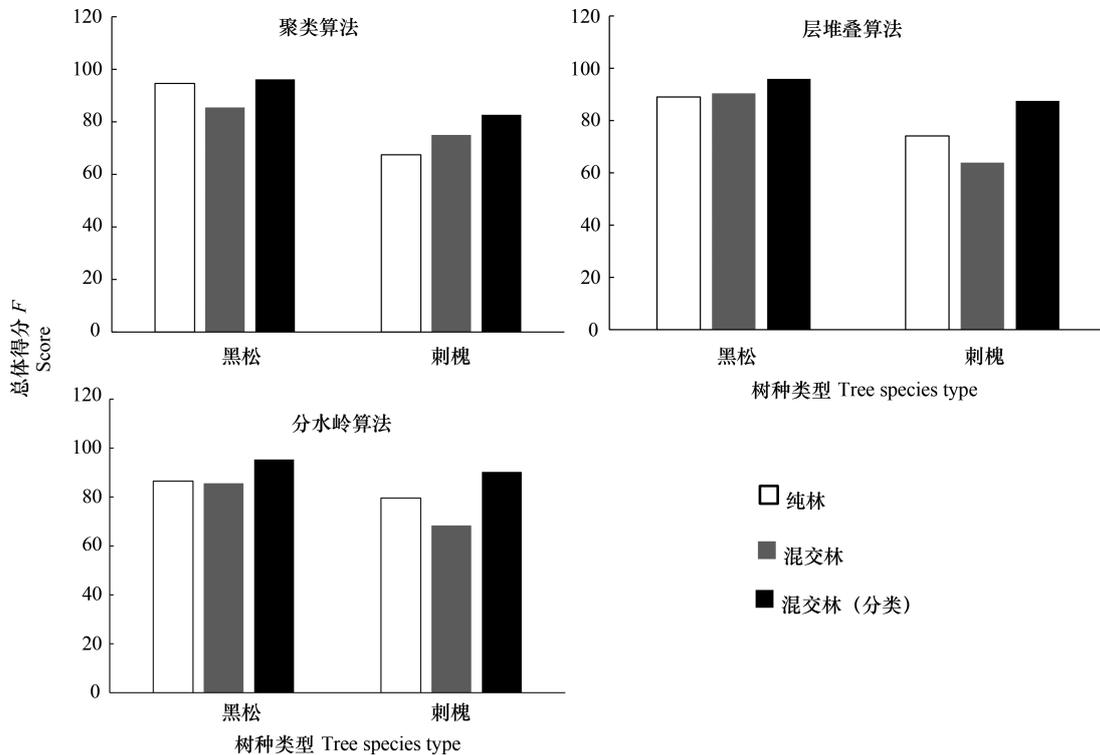


图 7 不同复杂度样地分割精度对比

Fig.7 Comparison of segmentation accuracy of different complexity samples

田横岛乔木以 20 世纪 60 年代人工栽培为主,林相相对整齐且种植具有一定的规律,但由于针叶林与阔叶林有着不同的生长习性,在后期的生长中两种林地具有一定的差异性^[29]。张俊艳等^[30]的研究表明,阔叶林森林结构复杂度大于针阔混交林大于针叶林,且单木分割的精度会随森林结构复杂度的增加而降低,本研

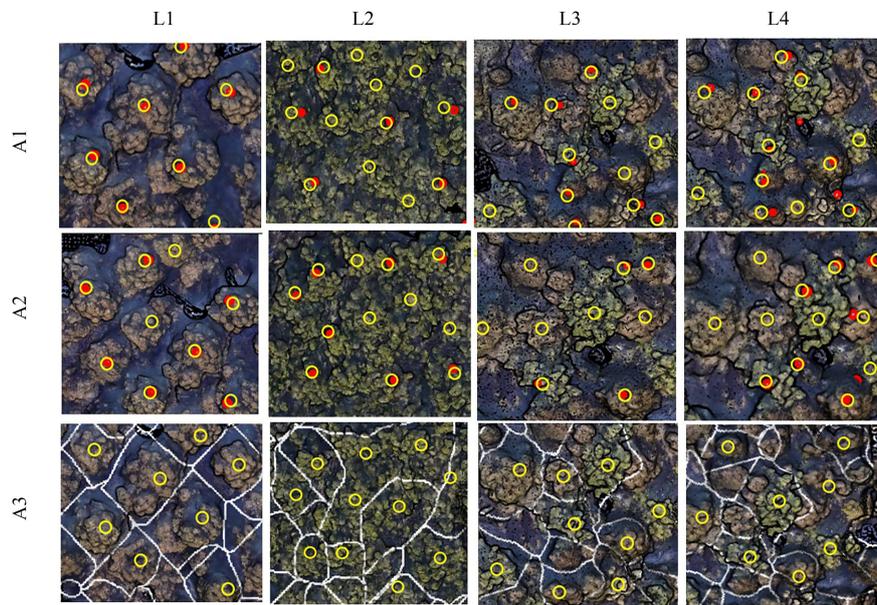


图 8 4 种样地 3 种单木分割算法精度对比

Fig.8 Comparison of accuracy of 3 single-wood segmentation algorithms in 4 plots

A1: 聚类算法 Clustering algorithm; A2: 层堆叠算法 Layer stacking; A3: 分水岭算法 Watershed algorithm; 黄色点: 树顶点真实位置 The true location of the tree vertice; 红色点: 算法识别树顶位置 The algorithm identifies the tree top location; 白色网格: 构建分水岭位置 Build watershed locations

究也显示出相同的规律。胡迎香等^[31]对比 3 种分割算法对亚热带针叶林的分割情况,证明分水岭算法对单木集群具有最高的分割精度。但在本研究中,图 6 显示在针叶林样地中聚类算法精度最高,原因可能是由于自然林较人工林在森林结构上更加复杂以及分割参数选取不同所致。但随森林复杂度的提升,分水岭算法相比于其他两种算法表现出更强的精确性与稳定性(图 8)。

本研究中 3 种单木分割算法对于阔叶林地区分割精度普遍偏低,后续可利用地基雷达实现冠层以下森林结构的扫描,结合无人机航拍点云数据和结合多光谱、高光谱、微波等传感器,可完成对其他地物更准确的分类和获取更为完整森林地区的三维结构信息^[32]。结合树种分类、单木分割等方法,获取精确的单木树高、胸径等参数,根据不同树种的异速生长方程计算单木生物量,将为海岛森林生物量、碳储量的调查提供方法。

5 结论

本研究在田横岛选取 3 块森林典型样地,对比分析了 3 种单木分割算法在不同复杂度样地的适用性,得出以下结论:(1)通过无人机正射影像与激光雷达数据可获得高分辨率海岛森林的冠层信息,结合单木二维、三维特征结合随机森林算法可有效对混交林中不同树种的树木进行分类。本研究中分类模型总体的精度为 94.51%,Kappa 系数为 0.9038(2)在结构简单的黑松林中,3 种单木分割算法均具有较高的精度,其中聚类算法总体精度最高($F=96.41$),其次为层堆叠算法($F=89.03$),分水岭算法精度最低($F=86.51$),3 种算法整体分割精度均在 80%以上;分水岭算法对复杂森林环境的适应能力最强,在刺槐林中大部分密集树冠被正确分割($F=79.59$),聚类算法($F=67.50$)和层堆叠算法($F=74.10$)存在部分错分与漏分的现象。(3)预先通过树木间不同的光谱与空间特征进行树种分类,可有效改善复杂林区的单木分割环境,根据分类后的数据提取混交林内单一树种点云,选取合适的分割参数逐步对不同的树种分割,最终叠加获取全部林区内的单木信息。相比于直接对混交林进行分割,该方法可改善灌木与高大乔木对其他树木遮挡以及参数选取的问题,聚类算法、层堆叠算法、分水岭算法对树木的分割精度均得到了不同程度的提高($\Delta F_1=10.06$, $\Delta F_2=9.51$, $\Delta F_3=12.6$)。

参考文献 (References):

- [1] 李秉璇, 周冰, 贺宣, 刘贺雄. 针对高光谱图像的目标分类方法现状与展望. *激光与红外*, 2020, 50(3): 259-265.
- [2] 苑全治, 吴炳方, 张磊, 李晓松, 曾源. 辅助数据在面向对象分类方法中的应用——以密云水库上游为例. *生态学报*, 2014, 34(24): 7202-7209.
- [3] Chen Q, Baldocchi D, Gong P, Kelly M. Isolating individual trees in a savanna woodland using small footprint lidar data. *Photogrammetric Engineering & Remote Sensing*, 2006, 72(8): 923-932.
- [4] Reitberger J, Schnörr C, Krzystek P, Stilla U. 3D segmentation of single trees exploiting full waveform LIDAR data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2009, 64(6): 561-574.
- [5] Ayrey E, Fraver S, Kershaw J A Jr, Kenefic L S, Hayes D, Weiskittel A R, Roth B E. Layer stacking: a novel algorithm for individual forest tree segmentation from LiDAR point clouds. *Canadian Journal of Remote Sensing*, 2017, 43(1): 16-27.
- [6] 杨秋丽. 机载 LiDAR 单木分割方法对比及其影响因素分析[D]. 乌鲁木齐: 新疆大学, 2018.
- [7] Bayati H, Najafi A, Vahidi J, Gholamali Jalali S. 3D reconstruction of uneven-aged forest in single tree scale using digital camera and SfM-MVS technique. *Scandinavian Journal of Forest Research*, 2021, 36(2/3): 210-220.
- [8] 段祝庚, 曾源, 赵旦, 吴炳方, 赵玉金, 朱建军. 机载激光雷达森林冠层高度模型凹坑去除方法. *农业工程学报*, 2014, 30(21): 209-217.
- [9] 李平昊, 申鑫, 代劲松, 曹林. 机载激光雷达人工林单木分割方法比较和精度分析. *林业科学*, 2018, 54(12): 127-136.
- [10] 李岩, 史泽林, 程坤, 甄贞. 运用激光雷达数据的单木树冠提取算法对帽儿山林场单木参数估测的影响. *东北林业大学学报*, 2019, 47(11): 59-65.
- [11] 刘新平, 姜帅, 程进明. 机载雷达点云在森林资源林分因子测算与评估中的应用. *现代测绘*, 2021, 44(5): 40-44.
- [12] 解宇阳, 王彬, 姚扬, 杨琅, 高媛, 张志明, 林露湘. 基于无人机激光雷达遥感的亚热带常绿阔叶林群落垂直结构分析. *生态学报*, 2020, 40(3): 940-951.
- [13] 谢天义, 潘洁, 孙玉琳, 郑光. 应用多维遥感信息协同的森林树种分类. *东北林业大学学报*, 2023, 51(3): 73-78.
- [14] 王效盖, 王健, 刘翔宇, 曹一. 改进的渐进加密三角网机载点云滤波方法. *北京测绘*, 2023, 37(2): 178-184.
- [15] Tian Y C, Huang H, Zhou G Q, Zhang Q, Tao J, Zhang Y L, Lin J L. Aboveground mangrove biomass estimation in Beibu Gulf using machine learning and UAV remote sensing. *Science of the Total Environment*, 2021, 781: 146816.
- [16] 李哲, 张沁雨, 邱新彩, 彭道黎. 基于高分二号遥感影像树种分类的时相及方法选择. *应用生态学报*, 2019, 30(12): 4059-4070.
- [17] 李桂林, 王红, 宋音, 杨世琪. 黄河三角洲孤岛刺槐林地上生物量的估算——基于无人机机载激光雷达. *中国农学通报*, 2018, 34(26): 52-57.
- [18] Meyer G E, Neto J C. Verification of color vegetation indices for automated crop imaging applications. *Computers and Electronics in Agriculture*, 2008, 63(2): 282-293.
- [19] Verrelst J, Schaepman M E, Koetz B, Kneubühler M. Angular sensitivity analysis of vegetation indices derived from CHRIS/PROBA data. *Remote Sensing of Environment*, 2008, 112(5): 2341-2353.
- [20] Gitelson A A, Kaufman Y J, Stark R, Rundquist D. Novel algorithms for remote estimation of vegetation fraction. *Remote Sensing of Environment*, 2002, 80(1): 76-87.
- [21] Torres-Sánchez J, Peña J M, de Castro A I, López-Granados F. Multi-temporal mapping of the vegetation fraction in early-season wheat fields using images from UAV. *Computers and Electronics in Agriculture*, 2014, 103: 104-113.
- [22] Haralick R M, Shanmugam K, Dinstein I. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 1973, SMC-3(6): 610-621.
- [23] Zhang Z Y, Kazakova A, Moskal L, Styers D. Object-based tree species classification in urban ecosystems using LiDAR and hyperspectral data. *Forests*, 2016, 7(12): 122.
- [24] Wang D Z, Wan B, Liu J, Su Y J, Guo Q H, Qiu P H, Wu X C. Estimating aboveground biomass of the mangrove forests on northeast Hainan Island in China using an upscaling method from field plots, UAV-LiDAR data and Sentinel-2 imagery. *International Journal of Applied Earth Observation and Geoinformation*, 2020, 85: 101986.
- [25] 熊艳, 高仁强, 徐战亚. 机载 LiDAR 点云数据降维与分类的随机森林方法. *测绘学报*, 2018, 47(4): 508-518.
- [26] Huang N T, Lu G B, Xu D G. A permutation importance-based feature selection method for short-term electricity load forecasting using random forest. *Energies*, 2016, 9(10): 767.
- [27] 张丽, 王健, 曲相屹, 邵为真. 机载激光点云单木分割方法对比及精度分析. *测绘与空间地理信息*, 2023, 46(5): 34-37, 42.
- [28] Li W K, Guo Q H, Jakubowski M K, Kelly M. A new method for segmenting individual trees from the lidar point cloud. *Photogrammetric Engineering & Remote Sensing*, 2012, 78(1): 75-84.
- [29] 王文杰, 王凯, 王媛媛, 温慧, 于景华, 王庆贵, 韩士杰. 东北针叶林与阔叶林乔灌草组成特征及碳汇功能对比研究. *北京林业大学学报*, 2022, 44(10): 52-67.
- [30] 张俊艳. 海南岛热带天然针叶林——阔叶林交错区的群落特征研究[D]. 北京: 中国林业科学研究院, 2014.
- [31] 胡迎香, 高红旗, 夏万求, 黄其欢, 陈志欣, 王德柱. 机载雷达点云亚热带针叶林单木分割探究. *应用激光*, 2021, 41(6): 1301-1309.
- [32] 艾萨迪拉·玉苏甫, 玉米提·哈力克, 阿不都拉·阿不力孜, 买尔当·克依木. 基于地面 LiDAR 数据的塔里木河下游胡杨林结构参数反演. *生态学报*, 2020, 40(13): 4555-4565.